Closed-World Semantics for Query Answering in Temporal Description Logics

Dissertation

zur Erlangung des akademischen Grades
Doktoringenieur (Dr.-Ing.)

vorgelegt an der
Technischen Universität Dresden
Fakultät Informatik

eingereicht von
Dipl.-Inf. Walter Forkel
geboren am 21. Oktober 1992 in Berlin

eingereicht am 6. Oktober 2020
verteidigt am 18. Dezember 2020

Gutachter:
Prof. Dr.-Ing. Franz Baader
Technische Universität Dresden

Prof. Dr. Alessandro Artale
Free University of Bozen-Bolzano

Dresden, im Dezember 2020

Abstract

Ontology-mediated query answering is a popular paradigm for enriching answers to user queries with background knowledge. For querying the *absence* of information, however, there exist only few ontology-based approaches. Moreover, these proposals conflate the closed-domain and closed-world assumption, and therefore are not suited to deal with the anonymous objects that are common in ontological reasoning. Many real-world applications, like processing electronic health records (EHRs), also contain a temporal dimension, and require efficient reasoning algorithms. Moreover, since medical data is not recorded on a regular basis, reasoners must deal with sparse data with potentially large temporal gaps.

Our contribution consists of three main parts: Firstly, we introduce a new closed-world semantics for answering conjunctive queries with negation over ontologies formulated in the description logic $\mathcal{ELH}_\bot$, which is based on the *minimal* universal model. We propose a rewriting strategy for dealing with negated query atoms, which shows that query answering is possible in polynomial time in data complexity. Secondly, we introduce $\mathcal{TELH}_\bot^{\diamondsuit,\mathsf{lhs}}$ a new temporal variant of $\mathcal{ELH}_\bot$ that features a *convexity* operator. We extend this minimal-world semantics for answering metric temporal conjunctive queries with negation over the lightweight temporal logic $\mathcal{TELH}_\bot^{\diamondsuit,\mathsf{lhs},-}$ and obtain similar rewritability and complexity results. Thirdly, apart from the theoretical results, we evaluate minimal-world semantics in practice by selecting patients, based their EHRs, that match given criteria.

Acknowledgments

I would like to thank my wife, Friederike, and our two kids, Georg and Theodor, for the joy they bring to my life every day.

I am thankful for my supervisor, Franz Baader, for giving me the opportunity to conduct this research. Furthermore, I would like to express my gratitude to Alessandro Artale for reviewing my thesis. I am indebted to Stefan Borgwardt, for all his advice, our fruitful discussions, and his sheer endless patience. I thank Andreas Ecke for all the interesting non-research related discussions during our coffee breaks. A special thanks goes to Monika Sturm for kindling my interest in theoretical computer science in many interesting lectures and seminars. I am grateful to Kerstin Achtruth for all the administrative works and the encouraging spirit she brings to our group.

Last, but not least, I thank my family, especially my parents, for their support and helpful advice in all my undertakings.

This work was financially supported by the Deutsche Forschungsgemeinschaft in the grant BA 1122/19-1 (GOASQ).

Contents

Abstract iii

Acknowledgments v

1. Introduction 1
1.1. Description Logics . . . 2
1.2. Ontology-Mediated Query Answering . . . 3
1.3. Incomplete Knowledge . . . 5
1.4. Temporalized Description Logics . . . 6
1.5. Outline and Contributions of the Thesis . . . 7

I. Theory 11

2. Preliminaries 13
2.1. Description Logics . . . 13
2.1.1. The Basic DL $\mathcal{ALC}$ . . . 13
2.1.2. Relation to First-Order Logic . . . 16
2.1.3. Extensions of $\mathcal{ALC}$ . . . 17
2.1.4. The Complexity of Reasoning in DLs . . . 18
2.1.5. The $\mathcal{EL}$ family of DLs . . . 19
2.2. Ontology-Mediated Query Answering . . . 21
2.2.1. Conjunctive Queries . . . 22
2.2.2. Complexity of OMQA . . . 23
2.3. Metric Linear Temporal Logic . . . 25
2.3.1. LTL with Binary Intervals . . . 27

3. Minimal-World Semantics for Conjunctive Queries with Negation 29
3.1. The Patient Selection Problem . . . 30
3.2. Existing Non-Monotonic Formalisms . . . 31
3.2.1. Standard Certain Answer Semantics . . . 32
3.2.2. Epistemic Logic . . . 32
3.2.3. Skolemization . . . 33
3.2.4. Datalog-based Ontology Languages . . . 33
3.2.5. Closed Predicates . . . 33
3.3. Minimal-World Semantics for $\mathcal{ELH}_\bot$ . . . 34
3.4. A Combined Rewriting for NCQs . . . 38
3.4.1. Correctness . . . 40
3.5. Minimal-World Semantics for more Expressive Horn Description Logics . 46
3.5.1. Cores . . . 47

3.5.2. Minimal universal models beyond $\mathcal{ELH}_\bot$ 48
3.5.3. Cores in Existential Rules 52
3.5.4. "Intuitively" Minimal Models for Horn-$\mathcal{ALCHOIQ}^{\mathrm{Disj}}_{\mathrm{Self}}$ 54
3.5.5. Queries beyond (N)CQs 54

4. Temporalizing $\mathcal{ELH}_\bot$ 55
4.1. Convexity Operators 56
4.2. The Temporal Description Logic $\mathcal{TELH}^{\diamondsuit,\mathsf{lhs}}_\bot$ 59
4.3. A Completion Algorithm 60
4.4. Discussion and Related Work 67

5. Temporal Minimal-World Semantics for $\mathcal{ELH}_\bot$ 69
5.1. Metric Temporal Conjunctive Queries with Negation 69
5.2. Minimal-World Semantics for MTNCQs 70
5.3. A Combined Rewriting for MTNCQs 73
5.4. Related Work and Discussion 77

II. Implementation and Experiments 81

6. Translating Criteria to Temporal Queries 85
6.1. Existing and Related Work 86
6.2. Methodology 86
6.2.1. Semantic Annotations 88
6.2.2. The Formal Queries 91
6.3. Experiments 93
6.4. Discussion and Future Work 94

7. Selecting Patients for Clinical Trials 97
7.1. QUELK: A Prototype For Temporal Query Answering 97
7.2. A Dataset for Patient Selection 102
7.3. Setup 106
7.3.1. Creation of the KB 106
7.3.2. Translation of the criteria 108
7.4. Experiments 109
7.5. Discussion and Related Work 111

8. Conclusion 113
8.1. Main Results 113
8.2. Future Work 115

Bibliography 117

Index 131

Chapter 1.

Introduction

Ontologies provide a semantic representation of a specific application domain and therefore play an important part in data access in many areas like medicine [RB08], industry [KMM+17] or the Semantic Web [BHL01; AGP14]. They can be used to model the terminology of an application domain in a machine-processable way. In contrast to traditional database systems, this terminological knowledge can be used to infer new information that were only implicitly available in the data. Moreover, the user can formulate a query based on the terminology and does not need to know the internal relational structure of the data. In *ontology-mediated query answering* (OMQA) this query is then answered over both the explicitly and the implicitly contained information. For many ontologies Description Logics (DLs), a family of logic-based knowledge representation languages, provide the semantic underpinnings [BHL+17; RN14]. They are based on the notions of *concepts* (**`cancer`** or **`liver`**) and *roles* (**`foundIn`**), which express relations between concepts (**`liver cancer`** is **`found in`** the **`liver`**) as well as concrete data (patient *1* is **`diagnosed with`** the disease **`liver cancer`**).

Usually, DLs employ the *open-world assumption*, which means that information that are not explicitly mentioned or cannot be inferred, are neither assumed to be true or false. This assumption is a classic example for *monotonic semantics*: If a fact can be inferred at some point, it will stay valid no matter how much additional knowledge is acquired. In scenarios like the Semantic Web, where a system cannot be assumed to have complete knowledge, the open-world assumption is very reasonable. In contrast, in traditional database systems the *closed-world assumption* is employed, which means that a fact that is not contained in the database is assumed to be false. These semantics are *non-monotonic*, because, if a fact is added to the database, a fact that was false before can become true. In many applications, both the open- and the closed-world assumption are too strong and something intermediate is needed. Many different semantics have been proposed that employ partial closed-world semantics [McC80; Rei80; Wol00].

Many applications, like the monitoring of sensor data [CCG10b] or the diagnoses of a patient[CT15; TSC11], contain an inherent temporal domain. In *temporalized DLs* a concept cannot only talk about the objects at a single time point, but also about objects at other time points. For example, we can express that a certain disease is incurable by a statement like 'if the disease is present now, then it will also be present at all future time points.' In temporalized OMQA, the query language is extended by operators from temporal logic, enabling also temporal information to be queried.

In this work we are concerned with temporalized OMQA over sparse temporal data that contain (almost no) negation. Such a setting can be found in the *patient selection task*, in which, for a given clinical study, eligible patients need to be selected based on their electronic health record. Such records are only available at sparse time points and contain mostly the positive findings. We propose novel closed-world semantics and a lightweight temporal logic that can be used in such settings. Moreover, based on the theoretical formalism, we evaluate our approach on real world patient data.

This chapter gives an informal introduction to the topic of query answering with non-monotonic semantics. In Section 1.1 we describe DLs and how they can be used for OMQA. We then describe different ways to deal with incomplete information in DLs in Section 1.3 and proceed to introduce temporalized OMQA in Section 1.4. In Section 1.5, we provide an overview of the following chapters and specify our contributions.

1.1. Description Logics

Description Logics (DLs) are a family of logic-based knowledge representation formalisms that can be used to express knowledge about a given domain of interest in a formally structured way. They can be seen as a decidable fragment of first-order logic and are closely related to modal logics.

DLs form the semantics basis of the Web Ontology Language OWL [HPV03] and are widely used for the definition of ontologies, such as the prominent biomedical ontologies SNOMED CT, GALEN[1], or the Gene Ontology.

DLs are based on the notions of *concepts* (e.g. **`Cancer`** or **`SkinStructure`**), *roles* (e.g. **`diagnosedWith`**, **`findingSite`**), and *individuals* (e.g. **`John`**, **`JohnsDiagnosis`**). More complex concepts can be built using the Boolean operators negation ($\neg$), disjunction ($\sqcup$), conjunction ($\sqcap$) and role restrictions. For instance, the role restriction $\exists\mathtt{diagnosedWith}.\mathtt{Cancer}$ is a concept that describes the class of objects that are diagnosed with cancer. Different concepts can be related to each other in *general inclusion axioms* (GCIs). For example, we can say that the class of **`SkinCancer`** is included ($\sqsubseteq$) in the class of objects that belong to **`Cancer`** and have a **`findingSite`** that is a **`SkinStructure`** with the following GCI:

$$\mathtt{SkinCancer} \sqsubseteq \mathtt{Cancer} \sqcap \exists\mathtt{findingSite}.\mathtt{SkinStructure}. \tag{1.1}$$

GCIs are collected in the *terminological box* (*TBox*), also called *ontology*. The *assertional box* (*ABox*) contains the data, i.e. knowledge about specific objects such as

$$\mathtt{Patient}(\mathtt{John}), \tag{1.2}$$

$$\mathtt{diagnosedWith}(\mathtt{John}, \mathtt{JohnsDiagnosis}), \tag{1.3}$$

$$\mathtt{SkinCancer}(\mathtt{JohnsDiagnosis}). \tag{1.4}$$

The first assertion states that **`John`** is a member of the concept **`Patient`**. **`John`** has a **`diagnosedWith`** relation to an object called **`JohnsDiagnosis`**, which by the third axiom belongs to the concept of **`SkinCancer`**.

A *knowledge base* (KB) consists of a TBox together with an ABox. Based on the explicit knowledge in the ABox and the TBox, additional implicit knowledge can be

[1]https://bioportal.bioontology.org/ontologies/GALEN

obtained from the KB. This implicit knowledge is implied by the formal semantics of DLs. For example, suppose the TBox contains the GCI

$$\texttt{CancerPatient} \equiv \exists \texttt{diagnosedWith}.\texttt{Cancer},$$

saying that belonging to the concept **CancerPatient** is equivalent ($\equiv$) to a diagnoses with cancer. By the formal semantics the implicit fact that **John** is a cancer patient is then entailed, even though this is not stated explicitly in the data. A foundational reasoning task is the question whether the KB is *consistent*, i.e. free of contradictions. Other reasoning tasks include the *subsumption* problem, where one asks if a concept is a sub-concept of another concept. For instance, in the above example it is the case that **SkinCancer** is a sub-concept of **Cancer**.

Depending on the application, more concept constructors can be allowed. For example, with inverse roles ($\cdot^-$) we can express that each child has something that is its parent by a GCI

$$\texttt{Child} \sqsubseteq \exists \texttt{parentOf}^-.\top$$

where $\top$ is a special concept that denote the class of all objects. The downside of added expressiveness is an increase in the *complexity* or cost of reasoning. For example, reasoning in the basic DL $\mathcal{ALC}$ [SS91], which is the smallest DL that is closed under Boolean operators, is already ExpTime-complete [BHL+17]. This means that no matter how optimized an algorithm is, there will always be certain classes of KBs for which the run time of the algorithm is exponential in the size of the KBs. Much of research in DL is devoted to a better understanding of the interactions between the expressiveness and the complexity of reasoning. An important member of the DL family are *lightweight DLs*, in which reasoning is tractable, i.e. can be done in polynomial time [BBL05]. A prominent member is the $\mathcal{EL}$-family of DLs [BCM+07; Baa03], which is used for reasoning over many medical ontologies, such as SNOMED CT[2], the Gene Ontology,[3] and the NCI Thesaurus.[4] In contrast to $\mathcal{ALC}$, complements and disjunctions are not allowed in $\mathcal{EL}$, which makes reasoning very efficient. In the course of this work, we will be mostly concerned with $\mathcal{ELH}_\bot$, a member of the $\mathcal{EL}$ family.

1.2. Ontology-Mediated Query Answering

In the beginning, the focus in DLs was mainly on classification tasks in the ontology. This shifted in recent years when data integration became a major issue in many areas. To manage and analyze large amounts of heterogeneous data, a common vocabulary is required, which can be captured by an ontology. It provides an abstraction of the underlying naming and relational structures of the heterogeneous data sources. In *ontology-mediated query answering* (OMQA) the ABox is viewed as a set of facts, similar to a relational database. In contrast to database queries, which evaluate the query directly over the data, in OMQA, the query is mediated through the ontology. This way the query is evaluated not only over the information explicitly contained in the ABox,

[2] `http://www.snomed.org/`
[3] `http://www.geneontology.org/`
[4] `https://ncit.nci.nih.gov/`

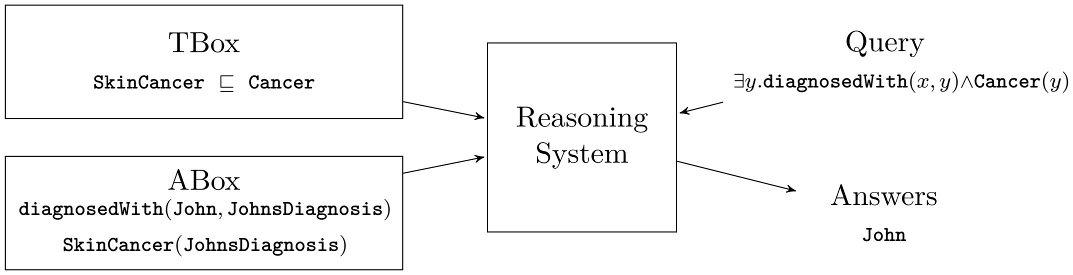


Figure 1.1.: An illustration of a OMQA system. To answer the example query the axiom **SkinCancer** $\sqsubseteq$ **Cancer** ("every object that belongs to **SkinCancer** also belongs to **Cancer**") in the TBox is taken into account, and the implicitly contained answers are returned.

but also over the implicit knowledge that is entailed by the ABox together with the ontology. To query the data, a popular choice are *conjunctive queries (CQs)*, which form a subset of the database query language SQL. A CQ is a conjunction of first-order atoms. Variables may be existentially quantified ($\exists y.$). All remaining variables are called *answer variables*, meaning that they are returned in the answer. For example, the CQ

$$q(x) = \exists y.\big(\texttt{diagnosedWith}(x, y) \land \texttt{Cancer}(y)\big) \tag{1.5}$$

asks for all patients x that are diagnosed with a disease y that is classified as a cancer. The variable y is existentially quantified and will not be returned as an answer, since we are not interested in the precise cancer diagnosis for each patient, but only need to ensure that there *exists* such a diagnosis. An illustration of this setting is shown in Figure 1.1. The reasoning system is aware of the TBox axioms and has access to the facts contained in the ABox. If the query would be evaluated directly over the ABox, it would not return any answers, since the facts alone do not indicate that **John** is **diagnosedWith Cancer**, but just with **SkinCancer**. Without the terminological knowledge contained in the TBox the system cannot know that **SkinCancer** is a form of **Cancer**. The reasoning system infers this implicit fact and therefore returns **John** as an answer to the query.

To obtain algorithms for answering CQs, a popular technique are *rewritings*. The idea is to take a CQ and rewrite it to a query in a target language, for example, a general first-order query. The advantage is that the rewritten query can be evaluated directly over the data using existing database systems, instead of implementing a completely new system. For example, suppose a query asking for a 'cancer that is found on the skin':

$$\exists y.\texttt{Cancer}(x) \land \texttt{findingSite}(x, y) \land \texttt{SkinStructure}(y) \tag{1.6}$$

If the query is evaluated directly over the example KB introduced in Section 1.1, it does not return any results. By the GCI in Equation (1.1) we know that every **SkinCancer** has some **findingSite** that is a **SkinStructure**, even though it is not mentioned explicitly in the ABox. A rewriting approach uses this knowledge and rewrites the query into

$$\texttt{SkinCancer}(x) \tag{1.7}$$

and evaluates this query (possibly among other rewritings) over the ABox, which because of the assertion in Equation (1.4) returns **`JohnsDiagnosis`** as an answer.

Depending on the underlying DL, different rewriting strategies can be employed: In $\mathcal{EL}$, a *combined rewriting* is necessary, in which not only the query is rewritten, but also the data is slightly modified [KLT+11]. Instead of rewriting to first-order queries, in a *Datalog-rewriting*, the query and the data are rewritten into Datalog [AHV95], a deductive database language.

The complexity of query answering is usually measured in two different ways: In *combined complexity*, the query, the ABox and the TBox are all assumed to be inputs. However, in most applications it is reasonable to assume that the query and the TBox are relatively small and do not change very often compared to the possibly millions of facts contained in the ABox, representing the potentially fast changing data. Therefore, in this work we are more interested in the *data complexity* of OMQA. Here only the size of the ABox is viewed as input, while the query and the TBox are assumed to be constant.

1.3. Incomplete Knowledge

In many real-world scenarios the knowledge cannot be assumed to be complete. Consider the medical examples from the previous section: Usually, only positive facts about diseases are available and we cannot be sure if **`John`** is suffering from another disease that is just not relevant for his cancer treatment and therefore was not mentioned in his electronic health record (EHR). Apart from that in some DLs like $\mathcal{EL}$ negation is also technically not available.

DLs usually employ the *open-world assumption* (OWA), i.e. it is assumed that the knowledge in the KB is incomplete. In scenarios like the Semantic Web, were the whole environment cannot be explored, the OWA is a reasonable choice. A fact is then entailed by a KB only if it holds in every possible world, i.e. each world that respects axioms in the ABox and the TBox. For example, a possible world for the KB defined in Equations (1.1) to (1.4) could contain many other patients and **`John`** could have many other diagnoses. It would even be possible that **`John`** and **`JohnsDiagnosis`** are names for the same object. Under the OWA we can conclude that **`John`** is a cancer patient, since every possible world has to contain that facts, but it is not possible to conclude that **`John`** has no other disease. The consideration of every possibility makes the semantics *monotonic*. If a conclusion can be drawn, it stays valid, no matter which additional knowledge the system obtains at a later point in time. On the one hand, this makes conclusions very robust but on the other hand, this limits the amount of conclusions that can be drawn. In many scenarios this assumption is too strong. For example, an EHR usually contains mostly positive findings of a patient. Intuitively, we would assume that a finding that is not mentioned in the EHR is not there, even though we cannot be sure. We assume that the knowledge in the EHR is complete to some extend and employ the *closed-world assumption* (CWA). A fact is only considered to be true, if it can be inferred from the KB, and otherwise false. With the CWA we can conclude that **`John`** has no other diseases except **`SkinCancer`**. When we learn later, that **`John`** is also suffering from the flu, our previous conclusions becomes invalid. Therefore, semantics based on the CWA are non-monotonic. The CWA is the quasi-standard in relational database systems [Lev96]. Such systems even assume a *closed domain*, i.e. there exist only the objects mentioned in the data and no further ones. When combined with existential restrictions like in Equation (1.1), this means that either **John** or **JohnsDiagnosis** have to be the **findingSite** of **JohnsDiagnosis**. This is certainly not what the user expects and therefore the CWA assumption is too strong as well. Many non-monotonic approaches have been introduced that are somewhere in between the OWA and the CWA, for example, epistemic logic [MD80; Wol00], Reiter's default logic [Rei80] or Circumscription [McC80]. The difficulty here is that a formalism can give intuitive results in a certain case, but unintuitive results in many others. Will discuss in detail in Section 3.2, why existing non-monotonic formalisms are not suitable to deal with anonymous individuals that are relevant in OMQA. To overcome this limitation, in Chapter 3 we introduce *minimal-world* semantics.

1.4. Temporalized Description Logics

Many application domains inherently contain a temporal dimension. For example, the EHRs of a patient certainly change over time and it is important to keep track of that. To deal with such scenarios, a plethora of temporalized DLs has been introduced in the literature, see for example [LWZ08; AKW+13]. In principle, a temporalized DL is a combination of a DL with a *temporal logic*. There exist many different temporal logics, however in this thesis we focus on *linear-time temporal logic* (LTL) [Pnu77], in which the flow of time is assumed to be discrete and linear, i.e. each point in time has exactly one successor.

When combining a temporal logic with a DL, the interactions need to be carefully controlled in order to stay decidable. For example, in $\mathrm{LTL}_{\mathcal{ALC}}$, the first temporalized DL introduced in [Sch93], temporal operators are allowed within concepts. For instance, the concept

$$\mathtt{Researcher} \sqcap \Diamond \exists.\mathtt{authors}.\mathtt{Publication}$$

describes researchers that eventually ($\Diamond$) author a publication. In contrast to non-temporal DLs, this concept talks not only about the objects at a given time point, but also involves objects at future time points.

In many scenarios, there exist certain properties that stay constant over time. We have already seen the example of an incurable disease, which once diagnosed stays valid at all future time points (in the live span of the patient). Other examples are properties like being a human: one can argue that an object that is a human will always stay a human and vice versa, an object that is not human will never become human. Such concepts and roles are called *rigid*. While very desirable, rigid roles are so powerful that already in the presence of a single rigid role, reasoning becomes undecidable, even when using $\mathcal{EL}$ instead of $\mathcal{ALC}$ [AKL+07]. Therefore a variety of restrictions has been investigated to regain decidability, for example by restricting temporal operators to occur only on the left-hand side of GCIs [GJK16].

In temporalized OMQA, temporal queries are answered over temporal data. An illustration can be found in Figure 1.2. Conceptually, the temporal data can be modeled as a sequence of ABoxes, containing all the assertions valid at a specific time point, while the TBox is assumed to be valid at all time points. A temporal query can ask not only

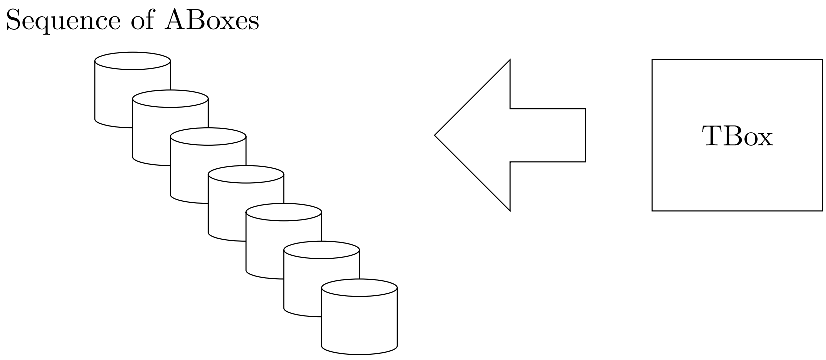


Figure 1.2.: A temporalized DL can be thought of as a single TBox together with an ABox for each time point. The axioms in the TBox are valid at every time point.

about relations between objects at a single time point, but also about objects at different time points. As a simple example, with a temporal CQ we can ask for all patients that have a '*history of cancer,*'[5] i.e. those that were diagnosed with cancer *at some point in the past*, using the temporal operator $\diamondsuit$:

$$\diamondsuit_{(-\infty,0]}(\exists y.\mathtt{diagnosedWith}(x,y) \land \mathtt{Cancer}(y)), \tag{1.8}$$

where the time interval $(-\infty, 0]$ refers to the whole history (relative to the current time point 0). More complex queries include '*type 1 diabetes with duration at least 12 months,*'[6] utilizing the temporal *always* operator $\Box$:

$$\diamondsuit_{(-\infty,0]}\Box_{[-12,0]}(\exists y.\mathtt{diagnosedWith}(x,y) \land \mathtt{Diabetes}(y)). \tag{1.9}$$

In this case, we are looking for some time point t in the past ($\diamondsuit_{(-\infty,0]}$) for which during the whole time interval $[-12, 0]$ relative to t (that is, $[t-12, t]$) the patient had a diagnosis of diabetes.

There are many applications in which data is collected in regular intervals. For example, in stream reasoning, a continuous stream of sensor data is assumed and needs to be processed in real time [CCG10b]. In other applications like EHRs, data are *sparse*. They are only available at irregular intervals that can be years apart. If this is the case, it is not optimal to artificially introduce all intermediate time points during reasoning. We put a special emphasis on developing a temporal OMQA approach that can deal with large temporal gaps in an efficient manner.

1.5. Outline and Contributions of the Thesis

In this work we are concerned with closed world semantics for temporalized OMQA over sparse temporal data that contain (almost no) negation. The thesis is separated into two main parts: In Part I we propose a novel closed-world semantics and a lightweight

[5] `https://clinicaltrials.gov/ct2/show/NCT00064766`
[6] `https://clinicaltrials.gov/ct2/show/NCT02280564`

temporal extension of $\mathcal{EL}$ that can deal with sparse temporal data and show how to answer temporal queries in this setting. In Part II we put the theory to practice by evaluating a prototypical implementation on real world data. In the following we give a brief outline of the remainder of the thesis section by section.

In Chapter 2 we give a short formal introduction to DLs. The syntax and semantics are introduced and the complexity of reasoning in DLs is discussed. In particular, the lightweight DL $\mathcal{ELH}_\bot$ is introduced, because this is the DL we are using in most of the thesis. It provides the formal bases for the SNOMED CT ontology, which we will use in our system for patient selection. We continue with a formal introduction of OMQA over $\mathcal{ELH}_\bot$-KBs and discuss the technique of combined rewritings. In the last part of Chapter 2 we introduce metric temporal linear logic. It is based on a discrete time line, which in our setting are the integers. In contrast to classic linear time logic, the temporal operators can be equipped with intervals, leading to a more concise notation.

In Chapter 3 we introduce a running example of the patient selection problem and define which patients would intuitively be expected to satisfy a given criteria. We then show that existing non-monotonic formalisms such as epistemic logics or closed predicates fail to return the expected results and therefore are not suitable for the patient selection task. We introduce *minimal-world* semantics in which only one model is used for reasoning. This model is required to be universal and at the same time minimal in the number of successors introduced per element. We show that every consistent $\mathcal{ELH}_\bot$-KB has a minimal universal model that is unique up to isomorphism. Then, we use the minimal-world semantics to answer conjunctive queries with negation (NCQs) over ontologies formulated in the description logic $\mathcal{ELH}_\bot$. We propose a rewriting strategy for dealing with negated query atoms, which shows that query answering is possible in polynomial time in data complexity. These parts are mainly based on [BF19a; BF19b; BF19c] where the first won the *best paper award* of the JELIA19 conference.

- Stefan Borgwardt and Walter Forkel: 'Closed-World Semantics for Conjunctive Queries with Negation over $\mathcal{ELH}_\bot$ Ontologies'. In *Proc. JELIA Conference*. Rende, Italy: Springer, 2019, pages 371–386. DOI: `10.1007/978-3-030-19570-0_24`
- Stefan Borgwardt and Walter Forkel: 'Closed-World Semantics for Conjunctive Queries with Negation over $\mathcal{ELH}_\bot$ Ontologies'. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence*. AAAI Press. 2019, pages 6131–6135. DOI: `10.24963/ijcai.2019/849`
- Stefan Borgwardt and Walter Forkel: 'Closed-World Semantics for Conjunctive Queries with Negation over $\mathcal{ELH}_\bot$ Ontologies (Extended Abstract)'. In *Proceedings of the 32nd International Workshop on Description Logics*. Volume 2373. CEUR Workshop Proceedings. 2019. URL: `http://ceur-ws.org/Vol-2373/paper-36.pdf`

We finish the chapter by looking at possible extensions of minimal-world semantics to more expressive Horn-DLs[7] that support transitivity, nominals and inverse roles. We discuss the relation of minimal-world semantics with *cores*, a concept from graph theory, that overlaps with our definition of minimality in the case of $\mathcal{ELH}_\bot$. Unfortunately,

[7]Roughly speaking, a logic is Horn if it disallows all forms of disjunctions.

cores deviate from the intuition behind minimal-world semantics when moving to more expressive Horn-DLs.

With minimal-world semantics, we have a way to deal with negation in the queries. In Chapter 4 we proceed by introducing the lightweight temporal logic $\mathcal{TELH}_{\bot}^{\diamondsuit,\mathsf{lhs}}$. It is a temporal extension of the tractable language $\mathcal{ELH}_{\bot}$, which features a new class of *convex diamond* operators that can be used to bridge temporal gaps in sparse data that occur in many domains. In the medical domain, for instance, data are only available as long as a patient is admitted to a hospital, but not in between. With the help of the diamond operators, such gaps can be interpolated by expressions like

$$\diamondsuit_2\mathtt{CancerPatient} \sqsubseteq \mathtt{CancerPatient},$$

i.e. if a person is known to be a cancer patient at two time points t, t' that are at most 2 months apart, then the patient was also a cancer patient at all time points in between t and t'. We develop a completion algorithm for our logic, which shows that entailment remains tractable. Moreover, we show that $\mathcal{TELH}_{\bot}^{\diamondsuit,\mathsf{lhs}}$ is suitable to deal with sparse temporal data: By the choice of the diamond operators, the behavior in every intermediate interval between time points occurring in the ABox can be captured by a single representative time point.

With the temporal logic $\mathcal{TELH}_{\bot}^{\diamondsuit,\mathsf{lhs}}$ at hand we want to employ it for temporal OMQA and introduce temporal conjunctive queries with negation (MTNCQs) as a temporal query language in Chapter 5. In order to answer MTNCQs we construct a temporal minimal model for a given consistent $\mathcal{TELH}_{\bot}^{\diamondsuit,\mathsf{lhs}}$-KB. Because it is not clear how the intuition behind minimality can be retained when temporal roles are present, we disallow temporal roles in this setting. Through a two step combined rewriting we are able to show that MTNCQs answering in $\mathcal{TELH}_{\bot}^{\diamondsuit,\mathsf{lhs}}$ w.r.t. minimal-world semantics is still tractable in data complexity. In the first step the atemporal rewriting defined in Chapter 3 is applied to the atemporal parts of the query. In the second step the temporal parts are rewritten to a finite representation by exploiting the property of $\mathcal{TELH}_{\bot}^{\diamondsuit,\mathsf{lhs}}$ that intermediate intervals can be described by a constant number of representatives. This and the previous chapter are based on the publications [BFK19; BFK20; BFK]

- Stefan Borgwardt, Walter Forkel, and Alisa Kovtunova: ‘Finding New Diamonds: Temporal Minimal- World Query Answering over Sparse ABoxes’. In *Proc. of the 3rd International Joint Conference on Rules and Reasoning (RuleML+RR'19)*. Bolzano, Italy: Springer, 2019. DOI: `10.1007/978-3-030-31095-0_1`

- Stefan Borgwardt, Walter Forkel, and Alisa Kovtunova: ‘Finding New Diamonds: Temporal Minimal- World Query Answering over Sparse ABoxes (Extended Abstract)’. In *Proceedings of the 33rd International Workshop on Description Logics (DL 2020)*. Volume 2663. CEUR Workshop Proceedings. 2020. DOI: `10.1007/978-3-030-31095-0_1`

- Stefan Borgwardt, Walter Forkel, and Alisa Kovtunova: ‘Finding New Diamonds: Temporal Minimal- World Query Answering over Sparse ABoxes’. In: Submitted to the Journal of Theory and Practice of Logic Programming (TPLP). DOI: `10.1007/978-3-030-31095-0_1`

This concludes the theoretical part of this work and we move to Part II, where we conduct two experiments to evaluate the suitability of our approach for the patient selection task, which we analyzed in [BBF18]

- Franz Baader, Stefan Borgwardt, and Walter Forkel: 'Patient Selection for Clinical Trials Using Temporalized Ontology-Mediated Query Answering'. In *Proc. HQA Workshop.* ACM, 2018, pages 1069–1074. DOI: `10.1145/3184558.3191538`

In Chapter 6 we discuss a prototypical implementation of a system for automatic translation of clinical trial criteria into MTNCQs. The system uses natural language tools and techniques to construct an intermediate structure containing the relevant information of a given criterion. A mapping is then defined that associates each intermediate structure to an MTNCQ. The approach is tested using criteria from ClinicalTrials.gov. For each criterion the translation provided by the system is evaluated manually. The evaluation shows that the system can provide correct translation to many criteria. This chapter is based on [XFB+19]

- Chao Xu, Walter Forkel, Stefan Borgwardt, Franz Baader, and Beihai Zhou: 'Automatic Translation of Clinical Trial Eligibility Criteria into Formal Queries'. In *Proc. of the 9th Workshop on Ontologies and Data in Life Sciences (ODLS'19), part of The Joint Ontology Workshops (JOWO'19).* CEUR Workshop Proceedings. 2019

In Chapter 7 we assume that the criteria are already translated to MTNCQs and we apply temporal OMQA to select patients. We shortly describe QUELK, our system for query answering, and the input formats it requires. Moreover, we discuss the parts in which the implementation differs from the theoretical algorithms we have developed before. QUELK is then tested on a small dataset for patient selection. For the evaluation we focus on criteria that contain temporal information and manually translate them to MTNCQs. Based on existing tools we extract medical concepts occurring in the EHRs and generate a temporal ABox automatically. We conduct two experiments to show the importance of temporal reasoning: In the first setting, all temporal information are ignored, effectively merging all EHRs to a single time point and using only NCQs for querying. This is evaluated against the temporal setting, in which the MTNCQs are evaluated over the temporal ABox. As expected the quality of the results is better when the temporal dimension is taken into account. We end this thesis by providing conclusions and discussing future work in Chapter 8.

Part I.

Theory

Chapter 2.

Preliminaries

In this chapter we give a short formal introduction into Description Logics (Section 2.1), Ontology-Mediated Query Answering (Section 2.2) and Temporal Logic (Section 2.3). For the reader familiar with these topics, this chapter may be skipped and used merely as a reference for the exact definitions that are used in the following chapters.

2.1. Description Logics

Description Logics (DLs) are a family of logic-based knowledge representation formalisms. In the following we give a short introduction to DL. For a more detailed introduction we refer the reader to [BCM+07; BHL+17].

Description Logics allow the modeling of the conceptual knowledge of a given application domain. Conceptual knowledge is represented by defining relevant *concepts* of the domain. This *terminology* can then be used to specify properties of objects and individuals that are part of the application domain. In contrast to some of their predecessors, DLs are equipped with formal, logic-based semantics. This allows for automated reasoning and therefore the inference of *implicit* knowledge from the *explicit* knowledge contained in the knowledge base [BCM+07].

In practice a user querying a knowledge representation (KR) system expects to get a positive or a negative answer in reasonable time. Different DLs offer different trade-offs between their *expressivity*, i.e. which features it supports for describing the application domain, and their *complexity* of reasoning.

2.1.1. The Basic DL $\mathcal{ALC}$

In the following we give a short introduction to the basic and widely used DL $\mathcal{ALC}$, first introduced in [SS91]. Its name stands for 'Attribute concept Language with Complement' and it supports all three Boolean operators, i.e. conjunction, disjunction and negation. While on the one hand it serves as the basis for many more expressive DLs, on the other hand many light-weight DLs can be seen as fragments of $\mathcal{ALC}$.

The basic building blocks of any DL are three sets: In the following, let N_C, N_R and N_I be disjoint, countable infinite sets of *concept names*, *role names* and *individual names*, respectively. The names from these sets can be used to define our application domain. We start out by defining how complex concepts can be build:

Definition 2.1 (Syntax of concepts). The set of $\mathcal{ALC}$ concepts is the smallest set satisfying the following:

- every $C \in N_C$ is an $\mathcal{ALC}$ concept,
- the *top concept* $\top$ and *bottom concept* $\bot$ are $\mathcal{ALC}$ concepts,
- if C, D are $\mathcal{ALC}$ concepts, then also $C \sqcap D$ (*conjunction*), $C \sqcup D$ (*disjunction*) and $\neg C$ (*negation/complement*) are $\mathcal{ALC}$ concepts, and
- if $r \in N_R$ and C is an $\mathcal{ALC}$ concept, then $\exists r.C$ (*existential restriction*) and $\forall r.C$ (*universal restriction*) are $\mathcal{ALC}$ concepts. $\diamondsuit$

We can use the available concept constructors of $\mathcal{ALC}$ to build complex concepts that describe the properties we are interested in. Consider for example the following concept description:

$$\texttt{Male} \sqcap \exists \texttt{livesIn}.(\texttt{House} \sqcap \exists \texttt{locatedIn}.(\texttt{Dresden} \sqcup \texttt{Bautzen})) \sqcap \forall \texttt{hasChild}.\texttt{Male} \quad (2.1)$$

This concept describes a male person that lives in a House in Dresden or Bautzen and has only male children.

In Description Logics formal semantics to such concept descriptions is defined model-theoretically using the notion of an interpretation.

Definition 2.2 (Semantics of concepts). An *interpretation* is a pair $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$, where the *domain* $\Delta^{\mathcal{I}}$ is a non-empty set and the *interpretation function* $\cdot^{\mathcal{I}}$ assigns

- to every $a \in N_I$ an element $a^{\mathcal{I}} \in \Delta^{\mathcal{I}}$,
- to every $C \in N_C$ a set $C^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}}$, and
- to every $r \in N_R$ a binary relation $r^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$.

The function is extended to complex concepts as follows:

- $(C \sqcap D)^{\mathcal{I}} := C^{\mathcal{I}} \cap D^{\mathcal{I}}$
- $(C \sqcup D)^{\mathcal{I}} := C^{\mathcal{I}} \cup D^{\mathcal{I}}$
- $(\neg C)^{\mathcal{I}} := \Delta^{\mathcal{I}} \setminus C^{\mathcal{I}}$
- $(\exists r.C)^{\mathcal{I}} := \{d \in \Delta^{\mathcal{I}} \mid \text{ there exists } e \in \Delta^{\mathcal{I}} \text{ with } (d,e) \in r^{\mathcal{I}} \text{ and } e \in C^{\mathcal{I}}\}$
- $(\forall r.C)^{\mathcal{I}} := \{d \in \Delta^{\mathcal{I}} \mid \text{ for all } e \in \Delta^{\mathcal{I}} \text{ if } (d,e) \in r^{\mathcal{I}} \text{ then } e \in C^{\mathcal{I}}\}$.

A concept C is said to be *satisfiable* if there is an interpretation $\mathcal{I}$ with $C^{\mathcal{I}} \neq \emptyset$. $\diamondsuit$

A possible interpretation $\mathcal{I}_1$ of the concept description in Equation (2.1) is the following:

$$\begin{aligned}
\Delta^{\mathcal{I}_1} &:= \{a, b, c, d\} \\
\texttt{Male}^{\mathcal{I}_1} &:= \{a, d\} \\
\texttt{House}^{\mathcal{I}_1} &:= \{b\} \\
\texttt{Dresden}^{\mathcal{I}_1} &:= \{c\} \\
\texttt{Bautzen}^{\mathcal{I}_1} &:= \emptyset \\
\texttt{livesIn}^{\mathcal{I}_1} &:= \{(a, b)\} \\
\texttt{locatedIn}^{\mathcal{I}_1} &:= \{(b, c)\} \\
\texttt{hasChild}^{\mathcal{I}_1} &:= \{(a, d)\}
\end{aligned}$$

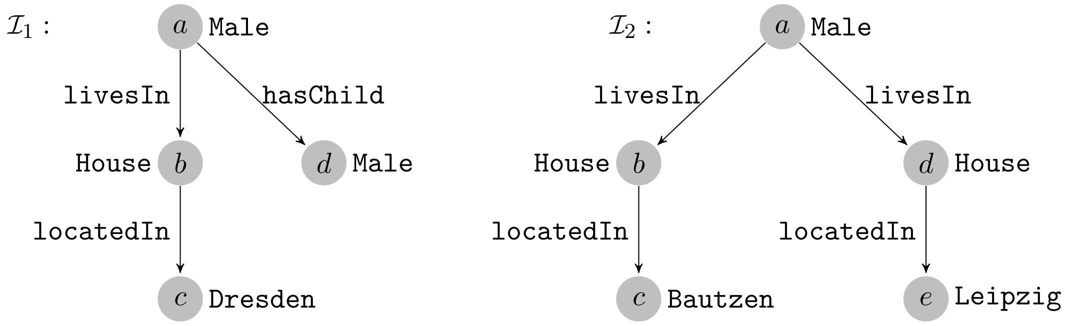


Figure 2.3.: A graphical representation of two possible interpretations for the concept description in Equation (2.1). Each object is denoted with a gray circle and labeled with the concept names it belongs to. Each edge between objects is labeled with the role names it satisfies.

In the interpretation $\mathcal{I}_1$ a male person has a **`livesIn`**-edge to an object that is a **`House`** that is connected with a **`locatedIn`**-edge to an object belonging to the concept **`Dresden`**. The male also has a **`hasChild`**-edge to an object that is **`Male`**. In Figure 2.3 a graphical representation of $\mathcal{I}_1$ and another interpretations $\mathcal{I}_2$ can be seen. The interpretation $\mathcal{I}_2$ might be more surprising, since there the male lives in two different houses, one in **`Bautzen`** and one in **`Leipzig`**. Still, this satisfies the concept, since it describes the existence of a house located in **`Bautzen`** (or **`Dresden`**). On the other hand, the constraint that all children should be male is satisfied trivially in $\mathcal{I}_2$, since there are no children.

In the following we make the *standard names assumption* (SNA), i.e. that each interpretation $\mathcal{I}$ assigns a standard name to each element $a \in N_I$, formally $a^{\mathcal{I}} = a$. The SNA is an assumption that is often made in DLs to simplify notation.

While concept descriptions already offer some expressivity, they alone do not allow to state hierarchical knowledge between different concepts, for example, that every person is a human.

Definition 2.4 (Syntax and Semantics of TBoxes). A *general concept inclusion* (GCI) is of the form $C \sqsubseteq D$, where C and D are concepts. A *TBox* is a finite set of GCIs. An interpretation $\mathcal{I}$ is a *model* of a GCI $C \sqsubseteq D$ if $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$. $\mathcal{I}$ is a model of a TBox $\mathcal{T}$, denoted by $\mathcal{I} \models \mathcal{T}$, if $\mathcal{I}$ is a model of all GCIs in $\mathcal{T}$. $\diamondsuit$

As usual, we define $A \equiv B$ as an abbreviation for the two GCIs $A \sqsubseteq B$ and $B \sqsubseteq A$.

Definition 2.5 (Syntax and Semantics of ABoxes). A *concept assertion* is of the form $C(a)$ and a *role assertion* is of the form $r(a, b)$, where $C \in N_C, r \in N_R$, and $a, b \in N_I$. An *ABox* is a finite set of concept and role assertions. An interpretation $\mathcal{I}$ is a *model* (i) of a concept assertion $C(a)$ if $a^{\mathcal{I}} \in C^{\mathcal{I}}$, (ii) of a role assertion $r(a, b)$ if $(a^{\mathcal{I}}, b^{\mathcal{I}}) \in r^{\mathcal{I}}$, and (iii) of an ABox $\mathcal{A}$, denoted by $\mathcal{I} \models \mathcal{A}$, if it is a model of all assertions in $\mathcal{A}$. $\diamondsuit$

In the following we often call GCIs, concept and role assertions simply *axioms*.

Definition 2.6 (Knowledge Base). A *knowledge base* $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ consists of a TBox $\mathcal{T}$ and an ABox $\mathcal{A}$. An interpretation $\mathcal{I}$ is a *model* of $\mathcal{K}$, denoted by $\mathcal{I} \models \mathcal{K}$, if $\mathcal{I}$ is a model of $\mathcal{T}$ and $\mathcal{A}$. $\mathcal{K}$ is *consistent* if it has a model. An axiom α is *entailed* by $\mathcal{K}$, denoted by $\mathcal{K} \models \alpha$ if every model of $\mathcal{K}$ is also a model of α. $\diamondsuit$

Example 2.7. The following example is based on an example from [Rud11]. Let $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ be a KB with $\mathcal{T}$ containing the following GCIs:

$$
\begin{aligned}
\mathtt{Healthy} &\sqsubseteq \neg\mathtt{Dead} && \text{(someone that is healthy is not dead)}\\
\mathtt{Cat} &\sqsubseteq \mathtt{Dead} \sqcup \mathtt{Alive} && \text{(a cat is dead or alive)}\\
\mathtt{HappyCatOwner} &\sqsubseteq \exists\mathtt{caresFor}.\mathtt{Cat} \sqcap && \text{(a happy cat owner cares for a cat and}\\
&\quad \forall\mathtt{caresFor}.\mathtt{Healthy} && \text{all beings he cares for are healthy)}
\end{aligned}
$$

Let $\mathcal{A}$ contain the assertion that Schrödinger is a happy cat owner:

$$\mathtt{HappyCatOwner}(\mathtt{Schrödinger})$$

It is easy to check that $\mathcal{K}$ is consistent by constructing a model of $\mathcal{K}$. As an example for entailment, it holds that $\mathcal{K} \models \mathtt{Cat} \sqcap \mathtt{Healthy} \sqsubseteq \mathtt{Alive}$, i.e. a healthy cat is also alive. The conclusion can be drawn because of the first two axioms in $\mathcal{T}$: Since each **`cat`** is **`Dead`** or **`Alive`**, and being **`Healthy`** implies not being **`Dead`**, the only possibility is that a healthy cat is **`Alive`**. ♢

Definition 2.8 (Signature). The *signature* sig of a concept C is the set of all concept and role names occurring in C. The signature of a KB $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ is defined as

$$
\begin{aligned}
\mathrm{sig}(\mathcal{T}) &= \bigcup_{C \sqsubseteq D \in \mathcal{T}} \mathrm{sig}(C) \cup \mathrm{sig}(D)\\
\mathrm{sig}(\mathcal{A}) &= \bigcup_{C(a) \in \mathcal{A}} \{C, a\} \cup \bigcup_{r(a,b) \in \mathcal{A}} \{r, a, b\}\\
\mathrm{sig}(\mathcal{K}) &= \mathrm{sig}(\mathcal{T}) \cup \mathrm{sig}(\mathcal{A})
\end{aligned}
$$
♢

2.1.2. Relation to First-Order Logic

As mentioned already DLs can be seen as decidable fragments of *first-order logic* (FOL). Concept and role names correspond to unary and binary predicates, respectively, while individual names correspond to constants. For example, the GCI

$$\mathtt{Parent} \sqsubseteq \mathtt{Person} \sqcap \exists\mathtt{hasChild}.(\mathtt{Male} \sqcup \mathtt{Female})$$

corresponds to the FOL formula

$$\forall x.\Big(\mathtt{Parent}(x) \implies \mathtt{Person}(x) \land \exists y.\big(\mathtt{hasChild}(x,y) \land (\mathtt{Male}(y) \lor \mathtt{Female}(y))\big)\Big).$$

The translation requires the use of at most two variables x and y. This also shows that $\mathcal{ALC}$, like many other DLs belongs to the two-variable fragment of FOL, i.e. those first-order formulas that can be written using only two variables x and y (and possibly requantifying them) [Bor96]. Formally, the translation function π_x is defined inductively by

$$\begin{aligned}
\pi_x(A) &= A(x) \\
\pi_x(C \sqcap D) &= \pi_x(C) \wedge \pi_x(D) \\
\pi_x(C \sqcup D) &= \pi_x(C) \vee \pi_x(D) \\
\pi_x(\exists r.C) &= \exists y.(r(x,y) \wedge \pi_y(C) \\
\pi_x(\forall r.C) &= \forall y.(r(x,y) \wedge \pi_y(C) \\
\pi_x(\neg C) &= \neg \pi_x(C).
\end{aligned}$$

The function π_y is analogously defined. The translation can be extended to ABoxes and TBoxes:

$$\begin{aligned}
\pi(\mathcal{T}) &= \forall v. \bigwedge_{C \sqsubseteq D \in \mathcal{T}} (\pi_x(C) \rightarrow \pi_x(D)), \\
\pi(\mathcal{A}) &= \bigvee_{C(a) \in \mathcal{A}} C(a) \wedge \bigvee_{r(a,b) \in \mathcal{A}} r(a,b)
\end{aligned}$$

Theorem 2.9 ([BHL+17]). Let $(\mathcal{T}, \mathcal{A})$ be an $\mathcal{ALC}$-KB. It holds that $(\mathcal{T}, \mathcal{A})$ is satisfiable iff $\pi(\mathcal{T}) \wedge \pi(\mathcal{A})$ is satisfiable.

2.1.3. Extensions of $\mathcal{ALC}$

DLs differ between one another by the concept and role constructors that are available. In the DL nomenclature additional constructors are denoted by concatenating their corresponding letters, for example, *number restrictions* ($\mathcal{N}$) allow to restrict the total number of role-successors. With number restrictions we can express by the axiom

$$\texttt{Parent} \sqsubseteq \geq 1.\texttt{hasChild}.\top$$

that a parent has at least 1 child. With *qualified number restrictions* ($\mathcal{Q}$), it is even possible to restrict the number of role-successors that belong to a certain concept, i.e.

$$\texttt{Car} \sqsubseteq \geq 4.\texttt{hasPart}.\texttt{Tire}$$

expresses that each car has a least four tires. If the DL supports *inverse roles* ($\mathcal{I}$), one can express that for each tire there is a car the tire is part of, formally

$$\texttt{Tire} \sqsubseteq \exists \texttt{hasPart}^{-}.\texttt{Car}.$$

If we allow both qualified number restrictions and inverse roles in $\mathcal{ALC}$, the logic we obtain is called $\mathcal{ALCIQ}$. The extension of $\mathcal{ALC}$ with *transitivity* is usually denoted by $\mathcal{S}$ because of its close relationship to the modal logic **S4**. If *nominals* are available then individual names can be used as concepts in the TBox. This can be used to restrict the members of certain concepts. For example, the axiom

$$\texttt{Element} \sqsubseteq \{\texttt{Earth}, \texttt{Water}, \texttt{Wind}, \texttt{Fire}\}$$

expresses that **Earth**, **Water**, **Wind**, or **Fire** are the only elements. With *role hierarchies* ($\mathcal{H}$) *role inclusion axioms* (RIAs) can be stated such as

$$\texttt{childOf} \sqsubseteq \texttt{descendantOf},$$

i.e. if b is the child of a, then b is also a descendant of a. Even more expressive are *role chains* ($\mathcal{R}$) which allow to express axioms such as

$$\texttt{hasChild}^{-} \circ \texttt{hasChild} \sqsubseteq \texttt{hasSibling},$$

saying that the child of someone I am a child of is my sibling. Transitivity can be expressed through role chains, for example the RIA

$$\texttt{descendantOf} \circ \texttt{descendantOf} \sqsubseteq \texttt{descendantOf}$$

make the role **descendantOf** transitive. The logic that allows all the constructors introduced above is called $\mathcal{SROIQ}$.

Throughout this thesis, we sometimes prefix some notions with the specific DL to make clear which DL is used to construct the concepts or axioms. For instance, we may write '$\mathcal{ALC}$-KB' to make clear that the KB is constructed using only concepts expressible in $\mathcal{ALC}$. If the DL under consideration is clear from the context, we omit this prefix for ease of presentation.

2.1.4. The Complexity of Reasoning in DLs

Most of the time we are interested not only in the explicitly stated axioms and assertions of a KB, but also in the implicit knowledge that is implied by the formal semantics of DLs. Given a KB $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ different reasoning problems have been investigated:

- the *consistency problem* asks if there is least one model of $\mathcal{K}$. If there are no models at all, the KB is inconsistent and contains a contradiction;
- the *satisfiability problem* is a bit more specific by asking if there is a model of $\mathcal{T}$ that also satisfies a given concept C;
- the *subsumption problem* asks if a concept C is contained in (a sub-concept of) a concept D in all models of $\mathcal{T}$, formally $\mathcal{T} \models C \sqsubseteq D$; and
- the *instance checking problem* also takes $\mathcal{A}$ into account and asks whether an individual a is a member of a concept C in all models of $\mathcal{K}$, formally $\mathcal{K} \models C(a)$.

In the following we may write $C \sqsubseteq_{\mathcal{T}} D$ as a notational variant to $\mathcal{T} \models C \sqsubseteq D$. The four reasoning problems have been shown to be closely related to each another:

Theorem 2.10 ([BHL+17]). Let $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ be an $\mathcal{ALC}$-KB, C, D concepts and a an individual name. Then the following hold:

1. $C \equiv_{\mathcal{T}} D$ iff $C \sqsubseteq_{\mathcal{T}} D$ and $D \sqsubseteq_{\mathcal{T}} C$.
2. $C \sqsubseteq_{\mathcal{T}} D$ iff $C \sqcap \neg D$ is not satisfiable w.r.t. $\mathcal{T}$.
3. C is satisfiable w.r.t. $\mathcal{T}$ iff $C \not\sqsubseteq_{\mathcal{T}} \bot$, where $\not\sqsubseteq_{\mathcal{T}}$ denotes the negation of $\sqsubseteq_{\mathcal{T}}$.

4. C is satisfiable w.r.t. $\mathcal{T}$ iff $(\mathcal{T} \cup \{C(a)\})$ is consistent for a new individual a.
5. $\mathcal{K} \models C(a)$ iff $(T, A \cup \{\neg C(a)\})$ is inconsistent.

The above theorem does not only hold for $\mathcal{ALC}$, but also for other DLs that support negation and conjunction in complex concepts. As a consequence it is sufficient to focus on the problem of consistency, since all others can be reduced to the consistency problem.

Of central interest to DL research is the question how efficient reasoning can be done in a given DL and how the efficiency changes when adding or removing certain constructors from the DL. This 'difficulty' is captured by the *computational complexity* of a problem. It gives measures for how hard it is to solve a given decision problem, and in particular to compute a solution to the problem. Every problem belongs to its own *complexity class*, which is defined by a (non)-deterministic Turing machine and a specific resource bound on space or computation time. In the following, we list some complexity classes that are relevant for this thesis, in increasing order w.r.t. set inclusion:

- PTime is the class of problems that can be solved in polynomial time by a deterministic Turing machine. Problems in PTime are also called *tractable* problems.
- NP is the class of problems that can be solved in polynomial time by a non-deterministic Turing machine.
- PSpace is the class of problems that can be solved in polynomial space by a (non)-deterministic Turing machine.
- ExpTime is the class of problems that can be solved in exponential time by a deterministic Turing machine.
- NExpTime is the class of problems that can be solved in exponential time by a non-deterministic Turing machine.

Additionally, *complementary complexity classes* exist, for example coNP is the class of problems whose complements are in NP. For a more in depth introduction to computational complexity, we refer the reader to [Pap94].

Given a complexity class $\mathfrak{C}$ a problem L is said to be $\mathfrak{C}$-*hard* if there is a reduction in polynomial time from a problem $L' \in \mathfrak{C}$ to L. A problem is $\mathfrak{C}$-*complete* if it is in $\mathfrak{C}$ and it is $\mathfrak{C}$-hard.

Lemma 2.11 ([BHL+17]). In $\mathcal{ALC}$ the consistency problem is ExpTime-complete.

2.1.5. The $\mathcal{EL}$ family of DLs

As we have seen even basic reasoning problems such as consistency checking are already not tractable in $\mathcal{ALC}$. This has lead to the investigation of fragments of $\mathcal{ALC}$ that trade some expressiveness for the benefit of tractable reasoning. In the following we introduce a lightweight DL called $\mathcal{ELH}_\bot$, a slight extension of the basic DL $\mathcal{EL}$ by role hierarchies and the bottom concept $(\bot)$.

An $\mathcal{ELH}_\bot$ concept is defined by the following grammar rule:

$$C ::= A \mid \top \mid \bot \mid C \sqcap C \mid \exists r.C$$

$$\text{T1}\ \frac{}{A \sqsubseteq \top} \qquad \text{T2}\ \frac{A_1 \sqsubseteq A_2 \quad A_2 \sqsubseteq A_3}{A_1 \sqsubseteq A_3}$$

$$\text{T3}\ \frac{r_1 \sqsubseteq r_2 \quad r_2 \sqsubseteq r_3}{r_1 \sqsubseteq r_3} \qquad \text{T4}\ \frac{A \sqsubseteq A_1 \quad A \sqsubseteq A_2 \quad A_1 \sqcap A_2 \sqsubseteq B}{A \sqsubseteq B}$$

$$\text{T5}\ \frac{}{\exists r.\bot \sqsubseteq \bot} \qquad \text{T6}\ \frac{A \sqsubseteq \exists r.A_1 \quad r \sqsubseteq s \quad A_1 \sqsubseteq B_1 \quad \exists s.B_1 \sqsubseteq B}{A \sqsubseteq B}$$

$$\text{A1}\ \frac{}{\top(a)} \qquad \text{A2}\ \frac{A(a) \quad A \sqsubseteq B}{B(a)} \qquad \text{A3}\ \frac{r(a,b) \quad r \sqsubseteq s}{s(a,b)}$$

$$\text{A4}\ \frac{A_1(a) \quad A_2(a) \quad A_1 \sqcap A_2 \sqsubseteq B}{B(a)} \qquad \text{A5}\ \frac{r(a,b) \quad A(b) \quad \exists r.A \sqsubseteq B}{B(a)}$$

Figure 2.12.: Completion rules for $\mathcal{ELH}_\bot$ knowledge bases, where A, B, A_1, A_2, A_3, B_1 are $\top$, $\bot$ or (normalized) $\mathcal{ELH}_\bot$ concepts from $\mathcal{K}$; r, s, r_1, r_2, r_3 are role names from $\mathcal{K}$; a, b are individual names from $\mathcal{K}$.

where $A \in N_C$ and $r \in N_R$. In addition to GCIs, an $\mathcal{ELH}_\bot$-TBox may also contain RIAs of the form $r \sqsubseteq s$ where $r, s \in N_R$.

In general members of the $\mathcal{EL}$ family do not allow negation or disjunction. However, by the $\bot$-concept it becomes possible to express *concept disjointness* through GCIs of the form $A \sqcap B \sqsubseteq \bot$, i.e. nothing can belong to A and B at the same time. This offers a limited type of negation, since $A \sqcap B \sqsubseteq \bot$ implies $A \sqsubseteq \neg B$ and $B \sqsubseteq \neg A$, which cannot be expressed directly in $\mathcal{ELH}_\bot$.

To simplify the description of algorithms, it is often useful to require the TBox to be in an appropriate *normal form*. A general $\mathcal{ELH}_\bot$-TBox $\mathcal{T}$ is in normal form if it contains only GCIs of the following form:

$$A_1 \sqcap \cdots \sqcap A_n \sqsubseteq B \qquad \exists r.A \sqsubseteq B \qquad A \sqsubseteq \exists r.B \qquad r \sqsubseteq s \tag{2.2}$$

where $A_1, \ldots, A_n, A \in N_C \cup \{\top\}$,$B \in N_C \cup \{\bot\}$, $r, s \in N_R$, and $n \geq 1$. Note that $\bot$ is allowed to occur only on the right-hand side, because GCIs with $\bot$ on the left-hand side are trivially satisfied and can therefore be omitted. For the same reason $\top$ is allowed only on the left-hand side of GCIs.

The transformation of a general TBox $\mathcal{T}$ into a normalized TBox $\mathcal{T}'$ can be done in polynomial time and introduces fresh concept names when needed, therefore the signature of $\mathcal{T}'$ possibly becomes larger [BHL+17]. For example, the GCI $A \sqcap \exists r.B \sqsubseteq C \sqcap D$ can be expressed by the following normalized GCIs

$$A \sqcap B_1 \sqsubseteq C \qquad A \sqcap B_1 \sqsubseteq D \qquad \exists r.B \sqsubseteq B_1$$

where B_1 is a fresh concept name. The transformation preserves consequences, i.e. it holds for all concepts $A, B \in \mathrm{sig}(\mathcal{T})$ that $\mathcal{T} \models A \sqsubseteq B$ iff $\mathcal{T}' \models A \sqsubseteq B$. In the following, if not stated otherwise we assume all $\mathcal{ELH}_\bot$-KBs to be normalized according to Equation (2.2).

In $\mathcal{ELH}_\bot$ the subsumption problem can be decided in PTIME by *consequence-based reasoning* algorithms: In a normalized $\mathcal{ELH}_\bot$-KB, there are only polynomially many GCIs (in normal form) that are possible consequences of the stated GCIs of the TBox. All of these GCIs in normal form can be generated in polynomial time by the rules T1-T6 in Figure 2.12. A rule is applicable, if the rule head (above the bar) can be satisfied by axioms in the KB. If the rule is applied the axioms in the body of the rule (below the bar) are added to the KB, if they were not present already. If the head of a rule is empty, it is trivially satisfied.

With the same approach the instance problem can be solved by applying the rules A1-A5 to saturate the ABox. For more details see [BHL+17].

While without negation and disjunction the reasoning tasks cannot directly be reduced to one another, the saturation procedure can also be used to check consistency in PTIME: If the KB is inconsistent, the GCI $\top \sqsubseteq \bot$ will be derived as a consequence. Similarly, a concept C is satisfiable if the saturation procedure does not derive the GCI $\bot \sqsubseteq C$.

2.2. Ontology-Mediated Query Answering

So far we have seen different reasoning tasks that give us basic information about the knowledge contained in a given KB. This can provide us with some insights, but often more advanced reasoning tasks are required. In *ontology-mediated query answering* (OMQA) the idea is to answer complex queries about the data and mediate them using the background knowledge contained in the TBox (ontology). In the following we introduce two query languages, namely general first-order queries and conjunctive queries, that we will use later to query $\mathcal{ELH}_\bot$-KBs.

In the following let N_V be a countably infinite set of *variables*. The set of *terms* is $N_T := N_V \cup N_I$, as the set of all variables and individual names.

Definition 2.13 (First-order query). A *first-order query* $\phi(\mathbf{x})$ is a first-order formula built from *concept atoms* $A(t)$ and *role atoms* $r(t, t')$ with $A \in N_C$, $r \in N_R$, and $t_i \in N_T$, using the Boolean connectives $(\wedge, \vee, \neg, \rightarrow)$ and universal and existential quantifiers $(\forall x, \exists x)$.

The free variables $\mathbf{x}$ of $\phi(\mathbf{x})$ are called *answer variables* and we say that ϕ is k-ary if there are k answer variables.

The remaining variables are the *quantified variables*. We use $\mathrm{Var}(\phi)$ to denote the set of all variables in ϕ.

A query without any answer variables is called a *Boolean query*. ♢

As usual, the semantics are given in terms of interpretations:

Definition 2.15 (Semantics of first-order queries). Let $\mathcal{I} = (\Delta, \cdot^\mathcal{I})$ be an interpretation. An *assignment* $\pi\colon \mathrm{Var}(\phi) \rightarrow \Delta^\mathcal{I}$ *satisfies* ϕ in $\mathcal{I}$, if $\mathcal{I}, \pi \models \phi$ under the standard semantics of first-order logic shown in Figure 2.14. We write $\mathcal{I} \models \phi$ if there is a satisfying assignment for ϕ in $\mathcal{I}$.

Let $\mathcal{K}$ be a DL-KB. A k-tuple $\mathbf{a}$ of individual names from $\mathrm{Ind}(\mathcal{K})$ is an *answer* to ϕ in $\mathcal{I}$ if ϕ has a satisfying assignment π in $\mathcal{I}$ with $\pi(\mathbf{x}) = \mathbf{a}$, where $\mathrm{Ind}(\mathcal{K})$ denotes the individual names occurring in $\mathcal{K}$.

We denote the set of all answers to ϕ in $\mathcal{I}$ by $\mathrm{ans}(\phi, \mathcal{I})$. ♢

Name	ϕ	$\mathcal{I}, \pi \models \phi$ **iff**
Concept atom	$A(x)$	$\pi(x) \in A^{\mathcal{I}}$
Role atom	$r(x, y)$	$(\pi(x), \pi(y)) \in r^{\mathcal{I}}$
Negation/Complement	$\neg\phi_1$	$\mathcal{I}, \pi \not\models \phi_1$
Conjunction	$\phi_1 \wedge \phi_2$	$\mathcal{I}, \pi \models \phi_1$ and $\mathcal{I}, \pi \models \phi_2$
Disjunction	$\phi_1 \vee \phi_2$	$\mathcal{I}, \pi \models \phi_1$ or $\mathcal{I}, \pi \models \phi_2$
Implication	$\phi_1 \rightarrow \phi_2$	$\mathcal{I}, \pi \models \neg\phi_1 \vee \phi_2$
Universal quantification	$\forall x.\phi_1$	$\mathcal{I}, \pi\{x \mapsto a\} \models \phi_1$ for all $a \in \Delta^{\mathcal{I}}$
Existential quantification	$\exists x.\phi_1$	$\mathcal{I}, \pi\{x \mapsto a\} \models \phi_1$ for some $a \in \Delta^{\mathcal{I}}$

Figure 2.14.: Semantics of first-order queries, where $x, y \in N_T$, $A \in N_C$, $r \in N_R$, π is an assignment extended by $\{a \mapsto a^{\mathcal{I}} \mid a \in N_C\}$, and $\pi\{x \mapsto a\}$ denotes the overwriting of the mapping of x in π.

Definition 2.16 (Certain Answer Semantics). Let $\mathcal{K}$ be a consistent DL-KB and let $\phi(\mathbf{x})$ be an n-ary first-order query. A k-tuple $\mathbf{a}$ of individual names from $\mathrm{Ind}(\mathcal{K})$ is a *certain answer* to ϕ in $\mathcal{K}$, if it is an answer to ϕ in all model of $\mathcal{K}$.

We denote the set of all certain answers to ϕ over $\mathcal{K}$ by $\mathrm{cert}(\phi, \mathcal{K})$. $\diamondsuit$

Example 2.17. Consider the following KB $\mathcal{K}_{\texttt{Wood}}$ containing the information that John owns a wooden table:

$$\begin{array}{ll} \texttt{WoodenTable}(\texttt{TableOfJohn}) & \texttt{Table} \sqsubseteq \texttt{Furniture} \\ \texttt{owns}(\texttt{John}, \texttt{TableOfJohn}) & \texttt{WoodenTable} \sqsubseteq \texttt{Table} \sqcap \exists \texttt{madeOf}.\texttt{Wood} \end{array} \quad \diamondsuit$$

Suppose we want to find persons that own furniture that is entirely made up of wood. We pose a first-order query with one answer variable x that should be assigned to the persons we are interested in:

$$\phi_1(x) = \forall y.\texttt{owns}(x, y) \wedge \texttt{Furniture}(y) \rightarrow \forall z.\texttt{madeOf}(y, z) \wedge \texttt{Wood}(z) \tag{2.3}$$

When evaluating ϕ_1 over $\mathcal{I}_1$ in Figure 2.18, we obtain `John` as an answer. However, `John` is not a certain answer to the query. While it is possible that `John` owns only a single wooden table and that table is made entirely out of wood, it is not certain, since, for example, $\mathcal{I}_2$ is also a model and provides the alternative possibility that some part of the table could be made of metal. When using certain answer semantics, as the name suggests, we only get answers that are certain, i.e. that hold in every possible model.

2.2.1. Conjunctive Queries

An important fragment of general first-order queries are *conjunctive queries* (CQs). While not being as expressive, they are powerful enough for many use-cases and often offer a lower complexity.

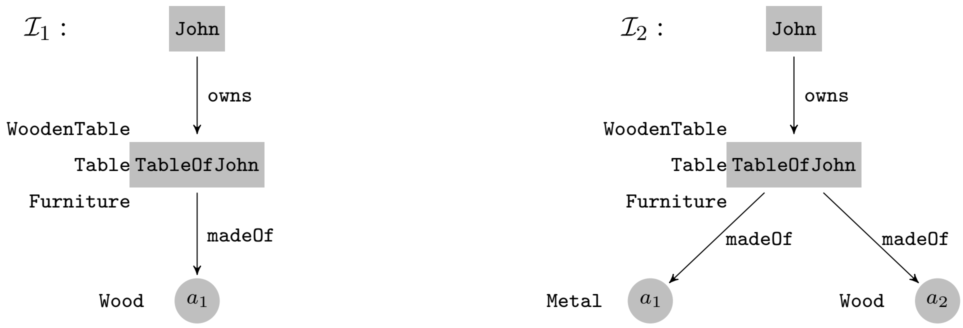


Figure 2.18.: Two possible models of the KB from Example 2.17.

Definition 2.19 (Conjunctive Query). A *conjunctive query* (CQ) $q(\mathbf{x})$ is a first-order query of the form $\exists \mathbf{y}.\, \varphi(\mathbf{x}, \mathbf{y})$, where φ is a conjunction of atoms.

Abusing notation, we write $\alpha \in q$ if the atom α occurs in q, and conversely may treat a set of atoms as a conjunction. The *leaf variables* x in q are those that do not occur in any atoms of the form $r(x, y)$. Clearly, q is satisfied in an interpretation if there is a satisfying assignment for $\varphi(\mathbf{x}, \mathbf{y})$, which is often called a *match* for q.

Two variables y, y' in q are *connected* if there exists a sequence of variables $(y_0, y_1, \ldots, y_n)$ in q with $y_0 = y$, $y_n = y'$ and for each $0 \leq i < n$ there exists $r(y_i, y_{i+1}) \in q$ or $r(y_{i+1}, y_i) \in q$ for $r \in N_R$. A CQ q is *connected* if every variable is connected to every other variable in q.

A CQ q is *rooted* if q has at least one answer variable and all quantified variables are connected to an answer variable in q. $\diamondsuit$

Continuing our example, suppose we modify our query and ask for persons that own some kind of furniture made out of wood. While $\phi_1(x)$ could not be expressed as a CQ, it is possible this time:

$$\phi_2(x) = \exists y. \exists z. \mathtt{owns}(x, y) \land \mathtt{Furniture}(y) \land \mathtt{madeOf}(y, z) \land \mathtt{Wood}(z) \tag{2.4}$$

For ϕ_2, **John** is a certain answer, since in all possible worlds (models), **John** owns a table that is at least made up of wood.

2.2.2. Complexity of OMQA

The complexity of OMQA depends on the underlying DL $\mathcal{L}$ and the chosen query language Q. For the purpose of analyzing the complexity of answering Q-queries over $\mathcal{L}$, OMQA can be viewed as a decision problem in the following way: Given a $\mathcal{L}$-KB $\mathcal{K} = (\mathcal{T}, \mathcal{A})$, a k-ary Q-query $\phi(\mathbf{x})$, and a k-ary tuple $\mathbf{a}$, does $\mathbf{a} \in \mathrm{cert}(\phi, \mathcal{K})$ hold?

The complexity of query answering is usually viewed in two different ways [Var82]:

- The *combined complexity* is measured as a function depending on the size of all inputs, i.e. $|\phi| + |\mathcal{T}| + |\mathcal{A}| + |\mathbf{a}|$.

- For *data complexity*, all inputs except the ABox (containing the data) are assumed to be constant. Therefore, they do not contribute to the complexity of query answering.

The key observation is that in practice the queries are usually very small compared to the amount of data contained in the KB. The same holds for the TBox, which might contain thousands of axioms, but is still negligible compared to millions of data assertions. For this reason, data complexity is generally considered to be the more useful complexity measure in the database world as well as the OMQA world.

In $\mathcal{ALC}$ answering CQs is in coNP in data complexity and ExpTime in combined complexity [OSE08]. As we will see in the following CQ answering in $\mathcal{ELH}_\bot$ is less complex.

$\mathcal{ELH}_\bot$ enjoys the *universal model property* (in the literature also referred to as *canonical model property*): There exist certain models that can be homomorphically mapped into any other model. Intuitively, these models satisfy all constraints of a given KB in the most general way [ORS11].

Definition 2.20 (Homomorphism and Universal Model). Let $\mathcal{K}$ be a consistent $\mathcal{ELH}_\bot$-KB and $\mathcal{I}, \mathcal{J}$ models of $\mathcal{K}$. A mapping $h : \Delta^\mathcal{I} \to \Delta^\mathcal{J}$ from $\mathcal{I}$ to $\mathcal{J}$ is a *homomorphism* if

- $h(a^\mathcal{I}) = a^\mathcal{J}$ for all $a \in N_I$,
- if $d \in C^\mathcal{I}$ then $h(d) \in C^\mathcal{J}$, for all $d \in \Delta^\mathcal{I}$ and $C \in N_C$, and
- if $(d, e) \in r^\mathcal{I}$ then $(h(d), h(e)) \in r^\mathcal{J}$, for all $d, e \in \Delta^\mathcal{I}$ and $r \in N_R$.

A homomorphism is said to be *strong*, if in the last two conditions the implications are replaced with equivalences, i.e. $d \in C^\mathcal{I}$ iff $h(d) \in C^\mathcal{J}$ and $(d, e) \in r^\mathcal{I}$ iff $(h(d), h(e)) \in r^\mathcal{J}$. An *endomorphism* of $\mathcal{I}$ is a homomorphism $h : \Delta^\mathcal{I} \to \Delta^\mathcal{I}$ from $\mathcal{I}$ into itself.

A model $\mathcal{I}$ of $\mathcal{K}$ is a *universal model* of $\mathcal{K}$, denoted by $\mathcal{I}_\mathcal{K}$, if for all models $\mathcal{J}$ of $\mathcal{K}$ there is a homomorphism from $\mathcal{I}$ to $\mathcal{J}$, denoted by $\mathcal{I} \to \mathcal{J}$. ♢

Because universal models can be homomorphically mapped into any other model, every match π in a universal model $\mathcal{I}_\mathcal{K}$ of a consistent KB $\mathcal{K}$ corresponds to matches in all other models of $\mathcal{K}$. A match in a given model $\mathcal{J}$ can be constructed by composing π with the respective homomorphism h from $\mathcal{I}_\mathcal{K}$ to $\mathcal{J}$. Many proposed reasoning procedures exploit this property to obtain optimal complexity bounds. The following lemma shows that in order to obtain certain answers to a given query, it is sufficient to evaluate the query exclusively over a universal model.

Lemma 2.21. Let $\mathcal{K}$ be a consistent $\mathcal{ELH}_\bot$-KB, $\mathcal{I}_\mathcal{K}$ a universal model of $\mathcal{K}$, and $q(\mathbf{x})$ a CQ. Then $\mathrm{cert}(q, \mathcal{K}) = \mathrm{ans}(q, \mathcal{I}_\mathcal{K})$.

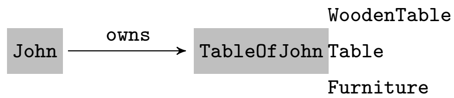
In general, universal models can be infinite so the query cannot be evaluated over them directly in practice. To overcome this it has been shown in [LTW09] that CQ answering over $\mathcal{ELH}_\bot$-KBs is *combined first-order rewritable*: For any CQ q and consistent KB $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ we can find a finite set of first-order queries $Q_\mathcal{T}$ and a finite interpretation $\mathcal{I}_\mathcal{K}^{\text{fin}}$ such that $\mathrm{cert}(q, \mathcal{K}) = \bigcup_{q_\mathcal{T} \in Q_\mathcal{T}} \mathrm{ans}(q_\mathcal{T}, \mathcal{I}_\mathcal{K}^{\text{fin}})$. Importantly, $\mathcal{I}_\mathcal{K}^{\text{fin}}$ is independent of q, i.e.

can be reused to answer many different queries, while $Q_{\mathcal{T}}$ is independent of $\mathcal{A}$, i.e. each query can be rewritten without using the (possibly large) dataset.

Consider query ϕ_2 from Equation (2.4). We know that if a person owns a wooden table, then this person also owns a piece of furniture that is made up of wood by the GCI $\texttt{WoodenTable} \sqsubseteq \texttt{Furniture} \sqcap \exists\texttt{madeOf}.\texttt{Wood}$ in Example 2.17. Therefore, ϕ_2 could be rewritten to the following query:

$$\phi'_2(x) = \exists y.\texttt{owns}(x, y) \wedge \texttt{WoodenTable}(y)$$

A finite interpretation $\mathcal{I}_{\mathcal{K}}^{\text{fin}}$ can be constructed in polynomial time, essentially by ignoring all anonymous objects from a canonical model $\mathcal{I}_{\mathcal{K}}$:



While the original query ϕ_2 does not yield any answers over $\mathcal{I}_{\mathcal{K}}^{\text{fin}}$, **John** is still returned as an answer to ϕ'_2 and is therefore a certain answer.

Regarding complexity, the finite interpretation $\mathcal{I}_{\mathcal{K}}^{\text{fin}}$ can be constructed in polynomial time in the size of a given KB $\mathcal{K} = (\mathcal{A}, \mathcal{T})$, and a single rewriting ϕ' of a given query ϕ can be constructed in polynomial time in the size of q and $\mathcal{T}$. The obvious NP algorithm can then be used to check whether a given tuple $\mathbf{a}$ is an answer, formally $\mathbf{a} \in \text{ans}(\mathcal{I}_{\mathcal{K}}^{\text{fin}}, q)$. Since the rewriting is done independent of the ABox, it does not influence the data complexity. As a consequence, CQ answering over a $\mathcal{ELH}_{\bot}$ ontology is PTime-complete in data complexity and NP-complete in combined complexity [Ros07a; LTW09].

The $\mathcal{EL}$ family of DLs belongs to the class of *Horn-DLs*. Their distinctive feature is the incapability of expressing disjunction, which means that Horn-DL KBs can be translated to the Horn-fragment of FOL. Because there is no disjunction, they enjoy the universal model property [BO15]. By exploiting this it has been shown that CQ answering stays tractable in data complexity even up to the very expressive Horn-$\mathcal{SROIQ}$ [ORS11].

2.3. Metric Linear Temporal Logic

In many applications a single static object dimension is not enough. Often things change over time, for example a patient's diagnoses or the values of a sensor. In such cases temporal logics need to be employed. In *linear-time* logics a linear flow of time is assumed, i.e. each time point has exactly one successor. A well-investigated such logic is *Linear-Time Temporal Logic* (LTL) [Pnu77]. There exist also temporal logics with a branching flow of time, for example *Computation Tree Logic* (CTL) [CE81]. However, in this thesis, we consider only combinations of DLs with LTL. In the following we briefly introduce LTL with and without intervals in binary.

Definition 2.22 (Syntax of propositional LTL). Let $\mathcal{P} = \{p_1, \ldots, p_n\}$ be a finite set of *propositional variables*. The set of *propositional* LTL *formulas* is the smallest set such that

- if $p \in \mathcal{P}$, then p is a propositional LTL-formula over $\mathcal{P}$; and

$$\mathfrak{W}_1: \cdots \longrightarrow \underset{-1}{\overset{\emptyset}{\bullet}} \longrightarrow \underset{0}{\overset{\{p_1,p_2\}}{\bullet}} \longrightarrow \underset{1}{\overset{\{p_2\}}{\bullet}} \longrightarrow \underset{2}{\overset{\{p_3\}}{\bullet}} \longrightarrow \underset{3}{\overset{\{p_1\}}{\bullet}} \longrightarrow \cdots$$

$$\mathfrak{W}_2: \cdots \longrightarrow \underset{-1}{\overset{\{p_1,p_2\}}{\bullet}} \longrightarrow \underset{0}{\overset{\{p_3\}}{\bullet}} \longrightarrow \underset{1}{\overset{\{p_1,p_2\}}{\bullet}} \longrightarrow \underset{2}{\overset{\{p_2\}}{\bullet}} \longrightarrow \underset{3}{\overset{\{p_3\}}{\bullet}} \longrightarrow \cdots$$

Figure 2.23.: Two propositional LTL-structures $\mathfrak{W}_1$ and $\mathfrak{W}_2$ for Example 2.24.

- if ϕ_1 and ϕ_2 are propositional LTL-formulas, then so are $\neg\phi_1$ (negation), $\phi_1 \wedge \phi_2$, (conjunction), $\bigcirc\phi_1$ (next), $\bigcirc^-\phi_1$ (previous), $\phi_1 \mathcal{U} \phi_2$ (until), and $\phi_1 \mathcal{S} \phi_2$ (since). $\diamondsuit$

In the following we omit the set $\mathcal{P}$ and talk about propositional LTL-formulas rather than propositional LTL-formulas over $\mathcal{P}$. We introduce the usual abbreviations of temporal logics:

- $\phi_1 \vee \phi_2$ (disjunction) as an abbreviation for $\neg(\neg\phi_1 \wedge \neg\phi_2)$,
- $\top$ (top) as an abbreviation for an arbitrary, but fixed propositional tautology $p \vee \neg p$ for some $p \in \mathcal{P}$,
- $\bot$ (bottom) as an abbreviation for $\neg\top$,
- $\Diamond\phi_1$ (eventually) as an abbreviation for $\top \mathcal{U} \phi_1$,
- $\Box\phi_1$ (always) as an abbreviation for $\neg\Diamond\neg\phi_1$,
- $\Diamond^-\phi_1$ (once) as an abbreviation for $\top \mathcal{S} \phi_1$,
- $\Box^-\phi_1$ (always in the past) as an abbreviation for $\neg\Diamond^-\neg\phi_1$.

The semantics is given by *LTL-structures*, an infinite sequence of *worlds*, $\mathfrak{W} = (w_i)_{i\in\mathbb{Z}}$, where $w_i \subseteq P$. We write

$\mathfrak{W}, i \models p$	iff	$p \in w_i$ if $p \in P$
$\mathfrak{W}, i \models \neg\phi_1$	iff	$\mathfrak{W}, i \not\models \phi_1$
$\mathfrak{W}, i \models \phi_1 \wedge \phi_2$	iff	$\mathfrak{W}, i \models \phi_1$ and $\mathfrak{W}, i \models \phi_2$
$\mathfrak{W}, i \models \bigcirc\phi_1$	iff	$\mathfrak{W}, i+1 \models \phi_1$
$\mathfrak{W}, i \models \bigcirc^-\phi_1$	iff	$\mathfrak{W}, i-1 \models \phi_1$
$\mathfrak{W}, i \models \phi_1 \mathcal{U} \phi_2$	iff	$\exists k \in \mathbb{N}$: $\mathfrak{W}, i+k \models \phi_2$ and $\mathfrak{W}, i+j \models \phi_1$ for all $0 \leq j < k$
$\mathfrak{W}, i \models \phi_1 \mathcal{S} \phi_2$	iff	$\exists k \in \mathbb{N}$: $\mathfrak{W}, i-k \models \phi_2$ and $\mathfrak{W}, i-j \models \phi_1$ for all $0 \leq j < k$

If $\mathfrak{W}, 0 \models \phi$, then we call $\mathfrak{W}$ a model of ϕ. A propositional LTL-formula ϕ is satisfiable if ϕ has a model.

Example 2.24. Let $\phi := p_1 \wedge (\bigcirc\neg p_1) \wedge (p_2 \mathcal{U} p_3)$ be a propositional LTL-formula. Consider the two propositional LTL-structures $\mathfrak{W}_1$ and $\mathfrak{W}_2$ that are depicted graphically in Figure 2.23. For $\mathfrak{W}_1$ we have $\mathfrak{W}_1, 0 \models \phi$, hence ϕ is satisfiable. $\mathfrak{W}_2$ is not a model at time point 0, but $\mathfrak{W}_2, -1 \models \phi$ and $\mathfrak{W}_2, 1 \models \phi$. $\diamondsuit$

Note that while usually the natural numbers are used as timeline in LTL, we use the integers here. Satisfiability checking in propositional LTL is known to be PSPACE-complete [SC85].

2.3.1. LTL with Binary Intervals

To make formulas more concise, we extend LTL by intervals given in binary and obtain $\text{LTL}^{\mathsf{bin}}$. Additionally, we generalize the propositional variables to arbitrary $\mathcal{Q}$-axioms, where $\mathcal{Q}$ is a logic and obtain $\mathcal{Q}$-$\text{LTL}^{\mathsf{bin}}$, which can be read as '$\text{LTL}^{\mathsf{bin}}$ over $\mathcal{Q}$-axioms'.

Definition 2.25. Let $\mathcal{Q}$ be a logic. $\mathcal{Q}$-$\text{LTL}^{\mathsf{bin}}$ formulas are of the form

$$\phi_1, \phi_2 := \alpha \mid \top \mid \neg\phi_1 \mid \phi_1 \wedge \phi_2 \mid \bigcirc\phi_1 \mid \phi_1 \mathcal{U}_I \phi_2 \mid \phi_1 \mathcal{S}_I \phi_2 \tag{2.5}$$

where α is a $\mathcal{Q}$-axiom and I is an interval of the form $[a, b]$, where $a \in \mathbb{Z} \cup \{-\infty\}$ and $b \in \mathbb{Z} \cup \{\infty\}$. If the boundaries are $\pm\infty$, the interval is open, but we use the same brackets. ♢

Definition 2.26. A *(temporal $\mathcal{Q}$-)interpretation* is a structure $\mathfrak{I} = (\Delta^{\mathfrak{I}}, (\mathcal{I}_i)_{i\in\mathbb{Z}})$, where each $\mathcal{I}_i = (\Delta^{\mathfrak{I}}, \cdot^{\mathcal{I}_i})$ is a $\mathcal{Q}$ interpretation over $\Delta^{\mathfrak{I}}$ (*constant domain assumption*) and $a^{\mathcal{I}_i} = a^{\mathcal{I}_j}$ for all $a \in N_I$ and $i, j \in \mathbb{Z}$. Validity of an $\mathcal{Q}$-$\text{LTL}^{\mathsf{bin}}$-formula ϕ at time point $i \in \mathbb{Z}$ w.r.t. a temporal $\mathcal{Q}$-interpretation $\mathfrak{I}$, denoted by $\mathfrak{I}, i \models \phi$, is defined inductively as follows:

$$\begin{array}{lll}
\mathfrak{I}, i \models \alpha & \text{iff} & \mathcal{I}_i \models \alpha \qquad \text{(where } \alpha \text{ is a } \mathcal{Q}\text{-axiom)} \\
\mathfrak{I}, i \models \neg\phi & \text{iff} & \mathfrak{I}, i \not\models \phi \\
\mathfrak{I}, i \models \phi_1 \wedge \phi_2 & \text{iff} & \mathfrak{I}, i \models \phi_1 \text{ and } \mathfrak{I}, i \models \phi_2 \\
\mathfrak{I}, i \models \phi_1 \mathcal{U}_I \phi_2 & \text{iff} & \text{there exists } k \in I \text{ such that } \mathfrak{I}, i + k \models \phi_2 \\
 & & \quad \text{and for all } 0 \leq j < k : \mathfrak{I}, i + j \models \phi_1 \\
\mathfrak{I}, i \models \phi_1 \mathcal{S}_I \phi_2 & \text{iff} & \text{there exists } k \in I \text{ such that } \mathfrak{I}, i - k \models \phi_2 \\
 & & \quad \text{and for all } 0 \leq j < k : \mathfrak{I}, i - j \models \phi_1
\end{array}$$

♢

With other words, the concept $\phi_1 \mathcal{U}_I \phi_2$ requires ϕ_2 to be satisfied at some point in the interval I, and ϕ_1 to hold at all time points before that. For instance, for $\mathfrak{W}_1$ in Figure 2.23 it holds that $\mathfrak{W}_1, 0 \models p_2 \mathcal{U}_{[0,2]} p_3$, but $\mathfrak{W}_1, 0 \not\models p_2 \mathcal{U}_{[0,1]} p_3$.

As usual, we define the following additional operators:

- the quantified *next* operator $\bigcirc_i \alpha := \top \mathcal{U}_{[i,i]} \alpha$, which states that α has to hold at the i-th next time point;
- the quantified *eventually* operator $\Diamond_I \alpha := \top \mathcal{U}_I \alpha$, which states that at some time point in the given interval I relative to the current time point α has to hold;
- the quantified *always* operator $\Box_I \alpha := \neg\Diamond_I \neg\alpha$, which states that α has to hold at all time points in the given interval I relative to the current time point; and

- unquantified variants of until, eventually and always:

$$\alpha \mathcal{U} \beta := \alpha \mathcal{U}_{[0,\infty)} \beta$$
$$\Diamond\alpha := \top \mathcal{U} \alpha$$
$$\Box\alpha := \neg\Diamond\neg\alpha.$$

- the symmetric past operators $\Diamond^-_I$, $\Diamond^-$, $\bigcirc^-{}_i$, $\Box^-_I$, and $\Box^-$ are defined similarly by using $\mathcal{S}_I$ instead of $\mathcal{U}_I$.

The binary intervals can be simulated using usual LTL with an exponential blowup [AH93]. While this blowup cannot be avoided in general, it is shown in [LWW07] that using intervals where the lower bound is always 0 does not increase the complexity compared to the qualitative case. For more information see [GJO16b; BBK+17].

Lemma 2.27 (based on [BBK+17]). Each formula in $\mathcal{Q}$-LTL$^{\mathsf{bin}}$ is equivalent to an $\mathcal{Q}$-LTL formula. Each interval can be expressed by exponentially many disjuncts and nestings of the $\bigcirc$-operator.

Chapter 3.

Minimal-World Semantics for Conjunctive Queries with Negation

Clinical trials play an important role in the evaluation of new medications and treatments. For example, to test a new treatment for a specific kind of cancer, this treatment needs to pass different tests in order to be approved. Eventually, it needs to be tested on humans. Consequently, after designing a study, the first main task is to find patients that are eligible for the study, i.e. that satisfy all *inclusion criteria* and do not satisfy any *exclusion criteria*. Unfortunately, in practice it is often a resource-intensive task to recruit enough patients to get statistically meaningful results. However, the increased usage of *Electronic Health Records* (EHRs) in hospitals offers a promising opportunity to improve the recruitment process by automating parts of it. An EHR contains information about the measurements, diagnoses, treatments and many other things of a patient. An actual EHR can be found in Example 7.4. As emphasized in [PCD+07], a major challenge lies in the fact that criteria are described on different levels of granularity, which range from quite specific to very general. This can often be bridged by using medical background knowledge that links broad categories ('*lung cancer*') to more specific ones ('*adenocarcinoma*') or even to more detailed descriptions ('*malignant neoplasm was found in the left lower lobe*'). Fortunately, medical ontologies such as SNOMED CT, formulated in $\mathcal{ELH}_\bot$, contain a large amount of medical knowledge that can be used in an OMQA setting. As we will see, what is missing are suitable semantics for this setting.

Therefore, in this chapter we develop new closed-world semantics and apply them to $\mathcal{ELH}_\bot$. In Section 3.1 we introduce a formal patient selection problem that will serve as motivation for our theoretical explorations. We then discuss existing non-monotonic formalisms that could be used to solve the patient selection problem in Section 3.2. While most of them yield intuitive results for instance queries and queries with only answer variables, the formalisms give unexpected results in the face of negation over anonymous individuals.

To overcome this we introduce *minimal-world semantics* in Section 3.3 and show that it gives answers to CQs with negation that are intuitive w.r.t. the patient selection problem. Through a first-order rewriting strategy, we are able to prove that reasoning is tractable in data complexity in Section 3.4. Finally, we discuss possible generalizations of minimal-world semantics to more expressive DLs in Section 3.5.

3.1. The Patient Selection Problem

The following patients are inspired by actual patients in the MIMIC-III[1] dataset. It is a de-identified dataset of hospital admissions collected in two different hospitals in the US. In total it includes data associated with over 40.000 patients who stayed in intensive care units.

We consider three patients: Patient p_1 (patient 2693 in the MIMIC-III dataset) is diagnosed with breast cancer and an unspecified form of cancer. This often occurs when there are multiple mentions of cancer in a patient's EHR, which cannot be resolved to be the same entity. Patient p_2 (patient 32304 in the MIMIC-III dataset) suffers from breast cancer and skin cancer ('[S]tage IV breast cancer with mets to skin, bone, and liver'). For p_3 (patient 88432 in the MIMIC-III dataset), we know that p_3 has breast cancer that involves the skin ('Skin, left breast, punch biopsy: Poorly differentiated carcinoma').

Since SNOMED CT does not model patients, we add a special role name **diagnosedWith** that connects patients with their diagnoses. One can use this to express diagnoses in two ways. First, one can explicitly introduce individual names for diagnoses in assertions like

$$\mathtt{diagnosedWith}(p_1, d_1), \qquad \mathtt{diagnosedWith}(p_1, d_2),$$
$$\mathtt{BreastCancer}(d_1), \qquad \mathtt{Cancer}(d_2),$$

implying that these diagnoses are treated as distinct entities under the standard name assumption. Alternatively, one can use complex assertions like $\exists\mathtt{diagnosedWith}.\mathtt{Cancer}(p_1)$, which allows the logical semantics to resolve whether two diagnoses actually refer to the same object. Since ABoxes only contain concept names, in this case one has to introduce auxiliary definitions like

$$\mathtt{CancerPatient} \equiv \exists\mathtt{diagnosedWith}.\mathtt{Cancer}$$

into the TBox. We use both variants in our example, to illustrate their different behaviors.

Example 3.1 (Patient Selection Task). We obtain the KB $\mathcal{K}_C$, containing knowledge about different kinds of cancers and cancer patients, together with information about the three patients. The information about cancers is taken from SNOMED CT (in simplified form):

$$\mathtt{SkinCancer} \equiv \mathtt{Cancer} \sqcap \exists\mathtt{findingSite}.\mathtt{SkinStructure} \tag{T1}$$
$$\mathtt{BreastCancer} \equiv \mathtt{Cancer} \sqcap \exists\mathtt{findingSite}.\mathtt{BreastStructure} \tag{T2}$$
$$\mathtt{SkinOfBreastCancer} \equiv \mathtt{Cancer} \sqcap \exists\mathtt{findingSite}.\mathtt{SkinOfBreastStructure} \tag{T3}$$
$$\mathtt{SkinOfBreastStructure} \sqsubseteq \mathtt{BreastStructure} \sqcap \mathtt{SkinStructure} \tag{T4}$$

Additionally, we add the following auxiliary definitions to the TBox:

$$\mathtt{CancerPatient} \equiv \exists\mathtt{diagnosedWith}.\mathtt{Cancer} \tag{T5}$$
$$\mathtt{SkinCancerPatient} \equiv \exists\mathtt{diagnosedWith}.\mathtt{SkinCancer} \tag{T6}$$
$$\mathtt{BreastCancerPatient} \equiv \exists\mathtt{diagnosedWith}.\mathtt{BreastCancer} \tag{T7}$$

[1] `https://mimic.physionet.org`

The EHRs are compiled into several assertions per patient yielding the following ABox:

$$\begin{aligned}
\texttt{Patient } p_1\texttt{:}\ & \texttt{BreastCancerPatient}(p_1) && \text{(A1)}\\
& \texttt{CancerPatient}(p_1) && \text{(A2)}\\
\texttt{Patient } p_2\texttt{:}\ & \texttt{SkinCancerPatient}(p_2) && \text{(A3)}\\
& \texttt{BreastCancerPatient}(p_2) && \text{(A4)}\\
\texttt{Patient } p_3\texttt{:}\ & \texttt{diagnosedWith}(p_3, c_3) && \text{(A5)}\\
& \texttt{SkinOfBreastCancer}(c_3) && \text{(A6)}
\end{aligned}$$

For example, skin cancers and breast cancers are cancers occurring at specific parts of the body (see T1 and T2), called '*body structure*' in SNOMED CT, and a breast cancer patient is someone who is diagnosed with breast cancer (see T7). This means that, in every model of $\mathcal{K}_C$, every object that satisfies **BreastCancerPatient** (in particular p_2 in A3) must have a **diagnosedWith**-connected object that satisfies BreastCancer, and so on.

For a clinical trial,[2] we want to find patients that have 'breast cancer', but not 'breast cancer that involves the skin.' This can be translated into an NCQ:

$$\begin{aligned}
q_B(x) := \exists y, z.\, & \texttt{diagnosedWith}(x, y) \wedge \texttt{Cancer}(y) \wedge \texttt{findingSite}(y, z) \wedge \\
& \texttt{BreastStructure}(z) \wedge \neg\texttt{SkinStructure}(z) \qquad \diamondsuit
\end{aligned}$$

We now describe which patients should intuitively be returned as answers. We know that p_1 is diagnosed with **BreastCancer** as well as **Cancer**. Since the former is more specific, we assume that the latter refers to the same **BreastCancer**. However, since we have no information about an involvement of the skin, p_1 should be returned as an answer to q_B.

We know that p_2 suffers from cancer in the skin and the breast, but not if the skin of the breast is also affected. Since neither location is implied by the other, we assume that they refer to distinct areas. p_2 should thus be an answer to q_B.

In the case of p_3, it is explicitly stated that it is the same cancer that is occurring (not necessarily exclusively) at the skin of the breast. In this case, the ABox assertions override the distinctness assumption we made for p_2. Thus, p_3 should not be an answer to q_B.

3.2. Existing Non-Monotonic Formalisms

We now evaluate existing semantics on the patient selection example task. Before we look into specific formalisms, consider Figure 3.2, which shows an interpretation $\mathcal{I}_C$ that is a model of $\mathcal{K}_C$ w.r.t. the standard semantics. The domain of $\mathcal{I}_C$ consists of the named individuals from the ABox, namely p_1, p_2, p_3 and c_3 together with seven anonymous individuals, depicted in circles. For each individual, the concept names next to it hold in $\mathcal{I}_C$. For example, p_2 is asserted to be a **BreastCancerPatient**, a **SkinCancerPatient**, and a **CancerPatient**. It can be easily checked that in $\mathcal{I}_C$ is a model. For instance, by

[2] https://clinicaltrials.gov/ct2/show/NCT01960803

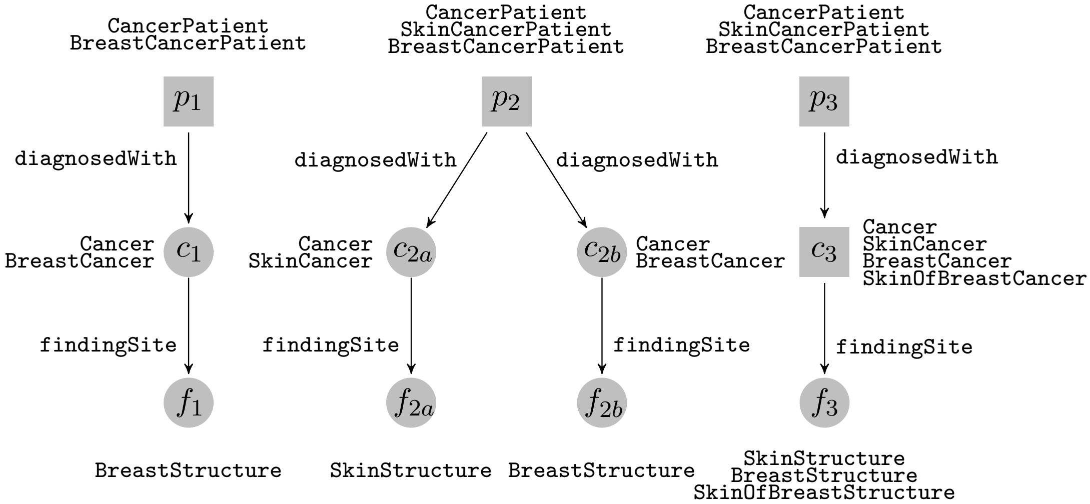


Figure 3.2.: A canonical model $\mathcal{I}_C$ of $\mathcal{K}_C$. Named individuals are depicted in squares, anonymous objects in circles.

axioms T1, T5 and T6 it is easy to see that $\mathcal{K}_C \models$ **SkinCancerPatient** $\sqsubseteq$ **CancerPatient**, hence p_2 satisfies **CancerPatient** in $\mathcal{I}_C$. Moreover $\mathcal{I}_C$ is canonical, i.e. it satisfies $\mathcal{K}_C$ in the most general way, for example, by introducing c_{2a} and c_{2b} instead of only one merged version of them. Having this model at hand we proceed now to discuss the behavior of different formalisms on the patient selection task.

3.2.1. Standard Certain Answer Semantics

Certain answer semantics as defined in Section 2.2 is clearly not suited here, because one can easily construct a model of $\mathcal{K}_C$ in which c_1 is also a skin cancer, and hence p_1 is not an element of cert$(q_B, \mathcal{K}_C)$. Moreover, under certain answer semantics answering CQs with guarded negation is already coNP-complete [GIK+15], and hence not (combined) rewritable. So clearly we need non-monotonic formalisms.

3.2.2. Epistemic Logic

As non-monotonic formalism, *epistemic logic* allows us to selectively apply closed-world reasoning using the modal knowledge operator **K**. For a formula $\mathbf{K}\varphi$ to be true, it has to hold in all 'connected worlds', which is often considered to mean all possible models of the KB, adopting an **S**5-like view [CDL+06]. For q_B, we could read $\neg$**SkinStructure**(z) as 'not known to be a skin structure', i.e. $\neg$**KSkinStructure**(z). Consider the model $\mathcal{I}_C$ in Figure 3.2 and the assignment $\pi = \{x \mapsto p_3, y \mapsto c_3, z \mapsto f_3\}$, for which we want to check whether it is a match for q_B. Under epistemic semantics, $\neg$**KSkinStructure**(z) is considered true if $\mathcal{K}$ has a (different) model in which f_3 does not belong to **SkinStructure**. However, f_3 is an anonymous object, and hence its name is not fixed. For example, we can easily obtain another model by renaming f_3 to f_1 and vice versa. Then f_3 would not be a skin structure, which means that $\neg$**KSkinStructure**(z) is true in the original model $\mathcal{I}_C$, which is not what we expected. This is a known problem with epistemic first-order logics [Wol00].

3.2.3. Skolemization

To enforce a stricter comparison of anonymous objects between models *Skolemization* could be used. The inclusion $\mathtt{SkinOfBreastCancer} \sqsubseteq \exists\mathtt{findingSite}.\mathtt{SkinOfBreast}$ could be rewritten as the first-order sentence

$$\forall x. \Big(\mathtt{SkinOfBreastCancer}(x) \to \mathtt{findingSite}(x, f(x)) \land \mathtt{SkinOfBreast}\big(f(x)\big)\Big),$$

where f is a fresh function symbol. This means that c_3 would be connected to a finding site that has the unique name $f(c_3)$ in every model. Queries would be evaluated over Herbrand models only. Hence, for evaluating $\neg\mathbf{K}\mathtt{SkinStructure}(z)$ when z is mapped to $f(c_3)$, we would only be allowed to compare the behavior of $f(c_3)$ in other Herbrand models. The general behavior of this anonymous individual is fixed, however, since in all Herbrand models it is *the* finding site of c_3.

While this improves the comparison by introducing pseudo-names for all anonymous individuals, it limits us in different ways: Since p_3 is inferred to be a $\mathtt{BreastCancerPatient}$, the Skolemized version of $\mathtt{BreastCancerPatient} \sqsubseteq \exists\mathtt{diagnosedWith}.\mathtt{BreastCancer}$ introduces a new successor $g(p_3)$ of p_3 satisfying $\mathtt{BreastCancer}$, which, together with the definition of $\mathtt{BreastCancer}$, means that p_3 is an answer to q_B since there is an additional breast cancer diagnosis that does not involve the skin.

3.2.4. Datalog-based Ontology Languages

In recent years many different variants of Datalog-based ontology languages with negation [HKL+13; AGP14] have been introduced. They are closely related to Skolemized ontologies, since their semantics is often based on the so-called *Skolem chase* [Mar09]. This is closer to the semantics we propose in Section 3.3, in that a single canonical model is used for all inferences. However, it suffers from the same drawback of Skolemization described above, due to superfluous successors. To avoid this, our semantics uses a special canonical model (see Definition 3.6), which is similar to the *restricted chase* [FKM+05] or the *core chase* [DNR08], but always produces a unique model without having to merge domain elements. To the best of our knowledge, there exist no complexity results for Datalog-based languages with negation over these other chase variants. A more detailed discussion will follow in Section 3.5.

3.2.5. Closed Predicates

Closed predicates are a way to declare, for example, the concept name $\mathtt{SkinCancer}$ as 'closed', which means that all skin cancers must be declared explicitly, and no other $\mathtt{SkinCancer}$ object can exist [LSW13; AOS16]. This provides another way to give answers to negated atoms as in q_B. However, as explained in the introduction, this mechanism is not suitable for anonymous objects since it means that only named individuals can satisfy $\mathtt{SkinCancer}$. In particular, in every model of $\mathcal{K}_C$ c_3 has to be the only individual satisfying $\mathtt{SkinCancer}$, since that is the only skin cancer that occurs in $\mathcal{K}_C$ explicitly.

When applied to $\mathcal{K}_C$, the result is even worse: Since there is no (named) **`SkinStructure`** object, no skin structures can exist at all and $\mathcal{K}_C$ becomes inconsistent. Closed predicates are appropriate in cases where the KB contains a full list of all instances of a certain concept name, and no other objects should satisfy it; but they are not suitable to infer negative information about anonymous objects. Moreover, CQ answering with closed predicates in $\mathcal{ELH}_\bot$ is already coNP-hard [LSW13].

3.3. Minimal-World Semantics for $\mathcal{ELH}_\bot$

We propose to answer NCQs over a special canonical model of the knowledge base. On the one hand, this eliminates the problem of tracking anonymous objects across different models, and on the other hand it enables us to encode our assumptions directly into the construction of the model. The model has to satisfy the following two assumptions:

Firstly, every anonymous individual has to be implied by the KB. If all we know is that a given patient suffers from skin cancer, there should only be diagnoses of skin cancer or super-concepts of skin cancer, but nothing else. It is unreasonable to consider arbitrary additional possibilities, for example, that the patient might also suffer from diabetes, even though it was never mentioned. This assumption is captured by universal models, which play a crucial role when answering CQs.

Secondly, the model should contain the minimum necessary number of anonymous objects since, unlike in standard CQ answering, the precise shape and number of anonymous objects has an impact on the semantics of negated atoms. If a patient is known to suffer from skin cancer, we assume that there is exactly one diagnosis of skin cancer and no other diagnosis of a generic cancer. If a patient is asserted to have diagnoses of skin cancer and generic cancer, we assume that the two assertions refer to the same diagnosis instead of two different ones. We do so because a diagnosis of skin cancer implies also a diagnoses of cancer. However when a patient has skin cancer and diabetes, the two diagnoses are unrelated and hence have to be different in a universal model. If redundancies are intended, they need to be modeled explicitly through the use of named individuals in the ABox.

When we combine the two assumptions, they imply that a model has to contain the 'minimal' number of individuals to still be a universal model.

Definition 3.3 (Minimal Universal Model). Let $\mathcal{K}$ be a consistent $\mathcal{ELH}_\bot$-KB. A model $\mathcal{I}$ of $\mathcal{K}$ is *minimal* iff every endomorphism is an isomorphism. A minimal model $\mathcal{I}$ of $\mathcal{K}$ is a minimal model of a model $\mathcal{J}$ of $\mathcal{K}$ if $\mathcal{I}$ is the image of an endomorphism of $\mathcal{J}$.

A model of $\mathcal{K}$ is a *minimal universal model* if it is universal w.r.t. $\mathcal{K}$ and minimal. $\diamondsuit$

We will see in Section 3.5 that minimality is closely related to cores, a concept from graph theory.

For the patient selection task (Example 3.1) this means that given $\mathcal{K}_C$, in contrast to the Skolemized semantics, we will not create both a generic **`Cancer`** and another **`BreastCancer`** successor for p_1, because **`BreastCancer`** is also a **`Cancer`**, and hence the first object is redundant. Therefore, in the minimal universal model of $\mathcal{K}_C$ depicted in Figure 3.2, for patient p_1 only one successor is introduced to satisfy the definitions of both **`BreastCancerPatient`** and **`CancerPatient`** at the same time. In contrast, p_2

has two successors, because **`BreastCancer`** and **`SkinCancer`** do not imply each other. Finally, for p_3 the ABox contains a single successor that is a **`SkinOfBreastCancer`**, which implies a single **`findingSite`**-successor that satisfies both **`SkinStructure`** and **`BreastStructure`**.

To detect whether an object required by an existential restriction $\exists r.A$ is redundant, we use the following notion of minimality. To obtain a clearer representation we assume w.l.o.g. that the KB contains no equivalent concept or role names, which can be checked in polynomial time in $\mathcal{ELH}_\bot$.

Definition 3.4 (Structural Subsumption). Let $\exists r.A, \exists t.B$ be concepts with $A, B \in N_C$ and $r, t \in N_R$. We say that $\exists r.A$ is *structurally subsumed* by $\exists t.B$ (written $\exists r.A \sqsubseteq^s_\mathcal{T} \exists t.B$) if $r \sqsubseteq_\mathcal{T} t$ and $A \sqsubseteq_\mathcal{T} B$.

Given a set V of existential restrictions, we say that $\exists r.A \in V$ is *minimal w.r.t.* $\sqsubseteq^s_\mathcal{T}$ *(in V)* if there is no $\exists t.B \in V$ such that $\exists t.B \sqsubseteq^s_\mathcal{T} \exists r.A$.

A CQ $q_1(\mathbf{x})$ is *structurally subsumed* by a CQ $q_2(\mathbf{x})$ with the same answer variables (written $q_1 \sqsubseteq^s_\mathcal{T} q_2$) if, for all $x, y \in \mathbf{x}$, it holds that

$$\bigsqcap_{\alpha(x)\in q_1} \alpha \sqsubseteq_\mathcal{T} \bigsqcap_{\alpha(x)\in q_2} \alpha, \text{ and } \bigsqcap_{\alpha(x,y)\in q_1} \alpha \sqsubseteq_\mathcal{T} \bigsqcap_{\alpha(x,y)\in q_2} \alpha,$$

where role conjunction is interpreted in the standard way [BCM+07]. $\diamondsuit$

In contrast to standard subsumption, $\exists r.A$ is not structurally subsumed by $\exists t.B$ w.r.t. the TBox $\mathcal{T} = \{\exists r.A \sqsubseteq \exists t.B\}$, as neither $r \sqsubseteq_\mathcal{T} t$ nor $A \sqsubseteq_\mathcal{T} B$ hold. Similarly, structural subsumption for CQs considers all (pairs of) variables separately. Structural subsumption ($\sqsubseteq^s_\mathcal{T}$) is stronger than standard subsumption ($\sqsubseteq_\mathcal{T}$):

Lemma 3.5. Let $\mathcal{K}$ be a consistent $\mathcal{ELH}_\bot$-KB. For existential restrictions $\exists r.A$ and $\exists t.B$ it holds that if $\exists r.A \sqsubseteq^s_\mathcal{T} \exists t.B$ then $\exists r.A \sqsubseteq_\mathcal{T} \exists t.B$.

Similarly, for CQs $q_1(\mathbf{x})$ and $q_2(\mathbf{x})$ with the same answer variables $\mathbf{x}$, it holds that if $q_1 \sqsubseteq^s_\mathcal{T} q_2$, then $\mathrm{cert}(q_1, \mathcal{K}) \subseteq \mathrm{cert}(q_2, \mathcal{K})$.

We use this notion to construct a canonical model. As we will see later, this model is a minimal universal model.

Definition 3.6 (Canonical Model Construction). Let $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ be an $\mathcal{ELH}_\bot$-KB. We construct the *canonical model* $\mathcal{I}_\mathcal{K}$ of $\mathcal{K}$ as follows:

1. Set $\Delta^{\mathcal{I}_\mathcal{K}} := N_I$ and $a^{\mathcal{I}_\mathcal{K}} := a$ for all $a \in N_I$.
2. Define $A^{\mathcal{I}_\mathcal{K}} := \{a \mid \mathcal{K} \models A(a)\}$ for all $A \in N_C$ and $r^{\mathcal{I}_\mathcal{K}} := \{(a, b) \mid \mathcal{K} \models r(a, b)\}$ for all $r \in N_R$.
3. Repeat:
   a) Select an element $d \in \Delta^{\mathcal{I}_\mathcal{K}}$ that has not been selected before and let $V := \{\exists r.B \mid d \in A^{\mathcal{I}_\mathcal{K}} \text{ and } d \notin (\exists r.B)^{\mathcal{I}_\mathcal{K}} \text{ with } A \sqsubseteq_\mathcal{T} \exists r.B,\ A, B \in N_C\}$.
   b) For each $\exists r.B \in V$ that is minimal w.r.t. $\sqsubseteq^s_\mathcal{T}$, add a fresh element e to $\Delta^{\mathcal{I}_\mathcal{K}}$, for each $B \sqsubseteq_\mathcal{T} A$ add e to $A^{\mathcal{I}_\mathcal{K}}$, and for each $r \sqsubseteq_\mathcal{T} s$ add (d, e) to $s^{\mathcal{I}_\mathcal{K}}$.

By $\mathcal{I}_\mathcal{A}$ we denote the restriction of $\mathcal{I}_\mathcal{K}$ to named individuals, i.e. the result of applying only Steps 1 and 2, but not Step 3. $\diamondsuit$

If Step 3 is applied fairly, i.e. such that each new domain element that is created in (b) is eventually also selected in (a), then $\mathcal{I}_\mathcal{K}$ is indeed a model of $\mathcal{K}$ (if $\mathcal{K}$ is consistent at all).

Lemma 3.7. Let $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ be a consistent and normalized $\mathcal{ELH}_\bot$-KB. Then, assuming fairness of rule application, the interpretation constructed according to Definition 3.6 is a model of $\mathcal{K}$.

Proof. We show that all axioms of the KB are satisfied in $\mathcal{I}_\mathcal{K}$ constructed according to Definition 3.6. It is easy to check that all ABox assertions are satisfied after Step 2 is applied. We make a case distinction for the TBox axioms:

- Suppose a GCI of the form $A_1 \sqcap \dots \sqcap A_n \sqsubseteq B \in \mathcal{T}$ with $n \geq 1$ and $d \in A_i^{\mathcal{I}_\mathcal{K}}$ for $1 \leq i \leq n$. If $d \in N_I$, then it is easy to check that $d \in B^{\mathcal{I}_\mathcal{K}}$ after Step 2. If $d \notin N_I$, then d was added because of some minimal $\exists r.B'$ in Step 3b). In this case we must have $B' \sqsubseteq_\mathcal{T} A_i$ for all $1 \leq i \leq n$ and hence also $B' \sqsubseteq_\mathcal{T} B$, which caused d to be added to $B^{\mathcal{I}_\mathcal{K}}$ in Step 3b).
- The case for RIAs of the form $r \sqsubseteq s \in \mathcal{T}$ follows the same argumentation.
- Suppose a GCI of the form $\exists r.A \sqsubseteq B \in \mathcal{T}$ and $d \in (\exists r.A)^{\mathcal{I}_\mathcal{K}}$. Then there exists $e \in \Delta^{\mathcal{I}_\mathcal{K}}$ with $e \in A^{\mathcal{I}_\mathcal{K}}$ and $(d, e) \in r^{\mathcal{I}_\mathcal{K}}$. If $e \in N_I$, then d must also be in N_I, and hence d was added to $B^{\mathcal{I}_\mathcal{K}}$ in Step 2. If $e \notin N_I$, then it was added in Step 3b) by some $B_1 \sqsubseteq_\mathcal{T} \exists r_1.A_1$ with $d \in B_1^{\mathcal{I}_\mathcal{K}}$ and $\exists r_1.A_1 \sqsubseteq^s_\mathcal{T} \exists r.A$. This implies $B_1 \sqsubseteq_\mathcal{T} B$ and hence d was already added to B, either in Step 2 if $d \in N_I$, or otherwise in Step 3b) when d was introduced.
- Suppose a GCI of the form $A \sqsubseteq \exists r.B \in \mathcal{T}$ and $d \in A^{\mathcal{I}_\mathcal{K}}$. At some point in the construction of $\mathcal{I}_\mathcal{K}$ d was picked in Step 3a). Let V be defined as in this step. Then V contains some minimal $\exists r_1.B_1$ with $\exists r_1.B_1 \sqsubseteq^s_\mathcal{T} \exists r.B$. Let d' be the individual introduced to satisfy $\exists r_1.B_1$. Then $(d, d') \in r_1^{\mathcal{I}_\mathcal{K}}$ and $d' \in B_1^{\mathcal{I}_\mathcal{K}}$ and because $r_1 \sqsubseteq_\mathcal{T} r$ and $B_1 \sqsubseteq_\mathcal{T} B$, it also holds that $(d, d') \in r^{\mathcal{I}_\mathcal{K}}$ and $d' \in B^{\mathcal{I}_\mathcal{K}}$. $\square$

Moreover, $\mathcal{I}_\mathcal{K}$ is universal, and therefore it holds for all CQs q that $\mathrm{cert}(q, \mathcal{K}) = \mathrm{ans}(q, \mathcal{I}_\mathcal{K})$.

Lemma 3.8. Let $\mathcal{K}$ be a consistent $\mathcal{ELH}_\bot$-KB and $\mathcal{I}_\mathcal{K}$ the canonical model obtained through the construction in Definition 3.6. Then $\mathcal{I}_\mathcal{K}$ is a minimal universal model of $\mathcal{K}$.

Proof. We first prove that $\mathcal{I}_\mathcal{K}$ is a universal model. Let $\mathcal{I}_0, \mathcal{I}_1, \dots$ be the interpretations obtained in the construction of $\mathcal{I}_\mathcal{K}$ before each application of Step 3, and let $\mathcal{I}$ be an arbitrary model of $\mathcal{K}$. We show by induction on i that there are $h_0, h_1, \dots$ such that h_i is a homomorphism from $\mathcal{I}_i$ to $\mathcal{I}$ and h_i and h_{i+1} agree on $\Delta^{\mathcal{I}_i}$, that is $h_{i+1}(d) = h_i(d)$ for all $d \in \Delta^{\mathcal{I}_i}$. The desired homomorphism is then obtained in the limit as $h = \bigcup_{i \geq 0} h_i$.

The homomorphism h_0 is defined by setting $h_0(a) := a^\mathcal{I}$ for all $a \in \mathrm{Ind}(\mathcal{A})$. Since $\mathcal{I}$ is a model of $\mathcal{A}$ it is easy to check that all conditions for a homomorphism are satisfied.

For the induction step, assume that h_i has already been defined. To define h_{i+1}, assume that $d \in \Delta^{\mathcal{I}_i}$ was picked in Step 3(a) and V is the set as defined in Definition 3.6. For each $\exists r.B \in V$ that is minimal w.r.t. structural subsumption, let A be a concept name that caused $\exists r.B$ to be in V, i.e. $A \sqsubseteq_\mathcal{T} \exists r.B$, and let e be the freshly introduced domain element. Then we know that $e \in (A')^{\mathcal{I}_{i+1}}$ for all $B \sqsubseteq_\mathcal{T} A'$ and $(d, e) \in s^{\mathcal{I}_{i+1}}$ for all $r \sqsubseteq_\mathcal{T} s$. Because h_i is a homomorphism, we must have $h_i(d) \in A^\mathcal{I}$ and since $\mathcal{I}$ is a model it is possible to find some $e' \in \Delta^\mathcal{I}$ with $(h_i(d), e') \in r^\mathcal{I}$ and $e' \in B^\mathcal{I}$. Clearly, $h_{i+1} := h_i \cup \{e \mapsto e'\}$ is a homomorphism from $\mathcal{I}_{i+1}$ to $\mathcal{I}$.

Next, we prove that $\mathcal{I}_\mathcal{K}$ is minimal. By definition this means that every endomorphism of $\mathcal{I}_\mathcal{K}$ has to be an isomorphism. To show this we prove the stronger statement that the only endomorphism on $\mathcal{I}_\mathcal{K}$ is the identify. We use a similar setup as before: Let $\mathcal{I}_0, \mathcal{I}_1, \dots$ be the interpretations obtained in the construction of $\mathcal{I}_\mathcal{K}$ before each application of Step 3. We show by induction on i that there are $h_0, h_1, \dots$ such that h_i is the only possible homomorphism from $\mathcal{I}_i$ to $\mathcal{I}_\mathcal{K}$ and h_i and h_{i+1} agree on $\Delta^{\mathcal{I}_i}$, that is $h_{i+1}(d) = h_i(d)$ for all $d \in \Delta^{\mathcal{I}_i}$. The endomorphism is then obtained in the limit as $h = \bigcup_{i\geq 0} h_i$.

By definition, the homomorphism h_0 has to map all constants onto itself, i.e. $h_0(a) = a$ for all $a \in \mathrm{Ind}(\mathcal{A})$ and is therefore the identity function. For the induction step, assume that h_i has already been defined. To define h_{i+1}, assume that $d \in \Delta^{\mathcal{I}_i}$ was picked in Step 3(a) and V is the set as defined in Definition 3.6. For each $\exists r.B \in V$ that is minimal w.r.t. structured subsumption, in Step 3(b) exactly one successor e is introduced. Suppose there would be another successor e' of d, introduced through some minimal $\exists r'.B'$ and e could be mapped to e'. This would imply that $\exists r'.B' \sqsubseteq^s_\mathcal{T} \exists r.B$, which is a contradiction, since we assumed $\exists r.B$ to be minimal. Hence, such an e' cannot exist and therefore the only possibility is to map e onto itself.

Therefore, the only possibility is to define $h_{i+1} := h_1 \cup \{e \mapsto e\}$, which is the identity function. □

We can adopt a result from graph theory and cores, namely Theorem 11 in [Bau95], which shows that if two structures are homomorphically equivalent then their minimal structures are isomorphic. Since all universal models are homomorphically equivalent by definition, this implies that every consistent $\mathcal{ELH}_\bot$-KB has a unique minimal universal model (up to isomorphism), which is why the canonical model is also 'the' minimal universal model of $\mathcal{K}$.

We now define the semantics of NCQs as described before, i.e. by evaluating them as first-order formulas over the canonical model $\mathcal{I}_\mathcal{K}$, which ensures that our semantics is compatible with the usual certain-answer semantics for CQs.

Definition 3.9 (Minimal-World Semantics). Let $\mathcal{K}$ be a consistent $\mathcal{ELH}_\bot$-KB. The *(minimal-world) answers* to an NCQ q over $\mathcal{K}$ are $\mathrm{mwa}(q, \mathcal{K}) := \mathrm{ans}(q, \mathcal{I}_\mathcal{K})$. ♢

For Example 3.1, we get $\mathrm{mwa}(q_B, \mathcal{K}_C) = \{p_1, p_2\}$ (see Figure 3.2), which is exactly as intended. Unfortunately, in general the canonical model is infinite, and we cannot evaluate the answers directly. Hence, we employ a rewriting approach to reduce NCQ answering over the canonical model to (first-order) query answering over $\mathcal{I}_\mathcal{A}$ only.

3.4. A Combined Rewriting for NCQs

We show that NCQ answering is combined first-order rewritable. As target representation, we obtain first-order queries of a special form.

Definition 3.10 (Filtered query). Let $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ be an $\mathcal{ELH}_{\bot}$-KB. A *filter* on a variable z is a first-order formula $\psi(z)$ of the form

$$(\exists z'.\psi^{+}(z, z')) \to (\exists z'.\psi^{+}(z, z') \wedge \psi^{-}(z, z') \wedge \Psi) \tag{3.1}$$

where $\psi^{+}(z, z')$ is a conjunction of atoms of the form $A(z')$ or $r(z, z')$ that contains at least one role atom, $\psi^{-}(z, z')$ is a conjunction of negated atoms $\neg A(z')$ or $\neg r(z, z')$, and Ψ is a (possibly empty) set of filters on z'.

A *filtered query* ϕ is of the form $\exists \mathbf{y}.(\varphi(\mathbf{x}, \mathbf{y}) \wedge \Psi)$ where $\exists \mathbf{y}.\varphi(\mathbf{x}, \mathbf{y})$ is an NCQ and Ψ is a (possibly empty) set of filters on leaf variables in φ. It is *rooted* if $\exists \mathbf{y}.\varphi(\mathbf{x}, \mathbf{y})$ is rooted. ♢

Note that every NCQ is a filtered query where the set of filters Ψ is empty. Furthermore, a filtered query can contain filters on answer variables as well as quantified variables.

We will use filters to check for the existence of 'typical' successors, i.e. role successors that behave like the ones that are introduced by the canonical model construction to satisfy an existential restriction. In particular, a typical successor does not satisfy any superfluous concept or role atoms. For example, in Figure 3.2 the element c_1 introduced to satisfy $\exists\mathtt{diagnosedWith}.\mathtt{BreastCancer}$ for p_1 is a typical successor, because it satisfies only BreastCancer and Cancer and not, e.g. SkinCancer. In contrast, the diagnosedWith-successor c_3 of p_3 is atypical, since the ontology does not contain an existential restriction $\exists\mathtt{diagnosedWith}.\mathtt{SkinOfBreastCancer}$ that could have introduced such a successor in the canonical model. With a filter we can check if a given patients diagnoses are typical cases or not. For instance the filter ψ

$$\begin{aligned}&(\exists y.\,\mathtt{diagnosedWith}(x, y) \wedge \mathtt{BreastCancer}(y)) \to \\ &\quad(\exists y.\,\mathtt{diagnosedWith}(x, y) \wedge \mathtt{BreastCancer}(y) \wedge \neg\mathtt{SkinOfBreastCancer}(y))\end{aligned}$$

is satisfied for all typical cases, i.e. p_1 and p_2, but not for p_3, which only has an atypical cancer diagnosis.

The idea of the rewriting procedure is to not only rewrite the positive part of the query, as in [EOŠ+12; BO15], but to also ensure that no information is lost. This is accomplished by rewriting the negative parts and by saving the structure of the eliminated part of the query in the filter. A filter on z ensures that the rewritten query can only be satisfied by mapping z to an anonymous individual in the canonical model, or to a named individual that behaves in a typical way.

Definition 3.11 (Rewriting). Let $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ be a KB and $\phi = \exists \mathbf{y}.\varphi(\mathbf{x}, \mathbf{y}) \wedge \Psi$ be a filtered query. We write $\phi \to_{\mathcal{T}} \phi'$ if ϕ' can be obtained from ϕ by applying the following steps:

(S1) Select a quantified leaf variable $\hat{x}$ in φ. Let $\hat{y}$ be a fresh variable and select

$$\begin{aligned}
\mathsf{Pred} &:= \{y \mid r(y,\hat{x}) \in \varphi\} \cup \{y \mid \neg r(y,\hat{x}) \in \varphi\} && \text{(predecessors of } \hat{x}\text{)},\\
\mathsf{Pos} &:= \{A(\hat{x}) \in \varphi\} \cup \{r(\hat{y},\hat{x}) \mid r(y,\hat{x}) \in \varphi\} && \text{(positive atoms for } \hat{x}\text{)},\\
\mathsf{Neg} &:= \{\neg A(\hat{x}) \in \varphi\} \cup \{\neg r(\hat{y},\hat{x}) \mid \neg r(y,\hat{x}) \in \varphi\} && \text{(negative atoms for } \hat{x}\text{)}.
\end{aligned}$$

(S2) Select some $M \sqsubseteq_{\mathcal{T}} \exists s.N$ with $M, N \in N_C$ that satisfies all of the following:
 (a) $s(\hat{y},\hat{x}) \wedge N(\hat{x}) \sqsubseteq^{s}_{\mathcal{T}} \mathsf{Pos}$, and
 (b) $s(\hat{y},\hat{x}) \wedge N(\hat{x}) \not\sqsubseteq^{s}_{\mathcal{T}} \alpha$ for all $\neg\alpha \in \mathsf{Neg}$.

(S3) Let $\mathcal{M}'$ be the set of all $M' \in N_C$ such that $M' \sqsubseteq_{\mathcal{T}} \exists s'.N'$ with $N' \in N_C$,
 (a) $\exists s'.N' \sqsubseteq^{s}_{\mathcal{T}} \exists s.N$ (where $\exists s.N$ was chosen in (S2)), and
 (b) $s'(\hat{y},\hat{x}) \wedge N'(\hat{x}) \sqsubseteq^{s}_{\mathcal{T}} \alpha$ for some $\neg\alpha \in \mathsf{Neg}$.

(S4) Drop from φ every atom that contains $\hat{x}$.

(S5) Replace all variables $y \in \mathsf{Pred}$ in φ with $\hat{y}$.

(S6) Add the atoms $M(\hat{y})$ and $\{\neg M'(\hat{y}) \mid M' \in \mathcal{M}'\}$ to φ.

(S7) Set the new filters to $\Psi' := \Psi \cup \{\psi^*(\hat{y})\} \setminus \Psi_{\hat{x}}$, where $\Psi_{\hat{x}} := \{\psi(\hat{x}) \in \Psi\}$ and

$$\psi^*(\hat{y}) := \big(\exists \hat{x}.\, s(\hat{y},\hat{x}) \wedge N(\hat{x})\big) \to \big(\exists \hat{x}.\, s(\hat{y},\hat{x}) \wedge N(\hat{x}) \wedge \mathsf{Neg} \wedge \Psi_{\hat{x}}\big).$$

We write $\phi \to^*_{\mathcal{T}} \phi'$ if there exists a finite sequence $\phi \to_{\mathcal{T}} \cdots \to_{\mathcal{T}} \phi'$. Furthermore, let $\mathrm{rew}_{\mathcal{T}}(\phi) := \{\phi' \mid \phi \to^*_{\mathcal{T}} \phi'\}$ denote the set of all rewritings of ϕ. ◇

Note that $\mathrm{rew}_{\mathcal{T}}(\phi)$ may be infinite. However, for rooted NCQs it is finite and we show later that even for non-rooted NCQs it suffices to consider a finite subset of $\mathrm{rew}_{\mathcal{T}}(\phi)$ (see Lemma 3.15). To see the former claim, observe that there is only a finite number of possible subsumptions $M \sqsubseteq_{\mathcal{T}} \exists s.N$ that can be used for rewriting steps. Additionally, in every step one variable ($\hat{x}$) is eliminated from the NCQ part of the filtered query. If the query is rooted, there always exists at least one predecessor that is renamed to $\hat{y}$, hence the introduction of $\hat{y}$ never increases the number of variables. Finally, it is easy to see that rewriting a rooted query always yields a rooted query.

The rewriting of Neg to the new negated atoms (via $\mathcal{M}'$ in (S6)) ensures that we do not lose important exclusion criteria, which may result in too many answers. Similarly, the filters exclude atypical successors in the ABox that may result in spurious answers. Both of these constructions are necessary.

Example 3.12. Consider the query q_B from Example 3.1. Using Definition 3.11, we obtain the first-order queries $\phi_B = q_B$, ϕ'_B, and ϕ''_B, where

$$\begin{aligned}
\phi'_B = {} & \exists y.\, \mathtt{diagnosedWith}(x,y) \wedge \mathtt{BreastCancer}(y) \wedge \neg\mathtt{SkinOfBreastCancer}(y) \wedge \\
& \Big(\big(\exists z.\, \mathtt{findingSite}(y,z) \wedge \mathtt{BreastStructure}(z)\big) \to \\
& \big(\exists z.\, \mathtt{findingSite}(y,z) \wedge \mathtt{BreastStructure}(z) \wedge \neg\mathtt{SkinStructure}(z)\big)\Big)
\end{aligned}$$

results from choosing z in (S1), $\texttt{BreastCancer} \sqsubseteq_{\mathcal{K}_C} \exists\texttt{findingSite}.\texttt{BreastStructure}$ in (S2), and computing $\mathcal{M}' = \{\texttt{SkinOfBreastCancer}\}$ in (S3), and

$$\begin{aligned}
\phi''_B = {} & \texttt{BreastCancerPatient}(x) \wedge \\
& \Big((\exists y.\, \texttt{diagnosedWith}(x,y) \wedge \texttt{BreastCancer}(y)) \to \\
& \quad (\exists y.\, \texttt{diagnosedWith}(x,y) \wedge \texttt{BreastCancer}(y) \wedge \neg\texttt{SkinOfBreastCancer}(y)) \wedge \\
& \quad\quad ((\exists z.\, \texttt{findingSite}(y,z) \wedge \texttt{BreastStructure}(z)) \to \\
& \quad\quad\quad (\exists z.\, \texttt{findingSite}(y,z) \wedge \texttt{BreastStructure}(z) \wedge \neg\texttt{SkinStructure}(z)))\Big)
\end{aligned}$$

is obtained due to $\texttt{BreastCancerPatient} \sqsubseteq_{\mathcal{K}_C} \exists\texttt{diagnosedWith}.\texttt{BreastCancer}$. We omitted the redundant atoms $\texttt{Cancer}(y)$ for clarity.

The finite interpretation $\mathcal{I}_{\mathcal{A}_C}$ can be seen in Figure 3.2 by ignoring all circle-shaped nodes. When computing the answers over $\mathcal{I}_{\mathcal{A}_C}$, we obtain

$$\operatorname{ans}(\phi_B, \mathcal{I}_{\mathcal{A}_C}) = \emptyset,\ \operatorname{ans}(\phi'_B, \mathcal{I}_{\mathcal{A}_C}) = \emptyset, \text{ and } \operatorname{ans}(\phi''_B, \mathcal{I}_{\mathcal{A}_C}) = \{p_1, p_2\}.$$

For ϕ'_B, the conjunct $\neg\texttt{SkinOfBreastCancer}(y)$ is necessary to exclude p_3 as an answer. In ϕ''_B, p_3 is excluded due to the filter that detects c_3 as an atypical successor, because it satisfies not only BreastCancer, but also SkinOfBreastCancer. Hence, both (S6) and (S7) are necessary steps in our rewriting. ♢

3.4.1. Correctness

In Definition 3.11, the new filter $\psi^*(\hat{y})$ may end up inside another filter expression after applying subsequent rewriting steps, i.e. by rewriting w.r.t. $\hat{y}$. In this case, however, the original structure of the rewriting is preserved, including all internal filters as well as the atoms $M(\hat{y})$, which are included implicitly by $\exists s.N \sqsubseteq M$, and $\{\neg M'(\hat{y}) \mid M' \in \mathcal{M}'\}$, which are included in Neg. We exploit this behavior to show that, whenever a rewritten query is satisfied in the finite interpretation $\mathcal{I}_{\mathcal{A}}$, then it is also satisfied in $\mathcal{I}_{\mathcal{K}}$. This is the most interesting part of the correctness proof, because it differs from the known constructions for ordinary CQs, for which this step is trivial.

Lemma 3.13. Let $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ be a consistent $\mathcal{ELH}_\bot$-KB and ϕ be an NCQ. Then, for all $\phi' \in \operatorname{rew}_{\mathcal{T}}(\phi)$,

$$\operatorname{ans}(\phi', \mathcal{I}_{\mathcal{A}}) \subseteq \operatorname{mwa}(\phi', \mathcal{K}).$$

Proof. Let $\phi' = \exists \mathbf{y}.(\varphi(\mathbf{x},\mathbf{y}) \wedge \Psi)$ and π be an assignment of $\mathbf{x},\mathbf{y}$ to N_I such that $\mathcal{I}_{\mathcal{A}}, \pi \models \varphi(\mathbf{x},\mathbf{y})$. Since $\mathcal{I}_{\mathcal{A}}$ and $\mathcal{I}_{\mathcal{K}}$ coincide on the domain N_I, we also have $\mathcal{I}_{\mathcal{K}}, \pi \models \varphi(\mathbf{x},\mathbf{y})$. Consider any filter $\psi(z) = \exists z'.\psi^+(z,z') \to \exists z'.(\beta(z,z') \wedge \Psi^*)$ in Ψ, where $\beta(z,z') := \psi^+(z,z') \wedge \psi^-(z,z')$. Then $\psi(z)$ was introduced at some point during the rewriting, suppose by selecting $M \sqsubseteq_{\mathcal{T}} \exists s.N$ in (S2). This means that φ contains the atom $M(z)$, and hence $d := \pi(z)$ is a named individual that is contained in $M^{\mathcal{I}_{\mathcal{A}}} \subseteq M^{\mathcal{I}_{\mathcal{K}}}$. By (S2), this means that $\mathcal{I}_{\mathcal{K}}, \pi \models \exists z'.\psi^+(z,z')$, and we have to show that $\mathcal{I}_{\mathcal{K}}, \pi \models \exists z'.(\beta(z,z') \wedge \Psi^*)$:

1. If $\mathcal{I}_{\mathcal{A}}, \pi \models \exists z'.\beta(z,z')$, then $\mathcal{I}_{\mathcal{K}}, \pi \models \exists z'.\beta(z,z')$ by the same argument as for $\varphi(\mathbf{x},\mathbf{y})$ above, and we can proceed by induction on the structure of the filters to show

that the inner filters Ψ^* are satisfied by the assignment π (extended appropriately for z').

2. If $\mathcal{I}_\mathcal{A}, \pi \not\models \exists z'.\beta(z, z')$, then we cannot use a named individual to satisfy the filter $\psi(z)$ in $\mathcal{I}_\mathcal{K}$. Moreover, since $\mathcal{I}_\mathcal{A}$ satisfies $\psi(z)$, we also know that $\mathcal{I}_\mathcal{A}, \pi \not\models \exists z'.\psi^+(z, z')$. Since $\psi^+(z, z') = s(z, z') \wedge N(z')$, this implies that $d \notin (\exists s.N)^{\mathcal{I}_\mathcal{A}}$. Hence, $\exists s.N$ is included in the set V constructed in Step 3(a) of the canonical model construction for the element $d = \pi(z)$. Thus, there exists $M' \sqsubseteq_\mathcal{T} \exists s'.N'$ such that $d \in (M')^{\mathcal{I}_\mathcal{A}}$, $d \notin (\exists s'.N')^{\mathcal{I}_\mathcal{A}}$, and $\exists s'.N' \sqsubseteq_\mathcal{T}^s \exists s.N$. By Step 3(b), $\mathcal{I}_\mathcal{K}$ must contain an element d' such that, for any $A \in N_C$ and any $r \in N_R$, we have $d' \in A^{\mathcal{I}_\mathcal{K}}$ iff $N' \sqsubseteq_\mathcal{T} A$ and $(d, d') \in r^{\mathcal{I}_\mathcal{K}}$ iff $s' \sqsubseteq_\mathcal{T} r$. Since $N' \sqsubseteq_\mathcal{T} N$ and $s' \sqsubseteq_\mathcal{T} s$, we obtain that $\mathcal{I}_\mathcal{K}, \pi \cup \{z' \mapsto d'\} \models \psi^+(z, z')$.

   We show that the assignment $\pi \cup \{z' \mapsto d'\}$ also satisfies $\psi^-(z, z') = \mathsf{Neg}$. Assume to the contrary that there is $\neg A(z') \in \mathsf{Neg}$ such that $d' \in A^{\mathcal{I}_\mathcal{K}}$ (the case of negated role atoms is again analogous). Then we have $N' \sqsubseteq_\mathcal{T} A$, which shows that all conditions of (S3) are satisfied, and hence M' must be included in $\mathcal{M}'$. Since the atoms $\{\neg M'(z) \mid M' \in \mathcal{M}'\}$ are contained in φ, we know that they are satisfied by π in $\mathcal{I}_\mathcal{K}$, i.e. $d \notin (M')^{\mathcal{I}_\mathcal{K}}$ and hence also $d \notin (M')^{\mathcal{I}_\mathcal{A}}$, which is a contradiction.

   It remains to show that the inner filters Ψ^* are satisfied by the assignment $\pi \cup \{z' \mapsto d'\}$ in $\mathcal{I}_\mathcal{K}$. Since we are now dealing with an anonymous domain element d', we can use similar, but simpler, arguments as above to prove this by induction on the structure of the filters. This is possible because the atoms $s(\hat{y}, \hat{x})$, $N(\hat{x})$ implied by $M(\hat{y})$ and the negated atoms induced by $\mathcal{M}'$ are present in the query even if the filter is integrated into another filter during a subsequent rewriting step. □

We can use this lemma to show correctness of our approach, i.e. the answers returned for the *union* of queries given by $\text{rew}_\mathcal{T}(\phi)$ over $\mathcal{I}_\mathcal{A}$ are exactly the answers of the original NCQ ϕ over $\mathcal{I}_\mathcal{K}$.

The proof is based on existing proofs for ordinary CQs [EOŠ+12; BO15], extended appropriately to deal with the filters.

Lemma 3.14. Let $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ be a consistent $\mathcal{ELH}_\bot$ KB and let $\phi(\mathbf{x})$ be an NCQ. Then, for all $\phi' \in \text{rew}_\mathcal{T}(\phi)$,

$$\text{mwa}(\phi, \mathcal{K}) = \bigcup_{\phi' \in \text{rew}_\mathcal{T}(\phi)} \text{ans}(\phi', \mathcal{I}_\mathcal{A}).$$

Proof. ($\supseteq$): By Lemma 3.13, we have $\text{ans}(\phi', \mathcal{I}_\mathcal{A}) \subseteq \text{mwa}(\phi', \mathcal{K}) = \text{ans}(\phi', \mathcal{I}_\mathcal{K})$.

Furthermore, there exists a sequence $\phi_0 \to_\mathcal{T} \cdots \to_\mathcal{T} \phi_n$ $(n > 0)$ with $\phi = \phi_0$ and $\phi' = \phi_n$. Hence it is sufficient to show that $\text{ans}(\phi_i, \mathcal{I}_\mathcal{K}) \subseteq \text{ans}(\phi_{i-1}, \mathcal{I}_\mathcal{K})$ for all $i, 1 \leq i \leq n$. Suppose the queries are of the following forms:

$$\phi_i = \exists \mathbf{y}_i.(\varphi_i(\mathbf{x}_i, \mathbf{y}_i) \wedge \Psi_i) \tag{3.2}$$

$$\phi_{i-1} = \exists \mathbf{y}_{i-1}.(\varphi_{i-1}(\mathbf{x}_{i-1}, \mathbf{y}_{i-1}) \wedge \Psi_{i-1}) \tag{3.3}$$

Let π_i be a satisfying assignment for $\varphi_i(\mathbf{x}_i, \mathbf{y}_i) \wedge \Psi_i$ in $\mathcal{I}_\mathcal{K}$. Suppose $\phi_{i-1} \to_\mathcal{T} \phi_i$ by

1. selecting variable $\hat{x}$ and introducing $\hat{y}$ in (S1) and

2. selecting $M \sqsubseteq_{\mathcal{T}} \exists s.N$ in (S2).

Let $\pi_i(\hat{y}) = d$. By Step (S6), $M(\hat{y}) \in \varphi_i$ and since π_i satisfies φ_i, it has to hold that $d \in M^{\mathcal{I}_\mathcal{K}}$. This implies that $d \in (\exists s.N)^{\mathcal{I}_\mathcal{K}}$. Since π_i satisfies the new filter $\psi_i^*(\hat{y})$ that is constructed in (S7), and by selecting $M \sqsubseteq_{\mathcal{T}} \exists s.N$ in (S2) the precondition of $\psi_i^*(\hat{y})$ is satisfied by π_i in $\mathcal{I}_\mathcal{K}$, there has to be an assignment $\pi_i \cup \{\hat{x} \mapsto d'\}$ that satisfies the conclusion of $\psi_i^*(\hat{y})$.

We define the assignment π_{i-1} of the variables of φ_{i-1} as follows

$$\pi_{i-1}(z) := \begin{cases} d' & \text{if } z = \hat{x} \\ d & \text{if } z \in \mathsf{Pred} \\ \pi_i(z) & \text{otherwise.} \end{cases} \tag{3.4}$$

Then π_{i-1} is a satisfying assignment for ϕ_{i-1} in $\mathcal{I}_\mathcal{K}$. To see this, first consider an atom α in φ_{i-1}. We show that π_{i-1} satisfies α in $\mathcal{I}_\mathcal{K}$.

If α contains $\hat{x}$, it can be of the following forms: $A(\hat{x})$, $\neg A(\hat{x})$, $r(y, \hat{x})$ or $\neg r(y, \hat{x})$ with $y \in \mathsf{Pred}$. For all of these cases, we know by Step (S7) that they are either implied by $s(\hat{y}, \hat{x}) \wedge N(\hat{x})$ or contained in Neg, with y replaced by $\hat{y}$. By the choice of d', we know that π_{i-1} satisfies each such atom.

If α does not contain $\hat{x}$, then φ_i contains the atom α' that is obtained from α by replacing all of the variables from Pred with $\hat{y}$. By construction, we know that $\pi_{i-1}(y) = \pi_i(\hat{y})$ for all $y \in \mathsf{Pred}$ and $\pi_{i-1}(z) = \pi_i(z)$ otherwise. Since α' is satisfied by π_i in $\mathcal{I}_\mathcal{K}$, α is satisfied by π_{i-1} in $\mathcal{I}_\mathcal{K}$.

What remains to show is that π_{i-1} satisfies Ψ_{i-1}. Consider any $\psi(z) \in \Psi_{i-1}$, and distinguish the following cases:

1. If $z = \hat{x}$, then $\psi(\hat{x}) \in \Psi_{\hat{x}}$. Since $\mathcal{I}_\mathcal{K}, \pi_i \cup \{\hat{x} \mapsto d'\} \models \Psi_{\hat{x}}$, we have $\mathcal{I}_\mathcal{K}, \{\hat{x} \mapsto d'\} \models \psi(\hat{x})$. Therefore, since $\pi_{i-1}(\hat{x}) = d'$, it holds that π_{i-1} satisfies $\psi(\hat{x})$ in $\mathcal{I}_\mathcal{K}$.
2. If $z \in \mathsf{Pred}$ we know that $\pi_{i-1}(z) = \pi_i(\hat{y}) = d$. Since $\mathcal{I}_\mathcal{K}, \pi_i \models \psi(\hat{y})$, it also holds that $\mathcal{I}_\mathcal{K}, \pi_{i-1} \models \psi(z)$.
3. Otherwise the filter is present in Ψ_i. In this case we know that $\mathcal{I}_\mathcal{K}, \pi_i \models \psi(z)$ and $\pi_i(z) = \pi_{i-1}(z)$. Hence, it must also hold that $\mathcal{I}_\mathcal{K}, \pi_{i-1} \models \psi(z)$.

($\subseteq$): Suppose that $\mathbf{a} \in \mathrm{mwa}(\phi, \mathcal{K}) = \mathrm{ans}(\phi, \mathcal{I}_\mathcal{K})$. We have to show that there exists a rewriting $\phi' \in \mathrm{rew}_{\mathcal{T}}(\phi)$ and a satisfying assignment π for ϕ' in $\mathcal{I}_\mathcal{A}$ such that $\mathbf{a} = \pi(\mathbf{x})$. To do this, we assign a *degree* (a natural number) to each satisfying assignment (including the existentially quantified variables of the NCQ part) such that a satisfying assignment with degree 0 does not use any anonymous individuals. We then show that for each satisfying assignment with a degree greater than 0, we can find a rewriting for which a satisfying assignment yielding the same answer, but with a lower degree, exists. In addition, for every such assignment π and for all filters $\psi(y)$ in ϕ' it should hold that,

$$\text{if } \pi(y) \in N_I, \text{ then } \mathcal{I}_\mathcal{A} \models \psi(\pi(y)), \tag{†}$$

i.e. all filters (at any stage of the rewriting) are satisfied within the confines of $\mathcal{I}_\mathcal{A}$.

For any element $d \in \Delta^{\mathcal{I}_\mathcal{K}}$, we denote by $|d|$ the minimal number of role connections required to reach d from an element in N_I, with $|d| = 0$ iff $d \in N_I$. Additionally, for any assignment π' in $\mathcal{I}_\mathcal{K}$, let

$$\deg(\pi') := \sum_{y \in \mathrm{dom}(\pi')} |\pi'(y)|. \tag{3.5}$$

Since $\phi \in \mathrm{rew}_\mathcal{T}(\phi)$, to prove the claim it suffices to show that whenever there is a filtered query $\phi_1 = \exists \mathbf{y}.\varphi_1(\mathbf{x}, \mathbf{y}) \wedge \Psi \in \mathrm{rew}_\mathcal{T}(\phi)$ such that φ_1 has a match π_1 in $\mathcal{I}_\mathcal{K}$ with $\mathbf{a} = \pi_1(\mathbf{x})$, $\deg(\pi_1) > 0$, and Equation (†) holds for π_1 and the filters in Ψ, then there exist ϕ_2 and π_2 with the same properties, but $\deg(\pi_2) < \deg(\pi_1)$.

Assume $\phi_1 \in \mathrm{rew}_\mathcal{T}(\phi)$ as above, and let π_1 be a match of φ_1. Since $\deg(\pi_1) > 0$ by assumption, there must exist a variable $\hat{x}$ of φ_1 such that $\pi_1(\hat{x}) \notin N_I$. Select $\hat{x}$ such that it is a leaf node in the sub-forest of $\mathcal{I}_\mathcal{K}$ induced by π_1. Note that $\hat{x}$ cannot be an answer variable.

We know that $\pi_1(\hat{x}) = d_{\hat{x}}$ was induced by some axiom $\alpha = M \sqsubseteq_\mathcal{T} \exists s.N$ and element $d_p \in M^{\mathcal{I}_\mathcal{K}}$ in Definition 3.6. By the construction of $\mathcal{I}_\mathcal{K}$, we know that

(i) $d_{\hat{x}}$ has just the one predecessor d_p, and

(ii) $d_{\hat{x}} \in A^{\mathcal{I}_\mathcal{K}}$ iff $N \sqsubseteq_\mathcal{T} A$ and $(d_p, d_{\hat{x}}) \in r^{\mathcal{I}_\mathcal{K}}$ iff $s \sqsubseteq_\mathcal{T} r$.

We obtain the query ϕ_2 from ϕ_1 through rewriting, by selecting $\hat{x}$ and introducing $\hat{y}$ in (S1), and selecting α in (S2). Let Pred denote the set of predecessor variables of $\hat{x}$ as defined in (S1). To see that this is a valid choice, the conditions in (S2) need to be verified:

(S2a) For any $A(\hat{x}) \in \varphi_1$, we have $d_{\hat{x}} = \pi_1(\hat{x}) \in A^{\mathcal{I}_\mathcal{K}}$, and hence $N \sqsubseteq_\mathcal{T} A$ by (ii). Consider any role atom $r(y, \hat{x}) \in \varphi_1$. From (i), the construction of $\mathcal{I}_\mathcal{K}$ (no inverse edges), and the fact that π_1 is a satisfying assignment for $r(y, \hat{x})$ in $\mathcal{I}_\mathcal{K}$, the only possibility is that $\pi_1(y) = d_p$. Therefore $(d_p, d_{\hat{x}}) = (\pi_1(y), \pi_1(\hat{x})) \in r^{\mathcal{I}_\mathcal{K}}$. By (ii), this implies that $s \sqsubseteq_\mathcal{T} r$.

(S2b) Consider any $\neg A(\hat{x}) \in \varphi_1$, for which we must have $d_{\hat{x}} \notin A^{\mathcal{I}_\mathcal{K}}$. From (ii) we know that $N \not\sqsubseteq_\mathcal{T} A$. Consider any $\neg r(y, \hat{x}) \in \varphi_1$. Since this is guarded by a positive role atom as above, again the only possibility is that $\pi_1(y) = d_p$. Hence $(d_p, d_{\hat{x}}) \notin r^{\mathcal{I}_\mathcal{K}}$. By (ii), this implies that $s \not\sqsubseteq_\mathcal{T} r$.

Therefore, we obtain a satisfying assignment π_2 for ϕ_2 in $\mathcal{I}_\mathcal{K}$ such that $\mathbf{a} \in \pi_2(\mathbf{x})$ (and $\deg(\pi_2) < \deg(\pi_1)$) by setting for all $z \in \mathrm{Var}(\varphi_2)$:

$$\pi_2(z) := \begin{cases} \pi_1(z) & \text{if } z \in \mathrm{Var}(\varphi_1) \\ d_p & \text{if } z = \hat{y}. \end{cases}$$

To see that π_2 satisfies ϕ_2, we argue why it satisfies the new atoms and filter from (S6) and (S7); the old atoms (possibly with renamed variables) remain satisfied.

The new atom $M(\hat{y})$ is satisfied since $\pi_2(\hat{y}) = d_p \in M^{\mathcal{I}_\mathcal{K}}$. Consider now an atom $\neg M'(\hat{y})$ with $M' \in \mathcal{M}'$ as specified in (S6); we have to show that $d_p \notin (M')^{\mathcal{I}_\mathcal{K}}$. Assume to the contrary that $d_p \in (M')^{\mathcal{I}_\mathcal{K}}$. By (S3), we know that $M' \sqsubseteq_\mathcal{T} \exists s'.N' \sqsubseteq^s_\mathcal{T} \exists s.N$. Moreover, $\exists s'.N'$ must be included in the set V in Step 3(a) of Definition 3.6, because otherwise we would already have $d_p \in (\exists s'.N')^{\mathcal{I}_\mathcal{A}}$, i.e. there would be a named individual b such that $(d_p, b) \in (s')^{\mathcal{I}_\mathcal{A}}$ and $b \in (N')^{\mathcal{I}_\mathcal{A}}$. Since $s' \sqsubseteq_\mathcal{T} s$ and $N' \sqsubseteq_\mathcal{T} N$, this would imply $(d_p, b) \in s^{\mathcal{I}_\mathcal{A}}$ and $b \in N^{\mathcal{I}_\mathcal{A}}$, i.e. $d_p \in (\exists s.N)^{\mathcal{I}_\mathcal{A}}$, which shows that the anonymous object $d_{\hat{x}}$ would not have been created. Since $\exists s'.N'$ is included in V and we assumed that $\exists s.N$ is minimal w.r.t. $\sqsubseteq^s_\mathcal{T}$, we must have $s \equiv_\mathcal{T} s'$ and $N \equiv_\mathcal{T} N'$. But then (S3b) directly contradicts (S2b).

We now consider the filters in ϕ_2. Suppose that Equation (†) holds for π_1 and all filters in ϕ_1. For the ones that are only copied from ϕ_1 (modulo renaming some variables to $\hat{y}$), the property is clearly preserved. For the new filter $\psi^*(\hat{y})$, assume that $\pi_2(\hat{y}) \in N_I$, and hence we need to show that $\mathcal{I}_\mathcal{A} \models \pi_2(\psi^*(\hat{y}))$. Assume that there exists an element $d' \in N_I$ such that $(d_p, d') \in s^{\mathcal{I}_\mathcal{A}}$ and $d' \in N^{\mathcal{I}_\mathcal{A}}$. But then in Step 3(a) in Definition 3.6, $\exists s.N$ could not have been added to V since $d_p \in (\exists s.N)^{\mathcal{I}_\mathcal{K}}$ already holds. Hence, the element $d_{\hat{x}}$ would have never been introduced, which is a contradiction. Therefore, in $\mathcal{I}_\mathcal{A}$ the precondition of $\psi^*(\hat{y})$ is never met, which makes the filter trivially satisfied.

Finally, to show that $\deg(\pi_2) < \deg(\pi_1)$, we make a case distinction on whether the set Pred is empty or not. If $\mathsf{Pred} = \emptyset$, then we essentially replace the variable $\hat{x}$ in φ_1 with a new variable $\hat{y}$ in φ_2 with $|\pi_2(\hat{y})| = |d_p| < |d_{\hat{x}}| = |\pi_1(\hat{x})|$. Since the remaining variables are not affected by the rewriting step, this shows that $\deg(\pi_2) < \deg(\pi_1)$. If $\mathsf{Pred} \neq \emptyset$, then we have $|\pi_2(\hat{y})| = |d_p| = |\pi_1(y)|$ for all $y \in \mathsf{Pred}$. Since the variables in $\mathrm{Var}(\varphi_1) \setminus \{\hat{y}\} = \mathrm{Var}(\varphi_1) \setminus (\mathsf{Pred} \cup \{\hat{x}\})$ are not affected and $|\pi_1(\hat{x})| > 0$, we conclude that

$$\begin{aligned}\deg(\pi_2) &= |\pi_2(\hat{y})| + \sum_{z \in \mathrm{Var}(\varphi_2)\setminus\{\hat{y}\}} |\pi_2(z)| \\ &< |\pi_1(\hat{x})| + \sum_{y \in \mathsf{Pred}} |\pi_1(y)| + \sum_{z \in \mathrm{Var}(\varphi_1)\setminus(\mathsf{Pred} \cup \{\hat{x}\})} |\pi_1(z)| \\ &= \deg(\pi_1). \end{aligned}$$

□

Under data complexity assumptions, ϕ and $\mathcal{T}$, and hence $\mathrm{rew}_\mathcal{T}(\phi)$, are fixed, and $\mathcal{I}_\mathcal{A}$ is of polynomial size in the size of $\mathcal{A}$.

For non-rooted queries the rewriting can get infinite, because we can always find a leaf variable that is not an answer variable. For example, suppose that the TBox $\mathcal{T}$ consists of the two GCIs $A \sqsubseteq \exists r.B$ and $B \sqsubseteq \exists r.A$. Let $\phi() = \exists x.A(x) \wedge \neg B(x)$ be a Boolean query. The rewriting algorithm would produce an infinite rewriting of the form:

$$\begin{aligned}
&\phi_0 = \exists x.A(x) \wedge \neg B(x) \\
&\phi_1 = \exists x.B(x) \wedge \underbrace{(\exists y.r(x,y) \wedge A(y) \to \exists y.r(x,y) \wedge A(y) \wedge \neg B(y))}_{\psi_1(x)} \\
&\phi_2 = \exists x.A(x) \wedge \underbrace{(\exists y.r(x,y) \wedge B(y) \to \exists y.r(x,y) \wedge B(y) \wedge \psi_1(y))}_{\psi_2(x)} \\
&\phi_3 = \exists x.B(x) \wedge \underbrace{(\exists y.r(x,y) \wedge A(y) \to \exists y.r(x,y) \wedge A(y) \wedge \psi_2(y))}_{\psi_3(x)} \\
&\dots
\end{aligned}$$

In the following, we show that such infinite rewritings can be avoided, since after a certain number of rewriting steps queries will not generate any new answers.

For a query ϕ with filters Ψ, where each $\psi(y) \in \Psi$ is of the form

$$(\exists y'.\psi^+(y,y')) \to (\exists y'.\psi^+(y,y') \wedge \psi^-(y,y') \wedge \Psi'),$$

we define the *nested filter depth* as follows:

$$|\phi| := \max_{\psi\in\Psi} |\psi| \qquad\qquad |\psi| := 1 + \max_{\psi'\in\Psi'} |\psi'|,$$

where the second expression is applied recursively to sub-filters.

Lemma 3.15. Let $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ be a consistent $\mathcal{ELH}_\bot$ KB and let ϕ be an NCQ. Then

$$\bigcup_{\substack{\phi'\in\mathrm{rew}_\mathcal{T}(\phi)\\ |\phi'|\le v+N_{C\mathcal{T}}^2\cdot N_{R\mathcal{T}}}} \mathrm{ans}(\phi', \mathcal{I}_\mathcal{A}) = \bigcup_{\phi'\in\mathrm{rew}_\mathcal{T}(\phi)} \mathrm{ans}(\phi', \mathcal{I}_\mathcal{A}).$$

where v denotes the number of variables in ϕ, and $N_{C\mathcal{T}}$ and $N_{R\mathcal{T}}$ denote the number of concept and role names in $\mathcal{T}$, respectively.

Proof. In the following, we assume a connected CQ ϕ with v variables. This is without loss of generality since the non-connected parts in the query can be dealt with separately.

Connectedness is preserved by the rewriting algorithm and hence there are two possible scenarios. In the first one, ϕ has only a finite number of rewritings. That happens if after at most v rewriting steps the rewriting is of the form $\exists\mathbf{z}.\varphi(\mathbf{z}) \wedge \Psi(\mathbf{z})$, where $\varphi(\mathbf{z})$ cannot be rewritten further. Otherwise, ϕ has infinitely many rewritings. In this case, after v rewriting steps, the rewriting and all further rewritings are of the form

$$\exists y.(A(y) \wedge \mathsf{Neg} \wedge \psi(y)), \tag{‡}$$

where $A \in N_C$ (recall that we assume a TBox in normal form), Neg is a conjunction of atoms of the form $\neg\hat{A}(y)$ for $\hat{A} \in N_C$, and ψ is a filter that is linearly extended in every further rewriting step.

Assume a query ϕ_0 of the form (‡) that has been further rewritten to ϕ_n in a sequence of $\phi_0 \to_\mathcal{T} \cdots \to_\mathcal{T} \phi_i \to_\mathcal{T} \cdots \to_\mathcal{T} \phi_n$, where $n \in \mathbb{N}$, every formula ϕ_j is of the form $\exists y.A_j(y) \wedge \mathsf{Neg}_j \wedge \psi_j(y)$ for $0 \le j \le n$, and there exists $1 < i < n$ with $A_n = A_i$, $\mathsf{Neg}_n \cong \mathsf{Neg}_i$ and $\psi_n^+ \cong \psi_i^+$, i.e. the filters of ϕ_n and ϕ_i have the same body (up to renaming of variables, denoted by $\cong$).

We show that if there is a satisfying assignment π_n for ϕ_n over $\mathcal{I}_\mathcal{A}$, then there is also a satisfying assignment for a query ϕ' with $|\phi'| < |\phi_n|$. Suppose the nested filters in ϕ_n are matched up to a nested filter depth of $0 \le d \le n$.

Case 1: If $d = 0$, the body of ψ_n is not satisfied, then π_n is also a match for $A_i \wedge \mathsf{Neg}_i$, since the $A_i = A_n$, $\mathsf{Neg}_i = \mathsf{Neg}_n$ and the body of ψ_i is the same as the body of ψ_n by assumption. Clearly $|\phi_i| < |\phi_n|$.

Case 2: If $d > 0$, then we construct a match for query ϕ_{n-1}. Suppose ϕ_{n-1} has been rewritten to ϕ_n by using a GCI $A \sqsubseteq_{\mathcal{T}} \exists r.B$. Then we know that

$$\begin{aligned} \phi_{n-1} &:= \exists y.B(y) \land \mathsf{Neg}_{n-1} \land \psi(y) \\ \phi_n &:= \exists x.A(x) \land \mathsf{Neg}_n \land (\exists y.r(x,y) \land B(y) \to \exists y.r(x,y) \land B(y) \land \mathsf{Neg}_{n-1} \land \psi(y)) \end{aligned}$$

Because $d > 0$, there has to be a satisfying assignment $\{x \mapsto a, y \mapsto b\}$ with $a, b \in \Delta^{\mathcal{I}_{\mathcal{A}}}$ for $r(x,y) \land B(y)$, which also satisfies $\mathsf{Neg}_{n-1} \land \psi(y)$. Then the assignment $\pi_{n-1} := \{y \mapsto b\}$ satisfies ϕ_{n-1}, and $\psi(y)$ will then be matched up to a depth of $d - 1$.

This result can be used to bound the nested filter depth of queries during rewriting: In each rewriting step the query is rewritten w.r.t. a GCI of the form $A \sqsubseteq_{\mathcal{T}} \exists r.B$ with $A, B \in N_{C\mathcal{T}}$ and $r \in N_{R\mathcal{T}}$. There can be at most $N_{C\mathcal{T}}^2 \cdot N_{R\mathcal{T}}$ different GCIs of this form. Suppose a rewriting of ϕ to ϕ' with $|\phi'| > v + N_{C\mathcal{T}}^2 \cdot N_{R\mathcal{T}}$, where v denotes the number of variables in ϕ. Then at least two rewritings between ϕ and ϕ' must start with the same expression $\exists x.A(x) \land \mathsf{Neg} \land (\exists y.r(x,y) \land B(y) \to \dots)$. By the arguments above, there are rewritings of ϕ of nested filter depth at most $v + N_{C\mathcal{T}}^2 \cdot N_{R\mathcal{T}}$ that yield the same answers as ϕ'. □

Hence, queries need to be rewritten only until this bound on the nested filter depth, which does not depend on $\mathcal{A}$. We obtain the claimed complexity result.

Theorem 3.16. Checking whether a given tuple **a** is a minimal-world answer to an NCQ ϕ over a consistent $\mathcal{ELH}_\bot$-KB $\mathcal{K}$ can be done in polynomial time in data complexity.

As discussed in Example 3.1 it can be useful to allow *complex assertions* in the ABox: The assertions of patient p_3 can then be stated without introducing a new constant for the disease by stating

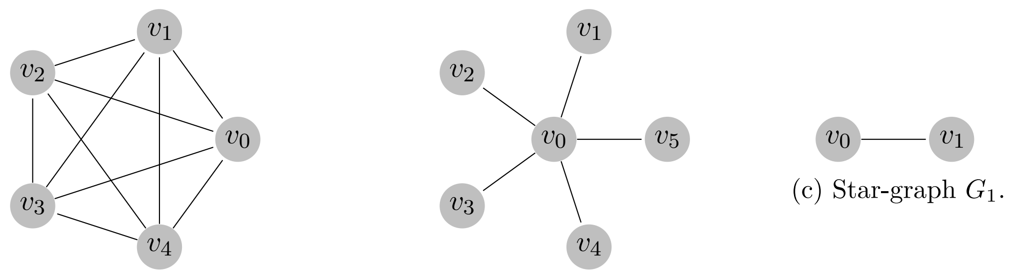
$$\exists \texttt{diagnosedWith}.\texttt{SkinOfBreastCancer}(p_3).$$

This leads to the introduction of additional acyclic definitions $\mathcal{T}'$, which are not fixed. The complexity nevertheless remains the same: Since $\mathcal{T}$ does not use the new concept names in $\mathcal{T}'$, we can apply the rewriting only w.r.t. $\mathcal{T}$, and extend $\mathcal{I}_{\mathcal{A}}$ by a polynomial number of new elements that result from applying Definition 3.6 only w.r.t. $\mathcal{T}'$.

What is more important than the complexity result is that this approach can be used to evaluate NCQs using standard database methods, e.g. using views to define the finite interpretation $\mathcal{I}_{\mathcal{A}}$ based on the input data given in $\mathcal{A}$, and SQL queries to evaluate the elements of $\mathrm{rew}_{\mathcal{T}}(\phi)$ over these views [KLT+11].

3.5. Minimal-World Semantics for more Expressive Horn Description Logics

We have seen that in consistent $\mathcal{ELH}_\bot$-KBs there always is a minimal universal model, which is unique up to isomorphism. In this section we want to extend minimal-world semantics to more expressive DLs.



(a) A complete graph with 5 nodes. (b) Star-graph G_5 with 5 nodes. (c) Star-graph G_1.

Figure 3.17.: Example graphs.

3.5.1. Cores

It turns out that the notion of minimality is closely related to the concept of *cores*, a concept first defined in finite graph theory [Fel82; HN92]. Intuitively, a minimal graph does not contain any redundant elements w.r.t. to homomorphisms. Formally, the core property has been characterized in different ways, all of which turn out to be equivalent for finite structures:

- A graph G is a core if every endomorphism of G is injective.
- A graph G is a core if every endomorphism of G is surjective.

Additionally, we say a sub-graph H of a graph G is a *core of* G if H is the image of an endomorphism of G. A simple example of a core is a complete graph $G = (V, E)$, where V denotes the set of vertices and E the set of edges. In a complete graph every pair of distinct vertices is connected with a unique edge. An example can be found in Figure 3.17a. In a complete graph it is not possible to map any two distinct vertices $v_1, v_2 \in V$ onto $v_3 \in V$ with some endomorphism h, because v_1 and v_2 are connected, formally $(v_1, v_2) \in E$, but v_3 is not connected to itself, formally $(h(v_1), h(v_2)) = (v_3, v_3) \notin E$. Therefore every complete graph is a core.

Now consider the a star-shaped graph $G_n = (V_n, E_n)$ where $V_n := \{v_i \mid 0 \leq i \leq n\}$ and $E = \{(v_0, v_i) \mid 0 < i \leq n\}$ for $n > 0$. Examples for $n = 5$ and $n = 1$ are shown in Figure 3.17. We can construct an endomorphism $h := \{v_0 \mapsto v_0\} \cup \{v_i \mapsto v_1 \mid 0 < i \leq n\}$ that is neither injective nor surjective. Therefore G_5 is not a core. The core of G_5 is obtained as the image of h, which yields the star-graph G_1. In fact, it can easily be checked that all star-graphs are homomorphically equivalent, which implies that G_1 is the unique core of all of them.

As we will see, the definitions are not equivalent anymore in infinite structures. To distinguish the different properties we adopt the notation of [Bau95]:

Definition 3.18. Let G be a relational structure.

- $\mathbf{I}(G)$ holds if any endomorphism of G is an injection.
- $\mathbf{S}(G)$ holds if any endomorphism of G is a surjection.
- $\mathbf{N}(G)$ holds if any endomorphism of G is a strong homomorphism.

$$\mathbf{ISN} \equiv \mathbf{SN} \equiv \mathbf{IN} \equiv \mathbf{IS} \equiv \mathbf{S} \equiv \mathbf{I} \rightarrow \mathbf{N}$$

(a) Implications in finite structures.

$$\begin{tikzcd}
 & & \mathbf{S} \\
 & \mathbf{IS} \arrow[ur] \arrow[dr] & \\
\mathbf{ISN} \equiv \mathbf{SN} \arrow[ur] \arrow[dr] & & \mathbf{I} \\
 & \mathbf{IN} \arrow[ur] \arrow[dr] & \\
 & & \mathbf{N}
\end{tikzcd}$$

(b) Implications in infinite structures.

Figure 3.19.: Implications that hold between **I**,**S**,**N** and their combinations thereof.

Let $\mathbf{X}$ be one of the above properties or any combination of them. We write $\mathbf{X}(G)$ to denote that all properties in $\mathbf{X}$ hold for G. G is called an $\mathbf{X}$-*core* if $\mathbf{X}(G)$ holds. A structure H is a $\mathbf{X}$-*core of* G if $H \subseteq G$, $G \rightarrow H$, and H is a $\mathbf{X}$-core. ♢

For finite relational structures, many properties (and their combinations) are equivalent as can be seen in Figure 3.19a. For example, if $\mathbf{I}(G)$ holds for some finite relational structure G, every endomorphism is an injection, which implies that the image of every endomorphism has to be G itself. Therefore, every endomorphism is also a surjection. Additionally, it holds for finite relational structures that (i) every finite structure has a core, which is (ii) unique up to isomorphism, and (iii) if two structures are homomorphically equivalent, their cores are isomorphic.

When moving to infinite structures, the situation changes and the different properties are not equivalent anymore, as can be seen in Figure 3.19b. For a more detailed introduction to core-like properties for infinite structures, we refer the reader to [Bau95].

In $\mathcal{ELH}_\bot$ we already have universal models that are not finite in general. We showed that for any consistent $\mathcal{ELH}_\bot$-KB there exists a unique universal model $\mathcal{I}_\mathcal{K}$ for which $\mathbf{ISN}(\mathcal{I}_\mathcal{K})$ holds (see proof of Lemma 3.8). Because of this property we were able to base minimal-world semantics on the minimal universal model.

3.5.2. Minimal universal models beyond $\mathcal{ELH}_\bot$

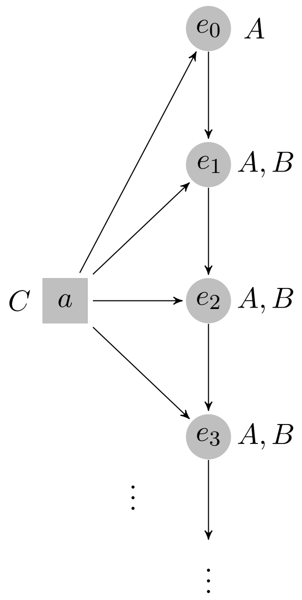
In the following we discuss problems that arise when moving to more expressive DLs. Since minimal-world semantics are strongly coupled to the existence of universal models we look into more expressive Horn-DLs. In particular, we discuss extensions of $\mathcal{EL}$ by nominals, inverses and transitivity. A summary can be found at the end of this section in Table 3.24. The following examples are inspired by the examples from [Bau95].

In the first example the universal model has to contain a ray of infinite length, which can be achieved with a combination of inverse roles and nominals:

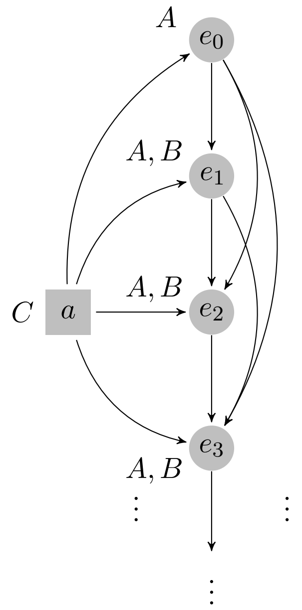
Example 3.21. Let the Horn-$\mathcal{ALCOI}$ KB $\mathcal{K}_1$ consist of the following axioms:

$$A \sqsubseteq \exists r.(A \sqcap B) \qquad B \sqsubseteq \exists r^-.\{a\} \qquad C \sqsubseteq \exists r.A$$
$$B \sqsubseteq \exists r^-.A \qquad C(a)$$
♢

In Figure 3.20 a universal model $\mathcal{I}_a$ of $\mathcal{K}_1$ is shown. Every universal model has to contain a ray with an initial element satisfying A, but not B. The constant a is connected to



(a) $\mathcal{I}_a$



(b) $\mathcal{I}_b$

Figure 3.20.: $\mathcal{I}_a$ and $\mathcal{I}_b$ are universal models of $\mathcal{K}_1$ and $\mathcal{K}_2$, respectively. All edges represent r-connections. Both models contain an infinite ray of anonymous elements starting from e_0 and a is connected to each elements on the ray.

every element in the ray, so that it does not have an influence on any endomorphism of the ray. Note that while infinite rays can be constructed in $\mathcal{ELH}_\bot$ as well by a simple cyclic TBox, it is not possible to have a be connected to every element of the ray. This makes a crucial difference. Because a is connected to every element on the ray, shifting along the ray becomes a possibility for endomorphisms. In fact, it is the only possibility because in order to map two consecutive elements $e, e' \in \Delta^{\mathcal{I}_a}$ to the same target $t \in \Delta^{\mathcal{I}_a}$, t would have to have an r-self-loop, formally $(t,t) \in r^{\mathcal{I}_a}$, which cannot be the case in any universal model of $\mathcal{K}_1$. Therefore, every endomorphism of $\mathcal{I}_a$ is of the form

$$h_k := \{a \mapsto a\} \cup \{e_i \mapsto e_{i+k} \mid i \in \mathbb{N}\} \qquad \text{for } k \in \mathbb{N},$$

which denotes the size of the shift along the ray. It is easy to check that h_k is injective for every $k \in \mathbb{N}$, hence for $\mathbf{I}(\mathcal{I}_a)$ holds. For any $k > 0$, the endomorphism is not surjective, since the first k elements of the ray are not part of the image of h_k. Non-relations are also not preserved, since e_0 will be mapped to e_k for any $k > 0$ and $e_0 \notin B^{\mathcal{I}_a}$, but $e_k \in B^{\mathcal{I}_a}$.

If $\mathbf{S}$ does not hold for $\mathcal{I}_a$ itself, does it hold for a substructure of of $\mathcal{I}_a$? It turns out that $\mathcal{I}_a$ does not even have an $\mathbf{S}$-core: Every substructure obtained as the image of $h_k(\mathcal{I}_a)$ with $k > 0$ can again be shifted by at least one element and so on. So every substructure we obtain with an endomorphism admits further endomorphisms that are not surjective. In contrast, $\mathcal{I}_a$ has an $\mathbf{N}$-core which can be obtained as the image of $h_k(\mathcal{I}_a)$ for any $k > 0$. Since the results are all isomorphic for any $k > 0$, $\mathcal{I}_a$ has an $\mathbf{N}$-core that is unique up to isomorphism. However, $h_k(\mathcal{I}_a)$ is not a model of $\mathcal{K}_1$, since the first element e_k in the ray of $h_k(\mathcal{I}_a)$ belongs to both B and A, but does not have an A-successor, which violates the axioms in $\mathcal{K}_1$. To summarize, for $\mathcal{K}_1$ there is a universal model, namely $\mathcal{I}_a$, that is an **I**-core, has no **S**-core, but an **N**-core that is not a model.

Unfortunately, we can construct examples where we get issues also with the **I** property. We modify the KB $\mathcal{K}_1$ slightly to the following:

Example 3.22. Let the Horn-$\mathcal{ALCOI}$ KB $\mathcal{K}_2$ consist of the following axioms:

$$A \sqsubseteq \exists r.(A \sqcap B) \qquad B \sqsubseteq \exists r^-.\{a\} \qquad C \sqsubseteq \exists r.A \qquad C(a) \qquad \diamondsuit$$

As before, $\mathcal{I}_a$ is a universal model for $\mathcal{K}_2$ and an **I**-core. Additionally, this time also $h_k(\mathcal{I}_a)$ is a universal model and $\mathbf{I}(h_k(\mathcal{I}_a))$ holds for all $k \in \mathbb{N}$. So for $\mathcal{K}_2$ there are two different universal **I**-core models that are not isomorphic and yield different answers to NCQs: For instance, to the query

$$q(x) := \exists y.r(x, y) \wedge A(y) \wedge \neg B(y),$$

a is an answer in $\mathcal{I}_a$, but not in any $h_k(\mathcal{I}_a)$ with $k > 0$. Since a we be the intuitive answer, using 'certain core answers', by returning only answers that are answers in every core model, is also not an option.

To circumvent such problems, uniqueness of cores seems to be a necessary condition. It has been shown in [Bau95] that the properties **ISN** and **SN** are the only combinations that preserves uniqueness of a core for infinite structures. However, while **ISN**-cores can guarantee uniqueness, we have already seen that for $\mathcal{I}_a$ not even an **S**-core exists, which implies that it does not have an **ISN**-core either. When nominals and inverse roles are allowed to interact, none of the properties satisfy all the requirements for a suitable core definition to base a minimal-world semantics on.

Transitivity

If transitivity of roles is available, we can construct a ray in a manner similar to $\mathcal{K}_2$, without using inverse roles or nominals:

Example 3.23. Let the Horn-$\mathcal{S}$ KB $\mathcal{K}_3$ consist of the following axioms:

$$A \sqsubseteq \exists r.(A \sqcap B) \qquad r \circ r \sqsubseteq r \qquad C \sqsubseteq \exists r.A \qquad C(a) \qquad \diamondsuit$$

A universal model $\mathcal{I}_b$ can be found in Figure 3.20. The only difference to $\mathcal{I}_a$ are the additional r-connections from every element e_i with $i \in \mathbb{N}$ in the ray to every element e_j with $j > i + 1$. As for $\mathcal{I}_a$, all endomorphisms are of the form of h_k with $k \in \mathbb{N}$. Without the use of an inverse role it seems impossible to require the initial element of the ray to satisfy only A and not B as it is the case in $\mathcal{K}_1$. Therefore, similarly to the case of $\mathcal{K}_2$, $h_k(\mathcal{I}_b)$ is a model of $\mathcal{K}_3$ for every $k \in \mathbb{N}$. This means that $\mathcal{I}_b$ has multiple non-isomorphic **I**-cores, which makes it unsuitable for minimal-world semantics. When using **IN**-cores, $\mathcal{I}_b$ is not a core, but $h_k(\mathcal{I}_b)$. This means that q would not yield any answers when evaluated over an universal **IN**-core model of $\mathcal{K}_3$, which is not the expected behavior.

Core Definition	Horn-$\mathcal{ALCOI}$	Horn-$\mathcal{S}$
I	universal models may have multiple non-isomorphic cores that are not necessarily universal models (see $\mathcal{K}_1$; $\mathcal{I}_a$); unintuitive query results (see $\mathcal{K}_2$ and q)	multiple non-isomorphic cores (see $\mathcal{K}_3, \mathcal{I}_b$)
S, **SN**, **IS**, **ISN**	core does not exist (see $\mathcal{I}_a$)	core does not exist (see $\mathcal{I}_b$)
IN, **N**	universal models may have cores that are not universal models (see $\mathcal{K}_1, \mathcal{I}_a$)	unintuitive query results (see $\mathcal{K}_3$ and q)

Table 3.24.: A summary of the different possibilities to define a core and the problems when queries should be answered over the resulting respective universal core models of either Horn-$\mathcal{ALCOI}$-KBs, which allow the use of nominals and inverse roles, or Horn-$\mathcal{S}$, which allows transitivity.

We have seen that for $\mathcal{ELH}_\bot$ minimal-world semantics can be defined without any complications. As soon as we extend it by either nominals and inverses or transitivity we fail to find a suitable definition of minimality, no matter which combination of properties we employ. The crucial difference between $\mathcal{ELH}_\bot$ and the extensions might be the degree of the universal models: In $\mathcal{ELH}_\bot$ there always exists a universal model that is *locally finite*, i.e. in which each element has only a finite number of successors. This property does not hold anymore in any of the two extensions we have looked at. $\mathcal{K}_1$, $\mathcal{K}_2$ and $\mathcal{K}_3$ are all KBs for which no locally-finite universal model exists, which leads to the difficulties we discussed.

3.5.3. Cores in Existential Rules

Recently, similar observations were made in the field of existential rules, which is a more general setting than $\mathcal{ELH}_\bot$. An *existential rule* ρ is a first-order formula of the form

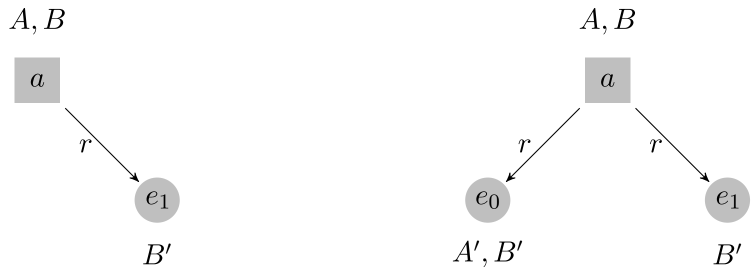
$$\rho = \forall \mathbf{x}, \mathbf{y}.\varphi[\mathbf{x}, \mathbf{y}] \rightarrow \exists \mathbf{z}.\psi[\mathbf{y}, \mathbf{z}],$$

where φ and ψ are conjunctions of atoms containing only constants or elements from mutually disjoint lists of variables $\mathbf{x}, \mathbf{y}, \mathbf{z}$. The left-hand side of the implication is called *body* and the right-hand side is called *head*. For simpler notation we omit the quantifiers in front of the rules from now on. For example, the axioms available in (normalized) $\mathcal{ELH}_\bot$ can be expressed by existential rules in the following way:

$$\begin{array}{rll}
 & \rightarrow A(x) & \qquad A(x) \\
 & \rightarrow r(x,y) & \qquad r(x,y) \\
\mathrm{A}(x) & \rightarrow B(x) & \qquad A \sqsubseteq B \\
\mathrm{A}(x) & \rightarrow A_1(x) \land A_2(x) & \qquad A \sqsubseteq A_1 \sqcap A_2 \\
r(x,y) \land \mathrm{B}(y) & \rightarrow A(x) & \qquad \exists r.B \sqsubseteq A \\
\mathrm{A}(x) & \rightarrow \exists y.r(x,y) \land B(y) & \qquad A \sqsubseteq \exists r.B \\
\mathrm{r}(x,y) & \rightarrow s(x,y) & \qquad r \sqsubseteq s
\end{array}$$

The semantics for existential rules are adopted from first-order logic. To construct models of a given finite set Σ of existential rules, so-called *chase algorithms* are employed: They construct a sequence of interpretations by consecutively applying rules if their body can be homomorphically mapped to some part of the current interpretation. The chase terminates when no more rules are applicable to any part of the model. If the chase algorithm terminates (which it does not necessarily do), the resulting interpretation is a model of Σ. Moreover it is universal in the sense we have defined before, which is why certain answers to CQs can be computed in this way. Since the order of rule application plays a crucial role in the termination of the chase, different chase algorithms have been proposed: The *standard chase* does not make any assumption on the order of rule application apart from fairness (every rule is applied at some step). In many cases this does not terminate, which is why Deutsch, Nash, and Remmel introduced the *core chase* [DNR08], where in the following core denotes the **IN**-core. In a singe step the core chase first applies all applicable rules simultaneously and then minimizes the result by computing its core. It has been shown that the core chase, in contrast to the standard chase, is complete for finding universal models for any finitely satisfiable rule set.

However, when moving to non-finitely satisfiable rules sets, infinite interpretations can have several non-isomorphic cores, or none at all, as is discussed in [CKM+18]. In [Krö20] it is observed that even though the core model has many good properties for finitely satisfiable rule sets, it is not supported by any major reasoning system, supposedly because the computation of the core in every step is expensive. To overcome this, special cases are identified in which the standard chase can be used to compute a core "incidentally". The class of standard chase sequences that are guaranteed to produce core models are characterized in terms of *alternative matches*.



(a) $\mathcal{I}'$ (b) $\mathcal{I}$

Figure 3.26.: Two models of Σ in Example 3.25.

Example 3.25. Consider the set Σ of existential rules given on the left, which correspond to the DL-axioms shown on the right.

$$\begin{aligned}
&\to A(a) && A(a) && (3.6)\\
A(x) &\to B(x) && A \sqsubseteq B && (3.7)\\
A'(x) &\to B'(x) && A' \sqsubseteq B' && (3.8)\\
B(x) &\to \exists y.r(x,y) \land B'(y) && B \sqsubseteq \exists r.B' && (3.9)\\
A(x) &\to \exists y.r(x,y) \land A'(y) && A \sqsubseteq \exists r.A' && (3.10)
\end{aligned}$$

$\diamondsuit$

Consider the two interpretations in Figure 3.26. Suppose the standard chase constructed model $\mathcal{I}$. During the chase $\mathcal{I}'$ was obtained through the application of a rule ρ (rule 3.9) by the mapping $h = \{x \mapsto a, y \mapsto e_1\}$ and the introduction of e_1. Then we can find a homomorphism $h' : \{x, y\} \to \Delta^{\mathcal{I}}$, that is identical with h on all variables occurring in the body (left-hand side) of ρ, but not identical for at least one variable that occurs only in the head (right-hand side) of ρ, by setting $h' := \{x \mapsto a, y \mapsto e_0\}$. A homomorphism that satisfies these conditions is called an *alternative match*. Krötzsch shows that whenever there are no alternative matches in a chase, then the result is a core model. However, even when there are alternative matches the result can still be a core model. Unfortunately, it is undecidable if some chase has alternative matches. To capture some cases where the application of a rule might introduce alternative matches for another rule that was applied earlier, a *restrainment* relation between rules is introduced that, similar to our structural subsumption, causes specific rules to be executed before (or instead of) others. Finally, non-monotonic negation in the core chase is investigated. In contrast to our setting, negation is not only allowed to occur in the query, but also on the left-hand side of TBox axioms. It is shown that this can lead to conflicts between non-monotonicity and cores in certain cases. Moreover, the focus for Krötzsch is on finite satisfiability, while we have infinite models and deal with this through a rewriting.

We have seen that, on the one hand, no matter which core definition is adopted, we can either guarantee uniqueness, but not existence, or we can guarantee existence, but lose uniqueness. On the other hand, the loss of good properties of cores for more expressive logics causes a discrepancy from the intuition behind minimal-world semantics: Arguably, a should be an answer to q in $\mathcal{K}_1$, $\mathcal{K}_2$ and $\mathcal{K}_3$ from Examples 3.21 to 3.23.

However, this cannot be achieved using cores in any obvious way, since, for example, in the sense of **IN**-cores, e_0 is a redundant element and therefore it can be merged into the chain of infinite length without it making a mathematical difference.

3.5.4. "Intuitively" Minimal Models for Horn-$\mathcal{ALCHOIQ}^{\mathsf{Disj}}_{\mathsf{Self}}$

While difficult to define formally, the intuition behind minimal-world semantics seems to be simple to grasp for a human. Based on this intuition to include just the necessary elements, a student work [Khy20] already proposed a construction for an (intuitively) minimal universal model for Horn-$\mathcal{ALCHOIQ}^{\mathrm{Disj}}_{\mathrm{Self}}$, an extension of Horn-$\mathcal{ALCHOIQ}$ with self-loops and role disjointness axioms. For the construction structural subsumption is extended to include also qualified number restrictions. Based on the subsumption relation, new successors are then introduced, starting from the currently minimal restriction α. If α is satisfied, the restriction is removed from the set and the next minimal restriction β is considered. Since α already introduced some of the successors required for β, only the difference needed to satisfy β is introduced. Additionally, nominals need to be taken into account when determining how many new successors need to be introduced. In [ORS11] it has been shown that CQ answering w.r.t. certain answer semantics over Horn-$\mathcal{SROIQ}$ can be reduced to CQ answering over Horn-$\mathcal{ALCHOIQ}^{\mathrm{Disj}}_{\mathrm{Self}}$. Similarly, it might be possible to show (after finding a suitable definition of minimality) that for a given Horn-$\mathcal{SROIQ}$-KB $\mathcal{K}$, there is a corresponding Horn-$\mathcal{ALCHOIQ}^{\mathrm{Disj}}_{\mathrm{Self}}$-KB $\mathcal{K}'$ such that the minimal universal model of $\mathcal{K}'$ corresponds to a minimal universal model of $\mathcal{K}$.

3.5.5. Queries beyond (N)CQs

Instead of moving to more expressive DLs, we could also use a more expressive query language.Suppose we want to select 'parents that have two kids'. With (N)CQs is not possible to count successors. One possibility for adding counting capabilities to NCQs would be to add inequalities, allowing us to pose the above query as

$$q(x) = \exists y_1.\exists y_2.\mathtt{parentOf}(x, y_1) \land \mathtt{parentOf}(x, y_2) \land y_1 \neq y_2.$$

Unfortunately, it has been shown that at least for certain answer semantics answering CQs with general inequalities over $\mathcal{EL}_\bot$ is undecidable [GIK+15]. In the case of minimal-world semantics a rewriting approach could be possible if the TBox axioms can be used to decide when a certain inequality can occur in the anonymous part of the canonical model. Additionally, the filters would have to be extended to deal with potentially untypical structures in the ABox. However, with general inequalities arbitrary variables in the query can be required to be not equal, no matter how many role connections they are apart from each other in the structure of the query. It could be difficult to take this into account, since the rewriting works by consecutively removing leaf variables in the query. Without inequalities, the leaf variable depends only on its predecessor variable(s). With inequalities, an additional layer of dependencies is introduced. Therefore, it might be necessary to make some restrictions on the use of inequalities.

Chapter 4.

Temporalizing $\mathcal{ELH}_\bot$

In the previous chapter we have introduced minimal-world semantics for $\mathcal{ELH}_\bot$ ontologies that can, for example, be used to give suitable semantics to NCQs for the patient selection task. Apart from negation, we need to be able to represent the temporal dimension: Many clinical trials contain temporal eligibility criteria [CT15], such as:

- "type 1 diabetes with duration at least 12 months"[1]; or
- "known history of heart disease or heart rhythm abnormalities"[2].

Moreover, measurements, diagnoses, and treatments in a patients' EHR are clearly valid only for a certain amount of time. Since EHRs only contain information for specific points in time, it is especially important to be able to infer what happened to the patient in the meantime. For example, if a patient is diagnosed with a (currently) incurable disease like `Diabetes`, they will still have the disease at any future point in time. Similarly, if the EHR contains two entries of `CD4Above250`[3] one week apart, one may reasonably infer that this was true for the whole week. Qualitative temporal DLs such as $\mathcal{TEL}^{\Diamond}_{\mathsf{infl}}$ [GJK16] can express the former statement by declaring `Diabetes` as *expanding* via the axiom $\Diamond\mathtt{Diabetes} \sqsubseteq \mathtt{Diabetes}$. We propose to extend this logic by adding a special kind of metric temporal operators, introduced in Section 4.1, to write

$$\Diamond\!\!\!\!\cdot\;_7\mathtt{CD4Above250} \sqsubseteq \mathtt{CD4Above250},$$

making the measurement *convex* for a specified length of time n (e.g. 7 days). This means that information is interpolated between time points of distance less than n, thereby computing a convex closure of the available information. The threshold n allows us to distinguish the case where two mentions of `CD4Above250` are years apart, and are therefore unrelated.

The distinguishing feature of $\mathcal{TEL}^{\Diamond}_{\mathsf{infl}}$ is that $\Diamond$-operators are only allowed on the left-hand side of concept inclusions [GJK16], which is also common for temporal DLs based on *DL-Lite* [AKW+13; AKK+15]. In Section 4.2 we introduce $\mathcal{TELH}^{\Diamond\!\!\!\cdot\,,\mathsf{lhs}}_\bot$, an extension of $\mathcal{TEL}^{\Diamond}_{\mathsf{infl}}$ by convex metric temporal operators.
We allow temporal roles like $\Diamond\!\!\!\!\cdot\;_2\mathtt{hasTreatment} \sqsubseteq \mathtt{hasTreatment}$, and provide a completion algorithm in Section 4.3 that can deal with the problem of having large temporal gaps in the data, e.g. in patient records. We show that reasoning in $\mathcal{TELH}^{\Diamond\!\!\!\cdot\,,\mathsf{lhs}}_\bot$ remains tractable. We end with some ideas for future work and discussion of related work in Section 4.4.

[1] `https://clinicaltrials.gov/ct2/show/NCT02280564`
[2] `https://clinicaltrials.gov/ct2/show/NCT02873052`
[3] The CD4 value is the ratio of T helper cells (which have the surface marker C4) to cytotoxic T cells. A reduced ratio is associated with reduced resistance to infections [YJY+15].

4.1. Convexity Operators

We first introduce the LTL$^{\mathsf{bin}}$ operators that we will use and analyze their properties. We consider only formulas built according to the syntax rule

$$\varphi ::= p \mid \varphi \wedge \varphi \mid \varphi \vee \varphi \mid \Diamond_I \varphi,$$

where $p \in P$ and I is an interval in $\mathbb{Z}$. More specifically, we only consider the following derived operators, where $n \geq 1$.

$$\begin{aligned} &\overset{\leftrightarrow}{\Diamond}\varphi := \Diamond_{(-\infty,\infty)}\varphi && \overset{\rightarrow}{\Diamond}\varphi := \Diamond_{[0,\infty)}\varphi && \overset{\leftarrow}{\Diamond}\varphi := \Diamond_{(-\infty,0]}\varphi \\ &\Diamond^{c}\varphi := \Diamond_{(-\infty,0]}\varphi \wedge \Diamond_{[0,\infty)}\varphi && \Diamond^{c}_n\varphi := \bigvee_{\substack{k,m\geq 0\\ k+m=n-1}} (\Diamond_{[-k,0]}\varphi \wedge \Diamond_{[0,m]}\varphi) \end{aligned} \tag{4.1}$$

The operator $\overset{\rightarrow}{\Diamond}$ is the 'eventually' operator of classical LTL, and $\overset{\leftarrow}{\Diamond}, \overset{\leftrightarrow}{\Diamond}$ are two variants that refer to the past, or to both past and future, respectively. The *convexity operator* $\Diamond^{c}$ requires that φ holds *both* in the past and in the future, thereby distinguishing time points that lie within an interval enclosed by time points at which φ holds. This can be used to express the convex closure of time points, as described in the introduction. Finally, the *bounded convexity operators* $\Diamond^{c}_n$ represent a metric variant of $\Diamond^{c}$, requiring that different occurrences of φ are at most $n-1$ time points apart, i.e. enclose an interval of length n.

To study the behavior of these operators, we consider their semantics in a more abstract way: given a set of time points where a certain information is available (e.g. a diagnosis), described by a propositional variable p, we consider the resulting set of time points at which $\hat{\Diamond}p$ holds, where $\hat{\Diamond}$ is a placeholder for one of the operators defined above (we will similarly use $\blacklozenge, \tilde{\Diamond}, \check{\Diamond}$ as placeholders for different $\Diamond$-operators in the following).

Definition 4.1. We consider the sets

$$\mathfrak{D}^{\mathsf{c}} := \{\Diamond^{c}\} \cup \{\Diamond^{c}_i \mid i \geq 1\}, \qquad \mathfrak{D}^{\pm} = \{\overset{\leftarrow}{\Diamond}, \overset{\rightarrow}{\Diamond}, \overset{\leftrightarrow}{\Diamond}\}, \text{ and } \qquad \mathfrak{D} := \mathfrak{D}^{\pm} \cup \mathfrak{D}^{\mathsf{c}}$$

of *diamond operators*. Each diamond operator $\hat{\Diamond} \in \mathfrak{D}$ induces a function $\hat{\Diamond}\colon 2^{\mathbb{Z}} \to 2^{\mathbb{Z}}$ with $\hat{\Diamond}(M) := \{i \mid \mathfrak{W}_M, i \models \hat{\Diamond}p\}$ for all $M \subseteq \mathbb{Z}$, with the LTL-structure $\mathfrak{W}_M := (w_i)_{i\in\mathbb{Z}}$ such that $w_i := \{p\}$ if $i \in M$, and $w_i := \emptyset$ otherwise. $\Diamond$

We will omit the parentheses in $\hat{\Diamond}(M)$ for a cleaner presentation. If M is empty, then $\hat{\Diamond}M$ is also empty, for any $\hat{\Diamond} \in \mathfrak{D}$. For any non-empty $M \subseteq \mathbb{Z}$, we obtain the following expressions, where $\max M$ may be ∞ and $\min M$ may be $-\infty$.

$$\begin{aligned} &\overset{\leftrightarrow}{\Diamond}M = \mathbb{Z} && \overset{\leftarrow}{\Diamond}M = (-\infty, \max M] && \overset{\rightarrow}{\Diamond}M = [\min M, \infty) && \Diamond^{c}M = [\min M, \max M] \\ &\Diamond^{c}_1 M = M && \Diamond^{c}_n M = \{i \in \mathbb{Z} \mid \exists j,k \in M \text{ with } j \leq i \leq k \text{ and } k-j < n\} \end{aligned}$$

In this representation, the convex closure operation behind $\Diamond^{c}$ becomes apparent. See Figure 4.2 for a graphical illustration of the diamond operators. We now list several useful properties of these functions.

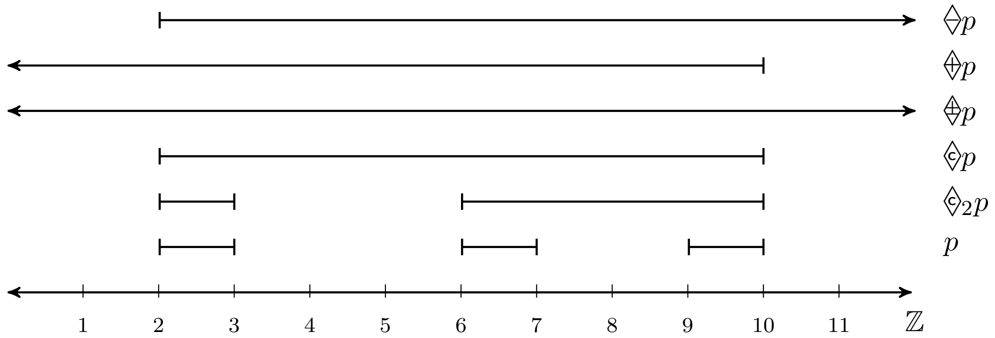


Figure 4.2.: A graphical illustration of the diamond operators. Suppose the predicate p is valid at the intervals given directly above the timeline. Then for each of temporal operator $\Diamond^{\ast}$, the interval is given at which $\Diamond^{\ast}p$ holds.

Lemma 4.3. Using the point-wise inclusion order $\subseteq$ on the induced functions, we obtain the following ordered set $(\mathfrak{D}, \subseteq)$, where $\mathrm{id}_{2^{\mathbb{Z}}}$ is the identity function on $2^{\mathbb{Z}}$:

$$\mathrm{id}_{2^{\mathbb{Z}}} = \Diamond^{c}_1 \subseteq \cdots \subseteq \Diamond^{c}_n \subseteq \Diamond^{c}_{n+1} \subseteq \cdots \subseteq \Diamond^{c} \begin{array}{c} \subseteq \Diamond^{+} \subseteq \\ \subseteq \Diamond^{-} \subseteq \end{array} \Diamond^{\pm}$$

Proof. We only show that $\Diamond^{c}_n \subseteq \Diamond^{c}_{n+1}$ for all $n \in \mathbb{N}$; the remaining inclusions are easy to check. If $i \in \Diamond^{c}_n$ then there exists $j, k \in M$ with $i \in [j, k]$ and $k - j < n$. The same choice of j, k is also valid for $\Diamond^{c}_{n+1}$, since $k - j < n$ implies $k - j < n + 1$. Hence $i \in \Diamond^{c}_{n+1}$. $\square$

We prove some additional technical lemmas.

Lemma 4.4. Each $\Diamond^{\ast} \in \mathfrak{D}$ is extensive and monotone, i.e. for all $M_1 \subseteq M_2 \subseteq \mathbb{Z}$, it holds that $M_1 \subseteq \Diamond^{\ast}M_1 \subseteq \Diamond^{\ast}M_2$.

Proof. Extensivity follows from Lemma 4.3 and monotonicity is obvious for most of the operators. For $\Diamond^{c}_n$ if $i \in \Diamond^{c}_n M_1$ by choosing $j, k \in M_1$, then since $M_1 \subseteq M_2$, the choice of j, k is also a valid choice in $\Diamond^{c}_n M_2$ and hence $i \in \Diamond^{c}_n M_2$. $\square$

Lemma 4.5. For all $\Diamond^{\ast} \in \mathfrak{D}^{\mathsf{c}}$, we have $\Diamond^{\ast}\{i\} = \{i\}$ for all $i \in \mathbb{Z}$. For all $\Diamond^{\ast} \in \mathfrak{D}^{\pm}$ and $M \subseteq \mathbb{Z}$, we have $\Diamond^{\ast}M = \bigcup_{i \in M} \Diamond^{\ast}\{i\}$.

Proof. The claim for $\Diamond^{\ast} \in \mathfrak{D}^{\mathsf{c}}$ is obvious. For $\Diamond^{-}$, we have

$$\Diamond^{-}M = (-\infty, \max M] = \bigcup_{i \in M} (-\infty, i] = \bigcup_{i \in M} \Diamond^{-}\{i\}.$$

The cases for $\Diamond^{+}$ and $\Diamond^{\pm}$ are similar. $\square$

The most important property is the following, which allows us to combine diamond operators without leaving the set $\mathfrak{D}$.

Lemma 4.6. The set $\mathfrak{D}$ is closed under composition $\circ$, point-wise intersection $\cap$, and point-wise union $\cup$, and for any $\Diamond^{\star}, \Diamond^{\bullet} \in \mathfrak{D}$ these operators can be computed as:

$$\Diamond^{\star} \cap \Diamond^{\bullet} = \inf_{(\mathfrak{D},\subseteq)}\{\Diamond^{\star}, \Diamond^{\bullet}\} \qquad \text{and} \qquad \Diamond^{\bullet} \circ \Diamond^{\star} = \Diamond^{\star} \cup \Diamond^{\bullet} = \sup_{(\mathfrak{D},\subseteq)}\{\Diamond^{\star}, \Diamond^{\bullet}\},$$

where $\inf_{(\mathfrak{D},\subseteq)}$ denotes the infimum in $(\mathfrak{D}, \subseteq)$, and $\sup_{(\mathfrak{D},\subseteq)}$ the supremum.

Proof. For the first claim, we distinguish two cases.

1. If $\Diamond^{\star} \subseteq \Diamond^{\bullet}$, then $\Diamond^{\star} \cap \Diamond^{\bullet} = \Diamond^{\star}$, and similarly for $\Diamond^{\bullet} \subseteq \Diamond^{\star}$.

2. If neither $\Diamond^{\star} \subseteq \Diamond^{\bullet}$ nor $\Diamond^{\bullet} \subseteq \Diamond^{\star}$, one of them must be equal to $\Diamond^{\downarrow}$ and the other to $\Diamond^{\ominus}$, and $\Diamond^{\downarrow} \cap \Diamond^{\ominus} = \Diamond^{c}$ holds by definition.

The result is exactly the infimum w.r.t. the relation $\subseteq$ from Lemma 4.3. The arguments for union are similar.

We show that $(\Diamond^{\bullet} \circ \Diamond^{\star})M = (\Diamond^{\star} \cup \Diamond^{\bullet})M$ holds for any $M \subseteq \mathbb{Z}$. The case where $M = \emptyset$ is trivial and we assume in the following that $M \neq \emptyset$. We distinguish three cases.

1. Suppose that $\Diamond^{\star} = \Diamond^{c}_m$ and $\Diamond^{\bullet} = \Diamond^{c}_n$ and $m \geq n$. By Lemma 4.4, we know that $\Diamond^{c}_m M \subseteq \Diamond^{c}_n(\Diamond^{c}_m M) = (\Diamond^{c}_n \circ \Diamond^{c}_m)M$.
   For the converse direction, let $i \in (\Diamond^{c}_n \circ \Diamond^{c}_m)M$. Then there exist $j, k \in \Diamond^{c}_m M$ with $j \leq i \leq k$ and $k - j < n$. This means that there have to be $a_1, b_1, a_2, b_2 \in M$ with $a_1 \leq j \leq b_1$, $b_1 - a_1 < m$, $a_2 \leq k \leq b_2$, and $b_2 - a_2 < m$.
   If $a_2 > b_1$, then $a_1 < b_2$ and $a_2 - b_1 \leq k - j < n \leq m$, and thus $\{a_1, b_1, a_2, b_2\} \subseteq M$ implies that $i \in [j, k] \subseteq [a_1, b_2] \subseteq \Diamond^{c}_m M$. Otherwise, $a_2 \leq b_1$ and $a_1 \leq j \leq k \leq b_2$, and hence the two intervals $[a_1, b_1]$ and $[a_2, b_2]$ overlap. Thus,

   $$i \in [j, k] \subseteq [\min\{a_1, a_2\}, \max\{b_1, b_2\}] = [a_1, b_1] \cup [a_2, b_2] \subseteq \Diamond^{c}_m M.$$

   For the case $n > m$, the arguments are similar, and we thus obtain

   $$(\Diamond^{\star} \circ \Diamond^{\bullet}) = \Diamond^{c}_{\max(n,m)} = (\Diamond^{\star} \cup \Diamond^{\bullet}).$$

2. Suppose that $\Diamond^{\star} = \Diamond^{c}_n$ and $\Diamond^{\bullet} \in \mathfrak{D}^{\pm}$. Then we know that $\min M = \min \Diamond^{\star} M$ and $\max M = \max \Diamond^{\star} M$, since only elements in between the already existing elements can be added. For the application of $\Diamond^{\bullet}$ this does not make a difference, hence we have $(\Diamond^{\bullet} \circ \Diamond^{\star}) = \Diamond^{\bullet} = \Diamond^{\star} \cup \Diamond^{\bullet}$. The case where $\Diamond^{\bullet} = \Diamond^{c}_n$ and $\Diamond^{\star} \in \mathfrak{D}^{\pm}$ is similar.

3. What remains is the case that $\Diamond^{\star}, \Diamond^{\bullet} \in \mathfrak{D}^{\pm}$. We only show the case of $\Diamond^{\star} = \Diamond^{\downarrow}$; the remaining cases follow the same arguments. If $\Diamond^{\star} = \Diamond^{\downarrow}$, then $\Diamond^{\downarrow} M = (-\infty, \max M]$ will be transformed by applying $\Diamond^{\bullet}$ to either $(-\infty, \max M]$ (if $\Diamond^{\bullet} = \Diamond^{\downarrow}$), or to $\mathbb{Z}$ (if $\Diamond^{\bullet} \in \{\Diamond^{\ominus}, \Diamond^{\uparrow}\}$). In both cases, the result is $(\Diamond^{\star} \cup \Diamond^{\bullet})(M)$. $\square$

4.2. The Temporal Description Logic $\mathcal{TELH}_{\bot}^{\diamondsuit,\mathsf{lhs}}$

We define a new temporal logic based on the operators in $\mathfrak{D}$. The main differences to $\mathcal{TEL}_{\mathsf{infl}}^{\diamondsuit}$ from [GJK16] are that $\diamondsuit_n$-operators may occur in concept and role inclusions, and ABoxes may have gaps, which require special consideration during reasoning.

Definition 4.7 (Syntax of $\mathcal{TELH}_{\bot}^{\diamondsuit,\mathsf{lhs}}$). Let N_C, N_R, N_I be disjoint sets of *concept*, *role*, and *individual names*, respectively. A *temporal role* is of the form $\diamondsuit r$ with $\diamondsuit \in \mathfrak{D}$ and $r \in N_R$. A $\mathcal{TELH}_{\bot}^{\diamondsuit,\mathsf{lhs}}$ *concept* is built using the rule

$$C ::= A \mid \top \mid \bot \mid C \sqcap C \mid \exists r.C \mid \diamondsuit C,$$

where $A \in N_C$, $\diamondsuit \in \mathfrak{D}$, and r is a temporal role. Such a C is an $\mathcal{ELH}_{\bot}$ *concept* (or *atemporal concept*) if it does not contain any diamond operators.

A $\mathcal{TELH}_{\bot}^{\diamondsuit,\mathsf{lhs}}$ *TBox* is a finite set of *concept inclusions (GCIs)* $C \sqsubseteq D$ and *role inclusion axioms (RIAs)* $r \sqsubseteq s$, where C is a $\mathcal{TELH}_{\bot}^{\diamondsuit,\mathsf{lhs}}$ concept, D is an atemporal concept, r is a temporal role, and $s \in N_R$. An *ABox* is a finite set of *concept assertions* $A(a,i)$ and *role assertions* $r(a,b,i)$, where $A \in N_C$, $r \in N_R$, $a, b \in N_I$, and $i \in \mathbb{Z}$.

The set of time points $i \in \mathbb{Z}$ occurring in $\mathcal{A}$ we denote as $\mathrm{tem}(\mathcal{A})$. Also we assume each time point is encoded in binary with at most n digits. A *knowledge base (KB)* $\mathcal{K} = \mathcal{T} \cup \mathcal{A}$ consists of a TBox $\mathcal{T}$ and an ABox $\mathcal{A}$. $\Diamond$

In the following, we always assume a KB $\mathcal{K} = \mathcal{T} \cup \mathcal{A}$ to be given. Moreover, we assume N_I to be non-empty.

Definition 4.8 (Semantics of $\mathcal{TELH}_{\bot}^{\diamondsuit,\mathsf{lhs}}$). A *temporal interpretation* $\mathfrak{I} = (\Delta^{\mathfrak{I}}, (\mathcal{I}_i)_{i \in \mathbb{Z}})$, is a collection of interpretations $\mathcal{I}_i = (\Delta^{\mathfrak{I}}, \cdot^{\mathcal{I}_i})$, $i \in \mathbb{Z}$, over $\Delta^{\mathfrak{I}}$. The functions $\cdot^{\mathcal{I}_i}$ are extended as follows.

$$\begin{aligned}
\top^{\mathcal{I}_i} &:= \Delta^{\mathfrak{I}} \\
\bot^{\mathcal{I}_i} &:= \emptyset \\
(\diamondsuit C)^{\mathcal{I}_i} &:= \{d \in \Delta^{\mathfrak{I}} \mid i \in \diamondsuit\{j \mid d \in C^{\mathcal{I}_j}\}\} \\
(\diamondsuit r)^{\mathcal{I}_i} &:= \{(d,e) \in \Delta^{\mathfrak{I}} \times \Delta^{\mathfrak{I}} \mid i \in \diamondsuit\{j \mid (d,e) \in r^{\mathcal{I}_j}\}\} \\
(C \sqcap D)^{\mathcal{I}_i} &:= C^{\mathcal{I}_i} \cap D^{\mathcal{I}_i} \\
(\exists r.C)^{\mathcal{I}_i} &:= \{d \in \Delta^{\mathfrak{I}} \mid \exists e \in C^{\mathcal{I}_i} \colon (d,e) \in r^{\mathcal{I}_i}\}
\end{aligned}$$

$\mathfrak{I}$ is a *model* of (or *satisfies*)

- a concept inclusion $C \sqsubseteq D$ if $C^{\mathcal{I}_i} \subseteq D^{\mathcal{I}_i}$ holds for all $i \in \mathbb{Z}$,
- a role inclusion $r \sqsubseteq s$ if $r^{\mathcal{I}_i} \subseteq s^{\mathcal{I}_i}$ holds for all $i \in \mathbb{Z}$,
- a concept assertion $A(a,i)$ if $a \in A^{\mathcal{I}_i}$,
- a role assertion $r(a,b,i)$ if $(a,b) \in r^{\mathcal{I}_i}$,
- and the KB $\mathcal{K}$ if it satisfies all axioms in $\mathcal{K}$.

This fact is denoted by $\mathfrak{I} \models \alpha$, where α is an *axiom* (i.e. inclusion or assertion) or a KB. An KB $\mathcal{K}$ is *consistent* if it has a model, and it *entails* α, written $\mathcal{K} \models \alpha$, if all models of $\mathcal{K}$ satisfy α. $\Diamond$

$\mathcal{K}$ is inconsistent iff $\mathcal{K} \models \top \sqsubseteq \bot$, and thus we focus on deciding entailment. In $\mathcal{ELH}_\bot$, this is possible in polynomial time [BBL05].

We do not allow diamonds to occur on the right-hand side of GCIs, because that would make the logic undecidable [AKL+07]. As usual, we can simulate GCIs involving complex concepts by introducing fresh concept and role names as abbreviations. For example, $\exists \diamondplus r.\diamondminus A \sqsubseteq B$ can be split into $\diamondplus r \sqsubseteq r'$, $\diamondminus A \sqsubseteq A'$, and $\exists r'.A' \sqsubseteq B$. Hence, we can restrict ourselves w.l.o.g. to GCIs in the following *normal form*:

$$\diamondstar A \sqsubseteq B,\ A_1 \sqcap A_2 \sqsubseteq B,\ \diamondstar r \sqsubseteq s,\ \diamondstar A \sqsubseteq \exists r.B,\ \exists r.A \sqsubseteq B, \tag{4.2}$$

where $\diamondstar \in \mathfrak{D}$, $A, A_1, A_2, B \in N_C \cup \{\bot, \top\}$, and $r, s \in N_R$.

Convex Names

When considering axioms of the form $\diamondstar A \sqsubseteq A$ for $A \in N_C$, we can first observe that the converse direction $A \sqsubseteq \diamondstar A$, although syntactically not allowed, trivially holds in all interpretations. Moreover, the following implications between such equivalences follow from Lemma 4.3:

$$\begin{tikzcd}
 & \boxed{A \equiv \diamondcross A} \arrow[dr] & & & \\
\boxed{A \equiv \diamondplus A} \arrow[ur] \arrow[dr] & & \boxed{A \equiv \diamondc_n A} \arrow[r] & \cdots \arrow[r] & \boxed{A \equiv \diamondc_1 A} \\
 & \boxed{A \equiv \diamondminus A} \arrow[ur] & & &
\end{tikzcd}$$

Since $\{A \equiv \diamondcross A, A \equiv \diamondminus A\}$ entails $A \equiv \diamondplus A$, it thus makes sense to consider the *unique strongest* such axiom that is entailed by $\mathcal{K}$ (for a given A). We call A

- *rigid* if $A \equiv \diamondplus A$ is the strongest such axiom,
- *shrinking* in case of $A \equiv \diamondcross A$,
- *expanding* for $A \equiv \diamondminus A$, and
- *(n-)convex* for $A \equiv \diamondc_{(n)} A$, i.e. whenever A is satisfied at two time points (with distance $< n$), then it is also satisfied at all time points in between.

1-convex concept names do not satisfy any special property, and are also called *flexible*. We use the same terms for role names.

4.3. A Completion Algorithm

We use the completion rules in Figure 4.9 to derive new axioms from $\mathcal{K}$. For simplicity, we treat $\top$ and $\bot$ like concept names, and thus allow assertions of the form $\top(a, i)$ and $\bot(a, i)$ here. It is clear that we cannot derive all possible entailments of the forms $\diamondstar A \sqsubseteq B$ or $A(a, i)$, because (1) $\mathfrak{D}$ is infinite, and (2) $\mathbb{Z}$ is infinite. Moreover, there may be arbitrarily many time points between two assertions in $\mathcal{A}$ (exponentially many in the size of $\mathcal{A}$ if we assume time points to be encoded in binary).

To deal with (1), we restrict the rule applications to the operators that occur in $\mathcal{K}$, in addition to $\diamondc$ and $\diamondplus$, which are the only elements of $\mathfrak{D}$ that can be obtained via $\cup$, $\cap$, or $\circ$ from other $\Diamond$-operators, namely from $\diamondcross$ and $\diamondminus$.

$$\text{T1}\ \frac{}{\Diamond^{c}_1 A \sqsubseteq A} \qquad \text{T2}\ \frac{}{\Diamond^{+} A \sqsubseteq \top} \qquad \text{T3}\ \frac{}{\Diamond^{c}_1 r \sqsubseteq r} \qquad \text{T4}\ \frac{\Diamond^{\star} A_1 \sqsubseteq A_2 \quad \Diamond^{\bullet} A_2 \sqsubseteq A_3}{(\Diamond^{\bullet} \circ \Diamond^{\star}) A_1 \sqsubseteq A_3}$$

$$\text{T5}\ \frac{\Diamond^{\star} r_1 \sqsubseteq r_2 \quad \Diamond^{\bullet} r_2 \sqsubseteq r_3}{(\Diamond^{\bullet} \circ \Diamond^{\star}) r_1 \sqsubseteq r_3} \qquad \text{T6}\ \frac{\Diamond^{\star} A \sqsubseteq A_1 \quad \Diamond^{\bullet} A \sqsubseteq A_2 \quad A_1 \sqcap A_2 \sqsubseteq B}{(\Diamond^{\star} \cap \Diamond^{\bullet}) A \sqsubseteq B}$$

$$\text{T7}\ \frac{}{\exists r.\bot \sqsubseteq \bot} \qquad \text{T8}\ \frac{\Diamond^{\star} A \sqsubseteq \exists r.A_1 \quad \Diamond^{\bullet} r \sqsubseteq s \quad \Diamond^{\dagger} A_1 \sqsubseteq B_1 \quad \exists s.B_1 \sqsubseteq B}{\Diamond^{\star} A \sqsubseteq B}$$

$$\text{T8}'\ \frac{\Diamond^{\star} A \sqsubseteq \exists r.A_1 \quad \Diamond^{\bullet} r \sqsubseteq s \quad \Diamond^{\dagger} A_1 \sqsubseteq B_1 \quad \exists s.B_1 \sqsubseteq B \quad (\Diamond^{\bullet} \cap \Diamond^{\dagger}) \in \mathfrak{D}^{\pm}}{((\Diamond^{\bullet} \cap \Diamond^{\dagger}) \circ \Diamond^{\star}) A \sqsubseteq B}$$

$$\text{A1}\ \frac{}{\top(a,i)} \qquad \text{A2}\ \frac{i \in \Diamond^{\star} A(a) \quad \Diamond^{\star} A \sqsubseteq B}{B(a,i)} \qquad \text{A3}\ \frac{i \in \Diamond^{\star} r(a,b) \quad \Diamond^{\star} r \sqsubseteq s}{s(a,b,i)}$$

$$\text{A4}\ \frac{A_1(a,i) \quad A_2(a,i) \quad A_1 \sqcap A_2 \sqsubseteq B}{B(a,i)} \qquad \text{A5}\ \frac{r(a,b,i) \quad A(b,i) \quad \exists r.A \sqsubseteq B}{B(a,i)}$$

Figure 4.9.: Completion rules for $\mathcal{TELH}_{\bot}^{\Diamond^{c},\mathsf{lhs}}$ knowledge bases , where A, B, A_1, A_2, A_3, B_1 are $\top$, $\bot$ or (normalized) $\mathcal{ELH}_{\bot}$ concepts from $\mathcal{K}$; r, s, r_1, r_2, r_3 are role names from $\mathcal{K}$; $\Diamond^{\star}, \Diamond^{\bullet}, \Diamond^{\dagger}$ are $\Diamond^{c}$, $\Diamond^{+}$ or elements of $\mathfrak{D}$ occurring in $\mathcal{K}$; a, b are individual names from $\mathcal{K}$; and i are values from $\mathrm{rep}(\mathcal{A})$.

For (2), we consider the set of time points $\mathrm{tem}(\mathcal{A})$ (of linear size). Additionally, consider a maximal interval $[i, j]$ in $\mathbb{Z}\backslash\mathrm{tem}(\mathcal{A})$ (where i may be $-\infty$ and j may be ∞). To represent this interval, we choose a single representative time point $k \in [i, j]$, which is denoted by $|\ell| := k$ for all $\ell \in [i, j]$. For consistency, the representative $|i|$ for any $i \in \mathrm{tem}(\mathcal{A})$ is defined as i itself. Moreover, for any $k \in \mathbb{Z}$ we denote by $\lfloor k \rfloor := \max\{i \in \mathrm{tem}(\mathcal{A}) \mid i \leq k\}$ the maximal element of $\mathrm{tem}(\mathcal{A})$ below (or equal to) k, which we consider to be $-\infty$ in the case that there is no such element, and similarly define $\lceil k \rceil$. Note that $\lfloor i \rfloor = i = \lceil i \rceil$ whenever $i \in \mathrm{tem}(\mathcal{A})$, and otherwise $\lfloor i \rfloor < i < \lceil i \rceil$. By restricting all assertions to the finite set of representative time points

$$\mathrm{rep}(\mathcal{A}) := \{|i| \mid i \in \mathbb{Z}\} \supset \mathrm{tem}(\mathcal{A}),$$

we can encode infinitely many entailments in a finite set. We also define the following abbreviations, for all $A \in N_C$, $r \in N_R$, and $a, b \in N_I$ ($\mathcal{K}$ refers to the KB after possibly already applying some completion rules).

$$A(a) := \{i \in \mathrm{rep}(\mathcal{A}) \mid A(a, i) \in \mathcal{K}\}$$
$$r(a, b) := \{i \in \mathrm{rep}(\mathcal{A}) \mid r(a, b, i) \in \mathcal{K}\}$$

Hence, we can write $\Diamond^{\star} A(a)$ in the completion rules to refer to the set of time points at which $\Diamond^{\star} A$ is inferred to be satisfied by a, given only the assertions in $\mathcal{A}$.

In the rules of Figure 4.9, we allow to instantiate A, B, A_1, A_2, A_3, B_1 by $\top$, $\bot$ or (normalized) $\mathcal{ELH}_\bot$ concepts from $\mathcal{K}$, r, s, r_1, r_2, r_3 by role names from $\mathcal{K}$, $\Diamond^{\star}, \Diamond^{\bullet}, \Diamond^{\dagger}$ by $\Diamond^{c}$, $\Diamond^{+}$ or elements of $\mathfrak{D}$ occurring in $\mathcal{K}$, a, b by individual names from $\mathcal{K}$, and i by values from $\mathrm{rep}(\mathcal{A})$, such that the resulting axioms are in normal form. The side conditions $(\Diamond^{\bullet} \cap \Diamond^{\dagger}) \in \mathfrak{D}^{\pm}$, $i \in \Diamond^{\star}A(a)$, $i \in \Diamond^{\star}r(a, b)$ can be checked in polynomial time. All rules also apply to axioms without diamonds since we can treat A as $\Diamond^{c}_1 A$.

If $\mathcal{K}$ contains all axioms in the precondition of an instantiated rule, we consider the axiom in its conclusion. If it is a new assertion, we add it to $\mathcal{K}$. If it is a concept inclusion $\Diamond^{\star}A \sqsubseteq B$, we check whether $\mathcal{K}$ already contains a GCI of the form $\Diamond^{\bullet}A \sqsubseteq B$. If not, then we simply add $\Diamond^{\star}A \sqsubseteq B$ to $\mathcal{K}$; otherwise, and if $\Diamond^{\star} \cup \Diamond^{\bullet} \neq \Diamond^{\bullet}$, we replace $\Diamond^{\bullet}A \sqsubseteq B$ by the new axiom $(\Diamond^{\star} \cup \Diamond^{\bullet})A \sqsubseteq B$, in order to reflect the validity of both axioms at once. RIAs are handled in the same way. For example, if we know that $\Diamond^{\perp}A \sqsubseteq B$ holds, and have just inferred that $\Diamond^{-}A \sqsubseteq B$ holds as well, then $\Diamond^{+}A \sqsubseteq B$ is a valid entailment, because $\Diamond^{+} \subseteq \Diamond^{\perp} \cup \Diamond^{-}$, and thus whenever an element satisfies $\Diamond^{+}A$, it must satisfy either $\Diamond^{\perp}A$ or $\Diamond^{-}A$. In this way, for any two concepts A, B, the KB always contains at most one axiom $\Diamond^{\star}A \sqsubseteq B$, and similarly for roles.

Let $\mathcal{K}^*$ be the KB obtained by exhaustive application of the completion rules in Figure 4.9 to $\mathcal{K}$, where we assume for technical reasons (see the proof of Lemma 4.11) that A2 and A3 are always applied at the same time for all $i \in \Diamond^{\star}A(a)$ and $i \in \Diamond^{\star}r(a, b)$, respectively.

This process terminates since we only produce axioms of the form $\Diamond^{\star}A \sqsubseteq B$, $\Diamond^{\star}r \sqsubseteq s$, $A(a, i)$, or $r(a, b, i)$, where $\Diamond^{\star}$ was already present in the initial $\mathcal{K}$ or it belongs to $\{\Diamond^{c}_1, \Diamond^{c}, \Diamond^{+}\}$, $i \in \mathrm{rep}(\mathcal{A})$, and A, B, r, s, a, b are from $\mathcal{K}$; there are only polynomially many such axioms.

To decide whether a concept assertion $D(a, i)$ follows from $\mathcal{K}$, we then simply look up whether $D(a, |i|)$ belongs to $\mathcal{K}^*$. For a concept inclusion $\Diamond^{\star}C \sqsubseteq D$ with $C, D \in N_C$, we check whether $\mathcal{K}^*$ contains an inclusion of the form $\Diamond^{\bullet}C \sqsubseteq D$ with $\Diamond^{\star} \subseteq \Diamond^{\bullet}$, which can be done in polynomial time (see Lemma 4.3). One can also check entailment of role axioms in a similar way, but we omit them here for brevity.

Proposition 4.10. $\mathcal{K}$ is inconsistent iff $\bot(a, i) \in \mathcal{K}^*$ for some $a \in N_I$ and $i \in \mathrm{rep}(\mathcal{A})$.

Now let $\mathcal{K}$ be consistent, C be a $\mathcal{TELH}_\bot^{\Diamond^{c},\mathsf{lhs}}$ concept, D be an $\mathcal{ELH}_\bot$ concept, and $\Diamond^{\star} \in \mathfrak{D}$. Then $\mathcal{K} \models \Diamond^{\star}C \sqsubseteq D$ iff either there is $\Diamond^{\dagger} \in \mathfrak{D}$ with $\Diamond^{\dagger}C \sqsubseteq \bot \in \mathcal{K}^*$, or there is $\Diamond^{\bullet} \supseteq \Diamond^{\star}$ with $\Diamond^{\bullet}C \sqsubseteq D \in \mathcal{K}^*$. Moreover, $\mathcal{K} \models D(a, i)$ iff $D(a, |i|) \in \mathcal{K}^*$. $\Diamond$

Before we prove Lemma 4.10, we first show some auxiliary properties of the set $\mathrm{rep}(\mathcal{A})$, which we formulate here only for concept assertions, but hold in the same way for role assertions. We use the following abbreviations, for $i \in \mathbb{Z}$ and $M \subseteq \mathbb{Z}$.

$$i^\uparrow := \{j \in \mathbb{Z} \mid |j| = i\}$$
$$M^\uparrow := \{j \in \mathbb{Z} \mid |j| \in M\}$$

The set $M^\uparrow$ extends M by all time points i represented by any $|i| \in M$.

Intuitively, the next lemma says that everything that holds between two adjacent elements $i < j$ of $\mathrm{tem}(\mathcal{A})$ must also hold for i and j.

Lemma 4.11. For all $B \in N_C$, $a \in N_I$, and $i \in \mathbb{Z}$, if $|i| \in B(a)$, if $-\infty < \lfloor i \rfloor$, then $\lfloor i \rfloor \in B(a)$, and, if $\lceil i \rceil < \infty$, then also $\lceil i \rceil \in B(a)$.

Proof. We show that this property remains satisfied throughout the completion process. In the beginning, this is trivial, because for all assertions $B(a,i)$ we have $\lfloor i \rfloor = \lceil i \rceil = |i| = i \in \mathrm{tem}(\mathcal{A})$. It remains to show that this property is satisfied whenever A2 is applied (the arguments for A3 are similar, and the arguments for A4 and A5 are simpler, because they only refer to one time point).

Let $|i| \in \Diamond^{\star} A(a)$ and $\Diamond^{\star} A \sqsubseteq B \in \mathcal{K}$, requiring us to add $B(a,|i|)$ to $\mathcal{K}$. If $i \in \mathrm{tem}(\mathcal{A})$, then the claim is trivial. If $i \notin \mathrm{tem}(\mathcal{A})$, then we need to show that also $B(a,\lfloor i \rfloor)$ and $B(a,\lceil i \rceil)$ are added to $\mathcal{K}$, i.e. that $\lfloor i \rfloor, \lceil i \rceil \in \Diamond^{\star} A(a)$. Recall that we assumed that A2 and A3 are always applied at the same time to all time points in $\Diamond^{\star} A(a)$ and $\Diamond^{\star} r(a,b)$, respectively. We make a case distinction on the form of $\Diamond^{\star}$.

- If $\Diamond^{\star} = \Diamond^{+}$, then there is $j \in A(a)$ with $j \geq |i|$. If $j = |i|$, then by our assumption we must also have $\lceil i \rceil \in A(a)$, and hence $\lfloor i \rfloor, \lceil i \rceil \in \Diamond^{+} A(a)$. If $j > |i|$, then $j \geq \lceil i \rceil$, which also yields the claim.
- If $\Diamond^{\star} = \Diamond^{c}_{n}$, then there are $j,k \in A(a)$ with $j \leq |i| \leq k$ and $k - j < n$. If the interval $[j,k]$ does not include $\lfloor i \rfloor$, then by our assumption we have $\lfloor i \rfloor \in A(a) \subseteq \Diamond^{c}_{n} A(a)$, and similarly for $\lceil i \rceil$. Otherwise, $\lfloor i \rfloor, \lceil i \rceil \in [j,k] \subseteq \Diamond^{\star} A(a)$.
- The other cases are similar. □

The next lemma shows that using $(\Diamond^{\star} A(a)) \cap \mathrm{rep}(\mathcal{A})$ as a representative for $\Diamond^{\star}(A(a))^{\uparrow}$ in A2 is correct, because expanding it via $\cdot^{\uparrow}$ yields the same result.

Lemma 4.12. If $\Diamond^{\star} \in \mathfrak{D}$ and $M = A(a)$ for $A \in N_C$ and $a \in N_I$, then $\Diamond^{\star} M^{\uparrow} = (\Diamond^{\star} M)^{\uparrow}$.

Proof. We show that $i \in \Diamond^{\star} M^{\uparrow}$ iff $|i| \in \Diamond^{\star} M$, by case distinction on the form of $\Diamond^{\star}$.

- $\Diamond^{\star} = \Diamond^{+}$: If $i \in \Diamond^{+} M^{\uparrow}$, then there is $j \geq i$ with $|j| \in M$, and thus $|j| \geq |i|$ and $|i| \in \Diamond^{+} M$. Conversely, if $|i| \in \Diamond^{+} M$, then there is $j \geq |i|$ with $|j| = j \in M$ since $M \subseteq \mathrm{rep}(\mathcal{A})$. If $|j| > |i|$, then $j > i$, and thus $i \in \Diamond^{+} M^{\uparrow}$. If $|j| = |i|$, then $i \in M^{\uparrow} \subseteq \Diamond^{+} M^{\uparrow}$ by Lemma 4.4.
- $\Diamond^{\star} = \Diamond^{c}_{n}$: If $i \in \Diamond^{c}_{n} M^{\uparrow}$, then there are $j,k \in \mathbb{N}$ with $j \leq i \leq k$, $k - j < n$, and $|j|, |k| \in M$. Thus, $|j| \leq |i| \leq |k|$. If $|i| = |j|$ or $|i| = |k|$, then $|i| \in M \subseteq \Diamond^{c}_{n} M$. Otherwise, we replace $|k|$ by $\lfloor k \rfloor$, and get $|i| \leq \lfloor k \rfloor$ and $\lfloor k \rfloor \in M$ by Lemma 4.11. Similarly, we replace $|j|$ by $\lceil j \rceil \in M$. Then we have $\lceil j \rceil \leq |i| \leq \lfloor k \rfloor$ with $\lfloor k \rfloor - \lceil j \rceil \leq k - j < n$, and thus $|i| \in \Diamond^{c}_{n} M$.
  If $|i| \in \Diamond^{c}_{n} M$, there are $j,k \in M$ with $|j| = j \leq |i| \leq k = |k|$ and $k - j < n$. If $|i| = |j|$ or $|i| = |k|$, then $i \in M^{\uparrow} \subseteq \Diamond^{c}_{n} M^{\uparrow}$. Otherwise, $j < i < k$, and thus $i \in \Diamond^{c}_{n} M^{\uparrow}$.
- The other cases are similar. □

We show soundness and completeness of Proposition 4.10 separately.

Lemma 4.13 (Soundness). If $\Diamond^{\bullet} C \sqsubseteq D \in \mathcal{K}^*$ and $\Diamond^{\star} \subseteq \Diamond^{\bullet}$, then $\mathcal{K} \models \Diamond^{\star} C \sqsubseteq D$. If $D(a,|i|) \in \mathcal{K}^*$, then $\mathcal{K} \models D(a,i)$.

Proof. If $\mathcal{K}$ is inconsistent, then it entails everything. Hence, we can assume that $\mathcal{K}$ is consistent. It suffices to prove that the following holds throughout the completion process: there is a model $\mathfrak{I} = (\Delta^{\mathfrak{I}}, (\mathcal{I}_i)_{i\in\mathbb{Z}})$ of $\mathcal{K}$ such that $D(a,i) \in \mathcal{K}$ implies $a \in D^{\mathcal{I}_j}$, for all $i \in \mathbb{Z}$, $j \in i^{\uparrow}$, $D \in N_C$, and $a \in N_I$, and similarly for role assertions. This is satisfied for all initial assertions $A(a,i) \in \mathcal{K}$ since $i \in \mathrm{tem}(\mathcal{A})$, and thus $i^{\uparrow} = \{i\}$.

We only discuss T8′ and A2, for the other rules one can use similar arguments.

For T8′, assume that $\Diamond_{\star} A \sqsubseteq \exists r.A_1$, $\Diamond_{\bullet} r \sqsubseteq s$, $\Diamond_{\dagger} A_1 \sqsubseteq B_1$, and $\exists s.B_1 \sqsubseteq B$ are satisfied by $\mathfrak{I}$ with $(\Diamond_{\bullet} \cap \Diamond_{\dagger}) \in \mathfrak{D}^{\pm}$, and consider any $d \in (\Diamond_{\ddagger} A)^{\mathcal{I}_i}$, where $\Diamond_{\ddagger} := ((\Diamond_{\bullet} \cap \Diamond_{\dagger}) \circ \Diamond_{\star})$. Then $i \in \Diamond_{\ddagger} M$, where $M := \{j \mid d \in A^{\mathcal{I}_j}\}$. For every $\ell \in \Diamond_{\star} M$, we get $d \in (\exists r.A_1)^{\mathcal{I}_\ell}$ since $\mathfrak{I} \models \Diamond_{\star} A \sqsubseteq \exists r.A_1$. Hence, there is an element $e_\ell \in \Delta^{\mathfrak{I}}$ with $(d, e_\ell) \in r^{\mathcal{I}_\ell}$ and $e_\ell \in A_1^{\mathcal{I}_\ell}$. Thus, $(d, e_\ell) \in (\Diamond_{\bullet} r)^{\mathcal{I}_j} \subseteq s^{\mathcal{I}_j}$ for all $j \in \Diamond_{\bullet}\{\ell\}$ and $e_\ell \in (\Diamond_{\dagger} A_1)^{\mathcal{I}_k} \subseteq B_1^{\mathcal{I}_k}$ for all $k \in \Diamond_{\dagger}\{\ell\}$. For every $k \in (\Diamond_{\bullet} \cap \Diamond_{\dagger})\{\ell\}$, we thus have $d \in (\exists s.B_1)^{\mathcal{I}_k} \subseteq B^{\mathcal{I}_k}$. Due to the fact that $(\Diamond_{\bullet} \cap \Diamond_{\dagger}) \in \mathfrak{D}^{\pm}$ and Lemma 4.5, we obtain $i \in \Diamond_{\ddagger} M = ((\Diamond_{\bullet} \cap \Diamond_{\dagger}) \circ \Diamond_{\star}) M = \bigcup_{\ell \in \Diamond_{\star} M} (\Diamond_{\bullet} \cap \Diamond_{\dagger})\{\ell\}$, thus $d \in B^{\mathcal{I}_i}$.

For A2, let $i \in \Diamond_{\star} A(a)$ and $\mathfrak{I} \models \Diamond_{\star} A \sqsubseteq B$. For $M = A(a)$, $M^{\uparrow} \subseteq \{j \in \mathbb{Z} \mid a \in A^{\mathcal{I}_j}\}$ by induction. Hence, by Lemmas 4.4 and 4.12, we have $a \in (\Diamond_{\star} A)^{\mathcal{I}_j} \subseteq B^{\mathcal{I}_j}$ for all $j \in i^{\uparrow}$, and thus we can safely add $B(a,i)$ to $\mathcal{K}$. $\square$

From this, it follows that $\bot(a,i) \in \mathcal{K}^*$ implies inconsistency of $\mathcal{K}$, and $\Diamond_{\dagger} C \sqsubseteq \bot \in \mathcal{K}^*$ implies $\mathcal{K} \models C \sqsubseteq \Diamond_{\dagger} C \equiv \bot$, and hence $\mathcal{K} \models \Diamond_{\star} C \equiv \bot \sqsubseteq D$. We now prove the remaining direction of Lemma 4.10.

Lemma 4.14 (Completeness). If $\mathcal{K}$ is inconsistent, then $\bot(a,i) \in \mathcal{K}^*$ for some $a \in N_I$ and $i \in \mathrm{tem}(\mathcal{A})$. If $\mathcal{K}$ is consistent and $\mathcal{K} \models \Diamond_{\star} C \sqsubseteq D$, then either $\Diamond_{\dagger} C \sqsubseteq \bot \in \mathcal{K}^*$ or $\Diamond_{\bullet} C \sqsubseteq D \in \mathcal{K}^*$ with $\Diamond_{\star} \subseteq \Diamond_{\bullet}$. If $\mathcal{K}$ is consistent and $\mathcal{K} \models D(a,i)$, then $D(a,|i|) \in \mathcal{K}^*$.

Proof. Assume that $\mathcal{K}^*$ does not contain assertions of the form $\bot(a,i)$. We construct a model $\mathfrak{I} = (\Delta^{\mathfrak{I}}, (\mathcal{I}_i)_{i\in\mathbb{Z}})$ of $\mathcal{K}$ s.t.

1. if there is no $\Diamond_{\dagger} C \sqsubseteq \bot \in \mathcal{K}^*$ or $\Diamond_{\bullet} C \sqsubseteq D \in \mathcal{K}^*$, $\Diamond_{\star} \subseteq \Diamond_{\bullet}$, then there are $i \in \mathbb{Z}$ and $d \in (\Diamond_{\star} C)^{\mathcal{I}_i}$ with $d \notin D^{\mathcal{I}_i}$, and
2. if $D(a,|i|) \notin \mathcal{K}^*$ then $a \notin D^{\mathcal{I}_i}$.

Let ${N_C}^+ := \{A \in N_C \mid \Diamond_{\dagger} A \sqsubseteq \bot \notin \mathcal{K}^*\}$. We define

$$
\begin{aligned}
\Delta^{\mathfrak{I}} &:= \big({N_C}^+ \times \mathbb{Z} \times \mathbb{Z}\big) \cup N_I, \\
B^{\mathcal{I}_i} &:= \big\{a \mid B(a,|i|) \in \mathcal{K}^*\big\} \\
&\quad\ \cup \big\{(A,j,k) \mid \Diamond_{\dagger} A \sqsubseteq B \in \mathcal{K}^*,\ i \in \Diamond_{\dagger}\{j,k\}\big\}, \\
r^{\mathcal{I}_i} &:= \big\{(a,b) \mid r(a,b,|i|) \in \mathcal{K}^*\big\} \\
&\quad\ \cup \big\{(a,(B,\ell,\ell)) \mid \Diamond_{\dagger} A \sqsubseteq \exists s.B \in \mathcal{K}^*, |\ell| \in \Diamond_{\dagger} A(a), \Diamond_{\ddagger} s \sqsubseteq r \in \mathcal{K}^*, i \in \Diamond_{\ddagger}\{\ell\}\big\} \\
&\quad\ \cup \big\{((A,j,k),(B,\ell,\ell)) \mid \Diamond_{\dagger} A \sqsubseteq \exists s.B \in \mathcal{K}^*, \ell \in \Diamond_{\dagger}\{j,k\}, \Diamond_{\ddagger} s \sqsubseteq r \in \mathcal{K}^*, i \in \Diamond_{\ddagger}\{\ell\}\big\}.
\end{aligned}
$$

Since ${N_C}^+ \times \mathbb{Z} \times \mathbb{Z} = \big\{(A,j,k) \mid \Diamond_{\dagger} A \sqsubseteq \top \in \mathcal{K}^*,\ i \in \Diamond_{\dagger}\{j,k\}\big\}$ and $N_I = \big\{a \mid \top(a,|i|) \in \mathcal{K}^*\big\}$ due to T2 and A1, in the following we can treat $\top$ like an ordinary concept name. The same holds for $\bot$ since $\mathcal{K}^*$ contains no assertions of the form $\bot(a,i)$ and the unnamed domain elements are restricted to ${N_C}^+$.

For any $a \in N_I$ and $A \in N_C$, let $M := A(a)$. Then we have $M^{\uparrow} = \{i \in \mathbb{Z} \mid a \in A^{\mathcal{I}_j}\}$, and therefore Lemma 4.12 yields that

$$a \in (\Diamond_{\star} A)^{\mathcal{I}_i} \text{ implies } |i| \in \Diamond_{\star} A(a) \tag{!}$$

for all $i \in \mathbb{Z}$ and $\Diamond^{\star} \in \mathfrak{D}$ (and similarly for role assertions).

We can now prove the claims. Property 2 holds by the definition of $\mathfrak{I}$. To verify Property 1, assume that there is no $\Diamond^{\dagger} C \sqsubseteq \bot \in \mathcal{K}^*$ or $\Diamond^{\bullet} C \sqsubseteq D \in \mathcal{K}^*$ with $\Diamond^{\star} \subseteq \Diamond^{\bullet}$. To show that $\mathfrak{I} \not\models \Diamond^{\star} C \sqsubseteq D$, we make a case distinction on the form of $\Diamond^{\star}$.

- If $\Diamond^{\star} = \Diamond^{+}$, then the rules cannot derive both $\Diamond^{-} C \sqsubseteq D$ and $\Diamond^{+} C \sqsubseteq D$, since otherwise $\Diamond^{+} C \sqsubseteq D \in \mathcal{K}^*$. Assume w.l.o.g. that $\Diamond^{-} C \sqsubseteq D \notin \mathcal{K}^*$. Then $(C, 0, 0) \in C^{\mathcal{I}_0}$ due to $\mathsf{T1}$ and Lemma 4.4, and thus $(C, 0, 0) \in (\Diamond^{+} C)^{\mathcal{I}_1}$, but $1 \notin \Diamond^{\bullet}\{0\}$ for any operator $\Diamond^{\bullet}$ with $\Diamond^{\bullet} C \sqsubseteq D \in \mathcal{K}^*$ ($\Diamond^{\bullet}$ cannot be $\Diamond^{-}$). Hence, by the construction of $\mathfrak{I}$ we have $(C, 0, 0) \notin D^{\mathcal{I}_1}$.

- If $\Diamond^{\star} = \Diamond^{-}$, then we cannot have $\Diamond^{+} C \sqsubseteq D \in \mathcal{K}^*$, but the strongest possible axiom is $\Diamond^{+} C \sqsubseteq D \in \mathcal{K}^*$. We can again use $(C, 0, 0)$ as a counterexample to refute $\mathfrak{I} \models \Diamond^{-} C \sqsubseteq D$.

- If $\Diamond^{\star} = \Diamond^{c}_{n}$, then $\mathcal{K}^*$ may only contain $\Diamond^{c}_{n-1} C \sqsubseteq D$. We have $(C, 0, n) \in (\Diamond^{c}_{n} C)^{\mathcal{I}_1}$, but $1 \notin \Diamond^{c}_{n-1}\{0, n\}$, and thus $(C, 0, n) \notin D^{\mathcal{I}_1}$.

- The other cases are similar.

We now show $\mathfrak{I} \models \mathcal{K}^*$, which implies that $\mathfrak{I} \models \mathcal{K}$. All assertions are satisfied by the definition of $\mathfrak{I}$.

- Consider a GCI $\Diamond^{\bullet} A \sqsubseteq B \in \mathcal{K}^*$. For all $(A', j, k) \in (\Diamond^{\bullet} A)^{\mathcal{I}_i}$, we have $\Diamond^{\star} A' \sqsubseteq A \in \mathcal{K}^*$ and $i \in \Diamond^{\bullet}\Diamond^{\star}\{j, k\}$. Since $\mathsf{T4}$ is not applicable to $\mathcal{K}^*$, we have $\Diamond^{\dagger} A' \sqsubseteq B \in \mathcal{K}^*$ with $(\Diamond^{\bullet} \circ \Diamond^{\star}) \subseteq \Diamond^{\dagger}$. Hence, $i \in \Diamond^{\dagger}\{j, k\}$, and thus $(A', j, k) \in B^{\mathcal{I}_i}$.
  For every $a \in (\Diamond^{\bullet} A)^{\mathcal{I}_i}$, by (!) we have $|i| \in \Diamond^{\bullet} A(a)$, and hence by $\mathsf{A2}$ we must have $B(a, |i|) \in \mathcal{K}^*$, i.e. $a \in B^{\mathcal{I}_i}$.

- Let $\Diamond^{\bullet} A \sqsubseteq \exists r.B \in \mathcal{K}^*$. For all $(A', j, k) \in (\Diamond^{\bullet} A)^{\mathcal{I}_i}$, there is $\Diamond^{\star} A' \sqsubseteq A \in \mathcal{K}^*$ with $i \in \Diamond^{\bullet}\Diamond^{\star}\{j, k\}$. By $\mathsf{T3}$ and $\mathsf{T4}$, there are $\Diamond^{c}_{1} r \sqsubseteq r, \Diamond^{\dagger} A' \sqsubseteq \exists r.B \in \mathcal{K}^*$ with $(\Diamond^{\bullet} \circ \Diamond^{\star}) \subseteq \Diamond^{\dagger}$ and $i \in \Diamond^{\dagger}\{j, k\}$. Since $i \in \Diamond^{c}_{1}\{i\}$, we have $\big((A', j, k), (B, i, i)\big) \in r^{\mathcal{I}_i}$. Note that $B \in {N_C}^{+}$ since any $\Diamond^{\ddagger} B \sqsubseteq \bot \in \mathcal{K}^*$ would yield $A' \notin {N_C}^{+}$ by $\mathsf{T7}$ and $\mathsf{T8}$. Moreover, by $\mathsf{T1}$, $(B, i, i) \in B^{\mathcal{I}_i}$, and hence $(A', j, k) \in (\exists r.B)^{\mathcal{I}_i}$.
  For all $a \in (\Diamond^{\bullet} A)^{\mathcal{I}_i}$, we have $|i| \in \Diamond^{\bullet} A(a)$ by (!). By $\mathsf{T3}$, we obtain $\Diamond^{c}_{1} r \sqsubseteq r \in \mathcal{K}^*$. Since $i \in \Diamond^{c}_{1}\{i\}$, this implies that $(a, (B, i, i)) \in r^{\mathcal{I}_i}$. Note that $B \in {N_C}^{+}$ since otherwise $A \notin {N_C}^{+}$, and thus $\bot(a, j) \in \mathcal{K}^*$ for some $j \in \mathbb{Z}$ with $a \in A^{\mathcal{I}_j}$, which contradicts our assumption. By $\mathsf{T1}$, it holds that $(B, i, i) \in B^{\mathcal{I}_i}$, and we conclude that $a \in (\exists r.B)^{\mathcal{I}_i}$.

- Consider $\exists r.A \sqsubseteq B \in \mathcal{K}^*$. For all $(A', j, k) \in (\exists r.A)^{\mathcal{I}_i}$, there exists (B', ℓ, ℓ) such that $\big((A', j, k), (B', \ell, \ell)\big) \in r^{\mathcal{I}_i}$ and $(B', \ell, \ell) \in A^{\mathcal{I}_i}$. Thus, there are $\Diamond^{\star} A' \sqsubseteq \exists s.B', \Diamond^{\bullet} s \sqsubseteq r, \Diamond^{\dagger} B' \sqsubseteq A \in \mathcal{K}^*$ with $\ell \in \Diamond^{\star}\{j, k\}$ and $i \in \Diamond^{\bullet}\{\ell\} \cap \Diamond^{\dagger}\{\ell\}$.
  - If $(\Diamond^{\bullet} \cap \Diamond^{\dagger}) \in \mathfrak{D}^{\mathsf{c}}$, then $(\Diamond^{\bullet} \cap \Diamond^{\dagger})\{\ell\} = \{\ell\}$ by Lemma 4.5, and thus $i = \ell$. By $\mathsf{T8}$, there is $\Diamond^{\ddagger} A' \sqsubseteq B \in \mathcal{K}^*$ with $\Diamond^{\star} \subseteq \Diamond^{\ddagger}$, and hence $i = \ell \in \Diamond^{\star}\{j, k\} \subseteq \Diamond^{\ddagger}\{j, k\}$, which shows that $(A', j, k) \in B^{\mathcal{I}_i}$.
  - If $(\Diamond^{\bullet} \cap \Diamond^{\dagger}) \in \mathfrak{D}^{\pm}$, then by $\mathsf{T8}'$ there is $\Diamond^{\ddagger} A' \sqsubseteq B \in \mathcal{K}^*$ with $((\Diamond^{\bullet} \cap \Diamond^{\dagger}) \circ \Diamond^{\star}) \subseteq \Diamond^{\ddagger}$. Together with Lemma 4.4, this yields $i \in (\Diamond^{\bullet} \cap \Diamond^{\dagger})\{\ell\} \subseteq (\Diamond^{\bullet} \cap \Diamond^{\dagger})\Diamond^{\star}\{j, k\} \subseteq \Diamond^{\ddagger}\{j, k\}$, which again shows that $(A', j, k) \in B^{\mathcal{I}_i}$.

For all $a \in (\exists r.A)^{\mathcal{I}_i}$, there exists $e \in A^{\mathcal{I}_i}$ with $(a, e) \in r^{\mathcal{I}_i}$.

- If e is of the form (B', ℓ, ℓ), then we proceed as above, using $\mathsf{T8}$ or $\mathsf{T8}'$ to get $\Diamond^{\ddagger} A' \sqsubseteq B \in \mathcal{K}^*$ with $|i| \in \Diamond^{\ddagger} A'(a)$. The only differences are that we have $|\ell| \in \Diamond^{\star} A'(a)$ instead of $\ell \in \Diamond^{\star}\{j, k\}$, and that we need to infer $|i| \in (\Diamond^{\bullet} \cap \Diamond^{\dagger})\{|\ell|\}$ from $i \in (\Diamond^{\bullet} \cap \Diamond^{\dagger})\{\ell\}$ in case that $(\Diamond^{\bullet} \cap \Diamond^{\dagger}) \in \mathfrak{D}^{\pm}$, which we can do by similar arguments as in Lemma 4.12. By $\mathsf{A2}$, we then obtain $B(a, |i|) \in \mathcal{K}^*$, and thus $a \in B^{\mathcal{I}_i}$.
- If $e \in N_I$, then $A(e, |i|), r(a, e, |i|) \in \mathcal{K}^*$, and thus by $\mathsf{A5}$ we have $B(a, |i|) \in \mathcal{K}^*$, and hence $a \in B^{\mathcal{I}_i}$.

- The other cases are similar. □

In the proof, we extend the standard construction of a universal model that satisfies exactly the axioms entailed by $\mathcal{K}$. To deal with $\Diamond^{\mathsf{c}}_n$, we add new domain elements of the form (A, i, j) with $|j - i| = n - 1$, which satisfy A at the time points i and j. These elements can be used to distinguish the behavior of $\Diamond^{\mathsf{c}}_n$ from that of $\Diamond^{\mathsf{c}}_{n-1}$.

We obtain the following result, where the lower bound follows from propositional Horn logic [JL76].

Theorem 4.15. Entailment in $\mathcal{TELH}_\bot^{\Diamond^{\mathsf{c}},\mathsf{lhs}}$ is PTIME-complete.

We now give a small example on how $\mathcal{TELH}_\bot^{\Diamond^{\mathsf{c}},\mathsf{lhs}}$ can be used to model temporal behavior in the medical domain.

Example 4.16. Consider *rheumatoid arthritis*, an autoimmune disorder that cannot be healed. In irregular intervals, it produces so called *flare ups*, that cause pain in the joints. We formalize this knowledge as follows:

$$\mathtt{RheumatoidArthritisPatient} \equiv \exists\mathtt{diagnosedWith}.\mathtt{RheumatoidArthritis} \tag{4.3}$$

$$\mathtt{FlareUpPatient} \sqsubseteq \mathtt{RheumatoidArthritisPatient} \tag{4.4}$$

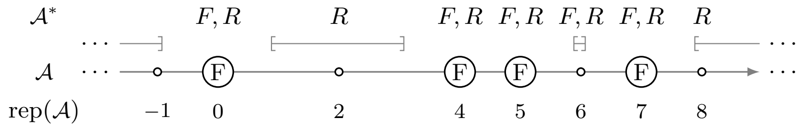
$$\diamondminus\mathtt{RheumatoidArthritisPatient} \sqsubseteq \mathtt{RheumatoidArthritisPatient} \tag{4.5}$$

$$\Diamond^{\mathsf{c}}_2\mathtt{FlareUpPatient} \sqsubseteq \mathtt{FlareUpPatient} \tag{4.6}$$

We make the assumption that a flare up is 2-month convex, hence if two flare ups are reported at most 2 months apart, we assume that they refer to the same flare up and hence the flare up also present in between the two reports. By applying Rule $\mathsf{T4}$ from the completion algorithm to axioms (4.4) and (4.5), we can add

$$\diamondminus\mathtt{FlareUpPatient} \sqsubseteq \mathtt{RheumatoidArthritisPatient}$$

to the KB. Suppose the ABox consists of the assertions $\mathtt{FlareUpPatient}(p_1, i)$, $i \in \{0, 4, 5, 7\}$, for a patient p_1. The completed ABox, denoted by $\mathcal{A}^*$, is illustrated in Figure 4.17, where for simplicity we omit the individual name p_1.



$\diamondsuit$

Figure 4.17.: An illustration of the $\mathcal{A}^*$ from Example 4.16. **`RheumatoidArthritisPatient`** and **`FlareUpPatient`** are abbreviated by their first letters, respectively. Representatives $-1, 2, 6$ and 8 have been introduced and the intervals they represent are illustrated in gray.

4.4. Discussion and Related Work

We have shown that with the current set of diamond operators reasoning in $\mathcal{TELH}_\bot^{\diamondsuit,\mathsf{lhs}}$ remains tractable. It seems possible to allow further diamond operators in $\mathcal{TELH}_\bot^{\diamondsuit,\mathsf{lhs}}$ axioms if they satisfy the relevant properties, i.e. $\mathfrak{D}$ remains closed under $\cap$, $\cup$, and $\circ$ and compatible with $\cdot^\uparrow$ (see Lemmas 4.3 and 4.6). For future research it would be interesting to identify such operators.

For a general overview of temporal ontology and query languages, see [LWZ08; AKK+17]. In the presence of a single rigid role, allowing the operator $\diamondplus$ on both sides of $\mathcal{EL}$ GCIs makes subsumption undecidable [AKL+07]. In [GJK16], a variety of restrictions is investigated to regain decidability, such as acyclic TBoxes and restrictions on the occurrences of temporal operators. In particular, allowing the qualitative operators $\diamondplus, \diamondminus, \diamondplus, \diamonddot$ only on the left-hand side of GCIs makes the logic tractable. However, these investigations do not cover metric temporal operators and employ the assumption that all timestamps are encoded in unary.

Similarly, combinations of $\mathcal{EL}$ with the branching temporal logic CTL have been shown to have undecidable subsumption problems, but they may become tractable under certain restrictions [GJL12; GJS14; GJS15]. Adding LTL operators to concepts was also investigated in other DLs, like $\mathcal{ALC}$ (without temporal roles) [WZ00; LWZ08] and *DL-Lite* [AKL+07]. Only recently, also metric variants of such logics were considered [GJO16a; BBK+17; Tho18].

Chapter 5.

Temporal Minimal-World Semantics for $\mathcal{ELH}_\bot$

In this chapter we extend minimal-world semantics to $\mathcal{TELH}_\bot^{\Diamond,\mathsf{lhs},-}$, a variant of $\mathcal{TELH}_\bot^{\Diamond,\mathsf{lhs}}$ without temporal roles. Our query language, which we introduce in Section 5.1, extends the temporal conjunctive queries from [BBL15b] by metric temporal operators and negation. For example, we can use queries like

$$\Box_{[-12,0]}(\exists y.\mathtt{diagnosedWith}(x, y) \wedge \mathtt{Diabetes}(y))$$

to query patients that were suffering from **Diabetes** during the last twelve months. In Section 5.2 we lift the construction of the minimal universal model from $\mathcal{ELH}_\bot$ to the temporal setting in $\mathcal{TELH}_\bot^{\Diamond,\mathsf{lhs},-}$ and show that also in this case such a model exists and is uniquely defined. Using a combined rewriting approach, we show in Section 5.3 that the data complexity of query answering is not higher than for positive atemporal queries in $\mathcal{ELH}_\bot$, and also provide a tight combined complexity result of ExpSpace. In the previous chapter we already mentioned that, unlike most research on temporal query answering [BBL15b; AKK+15], we do not assume that input data is given for all time points in a certain interval, but rather at sporadic time points with arbitrarily large gaps. The main technical difficulty is to determine which additional time points are relevant for answering a query, and how to access these time points without having to fill all the gaps. Currently, we do not support temporal roles, but conjecture in Section 5.4 that it is possible to remove this restriction in future work. We end the chapter with a dicussion of related work.

5.1. Metric Temporal Conjunctive Queries with Negation

Definition 5.1. *Metric temporal conjunctive queries with negation (MTNCQs)* are built by the grammar rule

$$\phi ::= \psi \mid \top \mid \bot \mid \neg\phi \mid \phi \wedge \phi \mid \phi \vee \phi \mid \phi\,\mathcal{U}_I\,\phi \mid \phi\,\mathcal{S}_I\,\phi, \tag{5.1}$$

where ψ is an NCQ, and I is an interval over $\mathbb{Z}$. An MTNCQ ϕ is *rooted*/*Boolean* if all NCQs in it are rooted/Boolean. ♢

MTNCQs can be seen as NCQ-LTL$^{\mathsf{bin}}$, and hence we employ the standard semantics as defined in Definition 2.26 in Section 2.3.

Recall that the *next* operator can be defined as $\bigcirc\phi := \top\,\mathcal{U}_{[1,1]}\phi$, and similarly $\bigcirc^{-}\phi := \top\,\mathcal{S}_{[1,1]}\phi$. We can also express

$$
\begin{aligned}
&\Diamond_I\phi := (\top\,\mathcal{S}_{-(I\cap(-\infty,0])}\phi) \vee (\top\,\mathcal{U}_{I\cap[0,\infty)}\phi) \text{ and}\\
&\Box_I\phi := \neg\Diamond_I\neg\phi,
\end{aligned}
$$

and hence, by (4.1), the $\Diamonddot_n$-operators from Section 4.1.

An *MTCQ* (or *positive MTNCQ*) is an MTNCQ without negation, where we assume that the operator $\Box_I$ is nevertheless included as part of the syntax of MTCQs.

Example 5.2. Consider the criterion 'Diagnosis of Rheumatoid Arthritis (RA) of more than 6 months and less than 15 years.'[1] This can be expressed as an MTNCQ as follows.

$$
\begin{aligned}
\phi(x) := {}&\Box_{[-6,0]}(\exists y.\mathtt{diagnosedWith}(x,y) \wedge \mathtt{RheumatoidArthritis}(y))\\
&\wedge \neg\Box_{[-180,0]}(\exists y.\mathtt{diagnosedWith}(x,y) \wedge \mathtt{RheumatoidArthritis}(y))
\end{aligned}
$$

The negated conjunct merely expresses that the data does not indicate an ongoing rheumatoid arthritis diagnosis for the past 15 years (minimal-world semantics), rather than that such a diagnosis is categorically ruled out by some TBox axioms (certain answer semantics). $\Diamond$

Definition 5.3. Let $\mathcal{K} = (\mathcal{T},\mathcal{A})$ be a $\mathcal{TELH}_\bot^{\Diamonddot,\mathsf{lhs}}$-KB, $\phi(\mathbf{x})$ an MTNCQ, $\mathbf{a}$ a tuple of individual names from $\mathcal{A}$, $i \in \mathrm{tem}(\mathcal{A})$, and $\mathfrak{I}$ a $\mathcal{TELH}_\bot^{\Diamonddot,\mathsf{lhs}}$-interpretation.

The pair $(\mathbf{a}, i)$ is an *answer* to $\phi(\mathbf{x})$ w.r.t. $\mathfrak{I}$ if $\mathfrak{I}, i \models \phi(\mathbf{a})$. The set of all answers for ϕ w.r.t. $\mathfrak{I}$ is denoted $\mathrm{ans}(\phi,\mathfrak{I})$.

The tuple $(\mathbf{a}, i)$ is a *certain answer* to ϕ w.r.t. $\mathcal{K}$ if it is an answer in every model of $\mathcal{K}$; all these tuples are collected in the set $\mathrm{cert}(\phi,\mathcal{K})$. $\Diamond$

5.2. Minimal-World Semantics for MTNCQs

In the following let $\mathcal{TELH}_\bot^{\Diamonddot,\mathsf{lhs},-}$ denote the fragment of $\mathcal{TELH}_\bot^{\Diamonddot,\mathsf{lhs}}$ in which no temporal roles are allowed. A discussion about the challenges involved in allowing temporal roles can be found in Section 5.4. In the definition of the model, we make use of entailment in $\mathcal{TELH}_\bot^{\Diamonddot,\mathsf{lhs},-}$, which can be checked in polynomial time as we have shown before (see Theorem 4.15). Thus, we can exclude w.l.o.g. equivalent concept and role names. Also, for simplicity, in the following we assume w.l.o.g. that all GCIs are in the following stronger normal form (cf. (4.2)):

$$
\Diamondblack A \sqsubseteq B,\ A_1 \sqcap A_2 \sqsubseteq B,\ r \sqsubseteq s,\ A \sqsubseteq \exists r.B,\ \exists r.A \sqsubseteq B,
$$

i.e. $\Diamond$-operators are allowed only in GCIs of the form $\Diamondblack A \sqsubseteq B$. In particular, disallowing GCIs of the form $\Diamondblack A \sqsubseteq \exists r.B$ allows us to draw a stronger connection to the construction in Section 3.3; see in particular Step 3(a) in Def. 5.4 below.

Definition 5.4. The *(temporal) canonical model* $\mathfrak{I}_\mathcal{K} = (\Delta^{\mathfrak{I}_\mathcal{K}}, (\mathcal{I}_i)_{i\in\mathbb{Z}})$ of a KB $\mathcal{K} = (\mathcal{T},\mathcal{A})$ is obtained by the following steps.

[1] `https://clinicaltrials.gov/ct2/show/NCT01198002`

1. Set $\Delta^{\mathfrak{I}_\mathcal{K}} := N_I$ and $a^{\mathcal{I}_i} := a$ for all $a \in N_I$ and $i \in \mathbb{Z}$.
2. For each time point $i \in \mathbb{Z}$, define $A^{\mathcal{I}_i} := \{a \mid \mathcal{K} \models A(a,i)\}$ for all $A \in N_C$ and define $r^{\mathcal{I}_i} := \{(a,b) \mid \mathcal{K} \models r(a,b,i)\}$ for all $r \in N_R$.
3. Repeat the following steps:
   a) Select an element $d \in \Delta^{\mathfrak{I}_\mathcal{K}}$ that has not been selected before and, for each $i \in \mathbb{Z}$, let $V_i := \{\exists r.B \mid d \in A^{\mathcal{I}_i},\ d \notin (\exists r.B)^{\mathcal{I}_i},\ A \sqsubseteq_\mathcal{T} \exists r.B,\ A, B \in N_C\}$.
   b) For each $\exists r.B$ that is minimal in some V_i
      i. add a fresh element $e_{rB}^{(i)}$ to $\Delta^{\mathfrak{I}_\mathcal{K}}$,
      ii. add $e_{rB}^{(i)}$ to $A^{\mathcal{I}_k}$ for each $\Diamond_{\star} B \sqsubseteq_\mathcal{T} A$ with $k \in \Diamond_{\star}\{i\}$ and $\Diamond_{\star} \in \mathfrak{D}$, and
      iii. add $(d, e_{rB}^{(i)})$ to $s^{\mathcal{I}_i}$ for each $\mathcal{K} \models r \sqsubseteq s$. $\Diamond$

We denote by $\mathfrak{I}_\mathcal{A}$ the result of executing only Steps 1 and 2 of this definition, i.e. restricting $\mathfrak{I}_\mathcal{K}$ to the named individuals. Since there are only finitely many elements of N_I, N_C, and N_R that are relevant for this definition (i.e. those that occur in $\mathcal{K}$), for simplicity we often treat $\mathfrak{I}_\mathcal{A}$ as if it had a finite object (but still infinite time) domain.

Lemma 5.5. Let $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ be a consistent $\mathcal{TELH}_\bot^{\Diamond,\mathsf{lhs},-}$-KB. Then the interpretation constructed according to Definition 5.4 is a model of $\mathcal{K}$.

Proof. We show that all axioms of $\mathcal{K}$ are satisfied in an interpretation $\mathfrak{I}_\mathcal{K} = (\Delta^{\mathfrak{I}_\mathcal{K}}, (\mathcal{I}_i)_{i \in \mathbb{Z}})$ constructed according to Definition 5.4. It is easy to check that the ABox assertions are satisfied after Step 2 is applied. We make a case distinction for the TBox axioms:

- Suppose a GCI of the form $\Diamond_{\star} A \sqsubseteq B \in \mathcal{T}$ and $i \in \Diamond_{\star} A^{\mathfrak{I}_\mathcal{K}}(d)$ for some $i \in \mathbb{Z}$ and $d \in \Delta^{\mathcal{I}_\mathcal{K}}$, where $A^{\mathfrak{I}_\mathcal{K}}(d) := \{i \in \mathbb{Z} \mid d \in A^{\mathcal{I}_i}\}$.
  If $d \in N_I$, then $i \in \Diamond_{\star}\{j \in \mathbb{Z} \mid \mathcal{K} \models A(d,j)\}$ and since $\mathcal{K} \models \Diamond_{\star} A \sqsubseteq B$ it also has to hold that $\mathcal{K} \models B(d,i)$.
  If $d \notin N_I$, then it was introduced in Step 3 at some point during the construction. Let $\exists r_1.B_1$ be the minimal concept in V_j for some $j \in \mathbb{Z}$ that caused the introduction of d. Let $\Diamond_{\bullet} \in \mathfrak{D}$ be the strongest diamond such that $\Diamond_{\bullet} B_1 \sqsubseteq_\mathcal{T} A$. Such a diamond $\Diamond_{\bullet}$ has to exist, otherwise $\Diamond_{\star} A^{\mathfrak{I}_\mathcal{K}}(d) = \emptyset$, which would be a contradiction. Then by Step 3bii) it holds that $A^{\mathfrak{I}_\mathcal{K}}(d) = \Diamond_{\bullet}\{j\}$. Therefore, have that $i \in \Diamond_{\star} A^{\mathfrak{I}_\mathcal{K}}(d) = (\Diamond_{\star} \circ \Diamond_{\bullet})\{j\}$. Moreover, $\Diamond_{\bullet} B_1 \sqsubseteq_\mathcal{T} A$ and $\Diamond_{\star} A \sqsubseteq_\mathcal{T} B$ imply that $(\Diamond_{\star} \circ \Diamond_{\bullet}) B_1 \sqsubseteq_\mathcal{T} B$ (see Rule T4 in Figure 4.9), hence d was added to $B^{\mathcal{I}_i}$ in Step 3bii).

The remaining cases follow the same argumentation as in the proof of Lemma 3.7. $\square$

In $\mathfrak{I}_\mathcal{K}$ each anonymous individual $e_{rB}^{(i)} \in \Delta^{\mathfrak{I}_\mathcal{K}}$ is connected to the rooted part of $\mathfrak{I}_\mathcal{K}$ only at i. Therefore, the behavior of e is entirely determined by its predecessor and i. In other words, the only place where relevant temporal interaction can occur (without temporal roles) is between named individuals. For this reason, in the following we consider only rooted MTNCQs, which can be evaluated only over the parts of $\mathfrak{I}_\mathcal{K}$ that are connected to the named individuals.

As in the atemporal case, we show that $\mathfrak{I}_\mathcal{K}$ is universal and can therefore be used to answer *positive* queries over $\mathcal{K}$ under certain answer semantics.

Lemma 5.6. Let $\mathcal{K}$ be a consistent $\mathcal{TELH}_\bot^{\Diamond,\mathsf{lhs},-}$-KB. Then for every rooted MTCQ ϕ, we have $\mathrm{cert}(\phi,\mathcal{K}) = \mathrm{ans}(\phi,\mathfrak{I}_\mathcal{K})$.

Proof. The inclusion $\mathrm{cert}(\phi,\mathcal{K}) \subseteq \mathrm{ans}(\phi,\mathfrak{I}_\mathcal{K})$ follows from the fact that $\mathfrak{I}_\mathcal{K}$ is a model of $\mathcal{K}$. For the other inclusion, consider any model $\mathfrak{J} = (\Delta^\mathfrak{J}, (\mathcal{J}_i)_{i\geq 0})$ of $\mathcal{K}$. We prove that $(\mathbf{a}, i) \in \mathrm{ans}(\phi,\mathfrak{I}_\mathcal{K})$ implies $(\mathbf{a}, i) \in \mathrm{ans}(\phi,\mathfrak{J})$ by induction on the structure of ϕ.

- If ϕ is a rooted CQ, then $\mathcal{I}_i \models \phi(\mathbf{a})$. Moreover, since ϕ is rooted, only the rooted part of $\mathcal{I}_i$, consisting of all elements connected to named individuals, is relevant for satisfying $\phi(\mathbf{a})$. It is easy to show that this part can be homomorphically mapped into $\mathcal{J}_i$, hence $\mathfrak{J}, i \models \phi(\mathbf{a})$.
- If $\phi = \phi_1 \vee \phi_2$, then $(\mathbf{a}, i) \in \mathrm{ans}(\phi_1,\mathfrak{I}_\mathcal{K})$ or $(\mathbf{a}, i) \in \mathrm{ans}(\phi_2,\mathfrak{I}_\mathcal{K})$, hence by induction $(\mathbf{a}, i) \in \mathrm{ans}(\phi_1,\mathfrak{J})$ or $(\mathbf{a}, i) \in \mathrm{ans}(\phi_2,\mathfrak{J})$, either of which implies that $(\mathbf{a}, i) \in \mathrm{ans}(\phi,\mathfrak{J})$.
- If $\phi = \phi_1 \mathcal{U}_I \phi_2$, then $(\mathbf{a}, i) \in \mathrm{ans}(\phi_1 \mathcal{U}_I \phi_2,\mathfrak{I}_\mathcal{K})$, hence there exists $k \in I$ such that $(\mathbf{a}, i+k) \in \mathrm{ans}(\phi_2,\mathfrak{I}_\mathcal{K})$ and for all $0 \leq j < k$ it holds that $(\mathbf{a}, i+j) \in \mathrm{ans}(\phi_1,\mathfrak{I}_\mathcal{K})$. By induction this implies that $(\mathbf{a}, i+k) \in \mathrm{ans}(\phi_2,\mathfrak{J})$ and $(\mathbf{a}, i+j) \in \mathrm{ans}(\phi_1,\mathfrak{J})$ for all $0 \leq j < k$ and therefore $(\mathbf{a}, i) \in \mathrm{ans}(\phi_1 \mathcal{U}_I \phi_2,\mathfrak{J})$.
- The cases of $\mathcal{S}_I$, $\Diamond_I$, and $\Box_I$ are similar, and therefore the claim also extends to $\Diamond_n$, $\bigcirc$, and $\bigcirc^-$. □

Additionally, $\mathfrak{I}_\mathcal{K}$ is a minimal model.

Lemma 5.7. $\mathfrak{I}_\mathcal{K}$ is minimal according to Definition 3.3, i.e. every endomorphism on $\mathfrak{I}_\mathcal{K}$ is an isomorphism.

Proof. We prove the stronger statement that the only endomorphism of $\mathfrak{I}_\mathcal{K}$ is identity.

Let $\mathfrak{I}_0, \mathfrak{I}_1, \ldots$ be the interpretations obtained in the construction of $\mathfrak{I}_\mathcal{K}$ by Definition 5.4 before each application of Step 3 with $\mathfrak{I}_i = (\Delta^{\mathfrak{I}_i}, (\mathcal{I}_i^{(k)})_{k\in\mathbb{Z}})$. We show by induction on i that all homomorphisms $h_0, h_1, \ldots$, where h_i is an endomorphism on $\mathfrak{I}_i$, are the identify function. The only endomorphism on $\mathfrak{I}_\mathcal{K}$ is then obtained in the limit as $h = \bigcup_{i\geq 0} h_i$.

By definition of homomorphisms, h_0 is determined to be of the form $h_0(a) := a^\mathfrak{I}$ for all $a \in \mathrm{Ind}(\mathcal{A})$, which is the identify function.

For the induction step, assume that h_i has already been defined. To define h_{i+1}, assume that $d \in \Delta^{\mathfrak{I}_i}$ was picked in Step 3(a) and V_k with $k \in \mathbb{Z}$ are the sets as defined in Step (a). For each $\exists r.B$ that is minimal in some V_k, let A be a concept name that caused $\exists r.B$ to be in V_k, i.e. $A \sqsubseteq_\mathcal{T} \exists r.B$, and let $e_{rB}^{(k)}$ be the freshly introduced domain element.

Since $\exists r.B$ is minimal in V_k and in no further construction step any successors will be added to d, we know that $e_{rB}^{(k)}$ is the only element in the domain of $\Delta^{\mathfrak{I}_\mathcal{K}}$ for which it holds that

1. for all $\Diamond\!\!\!\!\ast B \sqsubseteq_\mathcal{T} C$: $e_{rB}^{(k)} \in (C)^{\mathcal{I}_{i+1}^{(\ell)}}$ where $\Diamond\!\!\!\!\ast \in \mathfrak{D}$ and $\ell \in \Diamond\!\!\!\!\ast\{k\}$ and
2. for all $r \sqsubseteq_\mathcal{T} s$: $(d, e_{rB}^{(k)}) \in s^{\mathcal{I}_{i+1}^{(k)}}$.

Hence, the only possibility is to set $h_{i+1} := h_i \cup \{e_{rB}^{(k)} \mapsto e_{rB}^{(k)}\}$. □

Thus, the following *minimal-world* semantics is compatible with certain answer semantics for positive (rooted) queries, while keeping a tractable data complexity. Since the construction in Definition 5.4 yields the temporal minimal universal model $\mathfrak{I}_\mathcal{K}$ of a given consistent KB $\mathcal{K}$, the minimal-world answers to a given MTNCQ ϕ are obtained as the answers $\mathrm{ans}(\phi, \mathfrak{I}_\mathcal{K})$ to ϕ in $\mathfrak{I}_\mathcal{K}$.

5.3. A Combined Rewriting for MTNCQs

Since the minimal canonical model $\mathfrak{I}_\mathcal{K}$ may still be infinite, we now show that rooted MTNCQ answering under minimal-world semantics is also combined first-order rewritable [LTW09]. We proceed in two steps.

1. We rewrite ϕ into a *metric first-order temporal logic (MFOTL)* formula $\mathrm{rew}_\mathcal{T}(\phi)$, which combines first-order formulas via metric temporal operators; for details, see [BKM+15]. This query can be evaluated over $\mathfrak{I}_\mathcal{A}$ instead of $\mathfrak{I}_\mathcal{K}$. Hence, we reduce the infinite object domain to the finite set $N_I(\mathcal{K})$.
2. We then further rewrite $\mathrm{rew}_\mathcal{T}(\phi)$ into a three-sorted first-order formula (with explicit variables for time points), which is then evaluated over a restriction $\mathfrak{I}_\mathcal{A}^{\mathrm{fin}}$ of $\mathfrak{I}_\mathcal{A}$ that contains only finitely many time points (essentially those in $\mathrm{rep}(\mathcal{A})$, although we modify them slightly).

For the first step, we rewrite a rooted MTNCQ ϕ by replacing each (rooted) NCQ ψ with the first-order rewriting $\mathrm{rew}_\mathcal{T}(\psi)$ from Definition 3.11.[2] The result is a special kind of MFOTL formula $\mathrm{rew}_\mathcal{T}(\phi)$ [BKM+15], in which atemporal first-order formulas can be nested inside MTL-operators, similarly as in MTNCQs. The semantics is based on a satisfaction relation $\mathfrak{I}, i \models \mathrm{rew}_\mathcal{T}(\phi)$ that is defined in much the same way as for MTNCQs, the only exception being that $\mathfrak{I}, i \models \mathrm{rew}_\mathcal{T}(\psi)$ for a first-order formula $\mathrm{rew}_\mathcal{T}(\psi)$ is defined by $\mathcal{I}_i \models \mathrm{rew}_\mathcal{T}(\psi)$, using the standard first-order semantics. We can lift the atemporal rewritability result in a straightforward way to our temporal setting.

Lemma 5.8. Let $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ be a consistent $\mathcal{TELH}_\bot^{\diamondsuit\!\!\!\!\!\!\;\text{c},\mathsf{lhs},-}$-KB and ϕ be a rooted MTNCQ. Then $\mathrm{mwa}(\phi, \mathcal{K}) = \mathrm{ans}(\mathrm{rew}_\mathcal{T}(\phi), \mathfrak{I}_\mathcal{A})$.

Proof. We prove the claim by induction over the structure of ϕ.

Suppose that ϕ is a rooted NCQ. Since ϕ does not contain temporal operators, we can restrict our attention to a single atemporal interpretation $\mathcal{I}_i$ in $\mathfrak{I}_\mathcal{K} = (\Delta^{\mathfrak{I}_\mathcal{K}}, (\mathcal{I}_i)_{i \in \mathbb{Z}})$. Since ϕ is rooted, only the 'rooted' part $\mathcal{I}_i^r$ of $\mathcal{I}_i$, consisting only of those elements connected to named individuals via a sequence of role connections, is relevant for evaluating ϕ (and similarly for $\mathrm{rew}_\mathcal{T}(\phi)$). In the construction of $\mathfrak{I}_\mathcal{K}$ (Definition 5.4), we can observe that $\mathcal{I}_i^r$ is uniquely determined by the definition of $A^{\mathcal{I}_i}$ and $r^{\mathcal{I}_i}$ in Step 2. Moreover, $\mathcal{I}_i^r$ is isomorphic to the (atemporal) minimal canonical model of $(\mathcal{T}', \mathcal{A}_i)$ as defined in Definition 3.6, where

- $\mathcal{T}' := \{C \sqsubseteq D \mid \diamondsuit\!\!\!\!\!\!\;\star C \sqsubseteq D \in \mathcal{T}, \diamondsuit\!\!\!\!\!\!\;\star \in \mathfrak{D}^\pm\}$ and

[2] Strictly speaking, $\mathrm{rew}_\mathcal{T}(\psi)$ is a *set* of first-order formulas, which is however equivalent to the disjunction of all these formulas.

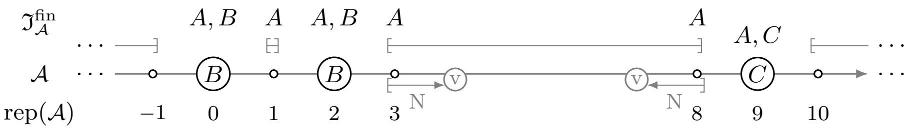
- $\mathcal{A}_i := \{A(a) \mid \mathcal{K} \models A(a,i)\} \cup \{r(a,b) \mid \mathcal{K} \models r(a,b,i)\}$.

In particular, one can observe that the temporal operators in $\mathcal{T}$ are irrelevant for the behavior of the anonymous elements in $\mathcal{I}_i^r$: In Step 3b of the construction an element e_{rB} is only connected to object d (selected in Step 3a) in $\mathcal{I}_i$ and hence belongs to $\mathcal{I}_i^r$, if e_{rB} is minimal in V_i. In that case e_{rB} is entirely determined by d at time point i since $\diamondplus C \sqsubseteq D$ entails $C \sqsubseteq D$. Therefore, we can restrict the attention to those assertions entailed for time point i. Hence, we can apply Lemma 3.14 to conclude that $(\mathbf{a}, i) \in \mathrm{mwa}(\phi, \mathcal{K}) = \mathrm{ans}(\phi, \mathfrak{I}_\mathcal{K})$ iff $\mathbf{a} \in \mathrm{ans}(\phi, \mathcal{I}_i) = \mathrm{ans}(\phi, \mathcal{I}_i^r) = \mathrm{ans}(\mathrm{rew}_\mathcal{T}(\phi), \mathcal{I}_{i,\mathcal{A}})$ iff $(\mathbf{a}, i) \in \mathrm{ans}(\mathrm{rew}_\mathcal{T}(\phi), \mathfrak{I}_\mathcal{A})$, where $\mathcal{I}_{i,\mathcal{A}}$ is the restriction of $\mathcal{I}_i$ to the named individuals.

For the remaining cases, it suffices to observe that ϕ and $\mathrm{rew}_\mathcal{T}(\phi)$ are built on the same structure of temporal operators, which have the same semantics for both MTNCQs and MFOTL queries. □

For the second rewriting step, we restrict ourselves to finitely many time points. More precisely, we consider the finite structure $\mathfrak{I}_\mathcal{A}^{\mathrm{fin}}$, which is obtained from $\mathfrak{I}_\mathcal{A}$ by restricting the set of time points to $\mathrm{rep}(\mathcal{A})$. By Lemma 4.10, the information contained in this structure is already sufficient to answer atomic queries. We extend this structure a little, by considering the *two* representatives i, j for each maximal interval $[i, j]$ in $\mathbb{Z} \setminus \mathrm{tem}(\mathcal{A})$. In this way, we ensure that the 'border' elements are always representatives for their respective intervals. The size of the resulting structure $\mathfrak{I}_\mathcal{A}^{\mathrm{fin}}$ is polynomial in the size of $\mathcal{K}$.

Example 5.9. Let $\mathcal{A} = \{B(a,0), B(a,2), C(a,9)\}$ and $\mathcal{T} = \{\diamondminus\Diamond\!\!\!\!\!\!{\scriptstyle c}\;_3 B \sqcap \diamondplus C \sqsubseteq A\}$. Below one can see the finite structure $\mathfrak{I}_\mathcal{A}^{\mathrm{fin}}$ over the representative time points $\{-1, 0, 1, 2, 3, 8, 9, 10\}$, where for simplicity we omit the individual name.



The rewriting from Lemma 5.8 can refer to time instants outside of $\mathrm{rep}(\mathcal{A})$. However, when we want to evaluate a pure first-order formula over the finite structure $\mathfrak{I}_\mathcal{A}^{\mathrm{fin}}$, this is not possible anymore, because the first-order quantifiers must quantify over the domain of $\mathfrak{I}_\mathcal{A}^{\mathrm{fin}}$. Moreover, since the query $\mathrm{rew}_\mathcal{T}(\phi)$ can contain metric temporal operators, we need to keep track of the distance between the time points in $\mathrm{tem}(\mathcal{A})$. Hence, in the following we assume that $\mathfrak{I}_\mathcal{A}^{\mathrm{fin}}$ is given as a first-order structure with the domain $N_I \cup \{b_1, \ldots, b_n\} \cup \mathrm{rep}(\mathcal{A})$ and additional predicates `bit` and `sign` such that $\mathtt{bit}(i, j)$, $1 \le j \le n$, is true iff the jth bit of the binary representation of the time stamp i is 1, and $\mathtt{sign}(i)$ is true iff i is non-negative.

Thus, we now consider three-sorted first-order formulas with the three sorts N_I (for objects), $\{b_1, \ldots, b_n\}$ (for bits) and $\mathrm{rep}(\mathcal{A})$ (for time stamps). We denote variables of sort $\mathrm{rep}(\mathcal{A})$ by t, t', t''. To simplify the presentation, we do not explicitly denote the sort of all variables, but this is always clear from the context. Every concept name is now accessed as a binary predicate of sort $N_I \times \mathrm{rep}(\mathcal{A})$, e.g. $A(a, i)$ refers to the fact that individual a satisfies A at time point i. Similarly, role names correspond to ternary predicates of sort $N_I \times N_I \times \mathrm{rep}(\mathcal{A})$. It is clear that the expressions $t' \bowtie t$ and even $t' - t \bowtie m$ for some constant m and $\bowtie \in \{\ge, >, =, <, \le\}$ are definable as first-order formulas using the natural order $<$ on $\{1, \ldots, m\}$.

Lemma 5.10. For $\mathrm{rew}_{\mathcal{T}}(\phi)$ there is a constant $N \in \mathbb{N}$ such that, for every sub-formula ψ of $\mathrm{rew}_{\mathcal{T}}(\phi)$, every maximal interval J in $\mathbb{Z} \setminus \bigcup\{[i-N, i+N] \mid i \in \mathrm{tem}(\mathcal{A})\}$, all $k, \ell \in J$, and all relevant tuples $\mathbf{a}$ over N_I, we have $\mathfrak{I}_{\mathcal{A}}, k \models \psi(\mathbf{a})$ iff $\mathfrak{I}_{\mathcal{A}}, \ell \models \psi(\mathbf{a})$.

Proof. We are going to prove a more specific statement here. Namely, let N_ψ be the sum of all interval bounds of temporal formulas in a sub-formula ψ of $\mathrm{rew}_{\mathcal{T}}(\phi)$ (except for ∞). Consequently, for the proof we consider instead every maximal interval J in $\mathbb{Z} \setminus \bigcup\{[i-N_\psi, i+N_\psi] \mid i \in \mathrm{tem}(\mathcal{A})\}$.

We show this by induction on the structure of ψ, but only consider three representative cases; the other cases are similar.

- If ψ is the rewriting of an NCQ, then $N_\psi = 0$ and the semantics of ψ depends only on the interpretation at a single time point. Since k and ℓ belong to the same maximal interval in $\mathbb{Z} \setminus \mathrm{tem}(\mathcal{A})$, by Lemmas 4.13 and 4.14 and the construction of $\mathfrak{I}_{\mathcal{A}}$, this interpretation behaves in the same way at k and at ℓ.
- If ψ is of the form $\psi_1 \mathcal{U}_{[c_1,c_2]} \psi_2$, then $N_{\psi_1} \leq N_\psi - c_2$ and $N_{\psi_2} \leq N_\psi - c_2$. Assume that $\mathfrak{I}_{\mathcal{A}}, k \models \psi(\mathbf{a})$. Then there exists $j \in [c_1, c_2]$ such that
$$\mathfrak{I}_{\mathcal{A}}, k+j \models \psi_2(\mathbf{a}) \text{ and } \mathfrak{I}_{\mathcal{A}}, m \models \psi_1(\mathbf{a}), \text{ for all } m \text{ with } k \leq m < k+j. \tag{5.2}$$
In case that $j = c_1 = 0$, we have $\mathfrak{I}_{\mathcal{A}}, k \models \psi_2(\mathbf{a})$. Since k and ℓ are farther than $N_\psi \geq N_{\psi_2}$ from the nearest element of $\mathrm{rep}(\mathcal{A})$, by induction we also have $\mathfrak{I}_{\mathcal{A}}, \ell \models \psi_2(\mathbf{a})$ and thus $\mathfrak{I}_{\mathcal{A}}, \ell \models \psi(\mathbf{a})$ in this case. Hence, we can assume in the following that $j \geq c_1 > 0$, and thus in particular $\mathfrak{I}_{\mathcal{A}}, k \models \psi_1(\mathbf{a})$.
Since both $k+j$ and $\ell+c_2$ are farther than $N_\psi - c_2 \geq N_{\psi_2}$ from the nearest element of $\mathrm{rep}(\mathcal{A})$, by induction we have $\mathfrak{I}_{\mathcal{A}}, \ell+c_2 \models \psi_2(\mathbf{a})$. Moreover, since $\mathfrak{I}_{\mathcal{A}}, k \models \psi_1(\mathbf{a})$ and k as well as all elements in $[\ell, \ell+c_2]$ are farther than $N_\psi - c_2 \geq N_{\psi_1}$ from the nearest element of $\mathrm{rep}(\mathcal{A})$, by induction we have $\mathfrak{I}_{\mathcal{A}}, m \models \psi_1(\mathbf{a})$ for all m with $\ell \leq m \leq \ell + c_2$. Hence, $\mathfrak{I}_{\mathcal{A}}, \ell \models \psi(\mathbf{a})$.
- If ψ is of the form $\psi_1 \mathcal{U}_{[c_1,\infty)} \psi_2$, then we have a similar situation as above, except that j is not bounded by c_2. We can again assume that $j > 0$ and $\mathfrak{I}_{\mathcal{A}}, k \models \psi_1(\mathbf{a})$. Let p be the maximal element of J. If $k+j > p+c_1$, then $k+j > \ell$ and the distance between ℓ and $k+j$ must be at least c_1. Moreover, by assumption 5.2 we have $\mathfrak{I}_{\mathcal{A}}, m \models \psi_1(\mathbf{a})$ for all m with $p < m < k+j$. Since $\mathfrak{I}_{\mathcal{A}}, k \models \psi_1(\mathbf{a})$ and all elements in J are farther than $N_\psi \geq N_{\psi_1}$ from the nearest element of $\mathrm{rep}(\mathcal{A})$, by induction we also have $\mathfrak{I}_{\mathcal{A}}, m \models \psi_1(\mathbf{a})$ for all m with $\ell \leq m \leq p$. Thus, $\mathfrak{I}_{\mathcal{A}}, \ell \models \psi(\mathbf{a})$.
We now consider the remaining case that $k+j \leq p+c_1$. Then both $k+j$ and $\ell+c_1$ are farther than $N_\psi - c_1 \geq N_{\psi_2}$ from the nearest element of $\mathrm{rep}(\mathcal{A})$, and thus by induction we have $\mathfrak{I}_{\mathcal{A}}, \ell+c_1 \models \psi_2(\mathbf{a})$. By similar arguments as above, we obtain $\mathfrak{I}_{\mathcal{A}}, \ell \models \psi(\mathbf{a})$. □

Hence, for evaluating sub-formulas of $\mathrm{rew}_{\mathcal{T}}(\phi)$, it suffices to keep track of time points up to N steps away from the elements of $\mathrm{rep}(\mathcal{A})$; this includes at least one element from each of the intervals J mentioned in Lemma 5.10, since every element of $\mathrm{tem}(\mathcal{A})$ is immediately surrounded by two elements of $\mathrm{rep}(\mathcal{A})$.

We exploit Lemma 5.10 in the following definition of the three-sorted first-order formula $[\psi]^n(\mathbf{x},t)$ that simulates the behavior of $\psi(\mathbf{x})$ at the 'virtual' time point $t+n$, where $n \in [-N,N]$. Whenever we use a formula $[\psi]^n(\mathbf{x},t)$, we require that t denotes a representative for $t+n$. Due to our assumption that each maximal interval from $\mathbb{Z} \setminus \mathrm{tem}(\mathcal{A})$ is represented by its endpoints (see Example 5.9), we know that t is a representative for $t+n$ iff there is no element of $\mathrm{rep}(\mathcal{A})$ between t and $t+n$. We can encode this check in an auxiliary formula:

$$\mathtt{rep}^n(t) := \neg\exists t'.\,(t+n \leq t' < t) \vee (t < t' \leq t+n).$$

Example 5.11. In Example 5.9, 3 and 8 are representatives for the missing time points 4–7, and we have $\mathfrak{I}_\mathcal{A}^{\mathrm{fin}} \models \mathtt{rep}^1(3)$ (with $N=1$). However, for $\phi_\mathcal{T} = \bigcirc\neg C(x)$, we have $\mathfrak{I}_\mathcal{A}, 3 \models \phi_\mathcal{T}(a)$, but $\mathfrak{I}_\mathcal{A}, 8 \not\models \phi_\mathcal{T}(a)$, i.e. the behavior at 3 and 8 differs. To distinguish this, we need to refer to the 'virtual' time point 4 (gray circled 'v') that is not included in $\mathfrak{I}_\mathcal{A}^{\mathrm{fin}}$, via the formula $[\neg C(x)]^1(a,3)$. By Lemma 5.10 it is sufficient to consider 4, because this determines the behavior at 5–7 . ♢

We now define $[\psi]^n(\mathbf{x},t)$ recursively, for each sub-formula ψ of $\mathrm{rew}_\mathcal{T}(\phi)$. If ψ is a single rewritten NCQ, then $[\psi]^n(\mathbf{x},t)$ is obtained by replacing each atemporal atom $A(x)$ by $A(x,t)$, and similarly for role atoms. The parameter n can be ignored here, because we assumed that t is a representative for $t+n$, and hence the time points t and $t+n$ are interpreted in $\mathfrak{I}_\mathcal{A}$ equally. For conjunctions, we set $[\psi_1 \wedge \psi_2]^n(\mathbf{x},t) := [\psi_1]^n(\mathbf{x},t) \wedge [\psi_2]^n(\mathbf{x},t)$ and similarly for the other Boolean constructors. Finally, we demonstrate the translation for $\mathcal{U}$-formulas (the case of $\mathcal{S}$-formulas is analogous). We define $[\psi_1 \,\mathcal{U}_{[c_1,c_2]}\, \psi_2]^n(\mathbf{x},t)$ as

$$\begin{aligned}&\exists t'.\bigvee_{n' \in [-N,N]} \Big((t+n+c_1 \leq t'+n' \leq t+n+c_2) \wedge \mathtt{rep}^{n'}(t') \wedge [\psi_2]^{n'}(\mathbf{x},t') \wedge \\ &\quad \forall t''.\bigwedge_{n'' \in [-N,N]} \Big(((t+n \leq t''+n'' < t'+n') \wedge \mathtt{rep}^{n''}(t'')) \rightarrow [\psi_1]^{n''}(\mathbf{x},t'') \Big)\Big),\end{aligned}$$

where c_2 may be ∞, in which case the upper bound of $t+n+c_2$ can be removed.

Lemma 5.12. Let $\mathcal{K} = (\mathcal{T},\mathcal{A})$ be a consistent $\mathcal{TELH}_\bot^{\lozenge,\mathsf{lhs},-}$-KB and ϕ be an MTNCQ. Then $\mathrm{ans}([\mathrm{rew}_\mathcal{T}(\phi)]^0(\mathbf{x},t), \mathfrak{I}_\mathcal{A}^{\mathrm{fin}}) = \mathrm{ans}(\mathrm{rew}_\mathcal{T}(\phi), \mathfrak{I}_\mathcal{A})$.

Proof. We show the following claim by induction on the structure of ϕ: for all $i \in \mathrm{rep}(\mathcal{A})$, all $n \in [-N,N]$, all relevant tuples $\mathbf{a}$, and all TNCQs ϕ such that if $\mathfrak{I}_\mathcal{A}^{\mathrm{fin}} \models \mathtt{rep}^n(i)$ then

$$\mathfrak{I}_\mathcal{A}^{\mathrm{fin}} \models [\mathrm{rew}_\mathcal{T}(\phi)]^n(\mathbf{a},i) \text{ iff } \mathfrak{I}_\mathcal{A}, i+n \models \mathrm{rew}_\mathcal{T}(\phi)(\mathbf{a}).$$

Since this includes the case where $i \in \mathrm{tem}(\mathcal{A})$, $n=0$, for which $\mathfrak{I}_\mathcal{A}^{\mathrm{fin}} \models \mathtt{rep}^0(i)$ holds, the statement of the lemma follows.

If ϕ is an NCQ, then

$$\mathfrak{I}_\mathcal{A}^{\mathrm{fin}} \models [\mathrm{rew}_\mathcal{T}(\phi)]^n(\mathbf{a},i) \text{ iff } \mathfrak{I}_\mathcal{A}, i \models \mathrm{rew}_\mathcal{T}(\phi)(\mathbf{a}) \text{ iff } \mathfrak{I}_\mathcal{A}, i+n \models \mathrm{rew}_\mathcal{T}(\phi)(\mathbf{a})$$

since i is a representative for $i{+}n$ and a single temporal variable t is used in $[\mathrm{rew}_\mathcal{T}(\phi)]^n(\mathbf{x},t)$ to denote 'current' time point in $\mathrm{rew}_\mathcal{T}(\phi)$.

For the Boolean constructors, the claim follows immediately from the semantics of first-order logic.

We now consider a formula of the form $\phi\,\mathcal{U}_I\psi$. By induction, $\mathfrak{I}_{\mathcal{A}}^{\mathrm{fin}} \models [\mathrm{rew}_{\mathcal{T}}(\psi)]^{n'}(\mathbf{a}, i')$ iff $\mathfrak{I}_{\mathcal{A}}, i' + n' \models \mathrm{rew}_{\mathcal{T}}(\psi)(\mathbf{a})$, for any time point i' with $i' + n' \geq i + n$, and $\mathfrak{I}_{\mathcal{A}}^{\mathrm{fin}} \models [\mathrm{rew}_{\mathcal{T}}(\phi)]^{n''}(\mathbf{a}, i'')$ iff $\mathfrak{I}_{\mathcal{A}}, i'' + n'' \models \mathrm{rew}_{\mathcal{T}}(\phi)(\mathbf{a})$ for all time points i'' and offsets n'' such that $i + n \leq i'' + n'' < i' + n'$ (assuming w.l.o.g. that ϕ and ψ have the same answer variables).

Hence, the formula $[\mathrm{rew}_{\mathcal{T}}(\phi)\,\mathcal{U}_I\,\mathrm{rew}_{\mathcal{T}}(\psi)]^n(\mathbf{a}, i)$ checks the conditions required for the satisfaction of the $\mathcal{U}_I$-expression for all time points in $\bigcup\{[i - N, i + N] \mid i \in \mathrm{rep}(\mathcal{A})\}$. However, Lemma 5.10 tells us that, if $\mathrm{rew}_{\mathcal{T}}(\psi)$ is satisfied in $\mathfrak{I}_{\mathcal{A}}$ at some time point $i' + n'$ with $n' > N$, then this is also the case for $n' = N$. Similarly, to check whether $\mathrm{rew}_{\mathcal{T}}(\phi)$ is satisfied at all time points between $i + n$ and $i' + n'$, it suffices to consider the time points up to N away from some element of $\mathrm{rep}(\mathcal{A})$. Hence, $\mathfrak{I}_{\mathcal{A}}^{\mathrm{fin}} \models [\mathrm{rew}_{\mathcal{T}}(\phi)\,\mathcal{U}_I\,\mathrm{rew}_{\mathcal{T}}(\psi)]^n(\mathbf{a}, i)$ iff $\mathfrak{I}_{\mathcal{A}}, i + n \models (\mathrm{rew}_{\mathcal{T}}(\phi)\,\mathcal{U}_I\,\mathrm{rew}_{\mathcal{T}}(\phi))(\mathbf{a})$. □

This lemma allows us to compute in polynomial time that patient p_1 from Example 4.16 is an answer to $\phi(x)$ from Example 5.2 exactly at time point 7. Below we summarize our tight complexity results, which by Lemma 5.6 also hold for rooted MTCQs under certain answer semantics.

Theorem 5.13. Answering rooted MTNCQs under minimal-world semantics over $\mathcal{TELH}_{\bot}^{\diamondsuit,\mathsf{lhs},-}$-KBs is ExpSpace-complete, and PTime-complete in data complexity.

Proof. ExpSpace-hardness is inherited from propositional MTL [AH94; FS08]. Moreover, first-order formulas over finite structures can be evaluated in PSpace [Var82]. Finally, the size of $[\mathrm{rew}_{\mathcal{T}}(\phi)]^0(\mathbf{x}, t)$ is bounded exponentially in the size of ϕ and $\mathcal{T}$: each rewritten NCQ $\mathrm{rew}_{\mathcal{T}}(\psi)$ may be exponentially larger than ψ, and each $[\psi_1\,\mathcal{U}_I\psi_2]^n(\mathbf{x}, t)$ introduces exponentially many disjuncts and conjuncts (but the nesting depth of constructors in this formula is linear in the nesting depth of $\psi_1\,\mathcal{U}_I\psi_2$).

For data complexity, hardness is inherited from atemporal $\mathcal{EL}$ [CDL+13]. Evaluating FO($<$, `bit`)-formulas is in DLogTime-uniform AC^0 in data complexity [Lin92], and the size of our rewriting only depends on the query and the TBox. By Lemmas 5.8 and 5.12 and since $\mathfrak{I}_{\mathcal{A}}^{\mathrm{fin}}$ is of size polynomial in the size of $\mathcal{A}$, deciding whether a tuple $\mathbf{a}$ is a minimal-world answer of an MTNCQ w.r.t. a $\mathcal{TELH}_{\bot}^{\diamondsuit,\mathsf{lhs},-}$-KB is possible in PTime. □

5.4. Related Work and Discussion

In the current setting with only temporal concepts, the temporal interaction between time points is limited to the named individuals. For the patient selection task this seems powerful enough, since making the concept `DiabetesPatient` expanding is sufficient: It is not important whether exactly the same object represents `Diabetes` at all involved time points, but merely that at each such time point some successor exists that represents the disease. In other settings temporal roles could be beneficial. While making the matter more complicated we conjecture that is possible to extend the current construction to deal with temporal roles such that the result is still a minimal universal model that is also universal. For a given individual a, the number of successors that need to be introduced to satisfy a given existential restriction depends not only on a single time point anymore, but on all the time points. The notion of structural subsumption needs to be expanded to be able to deal with the temporal dimension. For example, suppose a belongs to $\exists r.A$ at time point t and a belongs to $\exists r.A'$ at some later time point $t' > t$, and the KB entails

$$\begin{aligned} \diamondminus r &\sqsubseteq r & \qquad\qquad \diamondminus A &\sqsubseteq A \\ A &\sqsubseteq A' & \diamondminus A' &\sqsubseteq A' \end{aligned}$$

i.e. r, A and A' are expanding. Then only one fresh element e needs to be introduced satisfying $\exists r.A'$ from t on. If another element e' would be introduced to satisfy $\exists r.A$ at time point t' an endomorphism could map e' to e and hence the model would not be minimal anymore. One has to formalize the notion '$\exists r.A$ subsumes $\exists r.A'$ for individual a at all time points' in order to decide how many individuals need to be introduced. During the rewriting these steps would have to be reversed and the monotonicity of the diamond operators could probably be exploited to obtain an efficient algorithm.

Since we are looking for minimal models, another way of admitting smaller models could be to adopt *varying domains* instead of a constant domain. Generally, many individuals are introduced that are only connected to the rooted part of the model in one time point. So conceptually one could question why they should exist at other time points in the first place. However, as long as we only allow rooted MTNCQs, varying domains can be simulated by constant domains. With a rooted query there is no way to query information about individuals that are not connected to the named part of a model.

In this work we focus on a discrete timeline (over $\mathbb{Z}$) and data facts stamped with a single time point. However, in the literature there are other approaches how to incorporate temporal formalisms in an ontology and the data, e.g. dense timelines, like $\mathbb{Q}$ or $\mathbb{R}$ [CCG10a; RWZ19]; or interval-based data models, where facts are stamped with a pair of time points denoting the interval in which they are true [KPP+16; CCG10a]. Within the discrete time point-based approach, one can distinguish between formalisms with LTL temporal constructs and formalisms like $\mathrm{LTL}^{\mathsf{bin}}$ employing more refined metric temporal operators from Metric Temporal Logic (MTL) [AFH96].

Combining ontology-mediated query answering with LTL operators has been investigated in depth. In particular, similarly to the query language adapted in this work, there is a multitude of works [BBL13; BT15b; BBL15a; BBL15b] investigating the complexity of answering LTL-CQs that are obtained from LTL formulas by replacing occurrences of propositional variables by arbitrary CQs. Moreover, the research [BLT13; BLT15] focuses on the rewritability properties of LTL-CQs. An orthogonal approach for query rewriting over a temporalized *DL-Lite* ontology was proposed in [AKW+13; AKK+15]. Here the focus mainly lies on increasing expressivity of an ontology language by allowing a concept or a role to be prefixed with LTL-constructs. Only recently, also metric variants of LTL-CQs have been considered [GJO16a; BBK+17; Tho18].

Negation in queries with the classical open-world semantics results in non-tractable (mostly coNP or even undecidable) query evaluation [Ros07b; GIK+15]. Moreover, prior work [BT15a; BT20] on temporalized ontology-mediated query answering with negation shows that the high complexity of temporal query answering with negation is mostly due to the open-world assumption for negation in a query language. There are several approaches on how to introduce negation in ontology-mediated query answering without losing tractability which we discussed in detail in Chapter 3. However, we also showed that they are not suitable for dealing with negation over anonymous individuals. This observation was the motivation for the introduction of minimal-world semantics. Indeed, as we show by this work, by changing semantics for negation, we can apply efficient (in data complexity) algorithms for temporal query answering with negation.

Part II.

Implementation and Experiments

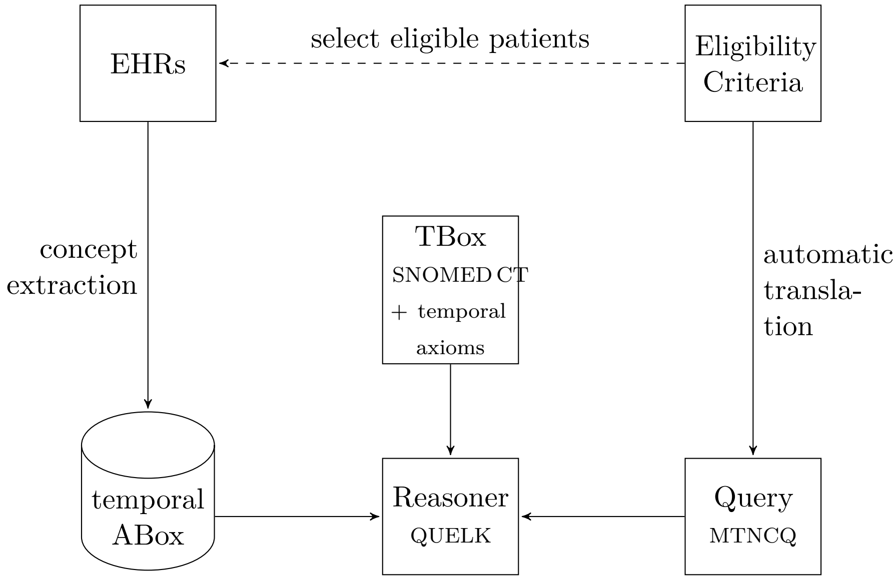


Figure 5.14.: The structure of an OMQA system for patient selection.

In Part I we were concerned with the theoretical development of temporal minimal-world semantics. Now we assess whether the theoretical work can be used in practice for the patient selection task introduced in Section 3.1. The following cite reiterates the practical importance of the task for medical researchers:

> *"Identifying patients who meet certain criteria for placement in clinical trials is a vital part of medical research. Finding patients for clinical trials is a challenge, as medical studies often have complex criteria that cannot easily be translated into a database query, but rather require examining the clinical narratives in a patient's records. This is time-consuming for medical researchers who need to recruit patients, so often researchers are limited to patients who either seek out the trial for themselves, or who are pointed towards a particular trial by their doctor. Recruitment from particular places or by particular people can result in selection bias towards certain populations (e.g., people who can afford regular care, or people who exclusively use free clinics), which in turn can bias the results of the study [...]. Developing NLP systems that can automatically assess if a patient is eligible for a study can both reduce the time it takes to recruit patients, and help remove bias from clinical trials [...]."* [3]

A system for patient selection that uses OMQA (in addition to natural language processing (NLP) techniques) has the structure illustrated in Figure 5.14. The overall goal is to select eligible patients for clinical trials based on their EHRs. In order to use an OMQA system, the following steps have to be taken:

[3] Cited from the website of the N2C2 2018 challenge (`https://portal.dbmi.hms.harvard.edu/projects/n2c2-2018-t1/`).

1. A terminology has to be chosen that can represent the necessary medical background knowledge. We base our vocabulary on the large biomedical ontology SNOMED CT, which is formulated in $\mathcal{ELH}_\bot$ and contains more than 300.000 different concepts. Since SNOMED CT was not developed for the representation of EHRs, we extend it slightly.
2. The eligibility criteria need to be translated to formal queries, in our case MTNCQs as introduced in Chapter 5. For this step it is crucial to have very accurate NLP tools.
3. Based on the vocabulary, the EHRs have to be transformed into temporal ABoxes. Here existing tools such as MetaMap can be used to find medical concepts in texts. Additional steps are required to extract the time stamps from EHRs. Once everything is lifted to the formal representations, a OMQA reasoner can be used to compute answers.

We have implemented tools for Steps 2 and 3: We describe and evaluate our system for the automatic translation of clinical trials to MTNCQs in Chapter 6. In Chapter 7 we introduce QUELK, our implementation of OMQA with the minimal-world semantics, and evaluate the feasibility of our approach on a gold-standard dataset for patient selection.

Chapter 6.

Translating Criteria to Temporal Queries

As the demand for (semi-)automated patient recruitment based on electronic health records (EHRs) becomes more and more urgent, the representation and formalization of eligibility criteria of clinical trials also have attracted considerable attention. To the best of our knowledge, however, there are no methods which can translate arbitrary eligibility criteria into logical expressions automatically (see Section 6.1 for related work).

Our goal in this chapter is to build a prototype system that can be evaluated in practice. The users of such a system would be medical researchers rather than logicians, hence the tool must be able to formalize eligibility criteria of clinical trials automatically. Since the available information is limited to EHRs, not all criteria can be evaluated by such a system, but it can support doctors in pre-selecting patients for a later, more thorough screening procedure.

In the following we present a prototypical implementation[1] that can automatically translate eligibility criteria into MTNCQs as introduced in Chapter 5. This can be seen as an instance of the larger field of translating natural language into a formal language with a precisely defined semantics.

Our translation is based on annotating eligibility criteria formulated in natural language by certain semantic roles and additional information. The semantic annotations we use focus on the kind of information that can be represented by our target query language, and hence can be seen as a filtering mechanism before the final translation to MTNCQs. Our prototype system uses existing natural language processing (NLP) techniques such as Word2Vec [MCC+13], Stanford NLP tools[2] [MSB+14], and MetaMap[3] [Aro01]. We evaluate our implementation on a random selection of criteria from clinicaltrials.gov,[4] which contains more than 3.000.000 criteria from over 250.000 clinical studies. We identify which kinds of criteria are easy or hard to translate. From this, we develop some suggestions on how to formulate eligibility criteria so that processing them automatically becomes easier and more accurate.

The remainder of this chapter is structured as follows: In Section 6.1 we discuss related work on the translation of eligibility criteria to a formal language. After that we introduce the methodology behind our approach in Section 6.2 and evaluate it on criteria from actual clinical trials in Section 6.3. Finally, in Section 6.4 we give a short

[1] Our prototype implementation with instructions on how to reproduce the results can be found at `https://github.com/wko/criteria-translation`.

[2] `https://nlp.stanford.edu/`

[3] `https://metamap.nlm.nih.gov/`

[4] `https://clinicaltrials.gov`

summary of our results and discuss how the automatic translation of criteria could be improved in the future.

6.1. Existing and Related Work

This work combines two strands of research, namely representation and formalization of eligibility criteria and automatic translation to formal languages.

Weng et al. [WTS+10] surveyed various representation methods of eligibility criteria and proposed a framework of five dimensions to compare them. According to different application scenarios, different representation methods for eligibility criteria are adopted. Bache et al. [BTM+15] proposed a general language for clinical trial investigation and construction (ECLECTIC) by analyzing 123 criteria from 8 clinical trials. Based on our own investigation of eligibility criteria, we propose MTNCQs as formal representation language since it covers a wide range of criteria, profits from existing medical ontologies and is based on a large body of research on (temporal) ontology-based query answering [BBL15b; BBK+17].

Previous work has already considered translation of eligibility criteria.
Tu et al. [TPC+11] proposed a practical translation method based on the ERGO annotation, which is an intermediate representation for eligibility criteria. However, ERGO annotation can only be done manually or semi-automatically. Milian et al. [MBT12; MT13] focused on breast-cancer trials and summarized 165 patterns, and used these patterns and concept recognition tools to structure criteria. After that, they generated a formal representation by projecting the concepts in criteria to the predefined query template. There is also some work about extraction and representation of *partial* knowledge in eligibility criteria. Zhou et al. [ZMP+06], Luo et al. [LYW11] and Boland et al. [BTC+12] focused on the recognition and representation of temporal knowledge. Huang et al. [HL07] and Enger et al. [EVØ17] proposed several methods for detecting negated expressions.

Weng et al. [WWL+11], Luo et al. [LYW11], Bhattacharya et al. [BC13], and Chondrogiannis et al. [CAT+17] classified the clinical trials into limited semantic classes by using semantic pattern recognition or machine learning methods, which is helpful for figuring out the most prominent kinds of information expressed in clinical trials.

In the field of NLP, automatic translation from a natural language into formal language, e.g. first-order logic formulas, is also known as *automatic semantic parsing*. Dong et al. [DL16] proposed an automatic semantic parsing method based on machine learning, different from traditional rule-based or template-based methods.

6.2. Methodology

The main idea is to use semantic annotations to bridge the gap between eligibility criteria and formal queries. The working of our system can be broadly divided into two stages: annotating the eligibility criterion, and then constructing a formal query from the semantic annotations. The outline of the system is shown in Figure 6.1.

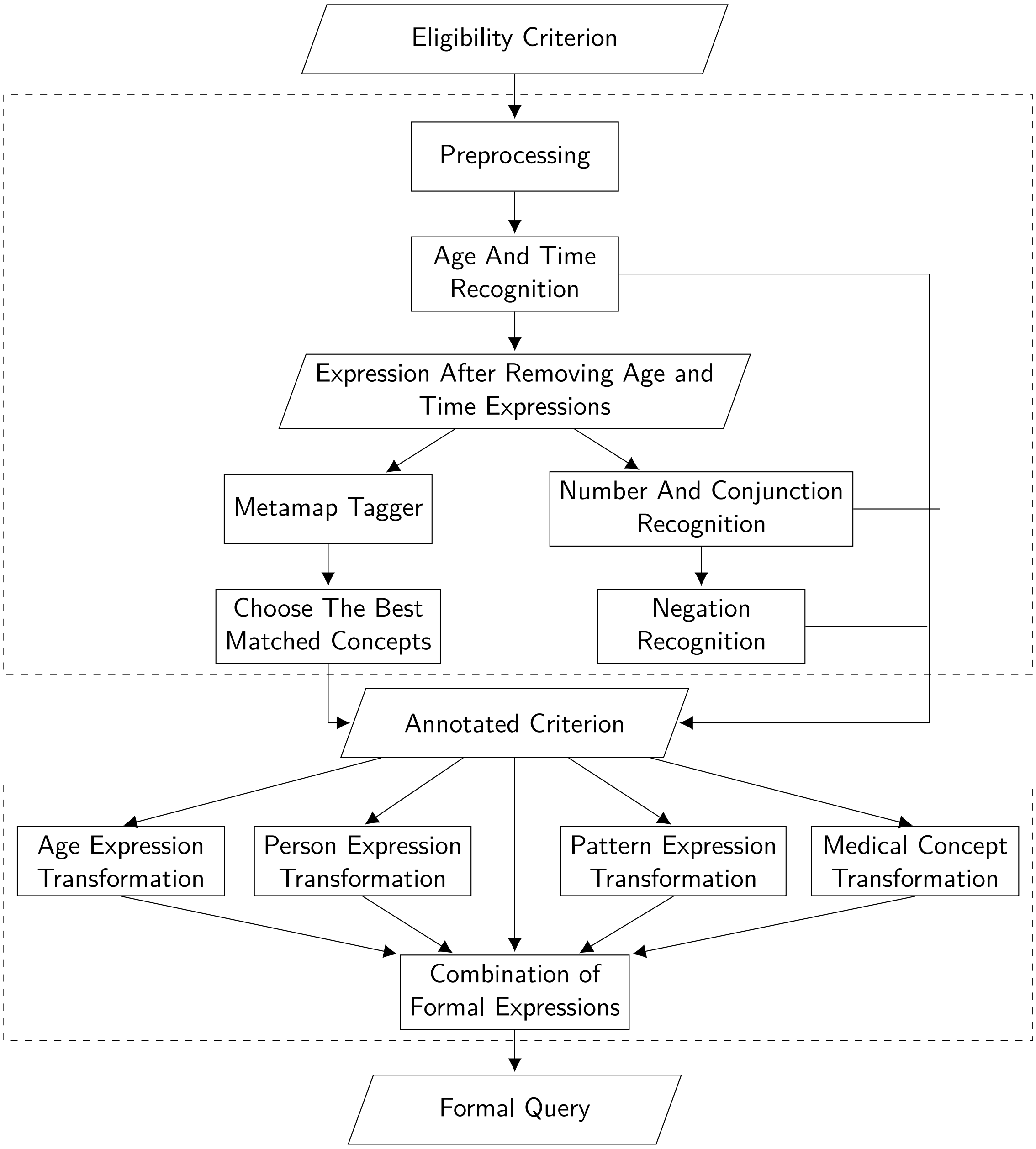


Figure 6.1.: Outline of the translation system

6.2.1. Semantic Annotations

Our annotations identify pieces of information that can be translated to MTNCQ constructors, such as temporal operators, negation, and medical concepts. The design of the annotations also incorporates knowledge about frequently occurring types of eligibility criteria, and takes into account whether it can be reasonably expected that the queried information can be found in EHRs. We use the MetaMap tagger to recognize medical concepts, and we use keyword matching to recognize other concepts. As a preprocessing step, we homogenize the NL criteria, e.g. replace ‘two’ with ‘2’ and replace ‘greater than’ with ‘>’.

The Selection of Semantic Roles

After looking at a number of eligibility criteria, we identified the following frequently requested types of information: *age, gender, diagnoses, medications, procedures, measurements*, and *temporal context* (e.g. ‘history of ...’). This analysis is consistent with the results of Weng et al. [WWL+11], Luo et al. [LYW11], Bhattacharya et al. [BC13], and Chondrogiannis et al. [CAT+17], which all rank this kind of information high in their lists of prominent semantic classes.

Our formalization is based on SNOMED CT, which contains 19 top-level and more than 350 second-level categories. Out of these, we identified 8 categories that correspond to the above-listed information: *clinical finding, observable entity, product, substance, procedure, unit, family medical history, person.* For now, we discard other semantic classes from SNOMED CT, such as *qualifier values* (‘severe’, ‘known’, ‘isolated’) or *devices*. This restriction helps to resolve some of the ambiguity of words or phrases. For example, in SNOMED CT ‘female’ can be mapped to ‘*Female structure (body structure)*’ or ‘*Female (finding)*’; and ‘scar’ can be identified as ‘*Scar (disorder)*’ or ‘*Scar (morphologic abnormality)*’. By excluding the types *body structure* and *morphologic abnormality*, we obtain a more uniform representation.

However, SNOMED CT only contains medical concepts, and we additionally consider the semantic roles *age, time, number, comparison sign, negation*, and *conjunction*. Table 6.2 contains an overview of all semantic roles with examples. In addition to the semantic role, we record additional information in the annotations, e.g. the precise concept from SNOMED CT or a time interval.

Our choice of semantic roles determines the *vocabulary* that we will use to formulate MTNCQs. More precisely, the concept names are restricted to the sub-concepts of the 8 categories in SNOMED CT identified above. We use the role names **`diagnosedWith`**, **`takes`**, and **`undergoes`** to connect patients to SNOMED CT concepts, but none of the role names from SNOMED CT itself. Additionally, we allow concrete domain predicates like **`hemoglobinOf`** that correspond to SNOMED CT *substances* (e.g. **`Hemoglobin`**) and *observable entities*, as well as **`ageOf`**. Finally, temporal information, negation, and conjunction are expressed by the logical connectives of our query language.

Semantic role	Examples	Representation
Age	age 18–70	$[\text{lower}, \text{upper}]$
Time	within 5 years	$[\text{start}, \text{end}]$
Comparison sign	greater than	$> \mid \geq \mid \leq \mid <$
Partial negation	other than	$\wedge\neg$
Main negation	no history of	$\neg$
Number	one, two, three, ...	Arabic numerals
Conjunction	and, or, defined by	$\wedge, \vee$
From SNOMED CT (e.g. clinical finding)	lung disease	Concept name

Table 6.2.: List of semantic roles and representations in the semantic annotation

Concept Recognition and Semantic Role Annotation

To illustrate the annotation process, we consider the criterion 'history of lung disease other than asthma';[5] Table 6.3 and the end result in Figure 6.4.

The first steps are to recognize and annotate age and temporal expressions using regular expressions. In our example, 'history of' is recognized by the regular expression

(a|any|prior|previous)(.*?)history of,

and then annotated by the semantic role *time* and the temporal interval $(-\infty, 0]$. We then remove the identified age expressions and temporal expressions from the EC. They form complete semantic units, and thus removing them does not affect the meaning of the remaining part of the EC, while it allows us to avoid accidental translation of these expressions into SNOMED CT concepts.

On the remaining criterion, we then run the MetaMap tagger [Aro01], a tool for recognizing concepts from the UMLS Metathesaurus, which subsumes SNOMED CT. Given a phrase or sentence, it returns the most likely phrase-concept matches. In our example, MetaMap does not identify any sub-phrases, and outputs the following concepts for the whole phrase 'lung disease other than asthma': *'Disorder of lung (disorder)', 'Lung structure (body structure)', 'Asthma (disorder)'*. By restricting the types as described in Section 6.2.1, we immediately rule out *'Lung structure'*.

A larger challenge, however, is to obtain more exact phrase-concept matches. For this, we split all sub-phrases returned by MetaMap into more sub-phrases using the Stanford NLP tools [MSB+14]. Then we try to find the best phrase-concept matches, by calculating a similarity value (in $[0, 1]$) of each sub-phrase to all candidate concepts using Word2Vec [MCC+13] and the Levenshtein distance; we also use the synonymous expressions provided by SNOMED CT to potentially obtain a higher similarity. To avoid spurious matches, we use a minimum threshold of 0.66 for the similarity. In our example, this excludes the words 'other' and 'than', because there is no candidate concept that is

[5] `https://clinicaltrials.gov/ct2/show/NCT02548598`

Stage	Output
Original Criterion	history of lung disease other than asthma
Age recognition	—
Time recognition	history of → *(time)*
Remove age/time	lung disease other than asthma
MetaMap	lung disease other than asthma → *Disorder of lung (disorder), Lung structure (body structure), Asthma (disorder)*
Restrict semantic roles	lung disease other than asthma → *Disorder of lung, Asthma*
Detect sub-phrases	lung disease, lung, disease, other, than, asthma
Compute most similar concept for each sub-phrase	(lung disease, *Disorder of lung*) : 0.91, (lung, *Disorder of lung*) : 0.81, (disease, *Disorder of lung*) : 0.89, (asthma, *Asthma*) : 1.0
Find best matches	lung disease → *Disorder of lung*, asthma → *Asthma*
Negation recognition	other than → *(negation)*
Other semantic roles	—

Table 6.3.: Example of the semantic annotation of an EC.

similar enough. The best matches for the phrases ‘lung disease’, ‘lung’, and ‘disease’ all refer to the same concept *Disorder of lung*, and we use the similarity values to choose the best of them, where we give preference to longer phrases.

It remains to recognize other semantic roles in the EC, i.e. *number, negation, comparison sign*, and *conjunction*. We mainly do this by keyword or pattern matching. The negation case is the most complex due to its various forms:

- explicit negation e.g. ‘not’, ‘except’, ‘other than’, ‘with the exception of’;
- morphological negation, e.g. ‘non-pregnant’, ‘non-healed’, ‘non-smoker’;
- implicit negation, e.g. ‘lack of’, ‘rule out’, ‘free from’.

In our prototype system, we focus on explicit negation, and consider two cases: either the whole sentence is negated (‘patient does not have . . . ’) or only part of it (‘. . . other than . . . ’). For conjunctions between parts of sentences, we use ‘$\vee$’ as default annotation, because there is no good way to map ‘and’ and ‘or’ in criteria to conjunction or disjunction exactly, e.g. in the criterion ‘. . . including cyclosporine, systemic itraconazole *or* ketoconazole, erythromycin *or* clarithromycin, nefazodone, verapamil *and* human immunodeficiency virus protease inhibitors’[6] both ‘and’ and ‘or’ have the same meaning.

The final semantic annotation for our example can be seen in Figure 6.4.

[6]`https://clinicaltrials.gov/ct2/show/NCT02452502`

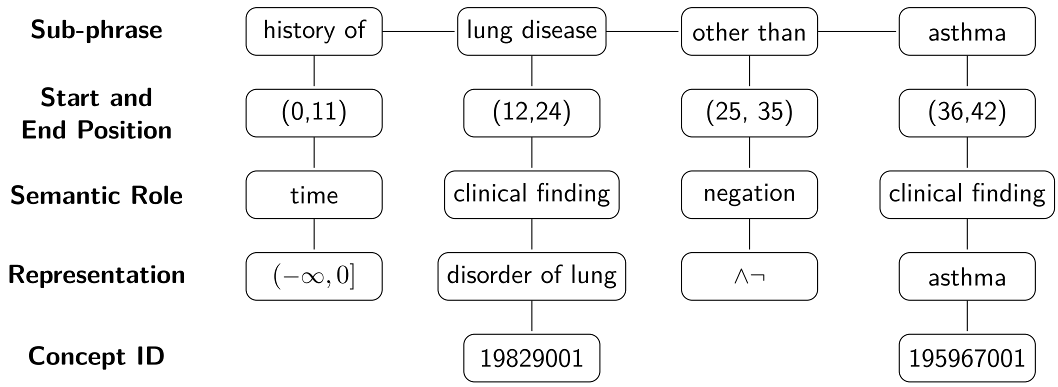


Figure 6.4.: The semantic annotation for our example.

6.2.2. The Formal Queries

To obtain the final MTNCQ, we combine the different annotated phrases according to the composibility of semantic roles and structural information. There are four kinds of basic sub-formulas: age formulas, person formulas, medical formulas and pattern formulas, and their translation is described in Table 6.5. For measurements, we detect patterns in the semantic annotation that correspond to a comparison of a *substance* or *observable entity* with a specific numerical value (including unit). Additionally, we group adjacent SNOMED CT *findings* together, to translate them into a set of atoms joined by $\vee$ inside the same $\exists y.\mathtt{diagnosedWith}(x, y) \wedge \ldots$ formula. We also translate negation between *clinical findings* into appropriate formulas, and do the same for *products* and *procedures*. In our running example, 'lung disease other than asthma' is formalized as

$$\exists y.\mathtt{diagnosedWith}(x, y) \wedge \mathtt{DisorderOfLung}(y) \wedge \neg\mathtt{Asthma}(y).$$

Finally, we combine these sub-formulas using the remaining connectives and negations and consider any time expressions. In our prototype system, we only express a single temporal operator of the form $\Diamond_{[-n,0]}$, which we found to be the most common in clinical trials. Such an operator is always applied to the whole formula, e.g. we obtain

$$\Diamond_{(-\infty,0]}(\exists y.\mathtt{diagnosedWith}(x, y) \wedge \mathtt{DisorderOfLung}(y) \wedge \neg\mathtt{Asthma}(y)).$$

If there is more than one temporal annotation, we choose the more specific one. For example, in 'history of myocardial infarction, unstable angina pectoris, percutaneous coronary intervention, congestive heart failure, hypertensive encephalopathy, stroke or TIA within the last 6 months'[7] there are 'history of' and 'within the last 6 months', and we choose the latter.

If there are no explicit connectives, we combine medical and measurement formulas by disjunction, and then combine them with age and person formulas by conjunction.

[7]https://clinicaltrials.gov/ct2/show/NCT00220220

Semantic role	**Formalization**	**Example**
Age	$\text{ageOf}(x) \geq \text{lower} \wedge \text{ageOf}(x) \leq \text{upper}$	$\text{ageOf}(x) \geq 18$
Time	$\Diamond_{[\text{start,end}]}$	$\Diamond_{[-12,0]}$
Person	$\text{conceptName}(x)$	$\text{Woman}(x)$
Clinical finding	$\exists y.\text{diagnosedWith}(x, y) \wedge \text{conceptName}(y)$	$\exists y.\text{diagnosedWith}(x, y) \wedge \text{HIV}(y)$
Product	$\exists y.\text{takes}(x, y) \wedge \text{conceptName}(y)$	$\exists y.\text{takes}(x, y) \wedge \text{Aspirin}(y)$
Procedure	$\exists y.\text{undergoes}(x, y) \wedge \text{conceptName}(y)$	$\exists y.\text{undergoes}(x, y) \wedge \text{Appendectomy}(y)$
Measurement pattern: substance/observable entity—comparison sign—number—unit		
Formula: $\text{conceptNameOf}(x)$ $(> \mid \geq \mid \leq \mid <)$ number conceptName		$\text{hemoglobinOf}(x) < 14g/dl$
Group pattern: clinical finding—clinical finding—. . .		
Formula: $\text{conceptName1}(y) \vee \text{conceptName2}(y) \vee \ldots$		$\text{HIV}(y) \vee \text{HepatitisC}(y)$
Negation pattern: clinical finding—partial negation—clinical finding		
Formula: $\text{conceptName1}(y) \wedge \neg\text{conceptName2}(y)$		$\text{Diabetes}(y) \wedge \neg\text{DiabetesType1}(y)$

Table 6.5.: Translation of basic query parts.

Eval.	Unans.	Ans.
1	282	119
2	254	147
3	237	164

Eval.	Good	Partial	Wrong
1	54	29	10
2	56	27	10
3	65	18	10

Table 6.6.: Experimental results. The right table shows the annotation of the translation quality for the 93 criteria that were marked as 'answerable' (Ans.) by all evaluators (Eval.).

6.3. Experiments

To the best of our knowledge, there are no gold standard datasets for the translation of criteria into formal language. Therefore, we evaluated our approach on real-world studies taken from clinicaltrials.gov.[8] During the design phase we used 24 randomly selected studies, which contained approximately 300 criteria. Our prototype system was optimized to cover as many of these criteria as possible.

For testing, we randomly selected criteria across all studies on clinicaltrials.gov and manually evaluated them. Due to time constraints, we managed to process 401 criteria. We defined the following metrics: A criterion is *answerable*, if a) it is possible for a human to translate it into an MTNCQ using only the vocabulary chosen in Section 6.2.1; and b) it can in principle be answered by only looking at the EHR of a patient. Hence, criteria that refer to the future ('during study phase'), or ask for subjective information ('in the opinion of the investigator', 'willingness to'), are not considered answerable for the purposes of our system. For each answerable criterion, we then evaluated the quality of the translation. The resulting MTNCQ is labeled as *good* if it contains all (necessary) information; *partial* if it represents at least parts of the criterion; and *wrong* otherwise. These metrics are clearly subjective to some extent. To get a more reliable evaluation and to quantify the amount of subjectivity, we let three evaluators (three of the authors of [XFB+19]) vote independently on the test data. The results can be seen in Table 6.6.

The results indicate that the judgment on whether a criterion is answerable or not differs between the evaluators. We found that the difference is mainly caused by two things: Firstly, it is sometimes difficult to judge whether a concept can be represented in SNOMED CT, because the concept name can differ significantly from the description in the text. Secondly, many criteria contain very specific phrases, for example 'Active bowels inflammatory disease ([Crohn], chronic, diarrhea...)'.[9] The word 'active' cannot be translated into SNOMED CT, and we could translate it into a temporal constraint only under some assumptions on the semantics of 'active'. Some might consider this to be not so important, while for others this renders the criterion unanswerable. Despite the differences, at least 60% of the criteria cannot be answered, even in the opinion of evaluator 3, who was the most optimistic. This is partially because of condition b) above. The second reason is that quite a number of criteria cannot be represented in our formal

[8] `https://clinicaltrials.gov/`
[9] `https://clinicaltrials.gov/ct2/show/NCT02363725`

language, either because of a lack of vocabulary in SNOMED CT, or because of missing semantic roles (see Section 6.2.1). While the former cannot be improved on, the latter offers room for future optimizations.

To compare the quality of the translations, we consider only criteria that have been marked as *answerable* by all evaluators. This leaves 93 criteria that are analyzed on the right-hand side of Table 6.6. The difference in the translation quality is again due to the varying opinions of the evaluators regarding how detailed a translation needs to be in order to be considered good. Our system is able to translate more than 50% of the (confidently) answerable criteria, which is a promising first result. In the following, we give examples for a good, partial, and a bad translation of our system:

'Has a history of diabetic ketoacidosis in the last 6 months.'[10]

$$\Diamond_{[-6,0]}\big(\exists y.\mathtt{diagnosedWith}(x,y)\wedge\mathtt{KetoacidosisInDiabetesMellitus}(y)\big)$$

'History of, diagnosed or suspected genital or other malignancy (excluding treated squamous cell carcinoma of the skin), and untreated cervical dysplasia.'[11]

$$\Diamond_{(-\infty,0]}\Big(\exists y.\mathtt{diagnosedWith}(x,y)\wedge \big(\mathtt{MalignantNeoplasticDisease}(y)\vee\mathtt{DysplasiaOfCervix}(y)\big)\Big)$$

'Primary tumors developed 5 years previous to the inclusion, except in situ cervix carcinoma or skin basocellular cancer properly treated'[12]

$$\Diamond_{(-\infty,0]}\Big(\exists y.\mathtt{diagnosedWith}(x,y)\wedge\big(\mathtt{CarcinomaInSituOfUterineCervix}(y)\vee\mathtt{SkinCancer}(y)\big)\Big)$$

The second translation is partially correct, because the temporal data and the main concepts have been recognized correctly, but 'excluding ... ' was not translated. The last translation is wrong since neither the temporal information, the negation, nor the main concept 'primary tumors' have been recognized correctly. For more examples, we refer the reader to the appendix in the extended version.

6.4. Discussion and Future Work

Formalizing eligibility criteria is a challenging task due to the gap between natural and formal language. We have presented an automatic translation method from eligibility criteria into formal queries, and developed a prototype system based on existing NLP tools. We have evaluated our prototype on 401 eligibility criteria. More than 50% of the answerable criteria have been translated correctly, which is an encouraging result that can be improved on by optimizing the translation process as we describe below. However, there remain certain criteria that are hard to translate (even for humans) due to their complex structure.

[10] `https://clinicaltrials.gov/ct2/show/NCT02269735`
[11] `https://clinicaltrials.gov/ct2/show/NCT01397097`
[12] `https://clinicaltrials.gov/ct2/show/NCT01303029`

While it is unreasonable to expect medical doctors to formulate clinical trial criteria directly as MTNCQs, we nevertheless identify a few key points that can be observed during the formulation of eligibility criteria to make the automatic translation easier:

1. Split criteria whenever possible, e.g. divide 'diagnosed with diabetes and hypertension' into 'diagnosed with diabetes' and 'diagnosed with hypertension.'
2. Formulate every criterion as an independent description that does not depend on other criteria or the background knowledge of clinical trials, like in 'Known hypersensitivity to any of the study drugs or excipients.'[13]
3. Avoid using nonadjacent words to express a concept, e.g. '... dermatologic, neurologic, or psychiatric disease'[14] should rather be formulated as 'dermatologic disease, neurologic disease, or psychiatric disease.'

We can improve the quality of our translation by collecting more regular expressions and custom mappings, or employing specialized techniques from the literature for the recognition of semantic roles like *comparison sign* or *negation*. Other obvious steps are the inclusion of more concept categories from SNOMED CT such as *devices, qualifiers*, and *events*. For example, the criterion 'severe aortic stenosis'[15] could be translated as

$$\exists y, z.\mathtt{hasDiagnosis}(x, y) \wedge \mathtt{AorticStenosis}(y) \wedge \mathtt{severity}(y, z) \wedge \mathtt{Severe}(z)$$

if we annotate 'severe' with the SNOMED CT concept *severe (qualifier value)* and detect the pattern *qualifier value—finding*. It is also straightforward modify our system to output a ranked list of multiple candidate translations that the doctor may choose from.

Another interesting direction for future work is to develop a *controlled natural language* [Kuh14] based on our semantic annotations. Criteria formulated in this way can then easily be transformed into MTNCQs as we have described. With appropriate editing support, creating new eligibility criteria that conform with this controlled natural language would be not much more difficult than writing them as free-form text. Of course, one should retain the possibility to add free-form criteria, which then have to be evaluated manually.

[13] https://clinicaltrials.gov/ct2/show/NCT01935492
[14] https://clinicaltrials.gov/ct2/show/NCT00960570
[15] https://clinicaltrials.gov/ct2/show/NCT01951950

Chapter 7.

Selecting Patients for Clinical Trials

In this chapter we assume that the criteria are already translated to MTNCQs and focus on applying temporal OMQA for the selection of patients. In Section 7.1 we describe QUELK, our system for query answering, and the input formats it requires. Moreover, we discuss the parts in which the implementation differs from the theoretical algorithms we have developed before. QUELK is then put to work on a small gold standard dataset for patient selection introduced in Section 7.2. We describe the preprocessing of the EHRs and the setup of the KB in Section 7.3. Based on existing tools we extract medical concepts occurring in the EHRs and generate a temporal ABox automatically. We also introduce criteria from the dataset that contain temporal information and provide manual translations to MTNCQs for them. We conduct two experiments in Section 7.4 to show the importance of temporal reasoning: In the first setting, all temporal information are ignored, effectively merging all EHRs to a single time point and using only NCQs for querying. This is evaluated against the temporal setting, in which the MTNCQs are evaluated over the temporal ABox. As expected the quality of the results is better when the temporal dimension is taken into account. In Section 7.5 we finish with a short discussion of our results relative to other systems and provide some related work.

7.1. QUELK: A Prototype For Temporal Query Answering

QUELK[1] is a prototypical system for answering MTNCQs over $\mathcal{TELH}_{\bot}^{\Diamond,\mathsf{lhs},-}$-KBs implemented in Java and Scala.

As input QUELK accepts an ontology in OWL functional-style syntax[2]. The temporal information are encoded using annotations on the GCIs, class and role assertions. For example, to make the concept **Patient** rigid and assert that individual **p101** is a **Patient** on **2173-02-04**, the following axioms can be used in a $\mathcal{TELH}_{\bot}^{\Diamond,\mathsf{lhs},-}$-KB:

$$\Diamond\!\!\!\!\!\;\mathtt{Patient} \sqsubseteq \mathtt{Patient}$$
$$\mathtt{Patient}(\mathtt{p101}, \mathtt{2173\text{-}02\text{-}04})$$

They can be expressed in OWL functional-style by a 'SubClassOf'-axiom and a 'ClassAssertion'-axiom, which can be read much like the original axioms in DL-syntax:

[1]The source code of QUELK can be found at `https://github.com/wko/quelk`.
[2]`https://www.w3.org/TR/owl2-syntax/#Functional-Style_Syntax`

```
SubClassOf(
  Annotation(time:diamond "rigid"^^xsd:string)
  :Patient :Patient)

ClassAssertion(
  Annotation(time:instant "2173-02-04T00:00:00"^^xsd:dateTime)
  :Patient :p101)
```

The saturation of the temporal KB is implemented using the saturation algorithm introduced in Section 4.3. To be able to deal with many facts, the saturated ABox is then stored in a relational database management system (RDBMS), since they are optimized for efficient storage and querying of huge amounts of data. We use our results from Chapter 4 and Chapter 5 and store just representative time points and the interval they represent. Interval bounds are allowed to be positive or negative infinity. The smallest time unit in QUELK are milliseconds, which is sufficient at least for the patient selection task.

When a MTNCQ ϕ with answer variable $\mathbf{x}$ is evaluated over a database $\mathfrak{I}_{\mathcal{A}}^{\text{fin}}$, it is rewritten to an SQL query that returns the correct answers when evaluated over $\mathfrak{I}_{\mathcal{A}}^{\text{fin}}$. We have shown in Chapter 5 that the rewriting of the NCQs inside ϕ is independent of the temporal operators in the TBox (see the proof of Lemma 5.8). This means for this step we can treat the $\mathcal{TELH}_{\bot}^{\diamond,\mathsf{lhs},-}$-TBox like an atemporal $\mathcal{ELH}_{\bot}$-TBox by omitting all diamond operators in any axiom. We implement the atemporal rewriting of the NCQs in ϕ using the algorithm in Section 3.4 and use the ELK reasoner[3] as an optimized black-box for computing subsumptions in $\mathcal{ELH}_{\bot}$.

In the second step, we deal with the remaining temporal parts in ϕ. Here our implementation differs from the theoretical rewriting described in Section 5.3, because we construct the SQL queries to work not on single time points, but on intervals instead: The SQL query corresponding to a given NCQ φ in ϕ selects *answer tuples* of the form $(\mathbf{a}, \mathbf{i})$, where $\mathbf{a}$ is a tuple of named individuals from $\mathfrak{I}_{\mathcal{A}}^{\text{fin}}$ (with the same arity as $\mathbf{x}$) and $\mathbf{i}$ is an interval at which $\mathbf{a}$ are answers to φ. Based on this representation all further temporal operators in ϕ can be taken care of.

In order for this representation to work, all computations on intervals have to be defined based on the bounds of the interval, since we never want to materialize the possibly infinitely many time points in a given interval. In the following an interval $\mathbf{i} = [a, b]$ is a closed interval over the integers in which a and b are also allowed to be $-\infty$ or ∞ (in which case the interval is open). Then $\mathbf{i}$ is empty (denoted by $\emptyset$) if $a > b$. Let $\mathbf{i}_1 = [a_1, b_1]$ and $\mathbf{i}_2 = [a_2, b_2]$ be intervals. The result of the intersection can always be represented by a new interval:

$$[a_1, b_1] \cap [a_2, b_2] = [\max(a_1, a_2), \min(b_1, b_2)]$$

[3] `https://github.com/liveontologies/elk-reasoner`

It is also easy to check that $\mathbf{i}_1$ and $\mathbf{i}_2$ are *overlapping* iff $\mathbf{i}_1 \cap \mathbf{i}_2 \neq \emptyset$. The result of the union of two intervals can be represented by a set of intervals:

$$[a_1, b_1] \cup [a_2, b_2] = \begin{cases} \{[\min(a_1, a_2), \max(b_1, b_2)]\} & \text{iff } [a_1, b_1] \cap [a_2, b_2] \neq \emptyset \\ \{[a_1, b_1], [a_2, b_2]\} & \text{otherwise} \end{cases}$$

We can also easily check if $\mathbf{i}_1$ is a sub-interval of (or contained in) $\mathbf{i}_2$, formally $\mathbf{i}_1 \subseteq \mathbf{i}_2$, by checking if $a_2 \leq a_1$ and $b_1 \leq b_2$.

Given a set of intervals $\mathbf{I}$ we sometimes want to check if a given interval $\mathbf{i}$ is contained in the intervals in the set. Generally, $\mathbf{I}$ can contain many overlapping intervals, which makes this check more complicated. Therefore, we introduce a "normalization" operation on $\mathbf{I}$, which merges all overlapping intervals in $\mathbf{I}$. In the database world this is known as the *gaps and islands problem* [DDL+11] and their exist many solutions to it. For us it is enough to know that the problem can be solved in SQL and the functions have the following semantics for $\mathbf{I}$:

$$\begin{aligned} \text{islands}(\mathbf{I}) &:= \{\mathbf{i} \mid \mathbf{i} \text{ is a maximal interval with } \mathbf{i} \subseteq \bigcup_{\mathbf{i}' \in \mathbf{I}} \mathbf{i}'\} \\ \text{gaps}(\mathbf{I}) &:= \{\mathbf{i} \mid \mathbf{i} \text{ is a maximal interval with } \mathbf{i} \cap \bigcup_{\mathbf{i}' \in \mathbf{I}} \mathbf{i}' = \emptyset\}, \end{aligned}$$

Example 7.1. An example of the results of applying the operations islands and gaps on a set of intervals $\mathbf{I} := \{[1, 5], [3, 9], [21, 28]\}$:

$$\begin{aligned} \text{islands}(\mathbf{I}) &:= \{[1, 9], [21, 28]\} \\ \text{gaps}(\mathbf{I}) &:= \{(-\infty, 0], [10, 20], [29, \infty)\} \end{aligned}$$

$\diamondsuit$

For easier notation we extend the two functions to sets of answer tuples. For a given set S of answer tuples of the form $(\mathbf{a}, \mathbf{i})$, islands can be computed independently for each answer $\mathbf{a}$ occurring in S and the union of all islands is returned. Gaps for answer tuples can be computed in a similar way.

Let $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ be a consistent $\mathcal{TELH}_{\bot}^{\diamondsuit,\mathsf{lhs},-}$-KB and $\mathfrak{I}_{\mathcal{A}}^{\text{fin}}$ a finite interpretation (stored in a relational database) containing the completion of $\mathcal{A}$ and just the representative time points (we already used this notation in Section 5.3). We define the function eval that computes the answers to an MTNCQ $\phi(\mathbf{x})$ in $\mathfrak{I}_{\mathcal{A}}^{\text{fin}}$ using the same semantic approach as the actual SQL queries constructed by QUELK. For an NCQ φ, eval selects all time points at which $\mathbf{a}$ is an answer to some rewriting of φ and lifts the result to the interval level by replacing each time point i by its represented interval, denoted by $i^{\uparrow}$, formally

$$\text{eval}(\varphi, \mathfrak{I}_{\mathcal{A}}^{\text{fin}}) := \{(\mathbf{a}, i^{\uparrow}) \mid \mathcal{D}, i \models \varphi'(\mathbf{a}) \text{ and } \varphi' \in \text{rew}_{\mathcal{T}}(\varphi)\}$$

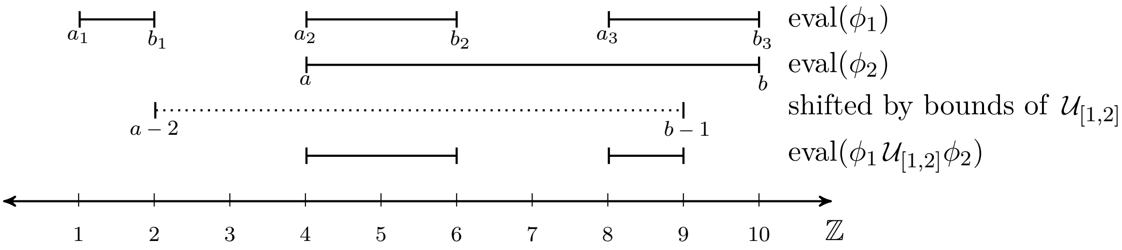


Figure 7.2.: An example for $\phi_1 \mathcal{U}_{[1,2]} \phi_2$. Suppose ϕ_1 is valid in the three intervals $[a_i, b_i]$ with $i \in 1, 2, 3$ and ϕ_2 in the interval $[a, b]$. The shifted interval $[a-2, b-1]$ is denoted with a dotted line. The intervals at which $\phi_1 \mathcal{U}_{[1,2]} \phi_2$ holds can be seen below the dotted interval. Note that the interval $[a_1, b_1]$ does not satisfy the condition that $b_1 \leq a-1$ and is therefore not part of the result.

For the remaining junctors eval is defined recursively as follows, where ϕ_1 and ϕ_2 denote MTNCQs with the same n answer variables:

$$
\begin{aligned}
\mathrm{eval}(\phi_1 \wedge \phi_2, \mathfrak{I}_{\mathcal{A}}^{\mathrm{fin}}) &:= \{(\mathbf{a}, \mathbf{i}_1 \cap \mathbf{i}_2) \mid (\mathbf{a}, \mathbf{i}_1) \in \mathrm{eval}(\phi_1, \mathfrak{I}_{\mathcal{A}}^{\mathrm{fin}}) \text{ and } (\mathbf{a}, \mathbf{i}_2, \mathfrak{I}_{\mathcal{A}}^{\mathrm{fin}}) \in \mathrm{eval}(\phi_2, \mathfrak{I}_{\mathcal{A}}^{\mathrm{fin}})\} \\
\mathrm{eval}(\phi_1 \vee \phi_2, \mathfrak{I}_{\mathcal{A}}^{\mathrm{fin}}) &:= \mathrm{eval}(\phi_1, \mathfrak{I}_{\mathcal{A}}^{\mathrm{fin}}) \cup \mathrm{eval}(\phi_2, \mathfrak{I}_{\mathcal{A}}^{\mathrm{fin}}) \\
\mathrm{eval}(\neg \phi_1, \mathfrak{I}_{\mathcal{A}}^{\mathrm{fin}}) &:= \mathrm{gaps}(\mathrm{eval}(\phi_1, \mathfrak{I}_{\mathcal{A}}^{\mathrm{fin}})) \\
\mathrm{eval}(\Box_{[\ell,r]} \phi_1, \mathfrak{I}_{\mathcal{A}}^{\mathrm{fin}}) &:= \{(\mathbf{a}, [a_1 - \ell, b_1 - r]) \mid (\mathbf{a}, [a_1, b_1]) \in \mathrm{eval}(\phi_1, \mathfrak{I}_{\mathcal{A}}^{\mathrm{fin}})\} \\
\mathrm{eval}(\Diamond_{[\ell,r]} \phi_1, \mathfrak{I}_{\mathcal{A}}^{\mathrm{fin}}) &:= \{(\mathbf{a}, [a_1 - r, b_1 - \ell]) \mid (\mathbf{a}, [a_1, b_1]) \in \mathrm{eval}(\phi_1, \mathfrak{I}_{\mathcal{A}}^{\mathrm{fin}})\}
\end{aligned}
$$

The metric until operator requires a bit more work. If the interval includes 0 it automatically includes all time points at which ϕ_2 holds. In all other cases it also depends on ϕ_1. Therefore, we make a case distinction, where $\ell > 0$:

$$
\begin{aligned}
\mathrm{eval}(\phi_1 \mathcal{U}_{[0,0]} \phi_2, \mathfrak{I}_{\mathcal{A}}^{\mathrm{fin}}) &:= \mathrm{eval}(\phi_2, \mathfrak{I}_{\mathcal{A}}^{\mathrm{fin}}) \\
\mathrm{eval}(\phi_1 \mathcal{U}_{[0,r]} \phi_2, \mathfrak{I}_{\mathcal{A}}^{\mathrm{fin}}) &:= \mathrm{eval}(\phi_1 \mathcal{U}_{[0,0]} \phi_2, \mathfrak{I}_{\mathcal{A}}^{\mathrm{fin}}) \cup \mathrm{eval}(\phi_1 \mathcal{U}_{[1,r]} \phi_2, \mathfrak{I}_{\mathcal{A}}^{\mathrm{fin}}) \\
\mathrm{eval}(\phi_1 \mathcal{U}_{[\ell,r]} \phi_2, \mathfrak{I}_{\mathcal{A}}^{\mathrm{fin}}) &:= \{(\mathbf{a}, [a_1, b_1] \cap [a_2 - r, b_2 - \ell]) \mid \\
&\qquad (\mathbf{a}, [a_1, b_1]) \in \mathrm{islands}(\mathrm{eval}(\phi_1, \mathfrak{I}_{\mathcal{A}}^{\mathrm{fin}})), \\
&\qquad (\mathbf{a}, [a_2, b_2]) \in \mathrm{eval}(\phi_2, \mathfrak{I}_{\mathcal{A}}^{\mathrm{fin}}), \text{ and} \\
&\qquad b_1 \geq a_2 - 1\}
\end{aligned}
$$

For the last case, a tuple $(\mathbf{a}, \mathbf{i})$ can only be a result if (i) $\mathbf{i}$ is a (sub)interval of $[a_1, b_1]$ at which $\mathbf{a}$ is an answer to ϕ_1; (ii) the time points in $\mathbf{i}$ are not too far away from time points in $[a_2, b_2]$ at which $\mathbf{a}$ is an answer to ϕ_2; this is guaranteed by the shifting of $[a_2, b_2]$ by the bounds of until; and (iii) the interval $[a_1, b_1]$ holds until at least the beginning of the interval $[a_2, b_2]$. Note that because we apply the islands function here, it is sufficient to look at a pair of intervals ($[a_1, b_1]$ and $[a_2, b_2]$) at the time. Without the islands function we would have to consider multiple intervals where $\mathbf{a}$ is an answer to ϕ_1 at the same time. An illustration of the evaluation of until is given in Figure 7.2. The metric since operator can be dealt with in a similar way.

Lemma 7.3. Let $\mathfrak{I}$ be a model of a consistent $\mathcal{TELH}_{\bot}^{\Diamond,\mathsf{lhs}}$-KB, $\phi(\mathbf{x})$ a rooted MTNCQ. Then it holds that $(\mathbf{a},[i,i]) \in \mathrm{eval}(\phi,\mathfrak{I}_{\mathcal{A}}^{\mathrm{fin}})$ iff $\mathfrak{I}_{\mathcal{A}}, i \models \phi(\mathbf{a})$.

Proof. Here we only provide the most complex case of until, the cases for the remaining operators can be done in a similar manner.

For easier notation, in the proof we write $(\mathbf{a},\mathbf{i}) \in \mathrm{eval}(\phi,\mathfrak{I}_{\mathcal{A}}^{\mathrm{fin}})$ to denote that there exists $(\mathbf{a},\mathbf{i}') \in \mathrm{eval}(\phi,\mathfrak{I}_{\mathcal{A}}^{\mathrm{fin}})$ with $\mathbf{i} \subseteq \mathbf{i}'$.

($\leftarrow$): Suppose $\mathfrak{I}_{\mathcal{A}}, i \models (\phi_1 \mathcal{U}_{[\ell,r]} \phi_2)(\mathbf{a})$ and $1 \leq \ell \leq r$. Then there exists $k \in [\ell,r]$ with $\mathfrak{I}_{\mathcal{A}}, i+k \models \phi_2(\mathbf{a})$ and for all $0 \leq j < k$: $\mathfrak{I}_{\mathcal{A}}, i \models \phi_1(\mathbf{a})$. By induction we have that $(\mathbf{a},[i+k,i+k]) \in \mathrm{eval}(\phi_2,\mathfrak{I}_{\mathcal{A}}^{\mathrm{fin}})$ and $(\mathbf{a},[i+j,i+j]) \in \mathrm{eval}(\phi_1,\mathfrak{I}_{\mathcal{A}}^{\mathrm{fin}})$ for all $0 \leq j < k$, which implies $(\mathbf{a},[i,i+k-1]) \in \mathrm{islands}(\mathrm{eval}(\phi_1,\mathfrak{I}_{\mathcal{A}}^{\mathrm{fin}}))$. This satisfies the last condition since $b_1 = i+k-1 \geq i+k-1 = a_2-1$. What remains to show is that

$$i \in [i,i+k-1] \cap [i+k-r, i+k-\ell].$$

The first interval contains i, since $1 \leq l \leq k$. For the second interval, since $k \in [\ell,r]$, we know that

$$i+k-r \leq i+r+r = i \qquad \text{and} \qquad i+k-\ell \geq i+\ell-\ell = i,$$

and hence, it also contains i. Therefore $(\mathbf{a},[i,i]) \in \mathrm{eval}(\phi_1 \mathcal{U}_{[\ell,r]} \phi_2, \mathfrak{I}_{\mathcal{A}}^{\mathrm{fin}})$.

($\rightarrow$): Suppose $(\mathbf{a},[i,i]) \in \mathrm{eval}(\phi_1 \mathcal{U}_{[\ell,r]} \phi_2, \mathfrak{I}_{\mathcal{A}}^{\mathrm{fin}})$. Then $i \in ([a_1,b_1] \cap [a_2-\ell, b_2-r])$ for some $(\mathbf{a},[a_1,b_1] \in \mathrm{eval}(\phi_1,\mathfrak{I}_{\mathcal{A}}^{\mathrm{fin}})$ and $(\mathbf{a},[a_2,b_2] \in \mathrm{eval}(\phi_2,\mathfrak{I}_{\mathcal{A}}^{\mathrm{fin}})$. By the third condition we know that $b_1 \geq a_2-1$ and we chose the smallest possible interval for $[a_1,b_1]$ by setting $a_1 = i$ and $b_1 = a_2-1$. Moreover, we consider the smallest possible case for $[a_2,b_2]$ by setting $b_2 = a_2$. We obtain

$$i \in ([i,a_2-1] \cap [a_2-r, a_2-\ell]).$$

Since i is contained in the intersection we must have $a_2 - r \leq i$ and $i \leq a_2 - \ell$. Hence, a_2 has to be somewhere in the interval $[i+\ell, i+r]$ and there exists $k \in [\ell,r]$ such that $(\mathbf{a},[i+k,i+k]) \in \mathrm{eval}(\phi_2,\mathfrak{I}_{\mathcal{A}}^{\mathrm{fin}})$ and $(\mathbf{a},[i,i+k-1]) \in \mathrm{eval}(\phi_1,\mathfrak{I}_{\mathcal{A}}^{\mathrm{fin}})$. By induction we obtain that there exists $k \in [\ell,r]$ such that $\mathfrak{I}_{\mathcal{A}}, i+k \models \phi_2(\mathbf{a})$ and for all $0 \leq j \leq k$: $\mathfrak{I}_{\mathcal{A}}, i+j \models \phi_1(\mathbf{a})$. $\square$

Based on these definitions QUELK constructs a SQL query that returns every possible answer tuple and the intervals at which it is valid. To select all answers at a given time point t, in the last step only answers are selected that are valid in an interval that includes t. At the time of this writing, the metric temporal operators $\mathcal{U}$ and $\mathcal{S}$ are not supported by QUELK. An implementation might be added in future versions.

7.2. A Dataset for Patient Selection

To evaluate our system we chose the dataset from Track 1 of the 2018 N2C2 cohort selection challenge[4] which was set out to answer the question: *"Can NLP systems use narrative medical records to identify which patients meet selection criteria for clinical trials?"* [SFS+19]. It contains anonymized EHRs from 288 patients with 2 – 5 records per patient. All patients in the dataset have diabetes and most have a high risk for heart disease. Since EHRs are formulated as free texts, they can vary in between each other. However, most of the time they have roughly the same structure. In the following we go through an example record given in Example 7.4 from the N2C2 dataset to give an impression of the information it contains and the way it is structured.

[4]https://portal.dbmi.hms.harvard.edu/projects/n2c2-2018-t1/

Example 7.4. An example patient record from patient 182 in the N2C2 dataset. Due to the length of the record some parts were omitted, denoted by [...]. The record is from an actual record, but all names and dates in the record were changed to guarantee anonymity.

```
Record date: 2133-02-11

[...]

Patient: Daniel Doherty
Medical Record Number: 767 02 38
Room: E5316Y
Date of Admission: 2/10/33
Attending Physician: Thomas Cotton
PCP: Robert Bailey (Central WV)
Code Status: FULL

Source: Patient unable to give history b/c of mental status, Medical Record

ID/CC: 68yoM h/o schizophrenia, medication non-compliance, presents s/p cardiac arrest.

History of Present Illness:
Per Dibble General Hospital Records, the patient was admitted on 2/04/33 after collapsing in the mall. His total time down is unclear, but the patient had an AED administered and he was shocked x2. When EMTs arrived, they found the patient to be in VF w/ agonal breathing. CPR was administered and he was shocked x3. After the 3rd defibrillation, the patient awoke and was in NSR. The patient could not recall any of the events.

[...]

Review of Systems:
Unavailable secondary to mental status.

Past Medical/Surgical History:
Hypercholesterolemia
Hypertension
Schizophrenia
H/o abnormal SPEP

[...]

Allergies:
NKDA

Medications at home:
Aspirin 81mg qd
HCTZ 25mg qd
[...]
Medications on transfer:
```

```
[...]
Social History:
Lives alone
TOB: quit 20y prior
ETOH: occasional
ILLICITS: unknown

Family History:
M: unknown
F: h/o MI

Physical Examination:
GEN: somnolent, easily arousable w/ manual stimulation, answers some questions,
 but not very interactive, breathing at 20, mildly labored

[...]

EKG SB at 57bpm, Q in v1, III/AvF, left axis, TWI in v3-v6, new from previous.
QTc 430ms (manually calculated)
CXR enlarged cardiac silhouette, L basilar opacity.

Impression/Plan:
68yoM h/o HTN/HL, schizophrenia, medication non-compliance presents s/p VF
arrest, NSTEMI, AMS. Patient likely had ischemic VT/VF. Per outside records,
the patient does not seem to have ongoing activity of his acute coronary
syndrome. His mental status continues to be poor, and is likely multifactorial;
 my suspicion is that he does have some level of anoxic brain injury as the
dominant pathology.

NSTEMI: npo p MN for cardiac cath in AM. Patient not able to consent at this
point, and does not have any clear family members that can do this for him.
Will follow his mental status exam.
       -aspirin, metoprolol, lipitor, heparin IV.
       -can add ACEi when he is taking po s and if BPs need further titration.
       -mucomyst for renal prophylaxis.
       -for risk factors, check a1c.
S/P VF ARREST: likely in the setting of NSTEMI
       -cath as above.
       -nl QTc
       -s/p amio IV load. C/w amiodarone po load. If reversible lesion, likely
       can d/c.
AMS: anoxic brain injury; current somnolence likely from narcotics.
       -appreciate neurology consultation
       -head CT
       -c/w home zyprexa. Haldol prn.

[...]

Raul Quilici, MD
Pager: #99591
```

Each EHR contains a record date (Line 1 in Example 7.4), which denotes when it was recorded. It is followed by some general information like the names of the patient and the physicians that treated the patient (Line 5-11). The "Chief Complaint" (CC) of the patient is given in Line 15: It is a 68 year old male with a history of schizophrenia. He is presented at the hospital with a "status post" (s/p) cardiac arrest[5]. The situation before he was admitted to the hospital is explained in Line 17-18. In the past the patient was suffering already from schizophrenia and other findings (Line 25-29). He has "no known drug allergies" (NKDA) (Line 34). The EHR contains also information about the medication the patient is taking regularly at home and the medications he got during transfer to the hospital (Line 36-41). The "Social History" of the patient tells us that he lives alone and stopped consuming tobacco 20 years ago (Line 42-46). For some diseases the family history is relevant, in this case the patients father has a history of myocardial infarction[6] (Line 48-50) (MI). The physical examination gives some details about his current condition and includes measurements that were taken (Line 52-58). Finally the paragraph "Impression/Plan" summarizes the impression the physician got of the patient, in particular that the patients cardiac arrest was likely a result of a certain kind of myocardial infarction (NSTEMI) (Line 60-75).

The given patient has 5 EHRs in the dataset, covering a time span of about six years.

Criteria

In the N2C2 challenge 13 criteria were chosen, inspired by real studies from ClinicalTrials.gov. Each patient in the dataset was then categorized by medical experts into met or not met for each criterion. Unfortunately, none of the criteria contains any form of negation. In our experiment we focused on 5 criteria that contain a temporal dimension. Their abbreviations and definitions are the following:

ADVANCED-CAD Advanced coronary artery disease[7] (CAD), where 'advanced' is defined as having two or more of the following:

- taking two or more medications to treat CAD
- history of myocardial infarction (MI)
- currently experiencing angina[8]
- ischemia[9] in the past or present

MAJOR-DIAB Major diabetes-related complication where 'major complication' (as opposed to 'minor complication') is defined as any of the following that are a result of (or strongly correlated with) uncontrolled diabetes:

- amputation

[5] Cardiac arrest is a sudden loss of blood flow resulting from the failure of the heart to pump effectively [Inc20b].

[6] A myocardial infarction (MI), also known as a heart attack, occurs when blood flow decreases or stops to a part of the heart, causing damage to the heart muscle [Inc20d].

[7] CAD is the most common type of heart disease and happens when the arteries that supply blood to heart muscle become hardened and narrowed [Med20].

[8] Angina, also known as angina pectoris, is chest pain or pressure, usually due to not enough blood flow to the heart muscle [Inc20a].

[9] Ischemia is a restriction in blood supply to tissues, causing a shortage of oxygen that is needed to keep tissue alive [Inc20c].

- kidney damage
- skin conditions
- retinopathy[10]
- nephropathy[11]
- neuropathy[12]

MI-6MOS MI in the past 6 months.

ASP-FOR-MI Use of aspirin to prevent MI.

KETO-1YR Diagnosis of ketoacidosis[13] in the past year.

For each criterion the dataset contains a *ground-truth (GT)-labeling*: Each patient is labeled with the criteria the patient satisfies, according to the medical experts that annotated the EHRs. For instance, patient 182 from Example 7.4 satisfies ADVANCED-CAD, MAJOR-DIABETES and MI-6MOS.

7.3. Setup

The overall setup of the experiment is illustrated in Figure 7.5. In the following we describe the construction of the $\mathcal{TELH}_{\bot}^{\Diamond,\mathsf{lhs},-}$-KB from the EHRs and the translation of the eligibility criteria to MTNCQs.

7.3.1. Creation of the KB

The TBox contains all of SNOMED CT. Additionally, we add the *GT-concepts* `ADVANCED-CAD`, `MAJOR-DIAB`, `MI-6MOS`, `ASP-FOR-MI` and `KETO-1YR`, the concept `Patient`, and the role `diagnosedWith` to the signature. The concept `Patient` is asserted to be rigid, formally $\Diamond\!\!\!\!\!\Diamond\,\mathtt{Patient} \sqsubseteq \mathtt{Patient}$. In a more general setting additional temporal axioms could be added, for example for patients diagnosed with some incurable disease. However, for the criteria we use in this evaluation temporal axioms are unnecessary.

The ABox is constructed by adding the diagnoses of each patient. Each EHR of each patient is processed by the medical term tagging tool MetaMapLite[14] which returns a list of SNOMED CT concepts found in the record. The information are then added to the ABox in the following way for each patient p and EHR e of p with timestamp i:

- p is asserted to be a patient at i by the assertion $\mathtt{Patient}(p, i)$.
- Then for each concept C found in e, the assertions $\mathtt{diagnosedWith}(p, a, i)$ and $C(a, i)$ are added, where a is a fresh individual.

[10] Retinopathy is any damage to the retina of the eyes, which may cause vision impairment [Inc20g].
[11] Nephropathy, is damage to or disease of a kidney [Inc20e].
[12] Neuropathy is a general term describing disease affecting the peripheral nerves, meaning nerves beyond the brain and spinal cord [Inc20f].
[13] Diabetic ketoacidosis is a serious complication of diabetes that occurs when your body produces high levels of blood acids called ketones [MR20].
[14] `https://metamap.nlm.nih.gov/MetaMapLite.shtml`

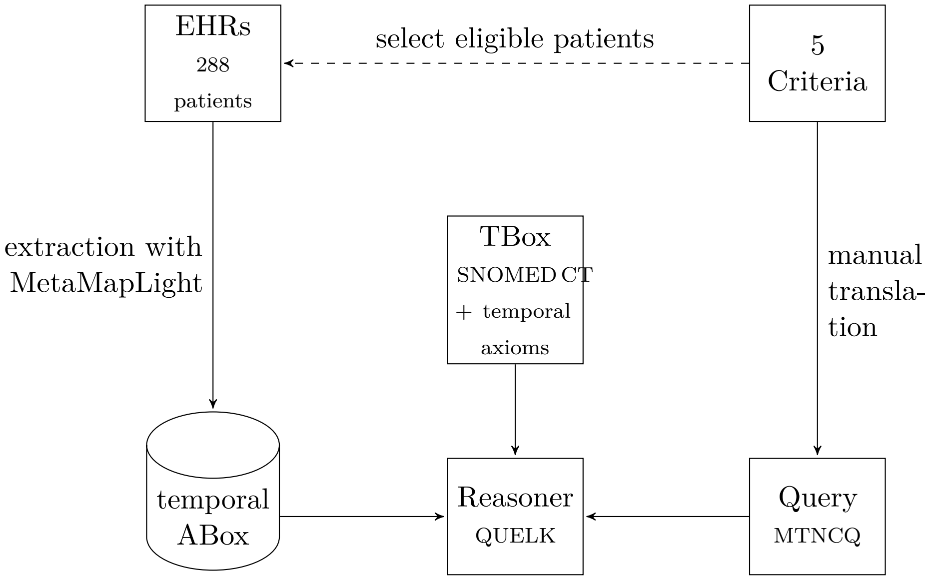


Figure 7.5.: An illustration of the setup of the experiment. The SNOMED CT terms in 288 patient records are tagged using MetaMapLite and transformed to temporal facts. SNOMED CT is used as a TBOX together with auxiliary axioms. 5 temporal criteria are manually translated to MTNCQs and then the minimal-world answers are computed. The results are evaluated against the GT-labeling of the data.

Currently there is no distinction between different roles. This could be improved in future versions, for example by connecting the patient to medications with a **`takes`**-role instead.

The GT-labeling assumes that the criterion is evaluated at the time point of the last available EHR for each patient. To simplify the queries, we shift the EHRs of all patients such that they share the last time point $\hat{i}$. The GT-concepts are used to represent the GT-labeling in the ontology.
For each $C \in \{\texttt{ADVANCED-CAD}, \texttt{MAJOR-DIAB}, \texttt{MI-6MOS}, \texttt{ASP-FOR-MI}, \texttt{KETO-1YR}\}$ and patient p an assertion of the form $C(p, \hat{i})$ is added if p satisfies C according to the GT-labeling.

7.3.2. Translation of the criteria

We translated the criteria manually to (disjunctions of) MTNCQs. A list of the SNOMED CT concepts that were used for the respective findings is given in Table 7.6. For example, ADVANCED-CAD corresponds to the following query:

$$\begin{aligned}\phi_{\texttt{ADVANCED-CAD}}(x) := {} & [\diamondminus\exists y.(\mathtt{Patient}(x)\wedge\mathtt{diagnosedWith}(x,y)\wedge\mathtt{CAD}(y))]\\ & \wedge\Big([\diamondminus\exists y.(\mathtt{Patient}(x)\wedge\mathtt{diagnosedWith}(x,y)\wedge\mathtt{MI}(y))]\\ & \vee[\diamondminus_{[-30d,0d]}\exists y.(\mathtt{Patient}(x)\wedge\mathtt{diagnosedWith}(x,y)\wedge\mathtt{Angina}(y))]\\ & \vee[\diamondminus\exists y.(\mathtt{Patient}(x)\wedge\mathtt{diagnosedWith}(x,y)\wedge\mathtt{Ischemia}(y))]\Big)\end{aligned}$$

To obtain better results, the query is relaxed in different ways: Firstly the criterion "taking two or more medications" is ignored, because MetaMapLite does not tag medications; secondly, it is sufficient if a patient satisfies one additional criterion indicating an advanced CAD. The original query requires two or more indicators. Thirdly, "currently experiencing angina" is relaxed to angina in the last 30 days. In the translation of MAJOR-DIAB the criteria were again assumed to be satisfied if they are satisfied at some time point in the past:

$$\begin{aligned}\phi_{\texttt{MAJOR-DIAB}}(x) := {} & [\diamondminus\exists y.(\mathtt{Patient}(x)\wedge\mathtt{diagnosedWith}(x,y)\wedge\mathtt{DiabetesMellitus}(y))]\\ & \wedge\Big([\diamondminus\exists y.(\mathtt{Patient}(x)\wedge\mathtt{diagnosedWith}(x,y)\wedge\mathtt{Amputation}(y))]\\ & \vee[\diamondminus\exists y.(\mathtt{Patient}(x)\wedge\mathtt{diagnosedWith}(x,y)\wedge\mathtt{KidneyDamage}(y))]\\ & \vee[\diamondminus\exists y.(\mathtt{Patient}(x)\wedge\mathtt{diagnosedWith}(x,y)\wedge\mathtt{SkinConditions}(y))]\\ & \vee[\diamondminus\exists y.(\mathtt{Patient}(x)\wedge\mathtt{diagnosedWith}(x,y)\wedge\mathtt{Retinopathy}(y))]\\ & \vee[\diamondminus\exists y.(\mathtt{Patient}(x)\wedge\mathtt{diagnosedWith}(x,y)\wedge\mathtt{Neuropathy}(y))]\\ & \vee[\diamondminus\exists y.(\mathtt{Patient}(x)\wedge\mathtt{diagnosedWith}(x,y)\wedge\mathtt{Nephropathy}(y))]\Big)\end{aligned}$$

MI-6MOS and KETO-1YR were translated directly:

$$\begin{aligned}\phi_{\texttt{MI-6MOS}}(x) &:= [\diamondminus_{[-183d,0d]}\exists y.(\mathtt{Patient}(x)\wedge\mathtt{diagnosedWith}(x,y)\wedge\mathtt{MI}(y))]\\ \phi_{\texttt{KETO-1YR}}(x) &:= [\diamondminus_{[-365d,0d]}\exists y.(\mathtt{Patient}(x)\wedge\mathtt{diagnosedWith}(x,y)\wedge\mathtt{Ketoacidosis}(y))]\end{aligned}$$

In the criterion ASP-FOR-MI the condition that aspirin should be used to prevent MI was relaxed to the following query asking for 'patients that have a history of MI and a history of (using) Aspirin':

$$\begin{aligned}\phi_{\texttt{ASP-FOR-MI}}(x) := {} & [\diamondminus\exists y.(\mathtt{Patient}(x)\wedge\mathtt{diagnosedWith}(x,y)\wedge\mathtt{MI}(y))]\\ & \wedge[\diamondminus\exists y.(\mathtt{Patient}(x)\wedge\mathtt{diagnosedWith}(x,y)\wedge\mathtt{Aspirin}(y))]\end{aligned}$$

Name of the finding	SCTID
CAD (Coronary artery disease)	`http://snomed.info/id/49601007`
MI (Myocardial infarction)	`http://snomed.info/id/22298006`
Angina	`http://snomed.info/id/194828000`
Ischemia	`http://snomed.info/id/52674009`
Diabetes Mellitus	`http://snomed.info/id/73211009`
Amputation	`http://snomed.info/id/81723002`
Kidney Damage	`http://snomed.info/id/40095003`
Skin Conditions	`http://snomed.info/id/422000003`
Retinopathy	`http://snomed.info/id/29555009`
Neuropathy	`http://snomed.info/id/386033004`
Nephropathy	`http://snomed.info/id/90708001`

Table 7.6.: The SCTIDs (SNOMED CT IDs) that correspond to the findings on the left.

7.4. Experiments

In the following let $\mathcal{K}$ denote the $\mathcal{TELH}_{\bot}^{\Diamond,\mathsf{lhs}}$-KB constructed from the EHRs in the way described in Section 7.3.

To see the impact of using temporal information and reasoning, we distinguish two different settings for each criterion: In the *temporal setting*, $\mathcal{K}$ and the queries are used with all their temporal information.

In the *atemporal setting*, all temporal information are ignored in $\mathcal{K}$ as well as in the query. More specifically, a given assertion $C(a, i)$ is interpreted as an atemporal assertion $C(a)$ and similar for role assertions, while in MTNCQs, all temporal operators are ignored. For instance, the atemporal version of $\phi_{\texttt{ASP-FOR-MI}}$ is just the conjunction of its NCQs:

$$\begin{aligned}&\exists y.(\texttt{Patient}(x) \land \texttt{diagnosedWith}(x, y) \land \texttt{MI}(y))\\ \land &\exists y.(\texttt{Patient}(x) \land \texttt{diagnosedWith}(x, y) \land \texttt{Aspirin}(y))\end{aligned}$$

For each query ϕ the results are evaluated using three measures. The *precision* of ϕ (in $\mathcal{K}$) is the ratio of the number of returned eligible patients to the total number of returned patients. The higher the precision of ϕ, the higher the probability that a selected patient is indeed eligible. The *recall* of a ϕ (in $\mathcal{K}$) is the ratio of returned eligible patients to the total number of eligible patients. A higher recall means that fewer eligible patients were missed by the query. The perfect system would have recall and precision of 1. To get a good system, both precision and recall have to be as high as possible. They are however connected: On the one hand, to increase the precision, we could focus on patients that are eligible with a very high probability. This comes at the risk of excluding many patients that are also eligible, and therefore often decreases the recall. On the other hand, if a criterion is relaxed, then the recall will increase, but the precision

	Atemporal			Temporal		
Criterion	**P**	**R**	F_1	**P**	**R**	F_1
ADVANCED-CAD	0.69	0.94	0.79	0.7	0.93	0.8
MAJOR-DIABETES	0.69	0.88	0.78	0.69	0.88	0.78
MI-6MOS	0.13	0.73	0.22	0.21	0.62	0.31
KETO-1YR	0.02	1	0.04	0	0	–
ASP-FOR-MI	–	–	–	–	–	–

Table 7.7.: Precision (P), recall (R) and the F_1-score for the five criteria.

will decrease. To keep both measures in balance the F_1-*measure* can be used. It is the harmonic mean of precision and recall, so a perfect system would have an F_1-score of 1.

To compute the measures we used the GT-concepts in the KB. For instance, all patients satisfying MAJOR-DIAB according to the GT-labeling can be selected by the GT-query for $\phi_{\texttt{MAJOR-DIAB}}$, which selects all individuals that belong to the GT-concept `MAJOR-DIAB`:

$$\phi_{\texttt{MAJOR-DIAB-GT}}(x) = \texttt{MAJOR-DIAB}(x)$$

Then for each query ϕ with GT-query $\phi_{\texttt{GT}}$, we can compute precision, recall and F_1-score of ϕ in $\mathcal{K}$ the following way:

$$\begin{aligned}
\texttt{precision}(\phi) &:= \frac{|\operatorname{mwa}(\phi,\mathcal{K}) \cap \operatorname{mwa}(\phi_{\texttt{GT}},\mathcal{K})|}{|\operatorname{mwa}(\phi,\mathcal{K})|}\\
\texttt{recall}(\phi) &:= \frac{|\operatorname{mwa}(\phi,\mathcal{K}) \cap \operatorname{mwa}(\phi_{\texttt{GT}},\mathcal{K})|}{|\operatorname{mwa}(\phi_{\texttt{GT}},\mathcal{K})|}\\
F_1(\phi) &:= 2 \cdot \frac{\texttt{precision}(\phi) \cdot \texttt{recall}(\phi)}{\texttt{precision}(\phi) + \texttt{recall}(\phi)}
\end{aligned}$$

The results for the different criteria and the atemporal and temporal queries are shown in Table 7.7. As expected using temporal queries instead of atemporal queries increases the precision, because a patients can be selected based on how long ago his diagnoses are in the past. This reduces the recall on the other hand. By looking at the F_1-score, which is generally higher for the temporal queries, we can conclude that the precision increases more that the recall decreases in comparison to the atemporal queries. So using temporal information in a patient selection system yields better results.

In ADVANCED-CAD and MAJOR-DIABETES both the atemporal and the temporal queries get high scores. A major reason for this might be that both criteria involve concepts MetaMapLite is very good at tagging in. In relative terms, there is not much difference between the two settings. When looking at the structure of the queries, we can see that the temporal query contains a lot of 'history of' constructions. A (sub-)query that starts with $\diamondminus$ is essentially equivalent to an atemporal query that is answered over a KB in which all time points are merged together. In ADVANCED-CAD the sub-query asking for ‘a diagnosis of angina within the last 30 days’ causes the slight difference between the temporal and atemporal settings.

The results for MI-6MOS are quite bad, even though the query has a very simple structure. On the one hand, this could be because the tagger was not configured well enough. There exist many different types and abbreviations for MIs[15] and MetaMapLite was not fine-tuned to detect all of them. With some tweaking of the tagging process the results could be improved. On the other hand, MIs are often not mentioned directly in the EHR, but just different symptoms that suffice to recognize a MI for a medical expert, but are very hard to infer for a computer. Here a more sophisticated modeling of the EHRs can make a difference. Apart from the low absolute scores for MI-6MOS, the relative difference between the temporal and the atemporal setting, especially in the precision, suggests that results are much better when the temporal dimension of the EHRs is taken into account when selecting patients.

KETO-1YR has the same structure as MI-6MOS, but the difference is that in the whole dataset there is just one patient that satisfies this criterion. The atemporal query selected the correct patient among 54, while the temporal query returned only 4 patients, but missed the correct one. This means the atemporal query just incidentally returned the correct patient, because the time point of the diagnoses of ketoacidosis must have been more than one year ago. Otherwise the temporal query would have returned the correct patient as well.

The last criterion ASP-FOR-MI did not give any results in any setting. As mentioned this is due to the missing medications in the KB. If MetaMapLite would be setup to also tag medications (which it probably can), then the results here would certainly improve. It remains unclear how big a further improvement could be if the KB contained causal relations between diagnoses and medications.

7.5. Discussion and Related Work

With our experiment we got promising first results for our prototypical system. Of course, a number of things could be improved in the future:

1. Since the tagger was not fine-tuned for the processing of EHRs, some concepts are never tagged, for example medications. In many cases the resolving of abbreviations can be improved, for example for the different kinds of MIs.
2. Currently, paragraphs like “past medical history”, “social history”, or “family history” (Line 25-50 in Example 7.4) are processed just like every other part, even though the concepts occurring inside the paragraph are not valid at the current time point, but before or refer to family members instead of the patient. A more fine-tuned tagging system should put the concepts found in these paragraphs in the right context.
3. If criteria are not translated manually, more temporal knowledge should be put into the KB. For our translation of the MAJOR-DIAB criterion, we used the knowledge that diabetes is incurable, so we formulated the query accordingly. If

[15]For the interested reader, all the following abbreviations refer to a kind of myocardial infarction: STEMI, NSTEMI, AMI, DMI, MI, NQMI, ACS.

the queries are translated automatically, then such background knowledge might not be available and should instead be contained in the KB.

In the N2C2 challenge the mean F_1-score for all submissions was 0.799 with the scores ranging from 0.2117 to 0.91. Stubbs, Filannino, Soysal, Henry, and Uzuner found that among the 10 best performing systems 4 were rule based and 6 used rules in a combination with machine learning [SFS+19]. Additionally, they found that systems performed better in general when a medical export was present in the team. With the resources available to us, we cannot beat the winning systems, yet. But the fact that rule based approaches performed very well on this problem supports our approach. For an in depth comparison of different statistical and rule-based approaches, we recommend the article accompanying the system of the winning team of the N2C2 challenge [OKK+19].

Previous work has already considered using ontologies for patient selection for clinical trials before. [PCD+07] worked with patient records from Columbia University Medical Center that were recorded using the MED ontology [CCH+94]. They mapped MED to SNOMED CT using a semi-automated approach that was guided by domain experts. The patient records were then integrated using a pattern matching rule-based approach. They showed that it is actually possible to find patient matches using an ontology, and were able to scale their approach to one year of patient data.

[BCZ+10] focused on 200 trials about prostate cancer and annotated them manually with UMLS concepts. As formal basis, they use OWL (which is based on DLs) together with SWRL rules[16], which allows them to add rules for temporal relations. They then load one patient at a time into the ontology and query the studies that the patient is eligible for. Their approach allows traceability of the results, which is a very desirable property. While they demonstrate that patients can be selected using their formal framework, they assume that the data are already formalized.

[TSC11] further analyzed and modeled the temporal patterns that occur in patient data. To represent them they introduce the OWL-based CNTRO 2.0 ontology for clinical narratives. Later, [CT15] classified most temporal statements occurring in descriptions of clinical trials and clinical guidelines into 16 basic temporal patterns that are expressible in CNTRO 2.0. Unfortunately, CNTRO 2.0 is not suitable for temporal query answering, since it can express temporal statements, but does not provide a temporal semantics and allows only rudimentary temporal inferences.

Other approaches to model temporal medical data use graph- or constraint-based formalisms to representing and reasoning with temporal statements [HZP+05; BJ18].

For a survey regarding also non-temporal, non-logical proposals for automated processing of EHRs and other medical data, see [KP14].

[16] `https://www.w3.org/Submission/SWRL/`

Chapter 8.

Conclusion

In the course of the thesis we have seen that the open-world assumption is not suitable in the medical setting, due to the absence of negative information. Apart from the medical setting, many users of semantic technologies that have no prior experience are surprised by the results obtained in pure open-world semantics. When they first try to model something, they are puzzled because different conclusions are entailed than they intended. That is because standard semantics are based on all possible models without any restrictions, which often includes many unintended models the user never thought of.

On the other end of the spectrum is the relational-database world, in which answers are computed over a single model represented by the entries in the database. While this concept is easier to grasp, it requires the user to explicitly model everything to make the data consistent. This is not always possible, especially when ontological knowledge is added. For instance, if a patient is a **`SkinCancerPatient`**, there has to exist an explicit **`SkinCancer`** object in the KB, because the domain of the interpretation is fixed.

Non-monotonic formalisms are usually somewhere in between these two assumptions: They identify a specific subset of 'intuitively' behaving models. What is considered intuitive is highly dependent on subjective factors, such as the application domain one wants to model and the user itself. Therefore, a plethora of different proposals for non-monotonic semantics exist and every single one of them can be motivated by some application scenario, in which the respective set of admissible models is the set of models a potential user intends. In this thesis, we have developed novel non-monotonic semantics and motivated them by the patient selection problem in which patients, that satisfy criteria from a given clinical trial, need to be selected, based on their EHRs. In Section 8.1 we provide a brief summary of the contributions of this thesis and mention some future work in Section 8.2.

8.1. Main Results

In this thesis we have introduced minimal-world semantics, which provide non-monotonic negation. The semantics have been motivated by the patient selection problem in Chapter 3, which requires queries to be answered over medical data that only rarely contain negative facts. In contrast to many other non-monotonic formalisms, our semantics have been tailored especially with OMQA in mind and hence can deal well with anonymous individuals.

In minimal-world semantics queries are answered over a *minimal universal model*, which we consider the 'intuitively expected' model of the user. The assumption underlying the minimality is that if two statements are present and the first is more specific than the second, then the second refers to the same individual as the first. For instance consider the statements: 'Alice has a cold.' and 'Unfortunately, Alice is sick.' With the minimal-world assumption we conclude that Alice is sick because she has a cold and not because of some other sickness that has not been mentioned before. We characterized minimality in terms of endomorphisms and showed that this is closely related to the notion of cores.

For $\mathcal{ELH}_\bot$ we have provided a construction of a minimal universal model and have thereby shown existence and uniqueness (up to isomorphism) of the minimal universal model for consistent $\mathcal{ELH}_\bot$-KBs. To answer NCQs we have extended the combined rewriting approach for CQs [EOŠ+12; BO15] by filters to take into account negation. We have shown that minimal-world answers to a given NCQ can be computed in polynomial time in data complexity.

In Chapter 4 we have introduced $\mathcal{TELH}_\bot^{\diamondsuit,\mathsf{lhs}}$, an extension of $\mathcal{ELH}_\bot$ by temporal operators on the left-hand side of GCIs. The temporal operators can be used to express that a given disease is not curable, i.e. if a patient is diagnosed once, then this diagnosis will be valid at all succeeding time points as well. Moreover, the logic features the convex diamond operators, both in an unbounded and a bounded variant to express statements like 'If a patients leg was broken now and 8 days ago, then the leg was also broken in between the two time points.' Apart from the novel temporal operators in $\mathcal{TELH}_\bot^{\diamondsuit,\mathsf{lhs}}$ the ABox is allowed to contain gaps, i.e. intervals of time points not occurring in the ABox. The behavior of objects can be interpolated with the diamond operators. To obtain an efficient reasoning procedure, we have shown that each gap can be represented compactly by one representative time point, that captures the behavior of all time points in the gap. We have provided a completion algorithm and have proven that entailment checking in $\mathcal{TELH}_\bot^{\diamondsuit,\mathsf{lhs}}$ is PTime-complete.

In Chapter 5 we have used minimal-world semantics to answer MTNCQs over $\mathcal{TELH}_\bot^{\diamondsuit,\mathsf{lhs}}$-KBs. We have shown that the temporal minimal universal model exists and is unique in this setting. Moreover, we have extended the rewriting to take into account the temporal parts as well which has led to a query answering algorithm that is tractable in data complexity, despite the additional expressiveness of the temporal operators.

In the last two chapters we have evaluated our theory on real world data with two experiments. Firstly, we have implemented and evaluated a system for the automatic translation of clinical trial criteria into MTNCQs in Chapter 6. The results have been promising and have shown that for many criteria an automatic translation is possible. In the second experiment, we have implemented a system for answering MTNCQs over $\mathcal{TELH}_\bot^{\diamondsuit,\mathsf{lhs}}$-KBs using the minimal-world semantics. We have evaluated it by computing answers to five manually translated criteria over a set of electronic health records. The results have showed the feasibility of the approach. While the system is not optimized enough to compete against other solutions yet, we have shown that temporal reasoning leads to better results than non-temporal reasoning.

8.2. Future Work

Some technical future work has already been discussed at the end of each chapter. Here we would like to give some more general remarks on future research.

As mentioned already an interesting next steps would be the extension of minimal-world semantics to more expressive Horn-DLs. In Section 3.5 we pointed out a number of issues that need to be solved when moving to more expressive DLs. On the other hand it would be interesting to see if minimal-world semantics can be simulated with other very general non-monotonic formalisms such as circumscription [McC80; BLW06; BLW09]. Most likely this would not yield optimal complexity bounds as more general formalisms usually have a higher complexity of reasoning, but it would open a new perspective on minimal-world semantics.

Technically, it seems possible to extend the construction of the temporal minimal universal model to temporal roles. In that case it would be very interesting to see how the rewriting would look like and what the final data complexity of query answering would be. Apart from the technical perspective it is not clear whether the answers in this setting are still consistent with the intuition behind minimal-world semantics. To further assess this, it could be beneficial to identify domains that are similar in structure to the medical domain. These domains could then be used to empirically test if the returned results are intuitive to the user or not.

Recently there has been a lot of interest in the explainability of the outputs of a given system both in logic based and in statistical approaches [Mil19; BHK+20; PK20]. For any AI system explainability is a crucial part, since otherwise the system will not be trusted. In our OMQA setting, it should not be a problem to extract the sequence of logical inferences that lead to the result. However, this probably long and complicated sequence is likely not acceptable to users, who mostly are non-logicians. It is subject to future research to find ways to present an explanation in an easy to understand way to users.

Bibliography

[AHV95] Serge Abiteboul, Richard Hull, and Victor Vianu: *Foundations of Databases*. Addison-Wesley, 1995. ISBN: 0-201-53771-0 (cited on page 5).

[AOS16] Shqiponja Ahmetaj, Magdalena Ortiz, and Mantas Simkus: 'Polynomial Datalog Rewritings for Expressive Description Logics with Closed Predicates'. In *Proc. of the 25th Int. Joint Conf. on Artificial Intelligence (IJCAI'16)*. AAAI Press, 2016, pages 878–885. DOI: `10.1016/j.artint.2019.103220` (cited on page 33).

[AFH96] Rajeev Alur, Tomás Feder, and Thomas A. Henzinger: 'The Benefits of Relaxing Punctuality'. In *Journal of the ACM* **43**(1): 1996, pages 116–146. DOI: `10.1145/227595.227602` (cited on page 78).

[AH93] Rajeev Alur and Thomas A Henzinger: 'Real-time logics: Complexity and expressiveness'. In *Information and Computation* **104**(1): 1993, pages 35–77. DOI: `10.1006/inco.1993.1025` (cited on page 28).

[AH94] Rajeev Alur and Thomas A. Henzinger: 'A Really Temporal Logic'. In *Journal of the ACM* **41**(1): Jan. 1994, pages 181–203. ISSN: 0004-5411. DOI: `10.1145/174644.174651` (cited on page 77).

[AGP14] Marcelo Arenas, Georg Gottlob, and Andreas Pieris: 'Expressive Languages for Querying the Semantic Web'. In *Proc. of the 33rd Symp. on Principles of Database Systems (PODS'14)*. ACM, 2014, pages 14–26. DOI: `10.1145/2594538.2594555` (cited on pages 1, 33).

[Aro01] Alan R Aronson: 'Effective mapping of biomedical text to the UMLS Metathesaurus: The MetaMap program.' In *Proc. AMIA Symposium*. American Medical Informatics Association. 2001, page 17 (cited on pages 85, 89).

[AKK+15] Alessandro Artale, Roman Kontchakov, Alisa Kovtunova, Vladislav Ryzhikov, Frank Wolter, and Michael Zakharyaschev: 'First-Order Rewritability of Ontology-Mediated Temporal Queries'. In *Proc. IJCAI*. AAAI Press, 2015, pages 2706–2712 (cited on pages 55, 69, 78).

[AKK+17] Alessandro Artale, Roman Kontchakov, Alisa Kovtunova, Vladislav Ryzhikov, Frank Wolter, and Michael Zakharyaschev: 'Ontology-Mediated Query Answering over Temporal Data: A Survey (Invited Talk)'. In *Proc. of the 24th Int. Symp. on Temporal Representation and Reasoning (TIME'17)*. Volume 90. Leibniz International Proceedings in Informatics. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2017, 1:1–1:37. DOI: `10.4230/LIPIcs.TIME.2017.1` (cited on page 67).

[AKL+07] Alessandro Artale, Roman Kontchakov, Carsten Lutz, Frank Wolter, and Michael Zakharyaschev: 'Temporalising Tractable Description Logics'. In *Proc. TIME*. IEEE Press, 2007, pages 11–22. DOI: 10.1109/TIME.2007.62 (cited on pages 6, 60, 67).

[AKW+13] Alessandro Artale, Roman Kontchakov, Frank Wolter, and Michael Zakharyaschev: 'Temporal Description Logic for Ontology-Based Data Access'. In *Proc. IJCAI*. AAAI Press, 2013, pages 711–717 (cited on pages 6, 55, 78).

[Baa03] Franz Baader: 'Terminological Cycles in a Description Logic with Existential Restrictions'. In *Proc. of the 18th Int. Joint Conf. on Artificial Intelligence (IJCAI'03)*. Morgan Kaufmann, 2003, pages 325–330. URL: http://ijcai.org/Proceedings/03/Papers/048.pdf (cited on page 3).

[BBF18] Franz Baader, Stefan Borgwardt, and Walter Forkel: 'Patient Selection for Clinical Trials Using Temporalized Ontology-Mediated Query Answering'. In *Proc. HQA Workshop*. ACM, 2018, pages 1069–1074. DOI: 10.1145/3184558.3191538 (cited on page 10).

[BBK+17] Franz Baader, Stefan Borgwardt, Patrick Koopmann, Ana Ozaki, and Veronika Thost: 'Metric Temporal Description Logics with Interval-Rigid Names'. In *Proc. FroCoS Symposium*. Springer, 2017, pages 60–76. DOI: 10.1007/978-3-319-66167-4_4 (cited on pages 28, 67, 78, 86).

[BBL13] Franz Baader, Stefan Borgwardt, and Marcel Lippmann: 'Temporalizing Ontology-Based Data Access'. In *Proc. of the 24th Int. Conf. on Automated Deduction (CADE'13)*. Springer, 2013, pages 330–344. DOI: 10.1007/978-3-642-38574-2_23 (cited on page 78).

[BBL15a] Franz Baader, Stefan Borgwardt, and Marcel Lippmann: 'Temporal Conjunctive Queries in Expressive Description Logics with Transitive Roles'. In *Proc. of the 28th Australasian Joint Conf. on Advances in Artificial Intelligence, AI'15*. Volume 9457. LNCS. Springer, 2015, pages 21–33. DOI: 10.1007/978-3-319-26350-2_3 (cited on page 78).

[BBL15b] Franz Baader, Stefan Borgwardt, and Marcel Lippmann: 'Temporal Query Entailment in the Description Logic $\mathcal{SHQ}$'. In *Journal of Web Semantics* **33**: 2015, pages 71–93. DOI: http://dx.doi.org/10.1016/j.websem.2014.11.008 (cited on pages 69, 78, 86).

[BBL05] Franz Baader, Sebastian Brandt, and Carsten Lutz: 'Pushing the $\mathcal{EL}$ Envelope'. In *Proc. of the 19th Int. Joint Conf. on Artificial Intelligence (IJCAI'05)*. Professional Book Center, 2005, pages 364–369. URL: http://ijcai.org/Proceedings/09/Papers/053.pdf (cited on pages 3, 60).

[BCM+07] Franz Baader, Diego Calvanese, Deborah L. McGuinness, Daniele Nardi, and Peter F. Patel-Schneider, editors: *The Description Logic Handbook: Theory, Implementation, and Applications*. 2nd. Cambridge University Press, 2007. DOI: 10.1017/CBO9780511711787 (cited on pages 3, 13, 35).

[BHL+17] Franz Baader, Ian Horrocks, Carsten Lutz, and Uli Sattler: *An Introduction to Description Logic*. Cambridge University Press, 2017. DOI: 10.1017/9781139025355 (cited on pages 1, 3, 13, 17–21).

[BTM+15] Richard Bache, Adel Taweel, Simon Miles, and Brendan C Delaney: 'An eligibility criteria query language for heterogeneous data warehouses'. In *Methods of Information in Medicine* **54**(01): 2015, pages 41–44. DOI: 10.3414/ME13-02-0027 (cited on page 86).

[BKM+15] David Basin, Felix Klaedtke, Samuel Müller, and Eugen Zălinescu: 'Monitoring Metric First-Order Temporal Properties'. In *Journal of the ACM* **62**(2): 2015, 15:1–15:45. DOI: 10.1145/2699444 (cited on page 73).

[Bau95] Benjamin Bauslaugh: 'Core-like properties of infinite graphs and structures'. In *Discrete mathematics* **138**(1-3): 1995, pages 101–111. DOI: 10.1016/0012-365X(94)00191-K (cited on pages 37, 47, 48, 50).

[BHL01] Tim Berners-Lee, James Hendler, and Ora Lassila: 'The semantic web'. In *Scientific american* **284**(5): 2001, pages 34–43 (cited on page 1).

[BCZ+10] Paolo Besana, Marc Cuggia, Oussama Zekri, Annabel Bourde, and Anita Burgun: 'Using Semantic Web Technologies for Clinical Trial Recruitment'. In *Proc. of the 9th Int. Semantic Web Conf. (ISWC'10)*. Volume 6497. Lecture Notes in Computer Science. Springer, 2010, pages 34–49. DOI: 10.1007/978-3-642-17749-1_3 (cited on page 112).

[BC13] Sanmitra Bhattacharya and Michael N Cantor: 'Analysis of eligibility criteria representation in industry-standard clinical trial protocols'. In *Journal of Biomedical Informatics* **46**(5): 2013, pages 805–813. DOI: 10.1016/j.jbi.2013.06.001 (cited on pages 86, 88).

[BO15] Meghyn Bienvenu and Magdalena Ortiz: 'Ontology-mediated query answering with data-tractable description logics'. In *Reasoning Web 11th Int. Summer School*. Volume 9203. Lecture Notes in Computer Science. Springer, 2015, pages 218–307. DOI: 10.1007/978-3-319-21768-0_9 (cited on pages 25, 38, 41, 114).

[BTC+12] Mary Regina Boland, Samson W Tu, Simona Carini, Ida Sim, and Chunhua Weng: 'EliXR-TIME: A temporal knowledge representation for clinical research eligibility criteria'. In *AMIA Joint Summits on Translational Science* **2012**: 2012, page 71 (cited on page 86).

[BLW06] Piero A Bonatti, Carsten Lutz, and Frank Wolter: 'Description Logics with Circumscription.' In *KR*. 2006, pages 400–410 (cited on page 115).

[BLW09] Piero A Bonatti, Carsten Lutz, and Frank Wolter: 'The complexity of circumscription in DLs'. In *Journal of Artificial Intelligence Research* **35**: 2009, pages 717–773. DOI: 10.1613/jair.2763 (cited on page 115).

[BJ18] Luca Bonomi and Xiaoqian Jiang: 'Patient Ranking with Temporally Annotated Data'. In *Journal of Biomedical Informatics* **78**: 2018, pages 43–53. DOI: 10.1016/j.jbi.2017.12.007 (cited on page 112).

[Bor96] Alex Borgida: 'On the relative expressiveness of description logics and predicate logics'. In *Artificial intelligence* **82**(1-2): 1996, pages 353–367. DOI: 10.1016/0004-3702(96)00004-5 (cited on page 17).

[BF19a] Stefan Borgwardt and Walter Forkel: 'Closed-World Semantics for Conjunctive Queries with Negation over $\mathcal{ELH}_{\bot}$ Ontologies'. In *Proc. JELIA Conference*. Rende, Italy: Springer, 2019, pages 371–386. DOI: 10.1007/978-3-030-19570-0_24 (cited on page 8).

[BF19b] Stefan Borgwardt and Walter Forkel: 'Closed-World Semantics for Conjunctive Queries with Negation over $\mathcal{ELH}_{\bot}$ Ontologies'. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence*. AAAI Press. 2019, pages 6131–6135. DOI: 10.24963/ijcai.2019/849 (cited on page 8).

[BF19c] Stefan Borgwardt and Walter Forkel: 'Closed-World Semantics for Conjunctive Queries with Negation over $\mathcal{ELH}_{\bot}$ Ontologies (Extended Abstract)'. In *Proceedings of the 32nd International Workshop on Description Logics*. Volume 2373. CEUR Workshop Proceedings. 2019. URL: http://ceur-ws.org/Vol-2373/paper-36.pdf (cited on page 8).

[BFK19] Stefan Borgwardt, Walter Forkel, and Alisa Kovtunova: 'Finding New Diamonds: Temporal Minimal- World Query Answering over Sparse ABoxes'. In *Proc. of the 3rd International Joint Conference on Rules and Reasoning (RuleML+RR'19)*. Bolzano, Italy: Springer, 2019. DOI: 10.1007/978-3-030-31095-0_1 (cited on page 9).

[BFK20] Stefan Borgwardt, Walter Forkel, and Alisa Kovtunova: 'Finding New Diamonds: Temporal Minimal- World Query Answering over Sparse ABoxes (Extended Abstract)'. In *Proceedings of the 33rd International Workshop on Description Logics (DL 2020)*. Volume 2663. CEUR Workshop Proceedings. 2020. DOI: 10.1007/978-3-030-31095-0_1 (cited on page 9).

[BFK] Stefan Borgwardt, Walter Forkel, and Alisa Kovtunova: 'Finding New Diamonds: Temporal Minimal- World Query Answering over Sparse ABoxes'. In: Submitted to the Journal of Theory and Practice of Logic Programming (TPLP). DOI: 10.1007/978-3-030-31095-0_1 (cited on page 9).

[BHK+20] Stefan Borgwardt, Anke Hirsch, Alisa Kovtunova, and Frederik Wiehr: 'In the Eye of the Beholder: Which Proofs are Best?' In *Proc. of the 33th Int. Workshop on Description Logics (DL 2020)*. 2020 (cited on page 115).

[BLT13] Stefan Borgwardt, Marcel Lippmann, and Veronika Thost: 'Temporal Query Answering in the Description Logic *DL-Lite*'. In *Proc. of the 9th Int. Symposium on Frontiers of Combining Systems, FroCoS'13*. Volume 8152. LNCS. Springer, 2013, pages 165–180 (cited on page 78).

[BLT15] Stefan Borgwardt, Marcel Lippmann, and Veronika Thost: 'Temporalizing rewritable query languages over knowledge bases'. In *Web Semantics: Science, Services and Agents on the World Wide Web* **33**: 2015, pages 50–70. DOI: 10.2139/ssrn.3199188 (cited on page 78).

[BT15a] Stefan Borgwardt and Veronika Thost: 'Temporal Query Answering in *DL-Lite* with Negation'. In *Proc. of the 1st Global Conf. on Artificial Intelligence (GCAI'15)*. EasyChair, 2015, pages 51–65 (cited on page 78).

[BT15b] Stefan Borgwardt and Veronika Thost: 'Temporal Query Answering in the Description Logic $\mathcal{EL}$'. In *Proc. of the 24h Int. Joint Conf. on Artificial Intelligence, IJCAI'15*. AAAI Press, 2015, pages 2819–2825 (cited on page 78).

[BT20] Stefan Borgwardt and Veronika Thost: 'Temporal Conjunctive Query Answering in the Extended *DL-Lite* Family'. In *arXiv:2003.09508*. 2020. arXiv: `2003.09508 [cs.LO]` (cited on page 78).

[CCG10a] Jean-Paul Calbimonte, Oscar Corcho, and Alasdair J. G. Gray: 'Enabling Ontology-Based Access to Streaming Data Sources'. In. ISWC'10: 2010, pages 96–111 (cited on page 78).

[CCG10b] Jean-Paul Calbimonte, Oscar Corcho, and Alasdair JG Gray: 'Enabling ontology-based access to streaming data sources'. In *International Semantic Web Conference*. Springer. 2010, pages 96–111. DOI: `10.1007/978-3-642-17746-0_7` (cited on pages 1, 7).

[CDL+06] Diego Calvanese, Giuseppe De Giacomo, Domenico Lembo, Maurizio Lenzerini, and Riccardo Rosati: 'Epistemic first-order queries over description logic knowledge bases'. In *Proc. of the 19th Int. Workshop on Description Logics (DL'06)*. Volume 189. CEUR Workshop Proceedings. 2006, pages 51–61 (cited on page 32).

[CDL+13] Diego Calvanese, Giuseppe De Giacomo, Domenico Lembo, Maurizio Lenzerini, and Riccardo Rosati: 'Data complexity of query answering in description logics'. In *Artificial Intelligence* **195**: 2013, pages 335–360. ISSN: 0004-3702. DOI: `https://doi.org/10.1016/j.artint.2012.10.003` (cited on page 77).

[CKM+18] David Carral, Markus Krötzsch, Maximilian Marx, Ana Ozaki, and Sebastian Rudolph: 'Preserving Constraints with the Stable Chase'. In *21st International Conference on Database Theory (ICDT 2018)*. Volume 98. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2018, 12:1–12:19. ISBN: 978-3-95977-063-7. DOI: `10.4230/LIPIcs.ICDT.2018.12` (cited on page 52).

[CAT+17] Efthymios Chondrogiannis, Vassiliki Andronikou, Anastasios Tagaris, Efstathios Karanastasis, Theodora Varvarigou, and Masatsugu Tsuji: 'A novel semantic representation for eligibility criteria in clinical trials'. In *Journal of Biomedical Informatics* **69**: 2017, pages 10–23. DOI: `10.1016/j.jbi.2017.03.013` (cited on pages 86, 88).

[CCH+94] James J Cimino, Paul D Clayton, George Hripcsak, and Stephen B Johnson: 'Knowledge-based approaches to the maintenance of a large controlled medical terminology'. In *Journal of the American Medical Informatics Association* **1**(1): 1994, pages 35–50. DOI: 10.1136/jamia.1994.95236135 (cited on page 112).

[CE81] Edmund M Clarke and E Allen Emerson: 'Design and synthesis of synchronization skeletons using branching time temporal logic'. In *Workshop on Logic of Programs*. Springer. 1981, pages 52–71. DOI: 10.1007/BFb0025774 (cited on page 25).

[CT15] Christopher L Crowe and Cui Tao: 'Designing ontology-based patterns for the representation of the time-relevant eligibility criteria of clinical protocols'. In *AMIA Joint Summits on Translational Science* **2015**: 2015, page 173. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4525239/ (cited on pages 1, 55, 112).

[DDL+11] Kalen Delaney, Louis Davidson, Greg Low, Brad McGhee, Paul Nielson, Paul Randal, and Kimberly Tripp: *SQL Server MVP Deep Dives, Volume 2*. USA: Manning Publications Co., 2011. Chapter 5. ISBN: 1617290475 (cited on page 99).

[DNR08] Alin Deutsch, Alan Nash, and Jeffrey B. Remmel: 'The chase revisited'. In *Proc. of the 27th ACM Symp. on Principles of Database Systems (PODS'08)*. ACM, 2008, pages 149–158. DOI: 10.1145/1376916.1376938 (cited on pages 33, 52).

[DL16] Li Dong and Mirella Lapata: 'Language to logical form with neural attention'. In *Proc. Annual Meeting of the ACL*. 2016. DOI: 10.18653/v1/P16-1004 (cited on page 86).

[EOŠ+12] Thomas Eiter, Magdalena Ortiz, Mantas Šimkus, Trung-Kien Tran, and Guohui Xiao: 'Query Rewriting for Horn-$\mathcal{SHIQ}$ Plus Rules'. In *Proc. of the 26th AAAI Conf. on Artificial Intelligence (AAAI'12)*. AAAI Press, 2012, pages 726–733 (cited on pages 38, 41, 114).

[EVØ17] Martine Enger, Erik Velldal, and Lilja Øvrelid: 'An open-source tool for negation detection: A maximum-margin approach'. In *Proc. SemBEaR Workshop*. 2017, pages 64–69. DOI: 10.18653/v1/W17-1810 (cited on page 86).

[FKM+05] Ronald Fagin, Phokion G. Kolaitis, Renée J. Miller, and Lucian Popa: 'Data Exchange: Semantics and Query Answering'. In *Theoretical Computer Science* **336**(1): 2005, pages 89–124. DOI: 10.1016/j.tcs.2004.10.033 (cited on page 33).

[Fel82] Wolf-Dietrich Fellner: 'On minimal graphs'. In *Theoretical Computer Science* **17**(1): 1982, pages 103–110. DOI: 10.1016/0304-3975(82)90135-9 (cited on page 47).

[FS08] Carlo A. Furia and Paola Spoletini: 'Tomorrow and All our Yesterdays: MTL Satisfiability over the Integers'. In *Proc. ICTAC*. Springer, 2008, pages 126–140. DOI: 10.1007/978-3-540-85762-4_9 (cited on page 77).

[GIK+15] Víctor Gutiérrez-Basulto, Yazmín Ibáñez-García, Roman Kontchakov, and Egor V. Kostylev: 'Queries with Negation and Inequalities over Lightweight Ontologies'. In *Journal of Web Semantics* **35**: 2015, pages 184–202. DOI: `10.1016/j.websem.2015.06.002` (cited on pages 32, 54, 78).

[GJK16] Víctor Gutiérrez-Basulto, Jean Christoph Jung, and Roman Kontchakov: 'Temporalized $\mathcal{EL}$ Ontologies for Accessing Temporal Data: Complexity of Atomic Queries'. In *Proc. IJCAI.* AAAI Press, 2016, pages 1102–1108 (cited on pages 6, 55, 59, 67).

[GJL12] Víctor Gutiérrez-Basulto, Jean Christoph Jung, and Carsten Lutz: 'Complexity of Branching Temporal Description Logics'. In *Proc. of the 20th Eur. Conf. on Artificial Intelligence (ECAI'12).* IOS Press, 2012, pages 390–395. DOI: `10.3233/978-1-61499-098-7-390` (cited on page 67).

[GJO16a] Víctor Gutiérrez-Basulto, Jean Christoph Jung, and Ana Ozaki: 'On Metric Temporal Description Logics'. In *Proc. ECAI.* IOS Press, 2016, pages 837–845. DOI: `10.3233/978-1-61499-672-9-837` (cited on pages 67, 78).

[GJO16b] Víctor Gutiérrez-Basulto, Jean Christoph Jung, and Ana Ozaki: 'On Metric Temporal Description Logics.' In *ECAI.* 2016, pages 837–845 (cited on page 28).

[GJS14] Víctor Gutiérrez-Basulto, Jean Christoph Jung, and Thomas Schneider: 'Lightweight Description Logics and Branching Time: A Troublesome Marriage'. In *Proc. of the 14th Int. Conf. on Principles of Knowledge Representation and Reasoning (KR'14).* AAAI Press, 2014, pages 278–287 (cited on page 67).

[GJS15] Víctor Gutiérrez-Basulto, Jean Christoph Jung, and Thomas Schneider: 'Lightweight Temporal Description Logics with Rigid Roles and Restricted TBoxes'. In *Proc. of the 24th Int. Joint Conf. on Artificial Intelligence (IJCAI'15).* AAAI Press, 2015, pages 3015–3021 (cited on page 67).

[HN92] Pavol Hell and Jaroslav Nešetřil: 'The core of a graph'. In *Discrete Mathematics* **109**(1-3): 1992, pages 117–126. DOI: `10.1016/0012-365X(92)90282-K` (cited on page 47).

[HKL+13] André Hernich, Clemens Kupke, Thomas Lukasiewicz, and Georg Gottlob: 'Well-founded semantics for extended datalog and ontological reasoning'. In *Proc. of the 32nd Symp. on Principles of Database Systems (PODS'13).* ACM, 2013, pages 225–236. DOI: `10.1145/2463664.2465229` (cited on page 33).

[HPV03] Ian Horrocks, Peter F Patel-Schneider, and Frank Van Harmelen: 'From SHIQ and RDF to OWL: The making of a web ontology language'. In *Journal of Web Semantics* **1**(1): 2003, pages 7–26. DOI: `10.1016/j.websem.2003.07.001` (cited on page 2).

[HZP+05] George Hripcsak, Li Zhou, Simon Parsons, Amar K. Das, and Stephen B. Johnson: 'Modeling Electronic Discharge Summaries as a Simple Temporal Constraint Satisfaction Problem'. In *Journal of the American Medical Informatics Association* **12**(1): 2005, pages 55–63. DOI: `10.1197/jamia.m1623` (cited on page 112).

[HL07] Yang Huang and Henry J. Lowe: 'A Novel Hybrid Approach to Automated Negation Detection in Clinical Radiology Reports'. In *Journal of the American Medical Informatics Association* **14**(3): May 2007, pages 304–311. ISSN: 1067-5027. DOI: `10.1197/jamia.M2284` (cited on page 86).

[Inc20a] Wikipedia Foundation Inc.: *Angina*. 2020. URL: `https://en.wikipedia.org/wiki/Angina` (visited on 30/09/2020) (cited on page 105).

[Inc20b] Wikipedia Foundation Inc.: *Cardiac Arrest*. 2020. URL: `https://en.wikipedia.org/wiki/Cardiac_arrest` (visited on 30/09/2020) (cited on page 105).

[Inc20c] Wikipedia Foundation Inc.: *Ischemia*. 2020. URL: `https://en.wikipedia.org/wiki/Ischemia` (visited on 30/09/2020) (cited on page 105).

[Inc20d] Wikipedia Foundation Inc.: *Myocardial Infarction*. 2020. URL: `https://en.wikipedia.org/wiki/Myocardial_infarction` (visited on 30/09/2020) (cited on page 105).

[Inc20e] Wikipedia Foundation Inc.: *Nephropathy*. 2020. URL: `https://en.wikipedia.org/wiki/Kidney_disease` (visited on 30/09/2020) (cited on page 106).

[Inc20f] Wikipedia Foundation Inc.: *Neuropathy*. 2020. URL: `https://en.wikipedia.org/wiki/Peripheral_neuropathy` (visited on 30/09/2020) (cited on page 106).

[Inc20g] Wikipedia Foundation Inc.: *Retinopathy*. 2020. URL: `https://en.wikipedia.org/wiki/Retinopathy` (visited on 30/09/2020) (cited on page 106).

[JL76] Neil D. Jones and William T. Laaser: 'Complete Problems for Deterministic Polynomial Time'. In *Theoretical Computer Science* **3**(1): 1976, pages 105–117. DOI: `10.1016/0304-3975(76)90068-2` (cited on page 66).

[KMM+17] Evgeny Kharlamov, Theofilos Mailis, Gulnar Mehdi, Christian Neuenstadt, Özgür Özçep, Mikhail Roshchin, Nina Solomakhina, Ahmet Soylu, Christoforos Svingos, Sebastian Brandt et al.: 'Semantic access to streaming and static data at Siemens'. In *Journal of Web Semantics* **44**: 2017, pages 54–74. DOI: `10.1016/j.websem.2017.02.001` (cited on page 1).

[Khy20] Ryny Khy: 'Report: Minimal Canonical Model for Horn-$\mathcal{ALCHOIQ}^{\mathrm{Disj}}_{\mathrm{Self}}$'. Technical report. Institute of Theoretical Computer Science, Chair of Automata Theory, 2020 (cited on page 54).

[KLT+11] Roman Kontchakov, Carsten Lutz, David Toman, Frank Wolter, and Michael Zakharyaschev: 'The Combined Approach to Ontology-Based Data Access'. In *Proc. of the 22nd Int. Joint Conf. on Artificial Intelligence (IJCAI'11)*. AAAI Press, 2011, pages 2656–2661. DOI: `10.5591/978-1-57735-516-8/IJCAI11-442` (cited on pages 5, 46).

[KPP+16] Roman Kontchakov, Laura Pandolfo, Luca Pulina, Vladislav Ryzhikov, and Michael Zakharyaschev: 'Temporal and Spatial OBDA with Many-Dimensional Halpern-Shoham Logic'. In *Proc. of the 25th Int. Joint Conf. on Artificial Intelligence (IJCAI'16)*. AAAI Press, 2016, pages 1160–1166 (cited on page 78).

[KP14] Felix Köpcke and Hans-Ulrich Prokosch: 'Employing Computers for the Recruitment into Clinical Trials: A Comprehensive Systematic Review'. In *Journal of Medical Internet Research* **16**(7): 2014, e161. DOI: `10.2196/jmir.3446` (cited on page 112).

[Krö20] Markus Krötzsch: 'Computing Cores for Existential Rules with the Standard Chase and ASP'. In *Proceedings of the 17th International Conference on Principles of Knowledge Representation and Reasoning (KR 2020)*. IJCAI, 2020 (cited on pages 52, 53).

[Kuh14] Tobias Kuhn: 'A Survey and Classification of Controlled Natural Languages'. In *Computational Linguistics* **40**(1): Mar. 2014, pages 121–170. DOI: `10.1162/COLI_a_00168` (cited on page 95).

[Lev96] Alon Y Levy: 'Obtaining complete answers from incomplete databases'. In *VLDB*. Volume 96. Citeseer. 1996, pages 402–412 (cited on page 6).

[Lin92] Steven Lindell: 'A Purely Logical Characterization of Circuit Uniformity'. In *Proc. of the 7th Annual Structure in Complexity Theory Conf.* 1992, pages 185–192. DOI: `10.1109/SCT.1992.215393` (cited on page 77).

[LYW11] Zhihui Luo, Meliha Yetisgen-Yildiz, and Chunhua Weng: 'Dynamic categorization of clinical research eligibility criteria by hierarchical clustering'. In *Journal of Biomedical Informatics* **44**(6): 2011, pages 927–935. DOI: `10.1016/j.jbi.2011.06.001` (cited on pages 86, 88).

[LSW13] Carsten Lutz, Inanc Seylan, and Frank Wolter: 'Ontology-Based Data Access with Closed Predicates is Inherently Intractable (Sometimes)'. In *Proc. of the 23rd Int. Joint Conf. on Artificial Intelligence (IJCAI'13)*. AAAI Press, 2013, pages 1024–1030. URL: `https://www.ijcai.org/Abstract/13/156` (cited on pages 33, 34).

[LTW09] Carsten Lutz, David Toman, and Frank Wolter: 'Conjunctive Query Answering in the Description Logic $\mathcal{EL}$ Using a Relational Database System'. In *Proc. of the 21st Int. Joint Conf. on Artificial Intelligence (IJCAI'09)*. AAAI Press, 2009, pages 2070–2075 (cited on pages 24, 25, 73).

[LWW07] Carsten Lutz, Dirk Walther, and Frank Wolter: 'Quantitative temporal logics over the reals: PSPACE and below'. In *Information and Computation* **205**(1): 2007, pages 99–123. DOI: `10.1016/j.ic.2006.08.006` (cited on page 28).

[LWZ08] Carsten Lutz, Frank Wolter, and Michael Zakharyaschev: 'Temporal Description Logics: A Survey'. In *Proc. of the 15th Int. Symp. on Temporal Representation and Reasoning (TIME'08)*. IEEE Press, 2008, pages 3–14. DOI: `10.1109/TIME.2008.14` (cited on pages 6, 67).

[MSB+14] Christopher Manning, Mihai Surdeanu, John Bauer, Jenny Finkel, Steven Bethard, and David McClosky: 'The Stanford CoreNLP natural language processing toolkit'. In *Proc. ALC Meeting*. 2014, pages 55–60 (cited on pages 85, 89).

[Mar09] Bruno Marnette: 'Generalized Schema Mappings: From Termination To Tractability'. In *Proc. of the 28th Symp. on Principles of Database Systems (PODS'09)*. ACM, 2009, pages 13–22. DOI: `10.1145/1559795.1559799` (cited on page 33).

[McC80] John McCarthy: 'Circumscription—a form of non-monotonic reasoning'. In *Artificial intelligence* **13**(1-2): 1980, pages 27–39 (cited on pages 1, 6, 115).

[MD80] Drew McDermott and Jon Doyle: 'Non-monotonic logic I'. In *Artificial intelligence* **13**(1-2): 1980, pages 41–72 (cited on page 6).

[MR20] Mayo Foundation for Medical Education and Research: *Ketoacidosis*. 2020. URL: `https://www.mayoclinic.org/diseases-conditions/diabetic-ketoacidosis/symptoms-causes/syc-20371551` (visited on 30/09/2020) (cited on page 106).

[Med20] U.S. National Library of Medicine: *Coronary Artery Disease*. 2020. URL: `https://medlineplus.gov/coronaryarterydisease.html` (visited on 30/09/2020) (cited on page 105).

[MCC+13] Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean: 'Efficient estimation of word representations in vector space'. In *arXiv preprint arXiv:1301.3781*: 2013 (cited on pages 85, 89).

[MBT12] Krystyna Milian, Anca Bucur, and Annette Ten Teije: 'Formalization of clinical trial eligibility criteria: Evaluation of a pattern-based approach'. In *Proc. BIBM Conference*. IEEE. 2012, pages 1–4. DOI: `10.1109/BIBM.2012.6392733` (cited on page 86).

[MT13] Krystyna Milian and Annette ten Teije: 'Towards automatic patient eligibility assessment: From free-text criteria to queries'. In *Proc. AIME Conference*. Springer. 2013, pages 78–83. DOI: `10.1007/978-3-642-38326-7_12` (cited on page 86).

[Mil19] Tim Miller: 'Explanation in artificial intelligence: Insights from the social sciences'. In *Artificial Intelligence* **267**: 2019, pages 1–38. ISSN: 0004-3702. DOI: `https://doi.org/10.1016/j.artint.2018.07.007` (cited on page 115).

[OKK+19] Michel Oleynik, Amila Kugic, Zdenko Kasáč, and Markus Kreuzthaler: 'Evaluating shallow and deep learning strategies for the 2018 n2c2 shared task on clinical text classification'. In *Journal of the American Medical Informatics Association* **26**(11): 2019, pages 1247–1254. ISSN: 1527-974X. DOI: `10.1093/jamia/ocz149` (cited on page 112).

[ORS11] Magdalena Ortiz, Sebastian Rudolph, and Mantas Simkus: 'Query answering in the Horn fragments of the description logics SHOIQ and SROIQ'. In *Twenty-Second International Joint Conference on Artificial Intelligence*. 2011 (cited on pages 24, 25, 54).

[OSE08] Magdalena Ortiz, Mantas Simkus, and Thomas Eiter: 'Worst-case Optimal Conjunctive Query Answering for an Expressive Description Logic without Inverses.' In *AAAI*. 2008, pages 504–510 (cited on page 24).

[PK20] Akash Kumar Panda and Bart Kosko: 'Random fuzzy-rule foams for explainable AI'. In *Fuzzy Information Processing 2020, Proceedings of NAFIPS-2020*. Springer. 2020 (cited on page 115).

[Pap94] Christos H. Papadimitriou: *Computational Complexity*. Theoretical computer science. Addison-Wesley, 1994. ISBN: 9780201530827. URL: `https://books.google.de/books?id=JogZAQAAIAAJ` (cited on page 19).

[PCD+07] Chintan Patel, James Cimino, Julian Dolby, Achille Fokoue, Aditya Kalyanpur, Aaron Kershenbaum, Li Ma, Edith Schonberg, and Kavitha Srinivas: 'Matching Patient Records to Clinical Trials Using Ontologies'. In *Proc. of the 6th Int. Semantic Web Conf. (ISWC'07)*. Volume 4825. Lecture Notes in Computer Science. Springer, 2007, pages 816–829. DOI: `10.1007/978-3-540-76298-0_59` (cited on pages 29, 112).

[Pnu77] Amir Pnueli: 'The temporal logic of programs'. In *18th Annual Symposium on Foundations of Computer Science (sfcs 1977)*. IEEE. 1977, pages 46–57. DOI: `10.1109/SFCS.1977.32` (cited on pages 6, 25).

[RB08] Alan L Rector and Sebastian Brandt: 'Why do it the hard way? The case for an expressive description logic for SNOMED'. In *Journal of the American Medical Informatics Association* **15**(6): 2008, pages 744–751. DOI: `10.1197/jamia.M2797` (cited on page 1).

[Rei80] Raymond Reiter: 'A logic for default reasoning'. In *Artificial intelligence* **13**(1-2): 1980, pages 81–132. DOI: `10.1016/0004-3702(80)90014-4` (cited on pages 1, 6).

[Ros07a] Riccardo Rosati: 'On Conjunctive Query Answering in $\mathcal{EL}$'. In *Proc. DL*. 2007, pages 451–458 (cited on page 25).

[Ros07b] Riccardo Rosati: 'The Limits of Querying Ontologies'. In *Proc. of the 11th Int. Conf. on Database Theory (ICDT'07)*. 2007, pages 164–178. DOI: `10.1007/11965893_12` (cited on page 78).

[Rud11] Sebastian Rudolph: 'Foundations of description logics'. In *Reasoning Web International Summer School*. Springer. 2011, pages 76–136. DOI: `10.1007/978-3-642-23032-5_2` (cited on page 16).

[RN14] Stuart Russell and Peter Norvig: *Artificial Intelligence: A Modern Approach*. 3rd edition. Pearson, 2014. ISBN: 978129202420 (cited on page 1).

[RWZ19] Vladislav Ryzhikov, Przemyslaw Andrzej Walega, and Michael Zakharyaschev: 'Data Complexity and Rewritability of Ontology-Mediated Queries in Metric Temporal Logic under the Event-Based Semantics'. In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI-19*. International Joint Conferences on Artificial Intelligence Organization, July 2019, pages 1851–1857. DOI: `10.24963/ijcai.2019/256` (cited on page 78).

[Sch93] Klaus Schild: 'Combining terminological logics with tense logic'. In *Portuguese Conference on Artificial Intelligence*. Springer. 1993, pages 105–120. DOI: `10.1007/3-540-57287-2_41` (cited on page 6).

[SS91] Manfred Schmidt-Schauß and Gert Smolka: 'Attributive concept descriptions with complements'. In *Artificial intelligence* **48**(1): 1991, pages 1–26. DOI: `10.1016/0004-3702(91)90078-X` (cited on pages 3, 13).

[SC85] A Prasad Sistla and Edmund M Clarke: 'The complexity of propositional linear temporal logics'. In *Journal of the ACM (JACM)* **32**(3): 1985, pages 733–749. DOI: `10.1145/3828.3837` (cited on page 27).

[SFS+19] Amber Stubbs, Michele Filannino, Ergin Soysal, Samuel Henry, and Özlem Uzuner: 'Cohort selection for clinical trials: n2c2 2018 shared task track 1'. In *Journal of the American Medical Informatics Association* **26**(11): 2019, pages 1163–1171. DOI: `10.1093/jamia/ocz163` (cited on pages 102, 112).

[TSC11] Cui Tao, Harold R. Solbrig, and Christopher G. Chute: 'CNTRO 2.0: A Harmonized Semantic Web Ontology for Temporal Relation Inferencing in Clinical Narratives'. In *AMIA Joint Summits on Translational Science Proceedings* **2011**: 2011, pages 64–68. URL: `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3248753/` (cited on pages 1, 112).

[Tho18] Veronika Thost: 'Metric Temporal Extensions of DL-Lite and Interval-Rigid Names'. In *Principles of Knowledge Representation and Reasoning: Proceedings of the Sixteenth International Conference, KR 2018, Tempe, Arizona*. AAAI Press, 2018, pages 665–666. URL: `https://aaai.org/ocs/index.php/KR/KR18/paper/view/18035` (cited on pages 67, 78).

[TPC+11] Samson W Tu, Mor Peleg, Simona Carini, Michael Bobak, Jessica Ross, Daniel Rubin, and Ida Sim: 'A practical method for transforming free-text eligibility criteria into computable criteria'. In *Journal of Biomedical Informatics* **44**(2): 2011, pages 239–250. DOI: `10.1016/j.jbi.2010.09.007` (cited on page 86).

[Var82] Moshe Y Vardi: 'The complexity of relational query languages'. In *Proceedings of the fourteenth annual ACM symposium on Theory of computing*. 1982, pages 137–146 (cited on pages 23, 77).

[WTS+10] Chunhua Weng, Samson W Tu, Ida Sim, and Rachel Richesson: 'Formal representation of eligibility criteria: a literature review'. In *Journal of Biomedical Informatics* **43**(3): 2010, pages 451–467. DOI: `10.1016/j.jbi.2009.12.004` (cited on page 86).

[WWL+11] Chunhua Weng, Xiaoying Wu, Zhihui Luo, Mary Regina Boland, Dimitri Theodoratos, and Stephen B Johnson: 'EliXR: An approach to eligibility criteria extraction and representation'. In *J. Am. Med. Inform. Assn.* **18**(Supplement_1): 2011, pages i116–i124. DOI: `10.1136/amiajnl-2011-000321` (cited on pages 86, 88).

[Wol00] Frank Wolter: 'First order common knowledge logics'. In *Studia Logica* **65**(2): 2000, pages 249–271. DOI: `10.1023/A:1005271815356` (cited on pages 1, 6, 33).

[WZ00] Frank Wolter and Michael Zakharyaschev: 'Temporalizing Description Logics'. In *Frontiers of Combining Systems 2*. Research Studies Press/Wiley, 2000, pages 379–402 (cited on page 67).

[XFB+19] Chao Xu, Walter Forkel, Stefan Borgwardt, Franz Baader, and Beihai Zhou: 'Automatic Translation of Clinical Trial Eligibility Criteria into Formal Queries'. In *Proc. of the 9th Workshop on Ontologies and Data in Life Sciencs (ODLS'19), part of The Joint Ontology Workshops (JOWO'19)*. CEUR Workshop Proceedings. 2019 (cited on pages 10, 93).

[YJY+15] YIN Yongmei, QIN Jie, DAI Yaping, ZENG Fanwei, PEI Hao, and WANG Jun: 'The CD4+/CD8+ ratio in pulmonary tuberculosis: systematic and meta-analysis article'. In *Iranian journal of public health* **44**(2): 2015, page 185 (cited on page 55).

[ZMP+06] Li Zhou, Genevieve B Melton, Simon Parsons, and George Hripcsak: 'A temporal constraint structure for extracting temporal information from clinical narrative'. In *Journal of Biomedical Informatics* **39**(4): 2006, pages 424–439. DOI: `10.1016/j.jbi.2005.07.002` (cited on page 86).

Index

F_1-measure, 110
$\perp$, *see* bottom concept
$\top$, *see* top concept
$\mathcal{ALC}$, 13
$\mathcal{ELH}_\perp$, 19
ExpTime, 19
NExpTime, 19
NP, 19
PSpace, 19
PTime, 19
coNP, 19

ABox, 2, 15
always, 27
answer, 21, 70
    certain, 22, 70
    minimal world, 73
assertion
    concept, 15
    role, 15
assignment, 21
axiom, 15

bottom concept, 14

CAD, *see* coronary artery disease
canonical model, *see* universal model
chase, 52
    core, 33, 52
    restricted, 33
    Skolem, 33
    standard, 52
closed predicates, 33
closed-world assumption, 5
combined first-order rewritability, 24
complement, 14
complexity, 13
    combined, 5, 23
    data, 5, 24
complexity class, 19
Computation Tree Logic, 25
computational complexity, 19
concept, 13
    name, 13
conjunction, 14
conjunctive query, 4
consequence-based reasoning, 21
consistency, 15, 18
controlled natural language, 95
convex, 55
convexity
    bounded, 60
convexity operator, 56
    bounded, 56
core, 47
coronary artery disease, 105
CQ, *see* conjunctive query, *see* conjunctive query
CTL, *see* Computation Tree Logic
CWA, *see* closed-world assumption

Description Logic, 2
disjunction, 14
DL, *see* Description Logic
domain, 14

EHR, *see* electronic health record
electronic health record, 5, 29
endomorphism, 24
entailment, 15
epistemic logic, 32
eventually, 27
existential restriction, 14
existential rule, 52
expanding, 55, 60
expressivity, 13

first-order logic, 16

flexible, 60

GALEN, 2
Gene Ontology, 3
general concept inclusion, 15
ground truth, 106
GT, *see* ground truth

homomorphism, 24
  strong, 24
Horn-DLs, 25

individual
  name, 13
instance checking, 18
interpretation, 14
  temporal, 59
inverse roles, 17

KB, *see* knowledge base
knowledge base, 2, 15

linear-time temporal logic, 6, 25
locally finite, 51
LTL, *see* linear-time temporal logic, 25

match, 23
MetaMap, 84
metric temporal conjunctive query with negation, *see* MTNCQ
MI, *see* myocardial infarction
MIMIC-III, 30
minimal, 34
minimal universal model, 34
  temporal, 70
model, 15
monotonic, 5
MTNCQ, 69
myocardial infarction, 105

natural language processing, 85
NCI Thesaurus, 3
negation, 14
next, 27
NLP, *see* natural language processing
nominals, 17
normal form
  $\mathcal{ELH}_{\bot}$, 20
  $\mathcal{TELH}_{\bot}^{\Diamond_{\mathsf{c}},\mathsf{lhs}}$, 60
  $\mathcal{TELH}_{\bot}^{\Diamond_{\mathsf{c}},\mathsf{lhs},-}$, 70
number restrictions, 17

OMQA, *see* ontology-mediated query answering, *see* ontology-mediated query answering
ontology, 2, 21
ontology-mediated query answering, 1, 3, 21
  temporalized, 1
open-world assumption, 5
OWA, *see* open-world assumption
OWL, *see* Web Ontology Language

precision, 109

qualified number restrictions, 17
QUELK, 97
query
  Boolean, 21, 69
  conjunctive, 22
  connected, 23
  first-order, 21
  rooted, 23, 69

recall, 109
relational database management system, 98
RIA, *see* role inclusion axiom
rigid, 6, 60
role
  name, 13
role hierarchies, 18
role inclusion axiom, 18

satisfiability, 14, 18
Semantic Web, 5
shrinking, 60
signature, 16
Skolemization, 33
SNA, *see* standard names assumption
SNOMED CT, 84
standard names assumption, 15
subsumption, 18

TBox, 2, 15

temporal logic, 6
term, 21
terminological box, *see* TBox
terminology, 13
top concept, 14
tractability, *see* PTIME
transitivity, 17
Turing machine, 19

universal model, 24
universal restriction, 14

variable, 21
  answer, 21
  leaf, 23
  quantified, 21
varying domains, 78

Web Ontology Language, 2