SYNTIENCE INC.
"We Understand"

171 Main St., Suite 276
Los Altos, CA 94022

T +1 408 802-7229

monica@syntience.com
syntience.com

A Radical Approach to Understanding Text

Monica Anderson, MSCS

The phenomenon of human intelligence can be divided into separate faculties for Reasoning and Understanding.

Reasoning is a conscious, symbolic, logical, Reductionist, context free, model based, and largely serial process that takes seconds to years.

Understanding is a subconscious, subsymbolic, intuitive, Holistic, context dependent, pattern based, and largely parallel recognition of objects, agents, concepts, abstractions and relationships that takes mere milliseconds.

Animals understand but do not reason; In humans, conscious, logical Reasoning is built on top of subconscious, intuitive Understanding. Both aspects were discussed at some parity until around 1955. With the advent of computers, Cognitive Science became dominated by programmers that concentrated their efforts on logic based, Reductionist theories. Since then, research has been overmuch preoccupied with Reasoning at the expense of Understanding. Without a foundation of Understanding, the Reductionist, logic based efforts at cognition were (and are) building castles in the air, since these systems have nothing to reason *about*.

Reductionism has been amazingly successful and effective. Since around 1650, the Reductionist approach to analyzing the world by building modular models of Structure and Causality has resolved thousands of important scientific questions. We are taught that in order to be "Scientific" we must use Reductionist methods and context free models. Computer Science, Mathematics, Physics, and Chemistry favor these approaches since they so often provide excellent results in these domains. But in other disciplines we have discovered that Reductionist methods are insufficient to attack the most important problems we are facing. Biology, Ecology, Psychology, Sociology, etc., and indeed all sciences that deal with *Life,* are forced to increasingly use alternative approaches in order to produce useful results. Precise measurements and reliable data are unavailable. Models become too complex and brittle to be useful. Context becomes more important and often dominates the problem statement.

Complexity theory, Systems Biology, Chaos Theory, and other sciences have outlined the borders; In my analysis[1], four major issues severely restrict the applicability of Reductionist models or even prevent their construction:

1. Chaotic systems – long term predictions cannot be made in systems exhibiting "Deep Complexity".
2. Irreducible systems – Reductionist methods do not work in open, inseparable, intractable, or constantly changing systems. For instance, the laws of Thermodynamics cannot be used in open systems.
3. Ambiguity – Many systems deal with ambiguity, incomplete information, opinions, hypotheses, lies and other misinformation, either internally or as their input. These cannot use Reductionist models since models require complete and correct input data in order to produce useful output.
4. Emergent effects – Things like Quality, Meaning, Health, Intelligence etc. cannot be modeled.[2]

Whenever any of these issues are encountered, Reductionist models cannot be created or used; encountering *all four* issues at once is surprisingly common. Systems with all four issues will be labeled "**Bizarre Systems**". [3] [4]

---

[1] http://artificial-intuition.com/bizarre.html . See http://videos.syntience.com for a video of Ms. Anderson's talk on Bizarre Systems.

[2] In simple cases, emergence can be *simulated* from faithful models of components but such simulations have little *explanatory power* – they provide no insights of the compact and context-independent form that Reductionist Science provides and expects.

[3] Dr. Kirstie Bellman first introduced the term at the IEEE International Conference on Systems Man and Cybernetics in Orlando FL in 1997. Dr. Steven Kercel, University of New England, Endogenous Systems Research Group popularized the term at ANNIE 1997 conference. Several other terms exist but we favor this one since it is memorable.

[4] See also Dr. Andrzej K. Konopka: Systems Biology, pp140-174

Some of the most important examples of Bizarre Systems:

"**The World is Bizarre**" – any attempt to make a dependable model of any significant aspect of the World will ultimately fail. Models of stock markets, economies[5], and social networks are always brittle and may easily fail catastrophically, as we have recently seen. Projects that attempt to model larger sub-domains of the world (like CYC, or the Semantic Web) may achieve partial success but will never reach a competence deserving the name "Intelligence".

"**Life is Bizarre**" – all life sciences deal with Bizarre Systems. The reason for their inability to make solid predictions for the phenomena they study (the way physics can) is that their domain is so complex, open, and full of emergent effects that long-term predictions are impossible.

"**Organisms are Bizarre**" – human physiology, gene expression, proteomics, the human interactome, drug effects and interactions etc. are all so difficult to model that "rational drug design" has remained an unreachable ideal.

"**The Brain is Bizarre**" – attempts to analyze and model brain functionality using higher resolution MRI and large computers will fail - at least if the goal is to build Reductionist models of "intelligence".

"**Language is Bizarre**" – any attempt to get to the *meaning* of language using Reductionist methods such as grammars will fail. Language is chaotic, irreducible, ambiguous, and emergent.[6]

Note that Bizarre Systems are not "bizarre" to humans; we effortlessly deal with Life, the World, our bodies, other people, and Language on a daily basis. This is because our brains do not use Reductionist Models at low levels.[7] Bizarre Systems are "Bizarre" only to those that attempt to build models of them, such as Reductionist Scientists and Reductionist programmers attempting to model life, the world, commerce, intelligence, or languages.

Cognitive Science should not be restricted to a Reductionist stance; it should have stayed closer to the Life Sciences, the way they were before 1950. According to one source[8], Cognitive Science is the only science still stuck in 1950's style Reductionism. One approach, used in other sciences, is to forego prediction and concentrate on description. Another is to use weaker models, such as statistical approaches, where the only model required is of the statistical distribution of the phenomena.[9] These approaches allow progress to be made in Bizarre Domains.

Yet another is to use **Model Free Methods** (**MFM**). Over a dozen MFMs can be easily identified. The simplest one is "Trial and Error", a.k.a. "Generate and Test". No matter how hard or simple the problem, no matter how stupid or smart an agent is, they can always use Trial and Error. All MFMs are *universally applicable*; you don't need intelligence to be able to select the correct and applicable Scientific Model, depending on the situation. Drugs are still discovered (using methods similar to "Trial and Error") rather than designed, and are tested in Model Free ways (administer and observe without preconceived notions about expected effects). The term "Model Free Methods" was first used in Genetics by L.S. Penrose around 1935.[10]

Starting from the simplest and going towards more powerful ones we find methods like "Trial and Error", "Enumeration", "Remember Failures", "Remember Successes", "Table Lookup", "Wildcarded Patterns", "Variant Solutions", "Determination of Saliency","Adaptation", "Selectionism", "Language", "Narrative", and "Consultation"[11]. At the core, most MFMs are about discovery, learning from experience, and from immediately applicable predictions, and range from trivial solutions to more general meta-solutions. Most of the simple MFMs have trivial implementations in computers; but much more advanced and effective algorithms also exist. Some are classified as Machine Learning algorithms among model based variants, but the distinction of being a MFM is important.

---

[5] http://nobelprize.org/nobel_prizes/economics/laureates/1974/hayek-lecture.html

[6] For examples, see Ms. Anderson's talk on Bizarre Systems at http://videos.syntience.com/ai-meetups/bizarresystems.html

[7] Our capability to build Reductionist models of simpler problems emerges at the higher layers of Holistic, sub-scientific intuitions

[8] http://shorst.web.wesleyan.edu/papers/Beyond_Reduction_F.doc p. 4

[9] Statistical methods have been adopted in the document understanding domain for web search engines etc.

[10] Penrose LS, The detection if autosomal linkage in data which consists of pairs of brothers and sisters of unspecified parentage, Ann Eugen 1935; 5:133-148

[11] This list is incomplete, and the more advanced methods deserve treatments outside of the scope of this paper. Some methods in the list have both model based and model free variants. Intermediate strategies such as Pseudo-Models and Non-Parametric Models also exist.

In a surprising reversal, the next generation of systems for computer based cognition may benefit from **totally restricting themselves to using only Model Free Methods** since their domain of discourse is Bizarre**.** This limitation might seem like a drawback but some MFMs are amazingly powerful. “Narrative” can be model free if the narration simply recounts what happened without supplying a moral or other kind of Reductionist analysis; "Consultation" is regarded as cheating in tests in school – if you ask someone else for the answer to a problem, then you can write down the correct solution but you have gained no insight that could help you solve another similar problem. All MFMs share this drawback. It is mitigated by use of “Wildcarded Patterns” and in other ways in advanced MFMs.

“Enumeration” was the MFM used to prove the Four-Color Theorem some years ago. This was the first time that the Mathematical community accepted that a computer based proof that no human could follow in detail was still a valid proof. But it was viewed as “inelegant”.

Some other common objections to MFMs: Fallible; wasteful of resources; inscrutable results; opaque processes; not repeatable; brute force; not optimal; no guarantee of completeness; results are too context specific; no transfer of knowledge to other problems; results have no general applicability[12]. But human minds *also* have *all* of these "problems". These properties have to be accepted in order to have any chance at creating intelligent/ intuitive systems.[13]

Reductionist Cognitive Science is divided: on one hand computer capabilities are expected to double every 18 months in order to supposedly reach "parity with human brains" by 2030. On the other hand, many decry MFMs as being "too inefficient". But if asked what they would do with a TeraByte of RAM, most programmers don't have a good answer. The recent availability of large computers is a key enabler of the use of MFMs since they require a lot of memory to store the rich contexts of past experiences; MFMs in turn will enable machine intelligence. In essence, MFM are the best way to use tomorrow’s large computers. Traditional Reductionist programming techniques were created for yesterday’s puny computers. Holistic programming using MFMs is really, really different.

When a problem is too complicated to be solved using any known Reductionist method, the Reductionist simplifies the problem by ignoring more of the context and by subdividing – reducing – the problem into simpler parts (if the problem is in fact reducible). In contrast, the Holist goes in search of *more* information in larger and larger contexts, hoping to discover patterns that match their prior experience. This can obviously be done even in irreducible domains.[14] To a Reductionist, context “dependency” is something bad, something to discard. To a Holist, **context is the *whole point***. These differences between Reductionist Science and Holistic problem solving in everyday life are very significant. A comparison of Reductionist approaches to machine based cognition against these new insights about Bizarre Systems and MFMs uncovers little overlap.

These new ideas are starting to take hold and provide results in the language understanding research community. The recent TextRunner project[15] uses unsupervised discovery-based learning. Google has won prizes for their Chinese <–> English and Arabic <–> English Machine Translation software. They use Non-parametric Models[16] which is a way to use statistics on populations of unknown distribution; the distribution is discovered and lerned over time, which means it is almost a Model Free Methods. Given our understanding of Bizarre Systems and MFMs it is unsurprising that this approach can provide prize-winning performance in machine translation – a task that definitely “requires intelligence”. Nobody on Google’s Chinese translation team spoke Chinese; they used no grammars or other models of language.

---

[12] General applicability at the meta level does not imply general applicability at the solution level. MFMs are always applicable. Their results are typically highly context specific. Human minds cannot at low levels generate portable solutions the way F=ma is portable; rather, we learn very small patterns that self-assemble into larger solutions, but this is a different kind of cross-domain applicability. Every new context requires the work of matching experience to situation to be completely re-done.

[13] http://artificial-intuition.com/tradeoff.html

[14] This is discussed in a talk by Ms. Anderson available at http://videos.syntience.com/ai-meetups/modelsvspatterns.html

[15] Prof. Oren Etzioni, University of Washington http://www.technologyreview.com/computing/22773

[16] Peter Norvig, Director of Research at Google discussed these at a talk at a MeetUp in Menlo Park, CA facilitated by Ms Anderson. A video is available at http://videos.syntience.com/ai-meetups/peternorvig.html

In the same spirit, Syntience Inc. has devised a novel algorithm named "**Artificial Intuition**".[17] It is a very advanced MFM built out of simpler MFMs. It is very different from other Machine Learning algorithms in that it is capable of independently identifying and learning nested higher order patterns in spatiotemporal sequences, such as text, without being given any a priori models of structure or semantics of these sequences. It learns language the same way children learn language, without grammars or other language models.

Such systems are capable of learning human languages using only unsupervised learning, starting from a blank slate – a substrate – that can learn. Powerful Artificial Intuition based systems for research purposes have been implemented in about 15,000 lines of Java code. This compiles to an empty system which can be given e.g. Jane Austen novels to read, character by character. The substrate is designed to encourage creation and self-assembly of nested patterns that capture the semantics of the provided text and thereby learn the language. Recurring automated tests for semantic level reading comprehension can quantify the learning.

The short term goal is to surpass current state of the art approaches to a specific problem called "English Language Word Segmentation": Remove all spaces from a page of text, and then request that the system insert spaces where appropriate[18]. This test can be completely automated since it starts from the known correct result. Dictionary based approaches and Markov chain based Models get to about 98% correct for English, but no more, since results beyond this level require bona fide understanding of the language[19], including at least a partial understanding of the world. Our new approach currently attains 80% correctness which is not very impressive but has surprised observers who had not realized that the new approach could work at all.[20]

Artificial Intuition is language agnostic (since it is model free, including grammar and dictionary free). Attaining of an ability to segment English means it will be able to segment Chinese after just reading a sufficiently large segmented Chinese corpus. This is economically important since Chinese does not use spaces to separate words; web search quality in Chinese (and the ability to sell well matched advertising) is directly dependent on the ability to correctly segment Chinese text into words. A significant improvement in Word Segmentation capability would enable dominance in the very important and rapidly growing Chinese web search advertising market.[21]

Success on this problem hinges on having achieved semantic competence; but once achieved, only small modifications to the surrounding “interface” code will be required in order to solve other problems in this category; competent semantic technology would immediately provide the key to a dozen quality issues in web search including spam, porn, noise, vandalism, hate speech, and threat detection. Beyond search, it would enable perfect speech recognition, semantic spelling correctors, perfect optical character recognition, machine translation, document summarization, automated customer support systems, and personal tutoring devices. Beyond that, systems capable of acquiring world knowledge by simply reading about the world (even on the web) could provide straight answers to questions instead of just providing documents as search hits. True understanding of document content allows much more competent automated personal information assistants and topic-based browsing of automatically organized archives of content. Competent semantic technology is the key to many billions of dollars worth of commerce spread over hundreds of markets, radically transforming thousands of existing processes, corporations, and organizations, and a major advance in intellectual capabilities and efficiency benefiting people worldwide. It will also enable future AI systems to learn about the world by reading about it.

Principal Researcher:	Monica Anderson, MSCS		
Mail:	Syntience Inc. 171 Main St., Suite 276, Los Altos, CA 94022		
E-mail:	monica@syntience.com	Phone:	+1 408 802-7229
Corporate:	http://syntience.com	Theory:	http://artificial-intuition.com
Blog:	http://monicasmind.com	Videos:	http://videos.syntience.com

[17] The word “Intuition” means “the art of guessing wisely”. It is **not** a mystical power, since it is fallible, just like the algorithm. Similarly, “wisdom” is a large, well tested set of salient micro-intuitions derived from experiences gathered over a lifetime

[18] Note that this is a trivial form of language generation; ability to generate correct language proves that the systems understands it

[19] Consider the decision whether “into” should be split as “in to”, plural ‘s’ before words that might be another word that starts with an ‘s’, etc.

[20] The comparison is complicated and is not fair. 80% correctness was attained after reading a few chapters of Jane Austen, about 20,000 words. This test was also a simplified variant of the true segmentation test. Google’s 98% result using Markov Chains required a 1.7 billion word training corpus; future evaluations will include actual industry standard test suites.

[21] Syntience Inc. aims to create a module with this competence (to be licensed to use in web search engines) as one of their first products