Biologically Inspired Cognitive Architectures (2014) 10, 24–29

Available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/bica

Counterfactual reasoning as a key for explaining adaptive behavior in a changing environment

Jaehyon Paik [a],*, Yunfeng Zhang [b], Peter Pirolli [a]

[a] Palo Alto Research Center (PARC), 3333 Coyote Hill Rd., Palo Alto, CA 94304, United States
[b] Department of Computer and Information Science, University of Oregon, 1202 University of Oregon, Eugene, OR 97403, United States

Received 7 November 2014; accepted 7 November 2014

KEYWORDS
Detecting changes;
Reinforcement learning;
Counterfactual reasoning;
ACT-R cognitive model

Abstract
It is crucial for animals to detect changes in their surrounding environment, and reinforcement learning is one of the well-known processes to explain the change detection behavior. However, reinforcement learning itself cannot fully explain rapid, relatively immediate changes in strategy in response to abrupt environment changes. A previous model employed reinforcement learning and counterfactual reasoning to explain adaptive behavior observed in a changing market simulation environment. In this paper, we used the same model mechanisms to simulate data from two additional tasks that require participants, who played the role of intelligence analysts, to detect the changes of a computer-controlled adversary's tactics based on intelligence evidence and feedback. The results show that our model captures participants' adaptive behavior accurately, which further supports our previous conclusion that counterfactual reasoning is a missing piece for explaining adaptive behavior in a changing environment.

Introduction

It is crucial for animals and humans to detect changes in their surrounding environment, which can happen either gradually or drastically. Animals' survival depends on how well they adapt to these environmental changes, and learn-

* Corresponding author.
E-mail addresses: jpaik@parc.com (J. Paik), zywind@cs.uoregon.edu (Y. Zhang), pirolli@parc.com (P. Pirolli).

http://dx.doi.org/10.1016/j.bica.2014.11.004
2212-683X/© 2014 Published by Elsevier B.V.

ing is perhaps the most powerful ability that animals posses to cope with these changes.

Studies on change detection argue that animals use reinforcement learning (Sutton & Barto, 1998) to detect environmental changes (Behrens, Woolrich, Walton, & Rushworth, 2007; Pearson, Heilbronner, Barack, Hayden, & Platt, 2011). Reinforcement learning is theoretically similar to linear operators, which have been shown, based on optimal foraging theory, to track the changes of a hidden environmental variable with probabilistic observations (McNamara & Houston, 1987). Several behavioral and neuroimaging studies showed that people seem to use reinforcement learning to detect changes, and their performance in the tasks approaches the performance of an ideal observer (Behrens et al., 2007; Nassar, Wilson, Heasly, & Gold, 2010).

Although reinforcement learning plays a key role in detecting changes, it alone cannot fully explain how some animals quickly switch to different task strategies because the error-learning rule for reinforcement learning can only explain a gradual transition between strategies in response to abrupt changes (Pearson et al., 2011). In a previous study (Zhang, Paik, & Pirolli, 2014), we showed that counterfactual reasoning, a cognitive strategy that considers what would happen if an option different from the selected option is carried out, might be the key to explaining rapid detection of environmental changes. In that study, we conducted an experiment and asked participants to try to earn as much money as possible by investing in a virtual market that periodically switches between a bear and a bull state. It was found that participants were able to make nearly optimal decisions about when to invest in the market and when to skip the investment opportunity to avoid likely losses. Furthermore, we found that a model that incorporated reinforcement learning and counterfactual reasoning was able to explain the behavioral data, whereas a model that only implemented reinforcement learning could not. In this study, we follow the same approach and develop an ACT-R model for two additional tasks to provide further evidence to support our hypothesis that counterfactual reasoning is a missing piece for explaining change detection behaviors.

Change detection tasks

The IARPA ICArUS program developed a series of five tasks which, collectively, are called TACTICS. These tasks simulate some common intelligence analysis missions, in which the analysts need to predict and sometimes counteract an opponent's actions. TACTICS is the successor to the ICArUS challenge tasks (Lebiere et al., 2013), and both projects were designed to drive the development of integrated neurocognitive models of sensemaking. In this paper, we report the results of two tasks, Mission 4 and Mission 5, which required participants to detect the changes of a simulated opponent's tactic based on intelligence evidence and feedbacks.

In TACTICS, a participant (Blue defense) operates against a computer agent (Red offense) over a series of trials in an area of interest using intelligence data depicted on a Geographic Information System display as can be seen in Fig. 1. Each trial involves a particular location (indicated by the blue dot and the green circle) in Blue's territory (outlined by the blue lines). Two pieces of intelligence information are given for each trial: (a) OSINT (open source intelligence, *P*), which indicates the probability that Blue will defeat Red (Blue's vulnerability); and (b) IMINT (imagery intelligence, *U*), which indicates the utility/payoff at stake in a showdown (Opportunity). Based on these two INTs, Blue can estimate the Red attack probability and then select either (a) *Divert*, to avoid a possible Red attack or (b) ∼*Divert*, to counter the Red attack. If Blue wins a show-

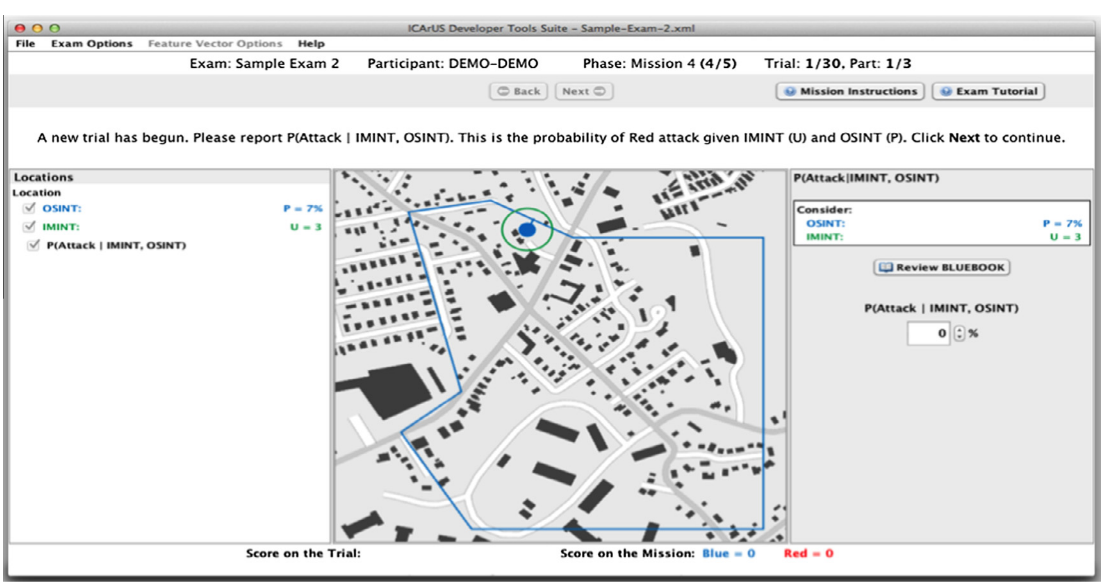

Fig. 1 A screenshot of the display of the TACTICS task. The left column displays the values of intelligence information (OSINT and IMINT) for a particular location on the map (middle column). The middle column provides geospatial information about an interest location, Blue's territory, and the density of buildings. The right column provides an input box that participants can enter the estimated Red attack probability based on two INTs.