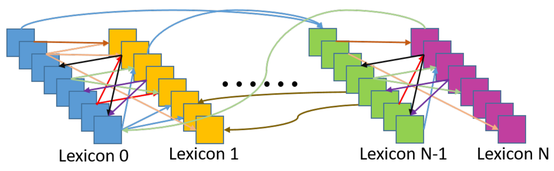
Probabilistic Inference in Neuromorphic Architecture: Applications and Implementations

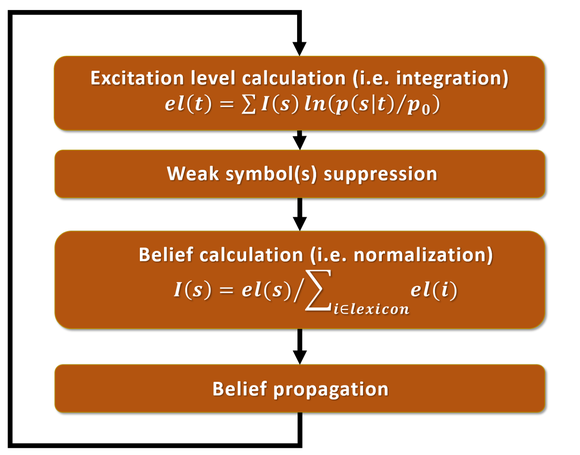
Zhe Li, Yanzhi Wang, Qinru Qiu
Department of Electrical Engineering and Computer Science, Syracuse University

Overview

- Connection-based cognitive computing model imitating neocortex
- Captures probabilistic distribution and correlations between features at the symbolic level as ***knowledge base***
- Collection of symbols is divided into categories known as ***lexicons***
- Symbols within a lexicon inhibit each other, at the same time excite symbols of different lexicons ***(belief propagation)***
- Support parallel implementation and real-time inference



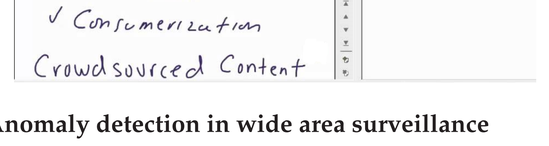
Example of lexicons, symbols and knowledge links



Algorithm flow for probabilistic inference

Applications

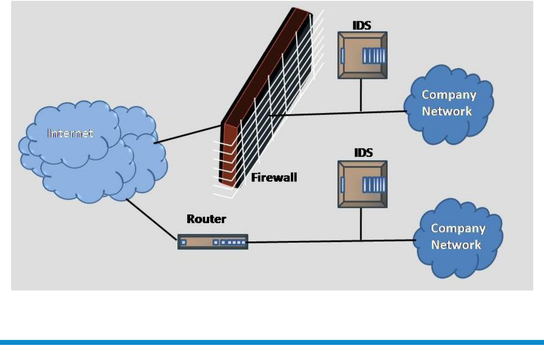
Document text recognition



Sentence completion

口号特别(震/振)撼人心(Slogan $\frac{excites}{shakes}$ people's mind)
锤炼(得/的)更坚强($\frac{temper\ it\ to\ be}{tempered}$stronger)
Nothing ? suspected in boy's ? on world's tallest water slide
Nothing ***criminal*** suspected in boy's ***death*** on world's tallest water slide

Offline Handwriting Recognition



Anomaly detection in wide area surveillance

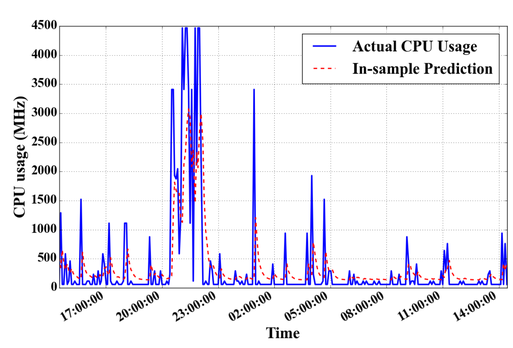
Network intrusion detection

Applications

Malicious program detection

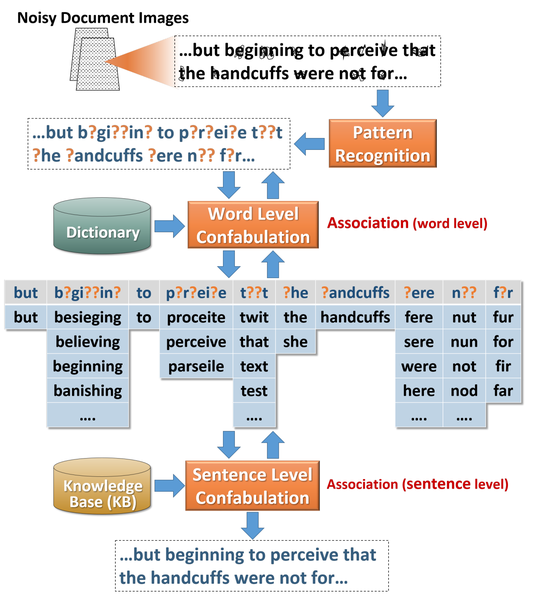
```
crash> bt -f
PID: 0      TASK: 1011000e030       CPU: 1   COMMAND: "sw
 #0 [101118b9e70] schedule at ffffffff8030bc9d
    101118b9e78: ffffffff8030bcf5
 #1 [101118b9e78] thread_return at ffffffff8030bcf5
    101118b9e80: 0000010208f317f0 0000000000000073
    101118b9e90: 000001011005e760 0000000000000000
    101118b9ea0: 0000000000000000 0000000000000000
    101118b9eb0: 0000000000000000 0000000000000001
    101118b9ec0: 0000000000000407 0000000000000002
    101118b9ed0: 0000000000000000 0000000000000008
    101118b9ee0: 00000101118b8000 0000000000000000
    101118b9ef0: 0000000000000018 ffffffff8010e749
    101118b9f00: 000001011000e030 000001011005e760
    101118b9f10: fffffffffffffa9 ffffffff8010e769
 #2 [101118b9f18] default_idle at ffffffff8010e769
    101118b9f20: 0000000000000010 0000000000000246
    101118b9f30: 00000101118b9f48 0000000000000018
    101118b9f40: 0000000000000001 ffffffff8010e7dc
 #3 [101118b9f48] cpu_idle at ffffffff8010e7dc
```

Virtual Machine resource usage prediction



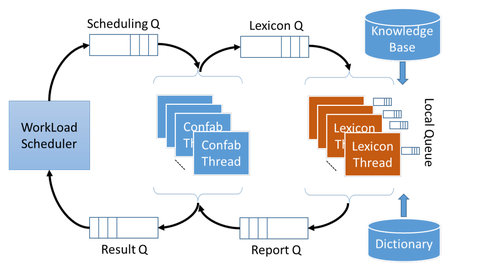
Implementations

Intelligent Text Recognition System (ITRS)



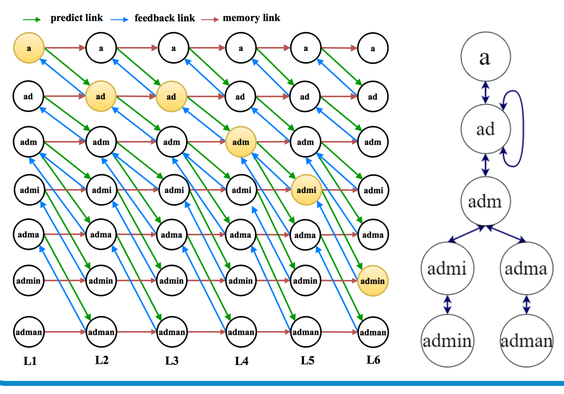
1. Pattern recognition recognizes text images with its best effort.
2. The word level confabulation provides all possible words that can be formed based on the recognized characters.
3. The sentence level confabulation finds the combination among those words that gives the most meaningful sentence.

Parallel Implementation for sentence confabulation



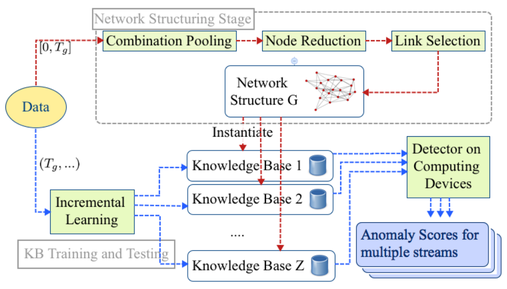
1. By introducing randomness from the race conditions in parallel processing, runtime is reduced and accuracy is improved
2. Further improvement is achieved by scheduling the lexicon processing intermittently
3. Reduced 93.4% computation time and 5% improvement in recall accuracy

Merge temporally repeated symbols to form recurrent network for sequence detection and recall



Implementations

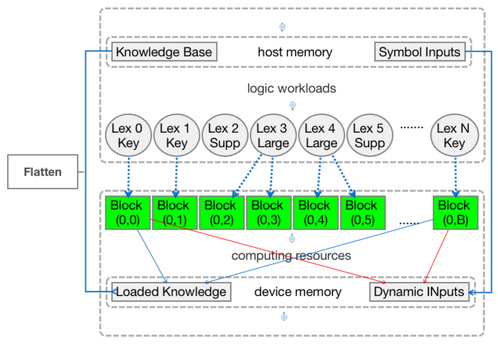
Autonomous Anomaly Reasoning and Detection (AnRAD) Framework



Self-structured network for AnRAD

1. Push the complexity from network architectures to the initial structuring stage and feature space
2. Simple network pipeline for online training and high concurrency
3. Combination Pooling:select relevant feature composition
4. Node Reduction:construct representative lexicons in feature space
5. Link Selection:connect correlated lexicons together
6. Network size reduction by 99%

GPU acceleration for Probabilistic Inference

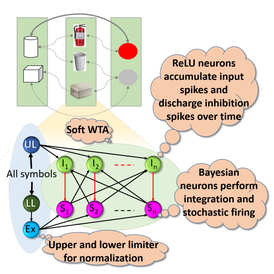


1. Dynamic workload mapping from lexicons to CUDA blocks
2. Fine-grained parallelism that maps each knowledge link to a CUDA thread
3. 1000X speedup to CPU version

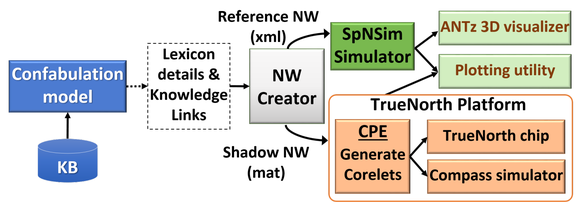
Neuromorphic Implementation

Computation Model of Spiking Neural Network (SNN)

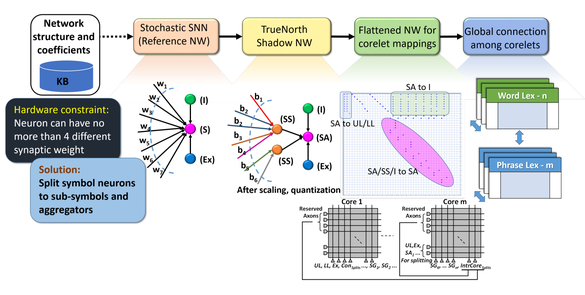
Computation Model	SNN model
Lexicons	Subnetwork
Symbols	Bayesian Neurons
Belief propagation	Excitatory links
Suppression	Soft winner take all (WTA) circuit
Normalization	Normalized WTA circuit



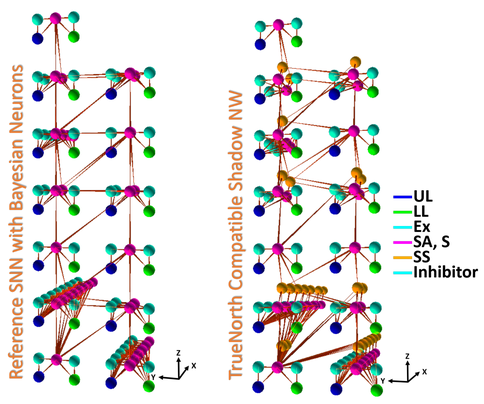
Design Environment for SNN



Mapping to TrueNorth Neurosynaptic Processor

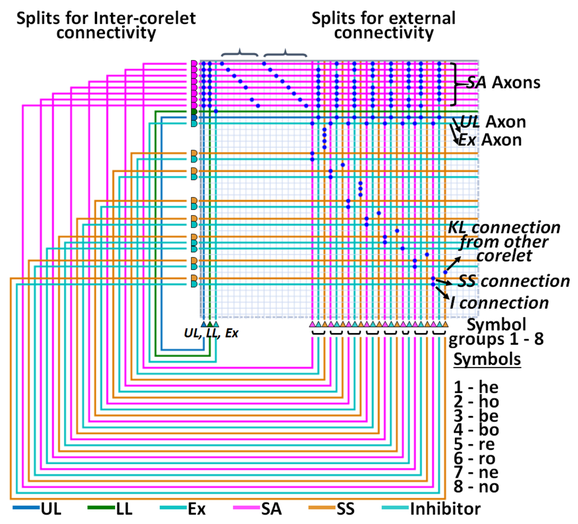


3D visualization of Examples of Spiking Neuron Networks



Neuromorphic Implementation

An example of TruNorth crossbar assignment and connectivity for one lexicon in sentence confabulation



Sentence Confabulation Testbench

- Input:multiple words from fuzzy image recognition
- Goal:grammatically and semantically correct sentence
- Accuracy: 88% Power: 0.205 mW (at 0.8V)

Reference

[1] Q. Qiu, Q. Wu, M. Bishop, R. E. Pino, and R. W. Linderman, "A parallel neuromorphic text recognition system and its implementation on a heterogeneous high-performance computing cluster," *Computers, IEEE Transactions on*, vol. 62, no. 5, pp. 886–899, 2013.

[2] Q. Qiu, Q. Wu, D. J. Burns, M. J. Moore, R. E. Pino, M. Bishop, and R. W. Linderman, "Confabulation based sentence completion for machine reading," in *Computational Intelligence, Cognitive Algorithms, Mind, and Brain (CCMB), 2011 IEEE Symposium on*. IEEE, 2011.

[3] Z. Li and Q. Qiu, "Completion and parsing chinese sentences using cogent confabulation," in *Computational Intelligence, Cognitive Algorithms, Mind, and Brain (CCMB), 2014 IEEE Symposium on*. IEEE, 2014.

[4] F. Yang, Q. Qiu, M. Bishop, and Q. Wu, "Tag-assisted sentence confabulation for intelligent text recognition," in *Computational Intelligence for Security and Defence Applications (CISDA), 2012 IEEE Symposium on*. IEEE, 2012, pp. 1–7.

[5] Q. Qiu, Z. Li, K. Ahmed, W. Liu, S. F. Habib, H. H. Li, and M. Hu, "A neuromorphic architecture for context aware text image recognition," *Journal of Signal Processing Systems*, pp. 1–15, 2015.

[6] Q. Chen, Q. Qiu, Q. Wu, M. Bishop, and M. Barnell, "A confabulation model for abnormal vehicle events detection in wide-area traffic monitoring," in *2014 IEEE International Inter-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA)*. IEEE, 2014, pp. 216–222.

[7] Q. Chen, Q. Wu, M. Bishop, R. Linderman, and Q. Qiu, "Self-structured confabulation network for fast anomaly detection and reasoning," in *2015 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2015, pp. 1–8.

[8] Z. Li, Q. Qiu, and M. Tamhankar, "Towards parallel implementation of associative inference for cogent confabulation," in *2016 IEEE High Performance Extreme Computing Conference (HPEC '16)*. IEEE, 2016.

[9] K. Ahmed, A. Shrestha, Y. Wang, and Q. Qiu, "System design for in-hardware stdp learning and spiking based probablistic inference," in *VLSI (ISVLSI), 2016 IEEE Computer Society Annual Symposium on*. IEEE, 2016, pp. 272–277.

[10] K. Ahmed, A. Shrestha, Q. Qiu, and Q. Wu, "Probabilistic inference using stochastic spiking neural networks on a neurosynaptic processor," in *Proc. of International Joint Conference on Neural Networks (IJCNN)*, 2016.

[11] K. Ahmed, A. Shrestha, and Q. Qiu, "Simulation of bayesian learning and inference on distributed stochastic spiking neural networks," in *Proc. of International Joint Conference on Neural Networks (IJCNN)*, 2016.

[12] Y. Li, Z. Li, and Q. Qiu, "Assisting fuzzy offline handwriting recognition using recurrent belief propagation," in *Computational Intelligence for Human-like Intelligence, 2016 IEEE Symposium on*. IEEE, 2016.

Contact Information

Qinru Qiu: qiqiu@syr.edu

Yanzhi Wang: ywang393@syr.edu

Zhe Li: zli89@syr.edu