Cognition 127 (2013) 159–176

Contents lists available at SciVerse ScienceDirect

Cognition

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

Bayesian learning and the psychology of rule induction

Ansgar D. Endress *

Universitat Pompeu Fabra, Center of Brain and Cognition, C. Roc Boronat, 138, Edifici Tanger, 55.106, 08018 Barcelona, Spain
City University London, Department of Psychology, London EC1V 0HB, United Kingdom
Massachusetts Institute of Technology, Department of Brain & Cognitive Sciences, 43 Vassar St, Cambridge, MA 02139, United States

ARTICLE INFO

Article history:
Received 25 October 2011
Revised 29 November 2012
Accepted 30 November 2012
Available online 1 March 2013

Keywords:
Bayesian learning
Process models
Rule learning
Perceptual or memory primitives

ABSTRACT

In recent years, Bayesian learning models have been applied to an increasing variety of domains. While such models have been criticized on theoretical grounds, the underlying assumptions and predictions are rarely made concrete and tested experimentally. Here, I use Frank and Tenenbaum's (2011) Bayesian model of rule-learning as a case study to spell out the underlying assumptions, and to confront them with the empirical results Frank and Tenenbaum (2011) propose to simulate, as well as with novel experiments. While rule-learning is arguably well suited to rational Bayesian approaches, I show that their models are neither psychologically plausible nor ideal observer models. Further, I show that their central assumption is unfounded: humans do not always preferentially learn more specific rules, but, at least in some situations, those rules that happen to be more salient. Even when granting the unsupported assumptions, I show that all of the experiments modeled by Frank and Tenenbaum (2011) either contradict their models, or have a large number of more plausible interpretations. I provide an alternative account of the experimental data based on simple psychological mechanisms, and show that this account both describes the data better, and is easier to falsify. I conclude that, despite the recent surge in Bayesian models of cognitive phenomena, psychological phenomena are best understood by developing and testing psychological theories rather than models that can be fit to virtually any data.

To recognize the taste of an apple, do we automatically think about the tastes of oranges as well as all other foods before we can know that we are eating an apple? According to a growing literature of Bayesian models, we make inferences (e.g., the kind of food we are tasting) by considering all possible situations (e.g., tasting apples, oranges, etc.) in addition to the situation we actually face, and then decide which of these situations is the most likely one. Bayesian inference models have been claimed to account for an impressive variety of cognitive phenomena, including visual grouping (Orbán, Fiser, Aslin, & Lengyel, 2008), action understanding (Baker, Saxe, & Tenenbaum, 2009), concept learning and categorization (Anderson, 1991; Goodman, Tenenbaum, Feldman, & Griffiths, 2008), (inductive) reasoning (Goodman, Ullman, & Tenenbaum, 2011; Griffiths & Tenenbaum, 2009; Kemp, Perfors, & Tenenbaum, 2007; Kemp & Tenenbaum, 2009; Kemp, Tenenbaum, Niyogi, & Griffiths, 2010; Lu, Yuille, Liljeholm, Cheng, & Holyoak, 2008; Oaksford & Chater, 1994; TTglás et al., 2011), judgment about real-world quantities (Griffiths & Tenenbaum, 2006), word learning (Frank, Goodman, & Tenenbaum, 2009; Xu & Tenenbaum, 2007), word segmentation (Frank, Goldwater, Griffiths, & Tenenbaum, 2010), and grammar acquisition (Perfors, Tenenbaum, & Wonnacott, 2010; Perfors, Tenenbaum, & Regier, 2011).

Despite this growing literature, various authors have criticized Bayesian approaches on theoretical grounds (Altmann, 2010; Bowers & Davis, 2012; Fitelson, 1999; Jones & Love, 2011; Marcus, 2010; Sakamoto, Jones, & Love,

* Address. Universitat Pompeu Fabra, Center of Brain and Cognition, C. Roc Boronat, 138, Edifici Tanger, 55.106, 08018 Barcelona, Spain. Tel.: +34 93 5421382.
E-mail address: ansgar.endress@m4x.org

0010-0277/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.cognition.2012.11.014

2008), and where Bayesian approaches have been explicitly compared to psychological models (e.g., in the case of causal inference), the non-Bayesian approaches typically explained the data better (e.g., Bes, Sloman, Lucas, & Raufaste, 2012; Fernbach & Sloman, 2009). Here, I add to this literature by taking a model in a domain that appears particularly suitable for Bayesian learning—rule induction, spell out its underlying assumptions as well as their predictions, and confront them with empirical data. Specifically, Frank and Tenenbaum (2011) recently proposed that infants acquire rules in Bayesian, optimal ways. I will compare this approach with an account of rule-learning based on simple, psychologically grounded mechanisms, and show that the latter approach provides a principled explanation for the data.

1. Bayesian approaches to cognition: what is optimal?

On a conceptual level, Bayesian inference is straightforward. For example, if we encounter an individual with a Red Sox cap, we conclude that she is more likely to come from Boston than from, say, New York. However, to draw this conclusion, we use our knowledge that the likelihood of somebody wearing a Red Sox cap is higher in Boston than in New York. Bayesian calculations allow us to turn the likelihood that somebody who is in Boston wears a Red Sox cap into the likelihood that somebody who wears a Red Sox cap is from Boston. Moreover, such calculations make "optimal" use of the available information.

Despite its conceptual simplicity, Bayesian inference is tremendously useful in domains from statistics (e.g., Gill, 2008; O'Hagan, 1994) to evolutionary biology (e.g., Huelsenbeck, Ronquist, Nielsen, & Bollback, 2001; Pagel, 1994). Further, natural selection can be formulated as a Bayesian optimization problem; as a result, Bayesian inference has given us important insights into the evolution of our mental abilities. For example, some researchers have shown that perceptual and cognitive mechanisms might be well adapted to the statistics of our natural environment (e.g., Brunswik & Kamiya, 1953; Elder & Goldberg, 2002; Geisler & Diehl, 2002, 2003; Sigman, Cecchi, Gilbert, & Magnasco, 2001; Weiss, Simoncelli, & Adelson, 2002).

However, when it comes to Bayesian models of learning and cognition, environmental statistics are generally lacking, forcing such models to be much more speculative and hard to verify. This problem follows directly from Bayesian claims to make "optimal" use of information in the environment, and our lack of understanding of what has been optimized over the course and under the constraints of evolution. In fact, not all behavioral traits are optimal, but some might simply be accidents of how a species has evolved. For example, in some monogamous animals such as Zebra finches, females seek extrapair copulations although this behavior is maladaptive for females. However, extrapair mating behavior might be selected for in females because it might be affected by an allele that is shared with males, for whom siring extrapair offspring is adaptive (Forstmeier, Martin, Bolund, Schielzeth, & Kempenaers, 2011). Hence, the seemingly maladaptive behavior might be due to the accidents of how this trait is encoded genetically, suggesting that it is extremely difficult to assess whether our cognitive mechanisms are optimal and, if so, what they have been optimized for.

2. An overview over Frank and Tenenbaum's (2011) models

Frank and Tenenbaum's (2011) model is representative of a large number of similar models, and is applied to a domain that is arguably well-suited to Bayesian approaches. (Frank & Tenenbaum (2011) present in fact three different models, but I will present the differences between these models as they become relevant for the current purposes.) They raise the question of how young infants learn rule-like patterns based on repetitions. For example, syllable triplets like *ba–li–li* follow an *ABB* pattern, where the last syllable is repeated; syllable triplets like *ba–ba–li* follow an *AAB* pattern, where the first syllable is repeated. Following Marcus, Vijayan, Rao, and Vishton's (1999) seminal demonstration that young infants can learn such patterns, repetition-patterns have become an important testing ground for rule-learning, both in humans (e.g., Dawson & Gerken, 2009; Endress, Dehaene-Lambertz, & Mehler, 2007; Endress, Scholl, & Mehler, 2005; Frank, Slemmer, Marcus, & Johnson, 2009; Gerken, 2010; Gómez & Gerken, 1999; Kovács & Mehler, 2008, 2009a; Marcus, Fernandes, & Johnson, 2007; Saffran, Pollak, Seibel, & Shkolnik, 2007) and in nonhuman animals (e.g., Giurfa, Zhang, Jenett, Menzel, & Srinivasan, 2001; Hauser & Glynn, 2009; Murphy, Mondragon, & Murphy, 2008).

According to Frank and Tenenbaum's (2011) model, infants try to figure out the "best" rule describing the stimuli they perceive. To do so, they come equipped with an innate inventory of elementary rules, and check whether what they hear (or see) is compatible with *all* of the rules in their inventory. For example, if they hear *AAB* triplets, they would not only think about *AAB* patterns, but also about *ABB* patterns and all other patterns Frank and Tenenbaum (2011) incorporated into their model, even if they never hear any of these alternative patterns. To choose a rule, Frank and Tenenbaum (2011) propose that infants assume that the probability that a stimulus has been generated by a rule is inversely proportional to the total number of stimuli that can be generated by the rule (Eqs. (2) and (3) in their first model; the other models make similar assumptions); this strategy has been called the *size principle* by Tenenbaum and Griffiths (2001).

Concretely, infants might encounter the triplets *pu–li–li* and *ba–pu–pu*, both following an *ABB* pattern. Hence, they encounter a total vocabulary of three syllables (i.e., *pu*, *li* and *ba*). According to Frank and Tenenbaum (2011), infants know (i) that the three syllables allow for a total of $3 \times 3 \times 3 = 27$ triplets; (ii) that 6 of these triplets follow an *ABB* pattern; and (iii) that 3 of these triplets follow an *AAA* pattern (where all three syllables are identical), even though infants have never heard any *AAA* triplets; infants know the number of triplets that are compatible with any other conceivable rule.

As a result, irrespective of any Bayesian computations, infants know that *AAA* patterns are a priori more unlikely