To appear in "Concepts and Fuzzy Logic"

Editors: Radim Belohlavek and George J. Klir

MIT PRESS

FINAL DRAFT

CHAPTER 8. Conceptual Combinations and Fuzzy Logic

James A. Hampton

References

8.1 What are Conceptual Combinations?

Suppose that an intelligent system has two concepts *A* and *B*. A conceptual combination will then be a third concept that is the result of combining *A* and *B* according to some principle. Some examples will make this clearer. If *A* is the concept of the action of FRYING and *B* is the concept of the class of objects called PAN, the two concepts could be combined in order to generate the complex concept FRYING PAN. The ability of humans to combine concepts is a vital part of our creativity. We can take concepts that have never been combined before and try them out together. So by recombining FRYING PAN and WASHING MACHINE, we could easily generate FRYING MACHINE and WASHING PAN. Sometimes such novel combinations will appear meaningless, but often enough new ideas will be formed.

Conceptual combination is not confined to the combination of actions and objects. In fact many concepts are freely combinable with others. However there must clearly be constraints on which combinations will make much sense. Concepts have to be the *right kind of thing* in order to combine successfully. Chomsky's famous sentence example of "colorless green ideas sleep furiously" contains conceptual combinations that are all quite problematic. Can an idea have a color? Can something be both colorless and green at the same time? Can an idea sleep, and is it possible for anything to sleep in a furious way?

8.2 Intersective and Non-Intersective Combinations

To bring some clarity to this general idea of combining concepts, we need to distinguish between different kinds of conceptual combination. It is sensible to distinguish those combinations which appear to have some recognizable *logical* form from those which rely on a wider range of heuristic processes to derive an interpretation. Intersective combinations are those which (at least at first blush) appear to correspond to set intersection of the categories designated by the meanings of the two words. Often, adjective+noun phrases will take this form, as in “green shirt” or “friendly neighbor”. A green shirt is green and it is a shirt. A friendly neighbor is both friendly and a neighbor. In terms of set membership, someone is a member of the set of all friendly neighbors if and only if they are a member of the set of all friendly people and also a member of the set of all neighbors. Conversely, noun+noun phrases combining two noun categories often fail to be intersective. A bottle opener is not both a bottle and an opener, and traffic cop is not both traffic and a cop. Note that even though the role of the first word can vary widely, the second word in the combination plays a more or less fixed role. It points to a given class of things. Traffic cops are cops, bottle openers are openers, and so forth. In linguistic terminology, the second noun is the *head noun*, while the first plays the role of *modifier*, qualifying the meaning of the head by restricting it to a subclass. So a traffic cop is the subtype of cop who manages road traffic.

From these examples, we can broadly generalise that a conceptual combination involves taking the second word (the head), and modifying or restricting its meaning in some way. Intersective modification will involve finding the intersection of the classes that are named by the two words, while non-intersective modification involves the discovery of the implicit link that explains just how the modifier works to modify the head noun meaning. Often enough, a non-intersective combination will still denote a subtype of the head noun category, such as a type of cop who directs traffic, or a type of opener designed for opening bottles. The difference is that the semantic relation needed to make the connection (e.g. “directs” or “designed for”) must be generated by the hearer before the meaning of the whole can be understood.

8.2.1 Non-Intersective Combination

Non-intersective conceptual combination has been widely studied. Building on work in linguistics (Gleitman and Gleitman 1970; Levi 1978), psychologists (Cohen and Murphy 1984; Gagné and Shoben 1997; Wisniewski 1997) have developed an account of the common ways in which such combinations are interpreted. It would appear that there is a set of around 6-10 fundamental semantic/thematic relations, which can be employed in either direction in order to generate interpretations. Thus the combination *AB* could be interpreted in different cases as

USE: *B* used for *A* (*cooking gas*), *B* that uses *A* (*gas cooking*);

MATERIAL: *B* made of *A* (*clay brick*), *B* used to make *A* (*brick clay*);

LOCATION: *B* in which *A* is found (*deer mountain*), *B* found in *A* (*mountain deer*);

CONTAINMENT: *B* contained in *A* (*pot noodle*), *B* in which *A* is contained (*noodle pot*);

and so forth. Other thematic relations include CAUSE, HAS, MAKES, and ABOUT.

These basic relations are reminiscent of the relations used in case grammars (Fillmore 1968). In modern Romance languages they are encoded with specific prepositions such as French á, de, pour, dans, sur, and en). Thus a *tasse de thé* in French is a cup of tea, while a *tasse à thé* is a tea-cup.

In addition to these interpretations based on thematic relations, people also employ other heuristic strategies to arrive at interpretations. One, discovered by Wisniewski

(1997) is to find a salient property of the modifier and to use that to modify the head noun. For example a “zebra mussel” is a mussel that has stripes like a zebra. Wisniewski calls this process Property mapping, and it involves a different strategy from that of finding a thematic relation. In particular, as Estes and Glucksberg (2000) demonstrated, a property mapping requires that the modifier noun has a highly salient property (such as the stripes of a zebra), and that this salient property corresponds to a variable dimension of the head noun (in this case the surface appearance of the mussel shell). Otherwise, the property mapping interpretation is unlikely to be available as a sensible meaning for the phrase.

In sum, the process of combining concepts can employ a wide range of knowledge structures, and can lead to interpretations that will defy the application of any given logical formalism. There is plenty of scope for ambiguity (a *criminal lawyer* may be a lawyer who represents defendants accused of crimes, or may be a lawyer who is him/herself a criminal) and creative extensions of word meaning (*couch potato* leading to *web potato*), all of which should warn the logician that this is territory not best traversed with the meagre representational formulae of logic, of whatever kind. As Murphy (1988, 2002) is at pains to point out, general conceptual combination is one of those problems in cognitive science/artificial intelligence, like non-monotonic reasoning, that cannot be solved without an open-ended access to world knowledge.

8.2.2 Intersective Combination – Adjectival Modification

In the case of adjectival modification, the problems of laying out a logical structure governing conceptual combination are at first sight more easily solved, in that many adjective-noun combinations are intersective, as defined above. The semantics of the combination are just a matter of finding the set intersection of the two categories denoted by the two words. A *green shirt* is green and it is a shirt.

The appearance of simplicity is however misleading as there are still major difficulties to be overcome. The first is that many adjectives change their meaning as a function of the noun that they are modifying. Murphy (1990) demonstrated this with a number of adjectives such as “open” or “fresh”. An open door, an open face, an open question, an open view, and an open golf championship all use very different senses of the word. Interestingly they all still retain a common sense, suggesting that this is not a simple matter of lexical ambiguity (as with words like “palm” or “bank”). Rather, the adjectival concept has become specialised, adapting a very abstract schematic idea (lack of boundaries, freedom of access) to the very different conceptual domains of a person’s friendly expression, an unanswered question or a competition with no entry restrictions. Such expressions do not always translate well between languages either, indicating that to a greater or lesser degree they rely on the language speaker learning the individual expressions as idioms. As Lakoff (1987) has pointed out, many extensions of the use of words can be justified *post hoc,* without having a predictive model of where extensions will occur. This fact is to be expected, given that languages evolve much like biological species, taking adaptive turns based on sequences of random shifts in usage within a population of speakers. Biology cannot predict that there should be creatures like zebras in Africa, but given that fact, the theory of evolution can provide a story of how they got there.

Another example of the polysemy of adjectives is the case of evaluative adjectives. Kamp and Partee (1995) point out that “John is a good violinist” can be used to infer

"John is a violinist" but does not entail "John is good". The sense of good here is tied to the concept of violinist, and cannot be treated independently.

A second problem for the logic of adjectival combinations is that even within a given narrow domain, the noun may still determine the applicability of the adjective. The words "big" and "small" can refer to a range of domains (minds, ideas, mistakes), but suppose we restrict ourselves just to physical size. Even then we have to know something about the normal range of size for the noun category in order to decide if an object is big or small. A large ant is much smaller than a small cat. Color terms such as "red" will similarly take different expected typical values when applied to different nouns such as hair, wine, face or car. Once again the meaning of the adjective has to be taken *within the context* of the noun. A "large ant" is an ant that is large *for an ant*. To evaluate the truth of the statement "that is a large *x*", we need to have knowledge of the distribution of sizes for the class of all *x*, and adapt the truth conditions accordingly.

A final problem (for now!) consists of so-called privative adjectives. Adjectives may include within them the power to change the truth value assigned to category membership in the noun category. A flawed proof is not a proof, in the sense that it does not prove anything. Other privatives are words such as "former", "fake", "pretend", "putative", "seeming" and "alleged". Nor does the adjective have itself to be privative for this change to occur. A chocolate egg, a plastic flower, and a wooden horse are none of them things of the right kind to be called respectively eggs, flowers and horses. It seems that our use of language allows us to take the name of a concept and extend it to refer to things that have the same appearance. A chocolate egg is an egg because it has the right shape. In fact language use includes a wide variety of ways in which meanings get extended. Words are often used figuratively or metaphorically. Capturing conceptual combinations with fuzzy logic will require a very strict means of identifying when such extended uses are being employed. Unfortunately, as with everything else, the question of when a meaning has been extended has no precise answer, so that the applicability of fuzzy logic to a segment of language may itself be a matter of degree.

8.3 Compositionality

The relevance of all of this to a book on fuzzy logic relates to the notion of compositionality. In a series of papers, Fodor and Lepore (2002) have argued for the importance of compositionality as a fundamental assumption in the representational theory of mind. Briefly, the principle of compositionality says that the meaning of a complex phrase should be based on the meaning of its components, the syntax of the language by which they are combined, and *nothing else*. Much has been written on compositionality, and ways in which the principle can be explicated (Machery et al. 2011). It should be clear from the foregoing discussion that one cannot reasonably expect to find compositional concept combination in much of everyday language. The adherent to the principle will therefore have to rely on a lot of additional linguistic/pragmatic theory of how the underlying compositional logic of a complex phrase can be derived from its non-compositional surface form. Even then, the challenge of providing a compositional semantics for natural language is daunting.

Putting aside these problems for now, let us consider the treatment of truth conditions for complex phrases that are apparently compositional, and hence suitable for a logical treatment.

8.3.1 Intersective Combination – Restrictive Relative Clause Constructions

In a series of empirical studies (Hampton 1987, 1988b, 1997) I investigated how people combine concepts when they are placed in an explicitly intersective linguistic form. A restrictive relative clause, such as “a sport that is also a game” is unambiguous in referring to the subset of recreational activities that are both sports and games. If we accept this interpretation, then we would expect the logical relation of conjunction to underlie the meaning of the combined phrase. In that case, (a) people should treat the phrase symmetrically (“a sport that is also a game” should apply to the same set of activities as “a game that is also a sport”), and (b) people should only consider the phrase to apply to an activity if they also agree that the activity is a sport and that it is a game.

Hampton (1988b) reported a number of studies using six different conjunctive combinations of this kind (others included “a tool that is also a weapon” and “a machine that is also a vehicle”). Participants in the studies were given a list of potential category members (e.g. tennis, archery, chess, trampolining, sky-diving, crosswords) and made three sets of category judgments. First they decided if they were sports, then if they were games, and then finally if they were “sports that are games”. Responses were given on a scale from −3 to +3, where a positive number indicated a “yes” decision with increasing confidence from +1 to +3, and a negative number indicated a “no” decision in the same way. A zero could be used in the case that the item was exactly on the borderline of the category, and the item could be left blank if it was unknown.

The data were analysed as follows. First, mean judgments were calculated for each item for each of the three category judgments (the constituents *A* and *B*, and their conjunction *A* that are *B*). Regression analysis was applied to these means to identify the strength of a linear relation between degree of belonging to the conjunction and degree of belonging to each constituent. Across three experiments, the fit of the regression function was extremely good, providing a positive interaction term was included. Multiple *R* (the correlation between predicted and observed values of membership for the conjunction) averaged around 0.97, where a value of 1 means a perfect match. *R* squared of 0.94 indicated that 94% of the variance variance in conjunctive membership across items could be explained in terms of items' membership for the conjunct categories. The success of the model suggests that a fuzzy logic function could be used to successfully predict membership degree in a relative clause conjunction from membership in the two conjuncts. The function was NOT however one of those usually associated with the intersection of fuzzy sets. The model used a function that was based on the product of the two conjunct memberships, but rescaled so that a hypothetical item that was on the borderline for each conjunct would also be on the borderline for the conjunction. Given that an item that was very clearly not in one of the conjuncts was clearly not in the conjunction, while items that were very typical of both conjunctions were very typical of the conjunction, a geometric mean may be the best approximation to the empirical function.

Further studies (Hampton 1997) confirmed this pattern of data, using different participants to provide the ratings for each category. In addition, combinations involving negated relative clauses were used (e.g. Games that are not Sports), which had the effect of reversing the sign of the negated conjunct in the empirical function, so that the negated conjunct and the interaction term both had negative regression weights. For these later studies, regression analysis applied to frequencies of positive responses (as opposed to mean scale ratings) confirmed the same pattern of data obtained.

It would seem then, that a fuzzy logic function based on a geometric average, with suitable scaling would provide a good model for the way in which people form this kind of conceptual combination. It is interesting that the function was anchored at the borderlines of each concept. Effectively one can suppose that the 0.5 point on the membership scale (which in the present case was the point where 50% of participants agreed that the item belonged) corresponds to a point of complete "quandary" to use Wright's term. It is the point where people are most likely to be in a perfect state of indecision about the question of the categorization of the item. The result of the experiment suggested the following principle for the logic of such states:

When a person is in a quandary about whether *x* should be in category *A* and also in a quandary about whether *x* should be in category *B*, then they are also in a quandary about whether *x* is in the conjunction of *A* and *B*.

This principle is intuitively plausible, corresponding to the propagation of a state of quandary about each conjunct to a state of quandary about the conjunction. (It would also apply to being in a quandary about the disjunction of the two categories – see below). The fuzzy logical averaging function then has the following consequences:

(a) *Overextension*: an item that is somewhat above 0.5 for *A*, but somewhat below 0.5 for *B* may still correspond to the state of quandary (0.5) for the conjunction *AB*. Hence the likelihood of an item being placed in the conjunction will be *greater* than the likelihood of it being placed in category *B*. This overextension of conjunctive membership is a direct consequence of the principle and the use of an averaging function (see Chapter 3, Section 3.4). The overextension has been found in a number of studies (Hampton 1988b, 1997; Storms et al. 1993, 1996, 1998, 1999).

(b) *Compensation*: an item's good degree of membership in one conjunct can compensate for its poor degree of membership for the other, in determining membership in the conjunction. It is as if in choosing a home one had two necessary requirements – a location within 1 hour's commute of work, and a minimum of 80 square metres of floor space. This conjunctive requirement determines the set of acceptable homes as the intersection of the two sets defined by each criterion. The averaging conjunctive rule allows that if one found a place that was only 30 minutes commute from work, one might still consider it acceptable, even though the floor space was only 70 square metres (Chater, Lyon and Myers 1990).

8.3.2 Overextension and Compensation

Overextension can be readily handled within fuzzy logic by the adoption of a suitable function such as the geometric average. One puzzle is why the function should not be closer to one of the standard functions for conjunction, such as the minimum rule or the product rule. A standard constraint on fuzzy logics for categorization is that membership in a conjunction cannot be greater than membership in a conjunct (see Chapter 3, Section 3.4). The truth of TWO statements together cannot be greater than the truth of each one considered individually. Yet overextension clearly breaks this constraint. The logician may not be too concerned. It may be considered sufficient to have shown empirically that membership in the conjunction can be accurately modelled with a fuzzy logic function of one kind or another. However this obviously leaves a lot still to be explained. In particular, why do we have the strong sense that the phrase "*A* that is a *B*" should be the overlap of *A*'s and *B*'s, and should be a subset of *A*'s, when in practice this is not how we categorize items within the complex set? This question suggests a worrying disconnect between the apparent (even transparent) logical form of an expression, and the categorization behavior that it invokes.

Compensation – the second consequence of an averaging function – can be even more of a problem. According to prototype theory (Hampton 1979, 1995; Rosch and Mervis 1975) degree of membership in a category can be most easily explained in terms of the similarity of an item to the prototype or idealised representation of the category's central tendency. Thus tennis is considered to be a sport because in terms of the properties that characterize typical sports (physical, skilled, involves exertion, competitive, has championships, has stars) it ticks all the boxes. Other activities may be marginal to the category because while they have some properties they lack others (for example scuba diving is normally non-competitive although it does involve skill and some physical exertion). Within the psychology of concepts, prototype theory has been superseded in a number of ways, but some form of property-based similarity remains the only account of why there are differences in typicality and why there are marginal members of categories. It is true that some have argued that borderline cases are owing to ignorance of the world, rather than semantic indeterminacy (e.g. Williamson 1994; Bonini et al. 1998). However there remains very little psychological evidence that this provides a general account of problems of the vagueness of meaning in language.

The difficulty for fuzzy logic lies in the fact that the similarity of items to the prototype of a category can continue to increase beyond the point at which membership in the category has reached a maximum. Osherson and Smith (1997) used this fact as a critique of the prototype account of graded membership. Typicality (as similarity to a prototype is usually termed) can not be simply identified with grades of membership, since the former can differ among items that all have full membership in the category. A robin is a more typical bird than a penguin (having the requisite features characteristic of birds in general), but both robins and penguins are fully birds, as judged by the 100% endorsement of the statements “A robin/penguin is a bird”. Thus the continuous truth value assigned to the statement “a penguin is a bird” must be 1, if the person doing the assigning has a firm belief that a penguin is definitely a bird. I argued in (Hampton 2007) that this is not after all a problem for prototypes and graded membership, since typicality and graded membership should be treated as two functions based on the same underlying similarity measure. Typicality is a monotonically rising function of similarity, whereas membership is a non-decreasing function of similarity that starts at 0, starts to rise at a certain point k_1 and then reaches a ceiling of 1 at further point k_2, where k_1 and k_2 are above the minimum and below the maximum values that similarity can take.

Why should this represent a problem for fuzzy logic accounts of conjunction? The difficulty comes when an increase in the typicality of items with respect to category *A* continues to affect membership in the conjunction *AB*, even when membership of the items in *A* is already at the maximum of 1. Consider two stimuli that are identical in shape, being composed of a figure half way between a capital A and a capital H. The vertical lines at the side of the H have been bent in at the top so that they could be taken to be an A or an H, and in fact when people have to choose, they are 50% likely to say one or the other. Now let both figures be colored red, so that everyone agrees 100% that they are both red. If asked if they are examples of a “red H”, they should therefore be inclined to agree around 50% of the time. Being red is unproblematic – everyone says they are red – so the only question relevant to their membership in the conjunction is whether they look more like an H or more like an A. Being identical in shape, there should be no difference in the degree of agreement. So far so good. Now let us suppose that the first figure, Stimulus 1, is a very bright prototypical red, while Stimulus 2 is a rather pale watery red with a slight hint of purple about it. So while both are 100% red in terms of the membership function, they are not equally typical.

In a study reported in (Hampton1996), I found that membership in the conjunction could be affected by differences in typicality, even when the point had been reached where membership was no longer in doubt. Typicality of a clear red could compensate for lack of match in the angle of the letter verticals. For fuzzy logic the challenge is to decide just how to define the membership function $c_A(x)$ in terms of the underlying similarity to prototypes. There are two possibilities, neither of which is without problems. If $c_A(x)$ is mapped to probabilities of category membership, then it will capture differences in the amount of disagreement and inconsistency in membership decisions. As a measure of "truth" it is intuitively most plausible to map $c_A(x)$ in this way. The function reaches a value of 1 at the point at which everyone assents to the truth of the statement, and a value of 0 at the point where no one assents to it. But then it is not possible to capture differences in typicality within the measure, since typicality continues to increase after the point at which $c_A(x)$ reaches a maximum (and continues to decrease after the point at which it reaches a minimum). Alternatively $c_A(x)$could be mapped to the typicality/similarity of an item. But then there would be the unintuitive result that although everyone accepts that x is an A, yet its membership in A is only (say) 0.9.

The phenomenon of compensation argues for the second mapping, since differences in typicality for one category once membership has reached a maximum *do* have an effect on degrees of membership in the conjunction of that category with another.

8.3.3 Disjunction

Further studies (Hampton 1988a, 1997) looked at other logical connectives as they are applied to natural language categories. In (Hampton 1988a), I looked at how judgments of category membership in two categories were related to judgments in their disjunction. Unlike conjunctions, disjunctions can be formed without their being any overlap between the two categories in question. For example one could form the disjunction “birds or trucks”, and in such a case it is plausible (although there is no empirical data on this question) to propose that the maximum rule would apply unproblematically. After all, anything that is at all close to being a bird is not going to have any chance of being a truck, and vice versa, so the two membership functions can simply be summed. A maximum rule and a sum will give the same values for the disjunction of *A* and *B* in the case that $c_A(x) \oplus c_B(x) = 0$ for all x, which for birds and trucks seems very probable. The more interesting case arises when the disjunction is formed of two categories that are semantically related and fall within the same domain, so that $c_A(x) \oplus c_B(x) > 0$, for some item x. Hampton (1988a) measured degrees of membership in two categories *A*, *B* and their disjunction $A \vee B$, using a selection of 8 pairs of categories, and showed that a regression function of the form

$$c_{A \vee B} = pc_A + qc_B - rc_A c_B, \tag{8.1}$$

where p, q and r are positive constants, could do a fair job of predicting disjunctive membership, with Multiple R varying from 0.95 to 0.99 in a within-subjects design and 0.86 to 0.97 in a between-subjects design.

In terms of the probability of an item belonging in the disjunction, given that it was judged to be in one of the disjuncts, there was a tendency for people to underextend. Just as people overextended conjunctions, they underextended the disjunctions. For example 90% of a participant group thought that a refrigerator was a House Furnishing, and 70% of a different group that it was Furniture, but for a third group only 58% agreed that it was either one or the other. Unlike the conjunctions, where the borderline of each conjunct appeared to anchor the borderline for the conjunction (the point of quandary), for disjunctions, an item that was borderline for both disjuncts tended to be excluded from the disjunction. In fact, a function closely fitting the disjunctive borderline was defined (for mean membership values on a scale from −3 to 0 to +3) by the equation

$$c_A + c_B = -2. \tag{8.2}$$

Thus to achieve a borderline membership in the disjunction an item had to have a summed value in the two disjunctions greater than some constant value. Items that were good members (+3) in one of the disjuncts were guaranteed to belong. But items that were only atypical members (+1) were only borderline to the disjunction if they were very poor members (−3) of the other disjunct.

The challenge for any type of logic, fuzzy or otherwise, is that the disjunction operation appears to break the constraints of classical logic. That is to say more people believe that the refrigerator is a House Furnishing than believe that it is either a House Furnishing or Furniture. The principle that says that if *A* is true then *A* or *B* is also true does not apply to these judgments.

Underextension was not the only problematic result in this study. Where a pair of categories divide a larger domain, such as Fruits and Vegetables, the judgments of disjunctive membership go the other way, exceeding the constraints of the maximum and *even* the sum rule. For example across the three groups of participants no one in the first group ever judged a mushroom to be a fruit, 50% of the second judged it to be a vegetable, while 90% of the third group judged it to be either a fruit or a vegetable. So given that no one apparently believes that a mushroom is a fruit, why is it more likely to be in the disjunction “fruits or vegetables” than in the class “vegetables” alone?

8.3.4 Negation

A further study (Hampton 1997) looked at the function of negation within restricted relative clause constructions. Negative concepts have sometimes been cited as examples of concepts which have no prototype (e.g. Connolly et al. 2007). Although it is easy to determine the membership of a class such as “not a fruit”, it seems that this determination does not involve similarity to some prototypical non-fruit. We therefore have to assume (along with fuzzy logic) that negation is part of the set of syntactic operations that can be applied to given positively defined concepts. In the context of conceptual combinations, however, negation can play a role in determining the attributes of a given concept. For example “non-alcoholic beverage”, or “non-smoking bar” are easily understood, and could even permit degrees of set membership. For example many non-alcoholic beers state that they contain less than 1% alcohol, which would make them less clear members of the category of non-alcoholic beverage than say lemonade or orange juice.

Hampton (1997) investigated the relation between membership of items in two constituent categories such as Sports and Games, and then membership either in the conjunction (Sports that are also Games) or a conjunction with negated relative clause (Sports that are not Games). Different groups were used to make each of the judgments. Judgments were made on a 7-point rating scale, in which a positive number (+1 to +3) was used to indicate that an item was in the category (with increasing typicality), and a negative number (-1 to -3) indicated that it was not in the category (with increasing unrelatedness). Data were analysed both in terms of mean scale ratings, and also in terms of the proportion of positive (versus negative) ratings given. For regression functions predicting either mean rated degree of category membership, or proportion of positive categorization responses, the results were consistent with Hampton (1988b). The negated conjunctions were predictable by a multiplicative function of the two constituent membership values, but with membership in the second category taking a negative weight. One notable difference from the earlier results (Hampton, 1988) was the occurrence of many more cases where an item was in the (negated) conjunction but not in the corresponding conjuncts. For example a Tree House was considered to be a Dwelling by 74% of one group, to be Not a Building by 20% of the second group, but to be a Dwelling that is not a Building by 100% of a third group. There was considerable overlap between the sets *A* that are *B*, and *A* that are not *B*. When the probability of an item being in one was added to the probability of being in the other, theoretically the sum should be the probability of being in set *A*, according to

$$c_{A \wedge B}(x) + c_{A \wedge \neg B}(x) = c_A(x). \tag{8.3}$$

In practice, the sum could reach values well above 1.0, indicating that (8.3) does not constrain judgments of membership in these categories. The actual value of the sum in (8.3) was predicted by a weighted average of $c_A(x)$ and the distance of $c_B(x)$ from a probability of 0.5.

8.4 Conceptual Combination – the Need for an Intensional Approach

By this point it should be apparent that truth-functional approaches to conceptual combination based on extensional measures of category membership find it very difficult to account for the actual data on how people combine concepts. It is not so much a matter of finding the *right* fuzzy logic function to define conjunction, disjunction or negation. Rather the function depends on the semantic contents or intensions of the concepts in question. Several papers in the psychological literature responded to the challenge of Osherson and Smith's (1981, 1982) papers by making the same point. Thus Cohen and Murphy (1984), Hampton (1987), and Smith and Osherson (1984) all came to the conclusion that the effect of forming a conjunction of two concepts *A* and *B* will vary depending on the relation between the intensions of the two concepts, and how they interact.

By an intension I mean the set of descriptive properties that are held to be generally true of a class, and thus provide a means of determining whether some novel item should belong in the class or not. Concepts have both extensions and intensions. For a concept like TRIANGLE, the extension is the set of all plane figures that fall under the term "triangle". Extensions are used for quantificational logic, and most truth-functional semantic systems. Hence the interpretation of a green shirt as the intersection of green things and shirts uses the extensions of each of the terms and forms the set of items that fall in both extensions. Intensions are of more interest for psychologists, since they reflect the way in which we represent the concept internally. Frege (to whom we owe much of this way of seeing things) pointed out that two concepts could have the same reference or extension (denote or refer to the same set of objects in the world) and yet have different senses or intensions. If a *triangle* is a closed plane figure with three angles, and a *trilateral* is a closed plane figure with three sides, the two concepts refer to the same set of mathematical objects, since as should be obvious every triangle has three sides, and every trilateral has three angles. But intensionally they are different, since it is possible for John (whose knowledge of geometry is very slight) to know that the angles in a triangle sum to 180˚, but not to know that the angles in a trilateral sum to 180˚. In order to explain how we combine natural concepts into conjunctions and disjunctions, it is necessary to look closely at the intensional information that represents each concept, and how the intension of one concept interacts with that of the other. We cannot get even close to an account of the logic of natural categories by looking at the extensions alone.

8.4.1 Contradictions and Fallacies

One of the key reasons for looking for an intensional model to explain conceptual combinations is the widespread occurrence of apparently contradictory or fallacious reasoning involving concepts. We have seen above how people consistently place items in the conjunction of two sets, while at the same time denying that the item is in one of the conjuncts. And people are reluctant to allow that an item is in the disjunction of two sets even though they will allow it in one of the disjuncts. Other similar effects have been reported in the literature, and all point to the fact that the human conceptual system is not based on a firm grounding in logic but has a different design with different purposes. I will briefly give some examples of these non-logical effects, starting with a study of my own in (Hampton 1982). Participants were asked to judge whether categorizations of the kind "*A* is a kind of *B*" were true or false. The study showed that this form of categorization may be intransitive. People said that clocks and chairs were furniture, and that Big Ben was a clock and a car-seat was a chair, but that Big Ben and car-seats were *not* furniture. A well reported effect of the same kind is Tversky & Kahneman's (1984)

Conjunction Fallacy. Just as people overextend conjunctive concepts, so they judge subjective probabilities in a way that overestimates the likelihood of conjunctive events. In this case people were told about Linda, who was a radical when in college, and then went on to judge it more likely that she was a feminist bank teller than simply a bank teller.

Similar non-logical effects have been reported by Sloman in studies of category-based inductive reasoning. In the premise specificity effect, people consider that an argument such as

"All apples are diocogenous therefore all Mcintosh apples are."

is stronger than

"All fruit are diocogenous therefore all Mcintosh apples are.",

even though both arguments are perfectly strong (given people's knowledge that McIntosh apples are apples, and that apples are fruit). They preferred an argument with a more specific premise. A similar effect was seen in the conclusion of an argument, with more typical conclusions being preferred, such as

"All animals have property *X* therefore all mammals do."

being considered a stronger argument than

"All animals have property *X* therefore all reptiles do."

Finally, Jönsson and Hampton (2007) reported an intensional equivalent of the Tversky and Kahneman fallacy, which they termed the Inverse Conjunction Fallacy. People were inclined to consider it more likely to be true that *All lambs are friendly*, than that *All dirty lambs are friendly*, paying no heed to the inclusion relation between lambs and dirty lambs.

8.4.2 An Intensional Theory of Concepts

From the earliest theories, psychologists have sought to model concepts in terms of their intensions. It makes intuitive sense to propose that people represent a class of things in the world by representing their typical characteristic properties. However this has not been the standard approach taken by logic, where the focus has been on the sets of objects in the world and the relations between those sets. Thus semantic theories in linguistics and philosophy have concentrated on describing the relation between statements in a language and the conditions of the world that would make those statements true or false (or in the case of fuzzy logic, true to some degree). Given a set of statements and their associated truth values, then the task of logics is to determine how truth is preserved as statements are combined through various syntactic operators. Psychological data however show that one cannot treat the truth of statements in a content-independent way. For example, when two related categories are combined in a disjunction, Hampton (1988a) showed that the function relating disjunct degrees of truth to truth of the disjunction varies across different pairs. While HOBBIES OR GAMES tended to be underextended (*beer drinking* was considered a hobby, but not to be in the category "hobbies or games"), FRUITS OR VEGETABLES were over-extended, with *mushroom* and *coconut* being near perfect members of the disjunction, although only partial members of one disjunct and not members at all of the other. To account for this dependence of the functions on the contents of the particular concepts being combined requires that they not be treated as logical atoms, but that the interaction of their intensions be considered.

Hampton (1988b) provided a detailed account of how one might combine concepts into conjunctions, using the intensions of each concept as the starting point. Classically a conjunction of extensions can be formed by taking a set union of the intensions. For example to pick out the class of green shirts, one has to take all the defining properties of green things and all the defining properties of shirts and form their set union, so that green shirts have the defining properties of both classes. Hampton's Composite Prototype Model (CPM) takes the same approach as a starting point, but with the proviso that the properties in question are not *defining* in the classical sense of being necessary and constitutive of the class, but are instead *prototypical*. In fact the properties vary from some which may be highly central and universally present in the class (e.g. *fish have gills)* to those which are almost incidental to the kind (e.g. *fish are eaten with French fries).* It is hypothesised that attributes have an "importance" which will reflect the degree to which they affect the similarity of an item to the concept prototype. Differences in attribute importance arise from statistical co-occurrence frequency, and from the degree to which attributes are embedded in causal dependencies with other attributes. Thus an attribute will be important if it has high predictive validity (most category members have the attribute, and most items with the attribute are category members), and if it is the cause or the effect of other attributes in the concept prototype.

The model for constructing a conjunction follows a number of steps (see Hampton 1988b, 1991 for details). In the model, properties are called "attributes"

a) all attributes of each concept are recruited into a composite prototype representation for the conjunction, with an importance based on their average importance for the two conjuncts

b) where an attribute has very low importance, it is dropped from the representation

c) where an attribute is of high importance for one conjunct, the attribute will have high importance for the conjunction

d) where an attribute of one concept is inconsistent with an attribute of the other concept, then a means of resolving the conflict is found, usually by dropping one or other of the attributes

e) new "emergent" attributes may find their way into the conjunction, either from accessing world knowledge (e.g. that pet fish live in tanks), or from attempts to improve the coherence of the new attribute set (e.g. that a blind lawyer is highly motivated – see Kunda, Miller and Claire 1990).

This model is based on empirical studies of how attributes are inherited in conjunctions with or without negation, and in disjunctions (Hampton 1987, 1988a, 1997). While importance for a conjunction was well approximated by an arithmetic average of the two constituent importances, it was also found to be subject to constraints, such that necessary attributes remain necessary, and impossible attributes remain impossible. One function that satisfies these constraints is defined by the formula

$$I(i, A \cap B) = \frac{I(i,A) \cdot I(i,B)}{I(i,A) \cdot I(i,B) + (1 - I(i,A)) \cdot (1 - I(i,B))}, \tag{8.4}$$

where $I(i, A)$ denotes the importance of attribute i for concept A. Clearly, $I(i, A \cap A) = 1$ when $I(i, A) = 1$ or $I(i, B) = 1$ and neither $I(i, A) = 0$ nor $I(i, B) = 0$. Similarly, $I(i, A \cap B) = 0$ when $I(i, A) = 0$ or $I(i, B) = 0$. Note also that if an attribute is necessary for one concept ($I(i, A) = 1$), but impossible for the other concept ($I(i, B) = 0$), then $I(i, A \cap B)$ is undefined, being zero divided by zero. Such a case would correspond to a conjunction with no members, the empty set.

8.4.3 Explaining the data

Intensional models such as Hampton's CPM provide a good account of the different non-logical effects reported in the literature. Overextension of conjunctions occurs because in forming a composite prototype typical attributes of each concept may be lost through the process of conflict resolution. A typical pet is warm, cuddly and affectionate, while a typical fish lives in a river, lake or ocean and is caught for food. Put these two concepts together and the result is a creature that is neither warm and affectionate, nor ocean living or eaten. The more central attributes of each concept have eliminated some of the less central attributes of the other concept, leaving a typical pet fish as a cold slimy water-living creature that lives in the home, has a name and an owner who cares for it and feeds it. Naturally enough, it is then very possible for some items (such as a guppy or goldfish) to be a much better fit to this new composite concept than they are to either of the original conjoined concept. A guppy is not a typical fish, or a typical pet, but it *is* a typical pet fish. Hampton's (1988b) discovery that not only are some items more *typical* of a conjunction than of each conjunct, but that they are also *more likely to belong* in the conjunctive class provides a strong endorsement of this account of concept combination.

There are of course intensional logics. They share with psychological intensional models the idea that concepts are concerned not just with the actual current world but also with the set of possible worlds. Modal notions of necessity and possibility are a part of intensional logic, just as they are a crucial element in psychological accounts of concept combination. In order, for example, for the composite prototype of a pet fish to be formed, it is not sufficient just to average out the importance of attributes for each component, but issues of compatibility, necessity and possibility also need to be addressed. Human thought is certainly not constrained to the here and now or even the actual. Our intensional concepts can be combined in ever more creative and imaginative ways to create fictional worlds of infinite possibility.

8.5 Fuzzy Logic and Conceptual Combinations

In the course of this chapter I have tried to lay out some of the complexities of the way in which concepts can be combined in everyday thought and language. On one hand, a continuous valued logic will be of key importance for understanding the psychological data, since there are indefinitely many situations in which the statements that we wish to make are neither completely true nor completely false (see Chapter 9). On the other hand a truth functional approach that seeks to find the "logical rules" by which the applicability or extension of complex concepts is determined by the extension of their parts and the syntax of their combination is unlikely to provide an answer to more than a very restricted range of human cognition. My approach has always been to collect empirical data first, and then to theorize in a way that attempts to account for those data. This approach has led to some surprising discoveries. Ways of speaking that appear to have a certain logical form turn out to work differently from the way we had supposed. But the approach has also led to some fruitful theorising that has brought together a range of phenomena under the general heading of "intuitive reasoning", to use Tversky and Kahneman's phrase.

References

Bonini, N., D. N. Osherson, R. Viale, and T. Williamson. 1998. On the psychology of vague predicates. *Mind and Language* 14: 377-393.

Chater, N., K. Lyon, and T. Myers. 1990. Why are conjunctive concepts overextended? *J. of Experimental Psychology: Learning, Memory, and Cognition* 16: 497-508.

Cohen, B. and G. L. Murphy. 1984. Models of concepts. *Cognitive Science* 8: 27-58.

Connolly, A., J. A. Fodor, L. R. Gleitman, and H. Gleitman. 2007. Why stereotypes don't even make good defaults. *Cognition* 103: 1-22.

Estes, Z. and S. Glucksberg. 2000. Interactive property attribution in concept combination. *Memory and Cognition* 28: 28-34.

Fillmore, C. J. 1968. The Case for Case. In *Universals in linguistic theory,* ed. E. Bach and R. T. Harms, 1-88. New York: Holt, Rinehart and Winston.

Fodor, J. A. and E. Lepore. 2002. *The compositionality papers*. Oxford: Clarendon Press.

Gagné, C. L. and E. J. Shoben. 1997. Influence of thematic relations on the comprehension of Modifier-Noun combinations. *J. of Experimental Psychology: Learning, Memory, and Cognition* 23: 71-87.

Gleitman, L. R. and H. Gleitman. 1970. *Phrase and paraphrase*. New York: Academic Press.

Hampton, J. A. 1979. Polymorphous concepts in semantic memory. *J. of Verbal Learning and Verbal Behavior* 18: 441-461.

Hampton, J. A. 1982. A demonstration of intransitivity in natural categories. *Cognition* 12: 151-164.

Hampton, J. A. 1987. Inheritance of attributes in natural concept conjunctions. *Memory and Cognition* 15: 55-71.

Hampton, J. A. 1988a. Disjunction of natural concepts. *Memory and Cognition* 16: 579-591.

Hampton, J. A. 1988b. Overextension of conjunctive concepts: Evidence for a Unitary Model of Concept Typicality and Class Inclusion. *J. of Experimental Psychology:*

Learning, Memory, and Cognition 14: 12-32.

Hampton, J. A. 1991. The combination of prototype concepts. In *The psychology of word meanings,* ed. P. J. Schwanenflugel, 91-116. Hillsdale, NJ: Erlbaum.

Hampton, J. A. 1995. Testing prototype theory of toncepts. *J. of Memory and Language* 34: 686-708.

Hampton, J. A. 1996. Conjunctions of visually based categories: Overextension and compensation. *J. of Experimental Psychology: Learning, Memory, and Cognition* 22: 378-396.

Hampton, J. A. 1997. Conceptual combination: conjunction and negation of natural concepts. *Memory and Cognition* 25: 888-909.

Hampton, J. A. 2007. Typicality, graded membership and vagueness. *Cognitive Science* 31: 355-383.

Jönsson, M. L. and J. A. Hampton. 2006. The inverse conjunction fallacy. *J. of Memory and Language* 55: 317-334.

Kamp, H. and B. Partee. 1995. Prototype theory and compositionality. *Cognition* 57: 129-191.

Kunda, Z., D. T. Miller, and T. Clare.1990. Combining social concepts: the role of causal reasoning. *Cognitive Science* 14: 551-578.

Lakoff, G. 1987. *Women, fire and dangerous things*. Chicago: University of Chicago Press.

Levi, J. 1978. *The syntax and semantics of complex nominals*. New York: Academic Press.

Machery, E., W. Hintzen, and M. Werning. 2011. *Handbook on Compositionality*.

Oxford: Oxford University Press.

Murphy, G. L. 1988. Comprehending complex concepts. *Cognitive Science* 12: 529-562.

Murphy, G. L. 1990. Noun phrase interpretation and conceptual combination. *J. of Memory and Language* 29: 259-288.

Murphy, G. L. 2002. *The big book of concepts*. Cambridge, MA: MIT Press.

Osherson, D. N. and E. E. Smith. 1981. On the adequacy of prototype theory as a theory of concepts. *Cognition* 11: 35-58.

Osherson, D. N. and E. E. Smith.1982. Gradedness and conceptual conjunction. *Cognition* 12: 299-318.

Osherson, D. N. and E. E. Smith.1997. On typicality and vagueness. *Cognition* 64: 189-206.

Rips, L. J. 1995. The current status of research on concept combination. *Mind and Language* 10: 72-104.

Rosch, E. R. and C. B. Mervis. 1975. Family resemblances: studies in the internal structure of categories. *Cognitive Psychology* 7: 573-605.

Smith, E. E. and D. N. Osherson. 1984. Conceptual combination with prototype concepts. *Cognitive Science* 8: 337-361.

Storms, G., P. De Boeck, I. van Mechelen, and D. Geeraerts. 1993. Dominance and non-commutativity effects on concept conjunctions: Extensional or intensional basis? *Memory and Cognition* 21: 752-762.

Storms, G., P. De Boeck, I. van Mechelen, and W. Ruts. 1996. The dominance effect in concept conjunctions: Generality and interaction aspects. *J. of Experimental Psychology: Learning, Memory, and Cognition* 22: 1-15.

Storms, G., W. Ruts, and A. Vandenbroucke. 1998. Dominance, overextensions, and the conjunction effect in different syntactic phrasings of concept conjunctions. *European J. of Cognitive Psychology* 10: 337-372.

Storms, G., P. De Boeck, J. A. Hampton, and I. van Mechelen. 1999. Predicting conjunction typicalities by component typicalities. *Psychonomic Bulletin and Review* 4: 677-684.

Tversky, A. and D. Kahneman. 1983. Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review* 90: 293-315.

Williamson, T. 1994. *Vagueness*. London: Routledge.

Wisniewski, E. J. 1997. When concepts combine. *Psychonomic Bulletin and Review* 4: 167-183.

Wright, C. 2001. On being in a quandary: Relativism, vagueness, logical revisionism. *Mind 110:* 45-98.