Symbol Grounding Problem

Angelo Loula
State University of Feira de Santana, Brazil
State University of Campinas (UNICAMP), Brazil

João Queiroz
State University of Campinas (UNICAMP), Brazil
Federal University of Bahia, Brazil

INTRODUCTION

The topic of representation acquisition, manipulation and use has been a major trend in Artificial Intelligence since its beginning and persists as an important matter in current research. Particularly, due to initial focus on development of symbolic systems, this topic is usually related to research in symbol grounding by artificial intelligent systems. Symbolic systems, as proposed by Newell & Simon (1976), are characterized as a high-level cognition system in which symbols are seen as "[lying] at the root of intelligent action" (Newell and Simon, 1976, p.83). Moreover, they stated the Physical Symbol Systems Hypothesis (PSSH), making the strong claim that "a physical symbol system has the necessary and sufficient means for general intelligent action" (p.87).

This hypothesis, therefore, sets equivalence between symbol systems and intelligent action, in such a way that every intelligent action would be originated in a symbol system and every symbol system is capable of intelligent action. The symbol system described by Newell and Simon (1976) is seen as a computer program capable of manipulating entities called symbols, 'physical patterns' combined in expressions, which can be created, modified or destroyed by syntactic processes. Two main capabilities of symbol systems were said to provide the system with the properties of closure and completeness, and so the system itself could be built upon symbols alone (Newell & Simon, 1976). These capabilities were designation – expressions designate objects – and interpretation – expressions could be processed by the system. The question was, and much of the criticism about symbol systems came from it, how these systems, built upon and manipulating just symbols, could designate something outside its domain.

Symbol systems lack 'intentionality', stated John Searle (1980), in an important essay in which he de-scribed a widely known mental experiment (*Gedankenexperiment*), the 'Chinese Room Argument'. In this experiment, Searle places himself in a room where he is given correlation rules that permits him to determine answers in Chinese to question also in Chinese given to him, although Searle as the interpreter knows no Chinese. To an outside observer (who understands Chinese), the man in this room understands Chinese quite well, even though he is actually manipulating non-interpreted symbols using formal rules. For an outside observer the symbols in the questions and answers do represent something, but for the man in the room the symbols lack intentionality. The man in the room acts like a symbol system, which relies only in symbolic structures manipulation by formal rules. For such systems, the manipulated tokens are not about anything, and so they cannot even be regarded as representations. The only intentionality that can be attributed to these symbols belongs to who ever uses the system, sending inputs that represent something to them and interpreting the output that comes out of the system. (Searle, 1980)

Therefore, intentionality is the important feature missing in symbol systems. The concept of intentionality is of *aboutness,* a "feature of certain mental states by which they are directed at or about objects and states of affairs in the world" (Searle, 1980), as a thought being about a certain place.[1] Searle (1980) points out that a 'program' itself can not achieve intentionality, because programs involve formal relations and intentionality depends on causal relations. Along these lines, Searle leaves a possibility to overcome the limitations of mere programs: 'machines' – physical systems causally connected to the world and having 'causal internal powers' – could reproduce the necessary causality, an approach in the same direction of situated and embodied cognitive science and robotics. It is important to notice that these 'machines' should not be just robots controlled

by a symbol system as described before. If the input does not come from a keyboard and output goes to a monitor, but rather came in from a video camera and then out to motors, it would not make a difference since the symbol system is not aware of this change. And still in this case, the robot would not have intentional states (Searle 1980).

Symbol systems should not depend on formal rules only, if symbols are to represent something to the system. This issue brought in another question, how symbols could be connected to what they represent, or, as stated by Harnad (1990) defining the Symbol Grounding Problem:

"How can the semantic interpretation of a formal symbol system be made intrinsic to the system, rather than just parasitic on the meanings in our heads? How can the meanings of the meaningless symbol tokens, manipulated solely on the basis of their (arbitrary) shapes, be grounded in anything but other meaningless symbols?"

The Symbol Grounding Problem, therefore, reinforces two important matters. First that symbols do not represent anything to a system, at least not what they were said to 'designate'. Only someone operating the system could recognize those symbols as referring to entities outside the system. Second, the symbol system cannot hold its closure in relating symbols only with other symbols; something else should be necessary to establish a connection between symbols and what they represent. An analogy made by Harnad (1990) is with someone who knows no Chinese but tries to learn Chinese from a Chinese/Chinese dictionary. Since terms are defined by using other terms and none of them is known before, the person is kept in a 'dictionary-go-round' without ever understanding those symbols.

The great challenge for Artificial Intelligence researchers then is to connect symbols to what they represent, and also to identify the consequences that the implementation of such connection would make to a symbol system, e.g. much of the descriptions of symbols by means of other symbols would be unnecessary when descriptions through grounding are available. It is important to notice that the grounding process is not just about giving sensors to an artificial system so it would be able to 'see' the world, since it 'trivializes' the symbol grounding problem and ignores the important

issue about how the connection between symbols and objects are established (Harnad, 1990).

BACKGROUND

The symbol grounding problem aroused from the notice that symbol systems manipulated structures that could be associated with things in the world by an observer operating the system, but not by the system itself. The quest for symbol grounding processes is concerned with understanding processes which could enable the connection of these purely symbolic representations with what they represent in fact, which could be directly, or by means of other grounded representations.

This represents a technological challenge as much as a philosophical and scientific one, but there is a strong interrelation between them. From one side there is the concern with the technological design and engineering of symbol grounding processes in artificial systems. On the other side, the grounding process is a process present in natural systems and therefore precedes artificial systems. Theories and models are developed to explain grounding and if consistent and detailed enough may in principle be implemented in artificial systems, which in return correspond to a laboratory for these theories, when their hypothesis are tested and new questions are raised, allowing further refinement and experimentation.

A first proposal for symbol grounding as made by Harnad (1990) in the same paper where he gave a definition for the 'symbol grounding process'. Harnad proposed that symbolic representations should be grounded bottom-up by means of non-symbolic representations: iconic representations – sensory projections of objects – and categorical representations – invariant features of objects. Neural networks were pointed out as a feature learner and discriminator, which could link sensory data with symbolic representations, after been trained to identify the invariant features. This would causally connect symbols and sensory data, but this proposal describes just a tagging system that gives names to sensed objects but does not use this to take actions and interact with its environment. A 'mental theater' is formed as Dennett (1991) defined, where images are projected internally and associated with symbols, but no one is watching it. Besides the symbols and the iconic representations are probably given by the systems operator and the system must learn them

Related Content

A Valid Scheme to Evaluate Fuzzy Definite Integrals by Applying the CADNA Library
Mohammad Ali Fariborzi Araghiand Samad Noeiaghdam (2017). *International Journal of Fuzzy System Applications (pp. 1-20).*
www.irma-international.org/article/a-valid-scheme-to-evaluate-fuzzy-definite-integrals-by-applying-the-cadna-library/189126

On Design and Development of QLIFEX: An Expert System for Social Area
Irena Atanasova (2018). *Intelligent Systems: Concepts, Methodologies, Tools, and Applications (pp. 1287-1303).*
www.irma-international.org/chapter/on-design-and-development-of-qlifex/205834

Multiplier for DSP Application in CPS System
Abhishek Kumar (2021). *Artificial Intelligence Paradigms for Smart Cyber-Physical Systems (pp. 229-250).*
www.irma-international.org/chapter/multiplier-for-dsp-application-in-cps-system/266142

Hybrid System based on Rough Sets and Genetic Algorithms for Medical Data Classifications
Hanaa Ismail Elshazly, Ahmad Taher Azar, Aboul Ella Hassanienand Abeer Mohamed Elkorany (2013). *International Journal of Fuzzy System Applications (pp. 31-46).*
www.irma-international.org/article/hybrid-system-based-rough-sets/101768

Role of Machine and Deep Learning Techniques in Diabetic Retinopathy Detection
Jagan Mohan Nagula, Murugan R.and Tripti Goel (2023). *Multidisciplinary Applications of Deep Learning-Based Artificial Emotional Intelligence (pp. 32-46).*
www.irma-international.org/chapter/role-of-machine-and-deep-learning-techniques-in-diabetic-retinopathy-detection/313342