Constructing Conceptual Spaces for Novel Associations

Kazjon Grace[1], Rob Saunders[1], and John Gero[2]

[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 [1, 2] 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 [3, 4] 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 relationship between representations must be provided with the representational components. Representation in analogy-making systems with a xed 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 addresses 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 that learns concepts to describe its world, learns how those concepts relate, constructs a space using those relationships, and then can nd 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 de nes 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 re ects 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 we derive 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

Modelling association as an interpretation-driven search has several bene ts 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 speci c to that interpretation. Each association embodies a new match, in that the association process produces a mapping between representations that was not previously known to the system: it is s-creative [5].

The interpretation process involves concurrent re-representation of the objects via a search of the system s experiences with them until a viable representation can be found. In a system governed by this idea of association it must be possible to produce many di erent representations of one object. We model this by allowing the concepts used to represent objects to have mutable meanings through a process analogous to conceptual slippage in the Copycat system [3]. In appropriate circumstances, the meanings of two concepts can slip together, allowing previously disparate objects to be matched. In Copycat, these slippages can only happen along prede ned paths and under prede ned circumstances. Our association system is freed from this constraint as it autonomously develops the relations that cue the slippage process between concepts.

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. This would allow the system to slip the meanings of concepts without prede ned paths along which to do so. 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 association. 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 a system s past. Our system re-interprets objects by drawing on this knowledge of related past experiences to nd another set of concepts that can be used to describe the object.

Conceptual spaces for analogy-making must contain rich and interrelated descriptions of the features that comprise objects. It is not su cient 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. Gardenfors theory of conceptual spaces [6] states that conceptual spaces are de ned 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 de nition 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 de nition states that proximal concepts are in some way similar. In our system the spaces are de ned by undirected multigraphs, 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 di erent kinds of relationships there is no notion of moving in a de ned direction in the space.

Concept-to-concept relationships can be learnt through how the system acquires and uses 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 use these relationships to reinterpret 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 nding another region of concepts that can be mapped onto this one. When two regions in conceptual space are mapped onto one another, 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 re ect an association between them.

If a structural similarity exists between the conceptual regions associated with two objects, then the ways the system models those two objects can be seen as alike. Once a mapping between the concepts in two 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. We can say that there are concepts in both regions playing similar roles within that group of concepts, 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 semantic information at the object level expressed as structural information within the conceptual space.

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 de nition and usage, c) construct a conceptual space embodying the relationships between concepts, d) nd a match between the structure of the regions in conceptual space that re ect the target object and a source and e) interpret the target and source objects to re ect 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 o er some insight into our theories.

3 A System for Constructing Spaces

We have developed a system capable of constructing conceptual spaces for 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, constructs a graph of those relationships and then searches for possible mappings within that graph.

The system operates in a very simple shape perception domain from which it receives 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 su cient 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 [7]. Discrimination-based

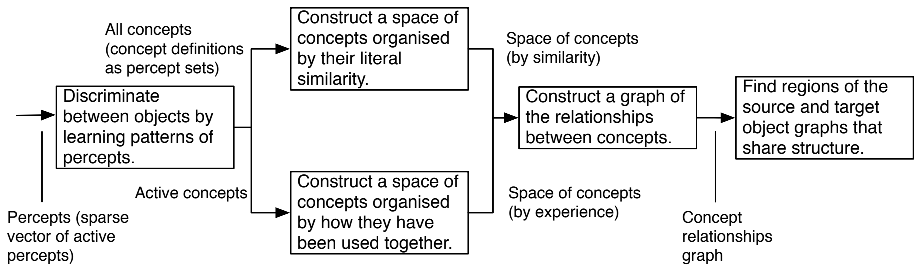

Fig. 1. A diagram of our system, showing the process from perceptual input on the left to the generation of possible matches on the right.

learning was chosen for its simplicity and prevalence as a reinforcement strategy in concept formation.

Similarity relationships between concepts are then calculated based on shared percepts, 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 described in Sarkar et al. [8]. This method extracts an underlying set of structurally important vectors from the concept usage and de nition data and then describes individual concepts in terms of those vectors. Concepts with similar composition in this singular value representation are similar in ways that are signi cant in the dataset. Concepts that are su ciently similar by either the literal or co-occurence 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 Example 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 each be connected to any other point orthogonally or diagonally adjacent to them. Figure 2 shows four objects in the size three line grid. Su cient 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 line grid shape 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.

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).

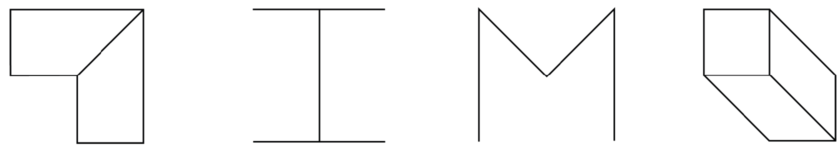

Fig. 2. A set of example objects in a 3x3 version of our line grid shape domain.

For example, in the 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 other concepts play the same "role" in their groups of related concepts as the source object's concepts do in its conceptual region. If a group of concepts can be found that shares structure with the group that describes the target, 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'.

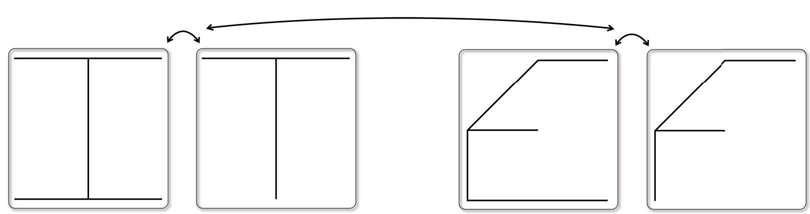

Fig. 3. Two examples of matches between pairs of objects in the domain that could be found by a complete analogy-making system and expressed as a proportional analogy of the form I is to T as F is to E .

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 for potential matches to be more varied. An Accuracy-Based Classifier System [9] modified to reinforce based on discriminative success was chosen as the concept learning algorithm. This algorithm was chosen 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 are useful for telling objects apart from their peers. Concepts use a similar representation to objects but are defined as trinary strings as each concept may require, forbid or not care about each edge in the grid.

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 [7], 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 it is a part of any discriminating set. 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 development 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 enough diversity existed in the population to subsume most new classifiers into existing more general ones and the population rapidly declined. 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.

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-occurence, or when two concepts are used together in discrimination tasks, and literal similarity, or when two concepts describe similar properties of objects. Experiential co-occurence relationships are designed to allow the association system to match between concepts that are "used" the same way: concepts that play a role in their group of concepts that is analogically equivalent to the role

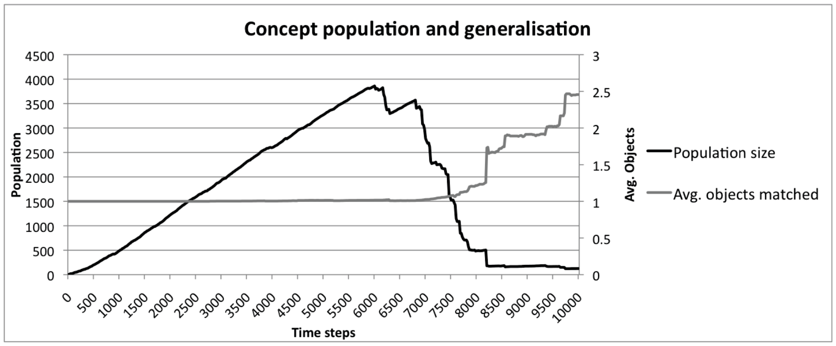

Fig. 4. The results of a run over 10,000 generations with a 4x4 grid and 20 random objects. The population of concepts is shown at the left, while the average number of objects that each concept can be used to describe is shown on the right.

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.

A conceptual space graph is formed where each relationship is described as either literal or experiential and is labelled by the difference between 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 can then be found that contain potential mappings between pairs of concepts that play the same "role" in their local area of conceptual space.

We employ Singular Value Decomposition (SVD), a linear algebra method with uses in statistical natural language processing, data mining and signals processing. In our work SVD calculates connections between the meanings and usages of concepts the system has learnt. The experiential co-occurence is calculated by running the SVD algorithm on a co-occurence 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-occurence". Future versions of this system will label edges by how the concepts differ. The current system is able to see patterns and structures in the body of concepts learnt by the system, but not the speci cs of how those patterns relate to each other beyond the kind and number of relationships involving each concept.

The correlations between concepts using the two metrics produced by the SVD algorithm are compared to a threshold and su ciently similar concepts are assigned an edge of the appropriate type. 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-occurence.

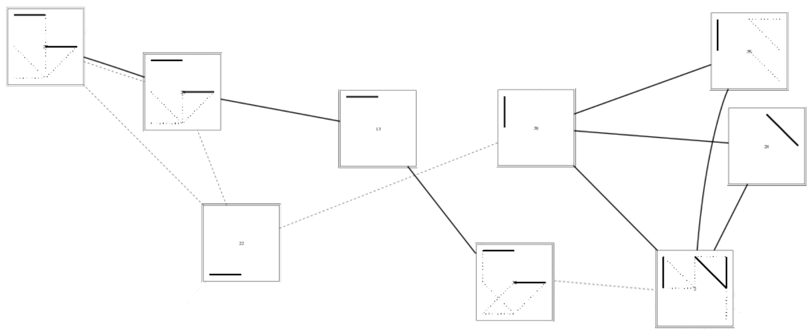

Fig. 5. Part of a graph describing relationships between concepts. Solid edges indicate concepts that are literally similar, while the dotted edges indicate co-occuring concepts.

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 constructed spaces, but producing more interesting associations in these spaces will require a richer description of concept relationships. The current system can only match between relations labelled as similarity or as co-occurence . 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 di ers between them, allowing for matches to other concept groups with a similar pattern of di erences.

Incorporation of a con dence attribute for relationships (the data for which exists in the SVD output) would allow the system to preferentially match between strongly related concepts but to search weaker links if no strong mappings were found. Association in the resulting conceptual space would then involve subgraph isomorphism between the labelled graphs; mapping between groups of concepts with similar patterns of relationships between them, with each relationship de ned by its type, strength and the speci cs of the di erence between its concepts.

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 di erent from its peers. As a result the graphs produced by our system represent the similarity between di erences and the co-occurrence of similar di erences . What is necessary for analogy-making is to extract common sub-components that when combined describe the objects themselves rather than describe the di erences between objects. Therefore discrimination-based concept formation may not be suitable for analogy-making systems.

We have described the bene ts of an analogy-making system that constructs its own conceptual spaces. In order to operate as a complete 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

1. French, R.: The computational modelling of analogy-making, *Trends in Computer Science*, **6**: 200–205, (2002).
2. Kokinov, B.: Analogy is like cognition: dynamic, emergent and context sensitive, *Advances in Analogy Research* NBU Press, Sofia, (1998).
3. 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).
4. 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).
5. 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).
6. Gärdenfors, P.: *Conceptual Spaces: The geometry of thought*, MIT Press, (2000).
7. Steels, L.: Perceptually grounded meaning creation, in *Proceedings of the Second International Conference on Mullti-agent Systems (ICMAS096)* Los Alamitos, CA, pp 338–344, (1996).
8. 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).
9. Wilson, S.W.: Classier Fitness Based on Accuracy. *Evolutionary Computation*, **3**(2): 149–175, (1995).