Priority Policies in Scheduling

Daniel Russo

March 18, 2020

Columbia University

Setup

- Suppose there are m queues and a single server.
- There are no new arrivals.
- A cost $g(i)$ is incurred per unit time per customer in queue i
- Service time dist. for a customer in queue i is $\text{exponential}(\mu_i)$
- After service is completed a customer in queue i
    1. Leaves the system with probability p_{i0}
    2. Joins queue $j > 0$ with probability p_{ij}.
- Costs are discounted exponentially at rate $\beta > 0$.

Main result:

A priority rule is optimal. Such a policy orders the m queues and, at each decision period, services the highest priority non-empty queue.

This result follows by using a few clever ideas to reduce the problem to a bandit and then applying the Gittins index theorem.

First idea: track the state of the customers not the queues

- There are initially n customers in the system, and we index them by $\ell \in \{1, \cdots n\}$.
- Let $i^{\ell}(t) \in \{0, \cdots, m\}$ be the state of customer ℓ at time $t \in \mathbb{R}_{+}$. The absorbing state 0 represents departure.
- $u(t) \in \{0, \cdots, n\}$ indicates which customer is being served.
  - Serving a customer in state 0 is feasible, but not optimal.
- The objective is to minimize

$$\mathbb{E}\left[\int_{t=0}^{\infty} \sum_{\ell=1}^{n} g\left(i^{\ell}(t)\right) dt\right]$$

where we set $g(0) = 0$.

Second idea: uniformization

As we have seen before, a trick called uniformization allows us to reduce discounted continuous time problems to discounted discrete time problems.

- Set $\mu = \max_i \mu_i$.
- Introduce random event times (potentially fictitious. . .):
  - $t_0 = 0, t_1, t_2, \cdots$
  - $\tau_1 = t_1 - t_0$, $\tau_2 = t_2 - t_1 \overset{\text{i.i.d}}{\sim} \text{exponential}(\mu)$
  - Set $i_k^\ell \equiv i^\ell(t_k)$.
- Modify costs and transition probabilities:
  - $\tilde{p}_{ii} = \frac{\mu_i}{\mu} p_{ii} + \frac{\mu - \mu_i}{\mu}$ and $\tilde{p}_{ij} = \frac{\mu_i}{\mu} p_{ij}$ for $j \neq i$.
  - $\tilde{g}(i) = \frac{1}{\beta + \mu} g(i)$.
  - Set $\alpha = \frac{\mu}{\beta + \mu}$ to be the effective discount factor.
- Our original continuous time problem is equivalent to:

$$\text{Minimize}_\pi \quad \mathbb{E}^\pi \left[\sum_{k=0}^{\infty} \alpha^k \left(\sum_{\ell=0}^{n} \tilde{g}(i_k^\ell) \right) \right]$$

Third idea: move expected future costs to the present period

Right now, our problem does not look like a bandit because we incur costs for customers that are not served. An accounting trick moves all costs to the period in which service is provided.

- Define $R(i,j) = \frac{\alpha}{1-\alpha}\tilde{g}(i) - \frac{\alpha}{1-\alpha}\tilde{g}(j)$
- Set $R(i) = \sum_{i=0}^{n} \tilde{p}_{ij} R(i,j)$.
  - Interpret $R(i)$ as the reduction in expected cost due to servicing a customer in queue i *if we were to assume this is the final service they receive.*

- One can show our objective is equivalent to that in a bandit problem:

$$\mathbb{E}^{\pi}\left[\sum_{k=0}^{\infty}\alpha^k\left(\sum_{\ell=0}^{n}\tilde{g}(i_k^{\ell})\right)\right] = \sum_{\ell=1}^{n}\frac{1}{1-\alpha}\tilde{g}(x_0^{\ell}) - \mathbb{E}^{\pi}\left[\sum_{k=0}^{\infty}\alpha^k R(i_k^{u_k})\right]$$

Justifying the accounting trick

Consider customer ℓ.

Let $T_1, T_2, \cdots$ denote the times at which it is played/serviced.

Let $i_1, i_2, \cdots$ denote its states at those times.
The total cost contribution from customer ℓ is:

$$
\begin{aligned}
&\mathbb{E}\left[\sum_{k=0}^{T_1} \alpha^k \tilde{g}(i_1) + \sum_{k=T_1+1}^{T_2} \alpha^k \tilde{g}(i_2) + \cdots\right] \\
=&\mathbb{E}\left[\left(\sum_{k=0}^{\infty} \alpha^k \tilde{g}(i_1) - \sum_{k=T_1+1}^{\infty} \alpha^k \tilde{g}(i_1)\right) + \left(\sum_{k=T_1+1}^{\infty} \alpha^k \tilde{g}(i_2) - \sum_{k=T_2+1}^{\infty} \alpha^k \tilde{g}(i_2)\right) + \cdots\right] \\
=&\mathbb{E}\left[\left(\sum_{k=0}^{\infty} \alpha^k \tilde{g}(i_1)\right) + \left(\sum_{k=T_1+1}^{\infty} \alpha^k \left(\tilde{g}(i_2) - \tilde{g}(i_1)\right)\right) + \left(\sum_{k=T_2+1}^{\infty} \alpha^k \left(\tilde{g}(i_3) - \tilde{g}(i_2)\right)\right) + \cdots\right] \\
=&\mathbb{E}\left[\frac{1}{1-\alpha}\tilde{g}(i_1) + \alpha^{T_1} \cdot \frac{\alpha}{1-\alpha} \cdot (\tilde{g}(i_2) - \tilde{g}(i_1)) + \alpha^{T_2} \cdot \frac{\alpha}{1-\alpha} \cdot (\tilde{g}(i_3) - \tilde{g}(i_2)) + \cdots\right] \\
=&\frac{\tilde{g}(i_1)}{1-\alpha} - \mathbb{E}\left[\alpha^{T_1} R(i_1) + \alpha^{T_2} R(i_2) + \cdots\right]
\end{aligned}
$$

Final step: applying the Gittins index theorem

We have the problem of maximizing $\mathbb{E}\left[\sum_{k=0}^{\infty} \alpha^k R(i_k^{u_k})\right]$, where when played in state i bandit ℓ's next state is j with probability $\tilde{p}_{ij}$.

- By symmetry, the Gittins index $G^\ell : \{0, \cdots, m\} \to \mathbb{R}$ does not depend on the bandit process (i.e. the identity of the customer). We have $G^\ell(\cdot) = G(\cdot)$ for all $\ell \in \{1, \cdots, n\}$
- The Gittins index theorem says the optimal policy selects at time k:

$$u_k^* \in \operatorname*{argmax}_{\ell \in \{1, \cdots, n\}} G(i_k^\ell).$$

  This is a priority rule, where the priority is determined by the values of $G(\cdot)$.