Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 November 2017 doi:10.20944/preprints201709.0062.v3

Article

Strong Cognitive Symbiosis: Cognitive Computing for Humans

Csaba Veres

University of Bergen, Fosswinckelsgate 6, Bergen, Norway

Abstract: *Cognitive Computing* has become somewhat of a rallying call in the technology world, with the promise of new smart services offered by industry giants like IBM and Microsoft. The 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. Much of the industry and media hype suggests that many traditional challenges have been overcome. On the contrary, our simple examples from language processing demonstrate that present day *Cognitive Computing* still struggles with fundamental, long-standing problems in AI. An alternative interpretation of *cognitive computing* 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 on this view suggests a distinction 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 introduce new software systems which were designed for cognitive symbiosis. We conclude that strong symbiosis presents a viable new perspective for the design of cognitive computing systems.

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

---

1. Introduction

The Gartner Hype Cycle for Smart Machines, 2017, names *Cognititve Computing* as a technology on the "Peak of Inflated Expectations" [1]. 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 been active in bringing Cognitive Computing to the commercial world for some years. Perhaps their earliest success was the computer 'Deep Blue' which beat the world chess champion after a six-game match on May 11, 1997 [2]. They then developed the computer 'Watson' which, it was claimed, 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! [3]. 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."

IBM's public promotional materials claim that "cognitive computers can process natural language and unstructured data and learn by experience, much in the same way humans do"[4]. This kind of extravagant language brings to mind the term 'strong AI' which describes systems that 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". On the other hand 'weak AI' proposes that the computer merely "enables us to formulate

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 November 2017 doi:10.20944/preprints201709.0062.v3

and test hypotheses in a more rigorous and precise fashion"[5]. 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. It is not clear if the current crop of *Cognitive Computing* systems claim to be strong AI, but the more extravagant claims appear not too far off.

Microsoft is another industry giant who has added *Cognitive Computing* to their repertoire, adding *Cognitive Services* to their Azure computing platform [6]. These are basically AI services which can be composed into an interactive application. The services include Vision, Knowledge, Language, Speech and Search.

In a similar vein, Google inc. is heavily involved in commercializing AI, particularly deep learning [7], an evolution of neural networks with many hidden layers [8] which are particularly good at image recognition tasks. Google demonstrated GoogLeNet, the winning application at the 2014 ImageNet Large-Scale Visual Recognition Challenge [9]. It should, however, be pointed out that Google does not specifically refer to cognitive computing by name.

The term *Cognitive Computing* has been in use since the 1980s, as can be seen in the Google Ngram Viewer (Figure 1). The early use of the term was associated with a strong growth in neural network research following a joint US-Japan conference on Cooperative/Competitive Neural Networks in 1982 [10]. In 1986 the backpropagation algorithm was detailed in the two volume publication: "Parallel distributed processing: Explorations in the microstructure of cognition"[11], which enabled networks to learn much richer associations than was previously possible. Neural network modeling became much more versatile and accessible to researchers, and resulted in a plethora of new research programs exploiting the connectionist paradigm.

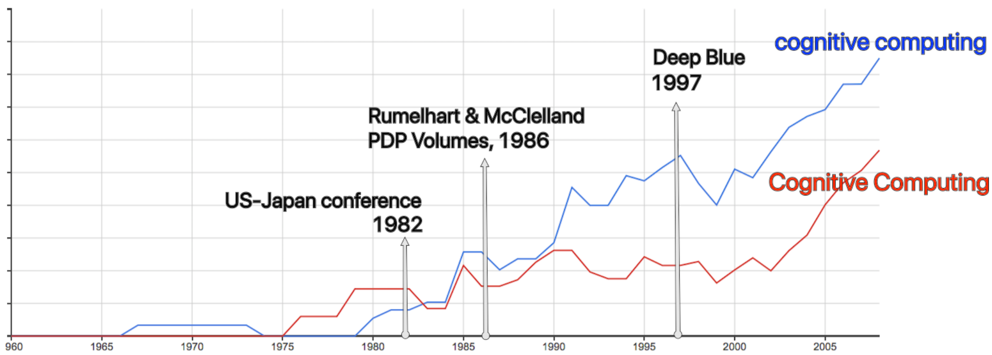

Figure 1. The use of the terms cognitive computing and Cognitive Computing according to Google's Ngram Viewer. Several key points in the evoultion of AI are also shown.

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*[12]. Finally, the mid-1980s also saw the advent of genetic algorithms which could be used to avoid local minima in learning systems[13]. In 1993 the state of the art could be summarized as: "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."[14]

While none of these technologies could decisively meet Searle's challenge for strong AI, it appeared that some of the research was heading in that direction. For example the claimed biological plausibility of neural networks was used to argue that connectionist models of cognition were more viable than theories based on symbol manipulation [11]. Similarly, neuro-fuzzy systems were supposed to operate in ways analogous to human cognition. According to Zadeh, "In the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 November 2017 doi:10.20944/preprints201709.0062.v3

final analysis, the role model for soft computing is the human mind."[12]. These technologies offered themselves as the foundation of programs that could indeed mimic human cognition. These sentiments are echoed in current claims that *Cognitive Computing* systems process information "as humans do".

Thirty years earlier Licklider was also contemplating a future with computers capable of human thought-like behaviour [15], in response to the bold expectations for AI by the U.S. Department of Defense (DOD). In the early 1960's, the DOD predicted that machines could take over from human operators by the 80's. But Licklider felt that the emergence of something like strong AI was not imminent, 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. Rather than build machines that mimic human reasoning, we should strive to understand how humans solve problems so that we can design programs that can take over those aspects of problem solving that are most mundane or difficult. The principles of human cognition must be well understood even if they can't be directly replicated, so computer programs can be written with precisely the functionality that is needed to enhance human cognition.

In this article 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 from the perspective of language processing and argues that the currently fashionable models do not accurately reflect human cognitive processes. Section 3. presents related work on human-computer symbiosis. 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 and Cognition

While the popular discourse about Cognitive Computing emphasizes the human-like properties, the scientific publications on the inner workings of Watson (perhaps the canonical example) clearly show the many non human-like aspects of the implementation. For example, during the initial 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[16]. 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. Mental processes almost certainly do not use SPARQL.

Noam Chomsky at the MIT symposium on "Brains, Minds and Machines" held in May 2011 [17] took modern AI to task, voicing the opinion that the currently popular statistical learning techniques cannot reveal causal principles about the nature of cognition in general, and language in particular. They are simply engineering tools which can perform very useful tasks, but they will not give insight into cognitive processes, and do not operate by the same principles.

Peter Norvig, a fellow speaker at the symposium and director of research at Google took up the challenge to argue that this is a false dichotomy and that Chomsky's proposed explanatory variables in linguistic knowledge are a fiction [18]. In his opinion predictive statistical models based on vast quantities of data are simply all there is to natural language cognition. Progress in Linguistics is to be made not by postulating hypothetical causal mental states and testing their consequences through intuition in the form of grammaticality judgment, 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 genuinely 'cognitive'), but if Chomsky is correct, then we must conclude that AI techniques and human cognition differ fundamentally. In this case we might

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 November 2017 doi:10.20944/preprints201709.0062.v3

expect the current approaches to run into difficulties under some circumstances. Our position is that if such differences are inevitable then it would be an advantage to know about them in advance, to design reliable and useful solutions which compensate for the deficits.

The fundamental theoretical divide is apparent in Chomsky's belief in linguistic *competence*, the tacit, internalized 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 judgments which are elicited in response to sentences constructed to test a certain theory about competence.

Norvig defends the use of corpora, while rejecting the use of grammaticality judgment as a form of linguistic evidence. He claims that elicited judgments do not accurately reflect real language use. He cites the famous example from Chomsky [19] 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, proving that grammar is not based on statistics:

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

Pereira [20] argues to the contrary and shows 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". The replication corroborates Pereira's findings. 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 about 1. and 2., but he is also wrong about the categorical distinction between grammatical and ungrammatical sentences: 1. is more grammatical than 2, but less grammatical than ordinary sentences, according to Norvig.

We disagree with these conclusions, and argue that the experiment in fact supports Chomsky's view. Suppose Norvig's concerns about the possible proliferation of Chomsky's sentence in the news corpus was in fact true, but it was true about 2. rather than 1. That is, sentence 2. becomes common in text. Perhaps a fundamentalist Chomskian government assumes power in the future and enforces a rule that every written newspaper 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 task where statistical methods have excelled is for lexical disambiguation, as summarized in [20] "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 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 riverside bank this afternoon, and if the line isn't too long, have a picnic by the nearby water feature.

 doi:10.20944/preprints201709.0062.v3

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 'riverside' which is a location, and gives us the clue that 'bank' must be some sort of bounded object that has a location property. We suggest that the resolution of ambiguity requires a suitable theory of compositional, structural lexical semantics (e.g.[21]) rather than statistical models. That is, some semantic elements like [location] and [physical object] would combine in some suitable account of compositional lexical semantics. In fact, even Watson uses a structured lexicon in question analysis and candidate generation[22].

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

5. I will go to the river bank this afternoon, and if the line isn't too long, 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 success stories 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" [23], 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" [24], 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. But the lack of semantics can also cause them to wildly mis behave. In the following section we review previous ideas about ways in which computers can augment human reasoning without necessarily trying to replicate it.

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" ([25]). The term *Intelligence Amplification* has been used in various guises since William Ross Ashby introduced the notion that human intelligence can be "amplified ... synthetically" [26] 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

 doi:10.20944/preprints201709.0062.v3

direct aid to an individual in comprehending complex situations, isolating the significant factors, and solving problems." [27].

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 [28] 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 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. An example is the visualization of the periodic table of elements conceived by Mendeleev in 1869, which can trigger novel human insight. 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 [28]. 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 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. The novelty of 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. [28] provide numerous examples of human cognitive properties and their implications for design of computer functionality. They base these cognitive properties on various proposals from the psychologist Daniel Kahneman, and therefore their proposals are predicated on a particular theoretical position [29].

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 cognitive computing implementations work together [30, 31]. 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 spoken requests by the users. Many sophisticated, interactive 2D and 3D visualizations are available, as well as speech output. CEL is a technologically sophisticated

 doi:10.20944/preprints201709.0062.v3

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, where the operations of computer systems are designed to have a deeper integration with 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. Thus, [31] 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" ([31], p.84). Clearly this is not a strongly theory driven approach.

Similarly, [32] argues that representations are the medium of cognition and are therefore key to supporting symbiosis. While the authors do not provide an implementation, they discuss 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 the artifacts and the players. However, the role played by the artifacts was simply 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, in intuition, as the CEL.

One view which presents IBM's Watson in a light closer to the Neo-Symbiosis view is shared by [33]. They argue that good results from cognitive systems can only come trough a symbiotic relationship where humans take charge of tasks in which the computers are deficient. In the case of Watson, this equates to the selection of the training corpus, which needs to be fine-tuned by humans because Watson cannot automatically infer which body of documents is likely to be relevant to a particular domain of interest. Another consideration is the kinds of data provided. Should the corpus include data catalogs, taxonomies and ontologies, or should the system be expected to discover these on its own? The decisions made by humans at this early stage of machine learning can significantly impact the overall performance of the system. A similar view is held by the CrowdTruth initiative which argues that semantic annotation should be spread among a large number of naive annotators, and that human disagreement should form an important input to cognitive learning systems [34]. In some places John Kelly also hints at this sort of interaction, claiming that computers must at some stage "... interact naturally with people to extend what either humans or machine could do on their own" [3].

A somewhat contrary but bold view of the consequences of Cognitive Computing can be seen in Dan Briody's post on IBM's "thinkLeaders" platform. He foresees a vastly changed business environment that has adapted to Cognitive Computing, and predicts that "New ways of thinking, working and collaborating will invariably lead to cultural and organizational change ..." [35]. Presumably these *new ways of thinking* are an adaptation to the human-like but not-quite-human cognitive assistants.

We will now describe our approach to cognitive symbiosis which does not rely on developing *new ways of thinking* but instead, intelligently supports *old ways of thinking* to achieve new results.

4. Towards a Strong Cognitive Symbiosis

The existing approaches to symbiosis stride the divide between two 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" 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

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

difference is that the two organisms are dependent on one another in the stronger WordNet definition, implying that there are functions that neither could perform without the other.

This distinction can be seen as a "symbiosis version" of *strong* versus *weak* AI. *Association* implies only that the machine can communicate and co-operate 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. That is, they share some key aspect of computation and representation which allows information exchange at an algorithmic level.

We can get a sense of this difference through the following two examples involving information representation in reasoning and decision making. In the book *Things that Make Us Smart* [36], 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. The information format of these external aids is critical for assisting particular kinds of reasoning. One example from the work of Ruth Day involves written notation about prescription drugs and the recommended doses. Figure 2 (a) shows 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 obvious that the two representations make certain tasks simpler, but there is no attempt to provide an explanation of this in terms of precise 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 [29] 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 [37]):

> 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:

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

[37] showed that 95 out of 100 physicians estimated the answer to be between 70% and 80%, which is in fact ten times higher than 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, [38] 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 analogy, "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. Using this reasoning they performed experiments in which the representational format was manipulated, and showed significant increases in answers corresponding to the Bayesian outcome. Consider the following, frequentist version of the previous problem.

> 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 researchers conducted several experiments and showed dramatic improvements in performance when the problem was presented in frequentist format. When presented in this format it is hard to ignore the large number of women (95) that will test positive even though they do not have breast cancer. The reasonable 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."

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 November 2017 doi:10.20944/preprints201709.0062.v3

These concluding comments support the strong notion of cognitive symbiosis. Our suggestion is that key interactions in the symbiotic system can be regarded as a hypothesis about cognitive functioning used to solve tasks. This hypothesis then determines the most useful information for assisting the problem solution. In other words, 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. [39]). However these successes are limited to very narrow domains and there are barriers which prevent similar success in others. This, in turn, implies that AI will be limited within the foreseeable future, just as it was in Licklider's time. Our suggestion is to adopt a strong view of cognitive symbiotic systems engineering in which the goal is to produce software systems whose interactions with people are optimized to tightly engage with empirically identified weaknesses in human as well as machine cognition.

Our concrete work on cognitive symbiotic systems has focused on applications which use predominantly natural language. In the area of natural language processing (NLP) and machine learning, semantic interpretation, or symbol grounding [40] pose one of the most difficult problems [34]. Two common NLP tasks which depend on semantic interpretation and therefore prove particularly difficult are keyphrase/term/word extraction and lexical disambiguation [41,42]. Yet these are tasks on which humans excel. Regarding lexical ambiguity, people are so efficient that they are typically unaware of alternative interpretations of ambiguous words and sentences [43]. The psycholinguist David Swinney has studied the time course of ambiguity resolution in sentence comprehension using the *cross modal priming* paradigm, His experiments have shown that humans can automatically resolve lexical ambiguity within three syllables of the presentation of the disambiguating information [44].

On the other hand humans are poor, but computers much more capable of storing and retrieving information. Jonides argues that memory is an essential component of thinking, and shows evidence that individual variations in working memory capacity correlate with performance on various reasoning tasks [45]. Limitations in working memory capacity result in deficiencies in reasoning. Minimizing the need to burden working memory ought to improve thinking.

The symbiotic applications we now describe were developed to exploit the human capacity for keyword selection and disambiguation, and combine it with the computer's capabilities to store, retrieve and discover vast amounts of text related to specific keyword indexes. We present this as an example of strong symbiosis, since each actor contributes to the result according to their respective cognitive strengths, and neither would be able to perform as accurately on their own.

LexiTags [46,47] is a social semantic bookmarking service in which users can save URLs of interest and annotate 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 called *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 (Figure 3). We call this *semantic tagging*. Semantic tagging therefore assigns unambiguous, user specific key topics to documents and other web resources. While sophisticated statistical algorithms exist for topic analysis (e.g.[48]), 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 [41].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 November 2017 doi:10.20944/preprints201709.0062.v3

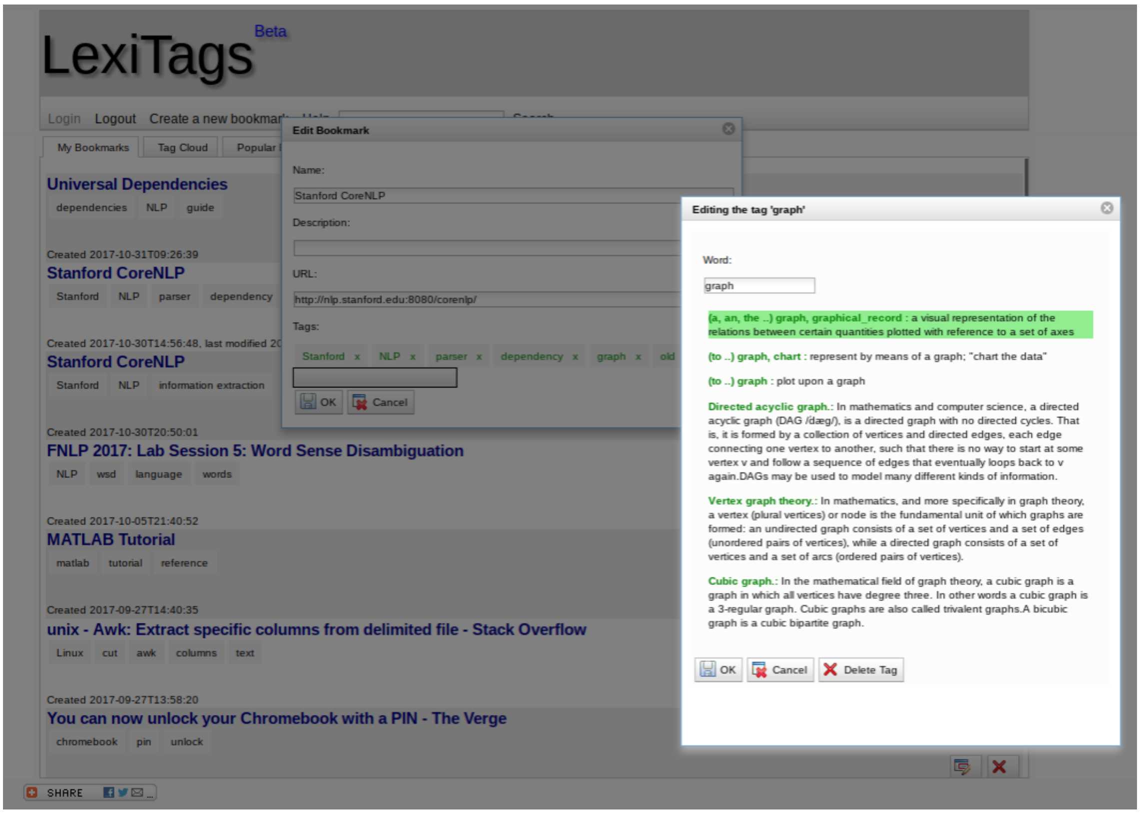

Figure 3. The LexiTags bookmarking interface. The bookmarks and editing interface are greyed out while the disambiguation screen is visible. Words are disambiguated with WordNet synsets or DBPedia identifiers.

In return for the additional step to disambiguate semantic tags, the user receives a range of benefits not available in traditional bookmarking services. Semantic tags facilitate accurate classification of the resource. This in turn makes it possible to identify other resources which are semantically related by precise relations such as taxonomy, meronymy, derivational relatedness, entailment or antonymy [49]. In addition, *word embeddings* can be used to identify statistically related semantic concepts [50]. Word embeddings can be made more precise and useful if disambiguation information is available. For example [51] forms ultradense representations with *AutoExtend* by using WordNet *synsets* and *lexemes* to create orthogonal transforms of standard word embeddings. To illustrate, Table 1 shows related words for the non disambiguated tag *suit* using word2vec, the state of the art tool for word embedding [52]. 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 these two disambiguated senses as encoded in AutoExtend, as well as the additional noun sense law*suit*. Clearly, recommendations of related items can be more accurate and varied when semantic tags are used. For example, the semantically disambiguated tag *suit#clothes* could recommend resources tagged with a rich set of the relevant tags *attire*, *garment*, *trousers*, *shirt*, *tuxedo*, *tux*, *pinstripe*, and not the more impoverished and mixed set from word2vec.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 November 2017 doi:10.20944/preprints201709.0062.v3

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#accomodate		meet, cater, adapt, provide, fit, oblige, satisfy
suit#lawsuit		lawsuit, countersuit, counterclaim, sue, violation, grievance, patent infringement, punitive damages, injunction

A second, related tool shows how disambiguated lexical tags can be used to perform a metadata reasoning task which might otherwise be very difficult. MaDaME[53] is a web application for developers who wish to mark up their sites with the http://schema.org classes and properties. 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. The tool allows users to highlight key words in their web site, and disambiguate them by selecting a sense from WordNet or DBPedia with a similar interface as LexiTags. 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. Figure 4 shows the higlighted word *novel* and its metadata mappings. 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.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 November 2017 doi:10.20944/preprints201709.0062.v3

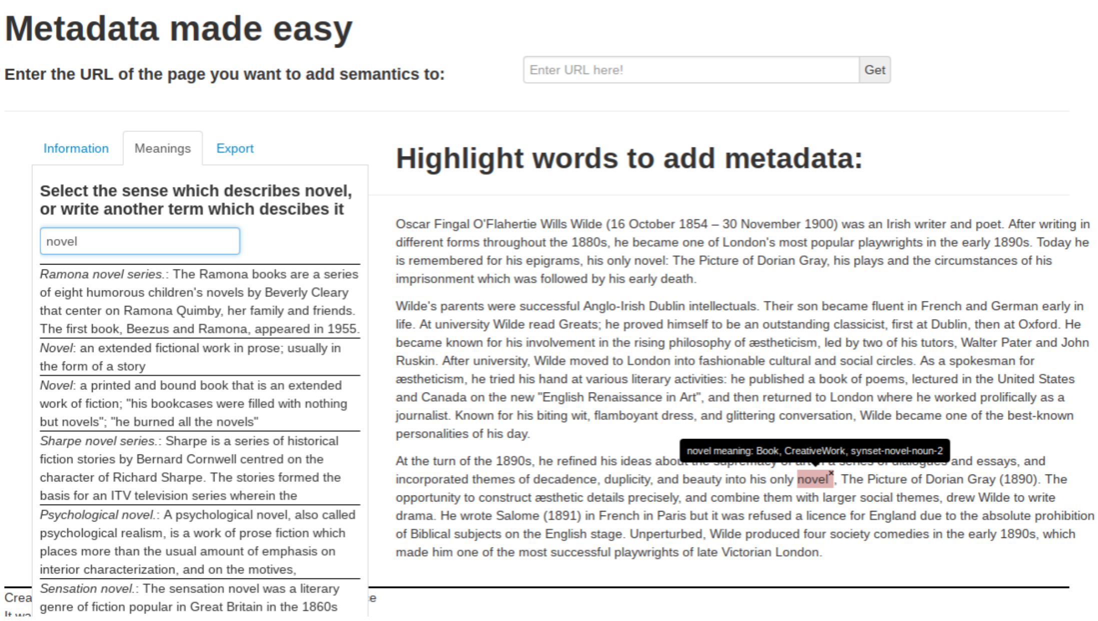

Figure 4. The highlighted word is disambiguated by the panel on the left. The WordNet synset is mapped to the closest schema.org type and SUMO concept, and the metadata is added to the HTML source code.

While Strong Cognitive Symbiosis is a new design principle pioneered in this publication, elements of the approach can be gleaned in other applications. For example [54] discuss visual analytic decision-making environments for large-scale time-evolving graphs. These pose difficulties for decision making because they describe phenomena where large volumes of inter related data are evolving in complex patterns. Current visualization techniques do not offer a solution for decision making with such complex data. The authors argue that designing decision-making environments for such complex tasks require systems which "work in symbiosis with humans" (p.85). These would require an understanding of human thought processes and incorporate those processes into the computational model to reduce human burden. To this end they propose three HCI principles for human-machine interaction in visual analytics regarding: *(i)Data and view specifications, (ii) View manipulations, and (iii) Process and provenance*. These principles essentially prescribe that graph browsing interfaces should allow users to select and navigate graph structures according to their specific needs and goals, and to retain traceability of states. In order to react to user requests with time-evolving graphs, the application has to solve some difficult computational problems in terms of data management, analytics and graph visualization. However, the computational problems almost exclusively involve formal properties of the graphs themselves rather than the way a human might process those graphs. For example, summarizing graphs involves the calculation of node-edge properties such as *journey, density, eccentricity, diameter, radius, modularity, conductance, reachability*, and *centrality measures*. Special techniques are needed for analysis, summary, and visualization of evolving graphs in which these formal properties are subject to change. The symbiotic aspect of the application is that the visualizations and summaries must be comprehensible for humans, and humans must be able to manipulate those representations to answer their questions.

5. Discussion

The rise in the awareness of Artificial Intelligence in public consciousness has been phenomenal in the past few years. Many leading technology companies have declared that "it's superior AI" are key to its continued success: Amazon, Google, Apple [55–57]. Russia's president Vladimir Putin has publicly declared that whoever masters AI will "rule the world" [58].

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* has profound warnings about what can happen if humans lose control of the machines [59–61].

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 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 and might stifle alternative approaches. For example if company *A* markets a fully automatic cognitive solution for managing unstructured data, 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. Computers are still very far from thinking like humans. It is therefore time to take a step back, and focus on systems which use modern AI techniques to realize a strong symbiotic relation between human and machine.

We acknowledge that Strong Cognitive Symbiosis is difficult to achieve because it requires a design in which the operation of the machine and human can interact at a deep algorithmic level. This is not typical of modern AI systems, especially those constructed around neural network or deep-learning frameworks. Such programs typically learn end-to-end generalizations from large data sets, and the focus is the input-output mappings they can learn. In the rare cases where an intervention is made at an algorithmic level, it is to the detriment of the result [23,24]. However, there is an emerging approach which is highly compatible with our suggestions, *Neural-Symbolic Learning and Reasoning* [62]. The goal of neural-symbolic computation is to integrate neural network learning and symbolic reasoning, for example by extracting logical expressions from trained neural networks, or using an independent feature space to enable *heterogeneous transfer learning*. The latter example is particularly interesting. [63] show how it is possible to train a network on an image clustering task where the training data is from a feature set that is different from the test set. In essence, they use an independent set of invariant image features derived from local image descriptors [64], to mediate between the training and test set. The technique works by computing co occurrence matrices between the invariant features F and an image space A, and between the features F and a second, text labeled image space W. Finally, [63] show a transfer of learning from text space W to image space A. The intriguing possibility for a strong symbiosis perspective is to use a similar technique in a domain where the invariant features are tuned through close interaction between human users and the computer, to obtain the best results for each individual user.

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

 doi:10.20944/preprints201709.0062.v3

term, domain restricted 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. http://www.cityam.com/270451/gartner-hype-cycle-2017-artificial-intelligence-peak-hype accessed 10/10/2017
2. http://www-03.ibm.com/ibm/history/ibm100/us/en/icons/deepblue/ accessed 10/10/2017
3. http://www.scientificamerican.com/article/will-ibms-watson-usher-in-cognitive-computing/ accessed 10/10/2017
4. http://www.research.ibm.com/cognitive-computing/#fbid=GZ_iDrBgajZ accessed 10/10/2017
5. Searle, J. R. (1980). Minds, brains, and programs. Behavioral and Brain Sciences, **3(03)**, 417–424. doi:10.1017/S0140525X00005756
6. https://azure.microsoft.com/en-us/services/cognitive-services/ accessed 10/10/2017
7. http://deeplearning.net accessed 10/10/2017
8. 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
9. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., et al. (2014). Going Deeper with Convolutions. Google Tech Report, arXiv.org.
10. http://cs.stanford.edu/people/eroberts/courses/soco/projects/neural-networks/History/history2.html accessed 10/10/2017
11. 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.
12. Zadeh, L. A. (1994). Soft computing and fuzzy logic. Software, IEEE, **11(6)**, 48–56. doi:10.1109/52.329401
13. Yen, J. (1999). Fuzzy logic-a modern perspective. Knowledge and Data Engineering, IEEE Transactions on, **11(1)**, 153–165. doi:10.1109/69.755624
14. Johnson, R. C. (1993). What is cognitive computing? Dr. Dobb's Journal, **18(2)**, 18–24.
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. 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
17. http://mit150.mit.edu/symposia/brains-minds-machines accessed 10/10/2017
18. http://norvig.com/chomsky.html accessed 10/10/2017
19. Chomsky, N. (1957). Syntactic Structures. The Hague: Mouton.
20. 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
21. Pustejovsky, James. (1995) The Generative Lexicon. MIT Press.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 November 2017 doi:10.20944/preprints201709.0062.v3

22. Deep parsing in Watson. (2012). IBM Journal of Research and Development, **56(3.4)**, 3:1–3:15. doi:10.1147/JRD.2012.2185409
23. Szegedy, Zaremba, Sutskever, Bruna: Intriguing properties of neural networks. (2013). arXiv:1312.6199 [cs.CV]
24. 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
25. Wiener, N. Cybernetics. John Wiley & Sons, New York; 1948.
26. Ashby, W.R., An Introduction to Cybernetics, Chapman and Hall, London, UK, (1956). Reprinted, Methuen and Company, London, UK, 1964.
27. Engelbart, D.C., Augmenting Human Intellect: A Conceptual Framework, Summary Report AFOSR-3233, Stanford Research Institute, Menlo Park, CA, October (1962)
28. 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.
29. Kahneman, D. (2011). Thinking Fast and Slow. Farrar, Straus and Giroux (Publisher).
30. 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
31. 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)
32. 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)
33. Judith Hurwitz, Marcia Kaufman, Adrian Bowles. (2015). Cognitive Computing and Big Data Analytics. Wiley. ISBN: 978-1-118-89662-4
34. Lora Aroyo, Chris Welty. (2014) The Three Sides of CrowdTruth. J. Human Computation. 1(1).
35. https://www.ibm.com/blogs/think-leaders/technology/cognitive-computing-new-vocabulary/ last accessed 10/10/2017
36. 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.
37. 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
38. Gigerenzer, G. & Hoffrage, U. How to improve Bayesian reasoning without instruction: Frequency formats. Psychological review 102, 684 (1995).
39. 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
40. 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.
41. 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
42. Agirre, Eneko; Edmonds, Philip; Word sense disambiguation: Algorithms and applications (2007) Springer Science & Business Media
43. Steven Lawrence Small, Garrison Weeks Cottrell, and Michael Tanennaus (Eds.) (1988) Lexical Ambiguity Resolution: Perspectives from Psycholinguistics, Neuropsychology, and Artificial Intelligence. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.
44. Swinney, David. (1979) A Lexical access during sentence comprehension: (Re)consideration of context effects. Journal of Verbal Learning and Verbal Behavior, volume 18, number 6, pages 645–659

45. Jonides, J. Working Memory and Thinking. In. Smith, E. E., and Osherson, D. N. (1995) An Invitation to Cognitive Science, Second Edition, Volume 3, MIT Press
46. 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.
47. 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
48. Blei, D. M. (2012). Probabilistic topic models. Communications of the ACM, **55(4)**, 77–84. doi:10.1145/2133806.2133826
49. Christiane Fellbaum (1998, ed.) WordNet: An Electronic Lexical Database. Cambridge, MA: MIT Press.
50. Tomas Mikolov, Scott Wen-tau Yih, Geoffrey Zweig (2013). Linguistic Regularities in Continuous Space Word Representations. Published In Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT-2013)
51. Rothe, Sascha and Schütze, Hinrich. AutoExtend: Extending Word Embeddings to Embeddings for Synsets and Lexemes. (2015) Proceedings of the ACL.
52. T Mikolov, I Sutskever, K Chen, GS Corrado, J Dean (2013) Distributed representations of words and phrases and their compositionality. Advances in neural information processing systems, 2013
53. 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.
54. Venna, Siva RamaKrishna Reddy & Gottumukkala, Raju & Raghavan, Vijay. (2016). Visual Analytic Decision-Making Environments for Large-Scale Time-Evolving Graphs. Handbook of Statistics . . 10.1016/bs.host.2016.07.002.
55. https://www.androidheadlines.com/2017/04/ai-key-amazons-future-success-says-jeff-bezos.html last accessed 10/10/2017
56. https://www.theverge.com/2016/4/21/11482576/google-ceo-sundar-pichai-cloud-ai-future last accessed 10/10/2017
57. http://bgr.com/2017/06/11/apple-ai-machine-learning-ar-wwdc-2017/ last accessed 10/10/2017
58. http://fortune.com/2017/09/04/ai-artificial-intelligence-putin-rule-world/ last accessed 10/10/2017
59. https://en.wikipedia.org/wiki/Technological_singularity last accessed 10/10/2017
60. https://goo.gl/JkudMV last accessed 10/10/2017
61. http://www.bbc.com/news/technology-30290540 last accessed 10/10/2017
62. Garcez, Artur d'Avila and Besold, Tarek R. and Raedt, Luc de and Foldiak, Peter and Hitzler, Pascal and Icard, Thomas and Kühnberger, Kai-Uwe and Lamb, Luis C and Miikkulainen, Risto and Silver, Daniel L, (2015), Neural-Symbolic Learning and Reasoning: Contributions and Challenges. Knowledge Representation and Reasoning: Integrating Symbolic and Neural Approaches: Papers from the 2015 AAAI Spring Symposium
63. Yang, Qiang and Chen, Yuqiang and Xue, Gui-Rong and Dai, Wenyuan and Yu, Yong. (2009). Heterogeneous transfer learning for image clustering via the social. http://dx.doi.org/10.1145/1687878.1687880
64. Lowe, David G. (2004) Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision, issn 0920-5691, volume 60, number 2, p. 91 - 110. http://dx.doi.org/10.1023/b:visi.0000029664.99615.94