In Proceedings of the Sixth International Conference on Cognitive Modeling, 406-407.
Mahwah, NJ: Lawrence Earlbaum.

Learning and Strategy Selection in Probabilistic Environments

Wolfgang Gaissmaier (gaissmaier@mpib-berlin.mpg.de)
Center for Adaptive Behavior and Cognition, Max Planck Institute for Human Development
Lentzeallee 94, 14195 Berlin, Germany

Abstract

Many decisions have to be made on the basis of knowledge about correlational structures in the environment. It has been found that people with a low working memory capacity perform better in a covariation detection task (Kareev, Lieberman, & Lev, 1997). This has been attributed to the assumption that they can only consider smaller samples which are more likely to bear a correlation parameter that exceeds the population parameter. Our data and results from a reinforcement learning model on an extended version of their task do not clearly support this account. As alternative explanations differences in reinforcement learning, hypothesis generation and strategy selection are considered. A very simple strategy (payoff maximization) is most successful in this environment. Differences in capacity and strategy selection are to be further studied in a cue based social categorization task, namely to predict political party preferences. Here, people have to apply knowledge about correlations acquired in the real world. It is hypothesized that people with a lower capacity use simpler strategies that could be, again, even more successful than more complex strategies because they exploit the structure of the environment. This could be reflected in behavioral data and in differential model fit to a variety of models like exemplar based models or Categorization by Elimination.

Keywords: Working Memory Capacity, Covariation Detection, Reinforcement Learning Model, Simple Strategies, Social Categorization

Introduction

We are concerned with the question how people learn correlational structures of the environment and how they make decisions on the basis of this knowledge. A major focus is on interindividual differences with regard to cognitive capacity. We are studying as well covariation detection in an artificial environment as a social categorization task in which knowledge about correlations has to be applied that participants acquired not in the lab, but in the real world.

Covariation Detection

The starting point for this project was the argument that cognitive limits are useful with regard to covariation detection (e.g., Kareev, 1995b). The observation is that it is more likely to find a correlation parameter that exceeds the population parameter in small samples. Therefore, people with a lower working memory capacity and hence a smaller sample size to consider should detect correlations earlier and perform better in a covariation task, which has been demonstrated empirically (Kareev et al., 1997).

Experiments

We conducted two experiments to test and model the original finding that low capacity people perform better in a covariation task. An additional hypothesis deduced from the small sample account is that they are also better in detecting a change in the correlational structure of the environment. As in the original experiment, working memory capacity was assessed with a digit span test. The original finding was replicated in the first but not in the second experiment, thus it seems to be a weak and unstable effect. It is worth noting that the probability of replicating a result at the same level of significance (and in the same direction) is only 50% (Goodman, 1992). Contrary to the predictions by the small sample account there was a high capacity advantage after a change in the first experiment. In the second experiment we did not find any differences between low and high capacity people, neither before nor after a change. Therefore, we focus on the first experiment with regard to modeling.

Modeling

Every model was fitted to each individual separately since we want to relate capacity to model parameters. A naïve window model that tries to translate the small sample idea directly could not capture the low capacity advantage. But we were able to model it with a reinforcement learning model (Camerer & Ho, 1999) with a decay, a sensitivity and an initial attraction parameter, where we forced the variance in each of the parameters separately by fixing the other two to their means. All three versions were able to capture the low capacity advantage on covariation detection, but only the initial attraction version was related to capacity and could predict behavior after a change.

Conclusions

The small sample account is not clearly supported by our data. First, the deduced hypothesis of a low capacity advantage after a change does not hold, we find either no effect or the opposite. Second, the model version with the decay parameter which has the strongest connection to memory has to be rejected. Instead, an initial attraction parameter model is successful, indicating a faster learning process of low capacity people in the beginning, but not later on. Still, faster learning can be interpreted as relying on smaller samples. But it is also congruent with the finding of Weir (1964) that children use the simple but most successful payoff maximization strategy (i.e. always choose the more frequent option given a color) earlier in a similar task because they are simply reinforcement driven. Adults, in contrast, develop complex hypothesis and apply complex strategies because they believe that there exists a perfect solution, but they end up worse. As capacity differs between children and adults (Kail, 1984) and plays an important role in hypothesis generation (Dougherty and Hunter, 2002) this could be an explanation for the low capacity advantage. Consistent with this explanation, a simple strategy model (win stay, lose shift) fits better to the behavior of low capacity people.

Predicting Political Preferences

Appropriate behavior in social situations depends strongly on what other people think, e.g. what they like or dislike. As it is often impossible really to know what they think we have to rely on inferences about their opinions based on cues. Imagine that you are having dinner with an important person that you have never met before. If you share similar political opinions, this may be a good topic of conversation, if not, it may be wiser to avoid it. Asking directly is risky, so you infer what the other person thinks based on, e.g., statements regarding other topics he or she made. This inference process can be seen as a social categorization process.

To be able to socially categorize others, e.g. according to party preference, one has to know correlated structures in the environment (e.g. Rosch, 1978). There is a broad literature reporting relations between preferences for particular political parties, the social category that interests us here, and a variety of variables like, e.g., self interest and economic beliefs (Allen & Ng, 2000).

We are interested in the accuracy of people's knowledge about these correlations and want to study which strategies they use when they apply this knowledge to a social categorization task. The hypothesis is that people with a lower cognitive capacity use simpler strategies that could even be more successful than more complex ones as most people whose party preferences have to be predicted do not have complex patterns of opinions but rather one dimensional ones (Gigerenzer, 1982). Furthermore, we want to asses each participant's own opinions and party preference, as it is known that people often overestimate homogeneity in outgroups (i.e. supporters of other parties), but not in their ingroup (Tajfel & Turner, 1986) which could have an influence on the accuracy of different strategies.

Task

The task will be to infer which political party people prefer. Target people are selected from the European Values Survey, a dataset which documents individual data on opinions and party preferences. They are successively presented to our participants who are allowed to ask questions like "Do you have confidence in the armed forces?" which serve as cues. Possible parties are those that are currently represented in the German parliament.

Modeling

We want to test which model best describes the actual behavior of people. In most studies concerning social categorization the categories are the independent rather than the dependent variable. But there are models of categorization that can be transferred to social categorization, e.g. exemplar based models (Nosofsky, 1988) or Categorization by Elimination (CBE, Berretty, Todd & Blythe, 1997). These models also reflect different strategies. Thus, a differential fitting to people's behavior as well as their behavioral data allows us to see which strategies individuals were most likely to use. It is probable that there are interindividual differences, and maybe they relate to memory capacity.

Acknowledgements

I want to thank my advisors Lael Schooler and Jörg Rieskamp for their great help and support.

References

Allen, M.W. & Ng, S.H. (2000). Self interest, economic beliefs and political party preference in New Zealand. *Political Psychology*, *21*, 323-345.

Berretty, P.M., Todd, P.M., & Blythe, P.W. (1997). Categorization by Elimination: A fast and frugal approach to categorization. In M.G. Shafto and P. Langley (Eds.), *Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society* (pp. 43-48). Mahwah, NJ: Lawrence Erlbaum Associates.

Camerer, C.F., & Ho, T. (1999). Experience-weighted attraction learning in normal-form games. *Econometrica*, *67*, 827-874.

Dougherty, M. R. P., & Hunter, J. E. (2002). Hypothesis generation, probability judgment and individual differences in working memory capacity. *Acta Psychologica*, *113*, 263-282.

Gigerenzer, G. (1982). Der eindimensionale Wähler. *Zeitschrift für Sozialpsychologie*, *13*, 217-236.

Goodman, S.N. (1992). A comment on p-values, replication and evidence, *Statistics in Medicine*, *11*, 875-879.

Kail, R.V. (1984). *The development of memory in children* (2nd ed.). New York: Freeman.

Kareev, Y. (1995b). Through a narrow window: Working memory capacity and the detection of covariation. *Cognition*, *56*, 263-269.

Kareev, Y., Lieberman, I., & Lev, M. (1997). Through a narrow window: Sample size and the perception of correlation. *Journal of Experimental Psychology: General*, *126*, 278-287.

Nosofsky, R. M. (1988). Similarity, frequency, and category representations. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *14*, 54-65.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. Lloyd (Eds.), *Cognition and Categorization*. Erlbaum: Hillsdale, NJ.

Tajfel, H. & Turner, J.C. (1986). The social identity theory of intergroup behaviour. In S. Worschel & W.G. Austin (Eds.), *Psychology of Intergroup Relations* (2nd ed.). Chicago: Nelson-Hall.

Weir, M. (1964). Developmental Changes in Problem-Solving Strategies. *Psychological Review*, *71*, 473-490.