A Computational Model of Novel Association

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

[1] University of Sydney, Sydney, Australia.
[2] Krasnow Institute for Advanced Study, Virginia, USA and University of Technology, Sydney, Australia.

Abstract. This paper presents a computational model for the construction of novel associations as a component of a larger project on analogy-making. Association-construction is driven by a reinterpretation-based model of subjective similarity. Associations are constructed by transforming the way the agent perceives the objects being associated so that they become similar. The paper describes a model where an agent develops associations by an iterative process of attempting to relate objects in its environment by building appropriate transformative interpretations. Possible methods for the learning of transformations are discussed. The capabilities and implications of the model are discussed through an example application to the domain of geometric proportional analogies.

1 Introduction

Computational models of analogy-making have paid significantly more attention to the creation of mappings between existing representations than to the issue of how such compatible representations arose. Reviews of the field of analogy-making (French 2002, Kokinov 1998) discuss the need for analogical research to explore the integration of analogy into cognitive processes for the production of relevant, useful and creative analogies. Creative analogies must by definition be based on associations that are novel. This paper presents a computational model for the construction of novel associations as a step towards the computational production of creative analogies.

This research defines an association as an interpretation of two objects which expresses a relationship between them. An analogy is defined as association-based but additionally incorporating the transfer of knowledge from the source to the target driven by the goals of the analogy-making attempt. A critical feature of associations (and by extension analogies) is that they are situated; they exist within a specific way of thinking about both the objects being associated and the world they are in. Associations between objects are highly subjective and dependent on an agent possessing particular representations of those objects.

In this paper we assume that the relationship between two objects is defined by the manner in which those two objects are being interpreted. It follows that the problem of constructing a new relationship between the objects can be modelled as the problem of constructing an interpretation of them such that they can be associated. In other words the critical problem in finding new associations becomes the search for appropriate interpretations rather than the search for possible mappings within interpretations.

A method for describing interpretation as a transformation is described with reference to a simple example in the domain of colour. A computational model for constructing novel associations, the "association game", is developed based on transformational interpretation and described formally. The model uses an iterative, experience-driven process of constructing interpretations of objects in which a new relationship can be perceived. An example association game is presented in the domain of geometric shapes. The place of this model in analogy-making and creativity research is discussed.

2 Interpretation-driven association

An interpretation is a perspective on the meaning of an object; a particular way of "looking" at an object. The process of interpretation involves the construction of that perspective within a situation. This process is described in this research using the theory of conceptual spaces (Gärdenfors 2000), in which concepts exist as regions in a conceptual space. A conceptual space is an abstract construct defined by a number of quality dimensions, or "aspects or qualities of the external world that we can perceive or think about" (Gärdenfors 1990), for example 'weight' or 'time'. Quality dimensions have associated spatial structures; weight is one-dimensional with a zero point and is isomorphic to the half-line of non-negative numbers, while hue is circular. These structures impart topological properties on regions; it is meaningful to talk about "opposite" colours but not "opposite" weights.

We represent interpretation as a transformation applied to a conceptual space. As concepts are represented by regions of the conceptual space, the application of a transformation to the space causes different features to be identified with objects within it. A new association is constructed when transformations are applied to one or both objects so that some shared features, that were not present before the transformations were applied, emerge. In transforming the conceptual space as it applies to the objects in the association the agent has created a new interpretation of the objects in which a mapping is possible.

The notion of learning a transformation to solve an association problem can be likened to the approach used in proportional analogy problems of the form $A : B :: C :?$, in which a transformation between A and B must be learned in a way that it can be generalised to C and used to produce D. This research focuses on the construction of the relationship on the left hand side of proportional analogy problems, but differs from previous proportional analogy systems (Evans 1964, Mullaly and O'Donoghue 2006) in that relationships are constructed rather than selected.

Figure 1 shows a very simple example of a transformative association in the Hue, Saturation, Brightness space of colours. In the unmodified diagram on the left the two colours, while both greenish in hue, are clearly different. In the reinterpreted diagram on the right the source object has been transformed by "swapping" the values in the saturation and brightness dimensions. Depending on the specificity of the concepts possessed by the agent they could now be both identified as "dark green". The interpretation that is being used to construct this association could be phrased in English as "treat saturation as brightness and brightness as saturation and the source and target can be thought of as alike".

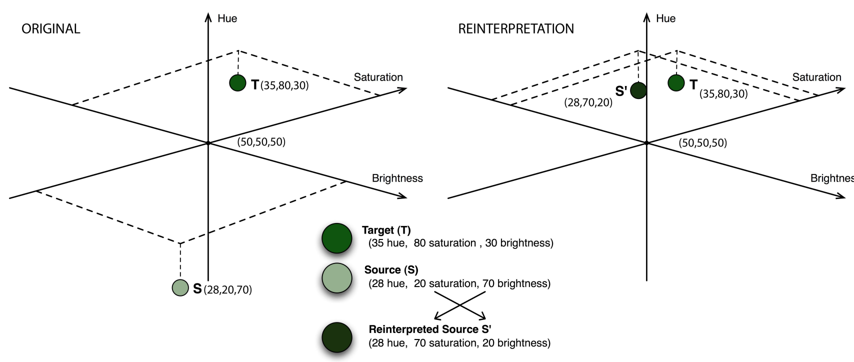

Fig. 1. An example of a transformation that enables an association. The Saturation and Brightness dimensions have been exchanged and the reinterpreted source can be related to the target.

3 The association game

Producing interpretation-based associations is an iterative process that produces multiple "failed" interpretations - those that do not produce appropriate mappings or do not produce mappings at all. Learning from these failed associations enables the construction of additional associations driven by past experiences. In this research we model association-construction through the construct of the "association game" an iterative experience-based process of reinterpretation. The game is based on an agent architecture developed by Steels (1996). Steels developed a way for agents to learn a grounded model of their world in the course of attempting to distinguish objects from their surroundings. Agents played a series of "discrimination games" in which they constructed feature detectors that were suited to identifying an object from its context. The agent constructs new feature detectors or refines existing ones as necessary to distinguish between objects. Through repeatedly learning to identify each object from its context the discrimination agent learns about all the objects in its world.

This model adapts the discrimination agent architecture for an alternative purpose: learning associations grounded in use. In this model an agent attempts to learn about its world by creating new ways to relate objects. The agent initially attempts to discriminate between objects, but if it is capable of telling two objects apart it will attempt to reinterpret them in such a way that a new relationship between them is created. The interpretation that enabled the new association persists the next time a game is played between the same two objects and as a result the discrimination phase of that next game will differ. Successive games played with the same objects will lead to sets of associations composing a coherent mapping between the source and target.

As multiple games occur with the same objects and new interpretations supersede old, the agent must evaluate whether the mappings enabled by the new interpretation warrant keeping. As the current model does not incorporate analogical transfer this evaluation cannot be performed based on the reason the agent is making the association. An association is kept if it is more valuable than the existing mappings (if any) according to evaluation criteria specific to the domain and problem. A preliminary model of this evaluation process will be based on the number of features that an association enables the mapping of.

The model as described here focuses on the relationships that can be constructed between objects, not the discovery of potential sources. The associations are generated between pairs of objects and over time the agent develops associations for possible pairs in its environment. The process of interpretation used by the system in constructing associations is situated in that interpretations are driven by a world view that is constructed from previous experiences. The agent attempts to reinterpret the source in ways that have worked in the past for similar problems. For each association game the agents memory constructs associations that are believed to be similar to the current problem and attempts to apply them. The features that the agent learns to detect are modelled as regions in the agents conceptual space and an association is modelled as a transformation within that space that produces an overlap between the features of two objects.

3.1 Formal description

An agent a exists in a world containing a set of objects $O = \{o_1, \ldots, o_n\}$. The agent possesses a set of sensory channels $\Sigma = \{\sigma_1, \ldots, \sigma_m\}$ which are functions over O. Each sensory channel σ_j defines a value for each object o_i. An agent has a set of feature detectors $S_a = (s_{a,1}, ..., s_{a,p})$ which is initially empty. A feature detector $s_{a,k}$ consists of a function $\phi_{a,k}$ over a sensory channel σ_j that is updated after each game. The feature set derived by applying the feature detectors of an agent a to an object o_i is defined as F_{a,o_i}. A distinctive feature set D^C_{a,o_i} is a set of features that serves to uniquely identify an object o_i to agent a from a context of other objects $C \subseteq O$. For formal definitions of F_{a,o_i} and D^C_{a,o_i} see Steels (1996). In this model the context C of an association game is always of order two, containing a source o_s and a target o_t.

An interpretation T_{a,o_i} is defined as a transformation applied to the feature detectors of agent a when they are applied to an object o_i, denoted $F_{a,T(o_i)}$. An association $\alpha(o_t, T(o_s))$ is a relationship between a source object o_s and a target object o_t such that when the interpretation T is applied to the source there are new common features detected in o_t and o_s. In other words $\alpha(o_t, T(o_s))$ implies that the set of mapped features $K \neq \emptyset$, where $K = \{f \mid f \in (F_{a,o_t} \cap F_{a,T(o_s)}), f \notin (F_{a,o_t} \cap F_{a,o_s})\}$.

E_a is the set of all experiences possessed by agent a. An experience $e \in E_a$ is a memory of an association $\alpha(o_t, T(o_s))$, containing distinguishing and common features of the source and target in addition to the transformation T. A situation $X_{a,g}$ is constructed from a set of experiences that are applicable to the current game g, defined as being within a threshold d given a similarity metric $\Delta_{X_{a,g}}$.

The association game $g = (a, o_s, o_t, \alpha)$, where a is an agent, o_s and o_t are objects chosen from O and α is the association that has been developed for this pair of objects (if any), proceeds as follows:

1. The agent determines the distinctive feature sets D^C_{a,o_t} and $D^C_{a,T(o_s)}$, where T is the interpretative transformation used in α (if any). For a more in depth discussion of selection criteria for distinctive sets see Steels (1996).
2. The distinctive feature sets and the common feature set $(F_{a,ot} \cap F_{a,T(o_s)})$ are used as cues to construct a situation $X_{a,g}$ from E_a. The transformation T' is constructed from experiences in E_a as applied to the current objects using $X_{a,g}$.
3. The agent determines whether to keep the new interpretation based on whether the application of the transformation T' yields more valuable mappable features than the previous transformation T. In principle the evaluation metric is determined by the purpose of the analogy, but a proof of concept implementation could be based on simple enumeration of mapped features.
4. If the reinterpretation using T' is successful in creating a new mapping the agent creates a new association $\alpha'(o_t, T''(o_s))$ where T'' is the new transformation T' applied to the existing transformation T. A new experience e' is added to E_a based on α'.

The association game can end in failure if one of the following conditions occurs:

1. $D^C_{a,o} = \emptyset, o \in C$. There are not enough distinctions in S_a to identify either o_t or o_s and therefore $\forall o_c \in C, F_{a_r}, o \subseteq F_{a_r,o_c}$. When this occurs, the agent will construct a new feature detector or modify an existing one, see Steels (1996) for the specifics of this process.
2. $X_a = \emptyset$. The agent has no experience applicable to the current context. In this case the agent analytically calculates a simple transformation of the source that will produce commonality between its distinctive features and those of the target. This transformation produces a mapping of the discriminating features of two objects between which no mapping previously existed.
3. $order(K') < order(K)$. The transformation that was applied results in a worse mapping than the one it supersedes. In this case the new transformation T' is discarded and T remains as the association in α. Currently no experiences are generated on a failure case.

4 An example association

The association game model can in theory be applied to any domain for which an appropriate conceptual space can be constructed. To illustrate the game process we present a simple example in the domain of geometric proportional analogies. Proportional analogy problems of the form $A : B :: C :?$ are suitable for transformation-based association because the relationships A to B and $A : B$ to $C :?$ can be easily understood in terms of transformations. This model formalises the transformations by applying them at the level of conceptual spaces rather than to the problem directly. The model at present focuses on the association between two objects and its application to this domain is limited to solving for $A : B$ by constructing a relationship between them.

An example proportional analogy problem can be seen in Fig. 2. In this example we assume that the agent has already learnt concepts suitable to describe and discriminate between the two figures A (the source) and B (the target). The conceptual dimensions used in this example, such as the recognition of attributes of sub-shapes within the figures, are for illustrative purposes only. Perceptual processes in an implementation of this model may produce entirely different conceptual spaces.

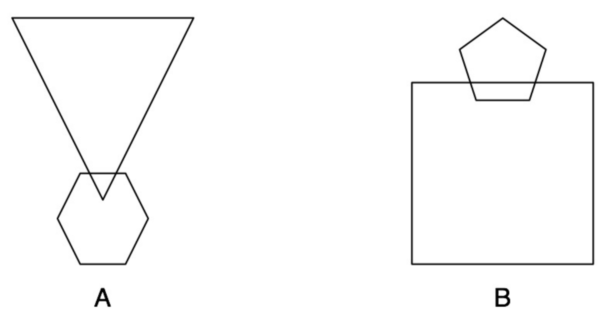

Fig. 2. A pair of figures as they are perceived by an agent playing an association game. Each figure has four features relevant to this example: the size and side counts for both the upper and lower shapes.

The game proceeds through the four steps outlined in Section 3.1. For ease of explanation we assume that the agent has encountered sufficient objects similar to these to possess the necessary conceptual lexicon. We also assume a very simple conceptual space for this example: a four dimensional space for each figure; the size and number of sides of both shape elements (starting with the upper shape), expressed $(size_1, sides_1, size_2, sides_2)$. Many other possible spaces exist for this and other proportional analogy problems, this space is not an example of the spaces that might be used by the agents but an explanatory aid.

The first step is to describe and discriminate between the two objects. The agent's feature detectors identify object A as having a top shape of size 4 with three sides, and a bottom shape of size 2 with six sides, denoted in this conceptual space as (4,3,2,6). Object B is identified as having a top shape of size 2 with five sides, and a bottom shape of size 4 with four sides, denoted (2,5,4,4). There are several possible distinctive feature sets between A and B, but we will assume that the agent discriminates based on the size of the upper shape - the agent states that the upper shape is "large" in figure A and "small" in figure B.

The second step is to try reinterpreting the source, A, by constructing an appropriate transformation that the agent believes will lead to a new relationship with B. Based on the feature sets and on the agent's prior experience with similar problems one possible transformation would be to treat the sizes of the two shapes within the figure as if they were reversed. The interpretation here would be "treat the size of the upper shape as if it were the size of the lower and visa versa". The problem can be seen in Figure 3 with the transformation applied.

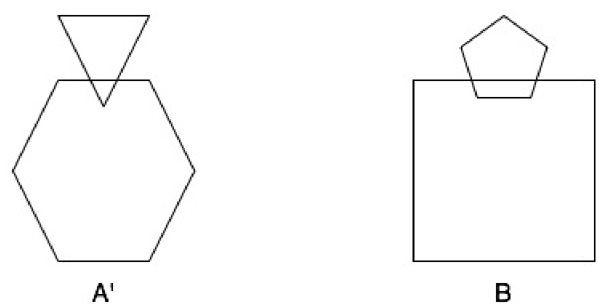

Fig. 3. The figures from Figure 2 after A has been reinterpreted to produce A'. A' is how the agent perceives A after its conceptual space has been transformed. An association with B can now be constructed.

The third step in the association game is to compare the reinterpreted source object with the target to determine In conceptual space the new co-ordinates of the source A would be $(2, 3, 4, 6)$. When this is compared to object B a new association could be constructed as the agent's feature detectors would now identify the agents by the same features on both the $size_1$ and $size_2$ dimensions. In other words when the agent "thinks of" figure A in this new way it is somehow "the same" as figure B - a relationship has been constructed where none was present before. Previously there was no association, so the agent views the re-interpretation as a success.

The last step in the association is to store the new association as an experience. This directly affects future association games between A and B as the stored transformation will be already applied and the new game will pick up from where this one left off. The experience also affects games between other objects as it can be used in the construction of other transformations. As a result of this experience and other association games the agent may learn to attempt to exchange dimensions between shapes within a figure.

A possibility in a future game applied to this example is that the agent attempts to create a mapping between the $sides_1$ and $sides_2$ dimensions. In other words the agent may apply its experience in swapping the sizes of the upper and lower shapes to the side counts. The number of sides is not an exact match as the shapes sizes were, with three and six sides for the shapes in figure A and five and four for the shapes in figure B. As a result of this the second mapping would depend on the specificity of the feature detectors that have developed for the number-of-sides dimensions. If the agent is specific enough to differentiate between 3 and 4 or 5 and 6 sided objects, then the second transformation would need to be more complex than a simple dimensional exchange for the reinterpreted source to be seen as "the same" in these dimensions. An informal English translation of the reinterpretation of shape A created by the application of both transformations would be "Treat the upper shape as if it were the lower and the lower shape as if it were the upper". The results of this second transformation can be seen in Figure 4.

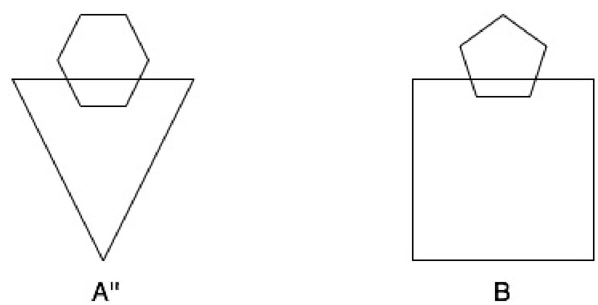

Fig. 4. The figures from Figure 2 after A has been reinterpreted twice to produce A". After a second association game is played starting from A', A" is produced by applying the same kind of transformation to different features.

5 Discussion

The model presented in this paper describes a method of constructing novel associations in which mappings are built over repeated iterations of the association game. The process of association is based on transformations applied to the conceptual space from which feature-based representations are generated. The model addresses the problem of representational rigidity suffered by symbolic association-based systems (Chalmers et al. 1994) by basing mappings on reinterpretations rather than symbolic identity. Through reinterpreting the source object the system is able to map between any two features, including those not present in the source before reinterpretation. This ability is mediated by an evaluation system that discards interpretations that enable the mapping of no additional useful features. Through iterative performance of the experience-driven association game the agent constructs an interpretation of the source object that enables mappings between those features that are appropriate within the interpretation. The association that is constructed between the source and target represents a relationship between them that is novel.

The agent playing association games learns about different ways of reinterpreting objects by learning different kinds of transformation and problems to which they can be applied. The agent learns transformations that are grounded in the specific set of objects it is exposed to and the sensory channels it possesses. In this paper we make the assumption that the learning system is capable of storing and applying transformations that have been useful to similar stuations where they may be similarly useful. We also assume that sufficient regularity exists in the environment that such appropriate transformations can be learnt and generalised to multiple associations in the domain. These assumptions are important points to be addressed in the development of systems that implement this model.

The model presented in Section 3 could be refined by better selecting what features are matched when no applicable transformations can be constructed from experience. In the current system the simplest possible discriminating features are chosen and an arbitrary transformation applied. The system currently relies on its capability to discard superficial mappings to ensure that appropriate matches can be found and others eventually discarded. This process could be improved by more tightly integrating the discrimination phase of the game with the goals of the association process. If the association was being made for a purpose the discriminating features could be selected based on their applicability to that purpose. The evaluation of whether a new interpretation is successful could also be integrated with the goal of the association, producing associations that may be useful to the agent in addition to being novel.

References

Chalmers, D., French, R. and Hofstadter, D.: High-level perception, representation and analogy: a critique of artificial intelligence methodology, *Journal of Experimental and Theoretical Artificial Intelligence*, **4**: 185–211.

Evans, T.: A heuristic program to solve geometric-analogy problems, in *Proceedings of the 25th Spring Joint Computer Conference*. Morgan Kaufman. (1964) pp. 327–338.

French, R.: The computational modelling of analogy-making. *Trends in Computer Science* **6**: 200–205.

Gärdenfors, P.: Induction, conceptual spaces and AI. *Philosophy of Science*, **57**: 78–95.

Gärdenfors, P.: *Conceptual Spaces: The geometry of thought*, MIT Press.

Kokinov, B.: Analogy is like cognition: dynamic, emergent and context sensitive *Advances in Analogy Research* NBU Press, Sofia.

Mullally, E.C. and O'Donoghue, D.P.: Spatial inference with geometric proportional analogies *Artificial Intelligence Review* **26**: 129–140.

Steels, L.: Perceptually grounded meaning creation, in *Proceedings of the Second International Conference on Mullti-agent Systems (ICMAS096)* Los Alamitos, CA. pp 338–344.