UNIVERSITÄT LEIPZIG

Onto-Med Report

An Analysis of Roles

Towards Ontology-Based Modelling
Master's Thesis

Frank Loebe

No. 6 – August 2003

Institute for Medical Informatics,
Statistics and Epidemiology (IMISE)

University of Leipzig

Liebigstr. 27
04103 Leipzig
Germany

-

Institute for Informatics (IfI)
Department of Formal Concepts

University of Leipzig

Augustusplatz 10/11
04109 Leipzig
Germany

Onto Med
Ontologies in Medicine

www.onto-med.de

Contact address authors

Frank Loebe

Institute for Medical Informatics,
Statistics and Epidemiology (IMISE)
University of Leipzig
Liebigstrasse 27
04103 Leipzig
Germany

Email: frank.loebe@imise.uni-leipzig.de

Imprint

Editors: Barbara Heller, Heinrich Herre, Markus Loeffler
Author: Frank Loebe
Institute for Medical Informatics,
Statistics and Epidemiology (IMISE)
Liebigstr. 27, 04103 Leipzig
Phone : +49 341 / 9716104, Fax : +49 341 / 9716130
Internet: http://www.onto-med.de/
Cover print by Buch- und Offsetdruckerei Herbert Kirsten
Copy deadline: 15. August 2003
© IMISE 2003
All rights reserved.
Reprint is only permitted with the explicit authorisation
of the authors and reference.
ISSN 1611-7352

An Analysis of Roles

Towards Ontology-Based Modelling

Master's Thesis[1]

Frank Loebe

Abstract. For decades, the concept of *roles* has frequently been discussed in the literature of various fields of computer science, such as conceptual modelling and knowledge representation, but an integrative definition of the concept has not yet been accomplished. On the basis of an extensive literature review, a general approach to roles is elaborated, which includes a system of classification. The motivation of this analysis comprises the integration of a well-founded concept of roles into the ontology-based knowledge representation language GOL (*General Ontological Language*).

The evaluation of the literature reveals two recurring concepts in connection with roles, called *player* and *context*. For instance, a `student` could be a role whose player is some `human` and whose context is determined by a relation to some `university`. The basis of the role approach being developed is formed by the hypothesis that roles characterise their player, while they themselves are determined by contexts. The GOL categories *relator*, *process* and *substance* can be identified as those ontological categories which cover all contexts in the literature reviewed. A distinct role type is postulated for each of these categories, which is then examined and formalised in particular with regard to the relation to the corresponding contexts. Discussions of problems as well as of properties of the representation based on these role types supplement these examinations. The extraction of common properties of all roles from the particular role types completes the approach.

In addition, the proposed role concept is compared to other ontological notions. A reinterpretation of some works selected from the reviewed literature shall demonstrate the expressiveness of the role approach. Referring to the motivation of this thesis, the axioms presented are proposed for integration into GOL. Further, an adaptation of the GOL representation of *material facts* is suggested in order to avoid problems which were previously shown to be solvable in terms of roles. Finally, a sample model from the domain of clinical trials shall clarify the application of the concepts developed. Moreover, it illustrates the potential precision of modelling based on these concepts.

Altogether, this thesis is intended as a contribution to knowledge representation on a clear ontological basis, paving the way for application-independent reuse of knowledge and other benefits.

---

[1] In May 2003, this report was submitted as a Master's Thesis to the Department of Computer Science at the University of Leipzig, Faculty of Mathematics and Computer Science.

Contents

List of Figures

List of Tables

List of Symbols

Standard Symbols

$\neg$	Negation
$\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$	Standard Connectives
$\dot{\vee}$	Exclusive-Or Connective
$\forall$, $\exists$, $\exists!$	Standard Quantifiers
$\sqcap$	Description Logics: Intersection
$\sqsubseteq$, $\equiv$	Description Logics: Subsumption, Equivalence
$\Box$, $\Diamond$	Modal Logic: Necessity Operator, Possibility Operator
$\mathbb{N}$	Set of Natural Numbers

General Ontological Language

$\mathscr{R}$, Relator	Category *Relator*
$\mathscr{S}$, Substance	Category *Substance*
$\mathscr{P}$, Process	Category *Substantial Process*
$\in$	Element-Of Relationship
$::$	Instantiation
$<$, $\leq$	Proper Part-Of Relationship, Reflexive Part-Of
$:\triangleright$	Containment
ontic	Ontical Connectedness
prb	Projection of a Process on a Time Boundary

Role-Related Symbols

$\mathscr{Q}$, Role	Category *Role*
$\mathscr{Q}_{rel}$, RelRole	Category *Relational Role*
$\mathscr{Q}_{proc}$, ProcRole	Category *Processual Role*
$\mathscr{Q}_{soc}$, SocRole	Category *Social Role*
SocSubst	Category *Social Substance*
$\rightsquigarrow$	Plays Relationship
$\succ$	Role-Of Relationship

Explicitly Defined Symbols

disjoint	Pairwise Disjointness
sip	Substance-in-Process Relationship
$\longrightarrow\!\triangleright$	Is-A Relationship

Special Text Styles

`Examples`	Typewriter Style
GOL-TERMS IN TEXT	Small Capitals
GOL-terms in formulae	Sans Serif Style

1 Introduction

Ontology is currently a catchphrase in some fields of computer science, such as knowledge representation and modelling. With the ascent of the Semantic Web[1], another catchphrase, ontologies became even more popular in the late 1990s than ever before. For example, users of the Internet have developed the need to depart from purely syntactical information retrieval, due to drawbacks such as producing huge amounts of irrelevant results. An ontology-related solution to such problems comprises the use of background knowledge about the queries posed by software applications, in order to control search processes or to select relevant results of such searches. Another main application for ontology is knowledge reuse: Humans are able to apply their knowledge in very different situations – a similar availability of knowledge, once formalised in a certain way in order to be reusable in different contexts, is extremely desirable. If the formalisation process is guided by a general methodology which might be related to the representation language employed, that would be another extremely helpful situation. In particular, software engineering would benefit significantly from such methodology since it would reduce the development and maintenance costs of software applications. Of course, this is not a new objective, and the first methodologies have already been developed and applied. However, current efforts with respect to ontology in computer science aim at the combination of the generality and reusability of represented knowledge, a representation on a formal basis, and methodological intuitivity in modelling and representation.

1.1 Motivation

It is this overall environment in which the GOL group[2] at the University of Leipzig emerged. Its aim is to provide a representation language with an underlying *top-level ontology*, i.e., an ontology pinning down the most basic categories of reality, in order to facilitate the development of domain-specific ontologies. This is not an isolated, purely theoretical approach, however. There is a strong association to medicine and therein the field of clinical trials, which serves as a source for case studies on the applicability and utility of the concepts[3] developed. Conversely, results (including intermediate results) in the development of the top-level ontology will be used in order to provide better information services for several clinical trial centers.

The concept of *roles* is one element of the GOL top-level ontology, which has only preliminarily been inserted and which has not been put on an elaborate, comprehensive footing. Some representative examples may establish a first intuitive impression on the subject-matter of roles. The most common example, which is alleged in almost every paper on roles, is the concept of `student`[4]. Other frequent sample roles are `project leader`, `food`, or more abstract terms like `agent` or `experiencer`.

[1] [Berners-Lee et al., 2001] contains a popular-science introduction. The Semantic Web is being developed by the World Wide Web Consortium (W3C), see http://www.w3.org/2001/sw/ for further information.

[2] GOL stands for General Ontological Language. See section 1.4 for a detailed introduction.

[3] Note that except where stated otherwise, the words *concept* and *term* are used in an intuitive sense throughout this thesis. *Concept* may be thought of as having a slight semantic connotation, while *term* owns a syntactic one. Both expressions are used to develop ideas, but they are not part of those ideas. Accordingly, *concept* and *term* should not be taken as technical expressions, and without any reference to a special meaning, such as, for example, in the field of terminology.

[4] Note that this change of letter style is always used to indicate examples. With regard to `student`, it is the standard example rather because it is familiar for everyone who has ever studied, instead of intrinsically being a particularly excellent role (according to whatever is associated with the term *role*).

Role terms are usually contrasted with natural, independent types of concepts, e.g., `human` or `apple`. Some evidence has to be found to support the current position of roles within GOL, or to explain different relationships between roles and other GOL categories. Altogether, a detailed analysis of this concept is necessary, and this leads to the inspiration for this thesis.

Regarding applications for such theoretical work, a first review of the literature yielded a promising number of approaches in different areas, all of which use some notion of roles. Since the late 1970s role concepts are introduced in such fields of computer science as knowledge representation and data/conceptual modelling. In the field of conceptual modelling, the main motivation for many approaches is the problem of capturing substantial, complex changes of represented entities while retaining their individual identity. For instance, if a `student` becomes an `employee`, the "underlying" `person` remains the same, although her/his properties change remarkably. Roles are examined in order to model such cases easily and efficiently. The questions associated with roles in knowledge representation differ from those in conceptual modelling. Examples for the former are questions on the connections between roles and relations, as well as differences of such representational elements as attributes, natural types, and roles. Nevertheless, on an intuitive level, there seem to be underlying real-world entities, which are referred to in terms of the same or similar examples in both fields. Moreover, the line which was just drawn between role issues in these fields is not sharp, as will be clear from chapter 2. According to these considerations, the ontological clarification of the concept of roles proposed in this thesis will have applications beyond the realm of GOL.

1.2 Objectives

The main objective of this thesis is to provide a detailed analysis of the concept of *roles*. From this it shall be possible to extract the ontological features of roles, which should either justify or change its position within the GOL top-level ontology. A corresponding proposal is made in section 5.1, where roles turn out to be distributed over the current top-level ontology, in such a way that they cannot be attached in general to one of the more specific GOL categories. This proposal is based on an axiomatised account of roles which is intended as an extension of the current axiomatisation of GOL. In the analysis of literature which precedes the development of our own approach, several individual formalisms as well as subareas of computer science are to be covered. Of course, this examination has to be restricted, since even within computer science a considerable number of role approaches exist. In addition, three "background areas" shall be taken into account which have stimulated computer science formalisms and which are likewise influential for this thesis.

Furthermore, an attempt will be made to find a comprehensive characterisation or a unifying model of roles for conceptual modelling in computer science. In the course of this examination, some problems and modelling issues with respect to roles will be elaborated, thereby providing methodological suggestions for representation and modelling. Additionally, the investigation of differences between the notion of roles and other generic modelling concepts shall result in an improved understanding of these concepts and their relationship to roles. A similar incentive suggests the re-interpretation of existing role approaches in terms of the notions developed, which, furthermore, demonstrates that major role approaches can be expressed. Finally, a domain model is to be developed, which is related to a software application in the context of clinical trial management. This sample application uses concepts derived from our notion of roles and some other elements of the GOL ontology. It is supposed to sketch the link between theory and reality.

1.3 Structure of the Thesis

The thesis is organised as follows. This introductory chapter concludes with an overview of the *General Ontological Language* (GOL) and some notes on *Formal Ontology* as a precursor. Chapter 2 presents various role concepts from the 1970s to current works. The focus is set on works in knowledge representation and conceptual modelling, and some remarks on disciplines outside computer science which have their own role notion and/or are relevant for one or the other role approach in computer science are given at the end of the second chapter.

In chapter 3 an approach to roles is developed as an attempt to grasp the commonalities and differences in what is presented in the state of the art (i.e., the second chapter). This does not mean that a single notion of roles has been found; rather, three different kinds of role concepts which may be regarded as ontologically pure are identified and discussed. Each notion is introduced in the form of an axiomatisation, which is subsequently elucidated. In addition, effects and implications of each of these role concepts are elaborated, and, finally, a more abstract notion of roles is extracted at the end of this chapter, comprising the common features of all role types.

In the fourth chapter, the discussion is extended to some other ontological notions in general (like *quality*, *function*, etc.) as well as to some of the works from chapter two. The goal of this part of the thesis is to describe the relationship of roles to similar notions frequently confused with roles, and to show inadequacies of existing approaches which mainly originate from a mixture of, or a focus on one of the three different role concepts addressed in chapter 3.

The fifth chapter applies the presented role approach. Firstly, an integration of those roles into GOL is proposed, followed by some remarks on how to understand GOL's concept of *material facts* in terms of one of the role subtypes developed. Thereafter, a small domain example is elaborated in the medical subdomain of clinical trials. This aims at bridging the gap between theoretical examinations and software development in practice.

The thesis concludes with chapter six which presents a summary of the main results and outlines a number of further directions of research that may be pursued, ranging from remaining tasks for the integration of the role concept(s) into GOL to the application of our approach in large-scale domain modelling, thereby testing its utility in practice.

1.4 The General Ontological Language (GOL)

At a special position compared to other formalisms, this section introduces the *General Ontological Language* (GOL). The choice not to include the GOL part in the state of the art chapter (i.e., chapter 2) was made for two reasons. On the one hand, GOL is introduced much more comprehensively than other formalisms with respect to the number of concepts not directly related to roles. The presentation of other works from the literature is limited to the idea of roles and the "surrounding" concepts which are necessary for understanding. The distinctions underlying the GOL top-level hierarchy, however, are to be seen as a basis of explanation for other works. Therefore a more comprehensive introduction is necessary. The second reason for excluding GOL from chapter two is that the notion of roles in GOL is rather preliminary and subject to debate in this thesis. Consequently and in contrast to the reviews in the following chapter, it hardly serves as a basis for the role approach developed herein.

Preceding the actual introduction of GOL, some remarks shall now adumbrate the research area in which it is developed. *Formal ontology* is a field of philosophy which has recently initiated a closely related field in knowledge representation, and thus within computer science. *Formal Ontology* and *Formal Ontology in Computer Science* are considered kindred yet distinct for two reasons. Formal ontology in computer science integrates present formal ontological principles, rather than engaging in (but not completely precluding) traditional, philosophical discussions of basic assumptions, or devising new ones in formal ontology. Moreover, the inspiration for both areas is somewhat different: the philosophical branch aims at a rigorous theory of "[...] the organisation and the nature of the *world* independently of the form of our knowledge about it." [Guarino, 1995]. Its counterpart in computer science is faced with problems originating from applications developed with implicit ontological commitments. For instance, software agents may use the same syntactic elements and seemingly be capable of communicating with each other. But there may be different intuitions or intentions of their developers, causing unexpected and erroneous behaviour. The goal is to exploit formal ontological results to solve such problems. One benefit may be more powerful and reliable applications; another could be reusability of domain models. Of course, the line just drawn between formal ontology and formal ontology in computer science is not fixed at all. Rather, both areas overlap and stimulate each other. A recent definition[5] of formal ontology shall conclude this short introduction to the field. It marks formal ontology as

> the systematic, formal, axiomatic development of the logic of all forms and modes of being. [Cocchiarella, 1991]

Nicola Guarino, a major proponent of the introduction of formal ontological principles into computer science, appreciates this definition in [Guarino, 1995] for the possibility of a double interpretation of *formal*: once referring to a *rigorous* treatment of the subject, and once associated with *all forms* of being.

As mentioned above, the development of GOL is an activity within formal ontology of computer science[6]. GOL is a formal language based on predicate logic, which is intended for building (domain-)ontologies. In order to do so, it exploits a top-level ontology called *General Formal Ontology* (GFO). The GFO is GOL's main component, offering language elements whose semantics is captured axiomatically. In the present stage, the GFO provides the first parts of a formalised and axiomatised top-level ontology. GOL altogether is still in development by the GOL group, but first papers which elaborate the framework are published in [Degen et al., 2001a,b; Guizzardi et al., 2002a,b]. Future work comprises the development of a new semantics in addition to the standard model-theoretic semantics of predicate logic, as well as the ongoing extension and evaluation of the ontology by application (first of all in medicine).

Figure 1.1 shows a taxonomy of several basic GOL categories according to the forthcoming GOL Reference Manual 1.0α [Degen et al., 2003], intended as a quick reference throughout this thesis.

[5]This definition is not to be seen as one of overall validity and acceptance. An account with the latter properties remains a matter of debate. See also [Guarino and Giaretta, 1995] for an account of several definitions of formal ontology and related terms.

[6]Hereafter, the terms *formal ontology* and *ontology* refer to their use in the computer science area. Note further, that *ontology* in philosophy exclusively refers to a discipline. In contrast, there may be various *ontologies* (with respect to computer science), i.e., descriptions of (parts of) the world, often expressed in a formal language. Put differently and as often cited, an ontology in computer science is an "explicit specification of a conceptualization." [Gruber, 1993] (whereas this work has also been criticised, e.g., in [Guarino and Giaretta, 1995]).

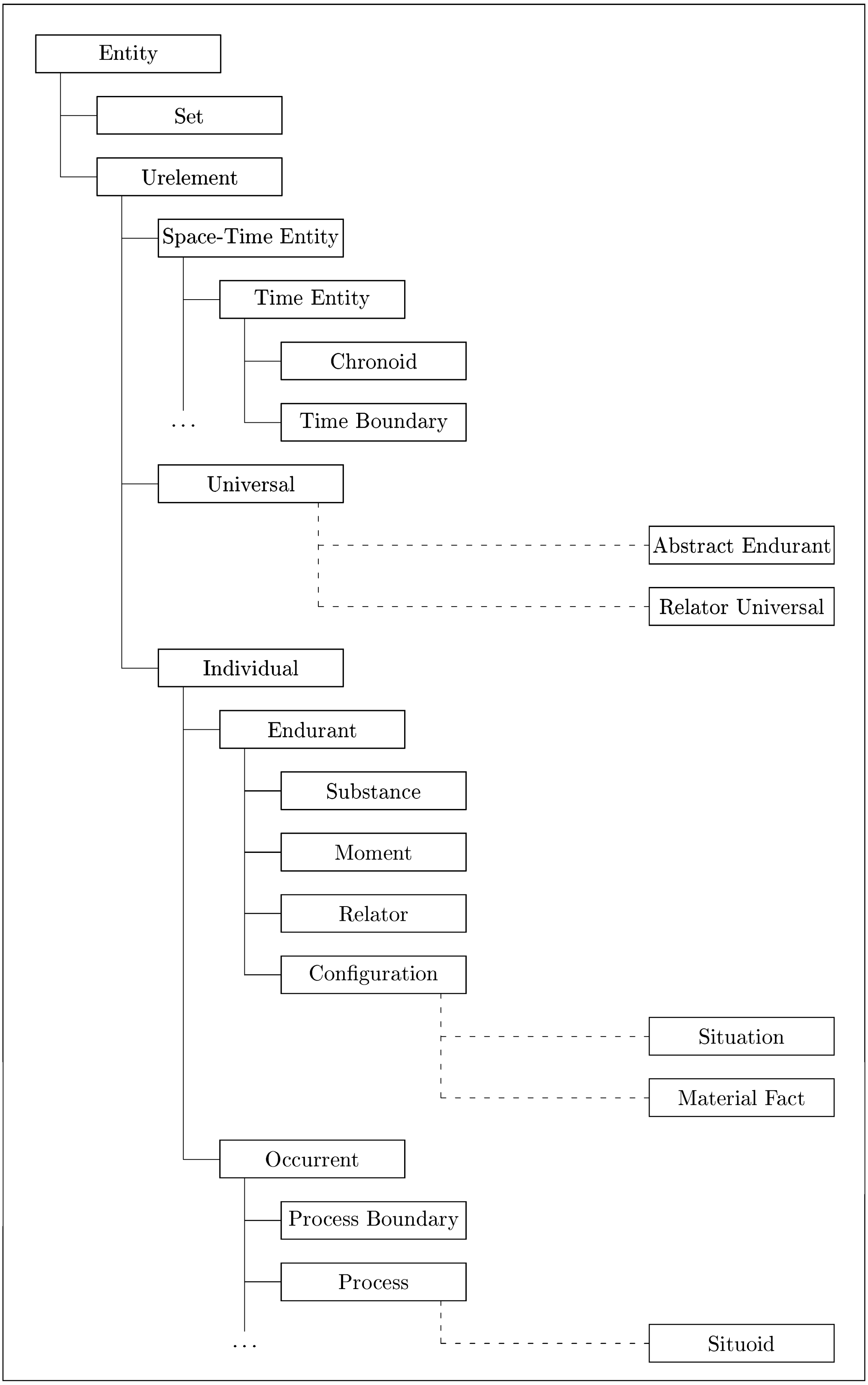

Figure 1.1: Taxonomy of Selected GOL Categories.

The ontological commitments behind these categories will be sketched below, since they are used as a reference model for those formalisms introduced in chapter 2. Note that Figure 1.1 is not to be understood as an exhaustive summary of the GOL Reference Manual; instead, it is limited to those GOL entities which are used in the sequel. Two different line styles are used in order to express the following difference with respect to the assumptions about cross-categorical relationships: the categories which are situated on the left-hand side of Figure 1.1, connected by solid lines, are viewed as being part of the top-level hierarchy of the GFO, whereas the five categories on the right, attached by dashed lines, are rather indirectly related to the categories they are attached to. The latter do not represent distinctions of the same importance as those on the left, and they need not even be mutually exclusive. These categories are included in Figure 1.1 in order to provide a common place of reference. In contrast, all categories on the same level of the top-level hierarchy are considered to be mutually exclusive[7]. Moreover, the first two levels of distinction below ENTITY are complete partitions of the corresponding supercategories. Ellipsis in subsequent levels indicate that further subcategories for the according level are described in [Degen et al., 2003]. Notice that there may be orthogonal partitions of each category, and that further levels of distinction are subject to debate as GOL is still being developed. Further, the top-level hierarchy is considered to capture the most basic distinctions within the GFO.

The realm of existing things is denoted by ENTITY[8], which is the top-node in Figure 1.1. It can be divided into SETS and URELEMENTS[9], where urelements are only limited by the condition that they cannot be analysed in set-theoretical terms, i.e., in particular, via the ELEMENT-OF relation. Sets have already been captured by means of the axiomatic system of Zermelo-Fraenkel, which is adopted for the set-theoretic part of GOL. Accordingly, most effort needs to be put into the classification of urelements, as well as a corresponding axiomatisation.

At the first level of distinction, urelements are separated into INDIVIDUALS, UNIVERSALS, and SPACE-TIME ENTITIES. Individuals are single entities, while universals are INSTANTIATED by several individuals[10]. The collection of all instances of a universal is a set, called the *extension* of a universal. In addition, there is an *intension* of each universal which can be captured in terms of axioms describing the relationships to other entities. With respect to individuals, the individuation principle in GOL is *location in space and time*, therefore space-time entities are not subsumed by either universals or individuals. Time and space are strictly separated in GOL, where spatial regions are called TOPOIDS and time intervals are named CHRONOIDS, both being basic entities in their own right, i.e., not defined in terms of other notions, for example a set of points. Each chronoid has exactly two TIME BOUNDARIES, which depend on this chronoid. Time boundaries may coincide, and two chronoids may stand in a PART-OF relationship. Further details about the connection between chronoids and other entities are propounded in section 3.1.2, since these are necessary for the understanding of our axiomatisation of role concepts in chapter 3.

[7]Note that there is a possible exception. In contrast to Figure 1.1, some relators may likewise be considered as moments, in the sense of relational moments. In the current version of GOL, however, relators are described separately from moments. Further, different formal relations apply to moments and relators. Therefore, both categories are shown as distinct subcategories of ENDURANT.

[8]As a remark on style, specific GOL terms are written in small capital letters if they are not present for longer passages within a section. Thereafter they appear in the usual text style.

[9]Sets and urelements are said to have type 0. In addition, on top of them a meta-mathematical superstructure can be built using CLASSES (in a mathematical reading) of type 0 and higher-order types. Herein, this class structure is not explicitly used.

[10]Intuitively, the basic ideas of *types* or, in object-oriented terms, *classes* are fairly similar to the concept of universals.

Concerning universals, hitherto no proper subclassification has been specified. Naturally, one possibility is induced by the subclassification of individuals that is established. Individuals are basically divided into ENDURANTS and OCCURRENTS, where the corresponding criterion refers to "interaction" with time. Endurants exist in time as any individual, but they do not have temporal parts. More precisely, an endurant exists completely at a time boundary. Contrariwise, occurrents are extended in time, i.e., they are not restricted to a single time boundary. There are several subtypes of occurrents, the most important of which herein are PROCESSES. These happen in time and have temporal parts. Examples of processes are a `running`, a `laughing`, a `discussion`, and so forth. Conforming to time boundaries, there are PROCESS BOUNDARIES which border processes. The remaining occurrent types in [Degen et al., 2003], e.g., CHANGE or HISTORY, shall not be discussed here. Note that there are two ways of considering processes – either with or without the endurants participating in a process. GOL follows the second approach, since endurants are to be found "within" process boundaries. Again, section 3.1.2 presents an account of this issue in greater detail.

Endurants themselves are separated mainly according to their degree of independence compared to other endurants. Properties of endurants are themselves endurants, called MOMENTS, and they INHERE in SUBSTANCES. The latter are a relatively independent type of endurant in which several moments inhere, and which does not inhere in any other endurant. For example, if a particular `colour` inheres in an `apple`, there is a dependence between the colour and the apple in the sense that the colour needs that apple for its existence, while the apple may change its colour. Section 2.3.1 also addresses the philosophical notion of existential dependence, but there is too much literature on this topic which prevents an exhaustive introduction herein. Intuitively, inherence may be understood as "having the property of". Notice that moments tentatively include the present GOL category of roles (cf. [Degen et al., 2003]), which shall not be further elaborated here, as this insertion of roles does *not* serve as a basis for this thesis.

The above example of an `apple` changing its `colour` touched on the area of identity problems, and therefore the notion of ABSTRACT ENDURANTS should be mentioned. As shown in Figure 1.1, abstract endurants are universals, providing for an account of diachronic identity. Diachronic identity means identity over time or identity at different time boundaries, respectively. The issue of how one may understand the connection between two "static" entities (like an `apple`) at different points in time has been a subject of debate in philosophy for hundreds of years. Matters are complicated by the fact that these entities may change due to their participation in processes. GOL provides its own solution to this problem on the basis of philosophical endurantism. The latter assumes that two types of entities exist, namely endurants and occurrents. Endurantism is usually contrasted with perdurantism as one of the most basic ontological choices. Perdurantism considers only one type of entity, whose instances are extended in space *and* time. Therefore the approach is sometimes called four-dimensionalism and its entities four-dimensional space-time worms. At the moment, endurantism is being explored in the development of GOL, but future versions shall accommodate several top-level ontologies which may be of either kind. Deferring a detailed introduction of the connection between endurants, occurrents, and abstract endurants to section 3.1.2, endurants exist at exactly one time boundary, while abstract endurants capture entities which persist through time[11].

[11] The prefix "abstract" may be omitted in the following chapters, if it is clear from the context, or irrelevant for an issue, respectively, which type of entity is referred to. Note further that there are, e.g., ABSTRACT SUBSTANCES and ABSTRACT MOMENTS, due to the hierarchical relationships shown in Figure 1.1.

On the basis of an understanding of processes and endurants, SITUATIONS and SITUOIDS can be introduced. The latter are special processes, whose boundaries are situations. Situations are specialisations of CONFIGURATIONS, which are complexes of interlinked endurants, i.e., they may CONTAIN substances, moments, and RELATORS (see below). In addition, certain universals are associated with a situation in order to cover a specific range of granularity and specificity. In this connection, a situation has to satisfy certain conditions of unity with respect to its associated universals. To a certain extent the notion of situations was influenced by the situation theory of Barwise and Perry [1983]. Similar to situations, situoids fulfill certain principles of coherence and continuity. They are entities with actual existence, considered to be the most independent entities. A `concert performance` in total may be considered as a situoid, i.e., comprising the `orchestra` with the `musicians` and their `instruments`, the `audience` and their `seats`, the `air` in the concert hall, the `hall` itself, and, of course, the `music`. On a certain level of granularity, a situoid cuts out a "sensible", connected part of reality. Note that situations and situoids are in a preliminary state of development to a greater extent than other categories. For instance, the association of universals and the principles of unity or coherence, respectively, have been sparsely elaborated yet. Therefore, either notion will not be granted a central position in this work.

Finally, there are entities which mirror connections between those entities from above. GOL introduces a theory of RELATIONS – entities establishing more complete wholes. The main feature to distinguish relations, in turn, is *formality*: a relation is *formal* if it holds of its relata as soon as these are given, i.e., without any reference to further entities. Otherwise a relation is called *material*. Examples of the latter are relations like `kiss`, `loves`, `owns` and so on. They are mediated by RELATORS, individuals which connect other entities on the basis of a FOUNDATION. For the sake of illustration, if `Paul owns a ring`, there must have happened some process (like a `buying`) whose (former) existence founds a relator mediating between `Paul` and the `ring`. Moreover, a relator and the entities it connects form a new individual, which is called a MATERIAL FACT. In order to establish a formal connection between the constituents of a material fact, the formal relation of HOLDING is introduced[12]. Formal relations do not rely on relators, but they are conceived as (mathematical) classes[13]. Further, certain general formal relations are called *basic* relations. They are employed in the axiomatisation of GOL without being explicitly defined in terms of other notions. The above mentioned concepts of INHERENCE and INSTANTIATION, for example, fall into this category, as well as the PART-OF relationship. The latter is an example of a formal relation with specialised subrelations, like TEMPORAL PART-OF, SPATIAL PART-OF, and CONSTITUENT PART-OF[14].

For further information about the ontology of GOL the reader is referred to the Reference Manual [Degen et al., 2003] or other works mentioned. The main ideas should have been sketched sufficiently to have a common ontological basis for the following presentation and analysis of formalisms. With the introduction of our approach in chapter 3 some required notions are described in greater detail.

[12] Section 5.1.2 suggests another approach in order to explain the connection between a relator and those other entities on the basis of one of the role types defined in chapter 3.

[13] With regard to footnote 9, therefore formal relations are not included in Figure 1.1, because they are not entities of type 0.

[14] This specialisation of PART-OF was introduced very recently into GOL. We do not use these specialised relations in the sequel, but only the general version PART-OF. This is not considered disadvantageous, however, since the current distinctions are yet easily derivable from the context.

2 State of the Art

The goal of this chapter is to provide an extensive overview of role concepts in several selected formalisms in the fields of knowledge representation and conceptual modelling, independently of the explicit use of the term *role*. Conceptual modelling is understood as referring to data modelling and modelling in software engineering (in particular modelling in the analysis stage of the software development process). In general, the classification of formalisms into either area is not to be taken too strict, because there are some borderline cases which may be assigned to either field. Although we focus on roles and adjacent concepts in each approach, omitting most other details, this overview can certainly not be an exhaustive treatment of *all* works on roles, as there are too many (cf. the bibliography of [Steimann, 2000b]).

In addition to works in computer science, other fields are to be taken into consideration, which have had influences on computer science formalisms. Philosophy, linguistics and a smaller field, *role theory*, overlapping sociology and psychology have been identified for this. In this connection, knowledge representation may be linked to mainly philosophical and linguistic sources, whereas few object-oriented formalisms appear to have drawn on ideas from sociology and psychology[15].

The chapter is structured as follows. At first, an overview of role concepts in computer science since the late 1970s is presented in the first two subchapters, which correspond to the fields of knowledge representation and conceptual modelling, respectively. The final subchapter introduces notions of roles or related concepts from philosophy, linguistics, and sociology/psychology in this order.

2.1 Knowledge Representation

Several formalisms in knowledge representation have introduced role concepts. Current examples are description logics (cf. [Baader et al., 2003]) as well as Sowa's conceptual graphs [2000]. In the latter, roles are described in an ontological framework, while there is no syntactic element of conceptual graphs explicitly reserved for the representation of roles. An earlier approach with a present application is that of *scripts* (developed by Schank and Abelson [1977]), in which roles occur in process descriptions. Moreover, Guarino et al. (cf. [Guarino, 1992; Masolo et al., 2002]) have pursued studies in formal ontology for at least ten years, which exhibit strong connections to philosophy. They have also developed a framework containing roles, which are put in relation to types and other metacategories [Guarino and Welty, 2000]. Finally, *topic maps* [ISO, 1999] are a recent standard of the International Organization for Standardization (ISO) with an interesting notion of roles.

2.1.1 Description Logics

The field of description logics[16] (also terminological logics) deals with formalisms closely related to predicate logic, but originating from works on semantic networks and frame-based languages, especially KL-ONE [Brachman, 1985]. Description logics have a Tarski semantics and are used to represent and reason over concept hierarchies[17]. In contrast to full first-order logic, description logics strive for an acceptable trade-off between expressiveness and computational costs.

[15] If at all noticeable, the influence of role theory is often implicit, i.e., hardly ever mentioned in object-oriented works.

[16] The most recent, comprehensive presentation of the field is [Baader et al., 2003]. [De Giacomo, 2001] may also serve as an overview, emphasizing the relation between description logics and modal logics.

[17] In this section, *concept* and *concept name* are specific terms according to the field of investigation.

Since a large variety of description logics and/or families of such logics exists, *ALC* [Schmidt-Schauß and Smolka, 1991] is chosen as a common representative. An alphabet in ALC consists of *concept names*, *role names* and *object names*. This corresponds to most other descriptions logics. As a full understanding is not necessary here, it is sufficient to note that concept names are interpreted as unary relations over the universe of objects, role names as binary relations over the same universe.

In regard to roles, there are two ways in ALC to define new concepts involving roles. Let R be a role name, C be a concept name. Then $\forall R.C$ and $\exists R.C$ denote new concepts which use R and C. Provided that $\mathscr{R}$ and $\mathscr{C}$ are first-order predicates corresponding to the ALC role and concept name, the application of the defined concepts to an object x corresponds to the following first order formulae (note the standard use of "iff" as an abbreviation of "if and only if"):

$$\forall R.C(x) \text{ iff } \forall y\,(\mathscr{R}(x,y) \rightarrow \mathscr{C}(y)) \tag{2.1}$$

$$\exists R.C(x) \text{ iff } \exists y\,(\mathscr{R}(x,y) \wedge \mathscr{C}(y)) \tag{2.2}$$

Other description logics allow further "operations" on role names, like inversion or the transitive closure, but details of these extensions are not to be elaborated here. From an ontological point of view, roles in description logics correspond to arbitrary binary relations. No further restrictions are given by the formalism, and only constraints on the role *filler*[18] can be directly expressed. It is interesting to note that, firstly, role names are usually not employed exclusively to express RELATIONS, but, additionally, to represent other categories in GOL, in particular MOMENTS. Secondly, role names are sometimes what one would expect to be a relation name, while sometimes they correspond to unary predicate names. Examples for the former may be `has` or `child-of`, for the latter `child` or `owner`. Illustratively, suppose that the concepts `Male` and `Human` are available. The concept `Father` is now to be defined using `Male` and somehow expressing a relation to a `Child`. Both of the following sample formalisations[19] are possible in ALC. From a methodological point of view, though, one of them should be preferred, because this should simplify modelling[20].

$$\texttt{Father} \equiv \texttt{Male} \sqcap \exists \texttt{has.Child} \tag{2.3}$$

$$\texttt{Child} \sqsubseteq \texttt{Human} \tag{2.4}$$

$$\texttt{Father} \equiv \texttt{Male} \sqcap \exists \texttt{child.Human} \tag{2.5}$$

Before concluding this section, attention shall be drawn to an early paper on KL-ONE [Brachman, 1985], a formalism which is a major source of description logics. In KL-ONE, roles are not only intended to model binary relations. Instead, they are equated with a notion of generalised attributes, used to model the conceptual components of an entity. They should represent a variety of entity types that appear in connection with a concept: "parts [...], inherent attributes of objects and substances [...], arguments of functions [...], and 'cases' of verbs in sentences [...]" [ibid., p. 208].

[18] This term refers to the y in formulae 2.1 and 2.2, respectively.

[19] The first formalisation comprises formulae 2.3 and 2.4, the second one only 2.5. $A \sqsubseteq B$ says "Concept A is subsumed by concept B", and $A \equiv B$ stands for the equivalence of A and B, i.e., $A \sqsubseteq B$ and vice versa. $A \sqcap B$ denotes the intersection of the concepts A and B. Description logics are used for reasoning about subsumption hierarchies defined in such a way.

[20] This kind of guidance is one of the open problems GOL attempts to tackle. Even if several variants are offered by a formalism, their interconnection should be contained implicitly in the formalism, such that there is no need to explicitly state this for each case in a domain.

Brachman further argues that such a generalised attribute has two major components. First, there is the *filler*, i.e., the entity which is the part, the value of the attribute, or the argument of a function in the above quotation. Secondly, there is the actual *functional role* in the conceptual complex the attribute belongs to. The syntactic element *role* captures both aspects by means of several *facets*. A part of an example from [Brachman, 1985] is the concept of an `arch` with a role `R1` that has the following facets: (1) the *name* `lintel`[21], (2) the *modality* `obligatory`, (3) a *value restriction* of the filler to the concept `wedge-brick`, and (4) a `cardinality` of one. Now, one realises that this use of roles is much more structured, while only the role name and the value restriction for the filler have survived in the notion of roles in description logics.

2.1.2 Scripts

Schank and Abelson, who may be assigned to the field of computational linguistics, developed *scripts* in order to represent common courses of events[22] within certain situations [Schank and Abelson, 1977]. Originally, scripts had a close relation to Schank's conceptual dependency theory [Schank, 1972], a theory of natural language understanding which exploits four conceptual primitives in order to uniformly capture the semantic structure of natural language expressions. Nowadays the main application of scripts still lies in natural language processing, for instance in structuring the knowledge base of a language processing system according to situations relevant for that system (cf. [Luger, 2001, p. 239]).

Scripts consist of several components: entry conditions, results, props, roles, and scenes. Props comprise things and facts, not only pure properties. Scenes are used to divide a script into smaller units. As an example, Luger [2001, p. 240–241] refers to a `restaurant` script. Props therein correspond to standard assumptions, for instance that a restaurant contains tables and offers a menu. Scene examples are `entering`, `ordering`, `eating`, and `exiting`. Entry conditions and results describe propositions that need to hold before and after the script is executed. Luger provides the following role definition in this context:

> *Roles* describe the actions of the single participants. The waiter receives the order, brings food and the bill. The guest orders, eats and pays. [ibid., p. 240, our translation]

It should be noted that in Luger's example there is also a role `owner` which does not occur in any scene, but only within a result. There it is defined as the one who gets the money. This example raises the interesting question of why an owner is "participating" though no owner is actually present in any scene. Section 2.3.1 provides the idea of ontological levels to address this question in general.

Recent work on processes is also based on scripts. Aitken and Curtis [2002] describe an extension[23] of scripts in the framework of the Rapid Knowledge Formation project. In this approach processes are formalised as scripts on a type-level, i.e., using UNIVERSALS. Additionally, the authors introduce a way of expressing how many entities of one type appear in a script. Then there are several relations connecting scripts, types, and roles, like *typePlaysRoleInScript* or *someInstancePlaysRole*. Formally, roles are used as binary predicates on the instance level, relating an event and an object.

[21] This name refers to a part of the `arch`.

[22] A course of events corresponds roughly to the notion of PROCESS in GOL. Thus, scripts can be understood as describing process universals.

[23] This extension is also based on CYC [Lenat and Guha, 1990], a well-known project in which a very large common sense knowledge base was developed in an ad hoc manner; this approach is often said to have failed, however.

One also finds an *actors* relationship, which "is used as the most general predicate relating events and instances" [Aitken and Curtis, 2002, p. 112]. In conclusion, every way of participation in a process is expressed by a corresponding role predicate in this formalism.

2.1.3 Roles in Sowa's Ontology

In 1984, John F. Sowa introduced a knowledge representation formalism called *conceptual graphs*. Conceptual graphs are a kind of semantic networks, the latter being developed in artificial intelligence and computational linguistics during the late 1960s. Sowa uses his formalism throughout the recent book on knowledge representation, [Sowa, 2000], which is the main reference work in this section. Conceptual graphs are defined as bipartite graphs whose nodes divide up into *concepts* and *conceptual relations* [p. 477]. Two or more concepts are linked to a conceptual relation, while concepts cannot be related directly by a graph's edges. See Figure 2.1 for an example. Conceptual graphs, like semantic networks in general, do not contain a syntactic element which is called or directly related to roles. Therefore, this intuitive account of syntax shall suffice. We now turn our attention to Sowa's top-level ontology which discusses roles.

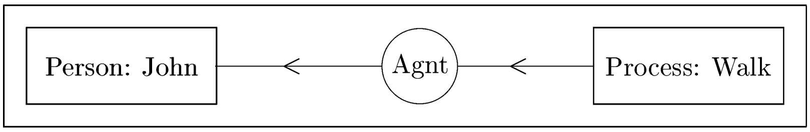

Figure 2.1: Example of a Conceptual Graph
Representing the Sentence "`John is walking`."

In his second chapter, entitled "Ontology", Sowa presents his own top-level ontology which is based on the work of several philosophers, but foremost on that of Peirce (cf. [p. 60 ff.]). The introduction of his ontology is limited herein to those elements directly related to his notion of roles (see Figure 2.2). Sowa inserts roles under a category called *actuality*. This category comprises entities which are *independent* and *physical*, the latter as opposed to *abstract*. Sowa's criterion to distinguish between physical and abstract entities is the distinction between "consisting of matter/energy" and being a "pure information structure" [p. 68]. Examples of entities under actuality could thus be the following: a `cat`, an `electromagnetic wave` or a `running elephant`. A `word` is an example of an abstract entity in this scheme, as it does not consist of matter/energy. It should be mentioned that Sowa makes a distinction between *objects* and *processes*, which are derived from the actuality category, in conjunction with one of the categories *continuant* (for objects) or *occurrent* (for processes). However, for Sowa the differentiation between continuants and occurrents is a question of viewpoint, which differs from the validity of the GOL distinction between ENDURANTS and OCCURRENTS[24], because the assignment of an entity to one of these categories is definite in GOL. Explaining his process category Sowa states:

> [...] Even a diamond could be considered a process when viewed over a long time period or at the atomic level of vibrating particles. [p. 73]

[24]This distinction is the best analogy to continuants and occurrents, although concrete endurants do not persist through time, like continuants. Moreover, Sowa's *objects* are understood as being comparable to SUBSTANCES, his *processes* to PROCESSES. Note that the relationship between ABSTRACT SUBSTANCES and the notion of SUBSTANCE PROCESSES (cf. [Degen et al., 2003]) reflects Sowa's approach of viewing an entity as an occurrent or a continuant.

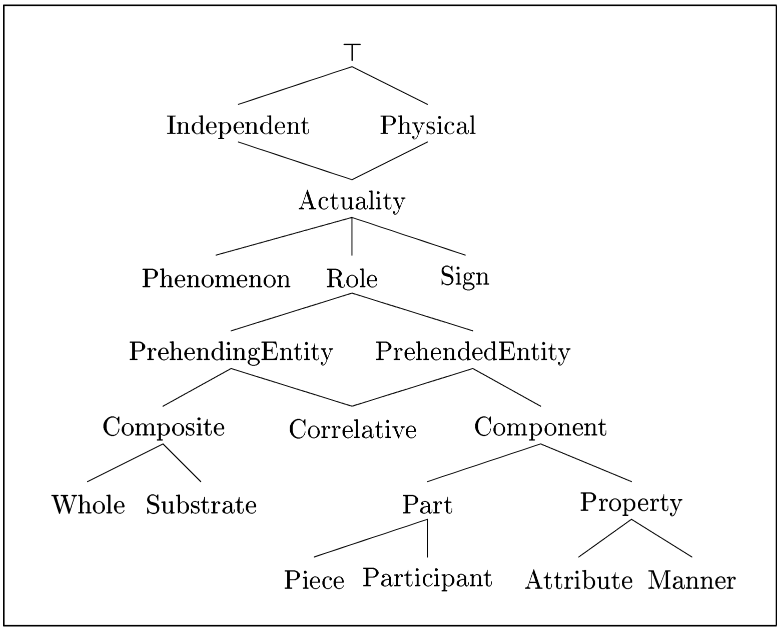

Figure 2.2: The Ontology of J.F. Sowa: Role-Related Part. (cf. [Sowa, 2000, pp. 87–89 and pp. 502–512])

Focussing on roles, these arise from the application of Peirce's trichotomy of Firstness, Secondness and Thirdness to actuality. In short, Firstness refers to "[...] properties that can be recognised without regard to any relationships to other entities.". Secondness comes into play if relationships to other entities are considered, and "[t]hirdness focuses on the mediation that brings the first and the second into relation." [p. 61]. Examples are `Woman`, `Mother` and `Motherhood`, respectively, where the first two refer to the *same* individual, once by a Firstness term, once by a Secondness term. "Motherhood [...] relates the mother and the child." [p. 62], and is thus an example of Thirdness. Obviously, `Motherhood` refers to a different individual compared to the First- and Secondness terms.

Applied to actuality, Sowa presents a distinction between *phenomenon*, *role* and *sign*. Phenomena are structural types which are characterised by inherent qualities and forms. Role types characterise entities referring to relationships to another entity. Finally, signs depend "[...] on the triadic relation of *representation*: a sign *x* represents something *y* to some agent *z*." [p. 81]. Interestingly, these categories can be applied to one and the same individual, as Sowa's examples show.

Having made this division of the class of actual entities, some properties of role types are given by Sowa, mainly on pages 80 and 81. Role types are not restricted to certain UNIVERSALS, i.e., instances of different universals can instantiate the same roles. Formally, they always depend on *binary* relations, though it is mentioned that in natural languages there is often a focus on one argument only. The classification of something in a role does not refer to the structure of an entity, which is the case for *phenomena*, and a *context* is essential for classifying an entity by role. Sometimes phenomenal types may be predicted for a role instance, but a strict implication is rarely applicable. Furthermore, [p. 267] allows the generalisation of an example of a `baby` and the same person as a `man`, i.e., one individual at different times. Sowa admits different roles for different stages in life, and the instances of some role types may be restricted to being in certain stages according to other universals.

Apart from properties that pertain to each role type, Sowa specifies a subclassification of role types (cf. Figure 2.2) based on several distinctions which are discussed in detail in chapter 4.2.3 below.

The criteria he adopts lead to further restrictions on roles. In particular, roles have to pass the *has test* [p. 84], i.e., for a relationship and its two roles X and Y, the sentence pattern "X has Y" needs to sound natural. This is a means to distinguish between the so-called *prehending* (the X in "X has Y") and the *prehended* entity (the Y), terms that stem from Whitehead, which stand for the subject and the datum of a given relationship [p. 64]. Suitable examples are "The car has an engine." and "A mother has a child.", while "The engine has a car." sounds odd [ibid.]. Consequently, role types can be found only in connection with a true subclass of the class of binary relations, which is restricted to relations whose roles pass the *has test*.

Sowa also bridges towards linguistics by inserting theta-roles under a subcategory of role [p. 511 ff.], to wit *participant*. This is interesting insofar as theta roles occur in connection with verbs (cf. section 2.3.2 herein). Verbs, though, often refer to processes, and not exclusively to relations. The appropriateness of this insertion depends on the question of how strict relations and processes are separated, where GOL advocates a clear distinction between both categories. See section 3.4.4 for a different approach to theta roles, and section 3.6.3 for a brief discussion of the connection between relations and processes.

Finally, Sowa has suggested other approaches towards roles in earlier works. [Guarino, 1992, p. 250] contains a quotation of Sowa from 1988, which distinguishes between natural types and role types. There, roles are "subtypes of natural types in some particular *pattern of relationships*" [italics added]. Hence, roles are not bound to a single relation and, contrary to what is quoted above, they are restricted to one type of the current *phenomenon* category.

2.1.4 Guarino's Meta-Classification of Properties

Formal ontology in computer science is a young research area which spans from philosophy to knowledge representation. One of its early ambassadors and current leaders, Nicola Guarino, has described it as a field "intended as a theory of *a priori* distinctions" along two lines (cf. [Guarino, 1995]). On the one hand, the entities of reality have to be classified (for instance, as physical objects, processes, regions, etc.). On the other hand, meta-level categories used in such world models (e.g., concepts, properties, roles, etc.) are to be distinguished.

Guarino has been looking for ontological and also linguistic concepts that may be helpful for structuring knowledge bases and for organising taxonomies. From the early days on, roles appear in his works on this topic in a fairly central position. During the last ten years his formal account of roles has remained quite stable, whereas the connotations of this account have changed slightly, as will be clear from what follows.

In [Guarino, 1992], he introduces two meta-properties of concepts[25] which are called *dependence* and *rigidity* in present works[26]. These notions are employed in a formal account of roles, whose latest definitions from [Guarino and Welty, 2001] are given below, differing only sparsely from the original formulations. Note that the quantified modal logic S5 supplemented by the Barcan formula is employed[27]. Further, the predicates P and C in Definition 2.1 denote a PART-OF and a constituency

[25]Note that the terms *concept* and *property* are used interchangeably in this section. Either is best understood as a UNIVERSAL in GOL.

[26]The term *dependence* has replaced that of *being founded* which was used in earlier papers.

[27]Due to space limitations, familiarity with modal logic and Kripke semantics is assumed. Note further, that due to

relationship, respectively. The variable t in Definition 2.2 refers to an explicit representation of time, i.e., $\phi(x,t)$ represents the fact that ϕ is a property of x at time t.

Definition 2.1 (Dependence [Definition 11, p. 59]) *A property ϕ is* externally dependent *on a property ψ if, for all its instances x, necessarily some instance of ψ must exist, which is not a part nor a constituent of x:*

$$\forall x \Box(\phi(x) \rightarrow \exists y\, \psi(y) \wedge \neg P(y,x) \wedge \neg C(y,x))$$

Definition 2.2 (Rigidity, Anti-rigidity [Definitions 1 and 3, p. 56])
A rigid property *is a property that is essential to* all *its instances, i.e.*

$$\Box(\forall xt\, \phi(x,t) \rightarrow \Box\forall t'\, \phi(x,t')).$$

An anti-rigid property is a property that is not essential to all *its instances, i.e.*

$$\Box(\forall xt\, \phi(x,t) \rightarrow \Diamond\exists t'\, \neg\phi(x,t')).$$

Definition 2.3 (Role) *A* role *is a property which is anti-rigid and dependent.*

The last definition is extracted from [Guarino and Welty, 2001]. One of Guarino's favourite examples to demonstrate the difference between rigid and anti-rigid properties, as well as between roles and *natural types* are the concepts `human` and `student`. The former does not depend on another concept and it is rigid, for which reason it is a natural type. Conversely, `students` (in a narrow, universitarian sense) depend on `professors` and `universities`, and a `student` is *not necessarily* a student, according to Guarino. This is not to be confused with a temporal interpretation, i.e., even if a student was a student for her/his whole life, this *being a student* is still not necessary. There may be a world where this object is not a student.

In [Guarino, 1992], roles are defined as externally dependent and *not rigid*, while the latter part is restricted to *anti-rigidity* in [Guarino and Welty, 2001]. In contrast to this rather small change in the formal account of roles, the intuitive description has developed more intensely. The starting point for the formalisation in [Guarino, 1992] is Sowa's suggestion that roles appear in "patterns of relationships". That means that relations are the basis to encounter roles, be they explicit as in `son` or implicit as in `pedestrian` [Guarino, 1992, p. 253]. The intuitive meaning of roles which is described in [Guarino and Welty, 2000] is as follows. Note that now an association with events appears more important than a link to relations.

> In general, roles are properties expressing the *part played* by one entity in an event, often exemplifying a particular relationship between two or more entities. All roles are anti-rigid and dependent (compare this with [Guarino, 1992]). [Guarino and Welty, 2000, p. 107, citation adapted]
>
> Intuitively, when a property is recognised to be a role, there should be some event that the role corresponds to. [ibid., p. 108]

this choice Guarino commits to a (philosophical) *possibilist* position, as contrasted to *actualism*. Basically, supporters of the latter contend that only actual entities exist. This is disputed by possibilists which agree that there are non-actual, but possible entities.

Moreover, [Guarino and Welty, 2000] contains a subclassification of roles. They are split into *formal* and *material* roles on the basis of their relation to identity. A property is said to carry identity if an identity criterion applies to all of its instances[28].

> In addition, *formal* roles do not carry identity, and intuitively represent the most generic roles that may form the top level of role hierarchies. [ibid., p. 107]
>
> Material roles are anti-rigid and dependent, but inherit identity conditions from some type. Material roles represent roles that are constrained to particular kinds of entities. [ibid., p. 108]

Interestingly, `patient` and `instrument` are given as examples of formal roles which is quite evocative of thematic roles (see section 2.3.2). Another such example, `being loved by John`, is used to show that the formal account given is not yet sufficient. This example is classified as a formal role, because it does not provide identity, in contrast to material roles. Thus identity seems to be an insufficient or even unsuitable criterion for the distinction between formal and material roles. Sample material roles are `student` and `food`. In this connection, the event `enroll` shall be seen as corresponding to `student`.

There are a number of other meta-categories introduced by Guarino and Welty, as well as suggestions for restrictions of the subsumption relation between them. Formal roles are proposed to be used for role taxonomies only. Material roles may be subsumed by roles and rigid properties, and may subsume only other material roles.

This section concludes with some remarks on recent works co-authored by Guarino, to wit the DOLCE ontology [Masolo et al., 2002]. Therein, a notion of roles is not explicitly mentioned, but the formal part introduces rigidity and several concepts of dependence which may be used for duplicating the account from above. On the other hand, there are categories which seem to subsume or to be related to role concepts, respectively. *Social Agent* and *Society* are outstandingly destined for this, not at least due to the examples given in [Masolo et al., 2002, p.10], like `(legal) person`, `contractant`, or `the Bank of Italy`. Either of these "leaf" categories[29] of DOLCE is subsumed by the following more general categories in the given order: *Agentive Social Object*, *Social Object*, *Non-physical Object*, *Non-physical Endurant*, *Endurant*, and, eventually, *Entity* which is the topmost node of the DOLCE taxonomy.

2.1.5 Topic Maps

Topic maps [Pepper and Moore, 2001] aim at providing (meta-)information about Internet resources. They are compatible with the Extensible Markup Language (XML) [Bray et al., 2000; Cowan, 2002]; indeed, topic map documents are to be serialised and exchanged using an XML syntax defined in the specification. A topic map captures knowledge about the subjects of resources and their interrelationships. The paradigm may be seen as being in line with the Semantic Web, which is taken as justification for including topic maps into the knowledge representation section. First introduced in 1993, the paradigm was acknowledged as an ISO Standard in 1999 [ISO, 1999].

[28] See [Guarino and Welty, 2001] for a formal account of identity, which is much more complex. Since it will not be used in the sequel, its specification is omitted.

[29] "Leaf" here refers to the tree-representation of the DOLCE categories.

A topic map consists of *topics*, *associations*, and *scopes*, the last not being relevant here. A topic represents any conceivable entity within a topic map system, while associations relate topics. A *role* describes the link between a topic and an association, or as defined in [Pepper and Moore, 2001], the concept refers to

> [t]he role that a topic plays as a member of an association; the nature of its involvement in that association.

With regard to this definition, roles in topic maps appear to be similar to UML *rolenames* (see section 2.2.2). However, associations in topic maps are specified free from order requirements, an approach hardly found anywhere else.

> There is no directionality inherent in an association. (Associations describe relationships: If A is related to B, then B must also be related to A. The issue is rather, what is the type of the relationship, and what roles are played by its members. The question of how to label a relationship is one of naming, not direction.) [ibid.]

With this characterisation of associations/relations, topic maps implement a remark in [Chen, 1976], which states that the order in tuples can be abandoned if roles are explicitly given (see also section 2.2.1). An example of an association shall illustrate the use of roles in the topic map formalism. Figure 2.3 comprises an association `written-by` between the topics `shakespeare` and `hamlet`. Abstracting away from the details of XML syntax, note the use of the XML-elements `<member>` in Figure 2.3. They represent the connection between a role, the corresponding association, and the topic which plays the role. The order of the `<member>` elements within the element `<association>` conveys no additional meaning, in contrast to the use of ordered tuples for representing associations, where explicit roles are neglected and the positions within a tuple are implicitly assigned to roles.

```
<association>
      <instanceOf><topicRef xlink:href="#written-by"/></instanceOf>
      <member>
            <roleSpec><topicRef xlink:href="#author"/></roleSpec>
            <topicRef xlink:href="#shakespeare"/>
      </member>
      <member>
            <roleSpec><topicRef xlink:href="#work"/></roleSpec>
            <topicRef xlink:href="#hamlet"/>
      </member>
</association>
```

Figure 2.3: Topic Maps: The `written-by` association consists of two roles, `author` and `work`. [Pepper and Moore, 2001]

2.2 Conceptual Modelling

This section deals with role concepts that belong rather to the field of software engineering than to knowledge representation. In particular, the focus is set on modelling aspects which occur mainly in the analysis and partially in the design stage of the software system's life cycle. The section is divided according to major paradigms. Although actually not being a modelling formalism, a data model specific to information systems in medicine is referred to in a separate section (section 2.2.5), because it comprises an unconventional and interesting notion of roles.

2.2.1 Data Modelling - Entity-Relationship (ER)-Modelling

Within data modelling, roles have been present since the late 1970s – [Bachman and Daya, 1977] is often cited as one of the first approaches with an explicit role concept (cf. [Steimann, 2000c, p. 84]). Even the standard entity relationship (ER) approach presented by Chen in 1976 already mentions a notion of roles. Note the interrelationship between the order of the arguments in a relationship and its roles:

> The role of an entity in a relationship is the function that it performs in the relationship. "Husband" and "wife" are roles. The ordering of entities in the definition of relationship [...] can be dropped if roles of entities in the relationship are explicitly stated as follows: $(r_1/e_1, r_2/e_2, \ldots, r_i/e_i)$, where r_i is the role of e_i in the relationship. [Chen, 1976, p. 12]

This remark on the substitutability of the order in tuples by means of specific roles seems to have long been ignored. Topic maps (see section 2.1.5) are the only exception that was identified to drop order requirements in favour of roles. Apart from this issue, Chen ascribes roles exclusively to one entity type, i.e., one UNIVERSAL.

David C. Hay presents data model patterns for ER modelling in [Hay, 1996], which match design patterns from the object-oriented paradigm (see below, p. 31). He encounters roles in two cases, namely as "work order roles" and as "contract roles" [Hay, 1996, p. 74 and 106, respectively]. Examples of the former are `project secretary` and `chief engineer`, of the latter `shipper` and `payee`. Prior to introducing these cases Hay models roles by means of relations. Later, however, when the need arises to add roles dynamically and to keep track of the history of the roles of an entity, roles are elevated to the realm of entities. There they exhibit the attributes `start` and `end date` and are related to a role type.

2.2.2 Object-Oriented Modelling

With the development of the programming language Simula-67, object-orientation started its way to become a widely used paradigm of programming and modelling, which it is nowadays. In contrast to knowledge representation, works that introduce notions of roles within this field have fairly technological motivations, often emerging directly from deficiencies of programming languages. Basic design choices in the development of object-oriented languages are closely related to roles. In particular, these choices include the questions of whether to allow multiple inheritance, multiple and dynamic instantiation, and/or polymorphism. While the acceptance of these concepts leads to a higher degree of expressiveness, several problems are caused especially by multiple inheritance and instantiation. Exemplarily, the question arises what to do in case a class inherits (or an objects instantiates) conflicting attributes or operations. A general resolution to this question seems hardly possible, because it is not clear in which way structural properties of classes could be exploited for these purposes. Computational tractability is also often alleged to as an argument against the acceptance of multiple inheritance and dynamic instantiation. On the other hand, since information systems are increasingly interlinked, there is the need to work with *evolving objects*. These may considerably change their attributes and behaviour which in turn complicates the assignment to one class only. Finally, it is not clear whether concepts like multiple instantiation are really sufficient to capture roles, even if such concepts are admitted. Many approaches refer to the lack of multiple, dynamic instantiation as their motivation for an introduction of roles (cf., for instance, [Fowler, 1997]), whereas papers like [Wieringa et al., 1994] elaborate a difference between roles and merely dynamically defined subclasses.

The following sections present a selection of object-oriented modelling formalisms containing role concepts. Emphasis is put on the current quasi-standard modelling language in object-orientation, the Unified Modeling Language (UML)[30]. Moreover, a recent comparison of roles within the object-oriented paradigm offers a considerable collection of requirements of what role modelling in object-orientation should allow for. The section concludes with some works that appeared after this comparison in order to include works not yet evaluated, as well as with some remarks on role-like approaches in the design patterns area.

2.2.2.1 Unified Modeling Language (UML)

The introduction of UML in 1997 can be considered a major step towards a standard way of modelling for object-orientation during different stages of the software development process. It is also strongly supported by and used in industry. Recently, voices appeared which call for a truly integrated UML and denounce UML as a loosely coupled selection of different modelling formalisms (cf. [Jähnichen and Herrmann, 2002]). Herein, we refer to version 1.4 of UML ([OMG, 2001], cf. also [Rumbaugh et al., 1999; Booch et al., 1999]), which provides the basis for the analysis. By detecting more roles than expected beforehand, we agree to the mentioned assessment of Jähnichen and Herrmann.

There are several types of UML diagrams encompassing different role concepts, some of which can be found as model elements whose names use the term “role”. Class diagrams, belonging to the static part of UML, contain *rolenames* and *qualifiers* in connection with associations. In the behavioural part, one finds *classifier roles* as a central model element of collaboration diagrams, as well as *association roles*. These diagrams are used to model the participation of several objects in a certain context. The same motivation in a broader sense refers to a distinction between types and roles in [Booch et al., 1999] and a connection to *interfaces*. Furthermore, *actors* in use case diagrams seem to mimic special roles. Let us now look at all of these conceptions in greater detail, since in chapter 4 they will be compared to the approach proposed in chapter 3.

Rolenames are linked to associations in UML, which themselves roughly correspond to RELATIONS (see [Guizzardi et al., 2002b]). The UML specification introduces rolenames as the name attribute of an AssociationEnd [OMG, 2001, p. 2-23]. This is commensurate to the following definition from the reference manual:

> *rolename* – A name for a particular association end within an association. [Rumbaugh et al., 1999, p. 414]

No restrictions on associations are given, either concerning their arity or any other restrictions[31], but still there are some conditions that influence the semantics of rolenames:

[30] We assume familiarity with the basic object-oriented notions of *objects*, *classes*, *attributes*, *operations*, and *inheritance*, as well as with UML in general and with those UML model elements which correspond to these basic notions in particular. Note that, ontologically, there are a lot of open problems with fundamental object-oriented concepts. [Guizzardi et al., 2002a] deals with some of them, but others remain open.

[31] Supposedly with regard to programming languages, several conditions according to objects have to be fulfilled, which are left out here if they clearly refer to implementational aspects, as in the case of *navigability*.

> All the rolenames in an association must be different. Within a self-association (an association involving the same class more than once), rolenames are necessary to disambiguate the ends attached to the same class. Otherwise, rolenames are optional, because the class names can be used to disambiguate the ends. [ibid.]

Another notion whose illustrations in UML documents suggest overlap with some role concept is that of *qualifiers*. This language element is defined as

> A slot for an attribute or list of attributes on a binary association, in which the values of the attributes select a unique related object or a set of related objects from the entire set of objects related to an object by the association. It is an index on the traversal of an association. [Rumbaugh et al., 1999, p. 398]

Two examples from [Rumbaugh et al., 1999, p. 326 and p. 399] may elucidate this definition. In the first example, an `officer-of` association between a `Club` class and a `Person` class is qualified (within an object diagram). At the association ends attached to class `Club`, two qualifier values appear: `treasurer` and `president`. At the opposite end, i.e., attached to the class `Person`, `officer` is given as a *rolename*. This example strongly suggests that roles (in an ontological sense) may appear as qualifiers in UML diagrams, and not only as rolenames. The second example qualifies an association between the classes `Bank` and `Person` with a qualifier `account #` at the `Bank` end. This seems to be a different use of qualifiers, since account numbers are not a prototypical role for persons.

In contrast to rolenames and qualifiers, *classifier roles* do not refer to relations. Rather, they may be understood as placeholders within the context described by a collaboration diagram. Collaborations may be considered as kindred to SITUOID or PROCESS UNIVERSALS in GOL. They describe a collection of objects together with the relevant relationships in this context, as well as interactions which occur in order to implement some behaviour. Thus collaborations usually refer to a *use case* or an *operation*, and can in this way be understood as contexts. If the participants within these contexts are named according to it, these names are classifier roles:

> A slot in a collaboration that describes the role played by a participant in a collaboration. [Rumbaugh et al., 1999, p. 194]

For example, a collaboration may refer to an `order` use case, i.e., the collaboration describes the internal connections between the participants of the use case and the actual procedure of an `order`. It may involve the classifier roles `client` and `contractor`, the instances of which are recruited from different, context-independent classes like `Person` and `Company`, respectively.

Of course, classifier roles are subject to several restrictions (cf. [Rumbaugh et al., 1999, p. 194] and [OMG, 2001, p. 2-115–2-122]). First of all, a classifier role always refers to other classifiers outside the collaboration, on which it defines a view by referring to those features only that are necessary in the context of the collaboration. These other classifiers are called *base classifiers*[32], for which the classes `Person` and `Customer` are examples given in the illustration referring to `orders` just above. It is possible to give multiplicity constraints stating how many objects of a base classifier may occur within

[32] Note that most other descriptions, including [Rumbaugh et al., 1999] and [Booch et al., 1999], refer to exactly *one* base classifier per classifier role, contrary to the specification where *several* base classifiers are accepted.

one instantiation of the collaboration. Hence, a classifier role refers to a set of objects in collaboration instances. Each classifier role is distinct from others in its unique context and only defined for that context. Consequently, classifier roles in two distinct collaborations do not literally represent the same role, although those roles may select the same features of the same base classifiers. Conversely, one base classifier can occur in several classifier roles, and different objects may fill classifier roles of distinct collaboration instances.

As classifier roles correspond to classes, there are also *association roles* which represent associations that are valid in the context specified by the collaboration diagram. If no base association is specified, they are valid in that context only. Taking up the above illustration again, a base association `customer-of` between the classes `Person` and `Company` may play an `orders-from` association role between the classifier roles `client` and `contractor`.

As mentioned above, collaborations can be used to model the implementation of *use cases*. The latter UML element is related to another element called an *actor*. Actors have some properties in common with roles, but other properties controvert an identification with roles. Conforming to [Rumbaugh et al., 1999, p. 144], an actor participates in one or several use cases and can be implemented by several physical objects. Thus one physical object itself may implement several actors. So far, these properties are compatible with roles, and examples like `salesclerk` and `customer` do not appear unorthodox in this respect. Nor is it a problem, that actors are restricted to abstractions of entities *outside* the system. However, "[a]n actor is an idealization with a focussed purpose and meaning and might not correspond exactly to physical objects," and each actor is said to define "a set of roles that users may assume when interacting with the system" [Rumbaugh et al., 1999, p. 144].

Finally, there is another distinction in UML, to wit between *interfaces*, *types*, and *roles*, which is given in the User Guide [Booch et al., 1999, p. 155–167]. There, these terms are described as follows.

> An *interface* is a collection of operations that are used to specify a service of a class or a component. A *type* is a stereotype of a class used to specify a domain of objects, [...]. A *role* is the behavior of an entity participating in a particular context. [Booch et al., 1999, p. 157]

The authors of the User Guide see the difference between types and interfaces in that types may have attributes. Further, roles are interfaces which appear in particular contexts. If dynamic types are to be modelled, one is advised to model types in a class diagram and to connect those to classes via generalisation.

In summary, UML contains a number of elements referring to roles which are not semantically equivalent from the UML point of view. There are rolenames and qualifiers, classifier and association roles, actors, as well as types and interfaces. Chapter 4 reconceives these UML notions of roles and reveals their relation to our concept of roles.

2.2.2.2 Steimann's Role Comparison

In 2000, Friedrich Steimann presented a comprehensive comparison of role models that had been proposed for object-oriented as well as data modelling [Steimann, 2000a,b,c]. Since one of his applications is a proposal for adaptations of UML (cf. [Steimann, 2000b,d]), his works are introduced within this section on object-orientation. Steimann has identified a number of properties of roles, reproduced in the list below, which are based on various approaches. These properties are partially contradictory, because they represent a mere collection from the literature instead of being a unified theory. However, this collection is consulted by Steimann in developing his own approach. In short, he states the following criteria in [Steimann, 2000c] (supplementary notes are omitted here; cf. section 4.2.4).

1. A role comes with its own properties and behaviour.
2. Roles depend on relationships.
3. An object may play different roles simultaneously.
4. An object may play the same role several times, simultaneously.
5. An object may acquire and abandon roles dynamically.
6. The sequence in which roles may be acquired and relinquished can be subject to restrictions.
7. Objects of unrelated types can play the same role.
8. Roles can play roles.
9. A role can be transferred from one object to another.
10. The state of an object can be role-specific.
11. Features of an object can be role-specific.
12. Roles restrict access.
13. Different roles may share structure and behaviour.
14. An object and its roles share identity.
15. An object and its roles have different identities.

Steimann goes on to present three common ways of representing roles, namely as *named places* of relationships, as *specialisation/generalisation*, and finally as *adjunct instances*. The first approach corresponds to UML rolenames. The main disadvantage of named places is, according to Steimann, that role-specific properties and behaviour cannot be expressed.

The representation of roles by specialisation and/or generalisation results in well-known issues like multiple inheritance and instantiation, as well as dynamic instantiation. However, Steimann does not agree that this models roles appropriately. He exposes the “paradox” that roles appear as supertypes from a static point of view, while from a dynamic point of view they are subtypes [Steimann, 2000c, p. 90]. This is demonstrated by the following example. He considers two role types, `Supplier` and `Customer`, and the natural type `Party` (in the sense of [Guarino, 1992], see also section 2.1.4), which comprises itself of the subtypes `Person` and `Organisation`. Steimann explains that, from a static point of view, every person and every organisation *can* be a customer and this argumentation is sometimes employed to justify, e.g., `Customer` as a superclass of `Person` and `Organisation`. Conversely, each customer is also a party (i.e., a person or an organisation) at any moment in time.

This in turn allows for the idea that `Customer` is a subclass of `Party`. Note that this argumentation cannot be supported within the GOL framework, where the taxonomic relationship is clearly defined in terms of instantiation: a UNIVERSAL A is then a subuniversal of B if and only if being an instance of A implies being an instance of B. Clearly, being an instance of `Person` does not imply being an instance of `Customer`. Therefore, we reject Steimann's argumentation on this issue, although we agree with the initial statement itself, i.e., the specialisation relationship is inappropriate to account for roles. Sections 2.3.1 and 3.5, respectively, present the notion of ontological levels, which forms the basis of our argument concerning this issue.

Solution three, the adjunct instances approach, is considered "unorthodox at least" [Steimann, 2000c, p. 94], because a real-world object would then be represented by several information objects some of which do not carry identity (*object-slicing*). In particular, each adjunct role is bound to the object which plays this role by means of a certain relation (which is often termed *plays* or something similar).

Steimann does not follow any of these three solutions. Instead, he considers separate type and role hierarchies and elevates the *plays* relationship to the type level, i.e., *plays* relates natural types and role types (instead of instances). Thus Steimann uses a second-order relationship. Further, roles are used to define relationships. In particular, each role occurs only within one relationship. Then it becomes possible to solve the problem of roles being static supertypes and dynamic subtypes, because the intension of a role type is specified twofold, consisting of an absolute part and a relative part. In addition, the intension is inherited twofold: the absolute part is inherited to a type standing in the *plays* relation to a role, while the relative part is only "inherited" to an instance of the type if this instance participates in the relation to which the role belongs. On the whole, Steimann creates association-dependent classes with an asymmetric dependence. For instances of those classes considered natural types the *plays* relationship is optional, while it is obligatory for role classes.

2.2.2.3 Works After Steimann's Comparison

[Dahchour et al., 2002] concentrate on dynamic change of role classes, multiple instantiation of the same class, and a role-specific access to objects. The latter two points cannot be realised by (even multiple) inheritance, because multiple instantiation of the same class is not related to inheritance issues, and object access according to specific superclasses is not provided by present inheritance mechanisms (cf. also [Gottlob et al., 1996]). In addition, Dahchour et al. consider different types of combinations of roles, e.g., *evolution* (gaining a new role while losing an old one) or *extension* (gaining a new role and keeping others). A distinction of class-level and instance-level requirements for roles is also new, although the requirements themselves match a subset of Steimann's. It is worth mentioning that roles in this approach have to be assigned to exactly one class.

A connection between *roles* and *purpose* can be found in [Fan et al., 2001]. The authors consider roles as "things that are, but only in the context of things that happen", as an intuitive description[33]. Roles are represented as types independent of *entities*, which apparently correspond to Guarino's natural types. The role hierarchy which is indicated contains subtypes close to frequent thematic roles (see section 2.3.2). The particularity of the approach is featured by capturing a notion of purpose as a

[33] More formally, Fan et al. refer to Guarino's early work [Guarino, 1992], although their emphasis on a reference of roles to *events* (i.e., PROCESSES) appears to be incompatible with Guarino's notion of roles based on *relations*, which are explicitly contrasted to events.

role intended to be played by certain entities. An example of a `shoe` playing the role of a `hammer` in a hammering event can then be expressed as follows ([Fan et al., 2001, p. 42]; $plays(x,y)$ states that entity x plays the role y in some event; misleadingly, the authors use the predicate *isa* to refer to instantiation).

$$\exists p,h\ isa(myShoe,Shoe) \wedge isa(h,Hammer) \wedge plays(myShoe,p) \wedge purpose(h,p)$$

Coulondre and Libourel [2002] pursue a completely different approach. They consider roles to be intermediate concepts between classes and single properties (attributes and methods). On this basis, classes are defined as a set of roles, whereas these roles are structured in an is-a hierarchy. In addition, classes themselves are organised in their own is-a hierarchy. A single role is defined by "a *criterion*, a type and a set of method signatures." [ibid.]. Unfortunately, the full intent of the criterion component of a role remains unclear, and therefore the actual notion of a role, too[34]. On the basis of their model of roles, Coulondre and Libourel suggest an extension of UML. The primary fact though which should be noted is the unique approach of considering classes, which may be understood as UNIVERSALS in the object-oriented paradigm, as being constructed from roles.

2.2.2.4 Design Patterns

During the last decade, *design patterns* have become popular among a community which is guided by implementational concerns rather than theories about modelling. Feeling a need for reuse and discovering similarities, people started to collect their solutions to software design problems in more abstract ways. [Gamma et al., 1994] represents a standard work of such collections, which were also introduced in data modelling (cf. section 2.2.1).

One may discover notions or at least properties of some design patterns which refer to roles. The *Decorator Pattern* [Gamma et al., 1994, p. 175] serves as an example, especially as [Costanza, 2000] highlights the connection of this pattern to roles. Explicit use of roles in design patterns has not been found very often. Two examples are the *role–actor–pattern* which Oestereich refers to [Oestereich, 1998] and the *Role Object pattern* [Bäumer et al., 2000], the latter apparently being closely related to the former.

> The Role Object pattern models context-specific views of an object as separate **role objects** which are dynamically attached to and removed from the **core object**. The resulting object aggregate represents one logical object, even though it consists of several physically distinct objects. [Bäumer et al., 2000, p. 15]

Accordingly, Bäumer et al. present a way of modelling roles independently from the objects playing them. Apart from all implementation-specific issues, they obviously consider an object with its roles as a logical unit, i.e., there is one instance of an object and several role classes. Again, there is a reference to the notion of *context*. Within their pattern, roles are connected to the core object via aggregation, such that one may draw parallels to [Coulondre and Libourel, 2002]. Hence, it implements what Steimann called the *adjunct instances* approach.

[34] A criterion is said to be "a restricted form of first-order logic (FOL) formulae *freely chosen* by the user, in order to reflect the desired semantics." [Coulondre and Libourel, 2002] – however, what is meant by "restricted" is not explicitly stated.

The Role Object pattern is also one of five variants of implementing the "trivial concept" of roles, as collected in [Fowler, 1997]. Mainly, Fowler suggests using that modelling approach among these five which suits best the application needs, but actually the question of how to represent roles would be solved if multiple, dynamic classification was at hand. Hence, he considers roles as UNIVERSALS which can be specialised and generalised like normal types.

2.2.3 Agent-Oriented Modelling

For about ten years, a version of distributed artificial intelligence has been emerging whose central notion is that of an "active object" or an object acting autonomously – the so-called *agent*. Frequently, several agents – in a multi-agent system – interact in order to produce some behaviour of the system. Despite the fact that up to now there has not been a fixed notion of an agent, there are a number of works on modelling as well as on methodologies for systems based on this philosophy.

[Iglesias et al., 1999] contains very short introductions to several agent-oriented methodologies. Within some of these, it is an important step to identify the roles of an agent in the application domain, more precisely in certain *scenarios* or *processes*, depending on the terminology of the specific formalism. [Depke et al., 2001] also reviews several notions of roles for agents. These span from use in protocol definitions over capturing goals, tasks, or functions to descriptions of organisational structures, with or without certain behaviour. The authors see protocols and social relationships as the main characteristics of roles. They propose a use of roles closely related to their work on the integration of roles into UML (cf. [Depke et al., 2000]; therein, a *role-of* relationship between classes is introduced, which is then discussed for several UML diagram types).

In order to get a better impression of the use of roles in agent frameworks, let us look at an example with structured role descriptions. In the framework of the *Gaia methodology* [Wooldridge et al., 2000], roles are analysed from an organisational point of view, i.e., societies and organisations are considered to consist of roles. Some of the criteria specified by Steimann are named, for instance that an entity may take on several roles. It should be mentioned that roles are not necessarily supposed to be played by persons, but also by departments within an organisation or organisations themselves (cf. [Wooldridge et al., 2000, p. 295]). Further, an interesting feature of this work is the structure of the descriptions of roles. Four attributes are used to capture the meaning of a particular role: *responsibilities*, *permissions*, *activities*, and *protocols*. Accordingly, responsibilities refer to the functionality of a role, hence roles always comprise a function. Permissions comprise the rights to use certain resources in order to fulfill the responsibilities. Finally, activities and protocols describe the actions an agent can exert in that role. They are distinguished by the criterion whether or not other agents are involved, where activities are agent-internal actions. Further details can be dispensed with, since the main idea about roles has been outlined – a lengthy introduction of the syntax can thus be avoided.

2.2.4 Systems Modelling - Object-Process Methodology (OPM)

Dov Dori attempts to overcome what he believes to be the predominance of objects in the object-oriented paradigm. In [Dori, 1995] he introduces a modelling formalism called *object-process modelling*, where objects and processes are on an equal footing. This approach has been refined and tested (cf. [Peleg and Dori, 1998, 2000]) and is comprehensively presented in a recent book[35] [Dori, 2002]. His approach comprises two representational formalisms, namely a diagrammatic language whose models are called *object-process diagrams* (*OPD*) and a textual language, *object-process language* (*OPL*). Each model using OPM may be expressed in either way.

Roles are not an explicit element in Dori's modelling approach. However, one finds a notion of *object roles* with respect to processes and to states, which will also be discussed in chapter 4. First of all, let us set OPM *objects*, *processes* and *states* in relation to GOL terms. Objects can be understood as ENDURANTS[36], whereas processes correspond to PROCESSES. States are different in that they resemble FACTS or even SITUOIDS, if one takes into account that compound states can be modelled [p. 332]. On this basis we can introduce OPM model elements the background of which refers to roles.

Procedural links [p. 93] "glue" objects or their states to processes. They are divided according to whether an object is transformed by the process or whether it is necessary to enable the process. In the former case, such an object is called a *Transformee*, in the latter it is an *Enabler*. Transformees may further be classified into *consumed*, *resulting* or *affected objects* in accordance with the relevant process. This distinction is aligned with the question of whether the object exists before, after or during the process. On the other hand, enablers are divided into *agents* and *instruments*, the definitions of which are:

> **Agent** is an intelligent enabler, which can control the process it enables by exercising common sense or goal-oriented considerations.
> **Instrument** is a non-human physical or informatical enabler. [p. 94–95]

Note that each of the above distinctions is exclusive. Hence, an object cannot be an agent and an affected object with respect to one particular process (cf. an example of a `student` which is a transformee rather than an agent if considered in the process of `studying`). Considering distinct processes, an object may take on different roles, i.e., it might be an enabler of one process and a transformee of another. So far one may allude to similarities between this classification and theta roles in subsection 2.3.2.

Cash exhibits **Owner**.
Bank and **Customer** play the role of **Owners** for **Cash**.

Figure 2.4: An OPM Example Containing a Role-Playing Sentence [Dori, 2002, p. 315].

In addition, two more issues of OPM are relevant with regard to roles. First, OPL contains *role-playing sentences*. These are used to point out that an entity appears in a model as an object as well as a state. Thus an entity which is an OPM object plays a role as an OPM state in another OPM object. This corresponds to an implicit representation of a relationship between two such entities. For instance, consider the OPL sentences in Figure 2.4. Clearly, the state[37] `Owner` is used for representing an

[35] Page references in this section refer to this book.

[36] However, OPM objects persist through time, which suggests an identification with ABSTRACT ENDURANTS. Nevertheless, those features predominate which support an understanding as ENDURANTS, like the exhibition of attributes, i.e., MOMENTS.

[37] The term "exhibits" indicates that `Owner` is modelled as a state.

`Owner` relationship between `Cash` and either `Customer` or `Bank`. In addition to role-playing sentences, OPM provides the notion of *structural relations* [Dori, 2002, section 6.1]. They correspond to GOL RELATIONS and to UML associations [p. 105]. Dori restricts his discussion to binary relations[38]. Relations are directed, but there is always a pair of a *forward* and a *backward* structural relation. If the forward and the backward structural relation between two objects are identical, they can be replaced by one *reciprocal* structural relation. Finally, there is no modelling element in OPM which would correspond to rolenames in UML.

2.2.5 HL7 - A Standard for the Clinical Domain

HL7, an organisation developing standards, methodologies, etc. "for interoperability of health care informations systems" [HL7, 2002a], also employ a notion of roles in their Reference Information Model (RIM, [HL7, 2002b]) which is an object-oriented model. It is worth-while to introduce this notion, because it provides a new aspect which appears in no other formalism.

The top-level classes of RIM are *Entity*, *Role*, *Participation*, *Act*, *RoleLink* as well as *ActRelationship*. Interestingly, this refers to a certain "top-level ontology" which is, of course, adapted to the domain of health care information systems, but which can still be compared to the top-level categories of GOL. Let us briefly describe these classes. *Entity* comprises physical things, organisations or places. The class *Act* refers to intentional actions in the HL7 business domain, be they clinical or administrative in nature. Note that Acts refer to both the execution of such an action as well as the resulting data that have to be stored. *Roles* and *Participations* mediate between Entities and Acts. In this connection, Roles represent qualifications or "categorizations of competency". Entities in those Roles can still participate in an Act in numerous ways, and therefore Participations describe the special way of how an Entity partakes in an Act. RoleLink states dependencies between roles, e.g., `has-authority`. ActRelationships are similar to RoleLinks, for they convey connections between Acts.

The description of Roles also contains some properties. Accordingly, an Entity can *play* several Roles at a time, i.e., *play* is an association between class Entity and class Role. Another association is *scope*, which restricts the scope of a Role as defined by another Entity which is not the player. For example, a `patient` is a Role which is usually played by a `person` and may be scoped by a `health care organisation` (the one where the `patient` is treated). Two attributes of the Role class are particularly interesting in our ontological context. First, there is the attribute *id* which is a unique identifier for a Role, actually identifying the Role-playing Entity. For instance, a `patient` may be identified via a `patient number` issued by the scoper of the Role, i.e., by an Entity which stands in the scope relation to this Role. Thus, the id attribute mimics UML qualifiers. Secondly, the attribute *certificateText* refers to a certificate that is issued by the scoper of the Role to the Role player. In the `patient` example, this may be an `attestation`.

Some of those requirements on Roles can also be found in other approaches, while the scope relation and the necessary certification of a Role are new. The connection of Entities and Acts via Roles *and* Participation is also uncommon. HL7 points out that the assumption was made that any Entity participating in a health care context takes on a particular Role.

[38] Though *n*-ary relations may indeed be reduced to *n* binary relations between each particular argument and the set of the remaining arguments, it is more convenient to be able to model and analyse *n*-ary relations directly. Chapter 3.3 contains further discussion on this issue as well as our way of handling relations in general, which even allows for a simple way of expressing anadic relations.

2.3 Roots of Roles in other Disciplines

In order to explore roles beyond the possibly limited scope in computer science, three different areas are briefly consulted. Some works are chosen due to being cited in some computer science papers. Further, the next chapter will implement some interesting ideas from the following sections, when we approach a classification and characterisation of roles.

2.3.1 Philosophy

Philosophy, as the field where formal ontology itself has its roots, offers sundry material. We focus on two authors, dealing with *ontological levels*, *relations* and with *social acts*. Roles themselves do not appear in the introduction of philosophical literature in this section.

2.3.1.1 Level Ontology

Johansson [Johansson, 1989, particularly ch. 2 and 14] describes an idea of *ontological levels*, which seems to have appeared quite early (e.g., in Aristotle's thinking), but which has been followed only by few modern philosophers (among others, Johansson mentions Ingarden and Bunge[39]). The main idea is that there are different levels within an ontology, the higher levels being existentially dependent on the lower, while the opposite does not hold[40]. These levels should not be confused with levels within a taxonomy. Rather, each ontological level contains its own taxonomy. Therefore, each of these levels is in a sense closed, i.e., things may be explained quite well within one level without the need to go to other levels. As an example, Johansson mentions a `refrigerator`. It is possible to explain its function without having to know the particular `cooling substance` in its atomic structure. These levels should likewise not be confused with pure part-whole granularity. For example, a `teacher` is also considered to be on a different level than a `human`. As higher levels depend on the lower levels, a certain `teacher` depends on a particular `human`. Johansson states that the entities of the different levels are coincident in space [Johansson, 1989, p. 135]. Thus the `teacher` occupies the same space as the `human` that forms its foundation.

2.3.1.2 Searle: Speech Acts and Social Reality

Searle has developed an account of social entities starting in the field of speech act theory. His approach is presented and discussed in [Smith, 2003] (cf. also [Searle, 1969, 1995]). The main contribution to be treated here is the *X counts as Y (in context C)* formula which describes so-called *constitutive rules*. Constitutive rules bring new forms of behaviour into existence, in contrast to *regulative rules* which merely regulate kinds of already existing behaviour. As an example for the latter, `polite table behaviour` is given, which regulates `eating`. `Signaling to turn left` and `playing chess` are examples of activities created by constitutive rules, i.e., certain artificial physical/material processes are defined as playing chess, for instance. The entity created by

[39]However, Bunge appears to be influential for computer science, because his works [Bunge, 1977, 1979] have a set-theoretic basis and are frequently cited by computer scientists. [Wand et al., 1999] may serve as an example.

[40]Johansson devotes a whole chapter, to wit ch. 9, to the concept of *existential dependence* and its further classification. We cannot account for a deeper introduction here, apart from the following definition [Johansson, 1989, p. 128]: "D9.1 A is existentially dependent on B if and only if it is logically impossible for A to exist, if B does not exist."

a constitutive rule is called a *Y term*, whereas the pre-existing entity is an *X term*. The *X counts as Y* formula can also be applied iteratively; for instance, a `human` may count as a `citizen` which in turn may count as a `judge`. Besides, Searle considers social entities to be physically identical to the underlying material entities.

There are also *free-standing Y terms*, for example a `company`. Free-standing Y terms refer to the idea that the existence of a Y term is acknowledged although no X term can be determined. Searching for an X term in the example, `signing a document` may *count as* the `founding of a company`, but then the company still exists when the founding process is completed. Consequently, the signing process cannot count as the company itself. This problem leads to a distinction between Y terms which exist *through* the corresponding X terms, and Y terms that are merely *represented* by some physical entity. In order to explain these, it is suggested that social reality emerges on the basis of facts rather than objects, since not every Y term refers to a physical object.

From our point of view, the conceptions of Johansson and Searle can be united in a complementary manner. Johansson has introduced different ontological levels, and Searle's approach provides a means of explaining the transition between two levels: a *counts-as* relation to reach social objects from physical and/or biological objects.

2.3.1.3 Johansson: Internal Relations

In the last part of the philosophical section, let us look at [Johansson, 1989, ch. 8], where Johansson distinguishes between *internal* and *non-internal* relations, which will be compared to Guarino's role definition on the basis of *rigidity* and *dependence* (cf. section 2.1.4).

In regard to Johansson, internal relations can be found between *A* and *B* if and only if *A* cannot exist unless *B* exists and vice versa [p. 117]. This resembles the dependence part of Guarino's definition of roles, and is a kind of existential dependence also in the sense of Johansson (cf. section 2.3.1.1). However, Johansson's examples are particularly interesting, which will be clear from what follows. Among these examples for internal relations are the pairs `teacher`–`pupil` and `priest`–`congregation`. Intuitively, the first example can clearly be identified as a pair of roles, which conforms with Guarino's definition since teachers and pupils depend on each other and both are anti-rigid "properties". The second example differs from the first one in that a `congregation` does not seem to be a role, whereas its members might still be recognised as such. Again, Guarino's criteria are sufficient to select the intuitively recognised role, because a congregation is a rigid concept and therefore not a role. But Johansson offers two other cases of internal relations, which may be problematic for the account of Guarino: internal relations between properties (i.e., MOMENTS, in contrast with Guarino's broader interpretation of this term as UNIVERSAL), and those between a substance and a property. As examples of the former, `length`–`breadth` and `colour`–`shape` are given. Intuitively, none of these terms seems to be a role. This case is also handled by Guarino's approach, because the rigidity criterion rules properties (i.e., MOMENTS) out, since any property is necessarily that property. The third type of internal relations, which relates substances and properties, becomes tricky, however. An example is the pair `commodity`–`price`. This issue refers to the philosophical idea of *essential properties* of a substance, i.e., by definition such properties need to be given for the substance to exist. The `price` is essential for any `commodity`. Unfortunately, essential properties are not excluded in Guarino's definition of dependence, where only *parts* and *constituents*[41] of the instances of the universal in ques-

[41] Intuitively, constituents refer to the material a substance is made of. Taken from [Guarino and Welty, 2001], `bricks`

tion are excluded from being a role. Now, if a universal depends on an essential property, anti-rigidity of the universal suffices to identify this as a role. According to this observation and Guarino's meta-classification (e.g., in [Guarino and Welty, 2001]), which allows for anti-rigidity of the categories of *phased sortals* and *attributions*, some of these may consequently be classified against our intuitions. Put differently, the statement that anti-rigid attributions and phased sortals are in every case not dependent becomes questionable in the light of essential properties.

2.3.2 Linguistics: Thematic Roles

Linguistics provides a notion which has served as a basis for role models in computer science. *Thematic roles*, which are also called thematic relations, theta roles/relations, semantic relations, or in early works "deep cases"[42], were introduced in the late 1960s (cf. [Fillmore, 1968; Jackendorff, 1991; Dowty, 1989]). [Parsons, 1990] provides a recent theory of how to explain the use of verbs in terms of events (PROCESSES), considering the English language as an example. In the course of this examination he discusses theta roles and presents a notable number of examples and open problems, some of which will be mentioned below. With regard to role approaches presented thus far, Steimann [Steimann, 2000c] and Sowa [Sowa, 2000] explicitly refer to thematic relations, while Dori [Dori, 2002] seems to have closely related concepts without explicit reference.

According to [Parsons, 1990], thematic relations are based on the idea that a finite set of relations holds between verbs and their argument positions. In addition, the syntactically defined notions of *subject* as well as *direct* and *indirect object* provide clues which restrict the possible relation between a verb and its arguments. In Parson's own theory, theta roles are also understood as relations between events and their participants. In contrast, Barker and Dowty [1992] define thematic relations more broadly, using them to explain lexical and semantic patterns in the behaviour of certain types of verbal predicates[43]. The arguments of a predicate fill those theta roles which are associated with it. Exemplarily, the verb `to walk` could have one obligatory theta role, the filler of which is the walking entity, and some optional theta roles for locations, such as the final destination of the walking. In the sentence "`John walks home.`", `John` is the filler of the obligatory role and `home` refers to the goal, the neglection of which still yields a meaningful sentence (cf. "`John walks.`").

Frequently, authors propose small collections of theta roles which are supposed to apply to a large number of verbs. For illustration, consider the following collection of thematic roles from different authors.

Thematic roles in [Sowa, 2000, p. 508–510]:

Agent An active animate entity that voluntarily initiates an action.

Beneficiary A recipient that derives a benefit from the successful completion of the event.

Completion A goal of a temporal process.

may constitute a `castle` or a `statue` may be constituted by `marble`. It is not clear to what extent constitution may resemble the level ontology approach.

[42]One should be aware of the fact that these terms are sometimes used synonymously in linguistics, while sometimes they bear slightly different connotations. Herein, synonymy is assumed.

[43]Predicate is to be understood in a syntactic meaning, ranging from verbs across adjectives and prepositions to nouns.

Destination A goal of a spatial process.

Duration A resource of a temporal process.

The remaining ones are *Effector*, *Experiencer*, *Instrument*, *Location*, *Matter*, *Medium*, *Origin*, *Path*, *Patient*, *PointInTime*, *Recipient*, *Result*, *Start*, and *Theme*. Some other thematic roles, taken from [DeArmond, 2002], are *Possessor*, *Trigger*, *Reason*, and *Purpose*.

It would be wrong to claim that there is a standardised set of thematic relations in the field of linguistics (cf. [Parsons, 1990, p. 73]), which is also demonstrated by the differences in the above examples. On the contrary, some authors support the opinion that each verb has its individual theta roles, which yields a vast number of thematic relations. There is not even an agreement on the issue to which word types theta roles should be assigned. While Sowa restricts this assignment to verbs, Barker and Dowty extend it even to nouns. There are further problems with respect to argument structure which are well-known in linguistics, for instance explanations of nominalisations and static verbs (like `to contain`, `to know`, etc.). Altogether, it is an open question, to what extent thematic relations are motivated by linguistic or ontological forces.

Nevertheless, discussions of theories employing theta roles are interesting for the present study on ontology, since many of the same problems are addressed. In addition, linguistic tests have often been used as a means to distinguish between (meta-)categories, e.g., by Guarino [Guarino, 1992] to determine whether a property is an attribute, as well as by Sowa (recall the *has test* from section 2.1.3).

Two approaches shall now be presented in order to illustrate effects and applications of theta roles. Let us start with [Dowty, 1991] and [Barker and Dowty, 1992], respectively. The former paper deals with a theory of thematic relations referring to verbs, while the latter extends them towards certain noun types. Barker and Dowty focus mainly on argument selection, i.e., for verbs, this is the problem of explaining which arguments take the subject, the direct object, and other positions. Therefore, when dealing with nouns, they restrict the discussion to relational, ultra-nominal nouns, i.e., nouns which are relational in nature but not derived from any verb. `Mother` is an appropriate example. They focus on two roles for verbs and two for nouns only, which are called proto-roles (analogous to the linguistic notion of *prototype*). For verbs these are *Proto-Agent* and *Proto-Patient*, while they are called *Proto-Part* and *Proto-Whole* for nouns. The following entailments for *Proto-Agent* and *Proto-Patient* from [Barker and Dowty, 1992] are similar to features of procedural links in Dori's approach (cf. section 2.2.4):

1. Proto-Agent entailments:
   (a) volitional involvement in the event (or state)
   (b) sentience and/or perception
   (c) causing an event or change of state in another participant
   (d) movement (relative to the position of another participant)
2. Proto-Patient entailments:
   (a) undergoes change of state
   (b) incremental theme
   (c) causally affected by another participant
   (d) stationary relative to movement of another participant

From an ontological point of view, these consequences for assigning a theta role to an argument still need to be expressed more clearly (most preferably, in a formalised manner). However, we will not provide an ontological account of a particular set of theta roles. Instead, section 3.4 will discuss a more general view of the ontological nature of theta roles.

Finally, a sample application of thematic relations can be found in [Parsons, 1990]. Parsons employs thematic relations for a limited kind of reasoning in natural language. With a stock of six theta roles (*Agent*, *Theme*, *Goal*, *Benefactive*, *Instrument/Performer*, and *Experiencer*), he formulates logical axioms which are directly derived from natural language sentences and shall allow one to draw inferences like the following, which exploits a connection between *into* and *in*.

> If Sam hits the 8-ball into the pocket, then the 8-ball ends up in the pocket. [Parsons, 1990, p. 79].

However, it appears questionable to us whether the field of linguistics can be employed in such a direct manner for reasoning[44], even if the above sentence seems to be an example of common sense reasoning, which is also strived for in the field of artificial intelligence.

2.3.3 Sociology & Psychology: Roles and Expectations

Within the social sciences one finds an area called *role theory*. It spans from social psychology to sociology and anthropology, starting around the early 1930s. Role theory deals with

> [...] patterns of human conduct; with expectations, identities, and social positions; and with context and social structure as well as with individual response. [Campbell, 1999]

There are various definitions for the term *role*, which, in general, is assumed to provide a theoretically useful connection between social structure and individual behaviour (cf. [Biddle and Thomas, 1966]). In addition, several specialised fields have developed, each dealing with slightly different aspects of role theory: *functional role theory*, *symbolic interactionist*, *structural*, *organisational*, and *cognitive role theory*. However, in order to provide a brief introduction of the role concept in the overall field, we focus on some interesting ideas described in [Campbell, 1999].

A first distinction is made between *static* and *dynamic aspects* of roles, the former referring to status and positions (like that of a `prime minister`), the latter to behaviour common for humans in these positions (the `appointment` of the members of the `cabinet` may belong to the dynamic aspects of a `prime minister`). A number of different dynamic roles are said to contribute to a single static role. Another new assumption (according to [Campbell, 1999] made by Sarbin, a symbol interactionist) is that social roles are not completely dependent on the individual, although, certainly, a social role is very closely related to the way an individual performs a particular role, but there are different levels of how much one identifies her-/himself with a role. Structural role theory focuses on the relations between roles and groups. Therein, a social group consists of roles necessary for its development and existence. A similar approach is organisational role theory, which is directed towards roles in hierarchical and task-driven organisations. Moreover, the concept of *role transition*

[44] One may think of a `hole` in the above `pocket`, or consider the question of what `in` means: completely in the pocket, more than one half of the 8-ball in the pocket, and so forth.

(recall role evolution and extension in [Dahchour et al., 2002], see p. 30) can be found in role theory. It refers to a shift in the expectations assigned to a role, an idea which has not been accounted for in computer science formalisms yet. In role theory, *role expectation* is one of several key concepts which are examined in individual subareas. Finally, a subdivision of *role expectations* shall be mentioned, namely into *norms*, *preferences*, and *beliefs*. Thus the specification of these three concepts may be used in the overall specification of roles.

On the whole, concepts and ideas from role theory hardly seem to be applied directly in computer science. [Kendall, 1999] is the only work in computer science we have encountered which mentions the existence of role theory, but even there it appears only marginally. Establishing a connection to philosophy (cf. section 2.3.1), it seems that role theory mainly deals with the link between the conceptions of roles on a social ontological level by individual humans being on a different ontological level.

2.4 Summary of Main Issues

Up to this point, many approaches involving roles in various ways have been presented, but what are the commonalities, if any, among them? Firstly, let us consider four recurrent terms which accompany the discussions of roles:

- *role*
- *player* or *filler* of a role
- role in a *context*
- *natural type* or *base classifier*

A tentative assignment of formalism-specific concepts to the first three terms from above yields the overview shown in Table 2.1, which covers all formalisms reviewed herein, apart from those of the background areas in section 2.3. Natural types are always discussed in connection with the player of a role and seem to depend on that notion. Therefore, they are not considered to be a conceptual primitive and not contained in Table 2.1. Note that, where possible, a corresponding GOL term or another synonym is used if the element of the formalism is denoted by the column heading. For instance, description logics speak about roles and fillers, but a description logic role corresponds to a binary relation, and the fillers to its arguments.

We identify the following questions as major issues in existing treatments of roles:

1. What is a context of a role, and how do role and context relate to each other?
2. Are roles exclusively UNIVERSALS, or do role INDIVIDUALS exist?
3. How is a role related to a player?
   (a) How many players does a role have?
   (b) How many roles can one player fill?
   (c) Is a player always bound to one natural type?

These questions will likewise form the basis of our own analysis and classification of roles in chapter 3. Anticipating the result of a detailed elaboration in the next chapter, we will concentrate on the first question, which appears to be the most neglected one in the literature, but which gives much information about roles.

Formalism	**Page**	**Role**	**Player/Filler**	**Context**
Description Logics	16	RELATION	RELATUM	—
Scripts	18	Action(s)	Participant	Script
Sowa	19	UNIVERSAL	Instance	Binary Relation
Guarino	21	UNIVERSAL	Instance	(Dependent Property)
Topic Maps	23	RELATUM	Resource	Association
ER	25	RELATUM	Entity	Relation
UML	26	Rolename	Object	Association
	27	Qualifier	Object	Association
	27	Classifier Role	Object	Collaboration
	28	Actor	Instance	Use Case(s)
	28	Role Type	Instance	*Context*
Steimann	28	Role Type	Instance	Relation
Coulondre & Libourel	31	View	Instance	Class
Role Object Pattern	31	Role Object	Core Object	*Context*
Agent Modelling (Gaia)	32	Social Substance	Agent	Society
OPM	33	(Procedural Link)	Object	Process
	33	State	Object	Object
HL7	34	*Role*	Entity	Scope

Table 2.1: Roles, Players, and Contexts of Reviewed Formalisms.

3 A Context-Centred Approach to Roles

The second chapter demonstrated that the concept of roles pervades a considerable number of modelling formalisms. Apart from roles themselves, two recurring, role-related concepts were identified, which are denoted by *player* and *context*. Further, a number of high-level questions were posed which should be answered by a characterisation of roles. Such a characterisation is the goal of this chapter.

Herein, this characterisation is to be approached via a different route than the player-centred access to roles which is pursued in the literature. We justify this idea in section 3.2, which then starts the analysis with the question of which ontological categories (on the basis of GOL) may appear as contexts for roles. Three such categories are identified, on the basis of which three special types of roles are introduced in sections 3.3, 3.4, and 3.5. Finally, we extract the commonalities of these role subtypes and discuss their interrelationships in section 3.6. This is the closest point we reach regarding a single, integrated role concept. The discussion now starts with a more detailed introduction of basic GOL terms which are employed in our formalisation.

3.1 Preliminaries

3.1.1 Signature

GOL (as described in the Reference Manual [Degen et al., 2003]) is the general framework within and for which axioms are developed herein. Accordingly, the formalisation is based on standard first-order logic with identity. Most GOL terms will appear as logical constants, since they are supposed to refer to unique entities in reality as well as it is necessary to refer to them directly in some axioms. FORMAL RELATIONS, though, need to be handled as an exception. They are introduced as predicates of the meta language. GOL itself further introduces unary predicates which mirror an assignment of entities to certain GOL categories (denoted by constants; as mentioned above)[45].

Table 3.1 contains the signature of our formalisation[46], which is identical in content to that part of GOL which is used in the sequel. The corresponding understanding of a symbol in terms of GOL is specified in the third column, where the entry *none* is given for notions which are not contained in GOL but introduced below. Column four refers to the sections where each notion is presented together with related concepts. The section listed may be either the general introduction in chapter 1, or the more detailed descriptions in one of the following sections. With regard to the symbols themselves, prefix notation is used for textual symbols, whereas infix notation applies to the non-textual ones.

Note that the following reading applies to non-standard predicates which are not explicitly defined. We follow the convention that capitals are chosen for variables which clearly denote UNIVERSALS or

[45] The language part of GOL itself is not applied in this thesis, as some GOL entities like those denoted by such constants are not available in the GOL object language. Notice further, that the categories of GOL themselves are viewed as UNIVERSALS herein, characterised by the descriptions and axioms in the Reference Manual [Degen et al., 2003]. In contrast, these categories may likewise be considered classes instead of universals. However, this is a choice without severe consequences for the presented account of roles itself. Rather, a few simple adaptations of some axioms suffice to refer to categories as classes with the axioms presented.

[46] Note that this signature is to be supplemented by several new terms in order to capture roles. This completion is performed in section 3.2 in order to clearly distinguish between established GOL concepts (plus some auxiliaries defined herein) and notions which are introduced in the present work in order to formalise roles. Finally, not all symbols of GOL are fixed at the time of writing; therefore some GOL entities may be denoted differently than in [Degen et al., 2003].

	Symbol	—	Corresponding GOL Concept	Section
Constants:	$\mathscr{R}$	—	RELATOR	1.4
	$\mathscr{P}$	—	SUBSTANTIAL PROCESS	3.1.2
	$\mathscr{S}$	—	SUBSTANCE	1.4
Unary Predicates:	Relator	—	RELATOR	1.4
	Process	—	SUBSTANTIAL PROCESS	3.1.2
	Substance	—	SUBSTANCE	1.4
	SocSubst	—	*none*	3.1.3
	disjoint	—	*none*	3.1.1
Binary Predicates:	$\in$	—	ELEMENT-OF Relationship	1.4
	$::$	—	INSTANTIATION	1.4
	$\longrightarrow\!\triangleright$	—	IS-A Relationship	3.1.1
	$<, \leq$	—	PART-OF Relationship	1.4
	$:\triangleright$	—	CONTAINMENT	3.1.2
	ontic	—	ONTICAL CONNECTEDNESS	3.1.2
	sip	—	*none*	3.1.2
Ternary Predicates:	prb	—	PROCESS BOUNDARY	3.1.2

Table 3.1: GOL-Based Signature of the Formalisation.

SETS. Otherwise lower case letters are used. Let x, y and U be variables. $x :: U$ says "x is an instance of U", $x < y$ denotes "x is a PART-OF y", $x\!:\!\triangleright y$ denotes that "x is contained in y", and $\mathsf{ontic}(x,y)$ represents the fact that "x is ontically connected to y". Moreover, the following definitions of predicates in Table 3.1 can be given.

- For two universals U and V, U IS-A V, denoted $U \longrightarrow\!\triangleright V$, iff being an instance of U implies being an instance of V.

$$\forall U \forall V\, (U \longrightarrow\!\triangleright V \leftrightarrow_{def} \forall x\, (x :: U \rightarrow x :: V)) \tag{3.1}$$

- A set S of universals is *disjoint* iff each two members of S have no instance in common.

$$\forall S\, (\text{disjoint}(S) \leftrightarrow_{def}$$

$$\forall U \forall V\, (U \neq V \wedge U \in S \wedge V \in S \rightarrow \neg \exists x (x :: U \wedge x :: V))) \tag{3.2}$$

- A substance s appears in a certain process p, symbolically $\mathsf{sip}(s,p)$[47], iff there is a process boundary e of p ($\mathsf{prb}(p,t,e)$, see section 3.1.2) which contains the substance.

$$\forall s \forall p\, (\mathsf{sip}(s,p) \leftrightarrow_{def} \mathsf{Substance}(s) \wedge \mathsf{Process}(p) \wedge \exists t \exists e\, (\mathsf{prb}(p,t,e) \wedge s\!:\!\triangleright e)) \tag{3.3}$$

[47] The predicate name is derived from the phrase "substance in process". Note that a formal relation called PARTIC was recently introduced in GOL which seems to resemble the notion defined by sip. However, PARTIC may also refer to ABSTRACT SUBSTANCES, cf. also section 3.4. Therefore we adhere to the definition of sip for our purposes.

3.1.2 Processes and Endurants in GOL

It is necessary to describe the connections between GOL's PROCESSES, ENDURANTS and ABSTRACT ENDURANTS in detail, because an understanding of these categories and their interplay is required for some of the axioms about processual roles presented in section 3.4. The discussion will also touch on CHRONOIDS and CONFIGURATIONS, but rather as auxiliary concepts.

Let us start with PROCESSES, sometimes more precisely called COHERENT PROCESSES in [Degen et al., 2003]. Coherent processes are extended in time and they do not have gaps; thus a single, connected chronoid c can be assigned to an arbitrary coherent process p (hereafter process). This chronoid is called the FRAMING CHRONOID, and it corresponds to the temporal extension of p. Furthermore, there is a projection relation, prb, such that a process is projected on a time boundary within its framing chronoid. The result of the projection of process p on time boundary t is named the PROCESS BOUNDARY of p on t, call it e. Formally, this is expressed by $\mathsf{prb}(p,t,e)$.

A process boundary, for its part, is an endurant. In general, this will be a CONFIGURATION, i.e., an entity "containing" several entities of the more basic endurant types SUBSTANCE, MOMENT, and/or RELATOR, which together form an integrated whole called a configuration. On the basis of the categories to which the entities in the boundary of a process belong, a classification of processes into *moment*, *relator*, and *substantial processes* is referred to in [Degen et al., 2003]. For our analysis, substantial processes are particularly important. They are defined as processes any boundary of which contains at least one substance. Substantial processes should not be confused with SUBSTANCE PROCESSES, however. The latter concept refers to an understanding of substances themselves as processes, and thus this concept approaches a perdurantistic perspective.

Returning to endurants, these are considered in a very strict sense in GOL, in that they exist only at exactly one time boundary, denoted by $\mathsf{at}(e,t)$ for an endurant e and a time boundary t. Nevertheless, an account of diachronic identity (cf. section 1.4) is provided by ABSTRACT ENDURANTS and a new basic relation, ONTICAL CONNECTEDNESS. [Degen et al., 2003] provides a sample explanation of how the proper name John is to be understood in GOL. John, viewed intuitively as *one* entity which persists over a period of time, refers to an abstract substance in GOL, call it $endur(\mathtt{John})$[48]. This is a universal, such that, at any time boundary t, there is at most one instance of it. These instances, which are determinable by the time boundary at which they exist, are actual GOL endurants and, intuitively, they correspond to John at time t. In addition to the universal $endur(\mathtt{John})$, a new basic relation is assumed to relate the Johns at different times t and t', as these are by definition different endurants. ONTICAL CONNECTEDNESS is the name of this relation, which is denoted by the binary predicate ontic.

Finally, Table 3.1 introduced the predicate sip which is intended to relate a substance s and a substantial process p if there exists a process boundary of p which contains s. The formal definition of sip can be found on page 43 (formula 3.3). This predicate serves as an "abbreviation" in the axiomatisation of processual roles in section 3.4.

3.1.3 Social Entities

All but one of the basic terms of Table 3.1 can be based solely on GOL. The predicate SocSubst is an exception. It is intended to describe a subcategory of SUBSTANCES, which is not yet accounted for in

[48] This denotation refers to a former version of the Reference Manual.

GOL, namely *social substances*. Examples of social substances are `universities`, `clinical trial centres`, `sports clubs`, `companies`, and `organisations` in general, but also `shares`, `licenses`, etc., as well as `deans`, `nurses` and the like. There are also other social entities, however, e.g., *social processes*, to name a very important one. The existence of social entities is philosophically supported by Johansson's level ontology as well as Searle's theory around the *X counts as Y formula* (see section 2.3.1). Further, other top-level ontologies in the field of computer science also contain that notion, for instance, the DOLCE ontology (section 2.1.4).

Of course, the social world is highly dependent on the realm of substances which are not predicated as social. The latter may be called *material entities*[49]. The social "part" of the world supervenes on or emerges from the material one[50]. Moreover, social substances really are substances, i.e., they are endurants, moments inhere in them and they can be connected by relators. *Abstract social substances* participate in processes – social processes, commonly. In order to exist, social substances have to be represented or acted by material entities.

Within human societies, i.e., where humans are the material entities acting for their social "counterparts", the existence of a social substance is often caused by an *act of declaration*. This means a certain material process is understood also as a social process and possibly as the inauguration of some new entity. Taking up a former example, `signing a document` may be viewed as the `foundation` of some `company`, which from then on exists "by definition". But there are also cases where social entities emerge, simply from interaction of a number of material entities. This holds for `ant colonies`, as well as for `groups of` interacting `humans`.

All parts of social substances are themselves considered to be social substances, conforming with the idea of relatively separated ontological levels. But why is the material entity `university building` not a material part of the social entity `university`? It appears preferable to state that a university `owns` that building and its interior for the following reasons. Consider the fact that, if a university changes buildings, it remains the same university. More precisely, it instantiates the same abstract social substance. One may argue that this is not a satisfactory argument because also material substances may lose parts over time and remain identical. Another argument involves moments inhering in material and social substances. It seems strange to ask for the `weight` or the `colour` of a university, if not figuratively. On the other hand, moments of these types inhere in the building, or material substances in general. Finally, the material "parts" of a university would not be spatially connected, which is rather common for material substances.

These arguments are some hints for a separation between material and social entities, an approach we will follow. It complements some parts of our analysis of roles, and it will, on the basis of roles, explain why a `dean` is a part of a `university` and of its parts, like `institutes` or `departments`, while no human is. Role-related concepts will be important for bridging the gap between the material and the social world.

---

[49]There are probably more ontological levels than two. However, this question is far beyond the scope of this work. The separation between what we call social and material entities seems to be important, though, and it should be intuitively comprehensible on the basis of our examples.

[50]Conforming with the individuation principle of substances, social substances are spatially and temporally located. Tentatively, we assume that they are co-located with their material constituents, which again corresponds to the DOLCE approach [Masolo et al., 2002]. However, it should be stated that this is a debatable decision. Maybe other individuation principles can be found, thus allowing for social substances without a direct relation to space.

3.2 Types of Roles

3.2.1 Justification for the Context-Centred Approach

Now that a basic framework has been established, the actual analysis of roles can be tackled. At the end of chapter 2, *player* and *context* were identified as concepts which are employed to describe roles in many approaches. The relationship between roles and their players is often addressed in great detail, while little attention is paid to the context in which a role exists. We call this situation the *player-centred approach* to roles. In contrast, we will follow a *context-centred approach*, where contexts are considered prior to players. Consequently, contexts will be used to classify roles, before the player relationship is briefly addressed in connection with a predetermined type of context.

It appears reasonable to discuss players on the basis of a fixed type of context because of the determination relationships between roles and contexts, as well as roles and players, respectively. A context determines which roles occur. Conversely, a role describes a player to a certain extent, instead of the idea that a player would characterise a role. For example, if `student` is said to be a role, this yields more precision if one adds that it is a role of the `studies-at` relationship. However, stating that `John` is a `student` describes `John`, but not the role `student`. Indeed, most accounts which were presented in the second chapter deal with the question of what the context of roles is. Usually, however, exactly one context is simply named, and then the relationship of roles and players is expounded further. Support for this observation will be presented in the following section. On the whole, it seems appropriate to determine those ontological categories which appear as contexts in the literature, and then to examine the player relationship for each category alone. Finally, we will attempt to extract commonalities between contexts of different categories.

3.2.2 A Collection of Contexts

The claim that players are usually unspecified, while contexts are ontologically fixed is based on Table 2.1 (p. 41). The context column of this table refers to a variety of categories, whereas the player column remains unspecific. In order to provide more sound support, we consider a mapping of terms from Table 2.1 into GOL categories.

The player column contains the following terms, arranged in descending order according to the absolute number of occurrences of each (given in parentheses): *Instance* (6), *Object* (5), *Entity* (2), *Relatum* (1), *Participant* (1), *Resource* (1), *Core Object* (1), and *Agent* (1). The first three and *Core Object*, covering 14 out of 18 approaches, are clearly understood as referring to INDIVIDUAL. *Resource* and *Agent* sound more specific than individual, but either term is equally unspecific with regard to its corresponding formalism as individual with respect to GOL (e.g., everything a topic map describes is considered a resource). Consequently, they are also mapped to INDIVIDUAL. The remaining two terms, *Relatum* and *Participant*, are attributed to a naming of concepts which is not compatible with GOL. This conception should be clear once the overall theory is established[51].

Turning to the column of contexts, one can find a greater variety. As we can see, there are clear ref-

[51] That means, the neglection of these terms cannot be well explained at the present stage, but it fits the theory with hindsight. As an anticipation, roles in our sense are implicit in description logics, due to the assignment of relations to single concepts. With respect to scripts, the notion of *actions* as roles corresponds to our subtype of processual roles, and in this connection *participant* is an appropriate term for the players of such roles.

erences to PROCESSES, RELATIONS, and SUBSTANCES, i.e., any context in Table 2.1 can be mapped to one of these categories. Among these, RELATIONS are clearly preferred. This is obvious from the terms *Association* (3) as well as (*Binary* (1)) *Relation* (2), and it is hidden in the *Object* (1) context in OPM role-playing sentences. Arguably, Guarino's approach using *Dependent Properties* (1) is added to this collection, conforming to his older explanations of basically one and the same formal account. The latest of his annotations (see section 2.1.4) may also be understood as directed towards PROCESSES. The category of processes needs to be understood rather comprehensively; for instance, complex processes like `clinical trials` or the `development of a country` over a certain period are also to be taken into account, comprising all their components[52]. The terms which can be mapped to processes more directly than Guarino's descriptions are *Scripts*, *Collaborations*, *Use Cases*, and of course, *Processes* of OPM (each of these occurring once). What is still missing are terms that defend the existence of contexts which are SUBSTANCES. Admittedly, these seem to be rare since this is only clear for *Society* (1) and *Scope* (1). However, those approaches which have not specified the notion of *Context* (2) further advocate roles as universals on the basis of assigning role-specific properties and behaviour to these roles (more precisely, role types). This corresponds to the specification of INHERENCE between instances of certain MOMENT and SUBSTANCE universals. Therefore, we think that these approaches support roles which can be interpreted as referring to SUBSTANCES, especially in connection with social substances, which are bound to certain societies. Finally, the context *Class* of Coulondre and Libourel as well as description logics have not been covered yet. For the latter, the problem is that the term *role* actually refers to a relation, which represents the context rather than the role (cf. also footnote 51). Similarly, Coulondre and Libourel seem to actually refer to *views* rather than *roles*, where the differences of both concepts are described in section 4.1. Therefore, the *Class* entry in Table 2.1 (p. 41) is not taken into consideration.

Note that the categories ascribed to contexts (in contrast to players) apply by no means exclusively for most approaches. Rather, the last column of Table 2.1 refers to the main category which is extracted from each approach, as can be verified on the basis of chapter two. In spite of the fact that mixtures of categories are common for contexts, we attempt to capture roles according to a single type of context in order to develop pure and distinct notions, such that seeming contradictions which would arise in one general approach can be resolved with respect to these specific concepts. Afterwards, interrelationships of roles according to different context types will be discussed and a common core will be extracted (cf. section 3.6).

3.2.3 Extension of the Signature

In the previous section, we identified the GOL categories RELATION, PROCESS, and SUBSTANCE as those categories to which all contexts of the notions of roles reviewed so far can be mapped. Therefore, we propose the following major classification of roles into the following subtypes:

- *relational roles* — on the basis of a dependence on relations
- *processual roles* — for roles describing participation in a process
- *social roles* — stressing role-specific properties and behaviour as well as membership of some society

[52] In some cases, the notion of SITUOIDS seems even more appropriate. However, as noted above, SITUOIDS are in a preliminary state of development, and therefore we refrain from employing them as basic terms herein.

These role subtypes are added to the signature on page 43, as well as a hypothetical category for all roles. Table 3.2 contains this extension of the signature, which further comprises two new basic formal relations in order to express the connections between roles and players, as well as roles and contexts, respectively. The former is called PLAYS, denoted by the symbol $\rightsquigarrow$, where $x \rightsquigarrow y$ represents the fact that an entity x is the player of a role y. With regard to contexts, the corresponding relation is called ROLE-OF, where $x \succ y$ stands for a role x being a role of some entity y which serves as a context. The motivation for choosing new relations instead of existing ones is that new relations do not exhibit dependencies to other notions, while they can still be identified with an existing relation after their introduction (cf. sections 3.6.2 and 5.1). The formalisation to be given in the remainder of this chapter is based on the union of the signatures in Tables 3.1 and 3.2.

	Symbol	—	**Intended Concept**	**Section(s)**
Constants:	$\mathcal{Q}$	—	Role	3.3
	$\mathcal{Q}_{\mathsf{rel}}$	—	Relational Role	3.3
	$\mathcal{Q}_{\mathsf{proc}}$	—	Processual Role	3.4
	$\mathcal{Q}_{\mathsf{soc}}$	—	Social Role	3.5
Unary Predicates:	Role	—	Role	3.3
	RelRole	—	Relational Role	3.3
	ProcRole	—	Processual Role	3.4
	SocRole	—	Social Role	3.5
Binary Predicates:	$\rightsquigarrow$	—	Plays Relationship	3.3–3.5
	$\succ$	—	Role-Of Relationship	3.3–3.5

Table 3.2: Signature Extension: Symbols for Representing Roles.

3.3 Relational Roles

Before starting the discussion of relational roles, the organisation of the following three sections, i.e., sections 3.3, 3.4, and 3.5, shall be outlined. Each of them starts with an introduction of the corresponding role type, which consists of an intuitive description of the intent behind the role type, thereby being related to the appropriate GOL entities. The introduction concludes with an axiomatisation of that role type. In the remainder of these sections, several implications and effects for modelling with each role type are discussed. Note that within these sections the term *role* is also used as an abbreviation for the corresponding specific role type, if this can be clearly resolved.

3.3.1 Introduction of Relational Roles

The context most often mentioned for roles is that of relations, as explained above. Approaches which capture this idea in a relatively pure form are UML rolenames and roles as defined in the topic map formalism. Intuitively, a relational role refers to a single entity or a single argument of a relation. The "plug-in point" for exactly one argument may be a metaphor to illustrate a relational role. The actual argument to be plugged in is the *player* of that role.

The translation into the GOL vocabulary starts with the term *relation*. Relations are considered to be restricted to MATERIAL RELATIONS, for two reasons. Firstly, domain-specific relations are assumed to be material relations in general, and secondly, if formal relations are also to be captured in terms of relational roles, this approach is threatened by an infinite regress, because at least the formal relations PLAYS and ROLE-OF will be necessary to tie players to roles, and roles to relations, respectively. Consequently, a regress arises if the arguments of, e.g., PLAYS are attached to PLAYS in terms of certain relational roles, since the attachment of these roles would again involve the PLAYS relationship. This resembles the infinite regress obtained by demanding relations which relate relations and their relata, which is well-known in philosophy and deliberately avoided in terms of formal relations in GOL. Of course, many of the following arguments may be directly applied to formal relations, however, unless the problem of the infinite regress can be solved, the discussion remains constrained to material relations.

Understanding relations as material relations has another advantage. Material relations do not rely on tuples within a set. Rather, they are mediated between INDIVIDUALS by RELATORS. Let us introduce some examples here. Consider the relations `parent-of` and `connects`. Sentences like the following:

“`Mary` is a `parent of John`.”
“The `bridge connects` the northern and the southern `river bank`.”

refer to individual relators, for instance the one mediating between `John` and `Mary`, call it `r`.

According to the relation between `John` and `Mary`, two relational roles can be detected. They refer to `John`’s *being the child* of `Mary`, and `Mary`’s *being the parent* of `John`. These are the two *roles of* the relator `r`. They are not to be confused with `John` and `Mary` themselves, which are the *players* of either role. Calling these roles `child` and `parent`, we suggest the following formalisation of the last example using the formal relations ROLE-OF and PLAYS:

$$\mathtt{child} \succ \mathtt{r} \quad \wedge \quad \mathtt{parent} \succ \mathtt{r} \tag{3.4}$$

$$\mathtt{John} \rightsquigarrow \mathtt{child} \quad \wedge \quad \mathtt{Mary} \rightsquigarrow \mathtt{parent} \tag{3.5}$$

On the basis of what has been said so far, an intuitive definition for relational roles can be given. Naturally, this definition is very similar to that of roles in topic maps (see p. 24). Note that it is indisputable in all reviewed approaches that there is always only one filler of an argument position of a relation. Further, to avoid giving a cyclic definition, the term *role* is replaced by *part* in an *intuitive* sense, which sounds common on a natural language level.

Definition 3.1 (Relational Role) *A* relational role *is a proper part of an arbitrary relator r. It refers to exactly one entity e and corresponds to the way e participates in r.*

In this definition, relational roles are implicitly INDIVIDUALS (or more precisely, ENDURANTS), because relators are endurants (and thus individuals), and parts of ENDURANTS are assumed to be ENDURANTS themselves. Of course, there are also relator universals, i.e., universals whose instances are relators with common features. Analogously, universals of relational roles can be postulated. Extending the previous example, if `Mary` has another child, some human plays another individual role `child`’ which is distinct from the `child` role above. However, both roles may instantiate the relational role universal `Child`, in contrast to the `parent` role. Indeed, what are specified in the models of most formalisms are relational role universals. Further examinations on the connection between relational role individuals and universals are presented in section 3.3.3.

Axiomatisation

In order to introduce roles more formally, we develop the following axiomatisation for relational roles. However, some axioms on roles in general need to be anticipated in order to avoid the reintroduction of similar axioms in subsequent axiomatisations of other role types.

- The universal ROLE is characterised by the relationships ROLE-OF and PLAYS.

$$\forall q\,(\mathsf{Role}(q) \leftrightarrow q :: \mathscr{Q}) \tag{3.6}$$

$$\forall q\,((\exists x\,(q \succ x)) \leftrightarrow \mathsf{Role}(q)) \tag{3.7}$$

$$\forall q\,((\exists x\,(x \rightsquigarrow q)) \leftrightarrow \mathsf{Role}(q)) \tag{3.8}$$

- Axioms on relationships between the universals *relational role*, *role* and *relator*.

$$\text{disjoint}(\,\{\,\mathscr{Q}, \mathscr{R}\,\}\,) \tag{3.9}$$

$$\mathscr{Q}_{\mathsf{rel}} \longrightarrow\!\triangleright\, \mathscr{Q} \tag{3.10}$$

$$\forall q\,(\mathsf{RelRole}(q) \leftrightarrow q :: \mathscr{Q}_{\mathsf{rel}}) \tag{3.11}$$

- Relational roles and relators are interdependent notions. Relators consist of at least two distinct roles.

$$\forall q \forall r\,(q \succ r \rightarrow (\mathsf{RelRole}(q) \leftrightarrow \mathsf{Relator}(r))) \tag{3.12}$$

$$\forall r\,(\mathsf{Relator}(r) \rightarrow \exists q \exists q'(q \neq q' \wedge q \succ r \wedge q' \succ r)) \tag{3.13}$$

- There is exactly one player for each relational role.

$$\forall q \forall x \forall y\,(\mathsf{RelRole}(q) \wedge x \rightsquigarrow q \wedge y \rightsquigarrow q \rightarrow x = y)) \tag{3.14}$$

- A relational role cannot be played by a relator or a relational role itself.

$$\forall q \forall x\,(\mathsf{RelRole}(q) \wedge x \rightsquigarrow q \rightarrow \neg\mathsf{Relator}(x) \wedge \neg\mathsf{RelRole}(x)) \tag{3.15}$$

It is worthwhile to discuss the non-trivial axioms. The reason for understanding roles, in general, and relators as universals with disjoint extensions in axiom 3.9 is that relators are mediating elements for *at least two* other entities. Conversely, a role only refers to one other entity, at least in this "independent" way (of course, a role itself depends on a context, but the context itself cannot be understood as a second independent entity).

Axiom 3.15 is postulated for relational roles, although especially object-oriented formalisms often claim that roles may play roles. From our point of view this claim refers to social roles (section 3.5), and is not reasonable in the case of relational roles. This appears as an example of a feature of roles which can be ascribed to some contexts, while it is not true for others, thereby backing the context-centred approach. The fact that relational roles are *dependent* entities supports axiom 3.15. If a relational role played another relational role, that would require a relator which mediates between two relational roles. The analogous implication holds for relators. In either case, this is considered counter-intuitive, because relators are supposed to connect entities with a certain degree of independence, like

SUBSTANCES or CONFIGURATIONS. Once all descendants of ENDURANT are identified, one may make allowance for a reformulation of axiom 3.15 such that the applicable categories are enumerated.

Finally, axiom 3.14 is uncontroversial in the literature. It is not necessary to allow for two players of a relational role, because the entity which actually mediates between the players is the relator. Consequently, this case is covered more appropriately by two relational roles, each with one player.

3.3.2 Material Relations & Predicate Logic

With this basic account of relational roles at hand, some issues of the representation of MATERIAL RELATIONS can be discussed. There are interesting insights even for predicate logic, if material relations are formalised in the same manner as formal relations, i.e., a predicate R denoting a material relation is introduced in order to express that two entities a and b stand in the relation R. The logical atom $R(a,b)$ describes what is to be expressed, which is the standard approach for formalising relations in predicate logic.

For this type of formalisation together with the standard semantics of first-order logic, the following problems shall be examined:

1. **Instantiation Problem:** Different "instances" of a relation between the same INDIVIDUALS cannot be expressed. This means, $R(a,b)$ always refers to one and the same RELATOR glueing a and b together, which is mirrored by the interpretation of $R(a,b)$ as a tuple of a *set*, which therefore cannot occur several times.

   Now, in GOL one can consider the following case. Let there be two relator UNIVERSALS R_1 and R_2 with disjoint extensions, which have some instances that mediate between the same entities, say between a and b. Assume further that both relator universals are specialisations of the above R. Then there are two distinct instances of R which mediate between a and b. In the set-theoretic interpretation of R, these two instances collapse into one tuple (a,b).

   The *foundation* (see section 1.4) of a relator is another source to recognise the instantiation problem. Two distinct relators instantiating the same universal and mediating between the same individuals may originate as a result of being independently founded.

2. **Extensionality Problem:** Two relation predicates R and R' may have the same set as their interpretation, which makes them indiscriminable within a model. But the intension may be differ, e.g., the foundations of R and R' may be unequal. (Analogously, universals may have different intensions, but the same extension; see [Sowa, 2000, p. 99]).

3. **Arrangement Problem:** The arrangement of the arguments of a relation is fixed, i.e., $R(a,b)$ and $R(b,a)$ are interpreted as different facts. For binary relations, this culminates in the concept of *inverse relations*, which is usually denoted by R^{-1}. However, intuitively, R and R^{-1} refer to the same relation in reality, which is duplicated once for every non-standard arrangement of its argument positions. Clearly, the problem applies also to relations with an arity greater than two.

   As a concrete example, consider the sentences "`Mary` is a `parent of` `John`." and "`John` is a `child of` `Mary`.", which refer to the same piece of reality. Further support for the existence of the Arrangement Problem can be derived from section 3.3.4.

GOL provides a solution to two of the three problems defined above, since relators already solve the Instantiation and the Extensionality Problem, because they are individuals and thus instantiate universals. However, the Arrangement Problem is not addressed by relators or any other GOL notion[53]. In the remainder of this section, we will show that relational roles can fill this gap.

Relational roles describe the way in which an entity participates in a relator. This kind of knowledge is also contained in the representation as $R(a,b)$. But here it is left implicit, and is attached to the different argument positions. Relational roles allow us to make this knowledge explicit, thereby solving the Arrangement Problem. Reconsider the example contained in formulae 3.4 and 3.5 on page 49. The Arrangement Problem does not exist in these formulae, because the connections between the arguments and the relator are spread over separate logical atoms. These, of course, are not bound to a particular order due to the commutativity of logical conjunction. This holds especially for the ROLE-OF atoms in formula 3.4, which correspond to the fixed positions in $R(a,b)$. The latter, though, are by definition not interchangeable and therefore exhibit the Arrangement Problem.

As a remark, it is necessary to have diverse relational roles, i.e., relational roles instantiating *different* universals. They are required in order to express the second aspect hidden in different argument positions: entities participate differently in relators (cf. the difference between `child` and `parent`). If this was not the case, one could drop relational roles completely, and establish the PLAYS relation directly between an entity and a relator. There even is a solution to differing participation in relators which renounces relational roles. Instead of one formal relation PLAYS, one might introduce a number of different formal relations such that only one of them directly glues an entity to a relator[54]. From an intuitive point of view, however, this should require as many formal relations as there are relational roles in our approach. This resembles a merge between a single relational role and the standard PLAYS relationship. Then, however, it is not possible to speak about roles as elements of the object language any more. Therefore, the solution advocating a multiplicity of *plays* relationships is rejected (see also section 3.4.4 for a similar discussion regarding processual roles).

Let us conclude this section with a remark on the need to assign the ROLE-OF and the PLAYS relationship to the individual level. UML rolenames, for instance, may suggest the opposite, because they are specified on a class level (which corresponds to universals). Following this approach in combination with relational roles, say for the ROLE-OF relationship, leads to the situation that individual roles and individual relators have no direct connection. Hence, it is unclear to which relator a relational role belongs, and consequently, the exclusive specification of ROLE-OF and PLAYS on a universal level needs to be rejected. Of course, this does not prevent the introduction of relations between relational role universals.

3.3.3 Restrictions of Role and Relator Universals

Apart from the insufficiency of declaring ROLE-OF and PLAYS relationships on a universal level, another issue is associated with the specification of relational roles on a universal level. The axiomatisation of relational roles introduced so far does not contain any restrictions between individual

[53] See section 5.1 for the identification of this problem in GOL itself, and a proposed solution.

[54] There may also be accounts which, e.g., reflect the order in the usual representation $R(a,b)$. This means that for any place n of a predicate, a distinct relationship *plays-nth-argument* is introduced. However, this idea is not desirable as argued above, for it leaves the knowledge about roles implicit.

relational roles and the universals they instantiate. Of course, it seems desirable to have such restrictions, because these would underline the ontological character of this approach, in contrast to purely set-theoretic accounts of relations where arbitrary unions can be formed, such that the result may lack any "natural meaning" with respect to reality.

One may suggest that every relator universal has some relational role universals associated with it, let us call them *natural role universals* with respect to a relator universal. The converse may be another option, i.e., at least some role universals should be associated with a relator universal. However, finding a suitable account of this issue is not as easy as it may seem at first glance. Difficulties arise from the fact that ROLE-OF and PLAYS are relationships between individuals. It is hardly possible to see a direct connection between a relator universal and a relational role universal, if generalisation (i.e., IS-A) hierarchies of either kind of universals are assumed, but the idea of most specific universals is rejected. In other formalisms, these connections are simply specified on a universal level.

As an example, UML rolenames are translated into relational role universals, in order to find a more abstract definition of natural role universals. Imagine a UML model which contains a self-association[55] `mother-of` of the class `Person`. Further, the rolenames `mother` and `child` are attached to the ends of `mother-of`. The question arises in which way, for instance, the relationship between the rolename `child` and the association `mother-of` should be understood. Mapping the latter to a relator universal $R_{\texttt{mother-of}}$ and the rolename `child` to a relational role universal $Q_{\texttt{child}}$, it would be wrong to think of a ROLE-OF relationship between those universals, because ROLE-OF as well as PLAYS are restricted to INDIVIDUALS. Hence, the universal level statement is more correctly understood as a short-hand notation for some domain specific axioms.

The following definition is created by an extraction of features from the description of UML rolenames. It is useful to introduce this definition for role universals in general, in order to avoid duplications for other role types, because analogous questions can be posed for the latter.

Definition 3.2 (Role Base for a Universal) *Let U be an arbitrary universal. A set S of role universals is called a* role base *of U, iff the following conditions are fulfilled:*

1. *The role universals in S have pairwise disjoint extensions,*

$$disjoint(S). \tag{3.16}$$

2. *Each role of an instance of U is an instance of exactly one role universal in S,*

$$\forall x \forall q\, (x :: U \land q \succ x \rightarrow \exists! Q\, (Q \in S \land q :: Q)). \tag{3.17}$$

3. *For each role universal $Q \in S$ it holds that*

   (a) *the instances of Q are bound to instances of U as their contexts,*

$$\forall x \forall q\, (q :: Q \land q \succ x \rightarrow x :: U), \tag{3.18}$$

   (b) *and each instance of U necessarily has at least one role instantiating Q,*

$$\forall x\, (x :: U \rightarrow \exists q\, (q \succ x \land q :: Q)). \tag{3.19}$$

[55] A self-association connects a class with itself. In this case, UML forces the specification of roles in order to distinguish the places in the association.

Applied to the above example and assuming a mapping of the rolename `mother` to the relational role universal $Q_{\texttt{mother}}$, one can easily comprehend that the set containing only the universals $Q_{\texttt{mother}}$ and $Q_{\texttt{child}}$ is a role base for $R_{\texttt{mother-of}}$, and that this notion of a role base is an attempt to capture the meaning of rolenames assigned on a universal level. The definition is already comparably strict – for example, a role base for a universal U cannot contain optional role types, i.e., role types whose roles occur only in instances of U, but not in *every* instance of it.

In spite of this stringency, we have not found a suitable axiom which would express a general connection between relator and role universals in the sense of a *natural* role base. For instance, one cannot claim that arbitrary collections of role universals need to be a role base for some relator universal. It is further not possible to look at the ROLE-OF relationship between two individuals, in order to infer role universals from these individuals, such that these role universals belong to a base for a certain relator universal. For clarification, consider the following corollary.

Corollary 3.1 *The singleton set containing only the relational role universal* $\mathscr{Q}_{\mathsf{rel}}$ *is a role base for the category of relators* $\mathscr{R}$*.*

Proof: Let $S = \{\mathscr{Q}_{\mathsf{rel}}\}$. The claim is that S is a role base for $\mathscr{R}$. Disjoint extensions of the role universals according to formula 3.16 are trivially given since S is a singleton. The requirements expressed by formulae 3.17 and 3.18 are true by axiom 3.12 and due to S being a singleton. More precisely, $\mathscr{Q}_{\mathsf{rel}}$ can be identified as the uniquely determined Q demanded in formula 3.17, since there is no other element of S and $\mathscr{Q}_{\mathsf{rel}}$ is explicitly bound to $\mathscr{R}$ by axiom 3.12. The latter likewise implies formula 3.18, which has to shown only for $\mathscr{Q}_{\mathsf{rel}}$. Further, requirement 3.19 follows immediately from axioms 3.13, 3.12 and S being a singleton. In greater detail, axiom 3.13 implies the existence of at least one role for each instance of $\mathscr{R}$, axiom 3.12 enforces that this role instantiates $\mathscr{Q}_{\mathsf{rel}}$, and, again, there are no further universals in S for which 3.19 needs to be shown. Consequently, all conditions are fulfilled by S. (q.e.d.)

If one selects an instance of $Q_{\texttt{child}}$ and one of $R_{\texttt{mother-of}}$, one has also found instances of $\mathscr{Q}_{\mathsf{rel}}$ and $\mathscr{R}$, respectively. There may be many other more or less general universals between, e.g., $Q_{\texttt{child}}$ and $\mathscr{Q}_{\mathsf{rel}}$, as well as less general universals than $Q_{\texttt{child}}$. Nothing indicates when the "same" level of generality has been reached for the role and the relator universal, if this is not known in advance.

The last idea to be discussed here is to demand the existence of a role base (hereafter also base, as a short hand) for each relator universal and require uniqueness of natural role bases. Likewise, a base for each relator universal can be defined in the same manner as in axiom 3.12. But this base is not very informative, and would not be considered to be *the* natural base. Formally, in contrast to the use of the definite article, there does not seem to be a unique natural base. Suppose we had a definition of a *natural* role base of a universal, and not only of a role base. For each universal Q in this natural base, one could refer to a universal P of the players of instances of Q. Then, in turn, one could speak of another role universal R, whose instances are determined by the fact that its players instantiate P. If this construction is extended to all members of the natural base, the property of being a natural base is transferred, by construction, to the resulting set of universals (one element of which is R). Of course, this construction is deliberately circular, and one may argue that R is nothing other than Q. However, we feel unable to assess whether there may be longer cycles which are not as transparent as the one constructed. As discussed in Kaplan's [2001] critique of Guarino's formalisations, it is often not useful to exclude single cases. In addition, GOL provides no details thus far about intensions of universals, thus R and Q may still be considered different.

Which conclusion should be drawn from this discussion on how to tag role universals to their intuitively clearly determined "natural" universals? We doubt that a sufficient criterion can be specified in terms of our formalisation as hitherto. A simple and effective way to accomplish this is to introduce a formal relationship between role universals and universals providing contexts. This relationship can explicitly mark natural pairings, thereby allowing for optional role types. Having gained this insight which applies to all role types, including processual and social roles, we proceed with further implications of relational roles.

3.3.4 Notions of Symmetric Relations

Relational roles provide new access to meta-level considerations of relations (or relators, respectively). Consider the notions of *symmetric* and *asymmetric* binary relations. These notions are usually defined in set-theoretic terms:

Definition 3.3 ((A)Symmetry of Extensional Relations) *A binary relation R is called* symmetric *iff*

$$\forall x \forall y\, (R(x,y) \rightarrow R(y,x)). \tag{3.20}$$

A binary relation R is called asymmetric *iff*

$$\forall x \forall y\, (R(x,y) \rightarrow \neg R(y,x)). \tag{3.21}$$

However, there also seems to be an intuitive account of these notions which refers to a replaceability of the arguments in the symmetric case, which is explicitly forbidden in the asymmetric case. Expressed in terms of relational roles and relators this means that a symmetric relator consists of two relational roles which instantiate one and the same universal. Thus the arguments can arbitrarily be assigned to those roles. In contrast, an asymmetric relator consists of two roles which instantiate universals with disjoint extensions, i.e., these universals are different and do not stand in an IS-A relationship. If there are different roles, that is why one has to pay attention to assign an argument to the "correct" position (in the conventional representation of relations). Revised versions of (a)symmetric relators may now be defined. It should be noted that this definition does not need to speak about any argument; all which is taken into account is the relator itself and its "internal structure" of roles, independently of any arguments. However, (a)symmetry has to be defined with respect to role universals, because this information is explicit if relational roles are used for representation.

Definition 3.4 (Intensional (A)Symmetry of Relators) *A binary relator r is called* intensionally symmetric *with respect to a relational role universal Q iff r consists of two roles instantiating Q.*
The relator r is called intensionally asymmetric *with respect to two distinct role universals Q' and Q'' with disjoint extensions iff r consists of two roles such that one instantiates Q', the other Q''.*

This definition refers to the level of individuals, whereas in the case of extensional symmetry it is not possible to introduce a definition on that level only. Conversely, there are two distinct ways to elevate Definition 3.4 to the level of relator universals:

Definition 3.5 (Intensionally (A)Symmetric Relator Universal) *A relator universal R is called* intensionally symmetric *with respect to some relational role universal Q iff every instance r of R is an intensionally symmetric relator with respect to Q.*
R is called intensionally asymmetric *with respect to two relational role universals Q' and Q'' iff every instance r of R is an intensionally asymmetric relator with respect to Q' and Q''.*

Definition 3.6 (Extensionally Symmetric Relator Universal) *A relator universal R is called* extensionally symmetric *with respect to two relational role universals Q and Q' iff every instance r_1 of R is an intensionally asymmetric relator with respect to Q and Q', and there is another instance r_2 of R with the players of r_1 playing interchanged roles in r_2, i.e.,*

$$\forall r_1 \forall q_1 \forall q_1' \forall a \forall b\, (r_1 :: R \wedge q_1 \succ r_1 \wedge q_1' \succ r_1 \wedge q_1 :: Q \wedge q_1' :: Q' \wedge a \rightsquigarrow q_1 \wedge b \rightsquigarrow q_1' \tag{3.22}$$

implies

$$\exists r_2 \exists q_2 \exists q_2'\, (r_2 :: R \wedge q_2 \succ r_2 \wedge q_2' \succ r_2 \wedge q_2 :: Q \wedge q_2' :: Q' \wedge b \rightsquigarrow q_2 \wedge a \rightsquigarrow q_2')) \tag{3.23}$$

Having these abstract definitions, it is natural to pose the question of their relevance in reality. Three examples demonstrate that there are cases where these fine distinctions are useful. First, let the binary relation `customer-of` express that someone is a customer of someone else. Let there further be a law enforcing that if x sells something to y, y also has to sell something to x. In the set-theoretic account of symmetry this would lead to a symmetric `customer-of` relation. However, intuitively, `customer-of` still consists of two different relational roles, `customer` and `seller`. In terms of relators and relational roles, `customer-of` is an *intensionally asymmetric* relator universal. What the law enforces in addition is that `customer-of` is also an *extensionally symmetric* relator universal. Indeed, the *intension* of the relation remains unaffected by the law, while its *extension* acquires symmetry.

A more natural example for this combination of intensional asymmetry and extensional symmetry may be found in biology. Consider a tree and a myccorhizal fungus which live together on the basis of a symbiosis. Among other relationships, the tree `supplies` the fungus with carbon hydrates, and the fungus `supplies` the tree with water, nitrogen, et cetera. Either is thus in a `supplier` role, and also in a `recipient` role. One instance of each of these disjoint role universals occurs in a relator of the relator universal `supplies`. Restricted to the symbiosis domain, `supplies` is intensionally asymmetric, but extensionally symmetric.

The final example refers to an *intensionally symmetric* relator universal. Imagine a situation where two children play with one another. In opposition to a child playing with a toy, there is no remarkable difference between each of the children participating in some game. Hence, the relator `play-together` which mediates between these children consists of two roles which instantiate the same universal `player`. This seems more appropriate than assuming two intensionally asymmetric relators to hold "in opposite directions" between both children.

3.3.5 Anadic Relations

The example of children who `play-together` also leads to the idea of anadic relations, i.e., relations whose arity is not fixed. GOL explicitly allows for anadic relations (cf. [Degen et al., 2003]).

Extending the example, there may be several situations[56] such that a different number of children play some game. Such a situation may found a relator which mediates between those children that occur in this situation. Assuming that all these relators instantiate the same universal `play-together`, this universal is called anadic.

Relational roles can describe such relations in a natural manner. With roles at hand one can easily say that the `play-together` relators always consist of an arbitrary number (greater than two and within a certain range) of roles, each of which instantiates the same universal `player`. It may also be stated that each role is different, or there may be roles with a fixed multiplicity which are combined with roles of variable multiplicities (according to different relators). The latter case can be illustrated by a `connection` relator universal, with a fixed role which instantiates the universal `connector` and, starting from two, a variable number of roles instantiating the universal `connectee`. For example, one bridge as a `connector` may connect two river banks, each of which is thus in the role of a `connectee`. Another bridge may connect three river banks, and both of these situations found instances of one and the same anadic relator universal. In addition to the advantages of relational roles presented in the previous section, descriptions of anadic relations can be stated more efficiently in terms of relational roles. Apart from the inadequacy of translations of instances of one relational role universal into axioms about extensional symmetry, such translations require much more representation. On an informal level, imagine stating that a relator consists of from one to ten roles, compared to specifying axioms that positions one and two behave symmetrically, as well as one and three, . . ., two and three, and so forth.

Of course, it may be argued that *n*-ary relations can be reduced to a collection of only binary relations [Dori, 2002, p. 106]. This is interesting from a theoretical point of view. However, modelling formalisms should allow a modeller to follow her/his intuitions. We think that anadic relations are used intuitively, as well as relational roles are an intuitive means of specification. The latter might be even more intuitive than relations, with regard to a remark on Aristotle who talks of *relatives* instead of relations [Mulligan, 1998, p. 328], although some examples of this notion, like `knowledge` and the `knowable`, may not reflect relational roles.

3.3.6 Uniqueness of Roles regarding Relators

There are also other axioms, the examination of which has not yet resulted in a decision for one of the choices available. One example is the question of whether a role is really bound to exactly one relator. At first glance, this appears quite reasonable and one may advocate this axiom:

$$\forall q \forall r \forall r' \, (\mathsf{RelRole}(q) \land q \succ r \land q \succ r' \rightarrow r = r') \tag{3.24}$$

However, an extension of the example of children `playing-together` suggests a different solution. Suppose that the situation which founds the anadic relator universal `playing-together` also founds a binary relator universal `co-player-of`. A relator of the former glues all children in a certain situation together, while relators of the latter are supposed to connect exactly two of those children. Instead of one relator per situation in the former case, there is a relator for any pair of children in the latter. On this basis it is imaginable to see a relationship between relators of the one and the

[56] An intuitive understanding of the term *situation* is apposite here, rather than a reference to certain SITUOIDS in the sense of GOL elements.

other type, as well as between the roles of these relators. In particular, if it can be assumed that in each case two individual roles belong to the `playing-together` relator as well as to one of the `co-player-of` relators, the relationship between those relators sounds very much like a PART-OF relationship. Accordingly, the following two axioms could replace the one from above[57]:

$$\forall r \forall r' (\mathsf{Relator}(r) \wedge \mathsf{Relator}(r') \wedge r < r' \rightarrow \forall q (q \succ r \rightarrow q \succ r')) \quad (3.25)$$

$$\forall q \forall r \forall r' (\mathsf{RelRole}(q) \wedge q \succ r \wedge q \succ r' \rightarrow (r' < r \vee r < r') \quad (3.26)$$

On the other hand, this mirrors a connection which is based on the foundation of those relators, rather than being internal to the relators themselves. Therefore the first variant may be preferred, so as not to confuse relators with their foundations. We do not have a solution to this problem. According to the literature this work is based on, relations (usually conceived as sets of tuples) do not have parts or the like, and are basic entities, in contrast to composed ones. Deferring further discussion of this issue, section 3.6.2 will take up this question after all role types have been introduced.

3.4 Processual Roles

3.4.1 Nature of Processual Roles

For some role approaches, PROCESSES were identified as the context of a role in section 3.2.2. The formalism of *Scripts* (p. 18) appears to be the closest to what is to be developed on the basis of GOL in this section. The GOL notion of (coherent) processes was introduced in section 1.4 and refined in the beginning of this chapter (see p. 44). In particular, CHRONOIDS are strongly related to processes, as well as ENDURANTS, which appear within the BOUNDARIES of processes. On an intuitive, natural language level, the latter are often called *participants* in a process.

Now, in which ways are processes usually dismantled in order to be understood and/or represented? Often, priority is given to the time aspect, i.e., such a process specification formalism mainly deals with some basic actions, among them a temporal order, concurrency issues and the like (Petri nets, UML statecharts, and possibly UML activity diagrams can be seen to belong to this kind of formalism). A little less frequently, participants are described (either in addition to the temporal structure or exclusively), in order to capture processes. Scripts and UML collaboration diagrams belong to those approaches which combine the specification of participants with time aspects of processes. For processual roles, the focus lies clearly on the side of participants. However, GOL provides other means to express temporal structures, e.g., the PART-OF relation, which can be applied to chronoids and processes.

In the literature, roles with a process context are usually seen to refer to *how* an entity participates in a process. In whatever way roles are understood, they are usually not considered to be processes themselves. But this is exactly the route to be taken herein. Following the argument that a process already contains the way a certain entity participates in it, a processual role is determined as that "part" of a process which refers only to this single participant. Hence, a processual role is an "excerpt" of a process, for which exactly one participant remains constant, because this participant is the determining factor. Consequently, processual roles may be said to constitute processes.

[57]The use of the formal relation PART-OF for relators may be unconventional, but as this is a tentative idea we avoid the introduction of a new formal relationship.

Any process can be dismantled into (its) processual roles. This approach to roles differs from all approaches introduced in chapter 2.

In contrast to the assumption that processes exhibit a relatively high grade of independence, processual roles cannot be considered as independent entities, in general. Instead, they are dependent on their complementing roles, with which they form independent, "compound" processes. "Compound" refers to the fact that there can be several participants in a process. In such a case the process is an independent whole, composed of several interdependent processual roles, namely one per participant. A single processual role cannot exist independently. However, there is an exception: if there is only one participant, for example, in case of a `sleep` process, there is also only one processual role and thus this role alone reaches the level of independence usually assigned to processes.

Let us explain two analogies in order to further clarify our idea of processual roles. Firstly, imagine a `film scene` showing several `people` and `things`. This scene could be edited so that all but one `person` or `thing` are removed. In length the scene remains the same as before, however. Another adequate illustration is contributed by `mimes`. In a strict variant, `mimes` do not use any utilities. Nevertheless, they try to behave as if they had other objects around them. This means they perform a processual role, while the usual complementing roles are missing.

It is difficult to give immediately understandable examples, as language does *not* provide terms with a direct reference to processual roles. Instead, there are terms which denote participants, i.e., endurants[58] which exert a certain processual role. An `opening` process described by a sentence like "`Mary` opens the `door`." may serve as an example. Language provides the term `opener`, but from our point of view this word refers to `Mary`, rather than `Mary's` actions or the way of participation in that opening process. In order to be able to speak about processual roles, the term `opener` is interpreted as the processual role played by `Mary` in the `opening` process. The role of the `door` is usually not named, but it may be formed as the `openee`, i.e., the role played by that entity which is opened as a result of `opening`.

Before an informal definition of processual roles can be introduced, the term *participant* needs to be interpreted in GOL. The notion of participant requires that something keeps its identity across time, i.e., during the chronoid as the process accumulates. Remembering the connection between endurants, abstract endurants, and processes as explained in section 3.1.2, it is only possible for abstract endurants to be identified with participants, because the existence of an endurant is temporally limited to a single chronoid boundary. We additionally assume that participants are limited to abstract substances, because abstract substances are exclusively supposed to provide identity across time[59]. Accordingly, the PLAYS relation connects processual roles and abstract substances.

Conforming with section 3.3.1, an informal definition is presented. Again, "part" is a suitable term to avoid a cyclic definition. On the whole, the similarity between this definition and Definition 3.1 is deliberately retained, since some commonalities of the different role types shall be gathered in section 3.6.

Definition 3.7 (Processual Role) *Let p be an arbitrary substantial process. A part q of p is called a* processual role*, iff there is an abstract substance S which participates in p, and q is equivalent to the way of participation of S in p.*

[58]Actually, abstract endurants according to the current GOL terminology, but this distinction will be discussed below.

[59]This does not contradict the notion of moment or relator processes in [Degen et al., 2003], since these are abstracted from substantial processes.

It depends on the interpretation of "equivalent to the way of participation" in the above definition, whether one allows for temporally shrunk processual roles. In the formal account of these roles, this should be clarified such that only complete participations can be identified as processual roles. Concerning the relation between processual roles and temporal parts, one should not influence the other. This means that if an entity is determined to be a temporal part of a process, it should comprise *all* abstract endurants which are actually present during that time. Accordingly, a processual role is not supposed to lose some phases of a process. Parts of processes which are cut along the time *and* the role dimension can still be described by combinations of these pure versions.

Axiomatisation

- Relationships between the universals *role*, *processual role*, and *process*.

$$\mathscr{Q}_{\mathsf{proc}} \longrightarrow\!\!\triangleright\, \mathscr{Q} \tag{3.27}$$

$$\mathscr{Q}_{\mathsf{proc}} \longrightarrow\!\!\triangleright\, \mathscr{P} \tag{3.28}$$

$$\forall q\,(\mathsf{ProcRole}(q) \leftrightarrow q :: \mathscr{Q}_{\mathsf{proc}}) \tag{3.29}$$

- Processual roles occur only in connection with (substantial) processes.

$$\forall p \forall q\,(q \succ p \rightarrow (\mathsf{ProcRole}(q) \leftrightarrow \mathsf{Process}(p))) \tag{3.30}$$

$$\forall p \forall s\,(\mathsf{Process}(p) \wedge \mathsf{sip}(s,p) \rightarrow \exists q\,(q \succ p \wedge \mathsf{sip}(s,q))) \tag{3.31}$$

- A processual role contains exactly one substance within each of its boundaries.

$$\forall q \forall t \forall e\,(\mathsf{ProcRole}(q) \wedge \mathsf{prb}(q,t,e) \rightarrow \exists! s\,(\mathsf{Substance}(s) \wedge s\!: \triangleright e)) \tag{3.32}$$

- All substances which are contained in the boundaries of a processual role are also contained in the boundaries of the corresponding process.

$$\forall q \forall p \forall t \forall e \forall s\,(\mathsf{ProcRole}(q) \wedge q \succ p \wedge \mathsf{prb}(q,t,e) \wedge s\!: \triangleright e \wedge \mathsf{Substance}(s) \quad \rightarrow$$
$$\exists e'\,(\mathsf{prb}(p,t,e') \wedge s\!: \triangleright e')) \tag{3.33}$$

- According to the process of which it is a role, a processual role is closed under ontical connectedness.

$$\forall q \forall p \forall s\,(\mathsf{ProcRole}(q) \wedge q \succ p \wedge \mathsf{sip}(s,q) \quad \rightarrow$$
$$\forall s'\,((\mathsf{sip}(s',p) \wedge \mathsf{ontic}(s,s') \rightarrow \mathsf{sip}(s',q)) \wedge (\mathsf{sip}(s',q) \rightarrow \mathsf{ontic}(s,s')))) \tag{3.34}$$

- The player of a processual role necessarily comprises all substances which appear in the boundaries of the role.

$$\forall q \forall S\,(S \rightsquigarrow q \rightarrow \forall s\,(\mathsf{sip}(s,q) \rightarrow s :: S)) \tag{3.35}$$

- The PART-OF relationship is passed on from processual roles to processes.

$$\forall p \forall q \forall q'\,(\mathsf{ProcRole}(q) \wedge \mathsf{ProcRole}(q') \wedge q \succ p \quad \rightarrow$$
$$(q' < q \rightarrow \exists p'\,(q' \succ p' \wedge p' < p)) \quad \wedge$$
$$(q < q' \rightarrow \exists p'\,(q' \succ p' \wedge p < p)) \tag{3.36}$$

Before these axioms are discussed, it is worth mentioning that *substantial layers* in [Degen et al., 2003] exhibit similar characteristics compared to those of processual roles. However, substantial layers may contain several substances, in contrast to processual roles. Consequently, the notion of processual roles may be used to define substantial layers. But this is a tentative suggestion, as both notions need to be compared in more detail yet.

Axioms 3.31 to 3.33 are all concerned with the relationship between a single substance, a process, and a processual role of that process. Note that axiom 3.31 can be stated in greater detail than the corresponding axiom of relational roles, 3.13, because GOL already provides a connection between substances and processes, namely the containment of substances in process boundaries. Should this connection be changed in the development of GOL, several axioms would need to be rewritten, though the underlying claims could be maintained. Section 3.4.3 debates this issue more intensely. Axiom 3.33 may be understood as a complement of 3.31, because the latter demands a role for each substance, while the former guarantees that roles are limited to the process they belong to and that no substances can be found in a role's boundaries which do not also belong to the process.

It is a little problematic that the axiomatisation relies on notions which have been defined as yet only tentatively in GOL. This holds in particular for the notion of ontical connectedness. With respect to axiom 3.34 it is useful to assume that ontical connectedness is limited to what is called object identity in the object-oriented paradigm. Illustratively, this means that if a `table` ceases to exist due to being decomposed into pieces, it is assumed that there are no "successor" substances which are ontically connected to that `table`. In question may be here the remaining parts of that `table`, for instance a `table-leg`. Certainly, this does not imply that such parts could not be ontically connected to their counterparts when the table still existed. The mentioned restriction of ontical connectedness is not discussed in [Degen et al., 2003], but further support for it originates from axiom 3.35), which is elaborately discussed in section 3.4.3.

Furthermore, the relations between temporal parts of processes and of processual roles is a complex topic, which becomes even more difficult if parts of substances are taken into additionally. Axiom 3.36 deals only with a tiny aspect of this topic, which can be expressed in the framework introduced so far. Further, more lavish axioms regarding the required formalism are deliberately omitted, because these would require a disproportional amount of explanations with regard to other themes herein. However, some ideas on their possible character are compiled in the following section.

3.4.2 Relations to Temporal Parts

Let us start with a recollection of the differences between a temporal part and a processual role via the example of `Mary` `opening` a `door`. This `opening` process may take two seconds, in the first of which `Mary` moves her hand towards the handle, then taking the second one to bring the door into a fully opened state. The partition of the process into two subprocesses of a length of one second clearly is a temporal partition. With respect to the above description, `Mary` as well as the `door` need to occur in either temporal part. Contrariwise, the processual role `opener` (cf. p. 59) contains `Mary's` actions within the whole interval (i.e., CHRONOID), but `opener` ignores the `door` completely.

According to this example, one may postulate an axiom which mirrors axiom 3.36, i.e., it would state that for any part p' of a process p with a role q, one finds another role q' which is a part of q and a role of

p'. However, consider "large" processes[60] like a `conference` process. There may be a varying number of participants, for instance a `guest speaker` who arrives in the middle of the `conference` and leaves immediately after the presentation. Hence, near the end of the `conference`, one may discover temporal parts without a role which is played by that `guest speaker`.

There is much more that needs to be said about parts of processes and their roles, as well as players and their parts[61], but this is left as future work. Rather, the point to be made here is that processual roles are an important link connecting processes in general, in which several entities may partake, and single participants.

It is an important feature of processual roles that they do not require a descent in temporal part-whole granularity in order to explain how an abstract substance participates in a process. Hence, there is no need to look at parts of a process to identify in which way a an abstract substance is involved in it. However, this also holds from the opposite perspective. One cannot gain much knowledge about the temporal structure of a process if only its processual roles are determined. More axioms than the one given (axiom 3.36) will apply; nevertheless these cannot say much about individual processes or processes of more fine-grained universals. In spite of this, processual roles provide a means to bring to an end the regress which arises if one starts to explain processes by specifying their temporal parts.

The fact that processual roles do not *require* descriptions of temporal parts does not imply that there is no relation between roles and temporal parts. Processual roles allow for a convenient representation of who fills which position within different temporal parts of a process. Consider the case of `Mary walking`. It is not difficult to guess that it is `Mary` who walks in the first half of her `walking`, as well as in the second half, because there is only *one* participant (more precisely, only one processual role is to be played). Now, `Mary` might `open` the `door` and `walk` through the `door`, which yields a more complex process, call it `crossing`. In case one knew only that Mary and the door participate in the crossing, the opening, and the walking – why should Mary not open the door, but the door walk through Mary? All that is known is that Mary as well as the door participate. By adding processual roles and stating that the role `opener` is a part of the role played by Mary in `crossing`, it can be derived that Mary also plays the `opener` role. This is due to the "construction" of processual roles as referring to exactly one substance. Certainly, there are also other answers to the problem just presented, but processual roles seem to be a powerful solution which may help to understand other approaches. In an extended theory of processual roles, it may also be possible to follow the part-whole structure of substances. For example, `pressing` the `handle` may then be considered to be a subprocess of `opening`, the processual roles of which are played by the `handle` and `Mary's arm`. Again, processual roles should be capable of properly specifying who or what participates how, conforming to parts and wholes.

The last remark concerning parts of processes and roles refers to properties of processes. As processual roles are also considered to be processes, properties of processes may be "inherited" to the roles these processes consist of. A notion of *state* may serve as an example. The definition of this concept will be different from the GOL definition of a STATE; therefore *uniform* is chosen as its predicate name.

[60] This point touches on SITUOIDS, which may be the more adequate concept of GOL to represent such "large" processes. It is an open question whether the notion of processual roles can be extended to situoids, but as noted before, we will not discuss situoid-related issues herein.

[61] More precisely, as players of processual roles are abstract substances, the term *part* actually refers to the parts of the instances of those players. These parts themselves instantiate abstract substances, thus being identifiable at different times.

Definition 3.8 *A universal P of coherent, substantial processes is called* uniform *iff all temporal parts of instances of P instantiate P themselves.*

$$\forall P\,(uniform(P) \leftrightarrow_{def} P \longrightarrow\!\!\triangleright\, \mathscr{P} \wedge \forall p \forall p'\,(p :: P \wedge p' \leq p \wedge \mathsf{Process}(p') \rightarrow p' :: P)) \tag{3.37}$$

For example, the process universal `sitting` is uniform, while `running` is not an example, because one will find small parts of a running which themselves cannot be considered as instantiating `running` (e.g., a single step). Intuitively, for any uniform process universal *P* their should be uniform role universals. Formally, though, this cannot be shown on the basis of the axiomatisation provided thus far, but an axiomatic extension to allow for the proof is left as future work since this is a minor point in this thesis. What should be illustrated by the above definition is that there are properties concerning the structure of processes, which become reflected in the world of processual roles.

3.4.3 Players of Processual Roles

Players are a little problematic in the account of processual roles based on the current GOL notion of endurants, which is the reason for the weak formulation of the only axiom in this regard, axiom 3.35. First of all, note that playing an individual role should not be confused with the fact that several entities may play the same processual role, i.e., each entity plays its own individual role, but all roles instantiate the same role universal. Instead of this, we refer to the problem that there are several possible players for a single, individual role, because players are supposed to be ABSTRACT SUBSTANCES. This is owed to the possibility of subuniversals of abstract substances, since the latter themselves are universals. Put intuitively, it is not clear whether `John` is involved in some `game`, or `John` restricted to a part of his life, like `the child John`, participates in it. If both abstract substances are admitted, this results in having several players for one role. Compared to relational and social roles as well as to most of the role approaches from chapter 2, processual roles should also be restricted to one player only. However, corresponding general attempts lead to considerations similar to those in section 3.3.3, mainly dealing with the question of choosing the appropriate universal, regarding the position in an IS-A hierarchy.

Two of three approaches which may solve this problem appear inappropriate. The first approach comprises dropping the limitation of the existence of endurants to a single time boundary in GOL, thus admitting endurants which persist through time. This solution would require us to restate most of the axioms on processual roles and it would further cause many changes to already established parts of GOL. The second approach is less radical but counter-intuitive. It is constituted by a reinterpretation of the PLAYS relation, such that it connects individual substances and processes. In this case, processual roles would have one player per role at a time, but this is already expressed by axiom 3.32 and leads to the fact that a processual role viewed as extended over time contains a vast number of different players.

The third idea is a more convenient and successful approach. Abstract substances themselves may be restricted, i.e., only abstract substances of a special kind can play processual roles. The basic relation of ONTICAL CONNECTEDNESS may serve the purpose of restriction. An abstract substance may be said to belong to this new kind if it is closed under the relationship of ONTICAL CONNECTEDNESS[62]. More concretely, suppose that *endur*(`John`) (cf. section 3.1.2) is such a special abstract substance, i.e.,

[62]Parenthetically, this special kind of abstract substances very much resembles the object-oriented notion of an *object*. A

its comprises John at every time of his life. Then, all these instances are ontically connected. Now consider a subuniversal of *endur*(John), e.g., *child*(John), which selects only instances when John was a child. If a processual role *q* is played by the child John, either abstract substance is a player of *q* so far. However, only *endur*(John) is closed under ontical connectedness, i.e., there are no individuals which are ontically connected to instances of *endur*(John) without themselves being instances of this abstract substance. This does not hold for *child*(John), assuming it is a true subuniversal with less instances. Being marked in this way, *endur*(John) can be said to be *the* player of *q*. Note that this idea requires the assumption that ontical connectedness does not connect intuitively different substances, a feature which is not treated in [Degen et al., 2003]. If it is added to the present account of ontical connectedness, a stronger version of axiom 3.35 can be formulated according to the idea outlined, admitting only one player.

The uniqueness of players of a role is also used to justify the introduction of individual roles (in contrast to roles as pure universals) in connection with the *counting problem* (cf. also [Wieringa et al., 1994]). Imagine a bus company and its need to analyse its customers. On the one hand, the company is interested in the total number of people they transport, but, even more frequently, the total number of passengers is relevant. Assuming the method of a special kind of abstract substances as suggested above, what is counted in the first case refers to players of processual roles, while the second case is best understood as counting the processual roles themselves, independently of their players. RELATIONAL ROLES are not regarded as a better choice in order to understand the *counting problem* due to its intrinsic temporal nature.

At last, two interesting connections between abstract substances and processual roles shall be mentioned. Firstly, it is possible to use processual roles themselves to define subuniversals of abstract substances. Suppose there is an abstract substance *S* which plays a processual role *q*. Then one can identify all instances of *S* which occur in process boundaries of *q*, thereby referring to a "new" universal which is defined on the basis of *q*. Secondly, the *life* of an abstract substance *S* could be defined in terms of processual roles, namely as the mereological sum[63] of all processual roles that are ever played by *S* (where one should realise the danger of giving a cyclic definition). Both observations are linked to the fact that the temporal structure of the life of an abstract substance is reflected in the taxonomic structure of the abstract substance itself. It is an open question whether this desirable.

3.4.4 Formal Relations and Theta Roles

As a main application, processual roles are suggested as a substitute for several formal relations formerly employed in GOL as introduced in, for example, [Degen and Herre, 2001]. This work contains the formalisation of a kissing process which employs the formal relations *does* and *suffers* to connect the participants to the process. Indeed, this approach is now replaced within GOL by the notion of process boundaries which CONTAIN endurants. Nevertheless, this is a frequent approach and it is motivated by or at least closely related to the linguistic notion of theta roles (cf. section 2.3.2). We hold that approaches with a fixed number of formal relations will not be sufficient. This claim shall be discussed here, followed by an interpretation of theta roles.

GOL substance may then be understood as an object at a certain time. UML statecharts (which describe the life-cycle of an object) may then be explained as describing subuniversals of objects translated to GOL abstract substances.

[63] Mereology deals with the analysis of the PART-OF relationship. The mereological sum forms a new whole from several entities which are then considered parts of the new whole.

Assume that *does* and *suffers* comprise all formal relations that may connect substances to processes[64]. To illustrate our concerns, we reflect on a `carrying` process that involves three substances `Mary`, `John`, and a `box`. Let `Mary` and `John` be connected to the `carrying` by the formal relation *does*, while the `box` is linked by the relation *suffers*. Further assume that the ways `John` and `Mary` carry are remarkably different. Using the above approach, it is impossible to distinguish between their ways of carrying from the point of view of the `carrying` process. By the construction of this example, however, both are different even from this perspective.

We see three approaches that would allow one to maintain the use of formal relations to connect participants and processes. First, it may be declared that each process can only connect one substance per formal relation. Assuming a relatively small and fixed number of formal relations, this declaration will be unacceptable due to the variety of processes. Relaxing the restriction to an arbitrary number of formal relations leads to the second solution, where a formal relation could correspond to a processual role. This case may be acceptable with respect to representational power, but it ignores the processual character of these roles. This is an impediment, e.g., because formal relations do not stand in a PART-OF relationship. Furthermore, if processual roles are contained in the ontology in addition, the approach of one formal relation per processual role results in duplicate parts of the ontology.

The third approach refers to a descent in granularity, i.e., in the above example one would look for parts of the process where one could assign `John` and `Mary` to different formal relations. However, this seems much less elegant, as well as showing the disadvantage of containing a regress of unknown depth. That means, in order to explain the connection between substances and processes on one level of granularity, another such level is necessary, which in turn has to be explained by a third one, and so forth. As noted in section 3.4.2, processual roles cut this unfortunate chain at the first level, for they do not leave it.

Before a final decision in favour of processual roles can be made, it is necessary to explain why the argument of the indiscriminability of participations, which is offered against formal relations in the above example, cannot be applied against processual roles themselves. The reason is that each individual process consists of its own processual roles, i.e., processual roles are individuals each of which instantiates a variety of universals. Stating that two substances have “the same role” within a process simply means that two processual roles of this process instantiate the same universal. In this sense the substances could replace each other, from the process perspective. However, there may be subuniversals of that “same role” which allow one to distinguish these roles with regard to more specific universals. Formal relations, though, comprise tuples which are not supposed to provide further information than that two entities are connected in this way. The attempt to imitate the above argument on formal relations results in the mentioned duplication of the hierarchy of processual role universals by a subset hierarchy of formal relations. This is not intended for formal relations, which are better understood as providing a determined, elementary level of specificity.

The use of formal relations to interlink substances and processes very much resembles theta roles from linguistics. Even the names which are given to the concepts are very similar, as can be seen, for instance, with *Instrument* and *Agent* links in OPM (cf. section 2.2.4 and [Dori, 2002, ch. 5]) and the theta roles *Instrument* and *Agent* (cf. section 2.3.2 and [Parsons, 1990, p. 73 f.]). Generally, [Parsons, 1990, ch. 5] discusses a theory of thematic relations in greater detail and with a more linguistic focus, where he uses six formal relations. Above, the idea of employing formal relations for the connection

[64]For simplicity, the term *substance* is here used instead of abstract substance, and sometimes for their instances, as well.

of substances and processes has been rejected. Nevertheless, theta roles have an existential eligibility and can be interpreted in terms of processual roles.

In particular, thematic relations can be understood as abstractions of a variety of processual roles on a very high level. They form a top-level hierarchy for processual roles. This view is also partially supported by Sowa[65] (cf. Figure 2.2, p. 20). Note that this understanding as top-level categories is the *opposite* idea, compared to a use of theta roles as a basis to explain the semantics of arbitrary verbs (or other language elements). As argued above, following the latter idea does not allow one to express differences, for example, between `barking` and `singing`. Both verbs may be allocated the same argument structure, i.e., the same patterns of theta roles, which may be a single `Agent` role in the case of `barking` and `singing`. Thus, it is not possible to distinguish between finer types of processes structurally (at least without going down to lower granularities concerning the part-whole relationship). Transferred to ontology, this corresponds to building a top-level ontology, and then claiming that anything in the world is composed of *only* those top-level categories. This is different from what top-level ontologies actually do, which is to provide a collection of categories intended to be capable of comprising all other categories. But `barking` and `singing` can be introduced as new and different concepts of a domain ontology. Their assignment to a top-level category may lead to some implications, but more specific axioms can always be added. This is the main difference in the understanding of theta roles, on the one hand viewing theta roles as a number of building blocks for processes (which is the formal relations approach), on the other hand viewing them as top-level categories of processual roles.

Taking the top-level perspective, it becomes possible to explain the variety of collections of thematic relations which have been developed, as well as the disagreement on a standard set of theta roles in the linguistic literature. Each such collection can be understood as a distinct, high level of processual role universals. Of course, there is overlap between these collections, and the question arises whether there is *one* best collection for all needs, which also holds for top-level ontologies in general. In spite of these competing approaches, ontology and modelling can benefit from works on thematic relations, as they contain implications for lower level processual roles (cf. section 2.3.2), thus providing a (first) theory of universals of processual roles which has already been examined in some detail. In addition, theta roles are frequently used as an efficient means of specifying ways of participation in a process. It seems that for human understanding of conceptual models it is sufficient to specify such high-level processual roles. Hence, an easily manageable modelling language may provide a collection of theta roles as a fixed set of processual role universals, internally translating them in a generic manner into processual roles of the corresponding processes.

Concluding this section as it started, some problems of theta roles shall be enumerated. Note that we have collected problems of theta roles being used in ontology, in contrast to their original application, which is a theory explaining the interface between the syntax and the semantics of natural languages. Thus the points below do not refer to linguistic aspects of theta roles. Ontologically interpreted as processual role universals, these problems show some kinds of cases in which it may be insufficient to specify high-level universals only. In addition, some difficulties refer to other problems, mainly caused by the semantic differences of the linguistically syntactic category of *verb*.

[65] See section 4.2.3 for a discussion of Sowa's notion of roles.

1. There may be multiple occurrences of a theta role, which require a more specific resolution.
   - On the way to Paris John went to Brussels. (Paris and Brussels are *Goals*.)
2. The scope of a theta role is usually limited to one verb. But several sentences may describe one and the same process. This can lead to different assignments of theta roles according to actually one process. The `car` in the following example is such a case.
   - The car comes around the corner. Fred drives the car.
   - Fred drives his car around the corner.

   (The car is an *Agent* and an *Instrument* in the first sentences, only an *Instrument* in the second.)
3. Not every verb refers to a process. This has also been stated by Dori [2002, p. 112], who gives further counter-examples of verbs rather referring to relations than to processes.
   - The bottle contains some water.
4. Theta roles do not take part-whole relations or ontological categories of the arguments into account.
   - John pressed the handle down with his hand. (The hand is a part of John, and an *Instrument*.)
   - Mary writes about processes. (No distinction between individuals and universals.)
   - The flight lasts for two hours. (A process which is itself an argument of a verb.)

3.5 Social Roles

3.5.1 Introduction of Social Roles

What has been dealt with so far are role variants that do not involve taking on new qualities for the corresponding player. However, for most modelling approaches it is a main characteristic of roles that they exhibit their own properties[66] (cf. ch. 2.2.2 and [Steimann, 2000c]). It does not appear reasonable to extend the notions of relational and processual roles in a way that playing one of these roles in every case implies a certain number of new properties, which are added to the player. Instead, we analyse these issues in the line of what has been presented in the introduction of SOCIAL SUBSTANCES (see section 3.1.3), because substances are that category of entities in which MOMENTS can inhere, and which are interconnected by RELATORS. Of course, it can then be formulated that such a substance needs to be connected by a relator of a certain type, or that an abstract substance needs to play a certain processual role.

Three examples shall guide the examination of social roles and social entities in general, namely a `driver`, a `professor` of a `university`, and an `ant colony` with its `queen`. `Driver`, `professor`, as well as `queen` are often considered as roles of substances instantiating the universals `human` and `ant`, respectively. Again, `driver`, `professor`, and `queen` are all understood as

[66] The term *properties* as used here includes the object-oriented notions of (static) properties (i.e., attributes) and (dynamic) behaviour.

entities with their own properties. In contrast to other role types introduced so far, the `driver` is not to be conceived as somebody currently driving a car, i.e., viewing `driver` as a processual role is not the interpretation which is referred to in this section. Instead, (in many countries) one becomes a legal driver after having taken driving lessons and passing a driving test. Accordingly, we do also not refer to a professional driver like a taxi driver. A concept of somebody being "entitled to drive" may come closer to the intended concept.

The reference to exactly one player is a feature of a social role contrasted with social substances in general. The `university` is not considered to be a social role, but the `professor` is, because one human is the basis for this social substance. The existence of exactly one player is usually indisputable, or not even mentioned in the literature. The requirement of having a single player is not really a restriction, because social entities themselves may acquire social roles as there is no reason to exclude this option. If it seems that several entities form the basis of a social role q, it is rather the case that they themselves form a group, i.e., a new social substance, and this group (as a single entity) plays the role q. Conversely, the relationship between social roles and their players is much more often examined with respect to the player side. Descriptions like "An object can play several roles at a time." refer to these issues.

Conforming to the idea that contexts are more determining for roles than players, what could the context of a social role be? This question is more interesting for social roles than it has been for relational or processual roles. For the latter two, the context is often explicitly mentioned. For social roles, this is less frequently the case, and an explicit context is hardly found beyond agent-oriented approaches. Agent-oriented methodologies speak about agent societies, which form the contexts for agent roles. Interpreting the term *society* as a complex social substance, the idea from agent-oriented methodologies can be transferred to the more general notion of social roles, where *social substances* form the context of a role. This approach results in a multiplicity of contexts, i.e., there is no unique context with regard to a certain role. Consider the `professor` example – a professor may be seen as a role in a `university`, but also in a `department` of that university. Hierarchical structures of organisations suggest the need for several contexts for one role. The `ant queen` may also be associated with the `ant colony` as its context. In either case, the context is not itself a role.

The `driver` is more difficult to understand. As a first idea, one may assume that the driver provides her/his own context, as there is no obvious social entity he could be assigned to. But consider the ants, whose colony comes into existence without any declarations in natural language, i.e., nobody (or nothing) founds an ant colony by declaring its existence to others, giving an outline of the goals, a name and so on. The latter way of coming into existence is usually the case with named organisations. In contrast, the colony *emerges* from the interaction of individual, material substances. This argumentation can also be applied to the driver example. A driver will not exist in isolation; in a narrow sense, he interacts with others in other roles of a traffic system. Understood more comprehensively, a driver may be seen as a role in a legal system like a `country` (not viewed geographically), a `state` and so forth. Taking the opposite perspective, there are two ways for a driver to come into existence: either as a role of an emergent group due to interaction of the role players, or by declaration. A final possibility of a social role without a context is that somebody declares himself a driver, but does not tell anybody else. However, this could better be interpreted as the exotic formation of a group with one member, with that group again providing a context.

Admittedly, much of the above argumentation is based on intuition as well as on some literature presented in section 2.3.1. In particular, current notions of GOL neither support nor reject social entities. Nevertheless, a definition together with some axioms are given in accordance with what has been said. The utility of this concept is to be shown in the sections below, as well as in chapter 5.

Definition 3.9 (Social Role) *A social* SUBSTANCE *q is called* social role, *iff there is exactly one other substance s which constitutes q, and there is another social substance c such that q can be considered as a part of c.*

Axiomatisation

- Axioms on relationships between the universals *social role*, *role* and *social substance*.

$$\mathcal{Q}_{\mathsf{soc}} \longrightarrow\!\!\triangleright \mathcal{Q} \tag{3.38}$$

$$\mathcal{Q}_{\mathsf{soc}} \longrightarrow\!\!\triangleright \mathcal{S} \tag{3.39}$$

$$\forall q\,(\mathsf{SocRole}(q) \leftrightarrow q :: \mathcal{Q}_{\mathsf{soc}}) \tag{3.40}$$

- Social roles are dependent on (complex) social substances, and vice versa.

$$\forall q \forall s\,(q \succ s \rightarrow (\mathsf{SocRole}(q) \leftrightarrow \mathsf{SocSubst}(s))) \tag{3.41}$$

- Social roles do not have parts or roles.

$$\forall q\,(\mathsf{SocRole}(q) \rightarrow \neg\exists x\,(x \succ q \vee x < q)) \tag{3.42}$$

- A PART-OF relationship between two social substances corresponds to an inclusion of the social roles of the part by the whole.

$$\forall q \forall s \forall s'\,(\mathsf{SocRole}(q) \wedge q \succ s \wedge s < s' \rightarrow q \succ s') \tag{3.43}$$

- Social roles are played by substances, and there is exactly one player for each social role.

$$\forall q \forall x\,(\mathsf{SocRole}(q) \wedge x \rightsquigarrow q \rightarrow \mathsf{Substance}(x)) \tag{3.44}$$

$$\forall q \forall x \forall y\,(\mathsf{SocRole}(q) \wedge x \rightsquigarrow q \wedge y \rightsquigarrow q \rightarrow x = y)) \tag{3.45}$$

Axioms 3.42 and 3.43 partially express the idea that social roles form an "atomic layer" for certain social substances[67]. `Organisations`, `groups` (of people), `companies` and so forth are all considered to consist of social roles, each one played by a `human`. Of course, overlaps in the part-whole-structure of social substances are necessary, but two social roles cannot overlap, because of their atomicity. From the perspective of social substances on the basis of single, separate roles, the restriction expressed by axiom 3.42 appears intuitive. Consider the role `professor`, which is a role of a `university`. It is not clear what a role of a `professor` could be – which is not to be confused with the role `professor` itself. "Role of a `professor`" is meant in the sense that the `professor` is the *context* for another role. This is likewise not to be confused with a role *playing* a different role, an issue discussed in section 3.5.3.

The uniqueness of a playing substance is guaranteed by axioms 3.44 and 3.45. Uniqueness has been discussed above, but so far a powerful motivation for the limitation of social roles to substances is

[67]Not all social substances need to consist of roles. This issue is discussed in greater detail in section 3.5.5.

missing. Unfortunately, the understanding of social entities in general, and that of social roles in particular is in an early state of development. Much effort will be necessary to reach an acceptably comprehensive account of the distinction between the social and material halves of the world, and the interconnections between them. The following sections attempt to make first steps towards this account and sketch closely related, open issues. In addition, section 5.2 which is based on this preliminary account of social roles will demonstrate the need for roles which are substances in a real-world example.

3.5.2 Multiple Inheritance and Instantiation

Social roles are most often realised by object-oriented modelling approaches, though they are sometimes mixed with relational or processual roles. In many cases role classes (in an object-oriented reading) are introduced, with role objects which are connected to objects which are not considered to represent roles (cf. [Dahchour et al., 2002; Depke et al., 2000; Kendall, 1999; Kristensen and Østerbye, 1996]). However, the motivation for distinct role classes is not the presupposition of a level ontology, but an insufficiency of major object-oriented programming languages, namely the handling of multiple inheritance and multiple, dynamic instantiation. For example, [Fowler, 1997] advocates the view that roles are an auxiliary, implementational notion, rather than a conceptual one. Concerning the identity of a role object, a combination of an object and its role objects is often seen as one logical unit, and therefore role identity is rarely granted.

Role identity questions correspond to questions raised in philosophy. For example, "John is in a room. There are also at least a waiter, a driver, a father, and an employee in that room. How many are in that room? Exactly one human." (a similar example can be found in [Smith, 2003]). Philosophy also provides support for a real difference between, for instance, the universal `human` and the social role universal `driver`. According to ideas from Johansson and Searle (see ch. 2.3.1), which we have already expounded on (sections 3.1.3, 3.5.1), a driver really is some new entity, not identical to the human. If this is the case, examining identity issues may lead to constellations which are problematic for the multiple inheritance and instantiation position.

Indeed, there are direct arguments against the absorption of roles in multiple inheritance and instantiation. Consider a `human` having two `driving licenses`, hence playing two social roles of the type `driver`. There are two different `license numbers`, and two different `rights` to drive cars of a certain class of weight. Each one of those properties is attached to one `driver` role. Firstly, it is difficult to imagine that one individual instantiates a `universal` twice, but this is a requirement often demanded for roles. Secondly, even if a double instantiation of one universal is allowed, the connection between the single properties becomes problematic. Understood as MOMENTS, each moment would simply inhere in the `human`, thereby losing the link between the `license numbers` and the `rights`. If these are kept together in some way, this corresponds to the introduction of a social role by the back door.

3.5.3 The Importance of Social Roles

A main advantage of separating social roles from their players is to have a separate substance, which is a kind of a part of other social substances. Returning to the former `ant–queen` example, on the one hand, a `queen` seems to be a part of the social system of an `ant colony`. However, if the `queen` and the `ant` are identical, there is a well-known problem due to the usually claimed transitivity of PART-OF. Accordingly, a part of the ant, e.g., a `leg`, would also be a part of the ant colony. In contrast, this is not a problem with respect to social roles, because the PLAYS relationship mediates between a player and its role, i.e., between the `ant` and the `queen` as different entities. Thus, the chain of PART-OF is interrupted, because the ant is not part of the colony, but the queen is, while the queen does not have any parts (cf. axiom 3.42).

Put differently, an `ant colony` consists of `ant positions` (among which `queen` is one type of position), and it does not consist of material `ants`. This may seem counter-intuitive at first glance. In particular, there are biological changes within the single `ant`, in order to become an `ant` playing the social role of a `queen`. Nevertheless, we argue that social properties are meaningless on lower ontological levels, for instance the *function* (compare also section 4.1) to care for the continued existence of the `colony`, which is attributed to the `queen`; the ability to `produce eggs` would rather be assigned to the `ant` itself. In general, the question of which properties remain if an `ant` is separated from the others may help with the distinction between social and material properties. For instance, `producing eggs` seems still biologically possible.

Another effect of employing social roles is a high flexibility in modelling. Especially within `organisations` it may be the case that one and the same social role is played by several SUBSTANCES one after the other. More precisely, an abstract social role may exhibit different players at different time boundaries. It could be reasonable to assume such a diachronic identity of the role, across several players.

3.5.4 Role-playing Roles

This section deals with the case of a role itself playing a role, which is not restricted by any axiom so far. Several approaches suggest this view, as can be seen in [Steimann, 2000c]. However, according to our analysis, there is another understanding of the question of whether roles can play roles. This could also mean that the player of one role also has to play another role. In this case, axioms expressing this constraint are considered to be a better representation of such circumstances. A frequent example are roles which are said to be played by `employees`, like the role of a `project leader`. More formally, the question is which of the following formulae is more appropriate, the first one about role-playing roles, or the second one with two roles enjoying equal rights with regard to a common player. (For clarity, we concentrate on the PLAYS relationship only.)

$$\forall x \forall q\,(x \rightsquigarrow q \wedge q :: \texttt{Project\_leader} \rightarrow x :: \texttt{Employee} \wedge \exists y\,(y \rightsquigarrow x \wedge y :: \texttt{Human})) \tag{3.46}$$

$$\forall x \forall q\,(x \rightsquigarrow q \wedge q :: \texttt{Project\_leader} \rightarrow \exists q'\,(x \rightsquigarrow q' \wedge q' :: \texttt{Employee})) \tag{3.47}$$

The first variant leads to a hierarchy of the PLAYS relationship. This looks appealing, because it is quite close to an IS-A hierarchy. The following situation can also be interpreted assuming either axiom. Consider `employees` and `freelancers` being allowed to play the role of a `project leader`. If a project leader, formerly a freelancer, becomes an employee, he may keep the very same project leader role. In the case of formula 3.47 there is no problem, because both roles are only indirectly linked. Representing this with role-playing roles (the first variant), the project leader role may remain identical, but it is played by different substances one after another. Otherwise, the project leader role ceases to exist and a new one needs to be set up. In conclusion, both possibilities remain admissible.

A question associated with role-playing roles results from the following example. Imagine a `human` which is `employed` as a `professor` at some university. Now consider him giving a `lecture`. First one may think of the material basis, i.e., there is one human `speaking` in front of several others. Now, this speaking *counts as* giving a `lecture`, thus referring to a social process. On the other hand, something must *count as* `working` of the `employee`. Similarly to the question of admitting role-playing roles, one may ask what counts as working: giving a lecture, i.e., this social process, or the material process of speaking. Intuitively, we consider the first option as more appropriate, because a professor is usually employed, among other reasons, to give lectures.

This example provides at least a little inkling which speaks against role-playing roles. If the latter are assumed, the professor seems to play the role of an employee when giving a lecture, instead of the employee playing the role of a professor. But the latter is the sequence which would correspond to the above example of `employee` and `project leader`. Remembering the problems arising from imposing an order on the arguments of relations where there is actually no order, formula 3.47 is considered more appealing.

3.5.5 Encountering Social Entities

Let us seize the last example of `professors` and `employees` of the previous section in order to point to some even more fundamental problems than the question of role-playing roles. In that example, the *counts as* relationship, which can be traced back to Searle (see section 2.3.1), is employed to relate social to material processes. For social roles, it is exactly this relationship which is mirrored by the PLAYS relationship. Indeed, section 3.1.3 introduces not only social substances, but social entities in general, comprising processes, moments, relators, and possibly others. This leads to the idea that these entities may also be considered to be roles, played by their material counterparts.

In order to illustrate this issue, a `chess game` shall be examined. Several entities are declared social roles, like `king`, `queen`, `bishop`, `knight`, and so forth. All of those are granted certain rights and powers. But what can be said about social actions like a `move`, which also have to be implemented on the material level. “Certain ways of moving something which plays the role of a chessman across something playing the role of a chessboard play the role of regular `moves` in a `chess game`.” sounds natural from a language point of view. According to the main features of roles, *player* and *context*, the player is clearly identifiable as the material process of moving a piece of wood. The context is somewhat harder to determine. First of all, it cannot be a complex social substance as is postulated for social roles, because an ENDURANT which exists at a single TIME BOUNDARY cannot provide a context for a process which is extended in time. Another option is the overall `game` as a social process. Of course, a move is also a part of the game, but the same applies for social roles and their contexts. However, consider the `game` itself, which is also “played by” some material process. Should the game be its own context, i.e., be a role of itself?

This question will not be tackled herein, as there are many other issues with regard to social processes which require a comprehensive treatment. One such issue is the discreteness of many social processes. For instance, there is not a half of a `move`. Therefore, a social `move` might be considered as a

CHANGE, i.e., an immediate transition between two situations, while the underlying material process is continuous. Clearly, there is a strong proximity to formal process descriptions with Petri nets, UML statecharts, et cetera. These should also be taken into consideration when dealing with this subject, but this would be far beyond the scope of this thesis. We expect that social roles as introduced above can be maintained, or that they are derivable from an integrated approach which might cover social entities in general.

Lastly, note the difference between social entities (including social substances) which are constituted by a number of roles, and those that are merely represented. This has already been described in the literature (cf. [Smith, 2003]). Examples of easily conceivable, representational entities are virtual entities occurring within the Internet. For instance, a `virtual marketplace` exists because of a representation, e.g., in several databases, but it does not have social roles because it is not played by anything. This is a different mode of the existence of social entities, which requires some other formal relations, for example, expressing *denotation* or *representation*. It is possible that this access to social entities will provide answers to the above questions with respect to social processes, thereby providing evidence for their rejection as roles.

3.6 Summary of Role Classification

This section briefly sketches the theory of roles developed in this chapter. We start with a summary of the main issues that have been treated. Then, common features of all role types are identified, and some remarks on the classification will be made. Finally, interconnections of relational, processual, and social roles will be discussed.

3.6.1 Results and Issues

In section 3.2, two main features of roles were identified which occur throughout the literature: *player* and *context*. The relation between players and roles has often been discussed, in contrast to that between roles and contexts. We have classified roles according to three major types of contexts identified on the basis of chapter 2. Contexts were chosen as the classification criterion, because they have a determining influence on roles. The differences between the axiomatisations provided for all three role types strongly support this approach of role classification.

Basic characterisations of these role types are introduced in terms of two new FORMAL RELATIONS: PLAYS ($\rightsquigarrow$) and ROLE-OF ($\succ$). PLAYS relates a player and a role, and ROLE-OF a role and its contexts. In addition, specific issues are discussed in each section. Table 3.3 on page 74 may serve as an overview, summarising the key notions and results.

Thus far, all three role types have been analysed in isolation, i.e., mainly without comparisons between each other. It remains for future work to transfer insights from one role type to others, where applicable. In particular, the discussion of relator and role universals in section 3.3.3 is vitally important for all role types, thus a possibly extended version should refer to roles in general. Seizing the idea of general features of roles, let us now look at commonalities of the role types examined, which can already be extracted from the existing analysis.

Topic	Section	Description
Relational Roles		
• Representation of MATERIAL RELATIONS in predicate logic	3.3.2	RELATORS and relational roles solve the *Instantiation*, the *Extensionality*, and the *Arrangement problems*.
• Relator and role universals	3.3.3	A general structural restriction for role universals and context universals cannot be given on the basis of the relations PLAYS and ROLE-OF, because they relate only individuals.
• Symmetry of relations	3.3.4	The comparison of *extensional* and *intensional symmetry* of relations results in the discovery of *extensionally symmetric*, *intensionally asymmetric* relators. These demonstrate a fine-grained analysis of relations by relational roles.
• Representation of anadic relations	3.3.5	Relational roles allow a simple representation of anadic relations, which is more efficient than the common use of predicate logic.
• Uniquely determined relators for relational roles	3.3.6	It remains an open question whether a unique relator can be assigned to each relational role.
Processual Roles		
• Temporal parts and processual roles	3.4.2	Processual roles are closely related to temporal parts, although they allow for the description of processes without resorting to temporal parts. Part-whole structures of a process are mirrored in part-whole structures of its roles.
• Players of processual roles	3.4.3	The notion of ENDURANT in GOL creates some specific requirements on the players of processual roles. In other frameworks this might be handled more easily.
• Use of formal relations to express participation in processes	3.4.4	The use of a small, fixed number of formal relations is rejected in favour of the notion of processual roles. In this connection, formal relations resemble *theta roles* from linguistics, which are interpreted as high-level processual role universals.
Social Roles		
• Multiple inheritance and instantiation	3.5.2	These notions cannot replace (social) roles, though this is often claimed in the literature.
• Advantages of social roles	3.5.3	Social roles solve a problem which is caused by the transitivity of PART-OF, namely that a part of a `human` is considered a part of some `group` because the `human` is a member thereof.
• Role-playing of roles	3.5.4	The question of whether social roles may themselves play social roles remains open, but we suggest that the evidence points in the direction of a negative answer.
• Roles and other social entities	3.5.5	Social roles may not be the only kind of roles in the realm of social entities, but for other social entities, there are too many open issues to be contemplated herein.

Table 3.3: Summary of Topics and Results in Role-Specific Sections.

3.6.2 Common Features of Roles

Besides the predetermined notions of *players* and *contexts*, there are a few features which can be assigned to all roles. We think this is the case because roles are a very abstract and thus flexible notion, which is likewise suggested by the frequent and flexible use of phrases like "plays a role of" and "plays an important role" in natural language. Of course, these phrases are not directly of ontological importance and should not be understood as such. Rather, our analysis of roles results in a widely non-uniform appearance of different role types. In the case of roles, these differences between role types correspond to a flexibility in language. According to this, they may be compared to the notion of *object* in object-orientation, which is used likewise flexibly.

Another fact is that roles are orthogonal to the established GOL hierarchy. That is, role individuals can be found in several disjoint categories of GOL, like SUBSTANCES and PROCESSES. In this respect they can be compared to the notion of PARTS. Indeed, there is even more support for a similarity of roles and parts. In order to avoid a cyclic definition, the term *role* could appropriately be replaced with the term *part* in all three intuitive definitions of relational, processual, and social roles. Still, these definitions sound natural, which is seen as support for the close relationship between roles and parts. Hypothetically, roles may be considered as a specialisation of parts, with an additional reference to a player. If this is assumed, the ROLE-OF relationship should be interpreted as PART-OF. However, firstly this requires a very extended notion of parts, and secondly, the two relationships are by no means equivalent, as PART-OF includes many other cases, like a part of a `body` or, in the social world, an `institute` as a part of a `university` which is not a role. Again, a more detailed consideration, e.g., of the compatibility with the axiomatisation of PART-OF, needs to be postponed.

Instead, one may choose the PLAYS relationship to initiate a comparison of roles to existing GOL entities. As every role depends on a player, at first glance this suggests a strong similarity with MOMENTS. This holds in particular for the case that a player is a substance, which seems to be the majority of all cases, by far. However, the analogy between a role and a moment is not appropriate, because, e.g., processual roles are not even ENDURANTS, and social roles are substances, which are disjoint from moments. What appears more reasonable in this connection is a search for different formal relationships replacing PLAYS, which apply for one of our role types. For processual roles, the formal relation of PARTICIPATION may be taken into account, which was very recently added to GOL and is in a early state of development. Hypothetically, in the case of relational roles, one could choose INHERENCE for an analogous purpose.

The following points can be summarised to hold for all roles:

- Each role is an INDIVIDUAL.
- Each role has exactly one player[68].
- There is at least one context for each role.
- If there are multiple contexts,
  - all of these belong to exactly one of the categories RELATOR, PROCESS, or SUBSTANCE thus far, but there may be more context-providing categories.
  - these contexts are in some way related by PART-OF.
- No role is a context of *another* role.

[68]Formally, processual roles cannot be said to have one player only (yet). However, an extension of GOL should account for this (see section 3.4.3).

Obviously, this collection of features applicable to all roles is rather abstract. It does not provide as many details as other role approaches, in particular with respect to the PLAYS relationship, as for instance in [Steimann, 2000c]. On the one hand, this is owed to the fact of providing a more comprehensive notion of roles, integrating seeming contradictions in the literature by separating specificities from a common, general *core*. The items above describe this core. Section 4.2 will demonstrate how major approaches can be interpreted in our role approach. On the other hand, more specific issues have been split into the previous sections. The applicability and utility of this separation is shown also in section 4.2. For instance, some properties of roles as collected in [Steimann, 2000c] can be restricted to certain role types, while they are not reasonable for others.

At this point, notice that the formalisations we have presented are, of course, strongly influenced by the GOL framework, which is itself still in development (cf. chapter 1), such that some aspects (cf. footnote 68) may be differently (and maybe more simply) formalised with respect to later GOL versions. Furthermore, it is clear from the above items that the classification provided so far cannot be considered as comprehensively exhaustive, but our subdivision aimed at capturing what was found in the literature. According to additional role types, one option is mentioned at the end of section 3.5.5, which may, however, be an extension of social roles instead. In general, new role types should be introduced with care, in particular with respect to the question whether there is an overlap with existing ones. This is a difficult question, indeed, a claim which is substantiated in the following section.

3.6.3 Interconnections between Role Types

The examples within the single sections on relational, processual, and social roles show that there are often difficulties in assigning natural language terms to role types. Take the so-called prototypical role, `student`, as an example. To which role type does this term refer? It is common knowledge, that analysis in computer science is strongly influenced by the purpose of the analysis, since abstractions are necessary in order to maintain the tractability of one or several models (cf. [Booch et al., 1999]). Accordingly, a `student` may be modelled as a relational role in one model, as a processual role in another, and as a social role in a third one. Clearly, there is a strong connection between all of these roles and models, from which the question arises whether one or the other role type is more basic.

First of all, in certain cases processual roles can be considered to be more basic than relational roles. This holds if a RELATOR simply *reflects* a PROCESS, because a process cannot be modelled directly in a language, as in ER modelling (cf. section 2.2.1). However, there are irreducible relators. These may be FOUNDED on processes, like an instance of the `mother-Of` relator universal which still holds after the corresponding process instantiating `birth` (assumed to be its foundation) is completed. Otherwise, an irreducible relator may not be directly associated with processes, as instances of the `contains` relator universal, which may mediate between instances of the substance universals `nail` and `box`.

Moreover, processual roles and processes are important in connection with social roles, because societies rely on interactions, i.e., processes, of their members. In addition, social entities depend on non-social entities playing certain processual roles from which social processes and thus societies emerge. Again, the understanding of social entities is certainly not very established yet, in particular as compared to the complexity of human societies which have had a long development up to now.

Further, other philosophical notions like *intentionality* will have to be analysed and integrated for a better understanding of social entities. Thus far, however, one can conclude that processual roles are of a very basic character, for social roles as well as for many relational roles, as explained in the last paragraph.

If we return to the `student` example, the answer to the question of which role type is the most appropriate is: none of them is, but all three role types are concepts of equal validity. A relational role universal `student` has different features, i.e., a different intension, than the processual or the social role universals, respectively, which may be denoted by `student`. Although it seems necessary for a social roles instantiating the social `student` universal to play a relational role as well as a processual role instantiating the corresponding relational and processual `student` universals, all three are different universals. Accordingly, the natural language term `student` conveys a several meanings, where the relevant one is unravelled by humans in terms of several factors, the examination of which is a part of linguistics. However, in language it is possible to switch easily between these meanings. What remains for a powerful modelling formalism, from our point of view, is to provide similarly smooth transitions between such closely related entities. GOL as well as UML contain approaches in this direction, which will be mentioned in sections 5.1 and 4.2.1, respectively.

4 Discussion

Two major issues shall be discussed in this chapter. First, the notion of a role is separated from other concepts which are sometimes presented as synonyms for the term *role* in the literature. Certainly, the analysis of these concepts cannot exhibit the same depth as in chapter 3, but some informative examples may serve the purpose of revealing differences. Secondly, the broad coverage of the theory introduced in the previous chapter is to be demonstrated by a reconsideration of some works introduced in chapter 2. This re-examination is intended to show the occurrences of mixtures of the identified role types, as well as where such mixtures have produced inappropriate restrictions according to our view.

4.1 Comparison to Other Notions

4.1.1 Function

An *intuitive* notion of *function* is the most favoured candidate of synonymy for roles. Notice that *function* does not refer to any mathematical meaning here, or to any philosophical account of that term (an example of the latter can be found in [Johansson, 1989]). Instead, many works in particular in the field of conceptual modelling use an intuitive understanding of role and function, thereby viewing them as equivalent. That this is common in natural language is also illustrated in chapter 5.2, where a more practical application of our role classification is presented.

However, there are differences between functions and roles. Let us start with the feature that a function can be *executed* or *fulfilled*, respectively. This means that there are PROCESSES in which the entity which has the function participates. These processes are considered to be the executions of the function. An example might be a `heart` and a `heart-function`. An individual heart fulfills the function of a heart if there exists an individual process within which the heart beats and blood is pumped through the blood vessels. However, even if the individual heart stops beating, the function of *a* heart remains. It seems that the term function actually refers to a process universal, such that the executions correspond to instances of that universal. On the other hand, a function is a feature *of* something. At this point, the question arises which kind of entities can have functions. We advocate universals of ABSTRACT SUBSTANCES to be this kind of entity (for simplicity, hereafter called substances throughout this section). Accordingly, the term *function* refers to a relationship between the substance universal `Heart` and the process universal `Beating`. If, instead of this, `Beating` was related to an individual `heart` as dependent entity, like a MOMENT, the `heart` would participate in the execution of one of its dependent features. This sounds unnatural, not at least because the `heart` itself is dependent on the process of the execution (cf. section 3.1.2), and therefore *function* is proposed as a relationship between universals.

More specifically, process universals which are functions for certain substance universals are universals of processual roles. This explains why they can be assigned to a single substance universal. Of course, the instances of these processual roles cannot appear without complementing roles such that the whole yields a process of several participants. The function, though, is executed in the form of a processual role. Switching to another kind of function, take the social substance universal $\texttt{Teacher}_{soc}$ as an example. Its function is to teach, i.e., the function is the processual role universal $\texttt{Teacher}_{proc}$. Of course, instances of $\texttt{Student}_{proc}$ are also necessary for a complete instance of the process universal `Teaching`, but $\texttt{Student}_{proc}$ does not belong to the function, because the activities of the students are clearly separated from the activities of the teacher. In this connection, another observation can be made. Functions refer to an agentive or instrument processual role (cf. the discussion on theta roles in section 3.4.4). The last example supports this view, as well as the `Heart` example. The functions `Teacher`$_{proc}$ and that assigned to `Heart` are subuniversals of the general processual role universal `Agent`. Other processual roles in these processes, like `Student`$_{proc}$, are not considered to be functions and have other theta roles assigned to them, with regard to the substance universal at issue.

Now one can start to ask for reasons for the commonly assumed equivalence of role and function. We think that this is mainly a restricted interpretation which is combined with the idea of social roles. The interpretation of functions in many approaches is limited to what may be called *social functions*. The `Teacher` example belongs to this category. But there are other types of functions which would cannot roles, such as *natural* functions (assigned to substance universals like `Heart` or `Lung`) or *instrumental* functions (such as those of `Hammers`, `Switches` and so forth). The problem with social functions is that there are universals of social roles which are exclusively or mainly characterised by a function. The social substance universal `Teacher`$_{soc}$, understood in a very broad sense, is only characterised by the function `Teacher`$_{proc}$. In this way, functions provide a transition from processual to social roles, and therefore it is tempting to confuse roles and functions.

4.1.2 State

If states are considered, two different interpretations have to be distinguished, which are called the *temporal* and the *atemporal* reading of state, respectively. The first version is to consider states as special OCCURRENTS, as is the case for GOL's STATES. Intuitively, temporal states are PROCESSES in which, from a certain point of view, no changes happen. On the other hand, the concept of an atemporal state is often understood as a condition, a view which corresponds to FACTS or SITUATIONS in GOL. In the latter case, states are either said to have no temporal aspects, or such aspects (like that of how long a fact holds) are neglected. Further, there is not necessarily a reference to processes in atemporal states, but MOMENTS may serve as their basis. With regard to an example from OPM (cf. section 2.2.4 and [Dori, 2002]), where states describe situations in which an object can be, a `car` may be in one of the states `new`, `used`, or `scrapped`.

A confusion of roles with states is less frequent than that of roles with functions. In particular, roles are not viewed as states, but states are sometimes thought of as roles. Viewing states as roles mainly originates from a synonymous use of certain words in natural language. A common example is the word `child`. On the one hand, it may refer to a relational role with the complementary role `parent`. On the other hand, a human is considered to be a `child` during a certain stage of life. However, in this case `child` is understood as a subuniversal of `human`, with regard to ABSTRACT SUBSTANCES. Instances of `child` are ENDURANTS, not OCCURRENTS.

Minor sources of confusion are superficial linguistic criteria, as well as implicit references to social roles. Linguistic criteria are quite popular as a means of distinguishing role universals from other types of universals (cf. [Guarino, 1992; Sowa, 2000]). For instance, relational roles appear often as opposites, such as `student` and `teacher`. But it would clearly be wrong to derive a rule from these cases which suggests that all opposites are (relational) roles. The pair `adult` and `child` may serve as an example. With regard to social roles, there are cases where certain terms can be interpreted as social roles or as states. In the literature one finds [Tamma, 2001, p. 127], where `teenager` is considered to be a role just as `student`. This may be appropriate, if `teenager` is considered to be a role in society, e.g., if it is described by expectations and norms (thereby being subject to *role theory* of the social sciences, cf. section 2.3.3). However, the term `teenager` may equally well refer to a human in a state. Note that instances of the latter can also be described in more detail, for example, with biological properties.

Finally, a short remark shall be made on [Marcos and Cavero, 2002], which refers to the atemporal reading of state. This work examines IS-A hierarchies, inheritance, role, and *state hierarchies* of (object-oriented) classes. In [Marcos and Cavero, 2002], subclasses of a class within a state hierarchy are required to be pairwise disjoint. In contrast, role hierarchies are said not to fulfill this requirement. According to Marcos and Cavero, this criterion may be used to distinguish between roles and states. However, we suspect that there are subclasses with regard to states which do not have disjoint extensions, especially in the atemporal reading of state. This suspicion is substantiated by the fact that states may refer to completely unrelated properties, each of which may be exhibited by an entity. An illustrative example adapted from [Marcos and Cavero, 2002] refers to the states `single` and `married` of persons, which can be arbitrarily combined with the states `tall` and `small`. In role hierarchies as well as in state hierarchies, subclasses can be chosen to be disjoint or overlapping. Thus the disjointness of subclasses is no criterion to distinguish between atemporal states and roles.

4.1.3 Quality

The fact that roles refer to exactly one player causes people to understand them as qualities of their players. However, the category QUALITY in GOL is a subcategory of MOMENT, and therefore it refers to intrinsic properties of SUBSTANCES. This means that quality instances have no reference to other INDIVIDUALS apart from the substance where they INHERE. That this is not appropriate for roles has already been mentioned in section 3.6. Now we return to this issue regarding social roles to provide more detailed reasons. For processual roles it should already be clear that they cannot be viewed as qualities, since they are OCCURRENTS, but qualities are ENDURANTS. Relational roles may be acceptable as MOMENTS, in which case the PLAYS relationship would collapse with INHERENCE for relational roles, but cf. also section 5.1 for that.

Social roles are two-edged entities: on the one hand they are themselves SUBSTANCES, on the other hand they depend on another substance which is their player. Hence, the question arises whether they should be considered as compound qualities or moments, i.e., moments in which other moments can inhere. In general, GOL does not explicitly ban moments inhering in other moments. Substances are defined as entities which themselves do not inhere in any other entity. So, why should not the assignment of social roles to substances be reconsidered in favour of moments? Firstly, in this case entities of one ontological level would inhere in substances on a different level. Intuitively, we disagree with a mixture of these levels. A second reason is that social roles are parts of other social substances, and a quality is usually not considered to be a part, not even for social objects. The final argument alludes to *role transfer*, i.e., the playing of one and the same role by different substances one after another. If role transfer is to be accepted, one either has to support the substance view of social roles, or the *non-migration principle* for moments needs to be dropped, because it demands that each moment is bound to exactly one substance. The latter option is contrary to GOL, in which the non-migration principle is postulated as an axiom. Thus one should agree with social roles as substances, not at least for a proper integration of our approach into GOL.

4.1.4 View and Interface

Borrowing mainly from object-oriented analysis and database modelling, we understand views as subsets of attributes of classes, while interfaces are subsets of operations of a class[69]. However, both notions belong to "implementational devices" in the context of computer science, i.e., they originate from implementation issues rather than questions of conceptual modelling. The distinction between *inheritance* and taxonomies as discussed in [Marcos and Cavero, 2002] is an appropriate analogue. Inheritance is motivated by the goal of code reuse in programming, while taxonomies belong to the field of conceptual modelling. Taxonomically arranged classes should support inheritance, but inheritance may be applied to classes which do not stand in an IS-A relationship. Views and interfaces can be understood in the same way as inheritance, whereas it is not clear what their counterparts in conceptual modelling are.

It is certainly possible to use roles, in particular social roles, in order to gain views on objects. This is also mentioned in [Oestereich, 1998, p. 175], where roles are understood as a view of an object[70]. [Coulondre and Libourel, 2002] seem to follow the same idea, as they compose classes from so-called roles, which seems to refer to views, actually. However, there is no one-to-one correspondence between roles and views or interfaces. Roles are ontologically restricted, but it seems to be an important feature of views and interfaces for programming, that attributes and/or operations can be *arbitrarily* composed. In addition, there may be other ontological categories which are implemented by means of views/interfaces.

In contrast to the just-mentioned distinction between implementational and modelling concepts, there are approaches trying to merge the notion of interfaces with that of roles. [Steimann, 2001] is a good example for that in UML, which is based on an analysis of roles. One conclusion which is drawn in this paper is that roles provide a "meaningful conceptual representation" for interfaces, and that merging roles and interfaces can reduce the number of modelling primitives. Further reorganisations of the UML meta-model have also found their way into [Steimann, 2001], which therefore retains its interesting character even if the equivalence of interfaces and roles is rejected herein. With regard to this amalgamation of concepts, we still support the idea of keeping modelling and implementational notions separated. Again, there may be other conceptual entities than roles giving rise to the use of interfaces in implementation.

4.2 Critiques

4.2.1 Unified Modeling Language (UML)

It is important to understand UML role concepts, because of the quasi-standard status of UML for modelling (above all in software engineering), and because it offers a variety of roles. Section 2.2.2 has exposed several UML modelling elements which, intuitively, appeared to be related to roles.

[69] There are more general approaches which allow for attributes in interfaces, in addition to operations, but this does not influence our argumentation.

[70] Interestingly, Oestereich suggests that roles, i.e., views should be understood as a property of some observer of the entity playing a role, instead of a property of the goal of the observation itself.

In short, there are *rolenames* and *collaboration roles*, the latter being divided into *classifier* and *association roles*, which explicitly mention the term role. There are even more implicitly related notions, namely *actors* in use cases, *qualifiers* in associations, as well as the methodological hint to use *interfaces* to model roles as types. In spite of this variety, there are numerous works which use a new notation combined with conventional UML models (cf. [Dahchour et al., 2002] or Agent-UML adaptations as in [Cranefield et al., 2002; Depke et al., 2001]), or which declare the integration of a different, explicit role concept into the UML meta-model (cf. [Depke et al., 2000; Reenskaug, 2001; Steimann, 2000d, 2001]). The present section maps UML role concepts to the role types presented herein, and discusses the properties of UML notions of roles. This paves the way for future research, like an inter-diagram comparison of modelling elements within UML, or an assessment of the innovation of various supplementary suggestions for UML, with regard to roles.

4.2.1.1 Relational Roles and Rolenames

Let us start with relational roles. *Rolenames* are the only UML element regarding roles which resemble relational role universals (hereafter, relational roles for short, which can be clearly resolved due to class diagrams referring to a universal level), provided associations correspond to RELATOR UNIVERSALS[71]. More precisely, we assume that associations can be understood as MATERIAL RELATIONS, which is supported by the fact that FORMAL RELATIONS like PART-OF and IS-A are granted special symbols within UML. Of course, rolenames are not completely equivalent to relational roles. There are three main problems which cause less expressiveness of rolenames compared to relational roles. First, the extension of an association is a set of tuples of object references. The second point is that UML cannot deal with anadic relations. Thirdly and finally, several restrictions are imposed on rolenames, which are unnecessary from the point of view of conceptually modelling relations.

With regard to the first problem, the set-theoretic interpretation of associations is disadvantageous, because two of the problems discussed for predicate logic in section 3.3.2 are introduced to the UML by this interpretation, namely the *Instantiation Problem* and the *Arrangement Problem*. According to the Arrangement Problem, this is even contrary to the fact that UML is a graphical formalism. Graphically, no order is imposed on an association with specified rolenames. However, an order is required in the definition of an association, in order to allow for the set-theoretic interpretation. The *Instantiation Problem* seems to be well-known to the authors of [Rumbaugh et al., 1999], but a proposal of a solution for this problem differs from employing relational roles. The suggestion of Rumbaugh et al. is to *reify* an association in order to "model a relationship that is a *bag* rather than a *set*" [Rumbaugh et al., 1999, p. 159, emphasis added], i.e., to create a class which then represents the relationship, instead of a UML association. Adversely, this leads to the situation that one can no longer assume a correspondence between UML associations and semantic relations, i.e., some relations may not be represented as associations, which in a sense contradicts the introduction of associations at all, if they can equally well be modelled as classes. What remains is the fact that each association models a semantic relation.

Remembering the third problem from section 3.3.2, the *Extensionality Problem* is not as difficult for UML, because so far UML is a mere representation language, i.e., without any reasoning support.

[71] A similar mapping, namely of associations to relations, has already been pointed out in [Guizzardi et al., 2002a]. At that time, the notion of relators had not yet been established.

Within a UML model, an association can be identified, e.g., by its name, such that associations which have the same set as their interpretation may be distinguished by name.

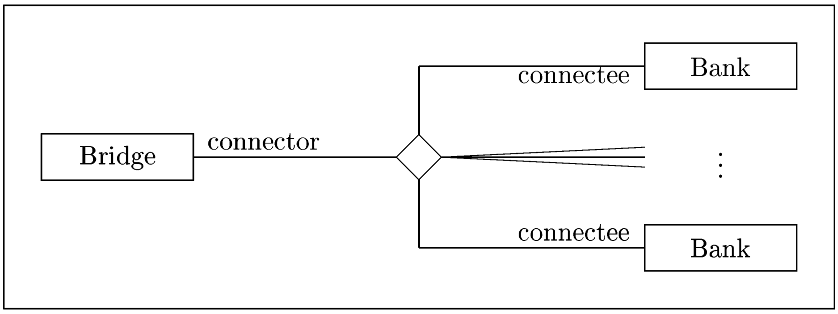

Figure 4.1: Improper Use of UML: Modelling Anadic Relations.

Another problem is that anadic relations cannot be represented in UML. We will not repeat section 3.3.5 here, but the reader may remember the `connection` example, where a `bridge` is related to a varying number of river `banks`. A UML-like representation for this example is given in Figure 4.1. Note that the standard UML model in Figure 4.2 cannot express all the subtleties contained in Figure 4.1, due to the fact that Figure 4.2 actually refers to a *different* relationship. It focuses on the relationship between a bridge and a river bank, rather than the relationship between one bridge and *all* river banks which that bridge connects. This difference leads to problems with the reduction of relations of higher arity to relations of lower arity, which cannot be solved without the introduction of entities different from `bridge`, `bank`, and the association itself, namely relational roles and relational role universals (cf. sections 3.3.5 and 3.3.6).

Figure 4.2: Standard Use of UML: No Substitute for Anadic Relations.

Concentrating on problematic requirements for rolenames, the denial of anadic relations is also contained implicitly in the condition requiring different rolenames for each association end within a single association (cf. p. 27, the second quote from [Rumbaugh et al., 1999, p. 414]). It becomes clear from the UML documentation that this requirement is owed to the fact that UML is usable in the analysis *and* the design stages of software engineering. In design models, rolenames are to be used for navigation between objects, which demands the uniqueness of rolenames according to [Rumbaugh et al., 1999]. However, from an analysis point of view, this restriction is unnecessary and should be dropped since there are deficiencies, e.g., *intensionally symmetric* relator universals (cf. section 3.3.4) cannot be expressed in terms of associations and rolenames. Accordingly, it seems desirable to allow for less restrictive analysis models, which may then be translated (supported by some CASE tool[72]) into design models which pay attention to such restrictions.

[72] CASE stands for Computer Aided Software Engineering.

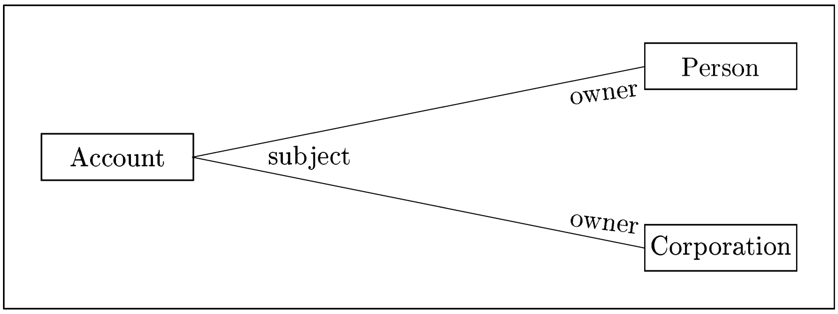

Figure 4.3: Multiple UML Associations with the Same Intended Meaning.

The navigability imposed on rolenames also requires that these are different across different associations, if the latter are attached to one and the same class with one of their ends[73]. Therefore, Figure 4.3 does not conform to UML. This is an adapted example borrowed from [Rumbaugh et al., 1999, p. 156]. In the original, the rolename `owner` is replaced by the rolenames `personalOwner` at class `Person` and `corporateOwner` at class `Corporation`. Distinct rolenames allow for a clear navigation from the `Account` class to one of the others. However, the relationship itself, without looking at the classes from where the players are recruited, is not different for these two cases. In particular, nothing in the model states the close relationship between the associations with the rolenames `personalOwner` and `corporateOwner`, respectively[74]. In contrast, having the same rolename and the same association still allows us to restrict association instances, so-called *links*, to those where the corresponding role is played by, e.g., instances of class `Person`. Therefore, this model is thought to be more natural, and the mixture of analysis and design issues is rejected once again. This is all the more important since it is tempting to find problematic ways around this restriction, like the introduction of an artificial superclass `PersonORCorporation` representing the union of the classes `Person` and `Corporation`, which more or less lacks support from reality. A second possibility is the introduction of an intermediate class `Owner` with three binary associations, namely to the `Person` class, the `Corporation` class, and class `Account`, respectively. This variant may be appropriate for a social role `owner`, but it should not be pursued for relational roles.

Moreover, notice that UML does not allow for rolename instances, although there are association instances. This is interesting because, in spite of the lack of rolename instances, links (which are instances of associations) can be specified in object diagrams, thereby using the same rolenames as in the corresponding class diagram. Obviously, rolenames are different in either case, once they have to be interpreted as individuals while in a class diagram they refer to relational role universals. Thereby, UML circumvents the problem of guaranteeing that relational roles of two links instantiating one association refer to the same relational role universal (or rolename, in UML terminology), as described in section 3.3.3.

[73] Note that this may be overlooked due to other UML documents. For example, the User Guide claims in a "**Note:** The same class can play the same or different roles in other associations." [Booch et al., 1999, p. 67]. Nevertheless, this restriction is contained in the specification, cf. the description of the *name* attribute of association ends in [OMG, 2001, p. 2-23].

[74] This holds without regard of the names of the associations, which are only useful for human readers. More precisely, a reflection of this connection in the meta-model or in terms of a formal interpretation appears desirable.

4.2.1.2 Processes and Processual Roles

From a UML perspective, notice that the term *process* is used in an intuitive sense in UML documents, but the Reference Manual defines two notions of processes, to wit as “[a] heavyweight unit of concurrency and execution in an operating system.” and as “[a] software development process [...]” [Rumbaugh et al., 1999, p. 395]. Obviously, these definitions are very close to implementational aspects, instead of being analysis concepts. Processes in the latter sense refer to a variety of modelling elements and diagrams, mainly the *behavioural* part of UML. This comprises statechart diagrams, activity diagrams as well as sequence and collaboration diagrams. We will only deal with the last of these diagram types, because the others mainly focus on the temporal structure of processes. Nevertheless, one can also find notions related to processual roles for these other diagram types. *Swimlanes* in activity diagrams may serve as an example. The *structural* part of UML also provides concepts related to processes and processual roles. In particular, these are *operations* of classes in class diagrams as well as *use case diagrams*[75]. However, use cases refer more directly to social roles and are therefore discussed in the next section.

The search for processual roles start with the concept of *operations*. The combination of behaviour and structure within classes is often praised by a certain community as one of the main contributions of object-orientation, in comparison to data modelling techniques like ER modelling. In contrast, critics of object-orientation feel a superiority of objects over processes in object-oriented paradigms [Dori, 1995, 2002], or they do not agree that behaviour and structure can always be bound together [Hay, 1999a,b]. We also think that operations lead to a mixture of static and dynamic aspects within classes, and that this is advantageous with regard to implementation rather than purposes of analysis. There are attempts to elevate the concept from the implementational level, for example by the distinction between *operation* and *method* (cf. [Rumbaugh et al., 1999, p. 369]), where the latter is supposed to refer to implementational issues. However, it is not obvious what an operation in the object-oriented paradigm is supposed to model in the real world, as will be shown in what follows.

A variety of definitions refer to operations. [Balzert, 1999, p. 30] introduces them in the context of analysis as “executable activity”. [Oestereich, 1998, p. 38] describes them as “behaviour of objects” and as “services which can be invoked by an object with a message, in order to bring about a determined behaviour” [ibid., p.351; all three citations are our own translations]. Further standard textbooks on UML provide even more versions which are fairly similar at first glance, but each version can be interpreted ontologically differently, as will be elaborated after all definitions will have been presented. In the Reference Manual [Rumbaugh et al., 1999, p. 369] operations are introduced as follows:

> An operation is a specification of a transformation or query that an object may be called to execute.

The User Guide [Booch et al., 1999, p. 51] describes operations slightly differently:

> An *operation* is the implementation of a service that can be requested from any object of the class to affect behaviour. In other words, an operation is an abstraction of something you can do to an object and that is shared by all objects of that class.

[75] Parenthetically, it seems questionable why use cases are assigned to the structural view, in particular as their implementation is described by collaboration diagrams in the behavioural half of the UML.

Finally, the specification [OMG, 2001] defines on page 2-45:

> An operation is a service that can be requested from an object to effect behaviour.

According to these definitions, on the one hand, an operation can be requested, which might correspond to invoking a process that an object carries out or controls in some way. An analogy with the idea of an *Agent* processual role universal may come to one's mind here. On the other hand, an operation is an abstraction of what one can do to an object – hence describing processes by which this object is transformed (remembering a *Patient* role universal). Another point of view might be that an operation is an arbitrary process of which an object is a participant, but this raises the problem of explaining why an operation is attached to exactly this class and not one of the other participants of the corresponding process. Interestingly, it seems that the last case of being an arbitrary participant is the one which suits best the understanding of operations. In terms of processual roles, this means that an operation (i.e., a process) is assigned to one of the objects playing a processual role of it. However, there seems to be no explicit, common criterion for the assignment of an operation to exactly one of several involved objects. This is the reason why complex operations should be based on UML collaborations, a view which is supported by [Booch et al., 1999, p. 378]:

> In many cases, you can specify the realization of an operation by going straight to code. However, for those operations that require the collaboration of a number of objects, it's better to model their implementation via collaborations before you dive into code.

Consequently, collaboration diagrams appear to be a better candidate for process representation in UML with regard to purposes of analysis. Indeed, the concept of *classifier roles* is a model element which comes very close to the notion of processual roles developed in chapter 3. However, note that a classifier role is not itself considered a process since it is "instantiated as an object" [Rumbaugh et al., 1999, p. 204]. Therefore, a classifier role would be better understood as a universal of players instantiating certain processual role universal. But this difference is motivated linguistically, as there are no natural language words for processual roles (cf. section 3.4.1). Apart from that, classifier roles are used to name a way of participation on the level of the overall collaboration as the following description shows, and this corresponds to the intention of processual roles.

> A collaboration represents a group of objects that work together to accomplish a goal. A role represents the part of one of the objects (or a set of the objects) in carrying out that goal. [...] [Rumbaugh et al., 1999, p. 194]

Of course, there are differences in the details. Two properties of classifier roles are considered incompatible with processual roles. Firstly, a classifier role is restricted to one *base classifier* (or a collection of these, *all* of which have to be instantiated), which would mean that only instances of this classifier could play a certain processual role. Having a liberal idea of processes, we do not agree with the necessity of constraining every kind of process to some definite universal. The second way of modelling to which we object here shows a lack of restrictions in UML for modelling entities of one and the same ontological category in a determined manner. For example, one can express anonymous classifier roles, i.e., roles which are only specified by a base classifier. If classifier roles are to be interpreted as processual roles, an empty processual role could only mean that the supposed player (which is the instance of the base classifier) actually does not participate in the process described by the collaboration. Certainly, this is neither the case nor can it be intended. The optional specification of a classifier role is only reasonable in two cases. Either the corresponding base classifier actually describes a processual role already, or the classifier role is unambiguously and easily derivable from the base classifier. The first case refers to the above mentioned lack of restrictions, because a processual role is represented as a class, in contrast to a classifier role. Furthermore, both cases refer to the close relationship between processual and social roles which is now at the centre of consideration.

4.2.1.3 Mixtures with Social Roles

Actors in use cases are to an even higher degree entangled in mixtures of processual and social roles than classifier roles are. In general, one can easily pose the question where the difference between collaborations and use cases (which are *realised* by collaborations[76]) can be established, apart from the artificially compelled restriction that use cases do not show any internal structure[77]. [Wegmann and Genilloud, 2000] is a recent article which, at a high level of abstraction, seems to push the notion of use cases even further towards collaborations. The authors allow several use cases to be represented within one diagram, and actors may appear in several use cases, with multiplicity constraints assigned to them. This is not far off track, as the concept of actors is very similar to the classifier roles of collaborations, anyway. However, actors are restricted to entities *outside* the modelled system, whereas the system is described by one or several use cases. Furthermore, relations between use cases (like *extension* and *inclusion*) and the fact that they are to be realised by collaboration diagrams speaks in favour of an understanding of actors as representing processual roles. On the other hand, actors are said to themselves define *several* roles that may be played during interaction with the system [Rumbaugh et al., 1999, p. 144]. Further, they may be modelled as classes with the stereotype ≪actor≫. As such, all characteristics of classes apply to actors, including attributes and operations. The combined features of actors resemble a mixture of social and processual roles. We think this is also the hidden driving force of [Wegmann and Genilloud, 2000]. The problem is that two types of classes are considered to be equal. Classes of the first type are constituted by behavioural aspects, namely collecting entities playing the same processual role. Classes of the second kind are based on abstractions due to common properties, which refers to a greater extent to social roles. We are not sure whether this could also be interpreted as an advantage, as the close connections within the role classification and a general need for a modelling language which allows for smooth transitions between concepts of different ontological categories has already been proposed (cf. section 3.6.3).

There is another mixture element which intermingles social and relational roles, namely *association classes*. These are associations to which attributes and behaviour can be assigned. Nevertheless, their instances are tuples of object references, and thus they do not have their own identity. From our point of view, two situations give rise to a model element like association classes. Firstly, the relationship modelled by the association may be FOUNDED by a process and characteristics of this process are

[76] A use case is understood as a specification of *what* happens, while collaborations specify *how* this what happens. A collaboration specifying how a use case is to be executed is said to realise this use case.

[77] This position is certainly influenced by works which claim that UML is a loose collection of separate modelling formalisms rather than one integrated whole, as stated in [Jähnichen and Herrmann, 2002], for example. If this is presupposed, model elements for very similar ideas may likely overlap to a high degree. Articles like [Wegmann and Genilloud, 2000] also support this hypothesis.

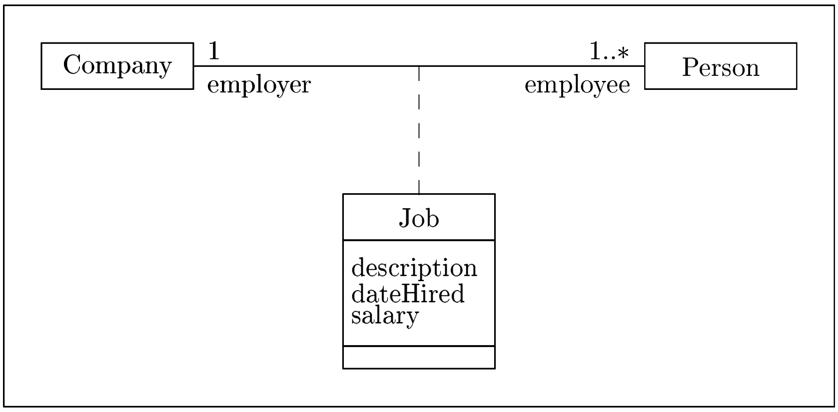

Figure 4.4: Example of an Association Class from [Booch et al., 1999, p. 148].

attached to the association. The attribute `dateHired` in Figure 4.4 seems to refer to this type of situation. The second kind of situation is present if a social role is mixed with a relational role of the association. The attributes `description` and `salary` would be better assigned to a social role, call it `Job`, which is in turn associated with the classes `Person` and `Company`. Assuming two binary relations, [Booch et al., 1999, p. 148] rejects this idea because "[t]hat wouldn't tie a specific instance of the `Job` to the specific pairing of `Company` and `Person`." We do not agree with this argument. If the maintenance of the mentioned pairing is absolutely necessary, one may simply use a ternary association for that effect. Apart from that, the association class in Figure 4.4 does not allow for a person having two jobs in the same company. Obviously, this is a general problem of association classes, which is solved by means of social roles.

Finally, *qualifiers* have to be treated. These also occur in connection with associations, but with the purpose of uniquely identifying objects being related by an association. Taking into account all the results up to this point, qualifiers are not considered to be directly related to roles. Rather, it seems that some social roles which may occur only once in a certain context can also be used to uniquely determine objects associated with others. In particular, this refers to the example with the qualifier values `treasurer` and `president` given on page 27. In this connection, these roles are abstracted to attribute values of objects of the class `Club`, where these objects provide the contexts for those roles in a social role reading.

4.2.1.4 Summary of Roles in UML

First of all, a general remark on process modelling with UML shall be made. There seems to be one cardinal problem which can be interpreted in two ways. Either there are too many possibilities for representing processes in UML, or the representational variety is not much worse, but the semantics of UML does not capture the relationships between different representations of one entity very well. With regard to the first case, one can model a process as a class, a use case, a collaboration, an activity diagram, et cetera. Among other questions, this depends on whether the process is behaviour of the modelled system, or merely represented as system data. It further depends on the question of in what detail a process is to be modelled. One problem with such a variety of ways to model entities of one ontological category is that different modellers develop different preferences of style and, even more problematically, very different models can be created for the same problem domain.

Concerning the lack of connections between different representations, neither is the user forced to model each perspective (which would be extremely expensive), nor can connections be expressed easily, although the Object Constraint Language (OCL) is a step towards a solution to this problem. Ontological clarifications may help to improve this situation.

Regarding roles, UML can be considered comparatively expressive. More specifically, no other formalism in chapter 2 embeds all three role types identified herein. Saying first that on a very specific, formal level, the relevant UML concepts and those herein differ, one can see that on an intuitive level relational roles correspond to rolenames, processual roles to classifier roles, and social roles to classes. Nevertheless, there are some impure UML notions with regard to roles, i.e., there are UML model elements which comprise a mixture of several types of roles. Collaboration diagrams, for example, may lead modellers to confuse processual and social roles. Preferably, the corresponding distinctions should be clarified within UML itself.

4.2.2 Roles in Works of Guarino

Our assessment of Guarino's formal approach regarding roles is twofold. On the one hand, the idea of a formal approach for capturing the essence of roles has to be welcomed. On the other hand, there are examples which do not fit his formal criteria, although they fit the descriptions given for these criteria. Therefore, one may be in doubt about the formalisation itself.

First of all, one needs to remember to which level of representation roles belong. According to Guarino, being a role is a meta-property of unary properties. In terms of GOL this leads to a classification of certain kinds of UNIVERSALS, hence role is a meta-universal with certain universals as instances. For example, the universal `student` is a role, and its instances are not role INDIVIDUALS, but individuals of some other kind, e.g., `human`. This does not correspond to our approach, because we consider roles themselves to be individuals (however, note that there are, of course, also role universals), which is an appropriate choice for each of the role types examined in chapter 3. For social roles, in particular, the underlying ontological levels justify role individuals. However, Guarino likewise seems implicitly to follow an ontological level approach. As described in section 2.1.4, the recent top-level ontology DOLCE, co-authored by Guarino [Masolo et al., 2002], contains a category which is a subcategory of individual social roles, namely "Social Agent", and some other categories appear to be good candidates of this sort, like "Non-agentive Physical Object" and "Non-agentive Social Object". The category "Agentive Physical Object" is disjoint from "Social Agent". Furthermore, DOLCE pursues a so-called "multiplicative approach", involving co-location of different entities and a notion of emergence. All of this suggests that DOLCE is based on an ontological level approach.

Apart from the question of roles being individuals or universals, the criteria *dependence* and *anti-rigidity*, used to distinguish roles from other properties, can be tested for each role subtype. In short, for a property ϕ to be a role, each instance of ϕ needs to be dependent on something else, which is not a part or a constituent of that instance, and which is itself predicated by some other property which needs to be equal for all instances of ϕ. Additionally, it must be *possible* for an instance of ϕ not to be an instance of ϕ, i.e., instances of ϕ are not *necessarily* instances of ϕ in order to exist (cf. section 2.1.4 for the formal definition).

For RELATIONAL ROLES, it is by definition true that they are dependent on other relational roles. What remains is the question of whether these are really never parts or constituents, but we have not found a counter-example. Anti-rigidity is a little harder to defend. Consider the property of being a child (on the basis of a parent-of relationship). Whether one can attribute anti-rigidity to child depends mainly on the question of when that property is lost. If one is still child after the death of her/his parents, it should be necessarily the case that all children really are children, even if the latest and most strict version in [Guarino and Welty, 2001] is applied, which clearly distinguishes the possibility of having a property and the temporal distribution of a property over the life of an instance of that property. As a consequence of this, child would be a rigid property, which contradicts the requirement of anti-rigidity for roles. So this is a problematic example for the type of roles accounted for by Guarino.

Nevertheless, it is interesting to test the criteria of anti-rigidity and dependence against processual and social roles. Concerning processual roles, even dependence seems inappropriate. This is the case for PROCESSES which involve only one ABSTRACT SUBSTANCE. In this case, no other abstract substance can be determined on which the process participant is dependent. A trivial solution might be to choose the process instead, but then the notion of dependence becomes irrelevant, as every instance of some first-order property depends on a process or SITUOID it is contained in. Further, the substance may be counted among constituents of the process, which excludes this process from being employed as a reference of dependence according to Guarino's formal definition of dependence. A circumvention of the problem of singular processual roles may be hidden in the example of pedestrian which bears an *implicit* dependence – unfortunately, the notion of implicit dependence in [Guarino, 1992] is not further elaborated as far as we know[78]. Anti-rigidity may be analysed similarly by posing the question of whether there are processes with processual roles which are, intuitively, considered roles in Guarino's sense, but which are necessary for some substance. If, for instance, breather (some breathing entity) is considered a role, it is not anti-rigid since anything which is breathing does this necessarily and all the time during its existence (on a mesoscopic temporal granularity).

Social roles fulfill the dependence criterion, as they require an entity of which they themselves are a part of. However, anti-rigidity has to be rejected in a direct fashion, e.g., a student is necessarily a student, as nothing remains when it ceases to exist. The human playing the role of a particular student is completely separate from the latter, as this human is on a different ontological level. Further, we do not see a reasonable way to express the anti-rigidity criterion in terms of the PLAYS relationship.

In summary, there are some open problems with Guarino's criteria for determining roles, i.e., one can find counter-examples of entities which are intuitively roles, but which are not identified as roles in the formal approach. In general, a universal predicated as a role in Guarino's meta-classification is a universal of players of some kind of relational roles. A final note concerns the formalisation in terms of modal logic. We do not think that this is a favourable approach, as well as the philosophical possibilist position in general. It is not clear to us why a student is not necessarily a student, but a human is necessarily a human. It seems equally conceivable that there are two possible worlds, one with an object Peter which is a human, the other with Peter being an arbor vitae[79]. Of course, this objection is only appropriate in a context-independent sense, i.e., if the modality is understood in a comprehensive manner. If, instead, this approach is supposed to refer to specific contexts (analogously to the use of modal logic in program verification, where possible worlds represent, e.g.,

[78] It may refer to social roles, viewing the pedestrian as part of a social system.

[79] Further, tales of magicians suggest that this is conceivable for many people.

states of execution), one has to admit that it is a flexible approach which re-identifies predetermined distinctions, e.g., between roles and natural types.

4.2.3 Sowa's Notion of Roles

At first glance, Sowa's conception of roles appears to be very similar to that of Guarino: most of his examples can be said to be universals of players of RELATIONAL ROLES, mainly as he assumes dyadic relationships between instances of his roles (which are universals). However, if the details of Sowa's subclassification of roles are taken into consideration, this first impression becomes blurred. Sowa makes many claims which will now be at issue, before another attempt can be made to identify correlations between Sowa's and our role concept.

In what follows in this section we refer to [Sowa, 2000] and to the introduction of his approach in section 2.1.3. In particular, we refer to the part of Figure 2.1 below *Role*. Sowa gives the following definition with regard to roles[80].

> **Role; Has(Entity, Entity).** Has is a general relation type used to define all roles. It represents the has-test of Section 2.4, which is used to determine the prehending entity and the prehended entity of a prehension. All relation types listed in Sections B.3 and B.4 are subtypes of Has. [Sowa, 2000, p. 505]

Accordingly, the first layer[81] of Sowa's role classification is formed by the distinction between a prehending and a prehended entity. If the *has test* (cf. [p. 84]) applies, i.e., if the sentence pattern "X has Y" sounds natural, X is identified as the prehending entity, and Y is the prehended one. Examples given in [Sowa, 2000] vary from "The car has an engine" to "wholes have parts". Thereby, no distinction is made between MATERIAL and FORMAL RELATIONS, like `engine-of` and `part-of`, respectively. Even if this distinction is considered irrelevant, it is not clear why `car` should be a role term.

In general, we do not agree with Sowa's substantiation of the *has test* [p. 84]. He refers to the category *Having* in the ontology of Aristotle, and the occurrence of the word "have" in many languages. Contrary to Sowa, we see the following objections to the *has test*. First, it selects a collection of relational roles in a somewhat arbitrary manner. Sowa himself points out that there are relations for which this test does not apply, like situatedness relations (cf. "The earth is situated beneath the sky." [ibid.]). If these relations do not justify the existence of roles, one can assume a kind of ontological siblings of roles for such relations, or it should be explained why there are none. The second problem is that the *has test* can also be applied to individuals, as in "Mary has a book." In spite of this, `Mary` or the `book` should not be considered as roles. Thirdly, we do not think that simple language tests provide good evidence for ontological questions, since languages seem to be too flexible to pinpoint ontological categories in a one-to-one assignment of syntactic elements. For example, nouns cannot simply be

[80] Note that the term *prehension*, which is borrowed from Whitehead, refers to a "physical relative" which relates two physical entities. Prehension is contrasted with *proposition*, the latter denoting entities which relate a *form* and an *actual* entity [Sowa, 2000, p. 70]. Unfortunately, we cannot introduce the complete ontology of Sowa as this would require a disproportional amount of space. Therefore the reader is referred to Sowa's book for a presentation of the categories *form* and *actuality*. It seems appropriate, though, that dyadic MATERIAL RELATIONS are included in Sowa's category of *prehension*.

[81] The ordinals of layers refer to the number of divisions made according to *Role*, and not directly to Figure 2.2. *Role* in the figure refers to layer 0 here.

assigned to the category of SUBSTANCES, if one thinks of nominalisations of verbs (cf. `flight`) and adjectives (cf. `height`)[82]. Finally, the greatest problem is that, even for universals, the *has test* does not distinguish whether a universal is a role universal or not. Consider the two sentences "A human has a mother." and "A child has a mother." – from our perspective, an account of roles should separate terms which refer to roles and those which do not. However, the *has test* declares `human` and `child` to belong to the same class, namely that of prehending entities.

The second layer below *Role* in Figure 2.2 is formed by a distinction which actually refers to a property of relations: a relation is called extrinsic, "[i]f either entity in a prehension could disappear without affecting the form or existence of the other, [. . .]" [p. 87]. Otherwise it is intrinsic. Only two of the three categories on this level are further divided in terms of Husserl's distinction of independence or dependence, which yields four categories on the subsequent layer: *Whole*, *Substrate*, *Part*, and *Property*. Interestingly, these terms seem to refer to the relational roles of the FORMAL RELATIONS PART-OF and INHERENCE[83]. The next subdivision refers to the top-level differentiation between, in GOL terms, SUBSTANCES and PROCESSES[84]. In Appendix B.5 of [Sowa, 2000] this gives cause for the insertion of theta roles below the category *Participant*. The latter can be understood to resemble our category of PROCESSUAL ROLES. However, there is a major discrepancy. We cannot agree that processual roles are a subcategory of relational roles, as the former belong to the realm of OCCURRENTS, while the latter are ENDURANTS. This problem originates from building a hierarchy of player universals based on roles, instead of a classification of the corresponding role universals themselves. More specifically, instances of Sowa's theta role categories play processual roles, which are themselves considered to play the (intuitive) relational role *Part* of the PART-OF relationship with respect to other processes. According to this overall approach, Sowa's category *Role* should have the same extension as ENTITY, because everything stands in some formal or material relation to something else, i.e., everything "has" anything else.

One conclusion which can be drawn is that the role approach presented by Sowa can be expressed in terms of our terminology. Sowa's category *Role* can be described as *player of some relational role*, assuming the extension of relational roles to FORMAL RELATIONS. However, the relations included by Sowa are restricted by the linguistic *has test* whose ontological adequacy and quality is rejected. Although parts of Sowa's hierarchy appear ontologically appropriate, the overall hierarchy veils the distinction between relational and processual roles due to the placement of high-level processual roles (theta roles are understood as such, cf. section 3.4.4 below relational roles in the hierarchy.

4.2.4 Steimann's Properties of Roles

In the present section we reconsider those properties ascribed to roles which were collected by Friedrich Steimann (cf. p. 28 or [Steimann, 2000c]) in his extensive comparison of object-oriented and data

[82] Admittedly, language tests similar to the *has test* appear also in linguistics. For instance, a test based on a so-called genitive *of-phrase* is claimed in [Barker and Dowty, 1992] to identify the semantically defined class of *relational nouns*. However, this test is likewise subject to exceptions, as noted by Barker and Dowty.

[83] The term *relational role* is used here on an intuitive basis, in spite of the fact that relational roles of FORMAL RELATIONS cannot be introduced formally because this would create an infinite regress. However, the restriction of relational roles to MATERIAL RELATIONS is a problem of formalisation rather than for the existence of these roles.

[84] Note that Sowa's description of this layer [p. 89] differs from Figures 2.11 and B.2 given in [Sowa, 2000], apart from the fact that both figures also differ from each other with respect to this layer. In the text, Sowa additionally distinguishes between *Participants* and *Stages* of a process, which resemble processual roles and temporal parts, respectively.

modelling literature. The following list presents some of these properties in more detail than in section 2.2.2, in order to point to some ontological problems afterwards, and to assign each property to the appropriate role types on the basis of our classification. Our comments follow the long dash (—).

1. "*A role comes with its own properties and behaviour.* This basic property suggests that roles are types. And indeed, only few approaches do not regard roles as types; [...]" — Interestingly, types, i.e., UNIVERSALS, are equated here with SUBSTANCES. The actual role statement (in italics) can only hold for social roles, because neither relational nor processual roles have MOMENTS, nor do they allow for new MOMENTS of their players. For social roles, moments can be found on the individual level.

2. "*Roles depend on relationships.* As suggested by the work of Sowa and Guarino, a role is meaningful only in the context of a relationship. [...]" — Obviously, this criterion refers to relational roles, for which one may add that (material) relationships also depend on relational roles, although relationships seem to be prior to relational roles. Processual and social roles are not directly dependent on relationships. However, as social roles hardly exist in isolation, they might also fulfill this property.

3. "*An object may play different roles simultaneously.*" — This point can be agreed to for all roles.

4. "*An object may play the same role several times, simultaneously.*" — First of all, here the term *role* has to be understood as role universal, as individual roles can only be played once (at the same time). This interpretation provided, the statement is correct for all role types, because a player can play several roles instantiating the same role universal, in general. However, one can likewise easily imagine cases where a player can only play one role instantiating a certain universal at the same time, e.g., the processual role `runner` does not allow for a simultaneous `run` elsewhere.

5. "*An object may acquire and abandon roles dynamically.*" — Although ostensibly adequate, this property requires different interpretations for processual and other roles, respectively. For example, acquiring a relational or a social role refers to two instances of an ABSTRACT SUBSTANCE, the temporally former without the role, the temporally latter exhibiting the role. In contrast, processual roles are directly played by abstract substances and accumulate over time since they are processes.

6. "*The sequence in which roles may be acquired and relinquished can be subject to restrictions.*" — Such restrictions can be expressed on the basis of GOL and our role theory in terms of domain-specific axioms. This property mainly seems to refer to social and processual roles, e.g., in the latter case this may reflect the temporal structure of a process comprising subsequent roles of one abstract substance.

7. "*Objects of unrelated types can play the same role.*" — This property holds for all roles. Nevertheless, constraints can be explicitly specified, e.g., if all players of certain roles are instances of one and the same universal (which is not based on roles).

8. "*Roles can play roles.*" — If at all, this property refers to social roles where it is an open question of our analysis (see section 3.5.4). It cannot hold for relational or processual roles, because the players of both belong to different ontological categories than the roles.

9. "*A role can be transferred from one object to another.*" — Role transfer is limited to social roles, as other role individuals are bound to one (in the case of processual roles: abstract) substance.

10. "*The state of an object can be role-specific.* The state of an object may vary depending on the role in which it is being addressed. Together with item 4, i.e., the possibility of one object playing the same role multiply at the same time, this seems to suggest that each role played by an object should be viewed as a separate instance of the object." — This requirement follows from mixing up social roles and their constituents, i.e., players. As these entities are different herein, this statement is not supported directly. However, it amounts to the conceptual formation of an entity integrating a player and various social roles.

11. "*Features of an object can be role-specific*. Attributes and behaviour of an object may be overloaded on a by-role basis, i.e., different roles may declare the same features, but realize them differently." — Cf. the comment of the preceding item.

12. "*Roles restrict access.* When addressed in a certain role, features of the object itself (or of other roles of the object) remain invisible. This corresponds to an object having different perspectives, facets, or aspects." — The note on different perspectives might be misunderstood, because roles are not the only way to come to different views (see section 4.1). Parenthetically, access-restriction is rather an implementational issue.

13. "*Different roles may share structure and behaviour.*" — This point mainly refers to social role universals, for which one can easily agree since these are substance universals, for which this property already holds. Relational roles do not have any structure or behaviour. Processual roles have a (temporal) structure, thus this property may partially apply to them.

14. "*An object and its roles share identity.*" — In contrast to this point, we advocate the position taken in the following, last item, which seems to belong to the minority in the object-oriented area. The approach of shared identity is not compatible with the separation of social roles and their material basis, and it is even less applicable for relational or processual roles in a strict sense of identity. Furthermore, Steimann himself mentions the *counting problem* (e.g., fewer people than passengers are transported over a week) addressed in sections 2.3.1 and 3.4.3, and thus he acknowledges a certain difference.

15. "*An object and its roles have different identities.*" — Cf. the comment of the preceding item.

On the whole, Steimann's collection appears to describe social roles in the majority of cases, although some items apply to all role types, with possibly differing interpretations. In general, these properties can be broadly interpreted due to the fact that they are fairly restrained, e.g., the use of "may" and "can" leaves the question open whether an item holds for *all* or only for *some* roles. Further, some items are better read as referring exclusively to role universals (like items 3, 4, and 7), others obviously describe role individuals (like items 9 and 15). Admittedly, some of these properties were not discussed in chapter 3, in particular those focussing on the PLAYS relationship. However, it is to be shown here that, firstly, these properties can be accounted for in terms of our notions and, secondly, that there can be differences between their degrees of validity if interpreted for distinct role types. The second point further supports the distinctions made.

Steimann's own approach to roles merges relational and social roles on the level of UNIVERSALS. A Steimann role has properties and behaviour, but there are no special instances. Instead, instances of roles can be found in certain types, specified by a *fills* relationship between roles and types. Some of the issues discussed in section 3.3 are answered, but without any reasons or discussion. For instance, a role can only occur once in a relationship, and the order independence is realised but dropped, i.e., in order "[...] to be able to use standard set notation in certain situations, a total order on the set of roles is assumed, [...]" [Steimann, 2000c, p. 95]. Both examples cause severe restrictions on the generalisation hierarchy of roles, for example, if a role (which is a type for Steimann) should comprise two disjoint roles of a relationship, this contradicts the idea of having different roles in a relationship. Many other points would require further examinations, but we conclude with the statement that Steimann's collection of properties as well as his approach can be expressed in terms of our framework.

4.2.5 Remarks on Other Formalisms

4.2.5.1 ER Modelling

The only role type that is supported in standard ER modelling is that of relational roles. This is the case because the formalism contains only *entity types*, *attributes* and *relationship types* as modelling elements. Processes, for instance, are not explicitly considered. This choice of primitives may be sufficient to simulate other role types, as it is possible to represent FORMAL RELATIONS like IS-A or PART-OF as relationships. However, such simulations do not contribute to the semantics of the formalism, and they cannot be considered a true substitute.

With regard to relational roles in original ER modelling, it should be said that Chen [1976] understands relationships as purely mathematical relations, i.e., as sets of ordered tuples. This is not obvious from the quote given on page 25. However, this reduction of relations to sets does not allow for equal roles in a relationship, as explained in section 3.3, because the difference between roles and tuple positions disappears in the set-theoretic framework. This may explain why Chen did not dwell further on roles and their properties.

4.2.5.2 Object Process Methodology (OPM)

As introduced in chapter 2, OPM is a systems modelling approach with an explicit process representation. In contrast to the UML separation of modelling aspects into structural and behavioural views, OPM aims at one integrated model for one level of specificity. In addition, it provides mechanisms to show the transition from one level of specificity to another. As it is the only formalism with an explicit process representation, it is of special interest whether OPM models processual roles explicitly.

Indeed, *procedural links* are model elements which are closely related to processual roles. More specifically, there is a fixed set of procedural links which can connect *objects* and *processes*[85]. These links correspond to a number of high-level processual roles, similar to our interpretation of theta roles in section 3.4.4. Of course, this does not allow us to represent all processual roles. However, the success of a few procedural links, which mirrors the attempts to manage the semantics of verbs with

[85] Note that procedural links between *states* and processes are not included here. Rather, a proper understanding requires a deeper analysis of the OPM notion of state, which at first glance resemble FACTS.

a limited set of theta roles, suggests that such a link *in conjunction with* a certain kind of process is in many cases sufficient to derive the actual processual role.

Notice that this hypothesis refers to one level of specificity/complexity. It becomes unsustainable for different levels, as can be seen in [Dori, 2002, chapter 9], entitled “Managing Systems’ Complexity”. This chapter introduces several ways of how the transition between specific and more abstract/less detailed diagrams is represented using OPM. In connection with such transitions, procedural links need to be merged if one abstracts from a detailed process, because only one type of procedural link is allowed to connect an OPM object and a process. A definition of preferences among the procedural links guides this merger. The need for these preferences is an interesting observation, which can be identified in the framework of processual roles. For example, consider a processual role which instantiates the universal `consumee`, with the meaning that the abstract substance playing this role is destroyed due to the process. This role can consist of two parts, the temporally first of which cannot also instantiate `consumee`, while the second one has to instantiate this universal. What has been said so far with regard to processual roles in OPM shall suffice as a remark suggesting further work. A more detailed analysis and comprehensive coverage of OPM processes and all their features is beyond the scope of this thesis.

In contrast to the representation of processual roles which is rare in other formalisms, there is a lack of explicit OPM modelling elements for relational roles, the latter being rather common modelling elements. Structural relations, the OPM equivalents of MATERIAL RELATIONS, are not provided with an analogue to UML rolenames, for instance. In addition to this, OPM distinguishes between *forward* and *backward* structural relations, which corresponds to the Arrangement Problem of predicate logic (cf. section 3.3.2). This also reveals the purely set-theoretic interpretation of relations in OPM. Notice that *role-playing sentences* can be employed in the sense of relational roles (cf. section 2.2.4), but they lack the semantic connection to structural relations. Therefore, relational roles can only be represented implicitly in OPM.

Finally, social roles are not distinguished from usual OPM objects. Objects are the proper modelling element to represent social roles, because objects comprise ABSTRACT SUBSTANCES and MOMENTS (and some other categories, cf. section 2.2.4) which is supported by many examples in [Dori, 2002] as well as the fact that the *exhibition* relation can connect objects. *Exhibition* is understood as comprising INHERENCE, but the complete ontological status with respect to GOL is unclear, because it can also connect processes with each other, as well as processes and objects.

5 Applications

This chapter covers a twofold application of the role classification developed so far. The first part discusses the incorporation of roles into GOL, whereas the focus is set on the effects for the top-level hierarchy of GOL, as well as on the status of the formal relation of HOLDING. The second part of this chapter deals with a sample model in the domain of clinical trials. In connection with the current development of the application SOP-Creator, roles in the domain of clinical trials are analysed and modelled exemplarily, which will then result in some practical guidelines.

5.1 Roles in the Framework of GOL

5.1.1 Issues of the Top-Level Hierarchy

The axiomatisation of RELATIONAL, PROCESSUAL, and SOCIAL ROLES which was developed in chapter 3 is intended to augment GOL with a clarified and sufficiently general concept of roles. Therefore, the axioms conform with the present GOL version as introduced in [Degen et al., 2003], and consequently, the axiomatisation is proposed to be added to the current set of basic axioms. More precisely, the signature given in Table 3.2 needs to be joined with the signature for basic axioms in [Degen et al., 2003], as well as some syntactic adjustments will be necessary, because the syntax of GOL was not completely determined when this work was written. However, this should not be problematic, as the axioms mainly deal with implications on the basis of symbols of our own signature. Nevertheless, it remains a future task to prove the consistency of the combined axiomatic system. Apart from the adoption of the axioms, it is a different question whether the category *Role* or its subcategories can be appropriately inserted into the "backbone" of the GOL hierarchy (cf. the left-hand side of Figure 1.1, p. 12). In the remainder of this section, possible ways of achieving this will be discussed, as well as their interplay with certain other GOL elements.

According to chapter 3, roles have turned out to be distributed over the current GOL top-level hierarchy, as they span from SUBSTANCES over PROCESSES to entities closely associated with RELATORS. This prevents a simple attachment of the category of *roles* to any of the existing categories, because the nodes on each level of the hierarchy are considered pairwise disjoint, and the first four levels[86] appear to be fully comprehensive. As was already noted in section 3.6.2, this is equivalent to the ontological character of *parts*. The concept of parts does likewise not appear as a node in the top-level hierarchy, and one can hold the view that roles should not be inserted either, not at least due to their strong similarity to parts.

A second variant of accomodating roles in the hierarchy amounts to a distributed insertion in terms of including the role subtypes under the corresponding nodes. This would result in positioning social roles as a subcategory of the category SUBSTANCE and processual roles as a subtype of the PROCESS category. Tentatively, relational roles would be placed under MOMENT, though this should be handled with caution as will be explained below. Discussing social roles first, this option exhibits a severe disadvantage. By definition, social roles require the distinction between social and non-social substances (p. 3.9, see also section 3.1.3). Furthermore, social substances can be separated into *role-based social substances*, i.e., social substances which consist of roles played by some material entity, and *repre-*

[86]The reference to four levels has to be understood relative to what has been specified yet. For clarity, the first level comprises only the category ENTITY, the fourth level, e.g., ENDURANT and OCCURRENT.

sented social substances, i.e., social substances which are merely represented in terms of material entities, thus relying on a denotation relationship which has not yet been introduced in GOL. Accordingly, a direct placement of social roles under substance would mix several higher-level distinctions. This suggests not elevating social roles to such a high level. Remembering the DOLCE ontology by Masolo et al. [2002] (cf. sections 2.1.4 and 4.2.2), their category *Social Agent* is also placed several steps below *Endurant*, where the latter may be viewed as a close relative of substance, because moments seem to be included in the DOLCE category *Quality* which is a sibling of their *Endurant*. We consider this further support for postponing the decision on where social roles should be inserted in the top-level hierarchy.

Processual roles can be evaluated similarly. They do not appear to be good candidates for a direct, highly relevant subclassification of processes, although they are a subcategory of processes. Note that a category may always be classified according to different criteria. However, the top-level hierarchy should only refer to criteria of major importance, arranged in a suitable order. The temporal structure of processes is expected to yield a classification of processes which is more relevant than the distinction between processual roles and non-roles, with respect to a balanced division of the overall process category. The existence of a huge number of modelling formalisms which allow one to specify the temporal structure of processes compared to the small number of formalisms dealing *only* with different participants is taken as an indication for the prior relevance of temporal structure. Thus, we suggest an insertion of processual roles on lower levels than those which are currently the matter of debate.

For relational roles, the situation is slightly different. Relational roles have not been assigned at a level of the hierarchy which is as specific as for social or processual roles, but only to ENDURANTS. Clearly, relational roles are neither substances, nor CONFIGURATIONS. By axiom 3.9 (p. 50), they are not relators. Hence, the most attractive category for relational roles is moment. As mentioned at the end of chapter 3, one may consider the formal relation PLAYS to be equivalent with INHERENCE with respect to relational roles. However, [Degen et al., 2003] presents an example of a relator which causes problems with regard to this equivalence. It amounts to the fact that a relator can likewise connect substances and UNIVERSALS, although one substance has to be among the relata. Thus, in general, one or more relational roles of a relator may be played by universals. However, moments are constrained insofar as they can only inhere in substances. Consequently, if relational roles are to be placed under moments, either this notion of GOL has to be adjusted, or one has to restrict relators in such a way that they connect only substances. The axiomatisation of relational roles presented in section 3.3 should not conflict with universals being players of these roles, however.

Let us mention another problem associated with relational roles as moments, which may emerge in the future. As developed herein, relational roles are restricted to relators and hence to material relations, but intuitively, they are equally conceivable for formal relations. If the problem of the infinite regress arising from the assumption of relational roles for all formal relations (cf. section 3.3.1) can be solved or restricted to a limited number of formal relations, the insertion of only those relational roles referring to material relations would cause very different positions of formal and material relational roles within the top-level hierarchy. This is not desirable because nothing seems to suggest tremendous differences between these subtypes of relational roles themselves, apart from the type of relation they are roles of.

Consequently, we offer two proposals to the insertion of roles into the GOL top-level hierarchy. The categories of processual and social roles should not be added, while relational roles may tentatively be placed under the category moment. However, if the presence of only one subtype of roles in the top-level hierarchy is not desirable, roles should not be inserted at all at the current state of development. For clarity, the presented axiomatisation of roles *is* suggested for complete integration into the axiomatic system of GOL.

5.1.2 A New Approach to Material Facts

The category of relational roles can be exploited in connection with the theory of RELATORS in GOL. More precisely, we suggest an understanding of MATERIAL FACTS on the basis of relational roles, instead of employing the formal HOLDING relation in order to describe material facts. Let us first introduce these terms in greater detail. [Degen et al., 2003, chapter 5] discusses the idea of a material fact being an INDIVIDUAL consisting of exactly one relator and all the entities which are connected by this relator. Material facts themselves are considered constituents of SITUATIONS, which, in turn, are certain collections of facts, being independent wholes. Consider a sentence like "`John owns that book.`" Assuming the abbreviations `j` for `John` and `b` for `that book`, this sentence refers to a relator r which mediates between `j` and `b`, where r mediates the material relation `owns`. This yields a material fact denoted by the term $\langle r : \mathtt{j}, \mathtt{b} \rangle$, which is itself kept together by the basic formal relation of HOLDING. In detail, GOL declares the entities r, `j`, and `b` "(*in this order*)" [Degen et al., 2003, emphasis added] to stand in the HOLDING relation.

With this definition, GOL inherits the Arrangement Problem described in section 3.3.2. In our example, the term $\langle r : \mathtt{b}, \mathtt{j} \rangle$ would be considered to denote a different fact, based on the order which is imposed on the arguments of the HOLDING relation. Analogously to the usual interpretation of predicate logic, one would understand this term as "`That book owns John.`", because there are implicit assignments of relational role universals to each position. However, in conformance with chapter 3, the following reading of $\langle r : \mathtt{j}, \mathtt{b} \rangle$ should be preferred: the relator r connects the entities `j` and `b`, which is equivalent to the idea that r connects `b` and `j`. Admittedly, thereby the information is lost which entity plays the `owner` role and which one the `ownee` role (where `ownee` refers to that entity which is owned), but this should explicitly be accounted for by relational roles.

The following approach supplements material facts with the information about relational roles, such that the Arrangement Problem is solved. We suggest replacing the HOLDING relation by PLAYS and ROLE-OF. In order to make the necessary relational role universals explicit, the denotation of material facts has to be supplemented by this information. In addition, once information about universals is contained in the notation, one can also specify a universal which is instantiated by the actual relator. The following is proposed as the denotation of the material fact from the example[87]:

$$\langle \mathtt{Owns}(r) : \mathtt{Owner}(\mathtt{j}); \mathtt{Ownee}(\mathtt{b}) \rangle \tag{5.1}$$

This is, by definition, equivalent with $\langle \mathtt{Owns}(r) : \mathtt{Ownee}(\mathtt{b}); \mathtt{Owner}(\mathtt{j}) \rangle$, because, actually, it is defined to represent the conjunction of the following formulae expressed in the signature used in chapter 3.

$$\mathtt{Owns} \longrightarrow\!\!\triangleright\, \mathscr{R} \wedge \mathtt{Owner} \longrightarrow\!\!\triangleright\, \mathscr{Q}_{\mathsf{rel}} \wedge \mathtt{Ownee} \longrightarrow\!\!\triangleright\, \mathscr{Q}_{\mathsf{rel}} \wedge \mathrm{disjoint}(\{\mathtt{Owner}, \mathtt{Ownee}\}) \tag{5.2}$$

$$r :: \mathtt{Owns} \wedge \exists! q \exists! q' (q \neq q' \wedge q \succ r \wedge q' \succ r) \tag{5.3}$$

$$\exists q\, (q :: \mathtt{Owner} \wedge q \succ r \wedge \mathtt{j} \rightsquigarrow q) \tag{5.4}$$

$$\exists q'\, (q' :: \mathtt{Ownee} \wedge q' \succ r \wedge \mathtt{b} \rightsquigarrow q') \tag{5.5}$$

[87] For clarity, universals start with a capital letter.

This representation makes explicit the implicit assumptions on the roles of a relator and is no longer dependent on a specific ordering. Further, it allows a simple representation of a generalised form of *intensionally symmetric* relators (cf. section 3.3.4), some roles of which instantiate the same relational role universal. For instance, if John (j) and Mary (m) together own a book c, mediated by the relator r', this can be denoted in the following way.

$$\langle \mathtt{Owns}(r') : \mathtt{Owner}(\mathtt{j},\mathtt{m}); \mathtt{Ownee}(\mathtt{c}) \rangle \tag{5.6}$$

Obviously, this also accounts for anadic relator universals, of which Owns is an example here. Note that j and m in the parentheses next to Owner are not considered to be ordered with respect to each other.

Incidentally, this way of representing material facts is a step towards a solution to the problems of the assignment between relational roles and relational role universals, which were discussed in section 3.3.3. However, it is not a complete solution to this problem, unless further assumptions are made. Concerning the traditional, order-based representation, these assumptions are also made, yet implicit, because relational roles implicitly derived from tuple positions can only be interpreted in dependence of the relator universal assigned to the tuple. Made explicit, this corresponds to two suppositions. Firstly, terms for material facts with the same leading relator universal have to exhibit the same role universals. This is a questionable constraint, because optional role universals cannot be represented and ambiguities arise from the representation of anadic relations unless further assumptions are made. Secondly, role universals can only occur with one and the same relator universal. This position is easier to comprehend for *certain* relational role universals rather than *all* of these, because, for example, one may be interested in role universals which are not only determined by a relator universal, but in addition by properties of their players. For such role universals it may be hard to find a corresponding relator universal, apart from falling back upon artificial definitions.

More formally, the following definition is given for the denotation of material facts. It generalises the above examples and reflects the implicit assumptions which can be attributed to the representation based on ordered tuples.

Definition 5.1 (Denotation of Material Facts) *Let f be a material fact yielded by an arbitrary relator r connecting all entities $e^i_{k_i}$ in $\bigcup_{1\leq i\leq n}\{e^i_{k_i} \,|\, 1 \leq k_i \leq m_i\}$, for the constants n ($n \in \mathbb{N},\ n \geq 2$) and $m_1, \ldots, m_n$ ($m_i \in \mathbb{N},\ m_i \geq 1$ for all $1 \leq i \leq n$). Further, let R be a relator universal, and $Q = \{Q_i \,|\, 1 \leq i \leq n\}$ a family of relational role universals. Then, f can be denoted by*

$$\langle\, R(r) : Q_1(e^1_1, \ldots, e^1_{m_1}); Q_2(e^2_1, \ldots, e^2_{m_2}); \ldots; Q_n(e^n_1, \ldots, e^n_{m_n})\,\rangle \tag{5.7}$$

if the following conditions are satisfied:

- *r instantiates the relator universal R,*
- *r has exactly $m = \sum_{i=1}^{n} m_i$ relational roles, i.e., the arity of r is m,*
- *for each pair of indices i and k_i ($1 \leq i \leq n,\ 1 \leq k_i \leq m_i$) there is exactly one relational role $q^i_{k_i}$ which is an instance of Q_i and a role of r, and which is played by $e^i_{k_i}$, as well as*
- *Q is a* role base *for R.*

Note that the last condition refers to the notion of a role base, which was introduced in section 3.3.3. Thereby, the specified conditions are rather strict (for instance, the last one excludes the representation of optional roles). Nevertheless, thus far two improvements with respect to the representation of material facts can be mentioned. Anadic relations can be specified easily and unambiguously, as well as there are no hidden implications in the order of the entities. That means, permutations of the arrangements of the arguments (apart from the fixed leading position of the relator) can denote one and the same material fact, even if different universals are chosen, because relational roles are INDIVIDUALS which may instantiate several universals.

In order to allow for a denotation of material facts without specific universals, an abbreviation is given in Definition 5.2. This abbreviation corresponds to the present representation of material facts in GOL, however, with explicit, generalised assumptions. Note that the fulfillment of the last condition in Definition 5.1 with respect to the equivalence established in Definition 5.2 was already shown in Corollary 3.1 (p. 54).

Definition 5.2 (Simple Denotation of Material Facts) *Let r be an arbitrary relator which connects all entities in the family $\{e_i \,|\, 1 \leq i \leq n\}$ $(n \in \mathbb{N},\ n \geq 2)$.*

$$\langle r : e_1, \ldots, e_n \rangle \tag{5.8}$$

is defined as an abbreviation of

$$\langle \mathsf{Relator}(r) : \mathsf{RelRole}(e_1, \ldots, e_n) \rangle \tag{5.9}$$

i.e., both denote the same material fact based on the connection of all e_i by r.

There are several other concepts which are based on material facts, like FACTUAL UNIVERSALS, FACTUOIDS et cetera[88]. One may thus expect influences on such concepts, because relators are supplemented with a fundamental structure of relational roles. However, a material fact still consists of a number of entities which are related, and of exactly one relator which, in a part-like sense, comprises its roles. The latter is the reason for the fact that material facts on the basis of relational roles are more informative and explicit, without loosing information which previously was only contained implicitly. Consequently, the HOLDING relation can be dismissed, since it is not used in connection with any other kind of entity.

[88] A proper introduction of these notions requires a number of GOL concepts which are otherwise only tangential for this thesis. Therefore, an introduction is omitted, and the reader is referred to the Reference Manual [Degen et al., 2003].

5.2 Roles in Standard Operating Procedures

The practical context in which this thesis is situated has already been mentioned in the Introduction (chapter 1). The Institute for Medical Informatics, Statistics and Epidemiology (IMISE) of the University of Leipzig develops software applications for the domain of clinical trials, and it is conducting the project TP2 “Telematics” of the Competence Network for Malignant Lymphoma (KML[89]). At this point, some remarks on the subject of clinical trials should be given. According to [Pocock, 1983], clinical trials are “planned experiments” which are conducted with patients in order to determine the efficacy and utility of new treatments in a reliable manner. The treatment and the relevant data of an appropriate number of patients are precisely documented and evaluated, where the latter is usually based on statistical methods. In general, this requires a considerable amount of organisational efforts. This contributes to the fact that clinical trials exhibit challenging organisational, administrative, and documentational aspects, as well as they comprise medical, ethical, and statistical/biometrical facets. With respect to knowledge representation and modelling for software systems in this complex domain, a broad range of knowledge representation issues have to be handled. Moreover, good modelling practices are elaborated at IMISE in order to facilitate the development of software applications and to gain reusable pieces of software.

One of the software projects at IMISE deals with the development of a web-based application called *SOP-Creator*. SOP stands for Standard Operating Procedure, which is a manual for professionals involved in carrying out clinical trials. With regard to [ICH, 1996], an SOP consists of “detailed, written instructions to achieve uniformity of the performance of a specific function.” Put differently, SOPs are a means of quality management, describing how to organise and implement a certain task that appears within clinical trials. Concerning their degree of specificity, SOPs are usually kept on a level which allows their reuse for different trials. There are documents which exceed SOPs in generality, like *guidelines*, for instance the Guideline of Good Clinical Practice [ICH, 1996]. In contrast, *working instructions*, a third type of documents, are more specific than SOPs, as they usually apply to one trial only, providing more detailed specifications of tasks described in SOPs. *Clinical trial protocols* are equally specific as working instructions, but they describe rather medical aspects of the trial, whereas the former types of documents refer to organisational aspects.

At present, SOPs are developed on paper or via basic text processing software. Even for a single study group, there are a considerable number of SOPs for a variety of tasks. In addition, multiple study groups endeavour to harmonise their SOPs in order to improve their quality and to achieve uniformity, which, in turn, shall contribute to easier comparisons between clinical trials of different study groups. Harmonisation seems possible due to the intermediate level of specificity which can be ascribed to SOPs. The development of new SOPs as well as harmonisation processes require a lot of effort and are fairly complex. Therefore, a software application which supports these tasks could be very beneficial.

The SOP-Creator is a software tool which will support the management and, in later stages, the development of Standard Operating Procedures. It is being developed as an XML-based application which relies on a content management system and a directory service based on LDAP[90]. SOPs are stored in an XML format enriched with metadata, which are used by the content management system, for

[89] The abbreviation is derived from the German name, “Kompetenznetz Maligne Lymphome”.
[90] Lightweight Directory Access Protocol, cf. [Howes et al., 1999].

instance, for high-quality retrieval functions. The construction of some of these metadata defines the scope of the examination on roles in SOPs.

5.2.1 Structure of SOPs and a Selection of Domain Roles

For the KML, the SOP GE01, “SOP-Struktur, Erstellung und Änderung”[91] [Mehl, 2002] describes the structure of SOPs, their creation and the creation of amendments. In general, an SOP is a document of a handful of pages, which consists of metadata at the beginning, a table of contents, a number of sections with the actual contents, and optional appendices. The metadata exist independently of any information system, i.e., they were not introduced in the course of the introduction of an electronic SOP management system. Rather, they are necessary because SOPs are subject to quality management. These metadata contain information about the author, the date of change, the release date, the version et cetera – among them a datum called *target group*[92].

It is desirable to compile a vocabulary of all those terms which may appear as an entry of the metadatum *target group*, in order to restrict such entries to terms with a standardised interpretation, instead of allowing for free textual style. Furthermore, this would allow one to check the consistency of these entries. Such a vocabulary list must certainly consist of role terms, because people participate in certain roles in a clinical trial. Instead of a long and unstructured list of role terms, the application of the role theory introduced above may help to organise that vocabulary. First, one has to identify which role universals occur in the domain, and to which role types these belong. Then, these role universals can be organised according to our role types and domain-specific criteria. The resulting structure within the collection of terms will thus have clear ontological commitments. This structure may also serve as a basis for the presentation of the metadatum in the user interface of the SOP-Creator, as proposed in section 5.2.3.

There are SOPs which themselves define domain roles (often under a different heading like *responsibilities* or *positions*, however), i.e., roles which may occur in clinical trials and in particular in SOPs. The intention of these SOPs is, similarly to the motivation above, to provide a shared understanding of the terms used in a set of SOPs. Due to the fact that the SOPs of the KML are still being developed, we will base our discussion on the SOP SP06 “Verantwortlichkeiten”[93] [Brosteanu, 2001] (hereafter denoted by [SP06]) of the Coordination Centre for Clinical Trials Leipzig (KKSL[94]).

[SP06] aims at describing *areas of responsibility* within those clinical trials that can be administrated by the KKSL. The term *function* is used synonymously with area of responsibility in [SP06], and there are *positions* and *tasks* assigned to each of these functions. Notice that here function differs clearly from the concept of function which was presented in section 4.1, where the latter was understood as a certain PROCESSUAL ROLE UNIVERSAL related to a SUBSTANCE UNIVERSAL. The term *position* in [SP06] suggests synonymy with the term role, which is supported by the fact that “several fields of duty or positions within a concrete clinical trial can be taken on by one person. [...]” [SP06, p. 2]. Therefore, the term role is used interchangeably with position in this section. All areas of responsibility/functions as well as all positions/roles from [SP06] are listed in Tables 5.1 and 5.2,

[91] All SOPs to which we refer are written in German, and we will translate specific terms into English. The original German term is mentioned where necessary due to non-standardised translations.
[92] in German: “Nutzer-/Zielgruppe”
[93] in English: “Responsibilities”
[94] in German: “Koordinierungszentrum für Klinische Studien Leipzig”

○ Steering Committee (Leitende Kommission)	• Data Management (Datenmanagement)
○ Advisory Board (Beratendes Komitee)	• Biometrics (Biometrie)
○ Independent Data Monitoring Committee (Datenüberwachungskomitee)	• Databases (Datenbanken)
○ Ethics Committee (Ethikkommission)	• On site Monitoring (On site Monitoring)
• Clinical Trial Coordination (Studienkoordination)	○ Reference Institution (Referenzeinrichtung)
• Clinical Trial Counselling Service (Medizinischer Konsiliardienst)	○ Trial Center (Prüfzentrum)

Table 5.1: Collection of **Functions** in Clinical Trials, extracted from [SP06].

respectively[95]. Although the mentioned tasks form the basis of some arguments in the sequel, they are omitted due to space limitations, and the reader is referred to the original SOP. In contrast to the assignment of roles to functions in [SP06], these are consciously separated into two tables, where the original order of appearance in [SP06] has been preserved. For some roles it is straightforward to find the association to one or the other function, but for some this is more difficult, e.g., because some of these are assigned to more than one function. It is not clear why the latter case occurs since all functions seem to be distinct and on the same level of (part-whole) granularity, i.e., no one function is a part/subfunction of another. Therefore, the direct assignment of roles and functions has not been specified here, which also allows us to analyse both lists in a more independent manner.

[95]Note that the translation of these terms is difficult because even within the medical community no consensus on this has been established. Nevertheless, we present a tentative translation together with the original German terms in parentheses. As an additional, anticipatory note, the use of different enumeration symbols indicates subtle distinctions between the terms in each table. Intuitively, empty circles indicate formally established groups and their members, respectively, while filled circles refer to groups emerging from single roles with similar tasks assigned to them. Moreover, underlined circles point to roles with a different context than the others. The following section presents a much more elaborate account of these issues.

- ○ Principal Investigator (Studienleiter)[a]
- ○̲ Domain Expert (Fachexperte)
- ○̲ Independent Physician/Medical Scientist (Unabhängiger Mediziner)
- ○̲ Independent Statistician (Unabhängiger Statistiker)
- ○̲ Independent Domain Expert (Unabhängiger Fachexperte)
- • Coordinating Investigator/Project Leader (Studienkoordinator/Projektleiter)
- • Clinical Trial Secretary (Sekretariatskräfte)
- • Counselling Investigator (Studienarzt)
- • Documentalist (Dokumentar)
- • Documentation Assistant (Dokumentationsassistent)
- • Data Assistant (Dateneingabekraft)
- •̲ other qualified staff (weiteres qualifiziertes Personal)
- • Biometrician (Biometriker)
- • DB-Administrator (DB-Administrator)
- • DB-Programmer (DB-Programmierer)
- • Monitor (Monitor)
- ○̲ Physician/Medical Scientist (Mediziner)
- ○̲ Medical Assistance Staff (Medizinisches Assistenzpersonal)
- ○ Investigator (Prüfarzt)
- ○ Clinical Trial Assistant (Studienassistent)
- × *Independent board of a university hospital or a state medical association.* (*Unabhängiges Gremium einer Universitätsklinik oder Landesärztekammer*)

[a] The translation of this term is particularly difficult, as the definitions of the term in either language, English and German, vary. Other possibilities are *Investigator* and *Sponsor-Investigator*. However, Principal Investigator appears most appropriate, according to the information provided in [SP06].

Table 5.2: Collection of **Roles** in Clinical Trials, extracted from [SP06].

5.2.2 Evaluation of Domain Roles

The questions below will guide the examination of roles and functions in Tables 5.1 and 5.2. These issues are very much interlinked and their answers are distributed over this section, in order to allow for a natural flow of argumentation. However, on page 108 we provide a summary referring to these questions.

1. Do all terms in Table 5.2 really denote ROLES? If so, which types of roles appear?
2. In which way are the functions in Table 5.1 related to the *contexts* of the roles?
3. Some terms in Table 5.1 seem to denote groups rather than areas of responsibility. Do these groups have their own MOMENTS that would allow them to be characterised as SOCIAL SUBSTANCES? In what way do they support the role types in Table 5.2?

Possibly, all but one of the terms given in Table 5.2 (p. 105) could refer to ROLES. The exception is printed in italics with a leading ×. This item refers rather to a more complex social substance (as suggested by the word "board"). For the remaining items, there is a reference to exactly one player, which is one of the features of roles in chapter 3, and therefore these items appear to be ROLES. Put more precisely, they refer to ROLE UNIVERSALS, where the players of the instances of these universals are humans. In order to further support the hypothesis of having found roles, one has to determine the contexts of these roles, which also allows one to classify them as relational, processual, or social roles.

Let us examine the functions/areas of responsibility of Table 5.1, since these seem to be intended to provide the context for the role candidates just identified. The first observation is a variation among the terms applied, which span from terms denoting groups, like `Steering Committee` and `Advisory Board`, over terms denoting true areas of responsibility (in a narrow sense, not understandable as a function) like `Biometrics` and `Databases`, to terms referring to functions within clinical trials like `On site monitoring`. The latter two families of terms can be unified on the basis of the tasks which are assigned to each of the items (cf. [SP06]). Hence, the leading symbols of the items in Table 5.1 indicate the remaining distinction between *group terms* (○ or <u>○</u>) and *function terms* (● or <u>●</u>). This distinction has also been transferred to the roles in Table 5.2 which belong to the corresponding items from Table 5.1, employing the same symbols.

A group term certainly refers to a complex SOCIAL SUBSTANCE, thus the corresponding roles can be easily identified as SOCIAL ROLES. For example, the `Steering Committee` consists of the social roles `Principal Investigator` and `Domain Expert`. Of course, one can also derive a function for these groups, from their names and the tasks ascribed to them, for instance. In order to give an example, the function of the `Steering Committee` could be *leading*. However, these functions seem to be assigned to the whole group, which suggests a high degree of *interaction* in carrying out the corresponding tasks, presented in Table 5.3 for the `Steering Committee` for illustration. Conversely, function(s) can also be performed by single persons [SP06, p. 2], and therefore a distinction between the social substance and its corresponding function is necessary. With respect to the example, if the function of *leading* is assigned to a single human this results in a social role which is called `Principal Investigator`. However, this role is different from that social role above, which is also called `Principal Investigator`, but which is by definition a role of a `Steering Committee`. Further, in the case of the existence of a `Steering Committee`,

- Initiation, coordination and supervision of the trial
- Final decisions on
  - Responsibilities within the trial
  - Trial protocol
  - Amendments to the trial protocol
  - Realisation of the trial
  - Ex- and inclusion of trial centres
  - Cessation of the trial
- Discussion and decision of all important trial issues
- Supervision of the implementation of decisions
- Responsibility during the preparation of reports and publications

Table 5.3: **Tasks** of the `Steering Committee` [SP06, p. 3].

the function *leading* is distributed over all roles of the `Steering Committee`, instead of being assigned exclusively to the `Principal Investigator`. Admittedly, natural language does not draw these distinctions, but humans are able to switch flexibly between such nuances of interpretation when speaking about either kind of `Principal Investigator`, for example. This is augmented by the fact that, in the domain of clinical trials, there are often verbal agreements instead of written contracts or the like. Nevertheless, inconsistencies can arise if assumptions are attached to general concepts, possibly overlooking unmentioned differences.

Moving on to a related issue, a closer examination of that notion of `Principal Investigator` is required whose instances are not associated with a `Steering Committee`. In spite of this lack of a `Steering Committee`, this kind of `Principal Investigator` is assumed to be a social role universal. Thus the question arises which complex social substance provides the context for roles instantiating that universal. One idea is to allow for a `Steering Committee` with only one member, which appears convenient from an “implementational” point of view. However, in this case the above differences are only shifted to the explanation of *how* function(s) and properties of a group are distributed over its roles. It may also appear counter-intuitive to assume groups with only one member if groups imply several members. Yet another context can be found, namely an emergent *trial society*. Similar to the example of the `ant colony` from section 3.5, a trial society emerges from a clinical trial. Actually, the issue is a little more complicated, as clinical trials are planned in advance, which amounts to a declaration before the actual process starts and a trial society can emerge. With regard to such a trial society, the single `Principal Investigator` is a role of this emergent social substance. Indeed, *all* roles from [SP06] can be recognised as roles of this trial society. Note that this does not contradict the fact that the `Principal Investigator` of the `Steering Committee` is a role of the `Steering Committee`, as the latter is a part of the trial society and axiom 3.43 is assumed.

The trial society is a comprehensive social substance which also provides a context for those role universals from Table 5.2 which are assigned to the *function terms* in Table 5.1 (recognisable at a preceding • or •). Therefore, instances of these role universals are likewise identified as social roles. The remaining question is in which way the corresponding functions are related to the trial society. These functions/areas of responsibility provide additional information, and it seems that each function is best understood as a PROCESS UNIVERSAL whose instances comprise several subprocesses of a clinical trial, among them the tasks assigned to each function. For instance, the `Databases` area of responsibility comprises such tasks (i.e., subprocesses of a clinical trial) as the `construction of the database`, its `validation`, `data security measures` and so forth. These subprocesses can be distinguished from each other relatively well, and, because they are also social processes, subsocieties of the trial society emerge which are fairly independent, in particular as concerns their agentive social substances. Now, one can assign the `DB-Administrator` to the society based on the `Databases` process universal, which is thus a more specific assignment than identifying it as a role of the trial society.

Having just mentioned *agentive social substances*, these point to another aspect of the terms given in Table 5.2: these terms are mainly derived from processes. Take the area of responsibility named `On site monitoring` with its unique role type `Monitor` as an example. Of course, there is also a processual role universal $\texttt{Monitor}_{PR}$, which corresponds to the *agent* role in the process universal `On site monitoring`. However, we do not think that it is this processual role universal which is contained in Table 5.2. Instead, it appears to be a major determinant of the social role `Monitor`, as it can be identified as the *function* of this social role (in the sense of section 4.1, i.e., every instance of `Monitor` should execute its function, which is nothing other than participating in a $\texttt{Monitor}_{PR}$ processual role). There is a severe danger of confusing these two notions, in particular if only tasks are given as further specification. In the case of the `Steering Committee`, at least some other properties are stated, like the `responsibility during the preparation of reports and publications` (cf. Table 5.3). In the framework of GOL, these properties correspond to MOMENTS which inhere in instances of the universal `Steering Committee`. However, this is not the case for `On site monitoring`, which seems to support the position that the term `Monitor` may refer to a processual role universal, because no special properties are specified. However, firstly, there are certainly some unspecified properties for roles like `Monitor` (like the `right to access certain data`) which allow one to recognise these roles as substances. Secondly, the fact that people are declared as a `Monitor` before they actually start acting as a $\texttt{Monitor}_{PR}$ speaks in favour of an understanding of the items of Table 5.2 as social roles.

Finally, an explanation for the underlining of several symbols in Table 5.2 has to be given. It refers to the observation that these terms do not seem to be specific to clinical trials. Rather, they appear to be required or desired *qualifications*. We will not examine whether qualifications may also be understood as (social) roles, because even if this was the case, these roles would be roles of a different context than the above-mentioned trial society (because, e.g., the notions of `Physician` or `Independent Statistician` would not be influenced by a participation in a clinical trial). Therefore, these terms in Table 5.2 should not be considered to be roles in the context of a clinical trial. Furthermore, they are the only terms which appear multiply with regard to different functions. Accordingly, they would require an additional specification of a function if they were to be used in the *target group* metadatum.

The answers to the questions posed at the beginning of this section shall now be summarised. Concerning the first question, only some items of Table 5.2 can be considered to refer to roles with a special meaning in the context of a clinical trial. These roles are preceded by ○ or • (without underlines), and have been identified as social roles. The emergent *trial society* can be named as a comprehensive context for these social roles. Secondly, asking for the connection between Table 5.1 and 5.2, a twofold answer can be given. Areas of responsibilities preceded by • correspond to fairly independent subprocesses of a clinical trial, thereby allowing for emergent subsocieties of the trial society which in turn serve as more specific contexts for the corresponding social roles. The remaining items of Table 5.1, preceded by ○, are better understood as pre-existing or declared social substances, which directly provide a more specific context for their corresponding roles. With this type of items we allude to the third question which asks for own MOMENTS of these *group* terms. A mainly negative answer has to be given on the basis of the contents of [SP06], apart from the names of these groups and one example in the case of the `Steering Committee`. The primary reason for this answer is that these group terms are also mainly defined via their function in the clinical trial, and thus the difference in naming mainly seems to indicate different degrees of interaction among the members of each of these subsocieties of the trial society. This assessment notwithstanding, in practice each group of Table 5.1 will have its own MOMENTS like addresses, phone numbers, and the like[96].

Table 5.4 presents a reorganised version of the assignment between functions and roles from [SP06], supplemented by an additional column for the groups executing these functions. This adds precision, although it may appear redundant in several rows. The terms in the *Group* column were consciously adopted from Table 5.1, since these are already established in the medical community. The *Group* column is the most appropriate one to accommodate all terms from Table 5.1. Without regard to such restrictions, one would better have an additional “Group” at the end of those entries of this column which do not already contain an analogous term like “Committee”, “Board”, et cetera. In the *Function* column of Table 5.4 new terms (written in italics) have also been admitted, either because they are missing in Table 5.1, or if they seem to describe the corresponding intention more precisely. Note that there are still differences in the relationship between these functions and groups. In particular, the last two groups, `Trial Centre` and `Reference Institution`, can multiply occur within a clinical trial (cf. multi-centre clinical trials). In the `Trial Centre` case, the roles of each trial centre remain reasonable for the overall community with the function of *Patient Servicing*. In contrast, the term `Reference Institution` from Table 5.1 may either be understood analogously to `Trial Centre`, because there may be several `Reference Institutions` with their own roles in a clinical trial, or `Reference Institution` itself might be better interpreted as a role. More precisely, it seems that usually pre-existing institutions or groups play a role which is associated with the function of providing reference examinations, whereas the internal structure (in particular, their roles) of these institutions is not relevant.

The last column contains a selection of terms from Table 5.2, namely those which have been identified above as social roles of the trial society[97]. First of all, note that we assume for every *Group* that its

[96]This can be seen, for instance, in clinical trial protocols where the contact persons of such groups and their addresses are listed, with regard to one specific trial.

[97]The remaining items from Table 5.2 are omitted in Table 5.4. This mainly refers to the distinction between qualifications and roles from above. We suggest providing information on required or possible qualifications for each role independently. In contrast, [SP06] only contains the general statement that functions have to be executed by qualified persons. Further, the structure of some groups (e.g., `Advisory Board`) is based on qualifications of the members instead of having underlying

Function	**Group**	**Specialised Role**
Leading	Steering Committee	Principal Investigator*
Advice	Advisory Board	
	Independent Data Monitoring Committee	
	Ethics Committee	
Coordination	Coordination	Coordinating Investigator*
		(Secretary)
Counselling	Clinical Trial Counselling Service	(Counselling Investigator*)
Data Management	Data Management	Documentalist*
		Documentation Assistant
		Data Assistant
Biometrics	Biometrics	(Biometrician*)
Database Administration	Databases	DB-Administrator*
		DB-Programmer
On site Monitoring	On site Monitoring	(Monitor*)
Reference Examination	(Reference Institution)	
Patient Servicing	(Trial Centre)	Investigator
		Clinical Trial Assistant

Table 5.4: Reorganised Selection of Functions, Groups, and Roles in Clinical Trials.

roles can be addressed by the neutral formulation “member of”, as in “all members of the `Steering Committee`”. This explains the empty fields in the last column, because for these groups no special roles could be found in [SP06]. This neutral formulation does not convey additional information, which is assumed to be the case for those terms not occurring in parentheses, instead. In contrast, terms in parentheses are considered to be defined only by their membership in a corresponding group. For example, each member of the group `Biometrics` is called `Biometrician`, each member of `On site Monitoring` is called `Monitor`. On the other hand, not every member of the `Steering Committee` can be called `Principal Investigator`. The awareness of this difference has implications, for instance, if variations in the naming of certain universals occur. Illustratively, if the members of `On site monitoring` are to be distinguished into `Monitor` and `Monitoring Assistant` at some point in the future, and the equivalence between `Monitor` and “member of `On site Monitoring`” is not deleted, this can lead to false conclusions or erroneous behaviour of software in which this equivalence is hard-coded.

Finally, a star (*) attached to a term indicates that this term is likewise used to denote such “single” roles as discussed above in the case of the two notions of a `Principal Investigator`. That means, these terms are used for roles which do not belong to a group, but have the function of the corresponding row directly assigned to them.

differences in properties internal to the group. This is considered a different kind of information, and it is therefore not maintained in Table 5.4.

```
TargetGroup ::= Entry (, Entry)*
      Entry ::= GroupEntry | RoleEntry

 GroupEntry ::= (All members of the Group [except for RoleList;]) |
                (DerivedRole [except for RoleList;])
      Group ::= BasicGroup | ComplexGroup

   RoleList ::= RoleEntry (, RoleEntry)*
  RoleEntry ::= BasicRole | DerivedRole
```

Table 5.5: Syntax for the *Target Group* Metadatum of SOPs.

5.2.3 A Proposal for Target Group Specification

On the basis of our evaluation of [SP06] we can now return to the actual goal of this examination, a clarified vocabulary of terms for the *target group* metadatum of SOPs. It appears most appropriate to allow one to use terms of the *Group* column from Table 5.4 as well as from the *Specialised Role* column. A formal grammar for the syntax of the entries of the *target group* field is defined in Table 5.5 using EBNF[98]. This syntax requires further definition of basic and derived roles as well as of basic and complex groups as terminal symbols of the corresponding non-terminals, e.g., `BasicRole`. [SP06] partially serves this purpose. Assuming the reorganisation of roles and groups in Table 5.4, the syntax can be completed with the following assignments.

- Define the terms of Table 5.4, column *Group*, as the selection of terminals for `BasicGroup`.
- Define all terms without parentheses from Table 5.4, column *Specialised Role*, as the selection of terminals for `BasicRole`.
- Define all terms with parentheses from Table 5.4, column *Specialised Role*, as the selection of terminals for `DerivedRole`.

Note that no `ComplexGroup` has been defined so far. Indeed, Table 5.4 does not contain groups ordered by an aggregation hierarchy. However, looking at several SOPs of the KKSL, one can find terms as entries of the metadatum *target group* which do not occur in [SP06], e.g., the term `Trial Office` (cf. [Kropf, 2001]). The structure of the trial office may vary for different trials, but usually it comprises several of the basic groups denoted by the terminals of `BasicGroup`. Accordingly, there is a relationship between the elements `BasicGroup` and `ComplexGroup`, and this should be explicitly captured in semantic constraints for these syntactic elements, thereby providing explicit knowledge of the domain. They can further be employed to express further restrictions on admissible entries, as a basis for consistency tests during the input of real-world data, or for powerful query processing with underlying background knowledge.

We propose the constraints below as a minimal account of a semantic supplement to the grammar in Table 5.5. $\mathscr{I}$ should be read like an interpretation function in this connection, mediating between the syntactic elements from above and real-world entities. In contrast to their specificity, these constraints contain a number of issues which cannot be decided herein, but they are subject to harmonisation

[98] Extended Backus Naur Form. For readability, different styles are used for `non-terminals` and *terminals*.

processes in the KML, for example. The next step for the SOP-Creator application is to translate the grammar and the constraints into an XML format which allows one to easily verify the validity of data. On this basis further tasks can be realised, e.g., uncovering redundant *target group* entries as well as a presentation in a normalised form.

(C1) Each `BasicGroup` g_{basic} denotes a SOCIAL SUBSTANCE UNIVERSAL $\mathscr{I}(g_{basic})$.

(C2) Each `BasicRole` q_{basic} denotes a SOCIAL ROLE UNIVERSAL $\mathscr{I}(q_{basic})$. There is a distinctive function which contributes to the characterisation of $\mathscr{I}(q_{basic})$, and which is executed by its instances. For this reason, among others, if q'_{basic} is another `BasicRole`, then $\mathscr{I}(q_{basic})$ and $\mathscr{I}(q'_{basic})$ have disjoint extensions.

(C3) Each `BasicRole` q_{basic} is associated with exactly one `BasicGroup` $C(q_{basic})$, called its context. Each instance of $\mathscr{I}(q_{basic})$ has to be a role of an instance of $\mathscr{I}(C(q_{basic}))$, while each instance of a `BasicGroup` g_{basic} needs to consist of roles such that each of these roles instantiates exactly one universal denoted by an item of `BasicRole`.

(C4) Let $g_{complex}$ be a `ComplexGroup`, which denotes a SOCIAL SUBSTANCE UNIVERSAL $\mathscr{I}(g_{complex})$. Then there is a set of items of `BasicGroup` $P(g_{complex})$ related to $g_{complex}$, such that each instance of $\mathscr{I}(g_{complex})$ can be dismantled into a set of PARTS, each of which is an instance of exactly one universal denoted by an element of $P(g_{complex})$[99].

(C5) For each `DerivedRole` $q_{derived}$ there is exactly one `Group` $G(q_{derived})$ such that each entity is an instance of the SOCIAL ROLE UNIVERSAL $\mathscr{I}(q_{derived})$ iff this entity is a role of some instance of $\mathscr{I}(G(q_{derived}))$.

Let us summarise the major arguments for this combined specification of the *target group* metadatum. The combination is mainly owed to a natural, expressive and efficient representation. A restriction to roles only is also possible, but this leads to the artificial creation of names which are not common in natural language. The “member-of” construction covers all these cases in a more elegant way. The decision to neglect the functions from Table 5.4 is based on two facts. Firstly, groups and functions are highly interdependent, mirroring each other. Secondly, groups are more flexible with regard to compositions. Further, some examples suggest that not all groups are formed on the basis of functions. In contrast, taxonomic hierarchies among functions can be reflected to a certain extent by the part-whole structure of the corresponding groups.

[99]This is an example of a constraint which is questionable in nature. It may turn out to be too restrictive, because one may wish to add single roles to a complex group which do not belong to any of the basic groups which constitute the complex group.

6 Conclusion

6.1 Summary of Results

The examinations in this thesis have aimed at a comprehensive characterisation of the notion of *roles*, a concept which occurs frequently in knowledge representation and conceptual modelling. Our approach was mainly motivated by the development of the *General Ontological Language* (GOL), in which roles have thus far been integrated only in a preliminary form. On the basis of a substantial literature review we have proposed a classification of roles which was subsequently applied to a number of important formalisms to demonstrate its expressiveness. Finally, concrete proposals for handling roles within GOL were made, followed by a sample application of the role classification in the field of clinical trials.

Let us recapitulate the important issues addressed herein in some detail, in order to adumbrate the whole picture. Chapter 2 introduced *roles* as they are touched on in a variety of formalisms. With hindsight, the approaches of Guarino and Sowa in knowledge representation, as well as UML in conceptual modelling are the most influential ones. There are several works which are directly based on the former, while UML exhibits a quasi-standard status in software engineering and, as scrutinised in chapter 4, covers all role types discussed herein with reservations concerning a number of details. In connection with all works reviewed, two basic notions were found to be closely related to roles, which we have called *player*[100] and *context*. These two concepts paved the way for our role classification, which was developed in the third chapter.

There, we have argued that most of the approaches reviewed mention only (if at all) the context of a role, while the relationship between role and player is at the centre of interest, in combination with the question of which role-independent type(s)[101] players are recruited from. This was called the *player-centred* access to roles. On the other hand, three top-level categories of GOL were identified to serve as context in all role approaches, i.e., in most formalisms one out of these three categories is assumed. Therefore, the so-called *context-centred* approach to roles advances the underlying hypothesis that contexts are a major determinant of roles, while roles determine players. Following this approach, we have focussed on the relationship between roles and each of the GOL categories RELATOR, PROCESS, and (SOCIAL) SUBSTANCE as contexts, which has led to the notions of *relational*, *processual*, and *social roles*, respectively. Afterwards, common features of roles in general were extracted. It turned out that roles are a highly flexible notion, and one should not necessarily consider the classification presented herein exhaustive. If something is considered a context, it should not be difficult in most cases to discover corresponding roles. For example, a new role type may arise from SITUOIDS. This generality is not considered unfavourable as it allows one to understand and explain various differing approaches. Analogously, the notion of parts is similarly flexible and general. Note in this connection that some properties of roles point in the direction of the idea that roles are a specialisation of parts[102].

In addition to the relationship to their contexts, some implications and open questions for each subtype of roles have been discussed. First of all, these role types are spread over the current GOL top-level hierarchy. Accordingly, relational roles are best viewed as a new subcategory of ENDURANT, processual roles are PROCESSES, and social roles belong to SOCIAL SUBSTANCES, a subcategory

[100] Also referred to as “filler” in many approaches.

[101] Often called “natural type”.

[102] If so, contexts would appear as a specialisation of *wholes*.

of SUBSTANCES[103]. Relational roles were shown to be beneficial with respect to problems which arise from the usual representations of MATERIAL RELATIONS in terms of ordered tuples, as well as regarding the representation of anadic relations. Further, relational roles allow for fine-grained distinctions, as demonstrated with the notions of *extensional* and *intensional symmetry* of relations in section 3.3.4. With regard to processual roles, the connection between temporal parts of these roles and their contexts was discussed, which results in strong parallels between roles and processes. Moreover, the use of several FORMAL RELATIONS to represent different ways of participation in a process was rejected, but an understanding of thematic roles from linguistics in terms of high-level processual role universals was introduced. Finally, some questions with respect to social roles have been answered, whereas others could only be identified herein. Among the former is a solution to the problem arising from the transitivity of PART-OF, applied to a social substance and a non-social substance.

With this theory of roles at hand, some of the formalisms reviewed were reconsidered in order to assign one or several role types to the notion(s) of roles contained in a formalism (cf. section 4.2). On the other hand, some deficiencies could be shown for the case that (actually specialised) notions of roles from the literature are viewed as referring to the general concept of roles. To mention only one example, there are singular processual roles, which do not directly refer to further roles. In contrast, relational roles are always associated with some other relational role, thus viewing relational roles as the general concept of roles leads to an assumption which is untenable in the theory elaborated herein. In order to classify the most important works with respect to our categorisation, Guarino's as well as Sowa's approaches to roles were identified to refer to relational roles, while UML contains all role types although spread over different diagram types without being connected to each other. In addition to such questions of individual formalisms, roles were distinguished from the notions of function, state, quality, view, and interface in section 4.1, because these terms are sometimes used as synonyms for roles, or roles are considered to be the conceptual modelling counterpart of the implementational concept of interfaces.

The application of this work (cf. chapter 5) is twofold. The theoretical application is a direct reply to our motivation, namely a discussion of how roles should be introduced in GOL. We propose the inclusion of the axioms developed in chapter 3 in GOL, but the insertion of a new top-level category of roles in the "backbone" top-level hierarchy does not appear fruitful because it is orthogonal to the classification criteria in this "backbone". Additionally, relational roles are suggested as a substitute for the formal relation of HOLDING in the definition of MATERIAL FACTS, since these roles provide more flexibility and solve the problems with ordered tuples discussed in section 3.3.2. The second part of the application, which is an analysis in the domain of clinical trials, illustrates the way in which roles and some of the assumptions given in this thesis can help to improve the representation of knowledge, based on explicit ontological commitments. In particular, a collection of terms originally viewed as "positions" is examined and reorganised, thereby clarifying the core of the definitions of these terms and separating different domain-specific role classes.

[103] In spite of these classifications, we do not suggest inserting these categories into the "backbone" of the GOL top-level hierarchy. See section 5.1.

Emphasising the major results and the main features of roles in a concise way, the following points can be summarised.

- Roles are determined by a context, and themselves determine their players.
- Most of the literature focuses on the relationship between role and player, tacitly assuming a single context category.
- On the basis of different context categories there are at least the following role subtypes:
  1. *Relational roles*, which are ENDURANTS with RELATORS as contexts,
  2. *Processual roles*, which are PROCESSES with PROCESSES as contexts, and
  3. *Social roles*, which are SOCIAL SUBSTANCES with contexts that are also SOCIAL SUBSTANCES.
- Each role is an INDIVIDUAL.
- Roles are different entities compared to *functions*, *states*, and *qualities*, and they should not principally be merged with the concept of an *interface* in programming.
- In general, all approaches in chapter 2 can be explained in terms of the theory herein, but usually with some differences in the details.
- Roles, in combination with the formal relations PLAYS and ROLE-OF, are proposed to replace the HOLDING relation in GOL.

6.2 Future Research

The present analysis has concentrated on the single modelling element of *roles*. In spite of this focus, several questions had to remain open, because their solution would have led to discussions the specificity of which would have conflicted with the broad scope of the present work. The literature review was kept deliberately broad in order to discover as many variants of roles as possible, and to develop a general and comprehensive theory. Apart from issues only referring to roles and their features, questions of the combination with further modelling elements have to be treated, since roles can hardly be applied in isolation. According to all of this, a considerable amount of work remains which could reasonably extend this thesis.

Let us start with the ontological framework of GOL. First of all, if the axioms of chapter 3 are included in GOL, it needs to be shown that the resulting axiomatic system is consistent. It is expected that consistency can be shown, as this work mainly relies on the two formal relations PLAYS and ROLE-OF which are not contained in GOL yet. However, there is some interplay with current GOL notions, such as the PART-OF relation or PROCESSES. In addition, it may be necessary to reconsider one or the other axiom on roles as GOL is still changing. Some points were mentioned in chapter 3 where minor adaptations of the existing framework are desirable from the point of view of roles. Furthermore, the development and extension of GOL, e.g., to notions like causality and intentionality, will require an analysis of the influences on the notions of roles presented. At first glance, for instance, causality may lead to new considerations about processual roles, in particular the processual role account of thematic roles.

Having noted thematic roles, the approach of interpreting these as processual role universals should be pursued in order to find further evidence for this interpretation. This may likewise be a valuable activity for linguists, not at least in connection with the actual use of theta roles as an interface between syntax and semantics. In the ontological area, it appears interesting to further examine Sowa's account on theta roles.

If the focus for extensions is set on the classification and axiomatisation, one may first tackle the question of the exhaustiveness of the classification presented. In chapter 3, it was mentioned that the notion of SITUOIDS is neglected in our considerations due to its early state of development. The category of situoids may also be a category providing contexts for roles, although this appears very similar and interlinked with PROCESSES and processual roles. Additionally, the axiomatisation of each role type may be further extended. Firstly, one may capture the relationship between players and roles in greater detail, where the literature can still be exploited. Secondly, several open questions with respect to social roles were mentioned in chapter 3, e.g., the question of role-playing roles or the relationship to other social entities. Interestingly, social roles may be the most important and most complex subtype of roles from a practical point of view, which can be seen in the application in section 5.2. The interconnection between roles of different types is closely related to this issue. The example of the natural language term `student`, which can refer to three role universals each of which belongs to one of the identified role subtypes, was given in section 3.6.3. It is possible that the connections between these role universals can be stated in general, because analogies to this example seem to hold for a considerable number of role universals. This is also clear from the application in section 5.2. As this issue is connected with different levels of abstraction, examinations advancing this problem may help to elaborate representations which can smoothly follow shifts in focus and level of details. A final theoretical problem which is not sufficiently addressed herein is the transition between statements on role universals and role individuals. Section 3.3.3 discusses some difficulties with this point, which is hardly addressed in the literature[104]. Most modelling formalisms specify relations between role universals instead of role individuals, which is appropriate for the efficiency of a modelling language, but the attempts made herein to cover the intuitively immediately clear implications either failed or did not provide the flexibility for which role individuals allow. On the whole, this issue is considered as one of great importance.

Last but not least, the theory established in this thesis needs to be implemented in software systems, which will provide feedback on the utility of the notions developed. For instance, the grammar and the constraints on the contents of *target group* entries in Standard Operating Procedures, which were developed in section 5.2, can be used as a basis for an XML representation, which will enable the software to automatically check the semantic validity of these entries to a certain degree, when they are stored in the database. Of course, this not a direct implementation of the role classification, on the other hand. The latter demands several addenda, which span from intuitive syntactic abbreviations of what would be long formulae with similar structures in predicate logic to evaluations of the complexity classes of algorithms for certain problems, like drawing inferences. The latter is considered to refer to the distant future, as we agree with the GOL approach of establishing an ontology without representational restrictions first, before complexity concerns determine what has to be excluded to maintain computational tractability. The relevance of the former point regarding syntactic abbreviations should

[104] This is due to the fact that many approaches only consider roles as universals, thereby excluding this problem from the outset.

not be underestimated either, as this increases the acceptance of models by people only using these models, but not developing them. If these people can read ontological models, they can provide hints which may help to uncover deficiencies.

Obviously, there are many directions to extend this work. The most important one from our point of view is the integration of the axiomatisation into GOL, as well as a proof of consistency. Further, several domain models should be built in order to test the current work, before additional axioms are devised. In spite of the remaining work, we hope that the ideas laid down in this thesis provide an expressive, integrative and extendible approach to roles, which includes beneficial aspects of various specialised works on this subject.

References

Aitken S, **Curtis J**, **2002**. *A Process Ontology*. In: Gómez-Pérez A, Benjamins VR, (editors), Proceedings of the 13th International Conference on Knowledge Engineering and Knowledge Management (EKAW'02). Berlin: Springer, vol. 2473 of *Lecture Notes in Artificial Intelligence*, pp. 108–113.

Baader F, **Calvanese D**, **McGuinness D**, **Nardi D**, **Patel-Schneider P**, (editors) **2003**. The Description Logic Handbook: Theory, Implementation and Applications. Cambridge (UK): Cambridge University Press.

Bachman CW, **Daya M**, **1977**. *The Role Concept in Data Models*. In: Proceedings of the Third International Conference on Very Large Databases. IEEE Computer Society, pp. 464–476.

Balzert H, **1999**. Lehrbuch der Objektmodellierung: Analyse und Entwurf. Lehrbücher der Informatik. Berlin: Spektrum Akademischer Verlag.

Barker C, **Dowty D**, **1992**. *Non-verbal Thematic Proto-roles*. In: Schafer A, (editor), Proceedings of North East Linguistic Society (NELS). Amherst: GSLA, vol. 23, pp. 49–62.

Barwise J, **Perry J**, **1983**. Situations and Attitudes. Bradford Book. Cambridge (US, MA): MIT Press.

Bäumer D, **Riehle D**, **Siberski W**, **Wulf M**, **2000**. *Role Object*. In: Harrison N, Foote B, Rohnert H, (editors), Pattern Languages of Program Design 4, Reading (US, MA): Addison-Wesley, chap. 2, pp. 15–32.

Berners-Lee T, **Hendler J**, **Lassila O**, **2001**. *The Semantic Web*. Scientific American 284(5):34–43.

Biddle BJ, **Thomas EJ**, (editors) **1966**. Role Theory: Concepts and Research. New York: John Wiley and Sons.

Booch G, **Rumbaugh J**, **Jacobson I**, **1999**. The Unified Modeling Language User Guide. Object Technology Series. Reading (US, MA): Addison Wesley.

Brachman RJ, **1985**. *On the Epistemological Status of Semantic Networks*. In: Brachman RJ, Levesque H, (editors), Readings in Knowledge Representation, Los Altos (US, CA): Morgan Kaufman Publishers, chap. 10, pp. 191–215.

Bray T, **Paoli J**, **Sperberg-McQueen CM**, **Maler E**, (editors), **2000**. *Extensible Markup Language (XML) 1.0*. [Internet]. 2nd ed. [Cambridge (US, MA)]: World Wide Web Consortium (W3C); [cited 2003 Jan 21]. (W3C Recommendation). Available from: http://www.w3.org/TR/REC-xml.

Brosteanu O, **2001**. *Verantwortlichkeiten*. Leipzig: Coordination Centre for Clinical Trials Leipzig (KKSL). Standard Operating Procedure SP06, Version F1.0.

Bruel JM, **Bellaséhne Z**, (editors), **2002**. OOIS 2002 Workshops, vol. 2425 of *Lecture Notes in Computer Science*. Berlin: Springer.

Bunge M, **1977**. Ontology I: The Furniture of the World, vol. 3 of *Treatise on Basic Philosophy*. New York: D. Reidel Publishing.

Bunge M, **1979**. Ontology II: A World of Systems, vol. 4 of *Treatise on Basic Philosophy*. New York: D. Reidel Publishing.

Campbell SJ, **1999**. *Role Theory, Foreign Policy Advisors, and U.S. Foreign Policymaking*. [Internet]. [New York]: Columbia International Affairs Online, Columbia University Press; [cited 2002 Jul 15]. (Prepared for presentation at the 40th annual meeting of the International Studies Association, Washington, D.C.). Available from: http://www.ciaonet.org/isa/cas01.

Chen PPS, **1976**. *The Entity-Relationship Model: Toward a Unified View of Data*. ACM Transactions on Database Systems 1(1):9–36.

Cocchiarella NB, **1991**. *Formal Ontology*. In: Burkhardt H, Smith B, (editors), Handbook of Metaphysics and Ontology, Munich: Philosophia Verlag, pp. 640–647.

Costanza P, **2000**. *Separation of Object Identity Concerns*. In: Workshop on Aspects and Dimensions of Concerns at ECOOP 2000, Cannes, France, June 2000.

Coulondre S, **Libourel T**, **2002**. *Towards a New Role Paradigm for Object-Oriented Modeling*. In: [Bruel and Bellaséhne, 2002], pp. 44–52.

Cowan J, (editor), **2002**. *Extensible Markup Language (XML) 1.1*. [Internet]. [Cambridge (US, MA)]: World Wide Web Consortium (W3C); [cited 2003 Jan 21]. (W3C Candidate Recommendation). Available from: http://www.w3.org/TR/xml11/.

Cranefield S, **Purvis M**, **Nowostawski M**, **Hwang P**, **2002**. *Ontologies for Interaction Protocols*. In: Proceedings of the Second International Workshop on Ontologies in Agent Systems, Bologna, Italy, 15-19 July, 2002 [Internet]. [Dunedin (NZ)]: University of Otago, Department of Information Science; [cited 2002 Nov 30]. Available from: http://cis.otago.ac.nz/OAS2002/Papers/oas02-16.pdf.

Dahchour M, **Pirotte A**, **Zimányi E**, **2002**. *A Generic Role Model for Dynamic Objects*. In: Pidduck AB, Mylopoulos J, Woo C, Ozsu MT, (editors), Proceedings of the 14th International Conference on Advanced Information Systems Engineering, CAiSE'02, Toronto, Canada, May 2002. Berlin: Springer, vol. 2348 of *Lecture Notes in Computer Science*, pp. 643–656.

De Giacomo G, **2001**. *Description Logics: Modal Logics for Class-Based Modeling*. In: Wolter F, Wansing H, de Rijke M, Zakharyaschev M, (editors), Advances in Modal Logic. Stanford: CSLI Publications, vol. 3.

DeArmond RC, **2002**. *Theta Roles*. [Internet]. [Burnaby (CA)]: Simon Fraser University, Department of Linguistics; [cited 2002 Dec 15]. Available from: http://www.sfu.ca/person/dearmond/322/322.theta.roles.htm.

Degen W, **Heller B**, **Herre H**, **2001**a. Contributions to the Ontological Foundation of Knowledge Modelling. Technical Report No. 2, Department of Computer Science, University of Leipzig.

Degen W, **Heller B**, **Herre H**, **2003**. General Ontological Language (GOL) Reference Manual, Version 1.0α. University of Leipzig. Forthcoming. Preliminary version of 2003 Apr 07.

Degen W, **Heller B**, **Herre H**, **Smith B**, **2001**b. *GOL: A General Ontological Language*. In: [Welty and Smith, 2001], pp. 34–46.

Degen W, **Herre H**, **2001**. *Contributions to the Axiomatic Foundation of Upper-Level Ontologies*. In: Stumme G, Maedche A, Staab S, (editors) Proceedings of the ONTO-2001 Workshop on Ontologies, Vienna, Austria, September 18, 2001 [Internet]. [Aachen (DE)]: CEUR Workshop Proceedings; [cited 2002 Jul 20]. (vol. 48 of CEUR Workshop Proceedings). Available from: http://sunsite.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-48/herre_degen.pdf.

Depke R, **Engels G**, **Küster JM**, **2000**. On the Integration of Roles in the UML. Technical Report 214, University of Paderborn, Paderborn (DE).

Depke R, **Heckel R**, **Küster JM**, **2001**. *Improving the Agent-Oriented Modeling Process by Roles*. In: Müller JP, Andre E, Sen S, Frasson C, (editors), Proceedings of the Fifth International Conference on Autonomous Agents (AGENTS 2001). New York: ACM Press, pp. 640–647.

Dori D, **1995**. *Object-Process Analysis: Maintaining the Balance between System Structure and Behaviour*. Journal of Logic and Computation 5(2):227–249.

Dori D, **2002**. Object-Process Methodology: A Holistic Systems Paradigm. Berlin: Springer.

Dowty D, **1989**. *On the Semantic Content of the Notion of 'Thematic Role'*. In: Chierchia G, Partee BH, Turner R, (editors), Properties, Types, and Meaning II, Dordrecht: Kluwer, pp. 69–129.

Dowty D, **1991**. *Thematic Proto-roles and Argument Selection*. Language 67(3):547–619.

Evans A, **Stuart K**, **Selic B**, (editors), **2000**. UML 2000 - The Unified Modeling Language, Advancing the Standard, Proceedings of the Third International Conference, York, UK, October 2000, vol. 1939 of *Lecture Notes in Computer Science*. Berlin: Springer.

Fan J, **Barker K**, **Porter B**, **Clark P**, **2001**. *Representing Roles and Purpose*. In: Proceedings of the First International Conference on Knowledge Capture (K-CAP 2001). pp. 38–43.

Fillmore CJ, **1968**. *The Case for the Case*. In: Bach E, Harms RT, (editors), Universals of Linguistic Theory, New York: Holt, Rinehart, and Winston, pp. 1–88.

Fowler M, **1997**. *Dealing with Roles*. [Internet]. [Chicago]: Martin Fowler; [cited 2003 Jan 20]. Available from: http://www.martinfowler.com/apsupp/roles.pdf.

Gamma E, **Helm R**, **Johnson R**, **Vlissides J**, **1994**. Design Patterns: Elements of Reusable Object-Oriented Software. Professional Computing Series. Reading (US, MA): Addison-Wesley.

Gottlob G, **Schrefl M**, **Röck B**, **1996**. *Extending Object-Oriented Systems with Roles*. ACM Transactions on Information Systems 14(3):268–296.

Gruber T, **1993**. *A Translation Approach to Portable Ontologies*. Knowledge Acquisition 5(2):199–220.

Guarino N, **1992**. *Concepts, Attributes and Arbitrary Relations: Some Linguistic and Ontological Criteria for Structuring Knowledge Bases*. Data & Knowledge Engineering 8(3):249–261.

Guarino N, **1995**. *Formal Ontology, Conceptual Analysis and Knowledge Representation*. International Journal of Human-Computer Studies 43(5/6):625–640. (Special Issue: Formal Ontology in the Information Technology).

Guarino N, **Giaretta P**, **1995**. *Ontologies and Knowledge Bases: Towards a Terminological Clarification*. In: Mars NJI, (editor), Towards Very Large Knowledge Bases: Knowledge Building & Knowledge Sharing, Amsterdam: IOS Press, pp. 25–32.

Guarino N, **Welty CA**, **2000**. *A Formal Ontology of Properties*. In: Dieng R, Corby O, (editors), Proceedings of EKAW-2000: The 12th International Conference on Knowledge Engineering and Knowledge Management. Berlin: Springer, vol. 1937 of *Lecture Notes in Computer Science*, pp. 97–112.

Guarino N, **Welty CA**, **2001**. *Supporting Ontological Analysis of Taxonomic Relationships*. Data & Knowledge Engineering 39(1):51–74.

Guizzardi G, **Herre H**, **Wagner G**, **2002**a. *On the General Ontological Foundations of Conceptual Modeling*. In: Spaccapietra S, March ST, Kambayashi Y, (editors), Conceptual Modeling - ER 2002, Proceedings of the 21st International Conference on Conceptual Modeling. Berlin: Springer, vol. 2503 of *Lecture Notes in Computer Science*, pp. 65–78.

Guizzardi G, **Herre H**, **Wagner G**, **2002**b. *Towards Ontological Foundations for UML Conceptual Models*. In: Meersman R, Tari Z, (editors), On the Move to Meaningful Internet Systems, 2002 - DOA/CoopIS/ODBASE 2002 Confederated International Conferences DOA, CoopIS and ODBASE 2002 Irvine, California, USA, October 30 - November 1, 2002, Proceedings. Berlin: Springer, vol. 2519 of *Lecture Notes in Computer Science*, pp. 1100–1117.

Hay DC, **1996**. Data Model Patterns: Conventions of Thought. New York: Dorset House Publishing.

Hay DC, **1999**a. *Object Orientation and Information Engineering: Object and Data Modeling*. The Data Administration Newsletter [Internet]. [cited 2002 Aug 8];(8). Available from: http://www.tdan.com/i008fe02.htm.

Hay DC, **1999**b. *There is No Object-Oriented Analysis*. Data To Knowledge Newsletter [Internet]. [cited 2002 Aug 12];27(1). Available from: http://www.essentialstrategies.com/documents/ooa.pdf.

HL7, **2002**a. *About HL7*. [Internet]. [Ann Arbor (US, MI)]: Health Level Seven, Inc.; [cited 2002 Nov 28]. Available from: http://www.hl7.org/about/.

HL7, **2002**b. *HL7 Reference Information Model (RIM) Version 1.21*. [Internet]. [Ann Arbor (US, MI)]: Health Level Seven, Inc.; [cited 2002 Nov 10]. Available from: http://www.hl7.org/library/data-model/RIM/C30121/rim.htm.

Howes TA, **Smith MC**, **Good GS**, **1999**. Understanding and Deploying LDAP Directory Services. Macmillan Network Architecture and Development Series. [Indianapolis (US, MA)]: Macmillan Technical Publishing.

ICH, **1996**. ICH Topic E6 Guideline for Good Clinical Practice. Step 5, Consolidated Guideline 1.5.96, The European Agency for the Evaluation of Medicinal Products (EMEA), London. (Post Step Errata: July 2002).

Iglesias C, **Garrijo M**, **Gonzalez J**, **1999**. *A Survey of Agent-Oriented Methodologies*. In: Müller J, Singh MP, Rao AS, (editors), Proceedings of the 5th International Workshop on Intelligent Agents V : Agent Theories, Architectures, and Languages (ATAL-98). Berlin: Springer, vol. 1555 of *Lecture Notes in Computer Science*, pp. 317–330.

ISO, **1999**. ISO/IEC 13250 – Topic Maps. Standard 13250, International Organization for Standardization (ISO), International Electrotechnical Commission (IEC).

Jackendorff R, **1991**. Semantic Structures. No. 18 in Current Studies in Linguistics. Cambridge (US, MA): MIT Press.

Jähnichen S, **Herrmann S**, **2002**. *Was, bitte, bedeutet Objektorientierung?* Informatik Spektrum 25(4):266–276.

Johansson I, **1989**. Ontological Investigations: An Inquiry into the Categories of Nature, Man and Society. New York: Routledge.

Kaplan AN, **2001**. *Towards a Consistent Logical Framework for Ontological Analysis*. In: [Welty and Smith, 2001], pp. 244–255.

Kendall EA, **1999**. *Role Modelling for Agent System Analysis, Design, and Implementation*. In: First International Symposium on Agent Systems and Applications ASA'99, Third International Symposium on Mobile Agents (MA'99), Palm Springs, October, 1999 [Internet]. [Frankston (AU)]: Monash University, School of Network Computing; [cited: 2002 Nov 20]. Available from: http://www.pscit.monash.edu.au/~kendall/selectedpapers.htm.

Kristensen BB, **Østerbye K**, **1996**. *Roles: Conceptual Abstraction Theory and Practical Language Issues*. Theory and Practice of Object Systems 2(3):143–160.

Kropf S, **2001**. *Behandlung fehlender und implausibler Daten*. Leipzig: Coordination Centre for Clinical Trials Leipzig (KKSL). Standard Operating Procedure BI09, Version F1.0.

Lenat DB, **Guha RV**, **1990**. Building Large Knowledge-Based Systems. Reading (US, MA): Addison-Wesley.

Luger GF, **2001**. Künstliche Intelligenz: Strategien zur Lösung komplexer Probleme, vol. 7002 of *Serie informatik*. Munich: Pearson Studium, 4. ed.

Marcos E, **Cavero JM**, **2002**. *Hierarchies in Object-Oriented Conceptual Modeling*. In: [Bruel and Bellaséhne, 2002], pp. 24–33.

Masolo C, **Borgo S**, **Gangemi A**, **Guarino N**, **Oltramari A**, **Schneider L**, **2002**. Wonderweb Deliverable D17. Preliminary Report Version 2.0, ISTC-CNR, Padova (IT).

Mehl C, **2002**. *SOP-Struktur, Erstellung und Änderung*. Cologne: Competence Network for Malignant Lymphoma (KML). Standard Operating Procedure.

Mulligan K, **1998**. *Relations: Through Thick and Thin*. Erkenntnis 48:325–353.

Oestereich B, **1998**. Objektorientierte Softwareentwicklung : Analyse und Design mit der Unified Modeling Language. Munich: Oldenbourg Verlag, 4. rev. ed.

OMG, **2001**. OMG Unified Modeling Language Specification. Version 1.4, Object Management Group, [Needham (US, MA)].

Parsons T, **1990**. Events in the Semantics of English: A Study in Subatomic Semantics. No. 19 in Current Studies in Linguistics. Cambridge (US, MA): MIT Press.

Peleg M, **Dori D**, **1998**. *Specifying Reactive Systems through the Object-Process Methodology*. In: Proceedings of the 1998 International Conference And Workshop: Engineering of Computer Based Systems (ECBS'98). IEEE Computer Society Press, pp. 29–36.

Peleg M, **Dori D**, **2000**. *The Model Multiplicity Problem: Experimenting with Real-Time Specification Methods*. IEEE Transactions on Software Engineering 28(8):742–759.

Pepper S, **Moore G**, **2001**. *XML Topic Maps (XTM) 1.0*. [Internet]. [place unknown]: TopicMaps.Org; [cited 2002 Dec 20]. TopicMaps.Org Specification. Available from: http://www.topicmaps.org/xtm/1.0/.

Pocock SJ, **1983**. Clinical Trials: A Practical Approach. Chichester: John Wiley & Sons.

Reenskaug T, **2001**. *Modeling Systems in UML 2.0: A Proposal for a Clarified Collaboration*. [Internet]. [Oslo]: University of Oslo; [cited 2002 Nov 13]. Available from: http://folk.uio.no/trygver/documents/2001/uml-20.pdf.

Rumbaugh J, **Jacobson I**, **Booch G**, **1999**. The Unified Modeling Language Reference Manual. Object Technology Series. Reading (US, MA): Addison Wesley.

Schank RC, **1972**. *Conceptual Dependency: A Theory of Natural Language Understanding*. Cognitive Psychology 3(4):532–631.

Schank RC, **Abelson R**, **1977**. Scripts, Plans, Goals and Understanding. Hillsdale (US, NJ): Erlbaum.

Schmidt-Schauß M, **Smolka G**, **1991**. *Attributive Concept Descriptions with Complements*. Artificial Intelligence 48(1):1–26.

Searle JR, **1969**. Speech Acts: An Essay in the Philosophy of Language. Cambridge (UK): Cambridge University Press.

Searle JR, **1995**. The Construction of Social Reality. New York: Free Press.

Smith B, **2003**. *John Searle: From Speech Acts to Social Reality*. In: Smith B, (editor), John Searle, Cambridge (UK): Cambridge University Press, Contemporary Philosophy in Focus, chap. 1. Forthcoming. See announcement, available from: http://books.cambridge.org/0521797047.htm [cited 2003 Jan 15].

Sowa JF, **2000**. Knowledge Representation: Logical, Philosophical and Computational Foundations. Pacific Grove: Brooks/Cole.

Steimann F, **2000**a. *Eine Grundlegung des Rollenbegriffs für die objektorientierte Modellierung mit dem Vorschlag einer Änderung von UML*. In: Ebert J, Frank U, (editors), Modelle und Modellierungssprachen in Informatik und Wirtschaftsinformatik: Modellierung 2000, Koblenz (DE): Fölbach-Verlag, pp. 55–69.

Steimann F, **2000**b. Formale Modellierung mit Rollen. Habilitation, Faculty of Electrical Engineering and Information Technology, University of Hannover.

Steimann F, **2000**c. *On the Representation of Roles in Object-Oriented and Conceptual Modelling*. Data & Knowledge Engineering 35(1):83–106.

Steimann F, **2000**d. *A Radical Revision of UML's Role Concept*. In: [Evans et al., 2000], pp. 194–209.

Steimann F, **2001**. *Role = Interface: a merger of concepts*. Journal of Object-Oriented Programming 14(4):23–32.

Tamma VAM, **2001**. An Ontology Model supporting Multiple Ontologies for Knowledge Sharing. PhD Thesis, University of Liverpool.

Wand Y, **Storey VC**, **Weber R**, **1999**. *An Ontological Analysis of the Relationship Construct in Conceptual Modeling*. ACM Transactions on Database Systems 24(4):494–528.

Wegmann A, **Genilloud G**, **2000**. *The Role of "Roles" in Use Case Diagrams*. In: [Evans et al., 2000], pp. 210–224.

Welty C, **Smith B**, (editors), **2001**. Formal Ontology in Information Systems: Collected Papers from the Second International Conference. New York: ACM Press.

Wieringa R, **de Jonge W**, **Spruit P**, **1994**. *Roles and Dynamic Subclasses: A Modal Logic Approach*. In: Tokoro M, Pareschi R, (editors), Proceedings of the 8th European Conference on Object-Oriented Programming, ECOOP'94. Berlin: Springer, vol. 821 of *Lecture Notes in Computer Science*, pp. 32–59.

Wooldridge M, **Jennings NR**, **Kinny D**, **2000**. *The Gaia Methodology for Agent-Oriented Analysis and Design*. Autonomous Agents and Multi-Agent Systems 3(3):285–312.

Acknowledgements

I am grateful to my supervisor, Professor Heinrich Herre at the Institute of Computer Science, University of Leipzig, who has fanned the flames of my interest for ontology, as well as for theoretical issues of computer science in general. He guided and supported the search for the topic of this thesis and was always prepared for productive and stimulating ontological discussions.

The Onto-Med research group at the Institute for Medical Informatics, Statistics and Epidemiology, University of Leipzig, first and foremost, Dr. Barbara Heller and Kristin Lippoldt, have provided a congenial working environment, for which I am thankful. Moreover, they have highlighted the needs of applications as well as problems which are easily overlooked when developing theories in the scientific ivory tower.

I thank Nguyen Thanh Hai for interesting discussions on various ontological questions, as well as Alexander Kovacs, who in addition supplied me with beneficial stylistic comments.

During the whole time of writing this thesis, Katrin Kühn has been a constant source of motivation and support. I thank her for this, as well as for her careful proof-reading of my last drafts.

Furthermore, I am deeply indebted to Dr. Evan Mellander for his willingness to discuss linguistics and language at any time, as well as his invaluable support in improving my English.

Last but not least, I am grateful to my family for their love and their support in every respect.