Andrew Loutfi

Doctor Kent D. Lee

CS 381

01/23/16

Reflection

Watson is a cognitive cloud based computing solution developed by IBM. The technology gained notoriety in 2011 for participating in the popular question-answer show *Jepordy!* where it won first place against two former show winners. Since then, the company has allocated more than one billion dollars into expanding and developing the program.

There are high expectations for Watson. The current CEO of IBM Virginia "Ginni" Rometty expects the program will generate ten billion dollars in revenue annually within ten years. More importantly, the Watson project is a catalyst through which cognitive computing can be advanced.

The Oxford dictionary defines cognition as "The process of acquiring knowledge and understanding through thought, experience, and the senses." This process should sound familiar, as it is how we as human beings derive answers to the same kinds of problems. Essentially, cognition is a way to learn. It is a process through which we answer questions where the response demanded are deeper than no or yes. IBM emulates this Cognitive Psychology to an extent with Watson. An engineer at IBM said the goal is to ultimately build Watson to learn and think as human like as possible.

The first thing that happens when Watson is learning about a new field is that it is taught the jargon and thought process of the specific field. For example, if Watson is being trained in medicine it will not only learn what oncology is, it will learn about cancer, its specific forms, and will also learn about each cancer's symptoms and the diseases affiliated with said cancer. Watson will learn the symptoms, causes, and complications of these diseases as well. Ultimately, Watson "reads" and "learns" information through algorithms developed for natural language processing.

Natural language processing (NLP) is a way to enable machines to derive meaning from human language. NLP is closely related to a discipline in computer science that is known as machine learning. Cognitive computing represents a branch of this discipline as well.

Natural language processing represents technologies that have been around for many years now. It is found in many digital assistant applications. However, the natural language processing algorithms used on Watson indicate a significant leap from being able to grab a friend's iPhone and telling Siri "text mom Hello World," thus directing the phone to send a text to that friends mother saying, "Hello World." Consider this clue taken from *Jepordy!* (Ferrucci):

> **Category:** Diplomatic Relations
> **Clue:** Of the four countries in the world that the United States does not have diplomatic relations with, the one that's farthest north.

To answer this clue, Watson will utilize slot grammar to build a parse tree of the clue. Slot grammar is a natural language processing technique where the computer is taught the rules of a specified language and then reads documents in the specified language to extract information from it. In the question above for example, the machine would first recognize that the language of the clue as English and then assign value to each word in the sentence using both the words definition and the rules of grammar conducive to the language.

This particular question has an inner and outer part to it. It requires one know the names countries that do not have diplomatic relations with the United States in addition their location. Watson would recognize this through slot grammar and use this knowledge to its advantage when deriving a hypothesis. It will query the inner subclue against information it has ingested first to determine possible hypotheses to the question. It will use the results of this analysis to generate several candidates to what the answer might be. Watson will use these candidates in conjunction with the category and the rest of the question to generate a more accurate hypothesis.

Inner subclue: The four countries in the world that the United States does not have diplomatic relations with
Answer to inner subclue: Bhutan, Cuba, Iran, North Korea

Having identified four candidates, Watson would then query the rest of the clue against itself to determine which of them is most likely to be correct. In this case, Watson has injected the candidates into the subclue.

Outer subclue: Of Bhutan, Cuba, Iran, and North Korea, the one that's farthest north.
Answer to outer subclue: North Korea

Watson will return the hypothesis that the country farthest north of the four countries that do not have diplomatic relations is North Korea, as well as a confidence metric displaying how confident it is that this is the correct answer. Engineers working on Watson generally consider anything above 80% to be acceptable. Unfortunately, the confidence metric for this query was unavailable, but Watson's answer to this question is correct.

After Watson has extracted information, it is reviewed by engineers and field experts. They will look for data that is erroneous, poorly regarded, or out of date. This process is known as curating data. After being curated, data is organized into metadata and data structures. This process is known as ingestion. Ingestion helps Watson responds more efficiently to complex questions. This is due to the fact that Watson will know exactly where the required data is after it has been mapped and organized.

While at IBM, I shadowed Nathan Lee of the Clinical Trial Matching ingestions team. Nathan enjoys his work a lot. He is very smart and was happy to answer any questions that I had for him. I enjoyed shadowing Nathan very much and I hope Nathan enjoyed having me shadow him!

After ingesting data, Watson is ready to run queries. It still needs to be trained how to efficiently access this data and interpret queries though. To do this, engineers use

machine learning techniques. One such method is to have the machine learn through established ground truths. To accomplish this, engineers will upload training sets in the form of question answer pairs to Watson. These pairs represent many of the basic truths in the field that Watson is studying.

After training sets are uploaded, Watson will use them to make deductions and reason when answering questions that it does not explicitly know. In addition to teaching ground truths, these training sets simultaneously teach Watson linguistic patterns. Consider the example of this question answer pair derived from Kieswetterian logic:

Question: Are all men mortal?
Answer: All men are mortal.

Imagine that this question answer pair has been uploaded to Watson as part of a training set. Imagine then that while extracting new information, Watson learns of an entity, named Caius. The natural language processing algorithms, such as the ones pertaining to slot grammar, will help Watson understand that Caius is a man. Then it will recall the ground truth it was taught: all men are mortal. Through this reasoning, Watson will deduce that Caius is also mortal.

Watson utilizes other principles derived from machine learning as well. One such principle is deep learning. Deep learning is branch of machine learning based algorithms that utilize large, unorganized sets of data to train the machine to recognize patterns and ultimately improve hypothesis generation.

One way Watson utilizes deep learning is through the use of its previous outputs. Routinely, software engineers will review Watson's interactions with humans as well as previously generated hypothesis. Data regarding these instances is then fed back into Watson, helping it to interpret information better. Through this large amount of data being fed to it, Watson will discover patterns in its behavior and is updated accordingly.

In addition to helping improve output and interaction, deep learning also helps Watson refine its other algorithms as well. In natural language processing for example deep learning will help Watson refine algorithms that have to do with discrepancies such as double entendres and human-made syntax or spelling errors. In establishing ground truth, deep learning will help Watson with reasoning. This is simply because it will have analyzed previously generated hypothesis and is unlikely to make the same deduction if it was deemed wrong before.

It is important to discern that Watson is not artificial intelligence. It may be faster than humans in some aspects, but it requires human intervention in almost all of its actions and is rudimentary in terms of traditional intelligence. That being said, Watson and cognitive computing represent many diverse problems related to artificial intelligence. The solutions that we discover to these problems will certainly aid with the advancement in many of the fields that will have a direct effect on artificial intelligence.

The field of artificial intelligence is complex and multifold. It involves various discipline of computer science such as machine learning, as well as disciplines such as behavioral psychology and neuroscience. It is a new, ever-changing field that indicates the coming of a new era in technology: machines that are to think, reason and function autonomously. The implications of such powerful technologies are enormous. Not only does the idea of artificial intelligence challenge many preconceived notions, but its arrival

will also hail a great paradigmatic shift in many different industries, sciences as well as the ideas of consciousness and life itself.

We must be cautious in the ways which we advance AI research. It is a technology that has the potential for virtually any application. We must also remember that there are applications that could not only have undesirable consequences, but also threaten the existence of the human race itself.

In early 2015, a letter highlighting the research goals regarding artificial intelligence was circulated throughout the scientific community. It includes the signatures of hundreds of software engineers from companies like IBM, Google, Apple, and Microsoft. There are many signatures from notable figures in the artificial intelligence community as well, including names like Stephen Hawking, Steve Wozniak, and Elon Musk.

The letter reasons that, "since everything that civilization has to offer is a product of human intelligence; we cannot predict what we might achieve when this (human) intelligence is magnified by the tools AI may provide, but the eradication of disease and poverty are not unfathomable. Because of the great potential of AI, it is important to research how to reap its benefits while avoiding potential pitfalls. (FLI)" The letter goes on to describe many concerns that will need to be addressed regarding artificial intelligence. These concerns relate to ideas such as whether or not to ban the development of artificially intelligent weapons to worst-case scenario handling of autonomous vehicles as well as legal liability. Above all, the letter contends that we must be mindful in our development of artificial intelligence.

The technology that is being created through Watson has great potential. It has, of course, already been utilized for many wealth generation and corporate oriented applications. These applications range from analysis regarding customers to analysis regarding an enterprise or field of business. Other applications for Watson range from talking children's toys to recommending treatments for various forms of cancer to physicians. There are many Watson applications currently being developed by IBM for research and health care.

The Oxford dictionary defines big data as "extremely large data sets that may be analyzed computationally to reveal patterns, trends, and associations, especially relating to human behavior and interactions." Over the past few years IBM has devoted much of its attention to projects that pertain to the management and analysis of big data. It comes as no surprise therefore that Watson has many applications that pertain to the use of big data. One of these applications is Watson's Discovery Advisor. This application uses the many algorithms developed using big data and deep learning as well as all the data Watson has ingested to analyze queries. These queries are the kind of queries where the answer could take researches days, even weeks, to figure out. Watson Discovery Advisor will then attempt to draw patterns between the query and the data it has ingested, returning the results almost instantly.

Another application that has recently become offered by IBM is Watson for Clinical Trial Matching (CTM). Clinical trials are the medium through which new treatments are tested and refined before becoming clinical practice. Clinical Trial Matching is the process of matching individuals who have volunteered to be part of a clinical trial to a specific clinical trial. If a patient or healthy volunteer meets the criteria to take part in a trial, then they have the option of undergoing a treatment that is potentially better than modern treatments as well as advancing research in the respective branch of medicine.

In the past, Clinical Trial Matching has been a very manual process. The work done entails a clinician going into the volunteer's medical records and comparing eligibility from the clinical trial with the record. Medical records and eligibility criteria are very lengthy, and there are potentially thousands of clinical trials to analyze. In addition, new trials are constantly becoming available, meaning clinicians might not have access to a complete list of all available trials. These medical records and clinical trials are examples of big data that is relatively unstructured.

When looking for patient volunteers there are many more factors to consider than when looking for healthy volunteers. The head of Oncology at the Mayo Clinic, Dr. Steven Alberts said, "Matching a patient to a clinical trial can be one of the most challenging parts of conducting clinical research." When dealing with patients, time also becomes a much greater issue. The manual extraction of information from hundreds of clinical trials as well as medical records is time consuming, meaning that many patients will not even have an opportunity to take part in a trial. With this in mind, as well as the demand for a constantly up-to-date list of clinical trials, it is logical that this process should be automated. This is the focus of the Watson for Clinical Trial Matching project.

First, a patient's medical record is queried against ingested clinical trials pertaining to the given cancer. Watson will then ask the clinician more questions based on the cancer as well as the patient. As work on Watson progresses, this will become more automated as well. Natural language processing will be used to extract this information if it is included in the patient's record or the clinical trial. After Watson has all of the information that it needs, it will list all of the trials that the patient is eligible to participate in. These trials are then listed in order based on a confidence metric.

Watson for CTM is being trained for breast, lung, colon, and rectal cancer. There will be more offerings as the project advances. Automating the Clinical Trial Matching process means that patients who are extremely sick will have access to the newest treatment options. It also means that the Clinical Trial process itself has the potential to become more streamlined. Trials that show significant promise would be able to be refined into clinical practices much sooner.

The ways in which Watson would contribute to the process, as well as medicine itself, are immense. It follows, naturally, that the undertaking of such a project is also immense. Therefore, engineers working on the Clinical Trial Matching project at IBM must be able to work and design software in a way that is efficient, robust, and adaptable.

Agile software development is a relatively new approach to software engineering. Despite this, it is being utilized field-wide. Agile software development is designed to promote teamwork, collaboration, and process adaptability for projects. It is intended for use on complex projects with dynamic, non-linear features, where many factors are often hard to accurately determine in early stages. There are many different methods that apply agile software development. Currently, the IBM Watson CTM project is utilizing such a method.

Agile development begins when a prototype of a perspective software product is shown to a customer. The customer then tells the company, in this case IBM, what the prototype has that they want, if anything is broken in the prototype, and what they would like to see in a fully developed product. Transparency is crucial to the optimization of agile software development, so managers and lead engineers of a project are involved when showing a product to a customer. After this, the engineers divide these newfound tasks

amongst themselves, working in small teams to finish each task by a set date.

The period of time in which these projects are worked on is known as a sprint. These sprints vary in length; decreasing the farther a project has come. **While I was shadowing the CTM project, the team received word that they had been given the go-ahead to move into production**, meaning that soon the team would be moving into shorter sprints. Feelings on the CTM team at IBM are mixed in regards to the length of sprints. Some engineers raise the concern that shorter sprints mean that they will not have enough time to push new updates. It should be noted that there are many legalities that must be observed when dealing with clinical trials and the confidentiality of individual patients. This means that time that is usually meant for coding is devoted to these overhead issues.

At the end of a sprint, the new updated software is shown to the perspective customer. From here, development becomes iterative. The perspective customer will make comments on the software, and another sprint will begin.

The benefits of using an agile software development derived approach to programming are immense. It means that large projects, such as Watson based projects, can be broken down into smaller, considerably more manageable projects. It means that projects become graph-able, and are therefore easier to organize and implement. Finally, agile software development makes more efficient use of time than traditional, linear, programming approaches. All of these factors mean that a project can be updated or changed faster and is designed from the ground up to do so efficiently and quickly, thus the word agile.

Almost all of the work done by these teams in agile based software development is independent. That being said, the components one team is working on could have any sort of effect on another team's components. This dependency means that communication is crucial to a projects success.

Problems faced by software teams will often require complex solutions. For this reason, much emphasis is placed on the formation of ideas and concepts conducive to a project. Software engineers will often hold meetings meet to discuss these new ideas, where they will talk about ideas themselves or even the ideation process. While shadowing the CTM project, I was able to sit in on one of these ideation meetings. This particular meeting was the first of the New Year, so the team outlined goals for ideation sessions, some ideas they had, and even discussed the ways each of them ideate best.

There are many other kinds of meetings held during software development as well. Some of these meetings are scheduled, other meetings are impromptu. These meetings that are impromptu often regard implementation issues, such as defects, that have been found. During these meetings issues are targeted and solutions are discussed. Sometimes these issues can be solved during the meeting, other times they are more complex and demand more time to solve. These issue solving meetings can be between team members, as well as between members of different teams. This style of impromptu meetings represent the majority of the meetings I sat in on during my shadowing opportunity.

Meetings that are scheduled often serve as guidance and planning sessions. Although ideation sessions and other sort of brain storming sessions fall under this category, their purpose is more to generate ideas than plans. Teams will often schedule general meetings amongst themselves to see where everyone is at in their work.

One kind of scheduled meeting is a scrum meeting, which occurs weekly. In scrum

style meetings team members go around and say what have done since the last scrum meeting, what they will be doing, and what could impede them from completing their task. In meetings between sprints, scrum meetings will also be used for planning work during the next sprint.

Communication within software engineering initiatives is crucial to the success of a project. This communication comes in many forms. Software development teams also utilize real time collaboration in order to communicate quickly. One solution from a coding perspective is to utilize a development environment that supports a Real Time Code interface. RTC (real time code) support allows for anyone working on the project to see what everybody else is working on. This is immensely helpful from development and implementation points of view, as it allows for bugs to be found easier, and allows everybody to be on the same page. It is RTC--along with an inter-department IM system--that is utilized by the team to aid with real time communication. It is these features, as well as agile software development and a team of extremely smart people that make development efficient, robust, and adaptable at Watson for Clinical Trial Matching.

My shadowing of IBM Watson for Clinical Trial Management was an amazing experience. Everyone whom I met and conversed with was happy to answer my questions and was very well informed. Not only did I learn about the new, powerful ways we are utilizing Watson, but also learned about how these large complex programs are worked on in a software engineering setting first hand. It was quite inspiring and motivating to see just how much energy and effort goes into peoples ideas on how to further technology. Most important of all though, I have learned that I would like to further immerse myself in the field of software engineering.

Works Cited

Ferrucci, David, et al. *The AI Behind Watson — The Technical Article* (n.d.): n. pag. 2010. Web. 20 Jan. 2016.

Future of Life Institute. "Research Priorities for Robust and Beneficial Artificial Intelligence." *FLI Future of Life Institute*. N.p., n.d. Web. 20 Jan. 2016. <http://futureoflife.org/ai-open-letter/>.

Grenning, James, et al. "Manifesto for Agile Software Development." *Manifesto for Agile Software Development*. N.p., n.d. Web. 20 Jan. 2016. <http://agilemanifesto.org/>.

McCord, et al. "Deep Parsing in Watson." (n.d.): n. pag. July 2012. Web. 20 Jan. 2016. <http://brenocon.com/watson_special_issue/03%20Deep%20parsing.pdf>.

Ng, Andrew. "Machine Learning." *Coursera*. N.p., n.d. Web. 20 Jan. 2016. <https://www.coursera.org/learn/machine-learning>.

"Watson for Clinical Trial Matching." *IBM*. N.p., n.d. Web. 20 Jan. 2016. <http://www.ibm.com/smarterplanet/us/en/ibmwatson/clinical-trial-matching.html>.

"Watson How It Works - Deep Dive (Narrative Version)." *Watson How It Works - Deep Dive (Narrative Version)*. IBM, n.d. Web. 20 Jan. 2016. <https://vimeo.com/ibmwatson/review/116549372/c7a8508ef5>.

All other information was taken from various IBM Watson for Clinical Trial Matching team members.