Technology Maturity for Massively Parallel Adaptive Computing

A Workshop to Chart Architectural Readiness of Real-World Applications

Dan Hammerstrom
Electrical And Computer Engineering
Portland State University

Intelligent Computing

- In spite of the transistor bounty of Moore's law, there is a large class of problems that computers still do not solve well
- These problems involve the transformation of data across the boundary between the real world and the digital world
  - "the digital seashore" Hiroshi Ishii, MIT Media Lab
- These "boundary" problems occur wherever a computer is sampling and acting on real world data
  - Which includes almost all embedded computing applications
- Our inability to adequately solve these problems constitutes a significant barrier to computer usage and to huge potential markets

- Examples include: computer vision, speech recognition, textual and image content recognition, robotic control, advanced automotive safety, intelligent power management, and data mining (making sense of massive quantities of seemingly unrelated data), ...
- These are difficult problems that require computers to find complex structures and relationships through space and time in massive quantities of low precision, ambiguous, noisy data
- Biological systems do this to various degrees – even simple insect brains …

Intelligent Signal Processing

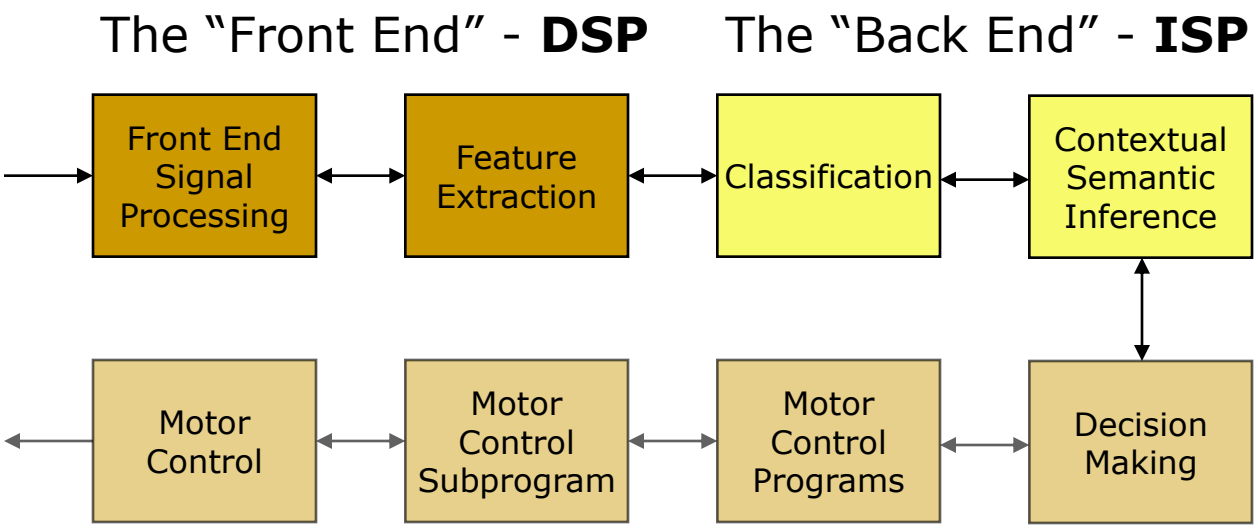
- The term Intelligent Signal Processing (ISP) has been used to describe algorithms and techniques which involve the creation, efficient representation, and effective utilization of large complex models of semantic and syntactic relationships

- ISP augments and enhances existing Digital Signal Processing (DSP) by incorporating such contextual and higher level knowledge of the application domain into the data transformation process

- Simon Haykin (McMaster U) and Bart Kosko (USC),
  - Special Issue issue of the *Proceedings of the IEEE*, “Intelligent Signal Processing,” November, 1998 (Vol. 86, No. 11)

A Prototypical Model of Intelligent Computing

- Although an over-simplification, this flow diagram is useful in discussing the functional components of this problem domain
  - Real applications tend not to partition so cleanly
  - Notice the bidirectional arrows between each block – data flow both ways
- Most IC applications have components in each of these blocks



The “Front End”

- We understand the front end pretty well, it is the realm of traditional digital signal processing (DSP)

- Front end algorithms tend to involve applying the same computation over a large array of elements, they are data parallel and communication is almost exclusively local
  - The neuromorphic, aVLSI, subthreshold CMOS, for example, has been used extensively for this kind of front end processing

  - Another example is the CNN architecture developed by Chua, Roska et al.
    - A 2D array of programmable, “analogic” processors, with local 2D grid interconnect

But Then There's The Back-End …

- In the early days of computing, "Artificial Intelligence" focused on the representation and use of contextual and semantic information
- Knowledge was generally represented by a set of rules
  - Logic operations (e.g., based on first order predicate calculus) were used to manipulate the rules and "infer" the properties and state of an environment from input data
  - Then from inferred state, logic operations then also enabled planning and decision making algorithms
- These systems were "brittle," exhibiting no real flexibility, generalization, or graceful degradation
  - Often catastrophic failure resulted if a circumstance occurred that was not covered by a rule

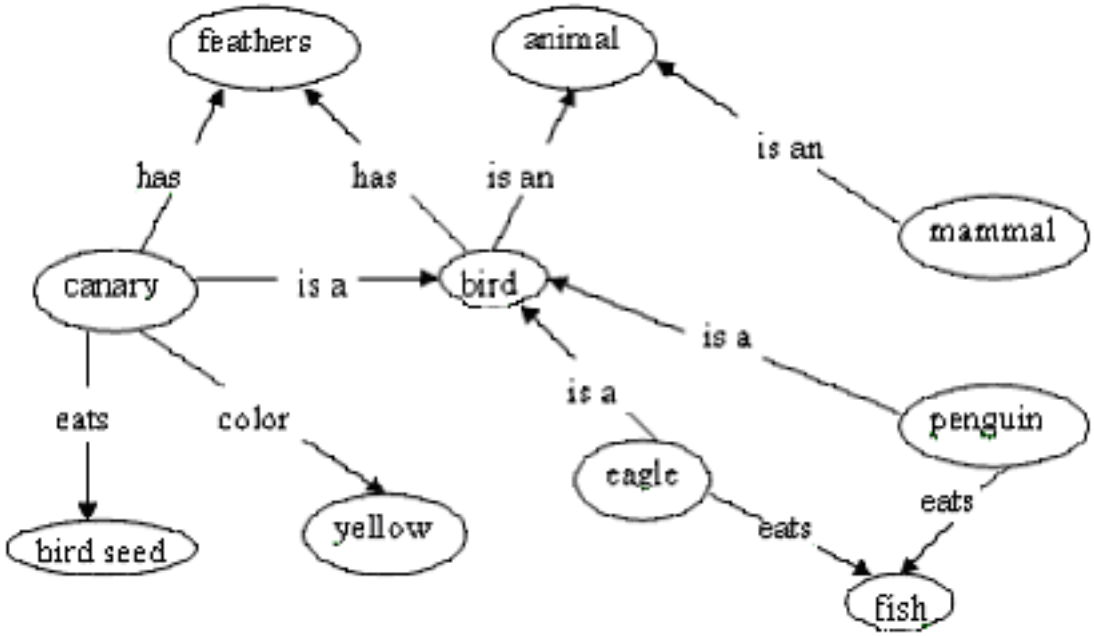
- And they were unable to adapt dynamically (learn) within the context of a real world application

- ANNs (Artificial Neural Networks) extended computational intelligence in a number of important ways, primarily by adding the ability to incrementally learn or adapt to complex non-linear functions
  - And they resulted in a whole new set of statistical and non-linear learning techniques

- However, ANNs tend to be focused on the lower levels of feature extraction and classification, and like AI before them, they have not scaled well
  - The lack of scaling also meant that traditional desktop computers (or clusters) gave the necessary performance applications and research, mostly eliminating the need for specialized hardware

What Does The “Back End” Do?

- Simplistically it captures information on “high order” relationships of “abstract” entities
- Which are represented via a graph structure
- These used to be “rules,” but more recently they have become probabilistic relationships -> “Bayesian” Networks
- Inference is then performed over these structures



Scaling

- The scaling limitations of both symbolic and traditional neural network approaches constitute one of their biggest shortcomings
- It is very likely that sheer size is a major component of the “secret sauce of cognition”
  - “Size matters!” (*Godzilla*, the movie, 1998)
  - **Consider the differences: hundreds of rules or thousands of nodes vs. billions of neurons**
- Therefore we need models and hardware/software platforms that scale to very large sizes

Learning / Adaptation

- Animals learn in real time from their surroundings and appear to do a kind of probabilistic inference from learned knowledge as they interact with their environment
- Consequently another important characteristic of real systems is incremental, integrative adaptation or learning during system operation
- In fact, systems configuration is more a process of self-organization and adaptation rather than of programming

cognition vs. Cognition

- Most mammals have neocortex
- This is particularly true of primates
- The cortices for these animals (excluding humans) are well developed and perform a fairly sophisticated “cognition” (with a small ‘c’)
  - Monkeys can do planning and learn to effect complex social behaviors

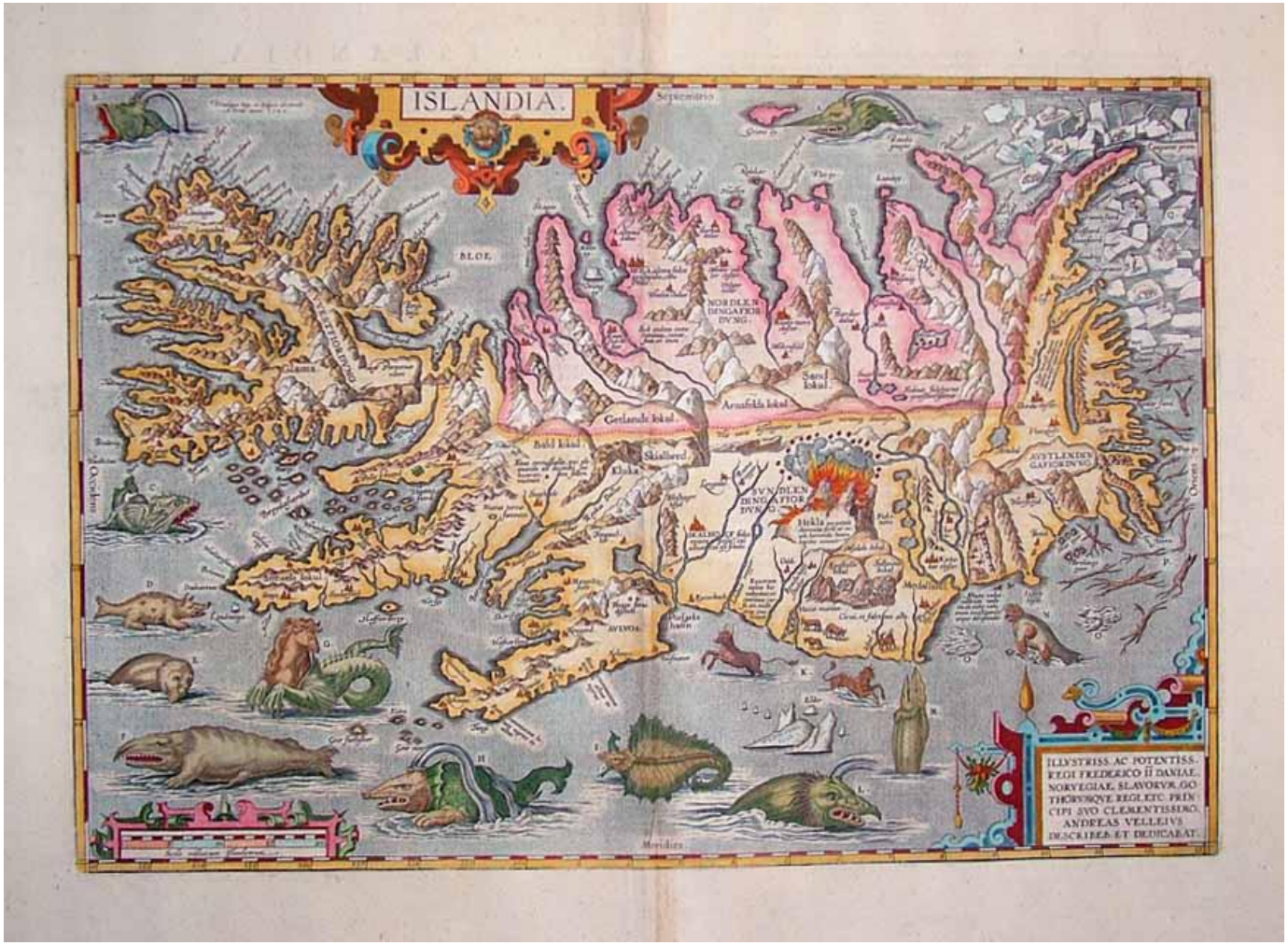
- Yet Humans, on the other hand, have an additional capability for reasoning, planning, and, in particular, for language – “Cognition” (with a capital ‘C’)
- Many researchers confuse this distinction, often referring to cognition when they really mean Cognition and vice-versa

- Traditional Artificial Intelligence, among other things, primarily went after Cognition without having solved cognition
- We believe that we need to solve cognition before we can address the Cognition problem and are focused on that
- And it is hard enough

ISP Toolbox – Still empty after all these years …

Now What?

- Neither AI, Artificial Neural Networks, Fuzzy Logic, nor Bayesian Networks have led to powerful, scalable ISP
- One potential, and maybe the only, source of inspiration for solutions to these challenges is biology
- Though there is a wide range of “natural” computation from genes to cells, the one that is most relevant here deals with neural circuits
- Now we enter the world of Computational (or Systems) Neuroscience



Hic sunt dracones! (Here be dragons)

- There has been remarkable progress in neuroscience in the 20 years since the second coming of ANNs
- This includes a sophistication in modeling and a systems orientation that did not exist in the field before that
- In fact, I believe that this is the most important result to come out of the ANN research efforts of the 80s and 90s
- Some of this is due to the computation bounty of Moore's law, and some of it to the movement of a number of physicists into neuroscience, who brought with them sophisticated mathematical tools, but there has also been a significant a cultural shift in neuroscience itself

The Ultimate Cognitive Processor: Mammalian Cerebral Cortex

Starting at the top …

- Neocortex is a folded, 2D, planar sheet
- In humans it is, when stretched out, about 2/3s of a square meter, 3-4 millimeters thick, and consists of roughly 30B neurons
- It has a fairly consistent 6 layer structure and is remarkably uniform, not only across all different parts of human cortex, but across almost all mammalian species
- Although we are a long ways from understanding the details of how it does what it does, some of the basic computations are beginning to take shape

- Nature has, so it appears, produced a general purpose computational device that is a fundamental component of higher level intelligence
  - Rewired cortex has been shown to adapt to any sensory input it is given
  - For example, for many blind people, the visual cortex ends up adapting and being used for reading Braille, a non-visual task
- Consequently in collaboration with the neuroscience community, several groups are looking to create increasingly more sophisticated abstract models of cortex, and then to apply these models to real applications
- In today's workshop you will hear about some of these efforts

- Neural models are asynchronous, low precision, stochastic, massively parallel, fault tolerant, and adapt rather than being programmed
- As we move to increasingly smaller transistors / devices, ultimately at the molecular scale, stochastic effects become ever pronounced,
- Consequently circuits and architectures based on neural models will be a better match to such silicon structures than traditional discrete, synchronous, deterministic, fault-intolerant computational models

Robert Noyce, Co-founder of Intel and Co-inventor of the Integrated Circuit ("The Next 100 Years," IEEE Technical Convocation 1984):

"Until now we have been going the other way; that is, in order to understand the brain we have used the computer as a model for it. Perhaps it is time to reverse this reasoning: to understand where we should go with the computer, we should look to the brain for some clues."

Sponsors

- Intel Corporation
- Portland State University
- The National Science Foundation
- The Office of Naval Research