KULeuven

Department of
Computer Science

ADVANCED CAPITA SELECTA ARTIFICIAL INTELLIGENCE (H02A8A)

Contemporary AI topics
Report

Jeroen Craps (r0292642)
Jorik De Waen (r0303087)

Academic year 2016–2017

I. Introduction

February 1996, **Deep Blue**[1] wins the first ever game of Chess as a computer against the world champion at the time Garry Kasparov. March 2016, **AlphaGo** defeats 18-time world champion Lee Sedol in a five-game Go match with a score of 4-1.

Originally the biggest test for an artificial intelligence was the Turing test. The test requires that a human being is unable to distinguish the machine from another human being during a conversation where it has to answer to questions. A limitation of this test is that there is no objective way to measure the progress towards the goals of AI [3].

Computers have been able to perform tasks better than humans like path planning, finding patterns and games. After 20 years of progress computers are now able to defeat the human mind at very complex games, but can it solve the same questions as asked in a 4th grade exam?

To conclude the answer is no, but current scientific research is trying to change this as this is being seen as a key component of any new measurement of artificial intelligence [3].

There are several reasons behind this. It has all the requirements of a test [3]: Accessible, Comprehendible, Measurable and offer a graduated progression for simple everyday things to deeper understanding of subjects. Also to answer these questions a significant improvement in language understanding and the modelling of the world are be required. In this report we will be discussing some of the current methods that are being used to improve the performance of AI on these kind of tests.

II. Overview

To learn the required knowledge artificial intelligence needs a source of data to work with. These data will almost exclusively consist of scientific papers and elementary school books. Currently two separate directions are being pursued, image recognition [1] and lexical analysis [4, 5].

Both streams will be explained in this report, but the focus will be put on the latter. The first will try to retrieve images from scientific papers and correctly label it by using the caption that is included with the table or figure. In the second method the main idea is to form a lexicon to which text will be mapped and converted into logical rules that can be interpreted by the AI. It can use these data to form Knowledge Graphs [2]. While answering questions the keywords of the text will be extracted and filled into the Knowledge graph. The multiple choice questions will be reformed to multiple true or false questions. Every answer will lead to a percentage which will define how true the statement was. The answer with the largest percentage is considered to be the best answer for this question.

III. Image retrieval

In current academic documents figures and tables are key sources of information, i.e. taking a look at a table in a paper can quickly summarize the work that has been done. However, these are not currently used in academic search engines. To better answer any kind of question the use of figure and tables is encouraged.

When there is a focus on scientific papers, it has been shown that **captions** are the key elements to indentify figures and tables. The difference between body text and captions has been proven to relatively easy to detect [1]. The search is done by looking for keywords that are likely to start a caption. Afterwards false positives are removed by applying a filter to them. This filter is focused on a particular format convention. This process of these filters is repeated until no false positives remain.

The second step is to classify every part of the paper as a certain region: caption, body

[1] The Chess playing computer designed by IBM.

text, graphical element or figure text. Caption regions are made from the captions and the following lines of text (if there are any). Poppler's algorithm is used to define body text or figure text. Additional heuristics are applied for improved accuracy. To find the graphical elements, the pdf is directly parsed. Internally PDF's make use of various "operators" that draw elements on the page. The algorithm uses the bounding boxes defined by the operators that the PDF would use to draw. The boxes can be merged if required to form the complete graphical region.

To assign captions/titles to figures or tables clustering is utilized. These clusters would be pruned and additional rules for increased performance. Questions that would match up with elements in a caption can now be linked to a corresponding image which might contain useful information to answer the question.

IV. ...

V. Conclusion

To conclude this report, we can say that the current scientific progress is hopeful and improving rapidly. Texts can be interpreted to reasonable extend and converted to logical rules. Images can be extracted with the corresponding title and caption. With every iteration the scores that the artificial intelligence is able to achieve increases.

References

[1] Christopher Clark and Santosh Divvala. Pdffigures 2.0: Mining figures from research papers. In *Proceedings of the 16th ACM/IEEE-CS on Joint Conference on Digital Libraries*, JCDL '16, pages 143–152, New York, NY, USA, 2016. ACM.

[2] Peter Clark, Niranjan Balasubramanian, Sumithra Bhakthavatsalam, Kevin Humphreys, and Kinkead. Automatic construction of inference-supporting knowledge bases. In *4th Workshop on Automated Knowledge Base Construction (AKBC)*, 2014.

[3] Peter Clark and Oren Etzioni. My computer is an honor student - but how intelligent is it? standardized tests as a measure of ai. *AI Magazine*, 2016.

[4] Jayant Krishnamurthy. Probabilistic models for learning a semantic parser lexicon. In *Proceedings of NAACL-HLT*, pages 606–616, 2016.

[5] Yang Li and Peter Clark. Answering elementary science questions by constructing coherent scenes using background knowledge. In *EMNLP*, pages 2007–2012, 2015.