

An Overview On IBM's Watson

Watson is a cognitive computer system developed by IBM that uses natural language processing to answer questions. The problem it addresses is answering questions with accuracy and in a timely manner. Classically, decision trees have been the cornerstone for handling these types of problems. But as technology progressed, the issue became a matter of scale and the need for more complex-decision making.

Watson was first introduced on the game show Jeopardy! where contestants answer a series of quiz questions, outperforming all other human contestants and winning first place. Its power comes from its ability to handle and process big data in an unprecedented manner. It is even more impressive when you consider that the fact that it synthesizes unstructured information to achieve this task [3]. Today, Watson is used commercially in various sectors, one of them being diagnosing and administering care for patients with cancer.

Watson is revolutionary in today's climate of computing due to its specialization in natural language processing, hypothesis generation and evaluation, and dynamic learning [1]. It uses highly sophisticated constraint satisfaction techniques, among many other AI techniques, to infer the context of what it is being asked.

Language is a challenging obstacle in machine learning due to its imprecise nature. There is plenty of room for a machine to misconstrue or misinterpret the ambiguity present throughout the English language alone. Watson is unique in this regard because it derives an answer based on the context of the question and the from its pool of knowledge. This pool of knowledge, also known as its corpus, consists of collections upon collections of data, ensuring its sheer amount of breadth and depth [1]. An answer is generated by inferring the context of the question in over a hundred different ways, also known as hypothesis generation [3]. It then uses evidence-based sources to rank all these potential answers, synthesizing them and then returns an answer with a confidence percentage [1].

Watson's ability to perform successfully depends on its level of accuracy, scope and scale [1]. Not only does it have to be accurate in inferring the context of the question presented but its domain must be comprehensive enough to return a confident answer. It generates its domain according to the context of the question [2]. By generating the domain in this manner, it tailors the possible response to be as accurate as possible. The way it contextualizes information and infers its meanings is essentially handled by a dictionary [2]. Big data is extremely important for Watson to effectively operate. Without a diverse and comprehensive domain and dictionary, it may not return an accurate answer.

With regards to the final project, it may be worthwhile to investigate the performance of a constraint satisfaction algorithm when applied to various logic problems. One could compare the runtimes of how quickly a problem is solved based on the size of the problem. Aspects to consider would be the number of constraints, domains generated, along with dictionary size.

References

- [1] High, Rob. “The Era of Cognitive Systems: An Inside Look at IBM Watson and How it Works,” *IBM Corp.*, 2012.
- [2] Chen, Y. and Argentinis, E. and Weber, G. “IBM Watson: How Cognitive Computing Can Be Applied to Big Data Challenges in Life Sciences Research,” *Clinical Therapeutics*, vol. 38, no. 4, pp. 688-701, 2016.
- [3] Ferrucci, D. and Lavas, A. and Bagchi, S. and Gondek, D. and Mueller, E.T. “Watson: Beyond Jeopardy!,” *Artificial Intelligence*, vol. 199-200, pp. 93-105, 2013.