Rethinking Machine Learning in the 21st Century: From Optimization to Equilibration

Sridhar Mahadevan
Autonomous Learning Laboratory
UMASS CS
SCHOOL OF COMPUTER SCIENCE
50 YEARS

Research funded in part by the National Science Foundation, NASA and AFOSR

National Science Foundation
WHERE DISCOVERIES BEGIN

NASA

AFRL
THE AIR FORCE RESEARCH LABORATORY
LEAD | DISCOVER | DEVELOP | DELIVER
AFOSR

(Almost) Dimension-Free
optimization

To

(Almost) Dimension-Free
equilibration

ALL

AUTONOMOUS LEARNING LABORATORY

1981-2014

http://all.cs.umass.edu/pubs.shtml

UMASS CS

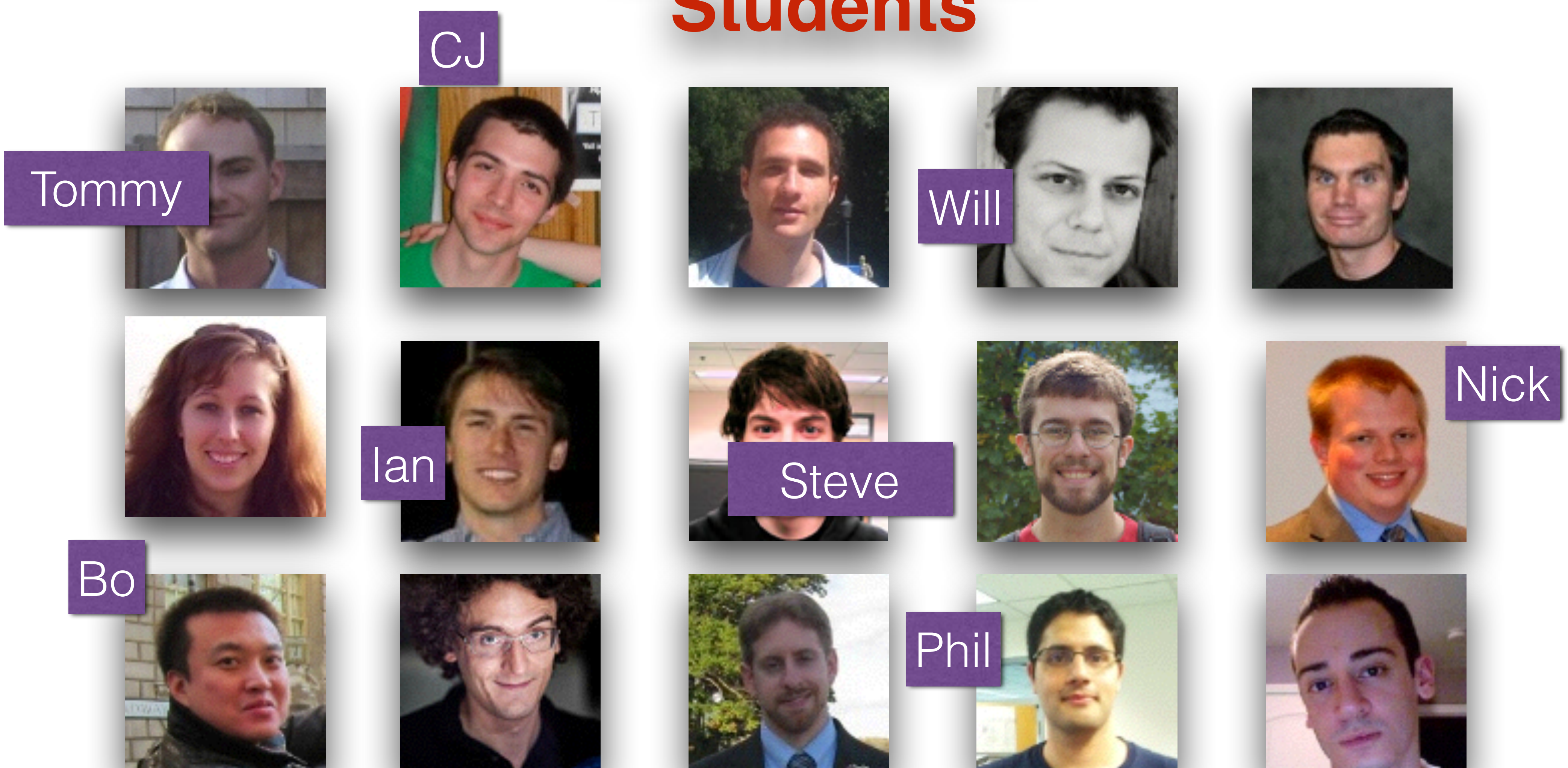
SCHOOL OF COMPUTER SCIENCE

50 YEARS

Barto (Emeritus)

Lab Directors

Current PhD Students



Thomas Boucher

CJ Carey

Bruno Castro da Silva

William Dabney

Stefan Dernbach

Kimberly Ferguson

Ian Gemp

Stephen Giguere

Thomas Helmuth

Nicholas Jacek

Bo Liu

Clemens Rosenbaum

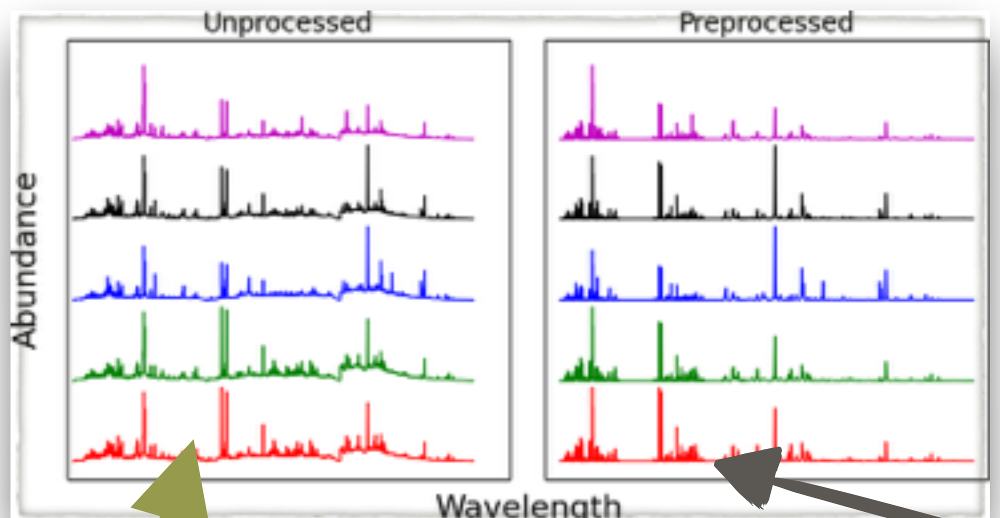
Andrew Stout

Philip Thomas

Chris Vigorito

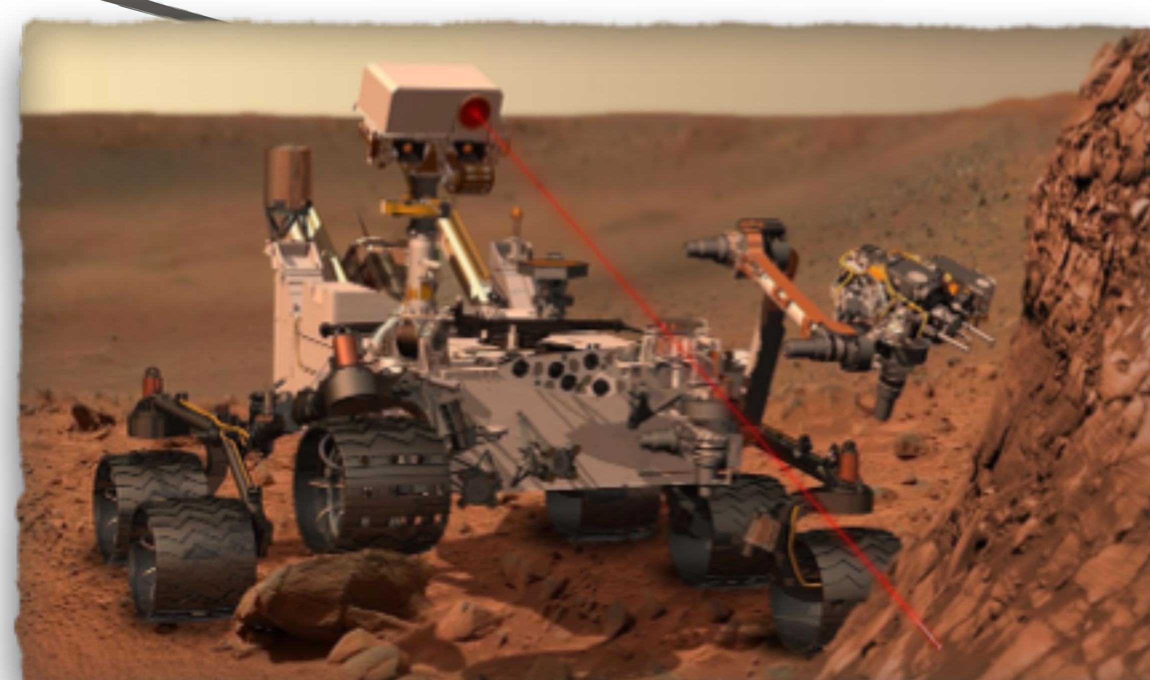
Transfer Learning on Mars

(Darby Dyar, Mount Holyoke; Thomas Boucher, Clifton Carey, Stephen Giguere, UMass)

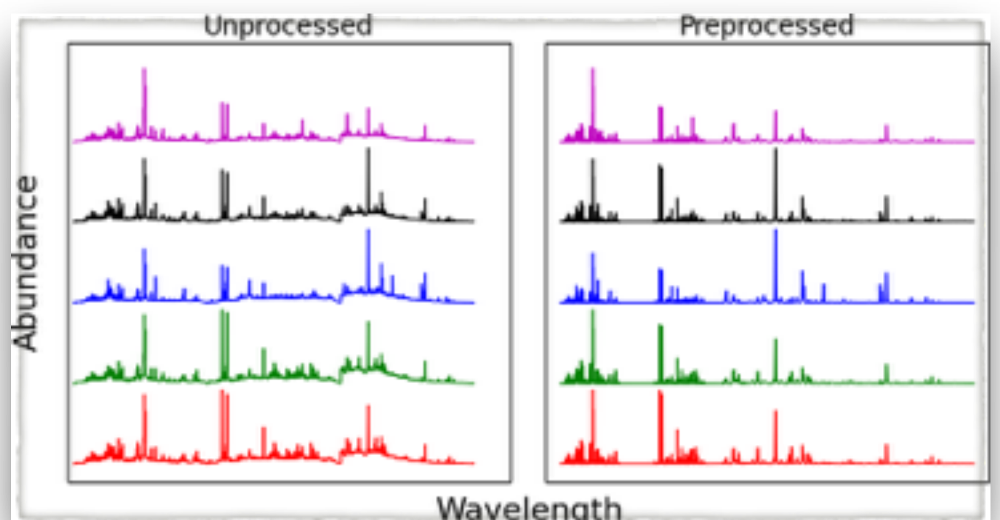


Curiosity zapping a rock with a laser

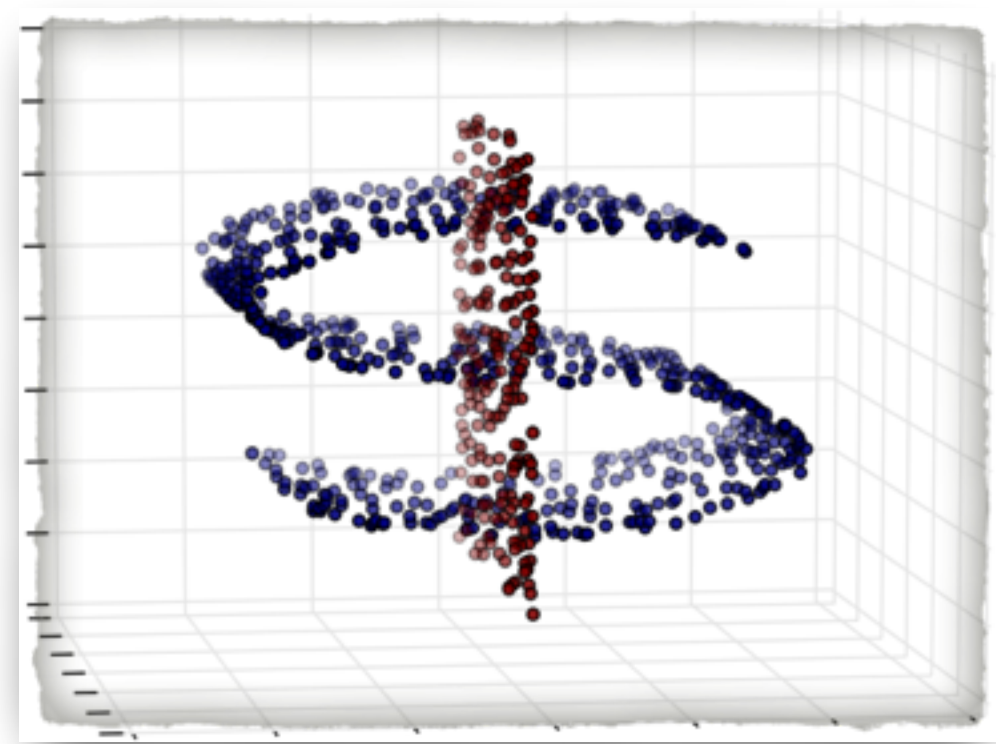
Same laser on Earth as on Mars



Low-Dimensional Representation Discovery



Original Data

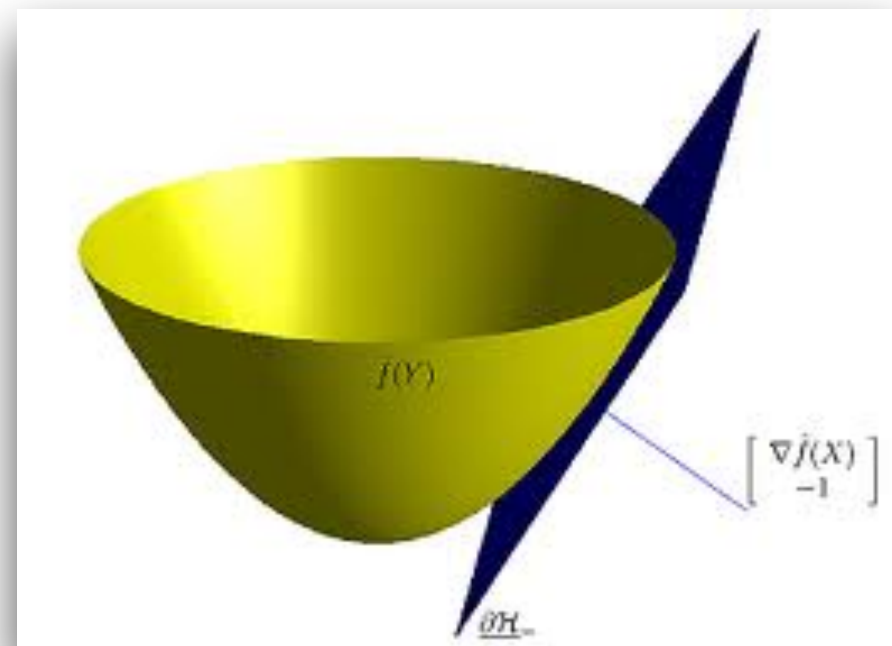


Mixture of low-dimensional (non-Euclidean) spaces

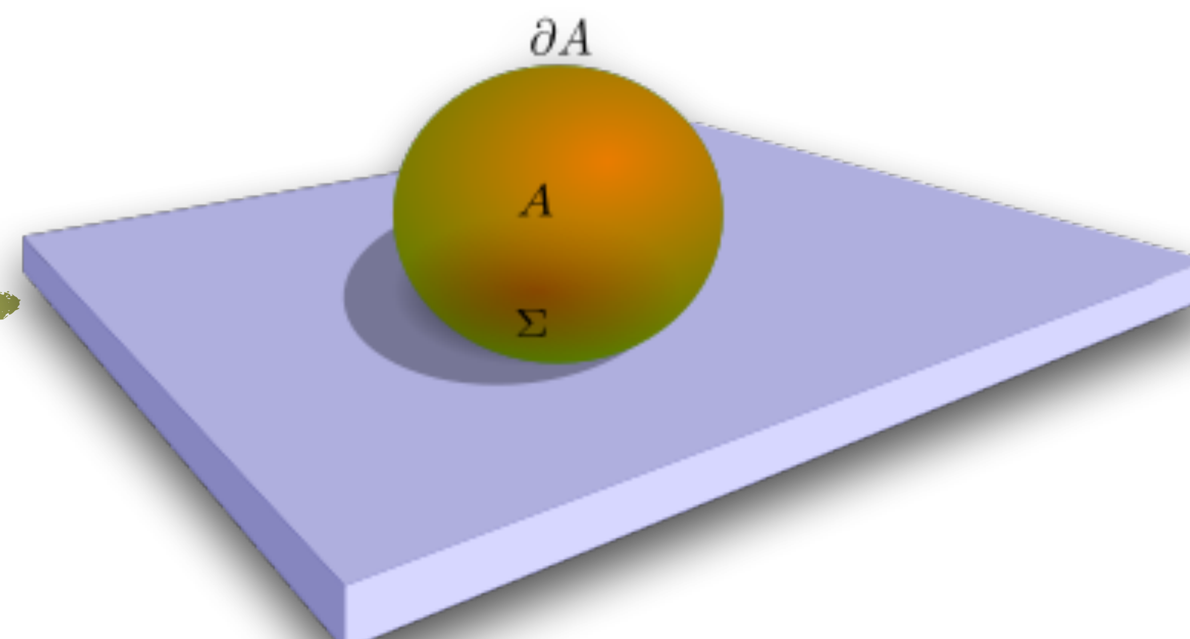
New dimensionality reduction method

Optimization

Equilibration



min f(x)
x in feasible set K



$$\langle F(x^*), x - x^* \rangle \geq 0, \ \forall x \in K$$

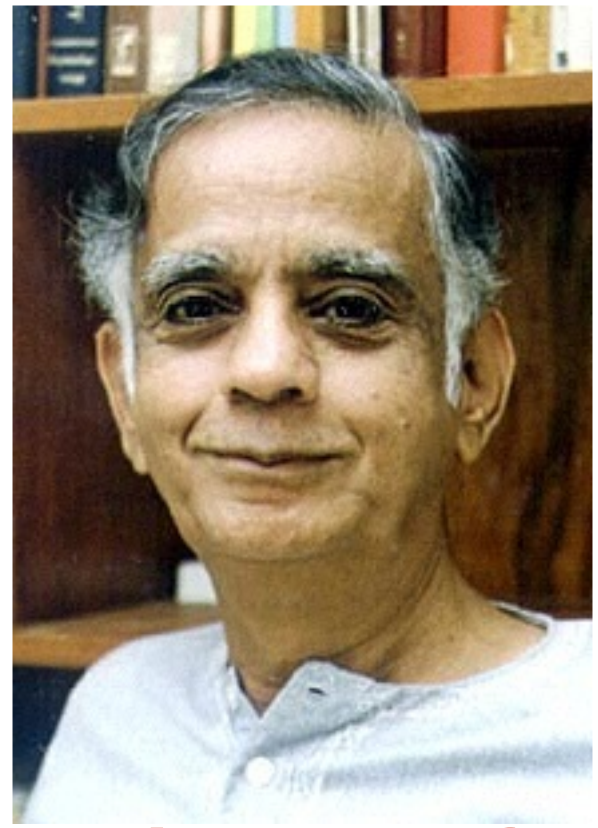
(Stampacchia, 1960s)

Part I: Some (Personal) History and Motivation

Iravatham
Mahadevan

ஐராவதம் மகாதேவன்

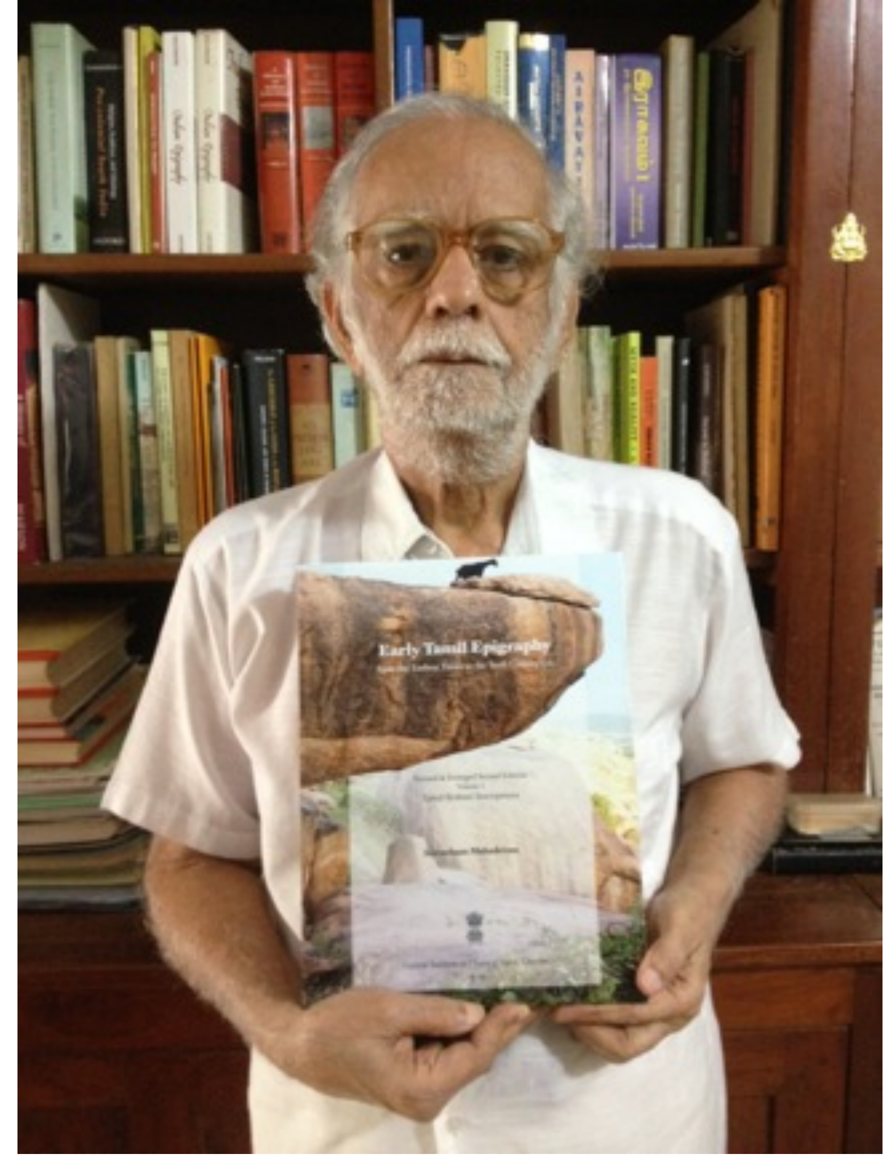
5th century BC
Tamil Brahmi



Early Tamil Epigraphy from the Earliest Times to the Sixth Century A.D
Edited and translated by Iravatham Mahadevan
Harvard Oriental Series 62

Iravatham Mahadevan, June 2014 at 85

A DRAVIDIAN ETYMOLOGICAL DICTIONARY



1904-2005

Obituaries | PASSINGS

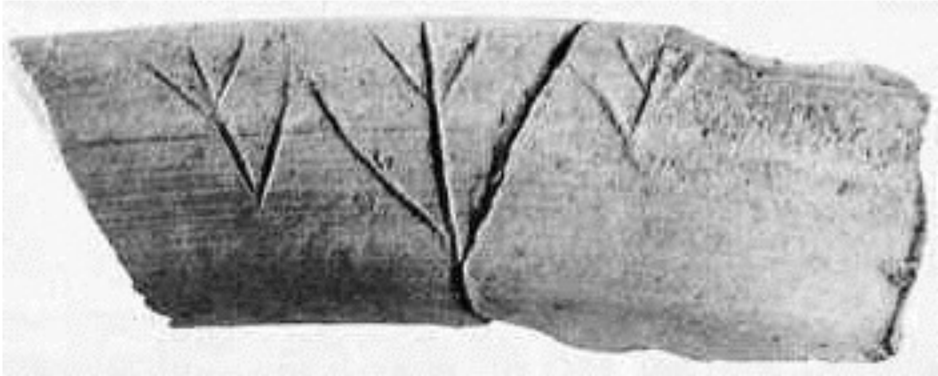
Murray Emeneau, 101; Founded UC Berkeley Linguistics Department

Murray Barnson Emeneau, 101, an expert in Sanskrit and Dravidian languages who founded the UC Berkeley Linguistics Department, died Aug. 29 in his sleep of natural causes at his Berkeley home.

2500
B.C.

Indus Script

Great Bath
Mohenjadaro

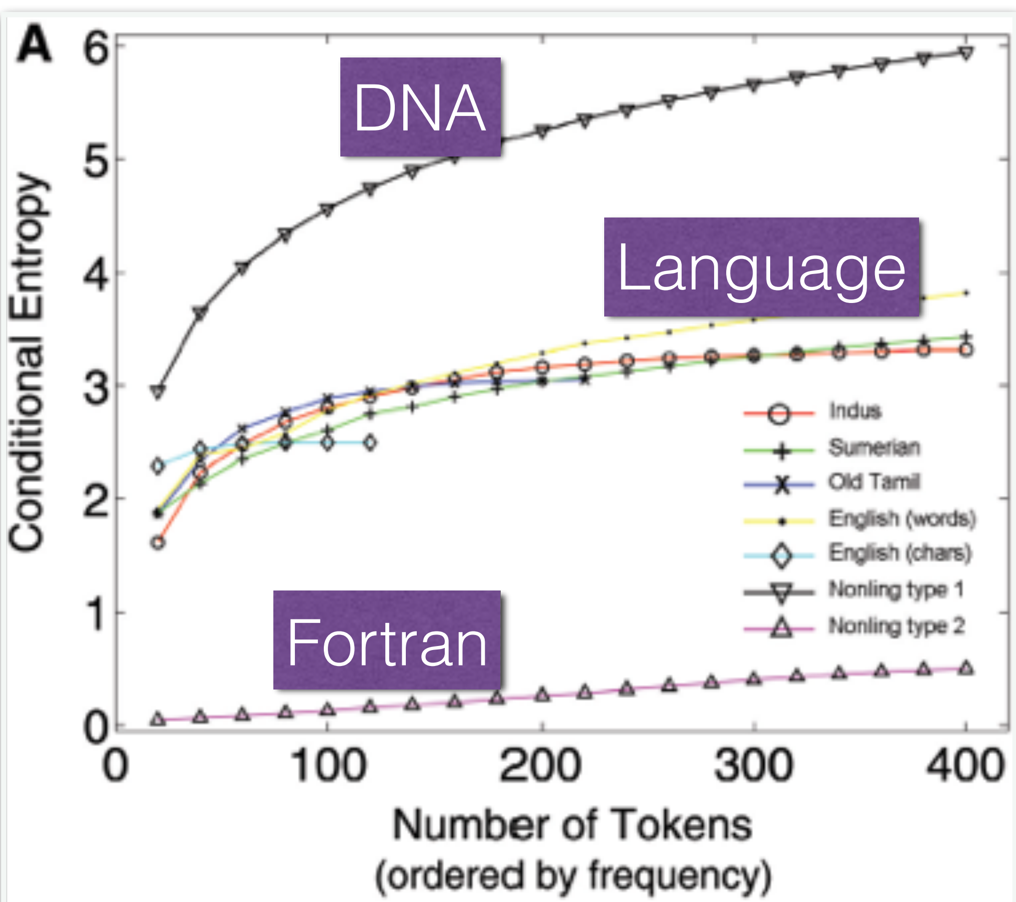


Published Online April 23 2009
Science 29 May 2009:
Vol. 324 no. 5931 p. 1165
DOI: 10.1126/science.1170391

- BREVIA

Entropic Evidence for Linguistic Structure in the Indus Script

. Rajesh P. N. Rao, Nisha Yadav, Mayank N. Vahia, Hrishikesh Joglekar, R. Adhikari, Iravatham Mahadevan



Is the Indus Script really a language?

Entia non sunt multiplicanda sine necessitate

Bertrand Russell's explication of Willam's philosophy

William of Ockham

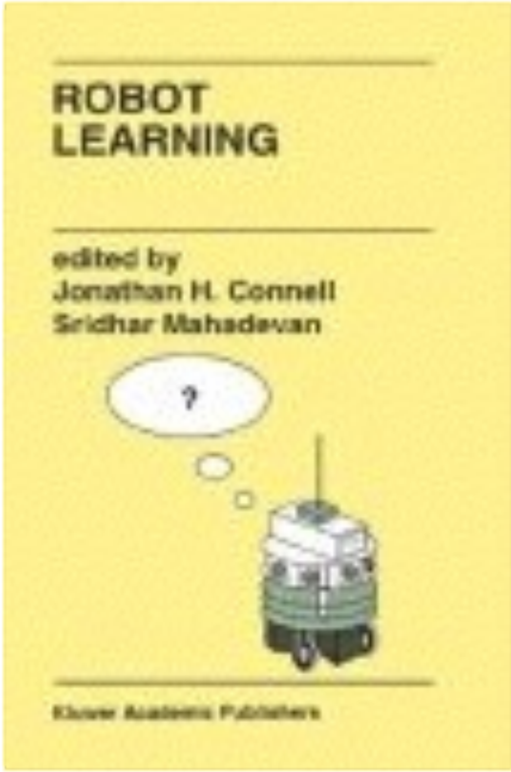
14th century logician and Franciscan friar

Why should machines learn?

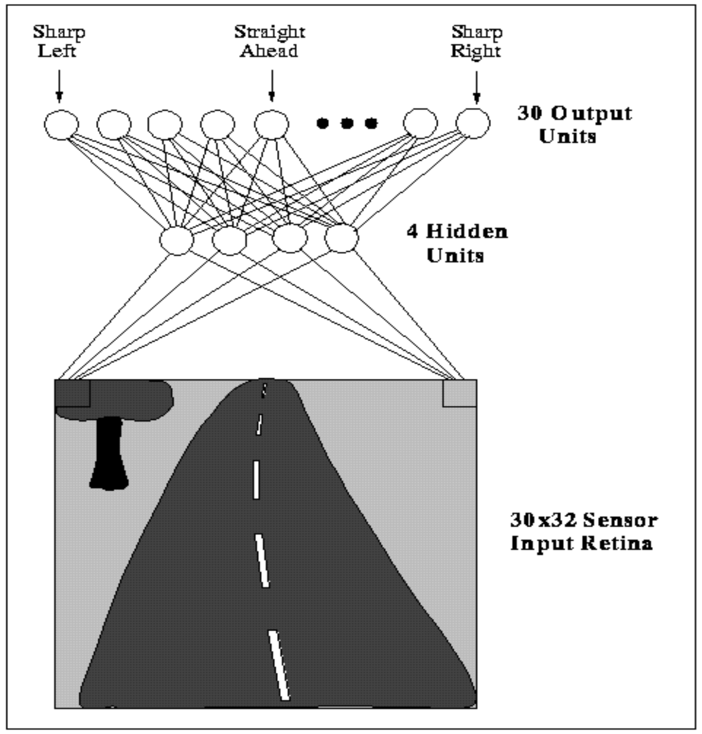
Herbert Simon

Learning denotes changes in a system that are adaptive in that they enable the system to do the same task or tasks drawn from the same population more efficiently and more effectively the next time around

Learning to Drive



ALVINN learns
from a human driver



Neural
Network

Can drive on actual highways at 65 miles per hour!

Søren Kierkegaard

19th century Danish philosopher

Life must be lived going forwards
but can only be understood
going backwards

Arthur Samuel's
Checker Playing Program
(IBM 700, 1956)

IBM's T.J. Watson said it would increase IBM stock by 10 points and it did!

First demo of ML!

MIT Press

Andy Barto

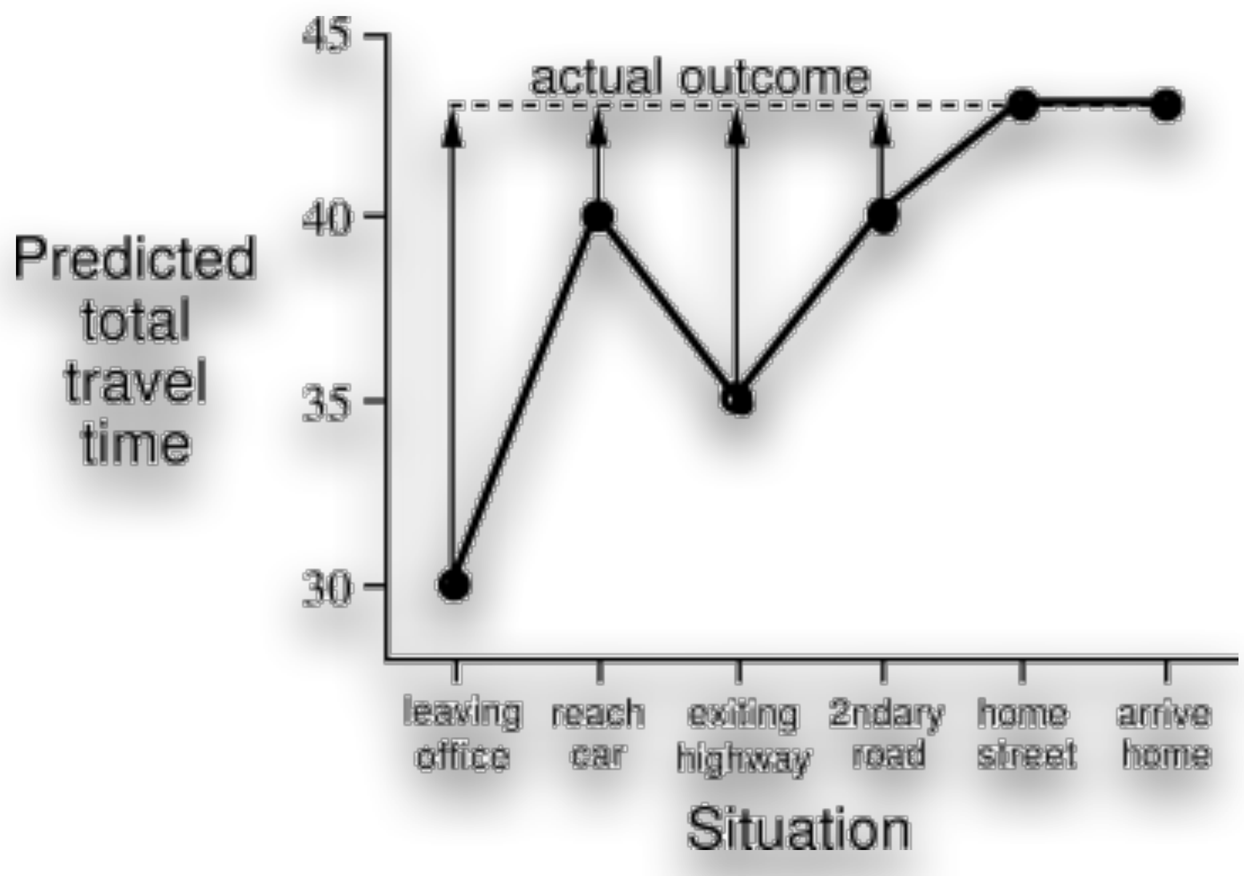
Richard Sutton

> 16,000 citations

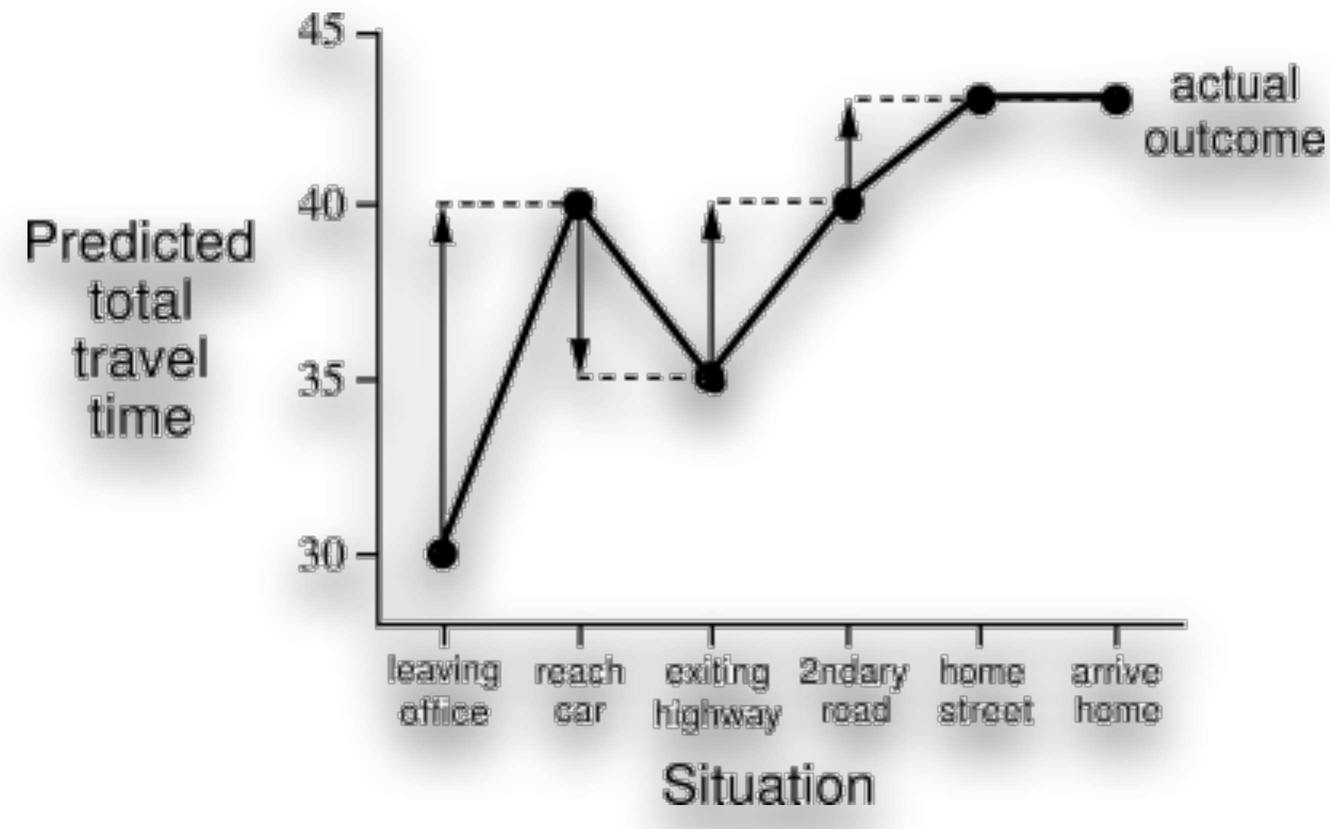
Temporal Difference Learning Algorithm

http://webdocs.cs.ualberta.ca/~sutton/book/ebook/node1.html

Least Squares



Temporal Difference Learning



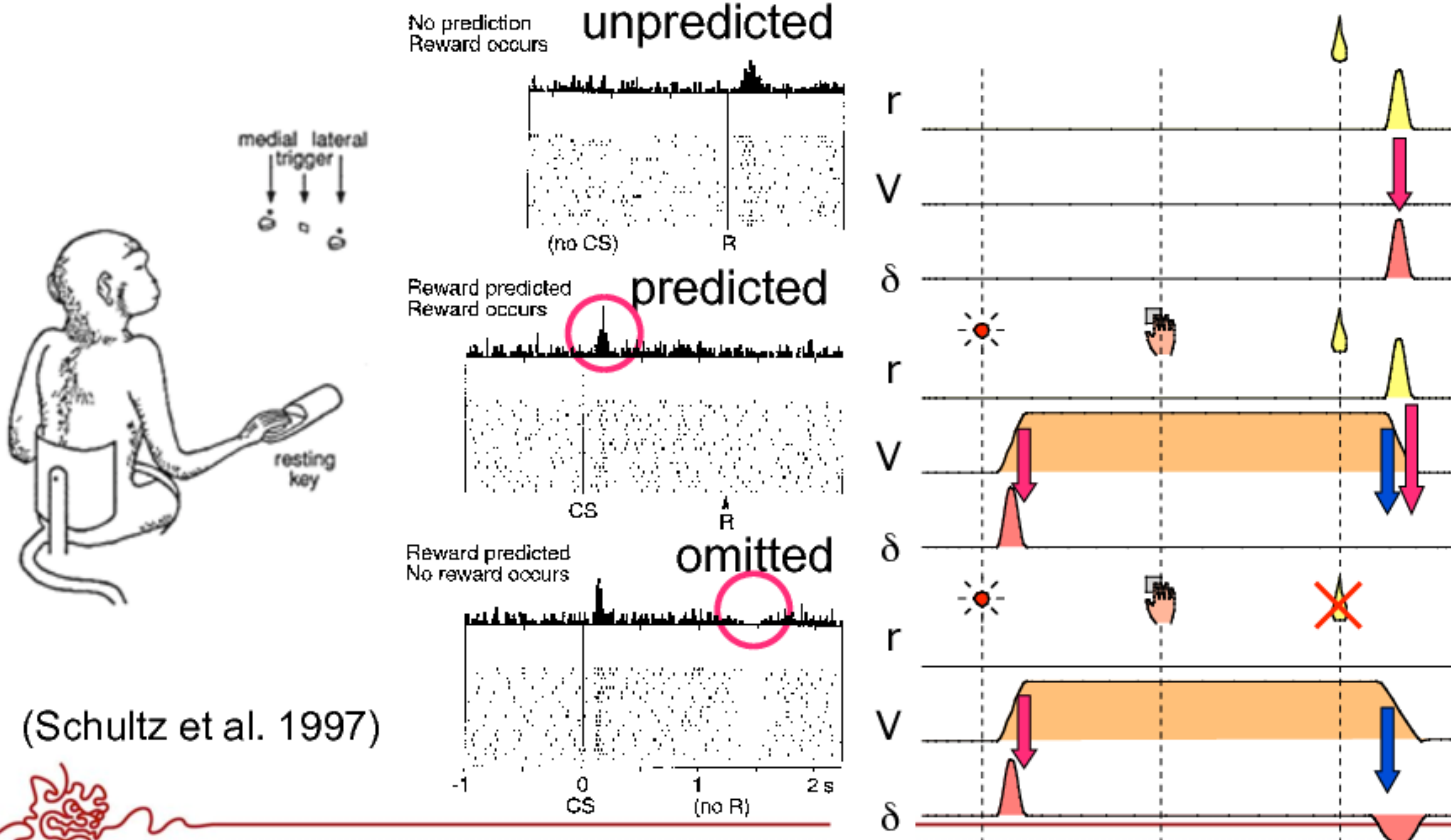
Algorithm 1 TD (1984)

(1) $\delta_t = r_t + \gamma \phi_t'^T \theta_t - \phi_t^T \theta_t$

(2) $\theta_{t+1} = \theta_t + \beta_t \delta_t$

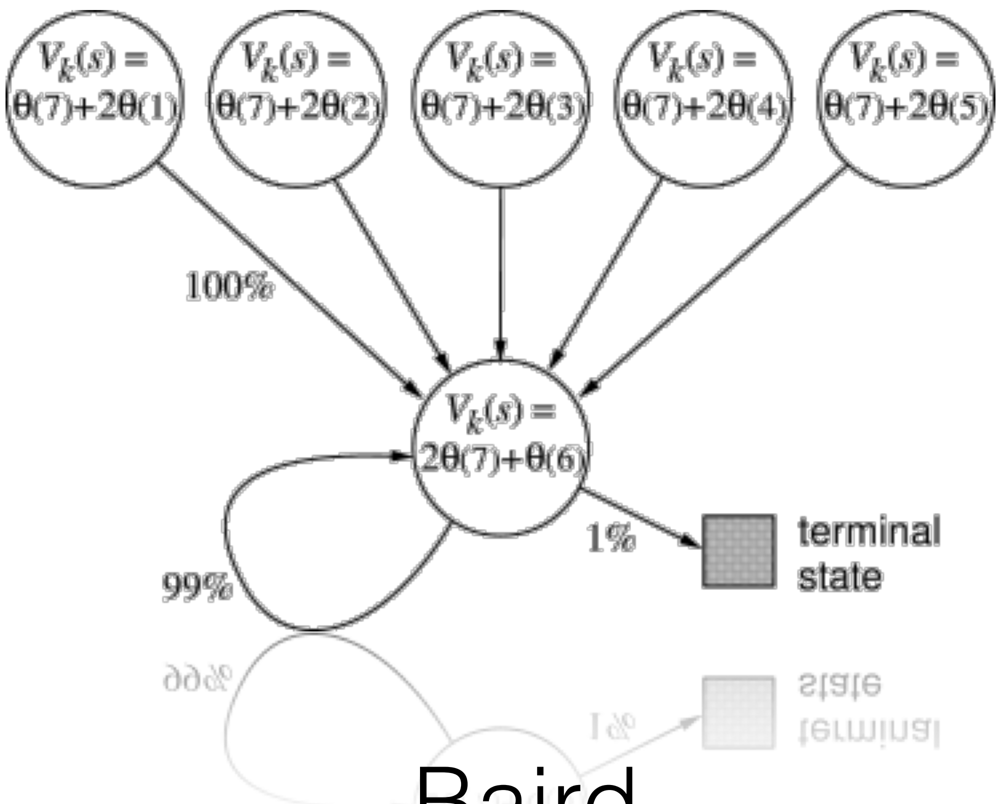
Dopamine Neurons Code TD Error

$$\delta(t) = r(t) + \gamma V(s(t+1)) - V(s(t))$$



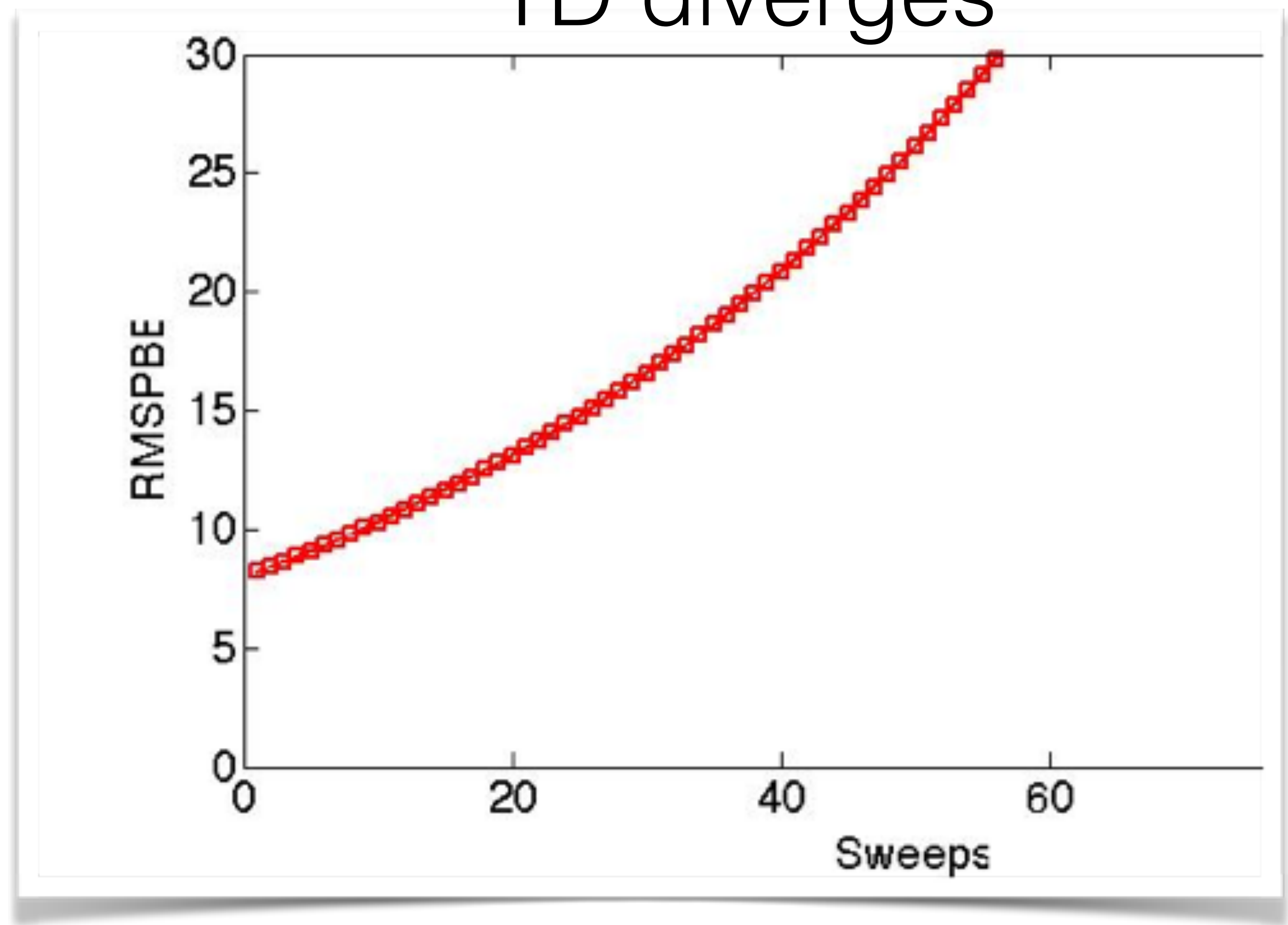
(Schultz et al. 1997)

TD-Learning Fails (not always, but predictably!)



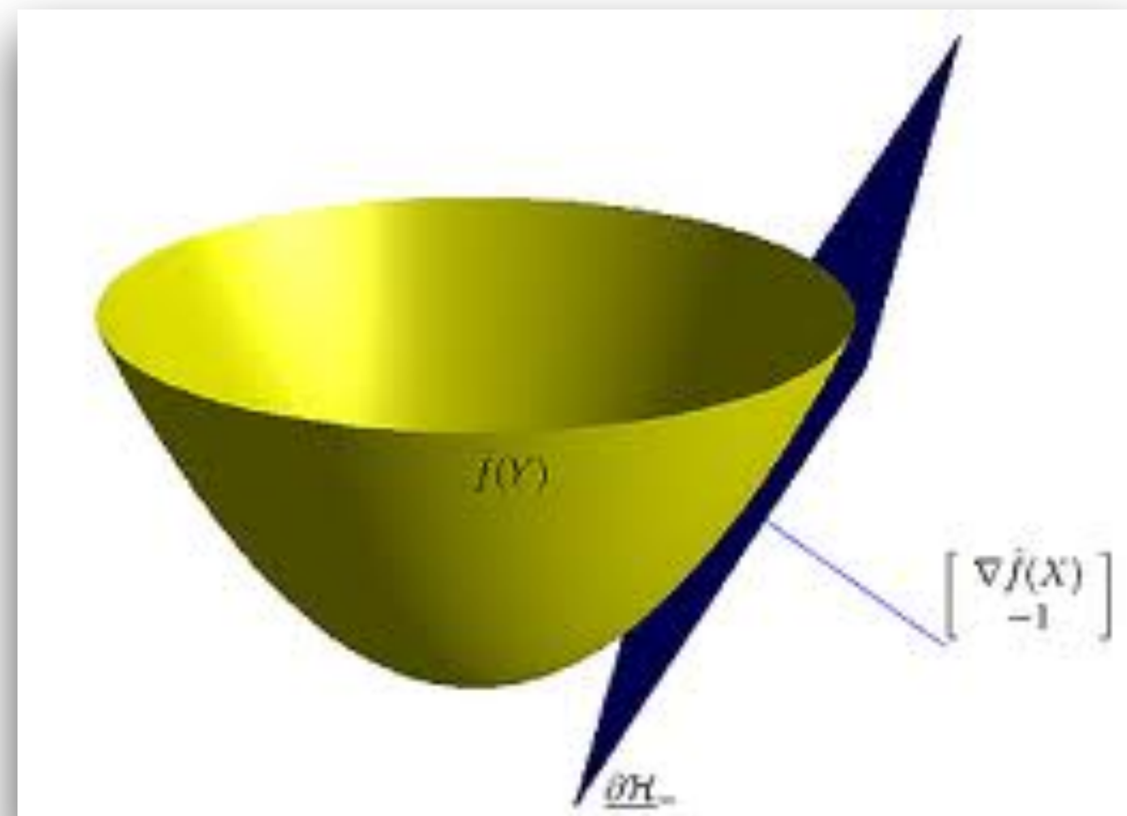
Baird
counter example

TD diverges



Optimization by Gradient Descent TD

Min f(x), x in feasible set K



1984-2014:
Can TD be converted into a "true" gradient method?

Latest attempt: Sutton, 2009
Introduces gradient objectives but solves without true gradient computation

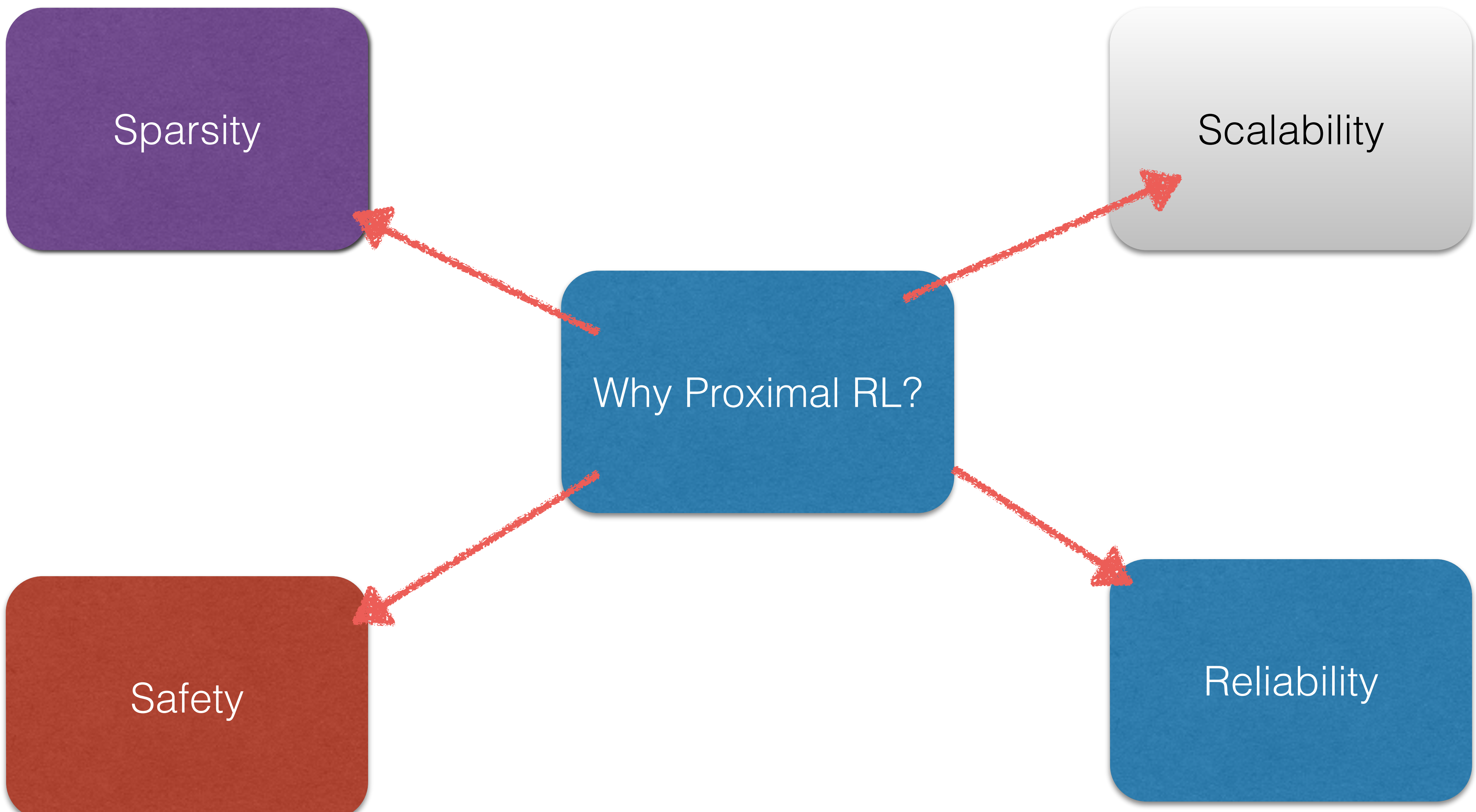
Part II:Equilibration in Reinforcement Learning

Proximal Reinforcement Learning: A New Theory of Sequential Decision Making in **Primal-Dual** Spaces,

Arxiv, May 26, 2014 (126 pages)

Sridhar Mahadevan, Bo Liu, Philip Thomas, Will Dabney,
Stephen Giguere,
Nicholas Jacek, Ian Gemp,
Ji Liu

Proximal Reinforcement Learning in Primal-Dual Spaces



Søren Kierkegaard

19th century Danish philosopher

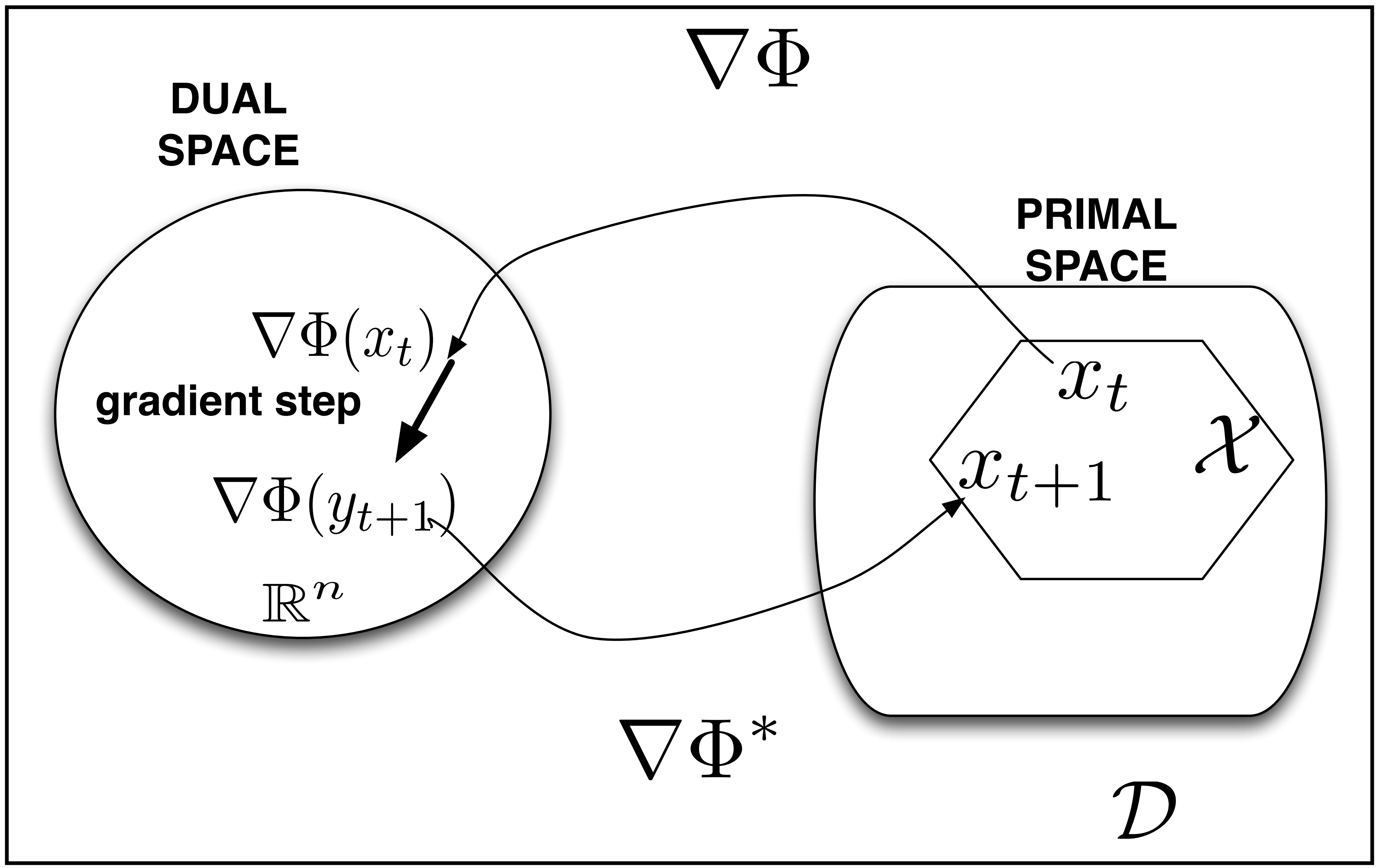
Life must be lived going forwards
but can only be understood
going backwards

In the Dual Space!

(Almost) Dimension-free optimization

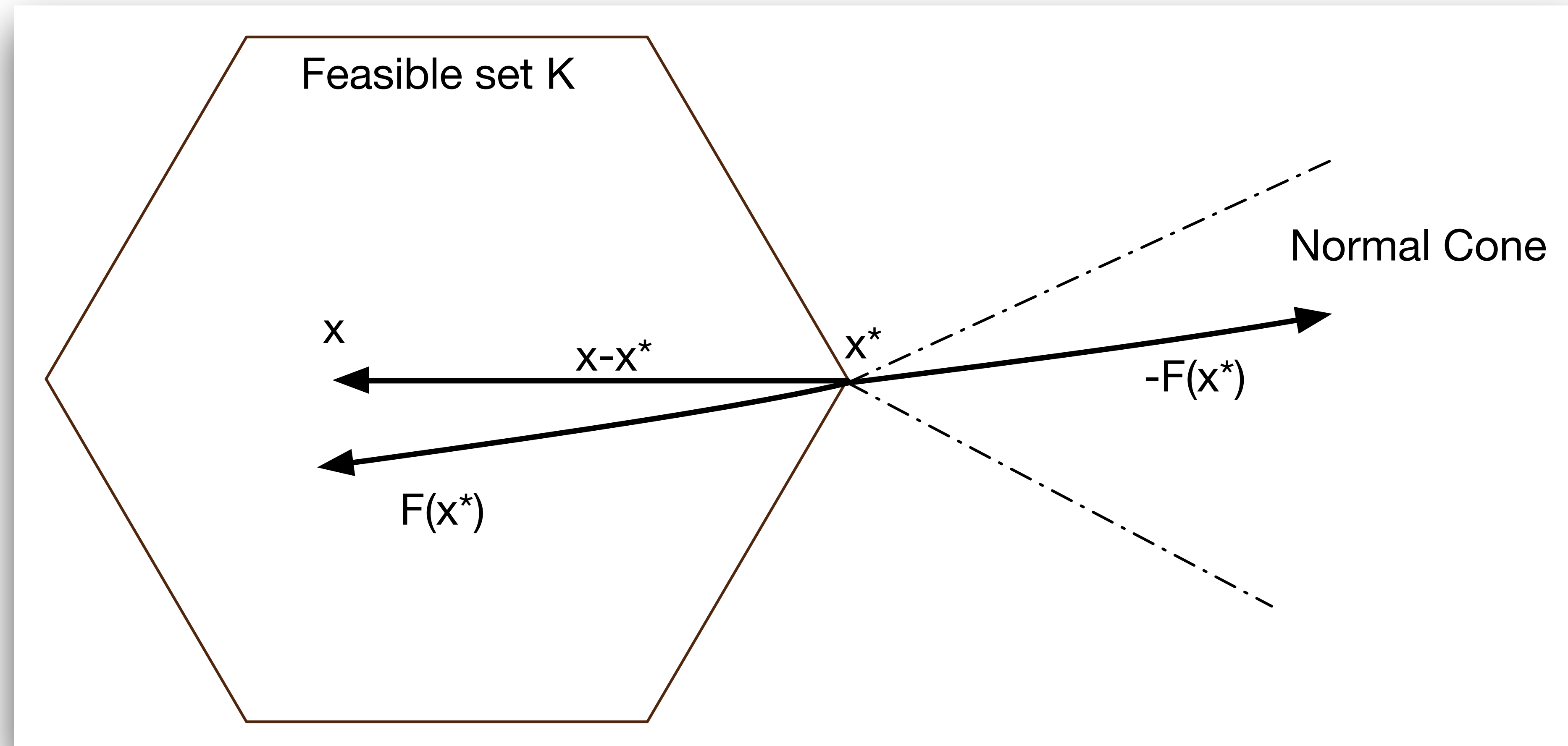
Mirror Maps

(Nemirovski and Yudin, 1980s; Bubeck, 2014)



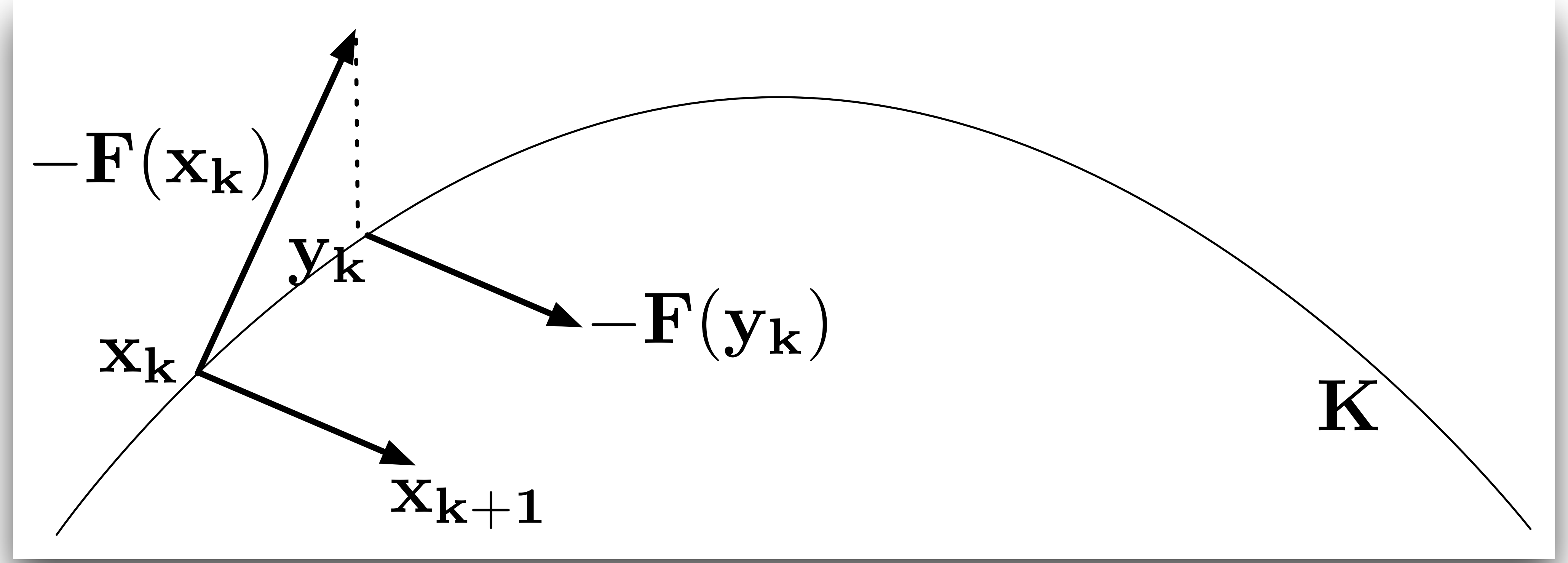
Variational Inequality

(Stampacchia, 1960s)



$$\langle F(x^*), x - x^* \rangle \geq 0, \ \forall x \in K$$

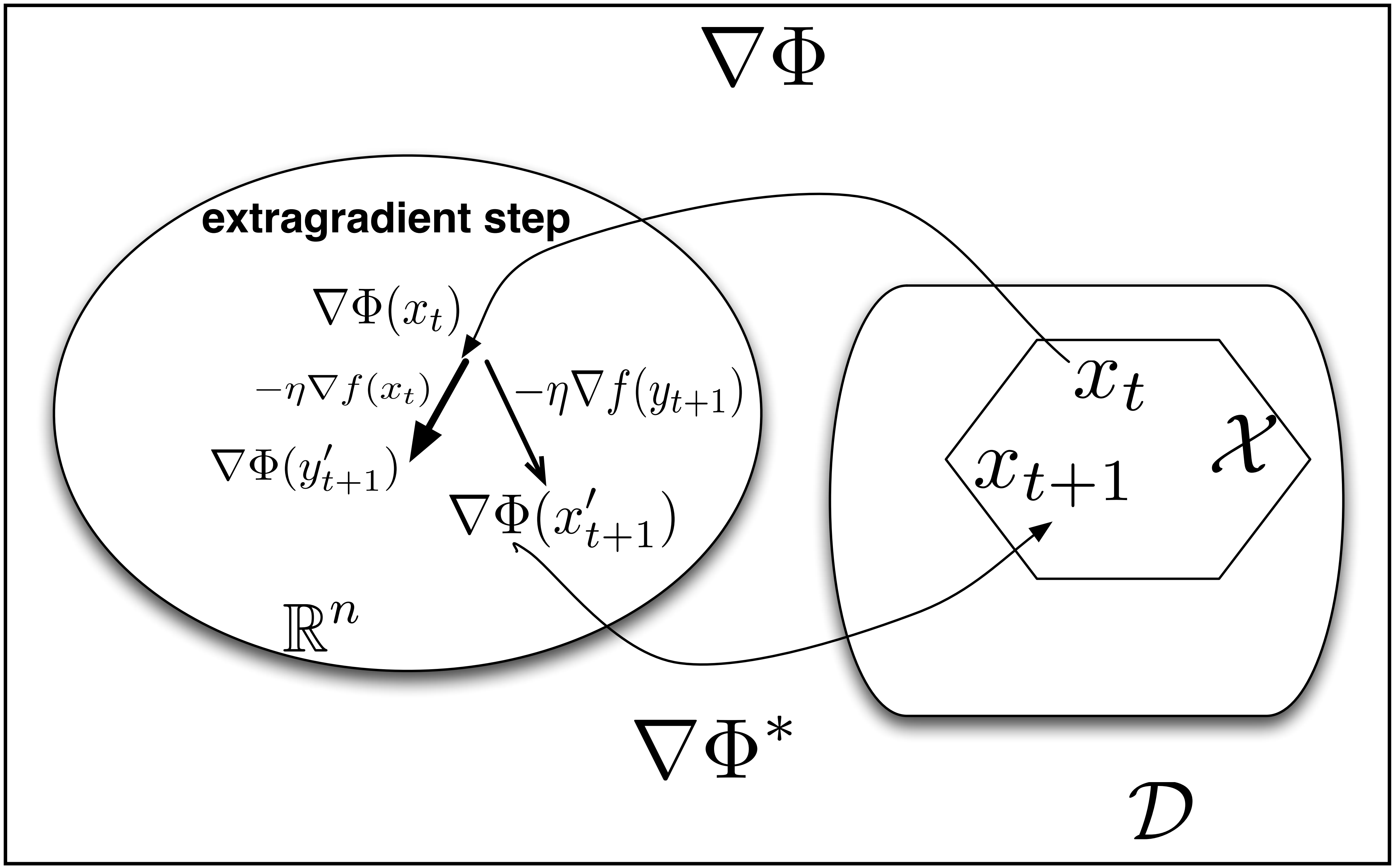
Extragradient Method



Korpolevich (1970s) developed the extragradient method

Mirror-Prox: Non-Euclidean Extragradient

(Nemirovski, 2005)



True Gradient TD-Learning: RL meets VI

Algorithm 2 GTD2-MP (2014)

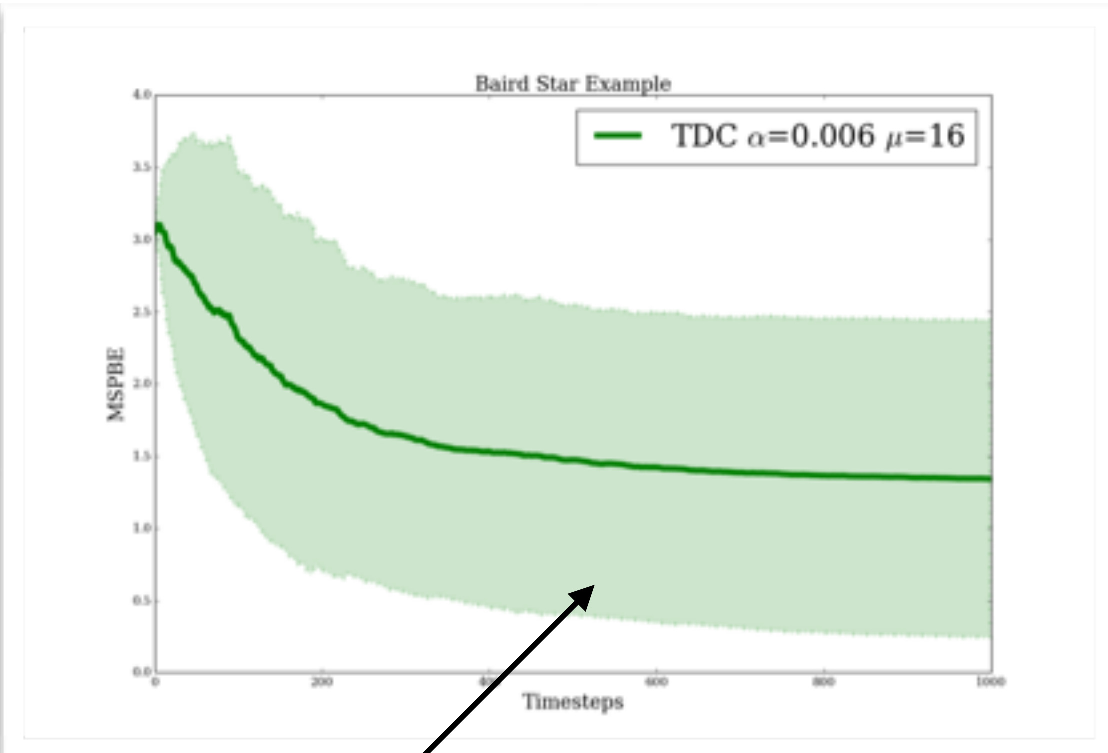
(1)
$$w_{t+\frac{1}{2}} = w_t + \beta_t(\delta_t - \phi_t^T w_t)\phi_t,$$
$$\theta_{t+\frac{1}{2}} = \text{prox}_{\alpha_t h}\left(\theta_t + \alpha_t(\phi_t - \gamma\phi_t')(\phi_t^T w_t)\right)$$

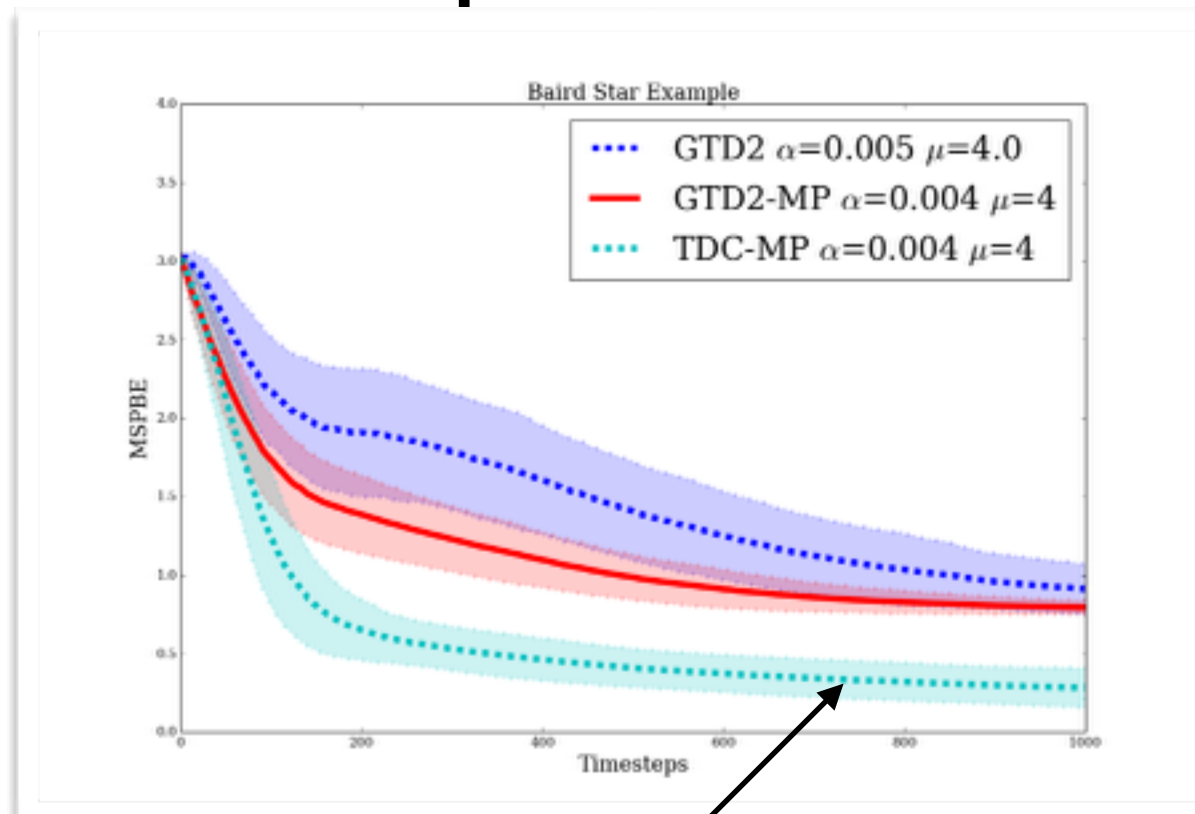
(2)
$$\delta_{t+\frac{1}{2}} = r_t + \gamma{\phi_t'}^T\theta_{t+\frac{1}{2}} - \phi_t^T\theta_{t+\frac{1}{2}}$$

(3)
$$w_{t+1} = w_t + \beta_t(\delta_{t+\frac{1}{2}} - \phi_t^T w_{t+\frac{1}{2}})\phi_t \ ,$$
$$\theta_{t+1} = \text{prox}_{\alpha_t h}\left(\theta_t + \alpha_t(\phi_t - \gamma\phi_t')(\phi_t^T w_{t+\frac{1}{2}})\right)$$

(Mahadevan et al., Arxiv 2014)

Baird counter example





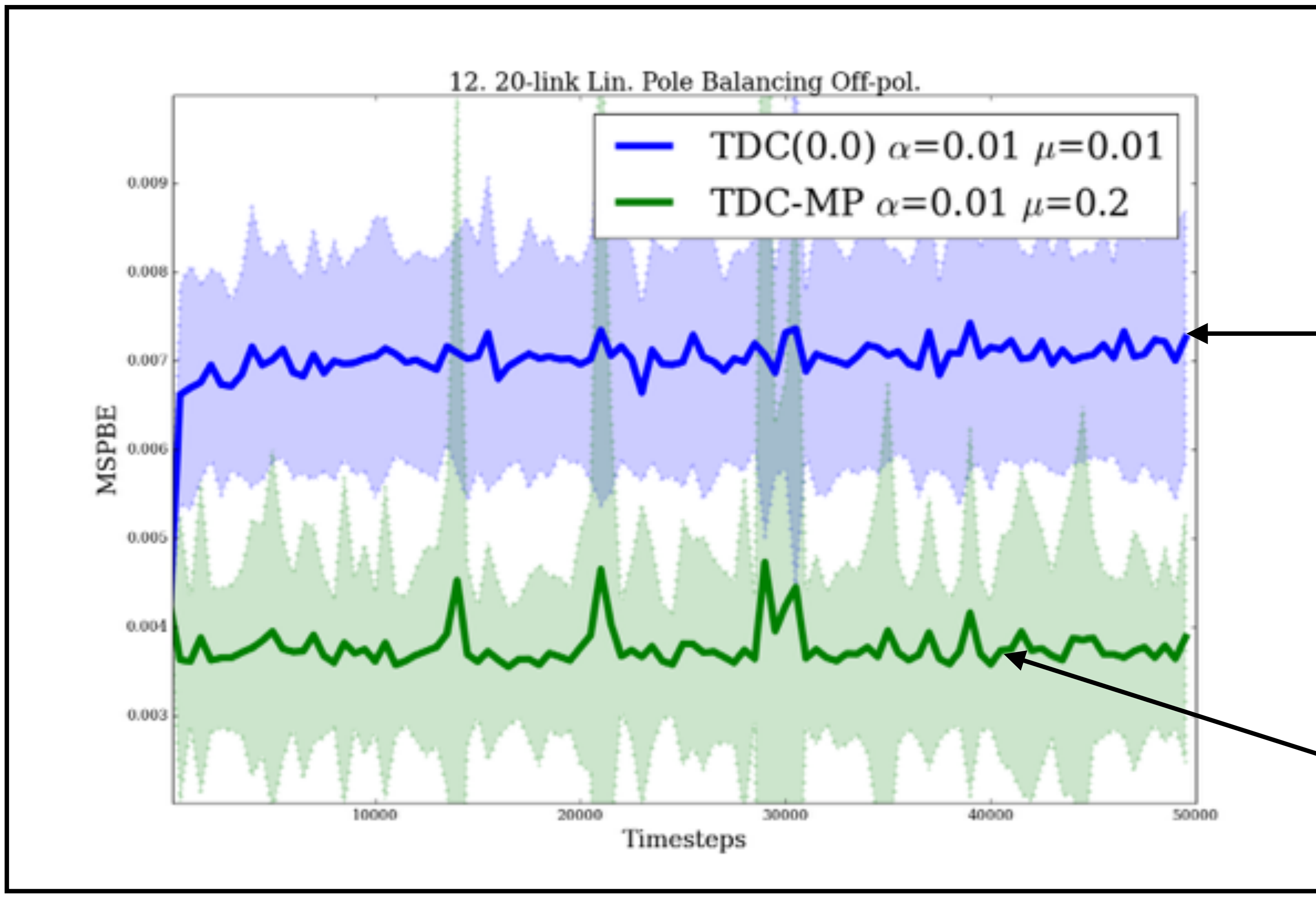
Sutton et al., 2009

Variance of previous methods: $\propto O(\frac{1}{\sqrt{n}})$

Our new methods 2014

Variance of our methods: $\propto O(\frac{1}{n})$

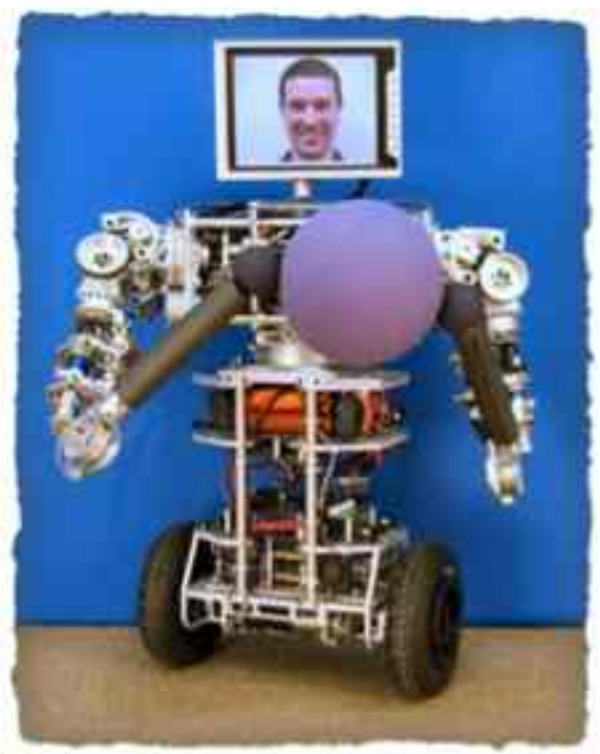
20-Dimensional Robot Arm



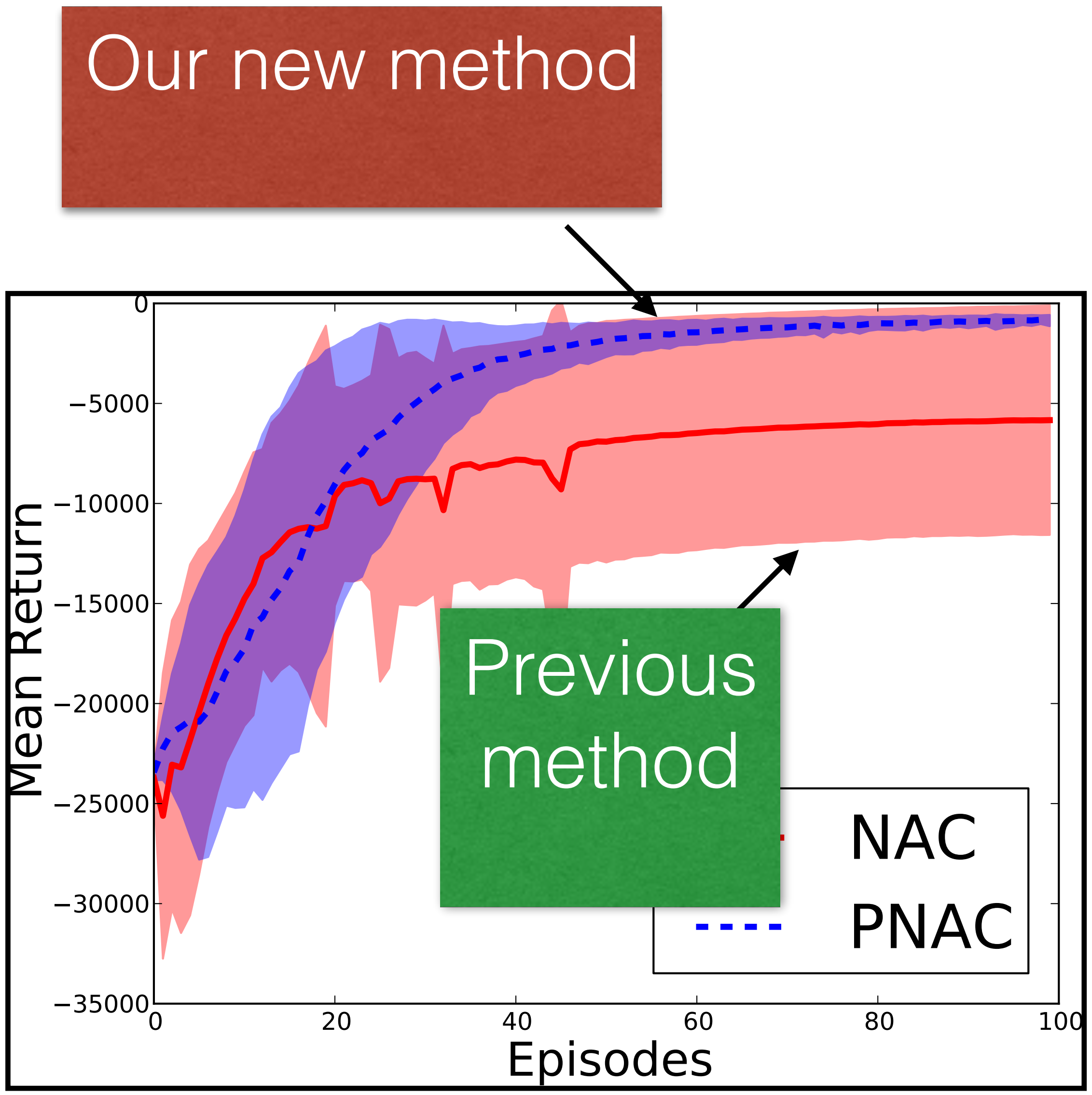
Sutton
et al., 2009

Our new
method

Safe Robot Learning



UBot, Laboratory of Perceptual Robotics



Thomas, Dabney, Mahadevan, Giguere, NIPS 2013

Mirror Descent = “Natural” Gradient
(Nemirovsky and Yudin; Amari, 1980s)

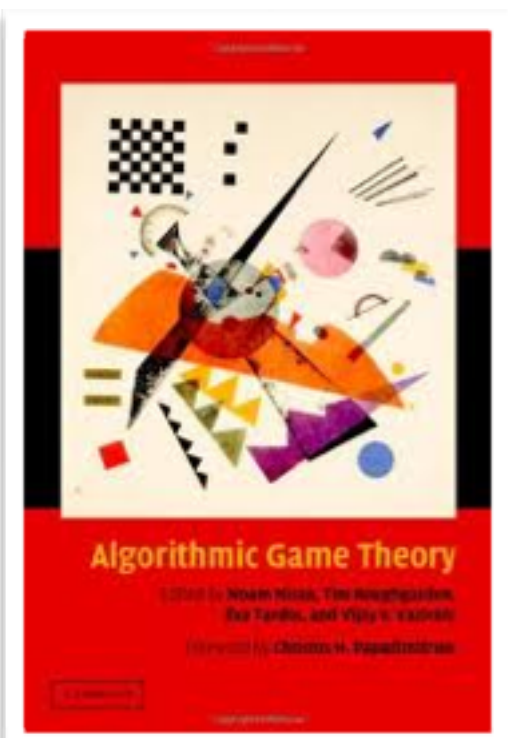
Thomas, Dabney, Mahadevan, Giguere, NIPS 2013

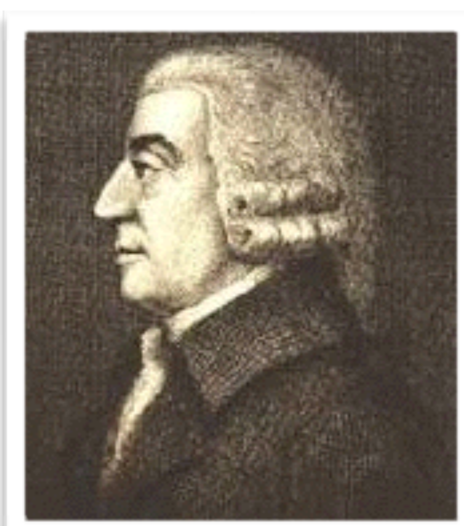
Part III: Equilibration Framework for ML, CS

“The Internet is an equilibrium — we just have to identify the game (Scott Shenker)”

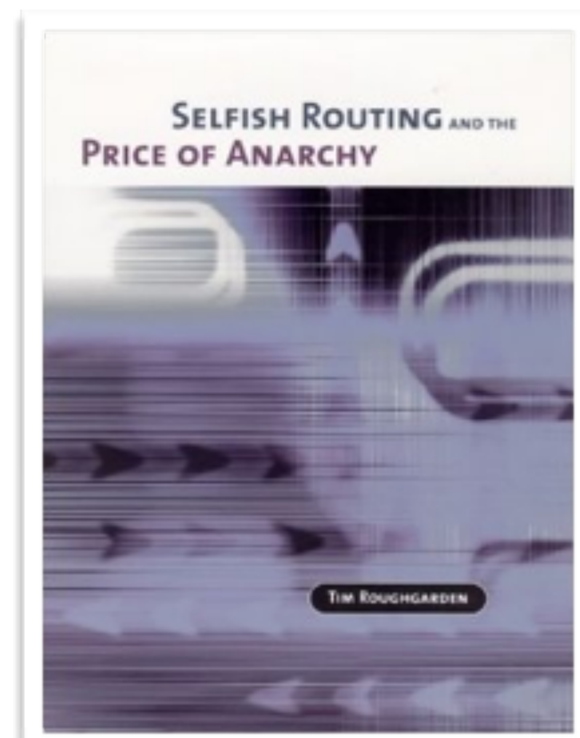
“The Internet was the first computational artifact that was not created by a single entity, but emerged from the strategic interaction of many (Christos Papadimitriou)”

The “Invisible Hand” of the Internet

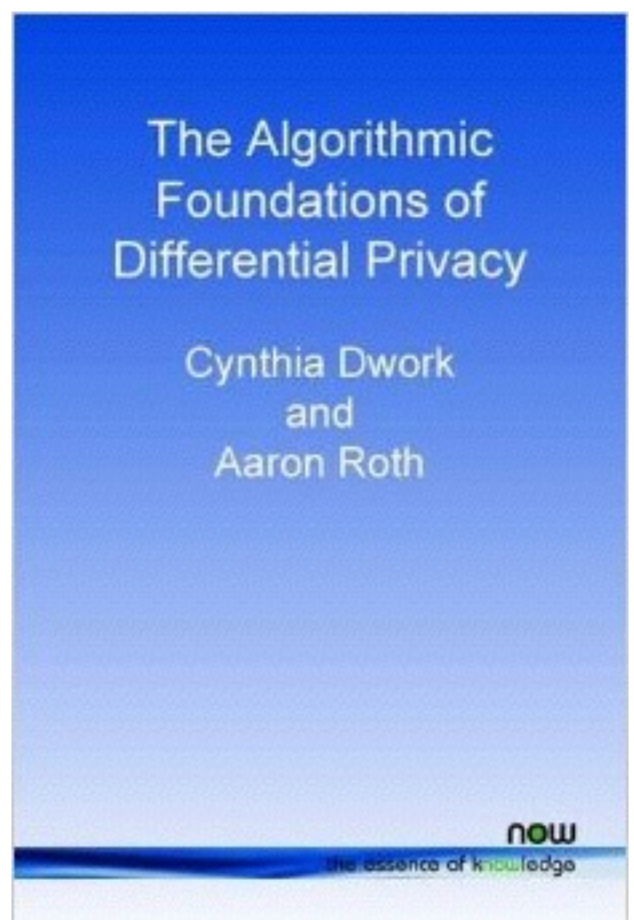


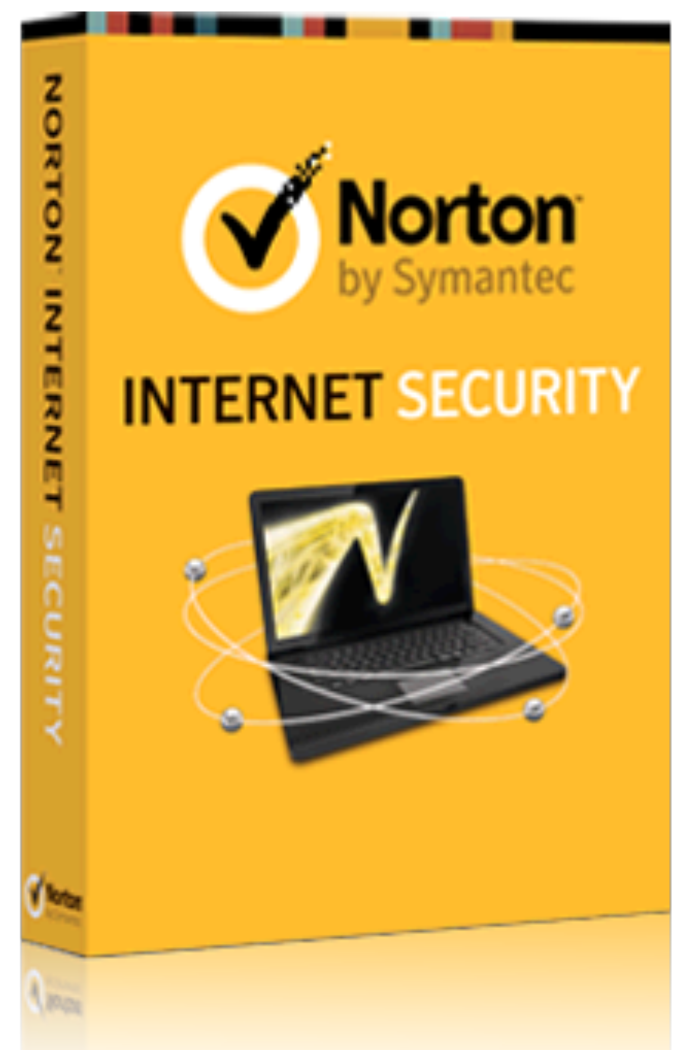


Adam Smith
The Wealth of Nations
1776



Competing Goals of the Internet: 1992-2014





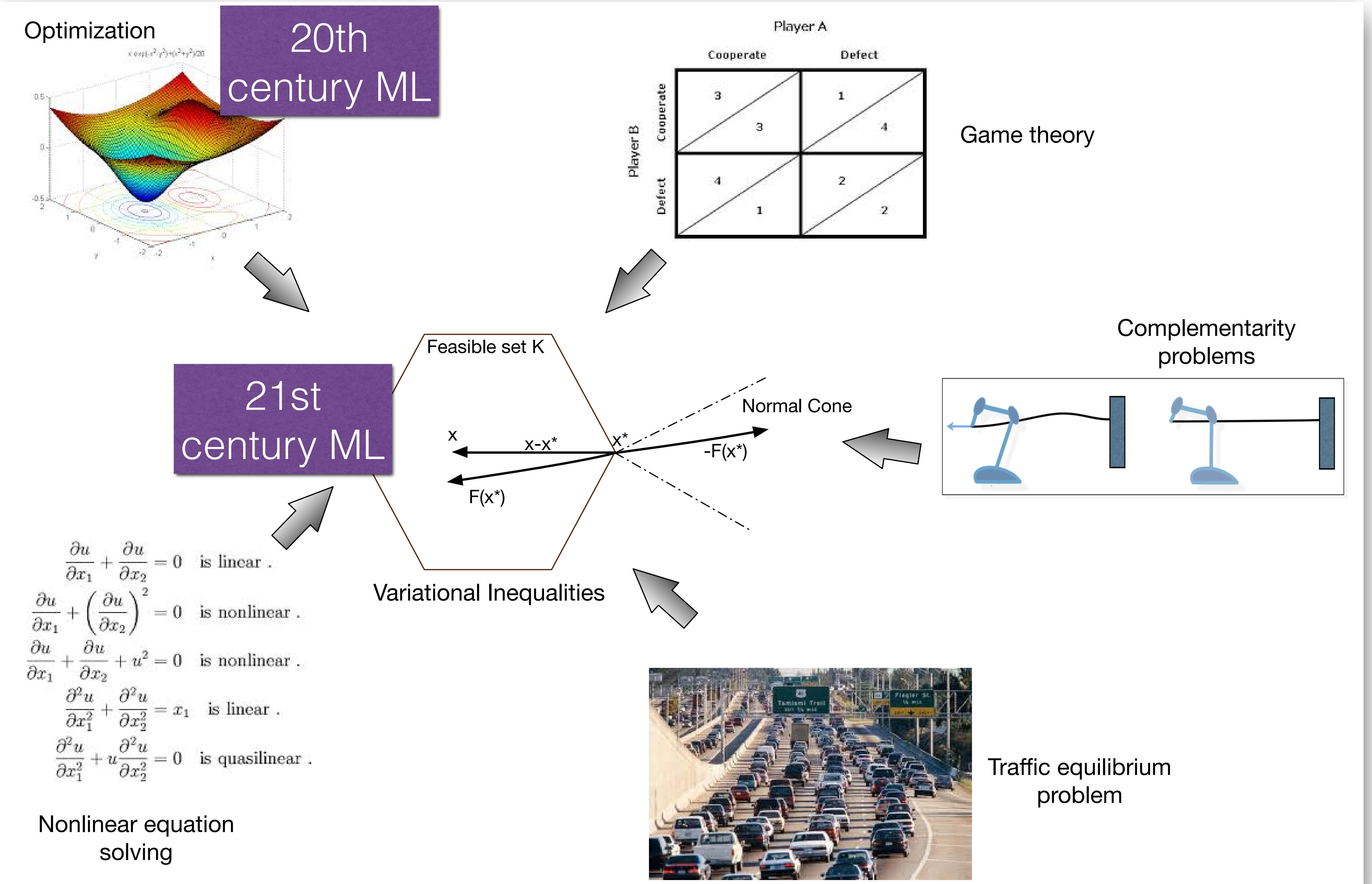
LA Times Story. June 05 2014

Verizon tells Netflix to stop blaming it for streaming issues

Verizon wants Netflix to stop blaming the Internet Service Provider for any streaming issues customers may be having when watching TV shows and movies.

"Netflix has been aware for some time that a few Internet middlemen have congestion issues with some IP Networks and nonetheless, Netflix has chosen to continue sending its traffic over those congested routes," said Verizon General Counsel Randal Milch. (Paul Sakuma / AP)

Part IV: New Algorithms, Applications, Results

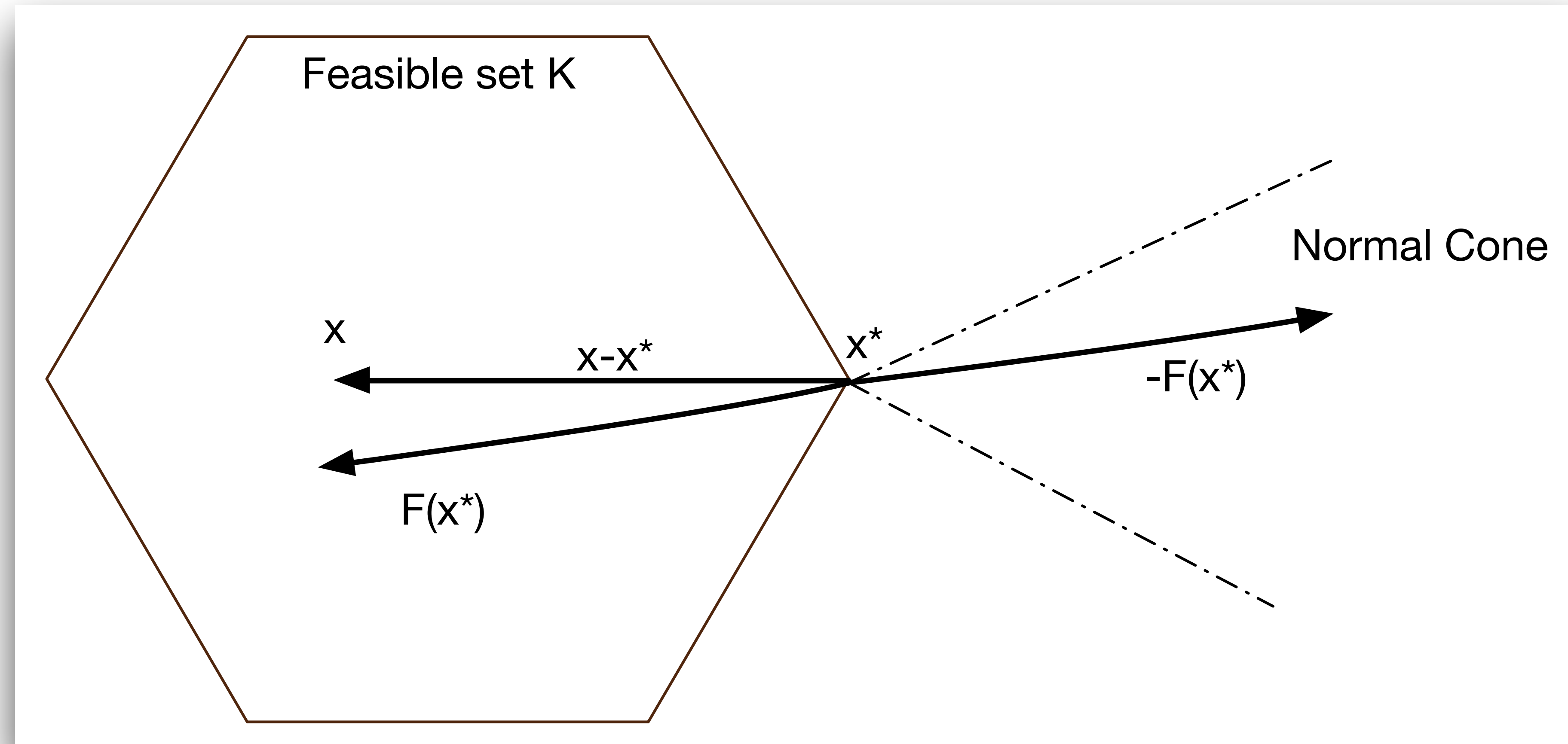

Optimization
20th century ML
Player A
Cooperate
Defect
Player B
Cooperate
Defect
3
3
1
4
4
1
2
2
Game theory
Feasible set K
21st century ML
x
x-x*
x*
Normal Cone
-F(x*)
F(x*)
Complementarity problems
Variational Inequalities
$\frac{\partial u}{\partial x_1} + \frac{\partial u}{\partial x_2} = 0$ is linear .
$\frac{\partial u}{\partial x_1} + \left(\frac{\partial u}{\partial x_2}\right)^2 = 0$ is nonlinear .
$\frac{\partial u}{\partial x_1} + \frac{\partial u}{\partial x_2} + u^2 = 0$ is nonlinear .
$\frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} = x_1$ is linear .
$\frac{\partial^2 u}{\partial x_1^2} + u\frac{\partial^2 u}{\partial x_2^2} = 0$ is quasilinear .
Nonlinear equation solving
Traffic equilibrium problem

(Almost) Dimension-Free
optimization

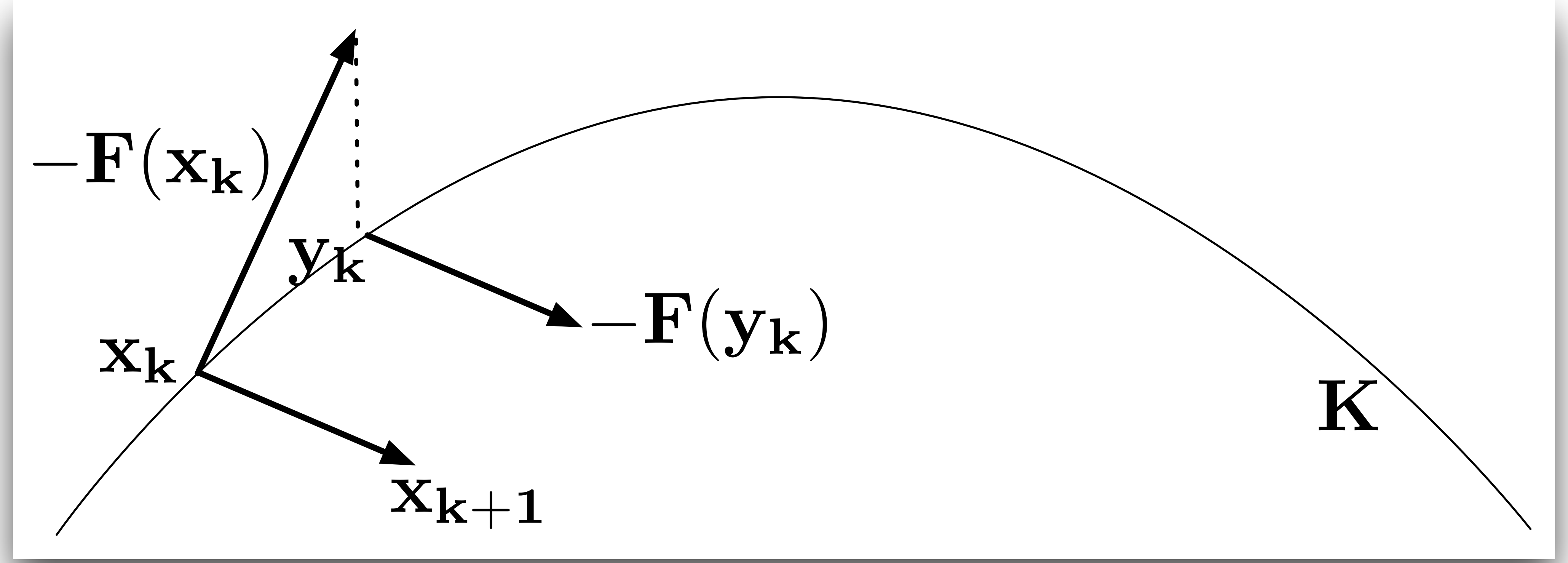
To

(Almost) Dimension-Free
equilibration

Fixed Point Formulation



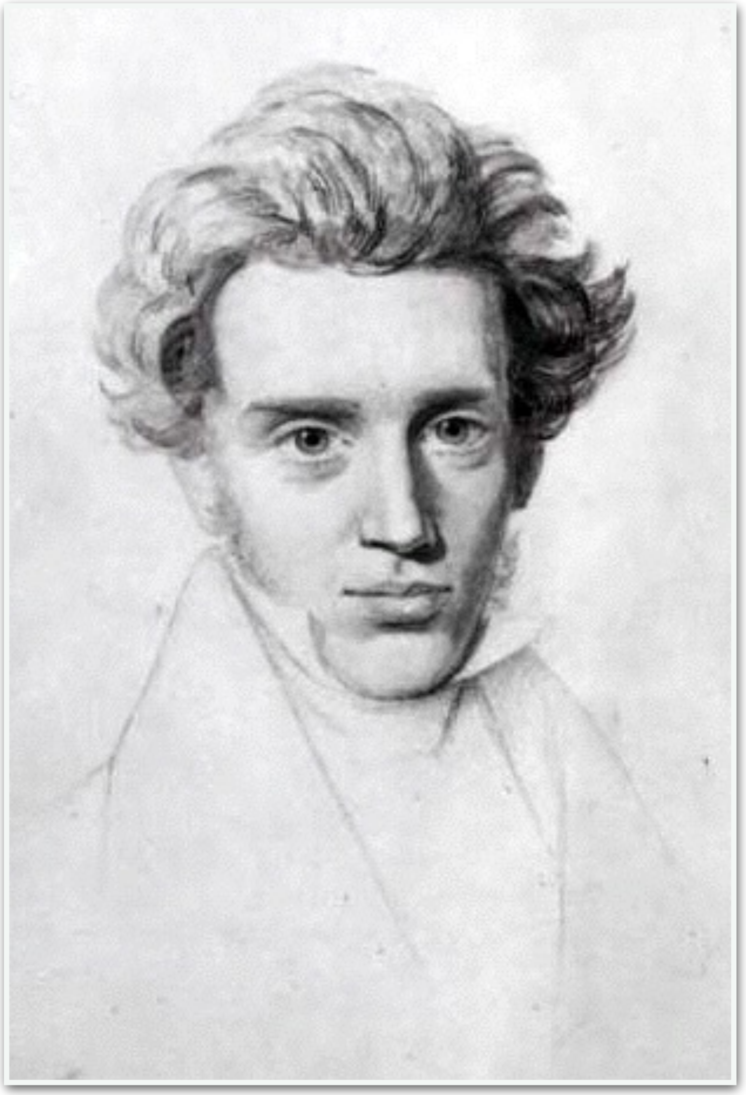
Extragradient Method



Korpolevich (1970s) developed the extragradient method

Søren Kierkegaard

Revisited



Life must be lived going forwards

many times in the Dual Space!

but can only be understood going backwards

many times in the Dual Space!

Our Latest VI Methods

(Gemp and Mahadevan, 2014)

General Runge-Kutta Mirror Descent (RKMDA)

$$
\begin{aligned}
k_1 &= \alpha_k F(x_k) \\
k_2 &= \alpha_k F(\nabla\psi_k^*(\nabla\psi_k(x_k) - a_{21}k_1)) \\
k_3 &= \alpha_k F(\nabla\psi_k^*(\nabla\psi_k(x_k) - a_{31}k_1 - a_{32}k_2)) \\
&\vdots \\
k_s &= \alpha_k F(\nabla\psi_k^*(\nabla\psi_k(x_k) - a_{s1}k_1 - a_{s2}k_2 - \ldots - a_{s,s-1}k_{s-1}))
\end{aligned}
$$

$$
x_{k+1} = \nabla\psi_k^*(\nabla\psi_k(x_k) - \Sigma_{i=1}^s b_i k_i)
$$

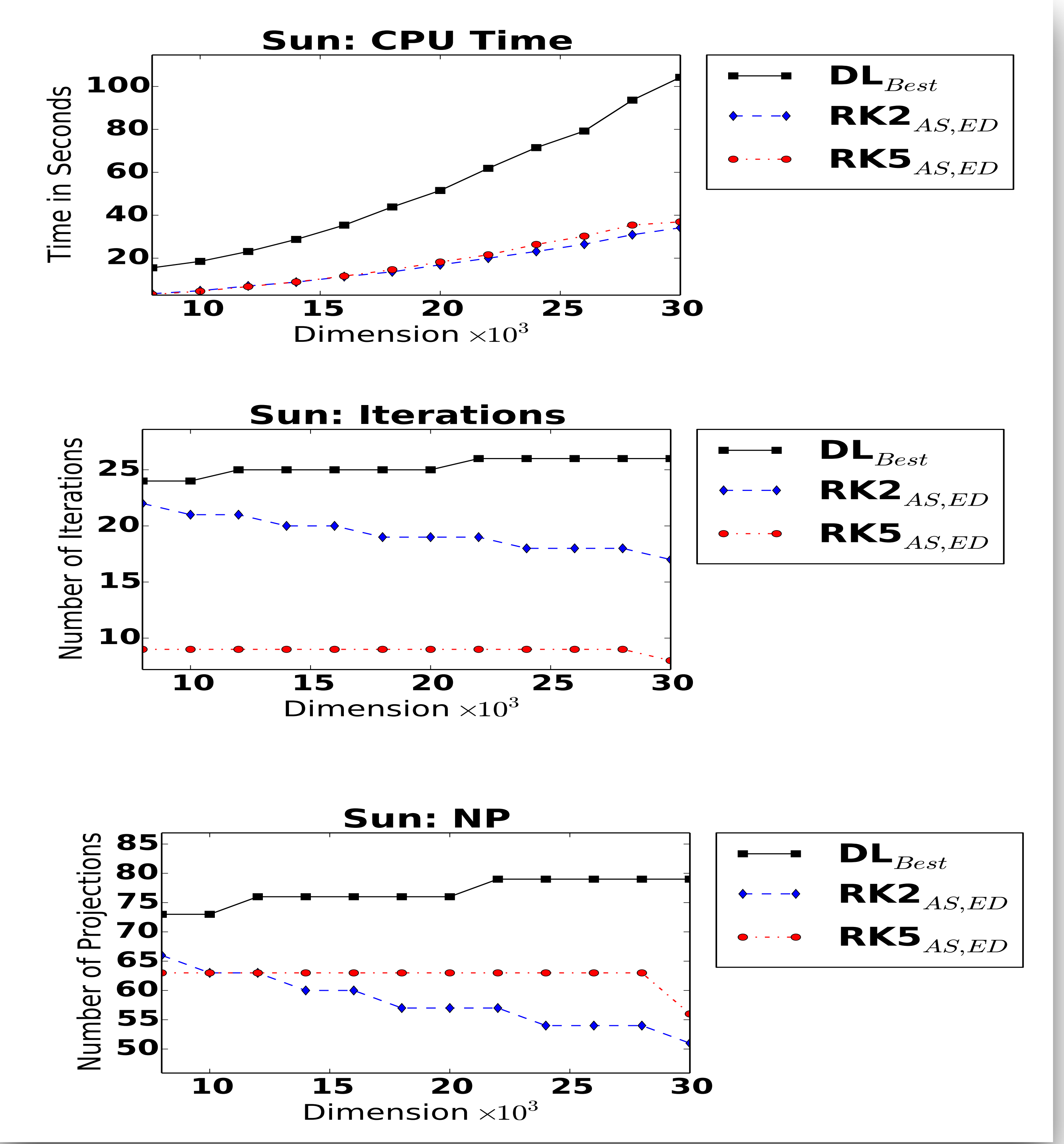
Benchmark VI Problem

Sun problem: This problem was proposed in [45]. The affine operator $F : \mathbb{R}^n \to \mathbb{R}^n$ is again given y $F(x) = Ax + b$, where

$$A = \begin{pmatrix} 1 & 2 & 2 & \dots & 2 \\ 0 & 1 & 2 & \dots & 2 \\ 0 & 0 & 1 & \dots & 2 \\ \vdots & \vdots & \vdots & \ddots & 2 \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

and $b = (-1, \dots, -1)^T$. The problem instances ranged from $n = 8000$ to $n = 30,000$.

Results on Benchmark VI



Next Generation Internet Model [Nagurney et al., 2014]

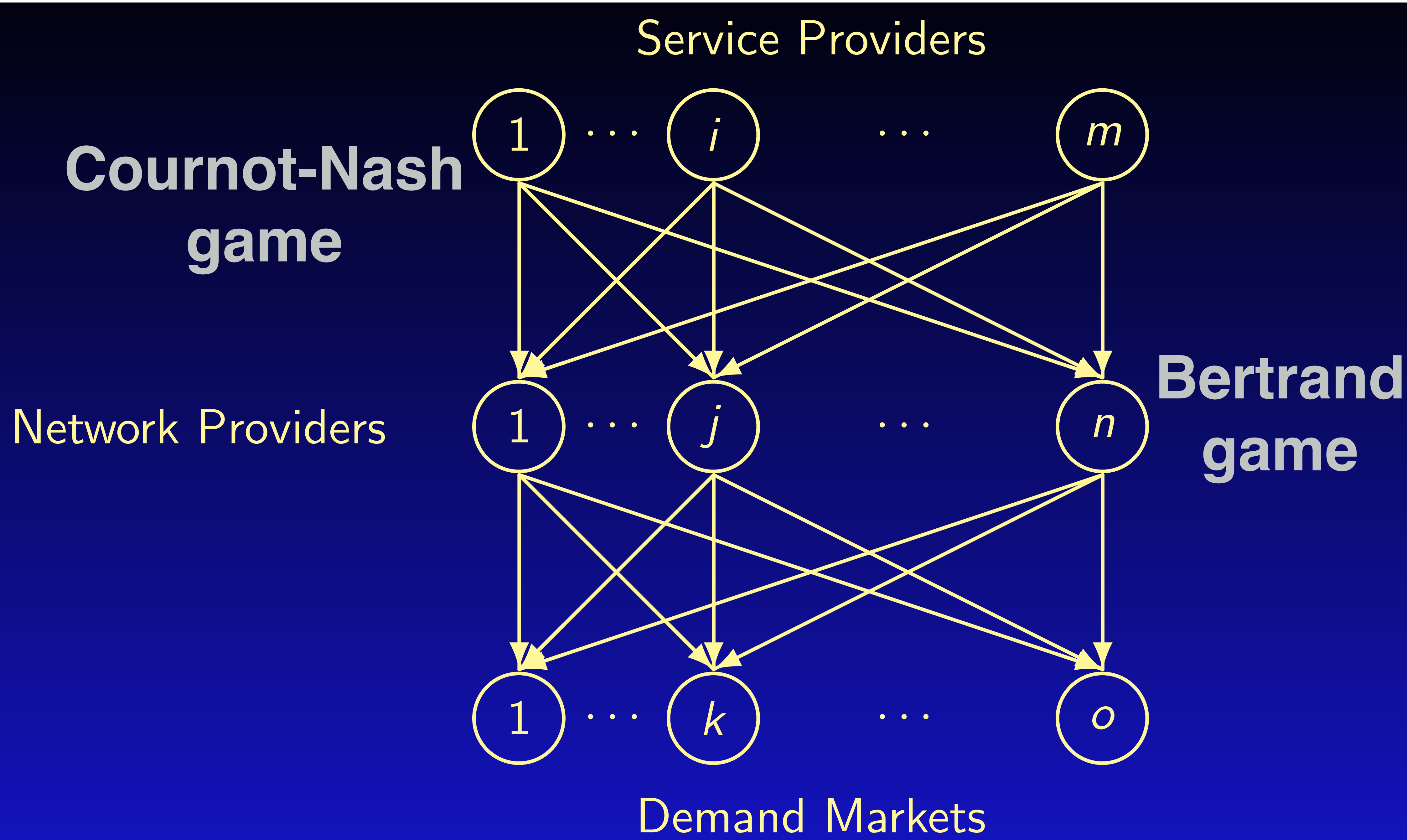


Figure 1: The Network Structure of the Cournot-Nash-Bertrand Model for a Service-Oriented Internet

Problem Formulation

Table 1: Notation for the Game Theoretic Cournot-Nash-Bertrand Model

Notation	Definition
Q_{ijk}	the nonnegative service volume from i to k via j. We group the $\{Q_{ijk}\}$ elements for all j and k into the vector $Q_i \in R^{no}_+$ and then we group all the vectors Q_i for all i into the vector $Q \in R^{mno}_+$.
s_i	the service volume (output) produced by service provider i. We group the $\{s_i\}$ elements into the vector $s \in R^m_+$.
q_{ijk}	the nonnegative quality level of network provider j transporting service i to k. We group the q_{ijk} for all i and k into the vector $q_j \in R^{mo}_+$ and all the vectors q_j for all j into the vector $q \in R^{mno}_+$.
π_{ijk}	the price charged by network provider j for transporting a unit of service provided by i via j to k. We group the π_{ijk} for all i and k into the vector $\pi_j \in R^m_+$ and then we group all the vectors π_j for all j into the vector $\pi \in R^{mno}_+$.
$f_i(s)$	the total production cost of service provider i.
$\hat{\rho}_{ijk}(Q, q)$	the demand price at k associated with service i transported via j.
$c_{ijk}(Q, q)$	the total transportation cost associated with delivering service i via j to k.
$oc_{ijk}(\pi_{ijk})$	the opportunity cost associated with pricing by network provider j services transported from i to k.

VI Fomulation

$$\langle F(X^*), X - X^* \rangle \geq 0, \quad \forall X \in \mathcal{K},$$

$$X \equiv (Q, q, \pi)$$

- Production cost function *f*(Q) - cost of providing a certain volume of content

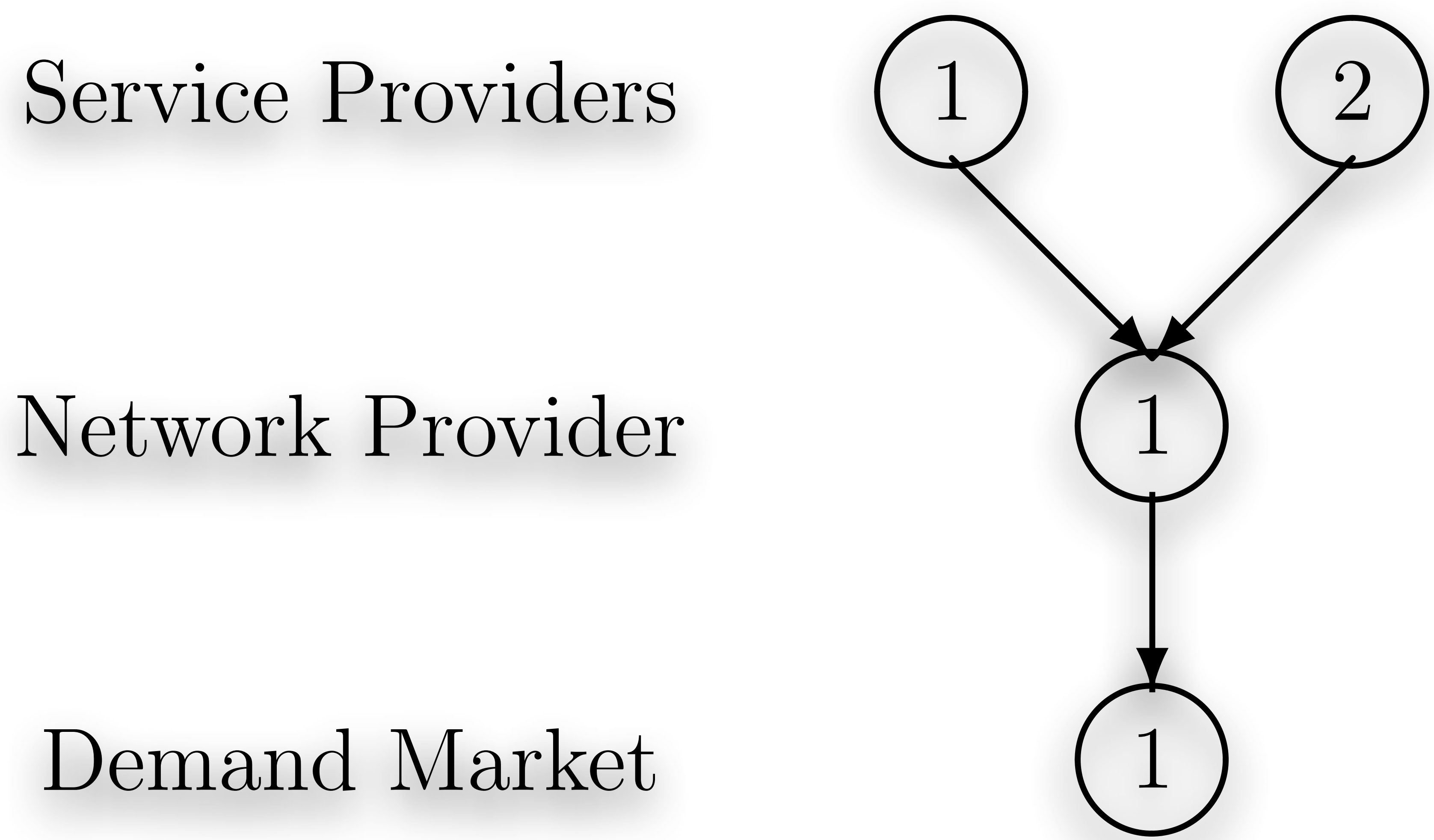
- Demand price function - user offer depends on content quality and market volume

$$F^1_{ijk}(X) = \frac{\partial \hat{f}_i(Q)}{\partial Q_{ijk}} + \pi_{ijk} - \hat{\rho}_{ijk}(Q,q) - \sum_{h=1}^{n} \sum_{l=1}^{o} \frac{\partial \hat{\rho}_{ihl}(Q,q)}{\partial Q_{ijk}} \times Q_{ihl},$$

$$F^2_{ijk}(X) = \sum_{h=1}^{m} \sum_{l=1}^{o} \frac{\partial c_{hjl}(Q,q)}{\partial q_{ijk}},$$

$$F^3_{ijk}(X) = -Q_{ijk} + \frac{\partial oc_{ijk}(\pi_{ijk})}{\partial \pi_{ijk}}.$$

Simple Example



Example

The production cost functions are:

$$\hat{f}_1(Q) = Q_{111}^2 + Q_{111}, \quad \hat{f}_2(Q) = 2Q_{211}^2 + Q_{211}.$$

The demand price functions are:

$$\hat{\rho}_{111}(Q,q) = -Q_{111} - .5Q_{211} + .5q_{111} + 100, \quad \hat{\rho}_{211}(Q,q) = -Q_{211} - .5Q_{111} + .5q_{211} + 200.$$

The transportation cost functions are:

$$\hat{c}_{111}(Q,q) = .5(q_{111} - 20)^2, \quad \hat{c}_{211}(Q,q) = .5(q_{211} - 10)^2,$$

with the opportunity cost functions being:

$$oc_{111}(\pi_{111}) = \pi_{111}^2, \quad oc_{211}(\pi_{211}) = \pi_{211}^2.$$

Example Results

$$F^1_{111}(X) = 2Q_{111} + 1 + \pi_{111} + Q_{111} + .5Q_{211} - .5q_{111} - 100 + Q_{111},$$

$$F^1_{211}(X) = 4Q_{211} + 1 + \pi_{211} + Q_{211} + .5Q_{111} - .5q_{211} - 200 + Q_{211},$$

$$F^2_{111}(X) = q_{111} - 20, \quad F^2_{211}(X) = q_{211} - 10,$$

$$F^3_{111}(X) = -Q_{111} + 2\pi_{111}, \quad F^3_{211}(X) = -Q_{211} + 2\pi_{211}.$$

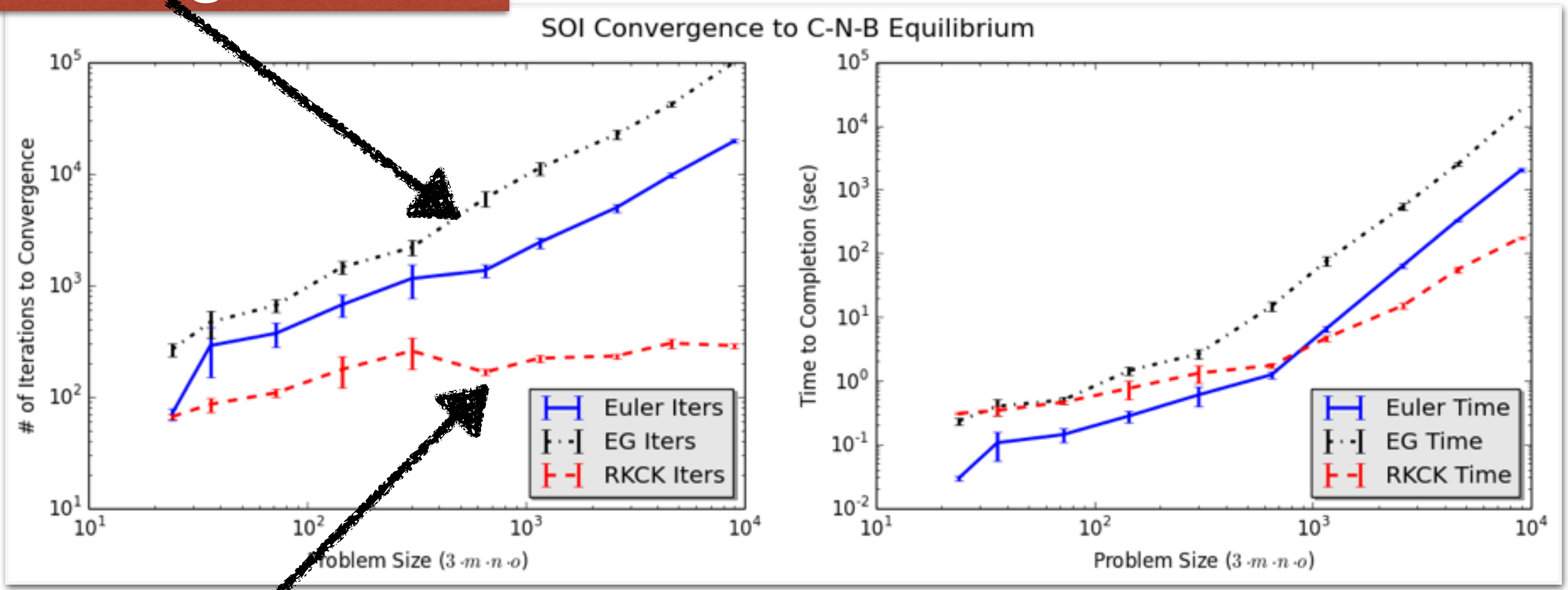
$$Q^*_{111} = 21.00, \quad Q^*_{211} = 30.00,$$

$$q^*_{111} = 20.00, \quad q^*_{211} = 10.00,$$

$$\pi^*_{111} = 10.50, \quad \pi^*_{211} = 15.00.$$

Results on Internet VI Problem

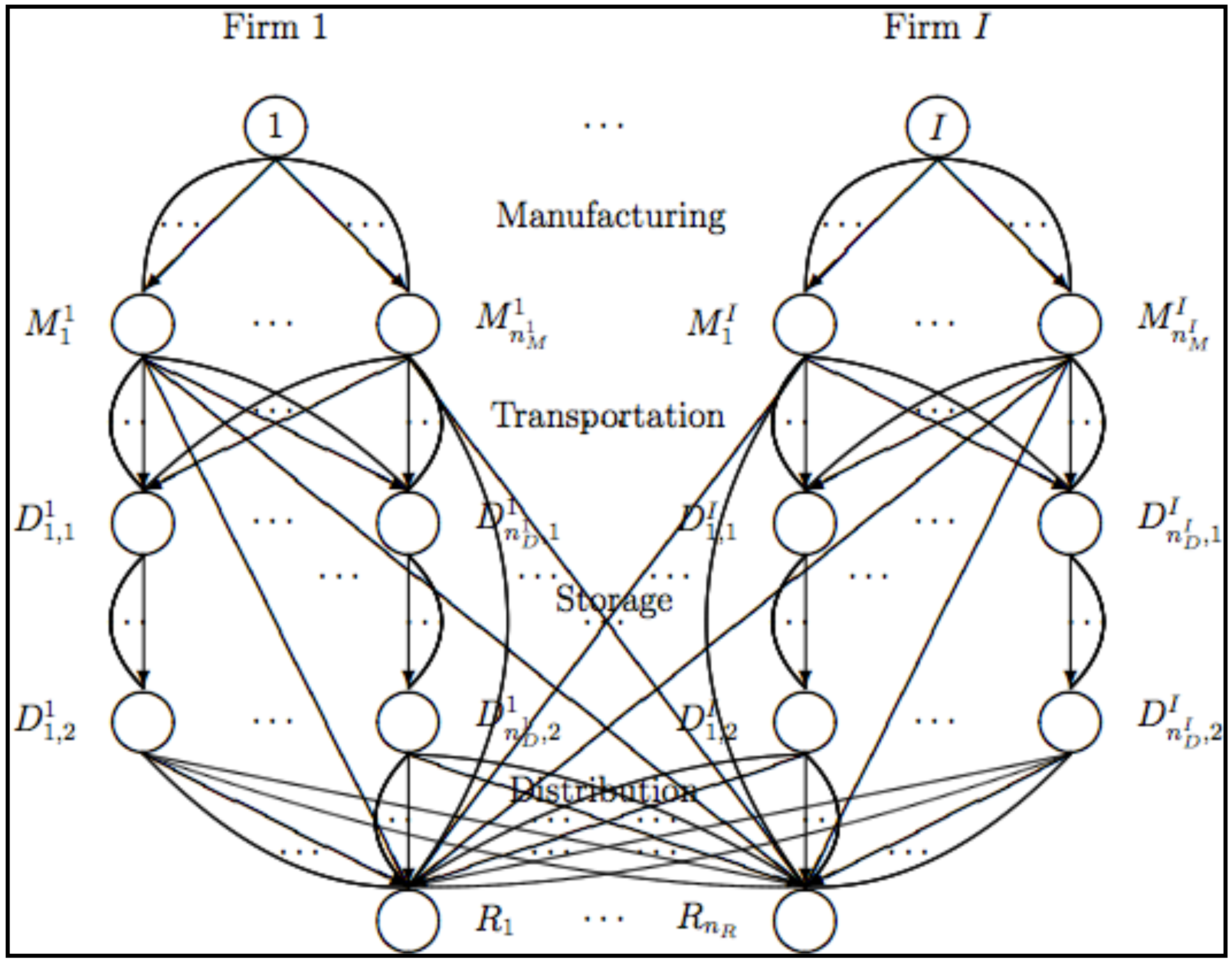
Extragradient



Our new algorithm

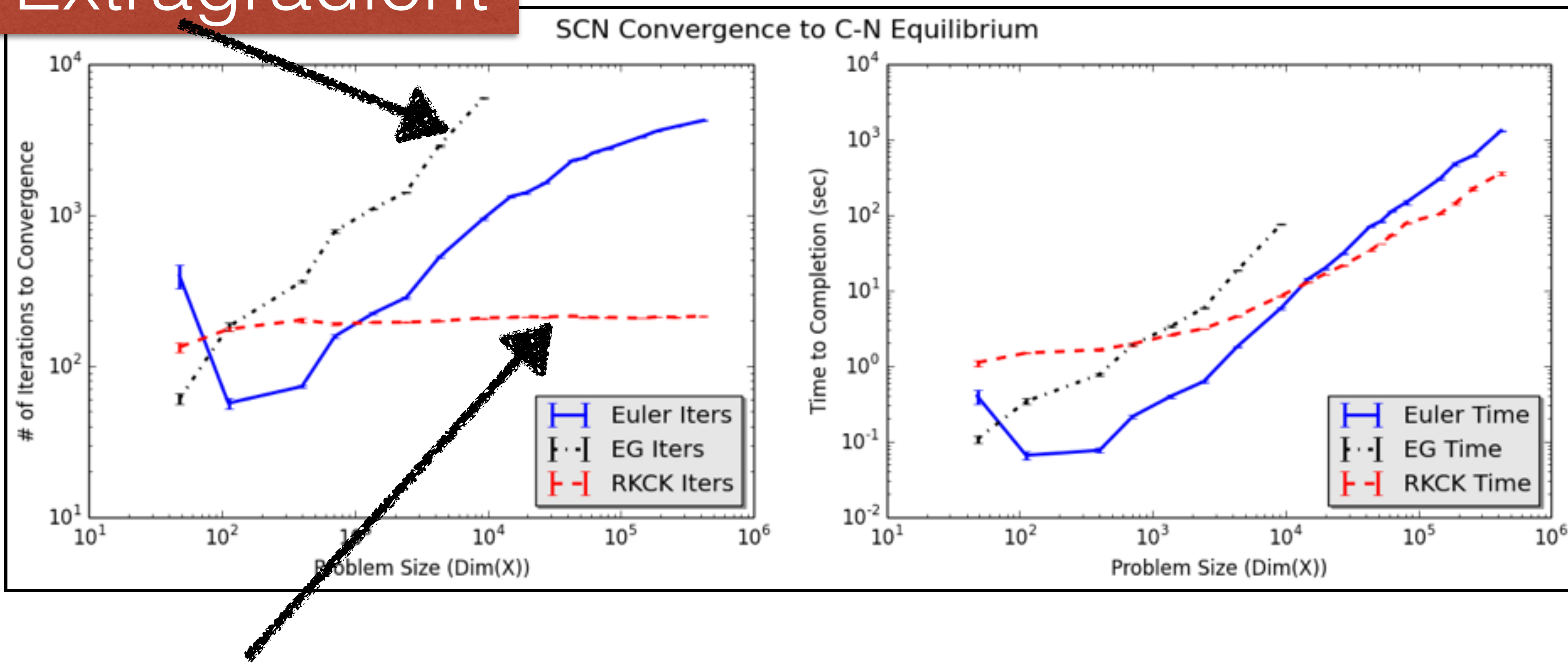
Sustainable Supply Chain

Nagurney et al.

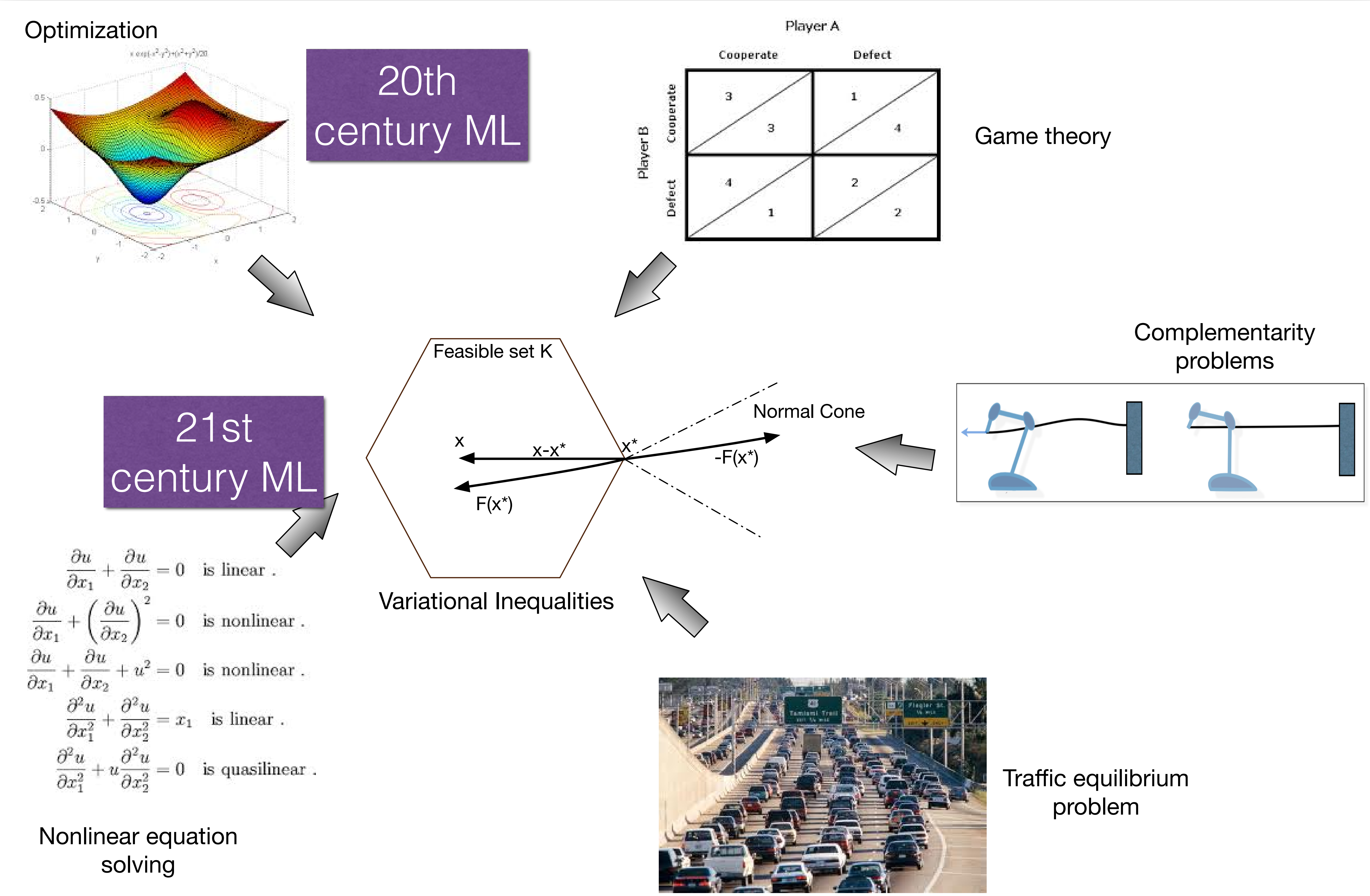


Results on Sustainable Supply Chain VI Problem

Extragradient



Our new algorithm


Optimization
20th century ML
Player A
Cooperate
Defect
Player B
Cooperate
Defect
3 3
1 4
4 1
2 2
Game theory
Complementarity problems
Feasible set K
x
x-x*
x*
Normal Cone
-F(x*)
F(x*)
21st century ML
$\frac{\partial u}{\partial x_1} + \frac{\partial u}{\partial x_2} = 0$ is linear .
$\frac{\partial u}{\partial x_1} + \left(\frac{\partial u}{\partial x_2}\right)^2 = 0$ is nonlinear .
$\frac{\partial u}{\partial x_1} + \frac{\partial u}{\partial x_2} + u^2 = 0$ is nonlinear .
$\frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} = x_1$ is linear .
$\frac{\partial^2 u}{\partial x_1^2} + u\frac{\partial^2 u}{\partial x_2^2} = 0$ is quasilinear .
Variational Inequalities
Nonlinear equation solving
Traffic equilibrium problem

Questions?