

Stochastic frames

Annika Schuster

Department of Philosophy
Heinrich Heine University Düsseldorf
annika.schuster@uni-duesseldorf.de

Corina Ströbner

Department of Philosophy II
Ruhr-University of Bochum
corina.stroessner@rub.de

Peter R. Sutton

Department of General Linguistics
Heinrich Heine University Düsseldorf
peter.sutton@uni-duesseldorf.de

Henk Zeevat

ILLC
University of Amsterdam
h.w.zeevat@uva.nl

Abstract

In the frame hypothesis (Barsalou, 1992; Löbner, 2014), human concepts are equated with frames, which extend feature lists by a functional structure consisting of attributes and values. For example, a bachelor is represented by the attributes GENDER and MARITAL STATUS and their values ‘male’ and ‘unwed’. This paper makes the point that for many applications of concepts in cognition, including for concepts to be associated with lexemes in natural languages, the right structures to assume are not merely frames but stochastic frames in which attributes are associated with (conditional) probability distributions over values. The paper introduces the idea of stochastic frames and three applications of this idea: vagueness, ambiguity, and typicality.

1 Background: Frames

Frames originated in Minsky (1974) and were further developed in the field of cognitive science by Barsalou (1992). Frames extend feature lists by a functional structure consisting of attributes and values. Petersen (2007) developed a precise formalisation of recursive frames, in which frames are connected directed graphs, with labeled nodes, labeled arrows and a central node. The labels on arrows are interpreted as partial functions over the domain (attributes) and the labels on nodes as classes of elements of the domain (values). A frame F applies to an object x if there is a function f that assigns x to the central node, and that assigns an object in the annotated class to each node and is such that whenever an arrow labeled a_i leads from node q_j to node q_k , $a_i(f(q_j)) = f(q_k)$. A function typ assigns an element from the set of types **TYPE** to every node (see Figure 1). Figure 2 demonstrates how the frame structure applies to a particular individual (the black cat Felix).

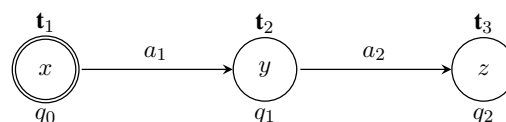


Figure 1: Schema for a frame, F , (a recursive attribute-value structure) with a central node with a value x such that: $a_1, a_2 \in A$, the set of attributes; $q_0, q_1, q_2 \in Q$, the set of nodes; $t_1, t_2, t_3 \in \mathbf{TYPE}$. F applies to x iff there is some function f such that $f(q_0) = x : t_1$, $f(q_1) = y : t_2$, $f(q_2) = z : t_3$ where $a_1(f(q_0)) = f(q_1)$ and $a_2(f(q_1)) = f(q_2)$.

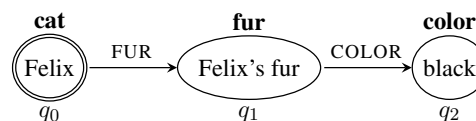


Figure 2: A partial frame of a specific black cat Felix, following the schema in Figure 1.

The type of a node contains the semantic information associated with that node. For example, the value for the attribute AGE is necessarily a time, and for humans it is a value between 0 and approximately 120 years. The set **TYPE** is usually accompanied by an ontology (Carpenter, 1992), which includes additional information about the relation between the classes. For example, it determines whether they are exclusive (‘red’ and ‘blue’) or one contains the other (‘crimson’ and ‘red’).

According to the frame hypothesis, frames are the “single general format of representations” in human cognition (Löbner, 2014, 23). However, this hypothesis requires restrictions on the set of admissible attributes and types to be empirically meaningful. The frame hypothesis can be empirically grounded if we assume that the types and functions are *natural* in human cognition. From

Binder et al. (2016), for example, one can conclude that smell, color, touch, texture, size, weight, sound, and shape are -in combination- vital for the recognition of many concrete natural objects (such as trees and rocks) and animate individuals (such as animals). From this, one can conclude that such attributes are natural and could be some of the attributes on which a natural frame is based.

Let us make a general remark on concepts that also applies to frames. Speaking of concepts, which are usually associated with lexemes, prejudices one into thinking that these concepts are autonomous parts of cognition. However, it seems equally correct to think of lexemes getting a high activation when a certain configuration in semantic memory is activated. In their role as what is expressed by natural language lexemes, concepts (or their frames) are not necessarily more than an isolatable chunk of mental life that does not exist independently of the processes from which it is isolated. The same remark holds for one of the questions we consider in this paper: what are the necessary ingredients of a (stochastic) frame? For many purposes, the explanation can be carried out with a limited notion of a particular frame, based on an abstraction over concrete cases or instances. In our terms, this means, for example, assuming a particular attribute-value frame structure in any given case. For implementation purposes, such limitations are essential, but that does not mean there is a realm of concept-like entities that exhibit these limitations.

2 Stochastic frames

Stochastic frames are the stochastic version of the frames defined in Petersen (2007). This incorporates the probabilistic semantics of Sutton (2015) (also see Cooper et al., 2015) in which the basic idea is that the nodes are associated probability distributions over possible values.

Formally, a (minimally) stochastic frame has a recursive attribute-value structure with a central node with a value x and where each node in the frame has a type from a type hierarchy, just as classical frames do. Where (minimal) stochastic frames differ from classical frames is that the values of attributes need not be categorical (given as particular entities). Instead, they may be probability distributions over entities/values of the relevant type. For example, for a stochastic frame F' that

contains an attribute **COLOR**, the range of this attribute is a probability distribution over entities of the type **Color** (points in the color space). If F' contains an attribute **HEIGHT**, the range of this attribute is a probability distribution over entities of the type **Height** (values on a measurement scale).

A simple example is given in Figure 3 for a cat, Felix, where the agent does not know what color fur Felix has. This contrasts with the non-stochastic, classical frame in Figure 2 in which a categorical value for Felix’s fur color is recorded. Some values in the stochastic frame can be categorical (technically, an assignment of probability 1 to a single value). For example, the value in the stochastic Felix frame for the attribute **FUR** is assumed to be categorical in this way if the agent knows that Felix has fur. However, in stochastic frames, values of attributes may be distributions over multiple values each with > 0 probability values. For example, a distribution over colors for the value of attribute **FUR** (such that this distribution may be generated by the agent’s experiences of the typical fur colors of cats, see section 3.3).

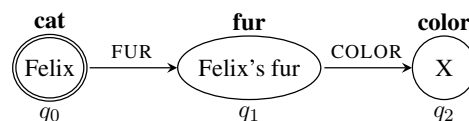


Figure 3: A partial stochastic frame for a cat Felix where X is a probability distribution over points in the color space.

This minimal conception of a stochastic frame is extended by the addition of constraints. Constraints are encoded as conditional probabilities that capture stochastic relations between the values that different nodes of the frame can obtain (see §2.2).

As a further extension, probabilities can also be embedded at the level of attributes, not just values. For example, the probability that cats have fur (that the cat frame has a **FUR** attribute), given that an agent may consider fur-less cats to be a possibility.

The minimal and extended characterisations of stochastic frames given above extend the notion of classical frames. Stochastic values (and attributes) allow us to model the uncertainty an agent has about a particular entity or class of entities. Combined with stochastic constraints on relations be-

tween nodes, as we will show, we can furthermore model the knowledge that agents have regarding the distributions of properties of entities of some particular class.

Stochastic frames and classical frames more or less converge when it comes to *ground frames*, i.e., frames for specific individuals in which every attribute has a categorical value. For example, the stochastic frame of a particular instance of ‘bachelor’ (say John) resembles the corresponding classical frame. (The difference between them is that, technically, the values of attributes in a stochastic ground frame are probability distributions over a single value in a classical frame.)

Where classical and stochastic frames diverge is with respect to uninstantiated frames and frames which are only partially grounded (where some attribute values are not given categorical values). A set of ground frames can be taken as the list of observations on which an uninstantiated stochastic frame or a partially grounded stochastic frame is based, where the list of observations corresponds to the probabilities the frame assigns. This is the case for the partially grounded stochastic frame in figure 3. The frame assigns categorical values to the referent of *Felix* and to the stuff that makes up *Felix*’s fur, but the distribution over colors of fur is based on observed instances of cats and the colors of fur they have. Thinking in terms of ground instances gives a simple transition from thinking in terms of belonging to classes with a given probability to thinking in terms of distributions over values.

2.1 Definitions and prototypes

In its non-stochastic form, the frame hypothesis on concepts fits best with a classical theory of concepts, where concepts are defined by necessary and jointly sufficient conditions of category membership. For example, the concept ‘bachelor’ is defined as a male, adult person, who is unmarried. The classical view can be traced back to Plato’s dialogues and was also fundamental in the early development of formal logic. Frames extend this view by the further demand that they are a quantifier-free conjunction of atoms of the form: x belongs to class C and attribute a_i maps x to y . Formulated in this way, the frame hypothesis seems to rule out any view of concepts in which they do not characterise necessary and sufficient

conditions, such as the prototype theory.

Starting with the writings of Wittgenstein (1953) the classical view began to lose credibility. The vagueness of concepts, as well as many other empirical results, can be taken as evidence that the classical view in which concepts provide necessary and sufficient conditions for their application is not on the right track (see Margolis and Laurence, 1999, 27).

Meanwhile, other approaches have been developed and have gained prominence. The most widely discussed one is the prototype theory of concepts, going back to Eleanor Rosch and her collaborators (Rosch, 1973; Rosch and Mervis, 1975; Rosch, 1978; Rosch et al., 1976). It explains the application of a concept to an instance in terms of its similarity to a so-called prototype. This prototype can be understood as a central instance, for example a focal color (Rosch, 1973), but it can also be an idealised representation of the concept (Rosch, 1978). In all variants of prototype theory, a central idea is that concepts are based on an overall similarity of instances rather than on defined features that are common to all instances. Conceptual spaces theory (Gärdenfors, 2000) is also based on similarity, which is understood as an inverse of geometric distance. In this approach, concepts are equated with areas in conceptual spaces and instances with points in these areas. Gärdenfors (2000) emphasizes the relation to the prototype view, according to which an instance is matched to C if it is similar to the the centroid (central point) of the geometric area covered by C . On this understanding, a prototype is a central point in the category. However, if the prototype is not seen as a central point but as a typicality weighted summary of properties one finds in the category, prototypes *are* stochastic frames: they express which properties are likely and in this respect typical.

2.2 Constraints

Constraints were already thought to be an important part of frames in Barsalou (1992), where positive dependencies are marked by a “+” and negative ones by a “-”. A good example is the concept of a bird. Birds have different principal modes of locomotion (flying, swimming, and walking), and birds also have different physical features such as webbed feet, or clawed feet. Flying birds with

clawed feet are more typical. However, there are correlations between the swimming, walking, and flying of birds to the feet type. While birds normally have clawed feet, the webbed structure is more expected for birds that swim. In a stochastic frame, the relations between properties of birds are captured as conditional probabilities. Figure 4 shows a partial stochastic frame annotated with such probabilistic constraints. These constraints, allow us not only to model the typical properties of birds, but also to reason about properties of entities on the basis of partial information, for example, that swimming birds have a high probability of having webbed feet.

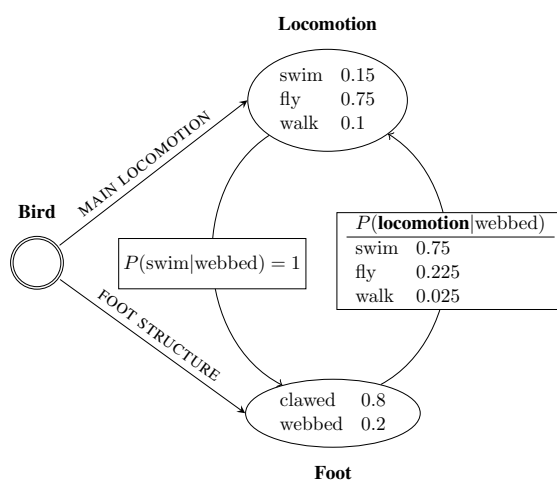


Figure 4: A partial frame for an arbitrary bird with a probabilistic constraint governing the connection between having webbed feet and the main means of locomotion.

The rest of the paper runs through three applications of stochastic frames, superficially, since these applications are covered elsewhere in greater depth. The point of including them here is to make it clear these applications need stochastic frames and that they require very similar characterisations of stochastic frames. This convergence together with the realisation that frames need to be replaced by stochastic frames if the frame hypothesis is accepted as true formed the basis of the authors’ cooperation for this paper.

3 Three Applications

3.1 Vague predication

Probabilistic models of vague expressions such as *tall* can capture how an utterance of a sentence such as (1) can reduce the uncertainty hearers

have about the way the world is, for example, John’s height, as well as about what contextual standards are in play regarding the meaning of *tall* (Lassiter, 2011).

- (1) John is tall for a basketball player

In this section, we outline, first, how stochastic frames can incorporate this insight of probabilistic models of vagueness. We then discuss why a frame-based analysis for gradable adjectives is advantageous when dealing with more complex varieties of adjectival modification than the example in (1).

We start with a derivation for *John is tall* (see Figure 5). The subject NP denotes a frame for *John* the central node of which is typed **Person**. This frame includes height information relative to this type (possibly affected by assumptions relating to gender etc.), namely an attribute **HEIGHT**, the value of which is a probability distribution over heights (we assume for convenience that the unit of measurement is centimetres). On the assumption that no other size information is known about John, and on the assumption that John is a man, this distribution should reflect the sizes of men and so have a mean value around 1.75m, the average height of men in the authors’ country of residence (we suppress gender information in Figure 5).

The interpretation of *tall*, $\llbracket \text{tall} \rrbracket$, we propose, is a function on the value of a **HEIGHT** attribute in a frame such that $\llbracket \text{tall} \rrbracket$ can compose with any frame that contains a **HEIGHT** attribute. We propose that $\llbracket \text{tall} \rrbracket$ furthermore encodes a function f_{tall} that is applied to the value of this attribute. Where the value of a **HEIGHT** attribute is represented as a tuple $\langle \mu, \sigma \rangle$ of the mean (μ) and standard deviation (σ) of a Gaussian distribution, the function f_{tall} is such that $f_{\text{tall}}(\langle \mu, \sigma \rangle) = \langle n\mu, m\sigma \rangle$ for some positive factor n and some negative factor m . In other words, we propose that $\llbracket \text{tall} \rrbracket$ shifts up one’s expectations as to average height (relative to the height expectations for the concept to which *tall* applies, and decreases the variance.

To derive *tall for a basketball player*, we propose that $\llbracket \text{tall} \rrbracket$ first modifies a basketball player frame (see Figure 5 for a schematic derivation). As with the modification of the **Person** frame (that was instantiated by John) in the previous case, the function f_{tall} applies to the value of the **HEIGHT** attribute. However, in this case, background knowledge about the heights of basketball players can

be different from the heights of people in general (we tend to know that the former are taller). On the assumption that basketball players are believed to be on average 200cm tall, the effect of applying f_{tall} is to shift this mean upwards and reduce the standard deviation. The effect of this is that the expected height of John, given the information in (1) is drawn from a different distribution over heights than if one were told that *John is tall*, without specifying a comparison class.

Finally, we propose a constraint on the felicitous use of comparison class *for*-PPs, namely that the type for the *for*-PPs (e.g., **BBP** (basketball player)) must be a subtype of the implicit type for the subject NP (e.g., **Person** in *John is tall for a basketball player*). This correctly predicts the oddity of sentences such as *John is tall for a bush*.

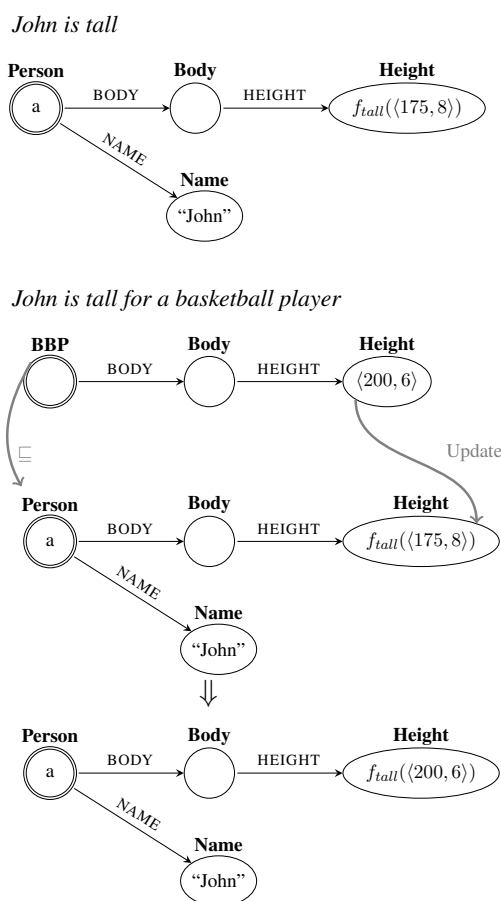


Figure 5: Schema for deriving *John is tall* and *John is tall for a basketball player* in stochastic frames.

The treatment for *tall* sketched above has in common with other probabilistic approaches to vagueness that vague adjectives convey probabilistic information that can be used to infer the probability of some object having observable properties, such as a particular height, given the knowledge and be-

liefs a hearer has about the way the world is. (See, among others, Sutton, 2015; Lassiter, 2011; Égré, 2017). An advantage of these approaches is that they can be formulated in such a way as to not assume the presence of hidden sharp boundaries in linguistic knowledge (a point above which entities are P and below which, they are not- P (see Sutton, 2018, for further discussion). It is unclear how any of these insights could be modelled within a classical frame.

The advantage of combining a probabilistic approach to vagueness with frame theory (using stochastic frames) is the incorporation of a theory of adjectival modification. For example, *red pen* can be naturally understood as meaning, among other things, a pen that writes in red, a pen that has a red casing, or a pen with a red lid. However, the value *red* must relate to some aspect of the pen (or of an object related to the pen in a specific context). In other words the locus for adjectival modification is underspecified but nonetheless constrained. Frame theory captures this in terms of the number of attributes of the right type there are in the frame. For example, $[[tall]]$ can apply only to HEIGHT attributes in an NP frame and $[[red]]$ can apply only to COLOR attributes in an NP frame.

3.2 Ambiguity

This section is a recapitulation of the relevant parts of Zeevat et al. (2015).

Words in natural languages can be studied for their contribution to the truth-conditions of the expressions they are part of in particular contexts. This gives rise to a bewildering number of non-equivalent readings (meanings in use) for high-frequency verbs with a long history: 84 for the verb “fall”, 71 for “run” where it is by no means clear that these are all the readings one needs, given that yet further uses are frequently reported. The situation is similar for many nouns and adjectives.

Stochastic frames can accommodate those readings, since they can deal with different possibilities with preferences due to the frequency with which the reading occurs, but mere accommodation is not what one wants. Human language users effortlessly disambiguate in these cases in linear time¹. This ability is modeled by stochas-

¹This follows from the empirical work on which the Mar-

tic frames in which the lexical ambiguity is not captured as list of readings, but as one integrated structure, with different preferentially weighted options, once an optimal use is made of the context as an additional resource in a unification like process.

Lexical stochastic frames for verbs will have many nodes that are presupposed from the context. The standard cases are the obligatory arguments of the verbs. Saying that such arguments are obligatory means that it is obligatory for the context to supply the relevant information. But talking of obligatory arguments is just one aspect of the more general case that many of the readings require values from noun phrases or prepositional phrases or from presupposition resolution to the linguistic and non-linguistic context to be possible. Interactions with the context will therefore deliver a particular version of the part of the concept that refers to given material. By unification, this will also change the part of the concept that contains the new information. In addition, the stochastic frame will contain alternatives with different probabilities and in disambiguation, the more probable alternative will be systematically preferred.

As reported in Zeevat et al. (2015), an approach of this kind was hand-tested on the verb “fall” (all readings in David Copperfield (Dickens, 2000)) with full success. The approach uses a logical representation of equivalence classes of stochastic frames (the ones that give the same inequalities). This allows the different users to learn their own probabilities, converging on the same equivalence class under enough exposure to uses of the word. The logical representation using (comparative) probabilistic preferences rather than full-fledged probabilities is human readable and can also be taken as the object that establishes the (near-)unity of the verbal meaning (what all meanings in use have in common) and of the different versions of the meaning of the same word that users learn.

The approach in Zeevat et al. (2015) is an implementation and extension of the approach to lexical ambiguity pioneered by Smolensky (1991) and further developed by Hogeweg (2009), which can be described as: take the maximal amount of content that fits the context. Coercion is part of the mechanism, not a separate process. The model

cus parser is based (Marcus, 1980)

gives a far more detailed picture than just a set of semantic features for lexical meanings. Going for strongest readings is what distinguishes it from approaches such as Asher (2011), which rely on contextual disambiguation and coercion only. Stochastic frames are more conservative than Casasanto and Lupyan (2015) in assuming that observed meanings are stored and serve as a basis for computing meanings in use. Stochastic frames can be learnt and meanings in use can be computed from them by methods that are within the current state of the art.

3.3 Typicality

In section 2.1, we pointed out that stochastic frames fit well to the prototype theory of concepts: understanding prototypes as a weighted sum of property probabilities means to take a stochastic frame *to be* the prototype of the concept. An understanding of prototype concepts as weighted attribute value structures has already been used by Smith et al. (1988) for explaining modifications such as “red apple”. Prototype frames extend this by explicitly using probabilistic weights. In this section, we aim to show that stochastic frames can be used to model one of the core phenomena of prototype concepts, namely the existence of typical and atypical category members. For example, apples are typical fruit, while avocados are not. The structures in (2) are partial frames with probability information for fruit, apple and avocado.

- (2) [fruit
 COLOUR: red 0.3 green 0.1 yellow 0.3 orange 0.2 other 0.1,
 TASTE: sweet 0.6 sour 0.3 other 0.1]
 [apple
 COLOUR: red 0.5 green 0.2 yellow 0.2 orange 0 other 0.1
 TASTE: sweet 0.8 sour 0.1 other 0.1]
 [avocado
 COLOUR: red 0 green 0.7 yellow 0 orange 0 other 0.3,
 TASTE: sweet 0 sour 0 other 1]

Probability information can be used to define diagnostic and frequent properties, i.e. attribute values V , such as sweet taste (Schurz, 2012):

- (3) A property V is frequent for a class C iff $P(V|C)$ is high
 A property V is diagnostic for a class C iff $P(C|V)$ is high

The latter is well-known as the notion of *cue validity*. It allows a definition of the diagnosticity of an attribute A in (4), where V_1, V_2, \dots, V_n are alternative values of the attribute A :

$$(4) \text{diag}(A, C) = \max(P(C|V_1), P(C|V_2), \dots, P(C|V_n))$$

The similarity Sim of the probability distributions of properties on one attribute (i.e., the frequency of the values) in a concept C and another concept, for example, a subcategory SC , can be compared in terms of (5):

$$(5) \text{Sim}(C, SC|A) = \sum_{i=1}^n \min(P(V_i|C), P(V_i|SC))$$

Sim can be used to express that the probability distribution of `COLOR` in ‘apple’ is quite similar to the one for ‘fruit’ ($0.3 + 0.1 + 0.2 + 0.1 = 0.6$) but not so similar for ‘avocado’ and ‘fruit’ ($0 + 0.1 + 0 + 0 + 0.1 = 0.2$).

Finally, the typicality of a subcategory is determined as the diagnosticity-weighted average similarity in all contributing attributes:

$$(6) \text{typ}(C, SC) = \sum_{i=1}^n \frac{\text{diag}(A_i|C)}{\sum_{i=1}^n \text{diag}(A_i|C)} \text{Sim}(C, SC|A_i)$$

With this formula, one can quantify how typical fruit apples or avocados are as a diagnosticity-weighted average of similarities in all contributing attributes.

4 Final remarks

The paper presents a notion of a stochastic frame that represents concepts and the linguistic knowledge of agents in terms of attribute-value structures in which values may only occur with some probability. We outlined how probabilistic constraints on stochastic frames facilitate reasoning about probable features (attribute values) in conditions of uncertainty. What comes out of this can be interpreted as a formalisation of the prototype theory of concepts in which all other theories of concepts can be understood as special cases. By trivialising the distributions, one obtains the classical view. (Products of) regions in conceptual spaces are obtained by deriving such regions from actual distributions (it becomes hard to see such an account of concepts as properly different from the prototype view).

Taking this common core, we also outlined three areas in which stochastic frames have obvious applications: vague predication, lexical ambiguity, and the typicality of kinds. A shared property of these phenomena is arguably that, in all cases, individuals must reason with complex, multifaceted concepts in conditions of uncertainty, be this uncertainty about the extension of a term (vagueness), uncertainty about the meaning of a term in use (lexical ambiguity), or uncertainty about properties of the instances (the typicality of kinds). For an explanation of all of these cases, we seem to need not only something along the lines of a probabilistic component to drive the reasoning process and model graded or fuzzy phenomena, but also a means of applying this reasoning tool to different aspects or properties of the entities being reasoned about. Representational structures such as frames give us the structure we need in this respect. Stochastic frames, therefore, give us the right combination of conceptual structure and a formal theory of reasoning.

Acknowledgements

This research was funded by the German Research Foundation (DFG) funded project: CRC 991 *The Structure of Representations in Language, Cognition, and Science*, specifically projects C09, D01 and a Mercator Fellowship awarded to Henk Zeevat. We would like to thank audiences at CoST 2019 at HHU Düsseldorf, the workshop on Records, Frames, and Attribute Spaces held at ZAS in Berlin, March 2018, and the Workshop on Uncertainty in Meaning and Representation in Linguistics and Philosophy held in Jelenia Góra, Poland, February, 2018.

References

- Nicholas Asher. 2011. *Lexical Meaning in Context: A Web of Words*. Cambridge University Press.
- Lawrence. W. Barsalou. 1992. Frames, concepts, and conceptual fields. In E. Kittay and A. Lehrer, editors, *Frames, fields, and contrasts: New essays in semantic and lexical organization*, pages 21–74. Erlbaum.
- Jeffrey R Binder, Lisa L Conant, Colin J Humphries, Leonardo Fernandino, Stephen B Simons, Mario Aguilar, and Rutvik H Desai. 2016. Toward a brain-based componential semantic representation. *Cogn Neuropsychol*, 33(3-4):130–174.

- Bob Carpenter. 1992. *The logic of typed feature structures*. Cambridge: Cambridge University Press.
- Daniel Casasanto and Gary Lupyan. 2015. All concepts are ad hoc concepts. In Eric Margolis and Stephen Laurence, editors, *The Conceptual Mind: New Directions in the Study of Concepts*, pages 543–566. MIT Press.
- Robin Cooper, Simon Dobnik, Staffan Larsson, and Shalom Lappin. 2015. Probabilistic type theory and natural language semantics. *LILT*, 10(4).
- Charles Dickens. 2000. *David Copperfield*. Modern Library, New York.
- Paul Égré. 2017. Vague judgment: A probabilistic account. *Synthese*, 194(10):3837–3865.
- Peter Gärdenfors. 2000. *Conceptual Space: The Geometry of Thought*. The MIT Press, Cambridge, MA.
- Lotte Hogeweg. 2009. *Word in Process. On the interpretation, acquisition and production of words*. Ph.D. thesis, Radboud University Nijmegen.
- Daniel Lassiter. 2011. Vagueness as probabilistic linguistic knowledge. In R. Nouwen, U. Sauerland, H.C. Schmitz, and R. van Rooij, editors, *Vagueness in Communication*. Springer.
- Sebastian Löbner. 2014. Evidence for frames from human language. In *Frames and concept types*, pages 23–67. Springer International Publishing.
- Mitchell P. Marcus. 1980. *Theory of Syntactic Recognition for Natural Languages*. MIT Press, Cambridge, MA, USA.
- Eric Margolis and Stephen Laurence. 1999. Concepts and cognitive science. In Eric Margolis and Stephen Laurence, editors, *Concepts: Core Readings*. MIT Press, Cambridge, MA.
- Marvin Minsky. 1974. A framework for representing knowledge. *MIT-AI Laboratory Memo*, 306.
- Wiebke Petersen. 2007. Representation of concepts as frames. In J Skilters, editor, *Complex Cognition and Qualitative Science. The Baltic International Yearbook of Cognition, Logic and Communication Vol. 2*, pages 151–170. Springer International Publishing.
- Eleanor Rosch. 1973. Natural categories. *Cognitive Psychology*, 4:328–350.
- Eleanor Rosch. 1978. Principles of categorization. In *Cognition and Categorization*, pages 27–48. Lawrence Erlbaum Associates.
- Eleanor Rosch and Carolyn B. Mervis. 1975. Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, 7(4):573–605.
- Eleanor Rosch, Carolyn B Mervis, Wayne D Gray, David M Johnson, and Penny Boyes-Braem. 1976. Basic objects in natural categories. *Cognitive psychology*, 8(3):382–439.
- Gerhard Schurz. 2012. Prototypes and their composition from an evolutionary point of view. In W. Hinzen, E. Machery, and Werning M., editors, *The Oxford Handbook of Compositionality*, pages 530–553. Oxford University Press, Oxford, UK.
- Edward E. Smith, Daniel N. Osherson, Lance J. Rips, and Margaret Keane. 1988. Combining prototypes: A selective modification model. *Cognitive science*, 12(4):485–527.
- Paul Smolensky. 1991. Connectionism, constituency and the language of thought. In M. Loewer and G. Rey, editors, *Meaning in Mind: Fodor and His Critics*, pages 201–227. Blackwell, Oxford.
- Peter R. Sutton. 2015. Towards a probabilistic semantics for vague adjectives. In Hans-Christian Schmitz and Henk Zeevat, editors, *Language, Cognition, and Mind*. Springer.
- Peter R. Sutton. 2018. Probabilistic approaches to vagueness and semantic competency. *Erkenntnis*, 83(4):711–740.
- Ludwig Wittgenstein. 1953. *Philosophische Untersuchungen*. Suhrkamp Verlag, Frankfurt am Main.
- Henk Zeevat, Scott Grimm, Lotte Hogeweg, Sander Lestrade, and E. Allyn Smith. 2015. Representing the lexicon: Identifying meaning in use via overspecification. In Kata Balogh and Wiebke Petersen, editors, *Proceedings of Workshop Bridging Formal and Conceptual Semantics (BRIDGE-14)*. Düsseldorf University Press.