



"Optimal Control"

Sulav Duwal

PhD Candidate

System pharmacology and disease control group

Freie Universität Berlin

Bioinformatics Social Meeting, HU Berlin

Outline

Optimization and Optimal Control

- Optimization

- Optimal Control in Nutshell

Optimal Control for Markov Process

- Disease and drug model - Markov Process

- Optimal control of Markov Process

- Brute Force Algorithm

- Branch and bound Algorithm

Outline

Optimization and Optimal Control

- Optimization

- Optimal Control in Nutshell

Optimal Control for Markov Process

- Disease and drug model - Markov Process

- Optimal control of Markov Process

- Brute Force Algorithm

- Branch and bound Algorithm

Optimization

- What is optimization ?

Optimization

- ▶ What is optimization ?
- ▶ 3 Ingredients

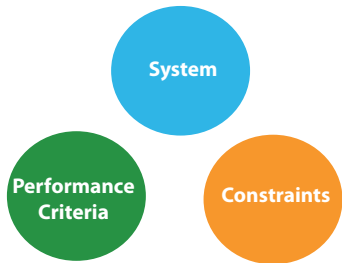


Figure: Ingredients of optimization

Optimization

► Example : Knapsack problem

- What is optimization ?
- 3 Ingredients

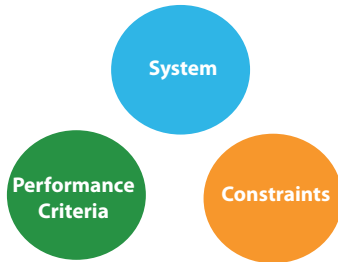


Figure: Ingredients of optimization

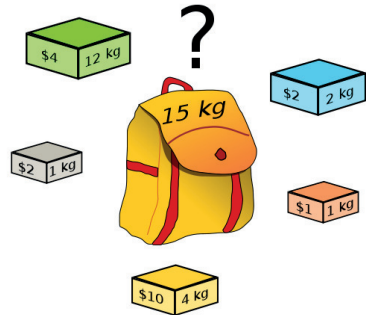


Figure: source:wikimedia commons

Optimization

► Example : Knapsack problem

- What is optimization ?
- 3 Ingredients

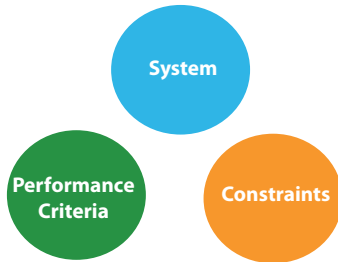


Figure: Ingredients of optimization

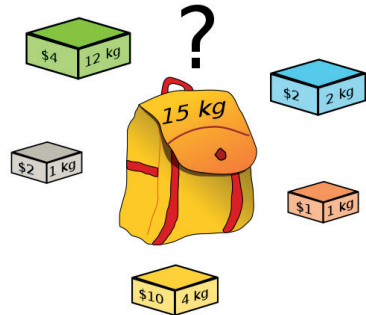
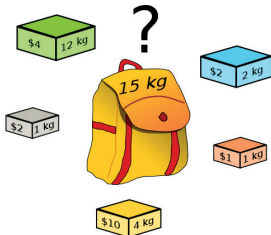


Figure: source:wikimedia commons

- **Description of system** : (weight , value)
- **Performance criteria** : Maximize total value
- **Constraints** : Total weights ≤ 15 Kg

Optimization and Optimal Control

- ▶ Static optimization
 - ▶ The system is **independent** of time.
 - ▶ Knapsack problem

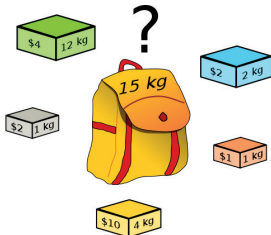


- ▶ Dynamic optimization
 - ▶ The system is **dependent** on time.

- ▶ Linear programming,
Integer programming

Optimization and Optimal Control

- ▶ Static optimization
 - ▶ The system is **independent** of time.
 - ▶ Knapsack problem

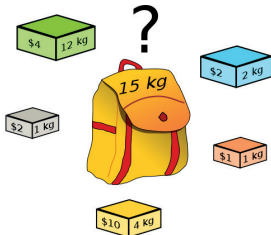


- ▶ Dynamic optimization
 - ▶ The system is **dependent** on time.
 - ▶ Minimum fuel problem

- ▶ Linear programming,
Integer programming

Optimization and Optimal Control

- ▶ Static optimization
 - ▶ The system is **independent** of time.
 - ▶ Knapsack problem



- ▶ Dynamic optimization
 - ▶ The system is **dependent** on time.
 - ▶ Minimum fuel problem
 - ▶ **A.k.a. Optimal Control**

- ▶ Linear programming,
Integer programming

Optimal Control in Nutshell

- System

$$\frac{dx(t)}{dt} = f(x(t), u(t)) \quad \text{and} \quad x(t_0) = x_0 \quad (1)$$

$x(t)$ is called **state variable** and $u(t)$ is called **control variable**.

- Performance criteria

$$J(x_0, u) = \int_{t_0}^{t_f} \mathcal{L}(x(t), u(t), t) dt + \phi(x(t_f), t_f) \quad (2)$$

$\mathcal{L}(x(t), u(t), t)$ is the **running cost** and $\phi(x(t_f), t_f)$ is the **terminal cost**.

Optimal Control in Nutshell

- System

$$\frac{dx(t)}{dt} = f(x(t), u(t)) \quad \text{and} \quad x(t_0) = x_0 \quad (1)$$

$x(t)$ is called **state variable** and $u(t)$ is called **control variable**.

- Performance criteria

$$J(x_0, u) = \int_{t_0}^{t_f} \mathcal{L}(x(t), u(t), t) dt + \phi(x(t_f), t_f) \quad (2)$$

$\mathcal{L}(x(t), u(t), t)$ is the **running cost** and $\phi(x(t_f), t_f)$ is the **terminal cost**.

- Optimal control

$$\begin{aligned} \min_u \quad & \int_{t_0}^{t_f} \mathcal{L}(x(t), u(t), t) dt + \phi(x(t_f), t_f) \\ \text{subject to} \quad & \frac{dx}{dt} = f(x(t), u(t)) \\ & x(t_0) = x_0 \end{aligned} \quad (3)$$

Optimal Control in Nutshell

- Optimal control

$$\begin{array}{ll} \min_u & \int_{t_0}^{t_f} \mathcal{L}(x(t), u(t), t) dt + \phi(x(t_f), t_f) \\ \text{subject to} & \frac{dx}{dt} = f(x(t), u(t)) \\ & x(t_0) = x_0 \end{array}$$

Optimal Control in Nutshell

- Optimal control

$$\begin{array}{ll} \min_u & \int_{t_0}^{t_f} \mathcal{L}(x(t), u(t), t) dt + \phi(x(t_f), t_f) \\ \text{subject to} & \frac{dx}{dt} = f(x(t), u(t)) \\ & x(t_0) = x_0 \end{array}$$

- Hamiltonian $H(\cdot)$ and adjoint equation $\lambda(t)$

$$H(x(t), u(t), \lambda(t), t) = \mathcal{L}(x(t), u(t), t) + \lambda(t) \cdot f(x(t), u(t)) \quad (4)$$

Optimal Control in Nutshell

- Optimal control

$$\begin{aligned} \min_u \quad & \int_{t_0}^{t_f} \mathcal{L}(x(t), u(t), t) dt + \phi(x(t_f), t_f) \\ \text{subject to} \quad & \frac{dx}{dt} = f(x(t), u(t)) \\ & x(t_0) = x_0 \end{aligned}$$

- Hamiltonian $H(\cdot)$ and adjoint equation $\lambda(t)$

$$H(x(t), u(t), \lambda(t), t) = \mathcal{L}(x(t), u(t), t) + \lambda(t) \cdot f(x(t), u(t)) \quad (4)$$

- Necessary condition of optimality

Optimal Control in Nutshell

► Optimal control

$$\begin{aligned} \min_u \quad & \int_{t_0}^{t_f} \mathcal{L}(x(t), u(t), t) dt + \phi(x(t_f), t_f) \\ \text{subject to} \quad & \frac{dx}{dt} = f(x(t), u(t)) \\ & x(t_0) = x_0 \end{aligned}$$

► Hamiltonian $H(\cdot)$ and adjoint equation $\lambda(t)$

$$H(x(t), u(t), \lambda(t), t) = \mathcal{L}(x(t), u(t), t) + \lambda(t) \cdot f(x(t), u(t)) \quad (4)$$

► Necessary condition of optimality

1. *Optimality Condition* at $u^* \Rightarrow$

$$\frac{\partial H}{\partial u} = 0$$

Optimal Control in Nutshell

► Optimal control

$$\begin{aligned} \min_u \quad & \int_{t_0}^{t_f} \mathcal{L}(x(t), u(t), t) dt + \phi(x(t_f), t_f) \\ \text{subject to} \quad & \frac{dx}{dt} = f(x(t), u(t)) \\ & x(t_0) = x_0 \end{aligned}$$

► Hamiltonian $H(\cdot)$ and adjoint equation $\lambda(t)$

$$H(x(t), u(t), \lambda(t), t) = \mathcal{L}(x(t), u(t), t) + \lambda(t) \cdot f(x(t), u(t)) \quad (4)$$

► Necessary condition of optimality

1. Optimality Condition at $u^* \Rightarrow$

$$\frac{\partial H}{\partial u} = 0$$

2. Adjoint equation

$$\frac{\partial H}{\partial x} = -\frac{d\lambda}{dt} \quad \text{and at } t = t_f \Rightarrow \lambda(t_f) = \frac{\partial \phi(x(t_f), t_f)}{\partial x}$$

Optimal Control in Nutshell

► Optimal control

$$\begin{aligned} \min_u \quad & \int_{t_0}^{t_f} \mathcal{L}(x(t), u(t), t) dt + \phi(x(t_f), t_f) \\ \text{subject to} \quad & \frac{dx}{dt} = f(x(t), u(t)) \\ & x(t_0) = x_0 \end{aligned}$$

► Hamiltonian $H(\cdot)$ and adjoint equation $\lambda(t)$

$$H(x(t), u(t), \lambda(t), t) = \mathcal{L}(x(t), u(t), t) + \lambda(t) \cdot f(x(t), u(t)) \quad (4)$$

► Necessary condition of optimality

1. Optimality Condition at $u^* \Rightarrow$

$$\frac{\partial H}{\partial u} = 0$$

2. Adjoint equation

$$\frac{\partial H}{\partial x} = - \frac{d\lambda}{dt} \quad \text{and at } t = t_f \Rightarrow \lambda(t_f) = \frac{\partial \phi(x(t_f), t_f)}{\partial x}$$

3. Dynamics of systems

$$\frac{\partial H}{\partial \lambda} = \frac{dx}{dt} = f(x(t), u(t)) \quad \text{and at } t = t_0 \Rightarrow x(t_0) = x_0$$

Optimal Control in Nutshell

► Necessary condition of optimality

1. *Optimality condition at u^* \Rightarrow*

$$\frac{\partial H}{\partial u} = 0$$

2. *Adjoint equation and Transversality equation*

$$\frac{\partial H}{\partial x} = -\frac{d\lambda}{dt} \quad \text{and at } t = t_f \Rightarrow \lambda(t_f) = \frac{\partial \phi(x(t_f), t_f)}{\partial x}$$

3. *Dynamics of systems*

$$\frac{\partial H}{\partial \lambda} = \frac{dx}{dt} = f(x(t), u(t)) \quad \text{and at } t = t_0 \Rightarrow x(t_0) = x_0$$

Optimal Control in Nutshell

► Necessary condition of optimality

1. *Optimality condition at $u^* \Rightarrow$*

$$\frac{\partial H}{\partial u} = 0$$

2. *Adjoint equation and Transversality equation*

$$\frac{\partial H}{\partial x} = -\frac{d\lambda}{dt} \quad \text{and at } t = t_f \Rightarrow \lambda(t_f) = \frac{\partial \phi(x(t_f), t_f)}{\partial x}$$

3. *Dynamics of systems*

$$\frac{\partial H}{\partial \lambda} = \frac{dx}{dt} = f(x(t), u(t)) \quad \text{and at } t = t_0 \Rightarrow x(t_0) = x_0$$

► *Optimal control problem is a **two-point boundary value problem**.*

A short recap

- ▶ Optimization - **system**, **performance Criteria**, **constraints**
- ▶ Dynamic optimization a.k.a **optimal control**
- ▶ Hamiltonian, Adjoint equation

A short recap

- ▶ Optimization - **system**, **performance Criteria**, **constraints**
- ▶ Dynamic optimization a.k.a **optimal control**
- ▶ Hamiltonian, Adjoint equation
- ▶ A broad field with history of 300 years. The theory was developed in parallel in USA and former USSR during cold war.

A short recap

- ▶ Optimization - **system**, **performance Criteria**, **constraints**
- ▶ Dynamic optimization a.k.a **optimal control**
- ▶ Hamiltonian, Adjoint equation
- ▶ A broad field with history of 300 years. The theory was developed in parallel in USA and former USSR during cold war.
- ▶ Application in biology : control of epidemics, treatment design etc. largely underexplored !!

Outline

Optimization and Optimal Control

- Optimization

- Optimal Control in Nutshell

Optimal Control for Markov Process

- Disease and drug model - Markov Process

- Optimal control of Markov Process

- Brute Force Algorithm

- Branch and bound Algorithm



RESEARCH ARTICLE

Optimal Treatment Strategies in the Context of ‘Treatment for Prevention’ against HIV-1 in Resource-Poor Settings

Sulav Duwal^{1,2}, Stefanie Winkelmann¹, Christof Schütte^{1,3}, Max von Kleist^{1,2*}

1 Department of Mathematics and Computer Science, Freie Universität Berlin, Germany, **2** Junior Research Group “Systems Pharmacology & Disease Control”, **3** Zuse Institute Berlin, Germany

Disease and drug model - Markov Process

- ▶ Stochastic process with memorylessness
- ▶ State space \mathcal{S} : The set of all possible states
- ▶ Action space \mathcal{A} : The set of all possible actions (treatment option).

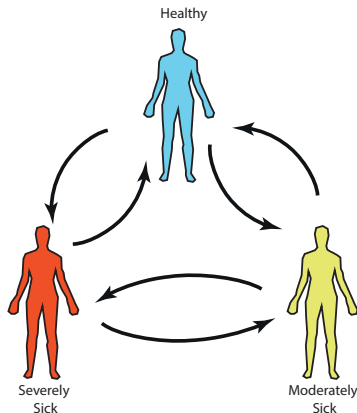


Figure: Markov Jump Process

Disease and drug model - Markov Process

- ▶ Stochastic process with memorylessness
- ▶ State space \mathcal{S} : The set of all possible states
 - ▶ $\mathcal{S} = \{Healthy, Mod.Sick, Severe.Sick\}$ and $X_t \in \mathcal{S}$
- ▶ Action space \mathcal{A} : The set of all possible actions (treatment option).

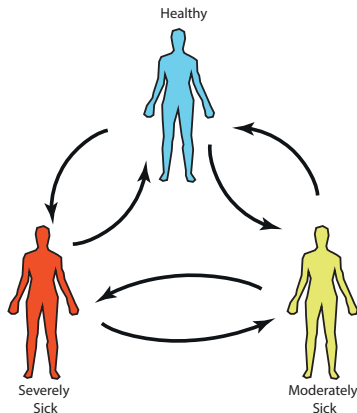


Figure: Markov Jump Process

Disease and drug model - Markov Process

- ▶ Stochastic process with memorylessness
- ▶ State space \mathcal{S} : The set of all possible states
 - ▶ $\mathcal{S} = \{Healthy, Mod.Sick, Severe.Sick\}$ and $X_t \in \mathcal{S}$
 - ▶ $p \in \Omega$ denote probability distribution vector on the state space \mathcal{S}
$$p[x](t) := \mathbb{P}(X_t = x)$$
- ▶ Action space \mathcal{A} : The set of all possible actions (treatment option).

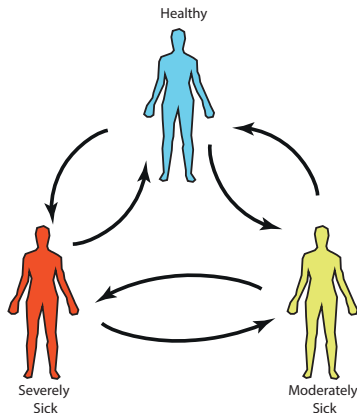


Figure: Markov Jump Process

Disease and drug model - Markov Process

- ▶ Stochastic process with memorylessness
- ▶ State space \mathcal{S} : The set of all possible states
 - ▶ $\mathcal{S} = \{\text{Healthy}, \text{Mod.Sick}, \text{Severe.Sick}\}$ and $X_t \in \mathcal{S}$
 - ▶ $p \in \Omega$ denote probability distribution vector on the state space \mathcal{S}
$$p[x](t) := \mathbb{P}(X_t = x)$$
- ▶ Action space \mathcal{A} : The set of all possible actions (treatment option).
 - ▶ Each action $a \in \mathcal{A}$ induces unique dynamics.

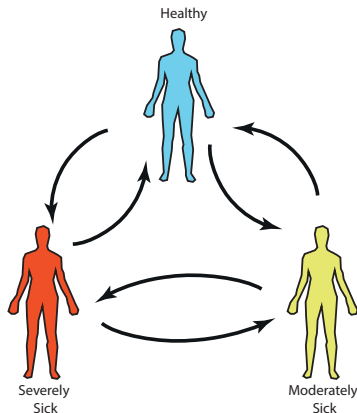


Figure: Markov Jump Process

Disease and drug model - Markov Process

- ▶ Stochastic process with memorylessness
- ▶ State space \mathcal{S} : The set of all possible states
 - ▶ $\mathcal{S} = \{Healthy, Mod.Sick, Severe.Sick\}$ and $X_t \in \mathcal{S}$
 - ▶ $p \in \Omega$ denote probability distribution vector on the state space \mathcal{S}

$$p[x](t) := \mathbb{P}(X_t = x)$$

- ▶ Action space \mathcal{A} : The set of all possible actions (treatment option).
 - ▶ Each action $a \in \mathcal{A}$ induces unique dynamics.
 - ▶ For a given time interval τ

$$p(t + \tau) = T_a \cdot p(t)$$

where T_a is the transition matrix specific to action a .

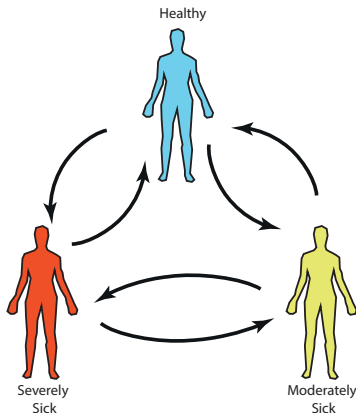


Figure: Markov Jump Process

Disease and drug model - Markov Process

- ▶ Stochastic process with memorylessness
- ▶ State space \mathcal{S} : The set of all possible states
 - ▶ $\mathcal{S} = \{\text{Healthy}, \text{Mod.Sick}, \text{Severe.Sick}\}$ and $X_t \in \mathcal{S}$
 - ▶ $p \in \Omega$ denote probability distribution vector on the state space \mathcal{S}

$$p[x](t) := \mathbb{P}(X_t = x)$$

- ▶ Action space \mathcal{A} : The set of all possible actions (treatment option).
 - ▶ Each action $a \in \mathcal{A}$ induces unique dynamics.
 - ▶ For a given time interval τ

$$p(t + \tau) = T_a \cdot p(t)$$

where T_a is the transition matrix specific to action a .

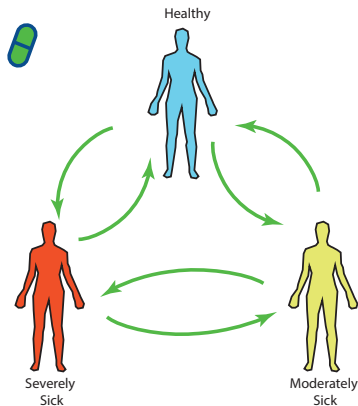


Figure: Markov Jump Process

Disease and drug model - Markov Process

- ▶ Stochastic process with memorylessness
- ▶ State space \mathcal{S} : The set of all possible states

- ▶ $\mathcal{S} = \{\text{Healthy}, \text{Mod.Sick}, \text{Severe.Sick}\}$ and $X_t \in \mathcal{S}$
- ▶ $p \in \Omega$ denote probability distribution vector on the state space \mathcal{S}

$$p[x](t) := \mathbb{P}(X_t = x)$$

- ▶ Action space \mathcal{A} : The set of all possible actions (treatment option).

- ▶ Each action $a \in \mathcal{A}$ induces unique dynamics.
- ▶ For a given time interval τ

$$p(t + \tau) = T_a \cdot p(t)$$

where T_a is the transition matrix specific to action a .

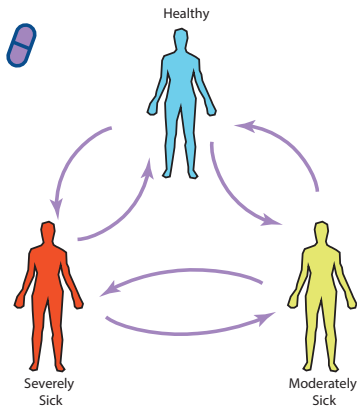
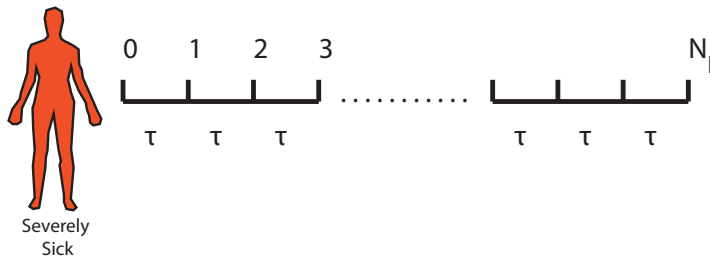


Figure: Markov Jump Process

Optimal control of Markov Process



- System : Discrete Dynamics

$$p_{j+1} = T_{u_j} \cdot p_j \text{ and } p(0) = p_0$$

where $u_j \in \mathcal{A}$

- Performance Criteria

$N_I :=$ is a number of intervals and $q_{u_j,j}, q_{N_I} \in \mathbb{R}^{|S|}$ are cost vectors.

$$J(p_0, u) = \sum_{j=0}^{N_I-1} q'_{u_j,j} \cdot p_j + q'_{N_I} \cdot p_{N_I}$$

Optimal control of Markov Process

► Optimal Control Problem

$$\begin{array}{ll} \underset{u \in \mathcal{U}}{\operatorname{argmin}} & \sum_{j=0}^{N_{\mathcal{I}}-1} q'_{u_j,j} \cdot p_j + q'_{N_{\mathcal{I}}} \cdot p_{N_{\mathcal{I}}} \\ \text{w.r.t} & p_{j+1} = T_{u_j} \cdot p_j \\ & p(0) = p_0 \end{array}$$

Optimal control of Markov Process

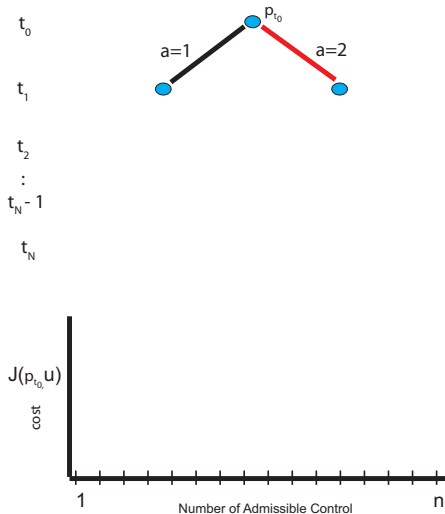
► Optimal Control Problem

$$\begin{aligned} \underset{u \in \mathcal{U}}{\operatorname{argmin}} \quad & \sum_{j=0}^{N_{\mathcal{I}}-1} q'_{u,j} \cdot p_j + q'_{N_{\mathcal{I}}} \cdot p_{N_{\mathcal{I}}} \\ \text{w.r.t} \quad & p_{j+1} = T_{u_j} \cdot p_j \\ & p(0) = p_0 \end{aligned}$$

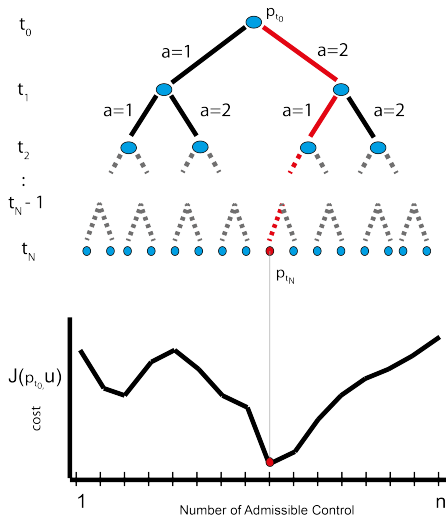
► Hamiltonian

$$H_j = \xi'_{j+1} \cdot T_{u_j} \cdot p_j + q'_{u,j} \cdot p_j$$

Brute Force Algorithm



Brute Force Algorithm

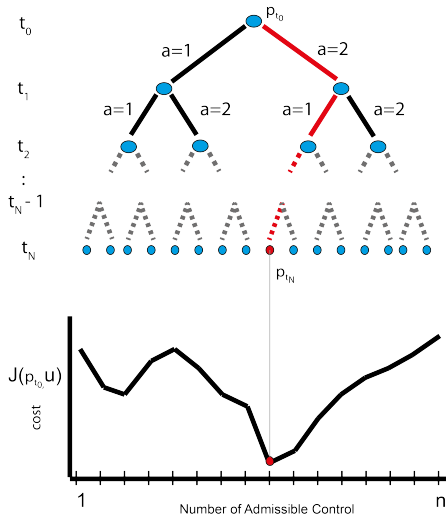


Brute Force Algorithm

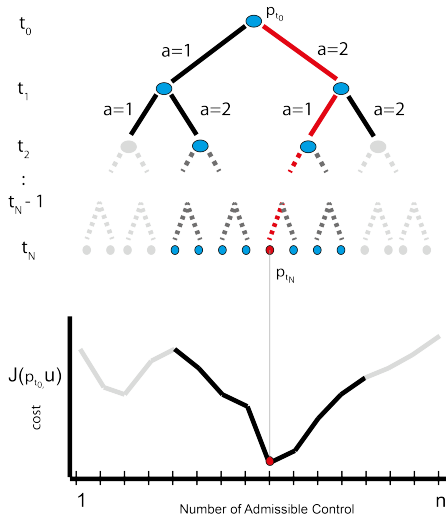
- Problem :
- ▶ Curse of dimensionality
 - ▶ Given $|\mathcal{A}|$ number of action options and $N_{\mathcal{I}}$ number of intervals
 $|\mathcal{A}|^{N_{\mathcal{I}}}$ number of controls available

Idea : ??

Branch and bound algorithm (Redundancy)



Branch and bound algorithm (Redundancy)



Santa's problem

Santa's problem

- ▶ 7 days before Christmas, Santa lost his beard except 1 strand.
- ▶ In 7 days, he has to grow exactly 999 strands. Otherwise

Santa's problem

- ▶ 7 days before Christmas, Santa lost his beard except 1 strand.
- ▶ In 7 days, he has to grow exactly 999 strands. Otherwise
- ▶ He has 2 options

Santa's problem

- ▶ 7 days before Christmas, Santa lost his beard except 1 strand.
- ▶ In 7 days, he has to grow exactly 999 strands. Otherwise
- ▶ He has 2 options
 - ▶ $a_1 : +3$ Naturally, he has +3 strands more in a day.

Santa's problem

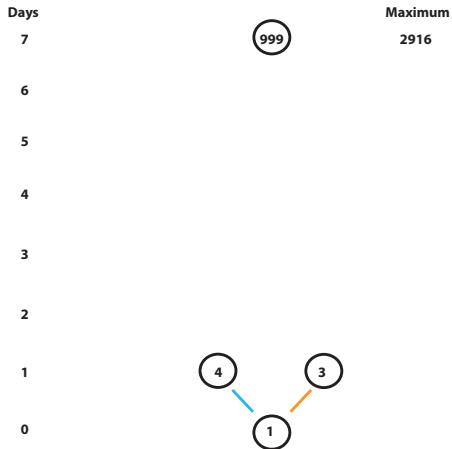
- ▶ 7 days before Christmas, Santa lost his beard except 1 strand.
- ▶ In 7 days, he has to grow exactly 999 strands. Otherwise
- ▶ He has 2 options
 - ▶ $a_1 : +3$ Naturally, he has +3 strands more in a day.
 - ▶ $a_2 : \times 3$ If he takes a magic drug, he multiplies his beard 3 times in a day.

Santa's problem

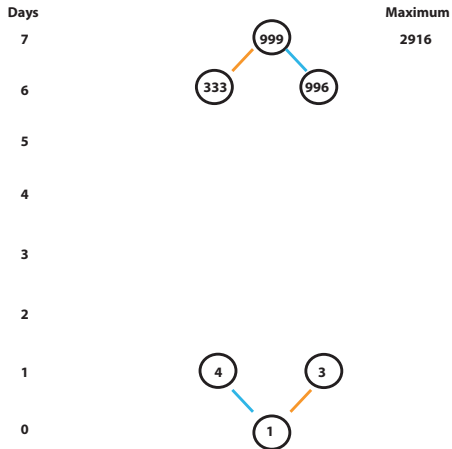
- ▶ 7 days before Christmas, Santa lost his beard except 1 strand.
- ▶ In 7 days, he has to grow exactly 999 strands. Otherwise
- ▶ He has 2 options
 - ▶ $a_1 : +3$ Naturally, he has +3 strands more in a day.
 - ▶ $a_2 : \times 3$ If he takes a magic drug, he multiplies his beard 3 times in a day.
- ▶ How can he get 999 beard in 7 days starting from a single beard ?

Santa's problem

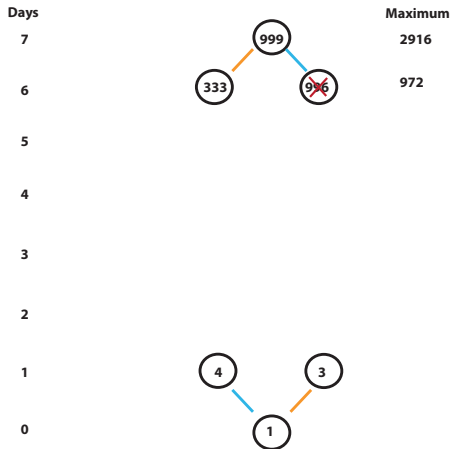
Santa's problem



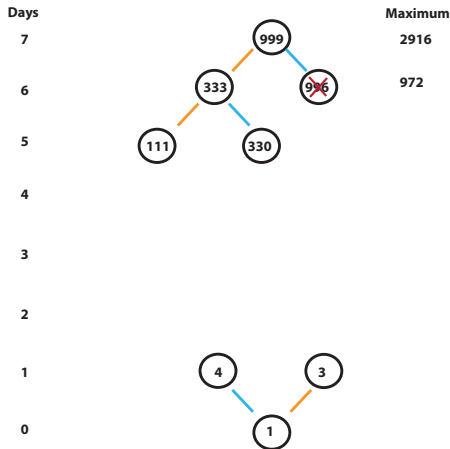
Santa's problem



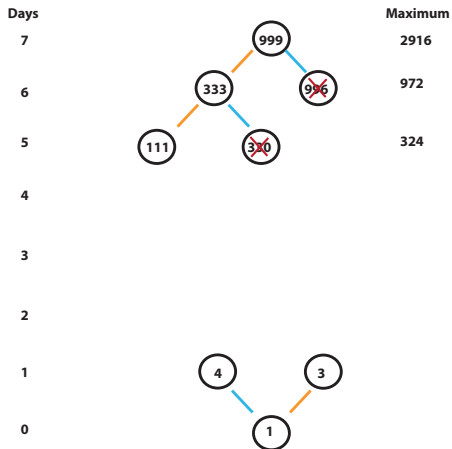
Santa's problem



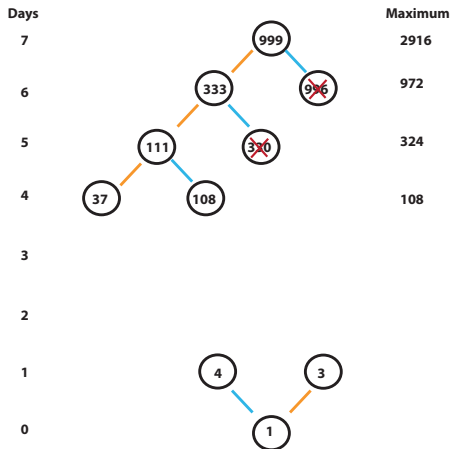
Santa's problem



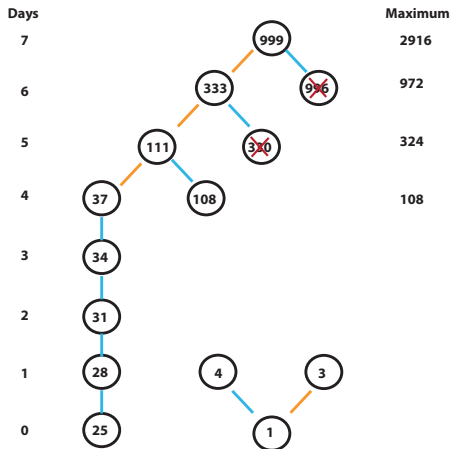
Santa's problem



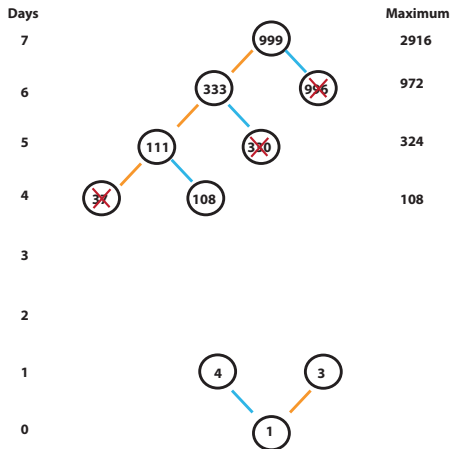
Santa's problem



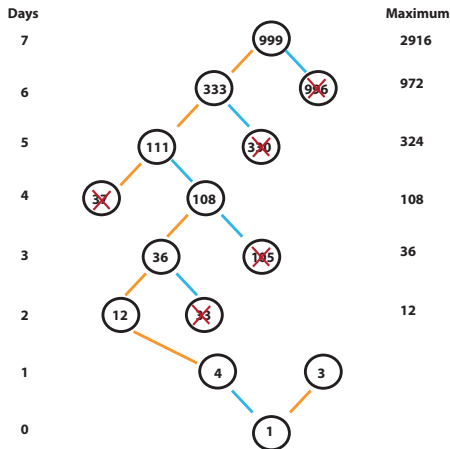
Santa's problem



Santa's problem



Santa's problem



Santa's problem

Dear Santa,
your drug schedule

- ▶ Day 1 no drug
- ▶ Day 2 magic drug
- ▶ Day 3 magic drug
- ▶ Day 4 magic drug
- ▶ Day 5 no drug
- ▶ Day 6 magic drug
- ▶ Day 7 magic drug

and you will have exactly 999 strands of beard.
Merry Christmas

Optimal Control for Markov Process

► Redundancy Test - Linear programming

$$\begin{aligned}
 \mu = \min_{\rho, \pi, \zeta} & \begin{bmatrix} \xi'_{j,1} & 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} \rho \\ \pi \\ \zeta \end{bmatrix} \\
 \text{w.r.t} & \begin{bmatrix} \xi'_{j,2} & 1 & -1 \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \\ \xi'_{j,n} & 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} \rho \\ \pi \\ \zeta \end{bmatrix} \geq \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \\
 & \sum_i \rho[i] = 1 \\
 & p_{min} \leq \rho \leq p_{max} \\
 & 0 \leq \pi \leq J_{max} \\
 & J_{min} \leq \zeta \leq J_{max},
 \end{aligned} \tag{5}$$

where p_{min}, p_{max} and J_{min}, J_{max} are lower and upper bounds for p_m^* and $J^*(p_0, u^*(p_0))$ respectively.

► if $\mu > 0$, the adjoint vector $\xi_{j,1}$ is redundant.



RESEARCH ARTICLE

Optimal Treatment Strategies in the Context of 'Treatment for Prevention' against HIV-1 in Resource-Poor Settings

Sulav Duwal^{1,2}, Stefanie Winkelmann¹, Christof Schütte^{1,3}, Max von Kleist^{1,2*}

¹ Department of Mathematics and Computer Science, Freie Universität Berlin, Germany, ² Junior Research Group "Systems Pharmacology & Disease Control", ³ Zuse Institute Berlin, Germany





Codes can be downloaded from <https://github.com/SulavDuwal/OptimalTreatmentStrategies>



@SulavDuwal

References

-  Suzanne Lenhart, John T. Workman
Optimal control applied to biological models
Book : Chapman & Hall/CRC Mathematical and Computational Biology
-  Robert F. Stengel
Optimal Control and Estimation
Dover Books on Mathematics