AN ALTERNATIVE ACCOUNT ON THE ORIGIN OF ANALOGICAL REASONING

Cesare Bianchi
UCD School of Computer Science and Informatics, Dublin, Ireland

Abstract

This work describes a model of analogical reasoning that diverges from all the “standard” models of analogical reasoning produced so far. The “standard” models are focused on the analogy between completed structures, while the model described here is focused on experiments where novel similar concepts are learned simultaneously, not necessarily starting from a well-known domain as a source of knowledge. The starting hypothesis is that analogy can be used between simultaneously learned categories. A subsequent hypothesis is that learning is usually split in two phases: some general "partial-categories" are formed first, which are then modified to arrive at the final categories. This model confirms the predictions originating from these hypotheses: 1. learning of similar categories is mutually related; 2. final learning is preceded by a phase of partial learning; 3. people prefer similar criteria over easily discernible ones; 4. people operate under memory constraints.

1. **Introduction**

Analogical reasoning is the ability to spot similarities between concepts or domains and to use those similarities to transfer knowledge from a concept to another (Gentner, 1981; 1983). The most common example of analogical reasoning is the completion of quadruplets like

Boat : Sea :: Aircraft : ?

Some remarkable cases of analogical reasoning can also be found in the history of science (Del Re, 2000; Hoffman, 1980), where it is used to discover new knowledge in a partially unknown domain being investigated. The typical example of this use is the analogy between the solar system and the atom. This case is representative also of the third common use of analogical reasoning: to teach a novel concept by using another well-known concept as the source of knowledge.

In real life novel concepts and domains are often learned simul-taneously, not necessarily starting from a well-known concept as a source. This early use of analogical reasoning can have advantages: instead of learning several concepts separately, the learning process can be unified, thus minimizing the effort both in terms of memory and of time.

Cognitive Systems 7-3, November 2012, 261 - 273

Therefore, we start from the hypothesis that analogy can be used between simultaneously learned categories. A subsequent hypothesis is that learning is usually split into two phases: a first phase of partial learning is followed by a phase of refinement; some general "partial-categories" are formed first, which subsequently are modified to arrive at the final categories. This minimizes memory and time needed, while maximizing the amount of information extracted in each stage.

Because of this advantage, it may be expected that people prefer to find classification criteria that have mutual similarity rather than easily discriminable criteria, even when both types are available. This contrasts with the standard views of categorization (Rosch, 1978; Nosofsky *et al.*, 1994; Ashby & Maddox, 2005), and also with machine–learning models (Lavrac & Dzeroski, 1994; Michalski, 1983), which predict that easily discriminable categories are more easily learned.

This two-step process may also give an emergentist explanation of the origin of analogy: concepts are similar because they stem from a common partial-category.

From these hypotheses originate a series of predictions: 1. learning of similar categories is related (i.e. time between learning of similar categories is less than time between learning of dissimilar ones); 2. before learning is complete, the errors are not random but they are more frequent inside the partial-categories; 3. when given alternative solutions, people find similar criteria instead of dissimilar ones.

To test these hypotheses and predictions, an appropriate task is needed in which novel similar concepts are learned simultaneously. One perfect candidate is a task of category learning in which some of the categories are similar.

It has been suggested before that analogical reasoning has a role in category learning, but only for the discovery of similarities between items of the same category (Gentner & Medina, 1998; Gentner & Namy, 1999; Sloutsky & Fisher, 2004). The case proposed here, i.e. when there are similarities between categories, has not yet been investigated.

In order to have categories with similarities between them (similarities that can be exploited only by using analogical reasoning), those categories must be defined not just by features, but by relations (Gentner & Kurtz, 2005). This is the case of categories such as "pilot", "captain", "driver", that are defined by the relation of their instances with the vehicles concerned, the relations with other people, the environment, and so on. Those relations are similar, thus one category can be easily mapped onto the others by using analogical reasoning.

2. **Experiments**

2.1. *First Experiment*

2.1.1. *Design*

The first experiment is a category-learning test, with categories defined by similarity relations. In order to avoid the influence of any previous knowledge, we used categories of a graphical abstract nature, composed of geometric shapes of various colours. Two categories were defined by relations and two by features. All the items of the categories were composed of coloured elements, randomly distributed over the items.

The two "relational" categories (indicated by R_1 and R_2) were defined by the numerical relation between the elements. In one category (R_1) the number of elements of one type was the same as that of another type, in the other category (R_2) it was twice as many. The two "feature" categories (indicated by F_1 and F_2) were instead defined only by the presence of a distinctive element. All the items also contained distracting elements. For each item the participant had to choose an answer by clicking on one of four buttons, then had to confirm his answers by clicking on another "Ok" button. After confirmation, a feedback was given: the buttons changed colour, showing the correct answer in green and the given answer, if wrong, in red.

2.1.2. *Results*

For each category a "moment of learning" was computed by using the numbers of correct answers. In order to evaluate the relation of learning of similar categories, the intervals between the learning of the two "relational" categories ($R_{1,}$ R_2) were compared to the average of the other intervals (OtherIntAvg). A sign test showed that R_1R_2 < OtherIntAvg ($p < 0.05$). The same happened for the intervals between the "features" categories (F_1F_2). This means that the two "relational" categories and the two "features" categories are related to each other. This result supports the prediction that learning of similar categories is mutually related, and thus the hypothesis that analogy can be used early in learning, between simultaneously learned cate-gories. Instead the sign test of R_1R_2 vs. F_1F_2 was not significant ($p> 0.05$).

In order to test the hypothesis that learning has two stages: a first phase of partial learning, followed by a phase of refinement; an analysis of errors (see below for an explanation) was performed. The incorrect answers given by a participant before learning of any category has occurred, can be a good indicator of how the learning process proceeds for that participant. Three kinds of mistakes were computed: classifying an item in the other "relational" category, classifying an item in the other "feature" category, and completely incorrect answers. The counts for each kind were compared to their random expected values (estimated by using the marginal means).

The pattern that emerges depicts a middle period of "partial learning" of both the relational and features categories. In fact the mistakes between similar categories occur more frequently than expected at random (sign test $p < 0.01$), while the mistakes between dissimilar categories (i.e. between different kinds of categories) are not different from random values ($p > 0.05$; power > 0.8). These results support the hypothesis that analogical reasoning starts to operate from the very beginning of learning, and helps in finding similarities, even between categories that are only partially learned.

2.2. *Second experiment*
2.2.1. *Design*

The second experiment introduces the factor of preference for similar or separable concepts. Thus participants were given two alternative solutions to the problem: one solution involved the use of criteria with similarities between them, across different categories. The alternative solution had different criteria for each different category. Participants needed just to find one solution; according to their behaviour, it had to be clear which solution they had found. The resulting design is summarized in Table 1. There were four different categories (plus the residual one), two of which were "dual", i.e. defined by two alternative criteria: one, the "analogical criterion" (represented by C_x, which stands for "common"), was structurally similar to the analogical criteria of the other categories, whereas the other "non-analogical criterion" (represented by A_x, which stands for "alternative") was unique to that category. Category A was introduced to balance the difficulty of the non-analogical criteria (A_1, A_2 and A_3) taken alone, and to assess their individual difficulty.

Category	*D1*	*D2*	*S*	*A*	*R*
Criteria	C1 + A1	C2 + A2	C3	A3	only distractors
	C1 + A1	C2 + A3	C3	A2	
	C1 + A2	C2 + A3	C3	A1	

Table 1: Category definition abstract criteria. Dual categories (D_1 and D_2) have two alternative criteria, one Common and one Alternative. Some others (S and A) have only one. D stands for dual, S for single, A for alternative, C for common. The residual category has no criterion.

Subjects could solve the task in different ways. If, for example, they find rules C_1, C_2, C_3 and A_3 (for the first group), it can be inferred that similarities are preferred over the separability of concepts. If otherwise they find rules A_1, A_2, C_3 , and A_3, a better and clearer separability is preferred.

Secondly, interactivity was introduced in order to make the test entertaining and resembling a game. A number of experiments on analogical reasoning which used static figures had already revealed several limits: they had little ecological plausibility, caused little interest, and raised poor attention in participants. Moreover, a task with interactive items could introduce, among others, causal and synchrony relations, which are essential in our everyday–life. More importantly, interactivity allowed to discover which criterion was found for each of the categories defined by two alternative criteria. In fact it was possible to record all the interactions (namely: clicks) of the user with the elements. Given that some elements pertained to one criterion and some other to the alternative criterion, when the participants clicked on an element, it was clear which criterion was shown. After a criterion was found for a category, for all the subsequent items the participant started to click randomly until the group of elements was found corresponding to that criterion, then the correct answer was given. Therefore the last click before an answer was given, was with high probability on an element according to the criterion discovered. The analysis of these last clicks can reveal which criterion (for the dual categories) the participants discovered.

As in the previous experiment, the items were composed by geometric shapes of different colours. Each item contained 3 groups of elements. Each group was formed by elements of the same shape and colour. There was aone-to-one correspondence between groups and criteria. So in a dual category, two groups were active, one for each alternative criterion. While for a single category, only one group was active. All the active elements when clicked elicited a reaction. The reactions could be of different kinds: a piece of music, or a tone could be played, some elements could move, or both. All the elements associated with the common criteria only played a piece of music, always the same piece of music for the same category. There were also other pieces of music in the test, which were sometimes played when subjects clicked on distractors. All the elements of the group for the A_1 criterion made the entire group react with the same action in synchrony. All the elements of the A_2 group also made the entire group react all together, but with heterogeneous actions. Finally, the A_3 criterion introduced synchrony between volume and movements: when an element was clicked, all the elements of the group gave the same action in turn and a tone was played, of which the volume was changed in synchrony with the action.

For each item the participant had to choose an answer, tagging the item with a label. A positive feedback was given if the answer was correct. Otherwise a negative feedback was given, and the participant could try with another label, until the correct answer was found. Then a "Next" arrow but-ton appeared, to go on to the next items. A “Back” arrow button was always available to go back to the previously seen (and correctly answered) items.

2.2.2. *Results*

The novel factor introduced in this experiment was the presence of alternative solutions. This allowed the possibility to test the prediction that when given alternative solutions, people find similar criteria instead of dissimilar ones. This prediction contrasts with the standard theories of category learning, which instead predict that the more the categories are separable in the representational space, the easier it is to discern them.

For the dual categories, after their learning points, the last clicks before answering were counted both for analogical and non-analogical elements. A sign test was done to confront the last clicks on analogical and non-analogical elements. The clicks on the analogical elements were significantly more numerous than on the non-analogical elements ($p < 0.01$), showing that the analogical criteria were found more often than the non-analogical ones. (The model behaves as the human participants, so for the analysis it is undistinguishable. That is, to discover the - possible - action associated to an element, the model has to "click" on that element, and only then it is told what "happens", exactly as for human participants.)

This is the new and most important result of this experiment: even if given an alternative, the participants found the analogical criteria. Therefore the drive to use similarities is stronger than the separability of concepts. This is another confirmation of our hypothesis that analogy can be used between simultaneously learned categories. Not only can analogy be used, but it is preferred, since in this way it is possible to minimize the constraint on memory, and the time required.

In order to test the relation of learning of similar categories, a method similar to the previous experiment was used. A sign test showed that the average of the intervals between analogical categories is less than the average of the other intervals ($p < 0.01$) in most of the cases, meaning that the analogical categories are related to each other.

In order to test the early use of analogical reasoning, a method similar to the previous experiment was used. Also in this experiment a middle period of "partial learning" emerged. In fact, the trends are the same as in the previous experiment, thus confirming those results.

It is also interesting to note that people who use more often the previously seen items (going back to compare them to the present items), solve the test more easily ($r = - 0.556$; $p < 0.01$). This result is consistent with the theory that some people use a more scientific strategy based on the falsification of hypotheses, and some others use instead a strategy based on reinforcement. Moreover, this result is a confirmation of the hypothesis that people operate under memory constraints.

2.3. *Third experiment*

2.3.1. *Design*

During the execution of the second experiment, doubts arose whether the analogical criteria were preferred because they were similar, or because music was a feature more salient than movement. This third experiment was performed in order to exclude that music was more salient than movement, and to confirm the result in a more general context. Thus, it is a generalization of the second one, and the analogical criteria are defined by movement and synchrony instead of music. The hypotheses tested in this third experiment are the same as in the second, as well as the predictions, the constraints that arise from them, and the analyses performed.

2.3.2. *Results*

The results confirm and generalize what was found in the second experiment. Although the preference for similar criteria over separable ones is weaker than in the previous experiment, it is still present. This means that in the previous experiment the preference was not caused by a greater salience of music. When the similar criteria are defined by movements, they are still preferred to separable criteria. Therefore the similarity between criteria is the crucial factor, and not some bias introduced by a stronger salience of some feature: the drive to use similarities is stronger than the separability of concepts.

3. **A Computational Model**

On the basis of the findings in the experiments, and according to the assumptions for the theory described in the introduction, a computational model was constructed. This model diverges from all models proposed until now (e.g. Dora, Seql, Lisa, etc - Doumas & Hummel, 2005; Kuene et al, 2000; Markman & Wisniewski, 1997). These models cannot explain the results of the present experiments, since they are based on different learning processes. Instead, it is more reasonable to adopt a model that can lead to a change of ideas, and *that has a unified process of reasoning*. It is a common process to reason in steps: first to make an assumption, then restrict it, then widen it again, include something, exclude something else. *When learning various things at the same time, use all the available information from all of them.* A good strategy then is to make a model that can lead to hypotheses, modify them if appropriate, and then use this information also *to create hypotheses for other categories*. It is not necessary that the first hypotheses are completely correct: they are retained even if they are only partially correct, so that they can undergo a process of modification.

The tests used for the performance of the computational model are very similarly to those for (human) subjects. For experiment 1, the model is shown one item, and is asked to give an answer, then it is given feedback about the right answer. For experiments 2 and 3, answers are to be given until the correct answer is found. In each case, from the feedback the model can learn the correct classification.

As stated, these changes can be in any direction, which is different from the other models. In this way partial categories can be created, and then subsequently refined to form the final categories. This simple mechanism can explain many results of the shown experiments, and can even give an emergentist account of some forms of analogical reasoning. In the simplest case, the model first finds the correct rule for one of the similar categories, then, using some random changes to this rule, it creates some other similar rules, until in a few steps it finds the correct rule for the other category. Not only is it computationally less “expensive” than finding both rules independently, but it is also less “expensive” than explicitly mapping knowledge. In order to use structure–mapping, some aspects of the target must be known, and this initial acquisition would only lengthen the process. A more interesting case is when some partial rule is found first: that is, a rule that is good for two (or more) categories, in other words a "partial category". After this initial stage of partial learning, the model tries to modify this rule until in a few steps it finds the two final rules, which are similar, for they descend from a common ancestor. As for the former case, no explicit mapping has taken place, but the process can be described as analogical reasoning.

3.1. *Model Architecture*

3.1.1. *Overview*

The general structure and working of the model is very simple. It has a memory for hypothesized rules, and another memory for discarded rules (in order that they will not be recreated, or used again). Each rule consists of a predicate, and a list of valid categories, and has a weight. This is the first main difference from the existing models: rules can be partial or final. Obviously, a rule that is valid for all the categories, or for none of them, is discarded. But it is possible to form temporary "drafts" of rules, valid for more than one category. Then they will be refined through modification, which is another salient new aspect of this model; one of the ways in which a rule can be created is by randomly modifying an existing rule.

The model accounts for all the possible ways a rule can be created:
1. it can be based on a shown item, taking some of its properties or hypothesizing a relational structure between them;
2. it can be created randomly;
3. it can be created with the modification of an existing rule;
4. it can be created with the unification (intersection) of two existing rules.

Therefore it can account for almost all kinds of strategies, that e.g. are based on experience, creativity, trial and error, and logic. The use of each method, as well as other aspects of the model, are parametrized, so with different parameter values the model can reproduce the behaviour of different participants

3.1.2. *Answering Phase*

The answering phase is quite simple. Starting from the oldest rule, the model seeks in memory a rule of which the predicate is true for the given item. If no rule is found, a random answer is given. If instead a rule is found, it can be final or partial (that is, valid for more than one category). In this latter case the answer is randomly chosen between the categories for which the rule is valid.

In the first phase, in which the rules are generic or wrong, the model is expected to give wrong or random answers. Then, if it finds partial rules, it will make mistakes similar to those of the participants, in their partial learning phase (i.e. giving answers inside the partial category). After the correct rules are found, and the partial rules deleted (as will be shown below), it will always give the correct answer at the first attempt.

3.1.3. *Learning Phase*

The learning phase is more complex, and consists of various steps. The model starts recording, for each rule in memory, if its predicate is true for the current item. Each rule can be good, bad, or neutral, for each category. It starts being neutral for all the categories, then, when an item is true for its predicate, the rule is marked as good for the item's category (unless it was previously marked as bad for that category). If instead the item is false for the predicate, the rule is marked as bad, with a probability proportional to a parameter. Upon the use of this parameter, it was decided because this kind of counter-factual reasoning is not common, and many participants probably would not use it. Finally, if the predicate is true for the current item, but the rule is final for another category, the rule is removed, since it would predict the wrong category as certain.

A second step of the learning phase is to check if there are final rules for some categories. If a final rule is found for a category, all the other partial rules are marked as bad for that category, so they would not produce.

The third step is discarding the useless rules, that is, the rules which are good for all or none of the categories, or the rules which are partial but too old (according to a parameter). The discarded rules are placed in the memory for the bad rules, the size of this memory is controlled by a parameter. For experiments 2 and 3, all of these recording steps are repeated also for some of the previously seen items, randomly chosen: the frequency of this "going back" is controlled by a parameter.

Finally, some new rules are created, with the methods previously explained (i.e. based on the shown item, randomly, modifying an existing rule, or intersecting two existing rules). If a rule is already in the memory for the bad rules, it is not created again.

As a last step, if the memories exceed their maximum sizes (controlled by two parameters), rules are randomly deleted from both memories (except for final rules in the memory for good rules) until their sizes are as required.

3.1.4. *Kinds of Predicates*

The model is open with respect to which predicates are actually used. A predicate is abstractly defined as something having two functions: "IsTrue" (referred to an item) and "ChangeRandom" (which returns a new predicate). Virtually all kinds of predicates can be implemented: they must simply provide these functions. In practice, only the kinds of predicates useful for the presented experiments are currently implemented, but the model can be extended with other predicates. It even can be [even] merged with some other model of discovery of relational structures, that can produce new predicates from scratch.

The predicates currently implemented are: 1. presence (or absence) of a feature or a set of features; 2. abstract number (a lot/a few) of elements with a given feature; 3. similarity between groups of elements (same/different colour, shape, number); 4. causal interaction (clicking on some element(s) causes reaction(s) of some kind); 5 intersection of two other predicates.

3.1.5. *Summary of the Parameters*

The parameters used by the model can be grouped into three categories: use of creation methods (one parameter for each method), available working memory, and use of counter-factual reasoning. By using a simple form of simulated annealing, it is possible to find the set of parameters that best reproduces a participant's performance. Starting from random sets of parameters, the space of the solutions is iteratively narrowed until a set of parameters is found which best reproduces the human behaviour.

In addition to these parameters, some values are computed concerning the effective use of some features. The actual use of these features probably changes with different tests, so it can be useful to know how much each feature was actually used. When the test ends, the final rules in memory are counted: how many were originally created from the shown items, how many randomly, and how many unifying two other rules. Also the number of times a rule was changed before becoming the final rule, was counted, as a measure of how much modification was used. Finally the average use of memory slots was computed.

3.2. *Model's Fit*

For the first experiment, the model was able to accurately reproduce the answers of 86% of the participants (significance of kappa < 0.05). For the second experiment, it was able to accurately reproduce the answers of 79% of the participants.

The sign tests of the learning intervals are significant ($p < 0.01$), and in the same direction as in the experiment. In other terms, also in the model's simulations, learning of one category helps learning of another category with a similar relational structure.

The results are the same as in the experiment, also for the sign tests of the given versus the expected number of partially incorrect answers during learning. They are better than expected ($p < 0.01$), meaning that also for the model the final learning of similar categories is preceded by a phase of partial learning.

As for the preference to learn similar criteria in comparison to learn dissimilar ones, the sign test of the clicks on analogical and non-analogical elements gives in the model the same result as in the experiments. The clicks on the analogical elements are significantly higher in number than on non-analogical elements ($p < 0.01$).

For what concerns the combinations of parameters found by the simulated–annealing algorithm, an analysis showed that the modification of rules is effectively used. Rules are changed, before becoming the final ones, an average of 1.82 times in the first experiment, and 3.68 times in the second and third experiments. Rules are created by using the shown items as the initial source almost all the times (99%). Instead the probability to find a good rule randomly is very low, as well as the probability of finding a good rule unifying two existing rules.

The variables that have the biggest impact on difficulty are the availability of memory ($r = -.152$; $p < 0.01$), the average use of modifications ($r = -.318$, $p < 0.01$), and the use of counter-factual reasoning ($r = -.346$, $p < 0.01$). These results confirm that people operate under memory constraints, and thus take advantage of the *modification heuristic* and of scientific reasoning.

3.3. *Summary*

The hypotheses on the use of analogical reasoning in category learning have been implemented in a model; it uses analogy between simultaneously learned categories, and it formulates partial rules for partial categories which are then refined. From these partial categories stem the similar final categories: this is an emergentist account of analogical reasoning.

With this new mechanism, this model can reproduce the results of the experiments, which were also predictions originating from the hypotheses: 1. learning of similar categories is mutually related; 2. final learning is preceded by a phase of partial learning; 3. people prefer similar criteria over easily discernible ones; 4. people operate under memory constraints.

4. **Conclusions**

This work describes a model of analogical reasoning that diverges from all the standard models of analogical reasoning produced so far (e.g. Gentner, 1981; Gentner, 1983; Holyoak & Koh, 1985). It is a more general model, based on the observation that normally, when learning things, people do not make a sharp distinction between known and unknown domains: they just do not know things. For that reason this study has focused on experiments where novel similar concepts are learned simultaneously.

For over twenty years the studies on analogical reasoning have focused on the analogy between completed structures, and they had to deal with high-complexity problems like mapping, concept alignment, and their computational complexity. On the contrary, it seems much more important to shift the attention toward the investigation of the role of analogy in conditions where different similar concepts are learned simultaneously. This is a more general phenomenon, which happens more frequently than the analogy between completed structures.

Constraints of time and memory have a big impact on analogical reasoning and generally on human learning. Differently from those previous studies, which avoided this problem, or dealt with it only marginally, this study attempts to make these constraints a central issue. Due to the limits imposed by restricted memory and time, the human mind resorts to analogy when novel concepts must be acquired. As already suggested by some earlier studies (e.g. Keane, 1995) and clearly demonstrated in this research: analogy is an extremely useful heuristic, that optimizes the amount of information that must be extracted and recorded, thus reducing the effort of memory and time required. It has implications for research on category-learning, especially where concepts have to be clearly distinguished (separated).

Given the salience of the discoveries presented in this work for the future development of cognitive science, some suggestions arise about what needs to be studied. Other interesting work could be done on the early use of analogy, in all the cases in which learning must take place in a limited amount of available time and resources. In those cases it is difficult to learn concepts separately and to use structure–mapping (as instead proposed by Kurtz & Loewenstein, 2007; Kurtz, Miao & Gentner, 2001). As far as category–learning is concerned, the theories which assume the need of a sharp separability of concepts need to be revised to account for the results of the present study.

Acknowledgments

This work was supported by the FP6 NEST Programme of the European Commission (ANALOGY: Humans – the Analogy-Making Species: STREP Contr. No 029088). Special thanks are due to Dr. Fintan Costello, who supervised this work, and to Dr. Isabella Proia, for her revision of a first version of this paper.

References

Ashby, F. G., & Maddox, W. T. (2005), Human Category Learning, *Ann. [u.] Rev. Psychol*, *56*, 149-78.

Del Re, G. (2000), Models and analogies in science, *HYLE–International Journal for Philosophy of Chemistry*, *6*(1), 5-15.

Doumas, L. A. A., & Hummel, J. E. (2005), A symbolic-connectionist model of relation discovery, In *Proceedings of the Twenty-Third Annual Conference of the Cognitive Science Society* (pp. 606-611).

Gentner, D. (1981), A structure-mapping theory of metaphor and analogy, *Bulletin of the psychonomic society*, *18*(2), 77.

Gentner, D. (1983), Structure-mapping: A theoretical framework for analogy, *Cognitive science*, *7*(2), 155-170.

Gentner, D., & Kurtz, K. J. (2005), Relational categories, *Categorization inside and outside the laboratory: Essays in honor of Douglas L. Medin, ed. WK Ahn, RL Goldstone, BC Love, AB Markman & P. Wolff*, 151–75.

Gentner, 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.

Hoffman, R. R. (1980), Metaphor in science, *Cognition and figurative language*, 393–423.

Holyoak, K. J., & Koh, K. (1985), The pragmatics of analogical transfer, *The psychology of learning and motivation: Advances in research and theory*, *19*, 59-87.

Keane, M. T. (1995), On order effects in analogical mapping: Predicting human error using IAM, In *Proceedings of the seventeenth annual conference of the Cognitive Science Society: July 22-25, 1995, University of Pittsburgh* (p. 449), Lawrence Erlbaum.

Kuehne, S. E., Forbus, K. D., & Gentner, D. (2000), SEQL: Category learning as progressive abstraction using structure mapping, *Proc. 22nd Ann. Conf. Cognitive Science Soc.*

Kurtz, K. J., & Loewenstein, J. (2007), Converging on a new role for analogy in problem solving and retrieval: When two problems are better than one, *Memory and Cognition*, *35*(2), 334.

Kurtz, K. J., Miao, C. H., & Gentner, D. (2001), Learning by analogical bootstrapping, *The Journal of the Learning Sciences*, *10*(4), 417-446.

Lavrac, N., & Dzeroski, S. (1994), Inductive logic programming, *Journal of Logic Programming*, *19*(20), 629-679.

Markman, A. B., & Wisniewski, E. J. (1997), Similar and different: The differentiation of basic-level categories, *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*(1), 54-70.

Michalski, R. S. (1983), A theory and methodology of inductive learning, *Machine Learning: an artificial intelligence approach*, *1*, 83–134.

Nosofsky, R. M., Palmeri, T. J., & McKinley, S. C. (1994), Rule-plus-exception model of classification learning, *Psychological Review*, *101*, 53-53.

Rosch, E. (1978), Principles of categorization, *Concepts: core readings*, 189-206.

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

Address author: UCD School of Computer Science and Informatics, Dublin, Ireland , E-mail: cesare@cesarebianchi.com

History of article: first published in Cruciani, M. (Ed). (2010), *Practices of Cognition: Recent Researches in Cognitive Sciences,* University of Trento, Italy, ISBN 978-88-8443-350-3