Final thesis

Designing for Collaborative Turn-Taking at the Digital Tabletop

in cooperation with Santa Anna IT Research Institute AB

by

Jonas Rybing

LIU-IDA-EX--2011/KOGVET-A–11/006--SE

2011-10-24
Designing for Collaborative Turn-Taking at the Digital Tabletop

by

Jonas Rybing


Supervisor : Dr. Magnus Bång  
Dept. of Computer and Information Science  
at Linköping University

Examiner : Prof. Arne Jönsson  
Dept. of Computer and Information Science  
at Linköpings universitet
Abstract

Collaboration technologies are difficult to design due to the complex myriad of social, cognitive, and communicative aspects of group interactions. New interaction technologies like multitouch sharable interfaces, such as digital tabletops, have lead to a renewed interest in designing collaborative technologies. This thesis focuses on turn-taking protocols as a coordinating mechanism during collaborative work with digital tabletops. The goal was to develop new conceptual designs and interactive mechanisms to support face-to-face collaborations of small groups. Inspired by ethnographical studies of collaborative work and theories in distributed cognition and related theories of language and action a model of collaborative turn-taking was developed. Moreover, the thesis presents five design concepts and interaction components for the digital tabletop that exemplifies the different properties of the model.

Keywords: Collaboration, turn-taking, distributed cognition, digital tabletops, design, communication, coordination
Acknowledgements

There are many people who helped me throughout the doings and writings of this thesis. First off, I would like to thank Yvonne Rogers, Paul Marshall, Per Linell, and Tomas Timpka for taking valuable time from your busy schedule in order to answer my questions. Secondly, I would like to thank everyone I’ve been in contact with during the ethnographical fieldwork, especially the very helpful staff at the ICU - Katarina Solnevik, Nicolas Wyon and Hanna Andersson among others.

Of course big thanks go out to all my fellow students, especially Robin Keskisärkkä, Christian Smith and Tomas Lundqvist for our many productive discussions. Special thanks to my instructor and supervisor Magnus Bång for all your guidance and help. Last and not least I would like to thank the Santa Anna Research Institute and Sture Hägglund for providing me with the possibility to write this thesis!

Thank you all!
Contents
Chapter 1

Introduction

The traditional single-user computer has become an indispensable tool for the many activities we are engaged in on a daily basis. We use them to find information, do bank errands, book our vacation and so on. The traditional ideas of window layout, mouse and keyboard interaction, and hierarchical navigation of file systems has set the standard for how we use and interact with the digital world (Lindell, 2009). This standard is however changing and new technological products are exploring different avenues of interaction. The Apple iPhone gave rise to a wave of so called “smartphones” which use direct touch interaction on a digital surface. Bigger horizontal table versions, such as the Microsoft Surface table, also use direct touch interaction, but are aimed towards groups interacting rather than single user interaction. By using digital tabletops and augmented reality techniques, the boundaries between the physical (real) world and the digital world (virtual) have started to be erased (Keller, 2006). These are all examples of alternative user interfaces that are becoming more and more popular, and they are most likely here to stay.

The rise of collaborative technology is due to the lack of support for collaborative processes and activities provided by traditional single-user technology (Grudin, 1991, 1994; Ellis et al., 1991). For instance, in the health-care domain many researchers have criticized the current approach of single-user computerization, and said that it may impose unnatural and
inefficient ways of working, and may add to the workload of these professionals (Berg, 1998; Sellen and Harper, 2003; Heath and Luff, 2000). In order to support collaboration with technology the field of Computer Supported Cooperative Work (CSCW) has evolved (Grudin, 1991; Ellis et al., 1991). Perhaps the biggest challenges for CSCW research is how to best understand collaborative work in order to aid design, and how best to design a collaborative systems for the purposes of supporting collaborative work (Fitzpatrick, 1998). Fitzpatrick phrases this challenge as “CSCW design is a ’wicked problem’, where there are no definitive solutions”.

Using traditional workspaces (fitted with tables, chairs, paper, pens and so forth) have long been considered the “ideal” collaboration environment (Scott et al., 2004). One can see why, large horizontal surfaces provides means for the spreading, piling, and organizing documents and sitting around a table is comfortable and relaxing. Tabletops workspaces provide high degrees of awareness and further use orientation and placement of objects to help mediate group interactions (Tang, 1991; Kruger et al., 2003; Scott et al., 2004). Therefore, many CSCW researchers have relied on them to inform the design on digital and virtual collaborative workspaces. One design issue is how digital tabletops can maintain the characteristics of traditional tables useful for collaboration.

The digital tabletop is a relatively new technology for groups interacting collaboratively, although several systems have already been developed and the first generation wave of digital tabletops has washed over us, with much research on both hardware components and software libraries (Morris, 2006; Lindell, 2009). Among these systems are the UbiTable (Shen et al., 2003), the Lumisight table (Matsushita et al., 2004), and the DiamondTouch(Dietz and Leigh, 2001) with its belonging software library DiamondSpin. These systems have taken some important steps to support and provide coordination among the collaborators and structuring of collaborative work, for instance by identifying what interactions belongs to which participant and providing support for private information. However, as the second wave of digital tabletops approaches (including the second generation of Microsoft Surface tabletops) these and other mechanics needs to be further developed.

Related and common for all groupware systems is the design issue of how turn-taking procedures should be managed (Shu and Flowers, 1992).
Shu and Flowers asks questions of whether everyone should be able to manipulate the objects simultaneously or should they have to take turns; if so, what would be an appropriate protocol for taking turns? In their studies, a simultaneous approach supported by independent points of view proved to be an appropriate solution. However, in digital tabletop studies conflicts and collaborative breakdowns have occurred when simultaneous and parallel interactions were used (Morris et al., 2004; Olson et al., 2011; Stewart et al., 1999).

In order to support co-located (face-to-face) collaborative turn-taking this thesis provides five design concepts for how turn-taking and collaborative work can be structured and coordinated through interaction with a digital tabletop. The designs are based on cognitive and linguistic theory and inspired by previous work with digital tabletops together with ethnographical data. During the course of the thesis an inspirational perspective of turn-taking was developed; illustrating that collaborative turn-taking can be viewed on different levels and how these levels provide different coordinating mechanisms. A simple classification of collaborative tabletop activities was also made to inform and guide the design of coordinating mechanisms.

1.1 Aims and Outline

This thesis takes a closer look at surface interaction technology supporting small group collaboration. The aim of the thesis is to generate theoretically grounded interaction component concepts for digital tabletops, with the purpose to support turn-taking protocols and coordination of collaborative work. The focus of this study is on turn-taking protocols as a coordinating mechanism for collaboration at a digital tabletop. The two main domains in focus are examples of where collaborative technologies may be used in the near future; the medical domain and the personal service domain (e.g. customer meetings). Further, in order to create interaction concepts for these domains an abductive research approach was used, where inspiration was sought from studying the current domain settings and theory.

To explore how collaboration and turn-taking operates within the current domain settings, shorter ethnographic field studies were performed;
and results were interpreted from a theory driven perspective. A model on collaborative turn-taking and collaborative activities was built up, inspired by the results from the ethnography and theories on distributed cognition, language and action, and the dialogical theory of communication. Finally, five general design concepts were created, grounded in the perspective on collaborative turn-taking and collaborative activities related to digital tabletops.

This thesis contributes to the design of future tabletop applications by considering important aspects for creating co-located collaborative technology (such as the nature of the collaborative activity) and considerations of tabletop interaction (e.g. how to use gestures and personal spaces). The thesis also contributes by providing five suggestions on how turn-taking protocols can be supported through the interface and through interaction with the digital tabletop surface. Further, the thesis provides a model and perspective on collaborative turn-taking based on empirical studies, previous works and theories.

This thesis is part of a project on user interaction with modern technology directed by Santa Anna Research Institute AB.

1.2 Delimitations

The suggested interface components and mechanics are not to be mistaken for detailed designs of complete tabletop applications. Rather, they are to be considered as concepts and notions on how these issues can be addressed through surface interaction and interface layout in co-located collaboration. The concept designs follow general principles and guidelines for design and tabletop interaction, but no formal evaluation of the concepts has been performed. The empirical studies performed are to be viewed as inspirational rather than analytical, as no strict theoretical analysis of the data was carried out. The model presented on collaborative turn-taking and collaborative activities is not supposed to be complete and mutually exclusive classifications or framework for turn-taking or activities in general, but a space of interpretation and a source of inspiration for designing collaborative mechanisms for a digital tabletop.
Chapter 2

Background

The background chapter will describe related scientific theories and introduce previous work on digital tabletops. The first section focuses on linguistic and cognitive theory related to groups of people working together; the section also introduces important terms such as groups, collaboration and coordination. The section will look closer at three theories; the dialogical theory of communication as put forth by Per Linell 1, the distributed cognition theory by Edwin Hutchins, and the theories of language and action by Herbert Clark. These theories can be seen to have a relation to each other, since they have ideas and themes in common. This will provide a base for the rest of the background, and in extent the rest of the thesis. The next section of the background moves on to introduce how technology can support collaboration and group activities. This section highlights work done with digital tabletops. The last section of the background will discuss the use of theory in design and lay down some principles and guidelines for designing digital tabletops and tabletop applications.

---

1Ideas traceable to the philosopher Martin Buber
2.1 Collaborative Work

“Unity is strength... when there is teamwork and collaboration, wonderful things can be achieved.” - Mattie Stepanek

Central to this thesis is the notion of two or more people working together towards a shared goal. This process of joint work is often referred to as collaboration or cooperation. Most of us have an intuitive feeling or assumption of what it means to collaborate or cooperate. Even though it may seem clear what this process is, a definition is in place. As Joseph McGrath and Andrea Hollingshead puts it: “Collaborative work in groups is a complex matter: with or without electronic technology” (McGrath and Hollingshead, 1993).

Before we try to get to grips with the notions of collaboration and cooperation, let’s first address a sometimes troublesome concept which is central when discussing group work: groups. There are many definitions of what a group is in the literature. One debated issue is whether dyads (two persons) are to be regarded as groups (Moreland, 2010; Williams, 2010). According to Stefan Jern (1998), “group” is a generic term, including many social activities with varying goals, contexts and preconditions. A certain type of groups is work groups. Jern defines work groups as a number of people who: 1) interacts with each other, 2) are psychologically aware of each other, 3) perceive themselves as a group, and 4) the group fulfills a formal function within an organization. I will use the term group rather pragmatically; including dyads and at least the first three of Jern’s characteristics of a work group - the fourth will probably apply for most situations. Another commonly used term is small-groups, often used for groups including up to 10 or 12 group members.

With the criteria for what a group is established, let’s now look at the processes of collaboration and cooperation. First, what is the difference between the two terms? According to Dillenbourg et al. (1996), collaboration is as a joint coordinated and synchronous process of solving some problem. Cooperation on the other hand involves an individual solving parts of the problem followed by a synthesis with the rest of the group. Further, according to their definition, cooperation may also be asynchronous. Henceforth I will use the term collaboration, but in a wide sense. The focus
still lies on joint coordinated synchronous processes, but I will not exclude the processes attributed to cooperation.

With key terms sorted out, let’s continue to look at what collaboration actually is. Cooperative and collaborative work is inherently distributed (Schmidt and Simone, 1996). Edwin Hutchins argues, from the perspective of distributed cognition, that there are many important aspects of any distributed social task (Hutchins, 1995). Among these aspects are concepts such as common ground, the nature of the activity, social relationships among the participants, the environment, and more. Perhaps the most important mechanic of collaboration is coordination among participants and between them and their surroundings. Coordination takes place all around us during a social task; among the individuals, between individuals and artifacts, between artifacts, and with the environment. Hutchins (1995) emphasizes that the knowledge required to coordinate actions is not all discretely contained inside the participating individuals; rather, much of it is intersubjectively shared among the members of the activity. In this way the human component acts as a sort of malleable and adaptable coordinating tissue. All labor, whether considered physical or cognitive in nature require distributed cognition in order to coordinate each participant’s effort. In this view even a simple system of two men driving a spike with hammers requires some cognition on the part of each to coordinate the collaboration.

One aspect of coordination involves deciding whose turn it is to act or interact with artifacts, tools or other controllers that are shared in the group (Shu and Flowers, 1992; Inkpen et al., 1997; Ellis et al., 1991; Chan et al., 2008; Olson et al., 2011). If collaborators are performing a task where they are structuring the work by taking turns, they are performing a sort of coordination; often referred to as turn-taking protocol. The following section will further elaborate on topics related to turn-taking protocols and coordination of collaboration.

2.1.1 Coordinating Collaboration

Coordination is something present in a winning sports team, during a well organized conference running smoothly or in any other successful activity controlled by a group of people according to Malone and Crowston (1994).
It’s often easier to observe the absence of coordination than there is to observe the process itself. There are a diverse number of definitions of coordination in the literature, which illustrates the lucidity of the concept. Malone and Crowston define coordination as:

*Coordination is managing dependencies between activities.* (Malone and Crowston, 1994, p.10)

In this rather inclusive definition of coordination they argue that it captures parts of related concepts such as cooperation, collaboration and competition as these processes can be viewed as ways of managing dependencies between activities. This definition is largely inspired by organizational theory, but can be applied to many different types of systems, for example biological, computer, or human centered systems. Malone and Crowston argue that similarities within these different coordination processes are greater than the differences.

**Perspectives on Coordination** When studying collaborative systems using technology one can view the systems from different angles. Edwin Hutchins takes on a system view of cognition that unlike traditional theories extends the reach of what is considered cognitive beyond the individual to encompass interactions between people, resources and materials in the environment (Hollan et al., 2000; Hutchins, 1995). Within distributed cognition the term cognitive system is commonly used; these are systems larger than the individual and they may have cognitive properties that cannot be reduced to the cognitive properties of individual persons (Hutchins, 1995). Distributed cognition shares many concepts and ideas with the dialogical theory of communication (Linell, 2009). Cognitive processes may be viewed as distributed across the members of a social group. Cognitive processes may involve coordination between internal (mental) and external (material or environmental) structure (Hollan et al., 2000).

“A process is not cognitive simply because it happens in a brain, nor is a process noncognitive simply because it happens in the interactions among many brains” (Hollan et al., 2000, p.175).
From a distributed cognition point of view, collaboration can be seen as coordination among persons and artifacts. Much of the sequence coordination is removed from the individual and given to the structure of the artifact or system with which one is coordinating (Hutchins, 1995). This is reflected in the overall definition of coordination within distributed cognition;

“...to set oneself up in such a way that constraints on one’s behavior are given by some other system.” (Hutchins, 1995, p.200).

To give a simple example, consider the game of backgammon. Both players’ behavior is constrained by the rules of the game and their actions are coordinated by the system of which they are both parts of. One does not roll the dice while the other player is making his or her move, since it would break the rules. One’s behavior is thus comfortably constrained by ongoing game state as coordinated by the game protocol.

Coordination also plays a big role in the theories of Herbert Clark (Clark, 1996). Clark presents in his book *Using Language* a theory of language use that is highly connected to actions and activities and that the study of language is an affair for both cognitive and social sciences. His theory is that language use and communication are really forms of joint action. A joint action is an action that is carried out by an ensemble of people acting in coordination with each other. In other words, language is carried out by a group of people acting in coordination with one another. Clark’s theory places emphasis on the fact that we are doing things together when we are using language; we are always engaged in some sort of joint activity. According to Clark, many researchers have not recognized this and thus treated many features of joint action as features of language; features including: coordination of action, cooperation, opportunistic action and common ground. Clark’s theory explores both social and cognitive aspects of language in use.

What then are the common denominators, the basic “processes” of coordination, in different systems and according to the different theories? Malone and Crowston (1994) emphasize two central processes: *group decision making* and *communication* - which they claim to be involved in almost all instances of coordination. According to Sharp et al. (2007), who are
inspired by distributed cognition, some of the central coordinating mechanisms include; communication; social aspects, in the form of schedules, rules and conventions; awareness; and shared external representations.

From a group psychological view on collaboration, communication may be regarded as the main coordinating processes and a core component of any collaborative task, according to (McGrath and Hollingshead, 1993). Collaboration is a complex process that involves more than the exchange of information; they further argue that collaborative work entails cognitive, emotional and motivational aspects. However, according to Linell (2009) language and communication provides more than coordination, it is further a way of solving the global project or activity (loosely used) the group is engaged in. Clark (1996) argues that coordination required to structure joint activities and actions can be achieved in many ways; including shared subtle clues, explicit agreements, social protocol, common ground and more.

The above paragraphs illustrate the lucidity of coordination. An underlying theme throughout the remainder of this chapter is to further fill in the concept of coordination from the different perspectives and theories.

**Awareness and External Representations** Being aware of our environment and what others are doing and keeping them aware of what you are doing are all important aspects of collaboration (Sharp et al., 2007). People use peripheral awareness to remain aware of what’s going on around them, even when engaged in another task. This is due to the combination of direct and peripheral monitoring that keeps us up to date on what’s going on around us. When people work together they coordinate their work based on the awareness of what the others are doing (Sharp et al., 2007). This is especially important for interdependent tasks, or what Hutchins (Hutchins, 1995) refers to as sequentially constrained procedures, where one person’s activity is required to carry out another task. This was also observed by Heath and Luff (Heath et al., 1992) who studied two controllers working in a control room in the London Underground. They found that the controllers often overheard or oversaw what the other was doing and acted accordingly - even though the other gave no explicit “order” to do so. This is a form of indirect non-verbal communication, as one’s action
sends signals to the collaborators that helps them to structure their work (Gutwin and Greenberg, 2000).

Shared external representations are commonly used to coordinate people, according to Sharp et al. (2007). These can be shared monitors, calendars, charts, checklists and so on. They provide external information on who is working on what, where, when, what happens next etc. - information that can be used to coordinate the group.

According to Clark (1996), in many joint activities the state of the activity can be traced by its external representation; for instance, it’s possible to assess how many moves has been performed in a given chess game through its representation, at least for a professional chess player. When performing joint activities these external representations are almost invaluable, imagine playing a chess game without a board or pieces. Figuring out what action to take next is made immensely easier when the external representation of any activity is available to us. In fact, many activities may become impossible without them. Ordinary external representations may be supported with additional information, the task of estimating the number of moves performed in a chess game would be much easier if the player would have written a list of all their moves.

Kirsh and Maglio (1994) and Kirsh (2001) gives examples of how we can use the environment to aid us through the act of organizing and coordinating work; by arranging the space surrounding us in an intelligent manner. These commonly used spatial arrangement techniques offload memory and reduce search. For instance, one could spatially arrange the parts of a bike in the order they are to be assembled, alternatively in the space where they are to be assembled. However, according to Hutchins (1995), the structure in the environment can provide much more in form of cognition than just external memory aid. External representations can change the entire task from a memory task to a visual task; for instance, instead of trying to remember a sequence of numbers, one can write them down on a piece of paper - transforming a memory task into perceptive task.

Space could also be used to coordinate a collaborative task by providing means for in-group communication or action turn-taking. This was observed by Bang and Tiimpka (2003; 2007) when they saw that clinicians placed patient records on a desk to form a shared public display with high information openness. With this display, clinicians were able to track
work-in-progress as the display represented the current problem state for the health-care team. The observations suggest that the patient records and other physical artifacts are used by clinicians in different ways to form cognitive tools that support joint attention and collaboration. As people use space to structure a task it is reasonable to say that they are using space to structure the order in which a certain task is to be performed, i.e. they use space to structure which action to be performed.

Typically, when groups are collaborating in a face-to-face setting seated around a table, they often use tools such as pens, pencils and paper. They share information and objects placed on this surface. People tend to use orientation in order to establish personal and group spaces, and to signal ownership of external objects and coordinate how shared objects can be used (Scott et al., 2004; Kruger et al., 2003). The research by Scott et al. (2004) reveal that collaborators use three types of tabletop spaces in order to coordinate their interactions: personal, group, and storage spaces. Further, Kruger et al. (2003) studied how participants in groups use orientation during collaborative activities; one activity studied was collaborative puzzle solving. They found that orientation can prove critical in how individuals comprehend information, communicate and coordinate their actions. Surprisingly, they noticed that participants often oriented puzzle pieces in a way that allowed the collaborators to read them. More on the roles of tabletop spaces and orientation will be presented in section 2.2.2.

Social aspects Social aspects of collaboration include the routine actions and interactions people engage in during everyday life, according to Sharp et al. (2007). Social aspects are coordinated through communication, awareness and other mechanics such as social protocols. Ellis et al. (1991) argue that collaborative work is governed by group protocols, which are mutually agreed upon ways of interacting. The group protocols may be partially built in into the hardware or software of the technology the group is interaction with, which is then called technological protocol. When control of the activity and interaction is left to the participants Ellis et al. refers to it as social protocol. When social protocol is used, control of the groups’ process is left to social etiquettes, which are (hopefully) mutually understood and agreed upon. Social protocol includes formal rules
or policies and less formal practices, such as polite turn-taking procedure and hand-raising protocol. People are generally good at social protocol in face-to-face collaboration and they tend to adapt these protocols when using technology - but sometimes protocols are not enough and may lead to breakdowns in collaboration, as observed by Morris (2006). How technological and social protocols can be used to support the collaborative process is further discussed in section 2.2.

Central to coordinating any collaborative activity is the common ground between participants, according to Clark (1996). Common ground is the set of knowledge, beliefs and suppositions that the participants believe they share. Any social activity takes place within common ground, in order for two people to coordinate their action they need to root these actions within their common ground. People often engage in a joint activity with great deal of presupposed common ground for how the activity is supposed to be carried out. For instance, common ground makes it possible for a speaker and a listener to coordinate on what the speaker means and what the listener understands the speaker to mean - it’s the coordinator between the speaker’s meaning and the listener’s (addressee’s) understanding.

Common ground affects coordination in different ways, in the way of what information is given to us in order to perform a task, or perhaps we share some common history together (Clark, 1996). If the task is for me and my partner to pick “the ball” and we get to choose from a bowling ball and a tennis ball we would probably both pick the tennis ball since it’s the most salient ball, and the task was to pick the ball. However, if we share the history of playing several years in a bowling team together we probably would pick the bowling ball, since it’s the most salient ball for us given our common ground. People tend to choose the solution to a coordination problem in regards to what they take to be theirs’ and their partner’s current common ground. This is what’s known as the principle of joint salience. Social protocol also requires or assumes a good amount of common ground between the participants (Clark, 1996).

Different situations and activities may cause a group to use different approaches and strategies - different “situational protocols”. Hutchins (1995) gives an example of how we use “rules of thumb” to coordinate our behavior, and he shows that they can sometimes cause difficulties when coordinating action in collaborative settings. To give a simple example, if you have a
2.1. Collaborative Work

set of patient charts you are supposed to arrange in order of how sick they are, which is the rule of thumb - a rather simple rule given the appropriate medical training. This task will be become more troublesome if it’s the patient charts are distributed among a group of individuals. Since you don’t know the full set of patient charts, you can’t know for sure where your patient chart is to be placed in the order. In this example a trivial problem since you could discuss in the group the state of the group’s patients and arrange thereafter, but in other situations the group may not have that privilege. This illustrates that rules of thumb may break down when used in an inappropriate situation. More on collaborative situations and activities will follow in section 2.1.4.

Coordinating with Communication  Communication is one of the most basic preconditions for any group to function (Jern, 1998; McGrath and Hollingshead, 1993; Linell, 2009; Malone and Crowston, 1994). Communication involves many things depending on from which perspective it’s viewed. Commonly involved notions are: bodily actions, expressions, signal systems, emotions, cognition, and of course speech and language. When studying let’s say two people playing the piano, we can analyze them as a social system, or as individuals (Clark, 1996). Clark argues that one can analyze their actions, their communication, their coordination, or what’s going on within their heads. All of these viewpoints are important, and if we are to truly understand the piano players as a joint coordinated system none of the perspectives will be sufficient alone, reasoning in line with the distributed cognition framework. The same goes for two people talking; therefore the study of language must be both a cognitive and social science, according to Clark (1996).

Different kinds of communication may take place within different kinds of interactive and communicative systems, according to McGrath and Hollingshead (1993). Technological systems developed (e.g. e-mail, video-conferences and digital tabletops) can be classified in different categories of communication based on their functional role within the group, ranging from in group communication to communication with information sources or external contacts. Important to note is that any technological system may serve more than one of these functions.
The coordinating device of communication is a bit too complicated to chew off in this section, and more on what communication might involve and how communication itself may be coordinated will be presented in the following two sections. This section can be concluded with saying that coordination and collaboration are two complex matters; involving perceptual, social, cognitive, communicative aspects; as well as aspects of the environment and the nature of the collaborative activity itself.

2.1.2 Language and Communication

Dialogical theory according to Linell is an alternative to mainstream theories in linguistics, psychology and social sciences (2009). To illustrate the dialogical theory in relation to communication it can be viewed in contrast with the more traditional view of communication: the monological theory of communication. The monological tradition is widely used in many fields of research and the observant reader will have picked up signs and mentions of this view in previous sections. So, let’s start with a short overview of the monological theory.

Monological theory can be called “monological” for the way the theory’s constituent theories tend to account for many aspects of communication, such as meaning and sense-making, as resulting from either one of two possible sources: the individual speaker or the language code (Linell, 2009). The monological view of communication is made up of a handful of constituent theories: the information processing theory of cognition (focusing on internal mental states and computation), the transfer model of communication (i.e. that communication is a transfer of a message from sender to receiver) and the code model of language (language consists of static signs; stable communications of expressions and fixed meanings). This can be seen in how context is generally treated as something external to thinking and communicating within the monological school of thought; sometimes even described as a collection of “nuisance variables”.

Linell (Linell, 2009) describes four points on which dialogical theory differs. First, communication is not perfectly clear; rather it’s sufficient for understanding. In other words, people are content with understanding each other sufficiently well to carry the conversation or other current doings forward. People have developed a set of skills and methods for
continued dialogue, including ways of repairing statements and follow up questions. Secondly, people seldom have univocal intentions, and they often have partly ambiguous or non-reflected goals with the statement. Speakers do not always know in advance what they are going to say, because new events may unravel while they are speaking. Utterances and interpretations of utterances are in this view coauthored. Thirdly, language does not contain stable links between expressions and meanings, and words do not have fixed unique meanings. Words instead are seen to have different ‘meaning potentials’ in different contexts. Lastly, cognition and communication are interdependent; it is not the case that the former determines the latter. Cognition and communication have in monological theory been seen as separate processes; cognition being prior and contained to the individual, while communication is something secondary taking place between people.

In the dialogical stance the cognitive process is being kept together with the interactional processes; they are seen as partly the same process rather than two different steps. Communication and cognition are closely related to each other. Dialogism also sees the distinction between doing and thinking as rather blurred, and they are both seen as actions - this view is also shared by Clark (1996), see section 2.1.4. The notion of dialogue covers both inner-dialogue (thinking) and outer-dialogue (interpersonal communication). A key concept in dialogical theory is that communication and cognition always involve interaction with others: other person, other systems, other artifacts, etc (Linell, 2009). A difference between the two is that communication per definition involves interaction. However, according to Linell cognition also involves interaction with the world (perception, thinking, acting and so on), albeit not always with other human beings. One research field in with features of dialogical theorizing is commonly used is cognitive dynamics, which entails the concepts of extended mind and distributed cognition according to Linell (Linell, 2009). So far the focus has indirectly been on verbal communication, but as we know communication may take other forms as well. As seen in the studies by Heath et al. (1992), communication can be very opportunistic and unintentional.

**Gestures** Bekker et al. (1995) categorized communicative *gestures* used
in collaborative co-located groups into four main categories. Kinetic: gesturing (movement) during all or part of an action performance - for instance gesturing how to use the coffee machine when explaining how it is used. Spatial: the movement indicates distance or location or size. The classical example would be someone saying “I caught a fish this big” and gesturing the size. Point: using fingers (usually) to point at a person, object or place. Often one points to a part of a document, part of the whiteboard or a position on a monitor to indicate some point of interest. Other: this category encompasses all gestures not included in the above categories, including those that merely serve to emphasize parts of a sentence or to attract attention. Gestures are sometimes used to indicate that a participant wants to have the next turn in speaking (Bekker et al., 1995; Mondada, 2007). Especially pointing gestures can be used as ways of projecting turns. This was observed by Mondada (2007) who further showed that pointing practices and turn-taking practices are deeply intertwined. Pointing was observed to be used as a method for projecting self-selection, designing transition spaces, and more generally defining speakership.

**Graphical Communication** Communication is not always taking place person-to-person, when viewing communication as dialogical or through the distributed cognition framework the environment and different artifacts can also be used as means of communicating (Linell, 2009; Hutchins, 1995). One way of communication through artifacts is by **graphical communication**. Graphical communication has often been interpreted as being a one-way communication (graphic to person), as discussed out by Fay et al. (2003). However, according to Fay et al. (2003) interactive two-way graphical communication is a frequent occurrence (e.g. collaborative map sketching). In their 2003 article, Fay et al. investigated the principles governing interactive graphical communication by comparing interactive graphical communication with verbal communication. In a set of experiments where participants played a Pictionary style game, Fay et al. found evidence suggesting several parallels between verbal and graphical communication. Just as in verbal communication, graphical ‘conversationalists’ first ground (Clark and Brennan, 1991) what they mean which aids both the production and comprehension of future representations.
2.1.3 Taking Turns Communicating

Traditionally, roles within conversation have been divided up into speaker and listener. Since both parts in a conversation normally have something to say the roles of speaker and listener change constantly. This constant interchange between speaker and listener within a conversation is highly coordinated and often referred to as turn-taking. A lot of work has been put into analyzing this conversational phenomenon, and the most well known work is the characterization of turn-taking by Sacks et al. (1974). They put forth a systematic model of turn-taking that characterizes the structure of human conversation. Conversation is structured as (1) an emergent property of (2) local decisions based on (3) prediction by the participants. In their view, turn-taking is locally managed and participant-administrated. Local management means “all the operations are ‘local’, i.e. directed to ‘next turn’ and ‘next transition’ on a turn-by-turn basis” (Sacks et al., 1974, p. 725). According to this model, “units” are used by speakers to construct a turn, or in other words to determine when to transition to the next speaker. These units are commonly known as turn-constructional units or TCUs for short. Recognizing that a particular sentence of type S is being uttered by a speaker, the listener can use his or her knowledge about sentence type S to predict when it ends, making it possible to take turns with no gaps.

The systematic model by Sacks et al. (1974) can be described in four steps according to Linell (In press). First, the current speaker at any given moment is privileged to speak until the end of the turn is reached - fulfilling a complete syntactical unit. Then the speaker reaches a potential turn-transition relevance place, or TRP, when others are expected to take the turn or at least give feedback. Second, the next speaker is the one directly addressed by the previous speaker; this is true especially when a question was asked. Third, if no next speaker has been addressed by the previous speaker, the turn is open to anyone present. If the turn is taken by someone else than the next addressed speaker the turn is said to be stolen. Fourth, if no other claims the turn, according to step two and three, the current speaker may proceed. Sometimes the speaker may feel compelled or forced to add something. The speaker may also prevent losing the turn. This process is often very fluid, and notable is that the time of changing speaker
is measured in tenths of a second. Aspects of the turn taking such as the allowed length of a pause may vary culturally (Linell, In press). The speaker may mark the end of a TCU syntactically, prosodically or pragmatically. Linell notes that not just the aspect of who has the turn varies; also the aspect of how a turn is built is worth exploring. A turn may consist of more than one TCU and are often shaped as identifiable sequences - the most common being utterances and replies (pairs of speech). This is a simplified overview of conversational turn-taking. The conclusion is that conversation can and has been analyzed as being structured in turn sequences.

According to Thórisson (2002), a mistake by Sacks et al. was to think only of how verbal utterances relates to the turn. What is really needed is a theory of how all communicative acts can be constructed in turns, including other communicative aspects as mentioned above. Another argument by Thórisson is that the model does not take into account the internal state of cognitive processing of the participants, which clearly also affects the way they respond to cues in the dialogue. Thórisson’s solution was to lump all of these ideas together into a bag she called context. When including more communicative actions, the problem of knowing when to switch turn re-enters the stage. According to Thórisson (2002), the difficulty of modeling turn-taking lies first and foremost in perception, because a participant has to infer what constitutes a valid turn-giving “signal” solely from perceptual information. Thórisson (2002) presented a model of turn taking called Ymir Turn-Taking Model (YTTM), which is not a descriptive model, but a generative model. YTTM address the full perception-action loop of real-time turn-taking, from (1) the basics of multimodal perception to (2) knowledge representation, (3) decision making, and (4) action generation for gaze, gesture, facial expressions and speech planning and execution. YTTM has been implemented in different working systems, receiving high scores in performance when reviewed by users, a most impressive feat to say the least.

2.1.4 Actions and Activities

Group interaction and performance is greatly affected by the type of task or activity that is performed by the group (McGrath, 1984; McGrath and Hollingshead, 1993). Virtually all researchers on groups would agree on
that group performance must be studied in relation to the group activity; group performance can (or should) in other words not be studied generically. To appreciate what effect technology will have on group work requires reckoning with the effects that various tasks have on group work. McGrath and Hollingshead (1993) divided tasks into four main categories; 1) generative, tasks including generating ideas or plans; 2) choosing, tasks including decision making and problem solving; 3) negotiating, this includes resolving conflicts of views and interests; 4) executive, this final category of tasks includes competitive tasks and other performance tasks. These main categories and their subtypes are illustrated in the group task circumplex, see figure 2.1. These types of tasks differ on a variety of aspects; such as amount of transmitted of information, degree of collaboration, and whether the nature of the task is cognitive or behavioral.

Figure 2.1 Task circumplex by McGrath (1984)
Sequentiality From the perspective of distributed cognition, the nature of the activity is central to how the activity is to be coordinated (Hutchins, 1995). Activities that take place in parallel are examples of simultaneous coordination in a functional system that transcends the boundaries of the individual actors. Buffering is an important ingredient in any activity in which the components are running in parallel, for it provides protection against destructive interference between processes running in parallel. For instance, a talker on the phone is an advanced buffer using her knowledge of conversational turn-taking to decide when to forward information and when to listen. Without this buffer the conversation would soon become chaotic.

Hutchins (1995) highlights an important aspect of any procedure or activity; that is whether the procedure is sequentially constrained or sequentially unconstrained. In a sequentially unconstrained procedure no operation will ever disable any other enabled operation. In this setting there is no communication needed between the agents other than their effects on their shared environment. Each agent can simply take action when the possibility for action arises. In a sequentially constrained procedure on the other hand, execution of any enabled operation will disable other enabled operations. This requires some control over the sequence of action, possibly through planning or backtracking.

Unconstrained procedures are easy to distribute since they can be run in parallel (Hutchins, 1995). One should try to take advantage of potential parallelism if possible, since it accommodates parallel work which may lead to increased performance. Sequentially constrained task on the other hand require coordination among the actions taken. There are many ways to coordinate these actions; or in other words, ways to coordinate the action-turns taken. Specifying in advance an overall global plan for sequential behavior is a solution to the turn sequencing problem, according to Hutchins. Although, Hutchins observed that it is possible to organize the collaborative actions without there being a global plan. He observed that each member only needs to know what to do when certain conditions are produced in the environment.

Often it’s desirable to turn a sequentially constrained task into an unconstrained task (Hutchins, 1995). One general technique for doing this is to manipulate the enablement conditions of various operations. Accord-
2.1. Collaborative Work

According to Hutchins, a simple rule for this is: “suppress the enablement of any operation that could disable another already enabled operation.” An example of this is that the starter motor in a car that will not turn unless the transmission is in neutral; here the operation of “starting the car” has been suppressed unless the car is in neutral. This problem will be re-encountered in section 2.2.2.

The external representation can also be used to alter whether the task is constrained or unconstrained. For instance, Zhang (1992 via Hutchins (1995)) showed the possibility of changing the sequential constraints of the Tower of Hanoi - problem by embodying some of the sequence-constraining “rules” in the physical representation of the problem.

**Joint activities** From the view of Clark’s theories (1996) we are always engaged in performing some sort of activity, and when we are working in a group we are engaged in a joint activity. These joint activities can be divided into hierarchically nested projects and subprojects consisting of joint actions or further joint activities. Clark uses the notion of activity instead of task and henceforth the term activity will be used. An activity is a fuzzy category, as it can either be a time bound task (e.g. a football game) or a continuous process (e.g. learning or living). The number of members within an activity can vary from one and up. These fuzzy activities are ultimately made up by physical actions, which in turn can be performed joint or individually.

Two people cannot carry out a joint activity without communicating with each other (Clark, 1996), and when people are using language to communicate they are performing a joint action. Coordination of joint activities involves transitions within a joint project and between joint projects. Bangerter and Clark (2003) propose in their paper that people use dialogue to coordinate these two kinds of transitions. Entering and exiting joint projects are referred to as vertical transitions; continuing within joint projects is called a horizontal transitions. Clark and Bangerter further show how people can coordinate vertical transitions using words like “okay” and “all right”, whereas they can coordinate horizontal transitions using words like “uh-huh” and “yeah”.

Clark (Clark, 1996) characterizes joint activities by using five dimen-
sions on which each activity-type differs. These dimensions are; (1) scriptedness, the extent to which an activity is scripted or spontaneous; (2) formality, how formal an activity is, ranging from a city council meeting to a gossip session; (3) verbalness, the degree to which speech is an integrated part of the activity; (4) cooperativeness, the cooperative nature of the activity, ranges from cooperative to competitive; (5) governance, the degree to which the activity is egalitarian (equal roles within the group) or autocratic (one has a dominant “leader” role). Clark points out that this is not an exhaustive list, but it’s enough to keep us occupied and give us a good initial categorization of different joint activity types. Many of these activity dimensions were first introduced by Levinson (1979).

Activities always have some goal or purpose, according to Clark (1996). The main goal of an activity Clark refers to as dominate goal. The initial dominate goal may be very vague (e.g. a goal to “gather materials”) and may evolve over time. There is also a certain task division among the participants that will mount up to the dominate goal. Participants in an activity often pursue more than one goal at the time. They will pursue the dominate goal, but the participants also often have procedural goals that govern how the dominate goal is to be reached. Participants also have social interpersonal goals, such as impressing co-workers or being polite. Further, participants may also have a private agenda, a private plan do deceive the others. All these goals are either private or public. The public goals are there for all participants to see and recognize; the dominate goal that the joint activity is trying to achieve is public. Private goals are on the other hand individual and concealed from the others’ view.

Clark (1996) makes a distinction between active participants and non-participants. The participants in a joint activity are the current actor (speaker in dialogue), the addressee and any side participants. Non-participants include bystanders and eavesdroppers, the difference between these two is that bystanders are openly present but not part of the ongoing conversation or activity and eavesdroppers listen in secret. The different participants present help shape how the speaker and addressees act towards each other - for instance, the speaker can try to act in a way that excludes any bystanders.
Joint Actions  The previous section stated that joint activities are ultimately made up of joint actions, but what does it mean for an action to be “joint”? According to Clark (Clark, 1996), joint actions are simply actions performed by ensembles of people. The differences between joint and individual actions may not be very clear, but a person’s processes are very different when he or she is performing an individual action in contrast to a joint action.

A simple example that illustrates this difference is whether one individual or two individuals are pushing a car that won’t start. If there is no coordination between the two individuals jointly pushing the car - i.e. if they act as two autonomous individuals unaware of each other - they would most likely lose control over the car sooner or later. Imagine that one of them pushes the car unaware of the other who has stopped pushing because the car is about to go over the edge of the harbor. Without any coordination between them they cannot affect each other’s actions and the car would end up soaked. With some coordination however they can push the car to the intended location, or at least stand a better chance of avoiding the water.

Still, when two people are working together they are still performing individual actions, these actions Clark refers to as participatory actions: an individual action performed as part of a joint action (Clark, 1996). Participatory actions do not have to be analogous, as they are in the example above when they are both performing the same task (i.e. pushing). For instance: when two people playing a duet were one of them are playing the piano and the other is playing the flute their respective participatory actions are quite different but highly coordinated with each other. Here one could argue that they are performing individual actions and not a joint action consisting of two participatory actions, as if they were playing alone in separate rooms. However, if you were to compare the isolated flute from the duet and the solo flute you would find them very different; since the flute in the duet version is coordinated to the sound of the piano. In the same way speaking and listening should be interpreted as participatory actions that make up the joint activity of language use (Clark, 1996).

Not every action is taken in order to move the individual or the group closer towards the intended goal. Kirsh and Maglio (1994) introduced a distinction of epistemic and pragmatic actions. Pragmatic actions have
been most studied and are the actions whose primary function is to bring an agent closer to his or her goal. A plan serves as a path of pragmatic actions from initial to goal state. When playing Tetris (as studied by Kirsh and Maglio (1994)) a pragmatic action would be to move a piece one step closer to its intended goal destination. Epistemic actions on the other hand are physical actions that do not bring the acting agent any closer to its goal. Rather, these actions serve to aid or simplify the agent’s current activity. In the Tetris example activity, players were observed to rotate pieces not to get closer towards the goal, but in order to support and simplify the decision-making of where to place it.

**Action Coordination** Joint actions move the joint activity forward and towards the domain goal of the activity (Clark, 1996). If the joint activity is one person buying an item from another, two involved joint actions may be deciding upon a price and exchanging the item for money. What makes these actions joint is the coordination of individual actions by the two persons. How coordination is achieved is one of the fundamental issues of language and action (Clark, 1996). Even in competing activities the participants are performing a form of coordination. Take for instance a tennis match, when one player hits the ball he or she wants the other player to not be able to hit the ball back, so he or she must figure out how to make it as hard as possible for the other player - competitive activities often involve this “anti-coordination”. There are two types of coordination involved in a joint action; coordination of content, what the participants intend to do; and coordination of process, the physical and mental systems they use in carrying out their intentions.

One important aspect of coordination is to *predict* what your collaborators will do, according to Clark (1996). In a series of games, constructed by Schelling (Schelling via Clark (1996)), where two players were to predict the other’s action. People are surprisingly good at this, even though they are given very little information to go on. The simplest game is simply the task of picking heads or tails in a coin toss, if they both pick the same side they win. This is a form of joint action since they are both performing a participatory action as part of an action by the pair of them governed by a coordination device. A coordination device is a focal point or clue that can
lead people to anticipate each other’s actions. In this case the coordination device is the prominence of “heads” - it’s simply the most common pick (given some common ground) and people use this clue to guess that the other picks heads.

The example above is just a game where the participants are not allowed to communicate with each other and still manage to perform some coordination, but what about more “real” situations? Clark (1996) proposes that people tend to use some conventional methods for coordination. One is to make an explicit agreement, e.g. to agree to meet by the library at 10 am or to agree to pick heads in the next coin toss. Conventions exists on a cultural level as well, for instance it’s convention to tip the personnel after dining at a restaurant or to shake hands when greeting somebody - many of these conventions make up the social protocols, as presented in section 2.1. One way to coordinate is to use signaling systems of various kinds which serves as bases for joint action. A signal system can be the traffic lights we see every day, since we all know what the colors mean the signals can effectively coordinate traffic - coordination when and where road users are supposed to start or stop. The meanings of the colors are used in a convention. Other signal systems can include orientation of objects, see section 2.2.2.

Most everyday coordination is continuous, not discrete (Clark, 1996). Continuity demands adaption with moment-by-moment decisions that don’t readily divide into discrete coordination problems. A handshake can be seen as a one-shot event, but when we view it as a joint action we must treat it as a process that unfolds in time. The added element is timing. Most coordination is aperiodic - that is to say not governed by cadence or rhythm. Further, it’s also most often unbalanced in the way that one participant takes some sort of leader role, and the rest follows.

Joint projects A joint activity can be divided into joint projects (Clark, 1996). A joint project is a joint action projected by one of the participants and taken up by the others. Joint projects can be of any size - an entire activity or sections of it. Many words and expressions that have been studied mainly as signals of listener feedback (back-channel signals) or turn-taking devices (acknowledgment tokens) can also be seen as navigators of joint
projects, according to Clark and Bangerter (Bangerter and Clark, 2003). They divide dialogue into two planes; on one plane, people create dialogue to manage the joint activities they are engaged in—for example, activities such as making dinner or designing a new phone application can be coordinated through dialogue. On the other plane, people coordinate the dialogue itself—here are features such as conversational turn-taking in the spirit of Sacks et al. (1974). Division onto two planes is not unique to dialogue. Rather, every joint activity requires coordination of both the content of the activity and the process by which the activity moves forward. Previous studies have focused mostly on the second plane of dialogue management. The planes are not independent, rather they are interdependent. Although, Bangerter and Clark claims that basic joint activities are primary, and dialogue is created to manage them (Bangerter and Clark, 2003).

2.2 Supporting Collaboration with Technology

"Computers will not appear everywhere, but the potential for supporting collaboration is limitless, and more of the potential is realized every day." (Grudin, 1991, p.32)

Technology has led to a revolution in human interconnectedness and collaborative possibilities. Every innovation, such as the telephone or the Internet, always seemed to take a step towards a better world. However, as new technologies of interconnectedness appeared there were surprising effects. For instance, electronic mail (email) seemed to be an effective and flexible way of communicating easily between geographically separated people. But this technology revealed problems of its own; first, by making it easy to communicate it became easy to become flooded with information (Malone et al. (1987) via Olson et al. (2001)); second, something about the medium affected the way people expressed themselves, the emails often appeared hostile or negative towards their recipients (Sproull & Kiesler (1991) via Olson et al. (2001)). To tackle repeated surprising effects like this, a lot of multidisciplinary scientific effort was mustered during the 1980s giving birth to many new conferences; CHI, CSCW, ECSCW and
INTERACT. In the following sections we’ll take part of some of the results and lessons learned for supporting collaboration with technology.

2.2.1 Computer-Supported Cooperative Work

When it comes to collaboration and technology there are several different terms being used. The term “Computer-Supported Cooperative Work” (CSCW) was first used by Paul Cashman and Irene Greif in 1984 (Grudin, 1991). They summoned researchers and developers to a series of ACM-sponsored conferences that examine how people work in groups and how technology can support them. The term “groupware” on the other hand has become a way of labeling products breaking away from the individual desktop confinement. Jonathan Grudin (1991) brings up “the paradox” of collaborative technology support in his introduction to collaborative technologies. What he refers to is the fact that people collaborate seemingly effortlessly in everyday settings, yet it’s hard to design technological support for these processes - why is that? The answer to the paradox, according to Grudin, lies in the fact that many things that we humans do effortlessly is very difficult for computers to do (e.g. speech and vision). We also need to better understand the complexity of the workspace and the nature of collaboration (Grudin, 1991; Gutwin and Greenberg, 2000).

There has recently been a great deal of interest in designing computer tools for group use (Olson et al., 2001). During the early 90s commercially available products started to pop up under the flag of CSCW or groupware. It appears likely that these systems will continue to evolve - in due time, heralding a paradigm shift in computer usage. Roughly, one can say that groupware developed in two separate directions; one is groupware for geographically separated collaboration, such as instant messaging or group organization; the other is groupware for co-located face-to-face group collaboration, such as conference rooms, or more recently - digital tabletops. Groupware can also be categorized by the factor time; that is, whether they support synchronous or asynchronous usage. The size of the groups intended to use the groupware is also relevant, because small groups function differently from large ones. Perhaps the most important factor to be considered is what task or activity is to be performed with the technology (McGrath, 1984).
Christer Garbis (2000) published two case studies in which he had studied collaboration and communication in control rooms using the distributed cognition framework. His focus was on the relationship between workgroups and artifact use. What he noted was that even artifacts intended for public and collaborative use are designed as single user artifacts. According to Grudin (1988), not only is single user technology insufficient in supporting co-located collaboration, but also groupware solutions seems to be insufficient since they often fail. The reason for this failure seems to be that groupware are mostly designed based on criteria for single user applications, thereby missing the potential collective benefits.

**Using Theories** One way of understanding collaboration and related cognitive and social phenomena is through the use of an appropriate theory. There are many theoretical approaches to CSCW, and the field has adopted a number of what Halverson (2002) refers to as theoretical constructs. These include activity theory, distributed cognition theory, conversation analysis, coordination theory and many more. Distributed cognition is now used as a common theoretical framework for research performed within the field of CSCW. Some of these theories relate directly to design, but far from all of them. CSCW often turns to other methods to support the design process. These methods include participatory design, contextual enquiry and user centered design.

How do these methods and theories shape the work within CSCW? Are they of any use? These, I would say, are well motivated questions. This is how Barthelmess and Anderson (via Halverson, 2002)) answers these questions:

*The value of any theory is not ‘whether the theory or framework provides an objective representation of reality’, but rather how well a theory can shape an object of study, highlighting relevant issues. In other words, a classification scheme is only useful to the point that it provides relevant insights about the objects it is applied to.* (Halverson, 2002, p.244)

Halverson (2002) sums this up elegantly: using a theory is like putting on a pair of dark glasses - it tints the world and brings some objects into
sharper contrast, while others fade away. How to choose a relevant theory is another question. It depends on what the intended use of the theory is, and every theory has different powers. If we are interested in designing a new system we want the theory to help us to make inferences about the world that may lead to design issues. The application aspect of how the theory applies to the real world often translates into our need to inform and guide system design. Since CSCW is a multidisciplinary effort, and consequently built on a bag of theories, it is important to reflect on whether any theory can be applied to all “dimensions” of CSCW - and if not: which dimensions are we interested in?

Coordination Mechanisms  What must be taken under consideration when developing technological support for co-located collaboration? After reviewing the literature and their own previous work, Gutwin and Greenberg (2000) argue that many usability problems for co-located groupware, e.g. digital tabletops, are caused by insufficient or mismatched support for the basic elements of collaboration. Contextual factors are also important, but many problems can be resolved by addressing the constituents of groupware collaboration in general. They call these the mechanics of collaboration, which are small-scale actions and interactions that group members must carry to complete an activity. Gutwin and Greenberg’s list of mechanics contains seven elements, ranging from aspects of communication to levels of protection for one’s own work - many of them relevant for successful turn-taking procedures.

The first mechanism is explicit communication, this is the information we intentionally receive and give away. Verbal and written communications are the cornerstones, but other channels such as gestures and deictic references plays important roles as well. In addition to explicit communication there is consequential communication, that is the information we unintentionally give away during interaction. Two types of consequential communication are artifact information and information from a person’s embodiment in the workspace.

People also try to coordinate their action so that they do not conflict with others, and with time people can also learn to predict others’ actions. Some tools and activities require that turns be taken or that actions take
place in an appropriate order. Poor coordination may lead to accidental bumps, overlap in work, or conflicts over shared resources. People can further plan their activities; as they go along with an activity they might among other things divide and re-divide the activity or reserve workspace areas. By monitoring the workplace people gather information which many of the other mechanisms rely on. A lot of the monitoring can be labeled as “workspace awareness”, as people monitor who are in the workspace, where they are working, and what they are doing. Apprentices perform a lot of monitoring in order to learn the tools of the trade.

By monitoring people detect others in difficulty and provide assistance. Assistance may be opportunistic, or may be requested; either way, appropriate assistance requires that people understand what others are doing. dangers of group work may be that others may inadvertently alter or destroy your work, therefore people must keep an eye open and sometimes prevent others action in order to protect one’s own work.

**Taking Turns** Coordinating actions can partly be done through turn-taking protocols. One way of managing turn-taking in (early) CSCW systems is through so called “floor control” protocols. Floor control is an example of a technological protocol as opposed to social protocols (Ellis et al., 1991), see section 2.1.1. These protocols are commonly used in geographically separated collaborative systems, such as conference systems (Lantz, 1986). When using the floor control protocols systems only process one user’s input requests at a time, and thereby impose a restricted group turn-taking process.

If a user has the floor for a specific document, parts of that document becomes locked to other users. If another user then requests access to the file a negotiation situation emerge instead of a “classical access denied”-message (Sacks et al., 1974). Having the floor is a collaborative mechanism that provides a scope of negotiation between users. Advantages of technological protocols includes that they ensures that a specific process is followed by providing more structure to the group’s activity and assistance for less experienced users (Ellis et al., 1991). They can on the other hand hinder collaboration by being overly restrictive; a group’s work strategy may not be supported by the protocol. The system can therefore constrain
a group that needs to use different processes for other activities, and they are particularly ill-suited for activities with high parallelism.

2.2.2 Digital Tabletops

Until recently, only a few display systems have supported co-located synchronous collaboration with parallel user input. The majority of current systems require turn-taking strategies as they provide only one input device and one active input channel (i.e. mouse or keyboard). Systems like these are collaborative in the sense that a group can gather around and discuss the digital information (Stewart et al., 1999), but only one person can manipulate digital artifacts at any given time and control must be passed in a turn-taking manner for a second user to interact. Parallel interaction is one of the virtues of modern groupware systems, such as the digital tabletop.

A digital tabletop is a type of sharable interface, which is an interface designed for use by more than one person at a time and often being able to handle multiple inputs from different users at the same time (Sharp et al., 2007). Sharable interfaces are sometimes designed to literally become a part of the furniture, e.g. a digital tabletop can be designed and used as an ordinary table. An underlying premise in these types of collaborative interfaces is that the natural way for people to work together is by gathering around tables or vertical boards. User studies have shown that horizontal surfaces, in contrast to vertical surfaces, further support turn-taking and collaborative work on co-located groups (Rogers and Lindley, 2004). Interacting with a digital tabletop reaches beyond the traditional WIMP (windows, icons, menus and pointer) framework which is the standard approach used for personal workstations (Sharp et al., 2007).

Digital tabletops have been used with a variety of interaction methods, ranging from pen based, touch and gesture interaction (Sharp et al., 2007) to more novel techniques such as multimodal interaction (Morris, 2006; Piper and Hollan, 2009) and interaction using tangibles (basically any physical object used for interacting with the tabletop) (Rogers et al., 2009; Olson et al., 2011). Touching and gesturing are considered a more natural way of interacting with an interface than the classical keyboard and mouse setup, according to Sharp et al. (2007). Different forms of interaction can provide users with more freedom for expression and manipulation;
a user can for instance use two hands in order to stretch and move an object.

**Tabletop Activities** Scott et al. (2003a) suggest that there exist four general classes of digital tabletop systems in the literature: digital desks, workbenches, drafting tables, and collaboration tables. Digital desks and drafting tables are mostly for individual use, integrating paper-based and digital media. Workbenches use virtual reality environments projected above the table. In focus here are collaboration tables, which include digital tabletops that support small-group collaborative activities, such as group design, story sharing, and planning.

According to Scott et al. (2003a), the design of digital tabletop systems is at a crossroads; the technology is here or on its way, but it is not clear which tabletop system configuration is suitable for each collaborative environment or activity. In their view, the creation of a collaborative task taxonomy is an important step in order to help determine which tasks and activities are suitable for tabletop collaboration. Morris (2006) contributed to this step by classifying three types of observed group work strategies used during different collaborative activities. These are: parallel (all group members perform similar actions in parallel), serial (all group members jointly focus on one item at a time), and assembly-line (all group members work in parallel on different aspects of the task). Morris further argue that no one of these strategies is more or less collaborative in general, although some might be more effective for a particular activity.

**Collaborative Breakdowns** Along with the benefits of co-located groupware, such as enhanced group productivity and collaboration, new problems and challenges arise. Allowing members of a group to simultaneously access a shared interface may cause members to change settings that impact the entire tabletop. As noted by Stewart et al. in their landmark paper on Single Display Groupware (Stewart et al., 1999, p.5): “new conflicts and frustrations may arise between users when they attempt simultaneous incompatible actions.”

Many of the tabletop studies presented in this section of the thesis are conducted by Meredith Morris and her collaborators, as presented in (Mor-
During their studies they supplemented their own observations by distributing surveys in order to gather opinions on how document-level conflict could and should be handled. The surveys presented scenarios in which two users (A and B) were in conflict over 1) a paper document on an ordinary table, and 2) a virtual paper document on the digital tabletop. Particularly noteworthy is the fact that the paper scenario elicited far more “social” responses than the digital one - rearranging the seats around the table, reading the document out loud to a partner, and explicit social negotiation. Morris speculate that these actions were left out from the digital tabletop setting because of the “newness” of the technology. Social solutions that would readily apply when interacting with traditional paper media may not come to mind in a novel technological setting.

This may be an explanation to why social protocols often break down during collaborative digital tabletop interaction (Morris, 2006) - as one would not simply snatch a physical paper from someone’s hand while they are reading it; such an act would be considered rude! Later findings also indicate this, in the recent studies by Olson et al. (2011) they studied children collaborating using digital tabletops. They found that the participating children seemed to have a clash of expectations in how to collaborate when using a digital tabletop. Some of the children assumed turn-taking procedures (asking to participate, passing the control, etc) while others assumed territories and fought for having sole control of the tabletop. Olson et al. speculated that the reason for this is that tabletop interaction was new to these children, and they therefore were unsure what set of scripts or norms applied.

Conflict Coordination Policies One conclusion drawn by Morris et al. (2004) is that social protocols may not always be enough to coordinate collaboration. In some cases social protocols are sufficient to mediate groupware collaboration, but certain actions and especially accidents and confusion are hard to govern with only social protocols (Gutwin and Greenberg, 2000). Morris et al. (2004) believe that good software design can help coordinate group actions to avoid situations as described above. Good design can further encourage equitable participation in group activities and increase awareness of important events.
Therefore, Morris et al. (2004) proposed a set of coordination policies in order to get to grips with conflicts that may arise at the digital tabletop. The policies aim to provide applications with more structure and predictability than social protocols, yet also allow for more flexibility than rigid access permissions. They focus foremost on how to structure and coordinate direct application manipulation on the tabletop to avoid and resolve conflicts amongst the users. They identified three types of conflicts: global, whole-document, and sub-element. Global conflicts affect the entire application, whole-document conflicts are conflicts when using a particular document, and sub-element conflicts occur when several users are editing the same item.

With these conflicts Morris et al. present initiative strategies for resolving the conflicts: proactive, reactive, and mixed-initiative. Proactive policies lets the outcome of the conflict be decided by the owner of the element, or the initiator of a global change. Reactive policies allows the other users to affect the outcome, and mixed-initiative policies takes in information from all parties involved in the conflict in order to resolve it.

For instance, one proactive policy for global conflicts can be the use of a “privileged object” (Morris, 2006); this can be a special shared menu for making global changes. A privileged object might make people more aware of the effect their interaction is going to have on other users. An example of a reactive policy for whole-document conflicts can be using a “duplicate policy” (Morris, 2006), where the contending user acquires a duplicate of the original document. Making documents private or public is another coordination policy for whole-document conflicts, but this brings the issue of what information should be private or public when and how.

**Sharing Objects** When a user shares a digital document on the tabletop it can be interpreted as the act of dynamically changing the accessibility of the document by transitioning from a personal access control policy (whereby only the document’s owner can access the document) to a public access control policy (where all users at the table access the document) (Morris, 2006). To support the process of sharing Morris suggests four general interaction techniques: release, relocate, reorient, and resize. All these techniques are inspired of how paper is used in natural table oriented
collaboration tasks.

The *release* action consist of a user holding a document by touching it with one finger, and when another user tries to take the document the holding user can then either keep holding the document to effectively deny the other access, or the user can lift her finger and thereby release the document and give control to the other user (Morris, 2006). The *relocate* action is performed when a user moves a document from a public or private area and then drops into an area governed by another access control policy. A user can also share a document within his area by *reorienting* it so it faces the center of the table and thereby providing public access to the document. Another way of providing access to documents within the user’s area is by *resizing* it. If the document exceeds a certain size threshold it becomes public; in other words, big documents signal public access and small documents signal private. Morris (2006) found from evaluation that relocate was the most frequently used technique by the participants to share a document. Relocate was also the fastest technique to perform, provided fewest (read 0) misunderstandings when sharing and rated as the technique most easy and natural to use by the participants.

*Orientation* and *position* of documents can be used to signal information of privacy and ownership of documents on the table (Kruger et al., 2003; Morris, 2006; Scott et al., 2004). In the studies by Morris (2006) she found indications that readability of oddly angled text may not be a major consideration for applications where the angled text consists of relatively short strings that remain constant throughout a session. Orientation was instead used as an intuitive method of managing document access and that people use the orientation of paper documents to indicate sharing, these observations are in line with previous work of Kruger et al. (2003), see also section 2.1.1.

Morris (2006) also found support for the strategies of using personal spaces as suggested by Scott et al. (2004); participants often treated areas near the edges as “personal” regions and in the center as a “group” region. Participants further avoided touching regions regarded as other group members’ personal spaces, and thus very few actions were performed outside the users’ own area. Morris relate this to informal social rules and protocol as indicated by the work of Kruger et al. and Scott et al.. Like human territories in general, tabletop personal spaces appear to help people
coordinate their task and social interactions.

An example of how spaces on the table can be used is suggested by Scott et al. (2004); by placing “global state controls” (i.e. privileged objects) in the center of the table the collaborators may be able to anticipate the use of such functionality, providing them with a opportunity to intervene if necessary.

The issue of object ownership has so far avoided the question of whether ownership should be explicit and restricting, or whether it should be indicating and serve to inform social protocols. According to Scott et al. (2004), explicit ownership of tabletop content can provide context by reminding people of who contributed items to the group product. People can also control the level of sharing as they can choose to allow others to “view” an item from his or her personal space, or allow them to modify or copy the item by placing the item in the group space. One example of restricting ownership can be found in the UBI table system (Shen et al., 2003) where visibility and privacy are separated. In their approach private data is not visible or accessible to anyone but the owner of the data, thus providing high levels of protection.

On the other hand, enforced ownership policies may interfere with fluid collaboration due to the extra actions required to grant others access to information (Scott et al., 2004). For this to infer with fluid collaboration requires that the participants are performing some parallel interaction. Rogers et al. (2004) observed that groups seldom interacted with the table at the same time even though they could; “Instead, the group members were very aware of each other and rapidly adopted a turn-taking method of interaction, coordinated with their ongoing decision-making” (Rogers et al., 2004, p.1273, (p.3)).

Morris (2006) on the other hand found different cases of parallel interaction based on the activities performed by the group. Traditionally, object level ownership is often mitigated by social protocol and not restricted by enforced ownership (Scott et al., 2003b), as we’ve also seen in the examples above.

**Tangibles** A digital tabletop may as mentioned also support tangible interaction in which various physical objects, e.g. bricks, balls, cubes and
so forth, are coupled with the digital representation. When a user interacts with the tangible object the digital representation on the table responds accordingly. For instance, a tangible block may be placed on the table to trigger an event, e.g. starting a new program. In another example, moving a tangible puck around on the table can cause digital events to take place. Much work with tangibles is still exploratory, aiming to encourage learning, playfulness and collaboration (Sharp et al., 2007).

Tangibles can be used in order to catalyze sharing and enhance turn-taking procedures. Recall the studies by Olson et al. (2011) mentioned above, where children argued and collaboration broke down as they used a digital tabletop. Olson et al. resolved the turn-taking and collaborative issues among the children by introducing a tangible control. Since the tangible affords removal from the table surface it can be separated from the tabletops territories. Olson et al. explains the success of the tangible in the following way:

“It introduced a familiar idea into an unfamiliar interactive setting, and, allowed children to recall turn-taking scripts: a ‘passing the baton’ protocol” (Olson et al., 2011, p.34)

Multimodality According to Morris (2006) and previously Stewart et al. (1999) the limited display space is a challenge for single display groupware (SDG) and thus digital tabletops, because they often need to display information for many users on one display. Tabletops often cause a lot of “visual clutter” on the screen when groups are engaged in collaborative work. This is partly because most systems only use the visual channel in conveying information to the group’s members. Ways of reducing clutter may include using table-space intelligently and using multimodal information and interaction.

Multimodal interaction follows the ‘more is more’ principle by providing enriched user interaction through different modalities; commonly through touch, sight and speech (Sharp et al., 2007). One example of a tabletop system using multimodal interaction is a digital tabletop using a Shared Speech Interface, as proposed by Piper and Hollan (2009). They used a digital tabletop equipped with speech recognition software to support hearing-impaired patients seeking medical advice from their doctor. As the
doctor spoke the tabletop interpreted and presented the speech in textual form on the table, providing the patient with a chance to read what the doctor said. Not having to look at an interpreter also allows the patient to have more eye-contact with the doctor.

The use of multimodal interfaces as solution to the personal vs. public information problem is a relatively unexplored area. Morris (2006) suggests the use of individually-targeted audio as a multimodal way of supplementing the visual information. They explored this solution in a series of studies where they provided each user around the table with an earpiece so they could access private audio. This was the private setting in the experiment and the public setting provided the group with loudspeakers, causing the group to listen to the same audio together. Results showed that the private setting encouraged all the group members to participate in the task. The public setting on the other hand created a setting in which one or two group members took a leadership role. Interesting to note is that the use of private vs. public audio in Morris’ studies (2006) changes the procedure of the activity and the collaborative strategy used by the group; moving from a parallel to a serial strategy. Additional results showed that groups actually talked more to each other in the private setting. This is probably due to the absence of a shared context (hearing the same thing).

Taking all of the factors mentioned in this chapter into account becomes a huge challenge for design. Thankfully, through the efforts of the researchers mentioned above different sets of guidelines for tabletop design has emerged. These guidelines will, together with a short introduction to interaction design, be presented in the following section.

### 2.3 Designing for Digital Tabletops

“Good software design can help coordinate group actions, encourage equitable participation in group activities, and increase awareness of important events.”(Morris, 2006, p.118)

Interaction design is about creating user experiences that augment and enhance the way people work, communicate and interact (Sharp et al., 2007). This can of course be done in many different ways, and it can be
2.3. Designing for Digital Tabletops

hard to say what interaction design is and what is it not. Here is an example of how interaction design can be defined: “designing interactive products to support the way people communicate and interact in their everyday and working lives” [my emphasis] (Sharp et al., 2007, p.8)

Kim Goodwin (2009) ties interaction design close to a product definition, which is about what functionality a product has. Goodwin’s view on design at large is that it is a craft - which is something in between art and science. One key question for interaction design, according to Sharp et al. (2007) is: how can interaction with a system be improved so that it supports and extends the users’ activities in effective, useful and usable ways? This involves among other things taking into account what people are good at and bad at, and considering what might help people with the way they are currently doing things. Further, one must understand what kinds of activities people are engaged in when interacting with products. Interaction design in the spirit of Sharp et al. (2007) is a wide discipline that is fundamental to all fields and approaches that are concerned with researching and designing computer-based systems. Interaction design is made up of several contributing approaches; academic disciplines (e.g. social sciences or cognitive science), design practices (e.g. graphic design), and interdisciplinary fields (e.g. CSCW). Several methodologies and approaches have been developed for designing, e.g. the Cooper goal-directed design approach (Goodwin, 2009). A designer’s toolbox contains a wide array of methods for generating and developing ideas and concepts. Some often used methods are affinity diagrams, requirements, personas, scenarios, storyboards, rapid paper prototyping and many more (Sharp et al., 2007). Which ones to use varies based on tradition and overall methodology, for instance the Cooper tradition emphasizes the uses of personas and scenarios, whereas Sharp et al. emphasizes prototyping and evaluation.

When designing for face-to-face collaborative environments it is important to realize that non-verbal communication plays an important role in augmenting conversation and collaboration (Sharp et al., 2007; Gutwin and Greenberg, 2000). One challenge for interaction design and CSCW is to design collaborative technologies to help co-located groups to communicate and work together when creating and sharing content. When designing coordination mechanisms it is important to consider how socially acceptable they are to people, according to Sharp et al. (2007). Otherwise, such
mechanisms may result in the users ignoring them, or worse abandoning the entire system. It is a question of balance between human and system coordination - a balance of social and technological protocol (see sections 2.2.1. As (Sharp et al., 2007, p.164) puts it: “Too much system control and the users will rebel; too little control and the system breaks down.”

2.3.1 Applying Theory to Design

Research theories, models and frameworks are used to inform and guide design (Sharp et al., 2007; Halverson, 2002; Hollan et al., 2000). Over the last 30 years, numerous theories have been used in human-computer interaction. Primary cognitive, social and organizational theories have been used. One of the main benefits of using such theories within interaction design is that such theories help to identify factors relevant to the design of a product. Recent models used within interaction design are user models that may predict what information a user wants in a given situation.

Hutchins, Kirsh and Hollan (2000) argue that the relevance and the value of the distributed cognition framework will increase along with the field of ubiquitous computing. The framework should be used when designing, evaluating and analyzing human computer interaction systems and artifacts. Especially since distributed cognition is concerned with how artifacts are actually used, not just how they are intended to be used. The distributed cognition approach can have many benefits for design, as ethnographic data gathering techniques can find new uses for old strategies, and find that techniques effective in one setting may be transferred to another. Due to the importance of the environment, cognitive ethnography (a collection of data collection and analysis techniques) has been developed as the most commonly used method for investigating cognition. Important tools in the ethnographer’s toolbox are methods for observation and interview. Halverson (2002) describes it as practitioners go to where the action is and observe how things really work. Researchers are then confronted with how (well or poorly) reality maps onto theoretical constructs.

Distributed cognition can be very useful for design due to its commitment to examining a broader socio-cultural-technical system, which is necessary for the collaboration between individuals mediated by artifacts (Halverson, 2002).
2.3.2 Design Principles and Guidelines

Years of experience together with scientific theories have led to some well established principles, i.e. generalizable abstractions, for design that are used by designers to aid their thinking during the design process (Sharp et al., 2007). More domain specific guidelines for designing digital tabletops have also been developed (Morris, 2006; Scott et al., 2003a). Let’s first consider some general principles for design.

**Design Principles** Design principles are derived products from a mix of theory-based knowledge, common sense and design experience (Sharp et al., 2007). Here follows a quick overview of the most commonly used design principles as described by Sharp et al.; these are: visibility, feedback, constraints, consistency and affordances.

The *visibility* principle predicts that the more visible a function is, the more likely it is that a user will be able to know what to do with it. For instance, knobs and switches are highly visible and intuitive devices for operating the lights in a room; whereas ‘activating zones’ are invisible and ambiguous. Related to visibility is the concept of *feedback*. Feedback is about sending information back to the user about what action has been done and what has been accomplished, which is central to coordinating an activity. Imagine if there was a delay when writing with a pen so that the ink appeared several seconds after the writing action was completed. That would be unbearable. Feedback can come in different forms such as: audio, tactile, visual, and in combinations.

*Constraints* are about restricting the different kinds of user interaction that can be performed at a given moment. One common way of doing this is by “graying out” choices in a menu when they are not available. *Consistency* refers to designing user interfaces where similar operations are used for achieving similar tasks. For instance, if using the left mouse buttons selects one type of icon, than using the left mouse button should select other types of icons as well.

Last but certainly not least is the concept of *affordances* that has a variety of meanings attributed to it, since it’s a concept used in different fields (Berglund, 2009). In the design field, an affordance is, according Sharp et al. (2007) an attribute of an object that allows people to know
how to use it; in this sense, “to afford” means “to give a clue”. This idea was unfortunately overly used and became a catch-all phrase, loosing much of its potency as a principle.

Problems can occur when applying more than one of these principles at the same time, since they often require trade-offs (Sharp et al., 2007). Some of the principles are also in themselves hard to apply, for instance rigid consistency must often be abandoned.

Tabletop Guidelines  In addition to pure design principles there are more specific guidelines for when designing a specific type of product or service. Scott et al. (2003a) have studied digital tabletops and come up with a general set of guidelines for designing tabletop applications and tabletop workstations. The guidelines state that digital tabletops must support 1) interpersonal interaction; supporting the fundamental mechanisms (Gutwin and Greenberg, 2000) that people use to mediate collaborative interactions, see section 2.2. It may not be necessary for co-located groupware to explicitly provide software support for each mechanic, but technology must not interfere with them. Interfering with these interactions can cause breakdowns in collaboration.

Further, digital tabletops must 2) support fluid transitions between activities; for instance, it’s important that all activities use the same input methods to ease the transitions between them (Scott et al., 2003a). In the same manner, tabletops must 3) support transitions between personal and group work. This guideline is very much in line with Morris (2006) discussion on private and public access to digital object. Providing users with distinct private areas on the table may facilitate the transitions between individual and group work. The next guideline states the need for 4) supporting transitions between tabletop collaboration and external work that exists in the environment beyond the tabletop boundaries (Scott et al., 2003a). This often involve supporting other types of none off-the-shelf software and providing possibilities to move information from the table to other computers (e.g. by using hyperdragging (Rekimoto and Saitoh, 1999)).

Tabletops must also be able to 5) support the use of physical objects on the table’s surface - tangibles (Scott et al., 2003a). The tabletops must support the familiar practice of placing both work-related and non-work-
related objects on tables, as well as providing additional digital features. This may allow users to apply the years of experience they have accumulated collaborating around traditional tables. In relation to the tangibles, the table needs to 6) provide \textit{shared access} to the physical and digital objects. Depending on the collaborative activity, participants may work with a single object, such as one large design sketch, or they may be working on a series of related objects. One must also 7) consider the \textit{appropriate arrangements of users}, as people can sit or stand around the table at different locations. This can cause a tradeoff between optimal viewing distance and comfortable personal distances.

The final guideline 8) states that tabletops must handle \textit{simultaneous user actions}. Collaborative activities can take on a variety of collaboration styles according to Scott et al. (2003a). The styles include working in parallel, working sequentially in tightly coupled activities, and working under different assumed roles (Scott et al., 2003b). Working individually can also be a sort of collaboration, as independent interaction is coordinated with a group; this is referred to as a divide-and-conquer collaboration style. Morris (2006) also observed how groups may use different interaction strategies based on the activity, see section 2.2.2.

An additional set of guidelines was established by Morris (2006) based on their experience developing and studying tabletop systems. These guidelines are grouped in accordance with seven main themes, each consisting of a set of fine-grained guidelines. The guidelines reflect the main themes presented in section 2.2.2. Here the most relevant themes will be presented\footnote{For the full set of guidelines, see Morris (2006)}.

The first group of guidelines Morris presents concerns \textit{table regions}; the center region of the table should be open for group use; visually designating personal and group regions aids users in understanding the different access permission policies; using copies of interaction elements and table contents can ease proxemic concerns; providing a region for “trash” items is often desired.

Related to regions on the table are the different \textit{access permissions} of items and regions. Permissions should be dynamically changeable which can facilitate shared document inspection and alteration. Restricted ac-
Background

Accessibility allows users to “guard” their data from other user accidental or intentional “interferences”. Further, making these accessibility visible can prevent confusion.

The next group of guidelines provides ways of managing group dynamics. The tabletop can provide explicit awareness information to help regulate participation levels, for instance by using subtle audio feedback or individually-targeted interaction visualizations can convince users to change their levels of activity. The placement of objects encourages different users to take responsibility for them. Users should be prevented from accidentally (or sometimes intentionally) dramatically altering the state of the tabletop without group consent, as this interrupts the workflow.

Groups may take on different work styles based on the activity. Combining private and shared data allows for smooth transitions between closely- and loosely-coupled group work (serial and parallel strategies). Other guidelines not explored here includes ways of managing and reducing visual clutter, utilizing cooperative gestures, and how to measuring usability - with focus on team processes rather than task completion speed.
Chapter 3

Method

A big part of the methodology leading up to the design concepts is the literature review performed, as reflected by the background chapter. This thesis shoulder a theory heavy approach to design with additional elements of ethnography and design methodology. The thesis is thus shaped by the presented theories of distributed cognition, dialogical theory and the language use theories by Herbert Clark. These theories are used due to the tight coupling between the tabletop and the users through interaction, and how they as a system is to structure turn-taking. Therefore it’s appropriate to view cognition, coordination and communication as something existing in-between the users, and in-between the users and the tabletop (and other present artifacts, such as tangibles).

This chapter will however focus on the undertaken ethnographic studies and the design methodology used within this thesis will be presented. First, the ethnographic studies at the two domains will be presented, also describing how the ethnography contributed to the resulting model on collaborative turn-taking. The overall work process is reflected in figure 3.1. In the figure many of the boxes representing the different phases of the work are not entirely sequentially placed, since many of the phases were overlapping. Often, backtracking to a previous phase was a necessary step.
3.1 Ethnographic Studies

In order to gain an insight in the daily work and activities of potential users of digital tabletop technology, and to understand the basic requirements for tabletop turn-taking, two ethnographic field studies was performed. The fieldwork was conducted with a twofold purpose; 1) how to interpret and think about collaborative turn-taking and activities; 2) to gather information and inspiration for the design of turn-taking concepts. Data gathered from the studies were coded and described using a grounded theory approach, influenced by previous work and theories.

3.1.1 Data Gathering

Ethnographic methods are often used as alternatives to laboratory experiments, and many groupware researchers advocate the use of such methods to consider culture and context. ethnographic methods were chosen since they are essential components in the cognitive scientist’s toolbox (Hollan et al., 2000). Ethnographic methods are also valued in interaction design where they are primarily used early in the design process in order to un-
3.1. Ethnographic Studies

Understand the users’ context, tasks and goals (Sharp et al., 2007). However, due to limited time for data gathering in design projects designers often shorten the ethnographic process. Shorter ethnographic studies such as these are sometimes called ‘quick and dirty ethnography’; in which brief ethnographic studies are undertaken to provide a general but informed sense of the setting, which is valuable for designers according to Hughes et al. (1994). Motivations for using ethnography within CSCW is basically twofold according to Hughes et al; 1) collaborative technology often fail because of lack of attention to the social context of work; 2) ubiquitous and distributed computing pose new problems for design which require the development of new methods which analyze the collaborative and social character of the workplace and its activities. During the course of this thesis two shorter ethnographic studies were conducted; one study at a local Intensive Care Unit (ICU) with focus on observationally gathered data, and a second shorter interview series at a local employment office.

Intensive Care Unit The first pre-study were conducted over one work week (48h) and contained observations together with unstructured interviews at the intensive care unit in a Swedish hospital. The focus of the observations was on how the clinicians collaborated with each other in different situations. For the most part I played the role of a passive observer, but during some situations I became involved in assisting with some basic task. The observations were recorded through pen and notebook since neither video nor audio recording was allowed or very practical because we moved around a lot. The observations were supplemented by interviews made continually as time and workload allowed them to be made. A lot of observational focus was placed on modalities of communication and coordination during an activity involving more than one participant.

Employment Office A series of three interviews were performed at the local employment office in order to acquire a general understanding of the work at the employment and how the job seeking process may look like, as well of how digital tabletop can be utilized by an employment office and ultimately how turn-taking can be structured in those possible situations. The interviews were of a semi-structured nature and three job coaches
participated. Areas of focus were threefold; understanding the process of a job seeker, current and future support for job seekers with little or no computer skills, and finally the role of a digital tabletop in supporting these processes. The interviews were mostly a source of inspiration due to the lack of supplementary observations.

3.1.2 Data Analysis

Data gathered from both studies was examined and categorized using a grounded theory inspired approach, an approach which is growing ever more popular in the field of interaction design (Sharp et al., 2007). Grounded theory encourages the practitioner to systematically analyze collected qualitative data in order to form new terms and concepts to describe the complex processes of the real world. After the process of coding and interpreting observational data other theories and previous work from the academic community can be considered, and synthesized with the interpretation. The end result of grounded theory is a theory that fits the set of collected data.

Central to grounded theory is the process of data coding; in this thesis, coding can generally be divided into two main points of interest; turn-taking and activities. The interpretation of the data was infused and synthesized with ideas and concepts from previous work with digital tabletops, theories of communication, and theories from the fields of cognitive science and linguistics. The grounded theory is derived during the ongoing process of gathering data and interpreting data, and the theory itself becomes a way of interpreting new gathered data, as illustrated in figure 3.1. In this way, the view of turn-taking, presented in this thesis, gradually evolved as data was gathered and analyzed.

During the development and coding of the theoretical perspective, and later the concepts themselves, a lot of mind maps and brainstorming procedures were used. An illustration of how such a mind map could look like is presented in figure 3.2\textsuperscript{1}. Gradually, the ideas and interpretations, presented in the following section, were fleshed out and refined.

\textsuperscript{1}Generated with the excellent online tool Wordle\textsuperscript{TM} - \url{http://www.wordle.net/}
3.2 Design procedure

According to Sharp et al. (2007), the process of conceptual design is concerned with developing a conceptual model that captures what a product will do and how it will behave. During the course of developing the design concepts no formal process model (such as the waterfall model) were followed. Rather, I used a mixture of methods, borrowing ideas from interaction design methods, guidelines, principles and procedures which I deemed useful and appropriate. All of the concepts are based in the inspirational theoretical perspective presented in the previous chapter. Important to my view of turn-taking is that any collaborative action be it communicative or otherwise takes place within an activity. Every design concept is therefore designed with one or more tabletop activity-types in mind. The concepts presented in the forthcoming sections are either explicit interactive components coordinating turn-taking around the digital tabletop, or more subtle interactive mechanisms and layout choices that may coordinate interaction around the tabletop. Many of the design methods are not directly applicable since the aim of the thesis is not to design a complete system or
product. Rather, the designs should be seen as parts of a future library of interaction concepts for managing participants using a digital tabletop.

There are many requirements for how the interaction techniques are to be shaped, e.g. the concepts are to be used for surface interaction on a digital tabletop. Many requirements are implicitly described through the perspective of collaboration and turn-taking used as given by the background. Using a digital tabletop comes with many requirements of its own; for instance a user-control (such as menus or buttons) should be placed and shaped so that it is within reach for the typical user. In order to structure and make the turn-taking model more vivid and inspirational for the design process, a set of simple scenarios were used. According to Bødker (via (Sharp et al., 2007, p.554)), scenarios can be used to aid and structure the design process. The scenarios used are basically hypothetical group activities taking place around a digital tabletop. Even though scenarios are hypothetic situations, they rest upon made observational data and previous work. All scenarios are based on the different types of activities identified, see section 4.1.2. Examples of scenarios will be presented in shorter forms together with the design concepts in chapter 4.

The scenarios were further elaborated on using storyboards; a storyboard is a sort of low-fidelity prototype (Sharp et al., 2007), which purpose is to explain the design concept. A storyboard is a scenario illustrated with simple sketches together with a descriptive text, the storyboards will be presented in chapter 4. During the design of the concepts several affinity diagrams were used to “map” each design idea, in order to get an overview and estimation of the ideas’ potential value and usefulness. A affinity diagram is often used as a way of grouping ethnographical data in order to inform the designers, often the affinity diagrams are made up of a large collection of post-its each representing some observed variable. The affinity diagrams used provided an overview of what aspects of turn-taking a particular design idea could be said to use or affect. The resulting design concepts and their corresponding storyboards will be presented in chapter 4.
Chapter 4

Results

In this chapter the model on collaborative turn-taking and the resulting design concepts are presented, together with a short review of the design process. The first section describes the model which lays the ground for the concepts. The following section introduces the five final design concepts in the form of design sheets and storyboards. Lastly, one of the design concepts will be presented as a higher fidelity mockup.

4.1 Conceptual Grounds and Design Space

Humans are generally good at conversational turn-taking, although this of course may vary from individual to individual. In our everyday life we fluently and rather effortlessly (for the most part) communicate with our fellow friends or colleagues. Generally, it would probably be difficult and rather fruitless to try to design support for conversation turn-taking for co-located collaboration, although there are examples of how a process like this could be supported for special needs groups (Piper and Hollan, 2009).

Turn-taking protocols (either social or technological) on the other hand describe the turn-taking processes when interacting with shared artifacts and are a means of structuring and coordinating group collaboration. These types of turn-taking procedures are more targetable for supporting interface
components and mechanisms at the digital tabletop. The breakdowns that are occurring in face-to-face collaboration when people are working with digital tabletops can further be related to breakdowns in turn-taking, see section 2.2.2.

In the following sections a model of tabletop related turn-taking and activities will be presented. The model represents my interpretation of turn-taking; a view based on ethnographical studies, previous work and theory from the previous chapter.

4.1.1 Turn-Taking Models

The view of turn-taking presented here divides turn-taking into two levels; communicative and collaborative. Turn-taking at the communicative level can be seen as a coordinator of communication, and turn-taking at the collaborative level can be seen as a coordinator of group collaboration in general. However, where the one ends and the other begin is somewhat blurry as they often are intertwined. Important to both levels of turn-taking is the activity in which they occur, since the type of activity performed shapes among other things; how artifacts are to be shared in the group, how procedures and labor can be divided amongst the members of the group, and sometimes communication within the group. Important to note is that this description of turn-taking is not an analytical model of turn-taking, rather a descriptive and design-inspiring framing of turn-taking. This description of turn-taking is also skewed towards turn-taking during activities at the digital tabletop, as this is the purpose of this thesis. Further, most of the focus is placed on the collaborative turn-taking level.

Communicative Turn-Taking

In the abstract, the phenomenon of turn-taking seems quite easy to define. The talk of one party bounded by the talk of others constitutes a turn, with turn-taking being the process through which the party doing the talk of the moment is changed. (Goodwin, 1981, p.2)

This level is concerned with how we take turns during communication, and how different types of communicative acts (explicit and consequential)
can coordinate turns at the digital tabletop. When defining a model for turn-taking on this level the choice of communicative theory will play a big part, whether it’s a traditional monological approach or a dialogical approach.

Turn-taking at what I call the communicational level has been well studied, and it therefore exists well defined models for these communicative mechanisms. The traditional model which is still commonly used is the systematic model by Sacks et al., see section 2.1.3. This conversational level of turn-taking can be said to follow some guiding principles for the organization of turn-taking. A turn at this level is very graspable and concrete. Sacks’ model can be defined in four concrete steps according to Linell, in contrast, what defines a “turn” becomes more abstract at the next level of collaborative turn-taking. However, communicative turn-taking does not only occur in conversation form.

People tend to classify turn segments from a number of sources, all the way from gaze to facial gesture to body stance. One example of a more encompassing communicative turn-taking model is suggested by Thórisson et al. The traditional roles speaker and listener used in turn taking literature is to narrow, it excludes other communicative channels. The terms content presenter and content interpreter fits better at the communicative level of turn-taking than speaker and listener, since the communicative level indeed includes more than conversation. The so called back-channel serves an important role in regards to communicative turn-taking, as it allows for the content interpreter to give feedback to the content presenter that indicates that he or she follows and understands what the presenter is trying to get across, without interrupting the turn.

One design issue is how to use interaction with the tabletop as means of producing appropriate implicit communication. For instance, using informative gestures when interacting with the table could provide a clear and open way of interaction that also serves the purpose of coordinating (communicating with) the group. For instance, moving an object should be performed with a clear natural moving gesture, not by the click of a button.

Indeed, communication in general at the ICU seems to be very opportunistic, using a range of available communicative channels; as it often involved pointing, gesturing, gaze following, speech, and artifacts in their
surroundings. For instance, they often used the mouse pointer to highlight information, often in combination with speech acts e.g. “look here”. They further made gestures using the mouse pointer, not circling an important set of data per se, rather circling as a gesture in order to emphasize a point they were trying to get across in speech. These kinetic, spatial and pointing gestures are examples of two of Gutwin and Greenbergs’ mechanisms of collaboration; explicit communication and consequential communication. People both communicate explicitly with each other (pointing, speech etc.) as well as pick up consequential communicative acts (as others interact with artifacts or with each other - as Gutwin and Greenberg puts it: “information given off by the characteristic actions of a person's embodiment in the workspace”.

Other forms of communication at the digital tabletop might be interactive graphical communication - that is, communicating through the layout of the program, the spatial arrangement of the content or the interactions of the other participants can inform participants and the turn-taking process. Mondada showed that pointing can be used to project turns, and that pointing practices and turn-taking practices can be deeply embedded. This was also observed at the ICU where turns in communication and collaboration were projected or seized by either gesturing or pointing.

The ongoing activity will influence communication. A well known distinction is between informal and formal communication, this is also used as a distinction of an activity itself by Clark. Formal communication may take place during a board meeting; during such activities, roles and types of conversational turns are often decided upon in advance. It’s common during these activities to have one person managing the turns - that is managing who speaks when, and what they are allowed to talk about. In contrast, informal communication is the type of communication going on when people socialize; during which the conversational turns are guided by rules such as those described Sacks et al.. At the ICU, both informal and formal communication seems to take place as the activities performed range from formal to informal.

Overall, communication is not a process that can be viewed separately from other cognitive and social processes, especially from the view of dialogical theory, distributed cognition and through the language theories of Clark. Therefore, these two levels of turn-taking presented here will not be
mutually exclusive, and some of the mechanisms and processes discussed in the following section could arguably be placed at the communicative level. This is due to that it’s not easy to define which actions should be considered communicative and which should not be. With that said, let’s move on to collaborative turn-taking.

**Collaborative Turn-Taking**

T: *J. said I could have a turn!*
J: *No, I didn’t!*
- (Transcript excerpt from Olson et al., 2011, p.32)

Non-communicative turn-taking is used to structure numerous activities in our daily lives. When we play a game of backgammon, stand in line at lunch or when we are sharing a computer we are performing various acts of turn-taking - various coordinating acts. Even at the digital tabletop collaborative activities may vary; for instance, one can imagine that turn-taking mechanics used during a formal meeting using a digital tabletop may be different from the turn-taking mechanics underlying an open brainstorming session. Although turn-taking can be said to take place in all of these activities, it seems as the role of turn-taking within the activities varies. These are very different collaborative situations or activities, governed by different social and situational protocols. This type of social turn-taking process involves a diverse set of cognitive processes or functions, ranging from low level functions, such as perception and memory, to higher level reflective processes, such as planning and reasoning. Moreover, cognitive processes are often interdependent and seldom appear isolated.

This type of collaborative turn-taking is not easily defined by general “rules” or guidelines for how one can or should take turns during collaboration; dialogical turn-taking may as seen be described in a set of rules. Collaborative turn-taking includes or borders to a wide variety of collaborative aspects, such as: coordination of private and public objects, social protocol, planning, collaborative schedules, situational protocol, and group dynamics to mention a few. This leaves us with a complex space of possible factors to consider. Some of these are included in Gutwin and Greenbergs set of collaborative mechanics - which from my point of view all play impor-
tant roles in collaborative turn-taking. In general, collaborative turn-taking is a way of coordinating who does what when during a collaborative activity.

At the ICU it was common that artifacts and objects were shared by participants. Often artifacts could be used by more than one person at a time and sometimes the clinicians had to take explicit turns as they passed the artifact back and forth. In some situations there were obvious “turn-leaders”; for instance, during a patient round were three clinicians partook. One of the clinicians stood for most of the conversation and further used a checklist in order to structure the round. The fact that this clinician held the checklist was an indication that this clinician “had the collaborative turn” in the activity structured by this external checklist. When this clinician was interrupted by a phone call the entire round came to a halt, but moments earlier the others were on the phone and the round did not stop. The collaborative activity was depending on this clinician in order carry the activity forward, this reflects some aspect of governance for collaborative turn-taking.

This view of collaborative turn-taking has much in common with Clark’s theories of language, in which language is a type of joint action and joint actions are deeply intertwined within joint activities that further shape the actions themselves. Clark’s view provides a framework of anchoring communication to action, and action to collaborative activities. The digital tabletop and its’ interface provides an instance of common ground for the current activity through external representations. In many ways, these external representations may be used to support the turn-taking processes (as with the checklist in the example above). An interesting question is how collaborative turn-taking relates to how activities in Clark’s view are made up of a hierarchy of projects and sub-projects. One can draw parallels between a project or sub-project and a “turn” for some types of activities. Bangerter and Clark demonstrated that conversation can be used as means for project navigation which is related to collaborative turn-taking. From the collaborative turn-taking point of view, project navigation can also be achieved through other actions and external representations on the digital tabletop.

This level of turn-taking borders other processes coordinating collaboration, for instance the structuring and planning of a collaborative task. The act of planning an activity can be organized and explicit during an a priori
stage of the activity (i.e. by some sort of planning process), or this can be done “on the fly” while the activity is ongoing - explicitly or implicitly. Whether activities are planned or structured in advance or structured on the fly can affect how turn-taking mechanisms are to be utilized during the activity. This is one of Gutwin and Greenbergs collaborative mechanisms, as they describe that during on the fly planning people tend to divide and re-divide the activity as they go along - for instance by reserving areas of the workspace for their personal use which further relates to personal spaces and object accessibility.

Furthermore, turn-taking mechanisms can vary from explicit to implicit in how they indicate “turn ownership” (governance) and how they regulate participant roles. When using a explicit turn-taking protocol it is explicit whose turn it is to act, and the user would in this way be assigned the role of “turn-owner” or “turn-controller”. An implicit turn-taking protocol only implies that it is a specific person’s turn to do something in a more subtle way. Gutwin and Greenberg’s notion of monitoring becomes an important cog during activities where turn-taking takes on a more implicit form. The distinction between explicit and implicit is however somewhat floating, and many turn-taking protocols may find themselves somewhere in between.

At the ICU examples of both explicit and implicit turn-taking acts were observed. For instance, a clinician could during patient rounds implicitly pass on the turn during rounds by calling the name of another patient - thus passing on the turn to the clinician responsible for this patient’s care. During other meetings the control of turns were controlled more explicit; for instance by seating one participant at the end of the table with access to the only computer, making this participant the overall turn-controller.

Closely related to explicit and implicit turn-taking protocol is the notion of whether the interaction is somehow restricted by the turn-taking protocol. If the protocol restricts the interaction the protocol can be said to be regulating, otherwise it’s open. Levels of regulation may vary depending on the ongoing activity and group performing it. One example of how interaction can be regulated is by only allowing one person at a time (the turn-owner) to interact with the tabletop.

Further, the regulation of interaction will affect how people are able to assist each other, which is one of the collaborative mechanics according to Gutwin and Greenberg. Opportunistic and informal assistance may be
hindered as interaction becomes restricted and groups must rely on more explicit forms of assistance and assistance requests. Interaction regulation also affects another collaborative mechanic: levels of protection. Higher level of protection for the turn-owner may follow from regulated interaction, depending on activity and tabletop layout.

To “have the turn” can be to interact with a certain shared object within an activity, or it can be to have the turn at controlling the entire activity. This introduces another type of turn-taking distinction within the collaborative level of turn-taking, and that is the distinction between activity and object level turn-taking protocols. Activity turn-taking is a way of globally coordinating who are to interact with the digital tabletop at any given time. This is relevant during sequential activities where users are to interact with the tabletop one at a time - i.e. taking turn using the table, similar to how people have to share the mouse when collaborating using one traditional desktop computer.

This also relates to how and when global state changes are allowed to occur, as previously discussed by Morris - see section 2.2.2. Loosely one could say that taking turns at the activity level is similar to taking turns being the “leader”. Turn-taking of object-type on the other hand coordinates whose turn it is to interact with any certain object placed on the table - whether tangible of digital. During tasks with multiple shared objects it’s not always enough to rely on social protocol to govern the interaction and turn-taking of shared digital objects. Object turn-taking thus relates to how private and public objects are to be managed by the system and also levels of protection for different objects. Again, the distinction between activity-type and object-type turn-taking is not entirely mutually exclusive; it can be a bit of both. Consider an activity where there is only one shared object on the tabletop, given this scenario both types of turn-taking will be more or less the same.

In many of the suggested activities and designs in the forthcoming sections there is a re-occurring (almost haunting) issue of whether the turn-taking mechanism should consider a layout that is based on personal copies of the content representations, or one shared group representation. When working with a shared representation it’s often located at the center of the table and may sometimes be movable. Using this setup user must often take turns when interacting with the representation and the activity therefore
reflects a sequential work strategy. When using copies of the content every user has a private copy of the content that they interact with freely; reflecting a parallel strategy. However, in many cases the copies can be linked to each other, causing interactions made with one copy to be reflected in the other copies. When using a linked setup, turn-taking procedures for who is to interact with the linked copies needs to be considered - i.e. activity level turn-taking. When the copies are not linked to each other a more parallel work strategy may be used.

With the turn-taking distinctions laid on the table it’s time to describe what possible activity-types a digital tabletop can be used for. Activities have been a re-occurring theme during the description of turn-taking here and in the literature, activities will therefore be important to take into account when designing support for turn-taking.

4.1.2 Collaborative Tabletop Activities

One assumption of collaborative turn-taking is that different types of joint activities are coordinated in different ways. Turn-taking mechanics used during a formal meeting are different from the turn-taking mechanics underlying an open brainstorming session - even though both activities are collaborative group activities. Coordination of conversation often remains the same in most activities, except for activities with strict regulations for conversation - e.g. in the court room. In order to develop design concepts for higher level turn-taking mechanisms the activities the mechanisms serves to coordinate must be identified and described.

There are many factors that define the nature of a collaborative activity. Clark characterizes activities with five dimensions or properties and McGrath and Hollingshead divides tasks in four major categories as presented in section 2.1.4. Clark’s dimensions are: scriptedness, formality, verbalness, cooperativeness and governance. This characterization is used to describe activities in general, whereas the focus here is on activities that can be performed with the digital tabletop. Clarks’ dimensions are usable for describing digital tabletop activities, but have been supplemented with the dimensions: purpose, parallelism and groupsize. These were added in order to better describe activities centered round a digital tabletop.
These additional dimensions will be described in the paragraph below. When presented together, these dimensions form what I will call an activity map. The activity map can be visualized with a simple radar chart, see figure 4.1.2. Important to note is that only one of the end-values is explicitly represented in the radar chart. The other end-value for each dimension is placed in the center (origo) of the radar chart.

**Figure 4.1** Example of a radar chart for visualizing activities

Additional dimensions According to McGrath in section 2.1.4, an activity can be cognitive or behavioral. I’ll use the terms epistemic and pragmatic activities inspired by Kirsh and Maglio, see section 2.1.4. From my view, epistemic activities are concerned with distributing and discussing information within the group in order to establish common ground to inform and simplify future actions and activities, whereas pragmatic activities are about doing or creating something. However, activities are neither purely pragmatic nor purely epistemic - it’s more likely that they contain a bit of both. This dimension I’ve chosen to call purpose, which chimes close to Clark’s notion of a dominate goal of an activity (as goals can take pragmatic or epistemic forms).

An activity can be viewed as constituted of a set of projects and sub-projects. The projects can sometimes be performed in parallel, but must sometimes be performed in a specific sequence. I suggest that the potential parallelism of an activity will be a useful dimension when describing activities at the digital tabletop for two reasons: 1) the layout of the interface can be shaped after how parallel the work is, and 2) the turn-taking coordination mechanism and amount of turn-taking required will differ with the parallelism of the activity. If the activity and all the actions within it can be performed completely in parallel, then collaborative turn-taking would not necessarily be required. When I use the term sequential it should be thought of group-sequential, i.e. that the entire group will focus on a certain part of the task together, and then move on the next. This can
be somewhat related to Clark’s dimension of governance, which describes to which degree an activity is dominated by one of the participants. Parallelism can be viewed from two different angles; whether the structure of the task is parallel and whether the user input can be conducted in parallel or not. When an activity can be performed in parallel, more users can participate by performing different parts of the activity when interacting with the digital tabletop at the same time. This relates to Hutchins characterization of sequentially constrained and sequentially unconstrained procedures. According to Hutchins, it’s often desirable to organize and represent procedures in a way that enables them to be performed in a sequentially unconstrained (parallel) manner since parallel action if often effective and requires less coordination. In Morris’ view, parallelism and tabletop activities can be further divided into parallel, serial and assembly-line strategies, as seen in section 2.2.2. From my perspective, these relate to which turn-taking styles are in use, such as object-level versus activity level turn-taking. Generally, the parallelism dimension reflects the parallel and assembly-line strategies which involve high levels of parallel interaction, and further that the activity itself allows for this type of interaction.

Another important factor of group collaboration is the size of the group, especially in relation to a digital tabletop due to the limited space available. This may not be a dimension of the activity per se, but some activities are more often performed by large groups rather than small groups - for instance a game of chess is seldom played by other than two players, and a league game of football (soccer) is played by 22 players. Although, we still call it football even when six players are playing - used carefully, groupsize can be an informative dimension of an activity.

This characterization is made for activities that may be performed at a digital tabletop. It’s not an exhaustive or mutually exclusive list - but it can be a useful tool for inspiring design and something to ground the designs in. It’s a tool and inspiration source for design, based on previous tabletop studies and my observations. These are activity classifications are made with the purpose to aid and inspire design based on data from observations and from literature. During the following pages I’ll present some common activities that can be performed with a digital tabletop.
**Presentation Activities** At the ICU, one recurring activity is the morning round where groups of clinicians’ distribute information regarding patient status and other important news relevant to the day-to-day operations. The major features of these types of activities include that they are performed by a relatively big group - or at least an activity with many side-participants or bystanders, rather sequential in their structure, there is high level of verbal activity, and they are autocratic from the view that one person at the time present information while the other participants listen - although the presenter may vary during the course of the activity. A serial strategy for interaction would be suitable for these types of activities.

Taking the turn in these situations often involve first requesting it - sometimes a turn-change is implied by some utterance, i.e. “next patient!” which refers to the next patient on the list which a certain clinician has information on - and thus this clinician gets the turn to present. During the interviews at the employment office this type of big group presentation activity was suggested as one which could be supported by a digital tabletop and other visual augmentations. The goal of these activities are often more informative or epistemic rather than pragmatic. If the activity is to be carried out using only a digital tabletop the groups cannot be too big if everyone is to see the table and have a chance of interacting with it. For pure presentation purposes an external vertical display may be used to support bigger group presentation - using the tabletop as a sort of presentation control. Turn-taking in these activities are involved with structuring who is to interact with the table at large, because of the high sequential, epistemic and autocratic nature of the task. A turn in this activity can reflect to the entire presentation by one participant. One specific problem is how to let the non-presenters illustrate something by interaction when they are for instance asking a question. Turn-taking is also concerned with what type of activity can be performed by the observers.

![Figure 4.2 Radar chart for presentation type activities](image)
4.1. Conceptual Grounds and Design Space

in parallel with the presentation (annotations for example).

**Group Meeting Activities** The distinction between a meeting and a presentation activity is not definite. This is much due to that meetings in themselves are hard to define and restrict. Meetings can range from formal to casual, from autocratic to egalitarian and the purpose of the meeting may vary. One signifying feature is that one or more participants pass information on to the other participant. For instance, the one participant informs and helps the “content receiver” about jobs (AF), sells a product to the receiver, or perhaps gives the receiver medical advice. Thus the activity contains a balance of epistemic and pragmatic elements, not as pragmatic as the joint work activity but less epistemic than presentations. Interactions are usually performed by one participant at a time, reflecting a serial interaction strategy. In this case, I’m focusing on the semi-formal doctor meetings (sometimes called the doctor round) observed at the ICU. During these doctor rounds, patient care and patient treatment plans were discussed. The differences between these meetings and the morning rounds include smaller groups and a less autocratic structure. During these meetings each doctor presented the patients in their care to the others and together discussed future treatment plans. When they were finished discussing a patient the current clinician or one of the participants moved on to the next patient, often through some verbal signal. In one example a participant called out the name of the next patient - “Erik?” (fictional name). The change of patient often also meant a change of the clinician who presented the patient, and interesting to note is that the change of patient was often invoked by someone other than the one who presented the previous patient. In other meeting situations participants can take more clear roles. In one observed meeting among clinicians it was obvious that one of the clinicians had a
Results

dominate role and controlled the artifacts that were used - this clinician had the turn most of the time during the activity. At one point the clinician having the turn was forced to answer the phone which lead to a freeze in the meeting; the other participants idled while waiting for the clinician to finish the call.

The customer meetings at AF are other good examples of this type of activity. The customer meeting activity is less scripted then a presentation activity, but is more sequential and more autocratic than a “typical” group meeting since one of the participants has a clear dominant role throughout the entire activity. The job coach may for instance guide the job seeker in using the computer system or give the job seeker council in the job seeking process.

**Joint Work Activities**  This is a fuzzy category of activities, referring to any more pragmatic than informative joint activity that can be performed jointly by a group seated at a table. The main difference here probably lies within the goal of the activity as these types of activities have a more pragmatic goal, whereas the previous activities are more epistemic - spreading information and coordinating the group. These activities include group brainstorming ideas, matching music to scenes seen in Morris, or designing a park as seen in Rogers. McGrath probably would place many of the joint work activities in the generative category in his distinction. The role of turn-taking within these activities are most likely to help coordinate whose turn it is to interact with a particular object (this also applies to fully parallel activities), turn-taking on the object level. It can also involve turn-taking at the table at large if the interaction for any reason should be restricted to one person interacting at a time.

These activities can often take advantage of parallel user interaction
more than the previous activities, and these activities are more egalitarian. Activities can be performed more or less in parallel leading up to a joint decision to be taken by the group, fitting the parallel strategy observed by Morris et al. Alternately, the entire task may have to be performed jointly and sequentially by the group, thus reflecting a serial strategy observed by Morris et al. The experiment by Morris, presented in section 2.2.2, illustrates this difference; when the participants listened to public audio the activity was sequential, as they only could listen to one song at the time; when they listened to private audio the activity became parallel since each participant could listen to a different song at the same time. In one observed situation at the ICU a nurse and a doctor were discussing a patient in the patient module. As they went along the nurse ordered tests and entered patient data into various computer systems; the doctors went through and filled in a patient information sheet. The doctor frequently told the nurse what values to enter and instructed on what test was to be ordered. At one point the instructions became too complicated and the nurse handed over control of the computer to the doctor and told the doctor to do it - giving up the turn at the computer.

**Parallel Work Activities** This can be an activity performed completely in parallel with no group synchronization at the end of the task, for instance reviewing a bunch of documents where every participant reads a separate document. The level of turn-taking in these activities is not concerned with structuring user interaction in the same way, since interaction needs to be in parallel. The activity itself can be structured in different ways however - introducing components and mechanisms that help user decide what task to do next. Here turn-taking may also be at the object-level, coordinating use of certain shared objects. Help users coordinate who does what and in what order - for instance by
using piles and areas of the table. Again the example by Morris can be mentioned, how a public sequential task can be turned into a set of private and parallel tasks by the use of multimodality, see section 2.2.2. Many activities also involve sequences of more or less parallel work, for instance when reviewing a bunch of documents the work may alternate between privately reading documents to collaboratively discussing the documents. These activities would fit the assembly-line strategy of user interaction.

4.2 Design Concepts

During the conceptual development guidelines and design principles were used to guide the design process. Through a process of iterating different concepts the following five ideas remained; the token concept, the list component concept, the interactive gesture concept, the turn-passing concept, and the copy-mechanics concepts. These concepts together illustrate the different aspects and ways collaborative turn-taking can affect and coordinate activities and collaboration when groups are using a digital tabletop. Many of these design concepts make use of different table regions, these regions should be visually indicated in accordance with the guidelines.

**Design Sheets** All designs are presented as “design sheets”, which basically is a page describing and mapping the design to the model of collaborative turn-taking. The idea behind the design sheet is to combine a general description of a concept together with the theoretical reasoning related to the concept and fit in into a single page. The design sheets describe four main aspects of the designs; first, a general *description* of the purpose and idea behind the concept; secondly, an *example* of how the concept can be used; thirdly, a general description of what types of *turn-taking* procedures and distinctions are involved in the concept; and lastly, I theorize about what types of *activities* the design might be best suited for.
Pass the Token

**General:** This concept uses a tangible token to indicate and control whose turn it is to perform some interaction. This is regulated by the spatial position of the token, i.e., in whose personal space on the tabletop the tangible token is currently located. This requires the system to map interaction on the table to a specific user, and further to recognize which user has the token. This is an explicit form of turn-taking protocol where having the turn and passing the turn takes the form of passing a physical object between the participants; having the token may grant global access, making the token (or baton) a “privileged object coordination policy”. The gesture of passing the token can be viewed as a pointing or kinetic gesture. The shape of the tangible could for instance be a ball - not too big since it could hinder interaction or visibility. How much of the tabletop interaction the token regulates may vary from task types and specific applications. The token could regulate the activity-type interaction and thus regulating turn-taking for the entire tabletop, or it could act as a turn-taking protocol on the object level. Overall, the token produces a high level of awareness of whose turn it is to interact and visibility of what that person does, since parallel interaction will be limited. Probably most usable when to control interaction with complex shared content (such as where one application interface covers the entire surface) or when using linked distributed copies of the content.

**Examples:** A group is performing a joint decision task where the group is supposed to design a collage for a friend. One participant has placed the token in his personal space on the table to indicate and enable that it’s his turn to place a picture. While he interacts he describes and narrates what he’s doing - “If we put this image here and move this picture to here...” [moving images around]. When he can’t find a good spot for the image he gives up the turn by saying - “Hmm, you have a go”. He then passes over the token to another participant by rolling it across the table. This participant places the token in his space, and then starts to move images around.

Another example inspired by observation at the ICU takes place during a morning patient round. Using linked copies of patient information, one user is presenting a patient by describing current patient status and treatment plan. Another user wants to illustrate a point and therefore says: “Pass the token over, I want to show you something”. When this user gets the token he places it in his personal space, giving him control over the linked copies. He demonstrates to the others his points by browsing through the patient data and highlighting features in the patient’s data not observed by the others.

**Turn-taking:** This design concept is best described as a collaborative turn-taking mechanism, regulating and coordinating collaboration and interaction. This is a type of explicit enforced turn-taking protocol for interaction that overrides ordinary social protocol. An alternative use of the token concept could be using the token without it regulating any interaction, only indicating who has the turn on a social level. Allowed interactions of participants not possessing the token may also depend on the activity, for instance during presentation activities it might be important for the listeners to continually make annotations.

**Activities:** This design concept is suitable for group-sequential activities where there is need of a high level of awareness and attention directed at the turn-taker. Suitable activities could therefore be meetings, presentations with different presenters, or joint decision tasks where the group has to take one unified decision.
Shared Representation List

**General:** This design concept is inspired by the patient rounds and the patient lists used during the rounds at the ICU. The idea behind this concept is to inform turn-taking through the content which is displayed on the table, a form of a priori planning. This is a general concept for entire applications, not just a supplementary turn-taking mechanism, which is meant to function as a patient list. During the rounds all patients are presented and discussed one by one, in the order they are listed on the patient list, by the group. Using this concept each patient would be listed on the digital patient list, and the “main active content” on the table would be the representation of the current patient. This list based concept can of course be used and is used during other activities besides patient round, consider meeting protocols or “Power Point-like” slides for instance. Every participant seated around the tabletop could have a copy of the list, or they could all share one common list. The main content on the tabletop which the list manages can also either be a shared representation in the center of the table, or copies distributed to each participant. This coordinates collaboration as a sort of coordinating schedule - a coordinating checklist. In this design concept the turn-taking structure is a natural built in part of the fabric and layout that makes up a tabletop application.

**Example:** As previously discussed; this design could be used during the doctor round for instance. In this example a clinician starts by presenting a patient to the others using the current patient representation on the tabletop. When the clinician is done introducing the patient, open discussion regarding patient health status and treatment plan take place. When the clinicians gathered around the table have reached satisfying plans and conclusions, it’s time to move on to the next patient. The clinician who presented this patient then presses the “Done”-button, alternatively presses the name of the next patient in the list, and the current content on the table is replaced with new content representing the next patient whose name becomes highlighted in the list. If the next patient is under the care of the same clinician; the clinicians continues to “control the table” and introduces the next patient. Alternatively, if the next patient is under the care of another clinician, the changed table content together with the highlighted list and social cues, serves as cues for the clinician “responsible” for the patient to introduce the patient to the other - similar to how the list structures the rounds today in other words. The patient representations are in this example linked copies which provide a common view of the patient information.

**Turn-taking:** This coordinates collaborative turn-taking of the implicit activity-type, by structuring what is to be presented when, and by whom. Collaboration and to some extent communication is coordinated implicitly, since the participant “responsible” for the active item on the list is hinted to take on the role of presenter. The design coordinates the shared view on the table and the roles of the participants (presenter or listener). It does not however restrict interaction, so the interaction can be regarded as open. For interaction social protocol will suffice, due to the low level of possible parallelism and high awareness of each other - especially of the presenter.

**Activities:** This concept can be used in any activity when different (or the same) participants have items that can be visualized on the tabletop to present for the rest of the group, or it can be used to structure a longer presentation. Meetings, such as the customer meeting at AF, or presentations could all be suitable activities.
4.2. Design Concepts

Gesturing for the Turn

**General:** As observed by both Bekker and Mondada, gestures can be used as a way of passing and asking for the turn during an activity. Gestures on the tabletop surface which triggers visual cues can be used to indicate that a participant wants the turn to interact with an object, or with the table at large. The gesture itself can be interpreted as a turn asking gesture, and the visual cue enhances the gesture. Hence, this design concept is not a fixed component on the table as the previous designs are. This can be regarded as a table version of the “hand-raising”-protocol when a participant wants to comment or as a question. This gesture can be a usable supplement social protocol as it can be used discretely to inform the presenter that a participant wants to comment or interrupt the presenter, or current turn-owner. When a turn-owner is highly focused on interacting with the tabletop contents, signals through the tabletop interface can inform without interrupting the turn-owner. Gestures coupled with visual signals like these can be a non intrusive signal to the current turn-owner that others want to interact.

**Example:** This example again takes place during a patient round at the ICU, where a group of clinicians are immersed in a discussion about a patient’s treatment plan. They are positioned around a digital tabletop and they are using digital patient representations placed on the tabletop surface to aid their discussions. One of the clinicians is currently illustrating the patient’s clinical status by showing and narrating patient symptoms using the representations. Another clinician wants to cut in but doesn’t want to interrupt the current content sender, so she signals that she wants in by performing the “turn-asking” gesture on the tabletop. The current turn-owner notices the visual cue, and when he has gotten to his point he passes over the representation to the gesturer.

**Turn-taking:** This design has the potential to inform and aid collaborative turn-taking at object and activity level. However, it’s not an explicit or regulated turn-taking protocol, which means that the indication signal can be ignored. It can also be argued that it’s a form of turn-taking at the communicative level, as this communication may coordinate further other communication. Another, regulating, version of this gesture-mechanic could automatically give table control to the gesturer - although this probably has a potential abuse factor.

**Activities:** This is useful in activities that are somewhat sequential or autocratic activities; or activities that involve central shared objects when focus tends to be aimed at the table contents and less on the other collaborators. It does not replace other social protocol, and should rather be seen as a supplement of those social mechanics. It can also be useful to help pass control when working with linked copies of an object. Suitable activities include: presentations, meetings and to some extent joint decision activities.
Passing the Turn

**General:** This design is not a visual component on the table, but rather subtle coordination mechanisms that can be applied to sharable digital objects on the digital tabletop to indicate turn-ownership. By moving an object into another’s personal space the object will rotate to face that user - and thus inviting him or her to interact. When using one finger while moving an object it will rotate, but when moving the object with two fingers it will not auto-rotate. The design concept is thus a combination of personal spaces, location and movement of objects, and orientation of objects - to indicate object-level turn-taking protocol. This concept is inspired by Morris’ relocate- and reorient-actions, where a user moves or rotates a document to regulate the documents access control policy. The differences are that auto-rotate is based on user space and that no explicit access control policy is changed, the turn-mechanism only indicates that access and control passes over to the next user. The idea is to take advantage of the tendency that users often form their own personal space on the digital tabletop, where they store objects which they intend to use. In this concept, spaces should be visually indicated to help users determine where personal space ends and where group space begins. Some objects are regarded as group objects in many activities and they are thus often located in the center of the tabletop - for instance, if a group is building a puzzle, then the puzzle itself will be built in the midst of the table, but the puzzle pieces may be spread out to each user. By making it possible to notice that one user finds it hard to collaborate (for instance by noticing hesitation or that objects are not moving out from his space) this mechanism can be used to invite the user to interact with an object - allowing for a on the fly planning approach.

**Example:** One situation could be at AF where a job coach is currently informing a job seeker to fill in a certain form. The job coach uses the orientation and placement of the forms to indicate which form the job seeker should fill in. By moving a form into the job seeker’s space while saying “Have a look at this” his gesture and the form itself indicates that it’s the job seeker’s turn to interact with the forms. For instance, if there are forms the job seeker is supposed to fill in for a job application, the job coach could move and place them in the job seeker’s personal space.

**Turn-taking:** This concept serves as a communication indicator and as collaboration coordinator. It’s an implied turn-taking concept that leaves the interaction possibilities open. It can be used as both activity-type turn-taking indicator, passing the turn to interact with the table at large to another participant, as well as object-type indicator passing the turn to interact with a certain object to another participant.

**Activities:** Seeing as this lightweight mechanism can be used as a turn-taking mechanism in several ways it could be applied in many different tasks and activities. Probably most suited for egalitarian activities with high levels of collaboration and possibilities for parallel interaction, such as the joint choice/decision activities where object control is likely to be shared. It can also be used during presentations or meetings or even during parallel activities, passing objects between tasks or subtasks.
4.2. Design Concepts

Copying Mechanics

**Description:** This concept is made up of two different copy mechanisms for digital objects on the tabletop. The first mechanism (1) is very similar to Morris’ duplicate policy used to resolve document-level conflicts. This concept is an interaction that creates a copy of an object on the tabletop, whether in use or not. By grabbing an object on the tabletop with three fingers a copy of the object can be created. This allows the owner of the original object to carry on his or her work without being interrupted. This is useful in certain tasks where many users at the same time want to study an object closely - for example a photo or a text document. Instead of passing around the object or placing it at the center of the table, where the distance and orientation will cause visibility problems for some users, all users can have a (perhaps temporary) working copy of the object.

The second mechanism (2), called “The Poker Dealer”, allows a user to fluently create multiple copies of an object by grabbing it with the palm of the hand, and performing a kinetic dragging gesture to the left or the right. During the dragging motion a “fan” of copies will be created, just as the poker dealer creates a fan of cards when shuffling a deck of cards. This is useful in activities where it’s obvious that there will be high focus on a certain object. If one user realizes that all users will want a copy of an object, he can swiftly create copies and pass them around. All copies should be marked as copies and should be easy to get rid of when they are no longer necessary according to the guidelines, or else they will produce unnecessary visual clutter on the tabletop. Creating copies can help with asymmetry of work by giving shy users a chance to interact with objects central to the activity. This can also be interpreted as a blocking behavior - preventing others from stealing the original object from the holder. Each copy can be restricted to allow access only to the user who created the copy (if it’s suitable for the activity) - thus becoming a policy for shared access.

**Example:** Let’s consider the customer meeting at AF where a group of job seekers are to be introduced to AF and the job-finding process by a job coach. As everyone gathers around the tabletop the job coach realizes that all participants might want a copy of the information - therefore he creates a set of copies using the Poker Dealer. He then hands them out to each participant and can start his presentation. This mechanism can also be used at the ICU during the rounds, where clinicians can create copies of patient or other data - during presentations of patients or open meetings. For instance, if one clinician drops in late for rounds, the clinician could easily grab a copy of the patient list - or have one “dealt”.

**Turn-taking:** The concept of copy-mechanisms can coordinate collaborative turn-taking first and foremost on the object-level. However, by creating copies the nature of the activity can shift from a sequential to a parallel activity - hence it coordinates and re-coordinates turn-taking at the activity level. Creating a copy can be interpreted as both a “take the turn”-action (grab a copy) and a “give away the turn”-action (poker dealer). Further, it’s mostly implicit and the copies can either utilize regulated or open interaction.

**Activities:** Most suitable for a joint work activity such as a collaborative decision task or creative design tasks where there are plenty of shared objects that otherwise would cause a lot of explicit turn-taking. Creating copies causes the users to collaborate in a more parallel fashion, thus makes parts of the activity more like a parallel work activity - as seen in Morris’ group strategies for collaboration (see section 2.2.2). It can transform a joint sequential activity into a more parallel activity and allows for users to structure the activity on the fly as they can divide and re-divide the activity as they go along.
Storyboards  The storyboards presented in this section were developed in order to more clearly explain and illustrate the design concepts. They serve as graphical supplements to the mostly textual based design sheets and the interaction examples described there. The storyboards are based on scenarios and made up of sketches of the ongoing interaction together with descriptions of what’s happening in each sketch. Metadata used in the sketches is sketched using a red color.
4.2. Design Concepts

Description:
Four users are gathered around a digital tabletop during the doctor round. User four has the token. They are using linked copies of the patient information and personal space on the table is annotated in blue.

Dialogue:
User four interacts with his copy, synchronizing the other copies to match his view. User one asks a question.

User four: "As we can see, temperature is stable, and..."
User one: "Is the patient showing signs of XXX?"

User one's question starts an open discussion which lasts for a while.

User one asks for the token in order to illustrate a point in her line of argument. User four gives the token to user one.

User one: "Can you pass the token?"

User four places the token inside her personal space, indicating that it’s her turn to interact with the tabletop. The tabletop recognizes the token and responds by indicating that her copy is now the master copy.

User one can now illustrate her reasoning by browsing the patient’s data, which synchronizes the others’ views.

Figure 4.6: Pass the Token storyboard
Results

Four participants are gathered at the doctor round. They are rounding one patient at a time using linked copies of patient information and digital patient lists.

User one presents the first patient. As she interacts with her copy, she synchronizes the others' views.

Dialogue:
User one: "The patient is showing high values of..."

When user one is finished with her patient, she moves on to the next patient in the list by touching the next patient's name.

Dialogue:
User one: "Next up is Erik"

User three recognizes the patient as one under her care. Thus, it's her turn to present the patient.

Dialogue:
User three: "Erik came to us early last week and received treatment for..."

User three continues to present the patient and her interaction with her copy synchronizes with the others.

Dialogue:
User three: "Recently his temperature has started to drop as we can see here..."

Figure 4.7: Shared representation list storyboard
4.2. Design Concepts

(a) Using gestures

Description:
During a customer meeting at AF the employment coach displays a bunch of documents and introduces the job seeker to AF's services.

Dialogue:
User two: “We will start by registering you in the system...
User two invites user one to interact with a document; when moved into user one’s space it orients itself to user one.

User two: “Have you filled in this application?”
User one: “No.”

User two continues his presentation. The orientation and position of the yellow document now signals that it’s user one’s turn to interact with the document in his space.

(b) Passing the turn

Description:
Three users are engaged in a task involving shared objects on the digital tabletop, for instance shared patient information. User two is currently engaged in interaction with one of the objects.

Dialogue:
User two: “If I move this around here...”
User one now wants to interact with the object. By rubbing the table surface with the palm of her hand she signals for attention. This causes the area of current interaction to light up—notify user two.

User one: “Can I try something?”
User two: “Okay”

Figure 4.8: Storyboards of using gestures and passing the turn
Three participants are engaged in an activity with a shared document; user two and three grab the shared documents.

User one also wants the yellow document that user two currently holds. User one thus uses the copy gesture by grabbing the document using three fingers.

When user one now retracts his hand, he creates a copy of the yellow document. User two holds the original.

During a customer group meeting at AF a employment coach (user one) wants to present a document for three job seekers.

Dialogue:
User one: “Can you all see the document?”
User three: “Not really”

User one decides to make copies of the document for each individual. By grabbing the document with the palm of her hand she can create copies of the document by performing a sweeping motion. When three copies are created she stops.

User one gives one copy to each participant and starts describing the document.

Dialogue:
User one: “Alright, what you want to do when filling in applications like this?”

During a customer group meeting at AF a employment coach (user one) wants to present a document for three job seekers.

Dialogue:
User one: “Can you all see the document?”
User three: “Not really”

User one decides to make copies of the document for each individual. By grabbing the document with the palm of her hand she can create copies of the document by performing a sweeping motion. When three copies are created she stops.

User one gives one copy to each participant and starts describing the document.

Dialogue:
User one: “Alright, what you want to do when filling in applications like this?”

Figure 4.9: Storyboards of copying mechanics

(a) Grab a copy

(b) Poker Hand copies
Mockup  In order to further illustrate how one of the concepts could look and feel a mockup was made. The mockup made of one of the designs is made with a higher fidelity, giving it a more realistic look. The concept “The Poker Hand” was further mocked up, mainly as it is not dependent on the application environment in the same way as many of the other concepts, where the design of the main representations are an equally big part. In the mockup in figure 4.10 three copies are made of a digital x-ray image placed on the digital tabletop, inspired by the ICU setting. As the image is dragged (as the arrow shows) copies are created with a regular interval. In this scenario the user is satisfied after three copies and thus stops the motion, creating the three copies.
Figure 4.10 Mockup of the “Poker Hand” concept.

(a) The above picture illustrates the palm grab gesture on a digital object on the tabletop surface. This gesture initiates the Poker Hand.

(b) In the picture bellow the hand has been moved to the right in a sweeping motion, as indicated by the arrow. The copies are placed in a sequence bellow the original object throughout the motion.
Chapter 5

Discussion

The discussion chapter can roughly be divided into three main topics; the model on collaborative turn-taking, the conceptual designs, and finally further research and conclusions. However, first I will quickly mention some aspects of working with digital tabletops and how these may relate to collaborative turn-taking, based on my own experiences.

Digital tabletops (i.e. horizontal displays) are excellent information output devices, and so are vertical displays. However, they are not as good when used as input devices depending on the activity. They can handle multiple user input at the same time and are good at semi-physical manipulation of objects, but they are not as good input devices in comparison to a physical keyboard. Digital tabletops can of course be equipped with e.g. external keyboards or pens in order to improve its input capabilities. Activities performed with a digital tabletop can be affected by which input devices are used. For instance, activities involving a lot of collaborative writing may not be suitable without some extended form of input. If one provides pens to be used as input devices one may ponder on how they can change the task itself, and affect the turn-taking protocols. If one for instance provides only one pen for input, the pen itself would become a passable token evoking a “pass the baton” protocol, similar to the passable token design concept.

Another experience from designing turn-taking mechanics for a digital
Discussion

Tabletop is that this effort circles an uncomfortable paradox; turn-taking among participants is required for sequential tightly coupled activities or when the technology does not provide possibilities for multiple-user-input (i.e. standard PC); and the digital tabletop has multiple-user-input and parallel work capabilities. Thus designing for turn-taking at the digital tabletop may seem like a paradox. However, the purpose for turn-taking at the digital tabletop is twofold: 1) there are certain activities suitable for a digital tabletop which are performed in sequential manner that does not encourage multiple users interacting at the same time, and 2) parallel interaction brings turn-taking related issues, such as object level accessibility (private and public objects) and coordination taking place in-between parallel interaction. For these reasons, collaborative turn-taking should be considered for digital tabletops.

5.1 The Collaborative Turn-Taking Model

As we’ve seen throughout this thesis, turn-taking at the digital tabletop does not fit in neatly at any one place. Rather, collaborative turn-taking considers a myriad of aspects and takes on several forms, everything from sequential one user interaction to parallel multi-user interactions. Before discussing the model itself, I’ll discuss some aspects of CSCW research in general.

There seems to be some ambiguity in the usages of the terms cooperation and collaboration in the literature. In my experience these terms are rarely defined and often used interchangeably, especially in the CSCW literature. The relation between collaboration and communication and how technology fits within this picture is often far from obvious; for instance, arguments in line with: “we have amplified communication and therefore we have amplified collaboration” are often put forth without any definition of either terms. I’m not saying that this line of reasoning is wrong; it’s just not clear why it’s correct. For instance, is there a situation in which supported communication does not lead to enhanced collaboration?

Many of the guidelines used when designing are on a general level and largely encompassed by the collaborative turn-taking perspective. For instance, one of the guidelines is to provide shared access to digital content.
which is something discussed throughout this thesis and the perspective; whereas other guidelines are too general and only usable when designing an entire digital tabletop systems. Many designs use personal spaces which should be visually indicated according to the guidelines. The guidelines generally provide notions on how the digital tabletop should be designed, and rightfully so. However, what these guidelines lack is a framework of evaluation - a way of using these guidelines as a heuristic measure of a design’s value.

It’s my opinion that CSCW lacks a common library of terms, guidelines and methods for research and design; a common framework. Lately, much work has been done under the parole of distributed cognition, but as Halverson (2002) points out: distributed cognition has no strong tradition of established terms. CSCW and research on digital tabletops would benefit from a establishing a framework, with roots in related fields of research, such as: cognitive science, linguistics, social sciences and group psychology. Distributed cognition is a good place to start, but more work needs to be done in theory and in practice. A framework for digital tabletops should place more emphasis on the nature of collaboration, collaborative activities, collaborative strategies, mechanics of coordination, communication, and how this type of knowledge can guide and shape design. The importance of establishing a classification of collaborative activities was previously stressed by Gutwin and Greenberg (2000). The lack of an established framework is what initially led me to view collaboration and collaborative turn-taking from my own perspective, influenced by a potpourri of theories and previous work.

From the collaborative turn-taking perspective, certain elements and aspects of collaboration, interaction and turn-taking are stressed; such as object-level interaction and activity-level interaction. These highlighted aspects followed from the activity classification and the aspects which are important to consider during the different collaborative tasks outlined in section 4.1.2. Considering these aspects will aid and inform the design process when designing collaborative technology. These aspects are further often indirectly mentioned and discussed in the vast literature on digital tabletops. The collaborative activities themselves followed from observations fueled by literature.

The perspective on collaborative turn-taking undertaken does make
some assumptions of turn-taking and collaboration. One of the underlying assumptions is that supporting turn-taking mechanisms can supplement social protocol to prevent and resolve conflicts. This assumption rests on Hutchins’, among others, view on coordination as something constraining one’s behavior through the surrounding system. By constraining socially unaccepted behavior and interactions, some of the conflicts and misunderstandings may be resolved as the user’s actions are constrained by the environment. Such constraints must be carefully considered not to interfere with the overall collaboration and workflow. One further assumption made is that activities affect how turn-taking protocols should be utilized, and activities are often seen as important factors in overall group performance, emphasized by McGrath and Hollingshead (1993).

The division of the model into two levels is not a given truth, as conversation itself can be views as a form of joint action, in accordance with Clark and Linell. Thus one level of action oriented collaborative turn-taking would suffice; one level describing the turns taken during a collaborative activity as actions. The division of the perspective into two levels is a way of grouping communicative aspects separately, as they are the biggest category of collaborative coordination mechanisms taking place above the table. Another reason was also to place more emphasis on other “non-communicative” (non-intentional communication at least) aspects of collaborative coordination. This level can surely be divided further, as it is a bag of different coordination mechanics that are very diverse - ranging from external representation to situational and social protocol. Important to note is that both levels of turn-taking are present and simultaneously active during a collaborative activity, and often they are hard to separate.

The activity classifications is not an ideal classification schema. In an ideal schema the categories should be mutually exclusive (one activity fits one and only one category), collectively exclusive (all possible tasks should fit some category), logically related to each other and the classification should of course should be useful (McGrath, 1984). I do not claim that this classification of activities is ideally exclusive, but it is very useful for designing and relating turn-taking to design at the digital tabletop. Neither is the number of used dimensions or the specific dimension chosen exhaustive in describing an activity, but they are effective at illustrating differences between the suggested activity types. When the dimensions are
5.2 Design Concepts

I’ve chosen a theory heavy path towards designing collaborative support. Such a path has its strengths and weaknesses. Theories shape how we think about and relate to what we are designing (and the world in general); with a theory we can describe and understand the world in order to bridge the gap between description and design. A good theory can thus shape our work towards something useful. Ideas and thoughts from both distributed cognition and the related dialogical theory are relatively new to western culture in general. Therefore, it can be somewhat of a tricky path to bring these ideas into the concepts of design. How does one take these aspects into consideration when designing interaction?

I believe that there exists some non-trivial problems to really understand how these theories of cognition and communication should be accounted for when designing. It’s a case of breaking outside the standard “box” of thought. It’s often easy to end up in the old loop of single user interfaces and standard design metaphors, just on new platforms with enhanced new visuals. Here a common extended framework for CSCW and co-located groupware could help. However, Halverson holds an optimistic view of the role and importance of theory and distributed cognition for design - which I agree with. The issues and difficulties I’ve experienced in “translating” lessons and ideas from distributed cognition into design practice may have been caused by my choice of method. If I would instead have performed longer ethnographical field studies follow by distributed cognition data analysis, the theory-design gap may have been bridged. Other frameworks for system analysis could also have been chosen, for instance a joint cognitive system view, in line with the work of Woods and Hollnagel (2006), would have emphasized other aspects of the human-tabletop system, possibly resulting in other design solutions. Another methodological approach towards designing these concepts may have been to adapt a stricter design practice and method, for instance by using the Cooper goal-oriented design method. This relies more on user interviews and the
creation of personas in order to clearly establish the user’s goals and how they could be supported through the designed product. This would have been a good approach if the tables were already in place and in use at the ICU or at AF, as the users would be more able to relate to the digital tabletop and their goals when using it. This method thus places a lot of focus on designing for a target group, whereas I’ve tried to keep the designs as general as possible and only using inspiration from observed group collaboration settings and activities.

The concept designs are examples of general tangible and digital mechanisms and components that can be used as a way of supporting turn-taking and coordinating the joint group activity at a digital tabletop. What’s important to emphasize is the notion that most of the activity coordination and communication still takes place above the table surface, i.e. between participants through social protocol, conversation and other means of communication. The design concepts are attempts at distributing some of the coordination mechanisms onto the digital surface, which enables the use of additional communicative and coordinative channels. Hopefully, these designs can support collaboration and help prevent conflicts caused by mismatched coordination.

These designs are all embedded with the surface interaction of the tabletop, thereby providing collaborative coordination through the tabletop interface. Something that must not be forgotten or ignored are the issues revealed by Morris and further illustrated in the studies by Olson et al. - the difference in attitude towards digital versus psychical objects. Why is it that using physical papers raise more social responses than digital; and why does sharing a digital toolbar cause collaboration to break down, while a tangible control works? How do we get people to engage a digital task with a more social approach? Both Morris and Olson et al. argues that the newness of the technology are the cause, but is that all that there is to it? One cannot ignore the fact that digital and physical objects provide different affordances and possibilities. For instance, what role does the fact that one can pick up a paper and hold it close play? If in fact these issues are caused by the newness of the technology, time will eventually mend these issues; if so, what can we do in the meantime to ease the transition and when would it be appropriate to use digital or physical objects? All of them, I think, are interesting questions and issues, deserving further
research.

5.3 Future Research

I would say that possible future work divides into three main branches of research. The first branch would be to prototype, evaluate, iterate and refine the suggested design concepts; and finally implement them for tabletop usage. Prototyping should be done with both lo-fi and hi-fi methods, ranging from simple paper prototypes to interactive versions on the tabletop. This branch also includes attempts to further generalize the suggestions in order to make them general and usable tools for digital tabletop collaboration and interaction. Especially interesting is to investigate how these gestures and mechanics for turn-taking will blend in to any given digital tabletop environment, and how to make them blend in efficiently and effortlessly. This entire branch of research is probably the most natural next step to take, as it deals with the direct results from this thesis.

The second branch of research on the other hand does not focus as much on the suggested design concepts. Rather, this branch is about following up on the ethnographical studies and the found interest for using digital tabletops in the observed domains. This branch would involve further ethnographical work and experimental studies to evaluate the usability of digital tabletops in these domains. The studies should pursue answers for what digital tabletops could be used in these domains - further designing how the tabletops should be used.

The third branch is more theory bound in its’ nature. This branch would be concerned with exploring, evaluating, extending and generally improving the suggested model on collaborative turn-taking and collaborative activities. If this is a fruitful venue for further research is hard to say, but the perspective does bring up interesting and important topics to the surface, such as how activities can shape digital tabletop applications. One of the goals in such an undertaking would be to corroborate the underlying assumptions made by the model. To do this I would suggest that an approach similar to the research activity suggested by (Hollan et al., 2000, p.181) be used, in which both ethnography and experiments play important parts.
These three branches are all topics that have more or less been touched upon during this thesis, and are all therefore more or less given topics for further research. Beyond these, there are topics for further research not explored in this thesis, for instance further studying the different aspects of using digital versus tangible objects for turn-taking control. Related to this is how ordinary physical objects can be enhanced with digital properties (for instance by adding labels or controls to physical documents placed on the tabletop) and how these digital properties may relate to turn-taking.

One further idea would be to investigate social relationships and social factors and how collaborative turn-taking, with or without technological protocols, can relate to it - this can also be viewed as an extension of the turn-taking model. An often cited aspect of “groupware acceptance” is that it’s affected by social factors, such as organizational culture, differences in personalities, and group dynamics.

From the view of distributed cognition, task performance is embedded in real human relationships and every action is not only part of completing the current task, but also a form of communication. Building and maintaining good social relationships becomes an important motive for competent performance (Hutchins, 1995). This is an important aspect of team performance to explore around the digital tabletop, especially since we’ve seen that people sometimes have problems when it comes to following social protocol in the tabletop settings. Can activity-appropriate turn-taking procedures support building and maintaining relationships by avoiding conflict and structuring interaction?

One can further speculate in what the cognitive system that is the digital tabletop and the users can do with the knowledge of who can and who currently is interacting with what. This information brings possibilities of providing enhanced dynamic user interfaces which adapt to the current situation - showing information to one user not relevant for the others. This can also be used in relation to the visual clutter problem often occurring with single display groupware. However, systems like these are hard to design and would require a lot of iteration and alternation between field studies and design evaluation - since information and controls not actively used may still serve other purposes.
5.4 Conclusions

Collaboration among group members with or without technology is a complex matter, and collaboration around digital tabletops and other groupware applications can cause unwanted breakdowns or conflicts when relying solely on social protocol. This thesis addressed this issue by exploring collaborative turn-taking protocols as coordinating mechanisms and the role of collaborative activities. A perspective on collaborative turn-taking emerged, and from this undertaken perspective five design concepts for turn-taking were created in order to support collaborative turn-taking. The designs are rooted in theory, previous work and guidelines from the literature together with inspirational ethnographical studies. An important next step to take involves prototyping and evaluating the suggested concepts.
Bibliography


of the SIGCHI conference on Human factors in computing systems: the CHI is the limit, (pp. 378–385) (ACM).


Williams, K. D. (2010). Dyads can be groups (and often are). Small Group Research, vol. 41:268.

Collaboration technologies are difficult to design due to the complex myriad of social, cognitive, and communicative aspects of group interactions. New interaction technologies like multitouch sharable interfaces, such as digital tabletops, have lead to a renewed interest in designing collaborative technologies. This thesis focuses on turn-taking protocols as a coordinating mechanism during collaborative work with digital tabletops. The goal was to develop new conceptual designs and interactive mechanisms to support face-to-face collaborations of small groups. Inspired by ethnographical studies of collaborative work and theories in distributed cognition and related theories of language and action a model of collaborative turn-taking was developed. Moreover, the thesis presents five design concepts and interaction components for the digital tabletop that exemplifies the different properties of the model.
Copyright

Svenska
Detta dokument hålls tillgängligt på Internet - eller dess framtida ersättare - under 25 år från publiceringsdatum under förutsättning att inga extraordinära omständigheter uppstår.

Tillgång till dokumentet innebär tillstånd för var och en att läsa, ladda ner, skriva ut enstaka kopior för enskilt bruk och att använda det oförändrat för ickekomersiell forskning och för undervisning. Överföring av upphovsrätten vid en senare tidpunkt kan inte upphäva detta tillstånd. All annan användning av dokumentet kräver upphovsmannens medgivande. För att garantera äktheten, säkerheten och tillgängligheten finns det lösningar av teknisk och administrativ art.

Upphovsmannens ideella rätt innefattar rätt att bli nämnt som upphovsman i den omfattning som god sed kräver vid användning av dokumentet på ovan beskrivna sätt samt skydd mot att dokumentet ändras eller presenteras i sådan form eller i sådant sammanhang som är kränkande för upphovsmannens litterära eller konstnärliga anseende eller egenart.

För ytterligare information om Linköping University Electronic Press se förlagets hemsida http://www.ep.liu.se/

English
The publishers will keep this document online on the Internet - or its possible replacement - for a period of 25 years from the date of publication barring exceptional circumstances.

The online availability of the document implies a permanent permission for anyone to read, to download, to print out single copies for your own use and to use it unchanged for any non-commercial research and educational purpose. Subsequent transfers of copyright cannot revoke this permission. All other uses of the document are conditional on the consent of the copyright owner. The publisher has taken technical and administrative measures to assure authenticity, security and accessibility.

According to intellectual property law the author has the right to be mentioned when his/her work is accessed as described above and to be protected against infringement.

For additional information about the Linköping University Electronic Press and its procedures for publication and for assurance of document integrity, please refer to its WWW home page: http://www.ep.liu.se/

© Jonas Rybing
Linköping, October 25, 2011