Analogy and Induction : which (missing) link ?

Antoine Cornuéjols & Jacques Ales-Bianchetti

Laboratoire de Recherche en Informatique (LRI), UA 410 du CNRS
Université de Paris-sud, Orsay
Bâtiment 490, 91405 ORSAY (France)
email : {antoine,ales}@lri.fr

Abstract : *In this paper, we argue that accounts of analogy should be consistent with the theoretical frameworks developed for related cognitive processes, such as induction. On one hand, this allows to more firmly anchor our theoretical perspectives on analogy, and, on the other hand, this may offer ways to improve on the current theories in the related fields. We propose some steps towards these goals.*

Keywords : Analogy, Theory and model for analogy, Machine Learning, Inductive learning theory, COPYCAT.

1. Introduction

The study of analogy confronts us with a formidable challenge. Its manifestations are seemingly ubiquitous : from perceptual processes responsible for recognizing concepts in "raw data", to categorization relying on perceived similarity, up to "higher" cognitive processes including communication through metaphors or creativity. It is definitively not to be ignored. But at the same time it is very difficult to study.

First of all, thanks to its multifarious aspects, it tends to be a slippery and hard to delimit notion. Many works (Indurkhya, 89) have made proposals to distinguish several types of analogies, emphasizing differences in purposes, a priori information and underlying processes. If some clarification results, at the price of complication, it remains to define precisely in each case both the goal of analogy (and the attached performance criteria) and the mechanisms involved.

Second, analogical reasoning is an unjustifiable (i.e. not logically valid) inference procedure. It goes beyond the deductive closure of the initial information and therefore cannot offer any warranty on its conclusions. But then what supports analogies ? What makes an analogy better than another one ? More concretely, why is it that it is so much used, apparently to the benefit of reasoning agents (as sanctioned by Evolution) ? Again, we encounter the problem of the evaluation criteria. More basically, the difficulty lies in the lack of firm referential system upon which to build and evaluate theories and models of analogy.

Responses to these problems have been twofold. One has been to seek some normative characterization of analogical reasoning whereby necessary conditions for sound inferencing are stated (Russell, 1987). Unfortunately this interesting approach so far has delivered very restrictive conditions that in effect exclude much of the subject matter. The other approach takes natural reasoning agents, prominently human ones, as standards against which to measure the quality of analogies and of the mechanisms that produce them. But of course, these natural yardsticks are subject to many parameters (perceived context, implicit goals, cultural background and so forth) that are impossible to securely control. Therefore this opens the door for endless arguments about the relevance and validity of each new experiment, and consequently of the tested models.

It is noteworthy that in this context, what is evaluated are not so much the end results of analogical inferencing, but rather the processes that are assumed to play a key role in their production. For instance, once it has been hypothesized that similarity judgments are at the core of analogical reasoning (and many other cognitive processes as well), theories, models, and arguments center on similarity measurements and what they involve, in effect evacuating the fundamental question of why a high degree of similarity between a source case and a target case should entail highly reliable transfers of information from one to the other (leaving aside both the important issue regarding the objective nature of similarity (Medin et al.,1993) and the question of the modus operandi of these transfers).

This overall situation : a subject matter concerned with an inferencing process both presenting seemingly many different facets and manifestations, and inherently lacking sound justification, is reminiscent of the situation faced by the students of induction ten to fifteen years back. There also, there were plenty of models for inductive reasoning that were assessed on the face of their measured performance on chosen benchmarks, and a corresponding need for an established theory. The situation has changed recently (mostly thanks to Vapnik (1995), Valiant and many brilliant co-workers of the COLT (Computational Learning Theory) community).

This apparent aside on induction points out a third potential way of approaching analogical reasoning. Since it is supposed, rightly, that it is a core component of many cognitive processes, it should not be an isolated point with regards to its internal working and its performance criteria. In other words, properties and principles uncovered in studying other fundamental cognitive processes should hardly be expected not to be shared, at least in part, with analogical reasoning. Consequently, any theory and model of analogy should be consistent with theories and models for other, related, faculties. This could, and should, provide for good anchor points on which to erect models of analogy.

This is indeed the track that we take in this paper. In a way, we are pursuing a very ambitious goal, that of uncovering some fundamental traits that would constitute the basis for an overall theory that would encompass several cognitive faculties, including of course analogy making. We propose not to find justifications for analogical inferencing, an hopeless pursuit, nor to assess the value of one's model by comparison with natural reasoning agents, something necessary but not sufficient and never to be completely satisfactory nor convincing, but to present a theory of analogical reasoning that both satisfies a reasonable criterion for analogy, and at the same time is consistent with existing theories of inductive learning, a process that we argue is intimately related to analogical inferencing.

This paper presents the current state of this endeavor. Section 2 argues that analogical reasoning and induction are intimately connected while at the same time being different in important aspects. It also sums up the current state of accepted theories of induction. In section 3, we present our own model of analogy, showing in which respects it is intuitively appealing and how it maintains closed links with theories of inductive learning. Section 4 demonstrates on a canonical example that the model yields realistic results. Finally, section 5 sums up the state of this project and points to directions for future research.

2. Analogy and induction : resemblance's and dissimilarities

Deeply rooted in analogy surely rests the notion of similarity. At the least, analogy induces similarity, sometimes totally unexpectedly, as in creative analogy. The objective nature of similarity is the object of active debate within psychological circles (Medin et al. 1993), but it undoubtedly underlies categorization too : similar things tend to be grouped together in cognition. Analogical reasoning also shares many common points with induction, as we see now.

2.1 A view on inductive learning and its theory

Figure 1 provides a flavor of what we are up to in inductive learning. A collection of examples, the *learning set*, is given, consisting of pairs $(x_i, f(x_i))$, and the goal is to infer what value would take the hypothetical and unknown function f on new points x_j. Generally, there is a cost associated with errors on $f(x_j)$, also called the *risk*, so that inductive learning consists in finding an hypothesis h such that the risk averaged[1] over the space of all possible instances, or the *expected risk*, be minimal.

[1] More precisely, the averaging is weighted by the distribution over the instance space, so that more weight is given to dense areas, where it is more likely to encounter future events.

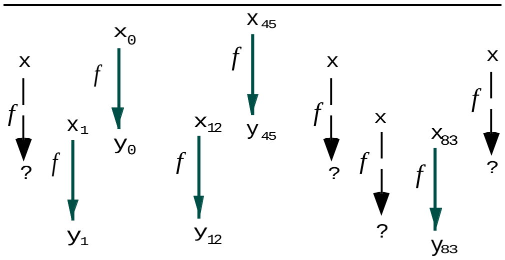


Figure 1. Inductive learning (in the supervised setting), consists in identifying a function *f* that "explains" the learning data (set of pairs $(x_i, f(x_i))$ and making the inference that the same *f* applies in unseen instances.

Before the large diffusion of theoretical studies of induction (Vapnik,1995), the common view was that the obvious learning strategy was to select an hypothesis minimizing the risk over the learning set, called the *empirical risk* since it is measurable, in order to automatically get the optimal hypothesis with respect to the expected risk (one that by nature is unknown). This belief has been formalized and given a name : the *Empirical Risk Minimization principle* (ERM for short). In essence, what this principle states is that the best account for the learning instances is ipso facto the best one also for yet to be observed events. Vapnik, and many other theorists in the last fifteen years, have disproved this naïve view.

Of course, the philosophers knew this all along. There cannot be any miraculous basis for inducing general laws from specific observations. But theorists of inductive learning have gone further, specifying sufficient conditions for induction to be a reliable source of inferences. Sketched in broad lines, the now "classical" theory of induction states that induction is possible and reliable in proportion that the set of potential candidate hypotheses considered by the learner is restricted[2]. In other words, a learner that is able to explain any data set is hence unable to make induction, while a learner that can only consider severely restricted classes of concepts, if with these it may explain the observed data (available in sufficient quantity), is justified to generalize to other, as yet unknown, cases. Given that there is no "free lunch", the problem is now to chose a priori the right set of hypotheses.

It is noteworthy that, according to these theories, the confidence that one may put in inductive learning only depends on statistical quantities characterizing the hypothesis set taken as a whole, as well as the distribution and the number of learning instances.

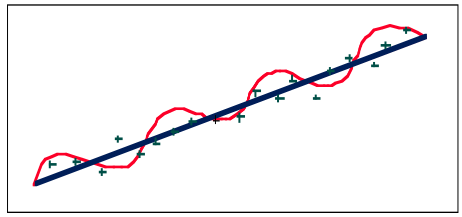


Figure 2. The best model for the data points is deemed to be the one that is at the same time "simple" and fits well to the data. Here, the linear model is simpler to specify than the polynomial one, and seems to fit equally well (or bad ?) the data points. Hence, following the MDLp, it should be preferred.

Other theoretical approaches to inductive learning share this property. These are the bayesian perspective on learning and the related *Minimum Description Length principle* (MDLp). Roughly, they prescribe to select the hypothesis which is maximally probable given the observed data and their a priori probabilities (something that is easily computed with Bayes formula). The MDL view replaces this principle by one where one should chose the hypothesis such that the sum of its code length (within some well chosen coding schema) and the length of the description of the data encoded with the hypothesis be minimal (figure 2 illustrates this). It is a remarkable fact that it can be proved that the Vapnik theory and the MDL principle, starting from widely different premises, are nonetheless tightly linked. A fact that reinforces the confidence in these theories.

This is all good and well, but does it have something to do with analogy ?

2.2 The same, yet different

As already noted, there are several types of analogies. Some involve the comparison of two given items (e.g. "abc" and "122333"), and some the completion of one item given the other (e.g. if "abc $\rightarrow$ abd", what should be the completion of "aababc $\rightarrow$?"). This last case (due to Hofstadter and his co-workers (Mitchell, 1993), (Hofstadter, 1995)) is a tricky one. We do not mean here that it might

[2] Technically, these restrictions concern the possible partitions of the instance space that are induced by the hypothesis set. They are measured via statistical quantities, the most famous one being the Vapnik-Chervonenkis dimension.

be difficult for the reader to infer the completion "aababcd", but that this is just a good example where one is made aware of the fact that much more has to be inferred. Indeed, nothing is given about the ways the strings (are they really ?) should be perceived, nor about the dependence relationship between "abc" and "abd" in the source case. Worse yet, the perception and interpretation of the source depends on the target probe. Had the last one be here "American Broadcasting Corporation → ?", that the source "abc → abd" would have been thought of completely differently. It is therefore evident that this type of analogy encompasses the former one where no completion, other than completion of interpretation, takes place. This is why we will consider this one type here.

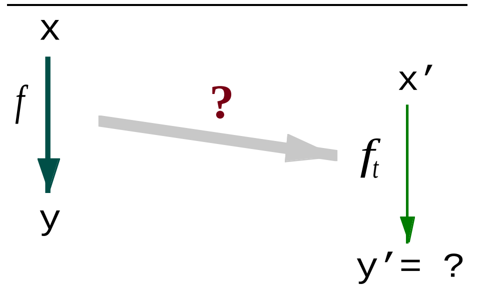


Figure 3. One view of analogy making enhances its inherent inferential aspect from limited information. Only x, $f(x)$, and x' are known to the agent. From these "raw data", the agent must infer their interpretation, the dependence relation f in the source, and the corresponding "transported" dependence relation f_t in the target. From this follows y'.

If now, we take a look at figure 3, it may strike us that analogy is but a limit case of induction where one has access only to one learning instance. Under this perspective, analogy and induction are the same. And this is why we argue that surely their respective theories should be consistent so that they merge in between where very few learning instances are available.

On the other hand, there exist significant differences that make problematic the simple extension of the classical theories of induction to analogy, but also, as we will see, offer the perspective of refining these existing theories beyond their current state. Here is a list of these differences.

- The prediction is to be performed on one point only, not on the whole potential instance space. The notion of expected risk is therefore undermined to say the least.
- Each item potentially has its own referential frame (as in "abc → abd"; "122333 → ?", or better yet in "abc → abd"; "American Broadcasting Corporation → ?"). This is in contrast to the unstated assumption in induction that the looked for hypothesis f is the same all over the instance space.
- The target plays an important role in analogy, shaping the interpretation of the source, while it does not intervene in any ways in existing theories of induction.
- Finally, may be as a consequence of the above points, it is strongly believed that the "distance" between the source and the target plays a key role in analogy. In contrast, there is no notion of distance between instances in inductive learning[3].

To sum up at this point. We believe that the study of analogy should deliberately take into account related cognitive processes, such as categorization and induction, and try to make contact with the theories therein. This would more firmly anchor tentative theories and models for analogy. At the same time, developing theories adapted to the specific demands of analogy offers the perspective to refine the theories of the related cognitive process. To be more specific, incorporating the notion of distance between instances, and/or of local referential frames, into the theory of inductive learning, in needed harmony with theories of analogy, should result in finer theories of induction. Theories that, for instance, would better predict which amount of information is needed in order to be able to learn, say, some classes of concepts.

This is in accordance with this philosophical outlook that we have undertaken to look for a theory of analogy, one that would be faithful to the phenomena, and be related to theories of inductive learning.

3. A proposal

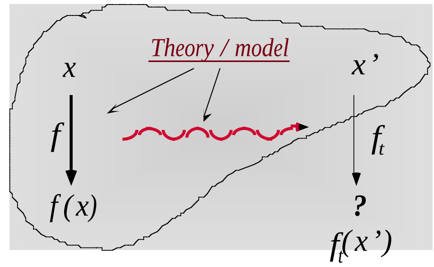
Let not be misled here, we are not, at this point, looking for the specification of a reasoning mechanism that would be a candidate for modeling analogy making, but we aspire to

[3] Beware not to confuse the notion of distance between instances, as in analogy, and between hypotheses or an instance and an hypothesis, as can be the case in induction.

find **a criterion for evaluating candidate analogies**, a criterion that the best analogy should optimize. Recalling figure 3, it is clear that this criterion must depend on what is known to the reasoning agent, i.e. the source : x and $f(x)$ (in the best of case including f itself), and the incomplete target : x'. It should also depend on prior knowledge which is the basis for the interpretation of the situations.

In addition to this, and following our policy, we should find a criterion that is consistent with the theory of induction. In particular, this criterion should take into account the "entropy" of the candidate hypotheses space, or, more intuitively, of the complexity of the candidate hypotheses. The idea being that the more underlying regularities are discovered in the data, the more its expression can be compressed. The MDLp is one expression of this general doctrine. We should therefore look for a measure of parsimony. The best analogy should correspond to the discovery of regularities both in the source and the target, regularities that should be as interrelated as possible. This last point being in agreement with a third desiderata : that the evaluation criterion reflects in some way our anticipation that analogy is linked to a notion of perceived similarity or distance between the analogs.

An evaluation criterion for analogy



In figure 4, we show how a version of the MDLp could be adapted to analogy. The best analogy should be the one that minimize the cost of the models or interpretations on which are based the perception of $(x, f(x))$ on the one hand, and, on the other hand, of x', while at the same time minimizing the cost of translating the interpretation of the source to the interpretation of the target. This is what is expressed in the following proposition.

Given M_S, M_T and f, it is easy to derive f_t by $f_t = \mathrm{pgm}_{MS\text{->}MT}(f)$, that is the transformation of the expression of f within the referential associated with M_S by the program that transforms referential M_S to referential M_T. Then $f_t(x')$ may be computed.

Proposition :

The set of models and descriptions M_S, M_T, x, f, x' that minimizes the formula[4] :

$$Total\_length = L(M_S) + L(x|M_S) + L(f|M_S) + L(M_T|M_S) + L(x'|M_T)$$

is the one associated with the best analogy between the source and the target.

[4] L is taken to be a function measuring the cost or length of its argument expressed in bits. We do not dwelve here in technical details about what that involves. We refer the reader to (Li & Vitanyi,1993) for a thorough introduction to algorithmic complexity theory on which our model is based.

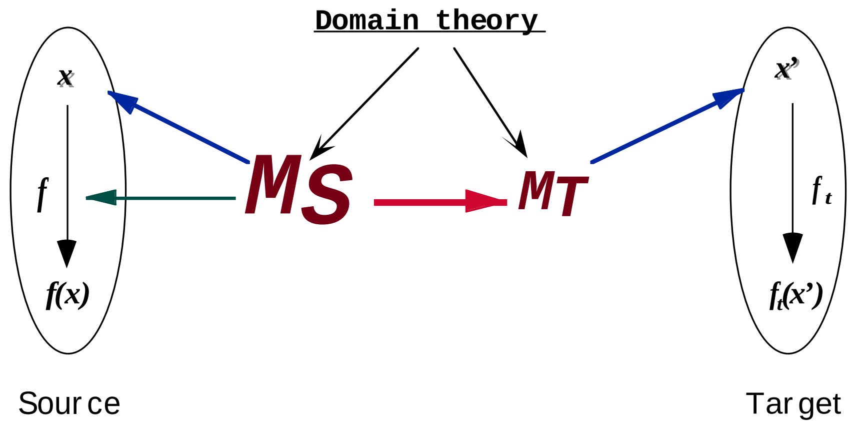


Figure 5. Following the theory presented here, any analogy involves interpretations or models, constructed from prior knowledge (the domain theory) that are local to the source : M_S, and to the target : M_T. From these, the specifics of each case can easily be reconstructed. At the same time, analogy making implies that a relationship be identified between M_S and M_T such that the two seem similar to each other. We submit that the best analogy is the one that minimizes the overall cost of specifying the models, there relationship (from M_S to M_T) and the derivation of the specifics of each case.

4. Illustration

This section intends to illustrate the above conceptualization. It is not meant to demonstrate its value as a model of the human ability in making analogies. This is beyond the scope of this short paper, and would require a careful discussion of representation primitives, suitable coding system, and hypothesized prior knowledge.

4.1 The domain

In order to keep things manageable, we have chosen a domain where it is easy to define representation primitives and theories, and yet which presents enough richness to be demonstrative of the wealth of issues in analogy-making. This domain is inspired from the microworld developed by Hofstädter et al. for the COPYCAT project (Mitchell, 1993).

The basic objects in this world are the 26 letters of the alphabet, but it would be straightforward to add numbers or geometrical shapes. The task consists in finding how a letter string is transformed given, as an example, another string and its transform. For instance, given that **`abc => abd`** (the source), what becomes of **`iijjkk => ?`** (the target). The problem, quite familiar in IQ like tests, is thus to identify the relevant aspects and transformation at work in the source that can best be mapped to the target problem. It is very easy to make up a whole variety of problems that test the range of analogy-making.

Following (Mitchell, 1993), the background knowledge or domain theory includes the basic representation primitives and the conceptual structures that allow to describe and highlight various aspects of the situations at hand (see table 1). In order for the quality criterion to be computable, each construct is associated with a number, that corresponds either to a prior probability from which it is easy to draw the related length using the relation $L=-log_2(P)$ (e.g. the concept of `string` is associated with the prior 1/8, hence is of length 3 bits), or directly to a length in bits (e.g. the concept of `nth` requires n bits). These numbers can be modified either manually or through learning to yield various biases corresponding to a variety of contexts or prior knowledge.

- Features describing the conceptual structures :
  - orientation (-> / <-) *1 bit*
  - cardinality or number of elements : n *$log_2(n)$ + 1 bits*
  - length : l *$log_2(l)$ + 1 bits*
  - starting or ending with element = x *L(x) bits*
- **Letter** *(1/2)*

  Particular letter (e.g. 'd') *(1/2.26)*
- **String** (orientation,elements) *(1/8)*

  L = 3 + L(orientation) + $\sum$ L(elements)

  e.g. L('a3bd' with orientation = ->) = 3 + 1 + $\log_2((1/2.26)^3$ + L(3)

  = 3 + 1 + 18 + 3 = 25 bits
- **Sequence** (orientation, type of elements, succession law, length, starting or ending with) *(1/8)*

  L = 3+ L(orient.) + L(type) + L(law)+ L(length) + L(start/end)
- Description and length of a **succession-law**

  succ(type-of-el.,n,x) ≡ the nth successor of the elt. x of type-of-el.

  L = L(type) + L(n (see below)) + L(x)

  L(n) = L(1/6) if n=1 or -1 (first successor or predecessor)

  L(1/3) if n=0 (same element)

  $L((1/3).(1/2)^p)$ otherwise (with p=n if n≥0, p=-n otherwise)
- First / last *1 bit*
- nth *n bits*

Table 1. List of some representation primitives with their associated description length either in bits or defined as probabilities.

Hence, the string **abc** could be described as:

'abc' ≡ **String**	(1/8)
orientation : ->	(1/2)
1st='A', 2nd='B', 3rd='C'	$(1/4.26)^3$
TOTAL Length :	**21 bits**

or else as :

'abc' ≡ **Sequence**	(1/8)
orientation : ->	(1/2)
type of elements = letters	(1/2)
succession-law :	
succ(elt(letter=x) = elt(succ(letter,1,x))	
L(letter) + L(1st succ) + L(x)	
= L(1/2.1/6.1) = 4 bits	
length = 3	3 bits
starting with element(letter='A')	(1/26)
TOTAL Length :	**17** bits

It is clear that the last description, which more fully represents the structure of the string **abc**, is the most economical one, even though it describes it more completely than the first description which corresponds to the perception of a set of three letters.

4.2 Experiments

We have tested the above scheme on a variety of analogy problems in order to see what rankings the criteria would give to various possible solutions. Limited space prevents us from giving a full account of the derivation of the complexity figures. The overall method is as follows. For each pair (Problem; Solution), we hypothesize associated models or perceptions. For instance, **iijjkk** can be perceived as a string of letters, or alternatively as a sequence of successive pairs of letters. Then, a program computes the algorithmic complexity of these constructs and of the transformation programs that allow to derive one description from another. The associated figures are reported in table 2.

Problem: **abc => abd ; iijjkk => ?**

Solutions :

S1 : *"Replace rightmost group of letters by its successor"* **iijjkk => iijjll**

S2 : *"Replace rightmost letter by its successor"* **iijjkk => iijjkl**

S3 : *"Replace rightmost letter by D"* **iijjkk => iijjkd**

S4 : *"Replace third letter by its successor"* **iijjkk => iikjkk**

S5 : *"Replace Cs by Ds"* **iijjkk => iijjkk**

S6 : *"Replace rightmost group of letters by D"* **iijjkk => iijjd**

5. Conclusion and perspectives

These experiments and calculations, cannot and do not pretend to be conclusive. They rely on many hunches and simplifications that would need to be more carefully set. Indeed, it is natural that such be the case, since this proves by the same token that our model nicely incorporate contextual effects and the possibility of learning (concepts and associations), and of the consequences these may have on analogy making. Still, these results show that the proposed scheme does not seem entirely unreasonable from the point of view of a comparison with natural cognition. But we also believe that most promising is the fact that this model is tightly linked with induction theory. Nonetheless, it remains unclear why a high degree of similarity, or the possibility of a simple interpretation of the analogs lends credit to the analogical inference. This is a question we actively study.

Else, one of our current research project is to better ground our calculations on the theory of algorithmic complexity, to maintain close links with inductive theory, while at the same time experimenting with many more examples from a variety of domains. We also study how mechanisms for the actual production of analogies (not only for evaluation) could be derived from this perspective.

	S1	S2	S3	S4	S5	S6
$L(M_S)$	10	9	11	11	12	11
$L(x\|M_S)$	8	18	18	18	22	15
$L(f\|M_S)$	4	4	3	7	8	3
$L(M_T\|M_S)$	5	0	0	0	0	17
$L(x'\|M_T)$	8	36	36	36	42	15
Length (bits)	**35**	**67**	**68**	**72**	**85**	**62**
Rank	1	3	4	4	6	2

Table 2 : The figures corresponding to the evaluation formula are reported for various solutions to the problem considered. Solution 1 emerges as a clear winner, which is also the choice of most human subjects when asked to rank these solutions.

References

• Hofstadter D. (1995). *Fluid Concepts and Creative Analogies*. Basic Books.

• Indurkhya B. (1989). Modes of Analogy. LNAI-397, Springer-Verlag.

• Li & Vitanyi (1993). *An introduction to Kolmogorov complexity and its applications*. Springer-Verlag.

• Medin, Goldstone & Gentner (1993). Respects for Similarity. *Psychological Review*, 1993, Vol.100, No.2, 254-278.

• Mitchell M. (1993). *Analogy-Making as Perception*. MIT Press.

• Russell S. (1989). *The Use of Knowledge in Analogy and Induction*. Pitman Publishing and Morgan Kaufmann, 1989.

• Vapnik (1995). *The Nature of Statistical Learning Theory*. Springer-Verlag.