ABSTRACT

In this article we perform some experiments to study how an automatic system learns a set of rules from its interaction with an artificial environment. In particular, we are interested in comparing these capabilities to the skills shown by humans to learn the same rules in similar conditions. We perform this analysis by conducting two experiments. On the one hand, we observe the evolution of the automatic learning system in terms of its performance along time. At the beginning, the system does not know the rules, but it can observe the positive/negative results of its decisions. As its knowledge about the environment becomes more precise, its performance improves. On the other hand, seventy students faced the same artificial environment in the same conditions, though this time the experiment was presented as a game. The objective of the game consists in gaining points, but the rules of the game are not known a priori. So, there is a clear incentive for finding them out. We use these experiments to compare the learning curves of both humans and automatic systems, and we use this information to analyze the similarities/differences between both learning processes. In particular, we are interested in assessing how close the automatic system is from passing the Turing test. [Article copies are available for purchase from InfoSci-on-Demand.com]

Keywords: artificial environment; automatic system; learning processes; Turing test

COMPARING LEARNING METHODS

It is well known that the inspiration for the scientific research is frequently found in the borders between scientific fields. In these borders, researchers from different fields find useful knowledge that belongs to the standard background of a community but is unknown for researchers of other communities. This is the case of Cognitive Informatics (Wang, 2002), which puts different areas in contact: On the one hand, Computer Science and, on the other hand, Neurology, Psychology, and other Sciences related to the human brain. Interesting results about this research area can be found in previous issues of this journal. Among them, we could highlight (Kinsner, 2007; Flax, 2007; Rajlich & Xu, 2007; López, Núñez & Pelayo, 2007; Encina, Hidalgo-Herrero, Rabanal, Rodríguez & Rubio, 2008), as they cover several different aspects of the most recent research in the area.

One of the main concerns of Computer Science in general and Cognitive Informatics
in particular is the relation between artificial and human cognitive processes. This issue has attracted the attention of Artificial Intelligence (AI) researchers for decades. One of the first approaches proposed to study this relation is a criterium that has become classical: The Turing test (Turing, 1950; Saygin, Cicekli, & Akman, 2000). Alan Turing proposed that a machine should be considered intelligent if its behavior is indistinguishable from that of a human being. This claim implicitly assumes that the human mind is a kind of optimal form of intelligence. The reason is not that we assume that the intelligence of humans is perfect, but that we do not have any other model to make this comparison. So, systems under assessment are compared with the only thing we know (almost by definition) that is intelligent.

Let us note that, during its evolution, the AI field has adapted itself to several practical problems. In the optimistic beginning, researchers tried to imitate the global human behavior in a wide sense. After their failure, the AI limited dramatically its scope to that it has currently, which consists in producing intelligent behaviors in very limited and restricted knowledge domains. As the goals of AI have changed along time, we think that those concepts proposed formerly to relate and compare human intelligence with artificial intelligence should be revisited and updated as well. In spite of the fact that achieving an artificial intelligent behavior in the wide sense (that is, with no domain restriction) is a hard task, creating machines whose behavior could be considered intelligent is not difficult when a specific domain is chosen. 

However, it is not clear that a simple preprogrammed system should be considered intelligent, regardless of whether it shows a deep knowledge about a specific topic. Since most current AI systems are constrained to work under specific domains, we postulate that nowadays it is not reasonable to compare the behaviors of humans and machines in a specific context, at least after they have finished their formative period. Let us note that showing a human behavior requires, in particular, to show the capability to learn. Hence, a machine that was endowed with the rules governing a domain but no learning skills could not be considered intelligent if, regardless of its high performance from the beginning, it does not learn in the long term and improves its performance. That is, it should not pass the Turing test. It is worth to point out that learning skills are considered by AI researchers to be one of the cornerstones of any intelligent system. Hence, we postulate that an additional condition should be added to the classical Turing test: The behavior of the machine under test should show a kind of learning skills in the long term (which could require performing a very long test). Let us note that in small domains it is specially clear that an AI system could show a great knowledge of a topic but still be non intelligent. Let us consider that a system is developed to show its knowledge in the topic “is 11 multiplied by 11 equal to 121?”. A system that answers “yes” to the only question fitting into that domain would be as accurate as any informed human. In most programming languages, the effort needed to develop that system is negligible. If we consider other more complex domains, a system could show a total or partial accurate knowledge because it could have been pre-programmed with a set of solutions or a set of rules allowing to infer them. Nevertheless, these systems could still be considered as non intelligent. Let us note that our reason to claim that thing is not related to the classical argument of the Chinese room (Searle, 1980; Searle 1990; Hauser, 1997), whose argument is as follows: A person who does not speak Chinese stays inside a room with a Chinese dictionary. People from the outside write Chinese symbols in a paper and put them into the room. Then, the person inside the room answers by matching that symbol in the dictionary. The argument says that the room does not know Chinese because the person inside does not. The typical reply is that the room does not know Chinese because the person inside does not. The typical reply is that the room, as a whole, does know Chinese. In contrast, our reason is that, even under a pure behaviorist point of view, such a system provided with a full knowledge from scratch
Related Content

Cognitive Processes by using Finite State Machines
Ismael Rodríguez, Manuel Núñez and Fernando Rubio (2009). Novel Approaches in Cognitive Informatics and Natural Intelligence (pp. 52-64).
www.irma-international.org/chapter/cognitive-processes-using-finite-state/27298/

The Cognitive Mechanisms and Formal Models of Consciousness
www.irma-international.org/article/cognitive-mechanisms-formal-models-consciousness/70574/

On Biological Computing, Information and Molecular Networks
Walter Riofrio (2010). Thinking Machines and the Philosophy of Computer Science: Concepts and Principles (pp. 53-64).
www.irma-international.org/chapter/biological-computing-information-molecular-networks/43690/

Testing the Behaviour of Entities in a Cognitive Language
www.irma-international.org/article/testing-behaviour-entities-cognitive-language/1552/

Decision-Making Theoretical Models and Cognitive Biases
www.irma-international.org/chapter/decision-making-theoretical-models-and-cognitive-biases/216763/