Modelling Typical Alphabetic Analogical Reasoning

Judith D.M. Grob (j.grob@ai.rug.nl)
Artificial Intelligence & Cognitive Engineering Institute, University of Groningen,
Grote Kruisstraat 2/1. 9712 TS Groningen, the Netherlands

Sharon Wood(sharonw@cogs.susx.ac.uk)
School of Cognitive and Computing Sciences, University of Sussex,
Falmer, Brighton, BN1 9QH, UK

Abstract

A study investigating the way in which people solve alphabetic analogical reasoning tasks (cf Copycat; Hofstadter & Mitchell, 1995), revealed that participants tend to use the same basic strategy, which was modelled in the cognitive architecture ACT-R. Performance evaluations indicate an average goodness of fit of 66.75% and a 100% goodness of fit on a subset of problems for which participants were significantly more likely to produce a single 'typical' response ($p<0.05$). The model is discussed in the context of various features of human analogical reasoning which were observed in the study, and in relation to Hoftstadter and Mitchell's (1995) discussion of 'elegant' solutions to problems.

Introduction

Analogical reasoning is about seeing similarities between different concepts, and making inferences based on these perceived similarities (Holyoak & Thagard, 1997). It plays an important role in problem solving, learning, social interaction (Hofstadter & Mitchell, 1995), scientific discovery and the arts (Boden, 1991). Observations of human development show that even before the age of two, children begin to develop the capacity for analogical thinking, without the aid of any formal training; evidently, the human mind is an analogical mind (Holyoak & Thagard, 1997).

Cognitive modelling has proven a valuable approach for investigating the cognitive underpinnings of analogy-making. Models such as Gentner's Structural Mapping Engine (SME)(1989), Holyoak and Thagard's (1997) Analogical Constraint Mapping Engine (ACME) and Hofstadter and Mitchell's Copycat (1995), successfully solve analogical problems in a human-like way, and have furthered our understanding of analogical thought.

The purpose of the work reported here was to investigate various aspects of analogical reasoning. A model for analogical reasoning was derived by ascertaining human responses in analogical reasoning tasks, and implemented in the cognitive architecture ACT-R (Anderson & Lebiere, 1998), which has proven a successful platform for modelling phenomena ranging from perception to problem solving, including analogical reasoning (Anderson & Lebiere, 1998).

Our work was inspired by the Copycat project (Hofstadter & Mitchell, 1995). Copycat operates in a highly idealised domain, consisting solely of alphabetic letter-strings, as a means of investigating reasoning about concepts and categories. According to its makers, by focusing on a highly idealised world, major properties of analogy-making can be isolated and explored more clearly, than when the focus is on real world problems (Mitchell, 1993). Focusing on this limited task, the Copycat program has proven to be capable of sophisticated analogical reasoning, and has even been credited with creativity (Boden, 1991).

Alphabetic Analogical Reasoning

An example of the alphabetic analogical reasoning tasks that were the focus of both the Copycat project and our work, is given in Figure 1. The problem can be read as: 'If it is known that "ABC" changes to "ABD", then what would "IJK" change to?'.

	left-hand-side		*right-hand-side*
source	**ABC**	→	**ABD**
target	**IJK**	→	**?**

Figure 1: Problem example.

Most people in our study, reported below, answer "IJL", by changing the last letter to its alphabetic successor. Other possible answers include "IJD","IJK" and "IJX". What answer is chosen, depends on the flexibility with which the source is perceived (compare "change the last letter to 'D'" with "change the last letter to any other random letter"). The range of responses Copycat can produce to this type of problem is compatible with answers given by human participants (Mitchell, 1993).

The empirical investigations carried out in conjunction with the Copycat project, presented human participants with five 'target' problems (see Table 1), and variations of these; corresponding to five clusters of alphabetic analogical reasoning tasks. In addition to the rather straightforward problems from cluster 1, the clusters address the effect of letter grouping (cluster 2), reversal of alphabetic flow (cluster 3) and encountering the alphabetic boundary (cluster 5) on answer-construction. The target problem from cluster 4, demonstrates tasks where the source consists of letter groups of different length; namely 1-2-3 ("M-RR-JJJ"). "ABC" can be regarded as 1-2-3 as well; being the first, second and third letter of the alphabet. This numerical congruence would be reflected by answering "M-RR-JJJJ", rather than "M-RR-KKK". The different types of tasks indicate that the same source ("ABC"→ "ABD") can be viewed in different ways, depending on the context provided by the target.

Table 1: Five target problems.

	Source	*Target*	*Type*
1	ABC→ ABD	IJK→ ?	Successor
2		IIJJKK→ ?	Grouped
3		KJI→ ?	Reversed
4		MRRJJJ→ ?	Numerical
5		XYZ→ ?	Boundary

Elegance & Typicality

In the Copycat project, rare and non-obvious responses to analogical reasoning tasks were referred to as more 'elegant' than their simpler counterparts. For example, answering items from cluster 3 in a numerical way (as described above) was seen as more ‘elegant’ than disregarding this numerical correspondence. Hofstadter and Mitchell (1995) seem to suggest that analogical questions that have been solved on a deeper conceptual level, are in a way 'better' than the alternative options, and therefore more interesting to research. In inspecting the data gathered by Mitchell (1993), however, it was striking that some answers given by participants occurred more frequently than others; and these did not correspond to the elegant answers.

Intrigued by the notion of 'elegance', the work discussed here repeated Mitchell's empirical investigations in order to obtain, firstly, a statistical analysis of the typicality of responses to the suite of analogical reasoning problems posed by Mitchell, and secondly, detailed protocols of the steps involved in the reasoning process. The question to be posed was whether an 'elegant' solution might otherwise be viewed as a non-typical solution, in statistical terms; further, might detailed protocol analyses provide evidence concerning how a single mechanism might be involved in producing these typical responses. As such, these typical answers might provide a natural ideal for the ACT-R model to emulate.

A Study of Analogical Reasoning Behaviour

Method

Participants: Forty people volunteered for the experiment. 21 were male and 19 female. Participants included both native (n=19) and non-native (n=21) speakers of English, with males and females being equally represented in both groups.

Procedure: Each person was given a computerised test. The test comprised 22 analogical reasoning problems previously used in Mitchell's (1993) study (see Table 2). The questions posed were distributed across the five clusters of analogical reasoning tasks, described above (see Table 1). Participants were asked to read the instructions, and then solve the problems in their own time. Following the test, participants were asked to type explanations to their solutions, where the answer given to a problem did not correspond to any of the answers previously found in Mitchell's study.

Results

The most typical answer, or solution, given to each question, is shown in Table 2[1]. A χ^2 test was conducted for each question to determine whether any answer was given significantly more frequently than any other.

Where a single response was significantly more frequent ($p < 0.05$), the response was considered to be a 'typical' answer. Using this criterion, 13 out of the 22 items were answered typically (see Table 2). Typical answers appeared predominantly for problems in clusters 1, 2 and 5, with proportionately fewer in clusters 3 and 4.

Table 2: Most frequent answers given by participants to 22 alphabetic analogical reasoning problems. T denotes whether the answer given could statistically be considered typical. Three answers predominated for problems in cluster 3(*).

Type	*N*	*Item*	*Most frequent answer(s)*	*T*
cluster 1	1	IJK	IJL	+
	2	XLG	XLH	+
	3	XCG	XCH	+
	4	ABCD	ABCE	+
	5	CDE	CDF	+
	6	CAB	CAC	-
	7	CMG	CMH	+
cluster 2	8	IIJJKK	IIJJLL	+
	9	HHWWQQ	HHWWRR	+
	10	LMFGOP	LMFGOQ	-
	11	LMNFGHOPQ	LMNFGHOPR	+
cluster 3	12	KJI	KJH*, LJI, KJJ	-*
	13	EDC	EDD*, FDC, EDB	-*
	14	CBA	CBB*, DBA, CBZ	-*
cluster 4	15	MRRJJJ	MRRKKK	-
	16	MRR	MRS	-
	17	MMRRRJJJJ	MMRRRKKKK	-
	18	RSSTTT	RSSUUU	+
	19	XPQDEF	XPQDEG	-
cluster 5	20	XYZ	XYA	+
	21	GLZ	GLA	+
	22	CMZ	CMA	+

[1] The full range of answers given by participants, is presented and discussed further in Grob (2002).

Interestingly, for each task in cluster 3, three alternative answers predominated. Collectively, these three answers (shown in Table 2) accounted for a significantly higher proportion (p<0.5) of the total set of responses than the remaining collection of answers.

Analysis of Findings

Almost all typical answers can be formed by applying the simple rule: 'Change the last item to its alphabetic successor'. with item denoting either a single letter, or a group of the same letters. If the last item is a 'Z' (as in cluster 5) then a continuous circular notion of the alphabet is applied, and successor 'A' is given. When the left hand target consists of a reversed letter string, such as in cluster 3, only one potential answer is formed by applying above specified rule; the other two equally frequent answers, take into account the reversed alphabetic order of the characters to which the rules 'Change the last item to its predecessor' and 'Change the first item to its successor' apply respectively.

The remaining tasks for which there was no 'typical' answer, also fitted the predominant rule 'Change the last item to its successor.' Only item 16 diverged from this pattern. Though the task "MRR" should be answered with "MSS" according to this main rule, only the last of the two letters is changed to its successor, yielding "MRS". Possibly, this is because the item has the exact same length as the left hand source, suggesting a mapping of the last "R" of "MRR" on the "C" of "ABC".

No significant differences in answer-patterns were found on basis of gender (p=0.811), native language (p=0.186) or the combination of the two (p=0.2).

The explanations of unusual answers that had been gathered in our study, described how many people answered by looking for an underlying pattern in the left-hand target. The most commonly applied strategy in less straightforward questions, was to count how far the letters in the target were removed from each other in the alphabet and use this information when forming an answer. Other strategies included basing the answer on acoustic qualities (rhyme), or referring to the adjuxtaposition of letters on a computer keyboard. Answers of this kind demonstrated the flexibility of the mechanisms involved in referring to a wider knowledge to solve problems. To gain further insight in how participants solve alphabetic analogical reasoning tasks, a second, more systematic study was conducted.

A Study of the Analogical Reasoning Task

To investigate the process of solving an alphabetic analogical reasoning task, a verbal protocol analysis was carried out.

Method

Two participants who had not taken part in the previous study were given a test involving the 5 target questions (see Table 1), plus an additional problem "XLG", which was included to test participants' reactions to an apparently random stimulus string. The participants were instructed to solve the problems in their own time and, whilst doing so, to recount verbally, the steps involved in their reasoning about how to solve the problems.

Results

The accounts of our participants, combined with the post hoc explanations obtained from our previous study, indicate that the steps involved in solving an alphabetic analogical reasoning task appear to be as follows:

1. Represent the left-hand source
2. Represent the right-hand source
3. Infer the source-rule
4. Represent the left-hand target
5. Map the left-hand target onto the left-hand source
6. Apply the source-rule
7. Give the answer

When mapping (step 5) is unsuccessful, target-representation (step 4) is applied recursively, until either a satisfying mapping has been established; or one runs out of patience. When mapping (step 5) an item from cluster 3 the source rule (which was formed in step 3) can also be adapted to reflect the 'oppositeness' between left-hand target and left-hand source.

An ACT-R Model of Analogical Reasoning

The derived protocol provided the specification for modelling analogical reasoning. The model was specified in the cognitive architecture ACT-R, which is both a specific theory on cognition, as well as a software environment for simulating human cognition[2].

Problem Representation

The model starts with representing the problem to be solved (See Figure 2). Representation takes place on two levels. On the syntactic level, the left or right hand side of the problem is represented in terms of its individual letters; and on the semantic level the relations that exist between the neighbouring letters are expressed. Consistent with the behaviour of the participants, it is assumed that when reading a letter string for the first time, people directly pick up on the repetition of letters (e.g. "A-A") as well as successorship (e.g. "A-B"). The model therefore encodes these relationships immediately they are encountered.

[2] See Anderson & Lebiere (1998) for a full account of ACT-R

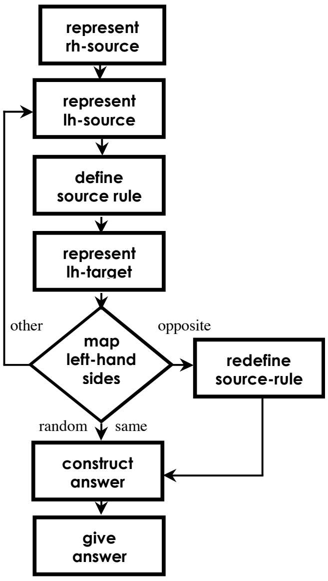

Figure 2: Flowchart of how the model solves a problem.

When a letter is repeated, this is represented by increasing a counter for that letter in the syntactic representation (see Table 3). In this way, groups of the same letters, are directly encoded as groups in the syntactic representation. No explicit representation of repeated characters is encoded at the semantic level; the 'sameness' between consecutive letters is implied by the counter following the letter itself. Thus, items in the semantic representation encode the relationship between neighbouring items in the syntactic representation, which can be either single letters or groups of the same letters.

When two consecutive letters are dissimilar, the alphabetic relationship between them is encoded in the semantic representation: a relationship can be labelled either 'next' (for successive letters) or 'other' for non-consecutive letters. When encoding of the entire string is complete, the program constructs a higher level semantic representation labelling the complete string as either a 'successorgroup', when a semantic representation consists of only 'next'-items (e.g. next-next), or an 'othergroup': when one or more of the items is 'other'. These labels, then, reflect the semantic relationships encoded at the lower level.

Table 3: Example of how the model represents string "AABL" (the | shows how far the model has processed the string)

Progress	**Syntactic**	**Semantic**
|AABL	-	-
A|ABL	A(1)	-
AA|BL -	A(2)	-
AAB|L	A(2)-B(1)	next-
AABL|	A(2)-B(1)-L(1)	next-other
		"othergroup"

Worked Example

To exemplify the process of solving an analogical reasoning task, Table 4 lists the steps the model takes when solving "ABC" □ "ABD" therefore "IIJJKK" □ "?". Representing both sides of the source is carried out in the manner described above. Following this, the source-rule is formed; this describes how the left-hand source has been transformed to form the right-hand source.

The source-rule is built in two stages from the syntactic representations produced during the previous stage of processing. In the case of "ABC"□ "ABD", first the A's are compared, then the B's and then the C with the D. This produces the source-rule 'same-same-next' which translates into 'change the *third* letter to its successor'.

Following this process would lead to the question "ABC"□ "ABD" therefore "IIJJKK"□ ?, being solved with "II*K*JKK". However, in accordance with our study findings, this rule is transformed into a higher level rule in which the string is encoded as two components: a 'body' and a 'last'. This produces the rule: 'body = same and last = next'.

After the source-rule has been formed, the left hand sides are mapped onto each other. Both "ABC" and "IIJJKK" are successor groups, and thus a perfect mapping is found. The answer can now be constructed.

The source-rule is applied to each letter in the syntactic representation of the left-hand target. The first letter, 'I' has a counter of 2 and is part of the body. As the rule for the body is 'same', the I is copied twice into the answer The 'J's are part of the body as well, so by applying the same rule, the answer now consists of "IIJJ". Finally the "K", with counter 2, is the 'last' element and here the source-rule states that the 'next'

Table 4: Overview of how the model solves the problem "ABC" -> "ABD", "IIJJKK" -> "?"showing the result of each consecutive step.

Step	**Result**
Represent left source	ls-syntactic = A(1)-B(1)-C(1) ls-semantic = "successorgroup" (next-next)
Represent right source	rs-syntactic = A(1)-B(1)-D(1) rs-semantic = "othergroup" (next-other)
Infer source-rule	source-rule = body(same)-last(next)
Represent left target	lt-syntactic = I(2)-J(2)-K(2) lt-semantic = "successorgroup" (next-next)
Map left sides	"successorgroup"="successorgroup" => same
Construct Answer	lt-syntactic / source-rule / application I(2)- / Body(same) / I-I- J(2)- / / J-J- K(2) / Last(next) / L-L
Give Answer	Output = I-I-J-J-L-L

letter needs to be taken. Therefore the successor of "K" is retrieved[3] and strung to the answer twice, giving the final result "I-I-J-J-L-L" as output.

Mapping Non-Standard Problems

In the example just described, mapping between left hand source and left hand target is immediately successful; however, for several problems, mapping does not proceed in this straightforward manner.

When encountering a reversed item like "KJI", or a 'random' numerical item like "MRRJJJ", the target is initially represented semantically as "othergroup". Mapping now becomes a problem, as the representations of left-hand target and source don't match ("othergroup"<>"successorgroup"). Therefore, the model tries to re-represent the left-hand target. Effectively, when the letters in the target are alphabetically very close to each other,[4] the program checks the target in reverse order from right to left. In the case of "KJI" this gives the semantic encoding: -next-next-, because the model is now reading it from right to left. This encoding is then transformed to "predecessorgroup" in the second stage of semantic processing. Attempting to map "predecessorgroup" onto "successorgroup", prompts the model to switch the source-rule, to reflect the fact that predecessor is the opposite of successor.

The model has two ways of inverting the source rule, both of which have a 50% chance of being selected. In one case, the source rule “take the successor of the last item” (body=same, last=next), is changed to “take the *predecessor* of the last item” (body=same, last= previous), yielding the solution “KJI”□ ”KJH”. In the other case the rule is changed to “take the successor of the *first* item” (first=next, tail=same), which gives “KJI”□ ”LJI”. If no predecessorgroup can be found either, as is the case for item "MRRJJJ", the left hand target is encoded as a "randomgroup" and the answer is formed by applying the original source-rule, which gives "MRRJJJ"□ "MRRKKK".

A final special case is encountered, when the model is given a problem from cluster 5, such as "XYZ". When the model wants to find the successor of 'Z', an error is retrieved. In accordance with participant behaviour, the model now recruits more general knowledge. Firstly it checks whether there is something 'special' about the letter 'Z' and retrieves that it is the 'last' letter in the alphabet. This enables the model to access the more general knowledge, that the successor of something which is last in a list, is the first item of that list. By retrieving the fact that the letter 'A' is the 'first' letter of the alphabet, the problem is now solved with "XYZ"□ "XYA". Of course, the same type of reasoning applies when looking for the predecessor of 'A'.

Evaluation

The model was run on the 22 problems shown in Table 2. The answers given by the model are presented in Table 5. To judge the performance of the model, goodness of fit was defined as the percentage of the answers that the model could give that were the same as the data to which it is compared. The model gives a 66,75% goodness of fit overall, when compared with the answers given to all 22 problems, by the 40 participants in our study.

The model was designed particularly to model the production of answers considered to be most 'typical' under our definition. When looking at the performance of the model on the tasks for which typical answers had been identified, the model shows a 100% goodness of fit. It should be noted, that though the items from cluster 3 weren’t typical accoring to our statistical definition, they each show an interesting pattern of three possible answers, of which the model is able to reproduce two.

Table 5: Answers given by the model on the 22 tasks

Type	*N*	*Item*	*Most Frequent*	*Model*	*Same*
cluster 1	1	IJK	IJL	IJL	+
	2	XLG	XLH	XLH	+
	3	XCG	XCH	XCH	+
	4	ABCD	ABCE	ABCE	+
	5	CDE	CDF	CDF	+
	6	CAB	CAC	CAC	+
	7	CMG	CMH	CMH	+
cluster 2	8	IIJJKK	IIJJLL	IIJJLL	+
	9	HHWWQQ	HHWWRR	HHWWRR	+
	10	LMFGOP	LMFGOQ	LMFGOQ	+
	11	MNFGHOPQ	LMNFGHOPR	LMNFGHOPR	+
cluster 3	12	KJI	KJH	KJH	+
			LJI	LJI	+
			KJJ	-	--
	13	EDC	EDB	EDB	+
			FDC	FDC	+
			EDD	-	-
	14	CBA	CBZ	CBZ	+
			DBA	DBA	+
			CBB	-	-
cluster 4	15	MRRJJJ	MRRKKK	MRRKKK	+
	16	MRR	MRS	MSS	-
	17	MMRRRJJJJ	MMRRRKKKK	MMRRRKKKK	+
	18	RSSTTT	RSSUUU	RSSUUU	+
	19	XPQDEF	XPQDEG	XPQDEG	+
cluster 5	20	XYZ	XYA	XYA	+
	21	GLZ	GLA	GLA	+
	22	CMZ	CMA	CMA	+

[3] Retrieval of successors and predecessors in the model is done via a re-implementation of the ALPHA model for letter retrieval (Klahr et al, 1983). For further details see Grob (2002)

[4] See Grob (2002) for precise details

Conclusion

The model presented here appears successful in answering the 22 problems in a typical, human-like way, but there are a number of ways in which its performance does not match that of the participants in our study. First, when a source and target do not map perfectly onto each other, the model merely checks to see whether the target has a reversed alphabetic letter ordering. Our participants, however, tried to re-represent the target in many ways, especially counting distances between letters, to arrive at a satisfactory answer. The full range and inventiveness of these observed strategies would be difficult to emulate; still, the model might at least be adapted to be able to apply numerical knowledge.

Secondly, when dealing with problems from cluster 3, the system can only choose to reverse the source rule, to reflect the reversed alphabetic flow in the target. However, some of the participants in our study chose to apply the rule: 'change the last letter(group) to its successor'. This, of course, is a rule the model is already able to deal with, and a simple adjustment to the model would enable it to capture this behaviour.

Finally, the present model has only been tested on 22 problems with source ABC ABD. Of course many more alphabetic analogical reasoning tasks can be imagined, and further research is needed to identify more precisely what the present model can and cannot handle, and why. Ultimately, the computational model should not only be able to describe typical answers, but predict them as well.

Discussion and Future Work

Whereas in Copycat (Hofstadter & Mitchell, 1995), a new modeling framework was built, the present work shows that typical alphabetic analogical reasoning can be successfully modelled in the more general cognitive architecture ACT-R. However, many features of ACT-R have of yet not been fully employed. Future research should make the model depend more heavily on ACT-R, adding to its explanatory strength.

Given the observed typicality of answers, the question arises, why some people give non-typical answers. Do they, for example, generate typical solutions, but ignore them in the quest for more interesting solutions, or are these non-typical solutions their first best guesses? In this context, we should note that only one individual in our study gave a typical answer to every problem. There appeared to be a continuum from individual participants who gave mostly typical answers and a few non-typical answers, to individual participants who gave mostly non-typical answers and a few that were typical.

The answers that were regarded as more elegant in the Copycat project, do not seem to be favoured by people such as the participants in our study. The work presented here suggests that typical answers are the product of reasoning on a lower representational level than in the production of elegant answers. Consideration of these more elegant answers, therefore appears unhelpful in understanding how humans generally deal with these types of questions.

However, Hofstadter and Mitchell (1995) link analogy making to creativity, by suggesting that answers based on a higher level of representation, are more 'creative'. It would be interesting to see whether a positive relationship between giving more elegant answers on the alphabetic analogical reasoning task and some general measure of creativity, does indeed exist. By extending the computational model to produce such 'elegant' solutions, we perhaps become better placed to understand what constitutes 'creativity' of this kind. This is an issue for further work.

Acknowledgements

The authors would like to thank Niels Taatgen for his useful advice on ACT-R throughout the project, and helpful comments on the writing of this paper. Thanks also to David Klahr, for providing background information on ALPHA, the computational model for alphabetic retrieval, and to Melanie Mitchell for her helpful comments on her empirical studies. This work took place at the University of Sussex with funding from NUFFIC and the VSB.

References

Anderson, J. and Lebiere, C. (1998). *The atomic components of thought.* Erlbaum, Mahwah, Nj.

Boden, M.A. (1991). *The creative mind: myths and mechanisms.* Abacus, London.

Gentner, D. (1989). The mechanisms of analogical learning. In S. Vosniadou and A. Ortony (Eds.), *Similarity and analogical reasoning*, Cambridge university press, Cambridge.

Grob, J.D.M. (2002). *Modelling alphabetic analogical reasoning in ACT-R.* Unpublished Masters Thesis. University of Sussex, Falmer. http://www.ai.rug.nl/~judith/dissertation.ps

Hofstadter, D.R. and Mitchell, M. (1995) Perspectives on Copycat. In D.R. Hofstadter and FARG(Eds.) *Fluid Concepts and creative analogies. Computer models of the fundamental mechanisms of thought.* Basic Books, New York.

Holyoak, K. and Thagard, P. (1997). The analogical mind. *American Psychologist*, 52, 35-44.

Klahr, D., Chase, W.G., and Lovelace, E.A. (1983). Structure and process in alphabetic retrieval. *Journal of experimental psychology: learning, memory and cognition,* 9, 462-477.

Mitchell, M. (1993). *Analogy-making as perception.* MIT press, Cambridge, Massachusetts.