Distributed Shared Memory and Machine Learning

CSci 8211
Chai-Wen Hsieh
11/5/2018

Agenda

Distributed Shared memory

- Architecture: Shared Memory & Distributed Shared Memory

Machine Learning

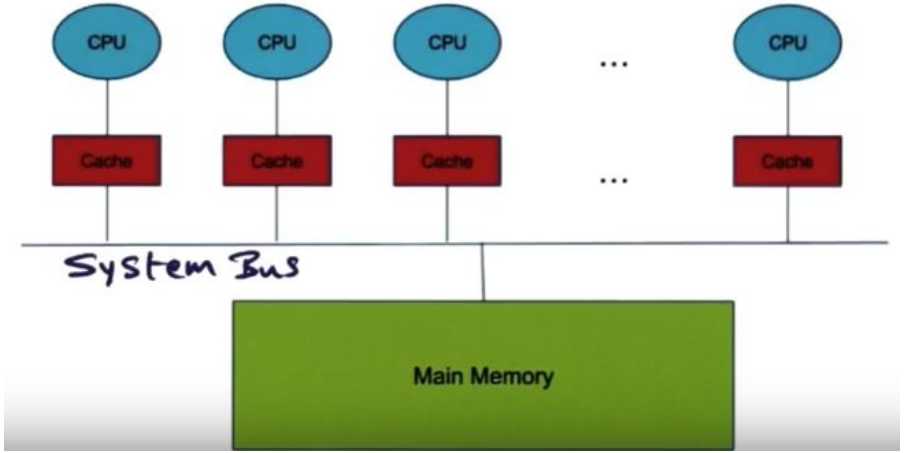
- Supervised, Unsupervised Training
- Gradient Descent
- Model/Data Parallelism

Topics

- Problems We Could Solve
- Distributed Shared Memory
- Deep Learning & DSM

Architecture - Shared Memory

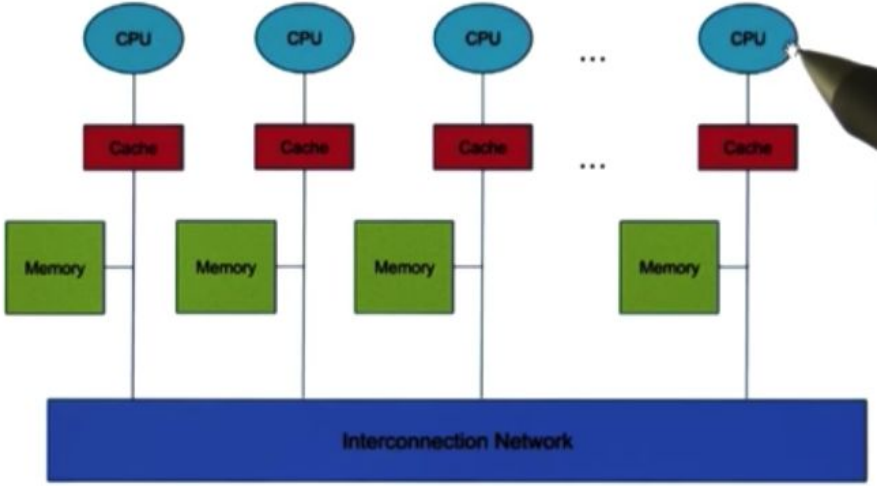
- Sharing one memory among several processors
- Communication through shared variables
- Architectures
  - SMP
  - NUMA
  - COMA



From [Advanced Operating Systems] - Udacity

Architecture - Distributed Shared Memory(DSM)

- Multiple independent processing nodes with local memory modules
- Models:
  Message Passing v.s. DSM
- Hidden data movement
- Locality of reference
- Provides large virtual memory space
- Cheaper than multiprocessor system
- Unlimited number of nodes



From Advanced Operating Systems - Udacity

DSM Issues

- Rewrite to shared memory aware program
- Cache coherence problem - maintaining coherence among several copies of data item
- Performance loss
  - Network
  - Synchronization: lock, barrier
- Failure of nodes
- “Shared memory machines scale well when you don’t share memory”
  -- Chuck Thacker

Machine Learning

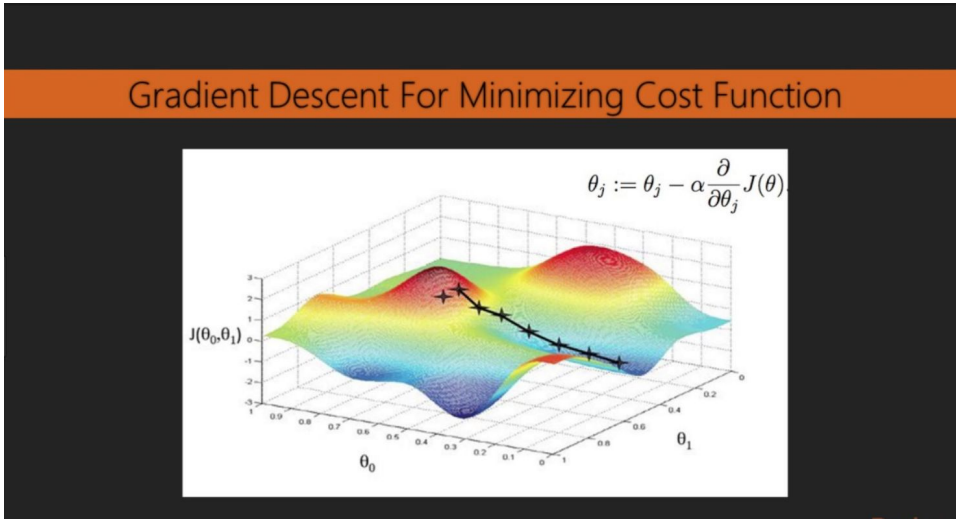
Supervised Learning

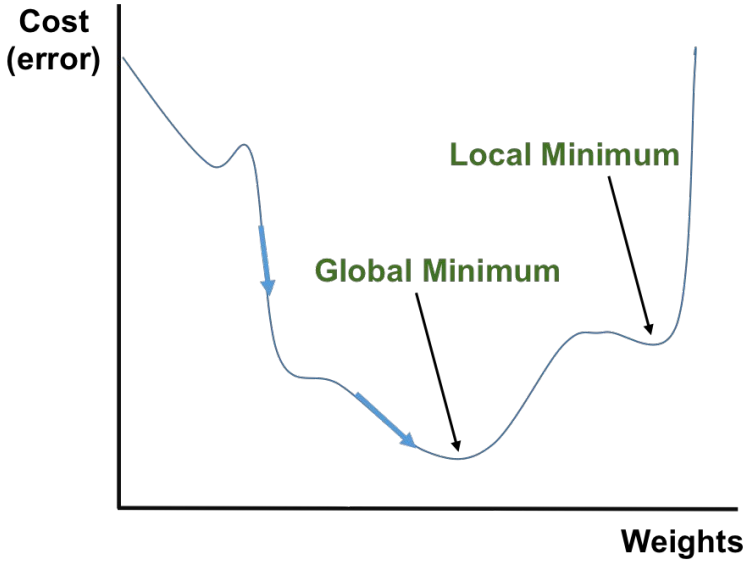
- Have input variables (X) and an output variable (Y) and you use an algorithm to learn the mapping function
- Problems:
  - Classification
  - Regression

Unsupervised Learning

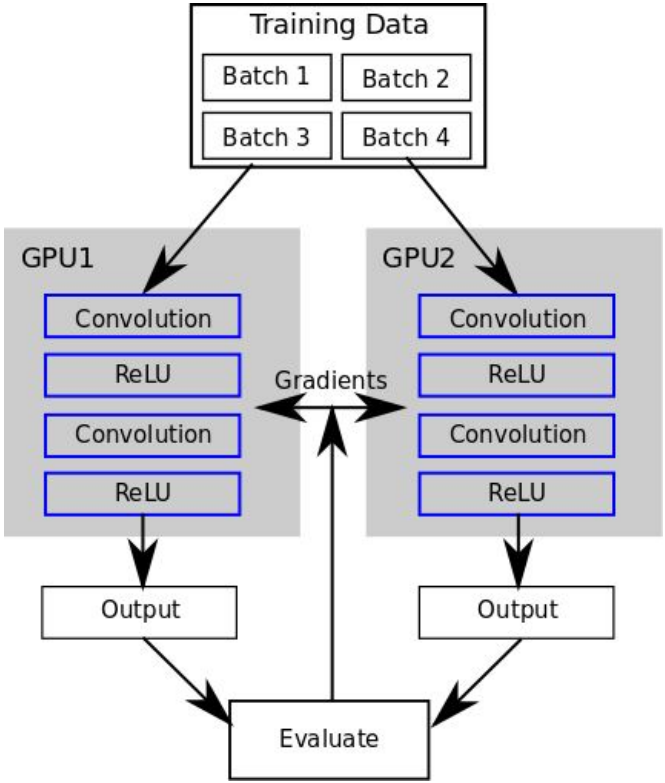
- Only have input data (X) and no corresponding output variables
- Problems:
  - Clustering
  - Association

Deep Learning - Gradient descent

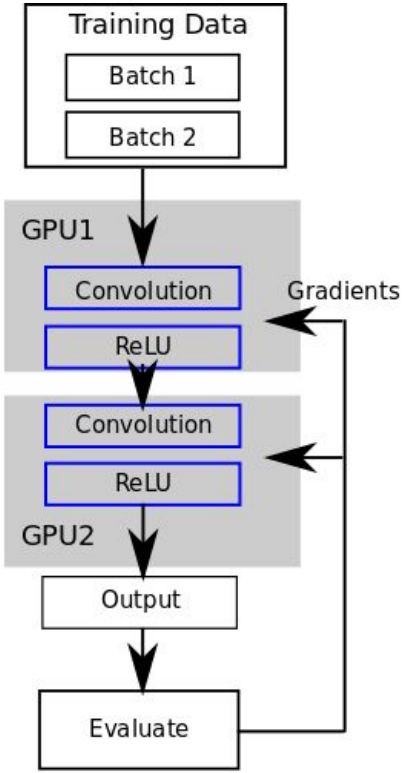




Multi-node Strategy: Data/Model Parallelism



Data Parallelism



Model Parallelism

Problems We Could Solve

1. Design a distributed shared memory framework that benefits machine learning training
2. Rewrite existing serial programs into parallel programs with ML
3. Adding nodes to a running system, where and when
4. Reduce overhead by prefetch, redistribution

需要選一個topic focus on it. Go deeper

Topics - Distributed Shared Memory

1. Z. Tasoulas, I. Anagnostopoulos, L. Papadopoulos and D. Soudris, "**A Message-Passing Microcoded Synchronization for Distributed Shared Memory Architectures**," in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*.
2. Fresno, J., Barba, D., Gonzalez-Escribano, A. et al. Int J Parallel Prog (2018). **HitFlow: A Dataflow Programming Model for Hybrid Distributed and Shared-Memory Systems.** https://doi.org/10.1007/s10766-018-0561-2
3. Yuji Tamura, Doan Truong Th, Takahiro Chiba, Myungryun Yoo, Takanori Yokoyama, **A Real-Time Operating System Supporting Distributed Shared Memory for Embedded Control Systems**, Information Science and Applications 2017. ICISA 2017. Lecture Notes in Electrical Engineering, vol

Topics - Deep Learning & DSM

1. Probir Roy, Shuaiwen Leon Song, Sriram Krishnamoorthy, Abhinav Vishnu, Dipanjan Sengupta, and Xu Liu. 2018. **NUMA-Caffe: NUMA-Aware Deep Learning Neural Networks.** ACM Trans. Archit. Code Optim. 15, 2, Article 24 (June 2018), 26 pages. DOI: https://doi.org/10.1145/3199605

2. Shinyoimg Ahn, Joongheon Kim, and Sungwon Kang. 2018. **A novel shared memory framework for distributed deep learning in high-performance computing architecture**. In Proceedings of the 40th International Conference on Software Engineering: Companion Proceedings (ICSE '18). ACM, New York, NY, USA, 191-192. DOI: https://doi.org/10.1145/3183440.3195091

Topics - Deep Learning & DSM - cont’

1. Amin Tootoonchian, Aurojit Panda, Aida Nematzadeh, Scott Shenker. 2018. **Tasvir: Distributed Shared Memory for Machine Learning.** SysML Conference. http://www.sysml.cc/doc/214.pdf

2. Wei Jinliang, “**Efficient and Programmable Distributed Shared Memory Systems for Machine Learning Training**”, PhD dissertation, Carnegie Mellon University, 2018.