Applications of Dynamic Games in Queues

Eitan Altman

INRIA Sophia-Antipolis
2004 route des Lucioles, BP 93
06902 Sophia-Antipolis Cedex

E-mail: altman@sophia.inria.fr

DEC. 2000, Adelaide

OUTLINE

I. To Q or not to Q

II. When to Q?

III. Where to Q?

IV. Where to Q: the Gaz station problem

V. Where to Q: queues with priority

VI. S-modular games, FTC, ATC

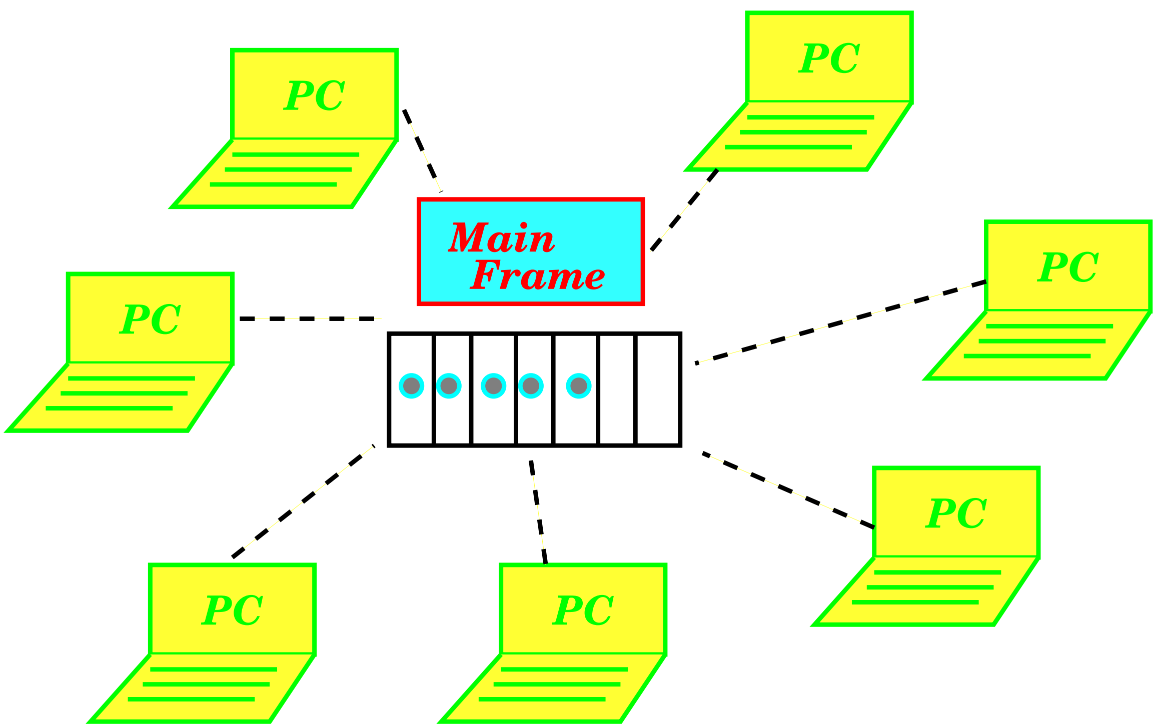
VII. Flow Control Models

Questions related to queueing:

- To Q or not to Q? (Admission-control)
- When should we arrive at a Q? (Scheduling)
- Where to Q? (Routing, Traffic-Assignment)
- Case of elastic demand: how much should we send to a Q? (Flow-control)

To Q or not to Q

Should we Q to receive a service from a shared service provider or use a dedicated one that does not require queueing?



When to Q?

Examples:

- When to arrive to the bank? A bank opens between 9h00 to 12h00. When should one come so as to minimize the expected wainting time?
- When to retry to make a phone call?

Where to Q?

Examples:

- Which path to take in a network?

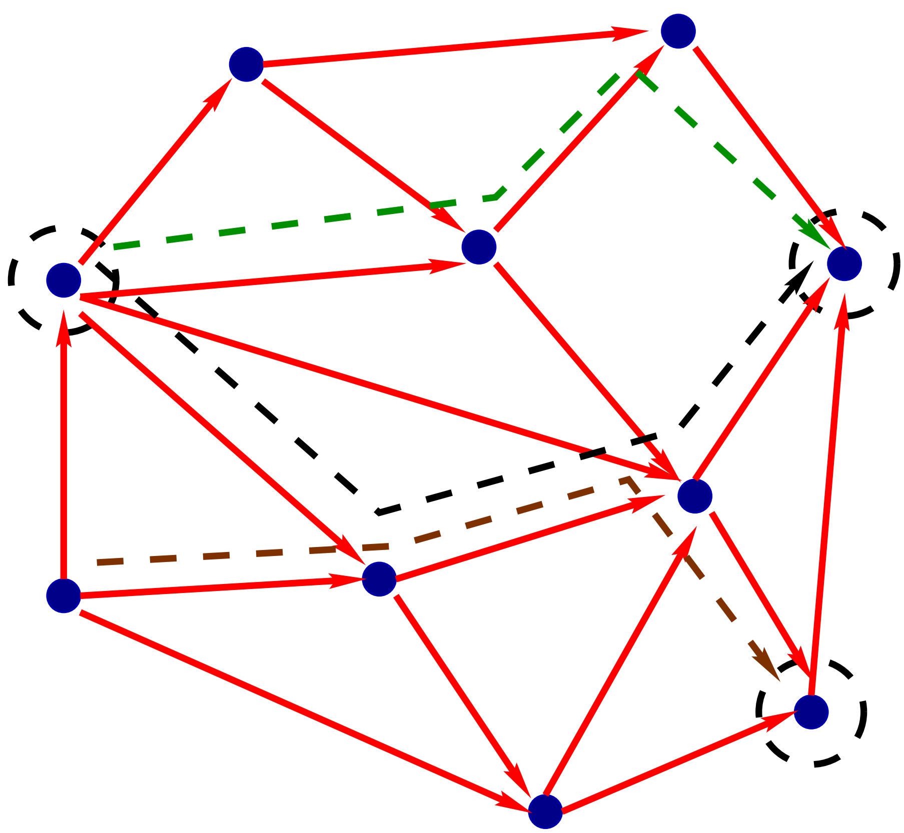


Figure 1: Competitive Routing

- The gaz station problem: a car arrives at a gaz station and observes a line of waiting cars. Should the car wait as well or should it continue to the next gaz station?

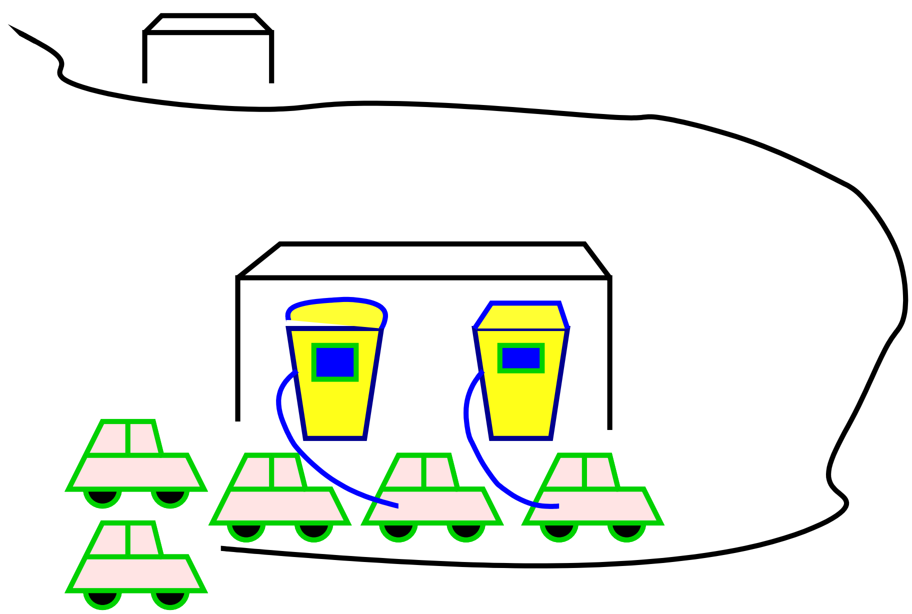


Figure 2: Competitive Routing

How much to Q?

Example: Flow-control

There are two main approaches to flow control in communications networks:

- Window-based flow control:

The sources receive acknowledgements from the destination on well-received packets.

A window is the number of packets that can be sent and not yet acknowledged.

The larger the window, the larger the throughput.

Ex: Internet (ftp, email etc).

- Rate-based flow control:

The source control its transmission rates.

It may receive information from the network on congestion.

Ex: ATM technology.

- Typical objectives: Minimize losses, maximize throughputs, minimize delays.
- The Internet is typically non-cooperative. The control is done at the end points by the users.
- Flow control in ATM networks can often be modeled as a team problem. The control is done within the network.

When to arrive to the bank?

Ref: A. Glazer and R. Hassin, "?/M/1: On the equilibrium distribution of customer arrivals", European J. of OR, 13 (1983), 146-150.

- A bank opens between 9h00 to 12h00. All customers that arrive before 12h00 are served that day.
- A random number X of customers wish to get a service.
- The service time are i.i.d. with exponential distribution.
- The order of service is FCFS (First Come First Served).
- Each customer wishes to minimize the waiting time in the bank.

- There exists a symmetric equilibrium distribution F of the arrival time with support $[T_0, 12h00]$ for all players, with $T_0 < 9h00$, such that if all customers follow F, our optimal policy is to use F.

- Note: if we eliminate the FCFS regime among those who arrive before the bank opens, then $T_0 = 9h00$. This can reduce average waiting time!

When to retry to phone?

Ref:

- R. Hassin and M. Haviv, "On optimal and equilibrium retrial rates in a busy system", *Prob. in the Eng. and Informational Sciences*, Vol. 10, 223-227. 1996.

- A. Elcan, "Optimal customer return rate for an M/M/1 queueing system with retrials", *Probability in the Engineering and Informational Sciences*, **8** (1994), 521-539.

The model

- At congestion times telephone calls might be blocked.

- A person whose call is blocked may typically retry calling.

- Assume that each retrial costs c, and the waiting time costs w per time unit.

- There is some equilibrium retrial rate that can be computed.

More precise model

- Call arrive according to a Poisson process with average rate λ.
- Service rates are i.i.d. with mean τ and finite variance σ^2. Let $S^2 := \tau^2 + \sigma^2$. $\rho := \lambda\tau$.
- Between retrials calls are "in orbit". Times between retrials of the ith call in orbit are exponentially distributed with expected value of $1/\theta_i$.

Socially optimal solution

- The expected time in orbit [Kulkarni, 1983] is

$$W = \frac{\rho}{1-\rho}\left(\frac{1}{\theta} + \frac{S^2}{2\tau}\right),$$

so the average cost per call is

$$(w + c\theta)W =$$

$$\frac{\rho}{1-\rho}\left(\frac{cS^2}{2\tau}\theta + \frac{w}{\theta}\right) + \frac{\rho}{1-\rho}\left(\frac{wS^2}{2\tau} + c\right).$$

This is minimized at

$$\theta^* = \frac{\sqrt{2w\tau/c}}{S}.$$

- THIS IS INDEPENDENT OF THE ARRIVAL RATE!

- If θ^* is used, the two terms that depend on θ are equal: the waiting cost and the retrial costs coincide.

- In the game case, one computes $g(\theta, \gamma)$ [Kulkarni, 83], the expected waiting time of an individual who retries at rate γ while all the others use retrial rate θ.

- This allows us to obtain the equilibrium rate:

$$\theta_e = \frac{w\rho + \sqrt{w^2\rho^2 + 16w\tau c(1-\rho)(2-\rho)/S^2}}{4c(1-\rho)}.$$

- θ_e monotone increases to infinity as λ increases (to $1/\sigma$).

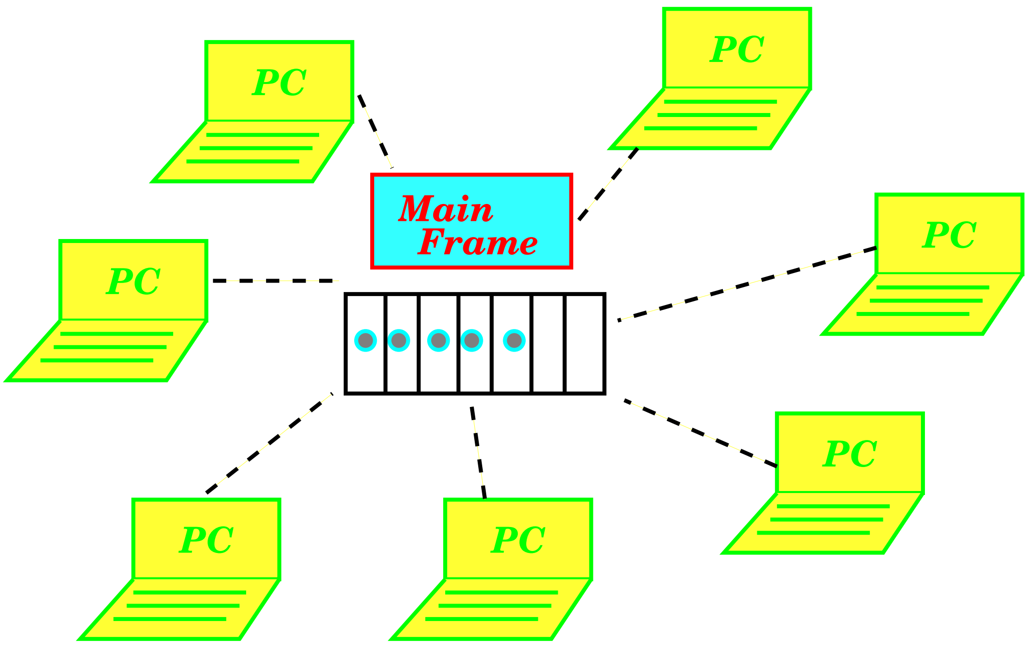
- Thus, the ratio between the equilibrium and the globally optimal cost tends to infinity.

- The equilibrium retrial rate is larger than the optimal retrial rate. They tend to coincide as $\rho \to 0$.

- Both equilibrium and optimal retrial rates are monotone decreasing in the variance of the service times.

To Q or not to Q: PC or MF

Ref: E. Altman and N. Shimkin, "Individual equilibrium and learning in processor sharing systems", *Operations Research*, vol. 46, pp. 776–784, 1998.



-
  - Requests for processing jobs arrive to the system.
  - Interarrival times are genereal i.i.d. with mean λ^{-1}.

- Upon arrival of a request, the user connects to MF and observes the load. Based on this information it decides whether to Q or not to Q there.

-
  - MF shares its computing capacity between all present users there. This is called Processor Sharing discipline.
  - The service at MF is exponentially distributed with rate $\mu(x)$. $x :=$ number of jobs queued there.
  - Typically: $\mu(x) = \mu$ and the service intensity per customer is $\nu(x)/x$.

- A PC offers fixed expected service time of θ^{-1}.

- $X(t)$:= the number of customers at MF at time t.

- $T_k,\ k \geq 0$:= arrival time of job C_k, where $0 = T_0 < T_1 < T_2 < \ldots$ ($t = 0$ is the arrival time of the first customer).

- The queue-length process $X(t)$ is defined to be left-continuous (thus $X(T_k)$ is the queue length just *prior* to the possible admission of customer C_k to MF).

- C_k must decide at T_k which queue to join, after observing $X(T_k)$.

- u_k:= strategy for C_k, is the probability of joining MF if $X(T_k) = x$.

- U:= the class of such maps, and

$\pi := (u_0, u_1, \ldots)$ denotes a multi-strategy of all customers.

Performance measure

- w_k:= service duration of customer C_k,
- $W_k(x,\pi)$:= expected service time of C_k, given that x customers are present at MF at his arrival.
- $V_k(x,\pi)$:= expected service duration of C_k at MF under the same conditions.

Then

$$W_k(x,\pi) = u_k(x)V_k(x,\pi) + (1-u_k(x))\theta^{-1},$$

- Observe that V_k depends on π through $\{u_l,\ l>k\}$, the decision rules of subsequent customers.
- Each customer wishes to minimize her own service time.
- To this end, she should evaluate her expected service time at the two queues, namely $V_k(x,\pi)$ and θ^{-1}, and choose the lower one.

Threshold Policies

For any $0 \le q \le 1$ and integer $L \ge 0$, the decision rule u is an $[L, q]$-threshold rule if

$$u(x) = \begin{cases} 1 & \text{if} \quad x < L \\ q & \text{if} \quad x = L \\ 0 & \text{if} \quad x > L \end{cases} . \tag{1}$$

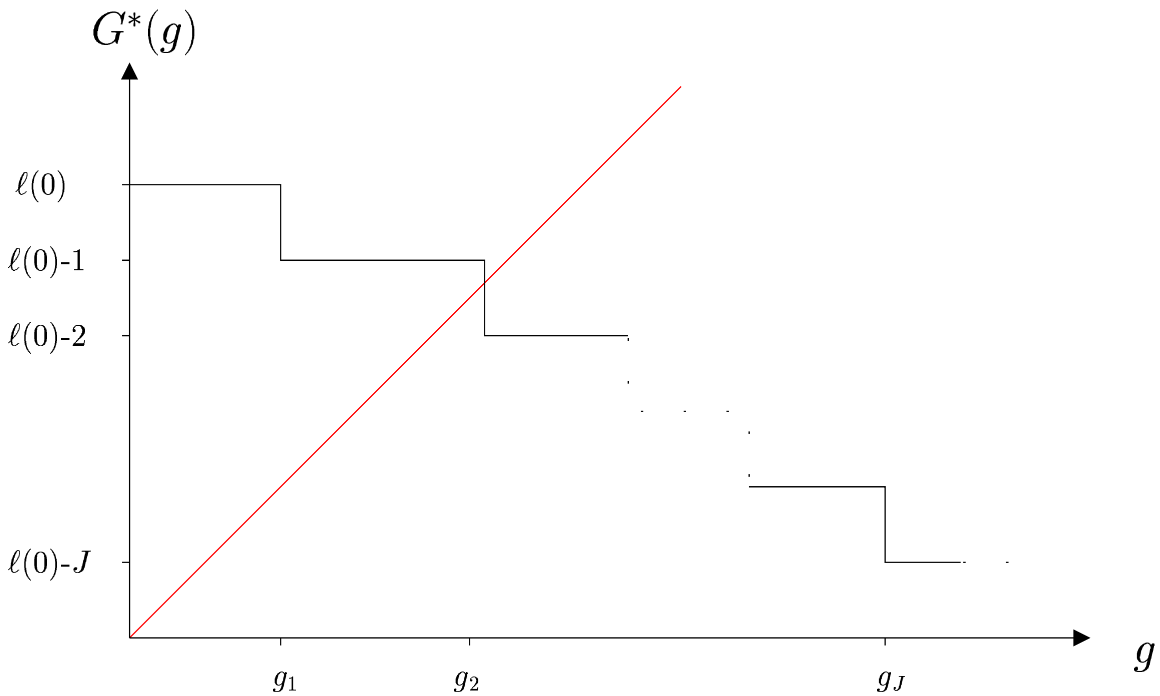
- A customer which employs this rule joins MF if the queue length x is smaller than L, while if $x = L$ she does so with probability q. Otherwise she joins PC.
- An $[L, q]$ threshold rule will be denoted by $[g]$ where $g = L+q$. Note that $[L, 1]$ and $[L+1, 0]$ are identical.

Theorem 1

(i) For any equilibrium policy $\pi^* = (u_0, u_1, \ldots)$, *each decision rule* u_k *is a threshold rule.*

(ii) A symmetric equilibrium policy $\pi^* = (u^*, u^*, \ldots)$ *exists, is unique, and* u^* *is a threshold rule.*

Basic steps of proof:



- (i) For every policy π and $k \geq 0$, $V_k(x, \pi)$ is strictly increasing in x.
- (ii) Assume that all jobs other than C_k use a threshold policy $[g]$. Then $V_k(x, [g]^\infty)$ is:

(i) strictly increasing in g, and

(ii) continuous in g.

Numerical Examples

- Consider $B = \infty, \theta = 10, \mu = 100$.

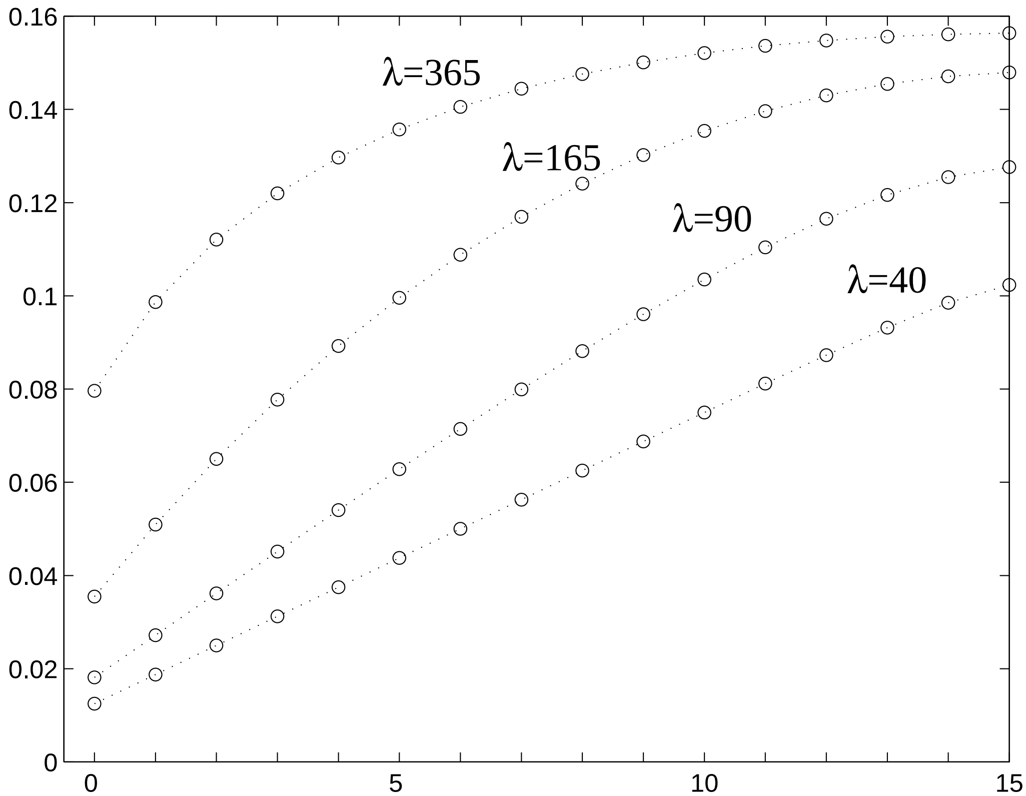


Figure 3: $V(x, [15, 1])$ as a function of x for various λ's

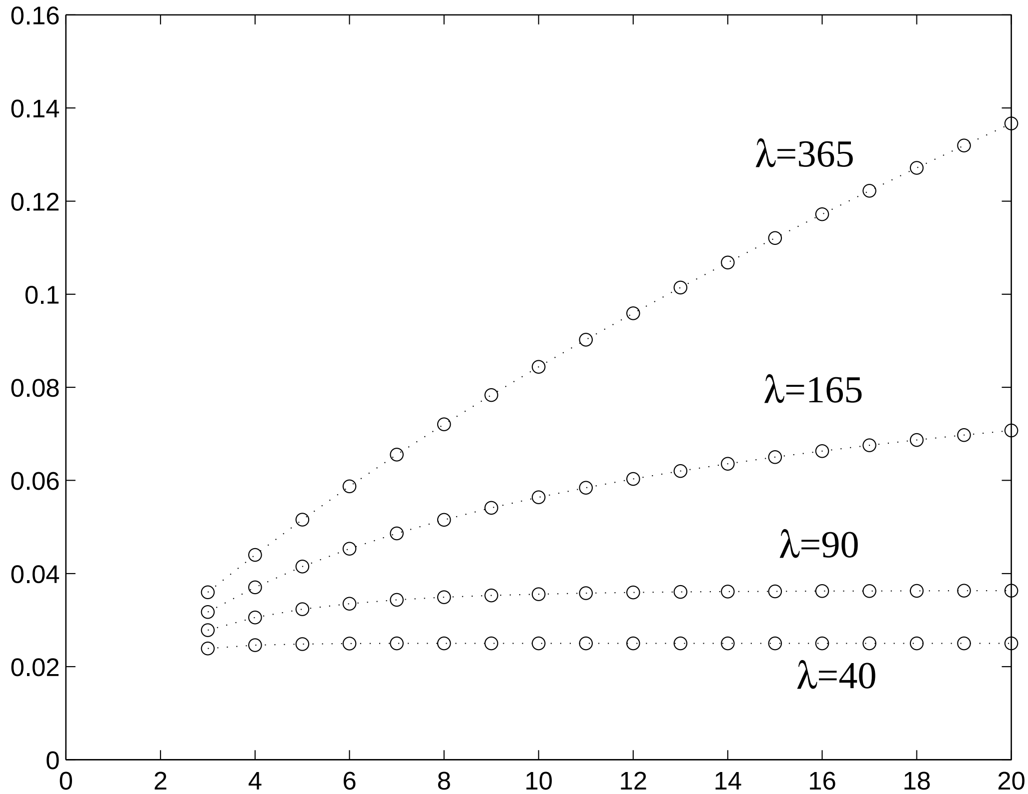


Figure 4: $V(3,[L,1])$ as a function of L for various λ's

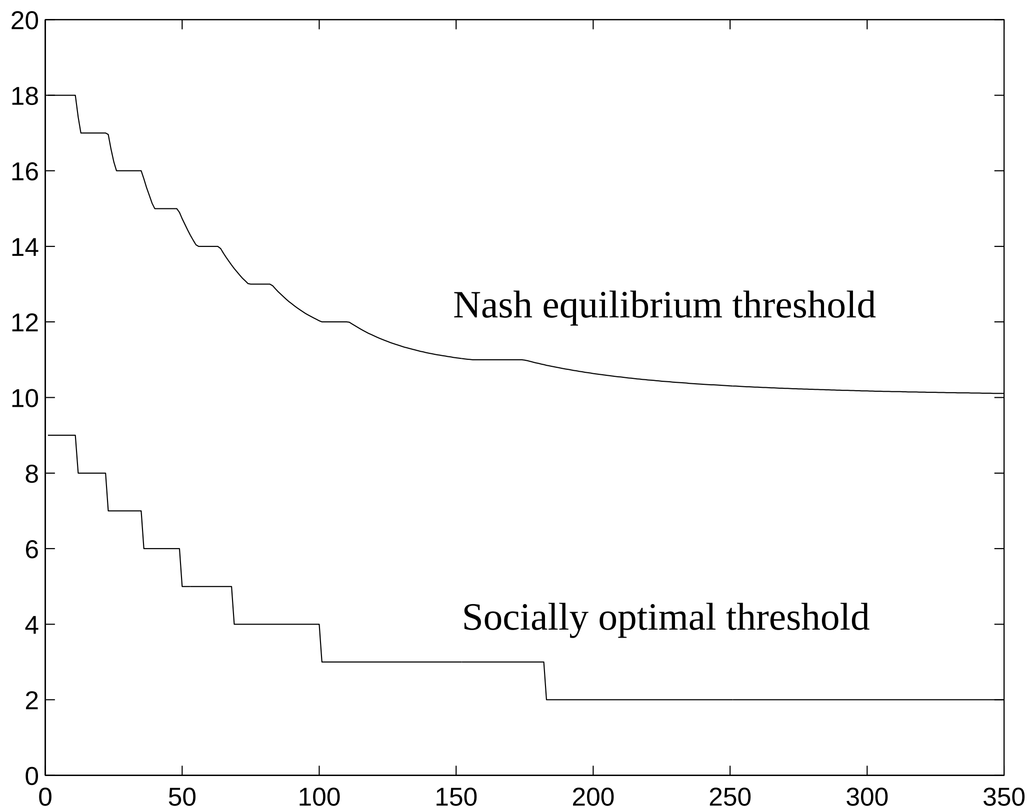


Figure 5: $g^* = L^* + q^*$ as a function of λ

Where to Q: the Gaz station problem

- R. Hassin. On the advantage of being the first server. *Management Science*, 2000.
- E. Altman, T. Jimenez, R. Nunez-Queija and U. Yechiali, Queueing analysis for optimal routing with partial information, 2000.

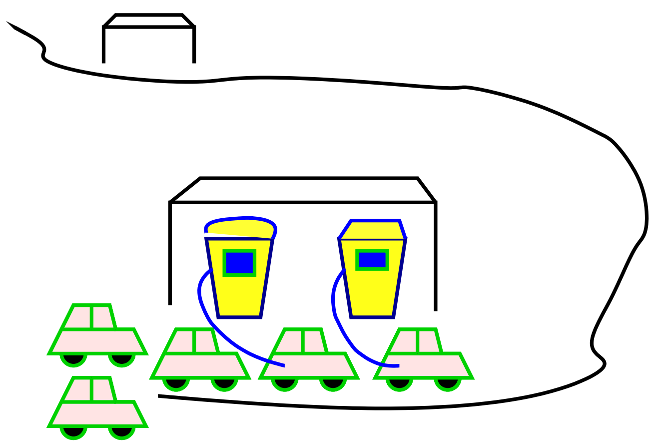


Figure 6: Competitive Routing

- We analyze the dynamic routing choices between two paths.
- When a routing decision is made, the decision maker knows the congestion state of only one of the routes; the congestion state in the second route is unknown to the decision maker.
- Applications in telecommunication networks: the state in a down stream node may become available after a considerable delay, which makes that information irrelevant when taking the routing decisions.
- Although the precise congestion state of the second route is unknown, its probability distribution, which depends on the routing policy, can be computed by the router.
- To obtain an equilibrium, we need to compute the joint distribution of the congestion state in both routes as a function of the routing policy.

- We restrict to random threshold policies (n, r):
  - if the the number of packets in the first path is less than or equal to $n-1$ at the instance of an arrival, the arriving packet is sent to path 1.
  - If the number is n then it is routed to path 1 with probability r.
  - If the number of packets is greater than n then it is routed to path 2.
- The delay in each path is modeled by a state dependent M/M/1 queue:
  - Service time at queue i is exponentially distributed with parameter μ_i
  - Global interarrival times are exponential i.i.d. with parameter λ.

- When all arrivals use policy (n, r), the steady state distribution is obtained by solving the steady state probabilities of the continuous time Markov chain:

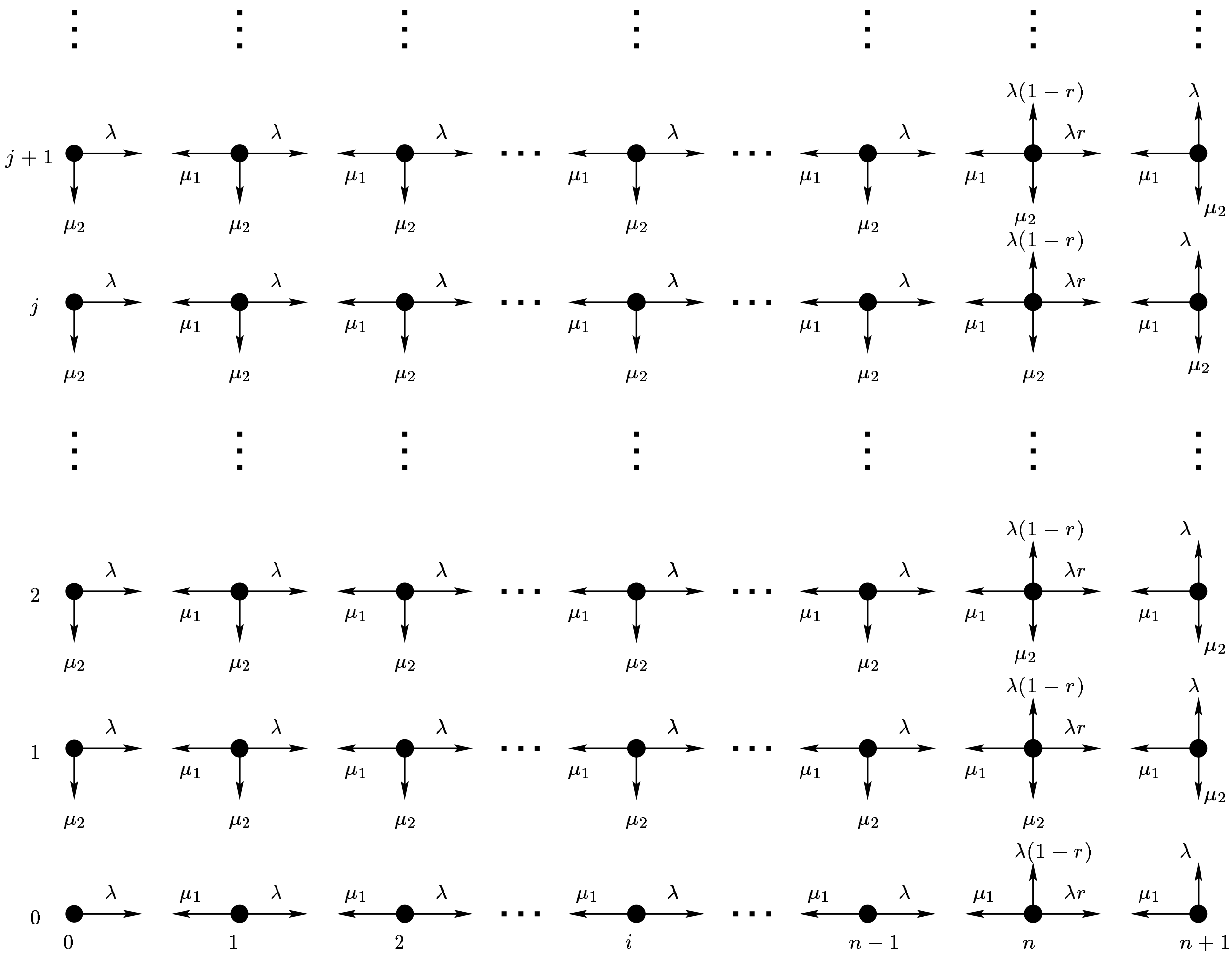


Figure 7: Transition diagram

- If an arrival finds i customers at queue 1, it computes

$$E_i[X_2] = E\left[X_2 | X_1 = i\right]$$

and takes a routing decision according to whether

$$T^{n,r}(i,1) := \frac{i+1}{\mu_1} \leq_? \frac{E_i[X_2]+1}{\mu_2} =: T^{n,r}(i,2).$$

- To compute it, the arrival should know the policy (n, r) used by all previous arrivals.

- If the decisions of the arrival as a function of i coincide with (n, r) then (n, r) is a Nash equilibrium.

- The optimal response against $[g] = (n, r)$ is monotone decreasing in g. This is the Avoid The Crowd behavior.

- Computing the conditional distributions, one can show that there are parameters $(\mu_1, \mu_2, \lambda, n, r)$ for which the optimal response to (n, r) is indeed threshold policy.

- Denote

$$\rho := \frac{\lambda}{\mu_1}, \qquad s := \frac{\mu_2}{\mu_1}$$

There are other parameters for which the optimal response to (n, r) is a two-threshold policy characterized by $t^-(n, \rho, s)$ and $t^+(n, \rho, s)$ as follows.

It is optimal to route a packet to queue queue 2 if $t^-(n, \rho, s) \leq X_1 \leq t^+(n, \rho, s)$ and to queue 1 otherwise.

At the boundaries t^- and t^+ routing to queue 1 or randomizing is also optimal if

$$T^n_{\rho,s}(i, 1) = T^n_{\rho,s}(i, 2)$$

Example

■ Consider

$$n = 3, \quad r = 1, \quad \rho = \frac{\lambda}{\mu_1} = 1 \quad \text{and} \quad s = \frac{\mu_2}{\mu_1} = 0.56$$

We plot $T^{n,r}_{\rho,s}(i,1)$ and $T^{n,r}_{\rho,s}(i,2)$ for $i = 0, 1, \ldots, 4$.

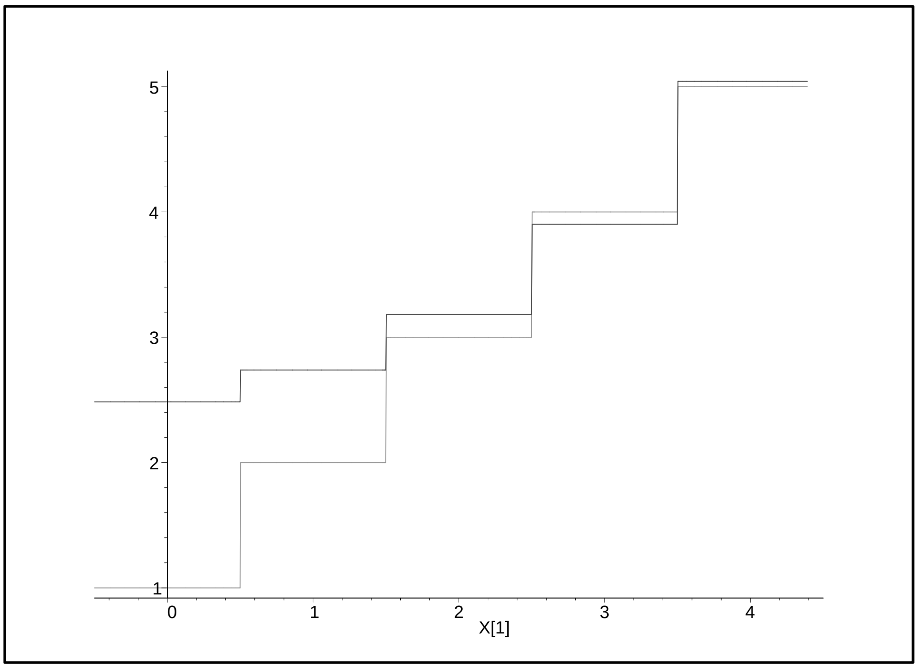


Figure 8: $T^{n,r}_{\rho,s}(i,1)$ and $T^{n,r}_{\rho,s}(i,2)$

Conclusions: for some parameters there may be no (n, r) equilibria!

Where to Q: queues with priority

- I. Adiri and U. Yechiali, "Optimal priority purchasing and pricing decisions in nonmonopoly and monopoly queues", *Operations Research*, 1974
- R. Hassin and M. Haviv, "Equilibrium thrshold strategies: the case of queues with priorities".

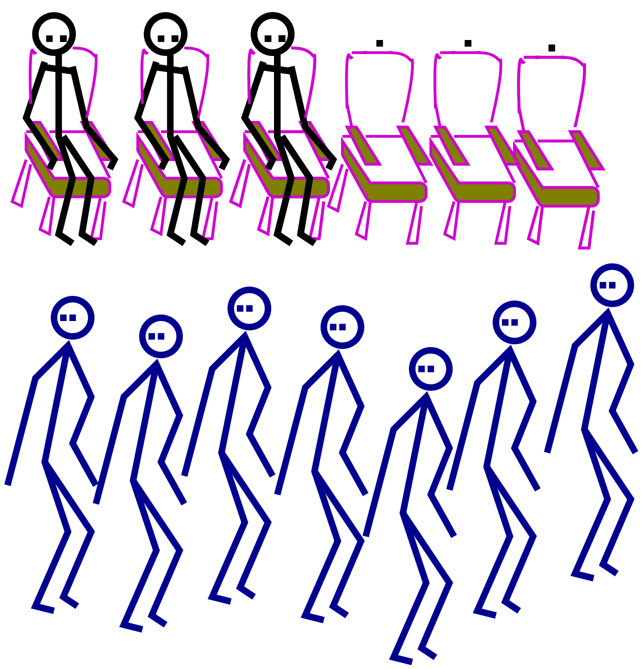


Figure 9: Competitive Routing

- 2 queues, single server.

- Poisson arrival process, rate λ.

Exponentially distributed service time, parameter μ; $\rho := \lambda/\mu$.

- Decisions: upon arrival, a customer observes the two queues and may purchase priority for a payment of an amount θ.

- The state: (i, j).

i:= number of high priority.

j:= number of low priority.

- Monotonicity: If for some strategy adopted by everybody, it is optimal for an individual to purchase priority at (i, j), then he must purchase priority at (r, j) for $r > j$.

- Lower dimensional state space: It follows that starting at $(0, 0)$ and playing optimally, there is some n such that the only reachable states are

$$(0, j), \ j \leq n, \quad \text{and } (i, n), \ i \geq 1.$$

- Indeed, due to monotonicity, if at some state $(0, m)$ it is optimal not to purchase priority, it is also optimal at states $(0, i)$, for $i \leq m$.
- Let $n-1$ be the largest such state.
- Then starting from $(0, 0)$ we go through states $(0, i)$, $i < n$, until $(0, n-1)$ is reached.
- At $(0, n)$ it is optimal to purchase priority. We then move to state $(1, n)$.
- The low priority queue does not decrease as long as there are high-priority customers.
- Due to monotonicity, it also does not increase as long as there are high-priority customers since at (i, n), $i \geq 1$ arrivals purchase priority! Therefore we remain at (i, n), as long as $i \geq 1$.

The Equilibrium

- Suppose that the customers in the population, except for a given individual, adopt a common threshold policy $[g]$. Then the optimal threshold for the individual is non-decreasing in g.

"Follow The Crowd" Behavior

- This implies Existence of an equilibrium
- No uniqueness! There may be up to

$$\left\lfloor \frac{1}{1-\rho} \right\rfloor$$

pure threshold Nash equilibria, as well as other mixed equilibria!

- Examples of multiple equilibria can be found at Hassin and Haviv's paper.

S-modular games, FTC, ATC

- D. Topkis, "Equilibrium points in nonzero-sum n-person submodular games", SIAM J. Contr. Optim., 17 (1979) 773-787.
- David D. Yao, "S-modular games with queueing applications", Queueing Systems 21 (1995) 449-475.

- Assume that the strategy space S_i of player i is a compact sublattice of R.
- Definition: The utility f_i for player i is supermodular iff
$$f_i(x \wedge y) + f_i(x \vee y) \geq f_i(x) + f_i(y).$$
- If f_i is twice differentiable then supermodularity is equivalent to
$$\frac{\partial^2 f_i(x)}{\partial x_1 \partial x_2} \geq 0.$$

Monotonicity of maximiers

- Let f be a supermodular function. Then the maximizer with respect to x_i is increasing in $x_j, j \neq i$.
- More preciesly, define

$$x_1^*(x_2) = argmax_{x_1} f(x_1, x_2).$$

Then $x_2 \leq x_2'$ implies $x_1^*(x_2) \leq x_1^*(x_2')$.

Monotonicity of the policy sets

- Consider 2 players. We allow S_i to depend on x_j

$$S_i = S_i(x_j), \quad i,j = 1,2, \quad i \neq j.$$

- Monotonicity of sublattices $A \prec B$ if for any $a \in A$ and $b \in B$,

$$a \wedge b \in A \qquad \text{and} \qquad a \vee b \in B.$$

- Monotonicity of policy sets We assume

$$x_j \leq x'_j \Longrightarrow S_i(x_j) \prec S_i(x'_j).$$

  This is called the Ascending Property. We define similarly the Descending Property.

- Lower semi continuity $x_1^k \to x_1^*$ and $x_2^* \in S_2(x_1^*)$ implies the existence of $\{x_2^k\}$ s.t. $x_2^k \in S_2(x_1^k)$ for each k, and $x_2^k \to x_2^*$.

Existence of Equilibria and Round Robin algorithms

Assume lower semi-continuity and compactness of the strategy sets.

- Supermodularity implies monotone convergence of the payoffs to an equilibrium. The monotonicity is in the same direction for all players. (We need the ascending property).
- Similarly with submodularity (for 2 players), but the monotonicity is in opposite directions. (We need the desceding property).
- In both cases, there need not be a unique equilibrium.
- Extensions to costs that are submodular in some components and supermodular in others. Extensions to vector policies.

Example of supermodularity: Qs in tandem

- A set of queues in tandem. Each queue has a server whose speed is controlled.
- The utility of each server rewards the througput and penalizes the delay.
- The players then have compatible incentives: if one speeds up, the other also want to speed up.

- Consider two queues in tandem with i.i.d. exponentially distributed service times with parameters μ_i, $i = 1, 2$. Let $\mu_i \leq u$ for some constant u.
- Server one has an infinite source of input jobs
- There is an infinite buffer between server 1 and 2.
- The throughput is given by $\mu_1 \wedge \mu_2$.
- The expected number of jobs in the buffer is given by
$$\frac{\mu_1}{\mu_2 - \mu_1}$$
when $\mu_1 < \mu_2$, and is otherwize infinite.

- Let
  - $p_i(\mu_1 \wedge \mu_2)$ be the profit of server i,
  - $c_i(\mu_i)$ be the operating cost,
  - $g(\cdot)$ be the inventory cost.
- The utilities of the players are

$$f_1(\mu_1, \mu_2) := p_1(\mu_1 \wedge \mu_2) - c_1(\mu_1) - g\left(\frac{\mu_1}{\mu_2 - \mu_1}\right)$$

$$f_2(\mu_1, \mu_2) := p_2(\mu_1 \wedge \mu_2) - c_2(\mu_2) - g\left(\frac{\mu_1}{\mu_2 - \mu_1}\right).$$

- The strategy spaces are

$$S_1(\mu_2) = \{\mu_1 : 0 \leq \mu_1 \leq \mu_2\},$$

$$S_2(\mu_1) = \{\mu_2 : \mu_1 \leq \mu_2 \leq u\}.$$

- If g is convex increasing then f_i are supermodular.

Example of submodularity: Flow Control

- There is a single queueing centre
- The rates of two input streams to the queueing centre are controlled by 2 players.
- Similar utilities as before.

More detailed example:

- Consider two input streams with Poisson arrivals with rates λ_1 and λ_2.
- The queueing center consists of c servers and no buffers. Each server has one unit of service rate.
- When all servers are occupied, an arrival is blcoked and lost.
- The blocking probability is given by the Erlang loss formula:
$$B(\lambda) = \frac{\lambda^c}{\lambda!} \left[\sum_{k=0}^{c} \frac{\lambda^k}{k!} \right]^{-1}$$
where $\lambda = \lambda_1 + \lambda_2$.
- Suppose user i maximimizes
$$f_i = r_i(\lambda_i) - c_i(\lambda B(\lambda)).$$
c_i is assumed to be convex increasing.

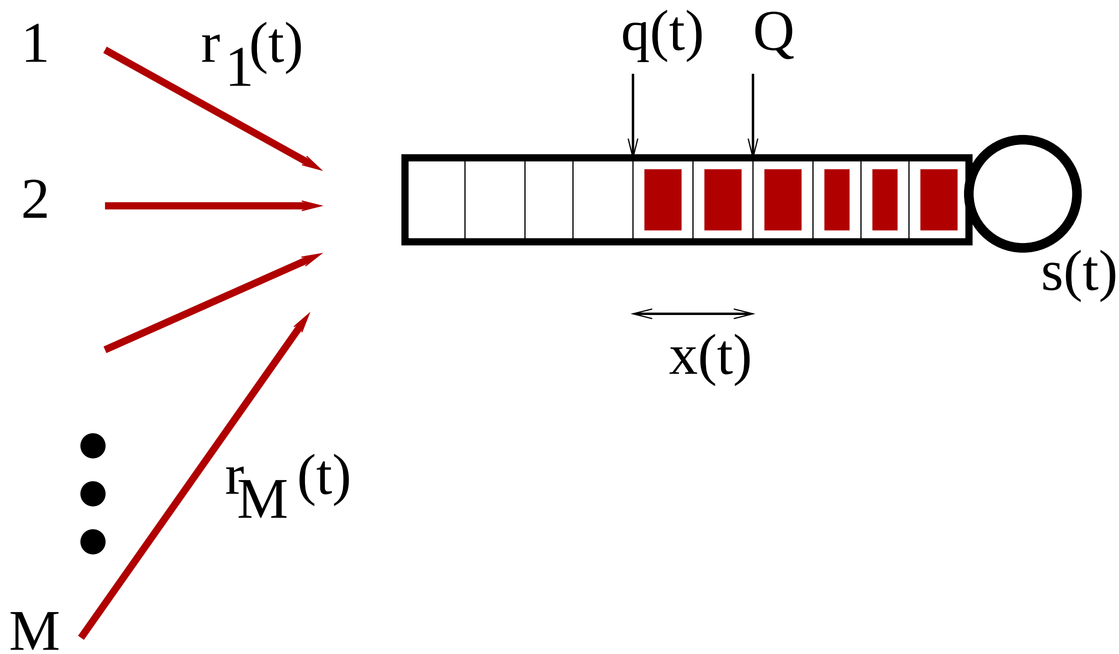
$\lambda B(\lambda)$ is the total loss rate.

Then f_i are submodular.

- Strategies: $\lambda_i \leq \overline{\lambda}$.
- Alternatively: $\lambda \leq \overline{\lambda}$. Then S_i satisfy the descending property.

Rate-based flow control: A Linear Quadratic Model

Ref: E. Altman and T. Başar, "Multi-user rate-based flow control", *IEEE Trans. on Communications*, pp. 940-949, 1998.



- M Users
- A single bottleneck queue
- Output rate: $s(t)$
- *Controlled* input rates: $r_m(t)$
- Queue length: $q(t)$. The target: Q. $x := q - Q$.

- Available bandwidth for user m is $a_m\, s(t)$, $\Sigma_m\, a_{m=1}^M = 1$.
- Define $u_m(t) := r_m(t) - a_m s(t)$.

The idealized dynamics:

$$\frac{dx}{dt} = \sum_{m=1}^{M} (r_m - a_m s) = \sum_{m=1}^{M} u_m \,, \qquad (2)$$

Policies and information:

$$u_m(t) = \mu_m(t, x_t), \qquad t \in [0, \infty)\,.$$

μ_m is piecewise continuous in its first argument, piecewise Lipschitz continuous in its second argument. The class of all such policies for user m is $\mathcal{U}_m$.

Objectives

– N1: the individual cost to be minimized by user m ($m \in \mathcal{M} = \{1, ..., M\}$) is

$$J_m^{N1}(u) = \int_0^\infty \left(|x(t)|^2 + \frac{1}{c_m} |u_m(t)|^2 \right) dt \,. \qquad (3)$$

– N2: the individual cost to be minimized by controller m ($m \in \mathcal{M}$) is

$$J_m^{N2}(u) = \int_0^\infty \left(\frac{1}{M} |x(t)|^2 + \frac{1}{c_m} |u_m(t)|^2 \right) dt \,. \qquad (4)$$

In case N2 the "effort" for keeping the deviations of the queue length from the desired value is split equally between the users.

Nash equilibria

We seek a multi-policy $\mu^* := (\mu_1^*, ..., \mu_M^*)$ such that no user has an incentive to deviate from, i.e.

$$J_m^{N1}(\mu^*) = \inf_{\mu_m \in \mathcal{U}_m} J_m^{N1}([\mu_m | \mu_{-m}^*]) \tag{5}$$

where $[\mu_m | \mu_{-m}^*]$ is the policy obtained when for each $j \neq m$, player j uses policy μ_j^*, and player m uses μ_m.

Similarly with J_2.

Main results

For case Ni $(i = 1, 2)$, there exists an equilibrium given by

$$\mu^*_{Ni,m}(x) = -\beta_m^{Ni} x \, , \; m = 1, ..., M,$$

where β_m^{Ni} is given by

$$\beta_m^{N1} = \overline{\beta}^{(N1)} - \sqrt{\overline{\beta}^{(N1)^2} - c_m}$$

where $\overline{\beta}^{(N1)} \stackrel{\text{def}}{=} \Sigma_{m=1}^{M} \beta_m^{N1}$, $i = 1, 2$, are the unique solutions of

$$\overline{\beta}^{(N1)} = \frac{1}{M-1} \sum_{m=1}^{M} \sqrt{(\overline{\beta}^{(N1)})^2 - c_m}$$

N2:

$$\beta_m^{N2} = \overline{\beta}^{(N2)} - \sqrt{\overline{\beta}^{(N2)^2} - \frac{c_m}{M}},$$

where $\overline{\beta}^{(N2)} \stackrel{\text{def}}{=} \Sigma_{m=1}^{M} \beta_m^{N2}$, $i = 1, 2$, are the unique solutions of

$$\overline{\beta}^{(N2)} = \frac{1}{M-1} \sum_{m=1}^{M} \sqrt{(\overline{\beta}^{(N2)})^2 - \frac{c_m}{M}} = \frac{\overline{\beta}^{(N1)}}{\sqrt{M}} .$$

Moreoevr,

$$\beta_m^{N1} = \beta_m^{N2}\sqrt{M}.$$

For each case, this is the unique equilibrium among stationary policies and is time-consistent.

The value

The costs accruing to user m, under the two Nash equilibria above, are given by

$$J_m^{N1}(\mu_{N1}^*) = \frac{\beta_m^{N1}}{c_m} x^2$$

and

$$J_m^{N2}(\mu_{N2}^*) = \frac{\beta_m^{N2}}{c_m} x^2 = \frac{1}{\sqrt{M}} J_m^{N1}(u_{N1}^*).$$

The Symmetric case

$c_m = c_j =: c$ for all $m, j \in \mathcal{M}$:

$$\beta_m^{N1} = \sqrt{\frac{c}{2M-1}},$$

and

$$\beta_m^{N2} = \sqrt{\frac{c}{M(2M-1)}}, \quad \forall m \in \mathcal{M};$$

The case of $M = 2$

General c_m's. we have for $m = 1, 2,\ j \neq m$,

$$\beta_m^{N1} = \left[-\frac{2c_j - c_m}{3} + 2\frac{\sqrt{c_1^2 - c_1 c_2 + c_2^2}}{3} \right]^{1/2}$$

$$\beta_m^{N2} = \frac{\beta_m^{N1}}{\sqrt{2}}.$$

If moreover, $c_1 = c_2 = c$ then

$$\beta_m^{N1} = \sqrt{c/3},$$

$$\beta_m^{N2} = \sqrt{c/6}.$$

Proof for (N1)

Choose a candidate solution

$$u_m^*(x) = -\beta_m x\,,\ m = 1, ..., M, \quad \text{where}$$

$$\beta_m = \overline{\beta} - \sqrt{\overline{\beta}^2 - c_m}$$

where $\overline{\beta} \stackrel{\text{def}}{=} \Sigma_{m=1}^M \beta_m$, are the unique solution of

$$\overline{\beta} = \frac{1}{M-1} \sum_{m=1}^{M} \sqrt{\overline{\beta}^2 - c_m}.$$

Fix u_j for $j \neq m$. Player m is faced with a LQ optimal control problem with the dynamics

$$dx/dt = u_m - \beta_{-m} x\,,\ \beta_{-m} = \sum_{j \neq m} \beta_j$$

and cost $J_m^{N1}(u)$ that is strictly convex in u_m. His optimal response:

$$u_m = -c_m P_m x,$$

where P_m is the unique positive solution of the Riccati equation

$$-2\beta_{-m} P_m - P_m^2 c_m + 1 = 0. \tag{6}$$

Denoting $\beta'_m = c_m P_m$, we obtain from (6)

$$\beta'_m = f_m(\beta_{-m}) \stackrel{\text{def}}{=} -\beta_{-m} + \sqrt{\beta_{-m}^2 + c_m}.$$

u is in equilibrium iff $\beta' = \beta$, or

$$\overline{\beta}^2 = \beta_{-m}^2 + c_m. \tag{7}$$

Hence

$$\beta_m = \overline{\beta} - \sqrt{\overline{\beta}^2 - c_m}.$$

Summing over $m \in \mathcal{M}$ we obtain

$$\Delta \stackrel{\text{def}}{=} \overline{\beta} - \frac{1}{M-1} \sum_{m=1}^{M} \sqrt{\overline{\beta}^2 - c_m} = 0$$

Uniqueness follows since

- Δ is strictly decreasing in $\overline{\beta}$ over the interval $[\max_m \sqrt{c_m}, \infty)$,
- it is positive at $\overline{\beta} = \max_m \sqrt{c_m}$ and
- it tends to $-\infty$ as $\beta \to \infty$. ∎

Greedy decentralized algorithms

Problem: Nash requires coordination, knowledge of individual utilities (c_m).

Solution: Try decentralized "best response" algorithms.

A greedy "best response" algorithm is defined by the following four conditions [Başar and Olsder, 1995]

- (i) Each user updates from time to time its policy by computing the best response against the most recently announced policies of the other users.
- (ii) The time between updates is sufficiently large, so that the control problem faced by a user when it updates its policy is well approximated by the original infinite horizon problem.
- (iii) The order of updates is arbitrary, but each user performs updates infinitely often.
- (iv) When the nth update occurs, a subset $K_n \subset \{1, \ldots, M\}$ of users simultaneously update their policies.

Proposed algorithms

– **Parallel update algorithm (PUA):** $K_n = \{1, ..., M\}$ for all n.

– **Round robin algorithm (RRA):** K_n is a singleton for all n and equals $(n+k)\text{mod}M + 1$, where k is an arbitrary integer.

– **Asynchronous algorithm (AA):** K_n is a singleton for all n and is chosen at random.

The initial policy used by each user is linear.
$\beta^{(n)}$:= value corresponding to the end of the nth iteration.
The optimal response at each step n:

$$\beta_m^{(n)} = \begin{cases} f_m(\beta_{-m}^{(n-1)}) & \text{if } m \in K_n \\ \\ \beta_m^{(n-1)} & \text{otherwise} \end{cases}, \qquad (8)$$

where

$$f_m(\beta_{-m}) \stackrel{\text{def}}{=} -\beta_{-m} + \sqrt{\beta^2_{-m} + c_m}. \qquad (9)$$

Convergence results

Consider PUA.
(i.a) Let $\beta_k^{(1)} = 0$ for all k. Then
- $\beta_k^{(2n)}$ monotonically decrease in n and
- $\beta_k^{(2n+1)}$ monotonically increase in n,

for every player k, and thus, the following limits exist:

$$\hat{\beta}_k \stackrel{\text{def}}{=} \lim_{n\to\infty} \beta_k^{(2n)}, \qquad \tilde{\beta}_k \stackrel{\text{def}}{=} \lim_{n\to\infty} \beta_k^{(2n+1)}.$$

(i.b) Assume that $\hat{\beta}_k = \tilde{\beta}_k$ (defined as above, with $\beta_k^{(1)} = 0$ for all k).
Consider now a different initial condition satisfying either

$$\beta_k^{(1)} \leq \beta_k \text{ for all } k,$$

(where β_k is the unique Nash) or

$$\beta_k^{(1)} \geq \beta_k \text{ for all } k.$$

Then for all k,

$$\lim_{n\to\infty} \beta_k^{(n)} = \beta_k.$$

Global convergence

If
– (ii.a) $M = 2$, and either

- $\beta_k^{(1)} \leq \beta_k$ for all k, or
- $\beta_k^{(1)} \geq \beta_k$ for all k;

or if
– (ii.b) $\beta_k^{(1)}$ and $c \stackrel{\text{def}}{=} c_k$ are the same for all k,

then $\beta^{(n)}$ converges to the unique equilibrium β^*.

Local convergence

For arbitrary c_k, there exists some neighborhood V of unique equilibrium β^* such that
if $\beta_k^{(1)} \in V$ then $\beta^{(n)}$ converges to the unique equilibrium β^*.

Numerical examples

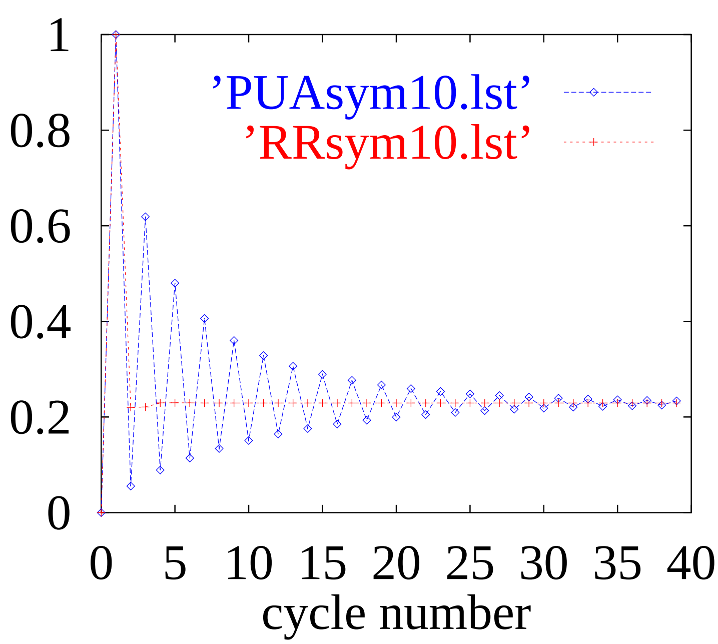


Figure 10: PUA versus RRA for $M = 10$

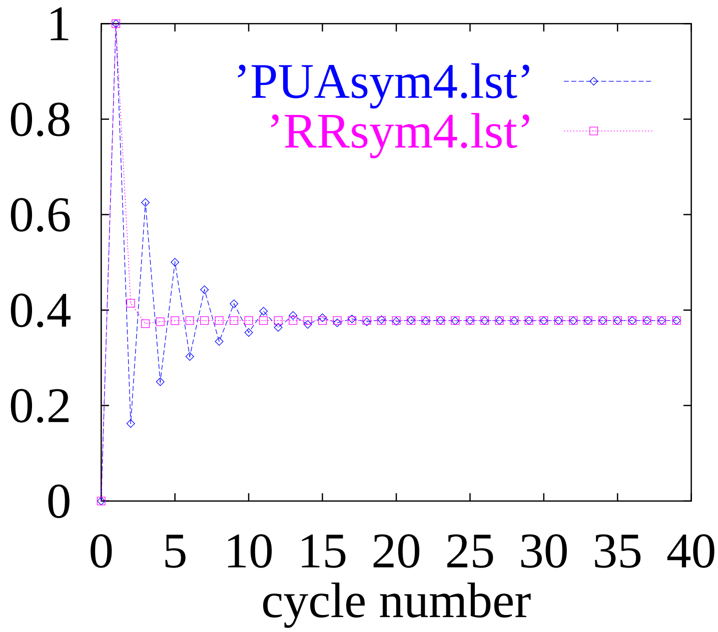


Figure 11: PUA versus RRA for $M = 4$

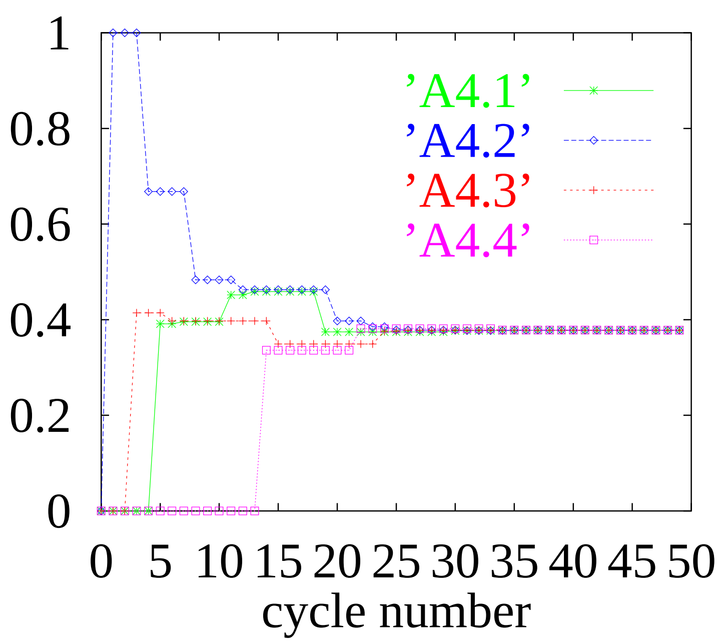


Figure 12: AA for $M = 4$

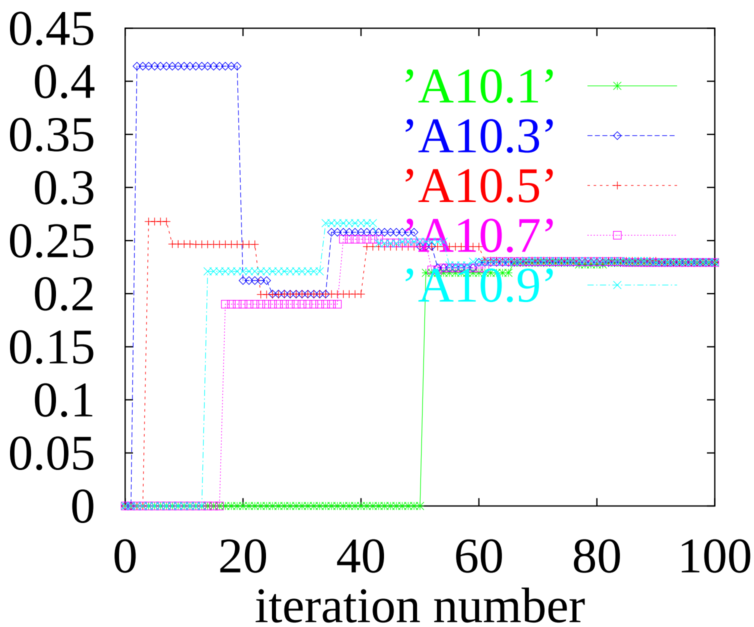


Figure 13: AA for $M = 10$