Kryat Vladislav

Student of Don State Technical University

IBM Watson supercomputer

If we remember the most intelligent machines created by man, then, definitely, Watson will take first place among them.

The story of Watson supercomputer began in 2006, when David Ferrucci, senior manager of IBM's semantic analysis, engaged in testing one of the most powerful supercomputers in the company, which occupies one of the top lines of the 500 most productive machines in the world. Ferrucci decided to try, how well the machine will deal with the tasks, formulated on "natural language". He suggested the machine to answer 500 questions from the american TV show Jeopardy. The results were disastrous: in comparison with the real players, the machine is not enough quick to respond, and the number of correct responses did not exceed 15%. Ferrucci interested in the causes of such behavior of supercomputer. As a result, in 2007, he was able to convince the leadership of IBM to give him a team of 15 people and 3-5 years for creation an efficient automated system, capable to answer on unformalized questions. Such system would be useful for all kinds of call centers, help systems and any other services which serving clients. IBM has already had a successful experience in creating machines, that can compete with human intelligence - a supercomputer Deep Blue, which won the world chess champion Garry Kasparov in 1997.

Machine, that recognizes a natural language, faces with difficult and complex rules of language with numerous distortions and deviations from that rules. The biggest challenge lies in the fact, that people communicate within their cultural and social context. Spoken language full of hints and allusions, references to general facts, concepts and phenomena certain for a particular social environment. For efficient processing of such information uses statistical algorithms, that allow you to link different concepts with each other by analyzing a variety of documents. Ferrucci team loads into memory of IBM Watson millions of documents of all kinds - textbooks, encyclopedias, reference books, fiction and religious literature. To analyze the questions Watson uses more than a hundred algorithms, offering hundreds of possible solutions. Then other algorithms evaluate the accuracy of the potential answers, eliminating the impossible and improbable.

After four years of study Watson made his debut in the American television show Jeopardy. Watson played against the champions of America. The result was stunning for humanity, Watson won with clear advantage.

What is Watson? At the time of debut it was 750 servers POWER7, united to the system with a volume of 15 terabytes of memory. Unique software is more important than the hardware. This software has been called DeepQA. It includes the integration of natural

language processing machine NLP, information retrieval IR, machine learning ML, multiple-parallel computing and knowledge representation and reasoning KP & R. The unique intellectual core of Watson includes 40 key programs. Watson's software written in various languages, including Java, C++ and Prolog and uses the Apache Hadoop.

Watson's hardware architecture allows distributed computing, i.e simultaneously work with multiple tasks in parallel mode. In addition, it allows to run programs at many devices connected to the head cluster of servers, implemented in the cloud. Watson has operating speed of supercomputers, which allows to solve the most complex problems in real time and thus maintain operative dialogue with the person without delay. Watson is able to work with super big data, ie, structured and unstructured data, including not only the text and tables, but also images, video and audio streams. All these disparate data formats integrate into a common information base and streams. The system has a powerful intellectual apparatus of multidimensional processing of any data. This apparatus uses not only statistical methods but algorithms, based on pattern recognition theory and other branches of mathematics. Watson implements the principle of machine learning. The results of computations are constantly compare with the reference answers, and, if there are errors, the system in the automatic mode rebuilds and improves itself so as to obtain more correct answers each subsequent iteration. The system is able to perceive a natural language, as in the form of a series of sounds or text, on the input and on the output. In this sense Watson is the most advanced expert system in the world, which able to engage in dialogue with the person in ordinary language.

After winning the show Watson went to study at Memorial Sloan Kettering Cancer Center. Here, the developers team with leading US oncologists trained Watson to diagnose the cancer. It was taught to help doctors diagnose patients and recommend appropriate treatment courses for each case. After careful study Watson started practical advice. Not so long ago, IBM and Memorial Center told the public that today Watson significantly increased the accuracy of diagnosis of cancer and can give advice to physicians of ordinary skill, that significantly increase the effectiveness of treatment of cancer patients by these doctors.

Simultaneously, Watson began to help in the public service. It is now known that the system is ready for use in the NSA and the FBI.

Also Watson proved himself as an perfect actuary, calculator of insurance risks and investment adviser. His recommendations for the formation of a portfolio of shares were much more profitable than the recommendations of the control group of investment advisers from Wall Street.

For the development of creative abilities of Watson its creators chose the culinary arts. This is very convenient testing ground: cooking is a very "human" intuitive process, poorly responsive for algorithmization and standardization. Millions of recipes, collected on the Internet, served as initial data for Watson. They were passed through a proven

natural language processing algorithms, that were used before. It has been extracted information from Wikipedia about the typical ingredients and reception processing specific to the different cuisines of the world. Finally, Watson received a thorough knowledge of the chemistry and physiology of human perception of taste and smell. New recipes are generated, based on the existing ones, using a genetic algorithm. As a suitability function, the values of novelty, pleasantness and compatibility were used.

As a result it was created an application in which you can specify a set of products, national style and variety of dishes, and then Watson gives a set of recipes, that can be sorted by degree of novelty, pleasantness and compatibility. In addition to individual dishes, Watson is able to create entire menus, ensuring diversity and the right combination of food through the use of thematic modeling. This method of constructing models of the collection of text documents, which devides the collection on themes and determines to what subject refers each document. Watson uses this model to the recipes. As keywords are the individual ingredients, as documents are recipes themselves.

All this has allowed to announce the start of the third stage in the life of Watson. He moved into the cloud and open to third-party application developers. Now Watson has an API, by which third-party programs and services may use the power of Watson cognitive computing.

The development teams can write their applications for specific subject areas, getting knowledge on Watson architecture and software solutions. Yet IBM limits the number of third-party developers. Those who want their applications to take advantages of Watson, should first contact the company to get approval and access to the cloud, and only then start working.

Cognitive computing in coming years will be in demand in various fields. First of all, they will affect those office areas where work is routine and it is very, very expensive. Watson offers great promise in law and consulting. Here, main part of the time is spent not on the search for a solution but on sorting vast amounts of information.

Watson also be able to help in the sphere of health care services. Undoubtedly, this is only the first steps of cognitive-computing business.

References:

1. http://habrahabr.ru/company/ibm/blog/202930/
2. http://medicena.ru/blogpost/v-ssha-nachal-rabotat-robot-vrach-ibm-watson/
3. http://www.ibm.com/developerworks/library/os-ind-watson/
4. http://old.computerra.ru/terralab/platform/541734/
5. http://www.ibm.com/smarterplanet/us/en/ibmwatson/what-is-watson.html