# Approximate Dynamic Programming (Via Linear Programming) For Stochastic Scheduling

Mohamed and Nelson
2.997 Project

May 12, 2004

# Outline

- Scheduling and stochastic scheduling problems
- Problem statement and formulation as an MDP
- Reformulation into a stochastic shortest path problem
- LP approach to approximate DP - Quick review
- Main result and outline of proof
- Questions and open issues

# Scheduling Problems

- Given a set of tasks and limited resources, we need to efficiently use the resources so that a certain performance measure is optimized
- Scheduling is everywhere: manufacturing, project management, computer networks, etc.
- Almost all interesting scheduling problems are computationally intractable
- Have to settle for near-optimal or approximate solutions

# Simple Example - 2 Machines

- $n$ jobs
- Processing time of job $i$: $p_i$, deterministic
- Objective: minimize the sum of completion times on 2 identical parallel machines
- Number of states is exponential $\Rightarrow$ Can't solve this problem by enumeration
- In fact, no polynomial algorithm is known
- Note that the problem is deterministic and yet remains quite hard

# Stochastic Scheduling - An example

# Stochastic Scheduling

- Many scheduling problems are plagued with uncertainties
- Stochastic scheduling problem: $p_i$'s follow some probability distribution
- Uncertainty $\Rightarrow$ Larger state space

# Problem Definition

- Set of jobs $N = \{1, \dots, n\}$
- 1 machine
- Processing time of job $i$: discrete probability distribution $p_i$
  - $p_i$ and $p_j$ pairwise stochastically independent for all $i \neq j$
- The jobs have to be scheduled *nonpreemptively*
- Objective: minimize

$$\gamma\left(C^{(1)}, \dots, C^{(n)}\right) = \frac{1}{n} \sum_{i=0}^{n-1} h\left(R_i\right)\left(C^{(i+1)} - C^{(i)}\right)$$

# Problem Definition - Continued

Objective: minimize

$$\gamma\left(C^{(1)},\ldots,C^{(n)}\right)=\frac{1}{n}\sum_{i=0}^{n-1}h\left(R_i\right)\left(C^{(i+1)}-C^{(i)}\right)$$

$C^{(i)}$ = time of the $i$th job completion, $C^{(0)}=0$

$R_i$ = set of jobs remaining to be processed at the time of the $i$th job completion

$h$ is a set function such that $h\left(\emptyset\right)=0$

Such an objective function is said to be *additive.*

# MDP formulation

- Finite horizon, finite state space
- State of the system:

$$x = (C_{max}(x), R_x) \in \mathcal{S}$$

$C_{max}(x)$ is the completion time of the last job completed

$R_x$ is the set of jobs remaining to be scheduled at state $x$

Note that the size of the state space is exponential in the number of jobs.

- Action at state $x$ is the next job to be processed: $a \in \mathcal{A}_x \subset R_x$

# MDP formulation - Continued

- Time stage costs:

$$g_a(x,y) = \frac{1}{n} h(R_x)(C_{max}(y) - C_{max}(x)).$$

- Transition probabilities:

$$P_a(x,y) = \begin{cases} p_a(t) & \text{if } R_y = R_x \setminus \{a\} \text{ and } C_{max}(y) = C_{max}(x) + t \\ 0 & \text{otherwise.} \end{cases}$$

# MDP formulation - Continued

- Solve for the finite-horizon cost-to-go function

$$J^*(x,n) = 0$$

$$J^*(x,t) = \min_{a \in \mathcal{A}_x} \left\{ \sum_{y \in \mathcal{S}} P_a(x,y) \left( g_a(x,y) + J^*(y,t+1) \right) \right\}, \quad t = 0, 1, \ldots, n-1$$

- Exponential state space ⇒ exact DP hopeless

- Approximate DP methods consider infinite horizon problems

  ⇒ Recast our problem as stochastic shortest path problem

# Reformulation into SSP

- Introduce a terminating state $\bar{x}$

  Only states with $R_x = \emptyset$ can reach $\bar{x}$ in one step

- Transition probabilities involving $\bar{x}$:

$$P_a(x, \bar{x}) = \begin{cases} 1 & \forall x \text{ such that } R_x = \emptyset \\ 1 & \text{if } x = \bar{x} \\ 0 & \text{otherwise} \end{cases}$$

- Time-stage costs involving $\bar{x}$:

$$g_a(x, \bar{x}) = \begin{cases} 0 & \forall x \text{ such that } R_x = \emptyset \\ 0 & \text{if } x = \bar{x}. \end{cases}$$

# Reformulation into SSP - Continued

- Cost-to-go function for SSP formulation:

$$J^*(x) = \min_u E\left[\sum_{t=0}^{T(x)} g_u(x_t, x_{t+1}) \middle| x_0 = x\right].$$

$T(x)$ = time stage when the system reaches the terminating state

- Every policy reaches the terminating state in a finite number of steps with probability 1

- $\Rightarrow$ The cost-to-go function for the SSP problem is the unique solution to Bellman's equation

# Approximate DP Via ALP

- Exact linear program (ELP):

$$\begin{array}{ll} \text{maximize} & c^T J \quad (c > 0) \\ \text{subject to} & TJ \geq J \end{array}$$

- Approximate linear program (ALP):

$$\begin{array}{ll} \text{maximize} & c^T \Phi r \quad (c > 0) \\ \text{subject to} & T\Phi r \geq \Phi r \end{array}$$

- $\tilde{r}$ is optimal solution to ALP $\Rightarrow$ obtain a (hopefully) good policy by using the greedy policy with respect to $\Phi\tilde{r}$

# Approximate DP Via ALP - Continued

Error bound for the ALP approach for discounted cost problems:

**Theorem 1.** *(de Farias and Van Roy 2003) Let* $\tilde{r}$ *be a solution of the approximate LP. Then, for any* $v \in \mathbb{R}^K$ *such that* $(\Phi v)(x) > 0$ *for all* $x \in \mathcal{S}$ *and* $\alpha H \Phi v < \Phi v$*,*

$$\left\| J^* - \Phi \tilde{r} \right\|_{1,c} \leq \frac{2c^T \Phi v}{1 - \beta_{\Phi v}} \min_r \left\| J^* - \Phi r \right\|_{\infty, 1/\Phi v}$$

*where*

$$(H \Phi v)(x) = \max_{a \in \mathcal{A}_x} \left\{ \sum_{y \in \mathcal{S}} P_a(x, y)(\Phi v)(y) \right\}$$

*and*

$$\beta_{\Phi v} = \max_x \frac{\alpha (H \Phi v)(x)}{(\Phi v)(x)}.$$

# Relaxed Stochastic Shortest Path Problem

- Relaxation: introduce discount factor $\alpha \in (0,1)$ at each time stage
- Call this formulation the *$\alpha$-relaxed SSP formulation*
- Cost-to-go function for relaxed formulation:

$$J^{*,\alpha}(x) = \min_u E\left[\sum_{t=0}^{T(x)} \alpha^t g_u(x_t, x_{t+1}) \middle| x_0 = x\right]$$

- For this relaxation, we show that the error of the ALP solution is uniformly bounded over the number of jobs to be scheduled.

# Main Result

**Theorem 2.** *Assume that the holding cost $h(S) \leq M$ for all subsets $S$ of $N$. Let $\tilde{r}$ be the ALP solution to the $\alpha$-relaxed SSP formulation of the stochastic scheduling problem. For $\alpha \in (0, 1)$,*

$$\|J^{*,\alpha} - \Phi\tilde{r}\|_{1,c} \leq \frac{2M \max_{i \in N} E[p_i]}{1 - \alpha}.$$

- The error is uniformly bounded over the number of jobs
- How amazing is that?

# Outline of Proof

The cost-to-go function is

$$J^{*,\alpha}(x) = \min_u E\left[\sum_{t=0}^{T(x)} \alpha^t g_{u(x_t)}(x_t, x_{t+1}) \middle| x_0 = x\right]$$

where

$$g(x_t, x_{t+1}) = \frac{1}{n} h(R_x)\left(C_{max}(x_{t+1}) - C_{max}(x_t)\right)$$

Recall $h$ is bounded from above by $M$. After some algebraic manipulation, this quantity is found to be

$$\leq \frac{M}{n} \sum_{i \in R_x} E[p_i]$$

# Outline of Proof - Continued

Let

$$V\left(x\right) = \frac{k}{n} \sum_{i \in R_x} E\left[p_i\right]$$

$V$ is a Lyapunov function

$\exists\ \beta < 1$ independent of $n$ such that $\alpha HV \leq \beta V$

Also,

$$\min_r \left\|J^{*,\alpha} - \Phi r\right\|_{\infty,1/V} \leq \frac{M}{k}$$

# Outline of Proof - Continued

Consider $c^T V$, $c$ some probability distribution over $\mathcal{S}$

We have the following uniform bound

$$\sum_{x \in \mathcal{S}} c(x) V(x) \leq k \max_{i \in N} E[p_i]$$

Combining these results,

$$\|J^{*,\alpha} - \Phi\tilde{r}\|_{1,c} \leq \frac{2M \max_{i \in N} E[p_i]}{1-\alpha}$$

# Conclusions and Remarks

- ALP approach has an error bound for our relaxed stochastic scheduling problem that does not grow with the number of jobs to be scheduled
- What about $\alpha = 1$? (original SSP formulation)
- Multiple machines?
- Computational experiments: how does ALP perform in practice?

Questions?