Article

Strong Cognitive Symbiosis: Cognitive Computing for Humans

Csaba Veres

Affiliation; University of Bergen, Fosswinckelsgate 6, Bergen, Norway

Abstract: *Cognitive Computing* has become a catchphrase in the technology world, with the promise of new smart services offered by industry giants like IBM and Google. Recent technological advances in Artificial Intelligence (AI) have thrown into the public sphere some old questions about the relationship between machine computation and human intelligence. While much of the industry and media hype suggests that many traditional challenges have been overcome, we show examples from language processing which demonstrate that present day Cognitive Computing still struggles with fundamental, long-standing problems with AI. An alternative conceptualization of artificial intelligence is presented, following Licklider's lead in adopting *man-computer symbiosis* as a metaphor for designing software systems that enhance human cognitive performance. A survey of existing proposals based on this view suggests that a distinction can be made between *weak* and *strong* versions of symbiosis. We propose a Strong Cognitive Symbiosis which dictates an interdependence rather than simply cooperation between human and machine functioning, and show two systems under development where the symbiotic relationship benefits both actors in achieving the task outcome.

Keywords: cognitive computing, cognition, AI, cognitive symbiosis, language, HCI

---

1. Introduction

The Gartner Hype Cycle for Smart Machines, 2014 named *Cognitive Computing* as a technology that is "on the rise".[1] By 2017, other Artificial Intelligence (AI) technologies such as *deep learning* and *machine learning* have joined it on the "Peak of Inflated Expectations".[2]. The IEEE Technical Activity for Cognitive Computing defines it as "an interdisciplinary research and application field" ... which ... "uses methods from psychology, biology, signal processing, physics, information theory, mathematics, and statistics" ... in an attempt to construct ... "machines that will have reasoning abilities analogous to a human brain".

The IBM corporation has invested heavily in bringing Cognitive Computing to the commercial world, starting perhaps with the computer Deep Blue which for the first time in history, on May 11, 1997, beat the world chess champion after a six-game match[3]. They then developed the computer 'Watson' which could process and reason about natural language, and learn from documents without supervision. In February 2011 Watson beat two previous champions in the "Jeopardy!" quiz show, demonstrating its ability to understand natural language questions, search its database of knowledge for relevant facts, and compose a natural language response with the correct answer. John Kelly, director of IBM Research, claims that "The very first cognitive system, I would say, is the Watson computer that competed on Jeopardy!"[4]. Kelly continues that cognitive systems can "understand our human language, they recognize our behaviours and they fit more seamlessly into our work–life balance. We can talk to them, they will understand our mannerisms, our behaviours - and that will shift dramatically how humans and computers interact."

---

1 https://www.gartner.com/doc/2802717/hype-cycle-smart-machines-
2 http://www.cityam.com/270451/gartner-hype-cycle-2017-artificial-intelligence-peak-hype
3 http://www-03.ibm.com/ibm/history/ibm100/us/en/icons/deepblue/
4 http://www.scientificamerican.com/article/will-ibms-watson-usher-in-cognitive-computing/

IBM's public promotional materials boldly state that "cognitive computers can process natural language and unstructured data and learn by experience, much in the same way humans do" and "interact naturally with people to extend what either humans or machine could do on their own."[5] This kind of extravagant language brings to mind the seminal thinkers about AI such as John Searle who coined the term 'strong AI' to describe systems which process information "in the same way humans do". Strong AI holds that "the appropriately programmed computer literally has cognitive states and that the programs thereby explain human cognition", which is on opposition to 'weak AI' where the computer merely "enables us to formulate and test hypotheses in a more rigorous and precise fashion"[3]. Searle argues against the possibility of strong AI with his famous Chinese room scenario, where he argues that an ungrounded symbol manipulation system lacks, in principle, the capacity for human understanding. The history of cognitive computing is sprinkled with claims approaching strong AI, though the topic is not explicitly broached as far as we are aware.

Google inc. has also been heavily involved in commercializing cognitive technologies, particularly deep learning[6], an evolution of neural networks with many hidden layers [1] which are particularly good at image recognition tasks. Google demonstrated GoogLeNet, the winning application at the 2014 ImageNet Large-Scale Visual Recognition Challenge[2].

The term 'cognitive computing' has been in use since the 1980s, as can be seen in the Google Ngram Viewer. The use of the term was associated with a strong growth in neural network computing following a joint US-Japan conference on Cooperative/Competitive Neural Networks in 1982[7]. In 1986 the backpropagation algorithm was detailed in the two volume manifesto: "Parallel distributed processing: Explorations in the microstructure of cognition"[4], which made neural network modeling much more versatile and accessible to researchers, and resulted in a plethora of new research programs exploiting the connectionist paradigm.

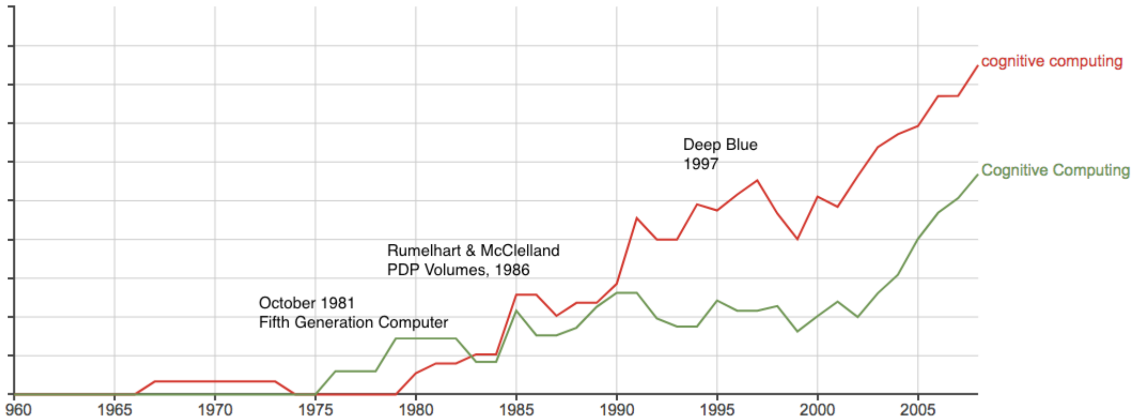

Figure 1. This is a figure, Schemes follow the same formatting. If there are multiple panels, they should be listed as: (**a**) Description of what is contained in the first panel. (**b**) Description of what is contained in the second panel. Figures should be placed in the main text near to the first time they are cited. A caption on a single line should be centered.

The advances in neural network computing also helped revive research in related fields such as Fuzzy Logic with the emergence of neuro-fuzzy systems which could learn parameters in a fuzzy system, leading to a set of methodologies that could perform imprecise reasoning, or *soft computing*[6]. Finally, the mid-1980s also saw the advent of genetic algorithms which could be used

5 http://www.research.ibm.com/cognitive-computing/#fbid=GZ_iDrBgajZ
6 http://deeplearning.net
7 http://cs.stanford.edu/people/eroberts/courses/soco/projects/neural-networks/History/history2.html

to avoid local minima in learning systems[5]. In 1993 the state of the art could be summarized by: "Cognitive computing denotes an emerging family of problem-solving methods that mimic the intelligence found in nature" ... "all three core cognitive computing technologies — neural-, fuzzy- and genetic-based — derive their generality by interpolating the solutions to problems with which they have not previously been faced from the solutions to ones with which they are familiar."[7]

While none of these technologies could decisively meet Searle's challenge for strong AI, it was pretty clear that the claimed biological plausibility of neural networks was to take us in that direction. Similarly, neuro-fuzzy systems were supposed to operate in ways analogous to human cognition: "In the final analysis, the role model for soft computing is the human mind."[6]. These technologies offered themselves as the foundation of programs that could indeed mimic human cognition.

Thirty years earlier Licklider was already contemplating a future with computers capable of thought like behaviour[15], in response to the bold investments in AI by the U.S. Department of Defense in the early 1960's. He imagined that the emergence of something like strong AI was not immediately imminent (contrary to the Air Force view that such machines would exist by the 80's), and there would be an interim period of "between 10 and 500 years" in which humans and computers would exist in a symbiotic relationship which would "bring computing machines effectively into the processes of thinking". He argued that for many years computer programs would not be able to mimic human thought processes, but instead work with humans as "dissimilar organisms living together in intimate association", enhancing the weaker parts of human cognition. On this view we should strive to understand how humans solve problems so that we can design programs which can take over those aspects of problem solving that are most mundane or difficult, rather than to design programs that mimic human reasoning. The principles of human cognition must be well understood even if they can't be directly implemented, so computer programs can be written with precisely the functions that are needed.

In the remainder of this paper we argue that the situation has not changed significantly since Licklider's seminal paper. Modern Cognitive Computing still falls short of realizing human-like thought. Section 2. considers the fundamentals of cognitive computing and argues that the current models do not reflect human cognitive processes. Section 3. presents related work. Section 4. develops our notion of a *Strong Cognitive Symbiosis* and discusses some applications which use these principles. Sections 5. and 6. conclude the paper.

2. Cognitive Computing?

While the popular discourse about Cognitive Computing emphasize their human-like characteristics, the scientific publications on the inner workings of Watson clearly show the many non human-like aspects of the implementation. For example, during the primary search phase Watson retrieves a large amount of potentially relevant data through a number of different techniques including the use of an inverted index in the Lucene search engine, and SPARQL queries to retrieve RDF triples from a triplestore[8]. This retrieves a huge volume of potentially relevant facts which are then further processed, often with statistical techniques. It is very unlikely that human reasoning would follow a similar process.

Noam Chomsky at the MIT symposium on "Brains, Minds and Machines" held in May 2011 [8] took modern AI to task more generally, voicing the opinion that the statistical learning techniques cannot reveal causal principles about the nature of cognition in general, and language in particular. They are useful engineering tools which can perform very useful tasks, but they will not give insight into cognitive processes, and will not operate by the same principles as those processes.

Peter Norvig, a fellow speaker at the symposium and director of research at Google argues, in a long essay on his web site, that this is a false dichotomy and that Chomsky's proposed explanatory

[8] http://mit150.mit.edu/symposia/brains-minds-machines

variables in linguistic knowledge are a fiction[9]. In his opinion predictive statistical models based on vast quantities of data are simply all there is to natural language cognition. Scientific progress is to be made not by postulating hypothetical causal mental states and testing their consequences through intuition in the form of grammaticality judgement, but by collecting vast quantities of language data and finding statistical models that best fit the data. If Norvig is correct then the current optimism about the possibilities of statistical models for cognitive computing are perhaps justified (and some of Watson's heuristics could be considered 'cognitive'), but if Chomsky is correct, then we might expect the current approaches to run into difficulties when human and machine processes differ. Our position is that if such differences are inevitable then it would be an advantage to know about them in advance to design solutions.

Chomsky believes in a distinction between linguistic *competence*, the tacit, internalised knowledge of language, and *performance* which is the observable manifestation of the former (speech acts, written texts, etc.). However, performance data is not a pure reflection of competence since linguistic productions are riddled with errors due to attention shifts, memory limitations and environmental factors. Chomsky therefore eschews corpus data as evidence for theory building, preferring instead grammaticality judgements which are elicited in response to sentences constructed to test a certain theory about competence.

Norvig defends the use of corpora, rejecting the use of grammaticality judgement as a form of linguistic evidence since it does not accurately reflect real language use. He cites the famous example from Chomsky [9] who claims that neither sentence 1 or 2 (or any part of the sentences) has ever appeared in the English language, and therefore any statistical model of grammaticalness will rule them as being equally remote from English. Yet it is clear to humans that 1 but not 2 is a grammatical sentence of English:

1. Colourless green ideas sleep furiously.
2. Furiously sleep ideas green colourless.

Pereira[10] demonstrates that modern statistical models of language prove Chomsky wrong. In fact, 1 is 200,000 times more probable than 2 in a large corpus of newspaper text. In his essay Norvig discusses a replication of the experiment on a different corpus "to prove that this was not the result of Chomsky's sentence itself sneaking into newspaper text", which corroborates the result. In addition, he finds that both sentences are much less probable than a normal grammatical sentence. Thus not only is Chomsky wrong about the statistical facts, but he is also wrong about the categorical distinction between grammatical/ungrammatical: 1 is more grammatical than 2, but less grammatical than ordinary sentences, according to Norvig.

We disagree with the conlcusions, and argue that this experiment in fact supports Chomsky's view. Suppose Norvig's concerns about the possible proliferation of Chomsky's sentence was in fact true, but it is true about 2 rather than 1. Perhaps a fundamentalist Chomskian government assumes power in the future and enforces a rule that every written text must be headed by Chomsky's "Furiously sleep ideas green colourless." to remind writers to use only grammatical sentences. Before long, the probability of 2 will exceed that of 1. But will 2 become more grammatical than 1, or will it just become annoyingly omnipresent? We think the latter, in which case the statistical theory would make the wrong prediction. To deny grammaticality judgment as a source of linguistic evidence in favor of corpora seems mistaken. There must be a principled criterion for what sort of observed strings should be counted as linguistic evidence.

One important use of statistical methods is for lexical disambiguation, as summarized in [10] "the co-occurrence of the words 'stocks', 'bonds' and 'bank' in the same passage is potentially indicative of a financial subject matter, and thus tends to disambiguate those word occurrences, reducing the

[9] http://norvig.com/chomsky.html

likelihood that the 'bank' is a river bank, that the 'bonds' are chemical bonds, or that the 'stocks' are an ancient punishment device". Norvig points out that 100% of the top contenders at the 2010 SemEval-2 completion used statistical techniques. However, the limitations of the approach can be easily demonstrated. Consider the following examples involving the ambiguous word bank.

3. I will go to the river bank this afternoon, and have a picnic by the water.
4. I will go to the local bank this afternoon, and afterwards have a picnic by the nearby water feature.

The word 'bank' in sentence 3 is clearly about "the land alongside or sloping down to a river or lake" (Oxford English Dictionary), while 4 is more difficult to interpret, but appears to be about the 'financial' interpretation of 'bank'. Both 3 and 4 contain words that are likely to co-occur with the 'sloping land' interpretation of 'bank' (i.e. picnic, water), which makes 4 misleading. But 4 also contains 'local' which is more likely to co-occur with the 'financial' interpretation, especially when they are strictly adjacent as in 'the local bank'. The interpretation of 'local bank' is something like 'the local branch of the bank', which is a sensible interpretation if the mental representation of the financial sense of 'bank' includes the fact that banks have 'branch offices'. We suggest that the resolution of ambiguity requires a suitable theory of compositional lexical semantics (e.g.[11]) rather than statistical models. In fact, even Watson uses a structured lexicon in question analysis and candidate generation[12].

We can push the example in sentence 4 a little further, by swapping the word 'local' with 'river':

5. I will go to the river bank this afternoon, and afterwards have a picnic by the water.

On first reading this seems odd, but suppose one was given as context that the person who uttered the sentence lived in a city which recently developed the previously neglected riverside into a business hub, and several banks were opened. With such knowledge the 'financial' reading of 'bank' becomes instantly clear, without a change in the a priori statistical distributions. As more people started talking and writing about the river branch of their bank then no doubt over time the statistical facts would come to reflect this usage. Statistical models completely miss the causal explanation for the change in the observed facts. Statistics does not drive interpretation: interpretation drives statistics. The current series of AI successes primarily involve statistical learning approaches which accomplish their specific tasks well, but lack the properties fundamental to aspects of semantic interpretation.

The semantic shallowness of cognitive computing by statistical learning has recently been illustrated through the construction of *adversarial* examples. In a paper titled "Intriguing properties of neural networks"[13], the authors show that slight (and hardly perceptible) perturbations in an image can cause it to be misclassified by a deep neural network. The manipulation involves changes in areas of the image that show points of maximum gradient in the trained network. A similar effect was shown in the paper "Deep Text Classification Can be Fooled"[14], where the authors showed that the *insertion*, *modification* and *removal* of hardly perceptible text snippets can cause text to be mis classified. In some cases, the insertion of a single key word can cause the text to be mis classified by a computer but remain correctly classified by the human. These examples show again that statistical techniques can perform semantic classification very accurately (99.9% accuracy before the insertion) without necessarily having representation of a semantics comparable to a human.

3. Related Work

The idea that technology can augment human cognition is an old one, and shared by many technical approaches. The engineering view of human thinking is central to the field of *cybernetics*, "the science of control and communication, in the animal and the machine" ([16]), and contributed to the emergence of the idea that human thinking could be improved artificially. The term *Intelligence Amplification* has been used in various guises since William Ross Ashby introduced the notion that human intelligence can be "amplified ... synthetically" [17] in his *Introduction to Cybernetics*. The use of computing devices to enhance human cognitive behaviors is of course a central theme of modern computing. Early attempts to harness the power of computers in this way can be seen in the work of Douglas Engelbart who founded the Augmented Human Intellect Research Center at SRI (Stanford Research Institute) International. He wrote: "The conceptual framework we seek must orient us toward the real possibilities and problems associated with using modern technology to give direct aid to an individual in comprehending complex situations, isolating the significant factors, and solving problems." [19]

While these early pioneers were concerned with how technology could help people solve complex tasks, it was the research field of Human Computer Interaction (HCI) which began directly investigating the interaction between humans and machines. Initially conceived during WWII as Human Factors Engineering, the goal was to discover principles which facilitated the interaction of humans and machines, in this case military hardware such as airplanes. As the investigations turned more specifically to human interaction with computing devices, other descriptors emerged to capture the subject matter more accurately: *cognitive systems engineering*, and *Human-Information Interaction* (HII) (see [18] for a historical review). Neo-Symbiosis is a new attempt to invigorate Licklider's notion of symbiosis in today's environment with our better understanding of cognition and more sophisticated computing resources. The insight of Neo-Symbiosis is that the human-computer interaction shouldn't be confined to simply augmenting the cognitive skills a person already has (e.g. with increased speed, memory, etc.), but to interact at a fundamental level to affect the reasoning process itself. A classic example of a visualization which can trigger human insight is the periodic table of elements conceived by Mendeleev. The Periodic table not only provided a simple display of known data but also pointed out gaps in knowledge that led to discoveries of new elements. It may have taken much longer to discover the gaps if the existing knowledge was coded in a different format [18]. Another example is the humble spelling checker which takes advantage of the computer's superior ability to reliably store and retrieve arbitrary data, in order to monitor any mistakes that a human might make in their spelling. Note that the interaction is genuinely symbiotic because the human can interact with the spell checker, instructing it to accept the correction, to ignore it, or even to learn a new alternative spelling if the person really did want to spell the word in a peculiar new way. These examples show that the basic principles behind Neo-Symbiosis are not necessarily new. Instead, the approach is to clarify known psychological principles in sufficient detail to specify functional allocations that are best performed by humans or computers. For example, human actions are frequently driven by context, such that a web search with the word "apple" would have a different intention if the person had previously searched for "orange" than if he had searched for "microsoft". Computer systems could therefore monitor cognitive state to determine intended context, and then use their powerful search capabilities to find relevant resources. As a related example, people often act differently in different contexts, but they might miss cues (or make mistakes) about the specific context in which they find themselves. A cognitive assistant could, for example, monitor a chat session in which a person is writing separately to their spouse and their boss, and issue a warning if they wrote an inappropriate message because they were inadvertently writing to the wrong person. Table 1 in [18] provides numerous examples of human cognitive properties and their implications for design of computer functionality. It should be noted that they base these cognitive properties on various proposals from the Nobel Prize winning psychologist Daniel Kahneman, and therefore the list is predicated on a particular theoretical position.

The IBM corporation's interpretation of *Symbiotic Cognitive Computing* is to immerse cognitive computing resources in a physical, interactive environment. They built a *Cognitive Environments Laboratory (CEL)* to explore how people and their cognitive computing implementations work together [20,21]. The CEL approach sees the role of the computer as a "super expert" which interacts with people, offering advice and information based on superior computational power. In the CEL environment the computer system follows individual users as they move about the environment, seamlessly connecting them to information sources. The system can perform functions like transcribing spoken conversations in order to preserve a recored of the discussion, and augment that with a record of all information that was on displays at the time. This can help decision makers re-trace their steps in case of disputes, for example. The environment can present information on one more of the large number of displays, based on vocal requests by the users. Many sophisticated, interactive 2D and 3D visualizations are available, as well as speech output. CEL is a technologically sophisticated environment in which researchers can study the interaction of humans and computers with state of the art speech and face recognition technologies.

The approach differs from Neo-Symbiosis, for example, where the operations of computer systems are designed to have a deeper relation to cognitive processes, rather than assume the role of intelligent assistants. The key observation is that Neo-Symbiosis uses specific theories about cognition to construct tools which support cognition at specific points of possible failure, whereas the CEL approach is to provide assistance during tasks which have been observed as difficult in work settings experienced over time. [21] propose five key principles of symbiotic cognitive computing: "context, connection, representation, modularity, and adaption." The principles are derived by "reflecting upon the state of human-computer interaction with intelligent agents and on our own experiences attempting to create effective symbiotic interactions in the CEL" ([21], p.84).

Finally, [22] argues that representations are the medium of cognition and are therefore key to supporting symbiosis. While the authors do not provide an implementation, they detail the *MatLab* programming competition which used a number of novel artifacts to communicate information about code snippets submitted by users, and to encourage the reuse of such code by other contestants using a rewards system. The authors argue that successful outcome was achieved through an symbiosis between artifacts and the players. However, the role played by the artifacts was to enable discovery and integration of the code snippets, and to provide an incentive mechanism to the players. As a symbiotic system, the MatLab game has a similar grounding as the CEL.

4. Towards a Strong Cognitive Symbiosis

The existing approaches to symbiosis stride the divide between two slightly different interpretations of the term. Mirriam Webster defines symbiosis as "the living together in more or less intimate association or close union of two dissimilar organisms (as in parasitism or commensalism)" or "a cooperative relationship (as between two persons or groups)". WordNet 3.1 gives a stronger interpretation as "the relation between two different species of organisms that are interdependent; each gains benefits from the other". The key difference is that the two organisms are dependent on one another in this stronger definition, implying that there are functions that neither could perform without the other.

This distinction can be seen as a symbiosis focused interpretation of *strong* versus *weak* AI, where *association* implies only that the machine can perform tasks which allow communication and co-operation at a level which is typically restricted to human-human interaction, whereas *interdependence* implies that the machine could not operate at some level without the human interaction. The former accords with a view of Cognitive Computing as the ability to perform tasks typically attributed to humans (e.g. natural language comprehension), without insisting that there was something in common with the way humans and machines performed the task. This appears to be the more common interpretation. The second view presupposes that computers can be programmed with algorithms which at some points of execution depends on the processes of human cognition.

We can get a sense of the difference in approaches through the following two theses on information representation in reasoning and decision making. In the book *Things that Make Us Smart* [23], Don Norman argues that the unaided human mind is "overrated" and much of what it has achieved is due to the invention of external aids that help overcome intrinsic limitations in memory capacity, working memory processing, and so on. In addition, the format of the external aids is critical for assisting particular kinds of reasoning. One example he cites from the work of Ruth Day involves written notation about prescription drugs and the recommended doses. Figure 2 (a) shows

Inderal	1 tablet 3 times a day
Lanoxin	1 tablet every AM
Carafate	1 tablet before meals and at bedtime
Zantac	1 tablet every 12 hours (twice a day)
Quinaglute	1 tablet 4 times a day
Coumadin	1 tablet a day

(a) Prescription suited for a doctor/pharmacist

	Breakfast	Lunch	Dinner	Bedtime
Lanoxin	✓			
Inderal	✓	✓	✓	
Quinaglute	✓	✓	✓	✓
Carafate	✓	✓	✓	✓
Zantac		✓		✓
Coumadin				✓

(b) Prescription suited for a patient

Figure 2. Two isomorphic views organized for different tasks

the longhand notation which is natural for prescribing doctors and contains valuable information for pharmacists filling the prescription. However, the format would not be easy for patients who are concerned with questions like "what pills should I take at breakfast?" These questions are much better answered by the representation in figure 2 (b). Notice in 2(b) that the medicine names have been re ordered so that they are now grouped according to the time of day to be administered. It seems intuitively indubitable that the two representations make certain tasks simpler, but there is no attempt to provide an explanation of this in terms of cognitive processes. Norman does make a distinction between *reflexive* and *experiential* thought, but these are not fleshed out in detail in terms of specific cognitive algorithms.

The second example concerns *cognitive illusions*, systematic problems of reasoning which result in errors of judgment (see [24] for a comprehensive review). A typical example is *base rate neglect*, which is supposed to show that the human mind lacks specific algorithms for naive Bayesian inference. For example, consider the following "mammography" problem (adapted from [25]:

> The probability of breast cancer is 1% for a woman at age forty who participates in routine screening. If a woman has breast cancer, the probability is 80% that she will get a positive mammography. If a woman does not have breast cancer, the probability is 9.69% that she will also get a positive mammography. A woman in this age group had a positive mammography in a routine screening. What is the probability that she actually has breast cancer? ______%

The correct answer can be calculated using then common formulation of Bayes' theorem (equation 1)

$$p(A \mid B) = \frac{p(B \mid A)\, p(A)}{p(B)} \tag{1}$$

which in this example evaluates to:

$$p(A \mid B) = \frac{(0.8)(0.01)}{(0.01)(0.80) + (0.99)(0.096)} = 0.078 = 7.8\% \tag{2}$$

[25] showed that 95 out of 100 physicians estimated the answer to be between 70% and 80%, which is in fact ten times the correct answer. This is an example of base rate neglect, since the error in reasoning is consistent with the claim that people ignore the relatively low background probability of having breast cancer (P(A) = 0.01). Thus, the nearly 10% probability of showing a false positive reading is quite high given the low background probability of actually having breast cancer, and drastically reduces the true probability that a person with a positive test reading has the illness.

However, [26] challenged the prevailing view that such experiments show that humans lack the appropriate cognitive algorithms to solve problems with Bayesian reasoning. Instead, they argue, humans do have the necessary procedures, but they operate with representations that are incompatible with the formulation of the problems. More specifically in the current example the problem formulation is in terms of *probability formats*, whereas the mental algorithms which would solve such problems operate on *frequency formats*. By way of illustrating their investigative strategy the authors suggest "assume that in an effort to find out whether a system has an algorithm for multiplication, we feed that system Roman numerals. The observation that the system produces mostly garbage does not entail the conclusion that it lacks an algorithm for multiplication. We now apply this argument to Bayesian inference."

Their general argument is that mathematically equivalent representations of information entail algorithms that are not necessarily computationally equivalent. While 0.01 + 0.01 might be mathematically equivalent to 1/100 + 1/100, the procedures for calculating the result are different. Using this reasoning they performed experiments in which the representational format was manipulated, for example, as in the following, and show significant increases in answers corresponding to the Bayesian outcome.

> 10 out of every 1,000 women at age forty who participate in routine screening have breast cancer. 8 of every 10 women with breast cancer will get a positive mammography. 95 out of every 990 women without breast cancer will also get a positive mammography. Here is a new representative sample of women at age forty who got a positive mammography in routine screening. How many of these women do you expect to actually have breast cancer? ______ out of ______

The frequentist representation suggests a startling conclusion, that humans might be able to draw Bayesian conclusions without attending to base rates. Consider how an old, experienced physician from an illiterate society with no formal statistical methods might solve the problem [26]. Imagine there is a terrible disease in the society which fortunately co occurs with a very distinctive symptom, albeit not with absolute certainty. In her lifetime the physician has seen 1000 people in total, and has seen the disease in 10 of those people. Of those 10, 8 showed the symptom. Of the remaining 990 people, 95 showed the symptom even though they were not afflicted. A new patient presents with the symptom .. what is the likelihood that he has the disease? That is, what is the likelihood that the patient has both the symptom and the disease? Based on past experience the doctor has seen a total of (95+8) = 103 people with the symptom, but only 8 had both the symptom and the disease. So the likelihood is 8/103 = 0.078. When the data is presented in frequency format, the number of hits and false alarms carry information about base rates, which do not need to be specifically evaluated. Thus, equation 2 corresponds to the mental calculation which is sufficient to provide the correct answer

$$p(H \mid D) = \frac{d\&h}{d\&h + d\& - h} \tag{3}$$

where *d & h* (data and hypothesis) is the number of cases with symptom and disease, and *d & -h* is the number of cases having the symptom but not the disease.

The authors conducted several experiments to determine human performance with different representations and showed how they relate to mental algorithms likely to be used in the different problem scenarios. Their conclusion is that "Cognitive algorithms, Bayesian or otherwise, cannot be divorced from the information on which they operate and how that information is represented", and this has a profound lesson for educators "... to teach representations instead of rules, that is, to teach people how to translate probabilities into frequency representations rather than how to insert probabilities into equations ..." and tutoring systems "... that enhance the idea of frequency representations with instruction, explanation, and visual aids hold out the promise of still greater success."

These concluding comments support the strong notion of cognitive symbiosis, where key aspects in the symbiotic system can be regarded as a hypothesis about cognitive functioning used to solve tasks, and the most useful information for assisting the problem solution. That is, the information exchanged between the human and computer in an effort to solve a problem are predicated on a hypothesis about what kind of cognitive algorithm will be used to solve the problem, and precisely what form of information and representation the algorithm requires.

Our vision of cognitive symbiosis is derivative of this approach. We assert that current approaches to AI are not sufficient to emulate the full range of human cognitive abilities, even though they do manage to perform *some* cognitive tasks at a level comparable to humans (e.g. [27]. However these successes are limited to very narrow domains and there are fundamental barriers which prevent similar success in others. This, in turn, implies that AI will be fundamentally limited is some respects and all effort to break through the limitations, fruitless. The most profitable avenue for technologists is to construct symbiotic systems in which the limitations of both humans and computers are mitigated as a result of the interaction.

Our work has focused on systems which use predominantly natural language, both in terms of data content and human-computer interaction. It is our contention that semantics, or symbol grounding[28], pose one of the most difficult problems for computational models of natural language understanding, and therefore the most productive point for a symbiotic interaction. More specifically, we propose that keyphrase/term/word extraction and lexical disambiguation are two semantic tasks which are particularly difficult to automate [33,34], and therefore most likely to benefit from human intervention. To this end we developed *LexiTags*, a semantic bookmarking service which demonstrates aspects of cognitive symbiosis.

LexiTags[29,30] is a social semantic bookmarking service in which users can save URLs of interest and tag them with disambiguated tags that are either WordNet senses or DBPedia identifiers. The service is very similar to http://delicious.com where users assign personal keywords (*tags*) to web sites of interest, and the service stores the URL together with the set of tags. The tags can then be used to *refind* the web sites. The additional step in LexiTags is that users have to disambiguate their tags by selecting one of the unambiguous choices offered through the user interface.

In addition to ambiguity resolution, LexiTags provides the identification of user specific key topics in a document through the nominated tags. While sophisticated statistical algorithms exist for topic analysis (e.g.[31]), the problem of allocating personalized, contextually significant topic(s) or tags to documents is more difficult because it relies on the subjective goals and beliefs of the reader [33]. In return for the minimal input, the user receives the benefit of precise automatic classification and search with semantically related terms, as well as precise, relevant recommendations from related data sources. Much of this work can be achieved through statistical word embeddings, which can be made more precise by extending the embeddings to WordNet *synsets* and *lexemes*. To illustrate, Table 1 shows related words for the non disambiguated tag *suit* using word2vec [], the state of the art tool for word embedding. The related words indicate that at least two distinct senses have been confounded, the noun *suit* of clothes and the verb to *suit*(his needs). The table also shows related words for the two senses individually, as well as the additional noun sense law*suit*, using AutoExtend [36] which forms ultradense sense representations based on WordNet.

Table 1. Ambiguous and disambiguated words and semantically similar words based on word2vec and AutoExtend.

Word	word2vec	AutoExtend
suit	suits, tailor, adapt, customize, conform, accommodate, tailored, meet, dress, cater	
suit#clothes		suit-of-clothes, attire, zoot-suit, garment, dress, trousers, pinstripe,shirt, tuxedo, gabardine, tux, pinstripe, costume, mumu
suit#lawsuit		lawsuit, countersuit, counterclaim, sue, violation, grievance, patent infringement, punitive damages, injunction
suit#accomodate		meet, cater, adapt, provide, fit, oblige, satisfy

A related tool shows how disambiguated lexical tags can be used to perform reasoning tasks which might otherwise be impossible. MaDaME[32] is a web application for developers who wish to mark up their sites with the schema.org classes and properties [10]. The tool allows users to highlight key words in their web site, and disambiguate them by selecting a sense from WordNet or DBPedia. The tool then automatically infers the most appropriate schema.org concepts and generates markup that adds schema.org as well as WordNet and SUMO identifiers to the HTML web page. The inference is currently performed via a mapping between WordNet synsets and schema.org classes; a tree search algorithm identifies the closest match between user selected synsets and the existing mappings. We are currently looking into replacing the classic search algorithm with one based on statistical methods.

Both of these tools generate curated metadata about web resources which can subsequently be used to automatically infer generalizations about, and relationships between web resources. The little human sourced semantics can go a long way in facilitating automated reasoning about the resources.

5. Discussion

The rise in the awareness of Artificial Intelligence in public consciousness has been phenomenal in the past few years. Every leading technology company has declared that "it's superior AI" are key to its continued success[11]. Russia's president Vladimir Putin has publicly declared that whoever masters AI will "rule the world"[12].

Together with this awareness have come warnings from prominent scientific and business figures about the dangers of an AI which becomes more powerful than the human mind. The so-called *singularity*[13] has profound warnings about what can happen if humans lose control of the machines[14].

We think that fears of singularity are overstated. While we are suitably impressed with recent progress in image recognition, text processing, and so on, we are also acutely aware of remaining

[10] http://schema.org is an effort originally proposed by a consortium of search engine providers to promote schemas for structured data on the Internet, on web pages, and in email messages

[11] Amazon: https://www.androidheadlines.com/2017/04/ai-key-amazons-future-success-says-jeff-bezos.html, Google: https://www.theverge.com/2016/4/21/11482576/google-ceo-sundar-pichai-cloud-ai-future

[12] http://fortune.com/2017/09/04/ai-artificial-intelligence-putin-rule-world/

[13] https://en.wikipedia.org/wiki/Technological_singularity

[14] Elon Musk: https://www.cnbc.com/2017/08/11/elon-musk-issues-a-stark-warning-about-a-i-calls-it-a-bigger-threat-than-north-korea.html, Stephen Hawking: http://www.bbc.com/news/technology-30290540

limitations. A technology which has difficulties with resolving lexical ambiguity, it seems to us, does not appear to be on the verge of attaining human-level cognition in the immediate future.

The biggest question of practical and commercial interest, then, is how to best use our human knowledge of statistical learning systems and AI in general, to construct computing platforms and information systems that can help humans perform complex cognitive tasks. What is the best way to benefit from *Cognitive Computing*?. A preconception that machines can perform tasks "just like humans" is counter productive if it is not true, because it sets up an industry expectation that cannot be fulfilled. For example if company A markets a fully automatic cognitive solution for managing unstructured data and semantics, then a competing company B will have a hard time developing a semi automated, symbiotic solution to the same problem, even if the symbiotic solution would prove more effective. In this paper we have argued that the preconception is in fact, not true. It is therefore time to take a step back, and consider principles which govern and delineate effective cognitive symbiosis between human and machine cognition.

Two applications were presented as an example of how the symbiosis might occur. Both rely on a human-computer interface at the level of lexical semantics, which is one of the difficult problems for current state of the art natural language processing systems. By providing information about semantics, the human user is rewarded by the vast computational and statistical inferencing capabilities of modern computing platforms. Both systems benefit from the interaction.

6. Conclusions

In conclusion we propose that, recent advances in deep neural network technology notwithstanding, we are no closer to predicting the arrival of "real" Artificial Intelligence than Licklider was 50 years ago. We are still in that interim period of "between 10 and 500 years". In this paper we argued that the false belief that we are in fact close to constructing computers with genuine cognitive abilities is disingenuous, for it diverts efforts away from investigating strong symbiotic systems which are constructed around their inherent but well understood cognitive limitations. We need to develop a principled framework which incorporates the shared and equal contribution of cognitive theories and technical solutions in programming smart machines, and not oversell short term, isolated engineering successes. *Strong Cognitive Symbiosis* is an attempt at such a framework.

Conflicts of Interest: "The authors declare no conflict of interest."

Abbreviations

The following abbreviations are used in this manuscript:

IEEE: Institute of Electrical and Electronics Engineers
AI: Artificial Intelligence
IBM: International Business Machines
SPARQL: SPARQL Protocol and RDF Query Language
RDF: Resource Description Framework
MIT: Massachusetts Institute of Technology
HCI: Human Computer Interaction
CEL: Cognitive Environments Laboratory
SUMO: Suggested Upper Merged Ontology
HTML: Hypertext Markup Language

References

1. Hinton, G. E., Osindero, S., & Teh, Y.-W. (2006). A fast learning algorithm for deep belief nets. Neural Computation, **18(7)**, 1527–1554. doi:10.1162/neco.2006.18.7.1527
2. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., et al. (2014). Going Deeper with Convolutions. Google Tech Report, arXiv.org.

3. Searle, J. R. (1980). Minds, brains, and programs. Behavioral and Brain Sciences, **3(03)**, 417–424. doi:10.1017/S0140525X00005756
4. David E. Rumelhart, James L. McClelland, and the PDP Research Group (Eds.). (1986). Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. 1: Foundations. MIT Press, Cambridge, MA, USA.
5. Yen, J. (1999). Fuzzy logic-a modern perspective. Knowledge and Data Engineering, IEEE Transactions on, **11(1)**, 153–165. doi:10.1109/69.755624
6. Zadeh, L. A. (1994). Soft computing and fuzzy logic. Software, IEEE, **11(6)**, 48–56. doi:10.1109/52.329401
7. Johnson, R. C. (1993). What is cognitive computing? Dr. Dobb's Journal, **18(2)**, 18–24.
8. Ferrucci, D., Brown, E., Chu-Carroll, J., Fan, J., Gondek, D., Kalyanpur, A. A., et al. (2010). Building Watson: An Overview of the DeepQA Project. AI Magazine,**31(3)**, 59–79. doi:10.1609/aimag.v31i3.2303
9. Chomsky, N. (1957). Syntactic Structures. The Hague: Mouton.
10. Pereira, F. (2000). Formal grammar and information theory: together again? Philosophical Transactions of the Royal Society of London a: Mathematical, Physical and Engineering Sciences, **358(1769)**, 1239–1253. doi:10.1098/rsta.2000.0583
11. Pustejovsky, James. (1995) The Generative Lexicon. MIT Press.
12. Deep parsing in Watson. (2012). IBM Journal of Research and Development, **56(3.4)**, 3:1–3:15. doi:10.1147/JRD.2012.2185409
13. Szegedy, Zaremba, Sutskever, Bruna: Intriguing properties of neural networks. (2013). arXiv:1312.6199 [cs.CV]
14. Bin Liang, Hongcheng Li, Miaoqiang Su, Pan Bian, Xirong Li & Wenchang Shi. Deep Text Classification Can be Fooled (2017). http://arxiv.org/abs/1704.08006
15. Licklider, J. C. R. (1960). Man-Computer Symbiosis. Human Factors in Electronics, IRE Transactions on, **(1), 4–11**. doi:10.1109/THFE2.1960.4503259
16. Wiener, N. Cybernetics. John Wiley & Sons, New York; 1948.
17. Ashby, W.R., An Introduction to Cybernetics, Chapman and Hall, London, UK, (1956). Reprinted, Methuen and Company, London, UK, 1964.
18. Griffith, D., and Greitzer, F. (2007) Neo-Symbiosis: The Next Stage in the Evolution of Human Information Interaction. International Journal of Cognitive Informatics and Natural Intelligence, 1(1), 39-52, January-March 2007.
19. Engelbart, D.C., Augmenting Human Intellect: A Conceptual Framework, Summary Report AFOSR-3233, Stanford Research Institute, Menlo Park, CA, October (1962)
20. J. O. Kephart and J. Lenchner, "A Symbiotic Cognitive Computing Perspective on Autonomic Computing," 2015 IEEE International Conference on Autonomic Computing, Grenoble, (2015), pp. 109-114. doi: 10.1109/ICAC.2015.16
21. Robert G. Farrell, Jonathan Lenchner, Jeffrey O. Kephart, Alan M. Webb, Michael J. Muller, Thomas D. Erikson, David O. Melville, Rachel K. E. Bellamy, Daniel M. Gruen, Jonathan H. Connell, Danny Soroker, Andy Aaron, Shari Trewin, Maryam Ashoori, Jason B. Ellis, Brian P. Gaucher, Dario Gil: Symbiotic Cognitive Computing. AI Magazine 37(3): 81-93 (2016)
22. Erickson, T.. A Distributed Cognition Perspective on Symbiotic Cognitive Systems: External Representations as a Medium for Symbiosis. AAAI Workshops, North America, mar. 2016. Available at: <https://www.aaai.org/ocs/index.php/WS/AAAIW16/paper/view/12616>. Date accessed: 28 Aug. (2017)
23. Donald A. Norman. (1993). Things that Make Us Smart: Defending Human Attributes in the Age of the Machine. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA.
24. Kahneman, D. (2011). Thinking, fast and slow. New York: Farrar, Straus and Giroux.
25. Eddy, D. M. (1982). Probabilistic reasoning in clinical medicine: Problems and opportunities. In Daniel Kahneman, Paul Slovic & Amos Tversky (eds.), Judgment Under Uncertainty: Heuristics and Biases. Cambridge University Press. pp. 249–267
26. Gigerenzer, G. & Hoffrage, U. How to improve Bayesian reasoning without instruction: Frequency formats. Psychological review 102, 684 (1995).
27. Kosinski, M., & Wang, Y. (2017, September 12). Deep neural networks are more accurate than humans at detecting sexual orientation from facial images.. Retrieved from psyarxiv.com/hv28a

28. Stevan Harnad, The symbol grounding problem, Physica D: Nonlinear Phenomena, Volume 42, Issue 1, (1990), Pages 335-346, ISSN 0167-2789, http://dx.doi.org/10.1016/0167-2789(90)90087-6.
29. Veres, C. (2011) LexiTags: An Interlingua for the Social Semantic Web. Social Data on the Web (SDoW) workshop at the 10th International Semantic Web Conference, Oct. 23 - 27, 2011.
30. Veres, C. (2013). Crowdsourced Semantics with Semantic Tagging: "Don't just tag it, LexiTag it!" in Maribel Acosta, Lora Aroyo, Abraham Bernstein, Jens Lehmann, Natasha F. Noy, Elena Simperl (Eds.): Proceedings of the 1st International Workshop on Crowd- sourcing the Semantic Web, Sydney, Australia, October 19, 2013. CEUR Workshop Proceedings
31. Blei, D. M. (2012). Probabilistic topic models. Communications of the ACM, **55(4)**, 77–84. doi:10.1145/2133806.2133826
32. Veres, C. (2013) Schema.org for the Semantic Web with MaDaME. in Steffen Lohmann (Ed.): Proceedings of the I-SEMANTICS 2013 Posters & Demonstrations Track, Graz, Austria, September 4-6, 2013. CEUR-WS.org 2013 CEUR Workshop Proceedings, Volume 1026.
33. Hasan, KS & Ng, V. Automatic Keyphrase Extraction: A Survey of the State of the Art. Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pp. 1262-1273 ACL (1) (2014). at https://pdfs.semanticscholar.org/f892/5b12b51cdbda6bc55da18b7e975c38446326.pdf
34. Agirre, Eneko; Edmonds, Philip; Word sense disambiguation: Algorithms and applications (2007) Springer Science & Business Media
35. P. Bojanowski, E. Grave, A. Joulin, T. Mikolov, Enriching Word Vectors with Subword Information. (2016) arXiv preprint arXiv:1607.04606
36. Rothe, Sascha and Schütze, Hinrich. AutoExtend: Extending Word Embeddings to Embeddings for Synsets and Lexemes. (2015) Proceedings of the ACL.