A Universal Framework for Sequential Decision Problems with Applications in Finance

Program in Digital Finance
University of Twente

February 4, 2024

Warren B Powell

Chief Innovation Officer, Optimal Dynamics

Professor Emeritus, Princeton University

Executive-in-Residence, Rutgers University


NASDAQ
NASDAQ COMPOSITE
S&P 500 Index
9533
AIR PRODUCTS
tell me more
www.airproducts.com
COVID 19
Coronavirus Vaccine
cargolux
E-COMMERCE

CHALLENGES

Virtually every problem in the domain of human processes combines decisions and uncertainty

- Engineering
- Physical science
- Economics
- Public health
- Health systems
- Laboratory sciences
- Energy systems
- Finance
- Manufacturing
- Personal transportation
- Supply chains
- E-commerce
- Freight transportation

Financial applications

Portfolio balancing

Managing cash balances

Pricing options

Tuning trading policies

Designing hedging contracts

Liquidating assets

Pricing for e-commerce

Supply chain finance

Lines of credit

Currency hedges

GOALS & OBJECTIVES

- Increase profits
- Reduce costs
- Increase yield
- Maximize productivity
- Minimize waste
- Maximize revenue
- Increase strength
- Reduce risk
- Reduce carbon footprint
- Reduce illness

Improving performance

If you want to run a better you have to make better decisions.

- What decisions do you have to make?
- How do we make better decisions?

OUTLINE

- ➜The seven levels of artificial intelligence
- ➜The universal modeling framework
- ➜Designing policies
- ➜Mutual fund cash balance optimization
- ➜Choosing the best policy
- ➜A new educational field: sequential decision analytics

OUTLINE

- ➜The seven levels of artificial intelligence
- ➜The universal modeling framework
- ➜Designing policies
- ➜Mutual fund cash balance optimization
- ➜Choosing the best policy
- ➜A new educational field: sequential decision analytics

Seven levels of artificial intelligence

???	Level 7: Creativity, judgment, reasoning	Science fiction
Next Gen AI	Level 6: Sequential decision problems	Optimization
1990's AI	Level 5: Deterministic optimization	Optimization
2020's AI	Level 4: Large language models	Machine learning
2010's AI	Level 3: Pattern recognition	Machine learning
1980's AI	Level 2: Basic machine learning	Machine learning
1970's AI	Level 1: Rule-based logic	Simple behaviors

https://tinyurl.com/7levelsofAI/

Seven levels of artificial intelligence

???	Level 7: Creativity, judgment, reasoning	Science fiction
Next Gen AI	Level 6: Sequential decision problems	Making computers smarter than humans
1990’s AI	Level 5: Deterministic optimization	
2020’s AI	Level 4: Large language models	Making computers behave like humans
2010’s AI	Level 3: Pattern recognition	
1980’s AI	Level 2: Basic machine learning	
1970’s AI	Level 1: Rule-based logic	

https://tinyurl.com/7levelsofAI/

Seven levels of artificial intelligence

???	Level 7: Creativity, judgment, reasoning	Science fiction
Next Gen AI	Level 6: Sequential decision problems	Optimization
1990's AI	Level 5: Deterministic optimization	Optimization
2020's AI	Level 4: Large language models	Machine learning
2010's AI	Level 3: Pattern recognition	Machine learning
1980's AI	Level 2: Basic machine learning	Machine learning
1970's AI	Level 1: Rule-based logic	Simple behaviors

https://tinyurl.com/7levelsofAI/

Seven levels of artificial intelligence

???	Level 7: Creativity, judgment, reasoning
Next Gen AI	Level 6: Sequential decision problems
1990's AI	Level 5: Deterministic optimization
2020's AI	Level 4: Large language models
2010's AI	Level 3: Pattern recognition
1980's AI	Level 2: Basic machine learning
1970's AI	Level 1: Rule-based logic

Science fiction

Optimization

Machine learning

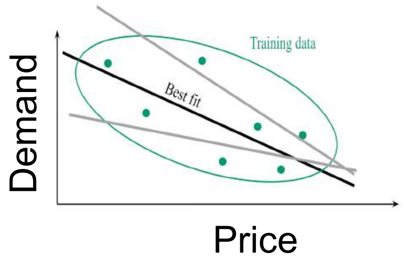
Simple behaviors

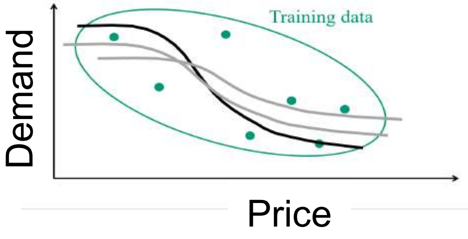
https://tinyurl.com/7levelsofAI/

Level 2

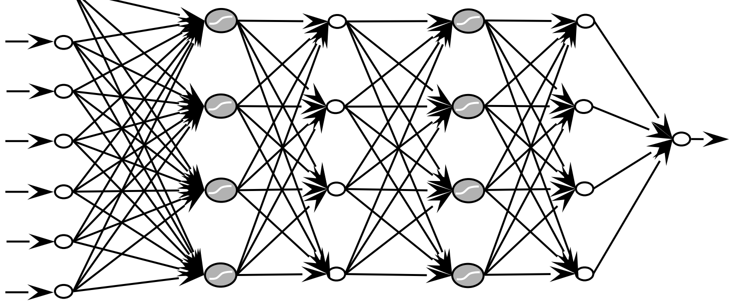
Basic machine learning (1900s)

- Sample applications
  - Estimation
  - Inference
  - Forecasting
- Major classes of methods
  - Parametric models
  - Time-series models
  - Early neural networks (10k-1m parameters)





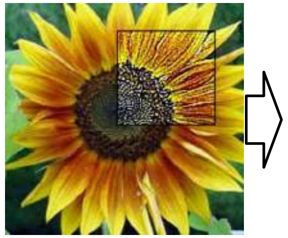
Neural network



Level 3

Pattern recognition (2010s)

- Image recognition
- Voice recognition
- Music identification
- Deep neural networks
  - 10m-100m parameters
  - First need for large labelled datasets



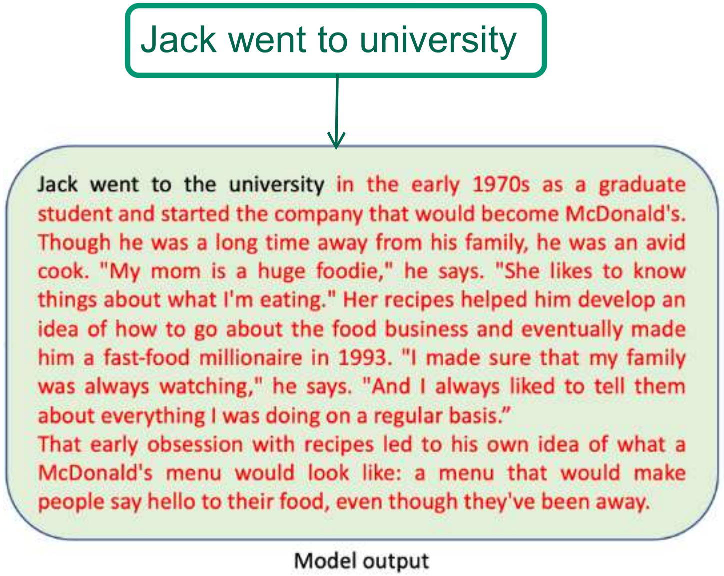
"sunflower"

"Take me home"

Level 4

Large language models (2020s)

- Text extension
  - From seed input to entire paragraphs or documents
- Natural language processing
- Colossal neural networks
  - 10 billion - 1 trillion parameters
  - Trained on massive text datasets



Optimal Dynamics

Seven levels of artificial intelligence

???	Level 7: Creativity, judgment, reasoning	Science fiction
Next Gen AI	Level 6: Sequential decision problems	Optimization
1990’s AI	Level 5: Deterministic optimization	
2020’s AI	Level 4: Large language models	Machine learning
2010’s AI	Level 3: Pattern recognition	
1980’s AI	Level 2: Basic machine learning	
1970’s AI	Level 1: Rule-based logic	Simple behaviors

https://tinyurl.com/7levelsofAI/


NASDAQ
NASDAQ COMPOSITE
S&P 500 Index
9533
AIR PRODUCTS
tell me more
www.airproducts.com
COVID 19
Coronavirus Vaccine
cargolux
E-COMMERCE

Decisions

What price to accept for a spot load?

Which load to accept now to move next week

Which driver should move a load?

What is the best policy for high-frequency trading?

What contracts to sign for raw materials?

Which vendor should supply each part?

What is the value of a financial optio

When should inventory be ordered?

Where should drivers be domiciled?

How many dedicated drivers should we have

How many syringes should be sent to each vaccination site, and when?

How much battery storage i needed?

What price should be charged

Which physician should handle a procedure?

How many nurses should we have to serve local hospitals?

When should I refill the customer’s tank with liquid nitrogen

When should gas turbines be scheduled?

Which nurse should visit this doctor’s office today?

Which customer tanks should we fill when we are in the area

Where should a patient be assigned for specific treatment?

How many suppliers should you have for a particular part, and where?

Which material handling jobs should be done by robots?

What bid should we place on Google for a set of ad-words?

How much energy should I purchase from the wind farm?

Which supplier should manufacture turbine blades?

How many jet engines should be made each day?

When should inventory be refilled at a fulfillment center?

Which fulfillment center should handle an order?

DECISIONS

Types of decisions.

Physical Decisions

» Managing inventories
» Scheduling nurses, assigning drivers and moving trucks
» Locating facilities and choosing suppliers

Financial Decisions

» Buying/selling assets
» Purchasing hedges
» Managing loans
» Pricing assets

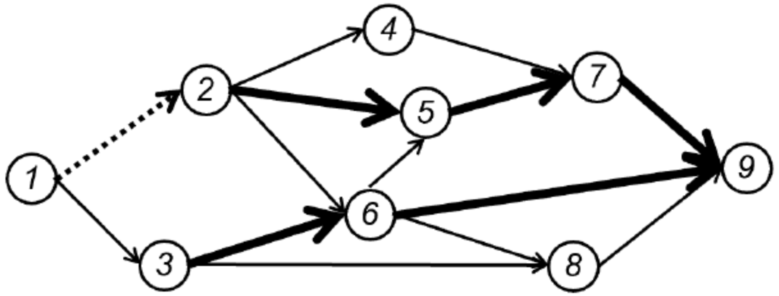
Informational Decisions

» Purchasing reports
» Testing new trading policies
» Trying new traders
» Investing in startups

DETERMINISTIC OPTIMIZATION

Low dimensional decisions

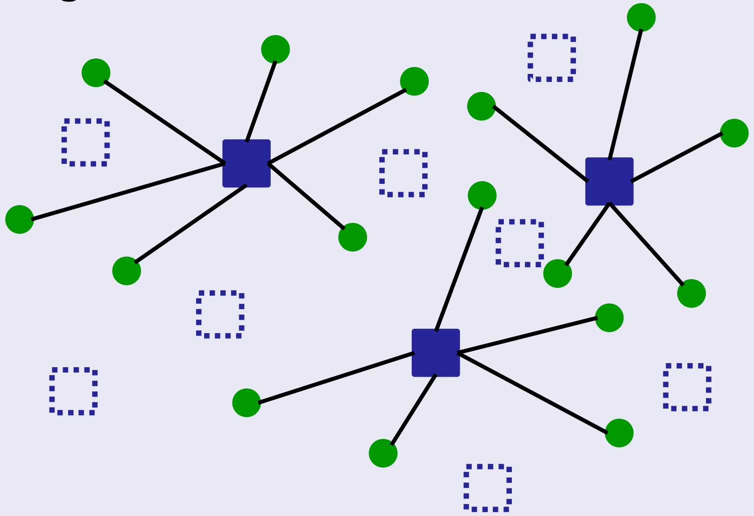
Planning a path to your destination



$$x_{ij} = \begin{cases} 1 & \textit{If we move from node i to node j} \\ 0 & \textit{Otherwise} \end{cases}$$

High dimensional decisions

Optimizing facility locations



$$x_i = \begin{cases} 1 & \textit{If we locate a facility at location i} \\ 0 & \textit{Otherwise} \end{cases}$$

DETERMINISTIC OPTIMIZATION

Airline scheduling

Airlines	Optimization Model	Airline Schedule

$$\min_x cx$$
$$Ax = b$$
$$x \geq 0$$

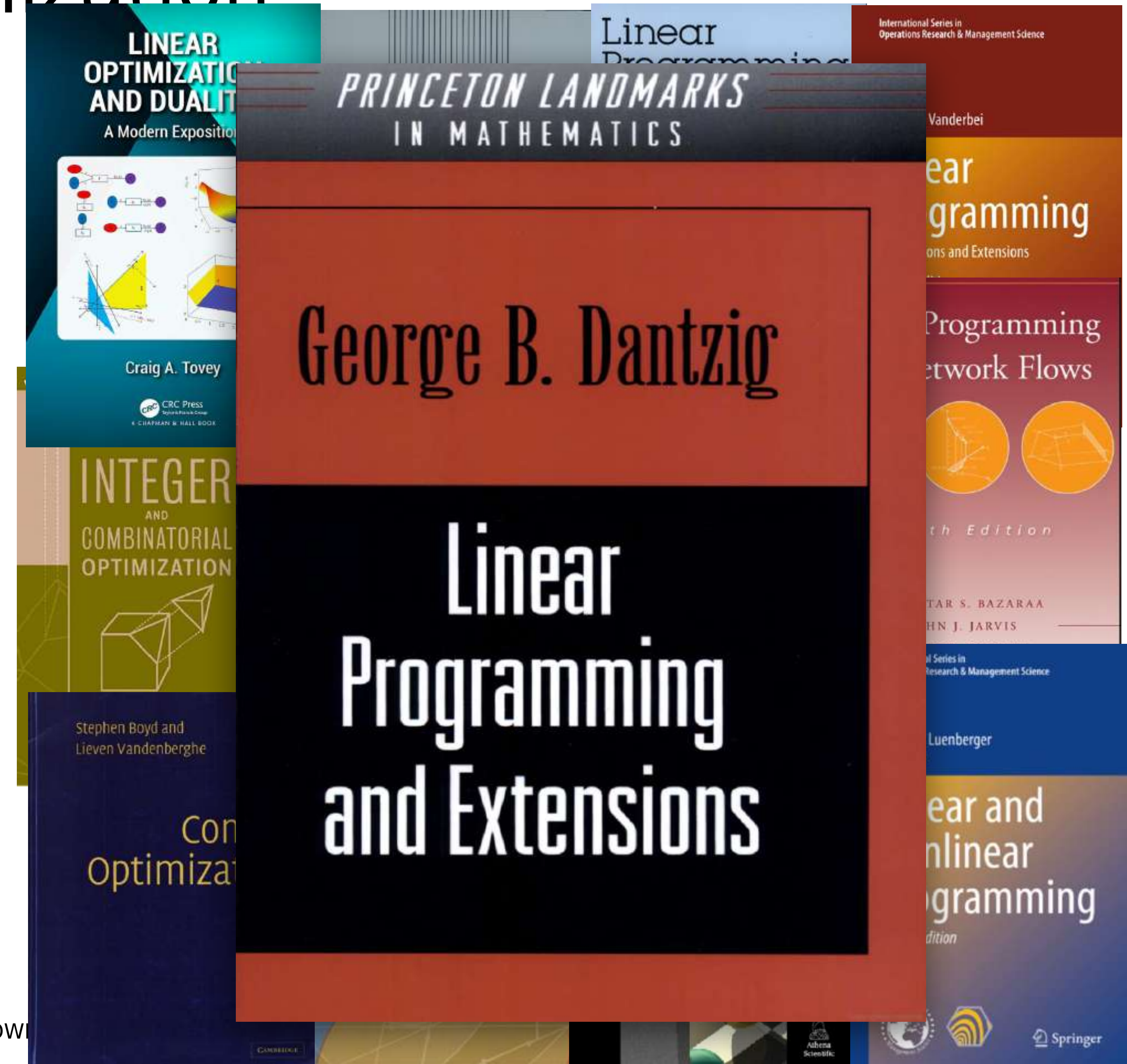
Airlines around the world use tools that depend on this mathematical model to perform strategic and operational planning.

Language of deterministic optimization

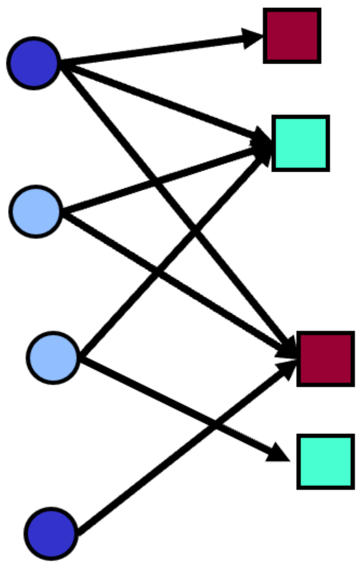
The language of deterministic optimization

$$\min_x cx$$
$$Ax = b$$
$$x \geq 0$$

- Spoken around the world.
- Many books communicate the same core theory
- Computer packages are available to solve realistic problems
- Many graduate programs producing thousands of students each year.



Deterministic optimization



- Standard modeling strategy is to pose the optimization problem at a point in time.
- The next step is to design some algorithm (e.g. simplex) to solve the decision problem.
- We then declare that the solution is *optimal!*

The language of classical optimization does not capture anything related to the evolution of the system over time.


Decisions


Linear programs

Seven levels of artificial intelligence

???	Level 7: Creativity, judgment, reasoning	Science fiction
Next Gen AI	Level 6: Sequential decision problems	Optimization
1990's AI	Level 5: Deterministic optimization	
2020's AI	Level 4: Large language models	Machine learning
2010's AI	Level 3: Pattern recognition	
1980's AI	Level 2: Basic machine learning	
1970's AI	Level 1: Rule-based logic	Simple behaviors

https://tinyurl.com/7levelsofAI/


NASDAQ
NASDAQ COMPOSITE
S&P 500 Index
9533
AIR PRODUCTS
tell me more
www.airproducts.com
COVID 19
Coronavirus Vaccine
cargolux
E-COMMERCE

Decisions

What price to accept for a spot load?

Which driver should move a load?

What is the best policy for high-frequency trading?

What contracts to sign for raw materials?

Which load to accept now to move next week

Which vendor should supply each part?

What is the value of a financial optio

Where should drivers be domiciled?

How many dedicated drivers should we have

How many syringes should be sent to each vaccination site, and when?

When should inventory be ordered?

How much battery storage i needed?

What price should be charged

Which physician should handle a procedure?

How many nurses should we have to serve local hospitals?

When should I refill the customer's tank with liquid nitrogen

When should gas turbines be scheduled?

Which nurse should visit this doctor's office today?

Which customer tanks should we fill when we are in the area

Where should a patient be assigned for specific treatment?

How many suppliers should you have for a particular part, and where?

Which material handling jobs should be done by robots?

What bid should we place on Google for a set of ad-words?

How much energy should I purchase from the wind farm?

Which supplier should manufacture turbine blades?

How many jet engines should be made each day?

When should inventory be refilled at a fulfillment center?

Which fulfillment center should handle an order?

Information

Market prices for spot freight

Driver requests for loads

Changes in asset prices and trading volume

Prices of raw materials by region

Quality of orders provided by a vendor

Offered loads by shipper, by lane

Production delays in order fulfillme

New COVID-19 cases by county; willingness to be vaccinated.

Employment rate; unemployment filings

Driver applicatior for jobs by region

Transportation and port unloadinig delays

Wind generation from a win farm

Competitor prices and availability

Patient arrivals and symptoms

Requests for nurse from doctor’s offices and hospitals.

Customer usage rate of liquid nitrogen and tank level.

Electricity prices on the grid and generator failures

Equipment failures at customer nitrogen tanks

Number of nurses calling in sick

Availability of specialists to treat a condition

Capacity shutdowns at suppliers due to labor or political problems

Flow of different parts to each machining station

Whether a bid wins an ad-click auction

The amount of energy that is generated from wind.

Lead times required by each manufacturer

Daily production of new jet engines

Flow of orders for a product by region around the country

Orders for a product from different regions

SEQUENTIAL DECISIONS

In most settings, decisions are made over time…

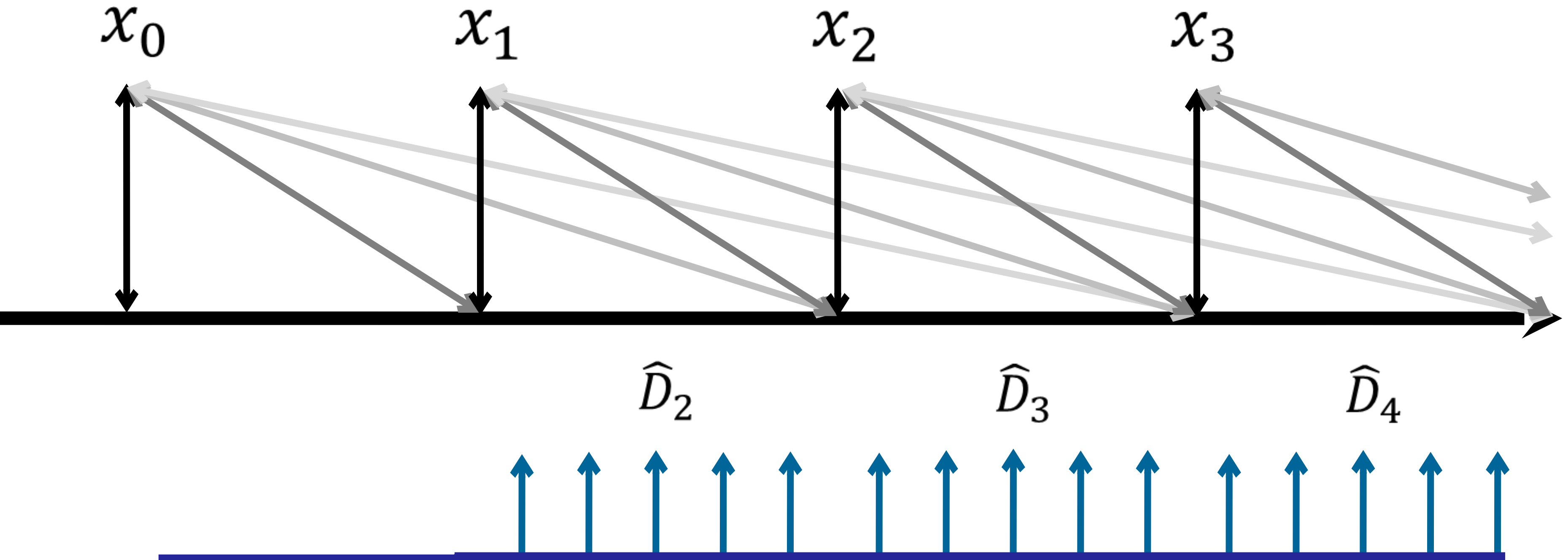
Decisions

Information that arrives after a decision is made is not known when we made the decision.

SEQUENTIAL DECISIONS

Cash management

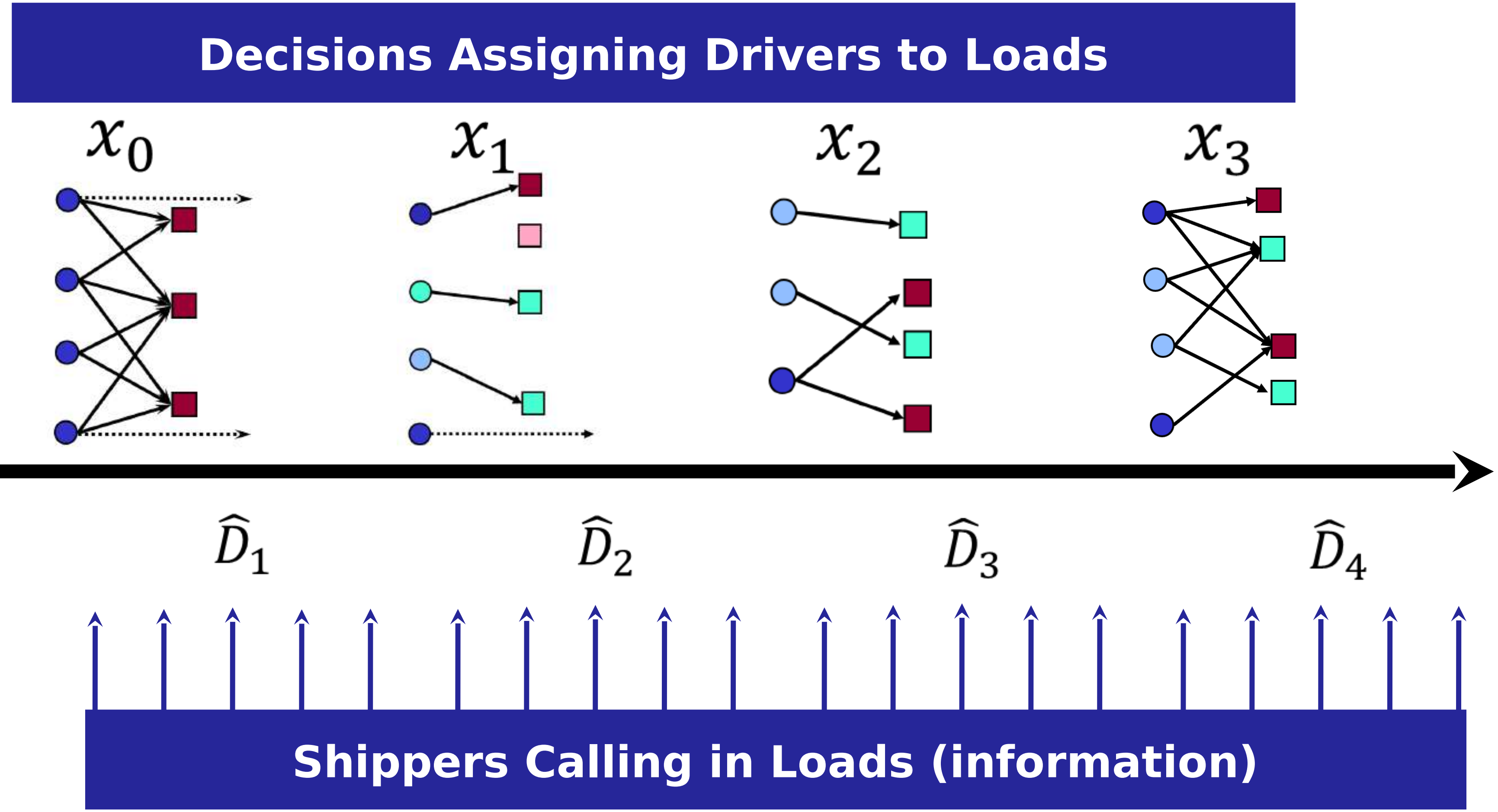
Cash inflows/outflows



Cash needs, market returns (information)

SEQUENTIAL DECISIONS

Driver dispatch for truckload trucking



SEQUENTIAL DECISIONS

Testing new vaccines

Vaccination Decisions (what dosage, which people)

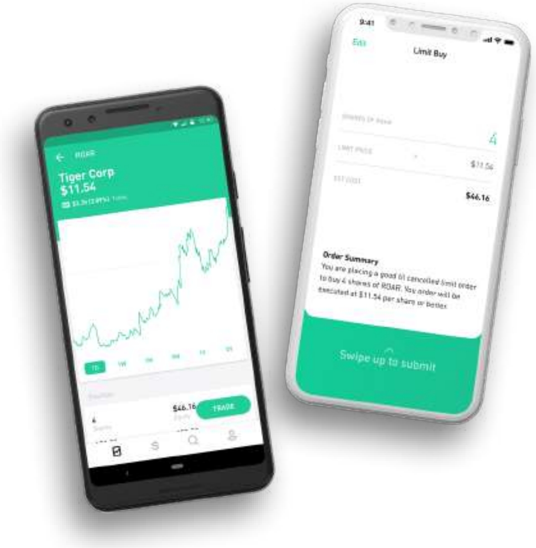
x_0 x_1 x_2 x_3

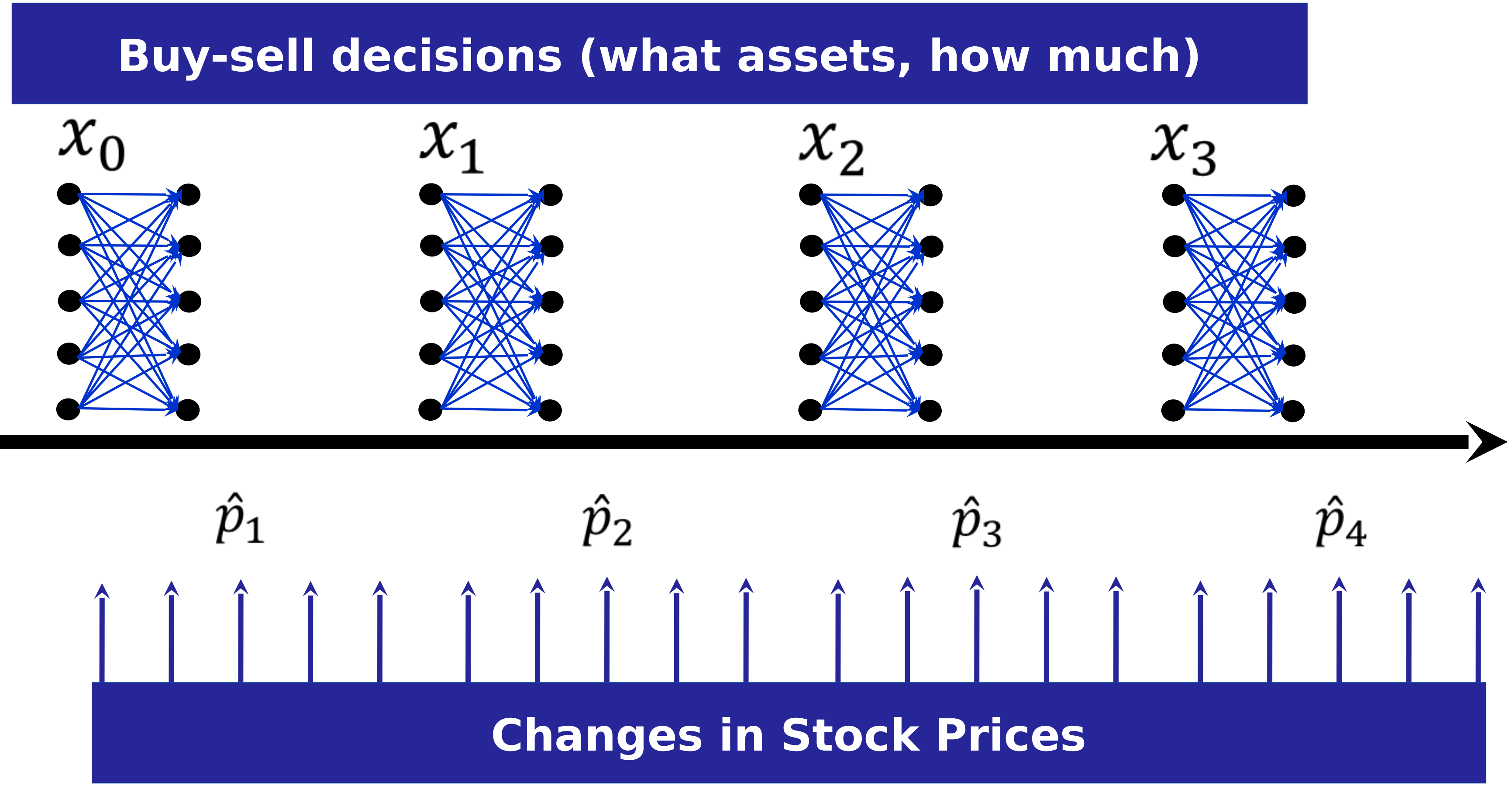
$\hat{I}_1$ $\hat{I}_2$ $\hat{I}_3$ $\hat{I}_3$

Patient Outcomes (information)

SEQUENTIAL DECISIONS

Financial Trading





OUTLINE

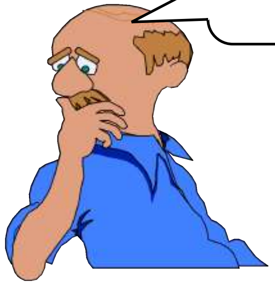
- The seven levels of artificial intelligence
- The universal modeling framework
- Designing policies
- Mutual fund cash balance optimization
- Choosing the best policy
- A new educational field: sequential decision analytics

Modeling sequential decision problems

The biggest challenge when making decisions under uncertainty is ***modeling***.

Everyone writing out a deterministic optimization model, or machine learning model, knows how to write out their problem mathematically…

$$\min E \left\{ \sum cx \right\}$$
$$Ax = b$$
$$x \geq 0$$



Mathematical model

Organize class libraries, and set up communications and databases

…we lack a standard modeling framework for sequential decisions.


Stochastic programming
Robust optimization
Decision analysis
Approximate dynamic programming
Simulation optimization
Dynamic Programming and control
Optimal learning
Model predictive control
Optimal control
Stochastic search
Bandit problems
Reinforcement learning
Markov decision processes
Stochastic optimization
Online computation
Stochastic control


John R. Birge
François Louveaux
Introduction to Stochastic Programming
Second Edition
Michael C. Fu Editor
Handbook of Simulation Optimization
Princeton Series in APPLIED MATHEMATICS
Robust Optimization
Introduction to Decision Analysis
A Practitioner's Guide to
SECOND EDITION
Approximate Dynamic Programming
Solving the Curses of Dimensionality
Warren B. Powell
VOLUME 2 • 4th EDITION
Dynamic Programming and Optimal Control
APPROXIMATE DYNAMIC PROGRAMMING
Dimitri P. Bertsekas
WILEY
INTRODUCTION TO STOCHASTIC SEARCH AND OPTIMIZATION
Estimation, Simulation, and Control
JAMES C. SPALL
Wiley Series in Probability and Statistics
Optimal Learning
Warren B. Powell
Ilya O. Ryzhov
MULTI-ARMED BANDIT ALLOCATION INDICES
SECOND EDITION
John Gittins, Kevin Glazebrook and Richard Weber
SECOND EDITION
Model Predictive Control
OPTIMAL CONTROL
Frank L. Lewis
Draguna L. Vrabie
Vassilis L. Syrmos
43
Jiongmin Yong
Xun Yu Zhou
Stochastic Controls
Hamiltonian Systems and HJB Equations
Modelling and Applied Probability
Reinforcement Learning
An Introduction
second edition
Richard S. Sutton and Andrew G. Barto
Markov Decision Processes
Discrete Stochastic Dynamic Programming
MARTIN L. PUTERMAN
LECTURES ON STOCHASTIC PROGRAMMING
Modeling and Theory
Alexander Shapiro
Darinka Dentcheva
Andrzej Ruszczyński
MOS-SIAM Series on Optimization
Online Computation and Competitive Analysis
Allan Borodin
Ran El-Yaniv

Modeling sequential decision problems

- Any sequential decision problem can be written:

What we know (or believe)

The decision

What we observe (or learn)

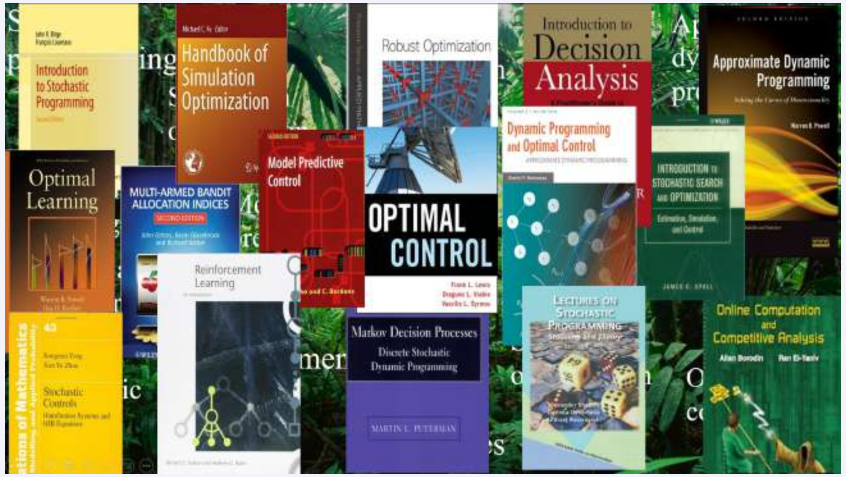
- Each time we make a decision, we receive a contribution
- Decisions are made with a method or *policy* which we design later.
- State variables evolve using a transition function: .
- The goal is to find the policy that maximizes expected contributions:

Optimal Dynamics

Modeling sequential decision problems

Every sequential decision problem can be modeled using 5 core components

1) State variables
   - Physical state , other information , beliefs .
2) Decision variables (or action , or control)
   - Decisions are determined by a policy
3) Exogenous variables
   - This is new information that arrives between and .
4) Transition function
   - This is how our state variable evolves given and .
5) Objective function for finding the best policy



This is the universal modeling framework

Optimal Dynamics

Modeling sequential decision problems

- Objective functions
  - » Cumulative reward ("online learning")

$$\max_\pi \mathbb{E}\left\{\sum_{t=0}^{T} C_t\left(S_t, X_t^\pi(S_t)\right) \vee S_0\right\}$$

  - » Final reward ("offline learning")

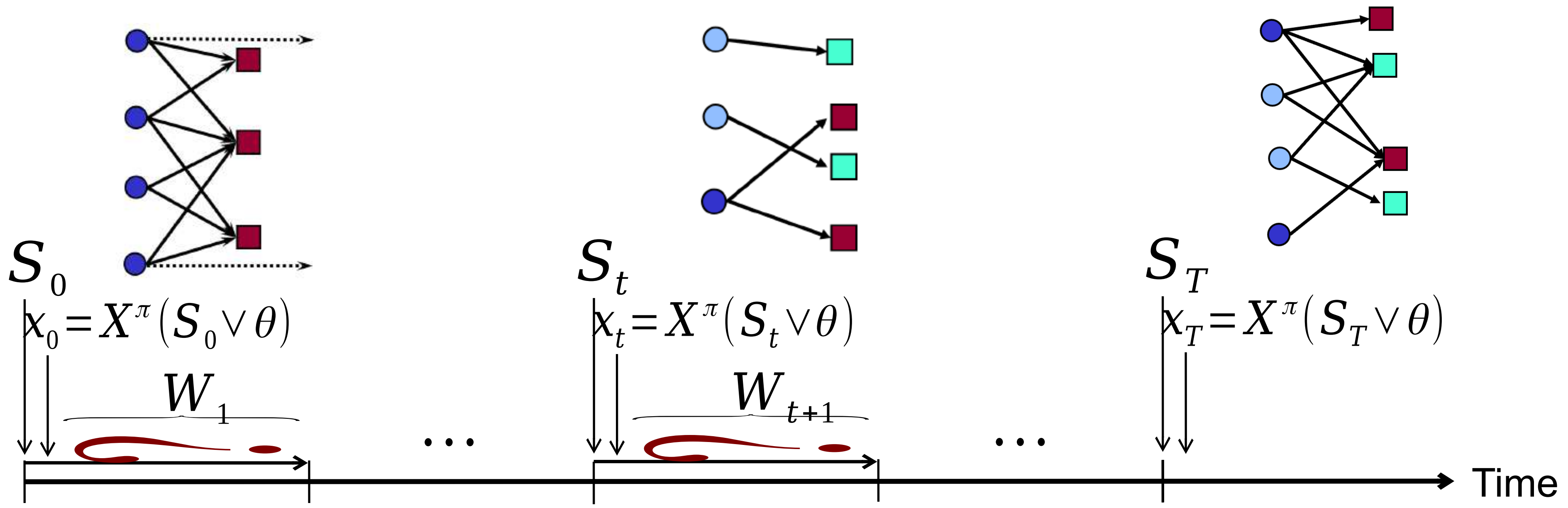
$$\max_\pi \mathbb{E}\left\{F\left(x^{\pi,N}, \widehat{W}\right) \vee S_0\right\}$$

  - » Risk

$$\max_\pi \underbrace{\mathbb{E}\left\{\sum_{t=0}^{T} C_t\left(S_t, X_t^\pi(S_t)\right) \vee S_0\right\}}_{\text{Base objective}} - \eta \underbrace{\rho\left\{C\left(S_0, X_0^\pi(S_0)\right), \ldots, C\left(S_T, X_T^\pi(S_T)\right) \vee S_0\right\}}_{\text{Risk metric}}$$

Deterministic optimization

Evaluating performance over time:



$$F^\pi(\theta) = C(S_0, x_0) + C(S_t, x_t) + \dots + C(S_T, x_T)$$

$$= \sum_{t=0}^{T} C\left(S_t, X^\pi(S_t \vee \theta)\right)$$ performance of policy over time

Modeling sequential decision problems

The objective function

We might average over simulations:

$$\overline{F}^{\pi} = \frac{1}{N}\sum_{n=1}^{N}\sum_{t=0}^{T} C_t\left(S_t\left(\omega^n\right), X_t^{\pi}\left(S_t\left(\omega^n\right)\middle|\theta\right)\right)$$

Modeling sequential decision problems

The objective function

.... which is an approximation of the expectation:

$$F^{\pi} = \mathbb{E}\left\{ \sum_{t=0}^{T} C_t\left(S_t, X_t^{\pi}\left(S_t \middle| \theta\right)\right) \vee S_0 \right\}$$

$$\overbrace{R_0 \, I_0 \, B_0}$$

State of the world

Investments
Cash
Loans
Notes
....

S&P 500
Inflation
Interest rates
Oil prices
....

Forecasts
Market values
Actual unemployment
Government policies
....

Optimal Dynamics

Modeling sequential decision problems

The objective function

.... which is an approximation of the expectation:

$$F^{\pi} = \mathbb{E}\left\{ \sum_{t=0}^{T} C_t\left(S_t, X_t^{\pi}\left(S_t \middle| \theta\right)\right) \vee S_0 \right\}$$

Evaluating policies

Theoretically
- Optimality proofs
- Asymptotic optimality
- Regret bounds

Computer simulation

Field testing

Modeling sequential decision problems

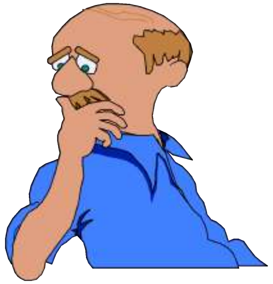
The objective function

We now want to find the best policy

$$\max_{\pi=(f\in F,\theta\in\Theta^f)} F^\pi = \mathbb{E}\left\{\sum_{t=0}^{T} C_t\left(S_t, X_t^\pi\left(S_t|\theta\right)\right) \vee S_0\right\}$$

Type of policy (structure of function)

Algorithmic tuning



$$\theta^{n+1} = \theta^n + \alpha_n \nabla_\theta F\left(\theta^n, W^{n+1}\right)$$

Optimal Dynamics

Modeling sequential decision problems

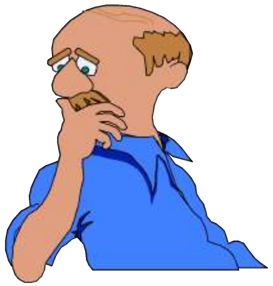
The objective function

We now want to find the best policy

$$\max_{\pi=(f\in F,\theta\in\Theta^f)} F^\pi = \mathbb{E}\left\{\sum_{t=0}^{T} C_t\left(S_t, X_t^\pi\left(S_t|\theta\right)\right)\vee S_0\right\}$$

Type of policy (structure of function)

Algorithmic tuning



$$\theta^{n+1}=\theta^n+\alpha_n\nabla_\theta F\left(\theta^n,W^{n+1}\right)$$

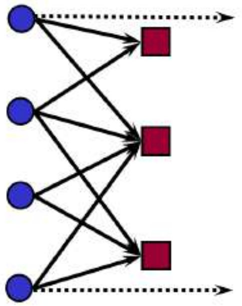
Evaluating policies

- » Solution quality
  - On average
  - Worst case performance
- » Computational
  - Speed
  - Reliability
- » Transparency
- » Methodological complexity
- » Data requirements

Modeling sequential decision problems

If our policy involves solving an optimization problem, we need to recognize two optimization problems:

The surrogate optimization model – used for making decisions



$$x_t = X^\pi(S_t \vee \theta)$$

The objective function, used for evaluating policies.

$$F^\pi(\theta) = \mathbb{E}\left\{ \sum_{t=0}^{T} C_t\left(S_t, X_t^\pi\left(S_t|\theta\right)\right) \vee S_0 \right\}$$

An optimal solution to the surrogate optimization model is almost never an optimal policy, and may perform quite poorly.

The problem solving process

A
p
p
l
i
c
a
t
i
o
n

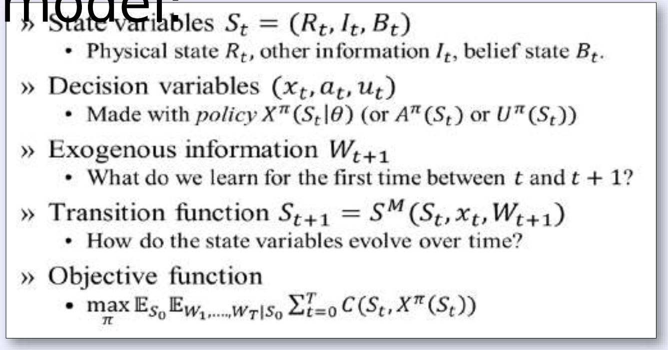
Step 1

Identify:

1. Performance metrics
2. Types of decisions
3. Sources of uncertainty

Step 2:

Mathematical model:



» State variables $S_t = (R_t, I_t, B_t)$
- Physical state R_t, other information I_t, belief state B_t.

» Decision variables (x_t, a_t, u_t)
- Made with *policy* $X^\pi(S_t|\theta)$ (or $A^\pi(S_t)$ or $U^\pi(S_t)$)

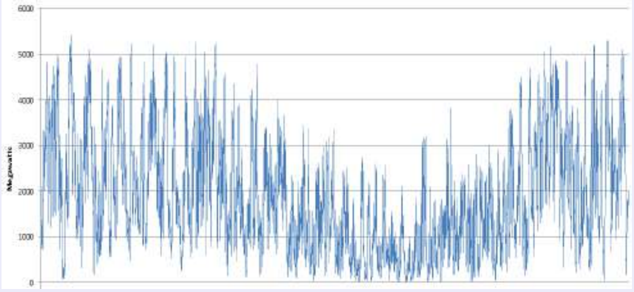
» Exogenous information W_{t+1}
- What do we learn for the first time between t and $t+1$?

» Transition function $S_{t+1} = S^M(S_t, x_t, W_{t+1})$
- How do the state variables evolve over time?

» Objective function
- $\max_\pi \mathbb{E}_{S_0} \mathbb{E}_{W_1,\ldots,W_T|S_0} \sum_{t=0}^{T} C(S_t, X^\pi(S_t))$

Step 3:

Uncertainty modeling



Step 4:

Designing policies

$$\max_\pi \mathbb{E}\left\{\sum_{t=0}^{T} C\left(S_t, X^\pi\left(S_t\right)\right) \vee S_0\right\}$$

Step 5:

Computer model

Step 6:

Implementation/analysis

The problem solving process

A
p
p
l
i
c
a
t
i
o
n

Step 1

Identify:

1. Performance metrics
2. Types of decisions
3. Sources of uncertainty

Step 2:

Mathematical model:

» State variables $S_t = (R_t, I_t, B_t)$
  • Physical state R_t, other information I_t, belief state B_t.

» Decision variables (x_t, a_t, u_t)
  • Made with *policy* $X^\pi(S_t|\theta)$ (or $A^\pi(S_t)$ or $U^\pi(S_t)$)

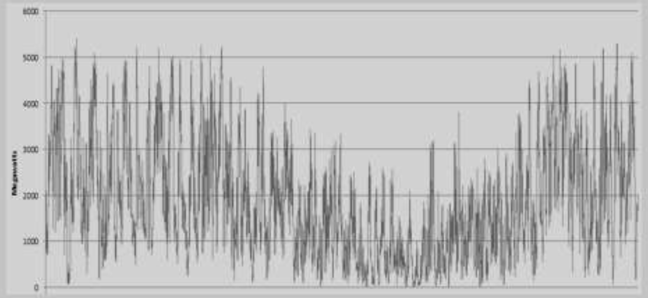
» Exogenous information W_{t+1}
  • What do we learn for the first time between t and $t+1$?

» Transition function $S_{t+1} = S^M(S_t, x_t, W_{t+1})$
  • How do the state variables evolve over time?

» Objective function
  • $\max_\pi \mathbb{E}_{S_0} \mathbb{E}_{W_1,\ldots,W_T|S_0} \sum_{t=0}^{T} C(S_t, X^\pi(S_t))$

Step 3

Uncertainty modeling



Step 4:

Designing policies

$$\max_\pi \mathbb{E}\left\{\sum_{t=0}^{T} C(S_t, X^\pi(S_t))|S_0\right\}$$

Step 5:

Computer model

Step 6:

Implementation/analysis

The problem solving process

Step 1

If you want to run a better you have to make better decisions.

The path to improvement starts by answering three questions:

1. What are the performance metrics?
2. What decisions do you have to make (and who makes them)?
3. What are the sources of uncertainty?

Download from https://tinyurl.com/DoingBetterppt/

OUTLINE

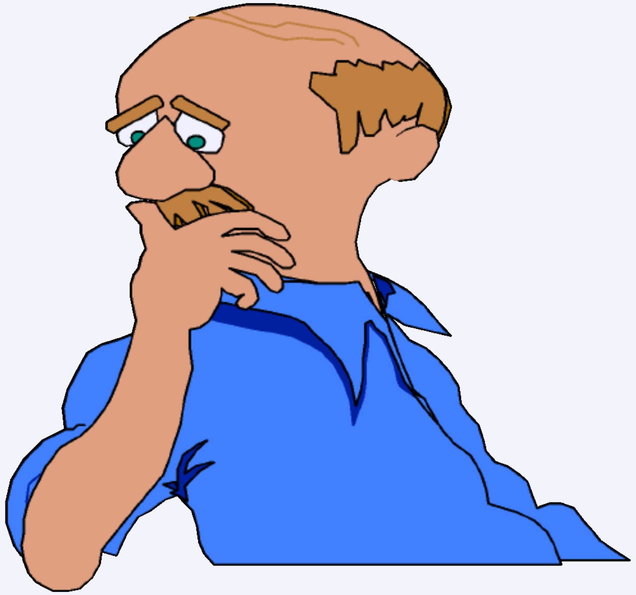
- ➜The seven levels of artificial intelligence
- ➜The universal modeling framework
- ➜Designing policies
- ➜Mutual fund cash balance optimization
- ➜Choosing the best policy
- ➜A new educational field: sequential decision analytics

Designing policies

What is a policy?

A policy is method that makes a decision using the information in the state variable.

… *any method.*



Designing policies

Policies and the English language

Algorithm	Format	Precedent
Behavior	Formula	Prejudice
Belief	Grammar	Principle
Bias	Habit	Procedure
Canon	Heuristics	Process
Code	Laws/bylaws	Protocols
Commandment	Manner	Recipe
Conduct	Method	Ritual
Control law	Mode	Rule
Convention	Mores	Strategy
Culture	Norm	Style
Customs	Orthodox	Syntax
Doctrine	Patterns	Technique
Duty	Plans	Template
Etiquette	Policies	Tenet
Fashion	Practice	Tradition
Way of life		

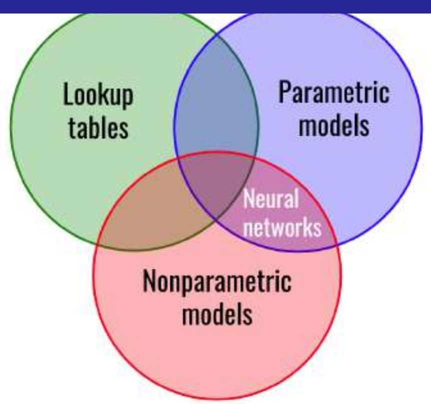
BRIDGING MACHINE LEARNING & SEQUENTIAL DECISIONS

Machine learning

$$\min_{f\in F,\theta\in\Theta^f} \frac{1}{N}\sum_{n=1}^{N}\left(y^n - f(x^n|\theta)\right)^2$$

Searching over functions

“Big dataset”



Sequential decisions

$$\max_{\pi=(f\in F,\theta\in\Theta^f)} \frac{1}{N}\sum_{n=1}^{N}\sum_{t=0}^{T} C(S_t^n, X^\pi(S_t^n|\theta))$$

$$S_{t+1}=S^M(S_t,x_t,W_{t+1})$$

Searching over policies

System model

Designing policies

There are two fundamental strategies for designing policies

Policy search – Search over a class of methods for making decisions to optimize some metric over time.

- » Finding the best class of policy.
- » Finding the best policy within the class.

Lookahead approximations – Approximate the impact of a decision now on the future.

- » The contribution from the first period, plus
- » An approximation of the sum of contributions in future time periods resulting from the first decision.

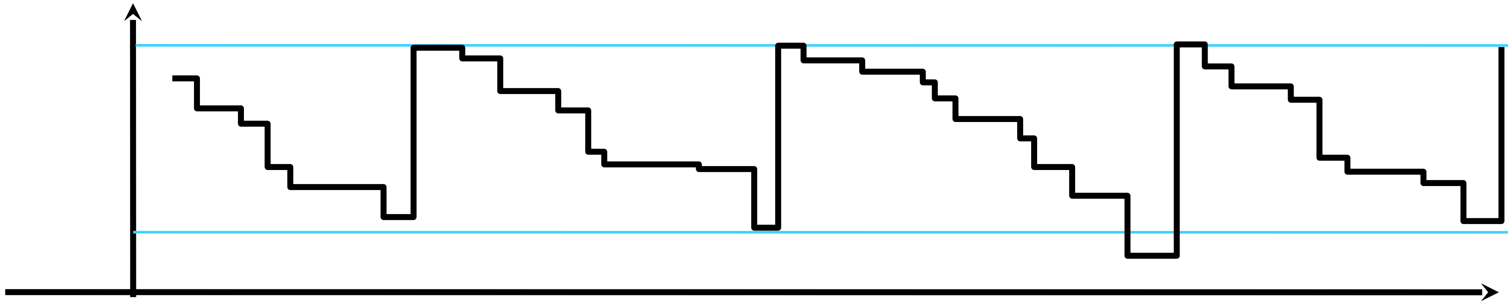
Policy search

1) Policy function approximation (PFA)

These are analytical functions that specify what to do given what we know.

Examples:

a) Order-up-to inventory policy



b) Buy when the price goes **below** and sell when it goes **above**

c) Lookup tables, linear/nonlinear models, neural networks, nonparametric models, … any function we might use in machine learning.

Policy search

2) Cost function approximations (CFAs)

These are parameterized optimization problems:

a) Find the shortest path to a destination, but add a buffer (e.g. 15 minutes) to make sure you arrive on time.

b) Schedule nurses for hours per week, which allows for unforeseen emergencies.

c) Advertise the product which solves:

Parametric CFAs are widely used in industry yet dismissed by the academic research community. This is actually quite a powerful strategy.

Portfolio optimization

Portfolio optimization:

Current and bought/sold assets

M assets

Current price

Transaction costs

Covariance matrix

$$X_t^{\pi}(S_t|\theta) = \arg\max_{x_t}\left(\sum_{i=1}^{M}\left((R_{ti}+x_{ti})(\bar{f}_{ti}(\theta)+p_{ti}) - c^{trans}|x_{ti}|p_{ti}\right) - \eta(R_t+x_t)^T\Sigma_t(R_t+x_t)\right)$$

$$\bar{f}_{ti}(\theta) = \sum_{f\in F}\theta_f f_{t,f,i}$$

Variance of portfolio returns

Adjustment based on assorted features of the economy

Portfolio optimization

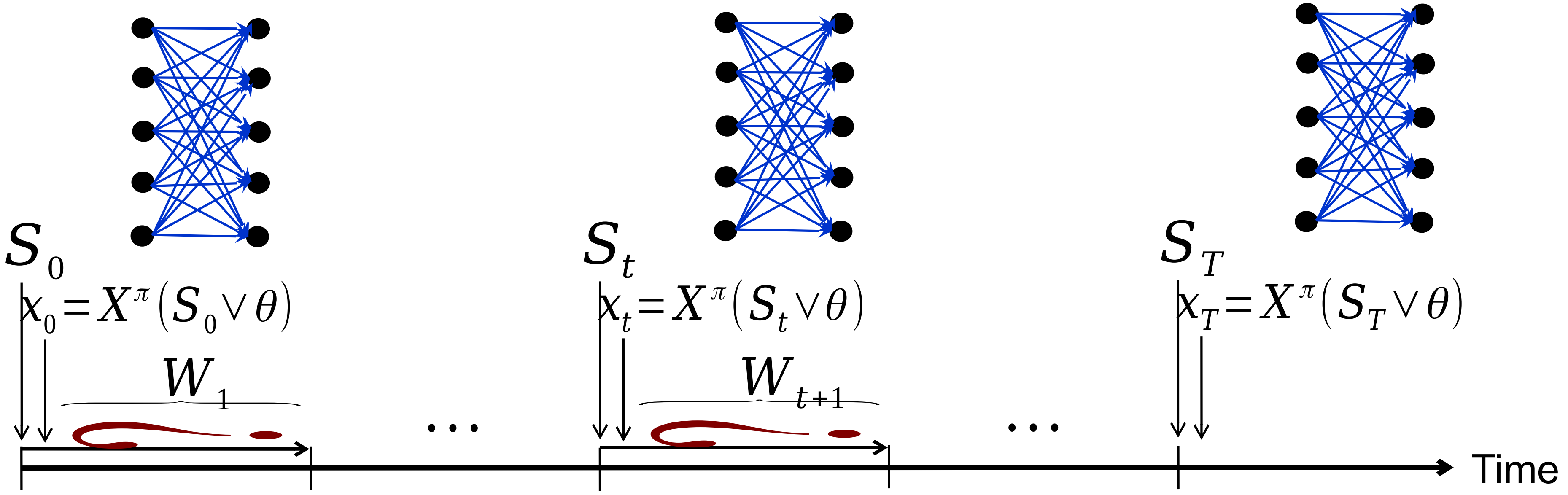
Portfolio optimization:

$$X_t^\pi(S_t|\theta)=\arg\max_{x_t}\left(\sum_{i=1}^{M}\left(\left(R_{ti}+x_{ti}\right)\left(\bar{f}_{ti}(\theta)+p_{ti}\right)-c^{trans}|x_{ti}|p_{ti}\right)-\eta\left(R_t+x_t\right)^T\Sigma_t\left(R_t+x_t\right)\right)$$

Tunable parameters

Portfolio optimization

Evaluating performance over time:



$$F^\pi(\theta) = C(S_0, x_0) + C(S_t, x_t) + \cdots + C(S_T, x_T)$$

$$= \sum_{t=0}^{T} C\left(S_t, X^\pi(S_t \vee \theta)\right)$$ performance of policy over time

Derivative-based stochastic search

- The policy (deterministic nonlinear program)

$$X_t^\pi(S_t|\theta) = \arg\max_{x_t} \left(\sum_{i=1}^{M} \left((R_{ti} + x_{ti})(\bar{f}_{ti}(\theta) + p_{ti}) - c^{trans} |x_{ti}| p_{ti} \right) - \eta (R_t + x_t)^T \Sigma_t (R_t + x_t) \right)$$

- The objective function (evaluated using simulation)

where

Historical prices

- Stochastic optimization problem

$$\max_{\pi=(f,\theta)} \mathbb{E}_W F^\pi(\theta, W)$$

- Stochastic gradient algorithm

$$\theta^{n+1} = \theta^n + \alpha_n \nabla F^\pi(\theta^n, \omega^{n+1})$$

(joint with Saeed Ghadimi)

Cost function approximations

- CFAs are widely used in industry (in an ad-hoc manner)

  » Airlines optimizing schedules with schedule slack to handle weather uncertainty.

  » Manufacturers using buffer stocks to hedge against production delays and quality problems.

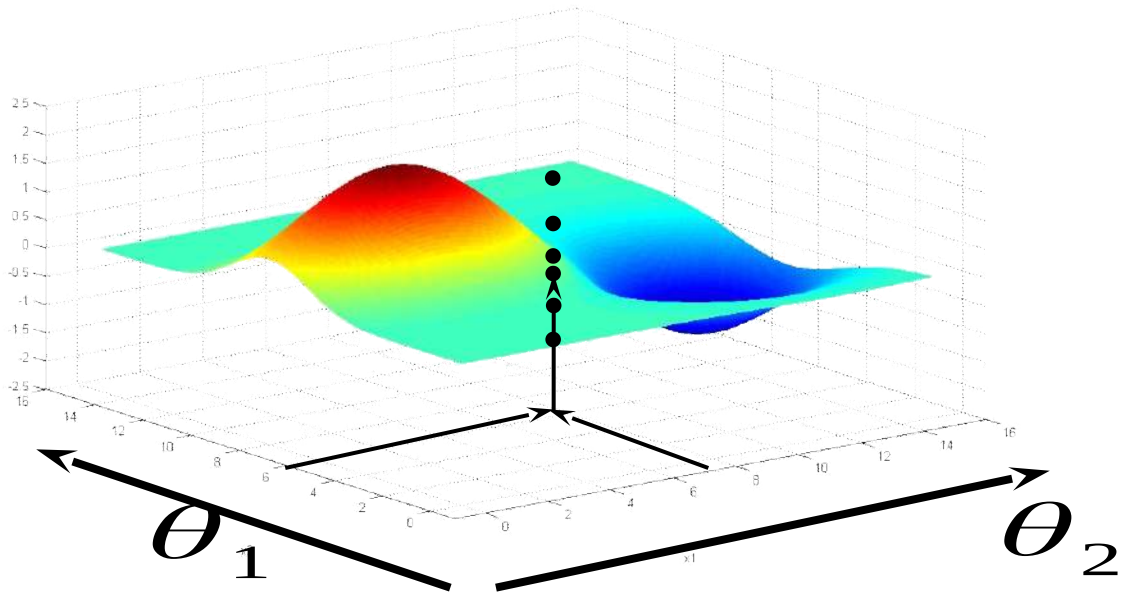
  » Optimizing the location of distribution centers and warehouses, allowing for the benefits of inventory pooling.

  » Grid operators scheduling extra generation capacity in case of outages.

  » See: https://tinyurl.com/cfapolicy for an introduction to CFA

Policy search

- Both PFAs and CFAs have tunable parameters which have to be tuned. We write this mathematically as



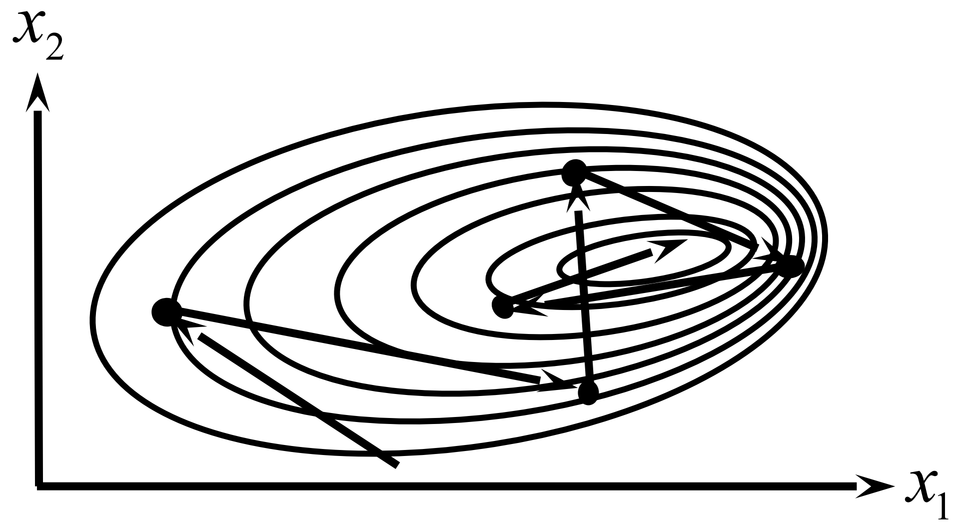
- There are two ways to evaluate a policy:
  - » In a simulator – This allows us to perform extensive testing in a controlled environment.
  - » In the field – This is “learning by doing”

Policy search

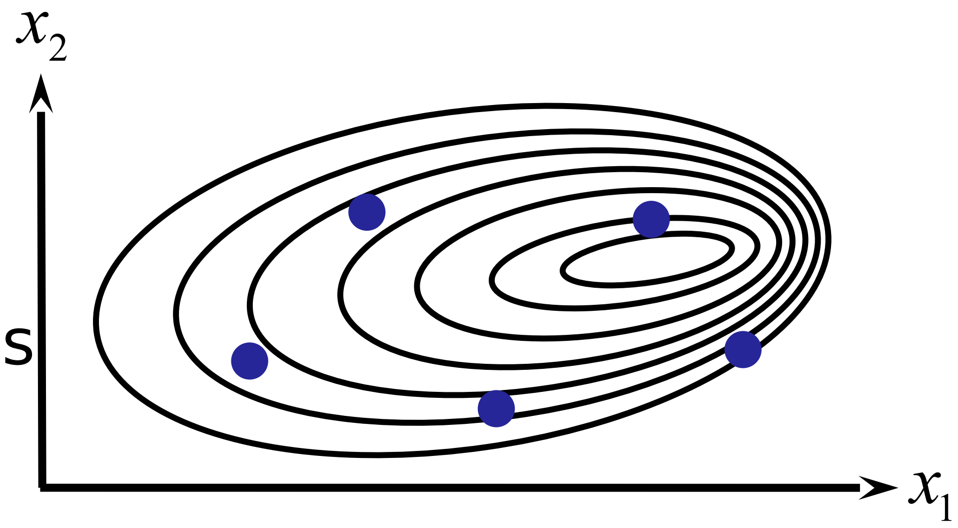
- How do we search for the best
  - » Derivative-based
    - Stochastic gradient methods:

$$\theta^{n+1} = \theta^n + \alpha_n \nabla_\theta F(\theta^n, W^{n+1})$$

Decision



  - » Derivative-free
    - Build a belief model that approximates our function.



Designing policies

There are two fundamental strategies for designing policies

Policy search – Search over a class of methods for making decisions to optimize some metric over time.

» Finding the best class of policy.

» Finding the best policy within the class.

Lookahead approximations – Approximate the impact of a decision now on the future.

» The contribution from the first period, plus

» An approximation of the sum of contributions in future time periods resulting from the first decision.

Lookahead approximations

- Lookahead approximations combine:
  - » The immediate contribution (or cost) of a decision made now…
  - » … and an approximation of future contributions (or costs)

Lookahead approximations

Lookahead policies are based on solving

$$X_t^*(S_t) = \arg\max_x \left(C(S_t, x_t) + \mathbb{E}\left\{ \max_\pi \left\{ \mathbb{E} \sum_{t'=t+1}^{T} C\left(S_{t'}, X_{t'}^\pi(S_{t'})\right) \mid S_{t+1} \right\} \mid S_t, x_t \right\} \right)$$

Contribution we receive now

Future contributions

- This looks like scary mathematics, but it is what all of us are doing when we make decisions now that consider what might happen in the future.
- The challenge is … *how to compute it!!!*

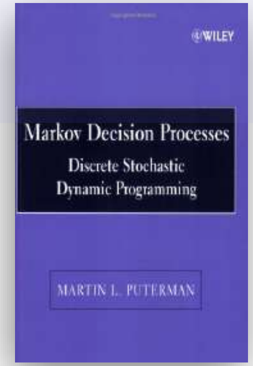
Lookahead approximations

Lookahead approximations

Approximate the impact of a decision now on the future

$$X_t^*\left(S_t\right)=argmax_x\left(C\left(S_t,x_t\right)+\mathbb{E}\left\{\max_\pi\left\{\mathbb{E}\sum_{t'=t+1}^{T}C\left(S_{t'},X_{t'}^{\pi}\left(S_{t'}\right)\right)\vee S_{t+1}\right\}\vee S_t,x_t\right\}\right)$$

3) Value function approximations (VFAs)



$$X_t^*(S_t)=\arg\max_{x_t}\left(C(S_t,x_t)+\mathrm{E}\left\{V_{t+1}(S_{t+1})\mid S_t,x_t\right\}\right)$$

$$X_t^{VFA}(S_t)=\arg\max_{x_t}\left(C(S_t,x_t)+\mathrm{E}\left\{\bar{V}_{t+1}(S_{t+1})\mid S_t,x_t\right\}\right)$$

$$=\arg\max_{x_t}\left(C(S_t,x_t)+\bar{V}_t^x(S_t^x)\right)$$

$$=\arg\max_{x_t}\bar{Q}_t(S_t,x_t)\quad(\text{"Q-learning"})$$

Lookahead approximations

Lookahead approximations

Approximate the impact of a decision now on the future

$$X_t^*(S_t) = argmax_x \left(C(S_t, x_t) + \mathbb{E}\left\{ \max_\pi \left\{ \mathbb{E} \sum_{t'=t+1}^{T} C\left(S_t, X_{t'}^{\pi}\left(S_{t'}\right)\right) \vee S_{t+1} \right\} \vee S_t, x_t \right\} \right)$$

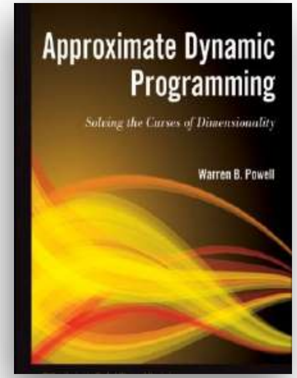
3) Value function approximations (VFAs)

$$X_t^*(S_t) = \arg\max_{x_t} \left(C(S_t, x_t) + \mathsf{E}\left\{ V_{t+1}(S_{t+1}) \mid S_t, x_t \right\} \right)$$

$$X_t^{VFA}(S_t) = \arg\max_{x_t} \left(C(S_t, x_t) + \mathsf{E}\left\{ \bar{V}_{t+1}(S_{t+1}) \mid S_t, x_t \right\} \right)$$

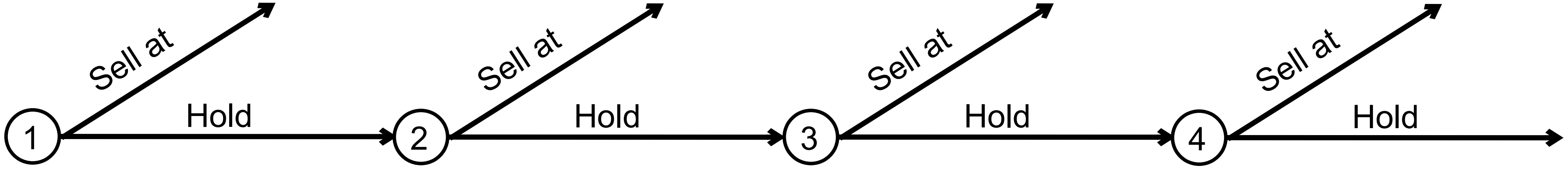
$$= \arg\max_{x_t} \left(C(S_t, x_t) + \bar{V}_t^x(S_t^x) \right)$$

$$= \arg\max_{x_t} \bar{Q}_t(S_t, x_t) \quad (\text{"Q-learning"})$$

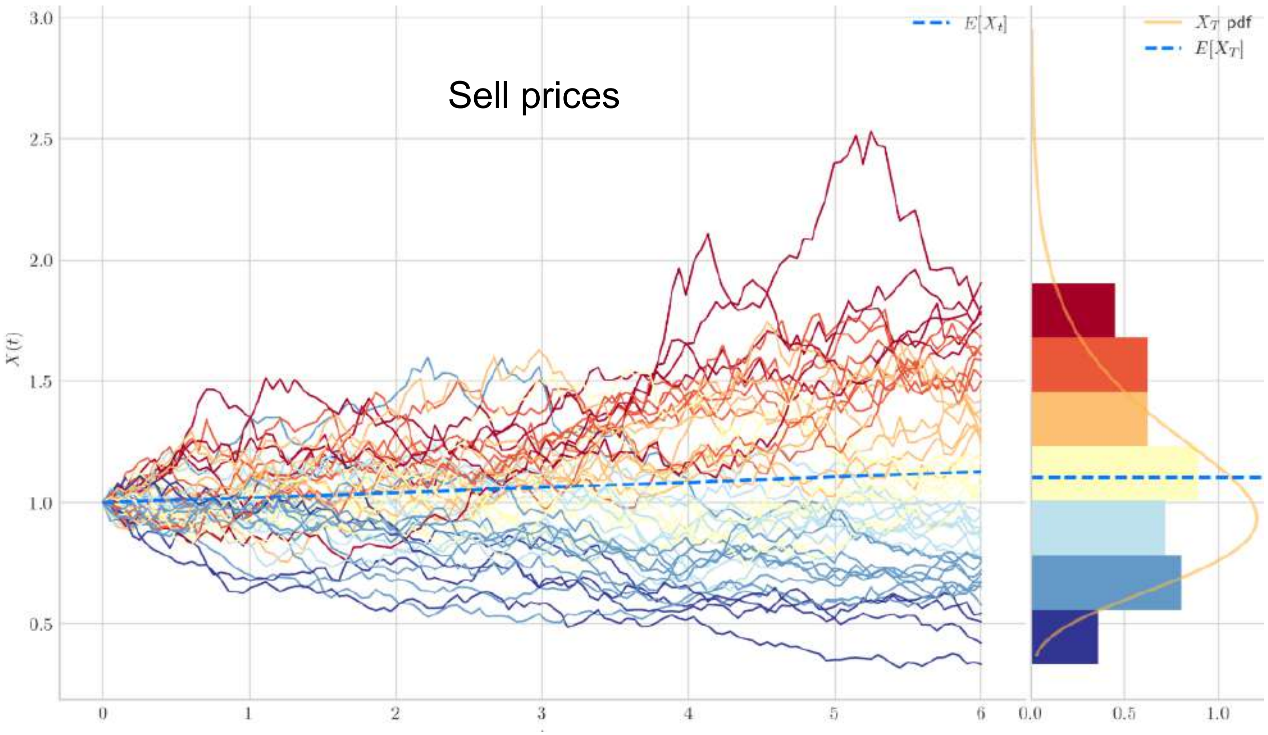


Value function approximations

Longstaff and Schwartz - Pricing an American option



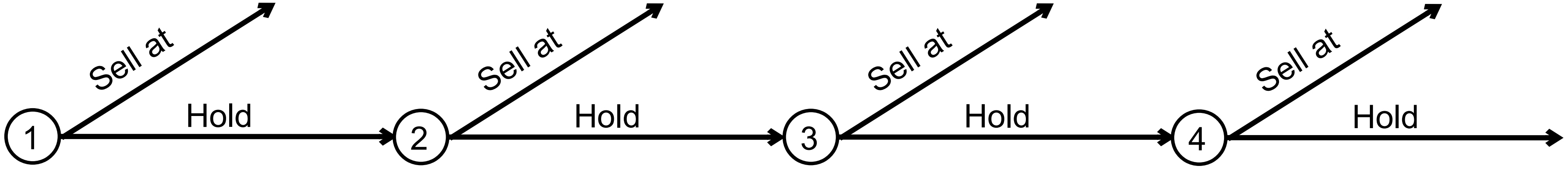
Stock prices				
	Time period			
Outcome	1	2	3	4
1	1.21	1.08	1.17	1.15
2	1.09	1.12	1.17	1.13
3	1.15	1.08	1.22	1.35
4	1.17	1.12	1.18	1.15
5	1.08	1.15	1.10	1.27
6	1.12	1.22	1.23	1.17
7	1.16	1.14	1.13	1.19
8	1.22	1.18	1.21	1.28
9	1.08	1.11	1.09	1.10
10	1.15	1.14	1.18	1.22



Optimal Dynamics

Value function approximations

Longstaff and Schwartz - Pricing an American option

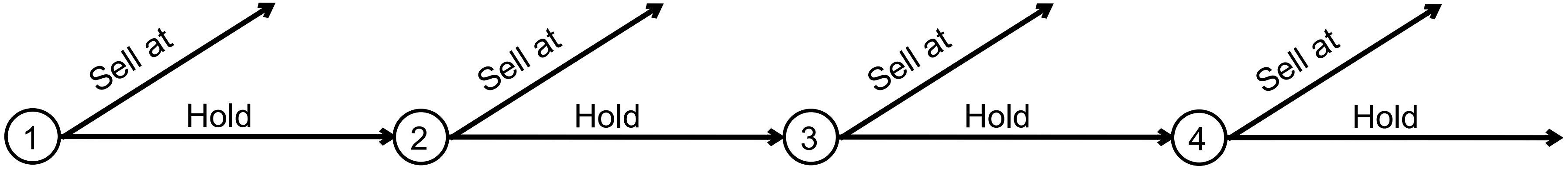


	Decision	
Outcome	Sell	Hold
1	0.03	0.04155 ×.95 = **0.03947**
2	0.03	0.03662 ×.95 = **0.03479**
3	0.00	0.02397 ×.95 = **0.02372**
4	0.02	0.03346 ×.95 = **0.03178**
5	**0.10**	0.05285 ×.95 = 0.05021
6	0.00	0.00414 ×.95 = **0.00394**
7	**0.07**	0.00899 ×.95 = 0.00854
8	0.00	0.01610 ×.95 = **0.01530**
9	**0.11**	0.06032 ×.95 = 0.05731
10	0.02	0.03099 ×.95 = **0.02944**

Prices at time 2

Value function approximations

Longstaff and Schwartz - Pricing an American option



	Decision	
Outcome	Sell	Hold
1	0.03	0.04155 × .95 = **0.03947**
2	0.03	0.03662 × .95 = **0.03479**
3	0.00	0.02397 × .95 = **0.02372**
4	0.02	0.03346 × .95 = **0.03178**
5	**0.10**	0.05285 × .95 = 0.05021
6	0.00	0.00414 × .95 = **0.00394**
7	**0.07**	0.00899 × .95 = 0.00854
8	0.00	0.01610 × .95 = **0.01530**
9	**0.11**	0.06032 × .95 = 0.05731
10	0.02	0.03099 × .95 = **0.02944**

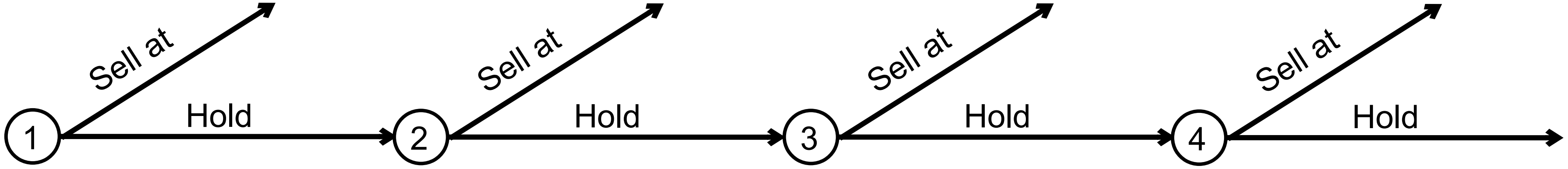
Expected value of holding the asset:

$$\overline{V}_3 = \theta_{t0} p_{t-2} + \theta_{t1} p_{t-1} + \theta_{t2} p_t + \theta_{t3} (p_t)^2$$

$$¿ 0.0056 - 0.1234 p_1 + 0.6011 p_2 - 0.3903 (p_2)^2$$

Value function approximations

Longstaff and Schwartz - Pricing an American option

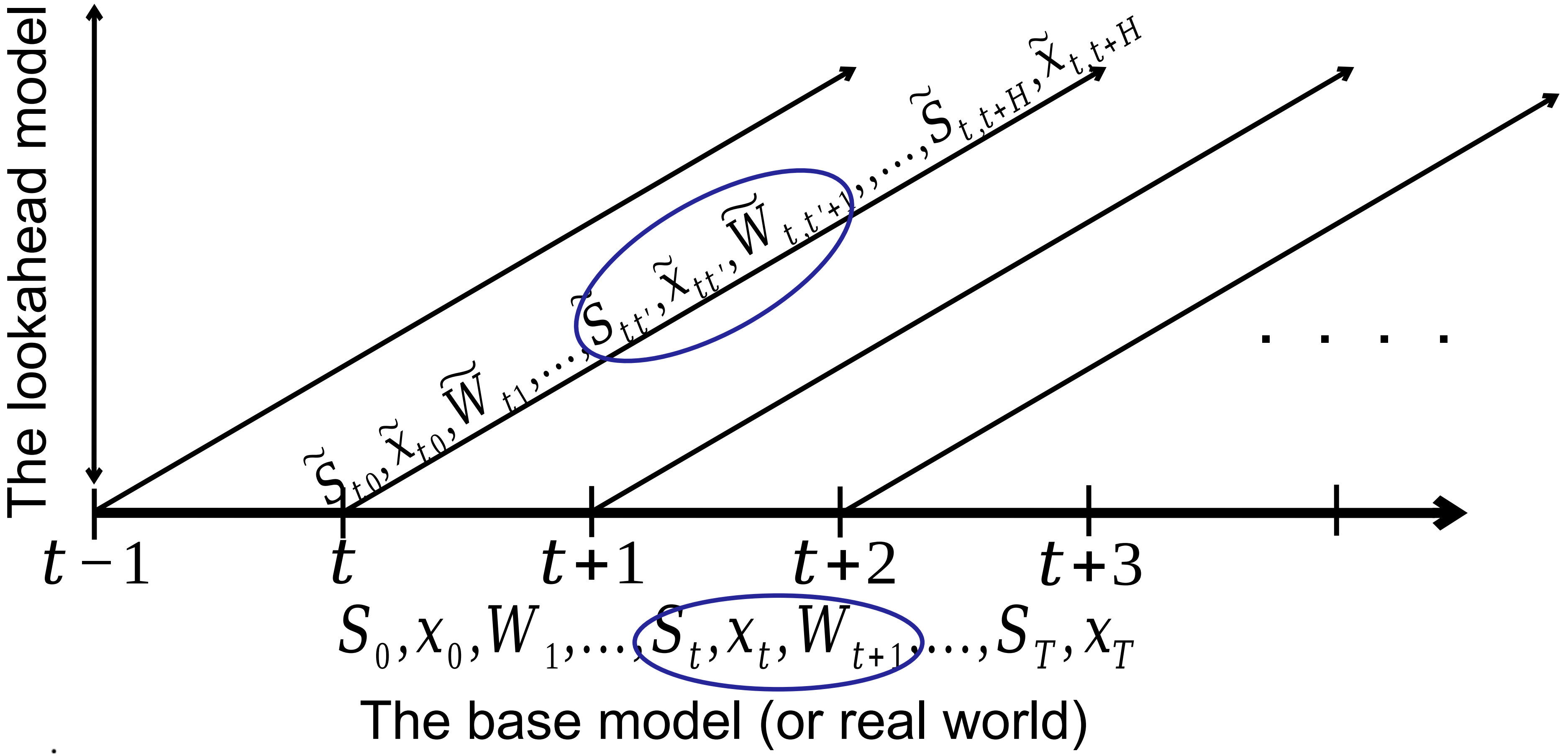


	Decision	
Outcome	Sell	Hold
1	0.03	0.04155 ×.95 = **0.03947**
2	0.03	0.03662 ×.95 = **0.03479**
3	0.00	0.02397 ×.95 = **0.02372**
4	0.02	0.03346 ×.95 = **0.03178**
5	**0.10**	0.05285 ×.95 = 0.05021
6	0.00	0.00414 ×.95 = **0.00394**
7	**0.07**	0.00899 ×.95 = 0.00854
8	0.00	0.01610 ×.95 = **0.01530**
9	**0.11**	0.06032 ×.95 = 0.05731
10	0.02	0.03099 ×.95 = **0.02944**

average of max value for each sample path

Direct lookahead policies

4) Direct lookahead policies (DLAs) – Here we create an approximation called the *approximate lookahead model:*



Direct lookahead policies

4) Direct lookahead policies (DLAs) – Here we create an approximation called the *approximate lookahead model:*

$$\left(\tilde{S}_{tt},\tilde{x}_{tt},\tilde{W}_{t,t+1},\tilde{S}_{t,t+1},\tilde{x}_{t,t+1},\tilde{W}_{t,t+2},\ldots,\tilde{S}_{tt'},\tilde{x}_{tt'},\tilde{W}_{t,t'+1},\ldots\right)$$

There are six classes of approximations we can introduce.

Our direct lookahead policy now requires solving:

$$X_t^{DLA}(S_t \vee \theta) = argmax_x \left(C(S_t,x_t) + \tilde{E}\left\{ \max_{\tilde{\pi}} \tilde{E}\left\{ \sum_{t'=t+1}^{t+H} C\left(\tilde{S}_{t'},\tilde{X}_{t'}^{\tilde{\pi}}(\tilde{S}_{t'})\right) \vee \tilde{S}_{t+1} \right\} ¿ S_t, x_t \right\} \right)$$

Sampled information process

Simplified policies

Restricted horizon

Limited decisions

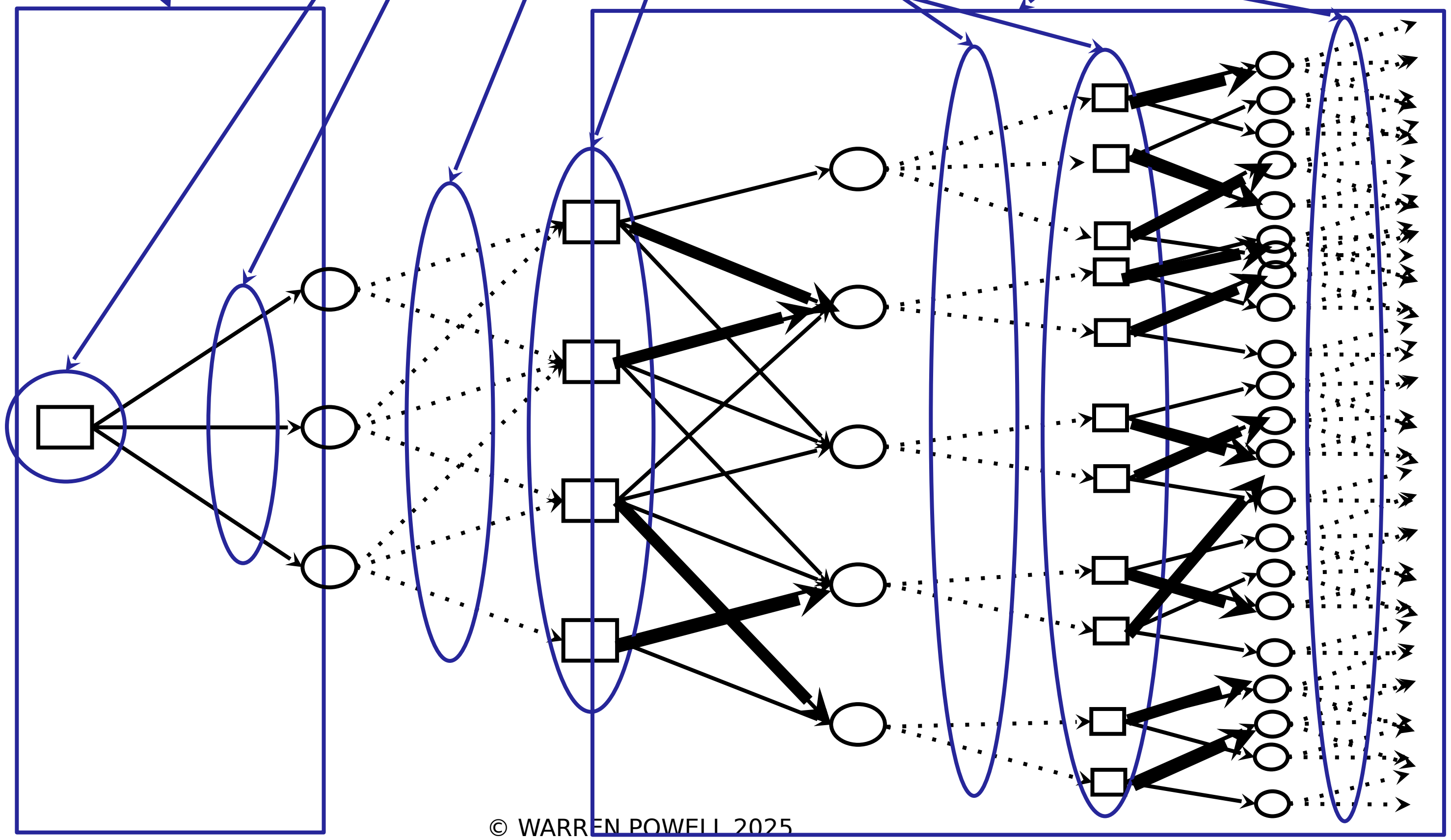
Reduced state variable

Optimal Dynamics

Lookahead approximations

Lookahead policies are based on solving

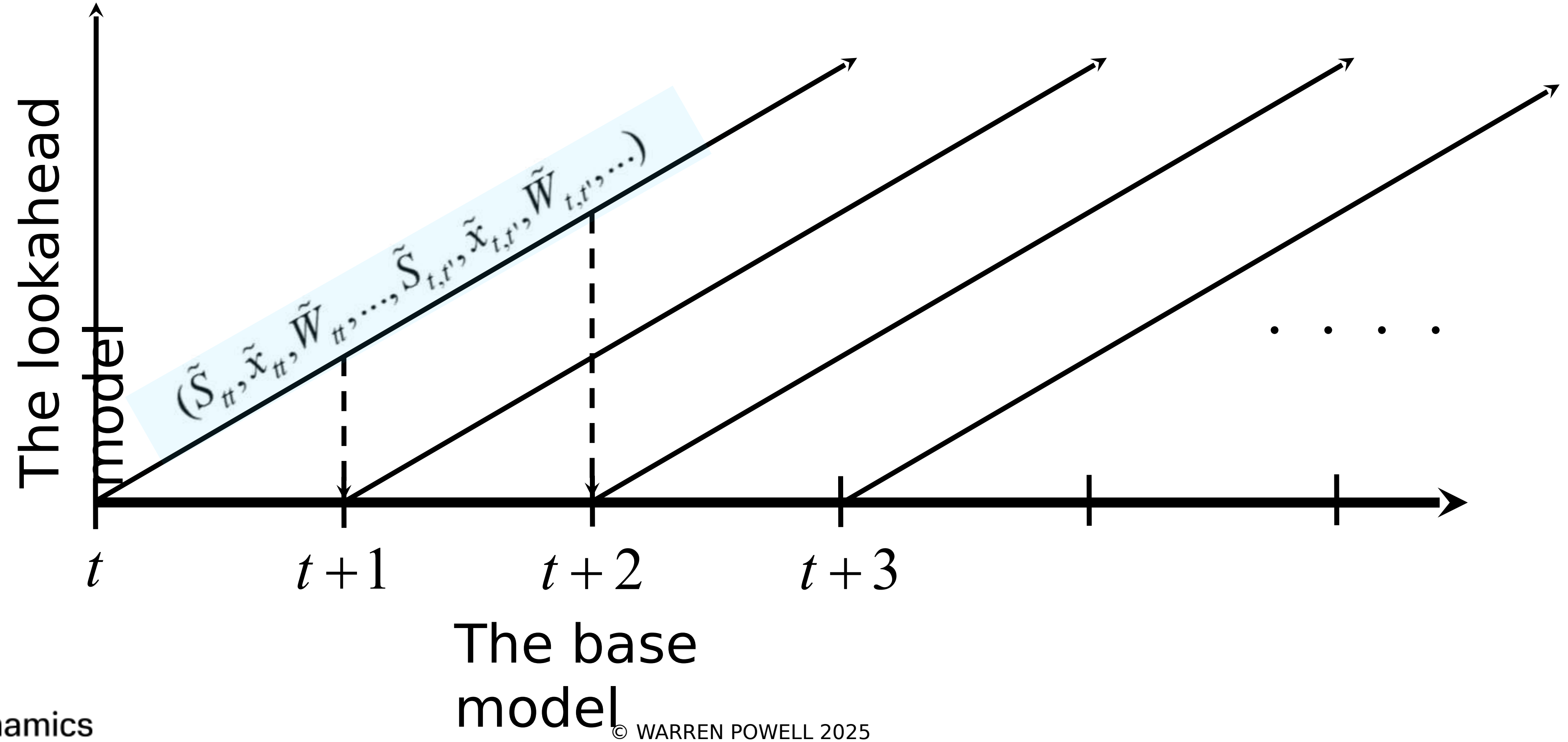
$$X_t^*(S_t) = \arg\max_x \left(C(S_t, x_t) + \mathbb{E}\left\{ \max_\pi \left\{ \mathbb{E} \sum_{t'=t+1}^{T} C(S_{t'}, X_{t'}^\pi(S_{t'})) \mid S_{t+1} \right\} \mid S_t, x_t \right\} \right)$$



Direct lookahead policies

Direct Lookahead Policies (DLAs)

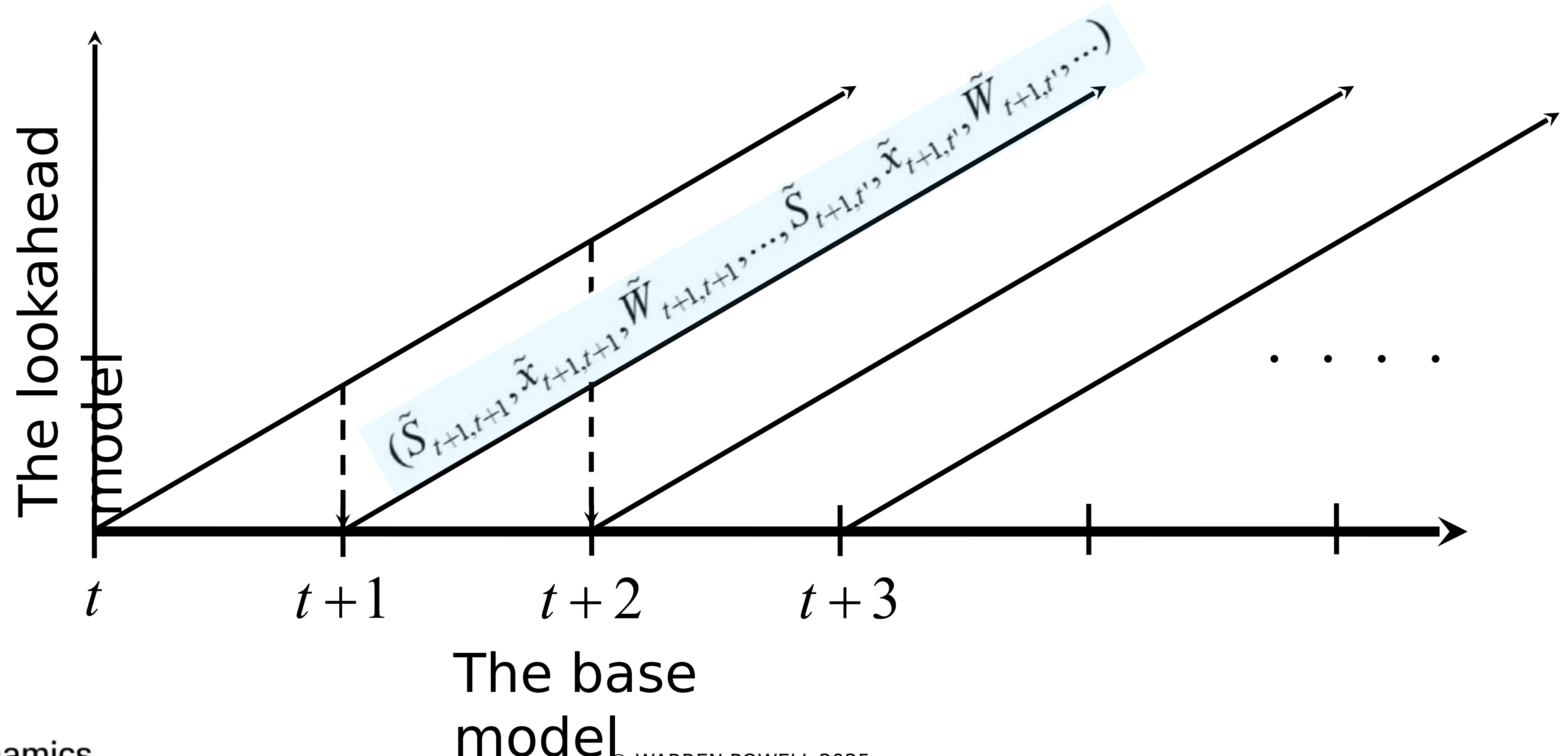
» Tilde variables are used to model approximate lookahead



Direct lookahead policies

Direct Lookahead Policies (DLAs)

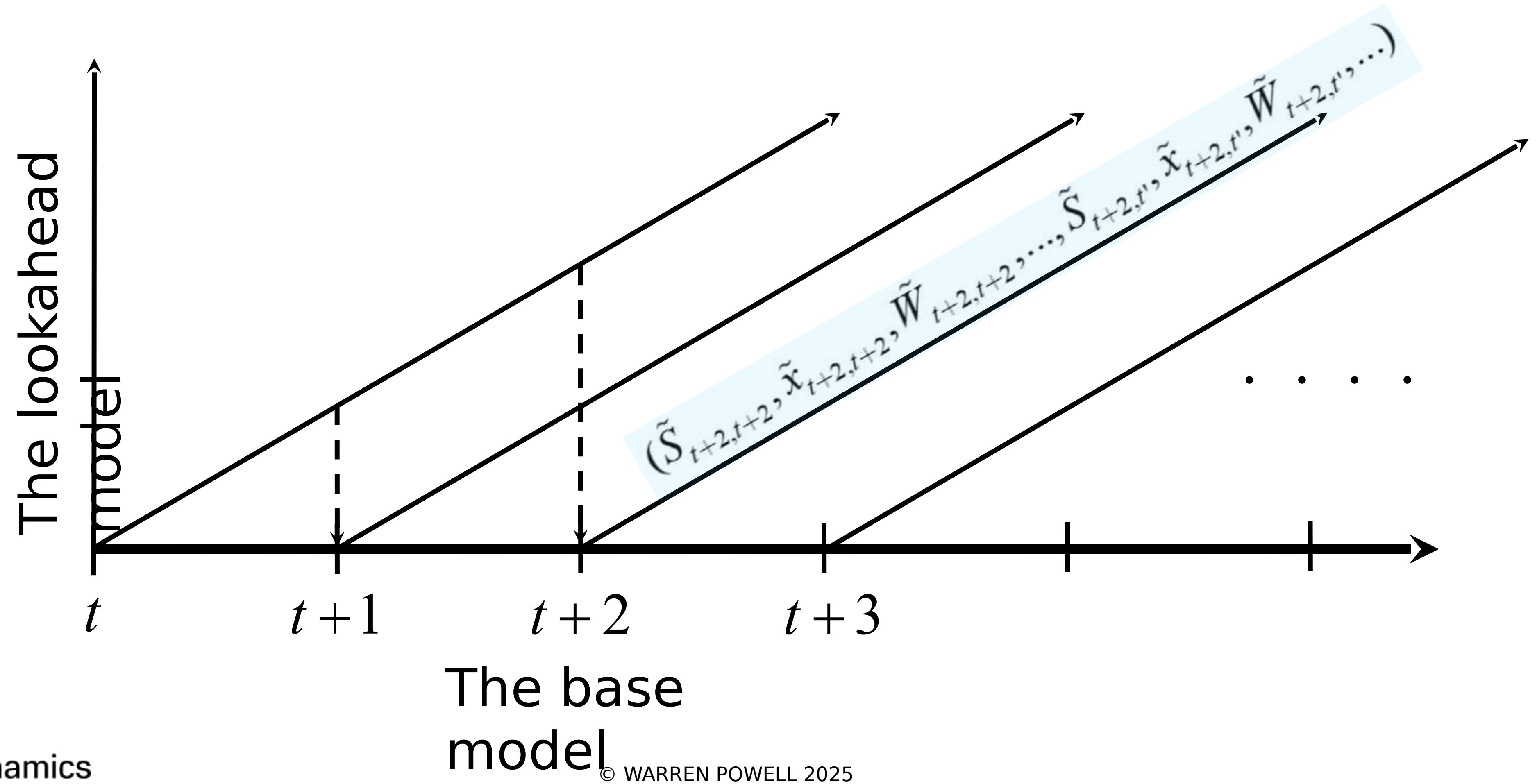
» Tilde variables are used to model approximate lookahead



Direct lookahead policies

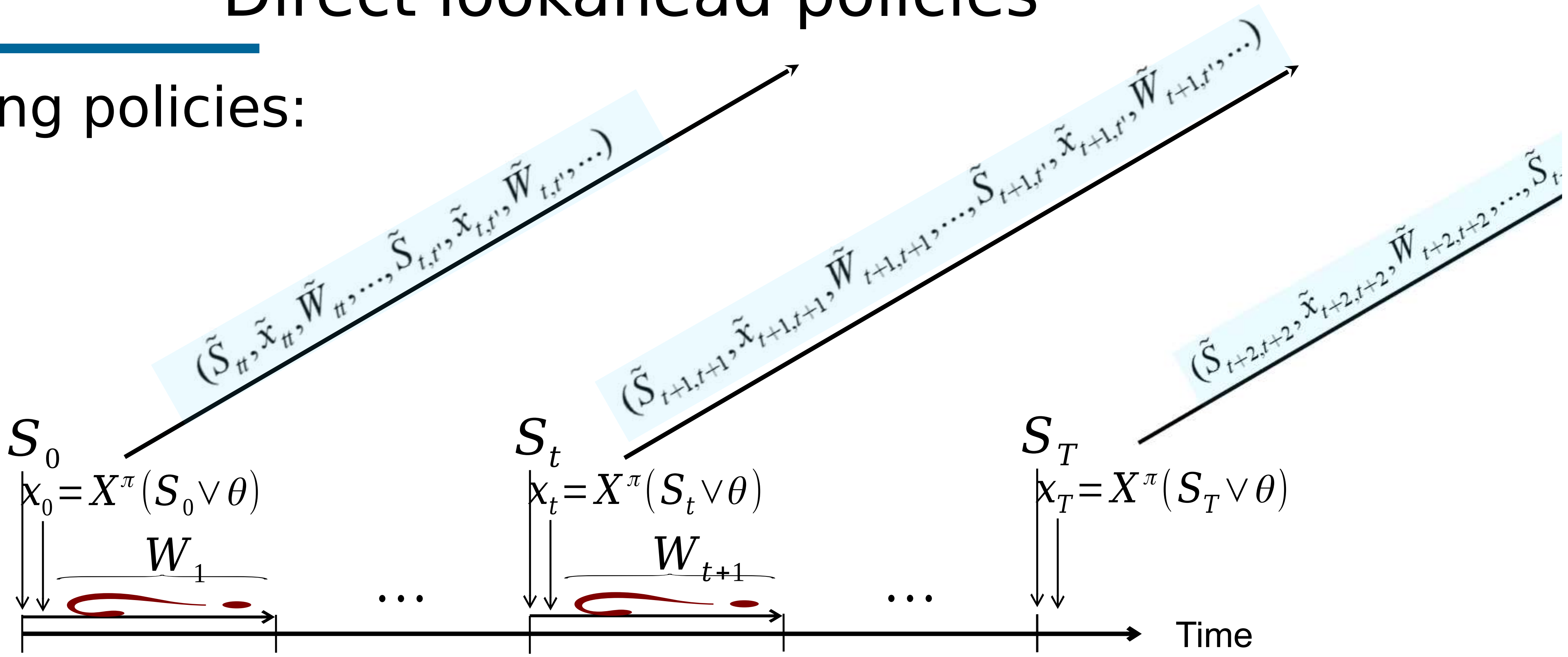
Direct Lookahead Policies (DLAs)

- Tilde variables are used to model approximate lookahead



Direct lookahead policies

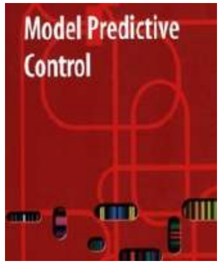
Evaluating policies:

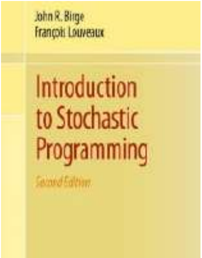


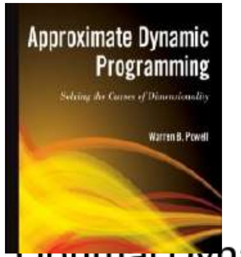
$$F^\pi(\theta) = C(S_0, x_0) + C(S_t, x_t) + \dots + C(S_T, x_T)$$

$$= \sum_{t=0}^{T} C\left(S_t, X^\pi(S_t|\theta)\right)$$ performance of policy over time

Direct lookahead policies







Examples of Lookahead Models

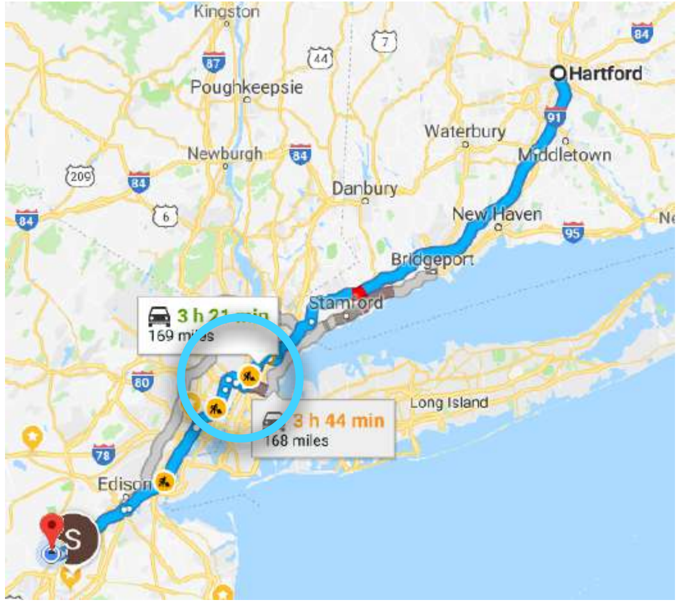
» **The deterministic lookahead model**
  - This is what is most widely used in practice.
  - Standard approach is to use a "best estimate" (which means deterministic) of travel times in the future.
  - This is often referred to as "model predictive control"

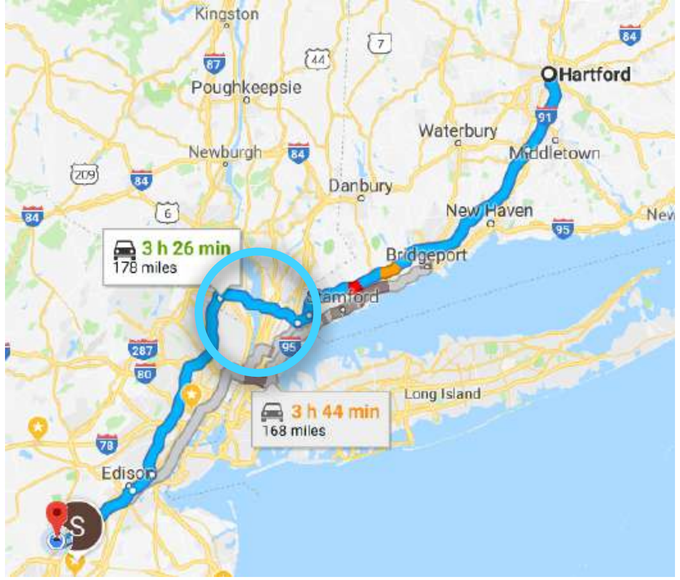
» **Robust optimization** - We could use the 90th percentile of travel times.

» **Stochastic programming** – We represent the future using, say, 20 samples.

» **Approximate dynamic programming applied to approximate lookahead model**

» **Chance constrained programming** – Impose constraint on the probability of





Designing policies

Policy search policies

Policy function approximations (PFAs)

» Simple rules, functions

» Examples:
- Order up to inventory policies
- Buy low, sell high
- Linear/nonlinear functions
- Any function that might be used in machine learning

Cost function approximations (CFAs)

» Parameterized optimization models

» Examples
- Schedule slack for scheduling models
- Buffer stocks in supply chain models
- Power reserves for energy planning

Lookahead policies

Value function approximations (VFAs)

» Making a decision now using the value of being in a future state

» Examples:
- The value of a truck driver
- The value of holding an asset

Direct lookaheads (DLAs)

» Models that optimize over a planning horizon (deterministically/stochastically)

» Examples:
- Deterministic lookahead: Google maps
- Stochastic lookahead: Decision trees

Designing policies

Policy search policies

Policy function approximations (PFAs)

- Simple rules, functions
- Examples:
  - Order up to inventory policies
  - Buy low, sell high
  - Linear/nonlinear decision rules

Cost function approximations (CFAs)

- Parameterized optimization models
- Examples
  - Schedule slack for scheduling models
  - Buffer stocks in supply chain models
  - Power reserves for energy planning

Lookahead policies

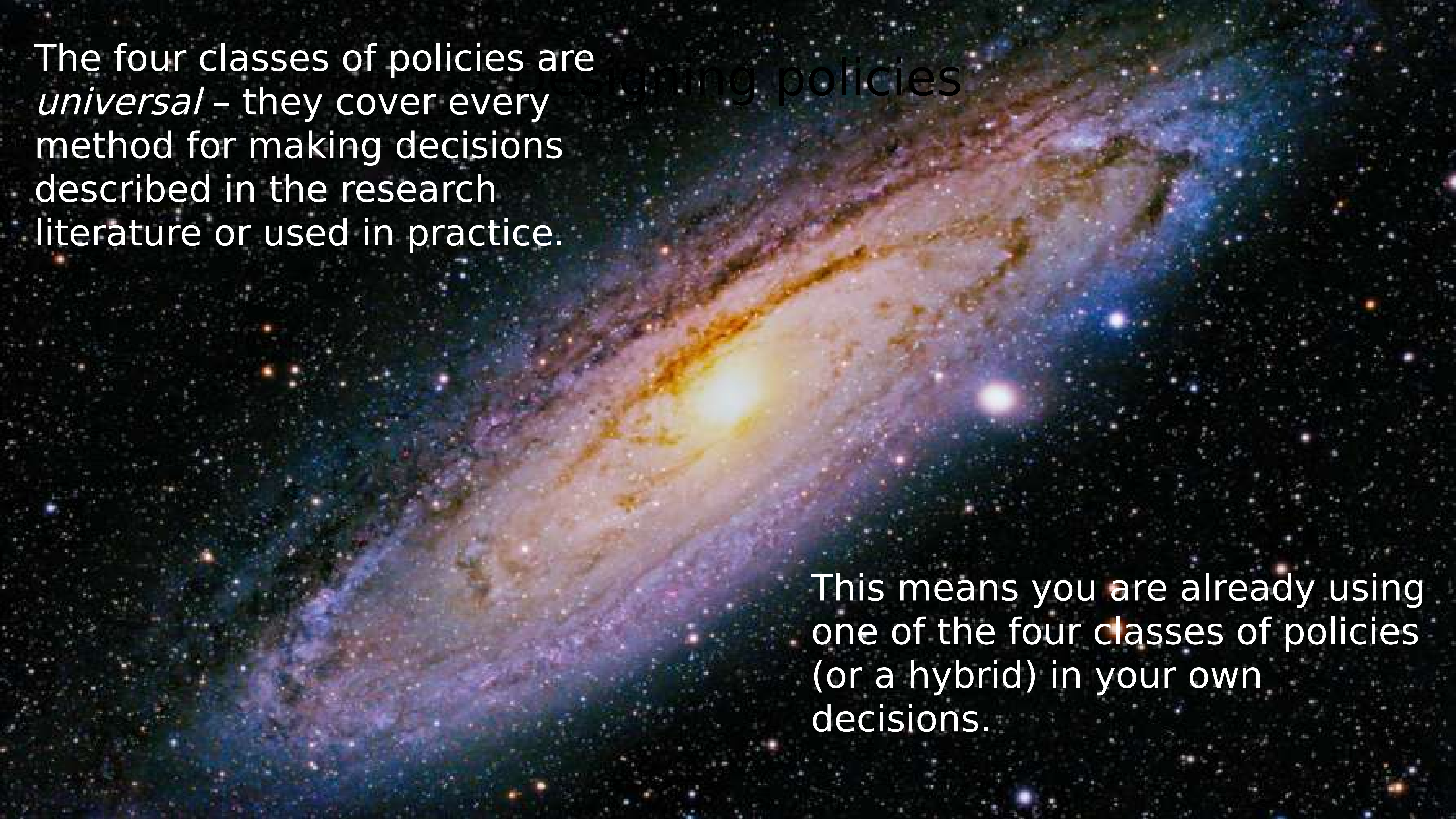
Value function approximations (VFAs)

- Making a decision now using the value of being in a future state
- Examples:
  - The value of a truck driver
  - The value of holding an asset

Direct lookaheads (DLAs)

- Models that optimize over a planning horizon (deterministically/stochastically)
- Examples:
  - Deterministic lookahead: Google maps
  - Stochastic lookahead: Decision trees

Three of the four classes of policies involve solving optimization problems.



Designing policies

The four classes of policies are *universal* – they cover every method for making decisions described in the research literature or used in practice.

This means you are already using one of the four classes of policies (or a hybrid) in your own decisions.

Designing policies

From the four classes, we can create hybrids:

- CFAs with PFAs
  - Choosing least cost supplier, but with rules to exclude high-risk companies.
- Lookahead (DLA) with VFA
  - Optimize seasonal production plan, with functions capturing value of ending inventories
- Parameterized deterministic direct lookaheads (DLA/CFA)
  - Plan seasonal production plan, using -percentile (say, 80^{th} percentile) demand forecasts.
- VFA policy using PFA
  - Distribution planning using VFAs to value inventory at each warehouse, but using rules (PFAs) to force deliveries to specific locations.
- VFA with CFA
  - Start with a VFA-based policy with a linear model, and then tune the parameters of the linear VFA to get the best results using a simulator.

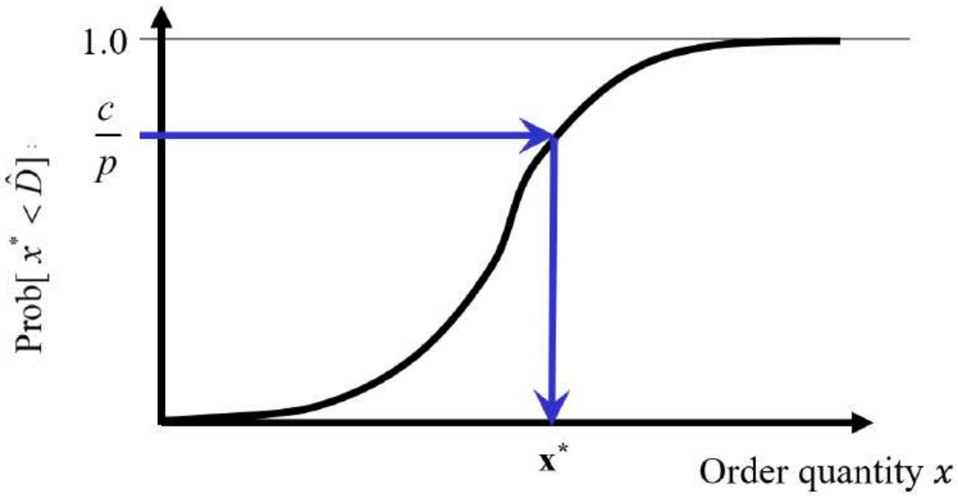
The only way to create effective hybrids is to know all four classes of policies.

OUTLINE

- The seven levels of artificial intelligence
- The universal modeling framework
- Designing policies
- Mutual fund cash balance optimization
- Choosing the best policy
- A new educational field: sequential decision analytics

The newsvendor problem

$$E\frac{dF\left(x^*,\hat{D}\right)}{dx} = p\,\mathrm{Prob}\left[x^* < \hat{D}\right] - c$$



Th

- The student was taught a stylized problem (the "newsvendor problem") which the professor solved in class.
- The student was not taught how to *think* about the problem in a way that allowed him to go after the real-world problems that arise in practice.

Th

Professor – The newsvendor inventory problem we learned in class looks like a way to optimize how much cash I should hold in the mutual fund I manage.

The problem

Mutual funds hold a certain percentage of their assets in cash to meet redemption requests from investors.

We can start by using their existing policy (a form of PFA):

$$X^{\pi}(S_t|\theta) = \theta^{asset} \cdot assets_t$$

… which can be tuned in a simulator or the field:

$$\max_{\theta^{asset}} \mathbb{E}\left\{\sum_{t=0}^{T} C\left(S_t, X^{\pi}(S_t|\theta^{asset})\right)|S_0\right\} \rightarrow$$

Computer simulation

Field experience

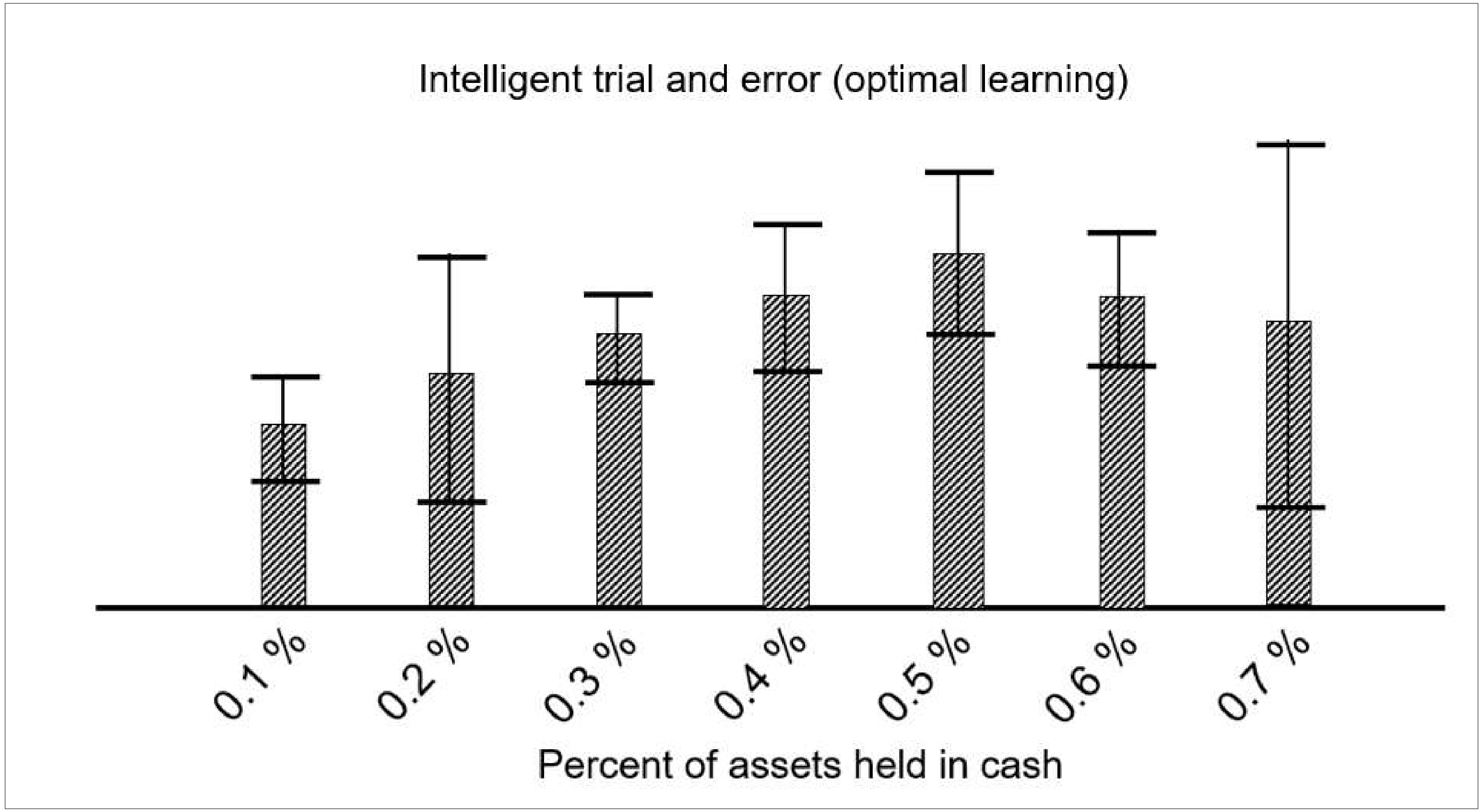
Computer simulations can be performed using historical data. Simulations, as with field testing, can be completely general. There is no need for simplifying assumptions.

Th

Professor – The newsvendor inventory problem we learned in class looks like a way to optimize how much cash I should hold in the mutual fund I manage.

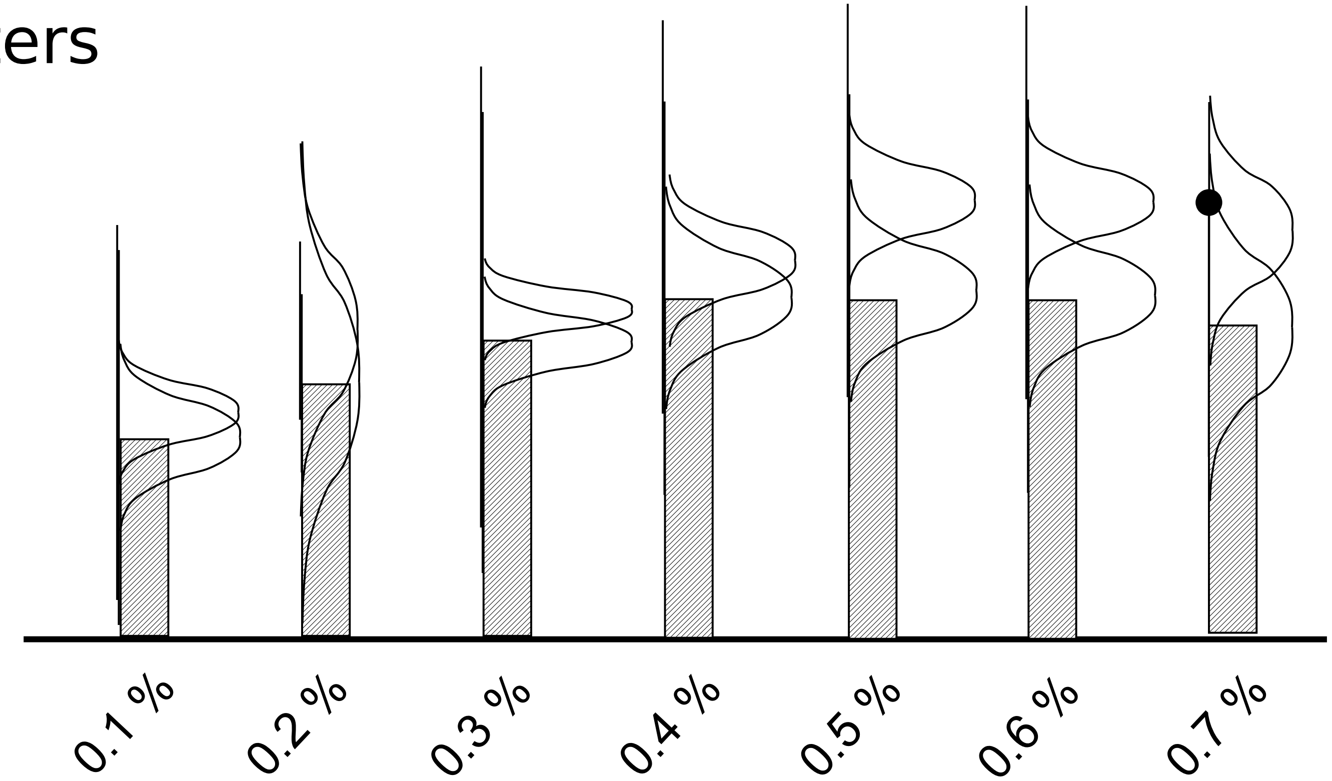
The problem

Mutual funds hold a certain percentage of their assets in cash to meet redemption requests from investors.



Cash management problem

Correlated beliefs: Testing one parameter value teaches us about parameters



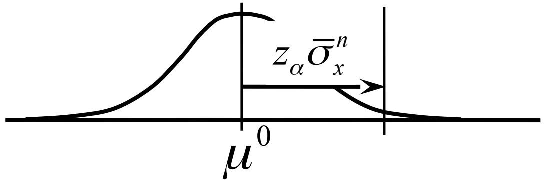
Cash management problem

Cost function approximations (CFA)

» Upper confidence bounding

$$X^{UCB}(S^n \mid \theta^{UCB}) = \arg\max_x \left(\bar{\mu}_x^n + \theta^{UCB} \sqrt{\frac{\log n}{N_x^n}} \right)$$

» Interval estimation

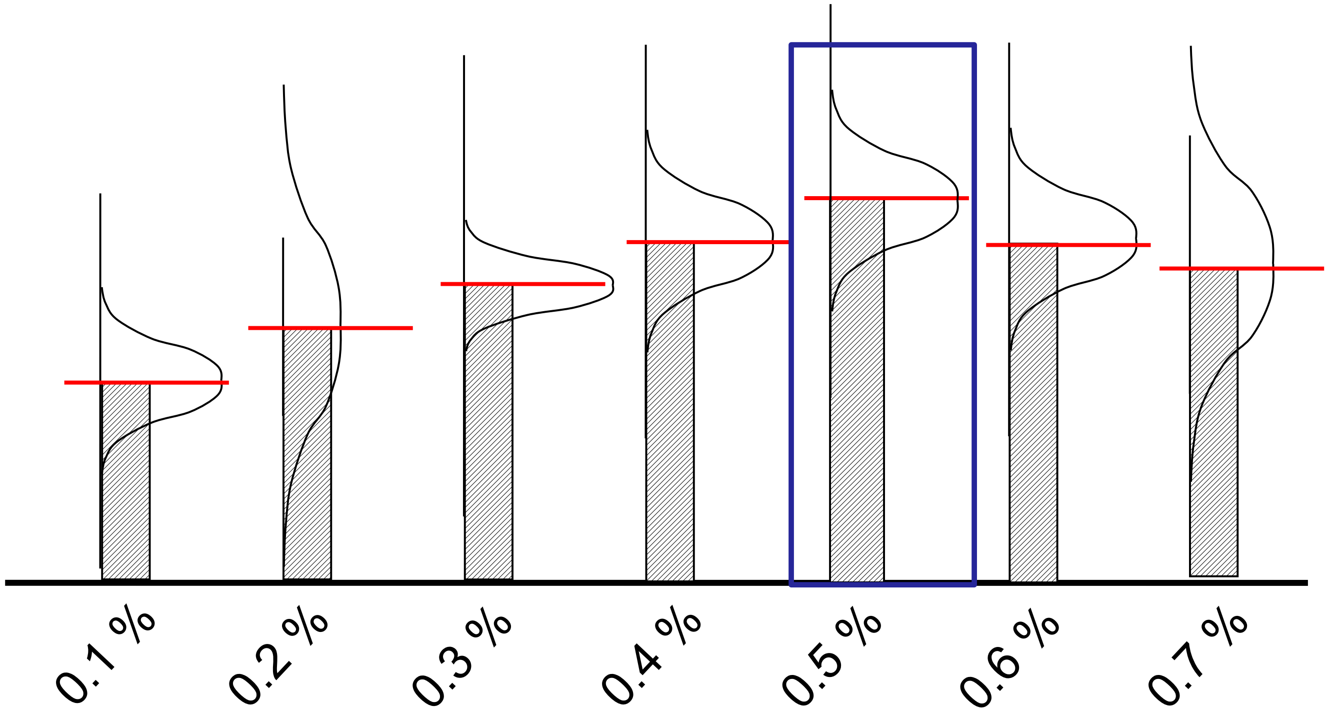


$$X^{IE}(S^n \mid \theta^{IE}) = \arg\max_x \left(\bar{\mu}_x^n + \theta^{IE} \bar{\sigma}_x^n \right)$$

» Thompson sampling

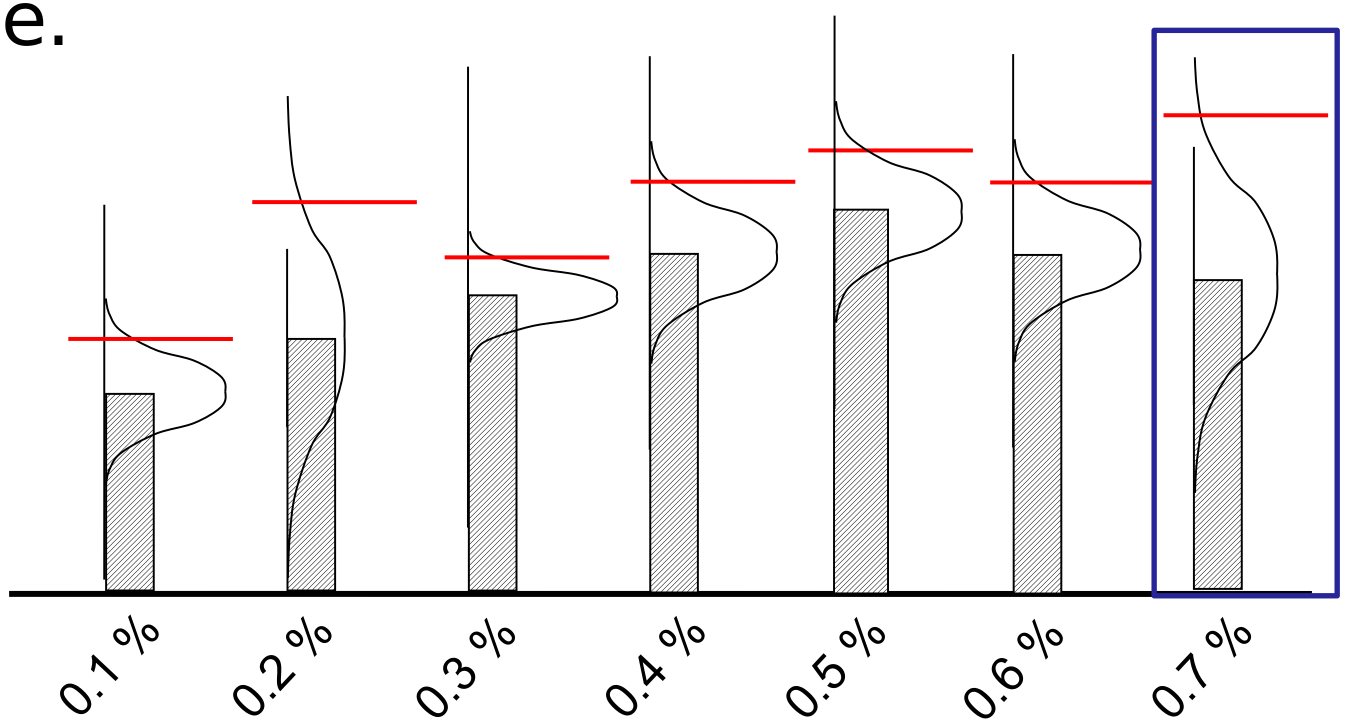
Cash management problem

Picking means we are evaluating each choice at the mean.



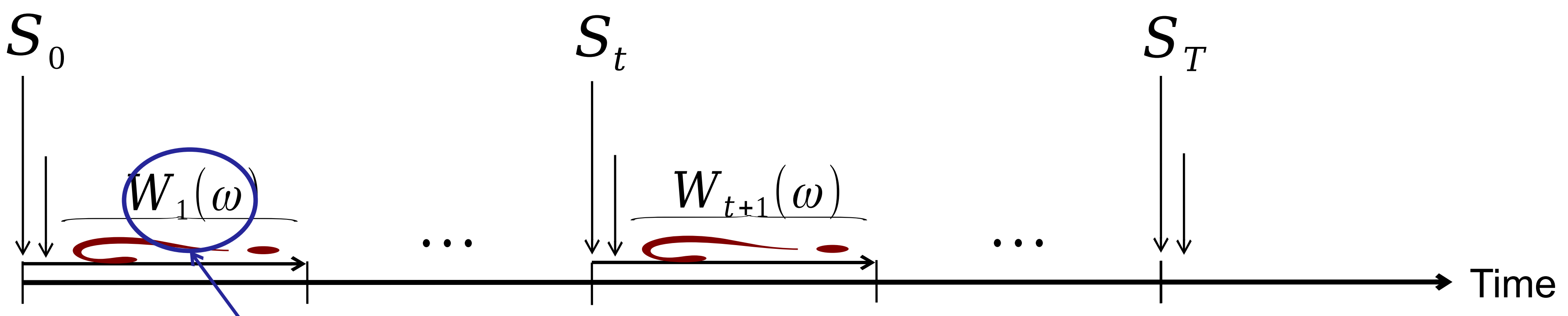
Cash management problem

Picking means we are evaluating each choice at the 95^{th} percentile.



Cash management problem

Evaluating performance over time:



$$\hat{F}^\pi(\omega) = C(S_0(\omega), x_0) + \cdots + C(S_t(\omega), x_t) + \cdots + C(S_T(\omega), x_T)$$

Deposits, withdrawals, market returns, interest rates, …

$$\hat{F}^\pi(\theta,\omega) = \sum_{t=0}^{T} C\left(S_t(\omega), X^\pi(S_t(\omega) \vee \theta)\right)$$ performance of policy over time following sample path where

$$\bar{F}^\pi(\theta) = \frac{1}{N}\sum_{n=1}^{N}\sum_{t=0}^{T} C\left(S_t(\omega^n), X^\pi(S_t(\omega^n) \vee \theta)\right)$$ average performance of policy over samples

Policy consists of:
- The use of a linear function
- The choice of parameter

Cash management problem

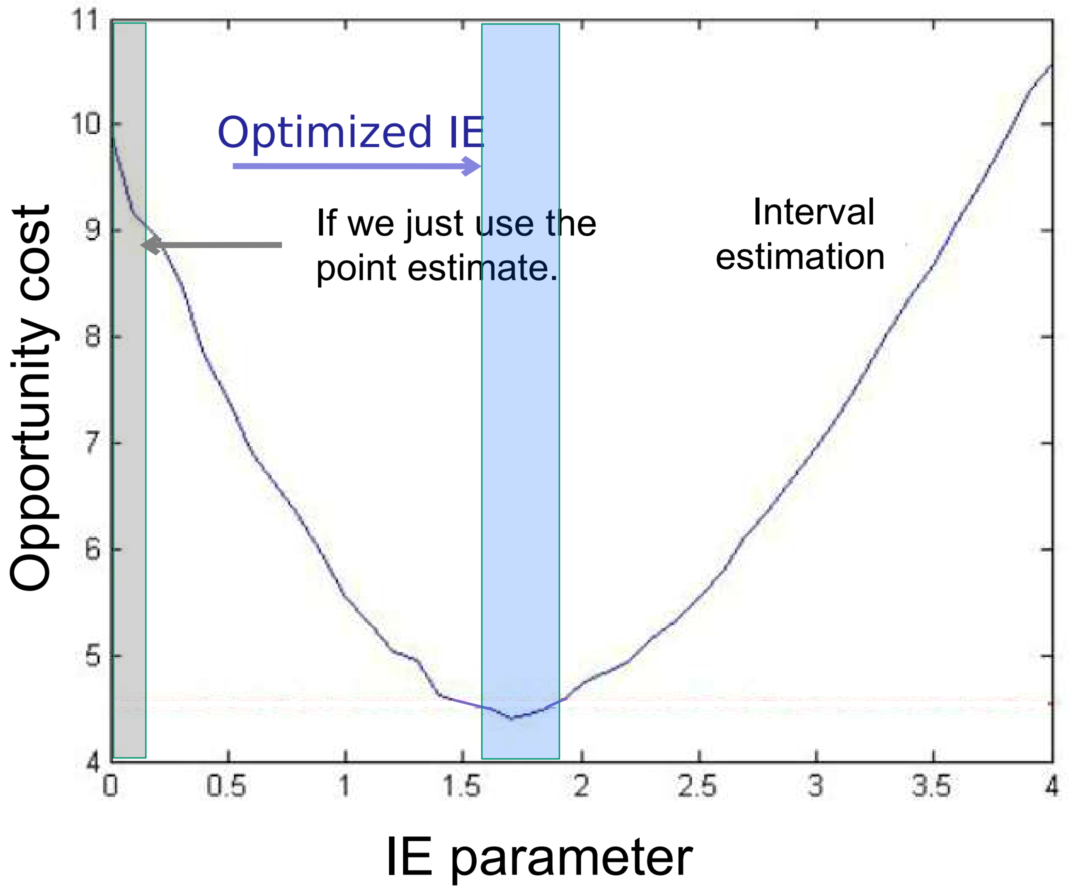
- Optimizing the IE policy
  - » We optimize to maximize:

$$\max_{\theta^{IE}} F(\theta^{IE}) = \mathbb{E} F\left(x^{\pi,N}, W\right)$$

where

$$x^n = X^{IE}(S^n \vee \theta^{IE}) = \arg\max_x \left(\bar{\mu}^n_x + \theta^{IE}\bar{\sigma}^n_x\right)$$

- Notes:
  - » Interval estimation is a form of cost function approximation (CFA)



The price of simplicity is tunable parameters … *and tuning is hard!*

Th

Professor – The newsvendor inventory problem we learned in class looks like a way to optimize how much cash I should hold in the mutual fund I manage.

The problem

Mutual funds hold a certain percentage of their assets in cash to meet redemption requests from investors.

We can start by using their existing policy…

$$X^{\pi}(S_t|\theta)=\theta^{retail}\cdot assets_t^{retail}+\theta^{inst}\cdot assets_t^{inst}$$

… which can be tuned in a simulator or the field:

$$\max_{\theta^{asset}}\mathbb{E}\left\{\sum_{t=0}^{T}C\left(S_t,X^{\pi}(S_t|\theta^{asset})\right)|S_0\right\} \rightarrow$$

Computer simulation

Field experience

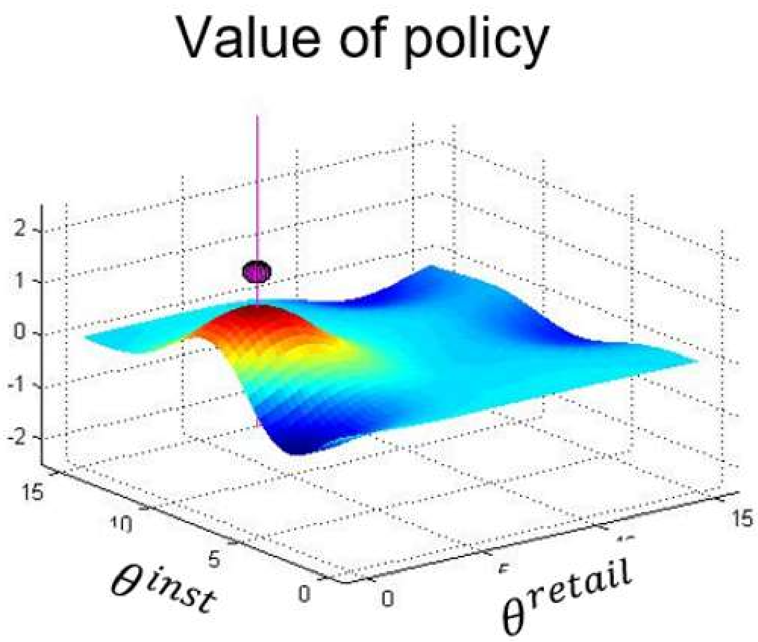
Computer simulations can be performed using historical data. Simulations, as with field testing, can be completely general. There is no need for simplifying assumptions.

Th

Professor – The newsvendor inventory problem we learned in class looks like a way to optimize how much cash I should hold in the mutual fund I manage.

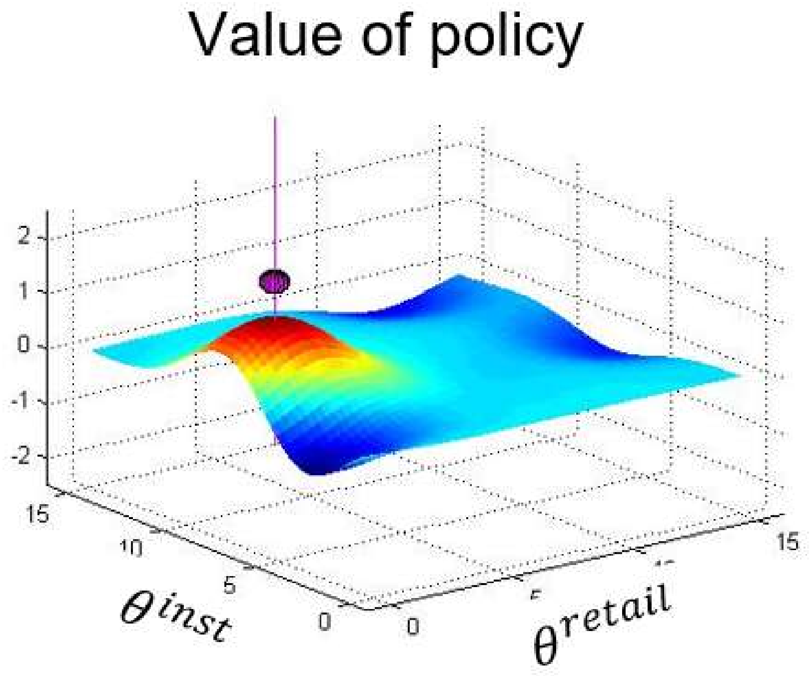
The problem

Mutual funds hold a certain percentage of their assets in cash to meet redemption requests from investors.



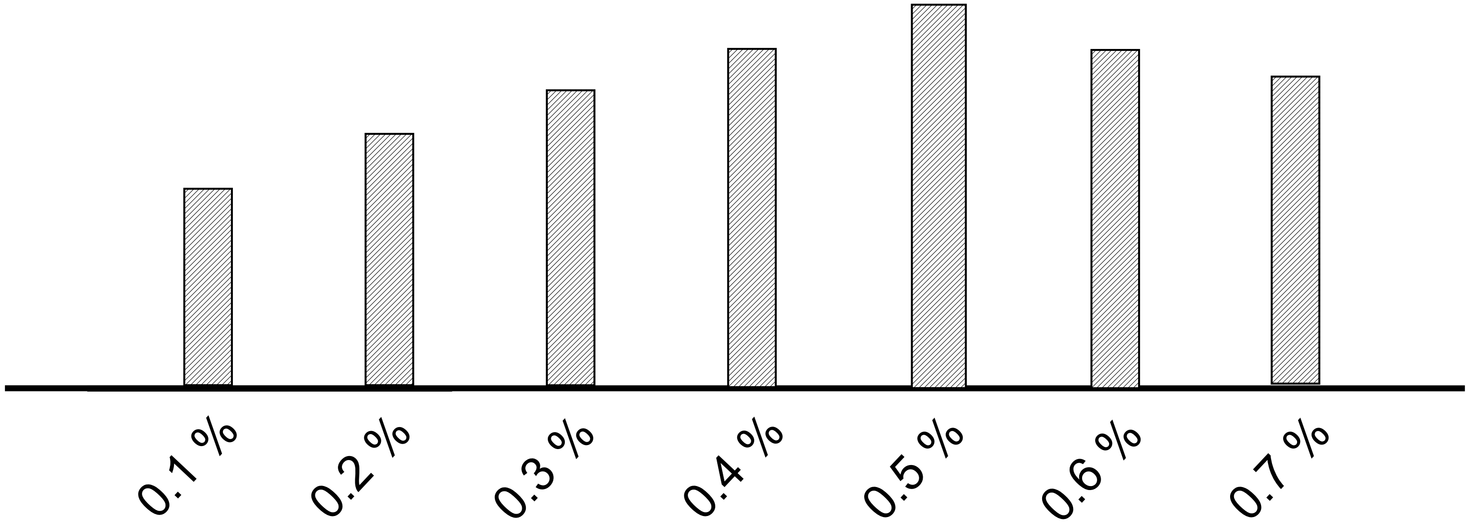
Cash management problem

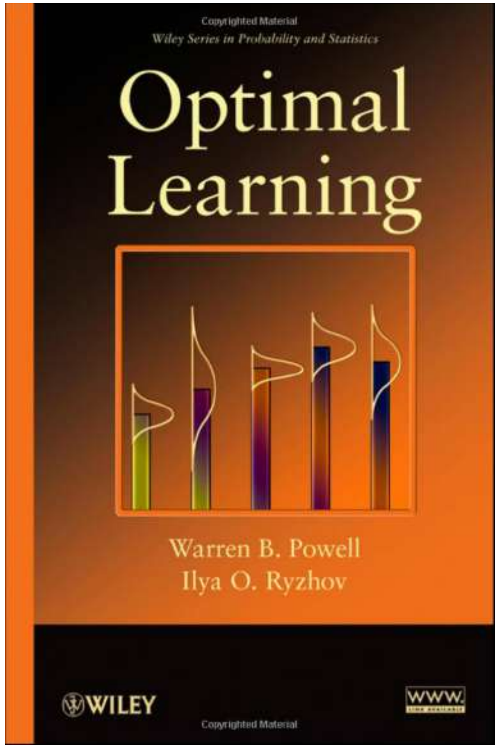
- To find we have to search over a two-dimensional, discretized space.
- A form of direct lookahead policy, called the *knowledge gradient*, estimates the value of information from an experiment.



Cash management problem

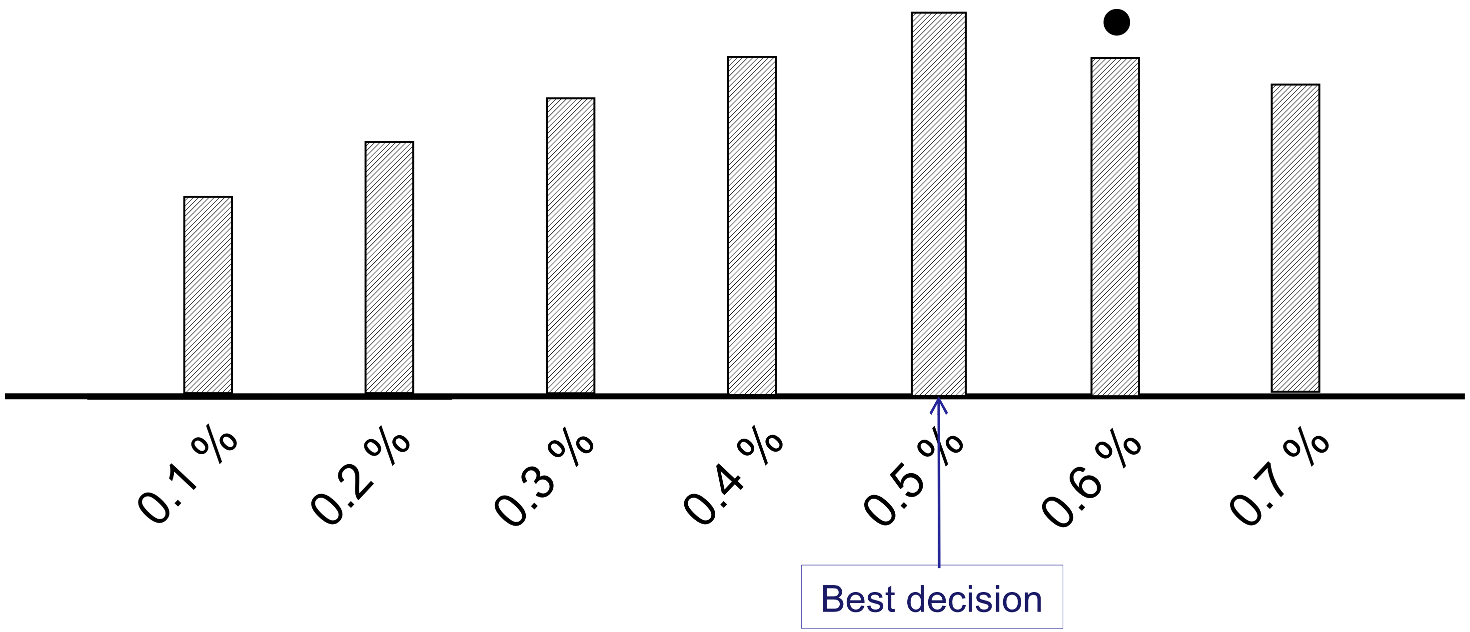
We have some (possibly large) set of choices, with uncertain beliefs about how well each performs.

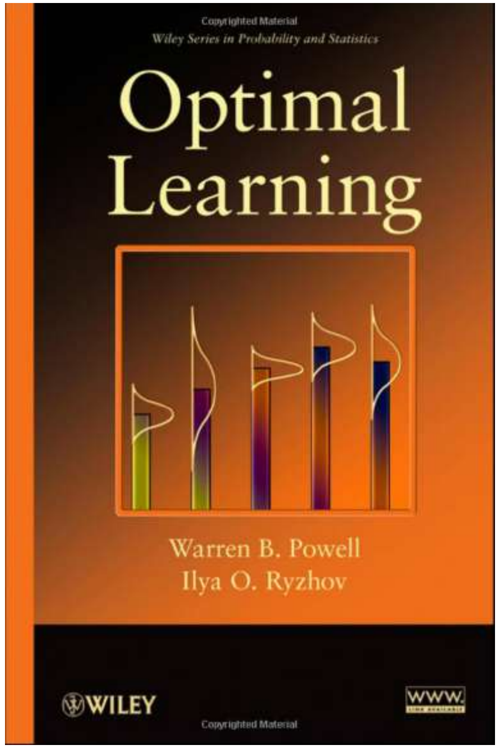




Cash management problem

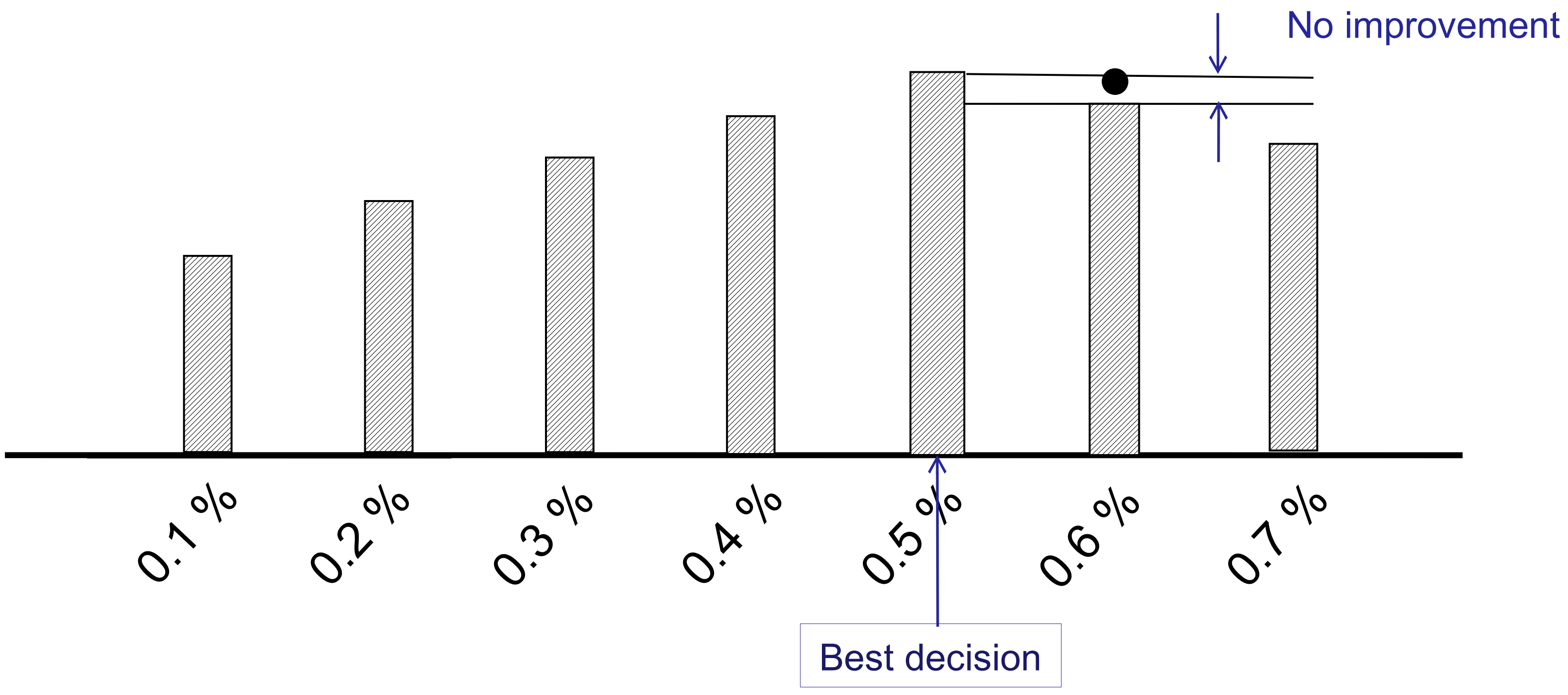
We have some (possibly large) set of choices, with uncertain beliefs about how well each performs.

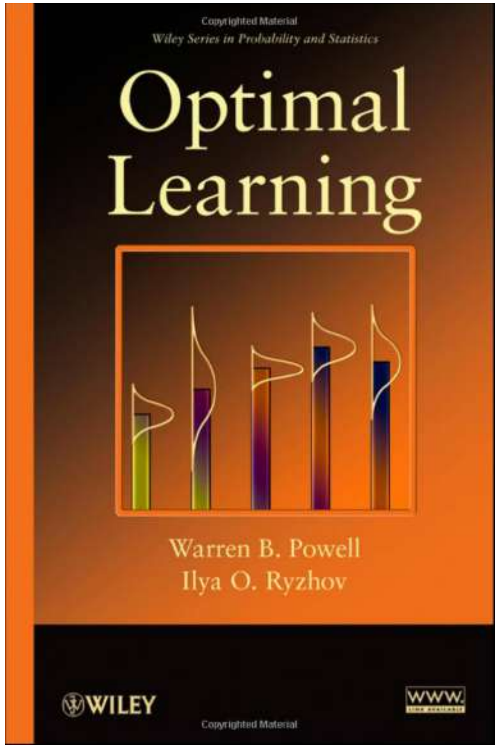




Cash management problem

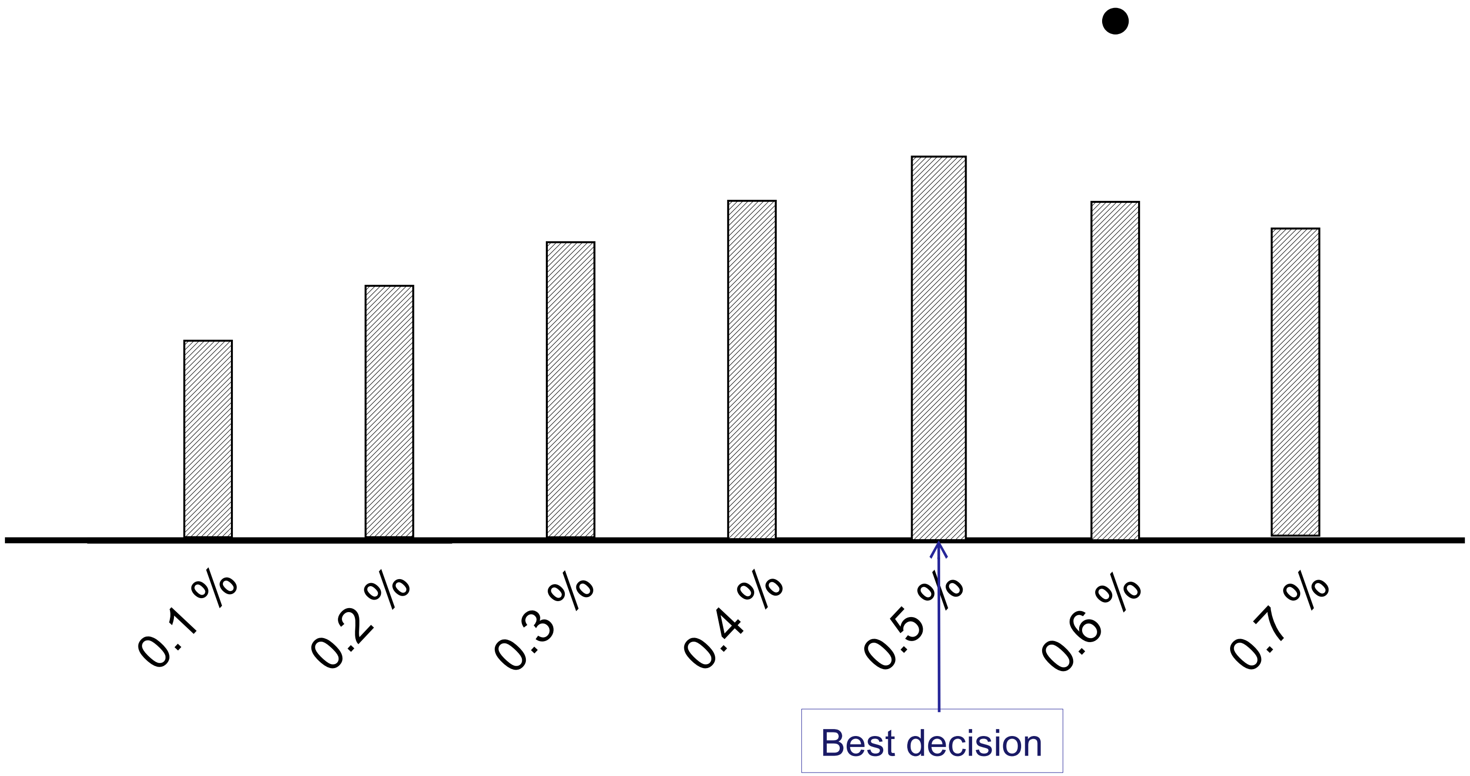
We have some (possibly large) set of choices, with uncertain beliefs about how well each performs.

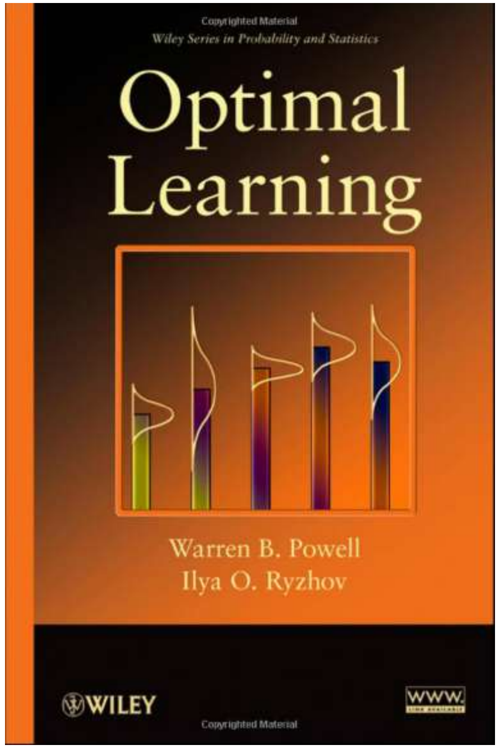




Cash management problem

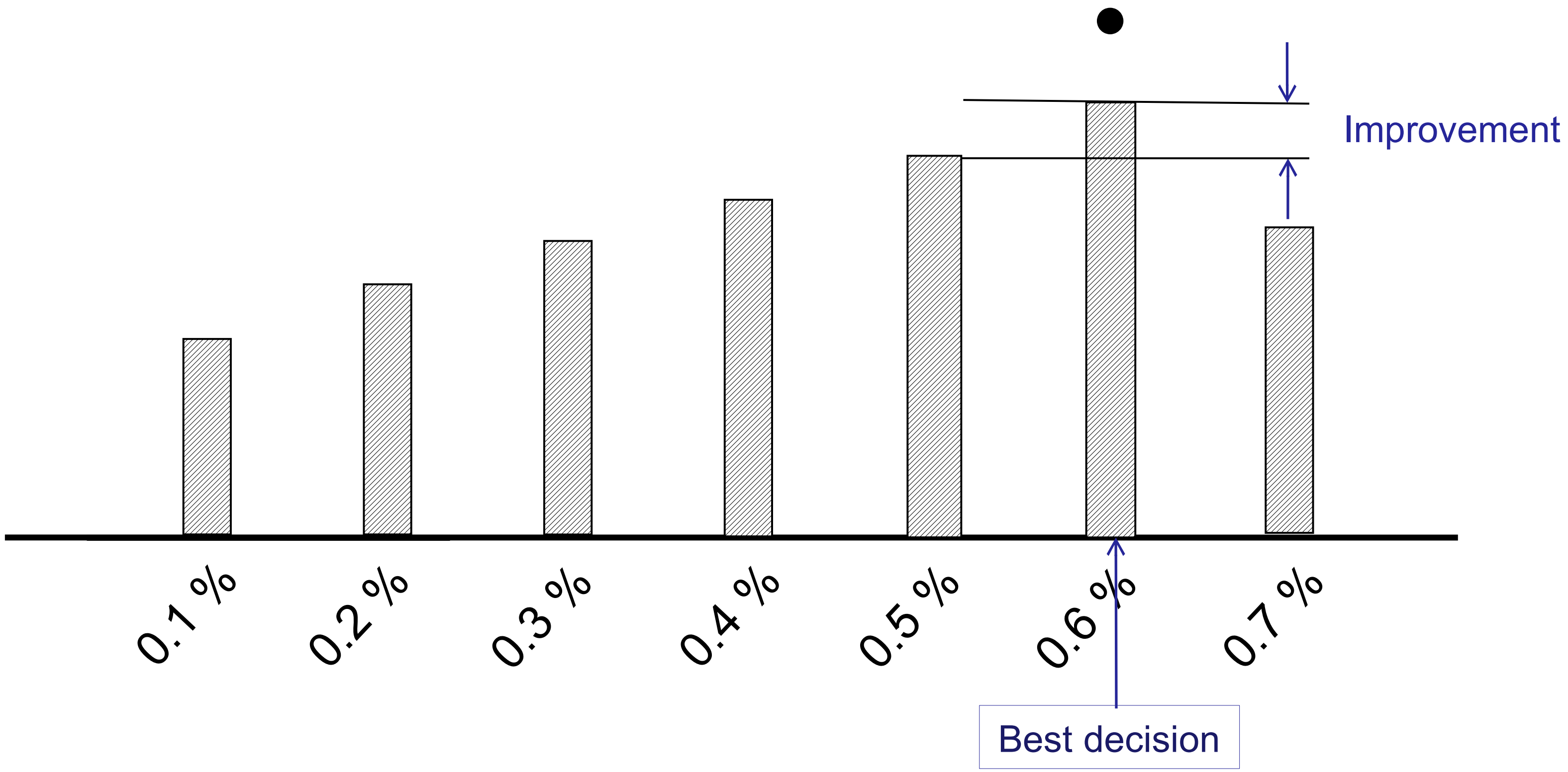
We have some (possibly large) set of choices, with uncertain beliefs about how well each performs.

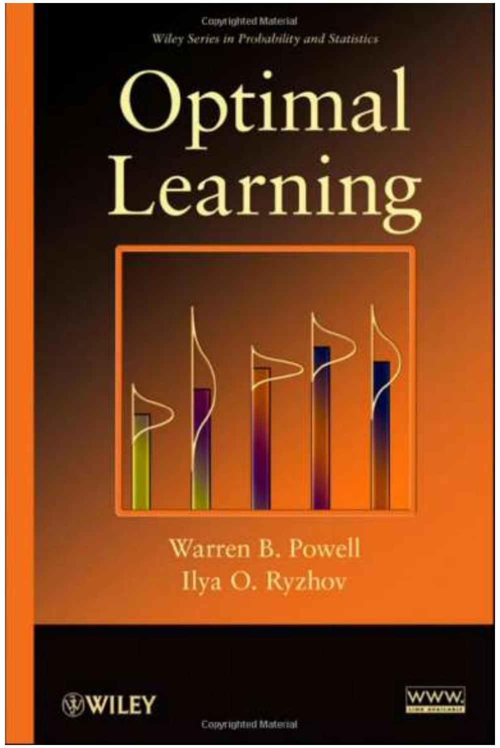




Cash management problem

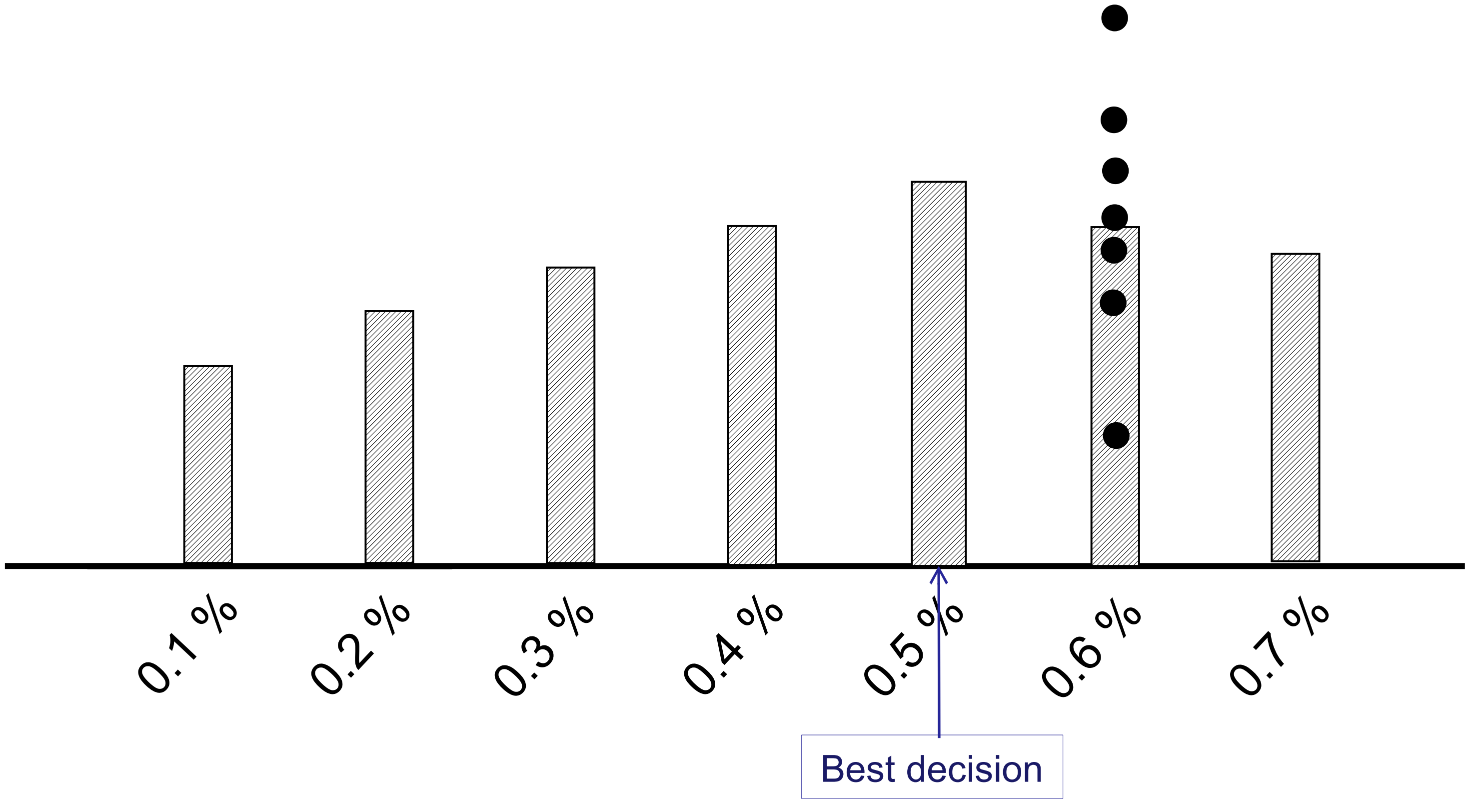
We have some (possibly large) set of choices, with uncertain beliefs about how well each performs.

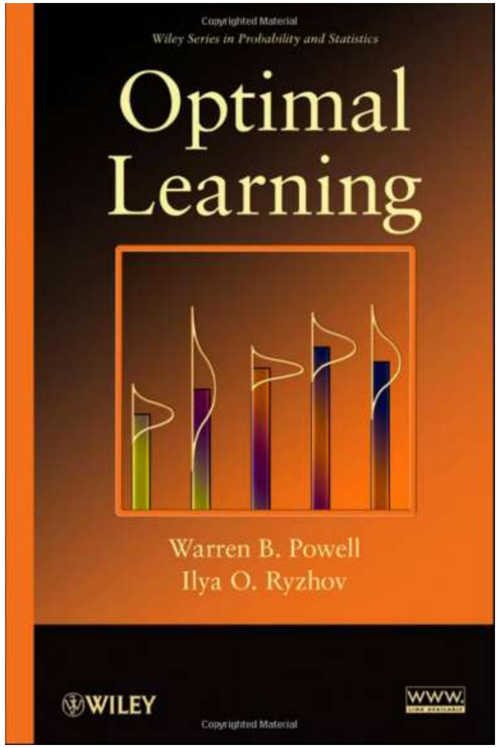




Cash management problem

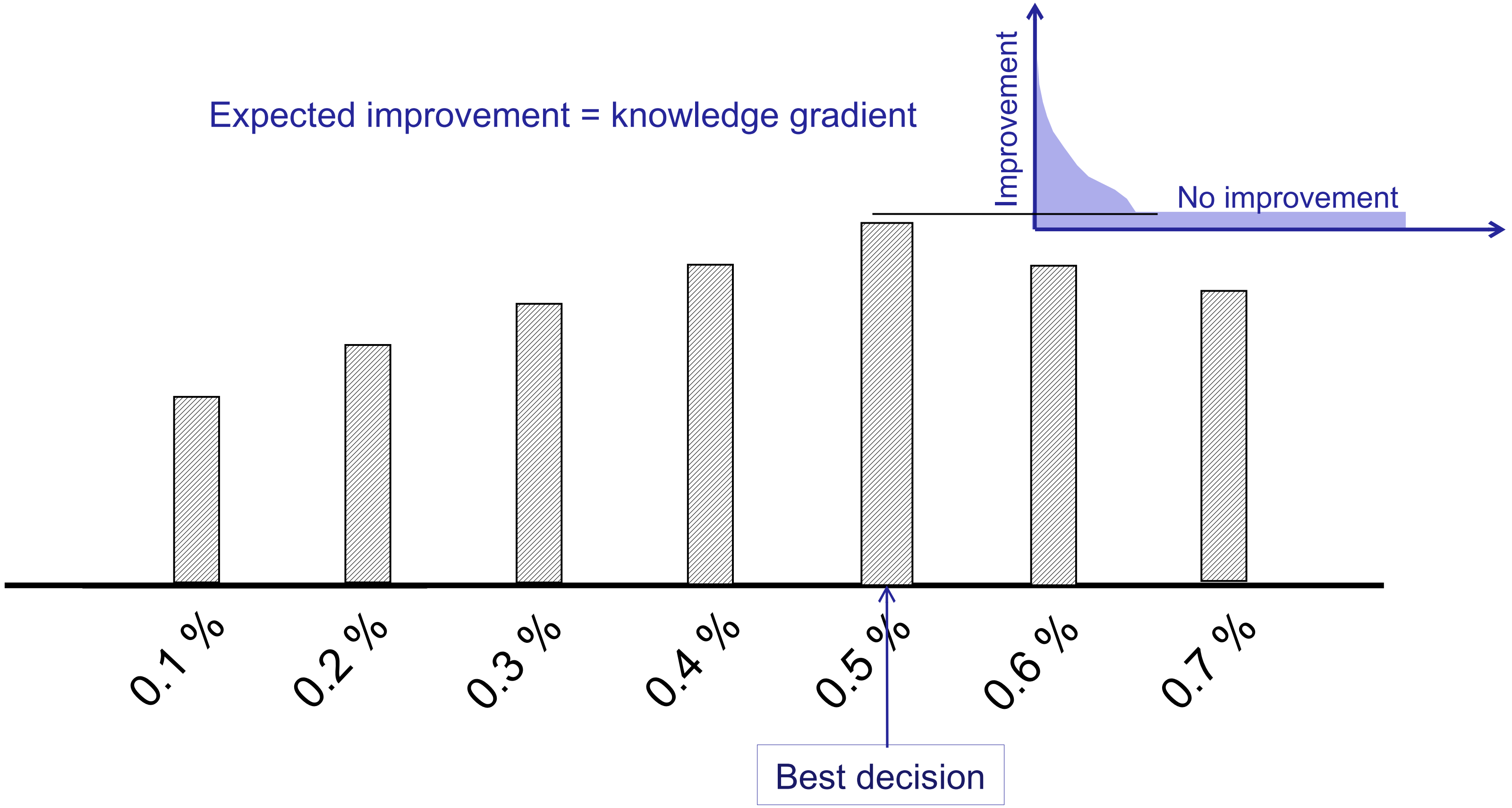
We have some (possibly large) set of choices, with uncertain beliefs about how well each performs.

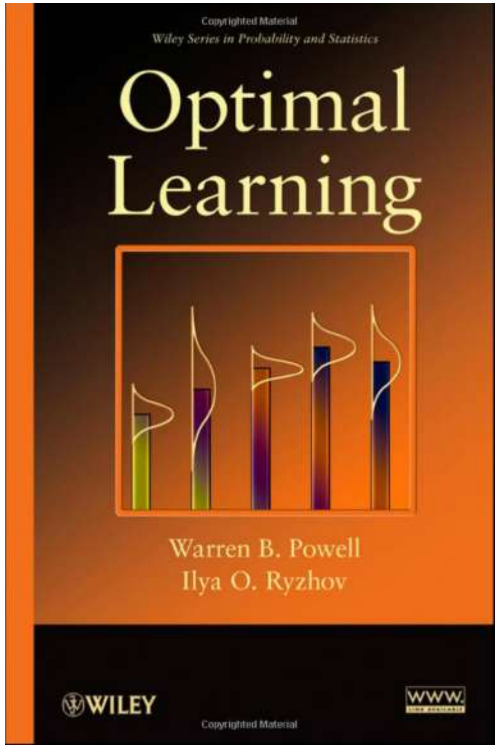




Cash management problem

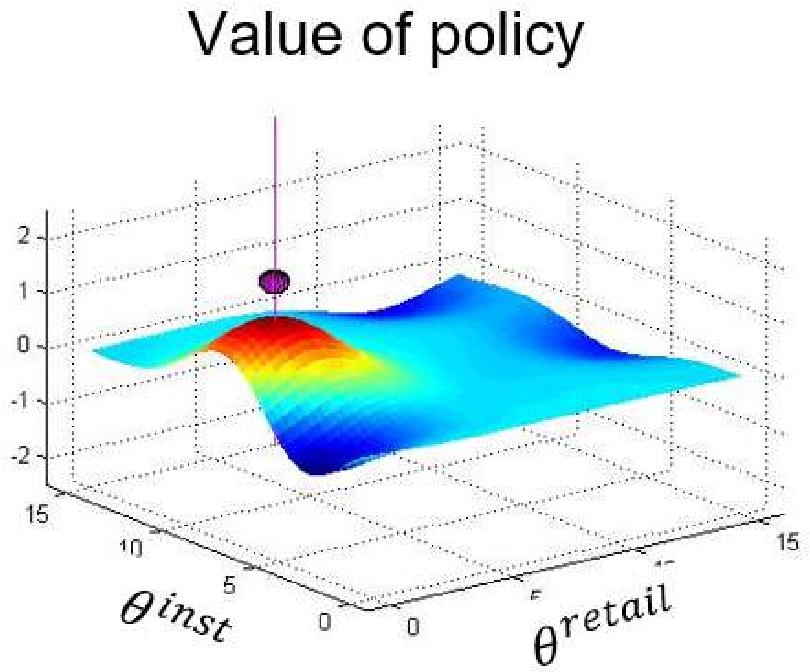
We have some (possibly large) set of choices, with uncertain beliefs about how well each performs.





Cash management problem

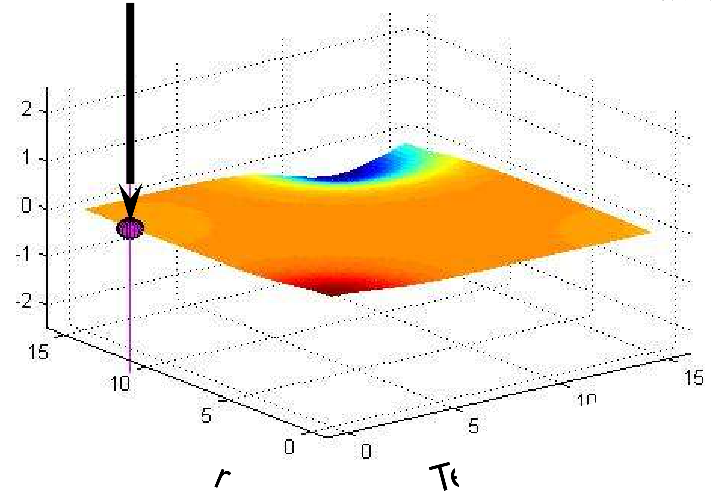
- Assume we discretize and into 50 intervals, we would have 2500 points to search over.
- We can easily calculate the knowledge gradient for each of the 2500 points, and then choose the point to observe that has the highest KG value.



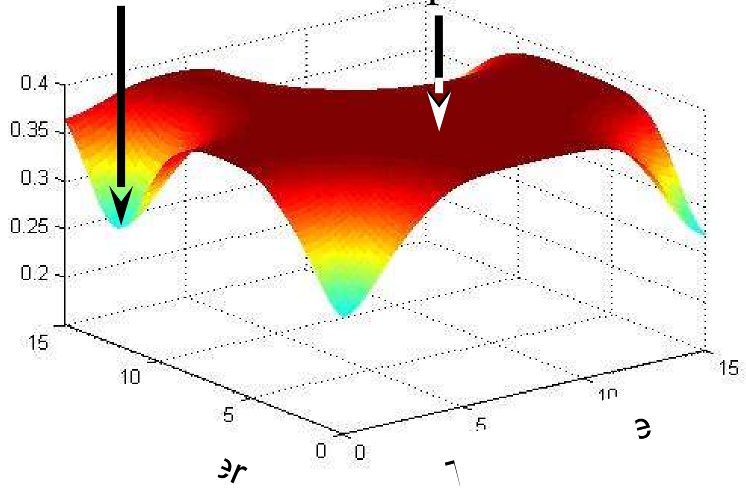
Cash management problem

- After four experiments:

Estimated surface



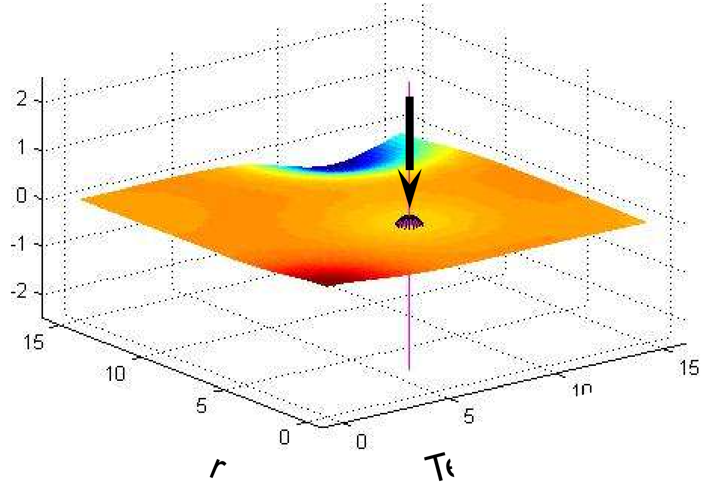
Knowledge gradient



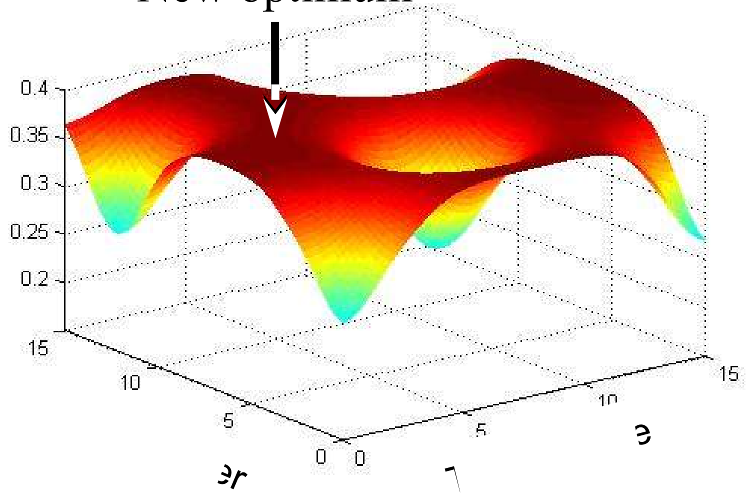
Cash management problem

- After five experiments:

Estimated surface



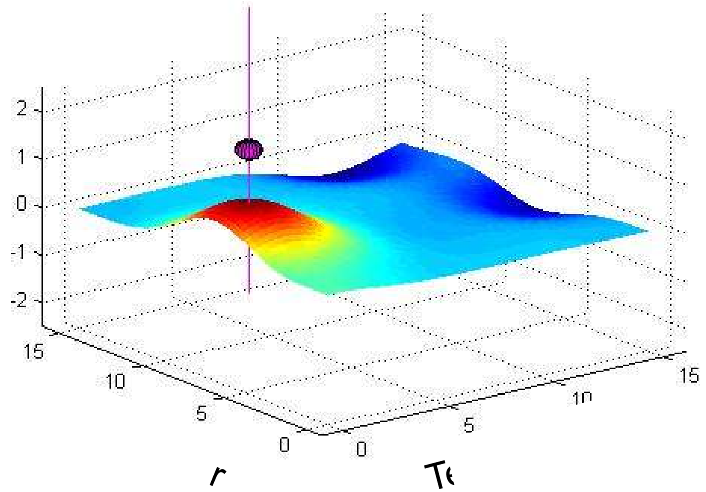
Knowledge gradient



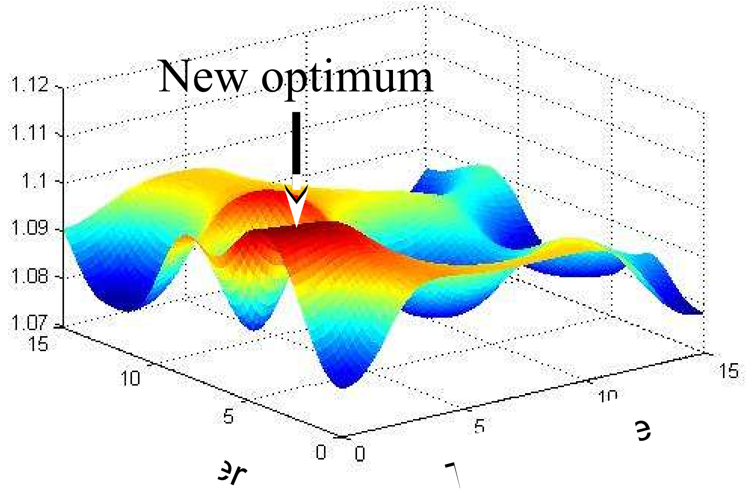
Cash management problem

- After seven experiments:

Estimated surface



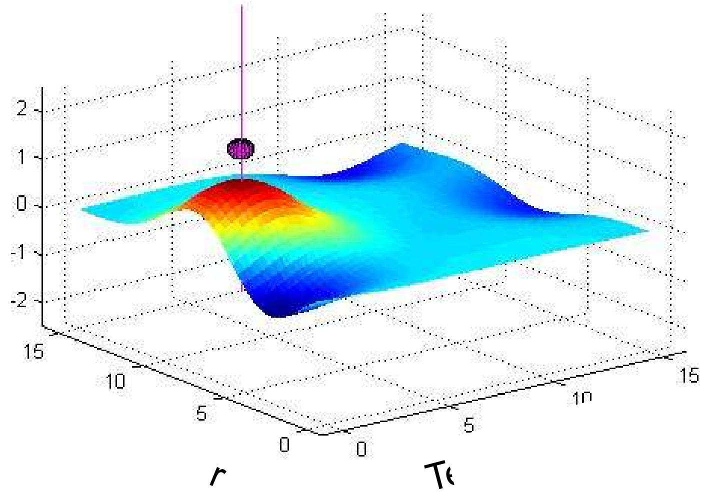
Knowledge gradient



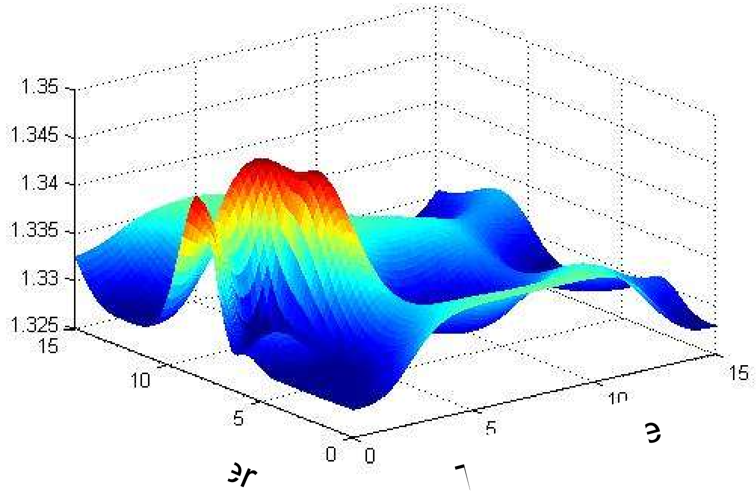
Cash management problem

- After nine experiments:

Estimated surface



Knowledge gradient



After nine experiments, we can identify the best out of 2500 combinations of wafer diameter and the temperature and the heating process.

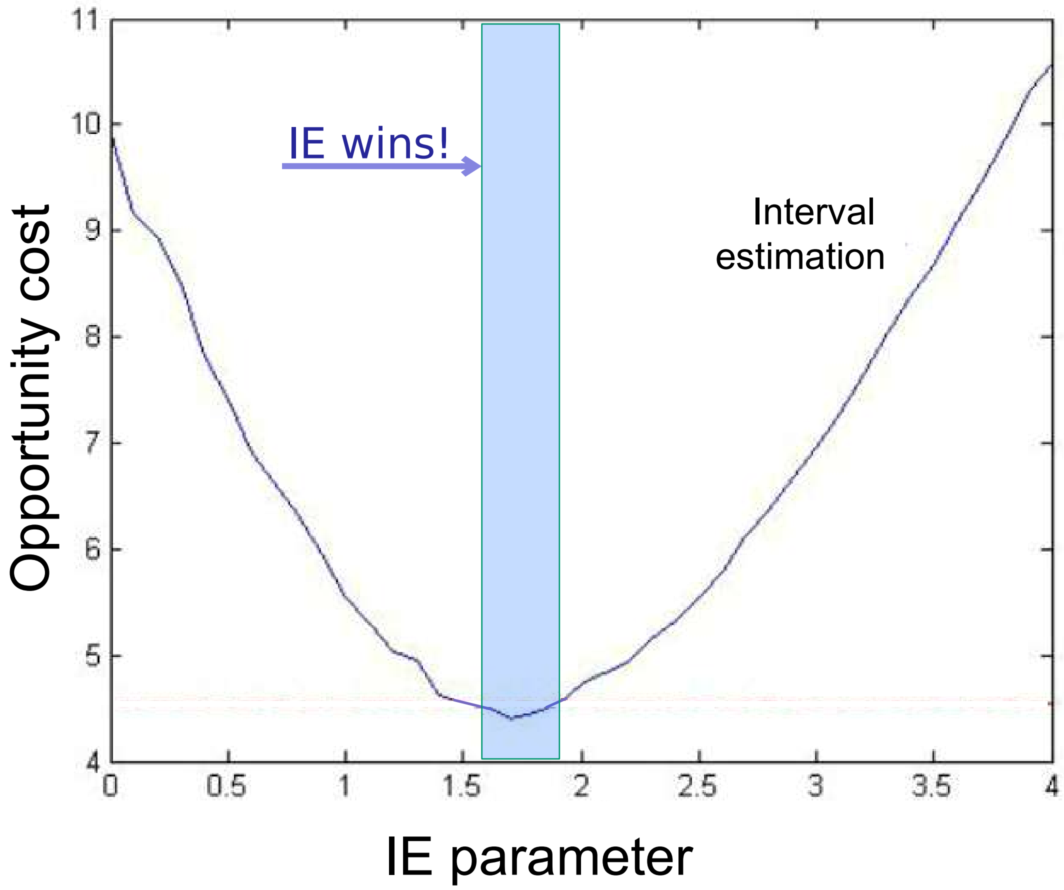
Cash management problem

- Optimizing the IE policy
  - » We optimize to maximize:

$$\max_{\theta^{IE}} F(\theta^{IE}) = \mathbb{E} F\left(x^{\pi,N}, W\right)$$

where

$$x^n = X^{IE}_{\square}(S^n \vee \theta^{IE}) = \arg\max_x \left(\bar{\mu}^n_x + \theta^{IE}\bar{\sigma}^n_x\right)$$



- Notes:
  - » Interval estimation is a form of cost function approximation (FAA)
  - » Knowledge gradient is a form of direct lookahead approximation (DLA), but requires no tuning.

The price of simplicity is tunable parameters … *and tuning is hard!*

Designing policies

Optimizing requires solving nested search problems:

» Searching for the best cash management policy

- PFA? Linear decision rules:

$$X^{\pi}\left(S_t \middle| \theta^{asset}\right) = \theta^{asset} \cdot asset\, s_t$$

$$X^{\pi}\left(S_t \middle| \theta^{asset}\right) = \theta^{retail} \cdot asset\, s_t^{retail} + \theta^{inst} \cdot asset\, s_t^{inst}$$

- CFA? VFA (Bellman)? DLA (stochastic programming)?

» Searching for the best parameter . Search policy:

- CFA (Interval estimation?)
- DLA (knowledge gradient?)

These policies have their own tunable parameters

Designing policies

Optimizing requires solving nested search problems:

» Searching for the best search policy
- CFA?
  - Interval estimation
- DLA?
  - Knowledge gradient

» Searching for the best search parameter :
- PFA? Stochastic gradient algorithm
- CFA? (UCB, Interval estimation, Thompson sampling)
- DLA? (Knowledge gradient, expected improvement)

These policies have their own tunable parameters

Cash management problem

- Notes:
  - » The underlying dynamic model can be completely general.
  - » We can simulate any stochastic process for redemptions, deposits, and the evolution of the market.

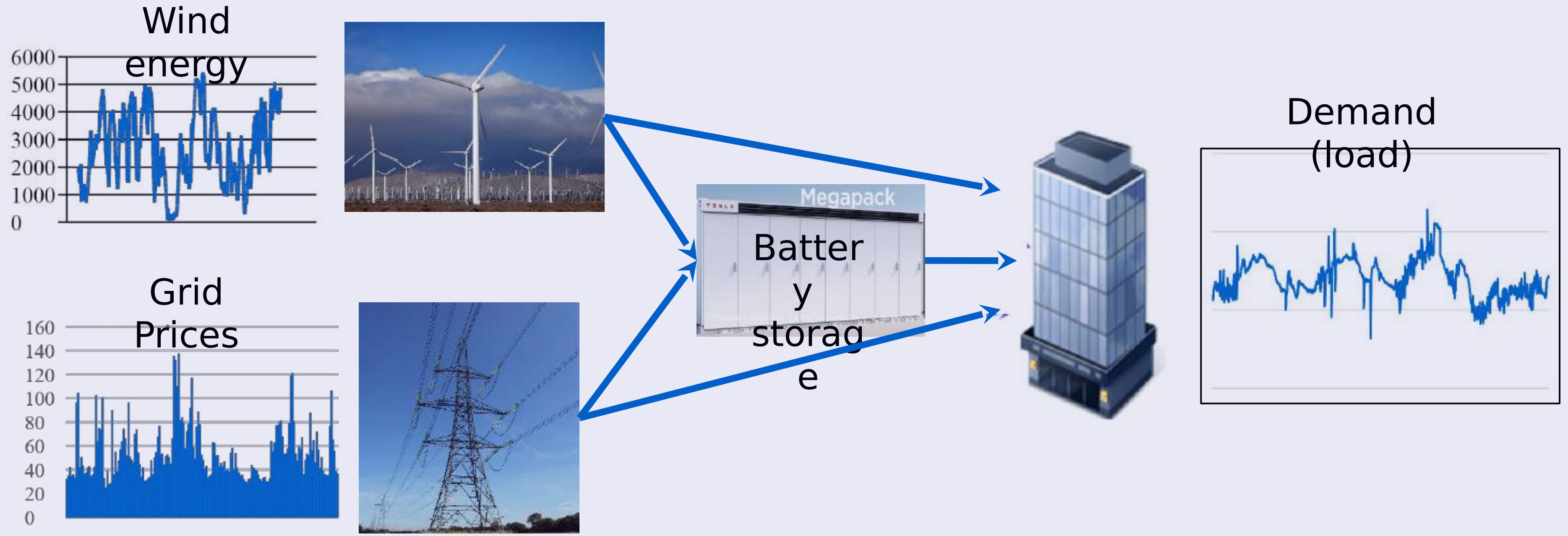
- Future research:
  - » We can try any of a range of parametric and nonparametric models for the policy:
    - Tree regression
    - Neural networks
  - » We need additional research to guide the process of learning in the field, where policy search is slow, noisy, and costs are cumulative.

OUTLINE

- ➜The seven levels of artificial intelligence
- ➜The universal modeling framework
- ➜Designing policies
- ➜Mutual fund cash balance optimization
- ➜Choosing the best policy
- ➜A new educational field: sequential decision analytics

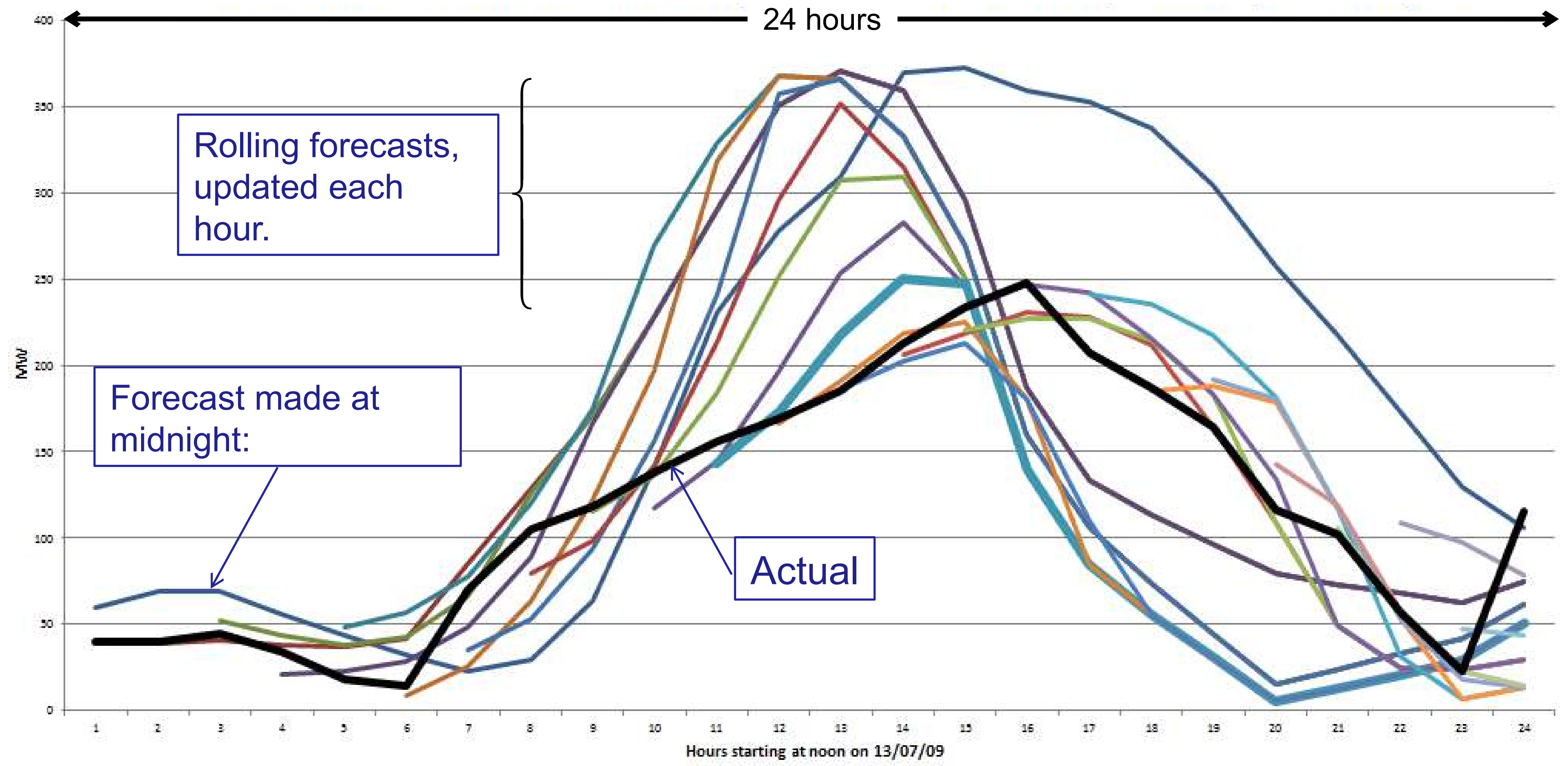
An energy storage application

Consider a basic energy storage problem



An energy storage application

Forecasts evolve over time as new information arrives:

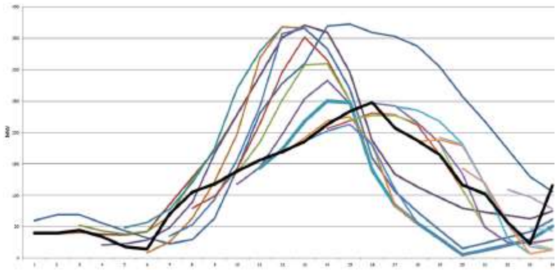


An energy storage application

- Benchmark policy – Deterministic lookahead

$$X_t^{D-LA}(S_t) = \underset{x_t,(\tilde{x}_{tt'},t'=t+1,\ldots,t+H)}{\arg\min}\left(C(S_t,x_t) + \left[\sum_{t'=t+1}^{t+H} \tilde{c}_{tt'}\tilde{x}_{tt'} \right] \right)$$

$$\tilde{x}^{wd}_{tt'} + \beta \tilde{x}^{rd}_{tt'} + \tilde{x}^{gd}_{tt'} \le f^{D}_{tt'}$$



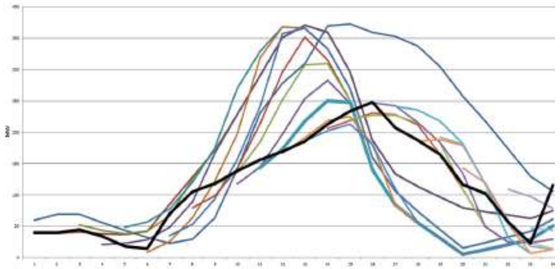
(joint with Saeed Ghadimi)

An energy storage application

- Benchmark policy – Deterministic lookahead

$$X_t^{D-LA}(S_t|\theta) = \underset{x_t,(\tilde{x}_{tt'},t'=t+1,\ldots,t+H)}{\arg\min}\left(C(S_t,x_t) + \left[\sum_{t'=t+1}^{t+H} \tilde{c}_{tt'}\tilde{x}_{tt'}\right]\right)$$

$$\mathcal{X}\left(S_t,\theta\right) = ¿\left\{ ¿\ \tilde{x}^{wd}_{tt'} + \beta \tilde{x}^{rd}_{tt'} + \tilde{x}^{gd}_{tt'} \le f^D_{tt'}\right.$$

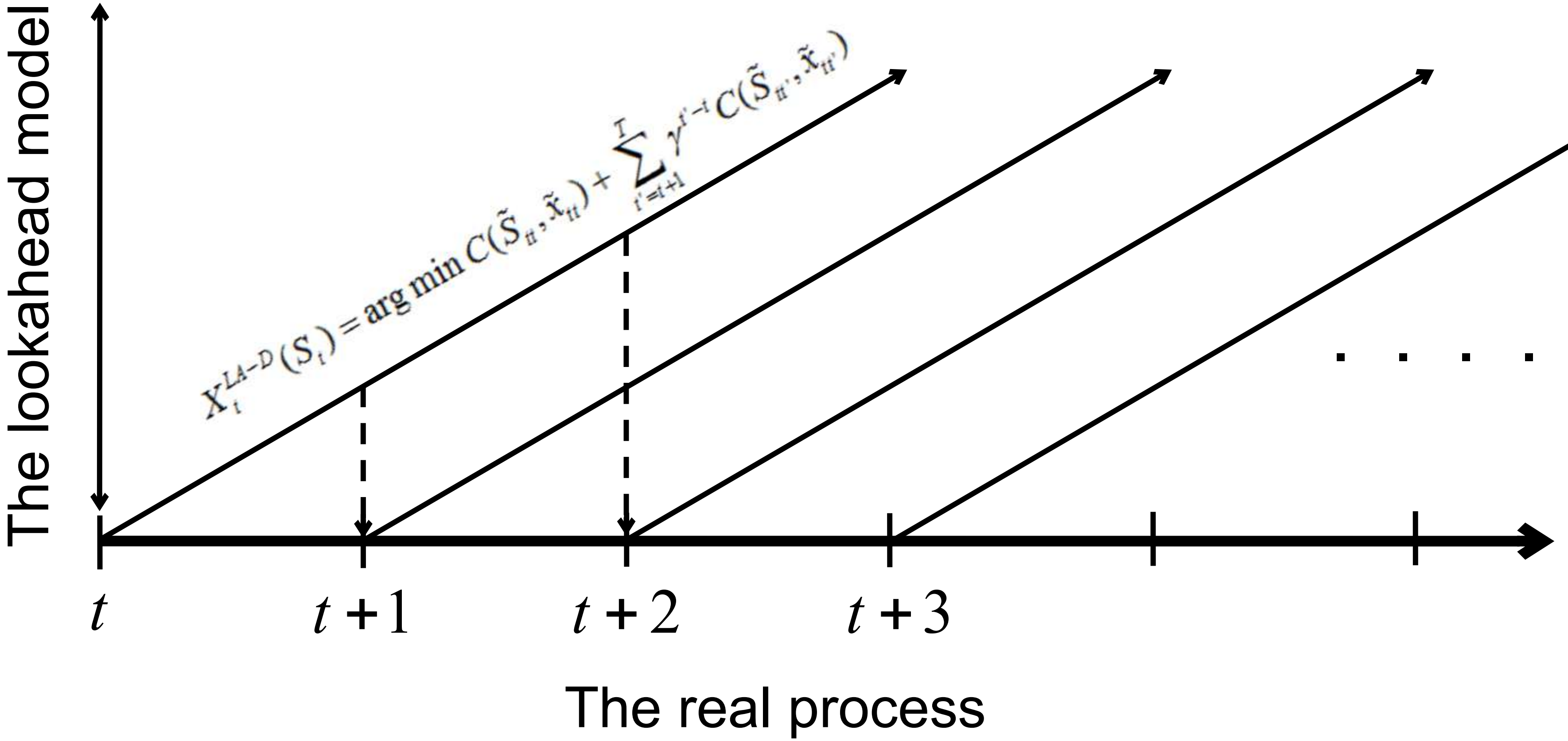


Optimal Dynamics

(joint with Saeed Ghadimi)

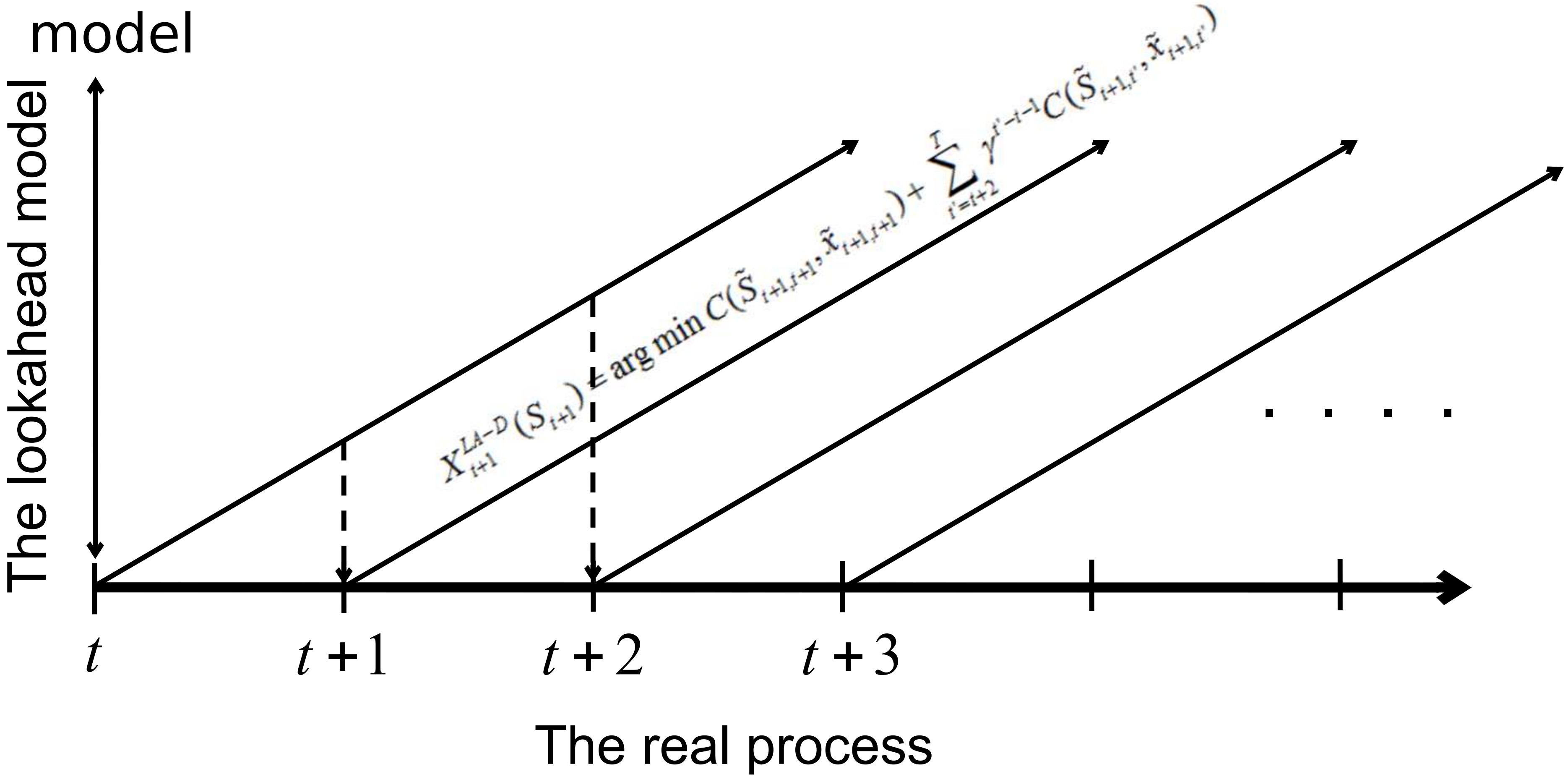
An energy storage application

- Lookahead policies peek into the future
  - » Optimize over deterministic lookahead model



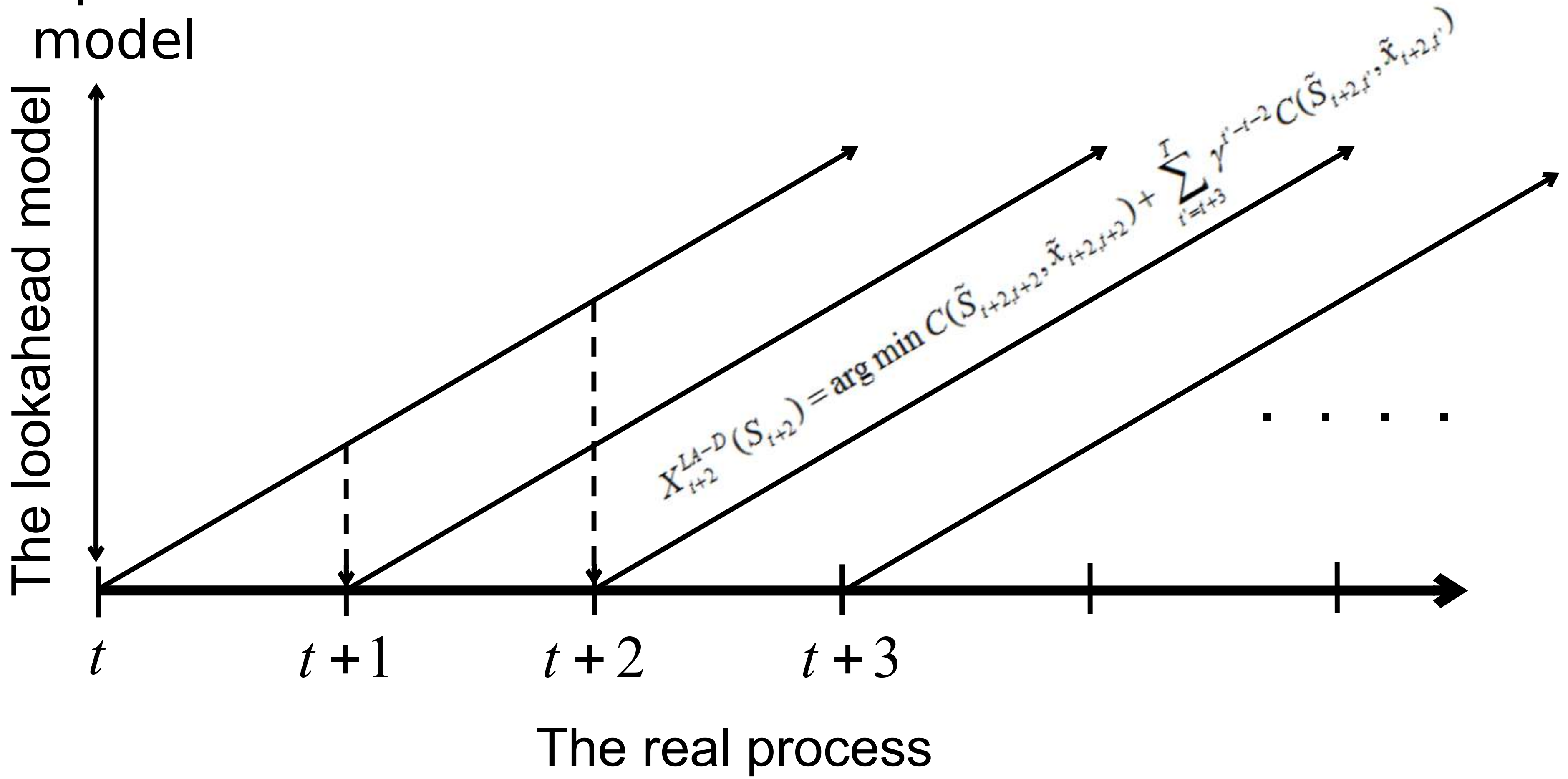
An energy storage application

- Lookahead policies peek into the future
  - » Optimize over deterministic lookahead model



An energy storage application

- Lookahead policies peek into the future
  - » Optimize over deterministic lookahead model



An energy storage application

- The policy (deterministic linear program)

$$X_t^{\pi}(S_t \vee \theta) = argmax_{x_t \in \mathscr{X}(S_t,\theta)} \sum_{t'=t}^{t+H} c_{t'} x_{t'}$$

- The objective function (evaluated using simulation)

where

- Stochastic optimization problem

$$\max_{\pi=(f,\theta)} \mathbb{E}_W F^{\pi}(\theta, W)$$

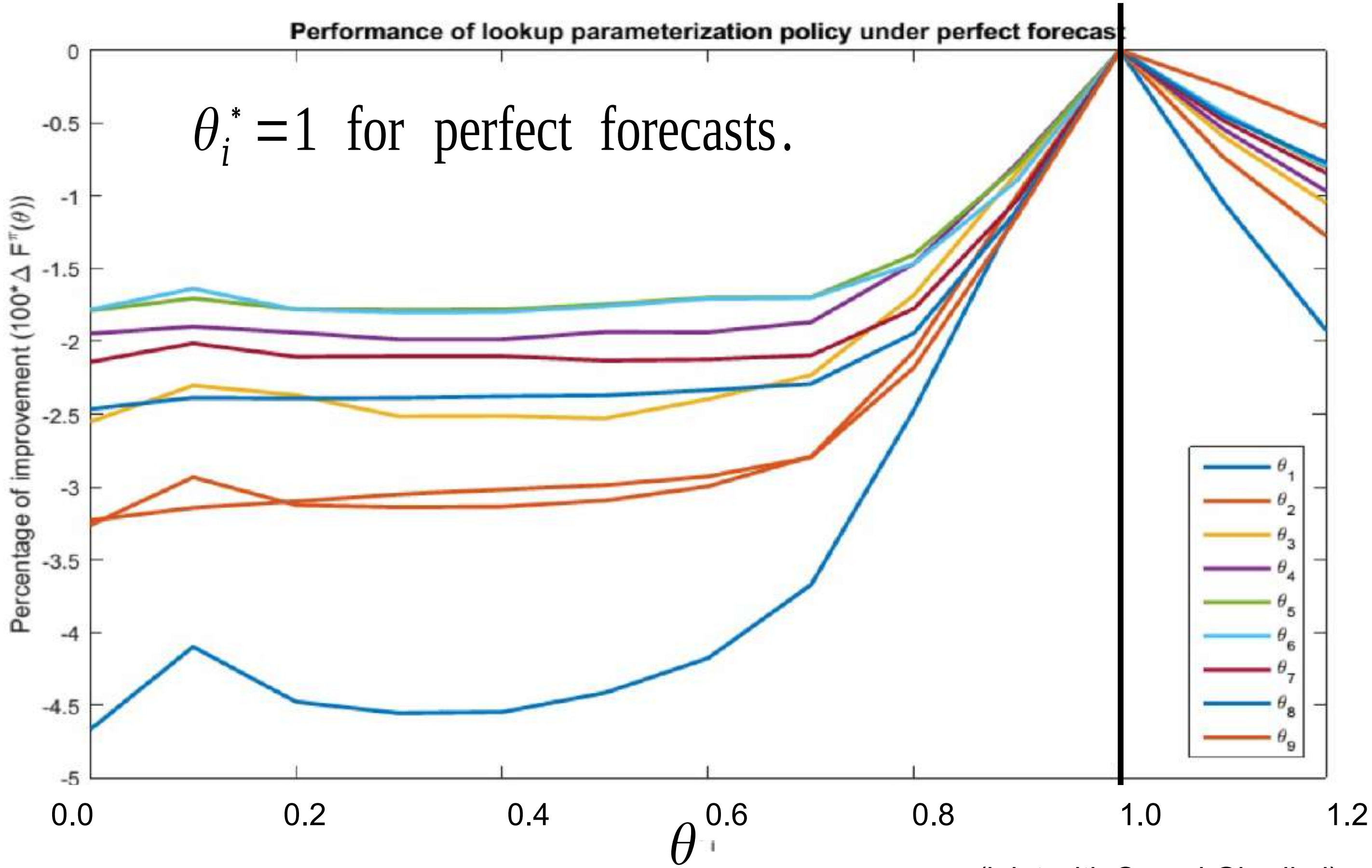
- Stochastic gradient algorithm (using Spall's SPSA algorithm)

$$\theta^{n+1} = \theta^n + \alpha_n \nabla F^{\pi}(\theta^n, \omega^{n+1})$$

(joint with Saeed Ghadimi)

An energy storage application

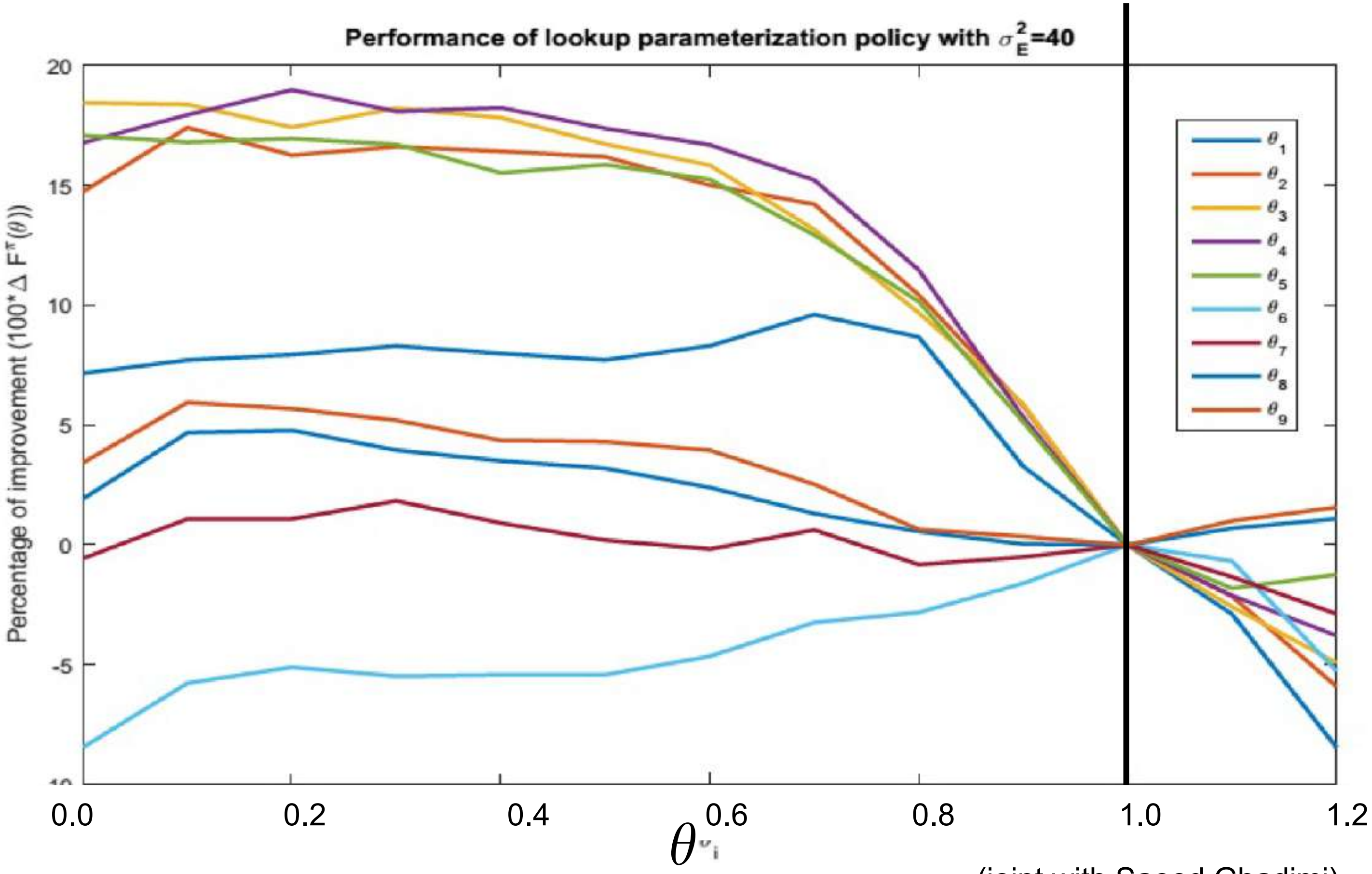
- One-dimensional contour plots – perfect forecast



(joint with Saeed Ghadimi)

An energy storage application

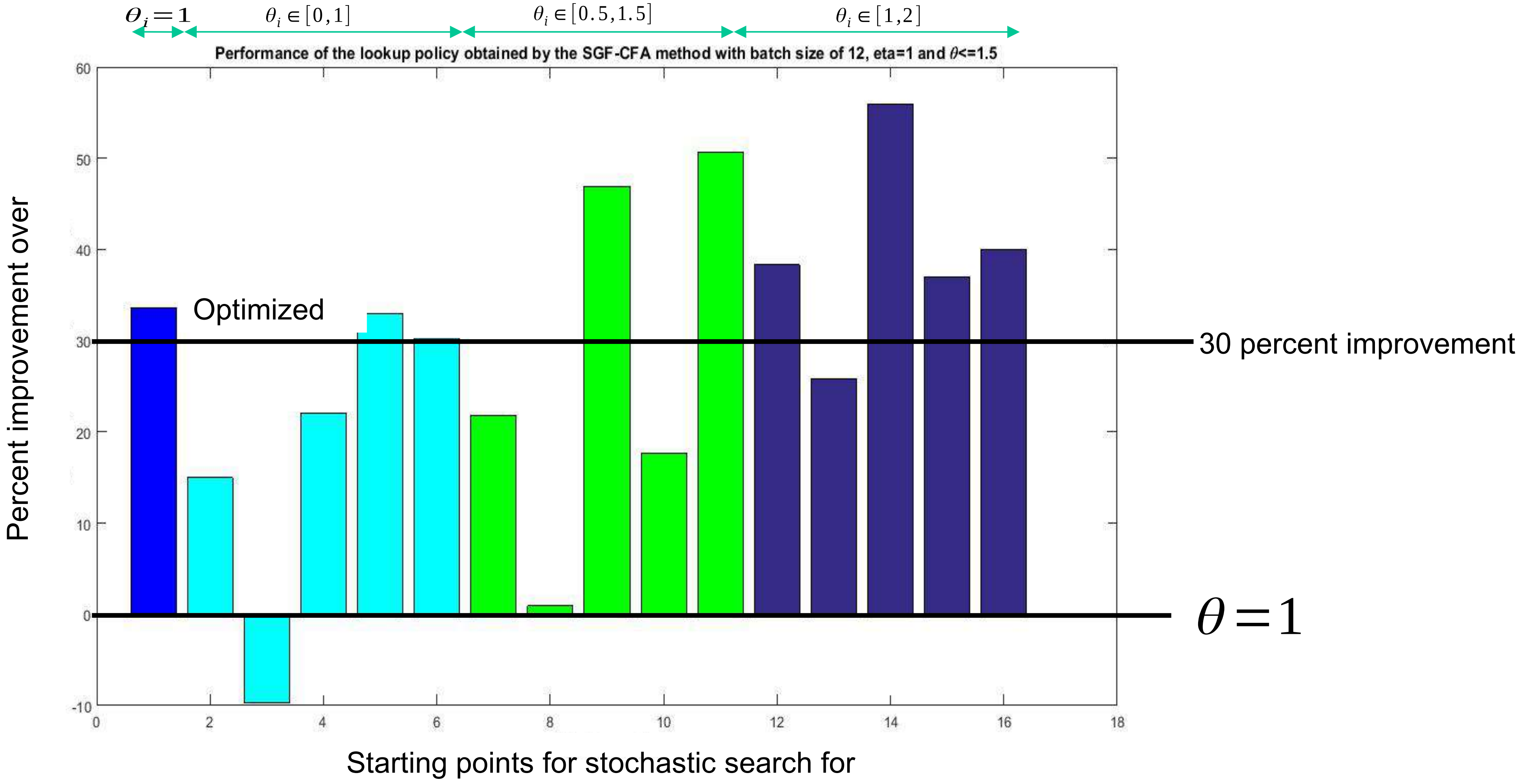
- One-dimensional contour plots - uncertain forecast



(joint with Saeed Ghadimi)

An energy storage application

- Tuning the parameters



An energy storage application

Each policy is best on certain problems

Problem:	Problem description	PFA	CFA	VFA	DLA	DLA/CFA
A	A stationary problem with heavy-tailed prices, relatively low noise, moderately accurate forecasts.	**0.959**	**0.839**	**0.936**	**0.887**	**0.887**
B	A time-dependent problem with daily load patterns, no seasonalities in energy and price, relatively low noise, less accurate forecasts.	**0.714**	**0.752**	**0.712**	**0.746**	**0.746**
C	A time-dependent problem with daily load, energy and price patterns, relatively high noise, forecast errors increase over horizon.	**0.865**	**0.590**	**0.914**	**0.886**	**0.886**
D	A time-dependent problem, relatively low noise, very accurate forecasts.	**0.962**	**0.749**	**0.971**	**0.997**	**0.997**
E	Same as (C), but the forecast errors are stationary over the planning horizon.	**0.865**	**0.590**	**0.914**	**0.922**	**0.934**

Joint research with Prof. Stephan Meisel, University of Twente, Netherlands.

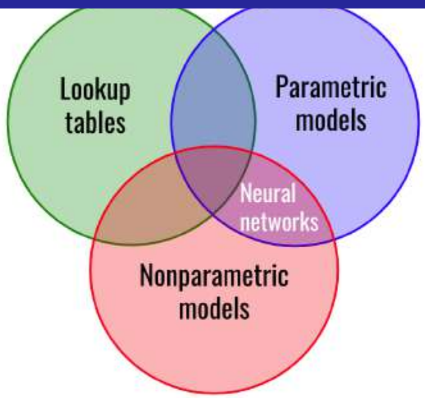
BRIDGING MACHINE LEARNING & SEQUENTIAL DECISIONS

Machine learning

$$\min_{f\in F,\theta\in\Theta^f} \frac{1}{N}\sum_{n=1}^{N}\left(y^n - f(x^n|\theta)\right)^2$$

Searching over functions

"Big dataset"



Sequential decisions

$$\max_{\pi=(f\in F,\theta\in\Theta^f)} \frac{1}{N}\sum_{n=1}^{N}\sum_{t=0}^{T} C(S_t^n, X^\pi(S_t^n|\theta))$$

$$S_{t+1}=S^M(S_t,x_t,W_{t+1})$$

Searching over policies

- Policy function approximations
- Cost function approximations
- Value function approximations
- Direct lookahead approximations

System model

- Analytical functions
- Optimization problem
- Optimization problem
- Optimization problem


John R. Birge
François Louveaux
Introduction to Stochastic Programming
Second Edition
Michael C. Fu Editor
Handbook of Simulation Optimization
Princeton Series in APPLIED MATH
Robust Optimization
Introduction to Decision Analysis
A Practitioner's Guide to
SECOND EDITION
Approximate Dynamic Programming
Solving the Curses of Dimensionality
Warren B. Powell
VOLUME 2 • 4th EDITION
Dynamic Programming and Optimal Control
APPROXIMATE DYNAMIC PROGRAMMING
Dimitri P. Bertsekas
WILEY
INTRODUCTION TO STOCHASTIC SEARCH AND OPTIMIZATION
Estimation, Simulation, and Control
JAMES C. SPALL
Wiley Series in Probability and Statistics
Optimal Learning
Warren B. Powell
Ilya O. Ryzhov
MULTI-ARMED BANDIT ALLOCATION INDICES
SECOND EDITION
John Gittins, Kevin Glazebrook and Richard Weber
SECOND EDITION
Model Predictive Control
nd C. Bordons
OPTIMAL CONTROL
Frank L. Lewis
Draguna L. Vrabie
Vassilis L. Syrmos
43
Jiongmin Yong
Xun Yu Zhou
Stochastic Controls
Hamiltonian Systems and HJB Equations
Modelling and Applied Probability
Reinforcement Learning
An Introduction
second edition
Richard S. Sutton and Andrew G. Barto
Markov Decision Processes
Discrete Stochastic Dynamic Programming
MARTIN L. PUTERMAN
LECTURES ON STOCHASTIC PROGRAMMING
Modeling and Theory
Alexander Shapiro
Darinka Dentcheva
Andrzej Ruszczyński
MOS-SIAM Series on Optimization
Online Computation and Competitive Analysis
Allan Borodin
Ran El-Yaniv


Michael C. Fu Editor
Handbook of Simulation Optimization
Springer
MULTI-ARMED BANDIT ALLOCATION INDICES
SECOND EDITION
John Gittins, Kevin Glazebrook and Richard Weber
WILEY
OPTIMAL CONTROL
Frank L. Lewis
Draguna L. Vrabie
Vassilis L. Syrmos
VOLUME 2 • 4th EDITION
Dynamic Programming and Optimal Control
APPROXIMATE DYNAMIC PROGRAMMING
Dimitri P. Bertsekas
SECOND EDITION
Approximate Dynamic Programming
Solving the Curses of Dimensionality
Warren B. Powell
43
Jiongmin Yong
Xun Yu Zhou
Stochastic Controls
Hamiltonian Systems and HJB Equations
Modelling and Applied Probability
Reinforcement Learning
An Introduction
second edition
Richard S. Sutton and Andrew G. Barto
WILEY
Markov Decision Processes
Discrete Stochastic Dynamic Programming
MARTIN L. PUTERMAN
LECTURES ON STOCHASTIC PROGRAMMING
Modeling and Theory
SECOND EDITION
Alexander Shapiro
Darinka Dentcheva
Andrzej Ruszczyński
MOS-SIAM Series on Optimization
WILEY
Wiley-Interscience Series in Discrete Mathematics and Optimization
INTRODUCTION TO STOCHASTIC SEARCH AND OPTIMIZATION
Estimation, Simulation, and Control
JAMES C. SPALL

John R. Birge
François Louveaux
Introduction to Stochastic Programming
Second Edition
Springer
Michael C. Fu Editor
SECOND EDITION
Model Predictive Control
E.F. Camacho and C. Bordons
Princeton Series in APPLIED MATHEMATICS
Robust Optimization
Aharon Ben-Tal
Laurent El Ghaoui
Arkadi Nemirovski
Introduction to Decision Analysis
A Practitioner's Guide to Improving Decision Quality
D C. SKINNER
IRD EDITION
LECTURES ON STOCHASTIC PROGRAMMING
Modeling and Theory
Alexander Shapiro
Darinka Dentcheva
Andrzej Ruszczyński
MOS-SIAM Series on Optimization
Online Computation and Competitive Analysis
Allan Borodin
Ran El-Yaniv
kov Decision Processes
Discrete Stochastic Dynamic Programming
MARTIN L. PUTERMAN
Vassilis L. Syrmos
Warren B. Powell
Ilya O. Ryzhov
43
Jiongmin Yong
Xun Yu Zhou
Stochastic Controls
Hamiltonian Systems and HJB Equations
Modelling and Applied Probability
An Introduction
second edition
Richard S. Sutton and Andrew G. Barto
gramming
Control
YNAMIC PROGRAMMING

Choosing a policy class

- It helps to identify five types of policies:
  - » 1) Policy function approximations (PFAs) – Simple rules, analytical functions.
  - » 2) Cost function approximations (CFAs) – Parameterized deterministic optimization models (typically static)
  - » 3) Policies based on value function approximations (VFAs) – Policies that use an approximation of the value of landing in a downstream state
  - » 4) Policies based on direct lookahead approximations (DLAs) – These should be divided into two subclasses:
    - 4a) DLAs using deterministic lookaheads (Det-DLA) – These may be parameterized.
    - 4b) DLAs using stochastic lookaheads (Stoch-DLA)

- So, which are the most useful?

Choosing a policy class

- We can divide the five types of policies into three categories from most to least popular:

  » Category 1 – This category consists of:
    - 1) PFAs – Rules/analytical functions
    - 2) CFAs – Parameterized det. optimization
    - 4a) Det-DLAs – Deterministic lookaheads

  By far the most widely used in practice. The choice among the three tends to be obvious from the structure of the problem.

  » Category 2 – This category consists of:
    - 4b) Stochastic direct lookaheads

  Useful for more complex problems where planning into an uncertain future is required, and risk is important.

  » Category 3 – This category consists of:
    - 3) Policies based on VFAs.

  A very powerful strategy for a very small number of specialized problems.

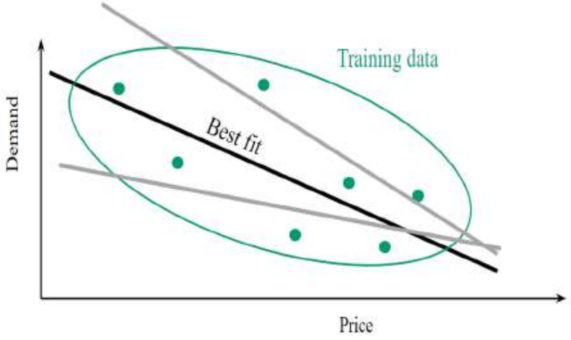
BRIDGING MACHINE LEARNING & SEQUENTIAL DECISIONS

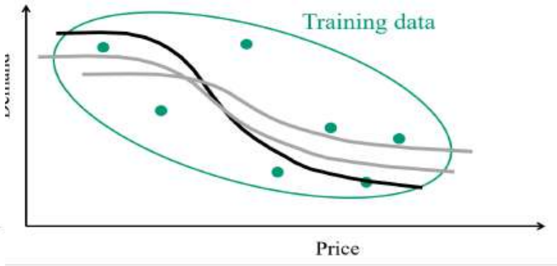
Machine learning

$$\min_{f\in F}\min_{\theta\in\Theta^f}\frac{1}{N}\sum_{n=1}^{N}\left(y^n - f\left(x^n|\theta\right)\right)^2$$

Searching over functions

Parameter tuning





Parameterized policies

$$\min_{f\in F}\min_{\theta\in\Theta^f}\frac{1}{N}\sum_{n=1}^{N}\sum_{t=0}^{T} C\left(S_t^n, X^\pi\left(S_t|\theta\right)\right)$$

$$S_{t+1}=S^M\left(S_t,x_t,W_{t+1}\right)$$

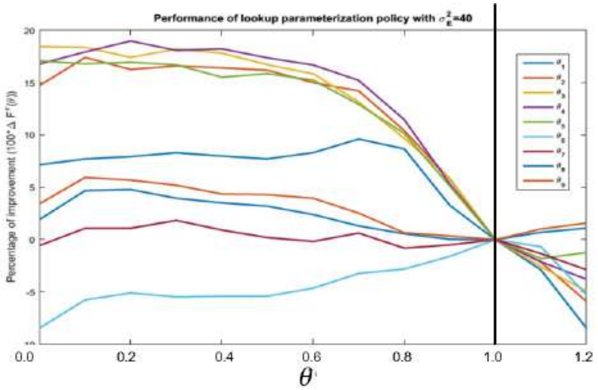
Searching over functions

Parameter tuning

PFAs

CFAs

Determ. DLAs



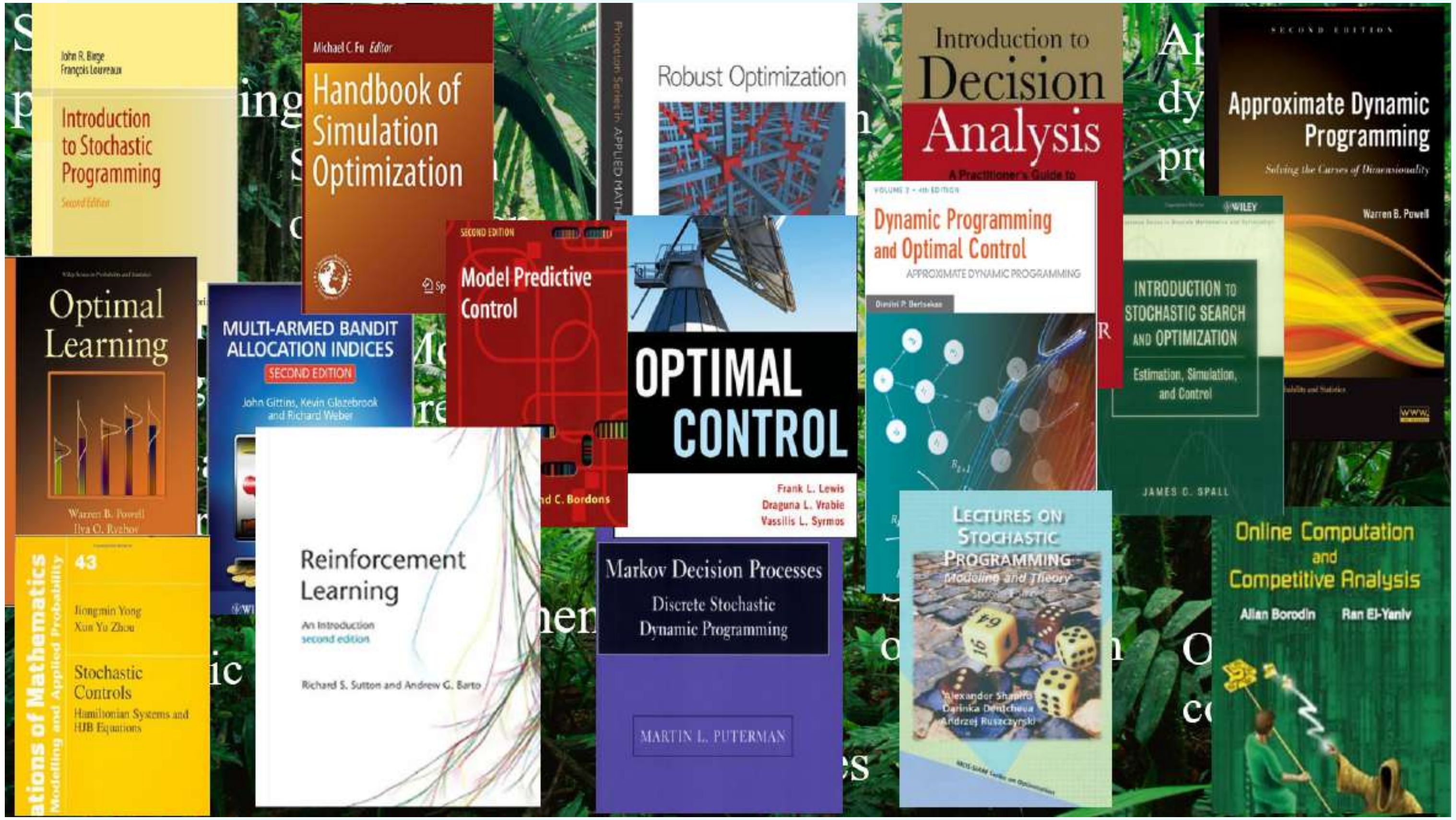
OUTLINE

- ➜The seven levels of artificial intelligence
- ➜The universal modeling framework
- ➜Designing policies
- ➜Mutual fund cash balance optimization
- ➜Choosing the best policy
- ➜A new educational field: sequential decision analytics

Teaching decision analytics

The fields that deal with decisions and uncertainty are completely fragmented.

- Sequential decision analytics is not a recognized field.
- There are 15 distinct communities that deal with decisions under uncertainty
- Each community offers tools that work only for specific problems
- Real applications require skills that span a wide range of problem settings.

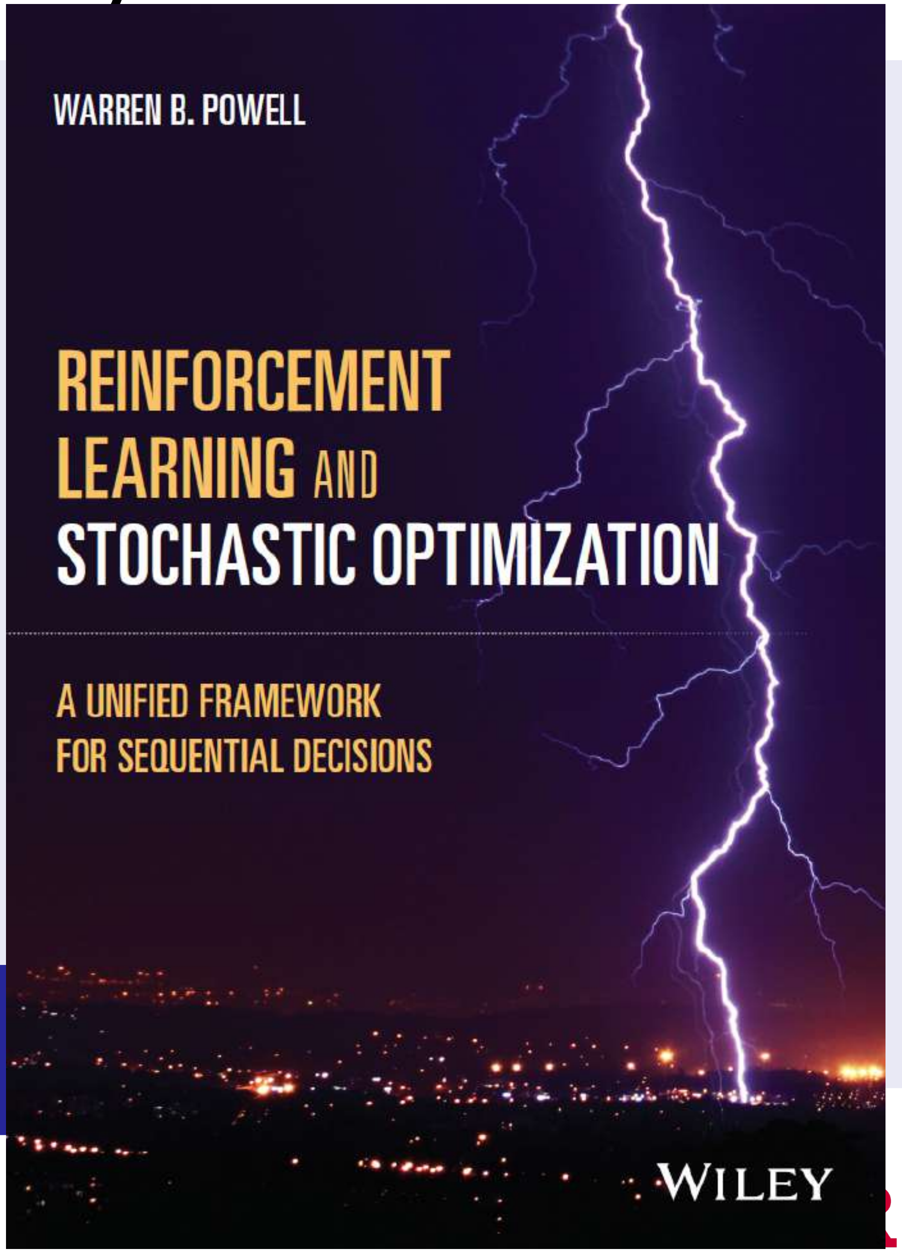


Teaching decision analytics

A graduate-level textbook (2022):

- First book to introduce a universal modeling framework, covering all four classes of policies.
- Describes the tools for modeling and solving *any* sequential decision problem, from simple learning problems to truckload fleets to complex supply chains.
- Aimed at a technical audience interested in writing software to develop models such as those described in this presentation.
- Provides the foundation for a new field we are calling *sequential decision analytics*.

http://tinyurl.com/RLandSO/

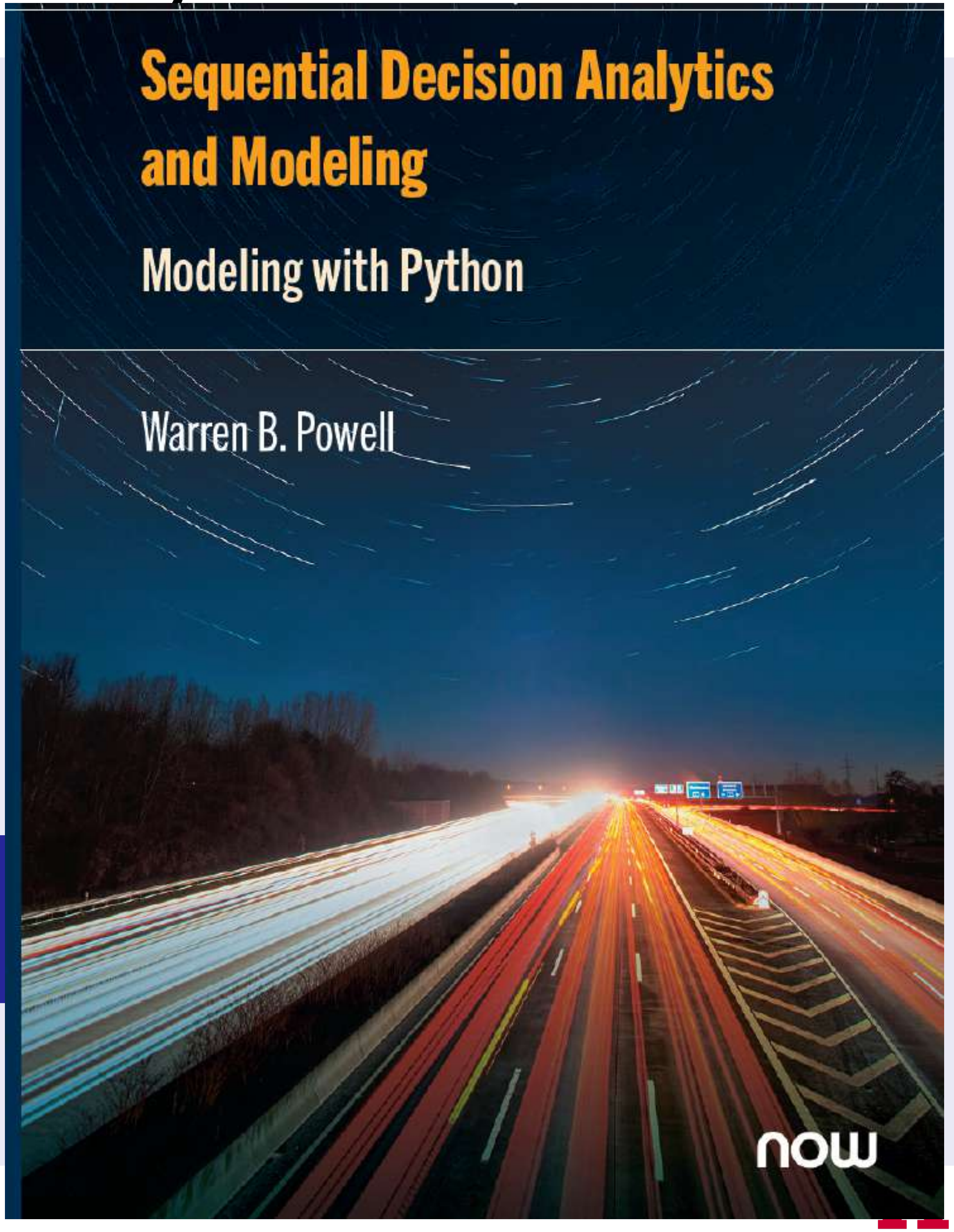


Teaching decision analytics

An introductory book (2022):

- Uses a teach-by-example style
- Illustrates how to model sequential decision problems using a rich set of examples
- Illustrates all four classes of policies
- Highlights uncertainty modeling
- Downloadable from:

http://tinyurl.com/sdamodeling



Teaching decision analytics

Chapter 1 - Introduction

Chapter 2 – An asset selling problem

Chapter 3 – Adaptive market planning

Chapter 4 – Learning the best diabetes medication

Chapter 5 – Stochastic shortest path problems - Static

Chapter 6 – Stochastic shortest path problems - Dynamic

Chapter 7 – Applications, revisited

Chapter 8 – Energy storage I

Chapter 9 – Energy storage II

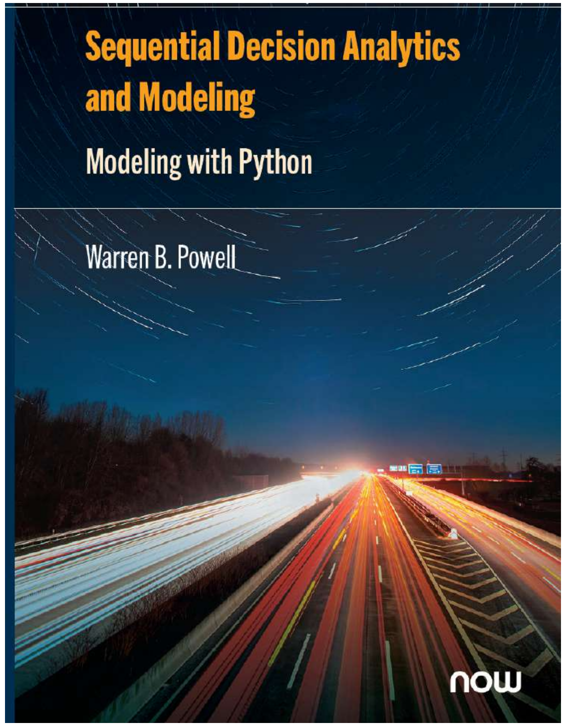
Chapter 10 – Supply chain management I – The two-agent newsvendor problem

Chapter 11 – Supply chain management II – The beer game

Chapter 12 – Ad-click optimization

Chapter 13 – Blood management problem

Chapter 14 – Optimizing clinical trials



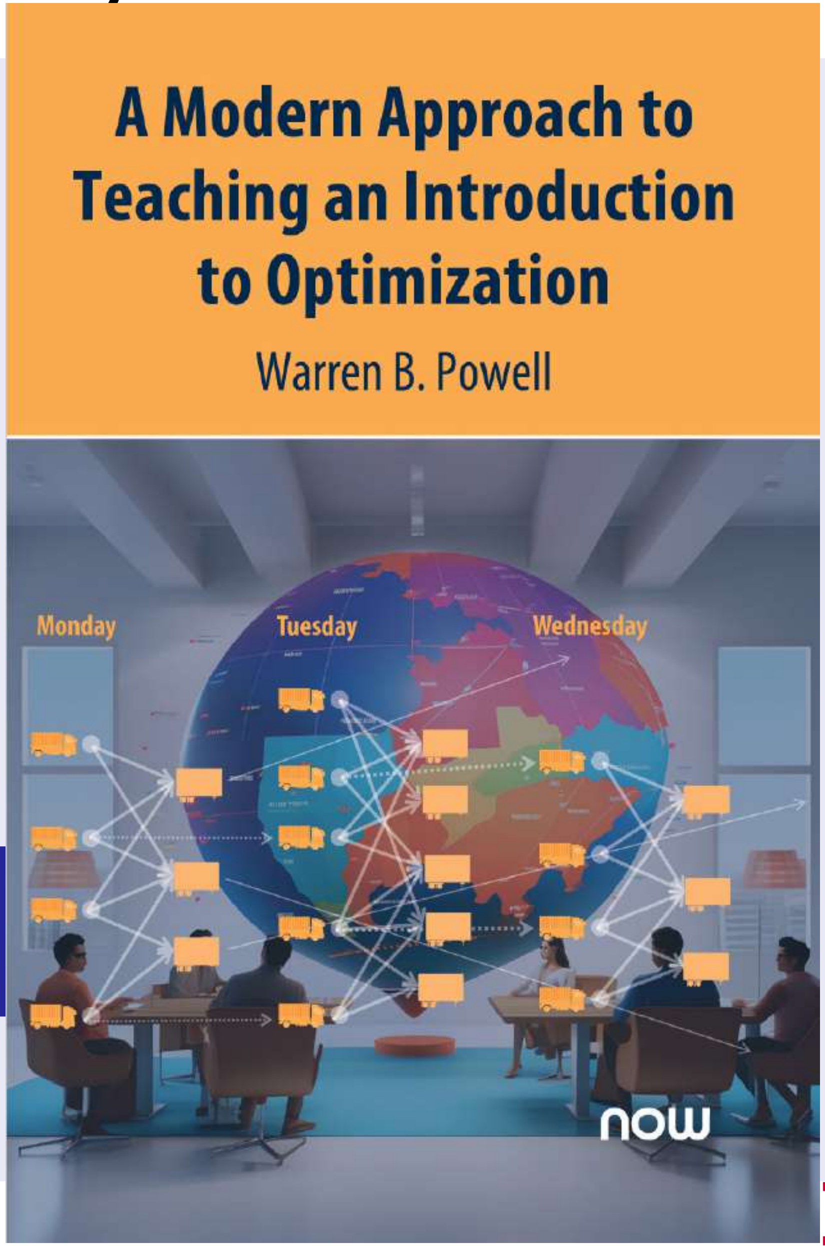
https://tinyurl.com/sdamodeling/

Teaching decision analytics

A book for instructors (2024):

- Fundamentally changes how to teach introduction to optimization
- Transitions from machine learning, through simpler sequential decision problems, to static and sequential linear, integer and nonlinear problems
- Illustrates both static and sequential decision problems
- Available as free download

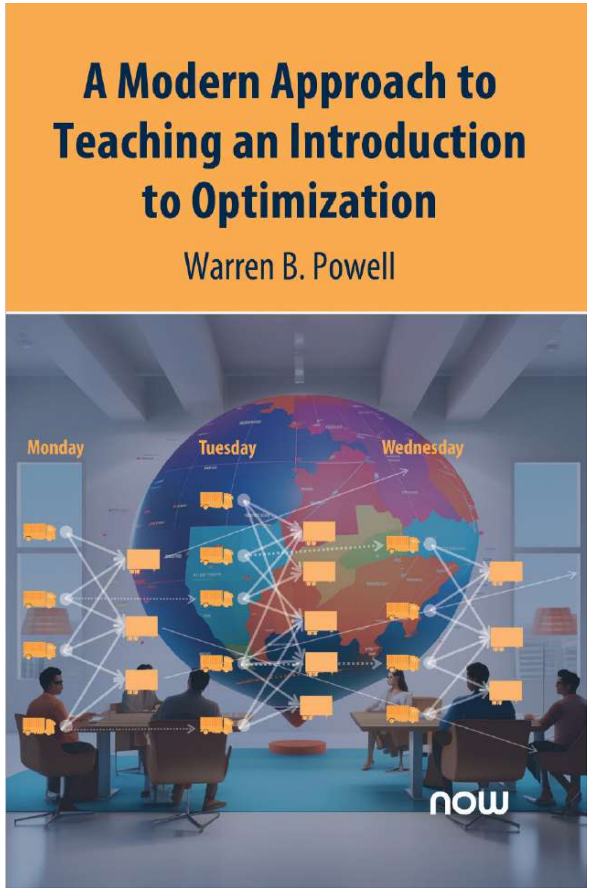
http://tinyurl.com/TeachingOpt



Teaching introduction to optimization

The course is divided into 11 topics:

Topic 1 – Machine learning
Topic 2 – Asset selling, inventory problems
Topic 3 – Newsvendor problem
Topic 4 – Optimal learning
Topic 5 – Shortest path problems (static, sequential)
Topic 6 – Overview of sequential decision problems
Topic 7 – Linear programming
Topic 8 – Dynamic inventory problem
Topic 9 – Integer programming I (static facility location)
Topic 10 – Integer programming II (dynamic facility location)
Topic 11 – Nonlinear programming (static,



https://tinyurl.com/TeachingOpt/

Thank you!

An information resource page for sequential decisions:

http://tinyurl.com/SDAlinks

This page includes links to:

» Books (introductory and advanced)
» Videos
» Educational webpages
» Courses and teaching methods
» A series of educational LinkedIn posts