Today

MICHIGAN ENGINEERING
UNIVERSITY of MICHIGAN ■ COLLEGE of ENGINEERING

- Stopping time problems
- Medical application of stopping time problems

Example: Stopping Time Problems

MICHIGAN ENGINEERING
UNIVERSITY of MICHIGAN ■ COLLEGE of ENGINEERING

The following stopping time problem was proposed in 1875

Problem: An employer seeks to hire someone to fill a position. There are N candidates for the position. The candidates are interviewed sequentially. Upon completion of each interview the employer decides whether to offer the job to the current candidate. If no offer is made the individual seeks employment elsewhere and is no longer available for hire.

Stopping Time Problems

Stopping time problems arise when there is a recurring decision to "quit" a process

Basic idea: Each decision epoch the decision maker faces the decision to continue a decision process or quit a receive a one time reward

Example: Parking spot selection

Stopping Time Problems

Epochs: $t = 1, \ldots, N$

States: $S = S' \cup \{\Delta\}$, where Δ is an absorbing state

Actions: $\mathrm{A_s} = \begin{cases} \{C, Q\}, s \in S' \\ \{\emptyset\}, \quad s = \Delta \end{cases}$

Reward if you **C**ontinue

Reward if you **Q**uit

Rewards: $\mathrm{r_t(s, a)} = \begin{cases} f_t(s), s \in S', a = C \\ g_t(s), s \in S', a = Q \\ \quad 0, \qquad s = \Delta \end{cases}$, $\mathrm{r_N(s)} = g_N(s)$

Transition Probabilities

$$\mathrm{p_t}(j|s, a) = \begin{cases} p_t(j|s), & s \in S', j \in S', a = C \\ 1, & s \in S', j = \Delta, a = Q \text{ } or \text{ } s = j = \Delta \\ 0, & otherwise \end{cases}$$

Stopping Time Problems

MICHIGAN ENGINEERING
UNIVERSITY of MICHIGAN ■ COLLEGE of ENGINEERING

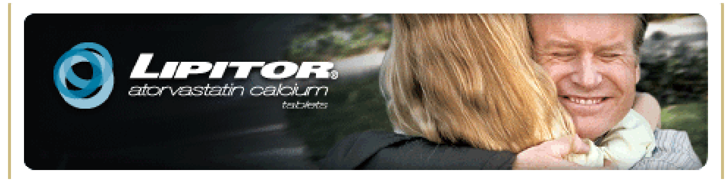
Optimality equations:

$$\mathrm{v_t(s)} = \max\left\{\mathrm{g_t(s), f_t(s,a)} + \lambda \sum_{\mathrm{j} \in S} p(j|s,C)v_{t+1}(j)\right\}, \forall s \in S$$

$$\mathrm{v}_N(\mathrm{s}) = \mathrm{g_N(s)}, \forall s \in S$$

Example: Statin Treatment

Some medical treatment decisions can be viewed as a "stopping time" problem:



- Statins lower your risk of heart attack and stroke
- Treatment has side effects and cost
- Patients receive annual cholesterol tests and make a decision to:
    - initiate statins
    - defer the decision for a year

Example: Treatment Initiation MDP

Optimality equations:

$$\mathrm{v_t(s)} = \max_{\mathrm{a}\in\{Q,C\}}\left\{\mathrm{g_t(s), r(s,C)} + \sum_{\mathrm{j}\in S} p(j|s,C)v_{t+1}(j)\right\}, \forall s \in S$$

$$\mathrm{v}_N(\mathrm{s}) = \mathrm{g_N(s)}, \forall s \in S$$

States are based on patient health status

Action C represents decision to defer statin initiation, Q denotes decision to start treatment

Reward $\mathrm{g}_t(s)$ is the expected survival for a patient that has initiated treatment

Absorbing state Δ represents health event that can be avoided by treatment or death from other causes

Markov Decision Process Model

MICHIGAN ENGINEERING
UNIVERSITY of MICHIGAN ■ COLLEGE of ENGINEERING

Stages:

- Time horizon: Ages 40-100
- Annual decision epochs

Decision:

- **Initiate** (Q) or **delay** (C) statin treatment

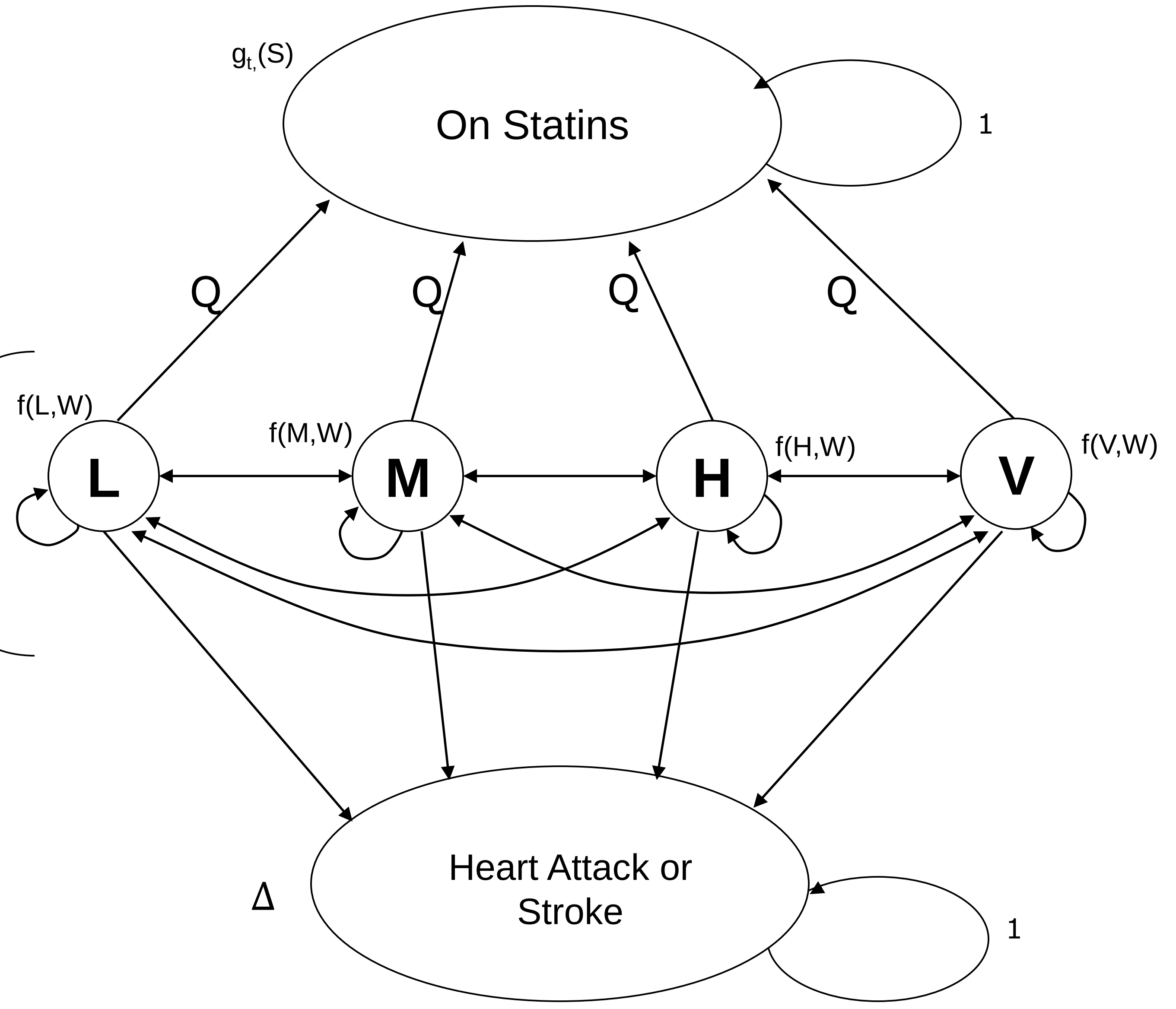
States:

- Metabolic: Total cholesterol and HDL (each can be L, M, H, V)
- Demographic: Gender, Race, BMI, smoking status, medical history

State Transition Diagram

Michigan Engineering
University of Michigan ■ College of Engineering

Metabolic States before an event has occurred.



Rewards

There are various types of reward functions used in health studies. The simplest definition for this problem is:

- $\mathrm{f_t}\ (s_t)$ is the time between decision epochs (e.g. 1 year)
- $g_t(s_t)$ is the expected future life years adjusted for quality of life on medication

Transition Probabilities

Transition probabilities between metabolic states:

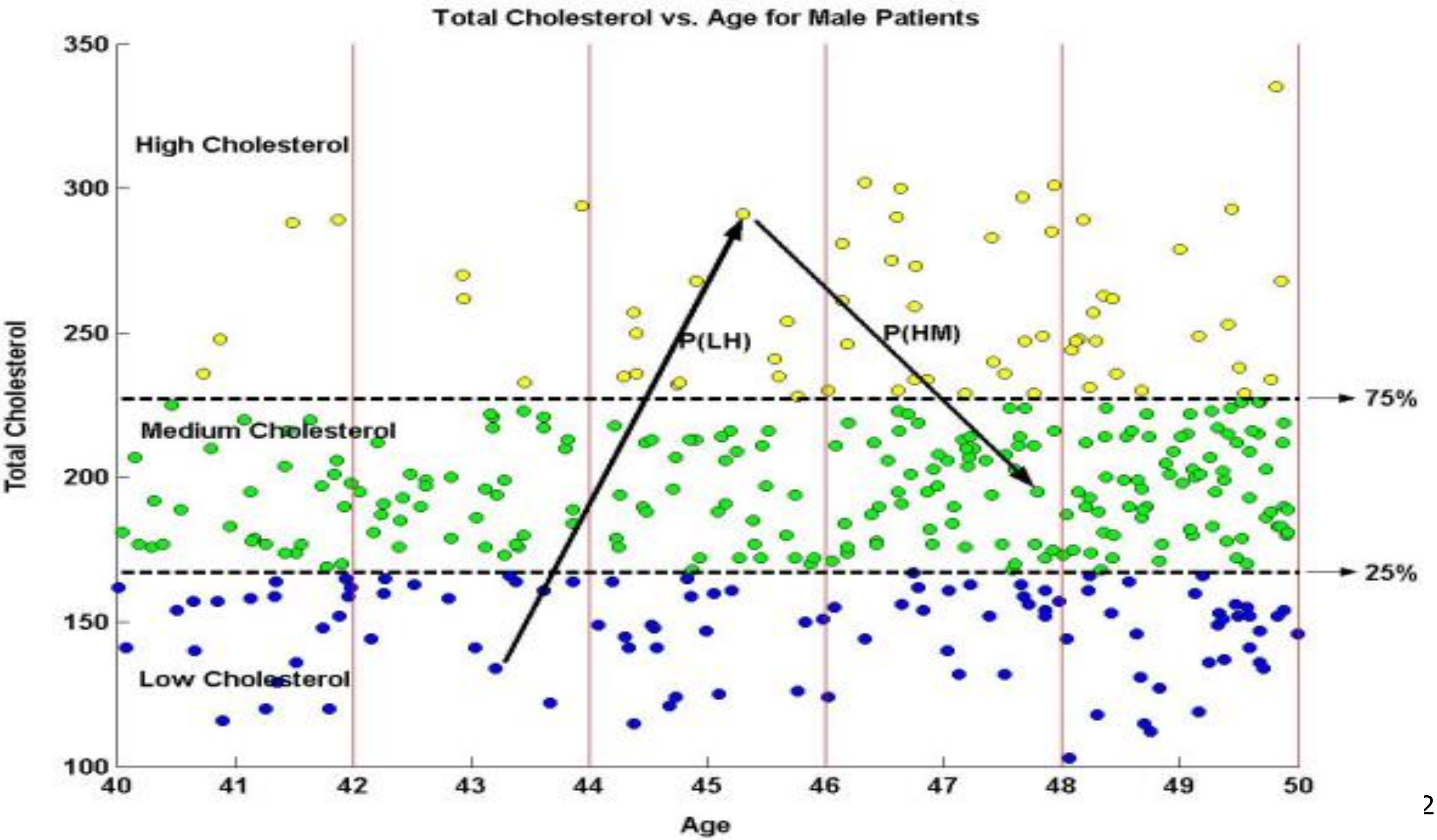
- Electronic medical record data for total cholesterol (bad cholesterol) and HDL (good cholesterol) levels for many patients

Transition probabilities from healthy states to complication state

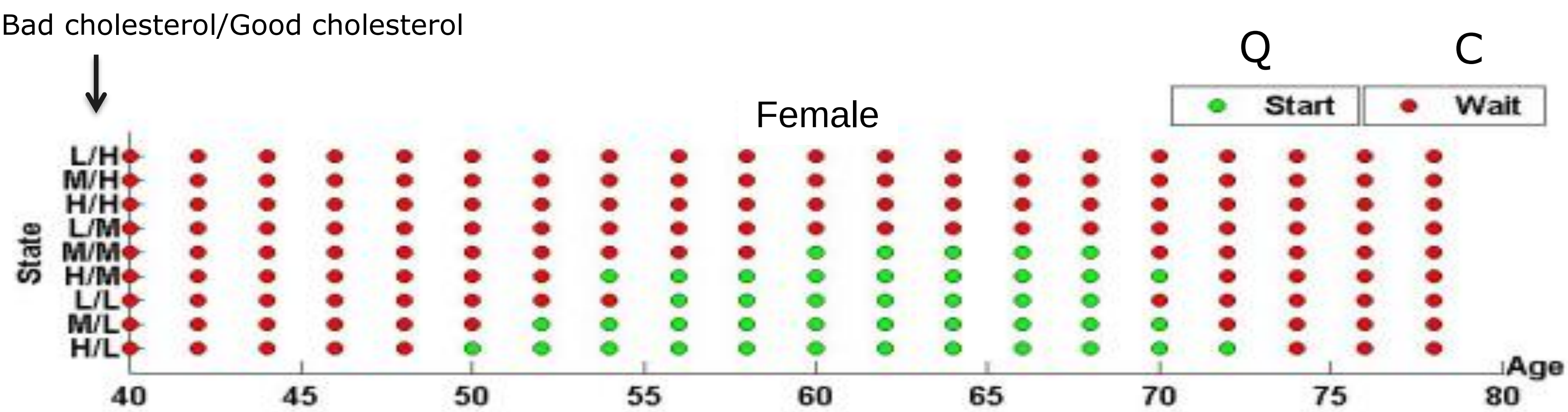
- Published cardiovascular risk models that estimate the probability of heart attack of stroke in the next year

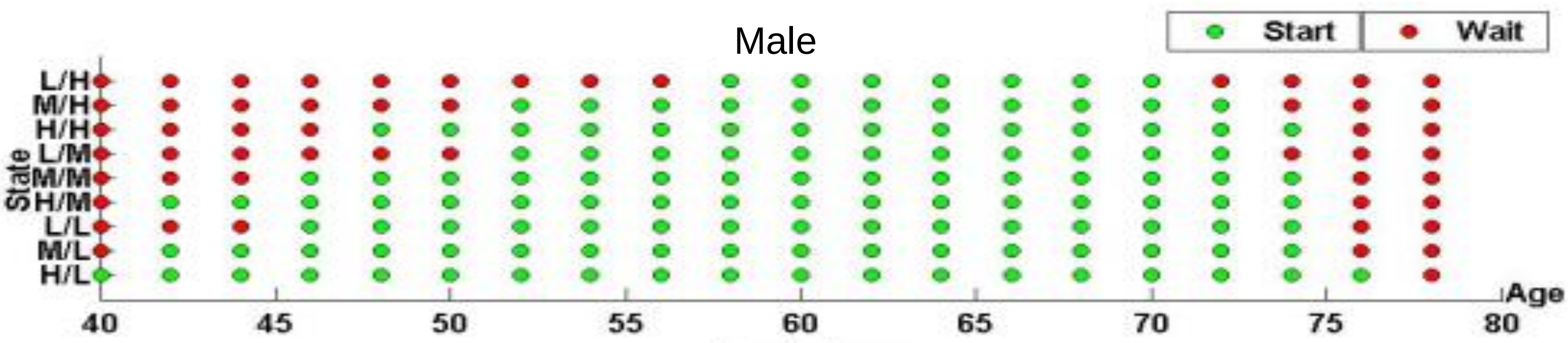
TP Example: Total Cholesterol

MICHIGAN ENGINEERING
UNIVERSITY of MICHIGAN ■ COLLEGE of ENGINEERING



Optimal Policy for Diabetic Patients





Follow-up

Learn more about these kinds of problems from the following papers in the background reading section on ctools:

- Kurt, M., Denton, B.T., Schaefer, A., Shah, N., Smith, S., "The Structure of Optimal Statin Initiation Policies for Patients with Type 2 Diabetes", IIE Transactions 1, 49-65, 2011

- Alagoz, O., Maillart, L.M., Schaefer, A.J., Roberts, M.S., "The Optimal Timing of Living-Donor Liver Transplantation," Management Science, 50(1), 1420-1430

Control Limit Policies

Optimality equations:

$$\mathrm{v_t(s)} = \max\left\{\mathrm{g_t(s)}, \mathrm{f_t(s,a)} + \lambda \sum_{\mathrm{j} \in S} p(j|s,C)v_{t+1}(j)\right\}, \forall s \in S$$

$$\mathrm{v}_N(\mathrm{s}) = \mathrm{g_N(s)}, \forall s \in S$$

Some stopping time problems have a **control limit** policy:

- Given certain conditions on the reward function and transition probabilities there may exist some ordering of states, S, and control limit, s^*, such that

$$\mathrm{a} = \begin{cases} C, if\ s \le s^* \\ Q, otherwise \end{cases}$$

Special Structured Policies

Policies with a simple structure like control limit policies are:

- Easier for decision makers to understand
- Easier to implement
- Easier to solve the associated MDPs

General structure of a **control limit policy**

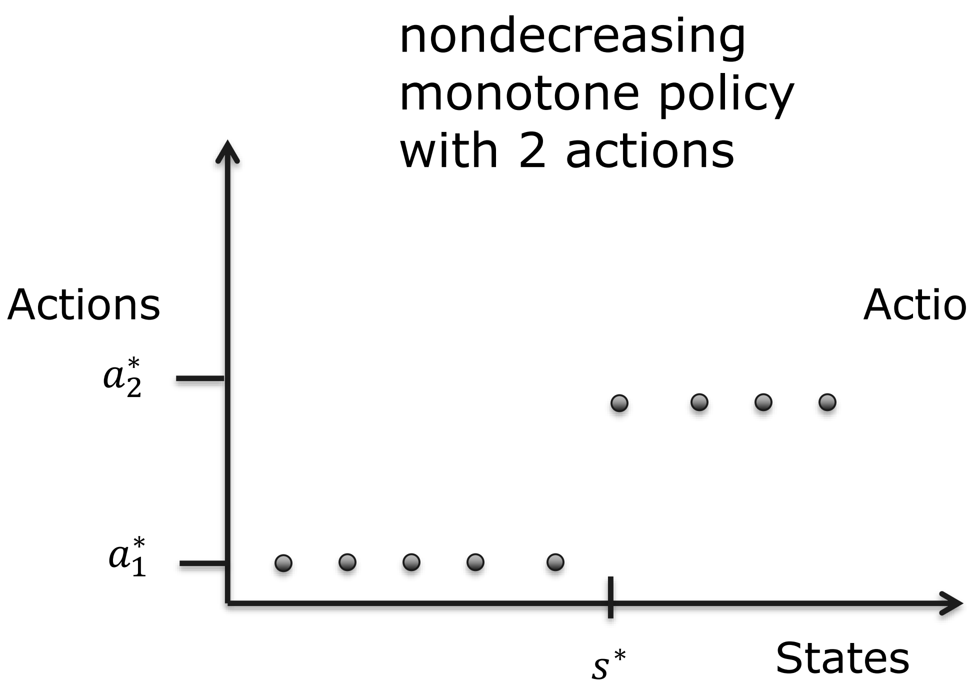
$$a_t(s_t) = \begin{cases} a_1, if\ s < s^* \\ a_2, if\ s \geq s^* \end{cases}$$

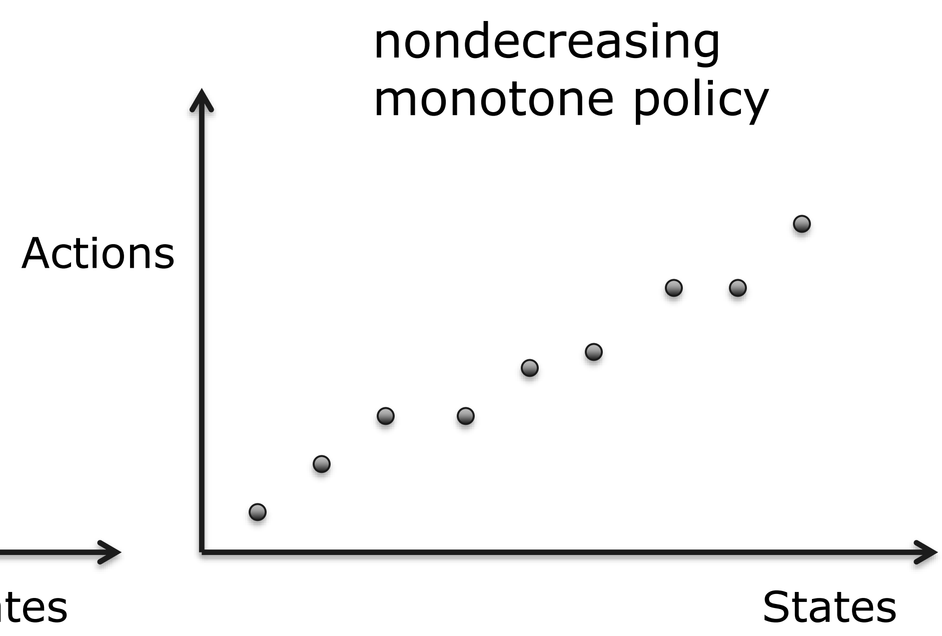
where a_1 and a_2 are alternative actions and s^* is a control limit.

Question: What conditions guarantee the existence of a control limit policy?

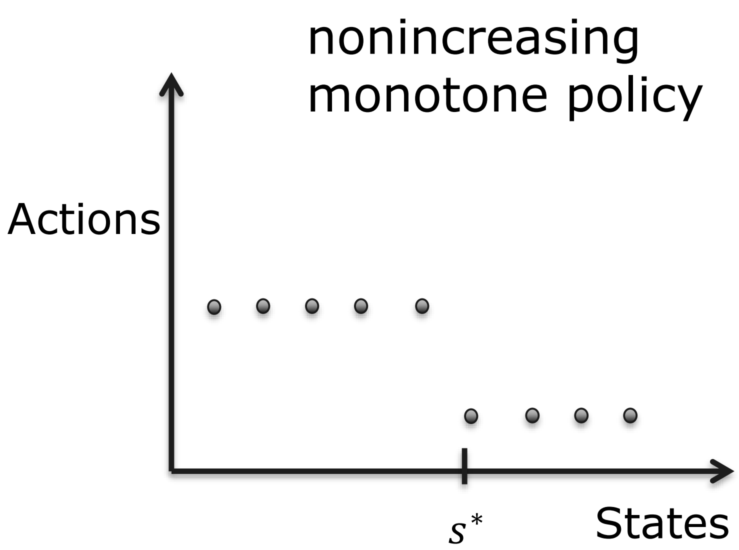
Examples of Monotone Policies

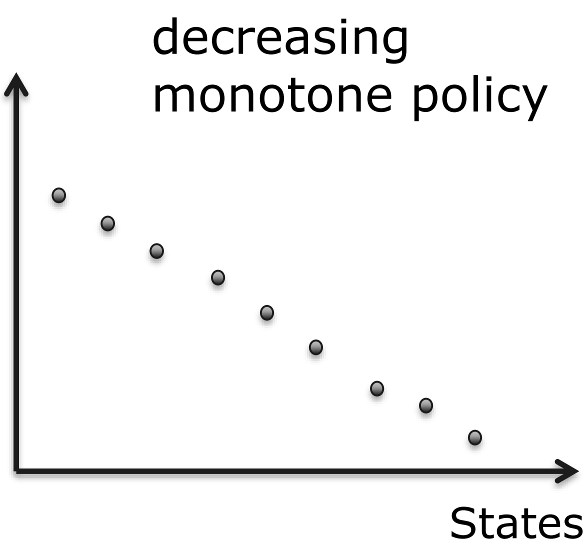
<u>Definition:</u> Control limit policies are examples of **monotone** policies. A policy is **monotone** if the *decision rule* at each stage is **nonincreasing** or **nondecreasing** with respect to a chosen ordering of the system states.



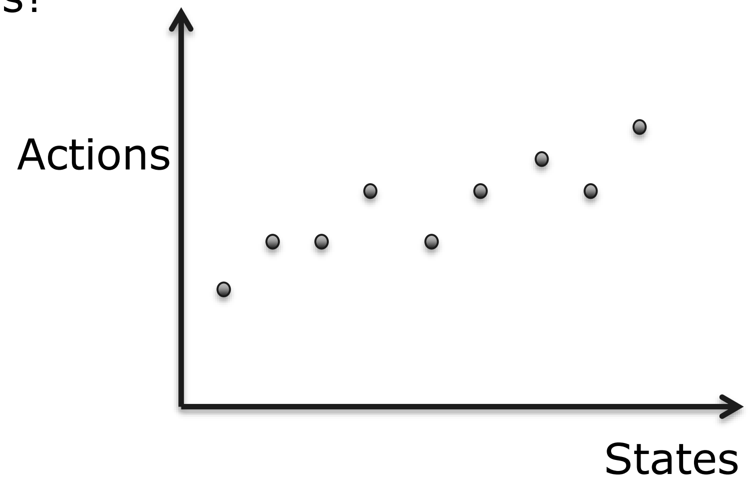


More Examples





Q. Which type is this?



Next Time

Michigan Engineering
University of Michigan ▪ College of Engineering

Next class we will:

- Define some useful concepts for analyzing MDPs
- Prove theorems about sufficient conditions for a monotone policy

Monotonicity: Optimal Policy

The following provides conditions under which the optimal policy is monotone.

Theorem (4.7.4 Puterman) Suppose for $t = 1, \ldots, N-1$

1. $r_t(s, a)$ is nondecreasing in s for all $a \in A$.

2. $q_t(k|s, a)$ is nondecreasing in s for all $k \in S, a \in A$.

3. $r_t(s, a)$ is **superadditive (subadditive)** on $S \times A$.

4. $q_t(k|s, a)$ is **superadditive (subadditive)** on $S \times A$, $\forall k$

5. $R_N(s)$ is nondecreasing in s.

Then there exist optimal decision rules, $d_t^*(s)$, which are nondecreasing (nonincreasing) in s for $t = 1, \ldots, N-1$.

Proof: See Puterman, p 107.

Concepts We Will Use

<u>Definition:</u> A function is **superadditive** if for $\mathrm{x}^+ \geq x^-$, where $\mathrm{x}^+, x^- \in \mathrm{X}$ and $\mathrm{y}^+ \geq y^-$, where , $y^+, y^- \in \mathrm{Y}$,

$$\mathrm{g}(\mathrm{x}^+,\mathrm{y}^+) + \mathrm{g}(\mathrm{x}^-,\mathrm{y}^-) \geq g(\mathrm{x}^+, y^-) + g(x^-, y^+)$$

If the reverse inequality holds then $\mathrm{g}(x, y)$ is **subadditive**.

<u>Definition:</u> A Markov chain has the *<u>IFR property</u>* if there is an ordering of states, $S \equiv \{1,2, \dots., n\}$, such that

$$q_t(k|s,a) = \sum_{j=k}^{n} p_t(j|s,a)$$

is nondecreasing in s for all k and a.

*Read about these concepts in Section 4.7 of Puterman