Two Computational Methods of Attributive Modification in Natural Language Semantics Compared

M.L. de Groot
6103677

Bachelor Thesis
Credits: 18 EC

Bachelor Opleiding Kunstmatige Intelligentie

University of Amsterdam
Faculty of Science
Science Park 904
1098 XH Amsterdam

Supervisor
Dr R. Blutner

Institute for Logic, Language and Computation
University of Amsterdam
P.O. Box 94242
1090 GE AMSTERDAM

January 18th, 2013

Abstract

In the field of linguistics, compositionality plays a big role. As Freges principle of compositionality stated: The meaning of a complete sentence must be explained in terms of the meanings of its constituent expressions. To understand the meaning of a sentence, all parts of the sentence must be merged to form one meaning. Attributive modification poses a subdomain of this problem and focusses on the meaning of a noun and adjective combined.

This thesis will compare two different methods on how to compute attributive modification in natural language. The first method is developed by Gärdenfors and states that in a semantic space, a correspondence occurs between two points and their position towards each other and their own boundaries of the corresponding concepts. The second method converts the semantic space to a vector space and uses the reduced tensor product to find the most prototypical composition of two vectors that form the concepts.

By aligning the two methods, a mathematical correspondence is found. Also, the conceptual differences are found that are relevant for modeling semantic spaces and concepts. Based on the mathematical correspondence and the conceptual differences, it is possible to choose one of the two methods based on the modeling of concepts within semantic spaces.

Keywords

Natural Language, Compositionality, Attributive Modification, Semantic Space, Meaning Space, Radial Projection, Reduced Tensor Product

Contents

1 Introduction

Linguistics in an important part of research within the field of artificial intelligence. Computers understand very little of the meaning of human language. This profoundly limits our ability to give instructions to computers, the ability of computers to explain their actions to us, and the ability of computers to analyse and process text [23].

A big problem within the computational linguistics is the problem of compositionality. Freges principle of compositionality states that the meaning of a complete sentence must be explained in terms of the meanings of its subsentential parts, including those of its singular terms [11]. This could also be explained in the following manner: the meaning of the whole is constructed from its parts, and the meaning of the parts is derived from the whole [12]. This concept is important within the field of artificial intelligence if we ever want the computer to understand the global meaning of entire sentences.

The problem of compositionality not only refers to the meaning of an entire sentence, but even the combination of an adjective and noun alone is hard to describe to a computer. The problem is that the noun, also spoken of as *head*, and the adjective, also known as *modifier*, both affect the meaning of the combination of the two words. For example, thinking of the colour *red*, one might think of the colour of a firetruck. But thinking of a *nose*, one assumes this concept has a skin colour. Then, when the terms are combined to *red nose*, excluding artificial clown noses, the colour we have in our minds immediately shifts to a reddish skin colour. This resulting colour however is (probably) not the most typical red colour, nor the most typical skin colour that springs to mind.

These cases, in which the head and the modifier together typically mean something else than the original terms did apart, are called *indirect composition*. However, it is also possible for the head and modifier not to affect each others meanings. These cases are called *direct composition* [24]. An example of these concepts could be a *yellow circle*, where the term *yellow* and *circle* are totally independent, since both terms do not have an overlapping meaning.

This thesis will cover two methods on how to overcome the problem of compositionality. The first makes use of *radial projection* [24] and the second applies a *reduced tensor product* [12].

The first method was developed by Gärdenfors, proposing an account of semantics as a mapping between individual meaning spaces based on the "meeting of minds" [24]. According to this view, people possess an inner world which can be modeled as a conceptual space. When we communicate, we try to align our representations of the world, to converge our worlds towards a smaller set of possible worlds. The moment we achieve, or at least when we *believe* to have achieved, mutual understanding, is when we reach a fixpoint in communication. This fixpoint represents the equilibrium in the communicative intent of the speaker and the understanding of the receiver; the result of an interactive, social process of meaning construction and evaluation.

One of the essential assumptions of this theory is the assumption of convexity. Based on earlier work of Gärdenfors himself, he assumes inner concepts are

represented as convex regions of mental spaces. Based on this mathematical propery, Gärdenfors and Warglien use radial projection to establish homeomorphism between two concepts, both represented by a convex set. Or in their own words: "Provided that the head and modifier spaces are compact and convex regions of metric spaces, a way always exists of re-scaling the distances in the modifier space to fit the constraints of the head space in a one-to-one correspondence."

There are two applications of this theory about the mapping of meaning spaces. The first applies to communication, the meeting of minds, where two mental representations of the same utterance align to reach a fixpoint in communication. The second is when we apply this theory to compositionality, where the combination of a head and modifier forms one concept.

The second method is based on the prototype theory. The basic idea of the prototype theory is that some members of a category are more central than others [23]. For example, a *blackbird* is a more typical instance of the concept *bird* than a *penguin*. So when a function of typicality assignes values to members of concepts, a blackbird will have a higher degree of 'birdiness' than a penguin [13]. In short, instances have a degree of membership to a concept.

In cognitive science, the prototype theory often makes use of vectors, providing a natural way for ordering degrees of membership for multiple instances [23]. Next, when we have two concepts modeled by typicality-vectors, mathematical functions can be applied as means to bind the two vectors. This method proposes to use the reduced tensor product to find the combined vector. Based on the two concepts, this combined vector will have combined degrees of membership for every instance, leading to the most prototypical combination of the two concepts.

In order to develop methods to solve the problem of compositionality, some sort of model is needed to represent the meaning of words. To compare the two methods in this thesis, both methods are explained seperately and colour models are used as a meaning space. Subsequently, the two methods will be compared to each other and these results will be discussed.

2 Composing by Radial Projection

This method of compositionality is based on Gärdenfors view on 'the meeting of minds,' where thoughts are shaped as convex sets [24]. When we, then, engage in conversation, we must align our thoughts and find the right point in the intersection of our individual thoughts. The point on which we then agree to have found a mutual understanding, is the fixpoint of our communication.

This problem of trying to achieve a mutual agreement between two spaces, also exists in language. When a modifier and head are merged, they have to find a combined meaning as well. For instance, the word *red* implies a red colour and the most typical one would perhaps be firetruck-red. However a word like *nose* implies a skin colour, so when the words are combined to *red nose*, suddenly a reddish skin colour is implied, which, usually, would not be firetruck-red.

To solve this problem of agreement between two spaces, Gärdenfors proposes that these meaning spaces are convex. This assumption is based on empirical findings of Jäger, who revealed that this can be the result of evolutionary dynamics of communicative strategies [9]. Moreover, the convexity assumption also gives a solid foundation for mathematical functions that allow us to map points existing in only one of the two sets to the intersection.

This section will start by addressing the assumptions of Gärdenfors' method, followed by a description of the method itself. Next, to illustrate the advantages and disadvantages of the theory, the method is applied to colour spaces, first in the RGB colour space, because this model is widely used, and then in the HSL model, to render a perhaps more intuitive result.

2.1 Assumptions

Gärdenfors makes certain assumptions in his approach:

- Mental spaces. First of all, Gärdenfors assumes people have a certain cognitive representation of the world, which can be modeled as a mental space.
- Convexity. The second assumption entails that not only we have mental spaces, but that these spaces are convex as well.
- Euclidean space. The last important feature of this method is that the calculations take place in Euclidean space, making straight lines the shortest distance between two points.

2.2 Method

Based on the assumptions mentioned in the previous section, and on the convexity assumption in particular, Gärdenfors solves both direct and indirect composition by using mathematical functions to map two sets onto each other.

Gärdenfors proposes that in the case of direct composition, where two associated domains are entirely independent - i.e. where there is no intersection -, a Cartesian product of the two spaces can be used. Or in the words of Gärdenfors:

"The meaning of *blue rectangle* is defined as the Cartesian product of the *blue* region of colour space and the *rectangle* region of shape space." This means that the meaning of the concept *blue rectangle* can be the result of every possible pairing of the term *blue* and the term *rectangle*, which can be calculated by using the Cartesian product of the two spaces. An important notion of this mathematical function is that the product of any compact and convex sets will again be a compact and convex set and thus preserves the structural properties of the conceptual spaces.

However, direct composition is rare, in most cases the space associated with the head affects the representation of the modifier. For instance, a white wine is not really white and a red wine does not have the colour of prototypical red.

In this case, where we have to find the overlapping space in which both the modifier and head are present, it is necessary to use radial projection. Radial projection is a mathematical procedure that takes a point in one of the two sets and calculates a representative replacement for this point within the intersection and as long as two sets are convex, compact *and* have a common interior point (the origo), homeomorphism between the two sets can always be achieved through radial projection [1]. In other words, as long as two convex sets share an intersection, one can always apply radial projection to linearly scale the sets onto the intersection of the two sets.

For instance, the combination of words in *red hair*; the term *red* implies all red colours of the colour space and the term *hair* only gives all possible (natural) haircolours. Next, a typical red colour is chosen and is projected onto the intersection with the possible haircolours. This is done by making a line from the intersection to the typical red colour, which gives crossings with the boundaries of both sets. In turn, the ratio of the distances between these four points render the most representative point for both sets within the intersection, ergo, this radial projection delivers the best match between the typical red colour and the possible natural haircolours.

In short, radial projection provides contextual re-scaling effects between two (partially) overlapping convex sets.

2.3 Radial Projection Applied to Colour Spaces

Physically, colour is derived from the perception of light, which is emitted in wavelengths. These wavelengths are measured in nanometers, where humans only percieve the wavelengths within the range of 400 nm to 700 nm. According to the theory of trichromacy, the human eye contains three types of cones that are sensitive to different wavelengths of light. These three cones are ordered by the wavelengths of the peaks of their spectral sensitivity: short, medium and long, which approximately correspond to blue, green and red lights [3] (shown in figure 1).

The first major distinction between colour spaces is device dependency. When the colour model is device independent, the colour coordinates for the same colour are the same for every output medium. The RGB space is device dependent, since the space is defined by specific primary colours, which are not the same on every device [15].

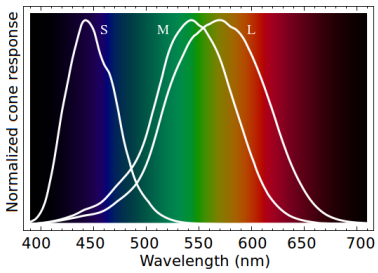


Figure 1: The normalized response spectra of human cones.

Another distinction is the intuitiveness of the colour spaces, the measurement of whether the global distances between the colours in the model actually correspond to the psychological distances between the colours. One of the more intuitive colour models is said to be the HSL colour model, because it is relatively easy to relate to the difference between colours in hue, saturation or brightness, making the local distances realistic.

Since compositionality is a computational problem, the RGB and the HSL colour models are chosen to apply Gärdenfors' method on. First the RGB model is demonstrated as this model is widely used in computers and therefore an important model in which the theory of radial projection should work. Next, the method is applied to the HSL colour space for this space is more intuitive.

2.3.1 The RGB Colour Space

The RGB colour space is defined by three primaries, red, green and blue. Each of these quantities indicate how much of this individual colour is included in the composite colour, or, in other words, the RGB colour model is an additive model. The value of each quantity is a natural number from 0 to 255. When all variables are set to zero, the outcome is black and when all quantities are set to 255, it will give a white result.

The RGB model also fulfills the principle of trichromacy, which means that all colours in this model are linear combinations of the three variables. The graphical model of the RGB colour space is shown in figure 2.

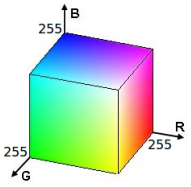


Figure 2: The RGB colour space.

Though the RGB colour space is a three-dimensional space and the method of radial projection works for n-dimensions, purely to keep this example easy to illustrate, the B-value of the RGB space is set to zero. This simplification allows us to plot a two-dimensional image as shown in figure 3.

When we apply radial projection, we need two convex sets. In this case one convex set is the entire space shown in figure 3, the second set is illustrated by the black square. The next step is to take an arbitrary *origo* in the intersection of the two convex sets, which we call *0*. Then, we select a point that we think is a typical green, for example, point x. To map this typical greenish colour to the intersection, we draw a straight line between the origo and this typical colour. This line crosses two boundaries, the one with the subset, where the crossing is called y_0 and the intersection with the outer bound, called x_0.

Consequentially, radial projection shows a correspondence between the distance from the origo to the inner boundary and the distance from the origo to the outer boundary, which gives the following formula:

$$\frac{d(O, y_0)}{d(O, y)} = \frac{d(O, x_0)}{d(O, x)} \tag{1}$$

leading us to point y.

In other words; when two convex sets have a common point, we map one convex set to another by scaling them linearly to fit one another. So when we have a point outside the intersection (x) and we map it to the inside of the intersection, we get the representative inner point (y) to replace the outlier x.

For this particular example, it would mean that the green colour, referred to by point x, with RGB-values $[20, 230, 0]$, would be represented by point y. This new colour has the following RGB-values: $[100, 160, 0]$. When we look at this colours side by side, they are not the same shade, but they are definitely both representing a green colour. If we were to choose a typical green colour from the inner convex set, it would even be sensible to choose the y-colour.

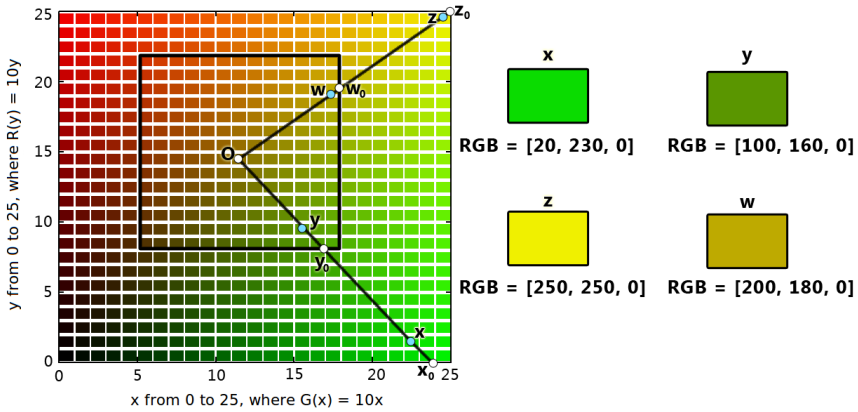


Figure 3: The RGB colour space, where $B = 0$, the origo is in the middle of the subspace, x represents a typical green colour and z a typical yellow colour. On the right of the plot, the first line shows the colours represented by x and y, respectively and the bottom two squares show the two colours represented by z and w, respectively.

However, there are a couple of sidenotes to address. First of all, the blue colour is missing in this example, by setting the B to zero. This might have led

to a different result; maybe the green colour represented by x would not be a typical green colour if we could add some blue.

Secondly, executing this method, choosing the origo is arbitrary, yet this does influence the outcome. If we choose the origo at a different location, the correspondences between the boundaries would differ and the gradient as well. This issue is illustrated by figure 4.

Still, both the outcome of figure 3 as well as the result of this second example are green. However, in this second example, the outcome is lighter and more orange than the first result.

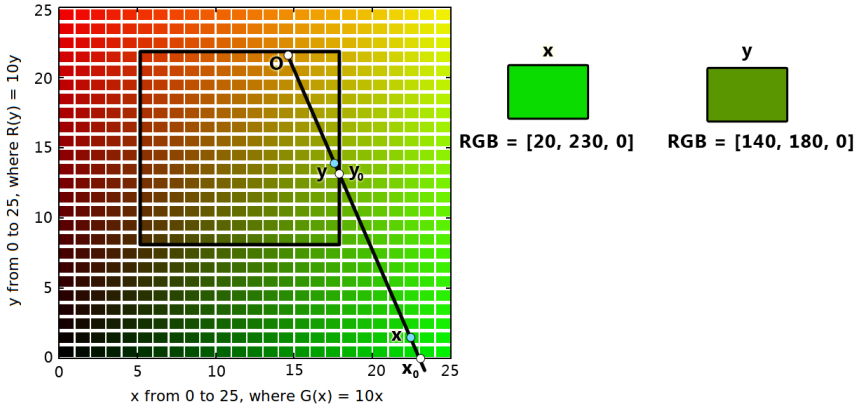


Figure 4: On the left the RGB colour space, where $B = 0$. On the right, the colours represented by x and y, respectively.

Finally, the distances between the colours in the RGB colour model are not psychologically intuitive. In the RGB space the distances are differences in the amounts of red, green or blue. To the naked eye, however, these distances do not seem natural since not all distances correspond to the same amount of difference between the colours, which is illustrated in figure 5, where the two colours on the right seem perceptually closer to one another than the two colours on the left even though the measured distance is the same.

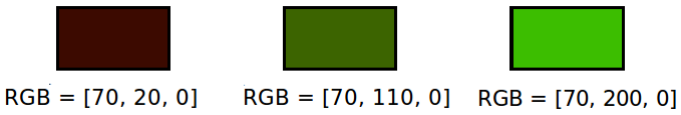


Figure 5: These three RGB colours differ only in the amount of green, where from left to right each colour has an added value of 90.

Nonetheless, it would be natural to say that, in both examples of radial projection, the colour represented in point y might have more similarity to the colour of the origo, than the colour represented by x. But it would not intuitively be said that that precise distance between the *origo*-colour and the y-colour in figure 3 is nine steps. Nor would people in general say that the x- and y-colour where the same colour, only mapped on a different space.

2.3.2 The HSL Colour Space

The abbreviation of HSL stands for hue, saturation and lumination. The hue is the value to which the colour seems similar to one, or to proportions of two other colours [3]. In other words, the hue of the colour yellow would be closer to the hue of orange than to the hue of blue since yellow and orange look more similar than yellow and blue. The saturation is the degree to which the hue differs from a neutral gray, running from 0% to 100%. If the saturation is lower, the colour will be closer to gray. Lastly, the lumination indicates the brightness of the colour, where at 0% the colour is completely black, 50% gives a pure colour and turns white at 100%.

The HSL colour space is also referred to as the *colour spindle*, where the radius is defined by the value of saturation. The height depends on the brightness of the colour and the rotation around the spindle depends on the hue. This is shown in figure 6.

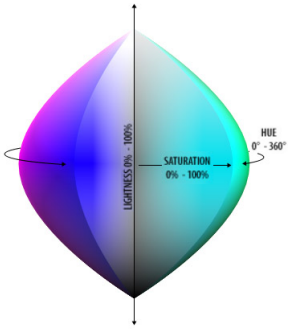


Figure 6: The graphical representation of the HSL colour model.

To easily illustrate this example, the value of the lumination is set to a half. Again two convex sets are needed, this time two circles. Then, when we take the centre of the inner circle as origo and a colour in the colour space, the resulting colour will always be a less saturated version of the same colour, because every line from the centre of both circles will always make a radius, which represents the saturation of the HSL colour space (figure 7). Furthermore, when a less saturated yellow than the displayed yellow in figure 7 would be chosen as starting point, the result would be even less saturated than the outcome shown in the picture.

So looking from the origo of the circle, this method seems intuitive; when we select a bright yellow and map this to grayisch colours, we end up with a grayish yellow. Also, when we were to choose a less bright yellow, we would finish with an even grayer form of yellow.

Nonetheless, despite of this intuitiveness, this colour model too, faces the problem of the origo. Since the origo is not a specified point, not all outcomes are logical. This problem is illustrated by figure 8, where the origo is not in the centre of both circles, which directly leads to a very different result.

Concisely, in the case of two circles of which one is placed exactly in de middle of the other circle and where the origo is chosen in the middle of both

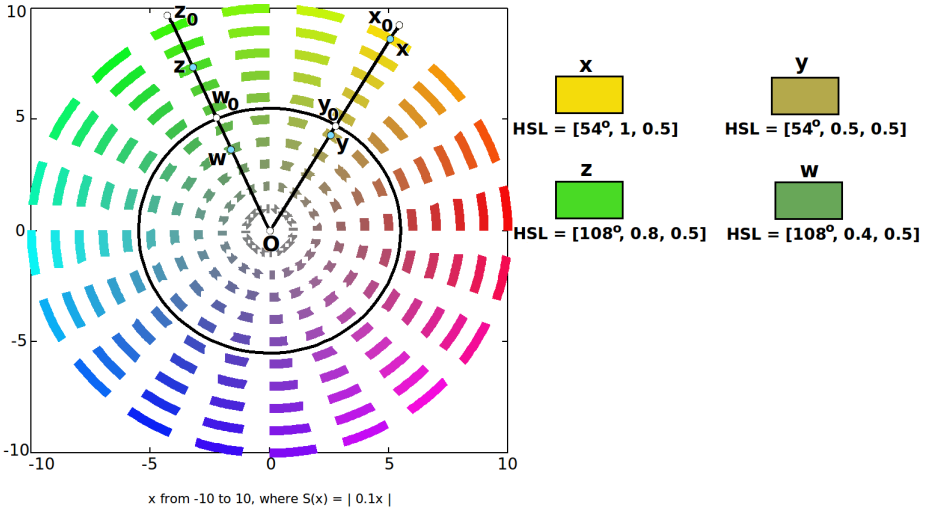


Figure 7: On the left the HSL colour space, where $L = 0.5$, x represents a typical yellow colour and z represents a typical green colour. On the right, on top the colours represented by x and y, respectively and below the two colours represented by z and w, respectively.

circles, the radius will give the same colour with less saturation, every other line will directly change the slope and move the representative inlier to a different colour, since the colours, which we believe are different from each other, are arranged by hue.

Furthermore, the other two objections mentioned in the previous section also withstand this colour model. This model too, is incomplete by fixing the lumination and lastly, although this model is more intuitive than the RGB colour space, these distances remain psychologically unintuitive. Even though it is intuitive to distinguish between colours based on hue, saturation or brightness, making the local distances natural, globally, the distances are still unintuitive. This problem is illustrated by figure 9, in which all colours differ the same amount in hue. However, the two colours on the right seem more similar than the two colours on the left even though the measured distance is equal to one another.

Also, in the HSL colour model, it would be natural to have a special distance metric since the difference in hue has more impact on the changes between the colours than the saturation. In figure 10 the first and the second colour differ two steps in hue and the second and last colour differ two steps in saturation. In this picture it is clearly shown that the first and second colour have a perceptually greater difference than the second and the third.

2.4 Points of Discussion

For both colour models the same three problems seem to arise; the fixing of one value, the psychologically unintuitive distances and the choosing of the origo.

First of all, both colour models are incomplete since in both cases one of the

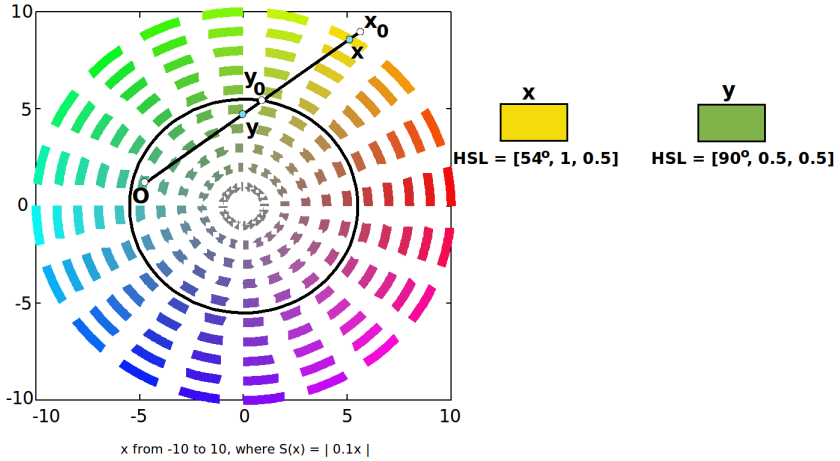


Figure 8: On the left the HSL colour space, where $L = 0.5$. On the right, the colours represented by x and y, respectively.

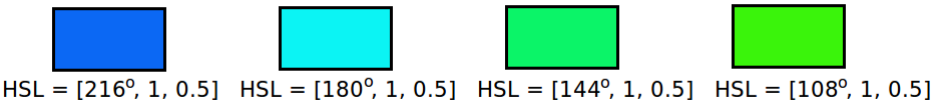


Figure 9: Four colours from the HSL space, from left to right all 36 degrees apart in hue.

three values has been fixed. However, the method should provide a solid theory for all combinations of two (partially) overlapping convex sets, so this should also be true for two convex sets that happen to lay in a two-dimensional plane of the colour space.

The next point of critique is the unintuitiveness of the colour models, which is a well-known problem for psychologists. This has led to more intuitive models like the Munsell colour space, where distances between colours correspond to real perception thresholds. Perhaps the theory would provide better results for different colour spaces. Nevertheless, the RGB colour model is relevant as the problem of compositionality is important in computer sciences. The RGB colour space is widely used in computers and therefore an important model in which the theory of radial projection should work.

Also, even though Gärdenfors' method is only tested on colour spaces, the theory should work for *all* conceptual spaces. However, the fact that colour spaces are not perfectly suitable for this theory, does not mean that the theory in general is faulty for every other conceptual space.

Finally, it might not even be possible to find a globally realistic solution using the Euclidian metric and maybe a different metric, like Riemannian metric, would provide more intuitive results.

Lastly, the most important point of improvement this theory needs is the choosing of the origo and the prototypical colour. When radial projection is

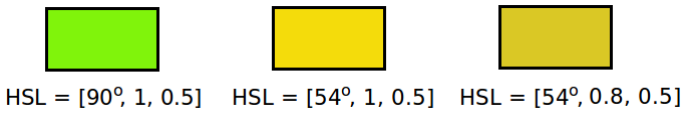


Figure 10: Three colours from the HSL space, where the first differs two steps in hue (equal to 36 degrees) from the second, which in turn differs two steps in saturation (0.2) to the third colour.

applied to colour spaces, the two points of reference are arbitrarily chosen.

For colour spaces it is probably the best solution to have the origo in the centre of the intersection, this will always generate unbiased results. Or, it might be a future expansion of the method to try all possible origos and then take the average over all outcomes. But even so, the point of the typical colour also remains in the eye of the beholder. For example, figure 11, where an *apple* could have all colours within the gray triangle and the colour *red* could be represented by all red colours within the blue triangle. Then, taking a *red apple* results in the intersection, where the origo is taken in the middle. Then there are two typical reds chosen, which both give different results, as shown in the figure, and strangely the darker shade of typical red gives a lighter shade of red for the apple than the other typical red does. This shows that radial projection is very subjective and therefore not a general method.

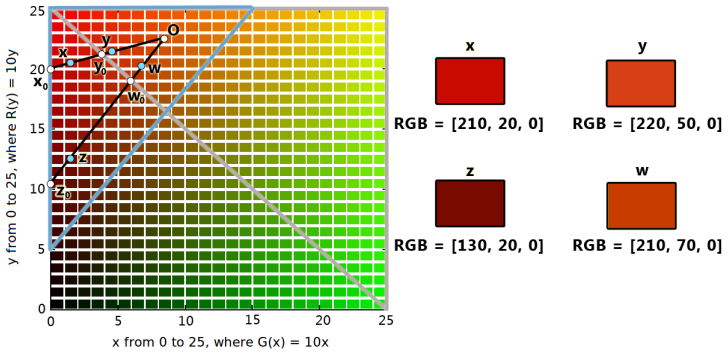


Figure 11: The red apple example. The gray triangle represents all possible apple colours and the blue triangle represents all red colours. Then the origo is set in the middle of the intersection and x and z represent two different typical red colours, after applying radial projection, resulting in the points y and w, respectively.

3 Composing with the Reduced Tensor Product

The theory of the reduced tensor product is based on the vector representation of concepts. Within this framework, concept similarity is a function of distance, where semantically related concepts are closer to one another than semantically less related words, which is also shown in figure 12 [11]. For example, the concept *mammal* will be more similar to *creature* than *lamp*.

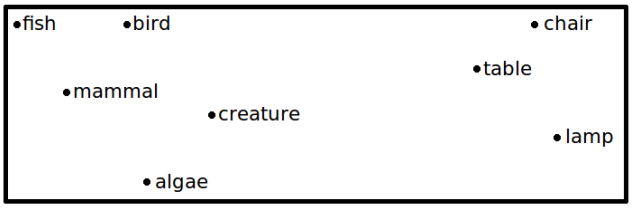


Figure 12: An example of semantic space, where similar concepts are closer than less similar concepts.

On a more detailed level, every concept is represented by its own vector of which every component represents the degree to which that component is representative to the concept, which is also called the cue validity [2]. More explicitely, the vector contains as many components as there are concepts in the represented world and the more similar the component is to the concept, the higher the value of that specific component. This vector is based on the prototype theory, where an object instances a concept only to the extent that it is similar to the prototype of the concept [13].

For instance, if our world contains the concepts *red*, *blue* and *green*, the concept *green* would be a vector with three components, namely the three colours, where the most similar colour would have the highest value. So the vector of the concept *green*, could have a vector like: $[0.3, 0.7, 1]$, where red is 0.3, and blue 0.7 representative to the concept *green*.

This factor of representiveness of every component is based on the distance between the component and the most prototypical instance of the concept. So, within a semantic space, mathematically speaking, a distribution is created, where the peak of the curve lies on top of the most prototypical concept and the curve fans out in every direction.

Next, these vectors can be used to calculate combinations of concepts. When two concepts are combined, the reduced tensor product is applied, resulting in a third vector, having the highest value for the most prototypical instance for the combination of the two concepts.

Originally, the tensor product representation was used by Smolensky for the connectionist representation of variable bindings [20]. However, the tensor product provided a growing dimensionality, as when a n-dimensional vector and a m-dimensional vector are combined through this direct outer product, the result is a nm-dimensional vector. To avoid this problem, Plate proposed to use holographic reduced representations, which uses circular convolution to associate items [16]. Unfortunately, this mathematical operation is, unlike our application in linguistics - a semantic space -, translation invariant. Several researchers have also proposed simple multiplication operations (cf. [2],[11]),

which is related to earlier proposals by Zadeh and Hajek in fuzzy logic where the relevant operation is called the product t-norm ([7], [25]).

The next section will describe the properties that are needed to use the reduced tensor product. Next, the method will be explained with the proper mathematical background. Also an elaboration on the term *reduced tensor product* will be given in this part.

3.1 Assumptions

There are a couple of properties this method requires:

- Semantic space. This method is applied to a semantic space, where semantically related concepts are more nearby than less related concepts.
- (Proto)typicality. The method is based on the prototype theory, which assigns a degree of typicality for every instance to a concept based on the similarity between the instance and the prototype.
- Euclidean space. All calculations take place in the Euclidian space.

3.2 Method

This account of semantics provides a framework for representing the meaning of word combinations in vector space, where points that are close together in this space are semantically similar and points that are far apart are semantically distant [12][23].

Mathematically, the typicality of instance i within concept A, written as $c_i(A)$, can be calculated using the following distance function:

$$c_i(A) = ne^{-\frac{1}{T}d(P_A,i)^k} \tag{2}$$

where the distance function d gives the Euclidean distance between the instance i and the prototype P_A of concept A. The prototype P_A also is one of the instances of A, resulting in the peak of the distribution. The terms T and k are used to scale the distribution, towards a clearer peak in the resulting curve; these two variables will be explained in depth in the next section. This calculation of the cue validity is used for every instance, resulting in a vector or matrix, depending on the meaning space. Finally, the entire concept is normalized, which is indicated by the term n.

When this formula is applied to two concepts, the resulting plot contains two peaks, which represent the two prototypes and both concepts have their own density distribution (figure 13). The x-axis represents the instances in the world, ordered semantically. The y-axis gives the degree of typicallity the instance holds.

Next, to combine two concepts, the two distributions are merged to a new distribution. In other words, the reduced tensor product is applied as means of binding one distribution to another to produce structured representations [12]:

$$(A \otimes B)_i = A_i B_i \tag{3}$$

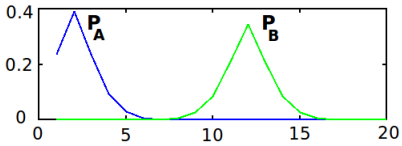


Figure 13: An example of vector space, where concept A has its prototype at (2), shown in blue, and the other concept, B at (12), shown in green. Also, the variables are set: $T = 2$ and $k = 1.5$.

However, the dimension of the tensor product is the product of the dimensions of the original spaces, so, in the case of the two vectors the resulting product gives a matrix. Because a growing dimensionality is not welcome, a *reduced* tensor product is taken. This means we do not use the entire resulting matrix, but only the diagonal of this matrix. This has the advantage that the dimensionality will stay equal. Then, this new vector will be normalized as well. This resulting vector also shows a peak, which represents the prototypical combination of the two previous prototypes, which we will call the *target*. When the two distributions shown in figure 13 are combined through this reduced tensor product, the following distribution arises (figure 14).

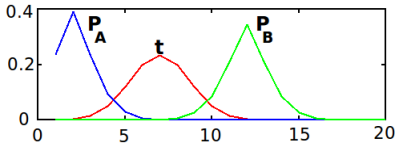


Figure 14: An example of vector space, where concept A has its prototype at (2) (blue) and the other concept, B, at (12) (green). Also, the variables are set: $T = 2$ and $k = 1.5$. The reduced tensor product then gives a third distribution, shown in red, which gives a prototypical combination, t, at (7).

3.3 Mathematical Background

3.3.1 Three Aligning Peaks

A very important property of this method is that the three peaks of the distibutions are aligned (figure 15).

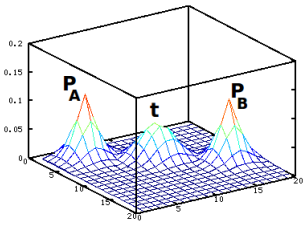


Figure 15: The vector space, where concept A has its prototype at (3,5) and the other concept at (10,15), $T = 2$ and $k = 1.5$. The reduced tensor product then gives a third distribution, of which the maximum is aligned with the two prototypes and indicated by the letter t.

When a line is drawn between the two prototypes and we were to look for the local maximum on this line, the point would have to lie in between the two prototypes, since the highest correspondence of both distributions lies there based on Euclidean distances. Next, if we were to deviate from this line, the distance to the prototypes will only grow bigger due to the triangle inequality, making the cue validity smaller and smaller. Therewith concluding that the local maximum on the line actually is the global maximum, because every movement off the line will only shrink the value of the cue validity.

Nevertheless, this feature does rely on a couple of properties of the distributions. The first property is the Euclidean geometry, where the shortest distance between two points is a straight line. Secondly, both distributions have to be symmetrical and monotonically decreasing. Fulfilling these properties cause the combined distribution to be on a straight line between the two prototypes of the concepts.

3.3.2 The Variables Explained

In the formula for calculating the distributions, two variables occur, namely T and k. The T-variable is a scalar, when the value increases the distribution will spread out more (figure 16).

In other words, the variable T defines the decay of the distribution. In effect, the higher the value of variable T, the lower the peak of the distribution. So when the value of T differs between concept A and B, the distributions will differ in height and decay, and the target will move towards the distribution with the higest decay (figure 17). However, it should be noted that T needs to have a positive value to conserve the properties of the distribution function.

The second variable used is the variable k, which is needed to make a distribution with a peak, which is necessary to point out the combined prototype. When k is set to one, the third distribution will be a constant instead of a distribution with a peak. Mathematically, the following formula is used when $k = 1$:

$$(A \otimes B)_{i,i} = n_A * n_B * e^{-d(P_A,i)} * e^{-d(P_B,i)} \tag{4}$$

where there are two concepts, A and B, which both have their own prototypes, P_A and P_B, respectively. Then, for every instance i the product of the normalization factors and the exponential distance functions are taken.

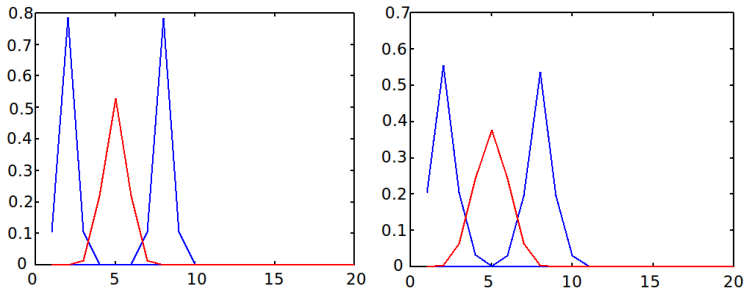
This function can be rewritten as:

$$(A \otimes B)_{i,i} = n_A * n_B * e^{-d(P_A,i)-d(P_B,i)} \tag{5}$$

However, since the distance between the two prototypes P_A and P_B is constant, the movement of i will only move the ratio between the two prototypes, but will not affect the outcome:

$$-d(P_A, i) - d(P_B, i) = -(i - P_A) - (P_B - i) = P_A - P_B \tag{6}$$

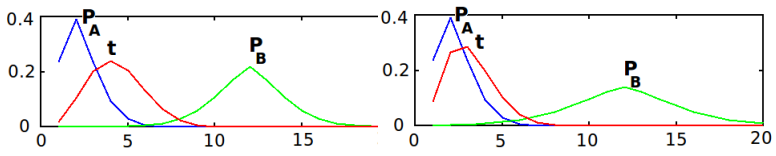
This problem can be solved by adding a power to which the distances are taken. This causes a difference between the distance from i to P_A and the distance from i to P_B, which makes for a distribution with a peak (figure 18). Also, if k differs between the two distributions, the target will move towards



(a) $T = 0.5$ (b) $T = 1$

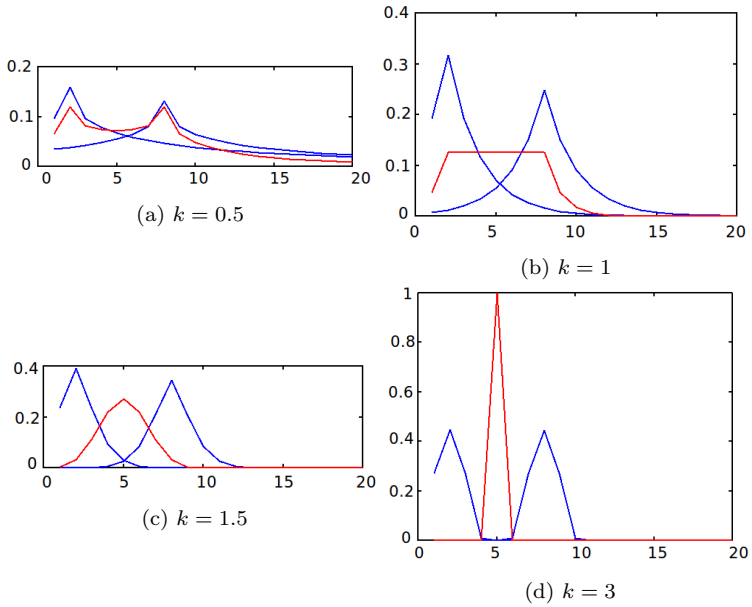
(c) $T = 2$ (d) $T = 5$

Figure 16: Examples of the vector space, where one concept has its prototype at (2) and the other concept at (8), $k = 1.5$ and T variates. The reduced tensor product then gives a third distribution, shown in red, which gives a prototypical combination at (5).



(a) $T_A = 2$ and $T_B = 4$. (b) $T_A = 2$ and $T_B = 8$.

Figure 17: Examples of the vector space, where concept A has its prototype at (2) (blue) and the other concept, B, at (12) (green). Also, the variables are set: k is equal for both distributions and set to 1.5, the value of T differs. The reduced tensor product then gives a third distribution, shown in red, which gives a prototypical combination, with its peak at t.



(a) $k = 0.5$ (b) $k = 1$ (c) $k = 1.5$ (d) $k = 3$

Figure 18: Examples of the vector space, where one concept has its prototype at (2) and the other concept at (8), $T = 2$ and k variates. The reduced tensor product then gives a third distribution, shown in red, which should give a prototypical combination at (5).

the distribution with the highest value of k, since the placement of the target depends on the difference between the distributions and the distribution with the highest power will have the highest outcome (figure 19). Furthermore, k should have a positive value as well. When k is negative, the distribution is upside down, meaning the cue validity of the prototypes will become zero, instances with normally the highest values suddenly have the lowest values and vice versa, resulting in a valley instead of a peak.

Another property of the two variables is that if T and k are equal for both distributions, the peak of the combined prototype will always lie in the middle of the two peaks. This can be explained with the following equations, where t stands for target, the combined prototype:

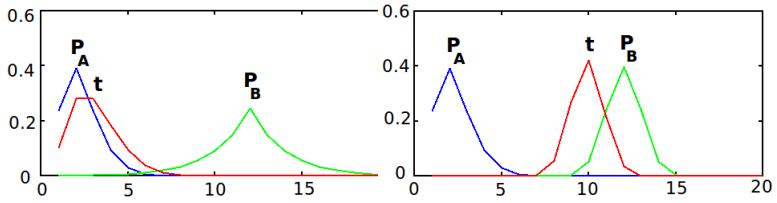
$$t = max_i(n_A * n_B * e^{-\frac{1}{T}d(P_A,i)^k - \frac{1}{T}d(P_B,i)^k}) \tag{7}$$

As discussed in the previous section, the three peaks are aligned, so P_A can be used as origin of the line, which entails that P_A can be replaced by zero and the distances can be rewritten as i and $P_B - i$ (figure 20).

$$t = max_i(n_A * n_B * e^{-\frac{1}{T}i^k - \frac{1}{T}(P_B - i)^k}) \tag{8}$$

To find the value of t, the derivative of this formula should be equal to zero. The instance i for which this formula is zero, will be the i with the maximum value, named target t:

$$\frac{\delta}{\delta t}(n_A * n_B * e^{-\frac{1}{T}t^k - \frac{1}{T}(P_B - t)^k}) = 0 \tag{9}$$



(a) $k_A = 1.5$ and $k_B = 1$. (b) $k_A = 1.5$ and $k_B = 2$.

Figure 19: Examples of the vector space, where concept A has its prototype at (2) (blue) and the other concept, B, at (8) (green). Also, the variables are set: T is equal for both distributions and set to 2, the value of k differs. The reduced tensor product then gives a third distribution, shown in red, which gives a prototypical combination, with its peak at t.

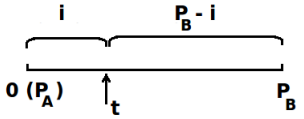


Figure 20: The graphical representation of the line between P_A and P_B.

$$e^{-\frac{1}{T}t^k - \frac{1}{T}(P_B - t)^k} * (-\frac{k}{T}t^{k-1} + \frac{k}{T}(P_B - t)^{k-1}) = 0 \tag{10}$$

Since the exponential part of the formula can not be zero, the second part of the equation needs to be equal to zero:

$$-\frac{k}{T}t^{k-1} + \frac{k}{T}(P_B - t)^{k-1} = 0 \tag{11}$$

These derivations lead to the following equations:

$$\frac{k}{T}t^{k-1} = \frac{k}{T}(P_B - t)^{k-1} \tag{12}$$

$$2t = P_B \equiv t = \frac{1}{2}P_B \tag{13}$$

This last equation tells us that, if the variables T and k are the same for both distributions, the distance from the prototype of concept A to the target is half the distance from the prototype of concept A to the prototype of concept B. So the target lies exactly in the middle of the line between P_A and P_B.

3.3.3 Variating the Variables

The previous section showed that the target always lies in the middle of the line between the two prototypes. However, this only happens when the variables T and k are equal for both distributions.

To broaden the scope of the reduced tensor product, it is more general to rewrite the formula:

$$t = \alpha P_B \tag{14}$$

Where α represents the ratio between the two concepts, causing the target to move from or to one of the two concepts. Right now, the ratio is fixed to

the middle of the line, causing α to be equal to a half. But if the variable T differs between the two prototypes, the target would still align with the two prototypes, but would also move towards the concept with the highest decay (the lowest value of T). The derivation of α can be calculated starting with the formula for t, also defined by equation (8), but now the values of T_A and T_B are defined seperately:

$$t = max_i(n_A * n_B * e^{-\frac{1}{T_A}i^k - \frac{1}{T_B}(P_B - i)^k}) \tag{15}$$

Then, the derivative of the formula should be set to zero to find the maximum value for i, namely the target:

$$e^{-\frac{1}{T_A}t^k - \frac{1}{T_B}(P_B - t)^k} * (-\frac{k}{T_A}t^{k-1} + \frac{k}{T_B}(P_B - t)^{k-1}) = 0 \tag{16}$$

Since the exponential expression at the beginning of the formula can not be equal to zero, the second part of the formula must be zero in order to solve the equation:

$$-\frac{k}{T_A}t^{k-1} + \frac{k}{T_B}(P_B - t)^{k-1} = 0 \tag{17}$$

This simplification leads to the following equation:

$$\frac{1}{T_A}t^{k-1} = \frac{1}{T_B}(P_B - t)^{k-1} \tag{18}$$

Next, the target can be isolated in the formula:

$$t = \frac{P_B}{1 + (\frac{T_B}{T_A})^{\frac{1}{k-1}}} \tag{19}$$

This equation gives us the formula for α, depending on the values of T_A, T_B and k:

$$\alpha = \frac{1}{1 + (\frac{T_B}{T_A})^{\frac{1}{k-1}}} \tag{20}$$

So if the value of T differs between the two distributions, the reduced tensor product still provides a linear function. Furthermore, if the value of T_A were to grow and the value of T_B would remain the same, the value of the denominator of α would have its limit at one and the target will end up at P_B. Because: the higher the value of T_A, the lower the decay of distribution A and the target will move towards the prototype of concept B. In short, the value of α depends on the ratio between T_A and T_B and when T_A grows to infinity or when T_B shrinks towards zero, the limit of α is one and when T_A shrinks towards zero or when T_B grows to infinity, the limit of α is zero. This means α will always have a value between zero and one, where zero represents the point of the prototype of concept A and one is equal to the position of the prototype of concept B. Or in other words, when T differs between the two distributions, the decay of the distributions differs, causing the target to move towards the distribution with the highest decay - i.e. the lowest value of T.

Another possibility is to differ the value of k instead of T. To make it easier to read, the k of the distribution of concept A will be called k and the power of concept B will be called m:

$$t = max_i(n_A * n_B * e^{-\frac{1}{T}i^k - \frac{1}{T}(P_B - i)^m}) \tag{21}$$

Then, to find the maximum value of t, its derivative has to be equal to zero:

$$e^{-\frac{1}{T}t^k - \frac{1}{T}(P_B - t)^m} * (-\frac{k}{T}t^{k-1} + \frac{m}{T}(P_B - t)^{m-1}) = 0 \tag{22}$$

Similar to the previous calculations, only the second part needs to be equal to zero:

$$-\frac{k}{T}t^{k-1} + \frac{m}{T}(P_B - t)^{m-1} = 0 \tag{23}$$

However, unlike the previous calculations for a differing T, the equations for differing the value of k lead to a complicated formula:

$$(\frac{k}{m})^{\frac{1}{m-1}} t^{\frac{k-1}{m-1}} + t = P_B \tag{24}$$

Unfortunately this equation can not be solved in a general matter. So, in summation, differing T between the two distribution gives a linear function that can be compared to the method of radial projection. However, differing the value of k between the two methods leads to an unsolvable formula and is therefore unwanted for the comparison of the two methods.

3.4 Points of Discussion

A couple sidenotes should be addressed using this method. Generally, there are two main problems: the initialization and the intuitiveness of the method.

Before the method can be executed, the variables have to be set and the prototypes need to be chosen, which can be very subjective. This is also shown in the colour models discussed in de previous chapter. So in order to make this method work, a consensus should be made regarding the prototypes and the variables as well, since they define the slope of the distribution.

Next, the intuitiveness could be improved. Modeling concepts to a intuitive space is a very difficult issue to solve. Even though it seems intuitive that semantically related concepts are more nearby than semantically unrelated concept, the precise distances between the concepts remain an obstacle.

Also, using this method, the distributions have two properties that might not intuitively be accurate. The first property is that the distributions fan out through the entire space, meaning that there are no boundaries between concepts. So even if two concepts are totally unrelated, if they both lie within the same conceptual space, they will still have a measure of typicallity. For example, in figure 21 the distribution of the concept *orange* would still provide a chance for the orange colour itself of being a green colour.

Finally, these distributions fan out equally in every direction. Figure 21 shows the flaw of this assumption; not all equal distances have the same equal

(un)relatedness semantically. Besides, if both T and k are equal for both distributions - i.e. both prototypes have the same weight -, the prototypical combination of two concepts always ends up in the middle of the two concepts. Of course it is also possible to differ the variables T and k, to move the outcome from or to one of the prototypes. However, a Boltzmann distribution, where the slope is not the same for every direction, would seem more natural. Or perhaps the Rasch model should be applied, where every concept is modeled according to the responses of subjects.

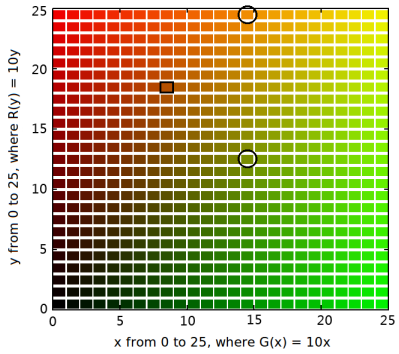


Figure 21: An example of the colour space, where the square represents a typical orange colour and both circles are at the same distance from the square.

4 Results and Discussion

4.1 Conceptual Comparison

Although the method of radial projection and the method of the reduced tensor product seem to be very different, there actually are just a few dissimilarities. The major difference between the two methods is that Gärdenfors' method is based on a meaning space in which concepts are convex sets and have boundaries. Whereas the method of the reduced tensor product uses distribution functions to formulate concepts throughout an entire semantic space. However, it should be noticed that this causes a similar constraint on both methods; to achieve radial projection, all concepts must be modeled using convex sets, but the reduced tensor product also knows a convexity assumption. When there are several prototypes in the same semantical space and we were to look at every instance seperately to see to which prototype they belong - i.e. to look for which prototype they have the highest cue validity -, and then assign the instances to one prototype, a Voronoi tessellation occurs. Since every cell of a Voronoi diagram is a convex set by definition, the reduced tensor product creates convex concepts as well [24].

The biggest semantic difference between the two methods is the notion of concepts having or not having boundaries. When we take the RGB colour space as an entire meaning space it seems logical to give the colour-concepts boundaries, because the concept *red* should not intervene with the concept *gray*. However, it is not possible to make a neutral gray colour without using some red and secondly, both concepts indicate a colour so they should be connected on some semantic level. On the other hand, looking at a colour model, the semantic definition that all instances within this model are indeed colours, was already established by choosing this meaning space. So perhaps the biggest semantic boundary was already chosen. Nevertheless, perhaps concepts indicated within colour spaces should have boundaries, whereas concepts within different meaning spaces should not have boundaries, or vice versa.

This notion actually leads to a psychological issue, whether concepts have boundaries and whether semantic models can be intuitive. Radial projection requires two points, an origo and a typical instance in one of the two concepts, both chosen subjectively. The second method also requires two subjectively chosen points, the two prototypes. Even if all people were to choose the same points for each method, the question of intuitiveness would still remain. This problem remains, because semantic models translate the difference in meaning into distances. The distances between meanings are not psychologically intuitive and perhaps Euclidean metric should only be used as a first approximation.

However, apart from these differences between the methods - the convex regions versus the distance distributions and concepts modeled with our without boundaries - the Euclidean distances along a linearly ordered line are used in both methods. Where the reduced tensor product uses distributions that result in a composed concept on the line between the two prototypes, radial projection draws a line between an arbitrary origo, through a prototype, to the outer boundary. Next, Gärdenfors' method uses the correspondence between the length of the line from the origo to the outlying prototype and outer bound-

ary and the length of the line from the origo to the inlying composed concept and the inner boundary. So basically, the placement of the composed concept on the line between two prototypes is either defined by the ratio of the boundaries of the concepts or by the ratio in weight of the prototypes of the concepts. The next section will cover the mathematical correspondence between the two proportions.

4.2 Mathematical Comparison

To compare the two methods, they both need to be reduced to a comparable form. Fortunately, both methods can be easily reduced to a straight line. In the method of radial projection, this line goes from the origo to the outer boundary of the entire space, crossing the target concept, the inner boundary and the typical outlying concept. Using the reduced tensor product, this line starts at the prototype of one concept and ends in the prototype of the second concept.

Because of the steady ratio between the two concepts and the combined concept caused by the reduced tensor product, this section will first recall the formula for this method. Next, this formula will be compared to the mathematical foundation of radial projection.

To compare the two methods, both lines need to be aligned. Firstly, the reduced tensor product can be described by a linear function, as explained in section 3.3.3, with the following formula:

$$t = \alpha P_B \tag{25}$$

However, it is important to repeat that this formula only has a differing T for both distributions, but that the k is equal, because we need a linear function to compare the two methods. Next, to align the method of radial projection with the proposed line of the reduced tensor product, it seems natural to replace the origo of radial projection with the prototype of concept A and the typical second concept, x, with the prototype of concept B (figure 22 and figure 23).

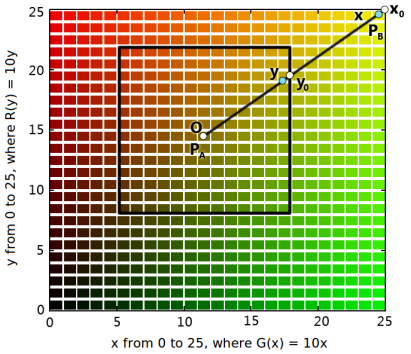


Figure 22: The RGB colour space, where blue = 0, the origo (P_A) is in the middle of the subspace and x (P_B) represents a typical yellow colour. Above the line are the terms that are replaced by the terms below the line.

Radial projection makes use of the following equation:

$$\frac{d(O, y_0)}{d(O, y)} = \frac{d(O, x_0)}{d(O, x)} \tag{26}$$

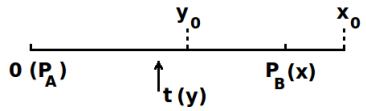


Figure 23: Representation of the line.

When the methods are aligned, and if the line of both methods begin in the same point, namely P_A, every distance from P_A to an arbitrary point on the line can be replaced by simply the placement of the arbitrary point on the line itself:

$$\frac{y_0}{t} = \frac{x_0}{P_B} \tag{27}$$

This leads to the following equation, ending in a simple formula for t:

$$t = \frac{P_B y_0}{x_0} \tag{28}$$

Now formula (28) can be compared to the equation for t in for the reduced tensor product, given in equation (25):

$$\frac{P_B y_0}{x_0} = \alpha P_B \tag{29}$$

Resulting in the following equation:

$$\frac{y_0}{x_0} = \alpha \tag{30}$$

This last formula shows the correspondence between the ratio of the inner and outer boundary and the ratio of the two concepts. So if the target lies in the middle of the two concepts - i.e. when α is equal to a half -, the inner boundary must be half the distance to the prototype of concept A compared to the distance from the outer boundary to the prototype of concept A. With this equation the results of the reduced tensor product can be used to calculate the boundaries for radial projection. Or, in short, the ratio between the inner and outer boundary in radial projection is equal to the ratio between the two prototypes of the concepts, depending on the values of T_A, T_B and k.

4.3 Discussion

The results regarding the comparison of the two methods bring the conclusion that, if attributive modification is calculated using Euclidean metric, the methods actually are interchangeable. In other words, starting with the reduced tensor product, the parameters of Gärdenfors' model can be reconstructed to find the same combined prototype and starting with radial projection it is possible to reconstruct the ratio in weights between the two prototypes of the reduced tensor product model. However, this is only possible under the assumption that the origo and typical instance of the model for radial projection are chosen at the same locations as the two prototypes from the reduced tensor product.

Then, the only remaining obstacles are the choosing of the prototypes or the origo and typical instance, the potential boundaries of the concepts and the intuitiveness of the semantic model itself. The choosing of all these points depend

on the individual and would require psychological research to achieve a common answer. Furthermore, if the prototypes were chosen based on popular opinions, a normal distribution with a mean and a deviation seems natural. However, people might still need boundaries for every concept instead of a distribution covering the entire meaning space.

Concisely, two domains can be further explored. The first is the modeling of a semantic space. It is possible that colour spaces are a bad example for attributive modification, or perhaps different colour spaces than the ones used in this thesis should be used, like the more intuitive model of Munsell. When attributive modification is tested on semantic models, the model should contain intuitive distances between its instances. To achieve this goal, the notion of using a different metric than the Euclidean metric system should be explored as well. However, when people do decide to use colour models for attributive modification, psychological research must be done to find more similarities between people in finding two prototypes and their combined colour. Also, if the prototypes are chosen based on votes, a distribution seems natural. However, if boundaries turn out to be a very natural mold as well, perhaps it is advantageous to combine the two methods.

The second domain that needs exploring is the mathematical point of view, which metric system should be used or which distribution functions would render the best results. Whether concepts do or do not have boundaries, both methods make use of convex sets and linearly ordered instances. Perhaps the two could be combined to distributions within convex regions with boundaries. Or perhaps a different distribution function, not being symmetrical or monotonically decreasing, would generate more intuitive results.

5 Conclusion

To understand language, pasting the meanings of seperate words together is not enough. When words are merged into a sentence, or even a part of a sentence, their meanings are merged as well. Focussing on the combination of a modifier and a head, multiple methods have been developed to solve attributive modification. Though there are two kinds of attribution modification; direct and indirect modification, most methods focus on the latter.

One of the most recent methods is the method of Gärdenfors. This method is based on the property that Gärdenfors ascribes to inner concepts of mental spaces: convexity. Based on this mathematical property, homeomorphism between two concepts can always be achieved through radial projection.

Making use of three assumptions - the existence of mental spaces, the convexity and Euclidean space -, two convex concepts with an overlapping part can be mapped into the intersection with a one-to-one correspondence. When two concepts overlap, an arbitrary origo is taken in the intersection and a typical instance is taken in one of the two concepts. Next, the distance from the origo to the typical instance with respect to the outer boundary of the corresponding concept is taken as a ratio for the distance from the origo to the resulting combination with respect to the boundary of the intersection.

To test this method, radial projection was applied to colour spaces, because these are the only modeled meaning spaces. Unfortunately, for both the RGB and the HSL colour space the method did not seem to render intuitive results. This problem could be due to the colour models, because in both cases one value of the model was fixed. Even though this should not cause a problem, perhaps using the entire colour spaces would provide different results. The second problem caused by the colour models is the unintuitiveness of the distances between colours, which might be solved by using different colour models or a different metric system. However, the perception of colours remains very subjective and therefore might not be suitable as an example for meaning spaces. Perhaps colour adjectives should even have its own behavioural class.

Finally, the most relevant problem this method encountered is the placement of the origo. Although the one-to-one correspondence between two sets with an overlap is possible for every arbitrary placement of the origo within the intersection, it does influence the placement and slope of the line from the origo to the typical instance. Furthermore, this typical instance is also subjective. In short, since both the origo and the typical instance are subjective, the result differs per person, making the method subjective instead of general.

The second method is based on the prototype theory and represented by a vector space. The prototype theory originates from the cognitive sciences and states that every concept has one member that is more characteristic than the others. When a concept is represented by a vector, this vector contains a degree of typicality for every instance of the concept. So if this representation is used in a semantic space, all instances will be ordered based on their meanings and the degree of typicality, called the cue validity, is defined as a function based on the distance between the instance and the prototype of the concept. Then, plotting the vectors, this will give a distribution with a peak on top of the prototype of the concept.

The next step is to find the prototypical combination of the two concepts, which is achieved by combining the two vectors that represent the concept of the head and the concept of the modifier. This method states that this should be done by taking the reduced tensor product between the two vectors. To stop the growing dimensionality of a regular tensor product, only the diagonal of this product is taken, hence the name *reduced* tensor product. The resulting vector also shows a peak, representing the prototypical combination of the two original concepts.

Nevertheless, this method depends on two variables; the cue validity is not only based on the distance between the prototype of the concept and every instance of the semantic space, but also influenced by the two parameters T and k. Whereas T represents the decay of the distribution, causing a higher or lower peak, the variable k represents power to which of the distance between the prototype and the instance is taken, which causes a bigger difference in magnitude of the cue validity.

The most important feature of this method, based on the symmetrical and monotonically decreasing distributions, is that the resulting peak aligns with the two original prototypes. Furthermore, when the parameters T and k are equal for both distributions, the combined vector has its peak exactly in the middle of this line between the prototypes.

However, this method also depends on subjectivity. First of all, the prototypes and the variables need to be set, which can differ between individuals. Next, the distance between semantically related and unrelated concepts forms an obstacle. Because the distributions are symmetrical and monotonically decreasing, an equal distance from the prototype means to have an equal semantic difference from the prototype, which is not an universal property of meaning spaces, as shown in the colour spaces. Also, however likely it seems that semantically similar meanings are closeby, exact distances are subjective as well. And finally, every instance in the semantic space will have a chance of having the same meaning as the prototype, because the distributions fan out through the entire space, without any boundaries.

Since both methods use Euclidean metric and are based on linearly ordered instances, they can be aligned mathematically. When the two points of radial projection, the origo and the typical instance, are chosen at the same location as the two prototypes of the reduced tensor product, a correspondence can be found between the two methods. This correspondence entails that the ratio between the inner and outer boundary of radial projection is equal to the ratio between the weights of the two prototypes. So if both prototypes have the same weight, the target lies in the middle of the two prototypes and for Gärdenfors' method the outer boundary of the concepts should be twice as far from the origo as the inner boundary.

Using this correspondence, only one difference between the methods remains: the modeling of concepts. The difference is the existence of boundaries; whereas the reduced tensor product uses the entire meaning space for every concept, radial projection chooses a convex region within the semantic space for every concept. The notion of each concept having boundaries or being related to every other concept is a psychological question and maybe even depends on the individual. Perhaps some concepts should have boundaries while others should not. Also, when a meaning space is already specialized to one aspect, like the

colour models, the biggest semantical boundary is already set.

In short, when the mental space is convex and ordered semantically, and two prototypical points are chosen, using Euclidean metric, both methods can be used and by adjusting the ratio between the boundaries of radial projection or the variables T and k of the reduced tensor product, the same result can be found. The only remaining difference would be the existence of boundaries of the concepts, which would depend on the psychological modeling of the concepts and thus would be in the eye of the beholder.

References

[1] C. Berge. *Topological Spaces*. Mineola: Dover, 1997.

[2] R. Blutner. Concepts and bounded rationality: An application of niestegges approach to conditional quantum probabilities. *Foundations of probability and physics*, 5:302–310, 2009.

[3] D. Bourgin. The color space faq, 2004.

[4] J.R. Busemeyer and P.D. Bruza. *Quantum Models of Cognition and Decision*. Cambridge University Press, 2012.

[5] C.P. Dolan and P. Smolensky. Tensor product production system: a modular architecture and representation. *Connection Science*, 1(1):53–68, 1989.

[6] P. Gärdenfors. *Concept combination: a geometrical model*. Lund University, 1998.

[7] P. Hájek. *Metamathematics of fuzzy logic*. Kluwer, Dordrecht, 1998.

[8] G. Hoffmann. Cie color space, 2002.

[9] Gerhard Jger. The evolution of convex categories. *Linguistics and Philosophy*, 30:551–564, 2007.

[10] E. Landa and M. Fairchild. Charting color from the eye of the beholder a century ago, artist albert henry munsell quantified colors based on how they appear to people; specializations of his system are still in wide scientific use. *American scientist*, pages 436–443, 2005.

[11] J. Mitchell and M. Lapata. Vector-based models of semantic composition. In *In Proceedings in ACL-08: HLT*, 2008.

[12] J. Mitchell and M. Lapata. Composition in distributional models of semantics. *Cognitive Science*, 34(8):1388–1429, 2010.

[13] D.N. Osherson and E.E. Smith. On the adequacy of prototype theory as a theory of concepts. *Cognition*, 9(1):35–58, 1981.

[14] B.H. Partee. *Privative Adjectives: Subsective plus Coercion*. Emerald Group Publishing Limited, 2010.

[15] D. Pascale. A review of rgb color spaces. *The BabelColor Company, Montreal, Tech. Rep*, 2003.

[16] T.A. Plate. Holographic reduced representations: convolution algebra for compositional distributed representations. *12th International Joint Conference on Artificial Intelligence*, pages 30–35, 1991.

[17] T.A. Plate. A common framework for distributed representation schemes for compositional structure. *Connectionist systems for knowledge representation and deduction*, pages 15–34, 1997.

[18] T.A. Plate. Structured operations with distributed vector representations. *Advances in Analogy Research: Integration of Theory and Data from the Cognitive, Computational, and Neural Sciences*, 1998.

[19] P. Smolensky. Connectionism, constituency and the language of thought. 1988.

[20] P. Smolensky. Tensor product variable binding and the representation of symbolic structures in connectionist systems. *Artificial Intelligence*, 46(1–2):159–216, 1990.

[21] S. Süsstrunk, R. Buckley, and S. Swen. Standard rgb color spaces. In *Proc. IS&T/SIDs 7th Color Imaging Conference*, pages 127–134, 1999.

[22] I. Tastl and G. Raidl. Transforming an analytically defined color space to match psychophysically gained color distances. In *Electronic Imaging 1998 Conference*, 1998.

[23] P. Turney and P. Pantel. From frequency to meaning: Vector space models of semantics. *Articial Intelligence Research*, 37:141–188, 2010.

[24] M. Warglien and P. Gärdenfors. Semantics, conceptual spaces, and the meeting of minds. *Synthese*, pages 1–29, 2007.

[25] L.A. Zadeh. A note on prototype theory and fuzzy sets. *Cognition*, 12:291–298, 1982.