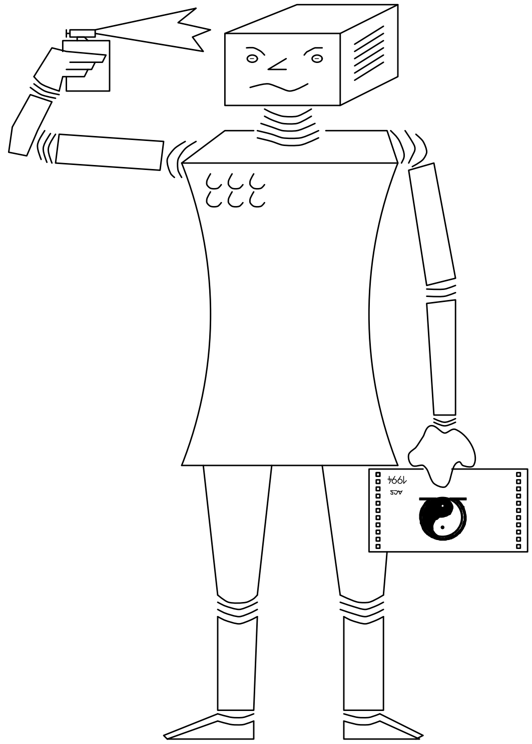
SEKI Report



Analogical Reasoning with Typical Examples

Manfred Kerber, Erica Melis, and Jörg Siekmann

SEKI Report SR-92-13

Analogical Reasoning with *Typical Examples*

Manfred Kerber Erica Melis
Fachbereich Informatik
Universität des Saarlandes
6600 Saarbrücken, Germany
{kerber|melis}@cs.uni-sb.de

Jörg H. Siekmann
DFKI
German Research Center for AI
6600 Saarbrücken, Germany
siekmann@dfki.uni-sb.de

Begriffe ohne Anschauungen sind leer
Anschauungen ohne Begriffe blind.

Immanuel Kant

Abstract

Typical examples, that is, examples that are representative for a particular situation or concept, play an important rôle in human knowledge representation and reasoning. In real life situations more often than not, instead of a lengthy abstract characterization, a *typical example* is used to describe the situation. This well-known observation has been the motivation for various investigations in experimental psychology, which also motivate our formal characterization of typical examples, based on a partial order for their typicality. Reasoning by typical examples is then developed as a special case of analogical reasoning using the semantic information contained in the corresponding concept structures. We derive new inference rules by replacing the explicit information about connections and similarity, which are normally used to formalize analogical inference rules, by information about the relationship to typical examples. Using these inference rules analogical reasoning proceeds by checking a related typical example, this is a form of reasoning based on semantic information from cases.

Keywords: typical examples, analogy, case-based reasoning, concept representation

Contents

1 Introduction

Human reasoning is to an extraordinary extent based on reasoning by examples and cases [21]. This is in sharp contrast to most systems of artificial intelligence (AI) that rely essentially on one kind of mechanism – namely rules combined with static facts. For these systems deductive calculi are an accepted theoretical basis. Although these systems are often adequate to solve the problems in a special domain and in particular they allow for efficient inference procedures for certain tasks, it is far from clear how these systems can be extended to the kind of reasoning patterns that rely on real world experience, such as reasoning by analogy, case-based reasoning or reasoning from typical examples as proposed in this paper.

A well-known criticism of classical AI by Dreyfus and Dreyfus [6] questions this rule-based approach: they claim that especially in the highest stage of knowledge processing (the expert level) examples are imperative. Of course we cannot remember all examples we have ever encountered, but instead we extract a few *typical* examples that are easy to remember, and that are sufficient to catch important aspects of the general case. Although the fields of analogical as well as case-based reasoning are both concerned with reasoning based on examples, surprisingly the importance of reasoning by *typical* examples, as for instance, investigated by Rosch [21] and Lakoff [19] has not been elaborated. However, an attempt of Winston [35] was based on statistic prototypicality. In this paper we derive analogical reasoning patterns based on typical examples. The idea behind our approach is to replace some symbolic rule-based information by explicit semantic knowledge that we assume to be directly represented in the memory (in a computer store) by neural nets, a case base or some other means. This is in line with current suggestions of e.g. Halpern and Vardi [12] as well as of Johnson-Laird and Byrne [13] elaborating model oriented reasoning patterns.

2 Typical Examples of *Typical Examples*

In the following we shall present three cases where reasoning by typical examples is an essential means to draw inferences.

2.1 Typical Birds Can Fly

In the context of non-monotonic logic the bird Tweety, and its disputed ability to fly, is a paradigmatic example[i].

Consider the following information: Typical birds can fly and Birdy is a bird, i.e. the set $\{[\forall x \ \mathtt{typex}_{\mathtt{bird}}(\mathtt{x}) \Longrightarrow \mathtt{can\_fly}(\mathtt{x})], \mathtt{bird}(\mathtt{Birdy})\}$, where $\mathtt{typex}_{\mathtt{y}}(\mathtt{x})$ represents the fact that $\mathtt{x}$ is a typical $\mathtt{y}$. If there is no additional information we want to conclude (tentatively) that

[i] Paradigmatic examples as a part of a scientific paradigm are introduced by Kuhn [18]. They are typical examples in a specific scientific context.

Birdy *can* fly, that is, `can_fly(Birdy)`. However if we have some knowledge that entails that Birdy cannot fly (for example because it has a broken wing) we want to conclude (definitively) that Birdy *cannot* fly, that is, ¬`can_fly(Birdy)`.

Reasoning by typical examples, we would describe the situation as follows: Suppose Birdy is a bird and also there is Tweety, which is a *typical* bird. Assume also that in addition to the object language (say first-order predicate logic) there is some means for the representation of a *typical bird*. This could be done in a semantic net or some frame structure or it could be represented in a neural net: the only formal requirement we have for this model representation – of which we assume that it is distinct and different form our (logical) object language – is that we obtain answers for certain questions. For example, the query `can_fly(Tweety)` should evaluate to `true`, if the representation of the typical bird Tweety contains the information that in fact Tweety can fly. In other words we have two different levels of information: the syntactic information, `bird(Birdy)`, which is stated in the object language, where (deductive) inferences are drawn and the semantic information level, where for example a typical case of a bird, namely `Tweety`, is represented. If `can_fly(Tweety)` evaluates to true, we tentatively conclude that Birdy can fly, too. Apparently, this procedure is similar to the way humans reason under these circumstances: The default knowledge is stored in the form of an example [23], from which the conclusion is drawn *by analogy* rather than by an explicit rule of deductive inference – be it monotonic or not. If we want to know something about an arbitrary bird, for instance, if it has teeth, we have no rule in mind like: "typical birds have no teeth" (and the myriad of other facts that are not the case), but we think of a typical representative for the concept bird and reason by analogy: Tweety has no teeth, hence Birdy has no teeth.

In the following we shall discuss examples from mathematics.

2.2 What is a Group?

When a teacher introduces a new concept in a math lesson, normally, she would give a definition, show some important properties, and above all she would give several well-chosen examples for this concept. Generally a student "understands" the concept only when he knows *all these facets* of the concept and not just the definition augmented by some properties [15]. The examples a good teacher presents to her students are well-chosen and typical: For instance, the concept of a group may be introduced by well-known examples like $(\mathbb{Z}, +)$ rather than the multiplication group of non-zero quaternions. Scanning some text books on algebra[i] we found in *all* text books, the integers $(\mathbb{Z}, +)$ and in all but one the symmetric group $(S(M), \circ)$ as examples for a group.[ii]

[i] Fischer, Godement, Kowalsky, Lang, Lipschutz, Lüneburg, Oeljeklaus–Remmert, van der Waerden

[ii] Similarly, when you are asked to quickly enumerate some tools of a craftsman, almost certainly "hammer" will be among the first examples. Similarly a violin is a typical musical instrument. In this psychological sense $(\mathbb{Z}, +)$ and "hammer" are typical examples.

In addition to these findings there is empirical evidence in the context of learning that for tasks that are usually considered purely deductive, people often use, although formally incorrect, a form of reasoning that exploits the typicality structure of their mental representation of concepts [3].

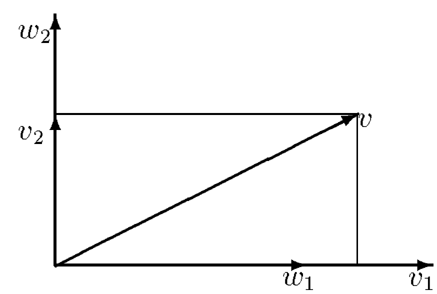
In the field of mathematical problem solving Cherniak [3] has found a deductive reasoning heuristic corresponding to Hadamard's observations [11] and Pólya's recommendation [24]: When trying to discover a proof of a theorem, it is often indispensable to start the work with a typical example that satisfies the initial conditions. Cherniak furthermore states that the more complex the deductive task, the more likely subjects will decide to use the typicality heuristic. We experienced that the typicality heuristics works in theorem proving too: proofs for typical problems can guide the search for proofs of more general problems (for example by analogy [14, 24] or via a proof plan [2]).

2.3 How to Prove a Theorem

In 2.1 and 2.2 typical examples of familiar concepts have been considered within the context of learning and common sense reasoning. Another important field where typical examples play an important rôle is in proving mathematical theorems. Just as concepts can be defined with the help of typical examples, some proofs can be found with the help of proofs for typical cases.

For instance, $\mathbb{R}^2$ is a typical example for a finite-dimensional Euclidean vector space with scalar product. The proof of the orthonormalization theorem[i] can be obtained from the proof for $\mathbb{R}^2$. The proof in the case of $\mathbb{R}^2$ is as follows:

Select a vector $v_1 \neq 0$ from $\mathbb{R}^2$ and define $w_1 := ||v_1||^{-1} \cdot v_1$. w_1 is an orthonormal basis of $\mathbb{R}^1$, which is a subspace of $\mathbb{R}^2$. Since $\mathbb{R}^2$ is two-dimensional, there is a vector v in $\mathbb{R}^2$ which is not in the one-dimensional space generated by w_1. We define v_2 to be the difference of v and the orthogonal projection of v on w_1, that is, $v_2 := v - \langle w_1, v\rangle \cdot w_1$ and $w_2 := ||v_2||^{-1} \cdot v_2$. $\{w_1, w_2\}$ is an orthonormal base of $\mathbb{R}^2$.



The induction step of the proof of the orthonormalization theorem for V proceeds analogously to the proof step for $\mathbb{R}^2$:

Let $\{w_1, \ldots, w_{n-1}\}$ be the orthonormal base of an $(n-1)$- dimensional subspace of a finitely dimensional vector space V with $\dim(V) > n-1$. Since V has a dimension greater

[i]Theorem: Let V be a finite-dimensional Euclidean vector space and $W \subset V$ a sub-vector space, then every orthonormal base $(w_1, \ldots, w_m)$ of W can be completed to a orthonormal base $(w_1, \ldots, w_m, w_{m+1}, \ldots, w_n)$ of V.

than $n-1$, there is a vector $v \in V$ which is not in the $(n-1)$-dimensional space generated by $\{w_1, \ldots, w_{n-1}\}$.

As in the case of $\mathbb{R}^2$, we define v_n as the difference of v and the sum of the orthogonal projections of v on w_i, that is, $v - \sum_{i=1}^{n-1} \langle w_i, v \rangle \cdot w_i$, and w_n as $||v_n||^{-1} \cdot v_n$. Then $\{w_1, \ldots, w_n\}$ is an orthogonal base of V.

Thus, proofs for typical cases can guide the search for proofs of arbitrary models.

3 Psychological Findings

The typicality of examples is a well-known research area in empirical psychology [19, 22, 25]. Rosch and Mervis [21] argue in their prototype theory of mental representations for the importance of representative instances of a concept. In particular they evaluate the membership problem of a concept in terms of the resemblance of an instance to a typical example.

The existence of a reproducible typicality rating that distinguishes typical examples from those that are are not is one of the main results, which has been interpreted and explained by a special memory organization that prefers typical examples [21]. Some of the important experimental methods for the extraction of typical examples are [19]:

- direct rating: Subjects are asked to rate (say on a scale from one to seven) how *representative* a given example (e.g., a robin or a chicken) is for a given category (e.g. bird). Typical examples are rated best, whereas atypical examples are rated least.
- reaction time: Subjects are asked to press a button to indicate true or false in response to a statement of the form "An [example] is a [category name]" (e.g., "A chicken is a bird"). Response times are shorter for typical examples.
- reproduction of examples: When asked to list or draw examples of category members, subjects are more likely to list or draw typical examples.
- asymmetry in similarity ratings: Less typical examples are often considered to be more similar to a typical example than the converse.
- generalization: Humans generalize more likely from typical examples than from arbitrary examples.

There is empirical evidence that all these methods extract similar ratings (for a discussion of the results see [19] and [26]). Generally, this rating is not identical with the set inclusion of properties [27]. Rosch and Mervis [21] investigated the concepts "furniture", "vehicle", "fruit", "weapon", "vegetable", and "clothing", and they found typicality ratings for each of them. While most investigations supporting the prototype theory have been made on natural concepts that are often not precisely definable, Armstrong et al. [1] carried out the same experiments as Rosch [27] for uniquely defined concepts like "odd number" and "grandmother". They found a typicality structure for these concepts too.

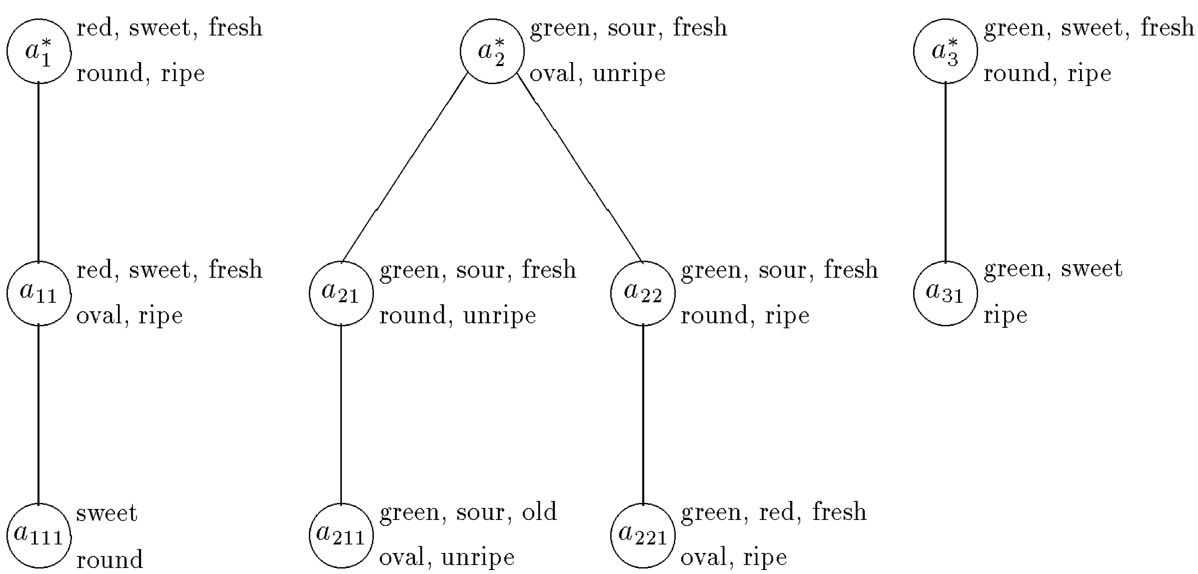
These psychological findings are a motivation and a legitimation for our formal approach, as all of these results suggest the existence of the notion of typicality. We shall represent typicality as a partial order $\sqsubset$ on the set of instances of a concept C and in the following we assume $\sqsubset$ to be a given partial order induced by the typicality rating. We take this partial order $\sqsubset$ on a concept C as given – as difficult as it may be, to determine it in a specific setting.

We fix the following notation (a more formal definition is given in section 4 below):
A *typical example of a concept* C is an example that is particularly representative for C, that is, it rates best under the assumed partial order of a typicality rating.

The following case is used as a point of reference for typical examples throughout this paper. Consider the concept of an "apple" which consists of the set of elements

$$\mathcal{E}_{apple} := \{a_1^*, a_{11}, a_{111}, a_2^*, a_{21}, a_{211}, a_{22}, a_{221}, a_3^*, a_{31}\}$$

each of which is a representation of a specific apple and hence an instance of the general notion of an apple. Apples have the properties colour, taste, age, shape, and ripeness. Every specific apple is characterized by certain property values, where knowledge may be incomplete as in the case of apple a_{111}, where the colour is not known. Correspondingly, we shall consider only partial interpretations in our formal treatment in section 4. The figure below displays each apple by a circle, inside of which is its name, e.g. a_{11}. Along with each apple we list its properties, e.g. red, sweet, fresh, oval, ripe in the case of apple a_{11}. The edges indicate the typicality rating, that is, a_{11} is less typical than apple a_1^*. Typical examples are marked by an asterisk.



Generally speaking, examples may be represented in different ways. For instance, Tweety could be represented as a neural net, the typical example for a group, i.e. the integers $(\mathbb{Z}, +)$, could be given as the `integer` data-structure in a programming language and so on.

The apple examples are represented as lists of properties, as this makes it easy to communicate these properties in a paper (such as this). Our formalism however is not committed to a particular means of representation for the typical examples (sub-symbolic by neural nets or such like or symbolic by frames, semantic nets, a programming language, a data base etc.), as long as this representation obeys certain very general properties. One important requirement is the decidability of whether or not an example is a typical example for a concept. Another one is the existence of a partial interpretation of formulae in the examples.

4 Formal Treatment

In this section we shall develop the first steps toward a hybrid computational approach of reasoning by typical examples. Our starting point is the notion of an *agent*.

Definition (Agent): An *agent* **A** is a pair consisting of a knowledge base **KB** and a reasoner, by which knowledge can be inferred from **KB**. A reasoner is a set of inference rules along with some means of their control.

These rules of inference may be classified as deductive, inductive, and abductive (as proposed by Ch. S. Peirce). In addition there is a special rule of inference, called *analogical reasoning by typical examples*, as defined in the next section. The knowledge base $\mathbf{KB} = (\Gamma, \mathcal{E}, \mathcal{I})$ itself is a triple consisting of a set of logical formulae Γ, a set of partially ordered sets of examples $\mathcal{E}_\kappa$, each of which with at least one element (non-emptiness), and a partial interpretation function $\mathcal{I}$. The elements $e \in \mathcal{E}_\kappa$, are called *examples* and the interpretation function $\mathcal{I}$ and these examples are connected in the following way: $\mathcal{I}$ is defined such that for some constants c_i in Γ, namely for the names of examples in $\mathcal{E}_\kappa$, we have $\mathcal{I}(c_i) = e_i$, where c_i is a constant and e_i the corresponding explicit representation of the example e_i in $\mathcal{E}_\kappa$. Since we consider only one agent at a time for the purpose of this paper we fix $\mathcal{I}$ without subscript.

The set of logical formulae Γ is stated in a sorted first-order logic $\mathcal{L}$, although the basic approach of this paper would remain valid for any arbitrary logic with model theoretic semantics.

Typical examples occur in many disguises: typical examples of concepts (such as a violin, a hammer or an apple above), as typical examples of situations (such as Schank's script of eating in a restaurant, opening a door), as typical examples of a proof for a theorem or as a typical example for a plan. Since most psychological investigations on mental representations have been restricted to concepts, we confine our treatment to concepts too (as a starting point in this paper), and represent them as *sorts* κ_i. For instance, a sorted formula is:
$\forall x : apple \;\; red(x) \Longrightarrow sweet(x)$.

For sorted logics compare e.g. [31, 29]. We do not consider typical examples for any arbitrary formula, but only for fixed sorts in our sorted logic. As usual constants and variables of sort κ are interpreted in a universe $\mathcal{D}_\kappa$ such that $\mathcal{E}_\kappa$ is a non-empty subset of $\mathcal{D}_\kappa$. For instance, constants of sort *apple* are interpreted in a universe $\mathcal{D}_{apple}$ of all apples. A subset $\mathcal{E}_{apple}$ of

$\mathcal{D}_{apple}$ is represented in the knowledge base of **A**, it contains all examples for apples, which are remembered by **A**. On each $\mathcal{E}_\kappa$ we assume a given partial order $\sqsubset_\kappa$, which is defined by the typicality rating. The interpretation function $\mathcal{I}$ is a partial function from constant symbols, function symbols (e.g. of sort $\kappa_1 \times \cdots \times \kappa_n \rightarrow \kappa$) and predicate symbols to the corresponding elements, functions (e.g. from $\mathcal{E}_{\kappa_1} \times \cdots \times \mathcal{E}_{\kappa_n}$ to $\mathcal{E}_\kappa$) and relations.

Definition (Concept): A pair $(\mathcal{E}_\kappa, \sqsubset_\kappa)$, where $\mathcal{E}_\kappa$ is a set of examples and $\sqsubset_\kappa$ is a partial order on $\mathcal{E}_\kappa$ (the typicality rating) is called a *concept* and denoted by $\mathcal{C}_\kappa$.

If there is no danger of confusion we write $\sqsubset$ instead of $\sqsubset_\kappa$ and $\mathcal{C}$ instead of $\mathcal{C}_\kappa$.

An example $e \in \mathcal{E}_\kappa$ of sort κ is called an *exception* if it is not comparable to any other example of $\mathcal{E}_\kappa$ (that is, if there is no $e' \in \mathcal{E}_\kappa$ with $e' \sqsubset e$ or $e \sqsubset e'$).

Definition (Typical Example): An example $e \in \mathcal{E}_\kappa$ of sort κ is called a *typical example* for κ if it is maximal (that is, there is no e' which $e \sqsubset e'$) and if it is not an exception. We write $\text{typex}_\kappa(e)$ or $\text{typex}(e)$, iff e is a typical example.

We assume that the interpretation $\mathcal{I}$ is extended to arbitrary formulae as usual and that the assignment $\xi[x \leftarrow e]$ is equal to the assignment ξ except for all free occurrences of x, where x is mapped to e. Since $\mathcal{I}$ is a partial function the interpretation of a formula may evaluate to `true`, to `false`, or to `undefined`.

The following definition captures the connection between the formulae in Γ and the examples in $\mathcal{E}$.

Definition (Example for a Formula): Let H be a formula in $\mathcal{L}$ with at most one free variable x_κ of sort κ. An example $e \in \mathcal{E}_\kappa$ of sort κ is called *an example* for H if $\mathcal{I}_{\xi[x \leftarrow e]}(H) =$ `true`. We say H is valid in e, $e \mid\!\sim H$ and define $e \mid\!\sim H := \mathcal{I}_{\xi[x \leftarrow e]}(H)$.

Definition (Relevant Validity): A formula $H \in \mathcal{L}$ with at most one free variable x of sort κ is said to be *relevantly valid in an example* $e \in \mathcal{E}_\kappa$ iff $e \mid\!\sim H$ and for all $e' \in \mathcal{E}_\kappa$ with $e' \sqsubset e$, $e' \mid\!\sim H$ is `true` or `undefined`. We write $e \mid\!\sim_{\mathbf{r}} H$.

Remark: If we want to know for a formula H and an example e whether $e \mid\!\sim_{\mathbf{r}} H$ holds, it is a priori necessary to scan the whole data base of examples. This of course is not the intention of reasoning with typical examples: The very idea is to check only a few typical examples and to avoid checking all the others. Hence, for efficiency reasons, relevancy should be explicitly integrated into the representation of a typical example. An alternative to the definition above is to introduce a higher-order relation "relevant" as done by Gentner [7], [8] or Winston [36].

Example: In the apple example we have: $\text{typex}(a_1^*)$, but not $\text{typex}(a_{11})$, $a_1^* \mid\!\sim_{\mathbf{r}} \text{sweet}(x)$, $a_1^* \mid\!\sim_{\mathbf{r}} \text{red}(x)$, but $a_1^* \not\mid\!\sim_{\mathbf{r}} \text{round}(x)$. Relevancy of features is important for reasoning by examples, in particular, for analogical reasoning. For instance, it is not reasonable to analogically infer that if an apple x is round then it is sweet, but it is reasonable (for the above apple examples) to infer that if x is red then it is sweet.

5 Analogical Reasoning with Typical Examples

In this section we assume a fixed agent and we shall discuss different reasoning patterns by analogy for this agent. These analogical reasoning patterns are then modified in order to incorporate reasoning by typical examples.

5.1 A General Inference Rule for Analogical Reasoning

Analogical reasoning is a form of reasoning that is heavily based on some background knowledge [28, 33, 10]. In the following, we assume the reader to be familiar with the mainstream work on analogical reasoning and just recall some basic notions; for a detailed discussion see for example [20, 28].

Generally speaking, analogical reasoning transfers of a property from a known *base example* to a partially unknown *target example*, if the base and the target are similar with respect to some other property. Later on we shall call these properties that establish the similarity between base and target *aspects*. Analogical reasoning is based on knowledge about a causal or another connection between properties (aspects) of examples. Such a connection is for instance that, "if two cars are built in the same year and have the same make, then their price will probably be the same." The two cars do not necessarily have to correspond in other properties (e.g. colour). These connections are part of an agent's experience and belong to its knowledge base, as argued in [28]. Connections can be represented as ordered pairs of properties. For a more general treatment of analogical reasoning, the notion "properties" has to be replaced by the more general notion of an "aspect" including for instance causes and effects. Hence, in our approach, a connection is represented as the pair $[A_1, A_2]$ of aspects. Our notion of connection subsumes the determinations [5], causal rules [4, 16], schemata [9], and abstractions [10].

Example: Let S denote the aspect "structure". Let h be a certain example of the concept "house". Then $S(h)$ is an $\mathcal{L}$-formula describing the structure of the house h. Let c be an example of the concept "car". Then $S(c)$ may be a totally different formula describing the structure of the car c.

The experience "the structure of an example determines the function of the example" is encoded in a *connection* [structure, function]. Often this connection is not given explicitly for all examples. But, from the connection [structure, function] one can *infer analogically* that if two examples b and t are similar with respect to their structure then they may correspond with respect to their function too.

Definition (Aspect): An *aspect* A is a partial function mapping the examples e of C_κ to non-tautological $\mathcal{L}$-formulae H with at most one free variable of sort κ, such that $e \mid\!\sim H$. If e is an example and A is an aspect then $A(e)$ is the $\mathcal{L}$-formula describing A in e.

Example: For the aspect "colour" in our apple example we have colour(a_2^*)=green(x) and colour(a_{221})=(green(x)$\wedge$ red(x)).

Definition (Analogical Inference): Let the *base* b be a (sufficiently) known example and the *target* t be an at least partially unknown example. Let A_1, A_2 be aspects.

An inference rule that is based on the similarity of b and t, with respect to an aspect A_1 (i.e., on $A_1(b) \Leftrightarrow A_1(t)$), and on a connection $[A_1, A_2]$, that results in a correspondence of $A_2(b)$ and $A_2(t)$ is said to be an *inference by analogy*. The base b is called an analogue of the target t with respect to the connection $[A_1, A_2]$.
The inference rule is expressed formally as:

$$(\mathbf{AN}) \qquad \frac{A_1(t) \Leftrightarrow A_1(b), [A_1, A_2]}{A_2(t) := A_2(b)}$$

That means, if the values of the aspect A_1 are equivalent for b and t, and there is a connection $[A_1, A_2]$ then we define A_2 for t as $A_2(t) := A_2(b)$. In particular it follows that $A_2(b)$ is hypothesized valid for t (with a justification depending on the justification of the connection $[A_1, A_2]$).

Elsewhere one of us distinguished several modes of analogical inference [20]: typical-example-based (shortly "typex-based"), schema-based, and theory-based. In this paper we are interested in the typex-based mode only and for this the general inference pattern (**AN**) has to be modified accordingly. The typex-based reasoning to be formalized below, can be stated informally as follows: given a base and a target example where both the base and the target are examples of the same concept, and furthermore if the target is in some sense less typical but still within the range of the base, then some properties may be transferred from the base to the target example.

In order to formalize this typex-based analogical inference, we have to introduce the notion of *relevancy* of an aspect.

Definition (Relevancy of an Aspect): An aspect A is said to be *relevant in an example* $e \in \mathcal{E}_\kappa$ if $A(e)$ is relevantly valid in e. We write $relevant(A, e)$.

Example: In the apple example we have $relevant(\text{colour}, a_1^*)$.

5.2 Inference Rules for Analogical Reasoning with Typical Examples

Restrictions have to be imposed on the above rule (**AN**) for typex-based analogical inference: Normally, connections as complicated as [structure, function] are based on generalizations of several examples (that is, these connections are schema-based or theory-based in the sense of [20]). In the typex-based analogical inference mode we just want to use one example, the typical one, and therefore we do not need any explicit information on connections. Hence, one of the prerequisites of the above (**AN**) rule, which requires the connection $[A_1, A_2]$, is not necessary and can be replaced by other information. However, it is necessary that the aspects A_1 and A_2 are *relevant* for the typical example b^* (the base). Else any formula that is valid in b^* could be inferred because of an irrelevant or accidental correspondence between b^* and t.

Example: Looking at the apple example, let round(x) be valid in t. Without the constraint on relevancy it is possible to infer sweet(x) to be valid in t by analogy.

Now we are going to modify (**AN**) for typex-based analogical inference. Let b^* and t be examples, i.e. elements, of a concept C, and let A_1 and A_2 be aspects. Assuming

(i) $A_1(t) \Leftrightarrow A_1(b^*)$, $relevant(A_1, b^*)$, $relevant(A_2, b^*)$, typex(b^*).

Then $relevant(A_1, b^*)$, $relevant(A_2, b^*)$, and typex(b^*) supports the connection $[A_1, A_2]$. Hence, from (i) follows

(ii) $A_1(t) \Leftrightarrow A_1(b^*), [A_1, A_2]$.

Because of (ii) and the inference rule (**AN**) we have

(iii) $A_2(t) := A_2(b^*)$.

That is, we have the permission to define A_2 for t by $A_2(b^*)$ if this is consistent with the previous knowledge about t.

The inference rule for examples b^* and t of a concept C expressed formally is:

$$(\mathbf{TYP1}) \qquad \frac{A_1(t) \Leftrightarrow A_1(b^*), relevant(A_1, b^*), relevant(A_2, b^*), \mathrm{typex}(b^*)}{A_2(t) := A_2(b^*)}$$

This is our first inference rule for typex-based analogical reasoning.

Example: If colour$(a_2^*) \Leftrightarrow$ colour(t), $relevant(\mathrm{colour}, a_2^*)$, $relevant(\mathrm{taste}, a_2^*)$, and typex$(a_2^*)$ then it is permissible to define the taste for t by
taste(t) := taste(a_2^*) = sour(x). In other words, we hypothesize $t \mid\!\sim \mathrm{sour}(x)$.

As the above mentioned experiments in empirical psychology have shown, people often transfer knowledge from a typical example to another example by an explicit use of the typicality rating. To model this, our second modification of (**AN**) uses the relation $\sqsubset$ explicitly. We restrict this modification to cases where typex(e) and $e_1 \sqsubset e$ implies the existence of an aspect A such that $relevant(A, e)$ and $A(e_1) \Leftrightarrow A(e)$. This restriction is not really very severe and in accordance with psychological findings, see e.g. [32].

Assuming that A_2 is an aspect, which is still undefined for t, and that b^* and t are examples of C, and assume

(i) typex(b^*), $t \sqsubset b^*$, $relevant(A_2, b^*)$.

From $t \sqsubset b^*$ follows by the above restriction that there is an aspect A_1 with $relevant(A_1, b^*)$, and $A_1(t) \Leftrightarrow A_1(b^*)$ Hence, we have

(ii) $A_1(t) \Leftrightarrow A_1(b^*), relevant(A_1, b^*), relevant(A_2, b^*), \mathrm{typex}(b^*)$.

Inference rule (**TYP1**) and (ii) together yield

(iii) $A_2(t) := A_2(b^*)$.

Now, (i) and (iii) together give our second inference rule expressed formally as:

$$(\textbf{TYP2}) \qquad \frac{\text{typex}(b^*), t \sqsubset b^*, relevant(A_2, b^*)}{A_2(t) := A_2(b^*)}$$

For the aspect A_2 and examples b^* and t of C, this second pattern of typex-based analogical inference expresses a permission to define A_2 for t, if that is consistent with the previous knowledge about t.

Example: Because of typex(a_1^*), $a_{111} \sqsubset a_1^*$, and $relevant(\text{colour}, a_1^*)$ it is allowed to define colour(a_{111}) as red(x).

Based on the two inference rules (**TYP1**) and (**TYP2**) our reasoner has two procedures for checking an example t:

1. Analogical reasoning can proceed according to (**TYP1**) for a target example t of a concept C (which is not completely known) in the following way:

Find some appropriate non-tautological formula H_1 with $t \mid\!\sim H_1$, which is relevantly valid in a typical example b^*. From another formula H_2, relevantly valid in b^*, hypothesize $t \mid\!\sim H_2$ by (**TYP1**).

Example: Consider again the apple example. For an example t with $t \mid\!\sim \text{sour}(x)$ and a still unknown $\sqsubset$-relationship one can hypothesize $t \mid\!\sim \text{green}(x)$, because of $a_2^* \mid\!\sim_{\mathbf{r}} \text{green}(x)$ and $a_2^* \mid\!\sim_{\mathbf{r}} \text{sour}(x)$. For a t with $t \mid\!\sim \text{green}(x)$ one can hypothesize $t \mid\!\sim \text{sour}(x)$ as well as $t \mid\!\sim \text{sweet}(x)$.

2. For a partially known target example t already integrated into the concept structure of C, the analogical inference is not necessarily based on the indirect mode via a formula H_1. A simpler and more natural procedure for typex-based analogical reasoning results from the pattern **TYP2**. In this case analogical reasoning proceeds for such a target example t of a concept C according to the following process:

Look for a typical example b^* of C with $t \sqsubset b^*$. Look for a formula H relevantly valid in b^*. Then hypothesize analogically $t \mid\!\sim H$ according to **TYP2**.

Example: In the apple example, because of $a_{111} \sqsubset a_1^*, a_1^* \mid\!\sim_{\mathbf{r}} \text{red}(x)$ the hypothetical validity of red(x) in a_{111} can be established.

This procedure is a more semantically oriented way of analogical reasoning that approximates human analogical reasoning by typical examples. It does not explicitly consider connections.

It is however interesting to note for both kinds of typex-based analogical reasoning, that several conditions of (**AN**) do not explicitly occur in (**TYP1**) and (**TYP2**), and hence we do not need to compute them explicitly. They are given implicitly by the concept structure and the typicality of b^*. In particular, connections and in (**TYP2**) the computation of the similarity of b^* and t can be avoided. This is an important advantage of reasoning by examples.

6 Conclusion

The particular kind of analogical reasoning that is presented in this paper is based on semantic information about concept structures and typical examples. Relevant properties are transferred from a typical example to another similar case. Based on this procedure it is inter alia possible to model certain aspects of human reasoning that are not rule-based per se.

There is related work: Winston's analogy model [34] and [35] explicitly preferred prototypical aspects of the target. Prototypical information is statistically determined for a certain class of examples. However, the statistical approach contradicts the current view of prototypicality. Johnson-Laird and Byrne [13] have developed a procedure based on model-checking of constructed (not necessarily typical) models. They incorporate non-monotonic aspects of human reasoning. In our approach non-monotonicity emerges as we refer to information about typical examples as a kind of default. We might have to withdraw some formulae, that are hypothesized to be true for a given example if we get further information about the example or the rating. Shoham's approach [30] for describing non-monotonic logics seems to be well suited for formally describing this procedure.

We have confined our treatment to typical examples of a concept only. This can be generalized by considering also typical examples b^* of a property (of an aspect A). Then similarity classes consist of examples b with values $A(b)$ similar or equal to $A(b^*)$. An appropriately generalized inference rule (**TYP2**) can be used in the second analogical reasoning procedure.

Our approach is open for using modal operators. For instance, relevancy could be stated as a modality and it would be interesting to investigate the relationship to possible world semantics [17]. Halpern and Vardi [12] constructed possible world semantics for model checking and they compared the complexity of deduction and model-checking. They found that one determinant the complexity of model-checking depends on, is the size of the Kripke structure. Our approach uses a heuristic that restricts the number of considered possible situations and ignores the rest. In addition, the complexity is limited by focusing the attention on relevant features only. This could be a good basis for further discussion: There is something afoot demanding for hybrid systems.

Acknowledgement

We like to thank Michael Kohlhase for thoroughly reading a draft of this paper.

References

[1] S.L. Armstrong, L. Gleitmann, and H. Gleitmann. What some concepts might be. *Cognition*, **13**:263–308, 1983.

[2] A. Bundy. The use of explicit plans to guide inductive proofs. In *Proc. of the 9th CADE*, Argonne, Illinois, USA, 1988. Springer Verlag, Berlin, Germany. LNCS 310.

[3] C. Cherniak. Prototypicality and deductive reasoning. *Journal of Verbal Learning and Verbal Behavior*, **23**:625–642, 1984.

[4] A. Collins and R. Michalski. The logic of plausible reasoning: A core theory. *Cognitive Science*, **13**(1):1–50, 1989.

[5] T. Davis and S. Russell. A logical approach to reasoning by analogy. In *Proc. of the 10th IJCAI*, pages 264–270, Milan Italy, 1987. Morgan Kaufmann.

[6] H. L. Dreyfus and S. E. Dreyfus. *Mind over Machine.* The Free Press, 1986. The Power of Human Intuition and Expertise in the Era of the Computer.

[7] D. Gentner. Structure mapping: A theoretical framework for analogy. *Cognitive Science*, **7**(2):155–170, 1983.

[8] 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.

[9] M. Gick and K. Holyoak. Schema induction and analogical transfer. *Cognitive Psychology*, **15**(1):1–38, 1983.

[10] R. Greiner. Abstraction-based analogical inference. In D. Helman, editor, *Analogical Reasoning*, pages 147–170. Kluwer Academic Press, Dodrecht, Netherlands, 1988. in Series: Perspectives of Artificial Intelligence, Cognitive Science, and Philosophy.

[11] J. Hadamard. *The Psychology of Invention in the Mathematical Field.* Princeton Univ. Press, Princeton, 1945.

[12] J. Halpern and M. Vardi. Model checking vs. theorem proving. In *Proceedings of the Second International Conference on Knowledge Representation KR'91*, pages 325–334, Cambridge MA, 1991.

[13] P. Johnson-Laird and R. Byrne. *Deduction.* Ablex Publishing Corp., 1990.

[14] M. Kerber. Some aspects of analogy in mathematical reasoning. In K. P. Jantke, editor, *Analogical and Inductive Inference; International Workshop AII '89, Reinhardsbrunn, GDR*, pages 231–242. Springer Verlag, Berlin, Germany, October 1989. LNAI 397.

[15] M. Kerber. On the representation of mathematical knowledge in frames and its consistency. In M. De Glas and D. Gabbay, editors, *WOCFAI '91 – Proc. of the First World Conference on the Fundamentals of Artificial Intelligence*, pages 293–301, Paris, France, 1991. Angkor.

[16] K. Konolige. Abduction versus closure in causal theories. *Artificial Intelligence*, pages 255–272, 1991.

[17] S. Kripke. Semantical analysis of modal logic I. *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*, **9**:67–96, 1963.

[18] T. Kuhn. *The Structure of Scientific Revolutions.* University of Chicago, 2nd edition, 1962/1970.

[19] G. Lakoff. *Women, Fire, and Dangerous Things – What Categories Reveal about the Mind.* The University of Chicago Press, 1987.

[20] E. Melis. Study of modes of analogical reasoning. TASSO Report Nr.5, GMD, St. Augustin, Germany, 1990.

[21] C. Mervis and E. Rosch. Family resemblance: Studies in the internal structure of categories. *Cognitive Psychology*, **7**:573–605, 1975.

[22] C. Mervis and E. Rosch. Categorization of natural objects. *Annual Review of Psychology*, **32**:89–115, 1981.

[23] H. de la Motte-Haber. Regelmaß und Einmaligkeit als ästhetische Prinzipien. In R. Wille, editor, *Symmetrie in Geistes- und Naturwissenschaften*, pages 17–29. Springer Verlag, Berlin, Germany, 1988.

[24] G. Pólya. *Mathematics and Plausible Reasoning.* Princeton University Press, Princeton, New Jersey, USA, 1954. Two volumes, Vol.1: Induction and Analogy in Mathematics, Vol.2: Patterns of Plausible Inference.

[25] L. Rips. Inductive judgements about natural categories. *Journal of Verbal Learning and Verbal Behavior*, **14**:665–681, 1975.

[26] L. Rips. Similarity, typicality, and categorization. In S. Vosniadou and A. Ortony, editors, *Similarity and Analogical Reasoning*, pages 21–59. Cambridge University Press, 1989.

[27] E. Rosch. Natural categories. *Cognitive Psychology*, **4**:328–350, 1973.

[28] S. J. Russell. *The Use of Knowledge in Analogy and Induction.* Morgan Kaufmann, San Mateo, California, USA, 1989.

[29] M. Schmidt-Schauß. *Computational Aspects of an Order-Sorted Logic with Term Declarations.* LNAI 395. Springer Verlag, Berlin, Germany, 1989.

[30] Y. Shoham. *Reasoning about Change – Time and Causation from the Standpoint of Artificial Intelligence.* The MIT Press, Cambridge, Massachusetts, USA, 1988.

[31] C. Walther. A many–sorted calculus based on resolution and paramodulation. In A. Bundy, editor, *Proc. of the 8th IJCAI*, pages 882–891, Los Altos, California, USA, August 1983. Morgan Kaufmann.

[32] E.J. Weiner. A knowledge representation approach to understanding metaphors. *Computational Linguistics*, **10**(1):1–14, 1984.

[33] J. Weitzenfeld. Valid reasoning by analogy. *Philosophy of Science*, **51**:137–149, 1984.

[34] P. Winston. *Learning Structured Descriptions from Examples.* PhD thesis, Institute of Technology, Cambridge Mass., 1970.

[35] P. Winston. Learning by creating and justifying transfer frames. *Artificial Intelligence*, **10**(2):147–172, 1978.

[36] P. Winston. Learning and reasoning by analogy. *Communications of the ACM*, **23**(12): 689–703, 1980.