The Computational Models for Inductive Reasoning

Milos Ilic[1], Vladimir Stankovic[2]
[1] College of Agriculture and Food Technology
Cirila i Metodija 1, 18400 Prokuplje
Serbia
{milos.ilic.pk@gmail.com}

[2] Faculty of Electronic Engineering at the University of Nis
Aleksandra Medvedeva 14, 18000 Nis' Serbia
{vladimir.stankovic@elfak.ni.ac.rs}

ABSTRACT: *In this review, we have analysed and critically evaluated the inductive reasoning models. In the inductive reasoning, we draw conclusions which are not logically valid. In general, it is applied to make likely but not certain predictions about how people will behave in new environments. We in this paper, discussed and described the computational models and the basics of inductive reasoning. We have discussed how to implement and present how the models work, and give their positive and negative sides.*

Keywords: Inductive Reasoning, Similarity Effects, Typicality Effects, Bayesian Model, Computational Modeling

Received: 19 September 2020, Revised 12 January 2021, Accepted 18 February 2021

DOI: 10.6025/jio/2021/11/2/35-40

1. Introduction

The practical problem of induction goes from childhood and does not disappear with adolescence. Adults face it every day whenever they make any attempt to predict an uncertain outcome. Inductive inference is a fundamental part of everyday life, and for cognitive scientists, a fundamental phenomenon of human learning and reasoning in need of computational explanation. Inductive reasoning is potentially an extremely large topic, especially because it is often defined as reasoning about problems that do not involve perfectly certain conclusions [1]. The class of problems that have perfectly certain conclusions is much more circumscribed, for example, it could be defined in terms of a set of logical rules about what conclusions must follow from a given set of premises. In comparison, the set of problems for which inductive reasoning applies is potentially "everything else," and that is indeed a large and varied set.

The paper is organized as follows. In section 2 we will get to know the basic effects of data in human inductive reasoning. We will present the *Similarity Effects*, *Typicality Effects* and *Diversity Effects*. The computational models of human inductive reasoning, based on these effects, will be presented in section 3. Here we will compare some of the computational models, and give references for modifications of some of the basic models.

2. Basic Effects of Data in Human Inductive Reasoning

2.1. Similarity Effects

The idea that similarity should guide inductive reasoning has a distinguished history. Some scientists argued that "what happens once, will, under a sufficient degree of similarity of circumstances, happen again." For example [1] if you want to buy a CD for your friend, if you know that she likes 1960s albums by Rolling Stones and does not like Celine Dion, the most promising strategy is no doubt to buy her a CD by a similar 1960s band rather than by someone else who sings like Celine Dion. Some research has contributed to the study of induction by describing structural relations between similarity and induction in a detailed mathematical form .What is crucial about these studies is the assumption that inductive reasoning can be accounted for in terms of a single measure of similarity [2]. Although these studies were successful at modeling induction by using just one kind of similarity, they did not attempt to describe reasoning about more than one kind of property. In the next section we will present a mathematical model that successfully derives its predictions from similarity measures obtained from other subjects, again pointing to the role of overall similarity in inductive reasoning.

2.2. Typicality Effects

This phenomenon is closely tied to categorization research, in particular the idea that not all category members are equal, but instead some are more prototypical than others [1]. Returning to the buying a CD problem, if your friend likes albums by Rolling Stones, a prototypical 1960s guitar-based rock band, there would seem to be a lot of similar 1960s bands to choose from. On the other hand, if you know that she likes albums by Moody Blues, a much less typical 1960s band that recorded with a symphony orchestra, it would seem harder to choose another 1960s band that she would like – she might only like rock bands that use classical music. There was an additional effect of typicality beyond what might be predicted based only on similarity. Intuitively, if a typical mammal, such as a horse, has a disease, then perhaps all mammals have it, that is, the property applies to the super-ordinate category. On the other hand, if mice have a disease, it might be restricted to a subcategory of mammals, such as rodents. In sum, the typicality effect is another robust phenomenon that must be addressed by the models of inductive reasoning.

2.3. Diversity Effects

The diversity effect is somewhat more elusive than similarity or typicality, but it, too, has a distinguished history. The diversity effect is also well illustrated in the example of buying a CD. If your friend actually likes both Rolling Stones and Celine Dion, then you might infer that she has broad tastes in music, and it would be safe to buy her one of many styles of music. On the other hand, if you know she likes Rolling Stones and The Who, another guitar-based 1960s band, you might infer that her musical tastes are fairly narrow after all, and you should not stray too far from similar bands.

3. Computational Models

3.1. Similarity-coverage Model

The similarity-coverage model (SCM) presented in [3] is perhaps the best known mathematical model of property induction. It predicts the strength of inductive arguments as a linear combination of two factors, the similarity of the conclusion to the premises and the extent to which the premises "cover" the smallest super-ordinate taxonomic category including both the premises and the conclusion [4]. For single-premise arguments, coverage more or less reduces to typicality, but for multiple-premise arguments, coverage gives something closer to a measure of diversity. Coverage is most easily explained with examples:

$$\frac{\text{Mice have property X.}}{\text{All mammals have property X.}} \tag{1}$$

$$\frac{\text{Horses have property X.}}{\text{All mammals have property X.}} \tag{2}$$

Hippos have property X. (3)
Rhinos have property X.

All mammals have property X.

For arguments (1) and (2), the lowest level super-ordinate that includes all the categories is mammal. Coverage is assessed in terms of the average similarity of the premise category to the members of the super-ordinate. To the extent that horses are more typical mammals than mice and therefore more similar to other kinds of mammals, argument (2) will have greater coverage than argument (1). This is how the model addresses typicality effects. When assessing similarity between members of the super-ordinate category and the multiple premises, only the maximum similarity for any one premise category is considered. So for argument (3), very large mammals tend to be similar to both hippos and rhinos, and small mammals tend not to be similar to hippos and rhinos. So including rhinos as a premise category does not add much information beyond just having hippos as a premise category alone. The model in [3] can be written out more formally, as shown in Eq. 1:

$$Strength = \alpha SIM(p_1, ..., p_n; C) + (1 - \alpha) \tag{4}$$

$$xSIM(p_1, ..., p_n; [p_1, ..., p_n; C])$$

Here, α refers to the relative influence of the similarity component (ranging from 0 to 1) and (1-a) is the influence of the coverage component. This equation applies when there are n premise categories P and one conclusion category C. When the premise and conclusion categories are all at the same taxonomic level (e.g. robins, blue-jays; sparrows), then SIM returns the maximum of the pairwise similarities between each P_i and C. When the conclusion category is at a higher taxonomic level than the premise categories then SIM is applied recursively to known c that are members of C and averaged over these c. Generally speaking, this model addresses a wide variety of structural phenomena in inductive reasoning and is particularly impressive in how it puts together information from multiple premises, because of the powerful combination of similarity and coverage components. Although the model does incorporate some information about categories and similarity, it does not address background knowledge effects, such as the differential use of similarity and properties in [2], exceptions to diversity in [5], or, more generally, any use of causal knowledge or causal reasoning. Some new research, described in [16], uses this model.

3.2. Feature-based Model

The feature-based model in [6] computes inductive strength as a normalized measure of feature overlap between the conclusion and the example categories. The author in [6] presents a quantitative comparison with the SCM: the results are not conclusive, but suggest that the model does not predict human judgments as accurately as the SCM. The model, however, predicts some qualitative phenomena that the SCM can not explain. More recently, authors in [7] have presented a feature-based approach to semantic cognition that uses a feed forward connectionist network with two hidden layers. This connectionist approach is more ambitious than any of the others we describe, and the authors apply their model to a diverse set of semantic phenomena. One of the applications is a property induction task where the model makes sensible qualitative predictions, but there has been no demonstration so far that the model provides good quantitative fits to human judgments. From our perspective, both feature-based models share the limitations of the SCM. Despite the range of applications in [7], it is not clear how either model can be extended to handle causal settings or other inductive contexts that draw on sophisticated domain knowledge. The models also include components that have been given no convincing justification. The model in [6] uses mathematical measure of feature overlap, but it is not clear why this should be the right measure to use. The authors in [7] provide no principled explanation for the architecture of their network or their strategy for computing the strength of inductive arguments, and their model appears to rely on several free parameters. The model relies on similarity effects since training and testing using similar input vectors will lead to strong outputs during testing. Input vectors and outputs are connected with an activation function witch is described in [1] as follows:

$$a(C \mid p_i, ..., p_n) = \frac{W(p_i, ..., p_n) \bullet C}{|C|^2} \tag{5}$$

In this function n is a set of given premise categories p with a conclusion category C. W represents a vector corresponding to the already-trained weights in the network after the premise categories have been learned. C is a vector corresponding to the future representation of the conclusion category. The dot product between W and C is computed, yielding a value corresponding to the similarity between the premise categories and the conclusion category. For example [1] donkeys and mules would have many features in common, and there would be a fairly high, positive dot product between the two vectors.

On the other hand, donkeys and ostriches would have fewer features in common and would give a lower dot product, perhaps close to zero. Finally, the activation is scaled in the denominator, by the squared length of the vector *C*, essentially a measure of the number of known features of *C*. If *C* corresponds to a well-known category, such as dogs, it will be relatively difficult to draw a new conclusion. If *C* corresponds to a poorly known category, such as ocelots, it will be easier to draw new conclusions about the category. The model described in [6], like the model described in [3], can account for many structural phenomena in inductive reasoning, but it does not address background knowledge effects and does not use knowledge about properties, which guides towards the use of similarity or information about causality.

3.3. Bayesian model

According to the Bayesian model described in [1,8], evaluating an inductive argument is conceived of as learning about a property, in particular, learning for which categories the property is true or false. The Bayesian model treats the premise or premises in an inductive argument as evidence, which is used to revise beliefs about the prior hypotheses according to Bayes' theorem. Once these beliefs have been revised, then the plausibility of the conclusion is estimated. The Bayes' theorem is shown in Eq. 3.

$$P(H_i \mid D) = \frac{P(H_i)P(D \mid H_i)}{\sum_{j=1}^{n} P(H_j)P(D \mid H_j)} \tag{6}$$

In applying Bayes' theorem in Eq. 3, the premise is treated as the data, *D*. The prior degree of belief in each hypothesis is indicated by $P(H_i)$. The task is to estimate $P(H_i \mid D)$, that is, the posterior degree of belief in each hypothesis given the data. The Bayesian model addresses many of the key phenomena in inductive reasoning. For example, the model predicts the similarity effect because novel properties would be assumed to follow the same distributions as familiar properties. The Bayesian model also addresses typicality effects under the assumption that according to prior beliefs, atypical categories would have a number of idiosyncratic features. In comparison, prior beliefs about typical categories would indicate that they have many features in common with other categories. The Bayesian model, unlike the previous, also addresses diversity effects with a rationale similar to that for typicality effects. This is good because if we take for example an argument with two similar premise categories, such as hippos and rhinos, this could bring a lot of idiosyncratic properties that are true just for large mammals. In a same way a novel property of hippos and rhinos might seem idiosyncratic as well. In contrast, an argument with two diverse premise categories, such as hippos and hamsters, could not bring to mind familiar idiosyncratic properties that are true of just these two animals. Instead, the prior hypotheses would be derived from known properties that are true for all mammals or all animals. In other way some of the authors in [8] showed that the Bayesian model addresses about the same range of structural phenomena in inductive reasoning as the similarity-coverage model and the feature-based model. A modification of the Bayesian model is given in [4]. The framework in [4] adopts a Bayesian approach similar to [8], but emphasizes the importance of modeling the form and the origins of appropriate priors. This framework includes two components: a Bayesian engine for inductive inference, and a language for specifying relevant aspects of domain theories and using those theories to generate prior probability distributions for the Bayesian inference engine. The Bayesian engine reflects domain-general norms of rational statistical inference and remains the same regardless of the inductive context. Different domain theories may be appropriate in different inductive contexts, but they can often be formalized as instances of a single unifying scheme: a probabilistic process, such as diffusion, drift or transmission, defined over a structured representation of the relevant relations between categories, such as taxonomic or ecological relations. More about this framework can be found in [4].

3.4. Models based on Support Vector Machine

These models deal with one kind of inductive reasoning argument such as: The person likes wine. The person doesn't like beer.

The person likes wine.
The person doesn't like beer.

The person likes champagne.

In this type of argument, its strength depends mainly on the entities in each sentence since these sentences share the same basic predicate. The study in [9] examines the impact of risk contexts on inductive reasoning. The previous models discussed the context-dependency of inductive reasoning argument and they have only addressed the issue with identical entity sets and by changing the predicate. They claim that the information required for similarity computation differs for different predicates, that is, different semantic contexts. This model however, reports that identical arguments are rated differently in different situational contexts. Findings from this model indicate that people modify the same similarity information necessary to rate argument strengths according to the given situational context, which results in different ratings. Inductive reasoning in risk contexts is best explained by a category-based model based on a Support Vector Machine (SVM) which adjusts the similarities for positive premise entities, negative premise entities, and conclusion entities. The processes of inductive reasoning addressed in this study are assumed to involve a kind of similarity-based temporal categorization that utilizes stable semantic knowledge. For example, the temporal category that “Mr. A likes” can be formed from positive and negative premise entities (e. g. “Mr. A likes steak”, “Mr. A doesn’t like Japanese noodle” → “steak” and “Japanese noodle”) and applied in making estimations about the likelihood of the conclusions (e. g., “Mr. A likes pork → highly likely) based on the similarity of “pork” with “steak” and the dissimilarity with “Japanese noodle”. This model is based on three assumptions [10]: internal representation assumptions, retrieval assumptions, and response selection assumptions. The internal representation assumption explains the way in which the stimuli and the contrasting categories are represented. The entities of premises and conclusions are assumed to be prototypes in a knowledge space. The retrieval assumption provides a description of the information that must be collected before a response can be made. In this model the similarities between the premise and conclusion entities are described by a nonlinear function of simple Euclidean distances. Kernel functions are assumed to be the nonlinear similarity functions between the premise and the conclusion entities. Thus, these proposed models show that people can temporally discriminate natural language concepts within a complex semantic structure according to various combinations of positive and negative premise entities. More about this can be found in [11]. The response selection assumption provides a description about how people select a response after all the relevant information has been collected. In this model, participants’ responses are assumed to be influenced by the desire to optimize response utility, that is, to choose a response that might not lead to score decreases. Since score decreases might cause the low evaluation of the participants’ ability, people try to avoid such a score decreasing risk by adjusting the relevant information collected for the task response. The response decision is assumed to be based on similarity estimations which are themselves biased by “situational” contexts that lead to the participant’s risk aversion strategies. Two kinds of models based on SVM are proposed in [9,12]. The first model processes feature-based representations, while the other processes category-based representations. The likelihood of a conclusion including entity is represented by the following discrimination function constructed from an SVM, based on Gaussian kernel functions:

$$v\left(N_i^c\right) = aSIM_+(N_i^c) + bSIM_-(N_i^c) \tag{7}$$

$$SIM_+(N_i^c) = \sum_j^{n^+} e^{-\beta dj^+} \tag{8}$$

$$SIM_-(N_i^c) = \sum_j^{n^-} e^{-\beta dj^-} \tag{9}$$

Depending on which model is used, dj^+ and dj^- have different representations. In feature-based version of models these parameters are functions for word distance based on the feature words [13, 14]. These words are denoted with A_k. In the category-based version of models these are word-distance functions based on the latent classes (C_k). In these models the mechanism underlying the risk context effects in inductive reasoning is explained by similarity adjustment based on risk aversion strategies toward social evaluation context. SVM was also used in [15], where we can find an extended use of the induction models based on SVM.

4. Conclusion

In this paper we gave a review of the basic models of inductive reasoning. The effects of data in human inductive reasoning, presented in the paper, present a key element of all the models of inductive reasoning. These effects connect the psychology of human reasoning and the computational models for inductive reasoning. All models in this paper are based on some effect or on multiple effects, and that was the reason we included exactly those models. Some of the models present a basis for more complex research in this field, and may be used like a good start for gaining new, upgraded models. In that sense a possible continuation of this work might be a computer implementation of some of the considered models of inductive reasoning.

References

[1] Sun, R. (2008). The Cambridge Handbook of Computational Psychology, Cambridge University Press, New York.

[2] Weber, M., Osherson, D. (2014). Category-based induction from similarity of neural activation, Cognitive, *Affective, & Behavioral Neuroscience*, 14 (1) 24-36.

[3] Holyoak, K., Seung Lee, H., Lu, H. (2010). Analogical and Category-Based Inference: A Theoretical Integration With Bayesian Causal Models, *Journal of Experimental Psychology*: General, 139 (4) 702–727.

[4] Tenenbaum, J., Kemp, C., Shafto, P. Theory-based Bayesian models of inductive reasoning, *Massachusetts Institute of Technology*.

[5] Medin, L., Coley, D., Storms, G., Hayes, K. (2003). A relevance theory of induction, *Psychonomic Bulletin & Review*, 10, 517–532.

[6] Griffiths, O., Hayes, B., Newell, B. (2012). Feature-based versus category-based induction with uncertain Categories, *Journal of Experimental Psychology*, 38 (3), 576.

[7] Rogers, T., McClelland, L. (2004). Semantic cognition: A parallel distributed processing approach, Cambridge, MA:MIT Press.

[8] Tenenbaum, J., Griffiths, T., Kemp, C. (2006). Theory-based Bayesian models of inductive learning and reasoning, *TRENDS in Cognitive Sciences*, 10 (7) 309-318.

[9] Sakamoto, K., Nakagawa, M. (2007). Risk Context Effects in Inductive Reasoning: An Experimental and Computational Modeling Study, *Tokyo Institute of Technology*, 2-21-1 Ookayama, Meguroku, Tokyo, Japan, p. 425-438.

[10] Ashby, G., Maddox, T. (1993). Relations between prototype, exemplar, and decision bound models of categorization, *Journal of Mathematical Psychology*, 37, 372–400.

[11] Sakamoto, K., Nakagawa, M. (2006). The Effects of negative premise on Inductive reasoning: A psychological experiment and computational modeling study, *In*: Proceedings of the Twenty-Eighth Annual Conference of the Cognitive Science Society, p. 2081–2086.

[12] Hayes, B., Fritz, K., Heit, E. (2013). The relationship between memory and inductive reasoning: Does it develop?, *Developmental Psychology*, 49 (5) 848-860.

[13] Ifenthaler, D., Seel, N. (2013). *Model-based reasoning, Computers & Education an International Journal*, Volume 64, 131–142, (May).

[14] Hayes, B., Heit, E. (2013). How similar are recognition memory and inductive reasoning?, Memory & Cognition, Volume 41 (5), p 781-795, 2013.

[15] Bharadwaj, A., Minz, S. (2014). Inductive-Analytical Learning based Stepwise Support Vector Machine (SVM) Model, *International Journal of Soft Computing and Engineering (IJSCE)*, 4, Issue-ICCIN-2K14, 2014.

[16] Heussen, D., Voorspoels, W., Storms, G. (2010). Can similarity-based models of induction handle negative evidence?, *Cognitive Science Society*, p. 2033—2038.