Creating Analogy-Based Interpretations of Blended Noun Concepts

Ahmed M. H. Abdel-Fattah and Ulf Krumnack
\{ahabelfatta | krumnack\}@uni-osnabrueck.de
Institute of Cognitive Science, University of Osnabrück,
Albrechtsstr. 28, 49076, Osnabrück, Germany.

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.

1 Motivation and Background
It has long been interesting how people create meanings for new compositions of arbitrary words, such as adjective-noun (e.g. RED NOSE) or noun-noun (e.g. BOOK BOX) compound forms. Humans possess an amazing ability of creating meaningful interpretations of novel compounds, even if the interpretations distantly differ from that of the already-known words they comprise (e.g. BUTTERFLY MILK).

A seemingly implausible suggestion to tackle the problem using artificial systems might be to build a dictionary module that defines noun compounds. Beside being language-dependent and unadaptable, such a (static) dictionary becomes even more a next-to-impossible task to achieve when ‘novel’ compounds are to be considered. More plausibly speaking, a module in a cognitively-inspired model of computational creativity may be suggested to interpret previously-unseen compounds à la human, but this necessitates deeper insights into how humans themselves develop interpretations of the compounds, based on their already-existing knowledge of the composing nouns.

The connection being made here concerns simulating the development of knowledge in a proposed model, though the nature of the interpretation construction problem has many applications in various domains (e.g. in natural language processing (NLP) and information retrieval (IR) (Gay and Croft 1990)). Investigating the way humans interpret previously-unseen compounds helps us endow cognitive agents in such a model with a simulated mechanism of interpretation. The current article tries to contribute to this interpretation mechanization by surveying the problem challenges from a cognitive science perspective, in order to suggest a plausible cognitively-inspired mechanization. This is achieved through developing concept representations that give the interpretations of the nouns and their compositions.

As concepts are essential entities in representing and building the knowledge of cognitive agents, nouns are represented in the proposed model as concepts that develop by knowledge acquisition and revision, aiming at ultimately allowing the model to employ cognitive processes in feasibly simulating forms of creativity and general intelligence (more on the thematic research trace is already given in (Abdel-Fattah, Besold, and Kühnberger 2012; Abdel-Fattah et al. 2012)). Note that cognitive psychologists often use the term conceptual combination to refer to what linguists might refer to as compound nominals. Here, both are referred to as conceptual compounds, or simply compounds.

The article is organized as follows. In the following section, we give a condensed overview of the problem challenges found in the literature, then quickly explain knowledge representation aspects of the proposed cognitively-inspired model. In addition, we address some modeling notions that simulate the development of compound interpretations using the already-existing knowledge in the proposed model. The article next expounds a way of realizing the aspects in an analogy-based framework and utilizing the notions in representing, whence interpreting, the compounds. The needed methodology to be employed, namely conceptual blending, will also be summarized. Later, a method to interpret novel noun-noun compounds is illustrated with a detailed example. The article concludes with some final remarks.

1.1 Problem Importance and Challenges
It is hard so far to find a thorough, cognitively-based account that models the interpretation problem of compounds, although both its importance and extreme difficulty are implied by the literature dedicated to solving it (Butnariu and Veale 2008; Gagné 2002; Gagné and Shoben 1997; Hampton 1997, to mention just a few). Proposals showing how interpretations (of concept combination) might be performed by humans can be found in (Mareschal, Quinn, and Lea 2010; Estes 2003; Keane and Costello 2001; Wisniewski and Gentner 1991, for instance). Limitations prevent us from
delving into the aforementioned cognitive science literature, but its influence on inspiring the design of the proposed model will be heightened as needed.

The meaning of a new compound depends not only on the corresponding meanings of the composing words, but also on the particular uses of such meanings, the surrounding context, and an implicit relationship between the composing words. The latter is a main challenge because the implicit relationships between a modifier and a head are extremely difficult to abstract. E.g. compare what “WOUND” contributes to in “HAND WOUND”, to what it contributes to in “GUN WOUND” (cf. (Coulson 2006; Hampton 1997)). A compound does not equal the sum of its parts, and its meaning is as sensitive to arbitrary changes as its underlying concepts, which can themselves develop over time. Contexts and artificial anecdotes, from which a deviated meaning can be inferred, influence the background of a person. Existing experiences also influence one’s comprehension (e.g. a DECOMPOSING COMPOUND to a chemist may differ from that to a linguist (Gagné 2002)). Some proposals related to deeper analyses in the literature are listed next. Beside acknowledging previous work, they help in further clarifying the inherently baffling nature of the problem, and further illustrate that no agreement among researchers on a ubiquitous solution exists. In a conceptual compound, comprehension requires the presence of relational inferences between the concepts. For instance, nine recoverably deletable predicates are given by (Levi 1978) to characterize the semantic relationships between the composing nouns (Coulson 2006). The abstract relations theory also indicates a limited number of predicates to relate a modifier with a head (Gagné and Shoben 1997). The dual process model (Wisniewski 1997) claims that attributive and relational combination are two distinct processes resulting from comparison and integration, respectively, whereas other linguistic models raise the possibility that a single-process integration model could account for all concept combinations (Estes 2003; Gagné 2002). Other works could also be mentioned (e.g. the composite prototype model of James Hampton, and the constraints theory of (Costello and Keane 2000)), but the conclusion remains valid: the challenge is hard (and there is no consensus).

Butnariu and Veale (2008) presents a concept-centered approach to interpret a modifier-head compound, where the acquisition of implicit relationships between the modifier and the head is captured by means of their linguistic relational possibilities. Unlike others, the approach indeed is concept-centered but, unlike ours, it is linguistic-oriented and English-based. In fact, the blending approach presented later in this paper (cf. section 3.2) does not use relational possibilities by means of both the modifier and the head. The outlined model simulates the emergence an interpretation of an unknown composition undergoes by employing cognitive mechanisms, such as analogy and conceptual blending, and by prioritizing past experiences (cf. section 2) in suggesting the relational inference.

Figure 1: The four-space model of CB: common parts of the SOURCE and TARGET concepts are identified, defining a GENERIC space and a BLEND.

1.2 Integration of Knowledge by Concepts

Conceptual blending (CB) has been proposed as a mechanism to facilitate the creation of new concepts by a constrained integration of available knowledge. Whence, CB is sometimes called ‘conceptual integration’. It operates by mixing two input domains, the mental spaces, to form a new one that depends on the mapping identifications between the inputs. The new domain is called the blend, which maintains partial structures from both inputs and presumably adds an emergent structure of its own.

The first step in generating a blend is the composition step, in which selective constituents from the input spaces are paired. In the second step, the emergence, a pattern in the blend is filled when structure projection matches long-term memory information (i.e already-existing knowledge). The actual functioning of the blend comes in the third step, the elaboration, in which a performance of cognitive work within the blend is simulated according to its logic (Coulson 2006; Fauconnier and Turner 2002).

Figure 1 is the prototypical model of CB, in which two concepts, SOURCE and TARGET, represent two input spaces. Common parts of the inputs are matched by identification, where the matched parts may be seen as constituting a GENERIC space. The BLEND space has an emergent structure that arises from the blending process and consists of some matched and possibly unmatched parts of the input spaces.

CB proved its usefulness in explaining cognitive phenomena and creating new theories (Fauconnier and Turner 2008; Coulson 2006; Fauconnier and Turner 2002). The formalization of the aspects of CB is expected to produce a significant development in artificial intelligence (AI) in general and in models of computational creativity in particular (Abdel-Fattah, Besold, and Kühnberger 2012). A challenge for the formalization of aspects of CB is always raised by the optimality principles, which are the guideline pressures that should derive the generation of a feasible blend and distinguish good blends from bad ones (Pereira and Cardoso 2003; Fauconnier and Turner 2002).

2 A Cognitively-Inspired Model

The model borrows notions from nature-inspired intelligence and belief revision to simulate the development in the
acquisition of knowledge. The needed assumptions are explained below.

2.1 Acquisition of Knowledge in Concepts

We explain a way in which our model helps its agents acquire, organize, and develop their knowledge using concepts, aiming at imitating aspects of knowledge emergence in humans that are based on their ability of acquiring facts from the environment, and using existing experiences in refining their understanding of repeatedly-encountered conceptions.

In a model like this, agents generally build and manipulate their knowledge base (KB) using a knowledge representation (KR) framework that stores world facts as (organized) beliefs in frames, which are expressed in a formal language (e.g. first-order logic). More frequently used knowledge parts will be reinforced, establishing a kind of experience which links to other knowledge where they are of use; whereas knowledge that is not in use will typically be less remembered.

In our model, the KB is denoted by $K_B$. An acquired experience is a knowledge frame, $F_i \subset K_B$, which can be thought of as a network of closely connected beliefs. The set of concept names $K_C$ is the model’s lexicon that contains the names that shorthand the concept representations, where a representation is a collection of frames, denoted $F^C = \bigcup F_i \subset K_B$. The sets $K_B$ and $K_C$ have different essences: the former is a set of beliefs organized in frames and expressed in the KR’s formal language, while the latter is a set of strings formed using an arbitrary alphabet (but for each concept (name) $c \in K_C$ there is a concept (representation) $F^c = \bigcup F_i \subset K_B$). A way in which beliefs are organized into concepts is widely explained in (Wang and Hofstadter 2006), where the system given there constitutes a model of categorization that is akin to ours, although it uses another representation formalism.

2.2 Development of Knowledge Concepts

Agents assign entrenchment values to their beliefs, which serve as mnemonics of belief occurrences and rank them according to importance and frequency: they depend on how recently, and how many times, the beliefs have been retrieved from the agent from the KB (e.g. when beliefs are retrieved in a concept formation process). The entrenchment value of a belief $b \in K_B$ contributes, in turn, to calculating the entrenchment level of any $c \in K_C$ with $b \in F^c$. In other words, the entrenchment values of the beliefs that underlie the ‘representation’ of a concept, contribute to the entrenchment level of its concept ‘name’. Thus, a simulation of knowledge development is done by mimicking the functioning of “pheromone trails” (see (Dorigo and Stützle 2004)) in building and organizing the knowledge frames of the concepts, and complying with the notions given by (Gärdenfors 1988) of not giving up beliefs that have high epistemic entrenchment.

The (overloaded) function $e_V : K_B \cup K_C \rightarrow [0, 1]$ is used for indicating both the entrenchment value of $b \in K_B$, and the entrenchment level of $c \in K_C$. A concept $c \in K_C$ is called a HELCO if $e_V(c) \geq \eta$ and is called a LEVCO otherwise, where $0 < \eta < 1$ is a threshold value. Some $c \in K_C$ are considered ‘innate’ (i.e. built-in concepts), with entrenchment level $e_V(c) = 1$. Others obtain by concept formation, as is the case in concept blending (cf. section 1.2), where $e_V(c)$ is initially less than $\eta$ in such a case.

In the described model, the interpretation of a novel compound by means of already-known nouns transfers to the process of forming new LEVCOs $c \in K_C$ and their corresponding representations $F^c \subset K_B$ by the conceptual blending of already-existing HELCOs. When HELCOs combine, LEVCOs result with entrenchment levels that depend on those of the composing HELCOs: for a recently blended LEVCO, $B \in K_C$, its entrenchment level $e_V(B)$ is a function in $e_V(S)$ and $e_V(T)$ of the composing HELCOs, $S, T \in K_C$. Details are not given in the current presentation about how these values are computed, rather about how the blends may develop.

3 Developing Conceptual Interpretations

An essential assumption taken by all blending approaches is the organization of knowledge in some form of domains. The design of the model’s KB is inspired by the language-of-thought hypothesis (Fodor 1983). Facts, frames, and concepts are therefore provided in modular groups to serve the blending process. Enough HELCOs that represent the already-known nouns are available at the agents’ disposal in the model to obey the case for humans, where “a person has a repertoire of available concepts and ways of combining those concepts into higher-order concepts and into propositions” (Hampton 1997).

3.1 The Model’s Framework

Heuristic-Driven Theory Projection (HDTMP) is used as the underlying modeling framework (Schwer et al. 2009). HDTMP is a mathematically sound framework for analogy making, together with the corresponding implementation of an analogy engine for computing analogical relations between two logical theories, representing two domains. Domain theories are represented in HDTMP as sets of axioms formulated in a many-sorted, first-order logic language. HDTMP applies restricted higher-order anti-unification (Krummack et al. 2007) to find generalizations of formulas, and subsequently proposes analogical relations between the source and target domains, that can later be used as basis for an analogy-based transfer of knowledge between the two domains (see Figure 2). The analogical transfer results in structure enrichment of the target side, which usually corresponds to the addition of new axioms to the target theory. Analogical transfer is desired in CB in order to create a new enriched domain while keeping the original target domain unchanged. In such cases the generalization, source, target, and enriched domains are interconnected by a blend. (For an expanded elaboration of HDTMP and its application domains see (Abdel-Fattah, Besold, and Kühlberger 2012; Abdel-Fattah et al. 2012; Martinez et al. 2011).)

3.2 Concept Blending using HDTMP

We are confined to handling modifier-head noun compounds, where such a combination is written in the form
The modifier adapts the meaning of the head and the noun-noun combination “SNAKE” can thus be interpreted by agents as a function application \(S(T)\), because \(S\) acts, in a sense, as an operator on \(T\) that, more or less, changes \(T\)’s meaning. The following method does not use function application, rather CB.

An axiomatization of the operator \(S\) is used as the SOURCE domain for HDTTP, and an axiomatization of the head \(T\) as the TARGET. When HDTTP is applied to the inputs, blends result that give possible interpretations of the compound. The transfer of knowledge, during analogical reasoning, is allowed in only one direction to pave the way for the “composition” and “emergence” steps of CB to come into play (cf. section 1.2).

Once \(S\) and \(T\) are appropriately represented in sorted, first-order logic, the blending starts by providing them to HDTTP, where an analogy is established and an explicit generalization, \(G\), is computed (cf. Figure 2), which can be a base for concept creation by abstraction. HDTTP proceeds in two phases: (1) in the mapping phase, \(S\) and \(T\) are compared to find structural commonalities (corresponding to the ‘identification’ between SOURCE and TARGET shown in Figure 1), and a generalized description is created that subsumes the matching parts of both domains, and then (2) in the transfer phase, unmatched knowledge in the source domain is mapped to the target domain to establish new hypotheses \(B_i\). Additional types of implicit relationships between the modifier and the head may later be suggested and established during the transfer phase.

3.3 Compound Interpretation: An Example

Consider the combination SNAKE GLASS, described by many human subjects in (Wisniewski and Gentner 1991) as a “tall, very thin drinking glass”. The example given below illustrates a possible blend of (partial formalizations) of the domains representing the source and target nouns SNAKE and GLASS, respectively (cf. Table 1). The blended domain, SNAKE GLASS, is an expansion of GLASS, the target, in which notions of ‘shape’ and ‘skin’ frames, taken from SNAKE, are added. In principle, the blended domain can be thought of as coming from enriching the first-order theory by which the target is represented with new notions (or frames) taken from the source, and then importing the axioms of the source into it (cf. Figure 3).

A formalization of a SNAKE concept should normally emphasize some salient SNAKE characteristics. A suggested formalization is given in Table 1, in which the common-sense emphasis is on a SNAKE having a length that is much bigger than its width, a curved body shape, and a skin that is covered in scales. Also, the characteristics that a typical GLASS exemplar must have, are its transparency and fragility. A GLASS object has dimensions determining its width and height. Note that, in our model, the represented characteristics of a given concept become more salient when their corresponding beliefs are more reinforced most of the (more recent) times the concept is retrieved. A blend (of the two concepts that represent SNAKE and GLASS) would, consequently, import the properties of a SNAKE that do

\[
\begin{align*}
\text{Source Axiomatization} & \quad S = “\text{SNAKE}” \\
\forall x \exists w \text{ Width}(x, w) & \quad (1a) \\
\forall x \exists l \text{ Length}(x, l) & \quad (1b) \\
\forall x \text{ Typical}_1(x) & \rightarrow \text{ Shape}(x, \text{curved}) \land \text{Skin}(x, \text{scaled}) \quad (1c) \\
\forall x \exists l \text{ Width}(x, l) \land \text{Width}(x, w) \rightarrow l > w & \quad (1d) \\
\text{Target Axiomatization} & \quad T = “\text{GLASS}” \\
\forall x \exists w \text{ Width}(x, w) & \quad (2a) \\
\forall x \exists h \text{ Height}(x, h) & \quad (2b) \\
\forall x \text{ Typical}_2(x) & \rightarrow \text{Transparent}(x) \land \text{Fragile}(x) \quad (2c) \\
\text{Blend} & \quad B_i = “\text{SNAKE GLASS}” \\
\forall x \exists w \text{ Width}(x, w) & \quad (3a) \\
\forall x \exists h \text{ Height}(x, h) & \quad (3b) \\
\forall x \text{ Typical}(x) & \rightarrow \text{Transparent}(x) \land \text{Fragile}(x) \quad (3c) \\
\forall x \text{ Typical}(x) & \rightarrow \text{Shape}(x, \text{curved}) \land \text{Skin}(x, \text{scaled}) \quad (3d) \\
\forall x \exists h \text{ Height}(x, h) \land \text{Width}(x, w) \rightarrow h > w & \quad (3e)
\end{align*}
\]

Table 1: Parts of suggested noun axiomatizations and a possible combination.

Figure 2: HDTTP’s overall approach to creating analogies and CB. \(S\) and \(T\) are source and target input theories, \(m\) represents the analogical relation between \(S\) and \(T\), and \(G\) is the generalization computed by anti-unifying \(S\) and \(T\). The dashed arrows \(S \rightarrow B\) and \(T \rightarrow B\) describe the (partial) injections of facts and rules from source and target to the blend space.

Figure 3: The form of the noun-noun blend that results from the transfer phase (cf. Table 1).
not conflict with the GLASS representation. In particular, a blend will indicate a relation between the dimensions of the SNAKE GLASS. In Table 1 and Figure 3, one can see that 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.

It is worth noting that the given framework does not function in the sense that two given nouns will only (or always) produce a unique result. 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 each time they encounter a similar combination (Mareschal, Quinn, and Lea 2010; Winsiewski 1997; Winsiewski and Gentner 1991). The framework rather enumerates alternative blends ranked by the complexity of the underlying mappings. Every SNAKE GLASS blend is intended to be represented (i.e. interpreted) by a LEVCO \(B_i \in \mathbb{R}_C\) with \(0 < e_V(B_i) < \eta\), such that the calculation of \(e_V(B_i)\) is affected by \(e_V(S)\) and \(e_V(T)\) of the source and target HELCOS. In our view, this is a desirable property because it (1) allows possible interpretations instead of just one, and (2) leaves space for experience to play a role in deciding whether or not a specific blend is favored over another. People also interpret novel combinations by drawing on past experience with similar combinations (Gagné 2002).

4 Conclusion and Final Remarks
Humans employ cross-domain cognitive mechanisms, especially analogy-making and concept blending, in developing their understanding of newly introduced conceptions that are basically combinations of already-known ones. Inspired by this claim, the article shows how it is possible to propose a computational model of creativity that employs both mechanisms. Finding a meaning of a (novel) combination is a difficult creative task, yet providing a computational account that simulates the task in human cognition is an even more difficult one, but it seems realizable. The challenges of the problem incite this contribution to solving it by proposing a computational, concept-based, cognitively-inspired, logic-based, and language-independent approach. The feasibility of computing a blend in the described manner exemplifies our cognitively-inspired suggestion of how this form of noun-noun combinations could be approached.

We argue that the presented method amounts to the development of conceptual interpretations of newly created concepts, where the emulation is inspired by the utilization of cognitive capabilities in creating analogies (Gentner, Holyoak, and Kokinov 2001). While the interpretation (of a new modifier-head compound) is evolving, a meaning is ‘invented’ online using a cross-domain creation process in which a virtual copy of the head is first imagined, which is analogical to the modifier in some sense. Then, particular traits of the modifier are picked up and added to this copy. The newly-created interpretation can be a combination of the characteristics of the two words appearing in the compound, depending on how much in common the two words have and on the existing background. Using our model’s terms, the frames that define the newly-created concept result from blending the salient frames defining the composing concepts. The resulting features depend on the organized beliefs of the modifier and head concepts, on the experiences, and on how a head may “look like” when it is attributed to the modifier (e.g. how may a BOX look like when it is attributed to a BOOK in the compound BOOK BOX, how may a GLASS look like when it is attributed to a SNAKE in the compound SNAKE GLASS, . . . etc.). Note that the “saliency” of a concept’s feature or trait results from enforcing, and re-enforcing, repeated experiences that are related to the concept’s defining frames. We believe that this agrees with people’s continuous re-conceptualization of their understandings of words as they are encountered over time (and in different contexts).

The principles given by (Keane and Costello 2001; Costello and Keane 2000), the developmental psychology literature in (Mareschal, Quinn, and Lea 2010; Lamberts and Shanks 1997, for instance), and the studies and experimental results of (Wisniewski 1997; Winsiewski and Gentner 1991) provide further motivation (and support from research in cognitive science) why a combination is proposed to evolve in this way. The method allows a form of blending that respects the dual process of comparison and integration, on which famous models are based (cf. (Estes 2003; Gagné 2002; Keane and Costello 2001; Levi 1978)), yet relational possibilities may only be suggested by the modifier, which is the source concept in our case (cf. (Gagné and Shoben 1997; Winsiewski 1997; Winsiewski and Gentner 1991)).

From a modeling point of view, the way analogy is made use of in identifying common parts of the source and target concepts of a modifier-head compound, in generalizing them, and creating blends, serves maintaining relational and attributive combinations at the same time. However, the implicit relational possibility that analogy provides us with between the head and the modifier still does not account for many of the different cases that can be encountered, though it is promising and could be improved by using the relationships between the underlying frames of the given concepts.

The method presented here is considered a first starting step towards the interpretation of noun-noun compounds using a new analogy-based perspective. It presumably overcomes some representation challenges that are usually faced in designing cognitively-inspired models of computational creativity; the model is promising and can be used in other applications. The encoding of rated experiences and the use of levels of entrenchment for the concepts can help in achieving solutions to other challenges, such as when concepts get changed or externally affected by newly observed facts.

References


