Cognition 157 (2016) 296–320

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Original Articles

Probabilistic representation in syllogistic reasoning: A theory to integrate mental models and heuristics

Masasi Hattori *

College of Comprehensive Psychology, Ritsumeikan University, Japan

ARTICLE INFO

Article history:
Revised 2 September 2016
Accepted 8 September 2016

Keywords:
Deduction
Probabilistic inference
Syllogistic reasoning
Mental representation
Information gain
Symmetry inference

ABSTRACT

This paper presents a new theory of syllogistic reasoning. The proposed model assumes there are probabilistic representations of given signature situations. Instead of conducting an exhaustive search, the model constructs an individual-based "logical" mental representation that expresses the most probable state of affairs, and derives a necessary conclusion that is not inconsistent with the model using heuristics based on informativeness. The model is a unification of previous influential models. Its descriptive validity has been evaluated against existing empirical data and two new experiments, and by qualitative analyses based on previous empirical findings, all of which supported the theory. The model's behavior is also consistent with findings in other areas, including working memory capacity. The results indicate that people assume the probabilities of all target events mentioned in a syllogism to be almost equal, which suggests links between syllogistic reasoning and other areas of cognition.

> All fools are poets; this the Prefect feels; and he is merely guilty of a non distributio medii in thence inferring that all poets are fools.
>
> [— Edgar Allan Poe, *The Purloined Letter* (1845)]

1. Introduction

Reasoning is intended to derive reasonable conclusions from premises. Given the assertions that "The Kyotoite are Japanese" and "The Japanese are Asian," it is reasonable to conclude that "The Kyotoite are Asian." In this case, the relation is *transitive*: if K → J and J → A, then K → A. However, if one knows that "The Kyotoite are suave," it is illogical to infer that "Suave people are Kyotoite." A relation is *symmetric* if $X \rightarrow Y$ implies $Y \rightarrow X$, but such symmetrical derivations are not licensed in logic. As such, some inferences are logically valid, and others are invalid; some are easy, and others are difficult. The difficulty of inference depends, at least partly, on its logical form, but an error-prone argument can sometimes be obvious with a slight change in wording (e.g., using familiar terms). At the same time, difficulty of inference must relate to other types of thinking, because if nothing else, reasoning must be carried out in working memory. Any comprehensive psychological theory of reasoning must address these issues, that is, why some inferences are difficult and how this relates to other areas of cognition. The current paper proposes one such attempt, along with the novel idea of *probabilistic representation*. Before going into detail, however, I first motivate two issues regarding cognitive architecture and inferential structure: mental representations and symmetry, which will feature strongly in what follows.

The current theory (*probabilistic representation theory* hereafter) proposes dual mental representations: probabilistic representations and individual-based mental models.[1] This is based on the hypothesis that people have several thinking modes. We sometimes take a *summary view* with probabilistic representations when, for example, we are seeking some rules or tendencies that are useful for predictive inference. In this heuristic mode, we think and talk about probable relations between classes of events or objects (e.g., "I know the Kyotoite are suave [by and large]."). In contrast, when critically testing a hypothesis or thinking counterfactually, we take a *distinctive view* with individual-based representations. In this analytic mode, we talk about stricter (i.e., more logical) rules, sometimes focusing on exceptions (e.g., "One of my acquaintances is Kyotoite, but he is not suave; so, I don't think this is true."). In this way, we can easily switch views according to factors such as context, situation, motivation, and purpose. These two distinctive views, I

---

* Address: College of Comprehensive Psychology, Ritsumeikan University, 2-150 Iwakura-cho, Ibaraki, Osaka 567-8570, Japan.
E-mail address: hat@lt.ritsumei.ac.jp

[1] This general word does not necessarily indicate the same thing that is meant by the mental model theory introduced later.

http://dx.doi.org/10.1016/j.cognition.2016.09.009
0010-0277/© 2016 The Author(s). Published by Elsevier B.V.

assume, depend on different representations: *continuous* (i.e., probabilistic) and *discrete* (i.e., individual-based).

Probabilistic representation theory supposes the summary view precedes the distinctive, because the summary view is based on heuristic processes but the distinctive view is based on deliberate processes. People first have a probabilistic intuition, and next construct mental models based on that intuition that serves for logical tests. As a result, the distinctive view is affected by the summary view, in that, people's probabilistic intuitions restrict how they test logical relations. In modeling this mechanism, a set of discrete mental models is a summary representation of a primitive continuous probabilistic model, and not the other way around. This implementation is unique among probabilistic approaches that have been proposed for deductive reasoning. One previous model (Chater & Oaksford, 1999) did not propose any internal representations, and the others (Guyote & Sternberg, 1981; Johnson-Laird, Legrenzi, Girotto, Legrenzi, & Caverni, 1999; Khemlani, Lotstein, Trafton, & Johnson-Laird, 2015) assume the priority of discrete models, introducing probabilistic behavior by allocating numerals (i.e., probability values) to discrete models.

This aspect of the theory is an extension of recent approaches to reasoning based on probability (e.g., Chater & Oaksford, 1999; Evans & Over, 2004; Oaksford & Chater, 2007) called the *new paradigm* in the psychology of reasoning (Elqayam & Over, 2013; Over, 2009). Although logic guides deductive reasoning, the idea that deduction depends on logic as a normative theory of human reasoning is now an "ancient proposal" (Johnson-Laird, Khemlani, & Goodwin, 2015, p. 201). After the 1990s, many researchers moved to probabilistic approaches to reasoning. In these approaches it is usually presupposed that degrees of certainty or belief correspond to subjective probabilities, and the validity of an argument is assessed via the probabilistic validity, or *p-validity*, proposed by Adams (1975): the uncertainty (i.e., the complement of the probability) of a *p*-valid conclusion does not exceed the sum of the uncertainties of the premises. This presupposition implicitly requires each proposition to retain its probability (at any time in any context, in principle) to enable probabilistic inference as follows:

The Kyotoite are suave.	(prob = 0.85)
The suave are ...	(prob = 0.43)
...	(prob = 0.05)
∴ The Kyotoite are ...	(prob = ...)

This actually places an excessive load on the working memory, especially when forming a chain of inferences, because an extra piece of information about probability must be retained for each statement. Moreover, even a couple of premises can result in innumerable (*p*-)valid (but vapid) conclusions (see, Johnson-Laird et al., 2015). Thus, a model based on a system of *p*-validity (as well as a standard binary logic) can generate serious concerns at the algorithmic level about the feasibility of a model implementing (deductive) reasoning. It seems reasonable to suppose that people discretize (i.e., simplify) their degrees of belief in each proposition at some point in time. For example, a statement with a probability of 95% or higher may be regarded as just a "true" statement somewhere in the course of the reasoning process. In the probabilistic representation model, this is done by constructing a discrete model (i.e., by generating a small number of samples) in accordance with a given probability distribution contained in the probabilistic representation.

The current theory also proposes that *symmetry inferences* are central to syllogistic reasoning performance. The symmetry inference is prevalent not only in syllogisms, but also in other areas. For example, a conditional "If *X* then *Y*" is often interpreted as if it also means that "If *not-X* then *not-Y*" or "If *Y* then *X*" at the same time (e.g., Geis & Zwicky, 1971; Staudenmayer, 1975). A logic-based account for this inference is that the conditional "$X \to Y$" is prone to be interpreted as a *biconditional* "$X \leftrightarrow Y$" (e.g., Johnson-Laird & Byrne, 1991). Similarly, if one is told that the probability of a woman who has breast cancer receiving a positive mammography is 80%, then one is apt to infer that the probability that a woman who tested positive actually has breast cancer is also about 80%, even if the answer clearly violates the Bayesian norm (Eddy, 1982; Gigerenzer & Hoffrage, 1995; Tversky & Kahneman, 1980). Many researchers have attributed this type of error to the *inverse fallacy*, a tendency to confuse a given conditional probability *P* (*symptom* | *disease*) with the inverse conditional probability, *P*(*disease* | *symptom*), that is to be judged (Braine, Connell, Freitag, & O'Brien, 1990; Gavanski & Hui, 1992; Hammerton, 1973; Koehler, 1996; Macchi, 1995; Villejoubert & Mandel, 2002; Wolfe, 1995). These are all examples of the symmetry inference.

One of the reasons why symmetry inference is important for a comprehensive theory of thinking is that this mode of inference has been argued to be distinctively human. Nonhuman animals such as chimps (Dugdale & Lowe, 1990, 2000), find symmetry inferences extremely difficult (e.g., D'Amato, Salmon, Loukas, & Tomie, 1985; Sidman et al., 1982). Many researchers have pointed to the relevance of symmetry to language processing (e.g., Dugdale & Lowe, 1990; Horne & Lowe, 1996; Oaksford, 2008) or to creativity (Hattori, 2008). The fundamental ability to perform symmetry inferences may be constrained by phylogenetic factors, and is closely related to other areas of cognition such as language and creativity that are only found in humans. Thus, appearance of symmetry inferences in syllogistic reasoning may be a reflection of our common cognitive architecture.

A theory with probabilistic representations may afford an insight into the nature of symmetry inferences. Hattori and Nishida (2009) hypothesized that people tend to regard two target classes of objects or events as almost equal in size (see Fig. 1). For example, when we think of a disease (e.g., breast cancer) and its symptoms (e.g., a positive mammography), we assume that the sizes of two target sets, one for the disease and the other for symptoms, are roughly the same. This default assumption results in the inverse fallacy. It is reasonable to assume that the target events have a similar probability, unless we know this is not the case (e.g., showing many false positives for a rare disease), because we then gain some information about the credibility of the test. Thus equating the sizes of two target sets (i.e., a *set-size balancing principle*[2]) is a reasonable model of ignorance and the simplest assumption. This principle is known to be maintained in other areas of human thinking, including causal induction (Hattori & Oaksford, 2007) and reasoning in the Wason selection task (Hattori, 2002). Therefore the current theory can reveal an important new link between deductive reasoning and other areas of thinking.

I now briefly introduce the syllogistic reasoning task and some of the terminology required to understand the literature and the current theory before turning to review of previous studies. I follow the orthodox Aristotelian classification in this paper, although there are several different forms of notation used in the psychological literature (see also Appendix A). Syllogisms are constructed with two premises and one conclusion. Each statement is one of four forms called *moods*. Traditionally, these are labeled A, I, E, and O:

---

[2] It is called the *equiprobability assumption* in Hattori and Oaksford (2007) and Hattori and Nishida (2009). However, some researchers use the same term with the different meaning that each individual possibility has the same probability (e.g., Johnson-Laird et al., 1999; Lecoutre, 1992). To avoid confusion, I adopt a different name here.