

Constructing conceptual spaces for novel associations

Kazjon Grace¹, Rob Saunders¹, and John Gero²

¹ University of Sydney, New South Wales, Australia.

² Krasnow Institute for Advanced Study, Virginia, USA.

Abstract. This paper reports on a system for computational analogy-making based on conceptual spaces. The system constructs conceptual spaces that express the relationships between concepts and uses them to build new associations. A case for this conceptual-space driven model of association making is made, and its advantages and disadvantages are discussed. A prototype space-construction system is detailed and one method by which such a system could be used to make associations is proposed. The system forms concepts that are useful to describe a set of objects, then learns how those concepts relate to each other. These relationships can then be used to construct analogies.

1 Introduction

The generally accepted frameworks (French 2002, Kokinov 1998) for computational analogy-making focus on three processes: representation, mapping and transfer. Representations of a source and target object are constructed, mappings are built between them and then knowledge is transferred from the source to the target. Existing models of the representation process (Hofstadter & Mitchell 1994, Kokinov 1994) build representations out of a set of provided components. Mappings produced by these systems must be constructed (by processes such as conceptual slippage or spreading activation) from relationships existing between those components. While the scope of representations in the system can be broad, all possible kinds of relationships between representations must be provided with the representational components. Representation in analogy-making systems with a fixed set of representational components is reduced to “choosing” between which of the pre-encoded relationships will underlie the mapping.

This research investigates an approach to computational association that overcomes this restriction: a system that constructs the conceptual space in which it performs representation. If a system builds the relationships between its concepts through use, then potential avenues for mapping between those concepts need not be pre-encoded. We detail a system which learns concepts to describe its world, learns how those concepts relate, constructs a space using those relationships, and then can find mappings through the reinterpretation of objects in that space. In other words, a system in which the associations made are not just expressed in the representations constructed but are situated

in the system’s experientially-derived conceptual space. Our hypothesis is that this increased autonomy in representation and mapping will aid in producing potentially creative analogies.

2 Association

This research defines association as the process of constructing a new mapping between two objects. The process involves identifying a match and building a mapping between the two objects that reflects that match. This process is fundamental to analogy-making, metaphor and other related tasks. We assume that pattern recognition makes recognising mappings in existing representations virtually automatic. From this assumption that associating two objects is fundamentally a process of re-representing the objects to express a connection between them. This is our notion of interpretation-driven association.

2.1 Interpretation-driven association

Association as interpretation-driven search has several benefits for an analogy-making system. Multiple associations between the same objects are possible through the development of multiple interpretations of those objects. Each association is situated within the interpretation used to construct it, and any knowledge learnt or transferred through that mapping is also specific to that interpretation. When we say that an association embodies a “new” match, we mean that the mapping between the representations produced by the association process is not previously known to the agent: it is s-creative (Suwa et al. 1999).

The interpretation process involves concurrent re-representation of both objects and a search of the agent’s experiences with both until a way of representing them can be negotiated. In a system governed by this idea of association it must be possible to produce many different representations of one object. We model this by allowing the concepts used to represent objects to have mutable meanings through a process analogous to the “conceptual slippage” in the Copycat system (Hofstadter & Mitchell 1994). In the right circumstances, the meanings of two concepts can “slip” together, allowing previously disparate objects to be matched. In Copycat, these slippages can only happen along predefined paths and in predefined situations. Our association system is freed from this constraint as it autonomously develops the relations between concepts that cue the “slippage” process.

Our goal is to produce an analogy-making system that builds representations out of concepts that it has learnt, but also to learn relationships between those concepts that allow it to “slip” their meanings without predefined paths. To do this requires the solution of two problems: we need to learn relationships between the concepts produced by the system, and we need to use those relationships to produce new interpretations and thus associations.

2.2 Conceptual spaces as a model of experience

In this research we use the notion of conceptual spaces to describe how concepts relate to each other and how those relationships can be used in reinterpretation. A conceptual space is an abstract construct in which all the concepts of a system are located. A conceptual space contains knowledge about how concept meanings relate to each other and about how concepts have been used in conjunction with each other. The conceptual space is an abstraction of a system’s experiences over the course of its operation and it can be used to put the act of perceiving an object in the context of the agent’s past. Our system re-interprets objects by drawing on this knowledge of related past experiences to find another set of concepts that can be used to describe the object.

Conceptual spaces for analogy-making must contain rich and interrelated descriptions of patterns and features that comprise objects. It is not sufficient to produce a conceptual space in which each object is represented by a single point as the space must express relationships between the concepts used to describe objects, not between the objects themselves. Gärdenfors’ (2000) “theory of conceptual spaces” states that conceptual spaces are defined by quality dimensions, or “aspects or qualities of the external world that we can perceive or think about”. If the relationships in a space can be expressed in terms of a few quality dimensions then any mapping produced within the space will be derived from those few qualities. Our definition of conceptual spaces does not imply that the spaces contain any globally coherent organisation.

The mechanism governing the location of concepts in space varies by implementation, but at minimum our definition states that proximal concepts are in some way similar. In our system the spaces are defined by undirected multi-graphs, with each node being a concept and each edge being a relationship. Some idea of the similarity between concepts can be gained through the edge distance between any two concepts, but as each edge can represent different kinds of relationships there is no notion of moving in a defined “direction” in the space.

Concept-to-concept relationships can be learnt through how the system acquires and uses the concepts. Relationships in conceptual space in our prototype take two forms; similarity between the meanings of concepts and similarity between the usage of concepts. We can then use these relationships to perform reinterpretation of objects.

2.3 Matching in conceptual spaces

Each object can be represented within the conceptual space as a set of nodes, one for each of the concepts that describe it. These concepts form a region in conceptual space that describes the object. Finding a way to reinterpret the concepts used in this representation involves finding another region of concepts that can be mapped onto this one. When two regions in conceptual space are mapped, one describing a source object and one describing a target object, it can be said that the concepts within those representations have had their meanings

“slipped” together. This results in representations of the two objects that reflect an association produced between them.

If a structural similarity exists between the conceptual regions associated with two objects, then we can say that it is possible to view the ways the system models those two objects as alike. Once a mapping between regions is found, we can produce an interpretation of one object using the concepts associated with the other. The structural similarity between two conceptual regions is indicative of how the system’s experiences with those two objects have had similar structure; there are concepts in both regions playing similar roles, and with similar patterns of relationships with their neighbours. This approach is syntactic in that it matches on the structure of conceptual space rather than its content, but that structure is learnt through the system’s interactions with its world. Therefore what is being mapped is, at the object level, still semantic information.

This research is concerned with developing a system that can both learn its own concepts and learn how those concepts relate to each other. The more removed the experimenter-provided data is from the analogies being made by the system, the more defensible is the claim that the system has autonomously constructed a new association. A system based on these principles would a) learn a set of concepts to describe the objects in its world, b) learn how those concepts relate to each other in both definition and usage, c) construct a conceptual space embodying the relationships between concepts, d) find a match between the structure of the regions in conceptual space that reflect the target object and a source and e) interpret the target and source objects to reflect the mapping that has been constructed between the concepts used to describe them.

We have developed a prototype of our approach to association construction that implements concept formation, conceptual interrelation, conceptual space construction and a limited form of matching. While this prototype does not yet produce compelling or interesting analogies, it serves as a proof of concept for our framework and its behaviours offer some insight into our theories.

3 A System for Constructing Spaces

We have developed a system capable of constructing conceptual spaces that is designed to be applicable to analogy-making. An overview of the system can be seen in Figure 1. The system takes a set of objects, learns concepts to describe them, learns relationships between those concepts based on literal similarity and co-occurrence, constructs a graph of those relationships and then finds a possible mapping.

The system operates in a very simple shape perception domain, from which it accepts symbolic perceptual input about objects. A future development goal is for the system to take lower level sensory input and learn its own perceptual representations of objects, but symbolic input is sufficient for the purpose of testing the construction of spaces. The system then learns a set of concepts that can uniquely describe each of the objects, using a method based on the discrimination agent architecture developed by Steels (1996) and described in Section

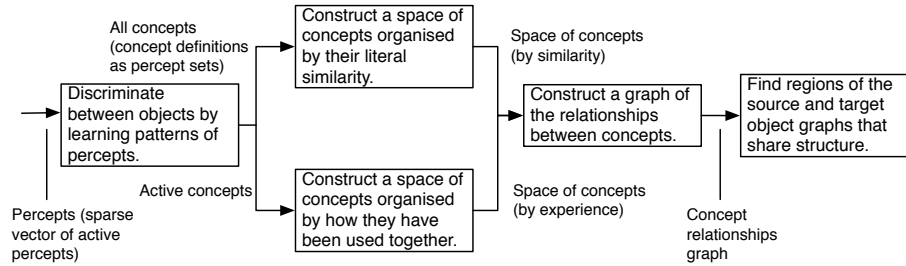


Fig. 1. A diagram of our system, showing the process from perceptual input on the left to the generation of possible matches on the right.

3.2. The similarity relationships between concepts are then calculated based on which percepts they include, while the experiential relationships between concepts are calculated based on which concepts co-occur with each other. These relationships are extracted from the set of concepts using the singular value decomposition process from Sarkar et al. (2009), which is described in Section 3.3. Concepts that are sufficiently similar by either of these metrics are judged to be related and an edge connecting them is added to the conceptual space graph. This graph can then be searched for matching sub-regions.

3.1 Our domain

The Line Grid domain used in this research is designed to be a simple visual way to investigate concept formation and space construction. The emphasis is not on the potential for interesting associations, but on the utility for testing conceptual space construction. A line grid of size n is an n -by- n grid of points that can be connected to any other point orthogonally or diagonally adjacent to them. Figure 2 shows four objects in the line grid of size three. Sufficient versatility exists in this domain to describe polygonal shapes, isometric depictions of 3D objects, line patterns and a simple but complete typeface of capital letters. A shape in a line grid is described by a binary string indicating which of the possible edges exist in that shape. Our system has been tested for size three and four line grids, which have twenty and forty-two possible edges respectively.

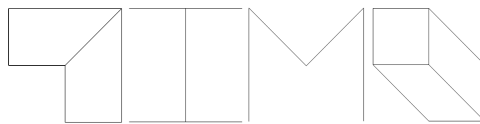


Fig. 2. A set of example objects in a 3x3 version of our line grid shape domain.

Concepts in this domain are patterns of edge presence and absence that exist in multiple shapes. Relationships between these concepts show how those concepts are similar (identifying similar patterns of edges), or how those concepts are used (identifying that they form discriminating sets together).

To take the example of a set of 26 objects representing the capital alphabet, these relationships include things such as “objects containing an enclosed space in the top half of the letter” being used together with “objects containing a stroke down the left side”, as in the letters P, B and R. These relationships would then be compiled into a conceptual space expressing the patterns of relationships between the concepts learnt by the system to describe the capital alphabet in the Line Grid domain. The system would then look for matches in the structure of regions of the conceptual space; areas in which different concepts play the same “role” in their groups of related concepts. If a group of concepts can be found that shares structure with the group that identifies with a target object, then another object that is described by that group may be a potential source. An example of a proportional analogy that could be made in the Line Grid domain by a complete analogy-making system is seen in Figure 3. Given letters in a consistent typeface, the system would find that similar structures existed between pairs of letters. In this case, the difference between the letters ‘I’ and ‘T’ could be considered analogous to the difference between the letters ‘F’ and ‘E’.

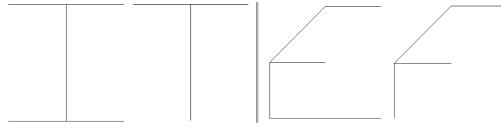


Fig. 3. Two examples of matches between pairs of objects in the domain that could be found by a complete analogy-making system and expressed as a proportional analogy of the form “I is to T as F is to E”.

3.2 Concept formation

Our prototype concept formation system is designed to produce sets of concepts that are suitable for association in conceptual space. It is desirable that each object is described by many concepts in order for conceptual spaces to be more interesting and potential matches to be more varied. An Accuracy-Based Classifier System (Wilson 1995) was chosen as the concept learning algorithm due to its ability to extract patterns from representations and thus produce many concepts per object. Concepts produced by the system represent patterns of percepts that exist in objects that are useful for telling those objects apart from their peers. Concepts use a similar representation to objects, except that they

are trinary strings as each concept may require an edge, require the absence of an edge, or not care either way.

The concepts are evolved to be able to discriminate an object from all others in the given set. Learning about a set of objects via attempting to tell them apart is a common approach to concept formation and is described in Steels (1996), where the discrimination occurs for the purpose of a set of agents trying to co-operatively learn language. The principle has been applied to an analogy-making system based on the idea that it must first be possible to tell objects apart before any interesting ways can be found to put them together. Concepts can be combined together to discriminate a chosen object from its context, with each concept discriminating that object from one or more other objects. This set-based reinforcement method means that each individual concept will be rewarded if a discriminating set exists that it is a part of. As the goal is to produce a rich set of general concepts, there are no limits on the size of each set or the number of discriminating sets that can be found: this promotes the reinforcement of multiple divergent approaches to discriminative success.

The classifier system was able to find a stable and compact set of general concepts to describe up to 100 objects in the 4x4 line grid domain. A plot of the system's performance over 10,000 generations on a twenty object problem in the 4x4 domain can be seen in Figure 4. The system reached 100% discriminatory success after 1,300 steps with approximately 600 concepts, but the population continued growing to 3,950 concepts after 6,000 steps. The system then reached a saturation point where no new ways of discriminating each object could be found and began rapidly finding more general discriminators. After approximately 8,000 steps the system had found 125 general concepts and maintained 100% discrimination rates. The generalisation can be seen in the second data series, with the average number of objects matched per concept rising to 2.5 with the generalisation process.

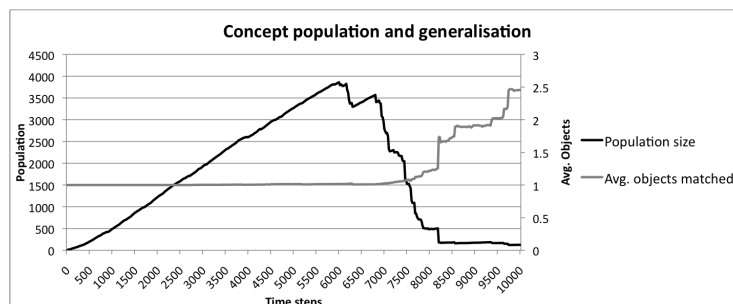


Fig. 4. The results of a run over 10,000 generations with a 4x4 grid and 20 random objects. The number of discriminating concepts is shown at the left, while the average number of objects that each concept matches is shown on the right.

3.3 Inter-conceptual relationships

The construction of conceptual spaces is dependent on the system’s ability to form relationships between the concepts that it has learnt. In our system we have identified two kinds of conceptual relationship to model: experiential co-occurrence, or when two concepts are used together in discrimination tasks, and literal similarity, or when two concepts are similar in the kinds of objects they describe. Experiential co-occurrence relationships are designed to allow the association system to match between concepts that are “used” the same way: to play a role in their group of concepts that is analogically equivalent to the role played by the source concept in its group. Similarity relationships are designed to allow the system to match between the pattern of differences that exist between the meanings of concepts in the two conceptual groups.

The intent is to form a conceptual space graph where each relationship is described as either literal or experiential, and is further labelled by the difference between the meanings of the concepts it connects. The structure of a region in conceptual space would then be described by the structure of differences between its concepts. Similarly structured regions could then be found that contain potential 1-to-1 mappings between concepts that play the same “role” in their local area of conceptual space.

We employ Singular Value Decomposition (SVD), an algorithm that has been used in statistical natural language processing to induce implicit connections between words in a corpus. In our work SVD calculates connections between the meanings and usages of concepts the system has learnt. The experiential co-occurrence is calculated by running the SVD algorithm on a co-occurrence matrix of concepts in discrimination sets. The literal similarity is calculated by running the SVD algorithm on a matrix of concept definitions in terms of which grid line edges they match and which they forbid. The advantage of the SVD approach in calculating literal similarity is that the algorithm is able to extract which grid lines represent important differences between concepts and reflect that accordingly, which the use of a literal distance measure would not do.

3.4 Constructing spaces

The space construction process takes the relationships identified by the SVD engine and compiles them into a coherent graph representation that can then be searched for matches. In the current prototype conceptual graph edges are labelled only as “similarity” or “co-occurrence”. Future versions of this system will label edges by how the concepts differ and matching will behave accordingly. The current system is able to see patterns and structures in the body of concepts learnt by the system, but not the specifics of how those patterns relate to each other, beyond the kind and number of edges assigned to each node.

The correlations between concepts using the two metrics produced by the SVD algorithm are compared to a threshold and sufficiently similar concepts are assigned an edge of the appropriate type. This binary threshold system is a placeholder for the ability for edges in conceptual space to reflect the strength of

the inter-conceptual relationship. This would allow the system to preferentially match between strongly related concepts but to search weaker links if no strong mappings were found. An example of part of a simple graph produced by the system can be seen in Figure 5. This graph shows some of the concepts learnt discriminating a small set of objects. There are two broad groups of literally related concepts connected by solid lines and between those groups are concepts connected with dashed lines indicating co-occurrence.

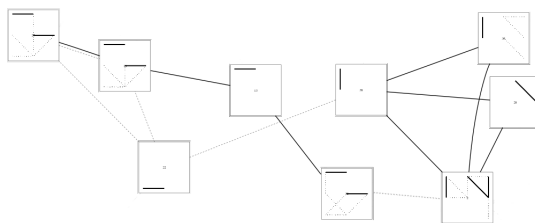


Fig. 5. Part of a graph describing relationships between concepts. Solid edges indicate concepts that are literally similar, while the dotted edges indicate co-occurring concepts.

4 Discussion

Conceptual relations and conceptual spaces can be constructed in the course of learning to describe a set of objects. We have performed simple matching between groups of concepts in the constructed spaces, but more interesting associations will require a richer description of concept relationships in conceptual space. The current system can only match between relations labelled as “similarity” or as “co-occurrence”. Much richer information about the nature of the relationships between concepts exists in the singular values produced by the SVD system. A detailed set of relations extracted from the singular values will permit a more complete labelling of edges in conceptual graphs. Edges between related concepts can be labelled by what differs between them, allowing for matches to other concept groups with a similar pattern of differences. Association in the resulting conceptual space will then involve subgraph isomorphism between the labelled graphs; mapping between groups of concepts with similar patterns of relationships between them.

Like many concept formation systems, learning of concepts in our prototype system is grounded in the ability to discriminate between objects. Our system produces a set of general concepts to identify each of a set of objects by how it is different from its peers. As a result the graphs produced by our system represent the ‘similarity between differences’ and the ‘co-occurrence of similar differences’. What is necessary for analogy-making is to extract common sub-components that when combined describe the objects themselves rather than describe the

differences between objects. Therefore discrimination-based concept formation may not be suitable for analogy-making systems.

We have described the benefits of an analogy-making system that constructs its own conceptual spaces. In order to operate as a functional analogy-making system the prototype described here requires additional features, most notably the ability to evaluate potential mappings both in terms of analogical quality and how they relate to previous analogies made by the system. With more detailed conceptual space construction and a revised concept formation process such a system could produce interesting and potentially creative analogies.

References

- French, R.: The computational modelling of analogy-making, *Trends in Computer Science*, **6**: 200–205, (2002).
- Gärdenfors, P.: *Conceptual Spaces: The geometry of thought*, MIT Press, (2000).
- Hofstadter, D. and Mitchell, M.: The copycat project: a model of mental fluidity and analogy-making, in Holyoak, K. and Barnden, J. (eds), *Advances in Connectionist and Neural Computation Theory*, **2**: 31–112, (1994).
- Kokinov, B.: A hybrid model of reasoning by analogy, in K Holyoak and J Barnden (eds), *Analogical Connections, Advances in Connectionist and Neural Computation Theory*, **2**: 247–318, (1994).
- Kokinov, B.: Analogy is like cognition: dynamic, emergent and context sensitive, *Advances in Analogy Research* NBU Press, Sofia, (1998).
- Sarkar, S., Dong, A. and Gero, J. S.: Design optimization problem (re)formulation using singular value decomposition, *Journal of Mechanical Design* **131**(8) (to appear), (2009).
- Steels, L.: Perceptually grounded meaning creation, in *Proceedings of the Second International Conference on Multi-agent Systems (ICMAS096)* Los Alamitos, CA, pp 338–344, (1996).
- Suwa, M., Gero J. S. and Purcell, T.: Unexpected discoveries and s-inventions of design requirements: A key to creative designs, In: J. S. Gero and M. L. Maher (eds) *Computational Models of Creative Design IV*, University of Sydney, Australia, pp 297–320, (1999).
- Wilson, S.W.: Classifier Fitness Based on Accuracy. *Evolutionary Computation*, **3**(2): 149–175, (1995).