

Semantic Preferential Subsumption*

Katarina Britz

KSG, Meraka Institute, CSIR
South Africa
arina.britz@meraka.org.za

Johannes Heidema

Dept of Mathematical Sciences
University of South Africa
heidej@unisa.ac.za

Thomas Meyer

KSG, Meraka Institute, CSIR
South Africa
tommie.meyer@meraka.org.za

Abstract

We present a general preferential semantic framework for plausible subsumption in description logics, analogous to the KLM preferential semantics for propositional entailment. We introduce the notion of ordered interpretations for description logics, and use it to define two mutually dual non-deductive subsumption relations \sqsubseteq and \sqsubseteq^* . We outline their properties and explain how they may be used for inductive and abductive reasoning respectively. We show that the preferential semantics for subsumption can be reduced to standard semantics of a sufficiently expressive description logic. This has the advantage that standard DL algorithms can be extended to reason about our notions of plausible subsumption.

Introduction

Consider the following simplified scenario: A clinician at a rural clinic examines a patient who displays a number of symptoms – severe headache, fever and nausea. The clinician, seeking plausible explanations for these symptoms, identifies malaria and meningitis as two plausible causes. She considers each of these in turn, checking for symptoms such as a stiff and painful neck, eye sensitivity and muscular aches. Alerted to a strong possibility of meningitis, she arranges for a lumbar puncture. The initial Xpert EV test result for viral meningitis is negative, hence she commences a treatment of antibiotics for bacterial meningitis without waiting the expected ten days for the conclusive test results to be returned to the clinic.

This scenario illustrates the reasoning required during medical diagnosis and drug administration: “Having observed certain symptoms in a patient, which syndromes would plausibly explain them?” This is abduction, the process of seeking plausible, partial explanations for observations. “Having identified meningitis as a possibility, which further symptoms should typically be present in the patient?” This is induction, the process of inferring plausible consequences from assumptions. “Most cases of viral

meningitis would give a positive Xpert EV test result. Viral meningitis is therefore unlikely, but cannot be ruled out.” Again an abductive inference. “The negative Xpert EV result would be typical in a case of bacterial meningitis. It is therefore prudent to start administering antibiotic treatment immediately.” Again an inductive inference.

These forms of defeasible reasoning can benefit greatly from an intelligent knowledge system that supports defeasible reasoning. The system can suggest plausible (though possibly wrong) explanations for observations, and can suggest further tests and treatment programmes.

The knowledge representation formalism that has the best potential to deal with these forms of defeasible reasoning in complex structured domains is description logics (Baader, Horrocks, and Sattler 2008). Description logics have already gained wide acceptance as underlying formalism in intelligent medical knowledge systems and other application domains (Rector 2003; Baader, Horrocks, and Sattler 2008). The expressive power of a description logic (DL) is determined by the constructs available for building concept descriptions. The nature of DL reasoning has traditionally been *deductive*, but there have been a fair number of proposals to extend DLs to incorporate some form of defeasible reasoning, mostly centered around the incorporation of some form of default rules, e.g. (Donini, Nardi, and Rosati 2002).

An alternative model for defeasible reasoning is *preferential reasoning* (Kraus, Lehmann, and Magidor 1990); this was motivated and used in the proposal of (Giordano et al. 2007) for a preferential semantics of concept inclusion.

We present a general preferential semantic framework for defeasible subsumption in description logics, analogous to the KLM preferential semantics for propositional entailment. Following (Peirce 1974; Britz, Heidema, and Labuschagne 2007), we distinguish between two forms of defeasible reasoning, namely induction and abduction. We accordingly propose defeasible subsumption relations that lend themselves to these forms of reasoning.

The rest of the paper is structured as follows: We first fix some standard semantic terminology on description logics that will be useful later on. After giving some background on rational preference orders, we introduce the notion of an ordered interpretation. We then define two mutually dual preferential subsumption relations \sqsubseteq and \sqsubseteq^* semantically, in terms of their satisfaction by a fixed, ordered interpre-

*This material is based upon work supported by the National Research Foundation under Grant number 65152. Copyright © 2008, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

tation (Definitions 3 and 4 respectively), and discuss their respective properties.

Next, we present a formal semantics of defeasible subsumption, but now in terms of entailment by a knowledge base, i.e. with the semantics now constrained by all ordered interpretations satisfying the knowledge base (Definition 10).

The subsequent section outlines some of the inductive and abductive reasoning tasks that may be carried out using preferential subsumption \sqsubseteq and its dual \sqsubseteq^* respectively. Finally, we relate our semantics of entailment of Definition 10 to a standard DL semantics. We show that our abstract semantic framework for defeasible subsumption translates to a standard semantics in which preferential subsumption and its dual can be reduced to classical concept inclusion (Theorem 14). We also relate both these semantics to a Hilbert-style axiomatisation of rational preferential reasoning in a suitably expressive description logic.

Preliminaries

In the standard set-theoretic semantics of concept descriptions, concepts are interpreted as subsets of a domain of interest, and roles as binary relations over this domain. An interpretation I consists of a non-empty set Δ^I (the *domain* of I) and a function \cdot^I (the *interpretation function* of I) which maps each concept name A to a subset A^I of Δ^I , and each role name R to a subset R^I of $\Delta^I \times \Delta^I$. The interpretation function is extended to arbitrary concept descriptions (and role descriptions, if complex role descriptions are allowed in the language) in the usual way.

A DL knowledge base consists of a *Tbox* which contains *terminological axioms*, and an *Abox* which contains *assertions*, i.e. facts about specific named objects and relationships between objects in the domain. Depending on the expressive power of the DL, a knowledge base may also have an *Rbox* which contains *role axioms*.

Tbox statements are *concept inclusions* of the form $C \sqsubseteq D$, where C and D are (possibly complex) concept descriptions. $C \sqsubseteq D$ is also called a *subsumption statement*, read “ C is subsumed by D ”. An interpretation I satisfies $C \sqsubseteq D$, written $I \models C \sqsubseteq D$, iff $C^I \subseteq D^I$. $C \sqsubseteq D$ is *valid*, written $\models C \sqsubseteq D$, iff it is satisfied by all interpretations.

Objects named in the Abox are referred to by a finite number of *individual names*. These names may be used in two types of assertional statements – *concept assertions* of the form $C(a)$ and *role assertions* of the form $R(a, b)$, where C is a concept description, R is a role description, and a and b are individual names. To provide a semantics for Abox statements it is necessary to add to every interpretation an injective *denotation function* which satisfies the unique names assumption, mapping each individual name a to a different element a^I of the domain of interpretation Δ^I . An interpretation I satisfies the assertion $C(a)$ iff $a^I \in C^I$; it satisfies $R(a, b)$ iff $(a^I, b^I) \in R^I$.

Rbox statements may include amongst others, role inclusions of the form $R \sqsubseteq S$ and role axioms that are used to define role properties such as transitivity.

An interpretation I satisfies a DL knowledge base \mathcal{K} iff it satisfies every statement in \mathcal{K} . A DL knowledge base \mathcal{K}

entails a DL statement ϕ , written as $\mathcal{K} \models \phi$, iff every interpretation that satisfies \mathcal{K} also satisfies ϕ .

Preferential semantics

In a preferential semantics for a propositional language, one assumes some order relation on propositional truth valuations (or on interpretations or worlds or, more generally, on states) to be given. The intuitive idea captured by the order relation is that interpretations higher up (greater) in the order are more preferred, more normal, more likely to occur in the context under consideration, than those lower down. (For historical reasons the order is often inverted in the literature, but we shall follow (Shoham 1988, p.74) in taking upwards as the direction of increased preference. This also respects the direction of accessibility relations in modal logic which is relevant to this paper.)

Our choice of generalisation of preferential semantics to more expressive languages is guided by one possible underlying intuition of the meaning of the preference order. Namely, we assume that some objects in the application domain are viewed as more typical than others. This leads us to take as starting point a preference order on objects in the application domain. Hence we assume that the domain of any interpretation I , Δ^I , is ordered by a preference relation. We make the preference order on the domain of interpretation explicit through the notion of an *ordered interpretation*: (I, \leq) is the interpretation I with preference order \leq on the domain Δ^I .

The notion of an ordered first-order interpretation is naturally analogous to an order on propositional valuations: The information characterising a propositional valuation is the truth or falsity of atomic propositions. In the context of description logics, the information characterising an object in a given domain of interpretation is its membership (or not) to each named concept. The link between description logics and modal logics yields a closely related analogy. Namely, the information characterising a modal possible world is the truth or falsity of atomic propositions in that world. So, in a description logic setting, we assume a preference order on objects; in a modal logic setting, this translates to viewing the accessibility relation to be a preference order on possible worlds.

Elaborating further on the analogy between truth valuations in propositional logic and objects in description logics, we note that semantic entailment of propositions corresponds to subsumption of concepts: $\phi \models \psi$ denotes the statement “every model of ϕ is also a model of ψ ”, whereas $C \sqsubseteq D$ denotes the statement “every object in C is also an object in D ”. In (Britz, Heidema, and Labuschagne 2007) classical propositional entailment \models is generalised in two mutually dual ways to preferential entailments \vdash and \vdash^* ; analogously, here \sqsubseteq will be generalised to preferential subsumptions \sqsubseteq and \sqsubseteq^* in Definitions 3 and 4 below.

To ensure that the subsumption relations eventually generated are *rational* (Freund, Lehmann, and Morris 1991; Lehmann and Magidor 1992), we assume the preference order to be a *modular partial order*, i.e. a reflexive, transitive, anti-symmetric relation such that, for all a, b, c in Δ^I , if a and b are incomparable and a is strictly below c , then b is

also strictly below c . Modular partial orders have the effect of stratifying each Δ^I into layers, with any two elements in the same layer being unrelated to each other, and any two elements in different layers being related to each other. (We could also have taken the preference order to be a total pre-order, i.e. a reflexive, transitive relation such that, for all a, b in Δ^I , a and b are comparable. Since there is a bijection between modular partial orders and total preorders on Δ^I , it makes no difference for present purposes which formalism we choose.)

We further assume that the order relation is *Noetherian* (and hence, in Shoham's terminology (Shoham 1988, p.75), bounded, which is the dual of well-founded, which in turn implies, in the terminology of (Kraus, Lehmann, and Magidor 1990), that the order relation is smooth), i.e., there is no infinite strictly ascending chain of objects. In the presence of transitivity, the Noetherian property is equivalent to the following condition: For every nonempty subset X of Δ^I and $a \in X$ there is an element $b \in X$, maximal in X , with b greater than or equal to a .

Satisfaction of preferential subsumption

We introduced the notion of an ordered interpretation above. We now develop a formal semantics for description logics using ordered interpretations. We first introduce the notion of satisfaction by an ordered interpretation, thereafter we relax the semantics of concept inclusion to arrive at a definition of satisfaction of preferential subsumption relation \sqsubseteq by an ordered interpretation. In the next subsection, we define the dual \sqsubseteq^* of this relaxation. We also outline important properties of both \sqsubseteq and \sqsubseteq^* relative to a fixed ordered interpretation. This subsection and the next then set the scene for a generalisation from "satisfaction relative to a fixed interpretation" to "entailment relative to a knowledge base". This will be addressed in the following subsection.

Definition 1 An ordered interpretation (I, \leq) consists of an interpretation I and a Noetherian, modular partial order \leq over its domain Δ^I .

Definition 2 An ordered interpretation (I, \leq) satisfies $C \sqsubseteq D$, written $(I, \leq) \Vdash C \sqsubseteq D$, iff I satisfies $C \sqsubseteq D$.

We extend the semantics of \sqsubseteq to include more pairs (C, D) by shrinking C^I to a smaller set, namely the set C^{I^-} of maximally preferred or typical objects in the extension of C . The preferential semantics of the resulting relation \sqsubseteq is defined as follows:

Definition 3 An ordered interpretation (I, \leq) satisfies the preferential subsumption $C \sqsubseteq D$, written $(I, \leq) \Vdash C \sqsubseteq D$, iff $C^{I^-} \subseteq D^I$, where

$$C^{I^-} = \{x \in C^I \mid \text{for no } y \in C^I \text{ is } x \leq y \text{ but } y \not\leq x\}.$$

We make no assumption about which concept or role constructors are part of the DL language under consideration, but assume that, if present, the constructors \sqcap , \sqcup and \neg are interpreted in the standard way, ignoring the order \leq on I . Some of the properties of \sqsubseteq listed below (and of \sqsubseteq^* in the following subsection) may therefore be irrelevant in some

DLs. For example, rational monotonicity is only relevant in a DL which can express negated concepts.

Preferential subsumption \sqsubseteq is *supraclassical*, *nonmonotonic* and *defeasible*, in the following senses of these terms:

Supraclassicality: If $(I, \leq) \Vdash C \sqsubseteq D$ then $(I, \leq) \Vdash C \sqsubseteq D$.

Nonmonotonicity: $(I, \leq) \Vdash C \sqsubseteq D$ does not necessarily imply $(I, \leq) \Vdash C \sqcap C' \sqsubseteq D$.

Defeasibility: $(I, \leq) \Vdash C \sqsubseteq D$ does not necessarily imply $(I, \leq) \Vdash C \sqsubseteq D$.

Note that any strictly supraclassical relation is also defeasible, but that in general a defeasible relation need not be supraclassical.

The following properties of \sqsubseteq are analogous to the familiar properties of rational preferential entailment (Kraus, Lehmann, and Magidor 1990; Lehmann and Magidor 1992).

Reflexivity: $(I, \leq) \Vdash C \sqsubseteq C$.

And: If $(I, \leq) \Vdash C \sqsubseteq D$ and $(I, \leq) \Vdash C \sqsubseteq F$ then $(I, \leq) \Vdash C \sqsubseteq D \sqcap F$.

Or: If $(I, \leq) \Vdash C \sqsubseteq F$ and $(I, \leq) \Vdash D \sqsubseteq F$ then $(I, \leq) \Vdash C \sqcup D \sqsubseteq F$.

Left logical equivalence: If $(I, \leq) \Vdash C \sqsubseteq D$ and $(I, \leq) \Vdash D \sqsubseteq C$ and $(I, \leq) \Vdash C \sqsubseteq F$ then $(I, \leq) \Vdash D \sqsubseteq F$.

Left defeasible equivalence: If $(I, \leq) \Vdash C \sqsubseteq D$ and $(I, \leq) \Vdash D \sqsubseteq C$ and $(I, \leq) \Vdash C \sqsubseteq F$ then $(I, \leq) \Vdash D \sqsubseteq F$.

Right weakening: If $(I, \leq) \Vdash C \sqsubseteq D$ and $(I, \leq) \Vdash D \sqsubseteq F$ then $(I, \leq) \Vdash C \sqsubseteq F$.

Cautious monotonicity: If $(I, \leq) \Vdash C \sqsubseteq D$ and $(I, \leq) \Vdash C \sqsubseteq F$ then $(I, \leq) \Vdash C \sqcap D \sqsubseteq F$.

Rational monotonicity: If $(I, \leq) \Vdash C \sqsubseteq D$ and $(I, \leq) \not\Vdash C \sqsubseteq \neg F$ then $(I, \leq) \Vdash C \sqcap F \sqsubseteq D$.

Cut: If $(I, \leq) \Vdash C \sqcap D \sqsubseteq F$ and $(I, \leq) \Vdash C \sqsubseteq D$ then $(I, \leq) \Vdash C \sqsubseteq F$.

Satisfaction of dual preferential subsumption

The semantics of $C \sqsubseteq D$ relaxes that of $C \sqsubseteq D$ by requiring only that $C^{I^-} \subseteq D^I$ instead of $C^I \subseteq D^I$. An alternative relaxation of concept inclusion is obtained by dilating D^I to D^{I^+} , for some appropriate choice of D^{I^+} . Since C^{I^-} is the set comprising only the most preferred or typical objects in the extension of C , we take D^{I^+} to be the dual notion, i.e. the set of all objects except for the most preferred or typical objects not in the extension of D . We then define the preferential semantics of the dual preferential subsumption relation \sqsubseteq^* as follows:

Definition 4 An ordered interpretation (I, \leq) satisfies the dual preferential subsumption statement $C \sqsubseteq^* D$, written $(I, \leq) \Vdash C \sqsubseteq^* D$, iff $C^I \subseteq D^{I^+}$, where

$$D^{I^+} = D^I \cup \{x \in \Delta^I \mid \exists y \notin D^I \text{ with } x \leq y \text{ and } y \not\leq x\}.$$

We may think of D^{I^+} as obtained by adding to D^I those objects outside of D^I that are not maximally preferred. The intuition underlying this form of subsumption is that, should there (against expectations) be an object in the extension of C which is not in the extension of D (i.e. a counterexample

to $C \sqsubseteq D$), then this object or counterexample is abnormal or exceptional, being not amongst the normal or typical objects not in D^I .

The properties of \sqsupset^* (in the form of their analogues for preferential propositional entailment) are discussed in (Britz, Heidema, and Labuschagne 2007), where the relation was first defined in the context of abductive reasoning in propositional logic. Briefly, \sqsupset^* is the dual of \sqsubseteq , where the operation $()^*$ is given by $(C, D) \in \sqsupset^*$ iff $(\neg D, \neg C) \in \sqsubseteq$, i.e., \sqsupset^* and \sqsubseteq are *contrapositives*. While \sqsupset^* is also supra-classical and defeasible, it is *monotonic*, belying the tacit assumption often made that defeasible inference relations are by definition also nonmonotonic. The dual properties of \sqsupset^* are as follows:

Reflexivity: $(I, \leq) \Vdash C \sqsupset^* C$.

Or: If $(I, \leq) \Vdash C \sqsupset^* F$ and $(I, \leq) \Vdash D \sqsupset^* F$ then $(I, \leq) \Vdash C \sqcup D \sqsupset^* F$.

And: If $(I, \leq) \Vdash C \sqsubseteq D$ and $(I, \leq) \Vdash C \sqsupset^* F$ then $(I, \leq) \Vdash (C \sqsupset^* D \sqcap F)$.

Right logical equivalence: If $(I, \leq) \Vdash C \sqsubseteq D$ and $(I, \leq) \Vdash D \sqsubseteq C$ and $(I, \leq) \Vdash F \sqsupset^* C$ then $(I, \leq) \Vdash F \sqsupset^* D$.

Right defeasible equivalence: If $(I, \leq) \Vdash C \sqsupset^* D$ and $(I, \leq) \Vdash D \sqsupset^* C$ and $(I, \leq) \Vdash F \sqsupset^* C$ then $(I, \leq) \Vdash F \sqsupset^* D$.

Monotonicity: If $(I, \leq) \Vdash D \sqsupset^* F$ and $(I, \leq) \Vdash C \sqsubseteq D$ then $(I, \leq) \Vdash C \sqsupset^* F$.

Cautious right weakening: If $(I, \leq) \Vdash C \sqsupset^* D$ and $(I, \leq) \Vdash F \sqsupset^* D$ then $(I, \leq) \Vdash C \sqsupset^* D \sqcup F$.

Rational right weakening: If $(I, \leq) \Vdash C \sqsupset^* D$ and $(I, \leq) \not\Vdash \neg F \sqsupset^* D$ then $(I, \leq) \Vdash C \sqsupset^* D \sqcup F$.

Cautious right strengthening: If $(I, \leq) \Vdash C \sqsupset^* D \sqcup F$ and $(I, \leq) \Vdash F \sqsupset^* D$ then $(I, \leq) \Vdash C \sqsubseteq D$.

We conclude this subsection by pointing out two special cases where satisfaction of \sqsubseteq , \sqsupset and \sqsupset^* are directly related to each other. Proposition 5 shows that both preferential subsumption and its dual reduce to classical subsumption if the preference order is the identity relation.

Proposition 5 Consider an ordered interpretation $(I, =)$ in which the order on Δ^I is the identity relation on this set. For any concepts C and D we have

$$(I, =) \Vdash C \sqsubseteq D \text{ iff } (I, =) \Vdash C \sqsupset D \text{ iff } (I, =) \Vdash C \sqsupset^* D.$$

Proposition 6 shows how background information A can be captured by a (crude) preference order \leq_A .

Proposition 6 Consider any concept A and ordered interpretation (I, \leq_A) for which $\{A^I, \neg A^I\}$ is a proper dichotomy of Δ^I (i.e., both sets are non-empty), while

$$\leq_A^I := \{(x, y) \in \Delta^I \times \Delta^I \mid x \in \neg A^I \text{ and } y \in A^I\},$$

i.e., every element of A^I is strictly preferred to every element outside A^I .

Now let C and D be two concepts such that C is consistent with A ($A^I \cap C^I \neq \emptyset$), and $\neg D$ is also consistent with A ($A^I \not\subseteq D^I$). Then we have

$$I \Vdash A \sqcap C \sqsubseteq D \text{ iff } (I, \leq_A) \Vdash C \sqsupset D \text{ iff } (I, \leq_A) \Vdash C \sqsupset^* D.$$

Entailment of defeasible subsumptions

Satisfaction for preferential subsumption \sqsupset and its dual \sqsupset^* is defined relative to a fixed, ordered interpretation. We now take this a step further, and develop a general semantic theory of entailment relative to a knowledge base using ordered interpretations. Note that this does not yield a preferential entailment relation. The entailment relation \models of Definition 10 below is deductive, monotonic and, in a sense that we shall make precise in Theorem 14, classical.

The modular partial order \leq on domain elements in an ordered interpretation (I, \leq) may be constrained by means of a knowledge base. Namely, we extend the knowledge base Abox to allow role assertions of the form $a \preceq b$, such that the interpretation of \preceq is constrained to be that of \leq :

Definition 7 An ordered interpretation (I, \leq) satisfies an assertion $a \preceq b$ iff $a^I \leq b^I$.

As in the previous subsections, we do not make any further assumptions about the DL language at present, but assume that concept and role assertions, concept and role constructors, and classical subsumption are interpreted in the standard way, ignoring the preference order of ordered interpretations (see Definition 2).

Definition 8 The preferential subsumption statement $C \sqsupset D$ is valid, written $\models C \sqsupset D$, iff it is satisfied by all ordered interpretations (I, \leq) .

The dual preferential subsumption statement $C \sqsupset^* D$ is valid, written $\models C \sqsupset^* D$, iff it is satisfied by all ordered interpretations (I, \leq) .

It turns out that validity of preferential and of dual preferential subsumption statements are not very interesting, since they reduce to validity of classical subsumption.

Proposition 9 For any concepts C and D ,

$$\models C \sqsubseteq D \text{ iff } \models C \sqsupset D \text{ iff } \models C \sqsupset^* D.$$

It is only when subsumption is induced by a non-tautological knowledge base that the semantics diverges from that of classical subsumption:

Definition 10 A DL knowledge base \mathcal{K} entails the preferential subsumption statement $C \sqsupset D$, written $\mathcal{K} \models C \sqsupset D$, iff every ordered interpretation that satisfies \mathcal{K} also satisfies $C \sqsupset D$.

\mathcal{K} entails the dual preferential subsumption statement $C \sqsupset^* D$, written $\mathcal{K} \models C \sqsupset^* D$, iff every ordered interpretation that satisfies \mathcal{K} also satisfies $C \sqsupset^* D$.

The following properties of \sqsupset and \sqsupset^* are direct consequences of their corresponding properties relative to a fixed, ordered interpretation:

\sqsupset and \sqsupset^* are both supra-classical: If $\mathcal{K} \models C \sqsubseteq D$ then also $\mathcal{K} \models C \sqsupset D$ and $\mathcal{K} \models C \sqsupset^* D$.

\sqsupset is nonmonotonic: $\mathcal{K} \models C \sqsupset D$ does not necessarily imply that $\mathcal{K} \models C \sqcap C' \sqsupset D$.

\sqsupset^* is monotonic: If $\mathcal{K} \models C \sqsupset^* D$ then $\mathcal{K} \models C \sqcap C' \sqsupset^* D$.

\sqsupset and \sqsupset^* are both defeasible: Neither $\mathcal{K} \models C \sqsupset D$ nor $\mathcal{K} \models C \sqsupset^* D$ necessarily implies that $\mathcal{K} \models C \sqsubseteq D$.

The other properties of \sqsubseteq listed earlier relative to a fixed, ordered interpretation (i.e. reflexivity, and, or, left logical equivalence, left defeasible equivalence, right weakening, cautious monotonicity, rational monotonicity and cut) extend analogously in the context of entailment relative to a knowledge base. For example, reflexivity of \sqsubseteq relative to \mathcal{K} reads $\mathcal{K} \models C \sqsubseteq C$.

Similarly, the properties of \sqsubseteq^* listed earlier (i.e. reflexivity, or, and, right logical equivalence, right defeasible equivalence, cautious right weakening, rational right weakening, monotonicity and cautious right strengthening) also extend analogously.

To recap then, we have defined a non-standard description logic with a preferential semantics, which we shall refer to as *RDL* in the remainder of the paper. We take *RDL* to be obtained from an arbitrary standard DL by allowing Tbox statements of the form $C \sqsubseteq D$ and $C \sqsubseteq^* D$, and Abox statements of the form $a \preceq b$.

Forms of reasoning

A central classical DL reasoning task is the establishment of whether a knowledge base \mathcal{K} entails a concept inclusion $C \sqsubseteq D$, i.e. whether $\mathcal{K} \models C \sqsubseteq D$. Here, the semantics of both the entailment relation \models and the subsumption relation \sqsubseteq are *deductive*, the former allowing no counterexample interpretation I that satisfies \mathcal{K} but does not satisfy $C \sqsubseteq D$, and the latter allowing no counterexample object in C which is not in D , and so strictly preserving truth from “ x is in C ” to “ x is in D ”.

A second classical reasoning task is the establishment of whether a specific individual object x is an instance of a concept C , i.e. whether $\mathcal{K} \models C(x)$. Again, the reasoning is *deductive* in two senses: no allowance is made for plausible concept membership because \models is deductive, and because \mathcal{K} contains no plausible concept inclusions.

Although the entailment \models of Definition 10 is deductive, alternative reasoning patterns to deduction are obtained by relaxing the semantic constraints on concept inclusion (i.e., rather than relaxing the semantic constraints on entailment from the knowledge base). Defeasible subsumption statements may be included in the knowledge base itself; whether or not to allow this is an engineering design decision.

Besides deduction, the 19th century philosopher Charles Sanders Peirce recognised two more forms of rational reasoning, called *induction* and *abduction* (Peirce 1974; Britz, Heidema, and Labuschagne 2007). Though plausible, these are defeasible inference relations and allow counterexamples. Inductive reasoning takes a stance focussing on the forward direction of reasoning from a premiss to plausible consequences of that premiss, from a fact or observation to predictions or prognoses, from, e.g., “We know that x is in C ” to “Is it plausible that for that reason x is then also in D ?”. Abductive reasoning, in contrast, aims in the converse direction, from an established fact, construed as a putative consequence, to plausible premisses, reasons, diagnoses that would yield or explain this consequence, from “We know that x is in D ” to “Is it plausible that a contributing reason, maybe even a cause, for this is that x is in C ?”.

We propose $\mathcal{K} \models C \sqsubseteq D$ as an apt constraint on “ C is plausibly subsumed by D ” in the context of an inductive stance on the instance checking question. If object x is a counterexample to $C \sqsubseteq D$ but not to $C \sqsubseteq^* D$, then it is an atypical object (with respect to the preference order as constrained by \mathcal{K}) of class C and lies outside class D . If we now defeasibly infer from the fact that x is in C that it is also in D , we presumably do not have full information about x and C , namely that although x is in C , it is not a typical element of C . Upon learning the full, sad truth about x , we would have to retract the conclusion that x is in D . This is the characteristic pattern of inductive reasoning: from correct but possibly incomplete information as premiss, to plausible, but possibly wrong, conclusions.

Dually, we propose $\mathcal{K} \models C \sqsubseteq^* D$ as an apt constraint on “ C is plausibly subsumed by D ” in the context of an abductive stance on the instance checking problem. Now any counterexample to $C \sqsubseteq D$ but not to $C \sqsubseteq^* D$ is an atypical object of class $\neg D$ which lies in class C . Suppose we observe that object y is in D , and we seek some plausible explanation for this observation. Suppose that the knowledge base endorses $C \sqsubseteq^* D$, is it then rational to consider “ y is in C ” as plausible reason for the fact that y is in D ? Yes, because all objects in C are, if not in D , at least atypical of elements outside of D . This is in accordance with Peirce’s view of abduction as the reasoning process of seeking partial, defeasible explanations for observations. Upon gaining more information, we may learn that y is in fact not in C , and would have to retract our hypothesis. This is the characteristic pattern of abductive reasoning: From an observation, construed as a conclusion, to plausible, but possibly wrong, partial explanations yielding that conclusion defeasibly. The monotonicity of \sqsubseteq^* allows for the accumulation of partial explanations to build towards a complete explanation of evidence. We now make these notions more concrete by means of an example.

Employing clinical data

Clinical patient records contain a wealth of information relating symptoms, diseases, and treatment programmes. To simplify the example, we do not consider treatment programmes here, though the effects of drug administration are also highly relevant to defeasible reasoning.

Consider the following simple entry in a clinical record, where A , B and C are symptoms, and Y and Z are medical conditions:

Number	Date	A	B	C	...	Y	Z
03154	17/06/04	1	-	0	...	-	1

This entry indicates that on 17/06/04 patient 03154 had symptom A , did not have symptom C , and was known to suffer from condition Z . It does not indicate whether or not the patient had symptom B , or suffered from condition Y .

For each medical condition, there are a number of symptoms deemed relevant to that condition. Suppose symptoms A and C are relevant to condition Z . Each completed entry of an (A, C, Z) profile in the medical record system, i.e. each entry having values of 0 or 1 for A , C and Z , contributes to the generation of a preference order on pa-

tient profiles. This preference order represents the relative likelihood of occurrence of the eight possible (A, C, Z) profiles in an arbitrary patient, from most likely (at the top) to least likely (at the bottom). A crude order can be obtained by counting the number of occurrences of each profile in the available clinical databases; this may be refined with input from medical domain experts, and using formal methods such as formal concept analysis (Baader et al. 2007).

Suppose the following order has been generated for the profile (A, C, Z) :

100	
010	
011	111
101	000
001	
110	

Note that profiles are objects in the domain Δ^I representing the medical profiles of categories of patients. Profile 100 is deemed most typical in this order, representing patients with symptom A , without symptom C , and who do not suffer from condition Z . Likewise, the least likely profile is that of patients with both symptoms A and C , and who yet do not suffer from condition Z .

For each medical condition, such a preference order on medical profiles relevant to that condition may be extracted and generated from the medical record system. We propose the use of these orderings in preferential reasoning to determine plausible explanations of, and predictions on, factors related to patient diagnosis and treatment.

To illustrate how this may work, consider the following question: ‘‘Given that a patient has symptom A , which medical conditions would plausibly explain or predict A ?’’ This question can now be approached by qualitative defeasible reasoning, different from quantitative probabilistic and statistical reasoning. In order to answer this question, we have to consider each medical condition relevant to symptom A in turn. In our example, one of these would be condition Z .

In what follows $(I, <)$ refers to the obviously relevant ordered interpretation(s). We note that A does not predict Z , nor does Z predict A (neither $(I, <) \Vdash A \sqsubseteq Z$ nor $(I, <) \Vdash Z \sqsubseteq A$ holds). So, although Z would explain A ($(I, <) \Vdash Z \sqsubseteq^* A$), the causal relationship between Z and A is weak.

Further, since $(I, <) \Vdash Z \sqsubseteq C$, we may inductively predict that a patient suffering from condition Z should also have symptom C .

The population from which the preference order is generated impacts strongly on the usefulness of subsequent preferential reasoning. One simple way to restrict the population is to consider the order relative to a given set of symptoms. For example, assuming A as given, we may consider only profiles of the form $(1, C, Z)$ in the preference order above. This yields the following preference order:

100
111
101
110

Now suppose a patient, manifesting A , tests negative for symptom C . Z can then no longer be offered as a plausible explanation of this symptom profile, since it is neither the case that $(I, <) \Vdash Z \sqsubseteq^* A \sqcap \neg C$, nor that $(I, <) \Vdash A \sqcap \neg C \sqsubseteq Z$. On the other hand, should the patient test positive for C , and noting that all of the following are satisfied by $(I, <)$: $Z \sqsubseteq^* A \sqcap C$; $Z \sqsubseteq A \sqcap C$; $A \sqcap C \sqsubseteq Z$, and even $A \sqcap C \sqsubseteq^* Z$, it is highly plausible that the patient suffers from condition Z .

In a more realistic setting, the relationship between symptoms and medical conditions would be structured by Tbox axioms in an ontology. Defeasible subsumption relations may also be added to the Tbox, relating symptoms, medical conditions and treatment programmes. How this may be translated to deductive DL reasoning is the topic of the next section.

Preferential reasoning in \mathcal{ALC}^\sim

The mutually dual preferential subsumption relations \sqsubseteq and \sqsubseteq^* are produced by a preference order on the intended application domain; their semantics are captured by Definitions 3 and 4 respectively, while the semantics of entailment (and hence, of knowledge base satisfiability) using ordered interpretations is captured by Definition 10. We now relate this semantics to a standard first-order semantics for description logics. The latter may be defined in at least two ways: The first option is to add constructors for defeasible subsumption as primitive relations to a suitable DL, and then develop special-purpose algorithms to reason about them. In essence, this is analogous to the approach taken in (Giordano et al. 2007). The second option is to express \sqsubseteq and \sqsubseteq^* as defined relations in a sufficiently expressive DL, preferably one for which algorithms have already been developed.

We adopt the latter approach. To our knowledge, a tailor-made DL for our purpose has not been proposed in the literature; neither do we construct a proof system or analyse the complexity of such a logic in this paper. We shall address the algorithmic aspects required by our semantic framework in a subsequent paper, restricting this paper to the presentation of a semantic framework that allows for a range of algorithmic alternatives and refinements.

The logic of interest to us in this section is an extension of the well-known DL \mathcal{ALC} (Schmidt-Schauß and Smolka 1991), hence we first introduce the relevant \mathcal{ALC} terminology. Concept descriptions are built from concept names using the constructors disjunction $(C \sqcup D)$, conjunction $(C \sqcap D)$, negation $(\neg C)$, existential restriction $(\exists R.C)$ and value restriction $(\forall R.C)$, where C, D denote concepts and R a role name. The semantics of the constructors are defined by:

$$\begin{aligned}
(\neg C)^I &= \Delta^I \setminus C^I; \\
(C \sqcup D)^I &= C^I \cup D^I; \\
(C \sqcap D)^I &= C^I \cap D^I; \\
(\exists R.C)^I &= \{x \mid \exists y \text{ s.t. } (x, y) \in R^I \text{ and } y \in C^I\}; \\
(\forall R.C)^I &= \{x \mid \forall y, (x, y) \in R^I \text{ implies } y \in C^I\}.
\end{aligned}$$

In the context of this paper, the order on objects in the application domain may be supplied by a domain expert. As such,

it may well not automatically have the properties we require of it. We constrain the interpretation of a preference order on objects in the language by means of a role axiom added to the knowledge base Rbox. Of course, one has to start with an intuitively sensible order on objects in the domain in order to obtain an intuitively sensible preference relation.

There is a natural bijection between modular partial orders (modular, reflexive, transitive, antisymmetric relations) and asymmetric, modular, transitive (hence also irreflexive and antisymmetric) relations over the same domain – the latter is just the strict counterpart of the former, while the reflexive closure of an asymmetric, modular, transitive relation is a modular partial order. We can constrain a role to be interpreted as a Noetherian, asymmetric, modular, transitive relation by adding a role axiom to the language of \mathcal{ALC} :

Definition 11 Let \mathcal{ALC}^\sim denote the extension of the DL \mathcal{ALC} obtained by allowing Rbox role axioms of the form: $Rational(R)$, where R is a role name. An axiom of the form $Rational(R)$ is true in an interpretation I , written as $I \models Rational(R)$, iff R^I is a Noetherian, asymmetric, modular, transitive relation.

We are now in a position to translate all statements expressible in the non-classical DL RDL into statements of a suitable classical DL, which we shall refer to as RDL^t . Let L be a specific instance of the language for RDL . That is, L generates a set of Abox, Tbox and Rbox statements, including statements of the form $C \sqsubseteq D$, $C \sqsubset D$, $C \sqsubset^* D$, and $a \preceq b$ (and possibly others), where C and D are concepts constructed from a fixed set of concept and role constructors, and with a fixed alphabet \mathcal{A} of object names, concept names, and roles names. Then we define L^t as the extension of the language of \mathcal{ALC}^\sim containing all concept and role constructors of RDL , and allowing all Abox, Tbox and Rbox statements of RDL , except for statements of the form $C \sqsubset D$, $C \sqsubset^* D$, and $a \preceq b$. The alphabet \mathcal{A}^t from which L^t is generated is obtained by extending the alphabet \mathcal{A} by adding to it a single new role name, say R , not occurring in \mathcal{A} .

Definition 12 For every statement ϕ in L , the translation ϕ^t of ϕ is defined as follows:

- If ϕ has the form $a \preceq b$ then $\phi^t = R(a, b)$;
- If ϕ has the form $C \sqsubseteq D$ then $\phi^t = C \sqcap \forall R. \neg C \sqsubseteq D$;
- If ϕ has the form $C \sqsubset^* D$ then $\phi^t = C \sqsubseteq D \sqcup \exists R. \neg D$;
- Otherwise $\phi^t = \phi$.

It is clear that the translations of statements in L are all statements in L^t . The role R is introduced to express preference statements of the form $a \preceq b$ as classical role assertions. Furthermore, in order to ensure that R is always interpreted as a Noetherian, asymmetric, modular, transitive relation, it is necessary to perform reasoning relative to a knowledge base containing at least the role axiom $Rational(R)$ in the Rbox. Moreover, both forms of defeasible subsumption are expressed as forms of classical subsumption. For \sqsubset this is achieved by exploiting the fact that, for any interpretation I of the RDL^t , $C \sqcap \forall R. \neg C$ is interpreted as the maximally preferred objects (relative to the preference order R^I) in the extension of C . So $(C \sqcap \forall R. \neg C)^I$ plays the role of the

shrunk set C^{I^-} in Definition 3. Similarly, for \sqsubset^* this is achieved by exploiting the fact that, for any interpretation I of the RDL^t , $D \sqcup \exists R. \neg D$ is interpreted as the set of objects in Δ^I that are not maximally preferred objects (relative to the preference order R^I) in the complement of the extension of D (i.e. $\Delta^I \setminus D^I$). So $(D \sqcup \exists R. \neg D)^I$ plays the role of the dilated set D^{I^+} in Definition 4.

We now proceed to prove that this translation is correct.

Definition 13 Let \mathcal{K} be a knowledge base expressed in L . The translation of \mathcal{K} , referred to as \mathcal{K}^t , is defined to be the knowledge base (expressed in L^t) obtained as follows: $\{\phi^t \mid \phi \in \mathcal{K}\} \cup \{Rational(R)\}$.

So \mathcal{K}^t is obtained by replacing every statement in \mathcal{K} by its translation, and adding the statement $Rational(R)$.

Theorem 14 For any knowledge base \mathcal{K} expressed in L and any statement ϕ in L we have that $\mathcal{K} \models \phi$ iff $\mathcal{K}^t \models \phi^t$.

Theorem 14 tells us that:

- checking whether $C \sqsubset D$ follows from \mathcal{K} can be reduced to checking whether $C \sqcap \forall R. \neg C \sqsubseteq D$ follows from \mathcal{K}^t in RDL^t .
- checking whether $C \sqsubset^* D$ follows from \mathcal{K} can be reduced to checking whether $C \sqsubseteq D \sqcup \exists R. \neg D$ follows from \mathcal{K}^t in RDL^t .
- checking whether $a \preceq b$ follows from \mathcal{K} can be reduced to checking whether $R(a, b)$ follows from \mathcal{K}^t in RDL^t ;
- checking whether any other statement ϕ in L follows from \mathcal{K} can be reduced to checking whether ϕ itself follows from \mathcal{K}^t in RDL^t .

Role transitivity has been investigated in (Schild 1991). Asymmetry of roles is included in the underlying DL language for OWL 1.1, $SR\mathcal{OIQ}$ (Horrocks, Kutz, and Sattler 2006), although we have no need for the full expressivity of that logic. To our knowledge, modular and Noetherian roles have not been investigated, and neither has the simultaneous enforcement of these two properties together with role asymmetry and transitivity. However, we do not expect such a restriction to impact negatively on the complexity of entailment since we are already concerned with DLs at least as expressive as the DL \mathcal{S} (i.e. \mathcal{ALC} plus transitive roles) for which checking entailment is EXPTIME-hard (Schild 1991).

Axiomatisation

The axiomatisation of propositional rational preferential reasoning was considered in (Britz, Heidema, and Labuschagne 2007), establishing modular Gödel-Löb logic as the underlying modal logic for rational inductive and abductive reasoning. This gives a syntactic Hilbert-style axiomatisation of rational preferential reasoning.

Modular GL is the tense logic with modal operators \Box and its converse \Box^c . It is obtained from the minimal tense logic \mathbf{K}_t (Blackburn, de Rijke, and Venema 2002, p.205) by adding the transitivity axiom, Löb axiom (Boolos 1993), and

weak modularity axioms for \Box and \Box^c :

$$\begin{aligned} \text{Modular GL} &= \mathbf{K}_t \oplus \Box X \rightarrow \Box \Box X \\ &\oplus \Box(\Box X \rightarrow X) \rightarrow \Box X \\ &\oplus \Box(\Box X \rightarrow Y) \vee \Box(\Box Y \rightarrow \Box X) \\ &\oplus \Box^c(\Box^c X \rightarrow Y) \vee \Box^c(\Box^c Y \rightarrow \Box^c X) \end{aligned}$$

These axiom schemas can all be expressed in \mathcal{ALC}^\sim with role converses, denoted R^- below (and usually referred to as role inverses in the DL community), added:

Transitivity: $\vdash \forall R.C \sqsubseteq \forall R.(\forall R.C)$;

Löb: $\vdash \forall R.((\exists R.\neg C) \sqcup C) \sqsubseteq \forall R.C$;

Weak modularity: $\vdash \exists R.(\forall R.C \sqcap \neg D) \sqsubseteq \forall R.(\exists R.\neg D \sqcup \forall R.C)$;

Converse weak modularity: $\vdash \exists R^-.(\forall R^-.C \sqcap \neg D) \sqsubseteq \forall R^-.(\exists R^-. \neg D \sqcup \forall R^-.C)$

This axiomatisation is a syntactic counterpart for both the ordered preferential semantics and the standard DL semantics related in Theorem 14.

The semantics of description logics with additional modal operators representing preferences lends itself naturally to many-dimensional structures (Gabbay et al. 2003) – an object dimension as is usual for the underlying description language, and one or more dimensions to represent preference orders. The specific choices within this semantic paradigm require further investigation. The authors would like to thank Ulrike Sattler for pointers in this regard.

Conclusion and future work

We have presented a non-standard semantic framework for preferential reasoning in description logics, based on the notion of ordered interpretations, and which can be expressed in the object language by virtue of the introduction of two mutually dual relations for plausible subsumption, \sqsubseteq and \sqsubseteq^* , as well as preference statements on objects of the form $a \preceq b$, leading to the definition of a non-classical description logic *RDL*. Both \sqsubseteq and \sqsubseteq^* are well-behaved in the sense that they satisfy desirable properties for rational, defeasible subsumption. An additional advantage is that reasoning in *RDL* can be reduced to reasoning in a sufficiently expressive classical DL, i.e. in extensions of \mathcal{ALC}^\sim .

While the work done in this paper focuses on a single mode of preference, the ordered semantics developed here extends readily to multiple preference orders on objects, as one would typically have in a DL. The notion of an ordered interpretation (I, \preceq) extends to an interpretation with a finite set of preference orders $(I, \preceq_1, \dots, \preceq_n)$. The defeasible subsumption relations \sqsubseteq and \sqsubseteq^* generalise accordingly to $\sqsubseteq_1, \dots, \sqsubseteq_n$ and $\sqsubseteq_1^*, \dots, \sqsubseteq_n^*$ respectively.

The algorithmic aspects, complexity analysis, and evaluation of a prototype system to determine the extent to which both inductive and abductive plausible subsumption can provide useful guidelines in a medical scenario such as that outlined in the introduction and example are planned as part of a wider research project. So, too, is the integration of the medical ontology SNOMED CT into the clinical medical record framework OpenMRS (Meyer et al. 2008).

References

- Baader, F.; Ganter, B.; Sattler, U.; and Sertkaya, B. 2007. Completing description logic knowledge bases using formal concept analysis. In Golbreich, C.; Kalyanpur, A.; and Parsia, B., eds., *Proceedings of OWLED 2007: Workshop on OWL: Experiences and Directions*. CEUR-WS.org.
- Baader, F.; Horrocks, I.; and Sattler, U. 2008. Description logics. In van Harmelen, F.; Lifschitz, V.; and Porter, B., eds., *Handbook of Knowledge Representation*. Amsterdam: Elsevier. 135–180.
- Blackburn, P.; de Rijke, M.; and Venema, Y. 2002. *Modal Logic*, volume 53 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge, UK: Cambridge University Press.
- Boolos, G. 1993. *The Logic of Provability*. Cambridge, UK: Cambridge University Press.
- Britz, K.; Heidema, J.; and Labuschagne, W. 2007. Entailment, duality, and the forms of reasoning. Submitted. URL <http://www.cs.otago.ac.nz/research/publications/OUCS-2007-01.pdf>.
- Donini, F.; Nardi, D.; and Rosati, R. 2002. Description logics of minimal knowledge and negation as failure. *ACM Transactions on Computational Logic* 3(2):177–225.
- Freund, M.; Lehmann, D.; and Morris, P. 1991. Rationality, transitivity and contraposition. *Artificial Intelligence* 52(2):191–203.
- Gabbay, D.; Kurucz, A.; Wolter, F.; and Zakharyashev, M. 2003. *Many-Dimensional Modal Logics: Theory and Applications*, volume 148 of *Studies in logic and the foundations of mathematics*. Amsterdam: Elsevier North Holland.
- Giordano, L.; Gliozzi, V.; Olivetti, N.; and Pozzato, G. 2007. Preferential description logics. In Dershowitz, N., and Voronkov, A., eds., *Logic for Programming, Artificial Intelligence, and Reasoning: Proceedings of the 14th International Conference (LPAR 2007)*, volume 4790 of *LNCS*. Berlin / Heidelberg: Springer. 257–272.
- Horrocks, I.; Kutz, O.; and Sattler, U. 2006. The even more irresistible *SRIOQ*. In Doherty, P.; Mylopoulos, J.; and Welty, C., eds., *Principles of Knowledge Representation and Reasoning: Proceedings of the Tenth International Conference (KR2006)*. Menlo Park, CA: AAAI Press.
- Kraus, S.; Lehmann, D.; and Magidor, M. 1990. Non-monotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence* 44:167–207.
- Lehmann, D., and Magidor, M. 1992. What does a conditional knowledge base entail? *Artificial Intelligence* 55:1–60.
- Meyer, T.; Seebregts, C.; Gerber, A.; Britz, K.; Pretorius, L.; Alberts, R.; and Moodley, D. 2008. The integration of SNOMED CT into the OpenMRS Electronic Medical Record System. In Cornet, R., and Spackman, K., eds., *Proceedings of KR-MED 2008: Representing and sharing knowledge using SNOMED*. CEUR-WS.org.
- Peirce, C. 1974. In Hartshorne, C., and Weiss, P., eds., *Collected Papers of Charles Sanders Peirce: 1931–1935*, volume I-VI. Cambridge, MA: Harvard University Press.

Rector, A. 2003. Medical informatics. In Baader, F.; Calvanese, D.; McGuinness, D.; Nardi, D.; and Patel-Schneider, P., eds., *The Description Logic Handbook: Theory, implementation, and applications*. Cambridge, UK: Cambridge University Press. 415–435.

Schild, K. 1991. A correspondence theory for terminological logics: Preliminary report. In *Proceedings of the 12th International Joint Conference on Artificial Intelligence (IJCAI'91)*. Menlo Park, CA: AAAI Press. 466–471.

Schmidt-Schauß, M., and Smolka, G. 1991. Attributive concept descriptions with complements. *Artificial Intelligence* 48:1–26.

Shoham, Y. 1988. *Reasoning about Change: Time and Causation from the Standpoint of Artificial Intelligence*. Cambridge, MA: The MIT Press.