Analogy As A Sub-Process Of Categorisation

Sarah Darrington, Toby Lingstadt and Michael Ramscar[†]
(sarahd, tobyl, michael@aisb.ed.ac.uk)
Department of Artificial Intelligence, University of Edinburgh,
80 South Bridge, Edinburgh EH1 1HN, Scotland

Abstract

Analogy has traditionally been defined in terms of a contrast definition: analogies represent connections between things which are distinct from the 'normal' connections determined by our 'ordinary' concepts and categories. In this paper we present empirical evidence which, when added to other findings, supports our argument that in the light of current knowledge, the distinction between the two is based more on folk-psychology than on empirically based theory.

Research into analogy is however, distinct from research into categorisation when it comes to the richness of its process models. A number of detailed, plausible models of the analogical process exist (Forbus, Gentner and Law, 1995; Holyoak and Thagard, 1995): the same cannot be said of categorisation. On the other hand, these analogical process models make a number of explicit and implicit assumptions regarding an 'external' categorical process. Whilst treating these processes as separate has been useful in constraining the scope of cognitive investigations, we argue that it ultimately confuses the relationship between analogy and categorisation and is hampering the progress towards further understanding of both.

Introduction

The belief that analogy and categorisation are distinct and separable cognitive processes has widespread appeal. In our ordinary lives we unquestioningly accept an ontology that distinguishes between literality saying what something 'really' is - and analogies and metaphors, which, however informative they may be, are nevertheless not considered to be real statements about the world. We might talk of "the *foundations* of a theory"; we might wish to "*buttress* a theory with more facts"; "theories that we *construct* can also *collapse*", but from our everyday viewpoints, an igloo and a castle and a skyscraper *really* are similar in a way that similarities between buildings and theories are not; we can talk of someone's foxy cunning without *really* meaning to directly equate the cognition of fox and human cunning. French (1995) describes an experience of suggesting to an academic audience that an upturned orange-crate, when covered with a cloth and laid out with a picnic, might *really* be described as a table. This met with the swift response, "An orange crate is an orange crate is an orange crate." The attachment to pre-theoretical intuitions is a strong one, even amongst those who seek to explore them.

Research into categorisation, analogy and metaphor has accepted this realism, as indeed for a large part has cognitive science in general. Holyoak and Thagard (1995) describe a world in which "we think we see things as they really are", and analogy is used in order to recycle our existing knowledge of the real world to formulate new bits of 'real' knowledge.

Ramscar and Pain (1996) questioned the basis of these everyday distinctions in theory by querying the formulations that are offered in their defence. Analogy and metaphor are defined in contrast to 'categories' - Holyoak and Thagard (1995) describe analogy and metaphor as something that "connects two domains in a way that goes beyond our normal category structure" (pp 217) - a definition that does nothing definite by way of defining when we consider the account that can be provided of what constitutes a 'normal category structure' (c.f. Komatsu, 1992).

Empirical evidence was introduced to support this dispute regarding current definitions: Ramscar and Pain examined whether analogy could be distinguished from categorisation by contrasting the two processes. They presented participants with Gentner, Ratterman and Forbus' (1993) classic analogy materials (the 'Karla the Hawk' stories) and asked them to categorise them. Given that Gentner *et al* define the analogical mechanism in terms of structure mapping theory (see below for a full account), Ramscar and Pain hypothesised that assuming that the 'analogical' process was not distinct from the basic categorisation process, structure mapping would determine categorisation.

Gentner *et al* assumed that match items with only structural similarities (i.e. analogues) should be considered as belonging to different categories. Ramscar and Pain predicted that they would be categorised together. They found that 79.5% of the groupings formed by participants in their study had only shared systematic structure (traditionally defined as analogy) as a common feature amongst members of the categories formed. In contrast, only 5% of groupings produced had common object descriptions as the common similarity across categories (i.e. the attribute matches often thought to be determinate of categorisation). To the 79.5% of structural congruity groupings could be added a further 8% of classifications where structural *additions* to otherwise structurally congruent representations caused them to be classed singularly. Ramscar and Pain concluded that that mechanisms normally considered to be analogical specifically the preference for mapping systematically similar structures - could also in fact support categorisation

[†] This ordering is alphabetical: the authors all contributed fully to the experiment and the composition of this paper.

tasks, and that in their study no discernible difference could be found between analogical and categorical behaviour.

Ramscar and Pain's study can be added to other theoretical and empirical evidence which casts doubt on a two-process account of literal (categorical) versus non-literal (analogical or metaphorical) reasoning, such as Hoffman and Kemper's (1987) review of a number of reaction time studies which also convincingly demonstrates the paucity of the evidence for the widely held belief that literal (intra-categorical) meanings are processed faster than metaphorical (inter-categorical) meanings (as well as the considerable evidence for the opposite effect; see also Récanati, 1995).

Despite the weight of evidence contra the two-process account - and the concomitant lack of evidence for it suggestions which violate the two processes-account still tend to meet with some incredulity (c.f. French, 1995, above). The belief that an orange crate is an orange crate is an orange crate holds great sway. Indeed, such is the two-process account's entrenchment in ordinary, pre-theoretical understandings of the world that a more than usual quantity of counter-evidence seems to be required simply for the two-process account to be subjected to the usual standards of scientific and theoretical justification. And, whilst Ramscar and Pain's (1996) study might apply to classifying *stories* their study provides little evidence that this classification of stories can be generalised to other categorisation behaviour.

Structural systematicity and 'normal' categories

The following experiment was designed to offer more evidence for Ramscar and Pain's (1996) claim that a sub-process rather than two-process view should be taken of the relationship between analogy and categorisation (and thus add further support to Medin, Goldstone and Gentner's (1993) contention that structure-mapping may play an important part in categorisation). The Ramscar and Pain study utilised the 'Karla the Hawk' story sets (Gentner at al, 1993). These were designed as materials for studying analogy, and comprised a number of scenarios, typically episodic stories, with a controlled variance of relations between the narrative features of each story set member. In order to demonstrate the generalisability of Ramscar and Pain's findings, we designed a new set of materials, based on the 'Karla the Hawk' sets, but which rather than being stories, were sets of descriptions of novel objects (again, with a controlled variance of relations between the narrative features of each set member; see Figure 1).

The principle reasoning behind this is straightforward: typical 'natural' categories - the kind of categories found and used in human societies - tend to concern objects - and other, more regular 'things in the world' - rather than stories (Rosch, 1978). Classifying objects is more akin to ordinary categorisation than classifying stories (though a set of descriptions of 'rituals' was also included to reflect the fact ordinary categories reflect a good deal more than just physical regularities). Thus, the resulting object descriptions fell neatly and clearly into 'normal' categories (Figure 2).

Ramscar and Pain showed that shared structural systematicity (Gentner, 1983; see below) - typically defined as analogy - was the key determinant in participants' categorising in their study. In the light of this finding, we hypothesised that because structural commonalities in the object description sets ran across the 'normal' categories embodied in the set's object descriptions, these 'normal' category boundaries would be ignored as participants categorised objects according to shared structure.

Base

A Karla is a novel type of cooking pot, used by the Timuni in Alnata.

The structure of the Karla is designed in order to reduce the heat inside, and therefore prevents the food getting burned in the scorching cooking fires.

Water is poured into a layer of the Karla during cooking, which cools the food.

Literal Similarity

The Valkri is a special kind of frying pan, used by the Jalpeni in Frodon.

The Valkri is created in such a way as to be able to reduce heat, thereby preventing meat being getting burned when using the extreme temperatures of the cooking fires.

A liquid is poured into the layers of the frying pan when cooking, which cools the temperature of the meat.

Structural Similarity Only

The Vubu is a special wall built by the Jakar tribesmen in Frodon.

The Vubu is built in such a way as to be able to reduce the heat within it, thereby preventing the Jakar from sweating too much in the extreme temperatures of the midday sun.

A liquid is pumped through the Vubu, which cools the stone and therefore prevents the Jakar within the walls from getting too hot.

Mere Appearance

The people of Frodon use a special type of frying pan, known as the Valkri.

The Valkri is designed in order to allow it to be handled by children, as this can be difficult.

Its handle is designed with a special U-shape, which enables it to be held by people with small hands.

Structurally Similar to MA

The Jakar tribesmen of Frodon have built a special wall known as the Vubu.

The Vubu's stone gates can be opened by elderly people, despite their heavy weight.

Handles set in the wall incorporate springs, which allow weaker people to open the gates.

Objects Only

A new type of cooking pot, called the Karla, is used by the people of Alnata. Karlas can be purchased in a range of colours. Food cooked in a Karla tastes great.

Figure 1 Example of an object description set used in the study.

Set Member Description						
Object Sets	B	LS	SSO	MA	SMAO	OO
SET 1	plant	plant	tribe	plant	tribe	plant
SET 2	country ruler	country ruler	leading animal	country ruler	leading animal	country ruler
SET 3	board game	board game	field game	board game	field game	board game
SET 4	animal	animal	tribe	animal	tribe	animal
SET 5	cooking utensil	cooking utensil	wall	cooking utensil	wall	cooking utensil
SET 6	animal	animal	priest	animal	priest	animal
SET 7	chant	chant	game	chant	game	chant
SET 8	food	food	drink	food	drink	food

Figure 2: 'Normal' categories embodied in each object description set.

Gentner's structure mapping theory of similarity

Gentner's (1983; Falkenhainer, Forbus and Gentner, 1989) Structure Mapping Theory (SMT) is well known, and we include only a brief account of the most salient aspects here.

SMT proposes that the mapping and inference between two representations can be achieved by assigning correspondences between objects and attributes and then mapping predicates with identical names. In order to do this, Gentner assumes a predicate-like representation distinguishing between *objects, object-attributes* and *relations*. Object-attributes are those predicates that have one argument and describe object properties. Relations are divided into a hierarchy of orders, with those predicates with two or more arguments which are used to describe relations between objects forming the lowest order, and those predicates describing different levels of relationships between relations forming the higher orders.

The theory itself comprises two parts: *mapping rules* and the *systematicity principle.* Mapping rules state that (a) attributes of objects are not mapped and (b) relations between objects are preserved. The systematicity principle requires that higher order relations are mapped preferentially, followed by the relations that constitute the higher order arguments. Ramscar and Pain (1996) showed that participants' classification of stories could be predicted and explained according to SMT.

The Experiment

Participants

20 volunteers participated in this experiment. The participants were a mixture of Artificial Intelligence and Psychology students from the University of Edinburgh.

Materials

The basic materials used for this study were 8 sets of 'Karla the Pot' novel object descriptions (see Figure 1 for examples). These were descriptions of objects created to replicate the framework used by Gentner, Ratterman and Forbus (1993) in the creation of the "Karla the Hawk" stories.

As with the materials used by Gentner et al, the following taxonomy of similarity relationships between the object descriptions was defined:

"Literal similarity" matches include both common relational structure and common object attributes;

- *"Surface matches"* are based upon common object attributes, plus some first order relations;

"Structural similarity" matches are based upon a common system of internal relations;

- *"First order"* matches only have first order relations as a common feature;

- *"Object only"* matches only have object matches in common between the object descriptions.

Each of our sets consists of a base (B), a literally similar object description (LS), an object description that shared the same structure as the base, but no object attributes (SSO), a mere-appearance object description, with surface and first order commonalities with the base (MA), an object description which shared structure with the MA, and object attributes with the SSO (SMAO), and an object only match object description, with only surface attribute commonalities with the base (OO). This allowed for a number of potential groupings to be formed, according to the classification strategy participants adopted.

We predicted that despite the fact that we were using novel object descriptions which embodied existing categories rather than Gentner et al's relatively 'category-neutral' stories, participants would again use structural similarity as their categorical similarity determinant, putting analogues and bases into the same categories (i.e. B, LS and SSO together), rather than grouping match items at the object level (i.e. grouping B, LS, MA and OO together; which also equated to existing category membership; see figure 2).

Procedure

Each participant was presented with eight envelopes, each containing a different set of six novel object descriptions, and was asked to work through them one set at a time. Sets were presented in random order, as were the object descriptions within them.

Participants were instructed to read through the object descriptions within a set several times, until they felt familiar with their contents. They were then asked to put the objects together into groups, grouping the things that fitted most naturally together in their judgement. Groupings could

range from putting all descriptions into the same group to having them all in separate groups as well as all variations in between.

When the categorisation decisions had been decided on, the participant pasted them onto a large sheet of blank paper and then circled each grouping using a marker pen.

Once all eight sets had been divided into groups using this procedure, participants were re-presented with their groupings a set at a time, and were asked to give any group containing two or more members a simple descriptive name.

(Participants were also asked to write a few sentences explaining what had led them to classify each named group of descriptions together, though this data will not be analysed here).

The experiment took around an hour to complete.

Results

For every object description set, the groups formed by each participant's classifications were analysed (with the results displayed in Table 1). Groupings which emerged fell into a number of broad patterns . These classification types are listed in Table 2, below. Similarities across groupings (i.e. similarity shared by every member of a two or more member group across a categorised object description set) which could be identified according to Gentner et al's taxonomy were found in 80% of groupings (in Types 1, 3, 4, 5 and 6).

The most common grouping pattern used was of Type 1 (groups divided into: 1. B-LS-SSO: 2. SMAO-MA: 3. OO, using a network of systematic causal relations), which accounted for 70% of all classified object description sets.

The next largest grouping, comprising 5% of the total was of Type 6. These sets were grouped using a largely structural criterion which resulted in the same grouping pattern as for Type 1 with the exception of the SSO object description which was grouped on its own, even though MA and SMAO were still grouped together.

Groupings which occurred due to participants using common first order relations (those of Type 3) occurred in 1.9% of cases.

Object description sets were grouped according to Types 4 and 5 in 3.1% of cases. The only similarity across groupings of these types is that the object descriptions in each group had only objects in common.

Other groupings worth mentioning were Types 7 and 8, in which the structured object descriptions were grouped according to a determinable pattern, (structure for Type 7, 4.4%) and object attributes (Type 8, 1.9%). but the OO descriptions were assigned according to features in Type 7 (where we would expect a separate grouping), and grouped separately in Type 8 (grouped with descriptions containing similar object attributes expected).

Only 0.6% (one occurrence) of groupings were of Type 2, where the base was put into a category of its own, with shared structure being the only similarity across groupings.

In 11.2% of groupings it was impossible to determine an overall criterion for determining the pattern produced; each of these groupings had only a single occurrence.

		Object Description Sets								Type 1 Total
		1	2	3	4	5	6	7	8	
Subjects	A	1	1	1	5	1	7	1	7	5
	B	1	1	1	1	1	1	1	1	8
	C	1	-	1	1	1	3	1	1	6
	D	-	1	1	1	1	8	-	1	5
	E	1	1	1	1	4	-	1	1	6
	F	1	1	1	1	1	1		7	6
	G	1	-	1	1	1	-	1	-	5
	H	-	1	1	1	1	1	1	1	7
	I	1	1	1	1	6	-	1	1	6
	J	1	1	2	1	1	-		1	5
	K	1	1	1	1	1	1	1	1	8
	L	1	1	1	1	1	-	1	1	7
	M	1	6	1	6	1	1	1	8	5
	N	8	1	1	1	1	1	4	1	6
	O	1	1	1	1	1	1	6	6	6
	P	1	6	1	1	1	7	1	7	5
	Q	1	1	1	1	1	1	1	1	8
	R	1	7	1		1	-	6	7	3
	S	9	1	1	6	1		9	-	3
	T	1	4	3	1	3		4	9	2

Table 1: Results for grouping patterns. Each participant was given 8 sets of object descriptions (each row represents one participant; each column an object description set): the type of grouping is indicated by the type number in the object description set column (see also Table 2).

Discussion

This study further examined the hypothesis put forward by Ramscar and Pain (1996) that categorisation judgements in humans can be determined more by shared structural systematicity than by shared object attributes (surface features) between the objects/ things/ rituals to be classified. The results show considerable evidence to support this hypothesis: 70% of the groupings were made in this way (had participants grouped randomly, mathematical combinatorics yield 213 possible groupings of the materials). In a further 10% of groupings (Types 2, 6 and 7), shared structure was clearly the criterion determining the participants' overall groupings, although a single object description was classified unaccountably (usually singly).

An interesting effect from the Ramscar and Pain study that intentionally - was not replicated in this experiment, was the production of a large number of Type 2 groupings. In their experiment, Ramscar and Pain left an extra structure (inserted by Gentner et al as part of their analogy study) in a subset of the base stories presented to participants. These base stories with extra structure then tended to be grouped singularly (see Type 2 in Table 2, below). Since the 'Karla the Pot' materials did not contain any extra structures in the Base, we did not expect significant numbers of Type 2 stories to be produced, and in the event, only 0.6% of groupings (1 out of 180) resulted in a Type 2 pattern, where the base was classified singularly in an otherwise structurally determined grouping pattern.

Classification Criterion	Number	% of Total
Systematic network of relations in common - Type 1 1 B LS SSO 2 SMAO MA 3 OO	112	70%
Systematic network of relations in common - Type 2 (Ramscar & Pain, 1996) (Base classified separately) 1 LS SSO 2 SMAO MA 3 B 4 OO	1	0.6%
First order relations in common - Type 3 1 B LS SSO SMAO MA 2 OO	3	1.9%
Only object similarities in common Types 4 & 5 1 MA LS B OO 2 SSO SMAO	4	2.5%
1 B OO 2 LS MA 3 SSO SMAO	1	0.6%
Largely systematic network of relations in common Type 6 1 B LS 2 SSO 3 MA SMAO 4 00	8	5%
OO 'Problems' Structure based - Type 7 Object attribute based - Type 8 1 B LS SSO OO 2 MA SMAO	7	4.4%
1 B LS MA 2 SSO SMAO 3 OO	3	1.9%
Type 9 and others - No clear pattern 1 B LS 2 SSO 3 MA 4 SMAO 5 OO	3	1.9%
Others	18	11.2%

Table 2: Output patterns from the categorisation task, showing the groups formed and criteria established. The object descriptions are labelled according to Gentner's taxonomy of similarity (defined above): B = Base; LS = Literal Similarity; SSO = Structural Similarity Only; SMAO = Structural Similarity with MA and Object Similarity with SSO; MA = Mere Appearance; OO = Object Only match.

Groupings that appeared to be formed on the basis of shared surface attributes only amounted to 3.1% of the total (Types 4 & 5). To these could be added another 1.9% of groupings (Type 8) in which shared features determined the overall groupings, although the OO object description - distinctive due to its complete lack of any systematic structure - was classified separately.

Of those object descriptions classified according to shared object attributes, only 2.5% (Type 4 groupings) reflect the 'normal' categories shown in Figure 2.

Clearly, structure appears to be the key determinant of participants' classifications in this study. Typically, categorisation models have tended to concentrate on object descriptions, making use of very representationally-simple attribute-value lists (see Murphy and Medin, 1985; Komatsu, 1992), whereas, analogy research has examined relationships between highly structured representations (considering the influence of attributes, relations and higher-order relations in judgements of similarity. The evidence of this study would appear to support the claim that more notice needs to be taken of the kinds of representations used - and the effects representations produce in categorisation studies (Medin, Goldstone, and Gentner, 1993; Ramscar and Pain 1996).

The results of the present study also support the broader findings of Ramscar and Pain (1996), who conjectured that the processes underlying analogy and categorisation are not as distinct as is usually proposed. Both their results and ours show shared structural systematicity (Gentner, 1983) as the main process underlying categorisation judgements in the particular experimental conditions. Ordinarily, structural systematicity has been considered the domain of analogy, rather than categorisation.

In this study, the influence of shared structural systematicity has been remarkable. Participants have preferred groupings between pots and walls, and walls and pans, to pots and pans and walls alone. Whilst we feel that these findings have strong implications for categorisation research, we also feel that they should cause some food for thought as regards the way that analogy is typically viewed. As noted earlier, there is a widespread acceptance in analogy research of the two-process view of analogical / metaphorical and literal understandings, whereby 'literal' (within category) understandings are external to non-literal (analogical or metaphorical) understandings, and are therefore assumed to be computed by separate cognitive processes.

The evidence of this study can join other theoretical and empirical evidence against a two-process account of literal (categorical) versus non-literal (analogical or metaphorical) reasoning: we mentioned earlier Hoffman and Kemper's (1987) review of reaction time studies, which convincingly demonstrates the meagre evidence for the widely held belief that literal (intra-categorical) meanings are processed faster than metaphorical or analogical (inter-categorical) meanings.

In spite of this, we do not want to say that analogy *is* categorisation. It is difficult to envisage how such a central cognitive process such as categorisation could be reduced to a single process (c.f. Goldstone, 1994). Given the difficulty inherent in characterising analogical, metaphorical and categorical reasoning (Wittgenstein, 1953; Ramscar, 1997), we are as dubious of the usefulness of the kind of identity statements made by Glucksberg and Keysar, (1990), who argue that metaphorical statements should be understood as

class-inclusion statements, as we are of the contrast definitions with which we started this account; we consider it plausible - even likely that a number of reasoning processes play a part in categorisation. Rather, like Ramscar and Pain (1996) we argue that - in the light of the evidence currently available - analogy is best viewed as a *sub-process* of categorisation, and not as a separate process. Ultimately, we believe that the adherence to the two-process account confuses the relationship between analogy and categorisation and is hampering progress towards further understanding of both (Ramscar, Pain and Cooper, 1997). Until there is a better empirical and theoretical basis to do so, we argue that it may be useful (and more honest) to keep an open mind as to whether an orange crate is an orange crate *can* be a table?

Acknowledgements

The authors would like to thank Helen Pain and Richard Cooper for their assistance in this work, and Dedre Gentner for providing the "Karla the Hawk" story sets. We are also grateful for the comments of two anonymous reviewers.

References

Falkenhainer, B., Forbus, K.D. and Gentner, D. (1989) The structure mapping engine: an algorithm and examples. *Artificial Intelligence*, **41**, 1-63

Gentner, D. (1983) Structure-mapping: a theoretical framework for analogy. *Cognitive Science*, **7**: 155-170.

Forbus K, Gentner, D, and Law , K (1995) MAC/FAC: A model of similarity based retrieval *Cognitive Science* **19**:2:141-205

French, RM (1995) *The Subtlety of Sameness*, MIT Press, Cambridge, Ma.

Gentner, D Ratterman, M.J. and Forbus, K. (1993) The roles of similarity in transfer. *Cognitive Psychology* **25**: 524-575

Goldstone, R.L. (1994) The role of similarity in categorization: providing a groundwork. *Cognition*, **52**: 125-157

Glucksberg, S. and Keysar, B. (1990) Understanding metaphorical comparisons: Beyond Similarity. *Psychological Review*, **97**: 1: 3-18

Hoffman, R R and Kemper, S (1987) What could reaction-time studies be telling us about metaphor comprehension? *Metaphor and Symbolic Activity* **2(3)**, 149 - 186

Holyoak, K.J. and Thagard, P. (1995) *Mental Leaps*. MIT Press, Cambridge, Ma.

Komatsu, LK (1992) Recent views of conceptual structure. *Psychological Bulletin*, **112(3)**, 500-526

Medin, D., Goldstone, R. and Gentner, D. (1993) Respects for similarity. *Psychological Review*, **100:2**:254-278.

Murphy, G.L. and Medin, D.L. (1985) The role of theories in conceptual coherence. *Psychological review*, **92**, 289 - 316

Ramscar, MJA & Pain, HG (1996) Can a real distinction be made between cognitive theories of analogy and categorisation? *(In Proc.) 18th Annual Conference of the Cognitive Science Society*. Lawrence Earlbaum Assoc., New Jersey.

Ramscar, MJA, Pain, HG & Cooper, R (1997) Is there a place for semantic similarity in the analogical mapping process? *(In Proc.) 19th Annual Conference of the Cognitive Science Society*. Lawrence Earlbaum Assoc., New Jersey.

Ramscar, MJA (1997) Wittgenstein and the nature of psychological categories *Similarity and categorisation: proceedings of SimCat 97*, Department of Artificial Intelligence Conference Proceedings, University of Edinburgh, Scotland, 205-211.

Récanati, F (1995) The alleged priority of literal interpretation *Cognitive Science* **19**:207-273

Rosch, E. (1978) *Cognition and Categorisation.* Lawrence Earlbaum Assoc., New Jersey.

Wittgenstein, L *translated by* Anscombe, E (1953). *Philosophical Investigations* Blackwell, Oxford.