Brain and Cognition 112 (2017) 84–91

Contents lists available at ScienceDirect

Brain and Cognition

journal homepage: www.elsevier.com/locate/b&c

To be precise, the details don't matter: On predictive processing, precision, and level of detail of predictions ☆

Johan Kwisthout *, Harold Bekkering, Iris van Rooij

Radboud University, Donders Institute for Brain, Cognition and Behavior, Nijmegen, The Netherlands

ARTICLE INFO

Article history:
Received 20 September 2015
Revised 15 February 2016
Accepted 21 February 2016
Available online 22 April 2016

Keywords:
Predictive processing
Precision
Level of detail
Structured representations
Formal modeling
Causal Bayesian networks

ABSTRACT

Many theoretical and empirical contributions to the Predictive Processing account emphasize the important role of precision modulation of prediction errors. Recently it has been proposed that the causal models used in human predictive processing are best formally modeled by categorical probability distributions. Crucially, such distributions assume a well-defined, discrete state space. In this paper we explore the consequences of this formalization. In particular we argue that the level of detail of generative models and predictions modulates prediction error. We show that both increasing the level of detail of the generative models and decreasing the level of detail of the predictions can be suitable mechanisms for lowering prediction errors. Both increase precision, yet come at the price of lowering the amount of information that can be gained by correct predictions. Our theoretical result establishes a key open empirical question to address: How does the brain optimize the trade-off between high precision and information gain when making its predictions?

1. Introduction

The predictive processing account has received widespread attention. Building on pioneering hierarchical predictive coding models of perception (Friston, 2005; Lee & Mumford, 2003; Rao & Ballard, 1999), recent literature proposes that the whole of perception, cognition, and action (Clark, 2013a) or even the entire operation of the brain (Hohwy, 2013) can be summarized by a simple, unifying principle. Rather than processing inputs in a mere bottom-up fashion, the brain is assumed to predict its inputs in a hierarchical manner by generative (causal) models and to process only that part of the input that is yet unexplained – the so-called prediction error. Sometimes prediction errors stem from the inherent stochastic nature of the world. To illustrate, take for instance, the observation of the outcome of a coin toss. We will have high confidence in our prediction that the coin will either land on heads or that it will land on tails, each event having a probability of 0.5; the observation of the actual outcome – while generating one bit of information – will normally not surprise us, as both events are fully consistent with our experience and knowledge of tossing a (fair) coin. One's generative models will therefore presumably not be changed as a consequence of this prediction error.

However, sometimes prediction errors are the result of an incomplete, immature, or just plain wrong generative model; think of trying an unknown dish in a restaurant or standing on skates for the first time. The uncertainty here is due to a lack of knowledge, and the prediction error *will* have impact: It allows the brain to update and improve its generative models (Payzan-LeNestour & Bossaerts, 2011; Yu & Dayan, 2005). These different roles of prediction errors, depending on the source of the uncertainty (*irreducible*, i.e., due to the inherent (known) stochastic nature of the world; or *reducible*, i.e., due to our lack of knowledge) are captured by the *precision* of the prediction error: A context-specific weighting of the prediction error that drives less or more attention to prediction errors. The net effect of the observation is thus a function of the precision of the prediction (capturing the uncertainty of the outcome) and the precision of the prediction error (capturing the model confidence).

Traditionally, computational operationalizations of the predictive processing account formulate the generative models (i.e., the stochastic relation between hypothesized causes and the predicted effects thereof) as Gaussian densities. Recently, however, Friston et al. (2015) propose to use *categorical* (discrete) probability distributions to describe the stochastic generative models that give rise to the predictions. An important distinction between Gaussian

☆ This research was funded by a NWO-TOP grant (407-11-040) awarded to Harold Bekkering and Iris van Rooij. We thank the participants of the Lorentz Workshop on Human Probabilistic Inference and two anonymous reviewers for helpful discussions and constructive comments on an earlier version of this paper.

* Corresponding author at: Donders Centre for Cognition, Montessorilaan 3, 6525 HR Nijmegen, The Netherlands.

E-mail address: j.kwisthout@donders.ru.nl (J. Kwisthout).

http://dx.doi.org/10.1016/j.bandc.2016.02.008

densities and categorical probability distributions is that in the latter the *state space granularity* (how detailed are the generative models and the predictions that follow from them) is crucial. Whereas the amount of uncertainty (or *precision*) in a Gaussian density can be adequately described by its variance, the precision in a categorical distribution must be described by its *entropy* (Shannon, 1948), which is a function of both the state space granularity and the nonuniformity of the distribution (Kwisthout & Van Rooij, 2015).

Note that this state space granularity is context-dependent. Crucial in the coin-tossing example is that we describe the outcome of 'tossing a coin' in terms of the side of the coin to land on top, disregarding all other information in the outcome (such as the amount of rotation of the coin in the plane) as irrelevant. Compare this with throwing a regular die. As all sides of the die are equally likely to land on top, one can expect an odd number to fall just as often as an even number. When the outcome of a die is predicted in terms of whether the number will be odd or even (and the result interpreted likewise), the precision of that prediction is equal to the precision of tossing a coin. However, if the outcome of a die is predicted in terms of the number of pips, and the result interpreted likewise, the prediction is more uncertain – simply, because there are more possible events ('1', ..., '6'; rather than 'odd' or 'even') and each event is equally likely – therefore, the prediction will have lower precision because a prediction was made (and the outcome interpreted) at a higher level of detail (Fig. 1). The precision of a prediction, hence, is indeed a function of both state space and nonuniformity.

Disentangling precision into level of detail and nonuniformity becomes necessary when cognitive (neuro) scientists aim to describe predictions and observable outcomes in terms of discrete, categorical events (Kwisthout & Van Rooij, 2015). Such outcomes may be the result of a coin flip (heads, tails) or of a die throw (odd, even; or 1...6, depending on how detailed our prediction is); they may describe the next action of a car in front of us (turn left, turn right, park, brake, or just keep driving), or of our spouse's emotions (sad, frustrated, happy, angry, bored; or whatever distinctions one makes); it may be a description of what one expects to see in a forest ('trees and other life forms'; or, when looking more closely, a chestnut tree, a squirrel, moss, bugs, etcetera). The appropriate level of detail that describes such outcomes is typically highly context-specific and depends on the epistemic and practical goals of the observing agent.

In this paper, we explore the computational and theoretical consequences of formalizing predictive processing in categorical probability distributions. After describing the predictive processing account more specifically, we introduce *level of detail* as a concept that intuitively captures the state space granularity, and together with the nonuniformity of the distribution describes its precision or entropy. We define the key computational processes in predictive processing in terms of (hierarchical, dynamical, multi-dimensional) causal Bayesian networks (Pearl, 2000). We show that manipulating the level of detail of generative models and/or predictions allows for the modulation of precision: For example, we can increase the precision of a prediction by decreasing the level of detail of the prediction. This, however, comes at the loss of information that can be gained by correct predictions. How this trade-off between predictions with high precision and predictions with high information gain is resolved in the brain is a key open theoretical (and empirical) question to address.

2. Predictive processing

The Predictive Processing (hereafter PP) account is becoming more and more popular as a unifying theory of what drives our cortical processes.[1] It encompasses key concepts such as the Bayesian brain (the brain encodes probability measures and balances prior expectations to sensory evidence according to the laws of probability theory, in particular Bayes' theorem; Knill & Pouget, 2004), the brain as prediction machine (the brain continuously makes predictions about future sensory evidence based on its current best model of the causes of such evidence; Dayan, Hinton, & Neal, 1995; Hohwy, 2007), the free energy principle (the brain minimizes overall expected prediction error as a proxy to minimize free energy; Friston, 2010) and the hierarchical organization of the brain (Friston, 2005, 2008). In particular it is claimed that the PP account applies to the entire cortex (Clark, 2013a) and that the same generic apparatus and mechanisms are used for both lower and higher cognition, e.g., both low-level vision and high-level intention attribution (Clark, 2013b; Kilner, Friston, & Frith, 2007; Koster-Hale & Saxe, 2013). However, to account for "higher cognitive phenomena such as thought, imagery, language, social cognition, and decision-making" there is still "plenty of work to do" (Hohwy, 2013, p. 3). In particular it is as yet unknown "What [...] the local approximations to Bayesian reasoning look like as we depart further and further from the safe shores of basic perception and motor control? What new forms of representation are then required, and how do they behave in the context of the hierarchical predictive coding regime?" (Clark, 2013a, p. 201).

PP can be understood as a cascading hierarchy of increasingly abstract (e.g., in time scale or space) hypotheses about the world, where the predictions on one level of the hierarchy are identified with the hypotheses at the subordinate level. The inference process, i.e., inferring assumed causes from stimuli, is presumed to be facilitated by having predictions (stemming from the generative, top-down process) at each level of the hierarchy, comparing these predictions with the observed (or inferred) observations, and using the prediction error to update both the current hypothesis and to learn for future predictions.

For example, in the action understanding domain, the hierarchy can include the actual visual, auditory, tactile, or olfactory *inputs*, like a series of visual inputs, at the lowest level; one level above we may situate the *kinematics*, like a grasping movement of the hand, followed by the more abstract object-oriented *actions* (picking up a cup). Eventually, the hierarchy may include complex social cognitive constructs such as future *intentions*, social conventions, world knowledge, context etcetera (Kilner et al., 2007). However, the PP account is currently computationally fleshed out predominantly at the basic perception and motor control level (Clark, 2013a; see also Hohwy, 2013). In particular, computational implementations of PP (typically grouped under the denominator *hierarchical predictive coding* or HPC), such as those suggested by Rao and Ballard (1999), Lee and Mumford (2003), and Friston (2005, 2010), reside at that level.

In a probabilistic interpretation, making a prediction based on the current hypothesis in any of the assumed levels corresponds to computing a posterior probability distribution $P(Pred|Hyp)$ over a space of candidate predictions, given the current estimated distribution over a space of hypotheses.[2] Computing (the magnitude of) a prediction error corresponds to computing the relative entropy

[1] An illustration of this might be the observation that a separate outlet (Cleeremans & Edelman, 2013) was created to allow for the large number of commentaries to Clark's (2013a) target article in *Behavioral and Brain Sciences.* Also indicative is that a search on "predictive coding" and "predictive processing" on Google Scholar together found about 2500 papers published in 2014.

[2] There appears to be some ambiguity in the literature about whether a prediction (hypothesis) refers to a distribution over candidate predictions (hypotheses), or the *mode* of that distribution; see, e.g., Kilner et al. (2007, p. 161), Hohwy, Roepstorff, and Friston (2008, p. 691), and Hohwy (2013, p. 61) for examples that suggest the latter. In this paper we adhere to the view (e.g., Knill & Pouget, 2004; Lee & Mumford, 2003; Friston, 2009) that suggests that whole distributions (or approximations thereof) are maintained, without claiming that this debate has fully settled yet. In the remainder, unless explicitly noted, *hypothesis* refers to a probability distribution over a space of candidate hypotheses, and similar for predictions.