# Computational Medical Diagnosis

John Foley

10 April 2014

**Abstract**

Computers are increasingly being used in the medical industry. They are used to provide structure and efficiency to record management and patient organization to both hospitals and medical labs, but the intelligence and efficiency of computers is also advanced enough to be able to analyze massive quantities of information and produce intelligent, reasonable conclusions. Of course in the context of diagnosis and medication, the data can consist of human medical records and genetic sequencing of genetic material and conclusions be diagnoses and care planning. Analysis of this information is currently conducted by one or more highly trained doctors or medical experts. The cost of these trained professionals is relatively high compared to the time it takes to perform the analysis per patient. Using computers, with advanced algorithm and high level cognitive science research, we are able to replicate the team of specialists and come to the same results in significantly less time and hard work, freeing those professionals to be able to perform harder problems. Another advantage to have expert diagnostic systems is having available oversight for work done by human doctors. Research has shown a connection to linguistic expression and diagnostic correctness. A computer that is programmed to diagnose can evaluate a doctor's work and catch costly misdiagnoses before they can occur.

# 1 Introduction

Computational intelligence and capability is quickly becoming noticeable in today's society as a helpful tool. People are beginning to experiment uses in a range of industries and corners of society. We have Siri on our Smartphones, robotics taking over as the manufacturing work horse, and computer diagnostic tools becoming smart enough to help us fix the very computer that they live on. Programs are more and more capable of reading in available information, digesting it and coming to a helpful conclusion. There are network diagnostic tools to figure out why your computer cannot connect to the internet, and antivirus programs that analyze every inch of memory to snoop out malware, spyware, or even simple corruption in the file structure. They do this multiple multitudes of times faster than their human predecessors and with significantly increased efficiency. Computer programs don't become bored or fatigued- the only problem they run into is that they are as powerful and effective as the programmer that designed their algorithm.

Diagnostic tools are at the point where problems such as network connectivity and search and find algorithms are child's play. They are faster and meaner than ever before and history, and AI researches are starting to experiment with ways of applying these tools to human diagnostics. A program can read in massive amounts of information and make sense of it, which is essentially what human doctors are paid and trained for years to do. There are recognizable patterns in the data that result in diagnoses. The patterns are distinguishable through medial histories and current symptoms, both are sources of information that computers can read from quickly and efficiently. The advantage of computer diagnostics are more readily seen when the source of information is so massive and complex that it would take a human specialist a long and costly time to find the pattern and resulting conclusion (in context of the medical field, a medication or treatment plan). An example would be in the form of DNA sequencing and finding deeply hidden and complex diseases or cancers.

Using computers to aid in the diagnostic and treatment in the human medical field will increase productivity and effectiveness overall. Of course programs and algorithms cannot be trusted completely for some time- the wrong diagnosis can be come to whether it's a human or a computer. The speed and sheer scaling power of using computers is an advantage that the medical field needs to take advantage of. Power artificial intelligence can aid in medical devices such an ultrasounds, aid in diagnoses of complex genetic diseases that would otherwise take a highly trained team of specialists, and even provide foresight and prevention of human mistakes in diagnosis.

## 2 Computer Aided Medical Tools

Humanity is by no means close to losing an entire profession. Doctors are still absolutely necessary for the health of a population, especially since production of AI and their respective robotics that pass the Turing Test, and can thusly replicate the feel of a human person, isn't within immediate site. Person to person contact is vastly preferred when it comes to health care rather than talking to a screen, even though I believe that a program could simulate that relationship well enough to be satisfactory. An advantage is algorithmically generated and thus highly personalized relationships with a medial device.

Human doctors are here to stay, so I digress. Doctors rely on their instruments and tools to accurately provide information so that they can make an accurate and well informed diagnosis. There is a moving trend towards high resolution, 3D imaging to provide information in a non-invasive, accurate way [8]. This method allows doctors to literally see what is going on within their patients. 3D imaging using ultrasound has no unwanted side-effects similar to radiology imaging and is incredibly accurate, especially in shallow areas [8].

### 2.1 Improved Imaging Analysis

Researchers are working on improving this technology using clusters of computers connected through a low latency network to increase computing power and speed up algorithm time. Images, and virtually all graphic algorithms, are better calculated in parallel because of the nature of the problem. By increasing calculation power, doctors can perform more detailed ultrasounds of their patients and potentially see results in real time [8].

## 3 Genetic Problem Solving

One of the areas of medical diagnosis that computers can help the most is in cancer and genetic disorder treatment and diagnosis. These illnesses are generally extraordinarily complex and hard to calculate treatments for, assuming that the human specialist needed to diagnosis and treat the problem was able to accurately do so. I should say hard for current specialists, and very expensive. Computers have always provided an excellent means to create an expert system, but unfortunately cancer treatment is so large and complex that the system has to be flexible and consider or weigh a staggering amount of variables [5].

Tumor classification based on Gene Expression Profiles (GEPs), which is of great benefit to the accurate diagnosis and personalized treatment for different types of tumor, has drawn a great attention in recent years. [5]. Because of the sheet amount of genetic material in the human genome, spotting health problems before they can even form is near impossible unless you know what you're looking for. One possible avenue in biology is to create correlation filters to find patterns in genetic material, no matter what different gene type they happen to be expressed in. Minimum Average Correlation Energy (MACE) and Optimal Tradeoff Synthetic Discriminant Function (OTSDF), are introduced to determine whether a test sample matches the templates synthesized for each subclass [5]. The template represents a known genetic pattern for a tumor type, and the test sample represents up to an entire gene (which contains an immense amount of genetic information). Each correlation filter attempts to take advantage of a computer's processing power to determine the correlation of a gene and a known tumor pattern.

Analyzing genetic information is further complicated by almost never having identically matching patterns. The lengths of genetic material that represents the testing samples are produced in chunks, and so test samples may overlap and mismatch. MACE and OTSDF have shown progress in being able to detect similarity of the overall pattern while ignoring common mismatches [5].

### 3.1 Quicker, Efficient Analysis

Computer intelligence armed with these power filtering algorithms have produced incredible results. They are able to find overall similar patterns in half the time it would take a team of geneticists to sift through results. Not only is the batch operation time better, but applications are optimizing the process by finding similarities in the smallest amount of genes possible [4]. The significance of this finding results in the ability to extract simple diagnostic rules to diagnose a problem without the need for classifiers. The largest, and most clear benefit is that now problems can be found with relatively few gene expressions rather than thousands to have to filter through.

Many types and subtypes of cancer can be linked directly to DNA mutations or corruption, so being able to find the problematic gene patterns early can reduce the risk of developing a large amount of lethal cancer cells. While prevention is preferred, treatment is obviously of high importance also. Treatment is very difficult because of all the variables that come into the play: diet, bodily reaction, medications, and underlying genetic conditions. A cancer treatment center has recently been developing a computer system that optimizes a simulation to accurately piece together effective treatment solutions [1].

### 3.2 Solving Problems

There are three primary reasons to use computers to aid in diagnosis and treatments in the medical industry. They reduce errors, provide economy and efficiency, and quantitative quality of probability [7]. One of the obvious problems with integrating computers into the diagnosis and treatment planning cycle is if physican to computer interaction is scarce, and the general limitation of hardware for such a daunting data processing task. The use of computers for record storage and retreival is immediately feasable, but the ability to process complicated biological and mathematical models concerned with probabilty and symptom to disease matching, which exists inherently in the problem, may not be within the common hospital's or medical lab's reach [7].

Another problem to solve is the interaction between humans and machines. Person to terminal interaction, as discussed before, is a challenge to overcome because of a person's ability to spot a fake human [2]. These problems are easy to overcome, but will have to come with the passage of time and the further integration of interactive technology with every day life.

# 4 Doctor Oversight

Doctors and medical professionals are human, and fall victim to mistakes just the same as anyone else. In the medical domain, this can lead to costly consequences. Clearly the process of decision making for diagnosis and reasoning must be better understood so that we may prevent or catch them before mistakes occur [3].

## 4.1 Linguistic Expression

Narrative recounts of past misdiagnosis have revealed a surpring amount of linguistic information. Analysis of the way that diagnosis were expressed revealed how sure a doctor was of their diagnosis [3].

# 5 Conclusion

Human illness is a very challenging problem to figure out, despite the patterns associated with each disease. The immensity of the information associated with the process, not to mention the enormity of the database held in DNA, has lead to a coming paradigm shift of using computers as diagnostic machines for humans. Analyzing these patterns and using probably models and DNA sequencing to calculate the most effective treatment plans is increasingly being done by artificially intelligent machines. Computers are less likely to commit preventable errors, and have vastly increased economy and efficiency in the diagnostic process. IBM's Watson was just recently released onto genome data to detect cancer in patients, soon this won't be news but everyday events.

## References


[1] Felipe F. Baesler and José A. Sepúlveda. Healthcare ii: Multi-objective simulation optimization for a cancer treatment center. In *Proceedings of the 33Nd Conference on Winter Simulation*, WSC '01, pages 1405–1411, Washington, DC, USA, 2001. IEEE Computer Society.

[2] G. O. Barnett. History of the development of medical information systems at the laboratory of computer science at massachusetts general hospital. In *Proceedings of ACM Conference on History of Medical Informatics*, HMI '87, pages 43–49, New York, NY, USA, 1987. ACM.

[3] Wilson McCoy, Jeff B. Pelz, Cecilia Ovesdotter Alm, Pengcheng Shi, Cara Calvelli, and Anne Haake. Linking uncertainty in physicians' narratives to diagnostic correctness. In *Proceedings of the Workshop on Extra-Propositional Aspects of Meaning in Computational Linguistics*, ExProM '12, pages 19–27, Stroudsburg, PA, USA, 2012. Association for Computational Linguistics.

[4] Lipo Wang, Feng Chu, and Wei Xie. Accurate cancer classification using expressions of very few genes. *IEEE/ACM Trans. Comput. Biol. Bioinformatics*, 4(1):40–53, January 2007.

[5] Shu-Lin Wang, Yi-Hai Zhu, Wei Jia, and De-Shuang Huang. Robust classification method of tumor subtype by using correlation filters. *IEEE/ACM Trans. Comput. Biol. Bioinformatics*, 9(2):580–591, March 2012.

[6] Kathryn Womack, Wilson McCoy, Cecilia Ovesdotter Alm, Cara Calvelli, Jeff B. Pelz, Pengcheng Shi, and Anne Haake. Disfluencies as extra-propositional indicators of cognitive processing. In *Proceedings of the Workshop on Extra-Propositional Aspects of Meaning in Computational Linguistics*, ExProM '12, pages 1–9, Stroudsburg, PA, USA, 2012. Association for Computational Linguistics.

[7] Stephen R. Yarnall and Richard A. Kronmal. Computer aids to medical diagnosis—problems and progress. In *Proceedings of the 1966 21st National Conference*, ACM '66, pages 269–274, New York, NY, USA, 1966. ACM.

[8] F. Zhang, A. Bilas, A. Dhanantwari, K. N. Plataniotis, R. Abiprojo, and S. Stergiopoulos. Parallelization and performance of 3d ultrasound imaging beamforming algorithms on modern clusters. In *Proceedings of the 16th International Conference on Supercomputing*, ICS '02, pages 294–304, New York, NY, USA, 2002. ACM.