Proposers’ Day Workshop

Monday, January 23, 2017

@srcJUMP, #JUMPpdw
Cognitive Computing

Mandy Pant
Academic Research Director
Intel Corporation
Today’s deep learning is still limited

Deep understanding of distributed architectures with brain like capabilities to optimize available resources

Current technologies do not scale to “brain-like” levels in terms of energy and performance

Create systems for boosting economic competitiveness and contributing to social good
Research Considerations

• Seeking **novel solutions** for cognitive systems that can
  • Learn at scale
  • Perform reasoning and decision making with purpose
  • Interact with humans naturally

• Explore alternatives to von Neumann model like
  • Stochastic computing
  • Approximate computing
  • Bio/brain-inspired models such as neuromorphic computing

➤ **A full-system approach required to achieve goals**
  • Include technology advances for fundamental leaps in performance, capabilities, and energy efficiency through improvements in programming paradigms, algorithms, architectures, circuits, and device technologies

**Grand Challenge:** Proposers expected to define a grand challenge in the Cognitive Computing space that the Center will address
Theme Requires

• Creating systems that, without explicit objectives, operate in natural world on their own
  • Exploration of multiple approaches for building machine intelligent systems with both cognitive and autonomous characteristics using von-Neumann, non-von-Neumann or combination of both

• **Fundamental improvements must be demonstrated** in performance, capabilities
  • Energy efficiency through breakthroughs (from devices to programming models)

• Benchmarking to prove potential of novel system
  • More than 100x better than traditional computing for same task
  • Demonstration of complete scalable compute system (with validation)

• **Research must span a gamut** of materials and devices, circuits and architecture, algorithms and programming theory
  • Cross-disciplinary effort
Research tasks of interest includes but are not limited to:

- **Knowledge generation, reasoning, prediction, planning, decision making:**
  - Create insights, actionable intelligence, knowledge, wisdom and machine assisted decision from raw sensor data
  - Anomaly detection for prediction, decision making, planning
  - **Architectures to accelerate tactical decision-making**
    - Computation on encrypted data to support safety of scalable decision making systems

- **Learning and perceiving:**
  - Pursue breakthrough advances in sensing technologies for cognitive systems
  - Develop universal algorithms using e.g. context-aware learning applicable to a wide range of cognitive workloads
    - Develop perception algorithms (which enable understanding of the environment), authentication enablers, and learning enablers
  - Develop hardware (architectures/circuits/technologies) and software to realize said universal learning system
Possible Research Tasks
Advancements in Cognitive Capabilities (2)

• Learning and perceiving (cont.):
  • Explore **fundamental limits and capabilities** of existing machine learning algorithms, for instance, deep neural network, for cognitive workloads
  • Develop new architectures and algorithms that allow hitting required test accuracy with **significantly reduced training set**
    • Handle big data and distributed analytics
    • Develop algorithmic and hardware speedup for sparse coding
    • Develop highly efficient inference systems
    • Develop energy-efficient, low-cost techniques for cognitive workloads

• Human-Machine Interface:
  • Develop **seamless Human-Machine Interface** for autonomous systems
  • Develop seamless Human-Machine Interface/collaboration for **sensing and feedback** (haptic, audio, vision, etc.) with high accuracy in HW and SW levels
Possible Research Tasks
Advancements in Cognitive Capabilities (3)

• Network of cognitive sub-systems:
  • Develop architectures for resilient self-optimizing and self-healing networks, memories and compute elements for connecting billions of devices in intelligent systems
  • Develop society-scale applications and intelligent data collections systems, which can interact with local cognitive and autonomous machine intelligence systems to optimize decision support

• Other Important Aspects:
  • Develop programming paradigms and languages that enable the cognitive capabilities
  • Investigate nature of malicious and/or destructive AI systems
Non-traditional computing embracing variations/randomness can provide significant benefits in energy, delay, error rate or enable scalable architectures on erroneous hardware layers

- Statistical computing
- Approximate computing
- Stochastic computing
- Shannon inspired computing

Applications:
- Cognitive/non-cognitive, logic based examples
  - Computation in projected spaces (increased and reduced dimensionality)
  - Application of Restricted isometry and dimensionality reduction for computing and sensing
  - Computing for classification and clustering

**Research tasks include but not limited to**

- Theoretical thrusts to address
  - Computational complexity (time and memory) of stochastic and random computing
  - Complexity of associate processors
  - Scalability

- Computing at variable accuracy covering both HW and SW
- Computing with device and hardware dynamic variations
- Statistical error correction
- Algorithmic noise tolerance
- Compressive sensing paradigms
Non-traditional computing founded on neuromorphic/neural networks theory can significantly advance energy efficiency, accuracy, scalability, and latency in cognitive computing

- Neural networks
- Neuromorphic/Brain-inspired (check definition; difference)

Applications
- Cognitive, non-cognitive, logic based examples

Research tasks include but not limited to:

- Develop the computational theory for various neural nets/neuromorphic/brain-inspired computing methods (enables applications to be implemented on h/w)
  - Computational Complexity (time and memory)
  - Self-Learning
  - Scalability

- Develop energy-efficient low-cost techniques

- Develop new architectures and algorithms that would enable improved accuracy using significantly reduced training

- Develop reconfigurable networks suitable for a class of NN applications (ANN, RNN, CNN, etc.)

- Develop materials and devices designed and optimized for neural network applications, such as artificial neurons and synaptic memory devices

- Mimic actions of a neuron through new memory elements and circuits,

- New topologies like spiking neural-nets and inference augmented neural-nets

- Develop compressive paradigms for classifiers and computing for classification and clustering
Demonstrate practical mechanisms for integration and verification of any new non-traditional methods with Von Neumann (traditional) systems to support realistic, complex workflows

Research tasks include but not limited to:

• The theoretical basis and support for non-traditional methods
• Breadth of applicability of proposed non-traditional methods
• Integration into relevant workloads and systems.
• Resiliency and reliability of proposed non-traditional methods
While Centers are not limited to students from these disciplines, graduate students working towards degrees in following areas of study are of particular interest

**MS / PhD**
- Applied Math, Mathematics
- Applied Physics, Physics
- Bioengineering
- Computational Neuroscience, Neuroscience
- Comp Sci. & Comp. Eng.
- Electrical / Electronics Engineering or Electrical and Computer Engineering
- Material Science or Engineering
- Systems Engineering