Conceptualizing Creativity: From Distributional Semantics to Conceptual Spaces

Kat Agres, Stephen McGregor, Matthew Purver, and Geraint Wiggins
School of Electronic Engineering and Computer Science
Queen Mary University of London
London E1 4NS UK
kathleen.agres, s.e.mcgregor, m.purver, geraint.wiggins (@qmul.ac.uk)

Abstract

This paper puts forth a method for discovering computationally-derived conceptual spaces that reflect human conceptualization of musical and poetic creativity. We describe a lexical space that is defined through co-occurrence statistics, and compare the dimensions of this space with human responses on a word association task. Participants' responses serve as external validation of our computational findings, and frequent terms are also used as input dimensions for creating mappings from the linguistic to the conceptual domain. This novel method finds low-dimensional subspaces that represent particular conceptual regions within a vector space model of distributional semantics. Word-vectors from these discovered conceptual spaces are considered, and argued to be useful for the evaluation of creativity and creative artifacts within computational creativity.

Introduction

This paper presents a computational-linguistic model for mapping lexical spaces populated by statistical representations of words to conceptual spaces defined in terms of feature dimensions of conceptual representations. This research has three main goals. The first is to compare the features of a distributional semantic vector space with the results of an empirical word-association task completed by human subjects. This empirical corroboration serves to demonstrate that the model can capture meaningful aspects of human conceptualizations of queried topics, which are "musical creativity" and "poetic creativity" in the present study. The second goal is to use novel methods inspired by computational linguistics to map terms from the linguistic domain to representations in the conceptual domain. To this end, the terms generated by participants are used as input parameters for our computational model that uses co-occurrence statistics and linear algebraic metrics to quantify conceptual proximity. The third motivation of this work concerns the evaluation of creativity. In the field of computational creativity (CC), the evaluation of creative output is often either subjective on the part of the developer/researcher or unsystematic. We offer our own fundamentally computational approach as a means of identifying facets of the investigated concept or domain. Put another way, our model can generate terms within a conceptual space that may be used to query different aspects of creative output or creative behavior.

Vector space models of distributional semantics are currently a popular approach for quantifying linguistic similarity, but many contemporary studies need grounding and external validation. Much of the work in this area compares model performance to semantic databases, but does not directly relate results to the cognitive performance of humans, or uses very restricted tasks, such as similarity judgments, rather than imploring subjects to elaborate on *concepts*. Because our aim is to elucidate how humans conceptualize creativity, sampling from people's own formulation of conceptual spaces is essential. Therefore, in the present work, our ground truth is derived from human responses stemming from direct queries about creative concepts. Because human response data is a limited and expensive resource, we hope that our comparison to human data will inform how conceptual spaces may be discovered as autonomously as possible in the future (that is, without the requirement of subjective user-input or parameter-tweaking). We also believe that this multidisciplinary and externally validated approach produces a more robust system.

In order to pinpoint the relationship between the output of our computational model and the results of our empirical study, we take the human-generated terms and investigate their situation within the multidimensional space of our distributional semantics model. We then determine the characteristic co-occurrence dimensions of sets of words associated with concepts, and apply appropriate methods to reduce the dimensionality of the space in order to map broader clusters of linguistic terms to conceptual regions. We argue that the online generation of a reduced lexical space corresponds to the contextualization inherent in the momentary way in which concepts are necessarily formed in response to situations in a cognitive environment. We expect that this methodology will be a useful applied approach to formalizing the geometrical representation of conceptual spaces.

Our research explores two related concepts: musical creativity and poetic creativity. There are several reasons for this choice. First, we are interested in computational creativity, and in particular in the evaluation of creative systems and their output. In order to evaluate creativity, it is necessary to characterize features of this concept using the expressive affordances of language. Our computational methods seek to capture these features of the conceptual space. Our model may also be used to discover conceptually-related terms that a human might not necessarily immediately consider. We hope this approach may be used to elaborate abstract concepts by elucidating an extensive set of terms that correspond to the queried conceptual spaces. We therefore offer this methodology as a novel approach for exploring and elaborating concepts, both for the evaluation of creative systems and for potentially contributing to creative pursuits themselves (such as poetry generation). Furthermore, we apply our method to a more concrete domain, extending a small subset of terms relating to the concept WILD ANIMALS in order to indicate the anticipated generality of our model.

The organization of the paper is as follows: first we offer a summary of computational approaches to conceptual creativity, situating our research within the field. This overview leads into a discussion of computational approaches to the topics of conceptual spaces and geometric representations of concepts. An explanation of our computational model is then provided, including a description of how we have modeled a lexical space populated by word-vectors. This is followed by a description of our empirical study with human participants, and findings from this questionnaire-based study are reported. Given this context, we then discuss two ways in which the participants' responses contribute to our computational approach. The first is a comparison of computationally-derived terms with human-generated terms. The second contribution will be to treat the salient features of the word-vectors corresponding to the most frequently reported human terms as an indication of the dimensions of a vastly reduced subspace of our distributional semantic model. We then discuss how terms that fall near the centroid of the positively valued surface of the discovered lexical spaces may be used for the evaluation of creativity.

CC and Concept Discovery

Computational approaches to creative conceptualization have provided a target that is both elusive and essential to the identity of a field that incorporates a particularly diverse range of topics. Creativity itself has been interpreted by Koestler (1964) as a kind of meshing of disparate conceptual schemes, by which expectations are violated in favor of interesting new combinations of frames of reference. Presciently, Koestler has couched his model of creativity in terms of vector spaces and transformations, an idea which is broadly shared by the model presented in the present paper. In the same spirit of conceptual exploration, Hesse (1963) argued that the formation of creative analogies is the essence of scientific discovery, an idea demonstrated by the primacy of analogical modeling in fields such as physics, where there is no realistic way to literally conceive of phenomena that occur on obscurely minuscule or vast scales.

In the specifically computational domain, Veale (2006) has proposed a system for the dynamic generation of new, non-literal conceptual categories based on a computational analysis of a taxonomical database such as WordNet. Likewise, other researchers are developing formal models of conceptual blending (Fauconnier and Turner, 2008) that seek to discover novel combinations of familiar ideas, targeting domains such as mathematical reasoning and story generation (Ontanón, Zhu, and Plaza, 2012). These approaches make clever and effective use of heuristics to pick out interesting new conceptual representations based on preconceived patterns identified by programmers. As such, the output of these methods is compelling and valid, but the conceptualization itself is arguably handed to the system in the prepackaged form of externally grounded symbols.

Elsewhere, Heath et al. (2013) have taken a more connectionist approach to conceptual creativity, combining human based word associations with statistical models of distributional semantics to design a system that infers conceptual categories from lists of terms, and likewise generates lists of terms from linguistic input that is interpreted conceptually. In a similar vein, Jäger (2009) has performed a statistical analysis on a set of human reported color terms and used this analysis to generate a geometric representation of certain consistencies in the ways that color is conceptualized across cultural linguistic boundaries. In their commitment to building models based on low level, non-symbolic observations about the world, these statistical approaches to creative conceptualization are in the same spirit as the work presented in the present paper.

The model described here has been designed to engage with the field of computational creativity on two different planes. Principally, our method seeks to implement a low level approach to the delineation of conceptual regions based on the geometry of a distributed semantic space. By viewing concepts as momentary and pragmatic phenomena, we are able to use *ad hoc* reductions of a high dimensional lexical space to map concepts creatively based on situational contexts which do not have to be preformulated in the design of the model. Furthermore, our target domains of musical and poetic creativity play nicely into a salient issue in the field of computational creativity: the analysis of creativity itself, a difficult procedure that necessarily involves some degree of conceptualization about creativity. This secondary aspect of the work, the potential for meta-analysis inherent in the question of whether our model's output will be useful for guiding an evaluative discussion of creative work elsewhere, is intended to give the work its own pragmatic grounding, in that this suggests a practical application for the creative output described in the following pages.

Spaces of Meanings

This project uses computational methods as a platform for exploring the relationships between words and concepts within the context of a cognitive system. In the pragmatic spirit of Wittgenstein (1953) and Grice (1969), language is presented as a system defined by its own functionality, with meaning emerging from the use of words in the course of accomplishing communicative goals. To the extent that language is used to communicate ideas, statements are formed contextually, with reference to expectations about how relationships between words will suggest hierarchies of categorization relative to a particular situation. Barsalou (1993) characterizes the relationship between words and concepts in terms of the *linguistic vagary* inherent in the application of names to ideas: words represent concepts in a way that is fleeting and mutable. Fundamentally, words stand as indices to concepts, and the relationship between language and ideas is best understood as a mapping between two separate domains. The project presented in this paper is therefore motivated by a desire to model the relationship between two different spaces, one of words and one of concepts, and to explore the ways in which these spaces might be aligned in terms of the computationally tractable elements of their geometries.

Gärdenfors (2000) has presented a spatial theory of concepts, by which the dimensions that determine the geometric situation of a conceptual region within a space of concepts correspond to the attributes which characterize that particular region. So, for instance, the concept RIPE BANANAS would occupy a region towards the higher end of the dimensions of curviness, yellowness, and sweetness within a conceptual space. This literal and factual quality of dimensions grounds conceptual spaces in low level observations about the world, giving regions within the space a geometric dynamism that lends itself to doing higher level work with the entities that emerge from the space as symbolic representations. In particular, well defined conceptual regions are characterized by convexity, a property that ensures that any intermediate point between two outlying extensions of a region will likewise belong to that domain.

Vector space models of distributional semantics, on the other hand, offer an approach to language modeling involving a distinctly unstructured computational analysis of linguistic data. In the tradition of Harris (1957), the distributional hypothesis holds that there is semantic information inherent in the statistical comportment of language: linguistic meaning can be found in the quantifiable contextual relationships between words. This insight has motivated a productive field of research, with computational analyses of large scale corpora yielding distributional semantic models in which the meanings of words, sentences, and documents are rendered in terms of mathematically tractable representations (Schütze, 1992; Landauer et al., 1997). Distributional semantic models treat words as vectors, with the dimensions of these vectors representing, either directly or abstractly, the likelihood of a word occurring in a given context. The closeness of vectors in a lexical space, which reflects the tendency of the proximal vectors to occur in similar contexts, has been shown as an indication of lexical similarity between the words tied to the vectors. In their most straightforward implementation, distributed semantic spaces are constructed by counting the frequency with which each word in the model co-occurs with all other terms in a base corpus (see Turney and Patel, 2010; Clark, 2015, for an overview).

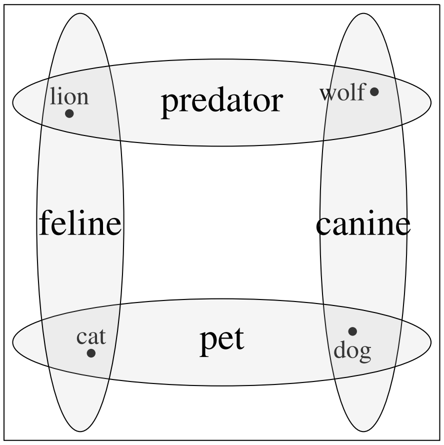
There is an important difference between lexical spaces and conceptual spaces: the dimensionally regimented quality of coherent domains within a conceptual space is not reflected in the distribution of vectors in a lexical space, where the dimensions of word-vectors correspond simply to the context in which those words are likely to occur, and therefore capture all the flexibility and uncertainty of language in use—the *linguistic vagary* of Barsalou's system of conceptual symbols. So, for instance, the other vectors in the proximity of the word-vector $\overrightarrow{pet}$ in a distributed semantic model cannot be expected to contain only terms corresponding to domesticated animals, not least because the word "pet" itself has other uses. In this sense, where conceptual spaces are marked by a tidy taxonomy facilitated by the clarity of a region's dimensional substrate, distributed semantic spaces embody the pragmatic messiness of language as it is encountered in its natural, operational environment. Therefore, while lexical spaces and conceptual spaces both utilize geometry as a vehicle for semantics, the arrangement of a lexical space is in an essential way less ordered.

An example of the difficulty of delineating conceptual regions within a lexical space is illustrated in Figure 1a. In the rudimentary distribution of words presented here, concepts are required to stretch and overlap in order to maintain their lexical constituencies. This simplified depiction of the potential uncertainty of conceptual membership doesn't demonstrate the even more fundamental problem of picking out salient words in regions that are littered with noise: in practice, in the densely and unevenly populated territory mapped out by a vector space model, many unwanted terms will be discovered in the region generally between two other terms. For instance, in the unrefined version of our model, there are 14 terms essentially between the word-vectors $\overrightarrow{cat}$ and $\overrightarrow{dog}$, including such unlikely candidates as "during", "eventually", and "featuring". There is thus an inherent patchiness to the mapping of concepts that might be read in an unrefined vector space model of distributional semantics.

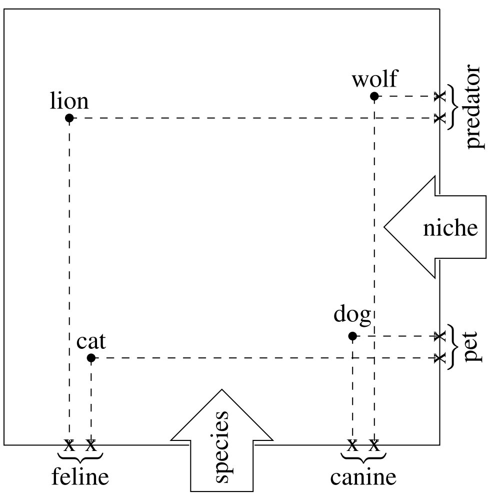
Here we propose a system for mapping lexical spaces to conceptual spaces by considering a conceptualization as a particular and temporary perspective on a space of distributed semantics. The idea behind this system is that, for any desired clustering of words corresponding to a particular conceptualization, there is some subset of a distributional space's dimensions that will render a subspace in which that clustering is realized. This intuition is illustrated in Figure 1b, where the conceptually entangled space of Figure 1a collapses into a particular conceptual regime depending on the axis along which the space is projected, which is to say, the perspective from which the space is considered. The task of our system is therefore to determine the dimensions which should be picked out of a higher order vector space model in order to realize a grouping of terms that is conceptually homogeneous by virtue of the contextualization imposed by a particular perspective on the space.

It is precisely the massive dimensionality of the space which facilitates the method's ability to pick out various successful conceptual perspectives on the space in a momentary and continuous way. With each additional contextual dimension introduced to a vector space, there is an exponential increase in the lower-dimensional combinations available to map corresponding spatial relationships of words to conceptual subspaces. Moving from the linguistic realm native to vector spaces back to the cognitive domain targeted by Gärdenfors, these dimensional perspectives might be construed as corresponding to a contextualized perception of a situation. In this respect, our system models the *haphazard* quality of conceptualization described by Barsalou, as well the *ad hoc* nature of concept formation discussed more recently by Allott and Textor (2012), who suggest that meaning is appropriated *in situ* to endow statements with contextually relevant implicature. This phenomenon of conceptualization arising from pragmatic communicative affordances is what our method seeks to computationally model.

Figure 1: Conceptual Perspectives of Vector Spaces



(a) In this simplified and unrefined distributional semantic space, the conceptual regions suggested by the spatial arrangement of terms are indeterminate. Each word is roughly equidistant from two other terms, either of which could be linked in a distinct linguistic depiction of a concept. The conceptual domains which are delineated by this arrangement of words are awkwardly elongated.



(b) If the same space illustrated above is considered from two different perspectives, the indeterminate arrangements of words collapse into lower dimensional spaces (one dimensional, in this simple example) in which the clustering of terms suggests straightforward conceptual domains. These perspectives effectively contextualize the meanings inherent in the distributional characteristics of the language model, and this context facilitates the mapping of the linguistic space to sets of conceptual regions.

A Literal Lexical Space

Our lexical model has been constructed based on the distribution of words found in the textual component of articles on the English language Wikipedia website.[1] The xml code of the site was downloaded and then parsed into a text-only format, eliminating images, tables, lists, captions, and section titles, leaving only the well formed sentences composing the content of the site's articles. Sentences were separated by identifying terminal punctuation followed by whitespace, then punctuation was removed and all characters were converted to lower case. Articles ("a", "an", and "the") were stripped from the text. Sentences containing less than five words were discarded. The resulting corpus consists of almost 60 million sentences, containing about 1.1 billion word tokens (individual words) corresponding to about 7.4 million word types (classes of words).

From this base corpus, we took the 200,000 most frequent word types to form our system's vocabulary. Our full lexical space is represented as a matrix $M_{w,c}$, where rows correspond to vectors representing words, and columns correspond to co-occurrence terms. The cell for word w and co-occurrence term c contains the *mutual information* $MI_{w,c}$ as described in Equation 1. Here $n_{w,c}$ is the frequency with which a term c is observed to co-occur within a context window of two words on either side of a vocabulary word w; n_w represents the total count of w in the corpus; n_c is the total count of c; and a is a smoothing constant.

$$MI_{w,c} = \log_2\left(\frac{n_{w,c} \times N}{n_w \times (n_c + a)} + 1\right) \tag{1}$$

The constant a reduces the undesirable effect of contextual words that occur very rarely throughout the corpus, but with a high frequency in the context of certain target words—we found 10,000 to be a good value for a based on trial and error. The value 1 is added to the probabilistic ratio in order to render all dimensions within the space positive: this means that the logged MI value of target words and context words that never occur together will be 0, and the value for terms that co-occur less frequently than would be expected in a random distribution will be between 0 and 1. Each word vector consists of a set of dimensions derived through this calculation, and each of these vectors is normalised to the scale of a unit vector. The result of this process is a distributional semantic space in which each of the 200,000 vocabulary terms sits in the positive region of the high-dimensional surface of a hypersphere.

One notable feature of our vector space is the literal correspondence of its dimensions to co-occurrence terms. In general, state-of-the-art systems apply some form of dimensional reduction to the overall space, either using linear algebraic transformations to perform a principal component analysis (Pennington, Socher, and Manning, 2014), or by weighted networks to train abstract lower-dimensional word representations that predict the context in which that word is encountered in the course of training (Mikolov et al., 2013). While these techniques certainly make the space less expensive to compute, and arguably improve results for a variety of semantic tasks, our system is specifically geared towards the identification of salient, literal co-occurrence dimensions, and as such the space is, for the purposes of our initial analysis, maintained in its raw high-dimensional form. The dimensionality of our space is therefore on the

[1]The December 8, 2014 dump, downloaded from http://meta.wikimedia.org/wiki/Data_dump_torrents on January 23, 2015, parsed into plain text using the "Wikipedia Extractor" software, downloaded on February 13, 2015 from http://medialab.di.unipi.it/wiki/Wikipedia_Extractor.

order of 7.4 million, as every word token in the corpus is a potential context for the 200,000 words in our vocabulary.

Empirical Validation

In order to provide grounding and validation for our computational model, human participants were asked to generate terms during a word association task, and to reflect upon how they would evaluate creativity in the musical and poetic domains. Provided terms were analyzed for comparison with the vector space model.

Method

Participants Twenty participants (avg age = 30 yrs, stdev = 5.2 yrs) volunteered to take part in the study, of which 11 were female. Sixteen individuals indicated that their career is either inherently creative or that they apply creativity to improve their job performance, and all but two of the participants engage in creative pursuits outside of work. Seven currently practice or perform music, and five individuals currently engage in creative writing.

Procedure After reading an information form and providing informed consent, participants were given a brief questionnaire to complete. Two of the questions consisted of a word association task in which participants were asked to list three terms they associate with "musical creativity" or "poetic creativity". The order in which musical or poetic terms were prompted was counterbalanced across subjects. The other two questions requested participants to write one sentence describing how they would evaluate whether a new piece of music or poetry sounds creative (the order of these questions were similarly counterbalanced across subjects). These results will only briefly be touched upon in the current paper; although certainly of interest, due to space constraints, an in-depth analysis of the evaluation sentences must be saved for an expanded version of this work.

After providing their responses, participants were given questionnaires requesting general demographic information (age, ethnicity, etc), and information about their past and current involvement in creative pursuits, e.g., "Do you currently play music or engage in creative writing?" Upon completing these, participants were debriefed as to the goals of the experiment and paid £2 each for their participation.

Results

For both musical and poetic creativity, participants' terms were placed into two lists: An exhaustive list of all terms provided for the concept, and a short list for terms cited by more than two participants (per concept). In the case of musical creativity, this yielded an exhaustive list of 52 distinct terms, and for poetic creativity, a set of 42 terms. The short list of musical terms included the following six terms: *innovation, sound, instruments, novelty, emotion, and expression*. The short list of poetic terms included these six terms: *emotion, rhythm, expression, structure, flow, and words*. We interpreted these concise lists of most frequent words to reflect dimensions of the concept that are more central to the conceptual space they populate. This resulted in discarding more peripheral terms such as "sensitive" that are undoubtedly related to creativity, but not cited frequently as an associated concept. Plurals and conjugations were considered to be the same category of term, e.g., "emotions", "emotional", and "emotion" were tallied together as "emotion". We continue the discussion of empirical findings in the next section, as we compare the model's performance with human results.

Mapping Words to Concepts

We began the exploration of mappings from our space of distributed semantics to a conceptual space with a top down approach, investigating the way our system reacted to the same kind of input that we presented to our human subjects. Along these lines, we examined the vectors for the word pairing "musical" and "creativity", and likewise for the pairing "poetical" and "creativity". In each instance, we calculated the mean value for each dimension that had a non-zero value for both words – that is, for each dimension corresponding to a term that co-occurred with both words at least once in the corpus – and returned a ranked list of average scores, running from high to low. Out of the 7.5 million co-occurrence features across the entire model, 4,772 were non-zero for both "musical" and "creativity", and 2,673 for both "poetic" and "creativity", statistics which highlight the sparsity of the base space. Our objective was to examine the nature of the terms that tended to come up in the context of our query as phrased for our human subjects. Results are listed in Table 1.

The first thing to note about these results is that they are, in a qualitative sense, coherent descriptions of properties typically associated with the two concepts being explored. To frame this more empirically, these results can be extended in order to discover how far down the list of top mean dimensions the terms reported by humans lie. Of the exhaustive list of terms reported by human subjects in response to the "musical creativity" query, 4 fall within the top 15 results generated by our model; likewise, 4 human responses fall within the top 15 mean dimensions for "poetic creativity" (these terms are italicised in Table 1). Considering that 200,000 words were used as the vocabulary of the model, yielding 4 of the top 15 dimensions in common with humans' responses for both concepts is quite compelling. This outcome may be interpreted as indicating that there is a high degree of mutual information between the query words and terms that humans would consider as conceptually descriptive of those queries. In other words, there is a high likelihood of conceptually relevant co-occurrence within the context of terms that summarize these creative conceptual domains.

These positive results do not hold up, however, for more concrete queries. For instance, when the mean dimensions for the query pair "wild" and "animal" are explored, top ranking results include some conceptually appropriate terms such as "boars", "deer", and "feral", but less directly relevant words like "skins" and "vegetable", and even antonymic terms like "domesticated" are also returned. It would seem that, in the case of words indexing more concrete concepts, the likelihood of co-occurrence in the conceptual context moves away from terms that generically describe components of the concept in question. This distinction is corroborated by Hill, Korhonen, and Bentz (2014),

"musical" & "creativity"	"poetic" & "creativity"
innovation	genius
imagination	*imaginative*
inventiveness	*metaphors*
improvisation	originality
talent	prose
talents	creativity
experimentation	artistry
versatility	craftsmanship
artistic	intuition
creativity	*imagery*
ingenuity	inspiration
aesthetics	talents
spontaneity	lyrical
individuality	talent
artistry	*self-expression*

Table 1: The top 15 dimensions with the highest mean scores between the word-vectors for each of our queries as given to human subjects. Terms in italics denote dimensions that were also cited by humans.

who have used computational analyses of both corpora and semantic graphs to illustrate a distinction between the way that abstract and concrete concepts are arranged in a cognitive linguistic system. It is hardly surprising, given the inherent ambiguity of language use – replete, as it is, with metaphor and implication – that simple co-occurrence probability statistics do not generally map neatly on to well defined conceptual spaces.

Projecting Words to Conceptual Subspaces

Motivated by this predictable shortcoming of a simple dimensional analysis, we developed a more sophisticated approach for delineating conceptual regions within dimensionally reduced subspaces of our language model. Our technique involves first hand-picking a small set of terms that might be considered as paradigmatic descriptions of components of a conceptual domain (in the present example, WILD ANIMALS). We perform an analysis similar to the one described above on these conceptual component terms, selecting the word-vector for each term and then extracting those features with non-zero values for all input terms. Once again, we compute a ranked list of these mean feature values and choose the co-occurrence dimensions which scored highest on average. These salient dimensions for the small set of words analyzed are again somewhat scattered: some of the highest mean dimensions correspond to relevant animal names, but the results also stray into the more conceptually ambiguous territory signified by words like "sightings", "chases", and "fat". There are 827 universally non-zero dimensions found between the word-vectors of the six input terms describing exemplars of WILD ANIMALS listed in the first column of Table 2.

We use these salient dimensions to define a drastically simplified subspace of our lexical model. Specifically, we reduce the model to the top 30 dimensions associated with the set of sample words (we arrived at the number 30 through trial and error; lower values tended to invite some unusual vectors into the crucial region of the subspace). After normalizing the new subspace, we then identify the central point on the surface of the positively valued quadrant of the reduced hypersphere—effectively the vector defined by 30 dimensions each with the the value $1/\sqrt{30}$. This positive centroid is then taken as the epicenter of a linguistic mapping of a new conceptual region, and we expand the region outward concentrically from this point, returning an ordered list of the points closest to the center of the positive surface of our space's low dimensional projection. Euclidean proximity is calculated by computing the square root of the sum of the squared feature-wise differences between the unit centroid and each of the 200,000 vocabulary words projected into the subspace. The top fifteen terms encountered using this method are reported in Table 2. Please note that the input terms for WILD ANIMALS are used as a preliminary test of the model's performance for concrete concepts; the input was hand-selected by the investigators, while the collection of terms relating to concrete concepts is the subject of ongoing emprical study.

This same technique for expanding conceptual regions through a dimensional reduction of a distributional language model is applied to our target domains of musical and poetic creativity, again with compelling outcomes. In this case we were able to make use of our results from our survey: for each of our two target domains, we choose all the terms that were reported by three or more human subjects and analyze these for their most salient dimensions of co-occurrence. Again, the system's output for these terms is not entirely unexpected, but also not conceptually completely cohesive. In the case of the highest mean dimensions for the human reported constituents of MUSICAL CREATIVITY, a number of predictable terms are returned, but somewhat less obvious dimensions such as "lab", "mere", and "shapes" also rank towards the top of the list.

Despite the conceptual uncertainty in the dimensional analysis, when a new subspace is constructed based on these dimensions, the central region of this space is replete with terminology appropriate to the example words at the base of the process. Interestingly, the original input words are only partially rediscovered in this new space, at least within the set of vectors most central to the positive surface of the new subspace. This indicates that some of the input word-vectors (used to select dimensions for creating the new subspace) are, in terms of the probability of regular co-occurrence with all the dimensions that underwrite the subspace, relative outliers which nonetheless make an essential contribution to the delineation of this linguistic representation of a conceptual region. It is also notable, and perhaps even remarkable, that in the case of the mapping of the conceptual region of poetic creativity, quintessential new terms such as "phrasing" and "inflection", arguably more intricately associated with the prosodic nature of the target domain than the original human generated terms, arise independently.

When examining how well the model captures relevant terms to a conceptual query, it is informative to cluster human responses into semantic categories (such as *emotion* and *structural elements*), and compare these results to apparent categories of output vectors. For example, considering the sentences that the 20 participants wrote about how

WILD ANIMALS		MUSICAL CREATIVITY		POETIC CREATIVITY	
human input	model output	human input	model output	human input	model output
lion	bobcat	innovation	novelty	emotion	phrasing
wolf	alligator	sound	liveliness	rhythm	intonation
coyote	raccoon	instruments	spontaneity	expression	musicality
alligator	opossum	novelty	innovation	structure	nuances
bear	armadillo	emotion	expressiveness	flow	timbre
snake	white-tailed	expression	refinement	words	sprightly
	anteater		nuance		rhythmical
	ocelot		ingenuity		nuance
	peccary		believability		expressiveness
	pronghorn		newness		rubato
	cougar		sophistication		instinctive
	cottontail		dynamism		bluesy
	rattlesnake		subtlety		directness
	skunk		vibrancy		modal
	boar		elusiveness		inflections

Table 2: The output vectors most central to the positive regions of the subspaces reduced in terms of the salient dimensions of a small set of conceptually exemplary input terms.

they would evaluate the creativity of a new song, many individuals referred to the notion of *novelty* in musical creativity, but used various terms to do so. In addition to explicitly using the term "novelty", participants made reference to "unexpected", "new", and "surprising elements" that were "like nothing else I'd heard before", as well as "melodic originality" and cases in which "known musical concepts or styles [are] combined in a novel/innovative way." Similarly, when considering the model's conceptual space of musical creativity, the words "novelty", "innovation", "ingenuity", "newness", "inventiveness", "distinctiveness", and "uniqueness" are found within the top 30 model output vectors. Although this is a qualitative assessment of the results, it does seem clear that the precise terms from humans and the model might not be exactly the same, but there is significant categorical or conceptual overlap between the two.

One may also note that the output vectors for poetic creativity appear to be rather "musical". This may reflect the fact that the input dimensions were provided by people who, overall, have significant musical experience - all but three of the participants have had musical training or have played music informally, whereas only four of the participants have experience with creative writing. People's experience with music might frame the way they think about other creative domains, or at the very least influencing the terms used to describe poetic creativity; consequently, this has led to a subspace that highlights the musical nature of this sample.

In light of our model's ability to find conceptually proximal terms, we propose that this method has the potential to be practically applied to the discovery of unexpected and valuable terms for the evaluation of creative output. Importantly, this approach may be applied to different corpora; for example, Wikipedia pages in different languages may be explored to address the difficult issue of identifying conceptually similar spaces across languages. The model's conceptual spaces concretely delineate evaluative terms that one person alone may not consider. For example, the terms "distinctiveness", "finesse", "artistry", and "stylization" were not cited by humans, but were within the top 30 output vectors discovered by the model. Future work may build on these findings, by using the model's discovered terms as criteria for subjective evaluation of creative output. In addition, discovering the geometry, flexibility, and contextual specificity of conceptual spaces may be very useful for assessing products or systems based on specific underlying concepts (or developed to address particular conceptual issues).

More generally, our method is presented as an implementation of the mapping of words to concepts: this approach charts a passage from a statistically tractable lexical space to the abstract but natively cognitive domain of ideas. The temporary and contextual aspect of this mapping is essential to its success: it is the flexibility of the model that allows for the bespoke generation of subspaces, just as it is the pragmatic frangibility of language that permits the ready-to-hand adaptation of meaning for unfolding expressive purposes. As can be seen in our results, the same terms arise in different constellations of meaning depending on the contextual perspective taken on the space. It is the strength of our language model that it can be adapted in this way, with the high dimensional arrangement of words allowing for their projection as multitudinous conceptual representations.

Conclusion

We investigated the terms and concepts that individuals most strongly associate with creativity in the musical and poetic domains, and described a computational methodology for modeling these conceptual relationships. Our multidisciplinary approach employs methods inspired by computational linguistics, as well as methods from empirical psychology. There were several outcomes of this work: the output from our distributional semantics vector space model was compared with human responses on a word association task. Human-generated terms were found within the top 15 dimensions of our model's lexical space, despite the model's very large vocabulary. This served as validation that the model discovers a lexical space that encapsulates the kind of terms humans use to describe these concepts.

Subsequently, the most frequently reported human terms were used as model input parameters for discovering conceptual spaces of lower dimensionality. Our model was able to find vastly reduced subspaces corresponding to MUSICAL CREATIVITY and POETIC CREATIVITY, which again captured semantically relevant terms, many corresponding directly to participants' terms, and others extending the list of terms to insightful new dimensions. In addition, by sampling word-vectors that fall near the centroid of the discovered conceptual mappings, we aimed to find potentially useful terms for the evaluation of creativity. Although computational and AI methods have generated many systems which aim to display creative behavior or produce creative artefacts, the evaluation of computational creativity remains distinctly problematic. Therefore, we offer our method and results as a formal approach to delineating conceptually-relevant criteria on which to base the evaluation of creativity and creative artefacts in future studies.

We saw, in terms of the most common dimensions in lexical space and the highest-mean word vectors in conceptual space, that the model is able to discover semantic categories and indices of concepts that are alligned to human conceptualizations. This said, the model did not capture all of the semantic categories cited by humans. The most noteworthy omission is in regards to emotion, as terms relating to affect and evoked emotional response were some of the most frequently cited terms for both musical and poetic creativity. Accordingly, future work will investigate why the model does not capture this cluster of emotion-related terms.

Further directions for the future include the application of this computational approach to other domains, such as "culinary creativity", both for the ontologically useful task of elaborating concepts themselves, and to create well-tailored terminology for the assessment of creative output from the corresponding domains. This methodology may also be used to approach the task of conceptual blending: rather than specifying input vectors that belong to only one concept, one may supply input dimensions from several. This could result in output terms discovered at the intersection of the lexical regions specified by the vectors' different input dimensions.

Acknowledgments

The projects Lrn2Cre8 and ConCreTe acknowledge the financial support of the Future and Emerging Technologies (FET) programme within the Seventh Framework Programme for Research of the European Commission, under FET grants number 610859 and 611733, respectively. This research is also supported by EPSRC grant EP/L50483X/1.

References

Allott, N., and Textor, M. 2012. Lexical pragmatic adjustment and the nature of ad hoc concepts. *International Review of Pragmatics* 4(2).

Barsalou, L. W. 1993. Flexibility, structure, and linguistic vagary in concepts: Manifestations of compositional system of perceptual symbols. In Collins, A. F.; Gathercole, S. E.; Conway, M. A.; and Morris, P. E., eds., *Theories of Memory*. Hove: Lawrence Erlbaum Associates. 29–101.

Clark, S. 2015. Vector space models of lexical meaning. In Lappin, S., and Fox, C., eds., *The Handbook of Contemporary Semantic Theory*. Wiley-Blackwell.

Fauconnier, G., and Turner, M. 2008. *The way we think: Conceptual blending and the mind's hidden complexities*. Basic Books.

Gärdenfors, P. 2000. *Conceptual Space: The Geometry of Thought*. Cambridge, MA: The MIT Press.

Grice, H. P. 1969. Utterer's meaning and intention. *The Philosophical Review* 78(2):147–177.

Harris, Z. 1957. Co-occurrence and transformation in linguistic structure. *Language* 33(3):283–340.

Heath, D.; Norton, D.; Ringger, E.; and Ventura, D. 2013. Semantic models as a combination of free association norms and corpus-based approaches. In *IEEE International Conference on Semantic Computing*, 48–55.

Hesse, M. B. 1963. *Models and Analogies in Science*. New York: Sheed and Ward.

Hill, F.; Korhonen, A.; and Bentz, C. 2014. A quantitative empirical analysis of the abstract/concrete distinction. *Cognitive Science* 38:162–177.

Jäger, G. 2009. Natural color categories are convex sets. In *Logic, Language and Meaning - 17th Amsterdam Colloquium*.

Koestler, A. 1964. *The Act of Creation*. : Hutchinson.

Landauer, T.; Laham, D.; Rehder, B.; and Schreiner, M. E. 1997. How well can passage meaning be derived without using word order? a comparison of latent semantic analysis and humans. In *Proceedings of the 19th Annual Conference of the Cognitive Science Society*, 412–417.

Mikolov, T.; Chen, K.; Corrado, G.; and Dean, J. 2013. Efficient estimation of word representations in vector space. In *Proceedings of ICLR Workshop*.

Ontanón, S.; Zhu, J.; and Plaza, E. 2012. Case-based story generation through story amalgamation. In *Proceedings of the ICCBR 2012 Workshops*, 223–232. Citeseer.

Pennington, J.; Socher, R.; and Manning, C. D. 2014. Glove: Global vectors for word representation. In *Conference on Empirical Methods in Natural Language Processing*.

Schütze, H. 1992. Dimensions of meaning. In *Proc. ACM/IEEE Conference*, 787–796.

Turney, P. D., and Patel, P. 2010. From frequency to meaning: Vector space models of semantics. *Journal of Artificial Intelligence Research* (37):141–188.

Veale, T. 2006. An analogy-oriented type hierarchy for linguistic creativity. *Knowledge-Based Systems*.

Wittgenstein, L. 1953. *Philosophical Investigations*. Oxford: Basil Blackwell. 3rd ed. Trans. G. E. M. Anscombe.