Biologically Inspired Cognitive Architectures (2014) 8, 120–131

Available at www.sciencedirect.com
ScienceDirect
journal homepage: www.elsevier.com/locate/bica

RESEARCH ARTICLE

Integrating perception, narrative, premonition and confabulatory continuation

Ladislau Bölöni *

Dept. of EECS, University of Central Florida, 4000 Central Florida Blvd, Orlando, FL 32816, United States

Received 14 February 2014; received in revised form 20 March 2014; accepted 24 March 2014

KEYWORDS
Cognitive architecture;
Integrated cognition

Abstract
Current state-of-the-art AI algorithms outperform humans on several well delimited tasks but have difficulty emulating general human behavior. One of the reasons for this is that human behavior, even in short scenarios, requires the integration of multiple cognitive mechanisms that are deployed simultaneously and are interacting with each other in complex and subtle ways. In this paper we show how a simple scenario of watching television requires at least four different cognitive mechanisms: perception, narrative, premonition and confabulatory continuation. We describe the general requirements of these mechanisms and outline the techniques through which the Xapagy cognitive architecture implements them.

Introduction

Recent years had seen significant progress in artificial intelligence algorithms. For instance, we have software that can beat humans at chess or Jeopardy, and in future years, a number of other specific domains will likely be conquered, that of medical diagnosis being a strong candidate. To achieve the original, broader goals of AI however, it is not sufficient to create a series of narrow systems. Human behavior does not consist of moving from the chess table to a game of Jeopardy, and then to a diagnosis task. Rather, in all the encountered situations, humans deploy a wide range of different cognitive behaviors. Memory, anticipation, emotions, goals, self interest, altruism and humor are applied simultaneously and interwoven in complex and subtle ways.

The field of cognitive architectures have spent the last 30 years working towards Alan Newell's vision of systems that integrate the whole (or at least large parts of) human cognition (Newell, 1994). Certainly, researchers in the 1990s underestimated the difficulty of building a human-equivalent cognitive system. However, the fact that

* Tel.: +1 407 823 2320.
E-mail address: lboloni@eecs.ucf.edu
URL: http://www.cs.ucf.edu/~lboloni

http://dx.doi.org/10.1016/j.bica.2014.03.008
2212-683X/

we now have AI algorithms that perform spectacularly well in a narrow domain, but fail trivially as soon as they leave it, validates the importance of the integrative vision of AI.

Let us consider how a relatively simple scenario of watching television requires the integrated deployment of several cognitive mechanisms. The main character Robby is either a human or an agent controlled by a cognitive architecture:

Robby has a previous experience in reading stories and watching movies involving duels between knights, warriors and gangsters. He is currently watching on the TV a dramatization of Homer's Iliad. On the screen he sees the fight between Hector and Achilles, while the voice-over narration comments on the story. Robby fears that the story will end in the death of one of the characters. Suddenly, the program is interrupted by a commercial. Frustrated, Robby tries to envision a way in which the story will end peacefully.

We will argue that in this simple scenario Robbie deploys at least four different cognitive mechanisms:

- **Perception:** the processing of a real-time, data rich input from a set of sensors.
- **Narrative following:** processing an input in the form of a high level narrative, roughly along the lines of a literary story.
- **Premonition:** the ability to judge that certain events are likely to occur in the immediate future.
- **Confabulatory continuation:** the ability to generate a fictional continuation of the ongoing series of events.

Of course, human behavior includes many other cognitive mechanisms: for instance, attention management and body control at the low level and problem solving, question answering and planning at the high level.

The objective of this paper is to illustrate a possible way in which these four cognitive mechanisms can be integrated in a common architecture. We start by a general, architecture independent description of the requirements of these four cognitive mechanisms. Then, we briefly introduce the Xapagy cognitive architecture and its Xapi language, and discuss the way in which these mechanisms can be implemented and integrated in that system. Finally, we describe the results of some experiments illustrating the proposed integration approach.

Four cognitive mechanisms

Perception

Perception is the cognitive mechanism that integrates the real time input of the sensors into the cognitive model of the agent. In this paper we assume that the sensor and associated processing units had already converted the low level input (pixels and audio waveforms) into a *perception stream* of higher level, symbolic inputs (objects, movement, spatial relationships, utterances).

One of the most important attributes of perception is that it happens in **real time:** the stream follows the temporal succession of events happening in the real world. While we can allow for a delay before the events reach consciousness as shown in the Libet experiments (Libet, Gleason, Wright, & Pearl, 1983), the perception would not include significant delays, nor present events out of order.

Another attribute of perception is its **specificity:** the perception stream is anchored in concrete, physical reality. It does not refer to abstractions: we do not perceive a ''dog-ness'', but see a concrete dog. Another issue is that many of the verbs that we use to describe our perception are not actually present in the perception stream, but are the internally generated results of interpretation or summarization processes. We do not hear ''hammering'', we only hear individual hammer strokes. We do not see ''fighting'': we only see two warriors performing specific movements. While our interpretation of the perception stream can often mislead us (we mistake playing for fighting) the specificity of the perception stream does not guarantee correctness. In some cases (optical illusions, phantom limb phenomena, hallucinations) it is possible that the low level processing makes mistakes in the segmentation of reality, or in assigning attributes to objects.

Finally, compared to the narrative, the perception stream is **dense**, contains a rich stream of details and is **unfilterable**. The latter attribute requires some explanation. A person might choose not to look at the television or to close his eyes. However, once he looks at a scene, he cannot choose which attributes and relations he perceives. For instance, he cannot look at Achilles and Hector without noticing which one is on the left side (spatial relationship), he cannot look at a red object without seeing it its color, or observe a human without observing his hands, legs, head and clothes.

Following a narrative

Humans cannot speak (write) or listen (read) fast enough to convey the full richness of the perception stream. Human communication happens at the level of the *narrative stream*, a list of statements through which salient aspects of the story are extracted and presented in a higher level, summarized and abridged form.

In contrast to the synchronous nature of perception, a narrative can be **asynchronous:** there is no obvious time bind between stories and the physical reality. A narrative might refer to events in the past, predict events in the future, describe events faster or slower than the time span they take in reality, or list events out of order.

A narrative is not bound to the truth as exists in the physical reality, and it can represent physical or even logical impossibilities. We can, for instance, say that Hector is simultaneously to the right and to the left of Achilles, a statement that cannot appear in the perception stream.

Compared to the richness of perception, a narrative stream is **sparse**; it can omit details that are necessarily part of a direct perception. For instance, we can narrate the fight of Hector and Achilles without specifying which one is on the right or the left side. One way in which sparseness can be achieved is by simple omission of details: we do not need to mention the color of Hector's armor. Another way is to use expressions that summarize longer ranges of perception: we say that we hear hammering, instead of