From Conceptual Links to Causal Relations

*Physical-Virtual Artefacts in Mixed-Reality Space*

Thomas Pederson
Abstract

This thesis presents a set of concepts and a general design approach for designing Mixed Reality environments based on the idea that the physical (real) world and the virtual (digital) world are equally important and share many properties. Focus is on the design of a technology infrastructure intended to relieve people from some of the extra efforts currently needed when performing activities that make heavy use of both worlds. An important part of the proposed infrastructure is the idea of creating Physical-Virtual Artefacts, objects manifested in the physical and the virtual world at the same time.

The presented work challenges the common view of Human-Computer Interaction as a research discipline mainly dealing with the design of “user interfaces” by proposing an alternative or complementary view, a physical-virtual design perspective, abstracting away the user interface, leaving only physical and virtual objects. There are at least three motives for adopting such a design perspective: 1) people well acquainted with specific (physical and virtual) environments are typically more concerned with the manipulation of (physical and virtual) objects than the user interface through which they are accessed. 2) Such a design stance facilitates the conceptualisation of objects that bridge the gap between the physical and the virtual world. 3) Many physical and virtual objects are manifested in both worlds already today. The existing conceptual link between these physical and virtual objects has only to be complemented with causal relations in order to reduce the costs in crossing the border between the physical and the virtual world.

A range of concepts are defined and discussed at length in order to frame the design space, including physical-virtual environment gap, physical-virtual activity, physical-virtual artefact, and physical-virtual environment.

Two conceptual models of physical-virtual space are presented as a result of adopting the physical-virtual design perspective: for the analysis of object logistics in the context of physical-virtual activities, and for describing structural properties of physical-virtual space respectively. A prototype system offering some degree of physical-virtual infrastructure is also presented.
Acknowledgements

I would like to thank my supervisor Lars-Erik Janlert for giving me the freedom to define my own topic, for acting as an excellent sounding board for ideas, and for coming up with many others to complement my own. Your joy and curiosity in discussing the state of the world “as it appears to be”, and your constant striving for clarity and simplicity when writing academic text, are two out of many aspects of the good research attitude I will do my best to bring with me and foster as a result from this six-year learning and co-operation period. Talking about co-operation, my closest colleagues Anders Broberg and Björn Bengtsson have played important roles both as critics and supporters, as has the Cognitive Computing Group (CCG) as a whole. Many are the research ideas that have been born, tested, and mercilessly killed in this forum.

I also want to thank Lennart Edblom for being the coolest head of department at the University, always willing to spend five minutes of his valuable time, keen and open for discussion, and of course most important of all: playing with unparalleled enthusiasm in the red floorball team every Wednesday. Go Big Red!

As for the specific work on this thesis, a number of persons have kindly chosen to spend time and effort on proof-reading drafts. Apart from my most diligent reader Lars-Erik Janlert, also Häkan Gulliksson, Michael Minock, Annabella Loconsole, and Bill Buxton have provided valuable input. Pointers to relevant information during the writing process has also been provided by Wendy Mackay, Tim Kindberg and Abigail Sellen. Thank you all for improving the quality of this thesis (from whatever level of quality it started...).

The theoretical part of this thesis is greatly inspired by the Magic Touch system developed by mostly computing science students as part of their master thesis projects. Without that work and without Magic Touch, this dissertation would be a completely different thing. Thank you (in order of appearance) Thomas Saathoff, Marie Lindström, Susanne Åberg-Ollander, Thomas Johansson, Fredrik Sandström, Daniel Granholm, Jonas Hägglund, Peter Landfors, Stefano Salmaso, and Stellan Jonsson. I hope it was as much fun for you as it has been for me!

Finally I would like to thank my parents and my brothers for their continuous support and encouragement. At the time of writing I am really looking forward to our traditional Christmas holiday together!

The really final thanks go to the person closest to me in all senses of the word. Bella, thanks for staying with me through rain and sunshine, through ice and snow, and through the writing of this PhD thesis.

Umeå, November 2003

Thomas Pederson
Sponsors
Apart from funding by the Department of Computing Science, financial support for this work has been provided by:

- Ericsson, Telia and SITI (Swedish Institute of Information Technology), sponsoring the research as part of the FIOL research programme.

- HUMlab at Umeå University, financing parts of the hardware used for developing Magic Touch as well as providing room for the system.
  http://www.humlab.umu.se

- Umeå Center for Interaction Technology (UCIT) and ABB.
  http://www.ucit.umu.se

- European Union / VITAL.
# Table of contents

1 Introduction
   1.1 The world is physical-virtual .................................................. 1
   1.2 Designing for the physical-virtual world ............................... 2
   1.3 Motivation .............................................................................. 3
   1.4 Summary of contributions ..................................................... 4
   1.5 Research method ................................................................. 5
   1.6 Thesis outline ................................................................. 7

2 HCI, society, and five new areas of research ......................... 9
   2.1 A brief history of modern computer use ................................ 9
   2.2 A brief history of Human-Computer Interaction as a research discipline. 14
   2.3 Emerging theoretical frameworks, research areas and methodologies .. 19
   2.4 Representative work within the relevant new HCI sub-areas ....... 22
   2.5 Summary ............................................................................. 33

3 Related work ........................................................................... 35
   3.1 The physical-virtual environment gap .................................. 35
   3.2 Physical-virtual activities ...................................................... 40
   3.3 Synchronised physical-virtual artefacts ................................. 43
   3.4 The physical-virtual design perspective ............................... 54
   3.5 Comparing the physical to the virtual world ....................... 58
   3.6 The situative model of physical-virtual space .................... 60
   3.7 The hierarchical model of physical-virtual space ................. 62
   3.8 The Magic Touch system .................................................... 66
   3.9 Summary ............................................................................. 73

4 Human activity in the physical and the virtual world ............ 75
   4.1 The physical and the virtual world ....................................... 75
   4.2 Activities ............................................................................ 78
   4.3 Physical and virtual activities .............................................. 81
   4.4 Human activities ............................................................... 82
   4.5 Summary ........................................................................... 88

5 Physical-virtual activities, physical-virtual artefacts ............. 89
   5.1 Human physical-virtual activities ........................................ 89
   5.2 A closer look at physical-virtual artefacts ........................... 94
   5.3 Distributed PVAs ................................................................. 99
   5.4 Summary ........................................................................... 104
Conclusions........................................................................................................... 261
  11.1 Summary ................................................................................................. 261
  11.2 Contributions ........................................................................................... 263
  11.3 Limitations ................................................................................................ 263
  11.4 Future work .............................................................................................. 264
  11.5 Closing remarks ....................................................................................... 267

Appendix A: English version of questionnaire (chapter 6) ...................... 269

Index ......................................................................................................................... 283

References ............................................................................................................... 285
Introduction

This chapter summarises the work presented in this dissertation, briefly explains the context in which it has been performed, the research method used, and provides a motivation.

1.1 The world is physical-virtual

Can you imagine an ordinary weekday without computers? Perhaps, if one limited oneself to the computer “archetype” of today: the Personal Computer (PC). But what if we take away embedded computing systems such as the ones inside alarm clocks, coffee makers, toasters, cars, radio tuners, code locks, cellular and wired telephones, cash registers, wrist watches and TVs? All in all, it would certainly make the world a bit different from what many of us are used to. Our de facto dependency on computing technology is vast and indisputable.

The emergence of the virtual world

During the past 20 years or so, many computers, small and big, have become (or are becoming) interconnected through a global network, the Internet, enabling people to take a peek inside many of these computing devices even if they happen to be far away geographically. Sometimes, it is even possible to change the state of these devices at distance, assuming that you have the proper equipment and are authorised to do so. The currently most wide-spread kind of interactive window towards this global network is that of the “web browser”, giving access to the World-Wide-Web (Berners-Lee, 1989), and which has made it very reasonable to conceptualise the Internet, or at least
a significant part of it, as a world full of different places, things and even people. I will
denote this world, which can be accessed from a growing number of interactive devic-
es (e.g. desktop and laptop PCs, PDAs, cellular phones, wearable computers) the virtual world throughout this thesis.

Physical-virtual activities
Both the embedded computers that act in the “background” and are not normally
conceived as computers, and the more tangible and interactive computer artefacts
(e.g. PCs, PDAs, cellular phones) play an increasingly important role in human activ-
ity. However, there are still things that most people would prefer to do “off-screen” in
the physical (real) world, such as having parties, reading long text documents, or
spending vacation. In this thesis I argue that there exists a third class of activities that
are neither physical or virtual, but “physical-virtual”. People not seldom switch be-
tween performing parts of an activity in the physical world (e.g. to proof-read a text
document under construction) and other parts of an activity in the virtual world (e.g.
to adjust paragraphs within “the same” document in a word processing environment)
or vice versa. It is likely that such activities, physical-virtual activities, continue to in-
crease as computers become part of an increasing set of human activities, and the sug-
gestion given in this dissertation is that environments should be designed with such
activities in mind.

1.2 Designing for the physical-virtual world
The need to take into account the physical world when studying and designing for
Human-Computer Interaction (HCI) has been expressed by an increasing number of
researchers in the field for the past 10 years or so. The work presented in this thesis
rides on the same wave of “HCI field expansion” but takes a quite extreme stance by
viewing the physical and the virtual world as entities existing in parallel and being of
equal importance for interaction. The models developed in some of the last chapters of
this thesis (chapters 8 and 9) are products of a tortuous but deliberate approach for
establishing a view of physical-virtual (mixed reality) space with as little bias towards
any of the two worlds as possible. Motivated by the observation that many objects
have representations both in the physical and virtual world, attempts are also made to
view these two parallel worlds as one single world, held together by one of the core con-
cepts of this dissertation: Physical-Virtual Artefact.

Perhaps the proposed physical-virtual design perspective is best described by pos-
ing a set of questions that it tries to answer or at least shed some light on:
• What if the physical world does (or would do) more to interaction with comput-
ers than just providing “context”? I.e. what if a state-change of the physical world
would actually be part of the result of performing the physical-virtual activity, and
not just irrelevant “interaction session left-overs”? Is the classical HCI term “user
interface” able to capture the interaction richness of this kind of settings?
INTRODUCTION

What if there are objects in the physical and the virtual world that are identical or have the same meaning? In that case, are there ways to strengthen this conception?

Why is it that physical-virtual activities often imply extra work, an overhead, at the point where the activity switches from being physical to being virtual, or vice versa? E.g., why can’t we just “copy and paste” the content of a paper fax in front of us into an e-mail we are currently writing?

Is it possible to model physical and virtual space in such a way that physical and virtual places, objects, and activities appear side by side? In that case, under what conditions?

In order to address questions such as the ones above, and in order to explore the design space that emerges from adopting a physical-virtual design perspective as described above, a set of theoretical concepts have been developed of which some have already been mentioned. These concepts are intended to form this thesis’ main contribution to the field of HCI, and to constitute an initial platform towards a more complete framework for designing environments better suited for human physical-virtual activities.¹

1.3 Motivation

Physical environments (e.g. an office, a shop floor, a sports stadium, or a house) and virtual environments (e.g. the environments offered by personal computers (PCs), digital assistants (PDAs), and cellular phones) are not viewed as completely separate entities by human agents when performing modern information technology-supported activities. On the contrary, physical and virtual objects are often related to each other by representing the same or similar things, by being used for the same purposes, or just by being used within the same activity context. This fact has led to the emergence of “physical-virtual activities” in society, e.g. activities in which human agents continuously and frequently are forced to, or choose to, alternate between the two worlds. However, bridging the gap between physical and virtual environments in the context of such physical-virtual activities is not for free in terms of cognitive load, time, and money.

The purpose of this thesis is to investigate the dimensions of the physical-virtual environment gap and to explore the possibilities of designing information technology for helping human agents bridging it.

¹ At times, the word “framework” will be used to denote the set of concepts described in this dissertation though the concepts are probably not enough coordinated and tested in practice in order to deserve such a “rigidity-promising” name. On the other hand, misuse of the term “framework” in this way has become more of a rule than an exception in much HCI literature.
1.4 Summary of contributions

The thesis’ main contribution is a set of theoretical concepts for modeling and designing physical-virtual environments, i.e. environments providing both real-world and virtual spaces in which human activity can take place:

**Definition of the physical-virtual environment gap (chapter 6)**
The border between the physical and the virtual world is investigated along four dimensions: “presentation”, “manipulation affordance”, “organisation/navigation”, and “causal dependency”. Different approaches for bridging the physical-virtual gap along these dimension are discussed both from the perspective of a human agent acting in a physical-virtual environment and from the perspective of a physical-virtual environment designer. A (limited) empirical study of the extent of the physical-virtual environment gap in a tele-communications company is also presented.

**Definition of physical-virtual activities (chapter 5)**
Set up by the definitions of human physical and virtual activities presented in chapter 4, the concept of physical-virtual activities is defined and elaborated on.

**Definition of synchronised physical-virtual artefacts (chapter 5)**
The concept of Physical-Virtual Artefact (PVA) is introduced as a way to describe a set of objects (in its simplest case, a set of two) that belong together although they are situated in the physical and the virtual world respectively. The possibility of keeping the two objects automatically synchronised, i.e. the design of synchronised PVAs, is also discussed from the aspects of design challenges and interaction challenges. The complexity arising from sharing “distributed PVAs” among several human agents and environments is also briefly discussed.

**A physical-virtual design perspective (chapter 8)**
The need for a more unified view on the physical and the virtual world is argumented for and limitations of such an approach is also discussed.

**A situative model of physical-virtual space (chapter 8)**
Definitions of “space” general enough to model both physical and virtual space is used as a basis for a situative physical-virtual space model intended to be used as a conceptual tool for modeling how both physical and virtual objects travel between the “physical-virtual world space”, the “observable physical-virtual subspace” and the “manipulable physical-virtual subspace” during the course of activity. A simple object ontology (“domain object”, “tool”, “agent”, “container”) is also introduced as well as the distinction between intra-manipulation (the change of an object’s internal state) and extra-manipulation (the altering of the relationship between an object and other surrounding objects).
A hierarchical model of physical-virtual space (chapter 9)
Based on assumptions of human everyday conception of “containment”, hierarchical models of both physical and virtual space are presented and later joined into a hierarchical model of physical-virtual space, clearly indicating where physical and virtual environments meet.

Supportive work
The above mentioned theoretical concepts are supported and grounded in additional work presented in this thesis:

- Definitions and distinctions made between physical and virtual human activity (chapter 4).
- A comparison of spatial and navigational aspects between instances of the physical and the virtual world (chapter 7).
- The development history, architecture, and functionality of Magic Touch 1.0, a physical-virtual environment prototype based on a novel approach for tracking location changes of objects in office environments using a wearable object identification and localisation unit (chapter 10).
- The architectural design of Magic Touch 2.x, a not yet fully implemented system that is intended to provide physical-virtual infrastructure for collaborative activities over the Internet (chapter 10).
- The context-providing chapters (number 2 and 3) which not report on the setting for this thesis work without taking a stance, but instead to some extent provide a subjective analysis and interpretation of HCI history; five new research areas; and specific work performed in these areas relevant to this dissertation.

On the generalisability
Although the framework has been developed in close interaction with the development of a specific physical-virtual prototype (Magic Touch, described in chapter 10), measures have been taken to make the concepts applicable for design and analysis of physical-virtual (e.g. Augmented/Mixed Reality) systems in general. — Many components of the framework have their origin in general phenomena observed in everyday modern human activity. The generalisability of the framework has however not yet been proven and is left for future work.

1.5 Research method
The work presented in this thesis has largely been conducted using methods common in “the technology exploration branch” of HCI discussed in the next chapter (page 19). As such, it has often been hard to empirically validate the resulting theoretical concepts and the developed interactive system. The design and choice of components of the proposed framework is instead throughout the thesis mostly argumented for, and motivated by, other factors such as established HCI theory and practice, em-
empirical studies of knowledge work performed by others, and experience from interacting with the prototype system presented in chapter 10 (Magic Touch). One empirical study of knowledge work conditions (presented in chapter 6) has also been performed in order to anchor the work in reality, and to motivate further exploration of the physical-virtual environment gap discussed in the same chapter.

Software, hardware and theory development — all together

The research method used in this thesis work can be described as a design process involving both software, hardware and theory development (the development of the proposed physical-virtual design framework). Figure 1 illustrates the main components of this design process and how they have influenced each other.

Although figure 1 is mainly intended to illustrate the methodology underlying the work presented in this thesis, I believe it to be quite representative for explorative basic research in HCI performed within the new technology-driven research areas (A/MR, U/PC, G/TUI, WC, and CA):

- It involves development activities in both the theoretical, software, and hardware domain.
• The results are as much theoretical, typically “conceptual frameworks” for analysis and design, as they are systems made of software and hardware.
• A working product is not expected (other than accidentally) but rather the hope for a better understanding of the challenges and possibilities of using emerging technology for HCI.
• The customers do not exist, other than as projections of today’s computer users (denoted in figure 1 as “application area”).
• There is plenty of room for testing new ideas (for interaction mechanisms, for activity support, for fun).

In contrast to the classical and in later years much criticised waterfall software life-cycle model (Royce, 1970), and to the currently more accepted spiral software life-cycle model (Boehm, 1988), the development model outlined in figure 1 does not have any “direction of development” or sense of incremental progress. This is because as a basic research activity (which the former two mentioned models were not actually designed to model, the comparison is made in the hope of making the message clearer), there are no pre-defined stages of maturity which the “product” will eventually reach. The closest thing to a classical product would be the scientific articles (or PhD theses for that matter!) usually published based on experiences from performing the work.

1.6 Thesis outline
• Chapter 2 discusses the general societal and research context in which the presented work has taken place.
• Chapter 3 discusses related work.
• Chapters 4, 5, 6, 8, and 9 develop a conceptual framework for design and analysis of physical-virtual environments, in short constituted by the concepts listed in section 1.4, page 4.
• Chapter 6 also presents a limited empirical investigation of the extent of the physical-virtual environment gap in a tele-communications company.
• Chapter 7 compares common virtual interaction paradigms with the physical world with regards to navigational and spatial aspects.
• Chapter 10 presents the system which has been developed in close interaction with the proposed conceptual framework.
• Chapter 11 provides a conclusion.
The overall aim of this chapter is to provide the “big picture” in which the work presented in this thesis has been performed. It gives a brief historical background to Human-Computer Interaction, both as phenomena in society and as a research field. Furthermore, recent trends in methodology, and the use of theory is presented. The chapter also briefly discusses some of the work that has set the stage for much of the HCI research performed historically, as well as representative work within five new HCI areas of high relevance for the work presented in this thesis, namely the areas of:

- Augmented/Mixed Reality
- Ubiquitous/Pervasive Computing
- Graspable/Tangible User Interfaces
- Wearable Computing
- Context(ual) Awareness

2.1 A brief history of modern computer use

The way computing technology has become a part of human everyday life in the (post-)industrialised world can (very simplified) be viewed as having taken place in two consecutive 25-year long phases and will be presented as such below. In addition, is a discussion of what might become a third distinctive phase in computing use from today and 25 years ahead.
HCI in the early days (1950-1975) — Computers as calculators

During the 50’s and 60’s, computers spread from being used within the military for calculating trajectories, for cryptography, etc. to be used by banks and governments for monitoring and computing large data sets such as economical transactions and election votes. Computers were still rare, big, expensive and used only by trained specialists. The data processed by the machines were input and stored locally in absence of computer networks. The material used in the computing machines for storing digital information evolved from being vacuum tubes to electric transistors to integrated circuits (large matrices of miniaturised transistors made of silicon). The embryo for what would become the personal computer as we know it today emerged in the second half of this era as a result of the general cost and size decrease for computing power combined with inventions of input and output devices still used today such as the graphical display (CRT) and the computer mouse.

HCI in the recent past (1975-2000) — Desktop PCs, LANs and WANs

The publicly available personal computers Xerox Star, Apple Lisa and IBM PC initiated the diffusion of computing power into the world of the non-specialists and at the same time, naturally, catalysed research in HCI. Suddenly, computers were to be used by people not trained in mathematics, software engineering, or electronics. The personal computers (PCs) first found their place in the manufacturing industry as a tool for design (i.e. CAD systems) and not much later as typewriter substitutes, soon to evolve into the general office tool and entertainment machine they have become. The affordability and versatility of PCs (not least as entertainment machine) made them enter also many homes during the 1990’s.

In parallel with the diffusion of PCs into the lives of ordinary people, another big change took place during these 25 years in the industrialised world as both public administration and industries inter-connected their computer systems. At first using local area networks (LANs) which in general made collaboration at each site more efficient, and later in wide-area networks (WANs), facilitating collaborative activities also among geographically disperse places and among different organisations. Pre-Internet WANs and the sharing of database information among institutions that are connected through them, has revolutionised the administration process of the governmental and industrial organisations that have chosen to utilise the possibility. One politically and socially difficult question has been (and remains to be) how to best regulate the extent to which information about human individuals should be allowed to be automatically transmitted and shared among the WAN-connected systems.

A third development trend that has affected everyday life in the industrialised world is the embedding of more specialised (compared to the PC) computers into everyday buildings, objects and even construction materials. Although they have surely

---

2. Although there are studies showing that “computerisation” per se does not automatically lead to increased productivity within organisations, e.g. Heikkilä (1995).
had an impact on the way we do things, they are deliberately designed to not appear as “computers” in the everyday sense of the word (hence “embedded”) but more often as remarkably “smart” versions of artefacts that were formerly non-computerised (e.g. cars, elevators, magnetic code locks, etc.). As such, they are under normal circumstances invisible and do their work very much in the background.

The most recent invention that has had a major impact on computer and computing use in the society is of course the World-Wide-Web (WWW) which made the power of WAN-computing practically and conceptually accessible also to non-computing-scientists in the mid-1990’s. The WWW has made WAN-supported human activities expand from being performed in industrial and governmental contexts, to include also individual and leisure-time activities. By doing so, the WWW has made the notion of a world parallel to the “real” world (i.e. the virtual world, Cyberspace, or Hyperspace) graspable not only to science fiction authors, philosophers and computing scientist (who had discussed the possibilities and challenges of such a parallel world for decades) but also to people in other academical fields and not least outside the academic world itself. At the time of writing, it is evident that for a large part of the population in the (post-) industrialised world (in particular the younger generations), Cyberspace as seen through the WWW is a natural companion to the “real” world, and a world that for some of them is almost as frequently visited as the real.

Of particular interest for this thesis, is the fact that the 25-year long “PC- and WAN-era” has brought with it not only a new world (Cyberspace) in which humans perform activities, but also, I will argue, an increasing set of human activities that in part take place in the “old” real world and in part in the new virtual world.

**The emergence of physical-virtual activities**

Personal Computers (and computers in general) are not the perfect tool for all activities although at times during the past 25 years, some of us might have been close to believe so, giving birth to for instance “the myth of the paperless office”. The fact is that certain things are (still) better done in the physical world rather than in the virtual, as provided by PCs. Consequently, as the PC was introduced into the everyday life of human agents, they started to alternate between acting in the physical world in one moment only to switch to the virtual world in the other, while still performing the one and same overall activity (e.g. writing a report). Physical-virtual activities became common, and as I will argue later in this thesis, they are here to stay. The main purpose of this thesis is in fact to define and study this modern phenomenon.

**HCI in the close future (2000-2025) — Wearable PCs, wireless LANs and WANs**

The development of computing technology, and above all its effect on society, is hard to predict. Based on existing state-of-the art computing systems and awareness of how computers did (and did not) find their way into the society in the past, one might however give it a try. The following prediction is by no means revolutionary (rather,
it is pretty much regarded as common sense among researchers and practitioners in the areas of Ubiquitous/Pervasive and Wearable Computing) but deserves to be made explicit here with the motivation that it describes the context in which some of the concepts and technologies discussed later in this thesis are imagined to be used and implemented.

The general assumption is that computers will continue to become faster, cheaper and smaller in size. This, combined with other emerging hardware technology such as ubiquitous wireless networks, will have two major effects with regard to everyday HCI: 1) embedded computers will continue to become increasingly ubiquitous and increasingly embedded, 2) the desktop PC will evolve into an always worn, always Cyberspace-connected and more individualised machine (some people in the Wearable Computer community compare this kind of future computer to a prosthesis) that will not only support tasks for which some of us today use wrist-watches, PCs, PDAs and cellular phones, but also for interacting with people and objects in the (embedded computer-augmented) real world in new ways. Against this cyborg-inspired vision stands, as always, the society’s willingness to accept the social and cultural changes that wide-spread use of this kind of technology would imply. While it is hard to say exactly to what extent we all will be cyborgs (as described above) by the year of 2025, there is definitely the chance that computer power will be as fundamental a resource as water in our everyday lives. Some might say that it already is today. The difference between whether the everyday-supporting computing power is supplied by technology that is worn on our bodies, or if it is embedded in the walls is probably irrelevant for most people except to the engineers that have to make them work and perhaps, to philosophers.

Adopting the engineering stance, the vision poses a set of questions including the one of how to design a standardised data communication infrastructure for all the real-world embedded and human-worn more or less powerful computers. From an HCI research viewpoint, the question of how to design the “user interface(s)” towards this potentially rather complex interwoven real-virtual environment is currently one of the most open, most challenging, and therefore most interesting questions.

**In search for better design frameworks, user models, and hardware**

During the past 25 years, a substantial number of models, guidelines and designs have been developed within the HCI community for the purpose of optimizing the “usability” of the PC. It is safe to say that without these efforts, the diffusion of computing into the society would have taken longer time and perhaps even a different path. During the past 10 years or so, a growing number of HCI researchers have started to search for new ways of modelling and designing HCI systems with the common belief that the “post-PC” era as described and envisioned earlier calls for new ways of looking at phenomena in Cyberspace and in the real world. This thesis presents one possible perspective, a physical-virtual design perspective, in which objects and places in the physical and the virtual world are described using the one and same terminology.
The research methodology used in this *avant garde* sub-field of the HCI research discipline is one where hardware development (once again) plays as big role as software development, and where reliable modelling of human behaviour in the real world is as important as it is difficult. The research communities within the areas of Ubiquitous/Pervasive Computing, Context Awareness, and Augmented/Mixed Reality are all struggling with limitations of the currently available real-world actuation and sensing technology, as well as the difficulty of making the right conceptual framing of relevant human activity. Typically, the “user modelling” efforts inherit the problematic characteristics of classical Artificial Intelligence research by being able to successfully interpret basic human behaviour and intentions in smaller, isolated, and carefully prepared settings, while facing enormous problems as soon as larger, more open-ended activities and environments are to be modelled. Unfortunately for the modellers (and perhaps fortunately for the people being modelled!) many of the actions in the real world tend to be open-ended, fuzzy, and hard to classify. The prototype system presented in this thesis (chapter 10) is no exception with regards to its sensing, actuation, and modelling limitation, adopting a pragmatic bottom-up modelling approach by tracking (and to some small extent) modelling one basic kind of action which humans tend to perform over and over again both in the physical and the virtual world: the action of moving objects from one place to another.

**Computing — only at the right side of the digital divide**

It has to be made clear that the discussion so far in this chapter on the diffusion of computing technology is only relevant and true for a certain group of people: the fortunate minority of the global population living in areas wealthy enough to invest in necessary infrastructure and education for making computing a publicly accessible resource. Furthermore, even within the currently most wealthy regions (such as for instance the USA), the level of “computer literacy” can be quite diverse and often correlate with economical status of individual families and local communities. Naturally, large-scale political efforts in providing access to computing in general and Cyberspace in particular (which have for instance been initiated in the EU during the 1990’s), can have a positive effect in regionally levelling the digital divide. Furthermore, the crack is also caused by factors such as level of education and age. The latter aspect increases in importance as the average age of populations increases (which for instance is the current trend in most states within the EU) and has made the “design for elderly” a high-priority theme within the growing research area and agenda of the

---

3. Throughout this thesis, the term “actuation technology” and “actuators” will be used to denote computer-controlled technology able to change human-perceivable properties of objects. Real-world actuation technology is actuation technology with potential of affecting physical objects. “Sensing technology” denotes computer-controlled technology enabling the computer system to perceive changes to objects external to the computing system. Unless specifically noted, sensors measure changes in the real (physical) world.
European-initiated “User Interfaces for All” (Stephanidis, 1995), and its North-American counterpart “Universal Usability” (Shneiderman, 2000). Both comprise a set of initiatives explicitly directed towards levelling the various aspects of the digital divide on a both regional and global scale.

I ask the reader kindly to bear in mind that the challenges and proposed solutions discussed in documents such as this thesis are relevant and of potential use only for the minor part of the global population that has adequate access to computing power in the first place.

2.2 A brief history of Human-Computer Interaction as a research discipline

Definition
The Curriculum Development Group of the ACM Special Interest Group on Computer-Human Interaction (SIGCHI) defines the research discipline of Human-Computer Interaction as follows in their report from 1992, “Curricula for Human-Computer Interaction”:

“Human-computer interaction is a discipline concerned with the design, evaluation and implementation of interactive computing systems for human use and with the study of major phenomena surrounding them.” (Hewett, Baecker, Card, Carey, Gasek, Mantei, Perlman, Strong & Verplank; 1992).

The term “HCI” criticised — who interacts with computers anyway?
The term “Human Computer-Interaction” has met criticism within the research community during the last decade or so for misleadingly emphasizing the role of the computer as a physical object with which humans interact. The extremely versatile, cheap, efficient, and (compared to pre-transistor days) reliable tool which the Personal Computer has become, has resulted in the fact that the general computer user no longer interacts with computers as special-purpose physical devices in particular physical contexts (which would motivate adopting the view of humans interacting with computers as physical objects) but, and which will be the view taken throughout this thesis, computers as providers of virtual environments in which various kinds of interaction can take place. In fact, the interaction of interest for the majority of HCI researchers has for some time now decreasingly been about the interaction between human and computer (other than for the sub-community exploring alternative input and output devices) and increasingly about the interaction between human and the virtual objects mediated through the computer in the role of a provider of an exponentially growing large virtual space sometimes referred to as Cyberspace, Hyperspace or the virtual world.
The way the work within the discipline is defined in the SIGCHI definition can be taken as an indication of the inadequacy of the term “Human-Computer Interaction” since the term “computer” has been substituted (if it ever was there?) by the more appropriate formulation “computing system” when describing the interaction counterpart towards which the human agent, as well as HCI practitioners and researchers, direct their interest.\(^4\)

**HCI research in the early days (1950-1975) — pre-PC/WIMP interaction paradigm exploration of possible computing hardware, software, and potential use context**

The research discipline of HCI emerged during the post-second-world-war years as a mix of established and emerging disciplines. Areas that are usually listed as having had fundamental influence are Ergonomy, Cognitive Psychology, and Computing Science (in particular computer graphics and operating systems research). A very important factor was also the series of inventions in electronics and electrical engineering (often appearing completely independent from HCI research) that paved the way for the research performed within the HCI community. It is important to acknowledge that without these hardware inventions (in particular the integrated circuit and the bit-mapped display), HCI and HCI research would look completely different. On a more theoretical level, important new work was done in modelling and measuring abilities and limitations of human cognition with respect to information-intense tasks (Cognitive Information Processing). Another important theoretical development that helped clarify and focus the work within the community was the various expressions of the idea to bring computers and humans closer to each other, of which Douglas Engelbart’s “augmentation of human intellect” project during the 60’s (e.g. Engelbart, 1963) is representative.

**HCI research in the recent past (1975-2000) — within-PC/WIMP interaction paradigm fine-tuning of the software to suit human “needs” and abilities in the context of desktop PC/WAN equipped office work**

If the early 25 years were the years of finding and defining the interaction paradigm and use context for novice computer users (as well as waiting for the hardware development to reach the necessary level of miniaturisation and acceptable cost levels), the

\(^4\) This thesis work belongs to Human-Computer Interaction. Acknowledging the shortcomings of the term “Human-Computer Interaction” as a slightly misleading term for denoting the kind of work commonly performed under its name today, the SIGCHI-definition quoted earlier frames the work presented in this thesis very well. The presented research does indeed discuss the design and implementation of computing systems for human use, and perhaps even more, various phenomena that appear during the course of human activity in relation to computing systems. The evaluation aspect (which is also a component of the HCI definition) of the presented work is this thesis’ weakest point, as will be discussed later in this chapter.
subsequent 10-15 years became the time of implementing this interaction paradigm in the society. The HCI community grew in size and importance in direct connection to this office-computerisation process when an exponentially growing population of computer-illiterate office clerks exposed themselves, and were exposed, to the new “augmented typewriters” on their desktops. The hardware part of the PC/WIMP system was left “as is” from the early years except for the somewhat late large-scale market introduction of the computer mouse (the mid-1980’s). The main focus was instead during this period (and still is!) on improving the usability of the virtual environment visualised on the computer screen and to expand the tools and the number of tools available so that every imaginable knowledge work task could be supported. I think it is rather safe to say that the major advances and the overall best solutions for supporting traditional office work within the PC/WIMP interaction paradigm have been found, and that the mission taken as a whole has been a great success.\(^5\) PC/WIMP systems are found on nearly all desktops today and play an important role in everyday knowledge work. Of course the credit (in lack of a better word) is not all the HCI community’s, but significantly so.

Among the most concrete contributions of HCI research during this period to the PC/WIMP paradigm of today, we find:

- The invention of the virtual desktop metaphor and the WIMP interaction paradigm (Windows, Icons, Menus, and Pointing device), often attributed to the people behind the Xerox Star PC, e.g. Smith, Irby, Kimball, Verplank Harslem (1982).
- The idea of creating the illusion of “direct manipulation” of virtual objects (Shneiderman, 1983).
- The exploration of HCI as a real-time dialogue, e.g. Donald Norman’s “gulfs of execution and evaluation” (Norman, 1988)
- The idea and implementation of local and global hypermedia. The original idea is usually attributed to Vannevar Bush (1945) while the credits of early implementation and exploration of local hypermedia on computers usually is attributed to Ted Nelson, Douglas Engelbart and Andries Van Dam for their work in the mid-1960’s. The inventor of the architecture for the first global hypermedia system (the World-Wide Web) was Tim Berners-Lee (1989).

Figure 2 illustrates some of the major focus areas within the HCI research discipline during the past 25 years, and also to some extent what has been considered external to core HCI research.

---

5. Of course, a certain amount of energy within the HCI community is also used for incorporating support for new knowledge work activities as a response to advances in hardware development as well as changes to the context in which the knowledge work takes place. However, these enhancements seldom challenge or question the PC/WIMP paradigm itself.

6. In particular, and once again, achievements in hardware development has been crucial.
HCI research in the close future (2000-2025) — post-PC/WIMP interaction paradigm exploration of hardware software, and potential use context

Looking back on the HCI work made during the period 1950-1975 and comparing it to the work performed during the 25 years after that it seems that while the first period was about finding and defining a working interaction paradigm for non-specialist computer use, the past 25 years has mainly been about polishing it. Why it is so is of course an open question. From a more balanced and less cynical viewpoint one might say that the reason that PC/WIMP is still the ruling computing paradigm is because all other tried alternatives (although few have actually been proposed) have been less suitable for the tasks at hand. The PC/WIMP paradigm was simply a great design from the start. Another reason might be that there has been no series of inventions in the field of electronics (corresponding to the ones in the early days) that could inspire new ways of interaction. Another important factor is the enormous “paradigm polishing” industry that has evolved during the mentioned time period and doing everything it can in order to ensure that the (for them) very lucrative paradigm stays.
My guess is that the PC-WIMP interaction paradigm depends on persistence of all four mentioned factors:
1. The major part of the HCI community has focused on “paradigm polishing” rather than searching for an alternative (in part because of a healthy lazy attitude),
2. the PC/WIMP design was amazingly good from the start,
3. there has been little (at least not discovered by the HCI community) revolutionary development in the field of electronics relevant for HCI during the last two decades, and finally because
4. a huge industry prefers people to continue to use mice, keyboards, screens, and WIMP-based operating systems as long as possible.

The end of the PC/WIMP polishing era?
Although the PC/WIMP interaction paradigm probably will continue to be the focus for the major part of the HCI community for another few years, recent advances in hardware technology and standardisation (in particular electronics miniaturisation, battery improvements, new sensor technology, and ubiquitous wireless network access) might soon trigger a paradigm shift just similar to how in the past the electronics development played an important role in determining the interaction paradigm of the early days of HCI research.

In fact, as portable/wearable computers, wireless networks and the new possibilities for making computers aware of events in the physical world emerged during the 1990’s, a growing number of academics within the HCI community left their “paradigm polishing” activities and began searching for an interaction paradigm that could better embrace the interaction possibilities enabled by this new hardware. Meanwhile, the established PC industry of course has continued to push for the use of the PC/WIMP paradigm also in this new technological context. Some of the “post-WIMP user interface” work that was done during the 1990’s will be discussed in the next chapter on related work. The subsequent chapters will be devoted to my own attempts in contributing to the search for a new interaction paradigm.

If the state-of the art hardware technology of today is enough for driving the paradigm shift, or if one or two more electronics inventions have to arrive first, is an open question. The increased interest in alternative interaction devices is an indication that the shift is at least approaching. There is of course also a slight possibility that the existing WIMP paradigm could be adapted and work well enough also for controlling computing in the new technological and social context as well. It is however unlikely, considering the basic assumptions of the WIMP paradigm such as

- the assumption that the human agent can dedicate all attention to the interaction with the virtual environment provided by the computer (e.g. does not bike or drive a car)
- the assumption that the real world environment in which the interaction takes place is always the same (quiet, in the shadow, etc.)
2.3 Emerging theoretical frameworks, research areas and methodologies

**Hardware rules (once again)**

The discipline of HCI depends and thrives on hardware technology almost as much as does computing science in general. HCI also has other “input channels” such as advances in the cognitive and social sciences. This means that the basis for HCI research is determined by on the one hand, advances in hardware development (often considered to be at the periphery of the discipline), and on the other, development in various fields of psychology (usually considered as core). While the latter research disciplines advance in a relatively steady incremental fashion in the traditions of classical natural and human sciences (at least from the perspective of a layman such as myself), the former field is one of constant yet unpredictable lane-changing and above all, running at a speed faster than most other research fields. The result is, I would argue, that although HCI is indeed a multi-disciplinary field of research, hardware advance has potentially greater impact on HCI (both as a phenomena in society and as a research discipline) than any other factor. Having pragmatically acknowledged this fact, a certain line of HCI research has developed a research methodology based at any given moment in time on the construction and evaluation of HCI system prototypes using and inspired by the very latest in hardware. Perhaps this trend is a logical reaction from the HCI research community after having been so focused on PC/WIMP interaction for the past 25 years, that hardware advance and innovation could not be followed and explored with sufficient curiosity?

**The technology exploration branch of HCI research**

As a science that in part studies and develops artefacts, HCI research in general does not search for “truth”, but for the optimal human-machine system in a given context. The explorative branch of HCI research (which includes some of the work presented in this thesis) does not even try to solve problems. It tries to find ways of utilising the just arrived technology from the never-ending stream of electronics innovation, for HCI purposes. The research methodology used might in Software Engineering terms be characterised as rapid prototyping with a sometimes very small and/or fuzzy requirements specification. As such, this particular branch of HCI research constitutes

---

7. Major differences between "artificial sciences" and natural science is discussed by Herbert Simon (1996).
an extremely explorative basic research activity which has to constantly fight in order to convince the more classically oriented research communities (both within and outside HCI) that the work performed is in fact science, and not applied systems development, undirected play, or art. A significant part of the work done in this fashion is performed within or in collaboration with high-tech industry whose interest of course is to find applications for the new technology. Perhaps the reason why the explorative methodology actually has found certain acceptance within the HCI community is the awareness that the field at large is very dependent on hardware and that if the community keeps a steady eye on the state-of-the-art, it has greater chances to affect the overall design of future products.\(^8\)

Five highly technology-driven sub-areas of HCI research which all emerged during the 1990’s (and which all will be discussed in more detail in the next chapter on related work) are:

- Augmented/Mixed Reality (A/MR)
- Ubiquitous/Pervasive Computing (U/PC)
- Graspable/Tangible User Interfaces (G/TUI)
- Wearable Computing (WC)
- Context(ual) Awareness (CA)

**On the lack of proper evaluation**

The most explorative kind of activities are particularly vulnerable to criticism from “hard-core” scientists as they fail to show falsifiable results. Typically, a) the tasks which the prototypes are designed to support are vaguely defined (or simply “new”) and therefore hard to compare with existing tasks and results from any previous studies, and/or b) the technology used is so immature that it cannot be tested on non-trained subjects and/or compared with more traditional ways of performing the task. As a result, many publications based on this kind of research activity tend to focus on quite isolated interaction aspects of the system as well as technology issues rather than presenting a statistical analysis of whether the new real-world setup A is better than the more traditional setup B. Other results from performing this kind of research include the unique possibility of developing interaction theory and models based on the experience of designing and using the prototypes. Often, the theoretical constructs proposed in such work are considered as (albeit weakly) validated as a result of being able to support and model the presented working computing system. Not seldom, the theoretical work is further validated and extended by constructing another prototype system, and so on, eventually proving the generalisability of the proposed theoretical

\(^8\). It has in fact been a long tradition for HCI practitioners to express their frustration over how late in the design process they are consulted and how this limits the possibilities of making a good design from an HCI perspective.
framework. To summarise, although this explorative work and pragmatic research method for good reasons is subject to constant criticism, it seems to have become a generally accepted part of HCI activity for other equally good reasons.9

**Back from the virtual desktop to the hardware exploration desk**

As will be discussed further on in this thesis at various locations, I believe it to be more beneficial than ever for the HCI community to have an active collaboration with various kinds of hardware designers (in this context including not only electronics engineers and researchers but also physicists, architects, industrial and mechanical designers), as a potential computing interaction paradigm shift approaches. It was in such a research context the currently dominating interaction paradigm was defined, and it is probably in this kind of research context that the new interaction paradigm will be shaped as well. To say that the field of HCI will be fundamentally transformed in the process (which is in fact what some people in the community have recently suggested, using terminology such as “The New HCI”) is perhaps a bit extreme since human cognitive skills and human needs can be regarded as constants and many theoretical design frameworks and methods developed until now within the realms of the HCI community will of course still be applicable in any future computing context.

The HCI research community is not facing a grand scientific paradigm shift à la Kuhn (1962) but one on a smaller scale, as the general interest focus gradually shifts away from the PC/WIMP interaction paradigm to something not yet defined. A reasonable guess is that it is the work which fundamentally relies on the PC/WIMP interaction paradigm that will decrease in importance, eventually leaving a gap to be filled with new theories and frameworks for the design of new kinds of interactive everyday computing systems.

**The industry detour for findings in physics**

The new exploratory HCI research areas are to some extent inherently contradictory as they on the one hand contain many elements of intuition-driven basic research, and on the other rely heavily on more or less purely market-oriented and applied development of electronics hardware. Of course there is a lot of basic research undertaken also within the field of computing hardware in the shape of various research areas of physics, but today, findings in these areas do in general take the detour through computing industry before being taken up and acknowledged by the HCI community. I believe

---

9. It should be mentioned that there are also many examples of explorative hardware-dependent work within HCI that actually is performed according to the standards of classical science because of luck or a “good nose” for what will work and what will not (stumbling into the “right” hardware and applying it to the “right” interaction context) or because of exceptional skills in setting up and performing experiments even though the experimental conditions are non-optimal. The utility of a common evaluation framework as a mean for establishing a “common ground” within these new research communities has also been proposed (e.g. Scholtz & Richter, 2002).
that the new exploratory HCI research areas would benefit from tighter collaboration with physicists and that knowledge in both HCI and physics will be very valuable in this line of research in the years to come.

The changing role of theory in HCI

“In an attempt to be more applied, many of the new approaches have sought to construct conceptual frameworks rather than developing fully-fledged theories in the scientific Popperian tradition.” (Yvonne Rogers, 2004.)

Yvonne Rogers notes (2004) that HCI theory developed and used before the late 1980’s often was targeted towards informative, predictive and prescriptive design and analysis of HCI (e.g. cognitive modelling as used in experimental cognitive psychology; Hutchins, Hollan & Norman’s (1986) conceptual framework of directness; Norman’s (1986) theory of action; Card, Moran & Newell’s (1983) Model Human Processor (MHP) and GOMS (Goals, Operators, Methods, Selection rules), while later theoretical constructs and frameworks tend to be used for other and more diverse purposes10, such as to:

• provide descriptive accounts (rich descriptions)
• be explanatory (accounting for user behavior)
• provide analytic frameworks (high level conceptual tool for identifying problems and modeling certain kinds of user-interactions)
• be formative (provide a lingua franca; a set of easy to use concepts for discussing design)
• be generative (provide design dimensions and constructs to inform the design and selection of interactive representations).

The last two kinds of “categories of theory” listed above, i.e. the formative and the generative, do very well describe the physical-virtual design framework proposed in this thesis. The conceptual framework is in fact intended to provide a terminology for describing an until now rarely discussed phenomena within the HCI community (to the best of my knowledge), i.e. that of physical-virtual activities, as well as to inspire for new interaction design by opening up the design space.

2.4 Representative work within the relevant new HCI sub-areas

Before introducing the reader to five relevant new HCI research areas by briefly discussing important and representative work in the respective areas, a general line will be drawn between those focusing on technology per se, and the ones focusing on the

---

10. As examples of the newer theoretical approaches Rogers (2004) mentions activity theory, ecological psychology, distributed cognition, external cognition, situated action, ethnomethodology, and various kinds of application-specific theories.
use of technology. Not everyone in the respective research communities would agree to such a distinction. A general consensus on how the areas relate to each other does not exist as the research areas are still relatively young. The following categorisation is the one found most appropriate and true in my eyes.

**Wearable & Ubiquitous/Pervasive Computing — building computing infrastructure because it is possible**

U/PC and WC are areas whose names indicate their focus on hardware technology. Within these research communities there is usually no hesitation as to whether it is a good idea to leave the desktop PC/WIMP paradigm or why computers should be worn or embedded in our physical environments. Instead, the utility of pursuing these paths is in many cases simply taken for granted. The main aim of much research in these fields is to make it technologically possible to successfully wear them and to build them into the walls. Finding application areas and demonstrating the utility is regarded as being on the periphery or outside of the areas themselves\(^{11}\), while the development of a reliable and flexible data-communication infrastructure is core. From an HCI perspective the positivistic engineering approach adopted might seem questionable but the fact is that these extremely technology-driven research fields can act as valuable “future probes” providing hints on what kind of hardware architecture will underlie HCI in the future. As discussed in the previous chapter, hardware technology does extensively decide the context of HCI research because technology is a key enabler of new interaction mechanisms and eventually new HCI interaction paradigms.

**Augmented/Mixed Reality, Graspable/Tangible User Interfaces & Context(ual) Awareness — building things to see how new technology can be used**

In contrast to WC and U/PC discussed in the previous section, A/MR, G/TUI, and CA are research areas that have more explicitly incorporated the search for application areas and use of technology into their research agenda. Very often, interaction mechanisms are developed, investigated and evaluated based on hardware infrastructure imported from work in the previously discussed sister research areas WC and U/PC. The significant property of work performed in these areas is that interaction is of primary concern, hardware and software technology of secondary. This does not mean that

---

\(^{11}\) This does not prevent U/PC and WC researchers from being active in the other new research areas as well. It is in fact very common that once a working wearable and/or ubiquitous computing platform has been designed (and while it is being designed) by an individual or group of individuals within the Wearable and/or Ubiquitous/Pervasive Computing field, they enter the areas of Augmented/Mixed Reality, Graspable/Tangible User Interfaces and/or Context(ual) Awareness in order to evaluate the computing infrastructure. The very frequent “cross-fertilisation” taking place between U/PC and WC on the one hand and A/MR, G/TUI, and CA on the other (making technology engineers think about HCI and HCI people think about technology engineering), and the fact that individual researchers and research groups tend to contribute to more than one of them at different points in time, sometimes makes it very hard to distinguish between the areas.
people active in these areas do not need to have knowledge in hardware design. It is in fact an important sub-task for researchers in these areas to keep themselves up to date with latest development in sensing and actuation technology as well as data communication. The main task is however in general to build HCI systems that combine these sensors and actuators in new interesting ways from an interaction designer’s perspective. All three areas attempt to involve the physical world as an integrated part of a HCI system, making real-world object identification and location tracking important common research issues. Real-world objects does in this context often include both “dead” artefacts, computing-augmented artefacts, as well as (human) agents.

**Augmented/Mixed Reality (A/MR)**

The research field of Augmented Reality (AR) is often defined as a field exploring the opposite HCI path to that of Virtual Reality (VR) (e.g. Wellner, Mackay & Gold, 1993), the latter being the field which also inspired the choice of the name. While VR investigates the challenges and possibilities of immersing human agents into the virtual world, AR tries to let objects and functionality of the virtual world “leak” into the real physical world. While AR in its widest sense denotes any kind of computer-augmentation of the real world, a certain sub-community focusing on visual world integration, usually based on location and orientation-tracked Head-Mounted-Displays (HMDs), has for some time preferred the term Mixed Reality (MR) instead of AR. Proposals for sub-ordinating AR as a special kind of MR has also been proposed (Milgram & Colquhoun, 1999), introducing another conceptual counterpart to AR, named “Augmented Virtuality”. The term MR has in the recent years gained acceptance in the AR community as a more general term for environments consisting of both physical and virtual phenomena because it is not, like AR, biased towards any of the worlds but quite neutral. Accepting that the relationship between AR and MR has still not reached a wide consensus, the terms will be used as synonyms in this thesis.12

**“Digital desk” prototypes**

Starting with Pierre Wellner’s (1993) Digital Desk, a series of AR/MR prototypes have been built and evaluated in office-like environments during the 1990’s by many different researchers and engineers. The setup usually consists of a projector mounted in the ceiling that projects a virtual environment onto a real-world desktop. A video camera is mounted in the ceiling as well for tracking objects and human activity on the desk. The desktop surface can be sensitive to touch, or to strokes of pens. The physi-

---

12. This thesis does in fact (and unfortunately) introduce yet another term for settings involving both real and virtual phenomena: “physical-virtual environments”. The motivation is that the proposed framework does not in any significant way stem from VR or is dependent on visual integration of the two worlds. Instead, it is a general, object-centred and world-neutral conceptual modelling stance: Both the physical and the virtual world contains objects and afford support for human activity. The places in which the worlds meet are viewed as physical-virtual environments.
cal-virtual environment which the real-world desktop becomes typically provides tools that cannot be made to appear without the projector/camera combination and allows for combined organisation and manipulation of both real and virtual documents.

**Augmenting REAL real-world settings**

The attempts to implement AR/MR technology outside the research labs has been the trademark of Wendy Mackay (e.g. 1998, 2002) paying particular attention to, and trying to preserve, many of the existing human practices at each real-world site instead of introducing a completely new interaction paradigm (typically PC/WIMP) into the environments. As a consequence from this approach, the existing real-world environments can remain largely the same, reducing the learning curve, but still offer unique functionality and activity support from the digital domain. As “AR” as AR can get.

**Visually co-locating physical and virtual objects**

The attempts of visually presenting virtual objects as being part of the real world was pioneered by Ivan Sutherland in the 1960’s (Sutherland, 1968). The basic principle has been the same since then: 1) An image of the real world is captured by one or two cameras mounted close to the human agent’s eyes. 2) Virtual objects are added to the real-world image previously captured and displayed on digital displays situated in front of the human agent’s eyes. By 3) tracking the position and orientation of the agent’s head in real time, the virtual objects can be made to appear as if they were actually situated in the real world, in particular if the objects are rendered in three dimensions using VR techniques. Alternatively, half-transparent displays can be used for merging virtual with physical instead of video cameras. The drawback with this latter approach is that it is harder to perfectly place the virtual objects in the real-world “scene” while the advantage is that the real world is viewable without any time delay since the real-world part of the presented image does not have to pass through the computing architecture.

This particular branch of AR/MR is very hardware-dependent and has benefited a lot from the miniaturisation of wearable display technology, advances in tracking sensor hardware, and the general increase of computing power (for the image processing typically involved). Although HMD is the most common display and viewpoint tracking choice (e.g. Bajura, Fuchs & Ohbuchi, 1992), other alternatives have also been investigated (e.g. Rekimoto & Nagao, 1995; Ullmer & Ishii, 1997; Siio, 2001; Tsang, Fitzmaurice, Kurtenbach, Khan & Buxton, 2002; Mackay, 2002).

A special case is when the virtual objects merged with the physical world image are not virtual independent (three-dimensional) objects but textual information about real-world objects currently in the field of view, i.e. “annotating” the real world. This latter case of more conceptual integration of the world belongs more to AR than the previously mentioned perception-based integration approaches that fall into the category of MR.
Ubiquitous/Pervasive Computing (U/PC)

While the origin of the term “Ubiquitous Computing” within the U/PC community is unisonally attributed to Mark Weiser (1991), the term “Pervasive Computing” appears to have gradually emerged at a later stage, and having a more anonymous originator. The two terms are in any case at the time of writing equally frequently used and denote exactly the same kind of research.

As discussed earlier, U/PC work typically focuses on the design of data communication infrastructure for HCI beyond the single-user, single computer PC/WIMP interaction paradigm. A common underlying design goal is to make a large part of the actual computing, and the computers themselves, as unobtrusive as possible, or even invisible.

“The most profound technologies are those that disappear.” (Mark Weiser, 1991)

Going beyond single computer single user PC/WIMP means to consider both a larger design space of potential input and output devices, an extended geographical space as part of the interactive system, as well as to a higher degree regard the social context in which HCI takes place.

Data communication infrastructure

As a representative example for research focusing on infrastructure design, Kindberg, Barton, Morgan, Becker, Caswell, Debaty, Gopal, Frid, Krishnan, Morris, Schettino, Serra (2000) base their data communication infrastructure on a WWW architecture by embedding web servers in various devices and letting these web servers act and react on people’s interaction with these devices. The system supports automatic or semi-automatic correlation between WWW pages on the one hand and “people, places and things” in the real world, on the other. The infrastructure is applied to various settings including the “Cooltown Museum” which is a (real-world) museum augmented with functionalities based on the presented data communication infrastructure, and the “Cooltown Conference Room” that facilitates collaboration in a meeting room. The system uses RF/ID (Radio Frequency IDentification) tags and infrared receivers and transmitters for obtaining necessary information about the real world. A Wireless Local Area Network provides the necessary data communication connection to the objects that are moving around (e.g. human agents).

As is typical for this line of research, a great deal of effort is put on designing a suitable data communication architecture. Kindberg et al. discuss at length the various challenges faced and design solutions made for assuring that the intended correlation between the numerous web servers, tracking systems and object identifiers is success-

13. Ironically, the first issue ever of the IEEE Journal of Pervasive Computing (2002), although being dedicated to the late Mark Weiser and his legendary article from 1991, does not mention or even touch on the subject of why the editors chose the name “Pervasive” instead of Weiser’s original “Ubiquitous” Computing!
fully managed by the system. Also typical, Kindberg et al. in addition discuss, although in perhaps less detail, the interaction and user interface issues that emerge in the environments using their infrastructure.

**Big-picture U/PC**

While most work within the realm of U/PC tends to focus on quite geographically limited indoor places, and sites where very specific human activity can be assumed to take place (e.g. a dedicated meeting room in an office building), attempts for constructing systems that encompass larger geographical space and less specific activities do exist. As an example, Norbert Streitz and his colleagues at GMD/Fraunhofer-IPSI have during the 1990s developed a range of U/PC systems and tools for supporting various collaborative knowledge work activities in office buildings (see Streitz, Tandler, Müller-Tomfelde & Konomi (2001) for some examples and an overview), making notable efforts in keeping “the big picture” of general collaborative knowledge work and the complete office building in mind. Several of the proposed systems and computing-augmented artefacts work together in flexible ways and with the intention of creating what they refer to as a “Collaborative Building”, a building that supports human-human collaboration by affording the best support the digital and the real world can currently bring for such activities. Although Streitz et al. do not use location tracking of individual human agents as a core component of their infrastructure, many other similar attempts in creating more “intelligent” office buildings do (e.g. Addlesee, Curwen, Hodges, Newman, Stegges, Ward & Hopper, 2001).

Another kind of “big picture” U/PC effort is constituted by the interest in constructing “Intelligent Family Houses” that almost exploded in the end of the 1990’s. Representative examples are the “Aware Home” project at Georgia Tech (Kidd, Orr, Abowd, Atkeson, Essa, MacIntyre, Mynatt, Starner & Newstetter, 1999) and “com-Home” at KTH/CID (Tollmar & Junestrand, 1998). Starting out with high expectations from the construction industry for a quick introduction of the IT solutions into existing and near-future dwellings, this kind of projects have in general revealed (and which should be regarded as one of their major contributions) the shortcomings of today’s technology and user interfaces in coping with the complex and flexible social environment a private home indeed is. While the acceptance for technology breakdowns and high thresholds in learning curves can sometimes be surprisingly high in professional contexts as found in office environments, the same can definitely not be said about the places where people spend their leisure time with family and friends. It seems like simple straight-forward artefacts, digitally augmented or not, are still preferred to less tangible and more ubiquitous “intelligent” systems. Perhaps frightening science-fiction pictures of central and ubiquitously “aware” computers such as the HAL 9000 make people reluctant? In any case, the fact remains that although the

---

U/PC community with theoretically sound arguments continues to argue for the benefits of connecting everything with everything, there are yet few examples that prove a significant gain in everyday practice outside the laboratories. If the reason is as simple as an immature data communication infrastructure, the solution will come from the U/PC community.

**Distributed user interfaces**

Although U/PC is in part about making computers disappear as objects from within the field of human perception, these invisible computers need to be addressed every now and then by human agents and thus have to be accessible through user interface devices. Weiser (1991) foresaw that in contrast to the uniform appearance of PCs, those interaction devices will be available in many different shapes and forms, as well as in many different places. This line of research typically assumes a data communication model based on data of interest hidden in a server somewhere in Cyberspace, and a multitude of input and output devices (I/O devices) potentially acting as clients towards this server when requested by human agents. Thus, infrastructure that simplifies the switching from one I/O device to another (e.g. Rekimoto, 1997; Fitzmaurice, Khan, Buxton, Kurtenbach & Balakrishnan, 2003) and handles a multitude of representations of the same virtual object (e.g. Myers, Malkin, Bett, Waibel, Bostwick, Miller, Yang, Denecke, Seemann, Zhu, Peck, Kong, Nichols & Scherlis, 2002) belongs to the investigated topics of this particular branch of U/PC research. Although the general aim of research in distributed user interfaces is often referred to as being to investigate alternative I/O devices to the ones found within the PC/WIMP paradigm, until this date most prototype systems in this category nevertheless seems to be based on visual displays of various sizes, mouse-like input devices and WIMP virtual environments.

**Graspable/Tangible User Interfaces (G/TUI)**

The G/TUI research can be viewed as taking the basic idea behind distributed UIs one step further by tearing the usually quite complex I/O devices within the distributed UI research paradigm (often fitted with both visual displays and input devices, i.e. providing “full-duplex” communication with the embedded computing system) into even smaller, more specialised, and less “intelligent” functional components. G/TUIs try to utilise real-world spatial and tactile dimensions in their design based on the idea that human agents as a result of Darwinian evolution and by simply growing up in the real world, possess skills in these domains without having to do additional training.

G/TUI components (common terms are “Bricks”, “mediaBlocks”, and “Phicons”) are concerned with interaction at a lower level of abstraction than distributed UIs (which provide similar components within the virtual domain instead of the real). Thus, they do typically only function as part of tailor-made applications made up of a set of G/TUI components and an application that frames and interprets the actions performed on those components similar to the way UI components in the WIMP par-
adigm communicate changes to software applications in the virtual world. Since the first fundamental publication on the subject (Fitzmaurice, Ishii & Buxton, 1995), the G/TUI approach has been applied in many different settings (see Ullmer & Ishii, 2000, for a review) of which the support of collaborative design (e.g. Arias, Eden & Fischer, 1997; Fjeld, Bichsel & Rauterberg, 1998; Underkoffler & Ishii, 1999) is probably the most common.

G/TUI interfaces are sometimes argued for as being more efficient and intuitive compared to WIMP-based user interfaces (“GUIs”, “Graphical User Interfaces”) although empirical evidence is in large lacking, and although their application areas do not completely coincide. The experiment of Fitzmaurice and Buxton (1997) is one out of very few that explicitly addresses the question and gives some credibility to the claim.

At the periphery of standard G/TUI research there are examples of graspable computationally augmented artefacts that do work in “isolation”, i.e. without a governing software application and without being used in spatial combination with other G/TUI components. For example, the person-to-person haptic tele-presence device “inTouch” (Brave, Ishii & Dahley, 1998) is able to transmit valuable information even through a minimalistic communication channel thanks to the intimacy of human touch. The device senses and actuates change of rotational position of three cylinder-shaped bars at each communication end point in real time, simulating that the two human hands that in reality might be thousands of kilometres apart, are only as far apart as the diameter of the cylinder rolls.

**Wearable Computing (WC)**

The possibility of carrying computing power with us all the time brings Bush’s (1945) and Engelbart’s (1963) visions of augmenting human intellect (see earlier in this chapter) to a completely new level.

> “Wearable computing pursues an interface ideal of a continuously worn, intelligent assistant that augments memory, intellect, creativity, communication, and physical senses and abilities.” (Thad Starner, 2001).

While the pioneers of modern WC had to carry heavy, voluminous, and extremely conspicuous equipment, the systems of today are getting so small, lightweight, decent-looking and cheap that powerful general-purpose wearable computers might very well become mass-market products within a few years from the time of writing. (See Steve Mann’s article (1997) for example photographs of both early and current state-of-the-art systems.) Apart from the core activities of the WC community of improving the designs of wearable computers (involving power consumption issues, heat dissipation, size, weight, I/O device design, etc.) WC researchers tend to have interests in many of the other new HCI research areas as well. As explained and exemplified very well by Thad Starner (2001), the “always on, always with you” characteristics of the ideal
wearable computer does indeed enable unique possibilities of trying out new interaction ideas within the fields of A/MR, U/PC, G/TUI and CA. Perhaps this is also the most important function of WC research within the discipline of HCI — to provide a concrete but not yet fixated system platform as an alternative to the PC platform, making it easier for explorative HCI researchers to think different by loosening up the conceptual ties to the PC/WIMP paradigm, ties that have been proven to be very hard to cut.

Wearable computers range from computationally simple identity badges transmitting infra red light to full-power desktop PC replacements. Also digital wristwatches and cellular phones are examples of wearable computers. While the less powerful wearable computers tend to act as not completely independent components of a ubiquitous computing system embedded in the environment (typically for enabling the identification and tracking of individual agents) the more computationally powerful wearable computers often act as the computation and storage core of a personal wearable network to which digital cameras, microphones, earphones, displays and keyboards are all attached depending on the intended use. Also these powerful wearable computers tend to communicate with the physical surroundings in order to frame the context of the actions taken by the human agent wearing the computer.

Ultra-Personal Computers

While the “always-on, always with you” computer might be a big step to take for most people at the time of writing, the current trend of miniaturisation of laptop computers into “Ultra-PCs” (UPCs) might be a more socially acceptable intermediate step. The Ultra-PC prototypes that have been shown by hardware developing industry in 2002 and 2003 are small enough to fit in a jacket pocket while still providing standard PC functionality (in contrast to PDAs) except for limited I/O user interface because of the lack of space for full-size keyboards and screens. The basic idea of the UPCs is that people should always bring with them their personal digital data and use the UPC “docked” to full-size screens and keyboards whenever such are available, and use the UPC as a stand-alone device only occasionally. In other words, UPCs are intended to be “always with you” but not “always on”, and should eliminate the need for PDAs as complement to standard PCs.

Wearable computing concepts somewhere in between the classical full-fledged wearable computer which is always accessible and the idea of UPCs that are made mainly for interaction in dedicated places, have also been proposed. Want, Pering, Danneels, Kumar, Sundar & Light (2002) present “the Personal Server” as a wearable, always-in-your-pocket-computer (even at times of interaction) that depends completely on external I/O devices to which it automatically and wirelessly connects whenever such devices happen to be available. The complete opposite approach in separating between computation hardware and user interface is exemplified by the idea of “thin clients” where the computing power is assumed to be omnipresent (e.g. built into the walls) and only components for interaction need to be made tangible to
(and wearable by) human agents. The thin client approach to WC is interesting because computing power can be relatively heavy (especially if the batteries needed for it is included), produce heat (which is always a challenging design issue for equipment close to the human body), and physically fragile. The major drawback of the approach, and which is probably the reason for why it is not really considered an alternative for most WC applications today, is its complete dependency on available computing power at the geographical locations in which the wearable thin client is supposed to be used. Perhaps the thin client approach will become more attractive as computing power and WLAN become more ubiquitous.

**Context(ual) Awareness (CA)**

As HCI systems expanded beyond the virtual environment presented on a computer screen during the 1990’s and started to encompass also real-world objects and places, the need to better conceptualise these new components of the system, as well as the intentions of the human agents currently operating the system, became pressing. The new mobility of computing also made it necessary to expand the interaction design space. New terms such as “situated interaction”, “situation awareness” and “contextual awareness” were introduced to describe this wider perspective in HCI. Context aware systems are different from traditional HCI systems not only because they tend to utilise the state of the physical world as part of interaction, but also because they do it *implicitly*. One might say that Context Aware systems provide computational functionality directly or indirectly tied to real-world events without adding more input devices in the classical HCI sense but by gathering information in other ways (typically through sensors of which the human is not necessarily aware of).

Although the term “context awareness” could in theory imply an interest for building systems that gather contextual information from any environment including digital (virtual) ones, almost all published work in the area of CA has so far been focused on sensing and modelling phenomena in the *real* world. A typical example is to have the context-aware computer system switch on the right computer display out of several possible alternatives, based on in what direction a specific person is headed, how many other persons there are in the room, what time of the day it is, and what have you. Instead, to infer intentions and use context based on simple measurable interaction primitives in the *virtual* world is something which has been worked on in the somewhat older research area of User Modelling. The two areas obviously face many common challenges and it is reasonable to believe that they will grow closer to each other, especially if the real and virtual worlds become increasingly intertwined as predicted by many researchers in the new explorative HCI research areas.

The currently most used term for describing a computing system’s ability to sense relevant real-world phenomena is “context awareness” and I will use this term throughout the thesis, although the term “context” is a word with many possible interpretations. Dey, Salber & Abowd (2001) provide a wide yet pragmatic definition:
“Context: any information that can be used to characterize the situation of entities (i.e. whether a person, place or object) that are considered relevant to the interaction between a user and an application, including the user and the application themselves. Context is typically the location, identity and state of people, groups and computational and physical objects.” (Dey, Salber & Abowd, 2001)

Computers can technically speaking provide an immense number of useful functions for an equally large number of everyday situations, especially if they are connected to Cyberspace. One major problem that the context awareness researchers address is how to make the right functions available and executed in the right way in specific situations. They try to identify and construct reliable heuristics for triggering easy selection (or automatic execution) of suitable actions. Viewed in this way, the connection to the classical computing science discipline of artificial intelligence (e.g. plan recognition) is very evident although surprisingly seldom spelled out in context awareness-related publications.

Another important aspect of context awareness research is the development, implementation and handling of sensors. It has in part been the emergence of cheap and powerful real-world sensing technology that led to the possibility of actually constructing context-aware computing systems. Based on sensor data, with the help of real-world models, context aware systems infer the state of the relevant part of the real world. Depending on the application this can include the state and location of human beings; dumb artefacts such as chairs; complex artefacts (more or less computerised machines) such dish washers and cars. To create environment and world models in a way so that a chain of deduction can be performed reliably (as to what state the world is really in, and what the human being really wants to happen), has been proven to be hard, but constitutes at the same time the very core of context aware research. As a result, HCI researchers active in the CA area have become increasingly interested in topics such as place, space, topology, spatial modelling, spatial cognition, location modelling, as well as for general everyday physics.

The expansion of HCI into the real world and the wish to sense the context of human activity has also given the social dimensions of HCI and human-human communication increased importance since social context is obviously an important part of the overall context in any HCI setting. The difficulty in sensing, modelling and interpreting phenomena in the social domain compared to the physical domain does however make designers of context aware systems focus on simple physical measures such as temperature, location, and time of the day.

As pointed out by Bellotti and Edwards (2001), challenges lie not only in making computers context-aware but also to make human agents aware of the nature of the computer system’s context awareness. Since a core idea of CA is to let the computer system draw conclusions based on activities of human agents that are not explicitly directed towards controlling the computer in the traditional HCI fashion, it becomes hard for human agents to understand the full implications of their actions.
As for representative examples of context-aware systems, almost all prototypes developed within the domains of A/MR, U/PC, G/TUI, and WC make use of some kind of context as defined by Dey et al. (2001). This indicates that CA should be viewed as a very general area of interest for the new explorative HCI research areas.

**Implicit output**

While typical CA research focuses on what could be denoted “implicit input”, examples of exploration of “implicit output” as an HCI component exist as well. The terms “foreground” and “background” (e.g. Fitzmaurice et al., 1995; Ishii & Ullmer, 1997) activities have been used to clarify the distinction between explicit and implicit interaction as described earlier in this section. In particular Ishii & Ullmer (ibid.) propose the idea of using “peripheral displays” as a means for informing human agents unobtrusively.

2.5 **Summary**

This chapter has presented a brief history of HCI, both as a phenomena in society and as a research field. Furthermore, representative work within five new explorative HCI research areas have been discussed: Augmented/Mixed Reality, Ubiquitous/Pervasive Computing, Graspable/Tangible User Interfaces, Wearable Computing, Context(ual) Awareness.
The related work presented in this chapter is structured in one of many different possible ways, namely according to how it relates to the concepts in the physical-virtual design framework proposed in this thesis. One aim of this chapter is to give the initiated reader a quick hint on what the following chapters will discuss, based on already published material in the research areas to which this dissertation attempts to contribute. Another aim of the chapter is to show that the concepts and ideas proposed in the rest of the thesis have not been shaped in isolation (should anyone believe so) but in fact as a partial result from inspiring and rewarding studies and discussions of other research.

For readers not well acquainted with research in the new explorative HCI areas, it is recommended to skip this chapter and get back to it after having read the rest of the thesis.

3.1 The physical-virtual environment gap

That there exists some kind of “gap” between the physical (real) world and the virtual world (Cyberspace) is a notion probably shared by many people that regularly use PCs. While most PC users naturally take the physical-virtual environment gap for granted as an unavoidable obstacle, it is (or should be) of central concern for everyone involved in the new explorative HCI research areas trying to integrate the two worlds.

15. It should be noted that the “objective” analysis in this chapter is occasionally interrupted with reflections and subjective statements related to the discussed work. The hope is that this presentation style makes the reading experience more interesting and does not significantly affect the validity of the survey.
Although it has been hard to identify other attempts to aggregate many different aspects of the gap into one concept, one gap (as done in chapter 6 in this thesis), a few quotes will be presented below in an attempt to show that particular dimensions of the gap has attracted attention from other people within the HCI community:

“One of the most significant issues confronting computer users, is the problem of bridging the gap between the physical and virtual worlds. For most activities, most current systems make it too difficult to move the artefacts back and forth between these two worlds, the physical and the virtual. Hence, the relevant documents, designs, etc. are isolated in one or the other, or split between the two.”

(Bill Buxton, 2001)

“... in today’s world, and probably for some time to come, many of the resources we use are both tangible and virtual. Most books currently reside outside the computer, and we value the affordances this provides, and certainly our colleagues reside outside our computers and most of discussions with them take place non-virtually. But use of tangible resources including people also carries a transaction cost for the modern writer. For now we must transfer the contents of those books, and conversations, as well as the contents of any notes and annotations we wrote down, from outside to inside our computer. No doubt an effective way to meet this challenge will be worked on for several years.” (David Kirsh, 2000)

We switch between physical and virtual environments because they are good for different things. David Kirsh continues:

“It is not a simple problem to be solved by digital scanners or video cameras. For our goal is to capture the content arising in the activity of composition as it is distributed over many spaces and many environments. Each of these environments – virtual and physical – has its own functionality and each has its own special resources that we have learned to use.” (David Kirsh, 2000)

However, as is discussed in chapter 6, the physical-virtual environment gap is not only constituted by “inertia” in information transfer but has other dimensions as well. One of them being the need to actively manage the two parallel sets of objects (the virtual and the physical), keeping them more or less synchronised and up-to-date.

“Users come to rely on the new features offered by the computer, but also maintain the paper artefact. They must thus manage two kinds of documents: those embodied as physical paper and those entirely on-line, with a new problem of how to manage the link between the two.” (Wendy Mackay, 2000)

16. Emphasis added by me.
The physical-virtual environment gap can be seen as a special case of the gaps that make switching between any kind of environments harder, also *within* the two worlds, e.g. the cost of taking the boat from the mainland to the island; the cost of having to re-type or re-draw a virtual object because the destination virtual environment (application) cannot automatically “import” the design from the original environment “as is”. The particularity of the physical-virtual environment gap is that many of its dimensions are caused by fundamental structural and material differences between the physical and the virtual world as explored in chapter 7 and as noted by others (Arias, Eden & Fischer, 1997; Sellen & Harper, 2002).

**Cognitive workflow and cognitive overhead**

Studying cognitive aspects of knowledge work, David Kirsh uses the term “cognitive workflow” to describe the way human activity is affected positively and negatively by structural properties in both physical and virtual environments (e.g. Kirsh, 2000). The physical-virtual environment gap can be seen as a phenomena that to a higher or lower degree disrupts the workflow during the performance of an activity that demands switching between physical and virtual environments, including the cognitive workflow as described by Kirsh. Analogously, physical-virtual activity overhead (definition 6.2.2, page 108) encompasses components of cognitive overhead in the same kind of physical-virtual settings, but also other non-cognitive aspects (see the gap dimensions listed in table 7, page 131). The different ways for environment designers to reduce the physical-virtual environment gap (discussed in section 6.4, page 123 and onwards) are in Kirsh’s terminology examples of approaches to “reduce interruption” (Kirsh, 2000) when switching between environments.

Arias *et al.* (1997) suggest that the cognitive workflow (although they use other terms) can be improved by integrating “physical and computational models” in their specific AR urban design prototype, allowing for a “continuity of argument” in collaborative settings.

**Bridging the gap needs motivation and some degree of synchronisation**

In the context of a “hands-on” exhibition where visitors were encouraged to interact with physical objects as part of the exhibition and at the same time were given a Personal Digital Assistant (PDA)-based exhibition guide, Fleck, Frid, Kindberg, O’Brien-Strain, Rajani & Spasojevic (2002) found that the costs of switching attention between the physical world (the exhibits, companions, surroundings, etc.) and the virtual world of content delivered through the handheld device were considerable, and was one of the main reasons for changing their real-time “exhibition guide”-design to a post-exploring “exhibition rememberer” design instead. They also observed cases of
“lost in hyperreality” when the synchronisation between the places in the real and the virtual world in the PDA at times was left to the user to ensure, while at other times it was performed automatically or semi-automatically by the computing system. From the perspective of the physical-virtual environment gap as described in chapter 6, the findings of Fleck et al. can be commented on as follows:

1) The specific exhibition was by default essentially an exhibition to be explored in the physical world. All experiments, all tryouts and all results of one’s attempts took place in the physical world (or at least, in the environment external to the virtual environment provided by the PDA). Since all environment switches are connected to costs, there has to be some motivation for performing them. The exhibition as a whole did basically not provide such a motivation other than the “wow”-factor mentioned by the authors, simply emerging from introducing a high-tech gadget. I would argue that the PDA-guide did not significantly enhance the experience of interacting with the exhibition because it was not part of it but provided static information about it, which is a significantly weaker relation. Why would anyone do a world switch from the physical to the virtual environment when everything of interest happens in the physical environment? If one for some technology-positivistic reason would like the visitors to use the PDA anyway, the exhibition designers would need to sufficiently integrate the virtual environment in the PDA with the rest of the exhibition to motivate an environment switch. For instance: a) By making the exhibition exploration less intuitive and thus forcing the visitors to consult the guide in order to get something out of the exhibition (a questionable approach, but in this way the virtual world would definitively provide added-value: the value taken away from the physical environment!). b) By strengthening the causal connection between phenomena in the two environments and let the objects in the PDA and the real-world exhibition objects provide experiences or perspectives that complement each other. Using the terminology of the framework presented in this thesis one would need to develop and place a sufficient number of interactive synchronised Physical-Virtual Artefacts (PVAs, see chapter 5) into the exhibition, which of course would be a project of a totally different magnitude in costs and development efforts, not to mention that the exhibition itself would probably become a different kind of exhibition.

2) The “lost in hyperreality”-phenomenon noted by the authors can be seen as a result of a malfunctioning synchronisation of a PVA consisting of the real-world exhibition space and the exhibits it contains (the physical PVA manifestation) on the one hand, and the collection of web pages (the virtual manifestation) inside the PDA on the other. If the “lost in hyperreality” phenomena would show to be a major obstacle for the visitors (Fleck et al. actually reports on only one single case), the designers of the system could consider removing the feature of being allowed to navigate the exhibition web structure without at the same time navigate the real-world exhibition cor-
respondingly.17 (Alternatively, provide mechanisms for automatically transporting the user between real-world exhibits as a result of browsing the web structure, but again, this would be a rather big project.)

**Synchronisation problems in incompletely shared physical-virtual environments**

Bowers, J., O’Brien, J. & Pycock (1996) report on particular gap-related problems in a collaborative distributed VR setting. While the participants in the meeting all share the same visualisation of the common space in the virtual world, including an audio space, they are (naturally) all immersed in local, non-shared, physical environments. Any phenomena in the local physical environments is perceivable only by the single participant residing at the specific physical place. Communication problems in the collaboration setting occur when local physical phenomena (e.g. if a meeting participant leaves her/his physical room temporarily without explicitly “signing off” to the others; if a person external to the distributed meeting engages in a conversation with one of the participants on site, drawing the attention of the specific meeting participant away from the meeting; or if the local network or computing environment experiences problems) since the other participants rely on the information available from the shared virtual space only, and this shared space is not completely in sync with each local physical environment at all times. Bowers et al. conclude the vulnerability of the setup as follows:

> If activities in the real world are adequately aligned with the embodiment’s activities in the virtual world, if—for the purposes at hand—the embodiment adequately displays its attentiveness to ongoing activity, if participants are adequately mutually focused on a common thread of activity, and if the machines are working, then the embodiments can be relied upon both by their ‘owners’ and by the others as trust-worthy resources for social interaction. Under such circumstances even just a subtle turning or an approach can influence the course of social interaction (e.g. by selecting who is next to speak or eliciting further details) or be closely meshed with it (e.g. by timing movements in close coordination with those of others or their speech). However, if users engage in activities in the real world which mean their virtual actions are either inappropriate (e.g. if a not-so-innocent bystander gives not-so-innocent ad hoc instruction) or non-existent when expected (e.g. if a call of the phone or of nature has had to have been answered), or if the embodiment cannot adequately portray a user’s attentiveness or readiness to participate (e.g. if an embodiment has ‘corpsed’), or if activity devolves into multiple threads (e.g. as some users are engaged in troubleshooting local problems,

17. To be able to act in the virtual world “un-synchronised” with the physical can of course also be viewed as a useful feature, allowing the human agent to try out possible scenarios (to explore “the possible worlds”) in a quicker trial-and-error fashion than what is possible in the physical world. Whether “un-synchronised activity” in physical-virtual space is to be regarded as a feature or an obstacle depends on the intention of the human agent performing the activity.
while—at the same time—others complain about the audio, and yet other parties are trying to focus the gathering on introducing themselves, or if a machine has crashed, then the embodiment cannot be trusted as a user’s representative in a virtual world and other methods will have to be resorted to (closely attending to or adapting one’s use of the audio link, scrutinising the text window, launching unix talk or making a phone call).” (Bowers, J., O’Brien, J. & Pycock, 1996)

Traversable Interfaces — making human bodies literally travel across the gap
Benford, Greenhalgh, Reynard, Brown & Koleva (1998) investigate various ways of letting human agents experience the boundary between physical and virtual environments. The work stands as contrast to most A/MR research as the approach emphasizes the presence of the border rather than attempts to eliminate it, and at the same time provides means to cross it. One of the main aims is to try to reduce the problems of shared VR environments as noted by Bowers et al. (1996) (see the brief description in the previous section) by forcing the human agents to actually leave the physical environment in which they might be initially, if they want to enter in a collaborative virtual space. The work is mentioned here partly because it is interesting in many ways (some of the prototypes are pieces of innovative performance art more than anything else), but mainly because the fundamental approach is completely opposite to the one taken in this thesis, which instead investigates the challenges and possibilities of integrating physical and virtual environments rather than separating them from each other.

3.2 Physical-virtual activities
As described in the introduction chapter, physical-virtual activities have emerged as a result of the general diffusion of virtual environment providers (PCs, PDAs etc.) into the everyday lives of human agents in the (post-)industrial society. Life is for many of us inherently physical-virtual as we frequently shift our attention from physical to virtual environments and back. Still, although the interest is growing, attempts to model, understand, and especially design for this shifting behaviour are quite few within the HCI community compared to the immense efforts in designing better virtual environments.

One important trend related to the study of physical-virtual activities has been the expansion of the HCI research discipline to incorporate social aspects of human-computer settings. Although infrastructure for support of “virtual communities” has provided means for social interaction also in the virtual world, it is safe to say that a considerable part of human social life will continue to take place in the physical world. Thus, human needs to socially interact is in fact a major guarantor for ensuring the existence of physical-virtual activities also in the future. An obvious weakness of this thesis is that it ignores social aspects of physical-virtual interaction; a deliberate deci-
sion made in order to narrow down and limit the scope of the presented work. As a result, related work in the social sciences will not be discussed here although such connections certainly exist.

Instead, some studies of physical-virtual activities in the context of knowledge work (Drucker, 1973) will be presented since it seems to be in this context physical-virtual interaction has been studied the most so far. Knowledge work, being a name for the performing of information- and knowledge-oriented activities as opposed to for instance activities for producing goods, is not an eccentric occupation these days but rather a common and growing one. Statistics from Statistics Sweden (Statistiska centralbyrån) show that the amount of service-oriented workers (who in most respects can be regarded as knowledge workers) constituted approximately 80% of the total workforce in Sweden year 2000 (Dahlbom, 2003).

Like many of the discussions and proposed design solutions in this thesis, the knowledge work-related studies presented below tend to be centred around physical and virtual objects, perceivable artefacts and structures to which human agents devote much of their attention. The fact that there are indications that everyday objects even essentially drive human activity (Suchman, 1987), makes an object-centric modelling approach a reasonable one.

Writing from multiple sources
The authoring of text documents is an important part of knowledge work that typically involves both physical and virtual activities and thus can be viewed as a physical-virtual activity. O’Hara, Taylor, Newman & Sellen (2002) have studied what they refer to as “the hybrid task of writing from multiple sources” and found that the human agents switched between physical and virtual environments based on current needs in specific sub-tasks (e.g. reading, searching, comparing). Even more interestingly, they report that their subjects tend to make use of multiple physical and virtual manifestations of the same document in parallel:

“Writers from multiple sources often use different media formats of the same document concurrently to provide a complementary set of affordances for interacting with and using information from multiple sources. For example, paper supported the quick flexible navigation of a document for more non-specific information search behaviours due to its tangibility and fixity of information with respect to a page. The electronic documents used supported more specific search behaviours and the ability to transfer information from one document to another through copying and pasting to see how it worked within the composition.”

(O’Hara, Taylor, Newman & Sellen, 2002)

This finding strengthens the credibility of the idea of creating an infrastructure for automatic synchronisation of physical-virtual artefacts (see chapter 5) such as text documents since human agents are (at least in this context) frequently making use of
multiple representations of the one and same object. Only the “synchronisation link” is missing. Furthermore, O’Hara et al. make another important point in the context of discussing electronic books:

“... it may be more important for these digital alternatives to focus more on the unique affordances that they can offer over paper rather than over investing effort into trying to emulate the affordances of paper for which they are likely to remain inferior.” (O’Hara, Taylor, Newman & Sellen, 2002)

Generalising from their statement we end up in one of the fundamental properties of physical-virtual artefacts as suggested in this thesis: that PVA manifestations in the physical and the virtual world should complement each other rather than just mimic each other’s affordances and appearances, in order to best make use of the unique characteristics of the two worlds. If the identification of corresponding PVA manifestations is supported by an infrastructure (which is also explored in this dissertation), the necessity for similarity in appearance decreases even further.

Organising large sets of documents

One particular sub-activity that knowledge workers seem to prefer doing in the physical world is the task of organising documents. Jacob, Ishii, Pangaro & Patten (2002) report that despite a potential “unbroken virtual chain” in the overall task of organising submitted papers in suitable conference sessions, the organisers preferred to switch to the physical world when grouping the documents although it (of course) implied costs when manually bridging the physical-virtual environment gap:

“Despite having excellent computer support at this meeting (held in a computer-enhanced meeting room at Microsoft Research), having all the relevant data already in electronic form, and needing to produce our final results in electronic form, the committee found it more effective to use manual/tangible interaction to perform this task this year, as in previous years.”

(Jacob, Ishii, Pangaro & Patten, 2002)

Drawing on the long tradition of the Tangible Media Group at MIT Media Laboratory of designing application-specific G/TUI systems, Jacob et al. constructed a prototype (the “Senseboard”) in an attempt to decrease the cost of bridging the gap in organising tasks such as the one described above. Comparing the support given in an experimental setup by a) the environment afforded by their physical-virtual prototype with b) pure physical and c) pure virtual environments, they found (although at weak statistical significance levels) that the task was completed faster using their prototype than in the other settings.

The work of Jacob et al. constitutes a concrete example of the general design approach outlined in this thesis by 1) identifying an existing physical-virtual activity or environment where a set of physical-virtual activities frequently occur, and then to 2)
design a physical-virtual system that helps human agents perform this activity more easily by helping them bridge the physical-virtual environment gap during the course of the overall activity.

**Learning as a physical-virtual activity**

The idea of seeing learners as knowledge workers acting in physical-virtual environments has been proposed in the light of the pedagogical theories of social constructivism and phenomenography (Milrad, Broberg & Pederson, 1999).

### 3.3 Synchronised physical-virtual artefacts

While concepts corresponding to “physical-virtual activities” have been hard to locate in the present HCI literature, concepts similar to “physical-virtual artefacts” are instead very common, particularly in the work performed within the new explorative HCI research areas. This is not surprising since any interest in involving the real world as part of an HCI system inevitably leads to the development of mechanisms for connecting something in the real world with something in the virtual, which is essentially what the concept of PVA is about.

The first main difference between the PVA concept presented in this thesis and most other related concepts is that it is intended to be used as a very general conceptual tool for the design and modelling of almost any kind of physical-virtual environment, while the latter tend to be more tuned towards specific applications or application domains. Consequently, the corresponding terms found in other HCI literature are more powerful in describing specific kinds of physical-virtual environments (e.g. collaborative design settings) but too narrow when discussing physical-virtual design on the rather high abstraction level of this dissertation. Full-fledged synchronised PVAs, i.e. artefacts whose manifestations in the physical and the virtual world share many attributes (see section 5.2, page 94), is a vision to strive for more than something we are able to construct using the technology of today.

The second main difference is that the concept of PVA is “world-neutral” and “world-symmetrical”. The first property means that the concept of PVA promotes a modelling perspective that does not consider one of the worlds as “point of departure” or “default environment” and the other as secondary, which is otherwise very common. The second property means that the causality relation between the manifestations of a PVA ideally is bi-directional, and not just one-way which is the most common relation between physical and virtual objects in HCI systems of today (i.e. from the physical to a virtual manifestation commonly referred to as “input” to a computing system, or from the virtual manifestation to a physical commonly referred to as “output” from the system). Used as a tool for user interface design, the PVA concept does in fact to some extent challenge the long-term dominant idea of constructing HCI systems on the basis of input and output devices as well as the concept of “user interface”, by offering the alternative approach of designing HCI systems based on
PVAs instead. As mentioned earlier, current limitations in sensing and actuation technology makes the concept more useful for theoretical work rather than practical since the construction of full-fledged PVAs is to date a very technologically challenging task.

**Graspable user interfaces**

Several important properties of graspable user interfaces have been pointed out by George Fitzmaurice and his associates (Fitzmaurice, 1996; Fitzmaurice, Ishii & Buxton, 1995), properties that are of high significance for PVAs and even the underlying motivation for developing PVAs. By eliminating what Fitzmaurice et al. denote the “acquire logical device”-part in the classical GUIs (e.g. to move the mouse pointer in a virtual environment to the object that should be pressed, dragged, or otherwise manipulated) graspable user interfaces allow for a more direct interaction with virtual objects since the physical (graspable) device and the logical (virtual) device are tied to each other at all times. Bill Buxton (1986) refer to this interaction paradigm based on tight physical-virtual connection between physical and virtual objects as “space multiplexing” as opposed to the “time multiplexing” interaction schema traditional GUIs are based on.

The PVA concept can be viewed as a generalisation of the graspable user interface concept as it in addition to the input channel from the physical to the logical device also encompasses an output dimension to the relation between the physical and the logical components, achieving “actuation symmetry”. That is, as a PVA manifestation, the virtual device is assumed to be able to affect the state of the physical device as well. Fitzmaurice et al. (1995) use very minimalistic non-computing-augmented physical artefacts, “Bricks”, as graspable devices. If one would like to transform the graspable user interface components into PVAs, these Bricks would either need to become more complex devices with basic computing power, wireless data communication and actuation capability (e.g. visual, tactile, and/or aural display components), or such actuation capability has to be induced from a system outside the specific components (e.g. by a so called “intelligent” environment of some sort).

Fitzmaurice (1996) also notes that graspable user interfaces are less general and better tuned for specific uses compared to time-multiplexed devices and that a well-balanced mix between the two kinds of interaction devices might be the best design in many settings. This insight applies also to PVAs since they share the space multiplexing strengths and weaknesses with graspable user interfaces. However there is a small point to be made in this regard: PVAs can potentially be very complex artefacts (e.g. whole PCs; even large constructions like buildings can be modelled as PVA manifestations) and allow for quite general interaction activities as well, all depending on the level of abstraction chosen by the modeller.
**Tangible user interfaces**

Having developed a long impressive series of graspable user interface prototypes starting in the second half of the 1990’s where visual displays (typically large-size projector-based) have played an important role, the Tangible Media Group at MIT Media Laboratory headed by Hiroshi Ishii have added system output to the original concept of graspable interfaces and at the same time given them another name: “tangible user interfaces” (TUIs). In their interaction model of TUI called MCRpd (see figure 3) Ulmer & Ishii (2000) distinguish between the three components control, representation and model. Notably, the control component is only present in the physical representation of the tangible user interface, and that is also the major difference between MIT Media Lab TUIs and the PVA concept proposed in this thesis. The latter also includes such a component on the digital side. The fact that the MCRpd model does not include the possibility of affecting the physical representation of the TUI from the digital world is not a weakness if the model is used for modelling the kind of TUIs built at the MIT Media Lab and in many other places where the technological possibility of implementing ideas might be almost as important as getting them. On the contrary, demonstrated by Ulmer and Ishii in a survey on TUIs (Ullmer & Ishii, 2000), the model can be successfully used for describing many existing TUIs. However, since the work in this thesis is of a more theoretical nature, more future-oriented, and less focused on implementation issues, the asymmetry of the MCRpd model and the narrowing of the design space that follows from excluding causal relationships in the direction from the virtual to the physical world, motivates the use of the more general PVA concept instead. Whether to include or not include the control component in the virtual world has philosophical implications. If it is included, human agents have to be modelled as potentially situated inside the virtual world as well to make any sense. Such a view is completely outside the scope of the traditional and ruling idea of “user interfaces” in HCI in which the human is always considered to “be” in the physical world. Whether it is realistic to model humans as being inside the virtual world or not

**FIGURE 3.** The MCRpd interaction model underlying many Tangible User Interfaces. (Ullmer & Ishii, 2000).
also indicates to what degree we believe human power of insight plays a role in inter-
action. Thus, the decision is of cognitive relevance as well. The physical-virtual design
perspective outlined in this thesis proposes a modelling stance in which human agents
can be imagined to be situated in the virtual world just as in the physical world.

The I/O Bulb

Another interesting idea related to PVAs is the “I/O Bulb” (Underkoffler, Ullmer &
Ishii, 1999). The I/O Bulb is a conceptual design for a device in the shape of a stand-
ard light bulb which both projects images onto a surface area and captures visual phe-
nomena taking place on the same surface area using an in-built video camera. The I/
O Bulb is typically connected to a computing system that can be assumed to interpret
the video stream and change the contents of the projected image. Used in a classical
TUI setting where human agents interact with physical objects located on the project-
ed surface, the I/O bulb does not act as a PVA but as a two-way “Inter-World Event
Mediator” (see table 2, page 91) since the manipulation of the I/O Bulb per se is not
of interest for the agents. Instead, the role of the I/O Bulb in such a setting is to me-
diate phenomena from the physical world to the virtual (using the video camera) and
vice versa (using the projector function). It is the physical graspable objects (“Bricks”,
“Phicons”, etc.) that together with their virtual corresponding manifestations visual-
ised on the projected surface that act as PVAs in this kind of setting.

Object categories

Categorising objects on the basis of their roles or on their intended purpose in a given
activity context is useful both in modelling and design since it helps framing the in-
teraction taking place. Thus, many designers of physical-virtual systems have created
ontologies based on the application domain and technological limitations. Interest-
ingly, there seem to exist some kind of minimal but sufficient set of 3-5 object types
which reappear again and again under different names but representing roughly the
same kind of entity in many physical-virtual systems. What can be denoted “tools”,
“data objects”, and “storage objects” are three of the most common types. As an exam-
ple of an object ontology related to PVAs, Holmquist, Redström & Ljungstrand
(1999) distinguish between “tools”, “containers”, and “tokens” in the context of TUI-
like artefacts more or less conforming to the MCRpd model discussed earlier. In short,
*tools* are physical objects used for manipulation of freely selectable virtual objects, *con-
tainers* are representations of freely selectable virtual objects, and *tokens* are tailor-made
representations “hard wired” to virtual counterparts. The former two categories main-
tain a space- or time-multiplexed relation (Buxton, 1986; Fitzmaurice, 1996; Fitz-
maurice *et al.*, 1995) between the physical and virtual manifestation depending on the
time-span of the analysis, while the latter is an object type that falls rather safely into
the space-multiplexing category since the physical and virtual manifestations are con-
sidered to be “eternally” connected to each other. As Holmquist *et al.* (1999) point
out, it is suitable to give time-multiplexed physical PVAs manifestations a relatively neutral appearance since you never know what kind of virtual object the user of the system might link to the physical artefact, while space-multiplexed physical PVA manifestation gain on a design that signals what virtual object is connected to it at the other end.

The object categories used in the context of defining the situative model of physical-virtual space (see section 8.3, page 171, and table 10, page 178) and later for framing interaction in the physical-virtual environment prototype presented in chapter 10, is a domain-specific ontology for framing object relationships in knowledge work activities. Nevertheless, it reminds of Holmquist et al.’s (1999) ontology, as well as to the categories used by Beaudouin-Lafon (2000) in the context of his Post-WIMP interaction model. Holmqvist et al. (ibid.) also briefly discuss the possibility of objects changing roles (and therefore object category) depending on the use context.

**Cutting the PC in pieces**

Being one of the strongest critic against the tendency of the HCI community to take the PC/WIMP paradigm for granted, Bill Buxton (2001) stresses the potential gain of almost literally cutting the PC in pieces. “Super-appliances” like the PC could in many cases be replaced with more specialised and more tailorable interactive devices that could be used at different dedicated physical places, providing better support than general tools can do. Using the terminology of this dissertation, the specialised computational devices Buxton suggests (taking an Internet radio device and a portable MP3 audio player as examples) would be viewed as “virtual environment providers”. However, if the devices would have virtual corresponding manifestations in an environment offered by some other virtual environment provider such as a PC (e.g. a specific internet radio application and a specific MP3 music player application respectively) and if some of the attribute states of the manifestations would be automatically kept synchronised (e.g. the radio station history and “bookmark” station list would be kept identical automatically, and both the physical and the virtual MP3 player manifestations would have access to the same pool of digital sound content at all times), the devices would instead be better modelled as physical manifestations of physical-virtual artefacts. In relation to Buxton’s line of thought, the concept of physical-virtual artefacts can be viewed as a design tool that promotes the design of environments where strong specific computational devices are made to co-exist with weak general devices: physical-virtual environments that allow human agents themselves to choose between interacting directly with a strong specific computational device, or accessing a corresponding artefact through a more general virtual environment provider instead, depending on personal preferences, the task at hand, or the specific physical-virtual activity context. For example, if you are already browsing the web on your office desktop PC you might prefer to use the MP3 player situated in the PC virtual environment. Strolling around in town, the portable MP3 player would probably be
your choice. The different MP3 player manifestations use exactly the same raw data (having access to the same audio file library), provide almost the same functionality (e.g. the purely virtual manifestation might be capable of distributing the audio to more than one individual through the PC loudspeakers while the portable device is limited to ear-phones), but make their appearance in a completely different physical package.

**Embodied User Interfaces**

While not explicitly arguing for a flora of diverse computational devices as Bill Buxton does (see the previous section), Fishkin, Moran & Harrison (1998) pragmatically note the recent emergence of computing devices and interaction paradigms alternative to the PC/WIMP (e.g. PDAs and G/TUIs respectively), and view these as steps towards the design of “invisible” user interfaces:

“There tends to be a progression towards tighter embodiments, more directness in manipulating the intended object, and more coincidence between input and output. We believe that this reflects a progression towards a more real-world interaction style, where there is no perceived mediation, i.e. an invisible user interface.”

*(Fishkin, Moran & Harrison 1998)*

Fishkin et al. propose the design of interactive computational devices that physically “embody” virtual objects much like the PC display devices do, but where the physical casing is assumed to be portable (e.g. handheld) and used as a tailor-made input device for the very specific task the device as a whole is intended to be used for. Incorporating the design principle of letting the interaction with the device as far as possible mimic corresponding interaction in the physical world (i.e. if the device is for reading virtual books, turning pages should be supported in a way similar to the way pages are flicked when reading a paper book), Fishkin et al. denote this emerging interaction paradigm “Embodied User Interfaces” and present a framework for analysis and design of devices conforming to this paradigm. In another paper, Fishkin, Gujar, Harrison, Moran & Want (2000) discuss the implementation of three prototypes belonging to the paradigm. In short, these prototypes are handheld display-fitted computer devices augmented with touch and motion sensors enabling new ways of manipulation of the virtual objects “embodied” by the devices.

In the eyes of the conceptual framework presented in this dissertation, Fishkin et al.’s (1998, 2000) Embodied User Interfaces are examples of (task-specific) virtual environment providers (and not PVAs) since their physical manifestation has a time-multiplexed (Buxton, 1986; Fitzmaurice, 1996; Fitzmaurice et al., 1995) relationship to the embodied virtual objects. For example, at one point in time, the virtual pages browsed in the Embodied UI device might belong to a completely different virtual book than it does at some other point in time. Furthermore, the concepts of intra- and extra-manipulation (see page 179) can be used to highlight the extended interac-
tional properties of Embodied UIs compared to standard PC/WIMP interaction. For instance, one of Fishkin et al.’s (2000) prototypes makes use of “tilting” movements of the embodied UI device for navigating the virtual space visualised on its display. The inclusion of extra-manipulation in the language of interaction of course stems from the fact that Embodied UIs, as described by Fishkin et al. (1998, 2000), are relatively small and portable devices.19

The link and the synchronisation
The design of synchronised PVAs involves the development of some kind of infrastructure, or mechanism, for tying physical and virtual manifestations to each other (or at least making use of an existing such infrastructure). Since the basic idea behind the concept is to automatically keep related manifestations “up-to-date” with each other, the infrastructure needs to constantly be aware of relevant state changes (typically caused by human manipulation) on PVA manifestations both in the physical and the virtual world. Basic requirements for a fully functional synchronised PVA infrastructure includes the ability to:

- identify the manipulated PVA manifestation
- identify the state change caused by the manipulation
- implement corresponding state changes on all corresponding PVA manifestations

All three basic requirements pose a set of technological challenges of which some are discussed in sections 5.2 and 5.3, page 94 and onwards. As a consequence, the physical-virtual systems developed today tend to fulfil the above three requirements only in part.

Barton and Kindberg (2001) points out the potential use of synchronised PVAs for bridging the physical-virtual environment gap (although they use other words), and highlights the crucial role of the linkage mechanism:

“Once we can establish virtual/physical links, the physical and virtual entities may play a variety of roles in augmenting their counterparts across the physical-virtual divide.” (Barton & Kindberg, 2001)

Three alternative designs for tying a light bulb to a light switch are discussed:

- hard-wired physical light switch
- networked linkage infrastructure with a virtual (web-based) light switch
- networked linkage infrastructure with a physical light switch manifestation

18. Whether a relation is time- or space-multiplexed sometimes depend on the level of abstraction used for modelling. Are the pages of a physical book time-multiplexed or space-multiplexed in relation to the book that contains them?

19. Size matters. When the display weighs 20 kilograms and the PC CPU 10 kilograms, and both are attached to the wall socket, only very imaginative and physically fit interaction designers would consider using extra-manipulation of them as means for HCI.
The latter two alternatives are technologically very similar and form the basis for the infrastructure used in the CoolTown system (Barton & Kindberg, 2001) which basically links physical objects and web links to services offered by PCs, printers, radios, projectors etc. The physical-virtual infrastructure provided by CoolTown is typical in the sense that it focuses on the support for physical→virtual actions (see figure 18, page 93), and also because the PVAs that are kept synchronised by the system tend to share a very limited set of attribute states. For example, the state of the light bulb (whether it is currently on or off) is not reflected by the state of the light switch (a bar code tag) to which it is conceptually connected. This is not to say that such asymmetry in general makes PVAs less useful. For example, as noted by Ljungstrand, Redström & Holmquist (2000) in the context of discussing their WebStickers system, the use of physical objects as “tokens” for immediate access to specific virtual objects and places can be useful in all its simplicity.

"Using physical tokens to access, organize and share bookmarks to the web have a number of advantages. It enables users to integrate their bookmarks with their physical workspace, and thus place them in different contexts. The benefit of a rich context is that it supports the user with many cues of when a bookmark was used, where it is and how it is related to other resources." (Ljungstrand, Redström & Holmquist, 2000)

Many G/TUIs also make use of token-based PVAs where the “internal state” of the physical manifestations rarely is used in the interaction. It is instead the physical manifestation’s external state (i.e. the state that changes as a result of an extra-manipulation of the object, for instance moving it from one place to another) which is monitored. Also Streitz et al. (2001) make use of token-based PV synchronisation mechanism ("passage") in their design of collaborative buildings.

As a last example of a system with a strong bias towards physical→virtual action support and one-way causal relationships between physical and virtual PVA manifestations, the Magic Touch system can be mentioned. However, the physical-virtual design perspective proposed in this thesis that underlie this system is intended to push developers of physical-virtual infrastructure towards also considering causality in the virtual→physical direction, even though the technological challenges are considerable. Related work that consider a two-way causal relation between physical and virtual PVA manifestations is discussed in the next section of this chapter ("The physical-virtual design perspective" on page 54).

**Making existing PVA manifestations smarter on their own (and not just act as “tokens”)**

Although current implementations of PVAs tend to let physical manifestations act as means for “input” and virtual for “output” as described above, there are some attempts to at least conceptually design more symmetrical PVAs, i.e. PVAs whose manifesta-
tions share many attributes and whose physical and virtual manifestations have the capability of acting both as “input” and “output” mediators. One such attempt is the TransWorld system and interaction model (Ito, Fujita, Shimazu, Nakajima & Yamada, 1999) that, similar to the Magic Touch system presented in chapter 10, focuses on paper and electronic documents. Being a system based on off-the-shelf hardware products, the TransWorld system uses printers and scanners to achieve (semi-)automatic synchronisation between physical and virtual PVA manifestations (documents). Identification of paper documents is done through the use of bar-codes fitted on the documents. The underlying PVA model is that the virtual manifestations are the “originals”, and the physical ones are “avatars” representing the original virtual documents at a specific point in time. Thus, the model has a virtual-world bias although it tries to incorporate the support of object manipulation and observation (input/output) both in the physical and the virtual world and to some extent does incorporate the idea of supporting physical-virtual activities by providing a PVA infrastructure. Instead of introducing special “monitoring” devices and making small adjustments (e.g. attaching identification tags) to existing otherwise “dumb” physical objects of interest (as done in the TransWorld and Magic Touch systems), one can create the PVA infrastructure by embedding the majority of the monitoring and communication intelligence into the objects themselves. The MediaCup (Beigl, Gellersen & Schmidt, 2001) serves as a good example of taking this approach (see figure 4).

The MediaCup is able to (by itself) sense relative movement and temperature, and communicate such data wirelessly to other devices nearby and to sensors built into the walls. Location, acceleration and temperature information (referred to as MediaCup “context”) can thus be received by virtual environment providers such as PCs.

---

20. Whether PVAs support one-way or two-way causality can have effect on how we value the different manifestations. Virtual document manifestations are in many contexts considered more valuable than physical ones largely because of such causal asymmetry. A physical document can easily be generated if you have access to the virtual. Not so the other way around.
(Beigl et al. have even built a wrist watch that communicates with the cup) and used as a basis for representation of the cups in the virtual world, resulting in the emergence of PVAs where the cup acts as physical manifestation. As noted by Beigl et al. (ibid.), the fact that a fair amount of computing, sensing, and communication technology is built into the physical artefact itself, makes it more autonomous compared to physical objects that to a larger degree depend on external technology for maintaining the link with the virtual world (such as paper documents in the TransWorld and Magic Touch systems discussed earlier).

The concept of PVA presented in this thesis does not distinguish between PVAs whose synchronisation mechanisms are built-in, and the kind of almost “IT-free” PVAs that remain functional only if supported by computation power from somewhere else. This is because both kind of PVAs have identical interaction properties from the perspective of a human agent (if all necessary infrastructural conditions are met, i.e. as long as the tracking system in the room is switched on, as long as the IR-transmitters are close enough, and as long as the battery in the “smart” artefact is charged, etc.).

**Distributed physical-virtual artefacts**

Work that makes use of concepts similar to that of distributed PVAs (section 5.3, page 99) can be found in computing science literature in general and in work on distributed systems in particular. Other more interaction-related “distributed models” can be found in the areas of Collaborative Virtual Environments (CVE) and tele-presence research. Published efforts in all these areas have been sources of inspiration for developing the concept of distributed PVAs as an extension to the “standard” PVA concept. The most obvious difference between distributed PVAs and models proposed in much of the before-mentioned work is that it expands the modelling of distributed virtual objects (which tend to be the dominating focus) to include also physical (computing-augmented) objects. As highlighted in section 5.3, page 99, challenges within the physical-virtual design space discussed in this thesis, arising from distributing object access to more than one human agent (such as ownership, version handling and privacy issues) are more or less directly inherited from the virtual design space of the before-mentioned research areas. Since these problem areas are vast and already subject to much investigation, they are not discussed in detail in this dissertation, but merely pointed out (e.g. figure 25, page 103).

**Physical-virtual artefacts are here**

The definition of physical-virtual artefacts (definition 5.2.1, page 94) should be viewed as an ideal that artefacts can comply with to a smaller or larger degree. Treated in this way, the majority of physical artefacts manufactured and sold by the industry today can be modelled as PVAs since detailed information about them, or at least their manufacturer, tends to be available in the virtual world, e.g. in the shape of web pages.
This is true for small simple objects such as rubber bands and pens, as well as for more complex artefacts such as cars and buildings. The fact that the link between physical and virtual manifestations is typically not maintained and supported by some automatic mechanism (i.e. the link might only exist in the minds of human agents using the products), and the fact that the PVA manifestations do not typically “share” any attributes (which would make them synchronised PVAs, see 5.2.3, page 96), surely makes these physical-virtual products weak examples of PVAs but nevertheless PVAs.

I argue that the mission taken up by many researchers in the new explorative HCI research areas is precisely to transform existing weak PVAs to strong ones, and to develop new strong PVAs from scratch, although the activities are seldom expressed in such terms. I also propose that the more successful we are in completing this mission, the more insignificant the physical-virtual environment gap will be.

Stronger links between physical and virtual objects is not only beneficial for so-called “end-users” of products, but also for the production industry itself by improving business-to-business and supply-chain automation. This further strengthens the idea that the physical and the virtual worlds will in fact (and not just in theory) become increasingly integrated by an increasing number of increasingly tighter coupled PVAs.

The now wide-spread use of digital identification technology for automatic tracking of physical objects in transport might be the first embryo to a global physical-virtual infrastructure. Financed by the industry, with roots in the logistics business, the Auto-ID Center sponsors research for (no more, no less) the design of “an open-standards-based system that connects all physical objects to the global Internet” (Cole & Engels, 2002).

**Making corresponding PVA manifestations appear as unified**

The design of PVAs is not only challenging from the engineering point of view (i.e. the problem of making the linkage and synchronisation infrastructure work) but also from the perspective of how to make PVA manifestations reveal their existence when a corresponding PVA manifestation is given interest, as well as to make PVAs appear as (more or less) single entities for human agents interacting with them. I.e. a paper document and a corresponding web page should ideally signal that they are part of a specific PVA. This challenge is briefly discussed in the section “Bridging the gap through design — the environment designer’s way” on page 125 and onwards, in which unified appearance of PVAs is suggested to be accomplished by maximising the number of coherently presented shared attributes among the specific PVA manifestation. This rather straight-forward approach could of course be complemented with other more subtle design approaches such as for instance Janlert & Stolterman’s (1997) suggestions on how to design “computer artifacts with character”. In particular, their observation that certain object characteristics tend to be coupled, is very in-
teresting in the context of PVA design since it could prove to be a way to overcome the fact that physical and virtual manifestations can never be complete “mirrors” of each other although this would have been desirable from a “sense of unity” standpoint.

“A character is a unity of characteristics, in our definition. We believe that much of the usefulness of the character concept comes from frequent couplings of certain characteristics; e.g. ‘big’ is usually associated with ‘heavy’ and ‘slow’, and with good reason, if we look to the terms’ physical interpretation.”
(Janlert & Stolterman, 1997)

This insight could be utilised in PVA design practice. For instance, whenever an attribute or characteristic of a physical PVA manifestation cannot be directly transferred to its virtual manifestation (or vice versa) by a PVA designer (e.g. because of fundamental differences between the physical and the virtual world, see chapter 7), perhaps the virtual manifestation could be given attributes that can be assumed to be coupled (in the mind of the human agents interacting with the PVA) to the attributes that for one reason or another cannot be represented? To what extent coupled attributes can be used as substitutes for the “real” attributes in the context of PVA design, as described above, is an open question.

3.4 The physical-virtual design perspective

The physical-virtual (PV) design perspective (see page 166) is a design and modelling stance that acknowledges the existence of a gap between the physical and the virtual world (see chapter 6) when human agents perform physical-virtual activities (see chapter 5).

At the same time, adopting the PV design perspective involves to disregard the PV environment gap at times because human agents tend to do so, at least on a higher conceptual level. That is, although most people performing physical-virtual activities are aware of obstacles in bridging the gap, it is often such an inevitable and routine part of the job to do this extra “gap bridging work” (e.g. retyping the hand-written notes into the virtual document under construction) that the gap seldom is reflected on, but instead taken for granted when discussing a physical-virtual activity (such as the authoring of a text document) in the context of other activities. Only if the activity is studied in enough detail to reveal the very operations that are involved in it, or the conditions triggering the switch from performing actions in the physical world to performing actions in the virtual (the core phenomena of physical-virtual activities), the gap becomes something tangible, possible to analyse, and possible to discuss.

The PV design perspective is a child of its time in the sense that it follows the trend set by the new explorative HCI research areas discussed in the previous chapter, all aiming at incorporating the physical world as part of the design domain. In the following, some theoretically and practically oriented conceptual frameworks will be discussed in order to place the PV design perspective in context, frameworks that also in
one way or the other attempt to model physical and virtual activities, objects and/or environments in new ways. As the expanded scope of HCI has not yet found its shape, numerous alternative views and models have been proposed and it is not possible to mention them all and it is even harder to rank them as many of them are actually modelling the future rather than the present. The aim here is not to present a complete survey but to frame the work presented in this thesis.

**Practically oriented models**

Stressing the need for “a wider view of interaction”, Janlert (2003) criticises the world-view that typically underlies the classical HCI and HCI research areas and proposes a simple tripartite ontology for describing some important cases of interaction as computing technology becomes more pervasive and commonplace. The model distinguishes between six basic types of interaction (see figure 5):

- idea/information-idea/information interaction
- human-idea/information interaction
- human-human interaction
- human-object interaction
- object-object interaction
- object-idea/information interaction

![FIGURE 5. The HIOI interaction model (Janlert, 2003).](image)
At a high level of abstraction, the PV design perspective differs from Janlert’s (ibid.) model as the PV design perspective tries to treat physical and virtual objects as being the same kind of objects, while the HIOI-model makes a clear distinction: virtual objects belong to the ontological category “idea/information” while physical objects are “objects”. However at a lower level of abstraction, e.g. when it comes to the design of (infrastructure for) synchronised PVAs, the “object-idea/information” interaction type in the HIOI model (see figure 5) plays a crucial role for the PV design perspective because it is exactly the existence of this interaction type that make synchronised PVAs possible. Furthermore, human agent manipulation of physical and virtual PVA manifestations within the realm of the PV design perspective corresponds to the “human-object” and “human-idea/information” interaction types in the HIOI model respectively. Although “idea/information-idea/information” and “object-object” interaction (the two cyclical arrows at the bottom of figure 5) come into play as the concept of distributed PVAs is introduced (on page 99), these interaction types do not belong to the core of the PV design perspective because the issue of interest when it comes to physical and virtual objects is inter-world dependencies (relations between physical and virtual objects), not intra-world interaction (relations among physical objects in separation from relations among virtual objects). The last type of interaction included in the HIOI model, “human-human” interaction, is also left out of the PV design perspective because human communication aspects (i.e. social interaction) would make the design perspective much more complex, and like the case of intra-world interaction mentioned earlier, such issues are not unique for the kind of physical-virtual environments the PV design perspective is intended to model. To conclude one might say that The PV design perspective proposed in this dissertation is concerned with the interaction types forming the triangle (not the cycles) of the HIOI model and that it explores the possibility of conceptually collapsing the “idea/information” and “object” constituents, motivated by the assumption that some objects are manifested as both physical and virtual objects (e.g. a web page and the print-out of the same) and still viewed conceptually as one by human agents.

In their aim of designing better workplaces, Streitz, Tandler, Müller-Tomfelde & Konomi (2001) propose a human-centric model based on four design perspectives or spaces: the mental space, the information space, the architectural space, and the social space (figure 6). The PV design perspective proposed in this thesis is less holistic and focuses on the information and architectural spaces. Also, it does not involve group and organizational aspects of human activity but models single agents interacting with a set of physical, virtual, and physical-virtual objects. Although Streitz et al. (ibid.) stress the necessity of paying concern to all four perspectives, they underline the importance of the gap between the physical and the virtual world:

“In summary, we have to consider all four design perspectives or spaces [...] where the implicit distinction between real and virtual worlds plays a special role. [...] we argue for a two-way augmentation and smooth transitions between real and
virtual worlds. Combining them in an integrated design allows us to develop enabling interfaces that build on the best affordances of everyday reality and virtuality in parallel. As designers of human-computer interaction, or rather human-information interaction, and human-human cooperation, we want to use the best of both worlds.” (Streitz, Tandler, Müller-Tomfelde & Konomi, 2001)

More fundamental frameworks

The physical-virtual design perspective outlined in this dissertation does not make any claim to be a full-blown theory of HCI. It should not (at its present state) either be considered as a complete framework for design of physical-virtual environments. Rather, it is a point of view, supported by a set of conceptual tools, that I believe to have the potential of becoming such a design framework if more work is done on fine-tuning and evaluating the concepts.

The PV design perspective has not been grounded on any particular existing theoretical HCI framework but has mainly emerged from the design and informal evaluation of the Magic Touch system (described in chapter 10) and literature studies of other related systems. Should one like to connect the proposed PV design perspective to work of a more fundamental nature however, involving also the social and cognitive aspects which the proposed perspective does not touch upon, Distributed Cognition (Hutchins, 1995; Hollan, Hutchins & Kirsh, 2000), Situated Action (Suchman,
1987), External Cognition (Scaife & Rogers, 1996) and Embodied Interaction (Dourish, 2001) offer theoretical constructs that are in line with the concepts presented in this thesis and/or tries to capture similar interaction-related phenomena. They all acknowledge the strong dependency between human activity on the one hand and the physical (and virtual) environments in which it takes place on the other. Furthermore they are all, at least to some degree, able to model physical and virtual objects in a uniform fashion although some of them, just like the perspective proposed in this thesis, do not hide the fact that important differences exist.

3.5 Comparing the physical to the virtual world

The comparison between the physical and the virtual world presented in chapter 7 is best viewed as a case study from which general properties of physical and virtual environments are derived and proposed. Lists of properties such as the ones summarised in table 8, page 161 have been proposed by others as well. Two cases will be presented below.21

In the context of collaborative design activities (Arias, Eden & Fischer, 1997)

Strengths of physical media (pp. 4-5)
- Direct, naive manipulability and intuitive understanding
- Tactile interaction
- Mediation of communication and social interaction (in cooperative work settings)
- Some degree of fidelity to reality (in their urban design settings) because “boundaries of the physical are enforced”, you can’t place things abnormally
  Many of these strengths are interrelated.

Weaknesses of physical media (p. 5)
- The models are passive, incapable of changing representation without intervention by users
- Behaviour is not easy to visualize: All interpretation of meaning has to come from users
- Automatic feedback on the consequences of a decision is not provided
- Fidelity to reality is limited due to problems such as scaling
- Alternate realities are not easy to model — it is not possible to do actions that are not possible in the physical world

21. Being shown in brief and compressed form, and ripped out of their original context, any criticism and/or misunderstanding should be attributed to me rather than the original authors who have kindly permitted the altered (shortened) reproduction of their work.
Management of information is difficult. Results generated by the game (descriptions, evaluations, and prescriptions reached by the players) must be transcribed into some other form for posterity and future use. Information from other sources that needs to be brought to bear on the problem is not available in the physical model.

**Strengths of computational media (p. 5)**
- Dynamic
- Affords simulation
- Provide access to contextualized information

**Weaknesses of computational media (p. 5)**
- Computational systems are often opaque. “In the box.”
- Users are often forced to “work the computer” rather than being able to focus on the task
- The decentralized control (or the natural ability to contribute) that is possible to in the physical media is often lost

**Paper vs. digital documents (Sellen & Harper, 2002)**

_The Affordances of Paper and of Digital Technologies for Reading (table 6.1, p151)_

Affordances of Paper:
- Quick, flexible navigation through and around documents
- Reading across more than one document at once
- Marking up a document while reading
- Interweaving reading and writing

Affordances of Digital Technologies:
- Storing and accessing large amounts of information
- Displaying multimedia documents
- Fast full-text searching
- Quick links to related materials
- Dynamically modifying or updating content

_Serious constraints imposed by typical PC interfaces (p149):_
- Input is indirect (via mouse or keyboard).
- Input is (largely) one-handed.
- Both input and feedback rely mainly on visual cues.


Affordances of Paper:
- Controlling access until information is “ready”
- Rich though inconsistent indexing of files
• Reminding by “flicking through”
• Reminding through physical presence
• Portability of files for meetings
• Jointly viewing and marking up while in discussion
• Quick access to work-in-progress files

Affordances of Document Management Systems:
• Storing large amount of information in a small place
• Widespread access to information store
• Remote access to information store
• Fast exhaustive searching of information store
• Flexible, systematic viewing and sorting
• Quick links to related materials
• Dynamically updating or modifying content

The natural places for paper and digital tools (p206)

Paper is good for:
• Typical “point-of-use” knowledge work activities
• Social processes such as face-to-face meetings

Digital tools are good for:
• Activities that support “point-of-use” knowledge work activities, e.g. access and organization of information prior to the real use. One of the exceptions is the common knowledge work activity of authoring, which usually is performed using digital tools in conjunction with paper.
• Finalizing, polishing and storing information after its use.
• Managing work flow and distribution of information.

3.6 The situative model of physical-virtual space

As pointed out by several investigators (e.g. Malone, 1983; Kirsh, 1995; Harrison & Dourish, 1996; Sellen & Harper, 2002) the management of space is an important part of knowledge work and probably human activity in general. The situative model of PV space (page 171) is intended to model basic spatial configurations, and space-related human activity. The particularity of the model is that it does not make any distinction between physical and virtual space or physical and virtual objects, which it achieves by being centred around what is perceivable by a human agent at a specific point in time, and what kind of roles object play (see “Object categories” on page 46) in the context of a specific activity. It is also a very blunt model spatio-structurally speaking as it does not explicate any detailed spatial relationships between modelled objects apart from their level of accessibility at a given point in time from the perspective of a particular human agent. Objects are considered to be either a) not observable, b) observable but not manipulable, or c) observable and manipulable (see figure 45, page 174).
The model is inspired by Sellen & Harper’s (2002) use of the terms “hot”, “warm” and “cold” objects for rating the importance of specific objects for a specific activity at a specific point in time, and the fact that this object property tends to coincide with the accessibility of objects. In short: human agents tend to continuously (consciously and unconsciously) place the most important objects close at hand during the activity. One of the main purposes of the model is to act as a framework for studying the way objects enter and disappear from the observable subspace during the course of a specific activity. The approach to merge physical and virtual space makes the proposed situative model of PV space unusual. As becomes evident in chapter 8, it is not a straight-forward procedure (mainly because geometrical and topological space are not directly comparable except under certain conditions and by adopting certain modelling perspectives) and it has been hard to find similar attempts.

One model that reminds a bit of the proposed situative PV space model is the “Dimension Space” model by Graham, Watts, Calvary, Coutaz, Dubois & Nigay (2000) although the intention is not to use it for studying the way objects enter and disappear from the observable subspace but rather how objects already inside it compete for attention from a specific human agent performing a specific task at a specific time. The model is intended to be applied on all objects (both physical and virtual) that play a role within an activity at specific points in time, resulting in a graphical “plot” for each object. Figure 7 shows a plot for a needle in the surgery setting modelled by the authors. The plots are analysed and used as implications for (re-)design of the specific physical-virtual environment based on three rules (figure 8). Although the purposes of the Dimension Space model and the situative PV space model proposed in this dissertation differ from each other in several ways (including their respective purposes), they are similar in one important aspect: they deliberately incorporate physical and virtual objects within the same design space in order to view a set of physical and virtual environments as one single (physical-virtual) environment.

<table>
<thead>
<tr>
<th>Rule 1:</th>
<th>If at a given time, two or more entities in the system require high attention from an actor, then the system should be redesigned to permit the actor to give simultaneous attention to those entities.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rule 2:</td>
<td>If an actor is not aware of which parts of a mixed physical/virtual system are physical and which are virtual, then all instruments the actor applies to that entity should have the same effects on the entity’s physical and virtual components.</td>
</tr>
<tr>
<td>Rule 3:</td>
<td>If multiple actors are involved in an interaction, the system must be designed to support the coordination protocols used by these actors.</td>
</tr>
</tbody>
</table>

**FIGURE 8.** “Example design rules that can be checked with the help of the Dimension Space.” (Graham, Watts, Calvary, Coutaz, Dubois & Nigay, L., 2000)

22. A model attempting to capture also spatio-structural properties of physical-virtual space is proposed in chapter 9.
3.7 The hierarchical model of physical-virtual space

While the situative PV space model presented in chapter 8 offers a joint physical-virtual perspective on the process of object extra-manipulation in the course of human activity (specifically of objects moving in and out of the observable physical-virtual subspace) it does not provide any detailed information on structural properties of the physical-virtual space, e.g. semantical properties that could be attributed to objects on the basis of their spatial location (apart from being “hot”, “warm”, or “cold”), or causal dependencies between objects caused by natural (in the physical world) and artificial (in the virtual world) laws.

23. The concepts of intra- and extra-manipulation are explained on page 179.
The hierarchical model of PV space developed in chapter 9 is an attempt to provide a way of describing such structural characteristics of physical-virtual space. Any such effort has to be centred around a structural property that can be identified (in one form or the other) in both the physical and the virtual world. In this case the concept of “containment” was chosen.

The hierarchical model of PV space, just as the situational model, has found inspiration in studies of how humans seem to manage space in knowledge work environments (e.g. Malone, 1983; Kwasnik, 1989; Mander, Salomon & Wong, 1992; Kirsh 1995, 2000; Sellen & Harper, 2002). Other important influences include space-related studies of CSCW settings (e.g. Harrison & Dourish, 1996) as well as Herbert Simon’s strong advocacy of hierarchies as a necessary cognitive tools for humans to use in order to cope with almost anything more advanced than a cup of tea:

“The fact then that many complex systems have a nearly decomposable, hierarchic structure is a major facilitating factor enabling us to understand, to describe, and even to “see” such systems and their parts. Or perhaps the proposition should be put the other way round. If there are important systems in the world that are complex without being hierarchic, they may to a considerable extent escape our observation and understanding. Analysis of their behavior would involve such detailed knowledge and calculation of their interactions of their elementary parts that it would be beyond our capacities of memory and computation.”

(Herbert Simon, 1996, p207.)

Another source of inspiration in the development of the hierarchical model has been the technological ability of the Magic Touch system to actually support it.

**Structural and semantical aspects of space**

Having an interest in how structural properties of environments affect (and sometimes perhaps even mirror) human cognition, Kirsch (2000) uses the term “activity space” to denote what in this thesis often is referred to as a physical-virtual environment. He aims at explaining how structural properties of such environments interplay with cognition and the way structure at the same time affords and constrains human activity.

“We shall think of an environment as an activity space – originally a physical space but now virtual spaces qualify as activity spaces as well – populated with resources, tools and constraints in which an agent operates. [...] At the same time that an environment represents a space of possibility it also represents a set of constraints. An environment is the space in which structures are created and actions have consequences. It is the substrate in which new structural and meaningful configurations (situations) can be created and the substrate which constrains the possibilities of creation. Not anything can be created. The work environment therefore constrains both what it is possible or acceptable to do, and what happens
as a result of performing actions. It is partly the product of an agent’s projections and partly the product of underlying causal realities.” (David Kirsh, 2000)

The hierarchical model of PV space is based on causal relations that are typically coupled with containment relations such as the fact that contained objects normally follow the object containing them if it is moved spatially. However, also the way human beings divide the world into objects (being places, containers, tools or what have you) is reflected in the hierarchical model. The nodes in the hierarchies refer to particular physical or virtual objects that are considered to be “objects” in the first place by virtue of structuring the environment into entities and relations that are distinguishable and meaningful to the user, or at the very least, the modeller (designer). Objecthood depends on meaningfulness. Each node in a physical-virtual hierarchical model inherits semantical value from the object it represents (after all, it is a model) like the purpose of the object or place, what it stands for, when it is typically used, etc. The model also expresses important structural properties of an object’s immediate surroundings, properties that substantially governs how it would be affected by changes in neighbouring objects. In short, one might say that the proposed hierarchical model at least to some degree captures both structural and semantical aspects of physical-virtual environments as described by Kirsch (ibid.) in the quote above.24

Place vs. space

Harrison & Dourish (1996) stress the importance of the semantical properties linked to spaces or spatial regions in the context of CSCW, arguing that “places, not spaces, frame appropriate behaviour” where a “place” stands for, in short, a space invested with meaning. They argue that it is places, not spaces, we should design or at least design for.

“Space is the opportunity; place is understood reality.”

(Harrison & Dourish, 1996)

In theory, the hierarchical PV space model could be used for modelling simple “space” in the terminology of Harrison & Dourish (ibid.). In practice however, even the most meaningless physical or virtual environment tends to be connected to some meaning. Human agents find it hard to tolerate the existence of meaningless spaces and will work more or less hard in order to invent a purpose. Thus, any specific hierarchical model of a particular physical-virtual environment will be biased based on the intentions of the modeller and the kind of activity taking place and can never be “objective” and semantically “empty” in a strict sense. Even more so because (as is explained in

24. Structure and semantics are treated in this discussion as relatively independent entities. This is an oversimplification since structure and semantics (i.e. the way the world is organised and the way we choose to interpret the world) of course affect each other. See the quote taken from Herbert Simon earlier (page 63).
chapter 9) the modeller is explicitly given a certain freedom in “chunking” objects together or to ignore them completely at wish when constructing the specific model. However, since the resulting hierarchy is based on properties of everyday containment (which seems to be an influential structural property of space from the viewpoint of human agents both in the physical and the virtual world) it is likely to say something useful about the modelled physical-virtual environment to most observers even though specific nodes in the hierarchies might mean slightly different things to different people.

As with the situative PV space model in chapter 8, the main feature of the proposed hierarchical model is its ability to offer an integrated view (albeit coarse) on physical and virtual space. While the situative model uses a) physical-virtual object “activity logistics” and b) limitations of human perception, as its integrative cement, the hierarchical model is based on the everyday concept of containment, a phenomenon which seems to not only be identifiable in both physical and virtual environments, but also influential since it has strong causal and semantical effect on the way environments behave and appear.

**Space models**

Within some of the new explorative HCI research areas, the modelling of physical space (and place) has become an almost standard part of systems design. Ontology, mereology, spatial cognition, and other theoretical frameworks applicable to spatial modelling have become “hot” subjects although models used by current systems so far tend to be designed on more pragmatical premises (the one maintained by the Magic Touch system presented in chapter 10 is no exception), being more influenced by available sensor technology than theories on how to best conceptualise space. It is reasonable to believe that as physical-world sensor technology advances, so will the use of more advanced (and more “true”) models of physical space.25

For the time being, many models (including the one developed in this thesis) are based on topology although alternatives of course exist, at least in theory. David O’Sullivan (2000) provides a continuum from a geographic perspective (figure 9).

---

25. However, as can be concluded from skimming the literature on interactive systems implemented the past 10-15 years within the new explorative HCI research areas, it is striking how even very simple spatial models can be sufficient if combined with the right application. Typically, the success of the spatial model also depends on the system designers’ ability to avoid falling for the temptation of inferring “too much” about human activity from “too little” sensor data.
3.8 The Magic Touch system

Seeing Wellner’s DigitalDesk (Wellner 1993) as a starting point, there has been a continuous interest in merging the physical and virtual worlds in office environments and in more specialised settings (e.g. Arias et al., 1997; Mackay et al., 1998). Although the DigitalDesk system differs from Magic Touch (see chapter 10) in many ways, it has stood role model for many subsequent systems that do possess characteristics comparable to the ones offered by Magic Touch and I will therefore start by mentioning the most evident differences between DigitalDesk and Magic Touch.

Visual vs. conceptual integration

The proposed system Magic Touch offers its physical-virtual functionalities throughout the whole office space while DigitalDesk affords physical-virtual mechanisms on a desktop. Furthermore, DigitalDesk merges virtual and physical environments visually, while Magic Touch focuses on conceptual and indirect integration. E.g., when the effect of manipulating a PVA is communicated by the Magic Touch infrastructure to also affect a corresponding PVA, the change of state of the PVA is only observable by the human agent in case the PVA happens to be in the observable physical-virtual subspace of the particular human agent, i.e. if the PVA is visible on the computer display and the human agent is watching it. In the case of DigitalDesk (and many other classical AR applications) physical-virtual causality is more direct since physical and virtual activity take place in the same space also visually. From this perspective, visual “in-place” integration is a better approach than the conceptual approach underlying systems like Magic Touch since causality and proximity tends to go hand in hand in nature. However, at least with the technology of today, the visual integration approach does not scale well. To pursue the visual integration approach for larger spaces than single “display” areas such as in the DigitalDesk case typically means to either invest in putting or projecting displays “everywhere” throughout physical space, or to make human agents wear motion-tracked Head-Mounted Displays (HMDs) that overlay the real and the virtual 3D environments in real-time. Neither approach is really feasible in the context of office environments for various reasons that will not be discussed further here, although experimental attempts exist.26

26. The approach I find most promising with regard to the problem of providing ubiquitous virtual environment displays within physical space is through an (ideally) ergonomically and socially unobtrusive wearable display, visualising or in other ways displaying a virtual environment that not necessarily is visually overlaid with the scene of the physical world but which in a simple fashion provides feedback, indicating the effect ongoing physical actions have on relevant parts of the virtual world.
The object identification mechanism — visual vs. radio tags

Magic Touch does not provide any visually integrated physical-virtual environment but focuses instead on providing an infrastructure that enables related objects in the physical and the virtual environment to exhibit stronger causal relationships than what is normally the case. Successful identification of physical objects play a crucial role for this kind of systems. Among the object identification technologies available to date, apart from the challenging AI approaches in making computers recognise physical objects “as they are”, visual (e.g. bar code systems, glyph/ceiling-mounted video camera-based) and radio-based (e.g. RF/ID) technology are the most commonly used within the explorative HCI research areas. The basic functionality is the same (they can all act as real-world “eyes” for computing systems) and it is instead differences in aspects such as cost, durability, and scalability that make their application domains differ slightly. See Want, Fishkin, Gujar & Harrison (1999) for a discussion on some advantages and drawbacks of the two alternative identification technologies.

In both cases, small digitally unique identifiers, “tags”, are attached to (or directly printed onto) the objects of interest and are recognised by “tag readers” or “tag detectors” connected to computing systems as soon as a detector and a tag come close enough to each other. Furthermore, the visual tag approach sometimes requires additional actions from the human agent (e.g. to pass the bar-code detector over the bar-code). The following discussion will by and large ignore such details and treat the two tagging approaches (visual vs. radio) and variations of them as identical in order to not obscure, at least from an overall interaction viewpoint, more important design concerns.

Where to place the tag detector

One design decision that has a large impact on the kind of interaction that will be possible to support (specifically: when, where and to some extent how physical objects will be recognised by the computing system) is the choice of where to place the detector(s), i.e. at what physical location the linkage mechanism will be accessible. Very roughly, one can distinguish between four different kinds of placements of tag detectors:

- stationary, e.g. the tag detector is fitted to a cash register
- handheld (including display), e.g. the tag detector is fitted to a PDA
- handheld (without display), e.g. a handheld bar-code reader
- wearable (without display), e.g. the Wearable Activity Tracker in Magic Touch

The following brief discussion on some example systems related to Magic Touch will be structured accordingly.
Fitted on stationary device

By attaching the tag detector to a device or object which is tied to a physical location for one reason or another, the linkage mechanism will be available only at that particular physical place. Thus, this detector placement constraints the size of the physical-virtual space of interaction. For example, the Senseboard (Jacob et al., 2002) discussed on page 42 earlier in this chapter forces the physical-virtual document organising activities to take place in direct vicinity of the Senseboard. Although the lack of space for performing this collaborative activity in front of the board was probably not the only problem with the prototype, it played a significant part in why the participants finally chose the traditional (physical) way of performing the task. The advantage of fixing the detector is that more advanced tag detectors can be used (more exact, larger tag reading distances etc.) resulting in more reliable and controllable capturing of object identities. The approach can be expanded to cover larger physical spaces by installing networked tag detectors at several physical locations and/or stationary devices which creates a stronger feeling of a physical-virtual environment rather than a physical-virtual device. However, this approach is relatively costly and still only offers the linkage mechanism at very dedicated places.

Fitted on a handheld device (with display)

By attaching the tag detector to a device or object that is not stationary, the link between physical and virtual objects can be utilised in all physical locations to which the human agent brings the portable tag-fitted device (and of course, as long as there is a tagged object around to be identified). A good example of a system based on this approach is the infrastructure presented by Want, Fishkin, Gujar & Harrison (1999). They use RF/ID tags attached to everyday objects (as does Magic Touch) and RF/ID readers fitted on PDAs and tablet PCs. Typically, the virtual objects linked to the physical tagged objects are shown on the display of the portable device as soon as the device is taken close enough to a tagged object. Thus, the action taken to invoke the linkage mechanism is reversed compared to the case presented in the previous section in which the tag detector has been attached to a stationary device. In the stationary case, tagged objects are put close to the tag detector-fitted stationary device. In the portable (handheld) case, the device is typically put close to the tagged object.

Fitted on a handheld device (without display)

If the tag detector is integrated into a small enough device to be held more comfortably in one hand compared to typical PDAs or tablet PCs, e.g. pen-like devices, the precision in which the tag detector can be pointed to objects is increased and ergonomy is also improved. The drawback is that it is hard to integrate a sufficiently large display (and sufficient computing power although this is changing) in such a small device and thus, in most cases, any presentation of virtual objects linked to tagged physical objects has to be offered by yet another computational device connected to the graspable
tag detector through wire-based or wireless data communication, e.g. a stationary or portable PC with full-size display. At least if the causal effect in the virtual world, stemming from the identification of the physical object, should be signalled to the human agent in real-time. The WebStickers system (Ljungstrand, Redström & Holmquist, 2000) can serve as an example of this approach. The tag detector is in this case an ergonomically designed bar-code reader connected to a PC with a standard-size PC screen. Ljungstrand et al. discuss some social and cognitive aspects of using sticky notes such as Post-Its to which bar-code tags have been attached, and to which web pages have been assigned. The drawback of this approach is that if the display device is needed in the interaction and is not wearable, it will either occupy one hand (and since the tag detector is held in the other this means that both hands are required for interaction) which can be limiting in some cases, or it will be stationary (like the scenario presented in the photographs in the WebStickers paper (Ljungstrand et al., 2000) forcing the use of the linkage mechanism to take place at dedicated physical places.

An example of a slightly more complex graspable tag detector-based prototype is PaperLink (Arai, Aust & Hudson, 1997) which uses a small video camera attached to a colour mark-up pen for text and shape recognition. PaperLink (among other things) allows commands to be executed and web pages to be shown as a result of moving the augmented pen over pre-defined keywords or shapes. Thus, it is intuitive to link several commands and web pages to the one and same paper page much like web pages can contain a lot of links to other web pages, only now the anchor page is a physical one. This differs from the proposed use of WebStickers (and many other systems based on tagged objects) in which one link per object is standard.

**Wearable, “handsfree”**

Magic Touch explores the possibility of taking the miniaturisation one step further by making the tag detector always attached to the hands of human agents. Such a solution allows for less obtrusive use of the linkage mechanism since a) the detector does not need to be held by hand (hence the term “handsfree”) and b) the solution does not demand explicit human action in order to function. Indeed, Schmidt, Gellersen and Merz (2000) refer to this kind of interaction as “implicit HCI” and present an RF/ID-based wearable tag detector designed for the purpose. The RF/ID antenna in their prototype is fitted to a glove and connected to RF/ID reader electronics worn on a belt (see figure 10). The solution is very similar to the one used in Magic Touch although in Magic Touch, the tag detector electronics are worn on the wrist instead of the belt, and the most recent antenna design (the one used in PVA Manipulation Tracker version 0.53) is small enough to be worn as a “plaster” strapped around one of the fingertips (see figures 56 and 58, page 218). Both designs are examples of an approach that makes use of proximal sensing (Buxton, 1997) to the extreme in the sense that the information the system acquires about the state of the physical world is
acquired by sensor technology situated very close to the human agent of interest. The logic behind the idea of wearing the sensors is indeed that they will always be “where the action is”. This is not the case to the same extent for any other of the earlier discussed places for fitting tag detectors.

**Object location tracking**

If the physical location of the tag detector is known to a computing system, then also the location of an identified tagged object is known (under the condition that a short-range tag detector is used). This is the basic and quite unique principle behind the object location tracking performed by the Magic Touch system. Since the application area has been limited to an office environment (instead of a whole office building, a soccer field, or a city), a positioning system with relatively high precision could be used. Because of the wish to monitor larger spaces, most other systems linking physical and virtual objects in open space (i.e. not on a desk, not on a “board”, etc.) tend to have a much lower precision in tracking objects. One of the first such systems was the Active Badge system (Want, Hopper, Falcao & Gibbons, 1992) and it has been followed by many others. Typically such systems are limited to room or sub-room (e.g. ±50 centimetres) level of precision. Furthermore, the size of the artefacts that have to be attached to objects to be tracked tend to be much larger than the kind of paper-thin RF/ID tags used in Magic Touch, making it only useful to tag sufficiently large objects (paper documents are in general out of the question). Taken together, these two limitations of common object tracking systems (lack of precision and unacceptable size of the position transponders) make it more reasonable to track the position of a few human agents rather than thousands of small objects. Consequently, it is typically the position of human agents that are tracked, and at a room- or sub-room level (e.g. at the doorstep, in the chair, close to the window, etc.).

**FIGURE 10.** a) RF/ID tags, b) RF/ID coils (antennas), c) the wearable tag reader. (Schmidt, Gellersen & Merz, 2000)
One of the few systems that actually comes close to the precision of the position tracker used in Magic Touch is the Sentient Computing System which also (contrary to Magic Touch) is able to track relatively large areas (Addlesee, Curwen, Hodges, Newman, Steggles, Ward & Hopper, 2001). The size and cost of the location transmitter, the “bat”, is still quite large however (at least compared to RF/ID tags which is the corresponding artefact in the Magic Touch system) and Addlesee et al. (ibid.) do in fact mainly describe the use of the positioning system when used for tracking human agents, i.e. the classical position-tracking application.

**Ceiling-mounted cameras + visual tags**

One of the most common technologies for tracking many small objects with high precision in “room-size” physical spaces (i.e. larger than the two-dimensional area of a desktop, smaller than a whole office building) is visual tags (e.g. bar-codes or “glyphs”) in combination with one or more video cameras mounted in the ceiling (e.g. Rekimoto, 1996). However, for reasons explained in chapter 10, the RF/ID approach was chosen for the design of Magic Touch.

**PVA infrastructure**

Infrastructure for linking physical objects to virtual ones is one of the most common features in any kind of system that tries to integrate physical and virtual environment. As with for instance the system described by Kindberg et al. (2000), the infrastructure typically takes the shape of a web-based client-server data communication model which allows for upscaling and easy spatial expansion. Any networked computer can act as server (containing data and functional information about PVAs) or client (enabling physical-virtual environments). New clients, and sensors/actuators connected to them, can because of this distributed communication model be relatively easily included and excluded from the infrastructure in an ad-hoc fashion.

Magic Touch 1.0 was a single-user, single-environment solution and served only as “proof of concept”. Magic Touch from version 2.0 and onwards is however not different from most other current U/PC infrastructures and is client-server and web-based as well.

**Real-world model**

Magic Touch uses a very simple model of the physical world based on rectangular volumes in space that can contain each other according to some basic rules. This way of modelling the physical world is very common in other related systems as well. Figure 11 shows how a human agent enters and leaves the elliptical space (which in Magic Touch terminology would be called an “active volume”, see page 221) in front of the computer screen residing in the physical environment, and how the Sentient system models the situation based on containment relationships.
The main difference between Magic Touch and the Sentient system (and most other location-awareness featured systems) is the level of model granularity. Magic Touch keeps track of (a) a larger number of (b) smaller objects, within (c) a smaller physical space (but not as small as typical G/TUI systems). Because of this more fine-grained model of the world, the containment relationships between objects tend to become more nested, at least in the context of office environments. For instance, a paper document can be located in a folder inside a drawer inside a cupboard hanging on the wall which is part of a room. Humans do simply not fit into all places where small objects such as paper documents do. It is perhaps because of this fact that many of the “location-aware” systems have not been designed to offer any hierarchical visualisation of physical environments in the way Magic Touch does, although some cases exist (e.g. Brumitt & Shafer, 2001). The world of interest for these systems is simply too “flat” to motivate hierarchical visualisations. Instead, the most common visualisations are either variants of two-dimensional bird’s-eye views (similar to the one in figure 11) or three-dimensional Virtual-Reality-kind of visualisations (similar to the ones pictured in figure 80, page 249).

**Other related work**

The support for searching artefacts in physical environments, which is one of the possible applications of the proposed system, has some similarities with the InfoClip (Mori, Kozawa, Sasamoto & Oku, 1999) solution. However, InfoClip does not involve any centralised knowledge database for where artefacts are located.

The Kimura system (Voida, Mynatt, MacIntyre & Corso, 2002) explores ways to support the organisation of documents in an office environment, but focuses on collections of *virtual* documents (Voida et al., 2002). Some physical-world context awareness functionality is however also provided.
3.9 Summary

This chapter has discussed work related to the concepts presented in this dissertation and to the system that has been developed in parallel with these concepts, namely:

- The physical-virtual environment gap
- Physical-virtual activities
- Synchronised physical-virtual artefacts
- The physical-virtual design perspective
- Differences between the physical and the virtual world
- The situative model of physical-virtual space
- The hierarchical model of physical-virtual space
- The Magic Touch system
This chapter constitutes a platform and point of departure from which the rest of the thesis will lift off. Definitions of what I mean with the physical and the virtual world are presented. Motivation for accrediting the virtual world the “world” status, and what it means for a human agent to act and “be” in the virtual world is also discussed. Furthermore, the concept of human activity is narrowed down and described at different levels of abstraction, in order to facilitate modelling efforts performed later on in this thesis.

4.1 The physical and the virtual world

**DEFINITION 4.1.1:** The *physical world* is the world built of and containing matter directly perceptible to humans, and whose state is defined by arrangements of such matter in places, constrained by and modified according to laws of nature, within a geometrical three-dimensional space, at any time instant partially perceptible by humans through their senses.

The term “physical world” is used in this dissertation in its most naive sense, denoting the “real” world as it appears to most humans in everyday situations. In the following, when discussing physical objects, physical environments or the physical world, I do not refer to these phenomena as physicists who for instance could be interested in describing interaction between matter on an atomic or sub-atomic level. On the contra-
ry, the major part of this thesis will be concerned with the kind of phenomena that is directly perceivable by a human agent as a meaningful event in, or attribute of, the physical world, e.g. that a book has been moved from the desktop to the book shelf and that the book has the colour blue. Furthermore a deliberate distinction is made between the physical and the “real” world, where the physical world stands for the subset of the “real” world that, in short, is observable both by human and artificial agents. (A more thorough discussion on what is meant with observable phenomena can be found in the section “Human activities” on page 82.)

DEFINITION 4.1.2: The virtual world is the world built of and containing digital matter (bits) that after transformation into physical phenomena becomes perceptible to humans, and whose state is defined by arrangements of such phenomena in places, constrained by and modified according to (human-designed) laws of logic, within a topological multi-dimensional space, at any time instant partially perceptible by humans through displays (possibly multi-modal and audio- visually up to three-dimensional) built into computational devices residing in the physical world.

The use of the term “topology” in this thesis
The expression “topological space” is used to denote spaces in which the relationships between objects and places that define the space are not necessarily of a geometrical nature but something more abstract. The topological space of the virtual world consists of objects and places connected to each other through (potentially non-geometrical) “links”. This should be contrasted to similar access-relationships between objects in the physical world that always implies some kind of geometrical relationship. As will be discussed more in detail later on in this thesis, the topological structure of the virtual world can in part be transformed into a geometrical structure. This does however not affect the global topology of the virtual world which is inherently non-geometrical.

The virtual world (Cyberspace) is a world
With the advent of the now within many regions of the globe publicly accessible World Wide Web and the Internet, some people say that what we are experiencing is a peak in an information revolution that started with the diffusion of real-time distance spanning communication technology about a hundred years ago such as telephone, radio and later television. Notably since the early 1990’s and at the time of writing quantitatively and qualitatively progressing in a speed hard to follow, many types of computational devices (and the virtual spaces they provide access to) expand our possibilities to share and exchange information, creating a persistent virtual information space so huge, topologically massively connected, and culturally active and diverse, that it from many aspects can be compared to the physical world. It is the size, the global objective persistency (Janlert, 1995), the large amount of dependencies be-
tween different virtual environments, and the increasing breadth of possible and performed human activities within this digital space, that makes it natural to treat it as a coherent world, comparable to the physical.

In some literature, the term “virtual” is used also to denote the mental, inner world, often as contrast to the real, external world. As is implicitly expressed in the definitions 4.1.1 and 4.1.2, this thesis does, for reasons of simplicity, not consider any world that exists solely in an individual mind, created and maintained by human cognition. In short, my intention is to model what could be regarded as objective phenomena in the physical and the virtual worlds.

There are of course other possible “worlds” besides the physical and the virtual as defined above. The “mental world”, the “world of books”, the “world of politics”, the “world of science”, and the “world of entertainment” are all relatively common terms used to narrow down and to describe phenomena from a certain perspective. Notably, but without getting into a philosophical discussion about it, it is easy to see that many of these other worlds, just like the virtual world, are used to describe and understand the physical.

The intention is not to model all human activity. In particular, social, political, cultural and other high-level human activities are excluded. Many such activities can however be broken down into lower-level activities which in turn can be modelled in the framework outlined in this thesis. Such decomposition of course trade off coverage (i.e. full understanding of the activity) for model simplicity (i.e. at least parts of the activity can be identified and reliably interpreted). Since my approach in integrating the physical and the virtual world to some extent relies on automatisation, the gain in model simplicity is given priority. Figure 12 illustrates human activities in the physical and the virtual world with high-level activities on top, everyday physical and virtual activities separated from each other in the middle layer, and the rules that govern the two worlds at the bottom. It is the middle layer and the everyday activities that are situated in it which is the focus of this dissertation.

**The physical and the virtual world exist in symbiosis**

The virtual world depends on the physical world because:
- Most virtual phenomena are to a smaller or larger degree representations of phenomena in the physical world. The virtual world is (still) in large a world of information about objects and activities in the physical world.
- Bits, the elementary particles of the virtual world, are made of energy from the physical world. Without electricity, the physical power plants that generate it, and the network that distributes it, the virtual world would be literally off.

The physical world depends on the virtual world because:
- Many human “physical” activities in the post-industrial world do, when studied more carefully, depend on — or are linked to — processes in the virtual world. Sometimes these virtual processes can be substituted by a more “purely physical”
one but in many cases the substitute would be so inefficient and costly that it in fact would not even be considered as an alternative; e.g. taking an aeroplane to your overseas colleague for a chat instead of phoning her/him. Or perhaps an even more demonstrative example: to calculate a weather forecast by hand instead of using high-performance computers. Physical activities tend to become increasingly dependent and/or inevitably tied to virtual processes, processes of which some are perceptible to the person(s) involved, many of them not. The bottom line is that if we would hypothetically decide to close down the virtual world, fundamental re-arrangements of many of the structures that make up the modern physical world of today would have to be made.

4.2 Activities
Before discussing activities in the two worlds in more detail, a terminology that serves the purpose of this thesis is needed. The following rather “quick and dirty” definitions of activity-related entities have been influenced by how human activity has been described in literature on Activity Theory, Task Analysis, Cognitive Science (compiled knowledge), Situated Actions, OOA (UML and the like), physical symbol systems (Simon, 1996).
A simple terminology

- An **object** is a physical or virtual entity having a set of attributes that all are associated to certain values at any given time instance. Any alteration of the values within this set of attributes is considered a state change of the object. Except in the case of agents (see below), object attributes typically only change values when directly or indirectly exposed to changes outside the object (E.g. a football does not enter the goal unless someone kicks it there; notes on a paper document does not appear unless someone writes them there, etc.). This approach on modelling objects as rather simple symbol systems conforms with common views within the area of Object-Oriented Analysis (OOA).

- An **agent** is an exceptional object possessing the power to initiate an activity (see below) seemingly autonomously. Humans are agents, and so are complex networked software applications.

- An **activity** is a (more or less) goal-directed sequence of state-changes within a set of objects, initiated by one or more agent(s) and directly or indirectly controlled and monitored by this/these agent(s) during the full lifetime of the activity. It is usually up to the agent(s) that initiated the activity to decide when to end it. Some of the actions that the activity consists of might be independent from each other in which case it is up to the agent(s) to decide which action to perform when, i.e. to order the sequence. In other cases dependencies constrain actions to be performed in a certain order. Activities can also be interrupted and taken up again, on initiative from the performing agent or because of outer circumstances.

- An **action** is a clearly goal-directed activity carried out by one or more agent(s) in order to improve the status of an overall, higher-level, activity. Actions are always adapted to, and/or even driven by (Suchman, 1987) environmental circumstances in the given situation. This definition of action reminds of how the term “task” is used in classical HCI literature, however with a certain emphasis on the dynamics in which actions respond to temporary affordances and constraints (Norman, 1988) given by the environment in which the action is performed.

- An **operation** is a functional sub-unit of an action that is carried out automatically (in the case of a human agent using what is called “compiled knowledge” in Cognitive Science literature) without adaptation to potential changes in the environment. Operations are typically directed towards changing the value of one or a few specific attributes of one single object. Operations become actions when the environment fail to meet the requirements that the operation presupposes, causing “breakdown” (Bødker, 1989) situations where adaptation unexpectedly becomes necessary.

The definitions are explicitly intended to adequately describe activities disregarding if they take place in the physical or the virtual world. Furthermore, although the interest is primarily directed towards human activities, the quite general term “agent” will be
used instead of one of the more traditional HCI terms such as “user”, “actor”, or “human” in order to leave an opening for also the modelling of activities performed by complex non-human objects (e.g. software applications and robots).

Example: Peeling an orange

Let us study the case of a person having a dinner. In particular, let us assume that she or he is about to finish the dinner by having an orange as dessert. Using the above described terminology, this situation can be modelled as an activity “having dinner” initiated by a human agent consisting of a set of actions such as for instance “having first course”, “having second course” and “having dessert”. In this particular case, the action “having dessert” involves an orange (the object) and can be modelled as consisting of sub-actions such as “choosing the orange among others”, and operations such as “to peel the fruit”, “to divide a set of orange pieces to chunks suitable for chewing”, “putting them in the mouth”, “to chew”, “to swallow” in a sequence that can be viewed as the action to actually “eat the fruit”.

Operation, action or activity?

There are at least three factors that influence whether a certain studied phenomena is best modelled as an operation, action, or activity.

• Agent domain knowledge: If the human agent has peeled oranges before, the peeling process can be regarded as an operation. If it is the first time, cognitive and perceptual resources have to be actively engaged in order to succeed. In this case, the peeling process is to be regarded as a conscious action (or activity, depending on the modeller’s interest and the chosen level of abstraction, see below).

• Breakdowns: If the orange has an exceptionally hard skin, or if its pulp is softer and therefore more inclined to resolve into an unwanted mashy consistence, or if it turns out at an early stage of peeling to be inedible, special attention is needed. This breakdown situation (Bødker, 1989) transforms what would normally be a (simple) operation into a conscious action. The reverse transformation does of course also occur in the moment when an action has been performed repeatedly to the extent that the control of it is moved from the consciousness to the unconscious in a learning process, e.g. after having peeled a certain number of oranges.

• The interest of the modeller: Independently from the studied situation, the person analysing and/or modelling (henceforth referred to as “the modeller” or “the analyst”) might override the two factors given above by forcing a given phenomena to be categorised as an operation, action or activity simply because of interest. In particular, different levels of abstraction in the analysis tell different things about the ongoing interaction. For instance, if we would be interested in the peeling technique of the agent, it might be useful to regard the peeling as an activity, the tearing-off of a piece of the skin as an action and the movement of one of the persons hand away from the orange as an operation.
Throughout this thesis, the term “activity” will be consistently used whenever it is irrelevant to deliberately categorise a specific process as an activity, action or operation. In cases when it is important to distinguish this general use of the term with the more specific, high-level meaning of it, it will either be clear by context or it will be explicitly expressed.

### 4.3 Physical and virtual activities

**Intra-world actions**

The most elementary kind of interaction between an agent and an object is probably the case where a physical agent, such as a human, manipulates a single physical object. Actions and operations are constrained by natural laws and human culture, first learned and practised as a child.\(^{27}\)

Actions performed by virtual agents on virtual objects (e.g. an e-mail-filtering agent deleting unwanted e-mails from the inbox) are almost as straightforward. One difference with the virtual agent-virtual object situation compared to the physical agent-physical object situation is that action constraints in the virtual case to a large degree depend on human-designed rules, while the majority of object manipulation constraints in the physical case are inevitably inherited from laws of nature.

**Inter-world actions**

Complexity increases when any of the agents performs actions in the “opposite” world, that is the world in which the agent is not situated. This occurs when a physical agent manipulates a virtual object or when a virtual agent operates on a physical object. The latter situation is still quite rare but will most likely become more common in the future as physical environments become more “intelligent” and automatised. In any case, these inter-world actions do by necessity rely on objects that mediate events from one world to the other. These objects will be called Inter-World Event Mediators (IWEMs), in order to have a general name for objects commonly referred to as sensors, actuators, input devices, output devices, interaction devices, interface devices, user interfaces, etc.

Table 1 illustrates the relationship between inter- and intra-world agent actions. The remaining part of this chapter will be concerned with the top row of the table, i.e. the kind of actions performed by physical agents. Specifically, the focus will be on human activities that involve actions and operations in both the physical and the virtual world.

---

\(^{27}\) See Schank & Rieger (1974) for a list of primitives originally developed for their Conceptual Dependency theory but easily transferable to physical actions and operations on physical objects.
4.4 Human activities

As mentioned earlier (page 77) purely cognitive activities are not included as a third kind of activity besides the physical and the virtual. I acknowledge that this limits the model (modern psychology for instance stresses the important relationship between the physical world and cognition) but claim that even without that part, the model can say a great deal about possibilities and limitations for the undertaking of integrating the physical and the virtual worlds. In any case, cognitive aspects will not be ignored but considered and discussed in the context of physical and virtual activities. It would be interesting to investigate the inclusion of a “mental world” as a third leg for the platform of my model but time and space allocated for this thesis does not permit such an extension.

**Meaningful and observable human activities**

Many human activities, except for our basic needs like eating and drinking, are motivated by social and psychological reasons. (These reasons might be derived from our basic needs through a long and complex chain of causal relationships but that is not the issue here.) This holds both for physical and virtual activities. Thus, one might argue that social and psychological issues should be corner stones of any model intended to describe environments in which humans act. While being hard to argue against this proposition on a humanistic-theoretical level, it is easier to do so from a more practically oriented engineering point of view. It is, and will continue to be for a long time, hard to construct artificial agents and systems able to reliably interpret the often very subtle and sophisticated language used by humans to consciously and unconsciously communicate both psychological and social aspects of their activities. Many times it is hard even among humans to unambiguously interpret the signals. It is clear that background knowledge that can frame any given situation, that can put it into a context, is vital for understanding the meaning of human activity. Furthermore, physiological status (heart beat, breathing etc.) of humans that take part in an activity can help an agent (human or artificial) to understand why and what is going on.

Since the main goal is not to understand and describe human activities *per se* (this would inevitably lead into lengthy psychological, social and philosophical discussions) but to facilitate and support these activities in the light of my own and others’ recent

<table>
<thead>
<tr>
<th>physical world</th>
<th>virtual world</th>
</tr>
</thead>
<tbody>
<tr>
<td>physical agent</td>
<td>immediate manipulation</td>
</tr>
<tr>
<td>virtual agent</td>
<td>manipulation through IWEMs</td>
</tr>
</tbody>
</table>

**TABLE 1:** How physical and virtual agents act in physical and the virtual world. Inter-world actions (shaded cells) need to be mediated through Inter-World Event Mediators (IWEMs), e.g. input/output devices, sensors, etc.
findings in how the physical and the virtual world seem to work, I find it reasonable to limit the discussions to activities that, at least ideally, possess all of the following properties: 28

a) **Clear meaning.** The activity is perceived as *meaningful* from an everyday human perspective; i.e. it has a purpose and it is distinguishable from other activities. The activity “to write a letter” would own this property while “to move a pen over a piece of paper releasing ink” would not.

b) **Observable by a human agent.** The activity is (technically speaking) perceptible by a human agent other than the one(s) directly involved in the activity. An ongoing soccer game would fulfil the requirement while a coach’s decision to substitute a player during the game would not, until the coach communicates the decision to someone else.

c) **Observable by an artificial agent.** The activity is identifiable as a whole, and important dynamic properties of the activity and their effects on objects involved is possible to capture for real-time interpretation and/or “offline” storage using technology existing today, or technology that can be predicted to be available in the near future. The activity “to park the car” would satisfy the claim (sensors exist, and the interpretation is easy) while “to flirt” is an activity that probably will be hard to frame artificially even in the close future.

**Evolution of the sensing capability of artificial agents**

It is important to note that the amount of artificially observable physical and virtual phenomena is constantly increasing both in quantity and quality. New sensor hardware add physical phenomena that were not before perceivable by artificial systems (smell is one of the newer sensor “modalities”) or that was not economically viable before, and new theoretical models of human behaviour help bringing the data gathered from both physical and virtual sensors (virtual agents monitoring virtual activity) up to an abstraction level closer to the level where humans consciously reflect on their activities themselves.

**Evolution of the sensing capability of humans**

Human biological senses are limited, if compared to the immensely broad spectrum of physical phenomena measurable by the union of available and close-future artificial sensors. If sensors are packed in wearable units, and reach a critical mass of diffusion into the society, humans suddenly (if viewed from the time perspective of biological

---

28. Other similar approaches for narrowing down and distinguishing between “sensed” and “non-sensed” physical phenomena in the context of Augmented/Mixed Reality systems design have been proposed. For instance Benford, Schnadelbach, Koleva, Gaver, Schmidt, Boucher, Steed, Anastasi, Greenhalgh, Rodden & Gellersen (2003) suggest to distinguish between sensible, sensable, and desirable movements of physical objects. Their definitions of “sensible” and “sensable” phenomena correspond quite well to phenomena with property a) and c) listed above, i.e. phenomena with clear meaning and phenomena observable by an artificial agent respectively.
evolution) have mutated themselves into being able to see, hear and feel things that a couple of years earlier they could not. (We might have to charge our new digital senses during the night but so we do also with our biological body.)

**Sensing and making sense of activity**

Since the sensing capability of human and artificial agents is continuously improving, the number of activities we intend to model (everyday meaningful activities observable by both human and artificial agents) is increasing as well. Furthermore, activities regarded as being “everyday” activities come and go as society develops, adding even more dynamics into the set of activities to be modelled.

This dissertation will neither deal with practical issues regarding the use of sensors, nor will it present any model of human physical and virtual activity that could be directly applied for interpreting sensor data. Instead, the discussions are based on differences and similarities between the physical and the virtual world (chapter 7), and includes the proposal of a general model for describing basic physical and virtual human activity using a common terminology (chapters 8 and 9). It is my hope that the proposed model can help future designers to link whatever sensors they might have access to, to whatever fine-tuned application-specific model of human activity they have designed.

In the remaining part of this section on human activities, common physical and virtual activities will be described from an interaction design perspective. The purpose is to make clear how the interaction between the human agent and the object of interest takes place in each case.

**Human physical activities**

One might rightfully argue that human activity is more than a mere manipulation of physical objects. In fact, as our society develops, *mental* activities such as making decisions, finding solutions, and learning tend to become increasingly important. However, mental activity will continue to be accompanied by manipulation of physical objects until the day we can communicate our ideas and intentions among ourselves and to our (computer) tools using telepathy. Conscious physical activity will also be part of all human life until the day we are able, and want to, automatise the fulfilment of our bodily needs. Until that day, if it comes, the physical world will play an important role in our existence.

**DEFINITION 4.4.1:** A *physical activity* is an activity initiated and controlled in the physical world only, by one or more physical or virtual agents. (There might be state-changes in the virtual world as a result of the activity but these are not considered to be relevant for the agent performing the activity.)
For reasons of simplicity, the discussion will mainly be restricted to activities performed by one single agent. Although the discussion in some cases holds also for activities involving several agents, the increased complexity that the interaction between agents themselves lead to would need extra attention. Extension of the model to incorporate also multi-agent interaction is left for future work.

**Physical actions**

Figure 13 shows a human agent manipulating a physical object. The action involves the agent and the object of interest (the pentagon-shaped entity in the figure). The agent’s attention is fixed on the object of interest and the action is guided by the perceived changes to the object. Examples would for instance be to kick a football at a goal on a football field, or, to write down a shopping list on a piece of paper. In the first case the football acts as object of interest while it is the combination of paper and pen in the latter.

**Physical actions with virtual side-effects**

What if the football contains sensors that track its absolute geographical position and velocity, forwarding the data to a game statistics database? Or if the paper and pen-action digitizes the shopping list as you write so that you could easily print it out again in case you lose it, or forget to bring it with you to the grocery store? In both cases, the agent performs a physical action as described in the previous paragraph but since information about the action is captured and stored, the action can have effects at another time and (typically virtual) place, as pictured in figure 14. The term side-effect is used to stress that (at the time and place of the action) little or no attention is given by the agent to the potential virtual effects of the action. Many activities within context-aware applications fall into this category where people and things are continuously and unobtrusively tracked, allowing physical activities to take place “as usual”. The captured data is often used in a different con-
text and for a different purpose in contrast to the context and agent intentions that influenced the action in the first place. Sometimes the virtual side-effect would be useful to perceive at the time and place of the performed action but is not presented because of technological and/or cost reasons. (Perhaps soccer shoes displaying the speed of the ball you just kicked will be standard in the future?)

**Physical actions with virtual, hidden, intervention**

Sometimes we think we operate the world directly, while in fact, we do not. Modern cars contain computer-controlled mechanical systems for facilitating the somewhat stiff action of turning the wheels using the steering wheel. This support makes it possible for even the weakest driver to turn the front wheels of even the biggest truck without much effort. However, the degree of “help” from the system is best adjusted to the speed of the specific vehicle so that the escalation of the driver’s force on the steering wheel is lower at high speeds and higher at low speeds, for instance when parking. Some vehicles allow drivers to adjust the scaling function to their personal preferences. The interesting thing here is that the action that under normal circumstances seems to be direct and “purely” physical is in fact intervened by a virtual system. This becomes evident if we try to turn the wheels with the power system off.

**Human virtual activities**

**Human-Computer Interaction as a dialogue**

Classic HCI literature often models human virtual activity as an act of communication, or dialogue, between the virtual world and a human agent. Initialisation and control of the activity is performed by the human agent by manipulating dedicated physical “input” devices (e.g. a keyboard or mouse) while monitoring the effects that the state changes of a sometimes invisible corresponding virtual device representation (e.g. mouse arrow or text cursor) has on objects in the virtual world. The state of the virtual world is presented to the human agent by a physical “output” device (e.g. a visual display, loudspeakers) that transforms virtual states (such as the position of the mouse or text cursor within the virtual world) to physical phenomena such as light and sound perceptible to humans.
DEFINITION 4.4.2: A virtual activity is an activity initiated and controlled only in the virtual world by one or more physical or virtual agents. (There might be state-changes in the physical world as a result of the activity but these are not considered to be relevant for the agent performing the activity.)

Virtual actions

Using the terminology from section 4.3, page 81, the classical HCI setting with a “user” interacting with a computer can be described as an intra-world activity canalized through IWEMs. Figure 16 illustrates this setting. The white boxes represents the (obligatory) IWEMs. Examples of virtual actions would be for instance to insert a character in a paragraph while editing a text document using a computer keyboard and screen, or to dial a phone number on a phone that displays the numbers as you dial.

If the delay is short enough between the human operation of manipulating the input device, and the perception of the effect it has on the virtual object, the dialogue fulfils one of the most important properties of a “direct manipulation” (Shneiderman 1983) dialogue. Direct manipulation situations gives human agents the illusion of being in direct engagement (Hutchins, Hollan & Norman, 1986) with virtual objects, similar to the feeling of directness afforded by objects in the physical world.

Time multiplexed input

Normally, the final state of the physical input device, including its relationship to other surrounding physical objects (e.g. the position of the computer mouse in relation to a telephone handset nearby), is irrelevant for the human performing the activity. One of the reasons for this is that dedicated input devices often have a time multiplexed (Buxton, 1986; Fitzmaurice, 1996; Fitzmaurice, Ishii & Buxton, 1995) relationship to their virtual counterparts, resulting in a physical object (the input device) that represents different virtual objects and functions at different times. Because of the general inability of physical matter to change shape autonomously, these devices tend to have a relatively neutral shape and do not provide information about in what state the object to which it is connected is in.29

In the following, when a human is said to perform an activity in the virtual world, virtual environment, or virtual place, I acknowledge the fact that the specific human activity is mediated through interaction devices as described above, but claim that (if not explicitly noted) this indirectness has very little or no relevance for my discussion.
This is true as long as the interaction device is fairly suitable for the particular action and/or the human agent is used to the device to the extent that handling it is an (unconscious) operation. Just as it is reasonable to regard the virtual world as a world of its own, it is in many situations (and especially in the case when the activity is mediated through a time multiplexed input device such as a computer mouse) fair to see virtual activities as autonomous and separate from the physical phenomena that has to take place (e.g. the moving of the mouse) in order for the virtual to happen.

As illustrated in table 1, page 82, humans perform activities in the virtual world indirectly through dedicated IWEMs.

**Human-Computer Interaction vs. Human-Virtual Environment Interaction**

Within the HCI community, the term “computer” or “computer system” is often misleadingly used instead of the more appropriate “virtual world”, “virtual environment” or something similar. In this thesis, the term “Virtual Environment Provider” will be used to denote interactive devices designed to enable human agents to access virtual objects and places. The hardware is in this case viewed as a medium for human virtual activity. The motivation is twofold: a) computer-literate human agents are more concerned with what is happening within the virtual environment of such digital devices than with the interaction hardware involved, b) it makes it possible to regard virtual objects as being directly accessible to humans, much as physical objects are. The latter point opens up for new possibilities in modelling and designing combined physical-virtual environments, as will be explored in chapter 8 and 9.

### 4.5 Summary

This chapter has presented fundamental definitions that underlie much of the work presented in the rest of the thesis, such as what is meant with the physical world and the virtual world as well as the distinction between activities, actions, and operations. The chapter has also exemplified and defined various types of human physical and virtual activities, phenomena that will be extended to physical-virtual activities in the next chapter.

---

29. Many traditional mechanical switches announce the current state of the machine or device they are controlling (e.g. they have an “on” and “off” state just as has the entity they are controlling). Instead, newer kinds of switches for controlling light, air flow, etc. in modern buildings are often microprocessor-controlled, triggered by simple touch or movement in the vicinity. Or they are semi-mechanical in the sense that you do notice a state change while operating the switch but after the operation is finished you cannot tell in what state it is in. These switches do not show if they (or the corresponding devices) are on or off unless the designer deliberately add this functionality.
Physical-virtual activities, physical-virtual artefacts

The notion of “physical-virtual activities” is introduced to describe settings where human agents switch between the physical and the virtual world while performing an activity. The existence and potential roles of objects that are present in both worlds (“physical-virtual artefacts”) is also discussed. The idea of creating automatically synchronised physical-virtual artefacts is introduced as a possible design approach for bridging the physical-virtual environment gap (discussed at length in the next chapter). The overall assumption is that by technically inter-linking artefacts residing in the two worlds, the worlds themselves become not only conceptually integrated with each other (as in the mind of human agents sometimes) but also (potentially) causally integrated.

5.1 Human physical-virtual activities

Physical-virtual action pairs

Human activities, such as “writing a text document”, can be divided into a time-wise ordered sequence of actions, e.g. “adjusting a paragraph”, “proof-reading”, or “printing it out”. Within some of these sequences of actions we can identify pairs of action in which the human agent switches from one world to the other.
DEFINITION 5.1.1: A physical-virtual action pair consists of two actions belonging to the same activity and often time-wise adjacent, where the first action is constrained (by lack of action support in the current environment) or chosen (e.g. based on individual preferences) to be performed in the physical world and the other action is constrained/chosen to be performed in the virtual world, or vice versa.

Physical-virtual artefacts

Within physical-virtual action pairs we can sometimes identify one or several information-mediating objects that are subject to indirect or direct human manipulation within both actions, objects that transcend the physical-virtual border since they have presentations in both the physical and virtual world. Such objects will in the rest of the dissertation be referred to as Physical-Virtual Artefacts, PVAs, and for denoting the presentations of them in the two different worlds, the term PVA manifestation will be used. A text document presented in both the physical (e.g. printed on paper) and the virtual world (e.g. within a word processing environment) would serve as a good example of a PVA, where each manifestation affords different kinds of manipulation. The virtual manifestation is highly dynamic and is a perfect object for manipulation during extensive writing actions, the physical has its strength over the virtual when it comes to extensive reading actions, where ergonomic issues and ease of navigation are important. Both writing and reading actions are important components of most text authoring activities, making modern text authoring an inherently physical-virtual activity.

Although corresponding PVA manifestations by necessity never can be identical in all senses (bits are not atoms), they often share some user-manipulable and/or perceptible attributes. This is because humans, by disregarding insignificant differences between the manifestations, view attributes of a corresponding manifestation (and sometimes whole manifestations) as “the same” attribute or object. For instance, text document manifestations, being physical or virtual, are often understood as being the same as long as the text and layout are fairly identical. It is of course important to acknowledge that without this everyday and many times unconscious human cognitive process of “phenomena characterization”, we could never justly describe an object presented in both worlds as being “the same” object.

The difference between IWEMs and PVAs

Both PVAs and IWEMs are artefacts with the capability of reducing the gap between the physical and the virtual world. It is up to the modeller to choose whether to treat a specific artefact as a PVA or an IWEM. Just as phenomena can be modelled as activities, actions and operations depending on the modeller’s interest and the level of abstraction intended (see the distinctions made between operations, actions, and ac-

30. Physical-virtual artefacts are discussed at length in the next section of this chapter.
tivities on page 79 and page 80), whether to treat a certain object as an IWEM or a PVA depends to some extent on the intentions of the analyst. Table 2 shows some rules of thumb. In the case of human activity, IWEMs are normally used to manipulate virtual objects or virtual manifestations of PVAs, while a manipulation of an IWEM on its own cannot constitute a meaningful action. However, a manipulation of a PVA can be a meaningful action with or without mediation through an IWEM. As an example, the mouse input device is under normal circumstances best modelled as an IWEM since manipulation of it (clicking one of its buttons, moving it around), without tying it to a virtual object (or virtual PVA manifestation) is not a meaningful action. However, to write something down on a piece of paper or to edit a text document in a word processor environment (potentially using the mouse device) would be a meaningful action. Another difference between IWEMs and PVAs is that the state of IWEMs seldom are relevant after the activity has finished while the changes to PVAs often are part of or the whole reason for performing the activity in the first place.

**Physical-virtual actions**

Physical-virtual activities do not necessarily have to contain physical-virtual action pairs. Some PVAs allow for a second and more direct way of crossing the border between the physical and the virtual. If the two PVA manifestations are computationally linked to each other, a human agent can perform alterations on one manifestation while monitoring the result on the other in real time. The underlying computer system has the responsibility of mediating the alterations made, just like in the case when a virtual object is manipulated through an IWEM: a virtual action. The difference between virtual actions and the kind of physical-virtual action in which the physical PVA manifestation is directly manipulated by the human agent monitoring the result-

<table>
<thead>
<tr>
<th></th>
<th>IWEM</th>
<th>PVA</th>
</tr>
</thead>
<tbody>
<tr>
<td>from the perspective of the human agent</td>
<td>is an instrumental tool for facilitating manipulation of other objects</td>
<td>is the object of interest</td>
</tr>
<tr>
<td>functionality</td>
<td>general-purpose, limited modifications possible</td>
<td>tailor-made, often highly modifiable</td>
</tr>
<tr>
<td>part of the outcome of the activity</td>
<td>no, except for experience-motivated activities</td>
<td>often yes</td>
</tr>
<tr>
<td>causal relationship between the physical and virtual manifestations</td>
<td>one-way-directed causality is often satisfactory</td>
<td>ideally bidirectional; the result should be reflected in both manifestations</td>
</tr>
</tbody>
</table>

**TABLE 2:** General differences between IWEMs and PVAs. (However, IWEMs can often be viewed as PVAs and vice versa, depending on the interest of the modeller.)
ing state changes of the corresponding virtual PVA manifestation, is that in the former case, the state change of the physical object (the IWEM) is not considered meaningful while in the latter case, the state change of the physical object (the physical PVA manifestation) has a meaning in the context of the ongoing activity also after the action has ended. (In both cases, the state change of the virtual object is of course considered meaningful.)

**DEFINITION 5.1.2:** A physical-virtual action is an action on a PVA where both the physical and virtual manifestations are directly controlled and/or monitored by the agent.

A typical case is when one PVA manifestation is used for alteration and the other manifestation is used for evaluating/acknowledging the result of the performed alteration. Automatically synchronised PVAs – a precondition for physical-virtual actions – are still rare and it is hard to find everyday examples. One example where the virtual manifestation of a PVA acts like primary input-device and the physical as primary output device is when a disc from a computer-mounted CD-player is ejected by pressing the “eject”-button in the virtual environment (the case shown in figure 17). Both manifestations change state: the physical manifestation has the disc ejected, and any navigation button on the virtual CD player is typically dimmed to signal that for the moment no music can be played. The term “primary” input and output device is used since the nature of physical-virtual actions does somewhat blur this traditionally very clear distinction. In the case of the physical-virtual CD-player, both manifestations act like potential output devices since they both immediately reflect the state changes. The physical manifestation is considered as primary output device assuming that the intention of the action would be mainly to eject the disc (perhaps in order to listen to another CD instead), and not for instance to dim the virtual navigation buttons.
As an example of a case where the physical manifestations act like primary input device and the virtual manifestation as primary output device (illustrated in figure 18) could be the mediaBlock movie clip organiser (Ullmer, Ishii & Glas, 1998). Here the human agent decides the order in which virtual movie clips will be played by arranging physical mediaBlocks. Also here the input/output relationship between the physical and virtual manifestations is non-trivial since the mediaBlocks do not only serve as input devices but give information on how the movie clips will be played based on the spatial relationships among the physical mediaBlocks, i.e. as a group they serve as an output device as well.

**Redundancy is our friend**

I propose, that in the ideal case, the synchronisation of the two PVA manifestations is so powerful that it narrows the causality gap depicted in figure 12, page 78, and lets the human agent decide more freely\(^\text{31}\) whether to use the physical or the virtual manifestation as “input” device and whether to evaluate the action by studying the physical or virtual manifestation. Figure 19 illustrates this situation. Obviously, not many PVAs allow for such complete freedom today, or will in the near future, but I believe that it should be the goal for designers of PVAs and physical-virtual environments at large, to provide human agents with PVAs that get as close as technologically and economically possible. One of the big challenges we are facing if we want to facilitate human activities that cross the border between the physical and the virtual world, is to be able to increase and control redundancy at the same time. This problem will be discussed at length later on in this dissertation.

Having categorised and illustrated the — for the purpose of this thesis — most interesting cases of human activity involving both the physical and the virtual world, another class of activity, “physical-virtual activity”, is defined as follows:

**DEFINITION 5.1.3:** A physical-virtual activity is an activity consisting of an ordered sequence of actions containing a) at least one physical-virtual action pair or b) at least one physical-virtual action.

---

\(^{31}\) The term “freely” is used to describe a situation where all properties and support that technologically can be implemented in both worlds, also are made available to the human agent. Some properties and support cannot be implemented in both worlds because of fundamental differences between the worlds and will still constrain agents to perform certain activities in one world and not the other.
CHAPTER 5

Why do people perform physical-virtual activities?

As will be discussed in the next section on the physical-virtual environment gap, physical-virtual activities are associated with costs for the human agent performing them. So why then do people bother switching between the physical and the virtual worlds? Obviously, they gain something by doing it. And if the gain is bigger than the cost, they have made the right choice. The main reason for switching is probably the (predicted) better support for the next action within a physical-virtual activity, compared to the support offered by the current environment. E.g. if you know that the next step in your physical-virtual activity “to author a document” is to proof-read it, you might choose to switch from the virtual to the physical world because the support for that action is better in the physical world than in the virtual. Other reasons might be various kinds of cultural reasons (work organisation standards, traditions), or plain personal preferences (e.g. you prefer to go and give your colleague comments on the latest e-mail you got from her/him in person, rather than to send an e-mail).

Strengths and weaknesses of the two worlds in different contexts, motivating switching between them (and thus, motivating physical-virtual activities) is the topic of chapter 7.

5.2 A closer look at physical-virtual artefacts

The concept of physical-virtual artefact has already been used as a theoretical tool for discussing human activity at the border between the physical and the virtual world earlier in this chapter without a proper definition. As the concept will be used even more frequent in the coming chapters it is suitable to define the concept more formally.

**DEFINITION 5.2.1:** A physical-virtual artefact is an abstract artefact that (1) is manifested in both the physical and the virtual environment, where (2) these manifestations to a large extent utilise the unique affordances and constraints (Norman, 1988) that the two different environments facilitate, and finally (3) where one manifestation of a specific physical-virtual artefact is easily identified if a corresponding manifestation in the other environment is known.
“Environment” will be defined in more detail in chapter 8 (definition 8.3.5, page 175). For now it is enough to regard it as an observable part of the physical or the virtual world.

**PVA notation**

The following notation will be used to abbreviate further discussion on physical-virtual artefacts (PVAs): While PVA refers to both (or all if they are more than two) manifestations of a PVA (that is, the PVA as whole), _PVA_ refers to the physical manifestation of a specific PVA and _PVA_ refers to the virtual manifestation of a specific PVA.

**Synchronised physical-virtual artefacts**

Except for the third somewhat fuzzy requirement in the PVA definition 5.2.1, there are no requirements for PVA manifestations to “be aware” of each other. Thus, the classification of objects in the physical and the virtual world as being part of (a manifestation of) a PVA depends on the interest of a human agent performing a specific activity, or on the designer/modeller that is analysing a specific physical-virtual environment setting.

It is however easy to imagine PVAs whose manifestations are “aware” of each others existence, and perhaps more interestingly, automatically keep one or more attributes of the corresponding manifestations synchronised.

**DEFINITION 5.2.2:** A *shared attribute* is an attribute represented and automatically updated in all physical and virtual object manifestations that have the attribute. (A minimum of two manifestations are to have the attribute.)

---

32. This definition is slightly adapted from Pederson (1999b). The original term “instantiation” has been substituted with the more appropriate term “manifestation”. The original stronger expression “equivalent manifestation” has also been changed to the weaker “corresponding manifestation”. 

**FIGURE 20.** A visionary illustration of the potential benefits from bridging the gap between the physical and the virtual worlds using physical-virtual artefacts. (Pederson, 1999a)
An example of a shared attribute could be the textual content of a document, shared between both a physical (paper) and virtual (e.g. a web page) PVA manifestations. The representations of the attribute can differ between different manifestations, i.e. the temperature of a PVA might be represented as a number in the PVA. The definition regards the causal connection between the manifestations and not the way the shared attribute is presented to a human agent.

**DEFINITION 5.2.3:** A *synchronised physical-virtual artefact* is a physical-virtual artefact where at least one of its physical manifestations and one of its virtual manifestations share at least one attribute.

An example of a synchronised PVA could be the mouse pointing device whose manifestations (the graspable physical part and the virtual arrow) share the attribute “relative motion”.

**Motives for creating synchronised PVAs**

*Distribution and synchronisation gains*

Different environments in the physical and the virtual world provide unique tools and mechanisms for manipulating PVA manifestations and therefore the PVA as a whole. By enabling the possibility of situating PVAs in both physical and virtual environments, human agents can choose to act in the environment that best support a particular kind of object manipulation. Furthermore, if changes made to shared attributes of a synchronised PVA are automatically transferred to the PVA (or vice versa), physical-virtual activities will be possible to perform with less overhead (as explored more thoroughly in the next chapter).

*It is becoming possible*

While the possibilities for a computer system to sense phenomena and cause changes in the virtual world are huge, physical world sensing and actuation is in general much harder. However, although there is still a long way to go if compared to the possibilities in the virtual world, recent advances in hardware development can be viewed as potentially decreasing the gap. Examples of such PVA-enabling technologies are: 1) cheap, small interaction devices connected through wireless and wire-based networks, 2) new types of cheap and more precise sensors of physical phenomena, 3) emerging physical-world actuators such as electronic ink and tactile displays.
Interaction design challenges posed by synchronised PVAs

*Intelligibility and magic*

In the physical world, which is the world most of us have as a reference, cause and effect within the context of single-human activity, is almost always based on changes in proximity between objects in the observable part of the physical world. Even more so, natural (non-artificial) physical causality often requires objects to be in physical contact with each other.

The kind of “distributed causality” created by synchronised PVAs challenges human cognition by not (as default) revealing the logic behind the causal relationship between synchronised PVA manifestations and thus produces:

- Unpredictable effects: The result of manipulating a synchronised PVA is potentially less predictable than any “natural” causal effect.
- Unobservable effects: The result of manipulating a synchronised PVA might potentially have effect outside the observable part of the physical-virtual environment (in case the corresponding PVA manifestation is not observable).

Making “magic” causality as described above understandable is not a unique design problem for the construction of PVAs but an important design challenge in all research related to digitally “augmented” physical environments (Belotti & Edwards, 2001).

However, before discarding the idea of non-“natural” causal relationships between PVA manifestations on the basis of difficulties in making them understandable, one has to bear in mind that “magic” causality is not a new kind of phenomena but has in fact become part of everyday life in particularly the industrialised part of the world. Electrical light switches that show no physical connection with the lamps they control, systems that transmit picture and sound around the globe in real-time, medical pills that helps the human body overcome or avoid diseases, vehicles that move by pressing a pedal. Very few people reflect on the infrastructure that these magical phenomena rely on, but this does not prevent them from using them. As long as the promised effect is delivered, who cares?

**Technological design challenges posed by synchronised PVAs**

Clearly, the development of synchronised PVAs involve several technological challenges. Some of the most obvious will be briefly discussed below:

**Automatic update of PVA when a PVA has been altered**

As discussed in chapter 7, laws of physics constrain the possibilities for computer systems to modify physical objects.

> “Unlike GUI icons, phicons cannot spontaneously disappear or ‘dematerialize,’ cannot instantly change position or instantly morph into different physical forms...”  
>  
> (Ullmer, 1997).  
>  

"Unlike GUI icons, phicons cannot spontaneously disappear or ‘dematerialize,’ cannot instantly change position or instantly morph into different physical forms...”

(Ullmer, 1997).
However, recent research related to this topic show promising results for particular kinds of physical objects. Certain materials can be reshaped\textsuperscript{33}, and paper-like displays painted with ink made up of special chemical substances can change content\textsuperscript{34} just by sending electric signals to them.

**Automatic update of PVA when a PVA has been altered**

The challenge lies in recognising and digitising the alterations made on the PVA. Also here there is promising development regarding some kinds of alteration. Specifically pen-based activities can be tracked using touch sensitive surfaces (e.g. the Crossboard from IBM) or by tracking the movements of the pen itself (e.g. the Anoto Pen from Anoto AB).

The special case when a specific PVA is moved from one place to another can be regarded as changing the environment in which the PVA is situated.\textsuperscript{35} Manipulations of this kind can be recognised by one of many existing position tracking technologies set up in the physical environment.

**Automatic revision control as synchronised PVAs become consecutively altered**

Revision control is a research area on its own but deserves to be mentioned as it is likely to become an important mechanism in all synchronised PVAs whose complexity exceed the basic 1-1 physical-virtual manifestation design.

**Stability and security**

Human agents have to be able to trust the causal relationships that their physical-virtual activities depend on. The causality cannot be allowed to “disappear” or be altered without clear signals to the human agent whose activities involve the affected PVA. Since the average acceptance of software instability is very high compared to the expectations we have on physical artefacts and systems (in the everyday context that is the concern of this thesis), you could argue that human agents would accept physical-virtual instability to the same degree as with software systems. However, I believe that as physical effects of activities within physical-virtual environments increasingly depend on a stable and secure physical-virtual infrastructure, acceptance of system failure will decrease to levels similar to level of failure-acceptance of purely physical artefacts and systems. Physical-virtual causality has to be a more reliable magic than virtual causality.

\textsuperscript{33} E.g. Microelectromechanical Systems (MEMS).
\textsuperscript{34} E.g. http://www.e-ink.com.
\textsuperscript{35} The manipulation of objects contained by other objects will be modelled in chapter 8 as intra- and extra-manipulation (page 179).
The general problem of transferring object attributes across the PV border

The PVA1s and PVA2s that together form a physical-virtual environment do to some extent possess attributes that for the corresponding manifestation has no meaning/cannot be represented/cannot be captured. It is fair to say that at the time of writing, only a small subset of all the attributes (and the states they can take) that the physical and virtual worlds afford, can be successfully recognized and represented in both worlds. It is one of the goals of my research efforts (and other researchers within areas such as Augmented Reality/Mixed Reality, Ubiquitous/Pervasive Computing) to increase the size of this subset, by means of IT infrastructure.

5.3 Distributed PVAs

So far, the focus of the discussion has been on synchronised PVAs in their simplest form, having one physical manifestation, one virtual manifestation, and some kind of physical-virtual infrastructure that guarantees synchronisation of attribute values between these manifestations. This situation is depicted in figure 21, where the PV infrastructure is symbolised by the large umbrella-like graphical object. The PVA concept can be used to model small simple artefacts such as a computer mouse, as well as large complex systems like a complete office building.

FIGURE 21. A simple synchronised physical-virtual artefact (PVA). The PVA1 object is the PVA’s single physical manifestation, the PVA2 object is its single virtual manifestation. The large umbrella-like area that connects the PVAs’ physical and virtual spaces represents the physical-virtual infrastructure that keeps the shared attributes of PVA1 and PVA2 synchronized.

Multi-manifested PVAs

Another direction in making the simple form of PVA pictured in figure 21 more powerful (in the sense of being able to model more kinds of objects in the two worlds) is to explore the possibility of PVAs having more than two manifestations. Sometimes there is more than one manifestation in each world that, for semantic reasons or other, are all best modelled as manifestations of the same PVA. An example might be an an-
annual budget report printed in 10000 full-length copies and 20000 short-version copies added to a multitude of virtual versions available on the web. It might be relevant to model all 30000+ versions of the report as manifestations of the same PVA “annual report 2003”. Figure 22 illustrates a synchronised PVA consisting of eight manifestations.

FIGURE 22. A multi-manifested synchronized physical-virtual artefact.

Not all manifestations have to be situated at the same physical (for the PVAs) or virtual (for the PVAs) place. Figure 23 shows the same “multi-manifested” PVA as in figure 22 but with cracks denoting that manifestation PVA3 is located at a different physical location than PVA1, PVA2, and PVA4; and that PVA8 is located at a different virtual location than PVA1, PVA2, and PVA3. Thus, the PVA shown in figure 23 can be denoted a “distributed multi-manifested PVA”.

FIGURE 23. A multi-manifested distributed synchronized physical-virtual artefact.
Now, imagine that $PVA_3$ is manipulated by a human agent. Figure 24a shows how the resulting state change of the manipulation is captured by the local PV infrastructure and is propagated to the PV infrastructure one step higher up in the “PVA hierarchy”. The higher-level PV infrastructure recognises the state-change (figure 24b) event and propagates it to all manifestations that belongs to the high-level PVA. The manifestations update themselves (figure 24c) and take any other appropriate actions based on the state change event they have received. Some manifestations might ignore the state change because they do not share the attribute that was changed, some might simply update their attribute and stop at that, others might generate a new state change that could be propagated up in the “PVA hierarchy” again.

**Synchronised PVA manifestations shared among human agents**

So far, the discussion has focused on physical-virtual artefacts manipulated and owned by single human agents. As with all distributed and networked system, complexity increases when distributed objects such as “multi-manifested” PVAs are shared among several human agents. However, since this kind of object sharing obviously is common in collaborative settings, modelling and infrastructure design efforts for physical-virtual environments such as the one presented in this thesis need to be able to support it. Figure 25 illustrates a situation where three PVAs owned by Anders, Ulrika, and Jonny (let us denote them $PVA_a$, $PVA_u$, and $PVA_j$ respectively) are all causally dependent on each other:

- $PVA_a$ and $PVA_u$ have manifestations that share attributes, namely $PVA_{a7}$ and $PVA_{u2}$.
- $PVA_u$ and $PVA_j$ also have manifestations sharing attributes: $PVA_{u6}$ and $PVA_{j7}$.

To make the situation more concrete we can imagine that $PVA_a$, $PVA_u$, and $PVA_j$ are all PVA’s assigned to current projects that Anders, Ulrika and Jonny are involved in respectively. The manifestations of interest that share attributes ($PVA_{a7}$, $PVA_{u2}$, $PVA_{u6}$, and $PVA_{j7}$) could be clones (copies) of the one and same information document about a product. In this hypothetical example, Anders’ role is that of a product developer who has been asked by the marketing director Ulrika to adjust the documentation of the product so that it incorporates all the new added features of the upcoming release of a new product version. She is at the same time with a potential buyer (Jonny) of the soon-to-be old version of the product.

What should happen when Anders manipulates one of the manifestations of his PVA ($PVA_a$) in a way so that it triggers an update of $PVA_{a7}$? The simple and default answer would be that the state change of $PVA_{a7}$ (belonging to Anders) triggers a state change in $PVA_{u2}$ (belonging to Ulrika) that in turn (depending on the logical configuration of the PV infrastructure in $PVA_u$) triggers a state change of $PVA_{u6}$. That state change in turn triggers a state change in $PVA_{j7}$ belonging to Jonny. To summarise:
FIGURE 24. A human agent manipulates PVA₃. (a) The local PV infrastructure captures the state change and sends an update event to the PV infrastructure higher up in the “PV infrastructure hierarchy”.

(b) The higher-level PV infrastructure recognizes the state-change event and propagates it to all manifestations that belongs to the PVA.

(c) The manifestations update themselves and take any other appropriate actions based on the state change event they have received.
among other things, Anders’ manipulation of his PVA\textsubscript{a} has effect on Jonny’s PVA\textsubscript{j}, because of the fact that Ulrika’s PVA\textsubscript{u} has manifestations that share attributes with manifestations of both Anders’ and Jonny’s PVAs.

- What if Anders is not aware of the fact that Ulrika’s PVA is also (in part) shared with Jonny? What if he in fact would not like his manipulation to be propagated to anyone else but Ulrika? (In the concretisation earlier, the developer Anders and certainly the marketing director Ulrika would not like Jonny to get information about the upcoming product enhancement at this stage of negotiation between Ulrika and Jonny.)

- What if Jonny has spent a long time setting up his PVA the way he wants it. Does he really like the change that is caused by Anders? What if the state of Jonny’s PVA implies life or death? Anders might not even know this person Jonny. (This

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure25}
\caption{Authorization challenges caused by PVA manifestations shared between several human agents. When should automatic state change propagation be allowed, and when does it need to be approved by someone?}
\end{figure}
case does not fully apply to the hypothetical situation outlined earlier, but can easily be imagined.) Answers to the questions above depend on the situation. The purpose of mentioning them here is just to show that a lot of issues other than technological emerge when synchronised PVAs belonging to multiple human agents share attributes. Conflicting interests of their owners, and differences in the context in which synchronised PVAs and PVA manifestations appear, create a challenging design space for shared distributed PVAs. This design space is already at the virtual side explored by research in Distributed Systems, Computer-Supported Cooperative Work, and Collaborative Virtual Environments. The question is, how does the presence of physical synchronised and distributed PVA-manifestations affect these existing research efforts in creating a safe and sound synchronisation infrastructure? Is it perhaps so that, as indicated by the scenario presented above, synchronisation of physical PVA manifestations is a more socially delicate issue compared to that of virtual PVA manifestations? These questions are left for future work to answer.

5.4 Summary

This chapter has defined two core concepts of the physical-virtual design framework proposed by this dissertation: the concepts of physical-virtual activity and physical-virtual artefact. Physical-virtual activities have been discussed in the light of consecutive physical and virtual actions, or physical-virtual action pairs, while the concept of physical-virtual artefact has been elaborated upon discussing shared states, synchronisation, and distribution in space.
The physical-virtual environment gap

The overall aim of this chapter is to analyse and define the gap between the physical and the virtual world from a human-computer interaction perspective. The general question is: Why cannot objects and humans travel without resistance between physical (real) and virtual (digital) environments? The mere existence of the gap, and the various shapes it takes, constitute main motives for pursuing the design approach proposed in this thesis. This chapter provides some evidence for the existence of the gap, takes a closer look on the gap from different angles, and describes ways in which it can be overcome.

6.1 Introduction

The modern human increasingly uses IT devices such as PCs, PDAs, cellular phones and other digital machines fitted with dynamic digital displays for everyday purposes. An observation of high relevance for this thesis is that although these digital machines together create and maintain a world that is strongly connected to the physical world when it comes to the “meaning” of presented objects, this virtual world is weakly connected to the physical in terms of directly perceivable causality. That is, although a specific pair of physical and virtual phenomena describe and represent roughly the same thing from the viewpoint of a human agent (e.g. a web page and a physical print-out of the same), changes to either would normally not affect the other automatically. The combination of similarity among roles that physical and virtual objects play on
the one hand, and lack of causal connection between the worlds on the other, will re-
erferred to as the “physical–virtual environment gap” and it is the purpose of this chapter
to further investigate its nature and importance.

Pragmatically one can say that two (or more) objects play the same role, or corre-
spond to each other if they are used to facilitate the same kind of human action (tool
similarity) and/or if they represent the same abstract entity (information correspond-
ence) within a specific human activity. It is in fact the set of physical and virtual objects
that play similar roles and/or correspond to each other within the one and same hu-
man activity that define much of the design space discussed in this thesis. Table 3
shows some examples. The proposed approach involves to reduce the physical–virtual

<table>
<thead>
<tr>
<th>tool similarity</th>
<th>physical</th>
<th>virtual</th>
</tr>
</thead>
<tbody>
<tr>
<td>pen &amp; eraser</td>
<td></td>
<td>word processor</td>
</tr>
<tr>
<td>tool similarity and information correspondence</td>
<td>paper document</td>
<td>web page</td>
</tr>
<tr>
<td>information correspondence</td>
<td>a specific car A</td>
<td>the instruction manual to car A</td>
</tr>
</tbody>
</table>

**TABLE 3:** Examples of objects that play similar roles and/or correspond to each other although they are situated in different worlds.

evironment gap by complementing similarity- and correspondence-relationships be-
tween physical and virtual objects with *causal* relationships. In addition, the approach
also leads towards designing new object manifestations into environments in which
the particular objects have not yet been manifested before. What could a virtual sta-
pler be good for?

I believe that humans spend an increasing amount of time and energy on bridging
the PVE gap by manually linking objects in the physical and virtual world together,
work that we as designers should try to automatise. Common “bridging” activities in-
volve (among others) the searching, identification, transformation, double-checking
and updating of objects, when switching from the physical world to the virtual world
and vice versa.

Questions arise such as:

- Why is it so natural for many of us to make a clear distinction between the world
  jointly presented to us by IT devices and the “real” physical world?

- Is it reasonable to even treat this collection of virtual “environments” as a coherent
  *world* on its own and by doing so ascribing it many attributes of the physical
  world which it probably never will possess?

- What exactly defines the border between the physical and the virtual world?
What makes the physical and the virtual world different, and what makes them similar?
How big is the problem of bridging the physical-virtual world gap in human everyday activities?
Can the gap be reduced by introducing new IT?
This chapter will address these questions.

6.2 Origin and dimensions of the gap
Most people with some experience of activities in the virtual world would intuitively agree that the physical and the virtual world are fundamentally different. (Perhaps they need a short introduction to the idea of seeing the virtual world as a world but that is another issue.) They would also agree that switching between the worlds is seldom without some extra efforts although they might not always reflect on it since the cost is taken for granted, “It has to be like that”, or because the amount of extra work seems small and insignificant. As engineers and human-computer interaction researchers, we know that there are many ways to build interactive systems and we also take it as an important part of our work to challenge previous designs. The design I challenge in this thesis is the design of interactive systems that separate the physical and the virtual worlds from each other. Some parts of the two worlds are hard to integrate, perhaps even impossible; other parts we would not like to integrate, or we don’t see the point. But, I believe there are parts of the two worlds that can be integrated, and among these, some that should be integrated motivated because of the potential gain for humans acting within the physical-virtual system, and because now, or very soon, there is technology to do it. Following the tradition from the late industrialisation phase to automatise routine work, time has come to automatise routine work performed at the border between the physical and the virtual worlds: the manual bridging of the physical-virtual environment gap.

DEFINITION 6.2.1: The physical-virtual environment gap is a name for the discrepancy between a) the strong conceptual connection between objects that play similar roles, or represent the same abstract entity, within the framework of a specific human activity, although they are situated in the physical and the virtual world respectively, and b) differences in behaviour as well as weak or lacking causal dependencies between these physical and virtual objects.

The physical-virtual environment gap creates a tension that in one way or the other has to be overcome when performing physical-virtual activities. Either the human agent takes extra steps in order to bridge the gap (e.g. retypes the text when paper-based information is to be transformed into digital form), or there is an underlying computer infrastructure that automatically reduces the “width” of the gap (e.g. the paper containing the information is automatically identified as available in electronic
form on Internet, and the system immediately retrieves and displays it in the virtual world when the human agent switches from the physical to the virtual world). The first solution is the standard today, while the second situation is something that has emerged as a possible alternative (or perhaps more realistically speaking, a complement) only recently. Sensors can inform the system about an ongoing activity before a switch between the physical and the virtual worlds takes place, so that when the agent enters the other world, it is all set for a continuation of the activity with minimal overhead effort needed from the side of the agent. Of course, automisation of this kind will for various not entirely eliminate cognitive overhead caused by the world switch, including the fact that the two worlds present themselves differently and that certain effort is needed to adjust one’s behaviour accordingly.

**DEFINITION 6.2.2:** *Physical-virtual activity overhead* is a comprehensive term for irrelevant, redundant and/or distracting actions directly imposed on agents by the physical-virtual environment gap, actions necessary in order to successfully continue the execution of a physical-virtual activity.

The term physical-virtual activity overhead is intended to be interpreted and used similar to how “cognitive overhead” is used within cognitive science; denoting the need for additional resources when solving a particular problem, usually having a negative effect on the activity. The cognitive overhead demanded by physical-virtual activities is an integrated part of the physical-virtual activity overhead since all human physical and virtual activity implies also cognitive activity.

**Gap dimensions**

When studied more closely, the physical-virtual environment gap can be viewed as a collection of differences between the physical and the virtual world, differences that in the light of a given physical-virtual activity are obstacles that have a negative effect on the performance. Each “difference category” or *gap dimension* measures the gap from a particular perspective and the idea is that by evaluating a specific physical-virtual environment along each gap dimension (assigning a value in the range from “very similar” to “very different” to the physical-virtual environment based on the different properties of the physical and virtual world in the specific setting and along the specific dimension), we get a reasonably good estimate of how big the physical-virtual environment gap is in a specific physical-virtual environment for performing a specific physical-virtual activity. Four important gap dimensions are:

- **Presentation**: The degree of difference between how corresponding physical and virtual objects/event are presented.
- **Manipulation affordance**: The degree of difference in what properties of objects that can be manipulated by human agents, and in how the manipulation is done.
- **Organisation/navigation**: The degree of difference in how collections of objects are organised in places and how you navigate the space.
• **Causal dependency:** The degree to which activity-relevant state changes in one world automatically has effect in the other, i.e. how strong the causal dependencies between objects and events in the two worlds are.

To measure the width of the physical-virtual environment gap for a specific physical-virtual activity is not only about summing the degree of difference found in each dimension\(^\text{36}\), but also to consider how each dimension influences the other and to understand which dimension that has greatest impact on the overall gap for a specific physical-virtual activity. For instance, different kinds of presentation allow for different kinds of manipulation, and problems caused by topological discrepancies can be reduced by strengthened causal dependencies. From an engineering and design standpoint, the economy and technological difficulty involved in decreasing the gap within a certain dimension has to be compared to the efforts needed to decrease the gap in other dimensions.

Table 7 on page 131 shows the four gap dimensions along with consequences of the physical-virtual differences we can identify within each. Column three and four in the table show some approaches to reduce or “bridge” the differences. These approaches will be discussed later in this chapter.

**The origin of, and cause for, the gap**

The physical-virtual environment gap finds its roots in a relatively stable set of fundamental differences in how objects can be presented, organised, manipulated, communicated, transformed and stored under influence of constraints from the laws of nature in the case of the physical world, and theories of computation in the case of the virtual world. (These fundamental world-unique properties are discussed in chapter 7.) But there are also other causes:

• **The mere existence of physical-virtual activities:** If people would use the two worlds for completely different and independent things, there would not be any gap. Or at least, it would not really matter.

• **The dominating single-world design perspective:** Although the HCI research community often suggests to incorporate many aspects of human activity in system designs, including physical world considerations, physical aspects of the environment (in particular the roles of physical tools and physical work objects) are rarely part of the design. HCI people and software engineers are (naturally) concerned with software design while designers and manufacturers of physical entities rarely are interested in virtual aspects. However, things are changing and integration activities (although not always seen as such) from both sides have increased slowly but steadily the past 10-15 years. From the software engineering perspective, the

\(^{36}\) The ultimate goal is of course to arrive to a point where physical-virtual environments can be numerically graded based on their capability of supporting particular physical-virtual activities. However, a set of well-defined numerical metrics are not available, so the measuring of the gap distance has to be based on qualitative methods.
increased interest in the physical world is often still considered to be a way to improve the “user interface” towards the virtual world. From the physical designer’s viewpoint, computing and data communication technology is considered a way to augment what is still primarily a physical thing. Sooner or later there will be designers that do not make such a sharp distinction between whether the objects of their designs are primarily physical or primarily virtual. And neither will the human agents that use them. A discussion on the adoption of an integrated physical-virtual design perspective is presented in chapter 8.

- The general problem of designing for flexibility: Even for intra-world activities it is hard to design environments that support “ill-defined” activities since deciding a way to perform it is part of the activity itself. There is a general trade-off between strong automisation (resulting in a limited set of possible ways of doing things) and weak automisation (giving potentially endless, but not efficient, ways of performing an activity). The trick is to allow for flexibility on the right level of abstraction for the human agent performing the activity. Since actions within an activity often vary between low and high level, optimal support at all abstraction levels is more of a utopia than something realistically achievable. One classic way within HCI to address the problem is to define a finite set of human agent categories (e.g. “beginner”, “experienced”, “expert”) and to adjust the system support based on what category the particular human agent is supposed to belong to. If we want to design environments that support physical-virtual activities in general, ways to support physical-virtual actions on many different levels of abstraction have to be found, or else the environment will force the human agents to either work at the “wrong” abstraction level (increasing physical-virtual activity overhead) or stop them from switching environment (leading to weak support for the next action within the activity).

- Technological sensing and actuation problems: For integration to take place, objects and events have to be transformed from phenomena in one world to phenomena in the other. Although sensor and actuation technology has advanced immensely in recent years there are still many phenomena that simply cannot survive a physical-virtual trip (from the physical world to the virtual or vice versa). And of those that can, most have to be transferred in ways that imply manual and/or cognitive efforts from humans, thus adding to the physical-virtual activity overhead and thus, in effect, are not helping much in bridging or decreasing the gap. The main problem when trying to automatically transform a physical phenomenon to virtual is to “frame” the physical action and to grasp all the tiny details of a manipulation. In the other direction, transforming virtual phenomena to physical, big problems arise for instance when trying to overcome the extreme inertia and rigidity of physical objects compared to the almost unlimited compliancy and plasticity of virtual ones. Physical-world actuators are to date very blunt and relatively expensive.
• **User modelling/AI problems:** In order to automatisate some of the actions that human agents have to do manually in order to bridge the gap\(^{37}\), computer systems would have to watch and interpret human activities “always”, and without asking or being asked. Such background processes often generate an immense amount of data that has to be fed into models of humans behaviour for interpretation and potential system reaction. These models are not easy to design without restricting the human agent to a very limited set of “allowed” actions. Since the strength of the physical world is in part its flexibility compared to most virtual environments, there is a tension, and a gap. And in the foreseeable future, many physical activities will probably continue to be just physical, not because the computer system is not able to notice them as changes in the data flow from the sensors, but because we are not yet able to design good enough models that make the system realise what kind of event this change in the never-ending bit-stream actually was about.

### 6.3 An empirical study of the gap in a knowledge work context

Virtual activities and physical-virtual activities appeared when personal computers started to diffuse into the society in the 1980’s. Being such a new phenomena, there are few studies made that can give a good estimation of the extent of the physical-virtual activity overhead in everyday life.

In recognition of this lack of empirical knowledge, a questionnaire-based study was designed and performed in 1999 in order to shred some light on the appearance of the physical-virtual environment gap in everyday professional activity.

**Overall question: conditions for knowledge work**

The study was designed and performed in collaboration with my colleague Anders Broberg, with the general aim of investigating conditions for knowledge work in the industry. Thus, the issues addressed by the study were manifold and did not only regard the physical-virtual environment gap. The study was explorative in nature, performed with the hope to find interesting results without actually putting up well-defined hypotheses on the outcome. In this chapter section I will give a short overview of the design rationale for the study in general, followed by a more specialised analysis and discussion of the part of the study that addressed the physical-virtual environment gap\(^{38}\). Some of the gap-related results reported here have been published elsewhere (Pederson, 2001b).

---

37. Common actions performed for bridging the gap manually are discussed later in this chapter (section 6.4, page 123).

38. A more complete analysis of the parts of the study not covered here can be found in Anders Broberg’s PhD thesis (Broberg, 2000).
Method
The study was performed by asking carefully chosen subjects to fill in an on-line WWW-based questionnaire. Other ways of obtaining data would have been for instance interviews or observation. The questionnaire-alternative was considered to be the best choice for economical and time reasons. Also, using the WWW as media would not be a limiting factor in the company where the study was performed since almost all potential subjects had access to the Internet. From an administrative perspective, the use of WWW as a media instead of paper and postal mail is believed to greatly have simplified the collection and compilation of the material. The study was performed in six steps, all which will briefly be discussed in the design section later on:
1. Observation-based pre-study
2. Design of pilot questionnaire
3. Pilot questionnaire data collection
4. Design of final questionnaire
5. Data collection in two phases
6. Analysis
Originally, complementary field observations and interviews were intended to be performed after the analysis phase, in case the results from the analysis would show to be hard to interpret or if the results encouraged extra investigations for other reasons. This optional part of the study was not performed for two reasons, 1) the results from the data collection were indeed considered to be quite satisfactory, 2) the collaboration within the FIOL consortium that sponsored this study (among others) was finished by the time the analysis of the data was performed, and thus, a continuation would have required extra efforts and fund raising.

Observation and interview pre-study
A handful of observations and informal interviews with professionals at the company helped us to concretize and tune the general questions (mentioned in the goal description earlier) towards the work environments observed. It should be noted that since we only observed people at one company site and since the subjects were gathered nation-wide at different sites, the working conditions did of course differ.

Design of pilot questionnaire
The pre-study helped us to filter out questions that were too abstract, or to reformulate them using more everyday terminology. Based on the experience from the pre-study, a first WWW-based questionnaire was designed.

Pilot questionnaire data collection
The first version of the questionnaire was filled in by 22 colleagues at the Department of Computing Science. They were not only asked to fill in the questionnaire but also to reflect on the quantity and quality of the questions in the questionnaire.
Design of final questionnaire

The pilot study caused some changes to the questionnaire. The parts of the revised and final version of the questionnaire that is relevant for the analysis of the physical-virtual environment gap will be repeated and discussed in the analysis section of this chapter. An English version of the complete questionnaire is included in this thesis as appendix A. The questionnaire consisted of 32 mainly closed questions divided into seven groups, all but one group addressing different aspects of knowledge work. The exception was the group of questions on personal data for acquiring information about the composition of the target group such as their age, sex, level of education, and experience of using personal computers. The other six question groups were addressing the following areas:

Tasks
The purpose with this group of questions was to acquire a better picture of what kind of work knowledge workers do. The focus was on what kind of personal qualities that they considered important for the kind of work they were doing, and how well these required qualities matched their own personal qualities. Additionally, one question regarded how much time they spent on tasks that they think they are hired to do, and how much time they spent on activities they want to perform, or enjoy.

Physical environment
This group of questions investigated environmental conditions such as where the work was performed, what factors they found important for a good working situation as well as how often they made changes to their physical working environments.

Computer support
The purpose of this group of questions was to see how and how much they utilised personal computers in their work, how big influence they had in deciding the level of computerisation in their work, and what kind of factors that governed their decisions concerning the level of computerisation of a work task.

Tools in general
This group of questions were aimed to make an inventory of tools (physical, virtual, mental) that the knowledge workers utilised in their work, and to find out for what purpose they used them and which features they valued within them. There were also questions in this group aimed to catch new or unusual ways of using well-known tools, and to discover needs for new tools or new kinds of support.

Sources of information
The purpose of this group was to examine how knowledge workers judged different kinds of well-known sources of information, with a focus on being able to cope with their tasks. There is also a question aimed to find out what factors that were important when they evaluated information sources.
Physical-virtual
The purpose with this group of questions was to investigate the general awareness and acknowledgement of the physical-virtual gap, to find out common problems related to it, and to see how the subjects coped with the gap.

Data collection
The questionnaire (or rather, the URL to it, and the invitation to fill it in) was sent out to a number of 130 carefully selected subjects that could be categorised as having one or several of the tasks/professions listed in table 4. It was a quite laborious and impossible task to find an equal number of subjects in each category (especially medical doctors are rare at telecom companies) but the resulting subject population can be regarded as fairly evenly spread. The selection of subjects was made in collaboration with contact persons within the company group and the subjects were also informed by their local managers that the time they spent on filling in the questionnaire, should they choose to do so, was to be considered as a fully valid part of their work. They were also briefly informed about the purpose of the study and why the company was interested in performing it, as well as the fact that the mapping between answers and identities of subjects would only be known to the researchers and kept confidential.

The data was collected in two phases since we initially did not get a satisfactory number of subjects (only 30 subjects responded in the first phase). The subjects in the second round were offered a cinema ticket as reward for their work in filling in the questionnaire. How this affected their response is an open question. The data collection ended with a total number of 81 subjects, giving a response frequency of 62%.

Results
Only the questions specifically related to the physical-virtual environment gap will be presented and analysed here: questions 16-19 and 30-32 (see the full questionnaire in appendix A for an overview of the questions asked).

TABLE 4: Task and profession categories of interest.
Q16: How big share of your work tasks do you perform completely without computer support, with some computer support, with large computer support, and using computer only?

As shown in table 5 and the corresponding bar chart pictured in figure 26, subjects report a fairly even distribution of tasks in the physical-virtual continuum (e.g. from tasks completely performed without computers to tasks performed completely with computers) with some peaks for tasks performed completely without computer support and for tasks performed with only some computer support. However, a large number of subjects also report a high share of tasks performed using computers only, or with large computer support. Almost no-one performs large amounts of non-computer-supported tasks. This can in part be explained by the fact that the company studied was a tele-communication company in which the process of task “computerisation” can be assumed to have gone relatively far within many profession categories.

Q17: Has the amount of computer-supported activities decreased, remained constant, or increased the past five years?

Figure 27 shows that a clear majority of the subjects are of the opinion that the amount of computer-supported tasks have increased during the period 1995-1999. This result only confirms what could be expected. The question does not as much address the physical-virtual environment gap as the prerequisites for it.

Q18: To what extent do you think you have the possibility to influence the element of computer support in your work?

The collective opinion towards this relatively vague question is pictured in figure 28 and does not provide a clear answer. There is a small majority thinking that the choice of using (or how to use) computers in their work is not up to them to decide. However, the slight overweight in this direction is not significant.
Q19: In the cases where you have the possibility to influence the degree of computer support, what factors determine your decision? Grade the following factors depending on how important they are for your decision.

Not surprisingly, table 6 and the corresponding figure 29 show that all mentioned factors are considered to be important for the decision whether to use computers or not. The only somewhat interesting result is perhaps that the “quality of the result” to such a high degree is agreed upon to determine the choice.

**FIGURE 26.** The share of work tasks performed with and without computer support. Data taken from table 5 (blank answers not shown).
FIGURE 27. Has the amount of computer-supported activities decreased, remained unchanged, or increased during the past five years? Two blank answers not shown.

FIGURE 28. To what extent do you think you can influence the element of computer support in your work? Two blank answers not shown.
**TABLE 6:** What factors determine computer use? See also figure 29.

<table>
<thead>
<tr>
<th></th>
<th>irrelevant</th>
<th>not so important</th>
<th>quite important</th>
<th>important</th>
<th>very important</th>
<th>blank</th>
</tr>
</thead>
<tbody>
<tr>
<td>personal interest</td>
<td>7</td>
<td>12</td>
<td>16</td>
<td>26</td>
<td>7</td>
<td>13</td>
</tr>
<tr>
<td>short-term efficiency</td>
<td>1</td>
<td>13</td>
<td>20</td>
<td>23</td>
<td>9</td>
<td>15</td>
</tr>
<tr>
<td>long-term efficiency</td>
<td>1</td>
<td>3</td>
<td>13</td>
<td>36</td>
<td>16</td>
<td>12</td>
</tr>
<tr>
<td>quality of the result</td>
<td>1</td>
<td>2</td>
<td>7</td>
<td>27</td>
<td>31</td>
<td>13</td>
</tr>
</tbody>
</table>
Q30: In some situations both material and computer-based tools are used in order to perform parts of a continuous work task, potentially leading to problems since the two environments don’t “speak the same language”.

Example 1: You have made notes on paper (approximately two A4 pages) during a meeting and would like to send them to a colleague using e-mail. In order to transfer the notes to the computer environment you need to rewrite them, taking up extra time.

a) Do you recognize the situation?

**Figure 30.** Physical-virtual gap-bridging situations, example 1. Do you recognize the situation?

Figure 30 shows that a clear majority of the subjects recognize the situation.

How do you handle the situation?

Figure 31 shows that a clear majority of subjects would retype the notes made on paper into the computer and e-mail them. However, a noticeable number of subjects would choose to send the notes as they are by fax, and a clear minority would also consider the phone. The fact that no-one would choose to skip notifying the colleague (!) is reassuring. Ordinary paper mail is not an option either.

Example 2: You have received a new important document from a remote colleague by e-mail, and you have the task to read it through (approximately 10 A4 pages) and give comments. Since it is quite urgent it would be good if you could send the document back together with your comments.

b) Do you recognize the situation?

Figure 32 shows that the second example scenario is even more recognised than the first.
FIGURE 31. How the first physical-virtual gap-bridging scenario is handled. Four blank answers not shown.

FIGURE 32. Physical-virtual gap-bridging situations, example 2. Do you recognize the situation?
How do you handle the situation?

**FIGURE 33.** How the second physical-virtual gap-bridging scenario is handled. Four blank answers not shown.

Figure 33 shows that e-mail is definitely the preferred communication channel also for returning the proof-read document but that the majority of subjects are divided into three almost equally large groups with regard to how the proof-reading action within the physical-virtual activity is performed. One group (23 subjects) prefer to read and write in the physical environment and to retype the hand-written comments into the virtual document manifestation before sending it back using e-mail. A second group (30 subjects) would do the same, except they would type the comments directly into the virtual document while reading the physical. A third group (21 subjects) would perform the entire activity in the virtual environment including the reading. Apart from the fact that the answer to this question to some extent proves the existence of physical-virtual activities as defined in this thesis among this particular subject population, it also points out the important fact that people choose to perform these activities in different places within the physical-virtual continuum. The physical-virtual design perspective proposed in this dissertation is based on the idea that environment designers should provide a physically-virtually redundant environment so that human agents acting within them have considerable freedom in the way they perform specific physical-virtual tasks.
Q31: How big problem do you think the information environment “gap” described in the examples above, between material and computer-based environments, is for you in your work?

As shown in figure 34, most subjects consider the aspects of the physical-virtual environment gap exemplified in Q30 to be a small problem, and no one thinks it is a big problem. However, the majority acknowledges it as a problem nevertheless.

Q32: How much of your work time do you estimate that you spend on bridging the “gap” between the material and the computer-based environment?

Figure 35 shows that almost everyone believes to spend at least some time in manually bridging the gap, and a majority estimates it to be between 1 and 15% of the total working time.

Conclusions

The professional activities performed by the studied subjects are estimated by themselves to be more or less evenly distributed in time spent on purely physical, purely virtual, and physical-virtual activities with a slight overweight for virtual activities (see figure 26, page 116). A majority has noticed a change towards computer-supported activities (i.e. virtual and physical-virtual) during the period of 1995-1999 but have an unclear opinion as whether they themselves can influence the element of computer-support in their activities. The overall recognition of two gap-bridging scenarios in which information has to be worked on and transferred between physical and virtual environments is very high. In both cases, but in particularly the second, subjects report preferences for different approaches in solving the situations, indicating the ex-
istence of individual physical-virtual work styles. A majority of the subjects see the physical-virtual environment gap, as exemplified by the two given scenarios, as a small problem on which they however spend a not negligible amount of time.

6.4 Bridging the gap

The two scenarios used in the empirical study presented above are examples where manual gap-bridging has to be performed in the causal dimension of the gap. In the remaining part of this chapter, the physical-virtual environment gap will be analysed and discussed from the perspective of how it can be bridged along that dimension and the other three dimensions already discussed in section 6.2.

As mentioned earlier, the gap can be bridged either manually by the human agent during the execution of a physical-virtual activity, or by environment designers before activities take place in the physical-virtual environment. The basic relationship between the two is that all “bridging” not performed by designers in the design phase is left for the human agent to overcome while performing the activity. We will now take a closer look at some components of the two ways of bridging the gap, organised around the four previously discussed dimensions.
Bridging the gap through situated actions — the human agent’s way

The presentation dimension

Before the introduction of graphical displays, representations of physical objects in the virtual world would always differ significantly from the “original”. GUIs and the direct manipulation interaction paradigm was a big step towards virtual representations that in many ways reminded of their physical correspondences. Physical representations of virtual objects on the other hand are often limited by physical laws of nature that makes it hard or impossible to present the dynamic properties of virtual corresponding manifestations (an actuation problem, see page 110). More or less the only way for human agents to cope is to learn the different presentations through trial-and-error.

The manipulation-affordance dimension

If a PVA attribute cannot be manipulated in the current environment but in the other, there is no other way to perform the desired manipulation but to switch environment.

The organisation/navigation dimension

The fundamental differences between how things in the physical and virtual worlds can be organised, forces human agents to adapt to different organisation and navigation styles when switching environment. Spatial memory is helpful both in the physical and the virtual world, as is a good cognitive model of how spaces and subspaces are organised, and why. To overcome differences in the topological dimension is a learning process.

The causal-dependency dimension

When PVAs are not automatically kept synchronised by a computer infrastructure (and today, usually they are not), the human agent has to assure that all different manifestations are up to date. At what point in time this manual synchronisation takes place can be a matter of personal preferences but the two probably most common situations are a) when a manipulation of one of the manifestation has just been finished, or b) when a manipulation of a manifestation is about to take place. In the first case, the gain is that the recently performed manipulations are fresh in memory and therefore easy to reconstruct in the other environment. The gain in the other case is that if further actions have to be carried out in the environment in which the manipulation took place, unnecessary extra switching between the worlds can be avoided by performing the update actions at a later stage. In any case, bridging the gap in the causality dimension always demands extra cognitive efforts and/or extra physical or virtual actions. Furthermore, cognitive resources are probably latently occupied with maintaining a coherent cognitive model of corresponding, potentially un-synchronised, PVA manifestations in the two worlds.
Examples of actions
To concretize the situated actions involved in bridging the gap, here are some examples:

- To find an object (e.g., a PVA manifestation that corresponds to a recently manipulated object in the other world). Example: To navigate the local file system for a word processor file that corresponds to a paper document just annotated with a pen.
- To confirm the identity of an object, i.e., to assure that the object is “the right” object. Example: To inspect identification sections (time-stamps, dates) in the opened file and to compare it with the same sections within the paper document.
- To transfer changes done to an object A in previous actions to object B in the other world. Example: To edit the file in the virtual environment based on the notes on the paper manifestation.
- To confirm that corresponding PVA manifestations are synchronised. Example: To check that one’s pocket diary is synchronised with the virtual calendar on the PC.

Bridging the gap through design — the environment designer’s way

The presentation dimension
HCI designers have been dealing with (mostly visual) presentations of virtual phenomena since the introduction of GUIs, trying to shape virtual objects and events so that they remind of more well-known physical artefacts, i.e., relying on metaphors. Sometimes the approach has been very successful, allowing fairly computer-illiterate users to interact with the objects without prior experience of them. The drawback is that in some cases, the approach leads to misconceptions about how the virtual world works at large, causing problems at a later stage. In any case, I subscribe to the idea of streamlining form and behaviour between physical and virtual manifestations of PVAs to the extent that the corresponding manifestations appear to belong together, whenever such design steps do not imply reducing the functionality of each manifestation. Furthermore, if the “belonging together” relationship between manifestations is strengthened by other means, for instance through stronger causal dependencies, human agents can probably cope with larger discrepancy in how they look.

In short I believe that designers of PVAs should try to:

a) make sure that all relevant properties assure that all relevant properties that can be expressed in both manifestations also are expressed, and preferably in a coherent form;

b) express the properties that are unique to the manifestation in a particular world in such a way that they on the one hand do not disturb the sense of unity created by the common manifestation properties, but on the other hand clearly expresses the unique properties as exactly that; unique properties that the PVA manifestation has in the particular world only.
Example: a coffee cup can be made to have a virtual manifestation that visually looks similar to the physical manifestation since the visual properties of the physical cup can be transferred to the virtual world in a pretty straight-forward fashion. However, there are other properties of the physical coffee cup that are harder to represent in the virtual world: what about weight, or temperature? In a standard two-dimensional WIMP virtual environment these physical properties have to be ignored when designing the virtual manifestation; they have to be represented in other ways. Following the design principles above, I would encourage to represent the properties even if they have to be made to appear differently than the original (physical) presentation. One alternative would be to show the temperature and weight of the cup in numbers. Since these numbers would correspond to properties that actually belong to the physical manifestation, and cannot be manipulated in the virtual world (assuming that no such actuators are built into the physical manifestation) they should not have a central place in the appearance of the virtual manifestation. Perhaps they should even be hidden, only to be revealed on explicit request from a human agent. A tempting solution for the representation of temperature would be to give the virtual manifestation different shades of red according to the temperature of the cup. I do not believe it to be a good design in general since it would disturb the “sense of unity” of the manifestations if one of the manifestations keep changing colour and the other one does not (unless the physical cup manifestation was made of a material that would do the colour changes as well, or the cup would be exposed to excessive heating that would bring out this behaviour, but these are exceptional circumstances).

“To me, one of the big problems of mixed reality is the potential for taking physical objects, which we understand intuitively (or through learning) and obey the laws of physics, and making them confusing. Virtual systems are much less understandable and predictable than physical ones - mixing them can make virtual systems more understandable or make physical systems more confusing.”

(Wendy Mackay, personal communication)

**Physical-virtual mode errors:** If physical and virtual object presentations become indistinguishable (which is a partial goal for research in Virtual Reality) at least in theory we run the risk as human agents to make wrong assumptions of what can and cannot be done with an object because we might have forgotten whether we are performing the manipulation in the physical or the virtual world at the specific moment. “What? Why can’t I make this paper document just vanish when one moment ago I did exactly that with another one that looked just the same?” In classical HCI terminology, this kind of misjudging the context of one’s actions is referred to as “mode error” (Norman, 1981). One might think that the risk for physical-virtual mode errors is insignificant in the foreseeable future, given the still easily discernible differences between the look-and-feel of immersive Virtual Environment and the real world. However, it is not far-fetched to imagine physical-virtual mode errors occur-
ring also in more subtle situations when physical and virtual objects are “augmented” with functionality and features from the corresponding world. A related mode error which is perhaps even more easy to imagine is that of mistakenly assume that an environment is augmented with cross-world functionality when it is in fact not. Because the spreading of integrated physical-virtual environments is likely to be non-uniform, such mode errors are likely to be a concern for a long time to come. As a designer of physical-virtual environments, perhaps these potential “physical-virtual mode errors” should be handled in the classical HCI way as well, by enriching objects with cues that consciously or unconsciously inform the human agent if it is a virtual or physical object that is currently manipulated, and if the environment is “augmented” or not.

The manipulation-affordance dimension
Whenever functionality and/or manipulation affordances differ between PVA manifestations, these “one-world” properties should not be hidden by the designer (motivated by the wish for perceptual and/or functional homogeneity among manifestations) but be clearly expressed. In other words, I do not propose a reduction of functionality and manipulation affordances to “the lowest common denominator” among the manifestations. First, this would obviously significantly limit what you can do with them. E.g. you would not be allowed to do a text search operation in your virtual document manifestation since this operation is not supported by the physical manifestation and you would not be allowed to punch holes in your paper version since there is no way to do the same (or something corresponding) in the virtual world. Secondly, for physical manifestations we do not have the power to restrict manipulation in the way that would be necessary: how do you prevent someone from crumpling up a paper document? Or to punch holes in a document? Physical and virtual objects will always have differences in what can be manipulated and how, and this fact is something to embrace rather than to reject. The most fruitful design approach within the manipulation-affordance dimension has to be a policy of expansion: to increase the amount of manipulation support in both worlds. This leads to fewer forced world switches for human agents and can as well inspire designers to invent new functionalities in their efforts to implement manipulation affordances found in one world in the other world. The approach is about moving (access to) virtual functionality to the physical world (e.g. to allow for alphabetical search for objects in the physical world) and vice versa (e.g. to let human agents knead virtual “matter” through the use of immersive VR equipment).

The organisation/navigation dimension
The fact that all organisation in the physical world is based on placing relatively rigid and heavy matter in a three- or four-dimensional space, and that virtual organisation is about placing lightweight bits in a multi-dimensional space, has great impact on how differently the two worlds can be navigated and organised. The virtual world affords cheap, immediate, and temporary organisation changes of the same set of objects
while similar processes in the physical world are slower and more costly. The number
of easily accessible objects in the virtual world, is significantly larger than the number
of things that are manipulable in the physical world. However, the viewable part of
the virtual world, at a certain time instant, is normally smaller than the directly per-
ceivable part in one's physical surroundings. Partly because of this fact, partly because
of the cheap destruction and reconstruction, objects in the virtual world tend to ap-
pear in, and disappear from the perceivable part of the virtual world much more fre-
quently than any physical objects do in the immediately perceivable and manipulable
physical environment. Furthermore, the virtual world is based on logic and discrete
data structures from the bottom up. Objects in the virtual world have unique identi-
ties that can be automatically and unambiguously arranged and rearranged according
to any system of organisation while physical objects need to be organised manually,
there is no guarantee for unambiguity, and the organisation scheme has to be chosen
with care since physical reorganisation costs.

As has been mentioned before, one way of bridging the gap is to transfer manipu-
lation support from one world to the other. In the organisation/navigation dimension
this can be done by creating an infrastructure that assigns a virtual manifestation to
each relevant physical object (i.e. to create a PVA), and to simulate physical reorgani-
sation and navigation activities by supplying the human agent with a virtual space for
reorganisation and navigation of the virtual manifestations of the PVAs. Although the
effect of the virtual activities might not be automatically transferred to the physical
world, the activities allow for valuable cheap trial-and-error organisation and naviga-
tion experiments by providing virtual functionality to a simulated physical environ-
ment.

Another way of bringing physical and virtual organisation/navigation closer to
each other, also based on a PVA infrastructure, would be to strengthen causal depend-
encies between physical and virtual places, i.e. to automatically “activate” and expose
manifestations of PVAs that are currently involved in an activity in the other world.
Example: If physical activities are performed in a specific physical place, involving a
certain set of physical PVA manifestations, a corresponding virtual place, containing
the corresponding virtual PVA manifestations can be automatically generated and pre-
sented to the user in the virtual world, in case the human agent would care for a world
switch. In this way, extra navigation and organisation actions when performing a
world switch can be avoided assuming that the switch is followed by continued ma-
nipulation of the same or roughly the same set of objects.

The causal-dependency dimension
Being what I believe to be one of the most important dimensions of the physical-vir-
tual gap, the causal-dependency dimension is also one which can dealt with in the
most straight-forward way: to automatically recognise relevant state-changes in PVA
manifestations and automatically update the corresponding manifestations accord-
ingly, i.e. to create synchronised PVAs (see definition 5.2.3, page 96). To do this in prac-
tice is not always easy but one should bear in mind that only a few decades ago it was
almost unthinkable. Sensors and actuators can act as “end-points” in these artificially
maintained causal relationships that are mediated through more or less complex mod-
els of how a state-change in one manifestation corresponds to another state-change in
another manifestation.

The design of synchronised PVAs has to face many challenges because many object
attributes cannot be part of a physical-virtual causal relation. In particular, it is phys-
ical-world sensing and actuation that is the hard nut to crack:

The physical-world sensing problem: In the case where an attribute is modified di-
rectly by a human agent, the performed changes could a) be too hard to sense by the
computer system (e.g. the fact that the heap of modelling clay that previously looked
like a car now looks like a boat), or b) they might have no meaningful representation
in the other world so even if captured, the utility is unclear (e.g. “So now the car is a
boat. But why?”). “Hard to sense” state-changes can be divided into two categories:
a1) the ones for which there are no physical or virtual sensor available for recognising
the state change (e.g. the clay example), and a2) the ones that are technically possible
to sense but hard to capture because there are no good models for correctly interpret-
ing the data that the sensors deliver (E.g. even if the shape change of the clay heap
could be captured, using a camera or laser scanner for instance, what is the meaning
of a heap of modelling clay in the shape which human agents would refer to as a
boat?39).

The physical-world actuation problem: Inter-world synchronisation can be hard
achieve also because of limitations of actuators. There might simply not be any actu-
ator available that can cause the state change in the desired way. Actuator limitations
are more evident in the physical world than the virtual and will probably continue to
be so in the foreseeable future. For instance, what combination of physical actuators
allow a physical ball to suddenly transform itself into a cube as a result of a human
agent’s transformation of its corresponding virtual manifestation from a ball to a cube?

In order to create more and stronger causal relationships between physical and vir-
tual objects, I would like designers to

• widen their design space to incorporate also the physical world (for software engi-
neers) and the virtual world (for physical designers), i.e. to adopt what is in this
dissertation referred to as a “physical-virtual design perspective”, and to increase
collaboration across the physical-virtual border;

39. Of course, the interpretation of human activity has to be done considering the context in which
it takes place. Context Awareness research has shown that physical activity context is hard to
frame, especially when activities are not tightly bound to dedicated places. In the case of the
modelling clay presented above, if it is a child imagining various ways of transport for no particu-
lar reason other than pure joy of creation, it is easy to see many possible interpretations of activity
context. How many of these interpretations would be of any use for a system as a basis for sup-
porting the activity?
• make use of sensors and actuators whenever there is the slightest chance that this could help creating synchronised PVAs that decrease the causal-dependency dimension of the physical-virtual environment gap;
• make more use of user modelling to decrease the costs of physical-virtual world switches by “setting the destination environment up” before the human agent arrives.

One general design problem is how to communicate to the human agents acting in the environments what properties of PVAs and the worlds in general that are automatically synchronised, and what properties that are not. I.e. the human agent should not expect an automatic physical-virtual manifestation synchronisation when it is not delivered by the system, and conversely, should not expect to be performing an intra-world action when in fact there are changes in the other world as well. (See the discussion on physical-virtual mode errors earlier; page 126.)

Table 7 summarises the discussion on the four dimensions of the physical-virtual environment gap.
<table>
<thead>
<tr>
<th>gap dimension</th>
<th>resulting in</th>
<th>human agent’s way of coping with the gap</th>
<th>designer’s potential way of decreasing gap</th>
</tr>
</thead>
<tbody>
<tr>
<td>physical-virtual presentation differences</td>
<td>things look, feel and behave different</td>
<td>general cognitive work, trial-and-error, learning</td>
<td>streamlining PV perceptual design if not interfering with function, strengthen PV identity by increasing causal dependencies</td>
</tr>
<tr>
<td>of objects or events although they play similar roles within a PV activity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>difference between what properties of objects that can be manipulated by human agents in each world</td>
<td>certain actions have to/are preferred to be performed in one of the worlds instead of the other</td>
<td>world switch</td>
<td>expand PV manipulation affordances as much as possible in both worlds, let manipulation affordances in one world inspire design in the other</td>
</tr>
<tr>
<td>organisation/navigation differences between how physical and virtual space is organised</td>
<td>different navigation methods</td>
<td>general cognitive work (incl. use of spatial memory), learning</td>
<td>PVA infrastructure, “intelligent” physical-virtual place synchronisation, the use of spatial metaphors</td>
</tr>
<tr>
<td>weak causal dependencies between physical and virtual PVA manifestations</td>
<td>physical and virtual objects “live separate lives” although they represent the same thing</td>
<td>manual synchronisation (incl. general cognitive work, extra physical or virtual actions), cognitive overhead for maintaining a coherent cognitive model of corresponding objects in the two worlds</td>
<td>adopting a physical-virtual design perspective, more collaboration between software engineers and physical designers, more use of sensors and actuators, better user modelling</td>
</tr>
</tbody>
</table>

**TABLE 7**: Four dimensions of the physical-virtual environment gap. The columns that show how human agents and how the designer address the gap dimensions (column 3 and 4) should be regarded as a non-complete collection of examples of approaches for minimising the gap in the particular dimension.
The promise of automatically controlled attribute redundancy

The approach taken in this dissertation for bridging the physical-virtual environment gap follows two complementary paths: 1) maximising redundancy so that activity-relevant object attributes in the two worlds are presented and preferably also manipulable in both worlds, 2) keeping these redundant attributes synchronised with minimal attention necessary from the side of human agents. One path cannot be successfully followed without the other.

It is reasonable to believe that physical-virtual systems that push in these two directions will offer a better environment for physical-virtual activities for two reasons: 1) they will reduce the number of forced physical-virtual environment switches since some of the actions that motivated them before are not necessary any more; 2) they will make some of the still necessary physical-virtual switches less expensive for human agents performing them, since some of the update-related actions connected to the world switch are now performed automatically.

Figure 36 illustrates a hypothetical situation where a human agent decides to switch to the virtual world in order to more efficiently perform the next manipulation action on the specific PVA. At command, the corresponding PVA is retrieved into the virtual environment (I will refer to this process as **physical-virtual cloning**) by the physical-virtual infrastructure. The virtual manipulation action is performed on the PVA by the human agent and the resulting state changes of shared attributes is automatically propagated to the PVA by the physical-virtual infrastructure. The human agent is free to choose in what environment (the physical or virtual) the activity involving the PVA is to proceed as long as the activity is about manipulating shared attributes.

When to synchronise

The time between shared attribute state change and complete state change propagation is denoted \( t_{\text{synch}} \) in figure 36. In many cases, \( t_{\text{synch}} \) would preferably be kept as small as possible, i.e. physical-virtual world synchronisation in real-time. Sometimes this will not be possible for technological reasons (e.g. one or several of the manifestations that should be affected by a state change might be temporarily “off-line” and thus out of reach for the physical-virtual infrastructure), or a human agent might want to try out or perform a series of manipulation actions before any synchronisation takes place (e.g. before manifesting the changes made to a local PVA to a corresponding PVA that is more publicly observable). Delayed synchronisation is in some cases a very useful or even vital in certain PVA manipulation contexts such as for instance in the case of the construction of a large bridge, where alternative designs is best evaluated in the virtual world only, before any synchronisation (to actually (re-)build the bridge) takes place. The cost-effectiveness of trying out different alternative attribute states in
the virtual world before manifesting it in the physical is probably the strongest argument for, and represents the most general case, where delayed synchronisation is preferred.

**Requirements for environment-independent task completion**

It is important to note that it seldom is necessary to replicate *all* attributes in *all* artefact manifestations in order to improve the support for “environment-independent tasks”. As long as attributes *crucial for the task* are shared among the distributed artefact manifestations, this will be enough to allow for environment-independent task completion. Example: In many cases, a PV text document’s background colour does not necessarily have to be replicated among the manifestations as long as the text (which in most cases has to be replicated) is perceptible.
Feasibility
To what extent the promise of automatically controlled attribute redundancy can be fulfilled in practice, and how big effect gap-bridging physical-virtual systems can have on human everyday activities, is an open question that can only be answered by experiments and case studies. It is quite clear at the time of writing that initial physical-virtual systems will only relieve human agents from a few of the obstacles involved in bridging the gap. The hope is nevertheless that a design approach based on the concepts and terms described in this thesis, combined with emerging technology will eventually lead to environments that are significantly better for human agents performing physical-virtual activities.

6.6 Summary
This chapter has provided a definition of the concept of physical-virtual environment gap and discussed its origin and its dimensions. Results from an empirical study of the gap in industry has been presented, indicating that manual gap-bridging is a relatively common and relatively time-consuming part of physical-virtual activities. Existing and potential ways of bridging the gap were also discussed, both from the perspective of environment designers and from the perspective of environment use. The chapter concluded with a theoretical motivation for the design of physical-virtual artefacts as a gap-bridging design approach.
Comparing the physical to the virtual world

This chapter takes a closer look on properties that make the physical and the virtual worlds different from each other, and properties that make them similar. Specifically, navigation and object manipulation aspects of the worlds are discussed. The world properties are categorised as belonging to the group of either fundamental or plastic properties depending on whether they are impossible or possible to alter by human design.

7.1 Comparison method

While the kind of human activities of concern for this dissertation (see section , page 81) in the physical world takes place in accordance with what can be viewed as one single interaction paradigm, activity in the virtual world takes place in one out of several more or less mutually related interaction paradigms. Different IT devices (virtual environment providers) are bound to different interaction paradigms and offer access to different more or less mutually isolated virtual subspaces. Luckily, the in-

40. See definitions 4.1.1 and 4.1.2, page 75.
41. For reasons of simplicity, the discussion will not take into account cultural and social practices connected to spaces (making them “places” in the terms of Harrison and Dourish (1996)) giving rise to differences in how otherwise objectively similar environments can appear and behave different, nor is it influenced by the possibility of various kinds of human perceptive and/or cognitive handicaps. Such considerations, although perfectly reasonable, would complicate the analysis in this chapter significantly.
teraction paradigms associated to these devices share interaction properties both within and across the device classes (as well as with the physical world), enabling human agents to make use of already acquired interaction paradigm knowledge when learning how to behave in the different environments, and when switching between them.

**The many faces of the virtual world**

Virtual environments are typically provided by interactive computational devices fitted with visual displays, such as PCs, PDAs and cellular phones already mentioned in chapter 4. A more extensive list of virtual environment providers popular at the time of writing is shown below:

- desktop PCs
- laptop/tablet PCs
- wearable PCs
- PDAs
- cellular phones
- digital (wrist) watches
- kitchen equipment such as microwave ovens
- electronic music instruments
- sound system devices such as radio tuners and amplifiers
- more or less “intelligent” display devices (TVs, LCD panels, digital projectors, CRT monitors)

Although other modalities are definitely used for mediating the state of virtual environments (e.g. sound and touch) visual displays will for reasons of simplicity be treated as the main media through which devices reveal the state of the virtual environment they provide.

Interaction paradigms are mapped to IT devices by the designers of the devices based on the task that the specific device is intended to support, hardware constraints (size, power consumption, input and output devices). Sometimes the interaction paradigm, or implementation thereof, is device independent, and can be chosen by the consumers themselves based on personal preferences. (Different paradigm implementations can sometimes compete directly with each other, having more or less the same target group, e.g. Apple MacOS, Microsoft Windows, and Linux.)

The diversity of appearances of the virtual world constrains any comparison between the physical and the virtual world as a whole to be done on a high level of abstraction and be based on properties of the virtual world that are present in most, if not all, of the present virtual interaction paradigms. These latter ubiquitous virtual properties will be denoted *common* virtual paradigm properties, and those that seem to appear in only one or few device classes for *paradigm-specific* virtual interaction properties. Both types will be discussed later in this chapter. The virtual interaction paradigms explored are:
• The WIMP paradigm (exemplified by Microsoft Windows 2000),
• The Command-Line Interface paradigm (exemplified by C shell in the UNIX clone SunOS 5.8),
• The PDA paradigm (exemplified by PalmOS 3.5.0),
• The cellular phone paradigm (exemplified by the cellular phone Ericsson T66)

Figure 37 shows examples of IT devices providing the interaction paradigms listed above.

Compared interaction paradigm properties

The differences and similarities between the physical and the chosen virtual interaction paradigms will be explored by studying five basic interaction-relevant properties of the specific different interaction paradigms:

• space size
• space navigation
• object presentation
• object manipulation
• object translation

The discussion on each interaction paradigm is preceded with a note on what kind of Inter-World Event Mediators (IWEMs, see section 4.3, page 81) that are commonly present on devices tied to the specific interaction paradigm. This is because, naturally, properties of specific IWEMs in general have big influence on the design of interaction properties of the different virtual interaction paradigms.
7.2 The physical world

Inter-World Event Mediators — no
Physical-world interaction is not built around IWEMs since humans act in the physical world performing only intra-world actions (see section 4.3, page 81), i.e. actions entirely taking place in the world in which the agent performing the actions is situated.
**Space size — practically infinite**

Although being possible to measure (the globe and its atmosphere define an upper limit), the number of accessible places can be regarded as larger than what is possible to visit for any single human being within a limited amount of time. Thus, physical space can be regarded as infinite for the purposes discussed in this thesis.

**Space navigation — smooth, slow, and geometrical**

Navigation in the physical world takes place in a three-dimensional geometrical space. This means that, among other things, navigation between two places in the physical world eventually implies travelling at least the geometrical distance defined by the straight line between the two places. Human agents navigate the space by translating their bodies, or part of bodies, to different positions in the space. The translation is in general smooth and inevitably continuous. Navigation can be speeded up through the use of agent-translation tools such as bikes, cars and aeroplanes.

**Object manipulation — local, not much “magic” involved**

The agent-manipulable part of the space is in general limited by the geometrical operative radius defined by the length of the agent’s limbs and other body parts used for interactive purposes, potentially expanded by tools. Objects within this operative radius can be manipulated through the application of physical force onto the object (e.g. pushing, grabbing, tearing, shaping, dropping it), or parts of it (e.g. pushing a button, underlining a phrase in a text document using a pen, hammering a nail into a wooden board).

**Object presentation — gravity, ageing, three-dimensional spatial extensions**

Being situated in the three-dimensional space of the physical world, all objects have a three-dimensional embodiment and a unique geometrical position. Laws of physics further assign properties such as inertia, weight, and light reflection properties to all objects. The agent-observable part of the space at any given point in time is in general defined and limited by the agent’s field-of-view, gaze direction, and the fact that objects closer to the agent occlude objects further away. The position of objects in the vertical dimension is constrained by the force of gravity, fundamentally shaping the way objects are organised by human agents and nature itself. Figure 38 illustrates the impact of the various laws of nature on an office environment. Furthermore, physical objects are subject to micro-physical ageing processes to a smaller or higher degree, processes that have the effect of changing objects’ colour, shape and sometimes even geometrical position as time passes.
Object translation — grab-translate, co-translation with self
Objects are basically translated in the physical space by human agents in a grab-translate (and potentially) -drop fashion. Tools enable human agents to translate larger objects larger distances. With the exception of the translation of objects within the directly operative radius of the human agent (see “object manipulation” earlier), object translation activities always include self-translation, i.e. space navigation.

7.3 The virtual world
As mentioned earlier, virtual environments are more heterogeneous than physical ones. The following general discussion on properties of the virtual world is based on common properties of the four virtual environments (figure 39) offered by the three IT devices selected earlier. The differences and unique features of each of these environments will be discussed in slightly more detail later on in this chapter.

FIGURE 39. The different “faces” of the virtual world explored in this chapter. A WIMP, CLI, cellular phone, and PDA paradigm environment.

Inter-World Event Mediators — interaction paradigm dependent

$P \rightarrow V$

- WIMP: pointing device (typically a “mouse”), keyboard
- CLI: alphabetical keyboard
- PDA: touch-sensitive screen, pointing device (a “pen”), additional mechanical and touch-sensitive buttons, alphabetical and numerical touch-sensitive input areas
- Cellular phone: Microphone, 12-key (alpha)numerical keyboard, Additional mechanical buttons
COMPARING THE PHYSICAL TO THE VIRTUAL WORLD

\[ V \rightarrow P \]

- **WIMP**: high-resolution pixel-based colour screen, polyphonic stereo sound or better
- **CLI**: text or pixel-based black&white or colour screen typically able to display a two-dimensional matrix of 80x25 characters
- **PDA**: low-resolution pixel-based black&white screen (the same as the one used for input), simple sound
- **Cellular phone**: low-resolution pixel-based black&white screen, low-volume loudspeaker for voice transmission, high-volume buzzer for ring tone, network status Light Emitting Diode (LED), vibration actuator.

**Space size — practically infinite (if connected to Cyberspace)**

Just like the physical world can be regarded as infinite with regard to the number of places it includes, so can the virtual world. Being constituted out of the sum of places created and maintained by topologically connected and disconnected virtual environment providers (through LANs such as corporate networks and WANs such as the Internet), and by being based on a finite set of matter and energy belonging to the (finite) physical world, the size of the virtual world is surely finite as well. However, taking a practical stance, the number of places in the virtual world as a whole is to date larger than what can be visited by any single human individual and can thus be regarded as infinite as well. The exception is when the virtual world is accessed through a virtual environment provider that is temporarily or constantly “disconnected” from the rest of the virtual world, in which case the size of the navigable space becomes significantly smaller and also very device dependent.

**Space navigation — manipulating the multifaceted prism**

Because of the fact that virtual space is multi-dimensional, and the fact that human agents’ perceptive and cognitive abilities are tuned to handle three-dimensional spaces, navigation in the virtual world becomes a more complex activity than in the physical. The basic navigation approach taken in most virtual interaction paradigms is to let the human agent control a geometrically two- or three-dimensional “interaction window” into which parts of the virtual world is transformed and presented. Because of the necessary space transformation from n to two or three geometrical dimensions, navigation cannot in general be described as merely a geometrical translation of the interaction window in the virtual space. One way to capture the relationship between the (perceivable) interaction window and the (unperceivable) virtual space is to see the interaction window as a flexible, multifaceted prism through which mutually topologically disperse virtual places and objects can be made to appear close to each other at wish. Navigation then, would be to trim the different prism facets in such a way that the set of virtual places and objects of interest are perceivable in the interaction window, at any given point in time. The prism metaphor also illustrates the remarkable
speed in which navigation in the virtual world can take place since navigation does not involve translation of heavy matter, but light. The absence of geometrical distances opens for more jumpy and discrete navigation than in the physical world.

**The importance of artificial structural constraints**

Since the virtual space is multi-dimensional, there is, in theory, always at least one dimension in which two places are adjacent. Thus, technically speaking, if no structural constraints would be enforced by environment designers, it would more or less be possible to reach anywhere from anywhere in no time. However, since total universal access in spaces is as inhibiting and useless as no access at all, assuming that the space is sufficiently large (and many virtual spaces are), designing the right, orientation-cueing, topological-structural navigation constraints becomes one of the most important and challenging problems in designing virtual environments.

**Screen real-estate**

The often limited size of the interaction window makes the art of careful situative manipulation of prism facets an important part of virtual activities. One can see this managing-screen-real-estate activity as a parallel human agent-controlled process latently applied onto the multi-dimensional virtual space.

**Object manipulation — highly structured**

Manipulation of virtual objects is normally constrained to take place in the interaction window. Thus, any object manipulation has to be preceded by a navigation and/or object translation action that brings the object of interest from the unperceivable part of the space into the interaction window. (Note: in some interaction paradigms this preceding navigation/translation action is hidden and implicitly included in the manipulation action. For example, in order to intra-manipulate a text document, e.g. to add a paragraph, one would normally need to first navigate to the “text document intra-manipulation environment” (a word processor) and from there translate the object of interest into the environment in which the manipulation can take place. However, most virtual environment automatically navigates the interaction window to an appropriate intra-manipulation place as a result of the human agent’s expressed wish to intra-manipulate an object (e.g. double-clicking on the objects iconic representation).

Although the way actual manipulation (state changes) of objects is done differently in different virtual interaction paradigms, the state space of objects are often much more structurally constrained compared to the state spaces of physical objects. The state of any virtual object can be broken down into a finite set of attribute-value pairs

42. The concepts of intra-manipulation and extra-manipulation are defined and discussed in the next chapter (page 179).
while although physical objects also can be described using such a model, it would not be a complete description because, if not for any other reason, quantum physics and chaos theory tells us that the physical world is all but stable.

Just like in the physical world, any more detailed object manipulation requires more or less specialised tools. Since virtual objects are of symbolic nature, the tools tend to specialise on different kinds of symbolic operations. However, as interaction paradigms operate on increasingly abstract representations of the underlying bits, some of the under-the-hood symbolical operations can be presented as seemingly physical operations on seemingly physical objects opening for what in the HCI literature is referred to as “direct manipulation” (Shneiderman, 1983).

Object presentation — highly dynamic

Just as virtual object manipulation is largely paradigm-specific, so is object presentation. There are some common properties however. Objects are usually represented as a two-dimensionally structured set of pixels that, if viewed as a whole, and in an interaction context, has some meaning to the human agent. Since the state of each pixel is independently controllable by the interaction device, there are relatively few technological restrictions on how objects can change in shape, position and colour compared to the physical world. This makes virtual objects and virtual environments in general potentially much more dynamic than the ones found in the physical world. Another related property of virtual objects in many interaction paradigms is the fact that they can be assigned a multitude of significantly different representations depending on interaction context. E.g. sometimes an object might be represented in text form, sometimes as an “icon” and sometimes as a “window”. Furthermore, although virtual objects in general never are affected by aging in any way similar to physical objects, they tend to be very sensitive to changes in the usually quite dynamic surrounding environment. At least in the more complex and advanced environment providers, the virtual environment surrounding virtual objects tend to not age but rather become younger! That is, more or less task-specific virtual tools and environments tend to be frequently updated with more functions and features which in turn make users want to more or less regularly “upgrade” their working environment. Large age differences between objects (domain objects, tools, etc.,) tend to create tensions which can only be overcome by constantly renewing the internal structure of older objects so that they are “compatible” to newer ones.

Object translation — inexpensive, largely geometry-independent

The support (and need for) translation of objects within virtual space is highly paradigm-dependent. Some virtual paradigms reduce the role of object translation by substituting it with navigation, while others have important functionalities tied to the translation of objects from one place to another. The more complex paradigms let nav-
igation and object translation complement each other or even offer redundant ways of achieving the same result: you might choose to navigate-grab-navigate-and-drop an object or you might stretch your long arm and translate the object from a distance.

Some virtual interaction paradigms designed for large interaction windows (large screen real-estate) allow for an additional kind of object translation only affecting temporal spatial relationships within objects in the interaction window, without necessarily affecting objects’ positions in the “real” long-term virtual space as such.

Just like navigation in the virtual world in general is inexpensive, so is translation. Furthermore, and unlike in the physical world, the translation of objects is a process tightly related to object cloning (or copying) which is also very inexpensive compared to the same process in the physical world. In fact, in practice, cloning an object can sometimes be even more cost-effective than moving it because the latter action requires the original object to vanish while the former can leave entirely as it is.

Unlike in the physical world, objects’ geometrical size (should they have one) do seldom affect the translation process. Instead, it is their size in bits (which in some sense can be regarded as the virtual correspondence to physical objects’ mass) that does. And just like in the physical world, the cost of translating heavy objects is paid within a trade-off framework having money, time and energy as its corner stones.

7.4 Specific virtual interaction paradigms

The following overview of some virtual interaction paradigms is not intended to be complete or very detailed in its presentation. The aim, rather, is to show the diversity as well as similarities among some representative existing paradigms through examples.

Windows, Icons, Menus, Pointing device paradigm (WIMP, Microsoft Windows 2000)

Inter-World Event Mediators

- P→V: pointing device (typically a “mouse”), keyboard
- V→P: pixel-based colour screen (high-resolution 1280x1024 pixels is common at the time of writing), polyphonic stereo sound or better

Space size

Most WIMP paradigm devices offer a constant or nearly constant access to the Internet. This makes the size of the accessible virtual space practically infinite. Even without access to WANs or LANs (which becomes increasingly uncommon situations), the navigable space is relatively large because WIMP devices are typically fitted with large local data storage (hard disks).
Space navigation
Since IT devices using the WIMP paradigm usually provide a relatively large interaction window (see figure 40), more than just one object can be inspected at a time. The WIMP paradigm provides an euclidean two-dimensional area typically referred to as the “desktop” where sets of objects can be viewed side by side, similar to how it can be done in the physical world. The desktop does to some extent afford also organisation in a third dimension but since objects tend to have only one single two-dimensional surface, and since this surface is always facing in the same direction, the desktop space is sometimes referred to as a 2.5-dimensional space. The possibility of organising some of the objects in the multi-dimensional space in two or (almost) three spatial dimensions calls for a different navigational metaphor than in the case when interaction paradigms basically only allows one object at a time to be inspected in the interaction window (the case in all three other paradigms discussed later in this section). Maybe the navigation activity in WIMP environments is best viewed as an inherently parallel activity where navigational actions on each object present in the interaction window (or more correctly, the 2.5-dimensional desktop space) is seen as an independent action, but where the sum of the navigational actions on all these objects define the overall virtual space navigation activity. It should be noted that the navigation actions given below can often be performed in alternative ways (i.e. using keyboard instead of mouse) in addition to the ones presented. However this does not change the presented navigation mechanisms and the effect they have in the virtual space.
• Icon→window cloning - Objects represented as icons (see object presentation below) can be cloned into window objects by “double-clicking” (pressing one of the buttons on the mouse PVA twice within a short time span while having placed the corresponding mouse PVA directly above an icon object) or single-clicking on them. The cloning is typically presented in one out of two ways: Either the icon object remains where it is, and its window clone appears on the uppermost layer in the z-dimension of the interaction window (e.g. when starting an application by clicking on its icon representation), or, the window object containing the clicked icon object is completely replaced with the cloned window object (e.g. when drilling down in a file hierarchy on the hard disk, or when following links on the WWW).

• List expansion - When list objects (“combo boxes”) are activated, they temporarily cover parts of the interaction window with a set of icon objects. Clicking outside the list object makes it disappear, putting the interaction window in the state it was before the list object was activated. Selecting an icon object situated within the list object also makes the list object disappear, but typically also changes parts of the content of the interaction window, as a result of the selection of the icon.

• Menu bar activation & menu item selection - Window objects, and the interaction window itself, typically contain “menu bars” statically situated somewhere in the upper or lower part of the two-dimensional space they occupy (the area in between is the area containing domain objects, i.e. the “work area”). These hierarchically structured sets of icon objects (menu items) mainly provide tools for manipulating objects situated inside the specific window object and the window object itself. In its deactivated mode, menu bars are typically presented as horizontally lined-up icons. A menu bar is activated by clicking on one of these “root” objects, revealing the unique descendant menu objects of that particular root object as a vertical list of icon objects. Clicking outside the menu bar makes all menu items disappear, except for the “root” objects, putting the environment in the state as it was before the menu was activated. Selecting (clicking on) one of the descendant icon objects of the menu bar also makes all descendant objects disappear, but typically also changes parts of the content of the interaction window, as a result of the selection of the menu object.

• Scrollbar dragging - Whenever object sets or single objects are become spatially too large to be completely shown within the area of a window object, scroll-bars on the right side and on the bottom of the window object allows for vertical translation of the contained object set relative the object window. Thus, window object are not constrained to contain a finite maximum number of objects, or only allow objects up to a maximum geometrical size, for mere spatial reasons.

• Desktop space layer navigation - The 2.5-dimensional desktop space can in general be regarded as a (very small) space independent from the (very big) rest of the virtual space. Only in very specific situations do these two spaces share object
relationships (see “object translation” below). Navigation in the desktop space is
done mainly by selecting what window object out of all present ones that should
be placed on the uppermost layer in the z-dimension. In practice this is done
either by traversing the objects by pressing “Alt+Tab” on the keyboard, or by
selecting the desired object on the always present “task menu bar”.

Object manipulation
The WIMP interaction paradigm has in practice been enabled by the invention of
pointing devices such as so called “light pens” and the computer mouse. The mouse
is a physical-virtual artefact that basically mirrors the spatial position of it’s PVA on a
flat surface in the physical world to the spatial position of it’s PVA within the interac-
tion window of a virtual environment. This PVA has in later years got competition
from other kinds of equivalent “pointing devices” (pads, pens, trackpoints) where the
original PVA has been substituted with artefacts possessing physical properties other
than the original. The PVA however, has more or less remained unchanged and is most
of the time represented as a small arrow in the interaction window, hovering above all
other objects in the depth (0.5) dimension. The PVA part of the pointing device also
typically offers one, two, or three buttons whose state changes are immediately medi-
ated to the virtual world as well. Since these buttons are fitted on, or in the direct vi-
cinity of the part of the PVA that is position-tracked, the general impression of
pressing the physical buttons is that the pressure exerted on them is directly mediated
to, and further on through, the PVA onto any virtual object or surface residing directly
below the PVA in the depth dimension. To maintain this physical-virtual illusion, the
WIMP paradigm typically provides visual and potentially auditory feedback whenever
the force is mediated. The overall result is that virtual objects in WIMP environments
can, in a limited way, be “directly” manipulated by human agents performing physical
actions mediated through the pointing device.

Complementing the mouse, a more traditional computer input device is used for
object manipulation: the alphabetical keyboard. Although the keyboard can play dif-
ferent roles in different interaction contexts (as can the pointing device), its main role
within the WIMP paradigm is to mediate text and text editing operations. In that par-
ticular interaction context, the mouse typically acts as a “global” text navigation and
selection device while the keyboard acts as a local text manipulation device. In general,
human agents could theoretically speaking choose to perform all WIMP activities
with only one of the two devices. However, because of their significant advantages in
specific interaction contexts, alternation and combination of the two is for most ac-
tivities much more efficient.

Object manipulation in the WIMP paradigm can be performed by human agents
in numerous ways. Tailor-made sub-environments (“applications”) have a certain free-
dom in offering their own manipulation paradigms although conformity among such
sub-environments has been regarded by designers as important since the beginning of
the WIMP paradigm’s existence. One general manipulation procedure is where the
object (or objects) to be manipulated are first selected using the pointing device, followed by the selection of a tool or function that should be applied to the selected object(s). Simple tools give immediate effect while more complex tools require manipulation of the tool itself before it can be applied. The tool selection action is performed by the human agent by using any of the navigation mechanisms mentioned earlier. Except for tools offered by the desktop object, most tools are only applicable within the window object where they were selected.

An example of a simple tool would be what can be denoted as the “text editing” tool. Having selected both the tool and the object to be manipulated (a text object), it transforms keystrokes on the keyboard into text appearing in the selected object. It also typically supports local navigation within the text object. An example of a more complex tool is what can be denoted the “import file” tool found in many menu bars of window objects, allowing to copy and transform an icon object situated almost anywhere in the virtual space into a window object situated within the realms of the window object in which the tool was invoked. This tool requires navigation actions (to locate the desired object to translate/clone) before the application of the tool can be successfully completed.

Some object manipulation actions, usually centred around pointing device operations, belong to the core actions of the WIMP paradigm and as such do not require any tool to be selected before the action is performed. The selection of objects (in menus, in windows, or any other place) by clicking on them is such an operation. The resizing of a window object by “dragging” its borders (clicking on one of the borders and keeping one of the buttons of the pointing device pressed while moving it) is an example of such a core manipulation mechanism.

Another core paradigm manipulation mechanism is the “pop-up” menu activation mechanism. Contrary to the menu bars statically present in window objects, pop-up menus are completely invisible when not active but instead activated by clicking with an alternate mouse button on a specific object. They are usually visualised in the vicinity of the point where the click was performed, and lists menu objects (tools) that can be used for manipulating or inspecting the specific object.

Object presentation
Although there are more object types within the WIMP paradigm, windows and icons probably play the most important role. Since most objects can take on both shapes, one might assume that the object properties available for inspection and manipulation should be roughly the same no matter what. Instead, the attribute manipulations afforded by icon representations and the ones afforded by window representations are fundamentally different. In fact, one reason for making the distinction is the combination of a need for different manipulable attributes of the one and same object in different interaction contexts, and the inability of the virtual environment provider to afford all necessary attribute manipulation possibilities all in one single object representation (like in the physical world). The reason for this inability is partly because of
technological limitations (screen size, 2D visualisation instead of 3D, relatively big pixels, interaction delays, limitations in IWEMs, etc.), partly because of the fact that virtual objects can potentially be viewed from more than three different “angles”, in contrast to physical ones. Many virtual objects simply possess more attributes than what can be represented in one single representation. The most evident differences between the two different representations is that the icon representation uses less space but reveals a very limited number of attribute states while the window representation allows for a highly detailed presentation of object attributes at the cost of screen real-estate. The latter factor can however be adjusted to some extent since window objects often are “elastic” (see “object manipulation” earlier) and thus allow for fine-tuned management of screen real-estate. The attributes presented by both the window and icon representations can be divided into core paradigm visual attributes and object specific attributes. The former attributes which most objects possess (e.g. all window objects have a “name bar”, many windows have resize bars; all icon objects have name area and/or a small picture associated to them) allowing for administration and manipulation of each object’s unique attributes and attribute values.

Object translation
The most straight-forward way to translate objects in the WIMP paradigm is probably by “dragging” an object (see “object manipulation earlier”) from one position to another within the interaction window. Although being possible to do with both icon objects and window objects, the meaning of the action differs. When window objects are dragged, it only affects the layout within the interaction window and does not change any relationships between objects in what can be called the “long-term” virtual space. Somewhat more effective in the long-term is the dragging and dropping of icon objects within the window object the dragged object was originally situated in. This operation permanently changes the spatial relationship between the dragged object and it’s “sibling” objects situated in the same window object. The translation operation generally having largest effect is probably the one when an icon object is translated from the window object it was originally situated in, to a different window object. In this case, depending on the type of window objects involved in the operation, at least two different outcomes are possible:

- If both window objects are visualising parts of the “long-term” virtual space, the dragged object will be either permanently moved or permanently copied from the virtual place represented by the origination window to the virtual place represented by the destination window.
- If the situation is as before, but the destination window is representing an application, the translated icon object will be copied and transformed into a window object within the realms of the destination window object.

As mentioned in the object manipulation section earlier, the same translation operation as the one described above (although it might perhaps be more of a copy operation in disguise rather than a translation) can be performed from within the
destination application using the “import file” tool present in most window objects of the “application” type. A third alternative way of performing the same thing would be to utilise the powerful “clipboard” mechanism that the WIMP paradigm offers. Almost any kind of icon object almost anywhere in the virtual space can be copied first into the clipboard and later from the clipboard to the destination window object. The strength of this mechanism is that it allows the human agent to perform the navigation to the destination window object after having initiated the translation operation which perhaps is more intuitive and certainly more “situative”. A mechanism related to object translation is also the possibility of creating “short-cuts” or “clones” of objects, place them somewhere else in the virtual space, and to let them point to the original object. The short-cuts or clones do not contain any information content except the virtual space coordinates (e.g. a Uniform Resource Locator, URL) of the object it represents. The strength of the short-cut mechanism includes the fact that it allows for less data redundancy by always pointing to the right object even if the destination object has been manipulated since the creation of the short-cut, and that it contribute to decrease the amount of data that needs to be handled both by human agents and by the environment providers. The backside is that the link is one-way, potentially leading to short-cuts that point to nowhere when the destination object has been moved or destructed without notice.

Command Line Interface paradigm (CLI, SunOS 5.8 C shell)

Inter-World Event Mediators
- P→V: alphabetical keyboard
- V→P: text or pixel-based black&white or colour screen typically able to display a two-dimensional matrix of 80x25 characters

Space size
Just like WIMP paradigm devices, also CLI-paradigm devices typically offer a constant or nearly constant access to the Internet. In fact, CLI paradigm environments are often presented as a places within WIMP spaces and “run” on the one and same hardware device. This makes the space offered by most CLI paradigm providers practically infinite as well. However, because of object presentation limitations in the paradigm (only text objects, or text manifestations of objects can be inspected), the CLI space becomes significantly smaller nevertheless.

Space navigation
Although seldom visualised as such, the space is organised as a semi-lattice (a slightly relaxed version of a hierarchical tree) and navigated by typing text-based commands on a keyboard. The typed letters are visualised at the input area at the bottom of the manipulation area and “scrolled” upwards and mixed with object information as the strictly dialogue-driven activity proceeds (figure 41). On the left side of the input area
(the “prompt”) orientation information is typically displayed, showing the path from the current location to the root of the semi-lattice. Moving upwards in the hierarchy (towards the root) is performed by the command sentence “cd ..” where cd stands for “change directory” and “..” for the place (directory) directly above the current place in the semi-lattice. Moving downwards in the hierarchy is done with the same command, having the desired name of the place as parameter to the command instead of “..”. Jumps can be made to any place in the space by giving the absolute or relative (to the current place) path as parameter to the “cd” command. In order to inspect the content of the current place, the command “ls” can be issued, displaying the list of objects at the current place.

Object manipulation
Objects are typically manipulated by issuing a manipulation command (the physical correspondence would be to grab a tool) followed by the name of the object that should be manipulated by it.
Simple tools, e.g. the ones who do not require extensive dialogue with the human agent during the course of performing the demanded manipulation, typically provides any feedback to the human agent by writing text in the manipulation area, before returning the “prompt” (the sign that the manipulation has been performed completely). No feedback information typically means that the manipulation was successful.

More complex tools might substitute the “normal” content of the interaction window with some tailor-made textual environment providing tailor-made manipulation commands and not seldom even a tailor-made interaction sub-paradigms. One such common sub-paradigm is one where object details are displayed at the major upper part of the interaction window and names of tailor-made manipulation commands at the lower minor part of the interaction window. Commands in these sub-paradigms do not necessarily have to be typed in as a whole, but are not seldom invoked by simply pressing a dedicated key on the keyboard.

**Object presentation**

Objects are by default presented by their (locally unique) names in text form whenever a place inspection command such as “ls” is issued. Whenever objects are to be inspected more closely, their “contents” can be viewed (also in text form) applying the command “more” on the specific object of interest. Then, the object is visualised in largest possible detail as an amount of text.

**Object translation**

Objects are moved from one place to another by applying the command “mv” (move) directly followed by the object of interest and subsequently followed by a destination path.

**Personal Digital Assistant paradigm (PDA, Palm OS 3.5.0)**

**Inter-World Event Mediators**

- P→V: touch-sensitive screen, pointing device (a “pen”), four mechanical buttons, one three-state mechanical navigation button, six touch-sensitive buttons, alphabetical character input area, numerical character input area
- V→P: pixel-based black & white screen (the same as the one used for input, low-resolution 160x160 pixel matrix is common at the time of writing), simple sound

**Space size**

To date, most PDA paradigm devices lack constant connection to WAN and LANs. They are also typically fitted with relatively small local data storage hardware. These two facts make the navigable space relatively small compared to WIMP and CLI paradigm spaces.
Space navigation

On the highest abstraction level, objects are organised in places depending on what “category” the objects belong to. Categories are defined and assigned to objects by the human agent. Among the default categories are “Main”, “Utilities”, “System” and “Unfiled”.

The space is navigated through the use of several interaction mechanisms

• Icon object expansion - Objects represented as icons (see “object presentation below” for a description and figure 42 for an illustration) are activated by “clicking” on them (physically touching the screen using the pointing device at the position of the object), replacing the contents of the interaction window with details of the object.

• List expansion - List objects (“combo boxes”) are activated, temporarily covering parts of the interaction window with a set of icon objects. Clicking outside the list object makes it disappear, putting the environment in the state as it was before the list object was activated.

• Menu activation & menu item selection - The touch-sensitive button called “menu” is activated, revealing a drop-down menu bar on the upper area of the interaction window, containing icon objects (menu items). Clicking outside the
menu makes it disappear, putting the environment in the state as it was before the menu was activated.

- Physical button navigation - Depending on what object that is currently fully expanded in the interaction window, the five mechanical buttons as well as the six touch-sensitive buttons can be assigned navigation functions. E.g., in the situation when the “application browser” object is fully expanded, i.e. interaction on the highest level of abstraction, the four mechanical buttons are associated to full-screen representations of specific virtual objects. Although in general, these objects can be activated without pushing these buttons residing outside the interaction window (by making use of navigation mechanisms present within the interaction window only) the idea is that the space-multiplexing factor (Buxton, 1986; Fitzmaurice, 1996; Fitzmaurice et al., 1995) that these physical buttons give rise to allows for faster navigation.

- Scrollbar dragging - Whenever object sets are too big to be completely shown within the area of the interaction window, a scroll-bar on the right side of the interaction window allows for vertical translation of the object set relative the interaction window. This virtual navigation mechanism is typically optionally availability through physical manipulation of the three-state mechanical navigation button as well.

Object manipulation

Object manipulation is in general performed by first activating the icon representation of the object of interest, then to navigate within the full-screen representation of the object to the property of the object that should be altered (i.e. to bring the property into the interaction window), and finally, to alter the property. All three steps can involve any of the interaction mechanisms listed in the “space navigation” section above. The last step might also involve text input as an alternative or only interaction mechanism. Text input is performed by “writing” letters one by one in one of the two character input areas by physically drawing the characters using the pointing device. The characters have to be drawn in correspondence to an alphabet especially designed for unambiguous automatic character recognition.

Object presentation

Objects do typically present themselves in two different ways depending on the interaction context: as icons or full-screen objects. The full-screen representation of an object can be viewed by activating the corresponding icon representation (“clicking” on it), revealing details about the objects utilising the full area of the interaction window. The icon-representation typically shows a small illustrative picture in conjunction to a short text describing the object. Sometimes only the picture or only the text is used. The icon-representation allows several objects to be simultaneously viewed in the interaction window.
Object translation

Objects on the highest abstraction level can be translated between different places in the virtual space, i.e. different “category-places” (see space navigation earlier), by activating the “Category” object in the application browser menu. A list of all objects and the category to which they belong (the manipulable property) is shown. When navigating back to the application browser object, any changes done by manipulating the category object is immediately evident.

Support for object translation in other interaction contexts, i.e. on a lower abstraction level, is low except for in the case of object manipulation using the text input mechanism in which text chunks from one place can be moved or copied to another place by the use of a general cut-and-paste mechanism. This mechanism is accessible through special two-stroke entries into the character input area: “/C” for copy, “/X” for cut, and “/P” for paste.

Cellular phone paradigm (Ericsson T66)

FIGURE 43. The virtual environment offered by a cellular phone.
One of the eight top-level places is the “phone book”.

Inter-World Event Mediators

- P→V: One two-way volume key, one microphone, and one 18-key keyboard consisting of a 12-key (alpha)numerical keyboard; left/up and right/down navigation keys; “yes” and “no” keys; one “alternative” key and one “delete” key.
- V→P: pixel-based black & white screen (low-resolution 101x65 pixel matrix), low-volume loudspeaker for voice transmission, high-volume buzzer for ring tone, network status Light Emitting Diode (LED), vibration actuator.
Space size

Many cellular phones available today, including the one studied here, offer a relatively limited access to the Internet in addition to the “local” virtual space. The Internet access is limited in four different ways: a) the connection process is in general complicated and slow, b) the bandwidth of the connection when established, is in general very low, c) the limited screen real estate constraints the kind of information possible to present, and d) the part of the Internet that is possible to navigate (provided through cellular paradigm designated communication protocols such as WAP) is significantly smaller than the part navigable by WIMP and CLI paradigm devices. Furthermore, the locally navigable space is in general very small and by and large dedicated to the storage of phone numbers and contacts, naturally. All in all, while the cellular paradigm space in theory is fairly large (larger than for instance the PDA paradigm space), it is in practice very small. In the following discussion on the paradigm, I have pragmatically chosen to omit commenting the Internet part of the virtual space because of the practical problems accessing and navigating it.

Space navigation

The virtual space is organised at the top level as a list of eight places where the last place is connected to the first, creating a cycle (see figure 43 for a “screen shot” of one of the places). Each place in turn contains another cyclic list of objects and so on. This hierarchical structure of cyclic lists is traversed using the “left/up” and “right/down” keys for navigating the current list and “yes” and “no” buttons for going down and up in the hierarchy respectively.

While in the “root” place (one of the eight top-level places and also the place which is the one shown when the device is turned on), many places are reachable by different kinds of short-cut mechanisms as well. I.e. by holding down an alpha-numerical key for more than a certain amount of time, the phone book object is immediately expanded and positioned depending on the key that was pressed; by keeping and holding the “yes” button, the call list object is expanded, revealing recently called persons and numbers, etc. Apart from these inbuilt “short-cut” key-to-place mappings, human agents can make their own as well.

Object manipulation

Depending on the complexity of object attributes, objects are manipulated in one out of two ways. Simple object attributes having a limited state space, such as for instance the “automatic key lock” attribute having the state space {on, off}, are changed by choosing the desired attribute values from a list using the space navigation mechanisms presented earlier. Whenever the state space of an attribute is too large for this approach, the attribute is manipulated by using the (alpha)numerical keyboard. A typical case is when the name or number of a new contact is to be added to the phone book.
Object presentation
The objects in the top cycle list are presented as a deck of cards where the top card covers all other cards except for their “tabs”. The tabs contain small illustrations intended to signal the content of it event when not on top. The top level card is presented with a combination of illustration and text. Objects and object attributes situated below the top level cycle are with a few exceptions presented as text elements in vertical cyclic lists. Some variations exist in the bottom layer of the hierarchy where a mix of graphics and text are sometimes used for specific purposes.

Object translation
The only object translation mechanism present is the customizable short-cut mechanism allowing human agent-selected objects (phone number entries) to be accessed from the “short-cut place” in addition to the standard place dedicated for storing these objects (the “phone book place”).

7.5 Summary: general differences and similarities
Based on the previous analysis, general differences and similarities between the physical and the virtual world can be derived. Some of the world properties (i.e. similarities and differences) can be regarded as fundamental and unavoidable properties of the specific world, like for instance gravity in the physical world, or the multi-dimensionality of the virtual world. These properties are not seldom also properties that define the essence of being in the specific world. Other properties are more plastic and possible to alter by design, such as the level of “presence” of virtual objects (using haptic, and other multi-modal IWEMs) or the level of perceivable inertia in navigating the virtual world which by default can be regarded as being zero. The distinction between fundamental and plastic world properties is important because it indicates what is possible and impossible to achieve as a designer in the physical-virtual design space. Some fundamental and plastic world properties will be discussed below, followed by a more complete list in table 8 on page 161.

Space size

Similarities
• Both worlds are practically infinite (unless constrained by infrastructure such as locked doors or lack of data communication network connection) although they in an objective sense are finite. The physical world is ultimately finite because the universe, according to the knowledge of our time, contains a limited amount of energy and matter. The virtual is finite because it is in strict physical sense built out of physical energy and matter. However, adopting a pragmatic human-centric perspective, both worlds are in practice infinite. This is because no matter what vehicles we choose to use when travelling the physical or the virtual world, we can
never during our lifetime, see it all. In this sense, the human capability of explor-
ing the worlds is more limited than the worlds themselves and thus, they become
in practice infinite for any single human being. The space size property is plastic
because it can be intentionally or unintentionally altered e.g. by limiting access to
parts of the physical world by locking doors, or by isolating a virtual environment
by disconnecting it from the rest of the virtual world.

Space navigation

Similarities

- Virtual “local” space shares many features with the physical world. This fact is
probably caused by the immense success of the idea to use the physical world as a
metaphor when designing the appearance of the virtual. The idea is fundamental
to the WIMP paradigm and has spread to many other virtual interaction para-
digms. The exception is the CLI paradigm whose “local” space has very little
resemblance with anything in the physical world. The CLI paradigm was also the
standard paradigm for many years before the WIMP paradigm appeared. Except
for niched computer professionals, the WIMP paradigm has become the de facto
standard because its relatively low novice user-threshold and its wider range of
support for different common virtual and physical-virtual activities. This is illus-
trated by the fact that the CLI paradigm environments today often are invoked
from within a WIMP environment only whenever niched “hard core” computer
administrative activities are to be performed.
- Short distance travel is inexpensive and relatively quick in both worlds.
- Navigation activities can be viewed in terms of object manipulation in both
worlds.

Differences

- One big difference between the physical and the virtual world is the structure of
the “hidden” space. The higher dimensionality of virtual “hidden” space affords
many of the features that motivates the existence of the virtual world. The global
virtual space can be faster, less expensive, and sometimes easier to navigate com-
pared to the physical global space. Furthermore, easy object translation and clon-
ing also affects navigation by affording low-cost temporary reorganisation of
the space to fit current needs within an activity. The navigational strength of the
physical world lies in its support for local navigation by affording the use of all
senses of the human to their full extent.
- The cost for “long distance” travelling is in general much higher in the physical
world than in the virtual.
- Navigation in the physical world is continuous and relatively slow compared to
navigation in the virtual world which tends to be quick and immediate. Naviga-
tion from one place to another in the virtual world can be done without the sen-
sation of actual travel. This property is plastic because navigation in the virtual world can be designed to mimic the “continuousness” of the physical world e.g. by using three-dimensional virtual interaction paradigms such as the ones common in Virtual Reality settings.

**Object manipulation**

Object manipulation in the physical world is more direct than in the virtual world. However, because of the ability of the human perception and cognitive systems to overlook weaknesses and contradictions in virtual object appearance and behaviour, some of the indirectness in virtual object manipulation can be disregarded. In fact, as explained in the end of chapter 4, this stance is taken throughout this whole thesis.

**Similarities**

- The one and same object categorisation can be found intuitively viable in both worlds. For instance, most objects in both worlds can be categorised as belonging to the group of tools, domain objects, or places.

**Differences**

- Laws of nature eventually restricts the behaviour of objects in the physical world. Virtual objects are in theory less constrained, but are nevertheless restricted to behave and look according to the “laws of nature” set up by the virtual interaction paradigm in which they are situated.
- The virtual world affords trial-and-error manipulation strategies by offering a perfect and inexpensive mechanism for reversing previous operations: “undo”.
- By representing data using electrons rather than heavy matter, the virtual world offers advanced, yet inexpensive, tools for symbolical transformations of large sets of data, compared to the physical world.

**Object presentation**

**Similarities**

- Objects playing similar roles in the physical and the virtual world can be made to visually look and behave in similar ways. This property is plastic to the extent that both physical and virtual objects are artefacts shaped by designers who have the power to make them look and behave similar within constraints posed by fundamental properties of the two worlds.

**Differences**

- Objects in the physical world do not and cannot transform their shape radically. On the contrary, virtual objects are often presented in more than one visually significant different form depending on interaction context.
- Physical objects can be touched, heard, thrown, eaten, smelled and squeezed. Virtual objects are in general only perceived through visual perception occasionally
complemented with audio. This property is to some extent plastic in the sense that it is possible to come close to physical-world experience levels using multi-modal IWEMs as used in immersive Virtual Reality settings.

**Object translation**

**Similarities**
- Virtual local object translation (i.e. within the perceivable/manipulable part of the virtual space) can be designed to look similar to its physical counterpart since both the physical space and the virtual (local) space is organised in two or three dimensions. (For this to be viable however, the screen real estate of the virtual environment provider has to be sufficiently large.) The property is plastic because from a technical standpoint, object translation in the virtual world does not per se have to be afforded in any way similar to how translation is done in the physical world because the fundamental properties of the physical world constraining physical object translation (mass, inertia, geometrical dimensions etc.) are not present in the virtual world.

**Differences**
- Virtual objects can in general be easily cloned or translated long distances while physical objects in general are hard or even impossible to clone or translate longer distances.

### 7.6 Related work

The analysis performed in the previous sections of this chapter has been based on a set of physical and virtual phenomena chosen with the interest of highlighting navigation and manipulation operations. Studies with somewhat more application-specific focuses have been performed in other contexts by other investigators of the physical-virtual design space. In order to broaden the perspective, but also to show that the very general similarities and differences between the physical and the virtual worlds presented in this chapter possess some ecological validity, brief summaries from two other related studies (Arias, Eden & Fischer, 1997; Sellen & Harper, 2002) are presented in chapter 3 (see “Comparing the physical to the virtual world” on page 58).

### 7.7 Conclusions

Navigation and object manipulation differ in many ways between the physical and the virtual world. One underlying reason is the low weight and compact geometrical dimensions of digital bits (when not inspected) if compared to physical matter. The virtual world becomes much more flexible because of the relative independence from laws of nature. On the other hand, it can be designed to mimic the physical world with many of its constraints at wish.
Both worlds are practically infinite with regard to the number of accessible places and objects.

The visually perceivable/manipulable part of the world is presented in two and/or three spatial dimensions.

Travelling short distances (e.g. between places in the perceivable/manipulable part of the worlds) is quick and inexpensive.

In both worlds, navigation activities can be seen as being performed through object manipulation operations. Sometimes it is oneself (or ones representation) that is manipulated, sometimes it is an “external” object.

A large number of objects in both worlds can be intuitively categorised as belonging to one of the categories of tools, domain objects, or places depending on the purpose of the analysis.

**TABLE 8:** Some general differences and similarities between the physical and the virtual world with respect to space size, space navigation, object manipulation, object presentation and object translation. The shaded properties are considered “plastic” and thus possible to affect by design. All other properties are regarded “fundamental” and thus, impossible or at least very hard to alter by design, should one want to do so.
The fundamental differences will remain because they deliver significant added value to the physical-virtual activity space. The fundamental differences listed in table 8 (the “unshaded” properties in the left column) signals why the physical and virtual worlds are likely to continue to look and feel different. Surely, efforts can be made to minimise these differences but only to the cost of reduced support for physical-virtual activities that rely on the unique affordances and constraints that the two worlds can offer. One could for instance choose to categorically reduce the multi-dimensional virtual space to three dimensions (removing any functions that allow “instant shortcuts” between distant places in the three-dimensional space) in order to reduce the difference in dimensionality between the worlds. Although this might be viable in some specific settings, the cost is obviously very high since the fast global navigation support is one of the most useful features of the virtual world.

43. Hypothetically, to reduce properties in the two worlds to some kind of “most common denominator”, only allowing phenomena that can be made to exist in both worlds, would lead to not only the discarding of the virtual world entirely, but also the discarding of the features of the physical world that actually makes it the physical world.
The fundamental similarities will remain because humans need them

The fundamental similarities (the “unshaded” properties in the right column in table 8) represent properties of the worlds that can be viewed as originating from how human object-oriented activities take place, based on cultural traditions as well as basic cognitive and perceptional abilities of human agents. These properties are the way they are because they have to in order to be useful and understandable by humans. E.g. we do not know how to visualise a space of higher dimension than three (neither in our minds nor in the world we perceive through our senses); we are used to organise space into more or less tailor-made places, containing tailor-made tools for manipulation of specific object types; we cannot act in a world where local navigation is slow.

The plastic properties are open for application-specific design considerations

The plastic properties (the shaded areas in table 8) represent similarities and differences that can be more or less altered by design. E.g. the size of the two worlds can be temporarily shrunk by closing doors or by disconnecting an IT device from the Internet; the navigation in the virtual world can be designed to become as smooth as in the physical world; physical and virtual objects can be made to look and behave in similar ways; virtual objects can even be made as tangible as physical ones.
This chapter presents a conceptual framework for modelling object-centred human activity as a phenomenon taking place in a unified physical-virtual space. The model is centred around a simple object categorisation and the way objects enter and disappear from “active” parts of space in the course of activities. The model is denoted “situative” because it is intended to model specific physical-virtual situations at a specific point in time. Being as much a conscious and intentional design stance as a tool for analysing basic human activity, it constitutes the platform for the more concrete model developed in the next chapter where structural properties of the physical-virtual space will be discussed. The development of the situative model in this chapter is presented in three steps where the first two act as a conceptual base for the third, as follows:

- A physical-virtual design perspective
- General world-neutral spatial definitions
- A situative model of physical-virtual space

As described in more detail in chapter 4 (page 81 and onwards), this thesis focuses on activities that, more or less, a) have a clear meaning, b) are observable by a human agent, and c) are (potentially) observable by an artificial agent. My interest in these “simple” human activities led me in the previous chapter to investigate the role of space navigation, object manipulation, and object translation in the physical and the
virtual world. This chapter tries to go further, by defining a conceptual framework able to describe these basic human activities uniformly no matter if they take place in physical, virtual, or as will be shown, physical-virtual space.

8.1 A physical-virtual design perspective

“The trick, then, is to consider both the affordances of paper and the digital alternatives to ask what together they could provide for the specific kinds of goals people have. With this knowledge in hand, designers can create combinations of the best of both the paper and digital worlds.” (Sellen & Harper, 2002, p. 143).

I propose a design perspective that acknowledges and is centred around the existence of physical-virtual activities. The main goal with the framework is to transform what has traditionally been the designer’s decision of in what world a specific action should be performed, into a situative decision made by the human agent while performing the specific activity. E.g. letting the human agent decide whether to read a document online or on paper at the time of use, rather than as a designer constraining the activity (at the time of environment design) to be performed in one or the other of the two environments.

Designing for physical-virtual in conjunction with physical and/or virtual

This framework can hopefully enable designers to expand their design/analysis space by adding physical-virtual actions (see definition 5.1.2, page 92) and activities (definition 5.1.3, page 93) to the existing kinds of phenomena they are concerned with today, i.e. either mainly physical or mainly virtual activities. Figure 44 illustrates the conceptual change from treating physical and virtual environments as separate design spaces to the perspective of designing for one single physical-virtual space.

Furthermore, the proposed framework can inspire to view physical phenomena from a virtual perspective and vice versa, potentially leading to new and improved activity support in future environments.

Common terminology

The framework operates on a high abstraction level, giving designers with a “physical” design background an easy entry into the design of virtual objects and properties and vice versa. Furthermore, the framework introduces a common terminology allowing a) physical and virtual designers to communicate designs to each other more easily, b) designers to a higher degree withhold decisions as to whether their designs should be implemented in the physical or the virtual world.
Limitations
I believe that the high abstraction level makes the framework relatively stable and useful for the foreseeable future. However, as new HCI paradigms emerge (and the framework is itself ironically a tool that pushes designers to explore alternative paradigms) some of the underlying assumptions might not hold. In particular, the framework does to some extent rely on a 2.5-dimensional spatial presentation of the virtual world such as offered by WIMP-based systems. Furthermore, the framework is probably best suited for the analysis and design of environments that are mainly used for physical-virtual activities. In the case of environment design in which outer circumstances inevitably constrain the activities to be performed completely in either one or the other world, the framework will prove to be of little use and a more classical design framework is a better choice. This is because the framework relies heavily on the possibility of “designing away” obstacles within an activity by moving actions (or rather, affording the possibility of moving actions) from one world to the other. If this option is not there because of outer circumstances, the framework loses its most powerful conceptual mechanism. Outer limiting circumstances can be for instance organisational policies and culture (“At McLawyer Inc. we (are to) handle all client communication using paper letter.”), technological limitations (no one has invented a way to reorganise the books in a bookshelf automatically yet), or plain economical reasons (the efficiency gain by redesigning all administration offices into highly integrated physical-virtual environments is perhaps not always large enough to motivate the investment). All three mentioned limiting factors are however also subject to changes as technology develops, becomes more affordable, and in turn forces organisational policies, cultures, and laws to change as well.
Another limitation is that since the framework operates on a high abstraction level relatively far away from implementation and real-world activity issues, some designs that look fine within the framework might not work well in practice. Thus, like any other high-level theoretical model, it has to be used as a design tool among others, including for instance empirical evaluation.

**Background**

The physical-virtual design perspective follows the spirit of several recent research trends within HCI and related areas. Some of them are briefly discussed in chapter 2. Apart from the general challenges and solutions discussed in these new research communities, the physical-virtual perspective has found its main inspiration in a few, seemingly uncontroversial observations and questions.

- An increasing amount of human activity seems to be physical-virtual, i.e. have components situated in both the physical and the virtual world.
- Physical (industrially produced) artefacts are almost always, and more or less substantially, represented in the virtual world. But, mapping the physical artefact to its virtual representation is still in large a manual (human) job.
- The border between the physical and the virtual world has many dimensions of which the differences between building blocks (atoms and bits) is only one. What are the other dimensions and can the gap between the worlds be reduced along some of the dimensions?
- The physical and the virtual worlds will continue to co-exist.
- Many human activities in the physical and virtual worlds seem to concern similar events, address similar problems, and are all more or less part of our wish to improve the order in the same multidimensional chaos.

**Why we need a different interaction model for physical-virtual environments**

**Physical objects are not just user interfaces**

Until recently, physical objects and the physical environment in general has been phenomena on the periphery or outside the scope for most people in the HCI community. At least those who have been specialising on interaction issues have often not included physical aspects in their design models except as being the outermost layer of an otherwise virtual “user interface”. While being natural and sufficient when the number of simultaneously used physical interaction devices is small, physical aspects become increasingly important as the number of interactive computational devices increase. More importantly, the physical interactive devices tend to become increasingly complex and start to embody activity-relevant states on their own, independent of the virtual objects they might have been designed to control in the first place. In short, they go from being part of a user interface towards the virtual world (or some part of it) to being objects useful (or fun) to interact with on their own. A good example is
the cellular phone which originally was a plain user interface component towards a telephone network that enabled sound waves to be transmitted from one location to another. Nowadays, making phone calls is only one of many functionalities offered by a cellular phone. The fact is, and this is the point, that the modern cellular phone can be useful even if it is not connected to the system it was once designed to be the user interface to: the phone network. The alarm clock function, the calendar, and the calculator are all examples of functional augmentations that has pushed the device to make the transition from belonging to the somewhat primitive category of user interface devices to the (higher-regarded?) category of independent objects.

**Physical objects are increasingly virtually observable**

As the state of physical objects become increasingly observable by computer systems, it is tempting to include existing previously “dumb” physical objects as “user interface components” towards virtual applications. This is the approach taken within the area of Augmented Reality. The affordances of the more physical interaction with the system is undoubtedly often beneficial and gives the human agent better control, fuller experience, etc. However, and this is the point I would like to make here, existing physical objects often embody an activity-relevant meaning which is changed when the objects are manipulated as part of a user interface. This change in meaning cannot be as easily cancelled, reversed, or ignored in the way manipulation of virtual user interface components can. Thus, unless the physical object is deliberately designed for being a user interface component only (such as in much of the work of Ishii and Ulmer) but has existed in the physical environment already (such as in the work of Wendy Mackay), it is misleading to view the object as being part of a user interface since the changed state of the physical object is often part of the intended result of the activity itself.

**Blurred border means blurred “user interface”**

To summarise my criticism against seeing physical objects as components of user interfaces towards virtual applications rather than objects on their own, I believe that the trend towards an increasing number of independently useful physical computationally enhanced devices, and the trend to include existing physical environments as parts of a physical-virtual system, demands an alternative way of modelling the objects we manipulate. The concept of “user interface” is suitable at the border between the physical and the virtual worlds only as long as the tool used for manipulation is in the opposite world of the object which is manipulated, e.g. the physical mouse and a virtual object dragged over the virtual screen area by keeping the mouse button down and moving it. In the physical-virtual environments I foresee, this separation is blurred since the state change that the manipulation tool is subject to, is an activity-relevant state change. This is not the case with for instance the mouse or the keyboard whose physical state changes are irrelevant as soon as the physical-virtual operation is completely done. The alternative view suggested in this thesis (discussed at length in chap-
ter 5), is to regard physical and virtual objects that are strongly related to each other as one single object, having physical and virtual manifestations that are partially dependent of each other.

8.2 General world-neutral spatial definitions

Before presenting the proposed situative and objective models of physical-virtual space in detail, important concepts such as “space”, “object”, and “containment” will be defined in a more abstract sense. These abstract definitions are intended to be intuitively valid in both the physical and the virtual world.

**DEFINITION 8.2.1:** A space (S) is an abstract entity defined as a finite set of objects and specific spatial relationships between these objects.

The spatial relationships between objects in a space can typically be quantitatively measured in units such as for instance meters (resulting in a geometrical space) or seconds (resulting in a temporal space). As will be discussed at length later in this chapter, virtual spaces are in general defined by topological object relationships of higher order than three. Virtual spaces will pragmatically be considered as topological spaces to highlight the greater flexibility of such spaces compared to geometrical ones although the geometrical space of the physical world certainly also can be viewed as having topology.

Complementary, space will also be viewed as an object attribute. Objects can possess spaces, and give access to (sub-)spaces (finite and intrinsically structured sets of (sub-)objects).

**DEFINITION 8.2.2:** An object is a physical or virtual entity that among other attributes possesses a space, and encloses it by boundaries technically observable by an agent.

The space boundaries afforded by objects make it possible for agents interacting with a specific object to understand the extension of the object, to distinguish one space from another, and to decide whether a specific object is situated within or outside a specific object’s space.

For stylistic reasons, the relationship between objects and “sub-objects” are made explicit:

---

44. Spaces of higher dimensionality than three are sometimes referred to as “hyperspaces” in mathematical and hypermedia literature.
45. Please see the discussion in relation to the definition of the virtual world (definition 4.1.2, page 76) for an informal description of how the term "topology" is used in this thesis.
46. This definition of “object” complements the previous informal definition given in chapter 4, page 79.
DEFINITION 8.2.3: An object B is contained by an object A if B is situated in a space possessed by A, denoted $S_A$.

Space boundaries do in general simplify the observable part of the world for an agent situated in the space by conceptually, perceptionally, and causally defining the spatial limits of the propagation of causal effects. Often, results of actions or events stop to have meaning or effect exactly at the boundary of spaces.

8.3 A situative model of physical-virtual space

Are there patterns in human behaviour with regard to where they put things, for how long, and why? Is it possible to search for these patterns in the physical and the virtual world using the same analysis tools? Do the physical and virtual “object use pattern” look the same or are they different? While not attempting to find final answers to these questions, this section will present a model that aims in this direction.

Object categories and situative space organisation

The (re-)organisation of objects is an important part of Knowledge Work activities (e.g. Malone, 1983; Kirsh, 1995; Sellen & Harper, 2002). Different objects play different specific roles in organisation-related actions. As a cognitive tool for understanding the roles of objects, it is helpful for designers of physical-virtual environments to categorise objects as domain objects, tools, agents, and containers in a given activity context. This ontological choice of categorisation has been inspired by Beaudouin-Lafon’s Post-WIMP UI interaction model (2000) based on the interplay between “domain objects” and “interaction instruments”, as well as the categorisation made by Holmquist, Redström and Ljungstrand (1999) to describe relations between objects involved when linking digital information to physical objects. The latter authors use an ontology based on “tokens”, “tools”, and “containers”. As will be pointed out in the following discussion, the proposed object categorisation is purpose-related in the sense that objects can play different roles at different times and in different activity context. Sometimes an object is best categorised as a tool, sometimes as a domain object.

DEFINITION 8.3.1: A domain object is a) an object that changes state when manipulated by agents performing an activity and whose change of state is considered to be (part of) the result of performing the activity itself, or b) an object that has the potential to act as abstract or concrete construction material for the development of new or revising existing domain objects during a specific activity.
Domain objects of type a) are characterised by a relatively easily-changed state within a large state space. Examples: A broken car in a car repair shop, an almost finished e-mail to a friend, or a chessboard with its chessmen in their carefully chosen positions during a game of chess. Examples of domain objects of type b) are nuts and bolts in a car repair shop; books, paper notes, and e-mails in an office environment.

**DEFINITION 8.3.2:** A *tool* is an object used by agents in order to facilitate the manipulation of one or several domain objects while performing an activity.

Tools range from being designed for manipulation of very particular domain objects in very specific tasks, to being designed as very general-purpose tools. Examples of specialised tools could be a course schedule for a particular recreational diving course set a particular week in a particular year, a stapler for a maximum of 20 sheets of paper, a left hooked ice hockey stick for children, or a software script for converting capital letters to lower-case in files conforming to the not well spread file standard “flkwqd”. Examples of more general tools would be a dollar bill, a stove, a hammer, a PC operating system, or the widespread convention of ordering letters in a particular sequence, called the alphabetical order. Of course, the generality and speciality of a particular tool depends heavily on the use context and the intention of the agent using it (Bробег, 2000).

**DEFINITION 8.3.3:** An *agent* is an object able to autonomously initiate and perform complex activities by manipulating domain objects with and/or without tools in order to achieve high-level goals.

Examples: humans, animals, artificially implemented complex and autonomous systems.

**DEFINITION 8.3.4:** A *container* is an object whose main purpose is to provide and define a space for object storage and scope of domain object manipulation, to make it possible to apply “macro”-functions on contained objects in one single operation (e.g. translation), or just to help an agent to separate objects and object collections from each other according to some scheme.

Containers generally provide support (activity context, manipulation affordances and constraints etc.) for agent activities. It is reasonable to believe that containers play an important structural role in human activity by for instance facilitating decision making enforcing physical or virtual constraints on possible actions.
The situativity of object roles

It is important to note that the same physical or virtual object can be viewed as for instance a domain object in one situation and as a container in another, depending on the type of activity, the interest of the human agent performing the activity, and on the interest of the analyst/designer. E.g. an office room might be viewed as a container when a person is performing a knowledge work activity in it and as a domain object in the context of constructing a building.

Hot, warm and cold domain objects and containers — defining object use ecology

The proposed framework tries to capture two important activity-supporting functions of containers: 1) to provide more or less structured space for long-term storage of “cold” objects, and 2) to provide a) space, and b) a selection of tools for manipulation of “hot” domain objects. The notion of “hot” and “cold” objects is borrowed from the empirical work of Sellen & Harper (2002) who use the terms to distinguish between physical objects assumed to be relevant for a currently ongoing activity, and those not. Although some container objects frequently play both roles, some are more tuned towards the storage function (“cold” containers) and others towards supporting the intra\textsuperscript{47}-manipulation of domain objects (“hot” containers). Whenever a distinction is needed, they are referred to as storage containers and workshop containers respectively.

Also other container types are imaginable. For instance complex physical artefacts like cellular phones could be seen as containers of the electronics inside, i.e. a kind of highly structurally constrained compositional container. Physical state change of this kind of container would imply rebuilding it according to very strict rules if it should still remain the same kind of object. In order to keep the model simple, such objects that in most settings are not treated as containers by the human agents, will be modelled as tools or domain objects depending on activity context.

Table 9 shows an example categorisation of some common everyday objects according to the previously defined object categories.

A situative physical-virtual space model

At any given point in time, a specific human agent is able to (visually) observe only parts of the physical and the virtual world. The majority of objects and space-defining object relationships remain hidden for reasons that will be discussed later on in this chapter. Furthermore, only parts of these observable “sub-worlds” contain objects that can be manipulated. Thus, at any given situation, and centred around a given human agent, three inclusion-related spaces can be distinguished in the physical and the virtual world respectively (assuming that at least one virtual environment provider is present in the physical space in which the agent is situated at the specific point in

\textsuperscript{47}. The concept of intra-manipulation is defined in the next section of this chapter.
time). By conceptually joining the physical and virtual counterparts of the three situative spaces, a simple physical-virtual (PV) space model can be constructed (figure 45).

**TABLE 9:** An example categorisation of objects in the physical and virtual worlds (Pederson, 2003).

<table>
<thead>
<tr>
<th>Domain Objects</th>
<th>Physical</th>
<th>Virtual</th>
</tr>
</thead>
<tbody>
<tr>
<td>tools</td>
<td>a screwdriver when mending a car</td>
<td>a clipboard</td>
</tr>
<tr>
<td>containers</td>
<td>a desktop on which you can find pens, a stapler, etc.</td>
<td>a word processor application window</td>
</tr>
<tr>
<td>storage</td>
<td>a refrigerator</td>
<td>a folder in a file hierarchy</td>
</tr>
<tr>
<td>agents</td>
<td>a human</td>
<td>a reminder-application</td>
</tr>
</tbody>
</table>

**FIGURE 45.** Relationships between the three spaces in the proposed situative physical-virtual (PV) space model: A small subspace of the physical-virtual world space is observable, and a part of that observable subspace is also manipulable, by a human agent at a given point in time.

Exactly how the observable physical and virtual subspaces appear together in a given situation, forming the physical-virtual (PV) observable and manipulable subspaces, will be discussed later in this chapter. For now it is enough to regard the physical and virtual observable subspaces as being always simultaneously exposed in parallel to the human agent.
Somewhat more formal, the two subspaces in the model can be defined as follows:

**DEFINITION 8.3.5:** The *observable subspace* \((S_{\text{obs}}^{PV})\), or *situative environment*, is the subspace of the world space \((S_{\text{world}}^{PV})\) which is perceivable with no or very small navigation effort from the side of a specific human agent acting in the world space at a given point in time. \(S_{\text{obs}}^{PV} \subseteq S_{\text{world}}^{PV}\)

In the physical world, examples of very small navigation effort would be to move one’s gaze or to turn one’s head. In particular, objects contained by other objects such as pens inside a closed drawer, food inside a closed refrigerator, or electronic components inside some electronic device are normally *not* part of the observable subspace most of the time of the activity. In the WIMP virtual world, examples of very small navigation effort would be to translate a partly obscured window from “behind” to the front of the 2.5-dimensional space offered by the virtual environment provider. Window objects that are *fully* obscured by other objects are considered to be outside of the observable virtual subspace.

**DEFINITION 8.3.6:** The *manipulable subspace* \((S_{\text{manip}}^{PV})\) is the subspace of the observable subspace \((S_{\text{obs}}^{PV})\) which is manipulable by a specific human agent with very small navigation effort by a human agent acting in the world space at a given point in time. \(S_{\text{manip}}^{PV} \subseteq S_{\text{obs}}^{PV}\)

The observable subspace is a proper subspace of the PV world space (see definition 8.3.5) because the world space can never be completely observable at a single point in time by a single agent. The major part of the world space is (except for divine agents) not part of the observable subspace. On the contrary, the manipulable subspace can at times coincide with the observable subspace and is thus not a proper subspace of the observable space, as indicated in definition 8.3.6.

In order to keep the model simple, the special case of non-observable yet manipulable objects (such as the wallet in your back pocket) has not been included. Compared to the vast number of observable-manipulable objects, there are few such object in most situations, and they will be treated as if they were observable in the proposed framework. In some cases, like the wallet in the back pocket mentioned above, the object might perhaps not be *visually* perceivable, but perceivable through other senses (e.g. touch or sound), and it is therefore reasonable to include the object in the observable subspace anyway. Although the use of the term observable subspace will have a certain emphasis on visual perception in the remaining part of the thesis (in part caused by the heavy emphasis on this modality in most computer user interfaces available today) I would like to stress that the observable subspace should be seen as the set of objects and relations among them which can be perceived through any human sense.
The complete relationship between the two PV subspaces and the PV world space can be formally described as:

\[
S_{\text{manip}}^{PV} \subseteq S_{\text{obs}}^{PV} \subseteq S_{\text{world}}^{PV}
\]

In the physical world, the natural point of departure for defining the manipulable subspace is the geometrical extension of the limbs of the human agent, relative the location of the objects in the physical environment. In the virtual world, differences between the observable and manipulable subspaces are more often based on authorisation rules, and access restrictions, as to who should be allowed to manipulate specific objects based on political or other socio-organisational reasons.

**Geometry and topology**

The observable PV subspace (and thus also the manipulable PV subspace) is a geometrical space. This means that the space-defining relationships between objects in this subspace is based on two- or three-dimensional geometry (see definition 8.2.1 earlier). As can be noted from the previous chapter (see table 8, page 161), several of the general similarities between the physical and the virtual world stem from the fact that the observable parts of them both are represented in a two- or three-dimensional space.

The PV world space on the other hand is a combination of geometrical and topological spaces. The geometrical spaces are, quite naturally, the physical world space and the observable virtual subspaces. The topological space is the part of the virtual world space which is not, at the chosen time of analysis, observable by the human agent, i.e. hidden. This space will be denoted \( S_{\text{world-obs}}^{V} \) *hyperspace*, a term borrowed from hypermedia literature\(^48\) where it has probably originated from the kind of immaterial links, hyperlinks, that in general make up the relationships among virtual objects. As described in the previous chapter, the fundamental structural difference between non-observable parts of the physical and virtual world is a factor that is a basis for many world-specific characteristics of the two worlds. Thus, it might at first seem counter-intuitive to conceptually join these two structurally different spaces into one space. The motivation for doing this is once again pragmatic: it seems that human agents for better or worse, think on also hidden virtual (topological) spaces as having geometry in the course of everyday activities.

As indicated in figure 45, I will for stylistic reasons sometimes refer to specific observable PV subspaces \( S_{\text{obs}}^{PV} \) as *situative physical-virtual environments*.

---

\(^{48}\) Hyperspace is sometimes in hypermedia literature used to denote the whole virtual world as a whole, i.e. synonymous to “Cyberspace”. In this thesis the term is used specifically to denote the, at a specific point in time, unobservable part of the virtual world.
Inter-space dynamics

The logistic process in the course of activities can be regarded as similar in both worlds in the sense that human agents in general can be assumed to translate objects of interest (“hot” objects) from the previously unobservable part of the PV world space \( S_{world-obs}^{PV} \) into the observable PV subspace \( S_{obs}^{PV} \). If the object is only to be “kept handy” at a specific point in time, or support the activity by its presence only, it may remain in the non-manipulable part of the observable PV space \( S_{obs-manip}^{PV} \) during a specific time period. Whenever the object is to be involved in any kind of manipulation, it will be translated by the human agent to the manipulable subspace \( S_{manip}^{PV} \).

Objects that are no longer relevant for a given activity will eventually be translated back to the unobservable part of the PV world space \( S_{world-obs}^{PV} \) again when the agent turns to other activities. Using the previously introduced terminology for describing relevance of objects, the relationship between PV (sub)spaces and the objects that can be found within them at any given point in time can be summarised as follows:

- Hot objects are located in the manipulable PV subspace \( S_{manip}^{PV} \),
- warm objects in the non-manipulable part of the observable PV subspace \( S_{obs-manip}^{PV} \),
- cold objects are located in the non-observable part of the PV world space \( S_{world-obs}^{PV} \).

However, while the general “object translation pattern” looks similar for both the physical and the virtual world as to where “hot” and “cold” objects can be found, there are differences when it comes to the frequency and dynamics of how object translation appears in the two worlds. In the physical world, objects more or less come to, and go from, the observable subspace in large “chunks” as the human agent navigates (moves around geometrically) in the physical world space. This is particularly true in artificial environments such as buildings where for instance leaving one room for another serves as a good example. In the virtual world, objects are instead often brought into and brought out from the observation subspace on a one-at-a-time basis, although the process might perhaps still best be described as space navigation. The studying of what could be called “situative physical-virtual object logistics” as described above seems to be an insignificantly explored area largely open for research.

Table 10 summarises the characteristics of the three (sub)spaces in the physical and the virtual world.

---

49. Since the observable and manipulable virtual space often coincide, objects that are to be kept handy in the virtual space tend to end up in the manipulable space anyway.
The physical location of virtual environments

The location-dynamics of virtual environments within physical environments depend on the degree of (potential) mobility/wearability and the correlated “presence” of the virtual environment providers. Here are some common devices and to what extent they dynamically can geometrically place the virtual space within the physical, described using the terminology of the situative PV space model:

- **Desktop PC**: the virtual environment is fixed to the position of the display within the physical environment. This means that the virtual environment is not present in the observable V space unless the physical environment that contains it is present in the observable P space. This also means that the manipulable V subspace completely disappears whenever the physical manifestation of the virtual environment provider’s P→V IWEM (e.g. a mouse, or a keyboard) leaves the manipulable P subspace, e.g. when the human agent turns around 180 degrees.

- **Laptop/tablet PC**: the virtual environment can appear in many different physical locations, but normally not during local physical navigation activities.

- **Handheld PC (PDA)**: same as laptop/tablet PC

- **Wearable PC + see-through HMD**: The virtual environment is “omni-present”, or ubiquitous, and is physical-positionally independent from the physical world. A significant difference compared to the desktop/laptop/tablet PC setting is that the virtual world afforded by the device does never disappear from the observable PV space. The human agent cannot look away from the virtual world. Thus, the virtual world is attributed a “presence level” comparable to the physical world.

### TABLE 10:
How the three (sub)spaces are defined in the physical and the virtual (WIMP) world. The technically unperceivable (for human agents) hyperspace is highlighted with a frame.

<table>
<thead>
<tr>
<th>PV world space $S_{PV}$</th>
<th>contains</th>
<th>defined in the physical world as</th>
<th>defined in the virtual (WIMP) world as</th>
</tr>
</thead>
<tbody>
<tr>
<td>manipulable PV subspace $S_{manip}^{PV}$</td>
<td>immediately manipulable objects</td>
<td>the geometrical space immediately accessible by the body of the agent</td>
<td>the geometrical two-dimensional display area of a virtual environment provider</td>
</tr>
<tr>
<td>observable PV subspace $S_{obs}^{PV}$</td>
<td>immediately observable objects</td>
<td>the geometrical space immediately perceivable through the senses of the agent</td>
<td>The 2.5-dimensional geometrical space possessed by the “desktop”</td>
</tr>
<tr>
<td>PV world space $S_{world}^{PV}$</td>
<td>potentially observable objects</td>
<td>the geometrical space technically possible to perceive through the senses of the agent</td>
<td>hyperspace: the multi-dimensional topological space based on “links” between virtual objects</td>
</tr>
</tbody>
</table>

The physical location of virtual environments

The location-dynamics of virtual environments within physical environments depend on the degree of (potential) mobility/wearability and the correlated “presence” of the virtual environment providers. Here are some common devices and to what extent they dynamically can geometrically place the virtual space within the physical, described using the terminology of the situative PV space model:

- Desktop PC: the virtual environment is fixed to the position of the display within the physical environment. This means that the virtual environment is not present in the observable V space unless the physical environment that contains it is present in the observable P space. This also means that the manipulable V subspace completely disappears whenever the physical manifestation of the virtual environment provider’s P→V IWEM (e.g. a mouse, or a keyboard) leaves the manipulable P subspace, e.g. when the human agent turns around 180 degrees.

- Laptop/tablet PC: the virtual environment can appear in many different physical locations, but normally not during local physical navigation activities.

- Handheld PC (PDA): same as laptop/tablet PC

- Wearable PC + see-through HMD: The virtual environment is “omni-present”, or ubiquitous, and is physical-positionally independent from the physical world. A significant difference compared to the desktop/laptop/tablet PC setting is that the virtual world afforded by the device does never disappear from the observable PV space. The human agent cannot look away from the virtual world. Thus, the virtual world is attributed a “presence level” comparable to the physical world.
- Wearable PC + see-through HMD + P→V position synchronisation points: Augmented/Mixed Reality setting where parts of the virtual world is translated/cloned into the observable V space depending on the physical location of the human agent.

- Wearable PC + see-through HMD + PV activity infrastructure (position + object ID): Augmented/Mixed Reality setting where the world is translated/cloned into the observable V space depending on the physical location of the human agent, and on the physical and virtual objects currently manipulated by the human agent.

## Containment hierarchies

As will be shown in the next chapter, both physical and virtual environments can be modelled as hierarchies based on the objects situated in them and containment relationships (see definition 8.2.3 earlier in this chapter) between those objects. However, because the physical and virtual worlds differ in their structure (the former is based on geometrical object relationships, the latter is based on topological relationships), “containment” does not, and to some extent cannot, necessarily mean exactly the same thing in both worlds. The differences between physical and virtual containment will be discussed in more detail in the next chapter. For the current discussion it is enough to regard them as equivalent. Furthermore, physical constraints ensure a very regular hierarchical tree structure for the physical world while in the virtual world, cheap “cloning” of objects on the one hand, and independency from physical laws of nature such as only three spatial dimensions on the other, opens up for a more irregular structure in the virtual world where for instance the same object can appear at more than one place. Thus, virtual containment hierarchies belong to the relatively relaxed class of hierarchies called semi-lattices. (See for instance Hirtle (1995) for an explanation of semi-lattices).

### Intra- and extra-manipulation

The effects of manipulating an object is propagated upwards and downwards in containment hierarchies according to mechanisms that will be denoted intra- and extra-manipulation, concepts inspired by Herbert Simon’s discussion on what he calls “nearly decomposable complex systems” (Simon, 1996).

An extra-manipulation of an object on one hierarchical level is considered to be identical to an intra-manipulation of that object’s parent-object one level above. Further, a container object’s internal state is equal to the set of external and internal states of that object’s children who in turn depend on the states of their children and so on. Since practically all objects can act as containers (e.g. when interested in the spatial relationship between fibres in an apple, the apple can act as container) and thus nesting is unavoidable, one object has to be chosen to act as a reference object (RO) whenever an analysis is to be done, to avoid confusion. Extra-manipulation of a RO changes the relationship between the RO and its surroundings (technically, its sibling and par-
ent objects in the containment hierarchy). An intra-manipulation of a RO changes the internal structure of the RO itself (technically, the relationship between RO’s children as well as the relationship between them and RO).

“Intracomponent linkages are generally stronger than intercomponent linkages. This fact has the effect of separating the high-frequency dynamics of a hierarchy — involving the internal structure of the components — from the low-frequency dynamics — involving interaction among components.”

(Herbert Simon, 1996, p204.)

**Short-term and long-term manipulation**

While it is sometimes hard in the physical world to clearly distinguish between short-term and long-term lasting object manipulations, such a difference is more evident in the virtual environment offered by the WIMP paradigm: Extra-manipulation of containers of the type “window” (see table 11) do seldom last for long and are seldom part of the result of the activity but instead motivated by the temporary management of screen real-estate. Such manipulations will be referred to as short-term manipulations. A physical example of short term extra-manipulation would be the placing of plates and cutlery on a dinner table, an organisation that lasts for a short period of time and is governed by spatial constraints more than other constraints. The passing of the salad bowl during dinner is perhaps an even better example: the position of the salad bowl at any given point in time is not (under normal dinner and modelling circumstances) associated with any particular meaning — it ends up where it is spatially, for the moment, most convenient.

**Linking the object categories to long- and short-term manipulation and the PV situative space model**

From the perspective of topological relationships between domain objects, workshop containers are dynamic “hot” places. A significant amount of manipulation and navigation in the spaces provided by workshop objects are of a temporary character whose effects are only useful and even valid during a short time-span. E.g. window-translations operations in the virtual “desktop” workshop, menu navigation within subspaces spanned by complex tools such as word processors, or a paper translation within a physical desktop workshop. Sometimes the effects of these “short-term” operations on domain objects can be made permanent by successive translation or cloning operations destined for some storage container in a colder area of the world space. These colder containers are more static and “long-term” with regard to the frequency of state changes appearing in the container space.

**Conditions for object manipulation in the physical and the virtual world**

Table 11 illustrates the differences between physical and virtual objects with regard to how they afford short- and long-term intra- and extra-manipulation operations. For reasons of space, only domain and workshop objects (see table 9) are included. Fur-
thermore, the virtual environment modelled in table 11 is the one presented by the WIMP interaction paradigm. Virtual environments offered by other interaction paradigms would look different. As discussed in the previous chapter, the WIMP paradigm is in general able to manifest virtual objects in three distinct forms which complicates the modelling since each object form allows for different manipulation opportunities, in contrast to the physical world whose objects rarely take on dramatically different shapes.

8.4 Summary

This chapter has presented a set of interrelated theoretical concepts, all aimed at describing physical and virtual space using a common terminology based on the assumption that simple human object-centred activity by and large take on similar shapes in both the physical and the virtual world.

The situative physical-virtual space model tries to capture and describe “object use ecology” in both worlds at a very high level of abstraction, i.e. the different roles objects play (e.g. domain objects, tools, and containers), and to some extent where they are situated relative to the human agent during activities in both worlds (by distinguishing between the PV world space, PV observable space, and PV manipulable space). This model also describes important properties of how manipulation of objects affect objects in the topological vicinity (the concepts of intra- and extra-manipulation) and to what extent different kinds of object manipulation have a long-term or short-term effect on the physical-virtual world space.

The design stance, the space definitions, and the situational model of physical-virtual space all serve as a platform for the development of the hierarchical models of physical-virtual environments developed in the next chapter.
<table>
<thead>
<tr>
<th>short-term</th>
<th>intra-manipulation</th>
<th>extra-manipulation</th>
<th>long-term</th>
<th>intra-manipulation</th>
<th>extra-manipulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Window RO</td>
<td>• DM-spatial translation of children objects(_W) (domain objects, tools) within the spatial boundaries of RO • to include an object as a child (“open”-menu item, “show tool x”-menu item)</td>
<td>• DM-spatial translation of RO within the spatial boundaries of the parent (the desktop) • DM-spatial resizing of RO • hide RO (“minimize”-button) • de-activate RO (“close”-button)</td>
<td>• to adjust workshop preferences settings such as simple/advanced menus; pics or points, etc.</td>
<td>• ?</td>
<td></td>
</tr>
<tr>
<td>Icon</td>
<td>• ?</td>
<td>• to select/deselect RO (by clicking (DM))</td>
<td>• to change file properties (pop-up menu)</td>
<td>• DM-cross-storage-container translation (or duplication) of RO • to rename RO</td>
<td>• ?</td>
</tr>
<tr>
<td>Virtual World (WIMP environment)</td>
<td>• DM-spatial translation of children objects (text etc.) within the spatial boundaries of RO (e.g. scrollbar, PgUp/PgDn, zoom)</td>
<td>• DM-spatial translation of RO within the spatial boundaries of the parent (the workshop object) • DM-spatial resizing of RO • transform RO from W to Wm (“minimize”-button) • de-activate RO (“close”-button)</td>
<td>• ?</td>
<td>• ?</td>
<td></td>
</tr>
<tr>
<td>Physical World</td>
<td>• DM-spatial translation of children objects within the spatial boundaries of the RO (e.g. to move a bottle of wine across the dining table) • to include an object as a RO child (e.g. to put a book on the desktop)</td>
<td>• DM-spatial translation of RO (e.g. to put your work bag (the RO) containing pens, papers and laptop on the seat beside you on the morning train)</td>
<td>• to repaint the walls of a living room (the RO)</td>
<td>• to move a desk (the RO) from one room to another</td>
<td>• ?</td>
</tr>
<tr>
<td>Any RO</td>
<td>• DM-spatial translation of children objects (e.g. book pages) within the spatial boundaries of RO (e.g. to turn the pages) • to write things on a blackboard (the RO)</td>
<td>• DM-spatial translation of RO within the spatial boundaries of the parent (e.g. to move a pawn (the RO) forward in chess) • DM-spatial resizing of RO (e.g. to roll-up the blinds of a window)</td>
<td>• to overline lines of text in a text document using a highlighter pen</td>
<td>• DM-translation of RO to a storage container (e.g. to move a book from a bag to a shelf) • DM-spatial resizing of RO (e.g. to crumple up a piece of paper)</td>
<td>• ?</td>
</tr>
</tbody>
</table>

**TABLE 11:** Examples of intra- and extra-manipulation afforded by physical and virtual objects. Legend: RO = Reference Object; DM = Direct Manipulation; object postfixes W, I, Wm = Window, Icon, and Minimized Window manifestation respectively. (Pederson, 2003)
A hierarchical model of physical-virtual space

This chapter presents a modelling framework, based on physical-virtual concepts presented in the previous chapter, intended to facilitate design and analysis of infrastructural support for human physical-virtual activities. The concept of everyday containment is analysed, formalised and used as a common denominator in order to describe structural properties of physical-virtual space. The presented physical-virtual model is intended to help designers construct more structurally integrated physical-virtual environments.

The fundamental differences between the physical and virtual worlds described in chapter 7 force any unifying modelling effort such as the one presented in this chapter into a series of challenging design trade-offs between the preservation of typical characteristics of one of the worlds and the loss of modelling power when describing the other. Some important such design decisions will be discussed in detail. The framework is centred around hierarchical modelling of the physical and virtual worlds. Hierarchies representing selected parts of the physical and the virtual world are constructed based on containment-relationships between objects.

While the previous chapter focused on human activity from the perspective of a specific human agent acting in a physical-virtual environment, i.e. adopting an human agent-centric view, this chapter will change viewpoint and study structural properties of physical and virtual environments at many times disregarding the impression they make on any specific human agent. The general space definitions from the previous chapter (page 170) are of course still valid but will in this chapter to some extent
be tuned and concretized differently for physical environments on the one hand, and virtual environments on the other. The aim is to provide a basis for “objective” description of physical and virtual environments and for constructing models of physical-virtual environments that represent physical and virtual objects and object relationships, i.e. spaces, in physical-virtually unified representations. The “objective” physical-virtual space model developed in this section relies on different containment definitions for physical and virtual environments, and will be presented in the following order:

- a hierarchical model of physical space
- a hierarchical model of virtual space
- a hierarchical model of physical virtual-space

Before developing and presenting these models however, an analysis of “containment” as it appears in everyday life, is undertaken.

### 9.1 Outset: Everyday physical containment

The strong correlation between geometry and how objects interplay between themselves and in interaction with human agents, makes geometrical boundaries of, and between, objects the probably single most important structural property of physical space. These boundaries can be modelled using the concept of containment. Containment as used in this thesis, is also tightly connected to the concept of “place” (Harrison & Dourish, 1996), helping human agents to structure their activities. I therefore argue that it is appropriate to use the concept of geometrical containment as a basis for developing models of physical space with the purpose of modelling human activity.

#### Properties of everyday physical containment

Many human activities in the physical world rely on functional and conceptual properties of containment. Apart from general aspects of physical containment, three other aspects of the relation will be discussed, namely conceptual aspects, interactional aspects, and object attribute aspects.

#### General aspects

In the physical world, the concept of containment is inevitably bound to the geometrical extensions of objects and their spatial position. Being constrained by laws of nature, the relation “physical containment” is:

- irreflexive - an object never contains itself (directly or indirectly)
- asymmetric - an object A that contains another object B is never contained (directly or indirectly) by object B
- transitive - if an object A contains an object B and object B contains an object C, then object A (indirectly) contains object C
- partial - if an object A does not contain an object B, this does not necessarily mean that object B (directly or indirectly) contains object A
Thus, physical (geometrical) containment is a relation belonging to the class of strict partial orders.

**Conceptual aspects**

Objects contained (directly or indirectly) by the same object(s) are usually (and especially in artificial environments) conceptually related to each other. One reason is that if a set of objects are often to be used at the same time, it is efficient to keep them in the same container (see below). Containment helps organising space so that it structurally reflects and supports human activity. A bookshelf is one example of an object whose containment role is primarily conceptual (semantic) (as will be explained later in this chapter, books on the shelves are considered to be “contained” by the shelves in the proposed modelling framework). In short, containment facilitates keeping order among related objects by making conceptually motivated distinctions between objects and object sets objectively perceivable. In a given activity context, if object A is not in the box together with objects B, C, and D, this typically means that 1) there is something about A which does not apply to B, C, and D; 2) B, C, and D have something in common that makes it meaningful (for a human agent) to distinguish them from A.

**Interactional aspects**

If an object A contains another object B, the effect on A of an external cause (e.g. some kind of extra-manipulation) is often propagated also to B, and any objects that B contains and so on. This is perhaps most evident in the case of object translation in which contained objects tend to be spatially “co-translated” together with their container(s), with respect to a point of reference external to them. Object translation is probably one out of several containment-pervasive operations that (dependent on the material properties of the container) is propagated to contained objects.\(^50\) The focus will be on translation propagation because of its ubiquity. The examination of the propagation of other containment-pervasive operations is left for future work.

Furthermore, containment defines conditions for object access by forcing human agents to traverse (with their whole body or with the part of the body intended to be used for manipulating the specific object) all objects that a) directly or indirectly contain\(^51\) the object that is to be accessed, and b) at the same time do not contain the human agent’s body (or the part intended to be used in the future manipulation operation) itself. This sequence of by necessity step-by-step traversable objects from a given point of reference will be called the “access-path” to an object, a term commonly used for describing a similar phenomena in virtual environments in contexts of file system navigation. Containment relationships between objects do also to a large degree influence how objects appear to human agents. The possibility of observing an

---

\(^{50}\) Other more complicated examples could be for instance deformation, the change of temperature, and exposure to fluids.

\(^{51}\) The concepts of direct and indirect containment will be defined later in this chapter.
object from a given vantage point depends largely on whether the object in question is contained by other objects, and is constrained by containment in a similar way as object access. To summarise, physical-world containment has effect on at least three important human agent ↔ object operations/relations:

- object observation/observability
- object access/accessibility
- object translation propagation

The terms “observation” and “access” are used here in a wide sense, as in “observable/manipulable with no or very small navigation effort”. See the definitions of the observable and manipulable physical-virtual subspaces (definition 8.3.5, page 175; definition 8.3.6, page 175) in the previous chapter for more details. In the following discussion, a distinction will be made between “opaque containment” and “solid containment”, special cases of containment that also are strict partial orders:

**DEFINITION 9.1.1:** An object \( A \) **opaquely contains** an object \( B \) iff \( A \) contains \( B \) and there exists no vantage point outside \( A \) from which \( B \) can be observed by a human agent.

Examples: Documents carried in a briefcase are typically opaquely contained by the briefcase. The breakfast cereals are opaquely contained in their package until you break the sealing. Objects such as transparent plastic bags offer extreme cases of non-opaque containment since objects contained by a specific transparent plastic bag are, by and large, observable from whatever point the bag itself is observable.

**DEFINITION 9.1.2:** An object \( A \) **solidly contains** an object \( B \) iff \( A \) contains \( B \) and there exists no point outside \( A \) from which \( B \) can be physically accessed by a human agent.

Some objects are containers that can be opened and closed with regard to observability and accessibility. For example, the pack of orange juice in the refrigerator is opaquely and solidly contained by the refrigerator while the refrigerator door is closed, but becomes both observable and accessible when the refrigerator (door) is opened.

Unfortunately, observability and accessibility as relations between points in space are both non-transitive: that point \( B \) is observable (accessible) from point \( A \), and point \( C \) is observable (accessible) from point \( B \) is no guarantee that \( C \) is observable (accessible) from \( A \). Objects are observable by a human agent only if in a line of vision of the agent. Only in the case of non-opaque containment and when the objects are oriented so that the “peepholes” line up is it possible to observe the object \( C \) contained in object \( B \) contained in object \( A \) from a point outside of \( A \).

The requirements for object accessibility are similar to those of object observation. Since the limbs used for accessing objects by human agents have a limited “bendability” and reach, the accessibility of object \( C \) (non-solidly) contained in object \( B \), which
is (non-solidly) contained in object A, from a point outside of A, can be approximated to depend on the object openings being lined up in a straight line within reach of the human agent. Given that A contains B and B contains C; if C is observable for a human agent from the outside of A, then necessarily B also is; and if C is physically accessible for a human agent from the outside of A, then necessarily B also is. In other words: you cannot “get at” C without “passing through” B.

Translation (movement) of an object is an important example of an operation that is generally propagated to the objects it contains. It is possible to move several objects at the same time, by moving an object that contains them. However, translating a contained object does not generally imply that objects that contain it are translated as well. Thus, translation propagation is not symmetric.

**Object attribute aspects**
Physical-world containment is also related to attributes of objects themselves, disregarding human activity. Two examples are geometrical size and weight. If object A contains object B then A (together with its content) is larger and heavier than B alone. (The weight of A is more exactly the sum of the weight of the container structure and weights of all directly contained objects.) While the cumulative property of contained object weight puts constraints on the building material of containers in the physical world, the relation between geometrical size and containment is sometimes useful in everyday physical-world search operations. We don’t look for the elephant in the sports car because it cannot possibly fit. On the other hand, small objects fit “everywhere” and as it happens, these are the ones hardest to locate once lost.

Container form, or shape, is also an attribute that influences the use of containers in the physical world. For instance, while bookshelves are very good for containing books, they do a bad job in containing plain paper documents (if the purpose is to support the organisation of them). An exploration of structural properties of containers (both physical and virtual) have been omitted in the presented work, except for how they in general relate to the concept of containment. However, such investigations could of course be interesting and is left for future work.

**Notes**

*What you see is what you can access*
In many cases when an object is observable in the physical world, it is also accessible (manipulable), and vice versa. One condition is that the objects are reachable with little or no physical navigation effort from the position of the specific human agent. In other words, it is more likely to be true in geometrically smaller environments such as a room or a building rather than in open space.
What makes containment containment?

An interesting question is whether we can accept losing a few of the containment properties mentioned in the previous section and still treat a specific relation between objects as containment, or at least something that strongly reminds of it. If so, what properties are necessary and what can be done without? The question is not only interesting from a philosophical standpoint but has in fact influenced how containment in the virtual world is defined later in this chapter. Without having done any empirical investigation of people’s conception of containment, I argue that it is still possible to say something on this matter on an intuitive basis. In the following discussion it should be noted that whether a property can be ignored or whether it is vital for the notion of containment is to some extent context dependent. However, among the properties that intuitively seem characteristic and therefore important (in a general context), I find:

- general containment irreflexivity - that an object does not contain itself,
- support for object translation propagation\(^{52}\) - that a contained object moves in space if the object that contains it does so,
- the existence of, and the need to traverse, an access-path - that a contained object in some way can and has to be accessed through the object(s) that contains it, and
- the asymmetry of geometrical size between containing and contained objects - that a geometrically smaller object does not contain a larger one.

It is reasonable to believe that in case one or more of these properties do not hold for a specific relation, it will (in general) be hard to conceive the relation as “containment”.

Exceptions

There are exceptions to the general influence of physical containment on human operations, and to the relationship between objects, as it has been presented above. Exceptions caused by transparent objects, fluids, objects made of elastic materials, artificial forces that cancel out gravity, electrically controlled actuators, etc. However, I do believe that the discussion correctly describes important relationships between most kinds of everyday objects of interest for this thesis; the way most objects respond to basic human operations; and the way most of them present themselves to human agents. Because my modelling goal is rather general and pragmatic, such exceptions can, at least for now, be ignored and treated as exceptions that prove the rule.

\(^{52}\) As mentioned earlier, there are probably also other kinds of manipulation that might under certain conditions be propagated down the “containment-chain” as well. For reasons of simplicity, but also because it is probably one of the most generally present “containment-pervasive operations”, I focus on translation.
9.2 A hierarchical model of physical space

Defining physical containment

For the purpose of creating hierarchical models based on containment, the exact meaning of physical containment within the modelling framework has to be defined. Physical space is used as the starting point. Based on the abstract definition of space presented earlier in the previous chapter (definition 8.2.1, page 170) physical space can be defined as follows:

**DEFINITION 9.2.1:** A physical space is a space made up of a finite set of physical objects related to each other geometrically in three dimensions.

Based on physical objects’ geometrical extension in space, and the spatial relationship between objects, a hierarchical model of a chosen part of the physical world (an “environment”) can be constructed.

In the proposed model of physical space, an object is said to be another object’s “child” in a hierarchical tree structure if it is spatially contained by that object in the modelled physical space. However, if “containment” is viewed in a strict mathematical sense, most physical environments would be modelled as very shallow hierarchical trees since far from all objects can or do spatially encompass other objects in all three dimensions. In order to be able to perform more detailed studies of physical environments, more relaxed constraints for physical containment are adopted, allowing for containment in two spatial dimensions as a complement to full three-dimensional containment, under certain conditions:

**DEFINITION 9.2.2a:** An object X is said to be basically physically contained by an object Y iff any spatial translation of Y also automatically implies a corresponding spatial translation of X, and the major part of X is spatially circumferenced by the space of Y in a) all three dimensions, or b) two dimensions when X is held in close contact with Y by gravity or some other force so that within the plane of contact, the major part of X’s surface is circumferenced by Y’s surface.

The first clause in the definition assures that the relationship persists also after translation operations of the containing object. The extension of the term “containment” in section b has important implications for how the relationships between objects, and between objects and their constituents, are modelled. A paper on a desktop is modelled as being contained by the desktop it resides on, a bookshelf hanging on a wall is modelled as being contained by the wall, and a desk is contained by the floor it stands on (which in turn is contained by the room the floor is a part of), etc. Furthermore, buttons and knobs fitted on complex devices such as radio receivers, remote control units, cellular phones, and cars are also treated as being contained by the respective
artefacts they are fitted on. A special case are objects whose main purpose is to physically connect two other objects with each other such as cables, wires, bridges, and roads. These objects are modelled as contained by all objects they are connected to, and thus can appear at more than one place in the (thus not so strict) object hierarchy.

In order to open for nested containment relationships, a second definition is needed:

**DEFINITION 9.2.2b:** An object X is *physically contained* by an object Y iff X basically physically is contained by Y or there is an object Z that basically physically contains X and is physically contained by Y.

Figure 46 shows a partial hierarchical model of my physical office, based on physical containment as defined above.

**Direct physical containment**

Because of definition 9.2.2b, containment properties are inherited from parent to child in the hierarchy in a (theoretically) endless manner. E.g. a page in a book is considered to be contained by a room if the book is contained by a bookshelf contained by the same room. Sometimes it is valuable to distinguish between this general and potentially indirect definition of containment and the more direct relationship between a parent object and its children present at strictly one level below in the hierarchical tree structure.

**DEFINITION 9.2.3:** An object X is said to be *directly physically contained* by an object Y iff X is physically contained by Y, and iff there exists no object Z that physically contains X and at the same time is physically contained by Y.

In the example given above, the book is directly physically contained by the bookshelf but not by the room.

**Part-of relationships**

It is up to the modeller/designer to decide whether perceptually noticeable entities that are physically inseparable from the object they are a part of should be modelled as contained by the object or if the entity should be ignored in the specific model. Example: The walls of a room can either be modelled as contained by the room, or be ignored. In the latter case, two objects residing on different walls will occupy the same topological position in the hierarchical tree, i.e. directly under the “room object” while in the former case, each of the two objects will be hierarchically positioned one additional level down in the hierarchy, each below their own unique “wall object” (in turn situated below the “room object”). This design decision should of course be based on whether the potential object has a significant meaning in the context of the specific activity of interest, or if it is irrelevant to separate it from its parent object. In the case
FIGURE 46. An example of a containment-based hierarchy of a physical environment. (A part of my office.)
of the walls in a room, the designer has to ask herself/himself if it matters if the model captures the difference between objects contained by the front wall and the back wall, or if the less fine-grained model of “any wall” or “the room” will provide sufficient granularity for describing the location of things hanging on the walls.

**Structural constraints imposed by the parent on the children**

Attributes of objects in the physical world (for instance colour or spatial position) change state in accordance with constraints imposed on them by the physical environment in which they are situated. Some of these constraints are fundamental ones imposed by the physical world (gravity, one object in one place, objects take up space, objects have weight, etc.) and have been discussed in chapter 7. Other constraints are deliberately designed into artefacts in order to both limit and guide interaction into what is believed to be the most efficient, most fun, or most simple way of configuring the specific artefact. A third class of constraints are caused by ignorance in the design process, resulting in bad support for a specific activity because it has not been considered by the artefact designer. In the proposed model, the structural state-related constraints (apart from the fundamental ones mentioned earlier) imposed on a set of objects are modelled as an attribute of the affected objects’ parent object. Examples:

- **Filing cabinets**: the fact that paper documents in a specific filing cabinet only can be placed in a one-dimensional sequence is due to the constraints built into the filing cabinet object.

- **Buttons and knobs**: The volume knob of an amplifier is typically constrained to be either turned clockwise/counter-clockwise (in the case of a knob), or slid along an incision (in the case of a slider). In both cases the result is a reduction from three to one degree of freedom in the interaction, caused by (in my model) the amplifier that is considered to be the spatial frame of reference and provides the necessary force to constrain the interaction with the knob/slider.

- **Desktops**: While being an object that at a first look does not impose large structural constraints on its children, even this object to some extent restricts the structure of the relationships between its children: objects can be freely arranged in the horizontal plane but is constrained to adopt a heap structure in the vertical dimension.

**Implicitly defined objects**

Spatial areas or volumes in-between a set of artefacts can sometimes (deliberately or by accident) come to define perceptually distinct spaces. Examples: the ground between two buildings, the floor between two bookshelves, the area lit up by a spot light. In the proposed model, also these in a way indirectly defined but perceptually bounded spaces can be modelled as objects in the hierarchical tree structure. Naturally, to include this kind of implicit objects in a specific model is probably best done when their spatial boundaries are defined by objects that can be assumed to not move or
change their spatial extension frequently. Else, the indirectly defined object acquires very dynamic spatial boundaries, increasing the complexity of the model. The motive for including indirectly defined objects in the spatial model is that, just like deliberately designed artefacts, any kind of perceptually bounded space can be used as a conceptually connotated container for other objects in the course of a human activity.

**Choosing the outer boundaries of the modelled space — the root of the tree**

A simple and natural way of deciding the boundaries of the space to be modelled, assuming that it is human activity that is of interest, is to imagine (or better still, empirically establish) the set of objects that perceptually shut the human agent(s) off from the rest of the physical world. These objects jointly define the border of the world that the human agent(s) can perceive, act upon, and react to. In artificially constructed environments such as buildings, this process is a trivial one, usually resulting in a bounding object set consisting of four walls, a floor and a ceiling. Eventually the root of the physical object tree becomes the “room” object.

**Perception loopholes**

However, rooms tend to have transparent windows through which “outside” events can be perceived visually, and half-open doors through which “outside” events can be perceived aurally, potentially making the modelling task somewhat more complex. These situations could be seen as special cases of non-opaque containment (see the definition of opaque containment; definition 9.1.2, page 186) where the human agent is the non-opaquely contained object “looking out” from the container, observing parts of the space surrounding the room or whatever the object is that directly contains her or him. In order to keep the physical spatial model simple and intuitively viable I suggest to treat objects such as these, that leak information about events to and from the otherwise well-defined space limited by the set of bounding objects, as objects contained by the previously defined set of bounding objects. Thus, for instance, a window on a wall is modelled as an object contained by the wall in which it is situated and everything perceivable which in reality goes on the other side of the window is treated as phenomena taking place **inside** the window, i.e contained by it. It is a dramatic simplification but a reasonable one as long as our main concern is phenomena inside the walls of a room (or whatever geometrical boundary found most suitable) rather than phenomena everywhere in the physical world, and if we keep to a human-agent-centric view.

**IT devices (P→V containment)**

Also IT devices such as telephones, TV, radio and PCs expand the observable subspace for human agents enabling varying levels and kinds of tele-presence. These should be treated in a similar way as the window and the open door above. The physical embod-
iment of these artefacts (the housing, the buttons and knobs, as well as display units fitted onto them) are modelled just like the constituents of any other physical object. The content which is made observable on any display object however, is viewed as phenomena in the virtual domain and should not be included in the object hierarchy at this stage. The modelling of virtual environments will be discussed in the next section, and the integration between physical and virtual object hierarchies is the subject of the section after that.

9.3 A hierarchical model of virtual space

Since the goal is to create containment-based hierarchical models also of virtual environments, containment in the virtual world has to be defined as well.

Searching for virtual containment

Containment in the virtual world is as intangible as the virtual world itself. Thus one might argue that there is no such thing as virtual containment. While this might be true from a low-level conceptual point of view, it is obviously not the case at the high abstraction level where most human agents interact with computers through WIMP-based interaction paradigms. The WIMP paradigm was in part designed to mimic the physical world, and containment (or parts of the properties of this relation at least) was one of the physical-world concepts that was imported to the “new world” although it is often described in other terms.

“Drag-n-drop”

For instance, popular WIMP operations such as “drag-n-drop” are not possible to support without an underlying model that incorporates a) the concept of co-translation (translation propagation) of objects that are contained by each other (for the “drag”-part of the operation), and b) the concept of geometrical containment (for the “drop”-part).

Hyperlinks

A containment-based structure that existed before the invention of the WIMP paradigm is that of the hierarchical file system. Data-files, applications and folders are linked to each other in the external memory (in terms of the von Neumann computer model) in a way so that they form a hierarchy. Until the WIMP-paradigm arrived, the “containment” lacked many properties of physical containment (e.g. the geometrical aspects) but was and is (this fact will be more thoroughly discussed later in this chapter) understood as containment by the human agents navigating the structure.

The hierarchical file system is an example of a structure based on the mechanism of hyperlinks. The hyperlink mechanism is here referred to as the general mechanism that makes a virtual object appear in the observable subspace as a result from intra-manipulation of an object already existing in the observable subspace. Examples apart
from the links within the file system hierarchy already mentioned are the links between menus and sub-menus in WIMP application menu bars, the links between icon- and window-representations of text files, the links between web pages. While the drag-n-drop kind of containment mimics the propagation of physical geometrical translation, the virtual hyperlink-mechanism implements the notion of physical object access. Virtual objects can be reached through other virtual objects almost in the way physical objects are accessed by first passing through boundaries of other physical objects that contain them in sequence. Human agents do not seldom refer to virtual objects hyperlinked to each other in terms of containment. Perhaps their conceptualisation of hyperlinks is based on physical containment as well?

There are probably other kinds of relations among virtual objects that could be viewed as containment but it seems that the two mechanisms briefly discussed above (drag-n-drop and hyperlink) are the most general, frequently occurring, and yet powerful containment-like virtual mechanisms. In order to keep the model simple, these two are considered enough for my modelling purposes.

**Defining virtual containment**

Since the two kinds of virtual containment complement each other and together to a certain extent fulfil the expectations on containment we have got from the physical world, the definition of virtual containment will be based on a combination of both the object co-translation relation (e.g. drag-n-drop) and access relation (hyperlink).

However, in order to define virtual containment, the space in which the containment takes place has to be defined first. In analogy with the way the concept of physical space was defined in the previous section, a definition of virtual space can be derived from the abstract space definition (definition 8.2.1, page 170) as follows:

**DEFINITION 9.3.1:** A *virtual space* is a space made up of a finite set of virtual objects related to each other topologically.\(^{53}\)

In order to define the concept of virtual containment, parts of the conceptual framework presented in the previous chapter will be used. Namely the notions of the observable subspace, hyperspace, and intra-manipulation.

**DEFINITION 9.3.2:** An object X is said to be *virtually contained* by an object Y iff a) X can be made to enter and/or disappear from the observable PV space as a result of manipulating Y or one of its virtually contained objects, or b) X will be correspondingly spatially translated (whether in the observable subspace or in hyperspace) as a result of any spatial translation of Y or one of its virtually contained objects.

\(^{53}\) See “The use of the term “topology” in this thesis” on page 76 for a short informal description of what is meant with the term topology in this thesis.
Whenever a distinction between the a)-type and the b)-type of containment in definition 9.3.2 has to be made, the first kind of virtual containment (definition 9.3.2 a) will be referred to as containment-by-access or containment\(_a\) and the second kind (definition 9.3.2 b) as co-translational containment or containment\(_{ct}\). The reason for incorporating them both in one single definition is that they both mimic (different) properties of physical containment and that they are both in general understood as "containment" by human agents.

Figure 47 shows a partial containment hierarchy of a virtual environment afforded by an instance of a Microsoft Windows 2000 operating system. Since the model in the figures are focused on a menu hierarchy (the "Task Bar") most of the containment-relationships pictured are of the type containment-by-access. But also cases of co-translational containment are present (e.g. the compact representation of Internet domain objects contained\(_{ct}\) by the application "Internet Explorer".

**FIGURE 47.** An example of an access-based (containment\(_a\)) partial containment hierarchy of a virtual environment. (A part of the environment offered by my PC running the Microsoft Windows 2000 OS.) The marked area is shown in more detail in figure 48.
FIGURE 48. A close-up of the partial virtual environment depicted in figure 47.
**Unrestricted transitivity across both kinds of containment**

The two occurrences of the phrases “its virtually contained children” in definition 9.3.2 does not make any distinction between containment and containment. This means that the two kinds of containment can appear in mixed sequences, i.e. there is nothing preventing an object A containing, an object B containing, and object C containing, and object D.

**Direct virtual containment**

If the scope of the containment definition is restricted to one hierarchical step only, in analogy with how direct containment was defined when modelling the physical world in definition 9.2.3, but basing it on the definition of virtual containment above, we get:

**DEFINITION 9.3.3:** An object X is said to be directly virtually contained by an object Y if X is virtually contained by Y, and if there exists no object Z that virtually contains X and at the same time is virtually contained by Y.

Note: The above definition implies that, contrary to physical objects, virtual objects can contain other objects without actually directly containing any object at all.

**Virtual containment as magic made normal**

The constraints put on physical containment by laws of nature are in general very hard to alter by design, should one like to do so. But not necessarily impossible from the point of view of human perception! Many magical tricks performed by professional conjurers are for instance exactly about creating the illusion of being able to override laws of nature, not least properties of everyday physical containment. “Where did that rabbit in the top hat come from?” However, phenomena such as these are (naturally) exceptional and not very common in everyday life. (If rabbits would normally jump out of empty hats, few people would pay to see it.)

Virtual containment instead, is a relation between objects that have to be shaped and controlled by software designers in order to appear as such. One might say that virtual object relations are “magical” as default, and remind only of well-known physical relations such as containment if they are deliberately designed to do so.

---

54. This is not the full story since virtual object relations might be (and probably is, in many cases) developed by software designers without using the concept of containment and still be conceptualised as such by a human agent exposed to it, simply because it is the most appropriate cognitive model of the phenomena, even if the phenomena fails to have all standard characteristics of everyday physical containment. The phenomena is simply understood as containment in absence of a better conceptual model.
There is more containment magic in hyperspace than in the observable space

Compliance with everyday physical containment is higher in the observable subspace than in the rest of the virtual space (hyperspace). The 2.5 dimensional observable space is forced to obey certain geometrical laws (for instance, an object just has to take up geometrical space, otherwise it cannot be in the observable space), and human agents have certain expectations on the objects in the space because objects do to some extent also have attributes typical for physical objects such as for instance shadows and depth. For instance, a file icon that has been dropped into a file browser window has to co-translate with that window, otherwise the illusion of containment is lost immediately. Constraints and expectations on object relations in hyperspace are weaker. Surprises and unexpected causality is more acceptable since hyperspace and its multi-dimensionality is already from scratch something structurally magical and hard to conceptually grasp.

Indirect containment magic seems easier to accept than direct

Object cannot contain themselves in the physical world because contain and geometrical containment is in general inseparable. This makes self-containment hard to accept for human agents in the virtual world as well. However, if the self-containment is intervened by other direct containment relations, it is more easy to conceptualise. Or perhaps more correctly, we do not need to conceptualise it because the “abnormality” does not show itself in its completeness but only in part. And even if we would reflect on why an object suddenly re-appears, we could explain the indirect self-containment phenomenon by imagining to have navigated in a circular path from one non-geometrical object to another and thus eventually arrive to the starting object. Since hyperspace itself does not offer any sense of direction (including the lack of notion of going in and out of objects), any path, except for a the minimal path consisting of only two objects (i.e. direct containment) can roughly be viewed as having any arbitrary shape including a circle that brings you back to where you came from. Alternatively, or complementary, our ability to accept indirect self-containment might also simply stem from limitations in human memory.

Cloned objects

One of the most “magical” phenomenon in the virtual world is the existence of “multi-contained” or “cloned” objects: objects having more than one direct parent. In the physical world, an object can sometimes seem to be at several different places at the same time, but it is not really true because the instances of the object are normally completely mutually independent with regards to the result of object intra-manipulation. 1000 copies of a text book will not change state when one line in one of them is underlined. Industrially multiplied objects such as the mentioned books are copies, not clones. The term “clone” is used to denote the case when a set of objects are in fact (ideally completely) mutually dependent. Watching a set of object clones of the one and same object should be regarded as viewing one single object but in different con-
texts. The possibility of allowing the view of the same object in several different contexts at the same time is one of the most powerful features of the virtual world and a feature that clearly distinguishes it from the physical world. There are more examples of cloning-like phenomena apart from the 1000 copies mentioned earlier. Mirrors and other artefacts that affect light can be used to give the illusion of watching several instances of exactly the same object. However, this “pseudo-cloning” is very limited since with all of the object instances except one, all you can do is watch. Virtual-world cloning typically poses no restrictions on the manipulability of the object clones.

**Cloned objects and the hierarchical model**

A quick look on the physical containment hierarchy in figure 46 and the virtual containment hierarchy in figure 47 shows that they both have a fairly regular structure with a slightly more chaotic touch in the virtual case. A closer look reveals that the irregularities in figure 47 are in fact caused by cloned objects. For instance, the “Control_Panel” object is contained by both the “My_Computer” object (not visible in figure 48 but present in figure 47) and the “Settings” object. The structural irregularity of the virtual containment hierarchy is in part caused by the mere plurality of parent objects, in part by the fact that the parent objects of a specific cloned object can reside at different levels in the containment hierarchy. Thus, as mentioned in the general section on containment hierarchies in the previous chapter (page 179) virtual containment hierarchies belong to the relatively relaxed hierarchy category of semi-lattices.

Since the virtual space does not limit the number of parents an object can have, i.e. how many places in which an object can be simultaneously situated, why not placing all objects in all places? Just like in the physical world, absence of objects can be as important as presence. It is safe to say that a well-designed containment hierarchy can efficiently guide human agents in their performing of activities both in the physical and the virtual world. It is probably true as well that unjustified object cloning in general has an inhibiting effect on this guidance by cluttering the environment with objects that obstruct access to more relevant ones and/or creates “semantic confusion” by delivering an unclear message to the human agent exactly what kind of activity the place is designed to support or is related to.

**Short-cuts**

One reason for cloning objects is the obvious fact that some objects are useful for more than one purpose or in more than one activity context. Sometimes cloning is the only alternative to “pure” object translation, for instance if a human agent would like to have access to an object in a context which the environment designer did not foresee. Some virtual environments allow the creation of “short-cuts” as a way to, at the same time 1) keep a standardised object hierarchy to simplify activities for virtual agents of
an operating system (e.g. all tools might have to reside in a certain “tool place” in the containment hierarchy) and 2) allow for customisation according to personal preferences of specific human agents.

**Ubiquitous objects**

Some objects are so general and/or frequently used that they are best kept in the observable PV space at all times. In the physical world, these objects are either made wearable (e.g. wristwatches) and thus always within the manipulable physical space, or distributed in numerous copies within an environment (e.g. chairs in an office building), thus always at least within the observable physical space. In the virtual world, such objects are made (by the environment designer or by the active human agent) directly contained by objects that are always in the manipulable virtual space (like the “task bar” container object in the interaction paradigm of Microsoft Windows 2000), or at least always in the observable virtual space (e.g. the “desktop” container object in the same interaction paradigm).

**Part-of relationships**

Just like for the physical containment hierarchy, it is up to the modeller/designer to decide whether perceptually noticeable entities that are inseparable from the object they are a part of (i.e. objects falling under the virtual containment type (b) in definition 9.3.2) should be modelled as contained by the object or if the entity should be ignored in the specific model. However, there is a certain group of objects that can be assumed to be left out in most modelling efforts, at least when modelling “high-level” human activities: objects acting as handles for controlling the short-term geometrical appearance of their parents within the virtual observable subspace. Examples of such handles in the Windows 2000 interaction paradigm are so called “active window borders”, “caption buttons”, and “title bars”. There are at least two general reasons for omitting such objects in the hierarchical containment model:

- they clutter the space model with too many irrelevant details that might lead the modeller to focus on the activity on a too low level of granularity
- they are substitutes for properties of objects in the physical world that are implicitly possessed by all physical objects, properties that are not modelled in the physical containment hierarchy, and therefore should not be modelled in the virtual containment hierarchy either

**Structural constraints imposed by the parent on the children**

Structural constraints can take many more forms in the virtual world than in the physical. While physical structural constraints mainly rely on laws of nature and how objects respond to those (e.g. an elephant does not fit in a suitcase), virtual structural constraints operate on a symbolical level and thus, are at least in theory much more flexible with regards to what object properties that are to govern the relationships be-
between contained objects. It is possible to distinguish between two classes of structural constraints in the virtual world, the class of constraint that guide interaction in the observable virtual subspace based on geometrical limitations, and the class of constraints caused by incompatibility among datatypes.

**Geometrical structural constraints (in the observable virtual subspace)**

These constraints affect the possibilities of performing operations that involve geometrical (two-dimensional) direct-manipulation operations on objects in the manipulable virtual subspace. The “dragging and dropping” of “icon” and “window” objects, “window bars”, and such cannot extend the geometrical area defined by the “desktop” workshop object. Furthermore, the fact that objects completely obscured by other objects in the 2.5-dimensional space afforded by the “desktop” workshop object cannot be involved in any direct-manipulation action is also regarded as a geometrical structural constraint enforced by the “desktop” workshop object.

**Datatype-based structural constraints**

As an effect of virtual structural flexibility, the structural constraints within different kinds of workshop containers in the virtual world are more diverse than structural constraints between physical workshop containers since the latter all have in common to be based on laws of nature. Most domain-specific virtual workshop containers (e.g. word processors, graphic design software, web browsers) in general only accept to contain the subset of domain objects that are low-level “compatible” with the workshop environment, i.e. domain objects whose internal structure conforms to the right domain object “type”. You cannot successfully inspect a text object in a music player workshop object. Text objects of a certain type might not be compatible with a certain word processor (workshop object). Such strong non-geometrical type restrictions on domain objects are not present in the physical world where it is for instance fully possible to put a cellular phone in such an incompatible workshop container as a washing machine. (The incompatibility strikes you later.) Another example of datatype-based structural constraints is the fact that the gravity in the 2.5-dimensional space offered by the “desktop” workshop object depends on whether objects are manifested as a “window” or “icon”. Only “icon” objects are affected by gravity and can be successfully translated into “window” container objects. Objects manifested as “windows” (both iconised and expanded) in general float mutually independent on top of each other in the depth dimension of the 2.5-dimensional “desktop” space.

**Physical-virtual artefacts (V→P containment)**

In the discussion of the physical containment hierarchy, IT devices were mentioned as physical objects containing not only physical sub-objects but virtual as well. Examples of virtual objects containing physical objects are somewhat harder to find. However, there are at least a few. For instance, the virtual (workshop-) object “CD Player” might directly virtually contain the physical object “Björk CD”. Since an intra-
nipulation of the CD Player (i.e. a click on the virtual eject button) can make the Björk CD enter the observable PV space (i.e. the CD is ejected from the physical CD drive). The virtual object “CD Player” and the physical CD drive do together in fact constitute a physical-virtual artefact (see definition 5.2.1, page 94) and it is likely that most occurrences of V→P containment arise in connection with such artefacts.

Implicitly defined objects

The observable virtual subspace is so dynamic that implicitly defined objects are not useful as placeholders in the same way as in the physical world (see “Implicitly defined objects” on page 192) where instead the gap between two objects can persist for longer time. Implicitly defined objects outside the observable virtual subspace are hard to imagine.

Choosing the outer boundaries of the modelled space — the root of the tree

While it is rather straight-forward to define the root of the physical containment hierarchy, defining the root of virtual containment hierarchy is a more open issue and depends on the context in which the modelling is performed. In the general case I suggest to choose for the root of the tree, the virtual object that acts as the main container for “hot” objects (Sellen & Harper, 2002) for the specific activity of interest.

9.4 A hierarchical model of physical-virtual space

Physical containment versus virtual containment

In defining physical and virtual containment above, the intention has been to rely on structural characteristics of the physical and virtual worlds that lie close to the conceptualisation of space navigation and manipulation created and maintained by human agents in everyday activities. A particular focus has been put on finding a common denominator through the concept of containment. The motive for defining the structural components of physical and virtual space in “down-to-earth” terms is that it allows for a more direct and intuitive modelling of human activity, compared to a more philosophical-physical definition perhaps more viable when searching for “the true” nature of physical and virtual space.

While I would argue that the definition of physical containment is very concrete and natural from the perspective of human everyday activity, the definition of virtual containment is more abstract because it is based on conceptual relationships between objects rather than geometrical. This slightly more abstract definition is necessary in order to topologically correctly model the structure of the virtual world. However, this does not mean that I subscribe to a view that human agents think about virtual containment in a significantly different way than physical containment. On the contrary, I believe that the everyday conceptual model of virtual space is very close to that of
Combining physical and virtual containment hierarchies

Although the previously presented physical and virtual hierarchies are based on different kinds of containment, there is nothing preventing us from joining them. (My definitions of both physical and virtual containment are grounded in what I believe to be the human everyday conception of containment, overriding any technical differences.) Figure 49 shows such a union of the physical and virtual containment hierarchies presented earlier in this chapter (figure 46 on page 191, and figures 47/48 on page 196 respectively). Figure 50 is a close-up of the section containing the inter-world container (a PC display unit) that connects the analysed physical and virtual environments to each other.

Examples of inter-world containment

Figures 49 and 50 illustrate the case where a virtual environment is directly physically contained (definition 9.2.3, page 190) by an object in a physical environment. The PC display contains the virtual “desktop” workshop object. Here are some other examples of inter-world containment:

- **Physical→virtual containment**: The physical button “calendar” on a Palm Pilot can directly virtually contain (definition 9.3.3, page 198) the virtual object “calendar” (a workshop object). If the button is pressed, the calendar workshop object enters the observable PV space.

- **Virtual→physical containment**: The virtual (workshop-) object “CD Player” might directly virtually contain (definition 9.3.3, page 198) the physical object “Björk CD”. Since a manipulation of the CD Player (i.e. a click on the virtual...
eject button) can make the Björk CD enter the observable PV space (i.e. the CD is ejected from the physical CD drive).

**Limitations**

The hierarchical model of physical-virtual space presented in this chapter should be regarded as a first step towards a rigorous and valid structural model. There are many aspects of the model that should be studied and developed in more detail in order to validate its usefulness as a tool for design.

As of now, the model has its main weaknesses in oversimplification of phenomena which, under certain conditions, probably can give a misleading picture of specific physical-virtual settings. This includes:

- The assumption that physical objects are geometrically “nice” in the sense that they possess one single continuous space.
- The assumption that the geometrical space possessed by the root object in the physical hierarchy is small enough to make all observable objects also manipulable.

**FIGURE 50.** A close-up of the physical-virtual containment hierarchy in figure 49. The thick dotted line represents the point where the physical direct containment relationship between the PC display unit and the “desktop” workshop object crosses the border between the physical and the virtual world.
(with small navigation effort) by the human agent. If this is not the case, the hierarchy does not give a correct picture of what objects that the specific human agent has (technically speaking) access to.

- The assumptions that contained objects in the physical world stick to the object that directly contains them in case their container “parent” is translated (or that they are not affected by the law of inertia, or that the frictional force is infinitely large inside containers, or whatever), and that they are completely detached from their container “parent” when they are translated themselves. If these assumptions cannot be taken for granted in a specific setting, the assumed transitivity of object translation is not valid.

9.5 Summary

Based on the concept of everyday containment, this chapter has presented an approach to the modelling of physical and virtual environments using hierarchies.

The proposed hierarchical structures that represent physical and virtual environments are intended to show important structural characteristics of spaces independently from a specific human agent and are in this sense “objective”. On the other hand, the construction of the hierarchies involves design decisions that depend on the kind of activity that is assumed to take place in the chosen physical-virtual environment, thus the hierarchies are in that sense not objective. Applying the model on any specific setting allows designer/analysts to view physical and virtual spaces involved in a human activity as one (large) hierarchy based on containment relationships between objects.

The modelling approach indicates that by choosing phenomena that exist and play important roles in both worlds (at least from the viewpoint of human perception) such as “containment”, structurally integrated models of physical-virtual space can be created.

The hope is that the model developed in this chapter can serve as a design tool for constructing environments and infrastructure aimed at supporting physical-virtual activities. The next chapter presents such a physical-virtual environment prototype that, although not entirely designed after the concepts presented in this chapter, has been inspired by, as well as inspired, the theoretical work previously discussed here and in several earlier chapters of this thesis.
This chapter adopts a different tone than previous chapters by describing the development and functionality of a prototype system.

The presentation of the system serves two major functions. First, it is intended to give a certain validity to the theoretical discussions in previous thesis chapters by demonstrating that it is indeed possible to implement systems that offer at least some support for bridging the physical-virtual environment gap, and that it is useful to model physical-virtual space using the concept of containment also in practice. Secondly, the presentation of the system gives a (modern) historical background to many of the theoretical constructs. Both related functions of this chapter brings the thesis down to earth, when it at points has admittedly been more up in the sky.

This chapter also serves as evidence (should one need any) that the research method of alternating between implementation and thought in an iterative manner is fruitful in the context of basic research. The physical-virtual framework presented earlier in this thesis on the one hand, and the Magic Touch prototype presented here on the other, have been developed literally in parallel. It is no exaggeration to say that one would not have existed without the other.
10.1 The system in short

Magic Touch is a system that tries to integrate a physical environment with a virtual ditto. Although the extent of the integration is actually very modest, it gives a hunch of future possibilities in doing a more extensive physical-virtual environment integration enabled by advances in sensing and actuation technology. The basic functions of the system is to track location changes of objects in an office environment and to display those changes on a computer display in real-time. Additionally, “added-value” mechanisms is offered to the human agent acting in the environment such as support for creating synchronised physical-virtual artefacts (see definition 5.2.3, page 96) and help in searching for objects in the physical environment.

A second phase of the system development explores the possibility of connecting several Magic Touch-fitted offices with each other through the Internet, allowing for (among other things) to share physical-virtual artefacts among several human agents and physical-virtual environments, and to “visit” physical offices from a distance. This second development phase is not concluded at the time of writing and is presented in brief and “as is”, with the intention to give the reader an idea of the possibilities and challenges lying ahead. The hope is also that the description of the transition from the first and the second development phase (Magic Touch 1.0 → Magic Touch 2.3) provides an insight into typical design decisions involved when designing physical-virtual environments in general and infrastructure for physical-virtual artefacts in particular. Figure 51 shows Magic Touch 1.0 in action.
10.2 Background

The motivation for developing Magic Touch originated as a response to the observation that human activity in knowledge work settings (Drucker, 1973; Malone, 1983; Kidd, 1994; Kwasnik, 1989; Mander, Salomon & Wong, 1992) has become an activity involving continuous and frequent alternation between performing tasks in the physical world on the one hand and the virtual on the other. Furthermore, although physical and virtual tasks (and objects involved in them) often are conceptually related, information about their states has to be manually communicated by human agents acting as messengers within something which can be seen as a physical-virtual system. Just like water between land makes us build bridges when the need to cross becomes

---

55. Statistically and qualitatively reliable empirical evidence for the existence of the gap, or rather, the extension of the gap was (and is) largely lacking. However, it was decided that the results from the study presented in chapter 6 taken together with my intuitive feeling that the gap a) existed and b) was frequently manually bridged by knowledge workers, was enough to motivate the development as a basic research activity.
frequent enough, the presence of the physical-virtual environment gap (see chapter 6) nourished the idea of creating some corresponding infrastructure for making these two worlds come closer as well, relieving human agents from at least parts of their role as “physical-virtual messengers”. The idea of physical-virtual artefacts (see chapter 5) became the conceptual bridge-pier on which much of the theoretical and practical gap-bridging activities were founded. The Magic Touch system presented in this chapter is no exception and the concept of physical-virtual artefacts shines through in many of its mechanisms.

10.3 Gearing down to the state of the art technology

As explained in chapter 6, the physical-virtual environment gap has many dimensions and it was early realised that building a system that bridges them all would, of course, be an impossible mission in the context of technology and knowledge available today.

Inspired by the importance of document organisation in knowledge work, identification and location tracking of paper documents soon became the focus.

Initial conceptual approach — ubiquitously fitted object identifiers

Initially, the idea was to fit object identification technology on places that had a clearly defined meaning in the context of knowledge work such as mail in- and outboxes, drawers in filing cabinets, waste baskets etc., and to let physical events (such as the appearance or disappearance of paper documents) at those places trigger events in the virtual world. The figures 52 and 53 show early sketches of such tailor-made physical-virtual “bridging points”. The object identification technology fitted onto these specific physical artefacts would provide a software system with information about the presence of a set of physical objects, and a very coarse model of where and in what state these physical objects were in, at a specific point in time.

The scaling problem

Although the ideas for tailor-made physical-virtual artefacts were interesting, it was obvious that the approach suffered from a scaling problem (fitting object identification technology onto physical objects in an office, not to mention a building, would be both costly and complex.

The problem of obtrusiveness

Of the identification technology surveyed at that time, there was no technology available that would not demand extra or significantly altered operations from the human agents acting in the environment if the necessary information about translation events should be obtained by the system.
Second thoughts — tracking individual objects instead?

One obvious solution is to let all objects in the environment carry position transmitters, whose signals are received by a motion tracker system. Such an approach gives both identification and position parameters accurately and continuously. Drawbacks are that the system is expensive if you want to track many objects, and the identification tags are at least to date fairly large. Another method is to put a camera in the ceiling and to attach visual tags, showing unique graphical patterns, to the objects. These patterns and the location of the tags are interpreted and calculated through analysis of the camera image (e.g. Rekimoto, 1996). In this case the tags are considerably cheaper compared to the other approach since it is possible to print them out on an ordinary
printer. Drawbacks include the necessity of free line-of-sight between the camera and the tagged objects and that the tags themselves can become fairly large if you want to identify many objects.

**Eureka! — Objects do not move by themselves!**
While searching for a suitable tracking method I discovered a fundamental fact: objects in office environments do not move by themselves! They move when they are moved by human agents’ hands. Put in another way, an object stays where it is until...
the agent grabs it in one or two hands, moves the hand(s) to a new location and drops the object. Based on this insight, I developed another object location tracking method which at least for my purposes proved to be more suitable than the other two mentioned tracking approaches.

As illustrated in Figure 54, the object location tracking system consists of (1) an office environment containing RF/ID-tagged artefacts, (2) wearable wireless tag readers, placed on each of the human agent’s hands\(^{56}\), identifying any tagged object the agent takes in her/his hand, and (3) a wireless location transmitter always aware of the positions of the agent’s hands.

The advantage of using RF/ID tags for my kind of application instead of other similar solutions like bar-codes is thoroughly discussed by Want, Fishkin, Gujar & Harrison (1999).

*The function that translates sensor data to “object translation” events*

As soon as the agent’s hand comes close to a tagged object, it is identified. If the human agent moves her/his hand and the reader still can read the tag it means that the agent has grabbed the tagged artefact. When the tag is no longer readable, the agent must have dropped the artefact, and the location of the agent’s hand at the point when the tag was last readable is regarded as the new location of the artefact.

*Limitations of the proposed location tracking approach*

With small and powerful enough readers and motion tracking technology I believe that this method delivers location changes of objects accurately and non-intrusively. However, the approach relies on several limiting assumptions. Here are the most evident (Pederson, 2001a):

- All objects that are to be tracked have to be tagged. (The proposed system shares this limitation with other similar systems however.)

---

\(^{56}\) So far, only one reader has been used at a time, meaning that only one of the hands is tracked. The general approach is however of course to track both hands, thus I say that two wearable units are used when describing the general object location tracking approach.
• No tagged object is allowed to change its own position by itself unless it is able to communicate the new position to the system by itself.
• The agent does not drop an artefact “in the air”, letting it fall down to its final location. If so, the system will store an incorrect height location.
• The agent moves only one artefact at a time. If many tagged artefacts are to be moved at once, some kind of “multi movement mode” has to be entered explicitly.
• The agent always moves artefacts directly with the hands and does not use any kind of tool to push or carry the artefacts “from a distance”. This would disable the identification mechanism since the tag reader has a limited reading range.
• Only agents that carry the wearable system equipment are allowed to move artefacts in the environment. If non-tracked agent hands are active in the environment the system’s artefact location database will become incorrect.
• The position of an object is based on the tag position, that is, one point in space. For small artefacts, it is a reasonable approximation, for larger objects the position approximation error will be more evident.

Some of the limitations mentioned above can be eliminated or at least alleviated by the introduction of additional mechanisms, sometimes also involving explicit human agent activity.

10.4 Development

From (working) chaos to components

The development process of the Magic Touch system has gradually changed from being purely exploratory and based on ad-hoc rapid prototyping performed by single individuals to a more controlled and systematic collaborative activity at points involving up to five software and hardware developers. The creation of use scenarios played an important role in the initial stages of the development. The main turning point in development strategy was the transition from a single-user, single-computer, single-environment solution (Magic Touch 1.0) to the client-server based distributed system model (Magic Touch 2.x). The more flexible client-server design demanded agreements on communication protocol standards as well as a more carefully thought-through separation of interests among the system components. As will be explained later in this chapter, the design of Magic Touch 2.x is also intended to enhance physical-virtual activity support and to fix ill-designed mechanisms present in Magic Touch 1.0 by including certain support for running the system, or parts of the system, on wearable computers.
The pool of developers

The people involved in the development apart from myself has mainly been undergraduate students at the department of computing science. Some system parts have been constructed by electrical engineering students and some parts by teaching colleagues at the department of computing science. I have not written a single line of code myself (!) but I have always been involved in important software and hardware design decisions as to what hardware to use, how it should be physically designed based on ergonomical issues, and the overall system architecture (high-level systems design, communication protocols, database design, user interface design).

System evolution year 2000-2002

Version 0.5 — Basic PVA manipulation tracking unit
MS Windows NT, MS Visual Basic 6.0, MS ActiveX.

Version 1.0 — Single-user, single-environment, desktop UI
MS Windows NT/2000, MS Visual Basic 6.0, MS Active X, MS Access 97.

Version 2.0-2.2 — Multi-user, multi-environment, wearable UI

Version 2.3 — Purified world model
Version 2.3 only exists as a conceptual design which differs significantly from earlier 2.x-versions only in the way PVA and PVA attributes are modelled. The change involves to model all kind of MT object types in the same way regardless of technology used for sensing PVA manipulation. All components of Magic Touch 2.x are not fully implemented at the time of writing.

10.5 Magic Touch 1.0-2.3 General description

General functionality of Magic Touch 1.0-2.3

Core functionalities
The system models physical space based on containment-relationships between objects and places

- The system tracks location changes of physical objects in real time
- The system maintains a link between monitored physical objects and virtual objects of choice (files), i.e. acts as infrastructure for user-designed PVAs
Added-value (end-user) functionality

The system provides added-value functionalities based on the PVA infrastructure such as:

- the possibility of locating PVA as if the corresponding PVA is identified (e.g. finding out where that annual report ended up three weeks ago),
- the possibility of automatic retrieval of corresponding PVA as if the PVA is indicated by the human agent (e.g. to view additional information about the coffee cup by putting it in the dedicated PVA inspection area of the desktop),
- and the possibility of visualising the monitored physical (office) environment as hierarchies or as (potentially filtered) three-dimensional visualisations, thus allowing to “browse” the physical space from alternative perspectives, and without leaving the computer workspace.

Magic Touch 2.x additional functionality

Magic Touch core system completely rewritten. Distributed Internet-based multi-user architecture not yet fully implemented.

- Allows human agents to monitor and interact with physical environments from a distance through the Internet.
- Unified hardware-independent object ontology: All objects have spatial extension, can be moved, and contain other objects.
- Magic Touch objects (and thus also spaces) can be shared among several users on different physical locations.
- Wearable system user interface allowing the user to control and command the PVA infrastructure without interaction with the PC running the local Magic Touch client.

General conceptual system architecture (Magic Touch 1.0-2.3)

Although the system was fundamentally revised in the transition between version 1.0 and 2.2, the core components, pictured in figure 55, have remained roughly the same:

- **The PVA Manipulation Tracker**: The PVA Manipulation Tracker captures information about PVA manipulation operations performed by a human agent, i.e. the manipulation of physical objects.

- **The PVA Database Management System (PVA-DBMS)**: The role of the PVA-DBMS is to keep PVA manifestations updated so that any manipulation of a specific PVA sensed by the PVA-Manipulation Tracker has effect on the corresponding PVA(s) and/or initiates some other pre-specified virtual event.

- **The PVA Configuration User Interface**: The PVA Configuration User Interface component allows human agents to define and inspect relationships between synchronised PVA and the PVA as well as any other logical functions and shared attributes bound to PVA.
• **Not implemented: The PVA Manipulation Tracker:** This component, intended to capture manipulation of virtual objects, is conceptually important for the symmetry of the system because a system providing physical-virtual infrastructure should ideally capture manipulation of both physical and virtual objects (PVA manifestations) and propagate state change events accordingly in both directions. However, since the tracking of human agent activity in the virtual world is relatively simple from a technological standpoint compared to the more challenging issue of tracking physical phenomena, the implementation of this component was kept at low priority. Furthermore, it was believed that “event-sniffing” software for tracking virtual activity of human agents probably already exists and thus could be added relatively easily whenever the rest of the system was up and running. Another not ignorable issue was also that if implemented, what should we do with the captured events? Physical actuation technology is still very immature and thus, the value of capturing virtual object manipulation in order to propagate changes to the physical world was estimated to give little payoff.
10.6 Magic Touch 0.5 — a (partially) wearable PVA Manipulation Tracker

Hardware evolution

- Version 0.50: Non-wearable unit consisting of RF/ID-reader electronics with integrated antenna + one location pod — proof of concept. Pictured to the left in figure 56.
- Version 0.51: Wearable unit consisting of RF/ID-reader electronics separated from a stiff external antenna placed in front of finger tip + two location pods, one attached on top of the finger tip, one on the forearm (see figure 57).
- Version 0.52: Same as 0.51 but with only one location pod on top of the finger tip. Pictured in the centre of figure 58 as well as in action in figure 67.
- Version 0.53: Wearable unit consisting of RF/ID-reader electronics separated from a small semi-soft external antenna placed below the finger tip + one location pod attached on top of the finger tip. Pictured to the right in figure 56.
- Version 0.54 (only planned, not yet implemented): wireless version of version 0.53 combined with three or four RF/ID tags acting as a wearable keyboard/input device.

FIGURE 56. Two versions of the wearable component of the Magic Touch PVA Manipulation Tracker. The electronics on the left paper sheet combined with the position pod in the center constitutes Magic Touch PVA Manipulation Tracker v0.50. The small black box is an RF/ID reader with integrated antenna and the larger circuit board is just a serial interface (RS232) board. The hardware on the right paper sheet constitutes v0.53 of the PVA Manipulation Tracker when the position pod in the center is glued on top of the tiny finger-worn white external RF/ID antenna. The larger black box contains the RF/ID electronics and is worn on the wrist.
FIGURE 57. Magic Touch PVA Manipulation Tracker v.0.51 in action. The photo shows (a) an RF/ID tag attached to a paper document, (b) the stiff antenna, and (c) position pods.

FIGURE 58. Three generations of external RF/ID antennas. The left most antenna is an early attempt based on the idea to fit a relatively large soft antenna on the backside of the hand. The one in the center is the one used as a component in PVA Manipulation Tracker v.0.52 (shown in action in figure 67), using a stiff smaller antenna placed directly in front of the finger tip. The right most antenna is part of PVA Manipulation Tracker v.0.53 (shown in figure 56) and is the most unobtrusive design. The antenna is hidden inside two layers of white elastic and placed directly below the finger tip. The position pod (fit on the v.0.51 version in the above picture) is normally glued on the “up” side of the elastic.
Software evolution

The “driver” software for the wearable unit has been continuously enhanced as part of the development of the general Magic Touch system. In Magic Touch 1.0 (described in the next section) the driver software is an integrated part of the whole Magic Touch system. In Magic Touch 2.x, the driver software is separated from the rest of the system and acts as a client towards the Magic Touch Object Server (described in the section on Magic Touch 2.0-2.3 later).

10.7 Magic Touch 1.0 — creating PVA infrastructure

Exploring the design space

Having proven that the chosen approach for tracking physical objects actually worked in practice (Magic Touch 0.5), more detailed requirements for a physical-virtual infrastructure were worked out. The process of combining technological possibilities with ideas for what functionalities that could be useful in knowledge work settings in office environments was made in an exploratory manner by imagining and designing future use scenarios and through brainstorming sessions. Several of the ideas were put on ice or discarded because of their demand for development resources, others because of technological limitations of the PVA Manipulation Tracker hardware.

All persons involved in this design phase were acquainted with several empirical studies on knowledge work (e.g. Malone, 1983; Kidd, 1994; Kwasnik, 1989; Mander, Salomon & Wong, 1992) and could without doubt also be regarded as belonging to this work category themselves as undergraduate and graduate university students. Thus, a certain understanding of knowledge work and for knowledge work activities was present within the group of developers. The concept of physical-virtual artefact (see section 5.2, page 94) served as a conscious (and probably also unconscious) conceptual anchor to which many ideas could be tied.

Basic system requirements

It was decided that as a first system requirement, the system should be able to capture and model basic object organisation activity found in empirical studies of knowledge work. Examples of human agent actions that should be covered (that is, both captured and in one way or the other facilitated) by the system should be:

- the placing of objects and containers in dedicated places based on their state in context of the (physical-virtual) activities the objects are involved in,
- the more or less temporary or improvised grouping of objects in piles,
- and the more conscious and long-term-intended packing of object sets into different kinds of containers.
Physical object ontology

For the purpose of capturing the previously mentioned basic knowledge work actions, a simple object ontology was developed on the basis of what information the Physical Activity Tracker could deliver by continuously tracking the picking-up and dropping of objects in the physical environment. Specifically, we found it useful and to some extent necessary to distinguish between objects based on:

- whether they are to be tagged / not tagged
- whether their extension in space can/should be taken into account or can/should be ignored
- whether they can be moved by the human agent or can be assumed to have a static spatial position
- whether they can be “piled”
- whether they can contain other objects

The object aspects listed are not mutually independent. For instance, in order to successfully track the translation of an object from one place to another, within the Magic Touch 1.0 system, usually the object has to be tagged as well.

Table 12 shows all 36 possible combinatorial combinations of the five object aspects listed above. Among the combinations, nine were considered of interest for capturing and modelling object organisation activities or for physical-virtual system control, based on use scenarios. Of these nine “candidate object types”, four were chosen to be implemented in a first development phase based on estimations of ease-of-implementation vs. utility:

- active volumes
- atomary objects
- containers
- implicit containers

The four chosen object types (shaded in table 12) were believed to provide the necessary minimum modelling support for creating a physical-virtual infrastructure that captures basic object organisation activity. Short descriptions of the four implemented object types follows below.

Active volumes — cutting through and structuring physical space

The fact that human agents in everyday situations often refer to physical space not by using coordinates but rather by dividing the space up in “places”, or geometrically defined three-dimensional “volumes”, was the starting point for the space model developed for Magic Touch 1.0. Since it was easy to imagine that activity in these volumes could trigger virtual (and potentially) physical actions by the PVA infrastructure, and since these volumes were imagined to be always monitored by the same system, these bounded geometrical spaces were given the name active volumes. The idea was that “active” signals that in these digitally augmented subspaces of the environment, anything can happen.
TABLE 12: Attributes and manipulation affordances of candidate object types in Magic Touch 1.0. The shaded columns (“is tagged”, “has geometry”) are attributes of importance from a system design perspective because they are products the technology used for implementing the PVA Manipulation Tracker. The unshaded columns (“can be moved”, “can form implicit container”, “can contain other objects”) are attributes that constrain object manipulation, and distinguishes object types from each other based on studies of human agent activity patterns in office environments. Of the nine object types identified in scenarios as potentially useful, four were selected for implementation (shaded in the table; “implicit container”, “active volume”, “atomary object”, and “container”).

<table>
<thead>
<tr>
<th>is tagged</th>
<th>can be moved</th>
<th>has geometry</th>
<th>can form impl. containers</th>
<th>can contain other objects</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>implicit container</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>active volume (untagged)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>spatial reference point</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>container that cannot form implicit container</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>atomary object</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>container (that can form implicit container)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>geometrical object that cannot form implicit container</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tagged active volume</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>geometrical object (that can form implicit container)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
However, the main purpose of active volumes is not to provide “digital magic”, but to act as an aid for the system in modelling physical space by defining the boundaries of physical places meaningful for the human agent acting in the environment. Active volumes are in their simplest outfit nothing but static, exactly geometrically positioned, sections of air that (hopefully) for human agent(s) are associated to some meaningful activity or state of activity. These meaningful sections of air normally coincide with (parts of) statically positioned physical objects. In office environments, typical such “place-defining” objects are bookshelves, drawers, desktops or parts of the walls or the floor. Figure 59 and 60 illustrate the total imperceptibility of active volumes if not supported by physical structure. Can you see the difference? Obviously, the human agent that defines a specific MT space has to make sure that every active volume is supported by some kind of perceivable structure, otherwise the space defined by the active volume cannot in general have a significant semantical meaning for the agent acting in the environment and is therefore not useful.

Active volumes can be nested (i.e. contain each other) endlessly. A room can contain a filing cabinet that can contain drawers and so on. Some objects like desktops and shelves do not have any significant perpendicular walls that support the notion of being able to contain anything. When defining the spaces that these objects provide, the geometrical coordinates of these objects are input to the system such that they in the eyes of the system actually do provide a three-dimensional space, by telling the sys-

**FIGURE 59.** An empty desktop.

**FIGURE 60.** The same desktop containing an active volume.
tem that the objects do in fact have a perpendicular geometrical extension. For in-
stance, a desktop area is typically defined as an active volume inputting one of the
coordinates 10-20 centimetres above the actual desktop surface. In this way, objects
put on top of the desktop will be interpreted by the system as being contained by the
desktop, which in the majority of the cases is the desired interpretation.

**Atomary objects — the simplest object type**

In Magic Touch 1.0, the majority of objects whose changes of position is tracked be-
long to the category of atomary objects. They are denoted “atomary” for two reasons:
1) Their individual extension in space (their “geometry”) is ignored by the system and
instead approximated as spheres of a fix and uniform radius with the centre positioned
where the tag is attached on the object. 2) They are, from the perspective of the sys-
tem, incapable of containing other objects. Both restrictions on atomary objects stem
from technological limitations of the wearable PVA Manipulation Tracker which can-
not sense the geometry of objects, and thus also not if an object is placed within the
real-world geometry of an atomary object. Typical examples of atomary objects in of-
fice environments would be paper documents, pens, and staplers.

**Containers — supporting manipulation of multiple objects**

Objects are not always moved around in the physical environment one at a time. Tech-
nical limitations of the PVA Manipulation Tracker prevents the system from success-
fully recognising more than one object at a time in the hand of the human agent.
Containers provide a somewhat constrained but still functional way of moving a set
of objects from one place to another in one single translation operation. Whenever a
container is moved, all objects contained by the container is assumed by the system to
move along as well. As can be seen in table 12, containers are atomary objects with the
added feature of being able to contain other objects. Examples of potential containers
in office environments are folders, binders, boxes, and coffee cups.

**Implicit containers — formalising strong semantical relationships between objects in
heaps and piles**

Implicit containers do not have a structural function in the physical world but formal-
ises strong relationships between objects that are placed close to each other. Implicit
container is the Magic Touch 1.0 correspondence to a pile. The term “implicit con-
tainer” comes from the idea that piled objects can be viewed as objects in a box with-
out the box. Implicit containers are abstract objects that the system generates on the
fly, whenever object’s are put closer to each other than a certain threshold radius. (A
threshold value of around ten centimetres was found to adequately distinguish piles
of paper documents from paper documents lying on the side of each other.) There are
no single physical objects in office environments that correspond to implicit contain-
ers, except that objects grouped together, e.g. in a pile, can be said to constitute the
building blocks of an implicit container that could (I speculate) be present in the mind of the human agent that has grouped the objects together, as a way to regard the object set as a single unit.

**The physical space as a containment-based hierarchy**

The model of physical space maintained by the Magic Touch 1.0 system is based on containment relationships between the four earlier mentioned object types. This was because 1) I believe that to structure space based on geometrical containment lies close to human everyday mental models of physical space, 2) it is a model that can relatively easily be maintained by computer systems.

**Expanding the concept of containment**

In order to fully capture the most important location-related semantics of office environments (e.g. that a paper document has been moved from the bookshelf to the desktop), and in order to cope for technological limitations of the PVA Manipulation Tracker (the fact that the extension in space of an atomary object could not be measured) the development team had to expand the notion of containment and use it in a broad sense in order to succeed in creating a useful model of physical space.

Having three-dimensional geometrical inclusion as a starting point, the concept of containment was expanded to include also two other kinds of object relationships:

- **Two-dimensional inclusion:** Small objects lying on top of larger objects, kept there by gravity, should be regarded as contained by the larger object (e.g. a paper document on a desktop). The same goes for smaller objects attached to larger objects in the vertical plane, counteracting gravity, (e.g. a painting on a wall).

- **Containment by vicinity:** Containers are objects designed to contain other objects, yet they do not have any geometrical extension in physical space as far as the system is concerned, because of the inability of the PVA Manipulation Tracker to provide such information. The Magic Touch 1.0 system gives all (from a system viewpoint) extension-less physical objects an imaginary spatial extension in the shape of a sphere. Whenever the imaginary spatial extension of an atomary object crosses the imaginary spatial extension of a container (implicit or explicit), the atomary object is considered to be contained by the container. Figure 61 shows an example. The containment-by-vicinity approach requires the real-world spatial extensions of container objects to be relatively uniform in order to act as a sufficient approximation of their real geometrical extensions. Being tuned towards office environments where containers for paper documents are dominant, and where papers are almost all of the same shape, this approach to “simulate” objects’ extension in space has been shown to work remarkably well as long as objects that are not supposed to engage in a containment-by-vicinity relation are not put too close to each other accidentally.

Figure 62 shows clear “standard” geometrical containment situations. Figure 63 shows situations that are more open for interpretation.
FIGURE 61. Containment-by-vicinity example. A transparent folder (container), a book, and a brochure (atomic objects). If the book and the brochure would be pushed slightly more to the left, so that their RF/ID tags come closer to the tag attached to the folder (i.e. enter the imaginary sphere with origo at the coordinates of the folder tag), they will be regarded as contained by the folder by the Magic Touch system.

FIGURE 62. Outside and inside the active volume represented in physical space by the red container and in the virtual space by the icon named “red box”. Two clear cases both for human agents and the system.
The important factor when defining active volumes has been shown to be the very exact input of the end points that define the object’s spatial extension. An error of just a few centimetres can make the system’s conception of whether an object is inside or outside an active volume to diverge from the view of a human agent. On the other hand, if defined with care, the system will interpret containment relations involving active volumes in a way that seems very intuitive. The cases shown in figure 63 are questionable cases both for the system and for human agents. Another important factor related to containment and active volumes is to be careful not to move the objects that act as “placeholders” for the active volumes. If such a placeholder object (the red box in figures 62 and 63 is translated or completely discarded, the human agent will have no perceptual support for understanding where the active volume is, or even that there is one in the first place. If the red box is taken away, and the phone book is placed approximately where it was in the situation pictured in the right photo of figure 62, we have the situation pictured in figure 64. For the system, the situation shown in figure 64 and the one pictured in the right photo of figure 62 are identical. Not so for a human agent.
The “rules” of Magic Touch 1.0 space

The intended use of the four different object types, and the extended definition of containment restricts how spatial relationships between objects can be interpreted. Table 13 shows what object types that can be contained by, and contain, what other object types.

<table>
<thead>
<tr>
<th></th>
<th>atomary objects</th>
<th>containers</th>
<th>implicit containers</th>
<th>active volumes</th>
</tr>
</thead>
<tbody>
<tr>
<td>atomary objects</td>
<td>can be contained by</td>
<td>no (1)</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>containers</td>
<td>can be contained by</td>
<td>no (2)</td>
<td>no (5)</td>
<td>yes</td>
</tr>
<tr>
<td>implicit containers</td>
<td>can be contained by</td>
<td>no (3)</td>
<td>no (6)</td>
<td>no (8)</td>
</tr>
<tr>
<td>active volumes</td>
<td>can be contained by</td>
<td>no (4)</td>
<td>no (7)</td>
<td>no (9)</td>
</tr>
</tbody>
</table>

**TABLE 13:** What object can be situated where? Possible combinations of containment as to Magic Touch 1.0.

**Allowed combinations — the seven “yes”**

Active volumes can contain any object type including themselves (illustrated by the right most column in table 13). Instead, atomary objects cannot contain any object at all. Containers can contain atomary objects only. Implicit containers can additionally also contain containers.

But what about all the relationships between objects that are considered invalid by the system (the nine “no” in table 13)? Certainly, the system cannot prevent the human agent from putting physical objects as she or he wishes.

**Prohibited combinations — the nine “no”**

Magic Touch 1.0 avoids and prevents the nine invalid object-relationship situations in the following ways (the list item numbers correspond to the “no”-numbers in table 13):
1. **Atomary objects cannot contain any atomary objects.** Atomary objects do not have an extension in real-world space (known to the Magic Touch 1.0 system) and cannot therefore contain any object by real-world space inclusion. This leaves containment by vicinity based on the object’s imaginary spatial extensions (see “Containment by vicinity:” on page 225 earlier). However, Magic Touch 1.0 enforces the creation of an implicit container as soon as two atomary objects enter each others’ imaginary extension “spheres”, and interprets both involved objects as contained by that new implicit container. Without the creation of implicit containers, a decision would have to be made by the system as to which of the objects contains which. A decision impossible to take with 100% certainty based on the information available.\(^{57}\) Also, as mentioned before, implicit containers formalises strong relationships between objects placed close to each other, which we believe could be helpful in human agent object organisation activities.

2. **Atomary objects cannot contain containers.** Whenever a container is placed in the vicinity of an atomary object, an implicit container is created by the system instead, containing both the container and the atomary object, in analogy to the description in list item 1 above.

3. **Atomary objects cannot contain implicit containers.** Implicit containers are abstract objects impossible to grasp and move by the human agent. Their location in the physical space model is controlled by the system only. Hence, there are no means through which a human agent could possibly put an implicit container inside an atomary object.

4. **Atomary objects cannot contain active volumes.** Active volumes are assumed by the system to never change location in the physical space model. Thus, there is no way for a human agent to move an existing active volume into (actually close to since vicinity-based containment is the conceptually only possible way for an atomary object to contain another object in the model) an atomary object.

5. **Containers cannot contain containers.** The reason for this constraint follows the same argumentation as for why atomary objects cannot contain each other (see list item 1 above).

6. **Containers cannot contain implicit containers.** Implicit containers are system-generated substitutes for “real” containers. Whenever a “real” container contains a set of objects, it makes no sense for the Magic Touch-system to generate an implicit container. Further, as already described in list item 3, implicit containers cannot

---

\(^{57}\) One could of course assume that an object A that is placed close to another object B is to be always interpreted as contained by B, based on the fact that physical objects do not fall through each other and that gravity constrains any placing of objects inside each other to take on the previously mentioned sequence (if you want to put the ball into the basket you move the ball, not the basket). This assumption fails however, as soon as objects are placed on top of other objects with the purpose of covering or hiding them, in which case the modelling approach would propose to regard the covered object to contain the covering object, which is counter-intuitive.
be moved by human agents and thus cannot be placed inside “real” containers anyway.

7. Containers cannot contain active volumes. Same argument as list item 4.

8. Implicit containers cannot contain implicit containers. As explained earlier (e.g. list item 3) implicit containers cannot be moved by human agents and thus cannot be placed inside other existing implicit containers.

9. Implicit containers cannot contain active volumes. Same argument as list item 4.

**Virtual PVA manifestations (PVAs)**

As mentioned earlier, the concept of physical-virtual artefact played an important role in the design of the Magic Touch system. Having decided how to categorise and model objects in the physical environment, representations and mechanisms for presenting their virtual counterparts had to be defined. During the development of a virtual visual representation of the containment relationships between objects in the physical environment, a simple “iconic” virtual representation of physical objects themselves became the natural choice because of the already ubiquitous presence of file hierarchies in virtual environments. Figure 65 shows a snapshot of parts of the Magic Touch 1.0 PVA Configuration User Interface including its hierarchical representation of a (minimalistic) physical office environment.

**Adding the second virtual manifestation**

At this point in the development, the system could be seen as a system that visualised physical object-translation in real-time by highlighting and moving a virtual representation of the translated physical object within a hierarchical representation of physical space. However, as examples of physical-virtual artefacts, the MT objects at that time were rather simplistic since while the physical objects were rich in their representations and manipulation support, the virtual manifestations were mere icons that did not respond to almost any kind of virtual manipulation.

A mechanism for adding a second virtual manifestation to the MT objects was added to the system by allowing the human agent to link a data file, application or web URL to the MT object. By introducing this second virtual manifestation, the MT objects became PVAs with one physical and two virtual manifestations. Although virtual manipulation of the linked object was not propagated to the other manifestations (which would have made them even better synchronised PVAs), at least the virtual representation became richer. Now it was for instance possible to link a web page about coffee to a physical coffee cup. Figure 66 shows the configuration of a typical atomic object.

In the spirit of the physical-virtual design perspective (see chapter 8) the mechanism for inspecting linked virtual objects (i.e. the web page or data file) was made possible to access both from the physical and the virtual environment and is described more in detail later in this chapter (in the section “Magic Touch 1.0 in use” on page 233).
Input and output
The flow of information about events and the causal propagation of events themselves between a human agent and a computer system is in classic HCI literature described in terms of input and output. In the case of Magic Touch, these terms do not quite capture the essence of some of the actions the system supports because input to, and output from the “system” are sometimes completely inseparable (e.g. when moving a

**FIGURE 65.** Parts of the PVA Configuration User Interface of Magic Touch 1.0 showing a hierarchical virtual representation of objects in a physical environment and containment relationships between them. The object most recently, or currently, physically grabbed is selected (highlighted). Active volumes, containers, implicit containers, and atomary objects are represented as green folder icons, yellow folder icons, pink folder icons (not shown) and white file icons respectively. (Pederson, 2001b)
Input to the Magic Touch system is in general given by the human agent in one out of three ways:

1. By performing physical operations that the wearable PVA Manipulation Tracker can sense, like moving a tagged physical object from one place to another.
2. By using the keyboard and mouse attached to the PC running the Magic Touch 1.0 system for fine-tuning the system or for changing properties of MT objects (PVAs) such as what file (PVA) should be linked to what physical object (PVA).

3. By using both input devices (if we regard keyboard and mouse as one and the PVA Manipulation Tracker as the other) in combination, such as when creating new “active volumes”, a process that will be explained later.

Output
The system reveals its state through the computer screen as well as through the use of sound in specific situations. Because the system in a way also consists of the monitored physical environment, also the physical environment can be regarded as an output “channel” or modality.

Magic Touch 1.0 in use
Having described the main components of the system I now turn to the interaction aspects of it, e.g. the way objects are introduced to the system, manipulated and discarded. During the discussion a certain emphasis will be on the determination of on which side of the physical-virtual border the human agent has access to the different functionalities.

Systems management
Before interaction takes place in the physical-virtual space supervised and monitored by Magic Touch, the system needs to have information about the identity of all physical and virtual objects in the physical-virtual space, as well as the containment relationships among the physical objects. The system allows for loading and saving such Magic Touch “spaces” to hard disk as well as creating spaces from scratch.

Since the positioning component of the wearable PVA Manipulation Tracker runs on battery, a mechanism for turning on and off the tracker is also available. As shown in table 14, these systems management functions are available in the virtual environment only.

Basic MT object administration

- **To create an active volume:** Information about structural properties of the physical space is input to the MT system by the human agent by creating a set of active volumes. Their absolute spatial position and size serve as a frame of reference for all other objects in the specific MT space. As indicated in table 14, and as will be explained more in detail later, active volumes are defined by performing a physical-virtual activity (see definition 5.1.3, page 93) initiated in the virtual environment and mainly made up of physical-virtual action pairs (definition 5.1.2, page 92) consisting of the specification of the corners of the new active volume.

58. If a Virtual Activity Tracker would have been implemented, also containment relationships between virtual objects would have to be determined.
To create an atomary object: Creation of atomary objects is a purely physical action since the system automatically creates a new MT object whenever the PVA Manipulation Tracker encounters an RF/ID tag which is not yet in the PVA-DBMS, i.e. when the human agent holds a tagged physical object not yet in the database. The new MT object is immediately shown in the hierarchical tree structure with a default name possible to change at wish.

To create an implicit container: The creation of implicit containers is also a physical activity but does not involve any new tagged object. Instead, implicit containers appear in the hierarchical tree structure as an indirect result of putting two or more atomary objects close to each other in the physical environment.

FIGURE 67. Magic Touch version 1.0 in action. Parts of the PVA Manipulation Tracker version 0.52 is shown in the foreground worn on the thumb by a human agent (note the stiff square RF/ID antenna sticking out as a big finger nail). The rectangular object attached to the paper document on the right side of the thumb is a thin RF/ID tag. In the background, the monitored physical environment is visualized as a hierarchical tree structure. The PVA manifestation situated in the hierarchy that corresponds to the PVA currently grabbed by the human agent (in this case the technical report in the foreground) is automatically highlighted in the hierarchical visualization.
• **To create a container:** The creation of new MT containers is a physical-virtual activity consisting of a physical-virtual action pair where the first action (physical) is to create a new atomary object (explained above), and the second is to transform the atomary object into a container (the virtual action).

• **To discard an MT object:** All MT objects except implicit containers are discarded from the system by performing a virtual activity (selecting “delete” in the Magic Touch application menu bar). Implicit containers are discarded automatically when the atomary objects that previously defined the implicit container by being close to each other change spatial position such that the distance(s) becomes larger than the threshold value. Thus, just like the creation of implicit containers, the discarding of them is triggered by physical activity. Table 14 shows two light grey entries to denote that the discarding of atomary objects and containers is planned to be supported also as physical actions. (The idea is simply to tie the “discard object” mechanism to a specific active volume, i.e. a trash can so that whenever a physical manifestation of an MT object is placed in it, the MT object as a whole is discarded.)

• **To (re-)name an MT object:** Naming and renaming of MT objects is done by clicking on their representation in the hierarchical tree structure.

**Advanced object actions**

• **To create/change/discard an MT object link:** The second (optional) virtual MT object manifestation, i.e. a local or distant virtual file, is administrated in the virtual environment only.

• **To inspect a file or URL belonging to an MT object:** Inspection of any file associated with an MT object can be initiated both in the physical and the virtual environment. The physical alternative is performed by placing an MT object’s physical manifestation into an “inspection volume”, an active volume dedicated for the purpose (see below). As a result, the virtual manifestation linked to the physical is immediately shown on the computer screen. In the virtual environment, the same result is achieved by “double-clicking” on the MT object’s representation in the hierarchical tree structure.

• **To create an “inspection volume”:** This action reminds of how containers are created by first creating an atomary object and then transforming it. Inspection volumes are defined by first creating an ordinary active volume (a physical-virtual activity) and then to specify (a virtual activity) that it should initiate an inspection mechanism when an object is placed into it.

• **To do a free-text search for MT objects in the hierarchical representation:** MT object’s icon representations can be located in the virtual hierarchical structure using a free text search mechanism invoked in the virtual environment.

• **To determine the location of an MT object’s physical manifestation:** Also physical MT object manifestations can be searched for. The mechanism is invoked by “right-clicking” on the object’s iconic representation in the hierarchy and choosing “locate
physical” from the pop-up menu. The rest of the activity is a physical-virtual activity consisting of moving one’s hand in the direction suggested by the Magic Touch system (using both sound and visual indicators). Eventually the hand will end up at the location of the searched physical object.

**MT object move**

- **To move the physical manifestation of an MT object**: This is probably the most basic physical Magic Touch activity. Since only spatial location changes of tagged objects can be tracked by the PVA Manipulation Tracker, only location changes (movement) of atomary objects and containers is supported.
### TABLE 14: Magic Touch-specific actions, their environment locality (physical or virtual), and what MT object types that can respond to them. Legend: The dark grey rectangles show in what environment each specific action takes place. Actions represented by filled rectangles that overlap the border between the physical and virtual environments are actions made up of obligatory operations situated in both environments, i.e. they are examples of “physical-virtual activities” (see definition 5.1.3, page 93). Actions represented by completely filled rectangles in both worlds are actions that can be independently performed in either world. The light grey rectangles represent action locality support that is planned to be implemented.
10.8 Magic Touch 2.x — platform for ubiquity

As mentioned earlier, Magic Touch 1.0 was a single-user, single-environment design to explore the possibility of creating a physical-virtual infrastructure using RF/ID as identification technology and infrared light/ultrasound as position tracking technology. Having proven that it was actually possible (although not necessarily without design and interaction limitations), ways of improving the system were considered.

For stability, flexibility and networking: client-server architecture and Java/CORBA

The first decision was to completely start from scratch with regards to the software developing environment, for two reasons: a) MS Visual Basic at least at the time did not allow for the system stability and flexibility needed, b) an analysis showed that the functionality of the system could conceptually very easily be broken down into one “client” part and one “server” part, a design choice that would increase the applicability of the system to a wider set of environments and situations. The vision was (and still is) to provide an infrastructure for internet-mediated shared physical-virtual environments as shown in figure 68. It was natural to place the client-server communication on an Internet-based platform. Sun Java was chosen as the default programming language and CORBA as the platform for data communication.

Figure 69 provides a general diagram of the client-server architecture and its components. The different kinds of clients and the design of the server will be discussed more in detail later.

**FIGURE 69.** An overview of the most important components of the Magic Touch 2.x client server model. The Virtual Activity Tracking Client and PVA Classification Client (marked with dotted borders) are not yet implemented at the time of writing.
From an interaction perspective, the client-server model makes it possible to support inspection and manipulation of PVAs at a distance. Naturally, PVA manipulation (i.e. to change the state of physical objects) at a distance remains a challenge, but at least manipulation of virtual objects could be possible to support from any Internet-connected computer that could run a Magic Touch 2.x client.

**Java Applet or not Java Applet**

The clients written in Java and executable as Java Applets have the huge advantage from a user-perspective to not require any explicit installation procedure since they are downloaded and executed automatically from within a web browser. Thus, a general design decision was to make as many of the Magic Touch 2.x components as possible...
executable as Java Applets. However, some clients have been implemented using other languages when found necessary. For instance, the MT VR PVA Visualisation Client pictured in figure 80 on page 249 was written in C++ in order to utilise a specific VR software library. Other clients (like for instance the MT PVA Classification Agent) could be written in Java but cannot be run as “Applets” because they need to access the local hard disk, and thus need to be installed using standard procedure. (Java Applets are in general not allowed to write to local hard disks directly).

**For tracking technology independence and ease of modelling: less complex world model (Magic Touch 2.3)**

The object ontology designed for Magic Touch 1.0 was tailor-made for the specific RF/ID and motion tracking system used (see “Physical object ontology” on page 221). However, as the idea of a distributed client-server-based architecture arose, so did the idea that some Magic Touch 2.3 clients might (or at least should be allowed to) obtain their information about object manipulation in other ways, e.g. by using visual tags and a ceiling-mounted camera, and still work fine with regard to the communication with the server. Furthermore, it was hard to theoretically motivate the design choice to limit the modelling of some physical objects based on their ontological type, e.g. active volumes were assumed not to be moved; atomary objects were not supposed to have any spatial geometry, and so on (see table 12, page 222). Such restrictions, originating from the chosen position tracking method, made modelling of the physical environment unnecessarily complex. Thus, it was decided that objects (whether classified as “active volumes”, “atomary objects”, “containers”, or “implicit containers”) should be modelled as far as possible as having the same set of attributes. E.g. all objects should have a physical extension in space, all objects should be possible to move around, all objects should be able to contain other objects, etc. Whether the values of these attributes are for example obtained using RF/ID tags, visual tags or some other method will be irrelevant for the model that is maintained in the Magic Touch 2.x server and just depend on the design of specific MT PVA Tracking Clients.

A UML class diagram of this purified PVA model is shown in figure 70. All system classes inherit from a common virtual base class MTObject so that a PVA can contain also any other kind of Magic Touch objects: other PVAs (PVArtefact); PVArs (VirtualArtefact); PVArs (PhysicalArtefact); and human agents (MTUser). This object architecture supports distributed PVAs (see section 5.3, page 99) whose different manifestations (physical, virtual, or physical-virtual) are updated according to the logic described in the property attribute (retrieved from an XML file as described in the next section) as a reaction to specific Magic Touch events.
For server-side execution of causal effects: logic attributes (Magic Touch 2.2)

The kind of added causal effect that the Magic Touch system creates as a result of a PVA manipulation, e.g. the effect that putting a PVA in an "activate" active volume brings up the corresponding PVA on the screen\footnote{See “Added-value (end-user) functionality” on page 216.} was written directly into the code governing also other parts of the system. This made it hard to support any wishes for some PVA types to behave differently than others. It also made it harder for the system to execute more complex operations such as considering the state of other PVAs in order to decide among a set of proper effects. The Magic Touch 2.2 PVA server includes a “logic” attribute in the set of attributes belonging to each PVA that enables PVA designers to specify more complex behaviour as a result of PVA manipulation of specific PVAs. Concretely, the logic attribute is a reference to a file expected to conform to the XML standard and contain a description of what the server should do if the specific
PVA goes from one particular state to another. The design of the XML-based language for describing this PVA-specific causality is not yet finished and is left for future work. The long-term goal is to provide easy-to-use tools (MT clients) that allow the end-user programming of physical-virtual environments.

**The Magic Touch 2.2 PVA Server**

Figure 71 provides an overview of the MT 2.2 PVA Server in the context of communicating with a Physical Activity Tracking Client (PAT client) and a PVA Manager Client (PVAM client). For each PAT client that connects itself to the server, a new instance of a PAT server object is created, responsible for the handling of the specific client. New PVAM client notifiers are created in the same fashion for each new PVAM client connecting to the server. However, there is only one single CORBA service object in the PVA server. Thus, all PVAM clients communicate with that CORBA object.

![Diagram of Magic Touch 2.2 PVA Server](image)

**FIGURE 71.** The architecture of Magic Touch 2.2 PVA Server. Adapted from Jonsson (2002).
MT PVA Manipulation Tracking Client

Figure 72 shows screenshots from the User Interface (UI) of a, very slimmed, PVA Manipulation Tracking Client. The user can start and stop the tracking of the combined hand position and object identification process, as well as connect to a PVA Server. Figure 73 shows the same PVA Manipulation Tracking Client UI in the mode of defining an active volume.

*Making the tracking user interface wearable as well*

The process of defining active volumes in Magic Touch 1.0 and 2.0 is awkward from an interaction perspective. The user needs to place the tracked hand at the corners of the intended volume and at the same time “lock” the points (pushing the “Set point n” buttons in figure 73) using the mouse and visual display. In short, only parts of the PVA tracker is actually wearable. The other part (the “explicit” UI) is spatially located
where the PC running the PVA Manipulation Tracking Client application is. Among all the possible design alternatives for providing interaction control at the point of interaction, the idea of using RF/ID tags as “interaction control” buttons was the easiest solution. This solution still means that parts of the explicit UI (the visual display offering feedback) remains stationary and forces the user to perform complex MT operations at locations where the screen is readable, but eliminates the need for also carrying a mouse and/or a keyboard.

The improved design (illustrated in figure 74) requires attachment of three RF/ID tags on the hand in positions reachable from the finger to which the RF/ID tag detector is worn, as well as implementation of a more “intelligent” PVA Manipulation Tracking Client able to filter out these tags as buttons rather than standard PVAs grabbed by the user.

![FIGURE 73. Defining an active volume using the (non-wearable) user interface of Magic Touch 2.0. (Jonsson, 2002)](image)

![FIGURE 74. The conceptual model of the Physical Activity Tracking Client supporting RF/ID tag-buttons for controlling basic Magic Touch interaction.](image)
Figure 75 shows the tag “buttons” intended to be worn on the hand and figure 76 shows a mode flow chart describing how interaction with the tags corresponds to PVA Manipulation Tracking Client commands.

As illustrated in the same figure (figure 76), also the deletion of PVAs is afforded by this display-less menu system. The exact placement of the tags on the hand, as well as how the probably glove-like material structure needed to actually keep them there

---

**FIGURE 75.** RF/ID tags as parts of the user interface of the PVA Manipulation Tracking Client. Each tag is approximately 4.5 x 4.5 centimeters in size and less than 1 millimeter thick. (Jonsson, 2002)

**FIGURE 76.** Modes and button-press sequences using the wearable user interface designed for the Magic Touch PVA Manipulation Tracker Client. (Jonsson, 2002)
should look, has not been decided at the time of writing. (Initial informal tests have shown however that the approach is certainly doable and improves interaction compared to the Magic Touch 1.0 approach.)

**MT PVA Manipulation Tracking Client (not implemented)**

Just like in the case of Magic Touch 1.0, the development of Magic Touch 2.x has given low priority to the implementation of components that monitor users’ manipulation of virtual objects. The motivation is as before that this problem is less challenging than trying to perform the same thing in the physical world. The architecture design as a whole does of course include such a component, following the physical-virtual design perspective outlined in this thesis.

**MT PVA Manager Client**

The purpose of the PVA Manager Client is to provide means for administrating PVAs and the PVA infrastructure offered by the PVA server. The most important functions include to allow the user to add, change, and remove links between PVAs to PVAs, and to manage the sharing of PVAs and active volumes among several users at potentially geographically disperse places. Screenshots from a simple minimalistic PVA Manager Client is shown in figures 77, 78 and 79.

![Image](image_url)

**FIGURE 77.** Peter logging in to the PVA server in order to get access to information about his PVAs, and the PVAs which he shares with other users. (Landfors, 2001)
MT Visualisation Client

The purpose of MT Visualisation Clients is to allow for presentation of PVAs and PVA sets in various ways and from various perspectives. The visualisation component in Magic Touch 1.0 visualises the layout of PVAs in physical space as a hierarchy based on physical-world containment relationships among the PVAs (see figure 65, page 231). The only MT 2.x Visualisation Client implemented to date instead presents a VR-based visualisation. Screenshots from this client is shown in figure 80. Of particular use is the ability of the VR visualisation client to make the otherwise completely invisible active volumes (see pictures 59 and 60 on page 223 earlier) virtually visible, enabling “debugging” of the physical world model maintained by the system.

MT PVA Classification Client (not implemented)

The idea of having a personal tool that automatically analyses activities monitored by the Magic Touch system, and provides more in-depth data-mining facilities than visualisation clients are intended to provide, resulted in a conceptual design of a PVA Classification Client (Salmaso, 2002). The main feature of this client is that it makes use of the fact that digital information about PVAs is already available to the system in the shape of links to virtual objects (e.g. web pages) as part of a Magic Touch ob-
ject’s virtual manifestation. The PVA Classification Client is intended to parse the virtual PVA manifestations (PVAs) in the background and relate them to each other using Latent Semantic Analysis (LSA).

The client is intended to store time-stamps of, and details about, PVA intra-manipulation and extra-manipulation (e.g. changes done to a digital document and the change of location of a physical object respectively) for later analysis and visualisation. A conceptual model of the client has been designed and a detailed analysis of how this client is to communicate with the Magic Touch 2.x client server has been performed. The PVA Server 2.2 and 2.3 has been prepared for the PVA Classification Client by including commands in the client-server communication protocol that are tailor-made for a future PVA Classification Client.

The flexible architecture

The splitting of Magic Touch 1.0 into separate client-server components makes it possible to integrate many different kinds of physical-world object identification and tracking systems and install them in any physical location in the world (assuming there is Internet access). The distributed architecture also allows for the development of a great variety of more or less tailor-made clients that combine PVA administration, visualisation, and PVA manipulation tracking in different ways. Figure 81 shows an example of MT clients that choose to integrate different client functions to a larger or...
FIGURE 80. A Virtual Reality MT Visualization Client in action. (Granholm & Hägglund, 2001)
smaller degree (the dark grey rectangles symbolise that the contained light grey components are integrated into one single application). Nothing in the architecture prevents running Magic Touch clients also on PDAs or modern cellular phones. It is in fact part of planned work to develop a PVA Manager Client, Visualisation Client, and MT PVA Activity Tracking Client that can run on very small, energy-efficient, and wearable computing systems. From the perspective of the PVA Server, these clients will of course not look different from the clients that have already been implemented and presented in this chapter.

10.9 Magic Touch in the eyes of the framework

Although terminology from earlier parts of the thesis has already been used to describe various aspects of Magic Touch in this chapter, a more focused analysis of the system from the perspective of the theoretical physical-virtual design framework will be presented below in order to point out the tight ties between theory and practice in the PhD project presented in this thesis. Since parts of Magic Touch 2.x is not yet completely implemented, the discussion applies only to Magic Touch 1.0 unless explicitly noted.

FIGURE 81. The Magic Touch 2.x PVA Server in data communication context. Different client configurations and connection to other PVA servers. Adapted from Jonsson (2002).
The PV design perspective — applied
MT development focus has been on the role of documents and relationships between documents in office environments inspired by the fact that documents afford different kinds of manipulation in the physical and the virtual world (e.g. Sellen & Harper, 2002). The fundamental idea behind the proposed physical-virtual design perspective to view the physical and the virtual environment as one, is manifested in the Magic Touch system as an infrastructure for synchronised PVAs and to some extent also by affording PVA manipulation mechanisms in environments they are not normally present, e.g. alphabetical search for the location of physical objects. The intention as system designer has been to decrease some dimensions of the physical-virtual environment gap that is present when manipulating physical and virtual documents. While basic synchronisation of PVAs is afforded when PVAs are manipulated, synchronisation that propagates PVA manipulation to corresponding PVAs has not been implemented. However, the architecture is designed for inclusion of such a mechanism at a later stage and can thus be said to be designed in accordance with the proposed physical-virtual design perspective that among other things promotes the idea of an infrastructure affording/capturing/actuating manipulation of PVAs in any of the two worlds, as the human agent wishes, and whenever it is technically and conceptually possible to allow for it.

The difference between the extensively integrated physical-virtual environment envisioned by the physical-virtual design perspective on one end (illustrated by figure 20, page 95; and figure 44, page 167), and the extent to which the actually implemented prototype system Magic Touch manages to comply to that vision (discussed in the present chapter) on the other, learns us two things: First, although Magic Touch is certainly not the most advanced physical-virtual environment prototype there is, it is one out of many comparable Augmented Reality/Mixed Reality systems that together establish the fact that state-of-the-art technology of today allows only for fulfilling small parts of the vision. Secondly, the difference between the vision and the state-of-the-art prototypes opens up a design space and can help direct future development and research.

Inter-World Event Mediators
A technological key issue for any physical-virtual environment is the quality and quantity of Inter-World Event Mediators (IWEMs) used by the system to interface the physical and virtual sub-environments (see “Inter-world actions” on page 81). Magic Touch uses only one IWEM apart from standard PC IWEMs such as mouse, keyboard and screen: the PVA Manipulation Tracker, which (if we disregard some obstacles caused by technological limitations) is a physical object that demands and gets no attention from the human agent. The fact that it is constantly mediating events, and that it is (ideally) completely unobtrusive, makes it different from most classical input devices since it does not demand attention from the human agent. The fact that it is constantly mediating events, and that it is (ideally) completely unobtrusive, makes it different from most classical input devices since it does not demand attention from the human agent wearing it. Regarding the quality aspect of the PVA Manipulation Tracker, it mediates only very basic
information about events related to extra-manipulation of physical objects. Taken together with the model of physical space maintained by the system however, the information has been proven to sufficient for keeping track of location changes of physical objects.

**Enabling synchronised PVAs**

All kinds of MT objects (atomary objects; containers; implicit containers; active volumes) are examples of more or less synchronised physical-virtual artefacts (PVAs, see page 95). Atomary objects are examples of distributed PVAs (page 99) since they can (optionally) have two virtual manifestations. Magic Touch 2.x additionally supports synchronised PVAs shared among multiple human agents (page 101).

The sensation of unity of corresponding PVA manifestations is probably strongest for atomary objects because as soon as a physical manifestation of an atomary object is grabbed by the human agent, its virtual manifestation in the hierarchical tree structure is highlighted. (This can also be viewed as an alternative search method to the free text search mechanism.). Taken together with the search mechanism for finding physical manifestations of atomary objects and containers, the system maintains an infrastructure that by and large fulfils the third requirement on physical-virtual artefacts as defined in definition 5.2.1, page 94, i.e. that “one manifestation of a specific physical-virtual artefact is easily identified if a corresponding manifestation in the other environment is known”.

One feature of the system is that the combining of physical and virtual objects into PVAs is not pre-defined by the system designers but something left to the human agent acting in the environment to perform as she or he wishes. Thus, Magic Touch provides an infrastructure that enables the end-user construction of PVAs rather than offers them “ready-made”.

The sensor and actuation technology used for capturing and actuating object manipulation is minimalistic. Events from manipulation of PVAs are (as mentioned earlier) mediated by the sensors built into the PVA Manipulation Tracker. No physical-world actuators are used by the system and hence there is no support for PVA synchronisation from the virtual world to the physical. In the virtual environment, only standard sensor and actuation mechanisms are used (e.g. object manifestations get “selected” by clicking on them etc.).

**Affording and demanding physical-virtual activities**

In general, the system focuses on capturing and facilitating “meaningful and observable” activities (see page 82), i.e. it makes no great attempt to infer user intentions (in contrast to the approach often taken in the Context Awareness research community where sometimes very complex user modelling is performed in order to guess what the human agent would like to do or means with her/his actions).
The physical-virtual activities (definition 5.1.3, page 93) that are enabled or facilitated in the physical-virtual environment provided by Magic Touch (other than “standard” physical-virtual activities that can take place in any modern office) are both activities involving the manipulation of domain objects and activities directed towards administrating the system itself. Below is a list of activities and a discussion on how they are related to the theoretical physical-virtual design framework outlined in previous parts of the thesis.

Physical action with virtual side-effects

The action to move a physical manifestation of an MT object from one location to another (specifically, from one active volume to another active volume) can be classified as a “physical action with virtual side-effect” (illustrated in figure 14, page 85). As soon as the PVA is dropped at its new location, the corresponding PVA in the virtual hierarchical model is immediately and correspondingly translated as well. The fact that this virtual translation takes place completely automatically and without demanding attention from the human agent makes it a “side-effect”. Complementary, the PVA translation action can be viewed as an extra-manipulation (see page 179) of the moved object and intra-manipulations of the active volumes involved.

Physical→virtual actions

The mechanism provided by the system for easy inspection of virtual manifestations of PVAs by moving a physical atomary object into an “inspection” volume, resulting in the display of the corresponding virtual manifestation in the virtual world, can be seen as a physical→virtual action (the case illustrated in figure , page 93).

Obligatory physical-virtual action pairs

Table 14, page 237, shows five cases of MT actions that are inherently physical-virtual because they depend on physical-virtual action pairs (definition 5.1.1, page 90), i.e. they cannot be performed without switching between the two environments: the creation of a new MT container; the creation of a new active volume; the transformation of an existing active volume into an “inspection volume”; to determine the location of the corresponding physical manifestation of a virtual atomary object and container respectively.

Physical-virtual redundancy

As shown in table 14, page 237, one of the MT actions can be performed independently in either environment (ideal case illustrated by figure 19, page 94), namely the inspection of the file or Internet URL belonging to an MT object. In the physical environment, the mechanism is invoked by putting the PVA into an “inspection volume”. In the virtual environment, it is invoked by “double-clicking on the MT object’s virtual manifestation in the hierarchical tree structure. Other redundant ac-
tion support could be relatively easily implemented as well (planned for future work), namely the support for discarding atomary objects and containers in the physical environment (marked light grey in table 14, page 237).

**Spanning the physical-virtual space**

All in all, Magic Touch 1.0 serves as an example of a system that offers object manipulation mechanisms in both the physical and the virtual world. As illustrated by the right most column in table 14, page 237, some operations are world-specific and can only be performed in one of the worlds (e.g., to start and stop the PVA Manipulation Tracker can only be done in the virtual world; the operation to move a physical MT atomary object from one place to another can only be performed in the physical world), other mechanisms can be invoked in any of the two worlds (e.g. to inspect a virtual file or web page associated with a specific MT atomary object), and a third class of activities are inherently physical-virtual, relying on physical-virtual action pairs (definition 5.1.1, page 90), and cannot be performed without performing parts of the activity in one world and other parts of the activity in the other world (e.g. the action of defining an active volume).

While the “purely physical” and “purely virtual” actions can be analysed and described using conventional physical and virtual spatial models respectively, the five physical-virtual MT actions calls for a unified model such as the proposed situative model of physical-virtual space (page 173 and on). Without a unified model of physical-virtual space, these actions cannot be described as homogeneous actions taking place in a single specific spatial location but only as a sequence of operations distributed in multiple physical and virtual locations independent from each other. Using the situative physical-virtual space model, domain objects involved in each of the five previously mentioned physical-virtual actions can be described as entering and leaving the manipulable and observable PV space disregarding if the objects happen to be physical, virtual, or physical-virtual and disregarding if operations involved in the actions take place in the physical or the virtual world.

**Bridging the physical-virtual environment gap**

Magic Touch addresses the following gap dimensions listed in table 7, page 131:

- The “manipulation affordance” dimension (described on page 127) and the “organisation/navigation” dimension (described on page 127) by 1) allowing semi-automatic search for objects in the physical world (an action normally only available in the physical world), 2) by providing a simple physical mechanism for inspecting PVA if the PVA is known (the “inspection container”), and 3) by automatically highlighting PVA manifestations of PVA currently being manipulated, making it easier to continue the manipulation of a PVA in the virtual world by eliminating the otherwise necessary manual search for the corresponding PVA when making the environment switch.
The “causal dependency” dimension by affording automatic synchronisation of some (few) properties between PVA manifestations, namely the synchronisation of the location of PVA with their corresponding PVA manifestations within the virtual hierarchical representation of physical space.

**Intra- and extra-manipulation of PVAs**

As mentioned earlier, the automatic synchronisation of PVAs is quite limited in the Magic Touch system. The capturing and modelling of object intra- and extra-manipulation (see page 179) is basically limited to object translation of physical objects, i.e. physical domain object extra-manipulation. Although the object translation-operations performed by human agents are measured by the system on a fine-grained level of granularity, the spatial model presented towards the human agent is based on containment relationships among the objects and more coarse-grained. This is however regarded as a feature because the idea is that the mental model of the space maintained by the user is assumed to be a high-level model as well: Being presented with a table of three-dimensional coordinates would be less useful in most cases. (Of course, the containment model also provides more information because objects and spaces are given names by the human agent as well.)

The system captures and models extra-manipulation of physical domain objects (denoted in the system as “atomic objects”), intra-manipulation of the type of containers denoted “active volumes”, and both intra- and extra-manipulation of containers categorised as “containers” and “implicit containers”. A more detailed analysis on intra- and extra-manipulation sensed and actuated by Magic Touch is shown in table 15.

**Space model based on containment and containment hierarchies**

- The use of containment and containment hierarchies (as discussed in chapter 9) by the Magic Touch system for modelling physical space has already been discussed in previous sections of the present chapter, specifically the sections “Physical object ontology” on page 221, and “The physical space as a containment-based hierarchy” on page 225. Since the module for tracking virtual object manipulation (the PVA Manipulation Tracker) is not implemented, no corresponding hierarchical model of the virtual space is maintained by the system at the time of writing. This also means that the system does not maintain any hierarchical model of physical-virtual space as described in section 9.4, page 203, either.

---

60. The system does however make use of hierarchical models of virtual space provided by external software (e.g. the local file hierarchy and web site hierarchies) for accessing virtual objects that the human agent chooses to include as an MT object PVA manifestation.
Many ideas presented in this thesis are based on experiences from developing and testing out Magic Touch. As inspiration for theory construction, the system has been very useful. The analysis of Magic Touch in the eyes of the framework indicates (not surprisingly) that the technology available for building physical-virtual environment as implicitly envisioned by the theoretical framework (e.g. the idea to support extensive \( V \rightarrow P \) PVA synchronisation) is still limited.

### On the absence of formal empirical studies

The system is unstable, takes time to learn (in particular the part of grabbing and dropping objects in a way so that the PVA Manipulation Tracker recognises it), is bulky (the cable would prevent subjects from working normally). In short, the prototype needs considerable refinements before any user studies can be conducted.

### TABLE 15: Intra- and extra-manipulation of physical-virtual artefacts sensed and actuated by Magic Touch 1.0. The shaded table elements represent the kind of object manipulation that is the main focus for the system. Legend: RO = reference object.

<table>
<thead>
<tr>
<th>framework object type</th>
<th>MT object type</th>
<th>physical world</th>
<th>virtual world</th>
</tr>
</thead>
<tbody>
<tr>
<td>domain objects</td>
<td>MT atomary objects</td>
<td>intra-manipulation</td>
<td>intra-manipulation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>extra-manipulation</td>
<td>extra-manipulation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>spatial translation</td>
<td>iconic manifestation: change name: change link to data file</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>with( \mathbf{S}_\text{obs}^V )</td>
</tr>
<tr>
<td>tools</td>
<td>MT system</td>
<td>system configuration</td>
<td>turning on/off the system</td>
</tr>
<tr>
<td>containers</td>
<td>MT active volumes</td>
<td>spatial translation of MT objects in and out of the container (RO) space</td>
<td>limited support (only discarding and (re-)defining, i.e. not translation)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MT containers</td>
<td>same as above</td>
<td>spatial translation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>same as above</td>
</tr>
<tr>
<td></td>
<td>MT implicit containers</td>
<td>same as above</td>
<td>spatial translation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>same as above</td>
</tr>
</tbody>
</table>

## 10.10 Results

Many ideas presented in this thesis are based on experiences from developing and testing out Magic Touch. As inspiration for theory construction, the system has been very useful. The analysis of Magic Touch in the eyes of the framework indicates (not surprisingly) that the technology available for building physical-virtual environment as implicitly envisioned by the theoretical framework (e.g. the idea to support extensive \( V \rightarrow P \) PVA synchronisation) is still limited.
Potential application areas “as is”

- As a help for visually disabled (to find things),
- parents to young children (to find the toys),
- for archiving-dense businesses handling important documents,
- for museums for keeping track of valuable items,
- for chemical and medical laboratories that should not accidentally mix the wrong substances.

Limitations and potential solutions

Limitations of the object location tracking approach has already been discussed earlier (page 213). In addition I would like to note the following limitations and potential solutions:

**Single-environment function access → dual**

The fact that some functions of the Magic Touch system are “single-environment functions”, i.e. they are accessible in only one of the two environments, limits and disturbs interaction. We should try to make them present in both worlds. One that seems easy is the discarding of MT objects which is a mechanism that could in part easily be made accessible in the physical environment as well (the two entries marked light grey in table 14).

**Limitations by using a desktop PC instead of a more mobile or wearable platform**

The inherently physical-virtual functions of the system suffers from the fact that the virtual environment is running on a desktop PC. More precisely, that the functions can only be accessed by the input devices connected to the stationary PC. As mentioned earlier in this chapter, it is for instance an obstacle to define active volumes because the physical sub-actions (pointing out the corners of the active volume) is, of course, location dependent. And the desktop computer is not necessarily reachable at all times. A mobile, or even better, a wearable system user interface would eliminate these obstacles by making the virtual environment accessible from anywhere in the physical environment. A step towards implementing a wearable user interface has already been taken (see the section on Magic Touch 2.x) and further work in this direction is needed if the PVA infrastructure should become accessible outside designated physical places.

10.11 Future work

**Personal PVA Server**

Apart from constructing various more portable Magic Touch clients, the idea of making also the PVA server wearable is also an interesting twist. This is a natural development from the perspective that the data stored in the server about PVAs and PVA manipulation typically belong to single individuals (apart from explicitly shared PVA
information). Combined with wearable PVA and PVA Tracking Clients this would enable the possibility to “share” ongoing local physical-virtual activities with geographically disperse people at wish, resulting in something that could be called a wearable or portable tele-presence system. Like cellular phones but including physical-virtual activities as mediated phenomena in addition to sound.

It is a natural solution also from the perspective that the same physical objects and places might be linked to different virtual object and places depending on individual interests and preferences. To carry a highly individual physical-virtual link ontology close to oneself, one that provides a personalised view of physical-virtual space, is as natural as having personal eye-glasses.

**PVA intra-manipulation tracking**

As for today, Magic Touch tracks only a specific kind of PVA extra-manipulation, namely object translation. Technology for tracking PVA intra-manipulation (e.g. hand writing onto paper documents) has emerged in recent years (e.g. the Anoto Pen from Anoto AB61) and could potentially be integrated into the Magic Touch system.

**PVA Manipulation Tracking**

To implement a component that tracks virtual activities would provide modelling symmetry and potential synergy effects.

### 10.12 Further reading

The presentation in this chapter has omitted many details about the system. I encourage the interested reader to study the set of master theses which have been produced in direct connection to Magic Touch. (All but the first one are in Swedish however.) Complete references can be found in the reference section. The theses in order of appearance:

- A Wearable Identification System for Physical-Virtual Artefacts (Saathoff, 2000).
- VR-baserat interface för realtidsavbildning av fysiska miljöer över Internet (Granholm & Hägglund, 2001).
- Design och implementation av klientserverarkitektur för distribuerad uppdatering och åtkomst av fysisk-virtuella artefakter (Landfors, 2001).
- Analys och design av ClassiFire — en komponent till Magic Touch-systemtet för automatisk klassificering av kontorstypiska fysisk-virtuella artefakter (Salmaso, 2002).
- Implementation av aktivitetsspärningsklient samt konceptuell design av nästa generations Magic Touch (Jonsson, 2002).

10.13 Curios: Magic Touch demonstrations and media appearances

Magic Touch has been presented in public media as a general system that can help people keep track of things.

**Media coverage after first demo in HUMlab, Umeå university and press release June 20 2000**
- Swedish national TV: SVT: Rapport, Rapport Morgon, Landet Runt, Hjärnkontoret
- Swedish regional TV: TV4: TV4 Botnia, Nordnytt
- Swedish national radio: SR: Vetenskapsradion P1
- Norwegian national radio (NRK)
- Swedish national print media: Svenska Dagbladet, Aftonbladet
- Swedish regional print media: Västerbottenskuriren, Västerbottens Folkblad, Metro Göteborg
- Magazines: Vetenskap — forum för svensk forskning

**Demonstrations at other occasions**
- Tekniska Mässan, trade fair, Stockholm, Sweden, 2000
- Science & Cyber, trade fair, Luleå, Sweden, 2001
Conclusions

"Filmmaking is not about the tiny details. It’s about the big picture!"\(^{62}\)

11.1 Summary

This dissertation has defined a set of concepts, and suggested a general design approach, for bridging the gap between the physical and the virtual world. It has explored the possibility of viewing the physical and the virtual world as, in many ways, similar to each other. Worlds that could (to some extent) and should (under certain conditions) be tighter linked to each other using information technology. The presented work challenges the common view of Human-Computer Interaction as a research discipline mainly dealing with the design of “user interfaces” by proposing an alternative view, a physical-virtual design perspective, in which the focus is on the design of physical-virtual artefacts. One motive for adopting this view is the observation that people well acquainted with specific (physical and virtual) environments are more concerned with the manipulation of (physical and virtual) objects themselves than the user interface through which they are accessed. Another motivation is that such a design perspective promotes the construction of systems and artefacts that could relieve

\(^{62}\) Lines served by actor Johnny Depp playing the character of passionate low-budget movie director Edward D. Wood Jr. (1924-1978) in Tim Burton’s biographical motion picture “Ed Wood” from 1994. Mr. Wood reassures the sceptical film sponsors that, although the cardboard headstone accidentally tipped over in the last shooting, the overall credibility of the graveyard scene and the picture as a whole is not affected. The resulting movie became his probably most well-known piece of art, “Plan 9 From Outer Space” (1959), viewed by many cineasts as one of the worst Hollywood movies ever made. Maybe details matter after all?
people of one of the overhead costs involved when performing physical-virtual activities, namely the cost of keeping related physical and virtual objects synchronised. Since many objects today tend to be manifested in both worlds, mechanisms for automatic synchronisation and identification would have a significant positive impact on the performance of physical-virtual activities.

The dissertation started out with a description of the societal and general research context. The section describing representative work in five relevant HCI research areas is especially motivated since these areas are so new that a general consensus as to what their major concerns are, or rather what their major concerns are not (!), has still to emerge.

The subsequent chapter presented fundamental definitions central to the physical-virtual design perspective: The distinction between the physical and the virtual world, as well as the common view on how human activity take place in these worlds. The crucial step is the abstraction of input devices, allowing the modelling of human-object interaction in the virtual world not very different from in the physical world. Based on this modelling “trick”, the concept of physical-virtual activity was defined for describing the increasingly common situation of people switching between physical and virtual environments in the middle of activities. As a way to facilitate this behaviour, the idea of designing physical-virtual artefacts is proposed and elaborated on.

In chapter 6, the concept of physical-virtual environment gap was defined and subsequently analysed along four dimensions. The results of an empirical study was also presented, indicating the presence of the gap in a telecommunications enterprise. Various ways of bridging the gap (both as an environment designer and as a person acting in the environment) were discussed, including the main strategy proposed and investigated in this dissertation, i.e. that of designing synchronised physical-virtual artefacts.

In order to somewhat frame the design space for the construction of more integrated physical-virtual environments, a simple comparison between structural properties of the physical and the virtual world was performed in chapter 7. Adopting the physical-virtual design perspective, a model of simple human activity in physical-virtual space was then proposed, based on the assumption that objects enter and leave the (at a specific point in time) observable part of physical and virtual environments in much the same way. The concepts of intra- and extra-manipulation were also introduced to distinguish between object manipulation altering the internal state of an object from object manipulation that primarily changes an object’s relation to other objects respectively. In the subsequent chapter (chapter 9), another physical-virtual model was developed and presented based on structural properties of the two worlds. Specifically, the everyday concept of “containment” was used to describe physical and virtual environments, allowing for visualisation of them as joint physical-virtual hierarchies.
Chapter 10 presented the prototype system Magic Touch which has served as a major source of inspiration for the theoretical constructs and discussions in the preceding chapters of the dissertation. The chapter highlights and exemplifies various design challenges and trade-offs faced when implementing infrastructure for the support of physical-virtual activities. It also serves as a kind of empirical proof that containment is indeed a powerful concept when modelling physical environments. Because a new distributed architecture of the Magic Touch is still under construction, the older “monolithic” version was given most attention although some features of the new version were also briefly discussed.

11.2 Contributions

- Definition of the physical-virtual environment gap
- Definition of physical-virtual activities
- Definition of synchronised physical-virtual artefacts
- A physical-virtual design perspective
- A situative model of physical-virtual space
- A hierarchical model of physical-virtual space

A more extensive version of the list above can be found in the introduction chapter (page 4), containing also short descriptions and mentioning also additional work.

11.3 Limitations

Lack of empirical evidence for the utility of the framework as a tool for design

The overall goal with the presented work is to facilitate the design of better physical-virtual environments. Evidence showing that this is actually the case is largely lacking. The only indication that the presented work has some qualities to that end is the fact that it has been developed in conjunction with an existing physical-virtual prototype infrastructure and is able to model basic human activity in that specific physical-virtual environment. It is reasonable to believe that adjustments and additions have to be made in order to make the physical-virtual design framework proposed in this dissertation applicable in practical design activities.

Weak empirical basis for the concepts

Some of the concepts defined in this thesis are based on my own belief of how people in general view the world. For instance, the assumption that “everyday containment” is used as a cognitive tool for people to conceptualise relationships between objects in the virtual world. Empirical proof for assumptions such as these would of course strengthen the credibility of the proposed design framework.
Not grounded on detailed task analysis of for instance knowledge work

The infrastructural support for bridging the gap is designed on the assumption that such support would have a general positive effect on physical-virtual activities performed in physical-virtual environments. How a physical-virtual infrastructure actually helps in specific real cases, or even that it actually does help, is not shown or validated. Thus, the design approach relies on presented logical reasoning and the intuition that such support would be beneficial.63

Physical-virtual artefacts are hard to build

Although the explosive development of sensor and actuation technology (or at least the interest of using it for the purpose of HCI) is a driving force behind research such as this, the challenges in building synchronised physical-virtual artefacts (PVAs), or physical-virtual infrastructure in general, are immense (see page 97 for a brief discussion). On the other hand I believe that basic research in HCI at least sometimes should point out potentially valuable directions for technology development and not the other way around which is usually the case. I also believe that the concept of physical-virtual artefact is a useful tool in modelling also environments with little or no infrastructural support for synchronising PVA manifestations, because large amounts of physical and virtual objects will continue to be related from a human perspective, disregarding if linkage mechanisms are present or not.

11.4 Future work

From a conceptual framework to a workable design methodology

One of the first activities has to be to evaluate the proposed concepts in practice and to find a method how to apply the framework systematically in design and for analysis purposes. This would include

- A survey of common existing physical-virtual activities, particularly in settings outside office environments (whose physical-virtual activity status is quite well known, e.g. the typically physical-virtual activity of authoring long text documents).
- Further development of PVA infrastructure and PVA prototypes. This would eventually lead to a more systematic application of the conceptual framework as a tool for design.

63. An extenuating circumstance is however that this lack of “hard” empirical evidence proving the utility is very common in the present line of research in Ubiquitous/Pervasive Computing. Perhaps it is necessary for this kind of basic research to be positive, having faith in technology as the solution to many problems in modern life?
• Development of metrics for measuring the size of the physical-virtual environment gap. Physical-virtual metrics could, as most metrics, be used in both design and analysis situations. The ultimate goal would be to be able to rate environments as to how well they are suited for specific physical-virtual activities.

Further development of the conceptual framework
Apart from changes and additions to the physical-virtual design framework that will probably emerge from assessing the framework in the way described in the previous section, room for improvement known today includes:

• Taking into account “nomadic” physical-virtual environments. Parts of the conceptual framework is based on the assumption that the modelled physical-virtual activities take place in a room, a building, or in another spatially limited and geographically fixed place. Wearable computers “always on, always with you” might call for adjustments or interesting additions to some of the proposed theoretical constructs. For instance, both physical and virtual sub-hierarchies in a specific hierarchical model of physical-virtual space would become very dynamic as users of wearable computers continuously change physical location and degree of connectivity to the Internet and other networks. Location-based services such as allowing people to attach virtual notes to physical places (“GeoNotes”; Espinoza, Persson, Sandin, Nyström, Cacciatore & Bylund, 2001), would probably fit into the proposed model without causing significant changes.64

• Multi-user physical-virtual activity. At this time, except for the brief discussion on distributed PVAs in chapter 5 (starting on page 101), the framework is mainly focused on single-user activity in physical-virtual space. A development of the framework in such a direction would obviously make use of previous work in areas such as Computer-Supported Cooperative Work (CSCW) and Telepresence.

• Version handling and temporary PVAs. Some PVAs have longer lives than others. In particular, some PVA manifestations are useful only during a very short time-period (e.g. a printed paper draft of a document in the context of rewriting it) compared to their corresponding manifestations in the other environment (e.g. the corresponding digital document which remains the “same” document for a longer time period). Version handling of PVAs and PVA manifestation has not been addressed so far.

64 One of the most valuable properties of the physical-virtual design perspective is the way it invites to “what-if” scenario-thinking based on physical-virtual symmetry (which it does not necessarily promote as an ideal case). In the case of GeoNotes, what if we constructed an inverse system that allowed for the placing of physical objects in virtual places instead? Is it possible? Would it be of any use? In that case, under what conditions?
• A general protocol or language that allows end-users to specify how they want the causal link between PVA manifestations to behave. This is an idea stemming directly from the development of the Magic Touch system.

• Further development of the hierarchical model of physical-virtual space. In particular it would be interesting to investigate how other structural properties of space (other than “containment”) afford and constrain physical-virtual activities.

• Further development of a hierarchical model of PVAs and PVA manifestations. The idea of modelling PVA manifestations as PVAs themselves opens up for new modelling challenges and possibilities. Challenges lie in the counter-intuitiveness of virtual objects having physical objects as their constituents; possibilities lie in the potential integration of the hierarchical physical-virtual space model and “hierarchical” PVAs.

• Development of general PVA design guidelines. Since the design of PVAs is a rather new endeavour, design guidelines would be something to strive for. In particular, how do we design corresponding PVA manifestations in a way so that they tell people that they belong together? How important factor is visual (or generally sensuous) correlation of physical and virtual, and specifically visual co-location, compared to temporal correlation and other perceived causal relations?

• Physical environments making magic. It would be interesting to investigate the possibility of using professional conjurer tricks (as a complement to the standard information technology approach) in order to create physical-virtual environments in which some laws of nature could be given the illusion to have been altered or twisted in order to expand the action space of physical-virtual environments in general, and physical environments’ conformity with the virtual world in particular. Specifically, relaxing physical world-constraints on containment would be very interesting. How far can you go? If these magicians can make whole human beings escape from seemingly closed boxes, why can’t we create systems that make paper documents vanish? Dexterous physical environments fooling everyone! Software designers have created virtual environments fooling millions of people that they have a desktop inside their computer screen! Isn’t it about time for physical-world architects, designers and why not physicists, to show what they can do? As an example, attempts to construct systems that trigger human senses to persuade people about the state of affairs, or to create stronger experiences, have been tried in a physical-virtual context by Koleva, Schnädelbach, Benford & Greenhalgh (2000).
11.5 Closing remarks

The presented work is to be regarded as initial work pointing out a direction of research, more than as a fully completed study. I hope that I through this dissertation have provided enough sound arguments for designing environments with focus on support for human physical-virtual activity, by adopting a physical-virtual design perspective, and by designing infrastructure for physical-virtual artefacts.

"Anything could be true. The so-called laws of Nature were nonsense. The law of gravity was nonsense. ‘If I wished,’ O’Brien had said, ‘I could float off this floor like a soap bubble.’ Winston worked it out. ‘If he thinks he floats off the floor, and if I simultaneously think I see him do it, then the thing happens.’"

(George Orwell, Nineteen Eighty-Four, 1948)
Appendix A. English version of questionnaire (chapter 6)

Inquiry on knowledge work and work environment

### Personal data

<table>
<thead>
<tr>
<th>[1/32] Age</th>
<th>16-20</th>
<th>21-25</th>
<th>26-30</th>
<th>31-35</th>
<th>36-45</th>
<th>46-55</th>
<th>56-64</th>
<th>65-</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>[2/32] Sex</th>
<th>female</th>
<th>male</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>[4/32] If you have performed academical studies at a university, state academical degree and achieved number of points.(^a)</th>
<th>16-20</th>
<th>21-25</th>
<th>26-30</th>
<th>31-35</th>
<th>36-45</th>
<th>46-55</th>
<th>56-64</th>
<th>65-</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^a\) In Sweden, higher academical studies are measured in points where each point corresponds to one full-time week of studies. For instance, BSc and MSc degrees usually requires 120 and 160 points respectively.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Work tasks

[6/32] What is your official position within the company?
[free text open question]

[7/32] Mark the professions (one or more) you think best describes your work tasks.

<table>
<thead>
<tr>
<th>Profession</th>
</tr>
</thead>
<tbody>
<tr>
<td>librarian / information manager</td>
</tr>
<tr>
<td>executive</td>
</tr>
<tr>
<td>controller</td>
</tr>
<tr>
<td>researcher/scientist</td>
</tr>
<tr>
<td>help desk / support</td>
</tr>
<tr>
<td>informant</td>
</tr>
<tr>
<td>purchaser</td>
</tr>
<tr>
<td>journalist</td>
</tr>
<tr>
<td>consultant</td>
</tr>
<tr>
<td>line man</td>
</tr>
<tr>
<td>medical doctor</td>
</tr>
<tr>
<td>information broker</td>
</tr>
<tr>
<td>planner</td>
</tr>
<tr>
<td>public relations officer</td>
</tr>
<tr>
<td>product developer</td>
</tr>
<tr>
<td>project leader</td>
</tr>
<tr>
<td>repairman</td>
</tr>
<tr>
<td>coordinator</td>
</tr>
<tr>
<td>secretary</td>
</tr>
<tr>
<td>superintendent</td>
</tr>
<tr>
<td>service / maintenance</td>
</tr>
<tr>
<td>seller</td>
</tr>
<tr>
<td>teacher / instructor</td>
</tr>
<tr>
<td>other</td>
</tr>
</tbody>
</table>

[8/32] Mark the 10 most important skills that you think are needed in order to perform your work tasks.

<table>
<thead>
<tr>
<th>Skill</th>
</tr>
</thead>
<tbody>
<tr>
<td>analytical</td>
</tr>
<tr>
<td>power of initiative</td>
</tr>
<tr>
<td>creative</td>
</tr>
<tr>
<td>linguistic</td>
</tr>
<tr>
<td>empathic</td>
</tr>
</tbody>
</table>
**[8/32]** Mark the 10 most important skills that you think are needed in order to perform your work tasks.

<table>
<thead>
<tr>
<th>Skill</th>
<th>Marked</th>
</tr>
</thead>
<tbody>
<tr>
<td>social (EQ)</td>
<td></td>
</tr>
<tr>
<td>authority</td>
<td></td>
</tr>
<tr>
<td>physiological</td>
<td></td>
</tr>
<tr>
<td>ability to plan</td>
<td></td>
</tr>
<tr>
<td>ability to generalise</td>
<td></td>
</tr>
<tr>
<td>oratoric</td>
<td></td>
</tr>
<tr>
<td>diplomatic</td>
<td></td>
</tr>
<tr>
<td>ability to sell</td>
<td></td>
</tr>
<tr>
<td>mediation</td>
<td></td>
</tr>
<tr>
<td>ability to manage stress</td>
<td></td>
</tr>
<tr>
<td>beauty</td>
<td></td>
</tr>
<tr>
<td>critical thinking</td>
<td></td>
</tr>
<tr>
<td>ability to come to decisions</td>
<td></td>
</tr>
<tr>
<td>ability to learn</td>
<td></td>
</tr>
<tr>
<td>ability to convince</td>
<td></td>
</tr>
<tr>
<td>ability to perform many tasks simultaneously</td>
<td></td>
</tr>
<tr>
<td>pedagogical</td>
<td></td>
</tr>
<tr>
<td>openness for new ideas</td>
<td></td>
</tr>
<tr>
<td>ability to inspire others</td>
<td></td>
</tr>
<tr>
<td>ability to establish and maintain contacts</td>
<td></td>
</tr>
<tr>
<td>to maintain ones own strong motivation</td>
<td></td>
</tr>
<tr>
<td>to be intuitive</td>
<td></td>
</tr>
<tr>
<td>to work independently</td>
<td></td>
</tr>
<tr>
<td>to be able to concretise</td>
<td></td>
</tr>
<tr>
<td>to cooperate</td>
<td></td>
</tr>
<tr>
<td>to be able to think in abstract terms</td>
<td></td>
</tr>
<tr>
<td>punctuality</td>
<td></td>
</tr>
<tr>
<td>multi-linguality</td>
<td></td>
</tr>
<tr>
<td>problem-solving</td>
<td></td>
</tr>
<tr>
<td>to be able to define and find problems</td>
<td></td>
</tr>
<tr>
<td>ability to communicate</td>
<td></td>
</tr>
<tr>
<td>openness for new technology</td>
<td>[free text]</td>
</tr>
<tr>
<td>other</td>
<td></td>
</tr>
</tbody>
</table>

**[9/32]** Mark the 10 most important skills that you think best describes your own personal qualities.

<table>
<thead>
<tr>
<th>Skill</th>
<th>Marked</th>
</tr>
</thead>
<tbody>
<tr>
<td>analytical</td>
<td></td>
</tr>
<tr>
<td>power of initiative</td>
<td></td>
</tr>
<tr>
<td>creative</td>
<td></td>
</tr>
<tr>
<td>linguistic</td>
<td></td>
</tr>
</tbody>
</table>
Mark the 10 most important skills that you think best describes your own personal qualities.

<table>
<thead>
<tr>
<th>Skill</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>empathic</td>
<td></td>
</tr>
<tr>
<td>social (EQ)</td>
<td></td>
</tr>
<tr>
<td>authority</td>
<td></td>
</tr>
<tr>
<td>physiological</td>
<td></td>
</tr>
<tr>
<td>ability to plan</td>
<td></td>
</tr>
<tr>
<td>ability to generalise</td>
<td></td>
</tr>
<tr>
<td>oratoric</td>
<td></td>
</tr>
<tr>
<td>diplomatic</td>
<td></td>
</tr>
<tr>
<td>ability to sell</td>
<td></td>
</tr>
<tr>
<td>mediation</td>
<td></td>
</tr>
<tr>
<td>ability to manage stress</td>
<td></td>
</tr>
<tr>
<td>beauty</td>
<td></td>
</tr>
<tr>
<td>critical thinking</td>
<td></td>
</tr>
<tr>
<td>ability to come to decisions</td>
<td></td>
</tr>
<tr>
<td>ability to learn</td>
<td></td>
</tr>
<tr>
<td>ability to convince</td>
<td></td>
</tr>
<tr>
<td>ability to perform many tasks simultaneously</td>
<td></td>
</tr>
<tr>
<td>pedagogical</td>
<td></td>
</tr>
<tr>
<td>openness for new ideas</td>
<td></td>
</tr>
<tr>
<td>ability to inspire others</td>
<td></td>
</tr>
<tr>
<td>ability to establish and maintain contacts</td>
<td></td>
</tr>
<tr>
<td>to maintain ones own strong motivation</td>
<td></td>
</tr>
<tr>
<td>to be intuitive</td>
<td></td>
</tr>
<tr>
<td>to work independently</td>
<td></td>
</tr>
<tr>
<td>to be able to concretise</td>
<td></td>
</tr>
<tr>
<td>to cooperate</td>
<td></td>
</tr>
<tr>
<td>to be able to think in abstract terms</td>
<td></td>
</tr>
<tr>
<td>punctuality</td>
<td></td>
</tr>
<tr>
<td>multi-linguality</td>
<td></td>
</tr>
<tr>
<td>problem-solving</td>
<td></td>
</tr>
<tr>
<td>to be able to define and find problems</td>
<td></td>
</tr>
<tr>
<td>ability to communicate</td>
<td></td>
</tr>
<tr>
<td>openness for new technology</td>
<td></td>
</tr>
<tr>
<td>other [free text]</td>
<td></td>
</tr>
</tbody>
</table>
### [10/32] How much of your work time do you spend on task that you...

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1-10</th>
<th>11</th>
<th>21</th>
<th>31</th>
<th>41</th>
<th>51</th>
<th>61</th>
<th>71</th>
<th>81</th>
<th>91-100</th>
</tr>
</thead>
<tbody>
<tr>
<td>... think you are employed to do</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>... really want to perform / enjoy doing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### [11/32] How much of your work time do you spend at fix locations (at home, at the office), in meeting rooms, recreational facilities, other places?

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1-10</th>
<th>11</th>
<th>21</th>
<th>31</th>
<th>41</th>
<th>51</th>
<th>61</th>
<th>71</th>
<th>81</th>
<th>91-100</th>
</tr>
</thead>
<tbody>
<tr>
<td>work place at home</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>work place in office building</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>in meeting room</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>travelling</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>at client</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>at recreation facility</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>other place</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>if other place, specify what place:</td>
<td>[free text]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### [12/32] How important are the following factors for you in order to feel comfortable with your work situation?

<table>
<thead>
<tr>
<th>social factors (relaxed informal conversation over a cup of coffee, organised social activities outside working hours, etc.)</th>
<th>very important</th>
<th>important</th>
<th>quite important</th>
<th>not so important</th>
<th>irrelevant</th>
</tr>
</thead>
<tbody>
<tr>
<td>possibility to choose how to solve tasks</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>possibility to choose kind of work tools</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>possibility to choose the degree of computer support in the work</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### [13/32] What other factors do you think affects your comfort at work?
State some factors and their degree of influence below.

[free text open question]
### Computer support at work

**[14/32]** How much of your work time do you spend on changing your work environment/situation? This includes ergonomical improvements, to try new tools (computer software, pens, work methods), to keep your self up-to-date, etc.  

<table>
<thead>
<tr>
<th>0</th>
<th>1-5</th>
<th>6-10</th>
<th>11-15</th>
<th>16-20</th>
<th>21-25</th>
<th>26-30</th>
<th>31-35</th>
<th>36-40</th>
<th>41-45</th>
<th>46-50</th>
<th>51-100</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**[15/32]** How much of your work time do you spend on education and competence development?  

<table>
<thead>
<tr>
<th>0</th>
<th>1-5</th>
<th>6-10</th>
<th>11-15</th>
<th>16-20</th>
<th>21-25</th>
<th>26-30</th>
<th>31-35</th>
<th>36-40</th>
<th>41-45</th>
<th>46-50</th>
<th>51-100</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**[16/32]** How big share of your work tasks do you perform completely without computer support, with some computer support, with large computer support, and using computer only?  

<table>
<thead>
<tr>
<th>0</th>
<th>1-10</th>
<th>11-20</th>
<th>21-30</th>
<th>31-40</th>
<th>41-50</th>
<th>51-60</th>
<th>61-70</th>
<th>71-80</th>
<th>81-90</th>
<th>91-100</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- completely without computer support
- with some computer support
- with large computer support
- using computer only

**[17/32]** Has the amount of computer-supported activities decreased, remained constant, or increased the past five years?  

<table>
<thead>
<tr>
<th>decreased</th>
<th>remained constant</th>
<th>increased</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
To what extent do you think you have the possibility to influence the element of computer support in your work?

<table>
<thead>
<tr>
<th>I have no possibility</th>
<th>I have small possibilities</th>
<th>I have large possibilities</th>
<th>I decide completely by myself</th>
</tr>
</thead>
</table>

In the cases where you have the possibility to influence the degree of computer support, what factors determine your decision? Grade the following factors depending on how important they are for your decision.

<table>
<thead>
<tr>
<th>very important</th>
<th>important</th>
<th>quite important</th>
<th>not so important</th>
<th>irrelevant</th>
</tr>
</thead>
<tbody>
<tr>
<td>personal interest in computers</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>short-term efficiency</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>long-term efficiency</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>the quality of the result</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Do you think there are other important aspects/factors that influences your choice between using/not using computer support? State these factors and their degree of influence below.

[free text open question]

Things used in work activities

Give some examples of material things you use in your work, both general and more tuned towards your specific tasks. State the things and the kind of activity they support.

Examples:
- paper & pen, quick sketches
- pocket diary, time planning
- Post-It notes, memory help
- colour mark-up pen, reading help
- telephone, personal communication
- binders & bookshelf, for archiving
- etc.

[free text open question]
[22/32] Give some examples of computer programs you use in your work, both general and more tuned towards your specific tasks. State the names of the programs and the kind of activity they support.

Examples:
Microsoft Word, word processing
Eudora, to handle e-mail
NetSpy Pro, computer network management
CU-SeeMe, video conferencing
RTF2HTML, text format conversion
etc.

[free text open question]

[23/32] Give some examples of working methods, mnemonic rules, habits etc. that you make use of in your work, both general and more tuned towards your specific tasks. State their names and the kind of activity they support.

Examples:
The three-second rule, to keep the right safety-distance towards the car in front to check in the pocket diary before you go home for the day, to prepare for tomorrow’s work LOKE, mnemonic method
etc.

[free text open question]

[24/32] In the cases where you have the possibility to choose between similar tools, for instance between paper & pen or to use a word processor, what factors determine your decision? Grade the following factors depending on how important they are for your decision.

<table>
<thead>
<tr>
<th>Factor</th>
<th>very important</th>
<th>important</th>
<th>quite important</th>
<th>not so important</th>
<th>irrelevant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Easy to understand the functionality — small or no hidden functionality.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Easy to use on different kinds of material — Example: it is possible to write on most materials using most pens.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Generality with regard to the method — Example: hammers are designed for beating, and can be used for nailing, crushing, dismounting, etc.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Combinability — that it is possible to combine the tool with other tools</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simplicity — it is simple in its construction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**[24/32]** In the cases where you have the possibility to choose between similar tools, for instance between paper & pen or to use a word processor, what factors determine your decision? Grade the following factors depending on how important they are for your decision.

<table>
<thead>
<tr>
<th>Flexibility — that it can be used for other purposes than it is designed for</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Personal interest</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Short-term efficiency</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Long-term efficiency</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The quality of the result</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**[25/32]** Tools are often designed for a specific field of application. In spite of this fact, one can sometimes find other ways in using them. Can you give some examples where you use tool for other purposes than what they originally were designed for?

Example: To use an adjustable spanner as hammer or tyre remover, or to use the cellular phone answering machine as dictaphone.

[free text open question]

**[26/32]** Are there work situations where you can identify a need for new kinds of support or aid (computer-based, material, mental, or combinations of these)? State situations and the kind of support/aid you feel a need for.

[free text open question]

**[27/32]** All individuals have their own view of computers and computer systems. Mark the/those alternative(s) that best describes your of computers and computer systems. The computer is for me a/an...

- automations
- communication channel
- coordinator
- expert
- fellow-worker
- game
- gold-mine
- knowledge bank
- oracle
- person
- person amplifier
- social actor
- library
Sources of information

How important are the following information sources for fulfilling your information need in your work?

<table>
<thead>
<tr>
<th>Source</th>
<th>very important</th>
<th>important</th>
<th>quite important</th>
<th>not so important</th>
<th>irrelevant</th>
</tr>
</thead>
<tbody>
<tr>
<td>colleagues</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>other personal contacts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>daily papers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>journals</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>books</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>newsgroups</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e-mailing lists</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>World-Wide-Web (WWW)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>libraries</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>manuals</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>professional journals</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>customers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>house magazines</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lunch rooms</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>staff cafeterias</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>other sources</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>if other sources, state what.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[free text]</td>
</tr>
</tbody>
</table>

All individuals have their own view of computers and computer systems. Mark the/those alternative(s) that best describes your of computers and computer systems. The computer is for me a/an...

typewriter
calculator
organiser
tool
other [free text]
**To use or not to use computer**

**[29/32]** Grade the following factors based on how crucial they are when you choose to use or not use an information source.

<table>
<thead>
<tr>
<th>Factor</th>
<th>very important</th>
<th>important</th>
<th>quite important</th>
<th>not so important</th>
<th>irrelevant</th>
</tr>
</thead>
<tbody>
<tr>
<td>time it takes to access/get the information</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>the language (swedish, english, spanish, etc.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>economical cost</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>language complexity (mathematical text, paper articles etc.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>short-term information need</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>long-term information need</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>your own experience of using the information source</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>information authority (author, the forum’s reputation, etc.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In some situations both material and computer-based tools are used in order to perform parts of a continuous work task, potentially leading to problems since the two environments don’t “speak the same language”.

**Example 1:** You have made notes on paper (approximately two A4 pages) during a meeting and would like to send them to a colleague using e-mail. In order to transfer the notes to the computer environment you need to rewrite them, taking up extra time.

<table>
<thead>
<tr>
<th>Do you recognize the situation?</th>
<th>yes, very much</th>
<th>yes, slightly</th>
<th>no</th>
</tr>
</thead>
</table>

How do you handle the situation?
— I choose to:

- type the notes into the computer
- skip notifying the colleague
- fax the notes
- mail the notes
- notify the colleague using phone
- notify the colleague in another way [free text]
In some situations both material and computer-based tools are used in order to perform parts of a continuous work task, potentially leading to problems since the two environments don’t “speak the same language”.

Example 2: You have received a new important document from a remote colleague by e-mail, and you have the task to read it through (approximately 10 A4 pages) and give comments. Since it is quite urgent it would be good if you could send the document back together with your comments.

<table>
<thead>
<tr>
<th>Do you recognize the situation?</th>
<th>Yes, very much</th>
<th>Yes, slightly</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>How do you handle the situation?</td>
<td>For read-ergonomical reasons you print the documents onto paper. You write your comments by hand directly on the paper and fax the revised document to your colleague.</td>
<td>For read-ergonomical reasons you print the documents onto paper. You write your comments by hand directly on the paper and transfer your hand-written comments to the electronic document and send it using electronic mail.</td>
<td>For read-ergonomical reasons you print the documents onto paper. You sit close to a computer when reading the paper document, write your comments directly into the electronic document and send it using electronic mail.</td>
</tr>
<tr>
<td></td>
<td>You read and comment the document directly in your word processing application and send the revised document using electronic mail.</td>
<td>I solve it in another way.</td>
<td>[free text]</td>
</tr>
</tbody>
</table>

How big problem do you think the information environment “gap” described in the examples above, between material and computer-based environments, is for you in your work?

<table>
<thead>
<tr>
<th>No problem</th>
<th>Small problem</th>
<th>Quite big problem</th>
<th>Big problem</th>
</tr>
</thead>
</table>

How much of your work time do you estimate that you spend on bridging the “gap” between the material and the computer-based environment?

<table>
<thead>
<tr>
<th>0</th>
<th>1-5</th>
<th>6-10</th>
<th>11-15</th>
<th>16-20</th>
<th>21-25</th>
<th>26-30</th>
<th>31-35</th>
<th>36-40</th>
<th>41-45</th>
<th>46-50</th>
<th>51-55</th>
</tr>
</thead>
</table>

[30/32]

[31/32]

[32/32]
Comments on the questionnaire

We are grateful for comments on the questionnaire. Is some question hard to understand? How long time did it take for you to fill it in? Etc.

[free text open question]
Index

Augmented/Mixed Reality ........................................................................................................24
containment ............................................................................................................................184
cloned objects .......................................................................................................................199
Context(ual) Awareness ........................................................................................................31
Graspable/Tangible User Interfaces .....................................................................................28
inter-manipulation ................................................................................................................179
inter-world action ...................................................................................................................81
inter-world event mediator (IWEM) ....................................................................................81, 91, 136
intra-manipulation ................................................................................................................179
intra-world action ...................................................................................................................81
manipulable physical-virtual subspace ..............................................................................175, 178
observable physical-virtual subspace ..............................................................................175, 178
physical activity ....................................................................................................................84
physical containment ..........................................................................................................189
physical space .......................................................................................................................189
physical world .......................................................................................................................75
physical-virtual action pair ..................................................................................................90
physical-virtual activity ........................................................................................................93
physical-virtual activity overhead ......................................................................................108
physical-virtual artefact (PVA) ..........................................................................................94, 90, 132
physical-virtual design perspective ....................................................................................166
physical-virtual environment gap ....................................................................................107
physical-virtual mode error .................................................................................................126
physical-virtual world space ..............................................................................................174, 178
space ..................................................................................................................................170
Ubiquitous/Pervasive Computing ......................................................................................26
virtual activity .......................................................................................................................87
virtual containment ...............................................................................................................195
virtual environment provider ..............................................................................................88
virtual space .........................................................................................................................195
virtual world .........................................................................................................................76
Wearable Computing ...........................................................................................................29
References


Mann, S. (1997). An historical account of the ‘WearComp’ and ‘WearCam’ inventions developed for applications in ‘Personal Imaging’. In Proceedings of Interna-


