Towards a Domain-Independent Computational Framework for Theory Blending

Maricarmen Martinez and **Tarek Besold** and **Ahmed Abdel-Fattah** and
Kai-Uwe Kuehnberger and **Helmar Gust** and **Martin Schmidt** and **Ulf Krumnack**
{mmartine | tbesold | ahabdelfatta | kkuehnbe | hgust | martisch | krumnack }@uos.de
University of Osnabrück, Institute of Cognitive Science
Albrechtstr. 28, 49076 Osnabrück, Germany

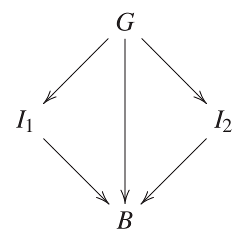
Abstract

The literature on conceptual blending and metaphor-making has illustrations galore of how these mechanisms may support the creation and grounding of new concepts (or whole domains) in terms of a complex, integrated network of older ones. In spite of this, as of yet there is no general computational account of blending and metaphor-making that has proven powerful enough as to cover all the examples from the literature. This paper proposes a logic-based framework for blending and metaphor making and explores its applicability in settings as diverse as mathematical domain formation, classical rationality puzzles, and noun-noun combinations.

Introduction

Since its initial development, the framework of conceptual blending has proved its importance in expressing and explaining cognitive phenomena, such as counterfactual reasoning and metaphor, as well as its usefulness in analogical reasoning (Fauconnier and Turner 2008; 2002). Although there is no commonly agreed definition of conceptual blending, it can roughly be described as the creation of new concepts by a principle driven combination of existing ones. In general, there are many ways to combine the same concepts but in any case the result maintains parts of their initial structure. Furthermore, the blending of two concepts is not a simple union, but may give rise to emergent structure.

While conceptual blending is considered an important part of cognition, there are still few formal or algorithmic accounts. One exception is the Divago system described in (Pereira 2007) that implements blending as a central mechanism of a cognitive architecture. An earlier discussion of computational aspects of blending can be found in (Veale and O'Donoghue 2000). The closest framework to our approach is the work by (Goguen 1999; 2006). Goguen's version of concept blending is formulated at an abstract level (in category theory) and can be described by the following diagram:



In this diagram, two input concepts I_1 and I_2 are related by correlations that are induced by a generalization G. The blend B is supposed to preserve the correlations. This may involve taking the morphisms to B to be partial, so that not all the structure from I_1 and I_2 is mapped to B.

Our work, which is based on the Heuristic-Driven Theory Projection framework (HDTP; Schwering et al. 2009), is a contribution to the effort of finding a general computational account of blending. Towards that end, we have explored several examples in order to test the adequacy of our framework, detect ways in which it may need to be adjusted, and cast light on general strategies of domain formation.

In this paper, we first discuss some lessons and challenges from our past work in this direction and then present a new example concerning the complex plane based on a historical case study. This new example, we think, sheds light on some general strategies that may be useful in order to recruit existing domains from a conceptual network while in the process of creating a new domain. To demonstrate that the techniques we propose are not restricted to mathematics, but rather are universal cognitive mechanisms, we will then sketch how they can be applied to other, maximal disparate problems. We will discuss how classical rationality puzzles might be expressed by our framework, and examine the interpretation of noun-noun composition using our account. The paper finishes with some concluding remarks and an outlook for future work.

We believe that the challenges posed by the examples discussed in this paper are not particular to our framework but in principle serve as a benchmark for any general (computational) account of blending.

Cross-Domain Reasoning

All approaches for blending seem to assume that knowledge is organized in some form of domains, i.e. the underlying

AAAI Fall Symposium 2011 "Advances in Cognitive Systems"

knowledge base provides concepts and facts in groups that can be taken as input to the blending process. In HDTP a domain will be described by a first-order axiomatization, and can be characterized as a consistent micro-theory about a certain aspect of the world. The different domains, though possibly incoherent and even mutually contradictory, are nevertheless not isolated but form a network organized by relations like generalization, analogy, similarity, projection, and instantiation. Besides intra-domain reasoning mechanisms, there also exist cross-domain mechanisms, that allow for the transfer of knowledge between domains and even for the establishment of entirely new domains.

In this section we introduce different cross-domain mechanisms. Since we can not spell out all mechanisms in their technical details here, we will provide the necessary references and illustrate them with examples. For coherence, all examples chosen are taken from mathematics and will also be a base for the historical case study presented in the following section. The general ideas are, however, not restricted to mathematics and we will provide applications in other fields in subsequent sections.

The literature on blending includes many examples of conceptual integration in mathematics, ranging from the most basic to the highly abstract (Fauconnier and Turner 2002; Alexander 2009; Lakoff and Núñez 2000). For example, (Lakoff and Núñez 2000) contains very detailed analyses of how complex mathematical notions may be grounded on a rich network of more basic ones via metaphorical mappings and conceptual blends[1]. At the bottom of this network they have grounding metaphorical mappings through which mathematical notions acquire meaning in terms of everyday domains. For instance, basic arithmetic can be understood metaphorically in terms of four everyday domains: object collections, motion along a path, measuring with a unit rod, and Lego-like object constructions. On top of basic grounding metaphors there are many layers of other (linking) metaphors and blends. These in turn can map domains that are not basic but either the target domains of metaphors or conceptual blends. The idea of the complex plane discussed below involves a blend of already non-basic algebraic, numerical, and geometric notions.

In our framework, domains are represented via finite first-order axiomatizations (defining domain theories) and intra-domain reasoning can be performed with classical logical calculi. Moreover, cross-domain reasoning can also be allowed, in at least the following ways.

1. *Analogies* have been identified as a central mechanism of cognition (Gentner, Holyoak, and Kokinov 2001). They are based on structural commonalities of different domains that provide an analogical relation. Based on this relation, knowledge can be transferred between domains by analogical transfer. HDTP applies the syntactic mechanism of anti-unification (Plotkin 1970) to find generalizations of formulas and to propose an analogical relation (cmp. Figure 1). Several of the mechanisms listed next are in fact supported by analogy-making and transfer.

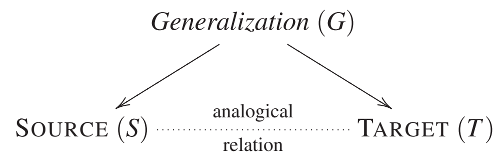


Figure 1: Analogy via generalization in HDTP

Analogical transfer results in structure enrichment. There are cases in which transfer is desired in order to create a new domain, while keeping the original target domain unchanged. In these cases the generalization, input, and enriched domains form a blend. An example found in (Guhe et al. 2011) discusses how to blend (partial formalizations of) the MEASURING STICK (MS) and MOTION ALONG A PATH (MAP). The chosen blended domain is an expansion of MAP in which a notion of unit (taken from MS) is added. This new domain can be thought of as a partial formalization of the positive real number line with markings for the positive integers. While we cannot go through the many details of the example here, we point out two important things suggested by it:

(a) Structure enrichment may involve not only the addition of new symbols to the language of a domain, but also the addition of new sorts. Conceptually, in the example MAP is enriched with a subsort of "whole units", and the axioms of MS are imported into the blended domain, in a form that relativizes them to the subsort.

(b) When possible, attaching to our theory morphisms semantic constraints about the sort of model-mappings that intuitively underpin these morphisms may turn out to be a valuable addition to our setting. The semantic constraints of a theory morphism relating domain theories T_1 and T_2 may say things such as "assume models M_1 and M_2 of T_1 and T_2 such that M_1 is embedded in M_2". The example in the next section will better illustrate what sorts of dynamics may be helped by the kind of semantic labelling of morphisms we are proposing.

2. *Cross-domain generalization*: the HDTP framework always forms a generalized domain in the process of finding analogies. As an example in mathematics (fully formalized in (Guhe et al. 2010)), one could start with partial axiomatizations of two of the basic grounding domains of Lakoff and Núñez mentioned above. These basic domains are analogous in various aspects[2] which are imported by HDTP into the generalization it forms when comparing the domains. As in Figure 2, calculating an analogy between this generalization and yet a third basic domain produces an even more generalized domain, one that gives an abstract approximation of naive arithmetic.

3. *Cross-domain specialization*: once a generalized domain G is obtained, the process of translating terms or properties from G to one of the domains that it generalizes can be thought of as a process of cross-domain specialization.

[1](Lakoff and Núñez 2000) and (Fauconnier and Turner 2002) treat blending in somewhat different but not incompatible ways.

[2]For example, joining object collections and putting together linear constructed Lego-objects are both commutative operations.

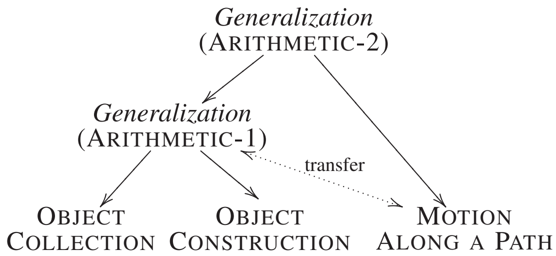


Figure 2: Computing generalizations from Lakoff and Núñez' grounding metaphors for arithmetics

The key is to identify conditions under which it is guaranteed that, say, the translation of any first-order proof done inside G (say Arithmetic-1 in Figure 2) is also a proof inside each of the domains G generalizes (say Object Collection in Figure 2). A discussion of the issues involved in obtaining such conditions, spelled out in the theory of institutions, can be found in (Krumnack et al. 2010).

4. *Detection of congruence relations*: the detection of weak notions of equality, namely structural congruences, is pervasive in mathematics. In (Guhe et al. 2011) we present an example related to the challenge of getting to the notion of fractional numbers based only on the four basic grounding domains proposed by (Lakoff and Núñez 2000). The example is a case of quotient calculation, and the key technical detail is that HDTP can detect when some relation in one of the input domains is analogous to the equality in the other.

   Due to the importance of quotient constructions in mathematics, we propose that the detection of an equality-like relation R provides an important cue that should trigger further processing when our setting is applied to exploratory reasoning in math. The extra processing consists of checking to see if R is a congruence. If successful, the process should result in the diagram of Figure 3.

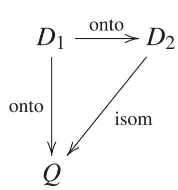


Figure 3: Diagram of a quotient Q over a domain D_1. Intuitively, at the level of models the elements of D_1 are mapped to their equivalence classes in the quotient Q.

Processes of quotient construction involve highlighting a notion of weak equality on a domain D_1, deriving from it a useful way to categorize the entities of D_1, and obtaining a theory of the space of categories. Each one of these aspects is important from the perspective of cognitive systems in general.

Take the case of novel representation systems. Through

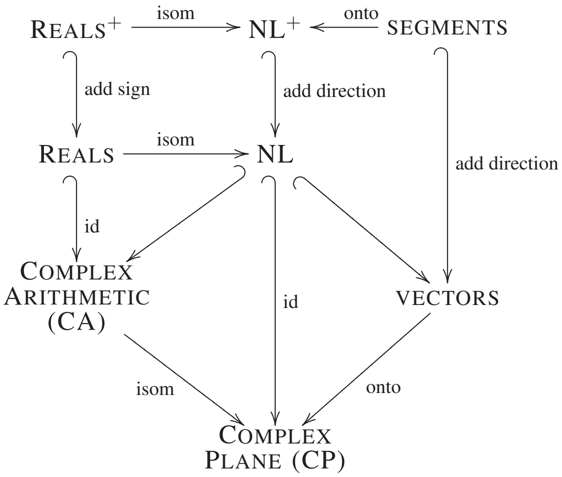


Figure 4: Network of domains involved in Argand's text. Arrows indicate the direction of (partial) theory morphisms. and arrow labels indicate constraints at the level of models. Curved tails are the same as an "injective" label on the arrow.

history humans have shown their ability for devising them. A mathematical example is that the positive real numbers can be used as a representation system for segments and some operations on them. The positive number line can be thought of as a quotient domain (Q) of the domain of physical segments (D_1) obtained thanks to the identification of segments of the same length (same image in the domain D_2 of positive reals) with the unique segment in the number line that has that length. By virtue of the isomorphism between the positive number line and the positive reals, one can also see the quotient Q as a geometric, more concrete model of the theory of the reals. We now contribute a new example that concerns the process of obtaining an intuitive interpretation of the complex numbers.

A Historical Example from Mathematics: the Complex Plane

For a long time mathematicians did not consider the complex numbers to be proper numbers, but rather mere syntactic entities that helped in performing calculations and obtaining useful results. A crucial breakthrough towards their acceptance as numbers (although this is not the whole story) was the discovery of the geometric interpretation known as the complex plane, made independently by J. Wallis and J.R. Argand in years as far apart as 1685 and 1806, respectively. In what follows we illustrate how Argand's ideas, as they appear in (Argand 1813) and (Stedall 2008), can be cast in our framework. Rather than trying to get detailed formalizations of the domains he used, we will formalize only the minimum necessary and instead highlight the dynamics by which old domains in the network are selected and used.

The network of main domains involved in Argand's reasoning, in our understanding of his text, is shown in Figure 4. The complex plain (CP) and the vector domain (VECTORS) are new nodes that Argand added to an already existent network of domains that included all the other nodes in the figure. Together with the complex arithmetic (CA) and number line (NL) domains they form the diamond shape of a blend that emerges from the whole process we set now to describe.

Argand started by pointing out it is easier to accept that numbers of a certain kind are "really numbers" if they correspond to some kind of "physical" magnitude. The negative numbers may not seem "real" if you only consider magnitudes such as the size of object collections, but they do make sense in the context of measuring weight in a two-plate scale, for example, where there is a reference point (an origin or "zero") and a displacement of a plate with respect to that point in one of two possible directions: upwards or downwards. This leads Argand to explicitly recruit the number line representation (NL), an already familiar domain that provides a geometric representation of the real numbers. Then he remarks that the "right side" of the number line (NL^+) is a geometric representation of the positive real numbers.

As discussed earlier, NL^+ is a quotient blend of the positive reals with the domain SEGMENTS of segments in the plane (with a designated unit segment). Argand's NL^+ includes notions of proportions (a:b::c:d) and geometric mean. The geometric mean of two segments a and b in NL^+ is the unique x such that a:x::x:b. It can be obtained algebraically as the solution of $x \cdot x = a \cdot b$ or by moving to the geometric domain of SEGMENTS, constructing the geometric mean, and mapping it to its NL^+ representation.

The next key point comes when Argand wonders what the notion of geometric mean may mean in NL. He notices that every real number is either a positive real or $-r$ for some positive real r, and that similarly each segment in the number line lies in the positive side of it or its (geometric) opposite[3]. He then proposes *as definition* that the magnitude of the geometric mean of two entities in NL is the geometric mean of their magnitudes, while the direction of it is the geometric mean of their directions, with the two possible directions being represented by +1 and -1. So Argand poses the problem of finding x such that 1:x::x:-1. The solution must have both algebraic and geometric readings, as does everything in NL. Algebraically (moving to the isomorphic domain of reals) this boils down to finding x such that $x \cdot x = -1 \cdot 1 = -1$. No such x exists in the reals (and therefore not in NL either), but the domain of complex numbers is recruited as an extension of the reals where there is a solution.

On the geometric side, by meta-analogy with the idea that adding direction to NL^+ leads to NL, Argands takes the domain of segments and enriches it with a notion of direction, thereby establishing a domain of directed segments, or vectors, with a designated unit vector. His newly established domain includes operations by which vectors can be stretched by a real factor, can be added, and rotated. He also considers projections. In Table 1 we give a very partial and redundant axiomatization of his knowledge about this domain. The $+_{2\pi}$ stands for real addition modulo 2π, for adding angles.

types
real, vector
constants
$\mathbf{0}, \mathbf{1}$: *vector*
$0, 1, 90, 180$: *real*
functions
$|\ |$, angle: $vector \to real$
$-$: $vector \to vector$
rotate, scale: $vector \times real \to vector$
$+_v$, *project*: $vector \times vector \to vector$
$+, \cdot, +_{2\pi}$: $real \times real \to real$
predicates
proportion: $vector \times vector \times vector \times vector$
geomean: $vector \times vector \times vector$
$\equiv$: $vector \times vector$
laws
$\forall v\,(rotate(rotate(v,90),90) = -v)$
$rotate(rotate(\mathbf{1},90),90) = -\mathbf{1}$
$\forall v\,(rotate(rotate(v,\alpha),\beta) = rotate(v,\alpha +_{2\pi} \beta))$
$90 +_{2\pi} 90 = 180$
$\forall v \exists a \exists b\,(v = scale(rotate(\mathbf{1},a),b))$
$\forall v \forall a \forall b\,(v = scale(rotate(\mathbf{1},a),b) \leftrightarrow (a = angle(v) \wedge b = |v|))$
$\forall v_1 \forall v_2 \forall v_3 \forall v_4\,(proportion(v_1,v_2,v_3,v_4) \leftrightarrow$
$\exists a \exists b\,(v_2 = rotate(scale(v_1,a),b) \wedge v_4 = rotate(scale(v_3,a),b)))$
$\forall v_1 \forall v_2 \forall v_3\,(geomean(v_1,v_2,v_3) \leftrightarrow proportion(v_1,v_3,v_3,v_2))$
(abelian group axioms for $(+_v, -)$ and $+_{2\pi}$)
(add also the missing vector space axioms for *scale* and $+_v$)
(congruence axioms for $\equiv$)
(imported axioms from the reals for $+, \cdot$)

Table 1: Partial formalization of the VECTORS domain

Argand's realization that the geometric mean of $\mathbf{1}, -\mathbf{1}$ must be a unitary vector perpendicular to $\mathbf{1}$ corresponds here to the fact that from our VECTORS axioms one can prove $geomean(\mathbf{1}, -\mathbf{1}, rotate(\mathbf{1}, 90))$. The other solution, $rotate(\mathbf{1}, -90)$, requires a few more axioms.

Finally, with the notion of geometric mean making sense algebraically in CA and geometrically in VECTORS, Argand investigates whether a geometric representation for the complex numbers can be obtained from the vector plane. In our terms, this boils down to finding a quotient blend space CP with inputs CA and VECTORS and the additional constraint that NL must be embedded in it. To see how this could go, consider the partial formalization of CA (the domain of complex numbers as objects of the form $a + bi$) given in Table 2.

When comparing the two axiomatizations of CA and VECTORS, the HDTP setup can detect that the first formulas in the domains are structurally the same. The generalization G formed by HDTP will include an axiom $\forall X (F(F(X,Y),Y) = -X)$ and the set of substitutions leading back from G to the CA and VECTORS will map the term $F(X,Y)$ to the terms $rotate(v, 90)$ and $i * z$ respectively. Given the intended meaning of our axiomatizations, this corresponds to detecting that rotating vectors by 90 degrees is analogous to the operation of multiplying by i in the

[3]In the future we would like to provide general conditions that applied to this example would allow us to say that since REALS^+ and NL^+ are isomorphic and constructed from REALS and NL in formally identical ways, then they must be isomorphic as well. This is one way in which we envision using semantic labels.

types
real, *complex* (with *real* being a subsort of *complex*)
constants
i: complex
$0, 1$: *real*
functions
re, *im*: *complex* $\rightarrow$ *real*
$+_c, *$: *complex* $\times$ *complex* $\rightarrow$ *complex*
$+, \cdot$: *real* $\times$ *real* $\rightarrow$ *real*
$-$: *complex* $\rightarrow$ *complex*
laws
$\forall z (i * (i * z) = -z)$
$i * i = -1$
$\forall z \forall a \forall b (z = a + b \cdot i \leftrightarrow a = re(z) \wedge b = im(z))$
$\forall z \forall z' (re(z +_c z') = re(z) + re(z'))$
$\forall z \forall z' (im(z +_c z') = im(z) + im(z'))$
$\forall z \exists a \exists b (z = a + b \cdot i))$
$\forall z \forall z' (re(z * z') = re(z) \cdot re(z') - im(z) \cdot im(z'))$
$\forall z \forall z' (im(z * z') = re(z) \cdot im(z') + im(z) \cdot re(z'))$
(abelian group axioms for $+_c$)
(add also the missing field axioms for $+_c$ and $*$)
(imported axioms from the reals for $+, \cdot$)

Table 2: Partial formalization of the CA domain

complex arithmetic domain. This is the key observation on which Argand initiated his construction of the complex plane. The domain G will of course include further things, such as the imported axioms for the reals and an operation G with axioms of abelian group that generalizes $+_v$ and $+_c$.

We believe this example illustrates how a general blending mechanism may benefit from a context given by a relevant part of a stored network of domains. One way in which such context helps is in providing the motivation for finding a particular blend between two spaces. In our case the motivation was a meta-analogy: can we get a geometric representation just as there is one in the network for the positive reals? We conjecture that meta-analogies like this are crucial in general as guides to do fruitful exploration and domain creation based in a network of domains like ours.

A second way in which the context of relevant domains contributes is by suggesting adequacy constrains to be imposed on a the blend to be created. In our example, there are two constraints that go beyond the general diagram of a blend. On the one hand, there is the constraint that (the relevant part of) NL must be embedded into the new blended domain CP. On the other hand, there is the constraint that CP must be the bottom domain of a quotient (i.e. play the role of Q in Figure 3). Interestingly, our example identifies a related issue to be addressed by a good computational account of blending. Namely, it may be that sometimes different adequacy constraints might need to be checked with respect to different blend diagrams. The way this shows in the example is that analogical mappings like the above that generalize the first axioms of the input domains, will also take the equality relations at both sides as analogous, but will fail to detect that the congruence relation of VECTOR is also analogous to the equality relation of CA.

Rationality Puzzles

One of the long-standing challenges in modeling human behavior and cognitive functions is the search for a feasible and widely accepted framework for rationality and rational behavior. Numerous approaches have been proposed, which can generally be clustered into logic-based models (cf. e.g. (Evans 2002)), probability-based models (cf. e.g. (Griffiths, Kemp, and Tenenbaum 2008)), heuristic-based models (cf. e.g. (Gigerenzer, Hertwig, and Pachur 2011)), and game-theoretically based models (cf. e.g. (Osborne and Rubinstein 1994)). Still, up to now, none of these models has really been widely accepted as a general solution to the problem, and the generality of each such class of frameworks has at the same time been challenged by psychological experiments.

In a recent paper (Besold et al. 2011), we proposed an alternate approach to creating a theory of rationality: Instead of formulating an abstract theory and then trying to fit it to empirical observations and psychological results, we suggest to use basic cognitive functions, namely analogy-making, as basis for a rationality framework, turning rationality and utility-maximization into emergent phenomena within this cognitive setting. Following this proposal, we want to focus on one of the most classical challenges to existing frameworks of rationality (mostly targeted as a counterargument to the probabilistic view), namely Tversky and Kahnemann's Linda problem (Tversky and Kahneman 1983) (cf. Table 3), and exemplarily indicate how the analogy-making engine HDTP would be capable of adressing this problem by means of its concept blending capabilities.

Linda-Problem:
Linda is 31 years old, single, outspoken and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations.

(F): Linda is active in the feminist movement.
(T): Linda is a bank teller.
(T&F): Linda is a bank teller and is active in the feminist movement.

Table 3: An abbreviated version of Tversky and Kahneman's Linda problem setting (Tversky and Kahneman 1983).

In the Linda problem it seems to be the case that subjects are amenable to the so-called conjunction fallacy: subjects are told a story specifying a particular profile about the bank teller Linda. Then, statements about Linda are shown and subjects are asked to order them according to their probability (cf. Table 3). 85% of subjects decide to rank the statement "Linda is a bank teller and active in the feminist movement" (T & F) as more probable than the statement "Linda is a bank teller" (T). This ranking contradicts basic laws of probability theory, as the joint probability of two events (T & F) is less or at most equal to the probability of each individual event. As a possible reason for the occurence of this fallacy, Tversky and Kahnemann argue that the decisive element when deciding on which statement to consider more or less probable may be representativeness: Given Linda's description, the feminist bank teller version (T & F) seems to be more representative of her than the mere bank teller option (T).

Following Tversky and Kahnemann's explanation, it has to be considered that "feminist" and "bank teller" are normally not only understood in the lexical sense (e.g., a bank teller is someone who works in a bank giving out and taking in money), but are read as prototypes or exemplars for a certain type of person, exhibiting typical properties normally associated with that class of people in a cliché-like manner. A formalized example for such a characterization of a bank teller and a feminist (based on a quick poll amongst staff of the Institute of Cognitive Science at Osnabrück[4]) can be found in Table 4.

Sorts:
person, *characteristic*, *trait*
Entities:
bt, *f* : *person*,
unemotional, *emotional* : *trait*,
accurate, *polite*, : *characteristic*,
educated, *strong-willed* : *characteristic*,
Predicates:
$has\_trait : person \times trait$,
$has\_characteristic : person \times characteristic$
Laws present in both characterizations:
$\forall x\,(has\_trait(x, unemotional) \leftrightarrow \neg has\_trait(x, emotional))$
Other laws of the bank teller characterization:
(β_1) $has\_trait(bt, unemotional)$
(β_2) $has\_characteristic(bt, accurate)$
(β_3) $has\_characteristic(bt, polite)$
Other laws of the feminist characterization:
(θ_1) $has\_trait(f, emotional)$
(θ_2) $has\_characteristic(f, educated)$
(θ_3) $has\_characteristic(f, strong\text{-}willed)$

Table 4: Example formalizations of cliché characterizations for a bank teller and a feminist.

Given these characterizations, HDTP can be used for computing a common generalization of both, yielding a generalized theory like given in Table 5.

$A : person, X : trait, Y : characteristic,$
$Z : characteristic$
(γ_1) $has\_trait(A, X)$
(γ_2) $has\_characteristic(A, Y)$
(γ_3) $has\_characteristic(A, Z)$

Table 5: Shared generalized theory of the cliché characterizations for a bank teller and a feminist.

[4]Subjects (which were taken from different research groups at the Institute of Cognitive Science, and were not part of the AI group) were asked to name the five most salient properties or character traits characterizing a bank teller and a feminist. The resulting lists of answers were standardized across subjects by replacing to a reasonable extent synonymous properties by one designated exemplar property. The three exemplar properties stated with the highest frequency for each of both types of person were then used to create the characterizations in the example.

This generalization, together with the original theories, serves as basis for a conceptual blend[5] between the distinct concepts of a bank teller and a feminist, yielding a feminist bank teller (depending on which of the multiple possible blends is returned by the system, the result will be closer to a cliché bank teller or a cliché feminist).

Tversky and Kahnemann's notion of representativeness in the HDTP framework breaks down to maximizing the number of shared elements between two theories, i.e. in this case the conceptual blend and the theory built concerning Linda when reading the description in Table 3 (a process which can be seen as related to the maximization of coverage between the generalized theory and the original domain theories used internally by HDTP when computing the generalization, cf. (Schwering et al. 2009). When doing so, it becomes obvious that each possible blend containing at least parts of the cliché feminist characterization has a higher overlap with Linda's description as does the mere cliché bank teller formalization: The *educated* characteristic fits to Linda being bright and having a degree in philosophy, the $strong-willed$ characteristic corresponds with Linda's outspokenness and the willignness to participate in demonstrations, and the *emotional* trait reappears in her being deeply concerned with issues of discrimination and social justice. On the other hand, none of the mere cliché bank teller traits or characteristics directly fits into the description.[6]

Thus, using HDTP for computing a blend between cliché characterizations of a bank teller and a feminist, and then (between the conceptual blend and a formalization of the description of Linda) performing an operation similar to the coverage maximization already internally done by HDTP when computing a generalization of two domain theories, the probabilistic conjunction fallacy found by Tversky and Kahnemann in human reasoning experiments can be explained, modeled and reproduced in a meaningful way.

Noun-Noun Combinations

Conceptual combination is the ability humans have of constructing meaning for previously unseen combinations of known words. We consider it as a special form of conceptual blending that has attracted the attention of cognitive scientists and computational creativity researchers. Such a process is commonly used by humans, and has to be performed quickly and easily in mind (Keane and Costello 2001). Conceptual combination has been extensively studied in linguistics and has, thus, a long history of experimental studies, model formalization proposals and debates in the literature. Linguists analyze, in particular, how noun-noun (e.g. BOOK BOX) and adjective-noun (e.g. RED NOSE) combinations are interpreted by humans, and propose several models of how such interpreta-

[5]As both concepts in its original cliché form are incompatible due to the constraint on the *emotional* trait, the result is a real blend, and not only a specialization or union of the two concepts.

[6]Of course, as Linda's description is only given in the form of natural language, before searching for the just outlined correspondences between theories – or the lack thereof – some automatic preprocessing has to be done.

tions could be performed (Wisniewski and Gentner 1991; Wisniewski 1997; Keane and Costello 2001; Estes 2003; Pereira 2007).

According to standard theory, a word is understood by the company of words it keeps (Firth 1957) or, according to the HDTP's jargon, by the *background knowledge* one possesses about them as well as about the context in which they appear. In the following, we confine ourselves to a specific set of noun-noun composites, where the way their meanings are interpreted by human subjects seem to be frequently encountered (Wisniewski and Gentner 1991; Wisniewski 1997; Pereira 2007), and show that HDTP has the potential to be used in computationally interpreting them.

We write a noun-noun combination in the form N="N_1 N_2", with the second noun N_2 being referred to as the *head*. The first noun N_1 functions as a *modifier* that adapts the meaning of the head. Therefore, in such cases, a combination "N_1 N_2" is interpreted by cognitive agents as a function application $N_1(N_2)$ (Wisniewski and Gentner 1991). Accordingly, we use an axiomatization of the operator N_1 as the SOURCE domain for HDTP, and an axiomatization of the head N_2 as the TARGET.

As a specific instance, consider SNAKE GLASS, which some humans described as a "tall, very thin drinking glass" (Wisniewski and Gentner 1991). Depending on the granularity of the representations (i.e. irrespective of whether or not other constituents are included in the representation), a formalization of the concept SNAKE should normally emphasize the existence of some *salient* SNAKE characteristics. A part of a suggested formalization is given in Table 6, in which the common-sense emphasis is on a SNAKE having a length that is much bigger than its width, a curved shape, and a skin that is covered in scales. The characteristics that a typical GLASS exemplar must have, among other things, are its transparency and fragility. A GLASS object also has dimensions determining its width and height. A blend of the two concepts would, consequently, import the properties of a SNAKE that do not conflict with the GLASS representation. In particular, the blend will indicate a relation between the *dimensions* of the SNAKE GLASS. That is, HDTP identifies $(1a)$ and $(1b)$ with $(2a)$ and $(2b)$, and infers from $(1d)$ that one of the dimensions of a SNAKE GLASS will be much larger than the other. A SNAKE GLASS would, in addition to the non-conflicting GLASS constituents, have a curved shape, as well as other non-conflicting constituents of a SNAKE (cf. Table 6).

Finding a meaning of a (novel) combination is a difficult task, and providing a computational account that simulates human cognition is an even more difficult one (Wisniewski 1997). Computing the blend in the described manner exemplifies our suggestion of how this form of noun-noun combinations might be handled. This framework is not functional in the sense that given two nouns it will not produce a unique result, but rather enumerate alternatives ranked by the complexity of the underlying mappings. Experiments show that humans too do not always agree on one meaning of the same given noun-noun combination, neither do they exactly follow one particular model. There exist some combinations that are unlikely to be encountered in life, such

Source Axiomatization N_1 ="SNAKE"

$$\forall x\ \exists w\ Width(x,w) \quad (1a)$$
$$\forall x\ \exists l\ Length(x,l) \quad (1b)$$
$$\forall x\ Typical_1(x) \rightarrow Shape(x, curved) \wedge Skin(x, scaled) \quad (1c)$$
$$\forall x \exists l\ \exists w\ Length(x,l) \wedge Width(x,w) \rightarrow l > w \quad (1d)$$

Target Axiomatization N_2 ="GLASS"

$$\forall x\ \exists w\ Width(x,w) \quad (2a)$$
$$\forall x\ \exists h\ Height(x,h) \quad (2b)$$
$$\forall x\ Typical_2(x) \rightarrow Transparent(x) \wedge Fragile(x) \quad (2c)$$

Blend N ="SNAKE GLASS"

$$\forall x\ \exists w\ Width(x,w) \quad (3a)$$
$$\forall x\ \exists h\ Height(x,h) \quad (3b)$$
$$\forall x\ Typical(x) \rightarrow Transparent(x) \wedge Fragile(x) \quad (3c)$$
$$\forall x\ Typical(x) \rightarrow Shape(x, curved) \wedge Skin(x, \text{scaled}) \quad (3d)$$
$$\forall x \exists h\ \exists w\ Height(x,h) \wedge Width(x,w) \rightarrow h > w \quad (3e)$$

Table 6: Parts of the noun axiomatizations and their combination. For $i \in \{1,2\}$, $Typical_i(x)$ can be defined in a variety of ways, depending on how concepts are represented.

as BOOK TIGER (Wisniewski and Gentner 1991), or have a meaning that differs from the meanings of the forming concepts (e.g. a PIT BULL is a breed of dogs). Of course, HDTP does not solve these challenges altogether, but it allows conceptual combinations in a way that respects the dual process of comparison and integration, on which famous models (in the conceptual combination literature) are generally based (Wisniewski 1997; Keane and Costello 2001; Estes 2003). The results depend on the complete representations of the nouns that are used as concepts, as well as on the (encoded) background knowledge.

Outlook

This paper attempts to show that such diverse cognitive abilities like the invention of new concepts in mathematics, the interpretation of certain noun-noun compounds in natural language, and the representation of classical rationality puzzles can be approached by a single formal framework that applies cognitive mechanisms such as metaphor and concept blending. We think that this small set of examples is just a tiny part of the large applicability of the mentioned mechanisms. In general, science seems to be a rich source of challenges for applying the formal methods presented. Here are some examples taken from physics.

- Einstein's relativity theory shifts the interpretation of certain measurement units from constants to variables. Mass and length of objects as well as time itself are no longer considered as invariants between different inertial systems, but as variable units (functions) that depend on the velocity of a moved inertial system (relative to an unmoved inertial system). This concept change can be modeled by a concept blend where parts of the original meaning of measurement units is blended with ideas from Lorentz ether theory, in particular, by the introduction of the factor $\sqrt{1-v^2/c^2}$.
- In a more concrete scenario, the principle of relativity has

as a consequence that it is impossible to determine absolutely the velocity of an inertial system for someone in this system, but only the relative velocity of this system to another inertial system. Assume you are sitting in a train and the train is accelerating slowly such that you do not recognize that you gain speed. If you see another train on the other track (and you see only this train and nothing else), it is impossible to determine whether this train or your train (or both trains) is (are) moving. In order to conceptualize this experience, it is necessary to blend two experiences: one in which you are sitting in the first train and one in which you are sitting in the second train. The indistinguishability of the two experiences yield to a relativized concept of movement.

- A classical example for blending two concepts in physics are approaches which merge two (sometimes competing) theories. Famous in this respect are relativistic versions of field equations. Examples are the Klein-Gordon equation or the Dirac equation.

There are many other examples that have been proposed for concept blending. A rich source is the problem solving domain, e.g. the Buddhist Monk puzzle (Goguen 2006) just to mention one famous example of such puzzles. Examples from physics, in problem solving tasks, or other domains just hypothesize that a conceptualization of certain new concepts can be modeled by concept blending in the present framework. A careful explanation of the underlying principles is necessary to decide whether such a modeling is sustainable for this broad applicability.

References

Alexander, J. 2009. Mathematical blending. Working paper, Case Western Reserve University.

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.

Besold, T. R.; Gust, H.; Krumnack, U.; Abdel-Fattah, A.; Schmidt, M.; and Kühnberger, K.-U. 2011. An argument for an analogical perspective on rationality & decision-making. In van Eijck, J., and Verbrugge, R., eds., *Proceedings of the Workshop on Reasoning About Other Minds: Logical and Cognitive Perspectives (RAOM-2011), Groningen, The Netherlands*, volume 751 of *CEUR Workshop Proceedings*, 20–31. CEUR-WS.org.

Estes, Z. 2003. A tale of two similarities: comparison and integration in conceptual combination. *Cognitive Science* 27:911–921.

Evans, J. 2002. Logic and human reasoning: An assessment of the deduction paradigm. *Psychological Bulletin* 128:978–996.

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

Fauconnier, G., and Turner, M. 2008. Rethinking metaphor. In Gibbs, R., ed., *Cambridge Handbook of Metaphor and Thought*. New York: Cambridge University Press. 53–66.

Firth, J. R. 1957. *Papers in linguistics 1934–51*. Oxford University Press.

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

Gigerenzer, G.; Hertwig, R.; and Pachur, T., eds. 2011. *Heuristics: The Foundation of Adaptive Behavior*. Oxford University Press.

Goguen, J. 1999. An introduction to algebraic semiotics, with application to user interface design. In *Computation for Metaphors, Analogy, and Agents*, volume 1562 of *Lecture Notes in Computer Science*. Springer. 242–291.

Goguen, J. 2006. Mathematical models of cognitive space and time. In Andler, D.; Ogawa, Y.; Okada, M.; and Watanabe, S., eds., *Reasoning and Cognition: Proc. of the Interdisciplinary Conference on Reasoning and Cognition*, 125–128. Keio University Press.

Griffiths, T.; Kemp, C.; and Tenenbaum, J. 2008. Bayesian models of cognition. In Sun, R., ed., *The Cambridge Handbook of Computational Cognitive Modeling*. Cambridge University Press.

Guhe, M.; Pease, A.; Smaill, A.; Schmidt, M.; Gust, H.; Kühnberger, K.-U.; and Krumnack, U. 2010. Mathematical reasoning with higher-order anti-unifcation. In *Proceedings of the 32st Annual Conference of the Cognitive Science Society*.

Guhe, M.; Pease, A.; Smaill, A.; Martinez, M.; Schmidt, M.; Gust, H.; Kühnberger, K.-U.; and Krumnack, U. 2011. A computational account of conceptual blending in basic mathematics. *Journal of Cognitive Systems Research* 12(3):249–265.

Keane, M. T., and Costello, F. J. 2001. Setting limits on analogy: Why conceptual combination is not structural alignment. In Dedre Gentner, K. H., and Kokinov, B., eds., *The Analogical Mind: A Cognitive Science Perspective*. Cambridge, Massachussets: MIT Press. 172–198.

Krumnack, U.; Gust, H.; Schwering, A.; and Kühnberger, K.-U. 2010. Remarks on the meaning of analogical relations. In Hutter, M.; Baum, E.; and Kitzelmann, E., eds., *Artificial General Intelligence, 3rd International Conference AGI*. Atlantis Press. To appear.

Lakoff, G., and Núñez, R. 2000. *Where Mathematics Comes From: How the Embodied Mind Brings Mathematics into Being*. New York: Basic Books.

Osborne, M., and Rubinstein, A. 1994. *A Course in Game Theory*. MIT Press.

Pereira, F. C. 2007. *Creativity and AI: A Conceptual Blending Approach*. Applications of Cognitive Linguistics (ACL). Mouton de Gruyter, Berlin.

Plotkin, G. D. 1970. A note on inductive generalization. *Machine Intelligence* 5:153–163.

Schwering, A.; Krumnack, U.; Kühnberger, K.-U.; and Gust, H. 2009. Syntactic principles of Heuristic-Driven Theory Projection. *Journal of Cognitive Systems Research* 10(3):251–269.

Stedall, J. 2008. *Mathematics emerging:a Sourcebook 1540–1900*. Oxford: Oxford University Press.

Tversky, A., and Kahneman, D. 1983. Extensional versus intuitive reasoning: The conjunction fallacy in probability judgement. *Psychological Review* 90(4):293–315.

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

Wisniewski, E. J., and Gentner, D. 1991. On the combinatorial semantics of noun pairs: Minor and major adjustments to meaning. In Simpson, G., ed., *Understanding Word and Sentence*. Elsevier Science Publishers B.V. (North-Holland).

Wisniewski, E. J. 1997. When concepts combine. *Psychonomic Bulletin & Review* 4(2):167–183.