Cognitive Psychology 95 (2017) 17–49

Contents lists available at ScienceDirect

Cognitive Psychology

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

Learning to allocate limited time to decisions with different expected outcomes

Arash Khodadadi *, Pegah Fakhari, Jerome R. Busemeyer

Indiana University, Department of Psychological and Brain Sciences, Bloomington, IN, United States

ARTICLE INFO

Article history:
Accepted 29 March 2017
Available online 19 April 2017

Keywords:
Sequential sampling models
Opportunity cost
Reward rate maximization
Speed-accuracy trade-off
Reinforcement learning

ABSTRACT

The goal of this article is to investigate how human participants allocate their limited time to decisions with different properties. We report the results of two behavioral experiments. In each trial of the experiments, the participant must accumulate noisy information to make a decision. The participants received positive and negative rewards for their correct and incorrect decisions, respectively. The stimulus was designed such that decisions based on more accumulated information were more accurate but took longer. Therefore, the total outcome that a participant could achieve during the limited experiments' time depended on her "decision threshold", the amount of information she needed to make a decision. In the first experiment, two types of trials were intermixed randomly: hard and easy. Crucially, the hard trials were associated with smaller positive and negative rewards than the easy trials. A cue presented at the beginning of each trial would indicate the type of the upcoming trial. The optimal strategy was to adopt a small decision threshold for hard trials. The results showed that several of the participants did not learn this simple strategy. We then investigated how the participants adjusted their decision threshold based on the feedback they received in each trial. To this end, we developed and compared 10 computational models for adjusting the decision threshold. The models differ in their assumptions on the shape of the decision thresholds and the way the feedback is used to adjust the decision thresholds. The results of Bayesian model comparison showed that a model with time-varying thresholds whose parameters are updated by a reinforcement learning algorithm is the most likely model. In the second experiment, the cues were not presented. We showed that the optimal strategy is to use a single time-decreasing decision threshold for all trials. The results of the computational modeling showed that the participants did not use this optimal strategy. Instead, they attempted to detect the difficulty of the trial first and then set their decision threshold accordingly.

1. Introduction

Suppose you are taking an exam. You have one hour to answer as many questions as you can. In addition, suppose that there are two types of questions, easy and hard. How much time should you spend on each question? For example, if the questions are presented sequentially and the first question is hard, would you be willing to spend 10 min on that question?

* Corresponding author at: Indiana University, Department of Psychological and Brain Sciences, 1101 E. 10th street, 47405-7007 Bloomington, IN, United States.

E-mail address: arash.khodadadi@gmail.com (A. Khodadadi).

http://dx.doi.org/10.1016/j.cogpsych.2017.03.002
0010-0285/

In this scenario, every moment that one spends on one question, less will remain for other questions and so fewer questions can be answered in limited time. On the other hand, by answering the questions too fast, the accuracy drops and one may be able to answer only a few questions correctly. This results in a trade-off between the speed and the accuracy.

This is an example of a more general problem in which a living organism has to allocate a limited resource to different courses of actions. Some examples of a limited resource are: energy, time, memory, attention and so on. Usually, spending more of the resource on a course of action results in more desirable outcome for those actions. However, spending more of the resource on some actions leaves less for the other actions, and this may result in lower *total outcome*. Therefore, the organism must allocate the resource "wisely" to obtain the maximum total outcome over all actions.

A situation in which this sort of trade-off arises naturally is perceptual decision making in which the animal has to make decisions based on noisy information. Usually, by spending more time the animal can make more accurate decisions which in turn lead to more desirable outcomes. A large amount of research has focused on explaining the relationship between the decision time and the accuracy in perceptual decision making, both theoretically and experimentally (Brown & Heathcote, 2005; Gold & Shadlen, 2002; Jones & Dzhafarov, 2014; Kiani, Hanks, & Shadlen, 2008; Khodadadi & Townsend, 2015; Ratcliff, 1978; Ratcliff, Van Zandt, & McKoon, 1999; Smith, 2000; Townsend & Ashby, 1983; Teodorescu & Usher, 2013; Usher & McClelland, 2001). The most popular theoretical framework for explaining the mechanism underlying this relationship is provided by a class of models known as *sequential sampling models*. A common assumption between different instantiations of these models is that the animal sequentially samples evidence favoring each of the possible decisions. Since these samples are noisy, a decision based on one sample will be very inaccurate. Instead, the brain accumulates these samples until the accumulated evidence favoring one of the decisions reaches a specific level called the *decision threshold*. Larger values of the decision threshold lead to slower but more accurate decisions. The rate at which the information is accumulated is proportional to the difficulty of the stimulus and so it is controlled by the experimenter and not the participant.

Experimental results together with computational modeling have shown that human participants adjust the value of their decision threshold in response to the emphasis on the speed or the accuracy in the instructions of the experiment (Forstmann et al., 2010; Luce, 1986; Ratcliff, 2002; Wagenmakers, Ratcliff, Gomez, & McKoon, 2008). This experimental paradigm, provides evidence that human participants can adjust their decision threshold when they are asked to do so. However, it does not show if this threshold adjustment will occur in order to maximize the outcome. Recently, some theoretical and experimental work has investigated this question. Gold and Shadlen (2002) proposed an experimental paradigm in which the participants had to make a sequence of decisions during a limited time. The participants received some rewards or punishments for their correct or incorrect decisions. Since time is limited, the participant should balance between her speed and accuracy to achieve the maximum amount of reward during the experiment. Bogacz, Brown, Moehlis, Holmes, and Cohen (2006) investigated the optimal strategies in this paradigm. Specifically, they showed the relationship between the optimal value of the decision threshold and the parameters of the experiment including the difficulty of the stimulus and the value of the reward and punishment. More recently, Simen et al. (2009), Bogacz, Hu, Holmes, and Cohen (2010), Balci et al. (2011) and Evans and Brown (2016) examined experimentally if human participants can learn the optimal decision threshold in this paradigm.

These studies have shed light on several aspects of the decision making mechanisms involved in balancing between speed and accuracy in information accumulation paradigms. However, many questions have remained unanswered. In this paper, we extend the previous research in several directions in order to investigate some of these questions. We outline these directions next.

1.1. A novel stimulus and decision paradigm

The speed-accuracy trade-off have been mainly investigated using perceptual decision making paradigms. These experiments are appealing because it is easy to manipulate the difficulty of the task to achieve a wide range of accuracy (from chance level to nearly perfect accuracy) and reaction time. However, using these stimuli for studying the properties of the decision thresholds have several drawbacks. First, in the tasks which are commonly used to study perceptual decision making, for example the random dot motion experiment (Britten, Shadlen, Newsome, & Movshon, 1992; Shadlen & Newsome, 2001), neither the accumulated information nor the decision threshold are directly observable. The only observable variables are the participants' choice and reaction time in each trial. Therefore, to infer the properties of the decision threshold in these experiments, one should either use the neuro-physiological data (Forstmann et al., 2010; Ivanoff, Branning, & Marois, 2008; Kiani et al., 2008; Shadlen & Newsome, 2001; Ratcliff, Hasegawa, Hasegawa, Smith, & Segraves, 2007), or computational modeling (Ratcliff, 1978; Ratcliff & Smith, 2004; Smith, 1995; Usher & McClelland, 2001). This makes the inference about the properties of the decision thresholds harder than if the decision threshold could be observed directly. Second, for the same level of task difficulty, there is usually a large amount of variations in the participants' performance. This is due to individual differences in perceiving the same stimulus. In the language of the sequential sampling models, for the same stimulus, the participants have different rate of information accumulation. For this reason, the properties of the optimal decision threshold will be different for different participants. Third, there is usually a large perceptual learning effect in these tasks. With experience, the rate of information accumulation increases for a participant (see for example Fig. 8 in Balci et al. (2011)). Therefore, the properties of the optimal decision threshold changes for a participant during the experiment. Fourth, the participants' average reaction time in these experiments is usually very short. As we will argue later, this may put some constraints on the shape of the decision thresholds.