

Characterizing Optimal Mixed Policies: Where to Intervene and What to Observe

Sanghack Lee and Elias Bareinboim
Columbia University

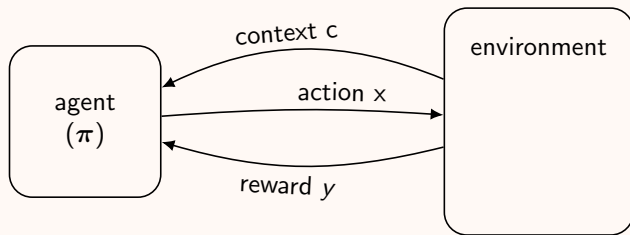
NeurIPS 2020
December 2020

executive summary

- ▶ Q: How can Causal Knowledge be utilized in Decision Making?
freedom to choose variables to intervene and observe
→ a large policy space
- ▶ Characterizations for an **efficient** and **effective** exploration:
non-redundancy no unnecessary intervention and observation
optimality capable to achieve an optimal reward
- ▶ The characterizations lead to *faster* convergence to an optimal policy.

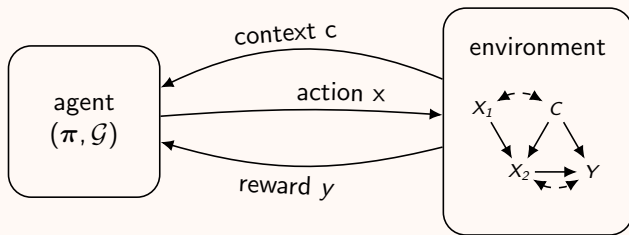
Decision Making, Policy, and Causal Structure

Agent \leftrightarrow Environment



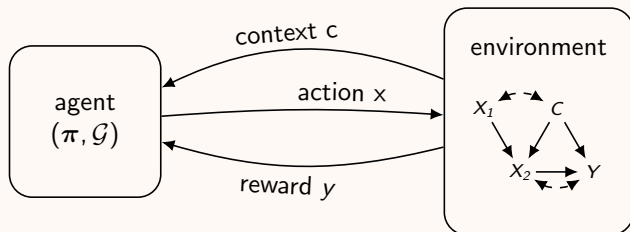
Decision Making, Policy, and Causal Structure

Agent \leftrightarrow Environment

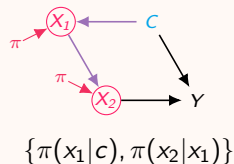
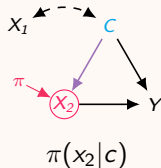
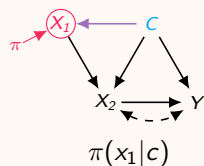


Decision Making, Policy, and Causal Structure

Agent \leftrightarrow Environment



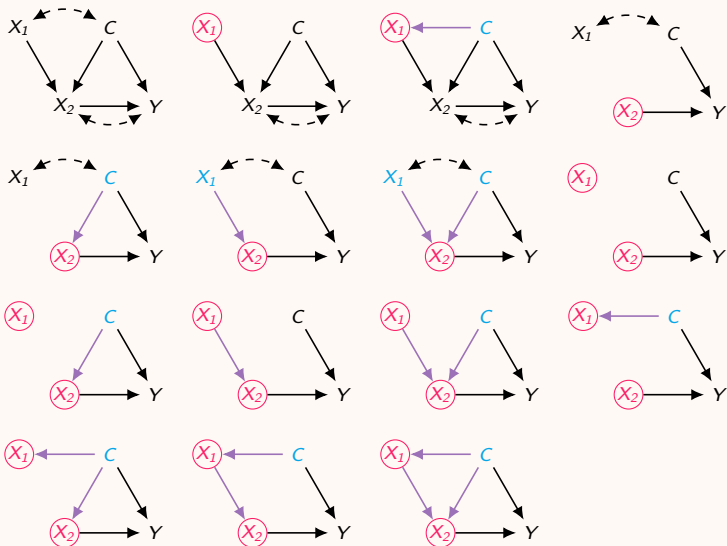
Mixed Policy Examples



\circledast for intervened variables, \rightarrow for policy-induced dependency.

The Modes of Interaction

Mixed Policy Scopes (as their induced graphs)



Characterizations: Non-redundancy & Optimality

- ▶ One scope has **fewer** actions or contexts than the other has, *yet* with the **same** optimal reward guarantee:



Non-redundancy: \nexists a smaller scope with the same optimal reward

- ▶ One's best policy is always **as good as** the other's:

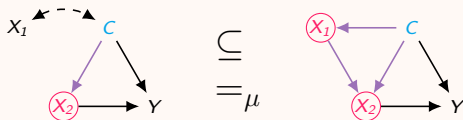


Optimality: \nexists a better scope wrt an optimal reward

\subseteq : fewer actions and/or contexts, $=_{\mu}$: rewards for their optimized policies are the same.

Characterizations: Non-redundancy & Optimality

- ▶ One scope has **fewer** actions or contexts than the other has, yet with the **same** optimal reward guarantee:



Non-redundancy: \nexists a smaller scope with the same optimal reward

- ▶ One's best policy is always **as good as** the other's:

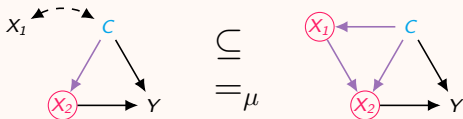


Optimality: \nexists a better scope wrt an optimal reward

\subseteq : fewer actions and/or contexts, \geq_{μ} : rewards for their optimized policies are the same.

Characterizations: Non-redundancy & Optimality

- ▶ One scope has **fewer** actions or contexts than the other has, yet with the **same** optimal reward guarantee:



Non-redundancy: \nexists a smaller scope with the same optimal reward

- ▶ One's best policy is always as good as the other's:

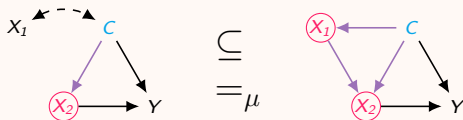


Optimality: \nexists a better scope wrt an optimal reward

\subseteq : fewer actions and/or contexts, $\stackrel{\mu}{=}$: rewards for their optimized policies are the same.

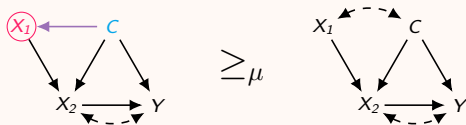
Characterizations: Non-redundancy & Optimality

- ▶ One scope has **fewer** actions or contexts than the other has, yet with the **same** optimal reward guarantee:



Non-redundancy: \nexists a smaller scope with the same optimal reward

- ▶ One's best policy is always **as good as** the other's:

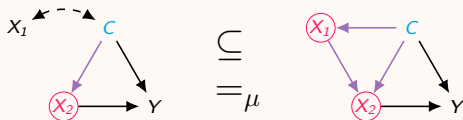


Optimality: \nexists a better scope wrt an optimal reward

\subseteq : fewer actions and/or contexts, $=_{\mu}$: rewards for their optimized policies are the same.

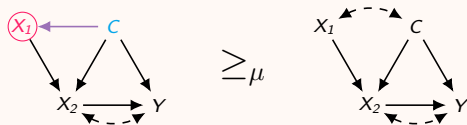
Characterizations: Non-redundancy & Optimality

- ▶ One scope has **fewer** actions or contexts than the other has, yet with the **same** optimal reward guarantee:



Non-redundancy: \nexists a smaller scope with the same optimal reward

- ▶ One's best policy is always **as good as** the other's:

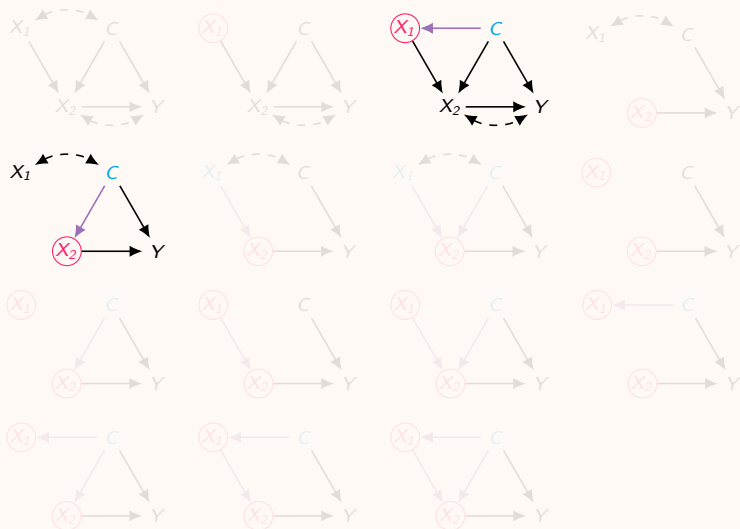


Optimality: \nexists a better scope wrt an optimal reward

\subseteq : fewer actions and/or contexts, $=_{\mu}$: rewards for their optimized policies are the same.

Characterizations → Efficient and Effective Exploration

Mixed Policy Scopes with *Non-redundancy* & *Optimality*



Characterizations → Efficient and Effective Exploration

Mixed Policy Scopes with *Non-redundancy* & *Optimality*

