

THE EMERGENCE OF ANALOGY
ANALOGICAL REASONING AS A CONSTRAINT SATISFACTION
PROCESS

Jan Van Dormael

Introduction

The recent interest in computational investigations of analogy has seen the rise of a distinction between so-called syntactic and pragmatic approaches towards the study of analogy (Keane, 1988). This distinction is essentially a distinction concerning the type of constraints placed upon the mapping process. In this paper we will point out that the thought process under study is justified reasoning by analogy and argue that the mapping constraints are introduced to select a 'valid' analogy. Understanding analogical thinking differently, we will argue that an analogy emerges progressively in thinking about the target situation as if it is the base situation. Furthermore, we will point out some of the differences between the "cognitive-computational" and a "cognitive-logical" perspective on analogical thinking.

1. Analogical thinking as reasoning by analogy

Analogical thinking is a form of thought frequently used in solving problems (e.g. Gick & Holyoak, 1980), learning (e.g. Winston, 1982), argumentation (e.g. Perelman & Olbrechts-Tyteca, L., 1958), understanding metaphor (e.g. Black, 1979), scientific theory formation (e.g. Hesse, 1966), explanation (e.g. Gentner, 1983) and case-based reasoning (e.g. Ashley, 1988). The main problem in understanding analogical thinking in these contexts is that this form of thought is not a reasoning process, i.e., if we understand reasoning as the act of verifying and proving propositions.

One way to solve this problem is to analyze analogical thinking as *reasoning by analogy*. A reasoning process in which an

analogy is not the result but its starting point. This reasoning process is generally described as a process of inferring that a property Q holds of a particular situation or object T (the target) from the fact that T shares a property or set of properties P with another situation or object S (the source or base) which has property Q. The set of common properties P is the similarity between S and T.

Given this interpretation of analogical thinking, the central problem is the *justification of an inference by analogy*. From the fact that T shares P with S and S has the property or set of properties Q it does not follow with logical necessity that T also has the property Q. As such the logical study of reasoning by analogy is predominantly concerned with the so-called justification problem: "*given two analogues x and y which resemble each other in a number of characters $B_1... B_m$, we want to know whether it is justified to infer that another character D of x belongs to y.*" (Hesse, 1966, 73). This means that we have to find a criterion which sufficiently establishes the truth of the conclusion (Davies, 1988, 229).

Different solutions, depending on whether reasoning by analogy is seen as akin to inductive (e.g. Mill, 1906) deductive (e.g. Weitzenfeld, 1984) or paradigmatic reasoning¹ (Davies, 1988) have been proposed. There is of course nothing wrong in trying to solve the justification problem. However it becomes a problem when analogical thinking is modelled on the process of justifying the use of analogy. In doing so we consider analogical thinking to be a reasoning process and neglect to explain the crucial question: how does an analogy gets formulated?

In Hesse's *Models and Analogies in Science*, for example, the question how analogies work in scientific theory formation is answered by (1) formulating the conditions justifying the use of analogies, (2) showing that reasoning by analogy provides a method of hypothesis selection which is justifiable on at least some of the recognized criteria for such a selection (Hesse, 1966, 55-6, 77). The conditions justifying the use of analogies are fulfilled by what Hesse calls the material analogy because these types of analogies lead to the formulation of strong falsifiable hypotheses (Hesse, 1966, p.128).

However, if we believe that an analysis of the conditions justifying the use of an analogy answers the question how analogies work in theory formation then we also assume that in some way or another the analogy already exists before it is used in theory formation. For Hesse (1966, 32) the analogy already exists in the form of a pre-scientific analogy. "*My whole point is that it is necessary to have these correspondences [the pre-*

scientific analogy] *before the theory, otherwise the theory is not predictive or falsifiable in the strong sense.*" (Hesse, 1966, 43) But, one of the characteristics of the formation of a new theory is precisely that an analogy is formulated during the process of transferring knowledge from one domain to another (see Schon, 1963; Boyd, 1979). How then can we use a not yet discovered analogy? The 'logic' of analogy becomes a 'logic' of the justification of analogy and the use of an analogy is reduced to the use of an analogy in the justification of a new hypothesis or solution.

It is our view that cognitive scientists, interested in explaining the thought process underlying the use of analogies, are confronted with the same problem for the same reasons. In order to construct correspondences between a source and a target domain it is necessary to assume an analogical representation of both domains and the introduced constraints are introduced for the sake of selecting necessary true correspondences.

2. Analogical thinking as constraint satisfaction

In cognitive science and A.I. it is generally agreed that analogical thinking is essentially a mapping process: the construction of orderly correspondences between the elements of a source analogue (S) and those of a target (T) (Carbonell, 1983; Kedar-Cabelli, 1988, Gick & Holyoak, 1980; Gentner, 1983; Holyoak & Thagard, 1989; Thagard, 1988). There is however less agreement on the relationship between mapping and the other subprocesses of analogical reasoning: the retrieval of a plausible useful source analogue, the elaboration of the source analogue and evaluation (e.g. Thagard, 1988; Kedar-Cabelli, 1988). Here, we will restrict ourselves to discuss the way the mapping process is described.

The main problem in computational descriptions of the mapping process (Keane, 1988; Holyoak & Thagard, 1989) concerns the reduction of possible matches which need to be considered in order to find an appropriate set of correspondences between the elements of source and target. If there are N elements in the source analogue and M elements in the target analogue the number of ways these situations can be paired is $N!/(N-M)!$, given that N is equal to or greater than M . To reduce the combinatorial explosion *constraints* are placed upon the mapping process. These constraints function as principles deciding the appropriateness of the mapping and implicitly define a set of candidate inferences that could be made about the target (Keane, 1988, 243; Holyoak & Thagard, 1989, 297).

Many cognitive scientists, particularly Gentner (1983), stressed the importance of structural or isomorphic correspondences as a criterion for satisfying analogical mapping. In Gentner's structure mapping theory, what gets mapped across situations depends solely on the syntactic properties of the represented base and target situations. In this view, the mapping is constructed according to three syntactic rules: (1) discard attributes of objects, (2) try to preserve relations between objects, and (3) give preference to relations that are part of a coherent system. The most important rule is the last rule, the principle of systematicity, people prefer to map systems of predicates that contain higher-order relations (such as "cause" and "implies") rather than to map isolated predicates.

In the computational implementation of this theory (Falkenheimer et al., 1986; Skorstadt et al., 1987) the algorithm used computes mappings between the various relations in both domains. From this a number of alternative structure-mappings are formed and any candidate inferences which follow from these structure-mappings are noted. These candidate mappings (called 'G-maps') are then evaluated to find the best possible mapping using three criteria: (1) evidence for the mappings in the G-map, (2) the number of mappings which can be made in the target domain using the G-map (the greater the number of mappings, the greater the acceptability of this particular G-map), (3) the number and relative size of connected components of the G-map (the bigger the better). Finally, the candidate with the highest systematicity is mapped into the target domain.

In Gentner's theory, the process of constructing an analogy (candidate G-maps) is interpreted as the construction of a structural analogy. A process which is explained on the basis of a mapping process satisfying structural constraints. But these structural analogies are precisely the conditions justifying the inference of a conclusion in reasoning by analogy. Weitzenfeld (1984), for example, argues that reasoning by analogy is valid deductive reasoning if the premisses are statements about an isomorphism between two situations. Seen from this perspective, the process of constructing an analogy is a process of justifying the use of analogy.

However, in contrast to a pure syntax-driven approach most accounts of mapping between analogues introduce besides syntactic constraints also semantic and pragmatic constraints. Holyoak and Thagard (1989), for example, developed a mapping theory that integrates syntactic, semantic and pragmatic constraints. The syntactic constraints are similar to those used by Gentner's structure mapping theory. The semantic constraint,

semantic similarity, supports correspondences between elements to the degree that they have similar meanings and the pragmatic constraint, role identity, restricts possible correspondences between elements to those that play identical roles in the two analogues (e.g. goal elements can only map to goal elements).

The introduction of these (additional) semantic and pragmatic constraints does not alter the main problem generated by considering the process of constructing an analogy as a process of justifying an analogy. In order to exploit these constraints we have to assume that the base and target situation are highly structured in advance. To exploit, for example, the pragmatic constraint we have to assume a representation of the target situation in terms of problems, plans and goals. In other words we must have some idea about the problem and how to solve it before correspondences between the base and target situation are constructed. But often we do not understand or have no clear idea about the problem. Precisely at this moment previous problems and their solutions are worth considering. It is only during the restructuring of the target situation in terms of a retrieved previous solved problem that we 'discover' or fully understand the new problem we are confronted with. The same goes for the other constraints: similarities and structural correspondences are 'discovered' in the restructuring of the target situation given a base situation.

A result of considering the construction of an analogy as a process of justifying the use of analogy seems to be that we can only construct an analogy between two situations if the representations of base and target situations are analogous. Given an analogical representation a mapping process by constraint satisfaction does not construct an analogy but *selects* an appropriate or 'good' analogy. It is our opinion that this approach towards the study of analogical thinking results from identifying analogical thinking with reasoning by analogy. According to us, another interpretation of analogical thinking is possible. An interpretation which does not model analogical thinking on reasoning by analogy.

3. *Analogical thinking as 'as if' thinking.*

Analogical thinking is a thought process in which we use our imagination to solve problems, understand metaphors, build new theories and learn more about the world. More specifically, in thinking analogical we use our imagination to think about something *as if* it is something else² and formulate judgments we

mistakenly take to be about the actual world. In *'Analogical reasoning: a logical inquiry in archaic thought'* (in press) I described the logic associated with as if thinking as a transdeductive, non-relational and irreversible logic. A logic in which judgments are juxtaposed not obeying the law of contradiction. In this archaic or egocentric logic I situated analogical thinking and described it as a kind of counterfactual thinking about a target domain given a source domain. The use of Stalnakers minimal change theory enabled me to formulate the truth functions of counterfactual statements and describe analogical thinking as a truth preserving process.

Let us suppose the following situation³. A doctor who has a patient with a tumor in his stomach wants to destroy the tumor using a kind of rays. If the rays reach the tumor all at once at a sufficiently high intensity, the tumor will be destroyed. But there is a problem. At a high intensity the healthy tissue that the rays pass through on the way to the tumor will also be destroyed. The doctor solves this problem by thinking about a story in which a general wanted to capture a fortress and for several reasons divided his army in small groups who from different directions simultaneously converged on the fortress. That is, the doctor, analogously, divides the rays into several low-intensity rays which from different directions simultaneously converge on the tumor. What is the thought process underlying the doctor's problem-solving behavior?

I argued that analogical thinking finds its expression in counterfactual statements. Counterfactual statements in which the antecedent is a counterfactual identity - an identity which is false given our actual set of beliefs - between an element of the base and target situation. The consequent is a statement obtained by substituting one for the other elements in the target. As such these counterfactuals are a formulation (in a conditional form) of an optimistic strengthening of Leibniz's Law of identity of indiscernibles⁴. Given our example, let us say that the target domain is represented by a set of propositions including 'a doctor wants to destroy a tumor with rays' and that the representation of the base domain contains the proposition 'a general wanted to capture the fortress with his army'. In counterfactual thinking about the source given the base the doctor formulates, for example, the following counterfactual: 'if I (the doctor) were the general then I would capture the fortress with my army'. In other words, in *imagining* he is the general the doctor *thinks* he wants to capture a fortress with his army.

Basically, analogical thinking is a progressive formulation of true counterfactual statements. Given the truth of a counter-

factual another counterfactual is formulated on the basis of the previous. (e.g. 'if the fortress were the tumor then I want to capture the tumor with my army') This process ends at the moment the problem is solved, the metaphor understood, etc... (e.g. 'if small groups were low-intensity rays then I capture the tumor by dividing the rays into low-intensity rays which from different directions simultaneously converge on the tumor'). Taking into account the fact that the conditionals are progressively formulated on the basis of a source domain given the target, the process may be schematically summarized as follows:

$$\begin{array}{l}
 (1) \quad S \quad \& \quad C_1 \quad \rightarrow \quad S_1 \\
 (2) \quad S_1 \quad \& \quad C_2 \quad \rightarrow \quad S_2 \\
 \quad \quad \cdot \quad \quad \quad \cdot \\
 \quad \quad \cdot \quad \quad \quad \cdot \\
 \quad \quad \cdot \quad \quad \quad \cdot \\
 (n) \quad S_{n-1} \quad \& \quad C_n \quad \rightarrow \quad S_n
 \end{array}$$

(S stands for the source domain, (C_1, \dots, C_n) are counteridenticals and (S_1, \dots, S_n) are the sentences expressing the progressive interpretation of the base situation on the basis of the formulated counteridenticals.

Given Stalnakers truth functional interpretation of these statements⁵ we can describe the analogical thought process as follows: after finding a counteridentical (C_1) on the basis of the source (S) and the problem solving situation a counterfactual is formulated and evaluated. That is, $S \& C_1 \rightarrow S_1$ is true in the actual world (w) if S_1 is true in a selected ($S \& C_1$) world. We add the true counterfactual to our stock of beliefs (our beliefs about the target and base situation) formulate and evaluate another counterfactual (C_2) on the basis of this changed set of beliefs. This process ends at the moment our set of beliefs is changed in such a way that it enables us to solve a problem, understand a metaphor, etc...

Let us assume that by adding $S_{n-1} \& C_n \rightarrow S_n$ to our stock of beliefs we are able, for example, to solve a problem. In Stalnakers conditional logic C_2 it is valid to infer from sentences (1) - (n):

$$(3) \quad S \& (C_1 \& C_2 \& \dots \& C_n) \rightarrow S_n$$

A counterfactual in which the antecedent clause may be interpreted as saying that on the basis of the source domain (S) a set of mappings (C_1, \dots, C_n) is constructed. Given the truth of the counterfactual and assuming the truth of its antecedent in the

actual world the problem is solved by inferring S_n .

$$\begin{array}{l}
 (4) \ S \ \& \ (C_1 \ \& \ C_2 \ \& \dots \ C_n) \ \rightarrow \ S_n \\
 \quad \ S \ \& \ (C_1 \ \& \ C_2 \ \& \dots \ C_n) \\
 \hline
 \quad \quad \quad S_n
 \end{array}$$

However, this inference scheme results from *an after the fact view* on analogical thinking. Only after changing our set of beliefs in such a way that it induces a way to solve the problem we can formulate, as a way of justifying our thinking, an inference based on the above mentioned inference scheme. But, in order to make this inference valid, we must assume the actual world to be one of those worlds in which the counterfactual is true. To solve the problem, the doctor assumes the actual world to be a world in which the antecedent is true and not true. Which is the same as saying that for the doctor the possible or fictional is as real as the actual.

If we compare our analysis with the study of analogical thinking as reasoning by analogy then the logical invalid inference scheme comes close to such an analysis. Apart from the fact that it is an *after the fact view* on analogical thinking, we understand the logical justification problem differently. The success of analogical thinking does not depend on the number of properties in common nor upon the existence of a structural similarity but on the lack of differentiating between planes of reality. The justification problem is not a problem of finding, for example, the necessary inductive support nor is it a problem of finding implicit premises which makes the inference valid. From our perspective the traditional logical problem is essentially a problem about relevant contradictions. How to permit contradictions without allowing everything to follow from these?

But in answering this question we do not answer a question about the way we think analogical. We do not assume that in analogical thinking a conclusion (S_n) is inferred from an analogy (for, example, C_1, \dots, C_n). Analogical thinking is a process in which judgments are juxtaposed. There exist no conscious implications or demonstrative links between the judgements. As such analogical thinking is not a process of *proving* a solution. It is even difficult to prove the solution by looking back because this thought process is an irreversible process of *finding* a solution. Irreversible in the sense that it is difficult to perform the return journey without deviating from the path. In starting from A and finding B one does not necessarily find A again or if one does, one will not be able to prove that the A found is the same

A. The evaluation of the different counterfactuals brings one in different possible worlds. The antecedent world which made the first conditional true is not necessarily the same as the antecedent world which makes the second counterfactual true, and so on. It is our opinion that precisely this irreversibility makes it difficult to find a logical model for analogical thinking.

This remark brings us to note the differences between our approach and the study of analogical thinking in cognitive science and A.I. A central difference between both approaches concerns (1) the constraints placed upon the representation of base and target situation and (2) the nature and function of the mapping constraints.

From our perspective an analogy between a target situation and a source situation is the (possible) result of thinking about something as something else. In thinking about x as *if* it is y , we *imagine* a world were x is y and *think* x is p (were p is a property of y). A thought process which finds its expression in counterfactual statements with counteridenticals. In thinking about the target *as if* it is the source situation we re-structure or interpret the target situation progressively in function of the source. Given a true counterfactual another is formulated, a major issue not addressed by theories of mapping. But also, there is no need to postulate an 'analogous' representation of the source and base situations. An analogical representation of the target situation *emerges* progressively in thinking about the target *as if* it is the base. It is the representation of the base domain which determines the analogical re-representation of the target situation through a process of formulating and evaluating counterfactual statements.

Looking at the mapping process from a conditional perspective another difference between both approaches is a difference concerning the constraints placed upon the antecedent. In computational approaches the mapping process is oriented towards finding appropriate mappings. The process of constructing correspondences is a process of finding antecedents which are also true in the actual world. That is, necessary true antecedents. Indeed, the formulation of true correspondences is necessary, but this only so if we want to conceive analogical thinking as justified reasoning by analogy. In other words, the process of finding identities, conceived as necessary, is a process of finding true antecedents enabling the inference of a valid conclusion. From our perspective, the process is not oriented towards finding necessary true antecedents. The correspondences are not necessarily true, they are counteridenticals. We imagine a world in which the antecedent is true in order to make the

consequent true. The identities which are counteridenticals are as conditions making the truth of the consequent possible. Seen from this perspective, analogical thinking is not an antecedent-but a consequent-driven process from its start. Instead of constructing analogical thinking as a two-step process in which first the necessary true correspondences are computed and then evaluated, the counteridenticals are progressively introduced and the constructed counterfactuals are immediately evaluated.

This does not imply that the process of constructing correspondences is free from constraints. But it says that the constraints are constraints concerning the *search* of correspondences and not constraints concerning the *selection* of necessary true correspondences⁶. Given the fact that finding a correspondence (a counteridentical) between an element of source and target situation is determined by the representation of both situation and the fact that the representation of the target situation is progressively restructured implies that the constraints *emerge* during the process itself and are not given in advance. When what type of constrain applies is determined by the re-structuring process itself.

4. Conclusion

As cognitive scientists we assumed that the process of evaluating analogical thinking paves the way to speak, (metaphorically?) about the cognitive process underlying analogical thinking. But in our approach we found that the semantic procedure for evaluating counterfactuals made it possible to speak about analogical thinking as a kind of *as if* thinking. A way of speaking which prevents us from looking at analogical thinking as a justification-oriented reasoning process. An approach in which the representation of base and target situation are not analogous representations and the justification-oriented constraints are not the cognitive wheels of analogical thought.

Seminar of logic and epistemology
Laboratory for applied epistemology
University Ghent

NOTES

1. This is the analysis of reasoning by analogy into first an inductive and then deductive step. An approach which goes

- back to Aristotle's syllogistic analysis of the 'paradeigma' (APr.II, 24, 68b37).
2. Black (1979) defends a similar approach towards metaphorical thought.
 3. This is a condensed version of an illustration of analogical problem-solving described by Gick & Holyoak (1980, 309)
 4. This law says that : if two terms 's' and 't' are imagined to be identical - are each other's images, under a certain mapping - you should substitute 's' for 't', in any sentence 'y'.
 5. Using the semantic systems for modal logic developed by Kripke, Stalnaker (1981, 45) specifies the following semantic rules (where $A \rightarrow B$ is a conditional - it may be a counterfactual - with antecedent A and consequent B):
 - $A \rightarrow B$ is true in w if B is true in $f(A, w)$;
 - $A \rightarrow B$ is false in w if B is false in $f(A, w)$
 A conditional is true in the actual world (w) when its consequent (B) is true in the selected $f(A, w)$ world.
 6. Compare for example the formal constraints the selection function must meet formulated by Stalnaker (1981, 46-7) with the constraints introduced by Holyoak and Thagard (1989, 304).

BIBLIOGRAPHY

- Ashley, K.D., 1988, 'Arguing by Analogy in Law: A Case-Based Model', *Analogical Reasoning*, ed. D.A. Helman, Kluwer Academic Publishers, Dordrecht, p.205-226.
- Black, M., 1979, 'More about Metaphor', *Metaphor and Thought*, ed. A. Ortony, Cambridge University Press, Cambridge, p.19-43.
- Boyd, R., 1979, 'Metaphor and Theory Change: What is "Metaphor" a Metaphor for?', *Metaphor and Thought*, ed. A. Ortony, Cambridge University Press, Cambridge, p.356-408.
- Carbonell, J.G., 1983, 'Learning by analogy: formalizing and generalizing plans from past experience', *Machine Learning*, eds., M.S.Michalski, J.G. Carbonell & T.M. Mitchell, Springer Verlag, Amsterdam, p.137-159.
- Davies, T.R., 1988, 'Determination, uniformity, and relevance: normative criteria for generalization and reasoning by analogy', *Analogical Reasoning*, ed. D.A. Helman, Kluwer Academic Publishers, Dordrecht, p.227-250.
- Falkenheimer, B, Forbus, K.D. & Gentner, D., 1986, 'Structure-mapping engine' Proceedings of AAI-86, Philadelphia.

- Gentner, D., 1983, 'Structure-mapping: A theoretical framework', *Cognitive Science* 7, p.155-170.
- Gick, M.L. & Holyoak, K.J., 1980, 'Analogical problem solving', *Cognitive Psychology*, 12, p.306-335.
- Hesse, M., 1966, *Models and Analogies in Science*. University of Notre Dame Press, Notre Dame, Indiana.
- Holyoak, K.J. & Thagard, P., 1989, 'Analogical Mapping by Constraint Satisfaction', *Cognitive Science*, 13, p.295-355.
- Kedar-Cabelli, S., 1988, 'Analogy - From a unified perspective', *Analogical Reasoning*, ed. D.A. Helman, Kluwer Academic Publishers, Dordrecht, p.65-103.
- Keane, M., 1988, 'Analogical mechanisms', *Artificial Intelligence Review*, 2, p.229-250.
- Mill, J.S., 1906, [1843], *A System of Logic*. Longmans, Green and Co., London.
- Perelman, Ch. & Olbrechts-Tyteca, L., 1958, *La Nouvelle Rhétorique*. *Traité de l'Argumentation*, Presses Universitaires de France, Paris.
- Schon, D.A., 1963, *Displacement of Concepts*, Tavistock Publications, London.
- Skorstadt, J., Falkenhainer, B., & Gentner, D., 1987, 'Analogical processing: a simulation and empirical corroboration', *Proceedings of AAI-87*, Seattle, WA.
- Stalnaker, R.C., 1981, 'A theory of conditionals' *Its*, eds. W.L. Harper et al., D.Reidel publishing company, Dordrecht, Holland, p.41-55.
- Thagard, P., 1988, 'Dimensions of Analogy', *Analogical Reasoning*, ed. D.A. Helman, Kluwer Academic Publishers, Dordrecht, p.105-124.
- Vandamme, F., 1987, 'Esquisses des strategies d'implementation pour des logiques d'analogies', *Capita Selecta uit de Toegepaste Epistemologie*, ed. F. Vandamme, Communication and Cognition, Ghent, p.77-93.
- Van Dormael, J. & Tessier, T., 1986, 'Analogies as a case study in Non-Fregean inference strategies', *Proceedings of the 2nd International Conference on Artificial intelligence*. Hermes, Paris, p.333-346.
- Van Dormael, J., (to be published), 'Analogical Reasoning: a logical inquiry about "Archaic Thought"', *From an empirical point of view*, eds. E.M. Barth & Van Dormael, J.
- Weitzenfeld, J.S., 1984, 'Valid reasoning by analogy', *Philosophy of Science* 51, p.137-149.
- Winston, P.H., 1982, 'Learning new principles from precedents and exercises', *Artificial Intelligence*, 19, p.321-350.