Online Discovery of Feature Dependencies

Alborz Geramifard - June, 2011
agf@mit.edu

Joint Work

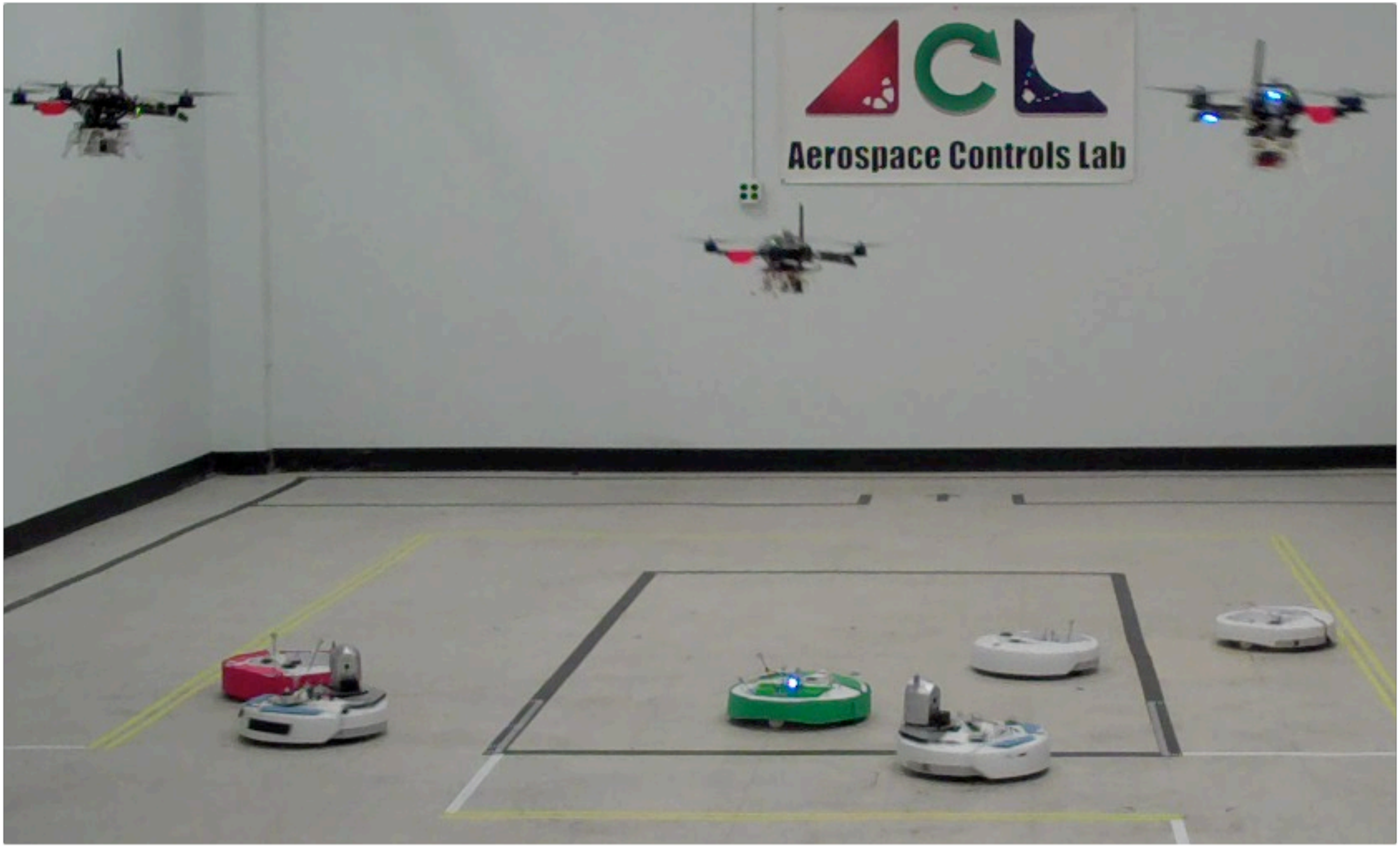
Finale Doshi

Joshua Redding

Nicholas Roy

Jonathan How

Problem



Why is it a hard?

- Unknown Model
- Stochastic Environment
- Large State Space
- Limited Online Computation

Why is it a hard?

- Unknown Model
- Stochastic Environment

Online Model-Free RL

- Large State Space
- Limited Online Computation

Existing Gap in the Literature

- Lack of Convergence [Rivest et al. 2003]
- Computational Complexity [Wu et al. 2004]
- Sample Complexity [Whiteson et al. 2007]
- Hand tuning many parameters [Kolter et al. 2009]

Existing Gap in the Literature

- Lack of Convergence [Rivest et al. 2003]
  Has convergence proof
- Computational Complexity [Wu et al. 2004]
- Sample Complexity [Whiteson et al. 2007]
- Hand tuning many parameters [Kolter et al. 2009]

Existing Gap in the Literature

- Lack of Convergence [Rivest et al. 2003]
  Has convergence proof
- Computational Complexity [Wu et al. 2004]
  Required < 4 ms per step
- Sample Complexity [Whiteson et al. 2007]

- Hand tuning many parameters [Kolter et al. 2009]

Existing Gap in the Literature

- Lack of Convergence [Rivest et al. 2003]

  Has convergence proof
- Computational Complexity [Wu et al. 2004]

  Required < 4 ms per step
- Sample Complexity [Whiteson et al. 2007]

  Scaled to large problems
- Hand tuning many parameters [Kolter et al. 2009]

Existing Gap in the Literature

- Lack of Convergence [Rivest et al. 2003]

  Has convergence proof
- Computational Complexity [Wu et al. 2004]

  Required < 4 ms per step
- Sample Complexity [Whiteson et al. 2007]

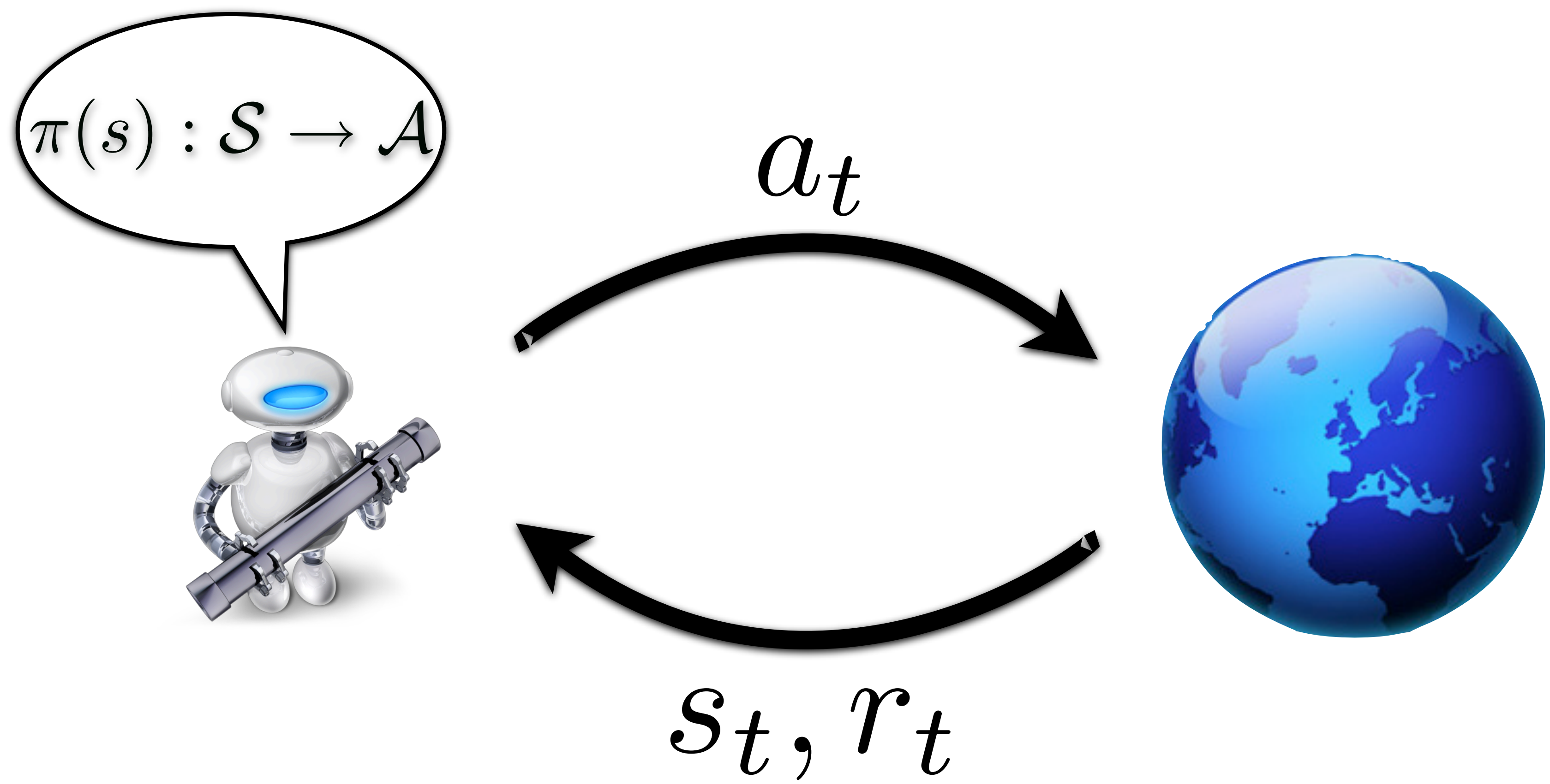
  Scaled to large problems
- Hand tuning many parameters [Kolter et al. 2009]

  Has one parameter

Contributions

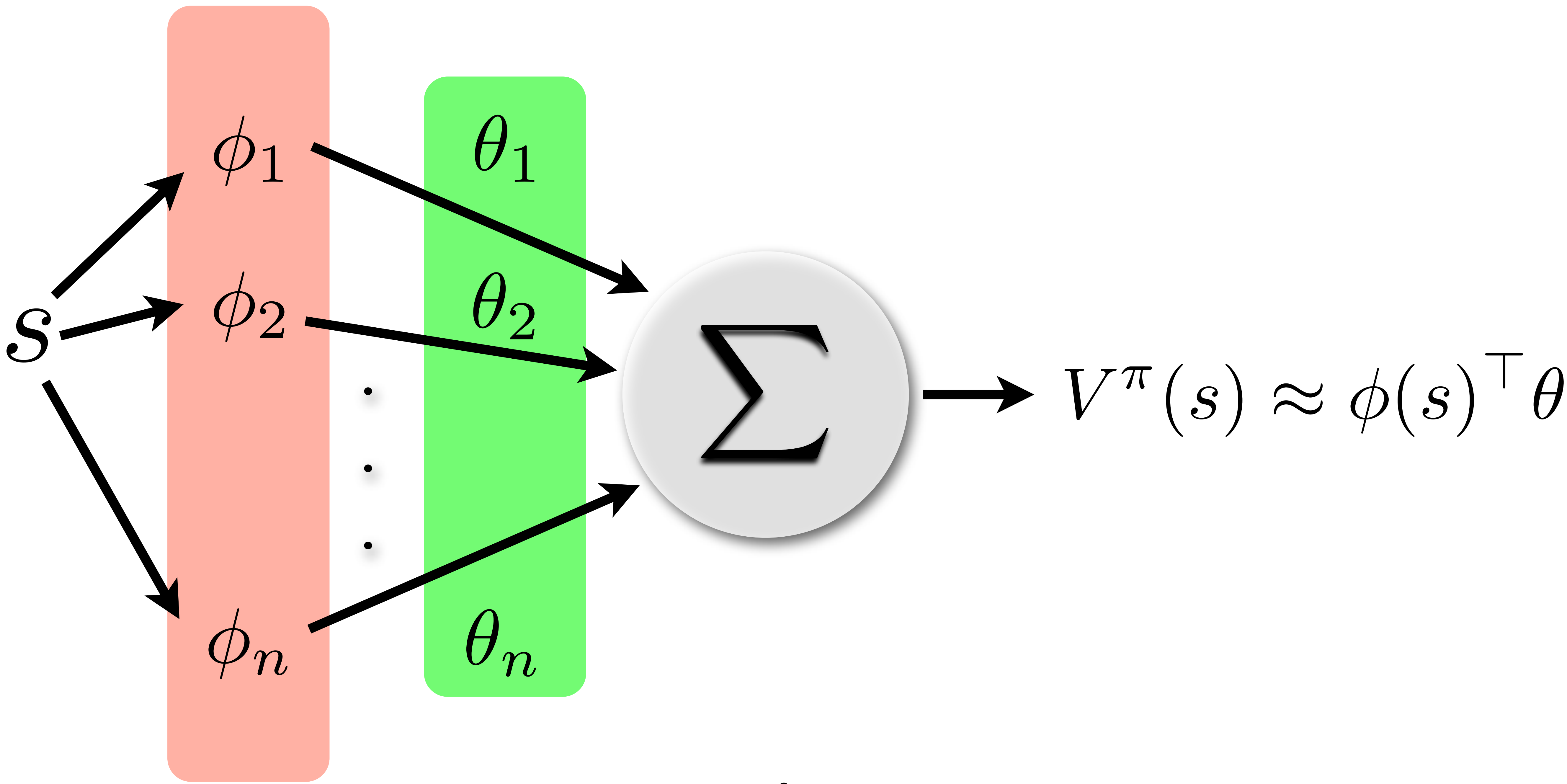
- Introduced Incremental Feature Dependency Discovery (**iFDD**) as a novel feature expansion method
- Provided asymptotic **convergence** analysis
- Empirically showed the **scalability** of the new approach in problems with $\approx 10^8$ possibilities

Reinforcement Learning

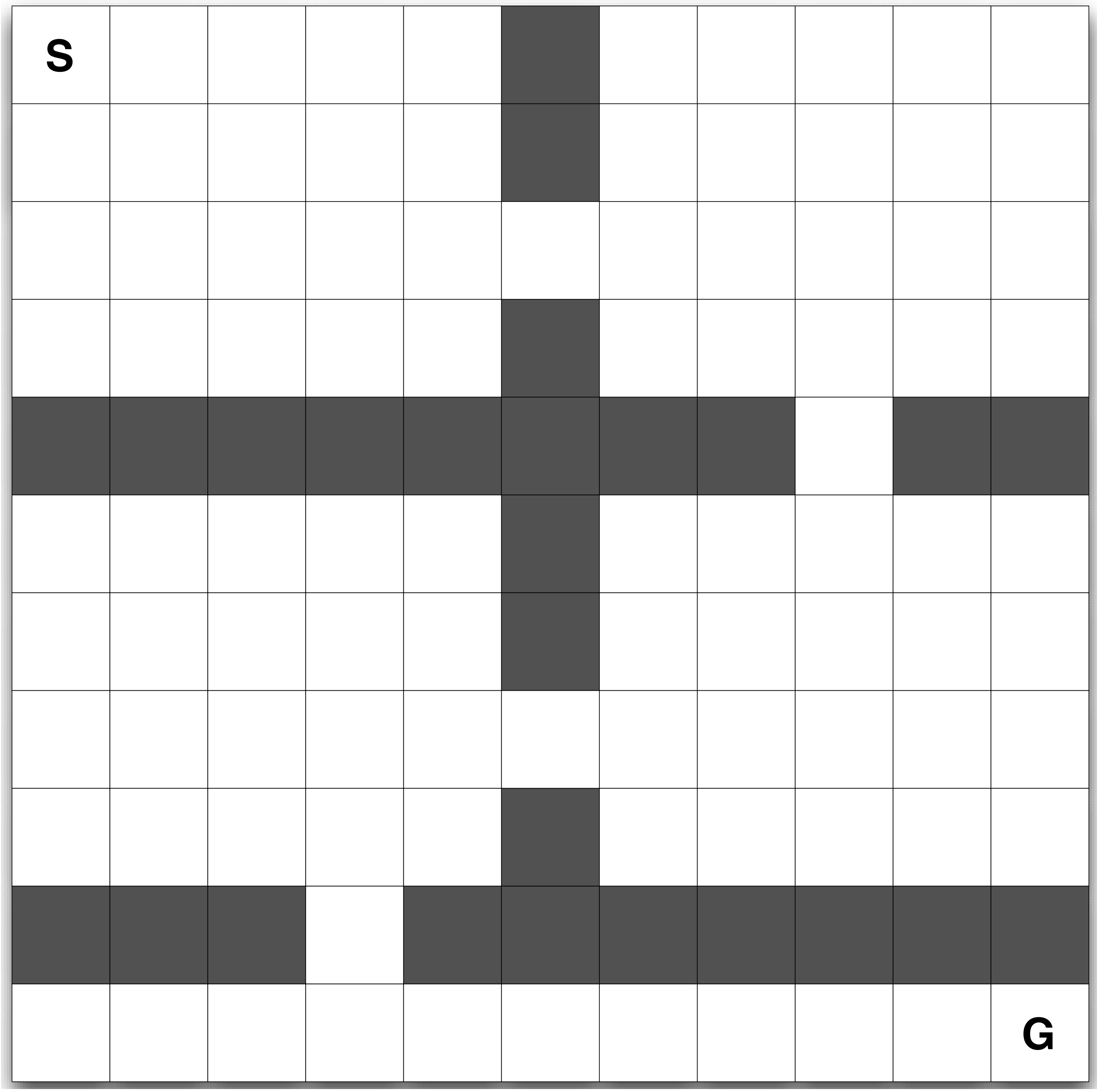


$$V^{\pi}(s) = E_{\pi}\left[\sum_{t=0}^{\infty} \gamma^t r_t \middle| s_0 = s\right]$$

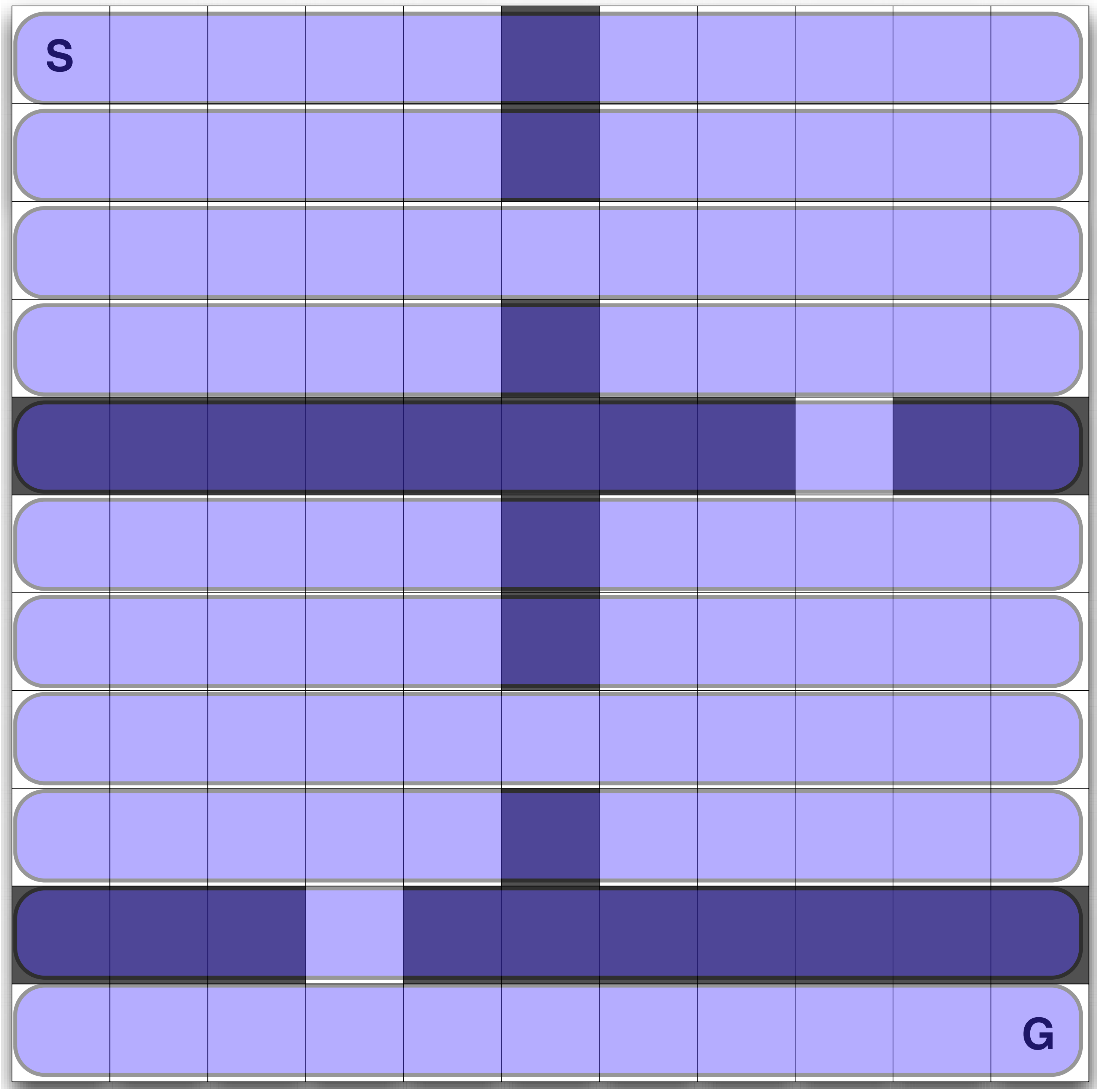
Linear Function Approximation



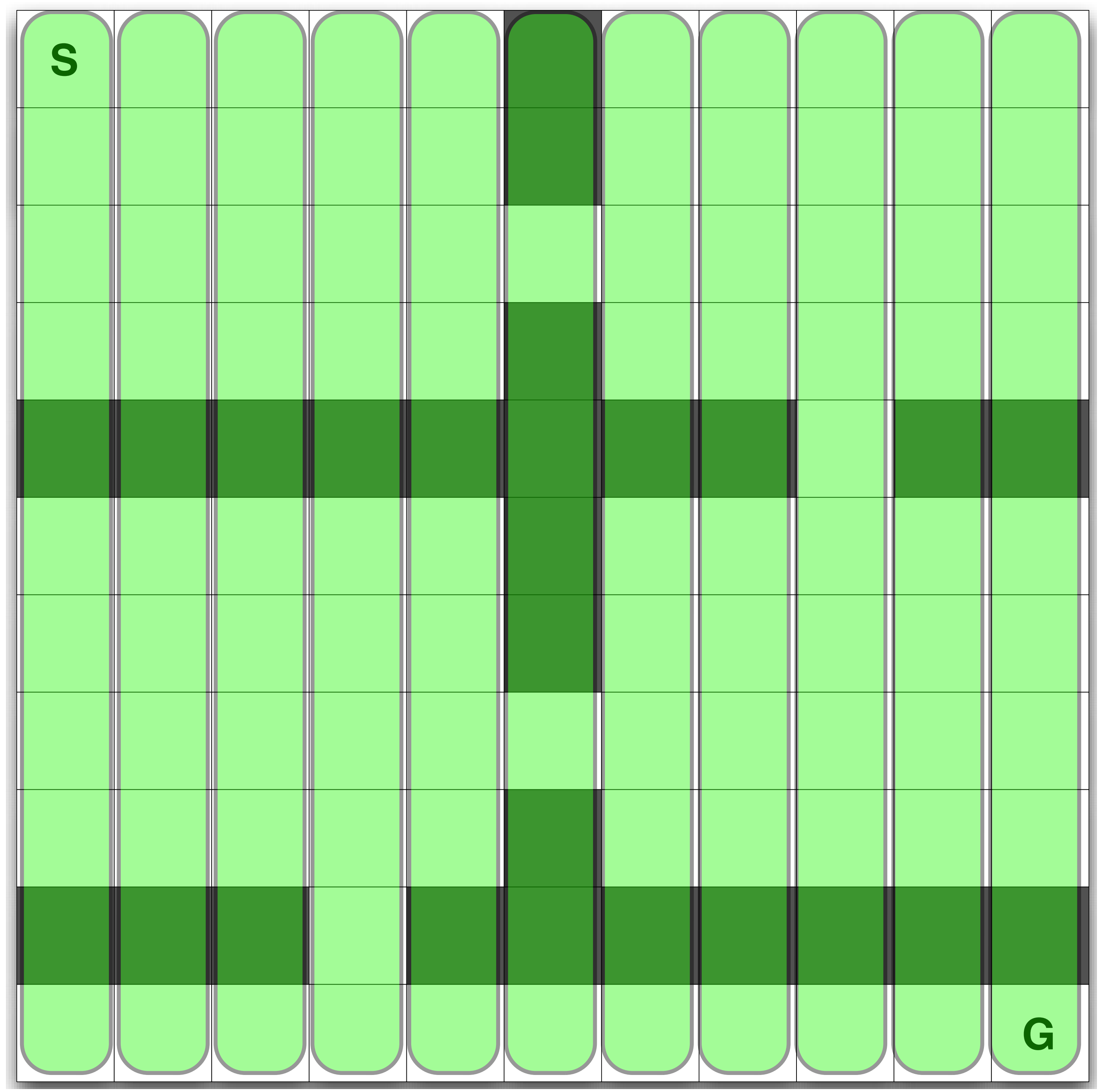
Why Features Expansion?



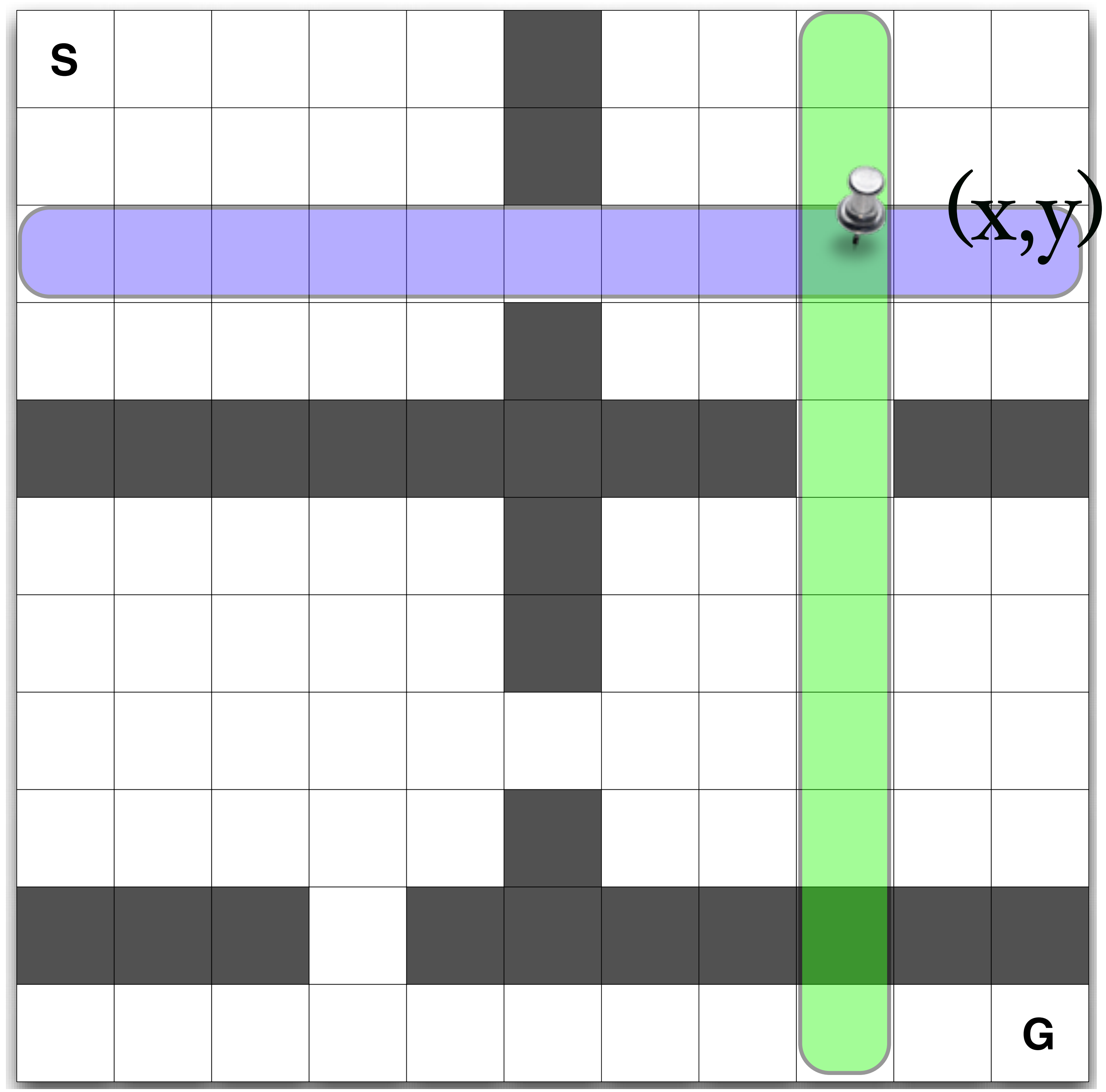
Why Features Expansion?



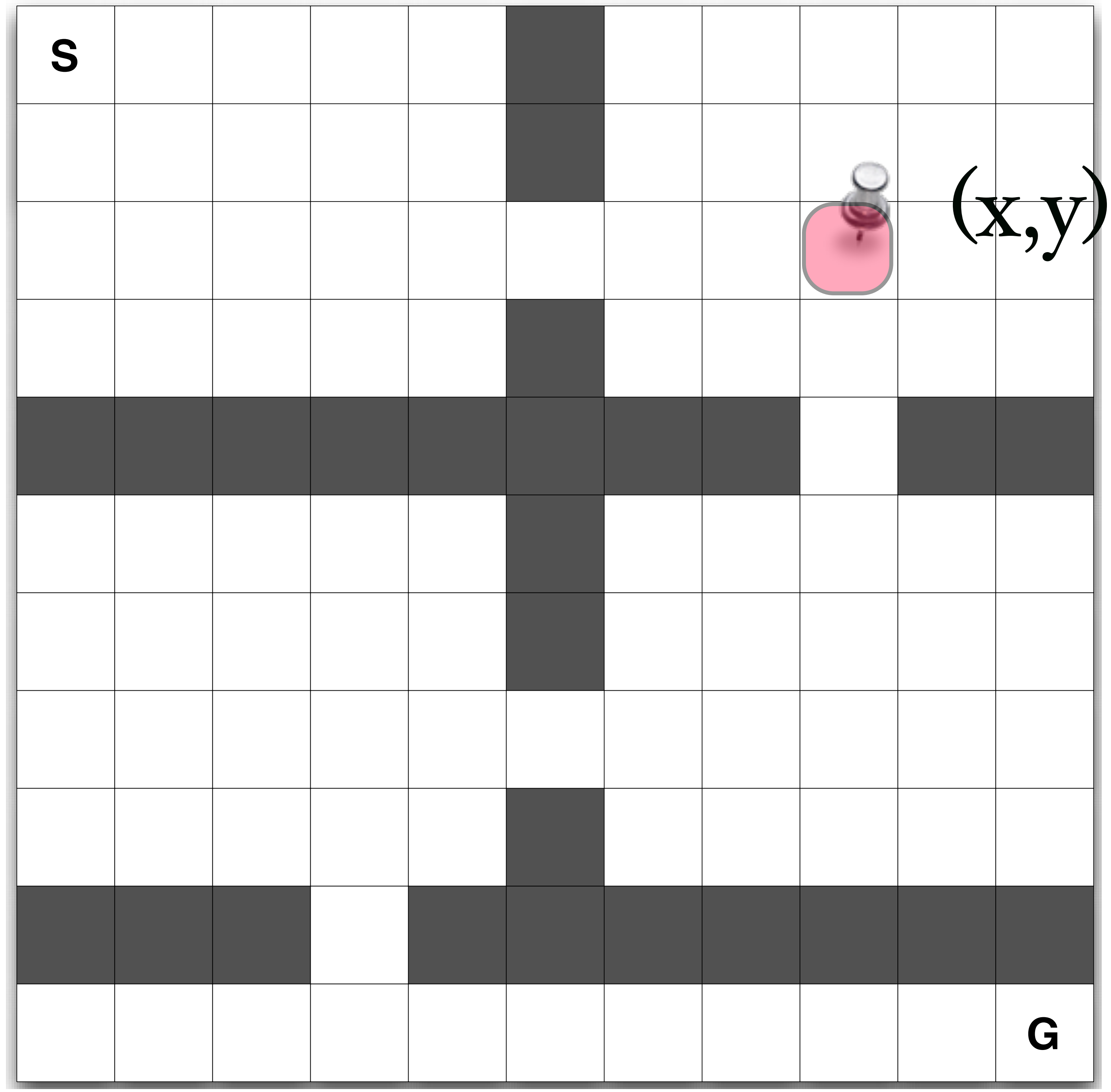
Why Features Expansion?



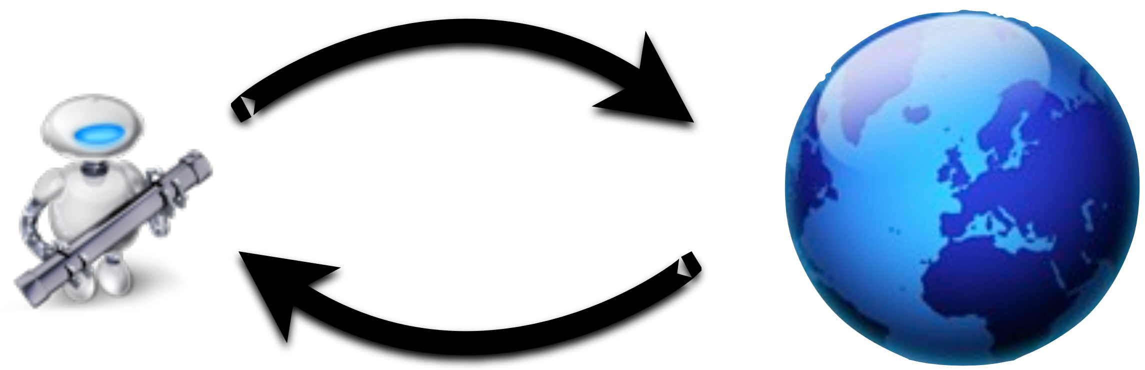
Why Features Expansion?

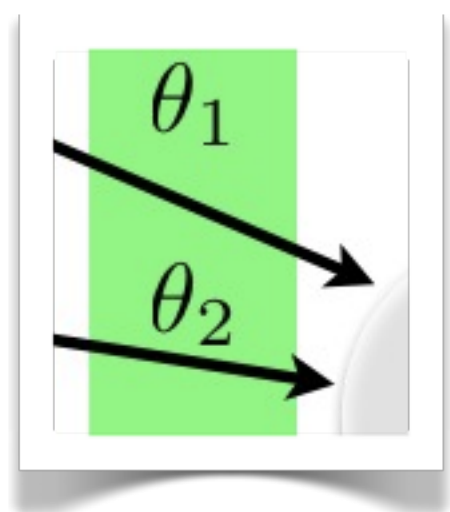


Why Features Expansion?

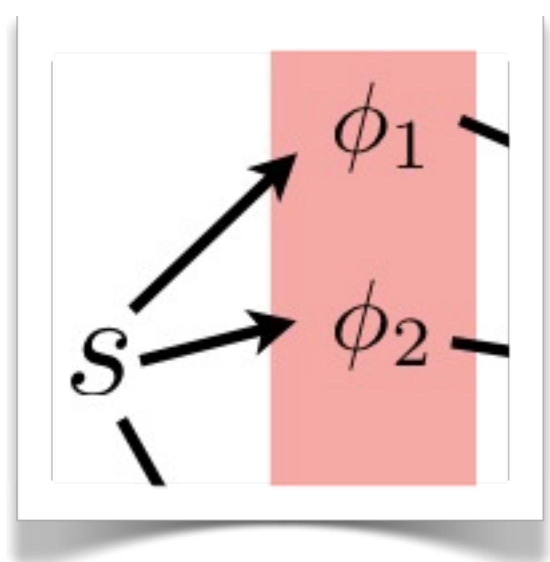


Control Loop:

1 

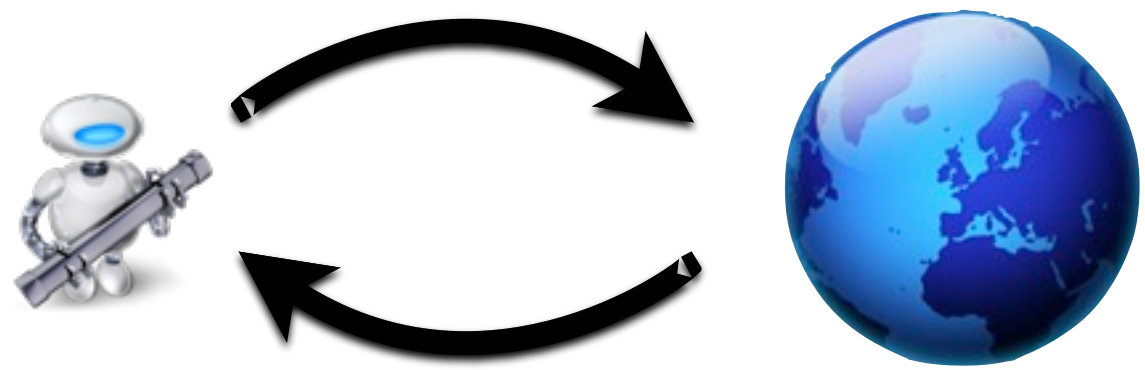
2 Update Weights 

Sarsa

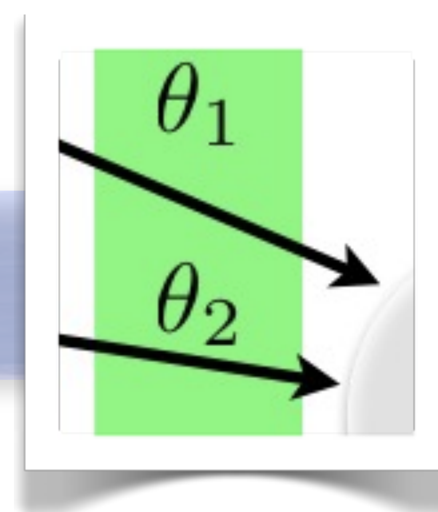
3 Update Features 

iFDD

Control Loop:

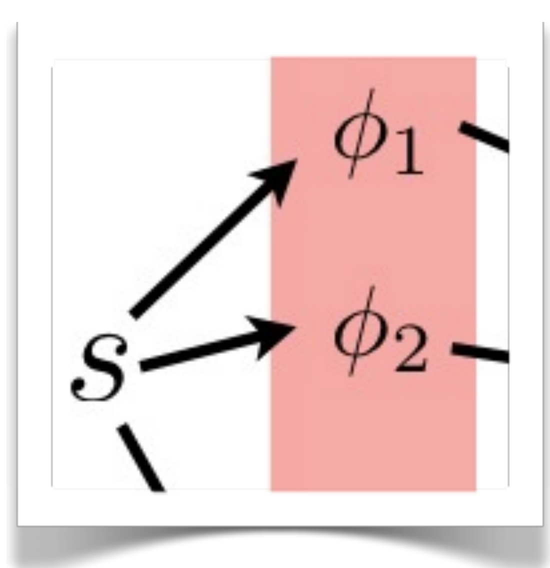
1 

2 Update Weights



Sarsa

3 Update Features



iFDD

Sarsa

$$s_t \xrightarrow{(\pi)\ a,r} s_{t+1}$$

- Temporal Difference (TD) Error

$$\delta_t = r_t + \gamma V(s_{t+1}) - V(s_t).$$

- Linear Function Approximation

$$\theta_{t+1} = \theta_t + \alpha_t \phi(s_t) \delta_t(V).$$

[Sutton 88]

Sources of TD Error

- Incorrect Weights
- Stochasticity
- Underpowered Representation

Sources of TD Error

- Incorrect Weights
- Stochasticity
- Underpowered Representation

Sarsa

Sources of TD Error

- Incorrect Weights
- Stochasticity
- Underpowered Representation

Model Based Methods

Sources of TD Error

- Incorrect Weights
- Stochasticity
- Underpowered Representation

iFDD

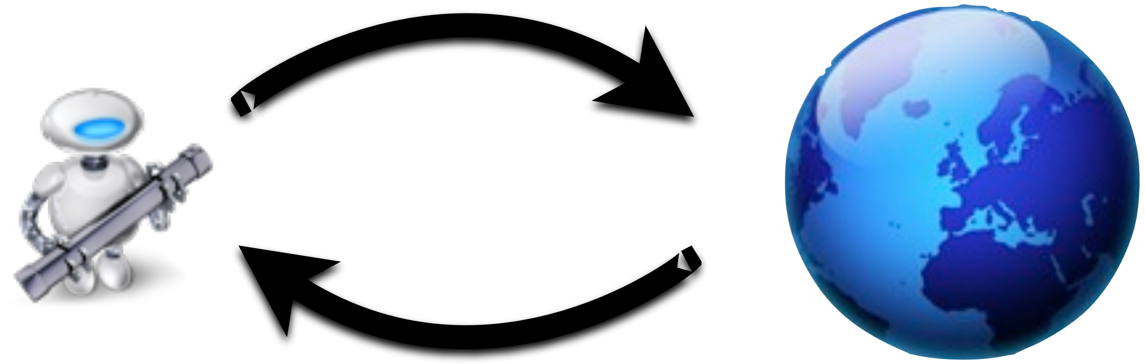
Sources of TD Error

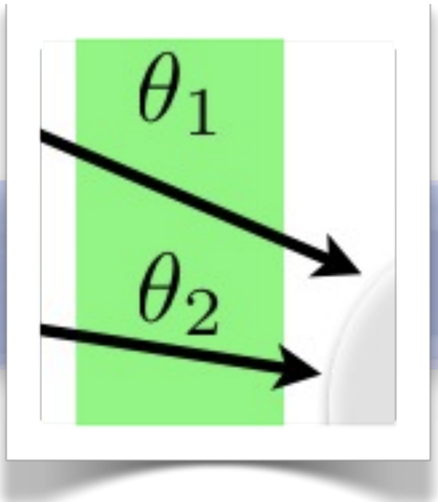
- Incorrect Weights
- Stochasticity
- Underpowered Representation

iFDD

Most **accumulated error** ⇒ where the representation should **grow**.

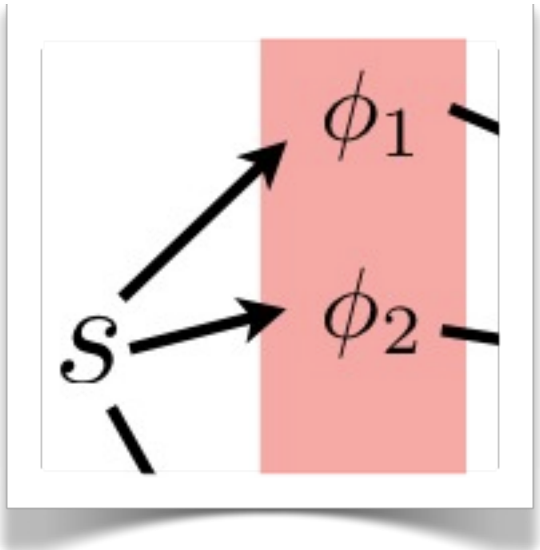
Control Loop:

1 

2 Update Weights 

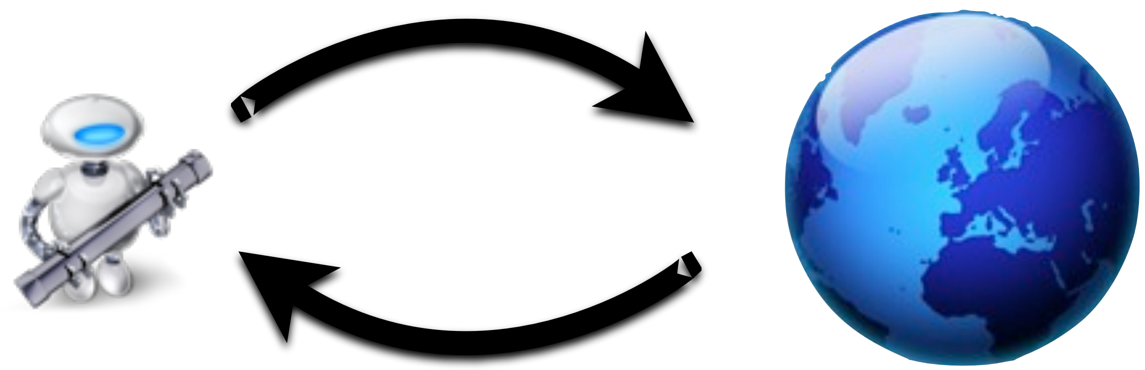
Sarsa

$\downarrow |\delta_t|$

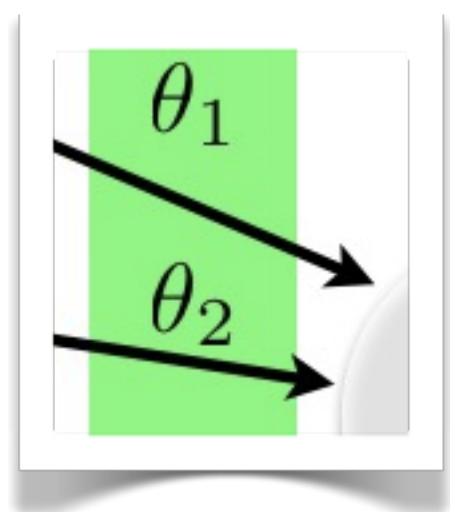
3 Update Features 

iFDD

Control Loop:

1 

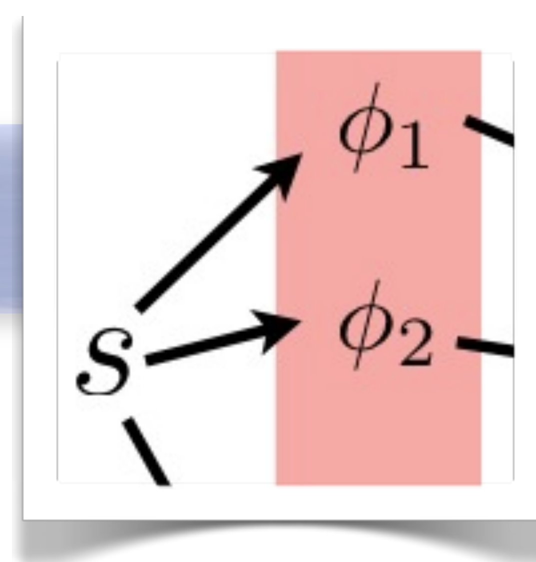
2 Update Weights



Sarsa

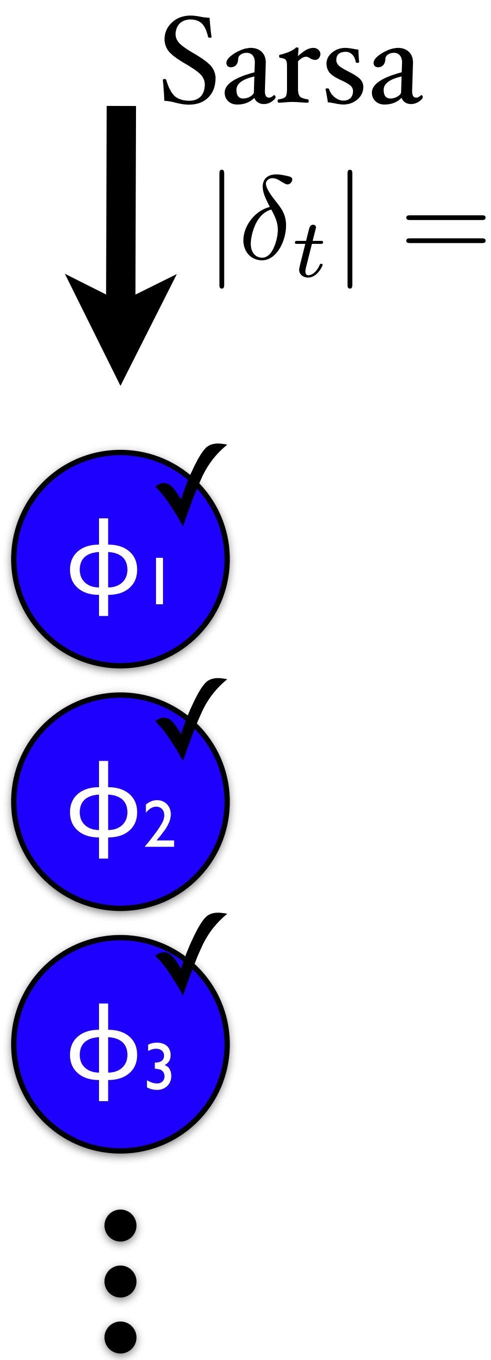
$|\delta_t|$

3 Update Features

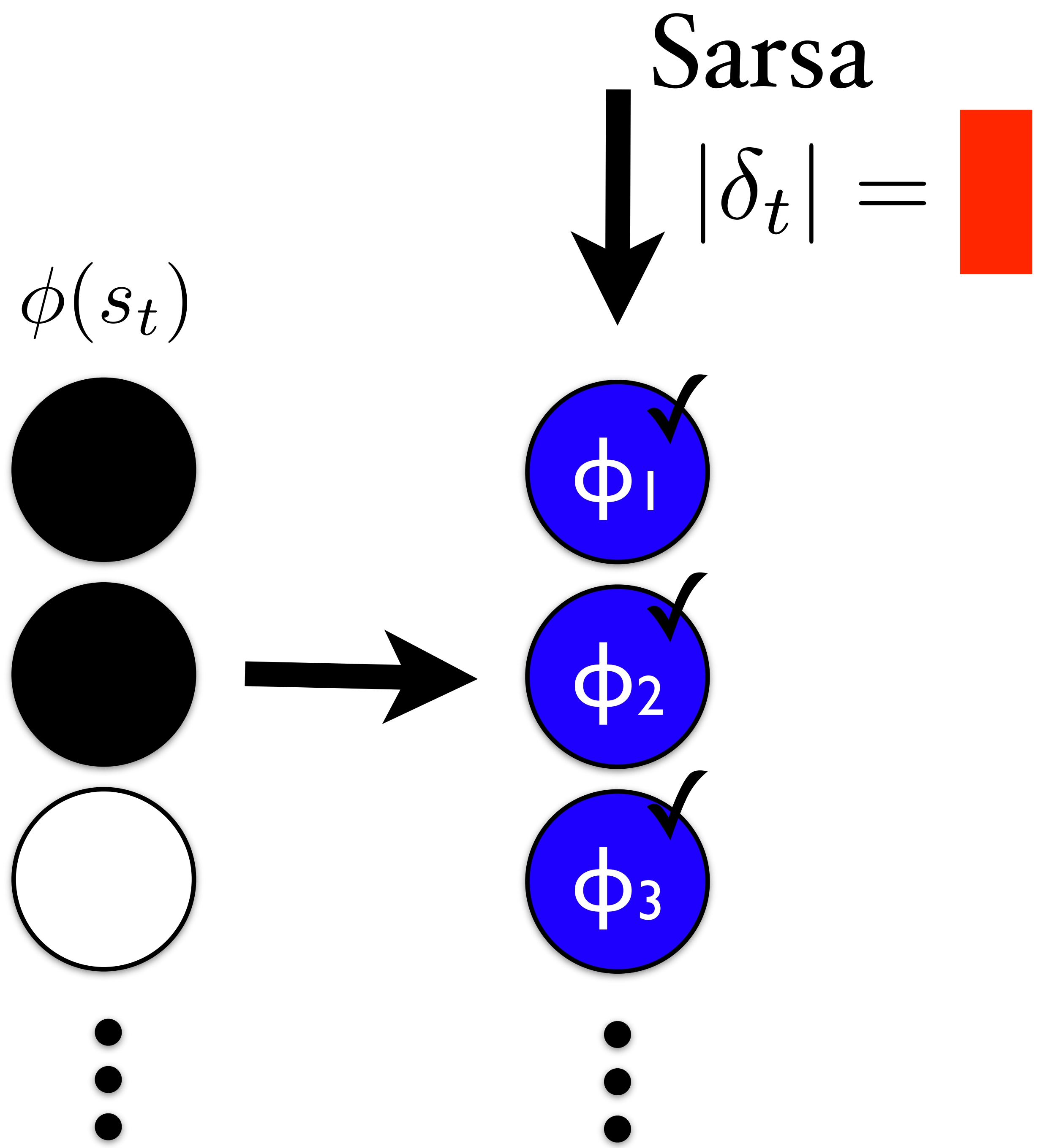


iFDD

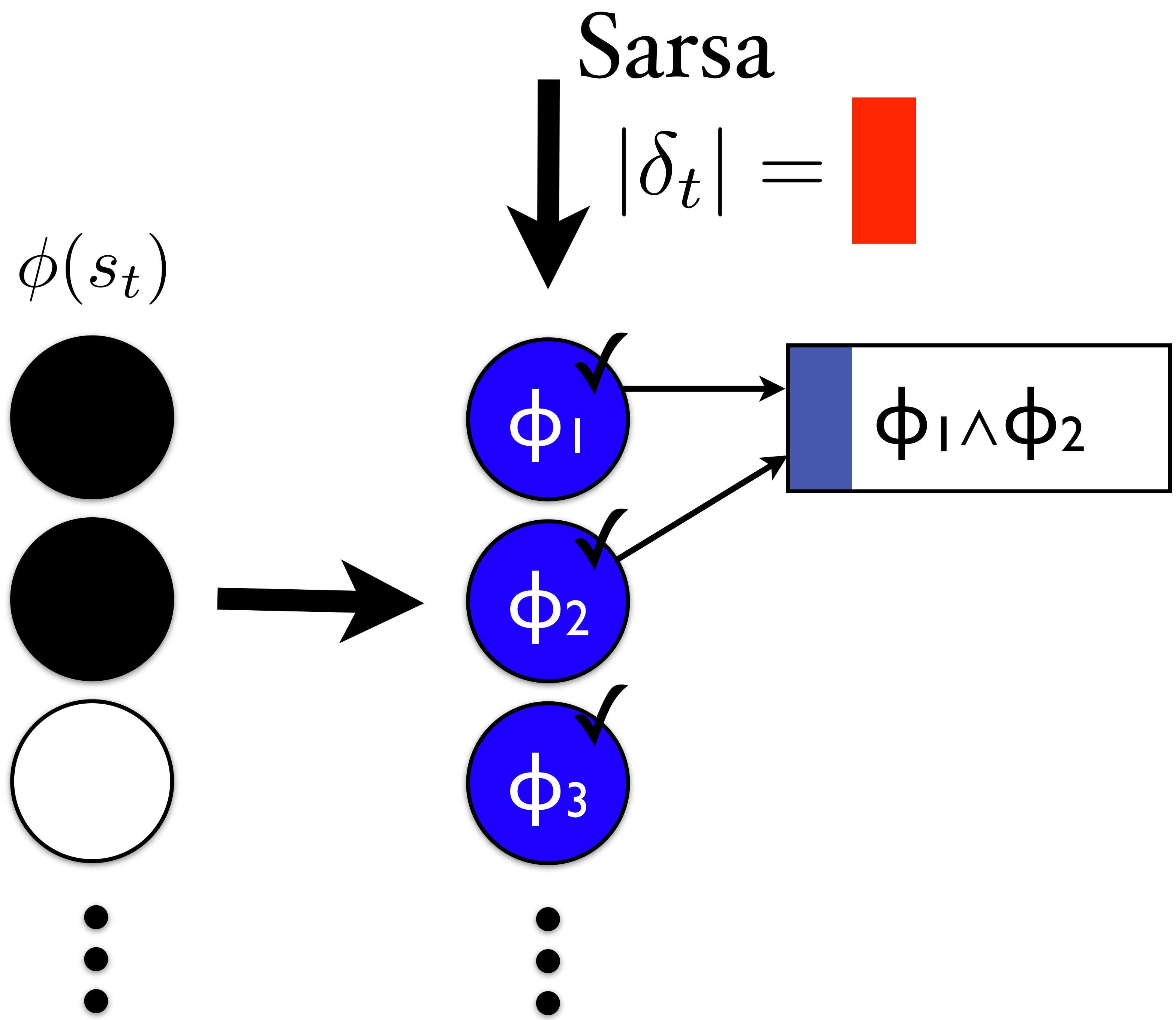
Incremental Feature Dependency Discovery



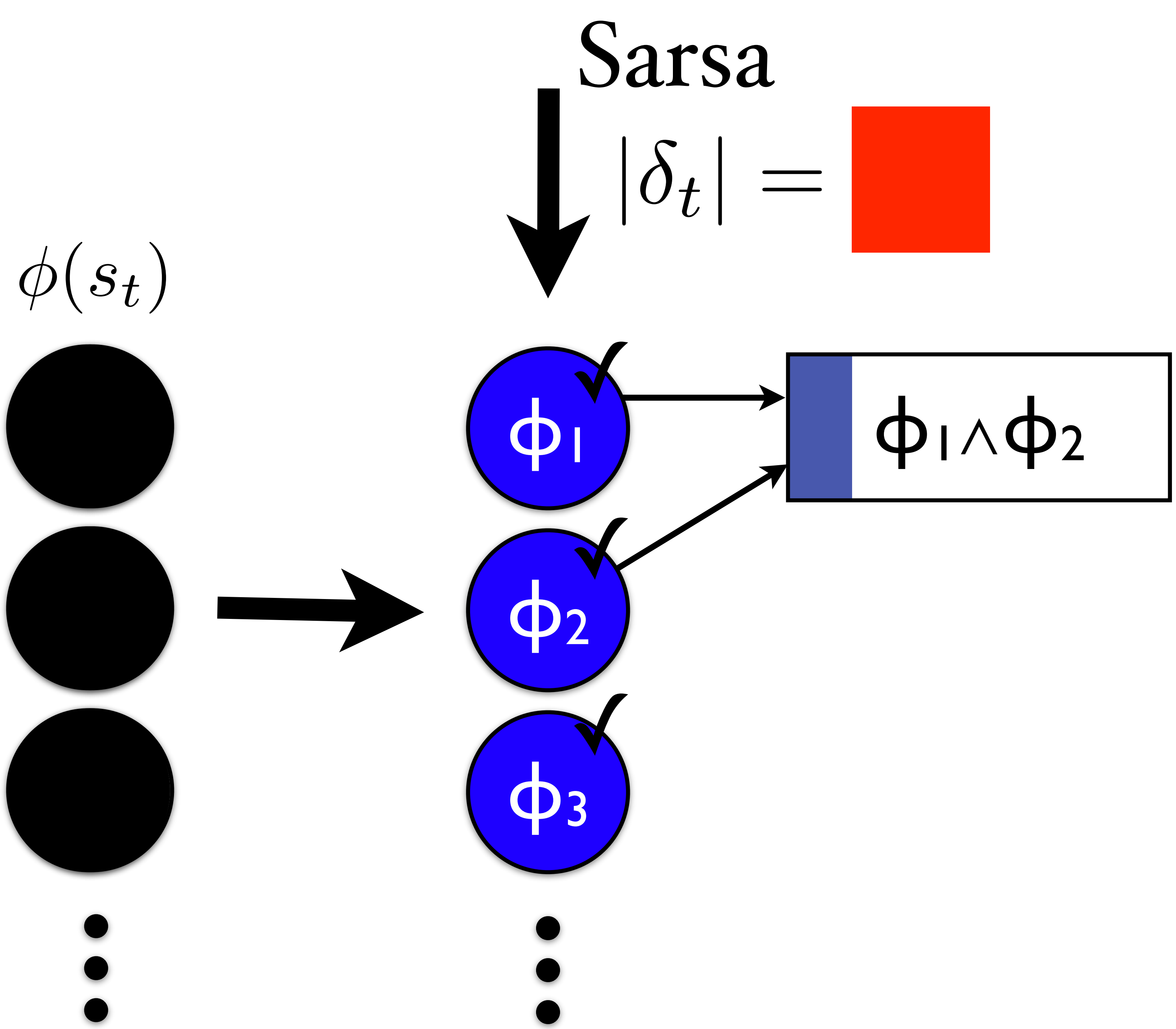
Incremental Feature Dependency Discovery



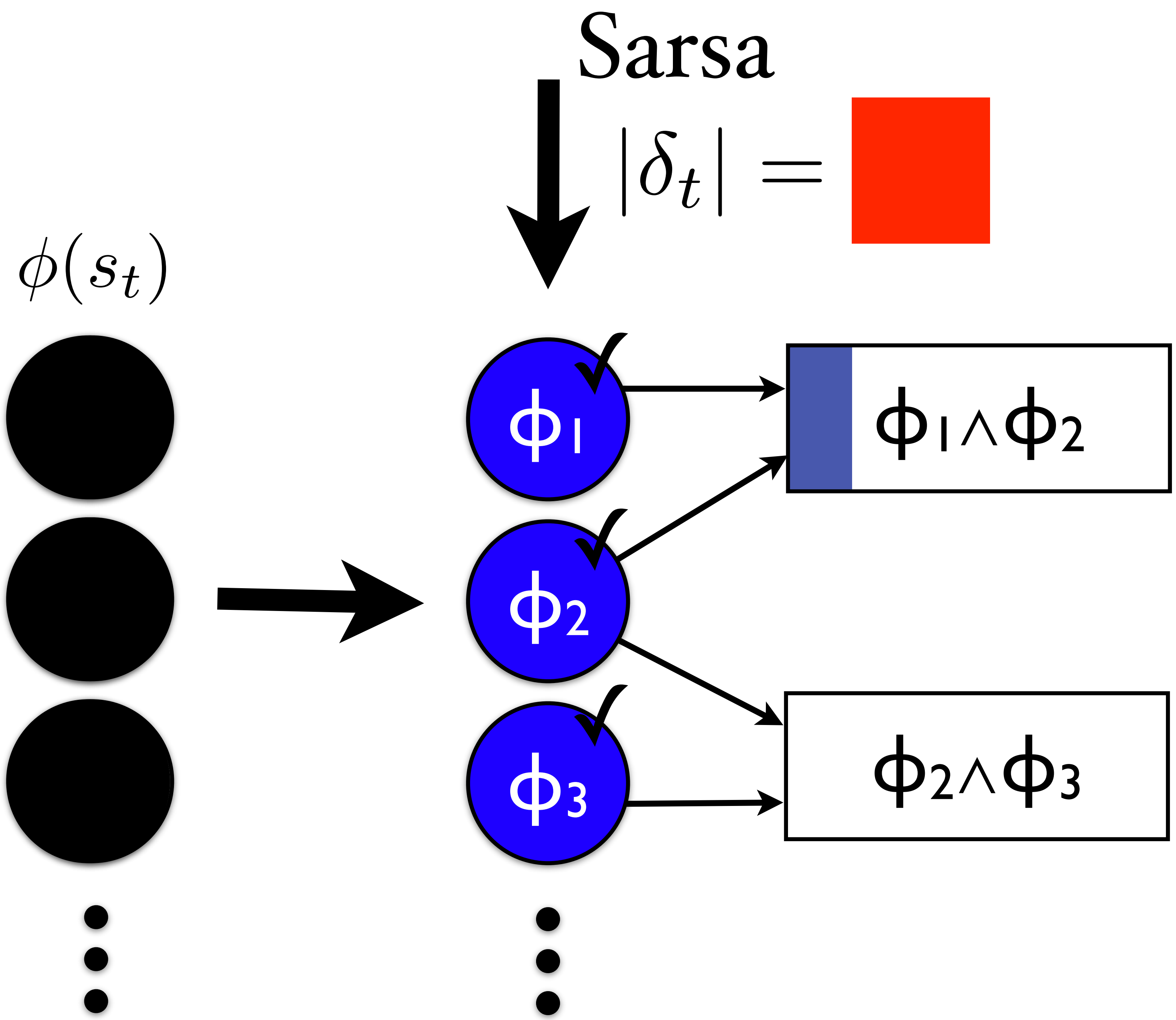
Incremental Feature Dependency Discovery



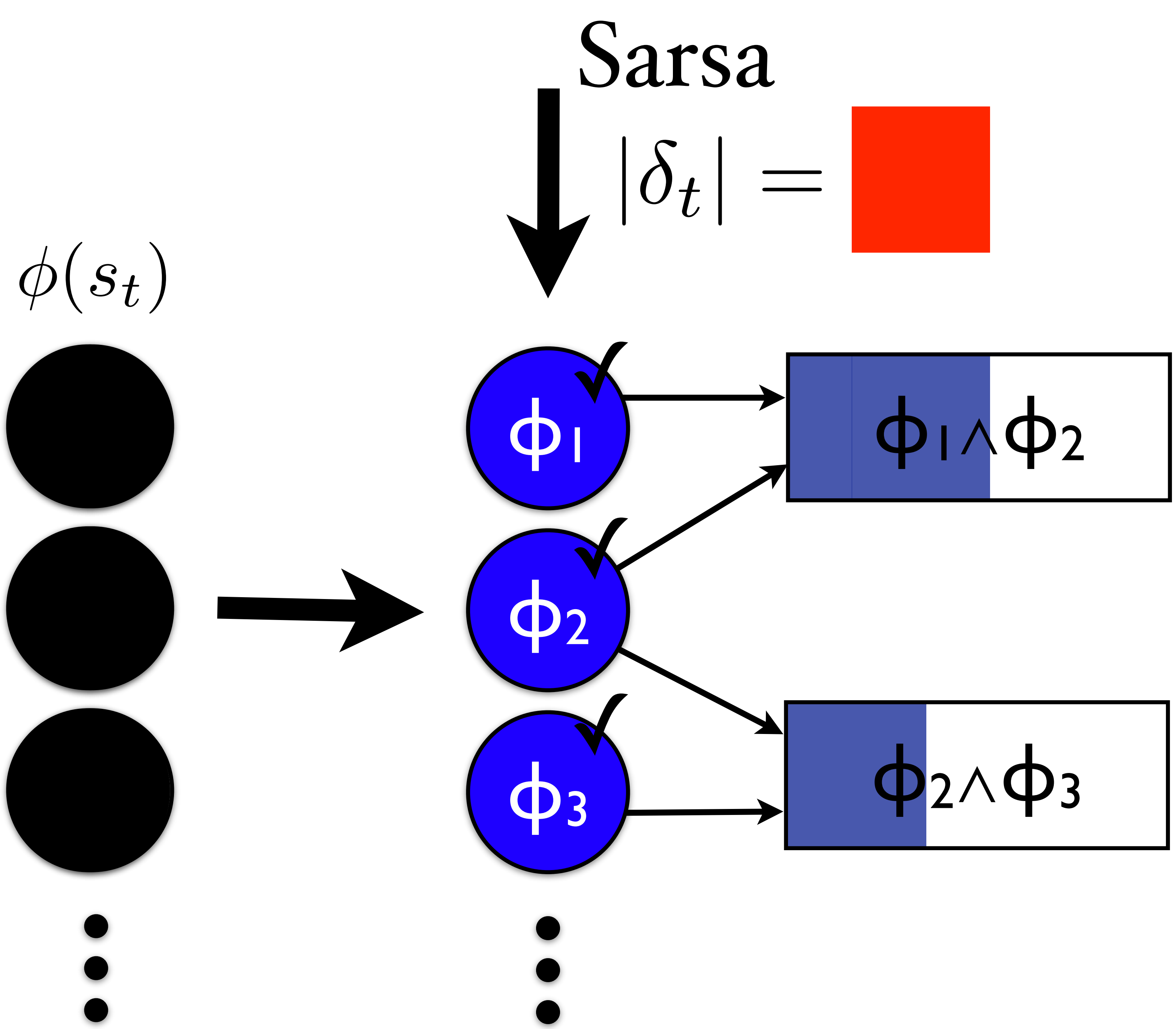
Incremental Feature Dependency Discovery



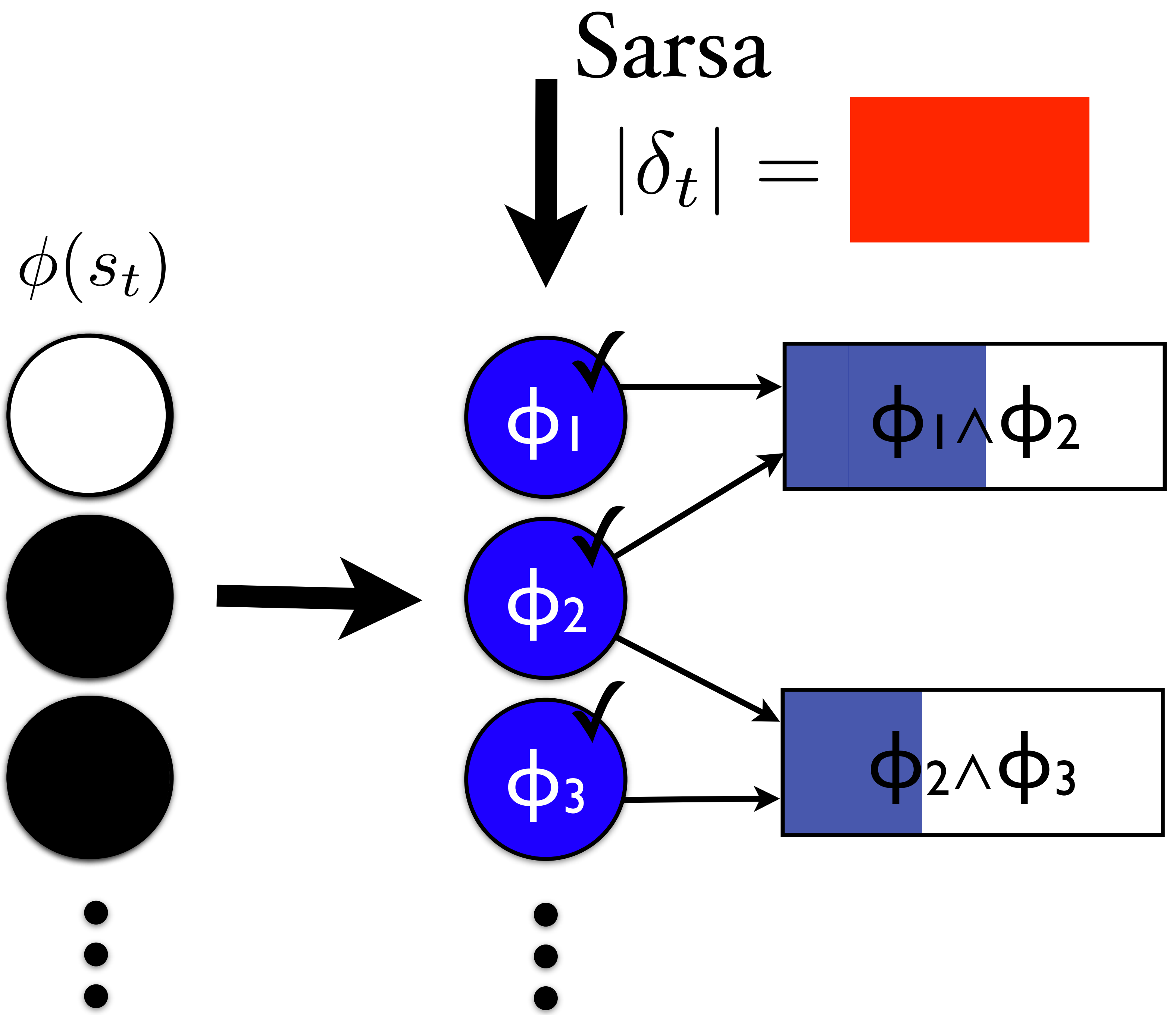
Incremental Feature Dependency Discovery



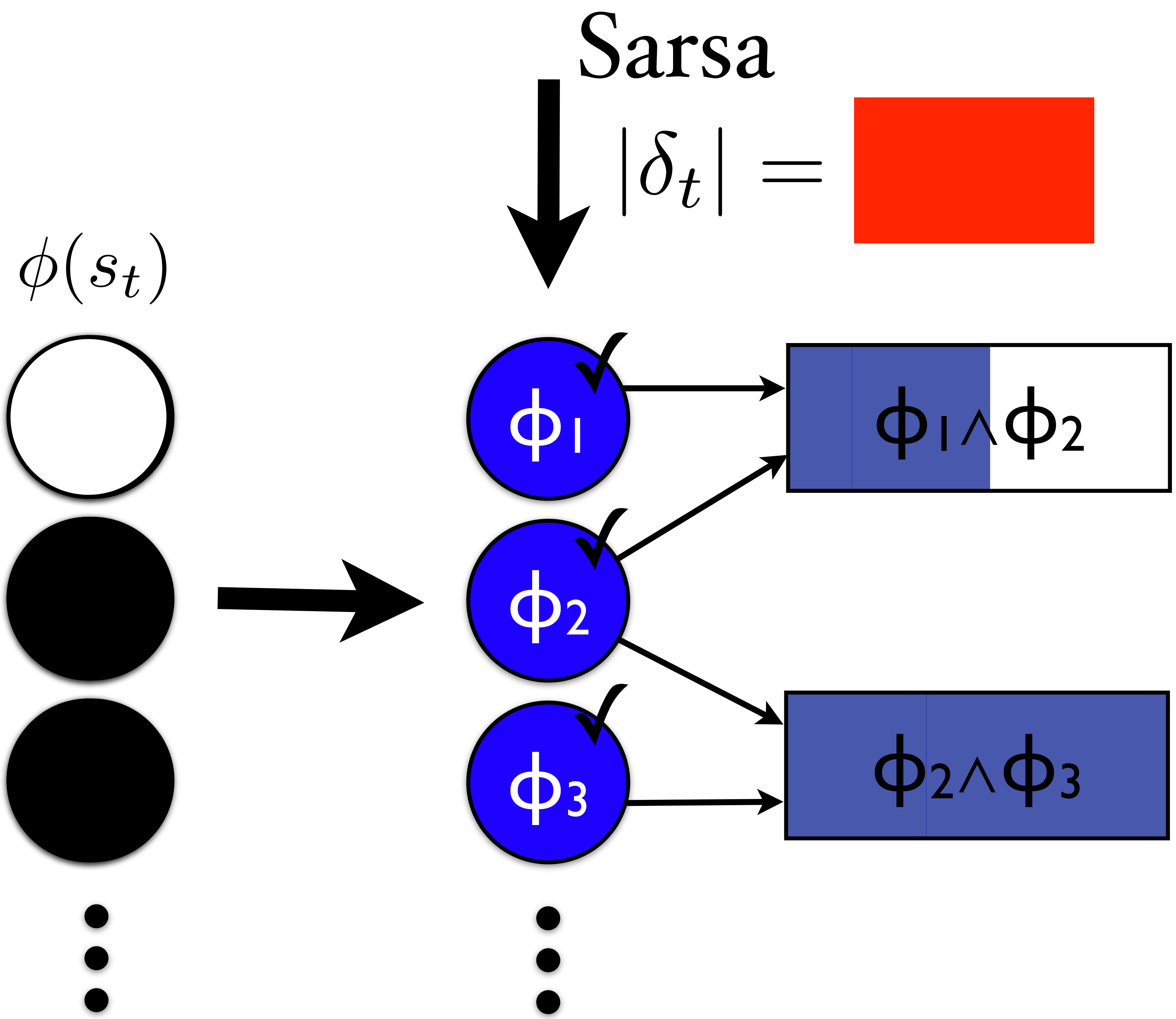
Incremental Feature Dependency Discovery



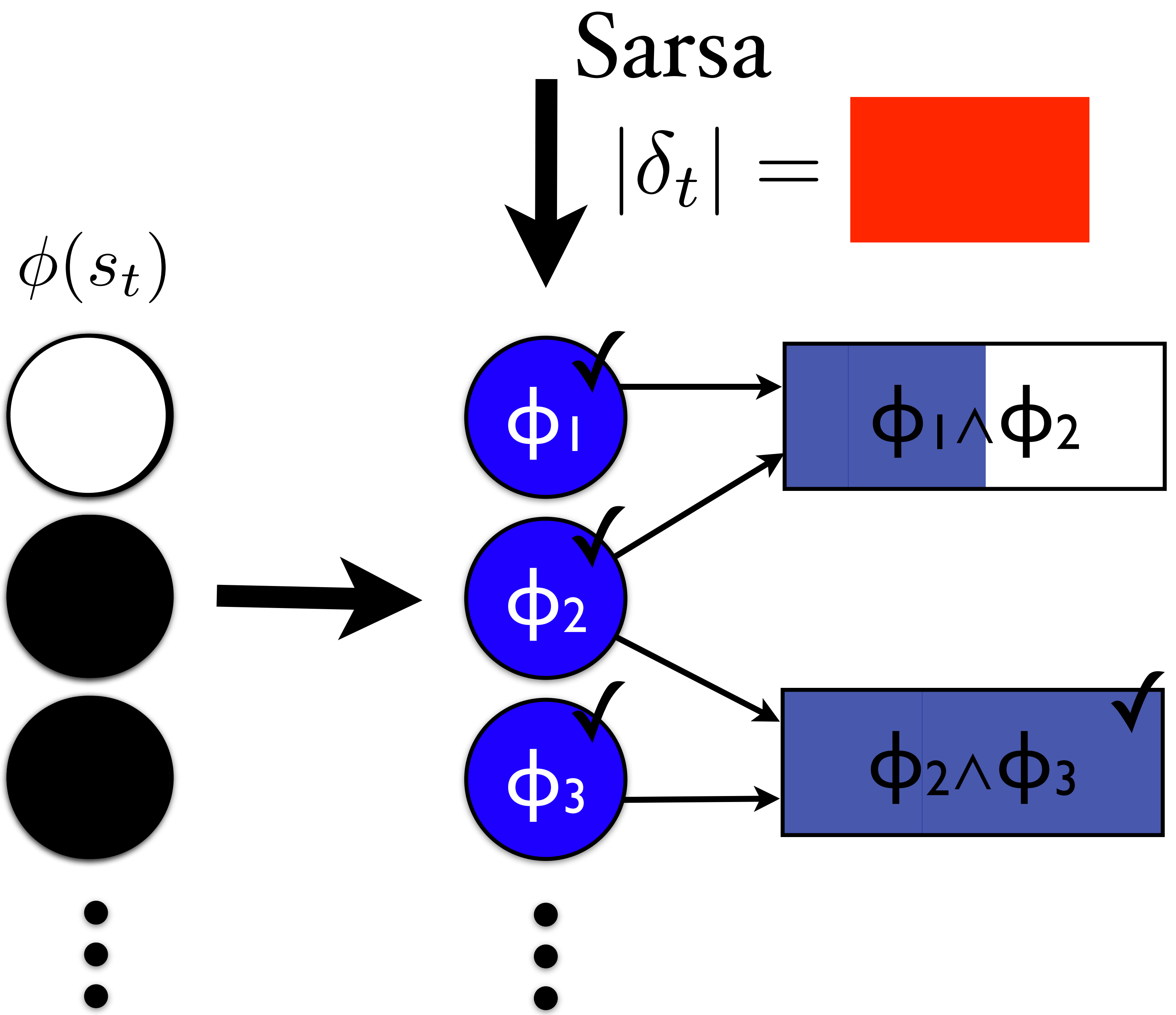
Incremental Feature Dependency Discovery



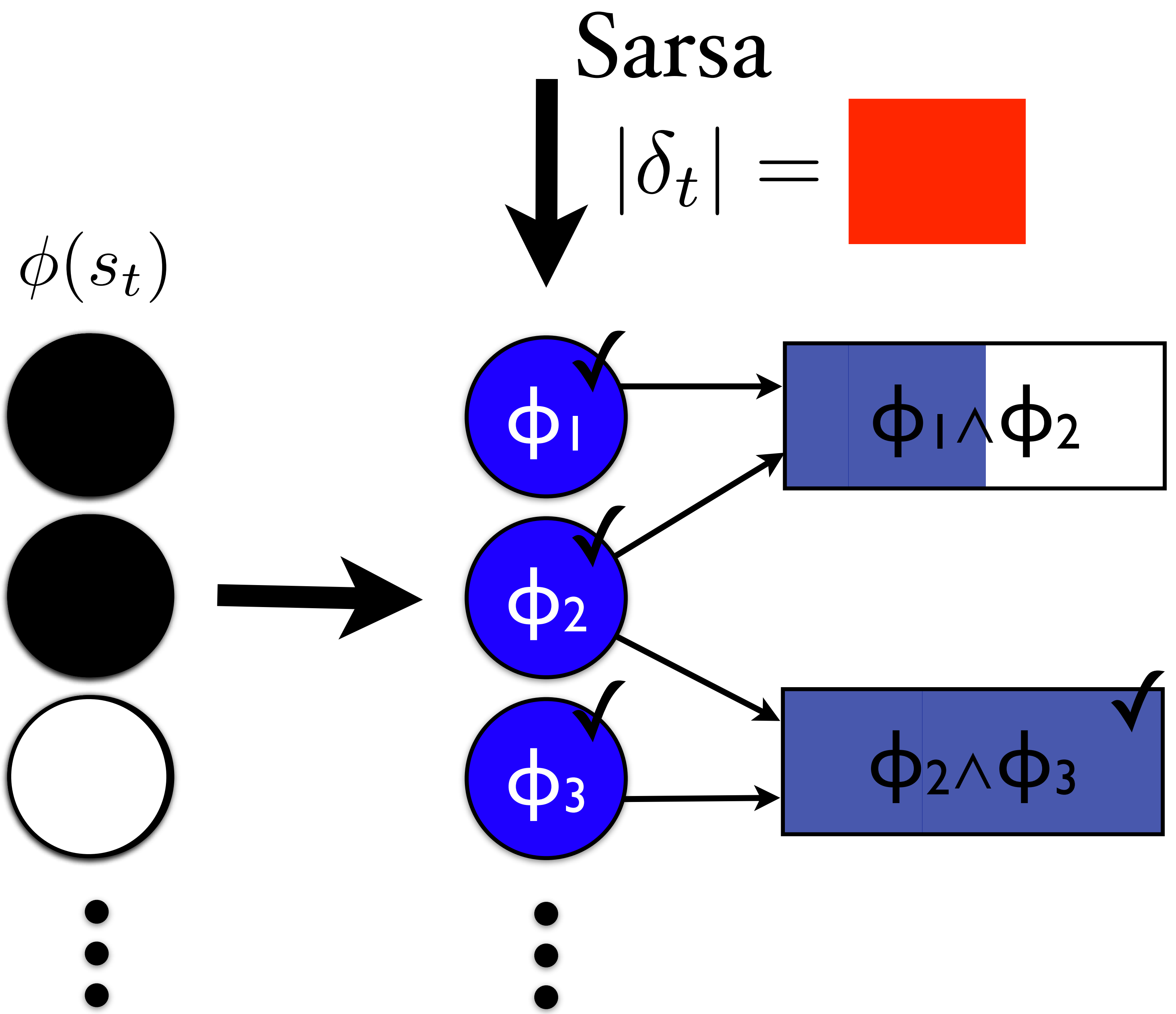
Incremental Feature Dependency Discovery



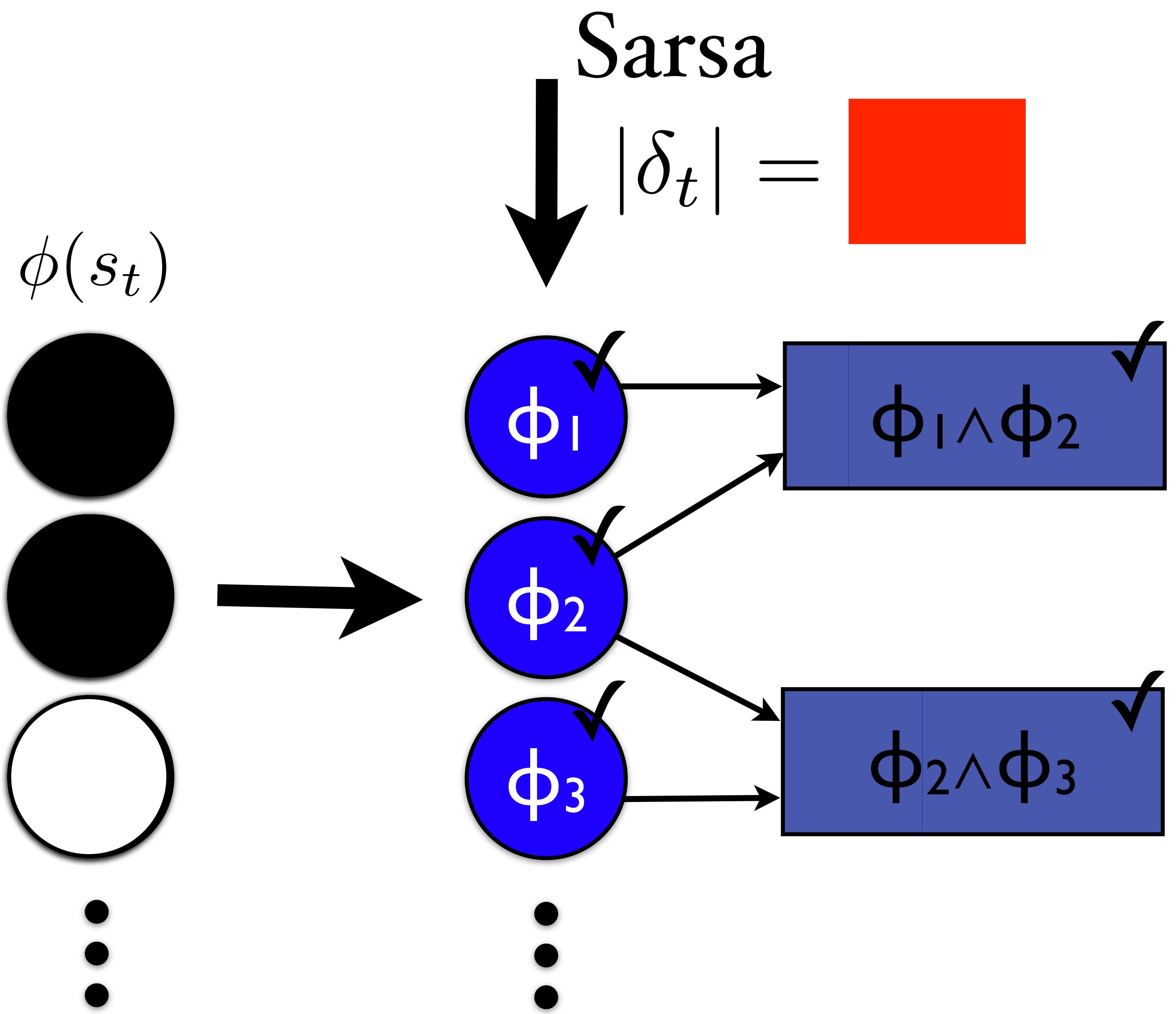
Incremental Feature Dependency Discovery



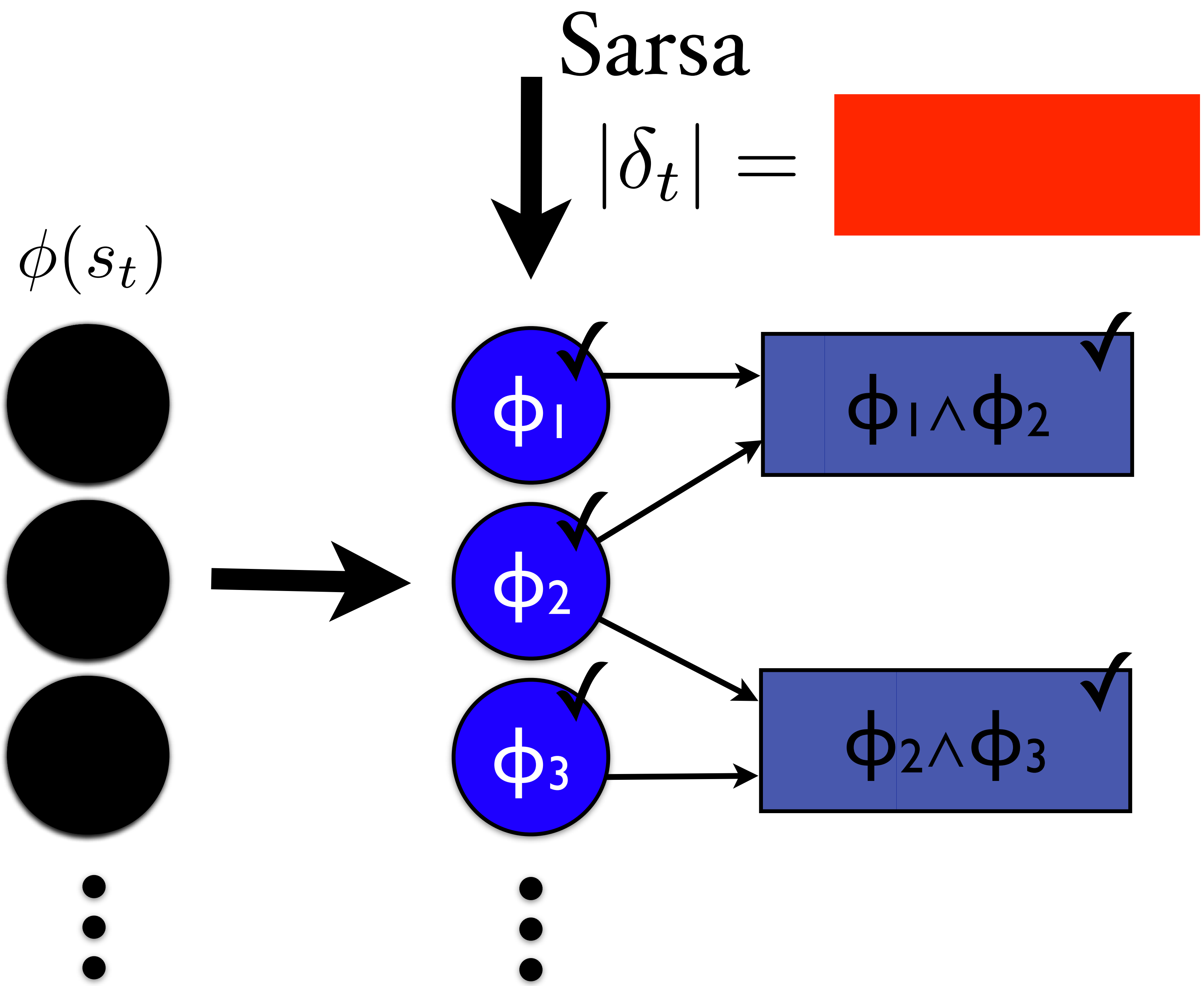
Incremental Feature Dependency Discovery



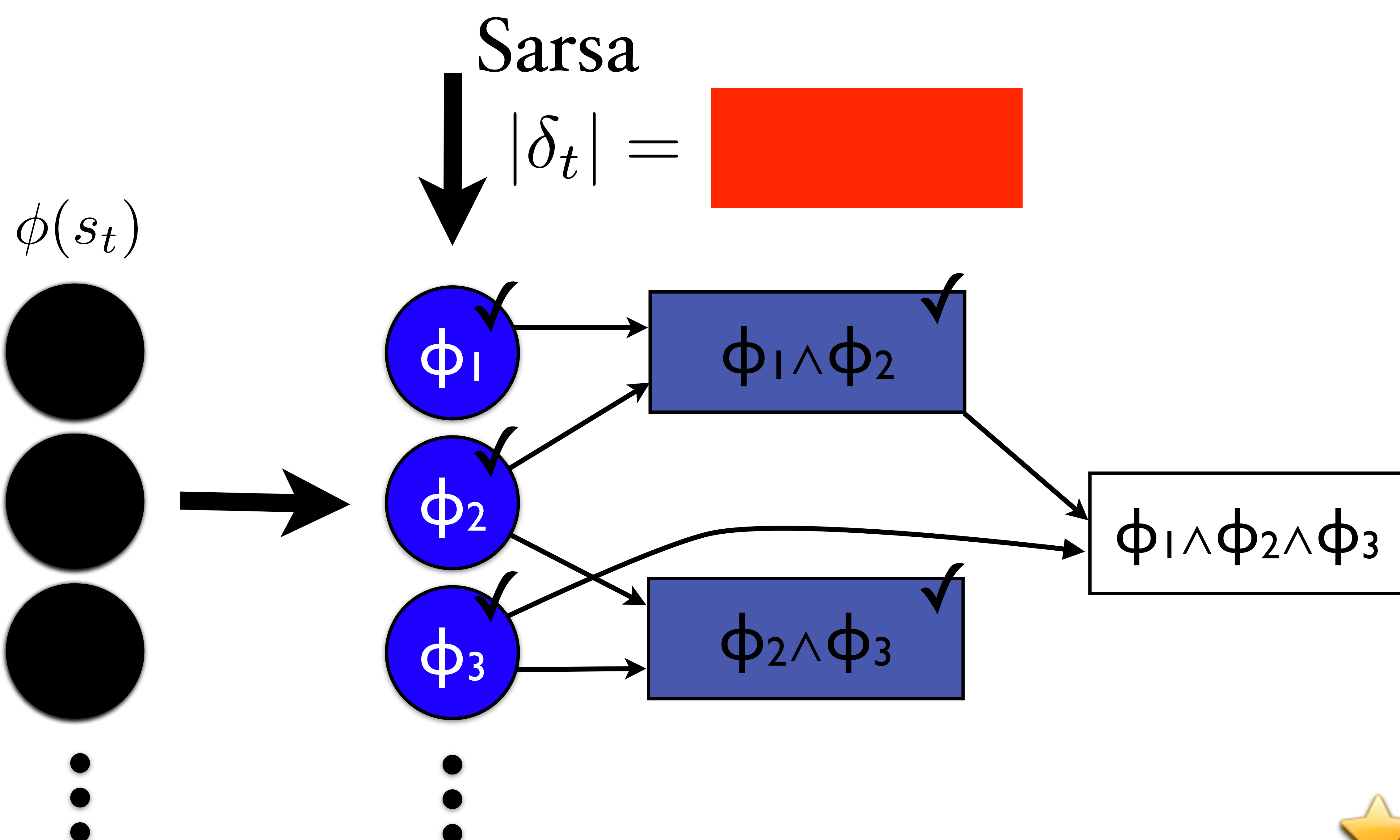
Incremental Feature Dependency Discovery



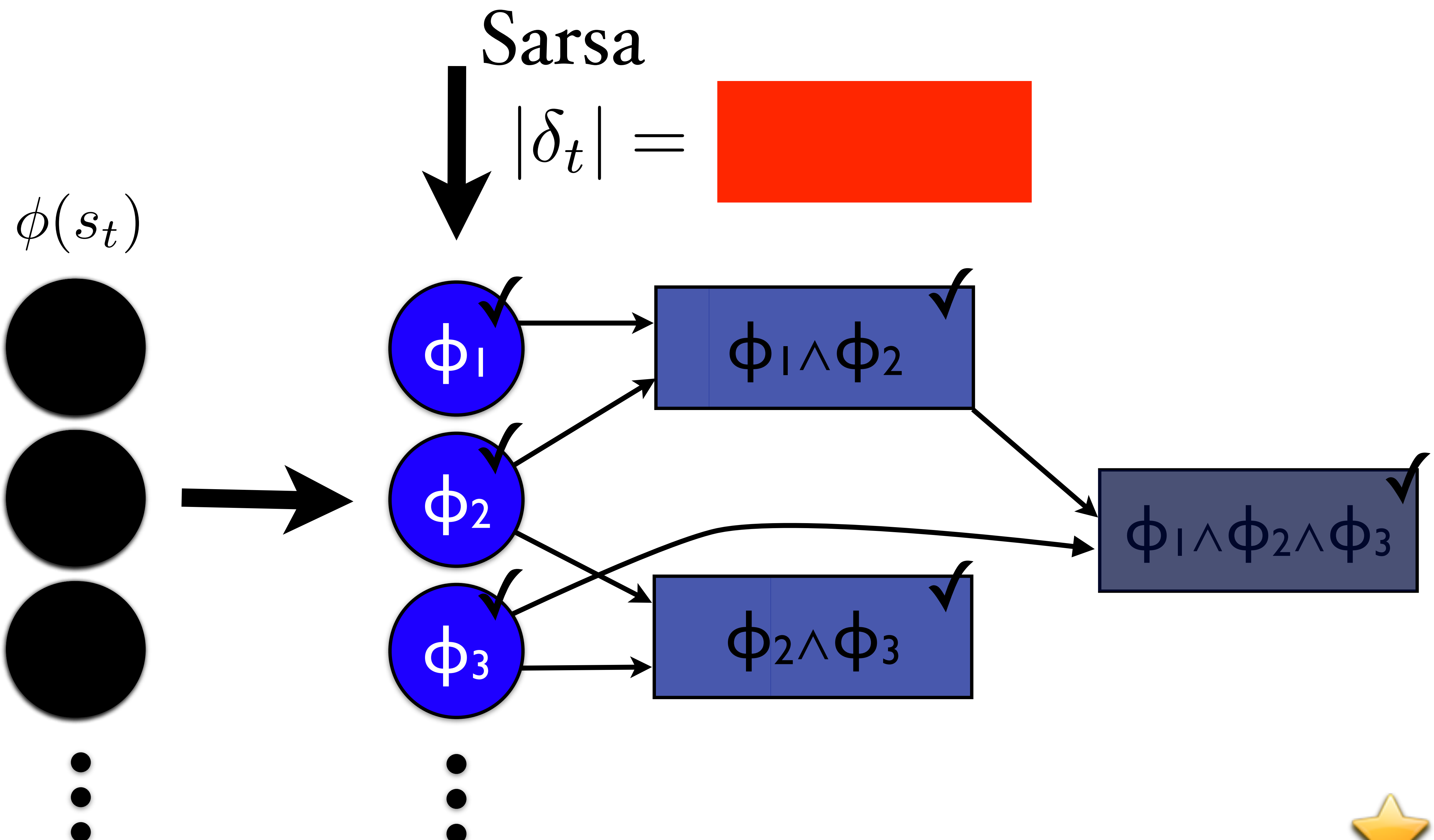
Incremental Feature Dependency Discovery



Incremental Feature Dependency Discovery



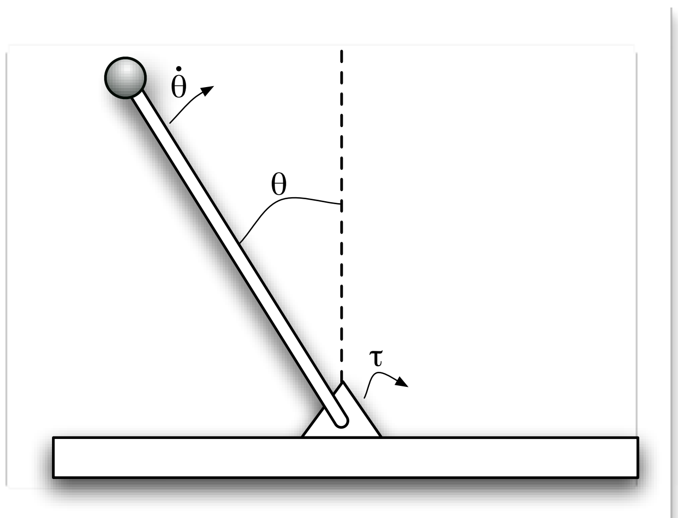
Incremental Feature Dependency Discovery



Empirical Results

- Representations used with Sarsa
  - initial
  - iFDD
  - ATC [Whiteson et al. 2007]
  - SDM [Ratitch et al. 2004]
  - Tabular

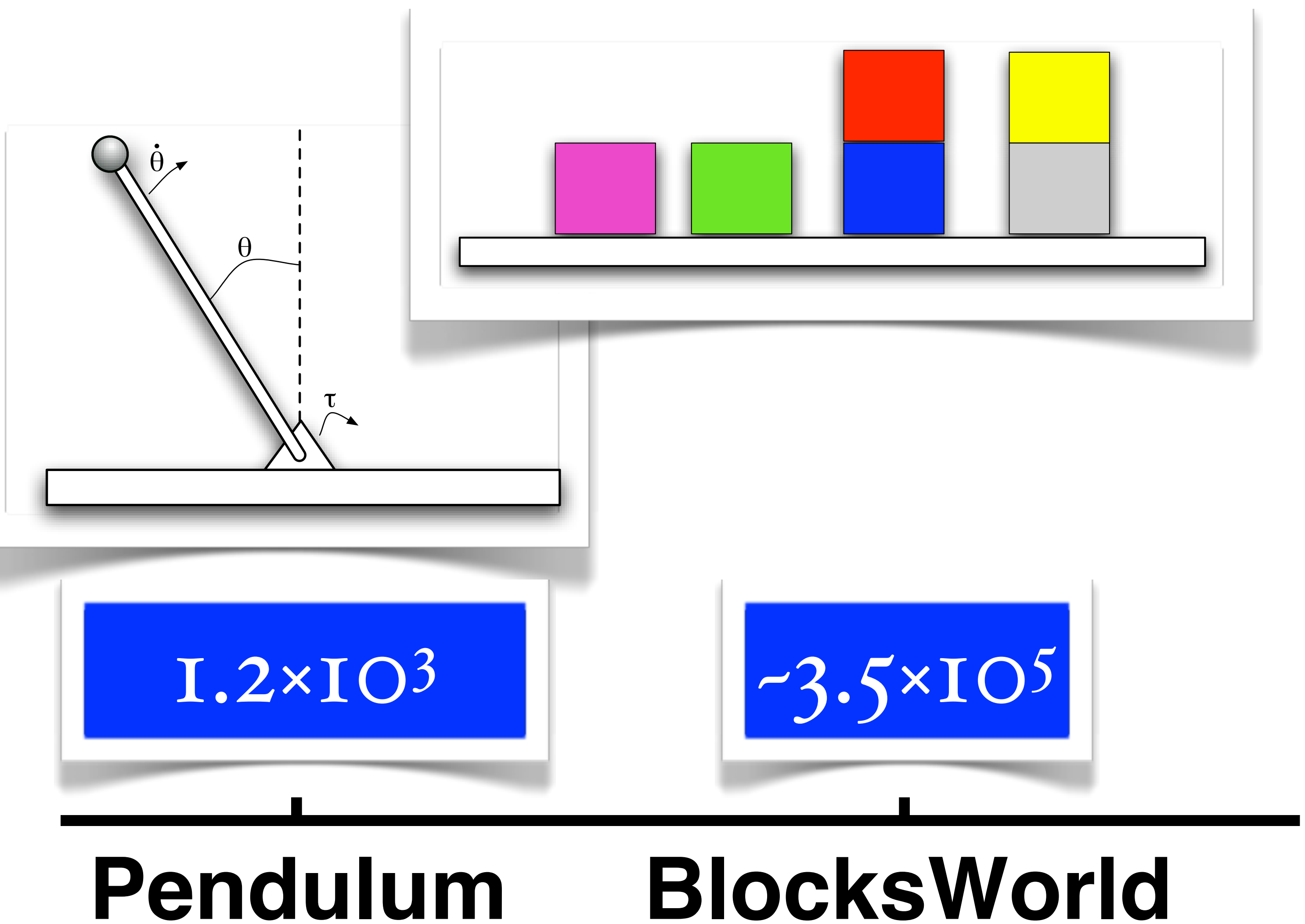
Domains



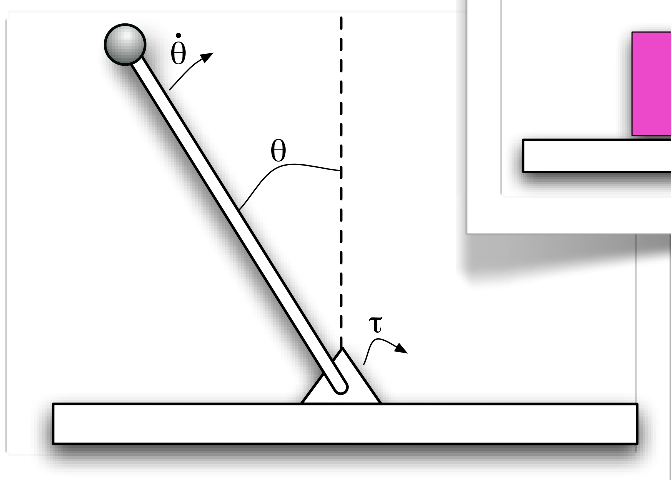
1.2×10^3

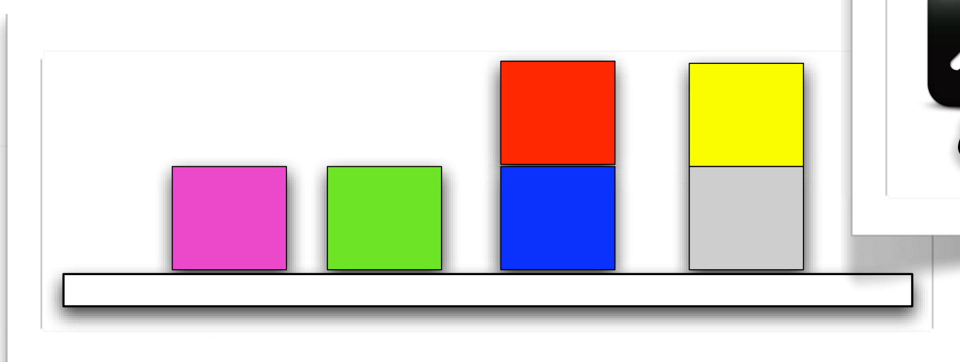
Pendulum

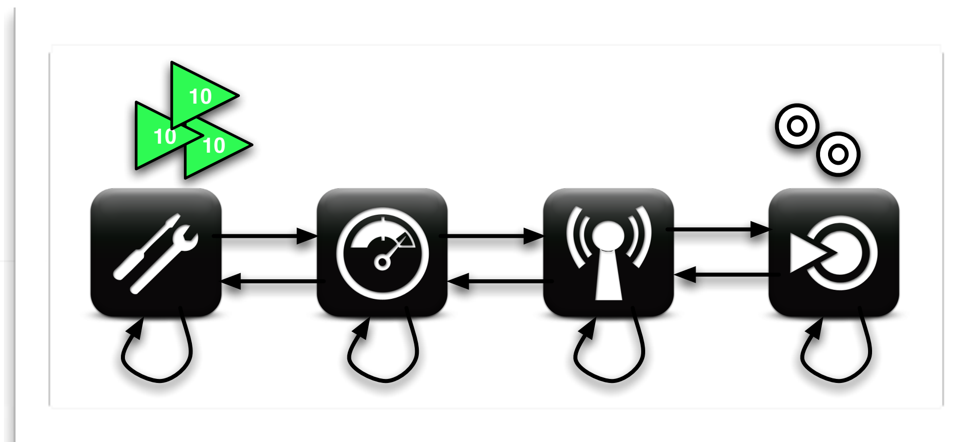
Domains



Domains







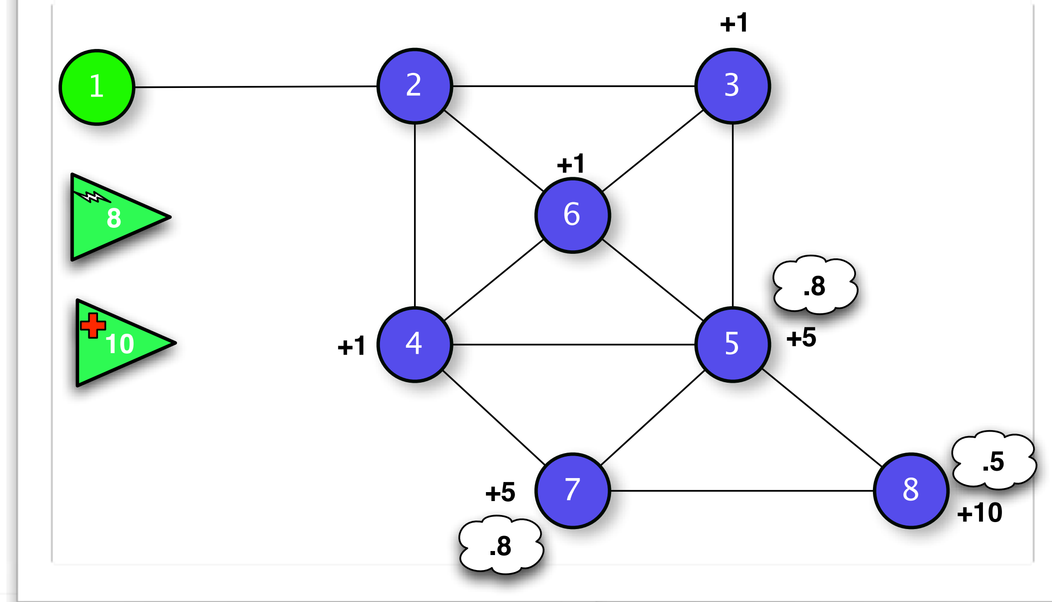
1.2×10^3

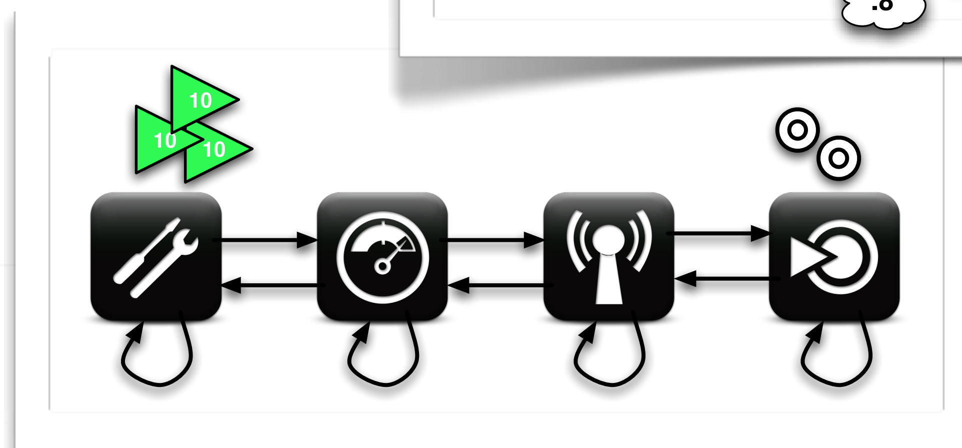
$\sim3.5\times10^5$

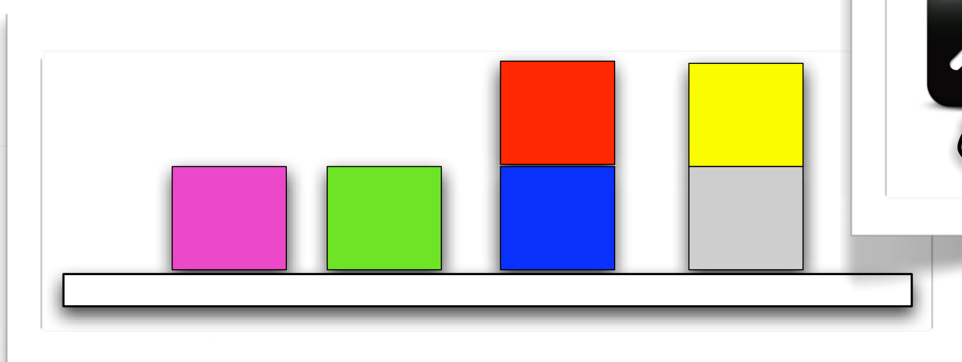
$\sim1.5\times10^8$

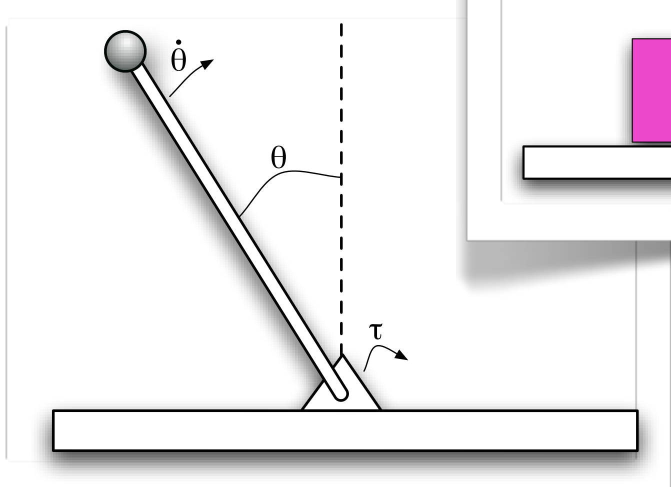
Pendulum **BlocksWorld** **PSM**

Domains





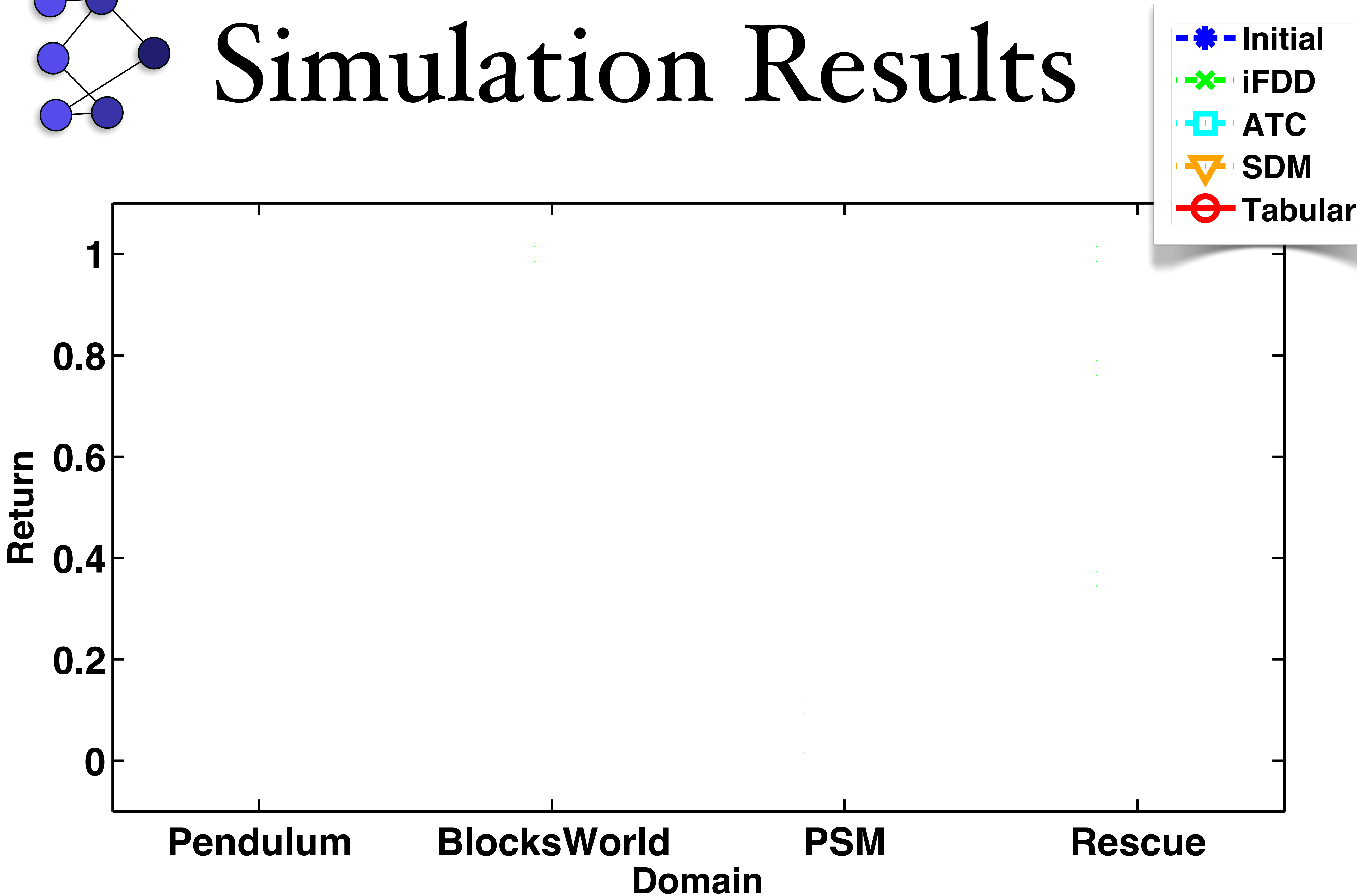




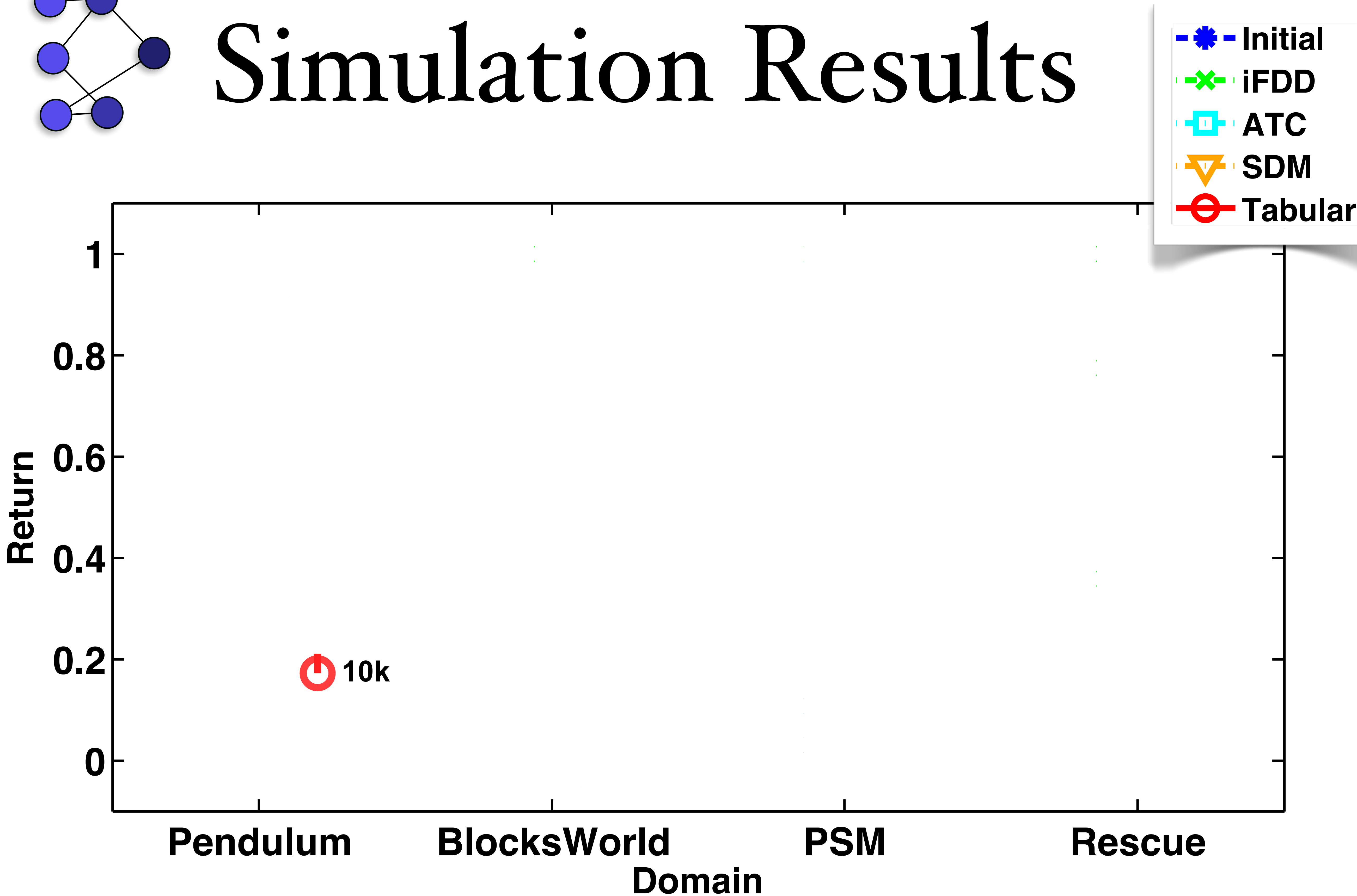
1.2×10^3 $\sim3.5\times10^5$ $\sim1.5\times10^8$ $\sim2\times10^8$

Pendulum **BlocksWorld** **PSM** **Rescue**

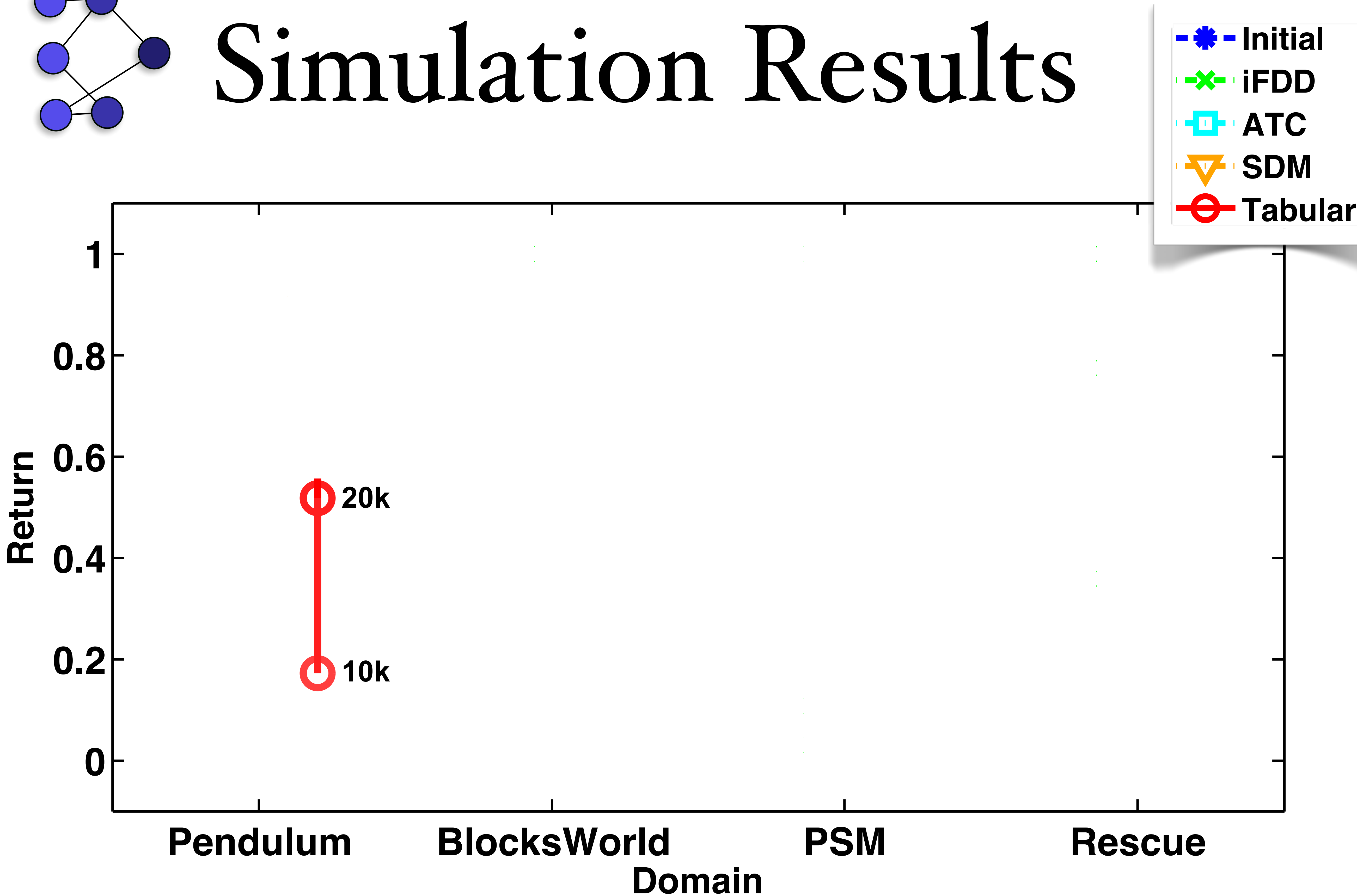
Simulation Results



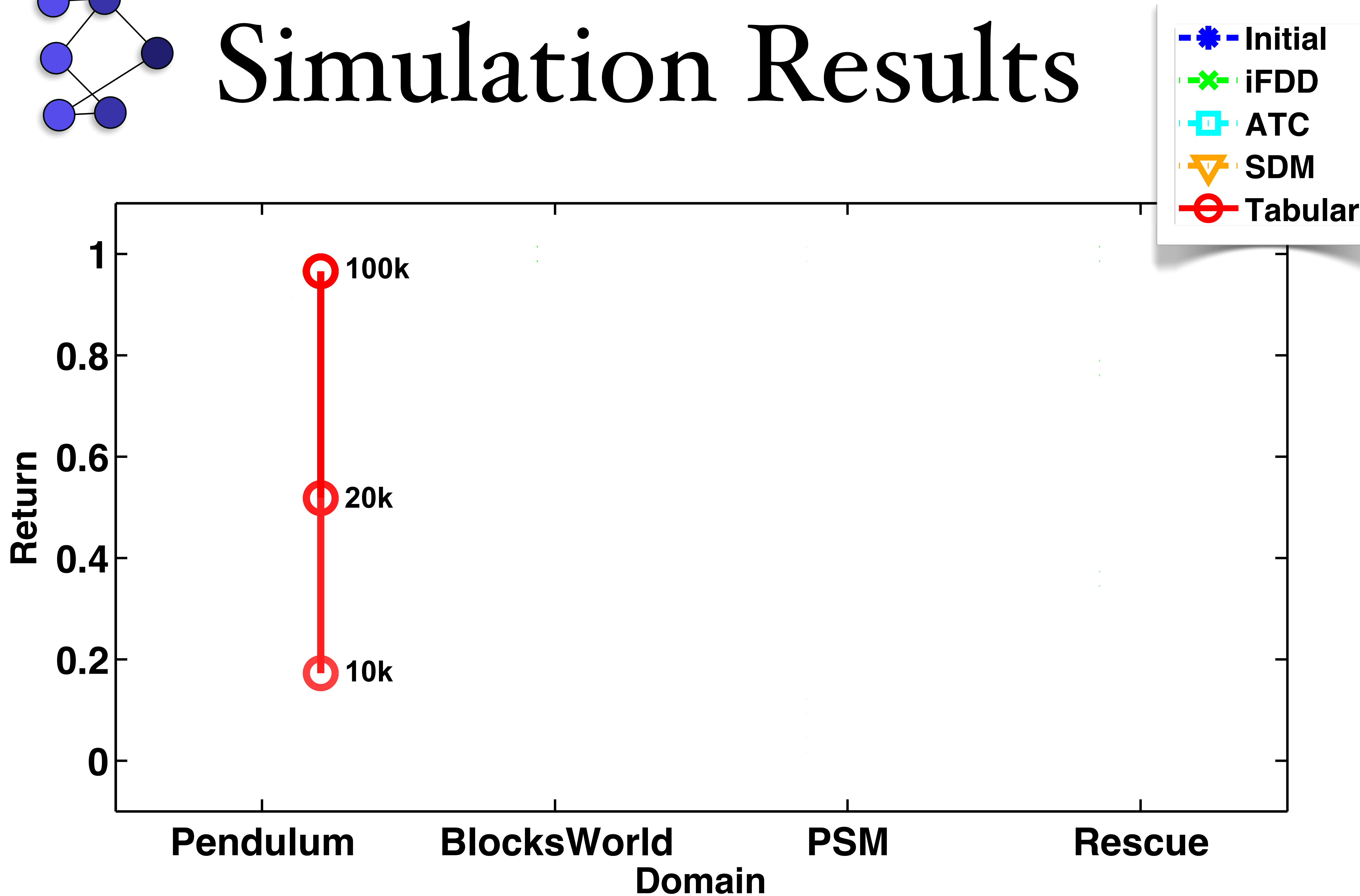
Simulation Results



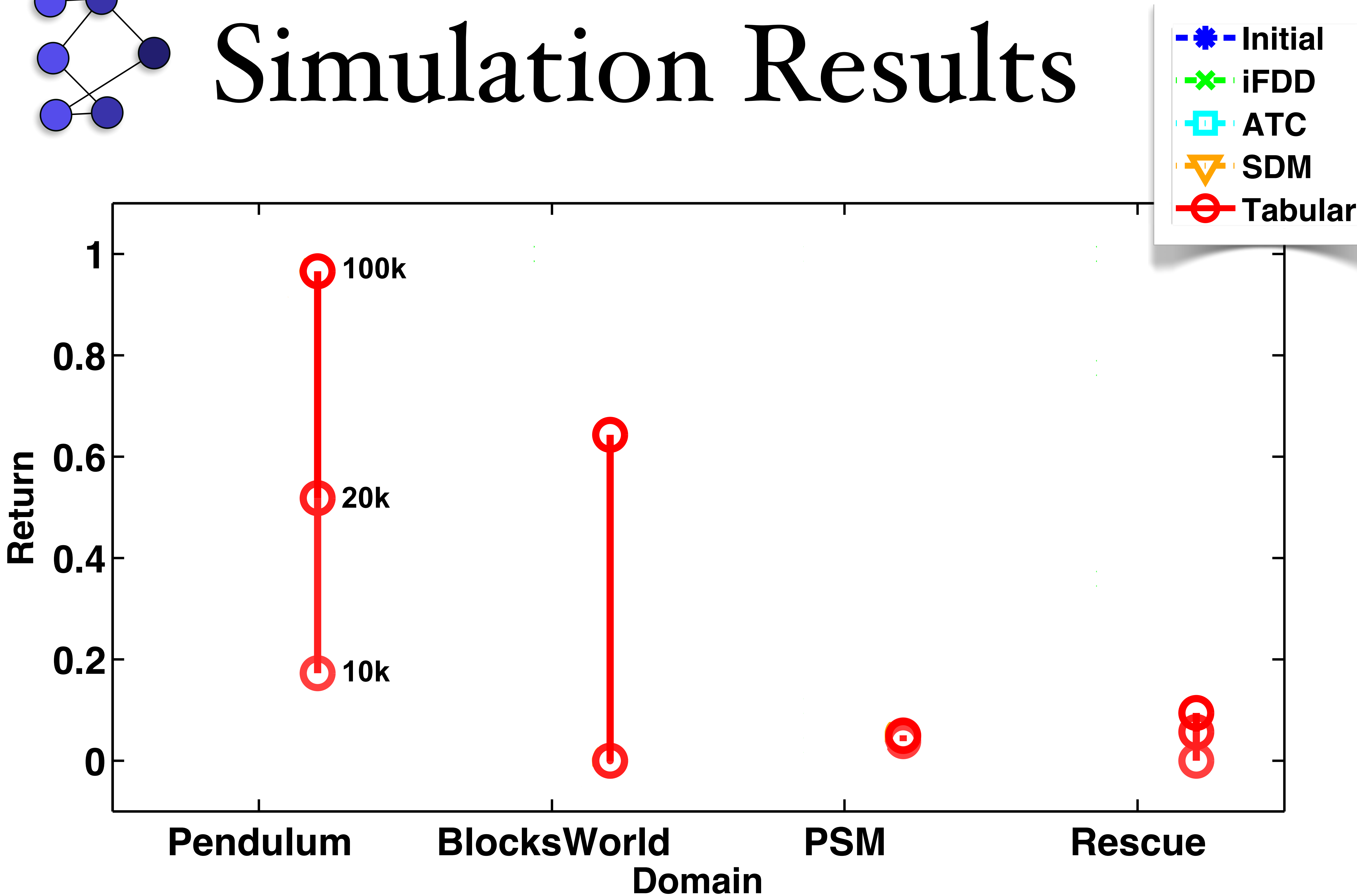
Simulation Results



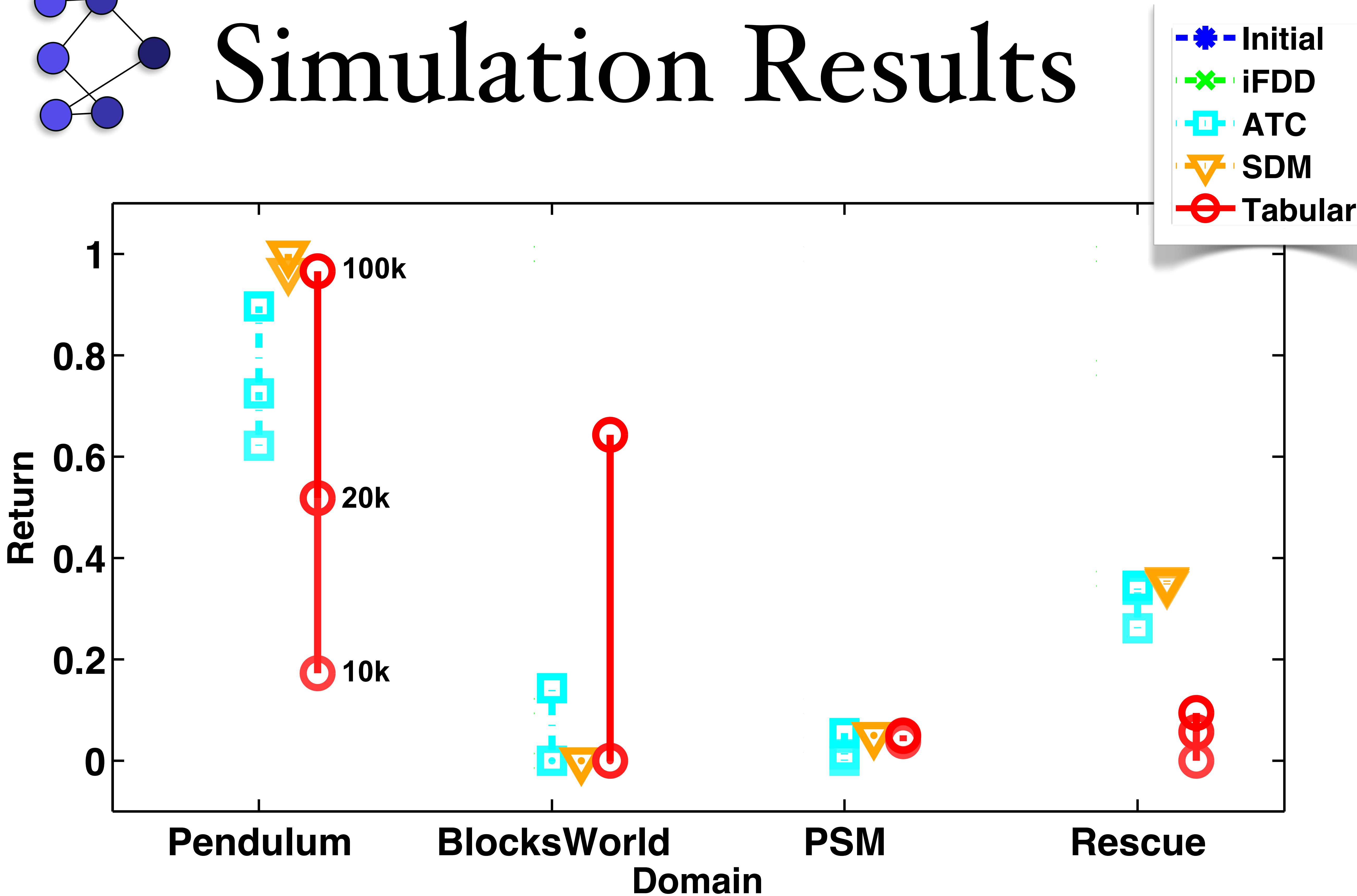
Simulation Results



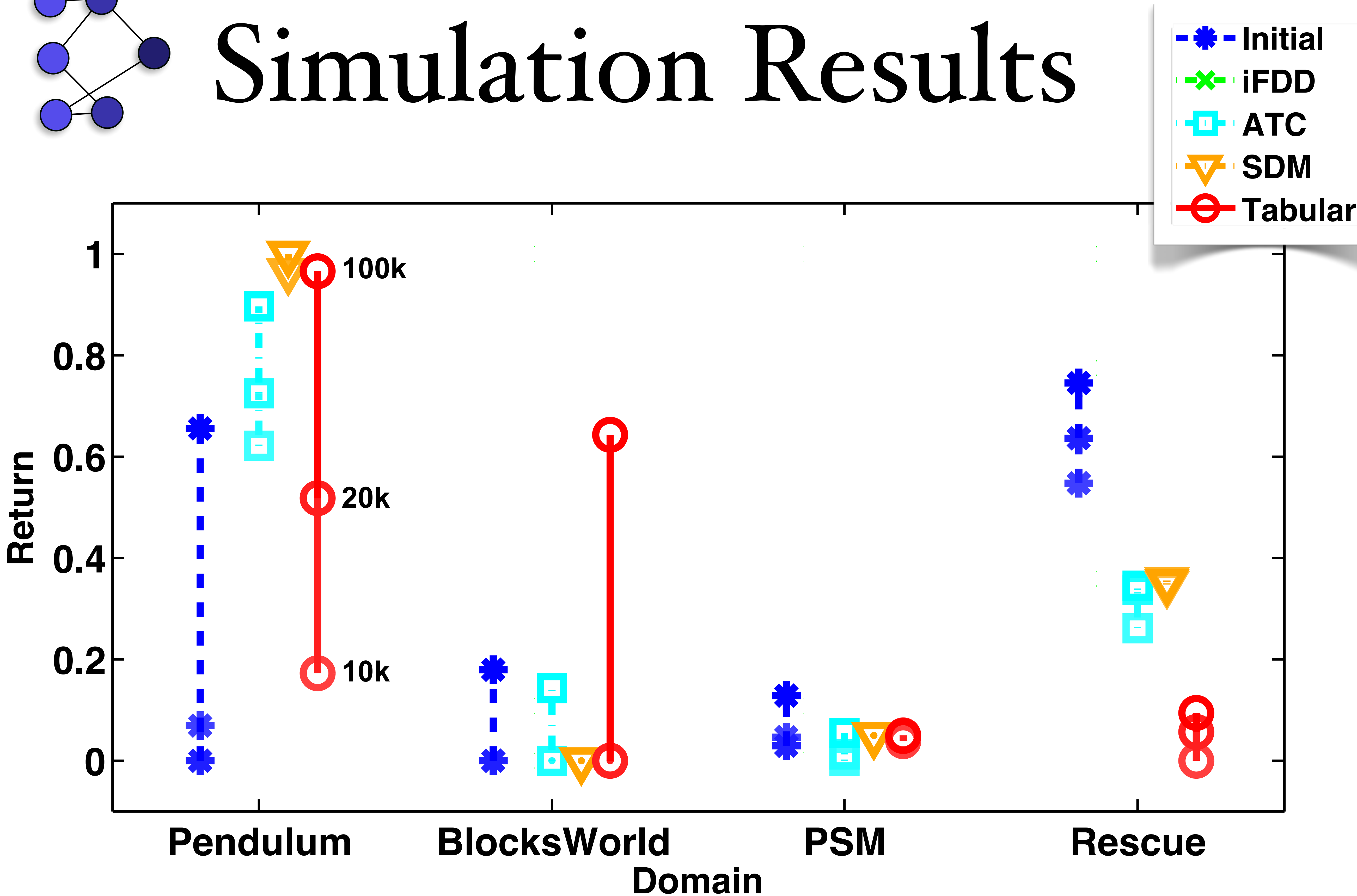
Simulation Results



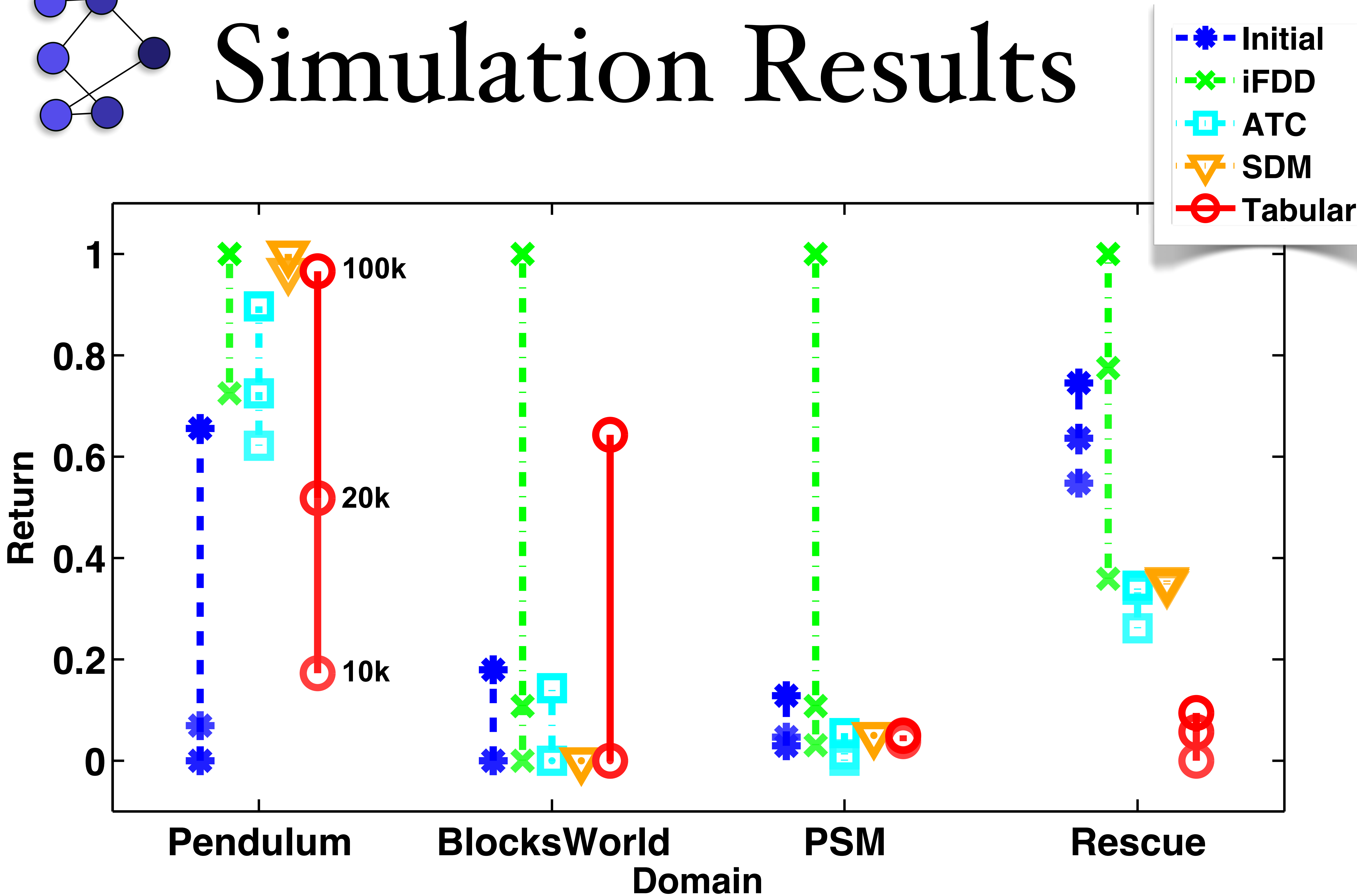
Simulation Results



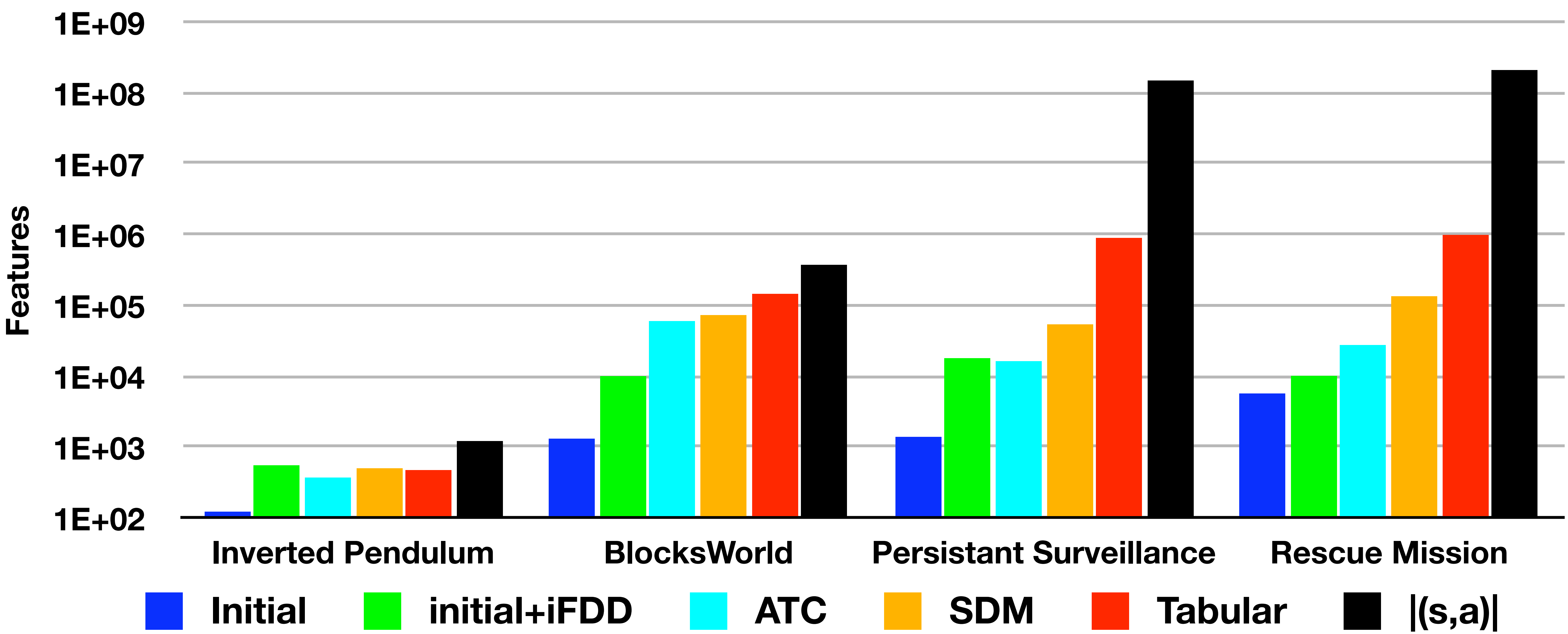
Simulation Results



Simulation Results



Simulation Results



iFDD Theory

- TD-iFDD will provide the **best** possible approximation given the initial set of features.
- Given initial features with sparse outputs, the per-time-step computational complexity of iFDD is **independent** of the total number of features.

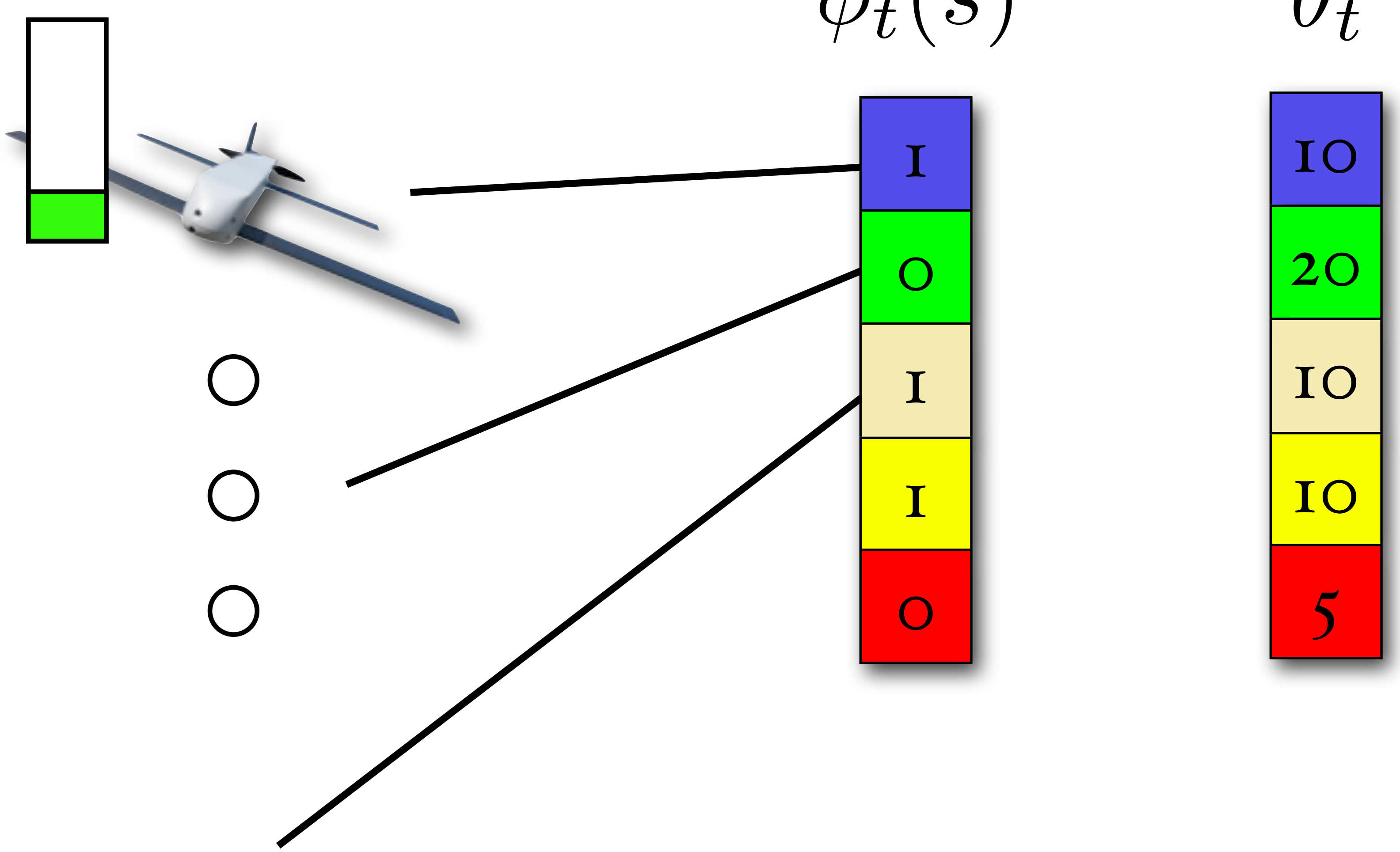
Contributions

- Introduced **iFDD** as a novel feature expansion method
- Provided asymptotic **convergence** analysis
- Empirically showed the **scalability** of the new approach in problem sizes $\boldsymbol{\approx 10^8}$

Backup Slides

LFA: Example

State | Feature $\phi_t(s)$ | Weight θ_t | Value



$$V(s) = 20+10+10 = 40$$

Algorithms

Algorithm 1: Discover

Input: $\phi(s), \delta_t, \xi, \mathbf{F}, \psi$

Output: $\mathbf{F}, \psi$

foreach $(g, h) \in \{(i, j) | \phi_i(s)\phi_j(s) = 1\}$ **do**

  $f \leftarrow g \wedge h$

  if $f \notin \mathbf{F}$ **then**

    $\psi_f \leftarrow \psi_f + |\delta_t|$

    if $\psi_f > \xi$ **then**

      $\mathbf{F} \leftarrow \mathbf{F} \cup f$

    end

  end

end

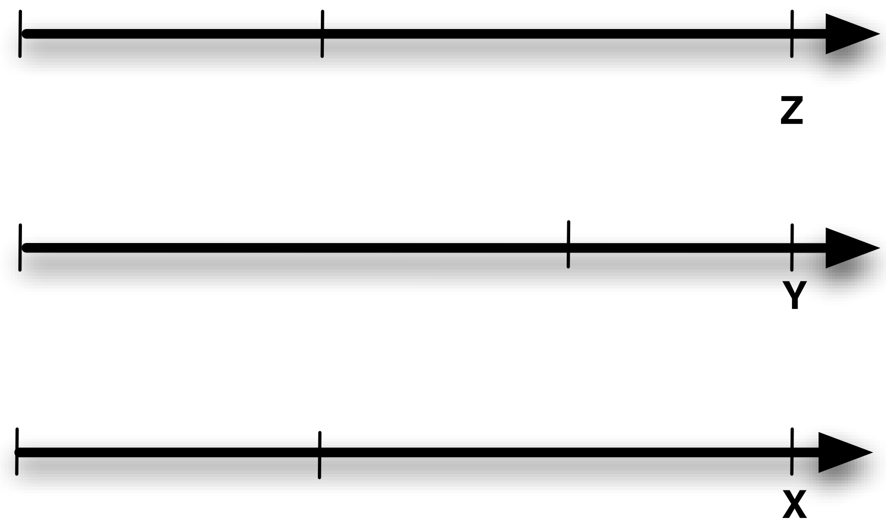
Algorithm 2: Activate Features

Input: $\phi^0(s)$, $\mathbf{F}$

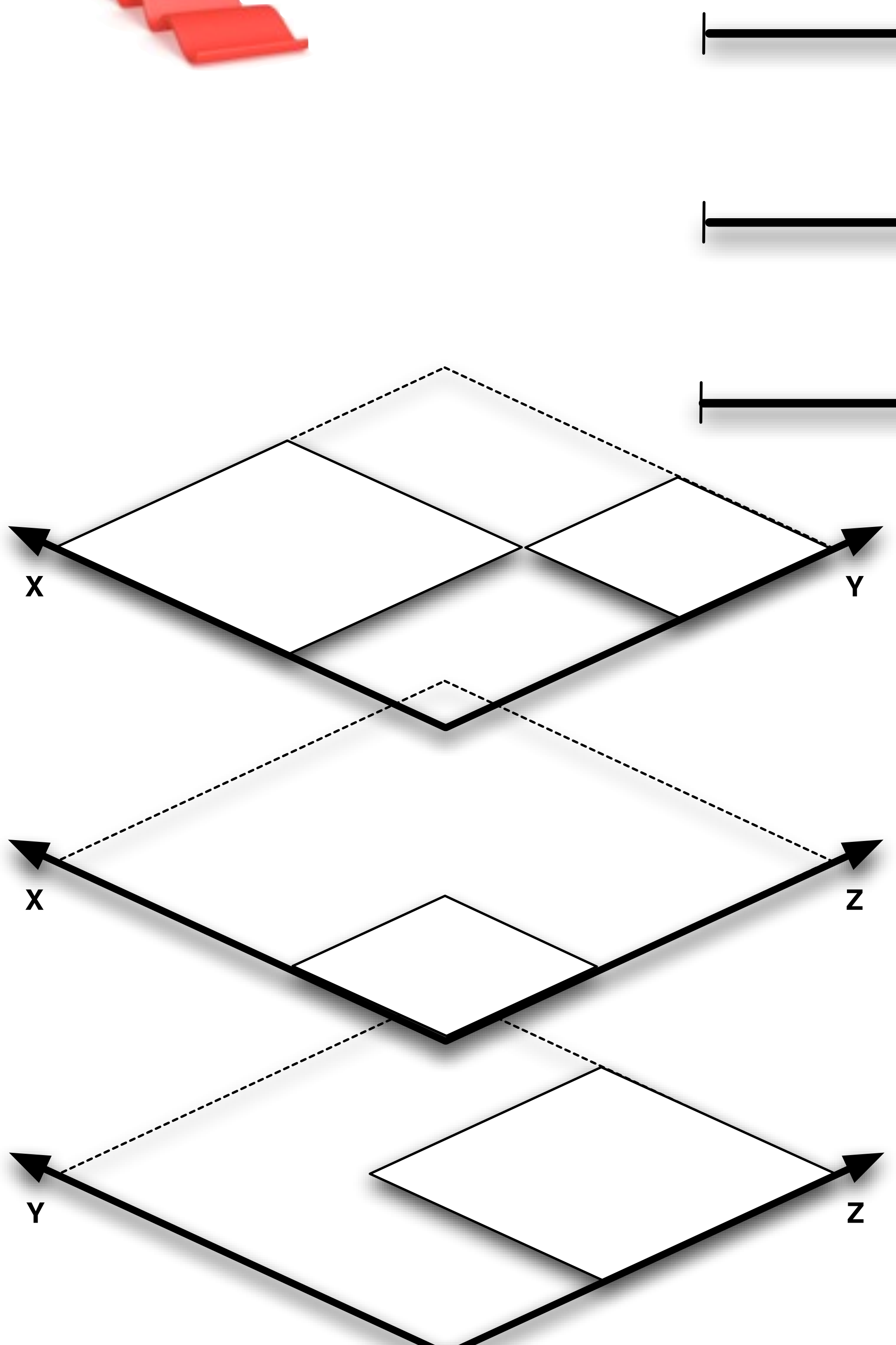
Output: $\phi(s)$

$\phi(s) \leftarrow \bar{0}$
$activeInitialFeatures \leftarrow \{i | \phi_i^0(s) = 1\}$
$Candidates \leftarrow \wp(activeInitialFeatures)$ (*sorted by set size)
while $activeInitialFeatures \neq \emptyset$ **do**
    $f \leftarrow Candidates$.next()
    if $f \in \mathbf{F}$ **then**
        $activeInitialFeatures \leftarrow activeInitialFeatures - f$
        $\phi_f(s) \leftarrow 1$
    end
end
return $\phi(s)$

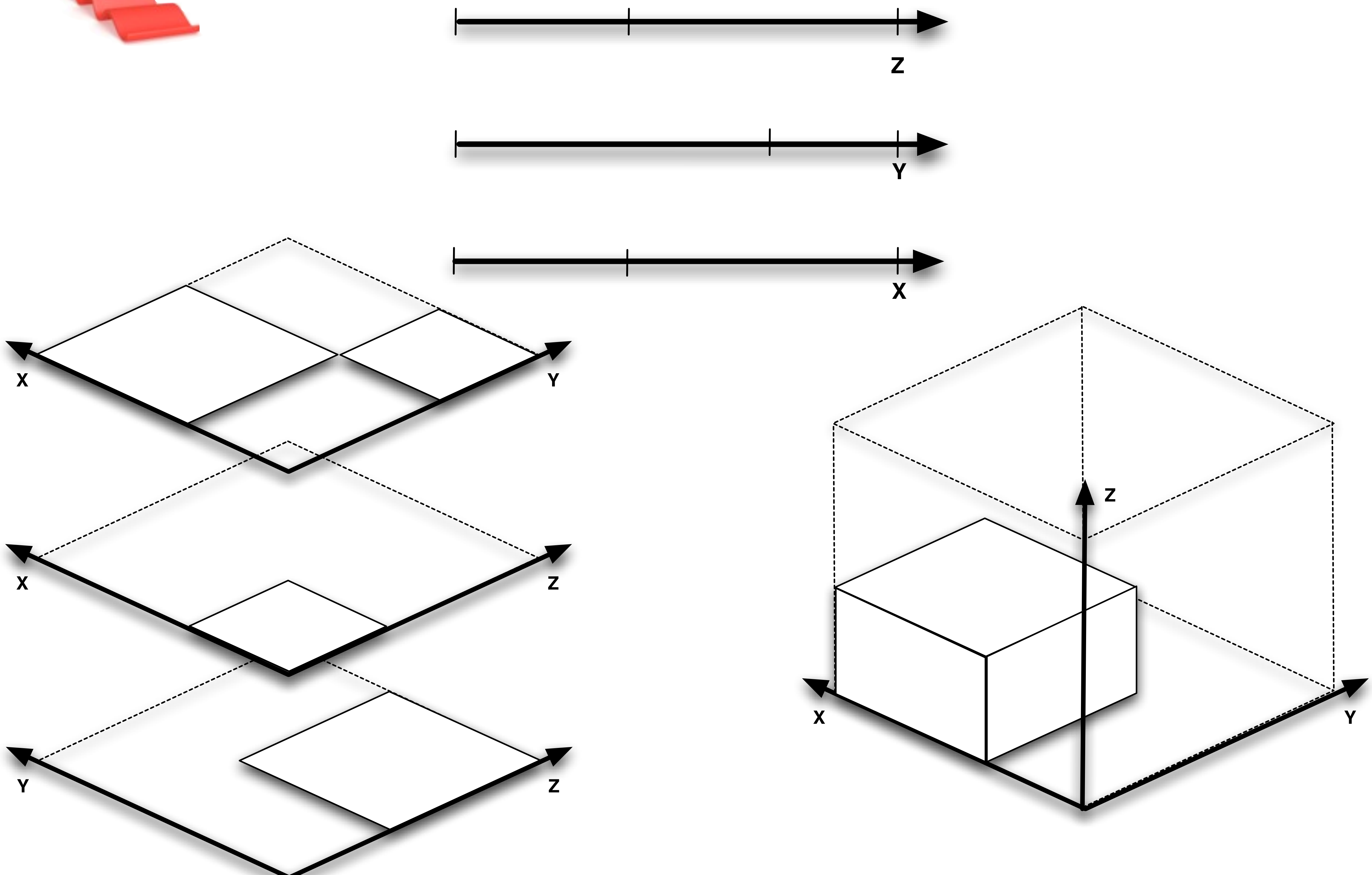
iFDD: 3D Example



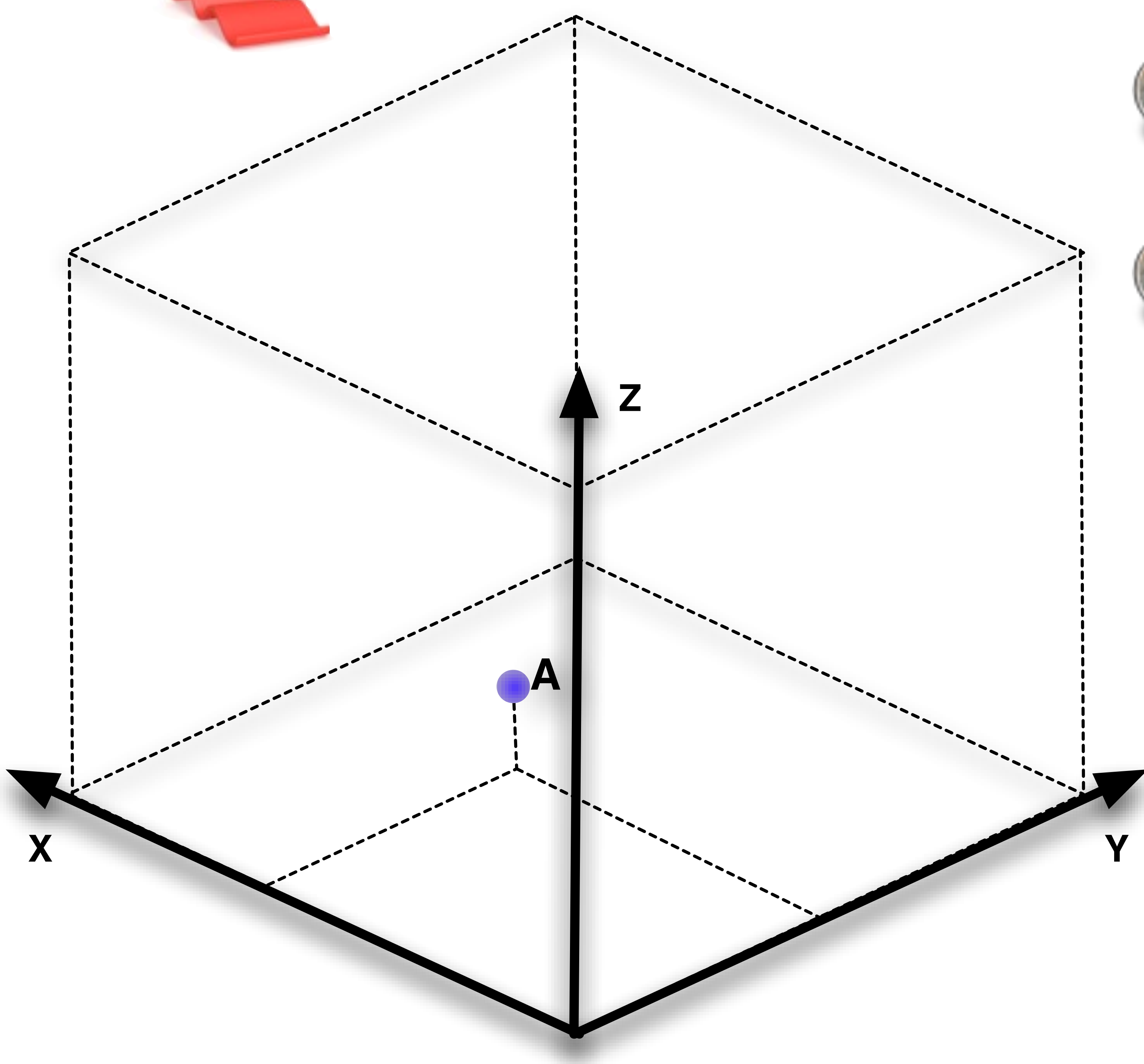
iFDD: 3D Example



iFDD: 3D Example



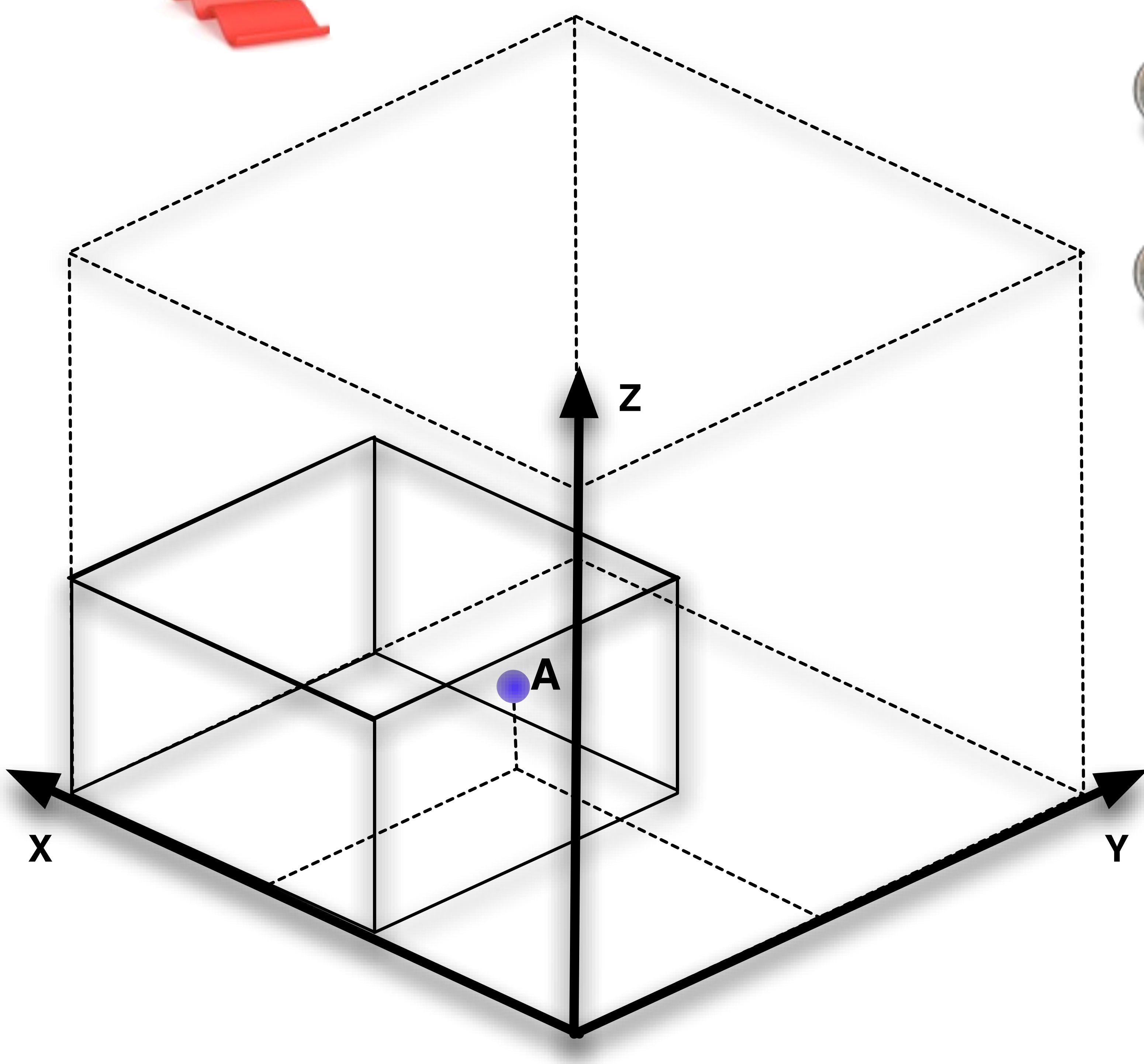
iFDD - Mapping



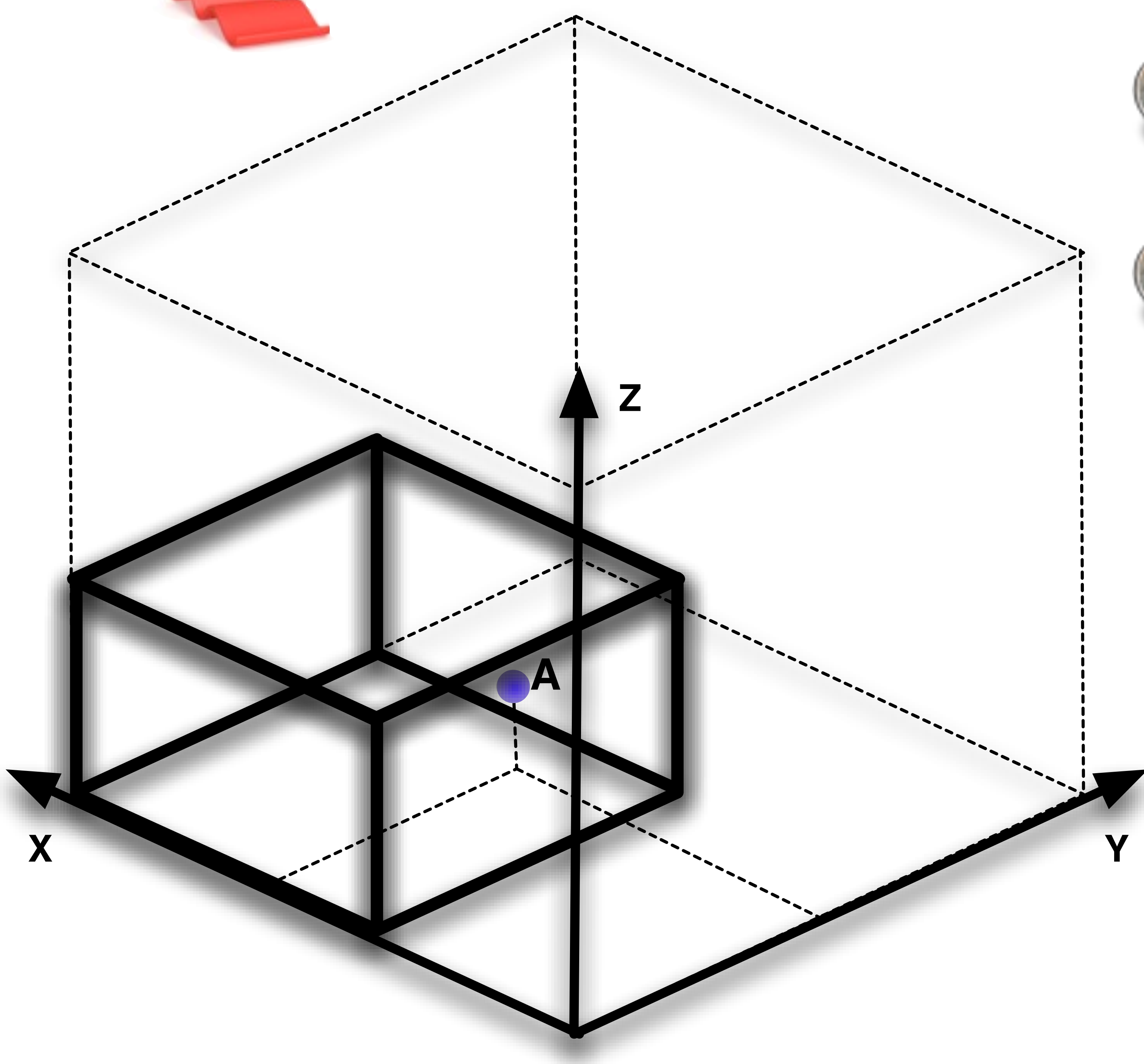
- Sort Layers
- Dropping a Stone

iFDD - Mapping

- Sort Layers
- Dropping a Stone

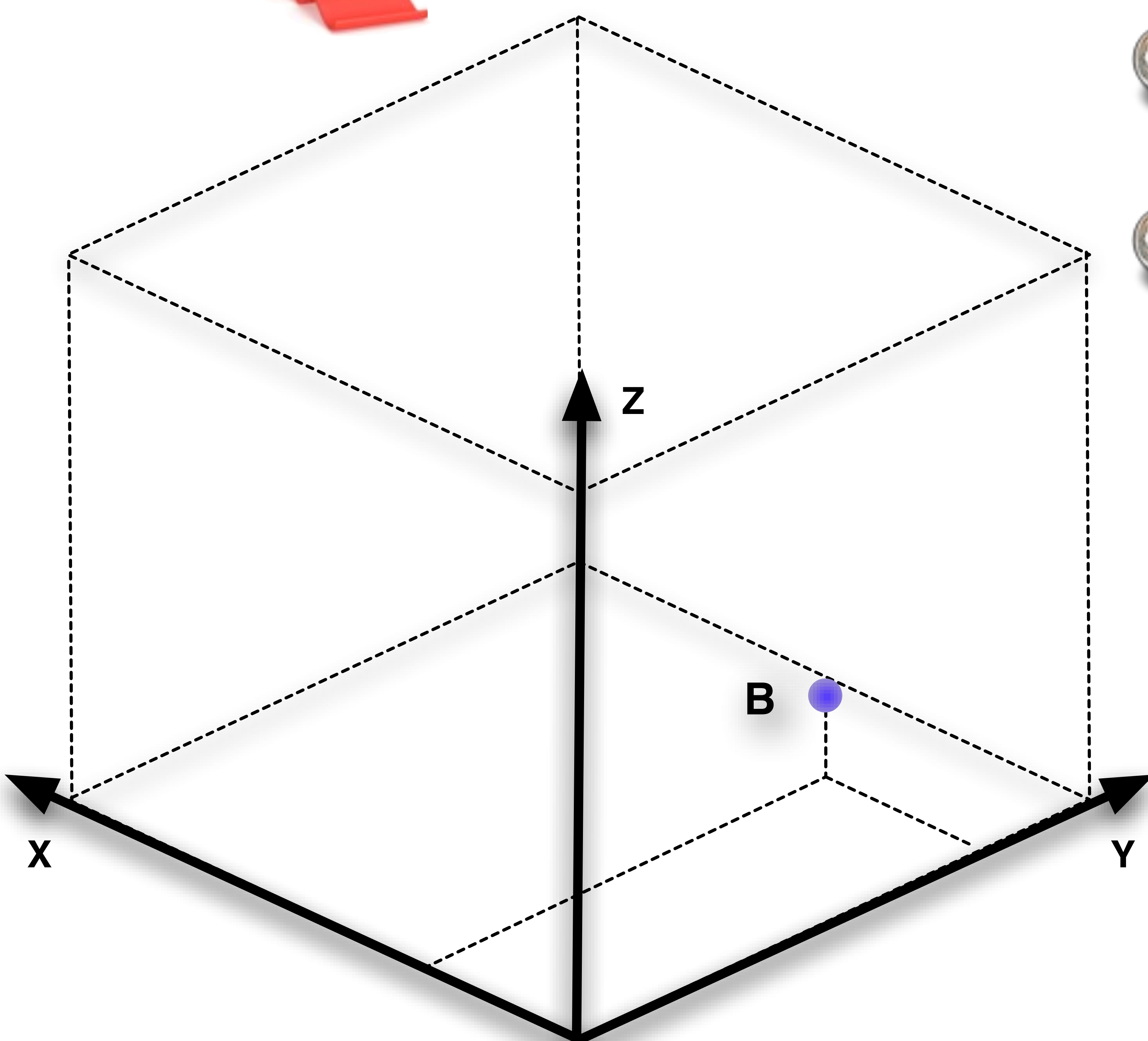


iFDD - Mapping



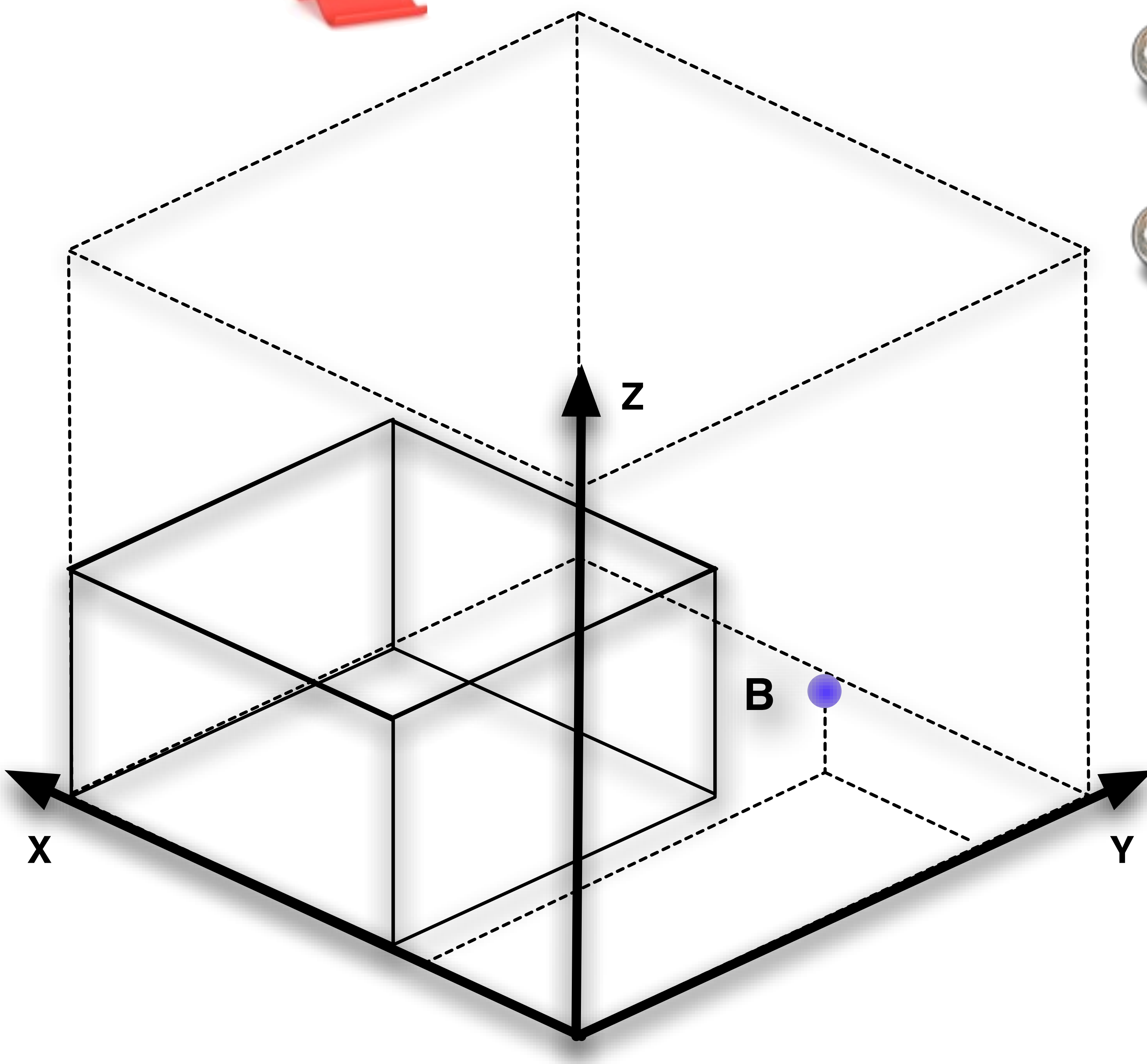
- Sort Layers
- Dropping a Stone

iFDD - Mapping



- Sort Layers
- Dropping a Stone

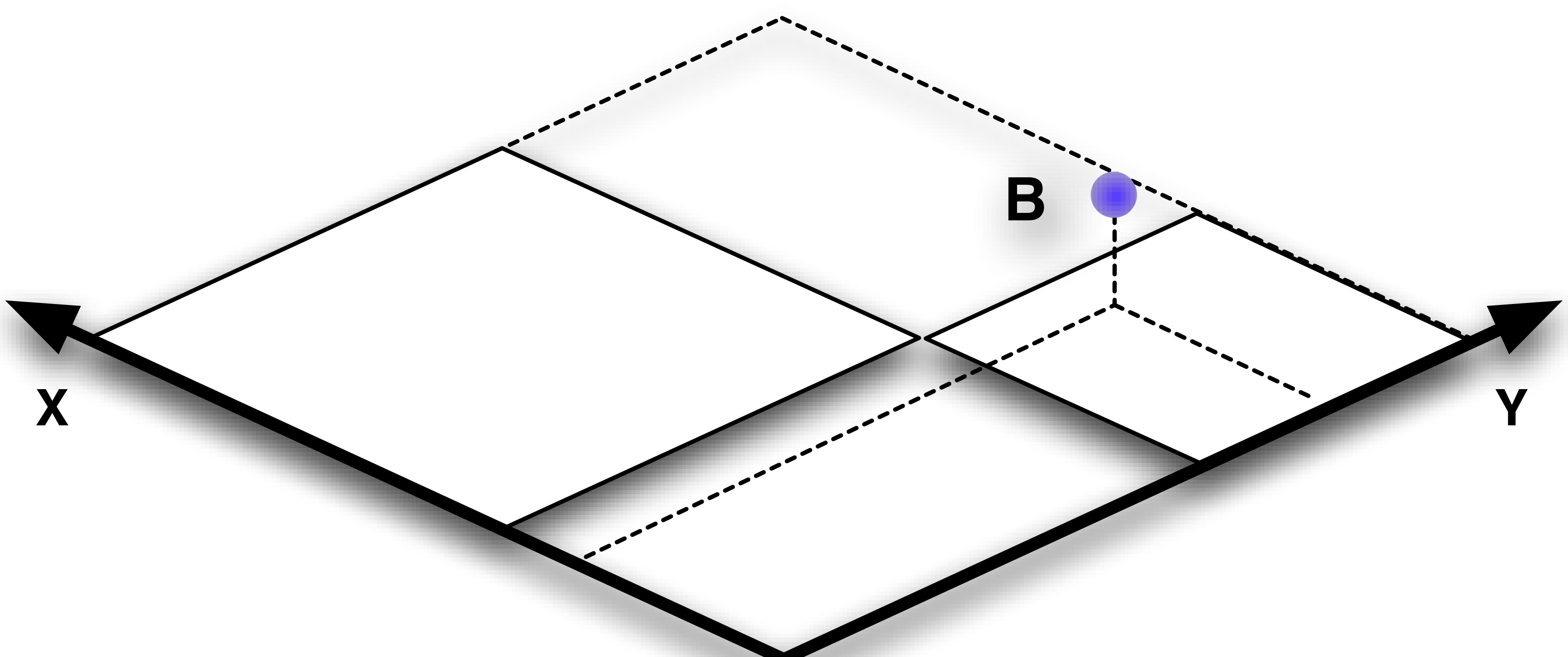
iFDD - Mapping



- Sort Layers
- Dropping a Stone

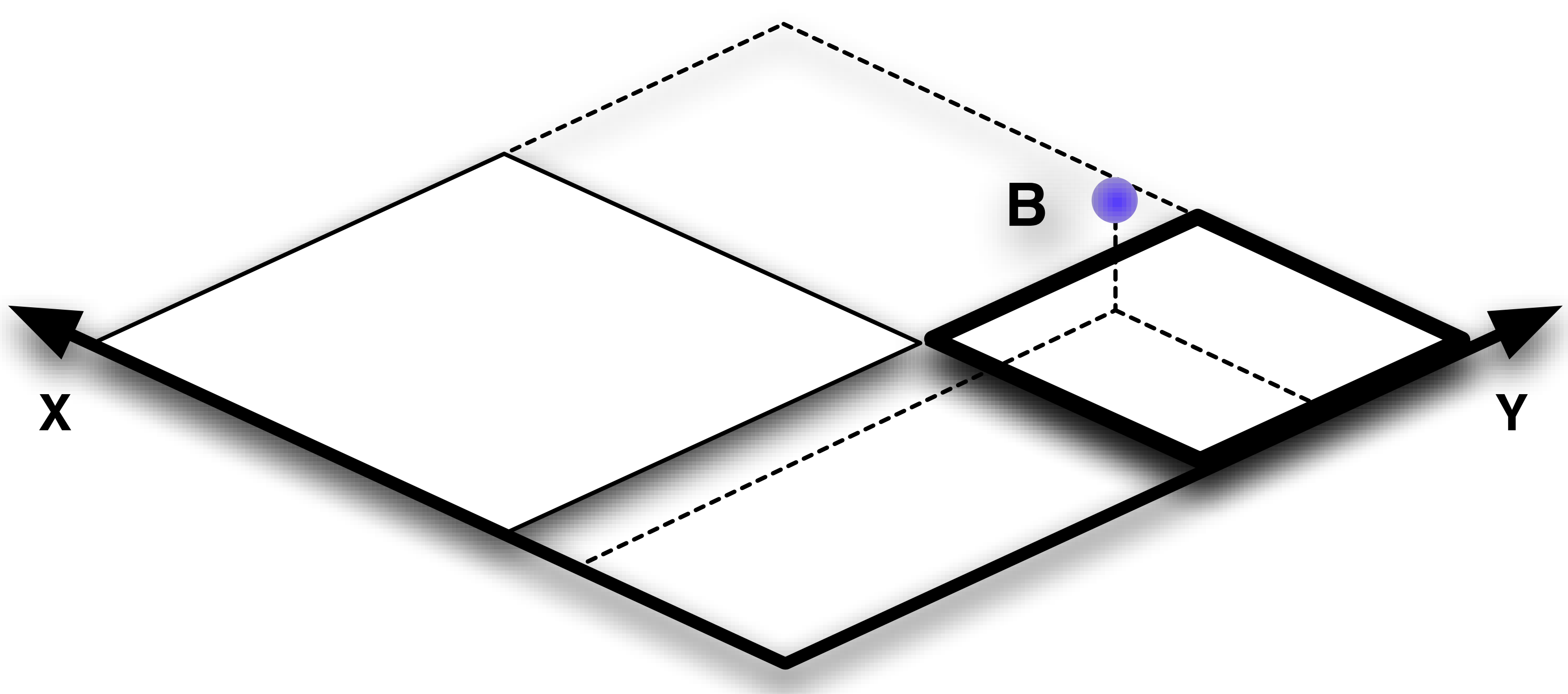
iFDD - Mapping

- Sort Layers
- Dropping a Stone



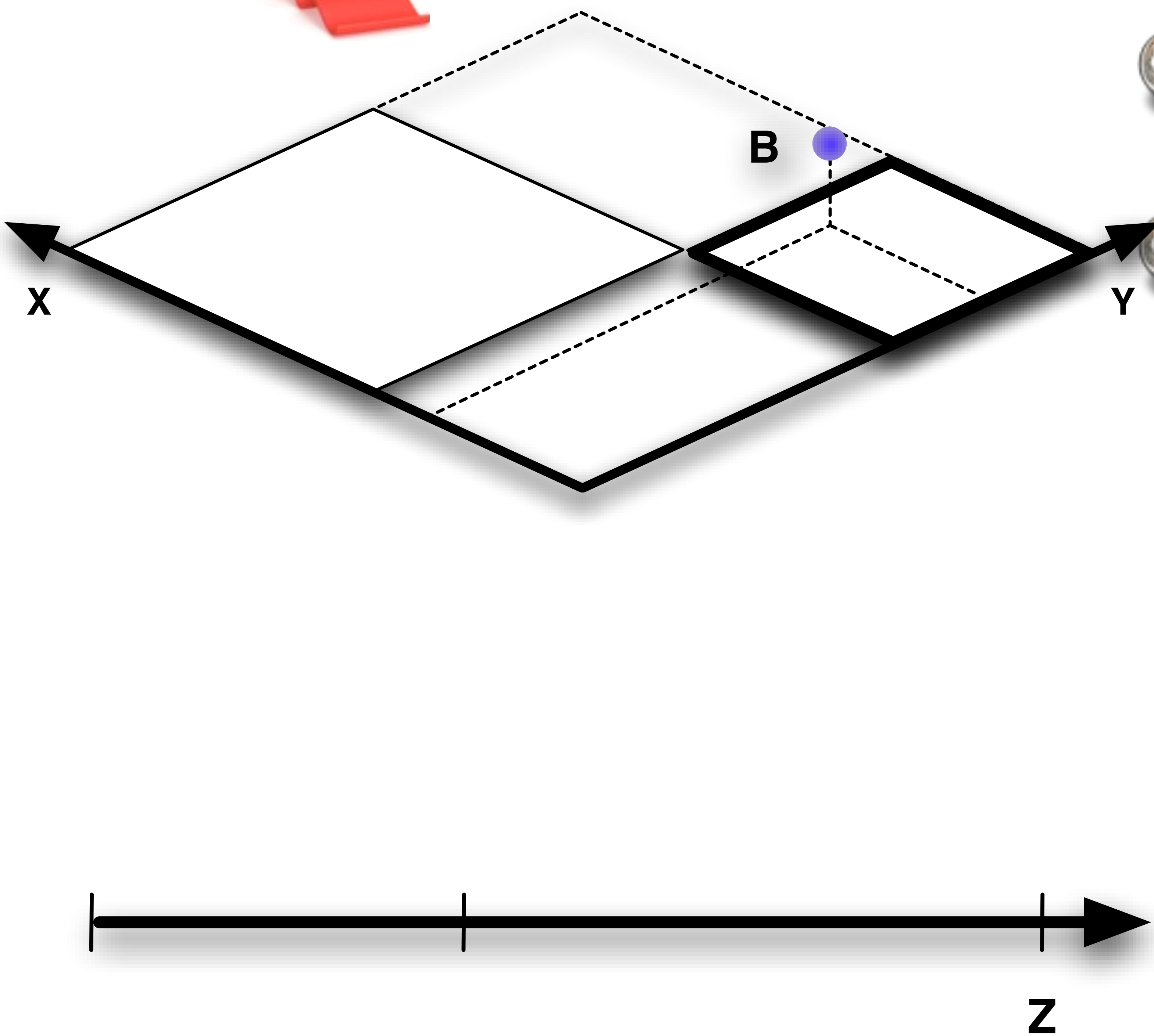
iFDD - Mapping

- Sort Layers
- Dropping a Stone

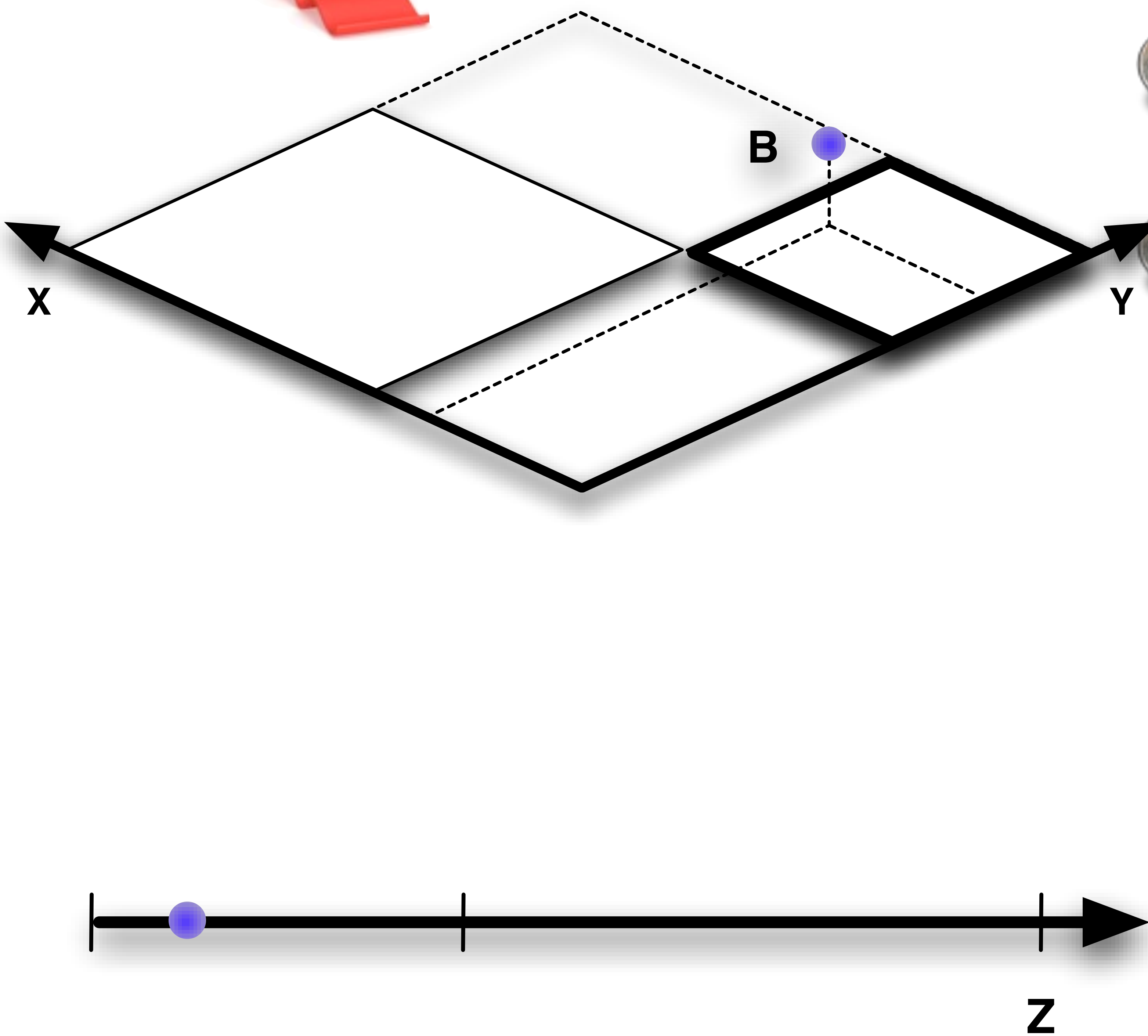


iFDD - Mapping

- Sort Layers
- Dropping a Stone



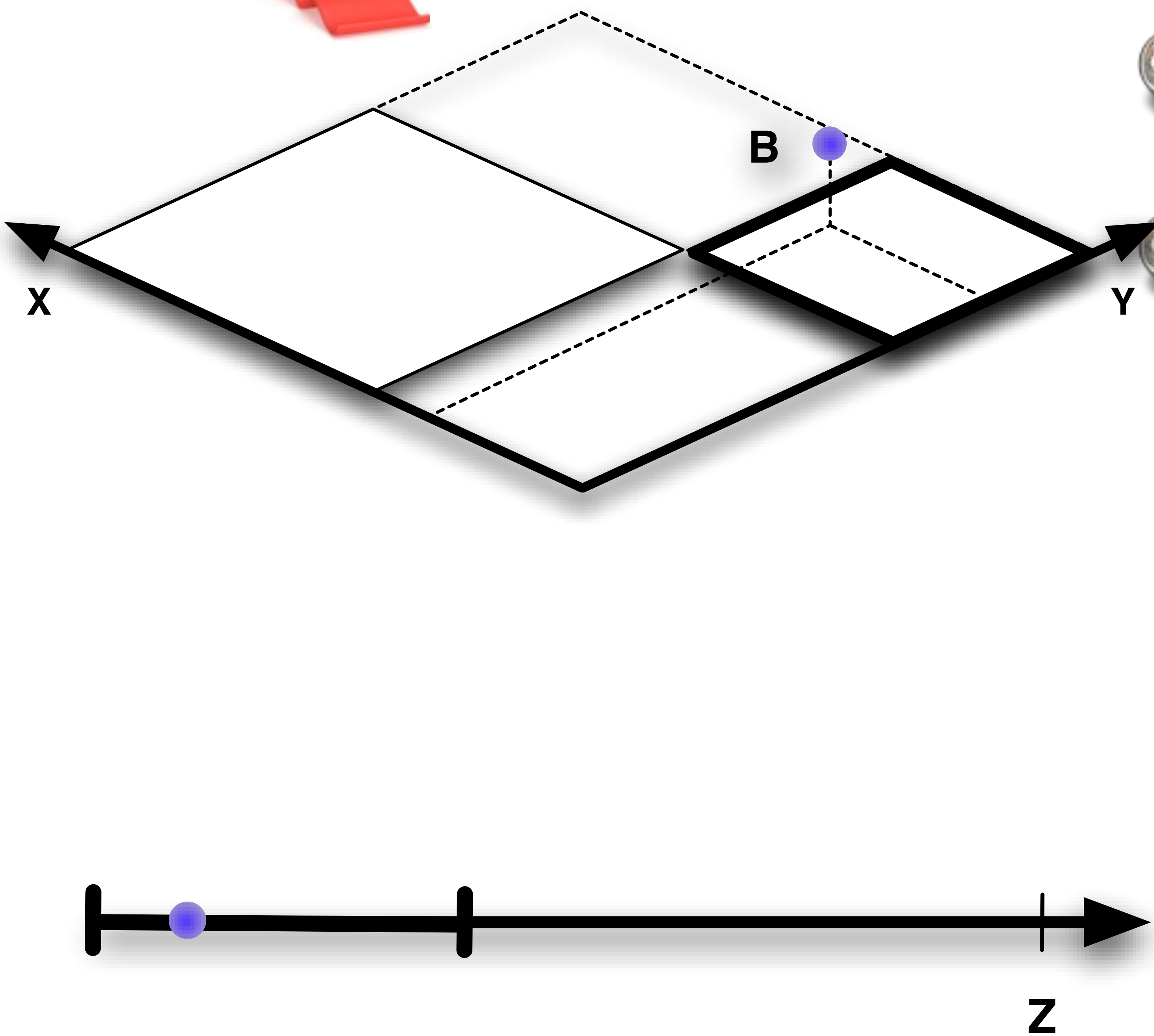
iFDD - Mapping



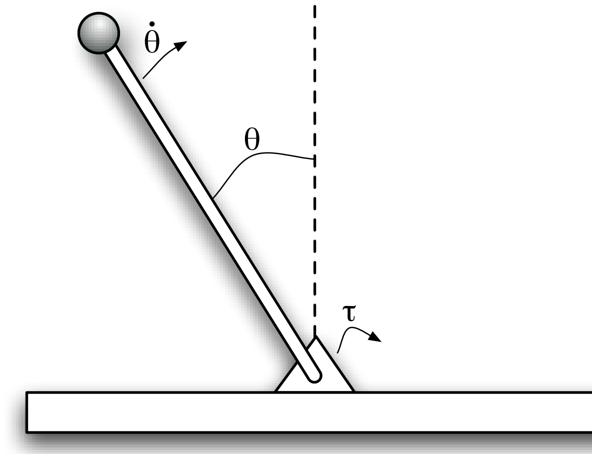
- Sort Layers
- Dropping a Stone

iFDD - Mapping

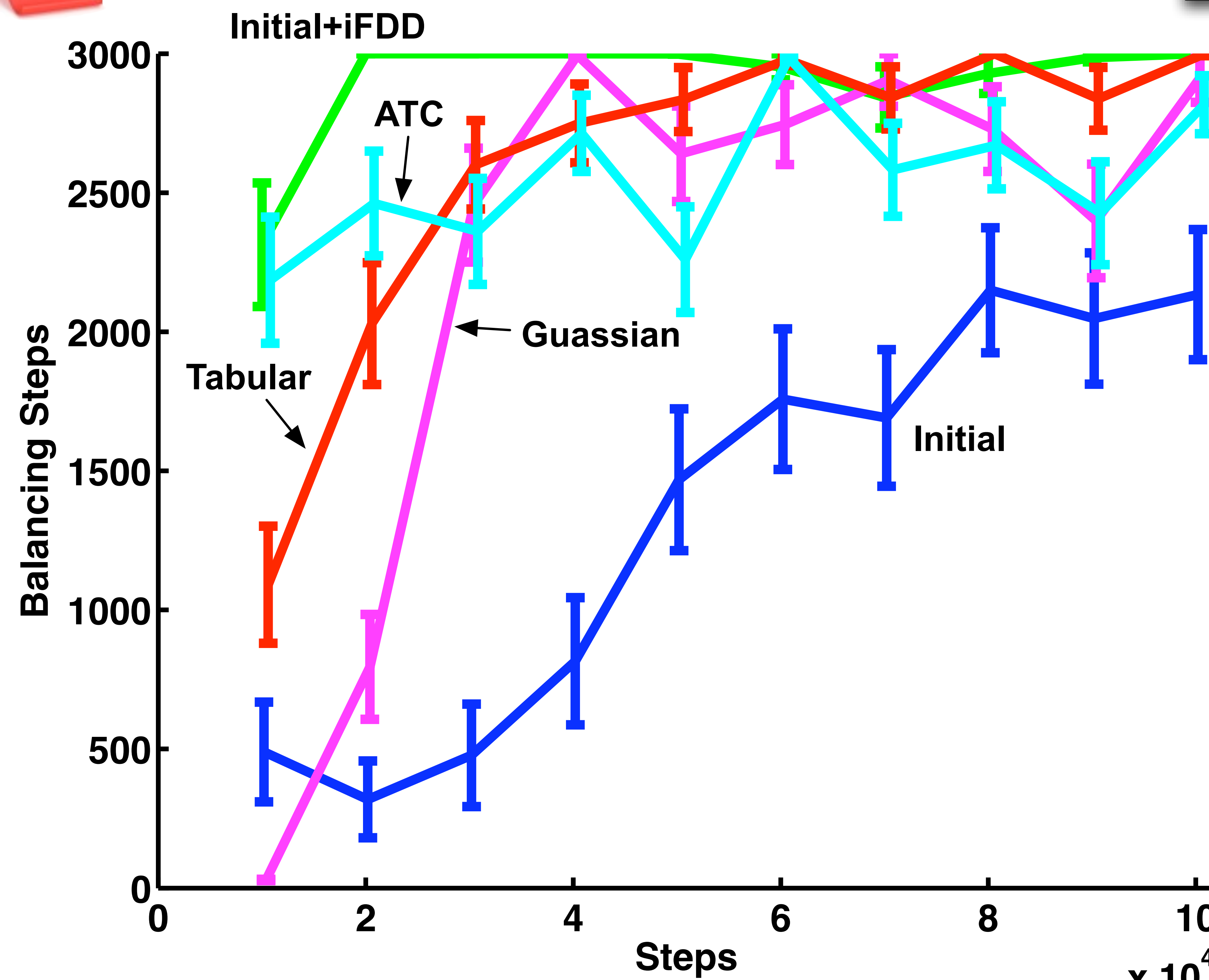
- Sort Layers
- Dropping a Stone



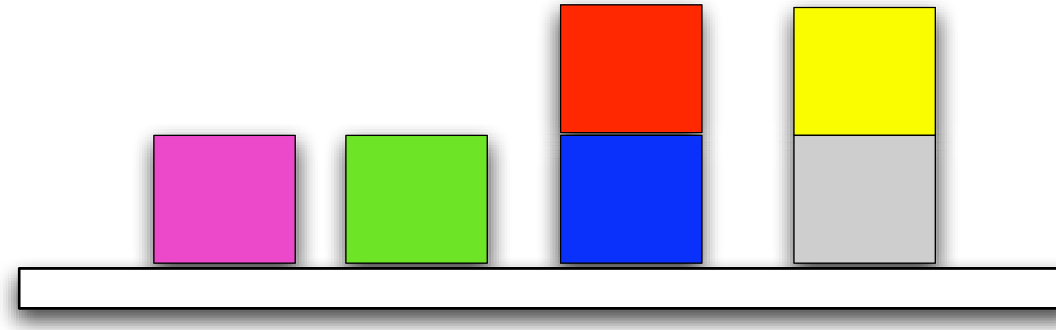
Detailed Results

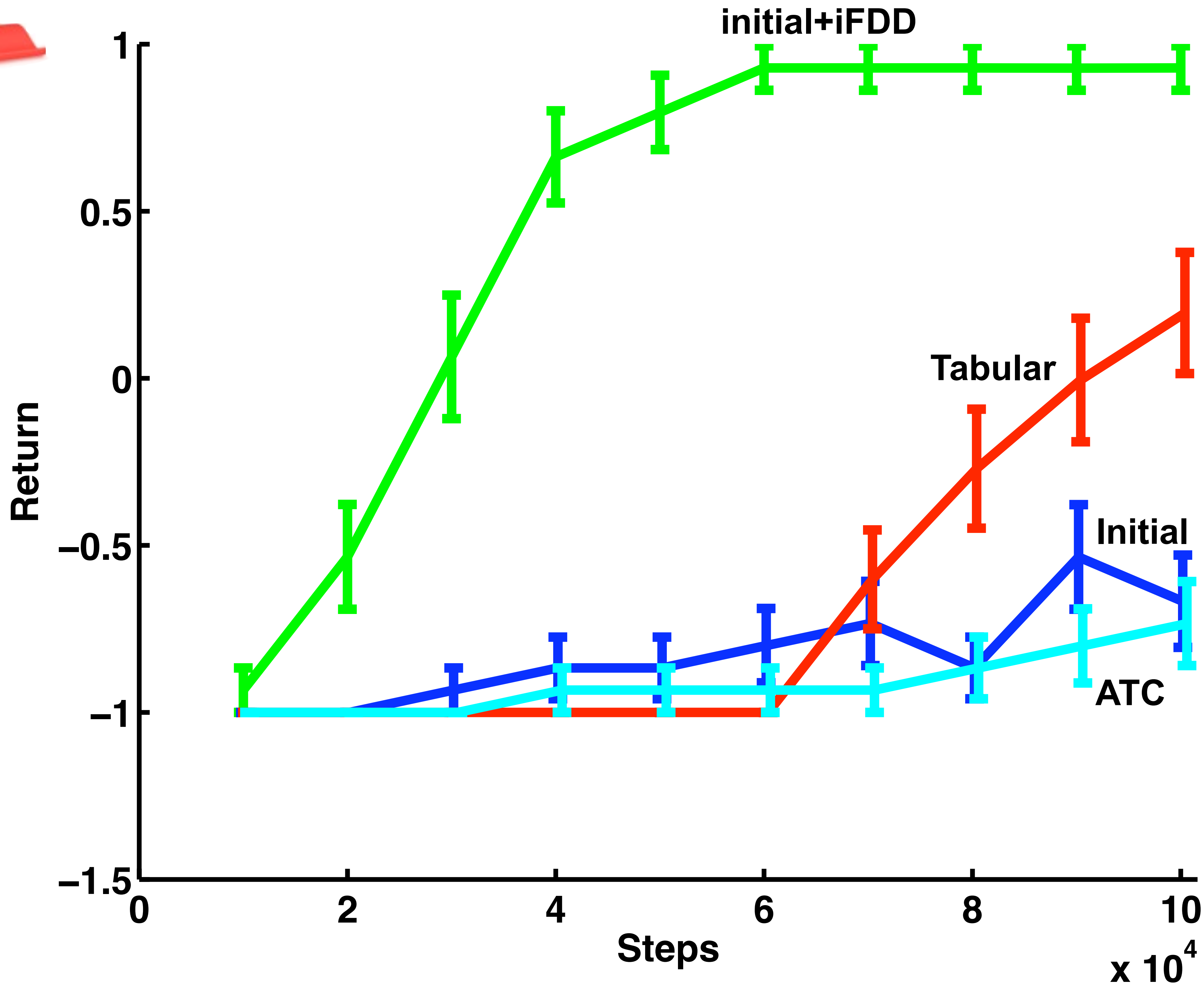


θ̇
θ
τ

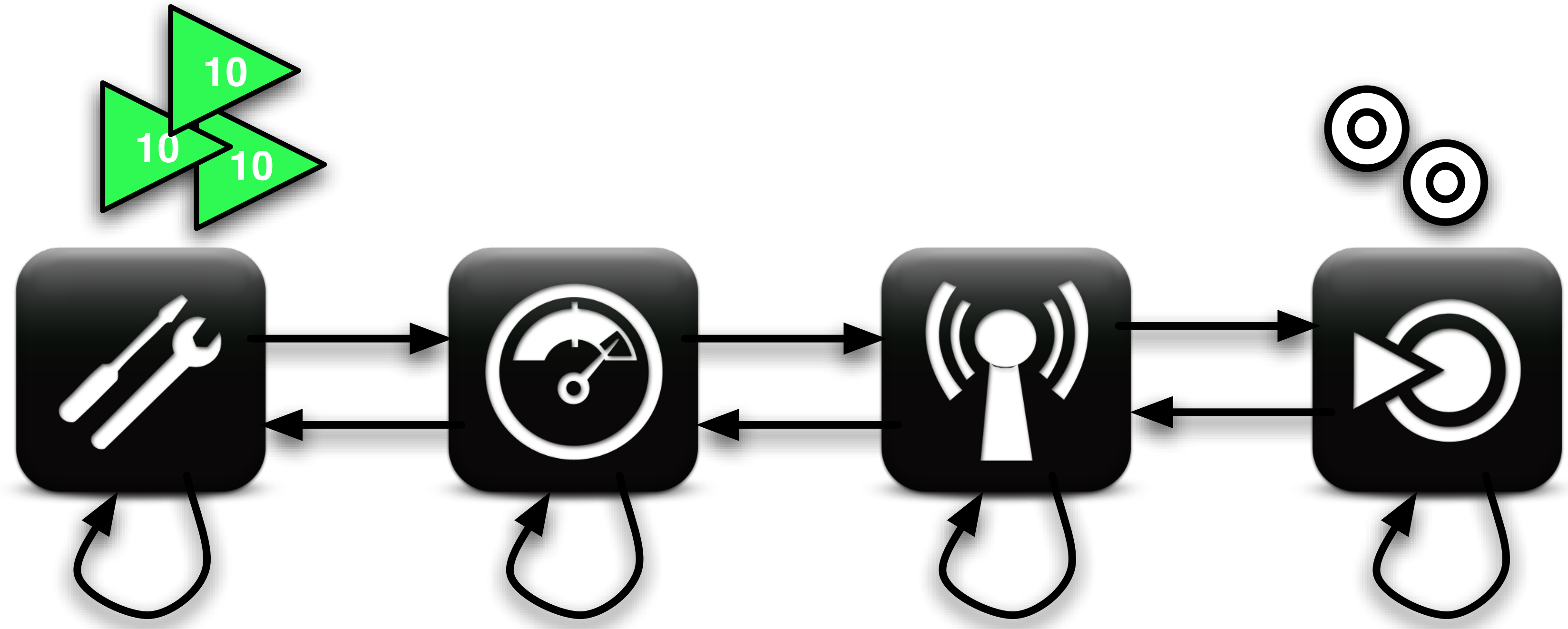


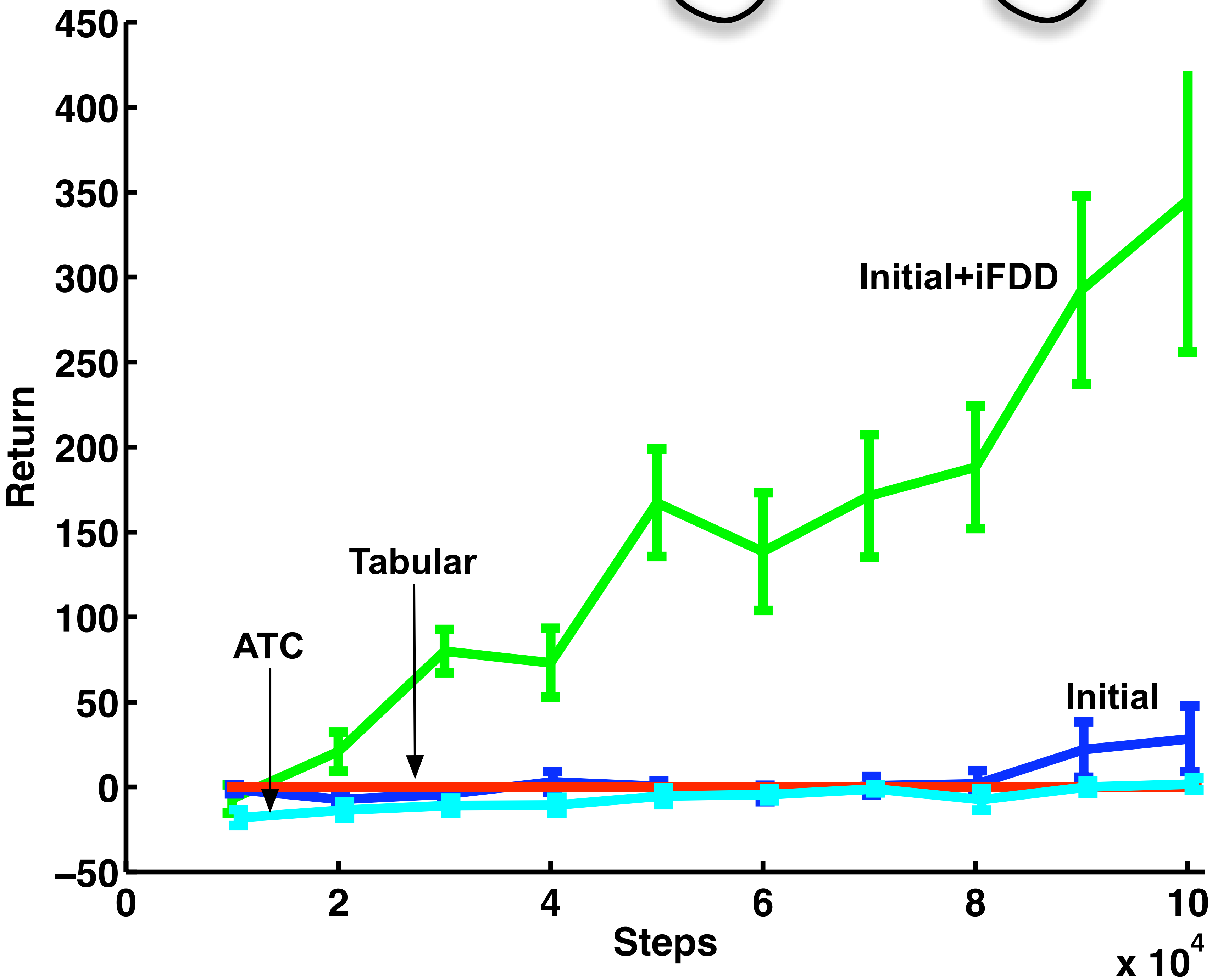
Initial+iFDD
ATC
Guassian
Tabular
Initial
Balancing Steps
3000
2500
2000
1500
1000
500
0
0
2
4
6
8
10
Steps
x 10^4

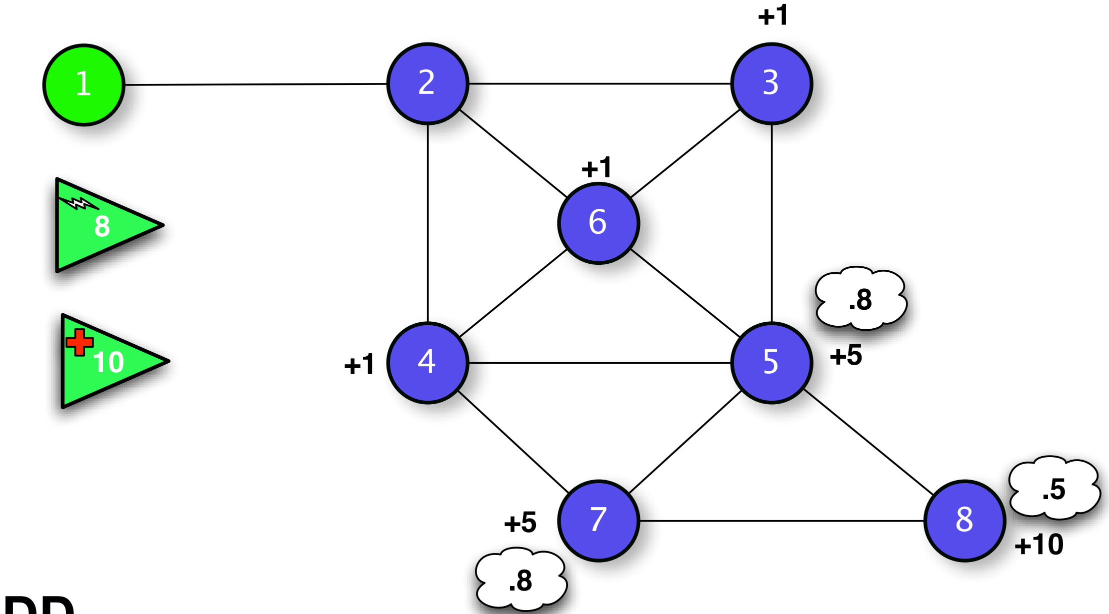




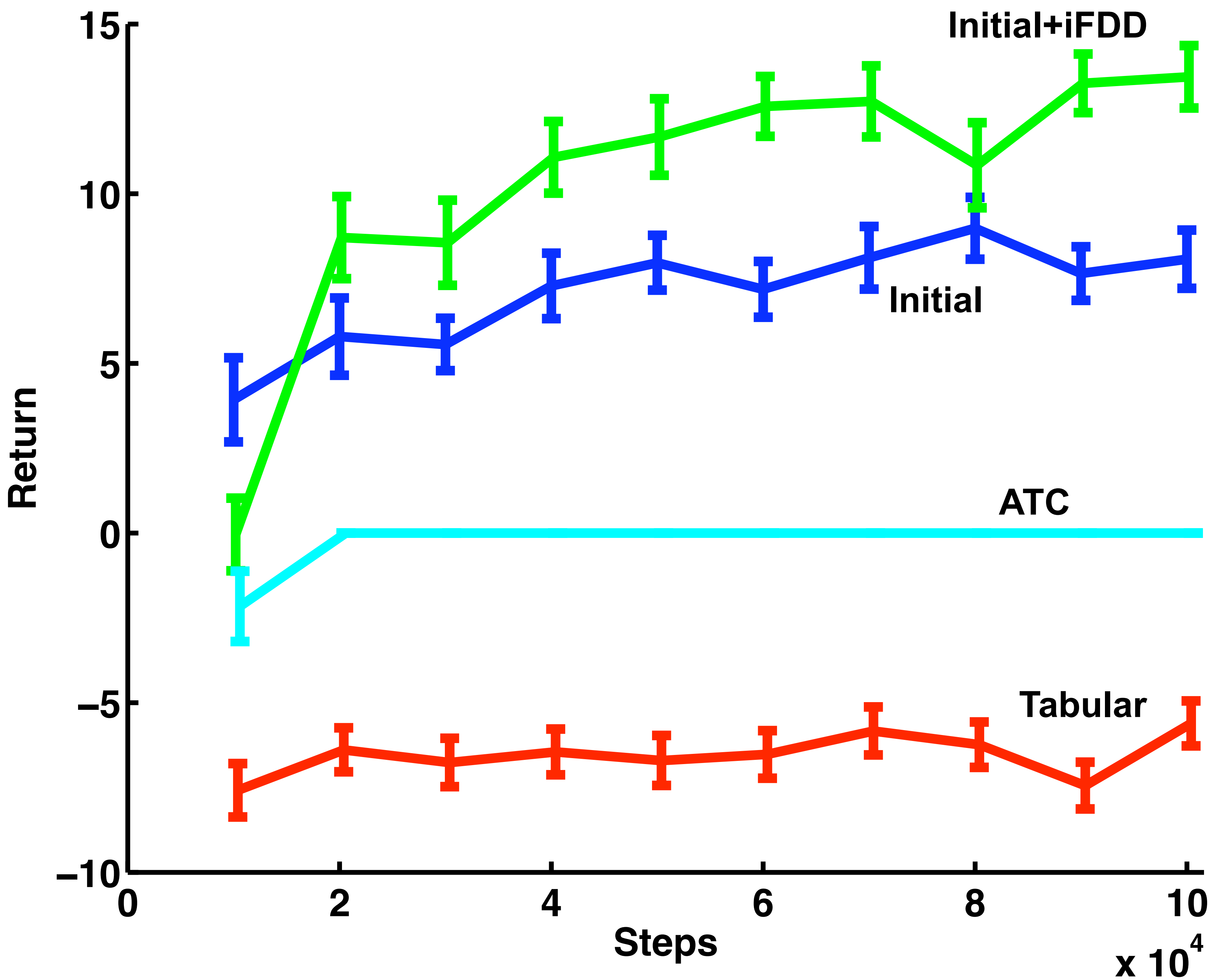
initial+iFDD
Tabular
Initial
ATC
Return
Steps
x 10^4
1
0.5
0
−0.5
−1
−1.5
0
2
4
6
8
10


10
10
10


Return
450
400
350
300
250
200
150
100
50
0
−50
0
2
4
6
8
10
Steps
x 10^4
Initial+iFDD
Tabular
ATC
Initial

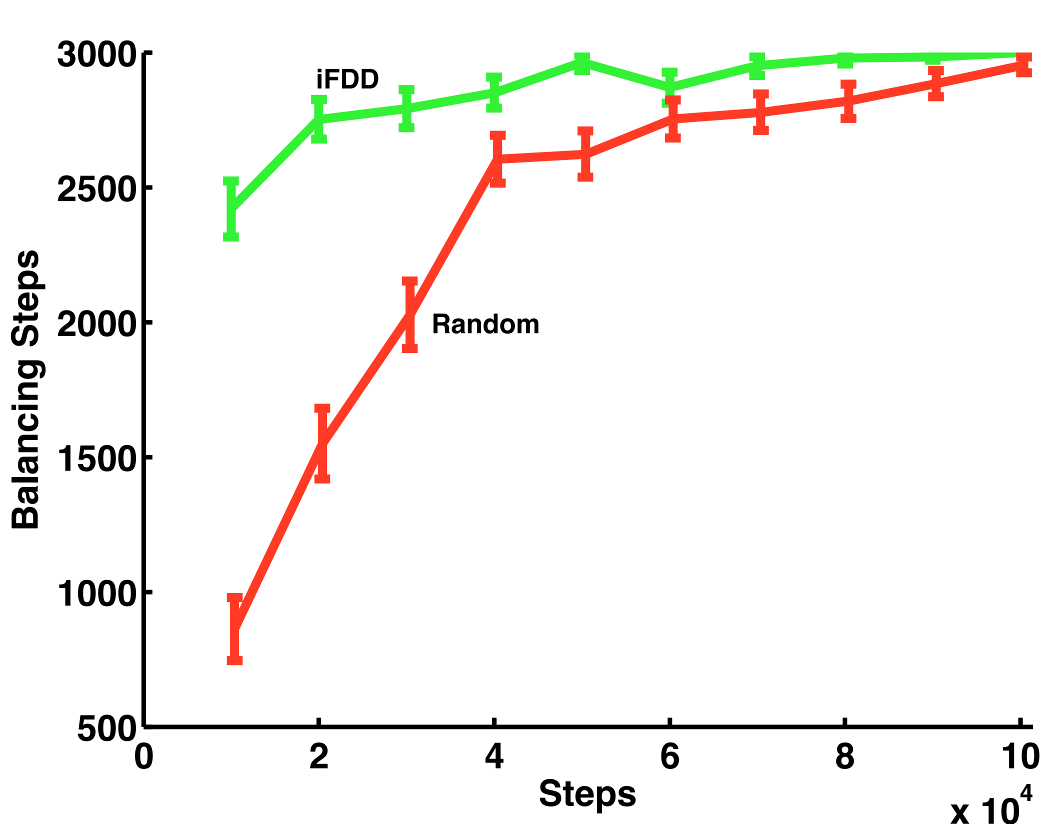


1
2
3
4
5
6
7
8
+1
+1
+1
+5
+5
+10
.8
.8
.5
8
10

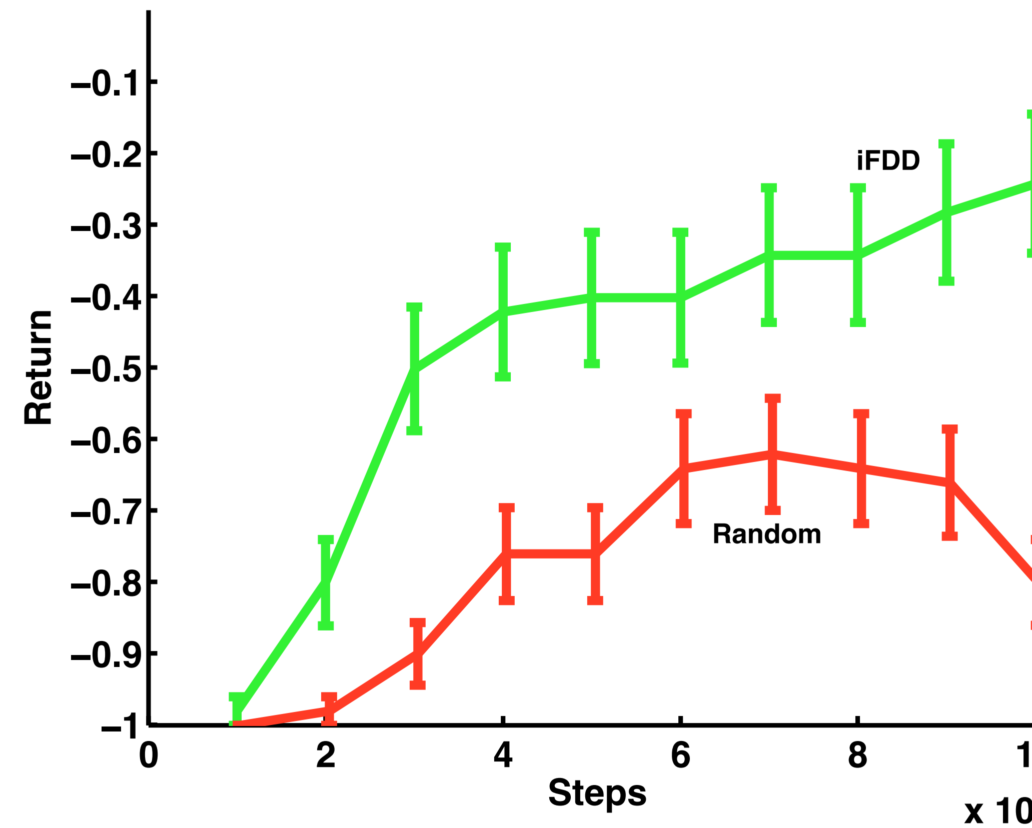


Initial+iFDD
Initial
ATC
Tabular
Return
Steps
x 10^4
15
10
5
0
−5
−10
0
2
4
6
8
10

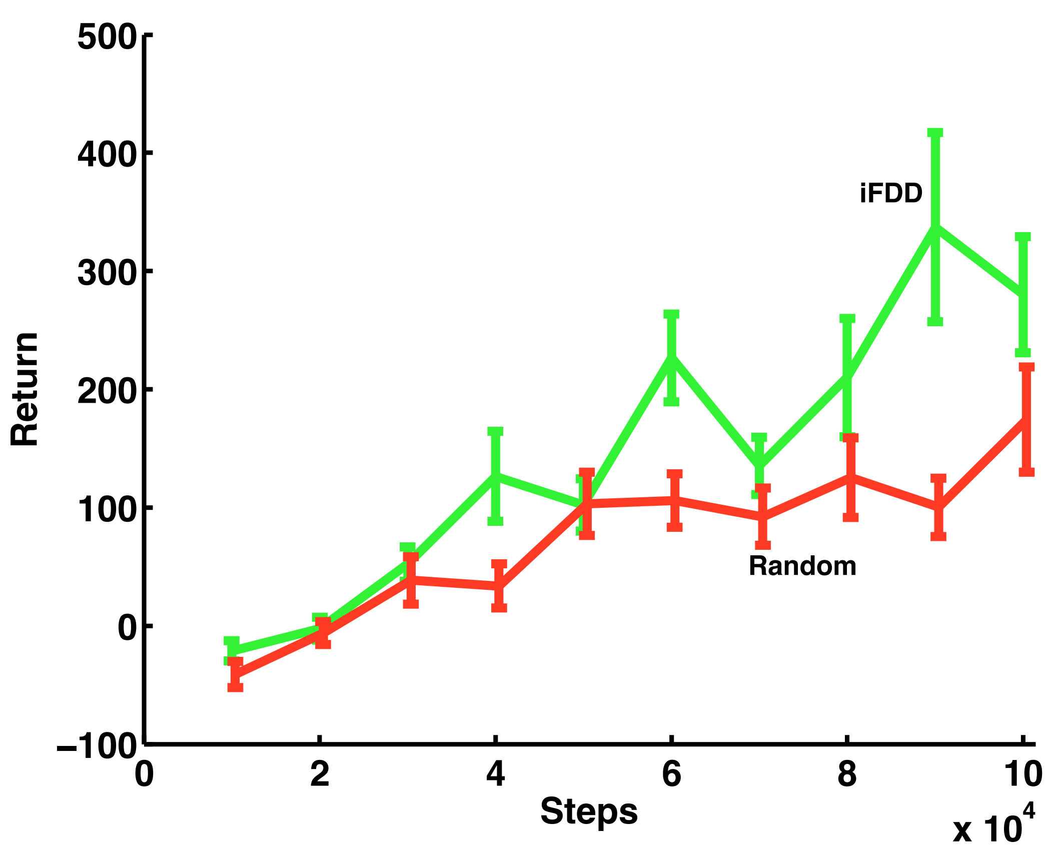
Comparison with Random Expansion



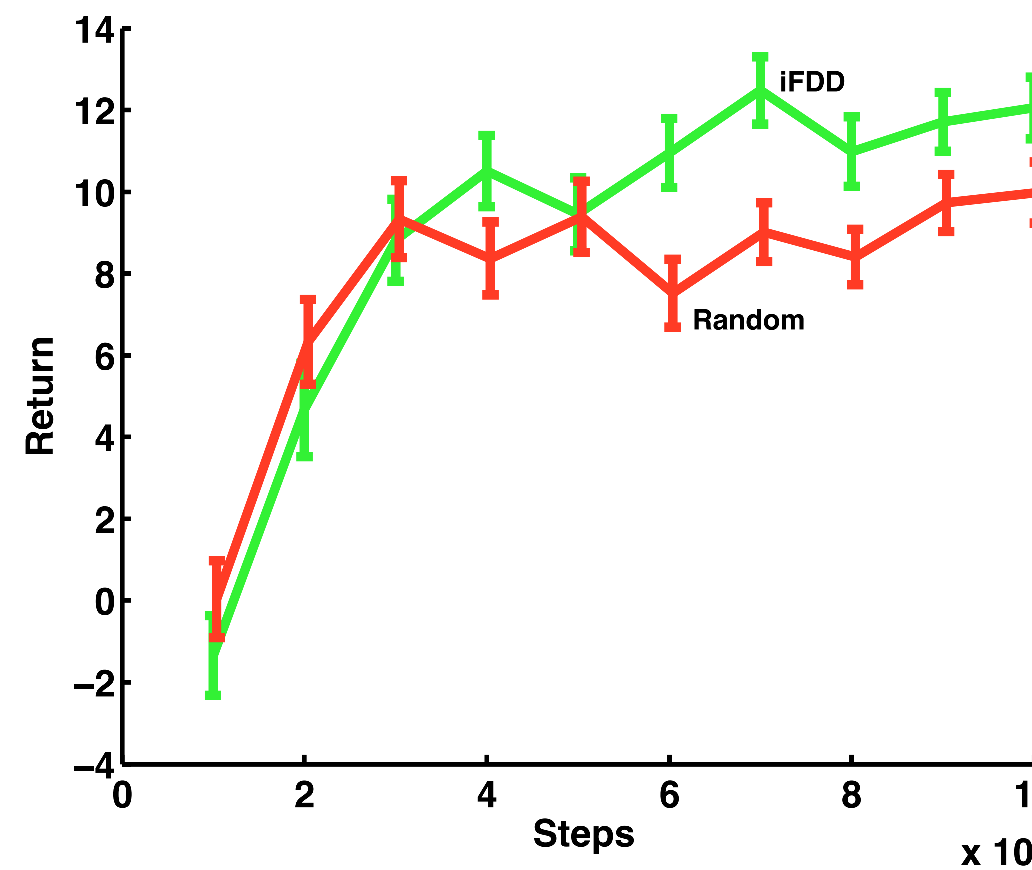
(a) Inverted Pendulum



(b) Blocksworld



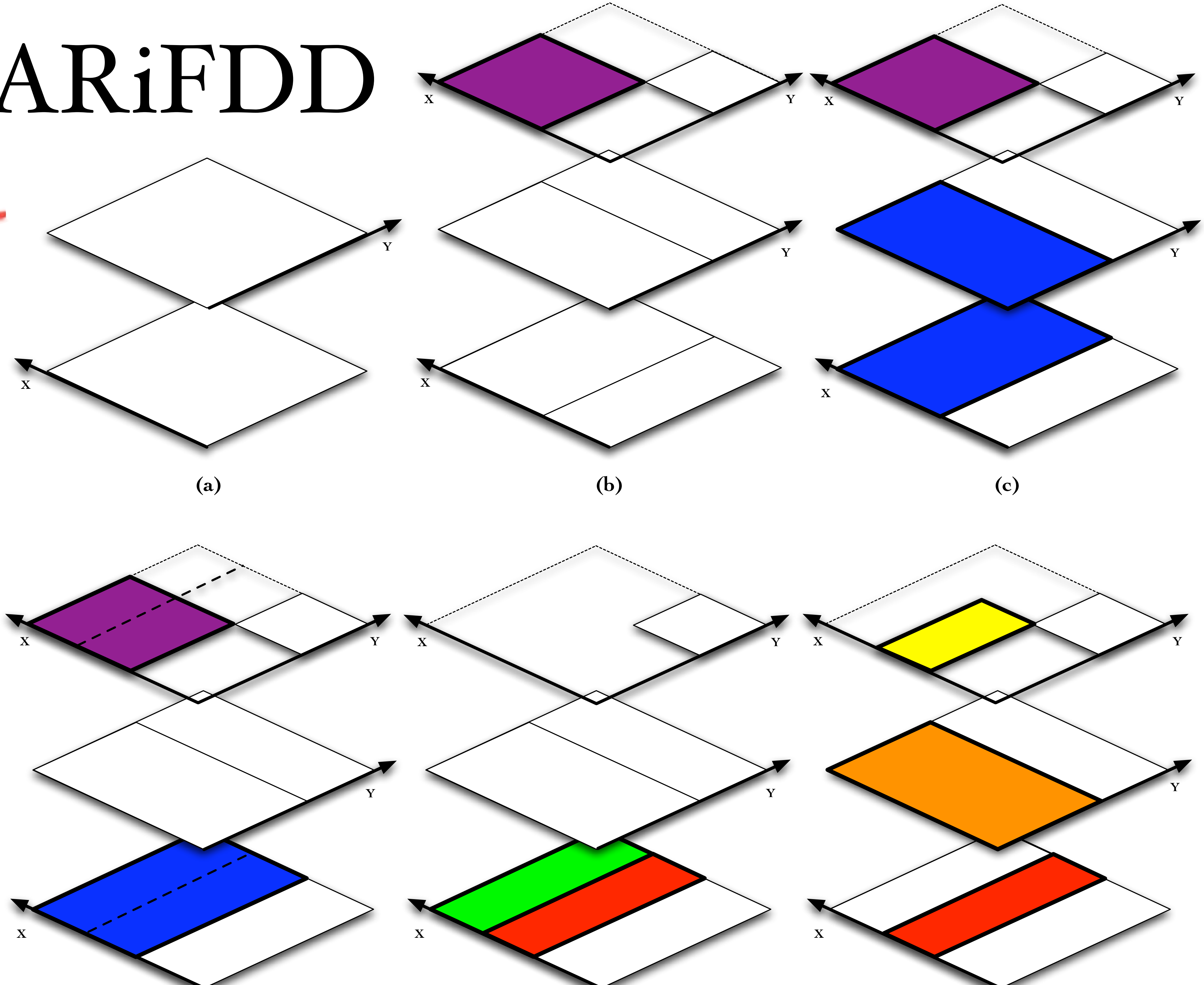
(c) Persistent Surveillance



(d) Rescue Mission

ARiFDD

ARiFDD



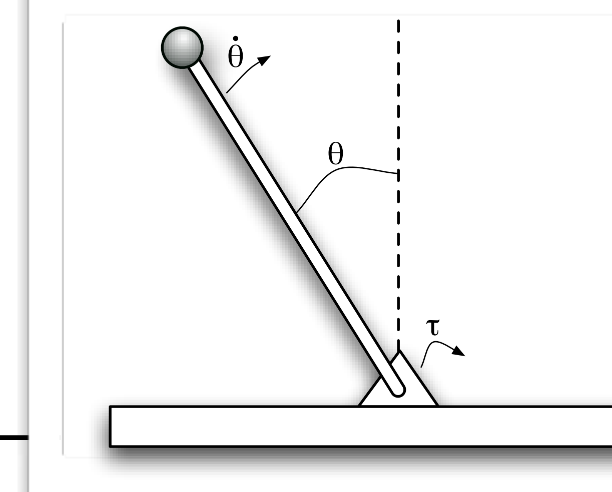
(a) (b) (c)

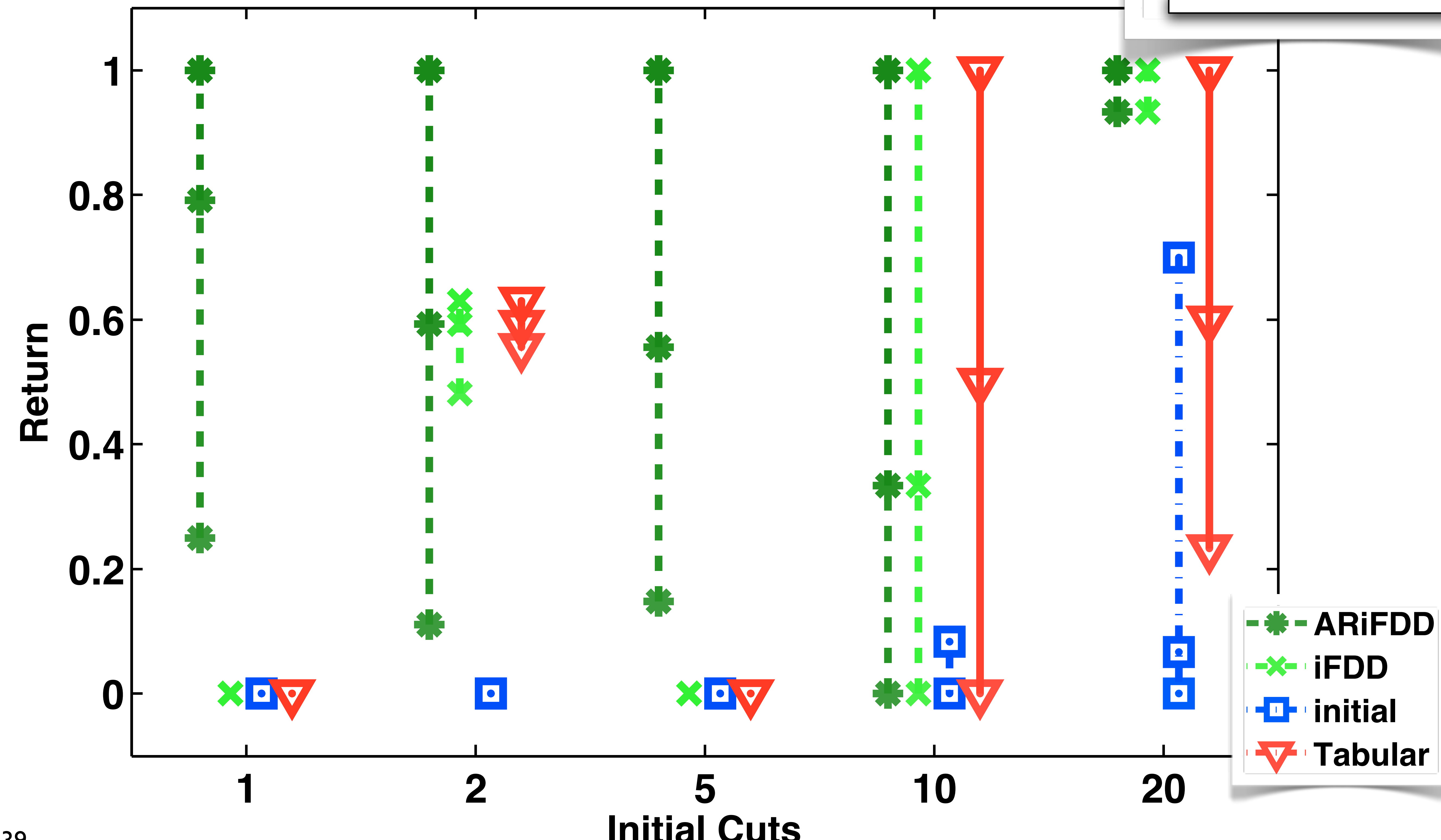
(d) (e) (f)

ARiFDD

- iFDD is ARiFDD with SplitThreshold of ∞.
- For each basic tile, weighted **μ** and **σ** are stored **incrementally**.
- Empirical results suggest cutting through the dimension with the **least variance** works best.

ARiFDD


θ̇
θ
τ


Return
1
0.8
0.6
0.4
0.2
0
Initial Cuts
1
2
5
10
20
ARiFDD
iFDD
initial
Tabular

Theory

Theorems

$$\tilde{\mathrm{V}} = \Phi\theta$$

- Consider a 4 state MDP with 2 binary features

$$\begin{array}{c} \mathrm{s1} \\ \mathrm{s2} \\ \mathrm{s3} \\ \mathrm{s4} \end{array} \begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

Theorems

$$\tilde{V} = \Phi\theta$$

- Consider a 4 state MDP with 2 binary features

$$\begin{array}{c} s1 \\ s2 \\ s3 \\ s4 \end{array} \begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

Rate of Convergence

$$||\mathbf{V}^* - \tilde{\mathbf{V}}|| = x > 0$$

$$\forall \boldsymbol{\phi}_f \in \mathbb{R}^n : \beta = \angle(\boldsymbol{\phi}_f, \boldsymbol{\delta}) < \cos^{-1}(\gamma)$$

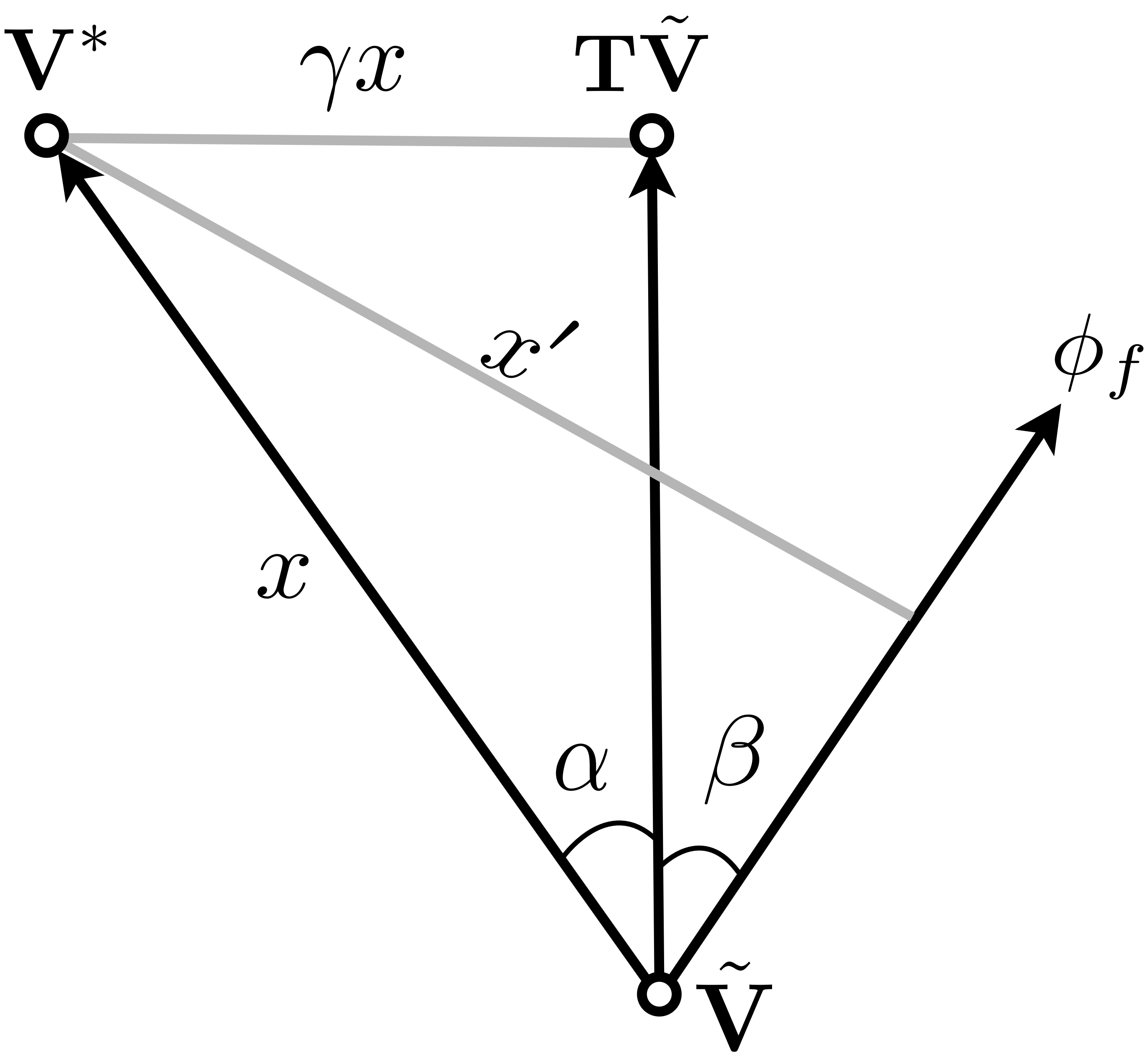
$$\exists \xi \in \mathbb{R} : ||\mathbf{V}^* - \tilde{\mathbf{V}}|| - ||\mathbf{V}^* - (\tilde{\mathbf{V}} + \xi\boldsymbol{\phi}_f)|| \geq \zeta x,$$

$$||\mathbf{V}^* - \Pi\mathbf{V}^*|| - ||\mathbf{V}^* - \Pi'\mathbf{V}^*|| \geq \zeta x$$

$$\zeta = 1 - \gamma\cos(\beta) - \sqrt{1-\gamma^2}\sin(\beta) < 1.$$

[Parr et al. 2007]

Proof Sketch



[Parr et al. 2007]

Selection Mechanism

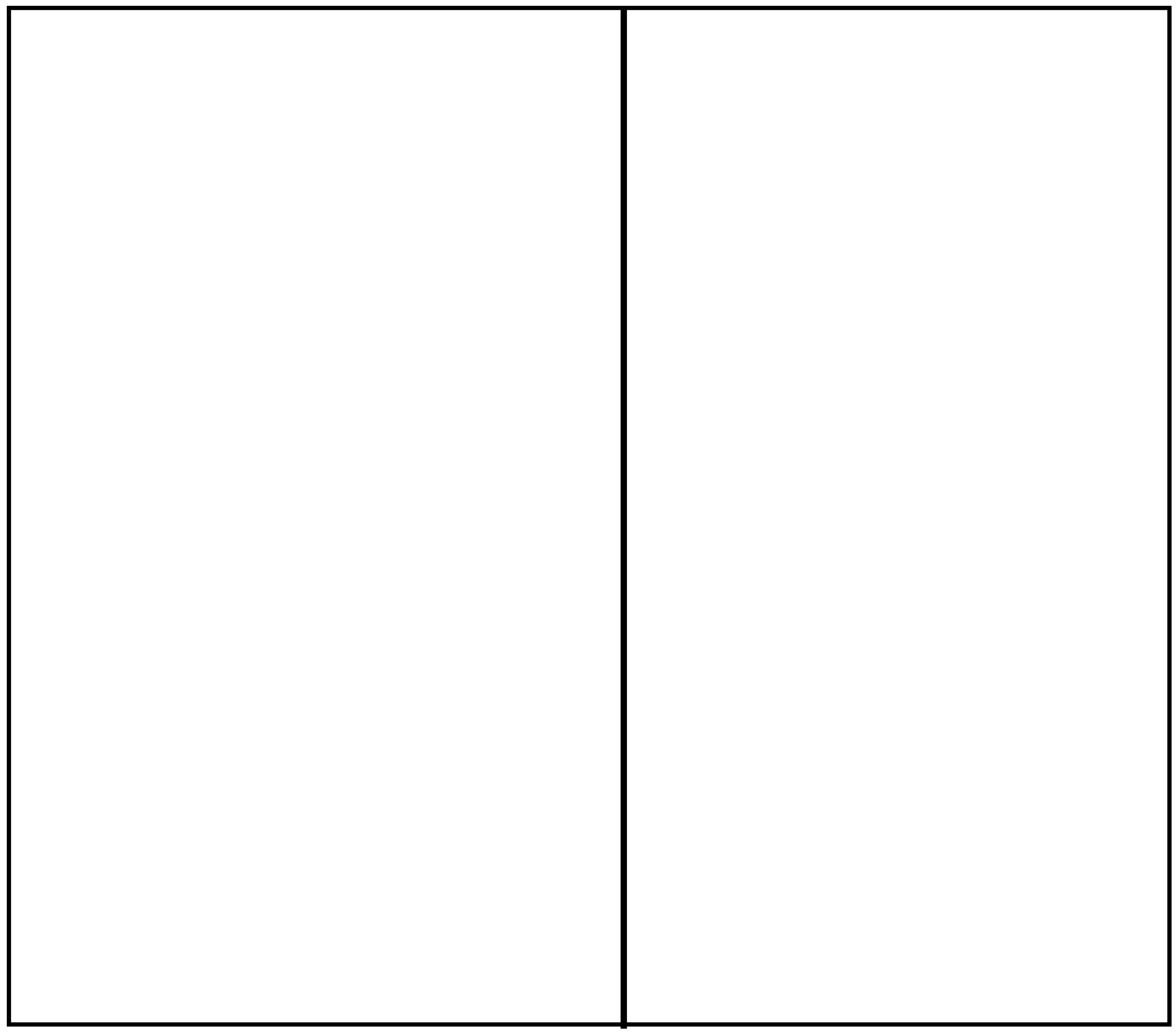
$$f^* = \mathrm{argmax}_{f \in pair(F)} \frac{\sum_{s \in Samples, \phi_f(s)=1} \delta(s)}{\sqrt{\sum_{s \in Samples} \phi_f(s) = 1}}$$

ATC and SDM

Adaptive Tile Coding

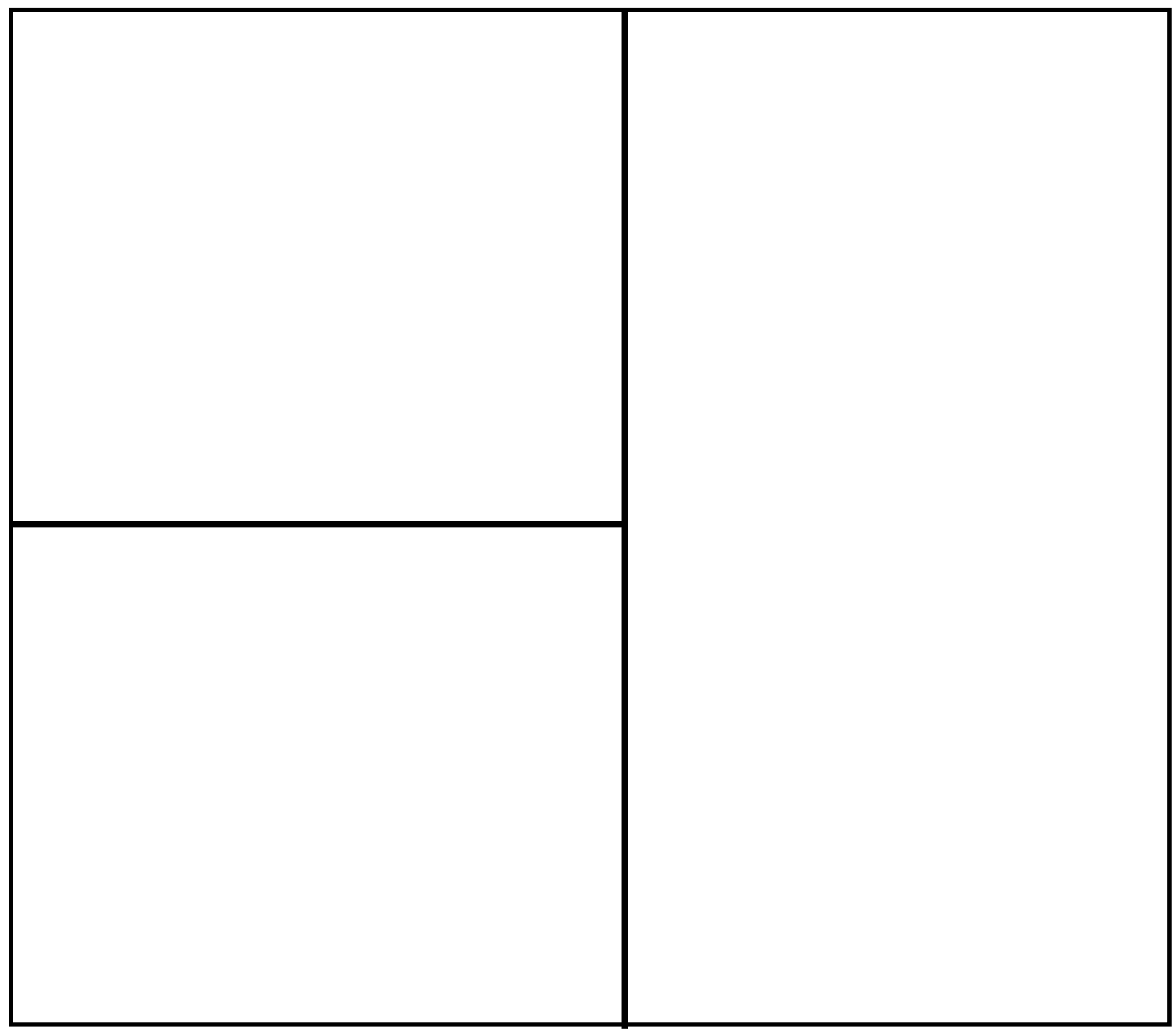
[Whiteson et al. 2007]

Adaptive Tile Coding



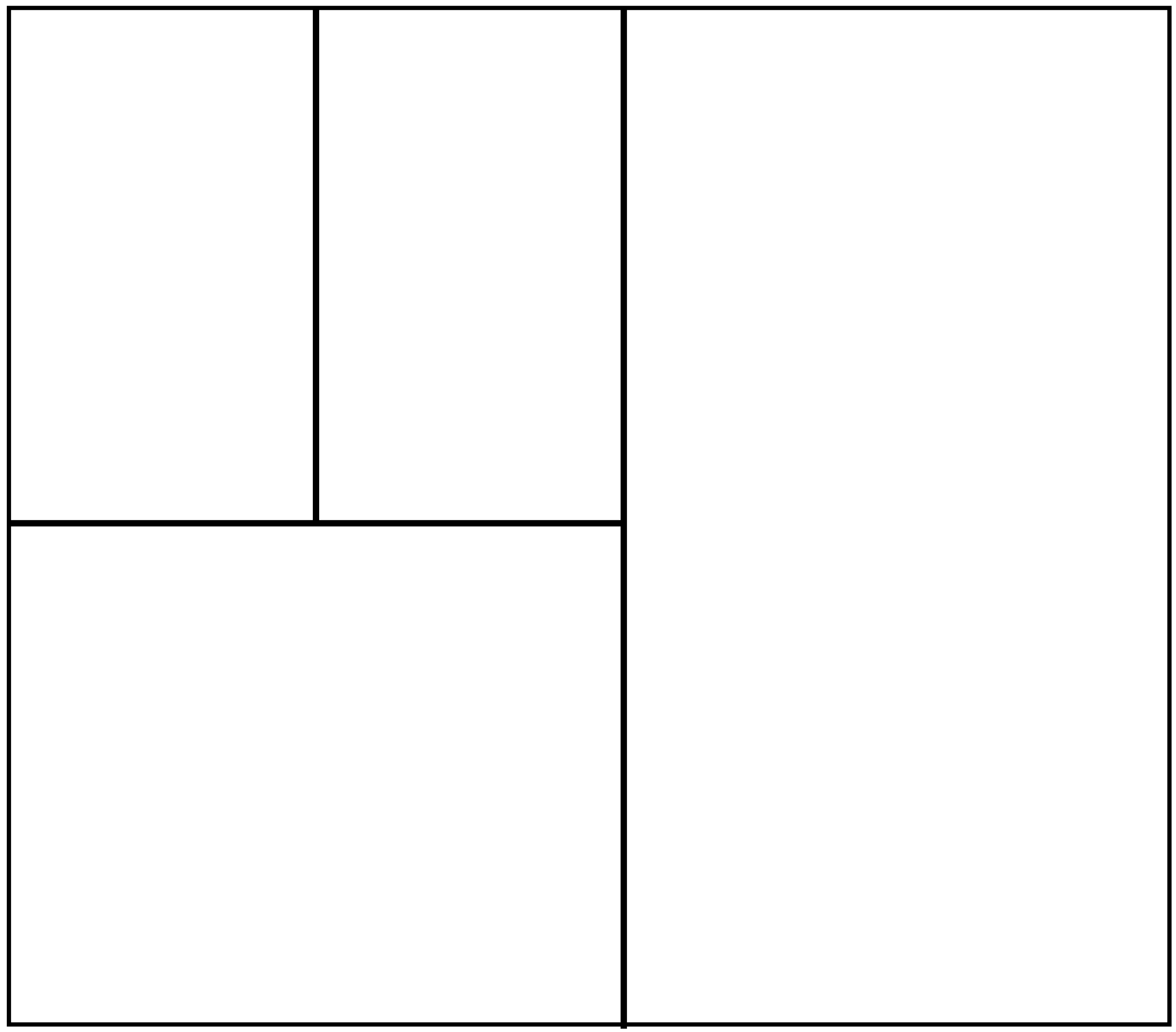
[Whiteson et al. 2007]

Adaptive Tile Coding



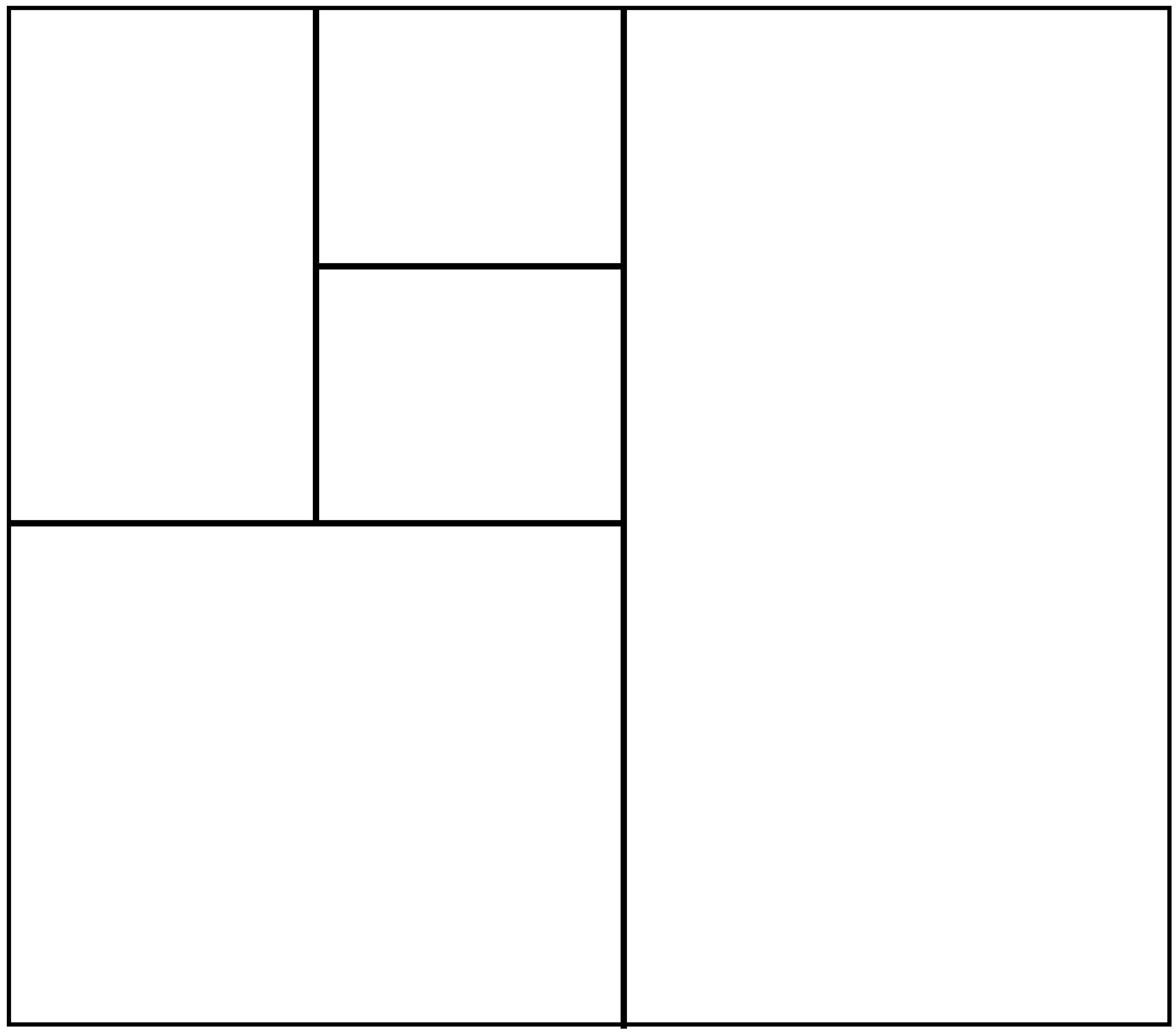
[Whiteson et al. 2007]

Adaptive Tile Coding



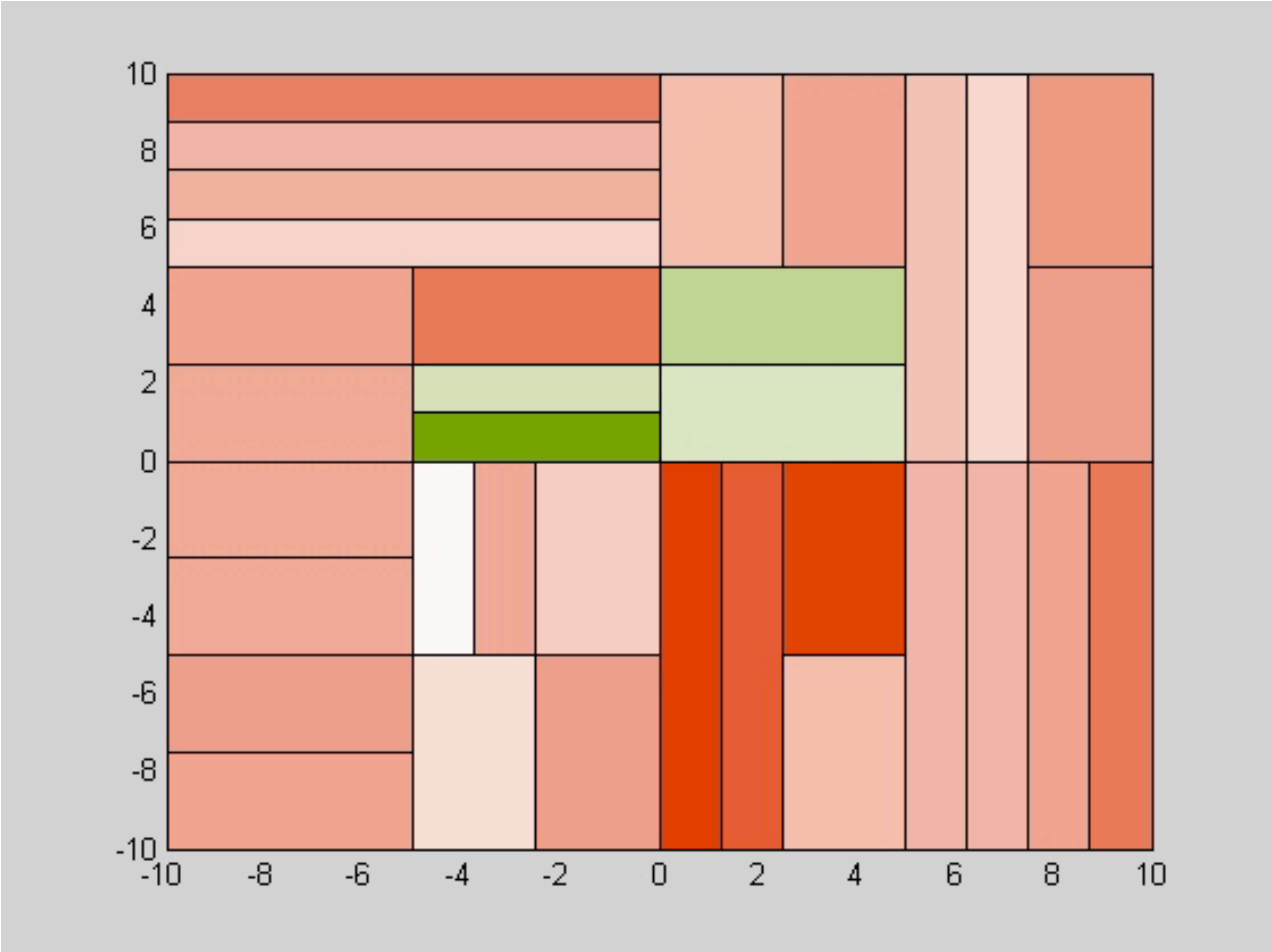
[Whiteson et al. 2007]

Adaptive Tile Coding



[Whiteson et al. 2007]

Adaptive Tile Coding

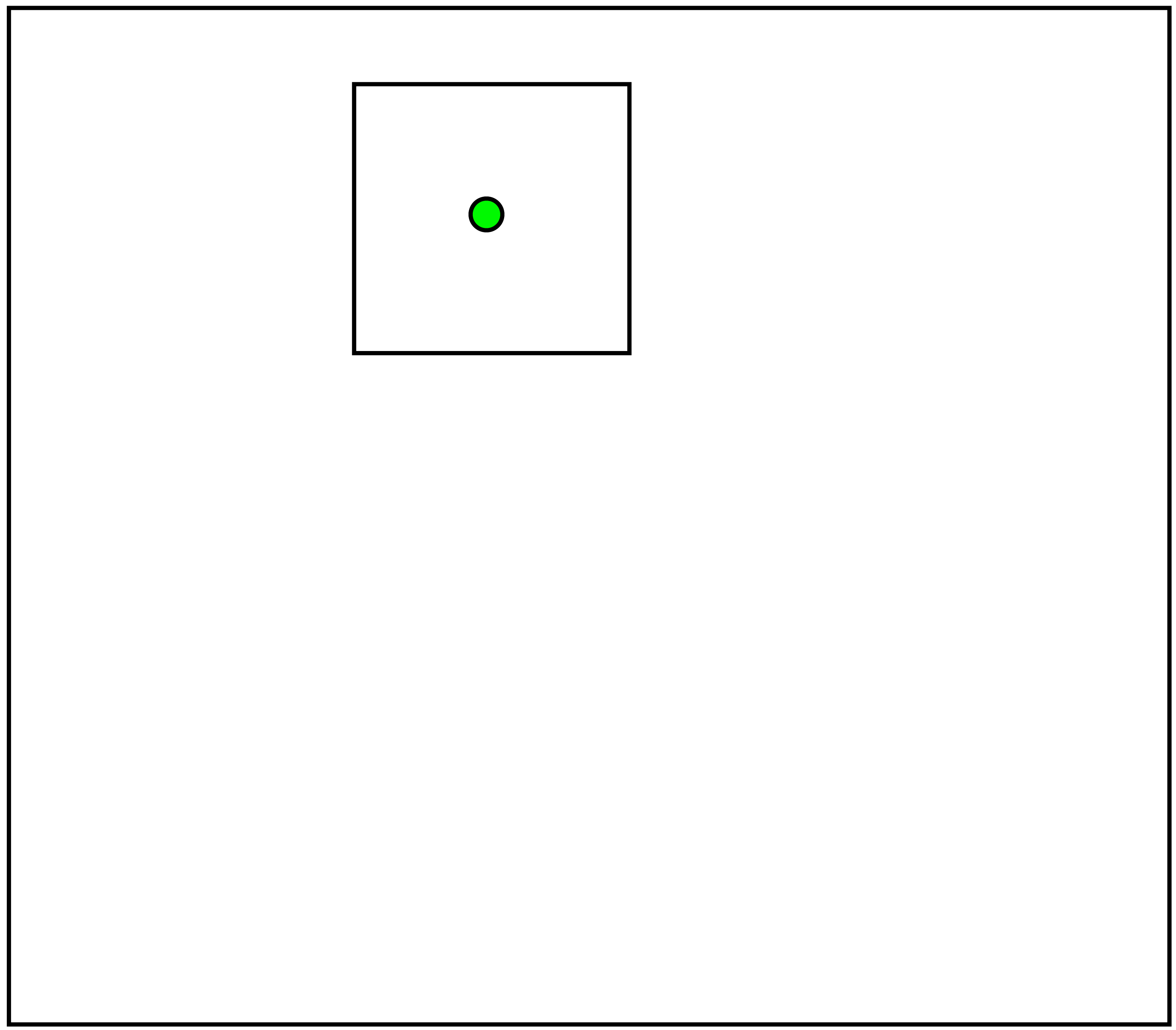


[Whiteson et al. 2007]

Sparse Distributed Memories

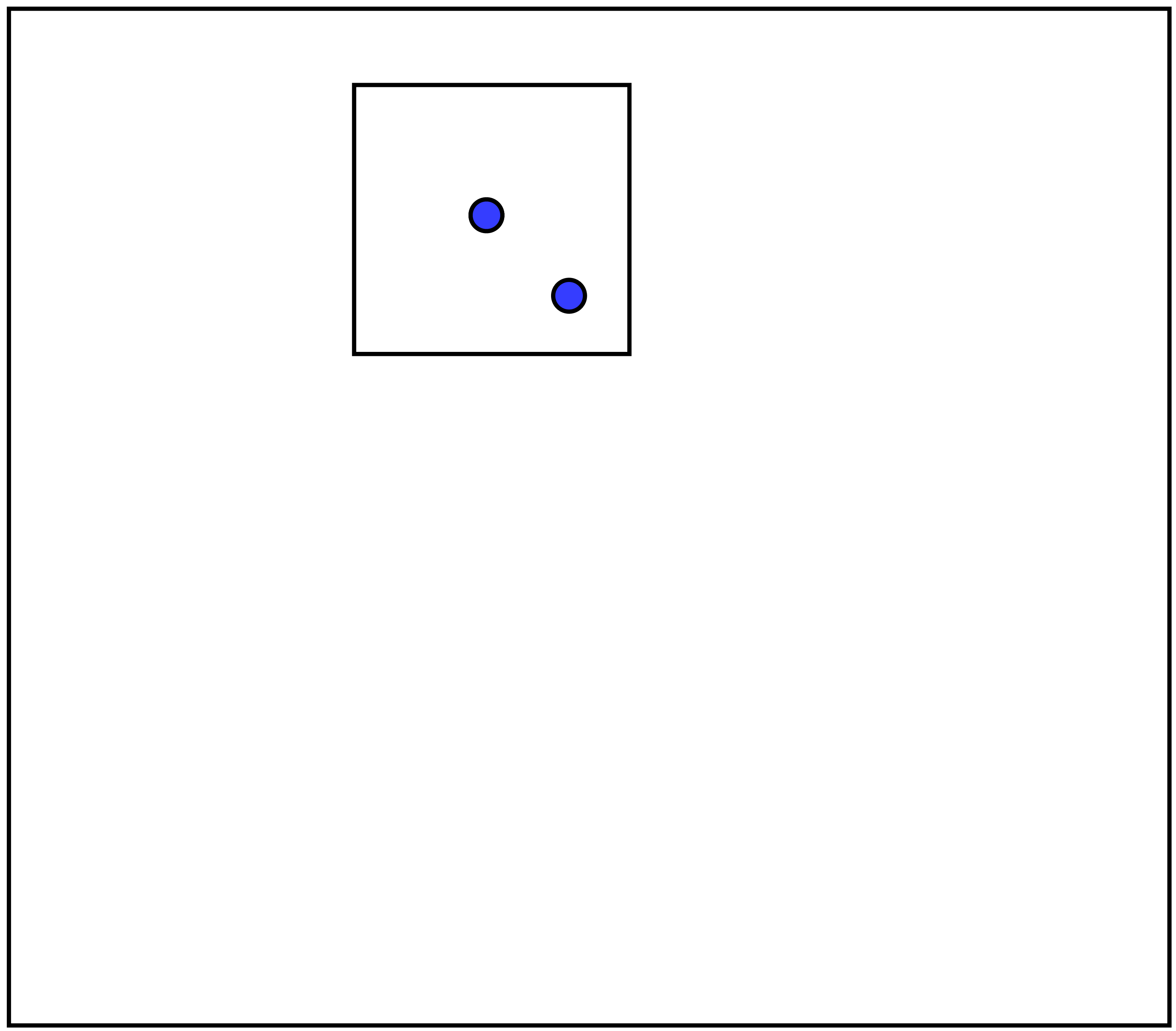
[Ratitch et al. 2004]

Sparse Distributed Memories



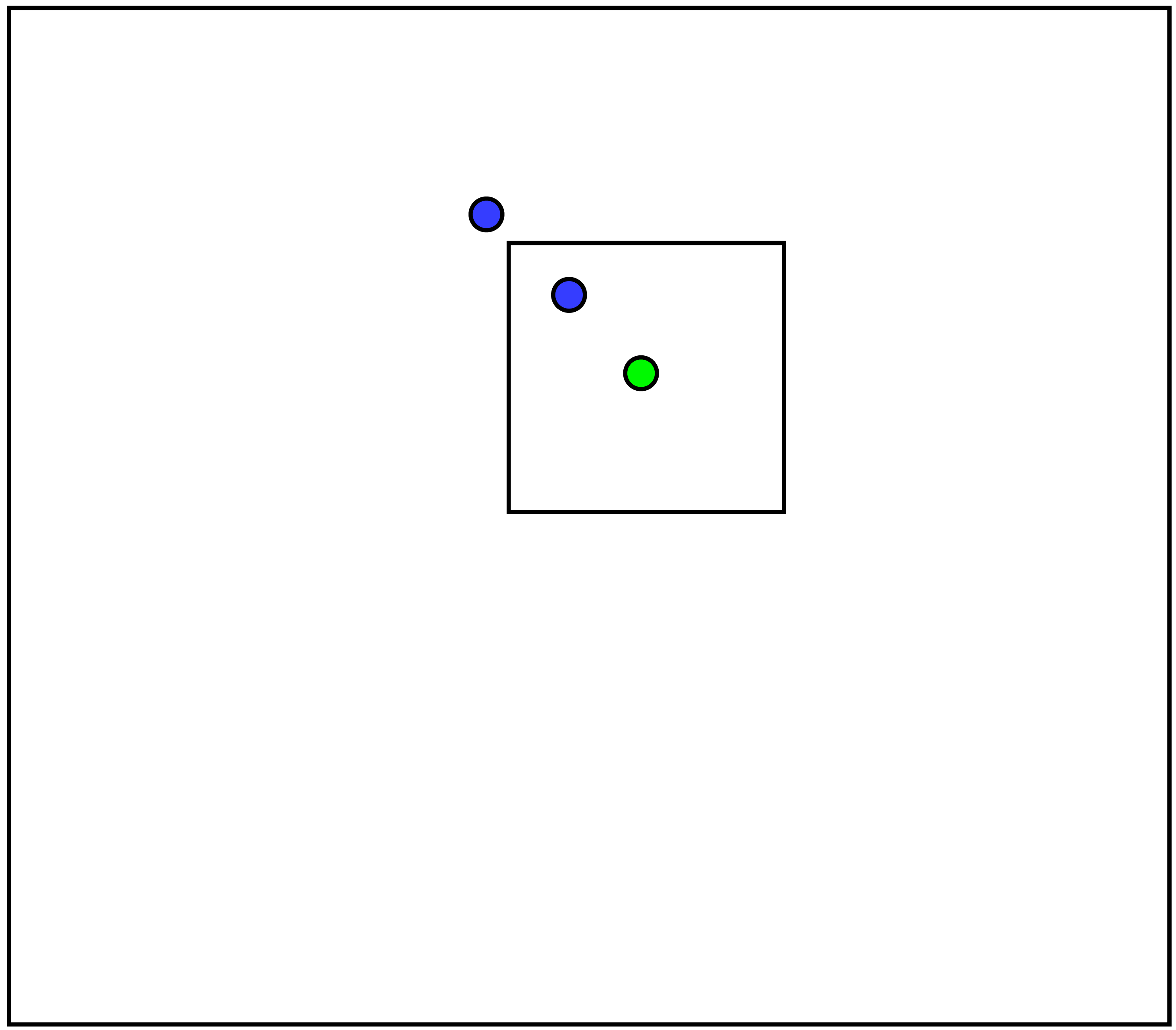
[Ratitch et al. 2004]

Sparse Distributed Memories



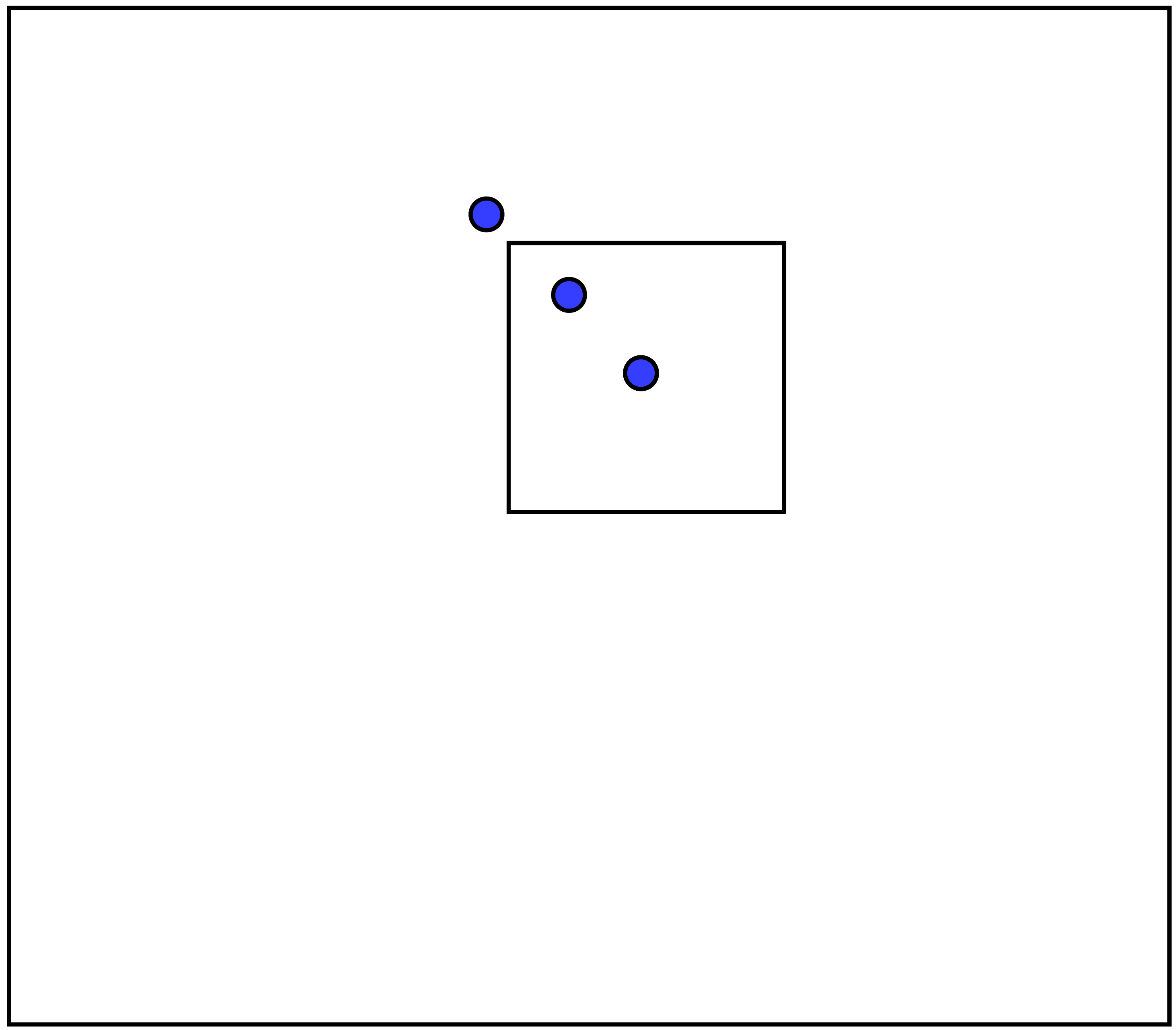
[Ratitch et al. 2004]

Sparse Distributed Memories



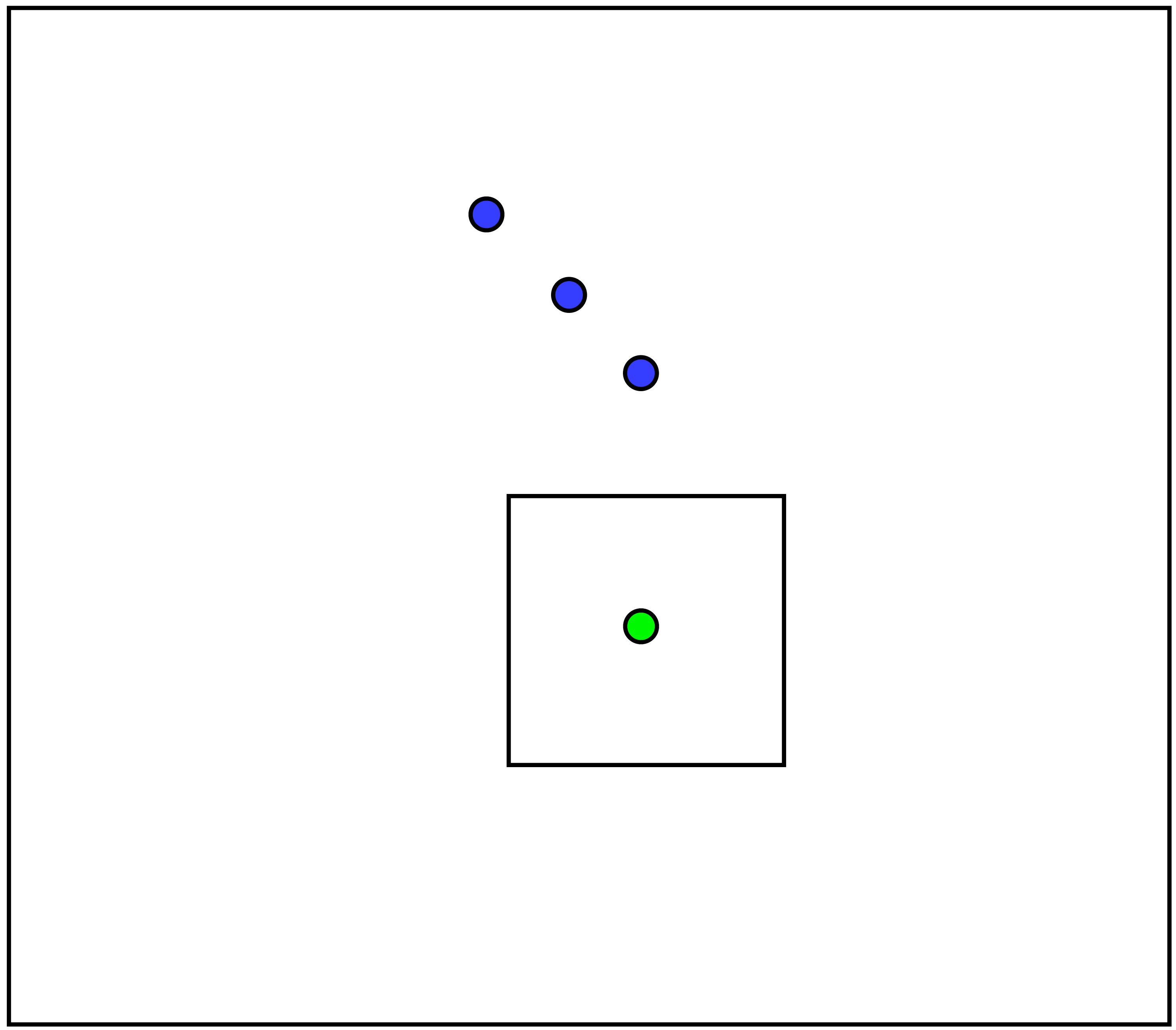
[Ratitch et al. 2004]

Sparse Distributed Memories



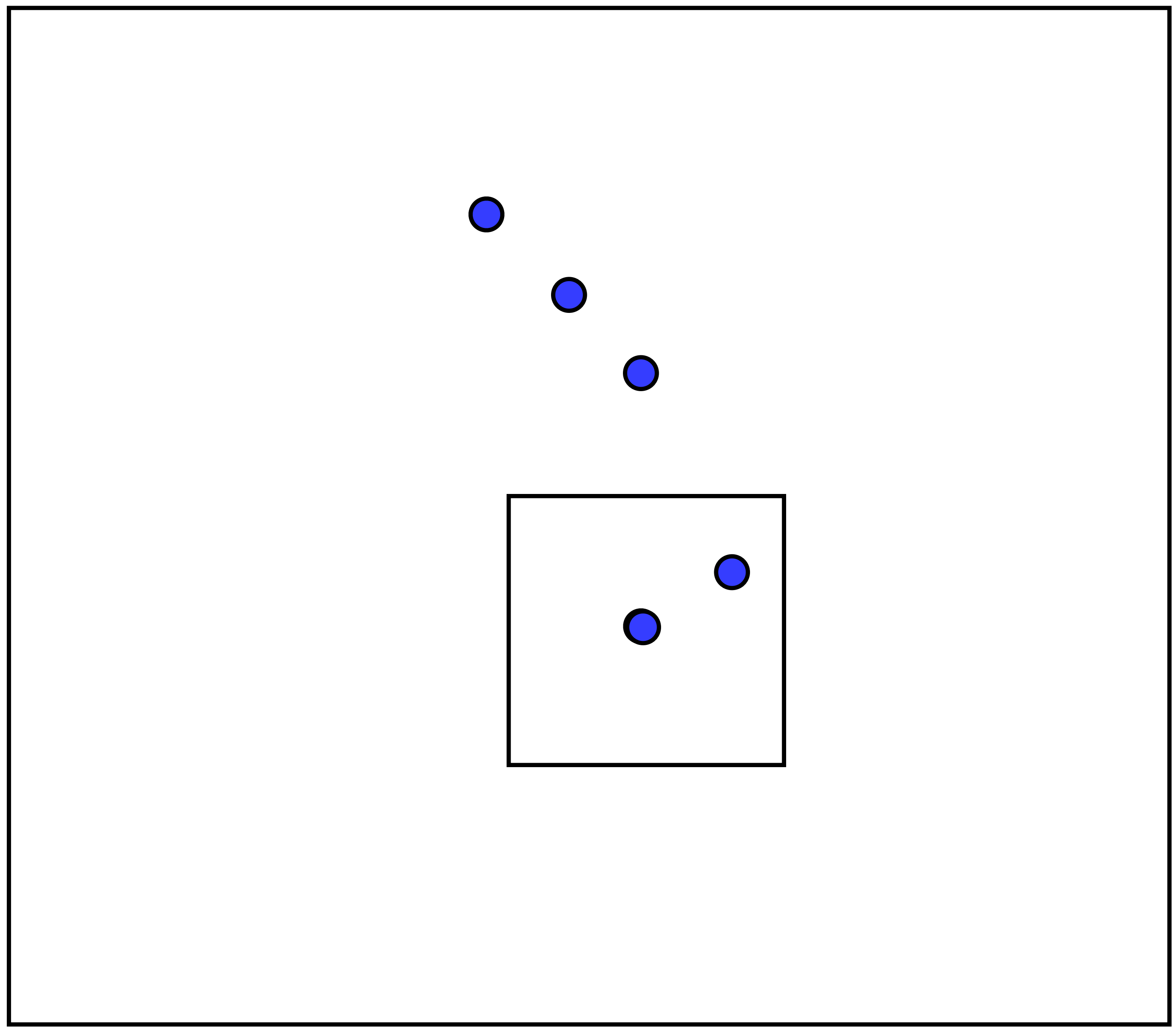
[Ratitch et al. 2004]

Sparse Distributed Memories



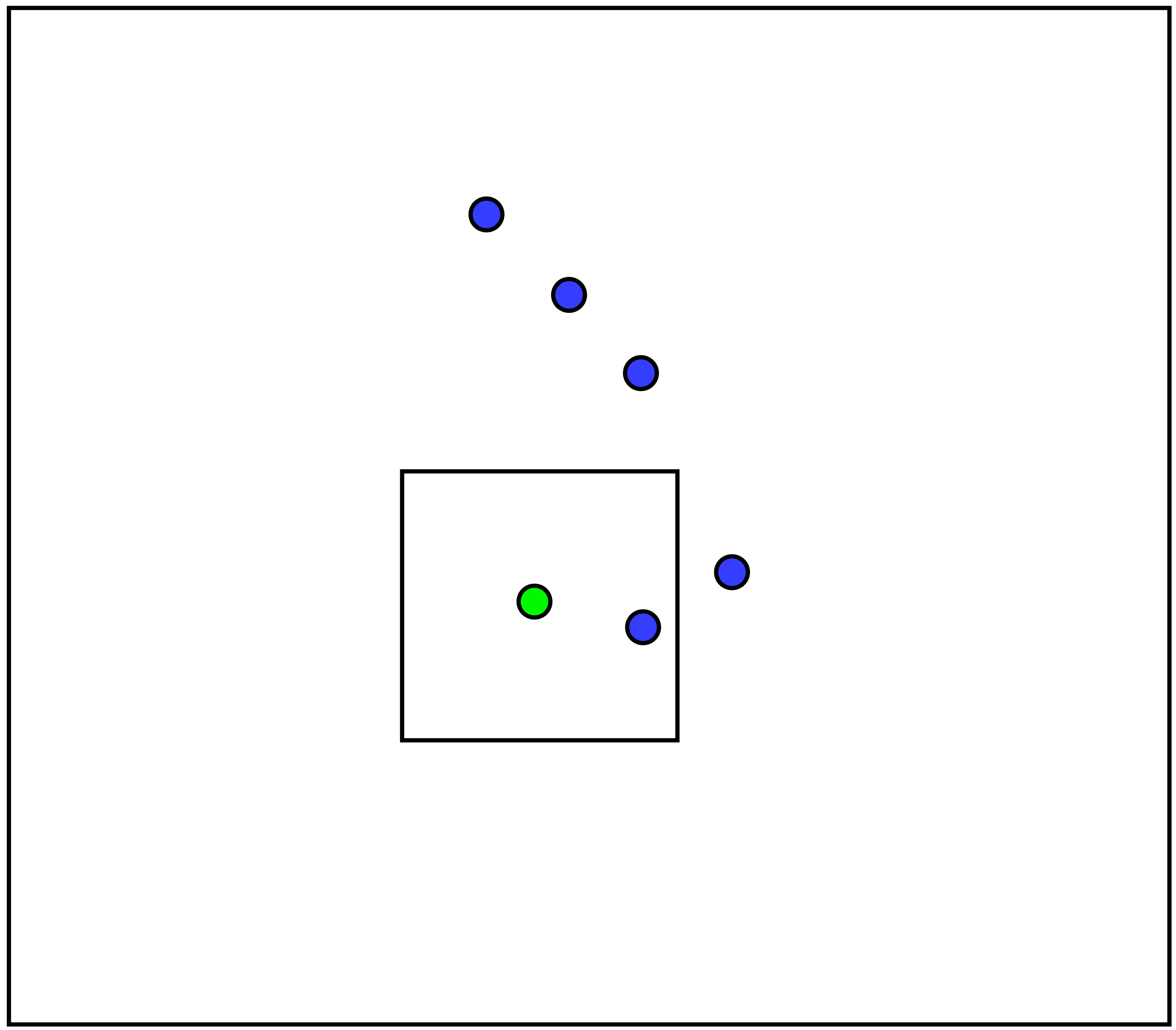
[Ratitch et al. 2004]

Sparse Distributed Memories



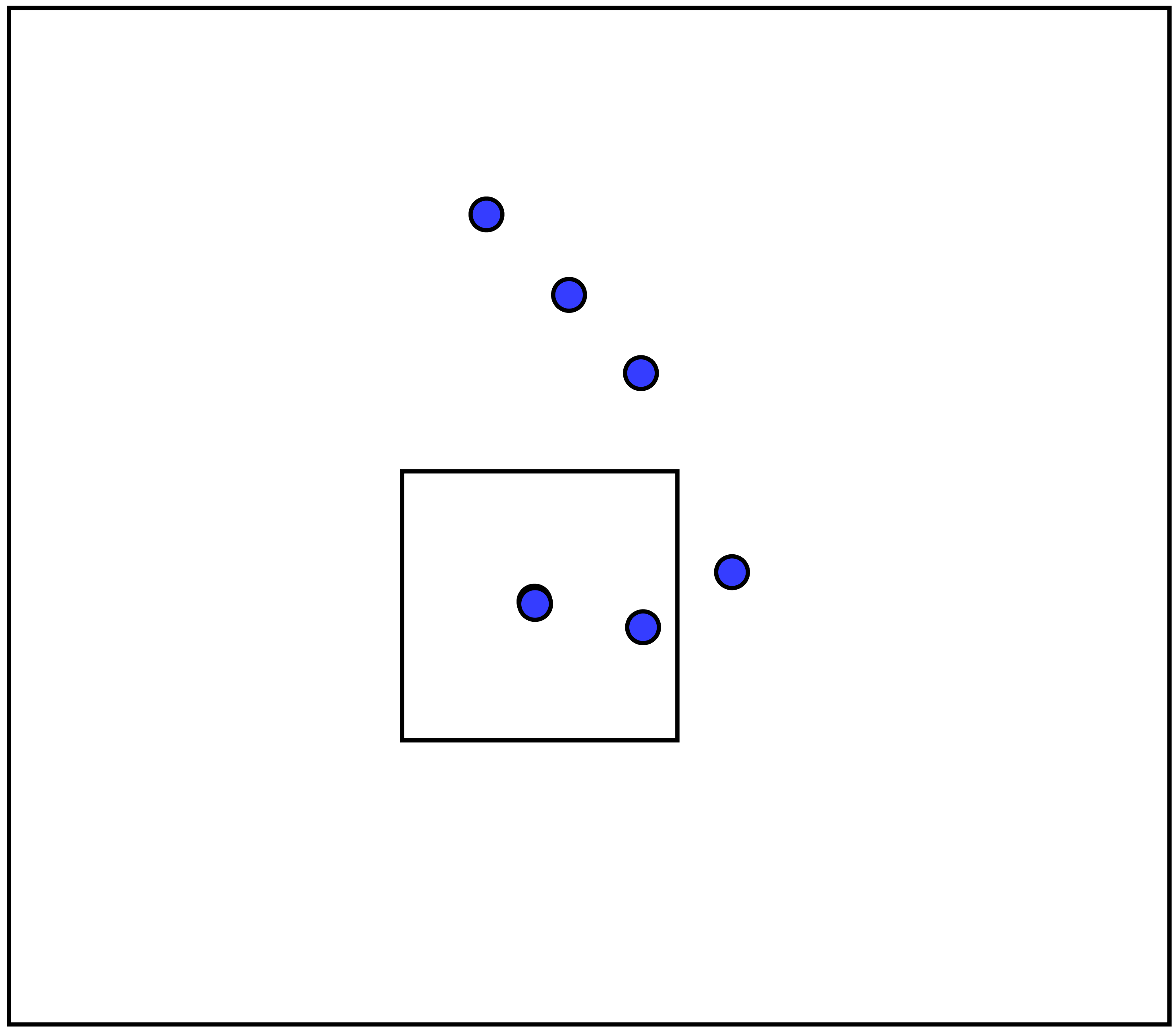
[Ratitch et al. 2004]

Sparse Distributed Memories



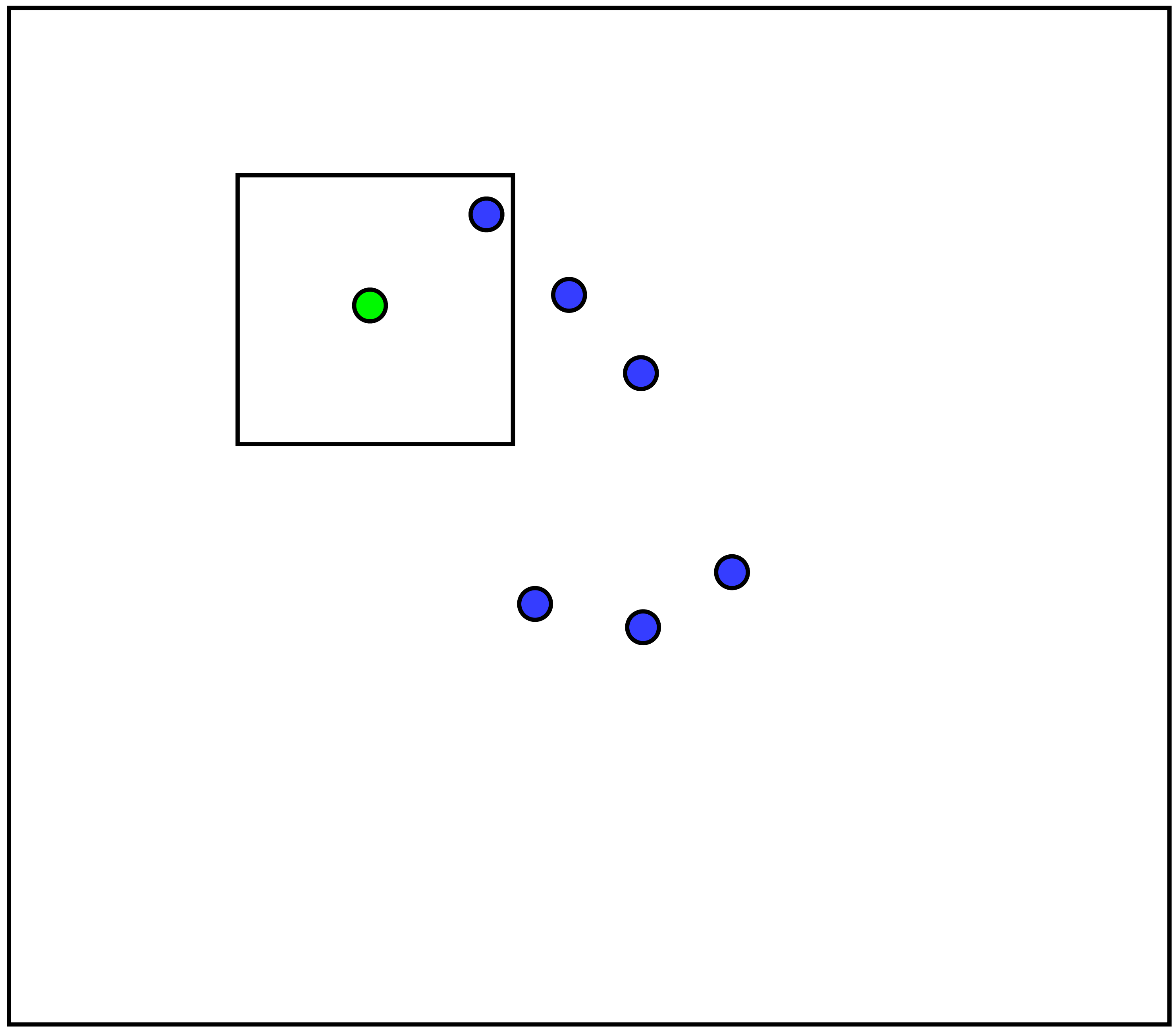
[Ratitch et al. 2004]

Sparse Distributed Memories



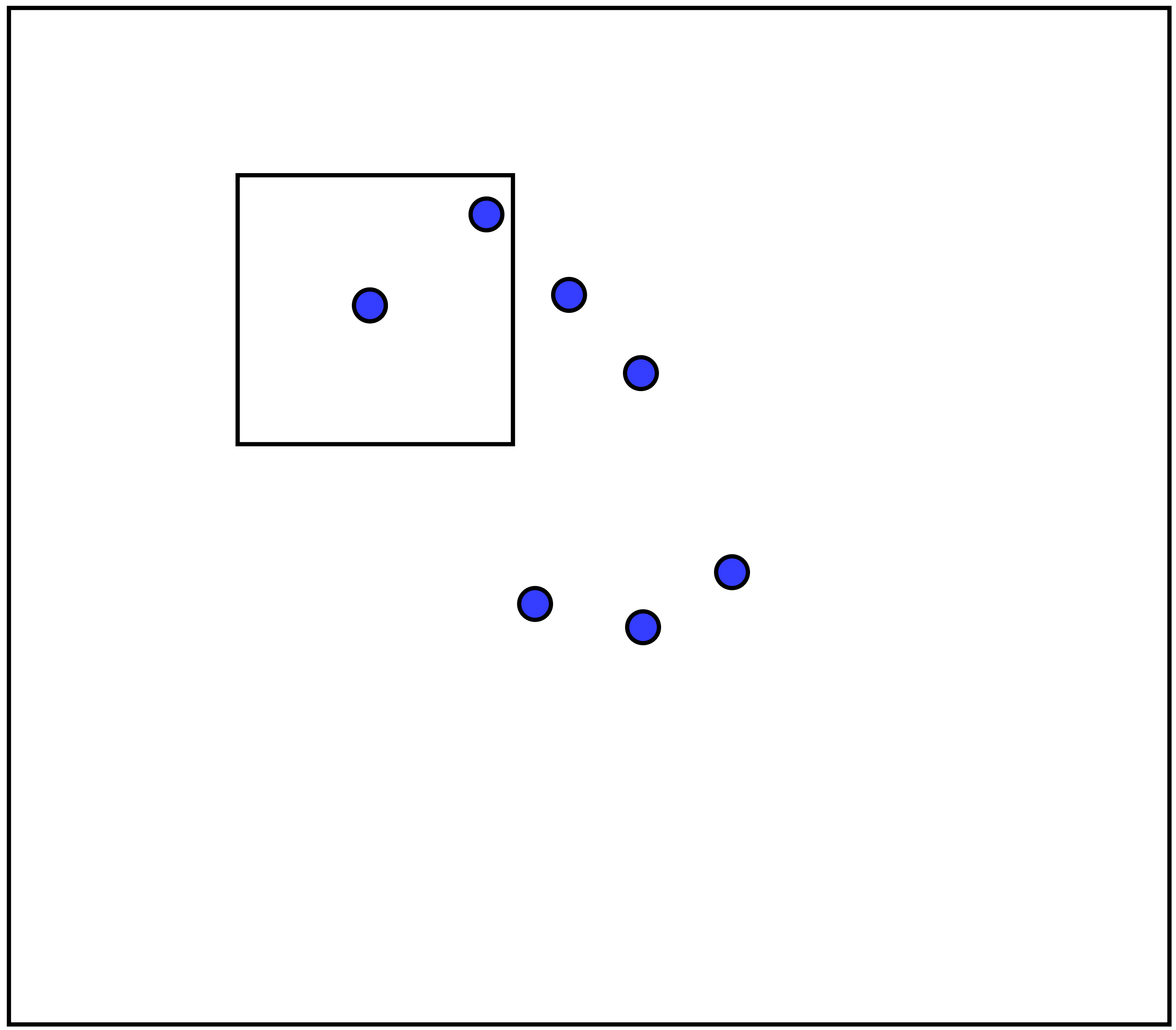
[Ratitch et al. 2004]

Sparse Distributed Memories



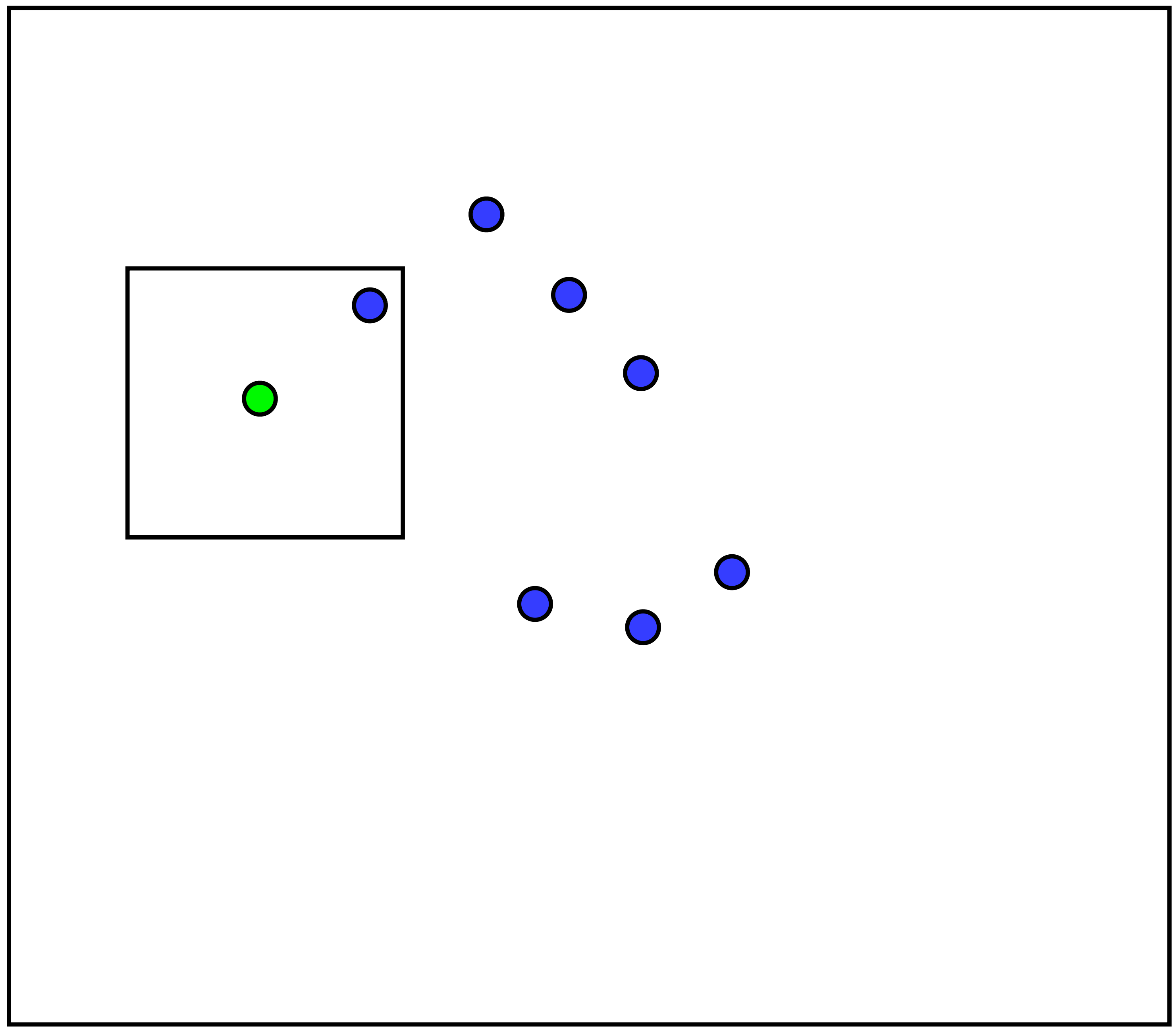
[Ratitch et al. 2004]

Sparse Distributed Memories



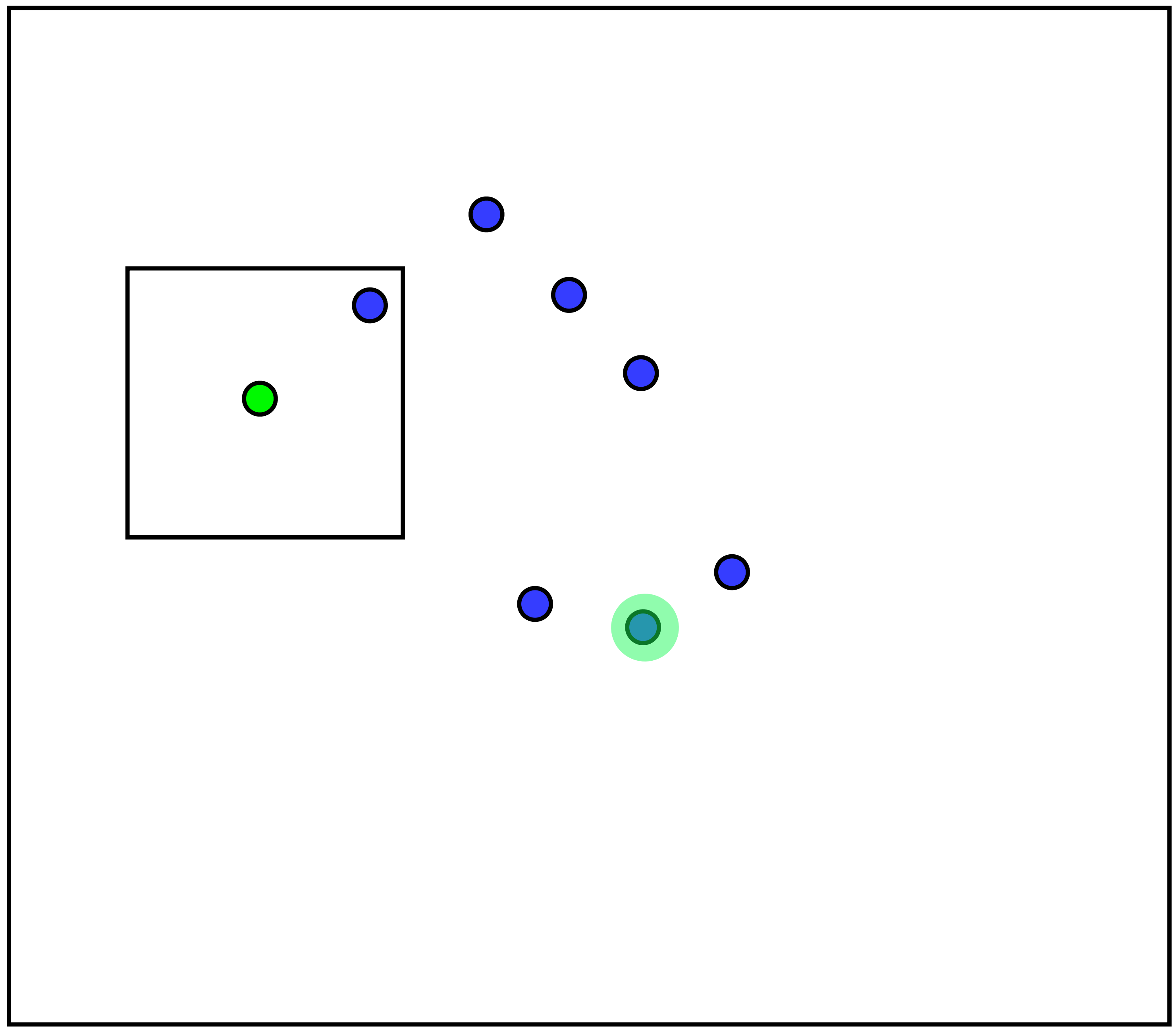
[Ratitch et al. 2004]

Sparse Distributed Memories



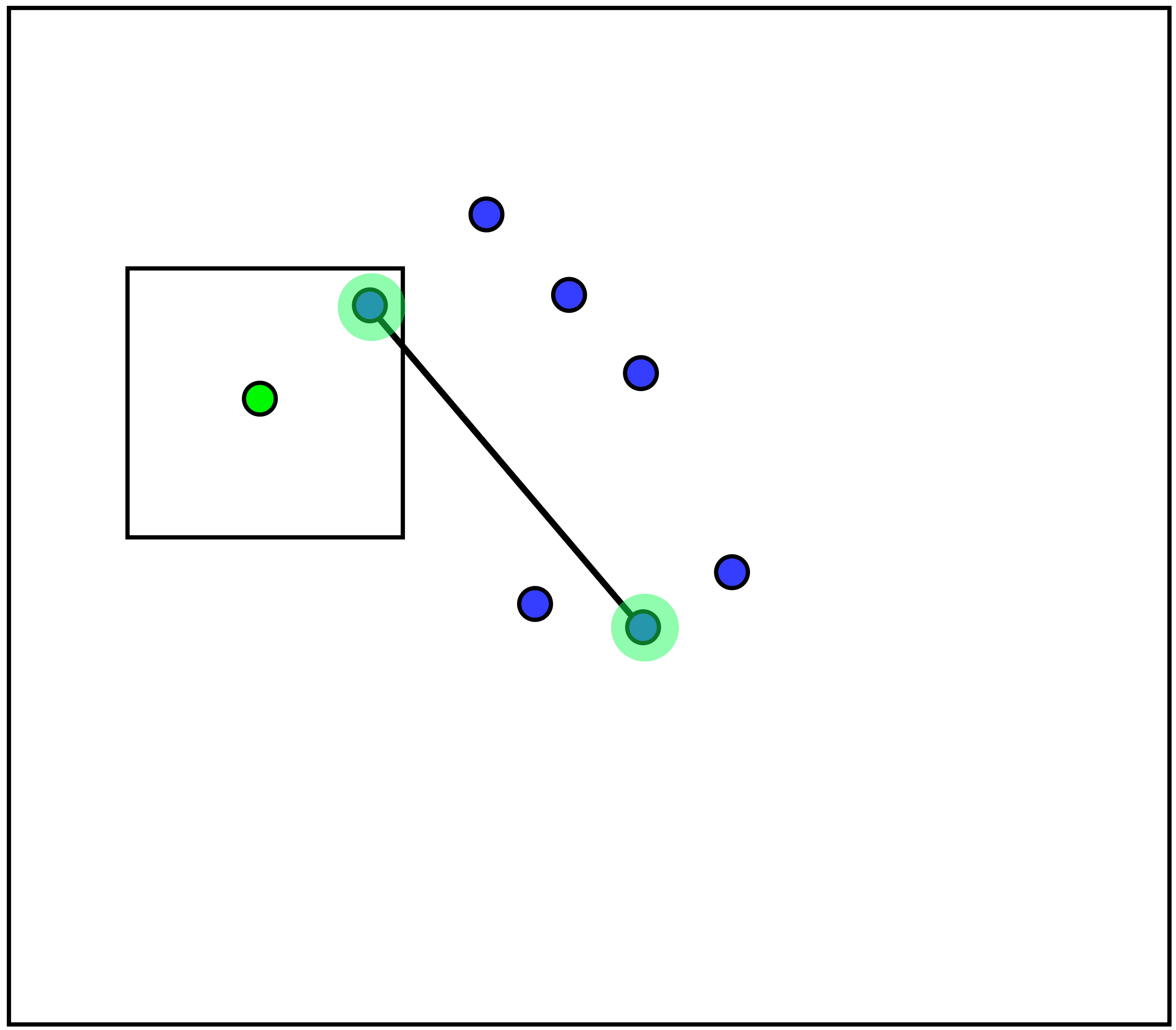
[Ratitch et al. 2004]

Sparse Distributed Memories



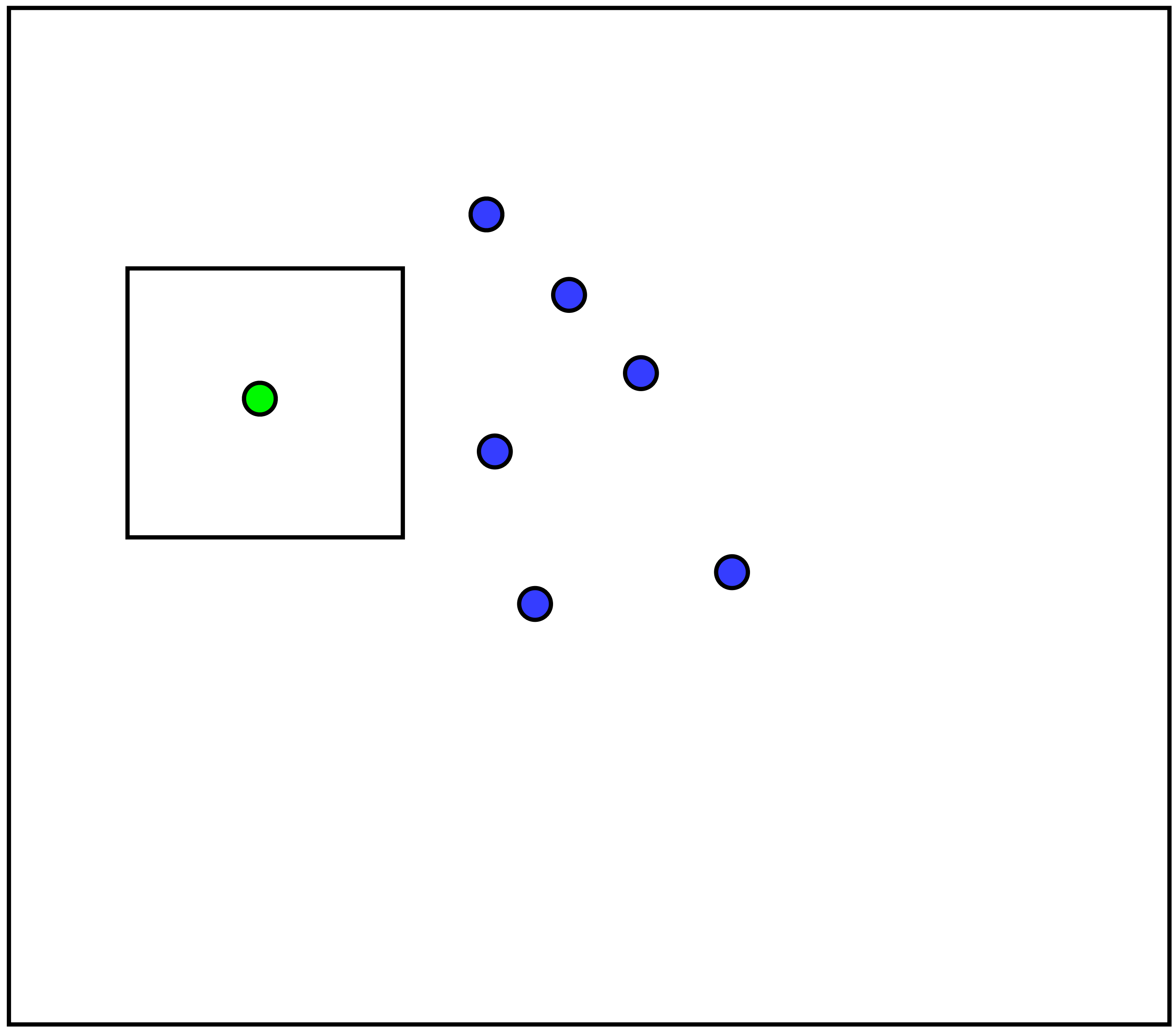
[Ratitch et al. 2004]

Sparse Distributed Memories



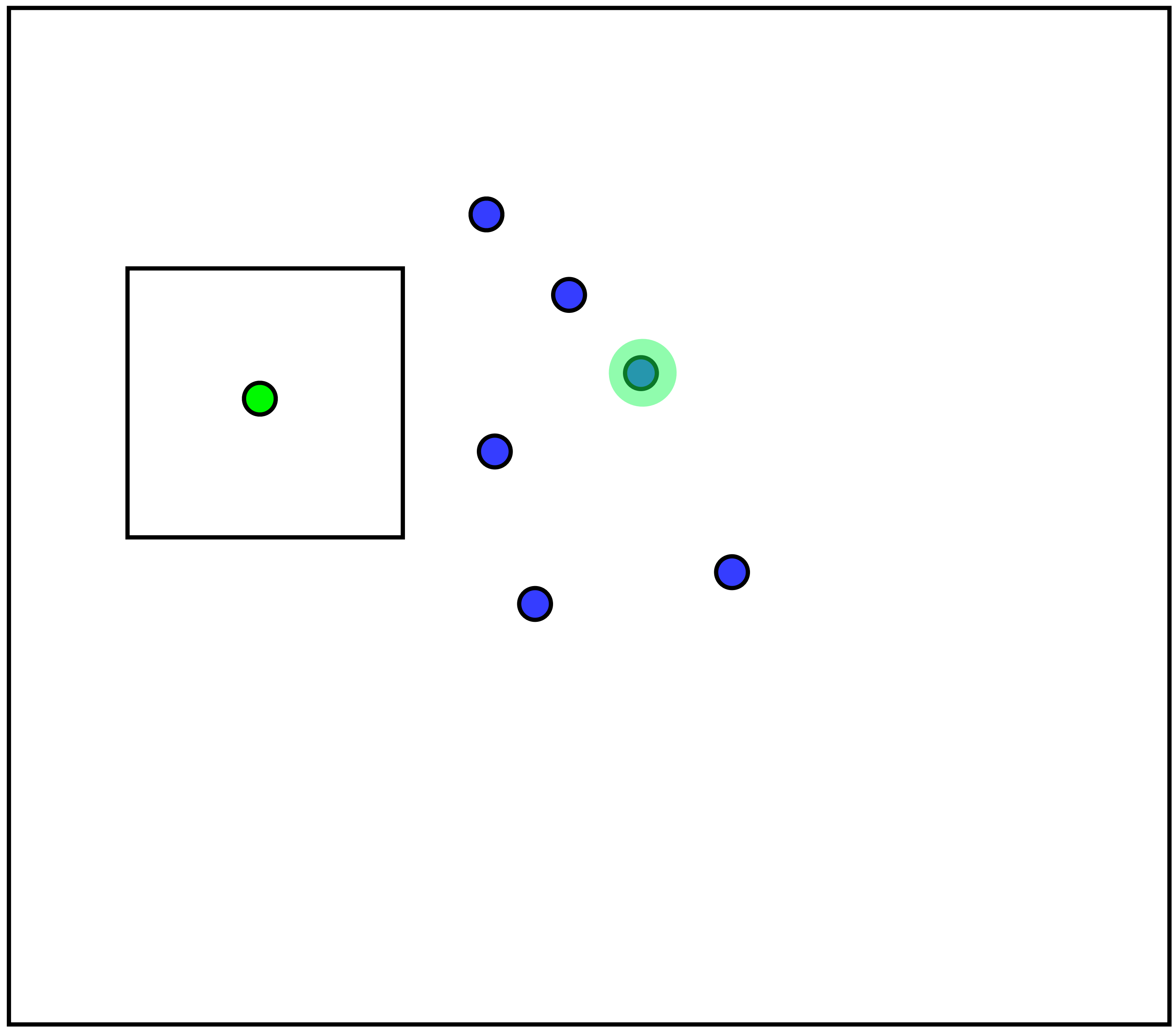
[Ratitch et al. 2004]

Sparse Distributed Memories



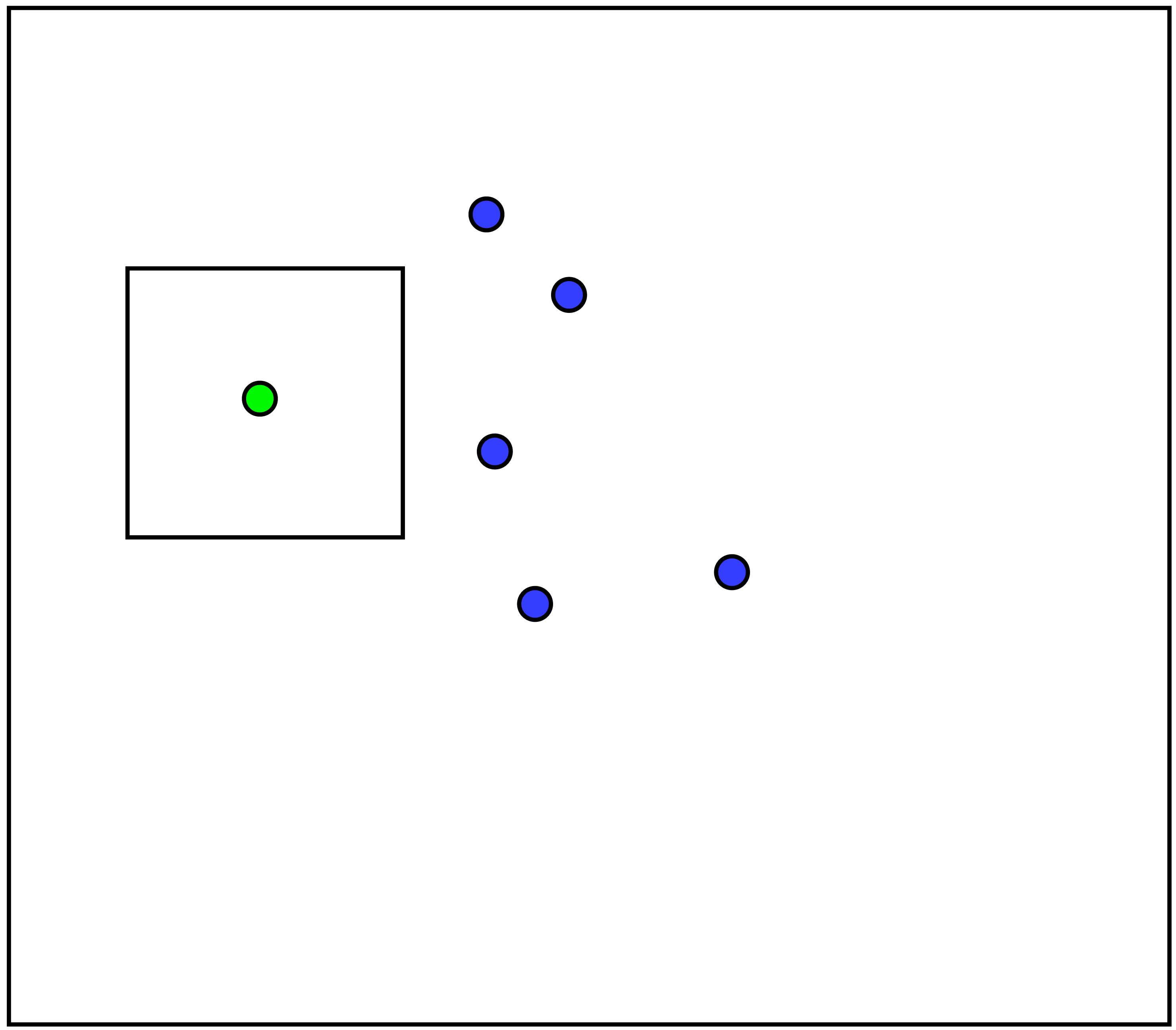
[Ratitch et al. 2004]

Sparse Distributed Memories



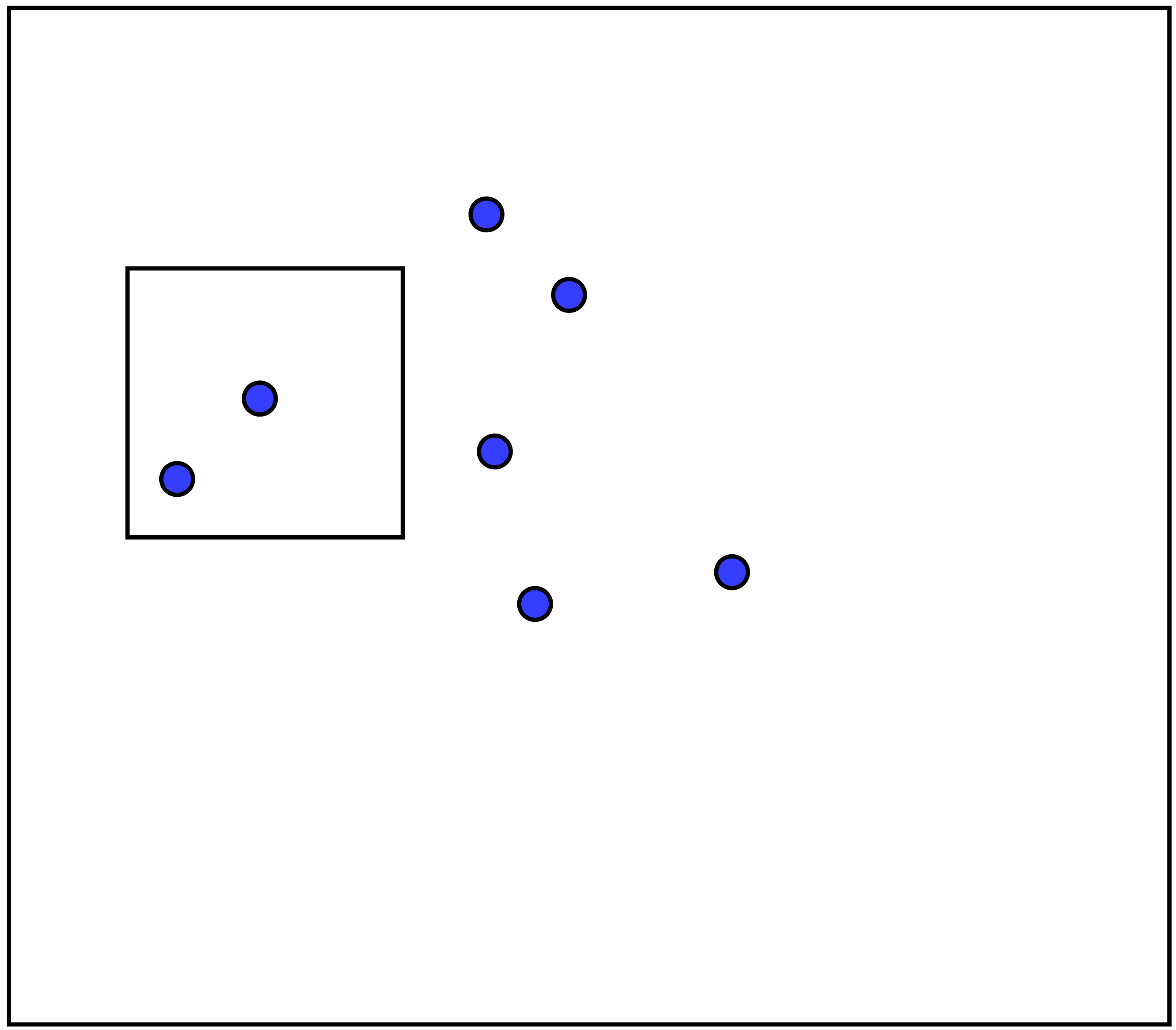
[Ratitch et al. 2004]

Sparse Distributed Memories



[Ratitch et al. 2004]

Sparse Distributed Memories



[Ratitch et al. 2004]