ARTICLE IN PRESS
YJETH:4253

ELSEVIER

Available online at www.sciencedirect.com

ScienceDirect

Journal of Economic Theory ••• (••••) •••–•••

JOURNAL OF Economic Theory

www.elsevier.com/locate/jet

Optimal categorization

Erik Mohlin

Nuffield College and Department of Economics, University of Oxford, New Road, Oxford OX1 4PX, United Kingdom

Received 12 December 2011; final version received 8 September 2013; accepted 27 February 2014

Abstract

This paper studies categorizations that are optimal for the purpose of making predictions. A subject encounters an object (x, y). She observes the first component, x, and has to predict the second component, y. The space of objects is partitioned into categories. The subject determines what category the new object belongs to on the basis of x, and predicts that its y-value will be equal to the average y-value among the past observations in that category. The optimal categorization minimizes the expected prediction error. The main results are driven by a bias-variance trade-off: The optimal size of a category around x, is increasing in the variance of y conditional on x, decreasing in the variance of the conditional mean, decreasing in the size of the data base, and decreasing in the marginal density over x.

JEL classification: D83; C72

Keywords: Categorization; Priors; Coarse reasoning; Similarity-based reasoning; Case-based reasoning; Regression trees

1. Introduction

Numerous studies in psychology and cognitive science have demonstrated the fundamental role played by categorical reasoning in human cognition. In particular, categorical reasoning facilitates prediction.[1] Prediction on the basis of categories is relevant in situations where one

E-mail address: erik.mohlin@nuffield.ox.ac.uk.

[1] For overviews of the voluminous literature, see e.g. Laurence and Margolis [26], or Murphy [33]. Regarding categorization and prediction, see Anderson [2]. Categorical thinking matters in economic contexts: Consumers categorize products [39], investors engage in "style investing" [5], rating agencies categorize firms w.r.t. default risk [9].

http://dx.doi.org/10.1016/j.jet.2014.03.007
0022-0531/

Please cite this article in press as: E. Mohlin, Optimal categorization, J. Econ. Theory (2014), http://dx.doi.org/10.1016/j.jet.2014.03.007

has to predict the value of a variable on the basis of one's previous experience with similar situations, but where past experience does not necessarily include any situation that is identical to the present situation. One may then divide the experienced situations into categories, such that situations in the same category are similar to each other. When a new situation is encountered one determines what category this situation belongs to, and the past experiences in this category are used to make a prediction about the current situation. These predictions can be computed in advance, thereby facilitating a rapid response.

Assuming that we use categorizations to make predictions, I ask which categorizations that are optimal, in the sense that they minimize prediction error.[2] The optimal number of categories is derived without imposing any exogenous costs or benefits of the number of categories. Instead costs and benefits arise endogenously from a bias-variance trade-off that is inherent to the objective of making accurate predictions. The advantage of fine-grained categorizations is that objects in a category are similar to each other. The advantage of coarse categorizations is that a prediction about a category is based on many observations.

The focus on optimal categorizations stems from evolutionary considerations.[3] Many categorizations are acquired early in life, through socialization and education, or because they are innate. From an evolutionary perspective we would expect humans to employ categorizations that generate predictions that induce behaviour that maximize fitness. It seems reasonable to assume that fitness is generally increasing in how accurate the predictions are. For instance, a subject encountering a poisonous plant will presumably be better off if she predicts that the plant is indeed poisonous, rather than nutritious. Hence, we would expect humans to have developed, and passed on, categorizations that are at least approximately optimal, in the sense that they tend to minimize prediction error in the relevant environments. Such categorizations will be called *ex ante optimal*.[4] Other categorizations are developed only after a data base of experiences has been accumulated. We would expect evolution to have endowed us with heuristics or algorithms that allow us to form categorizations that organize our experience in a way that tends to minimize prediction error, conditional on the data base. Categorizations that attain this goal will be called *ex post optimal*.[5]

As an example of a categorization that is acquired very early on, think of colour concepts. The subset of the spectrum of electromagnetic radiation that is visible to the human eye allows for infinitely fine-grained distinctions. However, in every day reasoning and discourse we employ a coarse colour classification, using words such as red and green. Presumably the colour categorizations that were developed and passed on to new generations were successful in the kind of environments that we faced.[6] As an example of categorizations that are formed after a data base has been accumulated, one may think of the many classifications that science has produced. The

[2] In Section 5.1 I argue that categorization-based prediction is less cognitively demanding than other forms of similarity-based reasoning, such as kernel-based estimation.

[3] One might suggest that a categorization is optimal if it is induced by a language that is optimal, in some sense. Language is undoubtedly important in shaping our concepts, but concepts seem to have come prior to language in evolution; there are animals who use concepts even though they do not use language (see e.g. Herrnstein et al. [18]), and children can use certain concepts before they have a language (see e.g. Franklin et al. [11]). This suggests that we need to explain the use of categories without reference to language.

[4] One might ask why the exact distribution of objects is not transmitted between generations. I will simply take it as an empirical fact that many categorizations are transmitted between generations. This indicates that there are some factors that make it infeasible or inefficient to transmit detailed information about the distribution.

[5] See Chater [8] on the relationship between simplicity and likelihood in perceptual organization.

[6] For inter-cultural comparisons, see Kay and Maffi [23] and references therein.

Please cite this article in press as: E. Mohlin, Optimal categorization, J. Econ. Theory (2014), http://dx.doi.org/10.1016/j.jet.2014.03.007