OPEN ACCESS

University of BRISTOL

Publisher's PDF, also known as Version of record

Link to published version (if available):
10.1016/j.artint.2016.04.008

Link to publication record in Explore Bristol Research
PDF-document

This is the final published version of the article (version of record). It first appeared online via Elsevier at http://www.sciencedirect.com/science/article/pii/S0004370216300492. Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research
General rights
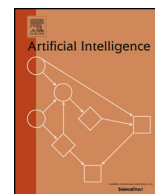
Hierarchical conceptual spaces for concept combination

Martha Lewis [a,*], Jonathan Lawry [b]

[a] *Department of Computer Science, University of Oxford, OX1 3QD, United Kingdom*
[b] *Department of Engineering Mathematics, University of Bristol, BS8 1UB, United Kingdom*

A B S T R A C T

We introduce a hierarchical framework for conjunctive concept combination based on conceptual spaces and random set theory. The model has the flexibility to account for composition of concepts at various levels of complexity. We show that the conjunctive model includes linear combination as a special case, and that the more general model can account for non-compositional behaviours such as overextension, non-commutativity, preservation of necessity and impossibility of attributes and to some extent, attribute loss or emergence. We investigate two further aspects of human concept use, the conjunction fallacy and the 'guppy effect'.

© 2016 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Humans undoubtedly have the ability to form new concepts by combining existing ones. The development of effective representational models of this phenomenon could potentially shed light on human cognition. Human-like reasoning has been argued to be important to artificial intelligence for its flexibility and robustness [6,29,44]. Further, a good representation of human concept use will aid us in considering problems of categorization and typicality, as argued by Freund [18]. Applications of AI that must interact with humans via natural language arguably need to be able to understand and to form for themselves novel combinations of concepts. Examples of theories proposed to account for such concept combination include prototype theory together with fuzzy set theory [51], conceptual spaces [19], and quantum probability [3, 9] approaches. Well-known counterexamples have been identified which suggest that fuzzy sets may not provide an appropriate formalisation in this context [25,27,40]. It is argued in [25] that the failure of fuzzy set theory to adequately model human concept combination results from its failure to consider the intension of concepts, i.e., the attributes that the concept possesses. In contrast, the conceptual spaces and the quantum approaches take intension into account, either by considering concepts as being comprised of a combination of properties,[1] which are themselves embedded in a space of quality dimensions, or by incorporating context into the model. Our proposed approach utilises a random set interpretation of membership so as to quantify an agent's subjective uncertainty about the extent of application of a concept. We refer to this uncertainty as *semantic uncertainty* [33] in order to emphasise that it concerns the definition of concepts and categories. Lawry and Tang [33] combine random set theory with conceptual spaces [19] and prototype theory [43], to give a formalisation of concepts as based on a prototype and an uncertain distance threshold, located in a conceptual space. We use this account of concepts to provide a framework for conjunctive concept combination which captures the effects seen

* Corresponding author.
  E-mail address: martha.lewis@cs.ox.ac.uk (M. Lewis).
[1] In the current paper, we use the terms 'attribute' and 'property' interchangeably.

in [25], including non-compositional behaviours such as overextension, non-commutativity, preservation of necessity and impossibility of attributes and to some extent, attribute loss or emergence.

An outline of the paper is as follows. Section 2 overviews a range of theoretical approaches to concept combination from the literature, and summarises the results from experimental studies that we aim to model. Section 3 describes a random set and prototype theory representational model for concepts within a conceptual space. This model provides the theoretical underpinning for our work. Section 4 introduces a framework for concept combination based on a hierarchy of conceptual spaces, and in which compound concepts are defined within Boolean spaces. We prove a number of results showing the properties of this framework and compare this approach to others in the literature. Section 5 provides a discussion of our results and indicates possible future directions.

2. Background

In this section, we describe a number of approaches to concept combination that have been proposed. We consider general set-theoretic approaches, supervaluation theory, prototype theory, fuzzy set theory, conceptual spaces theory, approaches from computational linguistics and quantum cognition approaches. We further describe some results from experimental studies with which we compare the theory we develop.

2.1. Set-theoretic approaches

Montague semantics [39] takes a model-theoretic approach to concepts and sentences. Concepts are defined using notions from set theory, and natural language expressions are modelled as functions or relations on these sets. This gives a description of how the semantics of a language interacts with the syntax, so that the meaning of a compound expression may be systematically derived from its parts. However, as discussed in [27,28], this is inadequate for modelling some types of adjectives. In [39], an adjective is viewed as a function from properties to properties. This allows sentences such as 'every small elephant is small' *not* to be branded as logically true, which is what we require. This enables various types of adjective to be modelled. *Intersective* adjectives are those where the application of that adjective may simply be viewed as an intersection of sets (such as 'red car'). Adjectives that are not intersective may be *subsective*, when the adjective-noun combination is a subset of the noun, or *non-subsective*, for example privative adjectives like 'fake', or 'former'. However, the theory of adjectives as a function of properties is inadequate, in particular because it doesn't account for comparatives, i.e. the ability to say that x is A-er than y. To account for this, Kamp introduces a theory of vague models, which are viewed as a nested sequence of partial models. In a partial model, a predicate is explained as assigning a value 1 to those objects which fall under the predicate, 0 to those that do not fall under the predicate, and no value to those for whom the predicate is indeterminate. These partial models may be completed in various ways, and the degree of truth of a sentence is related to the probability of a particular set of completions of a partial model of the sentence conditioned on all sets of completions of the model. This set of completed models forms the basis for Kamp's *supervaluation*, where a sentence has truth value 1 if it is true in all completions of the model, 0 if it is false in all completions of the model, and indeterminate if it is true in some and false in others.

Kamp's approach is similar to Fine's [17], in which the questions of the correct logic for vagueness and the correct truth conditions for a vague language are considered. Fine calls the possibility that logical relations hold between indefinite sentences *penumbral connection*, and truths that arise from such a connection *penumbral truths*, and argues that no natural truth-value approach respects such truths. He argues that differences in truth-value within penumbral truths concerning two predicates are essentially a difference in the way that these predicates can be made more precise. He describes a theory of super-truth, in which a sentence is true iff it is true in all admissible and complete specifications of the sentence.

Both these approaches use the idea that there are in fact precise ways of describing a concept, and that the truth value of a sentence using a vague concept is dependent on the different possible ways of making the sentence more precise. In what follows, we do not consider truth values of sentences but rather typicality of an item to a concept. However, consideration of logics using the fuzzy sets we develop would be an interesting line of future work.

Interestingly, [1] argue that adjective-noun combinations can be represented purely as set intersection between the adjective and the head noun. This is achieved by the use of typed sets. These are sets in which members are assigned types. So the adjective 'clever' is represented in the following way:

$$Clever = \{j : human, \ f : pet, \ f : policedog\} : clever$$

where the interpretation of j is 'John', and the interpretation of f is 'Fido'. [1] argue that by using this type of representation the problems of privative adjectives can be circumvented. An example is as follows. From the two sentences 'Maria is a former teacher' and 'Maria is a programmer', we do *not* wish to infer 'Maria is a former programmer'. The typed set representation is as follows:

$$Human = \{m : human, ...\} : human$$

$$Teacher = \{m : teacher, ...\} : teacher$$

$$Former = \{m : teacher, ...\} : former$$

$$Programmer = \{m : programmer, ...\} : programmer$$

Then, we can infer that $m \in Former \cap Teacher$, but not that $m \in Former \cap Programmer$. This is further extended to describe differences in scope when applying multiple adjectives. The approach described is interesting, and could presumably be extended to include some sort of typicality measure.

2.2. Fuzzy set theory and prototype theory

Prototype theory views concepts as being defined in terms of prototypes, rather than by a set of necessary and sufficient conditions. Elements from an underlying metric space then have graded membership in a concept depending on their similarity to a set of prototypical cases. There is some evidence that humans use natural categories in this way; see for example experiments reported in [43]. Fuzzy set theory [49] was proposed as a calculus for combining and modifying concepts with graded membership, and these ideas were then extended [51] to linguistic variables, these being variables taking words as values, rather than numbers. For example, 'height' can be viewed as a linguistic variable taking values 'short,' 'tall', 'very tall', etc.. The variable relates to an underlying universe of discourse Ω, which for the concept 'tall' could be $\mathbb{R}^+$. Then each value L of the variable is associated with a fuzzy subset of Ω, and a function $\mu_L : \Omega \to [0, 1]$ associates with each $x \in \Omega$ the value of its membership in L. Prototype theory gives a semantics for fuzzy set theory through the notion of similarity to a prototype, as described in [15]. In this context, concepts are represented by fuzzy sets and membership of an element in a concept is quantified by its degree of similarity to the prototype. Another possible semantic basis for fuzzy sets is random set theory (see [15] for an exposition). Here, the fuzziness of a set is a result of uncertainty about an underlying crisp set, i.e. semantic uncertainty. Fuzzy set theory seemed initially to be a natural formalisation of prototype theory, since it admits graded membership of concepts. However, work in this area has shown that it is inadequate as a model for human concept combination. A fuzzy set L is defined over a universe Ω via a membership function $\mu_L : \Omega \to [0, 1]$. Elements $x \in \Omega$ that are very good examples of the concept L have membership close to 1, whereas elements x that are bad examples of the concept have membership close to 0. The conjunction of two fuzzy sets is defined purely extensionally, for example $\mu_{L_1 \wedge L_2}(x) = min(\mu_{L_1}(x), \mu_{L_2}(x))$, where $min(a, b)$ indicates the minimum of the two values a and b. Then, overextension of conjunctions of concepts cannot be explained using standard conjunction operators within fuzzy set theory [25,27,40]. Two key examples of this are the conjunction fallacy [47] and the 'guppy effect' [40]. The conjunction fallacy is that humans often judge more specific conditions as more probable than more general conditions. For example, one might judge a bicycle that has been painted with polka dots to be more typical of the combined concept 'polka dot bicycle' than of the concept 'bicycle'. We discuss this further in section 4.4. The 'guppy effect' is introduced in [40], in which Osherson and Smith point out that a guppy, or goldfish, lacks many of the attributes of either a prototypical pet or a prototypical fish, whilst nonetheless being a prototypical example of a pet fish. These difficulties may be partly due to the failure of the fuzzy approach to account for the intension of concepts in the form of the attributes that the concept possesses. In contrast, conceptual space and quantum models are able to represent the intension of concepts, since in each case, a concept is viewed as being embedded in a multidimensional space, whose dimensions are, in some sense, the required attributes.

There is an important distinction between being a *typical example of* and being a *member of* a concept. It is entirely possible for something to be a member of a concept, but not typical of a concept. For example, a duck-billed platypus is a mammal, but it is not a typical mammal. This is explained by [45] by saying that concepts have defining and characteristic features, and that this is what determines concept membership. In [24,26] Hampton argues against this hypothesis. He argues that membership in a conjunction of concepts may be determined by placing a threshold on a judgement of similarity of an item to a composite prototype, so that the two notions of typicality and membership may be attributed to one common cause, and furthermore, that judgements of typicality are correlated with probability of categorization. We subscribe to Hampton's view, as will be seen, allowing for the different weighting of particular attributes, which can thereby contribute to the typicality of an item to a concept. We do not discuss the similarity threshold at which a judgement of membership in a concept should be made. However, we often use the notion of membership in a fuzzy set as a proxy for typicality, and in particular we use the terminology 'membership function' and 'membership value'. These should be seen as akin to typicality ratings in all that follows.

2.3. Conceptual spaces

Conceptual spaces are proposed in [19] as a framework for representing information at the conceptual level. Gärdenfors contrasts his theory with both a symbolic, logical approach to concepts, and an associationist approach where concepts are represented as associations between different kinds of basic information elements. Rather, conceptual spaces are geometrical structures based on quality dimensions such as weight, height, hue, brightness, etc. It is assumed that conceptual spaces are metric spaces, with an associated distance measure. This might be Euclidean distance, or any other appropriate metric. The distance measure can be used to formulate a measure of similarity, as needed for prototype theory, according to which similar objects are close together in the conceptual space and very different objects are far apart.

To develop the conceptual space framework, Gärdenfors also introduces the notion of integral and separable dimensions. Dimensions are integral if assignment of a value in one dimension implies assignment of a value in another, such as depth and breadth. Conversely, separable dimensions are those where there is no such implication, such as height and sweetness. A *domain* is then defined as a set of quality dimensions that are separable from all other dimensions, and a *conceptual space* is defined as a collection of one or more domains. Gärdenfors goes on to define a *property* as a convex region of a domain

Table 1
(a) Correlation strengths of 'red' and 'round' in 'apple'. (b) Correlation strengths of 'red' and 'round' in 'cricket ball'.

(a)

C^1	Red	Round
Red	1	0.5
Round	0.5	1

(b)

C^2	Red	Round
Red	1	1
Round	1	1

in a conceptual space. Finally, a *concept* is defined as a set of such regions that are related via a set of salience weights. This casting of (at least) properties as convex regions of a domain sits very well with prototype theory, as indeed Gärdenfors points out. If properties are convex regions of a space, then we can say that an object is more or less central to that region. Because the region is convex, its centroid will lie within the region, and this centroid can be seen as the prototype of the property.

There are a few approaches to defining concept composition based on conceptual spaces. Firstly, Gärdenfors proposed that when combining a pair of concepts as he defines them, properties in one concept are replaced by properties from the other, depending on the salience, or weighting, of each concept and each property. He goes on to introduce the notion of a contrast class. This has the effect that a particular property is restricted to a certain area. For example when talking about red wine, the concept 'red', determined by the contrast class 'wine', is a subset of the standard concept red. In order to model this, Gärdenfors maps the whole of the colour domain onto the subset of colours that can apply to wine. The formal rule for concept combination is then that the combination CD of two concepts C and D is determined by letting the regions for the domains of C confined to the contrast class defined by D, replace the values of the corresponding regions for D. So in the example of 'red wine', the space of colours has been restricted by the noun 'wine' to a subset of the full colour space with the same geometry. The colour of the wine is then taken to be the colour that is occupied by 'red' within the restricted colour space.

Consistent with this high level description of concept combination, we now describe below two more formal approaches based on conceptual spaces. Adams and Raubal [2] give a fairly straightforward formalisation of Gärdenfors's account, within which a conceptual space consists of a 6-tuple of domains, concepts, instances, contrast classes, contexts, and a similarity sensitivity parameter. Each domain is a set of quality dimensions. A concept is defined as a pair consisting of a set of convex regions of domains together with a prototypical instance P, and a property as a concept that includes only one domain region. A contrast class is defined as a region of a unit hypercube corresponding to a domain. Although it is not entirely clear why a unit hypercube is used rather than the domain itself, this is presumably as a way of normalising the dimensions of the domain before the contrast is applied. A context is defined as a finite set of salience weights. In [2], Adams and Raubal go on to define three types of concept combination: property–concept, concept–concept and contrast–class–concept combination and give algorithms for the implementation of each type of combination. However, they do not attempt to account for the fuzziness of natural concepts, or give any account of non-compositional features.

Another approach that gives a formal definition of conceptual spaces is described by Rickard et al. in [42]. This views concepts as a function from pairs of properties into a unit interval. These properties are defined as fuzzy sets in a domain Dom_i. A concept C is therefore a set of correlations between pairs of properties (a, b) where a, b belong to a set of properties. Each pair of properties (a, b) has a value C_{ab} in a concept, which gives the strength of the correlation of a and b in the concept. For example, consider the concept *Banana*. The property *yellow* and the property *sweet* are highly correlated, and the property *green* and the property *bitter* are highly correlated. A context is defined as a set of properties, and the similarity between two concepts C^1 and C^2 as the mutual subsethood of C^1 and C^2 relative to that context. The mutual subsethood functions as a way of determining the overlap of two concepts. It is defined by:

$$s(C^1, C^2) = \frac{\sum_{a,b} min(C^1_{ab}, C^2_{ab})}{\sum_{a,b} max(C^1_{ab}, C^2_{ab})}$$

where C_{ab} is the value of the correlation between properties a and b on concept C.

For example, suppose our context is the set {*red*, *round*} and our objects are $C^1 = apple$, $C^2 = cricket ball$. The strengths of the correlations are given in Tables 1a and 1b.

Then

$$s(C^1, C^2) = \frac{\sum_{a,b} min(C^1_{ab}, C^2_{ab})}{\sum_{a,b} max(C^1_{ab}, C^2_{ab})} = \frac{1 + 0.5 + 0.5 + 1}{1 + 1 + 1 + 1} = \frac{3}{4}$$

The role of the context is to determine which properties are relevant in calculating the similarity. The membership of an observation in a concept is defined as the similarity of the observation to a given concept, and the label is then assigned which has maximum membership for the given observation. Dynamics on the space are also introduced, which allow properties to be prioritised for attention. Composition of concepts is carried out by combining properties by taking the union of the property sets so that the resulting combined concept has properties that belong to both constituent concepts. Whilst Rickard et al. do model fuzziness, they do not attempt to account for non-compositional features of human concept use.

2.4. Computational linguistics and vector space models

Within the field of computational linguistics, vector-based models of word meaning have proved very fruitful. The meaning of a particular word is represented as a vector, where the basis of the vector space might be a chosen set of words (usually, the most common, excluding a list of stop-words such as 'a', 'the', 'and', and so on), or some other carefully chosen dimensions, and the entries in the vector are word co-occurrence statistics, or a relation between the words and the documents they occur in [10,31,37]. A comprehensive paper by Mitchell and Lapata [38] gives a comparison of various techniques for adjective-noun composition. Related approaches are given in [7,11], where adjectives are viewed as matrices, and nouns as vectors. Whilst these approaches have considerable merit, the underpinning space cannot be viewed as a conceptual space that describes features of the concept. The relationship of individual dimensions to the vectors are not attributes, but could be instances, parts of, or any other incidental relationship. The development of a suitable conceptual space for these models would be an interesting line of future research.

Another approach within the computational linguistics framework is the development of a family of 'microtheories' of word meanings. [41] develop a microtheory of adjectives whereby the analysis given in [27] is extended to examine what a number of linguists determine to be the taxonomy of adjectives. Words are represented as having syntactic and semantic types. The syntactic type describes how the word can be combined with others. The semantic type describes the semantic effect of making such a combination. So, for example, the adjective 'big' can be applied in the combinations Adj-Noun or Noun-Copula-Adj, and has the semantic effect that it can be applied to physical objects and limits the normalised value of the size property to greater than 0.75. Adjectives are divided into *scalar* – based on properties, *denominal* – based on object, and *deverbal* – based on processes. The distinctions between these types lie in their semantics, namely the different ways in which they combine with nouns to form a composite. In the work we present here we focus on intersective adjectives, since as pointed out in [27] even these require work to clarify how typicality functions in a composite concept, and therefore we do examine the difference in semantic types. Explaining these differences will be an interesting area for further work, however.

2.5. Quantum probability models

The quantum probability model introduced by Aerts [3], sees a concept as a quantum entity within a vector space, the dimensions of which are the contexts of the concept. When no context is present for the concept, the concept is in its ground state. The application of a context then changes this concept into the concept under that context. Typicality of an item to a concept changes with context. As such, problems such as the 'guppy effect' are accounted for by noting that the typicality of a guppy to the concept pet in its ground state differs from the typicality of a guppy to the concept pet in the context 'the pet is a fish'. Aerts et al [5] give a description of how the effects of contextuality, interference, entanglement and emergence may be seen in human concept use. Contextuality may be seen in the way that the typicality of an element to a concept changes with the context given. The phenomenon of interference in quantum vector spaces allows over- and under-extension to be modelled when combining concepts. Briefly, membership in a concept is modelled by the projection of the concept onto the subspace representing the item. When evaluating the membership of an item to the composite concept '*A* and *B*', it may be the case that an interference term needs to be introduced. This interference term accounts for over- or under-extension. This idea is explained in detail in [4], in which data from [23,24] is modelled. The phenomenon of entanglement is found to be present in data concerning the applicability of combinations of concept pairs. The concept of emergence is explained as the idea that a totally new concept has been introduced by forming a conjunction of concepts. To account for this, Aerts et al. propose the use of Fock space. In Fock space, an entity may be in a superposition of states. In the case of concept combination, one of these states is the completely new concept, and another is the concept as a combination of two concepts. Using these notions, the quantum probability model develops ways of modelling which account for both the fuzziness of human concept use and effects of non-compositionality. Within this paper, we aim to show that our approach can account for these aspects of human concept use within a simpler and more intuitive framework.

2.6. Experimental studies

Hampton [22] reports results from two experiments. The aim of the first was to generate a list of attributes for each of six pairs of concepts and their conjunctions. An example is the pair of concepts Sports and Games. A list of attributes was collected for each of these concepts, and for the conjunctions 'Games which are Sports' and 'Sports which are Games'. This was repeated for each of the six pairs of concepts. Based on these lists of attributes, the second experiment asked participants how useful each attribute was in defining the concept, as measured on the scale shown in Table 2 (the numerical value was imposed later rather than given by participants).

Averaging across subjects then gives the mean importance rating for each attribute in each concept. Some attributes have similar importance within pairs of concepts, and some differ. For example, the attribute 'Is used by people' has a mean rating of 3.00 for both Machines and for Vehicles. However, 'Replaces people' has a mean rating of 2.00 for Machines, and −1.00 for Vehicles. The challenge then is to predict the importance of an attribute for a combined concept such as 'Machines which are also Vehicles' from the attribute weightings of the constituent concepts. Hampton reports that using multiple regression to obtain weight coefficients for a weighted sum provides the best predictor of attribute weightings in

Table 2
Table of ratings.

Code	Importance	Numerical value
N	Necessarily true of all possible examples of the concept	4
A	A very important part of the definition	3
B	A fairly important part of the definition	2
C	Typically true but not very defining	1
X	Not usually true of all examples of the concept	−1
XX	Necessarily false of all possible examples of the concept	−2

the combined concepts, but that noncompositionality is also observed. For example, some attributes with low importance in the constituent concepts may have a high importance in the combined concept. This is termed 'attribute emergence' – the attribute 'Lives in a cage' has low importance for 'Pet' and for 'Bird', but high importance for 'Pet which is also a Bird'. A similar way in which noncompositionality manifests itself is in the preservation of necessary or impossible attributes. When an attribute is seen as necessary (or impossible) for one of the constituent concepts, that importance rating is carried over into the attributes for the combined concept. Therefore, there is no functional relationship between the importance of an attribute in the constituent concepts, and the importance in the combined concept. Rather, this depends on the particular concepts involved. Hampton finds that conjunction is not commutative in that the qualifying noun, i.e. the second noun in the conjunction, is given more weight. Lastly, dominance effects are also seen, in that concepts which bring more attributes to the conjunction tend to dominate.

Hampton therefore reports the following six main results:

- The attribute set for combined concepts is the union of the attribute sets of the constituent concepts
- The importance of attributes in the combined concept is usually a weighted sum of the importance of the attributes in each individual concept
- Necessity and impossibility are preserved
- Attribute loss or emergence is observed
- Conjunction is not commutative
- Dominance effects are observed

We will argue that our proposed model of concepts and concept combination can also account for these phenomena. Furthermore, our framework is a natural extension of the conceptual spaces model in which the importance of certain dimensions is related to their necessity as defined by possibility theory [16].

3. Formal model of concepts

In this section we outline our conceptual spaces based model of concepts which forms the theoretical underpinnings of our work. This model of concepts combines a prototype theory approach with random sets, capturing both typicality and semantic uncertainty, first outlined by Lawry and Tang in [33]. We will go on to build on this model of concepts to form a framework for concept combination.

3.1. A prototype and random set model of concepts

In this framework, agents use a set of labels $L = \{L_1, L_2, ..., L_n\}$ to describe an underlying conceptual space Ω which has a distance metric $d(x, y)$ between points.[2] If one of x or y is a set then we take the distance to be the minimum distance to any point in the set. For example, suppose Y is a set, then $d(x, Y) = \min\{d(x, y) : y \in Y\}$. Each label L_i is associated firstly with a set of prototype values $P_i \subseteq \Omega$, and secondly with a threshold ε_i, about which the agents are uncertain. The thresholds ε_i are drawn from probability distributions δ_i. Labels L_i are associated with neighbourhoods $\mathcal{N}_{L_i}^{\varepsilon_i} = \{x \in \Omega : d(x, P_i) \leq \varepsilon_i\}$. The neighbourhood can be seen as the extension of the concept L_i. The intuition here is that ε_i captures the idea of being sufficiently close to prototypes P_i. In other words, $x \in \Omega$ is sufficiently close to P_i to be appropriately labelled as L_i providing that $d(x, P_i) \leq \varepsilon_i$. This is illustrated in Fig. 1.

Given an element $x \in \Omega$, we can ask how appropriate a given label is to describe it. This is quantified by a membership function, denoted $\mu_{L_i}(x)$, corresponding to the probability that the distance from x to P_i, the prototype of L_i, is less than the threshold ε_i, as given by:

$$\mu_{L_i}(x) = P(\varepsilon_i : x \in \mathcal{N}_{L_i}^{\varepsilon_i}) = P(\varepsilon_i : d(x, P_i) \leq \varepsilon_i) = \int_{d(x, P_i)}^{\infty} \delta_i(\varepsilon_i) \mathrm{d}\varepsilon_i$$

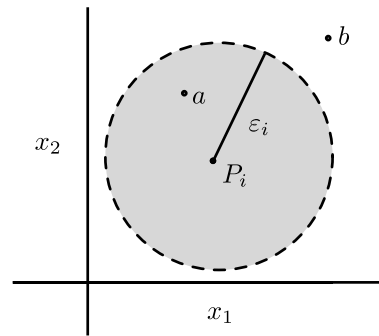[2] In fact, it is sufficient that $d(x, y)$ be a pseudo distance.

Fig. 1. Prototype-threshold representation of a concept L_i. The conceptual space has dimensions x_1 and x_2. The concept has prototype P_i and threshold ε_i. The uncertainty about the threshold is represented by the dotted line. Element a in the conceptual space is within the threshold, so we can say that a has property L_i. Element b is outside the threshold, so b does not have the property L_i. The neighbourhood $\mathcal{N}_{L_i}^{\varepsilon_i} = \{x \in \Omega : d(x, P_i) \leq \varepsilon_i\}$ corresponds to the shaded area.
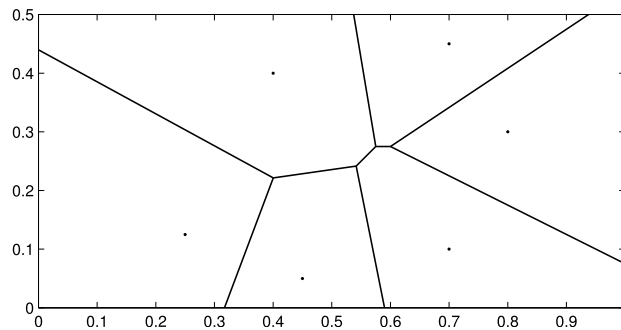

Fig. 2. Conceptual space divided into concepts according to a Voronoi tessellation around prototypes. Each part of the space corresponds to exactly one concept.

We also use the notation $\int_{\varepsilon_i}^{\infty} \delta_{\varepsilon_i}(\varepsilon_i)d\varepsilon_i = \Delta_i(\varepsilon_i)$, according to which $\mu_{L_i}(x) = \Delta_i(d(x, P_i))$. The above formulation provides a link to the random set formalisation of fuzzy sets. Random sets are random variables taking sets as values. If we view $\mathcal{N}_{L_i}^{\varepsilon_i}$ as a random set from $\mathbb{R}^+$ into 2^Ω, then $\mu_{L_i}(x)$ is the single point coverage function of $\mathcal{N}_{L_i}^{\varepsilon_i}$, as defined in [32].

Each label L_i is entirely defined by its prototype P_i, the distance metric in the space $d(x, y)$ and the distribution δ_i of the threshold ε_i. We can therefore, given a particular conceptual space Ω, use the notation $L_i = < P_i, d_i, \delta_i >$ to completely describe the label L_i.

The idea of a membership function presented here may be compared with the similarity relation that Gärdenfors uses. A similarity relation between points in a conceptual space may be defined as a decreasing function of distance in the space. Gärdenfors gives the example that the similarity $s(x, y)$ between two points x and y in the conceptual space is an exponentially decaying function of the distance $d(x, y)$ between the two points, i.e. $s(x, y) = \exp(-cd(x, y))$. In terms of the prototype-threshold approach outlined above, the membership of an element x in a concept L may then be defined as the similarity to the prototype P of L, where $\varepsilon \sim Exp(c)$. More generally, $s(x, y) = \Delta(d(x, y))$.

This approach is, however, in contrast to Gärdenfors' original approach which is to view the space as partitioned by a Voronoi tessellation. If this latter approach is taken, each individual point in the conceptual space is allocated to exactly one label. With a prototype-threshold approach, it is easy to accommodate the idea of an object being accurately described by more than one concept, or conversely, some points within the space not being assigned to any concept. This difference is illustrated in Figs. 2 and 3.

The Voronoi diagram approach to describing concepts can be extended to include graded boundaries, and such an approach is developed in [13,14]. We argue that a drawback to this type of representation is that every single point has been categorised. In contrast, in the label semantics approach there can be points which have not been categorised. This is desirable: imagine the first Western scientists to encounter a duck-billed platypus. It is not clear how to categorise this animal, and it could be modelled as being in a region of space that has not yet been assigned to a category. Another advantage of the neighbourhood model is that concepts can overlap which allows us specifically to refer to borderline regions. However, a benefit of the Voronoi tessellation approach to concept representation is that the membership of a point in a concept depends not only on the prototype of that concept but on the proximity of other prototypes. To integrate this aspect into the label semantics approach to concepts would be an interesting area for future research.
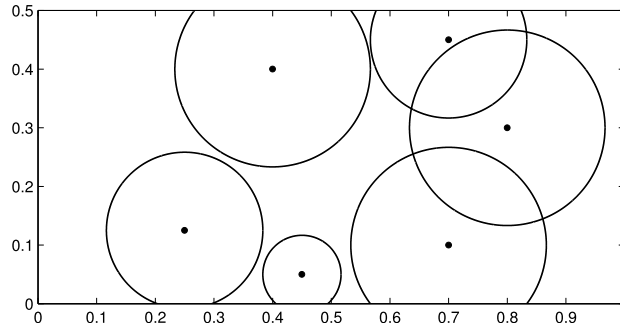
Fig. 3. Conceptual space divided into concepts according to a prototype-threshold approach. Some points in the space correspond to more than one concept, and some correspond to none.
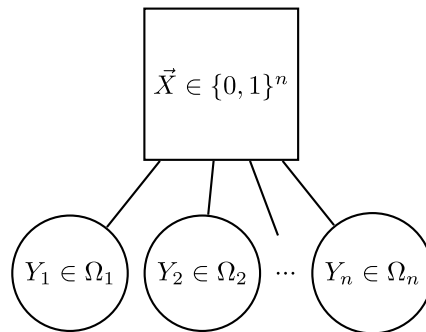

Fig. 4. Schematic of a hierarchical conceptual spaces model for combining concepts.

4. A hierarchical model of conjunctive composition

4.1. Background

As described in section 2, Hampton [22] gives a series of results on human understanding of conjunctive concepts, such as 'sports that are games'. It had already been shown [40,46] that standard fuzzy set-theoretical conjunctions and disjunctions do not adequately model human understanding of composite concepts. Hampton's work elicits data that could form the basis of a model of conjunction that more accurately reflects how humans understand conjunctive concepts.

4.2. A new approach to concept composition

An initial approach to modelling Hampton's data within the conceptual spaces framework would be to view individual attributes, for example 'Talks', 'Has fur', 'Has claws', as each forming a dimension of the conceptual space. However, these attribute dimensions are very different from the usual conceptual space dimensions in two ways. Firstly, they are mostly binary, unlike dimensions such as 'height', 'depth' or 'breadth'. Secondly, they are very complex in comparison to the types of dimensions proposed by Gärdenfors. For instance, having feathers seems to be a multidimensional concept in itself.

This motivates a new hierarchical formulation of conceptual spaces in which we model attributes as labels, each taken from individual domains. In Gärdenfors' terminology, each label would be a property from an integral domain. So an attribute like 'rounded' is seen as a label based in a space such as $\mathbb{R}^3$, or 'red' as based in the CIELab colour space. From this perspective, each of the attribute labels can form a *binary* dimension which are then combined to form the space $\{0, 1\}^n$ where n is the number of attributes. Within this binary space, we take the value 1 on a particular dimension to mean that an object has that particular property. Fig. 4 gives a schematic representation of this model, within which we treat the combination space $\{0, 1\}^n$ *itself* as a conceptual space with an associated metric. This enables us to apply the neighbour-based prototype model of concepts outlined in section 3 to form compound concepts made up of many properties. The motivation for treating this combination space itself as a conceptual space is that if we view each label as a property in an integral domain, then this is precisely a formalisation of the conceptual spaces that Gärdenfors proposes. Gärdenfors suggests both a weighted sum of properties and a weighted Euclidean distance metric in the property space. The formalism we propose corresponds to the weighted sum of properties, but generalises it. If we were to use a cube $[0, 1]^n \in \mathbb{R}^n$, then we might be able to give the weighted Euclidean distance as a special case. This is an area for further work, however.

In the sequel we formalise this idea and prove a number of key results concerning conjunctive concepts defined in this way. We show that if the threshold of the compound concept in the binary combination space is uniformly distributed,
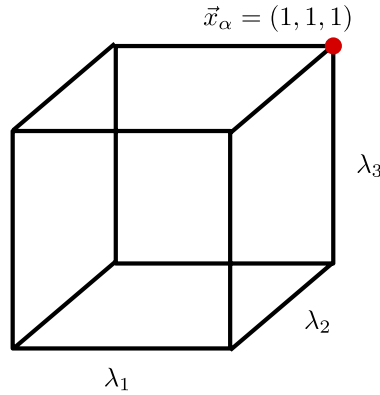
Fig. 5. The prototype for $\alpha = L_1 \wedge L_2 \wedge L_3$ is $\vec{x}_\alpha = (1,1,1)$. Each dimension i is weighted by λ_i.

the membership function for the compound concept is shown to be a weighted sum of the membership functions of the individual labels. This result nicely parallels Zadeh's operation of convex combination, and Gärdenfors' proposal that concepts should be seen as sets of properties related by salience weights. Lastly, we will show that under certain conditions, the importance of an attribute in a conjunction of two compound concepts can be calculated as the weighted sum of the importances of the individual attributes, directly mirroring Hampton's results.

A conjunctive label is defined in a binary space as follows. Consider a set of distinct integral domains $\Omega_1, \ldots, \Omega_n$, such as the CIELab colour space, size, and taste. We select a label from each domain for combination. So an apple might be described as red and sweet and medium sized. This gives us a set $LA = \{L_1, \ldots, L_n\}$ where $L_i \subseteq \Omega_i$ for $i = 1, \ldots, n$.

Each label L_i is defined by the triple $< P_i, d_i, \delta_i >$, as described in section 3, where the prototype $P_i \subseteq \Omega_i$, d_i is the distance metric in Ω_i, the threshold ε_i is a random variable into $\mathbb{R}^+$ and δ_i is a probability density on ε_i. We can then define a Boolean variable X_i into $\{0, 1\}$ with reference to a point $Y_i \in \Omega_i$ for $i = 1, \ldots, n$ as follows:

$$X_i = \begin{cases} 1 : d_i(Y_i, P_i) \le \varepsilon_i \\ 0 : \text{otherwise} \end{cases}$$

Here $X_i = 1$ means that the object being described has the property L_i, i.e., $X_i = 1$ iff $Y_i \in \mathcal{N}_{L_i}^{\varepsilon_i}$, where $\mathcal{N}_{L_i}^{\varepsilon_i}$ is the neighbourhood of L_i, as described in section 3. Also, $P(X_i = 1|Y_i) = P(d_i(Y_i, P_i) \le \varepsilon_i) = \mu_{L_i}(Y_i)$.

A vector $\vec{Y} \in \Omega_1 \times \ldots \times \Omega_n$ generates a Boolean vector $\vec{X}$ into $\{0, 1\}^n$. In this case, the probability distribution for X_i is determined by δ_i.

Now consider a *conjunctive concept* as being defined by a conjunction of labels or their negations, covering all labels in LA. It is therefore of the following form:

$$\alpha = \bigwedge_{i=1}^{n} \pm L_i$$

where $+L_i = L_i$ and $-L_i = \neg L_i$. Expressions of this type are referred to as atoms.

Each atom then naturally defines a point in $\{0, 1\}^n$ as follows:

$$\vec{x}_\alpha = < x_{1,\alpha}, x_{2,\alpha}, \ldots, x_{n,\alpha} > \text{ such that } x_{i,\alpha} = \begin{cases} 1 : \text{if } L_i \text{ appears in } \alpha \\ 0 : \text{if } \neg L_i \text{ appears in } \alpha \end{cases}$$

We think of the space $\{0, 1\}^n$ as the binary conjunction space, and of $\vec{x}_\alpha$ as the prototype of the conjunctive concept α. We also allow some deviation from the prototype by taking into account the different levels of importance of each label L_i. The differing importance of the labels is characterised by a weight vector $\vec{\lambda}$ which weights each dimension in the binary space. These ideas are illustrated in Fig. 5.

We can now consider membership in the conjunctive concept within the binary space. A conjunctive concept is defined by the triple $\alpha = < \vec{x}_\alpha, d, \delta >$ where α is an atom of LA, d is a distance on $\{0, 1\}^n$, ε is a random variable in $\mathbb{R}^+$ and δ is a probability distribution on ε. We say that an element $\vec{X} \in \{0, 1\}^n$ can be appropriately described by the concept α iff $d(\vec{x}_\alpha, \vec{X}) \le \varepsilon$ with the membership function defined by:

$$\mu_\alpha(\vec{X}) = \delta(\{\varepsilon : d(\vec{X}, \vec{x}_\alpha) \le \varepsilon\}) = \Delta(d(\vec{X}, \vec{x}_\alpha))$$

We can then relate membership in $\Omega_1 \times \Omega_2 \times \ldots \times \Omega_n$ to membership in the binary space $\{0, 1\}^n$ as follows.
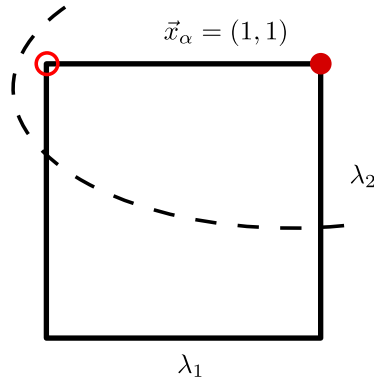
Fig. 6. Prototype for $\alpha = L_1 \wedge L_2$ and weighted dimensions in a two dimensional binary space, together with the threshold ε for α. The point $(0, 1)$, indicated by an open circle, can be considered to be an instance of the concept for which x_α is the prototype.

We define a binary random variable Z into $\{0, 1\}$ such that:

$$Z = \begin{cases} 1 : d(\vec{X}, \vec{x}_\alpha) \leq \varepsilon \\ 0 : \text{otherwise} \end{cases}$$

Clearly in this case $P(Z = 1|\vec{X}) = \mu_\alpha(\vec{X})$. Now by total probability we have that:

$$P(Z|\vec{Y}) = \sum_{\vec{X}} P(Z|\vec{X}, \vec{Y})P(\vec{X}|\vec{Y}) = \sum_{\vec{X}} P(Z|\vec{X})P(\vec{X}|\vec{Y})$$

We may assume that $\vec{Z}$ and $\vec{Y}$ are conditionally independent given $\vec{X}$, since $\vec{Z}$ is defined purely in terms of $\vec{X}$.

Letting $\mu_\alpha(\vec{Y})$ denote $P(Z = 1|\vec{Y})$ and assuming independence of the dimensions $i = 1...n$ we then have that:

$$\mu_\alpha(\vec{Y}) = \sum_{\vec{X}} \mu_\alpha(\vec{X})P(\vec{X}|\vec{Y}) = \sum_{\vec{X}} \mu_\alpha(\vec{X}) \prod_{i=1}^{n} P(X_i|Y_i)$$

More generally, we can define a *compound concept* with prototypical case $\theta = \bigwedge_I \pm L_i$, where $I \subseteq \{1, ..., n\}$, as a triple $\theta =< P, d, \delta >$ where:

$$P = \{\vec{x} =< x_1, x_2, ..., x_n >\in \{0, 1\}^n : x_i = 1, \text{ if } L_i \text{ appears in } \theta, x_i = 0, \text{ if } \neg L_i \text{ appears in } \theta\}$$

P is therefore a set of points which all have the same values on the dimension specified by the index set I, and cover all remaining possibilities across the dimensions not in I. This implies that where $I = \{1, ..., n\}$, P is a singleton.

In this case we have that:

$$\mu_\theta(\vec{X}) = \delta(\{\varepsilon : d(\vec{X}, P) \leq \varepsilon\}) \text{ where } d(\vec{X}, P) = \min\{d(\vec{X}, \vec{x}) : \vec{x} \in P\} \text{ and}$$

$$\mu_\theta(\vec{Y}) = \sum_{\vec{X}} \mu_\theta(\vec{X}) \prod_{i=1}^{n} P(X_i|Y_i)$$

We now define a distance metric in the binary space $\{0, 1\}^n$ based on Hamming distance and a weight vector.

Definition 1 *(One dimensional Hamming distance).*

$$\forall x, x' \in \{0, 1\}, \ H_1(x, x') = \begin{cases} 1 : x = x' \\ 0 : \text{otherwise} \end{cases}$$

Definition 2 *(Weighted Hamming distance).* For $\vec{\lambda} \in (\mathbb{R}^+)^n$,

$$\forall \vec{x}, \vec{x}' \in \{0, 1\}^n, \ H_{\vec{\lambda}}(\vec{x}, \vec{x}') = \sum_{i=1}^{n} \lambda_i H_1(x_i, x'_i)$$

The effect of this distance metric on membership in the binary space is illustrated in Fig. 6. Suppose that the concept 'bird' is characterised, for illustrative purposes, by two properties $L_1 =$ 'flies', $L_2 =$ 'has feathers'. The property L_1 may be relaxed, since there are birds which do not fly. So animals which have feathers, but do not fly, are still considered birds but

not typical birds. We characterise this using the weights in the binary space. Therefore in this case, the weight on the first dimension, λ_1 will be smaller than λ_2. The effect this has is to create elliptical neighbourhoods in the space.

We now outline a correspondence between our idea of a compound concept as a conjunction of attribute labels or their negations, and Hampton's account of concepts as combinations of attributes with individual weights. Firstly, as stated above, we model individual attributes as labels from conceptual spaces such as the colour space, or the taste space. We have a vector of weights, $\vec{\lambda}$, attached to the binary space which loosely corresponds to Hampton's attribute weights. However, the weights in Hampton's account range from 4, being necessary, to -2, being impossible. In contrast, our weight vector $\vec{\lambda}$ is always positive, and the idea of an attribute L_i being atypical or impossible is captured by the notion that the conjunction $\bigwedge_{i=1}^{n} \pm L_i$ includes $\neg L_i$. The extent of the atypicality of the attribute L_i is then given by the weight λ_i of the corresponding dimension.

4.3. Properties of the hierarchical model

We now give a series of results concerning the formulation and properties of compound concepts. We firstly give an example of how two properties may be combined. In section 4.3.1 we show that as a special case, the membership function of a compound concept reduces to a weighted sum of the membership functions of the individual constituent concepts, thereby giving a mathematical grounding to the ideas proposed in [19] of seeing a concept as a weighted combination of properties, or in [50] of forming complex concepts via a mechanism of *convex combination*. Lastly, section 4.3.2 shows how the conjunction of two compound concepts can again be modelled as a weighted sum. This result models Hampton's results [22], outlined in 4.1.

Example 3 (*Property–property combination*). Suppose we are combining the labels 'tall' and 'thin'. Let $P_{tall} = \{Y_1 \in \mathbb{R}^+ : x_1 > 175\}$ and $P_{thin} = \{Y_2 \in \mathbb{R}^+ : x < 24\}$. Define 'tall' $= < P_{tall}, || \bullet ||, \delta_{tall} >$ and 'thin' $= < P_{thin}, || \bullet ||, \delta_{thin} >$, where the distributions δ_{tall} and δ_{thin} are defined piecewise as Beta distributions on normalised variables for height and waist size within boundaries that we have chosen, and 0 otherwise, i.e.:

$$\delta_{tall}(d(x, P_{tall})/20) = \begin{cases} Beta(1, 5) \text{ if } d(x, P_{tall})/20 \leq 1 \\ 0 : \text{otherwise} \end{cases}$$

$$\delta_{thin}(d(x, P_{thin})/4) = \begin{cases} Beta(3, 1) \text{ if } d(x, P_{thin})/4 \leq 1 \\ 0 : \text{otherwise} \end{cases}$$

We use Beta distributions in this example, since this distribution produces various different shapes depending on the parameters used.

Suppose $\alpha = $ 'tall and thin' is mapped into $\{0, 1\}^2$ with $\lambda_1 = 0.4$, $\lambda_2 = 1 - \lambda_1 = 0.6$, and that the threshold ε is distributed according to $\delta = Beta(2, 1)$.

Then $\alpha = < (1, 1), H_{\vec{\lambda}}, \delta >$, and

$$\mu_{\text{tall and thin}}(\vec{Y}) = \Delta_\alpha(0)\mu_{\text{tall}}(Y_1)\mu_{\text{thin}}(Y_2) + \Delta_\alpha(1 - \lambda_1)\mu_{\text{tall}}(Y_1)(1 - \mu_{\text{thin}}(Y_2))$$
$$+ \Delta_\alpha(\lambda_1)(1 - \mu_{\text{tall}}(Y_1))\mu_{\text{thin}}(Y_2) + \Delta_\alpha(1)(1 - \mu_{\text{tall}}(Y_1))(1 - \mu_{\text{thin}}(Y_2))$$

where $\vec{Y} = (Y_1, Y_2) \in \mathbb{R}^2$.

This membership function is illustrated in Fig. 7.

4.3.1. Results for compound concepts using hamming distance

The example above shows that the membership functions generated within this framework can be very flexible. However, we show that by restricting the type of membership function used in the binary combination space, we can derive an expression for the membership function $\mu_\theta(\vec{Y})$ of the compound concept $\theta = \bigwedge_{i=1}^{k} \pm L_i$, $k \leq n$ as a weighted sum of the membership functions for individual domains $\mu_{\pm L_i}(Y_i)$. This grounds proposals in [19,30,50] that complex concepts can be built up as sums of weighted properties.

Theorem 4. *Let $\alpha = \bigwedge_{i=1}^{n} \pm L_i$ and $\lambda_T = \sum_{i=1}^{n} \lambda_i$. Let δ be the uniform distribution on the interval $(0, \lambda_T)$. If d is the weighted Hamming distance $H_{\vec{\lambda}}$ then:*

$$\mu_\alpha(\vec{Y}) = \sum_{i=1}^{n} \frac{\lambda_i}{\lambda_T} \mu_{\pm L_i}(Y_i)$$

Proof. W.l.o.g assume $\alpha = \bigwedge_{i=1}^{n} L_i$. Notice that $\Delta(u) = \frac{\lambda_T - u}{\lambda_T}$ for $u \in [0, \lambda_T]$, and therefore:
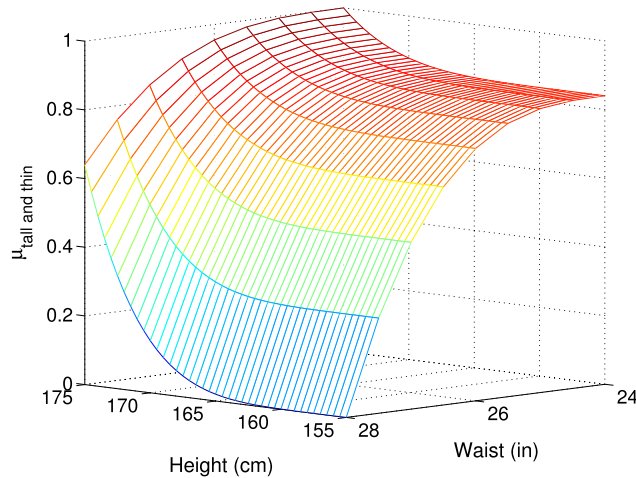
Fig. 7. $\mu_{\text{tall and thin}}(\vec{Y})$.

$$\mu_\alpha(\vec{X}) = \Delta(H_{\vec{\lambda}}((1,\dots,1),\vec{X})) = \Delta(\sum_{i:X_i=0} \lambda_i) = \frac{\lambda_T - \sum_{i:X_i=0}\lambda_i}{\lambda_T} = \frac{\sum_{i:X_i=1}\lambda_i}{\lambda_T}$$

From this we have that:

$$\mu_\alpha(\vec{Y}) = \sum_{\vec{X}} (\frac{\sum_{i:X_i=1}\lambda_i}{\lambda_T}) \prod_{j=1}^n P(X_j|Y_j) = \frac{1}{\lambda_T} \sum_{\vec{X}} \sum_{i:X_i=1} \lambda_i \prod_{j=1}^n P(X_j|Y_j)$$

$$= \frac{1}{\lambda_T} \sum_{i=1}^n \sum_{\vec{X}:X_i=1} \lambda_i \prod_{j=1}^n P(X_j|Y_j) = \frac{1}{\lambda_T} \sum_{i=1}^n \sum_{\vec{X}:X_i=1} \lambda_i \mu_{L_i}(Y_i) \prod_{j\neq i} P(X_j|Y_j)$$

$$= \frac{1}{\lambda_T} \sum_{i=1}^n \lambda_i \mu_{L_i}(Y_i) \sum_{\vec{X}:X_i=1} \prod_{j\neq i} P(X_j|Y_j) = \frac{1}{\lambda_T} \sum_{i=1}^n \lambda_i \mu_{L_i}(Y_i) = \sum_{i=1}^n \frac{\lambda_i}{\lambda_T} \mu_{L_i}(Y_i) \qquad \square$$

Theorem 4 grounds the idea that properties can be combined via a set of weights to form a concept, as proposed by Gärdenfors in [19], or the operation of convex combination proposed by Zadeh in [50]. The model sits particularly well with Gärdenfors's proposal since it uses a binary conceptual space as the mechanism for combination. Furthermore, the fact that we require specific conditions for the distribution of the threshold of the concept in the binary space is an advantage, since relaxing these conditions allows us to explain some of the characteristics of concept combination seen in psychological experiments, such as overextension or non-commutativity. We will discuss this further in section 4.3.3.

A key aspect of concepts in Gärdenfors's conceptual spaces is that they should be convex. In a space $\mathbb{R}^n$ with the Euclidean distance metric, convexity of a set S is defined by the property that $\forall x, y \in S$, every point on the line segment connecting x and y is also in S. A detailed discussion is given [19] citing experimental evidence for the fact that concepts as used by humans tend to be convex, and that the use of convex concepts requires less cognitive load. We give a definition of convexity for the binary combination spaces $\{0,1\}^n$.

Definition 5 (*Betweenness*). $\forall x, y, z \in \{0,1\}^n$ with distance metric $H_{\vec{\lambda}}$, z is between x and y, $B(x,y,z)$ iff $H_{\vec{\lambda}}(x,y) = H_{\vec{\lambda}}(x,z) + H_{\vec{\lambda}}(z,y)$.

Definition 6 (*Convexity*). A set $S \subseteq \{0,1\}^n$ is convex if $\forall x, y \in S$, every point z lying between x and y also belongs to S, i.e. $\{z : B(x,y,z)\} \subseteq S$.

We can now generalise Theorem 4 to the case where $\theta = \bigwedge_{i=1}^k \pm L_i$, $k \leq n$. In this case, the prototype P does not specify all values of $i = 1\dots n$. We firstly introduce some notation allowing us to talk about the set of dimensions in a prototype that remain invariant. We argue that if P does not specify all dimensions of $\{0,1\}^n$, then the weight vector $\vec{\lambda}$ must be such that only those dimensions contributing to the concept are weighted. For example, suppose $n = 3$, $P = \{(1,1,1),(1,1,0)\}$. Then $\vec{\lambda} = (0.4, 0.3, 0)$. We go on to prove that a similar result to that shown in Theorem 4 holds for θ.

We now introduce some notation to enable us to talk about the dimensions of a set of points $S \subseteq \{0,1\}^n$ that take a fixed value across the subset.

Definition 7 *(Fixed dimensions).*

- For $S \subseteq \{0,1\}^n$, $Z(S) \subseteq \{1, \ldots, n\}$ such that $Z(S) = \{i : \forall \vec{x} \in S, x_i = 0\}$ (the zeros of S)
- For $S \subseteq \{0,1\}^n$, $O(S) \subseteq \{1, \ldots, n\}$ such that $O(S) = \{i : \forall \vec{x} \in S, x_i = 1\}$ (the ones of S)
- For $S \subseteq \{0,1\}^n$, $E(S) \subseteq \{1, \ldots, n\}$ such that $E(S) = \{i : \forall \vec{x}, \vec{y} \in S, x_i = y_i\}$. Note that $E(S) = Z(S) \cup O(S)$ (the set of points in S that take fixed values)

Theorem 8. *Let* $\theta = <P, H_{\vec{\lambda}}, \delta>$ *where P is convex with respect to* $H_{\vec{\lambda}}$. *Furthermore, let* $\theta' = <P, H_{\vec{\lambda}'}, \delta>$ *such that:*

$$\lambda_i' = \begin{cases} \lambda_i : i \in E(P) \\ 0 : otherwise \end{cases}$$

Then $\mu_\theta(\vec{X}) = \mu_{\theta'}(\vec{X})$

Proof. Since P is convex we have that:

$$H_{\vec{\lambda}}(P, \vec{X}) = \sum_{i \in E(P)} \lambda_i = \sum_{i \in E(P)} \lambda_i' = H_{\vec{\lambda}'}(P, \vec{X})$$

Therefore,

$$\mu_\theta(\vec{X}) = \Delta(H_{\vec{\lambda}}(P, \vec{X})) = \Delta(H_{\vec{\lambda}'}(P, \vec{X})) = \mu_{\theta'}(\vec{X}) \qquad \square$$

So with respect to membership in θ in the binary space, the weight vector λ' produces the same results as the weight vector λ. We now show that an analogue to Theorem 4 holds.

Theorem 9. *Let* $\theta = <P, H_{\vec{\lambda}}, \delta>$ *with* $\delta = U(0, \lambda_T)$, *where* $\lambda_T = \sum_{i=1}^n \lambda_i$, *and* $\theta' = <P, H_{\vec{\lambda}'}, \delta'>$ *such that:*

$$\lambda_i' = \begin{cases} \lambda_i : i \in E(P) \\ 0 : otherwise \end{cases}$$

$\lambda_T' = \sum_{i=1}^n \lambda_i'$ *and* $\delta' = U(0, \lambda_T') = \frac{\lambda_T}{\lambda_T'} \delta$. *Then* $\mu_\theta(\vec{X}) = \frac{\lambda_T'}{\lambda_T} \mu_{\theta'}(\vec{X})$, *and* $\mu_\theta(\vec{Y}) = \frac{\lambda_T'}{\lambda_T} \mu_{\theta'}(\vec{Y})$.

Proof. Since P is convex, $H_{\vec{\lambda}}(P, \vec{X}) = H_{\vec{\lambda}'}(P, \vec{X})$.

So

$$\mu_\theta(\vec{X}) = \Delta(H_{\vec{\lambda}}(P, \vec{X})) = \Delta(\frac{\lambda_T'}{\lambda_T} H_{\vec{\lambda}'}(P, \vec{X}))$$

$$= \frac{\lambda_T'}{\lambda_T} \Delta'(H_{\vec{\lambda}'}(P, \vec{X}))$$

$$= \frac{\lambda_T'}{\lambda_T} \mu_{\theta'}(\vec{X})$$

Now,

$$\mu_\theta(\vec{Y}) = \sum_{\vec{X}} \mu_\theta(\vec{X}) \prod_{i=1}^n P(X_i | Y_i)$$

$$= \sum_{\vec{X}} \frac{\lambda_T'}{\lambda_T} \mu_{\theta'}(\vec{X}) \prod_{i=1}^n P(X_i | Y_i)$$

$$= \frac{\lambda_T'}{\lambda_T} \sum_{i=1}^n \frac{\lambda_i'}{\lambda_T'} \mu_{L_i}(Y_i) \quad \text{by Theorem 4}$$

$$= \frac{\lambda_T'}{\lambda_T} \mu_{\theta'}(\vec{Y}) \qquad \square$$

We can therefore convert any $\theta = <P, H_{\vec{\lambda}}, \delta>$ with $\delta = U(0, \lambda_T)$ into θ', such that $i \notin E(P) \to \lambda_i = 0$ via a suitable scaling of $\mu_\theta(\vec{X})$ and $\mu_\theta(\vec{Y})$.

These results show that a compound concept can be built up out of the weighted sum of individual concepts, provided that certain key conditions hold. We now go on to look at the behaviour of the conjunction of two such concepts.
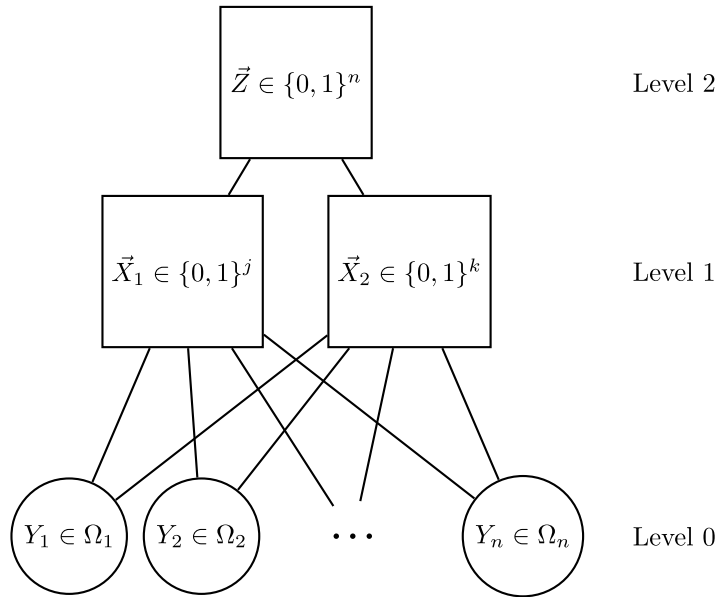
Fig. 8. Schematic illustration of the second level binary space.

4.3.2. Conjunctions of compound concepts

Up to now we have discussed how properties from integral domains may be combined to form concepts. The results in [22] concern how the weighting of these properties change under the conjunction of two such concepts. We extend our framework to take into account the conjunction of two compound concepts. To do this, we introduce a second level binary space, illustrated in Fig. 8, and combine the two concepts within the second level space using the same approach as in section 4.2, i.e. as if they are themselves properties.

Let $\theta = <P_1, d_1, \delta_1>$ and $\varphi = <P_2, d_2, \delta_2>$ be two compound concepts consisting of a conjunction of attribute labels from individual conceptual spaces Ω_i. As in section 4.2 we define the following two binary variables:

$$Z_1 = \begin{cases} 1 : d(\vec{X}, P_1) \leq \varepsilon_1 \\ 0 : \text{otherwise} \end{cases} \quad \text{and } Z_2 = \begin{cases} 1 : d(\vec{X}, P_2) \leq \varepsilon_2 \\ 0 : \text{otherwise} \end{cases}$$

Hence $\vec{X}$ naturally generates $\vec{Z} = (Z_1, Z_2) \in \{0, 1\}^2$. We now define the *conjunction of compound concepts* as the triple $\theta \wedge \varphi = (\{(1, 1)\}, d, \delta)$ where d is a distance metric on $\{0, 1\}^2$. In this case we can define a binary random variable C such that:

$$C = \begin{cases} 1 : d((1, 1), \vec{Z}) \leq \varepsilon \\ 0 : \text{otherwise} \end{cases}$$

We also define the membership function for $\theta \wedge \varphi$ as follows:

$$\mu_{\theta \wedge \varphi}(\vec{Z}) = \delta(\{\epsilon : d((1, 1), \vec{Z}) \leq \varepsilon\})$$

So that $\mu_{\theta \wedge \varphi}(\vec{Z}) = P(C = 1|\vec{Z})$. Now by applying the theorem of total probability we have that:

$$\mu_{\theta \wedge \varphi}(\vec{X}) = P(C|\vec{X}) = \sum_{\vec{Z}} \mu_{\theta \wedge \varphi}(\vec{Z}) \times P(\vec{Z}|\vec{X}) = \sum_{\vec{Z}} \mu_{\theta \wedge \varphi}(\vec{Z}) P(Z_1|\vec{X}) P(Z_2|\vec{X})$$

Then by a second application of the theorem of total probability we have that:

$$\mu_{\theta \wedge \varphi}(\vec{Y}) = P(C|\vec{Y}) = \sum_{\vec{X}} \mu_{\theta \wedge \varphi}(\vec{X}) P(\vec{X}|\vec{Y}) = \sum_{\vec{X}} P(\vec{X}|\vec{Y}) \sum_{\vec{Z}} \mu_{\theta \wedge \varphi}(\vec{Z}) P(Z_1|\vec{X}) P(Z_2|\vec{X})$$

$$= \sum_{\vec{X}} \prod_{i=1}^{n} P(X_i|Y_i) \sum_{\vec{Z}} \mu_{\theta \wedge \varphi}(\vec{Z}) P(Z_1|\vec{X}) P(Z_2|\vec{X})$$

We can now look at the behaviour of the weights in a conjunction of compound concepts.

Theorem 10. *Let* $\theta = <P_\theta, H_{\vec{\lambda}_\theta}, \delta_\theta>$, $\varphi = <P_\varphi, H_{\vec{\lambda}_\varphi}, \delta_\varphi>$, *where* $\sum_{i=1}^{n} \lambda_{\theta,i} = \lambda_{\theta,T}$, $\sum_{i=1}^{n} \lambda_{\varphi,i} = \lambda_{\varphi,T}$ *and* $\delta_\theta = Uniform(0, \lambda_{\theta,T})$, $\delta_\varphi = U(0, \lambda_{\varphi,T})$. *Let* $\theta \wedge \varphi = <\{(1,1)\}, H_{\vec{w}}, \delta>$ *where* $w_1 + w_2 = w_T$ *and* $\delta = U(0, w_T)$. *Suppose that if* $i \in E(P_\theta) \cap E(P_\varphi)$ *then* $P_{\theta,i} = P_{\varphi,i}$, *i.e. that the prototypes have no directly contradictory attributes. Then*

$$\mu_{\theta \wedge \varphi}(\vec{Y}) = \sum_{i=1}^{n} \left(\frac{w_1 \lambda_{\varphi,T} \lambda_{\theta,i} + w_2 \lambda_{\theta,T} \lambda_{\varphi,i}}{w_T \lambda_{\theta,T} \lambda_{\varphi,T}} \right) \mu_{\pm L_i}(\vec{Y})$$

So $\theta \wedge \varphi = <P_{\theta \wedge \varphi}, H_{\vec{\lambda}_{\theta \wedge \varphi}}, \delta_{\theta \wedge \varphi}>$ *where* $P_{\theta \wedge \varphi} = P_\theta \cap P_\varphi$, $\lambda_{\theta \wedge \varphi_i} = w_1 \lambda_{\varphi,T} \lambda_{\theta,i} + w_2 \lambda_{\theta,T} \lambda_{\varphi,i}$ *and* $\delta_{\theta \wedge \varphi} = U(0, w_T \lambda_{\theta,T} \lambda_{\varphi,T})$.

Proof. Suppose w.l.o.g. by Theorem 9 that $i \notin E(P_\theta) \implies \lambda_{\theta,i} = 0$, $i \notin E(P_\varphi) \implies \lambda_{\varphi,i} = 0$

$$\mu_{\theta \wedge \varphi}(\vec{Y}) = \sum_{\vec{X}} \prod_{i=1}^{n} P(X_i | Y_i) \sum_{\vec{Z}} \mu_{\theta \wedge \varphi}(\vec{Z}) \prod_{i=1}^{2} P(Z_i | \vec{X})$$

$$= \sum_{\vec{X}} \prod_{i=1}^{n} P(X_i | Y_i) \left(\frac{w_1}{w_T} \mu_\theta(\vec{X}) + \frac{w_2}{w_T} \mu_\varphi(\vec{X}) \right)$$

$$= \frac{w_1}{w_T} \sum_{\vec{X}} \prod_{i=1}^{n} \mu_\theta(\vec{X}) P(X_i | Y_i) + \frac{w_2}{w_T} \sum_{\vec{X}} \prod_{i=1}^{n} \mu_\varphi(\vec{X}) P(X_i | Y_i)$$

$$= \frac{w_1}{w_T} \sum_{i \in O(P_\theta)} \frac{\lambda_{\theta,i}}{\lambda_{\theta,T}} \mu_{L_i}(Y_i) + \frac{w_1}{w_T} \sum_{i \in Z(P_\theta)} \frac{\lambda_{\theta,i}}{\lambda_{\theta,T}} \mu_{\neg L_i}(Y_i)$$

$$+ \frac{w_2}{w_T} \sum_{i \in O(P_\varphi)} \frac{\lambda_{\varphi,i}}{\lambda_{\varphi,T}} \mu_{L_i}(Y_i) + \frac{w_2}{w_T} \sum_{i \in Z(P_\varphi)} \frac{\lambda_{\varphi,i}}{\lambda_{\varphi,T}} \mu_{\neg L_i}(Y_i)$$

Now, since if $i \in E(P_\theta) \cap E(P_\varphi)$ then $P_{\theta,i} = P_{\varphi,i}$ and $i \notin E(P_\theta) \implies \lambda_{\theta,i} = 0$, $i \notin E(P_\varphi) \implies \lambda_{\varphi,i} = 0$, we have that

$$\mu_{\theta \wedge \varphi}(\vec{Y}) = \sum_{i=1}^{n} \left(\frac{w_1 \lambda_{\varphi,T} \lambda_{\theta,i} + w_2 \lambda_{\theta,T} \lambda_{\varphi,i}}{w_T \lambda_{\theta,T} \lambda_{\varphi,T}} \right) \mu_{\pm L_i}(\vec{Y}) \qquad \square$$

So under certain conditions, the attribute weights in the conjunctive concept are a weighted sum of the attribute weights of the constituent concepts, which models one of Hampton's key findings as stated in section 4.1. Further, since $P_{\theta \wedge \varphi} = P_\theta \cap P_\varphi$, $E(P_{\theta \wedge \varphi}) = E(P_\theta) \cup E(P_\varphi)$, i.e. the attribute set for the conjunctive concept is a union of the attribute sets for the constituent concepts, modelling another aspect of Hampton's results.

Example 11 *(Concept–concept combination).* Suppose one concept, $SUCCESS$, is defined as a conjunction of labels $L_1 =$ 'rich', $L_2 =$ 'healthy', and suppose $HAPPY$ is defined as a conjunction of labels $L_2 =$ 'healthy' and $L_3 =$ 'busy'. Furthermore, suppose $SUCCESS$ is formed in the binary space $\{0,1\}^2$ with prototype set $P_s = \{(1,1)\}$, weighting $\vec{\lambda}_s = (0.75, 0.25)$ and $\delta_s = U(0,1)$, i.e. $SUCCESS = <\{(1,1)\}, H_{(0.75,0.25)}, U(0,1)>$ and similarly suppose $HAPPY = <\{(1,1)\}, H_{(0.5,0.5)}, U(0,1)>$. To combine these two concepts, we firstly expand the prototypes and weight vectors for each to encompass the dimensions of the other, i.e. they are embedded in the space $\{0,1\}^3$, where each dimension i indicates the presence or absence of the label L_i. So, for example, the point $(1,0,1)$ corresponds to 'rich and not healthy and busy'. Within this expanded space, $P_s = \{(1,1,x_3) : x_3 \in \{0,1\}\}$, i.e., $\{(1,1,0),(1,1,1)\}$ and $\vec{\lambda}_s = (0.75, 0.25, 0)$. Similarly, $P_h = \{(x_1,1,1) : x_1 \in \{0,1\}\}$, and $\vec{\lambda}_h = (0, 0.5, 0.5)$. Now, according to Theorems 4 and 9, $\mu_S(\vec{Y}) = 0.75 \mu_{L_1}(Y_1) + 0.25 \mu_{L_2}(Y_2)$, and $\mu_H(\vec{Y}) = 0.5 \mu_{L_2}(Y_2) + 0.5 \mu_{L_3}(Y_3)$.

We combine the expanded $\theta = SUCCESS \wedge HAPPY$ in the second level binary space $\{0,1\}^2$ with prototype $P_\theta = (1,1)$, $\vec{\lambda}_\theta = (0.5, 0.5)$, $\delta_\theta = U(0,1)$. Now, according to Theorem 10

$$\mu_\theta(\vec{Y}) = 0.5 \mu_{SUCCESS}(\vec{Y}) + 0.5 \mu_{HAPPY}(\vec{Y})$$

$$= 0.5(0.75 \mu_{L_1}(Y_1) + 0.25 \mu_{L_2}(Y_2)) + 0.5(0.5 \mu_{L_2}(Y_2) + 0.5 \mu_{L_3}(Y_3))$$

$$= 0.375 \mu_{L_1}(Y_1) + 0.375 \mu_{L_2}(Y_2) + 0.25 \mu_{L_3}(Y_3)$$

In many cases, we want to combine a concept with a property, such as 'red car', illustrated in Fig. 9. This can be implemented as a special instance of the concept–concept combination illustrated in Example 11, where the first concept has just one label.
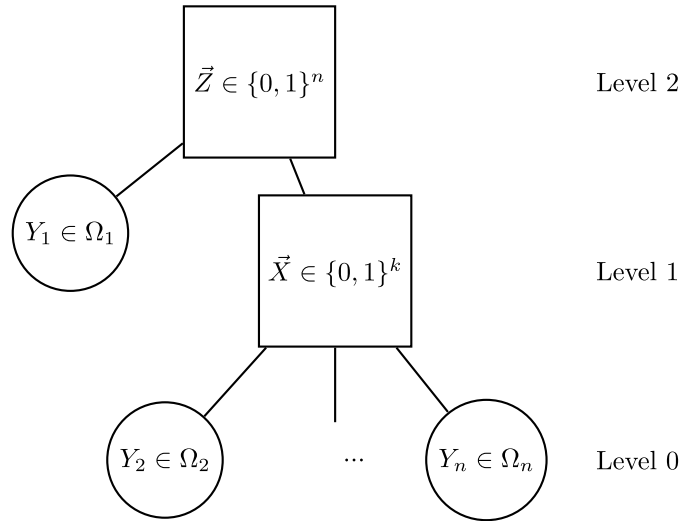
Fig. 9. Schematic illustration of property–concept combination.

Example 12 *(Property–concept combination).* Suppose $RED = L_1$, $CAR = L_2 \wedge L_3 \wedge L_4$, where $CAR = <(1,1), H_{\vec{\lambda}_C}, U(0,1)>$ with $\vec{\lambda}_C = (0.33, 0.33, 0.33)$ so that $\mu_C(\vec{Y}) = 0.33\mu_{L_2}(Y_2) + 0.33\mu_{L_3}(Y_3) + 0.33\mu_{L_4}(Y_4)$. We may combine $\theta = RED \wedge CAR$ in the space $\{0,1\}^2$ with prototype $(1,1)$, weight vector $\lambda = (0.5, 0.5)$ and boundary distribution $\delta = U(0,1)$. Then

$$\mu_\theta(\vec{Y}) = 0.5\mu_{RED}(\vec{Y}) + 0.5\mu_{CAR}(\vec{Y})$$
$$= 0.5\mu_{L_1}(Y_1) + 0.5(0.33\mu_{L_2}(Y_2)) + 0.33\mu_{L_3}(Y_3) + 0.33\mu_{L_3}(Y_4))$$
$$= 0.5\mu_{L_1}(Y_1) + 0.167\mu_{L_2}(Y_2)) + 0.167\mu_{L_3}(Y_3) + 0.167\mu_{L_3}(Y_4)$$

Two objections might be made to this approach. Firstly, the combination could be of the form 'apple green', where the resulting concept should be a colour, rather than an apple. The combination mechanism given in the example would return a green apple, rather than a colour. Secondly, as described here, this approach does not take into account cases like 'red wine', where the meaning of 'red' has been changed by the concept it is attached to. To answer the first objection, note that the resulting concept from a combination like 'apple green', or 'green apple' is indicated by the part of speech that each word belongs to. In English, this is indicated by word order, so we know that in the first case the resulting concept should be a colour, and in the second case the resulting concept should be a fruit. In the combination of properties and concepts, the resulting concept will be the original concept modified in the specific domain to which the property applies. So, in the case of 'green apple', the resulting concept is an apple in which the colour domain has been modified. In the case of 'apple green', we would choose just the domains from 'apple' that are relevant to the domains of 'green', i.e. colour, and form the combination 'apple green' using the approach outlined in Example 12. This approach is similar to that outlined by Gärdenfors [19].

In the second objection, the property 'red' in the combination 'red wine' refers to a particular set of shades of red which are not at all prototypical. Gärdenfors [19] has again addressed this, using the idea of contrast classes, which map the whole domain onto a subset of the domain determined by the that class. So in the case of wine, there is a range of distinctive shades of wine. When the whole colour domain is mapped down to this range, the places that the 'red' and the 'white' labels inhabited in the original labelling should now map onto the 'red' and the 'white' areas in the range of wine colours. Our approach is slightly different. As explained in the first objection, the concept 'wine red' can be obtained by forming a combination of 'wine coloured' and 'red'. The 'red' in 'red wine' is then understood to be an instance of 'wine red', rather than every day 'red'. Although this argument might appear to be circular, we argue that it is not, since the first time someone encounters the concept 'red wine', a mistake could be made about what colour the drink would be. Only after learning that the label 'red' in 'red wine' refers to the darkest colour of wine can they use it properly. This can be seen as an instance where the meaning of the term red, when applied to wine, is determined by convention rather than a systematic combination.

Furthermore, although when introducing this framework we have made a distinction between properties and concepts, this distinction is not really important in actually carrying out a combination. Increasingly complex concepts can be created and combined with other complex concepts or alternatively with simple properties utilising a single domain. The novelty of this approach is that the combination mechanism is itself characterised by a conceptual space. As a special case our framework entails that concepts may be characterised as weighted sums of properties, a characterisation of concepts proposed in [19,51]. Hampton shows that a majority of his data may be explained by a simple multilinear regression, which can be

modelled as in the example above. However, he also notes other non-compositional behaviours, which are key aspects of how humans use concepts. We show how our framework can model some of these non-compositional behaviours in the next section.

4.3.3. Non-compositional behaviours

In addition to the general rule that the importance of attributes in the conjunction is the weighted sum of the importance of attributes in the constituent concepts, Hampton identifies four additional behaviours: necessity and impossibility are preserved; attribute loss or emergence is observed; conjunction is not commutative; and dominance effects are observed. This section will discuss the capability of our model to capture these behaviours.

We consider firstly the ideas of necessity and impossibility. Hampton finds that necessity of dimensions is preserved, so that if an attribute is deemed necessary in a constituent concept, it is also deemed necessary in the conjunction. As outlined in section 4.2, necessity and impossibility are essentially the highest and lowest weights that can be assigned to an attribute in Hampton's experiments. Recall that we view the impossibility of an attribute L_i as equivalent to the necessity of $\neg L_i$, and we measure the necessity of an attribute or its negation using the notion of necessity from possibility theory as outlined in Definition 13.

We also introduce an alternative definition of the importance of an attribute, consistent with our random set based conceptual models. Rather than simply consider the weight $\vec{\lambda}$ to be the importance of the attribute, we use the idea of necessity from possibility theory [16]. Within possibility theory, the possibility $\pi(s)$ of a state of affairs s indicates to what extent this state of affairs is possible. $\pi(s)$ is a measure on the interval $[0, 1]$. The possibility of a set of states A is then defined in [16] to be:

$$\Pi(A) = \sup_{s \in A} \pi(s)$$

The necessity of an event A is then $N(A) = 1 - \Pi(A^c)$, i.e. $1 -$ the possibility that A does *not* occur.

Within our model, a state s is considered to be a particular point $\vec{X}$ in the binary combination space $\{0, 1\}^n$ and $\pi(\vec{X}) := \mu_\theta(\vec{X})$. To compute the necessity of a dimension i to a concept $\theta = <P, H_{\vec{\lambda}}, \delta>$, where $P \subseteq \{0, 1\}^n$, we consider the necessity of set $S_i = \{\vec{X} : X_i = p_i\}$, i.e. the set of all points that have equal value to the prototype on dimension i.

$$N(S_i) = 1 - \Pi(S_i^c) = 1 - \sup_{\vec{X} \in S_i^c} \mu_\theta(\vec{X}) = \inf_{\vec{X} \in S_i^c} (1 - \mu_\theta(\vec{X}))$$

$$= \inf_{\vec{X} \in S_i^c} P(d(x, P) > \varepsilon) = P(\mathcal{N}_\theta^\varepsilon \subseteq S_i) = \delta(\{\varepsilon : \mathcal{N}_\theta^\varepsilon \subseteq S_i\})$$

where $\mathcal{N}_\theta^\varepsilon$ is the neighbourhood of θ as defined in section 3. It may be the case that P contains both points with value 0 on dimension i and points with value 1 on dimension i. In this case we say that $S_i = \emptyset$, since S_i must satisfy $S_i = \{\vec{X} : X_i = 0 \wedge X_i = 1\}$.

Definition 13 (*Necessity of a dimension*). Given $\theta = <P, H_{\vec{\lambda}}, \delta>$, the necessity of a dimension i to θ is defined as the necessity of the set $S_i = \{\vec{X} : X_i = p_i\}$, where p_i is the ith dimension of $\{0, 1\}^n$, by $N(S_i) = \delta(\{\varepsilon : \mathcal{N}_\theta^\varepsilon \subseteq S_i\})$.

If P contains both vectors with value 0 on dimension i and vectors with value 1 on dimension i, then $N(S_i) = 0$, since $S_i = \emptyset$ and therefore $N(S_i) = \delta(\{\varepsilon : \mathcal{N}_\theta^\varepsilon \subseteq S_i\}) = P(\mathcal{N}_\theta^\varepsilon \subseteq \emptyset) = 0$.

Theorem 14. *For* $\alpha = <P, H_{\vec{\lambda}}, \delta>$, $N(S_i) = 1 - \Delta(\lambda_i)$.

Proof. $N(S_i) = P(\mathcal{N}_\alpha^\varepsilon \subseteq S_i) = P(\varepsilon < \lambda_i) = 1 - \Delta(\lambda_i)$. $\square$

In particular, when $\varepsilon \sim U(0, \lambda_T)$, $N(S_i) = \frac{\lambda_i}{\lambda_T}$.

We have introduced the concept of the necessity of a dimension in order to account for some of the non-compositional aspects of conjunctive combinations of concepts. To examine the necessity of a dimension in a conjunction of two compound concepts, we relate the distribution of the threshold ε in the higher level binary space to the neighbourhood in the first level binary space.

We begin by defining the neighbourhood of a conjunction of two concepts. Suppose that two concepts θ and φ are combined in the second level space $\{0, 1\}^2$, with weight vector $\vec{w} = (w_1, w_2)$ where w_1 is associated with θ and w_2 with φ. Suppose $w_2 \leq w_1$. Now, if the threshold ε in the second level space is less than w_2, then an element $\vec{X}$ of the first level binary space must belong to the neighbourhoods of both concepts θ and φ, i.e. $\vec{X} \in \mathcal{N}_\theta^{\varepsilon_\theta} \cap \mathcal{N}_\varphi^{\varepsilon_\varphi}$. When $w_2 \leq \varepsilon \leq w_1$, the points within ε in the second level space are $\{(1, 1), (1, 0)\}$, so $\vec{X}$ must belong to $\mathcal{N}_\theta^{\varepsilon_\theta}$, but does not have to belong to $\mathcal{N}_\varphi^{\varepsilon_\varphi}$. When $w_1 \leq \varepsilon \leq w_T$, where $w_T = w_1 + w_2$, $\vec{X}$ may belong to either neighbourhood, i.e. $\vec{X} \in \mathcal{N}_\theta^{\varepsilon_\theta} \cup \mathcal{N}_\varphi^{\varepsilon_\varphi}$. This is summarised in the following definition:

Definition 15. For a conjunction $\theta \wedge \varphi = <\{(1,1)\}, H_{\vec{w}}, \delta>$, where $\vec{w} = (w_1, w_2)$ and $w_2 \leq w_1$, the neighbourhood R of $\theta \wedge \varphi$ is:

$$R = \begin{cases} \mathcal{N}_{\theta}^{\varepsilon_{\theta}} \cap \mathcal{N}_{\varphi}^{\varepsilon_{\varphi}} : \varepsilon \leq w_2 \\ \mathcal{N}_{\theta}^{\varepsilon_{\theta}} : w_2 \leq \varepsilon \leq w_1 \\ \mathcal{N}_{\theta}^{\varepsilon_{\theta}} \cup \mathcal{N}_{\varphi}^{\varepsilon_{\varphi}} : w_1 \leq \varepsilon \leq w_T \end{cases}$$

This allows us to relate the necessity of a dimension to the concept $\theta \wedge \varphi$ to the necessity of the dimension to the constituent concepts θ, φ since $N_{\theta \wedge \varphi}(S_i) = \delta_{\theta \wedge \varphi}(\varepsilon_{\theta}, \varepsilon_{\varphi} : R \subset S_i)$.

Theorem 16. *For a conjunction of compound concepts* $\theta \wedge \varphi = <\{(1,1)\}, H_{\vec{w}}, \delta>$, *where* θ, φ *are defined as for* Theorem 10

$$N_{\theta \wedge \varphi}(S_i) = (1 - \Delta(w_1))N_{\theta}(S_i) + (1 - \Delta(w_2))N_{\varphi}(S_i) - (1 - \Delta(w_1) - \Delta(w_2) + \Delta(w_T))N_{\theta}(S_i)N_{\varphi}(S_i)$$

Proof.

$$\begin{aligned} N_{\theta \wedge \varphi}(S_i) &= \delta_{\theta} \times \delta_{\varphi}(\{\varepsilon_{\theta}, \varepsilon_{\varphi} : R^{\varepsilon, \varepsilon_{\theta}, \varepsilon_{\varphi}} \subset S_i\} \\ &= (1 - \Delta(w_2))\delta_{\theta} \times \delta_{\varphi}(\{\varepsilon_{\theta}, \varepsilon_{\varphi} : \mathcal{N}_{\theta}^{\varepsilon_{\theta}} \cap \mathcal{N}_{\varphi}^{\varepsilon_{\varphi}} \subset S_i\}) \\ &\quad + (\Delta(w_2) - \Delta(w_1))\delta_{\theta} \times \delta_{\varphi}(\{\varepsilon_{\theta}, \varepsilon_{\varphi} : \mathcal{N}_{\theta}^{\varepsilon_{\theta}} \subset S_i\}) \\ &\quad + (\Delta(w_1) - \Delta(w_T))\delta_{\theta} \times \delta_{\varphi}(\{\varepsilon_{\theta}, \varepsilon_{\varphi} : \mathcal{N}_{\theta}^{\varepsilon_{\theta}} \cup \mathcal{N}_{\varphi}^{\varepsilon_{\varphi}} \subset S_i\}) \\ &= (1 - \Delta(w_2))(N_{\theta}(S_i) + N_{\varphi}(S_i) + N_{\theta}(S_i)N_{\varphi}(S_i)) \\ &\quad + (\Delta(w_2) - \Delta(w_1))N_{\theta}(S_i) + (\Delta(w_1) - \Delta(w_T))N_{\theta}(S_i)N_{\varphi}(S_i) \\ &= (1 - \Delta(w_1))N_{\theta}(S_i) + (1 - \Delta(w_2))N_{\varphi}(S_i) - (1 - \Delta(w_1) - \Delta(w_2) + \Delta(w_T))N_{\theta}(S_i)N_{\varphi}(S_i) \quad \square \end{aligned}$$

Corollary 17. *If* $\delta = \text{Uniform}(0, w_T)$ *then* $N_{\theta \wedge \varphi}(S_i) = \frac{w_1}{w_T}N_{\theta}(S_i) + \frac{w_2}{w_T}N_{\varphi}(S_i)$.

Proof. $\delta = Uniform(0, w_T) \implies \Delta(w_1) = \frac{w_2}{w_T}$ and $\Delta(w_1) = \frac{w_1}{w_T}$

$$\begin{aligned} N_{\theta \wedge \varphi}(S_i) &= (1 - \Delta(w_1))N_{\theta}(S_i) + (1 - \Delta(w_2))N_{\varphi}(S_i) - (1 - \Delta(w_1) - \Delta(w_2) + \Delta(w_T))N_{\theta}(S_i)N_{\varphi}(S_i) \\ &= (1 - \frac{w_2}{w_T})N_{\theta}(S_i) + (1 - \frac{w_1}{w_T})N_{\varphi}(S_i) - (1 - \frac{w_2}{w_T} - \frac{w_1}{w_T} + 0)N_{\theta}(S_i)N_{\varphi}(S_i) \\ &= \frac{w_1}{w_T}N_{\theta}(S_i) + \frac{w_2}{w_T}N_{\varphi}(S_i) \quad \square \end{aligned}$$

This allows us to choose a boundary distribution δ which gives us the property that high necessity is carried through into the conjunction. This is illustrated in the following example.

Example 18. Suppose that $N_{\theta}(S_i) = 0.9$, $N_{\varphi}(S_i) = 0.6$, $w_1 = 0.2$, $w_2 = 0.8$. The weighted sum of the necessity of attribute i is then $0.2 \times 0.9 + 0.8 \times 0.6 = 0.66$. Now, if $\delta = \text{Uniform}(0, w_T)$ then $N_{\theta \wedge \varphi}(S_i)$ is equal to the weighted sum as shown in Corollary 17 and as reported by Hampton. However, if ε is distributed over a narrower range than the whole binary space, for example $\varepsilon \sim \text{Uniform}(0, 0.5)$ then:

$$N_{\theta \wedge \varphi}(S_i) = (1 - 0.6) \times 0.9 + (1 - 0) \times 0.6 - (1 - 0.6 - 0 + 0) \times 0.9 \times 0.6 = 0.96 - 0.144 = 0.744$$

which is closer to 0.9, even though less weight is given to that part of the conjunction. The necessity of this attribute has therefore been, if not entirely preserved, at least emphasised in this model.

The fourth aspect Hampton notes is that attribute loss or emergence is observed. When the distribution of ε is defined over a narrower range than $[0, w_T]$, the importance of attributes in the combined concept is higher than the importance of attributes in either constituent concept, as seen in Fig. 10. If the distribution of ε is over a wider range than the whole space, the importance can be lower. However, this does not account for how both phenomena can be observed together. In the example discussed and in Fig. 10, ε has a uniform distribution. Further work to investigate the behaviour of $N_{\theta \wedge \varphi}(S_i)$ is needed.

A further aspect uncovered in Hampton's results is that in general, the qualifying noun, i.e. the second concept in the conjunction, is given more weight than the first. Within our model, we can easily take account of this by setting weights in the binary combination space appropriately.

Lastly, Hampton finds that concepts with more attributes tend to have higher weightings. This can also be accounted for by weighting concepts θ relative to the cardinality of $E(P_{\theta})$.
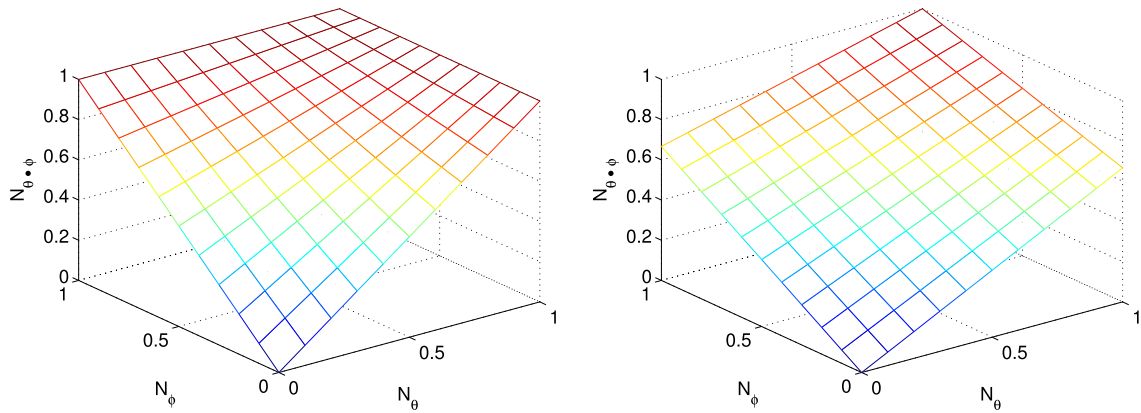
(a) $N_{\theta \wedge \varphi}(S_i)$ as a function of $N_\theta(S_i)$ and $N_\varphi(S_i)$ when $w_1 = w_2 = 0.5$, $\varepsilon \sim U(0, 0.5)$. Notice that when either $N_\theta(S_i) = 1$ or $N_\varphi(S_i) = 1$, so is $N_{\theta \wedge \varphi}(S_i)$.

(b) $N_{\theta \wedge \varphi}(S_i)$ as a function of $N_\theta(S_i)$ and $N_\varphi(S_i)$ when $w_1 = w_2 = 0.5$, $\varepsilon \sim U(0, 0.75)$. Notice that high necessity is preserved to a certain extent, but not as high as that in Fig. 10.

Fig. 10. Surfaces capturing the value of $N_{\theta \wedge \varphi}(S_i)$ as a function of $N_\theta(S_i)$ and $N_\varphi(S_i)$. It is possible for values of $N_{\theta \wedge \varphi}(S_i)$ to be higher than either $N_\theta(S_i)$ or $N_\varphi(S_i)$, modelling the phenomenon of attribute emergence.

4.4. Other aspects of human concept use

We give here some examples of how other cases of non-compositionality can be accounted for within our framework. A key example is the conjunction fallacy, in which the probability of an entity belonging to a conjunction of concepts is judged greater than the probability of that entity belonging to just one of the concepts. The example often cited is that of Linda the feminist bank teller, introduced by Tversky and Kahneman [47]. Linda is characterised as follows:

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

When asked to rank the probability of the statements 1) 'Linda is a bank teller' and 2) 'Linda is a bank teller and is active in the feminist movement', the majority of people rank 2) as more probable than 1), in violation of classical probability. This fallacy has been shown to dissolve when posed in frequentist terms [12]. The frequentist terms are given in the following way. Rather than being asked to rank the probability of Linda being a bank teller, participants are asked to consider 100 people who can be described as above, and decide how many of that number are bank tellers, and then how many of the 100 are feminist bank tellers. When described in this way, participants no longer give contradictory answers. However, the fallacy remains when asked for probability judgements. Within our framework we can consider the membership in a concept as the probability that a person would assert that concept to describe an object. As we see in the example below, our approach to membership in a conjunctive concept does not entail that membership in the conjunction is always less than or equal to membership in each of the constituent concepts.

Example 19. Suppose T = 'bank teller' is defined as a conjunction of L_1 = 'good with numbers', L_2 = 'medium intelligence', with $P_T = (1, 1)$, $\vec{\lambda}_T = (0.5, 0.5)$, $\varepsilon_T \sim U(0, 1)$ and F = 'feminist' is defined as a conjunction of L_3 = 'outspoken', L_4 = 'concerned with issues of discrimination and social justice' $P_F = (1, 1)$, $\vec{\lambda}_F = (0.5, 0.5)$, $\varepsilon_F \sim U(0, 1)$. Suppose we combine T and F in the space $\{0, 1\}^2$ with prototype $P = (1, 1)$, weight vector $\vec{\lambda} = (0.5, 0.5)$, threshold $\varepsilon \sim U(0, 1)$. Say $\mu_{L_1}(Linda) = 0.5$ $\mu_{L_2}(Linda) = 0.25$, $\mu_{L_3}(Linda) = 1$, $\mu_{L_4}(Linda) = 1$ Then:

$$\mu_T(Linda) = 0.5\mu_{L_1}(Linda) + 0.5\mu_{L_2}(Linda) = 0.375$$

$$\mu_{TF}(Linda) = 0.5\mu_T(Linda) + 0.5\mu_F(Linda)$$

$$= 0.5(0.5\mu_{L_1}(Linda) + 0.5\mu_{L_2}(Linda)) + 0.5(0.5\mu_{L_3}(Linda) + 0.5\mu_{L_4}(Linda))$$

$$= 0.6875$$

We argue here that rather than committing a fallacy, the participants are generating a new concept 'feminist bank teller' with a new prototype. So here the additional characteristics associated with being a feminist increase the membership of Linda in the conjunctive concept.

Another of the classic counterexamples relates to the fact that a goldfish, or a guppy, is a better example of a pet fish than it is of either a pet or of a fish, as introduced by Osherson and Smith [40]. The following example shows how we begin to address this.

Example 20. Suppose concept PET = 'pet' is defined by the conjunction of properties L_1 = 'lives in house', L_2 = 'furry', and $FISH$ = 'fish' is defined by the conjunction of properties L_3 = 'lives in water', L_4 = 'scaly'. We firstly form two binary combination spaces $\Omega_{PET} = \{0, 1\}^2$ for pet and $\Omega_{FISH} = \{0, 1\}^2$ for fish, and give each a weight vector, with say $\vec{\lambda}_{PET} = [0.6, 0.4]$ and $\vec{\lambda}_{FISH} = [0.5, 0.5]$.

In each case the threshold in the binary space is distributed according to $\varepsilon = U(0, 1)$ so that $PET = <[1, 1], H_{[0.6, 0.4]}, U(0, 1)>$ and $\mu_{PET}(\vec{Y}) = 0.6\mu_{L_1}(Y_1) + 0.4\mu_{L_2}(Y_2)$. Also $FISH = <[1, 1], H_{[0.5, 0.5]}, U(0, 1)>$ and $\mu_{FISH}(\vec{Y}) = 0.5\mu_{L_3}(Y_3) + 0.5\mu_{L_4}(Y_4)$. We then expand each of the spaces Ω_{PET} and Ω_{FISH} to $\{0, 1\}^4$ and expand the prototypes accordingly, setting the additional dimension weights to zero as described in Theorem 8. Suppose a goldfish (g) has membership values in the individual labels as follows: $\mu_{L_1}(g) = 1$, $\mu_{L_2}(g) = 0$, $\mu_{L_3}(g) = 1$, $\mu_{L_4}(g) = 0.25$. This gives:

$$\mu_{PET}(g) = 0.6 \times 1 + 0.4 \times 0 = 0.6$$

$$\mu_{FISH}(g) = 0.5 \times 1 + 0.5 \times 0.25 = 0.625$$

Suppose the two concepts PET and $FISH$ are combined in a higher level binary combination space with prototype $P = (1, 1)$, dimension weights $\vec{\lambda} = [0.5, 0.5]$ and $\varepsilon \sim U(0, 1.25)$. Now, from section 4.3.2 we have

$$\mu_{PET \wedge FISH}(g) = \sum_{\vec{X}} \prod_{i=1}^{n} P(X_i|g_i) \sum_{\vec{Z}} \mu_{PET \wedge FISH}(\vec{Z}) \prod_{i=1}^{2} P(Z_i|\vec{X})$$

where $\vec{X} \in \{0, 1\}^4$, i.e. the combination space at Level 1 in Fig. 8, $\vec{Z} \in \{0, 1\}^2$, i.e. the combination space at Level 2 in Fig. 8, and g_i refers to the ith dimension of 'goldfish', $i = 1, ..., 4$. Recall that $P(X_i|g_i) = \mu_{L_i}(g)$ if $X_i = 1$ and $P(X_i|g_i) = 1 - \mu_{L_i}(g)$ if $X_i = 0$. Looking at the values of $\mu_{L_i}(g)$, we see that we only need to sum over two values of $\vec{X}$, namely $\vec{X}_0 = (1, 0, 1, 0)$ and $\vec{X}_1 = (1, 0, 1, 1)$ (since in all other cases $\prod_{i=1}^{n} P(X_i|g_i) = 0$). We consider each of these two vectors in turn, firstly $\vec{X} = \vec{X}_0$. The product over individual labels, $\prod_{i=1}^{n} P(X_{0,i}|g_i) = \prod_{j:X_{0,j}=1} \mu_{L_j}(g) \prod_{k:X_{0,k}=0}(1 - \mu_{L_k}(g)) = 0.75$. Also,

$$\mu_{PET}(\vec{X}_0) = P(H_{\lambda_{PET}}(P_{PET}, \vec{X}_0) \le \varepsilon_{PET}) = 1 - 0.4 = 0.6$$

$$\mu_{FISH}(\vec{X}_0) = P(H_{\lambda_{FISH}}(P_{FISH}, \vec{X}_0) \le \varepsilon_{FISH}) = 1 - 0.5 = 0.5$$

Now consider the sum over $\vec{Z}$ in the Level 2 space:

$$\sum_{\vec{Z}} \mu_{PET \wedge FISH}(\vec{Z}) \prod_{i=1}^{2} P(Z_i|\vec{X}_0) = \mu_{PET \wedge FISH}((1, 1))\mu_{PET}(\vec{X}_0)\mu_{FISH}(\vec{X}_0)$$

$$+ \mu_{PET \wedge FISH}((1, 0))\mu_{PET}(\vec{X}_0)(1 - \mu_{FISH}(\vec{X}_0))$$

$$+ \mu_{PET \wedge FISH}((0, 1))(1 - \mu_{PET}(\vec{X}_0))\mu_{FISH}(\vec{X}_0)$$

$$+ \mu_{PET \wedge FISH}((0, 0))(1 - \mu_{PET}(\vec{X}_0))(1 - \mu_{FISH}(\vec{X}_0))$$

$$= 1 \times 0.6 \times 0.5 + 0.6 \times 0.6 \times 0.5 + 0.6 \times 0.4 \times 0.5 + 0.2 \times 0.4 \times 0.5$$

$$= 0.64$$

Similarly, $\prod_{i=1}^{n} P(X_{1,i}|g_i) = 0.25$ and $\sum_{\vec{Z}} \mu_{PET \wedge FISH}(\vec{Z}) \prod_{i=1}^{2} P(Z_i|\vec{X}_1) = 0.84$, giving

$$\mu_{PET \wedge FISH}(g) = 0.75 \times 0.64 + 0.25 \times 0.84 = 0.69 > \min(\mu_{PET}(g), \mu_{FISH}(g))$$

Also, consider a cat, for whom $\mu_{L_1}(cat) = 1$, $\mu_{L_2}(cat) = 1$, $\mu_{L_3}(cat) = 0$, $\mu_{L_4}(cat) = 0$. We want a cat to be a worse example of a pet fish than a goldfish is. Then $\mu_{PET \wedge FISH}(cat) = 0.6 < \mu_{PET \wedge FISH}(g)$. Similarly a cod with $\mu_{L_1}(cod) = 0$, $\mu_{L_2}(cod) = 0$, $\mu_{L_3}(cod) = 1$, $\mu_{L_4}(cod) = 1$ has membership $\mu_{PET \wedge FISH}(cod) = 0.6 < \mu_{PET \wedge FISH}(g)$.

There are a couple of points to mention regarding this treatment of the pet fish problem. Firstly, the result is reliant on the values chosen for e.g. $\vec{\lambda}_{PET}$. Choosing $\vec{\lambda}_{PET} = [0.75, 0.25]$ with the other values the same does not produce the same set of inequalities. Further analysis of the model is needed to elucidate these limitations, and modelling with actual data would be useful. Secondly, consider an entity for which $\mu_{L_i}(Y_i) = 0 \; \forall i$, i.e. all attributes for either pet or fish are lacking. This might be a tree, for example. With the choice of $\varepsilon \sim U(0, 1.25)$, the membership $\mu_{PET \wedge FISH}(tree) = 0.2$, when this value should clearly be 0, since a tree is not a pet fish in any way. This objection may be dealt with, however, by choosing a distribution for ε for which $\Delta(1) = 0$. This could be done piecewise, or by choosing a distribution that has this property

such as a $Beta(2, 1)$ distribution. Lastly, an entity for which $\mu_{L_i}(Y_i) = 1 \; \forall i$ will have a higher membership in 'pet fish' than the goldfish, since the prototype for a pet fish which we have specified here is something that lives in the house and is furry and lives in water and is scaly. The issue here is of interaction between dimensions, such as 'furry' and 'scaly'. We have not touched upon this type of interaction between dimensions in any examples thus far, and this is an area for further research.

5. Discussion and future directions

We have proposed here a formalism extending Gärdenfors' conceptual spaces theory, so as to incorporate the vagueness of natural language using a random set based prototype model. The framework we propose gives a mechanism for forming concepts as a combination of properties from integral domains. The innovation we have introduced here is that these properties are mapped into a binary combination space which itself is treated as a conceptual space. We have shown that the idea of concepts as weighted sums of properties, proposed by Gärdenfors and Zadeh, arises naturally as a special case of our framework, namely when the threshold of the concept in the binary combination space is uniformly distributed across the whole space (Theorem 4). We also characterise the combination weights as the necessity of a property to the concept, using the technical definition of necessity from possibility theory.

The framework we propose is hierarchical. Therefore, combinations may consist of multiple properties, as described above, or, as we describe in section 4.3.2, of two or more concepts which are themselves defined as combinations of properties. Again, under certain specific conditions, we can recover the results reported in [22] that the importance of properties in a conjunction of concepts is a weighted sum of the importance of the properties to the individual concepts. We also show how our framework can be applied to property–concept combination. In fact, the distinction between properties and concepts is somewhat artificial, and the conjunctive combination of any sort of concept can be performed within this framework.

A key element of human concept use, however, is that there are various instances where concepts cannot be adequately characterised as a simple weighted sum of properties. We can account for this by using our more general model, in which the threshold in the binary combination space does not have to be distributed according to the uniform distribution. A key result is that necessity and impossibility are carried through from constituent concepts into the combined concept. We have shown in Example 18 how our characterisation of the importance of an attribute in terms of the concept of necessity from possibility theory allows us to take account of this phenomenon. We have further shown how attribute loss and emergence, characterised as the diminished or increased importance of an attribute to a combined concept, may occur, as illustrated in Fig. 10.

We have also explained how non-commutativity and dominance effects may be modelled, by setting the weights used in the binary combination space. Hampton finds that in general, the second noun in the conjunction is given more weight than the first. This aspect could simply be built in to the weighting. Dominance effects are seen when one concept has more features than the other. Again, weightings could take this into account. We have given examples to show how two of the key problems in this area may be accounted for. These are the conjunction fallacy and the guppy effect, illustrated in Examples 19 and 20 respectively.

We therefore argue that our framework is better able to account for key aspects of human use of concepts than standard conceptual spaces approaches [2,6], which do not attempt to account for these non-compositional features. In contrast, the quantum approach has shown examples of how to account for this type of problem, and has done so successfully. We argue, however, that our approach is more conceptually straightforward than the quantum approach, with the formulation of the latter requiring high dimensional Hilbert spaces and the mathematics of quantum mechanics.

At present, our model does not distinguish between membership in and typicality to a concept. There is certainly a difference between these two notions, and an effective model of concepts should elucidate how these two ideas can be unified, or accounted for. Hampton [24] argues that membership and typicality may be subsumed within one approach, on the basis that membership may be decided by applying a suitable threshold to typicality. This type of approach could easily be incorporated into our model. We might ask where the threshold should be placed. However, we could simply place the threshold at 0.5 everywhere and then tune the weights of the model, rather than having separate thresholds for each concept.

The approach we have proposed here is particularly well-suited to concepts that can be described via a collection of attributes. However, some types of concepts are less well-suited to this type of description. For example, some concepts may be defined crisply in terms of necessary and sufficient conditions, such as the concept of even numbers. Another example, given by Barsalou [8] is that of goal-derived categories. These are defined by the extent to which they allow a goal to be reached. So attributes for 'things to eat on a diet' has as an ideal 'foods with zero calories'. These tend not to be central to a concept, although they could still be seen as prototypical in some way. Another example given is that of 'ad-hoc categories', in particular 'things to sell at a garage sale'. Barsalou argues that these cannot really be defined by a specific set of attributes.

With regard to goal-derived categories, we argue that to a certain extent, we can view the ideals as attributes to be achieved. Therefore, these could still be used as dimension in a conceptual space. Ideals may not be central tendencies, however, we do not require this in our model, since typicality to a concept is defined via a probability distribution based on distance to a prototype. Simply setting the probability to 0 on one side of the prototype allows for non-central ideals.

However, the example of ad-hoc categories does suggest further work. We may view categories defined via attributes as conjunctive in some sense: we use conjunctions of attributes. However ad-hoc categories are defined disjunctively: this old bicycle or these tins of paint etc. A movement from disjunctive to conjunctively defined categories might be viewed as a type of learning: children might see the concept 'cat' as defined by 'Tibbles or Patch', but later extract the attributes that allow generalisation. This would be an interesting area for further work. In particular, when using a semantic vector space approach, for example in statistical analysis of text corpora, the vectors created are formed of a mixture of instances of the concept, that would go into a disjunctive definition of the concept, and attributes, that would go into a conjunctive definition. An analysis of when a disjunctive or conjunctive definition is more appropriate, and the interaction between the two in, for example, learning concepts, would be of interest.

A further area for investigation would be to look at the questions raised by [1,27] concerning privative adjectives, i.e., those where the application of the adjective entails that the adjective-noun combination is no longer an instance of the noun. Examples are 'artificial', 'former', 'fake', and so on. Using the combination of the conjunctive and the hedged approach given in [36] might be of use here, to develop these more complex 'type 2' hedges, as Lakoff calls them [30]. An analysis of the effects of the hedges might be to say, for example, that the hedge 'fake' increases the weight of those attributes that are visually important, and decreases the weight of the other most important attributes. An alternative approach would be to tag 'definitional' attributes that the hedge 'fake' can then negate. More generally, the semantic differences between different types of adjectives that [41] discuss should be considered in further work. Adjectives that consist of modifying one property of a concept are well described by our model. Some of the more complicated adjectives such as 'abusive' may be more difficult to describe. Part of the work must be done by giving such an adjective an adequate representation in a conceptual space. Then the application of the adjective may consist of adding domains to its noun, or modifying existing domains.

The potential for application of this model should be discussed. As presented here, in its full generality, the model has a large number of free parameters, and a number of these need to be pinned down before they can be used. Various choices must be made before the model can be applied. For example, the type of threshold distribution is likely to be very dependent on the type of concept or property that is being described. Further, the type of normalisation that we might want to apply to the dimensions must be carefully chosen or learned from the data. We give here a number of possible applications of the model and describe how the parameters can be limited in those cases.

Firstly, the model can be applied as a model of concept combination in examining psychological data. In [34] the model is applied to a range of data from [24,23,22]. An example of the type of data is as follows. Participants were asked to rate the membership of instances such as 'Penguin', 'Dog', 'Cockatoo', and so on in concepts and their combinations – in this example, the concepts would be 'Birds', 'Pets', 'Birds which are Pets', and 'Pets which are Birds'.

Hampton's original analysis finds that mean typicality ratings can be systematically predicted using a multilinear regression. We apply the same analysis to the mean membership ratings of the items. The model we use maps each of the constituent concepts into a binary combination space $\{0, 1\}^2$. Each dimension of this space is weighted, with the weights summing to 1, resulting in weight vector $(\lambda, 1 - \lambda)$. The threshold ε in the binary space is distributed uniformly, $\varepsilon \sim U(0, b)$. We therefore fit just two parameters λ and b. This gives us the expression

$$\mu_{\theta \wedge \varphi}(x) = \mu_\theta(x)\mu_\varphi(x) + \Delta(\lambda)(1 - \mu_\theta(x))\mu_\varphi(x) + \Delta(1 - \lambda)\mu_\theta(x)(1 - \mu_\varphi(x))$$
$$+ \Delta(1)(1 - \mu_\theta(x))(1 - \mu_\varphi(x))$$

Hampton uses a multilinear regression given by:

$$\mu_{\theta \wedge \varphi}(x) = k_1\mu_\theta(x) + k_2\mu_\varphi(x) + k_3\mu_\theta(x)\mu_\varphi(x) + k_4$$

We therefore use fewer parameters than does Hampton to fit this data. The fit we achieve is comparable to Hampton's when using Akaike's 'An Information Criterion' for small samples (AIC_c).

Another way to reduce the number of parameters is to make the simplifying assumption that the threshold in the higher level space is distributed according to the assumptions of Theorem 4, i.e., that the combination of concepts is simply a weighted sum of membership values. This assumption was used in investigating the adoption of conjunctive concepts in a multi-agent simulation of language evolution [34,35]. In this application, agents are equipped with a range of basic concepts which they use to communicate about points in a conceptual space. Via iterated dialogues, agents converge onto a shared set of dimension weights that characterise the space. The weights to which agents converge is determined by the distribution of objects within the conceptual space.

Applications of the theory might also be possible in online classification tasks, for example in film. In an analogue to the 'pet fish' phenomenon, a film might not have typical characteristics of a horror film, nor of a comedy film, but be prototypical of a 'comedy horror'. In this cases, a weighted sum formulation on its own is not able to account for these phenomena. The fact that comedies and horror do not share many characteristics might be taken to be diagnostic of the fact that 'non-compositional' effects may be seen here. Further, we have not yet developed a theory for how to deal with contradictory attributes. These will be key to how non-compositional effects arise, and a simple weighted sum combination does not explain these in a satisfactory manner.

Within the framework we have discussed conjunction and disjunction of nouns and adjectives. There are, of course, many other operators and word types that should be captured in a full account of concepts. Gärdenfors and Warglien have started to develop conceptual spaces for verbs [21,48], and Gärdenfors has begun the development of a semantics for conceptual

spaces [20]. Other approaches to developing a full semantics use the notion of a semantic vector space, based on text corpora [11,38]. Our model could be extended to include other word types and composition types. Within our framework we have explicitly avoided the problem of conflicting prototypes. However, this is an area which must be developed. One important aspect of conflicting prototypes is that having conflicting attributes allows interesting phenomena to emerge. This goes hand in hand with another aspect that we have not discussed: the need for some sort of inference system.

A key element of our framework is the weights given to dimensions in the combination space. Results in [35] and ongoing work show that within a multi-agent model of language users, these weights can be related to the distribution of elements in the conceptual space, explaining why and how different dimensions should have different weights.

Further research will investigate how attributes may affect one another. For instance, in the pet fish example, the attributes 'furry' and 'scaly' might interact since they are, to some extent, incompatible. This is not examined so far, and indeed we require that the constituent concepts may not have contradictory prototypes. Further developing the framework for these cases is likely to allow even more effective modelling of non-compositional effects.

Acknowledgements

References

[1] Nabil Abdullah, Richard A. Frost, Adjectives: a uniform semantic approach, in: Advances in Artificial Intelligence, Springer, 2005, pp. 330–341.
[2] Benjamin Adams, Martin Raubal, A metric conceptual space algebra, in: Spatial Information Theory, Springer, 2009, pp. 51–68.
[3] D. Aerts, L. Gabora, A theory of concepts and their combinations I: the structure of the sets of contexts and properties, Kybernetes 34 (1/2) (2005) 167–191.
[4] Diederik Aerts, Quantum structure in cognition, J. Math. Psychol. 53 (5) (2009) 314–348.
[5] Diederik Aerts, Liane Gabora, Sandro Sozzo, Concepts and their dynamics: a quantum-theoretic modeling of human thought, Top. Cogn. Sci. 5 (4) (2013) 737–772.
[6] Janet Aisbett, Greg Gibbon, A general formulation of conceptual spaces as a meso level representation, Artif. Intell. 133 (1) (2001) 189–232.
[7] Marco Baroni, Roberto Zamparelli, Nouns are vectors, adjectives are matrices: representing adjective-noun constructions in semantic space, in: Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, 2010, pp. 1183–1193.
[8] Lawrence W. Barsalou, Ideals, central tendency, and frequency of instantiation as determinants of graded structure in categories, J. Exp. Psychol. Learn. Mem. Cogn. 11 (4) (1985) 629.
[9] Peter D. Bruza, Kirsty Kitto, B. Ramm, Laurianne Sitbon, D. Song, S. Blomberg, Quantum-like non-separability of concept combinations, emergent associates and abduction, Log. J. IGPL 20 (2) (2012) 445–457.
[10] John A. Bullinaria, Joseph P. Levy, Extracting semantic representations from word co-occurrence statistics: a computational study, Behav. Res. Methods 39 (3) (2007) 510–526.
[11] Bob Coecke, Mehrnoosh Sadrzadeh, Stephen Clark, Mathematical foundations for a compositional distributional model of meaning, arXiv preprint, arXiv:1003.4394, 2010.
[12] Leda Cosmides, John Tooby, Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty, Cognition 58 (1) (1996) 1–73.
[13] Lieven Decock, Igor Douven, What is graded membership?, Noûs 48 (4) (2014) 653–682.
[14] Igor Douven, Lieven Decock, What verities may be, Mind (2015), in press.
[15] D. Dubois, H. Prade, The three semantics of fuzzy sets, Fuzzy Sets Syst. 90 (2) (1997) 141–150.
[16] Didier Dubois, Henri Prade, Possibility Theory, Springer, 1988.
[17] Kit Fine, Vagueness, truth and logic, Synthese 30 (3) (1975) 265–300.
[18] M. Freund, On the notion of concept I, Artif. Intell. 172 (4–5) (2008) 570–590.
[19] P. Gärdenfors, Conceptual Spaces: The Geometry of Thought, The MIT Press, 2004.
[20] Peter Gärdenfors, The Geometry of Meaning: Semantics Based on Conceptual Spaces, MIT Press, 2014.
[21] Peter Gärdenfors, Massimo Warglien, Using conceptual spaces to model actions and events, J. Semant. (2012) ffs007.
[22] J. Hampton, Inheritance of attributes in natural concept conjunctions, Mem. Cogn. 15 (1) (1987) 55–71.
[23] J. Hampton, Disjunction of natural concepts, Mem. Cogn. 16 (6) (1988) 579–591.
[24] J. Hampton, Overextension of conjunctive concepts: evidence for a unitary model of concept typicality and class inclusion, J. Exp. Psychol. Learn. Mem. Cogn. 14 (1) (1988) 12.
[25] J. Hampton, Conceptual combinations and fuzzy logic, in: R. Belohlavek, G.J. Klir (Eds.), Concepts and Fuzzy Logic, The MIT Press, 2011.
[26] James A. Hampton, Testing the prototype theory of concepts, J. Mem. Lang. 34 (5) (1995) 686–708.
[27] H. Kamp, B. Partee, Prototype theory and compositionality, Cognition 57 (2) (1995) 129–191.
[28] J.A.W. Kamp, Two theories about adjectives, in: Edward Keenan (Ed.), Formal Semantics of Natural Language, Cambridge University Press, Cambridge, 1975, pp. 123–155.
[29] Sarit Kraus, Daniel Lehmann, Menachem Magidor, Nonmonotonic reasoning, preferential models and cumulative logics, Artif. Intell. 44 (1) (1990) 167–207.
[30] G. Lakoff, Hedges: a study in meaning criteria and the logic of fuzzy concepts, J. Philos. Log. 2 (4) (1973) 458–508.
[31] Thomas K. Landauer, Susan T. Dumais, A solution to plato's problem: the latent semantic analysis theory of acquisition, induction, and representation of knowledge, Psychol. Rev. 104 (2) (1997) 211.
[32] J. Lawry, Modelling and Reasoning with Vague Concepts, vol. 12, Springer, 2006.
[33] J. Lawry, Y. Tang, Uncertainty modelling for vague concepts: a prototype theory approach, Artif. Intell. 173 (18) (2009) 1539–1558.
[34] Martha Lewis, Modelling compositionality of vague concepts, PhD thesis, University of Bristol, 2015.
[35] Martha Lewis, Jonathan Lawry, Emerging dimension weights in a conceptual spaces model of concept combination, in: Proceedings of the 50th Anniversary Convention of the AISB, Society for the Study of Artificial Intelligence and the Simulation of Behaviour, 2014.
[36] Martha Lewis, Jonathan Lawry, A label semantics approach to linguistic hedges, Int. J. Approx. Reason. 55 (5) (2014) 1147–1163.
[37] Kevin Lund, Curt Burgess, Producing high-dimensional semantic spaces from lexical co-occurrence, Behav. Res. Methods Instrum. Comput. 28 (2) (1996) 203–208.

[38] Jeff Mitchell, Mirella Lapata, Composition in distributional models of semantics, Cogn. Sci. 34 (8) (2010) 1388–1429.
[39] Richard Montague, The Proper Treatment of Quantification in Ordinary English, Springer, 1973.
[40] D.N. Osherson, E.E. Smith, On the adequacy of prototype theory as a theory of concepts, Cognition 9 (1) (1981) 35–58.
[41] Victor Raskin, Sergei Nirenburg, Lexical semantics of adjectives, New Mexico State University, Computing Research Laboratory Technical Report, MCCS-95-288, 1995.
[42] John T. Rickard, Janet Aisbett, Greg Gibbon, Reformulation of the theory of conceptual spaces, Inf. Sci. 177 (21) (2007) 4539–4565.
[43] E. Rosch, Cognitive representations of semantic categories, J. Exp. Psychol. Gen. 104 (3) (1975) 192.
[44] Steven Schockaert, Henri Prade, Interpolative and extrapolative reasoning in propositional theories using qualitative knowledge about conceptual spaces, Artif. Intell. 202 (0) (2013) 86–131.
[45] Edward E. Smith, Edward J. Shoben, Lance J. Rips, Structure and process in semantic memory: a featural model for semantic decisions, Psychol. Rev. 81 (3) (1974) 214.
[46] E.E. Smith, D.N. Osherson, Conceptual combination with prototype concepts, Cogn. Sci. 8 (4) (1984) 337–361.
[47] Amos Tversky, Daniel Kahneman, Extensional versus intuitive reasoning: the conjunction fallacy in probability judgment, Psychol. Rev. 90 (4) (1983) 293.
[48] Massimo Warglien, Peter Gärdenfors, Matthijs Westera, Event structure, conceptual spaces and the semantics of verbs, 2012.
[49] L.A. Zadeh, Fuzzy sets, Inf. Control 8 (3) (1965) 338–353.
[50] L.A. Zadeh, A fuzzy-set-theoretic interpretation of linguistic hedges, J. Cybern. (1972).
[51] L.A. Zadeh, The concept of a linguistic variable and its application to approximate reasoning – I, Inf. Sci. 8 (3) (1975) 199–249.