Two Kinds of Non-Monotonic Analogical Inference*

Manfred Kerber[1] and Erica Melis[2]

[1] The University of Birmingham, School of Computer Science
Birmingham, B15 2TT, England
e-mail: M.Kerber@cs.bham.ac.uk
WWW: http://www.cs.bham.ac.uk/~mmk

[2] Universität des Saarlandes, FB Informatik
D-66041 Saarbrücken, Germany
e-mail: melis@cs.uni-sb.de
WWW: http://jswww.cs.uni-sb.de/~melis

Abstract. This paper addresses two modi of analogical reasoning. The first modus is based on the explicit representation of the justification for the analogical inference. The second modus is based on the representation of typical instances by concept structures. The two kinds of analogical inferences rely on different forms of relevance knowledge that cause non-monotonicity. While the uncertainty and non-monotonicity of analogical inferences is not questioned, a semantic characterization of analogical reasoning has not been given yet. We introduce a minimal model semantics for analogical inference with typical instances.

1 Introduction

Analogical reasoning is a process whereby similarities between a source and a target are used to infer the probable existence of further similarities. Thus, under certain conditions, an analogical inference can be employed to provide a description of an aspect of the target, if the description of the same aspect is known for the source.

What are the conditions that justify an analogical inference? There are several answers to this question, but most of them agree in characterizing some relevance knowledge as a justification for analogical inferences. Goebel [12], for instance, emphasizes some *relevant similarity* as the knowledge required for an analogical inference. Gentner [9] prefers by her systematicity criterion a mapping of predicates which are connected by a higher order relation explicitly to be known for analogical inferences. The higher order relations she examined, namely *cause* and *implies*, actually represent relevance knowledge of the form "aspect A_1 is relevant for aspect A_2".

* Published in FAPR'96. This work was supported by the Deutsche Forschungsgemeinschaft, SFB 314 (D2).

The maybe best-recognized proponent of justification is Russell, who defined the so-called determinations as the crucial relevance knowledge. His total determinations represent a strict form of connections between two aspects, expressing "the values of a predicate A_1 determine the values of a predicate A_2".

Let us look at an example: Knowing that the car car_S was produced in 1990, that its make is Honda-civic, and that its price is 13,000.00 DM you would infer analogically that a car car_T which was produced in 1990 and the make of which is Honda-civic would probably cost about 13,000.00 DM. The relevance knowledge that justifies this particular analogical inference is the determination of the price of a car by its make and year of construction. But what happens if you gain the additional knowledge that car_T is rusty and was involved in many severe accidents? Then you would no longer infer the price of 13,000.00 DM for car_T. This example provides a clue for the non-monotonicity of inferences called connection-based analogical inferences.

For understanding the non-monotonicity of another kind of analogical reasoning, remember the classical Tweety story of non-monotonic reasoning: In this story the non-monotonicity has a lot to do with the break down of an inheritance in the concept "bird". In (semantic net and frame) representations that use prototypes, such as the TypicalElephant of [4], certain analogical inferences provide new information by copying facts from the prototype to the individual. Even though known as "inheritance", actually this reasoning is another common kind of analogical reasoning, which we shall call typical-example-based. As the Tweety story shows, it is non-monotonic as well.

Analogical inference is commonly considered to be non-deductive, hypothetical, tentative, and non-monotonic. These features of analogical reasoning have been addressed explicitly in [12, 13, 24, 19] and implicitly presupposed in many approaches to reasoning by analogy. Other questions have usually been the center of attention of analogy research though. We discuss the additional justifying knowledge needed for inferring a further similarity from a similarity of a source and a target for two kinds of analogical reasoning. We propose how to represent this knowledge and present a semantic characterization for the analogical inference with typical instances that is actually designed for non-monotonic logics.

This paper is organized as follows: We first recall the basic ingredients of connection-based analogical reasoning. Then we introduce analogical reasoning based on typical instances, which makes use of an exemplary knowledge representation. We present the inference schemas for the two forms of analogical reasoning and discuss a semantic that is appropriate for the analogical inference with typical instances and which captures its non-monotonicity. At last we propose a hybrid framework integrating these two forms of analogical reasoning using a hybrid knowledge representation[3].

[3] Thereby we add to the debate about logic-based versus non-propositional knowledge representation in artificial intelligence, which dates back to McCarthy's Advice Taker [18] and Sloman's analogical representations [26].

2 Connection-Based Analogical Reasoning

An important concept in analogical reasoning is the so-called aspect. In many approaches to computational analogy, aspects are represented just by predicates or fixed formulae. In order to capture a wider range of analogical reasoning and to include inter-domain analogies, we define an *aspect* A as a partial function, mapping the individuals (instances c of a concept C) to non-tautological formulae with at most one free variable x such that $A\langle c\rangle[x/c]$ is `true`[4].

Informally speaking, if c is an instance and A an aspect then $A\langle c\rangle$ is a formula describing A of c. For example: The value of $\mathit{safety}\langle c\rangle$ of a car c might be $\mathit{airbag}(x) \wedge \mathit{antiblock}(x) \wedge \mathit{max\_speed}(x) < 100$. It is assumed that $\mathit{airbag}(c) \wedge \mathit{antiblock}(c) \wedge \mathit{max\_speed}(c) < 100$ is `true`. For a bicycle b, $\mathit{safety}\langle b\rangle$ might be $\mathit{frame\_diameter}(x, 3) \wedge \mathit{age}(x) < 10$.

The transfer of the value of an aspect A_2 from a source case s to a target case t based on the similarity of s and t with respect to another aspect A_1 is the *standard form of justified reasoning by analogy*, investigated, e.g. , in [7]. This kind of analogical inference requires for its justification some relevance knowledge expressing that A_1 is relevant for A_2. The relevance knowledge "A_1 is relevant for A_2", written as $[A_1, A_2]$, is usually represented explicitly and propositionally, as determinations [7], as schemata [11], as connections [19], or as similarity transforms [6]. In modeling human analogical reasoning, the relevance knowledge must allow for exceptions and uncertainty rather than being a logical implication or a total determination. In the following we use the most general notion, connection, that does not require a specific representation and may have exceptions.

A *connection* is a pair of aspects $[A_1, A_2]$. An example of such connections is $[\mathit{population}, \mathit{cars}]$ expressing "if two cities have the same number of inhabitants, then probably the same number of cars is registered in the cities". This connection is used in an analogical inference that yields a value of the aspect *cars* for the target instance *Rome*. This inference takes as inputs the similarity of *Rome* and another city, say *Madrid*, with respect to the number of inhabitants and the connection. It infers the correspondence of *Rome* and *Madrid* with respect to the aspect *cars*. Using the additional information of the actual value of $\mathit{cars}\langle \mathit{Madrid}\rangle$ the value of $\mathit{cars}\langle \mathit{Rome}\rangle$ can be inferred.

Utilizing connections, the connection-based analogical inference can formally be described by the schema

$$\textbf{(AR)} \qquad \frac{A_1\langle t\rangle = A_1\langle s\rangle, [A_1, A_2]}{A_2\langle t\rangle := A_2\langle s\rangle}$$

A connection-based analogical inference is confirmed only if the resulting target aspect $A_2\langle t\rangle$ does not contradict the knowledge inferable for the target by deduction. *The uncertainty of the connection and the consistency constraint*

[4] Note that $A\langle c\rangle$ is a formula with a free variable x. $A\langle c\rangle[x/c]$ denotes the formula $A\langle c\rangle$ in which the free variable x is substituted by c.

are responsible for the non-monotonicity of the connection-based analogical inference.

But, what happens, if such explicit connections are not available? The kind of analogical inference presented next allows for justified analogical reasoning that is not necessarily based on explicit connections.[5]

3 Analogical Reasoning Based on Typical Instances

Russell [24] tries to interprete analogical inference based on typical instances as a connection-based inference with "belonging to the same class" as the aspect A_2 of a connection $[A_1, A_2]$. However, psychological investigations provided evidence that typical instances are the only, or at least the preferred sources for analogical reasoning corresponding to inheritance among instances of a concept. Consequently, connection-based analogical inference does not cover analogical inferences which can have only a typical instance as the source rather than an arbitrary instance.

The typicality of instances of concepts is a phenomenon investigated in empirical psychology [17, 20, 23]. Reproducible typicality ratings that distinguish typical instances have been found. Some experimental methods [17] for the extraction of this typicality rating are the direct rating of representativity, the examination of the reaction time to decide whether an instance belongs to a category, the test of the reproduction of instances, and the use of instances in generalizations and in analogical reasoning. Hence you find the notation *structured concept* in the psychological literature that refers to a set of instances with a typicality relation which we denote[6] by $\sqsubseteq_C$. The notion "concept structure" denotes the structure of one single concept rather than the relationship between different concepts like in KL-ONE. The typicality relations $\sqsubseteq_C$ compare for each concept C the degrees of typicality (for example, a hammer is commonly considered a more typical tool than a compass saw) of similar instances. Within such concept structures we define typical instances:

DEFINITION: An instance $c \in C$ is called a *typical instance* if it is maximal (that is, there is no $c' \neq c$ with $c \sqsubseteq_C c'$), written as typex(c).

For example, a hammer is commonly considered a typical tool and a violin a typical musical instrument. An example of a concept structure is given in figure 1.

There are two kinds of aspects of typical instances: aspects that are important for the typicality of the instance (e.g. the size of a city) and aspects that are accidental (e.g. the number of research institutes of a city). Of course, justified analogical reasoning transfers only relevant information. This motivates the

[5] Although the field of analogical reasoning is concerned with reasoning based on examples, surprisingly the importance of reasoning by *typical* instances, as for example, investigated by Rosch [21] and Lakoff [17], has not been elaborated. Only an attempt of Winston [30] was influenced by statistic prototypicality.

[6] The subscript specifies the concept.

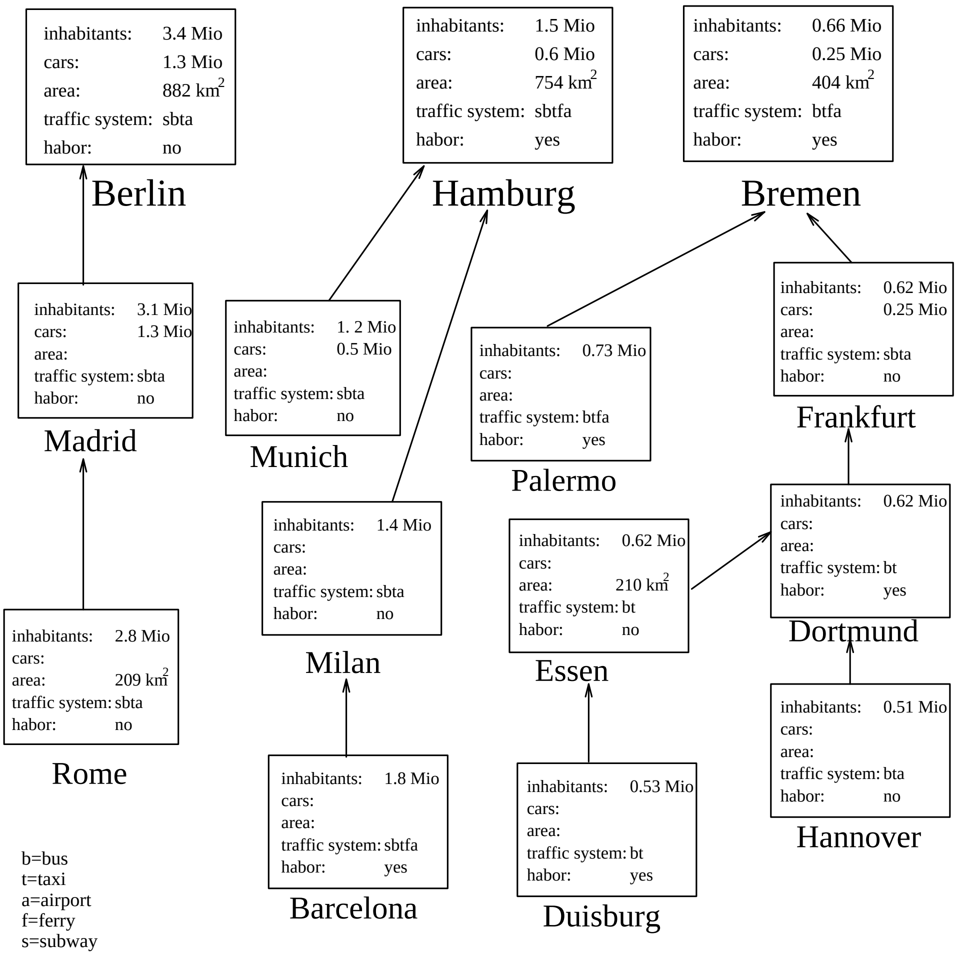

Fig. 1. A concept structure of cities

following definition that describes a relevance different from that in the previous section:

DEFINITION: An aspect A is called *relevant* for an instance c (written as $relevant(A, c)$) iff $A\langle c\rangle[x/c]$ is true and for all instances c' with $c' \sqsubseteq_C c$ holds $A\langle c\rangle[x/c']$ is either true or undefined.

Analogical inferences based on typical instances can formally be described by the schema

$$(\mathbf{AT}) \qquad \frac{\text{typex}(s), t \sqsubseteq_C s, relevant(A_2, s)}{A_2\langle t\rangle := A_2\langle s\rangle}$$

Again, an analogical inference based on typical instances is confirmed only if the result does not contradict the knowledge inferable for the target by deduction. *This consistency constraint and the uncertainty of the knowledge about the relevance of an aspect for a typical instance of a concept are responsible for the non-monotonicity of analogical inference based on typical instances.*

Relationship between AR and AT

AT can be heuristically justified in terms of the known justification of **AR**: Let s and t be examples, i.e. elements of a concept C, and let A_1 and A_2 be aspects. Assume:

(i) $\text{typex}(s)$, $t \sqsubseteq_C s$, $relevant(A_2, s)$.

According to Weiner [29] it follows heuristically that if s is a typical example and $t \sqsubseteq_C s$, then there is an aspect A_1 with $relevant(A_1, s)$ and $A_1\langle t\rangle = A_1\langle s\rangle$. Hence, we have

(ii) $A_1\langle t\rangle = A_1\langle s\rangle, relevant(A_1, t), relevant(A_2, t), \text{typex}(t)$.

Because of the definition of "relevant", $relevant(A_1, s)$, $relevant(A_2, s)$, and $\text{typex}(s)$ support the connection $[A_1, A_2]$ (at least for all instances rated under s). Hence, from (ii) follows

(iii) $A_1\langle t\rangle = A_1\langle s\rangle, [A_1, A_2]$.

Because of (iii) and the inference rule **AR** we have

(iv) $A_2\langle t\rangle := A_2\langle s\rangle$.

That is, we have the permission to define A_2 for t by $A_2\langle s\rangle$. Thus if **AR** is justified, then **AT** is heuristically justified.

A Semantics for Analogical Reasoning with Typical Instances

The common way to cope with the meaning of logical formulae and inferences is to find an appropriate semantics. For classical logic, the Tarski semantics has been defined, for modal logics, Kripke semantics proved to be appropriate. Several semantics have been developed for non-monotonic logics, e.g., Shoham's minimal model semantics. We shall relate the analogical inferences based on typical instances to this minimal model semantics by introducing interpretations which are compatible with the concept structure and by defining a partial order on these interpretations.

Let $\partial\mathcal{I}$ be a partial interpretation of formulae, that assigns to each pair (*instance*, *formula with one variable*) one of the values `true`,`false`,`unknown`. It can be thought of as the result of an inspection of the instances. In the case of our city example

$$\partial\mathcal{I}(Rome, no\_of\_inhabitants(x) \doteq 3million) = \texttt{true}$$

since a corresponding entry can be found in the concept structure. In contrast $\partial\mathcal{I}(Rome, no\_of\_cars(x) \doteq 1million) = \texttt{undef}$, since there is no entry for *cars* in the instance of *Rome*.

More precisely, for the following we assume a sorted (first-order) logic $\mathcal{L}$, where each sort can be viewed as a concept like *car* or *city*. We denote the sorts by lowercase Greek letters such as κ or μ. $\mathcal{E}$ is a set of sets $\{\mathcal{E}_\kappa\}_\kappa$, where each $\mathcal{E}_\kappa$ is called the *set of examples of sort* κ. The $\mathcal{E}_\kappa$ are such that their structure corresponds to the sort structure of $\mathcal{L}$, that is, if $\mu \sqsubseteq \kappa$ (i.e. μ is *subsort* of κ) then $\mathcal{E}_\mu \subseteq \mathcal{E}_\kappa$. $\mathcal{E}$ forms the *frame* (the collection of universes) for the partial interpretation of the terms. $\partial\mathcal{I}$ is a fixed partial interpretation function (corresponding to three-valued strong Kleene logic $\mathcal{L}^{\mathbf{K}}$ [16, 28]) in the frame $\mathcal{E}$. Each term t of sort κ is either interpreted by an example in $\mathcal{E}_\kappa$ or by the bottom element $\perp$. Formulae may be evaluated by $\partial\mathcal{I}$ to `true`, to `false`, or to `undef`. Furthermore, we assume that for every element $e^i_\kappa \in \mathcal{E}_\kappa$ there exists a constant c^i_κ of sort κ with $\partial\mathcal{I}(c^i_\kappa) = e^i_\kappa \in \mathcal{E}_\kappa$ and that there are only finitely many examples in $\mathcal{E}$.

The partial interpretation of composed formulae is defined as usual, based on the propositional connectives as defined by the following truth tables:

$\neg$	
`false`	`true`
`undef`	`undef`
`true`	`false`

$\vee$	`false`	`undef`	`true`
`false`	`false`	`undef`	`true`
`undef`	`undef`	`undef`	`true`
`true`	`true`	`true`	`true`

In order to fix the semantics of the universal quantifier, assignments ξ for the interpretation of the variables into the frame are necessary. If ξ is an arbitrary assignment, $\xi[x \leftarrow a]$ denotes the assignment equal to ξ for all variables except for x and $\xi[x \leftarrow a](x) = a$.

$$\partial\mathcal{I}_\xi(\forall x_\kappa \varphi) := \begin{cases} \texttt{true} & \text{if } \partial\mathcal{I}_{\xi[x\leftarrow a]}(\varphi) = \texttt{true} \text{ for all } a \in \mathcal{E}_\kappa \\ \texttt{false} & \text{if } \partial\mathcal{I}_{\xi[x\leftarrow a]}(\varphi) = \texttt{false} \text{ for one } a \in \mathcal{E}_\kappa \\ \texttt{undef} & \text{else} \end{cases}$$

The semantics of $\wedge$, $\Rightarrow$, $\Leftrightarrow$, and $\exists$ can then be defined in the usual manner.

Note that these definitions do not assume a concrete representation of the examples—the only requirement is that we get an answer to certain questions, thus fixing the interpretation function. In other words, $\partial\mathcal{I}$ has to be effectively computable for all ground formulae (i.e., variable free formulae) and, consequently, for all formulae, since we assume the number of examples in $\mathcal{E}$ to be finite.

As usual, we give an (extended) set theoretic semantics for a formula set. Our semantics is such that it is compatible with the examples. An interpretation of a knowledge base Δ is defined as an extension of the partial interpretation, given by $\langle\mathcal{E}, \partial\mathcal{I}\rangle$.

Definition ($\langle\mathcal{E}, \partial\mathcal{I}\rangle$-Interpretation): Let $\mathcal{E} = \{\mathcal{E}_\kappa\}_\kappa$ be a given set of example sets and let $\partial\mathcal{I}$ be a partial interpretation function in $\mathcal{E}$. An interpretation $\langle\{\mathcal{D}_\kappa\}_\kappa, \mathcal{I}\rangle$ is called an $\langle\mathcal{E}, \partial\mathcal{I}\rangle$*-interpretation* iff

– there are injective mappings $\mathsf{inj}_\kappa : \mathcal{E}_\kappa \hookrightarrow \mathcal{D}_\kappa$ with $\mathcal{I}(c_\kappa) = \mathsf{inj}_\kappa(\partial\mathcal{I}(c_\kappa))$ for all constant symbols c_κ with $\partial\mathcal{I}(c_\kappa) \neq \perp$. (When the sort is not important, we omit the index κ and simply write inj.)

– for all terms t and arbitrary ground instances $\sigma(t)$ with $\partial\mathcal{I}(\sigma(t)) \neq \bot$ holds $\mathcal{I}(\sigma(t)) = \mathsf{inj}(\partial\mathcal{I}(\sigma(t)))$.
– for all formulae φ and all ground instances σ of φ holds, if $\partial\mathcal{I}(\sigma(\varphi)) \neq \texttt{undef}$ then $\mathcal{I}(\sigma(\varphi)) = \partial\mathcal{I}(\sigma(\varphi))$

If $\mathcal{I}_\xi(\varphi) = \texttt{true}$ for all assignments ξ then $\mathcal{I}$ is called an $\langle\mathcal{E}, \partial\mathcal{I}\rangle$-*model* of φ. If φ has no $\langle\mathcal{E}, \partial\mathcal{I}\rangle$-model, it is said to be $\langle\mathcal{E}, \partial\mathcal{I}\rangle$-*unsatisfiable*. Γ $\langle\mathcal{E}, \partial\mathcal{I}\rangle$-*entails* the formula φ iff each $\langle\mathcal{E}, \partial\mathcal{I}\rangle$-model of Γ is an $\langle\mathcal{E}, \partial\mathcal{I}\rangle$-model of φ, too (i.e. $\Gamma \models_{\langle\mathcal{E},\partial\mathcal{I}\rangle} \varphi$).

So far all possible interpretations $\mathcal{I}$ have to be considered for modeling formulae. This is sufficient for capturing *deductive* reasoning as shown in [15]. However, in order to model any form of non-monotonic reasoning (in particular analogical reasoning), each conclusion drawn must potentially be withdrawn. In order to describe non-monotonicity semantically, Shoham [25] has introduced an order on the interpretations and weakened the notions of satisfiability, of consequence etc. His key idea is to only consider distinguished *minimal* models for the satisfiability and consequence relations. This approach is well-suited for analogical reasoning which is based on preferred instances—the typical instances of a concept.

Let us recall the notions of minimal model and preferential entailment in Shoham's approach:

DEFINITION: Assuming a partial order $<$ on interpretations, an interpretation $\mathcal{I}$ *preferentially satisfies* a formula φ (written as $\mathcal{I} \models_< \varphi$) if $\mathcal{I} \models \varphi$ and if there is no other interpretation $\mathcal{I}'$ such that $\mathcal{I}' < \mathcal{I}$ and $\mathcal{I}' \models \varphi$, that is, $\mathcal{I}$ is a minimal model of φ.

This definition attains a general semantic characterization of non-monotonic inferences. The only point to be specified is that of a partial order $<$ on the interpretation $\mathcal{I}$. Now we want to give a semantic characterization for the inference scheme **AT** of section 3. For the case of analogical reasoning with typical instances we define an interpretation to be smaller then another if it agrees in more individuals with the typical instances. Put formally,

DEFINITION: $\mathcal{I}' \leq \mathcal{I}$ if and only if[7,8]

[7] Nota bene: While the optimal models in Shoham's approach are *minimal*, the typical instances are *maximal* with respect to the typicality relation.

[8] In an alternative definition, the information is not taken from the typical instances, but from the instances immediately above the instance in consideration. In this case the definition looks as follows:
$\mathcal{I}' \leq \mathcal{I}$ if and only if

1. for all instances a with $\partial\mathcal{I}(a) = \bot$ and b is a least instance with $\partial\mathcal{I}(a) \sqsubseteq \partial\mathcal{I}(b)$ and $\partial\mathcal{I}(b) \neq \bot$ holds if $\mathcal{I}'(a) = \mathsf{inj}(\partial\mathcal{I}(b))$ then $\mathcal{I}(a) = \mathsf{inj}(\partial\mathcal{I}(b))$ and
2. for all predicates P with $\partial\mathcal{I}(P(a_1, \ldots, a_n)) = \texttt{undef}$, for all i, the b_i are least instances such that $\partial\mathcal{I}(a_i) \sqsubseteq \partial\mathcal{I}(b_i)$ and $\partial\mathcal{I}(P(b_1, \ldots, b_n)) \neq \texttt{undef}$ holds if $\mathcal{I}'(P(a_1, \ldots, a_n)) = \mathsf{inj}(\partial\mathcal{I}(P(b_1, \ldots, b_n)))$ then $\mathcal{I}(P(a_1, \ldots, a_n)) = \mathsf{inj}(\partial\mathcal{I}(P(b_1, \ldots, b_n)))$.

1. for all instances a with $\partial\mathcal{I}(a) = \bot$ and b is typical instance with $\partial\mathcal{I}(a) \sqsubseteq \partial\mathcal{I}(b)$ and $\partial\mathcal{I}(b) \neq \bot$ holds if $\mathcal{I}'(a) = \mathsf{inj}(\partial\mathcal{I}(b))$ then $\mathcal{I}(a) = \mathsf{inj}(\partial\mathcal{I}(b))$ and
2. for all predicates P with $\partial\mathcal{I}(P(a_1, \ldots, a_n)) = \mathtt{undef}$; for all i, the b_i are typical instances such that $\partial\mathcal{I}(a_i) \sqsubseteq \partial\mathcal{I}(b_i)$ and $\partial\mathcal{I}(P(b_1, \ldots, b_n)) \neq \mathtt{undef}$ and P is relevant for the b_i then holds:
   if $\mathcal{I}'(P(a_1, \ldots, a_n)) = \mathsf{inj}(\partial\mathcal{I}(P(b_1, \ldots, b_n)))$ then
   $\mathcal{I}(P(a_1, \ldots, a_n)) = \mathsf{inj}(\partial\mathcal{I}(P(b_1, \ldots, b_n)))$.
3. an analogous relation holds for functions f (replace $\mathtt{undef}$ by $\bot$.)

That is, in minimal models all *relevant* information that is not fixed by the knowledge base is transferred from typical instances.

Here is an example of how the semantics works: By definition, for a preferred model $\mathcal{I}$ with the typical instance *Berlin* and the instance *Rome* the information about *public_transportation* is transferred from *Berlin* to *Rome* if no additional information is known. Concretely, since $\partial\mathcal{I}(Rome, no\_of\_cars(x)) = \bot$, in a minimal model $\mathcal{I}^1_{\min}$ holds

$$\mathcal{I}^1_{\min}(Rome, no\_of\_cars(x)) = \mathsf{inj}(\partial\mathcal{I}(Berlin, no\_of\_cars(x))) = 1.3 \text{ Mio}$$

Of course this is only the case if $\mathcal{I}^1$ is a model of the first formula at all, that is, if there is no information that contradicts to the assumption that there are 1.3 Mio cars in *Rome*. It is easy to see that this semantics is non-monotonic: For instance, if the information that in *Rome* 2 Mio cars are one the roads, that is, $no\_of\_cars(Rome) \doteq 2$ Mio, is added to the knowledge base then $\mathcal{I}^1$ is no longer a model at all, and in particular no minimal model. Hence, the analogical conclusion cannot be inferred any longer.

The non-monotonicity can be concretely seen in the following form. Let $\varphi = no\_of\_cars(Rome) \doteq 1.3$ Mio and $\psi = no\_of\_cars(Rome) \doteq 2$ Mio, then we have:

$$\begin{array}{lll} \emptyset & \models^{\langle\mathcal{E},\partial\mathcal{I}\rangle}_{<} \varphi & \text{but} \\ \emptyset \cup \{\psi\} & \not\models^{\langle\mathcal{E},\partial\mathcal{I}\rangle}_{<} \varphi & \end{array}$$

4 Hybrid Framework

In order to *computationally* realize both kinds of analogy, a hybrid framework is needed, that provides propositional *and* exemplary knowledge representations and procedures to extract information from the non-propositional representation (see, e.g., [27, 22]). Our framework consists of three parts: a hybrid knowledge base, a reasoner, and procedures which deliver information from the knowledge base to the reasoner [14]. The knowledge base itself has two parts: a collection of propositions and non-propositional representations of concepts. The reasoner

3. an analogous relation holds for functions f (replace $\mathtt{undef}$ by $\bot$.)

consists of inference methods that operate on the propositional part of the knowledge base and of methods that use information contained in the conceptual part of the knowledge base. Figure 2 displays this framework.

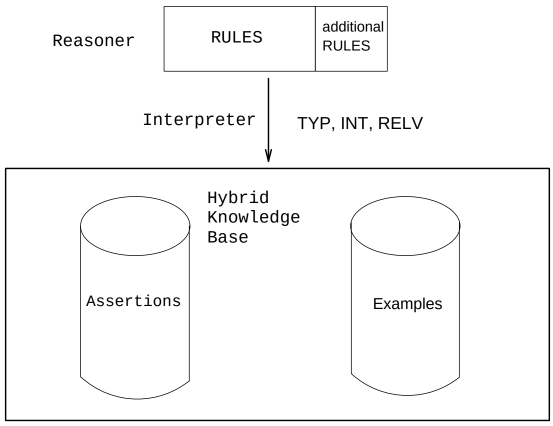

Fig. 2. System structure

The Propositional Subsystem

The propositional subsystem consists, as usual, of a set Γ of (sorted first order) formulae. Aspects, as mentioned above, can be defined in this subsystem. The propositionally represented connections of aspects—as far as they are available—belong to this subsystem as well.

The Conceptual Subsystem

We extend the knowledge base by a *conceptual* part consisting of *concept structures* which are non-propositional representations of concepts. We use directed acyclic graphs consisting of a set of instances and the typicality relation $\sqsubseteq_C$ as concept structures.

The instances themselves might be represented by neural nets, maps, diagrams or some other means including symbolic representations. The particular type of these representations is of no concern for the rest of the paper. We consider concepts C to be represented by a set of instances with a partial order $\sqsubseteq_c$. The elements of these concept structures represent concept instances. A concept structure is displayed as a directed acyclic graph as in figure 1. In this city example, the instances are not directly represented as maps, but we deal with another concept representation which is similar to that employed by Barwise and Etchemendy [3]. It encodes instances as tables. Nevertheless, the conceptual part of the knowledge base is non-propositional because of the concept structure.

The Reasoner

In order to integrate the conceptual subsystem into the framework, its informational content has to be accessible. The semantics corresponds exactly to the semantics given in section 3. Several inspection procedures work on the conceptual subsystem and provide the information that is needed by the rules of the reasoner:

- A TYP-procedure provides access to the structural content of the concept structures in that it finds a typical instance s with $t \sqsubseteq_C s$ for an instance t of a concept C. For example, TYP yields for the instance *Rome* the typical instance *Berlin* by looking up the concept structure *city*.[9]
- An interpreting partial procedure, called ASP, computes the values of aspects A for the *typical* instances c out of the representation of the instances. If there is a value it is required to be a formula $A\langle c\rangle$ with $\mathcal{I}(A\langle c\rangle) = \texttt{true}$. For example, ASP yields a value of the aspect *public_transportation* for the typical instance *Berlin*: $subway(x) \wedge bus(x) \wedge taxi(x) \wedge airport(x)$ by looking up the representation of *Berlin*.
- A RELV-procedure provides `true`/`false`-information about the relevancy of aspects for the typical instances that is encoded in the concept representation. An aspect A is *relevant* for an instance c iff $\mathcal{I}(A\langle c\rangle[x/c]) = \texttt{true}$ and for all instances $c' \sqsubseteq_C c$ holds $\mathcal{I}(A\langle c\rangle[x/c']) \in \{\texttt{true}, \texttt{undefined}\}$. This notion of relevancy corresponds to a kind of modal operator. We assume that every typical examples has at least one relevant aspect.

AR which finally provides information about a target case t takes as input:

- The similarity of a source case s and the target case t expressed by the equivalence w.r.t. an aspect A_1. For example, let be $s = Madrid$, $t = Rome$, and $A_1 = population$. The input is $population\langle Rome\rangle = population\langle Madrid\rangle$.
- A connection $[A_1, A_2]$ which belongs to the propositional part of the knowledge base. Such a connection is $[population, cars]$ which means that, if the populations of two cities agree, then probably the number of cars registered in these cities agrees too.

AR yields $A_2\langle t\rangle = A_2\langle s\rangle$. For the example input above: $cars\langle Rome\rangle = cars\langle Madrid\rangle$. Using the additional information about the actual value of $A_2\langle s\rangle$, which is explicitly given by the propositional subsystem, we can infer the value of $A_2\langle t\rangle$.

At a glance: Let the following information be given

[9] This example shows that typical examples are not independent of the culture background. For Italians in general, *Rome* would be the typical example and not *Berlin*.

$[population, cars]$	
$population\langle Madrid\rangle =$ $no\_of\_inhabitants(x, 3million)$	$population\langle Rome\rangle =$ $no\_of\_inhabitants(x, 3million)$
$cars\langle Madrid\rangle =$ $no\_of\_cars(x, 1million)$	$cars\langle Rome\rangle =$?

AR infers by analogy $cars\langle Rome\rangle = cars\langle Madrid\rangle$.
Hence it is possible to infer $cars\langle Rome\rangle = no\_of\_cars(x, 1million)$

The **AT** rule of the reasoner takes as inputs:

- a typical instance s with $t \sqsubseteq_C s$ which is computed by the TYP-procedure, and
- information about the relevancy of A_2 for this s. This information is either extracted by the RELV-procedure or explicitly represented in the propositional part of the knowledge base as, e.g., suggested by Gentner [10].

The **AT** rule then infers $A_2\langle t\rangle = A_2\langle s\rangle$. Using the additional information about the actual value of $A_2\langle s\rangle$, which is computed by the ASP-procedure, $A_2\langle t\rangle$ can be inferred.

Let us look at an example where the individual concept structure of the concept *city* is given. We want to compute $A_2\langle t\rangle = public\_transportation\langle Rome\rangle$ by analogy to a typical city. If there is no explicit connection for *public_transportation*, then **AR** cannot be applied and we proceed as follows:

- The TYP-procedure computes the typical instance *Berlin* as a typical city that is rated over *Rome*.
- RELV tests whether *public_transportation* is a relevant aspect for *Berlin*. Provided that the result is `true`, **AT** yields $public\_transportation\langle Berlin\rangle = public\_transportation\langle Rome\rangle$.
- With the additional information $public\_transportation\langle Berlin\rangle = (subway(x) \wedge bus(x) \wedge taxi(x) \wedge airport(x))$ which is provided by the ASP-procedure or explicitly given in the propositional subsystem, we infer $public\_transportation\langle Rome\rangle = (subway(x) \wedge bus(x) \wedge taxi(x) \wedge airport(x))$ by analogy.

At a glance: let the following information be given

$relevant(public\_transportation, Berlin)$ typex($Berlin$) and $Rome \sqsubseteq_{\mathrm{city}} Berlin$.	
$public\_transportation\langle Berlin\rangle =$ $(subway(x) \wedge bus(x) \wedge$ $taxi(x) \wedge airport(x))$	$public\_transportation\langle Rome\rangle =$?

AT infers by analogy

$$public\_transportation\langle Berlin\rangle = public\_transportation\langle Rome\rangle.$$

Hence it is possible to infer

$$public\_transportation\langle Rome\rangle = (subway(x) \wedge bus(x) \wedge taxi(x) \wedge airport(x)).$$

5 Conclusion

In this paper we discussed two different kinds of analogical reasoning, connection-based analogical reasoning and analogical reasoning based on typical instances. The non-monotonicity of the connection-based analogical inferences stems from the justifying connections that represent the relevance of an aspect for another aspect. On the other hand, the non-monotonicity of the analogical inference based on typical instances is caused by the knowledge about the relevance of an aspect for a typical instance of a concept. For the latter we presented an appropriate semantics that is a special case of Shoham's minimal model semantics.

The paper dealt with two different analogical inference schemas and with their computational realization in a hybrid framework that is equipped with different kinds of knowledge representation. Related work with typical instances has been done for the machine learning systems PROTOS and COBWEB [2, 8] that also use a representations of concepts with typical instances.

References

1. J. Arima and K. Sato. Non-monotonic reasoning. In K. Furukawa and F. Mizoguch, editors, *Knowledge Programming*, pages 189–214. Kyritso Shuppan Co. Ltd., 1988. in Japanese.
2. E.R. Bareiss, B.W. Porter, and C.C. Wier. Protos: An exemplar-based learning apprentice. *International Journal of Man-Machine Studies*, 29:549–561, 1988.
3. J. Barwise and J. Etchemendy. *Hyperproof.* CSLI Lecture Notes. University of Chicago press, Chicago, 1993.
4. R.J. Brachman and J.G. Schmolze. An overview of the KL-ONE knowledge representation system. *Cognitive Science*, 9:171–216, 1985.
5. G. Brewka. Belief revision in a framework for default reasoning. In Gerhard Brewka and Ulrich Junker, editors, *Aspects of Non-Monotonic Reasoning*, pages 2–18. GMD, St. Augustin, Germany, TASSO Report No.1, March 1990.
6. A.M. Collins and R. Michalski. The logic of plausible reasoning: A core theory. *Cognitive Science*, 13(1):1–50, 1989.
7. T.R. Davis and S.J. Russell. A logical approach to reasoning by analogy. In *Proceedings of the Tenth International Joint Conference on Artificial Intelligence*, pages 264–270, Milan Italy, 1987. Morgan Kaufmann.
8. D.H. Fisher. Conceptual clustering, learning from examples, and inference. In *Proceedings of the 4th International Machine Learning Workshop*, pages 38–49, 1987.
9. D. Gentner. Structure mapping: A theoretical framework for analogy. *Cognitive Science*, 7(2):155–170, 1983.
10. D. Gentner. The mechanisms of analogical learning. In S. Vosniadou and A. Ortony, editors, *Similarity and Analogical Reasoning*, pages 199–241. Cambridge University Press, 1989.

11. M.L. Gick and K.J. Holyoak. Schema induction and analogical transfer. *Cognitive Psychology*, 15(1):1–38, 1983.
12. R. Goebel. A sketch of analogy as reasoning with equality hypotheses. In K.P. Jantke, editor, *Analogical and Inductive Inference*, volume 397 of *Lecture Notes on Computer Science*, pages 243–253. Springer, 1989.
13. R. Greiner. Learning by understanding analogies. *Artificial Intelligence*, 35:81–125, 1988.
14. M. Kerber, E. Melis, and J. Siekmann. Analogical reasoning based on typical instances. in Proceedings of the IJCAI-Workshop on Principles of hybrid reasoning and representation, July 1993. Chambéry, France.
15. M. Kerber, E. Melis, and J. Siekmann. Reasoning with assertions *and* examples. AAAI Spring Symposium on AI and Creativity, März 1993. Stanford, California, USA.
16. S. C. Kleene. *Introduction to Metamathematics.* Van Nostrand, Amsterdam, The Netherlands, 1952.
17. G. Lakoff. *Women, Fire and Dangerous Things.* The University of Chicago Press, London, 1987.
18. J. McCarthy. Programs with common sense. In Marvin Minsky, editor, *Semantic Information Processing*, pages 403–418. MIT Press, Cambridge, Massachusetts, 1968.
19. E. Melis. Study of modes of analogical reasoning. TASSO-report 5, Gesellschaft für Mathematik und Datenverarbeitung, Birlinghoven, Germany, 1990.
20. C. Mervis and E. Rosch. Categorization of Natural Objects. *Annual Review of Psychology*, 32:89–115, 1981.
21. C.B. Mervis and E. Rosch. Family resemblance: Studies in the internal structure of categories. *Cognitive Psychology*, 7:573–605, 1975.
22. K.L. Myers and K. Konolige. Reasoning with analogical representations. In *Proceedings of the Third International Conference on Knowledge Representation KR'92*, pages 189–200, 1992.
23. L. Rips. Inductive judgements about natural categories. *Journal of Verbal Learning and Verbal Behavior*, 14:665–681, 1975.
24. S.J. Russell. *The Use of Knowledge in Analogy and Induction.* Pitman, London, 1989.
25. Y. Shoham. *Reasoning about Change - Time and Causation from the Standpoint of Artificial Intelligence.* MIT Press, Cambridge, Massachusetts, 1988.
26. A. Sloman. Interactions between philosophy and artificial intelligence: The role of intuition and non-logical reasoning in intelligence. *Artificial Intelligence*, 2:209–225, 1971.
27. A. Sloman. Afterthoughts on analogical representation. In *Proceedings of Theoretical Issues in Natural Language Processing*, 1975.
28. A. Urquhart. Many-valued logic. In D. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic*, chapter III.2, pages 71–116. D.Reidel Publishing Company, Dordrecht, The Netherlands, 1986. Volume III: Alternatives to Classical Logic.
29. E.J. Weiner. A knowledge representation approach to understanding metaphors. *Computational Linguistics*, 10(1):1–14, 1984.
30. P. Winston. Learning by creating and justifying transfer frames. *Artificial Intelligence*, 10(2):147–172, 1978.