A Unified Framework for Sequential Decisions under Uncertainty

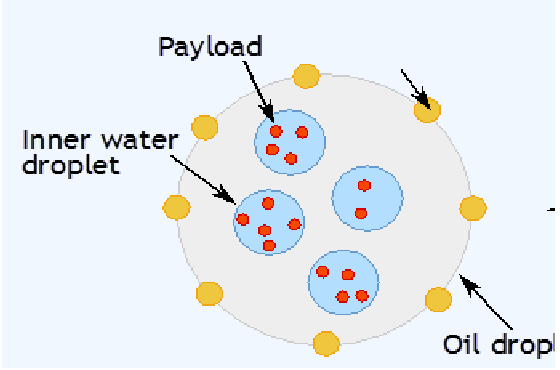
Warren B Powell
Chief Analytics Officer, Optimal Dynamics
Professor Emeritus, Princeton University

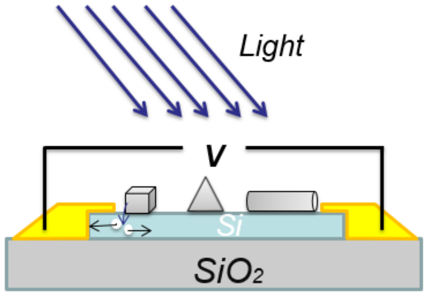
Institute of Data Sciences
Texas A&M University

January 24, 2022

Laboratory sciences

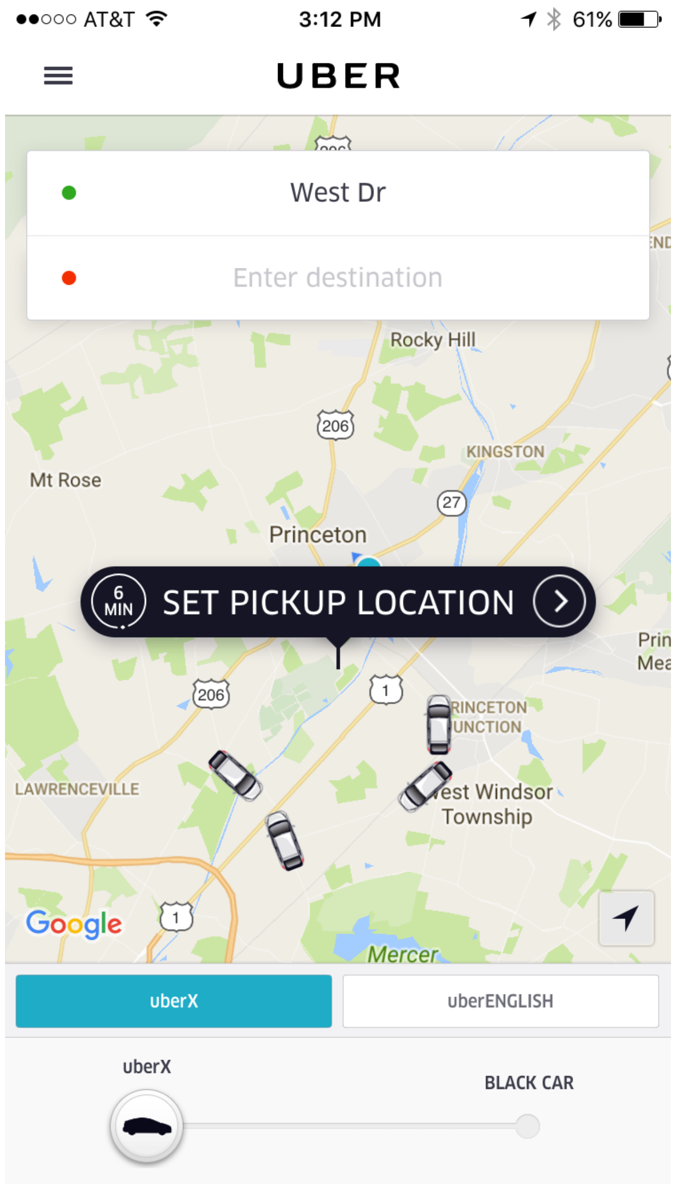
- Materials science
  - » Optimizing payloads: reactive species, biomolecules, fluorescent markers, …
  - » Controllers for robotic scientist for materials science experiments
  - » Optimizing nanoparticles to maximize photoconductivity





Managing ride-hailing fleets

- Uber
  - » Provides real-time, on-demand transportation.
  - » Drivers are encouraged to enter or leave the system using pricing signals and informational guidance.
- Decisions:
  - » How to price to get the right balance of drivers relative to customers.
  - » Assigning and routing drivers to manage Uber-created congestion.
  - » Real-time management of drivers.
  - » Pricing (trips, new services, …)
  - » Policies (rules for managing drivers, customers, …)

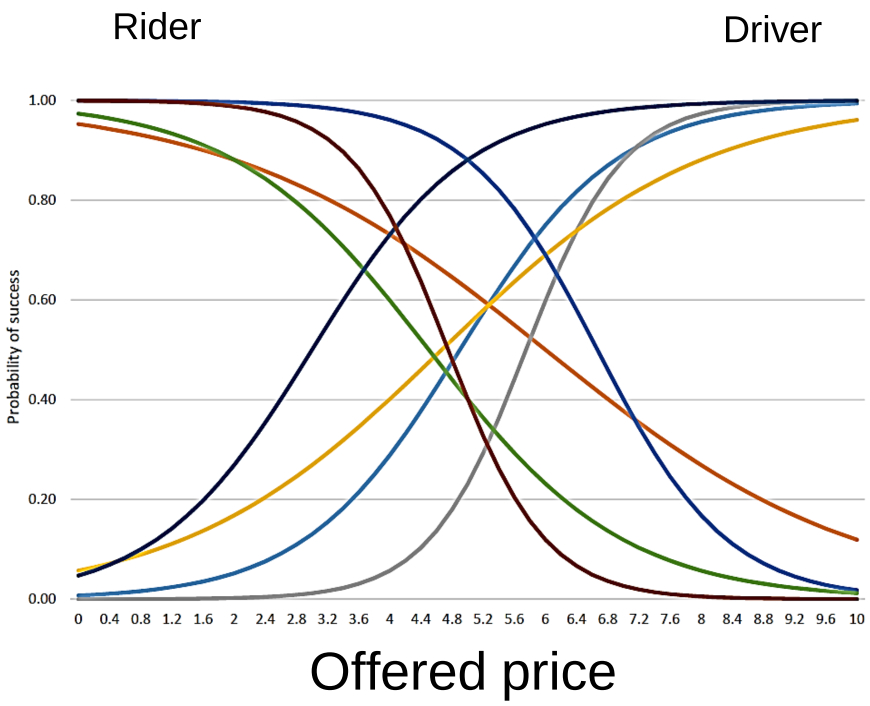


Managing ride-hailing fleets

- Now we have a logistic curve for each origin-destination pair (i,j)

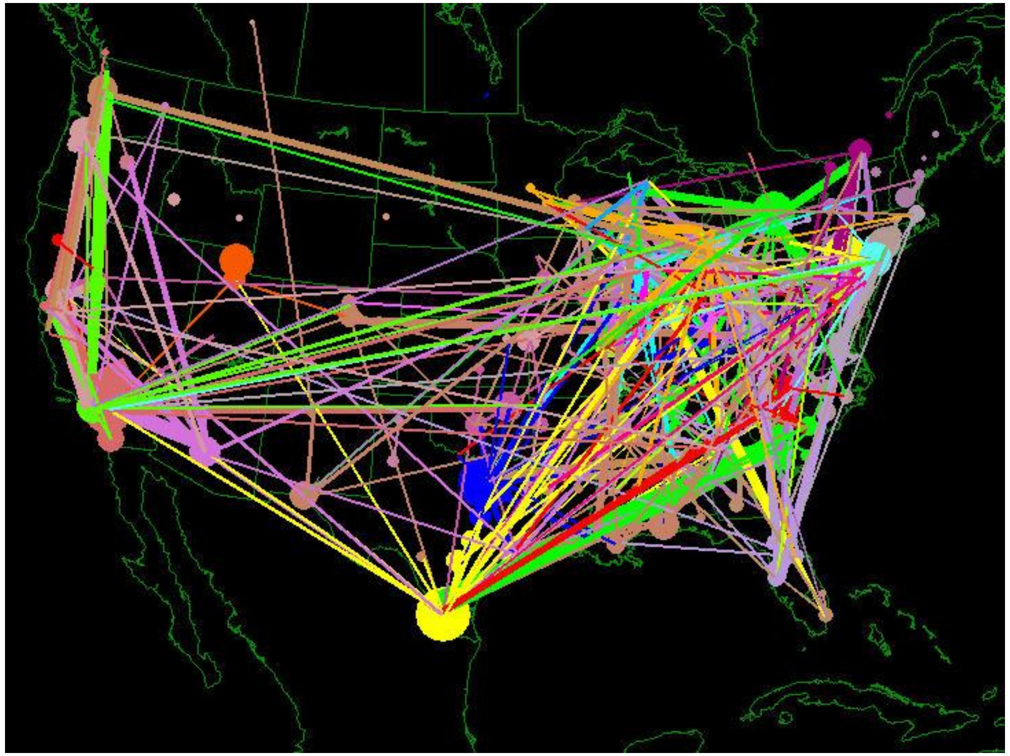
$$P^Y(p,a\,|\,\theta)=\frac{e^{\theta_{ij}^0+\theta_{ij}p+\theta_{ij}^a a}}{1+e^{\theta_{ij}^0+\theta_{ij}p+\theta_{ij}^a a}}$$

- Number of offers for each (i,j) pair is relatively small.
- Need to generalize the learning across hundreds to thousands of markets.



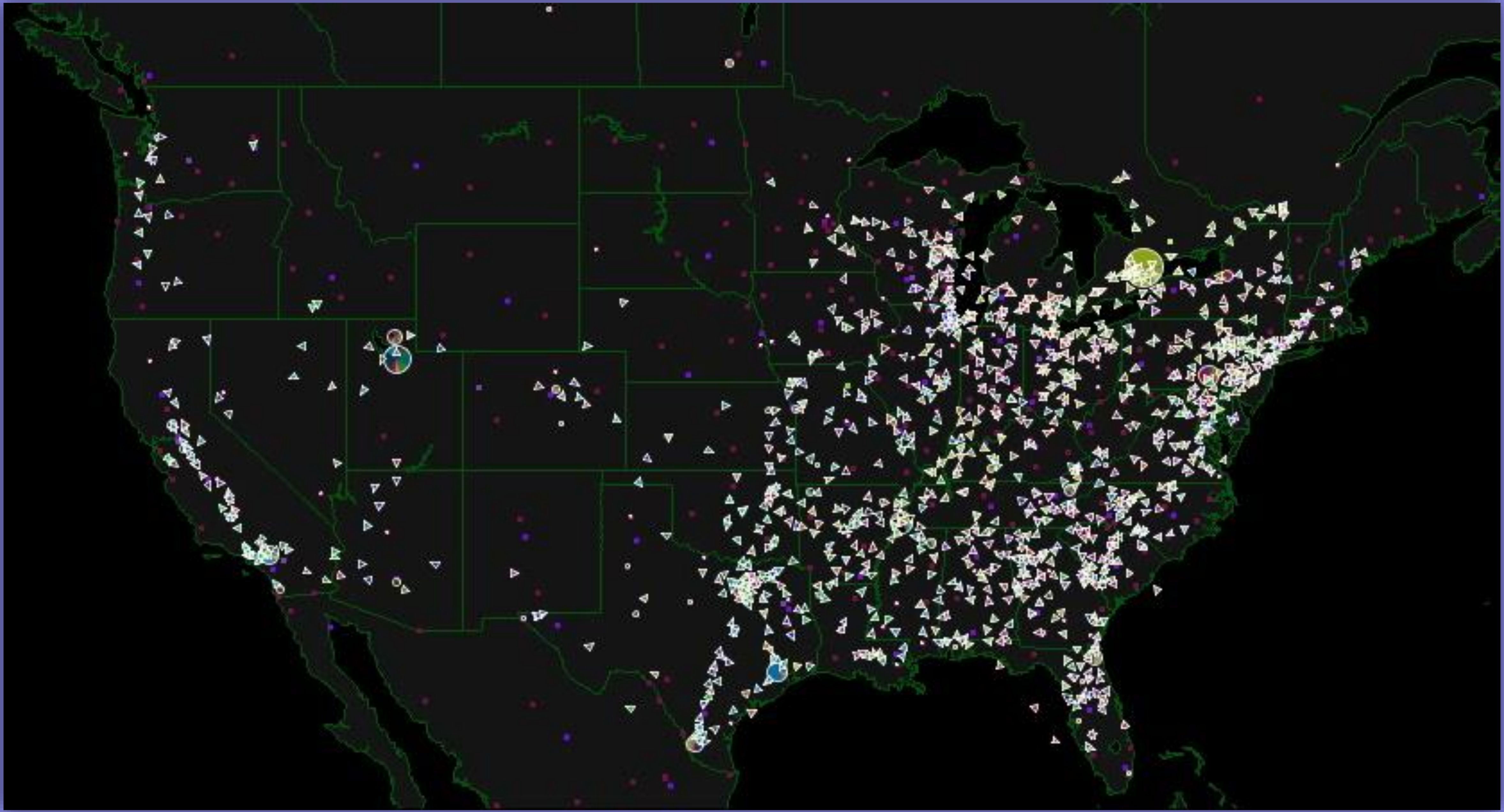
Optimal Dynamics

Fleet management

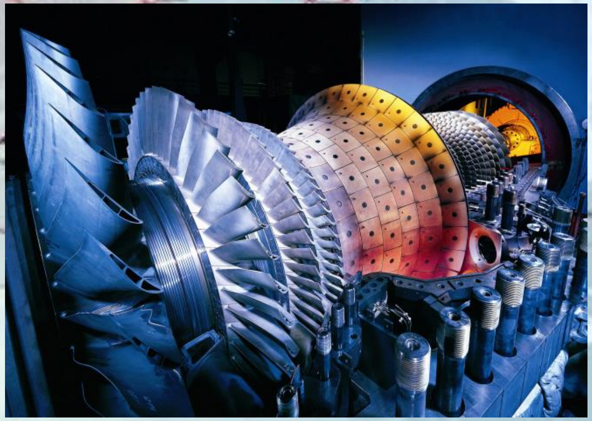


- Fleet management problem
  - » Optimize the assignment of drivers to loads over time.
  - » Tremendous uncertainty in loads being called in

Optimal Dynamics

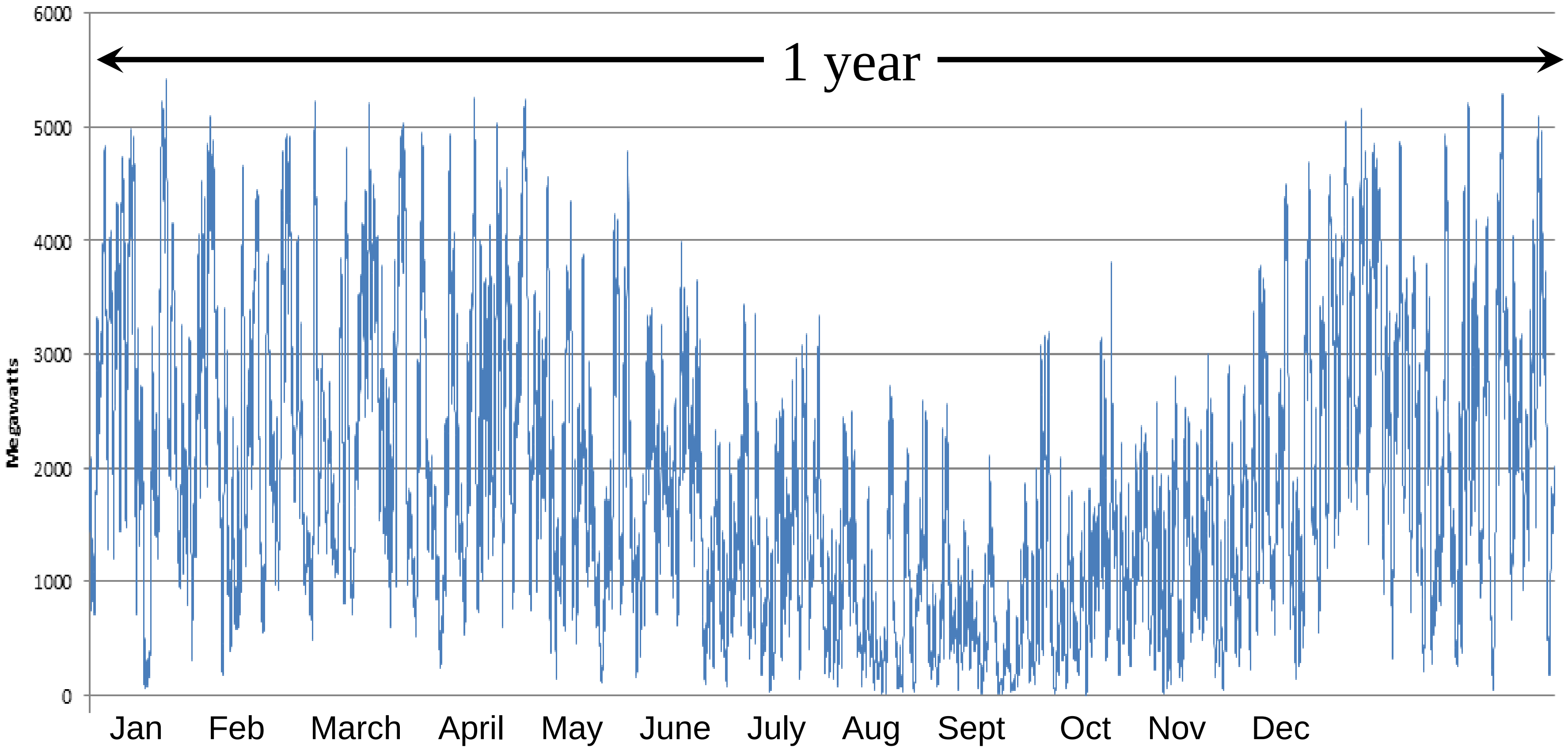


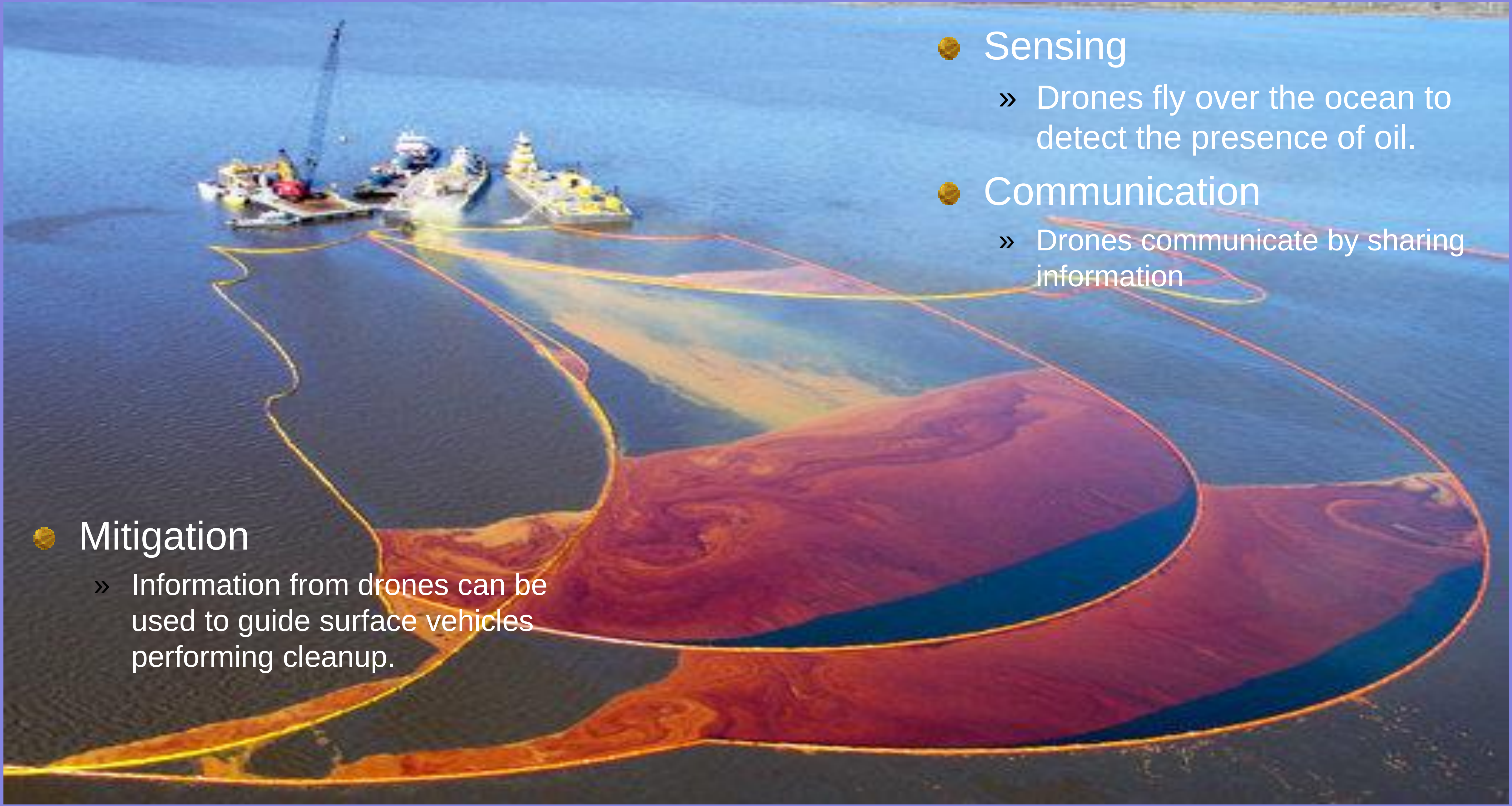
An energy generation portfolio



Energy from wind

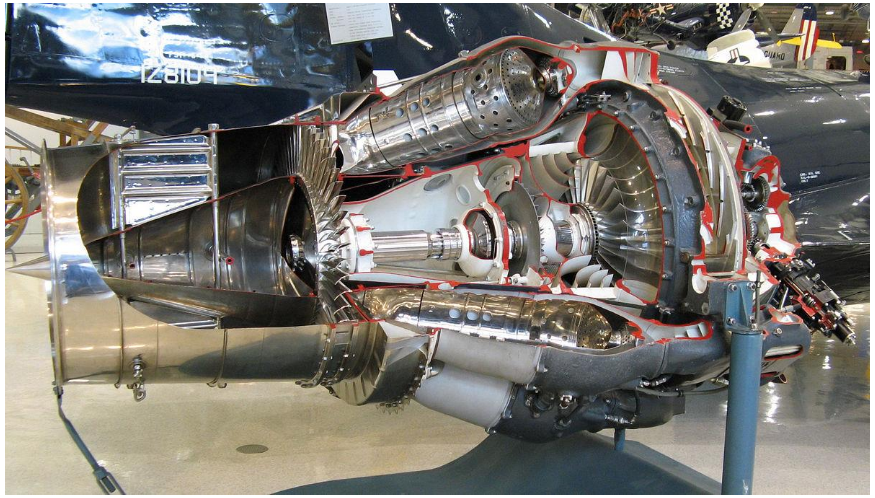
- Wind power from all PJM wind farms





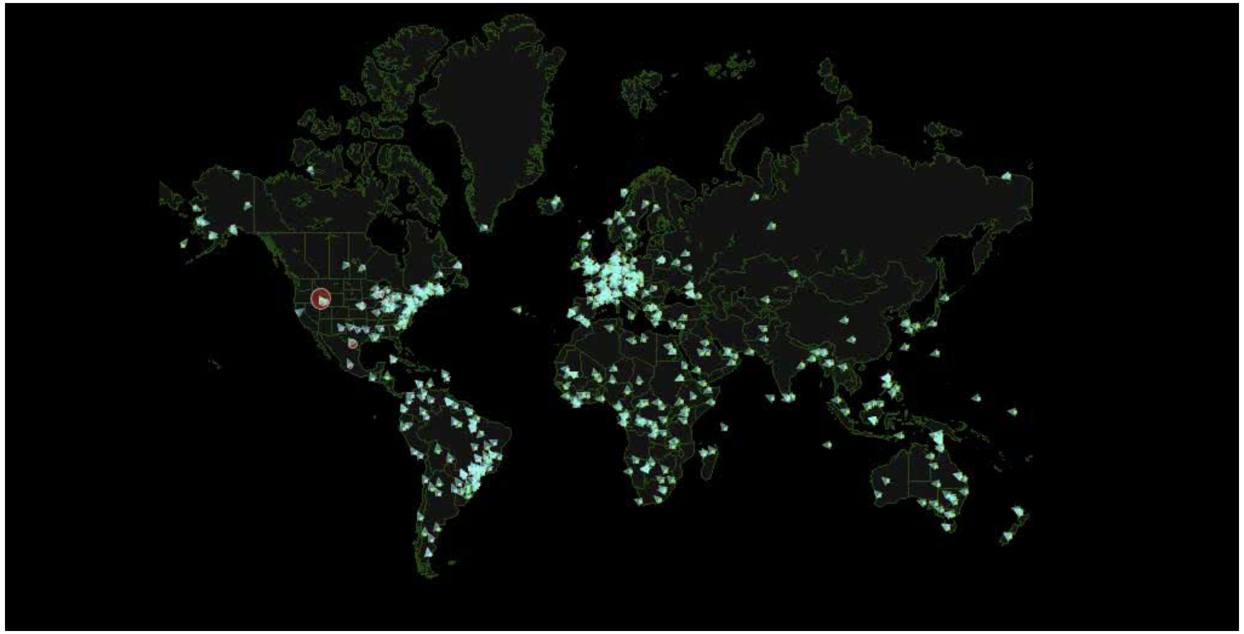
- Sensing
  - Drones fly over the ocean to detect the presence of oil.
- Communication
  - Drones communicate by sharing information
- Mitigation
  - Information from drones can be used to guide surface vehicles performing cleanup.

Multiagent supply chain management



- Pratt & Whitney jet engines
  - » Over 1,000 parts
  - » Median lead time for a part is 120 days. Some lead times are over 300 days.
  - » Parts often require reworking.

- Managing the supply chain
  - » Challenge is determining when to order parts given the long lead times, and production uncertainties.
  - » Suppliers work for multiple customers.




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

CHALLENGES

We are looking for opportunities for making better decisions where we have to deal with uncertainty

- Freight transportation
- Energy systems
- Personal transportation
- Health systems
- Finance
- Public health
- Business processes
- Logistics
- Manufacturing
- Engineering
- Laboratory sciences
- Supply chains
- Storm evacuations
- E-commerce

GOALS & OBJECTIVES

- Reduce costs
- Increase profits
- Improve reliability
- Minimize waste
- Maximize strength
- Improve health
- Reduce risk
- Increase yield
- Reduce carbon production
- Minimize lives lost

GOALS & OBJECTIVES

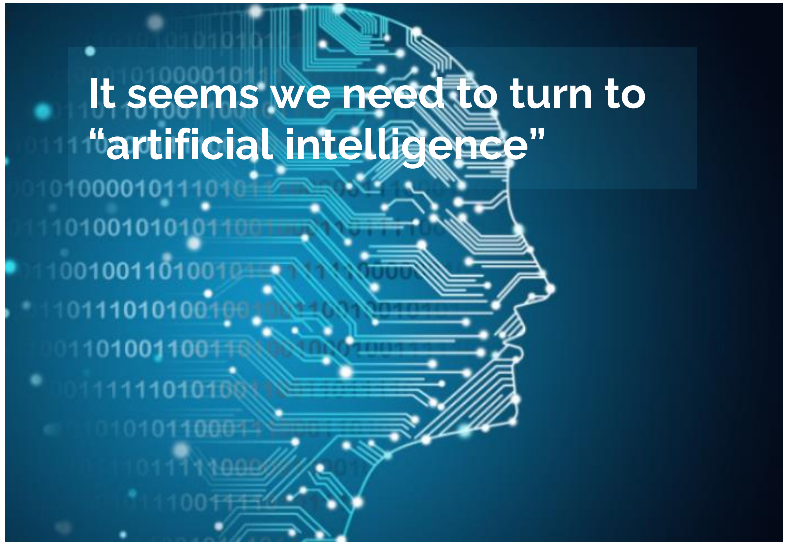
Improve performance by making better decisions.

1ST STEP

What decisions are you making?

2ND STEP

How do we make effective decisions?



Optimal Dynamics

Intelligent decisions

Artificial Intelligence

1. Making computers **behave like** humans
2. Making computers **smarter than** humans

1960s-1970s

Rule-based AI

- If eating meat, then drink red wine
- Given patient attributes, use this treatment.

1990s

Optimization

- Large scale linear & integer programming

2005

Machine Learning

- The new "AI"
- Neural networks

2015

Reinforcement Learning

- Making decisions
- Chess, Go, robots

2020+

???

Optimal Dynamics

WHAT'S NEXT IN AI?

Sequential decision problems, where we need to make decisions over time, as new information arrives.

I propose to unify 15 distinct fields that deal with dynamic decision making into a new field that I call *Sequential Decision Analytics.*

Sequential decision analytics includes all of reinforcement learning, but is broader, with a greater emphasis on uncertainty.

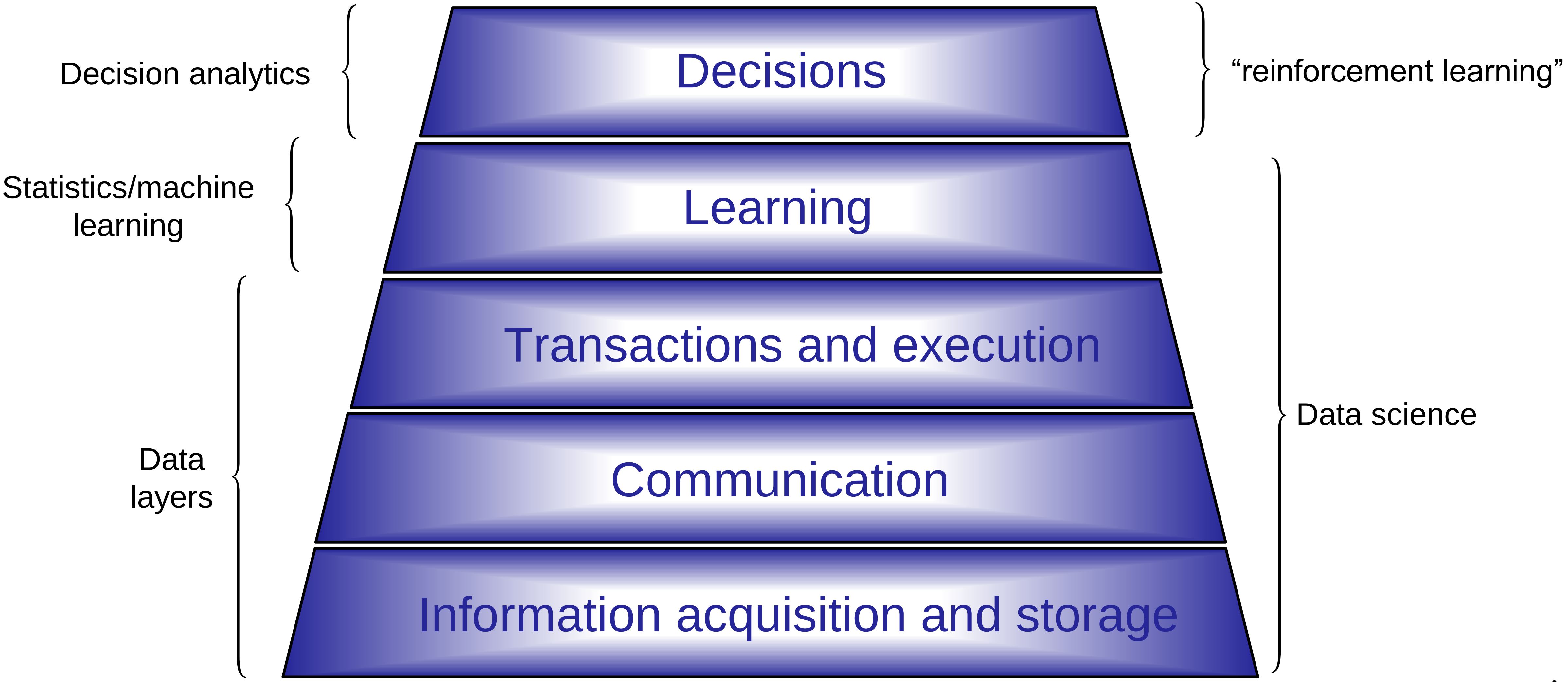
OUTLINE

- → The five layers of intelligence
- → Modeling sequential decision problems
- → Designing policies
    - → Policy function approximations
    - → Cost function approximations
    - → Value function approximations
    - → Direct lookahead approximations
- → A new educational field: sequential decision analytics

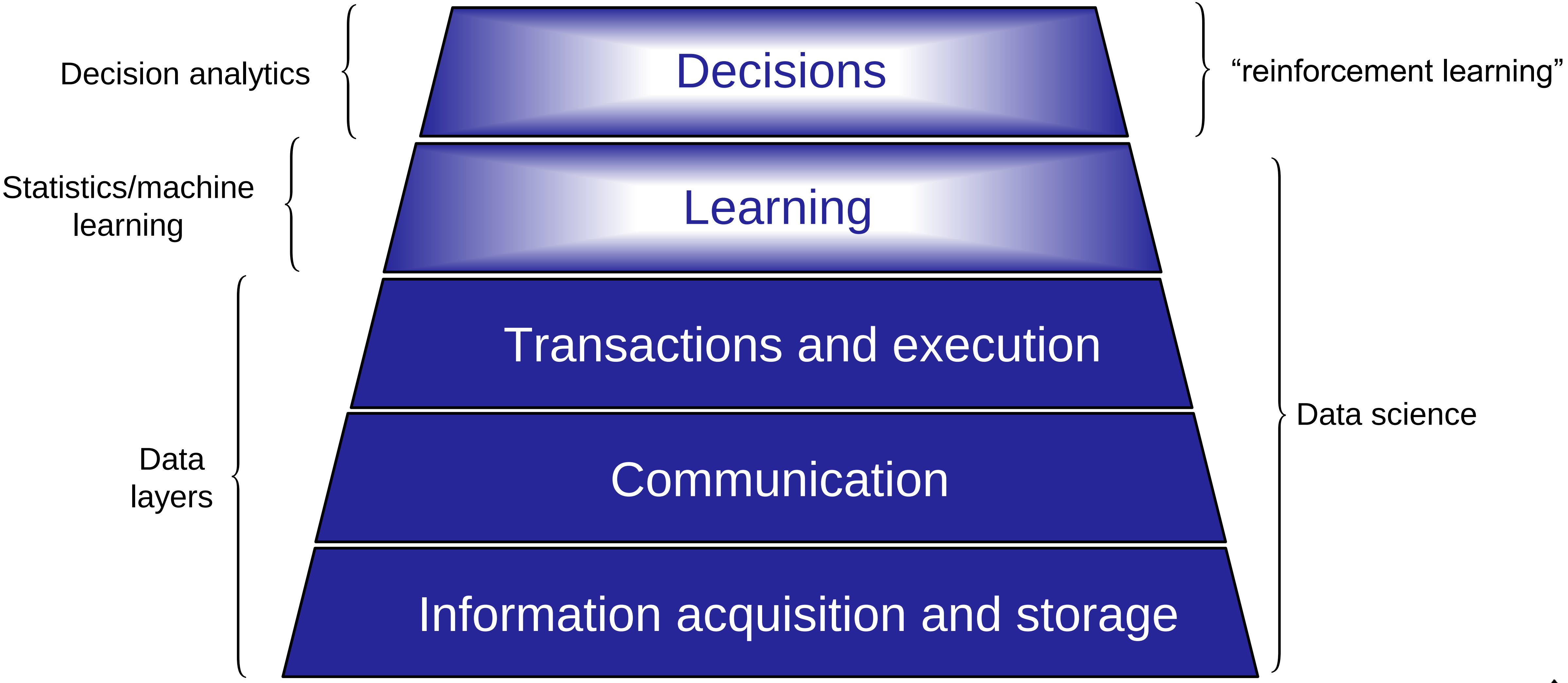
OUTLINE

- → The five layers of intelligence
- → Modeling sequential decision problems
- → Designing policies
  - → Policy function approximations
  - → Cost function approximations
  - → Value function approximations
  - → Direct lookahead approximations
- → A new educational field: sequential decision analytics

THE 5 LAYERS OF INTELLIGENCE

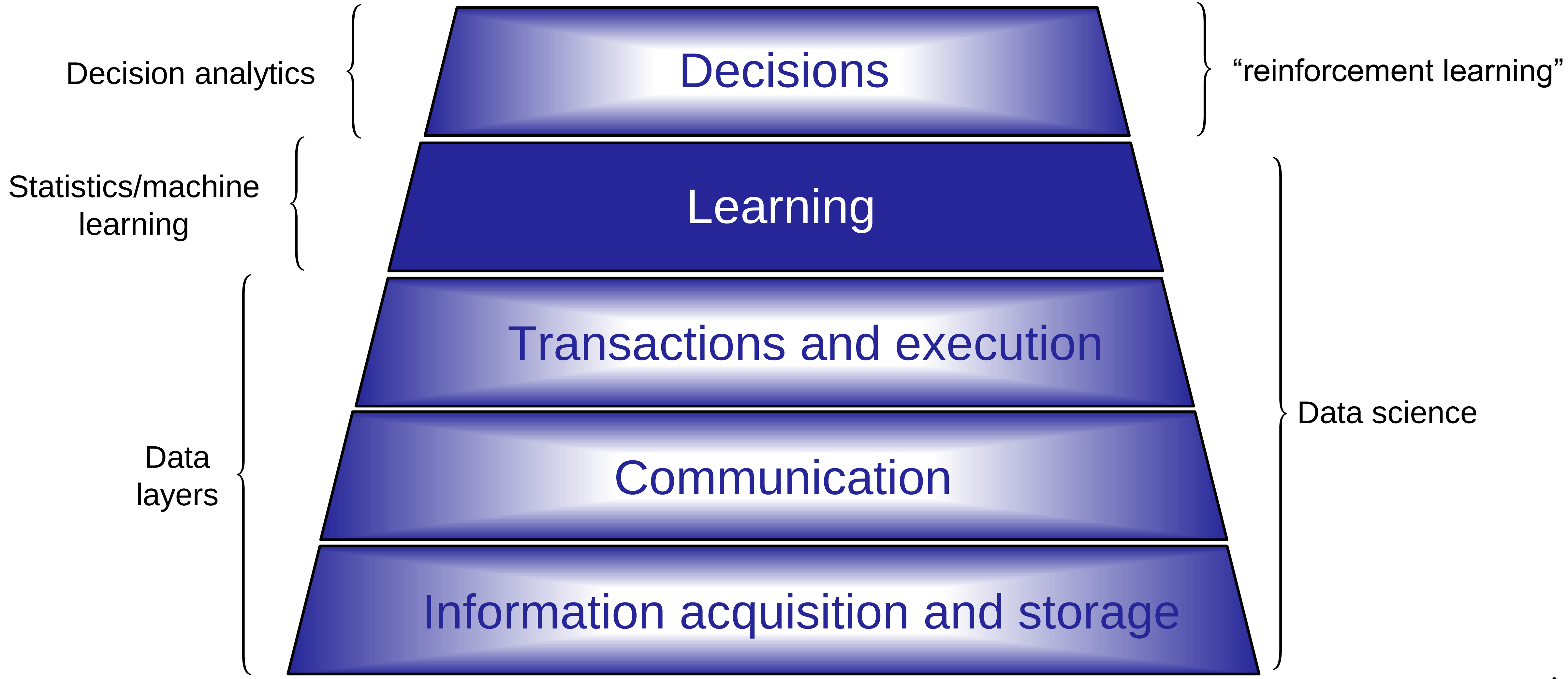


THE 5 LAYERS OF INTELLIGENCE



Optimal Dynamics

THE 5 LAYERS OF INTELLIGENCE



Optimal Dynamics

MACHINE LEARNING

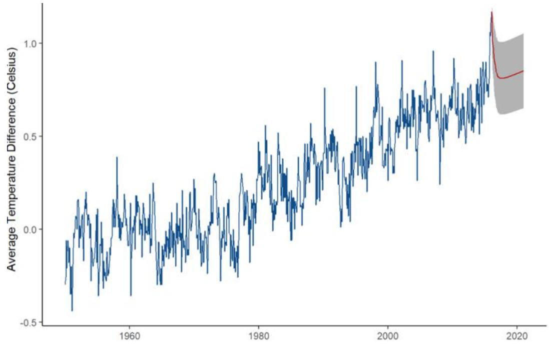
Types of Learning

Pattern Matching	Classification	Inference	Prediction

- What is the voice saying?
- What is in the picture?
- What is the email asking for?

- What product should I recommend for this customer?
- What treatment should I recommend for this patient?

- How will an increase in price affect market demand?
- What is the condition of a piece of equipment?

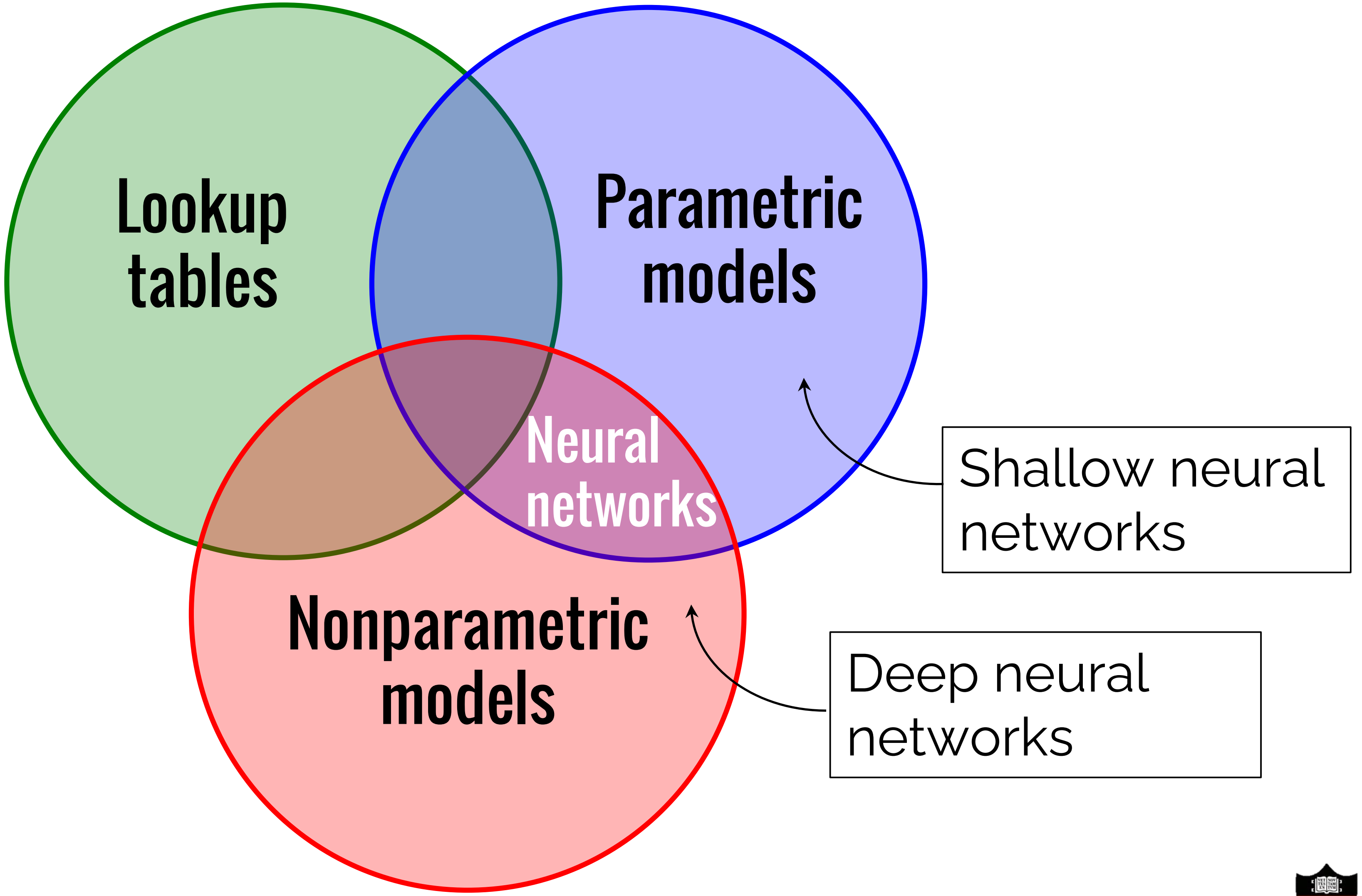


- What will the market demand be in three days?
- How many loads will the shipper need to move in a week?

Optimal Dynamics

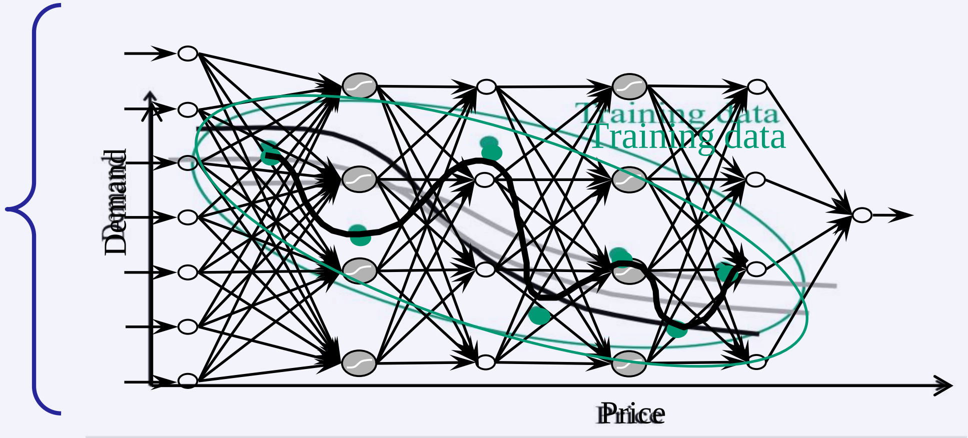
MACHINE LEARNING

Every single machine learning method falls in one of these three circles.



Machine learning as an optimization problem

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

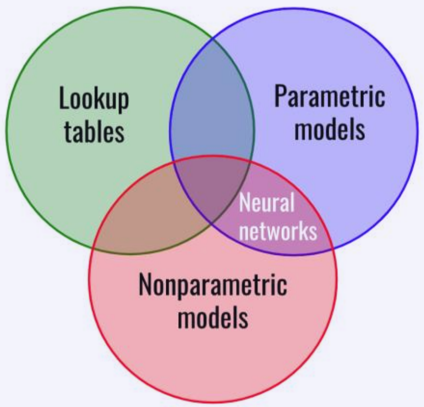


The first step is choosing a mathematical model that will do the best job of fitting the data (but be careful of overfitting with neural networks).

BRIDGING MACHINE LEARNING & SEQUENTIAL DECISIONS

Machine learning as an optimization problem

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



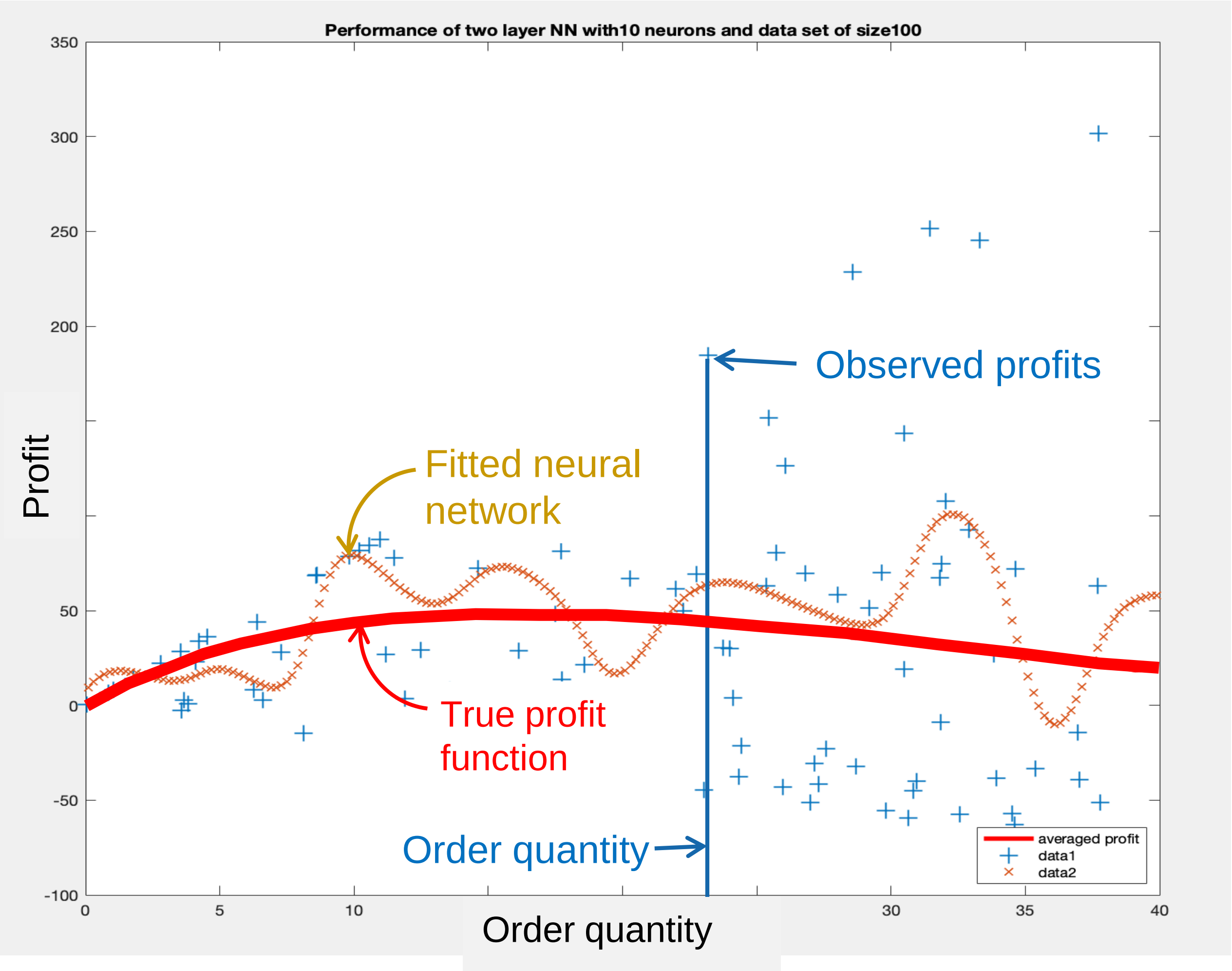
“Big dataset”

Searching over statistical models

These consist of functions $f \in \mathcal{F}$
and tunable parameters $\theta \in \Theta^f$

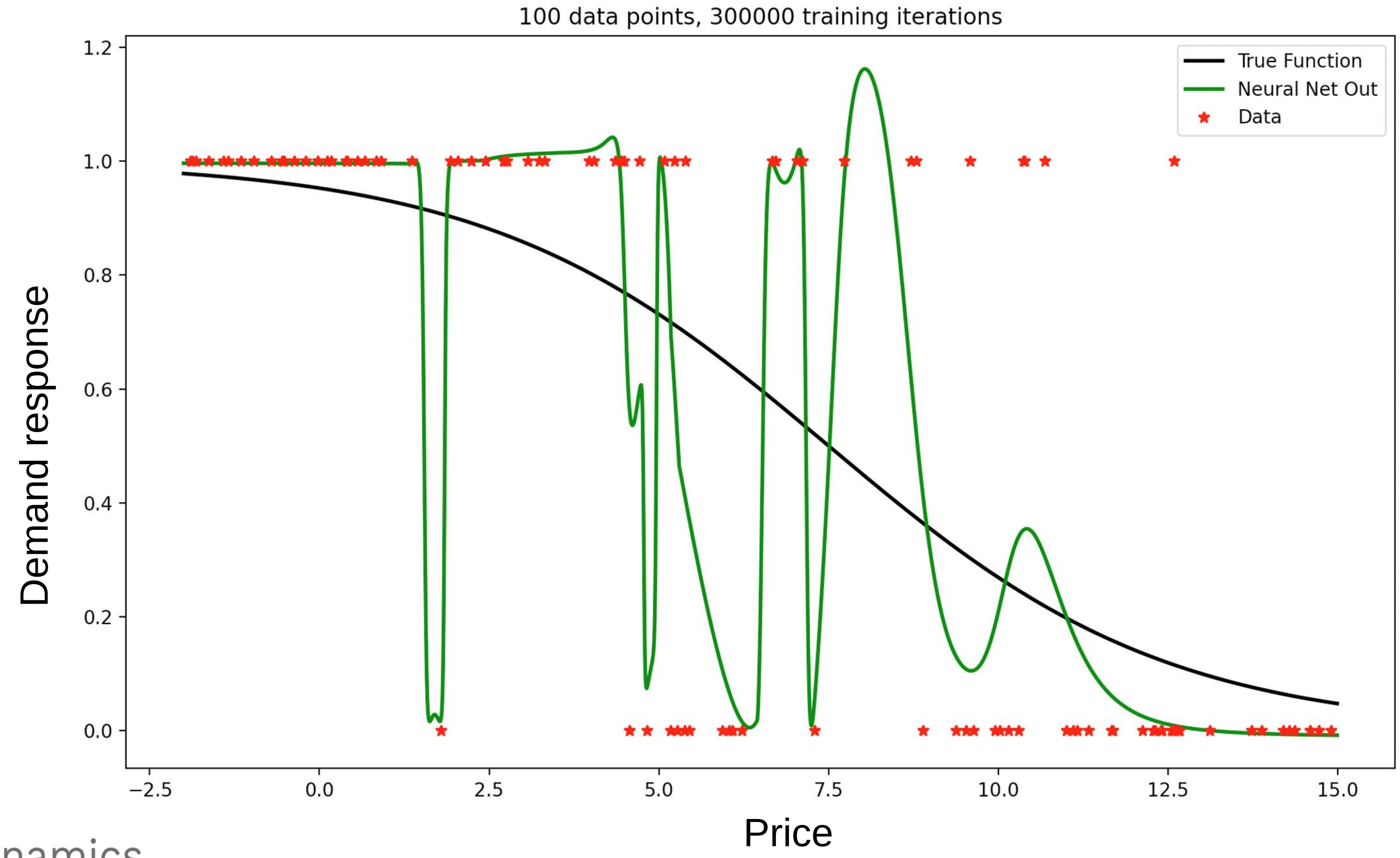
MACHINE LEARNING

- Neural networks struggle with:
  - Noise – Their high flexibility tends to fit the noise.
  - Structure – It is difficult to communicate structure:
    - Monotonicity – Higher price means lower demand
    - Concavity – As with the newsvendor problem to the right.



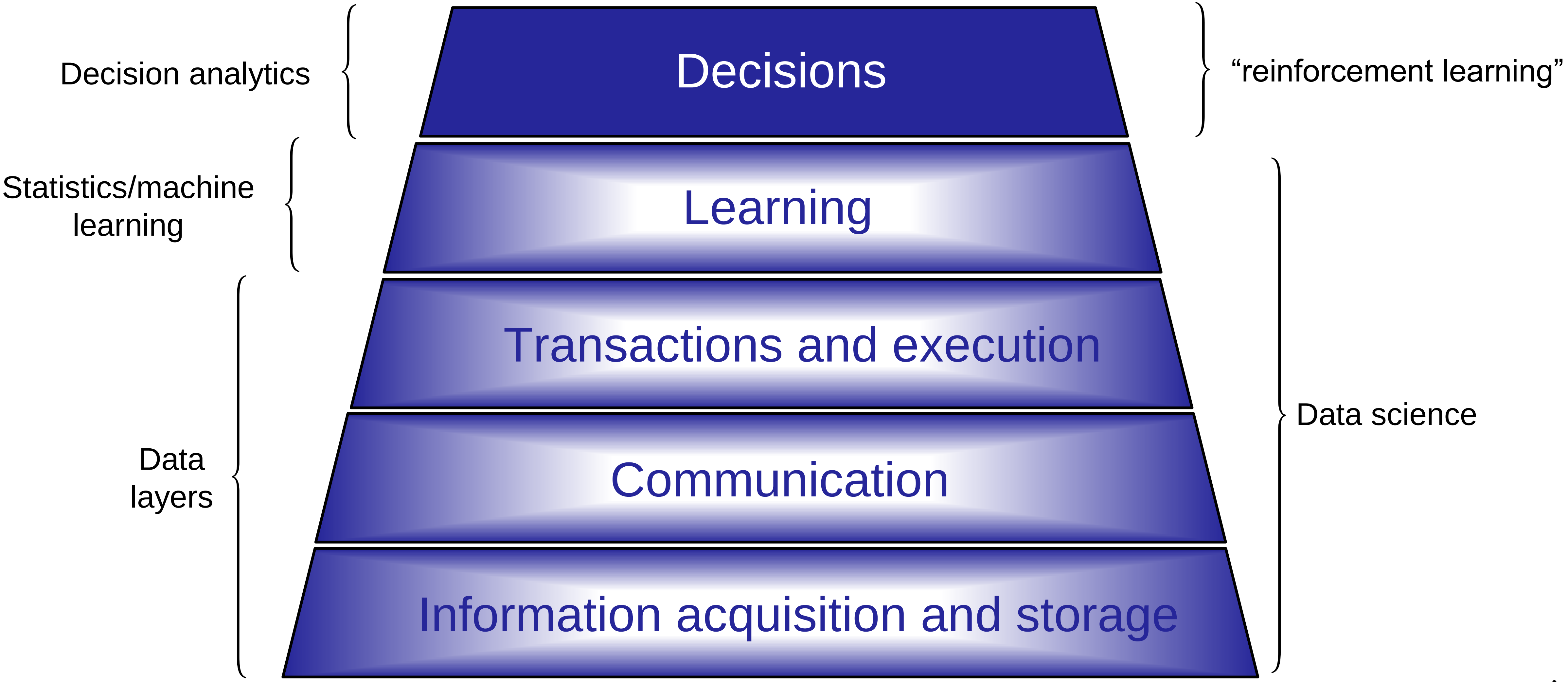
MACHINE LEARNING

- Neural network for demand response:



Optimal Dynamics

THE 5 LAYERS OF INTELLIGENCE

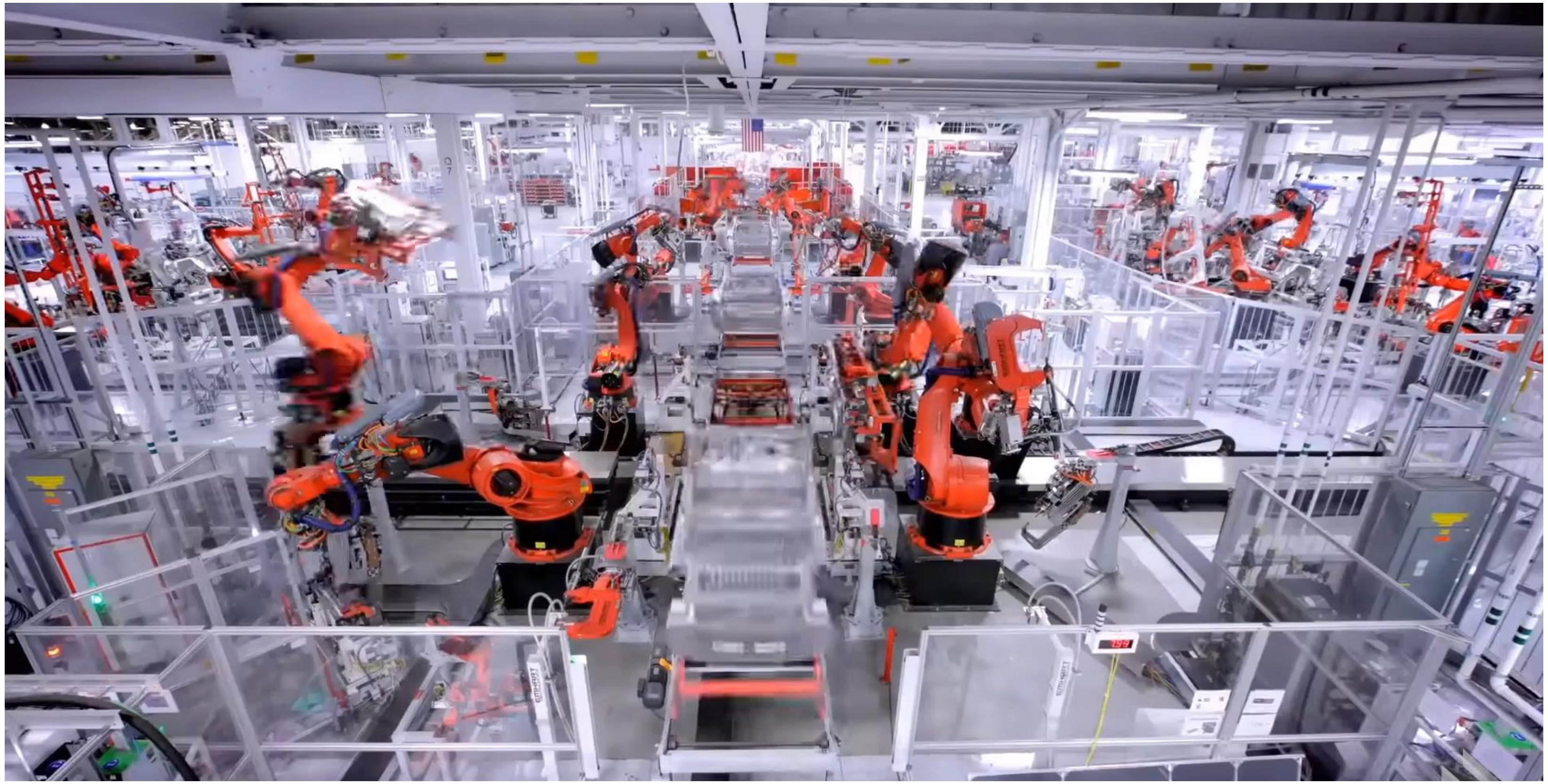


Optimal Dynamics

INFORMATION & DECISION PROCESSES

Information and decision processes

- There are parallels between the process of making “decisions” and a manufacturing line making “products”



We have to approach information processing and decisions like a manufacturing process.


NASDAQ
S&P 500 Index
COVID 19
Coronavirus Vaccine
AIR PRODUCTS
tell me more
www.airproducts.com
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

How many dedicated drivers should we have

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

What is the value o[f a] financial option?

Which vendor should supply each part?

When should inventory be ordered?

Where should drivers be domiciled?

Which physician should handle a procedure?

How many nurses should we have to serve local hospitals

How much battery sto[rage] needed to handle the variability of wind?

What price should be charged

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

Which nurse should visit this doctor's office today?

When should gas turbines be scheduled to handle drops in wind?

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, and which robot?

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

How many aircraft should I order for delivery in five years?

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?

Optimal Dynamics

DECISIONS

Types of decisions.

Physical Decisions

» Managing inventories
» Assigning drivers and moving trucks
» Scheduling nurses and energy generators

Financial Decisions

» Pricing decisions
» Insurance decisions
» Managing investments
» Hedging contracts

Informational Decisions

» Sending/receiving information
» Marketing and advertising
» Running experiments (lab or field)
» Testing drugs

Optimal Dynamics

THE TIME FRAMES FOR DECISIONS

Strategic planning and design – We simulate operational decisions so we understand how a system would respond to decisions far in the future:

- » How many gas turbines should a power grid have?
- » How should we design a building to withstand earthquakes?
- » What should the capacity of a levee or reservoir be?

Tactical planning decisions – We simulate operational decisions to help make decisions that impact the system in the near future,

- » How much energy generation should the grid plan for tomorrow?
- » How many gallons of water should be ordered in anticipation of a hurricane?
- » How to allocate traffic management personnel to handle storm evacuations?

Real-time decisions – These are decisions that impact the system now:

- » Making real-time ramping decisions for energy generators.
- » Notifying houses within a zone to begin evacuations before a storm.
- » Operating pumps to mitigate flooding during a storm.

THE TIME FRAMES FOR DECISIONS

Strategic planning and design – We simulate operational decisions so we understand how a system would respond to decisions far in the future:

We need to simulate decisions in the future that **do *not*** depend on the state of the system now.

Tactical planning decisions – We simulate operational decisions to help make decisions that impact the system in the near future,

We need to simulate decisions in the future that ***do*** depend on the state of the system now.

Real-time decisions – These are decisions that impact the system now:

We need to simulate the effect of a decision now
(which depends on the state of the system) on the future.

DETERMINISTIC OPTIMIZATION

Airline scheduling

Airlines	Optimization Model	Airline Schedule

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

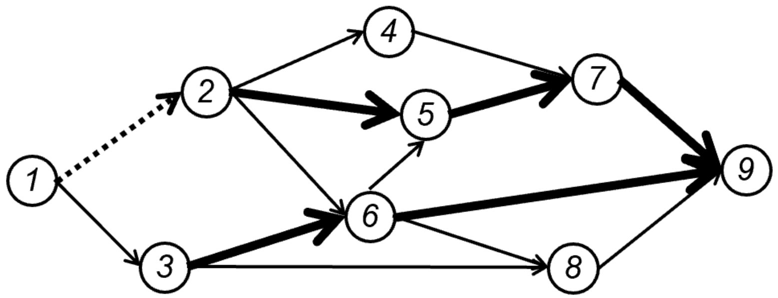
Airlines around the world use tools that depend on this mathematical model to perform strategic and operational planning.

Optimal Dynamics

DETERMINISTIC OPTIMIZATION

Low dimensional decisions

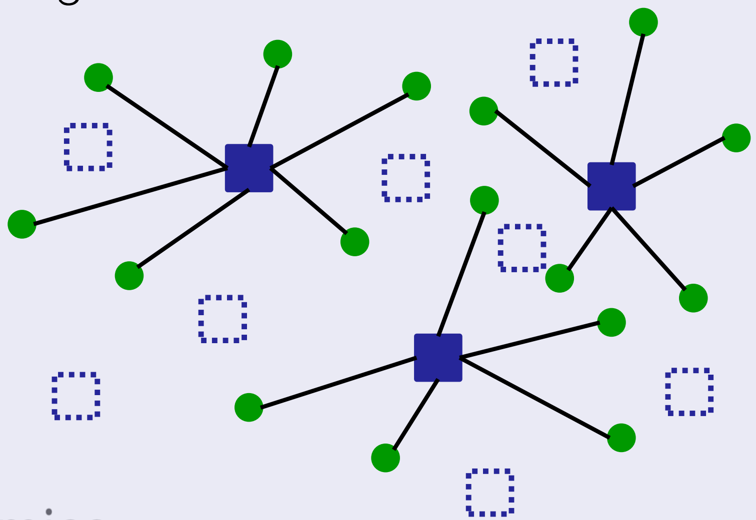
Planning a path to your destination



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

High dimensional decisions

Optimizing facility locations



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

SEQUENTIAL DECISIONS

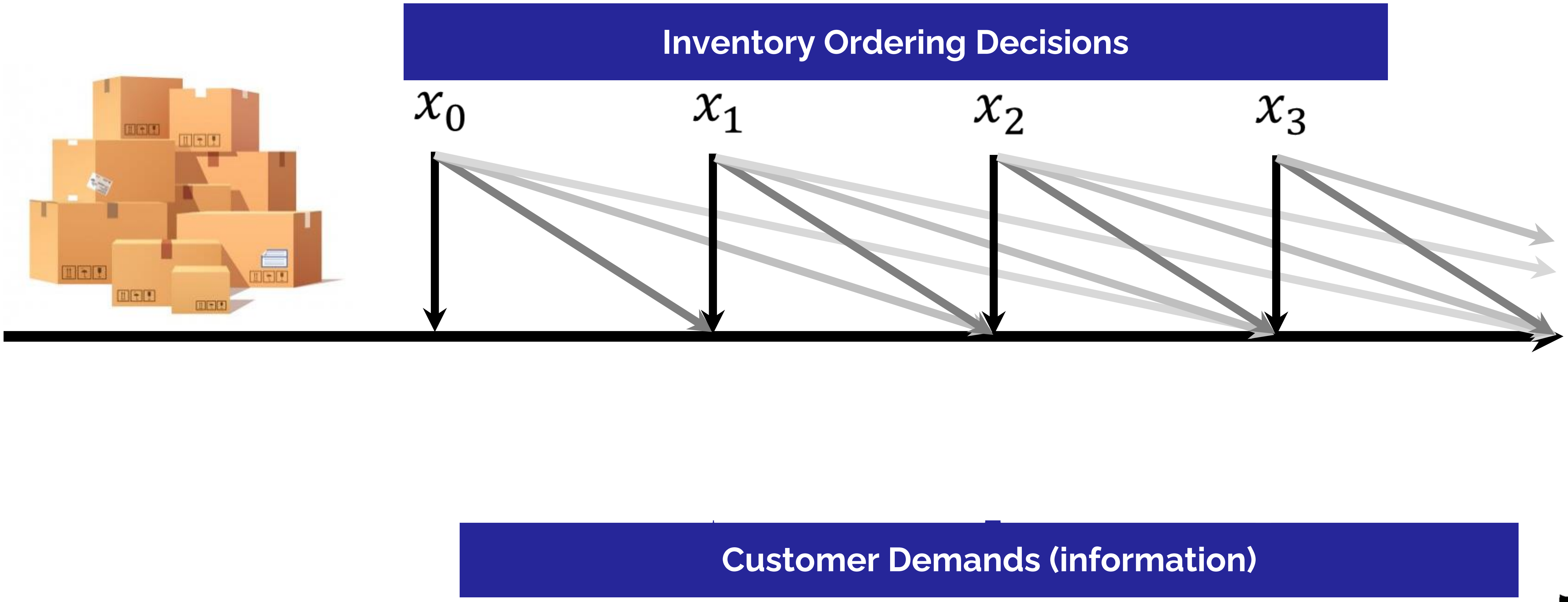
In most settings, decisions are made over time…

Decisions → Information → Decisions → Information → Decisions → Informatio

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

SEQUENTIAL DECISIONS

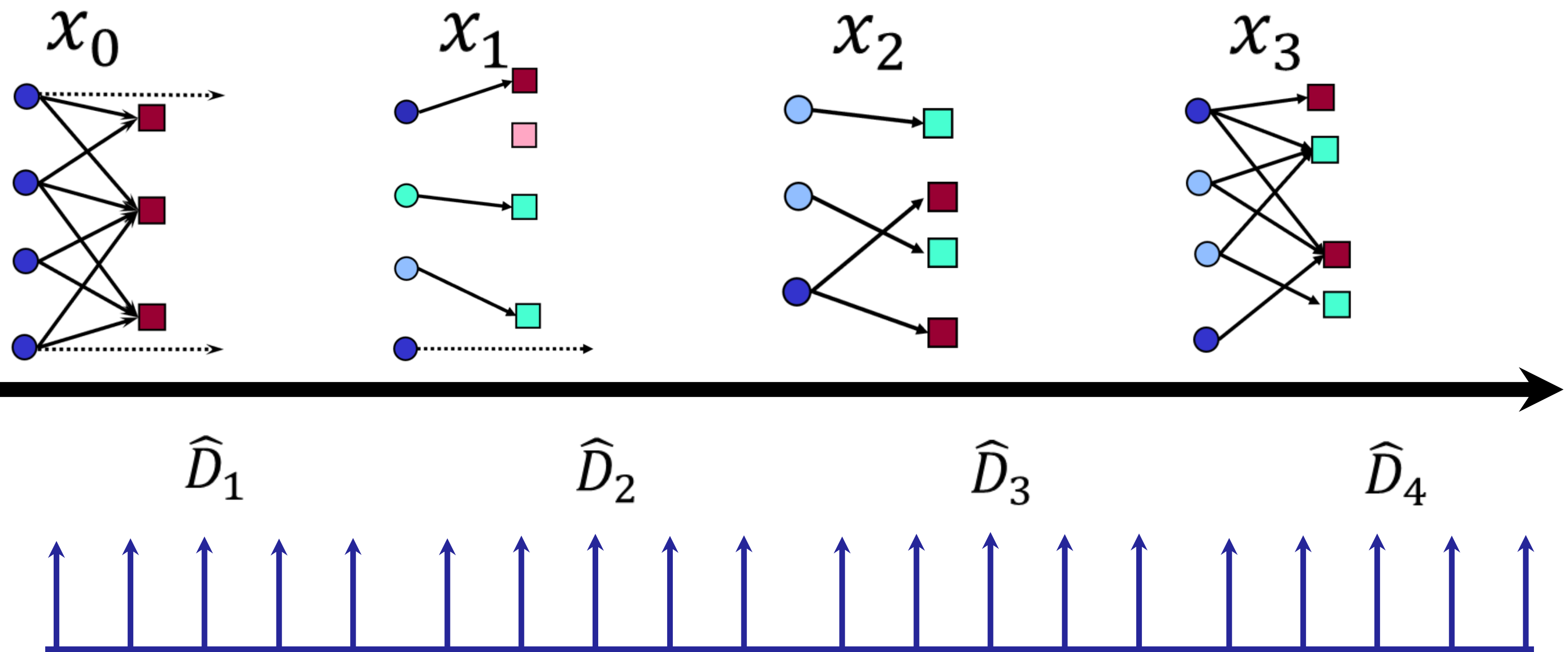
Inventory management



SEQUENTIAL DECISIONS

Driver dispatch for truckload trucking

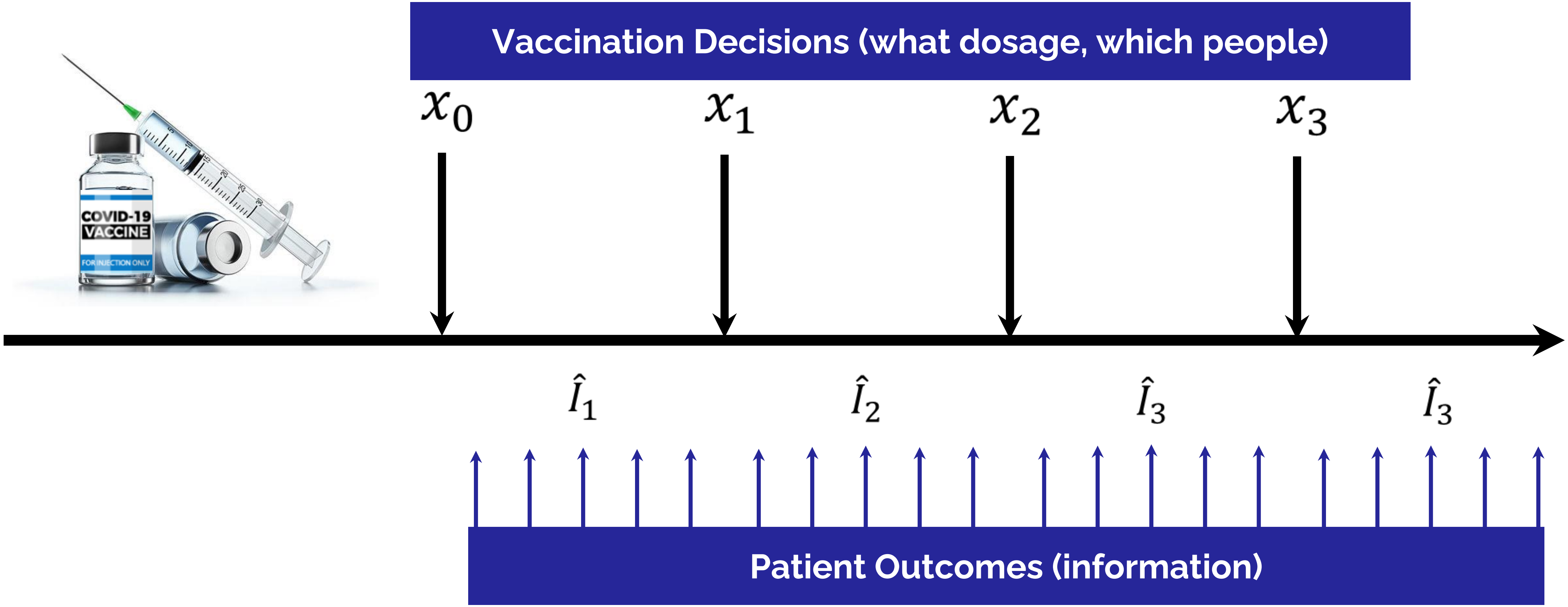
Decisions Assigning Drivers to Loads



Shippers Calling in Loads (information)

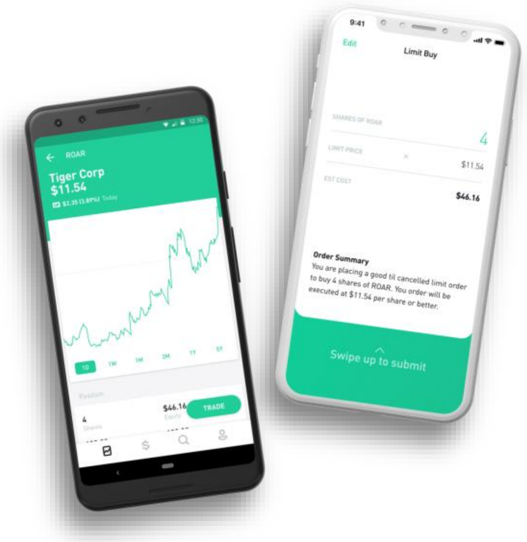
SEQUENTIAL DECISIONS

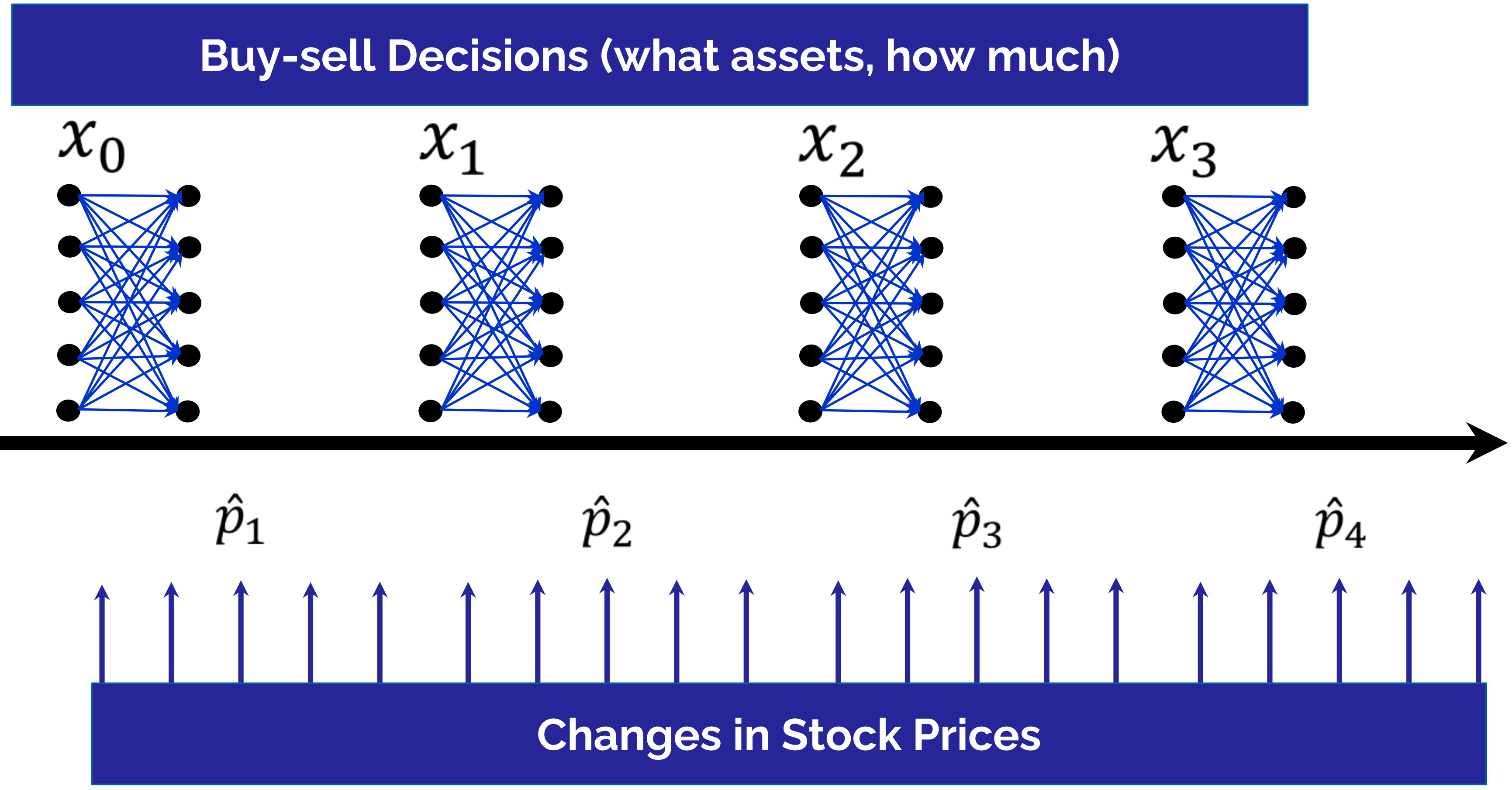
Testing new vaccines



SEQUENTIAL DECISIONS

Financial Trading






NASDAQ
AIR PRODUCTS
tell me more
www.airproducts.com
COVID 19
Coronavirus Vaccine
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 option

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 is needed to handle the variabilit of wind?

What price should be charged

Which physician should handle a procedure?

How many nurses should we have to se local hospitals and doctor's offices?

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

When should gas turbines be scheduled to handle drops in wind?

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, and which robot?

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

How many aircraft should I order for delivery in five years?

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?

Optimal Dynamics

INFORMATION

Market prices for spot freight

Driver requests for loads; time-at-home requests

Changes in asset prices

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

Employment rate; unemployment filings

Transit delays

Driver application for jobs by region

Competitor prices

Patient arrivals and symptoms

Wind generation from a wind farm

Requests for nurses from doctor's offices

Customer usage rate of liquid nitrogen

Electricity prices on the grid

Number of nurses calling in sick

Equipment failures at customer nitrogen tanks

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

Lead times required by each manufacturer

Availability of production capacity for new jets

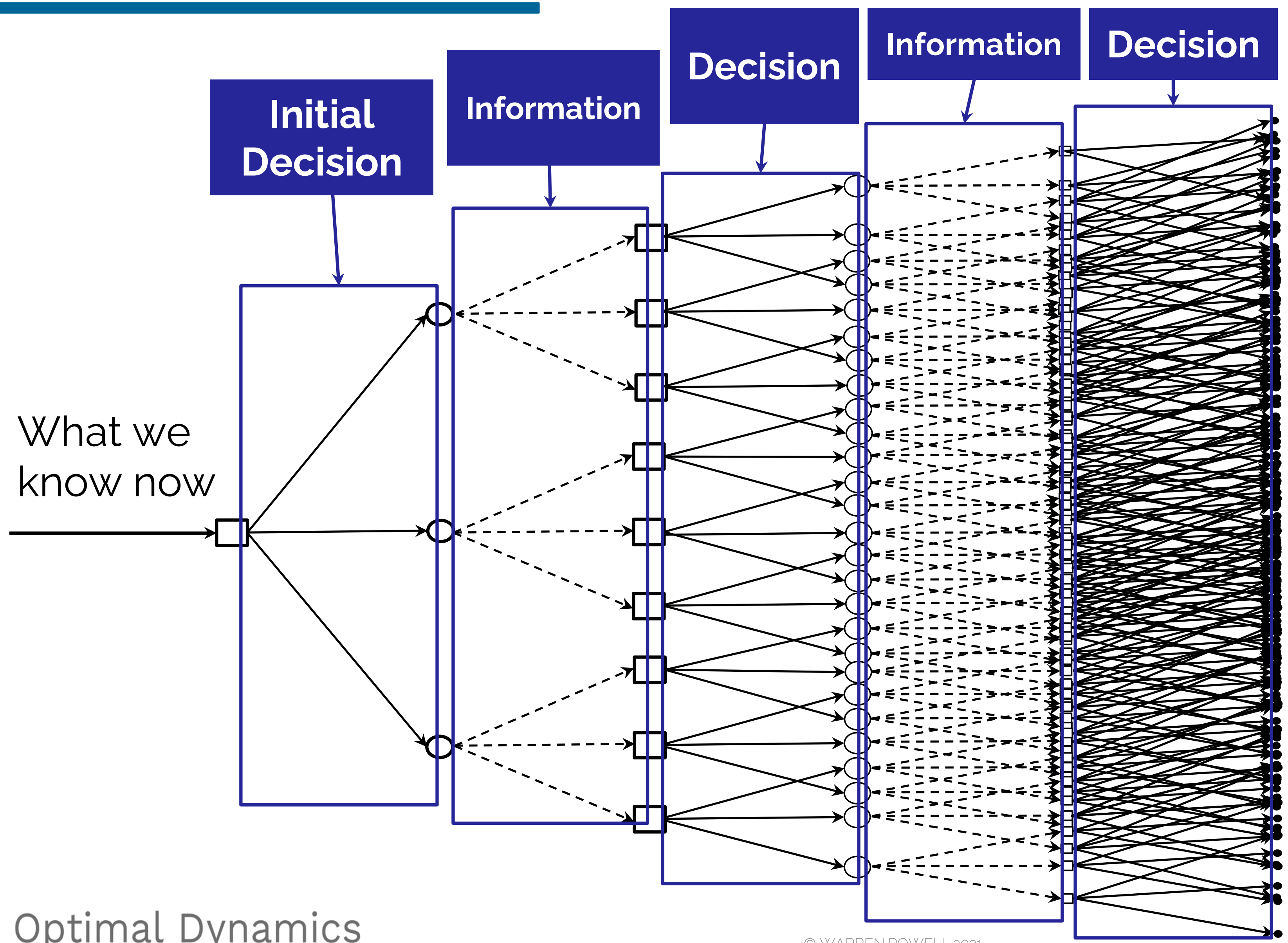
Daily production of new jet engines

Flow of orders for a product by region around the country

Orders for a product from different regions

Optimal Dynamics

SEQUENTIAL DECISIONS



Even small sequential decision problems explode dramatically as we plan into the future

OUTLINE

- → The five layers of intelligence
- → Modeling sequential decision problems
- → Designing policies
  - → Policy function approximations
  - → Cost function approximations
  - → Value function approximations
  - → Direct lookahead approximations
- → 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…

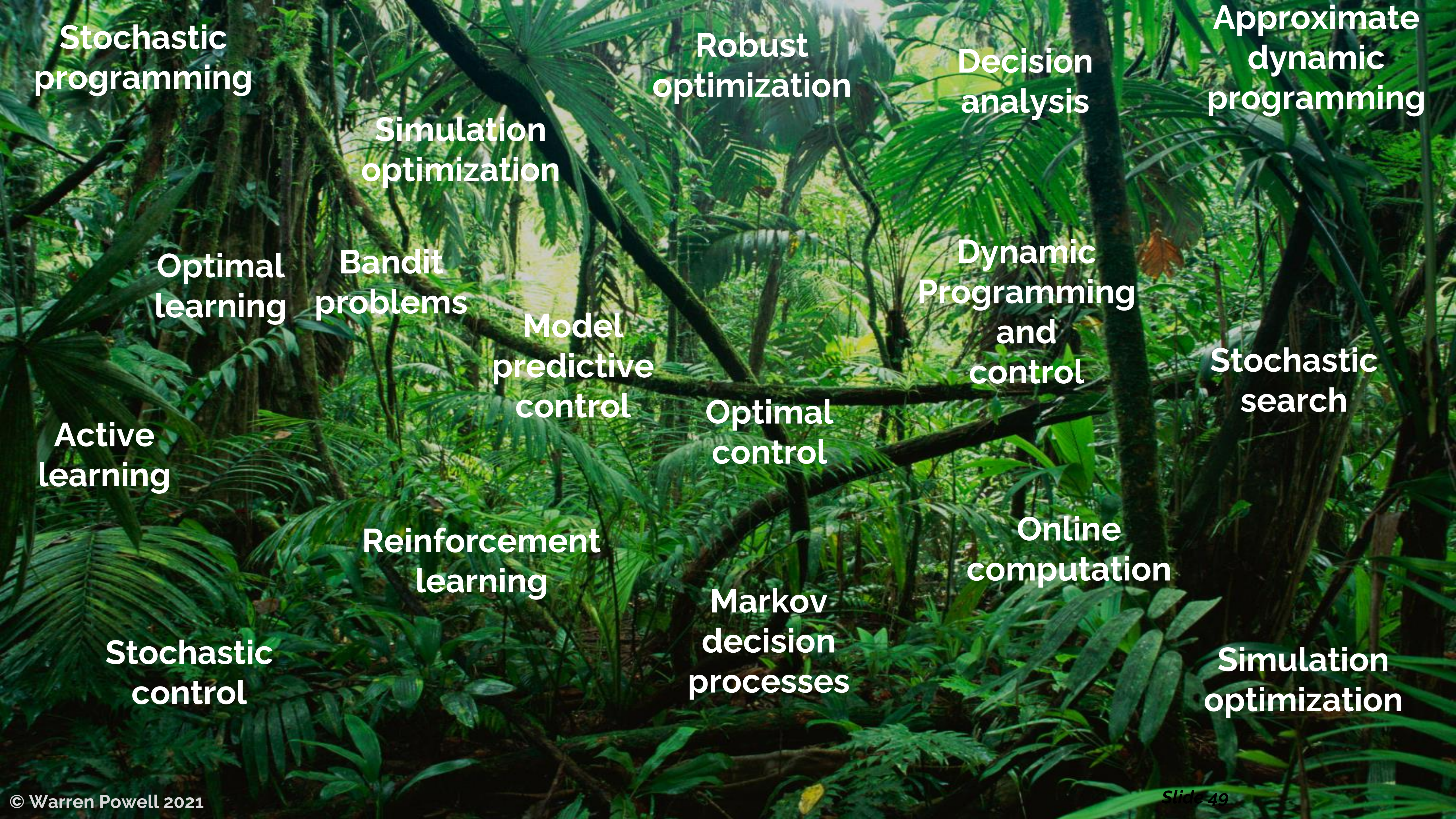
$$\text{Min } E\left\{\sum cx\right\}$$
$$Ax = b$$
$$x \geq 0$$

Mathematician

Organize class libraries, and set up communications and databases

…we lack a standard modeling framework for sequential decisions.

Optimal Dynamics


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


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 Improving Decision Quality
SECOND EDITION
Approximate Dynamic Programming
Solving the Curses of Dimensionality
Warren B. Powell
Wiley Series in Probability and Statistics
Optimal Learning
B. Powell
Ryzhov
MULTI-ARMED BANDIT ALLOCATION INDICES
SECOND EDITION
John Gittins, Kevin Glazebrook and Richard Weber
SECOND EDITION
Model Predictive Control
OPTIMAL CONTROL
VOLUME 2 • 4th EDITION
Dynamic Programming and Optimal Control
APPROXIMATE DYNAMIC PROGRAMMING
Dimitri P. Bertsekas
WILEY
Wiley-Interscience Series in Discrete Mathematics and Optimization
INTRODUCTION TO STOCHASTIC SEARCH AND OPTIMIZATION
Estimation, Simulation, and Control
JAMES C. SPALL
MORGAN & CLAYPOOL PUBLISHERS
Active Learning
Burr Settles
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
Online Computation and Competitive Analysis
Allan Borodin
Ran El-Yaniv
Vol. 1
STOCHASTIC SIMULATION OPTIMIZATION
An Optimal Computing Budget Allocation
Chun-Hung Chen • Loo Hay Lee

SEQUENTIAL DECISIONS

- Any sequential decision problems can be written:

$$(S_0, x_0, W_1, S_1, x_1, W_2, \ldots, S_t, x_t, W_{t+1}, S_{t+1} \ldots, S_T)$$

What we know (or believe)

The decision

What we observe (or learn)

- Each time we make a decision, we receive a contribution $C(S_t, x_t)$.
- Decisions are made with a method or *policy* $X^\pi(S_t)$.
- The goal is to find the policy that maximizes expected contributions:

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

Optimal Dynamics

MODELING SEQUENTIAL DECISION PROBLEMS

Every sequential decision problem can be modeled using 5 core components

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

These five elements describe any sequential decision problem.

Modeling supply chain problems

We start by identifying:

- What are the performance metrics you are focusing on?
- What decisions are involved?
- What are the sources of uncertainty and new information?
- What information is needed to compute metrics, make decisions, and model their evolution over time ("state variables")

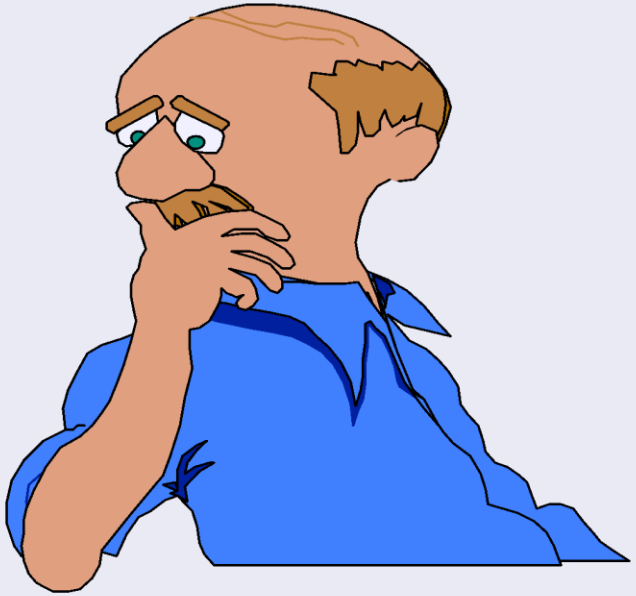
OUTLINE

- → The five layers of intelligence
- → Modeling sequential decision problems
- → Designing policies
  - → Policy function approximations
  - → Cost function approximations
  - → Value function approximations
  - → Direct lookahead approximations
- → A new educational field: sequential decision analytics

DESIGNING POLICIES

Policy Definition

A policy is method that maps a state variable into a decision … *any method.*



DESIGNING POLICIES

Policies and the English language

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

Designing policies

Every sequential decision problem can be modeled using 5 core components

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

These five elements describe any sequential decision problem.

Optimal Dynamics

Evaluating policies

1) Theoretically

- Optimality proofs
- Regret bounds
- Asymptotic convergence

2) Through numerical simulations

3) In the field

Optimal Dynamics

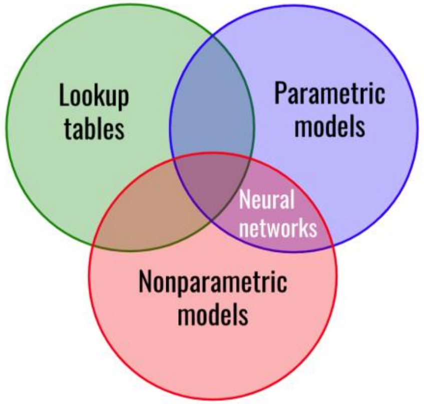
BRIDGING MACHINE LEARNING & SEQUENTIAL DECISIONS

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

Optimal Dynamics

DESIGNING POLICIES

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.

DESIGNING POLICIES

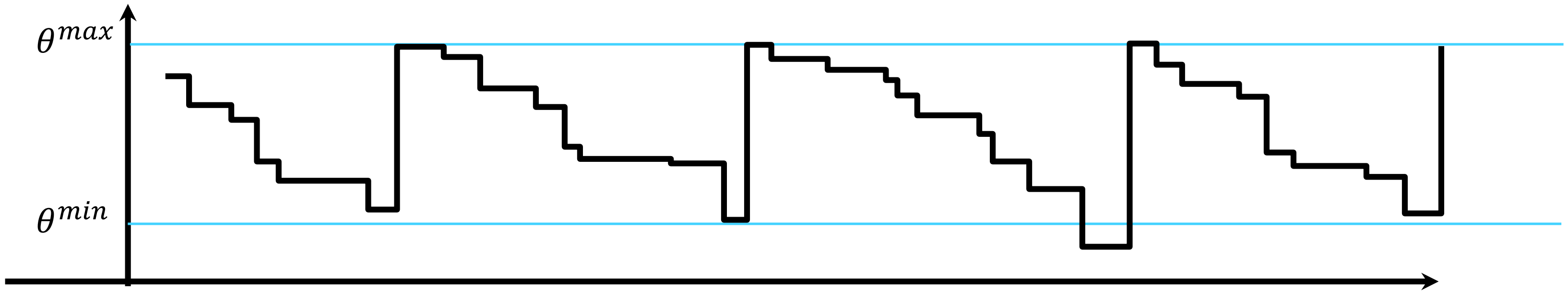
Policy search

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 $\theta = (\theta^{min}, \theta^{max})$



b) Buy when the price goes **below** $\boldsymbol{\theta^{min}}$
and sell when it goes **above** $\boldsymbol{\theta^{max}}$

c) Lookup tables, linear/nonlinear models, neural networks, nonparametric models, …

DESIGNING POLICIES

Policy search

2) Cost function approximations (CFAs)

These are parameterized optimization problems:

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

b) Optimize energy generation for tomorrow to meet forecasted demand, but add reserves θ in case of a generator failure.

c) Advertise the product x which solves:

$$X^{UCB}(S^n|\theta) = \max_x (Estimated\ revenue_x^n + \theta \cdot Standard\ deviation\ of\ estimated\ revenue_x^n)$$

Now solve: $\max_\theta \mathbb{E}\{\sum_{n=1}^N C(S^n, X^\pi(S^n|\theta))\,|S_0\}$

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

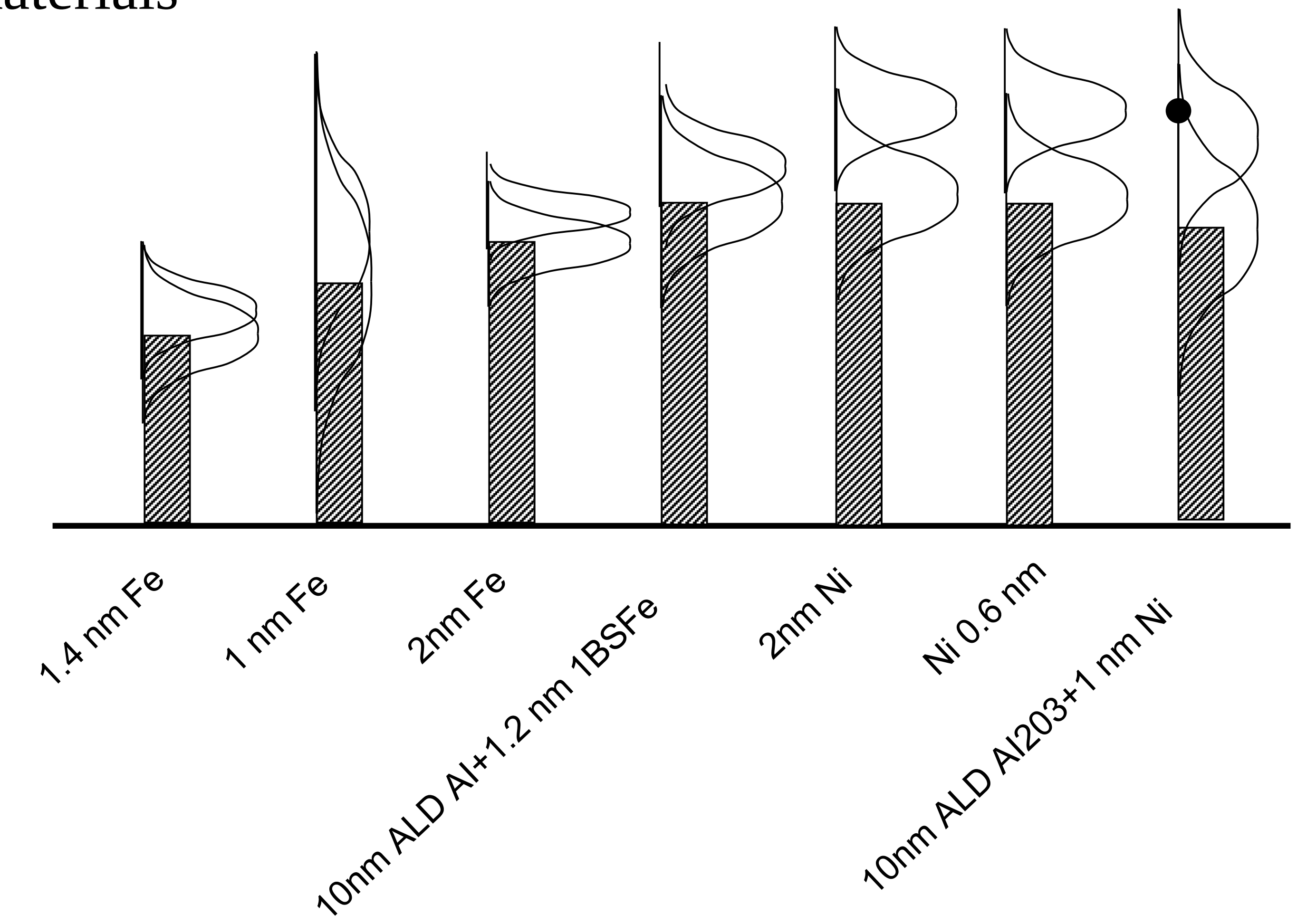
Cost function approximations

- Lookup table
  - » We can organize potential catalysts into groups
  - » Scientists using domain knowledge can estimate correlations in experiments between similar catalysts.

	1.4 nm Fe	1 nm Fe	2nm Fe	10nm ALD Al2O3+1.2 nm IBS Fe	2 nm Ni	Ni 0.6 nm	10nm ALD Al2O3+1 nm Ni
1.4 nm Fe	1	0.7	0.7	0.6	0.4	0.4	0.2
1 nm Fe	0.7	1	0.7	0.6	0.4	0.4	0.2
2nm Fe	0.7	0.7	1	0.6	0.4	0.4	0.2
10nm ALD Al2O3+1.2 nm IBS Fe	0.6	0.6	0.6	1	1	0.3	0
2 nm Ni	0.4	0.4	0.4	1	1	0.7	0.6
Ni 0.6 nm	0.4	0.4	0.4	0.3	0.7	1	0.6
10nm ALD Al2O3+1 nm Ni	0.2	0.2	0.2	0	0.6	0.6	1

Cost function approximations

- Correlated beliefs: Testing one material teaches us about other materials

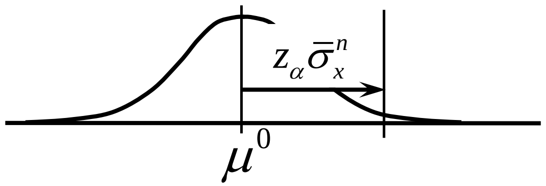


Cost function approximations

- 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



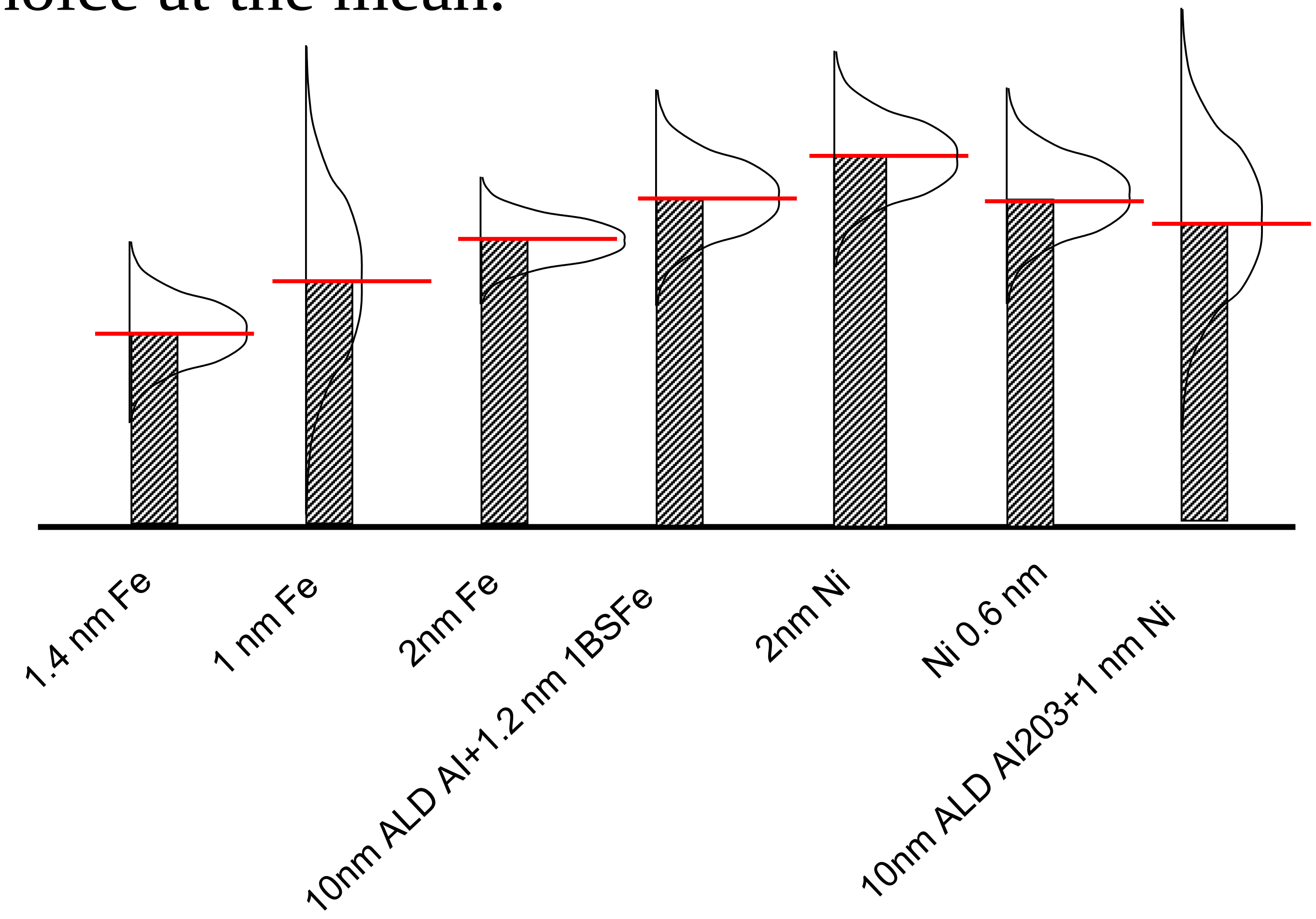
$$X^{IE}(S^n \mid \theta^{IE}) = \arg\max_x \left(\bar{\mu}_x^n + \theta^{IE} \bar{\sigma}_x^n \right)$$

  - » Thompson sampling

$$x^n = argmax_x \hat{\mu}_x^n \qquad \hat{\mu}_x^n \sim N(\bar{\mu}_x^n, \theta^{TS} \bar{\sigma}_x^{2,n})$$

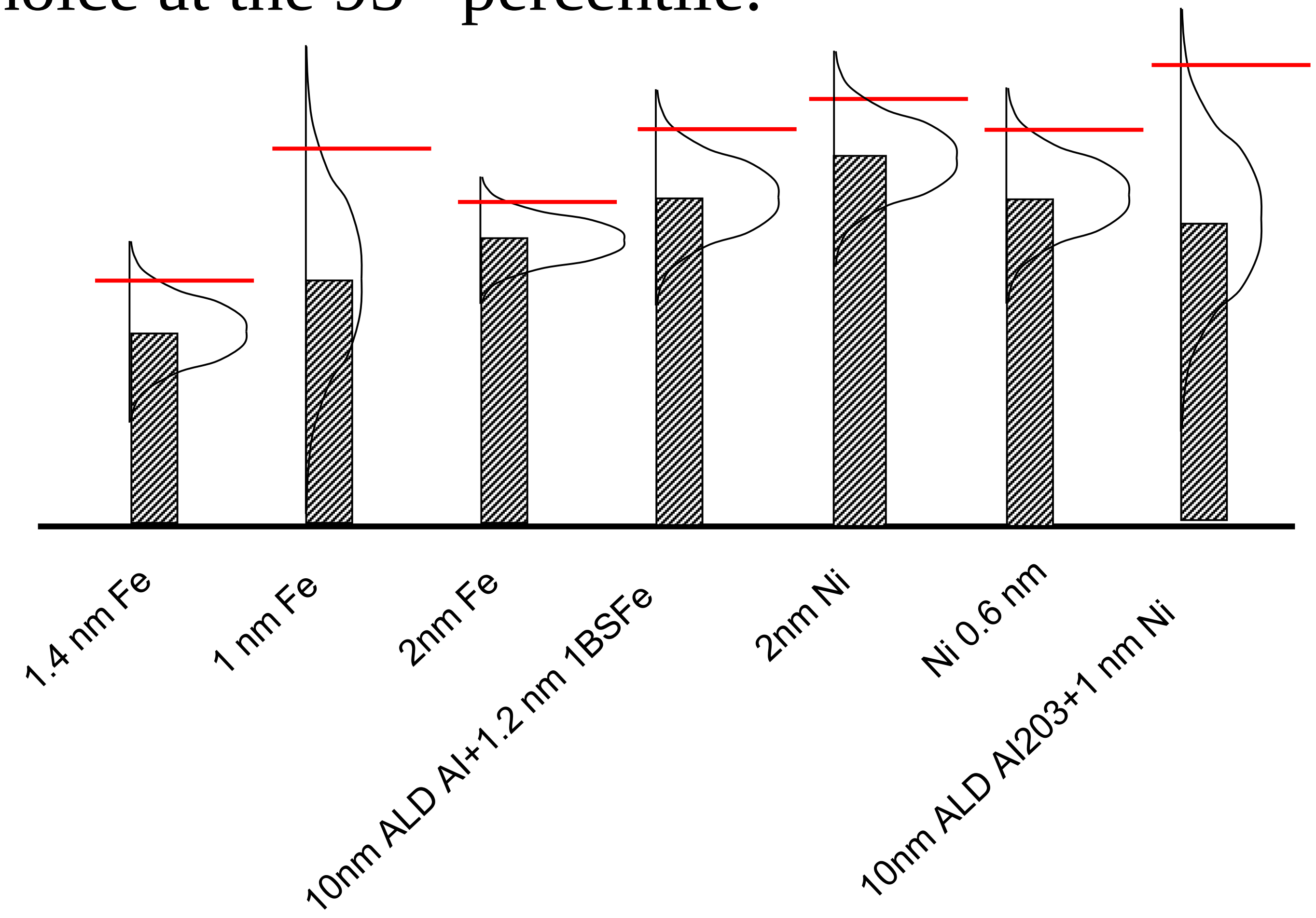
Cost function approximations

- Picking $\theta^{IE} = 0$ means we are evaluating each choice at the mean.



Cost function approximations

- Picking $\theta^{IE} = 2$ means we are evaluating each choice at the 95th percentile.



Cost function approximations

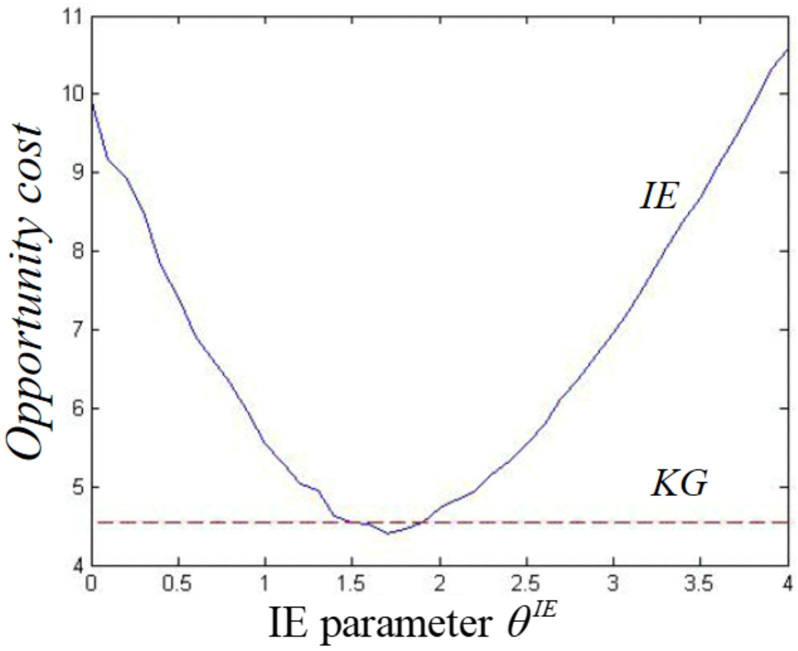
- Optimizing the policy
  - » We optimize θ^{IE} to maximize:

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

where

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

- Notes:
  - » This can handle any belief mode including correlated beliefs, nonlinear belief models.
  - » All we require is that we be able to simulate a policy.



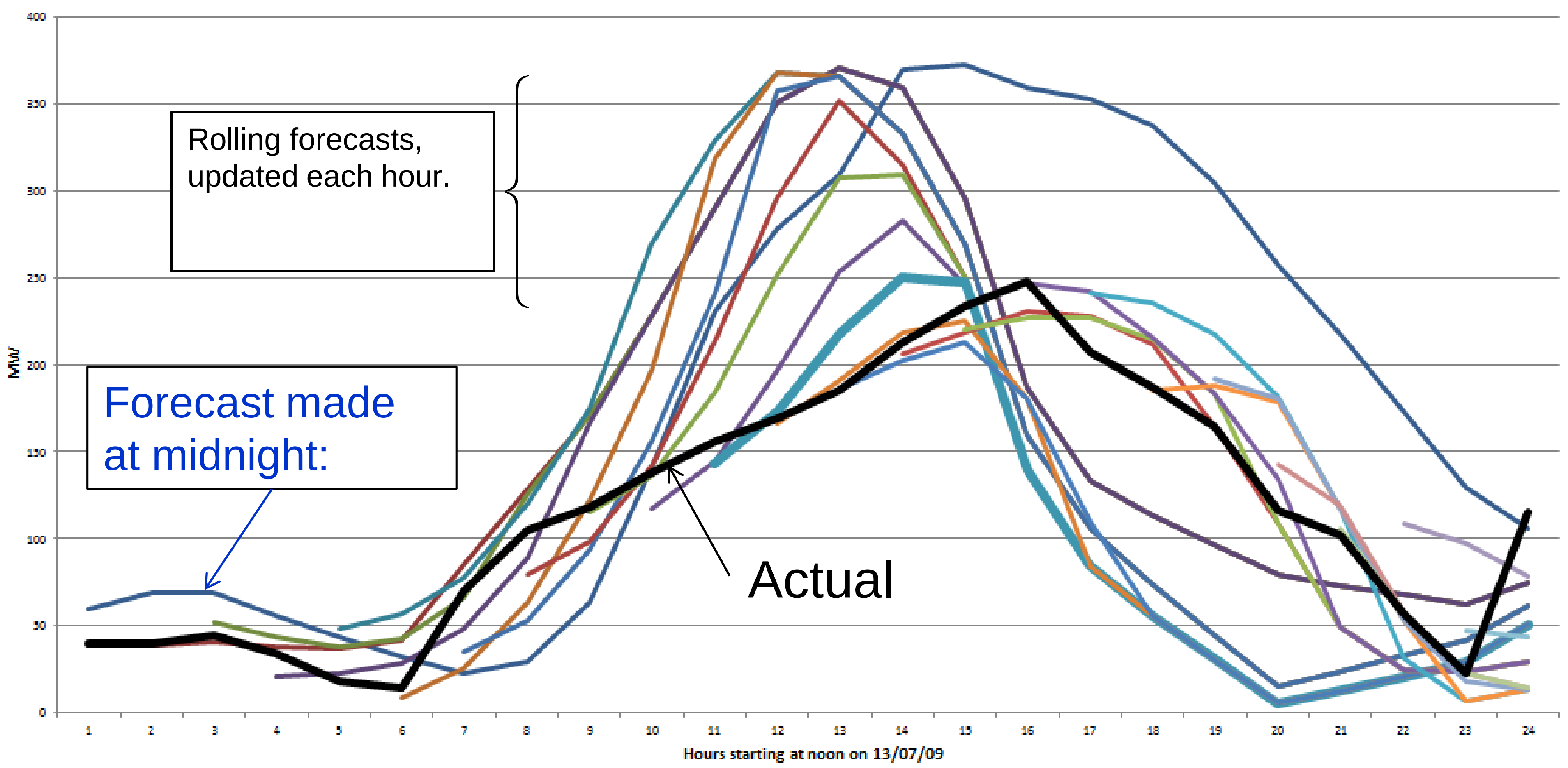
Hybrid direct lookahead/CFA

- An energy storage problem:

Optimal Dynamics

Hybrid direct lookahead/CFA

- Forecasts evolve over time as new information arrives:



Optimal Dynamics

Hybrid direct lookahead/CFA

- Benchmark policy – Deterministic lookahead

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

$$\begin{aligned}
\tilde{x}^{wd}_{tt'} + \beta \tilde{x}^{rd}_{tt'} + \tilde{x}^{gd}_{tt'} &\le f^{D}_{tt'} \\
\tilde{x}^{gd}_{tt'} + \tilde{x}^{gr}_{tt'} &\le f^{G}_{tt'} \\
\tilde{x}^{rd}_{tt'} + \tilde{x}^{rg}_{tt'} &\le \tilde{R}_{tt'} \\
\tilde{x}^{wr}_{tt'} + \tilde{x}^{gr}_{tt'} &\le R^{ma\tilde{x}_{tt'}} \\
\boxed{\tilde{x}^{wr}_{tt'} + \tilde{x}^{wd}_{tt'} \le f^{E}_{tt'}} \\
\tilde{x}^{wr}_{tt'} + \tilde{x}^{gr}_{tt'} &\le \gamma^{ch\arg e} \\
\tilde{x}^{rd}_{tt'} + \tilde{x}^{rg}_{tt'} &\le \gamma^{disch\arg e}
\end{aligned}$$

Hybrid direct lookahead/CFA

- Benchmark policy – Deterministic lookahead

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

$$\tilde{x}_{tt'}^{wd} + \beta \tilde{x}_{tt'}^{rd} + \tilde{x}_{tt'}^{gd} \leq f_{tt'}^{D}$$
$$\tilde{x}_{tt'}^{gd} + \tilde{x}_{tt'}^{gr} \leq f_{tt'}^{G}$$
$$\tilde{x}_{tt'}^{rd} + \tilde{x}_{tt'}^{rg} \leq \tilde{R}_{tt'}$$
$$\tilde{x}_{tt'}^{wr} + \tilde{x}_{tt'}^{gr} \leq R^{\mathrm{ma\tilde{x}}_{tt'}}$$
$$\boxed{\tilde{x}_{tt'}^{wr} + \tilde{x}_{tt'}^{wd} \leq \theta_{t'-t} f_{tt'}^{E}}$$
$$\tilde{x}_{tt'}^{wr} + \tilde{x}_{tt'}^{gr} \leq \gamma^{ch\,\mathrm{arg}\,e}$$
$$\tilde{x}_{tt'}^{rd} + \tilde{x}_{tt'}^{rg} \leq \gamma^{disch\,\mathrm{arg}\,e}$$

Hybrid direct lookahead/CFA

- Benchmark policy – Deterministic lookahead

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

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

$$\tilde{x}^{gd}_{tt'} + \tilde{x}^{gr}_{tt'} \le f^{G}_{tt'}$$

$$\tilde{x}^{rd}_{tt'} + \tilde{x}^{rg}_{tt'} \le \tilde{R}_{tt'}$$

$$\tilde{x}^{wr}_{tt'} + \tilde{x}^{gr}_{tt'} \le R^{\max\tilde{x}_{tt'}}$$

$$\boxed{\tilde{x}^{wr}_{tt'} + \tilde{x}^{wd}_{tt'} \le \theta_{t'-t} f^{E}_{tt'}}$$

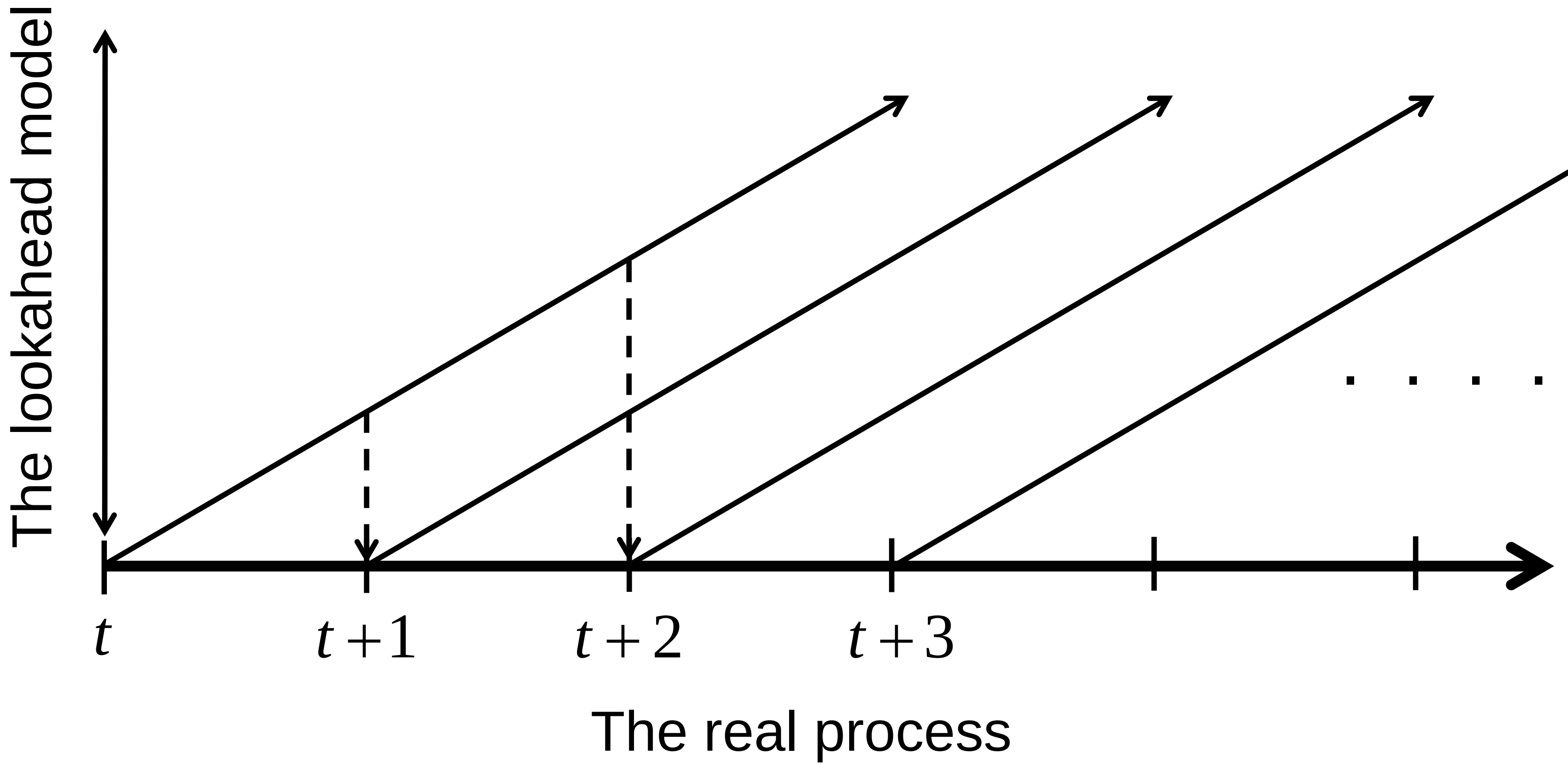
$$\tilde{x}^{wr}_{tt'} + \tilde{x}^{gr}_{tt'} \le \gamma^{ch\arg e}$$

$$\tilde{x}^{rd}_{tt'} + \tilde{x}^{rg}_{tt'} \le \gamma^{disch\arg e}$$

Optimal Dynamics

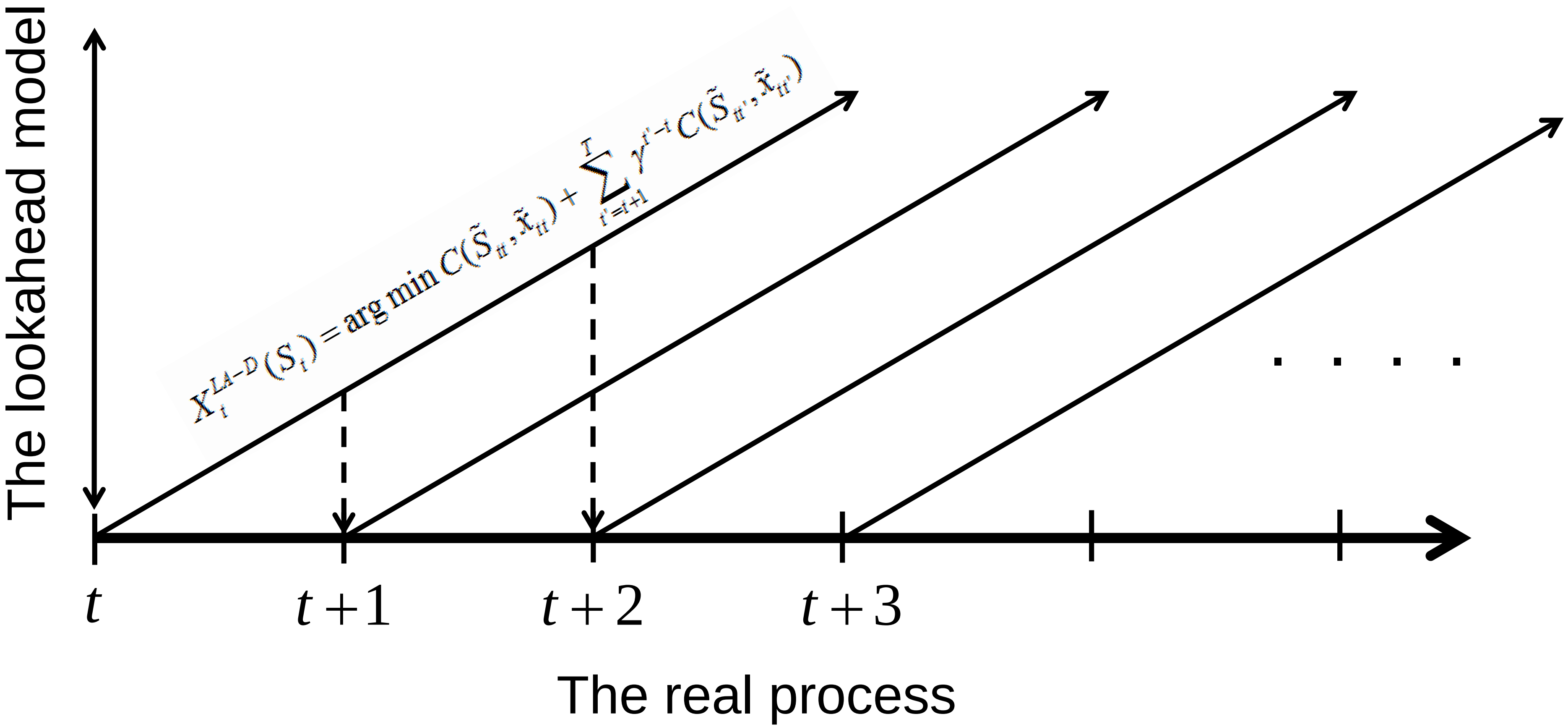
Hybrid direct lookahead/CFA

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



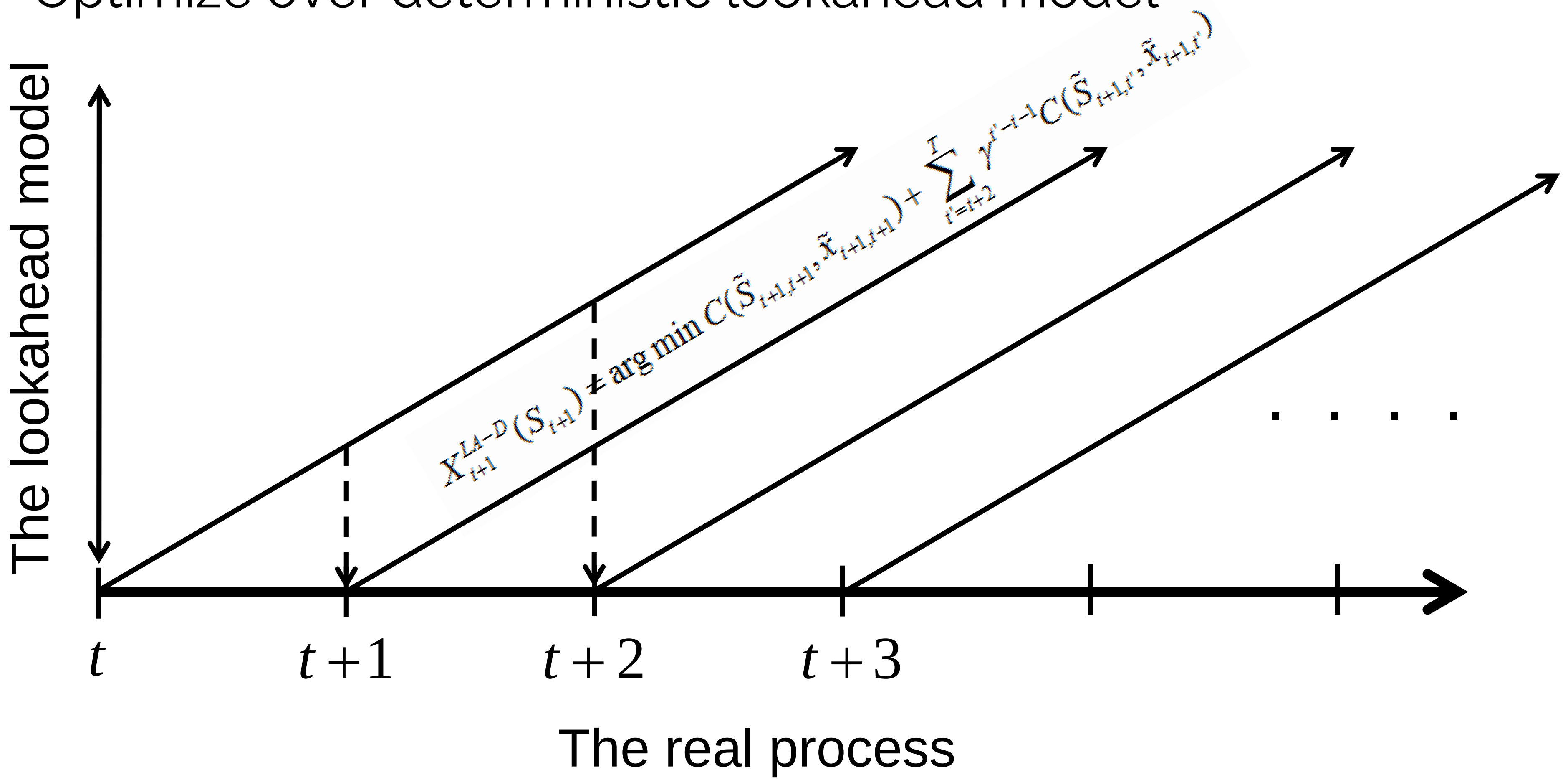
Hybrid direct lookahead/CFA

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



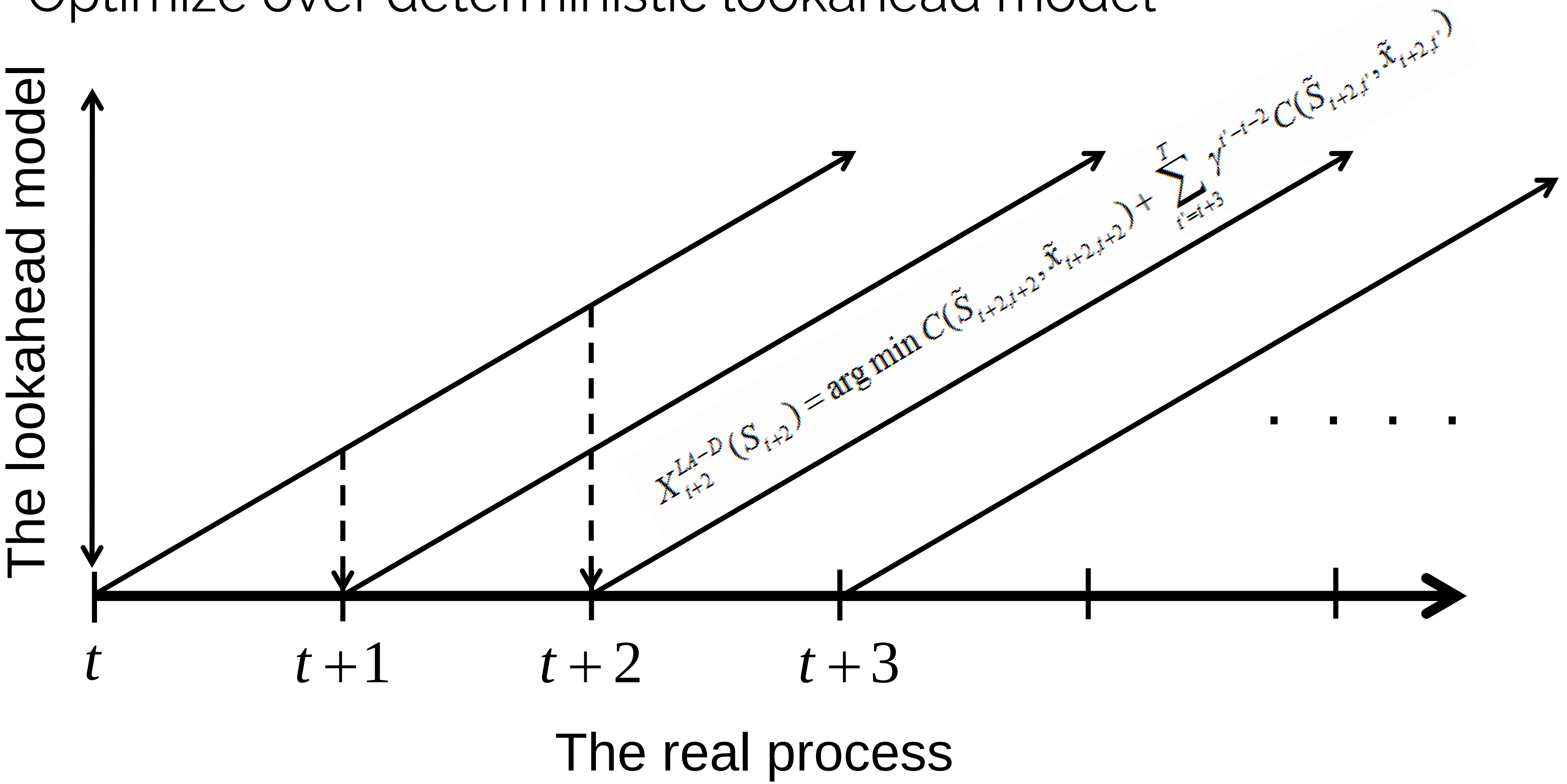
Hybrid direct lookahead/CFA

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



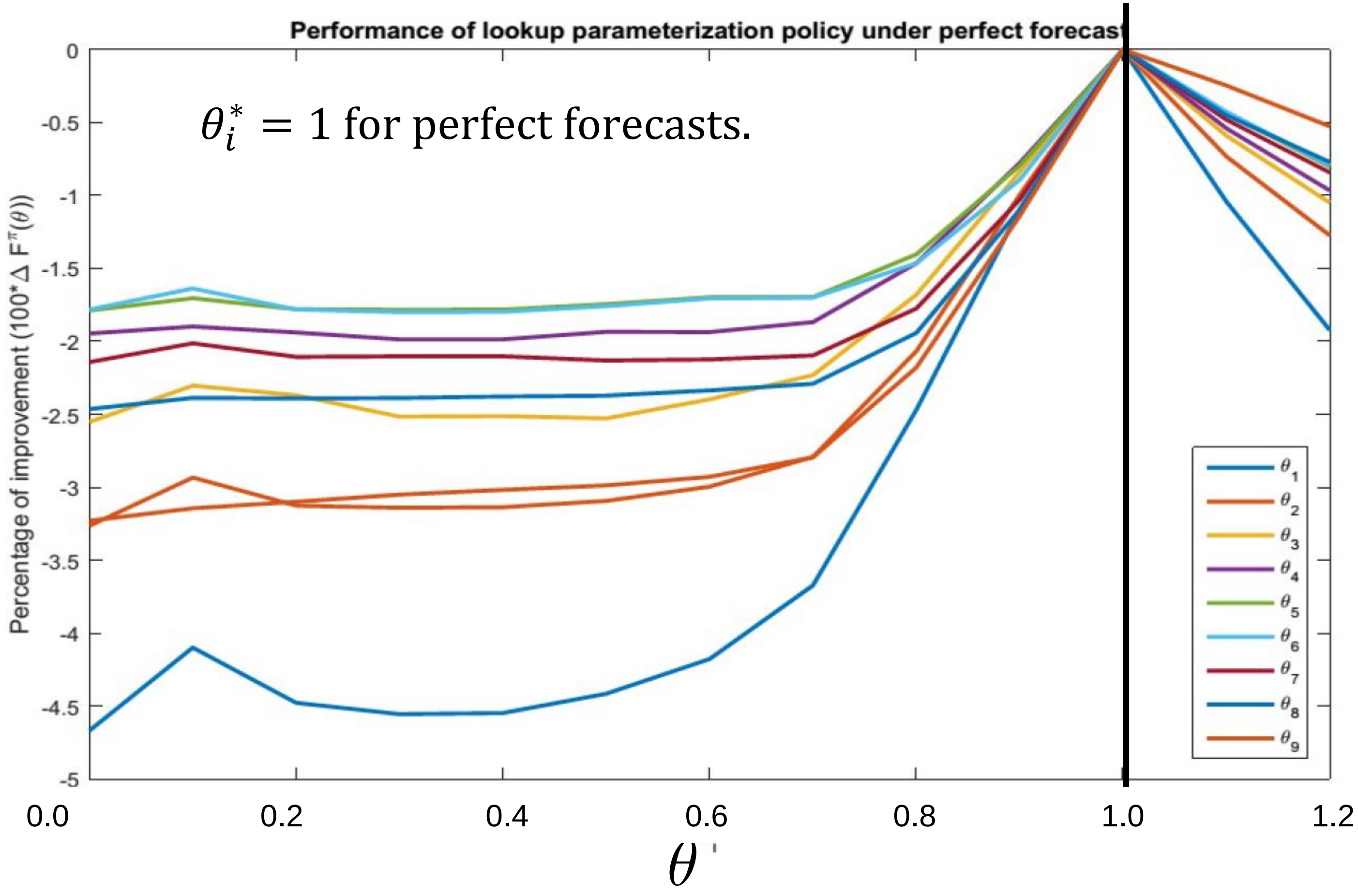
Hybrid direct lookahead/CFA

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



Hybrid direct lookahead/CFA

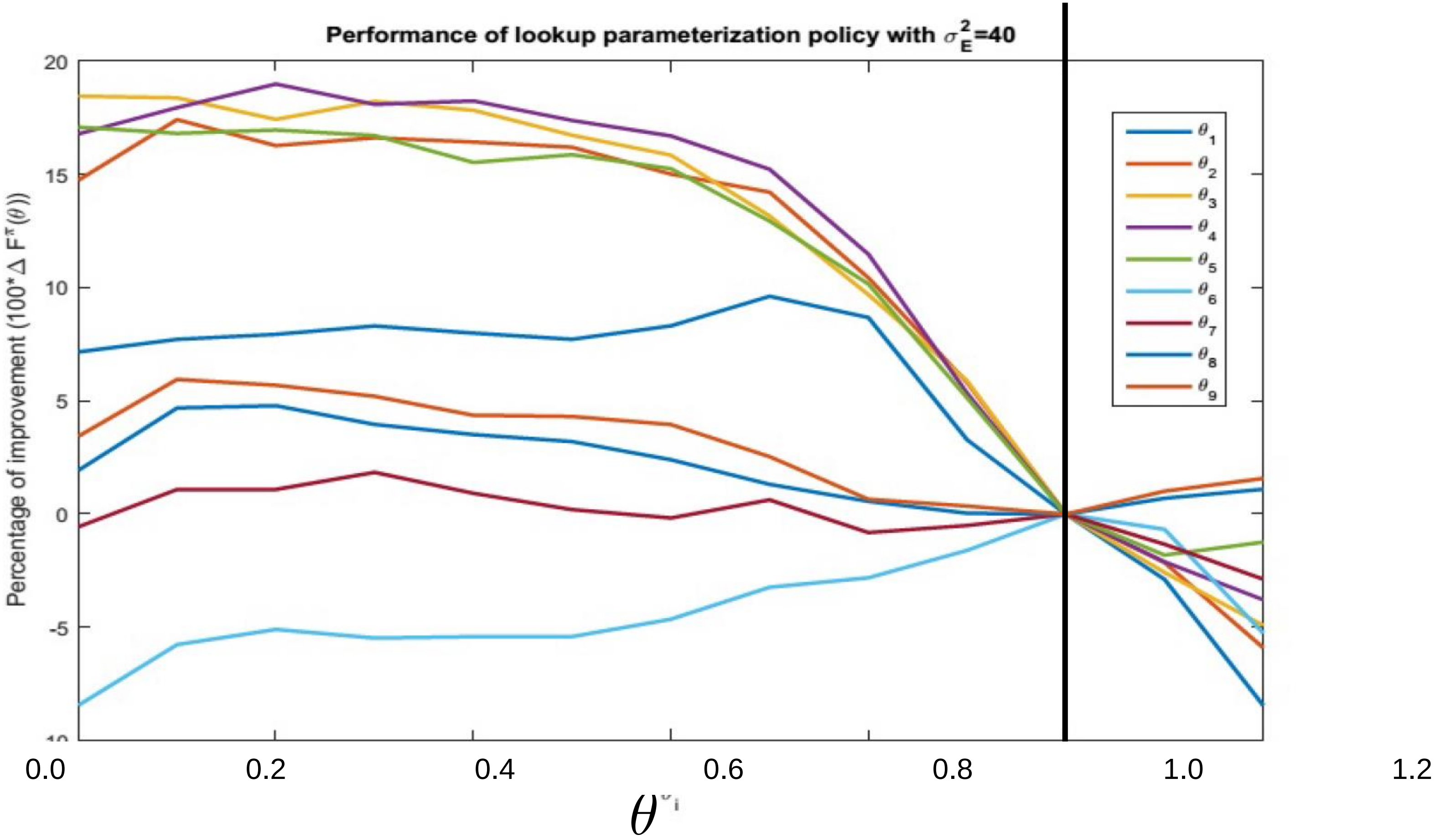
- One-dimensional contour plots – perfect forecast



Optimal Dynamics

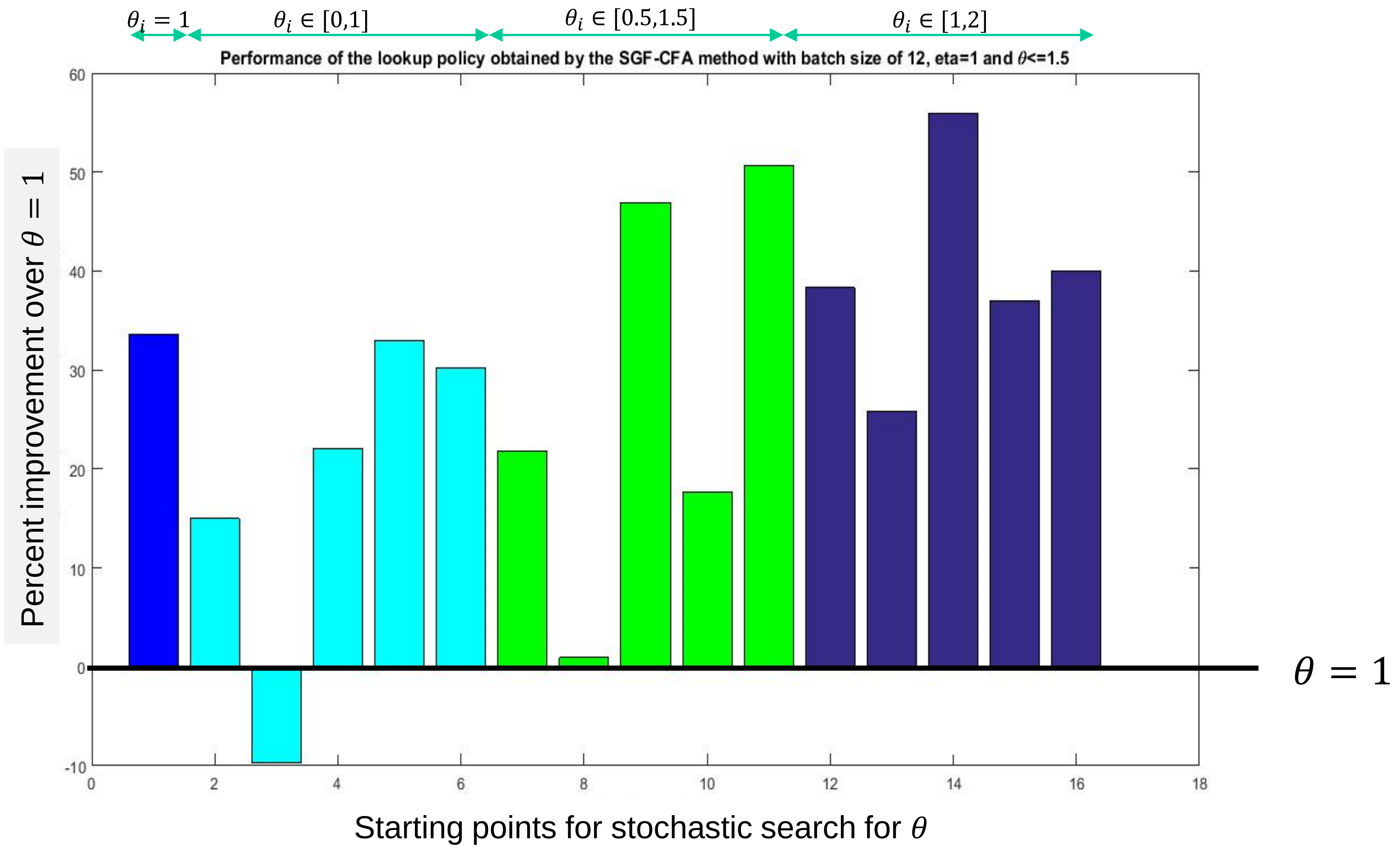
Hybrid direct lookahead/CFA

- One-dimensional contour plots-uncertain forecast



Energy storage optimization

- Tuning the parameters



Cost function approximations

- Other applications
  - » Airlines optimizing schedules with schedule slack to handle weather uncertainty.
  - » Manufacturers using buffer stocks to hedge against production delays and quality problems.
  - » Grid operators scheduling extra generation capacity in case of outages.
  - » Adding time to a trip planned by Google maps to account for uncertain congestion.

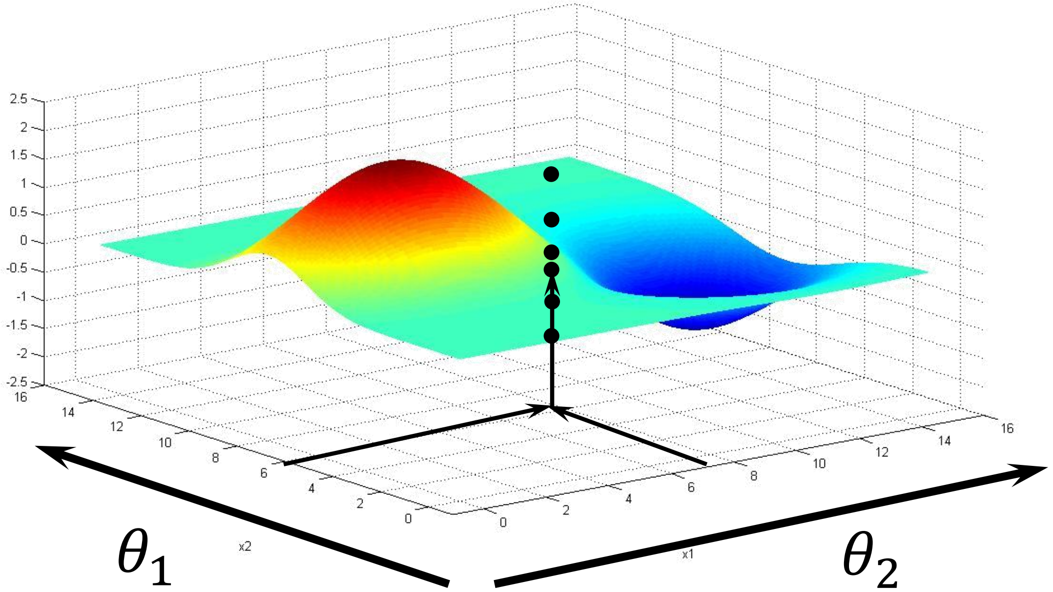
See http://tinyurl.com/cfapolicy for more on parametric cost function approximations.

Policy search

- Tuning the policy (PFAs or CFAs):
  - » We need to maximize

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

  - » We cannot compute the expectation, so we run simulations:

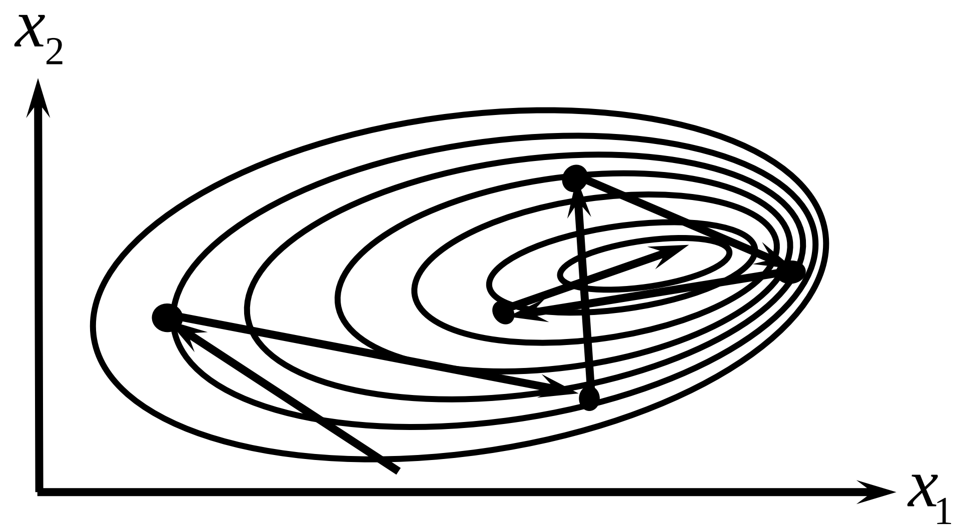


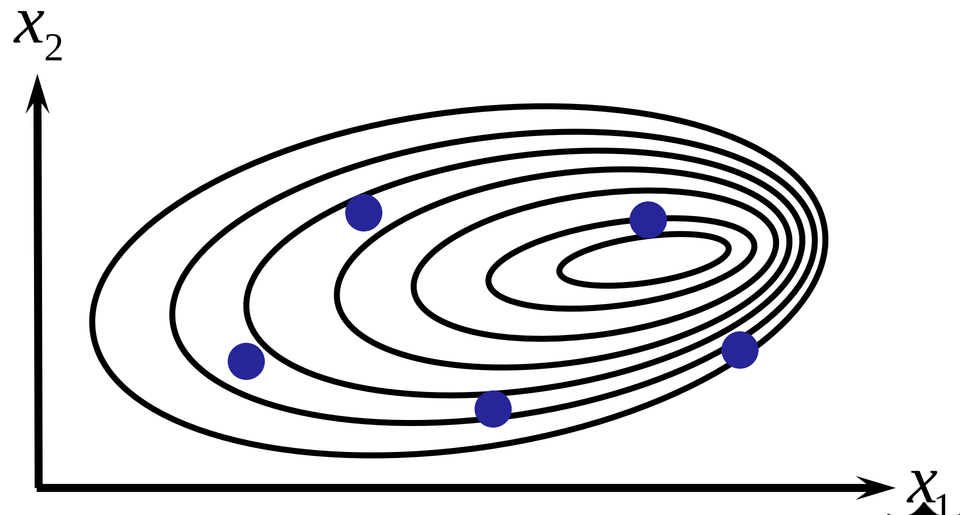
Policy function approximations

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

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

  - Derivative-free
    - Build a belief model $\bar{F}(\theta) \approx \mathbb{E}F(\theta, W)$ that approximates our function.
  - Both of these approaches are sequential decision problems!





DESIGNING POLICIES

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)

DESIGNING POLICIES

Lookahead policies are based on solving

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

Contribution we receive now

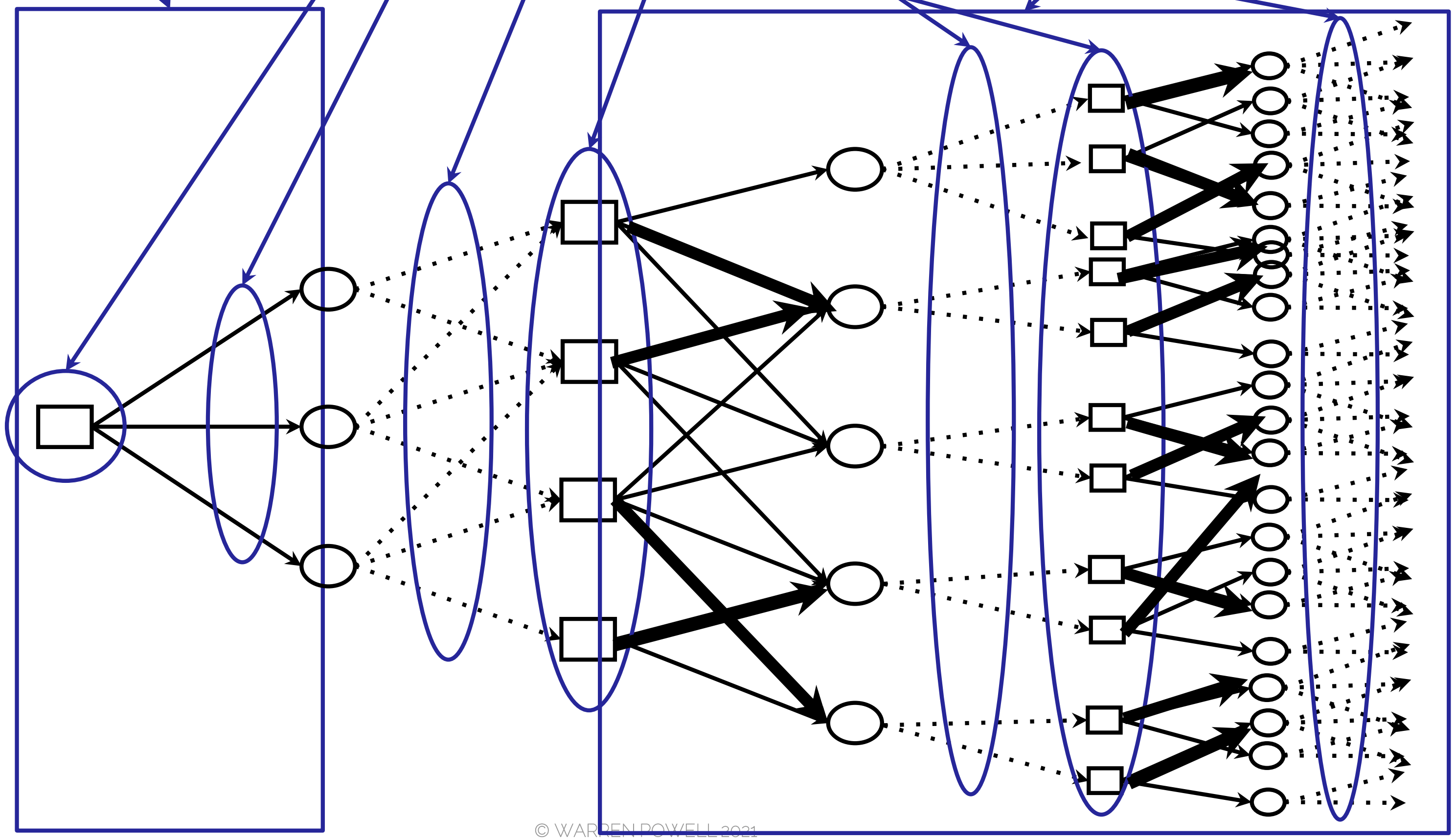
Future contributions

- 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!!!*

DESIGNING POLICIES

Lookahead policies are based on solving

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



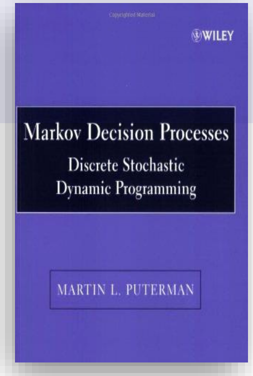
DESIGNING POLICIES

Lookahead approximations

Approximate the impact of a decision now on the future

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

3) Value function approximations (VFAs)



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

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

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

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

Optimal Dynamics

DESIGNING POLICIES

Lookahead approximations

Approximate the impact of a decision now on the future

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

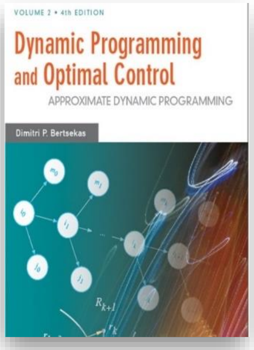
3) Value function approximations (VFAs)

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

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

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

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



DESIGNING POLICIES

Lookahead approximations

Approximate the impact of a decision now on the future

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

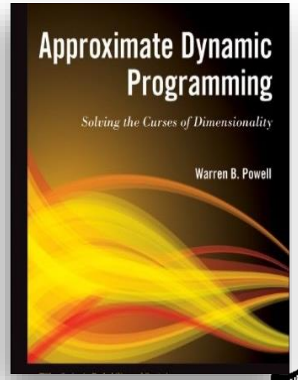
3) Value function approximations (VFAs)

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

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

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

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



Optimal Dynamics

DESIGNING POLICIES

Lookahead approximations

Approximate the impact of a decision now on the future

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

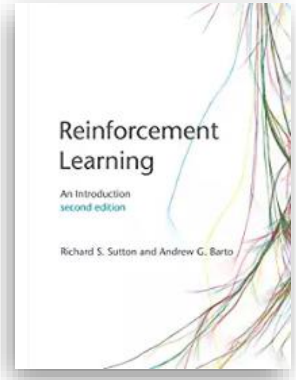
3) Value function approximations (VFAs)

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

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

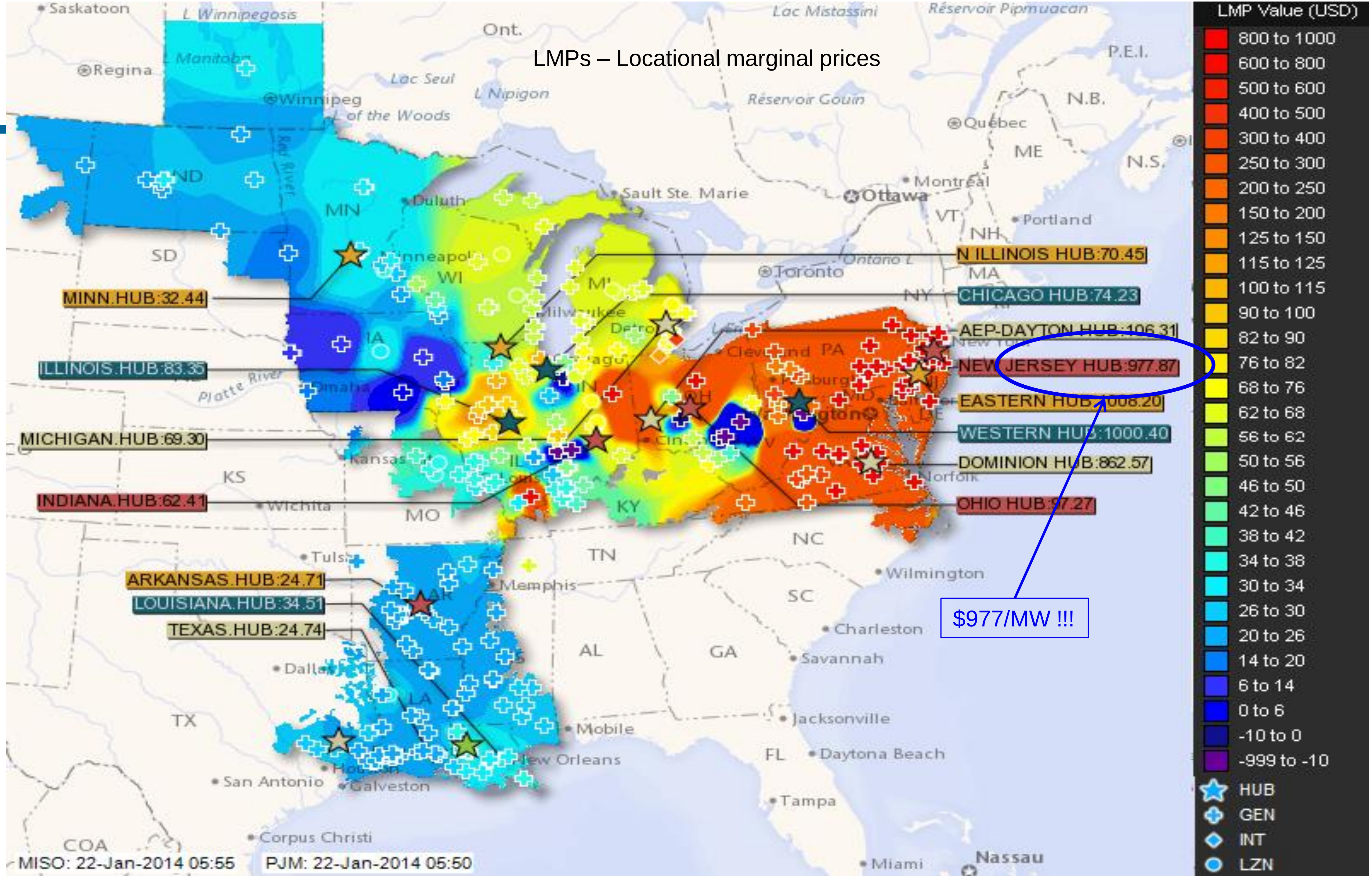
$$= \arg\max_{x_t} \left(C(S_t, x_t) + \bar{V}_t^x(S_t^x) \right)$$

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



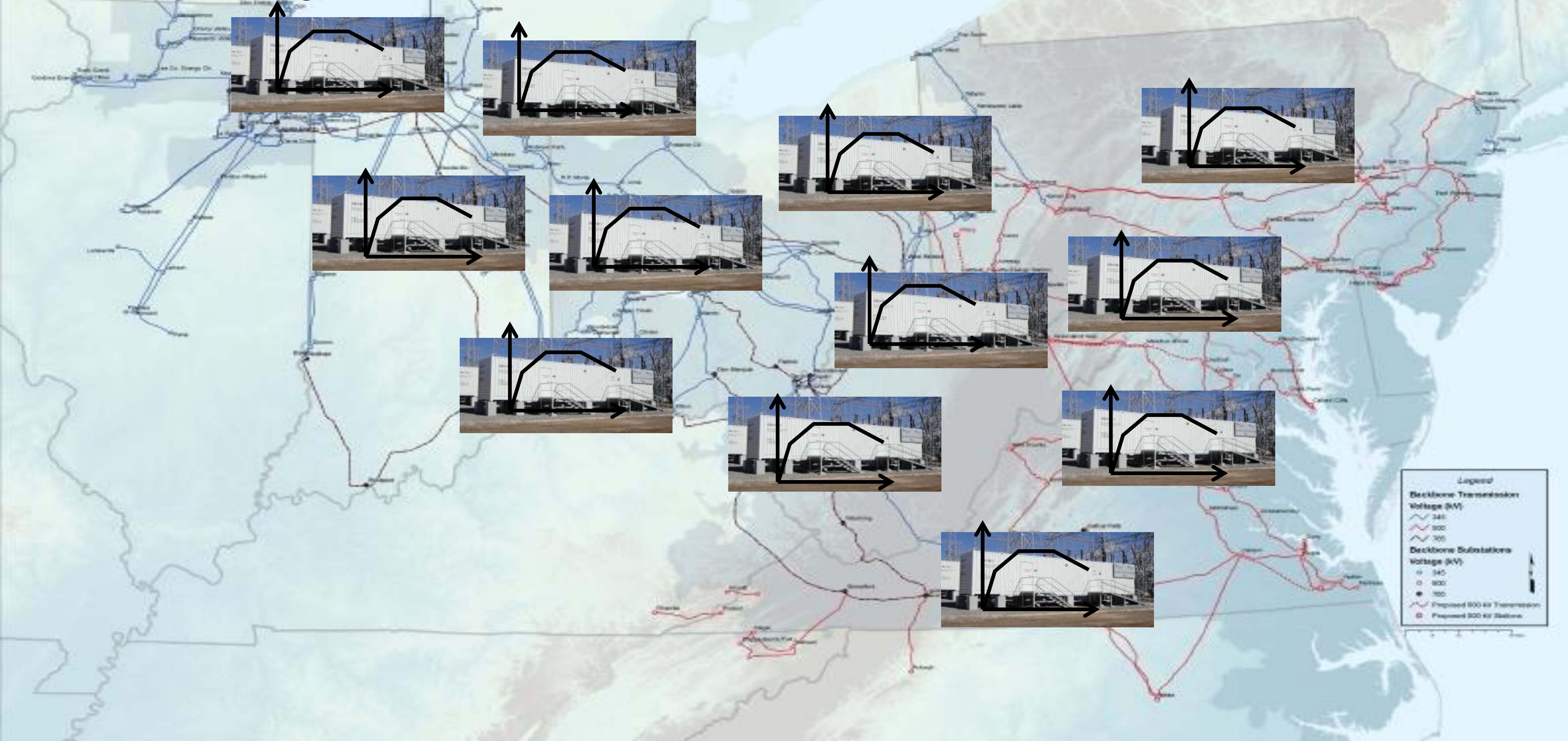
Optimal Dynamics

LMPs – Locational marginal prices

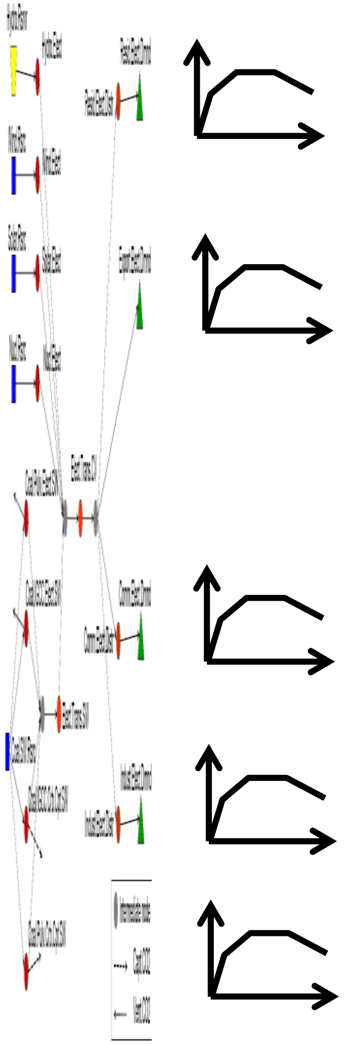


- Imagine 25 large storage devices spread around the PJM grid:

- Imagine 25 large storage devices spread around the PJM grid:



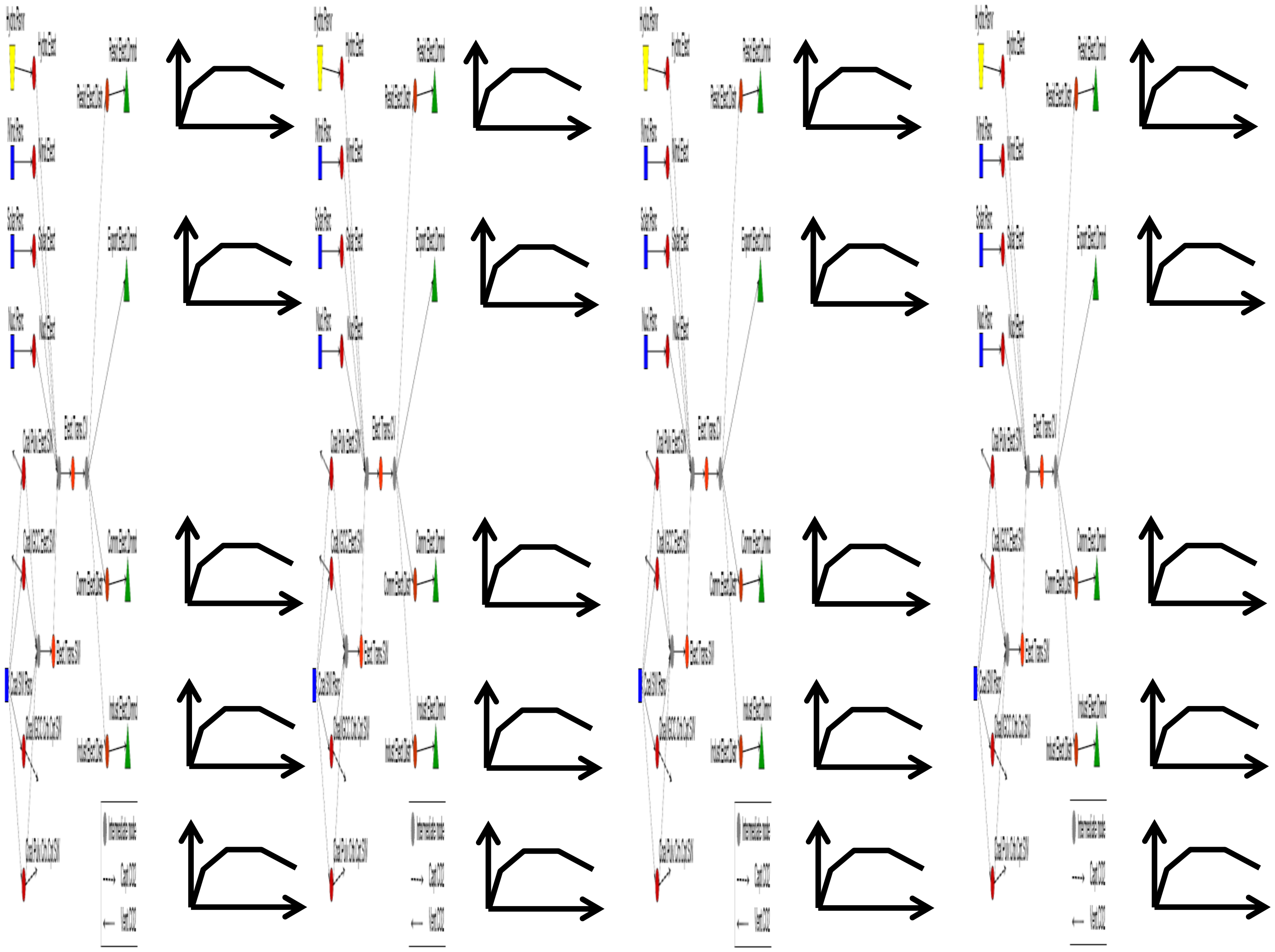
The value of grid level storage



The value of grid level storage

Monday

Time **:05** **:10** **:15** **:20**

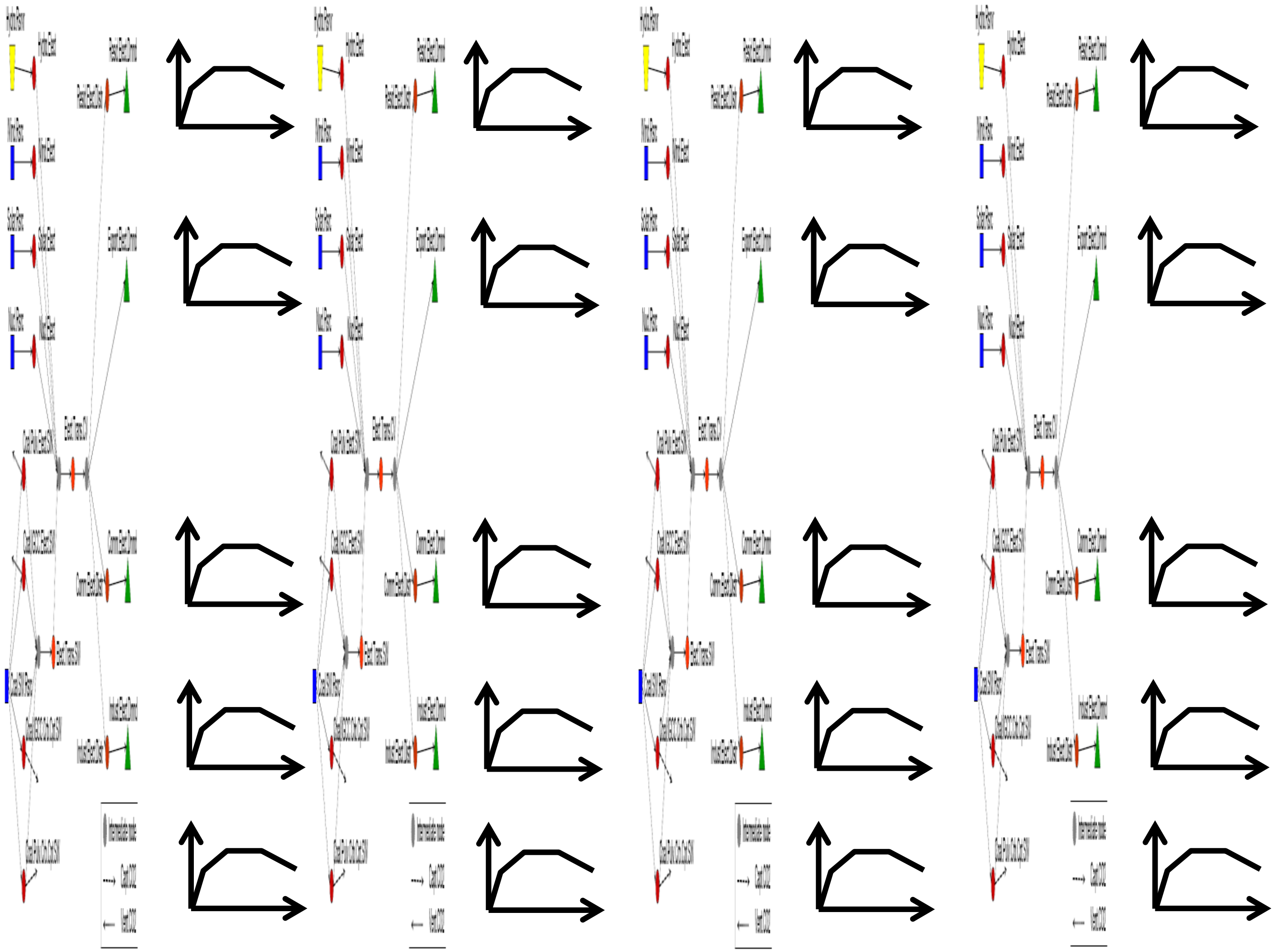


• • •

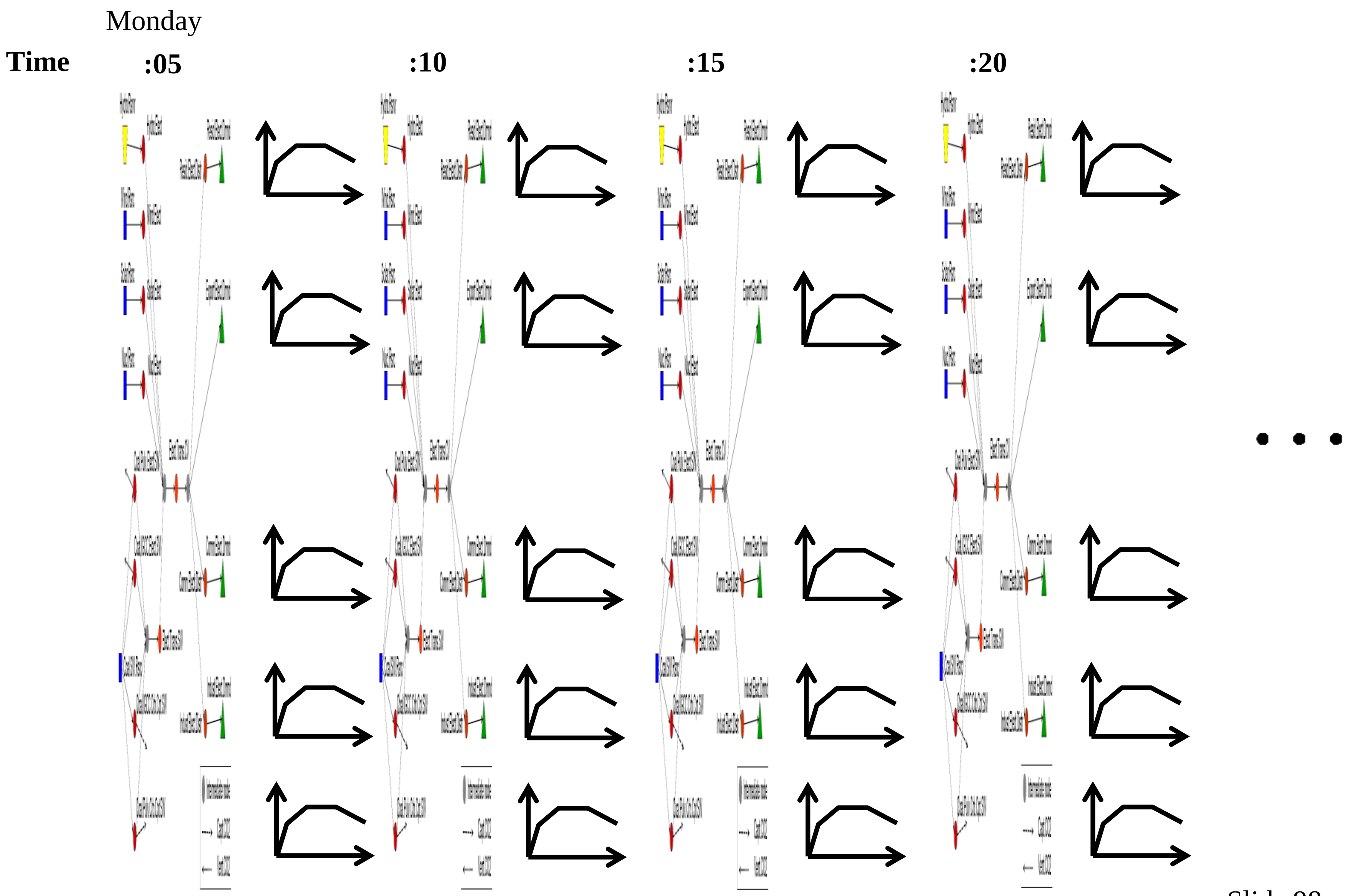
The value of grid level storage

Monday

Time **:05** **:10** **:15** **:20**

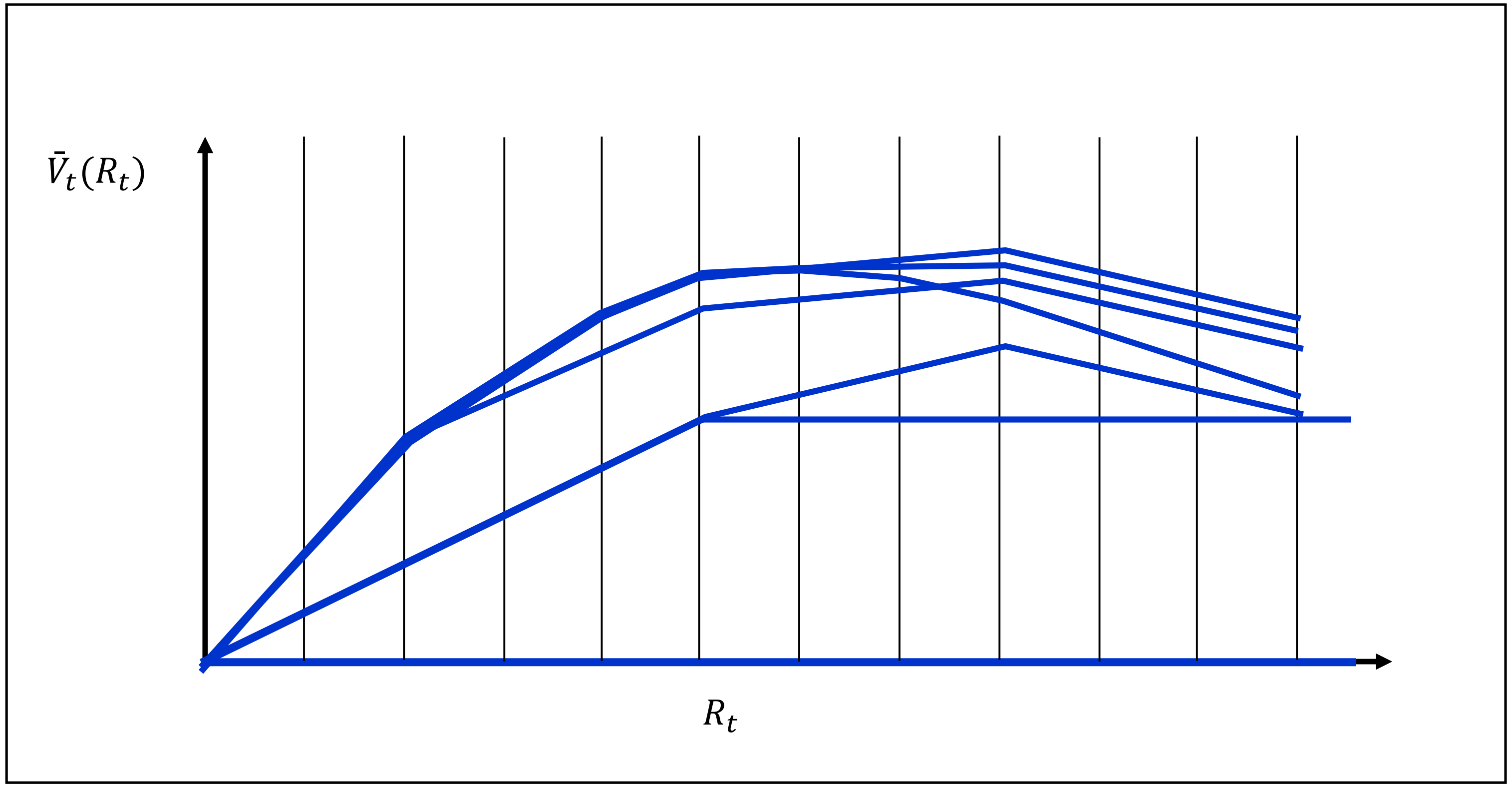


The value of grid level storage



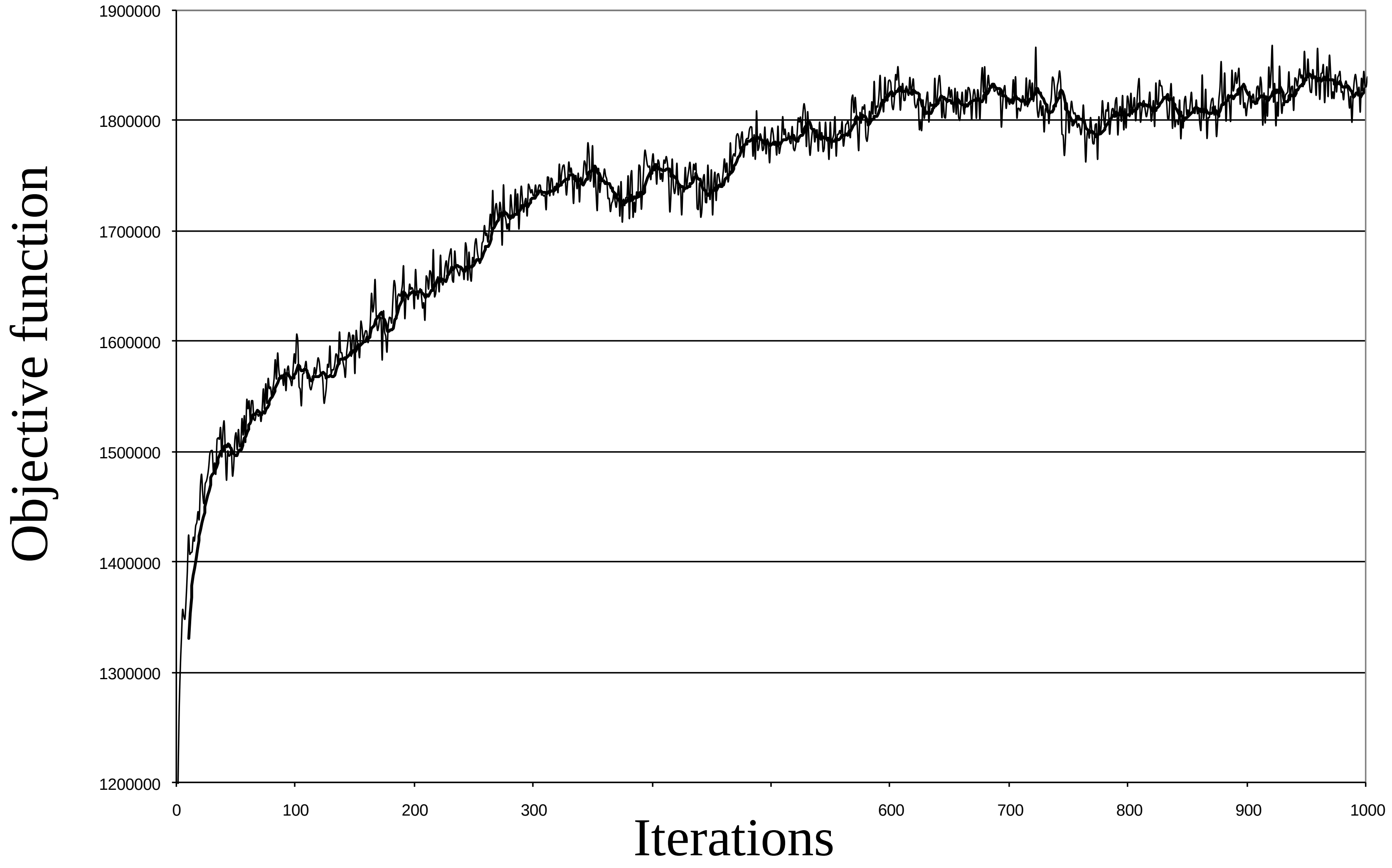
Approximate dynamic programming for energy storage

- Derivatives are used to estimate a piecewise linear approximation



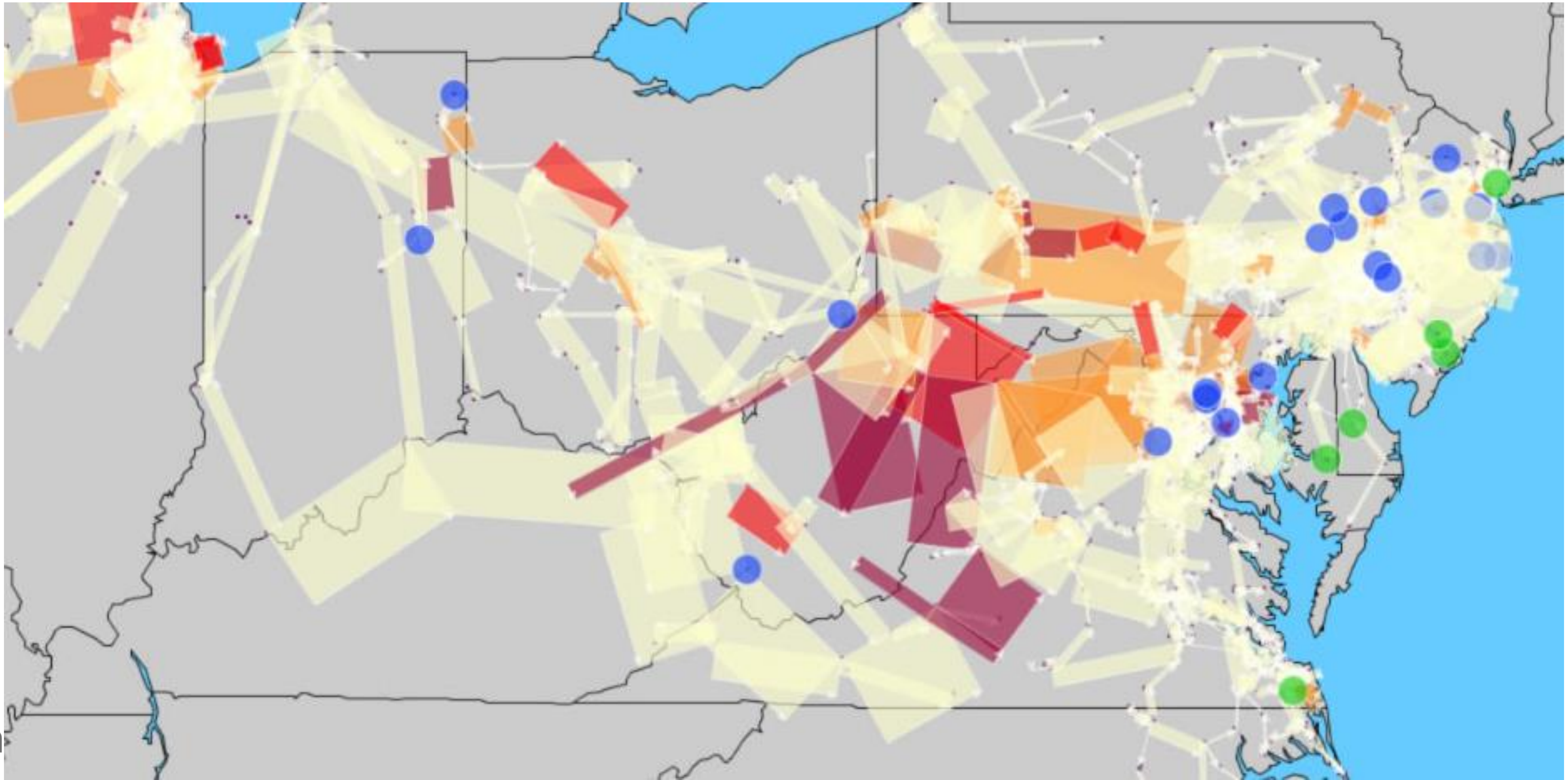
Approximate dynamic programming for energy storage

- With luck, your objective function improves



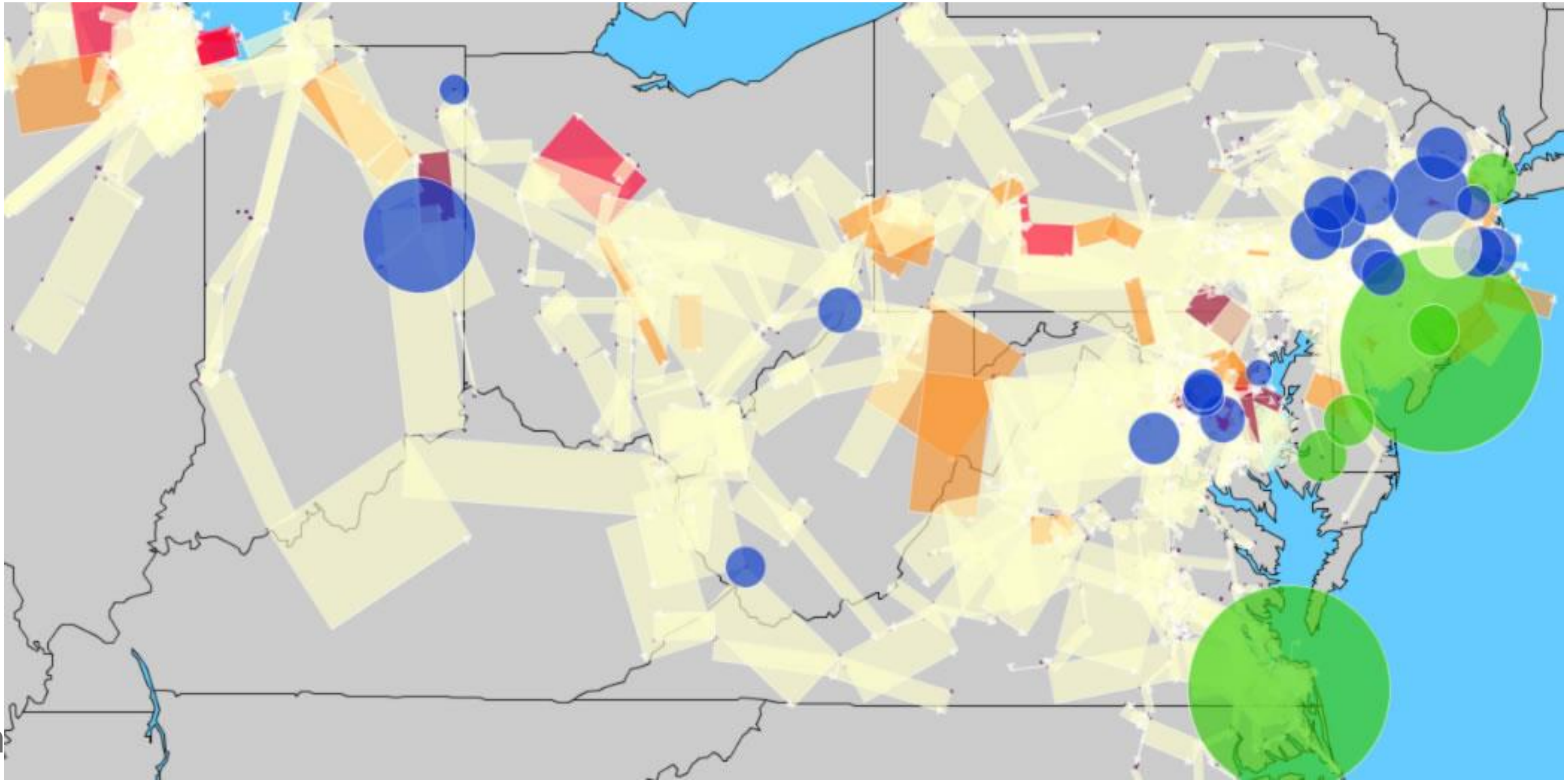
Approximate dynamic programming for energy storage

- Congested grid:
  - » Green and blue circles indicate energy storage



Approximate dynamic programming for energy storage

- Congested grid:
  - Green and blue circles indicate energy storage

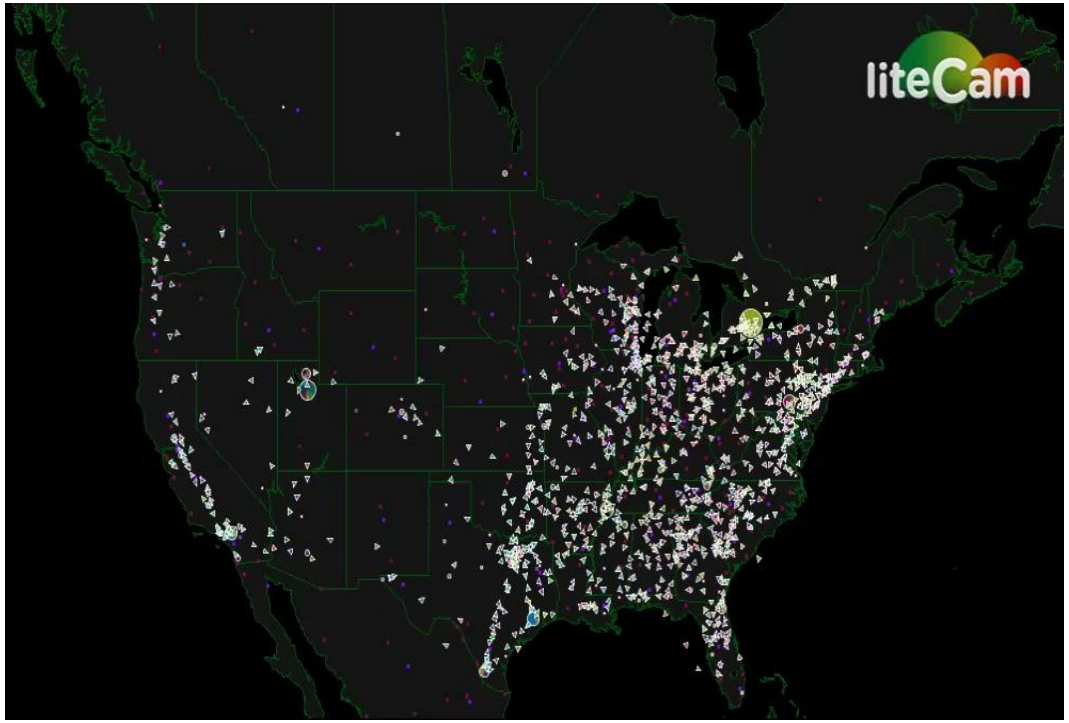


NOTABLE APPLICATIONS

Reinforcement Learning

» **Major achievement** — Playing Go

High-Dimensional ADP



» **Major achievement** — Optimizing a 5000 truck fleet

DESIGNING POLICIES

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

$$(\tilde{S}_{tt}, \tilde{x}_{tt}, \widetilde{W}_{t,t+1}, \tilde{S}_{t,t+1}, \tilde{x}_{t,t+1}, \widetilde{W}_{t,t+2}, \dots, \tilde{S}_{tt'}, \tilde{x}_{tt'}, \widetilde{W}_{t,t'+1}, \dots)$$

There are six classes of approximations we can introduce.

Our direct lookahead policy now requires solving:

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

Sampled information process

Restricted horizon

Limited decisions

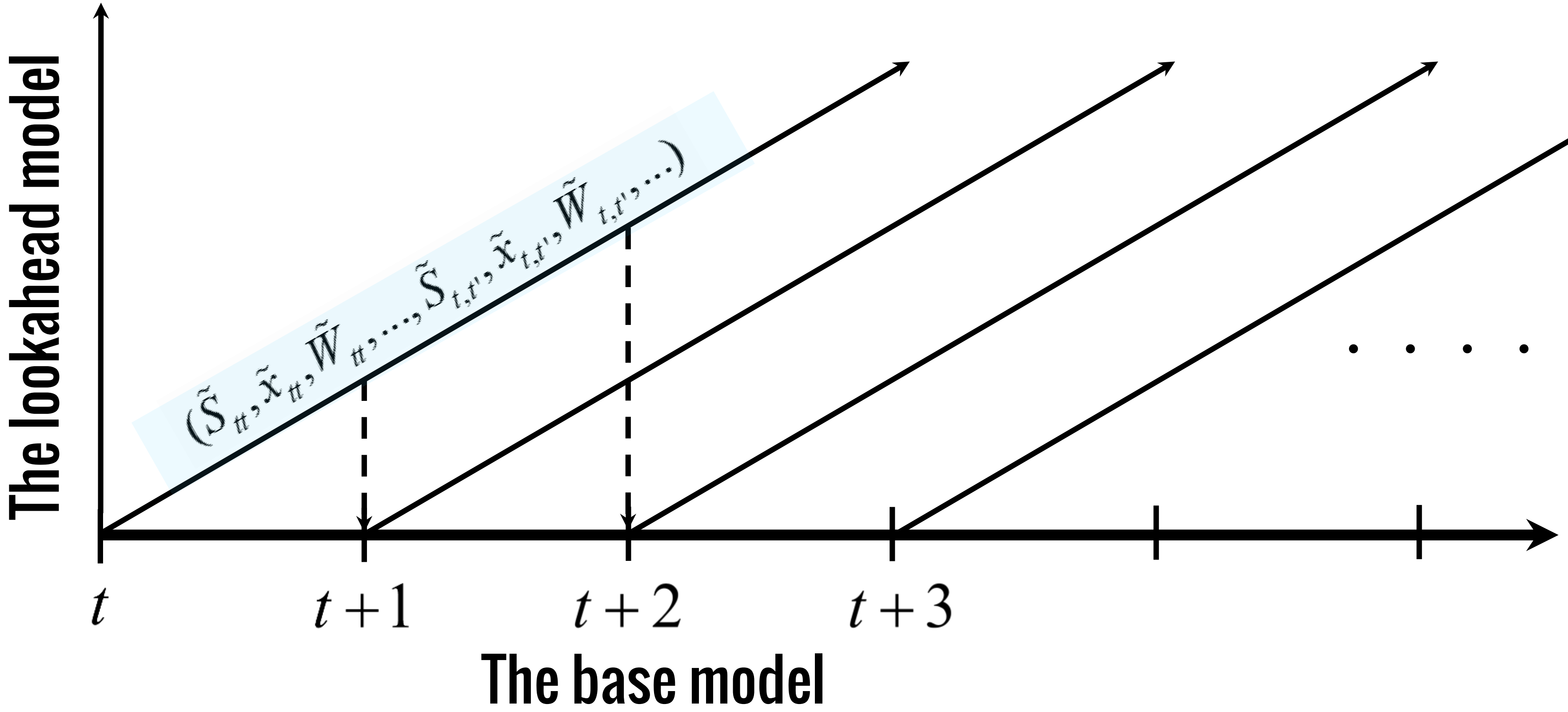
"Policy-within-a-policy"

Reduced state space

DESIGNING POLICIES

Direct Lookahead Policies (DLAs)

» Tilde variables are used to model approximate lookahead

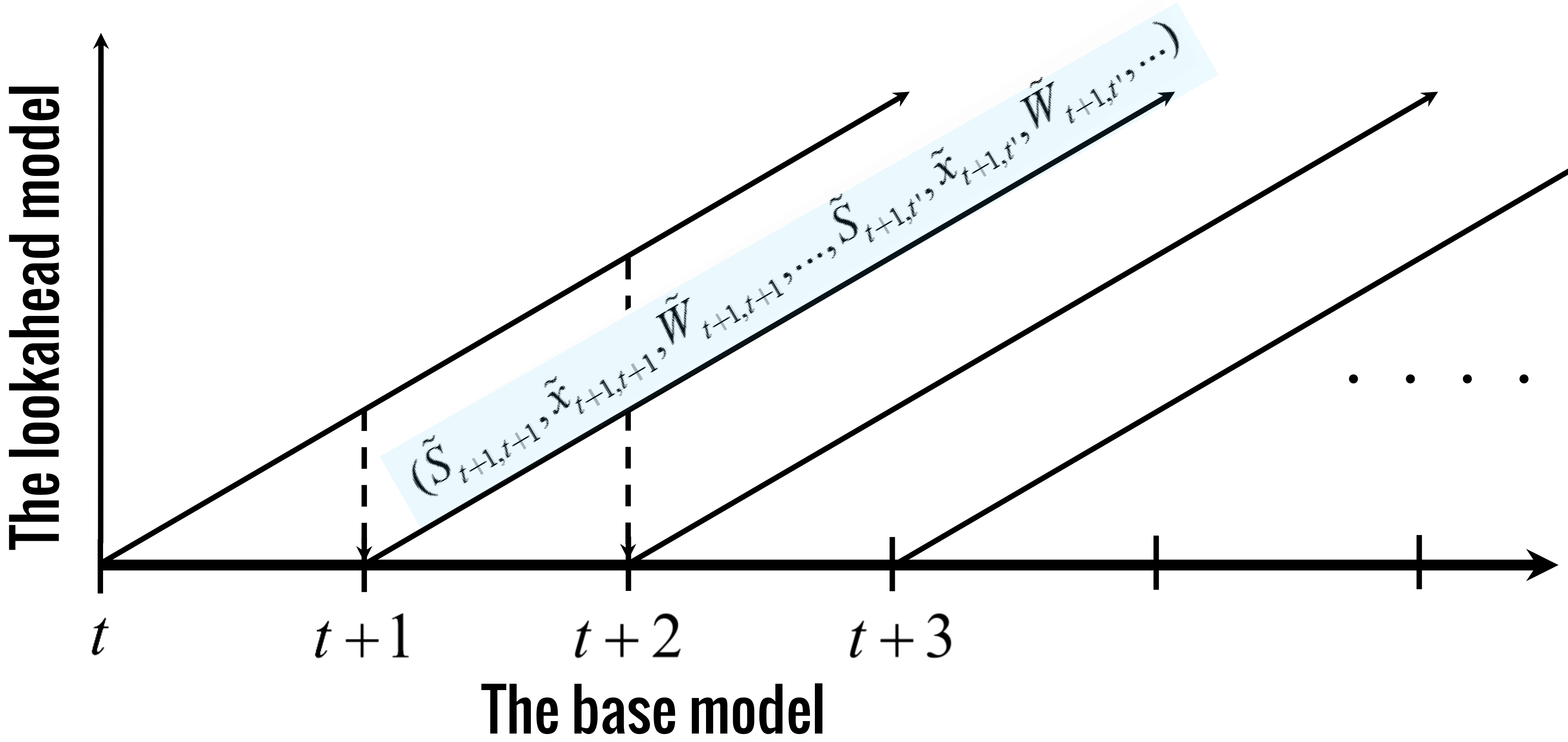


Optimal Dynamics

DESIGNING POLICIES

Direct Lookahead Policies (DLAs)

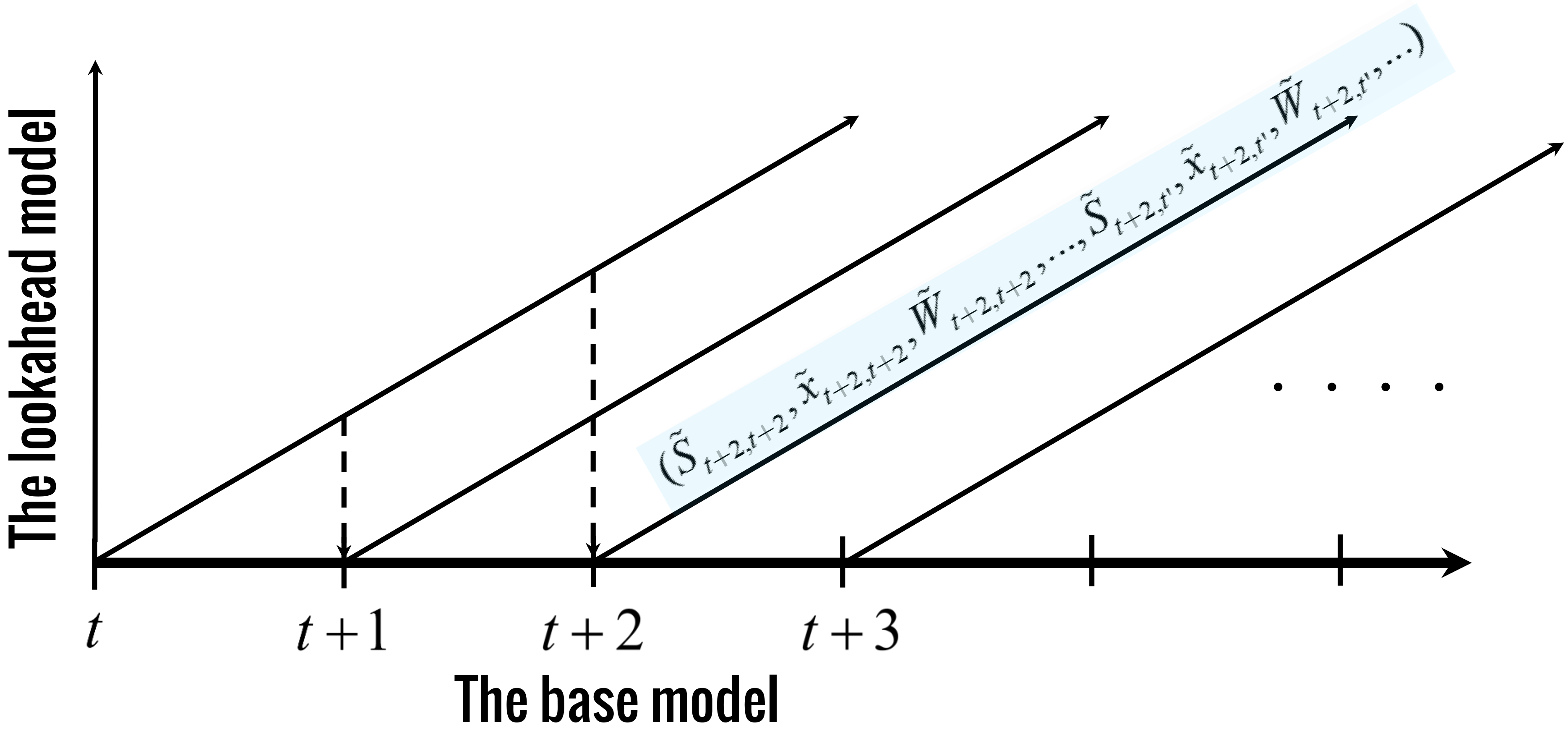
» Tilde variables are used to model approximate lookahead



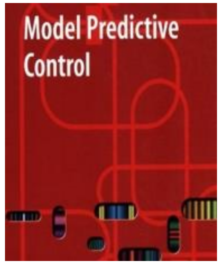
DESIGNING POLICIES

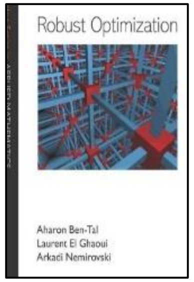
Direct Lookahead Policies (DLAs)

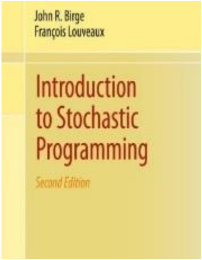
» Tilde variables are used to model approximate lookahead

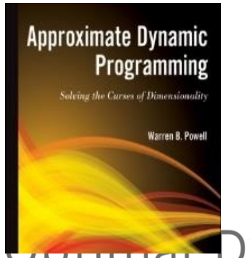


DESIGNING POLICIES









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"

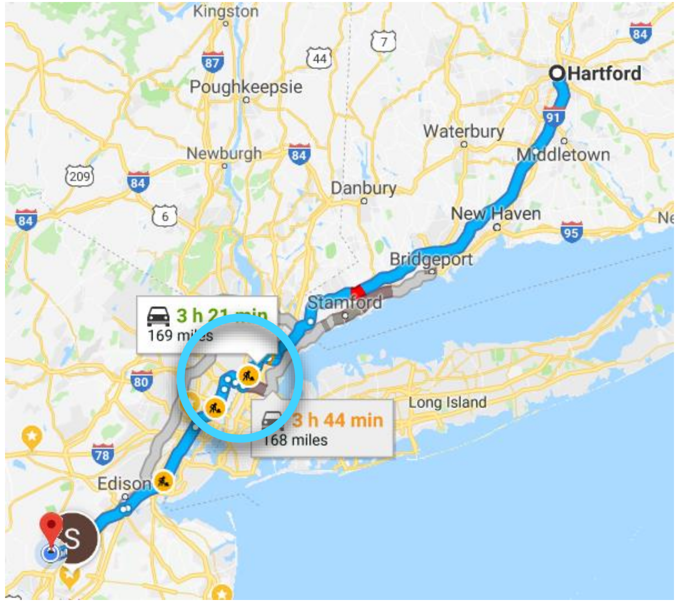
» **Robust optimization** - We could use the 90th percentile of travel times.

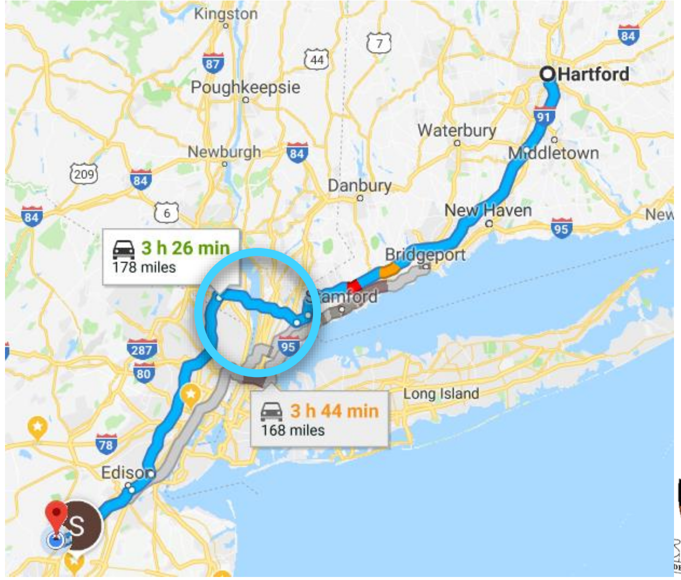
» **Stochastic programming** – We represent the future using, say, 20 samples.

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

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

See Chapter 19 at http://tinyurl.com/RLandSO

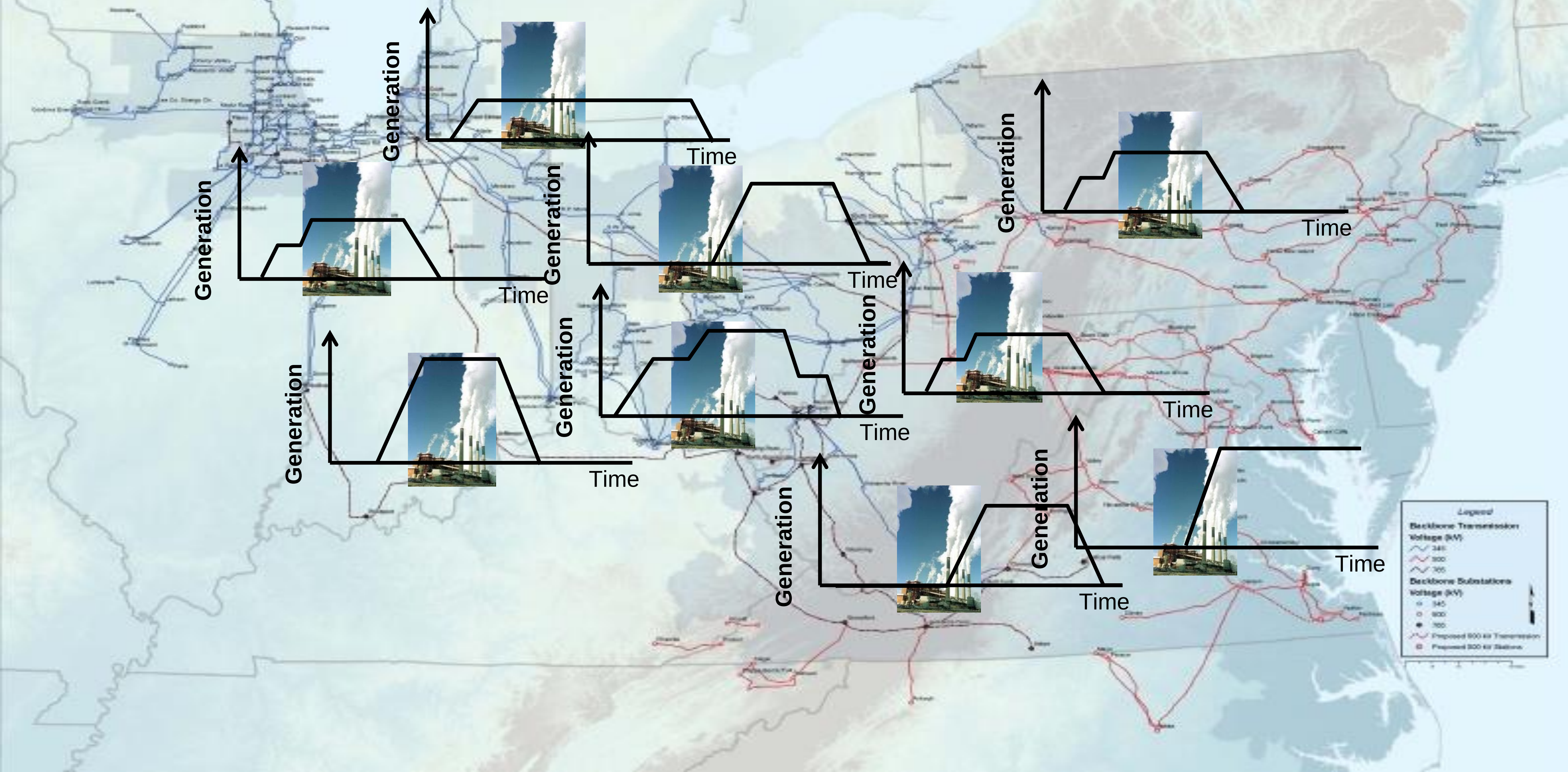




Optimal Dynamics

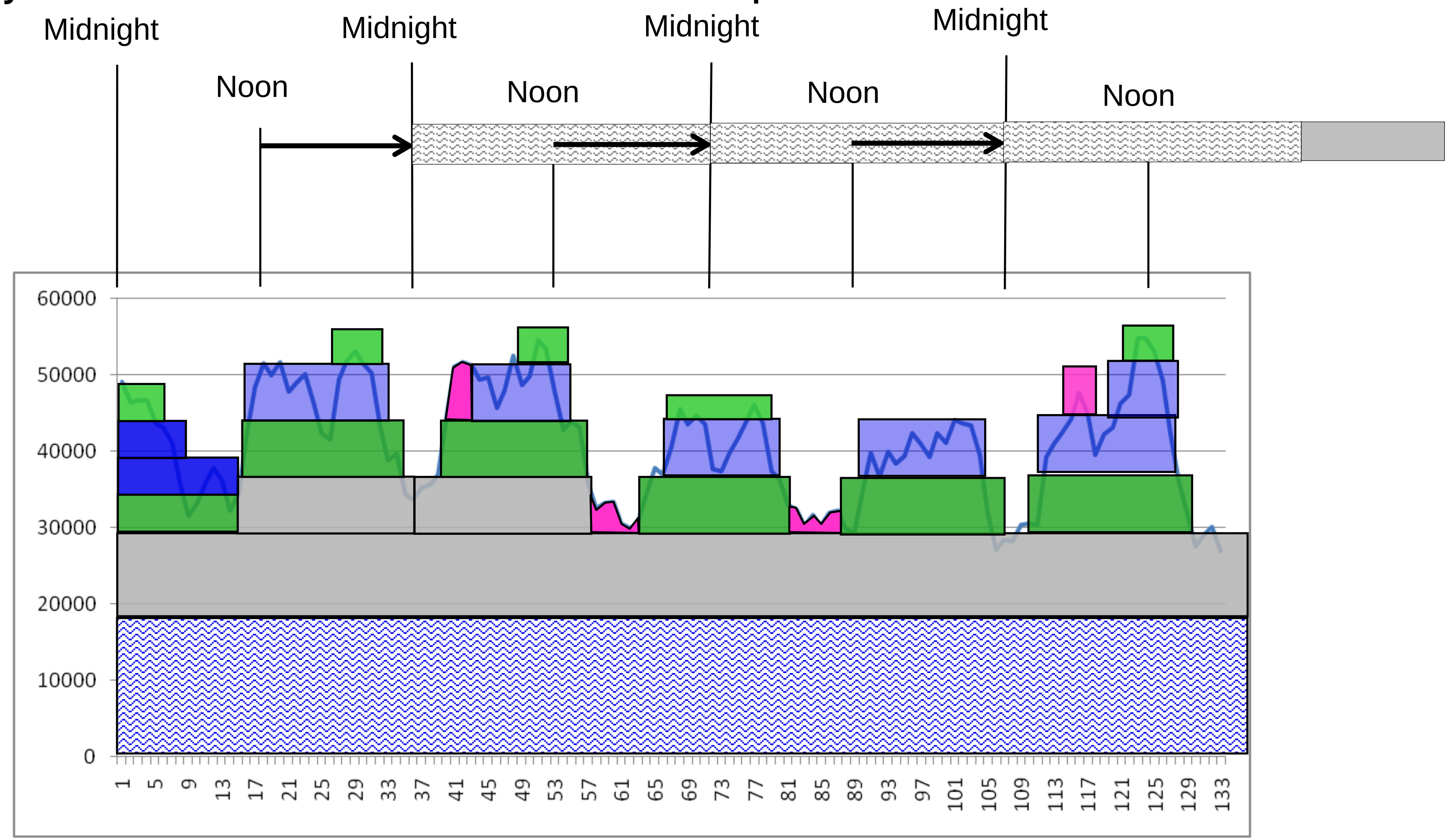
The unit commitment problem (for PJM)
Planning tomorrow's schedule

The unit commitment problem (for PJM)
Planning tomorrow's schedule



The timing of decisions

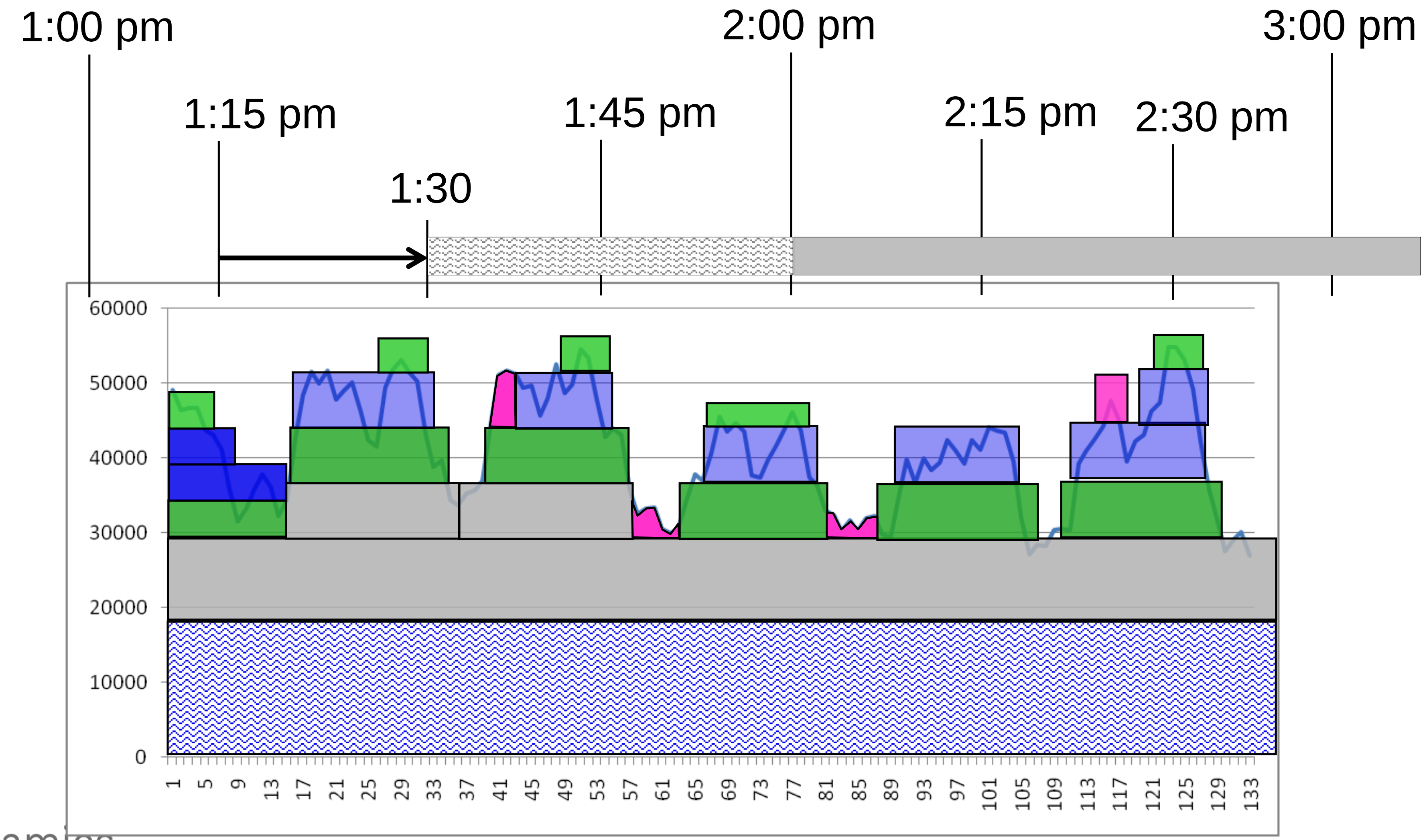
- The day-ahead unit commitment problem



Optimal Dynamics

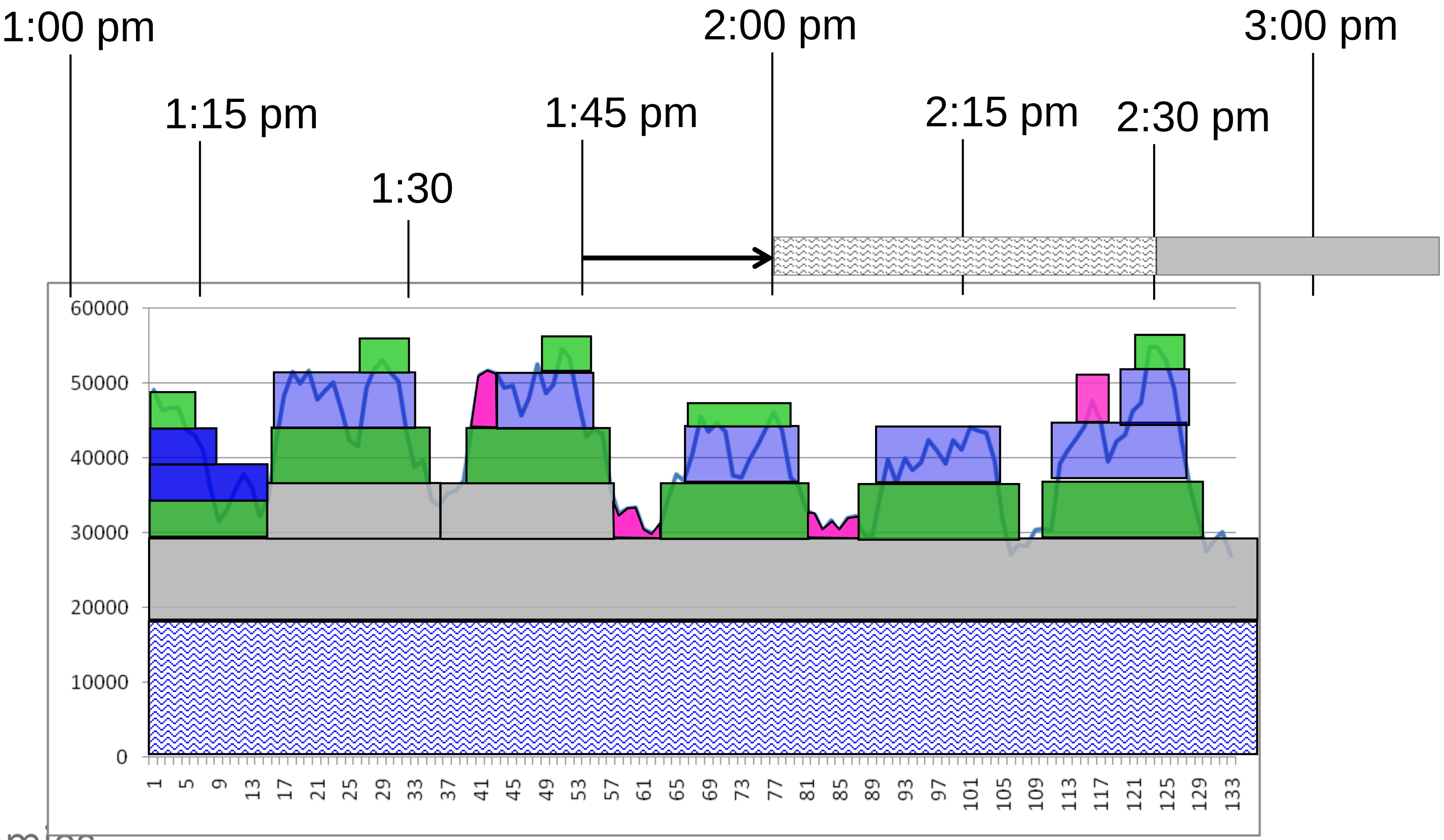
The timing of decisions

- Intermediate-term unit commitment problem

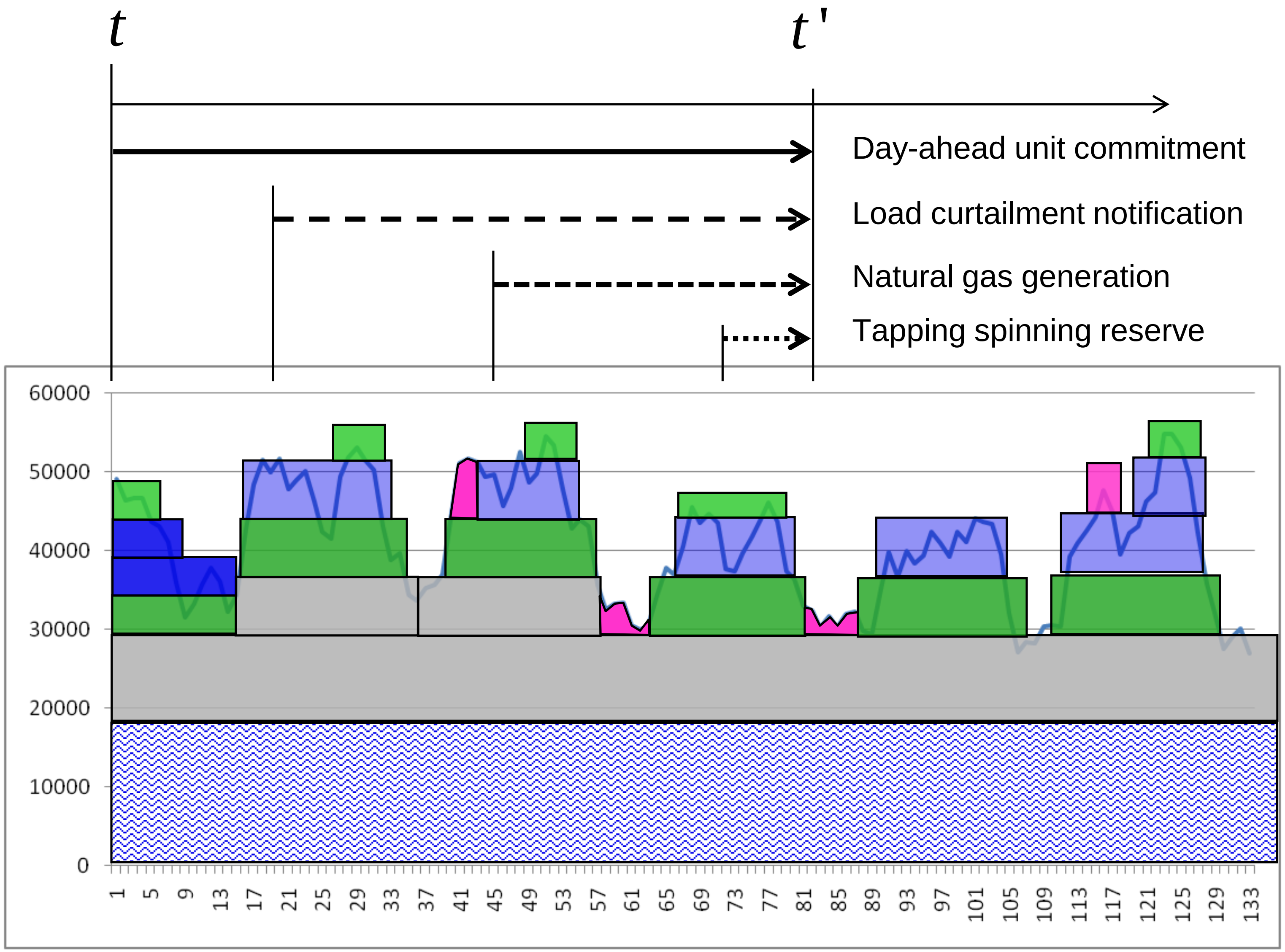


The timing of decisions

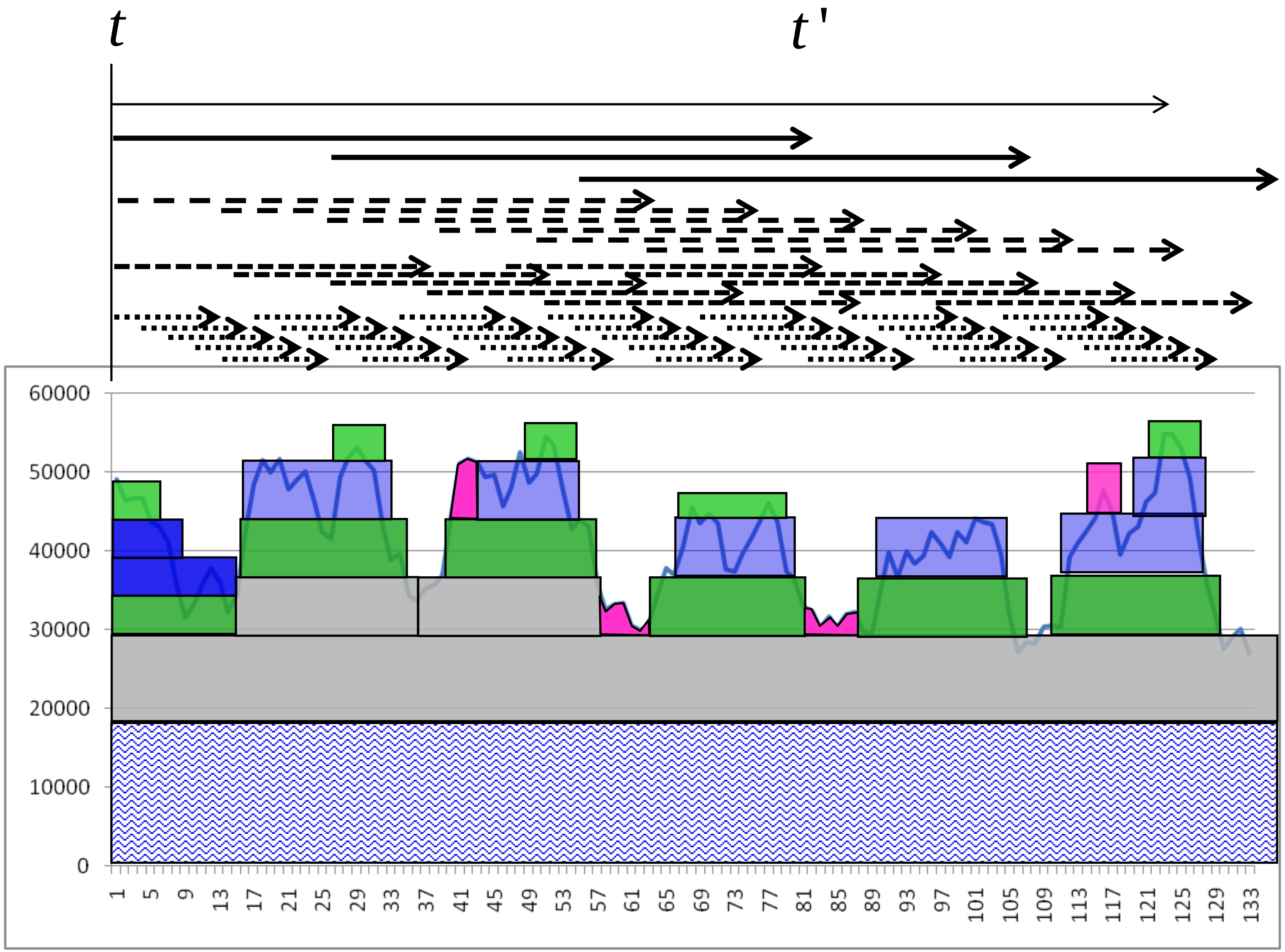
- Intermediate-term unit commitment problem



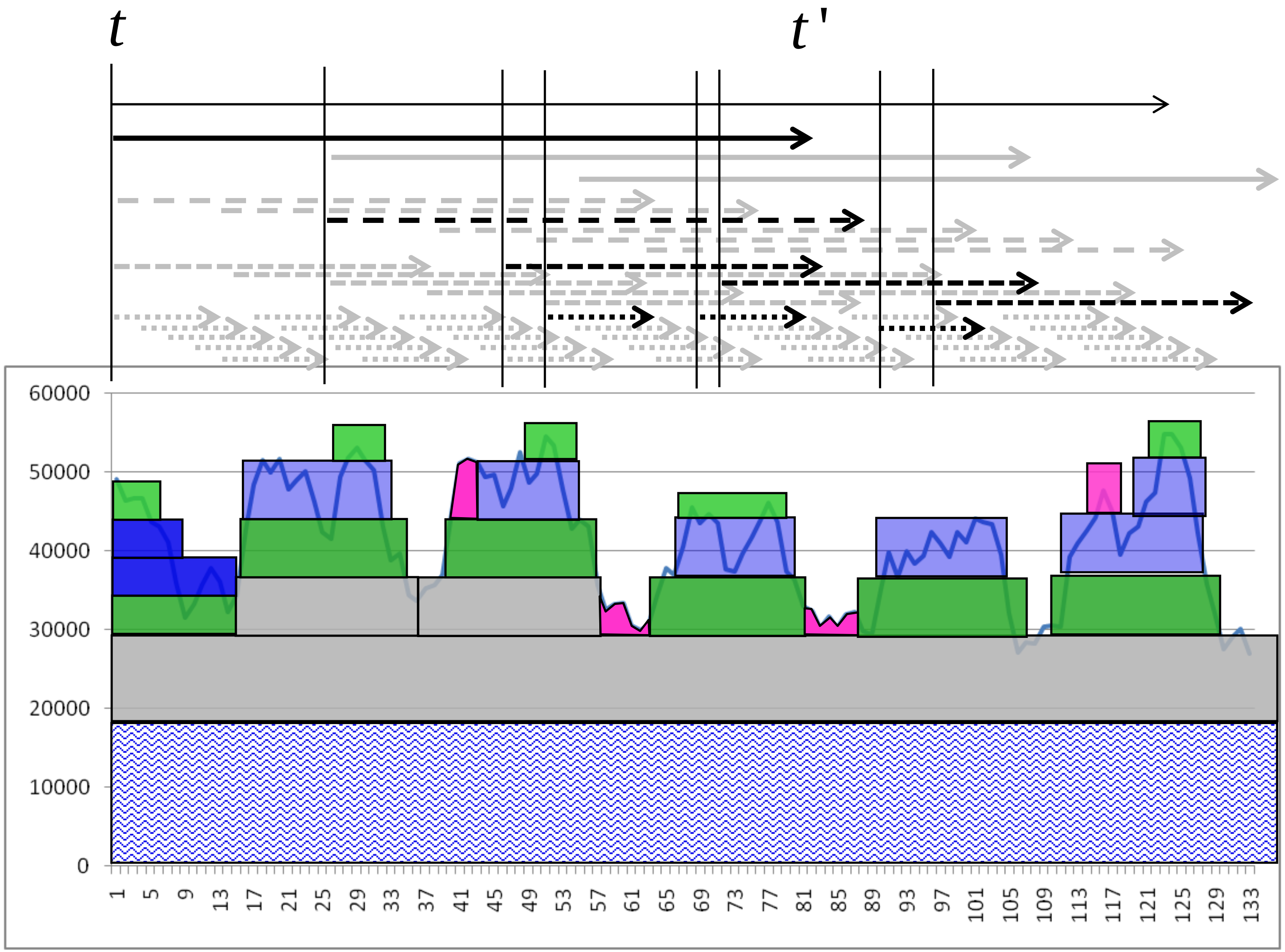
The unit commitment problem



The unit commitment problem



The unit commitment problem



DESIGNING POLICIES

Policy search policies

Policy function approximations (PFAs)

» Simple rules, functions

» Examples:

- Order up to
- Buy low, sell high

Cost function approximations (CFAs)

» Parameterized cost models

» Examples

- Schedule slack for trips
- Buffer stocks for inventory

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:

- Google maps
- Energy planning models

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.

But are you doing the best you can?

THE UNIVERSAL FRAMEWORK FOR SEQUENTIAL DECISIONS

Warren B. Powell, Princeton University

http://tinyurl.com/powelljungle

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

$$where \quad S_{t+1} = S^M(S_t, X^{\pi}(S_t), W_{t+1})$$

$$and\ given\ (S_0, W_1, W_2, \ldots, W_t, \ldots)$$

Policy

Policy Search

Lookahead Approximations

Policy Function Approximation(PFA)

$$X^{PFA}(S_t|\theta) = \begin{cases} If\ this\ then\ do \\ \sum_{f \in F} \theta_f \phi_f(S_t) \\ Neural\ network \end{cases}$$

Cost Function Approximation (CFA)

$$X^{CFA}(S_t|\theta) = \begin{cases} argmax_x c_t x_t + \sum_f \theta_f \phi_f(S_t) \\ argmax_x(\mu_{tx} + \theta^{IE} \bar{\sigma}_{tx}) \end{cases}$$

Value Function Approximation (VFA)

$$\begin{aligned} X^{VFA}(S_t|\theta) &= argmax_x(C(S_t, x) + \mathbb{E}\{V_{t+1}(S_{t+1})|S_t, x_t\}) \\ &= argmax_x(C(S_t, x) + \bar{V}_t^x(S_t^x)) \\ &= argmax\ Q(S_t, x) \end{aligned}$$

Direct Lookahead (DLA)

$$X^{DLA}(S_t|\theta) = argmax_x\left(c_t x_t + \sum_{t'=t+1}^{t+H} \tilde{c}_{tt'} \tilde{x}_{tt'}\right)$$

The four classes of policies (PFAs, CFAs, VFAs and DLAs) are *universal*. Any sequential decision problem will use one of these four classes (or a hybrid), including whatever you might be doing now.

The optimal policy (if we could solve it) is given by

$$X^*(S_t) = argmax_x\left(C(S_t, x) + \mathbb{E}\left\{\max_{\pi} \mathbb{E}\left\{\sum_{t'=t+1}^{t+H} C(S_{t'}, X^{\pi}(S_{t'}))\,|S_{t+1}\right\}|S_t, x_t\right\}\right)$$

DESIGNING POLICIES

- Policies in the "policy search" class are simplest
  - As a result, this is what you are most likely going to see used in practice.

    … but …

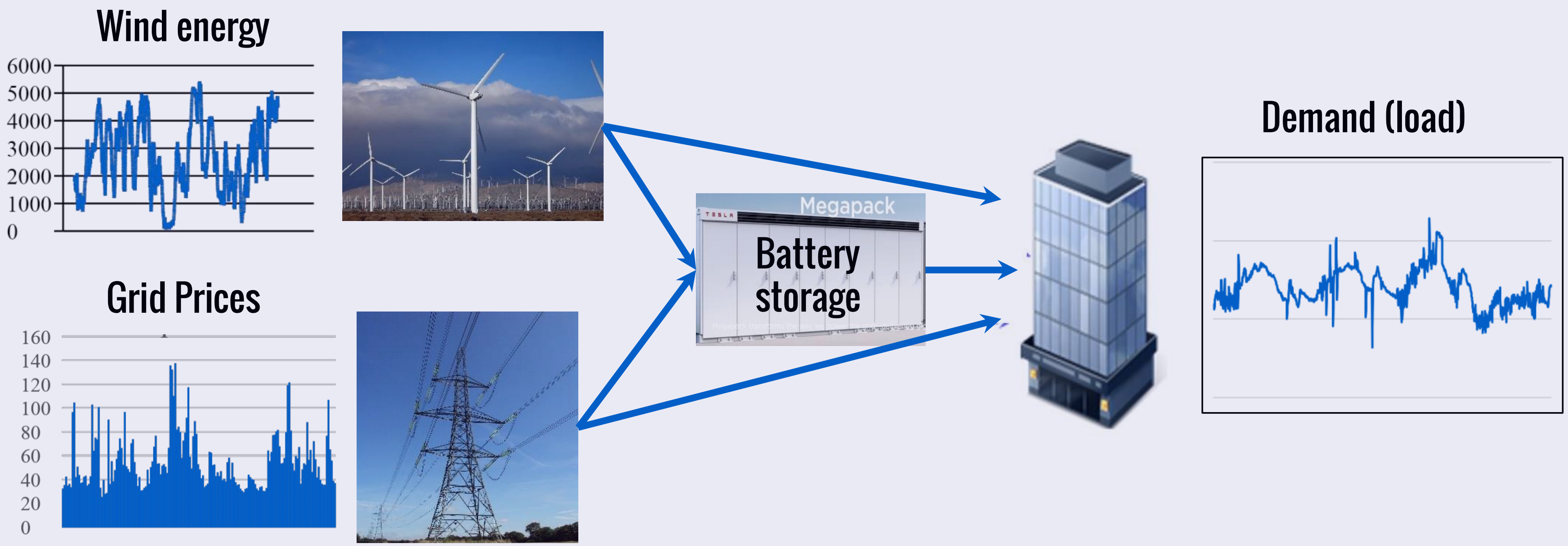
  - "The price of simplicity is tunable parameters"

    … and …

  - "Tuning is hard"!

AN ENERGY STORAGE PROBLEM

Consider a basic energy storage problem



We are going to show that with minor variations in the characteristics of this problem, we can make *each* class of policy work best.

AN ENERGY STORAGE PROBLEM

Each policy is best on certain problems

Problem:	Problem description	PFA	CFA Error correction	VFA	Determ. Lookahead	CFA Lookahead
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 Muenster, Germany.

» … any policy might be best depending on the data.

Optimal Dynamics

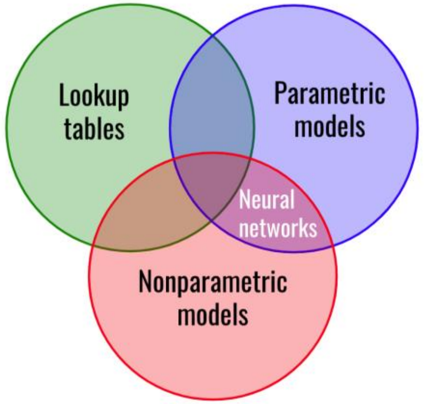
BRIDGING MACHINE LEARNING & SEQUENTIAL DECISIONS

Machine learning

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

Searching over functions

"Big dataset"



Sequential decisions

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

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

Searching over policies

System model

Policy function approximations	Analytical functions
Cost function approximations	Optimization problem
Value function approximations	Optimization problem
Direct lookahead approximations	Optimization problem

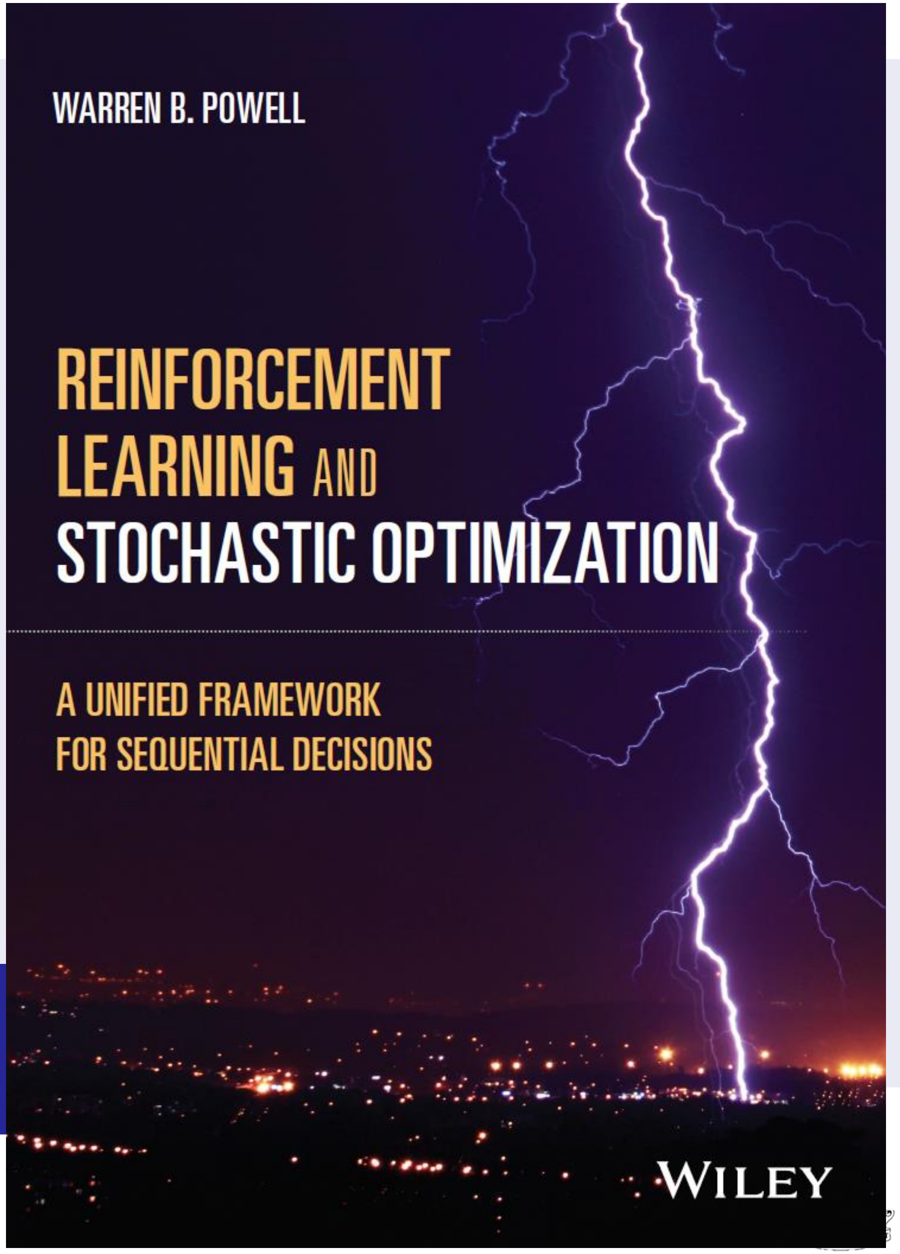
Optimal Dynamics

Choosing a policy class

A new book:

- 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.*
- To appear May, 2022.

http://tinyurl.com/RLandSO/

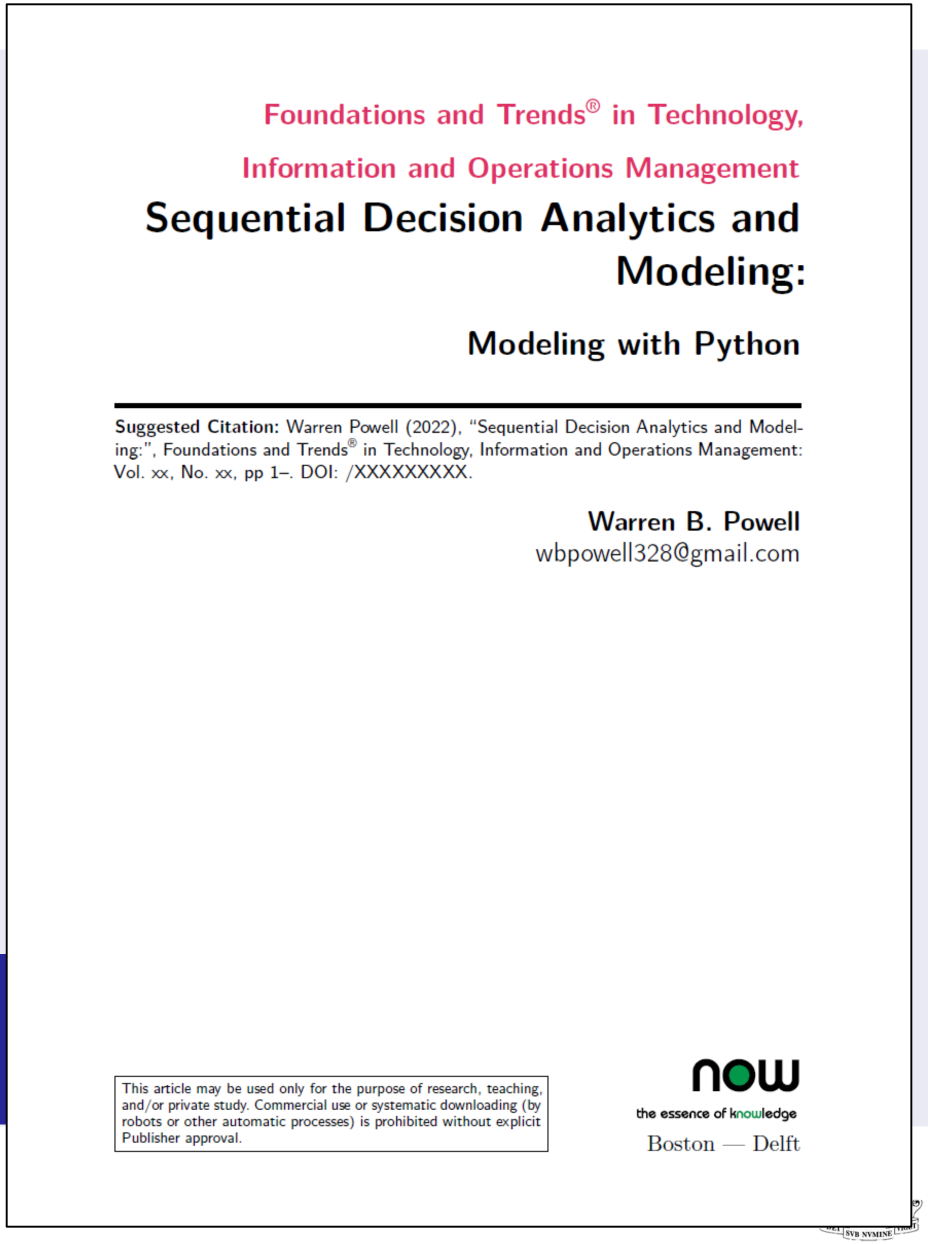


Choosing a policy class

An introductory book:

- 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

http://tinyurl.com/sdaexamplesprint



Foundations and Trends® in Technology, Information and Operations Management

Sequential Decision Analytics and Modeling:

Modeling with Python

Suggested Citation: Warren Powell (2022), "Sequential Decision Analytics and Modeling:", Foundations and Trends® in Technology, Information and Operations Management: Vol. xx, No. xx, pp 1–. DOI: /XXXXXXXXX.

Warren B. Powell
wbpowell328@gmail.com

This article may be used only for the purpose of research, teaching, and/or private study. Commercial use or systematic downloading (by robots or other automatic processes) is prohibited without explicit Publisher approval.

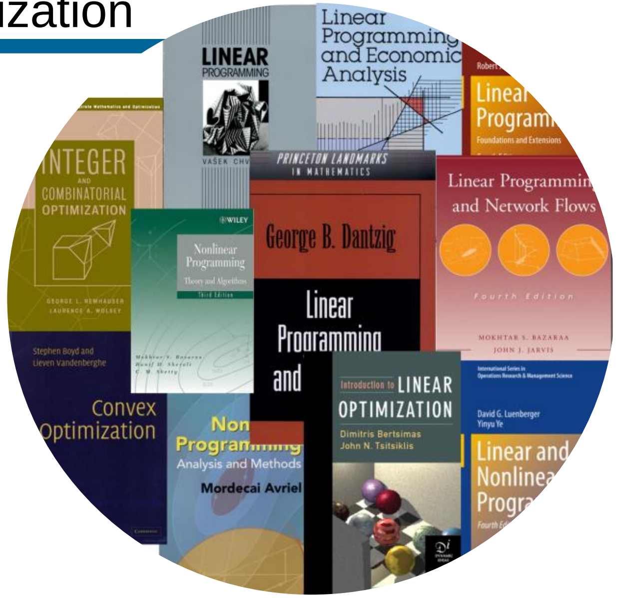
now
the essence of knowledge
Boston — Delft

OUTLINE

- → The five layers of intelligence
- → Modeling sequential decision problems
- → Designing policies
  - → Policy function approximations
  - → Cost function approximations
  - → Value function approximations
  - → Direct lookahead approximations
- → A new educational field: sequential decision analytics

The core disciplines of decision analytics

Optimization



Machine learning

Each of these fields have well-defined communities, using common notation and established tools.

There are widely used textbooks that cover common material, with standard notational frameworks..

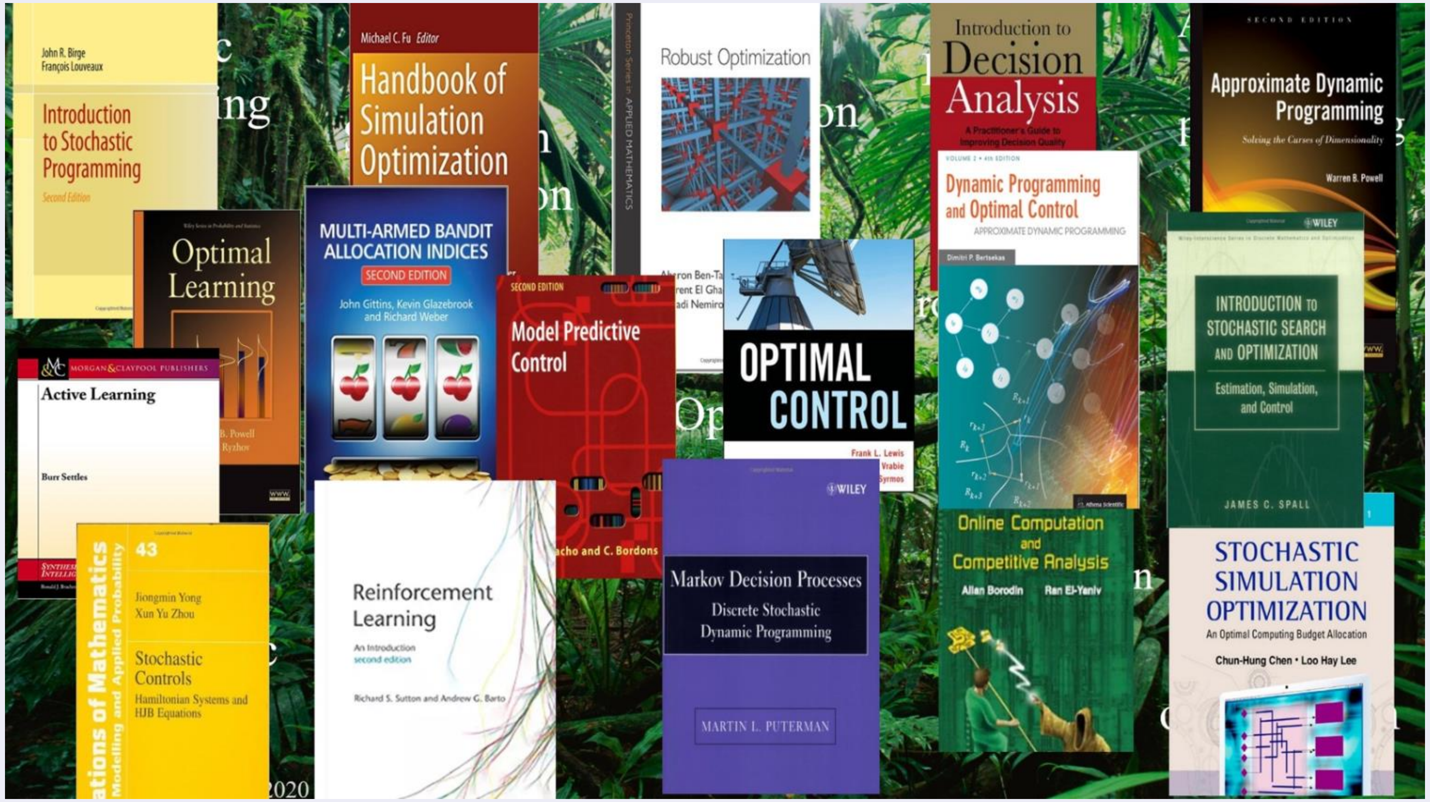
The concepts are taught in hundreds of academic programs, producing thousands of graduates each year which can be hired by industry.

Simulation

OPTIMIZATION UNDER UNCERTAINTY

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 narrowly defined problem classes.
- Real applications require skills that span a wide range of problem settings.



Optimal Dynamics

OPTIMIZATION UNDER UNCERTAINTY

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.

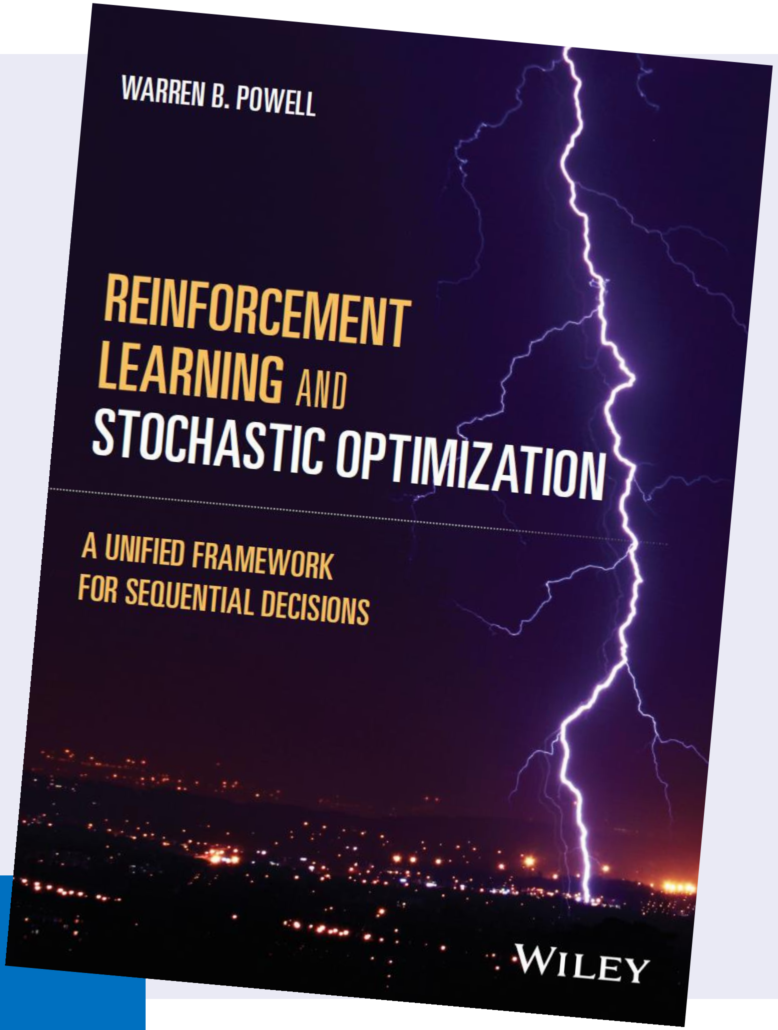
These will be the first books to present sequential decision problems and solution methods in a unified way.

SEQUENTIAL DECISION ANALYTICS

An academic proposal

- We need to establish academic programs in engineering focusing broadly on sequential decision analytics, comparable to existing programs in machine learning.

- As with machine learning, this program could be centered on methodology, or different problem domains that span engineering (all fields), the sciences (all fields), business, finance, logistics, energy and health.



FOR MORE INFORMATION, VISIT
http://tinyurl.com/sdafield

Optimal Dynamics

Thank you!

See

http://tinyurl.com/sdafield

for an introduction to a field I am calling

"Sequential Decision Analytics"

My new book is available at

http://tinyurl.com/RLandSO/