

Tomada de Decisão Sequencial

Diferentes critérios de decisão

Valdinei Freire, Karina Delgado

Señoritas Courier: Bike-entrega por mulheres



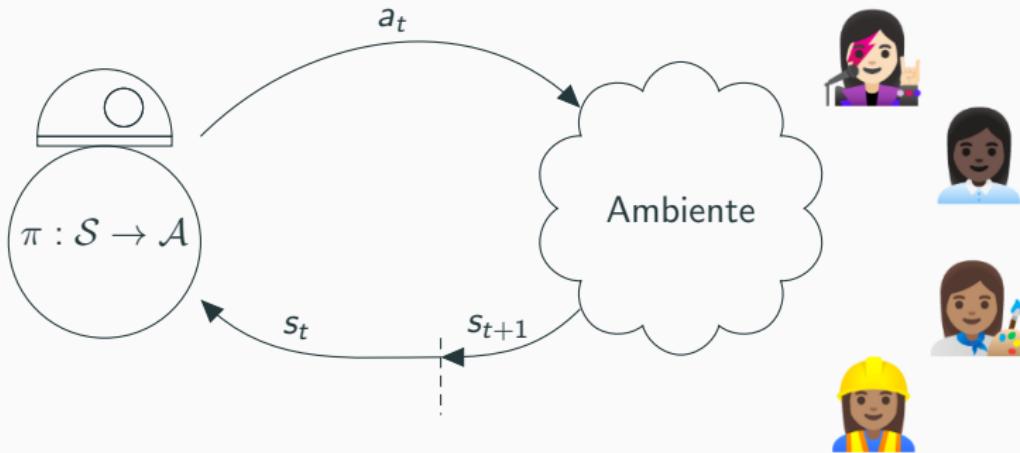
- Entregadoras utilizando bicicleta
- Entregadoras tem restrições: disponibilidade, área de entrega, capacidade.
- Vários pacotes a serem entregues: local, data, peso.
- **Decisões**
 - Qual preço cobrar?
 - Como distribuir pacotes para entregadoras?
 - Quais rotas seguir?
- **Objetivo:** maximizar o lucro de todas entregadoras.

Desafios para Tomada de Decisão



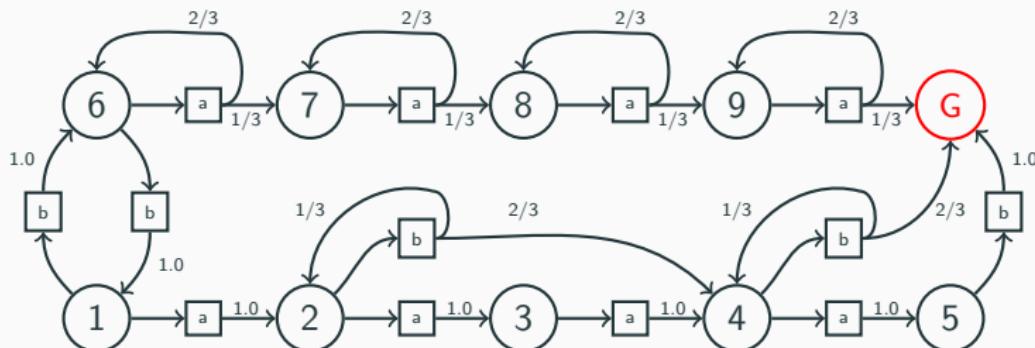
- Várias pessoas afetadas
- Incertezas no processo: aversão ao risco
- Incertezas no modelo: não estacionário, pessoas
- Poucos dados
- Meta vs Custo
- Representação de Preferências
- Níveis de descrição

Processo de Decisão Sequencial



1. o agente observa o estado s_t ;
 s_t : falta entregar pacotes 1, 2 e 3
2. o agente escolhe e executa uma ação a_t ;
 a_t : Joana, entregue o pacote 2
3. o ambiente transita para o próximo estado s_{t+1}
 s_{t+1} : falta entregar pacotes 1, e 3

Markov Decision Process (MDP)



- dinâmica probabilística
- estados, ações, função de transição, função de custo (recompensa)
- objectivo: alcançar a meta com o menor custo

Teorias Normativas de Racionalidade

- História (resultado, sequência, traço)



- resumo de histórias:
 - cada pessoa tem um resumo
 - custo (recompensa) acumulado: $\sum c_t = \sum c(s_t, a_t)$
 - atingiu a meta?
- ambiente estocástico + política $\pi : \mathcal{S} \rightarrow \mathcal{A}$
 - distribuição de probabilidade sobre histórias
 - custo acumulado esperado
 - variância do custo acumulado
 - probabilidade de atingir a meta

Axiomas de Justiça

Exemplo com duas pessoas: (C_1, C_2)

Como tratar equilíbrio entre resultados de cada pessoa?

- **Equidade:** quanto mais igualdade na distribuição melhor

$$(C_1 = 2, C_2 = 2) \succ (C_1 = 3, C_2 = 1) \succ (C_1 = 4, C_2 = 0)$$

- **Imparcialidade:** tratamento igual para iguais

$$(C_1 = 3, C_2 = 1) \sim (C_1 = 1, C_2 = 3)$$

- **Eficiência:** uma solução justa deve ser Pareto ótima

$$(C_1 = 3, C_2 = 1) \succ (C_1 = 3, C_2 = 3)$$

Axiomas de Decisões sobre Risco

Como comparar distribuições de probabilidades sobre histórias?

- **Ordenação:** \succeq é uma relação de ordem fraca sobre Π
- **Monotonicidade:** se $\pi' \succ \pi''$, então $\pi' \succ ([\alpha; \pi'], [1 - \alpha; \pi'']) \succ \pi''$ para todo $\alpha \in]0, 1[$;
- **Continuidade:** se $\pi' \succ \pi'' \succ \pi'''$, então existe $\alpha \in]0, 1[$ tal que $([\alpha; \pi'], [1 - \alpha; \pi''']) \sim \pi''$;
- se e somente se uma função de valor real V definida sobre Π existe tal que para qualquer $\pi', \pi'' \in \Pi$

$$\pi' \succ \pi'' \Leftrightarrow V(\pi') > V(\pi'')$$

e

$$V(\pi) = \sum_{h \in H} u(h) \Pr(h|\pi)$$

Perguntas de Pesquisa

1. Modelo Parametrizado

- qual função utilidade escolher?

$$u(C_1, C_2, \dots, C_k) = \sum_{i=1}^k e^{\beta C_i}$$

- outras teorias: $V(\pi) = \max C_i^\pi$

2. Algoritmo de Planejamento

- Programação Dinâmica
- Programação Linear
- Aprendizado por Reforço

3. Como escolher parâmetros

- pré-planejamento: encontrar parâmetros
- human in the loop: políticas interpretáveis
- pós-planejamento: non-dominated

4. Como descrever políticas para humanos

- política resumida
- histórias resumidas
- distribuições resumidas

Desafios para Tomada de Decisão



- Várias pessoas afetadas
- Incertezas no processo: aversão ao risco
- Incertezas no modelo: não estacionário, pessoas
- Poucos dados
- Meta vs Custo
- Representação de Preferências
- Níveis de descrição

Desafios para Tomada de Decisão

Pesquisas Passadas:

- Incertezas no processo: aversão ao risco
- Meta vs Custo

Pesquisas Atuais:

- Alexia Scheffer: várias pessoas afetadas, representação de preferências
- Igor Agostinho: níveis de descrição, incertezas no modelo
- João Victor da Silva: várias pessoas afetadas, poucos dados
- David de Andrade: várias pessoas afetadas, poucos dados
- Fernando Hsieh: poucos dados
- Guilherme Leite: poucos dados

Aversão ao Risco

REIS, FREIRE, DELGADO. Forward-PECVaR Algorithm: Exact Evaluation for CVaR SSPs. AAMAS, 2023.

FREITAS, FREIRE, DELGADO. Risk Sensitive Stochastic Shortest Path and LogSumExp: From Theory to Practice In: Brazilian Conference on Intelligent Systems. BRACIS, 2020.

PASTOR, BORGES, FREIRE, DELGADO, BARROS. Risk-Sensitive Piecewise-Linear Policy Iteration for Stochastic Shortest Path Markov Decision Processes. MICAI, 2020.

FREIRE, DELGADO. Extreme Risk Averse Policy for Goal-Directed Risk-Sensitive Markov Decision Process. BRACIS, 2016.

FREIRE. The Role of Discount Factor in Risk-Sensitive Markov Decision Processes. BRACIS, 2016.

MINAMI, FREIRE. Shortest Stochastic Path with Risk Sensitive Evaluation. MICAI, 2013.

CRISPINO, FREIRE, DELGADO. GUBS criterion: arbitrary trade-offs between cost and probability-to-goal in stochastic planning based on Expected Utility Theory. AIJ, 2023.

KUO, FREIRE. Probability-to-Goal and Expected Cost Trade-Off in Stochastic Shortest Path. ICCSA, 2021.

FREIRE, DELGADO, REIS. An Exact Algorithm to Make a Trade-Off between Cost and Probability in SSPs. ICAPS, 2019.

FREIRE, DELGADO. GUBS: a Utility-Based Semantic for Goal-Directed Markov Decision Processes. AAMAS, 2017.

Desafios:

- várias pessoas afetadas
- representação de preferências
- níveis de descrição

Projeto:

- Señoritas Courier
- Capacitated Vehicle Routing Problem
- Programação Linear com Janelas de Distribuição
- Medidas de Equidade
- Políticas Não-dominadas: equidade vs quilometragem

Desafios:

- níveis de descrição
- incertezas no modelo

Projeto:

- Señoritas Courier
- MDP Imprecisos
- Escolha de Histórias representativas
- História condicionadas em incertezas do modelo

Desafios:

- várias pessoas afetadas
- poucos dados

Projeto:

- trabalhadores com chegada e disponibilidade aleatórias
- alocação de tarefas sob demanda: repetidas vezes, possibilidade de falha na execução
- Programação Linear: equidade na execução de tarefas
- aprendizado do modelo relacionando trabalhadores e tarefas
- replanejamento com obtenção de novas experiências

Desafios:

- várias pessoas afetadas
- poucos dados

Projeto:

- trabalhadores com retorno e disponibilidade aleatórias
- alocação de tarefas sob demanda: uma única solução, possibilidade de falha na execução
- Contextual Multi-Armed Bandit
- Tarefas descritas em Linguagem Natural

Desafios:

- poucos dados

Projeto:

- Separar execução de operadores do efeito dos operadores
- Uso de Lógica para descrição de operadores
- Aprendizado por Reforço: execução de operador
- Descrição incompleta do efeito de operadores
- Planejamento: escolha de próximo operador

Desafios:

- poucos dados

Projeto:

- Aprendizado por Reforço
- Baseado em Modelos
- Life-Long Learning: aprendizado de conjunto de dinâmicas
- Inferência por Verossimilhança

Perguntas?