Solving Mathematical Puzzles: a Deep Reasoning Challenge

Position Paper

Federico Chesani, Paola Mello, and Michela Milano

DISI - Università di Bologna
V.le Risorgimento, 2, 40136 Bologna Italy
{federico.chesani | paola.mello | michela.milano}@unibo.it,

Abstract. This is the era of big-data: high-volume, high-velocity and high-variety information assets are being collected, demanding cost-effective information processing. Analytic techniques primarily based on statistical methods are showing astonishing results, but exhibit also limited reasoning capabilities. On the other end of the spectrum the era of big-reasoning is emerging with next-generation cognitive and autonomous end-to-end solvers. A problem description in terms of text and diagrams is given: problem solvers should automatically understand the problem, identify its components, devise a model, identify a solving technique and find a solution with no human intervention.
We propose a challenge: to design and implement an end-to-end solver for mathematical puzzles able to compete with primary school students. Mathematical puzzles require mathematics to solve them, but also logic, intuition and imagination are essential ingredients, thus calling for an unprecedented integration of many different AI techniques.

Keywords: Deep Reasoning, Artificial Intelligence, Mathematical Puzzles

1 Introduction

This is the era of big-data: high-volume, high-velocity and high-variety information assets that demand cost-effective information processing for enhanced insight and decision-making. Analytic techniques primarily based on statistical methods are showing astonishing results, but exhibit also limited reasoning capabilities. At the same time, we are assisting to a number of achievements in (almost all) Artificial Intelligence (AI) related fields, such as Natural Language Processing (NLP), Image and Video Recognition, Symbolic Reasoning, Neural Networks, Machine Learning and Data Mining, Query Answering (to cite a few). The number of achieved results is indeed impressive, and public debates (also in non-scientific communities) set on fire about the possibilities/risks and the social and ethical roles of AI in general.

Despite the noteworthy results, the human intervention is still essential in a number of steps when tackling complex problems with AI tools. For example, humans intervene in identifying problem components and the hidden knowledge in the description of the problem. Generally speaking, human intervention is still required to enable the transition from the description of the problem to a model and a solution approach. Recently, human intervention in problem solving has been related to Computational Thinking, that helps with these tasks and cognitive skills by defining distinctive problem-solving techniques and general intellectual practices [8].

In a long-term vision, however, next-generation artificial cognitive systems and robots will be autonomous end-to-end solvers that perform the whole problem-solving task starting from its description without any human intervention. Such autonomous intelligent agents will be pro-active and problem-solving driven in finding the right knowledge representation and encoding for modelling the problem and reasoning techniques for solving it.

This long-term vision of Artificial Intelligence results is still far from being achieved nowadays (and in the foreseeable future). Nevertheless, we strongly believe that considering and possibly solving single parts of the problem in a significant case study would be an important step forward, a way of starting a real-world discussion about the applications, implications and ethics of artificial intelligence and a source of inspiration for future artificial cognitive systems.

Since the dawn of Artificial Intelligence, challenges have been considered as a mean to push forward the limit of what computers can do autonomously, and to measure the level of "intelligence" achieved. Problems and issues have been identified in real-world applications and scenarios, but also thanks to contributions of visionary researchers like Turing, and from a specific human activity, i.e. the ability of playing games. Notably, two recent success in AI have been about games, and computers winning over human players (Jeopardy game and Watson[1], Go game and AlphaGO[1]).

In this paper, we argue for the opportunity of considering the *deep reasoning challenge*, where next-generation cognitive and autonomous end-to-end solvers will be able to deal with problems with no human intervention. A problem description in terms of text and diagrams is given: the end-to-end problem solver should automatically understand the problem, identify its components, devise a model, identify a solving technique and find a solution. To better clarify the deep reasoning challenge, we ground it on a (simple to be described) goal:

> **By the middle of the 21st century, (a team of) fully autonomous agent(s) shall win a mathematical puzzle competition against primary school students, winners of the most recent competitions.**

Mathematical puzzles are an integral part of recreational mathematics and in general are played by humans. They have specific rules as multi-player games have, but they do not usually involve competition between two or more players. Instead, one must find a solution that satisfies the given conditions. Puzzles

[1] http://www.ibm.com/smarterplanet/us/en/ibmwatson/

are described by text and images: to solve them, a human player uses understanding and intuition, as well as common sense, simple logical and mathematical/geometry knowledge, causal relations, and many others reasoning-related capabilities. The focus is on *deep reasoning* featuring (a) the extraction of comprehensive knowledge from multi-modal descriptions (texts, pictures and images, sound and speech, gesture, etc.); (b) the ability of determining both the proper model and the corresponding reasoning capability; and (c) the capability of effectively solving the problem (possibly with a feedback from/to the two previous steps). Moreover, being usually played by humans, mathematical puzzles involve small quantity of data.

2 Inside the Deep Reasoning Challenge

In a computer-aided problem solving process, there is always a substantial human intervention that enables the encoding of a problem described by text and diagrams in a model and a solution algorithm. Human intervention is essential for identifying problem components (decision variables, constraints, logical relations, objective functions), and the hidden knowledge in the description of the problem. Usually, a domain expert reads the text and diagrams of the problem, he/she decides the modelling and solving approach based on his/her experience, and frames the model. At this point, an automatic problem solving procedure completes the jobs producing one or many solutions if they exist[2].

A road to achieve deep reasoning could consist of gradually removing the human intervention and let the computer perform the whole task autonomously. Possible steps for automatic problem solving could be the following ones:

1. Read and "understand" relevant text and diagrams of a mathematical puzzle.
2. Identify relevant components/sub-problems the original problem can be decomposed in.
3. Identify a modelling and solving technique(e.g., logic and resolution, constraint satisfaction, planning, mathematical equations, heuristic search).
4. Identify problem components and hidden knowledge, and suitable encodings (possibly guided by the chosen modelling technique).
5. Frame the problem model - represent the original problem and its components by means of an equivalent, machine-understandable model, suitable for reasoning.
6. Solve the problem - by running an automated problem solving procedure.

For the sake of understanding, we introduce these steps as sequential and consecutive. However, it is a matter of discussion if they are in the right order, if there should be one or more iterations, and how much each step influences the other ones.

These steps resemble those proposed in 1945 by the mathematician George Polya in "How to Solve It", and then addressed by Minsky[5]. Moreover, they can

[2] Note that in mathematical puzzles the solution always exists and it is unique.

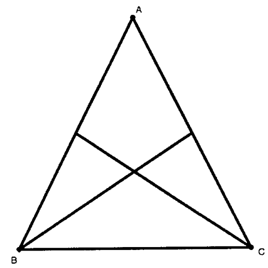

Fig. 1. Triangles

be related to the four key techniques of Computational Thinking: *de-composition* - i.e. breaking down a complex problem or system into smaller, more manageable parts; *pattern recognition* - i.e. looking for similarities among and within problems; *abstraction* - i.e. focusing on the important information only, ignoring irrelevant detail; and *algorithms*, that is substituted here with the more general terms *problem solving* or *inference engine* peculiar of AI applications.

Hence, in our context, Computational Thinking becomes the more general process of Artificial Intelligence Thinking:

> "AI thinking is concerned with frameworks, skill sets, and general tools that are distilled from AI research and practice and are of general interest to everyone, not just AI researchers. Compared against computational thinking, AI thinking goes beyond the logic- and algorithm-based perspectives and should emphasize items such as how to leverage knowledge bases and case bases in problem solving, capture and reason about common sense, enable processing of semantics and contexts, and deal with unstructured data, among others". [9]

To make our challenge more practical, we propose the following few examples.

Example 1: A confused bank teller. A confused bank teller transposed the dollars and cents when he cashed a check for Ms Smith, giving her dollars instead of cents and cents instead of dollars. After buying a newspaper for 50 cents, Ms Smith noticed that she had left exactly three times as much as the original check. What was the amount of the check?

Example 2: Triangles. How many triangles are in Figure 1?

Example 3: Knights and Knaves (from [7]). In a village there are two inhabitants Frank and Jacob. Either Frank is a knight, or Frank is a knave. Either Jacob is a knight, or Jacob is a knave. Knights always say the truth. Knaves always say the false. Frank says that Frank is a knave, or Jacob is a knight. Is Frank a knight or a knave? Is Jacob a knight or a knave?

Apparently there is no common feature underlying these problems, but they certainly require skills as natural language understanding, diagram understanding, basic mathematical knowledge, hidden knowledge discovery, common sense knowledge identification (1 dollar corresponds to 100 cents), problem modelling and solving capabilities. Example 1 can be addressed by means of Constraint Programming techniques, while Example 3 would call for some logic formalism and a corresponding procedure (e.g., propositional logic and resolution). Example 2 instead would call for a tight integration between text understanding and image recognition algorithms.

One cornerstone of the challenge is that these skills should be problem-solving driven. The fact that a bank teller is confused, the name of Ms Smith and the newspaper bought do not help the solution process and should be discarded. The fact that we are talking about the two inhabitants, Jacob and Frank, does not play any important role in the problem solving activity. They could be as well two friends, students or anything else without changing the problem model. This requires a paradigm shift with respect to traditional natural language processing techniques that try to extract contextual knowledge for every word in the phrase.

Notice also that, even if Figure 1 is quite easy to understand (at least from a human viewpoint), this is not generally true. Accompanying pictures could require a deeper understanding, and the same holds for the textual part. From a computer perspective, shape recognition algorithms could easily achieve the correct answer. However, deep reasoning would call here for the right mix between natural language processing and image recognition.

3 Discussion

A number of similar challenges have been proposed. In the context of the Aristo Project, a challenge [2] has been proposed, with the goal of having the computer pass elementary school math and science exams. Natural language comprehension, as well as images and pictures understanding are required as fundamental steps, together with some form of inference and algebraic/mathematical solving techniques. The Euclid project [6] is investigating the fundamental steps to automatically solve geometric problems. These challenges have a strict relation to our proposed notion of deep reasoning, but they mainly focus to provide end-to-end problem solvers stuck to one specific solution method (e.g., mathematical equations, logic), within well-specified domains (such as geometry problems). Deep reasoning should be able to cope with problems described through a number of diverse input media (natural language, still images, diagrams, moving pictures, and sounds) and, more fundamental, it should comprise also the choice of the solution method, as well as the choice of the proper problem model.

Given the breath and the complexity of the challenge, the reader might question if it has the right “size” to foster research advances w.r.t. existing state-of-the-art solutions. An important feature of research challenges is the possibility of facing them is a stepwise fashion, thus providing short-term goals, as well as long-term ones. The proposed challenge offers a number of different intermediate steps: for example, we could approach problems at increasing complexity levels. Moreover, we could focus on problem solvable with a single technique, and then moving towards the use of an ensemble of techniques. Also, in the beginning human collaboration might be envisaged as well, fostering a research direction on how this collaboration could be achieved, as proposed in [4, 3]. A further research question would be on how to measure the advancements: the number of correctly solved puzzles, the quality of the solution, the required time, and the autonomy level, are all dimension that would provide measurable outcomes.

We are aware that the proposed challenge is hard and of difficult solution nowadays, but we strongly believe that even studying and solving only single parts of the problem would be an important step forward, and a source of inspiration for future Artificial Intelligence researches and applications. In addition, on the road to autonomy, it would be interesting to study which level of human intervention and interaction with the machine are needed to effectively collaborate to solve the problem.

References

1. AlphaGo versus Lee Sedol. https://en.wikipedia.org/wiki/AlphaGo_versus_Lee_Sedol, accessed: 2016-09-20
2. Clark, P.: Elementary school science and math tests as a driver for AI: take the Aristo Challenge! In: Bonet, B., Koenig, S. (eds.) Procs. of the 29th AAAI Conf. on Artificial Intelligence, Jan 25-30, 2015, Austin, Texas, USA. pp. 4019–4021. AAAI Press (2015)
3. Gal, Y., Grosz, B.J., Kraus, S., Pfeffer, A., Shieber, S.M.: Agent decision-making in open mixed networks. Artif. Intell. 174(18), 1460–1480 (2010), http://dx.doi.org/10.1016/j.artint.2010.09.002
4. Grosz, B.J., Hunsberger, L., Kraus, S.: Planning and acting together. AI Magazine 20(4), 23–34 (1999), http://www.aaai.org/ojs/index.php/aimagazine/article/view/1476
5. Minsky, M.: Steps toward artificial intelligence. In: Computers and Thought. pp. 406–450. McGraw-Hill (1961)
6. Seo, M.J., Hajishirzi, H., Farhadi, A., Etzioni, O.: Diagram understanding in geometry questions. In: Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence, July 27 -31, 2014, Québec City, Québec, Canada. pp. 2831–2838 (2014)
7. Smullyan, R.: What is the Name of this Book?-The Riddle of Dracula and Other Logical Puzzles. Prentice-Hall (1978)
8. Wing, J.M.: Computational thinking. Commun. ACM 49(3), 33–35 (2006), http://doi.acm.org/10.1145/1118178.1118215
9. Zeng, D.: From computational thinking to AI thinking. IEEE Intelligent Systems 28(6), 2–4 (2013), http://dx.doi.org/10.1109/MIS.2013.141