# On the Logical Formalization of Analogies and Theory Blending in the HDTP Framework

Dissertation
zur Erlangung des Doktorgrades
des Fachbereichs Humanwissenschaften
der Universität Osnabrück

vorgelegt
von

**Ulf Krumnack**

aus
Bonn

Osnabrück, 2015

This is a cumulative doctoral thesis that includes articles published in peer-reviewed conference proceedings and books. The articles are not included in this electronic version of the dissertation due to copyright reasons. Instead, only the abstracts are embedded into this synopsis, but the full bibliographical reference and (where available) the corresponding URL of each article are provided.

Supervisor and first examiner: Prof. Dr. Kai-Uwe Kühnberger
Co-examiner: PD Dr. Helmar Gust
Co-examiner: Prof. Dr. Ute Schmid

# Abstract

Analogies are considered a cognitive core mechanism, that is applied in many everday reasoning processes. Heuristic-driven theory projection (HDTP) is a logic-based framework that allows to model analogies with the aim of making analogical reasoning techniques available for artificial intelligence systems.

The formal properties of HDTP are investigated, refining and extending some of the original ideas. A special form of restricted higher-order anti-unification is proposed as a means for the generalization process, allowing to account for flexibility in the mapping while staying computationally tractable. Concerning the semantics, it is argued that a sensible interpretation can be given to the syntactic processes, based on an understanding of the involved mappings as a decent type of theory morphisms. The logical nature of HDTP also allows for a notion of re-representation that is discussed from a theoretical and algorithmic point of view. Moreover, the framework of HDTP is also analyzed from the abstract perspective of institution theory, suggesting that the main ideas can be spelled out in other logical formalisms as well.

To collect support for the practical utility of HDTP, it is applied to different fields in a series of studies. The domain of geometric analogy serves as an arena to demonstrate the operation of HDTP, including the treatment of ambiguous problems based on thoughts from Gestalt psychology. Another line of research explores how the idea of conceptual blending can be related to analogies and a formalization building on HDTP is presented, leading to the notion of theory blending. These ideas are applied to a classical problem of the field, the interpretation of noun-noun compounds, but they proof to be applicable in other areas as well, demonstrated by a framework for counterfactual reasoning. Furthermore, applications of analogical reasoning and theory blending in mathematics are discussed, including the formal modeling of an example from the history of mathematics and a framework to support mathematical discovery.

# Contents

# Part I

# Synopsis

# Chapter 1

# Prolegomenon

This doctoral thesis is concerned with the logic-based modeling of analogies and related cognitive mechanisms. It reflects my work in the AI group at the institute of cognitive science in Osnabrück during the last years and is based on a selection of 11 peer-reviewed papers, published as journal articles, book chapters or in conference proceedings.

This text consists of two parts. In this first part I give a concise overview over the topics of the thesis. Following these introductory remarks, chapter 2 provides a short survey of research in analogies, analogical reasoning and modeling of analogies with the intent to clarify the place of this work within the greater fields of cognitive science and artificial intelligence. In the following two main chapters 3 and 4, theoretical and practical aspects of Heuristic-Driven Theory Projection (HDTP), a framework originally presented in Gust et al. (2006), are analyzed. Chapter 3 is concerned with formal aspects of HDTP, covering my own, original work on the topic. After a short presentation of the framework, the notion of generalization, the problem of re-representation, semantic issues, and an abstract perspective are investigated. Chapter 4 describes applications of HDTP in different fields, reporting work that I have conducted in collaboration with colleagues from the AI group in Osnabrück and from the University of Edinburgh. The presentation starts with geometric analogy problems, then turning to the treatment of blending processes, analyzing noun-noun compounds, counterfactual reasoning, and finally the role of analogies in mathematical thinking. Chapter 5 concludes the work by summarizing the main achievments and sketching possible future developments.

All sections include a brief introduction into the respective topic, providing references to related research before summarizing my own work, which the interested reader can find in full length in the publications collected in the second part. The papers presented there have been choosen to give a representative impression of the work that has been conducted, trying to cover the relevant topics while minimizing repetition and redundancies. Due to the interdisciplinary breadth of cognitive science, different foci and styles of presentation are unavoidable, then again indicating the relevance and applicability of the approach in the whole field. Some results, which have not been published yet, have been put into an appendix, to prevent the main text from being disrupted by detailed technicalities.

# Chapter 2

# Analogies in Cognitive Science and Artificial Intelligence

An analogy is a structural correspondence between two domains, usually referred to as source (or base) and target. It can function as a basis for translating and transferring concepts, allowing to apply knowledge from a well-known source domain to a lesser-known target domain, establishing new ideas there. This process of analogical inference differs from other classical ways of reasoning like deduction, induction, and abduction. It is abundant in everyday life, reaching from children's imitation games to political argumentation.

It has to be emphasized that analogical inference is not a sound reasoning procedure, i.e. unlike logical deduction, starting from valid premisses there is no guarantee that the inferred conclusions actually hold in the target domain. Although the identification of criteria that would allow to assess and support the reliability of anlogical inferences has been a recurring theme in research on analogies there seems to be a consensus now that in general no such principles exist (Bartha, 2010). Nevertheless, analogical reasoning allows the introduction of new ideas and hence is considered a valuable tool, provided that analogical inferences are recognized as mere hypotheses, that have to be supported by additional means to claim validity. The power to introduce new ideas into a domain makes analogical reasoning a decent mechanism to explain phenomena of creativity.

Also in science analogies are considered an indispensable tool. For example in physics, Oppenheimer (1956) argues in that direction, pointing out that much of modern quantum physics has been developed based on analogies to models and formalisms of classical physics, an example elaborated in much more detailed by Darrigol (1992). But not only for analyzing new phenomena but also for communicating ideas analogies are an apt tool, as argued for instance by Maxwell. A well choosen analogy allows to quickly access a new domain, explaining new phenomena by relating them to already familiar concepts, making it well suited for educational purposes. A standard example, usually attributed to Rutherford (1911) – although main ideas have been employed before – is the analogy describing the structure of an atom by relating it to the solar system.

Even though the details of this model proved untenable, the analogy is still used in education, probably due to its accessibility and memorability.

Analogies are applied and investigated in many different fields leading to an abundance of research which makes an exhaustive overview unrealistic. Even Guarini et al. (2009), who compiled a list of more than 1,400 references to works on analogy, mostly journal articles and monographs from the last decades, had to focus their collection on English language works from Western academic tradition.

The omnipresence of analogies in everday practice has also attracted the attention of researchers in cognitive science. There, analogies are often considered a core mechanism of cognition (Chalmers et al., 1992; Forbus et al., 1998; Hofstadter, 2001; Gentner, 2003; Gust et al., 2008). The process of mapping two given structures and the subsequent transfer of information can function as base for various cognitive capabilites. Examples include memory and adaption, (common sense) reasoning, learning by transfer and by generalization, and creativity.

A mile stone was set with the presentation of the *structure mapping theory* (SMT) by Gentner (1983). Assuming a structured representation of knowledge, this theory provides a set psychologically motivated principles – like systematicity and a 1:1 matching constraint – for an analogical relation between two domains. Falkenhainer et al. (1986, 1989) provide an implementation of SMT called *structure mapping engine* (SME), that employs these principles to compute mappings between two domains given in a graph-based representation. SME has undergone further developments and is still used in applications like NuSketch (Forbus et al., 2004). The SMT was challenged by Chalmers et al. (1992) who criticized the static representation used in SMT and emphasized the importance of adaptation and and change of representation as part of the mapping process, a topic that will be taken up in section 3.2 or this thesis. Hofstadter and the Fluid Analogies Research Group (1995) propose to understand analogy as the restructuring of a fluid representation in the presence of another, also fluid counterpart and propose the CopyCat system that operates along these lines. A lot of other approaches to model analogical thinking have been proposed, differing in the representation formalisms and mapping strategies applied. French (2002) and more recently Gentner and Forbus (2011) have compiled surveys.

Also in artificial intelligence research on analogy has a long history. A first system developed by Evans (1962) was focusing on proportional geometric ananalogies (cmp. section 4.1). While this early research was directed towards systems that can solve certain explicitly stated analogy problems, more recently the interest has shifted towards analogies as a general reasoning mechanim that can be employed in different contexts, especially in the new branch of *artificial general intelligence* (AGI). AGI focuses on the understanding and implementation of intelligence at a human level.[1] It thereby takes up the original goals stated by the founders of AI, which seem to have come out of sight during the further evolution of the field. While many techniques and systems have been developed which show impressive, human-like or even super-human performance in certain applications, they usually fail to generalize to other fields. Therefore,

---

[1] There are annual conferences on AGI since 2008 and the *Journal of Artificial General Intelligence*

one focus of AGI research lies in identifiying instruments that can be applied broadly to different domains and tasks. Analogies and similar cognitively inspired mechanisms are promising candidates and the present work can be seen as one effort to push them further forward.

# Chapter 3

# The Formal Framework of HDTP

I will follow the path set up by Gust et al. (2006), who introduce *heuristic-driven theory projection* (HDTP). This framework defines a way to specify domains as well as an algorithm to compute analogical relations and to propose analogical inferences. Source and target domains are described by logical theories given as first-order axiomatizations. The system constructs a mapping of source and target axioms based on several heuristics, which aim at finding a tradeoff between a simple mapping and a large coverage of axioms. Unmatched axioms from the source domain are candidates for analogical transfer. Based on the analogical mapping, they are translated into the target language and serve as candidates for analogical inferences, which have to be sanctioned by some oracle. Although the system operates on a purely syntactic basis, the authors indicate that the established mapping creates a correspondence on the semantic level as well, allowing to associate given models for the source and target theory, and to adapt the target model after a successful transfer.

Before starting, I may briefly explain the choice of HDTP for this work. This should not be understood as judging HDTP to be superior to other frameworks, which are abundant, some of which much better-known. Many ideas presented in this thesis could probably be applied, more or less directly, to other approaches as well. However, the logic-based nature of HDTP is especially well suited for a formal analysis. It contrasts with the often informal or purely operational descriptions of other systems, that make it hard to directly compare different frameworks. The present work also aims at improving this situation.

## 3.1 Analogy by Generalization

The theme of HDTP is analogy by generalization. Following the SMT tradition, an analogy indicates corresponding structures in source and target domain. These structures can be viewed as an abstract core, that is instantiated in different ways in both domains. The analogical relation reflects this abstraction by mapping entities that are instances of the same abstract idea (cmp. figure 3.1).

This idea can be applied to most formal frameworks for analogy making, although it is rarely spelled out. HDTP makes this abstraction explicit by com-

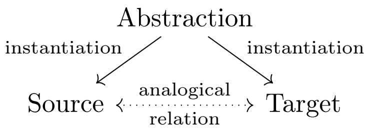

Figure 3.1: Analogy by generalization: common structures of source and target domain can be seen as instances of an abstracted domain

puting a generalized theory which can be "projected" into both of the original domain theories. It should be mentioned that these projections are in general not surjective, i.e. there may be parts of the input theories that do not have corresponding parts on the other side, and hence should not be reflected in the generalization. Especially in analogical reasoning, one assumes that the source domain contains additional knowledge, that can be transferred to the target side in the reasoning step. I will call those parts of the domains, that lie in the image of projections, *covered* by the analogy, while the remainder is *uncovered.*[1]

The problem of generalization in the context of first-order logic has been investigated by researchers in Edinburgh around 1970 (Plotkin, 1970, 1971; Reynolds, 1970; Popplestone, 1970). Plotkin (1970) and Reynolds (1970) independently develop an algorithm that computes a generalization for a given set of terms by comparing their structures and replacing conflicts by fresh variables. Reynolds (1970) coined the notion of *anti-unification* to express the duality to the well-known operation of *unification* (cmp. figure 3.2):[2] while unification aims at finding the most general unifier (mgu) for a set of terms, in anti-unification one is interested in the least general generalization (lgg). In case of two terms $t_1$ and $t_2$, a generalization is a term $g$ together with substitutions $\sigma_1$ and $\sigma_2$ such that $g\sigma_1 = t_1$ and $g\sigma_2 = t_2$. A generalization is said to be least general, if there is no other generalization $g'$ that is more specific, i.e. an instance of $g$. A lgg retains the structure common to both terms and introduces variables only in case of mismatch. Anti-unification can be seen as an approach towards a formal counterpart of the human ability of abstraction.

Although anti-unification has attracted by far less attention then unification, it has been investigated in a number of fields. Walker et al. (1987) study generalization in the context of inductive learning and present a refined algorithm for anti-unification, Knight (1989) has contrasted it with unification, Pfenning (1991) uses anti-unification for proof generalization, while Feng and Muggleton (1992) induce higher-order rule templates for automatic program transformation. Furtado (1992) already discusses the potential of anti-unification for analogy by generalization and Hasker and Reddy (1992) also see a

---

[1] The covered parts of the domains roughly correspond to the notion of positive analogy in Hesse (1966). The uncovered parts are subdivided further into negative analogies, which are known not to hold, and neutral analogies, whose status is not yet known and which are candidates for analogical transfer.

[2] Although anti-unification is often referred to as the "dual" operation of unification, it should be emphasized, that these two notions are not fully symmetric: for unification there is a distinction between weak and strong unification, whereas such a distinction is meaningless for anti-unification. Also in the standard setting of first-order terms and substitutions, a set of terms may not be unifiable, while it will always be anti-unifiable. In general, results on unification can often not simply be "dualized" to anti-unification or vice versa.

$$\begin{tikzcd}
t_1 \arrow[dr, "\mu_2"'] & & t_2 \arrow[dl, "\mu_1"] \\
& \mathrm{mgu} &
\end{tikzcd}$$

(a) (weak) unification

$$\begin{tikzcd}
& \mathrm{lgg} \arrow[dl, "\sigma_1"'] \arrow[dr, "\sigma_2"] & \\
t_1 & & t_2
\end{tikzcd}$$

(b) anti-unification

Figure 3.2: Formal duality of unification and anti-unification: a most general unifier (mgu) is an instance of two terms $t_1$ and $t_2$, while the least general generalization (lgg) has $t_1$ and $t_2$ as instances.

relation to analogies when they use anti-unification as a tool to analyse proof modification in automated proof systems. Further applications can be found in domains like similarity-based learning, software code refactoring, or program replay (Hasker, 1995). Many of these works leave the classical setting of first-order terms and apply the idea to different systems, as Armengol and Plaza (2000) focusing on feature terms, Lu et al. (2000), who study anti-unification in the $\lambda 2$ calculus,[3] and Burghardt (2005), who considers anti-unification modulo an equational theory. More recently Kutsia et al. (2011) investigate the anti-unification of unranked terms and hedges and Baumgartner et al. (2012) compute a higher-order pattern generalization with analogical programming and analogical theorem proving in mind.

Already Gust et al. (2006) point out, that simple (first-order) anti-unification does not seem to be sufficient for analogy making. A problem of first-order anti-unification is its failure to capture certain commonalities of terms. Whenever the topmost symbols of two terms differ, the entire terms are generalized to a single variable $X$ and any similarity between the subterms is ignored. Figure 3.3 shows some examples of overgeneralization, for which first-order anti-unification would not reflect commonalities in pairs of terms and hence would fail to identify analogous structures.

Some of these cases can be addressed by introducing a meta function "apply" (and reformulating terms, e.g. "$f(a)$" to "$\mathrm{apply}(f, a)$") but this is insufficient for the examples of figure 3.3b–3.3d. Hence a form of higher-order anti-unification seems suitable for analogy making, i.e. a formalism allowing for generalized function symbols that can be instantiated by complex functions. However, it is well-known that – similar to higher-order unification – higher-order anti-unification suffers from severe theoretical problems. Not only uniqueness of the least general generalization is lost but, as demonstrated by Hasker (1995), a general notion of higher-order anti-unification is not even well-defined, since one can construct infinite chains of ever more and more specific generalizations.

To be useful for analogy making, higher-order anti-unification should be restricted. At least two ways to achieve this have been discussed in the literature. One way is to restrict the form of generalized terms, as proposed by Pfenning (1991), who allows only so-called "higher-order patterns". A merit of this formalism is that the uniqueness of the least general generalization can be regained. However, higher-order patterns seem to be too rigid for analogies, as they still overgeneralize in some cases. Another way is gone by Hasker (1995), who uses

---

[3] The $\lambda 2$ calculus is an extension of the simply typed $\lambda$ calculus, that allows type variables, see e.g. Barendregt (1991).

$$\begin{tikzcd} & X \arrow[dl] \arrow[dr] & \\ f(a) & & g(a) \end{tikzcd}$$

(a) different function symbols

$$\begin{tikzcd} & X \arrow[dl] \arrow[dr] & \\ f(a) & & g(h(a)) \end{tikzcd}$$

(b) nested functions

$$\begin{tikzcd} & X \arrow[dl] \arrow[dr] & \\ f(a) & & g(a,b) \end{tikzcd}$$

(c) additional arguments

$$\begin{tikzcd} & X \arrow[dl] \arrow[dr] & \\ f(a,b) & & g(b,a) \end{tikzcd}$$

(d) permuted arguments

Figure 3.3: Examples for terms with common structure that is not reflected by first-order anti-unification, resulting in overgeneralization: a single variable $X$.

combinator terms instead of $\lambda$-terms to express functions. The advantage of combinators is, that they allow for a better control how function arguments are used in terms and by restricting the set of permissible combinators (to so called "relevant" combinators), versions of higher-order anti-unifications with good properties can be defined. I will present a restricted form of anti-unification for HDTP, which is similar in spirit to this second approach.

In the context of analogies, there is a plausible criterion that helps to constrain the process of anti-unification: generalizations should not be more complex than the original expressions in the source and target domain, from which they abstract. The rationale behind this is that there should be nothing in the common core that is not supported by both input domains. In [Publication 1] we use this criterion to introduce an approach to higher-order anti-unfication that restricts the form of substitutions that can be applied.

The basic idea is to allow only substitutions, that can be constructed as a chain of selected "basic substitutions". Figure 3.4 lists the four types of basic substitutions proposed in [Publication 1]. Applying one of these substitutions will never reduce the number of symbols in a term and therefore guarantees that a generalization only contains variables that are backed by both input domains. Furthermore, the substitutions are choosen in a way that they can account for the problems shown in figure 3.3.

This formalism seems to be suitable in the context of HDTP as it achieves at least three goals. First, it provides a notion of higher-order anti-unification with least general generalizations that are more specific than in the first-order case, i.e. it allows to identify analogous structures for given terms that would be ignored in the simple case. Second, although such generalizations are no longer unique, the formalism guarantees that their number is finite and restricted by the size of the input terms, making them effectively computable. A simple algorithm is presented in [Publication 1], and a refined version, has been published by Schmidt (2010). Third, it allows for a stepwise reduction of the complexity by excluding some types of basic substitutions. Robere and Besold (2012) have analyzed the complexity of the algorithm and identified permutation and renamings as two potential sources of problems when domains become large. It should be remarked that one can introduce a normal form for chains of basic substitutions, by first applying permutations, then argument insertions, followed by fixations and finally renamings, putting the problematic cases at the beginning and the end, where they can be treated in a special way. The basic substitution of renaming is only needed for cases where the input terms contain variables, and they can be dropped completely, if those variables are named in a common way before starting anti-unification. The topic of variables will be dealt with in

1. A *renaming* $\rho^{F,F'}$ replaces a variable $F \in \mathcal{V}_n$ by another variable $F' \in \mathcal{V}_n$ of the same arity:
$$F(t_1, \dots, t_n) \xrightarrow{\rho^{F,F'}} F'(t_1, \dots, t_n).$$
2. A *fixation* $\phi_c^V$ replaces a variable $F \in \mathcal{V}_n$ by a function symbol $f \in \mathcal{C}_n$ of the same arity:
$$F(t_1, \dots, t_n) \xrightarrow{\phi_f^F} f(t_1, \dots, t_n).$$
3. An *argument insertion* $\iota_{V,i}^{F,F'}$ with $0 \leq i \leq n$, $F \in \mathcal{V}_n$, $V \in \mathcal{V}_k$ with $k \leq n - i$, and $F' \in \mathcal{V}_{n-k+1}$ is defined by
$$F(t_1, \dots, t_n) \xrightarrow{\iota_{V,i}^{F,F'}} F'(t_1, \dots, t_{i-1}, V(t_i, \dots, t_{i+k-1}), t_{i+k}, \dots, t_n).$$
4. A *permutation* $\pi_\alpha^{F,F'}$ with $F, F' \in \mathcal{V}_n$ and $\alpha : \{1, \dots, n\} \to \{1, \dots, n\}$ bijective, rearanges the arguments of a term:
$$F(t_1, \dots, t_n) \xrightarrow{\pi_\alpha^{F,F'}} F'(t_{\alpha(1)}, \dots, t_{\alpha(n)}).$$

Figure 3.4: The four basic substitutions introduced in [Publication 1]

more detail in section 3.3 when considering semantic aspects.

The use of basic substitutions has the additional advantage that they provide a means to assess the complexity of a substitution. While Gust et al. (2006) mention simplicity of substitution as one heuristics for finding good analogies, they give only a vague description of how it should be determined. By assigning cost values to individual basic substitutions, one gets gets a complexity measure for substitutions that allows to compare different solutions. This measure also provides a way to to select a preferred generalization, in case that multiple such generalizations exist. Even more, as demonstrated in [Publication 1], it can be used to encourage the reuse of substitutions, when not individual terms but two sets of terms shall be anti-unified. This is especially interesting in the context of HDTP, which aims at generalizing two domain theorie, given by sets of axioms.

## 3.2 Re-representation

The ability to see things in a new way, i.e. to change representations, can be seen as a key element of cognitive abilities like understanding, learning, reasoning, and creativity. A famous example from physics is the conceptualization of light as either waves or particles, depending on the phenomenon to be explained. But re-representation is not limited to the selection from a set of given interpretations. Depending on the reasoners goals, she can also introduce new ideas to make sense of the current situation. Such processes are for example examined by Kotovsky and Gentner (1999) studying concept learning and development of children and Schön (1963), who describes discussions in a development team that passes through different conceptualizations of their product, an example later picked up by Indurkhya (1989) to argue for an interactive conception of analogy.

Establishing an analogical relation presupposes the existence of compatible structures in the source and target domain. What exactly compatibility means in this context depends on the particular analogy model in use, but some variant of this notion should exist in every model. However, the assumption that parallel structures are already given, and that the task of an analogy engine only consists in finding the best alignment between the given descriptions, is somewhat artificial and unsatisfactory in many contexts. While there might be situations, where the pure mapping task is sufficient, it is often considered more realistic to include the preparation of a description that allows for finding an alignment as part of the analogy making process.

A fundamental critique of ready-made representations in the context of analogies, like those used by the SME, is brought forward by Chalmers et al. (1992), who doubt that such models adequately reflect cognitive processes. They rather propose to consider analogy making as a form of high-level perception, in which conceptual structures are steadily adapted according to the stimuli presented. These ideas are reflected in the Copycat architecture (Hofstadter and Mitchell, 1994; Hofstadter and the Fluid Analogies Research Group, 1995) by spreading activation between codelets in the slipnet: processes running in parallel result in competing or cooperating temporary perceptual constructs, that can be created, destroyed, or varied probabilistically, until a satisfying match is found. Similar in spirit is the approach of Kokinov and Petrov (2001), who in their AMBR model represent knowledge by a network of micro-agents. Within that network structure the individual activity of the agents also influences their contribution and thereby the operation of the system. Changing the activation pattern can hence be seen as a kind of re-representation. As the initial activation is determined by the environment of the system, this mechanism allows to account for context effects. Both of these approaches can be characterized as hybrid in the sense that – besides the symbolic structure – they employ an additional dimension of activation to realize a change of representation.

But also proponents of purely symbolic systems acknowledge the need for re-representation. Already Holyoak and Thagard (1989) identify re-representation as a major issue to be considered, although no details on how it could be implemented in their ACME system are given. Yan et al. (2003) extend the SME by an additional component that is introduced after the mapping phase. If the analogical mapping is unsatisfying, the original representations can be changed based on the detected shortcomings. A set of different techniques can then be applied, like (logical) transformation (changing LESSTHAN to GREATERTHAN) and semantic decomposition (changing RUN to MOVE) that help to relax the *identicality* constraint for relational matches. In addition *entity splitting* or *entity collection* can help to overcome problems with the *one-to-one* constraint.

The algebraic interactionist model of Indurkhya (1992) focuses on re-representation as an integral part of the analogical mapping process. *Analogy by rendition* is characterized by the interaction of source and target to create a description of the target in terms of the source, thereby adapting the descriptions of both domains. Based on these ideas, Dastani (1998) develops an algebraic language of perception that allows to describe geometric figures in different ways. In the context of proportional geometric analogies (of the form $A : B :: C : D$), the comparison of structures of the given figures can guide a re-representation process that allows for describing the relation between the source figures $A$ and $B$ as well as the analogical relation between the source and the target side (geometric proportional analogies are covered in more detail in section 4.1).

The logical framework of HDTP in a certain sense resembles the algebraic setting. Domains are represented by axiomatizations, i.e. by finite sets of sentences of a logical language. Each axiomatization spans a theory, i.e. an infinite set of sentences, that can be inferred from the axioms by logical derivation. A different selection of axioms, that span the same theory (or some subtheory of it), may be seen as an alternative representation of the domain. As the analogical relation computed by HDTP is based on a mapping of the provided axioms, such a re-representation may lead to other mappings and an improved analogy.

An obvious problem of this approach is that the space of possible re-representations is infinite and hence without further guidelines not useful in practical applications. In [Publication 2] we propose an algorithm that applies some simple heuristics to restrict the search space. It works on logical formulae given in clause form and only allows these clauses to become smaller by refuting some of their literals within the given domain theory. Hence, only a small number of additional formulae are generated as candidates for the analogical mapping, i.e. re-representation does not dramatically increase complexity with respect to the number of formulae to be generalized.

The algorithm given in [Publication 2] is not the only way to address the problem. Another approach is, once some initial mapping has been established, to use that mapping to transfer unmatched formulae from one domain to the other and then apply an automatic theorem prover to check if the transferred formulae lie within the respective theory. If so, it can be included in the generalized theory.

A critical point of this kind of re-representation is that it depends on logical derivation and hence requires automated prove procedures. This may be regarded as problematic due to the semi-decidability of the underlying logic. However, this is not necessarily a knock-out argument, as one can always cut proofs after a specified length. This may exclude some possibly interesting mappings but it allows the overal system to finish with some analogy within a limited time. By increasing the maximal proof length, a system may systematically explore larger parts of a domain theory and thereby eventually discover novel and better mappings.

These logical approaches to re-representation operate on the syntactic level. They allow to describe a given situation by other sentences while still relying on the same words with the same meaning, i.e. the semantics is not touched. It may be argued by proponents of a stronger notion of re-representation that this is not enough, but that in fact what has to be changed when establishing an analogical mapping is the underlying model. This line of thought has not been further investigated yet. Up to now, in HDTP the only non-conservative adaptation of a domain theory can occur by analogical transfer, when sentences from the source side are projected into the target theory. Those transfers then have to be sanctioned by additional means, as explained in Gust et al. (2006).

To conclude, it can be stated that the problem of re-representation is a hard one that adds complexity to an analogy system. Wareham et al. (2012) investigate the computational complexity of the re-representation processes proposed for SME and show them to be computationally complex if not intractable. However, they point out that by restricting certain parameters the situation can be improved. This holds similarly for HDTP. It seems to be a good strategy to design a system in a fashion that it can act well on simple analogies, e.g. on those for which the domains are given in parallel, and that in addition is able to spent more effort to adapt the representation, whenever that is desired. This seems to be on par with humans, who can easily grasp situations where the analogous structures are presented in an obvious way, but who usually need a long time to discover analogies which require some adaptation of their conceptual system. Nevertheless, such hard analogies are often the key to new insights and developments, and once discovered they can be communicated and preserved within a society.

## 3.3 Semantic Aspects of HDTP

A point that is somewhat unsatisfactory from a semantic perspective is the logical status of the "variables" introduced by anti-unification. They are problematic in at least two respects: first, due to the use of higher-order anti-unification, the resulting generalized formulae may no longer be first-order, leaving the realm marked by the input theories. Second, as anti-unification introduces fresh variables, they will occur free in the generalized theory. This poses the question of how these variables should be handled semantically. An implicit universal quantification seems odd as it would exclude models of the domain theories to act as models for the generalization. An existential reading seems more suitable, and it comes close to the idea to consider the "variables" as new symbols for the generalized signature, as proposed in [Publication 3].

This interpretation comes fits to the intuition of analogies as a multi-domain mechanism and a more natural formulation of it can be reached using the notion of signature morphisms. A signature morphism $\Sigma_1 \to \Sigma_2$ maps symbols (constants, function, and predicate symbols) from a signature $\Sigma_1$ to symbols of another signature $\Sigma_2$, preserving their arity and types. It thereby induces mappings on the syntactic and semantic level: a $\Sigma_1$-expression is translated into a $\Sigma_2$-expression, and an interpretation for $\Sigma_2$ can be used to construct an interpretation for $\Sigma_1$, in a way that is compatible with logical satisfaction. This observation is the core of the institution theory (Goguen and Burstall, 1984), treated in more detail in section 3.4.

Using this terminology, one can describe the idea of HDTP by saying that it computes a set of axioms stated in a generalized language, i.e. a language formed over a generalized signature, together with a pair of signature morphisms, that translate the generalized language into the languages of the source and target domain, respectively. This construction is performed in a way that generalized axioms are translated into axioms of the original domains, or at least into formulae that can be derived from these axioms, when re-representation is taken into account. In the terminology of Goguen and Burstall (1984) this amounts to saying that these signature morphisms are theory morphisms. Stated compactly, what HDTP does is computing for given input theories $T_1$ and $T_2$ a *correspondence*[4], i.e. a diagram of the form

$$\begin{tikzcd} & T_G \arrow[dl, "\sigma_S"'] \arrow[dr, "\sigma_T"] & \\ T_S & & T_T \end{tikzcd}$$

---

[4] also sometimes called *span* or *roof*

in the category of first-order logic theory morphism. This resembles figure 3.1 sketching the idea of analogy by generalization, however, now having a clearly defined formal meaning.

So, instead of considering the domain theories as syntactic instantiations of the generalized theory described by substitutions, they should rather be seen as independent theories, into which the generalized theory can be projected using a suitable theory morphism. This perspective has the advantage of coming with a clear description of what happens on the semantic side and it underlies the analysis of the semantic aspects of HDTP in [Publication 5]. However, the above description is somewhat simplified, as it neglects the fact, that a substitution can replace a variable by arbitrary terms and restricted higher-order substitutions can establish even more complex mappings. This is not captured by the notion of a signature morphism, that by definition just maps symbols of signatures. This problem is addressed in [Publication 5], similarly to Gust et al. (2006), by extending the domain signatures with additional symbols and the domain theories by complex definitions for these symbols. It is then shown, that these definitions do not actually affect the semantic side.

This somewhat cumbersome treatment can be simplified using the concepts of *derived signature* and *derived signature morphism* (Goguen et al., 1978): a derived signature extends the underlying base signature by adding complex terms as new (atomic) symbols. The crucial point is that these new symbols do not change the underlying logic, i.e. sentences over the derived signature can be simplified to sentences over the original signature and a model for the original signature already provides a unique interpretation for the derived symbols, in a way that logical satisfaction is preserved. A standard derivation $\Sigma^\delta$ for a first-order signature $\Sigma$ (mentioned in Diaconescu, 2008; Sannella and Tarlecki, 2012) can be informally summarized as follows:

1. $n$-ary function symbols of the derived signature are $\lambda$-terms of the form $\lambda x_1 \dots x_n.t$ with $t$ being a $\Sigma$-term from $T_\Sigma(\{x_1, \dots, x_n\})$.

2. An $n$-ary function symbol $f$ from $\Sigma$ can be embedded into the derived signature as $\lambda x_1 \dots x_n.f(x_1, \dots, x_n)$.

3. Sentences over the derived signature can be translated to $\Sigma$-sentences by $\beta$-reduction.

4. Given a $\Sigma$-interpretation $I$, a derived symbol $\lambda x_1 \dots x_n.t$ is understood as a function that maps $(a_1, \dots, a_n)$ to $I(t)[x_i \mapsto a_i]$.

Now a derived signature morphism from $\Sigma_1 \to \Sigma_2$ is simply defined as an ordinary signature morphism from $\Sigma_1 \to \Sigma_2^\delta$. The important point is that by the properties of a derivation $\Sigma_1$-sentences can be translated into $\Sigma_2$-sentences along a derived signature morphism and $\Sigma_2$-models can be reduced to $\Sigma_1$-models in a way that respects logical satisfaction, just as it is the case for ordinary signature morphisms.

It is easy to see, that for the basic substitutions of HDTP introduced in section 3.1, one can provide corresponding derived signature morphisms that induce the same mapping on the level of formulae as an application of the respective substitution would do. Hence, the above formal description can be

$$\begin{tikzcd}
\mathbf{Sen}(\Sigma_1) \arrow[r, "\mathbf{Sen}(\sigma)"] \arrow[d, dash, "\models_{\Sigma_1}"] & \mathbf{Sen}(\Sigma_2) \arrow[d, dash, "\models_{\Sigma_2}"] \\
\mathbf{Mod}(\Sigma_1) & \mathbf{Mod}(\Sigma_2) \arrow[l, "\mathbf{Mod}(\sigma)"']
\end{tikzcd}$$

Figure 3.5: A corridor from a room $(\mathbf{Sen}(\Sigma_1), \mathbf{Mod}(\Sigma_1), \models_{\Sigma_1})$ to another room $(\mathbf{Sen}(\Sigma_2), \mathbf{Mod}(\Sigma_2), \models_{\Sigma_2})$ consists of a pair of mappings $(\mathbf{Sen}(\sigma), \mathbf{Mod}(\sigma))$ fulfilling the satisfaction condition: for every sentence $\phi \in \mathbf{Sen}(\Sigma_1)$ and every model $m \in \mathbf{Mod}(\Sigma_2)$ it holds $\mathbf{Mod}(\sigma)(m) \models_{\Sigma_1} \phi$ iff $m \models_{\Sigma_2} \mathbf{Sen}(\sigma)(\phi)$

.

adapted, summarizing the operation of HDTP as the construction of a *correspondence* in the category of first-order theories and derived theory morphisms for a pair of given input theories.

## 3.4 An Institutional Perspective on HDTP

Although the framework of HDTP is spelled out for first-order logic, many ideas seem to be applicable independent of this specific choice of logic, or at least may be reformulated for other formal languages. This raises the question, whether the central ideas of HDTP can be described in an abstract way that would allow to easily apply it to a given formalism. Such an abstract formulation may have other advantages as well, as it may shed light on the central notions without referring to technical peculiarities of a given formalism (e.g. the special form of terms in first-order logic). An abstract description may in fact be simpler and easier to understand.

The theory of institutions has been introduced by Goguen and Burstall (1984, 1985) to formalize the intuitive notion of a logical system. The theory uses the language of category theory (e.g. MacLane, 1998) that facilitates a quite abstract description of many ideas, while still being clear and precise, allowing strict formal statements and proofs. It has been shown that many results from various types of logic can be stated and proven on this abstract level of institutions (e.g. Diaconescu, 2008).

At the core of many logical systems are the interrelated notions of (syntactic) sentence, (semantic) model and satisfaction. Intuitively, a set of sentences determines via the satisfaction relation a class of models, in which these sentences hold, and vice versa, a class of models determines a set of sentences that are satisfied in these models. Goguen and Burstall (1985) call this situation a *room*. Of special insterest to institution theory is the possibility to move from one room to another, thereby respecting the satisfaction relation, an idea formalized in the notion of a corridor, depicted in figure 3.5. Together, rooms and corridors form a category (details are provided in appendix A).

In institution theory, rooms are based on signatures in a systematic way. The intuition of a signature is to provide a vocabulary, from which a set of sentences is constructed and which has to be interpreted to get a model. The basic axiom for institutions is that truth (i.e., satisfaction) is invariant under change of notation. This is expressed by defining an institution to be a functor from the category of signatures and signature morphisms to the category of rooms and corridors. One can (with some effort) check that first-order logic with its traditional syntax and semantics forms an institution. But also other logics, like propositional logic, horn clause logic, equational logic, second-order logic, modal logic, intuitionistic logic, and many more can be shown to be institutions (Diaconescu, 2008).

The theory of institutions seems especially suitable for HDTP, as many of the central notions, like signatures and signature morphisms, axioms (presentations) and theories, as well as the distinction of a syntactic and a semantic level, are available in this framework. [Publication 8] introduces the central notions of HDTP, like generalization, analogy, and coverage using the general language of institutions, showing that this leads to sensible notions on the syntactic and semantic level. It is further shown that the classical, first-order HDTP framework can actually be seen as an instance of this general formulation.

A point that is left open is a proper treatment of derived signature morphisms on the institutional level. Although the notion of signature morphism is central to the theory of institutions, the notion of a derived signature morphism has so far been neglected. Krumnack et al. (2014) try to close this gap by proposing three ways to abstractly define derived signature morphisms in an institution-independent way. The variant that makes use of morphisms in a Kleisli-category over a suitable institutional monad is spelled out in appendix A. It seems especially appropriate in the context of HDTP, as it guarantees that such morphisms induce a corridor, i.e. a sentence translation mapping with an associated model reduction. The derived signature morphisms for first-order logic discussed above are one example, but one could also use other mappings, e.g. mapping predicate symbols to complex formulae. Also in other institutions, derivations of this type can be found. Using this definition allows the $\Sigma$-substitutions, that are used in [Publication 8] to define the notion of generalization, to be replaced by the notion of abstract derived signature morphism. In that way, the abstract formulation of HDTP becomes almost literally identical to the first-order version, just replacing first-order notions by their institution-independent counterpart.

## 3.5 Conclusion

Different extensions and refinements to the formal framework of HDTP, originally introduced by Gust et al. (2006), have been developed. Some points that remained sketchy there have been worked out and put on solid theoretical grounds, showing the original intuitions to be justified. The notion of generalization is clarified by introducing a form of restricted higher-order anti-unification. An approach towards re-representation by deduction is proposed that fits well into the overall logical framework. It is argued that the best way to understand the mappings between the domains is as a translation between signatures, i.e. signature morphisms, that induce coherent mappings on the syntactic and semantic level. The approach can be extended to derived signature morphisms to account for more complex mappings. The theory of institutions is proposed as a suitable formalism to describe the overall framework, leading to a clear description and allowing to generalize the ideas to other logical systems besides first-order predicate logic.

The focus of this work is the analysis of analogical mappings in a logical framework while the second step, the analogical transfer, is not further investigated here. Nevertheless, the logical presentation suggests that there are different types of analogical transfer, depending on how the target domain is extended. Already the simplest form, that just transfers formulae from the source to the target theory, may well have semantic impacts by invalidating certain models. More complicated scenarios arise, when new symbols are introduced into the target signature. These new symbols have to be accounted for on the semantic level either by interpreting them using objects already available, or by extending the underlying universe. A general analysis of analogical transfer may profit from such considerations.

# Chapter 4

# Applications of HDTP

The formal framework of HDTP, as presented in chapter 3, has been developed with applications from cognitive science in mind. As argued in chapter 2, analogy can be understood as a core mechanism of cognition with the potential to explain various cognitive abilities in different fields. This chapter intends to substantiate this claim by presenting a row of studies that ground a wide range of capabilities on analogy making, using HDTP to provide formel models.

Already [Publication 3] shows how HDTP can be used to formalize some well-known analogies that have previously served to demonstrate principles of analogical reasoning in different analogy models, including the Rutherford analogy, the heatflow analogy, and a comparison between different political systems, among others. In the following sections I will describe further studies that focus on less classical applications. It should be noted that all these works are exploratory in nature. They do not aim at providing an optimized system for a particular task. Neither it is claimed that the particular way of modeling is the only or even the best way of adressing the problems. These works should be understood as proofs of concept showing the suitability of the framework in different contexts.

## 4.1 Geometric Proportional Analogies

A proportional analogy is a pattern of the form $(A : B) :: (C : D)$, consisting of four elements $A$, $B$, $C$, and $D$ with the property that $A$ relates to $B$ in the same (or in an analogous) way as $C$ relates to $D$. In contrast to other types of analogies, which are usually characterized as cross domain mappings, the source and target elements of a proportional analogy are typically from the same domain. Being described already by Aristotle, proportional analogies have a long tradition, originally considering numerical relations, but later also focusing on other domains like concepts or geometric figures. Since Spearman (1923) has proposed proportional analogy problems – in which subjects have to reason about a suitable $D$ for given $A$, $B$, and $C$ – as one possible method to assess cognitive abilities, they have often been used for designing tests. Especially geometric problems have become popular, as they do not presuppose language, reading or writing skills and hence are also suitable for children and illiterates (and computers). A well-known example is Raven's Progressive Matrices Test

(1936) that – although not being strictly of the form of a proportional analogy problem – requires similar analysis and mapping strategies.

Due to their alleged association with intelligence, such geometric analogy problems have attracted the interest of AI researches quite early. Evans (1962) develops a system named ANALOGY that is able to solve geometric analogy problems which have been in use in the 1942 and 1943 American Council on Education Examination (ACE) test. Geometric analogy problems seem attractive for AI research, as they are quite formal in nature and can be considered as a closed-domain problem that does not rely on a huge amount of background knowledge. A core element for every system addressing this kind of problem is a suitable representation mechanism for geometric figures. In the context of analogy it is important to have a structured representation, i.e. a formalism that does not just list dots and lines but accounts for geometric shapes, spatial, metric, and topological relations as well as transformations like mirroring, stretching, and rotation.

Such a structured representation seems also plausible from a cognitive point of view. Following Gestalt psychology (Wertheimer, 1912; Köhler, 1929; Koffka, 1935), human perception is holistic, resulting in a consistent and meaningful overall picture of the environment. Unstructured visual input is transformed into a structured representation of coherent shapes and patterns being formed according to some Gestalt principles. The laws for grouping elements, including principles of proximity, similarity, symmetry, and good continuation, guarantee the formation of a good Gestalt, that is as regular, simplistic, ordered, and symmetric as possible. It should be noted that these principles may interfere, and hence perception my vary among people and also depending on the context, due to application of different Gestalt principles. In the case of geometric proportional analogies, the perception of an ambiguous figure may be constrained by the context provided by the other figures of the problem.

In contrast to classical geometric analogy problems, which usually have one preferred ("correct") solution, Schwering et al. (2009a) describe an experiment focusing on ambigous problems, which allows for different plausible outcomes. In the experiment, geometric proportional analogy problems are presented to human subjects who are asked to construct geometric figures as solutions. An advantage of this construction method is that subjects are not constrained in their reasoning process by considering pre-defined solutions. In addition, this setup is intended to inspire subjects to reflect explicitly about the transformations to be applied, hence suggesting the use of relational structure instead of superficial features. The experiment has shown that indeed different plausible solutions are constructed by the test subjects, which can be explained by different Gestalt principles taking effect. In addition, the context of the other figures in a given analogy problem has an impact on the structuring of a figure.

Already Schwering et al. (2007) apply HDTP to ambiguous geometric analogy problems, and show the advantage that a logic-based representation has in such a scenario, allowing to consider different structures for the figures by applying re-representation. This work is refined in [Publication 4], which presents a formalization of the analogies used in the aforementioned experiment. It uses a logical version of VREG, a language of perception developed by Dastani (1998), that allows to express different Gestalt perceptions for a visual stimulus. Gestalts representations can be computed from 2-dimensional geometric figures by triggering different Gestalt principles allowing HDTP to propose analogical inferences at different levels of cognitive plausibility.

## 4.2 Theory Blending

The idea of conceptual blending, also known as conceptual integration, has been studied in cognitive linguistics, most promiently by Fauconnier and Turner (1998, 2002). Broadly speaking, blending allows the construction of new ideas based on the combination of already established knowledge. An often quoted example is the combination of the concepts HOUSE and BOAT to create conceptual blends like HOUSEBOAT and BOATHOUSE among others. However, Fauconnier and Turner (2002) do not restrict the topic to such simple examples, but argue that conceptual blending is a fundamental mechanism, which underlies several complex high-level cognitive capabilities, supporting their claim by collecting evidence from different fields.

Conceptual blending resembles analogy in that it is a multi-domain mechanism, but it goes beyond analogies, as the blending process not only involves the relation between two given domains and the mapping of knowledge from one domain to the other, but it rather establishes an entirely new domain.[1] This new domain, often called *blend space* or simply *blend*, maintains partial structures from both inputs and may add an emerging structure of its own. It is supposed to coexist with the original domains: one still wants to maintain the concepts of HOUSE and BOAT, even when one is aware that a BOATHOUSE has some relation to both HOUSEs and BOATs. If one domain is updated, this will also be reflected in the blend spaces based on that domain.

Already the HOUSEBOAT/BOATHOUSE example demonstrates that multiple blends can exist for a given pair of input spaces.[2] Nevertheless, it is still an open question what criteria should be used to assess the quality of a blend, or how to compare two different blend candidates. Fauconnier and Turner (2002) have formulated various competing optimality principles.

Although the assumption about the importance of blending mechanisms within human cognition suggests that an implementable formalization of conceptual blending could trigger significant development, especially on human-level artificial intelligence, only few approach to such systems exist. An early computational model named Sapper is described by Veale and O'Donoghue (2000). Another approach is described by Lee and Barnden (2001) who present a rule-based system for counterfactual reasoning in natural language. Pereira (2007) presents a system that computes conceptual blends, employing a parallel search engine based on genetic algorithms. In Pereira's account, mappings do not have to rely on similarity, as they can also present conflicts that are striking, surprising, or even incongruous.

Similar to analogies, the development of a formal framework for blending heavily depends on the formalism choosen for knowledge representation. In addition, the formalism also influences how the optimality principles which guide the creation of sensible blend spaces and which help to distinguish good blends

---

[1] The original publications speak of "concepts", "conceptual domains", or "conceptual spaces" rather than of "domains", the expression choosen here to emphasize that what is meant is essentially the same as domains in the context of analogies.

[2] Goguen and Harrell (2010) present some additional, less obvious blends that can be constructed from HOUSE and BOAT.

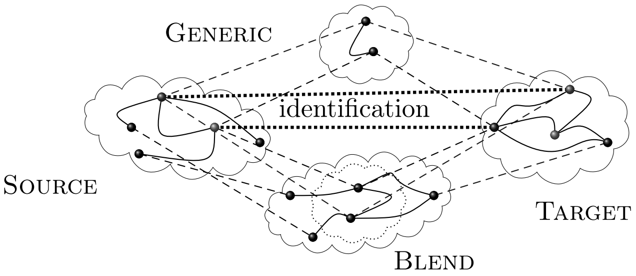

Figure 4.1: The four-space model of conceptual blending: corresponding parts of the Source and Target concepts are identified, defining a Generic space and a Blend (from [Publication 9])

.

from bad ones can be realized (Pereira and Cardoso, 2003). An early formal approach towards blending – that is of special interest in the context of HDTP – has been put forward by Goguen (2006). More recently, also Kutz et al. (2015) sketch a formal account to blend logical theories within a distributed ontology setup.

It has often been remarked that there seems to be a close relation between blending and analogical mapping. When combining two conceptual spaces, elements and structures from the two input spaces are identified in the blend space. This identification gives rise to a mapping between elements of the input spaces that resembles an analogical mapping, constrained by principles like systematicity. Turning perspective, one can consider such a mapping as a starting point to establish a blend space by identifying mapped structures from the input spaces to form the core of the blend into which then additional parts from the input spaces can be imported when appropriate. This leads to the four-space model of conceptual blending depicted in figure 4.1.

There are two extreme cases of blends: when the analogical mapping is empty, no identification takes place, and the blend space is basically the disjoint union of the two input spaces. The other extreme case is given by two isomorphic input domains, that are completely covered by the mapping. In this case, the blend is isomorphic to both of the input spaces. Analogical inferences can in this setting be interpreted as a special kind of blend: all of the target domain is included in the core blend, while unmapped parts of the source domain can be added as analogical inferences.

The foure-space view opens a way to use HDTP for blending as described for example in [Publication 6]. Starting with two input domains, an identification is computed as an analogical mapping via the usual generalization procedure. Then, in a second step, instead of mapping knowledge from the source to the target domain, a blend is constructed, i.e. a new logical theory, consisting of the axioms covered by the analogy and potentially also some uncovered axioms from both sides. When considering logical theories to represent the input domains, the analogical mapping is not the only hard constraint that restricts the set of possible blends. Logical consistency is a further requirement to ensure that blends can be given a sensible interpretation. This is emphasized by using the notion *theory blending.*

In [Publication 11] we present a quite general analysis of theory blending.

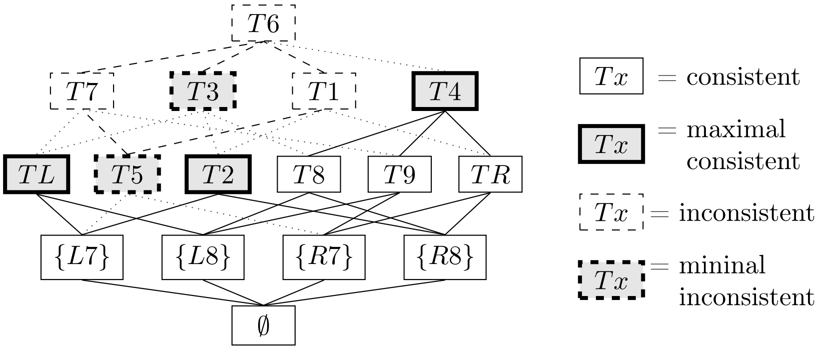

Figure 4.2: from [Publication 11]: a lattice of 'blend candidates', defined by different axioms to be included into the blended theory. The lattice is based on an example presented in that paper.

For a given analogical relation, the set of possible blends can be considered as a lattice, formed by the subsets of the maximal blend, i.e. the one that integrates all unmatched axioms from the source and target domain. This maximal blend may, however, be inconsistent and hence of little interest. The consistent blends will form a lower set in this lattice, that can be described by its maximal elements, forming a maximal antichain, i.e. a subset of incomparable blends (figure 4.2). [Publication 11] contains an algorithm that can effectively compute these maximal consistent blends, assuming a theorem prover that is able to check consistency for blend candidates. The algorithm tries to minimize the number of consistency checks, because those are considered to be the most expensive operations.[3] The ideas are demonstrated by blending theories describing a discrete and a continous concept of numbers.

In this approach, there may exist multiple maximal consistent blends based on the same analogical mapping. Starting with another mapping can result in a further set of maximal consistent blends. As demonstrated in [Publication 11], an analogical mapping can also be relaxed by dropping some identifications, if the resulting blends are not satisfactory. This may allow to incorporate additional axioms into the blend.

## 4.3 Theory Blending for Noun-Noun Composition

The literature on blending prominently features noun-noun compounds, such as the previously mentioned word pair HOUSEBOAT and BOATHOUSE, as examples of conceptual blends. Although the linguistic details of such compounds may differ slightly within a language, or between different languages, the rules to interpret such compounds seem quite universal, indicating that they are based on the same cognitive mechanisms. Many noun-noun compounds are lexicalized, having an established meaning, probably including some understanding

---

[3] In certain situations, the method for theory combination by Nelson and Otten (1979) may by applied to find a decision procedure for a blend space. However, in general such a method will not be applicable due to the non-trivial nature of associations in the blended theory.

on how they relate to the meanings of their components. A formal framework for conceptual blending would allow to adequately represent their semantics in computational dictionaries. But humans can also produce and understand new noun-noun compounds on the fly, and their interpretation then often depends on the current context. A plausible computational model of human language skills should account for these observations as well.

As many noun-noun compounds obviously combine the underlying concepts, blending seems an appropriate mechanism for tackling this phenomenon. However, the linguistic nature of noun-noun compounds is usually not as symmetric as the general blending framework presented above would suggest. In fact, many such compounds consist of a modifier and a head, the latter determining the category of the newly created concept, which has often led to analyze them as some kind of function application. A boathouse is a kind of house and a houseboat a kind of boat (even though it may also be considered as a kind of house). For an analysis based on blending, this leads to a more analogical treatment which takes the target domain (head) as the core of the blend, into which features from source (modifier) can be intergrated.

[Publication 9] presents a computational model for the interpretation of noun-noun compounds that is based on HDTP and roughly follows the two step approach described above. Concepts are formalized in a frame-like style, providing first-order axioms used as input domains. Slots can be annotated with an entrenchment value which reflects the saliency of the property based on general knowledge and the current context. In the mapping phase, structural commonalities are explored using the standard analogical generalization, without taking entrenchment values into account. In the second step, when the blend is constructed, the approach does not simply try to find a maximally consistent blend, but instead considers the specificities of noun-noun compounds. Starting with a virtual copy of the head noun, particular traits of the modifier are added. In case of conflict, the entrenchment values can be used to decide about the inclusion into the blend. The idea is exemplified using the compound SNAKE GLASS, that has been used by Wisniewskia and Gentner (1991) in an empirical study on the combinatorial semantics of noun pairs, where many human subjects describe it as a “tall, very thin drinking glass”. Here, features like a longish shape for SNAKE or transparency for GLASS of nouns may be highly salient and hence enter the blend.

On the positive side, this approach shows that HDTP can be used as a framework to study noun-noun compounds as blending phenomena. The approach is also pluralistic in the sense that it allows for different blends, depending on the choice of entrenchment values. This is considered desirable as it can explain experiments, showing that humans do not always agree on the meaning of a given noun-noun compound. On the other side it has to be mentioned that this study is only a first step, being exploratory, using a quite ad hoc representation of concepts that surely deserves a deeper treatment. It also shows that – despite the logical framework of HDTP being capable of modeling blending for noun-noun compounds – on the algorithmic side there seems to be the need for a specific type of blending process, adapted to this particular setting, an observation that will recurr for other applications of theory blending described in the following sections.

## 4.4 Counterfactual Conditionals Based on Theory Blending

Another application of the blending framework is the analysis of counterfactual reasoning. A counterfactual conditional (CFC) is a conditional statement of the form "if $A$ then $C$" characterized by the property that the antecedence $A$, and usually also the conclusion $C$, are assumed to be counterfactual (i.e. false). An often used example from Quine (1960) is the following statement:

(CFC-1) "If Julius Caesar was in command during the Korean war, then he would have used the atomic bomb."

An analysis of such a conditional based on classical logic would come to the result that is it trivially true, due to the principle *ex falso quodlibet*. In practice however, people can argue about such conditionals and judge them to be more or less plausible. A statement like (CFC-1) could also be rejected and countered by another one like:

(CFC-2) "If Julius Caesar was in command during the Korean war, then he would have used the catapult."

Counterfactual conditionals can be found in diverse fields like learning, moral judgement, or decision making under risk and uncertainty and have for a long time attracted attention (e.g. Ramsey, 1929; Goodman, 1947). In philosophy, probably the most influential treatment goes back to Lewis (1973), who proposes to apply a similarity measure between possible worlds. A counterfactual argument is more plausible, if the world in which it holds is closer to the real world. More recently, cognitive scientists have explored human behaviour in dealing with counterfactuals, like Byrne (2005), who studies the mental creation of alternatives to reality and, in particular, the verification of CFCs. Pearl (2011) presents an algorithmic approach to assess CFCs, which systematically adapts a symbolic model of the factual world in a way that the counterfactual antecedence and conclusion hold in the modified model. The operations required for adjustment are used as a base to evaluate the plausibility of the CFC.

Human proficiency in counterfactual argumentation seems to require no special training and hence one may conjecture that it makes use of basic cognitive mechanisms. In the tradition of cognitive linguistics it has been argued that conceptual blending can be used for the construction of a counterfactual world, in which reasoning is performed (e.g. Coulson, 2001). Such a blend combines knowledge about the factual world with some counterfactual idea. Lee and Barnden (2001) have taken up this idea and present a system that can create a blend space to represent and analyze CFCs. This system, originally designed for uncertain metaphor-based reasoning, provides mechanisms for conservative conflict resolution. In [Publication 10] we apply the idea to use blending for assessing CFCs in the context of HDTP, however providing a different algorithm to create the blend space from an analogical mapping. The form of a CFC gives some guidance to decide what knowledge should enter a counterfactual blend candidate (CFB), expressed by four principles:

(P1) Counterfactuality: A CFB should satisfy the antecedent of the given CFC.

(P2) Choice: For every matching pair, one alternative is allowed to be imported into a CFB.

(P3) Consistency: A CFB should sustain (logical) consistency.

(P4) Maximality: A CFB should contain as many imported instances of the original axioms as possible.

Using the examples (CFC-1) and (CFC-2), in [Publication 10] we demonstrate how these principles can control the blending process to give meaningful explanations for both of them. This blending approach seems especially obvious for metaphorical CFCs, i.e. situations where counterfactuality is caused by combining two concepts from different parts of the factual word, but other cases may be handled as well. It provides a way to introduce conterfactual reasoning into the context of artificial cognitive systems, a field where it has often been neglected. In [Publication 10] we argue that this lack of interest seems inappropriate in view of the wide use of counterfactual arguments in human thinking and propose to add CFCs as a benchmark problem for comparing and evaluating such systems.

## 4.5 Analogies and Blending in Mathematics

Mathematics is traditionally considered as a deductive discipline, presenting its results in form of proofs following the strict rules for inference prescribed by classical logic, leaving little space for non-classical reasoning mechanisms. But even though rarely studied systematically, analogies are not unusual in mathematics. A simple, relatively common example are dual situations, where a statement or proof is formulated only for one case with the additional remark that the other case is “analogous”, meaning that it can be achieved by repeating the original argument and replacing signs, adding transpositions, or reverting arrows in a systematic way, wherever appropriate.

Analogous situations may also motivate the development of a mathematical description of the analogical relation involved. A prime example are the categorical notions of *functor* and *adjoint*, that evolved from the investigation of corresponding concepts and constructions in algebra and topology. A functor postulates the perservation of structure, an adjoint indicates an even closer relationship between two categories. The recognition of analogous situations may also be a basis for the creation of more abstract concepts, i.e. the abstract core hidden in the analogical mapping, is made explicit in form of a generalized theory. Many modern algebraic concepts, like groups, rings, and categories can be considered results of such a process. The advantage of such general notions is at least twofold: proofs can be carried out at an abstract level and will then apply automatically for all more basic cases instantiating the general theory. And abstract proofs may even be simpler, as they are not blurred by the pecularities of a specific domain, and thereby allow to exhibit the structure of an argument more clearly (on the other hand, they can by harder to find, as working with abstract concepts may require additional training).

But the analysis of the role of analogies in mathematics should not be restricted to those evidences that can be found in the published body of mathematical knowledge. There is good reason to believe that behind the scenes analogies play a major role in the formation of mathematical concepts, conjectures, and initial proof attempts. Polya (1954) provides a rich set of examples from the history of mathematics as well as guidelines for finding plausible mathematical arguments and proofs based on analogy. Also blending processes seem not to be uncommon in mathematics. Obvious cases are concepts like topological groups, which combine the axioms of a topological space and an algebraic group to gain a new structure. However, blending may occur in other forms as well. Fauconnier and Turner (2002) discuss two cases: the indirect proof as a blend of positive knowledge and the negated hypothesis, from which then a contradiction has to be derived. The main example however is the development of different number systems which is described as a series of recursive blends. Alexander (2011) presents a formal treatment of these constructions, describing blending as a kind of pairing followed by a quotient construction: negative numbers can be constructed from pairs of positive numbers, identifying thus pairs which have the same difference. A more common example are the rational numbers that are constructed as pairs of integers (fractions), identifying those with identical ratio. In the same manner, other domains like reals and imaginary numbers can be constructed using this blending procedure. To introduce operations for a blend space, these have to be defined for the pairs in a way compatible with the quotient mapping.

Another example, with major historical impact, is the establishment of the complex plane by Argand (1813). Although calculations with imaginary entities have been carried out before, Argand's discovery, that the algebraic operations can be associated with a geometric meaning, is considered a major breakthrough by which these entities have been accepted as proper numbers, inspiring further developments. Martinez et al. (2011) present a formal analysis of Argand's discovery, that we elaborate further in [Publication 7]. It is described as a multi-domain reasoning process, comparing algebraic and geometric notions. In this process, analogical mappings are established, e.g. between the sign of a number and the orientation of a line segment, and based on these a blend space is constructed that assigns algebraic and geometric properties to its elements. This idea is spelled out giving axiomatizations for the algebraic and geometric domains in first-order logic and then applying HDTP to compute the analogical relations. Based on these a suitable quotient space is constructed, leading to the desired blend.

Lakoff and Núñez (2000) propose a more fundamental understanding of the role of analogies in mathematical cognition. They argue that mathematical concepts are related by "metaphorical" mappings[4] to more basic ideas, thereby establishing a hierarchical network of mathematical domains. At the bottom level, mathematical notions are grounded in bodily experience. They exemplify this grounding for the case of basic arithmetic of natural numbers by providing four "grounding metaphors" that relate numbers and operations on them to everyday experiences. The OBJECT COLLECTION metaphor relates a number to a collection of (physical) objects of the respective cardinality. Addition and substraction of numbers corresponds to adding or removing objects from the collection. The OBJECT CONSTRUCTION metaphor considers objects that are constructed from smaller parts. The MEASURING STICK metaphor interprets numbers with the help of lines created by putting equally long sticks

---

[4] The "metaphors" from cognitive linguistics can in our context be understood as analogies.

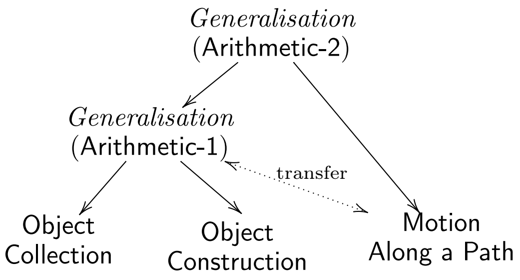

Figure 4.3: Generalization of grouding domains (from [Publication 7])

end-to-end. Numerical relations like greater and less map to physical relations of longer and shorter lines. MOTION ALONG A PATH relates the experience of moving forward or backward on some path to the arithmetic notions of addition and substraction. Lakoff and Núñez base their analysis on linguistic arguments, pointing out that the vocabulary used when talking about arithmetic is actually metaphorical, getting a literal and in the end embodied meaning when interpreted along a suitable grounding metaphor. Although all four metaphors ground basic arithmetic notions, each metaphor brings in different aspects.

In a series of studies we have taken up this idea with the aim of creating a formal model of these kinds of metaphorical relations between mathematical domains. Such an effort, if successful, would have multiple benefits: it would provide a way to deal with the symbol grounding problem, allowing to couple symbols in a formal system with sensorimotor contingencies, allowing to investigate abstract (algebraic) reasoning by specific (physical) actions. Furthermore, if Lakoff and Núñez's 2000 claim is correct and mathematical ideas are organized via metaphorical relations, such a framework would allow to model these structures by making the relations explicit using a precise language, helping to understand the interdependence and development of mathematical ideas. In the end, such a framework could provide a way to support the creation of mathematical ideas and, interoperating with automatic theorem provers, foster mathematical research.

Guhe et al. (2009a,b) describe how the grounding domains of Lakoff and Núñez (2000) can be linked to basic arithmetic notions using the infomorphisms of the Information Flow theory (Barwise and Seligman, 1997), which are employed to formalize metaphorical relations. Guhe et al. (2010) continue this work, now applying HDTP to relate and generalize the grounding domains, to compute an abstract domain of arithmetics. Here, it is observed, that repeated generalization will result in only more and more abstract theories, as aspects not found in all grounding domains will be removed from the generalization. To overcome this problem, intermediate theories of arithmetic can be enriched by analogical transfer from other ground domains (see figure 4.3). For example, the origin in the MOTION ALONG A PATH has no corresponding concept in the generalized Arithmetic-1 domain. This can trigger the introduction of an analogon for origin, i.e. a concept behaving as 0, into Arithmetic-1.

[Publication 6] develops this idea further, stating that metaphorical relations and generalizations are not sufficient to describe the formation of new mathematical domains, but that some form of blending should be considered as well. We describe a three-step process for mathematical discovery: first, in the *exploration* phase, different domains are compared, leading to analogical mappings and generalizations. This can be realized employing the analogical mapping of HDTP and it is examplified in the paper using the four grounding domains of Lakoff and Núñez (2000). Here, it is essential that the domains are not fully isomorphic so that the mapping is not complete but leaves some domain specific pecularities. In the second step, called *recognition and goal formulations*, these mappings are analyzed, and especially the differences between domains can inspire the formulations of goals like the introduction of new conecpts into a domain by analogical transfer, or the creation of a new domain by blending. In the following third step, actual *discoveries* can be achieved by working out the goals, i.e. the proposed domains are assembled so that different hypotheses can be analyzed. An example described in [Publication 6] concerns the questions, if different fractions in the OBJECT CONSTRUCTION domain which are mapped to the same point in the MOTION ALONG A PATH domain can actually be identified, thereby leading to a new notion of equality for fractions (a process that underlies the idea of rational numbers).

## 4.6 Conclusion

In addition to the application scenarios presented in this chapter, HDTP has been used in other domains as well. Schwering et al. (2009b) investigate the interpretation of visual metaphors in advertisements. Besold et al. (2013) describe applications of computational analogy engines in general and specifically HDTP for analyzing questions and creating models in education and teaching research. Besold (2013) considers the concept of rationality and presents an analogy-centered account and an architectural sketch for applying HDTP to explain certain aspects of rational behaviour. Furthermore, Abdel-Fattah et al. (2012) discuss the potential of HDTP in the context of computational creativity and Abdel-Fattah et al. (2013) propose to use HDTP for inductive learning.

In summary, HDTP can not only be applied in the context of classical analogy problems, but also seems suitable to adress a range of tasks from different fields in artificial intelligence. This diverse applicability is a strong support for the claim that analogies can function as a central cognitive mechanism. In addition, it also shows that a symbolic analogy framework can be employed to perform various forms of non-classical reasoning, paving the way to equip traditional, logic-based systems with additional capabilities.

However, while processes of analogical reasoning seem to be relatively well understood and stable, allowing the formal framework to be directly applied in various situations, the situation for blending is less clearcut. Starting a blending process from an analogical mapping is relatively widely applicable. However, the construction of the actual blend requires additional means. Already the examples presented in this chapter apply quite different strategies for the assembly of the blended domain. Consistency is an obvious precondition for logic-based systems, but in general it seems insufficient to identify interesting blends. While the principles listed by Fauconnier and Turner (2002) seem plausible, their informal nature gives much room for interpretation when developing formal systems, offering a great opportunity for further research.

# Chapter 5

# Epliogue

The framework of HDTP is a logic-based approach to compute analogical mappings and propose analogical inferences. Different aspects of HDTP have been analyzed, providing some refined notions and leading to a better understanding of the underlying logical processes. From this background it is argued, that many of the ideas of HDTP are not specific to first-order logic, but can be formulated on an abstract level, so that they become applicable in other settings as well. Following the understanding of analogy as a core cognitive mechanism, HDTP can serve as a basis to provide various higher-level capabilities.

As argued by Abdel-Fattah (2014), cross-domain cognitive mechanisms, like analogies and blending, should be central in the development of AGI systems, as they allow to address challenges like understanding newly introduced concepts by considering them as combinations of already-known ones, transferring ideas from one domain to another to solve problems there or to even actively develop new plans and ideas. HDTP provides the means to describe such processes in a logical setting, paving the way for a formal analysis and an algorithmic treatment. As logic is still the standard approach to knowledge representation, HDTP opens a way to combine well-established AI techniques with non-classical reasoning procedures. This substantiates the argument from (Gust et al., 2009) that logic is suitable as a “lingua franca for AGI systems”.

A central but often neglected problem, that has also not been adressed in this thesis, is the question of how domains are established. Although crucial for analogical reasoning in particular, and cross-domain methods in general, it is usually just assumed that domains are provided in a ready-made fashion. This is probably not the case in most existing large scale knowledge bases, consisting of a monolithic ontology or a long list of facts and rules. One possible answer to this question could be to create domains on the fly, by selecting subsets of a given knowledge base that are relevant in the current situation. Such an approach would allow to take context effects into account, as demonstrated in the DUAL system (Kokinov, 1994). Another possibility would be to establish a new paradigm for knowledge representation, that employs domains and inter-domain relations as organizational units. Such a knowledge structure would allow for different kinds of reasoning processes and would also support forms of learning that acquire new domains and relations based on analogies, generalization, and blending processes. HDTP is a promising base for starting such an endeavor.

# Appendix A

# Derived Signature Morphisms

This appendix presents an institution-independent approach to derived signature morphisms. This approach, that was developed in joint work with Till Mossakowski from the University of Magdeburg, is one of three different proposals contained in a yet unpublished paper, that has been worked out in succession of the work shop contribution Krumnack et al. (2014). I just give the formal definitions here, referring to the main text (section 3.4) for motivation and background information. I will introduce institutions via rooms and corridors (in the terminology of Goguen and Burstall, 1986), which capture the Tarskian notion of satisfaction of a sentence in a model, as this leads to a very compact development of the central ideas:

**Definition 1.** *A* room $\mathcal{R} = (S, \mathcal{M}, \models)$ *consists of*

- *a set of* $S$ *of* sentences*,*
- *a category* $\mathcal{M}$ *of* models*, and*
- *a binary relation* $\models \subseteq |\mathcal{M}| \times S$*, called the* satisfaction relation*.*

Then, morphisms between rooms are of course called corridors:

**Definition 2.** *A* corridor $(\alpha, \beta)\colon (S_1, \mathcal{M}_1, \models_1) \to (S_2, \mathcal{M}_2, \models_2)$ *consists of*

- *a sentence translation function* $\alpha\colon S_1 \to S_2$*, and*
- *a model reduction functor* $\beta\colon \mathcal{M}_2 \to \mathcal{M}_1$*, such that*

$$M_2 \models_2 \alpha(\varphi_1) \text{ if and only if } \beta(M_2) \models_1 \varphi_1$$

*holds for each* $M_2 \in |\mathcal{M}_2|$ *and each* $\varphi_1 \in S_1$ *(*satisfaction condition*).*

Since corridors compose and there are obvious identity corridors, rooms and corridors form a category $\mathbb{R}oom$. Then, an *institution* can be introduced as a functor:

**Definition 3.** *Given a category* $\mathbb{S}ign$ *of signatures and signature morphisms, an* institution *is a functor* $\mathcal{I}\colon \mathbb{S}ign \to \mathbb{R}oom$*.*

Spelled out in more detail, the notion of an institution consists of

- a functor **Sen**: **Sign** $\to \mathcal{S}et$,[1] giving a set $\mathbf{Sen}(\Sigma)$ of $\Sigma$-*sentences* for each signature $\Sigma \in |\mathbf{Sign}|$, and a function $\mathbf{Sen}(\sigma)\colon \mathbf{Sen}(\Sigma_1) \to \mathbf{Sen}(\Sigma_2)$, denoted by $\sigma(\_)$, that yields $\sigma$-*translation* of $\Sigma_1$-sentences to $\Sigma_2$-sentences for each signature morphism $\sigma\colon \Sigma_1 \to \Sigma_2$;
- a functor **Mod**: $\mathbf{Sign}^{op} \to \mathbb{CAT}$,[2] giving a category $\mathbf{Mod}(\Sigma)$ of $\Sigma$-*models* for each signature $\Sigma \in |\mathbf{Sign}|$, and a functor $\mathbf{Mod}(\sigma)\colon \mathbf{Mod}(\Sigma_2) \to \mathbf{Mod}(\Sigma_1)$, denoted by $\_|_\sigma$, that yields $\sigma$-*reducts* of $\Sigma_2$-models for each signature morphism $\sigma\colon \Sigma_1 \to \Sigma_2$; and
- for each $\Sigma \in |\mathbf{Sign}|$, a *satisfaction relation* $\models_{\mathcal{I},\Sigma} \subseteq |\mathbf{Mod}(\Sigma)| \times \mathbf{Sen}(\Sigma)$. The pair $(\mathbf{Sen}(\sigma), \mathbf{Mod}(\sigma))$ associated with a given signature morphism $\sigma : \Sigma_1 \to \Sigma_2$ fulfills the *satisfaction condition.*

Relationships between institutions (and entailment systems) are captured mathematically by ‘institution morphisms’, of which there are several variants, each yielding a category under a canonical composition. For the purposes of this appendix, the institution morphisms of Goguen and Roşu (2002) seem technically most convenient. For the notion of an institutional monad introduced below, we also need 2-cells between institution morphisms, called modifications.

**Definition 4.** *Given institutions* $I_1\colon \mathbb{S}ign_1 \to \mathbb{R}oom$ *and* $I_2\colon \mathbb{S}ign_2 \to \mathbb{R}oom$, *an* institution morphism $(\Phi, \rho)\colon I_1 \to I_2$ *consists of a functor* $\Phi\colon \mathbb{S}ign_1 \to \mathbb{S}ign_2$ *and a natural transformation* $\rho\colon I_2 \circ \Phi \to I_1$.

*Given institution morphisms* $(\Phi, \rho)\colon I_1 \to I_2$ *and* $(\Phi', \rho')\colon I_1 \to I_2$, *an* institution morphism modification $\theta\colon (\Phi, \rho) \to (\Phi', \rho')$ *is just a natural transformation* $\theta\colon \Phi \to \Phi'$ *such that* $\rho = \rho' \circ (I_2 \cdot \theta)$.[3]

*This leads to a 2-category* $\mathbb{I}ns$ *of institutions, morphisms and modifications.*[4]

These notions are demonstrated now using a simple example. Here $FOL^=$ denotes the institution of first-order logic with equality and *Prop* is the institution of propositional logic.

**Example 5.** *There is an institution morphism* $\mu_1 : FOL^= \to Prop$. *From a first-order signature, it only keeps the nullary predicates, which become propositional variables. Also from a first-order model, only the interpretations of the nullary predicates are kept. Moreover, there is an obvious inclusion of Prop-sentences into* $FOL^=$*-sentences. The satisfaction condition is easily shown.* □

**Example 6.** *Another institution morphism* $\mu_2 : FOL^= \to Prop$ *keeps* all *predicates from a first-order signature as propositional variables. From a first-order model, extract a valuation by mapping a predicate to true iff it is universally true. A propositional variable is translated to a sentence stating that the corresponding predicate holds universally. Again, the satisfaction condition is easily shown.* □

**Example 7.** *The inclusions* $\iota_\Sigma : (\mu_1)_\Sigma \to (\mu_2)_\Sigma$ *form a institution morphism modification* $\iota : \mu_1 \to \mu_2$. □

---

[1] The category $\mathcal{S}et$ has all sets as objects and all functions as morphisms.

[2] $\mathbb{CAT}$ is the quasi-category of all categories, where “quasi” means that it lives in a higher set-theoretic universe.

[3] The original notion from Diaconescu (2002) is a lax variant of this: a morphism $\rho \to \rho' \circ (I_2 \cdot \theta)$ is given instead of equality.

[4] See Lack (2010) for more information on 2-categories.

Using general notions from the theory of 2-categories, one can introduce the notion of an institutional monad.[5] The essential point is that a monad on signatures needs to interact with the structure of the institution:

**Definition 8.** *An* institutional monad $\mathcal{T} = (T, \eta, \mu)$ *is a monad in the 2-category* $\mathbb{I}ns$*, which amounts to*

- *an institution* $\mathcal{I}$*,*
- *an institution morphism* $T = (\Phi^T, \rho^T) \colon \mathcal{I} \to \mathcal{I}$*,*
- *an institution morphism modification* $\eta \colon id \to T$*, and*
- *an institution morphism modification* $\mu \colon T \times T \to T$*,*

*such that the usual laws of a monad are satisfied:*

$$\begin{tikzcd}
T \arrow[r, "\eta_T"] \arrow[d, "T\eta"] \arrow[dr, equal] & T \times T \arrow[d, "\mu"] \\
T \times T \arrow[r, "\mu"] & T
\end{tikzcd}
\qquad\qquad
\begin{tikzcd}
T \times T \times T \arrow[r, "\mu_T"] \arrow[d, "T\mu"] & T \times T \arrow[d, "\mu"] \\
T \times T \arrow[r, "\mu"] & T
\end{tikzcd}$$

By selecting the signature component only, an institutional monad $\mathcal{T}$ gives rise to an ordinary monad in the category $\mathbb{S}ign$, which is here denoted by $\mathcal{T}^{\mathbb{S}ign}$. The idea is that a suitable monad describes the construction of a *derived signature*, i.e. how to systematically introduce a *signature of terms* $\Phi^T(\Sigma)$ for a given signature $\Sigma$ from $\mathbb{S}ign$. For first-order logic, one example is the standard derivation already presented in the main text (section 3.3), which in fact constitutes an institutional monad:

**Example 9.** *Let* $T = (\Phi^T, \rho^T) \colon FOL^{=} \to FOL^{=}$ *with* $\rho^T_\Sigma = (\alpha^T_\Sigma, \beta^T_\Sigma)$*.*

- $\Phi^T(\Sigma)$ *adds terms* $\lambda x_1 : s_1, \dots x_n : s_n.t$ *as* $n$*-ary operations and terms* $\lambda x_1 : s_1, \dots x_n : s_n.\varphi$ *(where* $\varphi$ *is a formula) as* $n$*-ary predicates;*
- $\alpha^T_\Sigma \colon Sen(\Phi^T(\Sigma)) \to Sen(\Sigma)$ $\beta$*-reduces all application of* $\lambda$*-term operations and predicates;*
- $\beta^T_\Sigma \colon Mod(\Sigma) \to Mod(\Phi^T(\Sigma))$ *interprets* $\lambda$*-term operations and predicates in* $\beta^T_\Sigma(M)$ *as* $(a_1, \dots, a_n) \mapsto M(t)[x_i \mapsto a_i]$
- $\eta_\Sigma \colon \Sigma \to \Phi^T(\Sigma)$ *is the obvious inclusion;*
- $\mu_\Sigma \colon \Phi^T(\Phi^T(\Sigma)) \to \Phi^T(\Sigma)$ *collapses two levels of* $\lambda$*-terms into one.*

Given such a derivation, a derived signature morphism $\sigma \colon \Sigma_1 \to \Sigma_2$ is defined as an ordinary signature morphism $\sigma \colon \Sigma_1 \to \Phi^T(\Sigma_2)$. In fact, one can also introduce a composition for these morphisms, leading to the notion of a *Kleisli institution*:

**Definition 10.** *Given an institutional monad* $\mathcal{T} = (T \colon \mathcal{I} \to \mathcal{I}, \eta, \mu)$*, its* Kleisli institution $\mathcal{I}_{\mathcal{T}}$ *is the Kleisli object of* $\mathcal{T}$ *in* $\mathbb{I}ns$*, which amounts to*

- *the signature category of* $\mathcal{I}_{\mathcal{T}}$ *is the Kleisli category of the monad* $\mathcal{T}^{\mathbb{S}ign}$*,*

---

[5] Again, see Lack (2010) for the notion of monad in a 2-category.

- *given a signature* $\Sigma \in \mathcal{T}^{\mathbb{S}ign}$, $\mathcal{I}_{\mathcal{T}}(\Sigma)$ *is just* $\mathcal{I}(\Sigma)$*, and*
- *given a signature morphism* $\sigma\colon \Sigma_1 \to \Sigma_2 \in \mathcal{T}^{\mathbb{S}ign}$ *(which is a signature morphism* $\sigma\colon \Sigma_1 \to \Phi^T(\Sigma_2)$ *in* $\mathcal{I}$*),* $\mathcal{I}_{\mathcal{T}}(\sigma)$ *is given by*

$$\begin{tikzcd} \mathcal{I}(\Sigma_1) \arrow[r, "\mathcal{I}(\sigma)"] & \mathcal{I}(\Phi^T(\Sigma_2)) \arrow[r, "\rho^T_{\Sigma_2}"] & \mathcal{I}(\Sigma_2) \end{tikzcd}$$

  *providing sentence translation and model reduct for Kleisli morphisms.*

For a given institution and a derivation (provided by an institutional monad) the Kleisli institution consists of the same signatures, connected by derived signature morphisms. The institution independent notions of logical consequence, theory etc. and corresponding results of course also apply to the Kleisli institution; in particular, Kleisli theory morphisms preserve logical consequence.

# Bibliography

Abdel-Fattah, A., Besold, T., Gust, H., Krumnack, U., Kühnberger, K.-U., Schmidt, M., Schneider, S., and Schwering, A. (2013). Sketch learning by analogy. In Kutz, O., Bhatt, M., Borgo, S., and Santos, P., editors, *Proceedings of the Second Interdisciplinary Workshop on The Shape of Things (Shapes 2.0)*, volume 1007 of *CEUR Workshop Proceedings*, pages 48–58.

Abdel-Fattah, A. M. H. (2014). *Utilizing Cross-Domain Cognitive Mechanisms for Modeling Aspects of Artificial General Intelligence.* PhD thesis, Institute of Cognitive Science, University of Osnabrück.

Abdel-Fattah, A. M. H., Besold, T., and Kühnberger, K.-U. (2012). Creativity, cognitive mechanisms, and logic. In Bach, J., Goertzel, B., and Iklé, M., editors, *Artificial General Intelligence*, volume 7716 of *Lecture Notes in Computer Science*, pages 1–10. Springer.

Alexander, J. C. (2011). Blending in mathematics. *Semiotica*, 187:1–48.

Argand, J.-R. (1813). Philosophie mathématique. essay sur une manière de représenter les quantités imaginaires, dans les constructions géométriques. *Annales de Mathématiques pures et appliquées*, 4:133–146.

Armengol, E. and Plaza, E. (2000). Bottom-up induction of feature terms. *Machine Learning*, 41(3):259–294.

Barendregt, H. (1991). Introduction to generalized type systems. *Journal of Functional Programming*, 1(2):125–154.

Bartha, P. F. A. (2010). *By Parallel Reasoning: The Construction and Evaluation of Analogical Arguments.* Oxford University Press.

Barwise, J. and Seligman, J. (1997). *Information Flow: the Logic of Distributed Systems.* Cambridge University Press.

Baumgartner, A., Kutsia, T., Levy, J., and Villaret, M. (2012). A variant of higher-order anti-unification. RISC 12-19, Johannes Kepler University Linz.

Besold, T. R. (Besold:2013). Rationality in context: An analogical perspective. In Brézillon, P., Blackburn, P., and Dapoigny, R., editors, *Modeling and Using Context*, volume 8175 of *Lecture Notes in Computer Science*, pages 129–142. Springer.

Besold, T. R., Pease, A., and Schmidt, M. (2013). Analogy and arithmetic: An HDTP-based model of the calculation circular staircase. In *Proceedings of the 35th Annual Conference of the Cognitive Science Society*, pages 1893–1898, Austin, TX. Cognitive Science Society.

Burghardt, J. (2005). E-generalization using grammars. *Artificial Intelligence Journal*, 165(1):1–35.

Byrne, R. M. J. (2005). *The Rational Imagination: How People Create Alternatives To Reality.* Bradford Books. MIT Press.

Chalmers, D. J., French, R. M., and Hofstadter, D. R. (1992). High-level perception, representation, and analogy: A critique of artificial intelligence methodology. *Journal of Experimental & Theoretical Artificial Intelligence*, 4(3):185–211.

Coulson, S. (2001). *Semantic Leaps: Frame-Shifting and Conceptual Blending in Meaning Construction.* Cambridge University Press.

Darrigol, O. (1992). *From c-Numbers to q-Numbers.* University of California Press.

Dastani, M. M. (1998). *Languages of Perception.* PhD thesis, University of Amsterdam.

Diaconescu, R. (2002). Grothendieck institutions. *Applied Categorical Structures*, 10(4):383–402.

Diaconescu, R. (2008). *Institution-independent Model Theory.* Studies in Universal Logic. Birkhäuser, Basel.

Evans, T. G. (1962). A heuristic program to solve geometric analogy problems. AI Memos at the MIT (AIM-64).

Falkenhainer, B., Forbus, K. D., and Gentner, D. (1986). The structure-mapping engine. In *5th National Conference on Artificial Intelligence*, pages 272–277.

Falkenhainer, B., Forbus, K. D., and Gentner, D. (1989). The structure-mapping engine: Algorithm and examples. *Artificial Intelligence*, 41:1–63.

Fauconnier, G. and Turner, M. (1998). Conceptual integration networks. *Cognitive Science*, 22(2):133–187.

Fauconnier, G. and Turner, M. (2002). *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities.* Basic Books.

Feng, C. and Muggleton, S. (1992). Towards inductive generalization in higher order logic. In Sleeman, D. H. and Edwards, P., editors, *Proceedings of the Ninth International Workshop on Machine Learning (ML'92)*, pages 154–162. Morgan Kaufmann.

Forbus, K. D., Gentner, D., Markman, A. B., and Ferguson, R. W. (1998). Analogy just looks like high-level perception: why a domain-general approach to analogical mapping is right. *Journal of Experimental & Theoretical Artificial Intelligence*, 10(2):231–257.

Forbus, K. D., Lockwood, K., Klenk, M., Tomai, E., and Usher, J. (2004). Open-domain sketch understanding: The nuSketch approach. In *AAAI Fall Symposium on Making Pen-based Interaction Intelligent and Natural*, pages 58–63, Washington, DC. AAAI Press.

French, R. M. (2002). The computational modeling of analogy-making. *Trends in Cognitive Sciences*, 6(5):200–205.

Furtado, A. L. (1992). Analogy by generalization—and the quest of the grail. *ACM SIGPLAN Notices*, 27(1):105–113.

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

Gentner, D. (2003). Why we're so smart. In Gentner, D. and Goldin-Meadow, S., editors, *Language in mind: Advances in the study of language and thought*, pages 195–235. MIT Press, Cambridge, MA.

Gentner, D. and Forbus, K. D. (2011). Computational models of analogy. *Wiley Interdisciplinary Reviews: Cognitive Science*, 2(3):266–276.

Gentner, D., Holyoak, K., and Kokinov, B., editors (2001). *The Analogical Mind: Perspectives from Cognitive Science*. MIT Press.

Goguen, J. and Harrell, D. F. (2010). Style: A computational and conceptual blending-based approach. In Argamon, S., Burns, K., and Dubnov, S., editors, *The Structure of Style: Algorithmic Approaches to Understanding Manner and Meaning*, pages 291–316. Springer, Berlin.

Goguen, J. and Roşu, G. (2002). Institution morphisms. *Formal Aspects of Computing*, 13(3–5):274–307.

Goguen, J. A. (2006). Mathematical models of cognitive space and time. In Andler, D., Ogawa, Y., Okada, M., and Watanabe, S., editors, *Proceedings of the Interdisciplinary Conference on Reasoning and Cognition*. Keio University Press.

Goguen, J. A. and Burstall, R. M. (1984). Introducing institutions. In Clarke, E. and Kozen, D., editors, *Logics of Programs*, volume 164/1984 of *Lecture Notes in Computer Science*, pages 221–256. Springer.

Goguen, J. A. and Burstall, R. M. (1985). A study in the foundations of programming methodology: Specifications, institutions, charters and parchments. In Pitt, D., Abramsky, S., Poignè, A., and Rydeheard, D., editors, *Category Theory and Computer Programming*, volume 240 of *LNCS*, pages 313–333. Springer Verlag.

Goguen, J. A. and Burstall, R. M. (1986). A study in the foundations of programming methodology: Specifications, institutions, charters and parchments. In Pitt, D., Abramsky, S., Poigné, A., and Rydeheard, D., editors, *Category Theory and Computer Programming*, volume 240 of *Lecture Notes in Computer Science*, pages 313–333. Springer.

Goguen, J. A., Thatcher, J. W., and Wagner, E. G. (1978). An initial algebra approach to the specification, correctness and implementation of abstract data types. In Yeh, R. T., editor, *Current Trends in Programming Methodology – Vol. IV (Data Structuring)*, pages 80–149. Prentice-Hall.

Goodman, N. (1947). The problem of counterfactual conditional sentences. *Journal of Philosophy*, 44(5):113–128.

Guarini, M., Butchart, A., Smith, P. S., and Moldovan, A. (2009). Resources for research on analogy: A multi-disciplinary guide. *Informal Logic*, 29(2):84–197.

Guhe, M., Pease, A., Smaill, A., Schmidt, M., Gust, H., Kühnberger, K.-U., and Krumnack, U. (2010). Mathematical reasoning with higher-order anti-unification. In *Annual Meeting of the Cognitive Science Society (CogSci 2010)*, pages 1192–1197.

Guhe, M., Smaill, A., and Pease, A. (2009a). A formal cognitive model of mathematical metaphors. In *KI 2009: Advances in Artificial Intelligence*, volume 5803 of *Lecture Notes in Computer Science*, pages 323–330.

Guhe, M., Smaill, A., and Pease, A. (2009b). Using information flow for modelling mathematical metaphors. In *Proceedings of the 9th International Conference on Cognitive Modeling*, pages 258–263.

Gust, H., Krumnack, U., Schwering, A., and Kühnberger, K.-U. (2009). The role of logic in agi systems: Towards a lingua franca for general intelligence. In Goertzel, B., Hitzler, P., and Hutter, M., editors, *Artificial General Intelligence. Second Conference on Artificial General Intelligence (AGI 2009)*, pages 43–48. Atlantis Press.

Gust, H., Kühnberger, K.-U., Krumnack, U., and Schwering, A. (2008). Analogical reasoning: A core of cognition. *Zeitschrift für Künstliche Intelligenz (KI)*, 1/08:8–12. Themenheft KI und Kognition.

Gust, H., Kühnberger, K.-U., and Schmid, U. (2006). Metaphors and heuristic-driven theory projection (HDTP). *Theoretical Computer Science*, 354(1):98–117.

Hasker, R. W. (1995). *The replay of program derivations.* PhD thesis, University of Illinois at Urbana-Champaign, Champaign, IL, USA.

Hasker, R. W. and Reddy, U. S. (1992). Generalization at higher types. In Miller, D., editor, *Proceedings of the Workshop on the Prolog Programming Language*, pages 257–271. University of Pennsylvania. Available as UPenn CIS technical report MS-CIS-92-86.

Hesse, M. B. (1966). *Models and Analogies in Science.* Notre Dame University Press, Notre Dame, IN.

Hofstadter, D. and the Fluid Analogies Research Group (1995). *Fluid concepts and creative analogies.* Harvester Wheatsheaf, New York.

Hofstadter, D. R. (2001). Analogy as the core of cognition. In Gentner et al. (2001), pages 499–538.

Hofstadter, D. R. and Mitchell, M. (1994). The copycat project: A model of mental fluidity and analogy-making. In Holyoak, K. and Barnden, J., editors, *Advances in Connectionist and Neural Computation Theory Volume 2: Analogical Connections*, pages 31–112. Ablex Publishing Corporation, Norwood NJ.

Holyoak, K. J. and Thagard, P. (1989). Analogical mapping by constraint satisfaction. *Cognitive Science*, 13:295–355.

Indurkhya, B. (1989). Modes of analogy. In Jahnke, K. P., editor, *Analogical and Inductive Inference*, volume 397 of *Lecture Notes in Computer Science*, pages 217–230. Springer-Verlag.

Indurkhya, B. (1992). *Metaphor and cognition: an interactionist approach.* Kluwer Academic Publishers, Dordrecht.

Knight, K. (1989). Unification: a multidisciplinary survey. *ACM Computing Surveys*, 21(1):93–124.

Koffka, K. (1935). *Principles of Gestalt Psychology.* Harcourt, New York.

Köhler, W. (1929). *Gestalt Psychology.* Liveright, New York.

Kokinov, B. (1994). The context-sensitive cognitive architecture DUAL. In *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society*, pages 502–507, Hillsdale, NJ. Erlbaum.

Kokinov, B. and Petrov, A. (2001). Integration of memory and reasoning in analogy-making: The AMBR model. In Gentner et al. (2001), pages 59–124.

Kotovsky, L. and Gentner, D. (1999). Comparison and categorization in the development of relational similarity. *Child Development*, 67(6):2797–2822.

Krumnack, U., Mossakowski, T., and Maibaum, T. (2014). What is a derived signature morphism? In Diaconescu, R., Codescu, M., and Ţuţu, I., editors, *WADT 2014: 22nd International Workshop on Algebraic Development Techniques*, pages 50–51.

Kutsia, T., Levy, J., and Villaret, M. (2011). Anti-unification for unranked terms and hedges. In Schmidt-Schauß, M., editor, *Proceedings of the 22nd International Conference on Rewriting Techniques and Applications (RTA 2011)*, volume 10 of *LIPIcs*, pages 219–234. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik.

Kutz, O., Bateman, J., Neuhaus, F., Mossakowski, T., and Bhatt, M. (2015). E pluribus unum. In Besold, T. R., Schorlemmer, M., and Smaill, A., editors, *Computational Creativity Research: Towards Creative Machines*, volume 7 of *Atlantis Thinking Machines*, pages 167–196. Springer.

Lack, S. (2010). A 2-categories companion. In Baez, J. C. and May, J. P., editors, *Towards Higher Categories*, volume 152 of *The IMA Volumes in Mathematics and its Applications*, pages 105–191. Springer.

Lakoff, G. and Núñez, R. E. (2000). *Where Mathematics Comes From.* Basic Books, New York.

Lee, M. and Barnden, J. (2001). A computational approach to conceptual blending within counterfactuals. Cognitive Science Research Papers CSRP-01-10, School of Computer Science, University of Birmingham.

Lewis, D. K. (1973). *Counterfactuals*. Harvard University Press, Cambridge, Mass.

Lu, J., Mylopoulos, J., Harao, M., and Hagiya, M. (2000). Higher order generalization and its application in program verification. *Annals of Mathematics and Artificial Intelligence*, 28(1–4):107–126.

MacLane, S. (1998). *Categories for the working mathematician*. Springer, New York, second edition.

Martinez, M., Besold, T. R., Abdel-Fattah, A., Kühnberger, K.-U., Gust, H., Schmidt, M., and Krumnack, U. (2011). Towards a domain-independent computational framework for theory blending. In *AAAI Fall 2011 Symposium on Advances in Cognitive Systems*, pages 210–217. AAAI Press.

Meltzer, B. and Michie, D., editors (1970). *Machine Intelligence*, volume 5. Edinburgh University Press, Edinburgh.

Nelson, G. and Otten, D. C. (1979). Simplification by cooperating decision procedures. *ACM Transactions on Programming Languages and Systems*, 1(2):245–257.

Oppenheimer, R. (1956). Analogy in science. *American Psychologist*, 11(3):127–135.

Pearl, J. (2011). The algorithmization of counterfactuals. *Annals of Mathematics and Artificial Intelligence*, 61(1):29–39.

Pereira, F. C. (2007). *Creativity and artificial intelligence: a conceptual blending approach*, volume 4 of *Applications of cognitive linguistics*. Mouton de Gruyter.

Pereira, F. C. and Cardoso, A. (2003). Optimality principles for conceptual blending: A first computational approach. *AISB Journal*, 1(4).

Pfenning, F. (1991). Unification and anti-unification in the Calculus of Constructions. In *Sixth Annual IEEE Symposium on Logic in Computer Science*, pages 74–85.

Plotkin, G. D. (1970). A note on inductive generalization. In Meltzer and Michie (1970), pages 153–163.

Plotkin, G. D. (1971). A further note on inductive generalization. *Machine Intelligence*, 6:101–124.

Polya, G. (1954). *Mathematics and Plausible Reasoning: Induction and Analogy in Mathematics*, volume 1. Princeton University Press, Princeton.

Popplestone, R. J. (1970). An experiment in automatic induction. In Meltzer and Michie (1970), pages 203–215.

Quine, W. V. O. (1960). *Word and Object.* MIT Press, New York/London.

Ramsey, F. P. (1929). General propositions and causality. In Mellor, H., editor, *Philosophical Papers*, pages 145–153. Cambridge University Press, Cambridge.

Raven, J. C. (1936). Mental tests used in genetic studies: The performance of related individuals on tests mainly educative and mainly reproductive. Msc thesis, University of London.

Reynolds, J. C. (1970). Transformational systems and the algebraic structure of atomic formulas. In Meltzer and Michie (1970), pages 135–151.

Robere, R. and Besold, T. R. (2012). Complex analogies: Remarks on the complexity of HDTP. In Thielscher, M. and Zhang, D., editors, *AI 2012: Advances in Artificial Intelligence*, volume 7691 of *Lecture Notes in Computer Science*, pages 530–542, Berlin. Springer.

Rutherford, E. (1911). The scattering of $\alpha$ and $\beta$ particles by matter and the structure of the atom. *Philosophical Magazine Series 6*, 21(125):669–688.

Sannella, D. and Tarlecki, A. (2012). *Foundations of Algebraic Specification and Formal Software Development.* Monographs in Theoretical Computer Science. Springer, Berlin.

Schmidt, M. (2010). Strategies in structural alignment for heuristic-driven theory projection. Master's thesis, University of Osnabrück.

Schön, D. A. (1963). *Displacement of Concepts.* Humanities Press, New York.

Schwering, A., Bauer, C., Dorceva, I., Gust, H., Krumnack, U., and Kühnberger, K.-U. (2009a). The impact of gestalt principles on solving geometric proportional analogies. In Kokinov, B., Holyoak, K., and Gentner, D., editors, *New Frontiers in Analogy Research. Proceedings of the Second International Conference on Analogy*, pages 404–413. New Bulgarian University Press.

Schwering, A., Krumnack, U., Kühnberger, K.-U., and Gust, H. (2007). Using gestalt principles to compute analogies of geometric figures. In McNamara, D. S. and Trafton, G., editors, *Proceedings of the 29th Annual Conference of the Cognitive Science Society*, pages 1485–1490, Austin, TX. Cognitive Science Society.

Schwering, A., Kühnberger, K.-U., Krumnack, U., Gust, H., Wandmacher, T., Indurkhya, B., and Ojha (2009b). A computational model for visual metaphors: Interpreting creative visual advertismens. In Filipe, J., Fred, A. L. N., and Sharp, B., editors, *Proceedings of International Conference on Agents and Artificial Intelligence (ICAART 2009)*, pages 339–344. INSTICC Press.

Spearman, C. (1923). *The nature of intelligence and the principles of cognition.* Macmillan, London, UK.

Veale, T. and O'Donoghue, D. (2000). Computation and blending. *Computational Linguistics*, 11(3–4):253–282. Special issue on Conceptual Blending.

Walker, A., McCord, M., Sowa, J. F., and Wilson, W. G. (1987). *Knowledge systems and Prolog*. Addison-Wesley, Reading, MA.

Wareham, T., Robere, R., and van Rooij, I. (2012). A change for the better? assessing the computational cost of re-representation. In Rus̆winkel, N., Drewitz, U., and van Rijn, H., editors, *Proceedings of the 11th International Conference on Cognitive Modeling*, pages 111–116, Berlin. Universitätsverlag der TU Berlin.

Wertheimer, M. (1912). Experimentelle Studien über das Sehen von Bewegung. *Zeitschrift für Psychologie*, 61(1):161–265.

Wisniewskia, E. J. and Gentner, D. (1991). On the combinatorial semantics of noun pairs: Minor and major adjustments to meaning. In Simpson, G. B., editor, *Understanding Word and Sentence*, volume 77 of *Advances in Psychology*, chapter 10, pages 241–284. Elsevier.

Yan, J., Forbus, K., and Gentner, D. (2003). A theory of rerepresentation in analogy matching. In Alterman, R. and Kirsh, D., editors, *Proceedings of the Twenty-Fifth Annual Meeting of the Cognitive Science Society*, pages 1265–1270, Boston, MA. Cognitive Science Society.

# Part II

# Referenced Published Articles

# Publication 1 (Krumnack et al., 2007): *Restricted Higher-Order Anti-Unification for Analogy Making*

**Abstract:**

Anti-unification has often be used as a tool for analogy making. But while first-order anti-unification is too simple for many applications, general higher-order anti-unification is too complex and leads into theoretical difficulties. In this paper we present a restricted framework for higher-order substitutions and show that anti-unification is well-defined in this setting. A complexity measure for generalizations can be introduced in a quite natural way, which allows for selecting preferred generalizations. An algorithm for computing such generalizations is presented and the utility of complexity for anti-unifying sets of terms is discussed by an extended example.

**Originally published as:**

Ulf Krumnack, Angela Schwering, Helmar Gust, and Kai-Uwe Kühnberger. Restricted higher-order anti-unification for analogy making. In Mehmet A. Orgun and John Thornton, editors, *AI 2007: Advances in Artificial Intelligence*, volume 4830 of *Lecture Notes of Artificial Intelligence*, pages 273–282. Springer, 2007.

**DOI:**
10.1007/978-3-540-76928-6_29

# Publication 2 (Krumnack et al., 2008): *Re-representation in a Logic-Based Model for Analogy Making*

**Abstract:**

Analogical reasoning plays an important role for cognitively demanding tasks. A major challenge in computing analogies concerns the problem of adapting the representation of the domains in a way that the analogous structures become obvious, i.e. finding and, in certain circumstances, generating appropriate representations that allow for computing an analogical relation. We propose to resolve this re-representation problem of analogy making in a logical framework based on the anti-unification of logical theories. The approach is exemplified using examples from qualitative reasoning (naive physics) and mathematics.

**Originally published as:**

Ulf Krumnack, Helmar Gust, Kai-Uwe Kühnberger, and Angela Schwering. Re-representation in a logic-based model for analogy making. In Wayne R. Wobcke and Mengjie Zhang, editors, *AI 2008: Advances in Artificial Intelligence, 21st Australasian Artificial Intelligence Conference*, volume 5360 of *Lecture Notes in Artificial Intelligence*, pages 42–48. Springer, 2008.

**DOI:**

10.1007/978-3-540-89378-3_4

# Publication 3 (Schwering et al., 2009): *Syntactic Principles of Heuristic-Driven Theory Projection*

**Abstract:**

Analogy making is a central construct in human cognition and plays an important role to explain cognitive abilities. While various psychologically or neurally inspired theories for analogical reasoning have been proposed, there is a lack of a logical foundation for analogical reasoning in artificial intelligence and cognitive science. We aim to close this gap and propose heuristic-driven theory projection (HDTP), a mathematically sound framework for analogy making. HDTP represents knowledge about the source and the target domain as first-order logic theories and compares them for structural commonalities using anti-unification. The paper provides an overview of the syntactic principles of HDTP, explains all phases of analogy making at a formal level, and illustrates these phases with examples.

**Originally published as:**

Angela Schwering, Ulf Krumnack, Kai-Uwe Kühnberger, and Helmar Gust. Syntactic principles of Heuristic-Driven Theory Projection. *Special Issue on Analogies - Integrating Cognitive Abilities. In: Journal of Cognitive Systems Research*, 10(3):251–269, 2009.

**DOI:**

10.1016/j.cogsys.2008.09.002

# Publication 4 (Schwering et al., 2009): *Spatial Cognition of Geometric Figures in the Context of Proportional Analogies*

**Abstract:**

The cognition of spatial objects differs among people and is highly influenced by the context in which a spatial object is perceived. We investigated experimentally how humans perceive geometric figures in geometric proportional analogies and discovered that subjects perceive structures within the figures which are suitable for solving the analogy. Humans do not perceive the elements within a figure individually or separately, but cognize the figure as a structured whole. Furthermore, the perception of each figure in the series of analogous figures is influenced by the context of the whole analogy. A computational model which shall reflect human cognition of geometric figures must be flexible enough to adapt the representation of a geometric figure and produce a similarly structured representation as humans do while solving the analogy. Furthermore, it must be able to take into account the context, i.e. structures and transformations in other geometric figures in the analogy.

**Originally published as:**

Angela Schwering, Kai-Uwe Kühnberger, Ulf Krumnack, and Helmar Gust. Spatial cognition of geometric figures in the context of proportional analogies. In Kathleen Stewart Hornsby, Christophe Claramunt, Michel Denis, and Gérard Ligozat, editors, *Conference on Spatial Information Theory (COSIT'09)*, volume 5756 of *Lecture Notes in Computer Science*, pages 18–35. Springer, 2009.

**DOI:**
10.1007/978-3-642-03832-7_2

# Publication 5 (Krumnack et al., 2010):
## *Remarks on the Meaning of Analogical Relations*

**Abstract:**

Analogical reasoning plays an important role in the context of higher cognitive abilities of humans. Analogies can be used not only to explain reasoning abilities of humans, but also to explain learning from sparse data, creative problem solving, abstractions of concrete situations, and recognition of formerly unseen situations, just to mention some examples. Research in AI and cognitive science has been proposing several different models of analogy making. Nevertheless, no approach for a model theoretic semantics of analogy making is currently available. This paper gives an analysis of the meaning (the semantics) of analogical relations that are computed by the analogy engine HDTP (Heuristic-Driven Theory Projection).

**Originally published as:**

Ulf Krumnack, Helmar Gust, Angela Schwering, and Kai-Uwe Kühnberger. Remarks on the meaning of analogical relations. In Eric Baum, Marcus Hutter, and Emanuel Kitzelmann, editors, *Third Conference on Artificial General Intelligence (AGI'10)*, pages 67–72, Amsterdam, 2010. Atlantis Press.

**DOI:**
10.2991/agi.2010.8

# Publication 6 (Guhe et al., 2011):
# *A Computational Account of Conceptual Blending in Basic Mathematics*

**Abstract:**

We present an account of a process by which different conceptualisations of number can be blended together to form new conceptualisations via recognition of common features, and judicious combination of their distinctive features. The accounts of number are based on Lakoff and Núñez's cognitively-based grounding metaphors for arithmetic. The approach incorporates elements of analogical inference into a generalised framework of conceptual blending, using some ideas from the work of Goguen. The ideas are worked out using *Heuristic-Driven Theory Projection* (HDTP, a method based on higher-order anti-unification). HDTP provides generalisations between domains, giving a crucial step in the process of finding commonalities between theories. In addition to generalisations, HDTP can also transfer concepts from one domain to another, allowing the construction of new conceptual blends. Alongside the methods by which conceptual blends may be constructed, we provide heuristics to guide this process.

**Originally published as:**

Markus Guhe, Alison Pease, Alan Smaill, Maricarmen Martinez, Martin Schmidt, Helmar Gust, Kai-Uwe Kühnberger, and Ulf Krumnack. A computational account of conceptual blending in basic mathematics. *Cognitive Systems Research*, 12(3–4):249–265, 2011.

**DOI:**

10.1016/j.cogsys.2011.01.004

# Publication 7 (Martinez et al., 2012): *Theory Blending as a Framework for Creativity in Systems for General Intelligence*

**Abstract:**

Being creative is a central property of humans in solving problems, adapting to new states of affairs, applying successful strategies in previously unseen situations, or coming up with new conceptualizations. General intelligent systems should have the potential to realize such forms of creativity to a certain extent. We think that creativity and productivity issues can be best addressed by taking cognitive mechanisms into account, such as analogy-making, concept blending, computing generalizations and the like. In this chapter, we argue for the usage of such mechanisms for modeling creativity. We exemplify in detail the potential of such a mechanism like theory blending using a historical example from mathe- matics. Furthermore, we argue for the claim that modeling creativity by such mechanisms has a huge potential in a variety of domains.

**Originally published as:**

Maricarmen Martinez, Tarek Richard Besold, Ahmed Abdel-Fattah, Helmar Gust, Martin Schmidt, Ulf Krumnack, , and Kai-Uwe Kühnberger. Theory blending as a framework for creativity in systems for general intelligence. In Pei Wang and Ben Goertzel, editors, *Theoretical Foundations of Artificial General Intelligence*, volume 4 of *Thinking Machines*, pages 219–239. Atlantis Press, 2012.

**DOI:**

10.2991/978-94-91216-62-6_12

# Publication 8 (Krumnack et al., 2013): *Formal Magic for Analogies*

**Abstract:**

During the last decades, a number of different approaches to model analogies and analogical reasoning have been proposed, that apply different knowledge representation and mapping strategies. Nevertheless, analogies still seem to be hard to grasp from a formal perspective, with no known treatment in the literature of their formal semantics. In this paper we present a universal framework that allows to analyze the syntax and the semantics of analogies in a logical setting without committing ourselves to a specific type of logic. We then apply these ideas by considering an analogy model that is based on classical first-order logic.

**Originally published as:**

Ulf Krumnack, Ahmed M. H. Abdel-Fattah, and Kai-Uwe Kühnberger. Formal magic for analogies. In Ahmed M. H. Abdel-Fattah and Kai-Uwe Kühnberger, editors, *Proceedings of the Workshop Formalizing Mechanisms for Artificial General Intelligence and Cognition (Formal MAGiC)*, volume 1-2013 of *Publications of the Intitute of Cognitive Science (PICS)*, pages 44–51, 2013.

**URL:**
http://ikw.uni-osnabrueck.de/de/node/845

# Publication 9 (Abdel-Fattah & Krumnack, 2013): *Creating Analogy-Based Interpretations of Blended Noun Concepts*

**Abstract:**

An analogy-based approach is explained that suggests possible interpretations of previously-unseen modifier-head noun compounds. The approach utilizes an analogical relation between the modifier and the head, viewing them as source and target concepts in the analogy, in order to suggest a relationship between the constituent nouns in the compound. The approach interprets the novel composition by employing the conceptual blending mechanism to create new concept representations in a proposed model of computational creativity.

**Originally published as:**

Ahmed M. H. Abdel-Fattah and Ulf Krumnack. Creating analogy-based interpretations of blended noun concepts. In Georgi Stojanov and Bipin Indurkhya, editors, *AAAI Spring Symposium: Creativity and (Early) Cognitive Development: Papers from the 2013 AAAI Spring Symposium*, pages 2–7, Palo Alto, CA, 2013. The AAAI Press.

**URL:**
https://www.aaai.org/ocs/index.php/SSS/SSS13/paper/view/5718

# Publication 10 (Abdel-Fattah et al., 2013): *The Importance of Two Cognitive Mechanisms in Analyzing CFCs*

**Abstract:**

We aim at covering three facets of the problem of analyzing counterfactual conditionals: emphasizing that the problem is crucial for cognitive systems to consider, investigating the cognitive mechanisms responsible for reasonable analyses of counterfactuals, and proposing how to computationally contribute to solving the problem by cognitive systems. Aiming at an implementation-oriented explication, the challenges that an artificial system may encounter in computationally solving this problem are discussed. The utilization of two particular computationally-plausible cognitive mechanisms is shown to be helpful in overcoming these challenges: we argue that the operational utilization of analogical mapping and conceptual blending is possible, and leads to reasonable analyses of counterfactual conditionals in artificial cognitive systems.

**Originally published as:**

Ahmed M.H. Abdel-Fattah, Ulf Krumnack, and Kai-Uwe Kühnberger. The importance of two cognitive mechanisms in analyzing counterfactuals: An implementation-oriented explication. In Glen Hunt and Ben Meadows, editors, *Second Annual Conference on Advances in Cognitive Systems*, pages 1–18, 2013.

**URL:**
http://www.cogsys.org/posters/2013

# Publication 11 (Martinez et al., 2014):
# *Algorithmic Aspects of Theory Blending*

**Abstract:**

In Cognitive Science, conceptual blending has been proposed as an important cognitive mechanism that facilitates the creation of new concepts and ideas by constrained combination of available knowledge. It thereby provides a possible theoretical foundation for modeling high-level cognitive faculties such as the ability to understand, learn, and create new concepts and theories. This paper describes a logic-based framework which allows a formal treatment of theory blending, discusses algorithmic aspects of blending within the framework, and provides an illustrating worked out example from mathematics.

**Originally published as:**

Maricarmen Martinez, Ulf Krumnack, Tarek Richard Besold, Alan Smaill, Martin Schmidt, Ahmed M. H. Abdel-Fattah, Helmar Gust, Kai-Uwe Kühnberger, Alison Pease, and Markus Guhe. Algorithmic aspects of theory blending. In Gonzalo A. Aranda-Corral, Jacques Calmet, and Francisco J. Martín-Mateos, editors, *Artificial Intelligence and Symbolic Computation*, volume 8884 of *Lecture Notes in Computer Science*, pages 180–192. Springer, 2014.

**DOI:**
10.1007/978-3-319-13770-4_16