Optimal Control in Stochastic Systems: Challenges for the Future

Workshop YEQT IV

Chelsea C. White III
Schneider National Chair of Transportation & Logistics
Georgia Institute of Technology

On leave of absence to Khalifa University, Abu Dhabi

Eindhoven, 25 November 2010

Outline

- Motivation & context
- Models of sequential decision making – MDP
- Applications of the standard MDP
    - Inventory replenishment under supply uncertainty
    - Inventory & supply systems
- Extensions of the standard MDP
    - Partially observed MDP
    - In-transit perishable product inspection
    - Value of traffic information for freight transport
- Sequential games, supply chain design, & risk mitigation (future research directions)
- Q&A

Outline

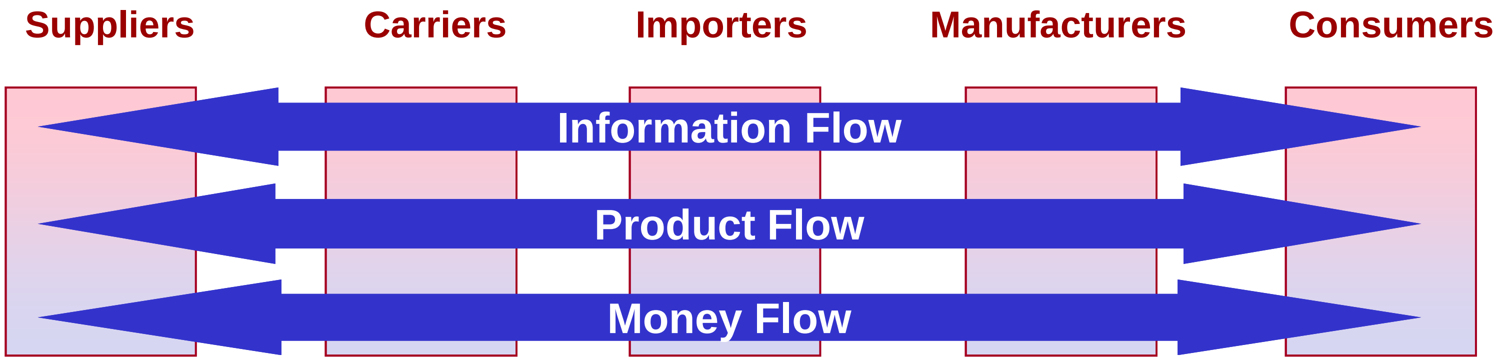
- Motivation & context
- Models of sequential decision making – MDP
- Applications of the standard MDP
  - Inventory replenishment under supply uncertainty
  - Inventory & supply systems
- Extensions of the standard MDP
  - Partially observed MDP
  - In-transit perishable product inspection
  - Value of traffic information for freight transport
- Sequential games, supply chain design, & risk mitigation (future research directions)
- Q&A

Sequential decision making

- Sequential decision making (& data collection)
  - Take an action, based on what we currently know
  - Observe the consequences of the action
  - Take the next action, based on what we know and what we've just learned
- Actions taken are intended to achieve an objective(s)

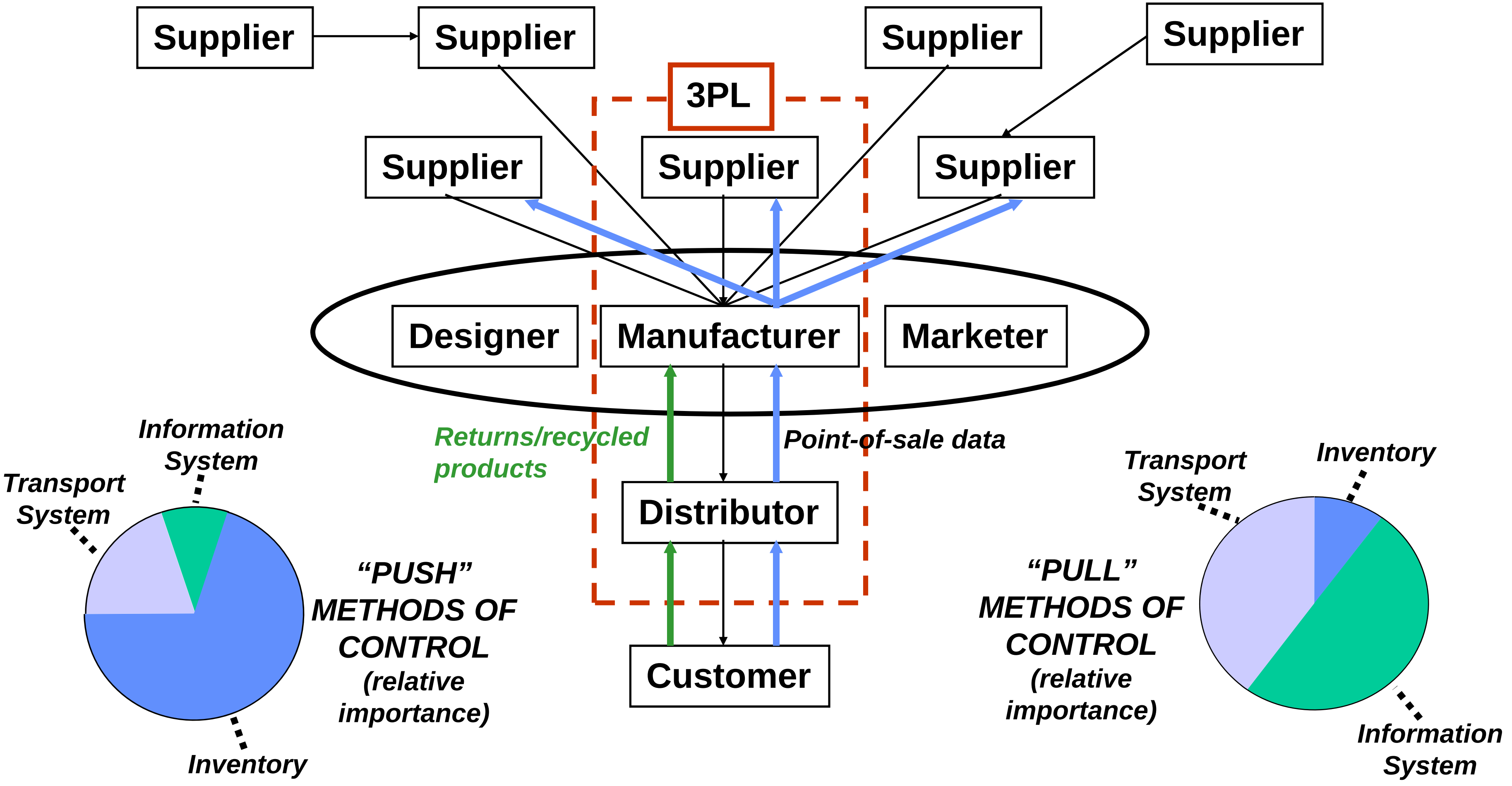
The extended enterprise

- The extended enterprise – network of independent companies with intent to respond to customers with better, less expensive products and technologies, faster to market. *Better, cheaper, faster*
- Supply chain – the flows of goods, information, and money in this network (definitions vary)



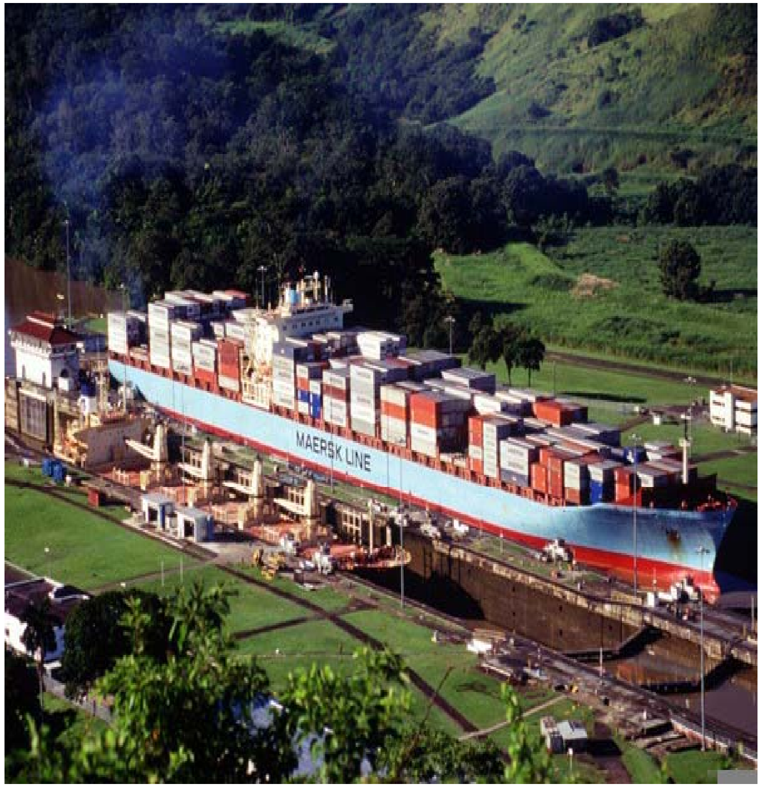
Source: Bearing Point

Supply chain systems



Supply chain control

- Criteria options: efficiency & profit, environmental impact, resiliency, risk (various types, e.g., enterprise, legal, operational), various combinations. Implication of multi-objective approach involving additive and multiplicative criteria.
- Explicit consideration of uncertainty:
  - at demand/supply level
  - at network structure level (assoc w/ major disruptions, resiliency).
- Congestion (on roads, at ports, terminals, railheads, etc.) increases both lead-time mean and variance.



Bottlenecks & delays can occur:

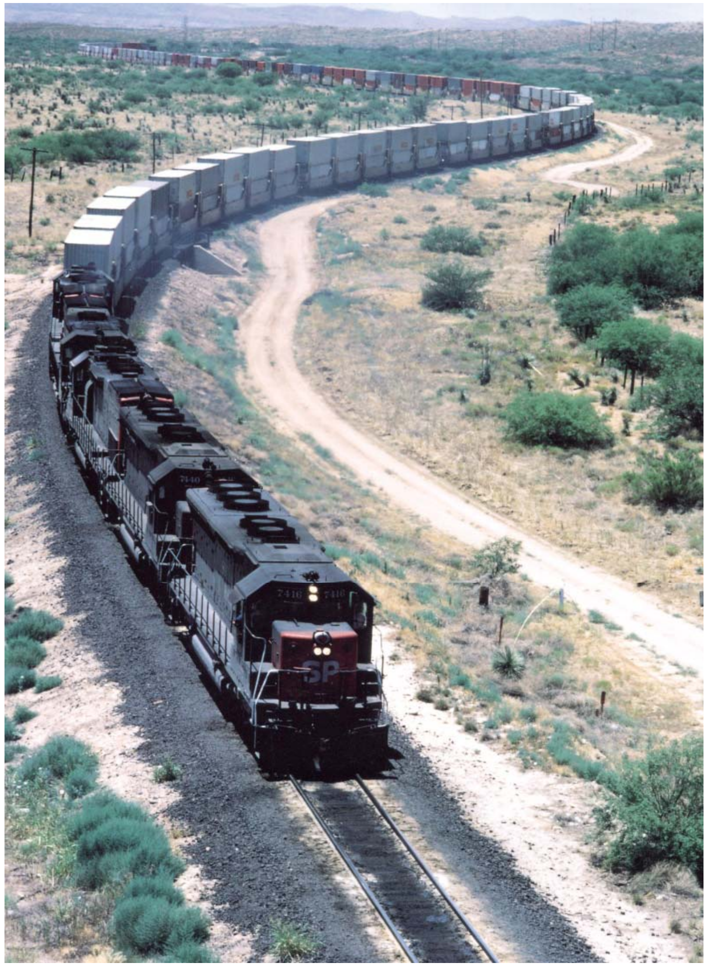
- In production
- In transit
- At canals
- At truck terminals, sea & air ports, rail heads

Real time information

- Real time control of supply chains, based on real time data.
- IT issues – noise corrupted sensor output, data transmission. Transmission delays. Sensor placement
- *Information pattern* – who knows what & when? Centralized & decentralized control. Multi-agent systems, each agent with a different information pattern. What is best information pattern design, given cost of communication? Incentive alignment so decentralized system can behave ‘almost’ like a centralized system
- Partnership trust & information sharing. Can ‘trust’ be quantified?
- *Value of information* – its role in strategic decisions, e.g., IT systems purchase. How to extract the value of information operationally?
- When is the value of information negative?
- Temperature sensitive supply chains – more likely to involve IT?

Where do the data come from?

- Inventory levels
- Production rates
- Vehicle, vessel, or trailer
  - Position
  - Speed
  - Direction
  - Temperature
  - Oil or air pressure
- Driver alertness
- Traffic congestion sensors
- Weather
- Freight status & visibility (RFID)



Outline

- Motivation & context
- Models of sequential decision making – MDP
- Applications of the standard MDP
  - Inventory replenishment under supply uncertainty
  - Inventory & supply systems
- Extensions of the standard MDP
  - Partially observed MDP
  - In-transit perishable product inspection
  - Value of traffic information for freight transport
- Sequential games, supply chain design, & risk mitigation (future research directions)
- Q&A

Sequential decision making

- Models
  - Optimal control models – continuous state, linearity, difference/differential equation state dynamics of physical systems
  - Markov decision processes – discrete state & action, conditional probability for state dynamics, information/discrete systems
  - Decision analysis – discrete action & consequence, implicit notion of a state
  - Sequential games – non-cooperative games with partial information about the other agents' information patterns & objectives
  - Etc.

MDP problem definition

- Discrete time t, state s(t), action a(t)
- *State dynamics*: P[s(t+1) | s(t), a(t)]
- *Information pattern*: a(t) = d[s(t), t] for all t
- Cost per period: c[s(t), a(t)]
- *Criterion*: expected total discounted cost over the problem horizon (finite, random, infinite)
- Problem *objectives*:
  - Find a *policy* {d[s(t), t], t = 0, 1, …} that minimizes the criterion
  - Find the value of the criterion generated by an optimal policy

Applications

- Inventory control (Atan)
- Scheduling hospital admissions (Bekker)
- Production systems (Buyukkaramikli)
- Call centers (Gurvich, Spieksma)
- Traffic control (Haijema)
- Wireless communication systems (Hoekstra)
- Consumption-investment problems (Rieder)
- Transshipment (Wijk)

Solving the MDP

- The *optimality equation* (infinite horizon):

  $$v[s(t)] = \min \{c[s(t),a(t)] + \beta E\{v[s(t+1)]\}: a(t)\}$$

- Results:
  - There exists a unique solution to the OE, v*
  - This solution is the value of the criterion generated by an optimal policy
  - A policy that achieves the ‘min’ in the OE is an optimal policy
  - Various computational procedures for finding v*
  - ‘Curse of dimensionality’ for large problems

Extensions & limitations

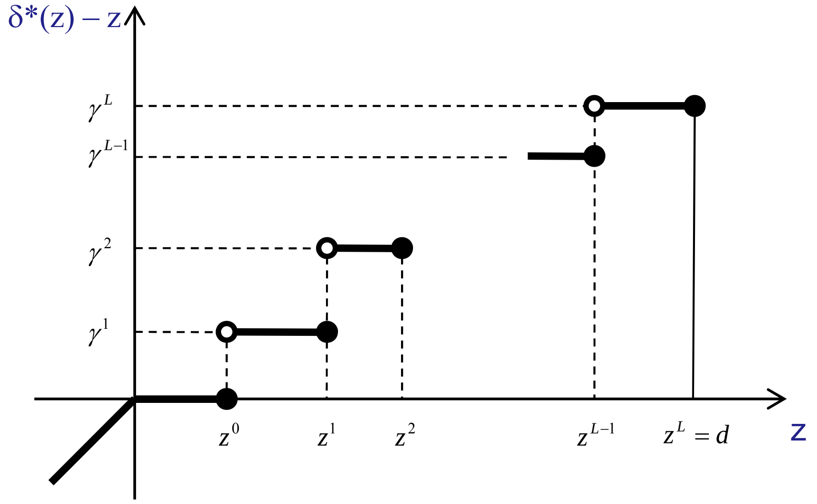
- The MDP is a robust model – many applications, but also many limitations:
  - State may not be completely observed – partially observed MDP (POMDP)
  - Multiple agents, each with a different information pattern & objectives, perhaps only partially known to the other agents. Different sources of uncertainty (Nature, benevolent, malevolent) – cooperative & non-cooperative sequential games
  - Multiple/vector criteria
  - Imprecise parameter values; various knowledge representations
  - Non-additive criteria, risk, & the violation of the ‘principle of optimality’
  - Data transmission may be delayed
  - Always computational challenges

Outline

- Motivation & context
- Models of sequential decision making – MDP
- Applications of the standard MDP
  - Inventory replenishment under supply uncertainty
  - Inventory & supply systems
- Extensions of the standard MDP
  - Partially observed MDP
  - In-transit perishable product inspection
  - Value of traffic information for freight transport
- Sequential games, supply chain design, & risk mitigation (future research directions)
- Q&A

Inventory replenishment under supply uncertainty

- Assumptions:
  - No backlogging
  - Single period demand, d, is stationary and deterministic
  - No shrinkage once an item is placed in inventory
  - If a units are ordered, then α units are placed into inventory with probability $P(\alpha|a)$, $\alpha \leq a$
- Results: if x is the inventory level & $z = d - x$, then there is a computationally desirable optimal "staircase" policy δ^* such that:
  - if $z \leq 0$, then $\delta^*(z) = 0$
  - if $z \geq 0$, then $\delta^*(z) \geq z$
  - if $z \geq 0$, then $\delta^*(z) - z$ is monotonically non-decreasing



Outline

- Motivation & context
- Models of sequential decision making – MDP
- Applications of the standard MDP
  - Inventory replenishment under supply uncertainty
  - Inventory & supply systems
- Extensions of the standard MDP
  - Partially observed MDP
  - In-transit perishable product inspection
  - Value of traffic information for freight transport
- Sequential games, supply chain design, & risk mitigation (future research directions)
- Q&A

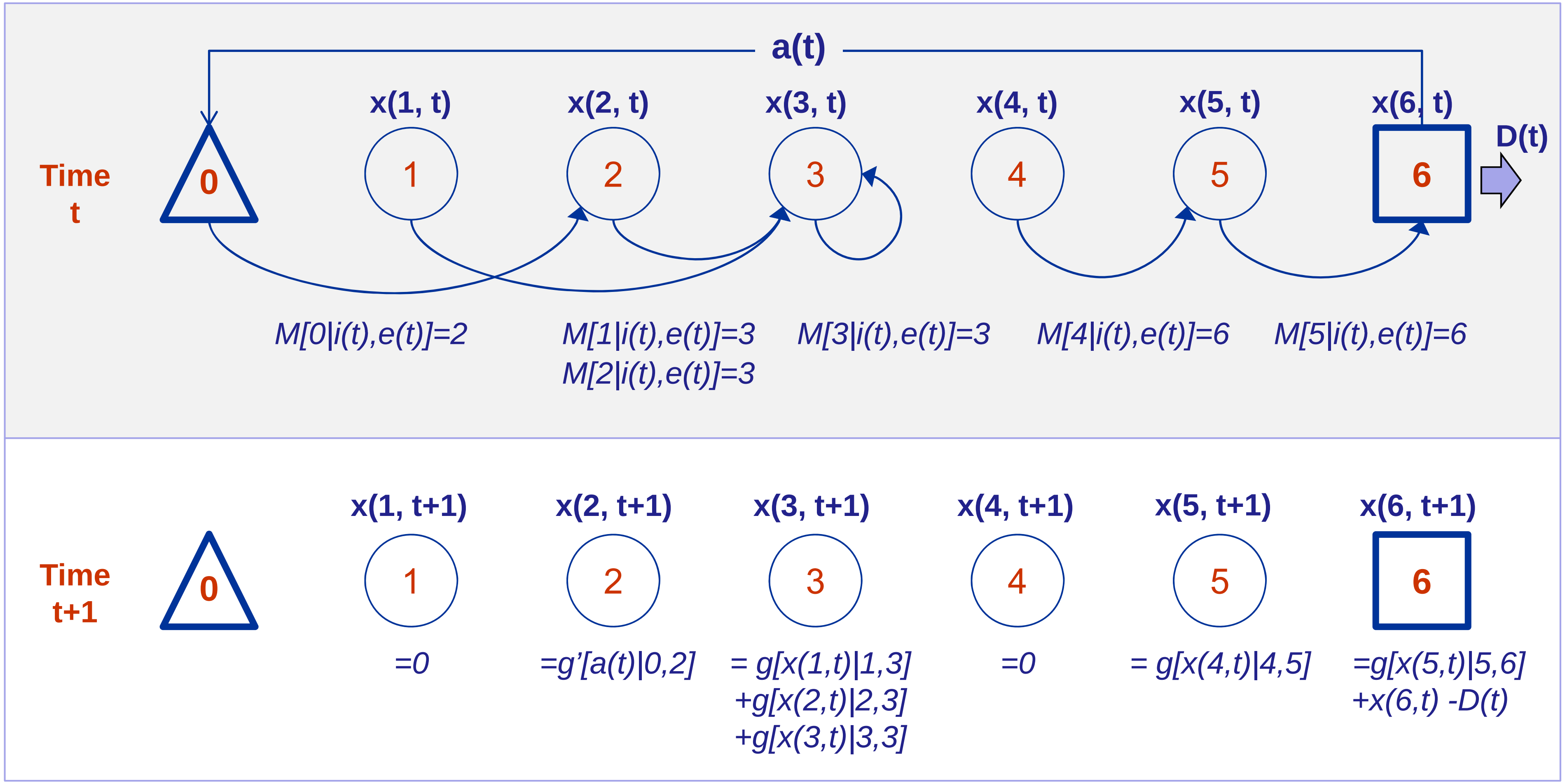
Inventory & supply systems

- Origin (0)
- Destination (N)
- Intermediate locations (1, …, N-1) - the *supply system*
- Number of items ordered at time t: a(t)
- Number of items at location n at time t: x(n, t)
- Demand (i.i.d.) at the destination at time t: D(t)
- Movement from k to M[k| i(t), e(t)], where:
  - i(t+1) = Q[i(t), e(t)], the state of the network
  - e(t) are i.i.d.
  - no ‘cross over’; M monotone

Inventory & supply systems

- Shrinkage & state dynamics:
  - $g'[a(t)|i(t), e(t)] \leq a(t)$ arrives at $M[0|i(t), e(t)]$ at $t+1$ (random yield)
  - $x(n, t+1) = \Sigma\, g[x(k, t)|\, k, n]$ for $n = 1, \ldots, N-1$, where Σ is over all $k > 0$ such that $M[k|\, i(t), e(t)] = n$ (deterministic shrinkage)
  - $x(N, t+1) = x(N, t) + \Sigma\, g[x(k, t)|\, k, N] - D(t)$, where Σ is over all $k > 0$ such that $M[k|\, i(t), e(t)] = N$ (backlogging allowed)
  - $g[g(x|k, m)|m, n] = g(x|k, n)$
  - $x(t+1) = f[x(t), i(t), a(t), e(t)]$ – large state space

Inventory & supply systems



Inventory & supply systems

- Cost structure:
  - fixed order cost K
  - per unit wholesale cost c
  - per item per period holding cost h
  - backlogging penalty per item σ
- Cost accrued in period (t, t+1), where $\delta(a)$ is the indicator function & Σ is over all n = 1, …, N:

$$K\delta(a) + ca + h\,\Sigma\, x(n, t+1) - (h + \sigma)\min\{0, x(N, t+1)\}$$

Inventory & supply systems

- Results:
  - There exists a computationally desirable & easy to implement sufficient statistic, the *effective pipeline inventory position*, I(x)
  - $I(x) = \Sigma\, g[x(n)|n, N]$, where Σ is over all $n = 1, \ldots, N$
  - There is an optimal policy with (s, S) structure
  - There are good suboptimal policies based on this sufficient statistic when yield is also random in the supply system

Outline

- Motivation & context
- Models of sequential decision making – MDP
- Applications of the standard MDP
  - Inventory replenishment under supply uncertainty
  - Inventory & supply systems
- Extensions of the standard MDP
  - Partially observed MDP
  - In-transit perishable product inspection
  - Value of traffic information for freight transport
- Sequential games, supply chain design, & risk mitigation (future research directions)
- Q&A

Markov decision process – partially observed problem definition

- Discrete time t, state s(t), action a(t), observation z(t)
- *State & observation dynamics*: P[z(t+1), s(t+1)| s(t), a(t)]
- P[z(t+1), s(t+1)| s(t), a(t)] = P[z(t+1)|s(t+1), s(t), a(t)] x P[s(t+1)| s(t), a(t)]
- *Information pattern*: a(t) = d[h(t), t] for all t, where h(t) = {z(t), z(t-1), …, a(t-1), a(t-2), …} is *history*
- Cost per period: c[s(t), a(t)]
- *Criterion*: expected total discounted cost over the problem horizon (finite, random, infinite)
- Problem *objectives*:
  - Find a *policy* {d[h(t), t], t = 0, 1, …} that minimizes the criterion
  - Find the value of the criterion generated by an optimal policy
- Basis for determining the *value of information*

POMDP – finding a solution

- $x(t)$ is a sufficient statistic, where $x(i, t) = P[s(t) = i \mid h(t)]$.
- State space of $x(t)$ is uncountable
- The *optimality equation* (infinite horizon):

$$v(x) = \min \{xc(a) + \beta\Sigma\sigma(z, x, a)\, v[\lambda(z, x, a)]: a\}$$

- where:
  - $xc(a)$ is $E[c(s(t), a) \mid x(t) = x]$
  - sum Σ is over all observations z
  - term $\sigma(z, x, a) = P[z(t+1) = z \mid x(t) = x, a(t) = a]$
  - jth element of $\lambda(z, x, a) = P[s(t+1) = j \mid z(t+1) = z, x(t) = x, a(t) = a]$, the posterior to the prior $x(t)$ (Bayes' Rule)

POMDP – finding a solution

- Results:
  - There exists a unique solution to the OE, v*
  - This solution is the value of the criterion generated by an optimal policy
  - A policy that achieves the ‘min’ in the OE is an optimal policy
  - v* is concave & (usually) piecewise linear in x; thus, v* (usually) has a *finite* representation but number of facets can be enormous
  - Various computational procedures for finding v*
  - Improved observation quality improves optimal policy performance.
  - Improved observation quality does *not* necessarily improve sub-optimal policy performance.

Outline

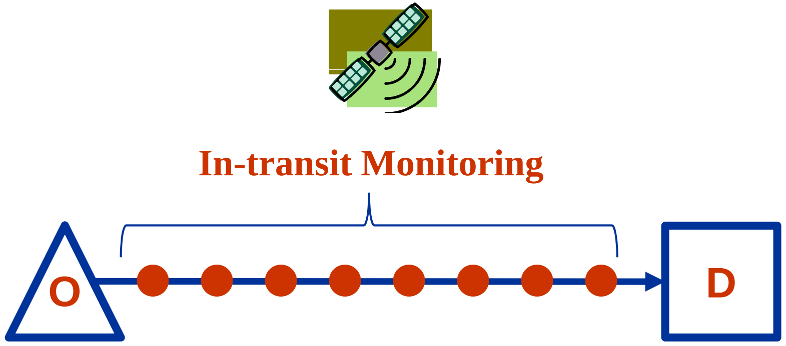
- Motivation & context
- Models of sequential decision making – MDP
- Applications of the standard MDP
  - Inventory replenishment under supply uncertainty
  - Inventory & supply systems
- Extensions of the standard MDP
  - Partially observed MDP
  - In-transit perishable product inspection
  - Value of traffic information for freight transport
- Sequential games, supply chain design, & risk mitigation (future research directions)
- Q&A

In-transit perishable product inspection

- Current practice in industry

  - No temperature logger deployed
  - Temperature logger deployed, but only revealed upon arrival at the destination
    ➔ Decision made at the destination only

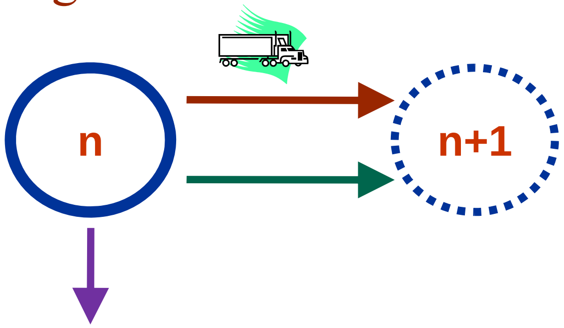
- Suggested model: incorporation of temp sensor technology in a cold chain



  - Temp sensor technology deployed
  - Temp information accessible by the decision maker when requested
  - Availability of such information in transit can lead to ***decisions made in transit*** that can improve supply chain efficiency

In-transit perishable product inspection

- Single vehicle transports perishable goods from an origin (0) to a destination (N) through locations (1, …, N-1)
  - Freight is subject to quality degradation as the vehicle travels toward the destination
- Possible actions at each location:
  - Continue to next location & (perfectly) observe freight status for a cost
  - Continue to next location & do not observe
  - Alternative action
- Possible alternative actions:
  - Dispose of goods, return to origin, re-load
  - Dispose of goods, expedite
  - Sell goods at a secondary market
- When should we observe, not observe, or take alternative action? What value is there for being able to monitor freight status in transit?



In-transit perishable product inspection

- Let:
  - $v(n, x)$ is the solution of the OE
  - $v\#(n, x)$ is the number of facets needed to describe $v(n, x)$
  - $d^*(n, x)$ is an optimal policy
- Results:
  - $v\#(N, x) = 1$; $v\#(n, x) \leq v\#(n+1) + 2$
  - If x is 'at least as fresh' as x', then $v(n, x) \leq v(n, x')$
  - As x becomes 'fresher', then d* tends to move from "alternative" to "observe" to "don't observe"
  - When x is 'spoiled', then select "alternative"
  - When observations are free, then $d^*(n, i)$ has a control-limit that is non-increasing in n

In-transit perishable product inspection

- Implications:
  - Results can lend insight into whether or not to acquire the capacity to observe freight status in transit
  - How to extract value of this capability
  - How much value can be extracted

Outline

- Motivation & context
- Models of sequential decision making – MDP
- Applications of the standard MDP
  - Inventory replenishment under supply uncertainty
  - Inventory & supply systems
- Extensions of the standard MDP
  - Partially observed MDP
  - In-transit perishable product inspection
  - Value of traffic information for freight transport
- Sequential games, supply chain design, & risk mitigation (future research directions)
- Q&A

Value of traffic information for freight transport

- Urban freight transport, using (near) real-time traffic congestion information
- Modeling tools: POMDP & road network models
- *Dynamic route determination*: what road link to take next, based on real-time traffic congestion information
- *Dynamic tour determination* (the dynamic TSP): what stop to visit next, based on real-time traffic congestion information
  - Use of (lower bound) deterministic TSP solutions to guide AI-based search procedures
- Both dynamic approaches show promise for freight transportation productivity improvement
- How is the value of information affected by delays in traffic data collection, analysis, & transmission?

Outline

- Motivation & context
- Models of sequential decision making – MDP
- Applications of the standard MDP
  - Inventory replenishment under supply uncertainty
  - Inventory & supply systems
- Extensions of the standard MDP
  - Partially observed MDP
  - In-transit perishable product inspection
  - Value of traffic information for freight transport
- Sequential games, supply chain design, & risk mitigation (future research directions)
- Q&A

Supply chain design & risk mitigation

- Consider a (non-cooperative) sequential game between attackers and a defender that pursue opposite objectives.
- The attackers must choose one of several possible targets.
- The defender must decide how to defend them.
- The defender has available assets to protect the targets & must distribute the assets among the targets.
- The attackers communicate, and may coordinate or compete, with each other.
- When the defender can manipulate the information available to the attackers:
  - What is the value of disrupting communication between attackers?
  - What is the value of deception (mis-information)?

Outline

- Motivation & context
- Models of sequential decision making – MDP
- Applications of the standard MDP
  - Inventory replenishment under supply uncertainty
  - Inventory & supply systems
- Extensions of the standard MDP
  - Partially observed MDP
  - In-transit perishable product inspection
  - Value of traffic information for freight transport
- Sequential games, supply chain design, & risk mitigation (future research directions)
- Q&A

Thank you!