Natural language processing in artificial intelligence (NLP AI) and natural language processing algorithms relating to grammar as a foreign language

Article · October 2018

CITATION
1

READS
11,513

1 author:

Javier Julian Enriquez
Universitat Politècnica de València
84 PUBLICATIONS **29** CITATIONS

SEE PROFILE

Natural language processing in artificial intelligence (NLP AI) and natural language processing algorithms relating to grammar as a foreign language

Abstract

Universal Artificial Intelligence (UAI) within an algorithm design framework might play a major role in pointing up important technological contribution and innovative development that could be useful for content analysis (semantic) of large databases, for example. In view of that, this paper aims to provide a background for further research based on other authors' analysis such as James Le, Luonge, Maldonado Fonken, Sutskever, or Vinyals, whose results may be the epistemological framework in Applied Linguistics and trends of theoretical language research in comparative form between native and foreign languages.

Key words: Universal Artificial Intelligence (UAI), algorithm design framework, content analysis (semantic), theoretical framework, Applied Linguistics, comparative form.

INTRODUCTION

Hutter (2004) undertakes highly focused research on Universal Artificial Intelligence (UAI), thus pointing out that we need to take closer examination of this approach looking for theories that can guide our search for intelligent algorithms, which measure an agent's ability to achieve goals in a wide range of environments, whose relevant research fields are:

- Computer science (artificial intelligence, machine learning);
- Engineering (information theory, adaptive control);
- Economics (rational agents, game theory);
- Mathematics (statistics, probability);
- Psychology (behaviorism, motivation, incentives),
- Philosophy (reasoning, induction, knowledge).

Progressing in the exposed sense, according to Everitt and Hutter (2018), Universal Artificial Intelligence (UAI) is an increasingly well-studied foundational theory for Artificial Intelligence (AI), based on ancient principles in the philosophy of science and modern developments in information and probability theory, which also offers a deeper appreciation of fundamental problems such as the induction problem and the exploration-exploitation dilemma. UAI, then, as indicated by these authors, is a formal, foundational theory for AI that gives a precise answer to the question of what is the optimal thing to do for essentially any agent acting in essentially any environment. In doing so, the UAI theory is composed of the following four components underpinned by foundational theories that have contributed greatly to scientific progress in many fields, as Everitt and Hutter (2018: 18) point out:

- Framework. Defines agents and environments, and their interaction.
- Learning. The learning part of UAI is based on Solomonoff induction. The general learning ability this affords is the most distinctive feature of UAI.
- Goal. In the simplest formulation, the goal of the agent will be to maximize reward.
- Planning. (Near) perfect planning is achieved with a simple expectimax search.

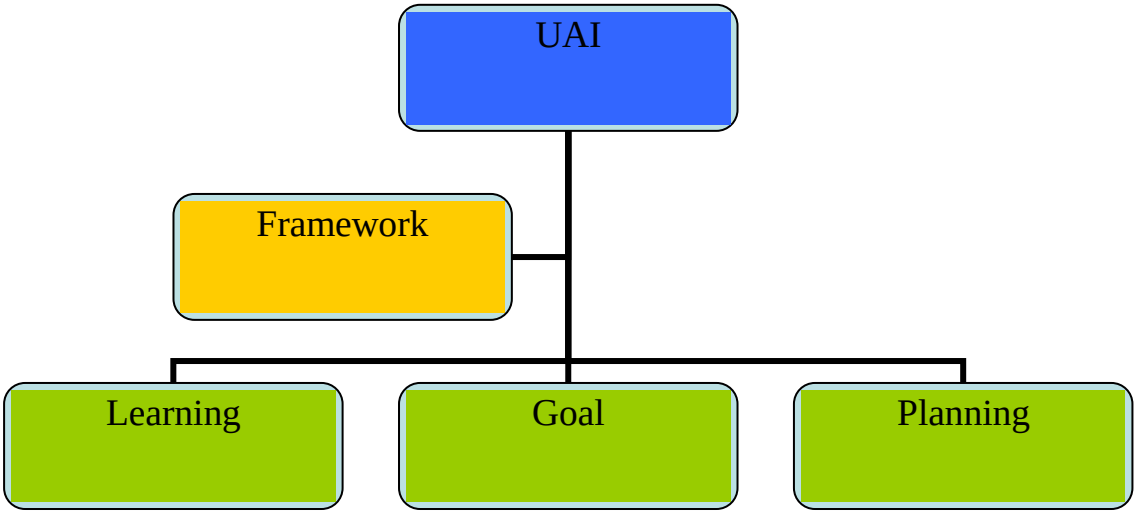

Figure 1. The UAI theory (Adaptation based on Everitt and Hutter's study, 2018)

As a result, this epistemological model may provide us with a framework for further research; particularly, its study might include, for example, the use of the *Master Algorithm*, conceived of as *machine learning* and *new switchboard for Higher Education*. In this respect, we should underline the importance that language, as a system, is acquired from linguistic data presented in natural settings, concerned as well with the path followed by linguistics proper to the configuration of a pragmatic or communicative paradigm from a Blending and Online Learning perspective (Tom Vander Ark, 2015).

Similarly, we believe that we could endeavor to include such epistemological principles within an algorithm design framework from the knowledge, procedure, skill, value, and critical analysis perspective underpinned by the aforementioned foundational theories that would also contribute greatly to scientific progress in many fields by utilizing the *application of holistic strategic management PTT Know How* for future education, for example. Appropriately, several interesting implications may arise from these approaches from our point of view:

1. A breakthrough and groundbreaking piece of research for development of solutions, thus focusing on the study of cultural deficiencies and its impact on leadership, strategic management and governance;
2. Challenges by implementing creative and pioneering solutions for future technology;
3. Rigorous methods and procedures, thus offering a sophisticated insight into instructional as well as educational approaches;
4. Application of technology in instruction, computer-assisted language learning, and learner-centered instruction;
5. A relationship between education, literacy, and identity within a sociocultural context at the cutting edge of discussions about what matters as humankind and civilization learn.

Fittingly, Vinyals et al. (2015) undertake highly focused research on Grammar as a Foreign Language by using a Recurrent Neural Network with attention

mechanism to generate sentence parse trees. Hence, according to Vinyals et al. (2015), syntactic constituency parsing is a fundamental problem in linguistics and natural language processing that has a wide range of applications. As stated, then, by Vinyals et al. (2015), this problem has been the subject of intense research for decades, and as a result, there exist highly accurate domain-specific parsers. These authors, thus, highlight that the computational requirements of traditional parsers are cubic in sentence length, and while linear-time shift-reduce constituency parsers improved in accuracy in recent years, they never matched state-of-the-art. Furthermore, Vinyals et al. (2015) emphasize that standard parsers have been designed with parsing in mind; they underline that the concept of a parse tree is deeply ingrained into these systems, which makes these methods inapplicable to other problems.

In doing so, they show that generic sequence-to-sequence approaches can achieve excellent results on syntactic constituency parsing with relatively little effort or tuning. In addition, while they find the model of Sutskever et al. (2014) not to be particularly data efficient, they discover that the attention model of Bahdanau et al. (2014) is highly data efficient, as it matches the performance of the Berkeley Parser when trained on a small human-annotated parsing dataset. Finally, they show that synthetic datasets with imperfect labels could be highly useful, as their models have substantially outperformed the models that have been used to create their training data. Accordingly, they suspect it is the case due to the different natures of the teacher model and the student model: the student model is likely viewed the teacher's errors as noise, which it is able to ignore. As a result, Vinyals et al. (2015) point out that such an approach is so successful that they obtain a new state-of-the-art result in syntactic constituency parsing with a single attention model, which also means that the model is exceedingly fast. Consequently, these authors conclude that this work showing domain independent models with excellent learning algorithms can match and even outperform domain specific models.