The Ship of Theseus: Using Mathematical and Computational Models for Predicting Identity Judgments

Tuna Cakar (e170996@metu.edu.tr)
Cognitive Science Department, Informatics Institute
Middle East Technical University, Ankara, TURKEY

Annette Hohenberger (hohenber@metu.edu.tr)
Cognitive Science Department, Informatics Institute
Middle East Technical University, Ankara, TURKEY

Keywords: identity judgments; paradoxical reasoning; Ship of Theseus; modeling of decisions.

Introduction

Reasoning processes have been one of the central targets for cognitive modeling. Modeling of reasoning processes appears as an even harder challenge during paradoxical conditions such as the Ship of Theseus paradox. This work attempts to model empirical data from a behavioral study on paradox resolution with different modeling techniques: discriminant analysis (DA), decision tree analysis and neural networks. While each method has its own advantages and disadvantages, this paper attempts to compare and to contrast these methods trying to select the best model for future work.

Identity judgments have long been at the center of philosophical debates, e.g., is a car still the same after being fixed after a serious accident? Beyond the philosophical debates on the nature of objects and the concept of identity, it has also been a matter of interest how laymen respond to the identity question under different circumstances. The present study focuses on a famous paradox from ancient Greece, the Ship of Theseus (Hall, 1998). Answers to this paradox have been predicted by a Conceptual Tendency Test (CTT) tapping the concept of "sameness". The main aim of this paper is to present and compare the results of three predictive models in terms of their accuracy (predictive success) and to discuss the theoretical basis of the findings.

The initial, empirical part of this work aims to understand how participants reason during resolution of a given paradox, namely, the Ship of Theseus (Clark, 2002). In a nutshell, a ship owned by Theseus has been renewed part by part over time. At the end, all of the parts of the old ship have been renewed (Ship A) and the removed parts were reassembled to build another ship (Ship B). Thus, there are two ships finally. The classical paradox is: Which ship is the ship of Theseus, the renewed one (Ship A) or the one that has been reassembled with the old parts (Ship B)? Ship A responses seem to reflect a "functionalist" position, i.e., the function of the ship has been preserved; whereas ship B responses seem to reflect an essentialist position, i.e., the physical essence of the ship has been preserved. The problem has been discussed by several philosophers and related to the concept of "sameness" or "identity" (Wiggins, 2001). It is plausible to assume that participants' decisions are determined by several dimensions involved in the critical concept at stake, among them spatiotemporal considerations: how long did the renewal and reassembly process take (short or long) and where did it take place (at a proximal or distal place)? (Rips *et al*., 2006; Scholl, 2007). Functionalist and essentialist positions could be affected by these parameters differently. Participants initially performed a Conceptual Tendency Test (CTT) in which they were asked to rate a set of propositions which are directly related to the core concept of "sameness/identity" involved in the paradox before answering the paradox (see method).

In this current work, we focus on two main research questions: (1) Do the identity judgments in the CTT contain conceptual cores that are influential during the reasoning process on the paradox? (2) Can the final decision of a participant be predicted by the CTT? Our hypotheses on paradox resolution are as follows:

H_1: Participants take spatiotemporal features into account while making decisions about judgments on identity of an object over time.

H_2: The final decisions of the participants to the paradox can be predicted by their response to the CTT.

Experimental Design & Data Collection

50 undergraduate and graduate students (25 female; age-range 19-28 years) were allocated to the two experimental conditions – high vs low spatiotemporal proximity (STP) – randomly. Fourty-eight propositions were prepared as image files and were randomly presented for 15 seconds on the computer screen by E-Prime 1.0. Participants were initially asked to rate a set of propositions (the CTT) which are directly related to the core concept involved in the paradox. Participants responded on a scale from 1 to 5 (where 1 corresponds to total agreement, 3 to neutral, and 5 to total disagreement). There were 24 proposition pairs half of which were phrased in terms of "same" ("A" for Turkish "ayni" ("same")), and half in terms of "different" ("F" for Turkish "farkli" ("different")), e.g., A: "A bicycle that has its pedals removed is the same"; F: "A piece of paper bent over 3 times is different". Each proposition was presented for 15 seconds. After the presentation of all propositions, participants were then presented with the Ship of Theseus

paradox in one of the STP conditions: in the high STP condition, the ship was renewed/reassembled over a short period of time (5 years) and at a neighboring port; in the high STP condition it was renewed/reassembled over a long period of time (50 years) and at a distant port. Participants were given unlimited time to respond on a 5-point Likert scale, where ratings of 1 and 2 were considered as strong and weak "ship A" ratings, ratings of 3 as "undecided" and ratings of 4 and 5 as weak and strong "Ship B" ratings.

Results & Discussion

Behavioral Results

The obtained responses to the Ship of Theseus question revealed a bimodal, M-like distribution (20 cases for ship A and 24 cases for ship B) for the paradox. The M-shape indicates that few subjects would take a strong stance and respond with the value 1 for A or 5 for B, respectively, but rather take a weak stance (2 or 4). In the middle of the distribution were 6 undecided participants. This result shows that participants avoided strict positions but rather stayed in a flexible zone while reasoning about the paradox. Responses for the two conditions (high STP vs low STP) were almost equally distributed. In other words, there was no effect of condition, contrary to our hypothesis. The initial statistical analysis on the Conceptual Tendency Test (one-way ANOVA) revealed that 4 different propositions (P17F, P21A, P6F, and P2F) reached significance and 6 further propositions (P9F, P10A, P10F, P14F, P17A, P6A) reached marginal significance ($p<.08$). Among the significant propositions was P17F, stating that a robot that had been disassembled and reassembled was different now; P2F: that two birds with identical genetic and behavioral features were different; P6F: that a piece of paper bent over 3 times was different; and P21A: that a robot with memory problems after a memory chip transplantation was the same.

Modeling Results

As a first mathematical model, Discriminant analysis (DA) was used to classify the responses to the paradox relying on the responses to the CTT. In DA groups of participants are discriminated based on linear combinations of variables. The initial discriminant analysis was run with 2 variables for the response to the paradox (Ship A or Ship B). Strong and weak positions for Ship A (1,2) and for Ship B (4,5) were therefore collapsed and intermediate positions (3) were eliminated in order to meet the statistical assumptions (Box's M-Test). Wilks' lambda was significant for the single function that the DA had computed (V=0.571, $\chi^2(9)$=21.020, p=.013). The canonical correlation that is a function of the eigenvalue was .655 for this function whereas the eigenvalue had the value of .751. 79.5% of the originally grouped cases were correctly classified. 6 out of 9 (66.6%) misclassified responses stemmed from weak positions (2 or 4) that were obviously harder to classify than strong positions (1 and 5).

As a second model, a decision tree analysis was performed for the same data based on two core propositions, namely P10F and P17F (both at significance level p=.002, Bonferroni-adjusted), resulting in 77.3% predictive success This decision tree consists of 5 nodes (3 of which are terminal nodes) and has the depth of 2. It is important to note that both of the propositions are 'different' (F) statements. This finding indicates that participants responded differently to the 'same' (A) versus 'different' (F) propositions. Interestingly enough, propositions with 'different' status were found to be more critical in predicting identity judgments.

As a third and last model the identity judgments were modeled with the neural network modeling technique (multilayer perceptron) relying on the same critical propositions of the CTT. 70% of the cases were used as training items and 30% as test items. The model was run with two units in a single hidden layer and the activation function was hyperbolic. The obtained Neural Network Model classified test cases with 88.9% predictive success.

Conclusion

Our modeling results revealed that the use of mathematical models like DA is beneficial in order to understand and explain the reasoning processes during paradox resolution like in the Ship of Theseus paradox – despite the fact that a computational model like the neural network model could predict the same data better. The present work demonstrates that the final judgments of the participants to the paradox could be predicted with a relatively high predictive success (>77%) solely relying on some critical propositions of a previously designed Conceptual Tendency Test (CTT). Participants tend to rely on a conceptual core about the target concept "sameness" which guides them through their reasoning process. Moreover, 'different' statements seem to play a more critical role in identity judgments when compared to 'same' statements. This finding suggests that these are two distinct cognitive processes even though they appear to be similar and participants were not aware of the fact that they responded differently to 'same' and 'different' statements. In conclusion, modeling is a worth-while methodology in order to better understand higher cognitive processes such as reasoning about paradoxes.

References

Clark, M. (2002). *Paradoxes from A to Z*. TJ International Ltd: New York.

Hall, D. G. (1998). Continuity and the persistence of objects: When the whole is greater than the sum of the parts. *Cognitive Psychology*, *37*, 28-59.

Rips, L. J., Blok, S., & Newman, G. (2006). Tracing the identity of objects. *Psychological Review*, *113*, 1-30.

Scholl, B. J. (2007). Object persistence in philosophy and psychology. *Mind & Language, 22*, 563-591.

Wiggins, D. (2001). *Sameness and Substance Renewed*. Cambridge: Cambridge University Press.