Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Part B 41 (2007) 772–783

TRANSPORTATION
RESEARCH
PART B

www.elsevier.com/locate/trb

Parametric action decision trees: Incorporating continuous attribute variables into rule-based models of discrete choice

Theo Arentze, Harry Timmermans *

Eindhoven University of Technology, Department of Building and Architecture, Urban Planning Group, P.O. Box 513, 5600 MB Eindhoven, The Netherlands

Received 10 January 2006; received in revised form 20 October 2006

Abstract

Rule-based models, such as decision trees, are ideally suited to represent discontinuous effects of independent variables on discrete choice behavior in transport or spatial systems. At the same time, however, the models require that continuous attributes, such as for example travel time and travel costs, are discretisized, which may decrease the sensitivity of predictions for policy measures that involve these attributes. To overcome this problem and combine the specific strengths of the rule-based and parametric modeling approaches, this paper introduces a hybrid approach. The so-called parametric action decision tree (PADT) replaces the conventional action-assignment rule of the decision tree by a logit model or any other parametric discrete choice model. The PADT includes alternative-specific constants to take the impact of leaf-node membership into account in addition to terms for the continuous attributes. As an illustration, we show how the approach can be used to incorporate travel-costs sensitivity in Albatross – a rule-based model of activity-travel choice. The results indicate that the enhanced, hybrid model can reproduce realistic ranges of price elasticities of travel demand.

Keywords: Discrete choice models; Rule-based models; Decision-tree induction; Activity-based models

---

1. Introduction

The problem of dealing with non-compensatory decision rules within a random-utility-maximizing framework of discrete choice has received continuous attention in travel-behavior research since Tversky's seminal work on behavioral decision making (Tversky, 1972). Swait (2001) proposed a non-compensatory choice model incorporating attribute cut-offs into the decision model and tested it as an extension of the traditional compensatory utility maximization framework. Cantillo and Dios Ortúzar (2005) proposed a hybrid semi-compensatory two-stage model incorporating thresholds for the acceptance of attributes in the process of

---

* Corresponding author. Tel.: +31 40 247 3315; fax: +31 40 243 8488.
E-mail address: eirass@bwk.tue.nl (H. Timmermans).

doi:10.1016/j.trb.2007.01.001

discrete choice, similar to earlier work by Timmermans et al. (1986). The rule-based formalism, such as the production system, provides an alternative framework, which is particularly suitable to represent discontinuous effects of attributes on choice behavior (Gärling et al., 1994). Rule-based models are founded on theories of bounded rationality (Newell and Simon, 1972) and can be extracted from choice observations through rule-induction methods that emerged from recent work in artificial intelligence, statistics and other disciplines. Several studies demonstrated the use of rule-induction methods to model spatial and travel choice behavior (Thill and Wheeler, 2000; Gahegan, 2000; Arentze et al., 2000; Wets et al., 2000; Yamamoto et al., 2002; Xie et al., 2003; Moons et al., 2005).

Compared to parametric methods, the strength of the rule-based approach is that it allows one to represent various types of interactions between variables, such as conditional relevance and conditional classification (Van der Smagt and Lucardie, 1991). Conditional relevance occurs if an attribute variable is relevant conditional upon some other variable. For example, in an elimination-by-aspects heuristic, the relevance of a certain attribute (i.e., aspect) is dependent on the performance of alternatives on a higher-level attribute. Only if the choice alternatives perform equally well on the first attribute, a second attribute is taken into consideration. The same holds for the next attribute in the hierarchy and so on. Conditional classification, on the other hand, is related to the phenomenon that an individual may be indifferent for variation on an attribute within certain sections of the domain of the attribute. In such cases, the number and position of cut-off points on the range of the variable determines how the individual classifies an outcome of the attribute. The classification may be dependent on other variables. An example is the perception of travel time of a trip in terms of labels such as short and long. The perception may be dependent on the transport mode used for the trip or situational settings such as the weather or time pressure. Conditional relevance/classification gives rise to discontinuous effects of attribute variables on choice probabilities. A change in situation or attributes of choice alternatives may have less or more than proportional impacts depending on whether or not they lead to a qualitative shift in perception.

This strength of rule-based approaches may at the same time be a weakness. The discrete representation of variables is not only a power but also a restriction of the formalism for variables measured on a continuous (interval or ratio) scale.[1] Some rule-induction methods are able to optimize the choice of cut-off-points simultaneously with constructing the tree.[2] However, even if classifications are flexible, the assumption of indifference between outcomes within classes is still imposed on the model. The property is a power if such limited sensitivity accurately describes the choice behavior modelled. However, if classifications underlying choice behavior in a population studied differ between individuals or situations not captured by condition states in the tree, the response behavior of the population may be better approximated by allowing for some sensitivity across the continuous range. In transport and spatial systems, continuous service-level variables, such as travel time and travel costs, are often important explanatory variables for choice facets of trips such as frequency, destination, transport mode and even departure time. If (condition-dependent) ranges of indifference on such service-level variables are not uniformly distributed in the population, the response to changes may vary in an approximately continuous fashion at an aggregate level and the rule-based model would fail to reproduce this. Furthermore, elasticities (i.e., marginal substitution rates) related to continuous variables, which are often relevant for policy making and theory development, are not revealed when their values are discretisized.[3]

To address this problem and combine the specific strengths of rule-based and parametric methods, we propose a more flexible formalism defined as an extension of the decision tree. In principle, several formalisms including decision trees, production systems, association rules, causal networks and decision tables can be applied (Witten and Frank, 2005; Keuleers et al., 2001; Janssens et al., 2004, 2005). The decision tree has

---

[1] Some rule-induction methods, such as CART (Breiman et al., 1984) and M5 (Quinlan, 1992), do allow *dependent* variable to be continuous and to be treated as continuous. Our argument here, however, is related to *independent* variables.

[2] An example of a rule-induction method that optimizes the discretization of attribute variables simultaneously with the structure of the tree is C4.5 (Quinlan, 1993).

[3] This is not to say that no elasticity information can be derived from rule-based models. Simulation methods have been proposed to extract elasticity information from rule-based model in a pre-processing step (Arentze and Timmermans, 2003a). However, for continuous variables elasticity information obtained in that way is limited if these variables are discretisized.