Compositionality Meets Belief Revision: a Bayesian Model of Modification

Corina Strößner¹

Published online: 9 May 2020
© The Author(s) 2020

Abstract

The principle of compositionality claims that the content of a complex concept is determined by its constituent concepts and the way in which they are composed. However, for prototype concepts this principle is often too rigid. Blurring the division between conceptual composition and belief update has therefore been suggested (Hampton and Jönsson 2012). Inspired by this idea, we develop a normative account of how belief revision and meaning composition should interact in modifications such as “red apple” or “pet hamster”. We do this by combining the well-known selective modification model (Smith et al. *Cognitive science* 12(4):485–527 1988) with the rules of Bayesian belief update. Moreover, we relate this model to systems of defeasible reasoning as discussed in the field of artificial intelligence.

1 Introduction

The prototype theory of concepts, which goes back to various influential papers by Eleanor Rosch and colleagues (cf. Rosch 1973, 1978; Rosch et al. 1976; Rosch and Mervis 1975), emphasises that most natural language concepts are not captured by universal definitional features but in terms of similarity to salient, that is, typical exemplars, or in terms of typical properties, which most category members possess. Note that the notion of typicality is ambiguous between an extensional and an intensional reading. It can be applied to category members (a typical fish) or properties (a typical shape of fish). The grammatical context makes clear what is meant. In

This research was generously funded by the German research foundation DFG (SFB 991, D01).

✉ Corina Strößner
stroessner@phil.hhu.de

¹ Duesseldorf Center for Logic and Philosophy of Science, Heinrich Heine University Düsseldorf, Building 24.52, Universitätsstraße 1, 40225 Düsseldorf, Germany

our own investigation, we will be concerned with the typicality of properties, that is, intensional typicality.

There are several versions of prototype theory and there is no unique understanding of what a prototype is. Rosch (1973, 330) defines “natural prototypes” as salient exemplars around which categories are organised. This is plausible with regard to perceptual categories (such as colours and shapes). Later Rosch extends the notion: “Other natural semantic categories (e.g., categories such as “fruit,” “bird”), although unlikely to possess a perceptually determined natural prototype, may well have artificial prototypes” (Rosch 1973, 349). Nowadays, these “artificial prototypes” are usually understood as structures that distinguish central features of a concept’s instances, which can be represented as feature list but also in other representational models (cf. Hampton and Jönsson 2012, 386). One can, for example, use functional attribute value structures like Smith et al. (1988) in their well-known selective modification model or Barsalou (1992) in his frame account. These functional structures use attributes (COLOUR, SHAPE, and so on) and assign appropriate values (“red” for COLOUR, “round” for SHAPE, for instance). Attributes are always functional. That means, every entity takes exactly one value. In functional models, features are not just listed, but structured by the attributes. For example, an unwed man is characterised by having the value “male” on SEX, and “unwed” on MARITAL STATUS.

Osherson and Smith (1981) were the first to point out a major problem with prototype theory when it comes to building complex categories. For example, a guppy is an atypical fish and an atypical pet but a typical pet fish. Prototypes, it seems, can violate the principle of compositionality, henceforth PC. This principle is usually associated to Frege’s philosophical papers, published in Frege (2007), and can be phrased as follows:¹

The meaning of complex phrases must be determined by the meaning of their basic components and the way they are composed.

For example, a sentence derives its meaning from its syntax and the meaning of its constituent words. Applied to *concepts*, compositionality means that the content of composed concepts is determined by the atomic concepts and the way they in turn are composed. It has been argued that compositionality is the best explanation for the ability to form and understand a multitude of novel phrases. Therefore, it is also viewed as an important desideratum for a successful theory of concepts (cf. Laurence and Margolis 1999). Prototypes, however, do not compose in an obvious way, as frequently pointed out by Jerry Fodor, who attacked the prototype view of concepts on these grounds (Fodor 1998; Fodor and Lepore 1996; Connolly et al. 2007): How can concepts be prototypes if they cannot be combined in a clear way?

In response, Hampton and Jönsson (2012) argued that prototype structures are indeed not fully compositional in the strict sense and proposed a weakened rule of compositionality, PC*:

¹ Compositionality is often called Frege’s principle. However, the principle was never formulated by him and it is not even clear whether he subscribed to such a claim (cf. Pelletier 2001).

The content of a complex concept is completely determined by the contents of its parts and their mode of combination, *together with general knowledge*. (Hampton and Jönsson 2012, 386)

Jönsson and Hampton argue that prototypes are weakly compositional, in the sense of PC*. They claim that in the case of prototype concepts, it is not the set of typical instances (extension) that is combined, but rather the information on the prototype, which yields a set of typical instances of the combined concept that is usually different from an intersection of the component extensions. Let us again consider “pet fish” as a toy example. The pet prototype gives the information that pet fish are likely to be found in homes and the fish concept contributes the information that pet fish live in water. This makes the feature “live in aquariums” typical for pet fish. The typical pet fish are thus animals that we find in aquariums, such as gold fish or guppies.

Like Hampton and Jönsson (2012), we base our discussion on the idea that semantic composition and the merging of information should go together. However, we deviate from their illustration in several aspects. First, they emphasise the extensional question about (vague) category membership: which entities will be perceived as (typical) members of composed concepts? They suggest approaching the question by first merging the prototype information. We will focus on the merging process itself and disregard the extensional question. Second, Hampton and Jönsson (2012) do not suggest a formal model with computable results. Rather, they describe several empirical aspects of concept composition as a non-logical form of reasoning. In contrast, we will suggest a normative model for merging information from a probabilistic point of view. This model might depart from *empirical* results but can be connected with *normative* approaches in artificial intelligence.

This paper will focus on modification – a very common form of concept combination. In our work, modification will be broadly understood as an operation that builds a subcategory from a given category by specification of a property. In natural language, this is commonly realised in adjective noun compounds like “black bird”, “fresh apple”, or “new house”, or in noun noun compositions such as “language school”, but is also found in relative clauses such as “a school where languages are taught”. Modifications are an especially salient case for compositionality: they are highly common, comparatively simple, and empirically well-researched (cf. Connolly et al. 2007; Gagné and Spalding 2011, 2014; Hampton et al. 2011; Jönsson and Hampton 2006, 2012; Spalding and Gagné 2015).²

In a very simplified formal understanding, modifications build set-theoretic intersections. If *Black* is a set with all black things and *Bird* is a set of birds, then the meaning of “black birds” is just $Black \cap Bird$. This view was also a source of the reservations against prototype composition in Osherson and Smith (1981). However, the intersection approach to modification has many limitations. Not only does it fail in privative modifications (such as “fake money”), but also in the large class of cases

²The key result of these studies is that modifiers influence typicality ratings in a very robust way. Subjects rate ascriptions of typical properties (“Ravens are black”) as highly likely to be true. However, the likelihood decreases if the noun is modified. The decrease is more pronounced for atypical modifiers (“Jungle ravens are black”) than for typical ones (“Feathered ravens are black”).

in which modifiers are relative to the nouns (such as “old” in “old castle” versus “old tablet”). The conjunctive reading of modification has therefore lost its prominence in formal semantics (Pelletier 2017).

The selective modification model of Smith et al. (1988) (henceforth SMM) is a more sophisticated approach to modification, which is specially construed for prototype concepts. It represents prototypes as weighted attribute value structures of categories. That is, prototype concepts (such as “strawberry”) provide information about the importance of an attribute (for instance COLOUR) and the prominence of particular values (such as “white”, “green”, “red”).

The basic plan of the current paper is to enrich the SMM with rules for Bayesian belief update. We elaborate a *normative* model of composition with background beliefs, based on the rules of probability theory. Though we discuss some recent empirical results in the conclusion, our primary goal is not to explain how humans actually deal with modifications. Rather, we aim to contribute to a normative account of prototype reasoning. Such work is still useful for the empirical investigation of prototype composition insofar as it delineates the *rational aspect* in the likelihood ratings of modified and unmodified sentences. It can offer a normative background against which empirical results can be compared. For example, Strößner and Schurz (2020) investigate the extent to which empirical results on modification are a result of probabilistic reasoning. Our goal is thus to provide a connection between psychological modelling and idealised belief representation in formal epistemology and artificial intelligence.

The paper proceeds as follows: in the next section, we describe SMM in more detail and reformulate it into a probabilistic model of composition. The subsequent section discusses belief constraints, formalised as conditional probabilities. We argue that the resulting Bayesian model of modification combines compositionality and belief revision in a fruitful way. In the concluding section, we reconsider the merits and limitations of our model.

2 A Probabilistic Selective Modification Model

The SMM represents concepts as attribute value structures. An attribute is a function, such as TASTE or COLOUR, which can be saturated with appropriate values, for instance “sweet” or “red”. Both, attributes and values, are furnished with numerical information. Attributes have an importance weighting, called *diagnosticity*, and values have a *salience* weighting, given in terms of votes, indicating the typicality of a particular value (cf. Smith et al. 1988, 489). According to Smith et al. (1988), votes represent subjective frequencies, biased by perceptibility, which puts them very close to subjective probabilities.³ Figure 1a shows the representation for “apple”.

In the SMM, modification is a strictly selective process, the effect of which is limited to one attribute. The modifier selects the attribute the adjective addresses, shifts all votes to its value, and increases the importance of the attribute (cf. Smith et al.

³Note that, in their experimental part, votes are numbers of mentions (Smith et al. 1988, 498). These notions do not necessarily coincide.

Fig. 1 Modification in SMM, following Smith et al. (1988, 490,494)

1988, 492).⁴ Figure 1b shows how the modifier “red” for “apple” operates as value of the COLOUR attribute: all votes go to “red” and the importance of COLOUR is increased. Similar to frames, which became widely known with the work of Barsalou (1992), Smith et al. (1988) presuppose a functional attribute value structure. Thus, an instance of a concept realises one of the possible values that are associated with the attribute. This is why the modifier takes the votes from alternative values: green apples are not red. In our view, the functional structure is very useful. It allows for modelling how a modifier directly affects the set of alternative values – for instance, how a colour modifier acts on the whole COLOUR attribute. One might object that the assumption of functionality makes the model unrealistic. Apples, for example, mostly have more than just one colour. The colour attribute in Fig. 1 can therefore only refer to the dominant colour of the apple, or more exactly, the dominant colour of its peel. Note, however, that the functionality assumption does not impel coarse-grained representations, though they are often convenient. An extremely fine-grained representation, with values such as “green with yellow sprinkles” or “red with a brown spot” is possible but in most cases needlessly complex. The functional attribute value structure is thus sufficiently flexible to allow for such variations, though, normally, simplified examples are adhered to.

Smith et al. (1988) are aware of the fact that a known dependency between the values of different attributes can influence the modification. They defer possible adjustments to a post-compositional process, which is slower and primarily driven by general background knowledge (cf. Smith et al. 1988, 525). In this respect, the model deviates from the weakened compositionality principle PC* of Hampton and Jönsson (2012), which would explicitly ask to take common knowledge into account during the composition. In the current paper, we will show how this can be done in the SMM. But before we make room for including background knowledge in the modification process, we will modify the SMM into a probabilistic model.

⁴ Note that the SMM cannot deal with privative modifications, like “stone lion” or “married bachelor”, which lead to coercion and metaphorical use. They confront us with their own obstacles, the reasons for which do not lie in the compositionality of prototypes. With a classical definition view of concepts, it is even more puzzling how competent speakers interpret “stone lion”.

2.1 Typicality and Probability

Probability-based approaches to categorisation have a longstanding tradition and have gained importance with the rising prominence of the prototype theory. According to Eleanor Rosch, the ‘founding mother’ of prototype theory, prototypes select property clusters from persisting covariances in the world (cf. Rosch 1978). The existence of such clusters as well as their cognitive representation in humans can be explained by a twofold evolutionary argument. Evolution, Schurz (2012) argues, leads to an environment replete with prototype structures. Selective reproduction leads to classes of entities with many similar features. Since there is also variation, for instance through mutation, these features will not become universal but will be very frequent. The cognitive capabilities of humans evolved in these evolutionary shaped surroundings, and prototype representations are an adaptation to this environment.⁵ From this point of view, prototype reasoning is a naturally evolved form of statistical reasoning. As such, it can be expected that typicality and probability are closely related. This is actually what Schurz (2012) claims. He distinguishes typical properties *in the wide sense* from typical properties *in the narrow sense*, both being highly frequent on account of having been selected and reproduced in the evolutionary history of the class. Typical properties *in the narrow sense* are also diagnostic because they are specific to the category and indicate category membership. Such properties evolved in the unique evolutionary history of the class and are not shared by contrast classes. For example, having a heart is only typical in the wide sense for birds; having a beak is also typical in the narrow sense. Typicality generally helps in inferring likely properties, but only typicality in the narrow sense also allows for rapid categorisation of entities.

In probabilistic terms, if V is typical in the wide sense for C , then $P(V|C)$ is high. If V is also typical in the narrow sense, $P(C|V)$ is also high. The latter component of typicality in the narrow sense is the well-known notion of cue validity (Rosch and Mervis 1975). Typicality in the wide sense is known as category validity (Murphy 1982). The notion of typicality in the narrow sense is thus a combination of these. In this respect it resembles the category-feature collocation (Jones 1983), which is the product of cue validity and category validity, though Schurz (2012) does not suggest a specific way to combine the two.⁶

Is high diagnosticity a necessary condition for properties to be found typical? If we look at the feature lists of Cree and McRae (2003), who collected frequently associated properties for common nouns, we find that this is not the case. For example, having seeds and being red are listed as typical properties of strawberries. The first is almost universal for fruits, and the second is at least frequent in contrast categories. Neither property is specific for strawberries. This speaks against a narrow reading of typicality and in favour of a wide one, that is, of an interpretation of typicality as high probability of a property. Further evidence comes from Strößner et al.

⁵Note that Schurz explicitly views culture as a domain of evolution, namely cultural evolution. The argument in favor of prototype theory, however, goes through without this assumption, since basic conceptual abilities of humans arguably evolved before cultural evolution played a major role.

⁶For an overview of the problems and merits of these measures, see Murphy (2004, 215f).

(2020), who show that there is a high overlap between probability and typicality ratings. This notion of typicality is admittedly simplified.⁷ However, this simplification not only builds a fruitful bridge to probability theory, but also acknowledges the fact that a very close connection between likelihood and typicality is already assumed in empirical research. For example, studies on the modifier effect generally use likelihood ratings when testing the degree to which typical properties are accepted in subcategories.

The formal foundation of our probabilistic approach is a finite set of possibilities D (which contains, for instance, conceivable apples with their particular tastes, colours and forms but can also contain non-existing entities like flying potatoes) and an algebra $\mathcal{A}$, where $\mathcal{A} \subseteq \mathfrak{P}(D)$ (a subset of the power set of D). Elements of $\mathcal{A}$ are, for example, sets of all green things, all potatoes, all flying things, and the set of all cognitively conceivable entities (D). Roughly speaking, the algebra $\mathcal{A}$ separates the elements of D into categories.⁸ The probability P is a function from $\mathcal{A}$ to the set of real numbers. P follows three axioms: (1) for all $V \in \mathcal{A}$: $0 \leq P(V) \leq 1$, (2) $P(D) = 1$, and (3) for a set of mutually exclusive $V_1, V_2, \dots, V_n \in \mathcal{A}$: $P(V_1 \cup V_2 \cup \dots \cup V_n) = P(V_1) + P(V_2) + \dots + P(V_n)$.⁹ The important notion of conditional probability is defined as $P(V_1|V_2) = \frac{P(V_1 \cap V_2)}{P(V_2)}$.

For the application to a frame-like structure, like the SMM, this means that a concept corresponds to an element from the algebra $\mathcal{A}$, for example, to the set of all apples. An attribute divides this set into exclusive and mutually exhaustive subcategories by specifying possible values, for example, the set of green apples, red apples, yellow apples and so on. The probability we give to these values specifies the relative probability of these properties. It is thus by definition a conditional probability of a property given the (superordinated) concept. By fitting this probabilistic model of belief representation into the functional structure of the SMM (or Barsalou's frames, respectively), we encounter a further constraint. Because the possible values $V_1, V_2, \dots, V_n$ of one attribute are mutually exhaustive and exclusive, $P(V_1|C) + P(V_2|C) + \dots + P(V_n|C) = 1$.

In most contexts, typicality is a discrete concept. That is, we usually do not measure typicality quantitatively, other than likelihood. Psychological experiments on the effect of modification for typicality statements starting from Connolly et al. (2007) also use a binary notion of typicality. A quantitative understanding of typicality, as in our probabilistic approach, should offer a bridge principle to the discrete notion of typicality. It is hard to determine a specific probabilistic threshold t a property has to meet to be perceived as typical, but it is rather uncontroversial that this threshold is higher than 0.5 and lower than 1. The first criterion ensures that contrary values of an attribute cannot both be typical and the second ensures that typicality does not trivialise to universal truth:

⁷For example, it applies to properties only, not to the typicality of category members. The fact that apples are typical fruits is much harder to capture in probabilistic terms.

⁸Since $\mathcal{A}$ is only a subset of the power set of D , not all possible collections of subsets have to be part of it. Only D , $\emptyset$ as well as complements and unions of elements of $\mathcal{A}$ are necessarily included in $\mathcal{A}$.

⁹These are the Kolmogorov axioms for (countably) additive probability functions. For a more detailed explication, see Huber (2016).

A property V is *typical* for a concept C if and only if $P(V|C) \geq t$ with $1 > t > 0.5$.

It is not self-evident what the opposite of typicality is. Typicality can be negated in two different ways: internally and externally. This makes the notion of non-typicality slightly ambiguous. “ C are not typically V ” (external negation) and “ C are typically not V ” (internal negation) are two different claims, where the latter is strictly stronger than the former. We will use the term “atypical” solely in the stronger sense, that it, as a contrary opposite of “typical”. The term “neutral” will be used for other non-typical properties:

A value V is *atypical* for a concept C if and only if $1 - P(V|C) \geq t$ with $1 > t > 0.5$.

A value V is *neutral* for a concept C if and only if it is neither typical nor atypical.

Atypical values are often in contrast with an alternative value that is typical. For example, mammals typically have legs as their means of locomotion. Therefore, fins are atypical. In this case, atypicality indicates exceptionality. However, values can also be atypical because there are so many possibilities that each is unlikely. Intuitively, we would not want to view all these properties as exceptional. The problem ultimately depends on the granularity of the representation. For example, the LOCATION attribute (for instance, as the destination of a journey) has a moderate number of values if measured in terms of continents, but very many if measured in terms of GPS coordinates. In the latter case, all of these coordinates have a rather low probability but are not unexpected.

There is an analogy here with debates in formal epistemology. The granularity of the partition is of crucial importance to the question of whether logical norms of qualitative belief, such as deductive closure, can be reconciled with probabilistic notions of subjective degrees of belief. In coarse-grained (context-dependent) partitions, this can be achieved (Leitgeb 2017). However, in more fine-grained partitions, the probabilistic threshold for belief would have to be extremely high (Schurz 2019). The determination of appropriate levels of granularity is thus an important, though often underestimated, issue in research on conceptual representations.

2.2 SMM with Probabilities

In our model, we adopt the representation of concepts in terms of attribute value structures from Smith et al.’s SMM, but instead of votes, we assign probabilities to the possible values. This step is not a drastic change from the initial SMM. One can easily gain the probabilities from the ratio of the votes among the attributes. For example, in Fig. 1, we would assign $\frac{5}{6}$ to “red” and $\frac{1}{6}$ to “green”. The skeleton of the resulting model is illustrated in Fig. 2, which represents a concept C with its attributes (A and B) and their values (V_1 , V_2 and W_1 , W_2). Note that these probabilities are already conditional on the concept C . As explained before, the probabilities that are assigned to the values of one attribute ($P(V_1)$, $P(V_2)$ for A and $P(W_1)$, $P(W_2)$ for B) always have to add up to 1 since the values are exclusive and mutually exhaustive. Whether a value, such as V_i is typical can be specified by the threshold t discussed

Fig. 2 Basic model

above. Within the attribute value structure, rules of typicality become obvious. For example, if value V_1 is typical, then any of its alternatives is atypical. As already stated, the reverse is unfortunately not the case: the existence of an atypical value does not imply that there is a typical alternative value. In principle, all values of an attribute can be atypical if there are many alternatives.

Note that we will not discuss importance weights in our probabilistic model. The attribute itself is a function and its importance can thus not be understood as a probability, simply because the attribute cannot be true. Attribute importance is not part of our (factual) beliefs about the world but about our attention. We agree with Smith et al. (1988) that modified attributes become more important. Also, our account is fully compatible with their suggested importance raising. We omit the parameters of attribute importance not because we find them insignificant but because our model does not suggest any restrictions in this respect. However, what we suppose and what is essential about giving attention to the modified attribute is that it must become part of the representation, even if it is usually not. For example, when thinking about an apple, we usually do not represent age. However, in understanding “old apple” it is inevitable to retrieve the AGE attribute for the noun “apple” first.

Modification is, *prima facie*, the same as in the selective modification model. If V_1 is the modifier, as shown in Fig. 2b, it becomes the only possible value for this attribute and the probabilities for attribute B remain the same. Thus, our model proposes that typical properties are inherited by subcategories, at least if no background knowledge intervenes. This influence of knowledge constraints is the subject of the next section.

3 Belief Constraints

3.1 Constraints

Often we know that the values of different attributes are correlated with each other. For example, most people know that the colour of a strawberry is related to its taste. The SMM does not offer the tools for capturing such beliefs. Frames, as described

Fig. 3 Positive constraint between “bumpy” and “brown”

in Barsalou (1992), have the functional attribute value structure of the SMM. Additionally, he introduces constraints to his frames to model common beliefs about dependencies between values. According to him “[t]he central assumption underlying constraints is that values of frame attributes are not independent of one another” (Barsalou 1992, 37). Barsalou distinguishes different kinds of constraints. *Value constraints* are concerned with the relation between two specific values of different attributes: climbing as the purpose of travel influences the destination, while relaxing as the purpose of travel seems unrelated to a specific location. *Attribute constraints*, on the other hand, hold across all the possible values of two attributes. For example, the speed of a chosen vehicle influences the duration of the journey *in general*. Constraints are positive or negative, depending on the direction of the correlation.

Since frames share the general attribute value structure of the SMM, constraints can easily be transferred to it. If we reconsider the example of apples, which is used by Smith et al. (1988, 494), we are not only able to name the possible colours and textures of apples, but also the belief that the texture and colour of apples coincide. Figure 3 represents the belief that brown apples are more likely to be bumpy than apples in general.

In our probabilistic model, we propose to illustrate and quantify constraints by conditional probabilities that represent not only the direction but also the strength of the relation. A constraint from V_1 to W_1 is expressed by $P(W_1|V_1)$. Figure 4a shows

Fig. 4 Constraints as conditional probabilities

how the simple SMM model in Fig. 2a is enriched by a constraint. The direction (positive or negative) is determined by comparing $P(W_1)$ and $P(W_1|V_1)$. The constraint is positive if $P(W_1|V_1) > P(W_1)$ and negative if $P(W_1|V_1) < P(W_1)$. Other than in Barsalou, the strength of a constraint can be quantified, namely as the difference between $P(W_1)$ and $P(W_1|V_1)$. More precisely, it is measured as $\frac{P(W_1|V_1)}{P(W_1)}$, which is the factor needed to adjust the unconditional probability to the conditional one, that is to the value of the constraint. For negative constraints, this factor is smaller than 1 and for positive constraints, it is greater than 1.

Let us illustrate this with an example from Connolly et al. (2007). Subjects in their study evaluated, *inter alia*, the truth likelihood of “Hamsters live in cages” and “Pet hamsters live in cages”. In contradiction to the general trend of assigning lower ratings to sentences with modifiers, the second sentence was perceived as more likely to be true than the first. On a scale from 1 (very unlikely) to 10 (very likely), the mean was 6.7 for the unmodified sentence and 7.9 for the modified one. For the category of hamsters, being a pet and living in a cage are positively correlated. The representation in Fig. 4b on page 10 illustrates this connection.

Our model can be easily extended to incorporate more complex structures, which are not found in the selective modification model but part of *recursive* frames in the sense of Barsalou (1992). Figure 5 gives an example of a partial bird frame, where attributes are applied to a value of another attribute. On the left side, we apply the attribute LOCOMOTION. The values are either a state of rest or a state of movement, where the latter becomes the argument of further attributes (MODE and SPEED). The probabilities of movement and rest model how likely each of these states is for the

Fig. 5 A partial representation of “bird” with a hierarchical structure and constraints between different levels

animals. For a sloth, the probability of a resting state would be very high. This probability of a value to which further attributes are applied should not be mistaken as importance of outgoing attributes. For example, the moving state has a low probability for the sloth, but specifications of the movement, especially of SPEED are nevertheless highly important. In our representation of “bird”, we assume that states of rest and movement are balanced for birds. The event of movement can be further characterized by the attributes MODE and SPEED. The probabilities of the values on this second level, e.g. “swimming”, “flying” etc. are always conditional on the upper value, that is that the animal is moving. There are two constraints in the model. First, we assume that movement by flying is necessarily fast. This is a constraint between values on the same level. A second constraint concerns different levels in the representation. Typically, birds have clawed feet and they usually move by flying. Some birds, however, have webbed feet, which are an adaptation to living in water and which enhance swimming. Thus, learning that a bird has webbed feet should increase the expectation that a movement of the bird will be swimming. The mentioned conditional probability replaces the probability of “swims if it moves” and is thus not only conditional on the constraining value “webbed” but also implicitly conditional on the event of movement. For sake of simplicity, we will focus on non-recursive models in our further discussion of constraints though recursive structures and nested constraints are possible.

3.2 Modification with Constraints

For constraining modifiers, the modification process is no longer as strictly selective as was illustrated in Fig. 2. One needs to take the constraint into account. In the spirit of Hampton and Jönsson (2012), common knowledge, represented in the form of a constraint, influences the meaning of V_1C – namely, by determining the probability of W_1 as $P'(W_1) = P(W_1|V_1)$. The probability of alternative values of W_1 will also have to be adjusted. If no further constraints are involved, then the probabilities of W_i , the alternative values of W_1 , reflect their initial proportion: $P'(W_i) = P(W_i) \frac{1-P(W_1|V_1)}{1-P(W_1)}$ if $P(W_1|V_1) < 1$ and $P'(W_i) = 0$ otherwise. Figure 6 shows the procedure for the basic model and for the example “pet hamster”. The readjustment of alternative values is rather trivial in this case since there is only one alternative value.

Fig. 6 Modification with constraints

Fig. 7 Birds and web-footed birds

Figure 7 gives a simplified representation of “bird”, where we do not consider specific movements and their mode but the preference for a type of locomotion. The commonly known correlation between webbed feet and swimming is represented by a positive constraint from the value “webbed” for foot structure to the value “swimming” as preferred method of locomotion. This is illustrated in Fig. 7a. In the modification “web-footed bird” the constraint is activated, leading to the new dominant value “swimming” for preferred locomotion with the probability of 0.75. The probability of the alternative values is calculated by multiplying the previous probability by $\frac{0.25}{0.85}$.

3.3 The Spreading of Constraints

The examples above clearly show that the influence of constraints spreads. A constraint from V_1 to W_1 has an influence on W_1 ’s alternatives. Constraints always exert an influence on more than one value of the attribute. Furthermore, our probabilistic constraints are bidirectional. The constraint from V_1 to W_1 comes with a reverse constraint from W_1 to V_1 , which can be calculated using Bayes’ theorem ($P(V_1|W_1) = \frac{P(W_1|V_1) \cdot P(V_1)}{P(W_1)}$).¹⁰

As an illustration, let us reconsider the constraint between webbed foot structure and swimming as preferred method of locomotion in birds. We assume that $P(\text{webbed}) = 0.2$, $P(\text{swimming}) = 0.15$ and the constraint $P(\text{swimming}|\text{webbed}) = 0.75$. Using Bayes’ theorem, we can infer the reverse constraint:

$$P(\text{webbed}|\text{swimming}) = \frac{P(\text{swimming}|\text{webbed}) \cdot P(\text{webbed})}{P(\text{swimming})} = \frac{0.75 \cdot 0.2}{0.15} = 1.$$

Thus, the initial constraint from foot structure to locomotion entails that swimming birds definitely have webbed feet. From that, it follows that claw-footed birds do not use swimming as their preferred mode of locomotion. We therefore have a derived constraint $P(\text{clawed}|\text{swimming}) = 0$, which immediately entails $P(\text{swimming}|\text{clawed}) = 0$. These derived constraints are illustrated in

¹⁰Note that bidirectional relevance is due to the epistemic interpretation. For probabilistic models of causality such as that of Pearl (1988) it is sometimes not plausible. In our discussion, we may safely assume symmetric relevance, because modification is an epistemic process rather than a causal one.

Fig. 8a. By these derivations, the constraint $P(\text{swimming}|\text{webbed}) = 0.75$ also has consequences for “clawed” as modifier, as is shown in Fig. 8b.

The spreading of constraints is notable for two reasons. First, attribute constraints and value constraints are strictly separated in Barsalou (1992), where the latter are limited to some specific values of certain attributes. In our probabilistic framework, every constraint is global. In a nutshell, if V_1 increases the likelihood of W_1 , then it decreases the probability of its alternatives and *vice versa*. Finally, W_1 increases the likelihood of V_1 and decreases the likelihood of V_2 if and only if V_1 increases the likelihood of W_1 . Note that the aforementioned strength of a constraint – measured as a factor between the unconditional and the conditional probability – is identical for the constraint and its reversal, namely: $\frac{P(W_1|V_1)}{P(W_1)} = \frac{P(V_1|W_1)}{P(V_1)}$. In our bird example, this is $\frac{0.75}{0.15} = \frac{1}{0.2} = 5$, which also happens to be the probabilistically maximal association for this initial distribution of values. This leads us to an important point: constraints are themselves constrained by the laws of probability theory.

3.4 Constraining Constraints

Constraints are restricted. For example, a constraint from a typical value to an atypical value cannot turn it into a typical one. It is not consistent to assume that hamsters typically have fur and teeth, but that furred hamsters are typically toothless. Where are the general borderlines? The crucial value that determines the possible impact of a constraint from V is $\frac{1}{P(V)}$, which is the factor that is needed to make V maximally probable – that is, certain. As shown in our example above, this is also the maximal possible association between V and a value of another attribute because $P(W|V) \leq \frac{P(W)}{P(V)}$. The maximal positive constraint is thus $P_{\max}(W|V) = \min(1, \frac{P(W)}{P(V)})$. Similar considerations determine the maximal negative constraint. A negative constraint to W can be understood as a positive constraint to $\neg W$. From $P(\neg W|V) \leq \frac{P(\neg W)}{P(V)}$ one can infer $1 - P(W|V) \leq \frac{1 - P(W)}{P(V)}$ and further $P(W|V) \geq 1 - \frac{1 - P(W)}{P(V)}$. Thus, we specify the minimal constraint as $P_{\min}(W|V) = \max(0, 1 - \frac{1 - P(W)}{P(V)})$.

Table 1 illustrates how the rules restrict the effect of constraints for some exemplary values, namely 0.9 as the probability of a typical value, 0.5 as the probability of a neutral value, and 0.1 as the probability of an atypical one. The crucial role of the initial probability of the modifying value V becomes apparent. If V is atypical,

Fig. 8 Birds and webbed footed birds

Table 1 Possible constraints depending on the probability of the constraining value V and the constrained value W

	$P(V)$	$P(W)$	$P_{min}(W V)$	$P_{max}(W V)$	Possible loss	Possible gain
Atypical to atypical	0.10	0.10	0.00	1.00	-0.10	0.90
Atypical to neutral	0.10	0.50	0.00	1.00	-0.50	0.50
Atypical to typical	0.10	0.90	0.00	1.00	-0.90	0.10
Neutral to atypical	0.50	0.10	0.00	0.20	-0.10	0.10
Neutral to neutral	0.50	0.50	0.00	1.00	-0.50	0.50
Neutral to typical	0.50	0.90	0.80	1.00	-0.10	0.10
Typical to atypical	0.90	0.10	0.00	0.11	-0.10	0.01
Typical to neutral	0.90	0.50	0.44	0.56	-0.06	0.06
Typical to typical	0.90	0.90	0.89	1.00	-0.01	0.10

The table shows that constraints from typical modifies, the last three lines, are usually ineffective. The value of W can deviate by about 0.1 at most. Atypical values, on the other hand, can have extremely strong constraints that set the probability of W to a totally new value, independent of its initial likelihood, as seen in the first three lines

as in the first three lines, then the constraint can change the new distribution of votes greatly. A typical modifier V , on the other hand, has only a limited potential to alter the initial probability. This is hardly surprising: learning things that we have already been suspected should not change our beliefs enormously.

Besides the formal considerations, there are often positive correlations between typical properties, especially if functionality is involved. In particular, patterns of typical functional properties are to be expected for categories with an evolutionary history, whether biological or cultural. For example, the typical shape of a tool is adjusted for its typical use. For the formal reasons explicated above, these constraints between different typical values do not lead to a high variability by typical modifications themselves. However, their counterparts can be quite effective. Learning that a particular human is a biped, for example, has little effect on expectations about her locomotive abilities, but learning that she has only one or no legs makes a crucial difference.

4 Selectivity and Uncertainty

We have presented a model of modification that is based on meaning composition according to the SMM, together with probabilistic rules of belief revision. The rationality of the Bayesian rule and its influence on modification should be uncontroversial. Apart from that, we generally defend a *default inheritance* strategy: if no constraint is involved, values are inherited. An agent can change her prototype assumptions about a subcategory from that of the category only if she has beliefs that impel her to do so. This is not so uncontroversial, because this strategy involves uncertainty.

4.1 The Objection on the Grounds of Uncertainty

Objections on the grounds of uncertainty were most prominently raised by Connolly et al. (2007) who say that “ S are T , thus MS are T ” is probabilistically invalid, since $P(MS) \leq P(S)$ follows from $MS \subseteq S$. As already remarked by Schurz (2012), this argument is flawed. To model an inference of properties from categories to subcategories, we must not consider the probabilities of the categories, but the conditional probabilities of the typical properties given category membership – that is $P(T|S)$ and $P(T|MS)$ – which are not related in this way. Depending on the probability of M and dependencies between M and T , $P(T|MS)$ can be even larger than $P(T|S)$. However, it is true that $P(T|MS)$ is not *guaranteed* to be as high as $P(T|S)$. An inference from “ S are T ” to “ MS are T ” is therefore risky (see also Ströbner and Schurz 2020). Figure 9 gives an impression of this uncertainty, showing how $P(T|MS)$ can deviate from $P(T|S) = 0.95$, relative to $P(M|S)$. The borderlines of uncertainty are determined by the same probabilistic considerations as maximal and minimal constraints.

In Fig. 9, we see that the possible probabilistic *gain* of $P(T|MS)$, compared to $P(T|S)$, is low, because $P(T|S)$ is already very high. The possible loss of $P(T|MS)$, compared to $P(T|S)$, depends on the likelihood of M . It is very small for likely M , but high for unlikely M . In other words, default inheritance to small subcategories is probabilistically very risky. This underpins the objection to default inheritance for atypical modifiers. The same line of argument can be brought against default inheritance with multiple modifiers, because with each additional modifier the subcategory becomes smaller. Though the general probabilistic argument by Connolly et al. (2007) is invalid, the following objection should nevertheless be taken seriously:

Fig. 9 Probabilities of property T in the subcategory MS given that T has an initial probability of 0.95 in the category S . The low red line shows the least probability, which remains very close to 0.95 if the modifier is likely or at least not unlikely but drops rapidly for increasingly unlikely modifiers (especially for values below 0.2)

Uncertainty under composition (UC): The more concepts enter novel combinations, the less certain one is about the stereotypic properties of the instances of the combined concept (Connolly et al. 2007, 13).

At this point, discussion of prototype composition is strongly linked to results on non-monotonic reasoning generated in the field of artificial intelligence. We will briefly sketch this relation.

4.2 Nonmonotonic Reasoning

The general idea behind nonmonotonic reasoning is defeasibility. Classically valid conclusions are not sensitive to the addition of further premises – that means, they satisfy monotony. Moreover, classical conditionals, either material or strict, are monotonic in the sense that they cannot become false by the addition of further premises: “If A then (necessarily) C ” entails “If $A \& B$ then (necessarily) C ”, no matter what B says. Normality assumptions of the form “If A then normally C ” ($A \rightsquigarrow C$) do *not* satisfy this principle. “If $A \& B$ then normally C ” ($A \& B \rightsquigarrow C$) does not necessarily follow. For example, soil will normally become wet if it rains, but not if the soil is covered. A central theme of nonmonotonic reasoning is to find restricted versions of monotony that are still plausible. The following weakened versions of monotony are discussed by Kraus et al. (1990):

Cautious monotony: $A \rightsquigarrow B, A \rightsquigarrow C \vdash A \& B \rightsquigarrow C$

and the stronger

Rational monotony: $\neg(A \rightsquigarrow \neg B), A \rightsquigarrow C \vdash A \& B \rightsquigarrow C$

Cautious monotony can be easily grounded on probabilistic considerations (cf. Schurz 2005). As we saw in Fig. 9, the possible loss in probability is small for incoming information that is already expected. The riskier rational monotony, which allows us to infer monotonically in situations that are not exceptional, leads to moderate increases in uncertainty, which are not very high but more than negligible.

However, there are even more courageous accounts than rational monotony. Pearl (1990) suggests system Z, where monotony holds by default: $A \rightsquigarrow C \vdash_z A \& B \rightsquigarrow C$. The inference itself is defeasible and can become invalid if a conflicting normality rule is added: $A \rightsquigarrow C, B \rightsquigarrow \neg C \not\vdash_z A \& B \rightsquigarrow C$. Simulation studies by Schurz and Thorn (2012), Thorn and Schurz (2014, 2016) have shown that default inheritance, as implemented in system Z, usually provides a good balance between the number of inferences that can be drawn and their reliability.

Pearl (1990) also sympathises with making default inferences to *exceptional* subcategories. According to him, it should be legitimate to infer “Dark-haired Swedes are well-mannered” from “Swedes are well-mannered” and “Swedes are blond” if there is no known relation between manners and hair colours (cf. Pearl 1990, 130). This inference pattern is *not* implemented in Z: the mere exceptionality of dark-haired Swedes blocks the inheritance. To enable the inference, one can use the maximum entropy approach, which restricts consideration to probability distributions that are

as neutral as possible insofar as they only give high conditional probabilities if there are known dependencies (Pearl 1988, 1990). This is exactly what our enriched selective modification model does: it assumes independence where no dependence is known.

From the viewpoint of a possible increase in uncertainty, default inference to exceptional subcategories is extremely risky. However, recent simulation studies by Thorn and Schurz (2018) give a more differentiated picture. Categories that are defined by one necessary and sufficient atomic property alone (for example, the category of green things) do indeed make default inheritance very unreliable. The situation is different for categories that are constructed by maximalizing internal similarity in different properties. These categories resemble prototype categories insofar as they are structured by being close to a centroid, that is, the central point of the category. For these so-called *fitted classes*, inheritance inferences to subclasses are reasonably reliable: error rates do not exceed 0.1 in different scenarios. Such results provide evidence for a more optimistic view of default inheritance.

4.3 Defending Selectivity

Our extended modification model restricts selectivity only if constraints are involved. In terms of nonmonotonic logic, the model thus exceeds Z and resembles the maximum entropy approach insofar as we sanction inference to exceptional subcategories. In our view, any model of compositionality should implement this inference. To deny default inheritance for atypical modifiers would go far beyond the hypothesis that composition interacts with common knowledge and would relinquish the central ideas of compositionality. In other words, without default inheritance, the model would hardly represent any kind of compositionality. Of course, this is not convincing for those who hold that prototypes are not compositional precisely because default inheritance is faulty. However, in view of research on nonmonotonic reasoning, we support a more optimistic view of default inheritance:

- Intuitively, default inheritance (as in system Z) is generally a preferable feature of nonmonotonic systems. Simulations tend to support this intuition.
- For typical modifications, default inheritance also comes with little increase in uncertainty.
- For atypical modifications, there is an enormous degree of possible uncertainty. However, at least for categories with internal similarity, that is, prototype concepts, the reliability of default inheritance is still good.

Note that these arguments in favour of default inheritance also support selectivity, which commands the transfer of properties from the category to the subcategory. Selectivity is thus the compositional counterpart of minimally restricted monotony.¹¹

¹¹ A difference is that monotony deals mainly with typical properties, while selectivity concerns the behaviour of typical, neutral, and atypical properties alike.

5 Conclusion

The last few sections presented a model of compositionality that takes background knowledge into account. We developed and defended it as a *normative* model of prototype composition and reasoning – that is, a model of how a rational agent *should* deal with the interaction of background knowledge and compositionality in modifications of prototype concepts.

As mentioned earlier, there is also much empirical research on how humans deal with modification tasks. Though we have mainly aimed to contribute to the *logic* of typicality, we briefly address this work here. In all studies of this effect, subjects tended to rate a sentence with a typical modifier (for example, “Feathered ravens are black”) as more likely to be true than a sentence with a non-typical one (for example, “Young ravens are black”). A possible explanation in purely probabilistic terms would be that the uncertainty is smaller with inheritance to large subcategories. In our model, we can also offer a further explanation in terms of constraints. Typical values have no highly influential constraints on each other. Sometimes they are related by weak positive constraints, which in turn might lead to highly influential negative constraints on alternative atypical values.

There are, however, empirical findings in conflict with an explanation in terms of rational reasoning as we aimed to model it. For example, Gagné and Spalding (2011, 2014) observed decreased likelihood ratings for meaningless pronounceable modifiers, which is not explainable by background knowledge. Findings from Strößner and Schurz (2020) support the thesis that human ratings of modified typicality statements are to a large extent influenced by pragmatics, which might not even be conscious. Note however that such effects are not exclusively found for probabilistic reasoning. Hampton et al. (2011) showed that decreased ratings for likelihood are even found for properties that are analytically true. A sentence such as “Young jungle ravens are birds” is judged as less plausible than “Ravens are birds”, even though the sentences are not only highly probable but certain. This fits also with findings that individuals judge equally fallacious, when dealing with modified *universal* statements: Jönsson and Hampton (2006) found that sentences such as “All handmade sofas are comfortable” are perceived as less likely than “All sofas are comfortable”, though the former is clearly a logical consequence of the latter. Inasmuch as this empirical evidence does not disprove the validity of classical logic, the occurrence of modifier effects for prototype concepts should not be seen as disproving probabilistic models. Thus, we can remain optimistic that probabilistic logic can serve as the normative counterpart of prototype theory, as deductive logic has served, and still serves, as the formal counterpart of the classical theory of concepts. However, not surprisingly, mathematical, normative models and psychological explanations part company at some point, no matter which theory of concepts one assumes.

In conclusion, let us make a more general point about compositionality. The requirement of compositionality has often been viewed as an important, if not the most important, criterion for what concepts can be. It is usually postulated as the explanation for the creativity of human thinking and understanding. Thoughts and sentences can be created from existing concepts and words: this is indeed an essential part of what concepts are. However, that does not mean that the principle of

compositionality should decide over what concepts can or cannot be. It is rather that appropriate rules for composition depend on what concepts are. The needed procedures can go beyond set-theoretic operations or classical logic.¹² In view of many phenomena in natural language, it is hardly plausible that composition is always straightforward. In general, the fact that background knowledge is part of conceptual content never constitutes an obstacle. The results of belief changing actions are derivable from what is known and what is learned. In this sense belief revision itself is compositional.

Acknowledgments The work was generously founded by the German Research Foundation DFG. Grant SFB 991 D01. Open access is provided via the project DEAL. I am grateful to my project colleagues Annika Schuster, Gerhard Schurz and Paul Thorn as well as Peter Sutton and Henk Zeevat with whom I discussed probabilistic versions of frames while working on this paper. Moreover, James Hampton and Steven Verheyen provided many helpful comments on earlier versions of the paper.

Funding Information Open Access funding provided by Projekt DEAL.

Compliance with Ethical Standards

Conflict of interests I have no conflicts of interest to declare.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Barsalou, L.W. 1992. Frames, concepts, and conceptual fields. In *Frames, fields, and contrasts*, eds. A. Lehrer, and E. Kittay, 21–74. Hillsdale: Lawrence Erlbaum Associates Publishers.
- Connolly, A.C., J.A. Fodor, L.R. Gleitman, and H. Gleitman. 2007. Why stereotypes don't even make good defaults? *Cognition* 103(1): 1–22. <https://doi.org/10.1016/j.cognition.2006.02.005>.
- Cree, G.S., and K. McRae. 2003. Analyzing the factors underlying the structure and computation of the meaning of chipmunk, cherry, chisel, cheese, and cello (and many other such concrete nouns). *Journal of Experimental Psychology: General* 132(2): 163. <https://doi.org/10.1037/0096-3445.132.2.163>.
- Fodor, J. 1998. *Concepts: Where cognitive science went wrong*. Oxford: Oxford University Press.
- Fodor, J., and E. Lepore. 1996. The red herring and the pet fish: why concepts still can't be prototypes. *Cognition* 58(2): 253–270. [https://doi.org/10.1016/0010-0277\(95\)00694-X](https://doi.org/10.1016/0010-0277(95)00694-X).
- Frege, G. 2007. *Funktion, Begriff, Bedeutung: Fünf logische Studien*. ed. G. Patzig. Göttingen: Vandenhoeck & Ruprecht.
- Gagné, C., and T. Spalding. 2011. Inferential processing and meta-knowledge as the bases for property inclusion in combined concepts. *Journal of Memory and Language* 65(2): 176–192. <https://doi.org/10.1016/j.jml.2011.03.005>.

¹²An impressive example of the formal complexity of prototype composition is Lewis and Lawry (2016), who formally capture experimental results on prototype composition from Hampton (1987) using random set theory and conceptual spaces.

- Gagné, C., and T. Spalding. 2014. Subcategorisation, not uncertainty, drives the modification effect. *Language, Cognition and Neuroscience* 29(10): 1283–1294. <https://doi.org/10.1080/23273798.2014.911924>.
- Hampton, J. 1987. Inheritance of attributes in natural concept conjunctions. *Memory & Cognition* 15.1: 55–71. <https://doi.org/10.3758/BF03197712>.
- Hampton, J.A., and M.L. Jönsson. 2012. Typicality and compositionality: The logic of combining vague concepts. In *The Oxford handbook of compositionality*, eds. W. Hinzen, and E. Machery, 385–402. Oxford: Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199541072.013.0018>.
- Hampton, J.A., A. Passanisi, and M.L. Jönsson. 2011. The modifier effect and property mutability. *Journal of Memory and Language* 64(3): 233–248. <https://doi.org/10.1016/j.jml.2010.12.001>.
- Huber, F. 2016. Formal representations of belief. In *The Stanford encyclopedia of philosophy* (Spring 2016 edn), ed. E. N. Zalta. Stanford: Metaphysics Research Lab, Stanford University. <https://plato.stanford.edu/entries/formal-belief/>.
- Jones, G.V. 1983. Identifying basic categories. *Psychological Bulletin* 94(3): 423–428. <https://doi.org/10.1037/0033-2909.94.3.423>.
- Jönsson, M.L., and J.A. Hampton. 2006. The inverse conjunction fallacy. *Journal of Memory and Language* 55(3): 317–334. <https://doi.org/10.1016/j.jml.2006.06.005>.
- Jönsson, M.L., and J.A. Hampton. 2012. The modifier effect in within-category induction: default inheritance in complex noun phrases. *Language and Cognitive Processes* 27(1): 90–116. <https://doi.org/10.1080/01690965.2010.544107>.
- Kraus, S., D. Lehmann, and M. Magidor. 1990. Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial intelligence* 44(1–2): 167–207. [https://doi.org/10.1016/0004-3702\(90\)90101-5](https://doi.org/10.1016/0004-3702(90)90101-5).
- Laurence, S., and E. Margolis. 1999. Concepts and cognitive science. In *Concepts: core readings*, 3–81. Cambridge: MIT Press.
- Leitgeb, H. 2017. *The stability of belief: how rational belief coheres with probability*. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780198732631.001.0001>.
- Lewis, M., and J. Lawry. 2016. Hierarchical conceptual spaces for concept combination. *Artificial Intelligence* 237: 204–227. <https://doi.org/10.1016/j.artint.2016.04.008>.
- Murphy, G. 2004. *The big book of concepts*. Cambridge: MIT press.
- Murphy, G. 1982. Cue validity and levels of categorization. *Psychological Bulletin* 91(1): 174–177. <https://doi.org/10.1037/0033-2909.91.1.174>.
- Osherson, D.N., and E.E. Smith. 1981. On the adequacy of prototype theory as a theory of concepts. *Cognition* 9(1): 35–58. [https://doi.org/10.1016/0010-0277\(81\)90013-5](https://doi.org/10.1016/0010-0277(81)90013-5).
- Pearl, J. 1988. *Probabilistic reasoning in intelligent systems: networks of plausible inference*. San Francisco: Morgan Kaufmann. <https://doi.org/10.1016/C2009-0-27609-4>.
- Pearl, J. 1990. System z: a natural ordering of defaults with tractable applications to nonmonotonic reasoning. In *Proceedings of the 3rd conference on theoretical aspects of reasoning about knowledge*, 121–135. San Francisco CA: Morgan Kaufmann.
- Pelletier, F.J. 2001. Did Frege believe Frege's principle? *Journal of Logic, Language and information* 10(1): 87–114. <https://doi.org/10.1023/A:1026594023292>.
- Pelletier, F.J. 2017. Compositionality and concepts - a perspective from formal semantics and philosophy of language. In *Compositionality and concepts in linguistics and psychology*, eds. J.A. Hampton, and Y. Winter, 31–94. Cham: Springer. https://doi.org/10.1007/978-3-319-45977-6_3.
- Rosch, E. 1973. Natural categories. *Cognitive Psychology* 4(3): 328–350. [https://doi.org/10.1016/0010-0285\(73\)90017-0](https://doi.org/10.1016/0010-0285(73)90017-0).
- Rosch, E. 1978. Principles of categorization. In *Cognition and categorization*, eds. E. Rosch, and B. Lloyd, 27–48. Hillsdale: Lawrence Erlbaum Associates.
- Rosch, E., and C.B. Mervis. 1975. Family resemblances: studies in the internal structure of categories. *Cognitive Psychology* 7: 573–605. [https://doi.org/10.1016/0010-0285\(75\)90024-9](https://doi.org/10.1016/0010-0285(75)90024-9).
- Rosch, E., C.B. Mervis, W.D. Gray, D.M. Johnson, and P. Boyes-Braem. 1976. Basic objects in natural categories. *Cognitive Psychology* 8(3): 382–439. [https://doi.org/10.1016/0010-0285\(76\)90013-X](https://doi.org/10.1016/0010-0285(76)90013-X).
- Schurz, G. 2005. Non-monotonic reasoning from an evolutionary viewpoint: ontic, logical and cognitive foundations. *Synthese* 146(1–2): 37–51. <https://doi.org/10.1007/s11229-005-9067-8>.
- Schurz, G. 2012. Prototypes and their composition from an evolutionary point of view. In *The Oxford handbook of compositionality*, eds. W. Hinzen, and E. Machery, 530–553. Oxford: Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199541072.013.0026>.

- Schurz, G. 2019. Impossibility results for rational belief. *Noûs* 53: 134–159. <https://doi.org/10.1111/nous.12214>.
- Schurz, G., and P.D. Thorn. 2012. Reward versus risk in uncertain inference: theorems and simulations. *The Review of Symbolic Logic* 5(4): 574–612. <https://doi.org/10.1017/S1755020312000184>.
- Smith, E.E., D.N. Osherson, L.J. Rips, and M. Keane. 1988. Combining prototypes: a selective modification model. *Cognitive Science* 12(4): 485–527. [https://doi.org/10.1016/0364-0213\(88\)90011-0](https://doi.org/10.1016/0364-0213(88)90011-0).
- Spalding, T., and C. Gagné. 2015. Property attribution in combined concepts. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 41(3): 693–707. <https://doi.org/10.1037/xlm0000085>.
- Strößner, C., and G. Schurz. 2020. The role of reasoning and pragmatics in the modifier effect. *Cognitive Science* 44: e12815. <https://doi.org/10.1111/cogs.12815>.
- Strößner, C., A. Schuster, and G. Schurz. 2020. Modification and default inheritance. In *Concepts, frames and cascades in semantics, cognition and ontology*, eds. T Gamerschlag, T Kalenscher, S Löbner, M Schrenk, and H Zeevat.
- Thorn, P.D., and G. Schurz. 2018. Inheritance inference from an ecological perspective. CoSt18 talk. <https://cognitive-structures-cost18.phil.hhu.de/wp-content/uploads/2018/08/Thorn.Inheritance-Reasoning-from-an-Ecological-Perspective.pdf>.
- Thorn, P.D., and G. Schurz. 2014. A utility based evaluation of logico-probabilistic systems. *Studia Logica* 102(4): 867–890. <https://doi.org/10.1007/s11225-013-9526-z>.
- Thorn, P.D., and G. Schurz. 2016. Qualitative probabilistic inference under varied entropy levels. *Journal of Applied Logic* 19: 87–101. <https://doi.org/10.1016/j.jal.2016.05.004>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.