Misalignments of Social and Numerical Identity — an Ontological Analysis

Birger Andersson, Maria Bergholtz, and Paul Johannesson

Department of Computer and Systems Sciences, Stockholm University
ba.birger.andersson@gmail.com,{maria,pajo}@dsv.su.se

Abstract. Numerical identity is the relation that an individual thing bears to itself and only itself, which is not dependent on any other object. However, in the discourse of identity management, identity is often something else: an identifier that an organization has assigned to some entity, here called social identity. These two notions of identity are closely related, as social identities are designed to mirror numerical identities from an organisational point of view. But this mirroring can easily break down or be misaligned. This paper offers an ontological analysis of the relationships between numerical and social identity, with a focus on identifying different forms of their misalignments and potential causes for these. For the analysis we rely on the Unified Foundational Ontology (UFO), and for the conceptual modelling we use OntoUML. The result of the ontological analysis takes the form of a conceptual model. We envisage that this model can not only clarify theoretical concepts related to identity, but also have practical applications in addressing issues of rights and agency in digital identity management.

Keywords: Ontology · Conceptual Model · Social identity · Numerical identity

1 Introduction

As digital infrastructures and ecosystems continue to expand, maintaining security, trust, and personalization has become increasingly crucial. A key instrument for this purpose is the digital identity. It refers to the online representation of an individual, organization, or entity in the digital world. However, this idea of identity differs from both our everyday and philosophical understandings of identity. In this paper, we will analyze and compare these differing notions of identity, explore the challenges that arise when they overlap, and propose an ontological analysis as a way of disentangling them.

Two things are identical if and only if they are the same; the notions of identity and sameness are the same. However, there are different kinds of identity, and a common distinction is the one between qualitative and numerical identity. Qualitative identity refers to the similarity in properties between two entities, which can result in varying degrees of identity. For instance, a Volvo and a Mercedes share the property of being a car, which makes them, to some extent,

qualitatively identical. In contrast, "numerical identity requires absolute, or total, qualitative identity, and can only hold between a thing and itself" [12]. A related notion is that of a principle of identity, which has been defined by [8] as "A principle of identity supports the judgment whether two particulars are the same, i.e., in which circumstances the identity relation holds".

Another way of conceptualizing identity, here called "social identity", and common in the field of identity management [2], involves organizations providing identities to individuals or physical entities. Those identities are typically based on some identifier, which is a code or another specific piece of information assigned to uniquely identify an entity within a particular organization or context, e.g. a social security number, an employee number or a vehicle identification number [4]. As expressed by [13], "Most of what we call "identity" isn't. It's identifiers. It's how some organization identifies you: as a citizen, a driver, a member, a student. Those organizations may issue you an "ID" in the form of a passport, license, or membership card, but that isn't your identity. It's their identifier". Provisioning social identities within organisations serves administrative, security as well as compliance purposes, thereby facilitating efficient communication and accountability within an organisational structure.

These two notions of identity, numerical identity and social identity, are closely related, as the purpose of social identities is to mirror numerical identities from an organisational point of view. Nevertheless, the notions are distinct, which in practice becomes apparent when the mirroring breaks down. A number of such breakdowns, or misalignments, are discussed below in the form of vignettes.

The goal of the paper is to offer an ontological analysis of the relationships between numerical and social identity, with a focus on identifying different forms of their misalignments and potential causes for these. The paper is structured as follows. First, Section 2 provides some motivating examples through a number of vignettes. Then, Section 3 presents the ontological analysis of numerical and social identity, in the form of a UFO-based conceptual model. Finally, Section 4 concludes with some open questions and directions for future research.

2 Vignettes of Identity Misalignment

The Dishonest Applicant A person presents themselves at the citizen registration office with documents from another country, claiming they are sufficient to obtain citizenship. The responsible officer thoroughly examines the documents and deems them satisfactory, ultimately granting citizenship. However, it was later discovered that the documents were fraudulent. In light of this, is the citizenship valid? Can the person be considered a citizen?

The Deceased Citizen A person who is a citizen passes away, but the citizen registry is not updated to record this fact. Although the person does not exist anymore, various rights related to the citizenship still exist, e.g. allowances and the right to vote. And other persons can actually exercise these rights, even if unlawfully. Does a citizenship and a citizen still exist?

The Corrupt Officer A record of a new citizenship is fraudulently entered into the citizen registry by an officer at the registration office. This record is then sold to a known criminal who allows multiple individuals to use the citizenship. No one ever discovers these machinations, and several people actually use the citizenship. But is there a new citizenship? Is there a new citizen?

The Involuntary Voter During a meeting where people are voting, a newcomer enters the room. The chairman requests everyone to vote by raising their hands. However, the newcomer is unaware of the voting process and simply raises his hand to shoo away a fly. The chairman counts the newcomer's raised hand as a vote. Did the newcomer actually cast a vote? (This vignette is in fact not about identity misalignment but serves to generalize the phenomenon.)

3 A Conceptual Model of Numerical and Social Identity

3.1 The Unified Foundational Ontology (UFO)

In this paper we rely on UFO [8] as a top-level ontology modelled by means of OntoUML [9]. OntoUML is an extension of UML that incorporates the basic ontological distinctions made in UFO in the form of UML stereotypes. Stereotypes (enclosed between « » symbols) indicate the meta-category (kind of universal) to which a certain UML class belongs, constraining in this way its semantics according to the UFO ontology.

The basic kinds of universals we use in our model are *kind*, *role* (both subclasses of sortal universal), *category*, and *rolemixin* (both subclasses of mixin universal). In addition, we shall use the stereotypes *relator*, *event*, *mode*, *quality*, and a few others.

Intuitively, we can see qualities as specific aspects of things we can use to compare them. Qualities inhere in things, where inherence is a special kind of existential dependence relation. Qualities are considered as endurants in UFO. In UFO, qualities belong to the more general class of intrinsic moments which also includes modes such as a thought, a belief, or an intention. Among modes, there are externally dependent modes, such as a particular mental attitude towards another person or object, which are existentially dependent on something else besides their bearer (i.e., the entity they inhere in).

Concerning relations and relationships, while a relation is usually intended as a set of tuples, a relationship should not be considered a tuple (i.e., an ordered set of objects) but rather an object in itself that needs to exist in the world in order for a relation to hold. Relations hold (i.e., relational propositions are true) in virtue of the existence of a relationship; relationships are therefore truthmakers of relations.

A relevant class of relations are extrinsic relations which can not be derived from the intrinsic properties of their relata — 'married with' is the prototypical example. Extrinsic relations have corresponding relationships that can be understood as mereological sums of relational qualities: for instance, a marriage can

be understood as a sum of mutual commitments and obligations. An advantage of such a position is that, since qualities are assumed to be endurants (i.e., entities that may change in time while maintaining their identity), relationships are endurants as well, whose behavior in time accounts for the way a relation holds in time. In UFO, such relationships are called *relators*.

3.2 The Conceptual Model

The conceptual model of figure 1 contrasts the notions of a social identity with the numerical identity of an agent. It intends to capture the mirroring between entities that need to be managed by a social system and those created by that system for identifying them. In particular, it intends to make explicit how rights are related to and exercised by these different kinds of entities. For the sake of brevity, we will focus on entities that are agents.

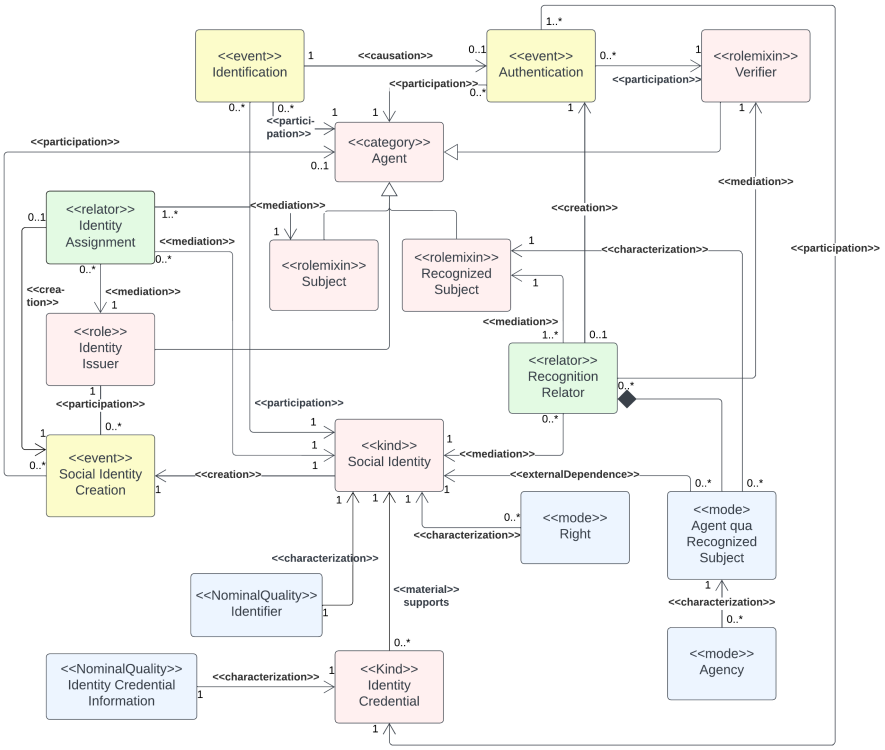

Fig. 1. OntoUML Model of Social Identity and Related Concepts

Agents and Subjects — the Participants of Social Systems For the notion of a social system, we take a broad view, including any structured society or community, such as nation-states, professional groups, corporations, and communities based on established standards.

An *agent* is an entity that can perceive its environment, process information, make decisions and take actions to achieve specific goals. An agent can be a human or a machine; thus, the class has no uniform principle of identity and is stereotyped as a category.

A *subject* is an agent that is part of a social system and is influenced, affected, or governed by it, e.g. a citizen, an employee or a physician. A subject is dependent on a social system, and the class is therefore stereotyped as a rolemixin.

Social Identities — Identifying Subjects In order to be recognized by other participants of a social system, a subject needs to be identifiable. It is the responsibility of *identity issuers* to create and maintain such identifiable entities. Identity issuers thereby play a crucial role in establishing trust and enabling secure interactions within the social system. An identity issuer is a role, which is typically played by a trusted organization such as a government agency, a financial institution, or a major tech company, e.g. the Swedish government, Crédit Suisse, or Google.

A *social identity* is a semiabstract entity, meaning that it has no extension in space but a beginning and end in time. It is created by an identity issuer, and its purpose is to serve as a unique, stable and authorized reference for a subject. Thus, a social identity is related to an identifier which is stereotyped as a NominalQuality [14], e.g. a social security number or an employee number.

To represent that a social identity is assigned to a subject by an identity issuer, a relator *identity assignment* is used. An identity assignment, as well as a social identifier, are created by means of an event called *social identity creation*, in which an identity issuer declares that a social identity with a specific identifier has been created. The identity issuer also declares that the subject has been assigned that social identity, similar to a baptism event. Ideally, every social identity should be related to an identity assignment, but mistakes or fraud during the social identity creation event can result in the absence of such an assignment, as demonstrated in the Corrupt Officer vignette. Another reason is that a subject's social identity may become unlinked to its identity assignment if the subject ceases to exist, while their social identity remains, as shown in the Deceased Citizen vignette.

In order to verify a subject's social identity, i.e., the social identity assigned to the subject, there is a need for information that can be used to determine the subject's (numerical) identity. Such information can be a username, a password, a security question or biometric data, such as fingerprints and facial scans. We model this through the classes *Identity Credential* and *Identity Credential Information*.

Social Identities — Serving as a Nexus of Rights We here base the notion of a *right* on the work by Hohfeld [10], which was further developed by Alexy [1], and applied in UFO-L [7]. Hohfeld distinguishes between different kinds of rights, including privileges, claims, and powers. A person has a privilege to perform an action if they are free to do so in accordance with the rules of a social system. A person has a claim on another person if the latter is required to act in a certain way for the former's benefit. A power is the ability of a person to create or modify claims, privileges or powers. These notions can be extended also to organizations and, more controversially, to machines. However, we choose to model a right as being a mode inherent in a social identity, the reason being that the same rights can be claimed and exercised by different agents. In this way, a social identity can serve as a nexus of rights, where different rights come together in a focal point and can be recognized so that an agent can exercise those rights; more on this will be discussed below.

Authentication and Recognition — From Right to Agency An *identification* is an event in which an agent claims that it is assigned to a social identity. For example, when an agent provides their name, username, or email address, they are identifying themselves. An identification does not prove that the claimed social identity is assigned to the agent; it is only the agent's assertion that this is the case.

A *verifier* is an agent that verifies whether the claim in an identification event is valid or not. A verifier is stereotyped as a rolemixin.

An *authentication* is an event in which a claimed social identity is to be verified, i.e. when an agent claims that it is assigned to a social identity, a verifier confirms whether this claim is valid or not. An authentication involves presenting evidence that the agent is actually who they claim they are, e.g. a fingerprint, a password or an answer to a security question. The verifier then compares this evidence with the identity credential of the claimed social identity to assess the agent's claim.

If an authentication event verifies an agent's claimed social identity, a *recognition relator* is created. This relator expresses that the social system recognises the agent with their claimed social identity and is prepared to acknowledge their rights accordingly. This relator is between a *recognized subject*, which is an agent, and a social identifier. For example, a check-in clerk at an airport, recognising a person with their claimed social security number, will give access to flights booked by that social identity. Thus, the recognized subject is able to exercise (some of) the rights associated with the claimed social identity. This situation is modelled by means of a qua object, *agent qua recognized subject*, which is a mode that characterizes a recognized subject and is externally dependent on a social identity.

This qua object has a number of agencies that correspond to the rights of that social identity, i.e., it is able to act because the verifier has acknowledged those rights. In the aforementioned example, the person can board the airplane, choose a seat, request assistance from the flight crew, and so on. It is possible that a

recognized subject with a social identity is different from the subject assigned to that social identity; in this case, the recognized subject has fraudulently claimed a social identity it did not have. From the perspective of the social system, this is clearly an undesirable situation. However, even when this is the case, the recognized subject will still be able to use their agencies.

3.3 The Vignettes Revisited

In the following we illustrate how the conceptual model mentioned earlier can help elucidate the first three vignettes in Section 2. Figure 2 shows a small case study in the form of a domain model that has been tailored specifically to the topic of citizenships.

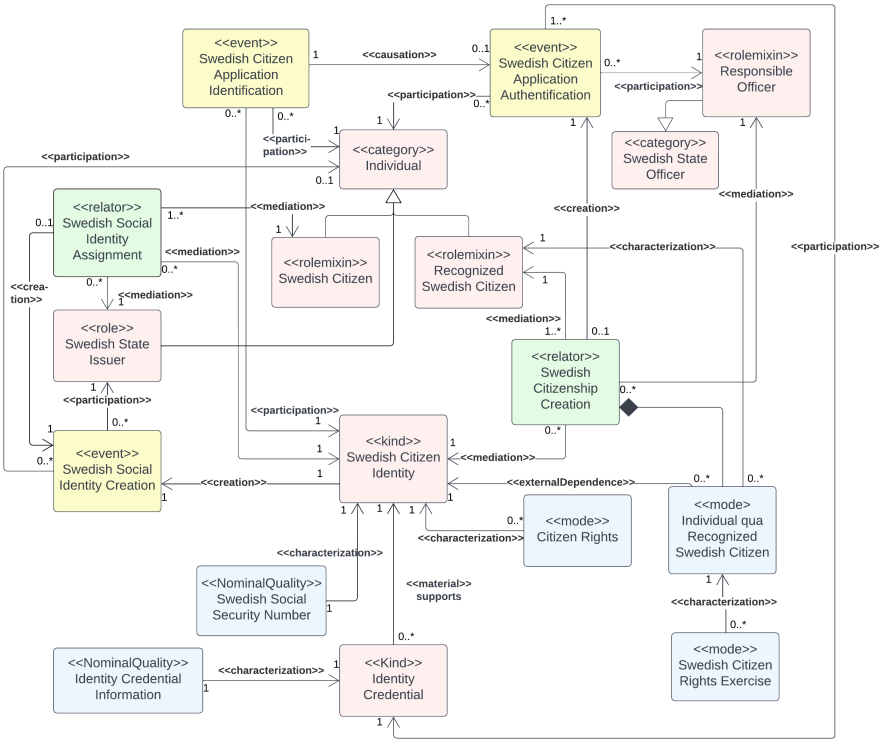

Fig. 2. The Conceptual Model Specialized to Citizenships

The Dishonest Applicant In the Swedish Social Identity Creation event, a social identity for a Swedish Citizen Identity is created, which will work for recording rights and recognizing a Recognized Swedish Citizen. For these

purposes, it doesn't matter that the event was based on faulty evidence. However, because of this faulty evidence, it can be questioned whether a Swedish Citizen was actually created — the answer to this question depends on the specific circumstance of the case (e.g. whether some error of judgment by the officer was made) and the applicable regulations, as discussed further in Section 4.

The Deceased Citizen In this case, a Swedish Citizen Identity still exists and is available for claiming and exercising the associated rights. But there is no person and, therefore, no Swedish Citizen nor a Swedish Identity Assignment.

The Corrupt Officer In contrast to the previous case, there has never been any Swedish Identity Assignment. But there is a Swedish Citizen Identity created by the corrupt officer, which is the basis for claiming and exercising the associated rights. And multiple agents can claim and exercise these rights.

4 Concluding Discussion, Open Issues and Future Work

We have proposed a conceptual model, based on UFO, for clarifying the relationships between numerical and social identity. In particular, we have investigated how social identities are assigned to subjects of social systems, how rights are related to social identities, and how authentication events provide subjects with agencies that let them exercise rights. Most of the model's concepts are based on work in the area of identity management, [2], and Identity Management and Access (IAM) systems [11]. There is also work in the information systems area that has acknowledged the need for social (also called institutional) identities that correspond to physical entities, for example [6] and [5]. One distinguishing feature of our work is that we, through the use of a conceptual model, have made explicit the assignment of rights and the verification of social identities, thereby providing a basis for characterizing misalignments between numerical and social identity.

We have only considered rights as being associated with a social identity. Thus, the exercise of these rights requires authentication of that social identity. However, rights can also be related to physical or informational entities, where a subject only must have custody of those entities in order to exercise those rights. For example, a banknote can grant someone the power to purchase items, but they can only make the purchase if they have the banknote in their custody. As pointed out already by [10], if an individual has custody of an entity, they will always have the power that comes with it, but they may not have the privilege to use that power if the entity was acquired illegally — such as if the banknote was stolen. This situation is analogous to a subject claiming a social identity and being recognized for it, even if they were not assigned to it. Thus, these different kinds of violations should be possible to address in a uniform way.

The model does not include physical substrates, or physical carriers, of information entities, such as identifiers, credential information and right records. In particular, it does not include physical credentials, such as passports or access cards, which verifiers would typically use in order to authenticate agents.

To address this limitation, the model could be extended by concepts from the Information Artefact Ontology [3].

In the proposed conceptual model, we have focused on persons, but as digital ecosystems become increasingly important, there is a need also to cater for organizations, systems, and things. We envisage that in such an extension of the model, it would be critical to investigate the rights and responsibilities of people, as well as entities such as organizations and systems, in order to ensure fair and ethical practices within the digital ecosystems.

A characteristic of social agency, in contrast to physical agency, is that a subject is only able to exercise its agency when it is recognized by some other agent. Thus, the subject's agency is temporary and only lasts for as long as they are recognized. For example, a person as a bank customer can only carry out transactions for a session, a limited time period. In the model, this fact is partially captured by a recognition relator associating a subject with a verifier and a social identity. However, this way of modelling is not explicitly indicating that the agency of a subject is temporary.

As illustrated in the vignette of the Corrupt Officer, an identity issuer could create a social identity without a corresponding identity assignment — a personal number for a Swedish citizen could be created though there is no corresponding citizen. This situation will come about if some rule of the process for creating a Swedish citizen is violated. And these rule violations may be caused by ordinary mistakes or even fraud. A similar situation may occur when a subject claims a social identity to which it is not assigned, but a verifier still accepts the claim.

One concern is about the consequences of such undesirable situations, i.e. statements about identity not being correct. If they go undetected, numerous social transactions and relationships may be dependent on them without anyone realizing that anything is amiss, and thus they are simply accepted without question. Still, it is (almost) always possible to question those statements, and if they are deemed incorrect and then corrected, the issue arises of what should happen to those transactions and relationships that were dependent on the incorrect statements. There is clearly no general answer to this issue, but in certain cases, they may be partially answered by laws and policies. It might be worthwhile to investigate how such regulations could be modelled.

An additional concern pertains to the recurrent potential for questioning the correctness of identity statements, i.e. if it is always possible to question those statements or if there is some limitation. One peculiar fact of those statements is that there is often some institution with the authority to determine their correctness, e.g. a supreme court. For instance, if someone's Swedish citizenship is in doubt, the case can be taken to various courts, but the ultimate decision of the Swedish supreme court is final — it can no more be questioned. It could be valuable to explore how such regulations could be modelled.

The two concerns above, about the consequences of identity statements and the potential for recurrently questioning them, are not limited to identity statements but apply to all social facts. The vignette of the involuntary voter provides an example of this, where a physical action is interpreted as a social one, even

though the person who performed the action did not intend it to be. Thus, the approach suggested in this paper can be expanded and adjusted to address instances of fraud and epistemic confusion in other scenarios.

Acknowledgements

Many thanks to Giancarlo Guizzardi for fruitful discussions on the topics of fraud and epistemic confusion.

References

1. Alexy, R., Rivers, J.: A Theory of Constitutional Rights. Oxford University Press, 2 edn. (Oct 2002)
2. Bertino, E., Takahashi, K.: Identity Management: Concepts, Technologies, and Systems (Information Security & Privacy). Artech House Publishers (Sep 2011)
3. Ceusters, W., Smith, B.: Aboutness: Towards foundations for the information artifact ontology. In: Proceedings of the Sixth International Conference on Biomedical Ontology (ICBO). CEUR vol. 1515. ceur-ws.org (2015)
4. Clarke, R.: A reconsideration of the foundations of identity management. 35th Bled eConference Digital Restructuring and Human (Re) action (2022)
5. Eriksson, O., Ågerfalk, P.J.: Speaking things into existence: Ontological foundations of identity representation and management. Inf. Syst. J. (isj.12330) (Mar 2021)
6. Eriksson, O., Johannesson, P., Bergholtz, M.: Institutional ontology for conceptual modeling. J. Inf. Technol. (Feb 2018)
7. Griffo, C., Almeida, J.P.A., Guizzardi, G.: Conceptual modeling of legal relations. In: Conceptual Modeling. pp. 169–183. Springer International Publishing (2018)
8. Guizzardi, G.: Ontological foundations for structural conceptual models. PhD Thesis, Enschede, the Netherlands (2005)
9. Guizzardi, G., Fonseca, C.M., Benevides, A.B., Almeida, J.P.A., Porello, D., Sales, T.P.: Endurant types in Ontology-Driven conceptual modeling: Towards OntoUML 2.0. In: Conceptual Modeling. pp. 136–150. Springer International Publishing (2018)
10. Hohfeld, W.N.: Some fundamental legal conceptions as applied in judicial reasoning. Yale Law J. **23**(1), 16–59 (1913)
11. Indu, I., Anand, P.M.R., Bhaskar, V.: Identity and access management in cloud environment: Mechanisms and challenges. Engineering Science and Technology, an International Journal **21**(4), 574–588 (Aug 2018)
12. Noonan, H., Curtis, B.: Identity. Metaphysics Research Lab, Stanford University, fall 2022 edn. (2022)
13. Preukschat, A., Reed, D.: Self-Sovereign Identity. Manning (Aug 2021)
14. Suchánek, M.: Quality — OntoUML specification documentation. `https://ontouml.readthedocs.io/en/latest/classes/aspects/quality/index.html` (2018), accessed: 2023-7-31