Proceedings of the 19th Irish Conference on Artificial Intelligence and Cognitive Science (2008) 3-12

Analogical Reasoning helps learning of Similar Unknown Concepts: the use of Analogies between Categories in Category Learning

Cesare Bianchi[1], Fintan Costello[1]

[1]UCD School of Computer Science and Informatics, Dublin, Ireland
{cesare.bianchi, fintan.costello}@ucd.ie

Abstract. Analogical reasoning is used in various areas of cognition. Current accounts of analogy in category learning focus on how analogies between members of a single category can help participants learn that category. In this study we ask whether analogies between members of two different categories can help participants simultaneously learn those two categories. A test with 4 categories, 2 groups of 2 similar categories, was presented to 3 groups of participants: one group sees pairs of objects from two similar categories, another pairs of objects from two dissimilar categories, and the third group sees only one object at the time. Time to finish the test, number of correct answers, and differences in learning the various categories are analyzed. Results suggest that analogy between categories is an important part of category learning.

Keywords: Analogical Reasoning, Category Learning, Paired Examples

1 Introduction

Analogy is a process allowing us to identify similar relational structures occurring in different contexts and transfer relational information from one context to another. This process is pervasive in human thought, occurring in famous scientific discoveries, in everyday adult thinking, and in even very young children (Goswami, 2001). Analogy plays a role in reasoning, perception, problem solving, language use, argumentation, learning, categorization, and many (if not all) other areas of thought (Gentner & Holyoak, 1997; Holyoak & Thagard, 1995). "Analogy pervades all our thinking, our everyday speech and our trivial conclusions as well as artistic ways of expression and the highest scientific achievements" (Polya, 1957; in Goswami, 2001).

Analogy is typically assumed to involve comparison between one domain whose relational structure is well understood (the Source), and another less well understood domain (the Target). On the basis of this comparison, relational structure is transferred from Source to Target, facilitating reasoning, problem solving and learning in the target domain (Markman & Gentner, 1993; Falkenhainer et al, 1989). Most theories of analogy are based on this imbalance between Source and Target: Gentner's Structure Mapping Theory, for example (Gentner, 1983) assumes that the relational structure of the Source is available and that analogy takes place on the basis of that structure (see the left side of Figure 1).

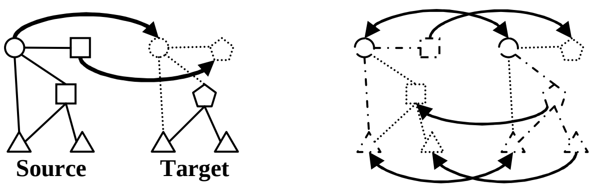

Fig1. *On the left*: analogical transfer of knowledge from a well known domain to a partially known one; *on the right*: reinforcement and completion of two partially understood domains.

The formation of analogies between a well-understood Source and a partially-understood Target is explained in accounts like Structure Mapping (Gentner, 1988; Gentner, 1983). The formation of analogies between two partly-understood domains (right side of Figure 1) is not easily explained in such accounts. We investigate analogies between such domains, asking whether analogy can take place even when neither the structure of Source domain nor Target domain have yet been learned.

1.1 Analogies *within* Categories and Analogies *between* Categories

A number of studies (Gentner & Medina, 1998; Gentner & Namy, 1999; Sloutsky & Fisher, 2004) have investigated the role of analogy in category learning, finding that analogies between members of a given category help people to find the similarities (and dissimilarities) between members of that category and aid the discovery of the features common to members of that category. We refer to this as the *within-category* use of analogy in category learning. Very little has been done, however, to investigate the role of *between-category* analogies during category learning. Our experiment aims to address this lack by asking whether analogies between members of two different categories can help participants as they simultaneously learn both categories. If we find that between-category analogy does have an influence on learning, this would demonstrate that analogical reasoning is used not only to transfer knowledge from an already well known domain to a less known one, but also to find similarities between domains that are both still being learned.

1.2 Analogical Reasoning in Learning of Similar Concepts in Completely Unknown Domains

This study investigates analogical reasoning between two ‘unknown’ domains, using a Category Learning task in which participants had to learn four categories. Two of these categories were designed to have complex analogies between them and two to have simpler analogies between them. These categories were formed from exemplars made up of various colored geometric shapes and defined by logical rules. This design ensures that no previous category knowledge is available to participants. In the study we ask whether partial learning of a given category aids participants in learning its analogous pair: that is, whether participants learn analogous categories together. If there is a pattern of linked learning between analogous categories even before those categories have been fully learned, this will demonstrate the use of analogy between two unknown domains (two unknown categories).

2 Experiment

The experiment tests two hypotheses at once: 1. that similarities between categories are used to learn both categories and 2. that these similarities are used to transfer knowledge from between those two analogous categories.

To test both hypotheses the experiment is designed with four different categories, similar two by two (Complex categories A and B have a similar relational structure, Simple categories C and D are based just on the presence of a distinctive element). Learning one Complex category may aid learning of the other, analogous, category, but should not affect learning of the other non-analogous categories. Participants in this experiment were asked to learn to correctly identify members of these 4 categories, to some criterion (80% correct). Participants learned the categories using a standard exemplar-presentation procedure, where participants are presented with examples and initially asked to guess which category those examples were members of. Participants were given feedback telling them the correct category of the example shown, and after repeated exposure to different examples, learn to identify category members.

The main novelty in this experiment is that some participants were presented with pairs of examples simultaneously during the category learning task, and had the opportunity to benefit from the comparison of the presented examples. There are three groups of participants in the experiment. The three groups differ in the materials they receive during the category learning task: one group (Paired) is most often presented with two examples from similar categories, the second group (Unpaired) sees more often examples taken from dissimilar categories, while the third control group (Single) is shown only one example at a time, and thus has no possibility of direct comparison. We expect that if analogies are detected and used during learning, the Paired group should find the test easier and thus finish in less time and using fewer examples, while the Unpaired group, which gains no advantage from the comparison of paired examples, should have a performance level equal to or worse than the control group, since the unmatched pairs may slow or impede learning.

The second hypothesis is tested by comparing the amount of time and number of examples needed to learn each category (i.e. to reach an accuracy of 75%). If learning one category helps learning the other analogous category (but not the other two, which have different structures), the time and examples elapsed between the learning of the two analogous categories should be less than the time and examples elapsed between the learning of those categories and the other two. This would demonstrate that there are two distinct learning processes that don't influence each other: one for the two Complex categories, the other for the other two. Moreover, if the difference between the learning of the two Complex categories is less than the difference between the learning of the two Simple categories, another hypothesis could be inferred: that a strong relational similarity facilitates transfer of knowledge.

2.1 Method

Participants. Participants were 30 volunteers from the School of Computer Science and Informatics of the University College Dublin: 20 Males and 10 Females, average age 26.2.

Stimuli. Stimuli were examples of the four categories, composed of colored shapes arranged in random order inside a container (a grey circle), shown over a white background. The number of presented elements in each example could vary from 1 to 12. Since there were 5 different shapes (circles, squares, triangles, crosses and stars) and 5 different colors (blue, red, yellow, green and pink), there were 25 distinct kinds of elements in total. Each example contained from 1 to 3 different kinds of elements. Each category was defined by a different rule, and the examples were randomly generated by the computer according to those rules.

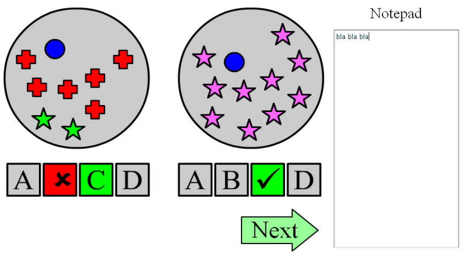

Fig2. The Experimental setup.

Rules for categories A and B (Complex categories) were based on the number of elements of the same shape and different colors (the shape and the 2 colors were chosen randomly at the beginning of the test and remained fixed). For category A, the number of elements with the given shape had to be the same for the two different colors, for category B, the number of elements with the given shape had to be twice the number for one color than it was for the other. For example, an exemplar of category A could be defined by 2 red circles and 2 green circles, while an exemplar of category B by 2 red circles and 4 green circles (or the opposite). The remaining elements were randomly added and had no role: they were just distractors.

Rules for categories C and D (Simple categories) were instead based on the simple presence of a distinctive element, of the same shape but different colors for the two categories. For example category C could be defined by the presence of a pink square, while category D by a blue square. The remaining elements were randomly added.

This design ensured that it was impossible for participants to guess the correct answer by exclusion: instead all the four categories had to be correctly identified. In particular, it wasn't enough for participants to identify the two macrocategories (Complex or Simple), since this would lead them to mistake A for B or C for D. Moreover, even though the C and D categories are defined just in terms of features, to correctly identify A and B participants have to discover their internal relational structures. Since the relational structures of A and B were analogous, we expected learning of one of these categories to aid learning of the other category, but to be independent from C and D, which had a completely different structure (and which could in turn help each other).

Design. Participants were assigned randomly to the Paired, Unpaired or Single groups. Participants in the Paired condition saw two examples on each presentation; 5 times out of 6 these two examples were both of the same kind (i.e. both Simple or both Complex, e.g. A and B, A and A, C and D, etc.), and only 1 of different kinds (e.g. A and C). Participants in the Unpaired condition also saw two examples at a time, but more often of different kinds (one Simple, one Complex). Finally, participants in the Single condition saw only one example at a time. The presentation order was random, but balanced in cycles of 24 examples, so every 12 steps (or 24 in the single group) participants were shown 6 examples for each category.

Procedure. At the start of the test, participants read detailed instructions about the task, then after filling a form with questions about their age and educational background, they started the task. According to the group, examples were presented paired, one on the left and the other on the right of the screen, or singly at the center of the screen. Beneath each example were four buttons with the four letters (A, B, C and D) that had to be clicked to choose the correct category. The answer(s) could be modified, and had to be confirmed by clicking an "Ok" button. Given answers were then recorded, and feedback was given through the buttons themselves: if the answer was correct, the clicked button became green with a "tick" symbol, otherwise it became red with a "X" symbol and the correct button became green. Subjects could spend as long as they wanted to learn from the feedback, and had a "Next" button to go on and be presented other examples (Figure 2). The test ended when subjects reached an accuracy of 80% in every category.

An electronic notepad of maximum 500 characters was available to take notes at the side of the screen, and its content was recorded at every step. Subjects were not permitted to use pen and paper. To better understand the learning process, when subjects reached an accuracy of 40% and 60% in every category, and at the end of the test, they were asked if they had found a rule for any of the categories and, if they had, what those rules were. At the end of the test a debriefing question asked what technique they used to solve the test. The test was implemented in Macromedia Flash and administered in a dedicated computer lab. All answers were timed with an accuracy of 1 millisecond and recorded on a server.

2.2 Results

The actual experiment was preceded by a preliminary test on 45 volunteers from outside the university, to check if it was solvable and to assess the expected trends in results. In this pre-test there were just the Paired and Unpaired groups, to which participants were random assigned.

Ease of Category Learning

Pre-Test. Since a preliminary analysis in the pre-test showed that ease of category learning was significantly different between the two groups (time in minutes to finish the experiment: Paired 34.2, Unpaired 52.9, F(1,15)=8.19, $p<.02$; number of examples to finish the experiment: Paired 156, Unpaired 456, F(1,15)=9.76, $p<.01$), we decided to proceed with the actual experiment.

Experiment. In the actual experiment that significant difference wasn't found (time: Paired 22.9, Unpaired 28.7, Single 25.7, F(2,27)=.41, $p>.60$; examples: Paired 138, Unpaired 193, Single 133, F(2,27)=.48, $p>.60$), although there is the same trend (Figure 3).

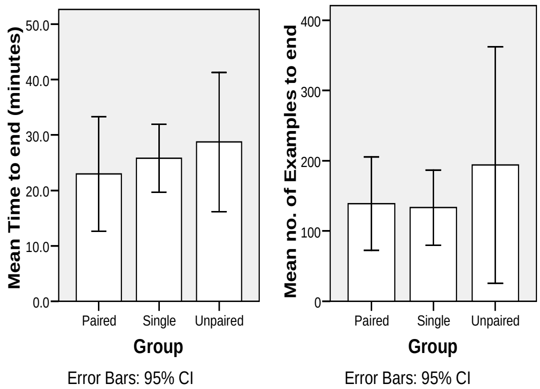

Fig3. Time and number of examples required to finish the test in Experiment

Comparing the results from the pre-test and the experiment, we discovered that in the experiment the participants took significantly less time and fewer examples to finish the test (time: F(1,45)=13.29, $p<.001$; examples: F(1,45)=4.80, $p<.05$). This difference suggests that the participants from the university were better able to solve these problems, probably because of intensive training in problem solving.

Facilitation between Similar Categories

Experiment. To ask whether analogy was used to transfer knowledge between categories during learning, we analyzed the differences between the time taken (and number of examples used) to learn different categories (to respond with an accuracy of 75% - see Figure 4). For each participant we calculated the time taken to successfully learn the first Simple category (to the 75% criterion) and the time taken to learn the second Simple category. The difference between these two times was calculated as DiffSS. The difference in the time taken to learn the two Complex categories (DiffCC) was calculated in the same way. Finally, the mean of all other differences (all Simple versus Complex) was calculated as DiffCS.

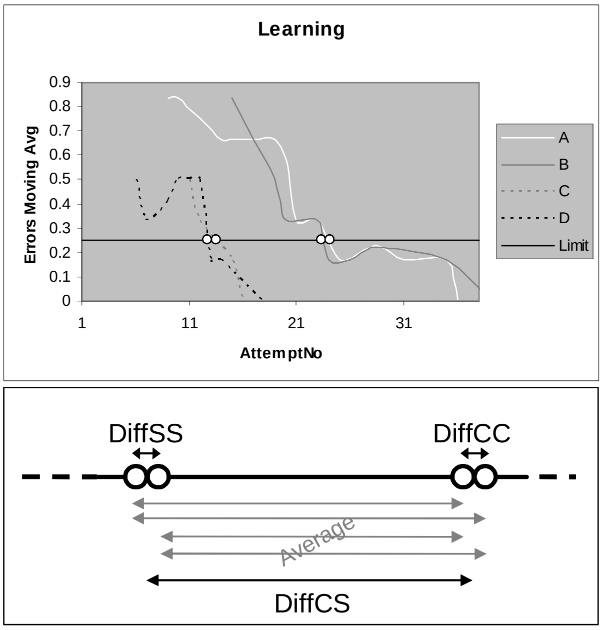

Fig4. Differences in Learning: A and B are Complex categories, C and D are Simple categories.

If learning happens independently for each category, all those differences should be the same. If learning of one category helps participants to learn the other category of the same kind (but not the two of the other kind), DiffCC and DiffSS should be less than DiffCS (see Figure 5). A Repeated Measures ANOVA showed a significant effect of the Difference factor, both for time and number of examples (time: $F(2,54)=22.53$, $p<.001$; examples: $F(2,54)=25.32$, $p<.001$) while the Group factor and the Difference * Group interaction were not significant ($p>>.20$).

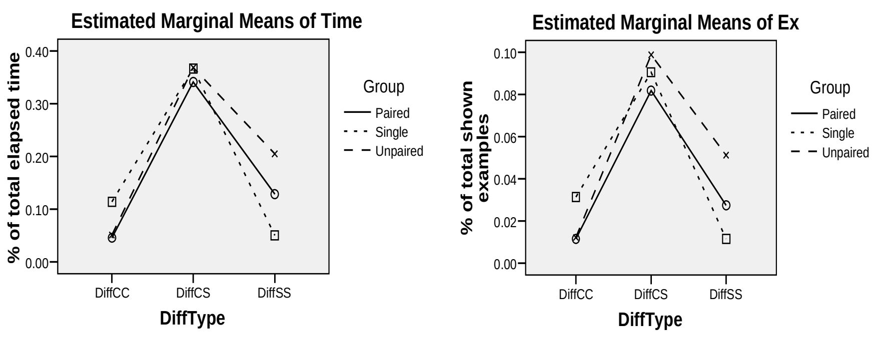

Fig5. Mean differences in Learning

Pairwise Comparisons showed a significant difference between DiffCC and DiffCS, and between DiffSS and DiffCS (for all: $p<.001$) while the differences between DiffCC and DiffSS weren't significant (*p*>*.10*).

Since both DiffSS and DiffCC are significantly less than DiffCS, the similarities between the two categories of each kind help learning of the other category of the same kind, thus showing that analogy between categories has an influence on category learning.

Formation of Macrocategories

Experiment. We investigated the degree of gradual refinement in participant's category learning by asking whether they first learned general macrocategories (i.e. Simple vs Complex) and only subsequently learned the final specific categories. We analyzed the percentage of correct answers at 3 different points in the test, going backward from the moment at which the first category was learned (i.e. accuracy of 75% on at least one category). Calling that moment *t*, we calculated the percentage of correct answers at 0.33*t*, 0.66*t* and *t*.

Table 1: Classification of Answers

		Given Answer			
		A	B	C	D
Correct Answer	A				
	B				
	C				
	D				

We calculated not only the percentages of correct answers (over the total given answers) for each category, but also the percentage of correct answers for each macrocategory (i.e. Simple or Complex - see grey cells in the Table 1) and the percentage of completely wrong answers (wrong specific category and wrong macrocategory: black cells in Table 1).

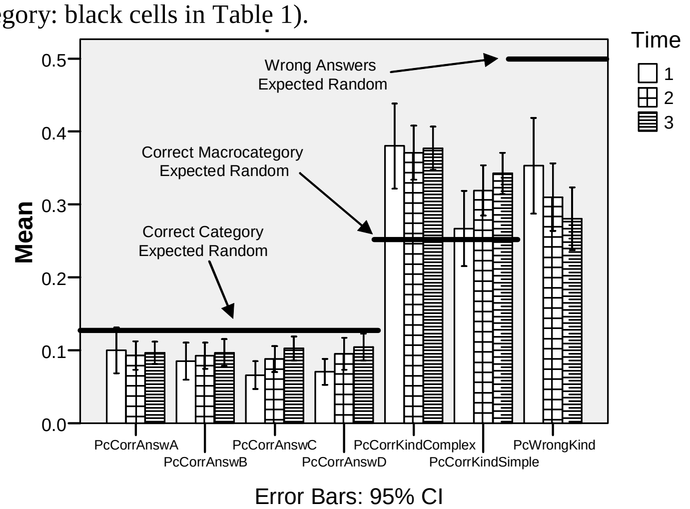

Fig6. Mean percentages of answers at 3 differentpoints in the test,comparing actual occurrence with expected random values.

If the answers were completely random, the frequency of exactly correct answers should be 12.5%, the frequency of answers inside the correct macrocategory should be 25%, while the frequency of completely wrong answers should be 50%. A t-test against those expected values showed (Figure 6) that well before the first category was learned, the number of completely wrong answers was less than the random value ($p < .001$), while the number of correct answers inside the macrocategories was more than the random level ($p < .001$), except, in the very beginning, for the Simple macrocategory. This situation remains the same through the first phases of learning. Results are the same across the two experiments, showing a reliable trend. We deduce that the formation of categories happens through consecutive refinement, and that the presence of analogical relations inside the categories, instead of making the task more complex, facilitates this refinement and so simplifies the learning task.

Other factors. Counting the number of entries (per minutes and per example) participants wrote into the notepad, we estimated its usefulness and found that it is indeed helpful in solving the test. We suspect that it shadowed the presentation factor, since it allowed participants to increase the span of comparison between examples.

2.3 Discussion

This study investigated the role of Analogy in a task in which similar categories are learned simultaneously. Our hypothesis, which we tested in different ways, was that analogy is indeed used in this situation, as in many other learning tasks. We found that learning of one category helped learning of another analogous category, and that in the case in which there are categories similar to each other, a common partial macrocategory is discovered well before the final categories are found. The discovery of these partial macrocategories shows that learning in our task was a process of continual refinement and that analogical reasoning is continuously at work, transferring partial knowledge.

Our theoretical proposal is that, since there is no prior category knowledge available at the start of the task we gave to participants, and since categories are all learned simultaneously during that task, Structure Mapping (Gentner, 1983) cannot successfully explain our results. Structure Mapping assumes an existing well known Source from which it maps knowledge to a less known Target, but in our case all concepts are learned simultaneously and so no well-known source is available.

To explain the use of analogy in our task, there is need of another, less deterministic and more automatic and basic process, which tries to find rules for the categories, even if initially those rules are only partial hypotheses to be further refined. If this basic process is also able to modify existing rules and see if the modified rules apply better, it can account for the emergence of Analogical Reasoning, without the need of more complex and advanced processes.

Further development: a Computational Model. Based on our results a Computational Model of Analogical Reasoning in Category Learning has been developed, to account for the progressive refinement of concepts and continuous transfer of partially understood knowledge. It is based on the idea that partial rules are formed, which could apply to more than one category, and which can be further refined. The processes involved are very simple, including the creation of rules from the presented example(s), the modification of other existing rule(s), and the testing of the existing and newly created rules to reinforce successful rules and reject poor rules. The model is indeed able to solve the task and exhibits patterns of behavior similar to those shown by participants. Although at present there are only preliminary results, the model shows interesting features and will be further developed and investigated, and a wider description and discussion of the model will be provided, with more definitive results.

Acknowledgments. This work was supported by the FP6 NEST Programme of the European Commission (ANALOGY: Humans – the Analogy-Making Species: STREP Contr. No 029088).

3 References

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

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

Gentner, D. (1988). Metaphor as Structure Mapping – the Relational Shift, *Child Development 59 (1)*: 47-59

Gentner, D. & Holyoak, K.J. (1997), Reasoning and learning by analogy – Introduction, *American Psychologist 52 (1)*: 32-34

Gentne,r D. & Medina, J. (1998). Similarity and the development of rules. *Cognition, 65 (2-3)*, 263-297

Gentner, D. & Namy, L.L. (1999). Comparison in the development of categories. *Cognitive Development, 14 (4)*, 487-513

Goswami, U. (2001). Analogical reasoning in children. *In Gentner, D., Holyoak, K. J., and Kokinov, B. N. (Eds.) The analogical mind: Perspectives from cognitive science*, pp. 437-470. Cambridge, MA: MIT Press.

Holyoak, K. J. & Thagard, P. (1995). Mental leaps: Analogy in creative thought. *Cambridge, MA: MIT Press*

Markman, A.B. & Gentner, D. (1993). Structural Aligment during Similarity Comparisons. *Cognitive Psychology, 25*, 431-467

Polya, G. (1957). How to solve it. Princeton, NJ: University Press.

Sloutsky, V.M. & Fisher, A.V. (2004). Induction and categorization in young children: A similarity-based model. *Journal of Experimental Psychology-General, 133 (2)*, 166-188