Available online at www.sciencedirect.com

ScienceDirect

Cognitive Systems Research 40 (2016) 144–160

Cognitive Systems RESEARCH

www.elsevier.com/locate/cogsys

Exploration of cognition–affect and Type 1–Type 2 dichotomies in a computational model of decision making

Action editor: Luiz Pessoa

Michael Anton Palkovics [a,*], Martin Takáč [b]

[a] MEi:CogSci Program, Faculty of Philosophy and Education, University of Vienna, Austria
[b] Department of Applied Informatics, Faculty of Mathematics, Physics and Informatics, Comenius University Bratislava, Slovakia

Received 29 September 2015; received in revised form 2 June 2016; accepted 9 June 2016
Available online 15 June 2016

Abstract

This paper studies the role of cognition and affect in decision-making as well as notions of Type 1 and 2 processes and behaviors typically used in dual process theories. In order to demonstrate that there is no 1:1 correspondence between types of observed behavior and internal processes causing them, and that Type 1 and Type 2 processes can be produced by a single system, we implemented a computational model integrating affective and cognitive processing. Our model is based on the model of Marinier, Laird, and Lewis (2009). We modified it by increasing the agent's visual field, adding a GOFAI-style cognitive module (sub-goal management) and expanding the environment by a high-threat tile, to which the agent responds with a hard-wired automatic reaction. This allowed us to generate and observe different types of behavior and study interesting interactions between cognitive and affective control. By comparing our re-implementation to the modified agent, we demonstrated clear cases of Type 1 (fast, automatic) and Type 2 (slow, deliberative) behavior, providing further evidence for the "single-system, two processes" hypothesis.

Keywords: Affective computing; Dual process theory; Decision-making

1. Introduction

Decision-making has traditionally been approached from a normative perspective that developed out of behavioral economics. In their attempts to show that the assumptions of the homo economicus were incorrect, Tversky and Kahneman (e.g. 1974) established that humans use heuristics, and are, more often than not, subject to cognitive biases that lead them to irrational decisions. Because of that, Kahneman considered the use of heuristics and biases inherently "bad". The opposing view, held for instance by Gigerenzer and Selten (2002), is that heuristics and biases are necessary and unavoidable, but not inherently "bad", because behavior adhering to heuristics or biases is not necessarily irrational. Rationality can either be understood as approaching the highest possible outcome – which is the traditional view held by Tversky and Kahneman (1974) and Byrne and Johnson-Laird (2009), or as reaching a sufficiently good outcome – called bounded rationality, as proposed by Simon (1956) and Gigerenzer and Selten (2002).

This distinction lead Kahneman (2011) to interpret his empirical findings in the light of the so called dual process framework (Evans & Stanovich, 2013; Sloman, 1996; Stanovich & West, 2000), wherein judgment, reasoning and decision-making are governed by two (conceptually and neurologically) distinct systems, called System 1 and System 2. While the former is conceived as being fast,

* Corresponding author at: MEi:CogSci Program, Faculty of Philosophy and Education, University of Vienna, Universitätsstraße 7, 1010 Vienna, Austria.

http://dx.doi.org/10.1016/j.cogsys.2016.06.001
1389-0417/

parallel, association-based, and requiring little to no mental effort, the latter is thought to be slow, exact, and requiring conscious effort. In Kahneman's conceptualization, these two systems and their processes are directly related to observable behaviors. Well-known and documented effects such as framing, anchoring, or the availability heuristic are behaviors caused by System 1 and its processes. To overcome them, one has to consciously invoke System 2 and expend mental effort. Note that Kahneman's methodology is behavioral: processing by System 1 or 2 is detected based on externally observable behavior, hence the 1:1 relation between observable behaviors and internal systems that cause them. As will become evident later in this paper, we will adopt a distinction between *internal processes* and *external*, observable *behaviors*, but we will use a different methodology that will allow us to question their 1:1 correspondence – computational modeling.

Computational modeling has been glossed as "Understanding by building" (Pfeifer & Scheier, 2001). Theories have to be specified in enough operationalized detail so that they are amenable to algorithmic description and implementation. The implemented model is then simulated and the simulation results are compared to empirical data. A clear advantage of modeling is that, unlike in people, internal processes of the model are accessible to observation and internal parameters to manipulation/control.

One of the goals of this paper is to study the processes behind observable behaviors traditionally attributed to System 1 and 2 in a computational model of human decision-making behavior. Before we introduce the other goal – to explore how these processes relate to the cognition – affect distinction, we need to refine the notions introduced so far.

The original dual process conception has undergone several modifications. Among other changes, System 1 was re-conceptualized as The Autonomous Set of Systems (TASS), and is not anymore thought to be one single system rather than an agglomeration of systems, each being responsible for a different set of (Type 1) behaviors (Evans & Stanovich, 2013). Evans and Stanovich also suggest that it would probably be more appropriate to speak of Type 1 and Type 2 *processes* since that notion does not imply the necessity of two distinct *systems*. Similarly, Osman (2004) theorized that instead of assuming two systems working on problems as they arise, there is only one system, activating appropriate processes depending on the representation of the problem. Such an approach, also supported by Gigerenzer and Selten (2002), despite only employing one system, can account for the fact why humans tend to first think of responses that are subjects to heuristics and biases – just like originally suggested by Kahneman. In order to generate a fast response, there is no need for sophisticated mental representations of a problem, thus making biased results, or those based on heuristics, earlier available in the process of response-generation. Once the representation is sophisticated enough, Type 2 processes can analyze the problem appropriately.

These alternatives to Kahneman's original model also essentially question the direct correspondence between internal processes and observable behavior. First and foremost, if there is only one system, it must cause both behaviors. The next question that must be raised then (and we will try to answer it in the paper), is whether Type 1 and Type 2 *processes*, with respect to causing *behaviors*, can be still be separated as clearly as originally suggested by Kahneman.

How is all of this related to emotions? Are the dichotomies of Type1/Type2 processing and affective/cognitive processes orthogonal or overlapping? In terms of dual process models, emotions were traditionally considered to be exclusively related to Type 1 processes. They were supposed to provide immediate feedback about situations and were seen to be one of the main sources of our irrationality (within a rationality framework of Kahneman or Johnson-Laird). Stanovich (2009) explicitly links emotions to TASS (formerly System 1) and their processes and attributes behavioral regulation to them. Put simply, imagine a complex situation evoking multiple (parallel or serial) emotions. While Type 2 processes take considerably longer to provide an evaluation of the situation, Type 1 responses are triggered automatically and involuntarily. This allows them to contribute a first coarse evaluation of the situation until Type 2 processes have assessed the situation. However, this does, and should, not automatically mean that the influence of Type 1 responses only pertains to Type 1 processes. Considering Osman's single-system model (2004), where processes are invoked depending on the representation, emotions could potentially affect processing beyond Type 1, simply because of their role in building original representations of a situation. Additionally, if Type 1 processes are responsible for early forms of representations and Type 2 processes modify them, it would be negligent to say that Type 2 behavior is only based on Type 2 processes.

Pessoa (2008, 2010) argues that the whole division to (purely) cognitive and (purely) affective processes might be problematic, as well as the classical conceptualization of certain brain areas as cognitive and affective centers. Drawing on neuroimaging studies, he argues that rich anatomical and functional connectivity between areas traditionally considered as emotional (e.g. amygdala and anterior insula) and PFC areas associated with cognitive control suggests a huge potential for cognitive-emotional interactions. More specifically, given the limited processing capacity in perception and control, stimuli compete for neural resources and emotions effectively modulate this competition by ascribing *value* to different stimuli. In perception, some brain areas (frontal eye field and parietal cortex) seem to contain "priority maps" of salience/relevance of different spatial locations, modulating the attention in favor of salient areas. In executive control, processing is mildly biased in favor of emotional items if their threat content is low, while processing resources are diverted toward a highly threatening item in much more dramatic fashion.