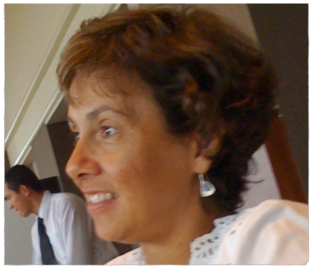


Machine Learning Applied to Systems Research

Una-May O'Reilly

Evolutionary Design and Optimization Group

EvoDesignOpt

CSAIL

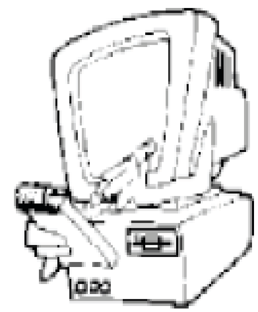
Agenda

- **My group's technical focus**
  - **Bio-inspired intelligence … evolutionary algorithms**
    - » **Machine learning: regression, classification, optimization**
  - **Different "application" areas**
    - » **Flavor design, wind resource prediction, wind farm layout optimization, network coding, analog CAD …**
    - » **Systems:**
      - **meta-heuristic optimization**
      - **auto-tuning**
- **Optimizing Sparse Matrix Algebra**
- **Learning Quality of Service Models for VMs**

New Approaches… with Artificial Intelligence

Bio-Inspired Intelligence

$$A = \begin{array}{c} \\ d_{-1} \\ d_{-2} \\ \cdot \\ \cdot \\ \cdot \\ d_{1-n} \end{array} \begin{array}{c} d_0 \quad d_1 \quad d_2 \quad \cdot \quad \cdot \quad \cdot \quad d_{n-1} \\ \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdot & \cdot & \cdot & a_{1n} \\ a_{21} & a_{22} & a_{23} & & & & \cdot \\ a_{31} & a_{32} & a_{33} & & & & \cdot \\ \cdot & & & \cdot & & & \cdot \\ \cdot & & & & \cdot & & \cdot \\ \cdot & & & & & \cdot & \cdot \\ a_{n1} & \cdot & \cdot & \cdot & \cdot & \cdot & a_{nn} \end{bmatrix} \end{array}$$



Machine Learning

Evolutionary Algorithms
Genetic Algorithms,
Genetic Programming

EvoDesignOpt

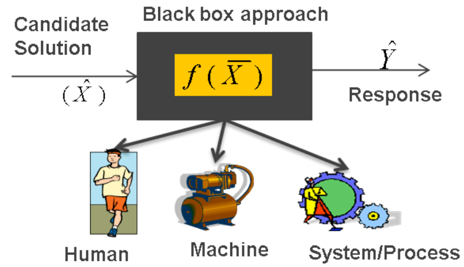
Evolutionary Design & Optimization Group

CSAIL

Why Evolutionary Machine Learning?

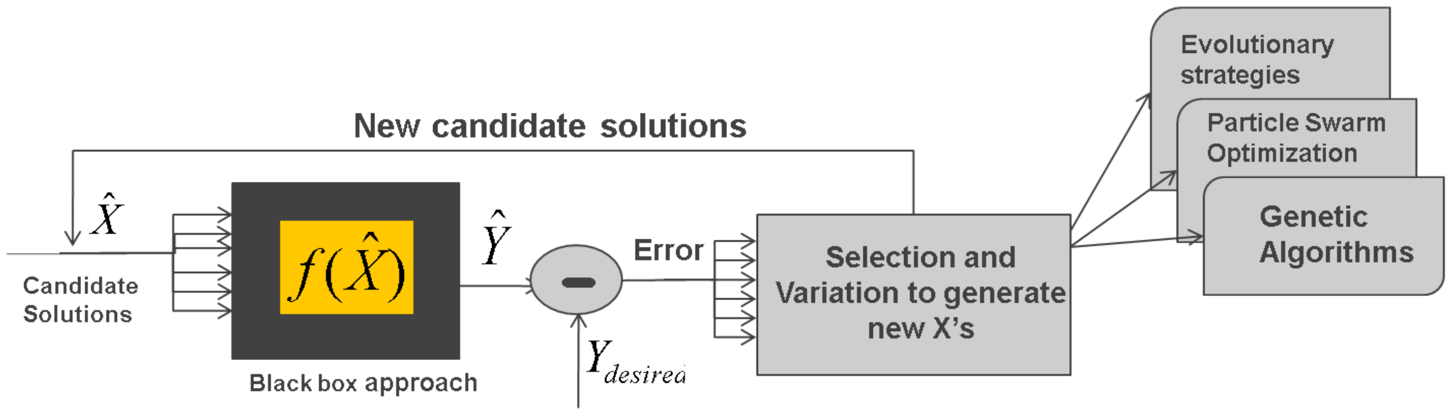
- Optimization Goal : $\arg\min_{\hat{x}} f(\hat{X})$
- Traditional methods
  - Gradient based methods
  - Convex Optimization
  - Linear programming
- Types of problems
  - Continuous valued
  - Integer problems
  - Combinatorial problems
- Cannot work with
  - Non differentiable functions
  - When no analytic expression is available
  - Non convex
  - Large scale complex systems

- Black box approach
- Ability to model and optimize
  - Systems with human in loop
  - Machines
  - System of systems



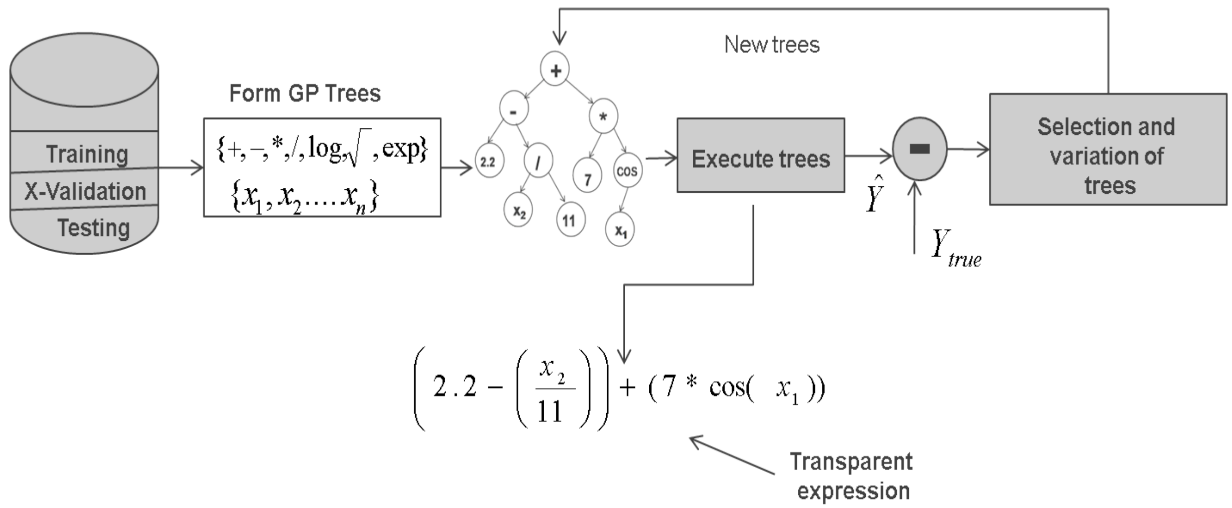
- Other Advantages
  - Parallelizable
  - No gradient needed

Evolutionary Optimization



Evolutionary Regression and Classification

FlexGP
EvoDesignOpt



Symbolic regression

- Similar to Neural Nets
- Iteratively reduces the errors by choosing better solutions
- The output is a mathematical expression that captures non-linear interactions
- The final solution is transparent!
- Capability to produce many alternate explanations

Optimizing Sparse Matrix Algebra

“Smart”

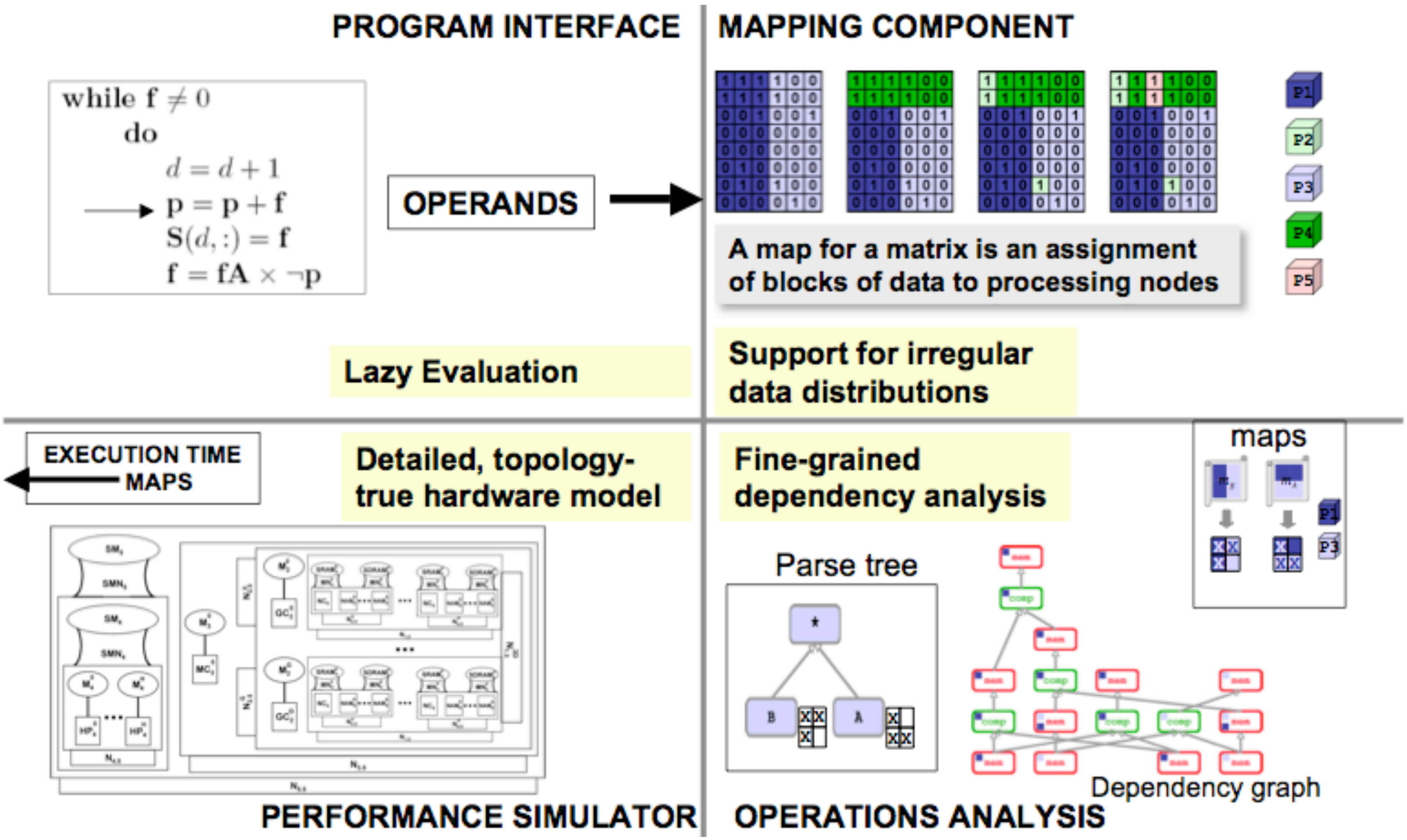
“More”

Project

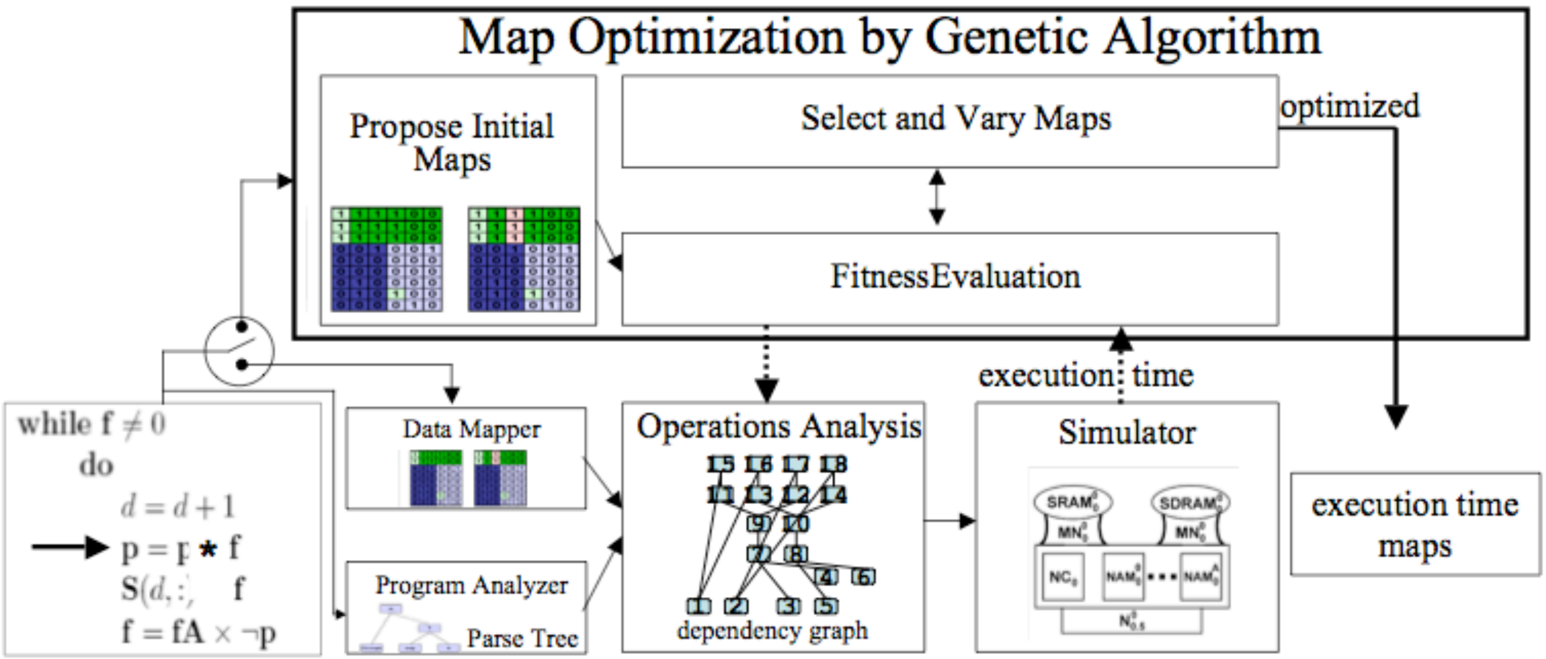
2007-2010

Evolutionary Design & Optimization Group

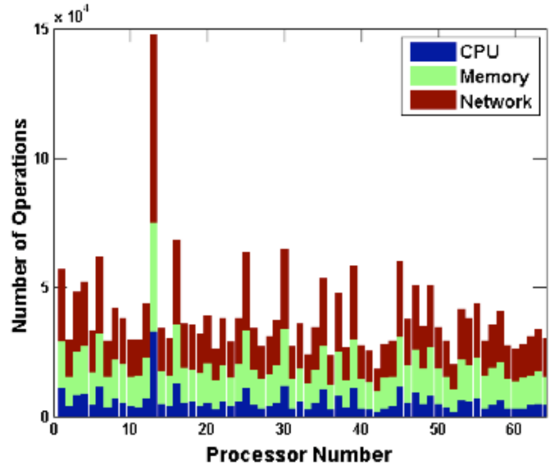
MORE Framework

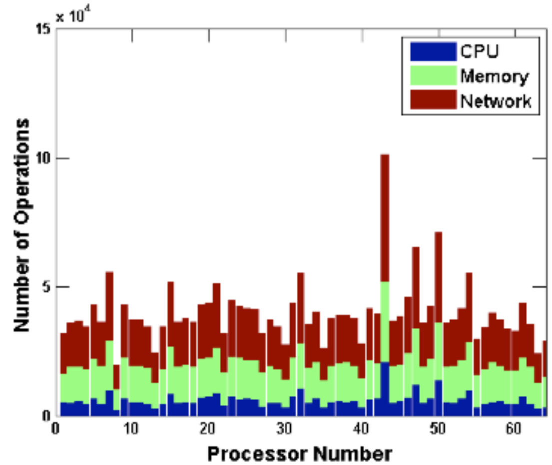


Learning High Performance Maps



Mapping Operator Balance Results





Random Map | **Optimized Map**

	Mean (SD) Best of Run (OP/s)	Best of Runs (OP/s)	Relative to ADBC
ADBC	3.17E+09 (6.39E+08)	3.71E+09	1.0
RANDMU	8.06E+09 (3.23E08)	8.68E+09	2.54X
BALANCINGMU	9.47E+09 (1.70E08)	9.78E+09	2.99X
BALANCINGMU + RANDSWAP	9.56E+09 (1.45E08)	9.92E+09	3.01X

Machine Learning for Virtualization

Figure 4. Hosts, Clusters, and Resource Pools

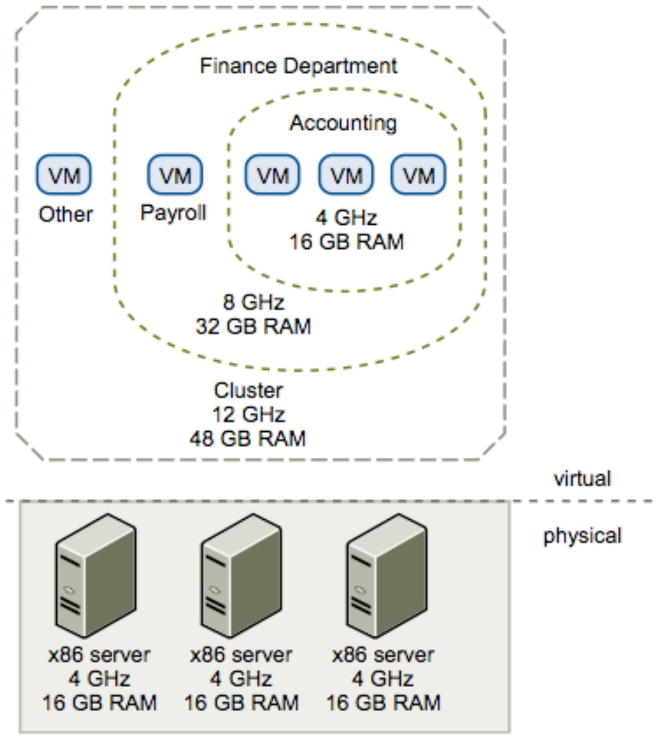
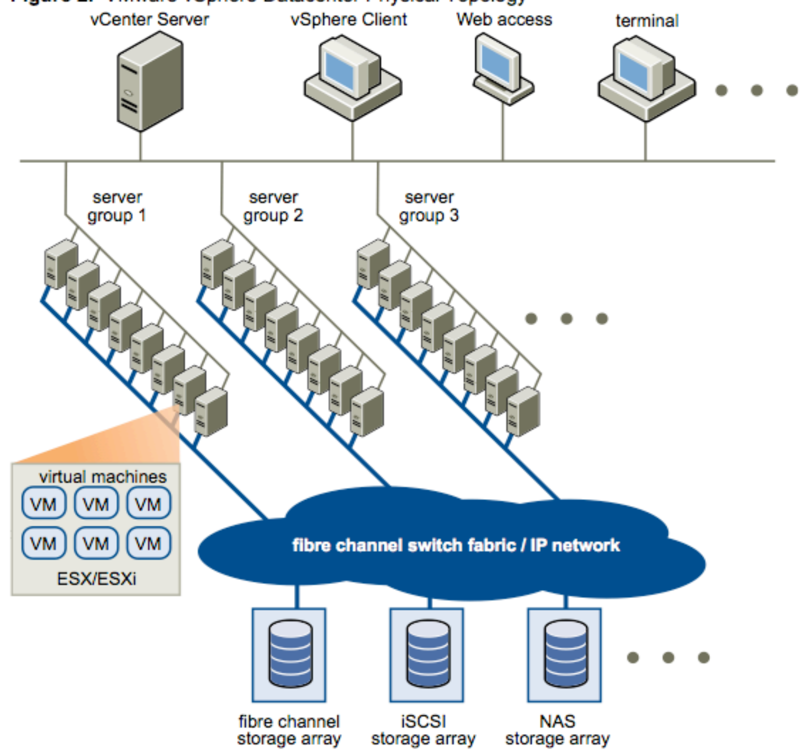


Figure 2. VMware vSphere Datacenter Physical Topology



Evolutionary Design & Optimization Group

Detecting Breakdown Points of VMs

Current Paradigm

- Client demands a certain web server response time
- Sysadmin heuristically constructs a resource allocation configuration that should satisfy the SLA
- Generous resource overprovisioning because:
  - Service levels are complex and hard to model
  - Resource allocations are static
  - Resource sharing is difficult to optimize manually
  - Shared resources don't translate linearly to service levels

Can we do better?

- Model application performance/SLA from resource use
- Potentially a means toward dynamic resource allocation
  - Throttling each VM for power savings
  - migration out – to give more resources
  - migration in – consolidation for power savings
  - Replication to increase throughput