International Journal of Approximate Reasoning ••• (••••) •••–•••

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

www.elsevier.com/locate/ijar

Plausible reasoning and plausibility monitoring in language comprehension ☆

Maj-Britt Isberner [a], Gabriele Kern-Isberner [b]

[a] Department of Psychology, University of Kassel, Germany
[b] Department of Computer Science, University of Technology Dortmund, Germany

ARTICLE INFO

Article history:
Received 27 November 2016
Received in revised form 8 May 2017
Accepted 9 May 2017
Available online xxxx

Keywords:
Plausible reasoning
Belief revision
Language comprehension
Spohn's ranking functions
c-representations
c-revisions

ABSTRACT

In psychological research on language comprehension, so-called epistemic Stroop effects illustrate how implausible information can interfere with human action decisions, i.e., actions with positive goals can be delayed after implausible information, and vice versa. The basic assumption here is that humans reason from suitable situation models that are built upon background beliefs. In this paper, we present formal models that are apt to simulate cognitive processes that are relevant for language comprehension and these epistemic Stroop effects. Since background knowledge is crucial for the situation model, we use the inductive methods of c-representation and c-revision that are capable of processing explicit (conditional) knowledge bases to make plausible reasoning in the experimental tasks transparent. We argue that the delays in response time are partially caused by belief revision processes which are necessary to overcome the mismatch between plausible context (or background resp. world) knowledge and implausible target words. We also present first tentative results that different types of knowledge may induce different processing patterns.

1. Introduction

Nonmonotonic logics (cf., e.g., [1]) have been devised to overcome the limitations of classical logics with respect to handling rules with exceptions. As is usually the case in Artificial Intelligence, human reasoning has provided the paradigm for reasoning in those logics, and formal systems to axiomatize nonmonotonic logics, like system P [2], have often been considered also as rationality postulates that human reasoning follows. Likewise, the AGM postulates of belief revision [3] have been motivated by considerations of which belief change operations humans would deem to be rational. While there is an increasing interest in evaluating how well the rationality postulates from nonmonotonic reasoning and belief revision are suited to model human reasoning (cf., e.g., [4,5], research that looks more closely and empirically into the relationships between formal logic-based models of reasoning on the one hand, and commonsense human reasoning on the other hand is still rare. To date, we do not know much about how useful our formal models of reasoning are in fact to at least describe how humans reason, whether humans distinguish between different types of knowledge, e.g., such as causal or normative knowledge, or how the plausibility of incoming information is evaluated to produce the information that a human would be willing to accept for revision processes, and what influence perceived (im)plausibility has on the reasoning of humans,

☆ This paper is part of the Virtual special issue on Uncertainty Reasoning, Edited by Robert E. Mercer and Salem Benferhat.

E-mail address: gabriele.kern-isberner@cs.uni-dortmund.de (G. Kern-Isberner).

http://dx.doi.org/10.1016/j.ijar.2017.05.003
0888-613X/

or even on their action decisions. The BDI[1]-model of agent theory [6] neatly distinguishes between the modules in which beliefs, desires, and intentions are processed, and how this leads to actions. While there are numerous interactions between all three modules, it is usually assumed that these interactions occur via interfaces, i.e., beliefs have influences on intentions but the ways in which beliefs are produced are irrelevant.

However, different effects have been observed in psychology: In a seminal study, Stroop [7] showed that when people were asked to name the color in which color words were printed, a mismatch between the color and the meaning of the word (e.g., the word "blue" in red font or vice versa) resulted in slower and more erroneous responses. This is generally taken as evidence that the process of reading is so strongly automatized that it cannot be suppressed even though it is irrelevant for – and interferes with – the actual task of naming the color. In other words, the automaticity of a process is established by demonstrating its interference with performance in an unrelated task for which the process is not required. This casts general doubts on how well the BDI agent model with its clear modular structure fits the way humans base their action decisions on cognitive processes. The empirical insights provided by Stroop and others seem to suggest that it is not only the result of reasoning that is decisive for human action decisions but also that irritations during the process of reasoning itself can influence a human's disposedness for actions.

Plausibility of observations or perceived information (in particular, through reading), and how information is processed play a crucial role in investigating these Stroop-like effects. This creates an interesting connection between knowledge representation in artificial intelligence and language comprehension in psychology. Most modern theories of language comprehension agree that to understand a text, readers need to integrate text information with their knowledge about the world to construct a situation model of what the text is about [8–10]. An important but generally overlooked implication of this assumption is that the process of constructing a situation model must be sensitive to the goodness of fit between incoming information and world knowledge [11]. Therefore, Isberner and Richter [12,13] proposed that knowledge-based plausibility must be routinely monitored during language comprehension. They tested this assumption with a reaction time paradigm in which an assessment of plausibility was irrelevant or even detrimental to performance on the actual experimental task. In three experiments using different experimental tasks, they found interference of task-irrelevant plausibility with task performance, which constitutes evidence that readers cannot actually comprehend information without also assessing its consistency with their plausible beliefs about the world.

In this paper, we elaborate on the relations between formal models of plausible reasoning and belief revision on the one hand, and plausibility monitoring in language comprehension on the other hand. While Isberner and Richter [12,13] are mainly interested in demonstrating Stroop-like effects and measure the impact of plausibility only implicitly, their empirical work provides nevertheless deep insights into the role background knowledge plays for human reasoning. This is particularly visible and impressive in [13] where the authors explicitly distinguish between items involving high or low knowledge. As a main contribution of this paper, we propose formal models of the cognitive processes that may happen in the reader when he or she encounters plausible and implausible information of the kind used by [12,13] in their experiments, and discuss to what extent this model can account for their empirical findings. These formal models allow for plausible reasoning and belief revision while taking commonsense background knowledge explicitly into account, in order to comply with this crucial aspect of Isberner and Richter's work. As a suitable framework, we choose Spohn's ordinal conditional functions, *OCF* [14,15], and the approach of c-representations and c-revisions [16,17] because this combination is able to provide all methods necessary for a framework of plausible, inductive reasoning from background knowledge and iterated belief revision in the spirit of [3,18]. C-representations allow for (inductive) nonmonotonic reasoning of a very high quality, meeting basically all standards which have been proposed for nonmonotonic logics so far (cf. [16,17]). Moreover, c-revisions generalize c-representations, so that we can take advantage of a seamless methodological framework for all reasoning activities that we consider in the experiments. This is important both from a formal and a psychological point of view because such a unifying theory adequately models the close link between uncertain reasoning and belief revision (cf., e.g., [16]) which has also been pointed out in the psychological literature (cf., e.g., [19]). However, we would like to emphasize that the focus here is on the formal reasoning activities themselves (inductive conditional reasoning, plausible reasoning, and iterated belief revision) as potential causes for observed delays. This means that, unlike [19], not the exact methodologies according to which plausible reasoning and belief revision are performed are in the focus of this paper but how suitable formalisms can simulate and explain psychological findings on human thinking in general. Conceivably, other unifying frameworks of plausible reasoning that provide all the mentioned reasoning activities might work as well. It might be an interesting topic of future work to compare different specific formalisms with respect to their adequacy of modeling significant features of human reasoning.

The basic idea is to simulate the test persons' reasoning by first setting up a knowledge base of conditionals which express the relevant beliefs for the situation under consideration in a task within an experiment. Instead of using some kind of plausibility distribution right away, we thereby aim at making plausible beliefs which form the relevant background knowledge that the test person may use for processing the information shown in the tasks as explicit and transparent as possible. Then, an OCF-c-representation is built up which serves as an epistemic model of this background belief base, making the test person ready for responding to the respective task the appertaining (but unrelated) information of which may require a revision process. Our claim is that this revision takes more or less time and needs more or less effort, depending on how compatible the new information is with the contextual epistemic state, and thus may cause delays or

[1] BDI = Beliefs, Desires, Intentions.

Please cite this article in press as: M.-B. Isberner, G. Kern-Isberner, Plausible reasoning and plausibility monitoring in language comprehension, Int. J. Approx. Reason. (2017), http://dx.doi.org/10.1016/j.ijar.2017.05.003