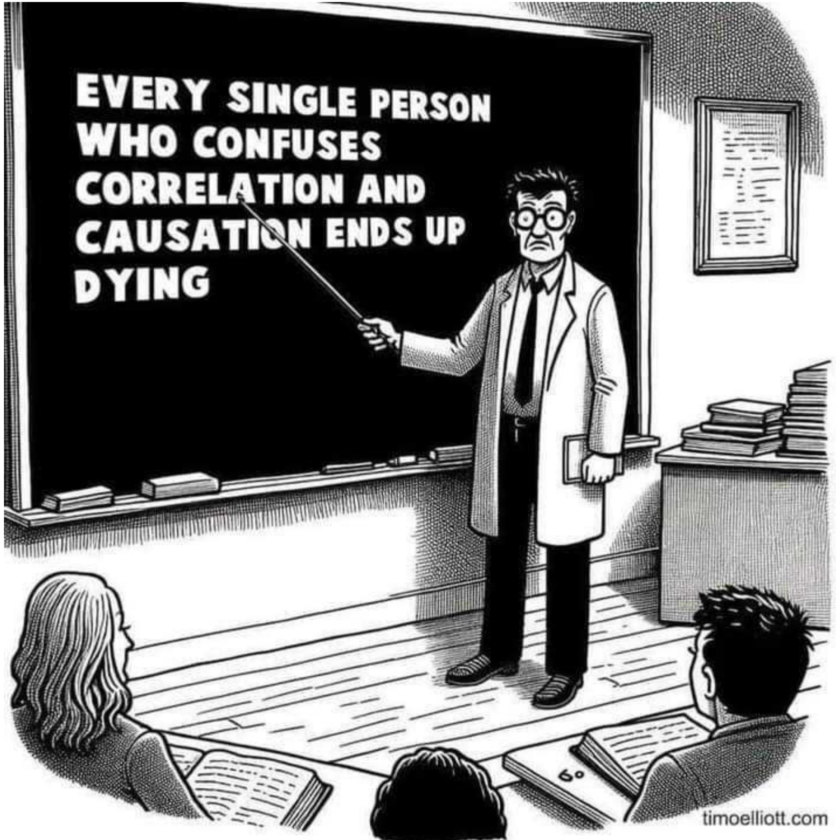
FACULTY OF INFORMATION TECHNOLOGY CTU IN PRAGUE | DATA SCIENCE LABORATORY

Introduction to Causality
AML Spring 2024

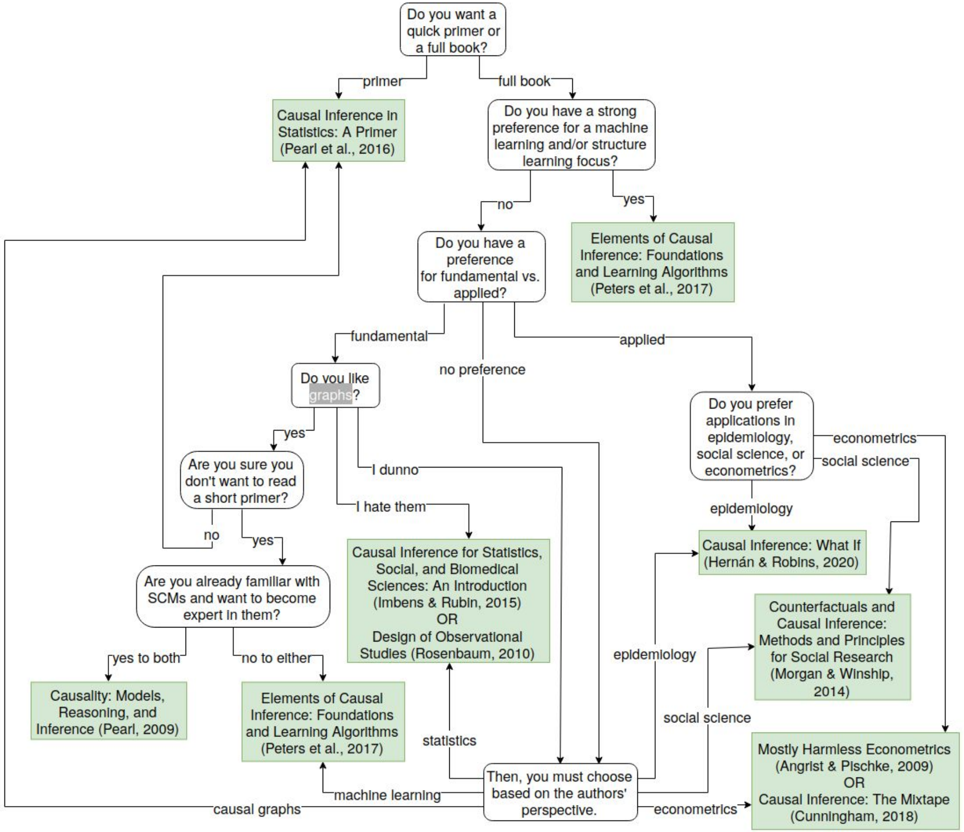


Causal Inference

- Inferring the effects of any treatment/policy/intervention/etc.
- Examples
  - Effect of treatment on a disease
  - Effect of climate change policy on emissions
  - Effect of social media on mental health
  - Many more (effect of X on Y)

Causality

- Concept that could be approached from various standpoints
- Used in field like
  - Econometrics
  - Social science
  - Epidemiology
  - Statistics
  - Machine Learning
  - (Multi-agent) Reinforcement Learning



Simpson's Paradox

FACULTY OF INFORMATION TECHNOLOGY CTU IN PRAGUE | DATA SCIENCE LABORATORY

- Hypothetical disease with two possible treatments
- Table showing mortality rate

Treatment	Condition: Mild	Condition: Severe	Condition: Total
A	15% (210/1400)	30% (30/100)	**16%** (240/1500)
B	**10%** (5/50)	**20%** (100/500)	19% (105/550)

- Apparent paradox:
    - If condition is not know, treatment A is better
    - If condition is known, treatment B is better

Simpson's Paradox

Which treatment is better depends on the **causal** structure of the data

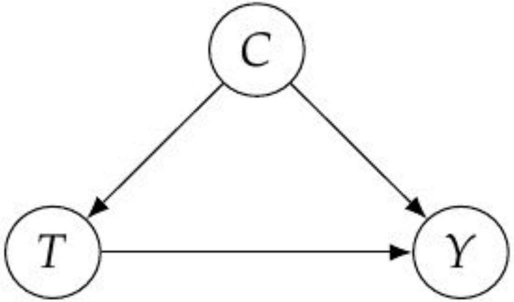


Figure 1.1: Causal structure of scenario 1, where condition C is a common cause of treatment T and mortality Y. Given this causal structure, treatment B is preferable.

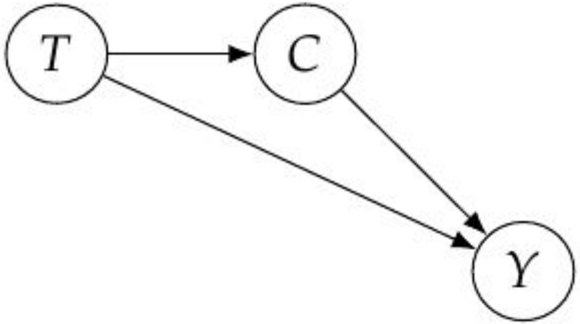
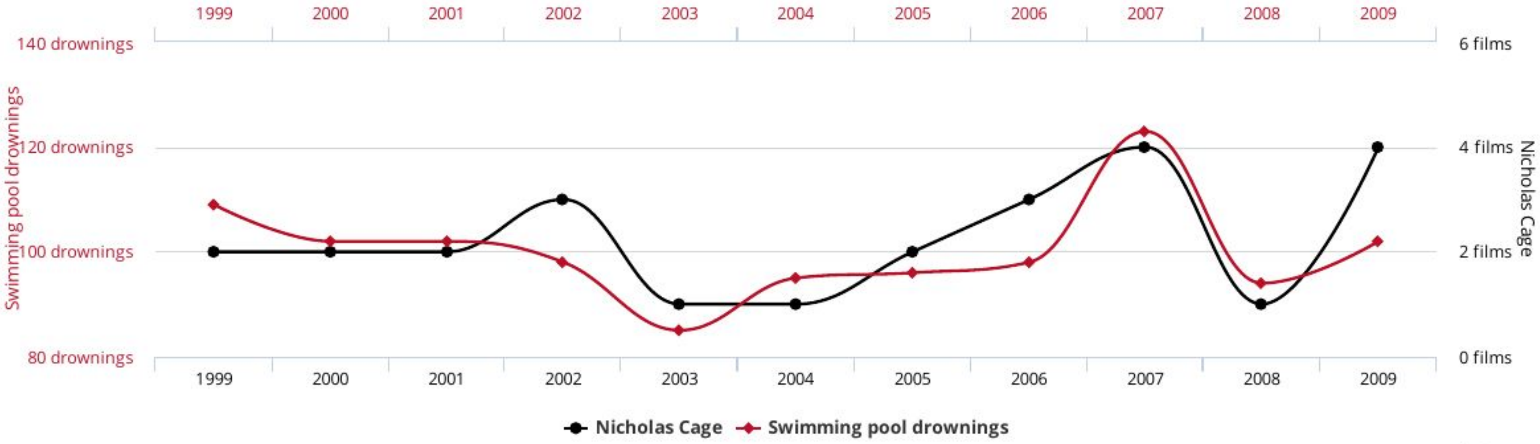


Figure 1.2: Causal structure of scenario 2, where treatment T is a cause of condition C. Given this causal structure, treatment A is preferable.

Correlation is not causation



Association is Not Causation

- Correlation is meant statistical dependence
- Technically, it is measure of linear dependence, better term should be association
- Total association is no all or none, could be combination of
  - Spurious (correlation)
  - Confounding (hidden common cause)
  - Causal association

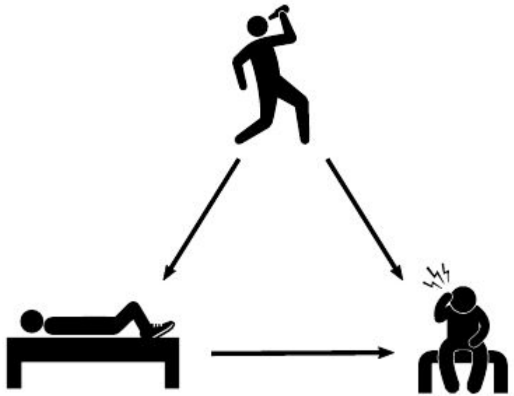


Figure 1.4: Causal structure, where drinking the night before is a common cause of sleeping with shoes on and of waking up with a headaches.

Counterfactuals

- Alternatives scenarios that did not actually happened but could happen under different circumstances
- Humans as Counterfactual Reasoning Machines
  - Constantly evaluating alternative scenarios
  - Imagining outcomes of different actions
- Counterfactuals in Everyday Life
  - Informed decision-making based on "what-if" analysis
  - Learning from past experiences and mistakes
- Regret Minimization in Human Behavior
  - Comparing outcomes of taken and untaken actions
  - Guiding future decisions to minimize regret

Potential Outcomes Framework

- person has a headache and decides
  - Take a pill (treatement)
  - Not take a pill (control)
- Potential outcome: will headache persit
  - $Y_i(1)$ severity of headache hour after taking the pill
  - $Y_i(0)$ severity of headache hour after (not taking a pill)
- Individual treatment effect
  - tau = $Y_i(1)$ - $Y_i(0)$

$$\tau_i = Y_i(1) - Y_i(0)$$

- We can observe only one of outcomes $Y_i(1)$ *and* $Y_i(0)$
- How to compute the treatment effect?
  - Fundamental problem of causal inference

Average treatment effect

- To estimate the average causal effect of the pill, we can use a sample of individuals who took the pill and another sample of individuals who did not.
- Average treatment effect

$$\tau = \frac{1}{N_1}\sum_{i=1}^{N_1} Y_i(1) - \frac{1}{N_0}\sum_{i=1}^{N_0} Y_i(0)$$

- How to average question marks?

i	T	Y	$Y(1)$	$Y(0)$	$Y(1) - Y(0)$
1	0	0	?	0	?
2	1	1	1	?	?
3	1	0	0	?	?
4	0	0	?	0	?
5	0	1	?	1	?
6	1	1	1	?	?

Ignorability

- What makes it valid to calculate the ATE by taking the average of the Y(0) column, ignoring the question marks, and subtracting that from the average of the Y(1) column, ignoring the question marks?
- Ignoring the question is called ignorability
  - ignoring how people ended up selecting the treatment they selected
  - and just assuming they were randomly assigned their treatment;

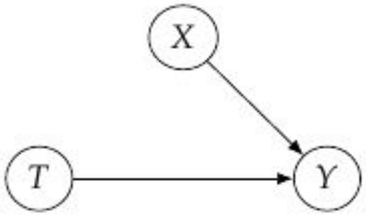


Figure 2.2: Causal structure when the treatment assignment mechanism is ignorable. Notably, this means there's no arrow from X to T, which means there is no confounding.

Controlling for Confounding

- Confounding Factor
  - Variables that affect both the treatment and the outcome
  - Can lead to biased estimates of causal effects
- Importance of Controlling for Confounding
  - Obtain unbiased and accurate estimates of causal effects
  - Improve decision-making based on observational data
- Methods to Control for Confounding
  - Matching
  - Stratification

Control for confounding

- Matching
  - Attempt to create comparable groups of individuals who took the pill and those who did not
  - Match based on confounding variables (e.g., age, gender, baseline health)
  - nearest neighbor matching
  - directly pairs treated and control individuals based on their similarity in confounding variables,
- Stratification
  - divide the population into strata based on the confounding variables
    - Intial headache severity
  - estimate the causal effect within each stratum
  - combine these estimates to calculate the overall average causal effect
  - weighting the estimates by the proportion of individuals in each stratum.
  - divides the population into groups based on the values of confounding variables and estimates the causal effect within each group.

Deep End-to-end Causal Inference (DECI)

- Python package by Microsoft Research
- Causal discovery
- ATE and CATE

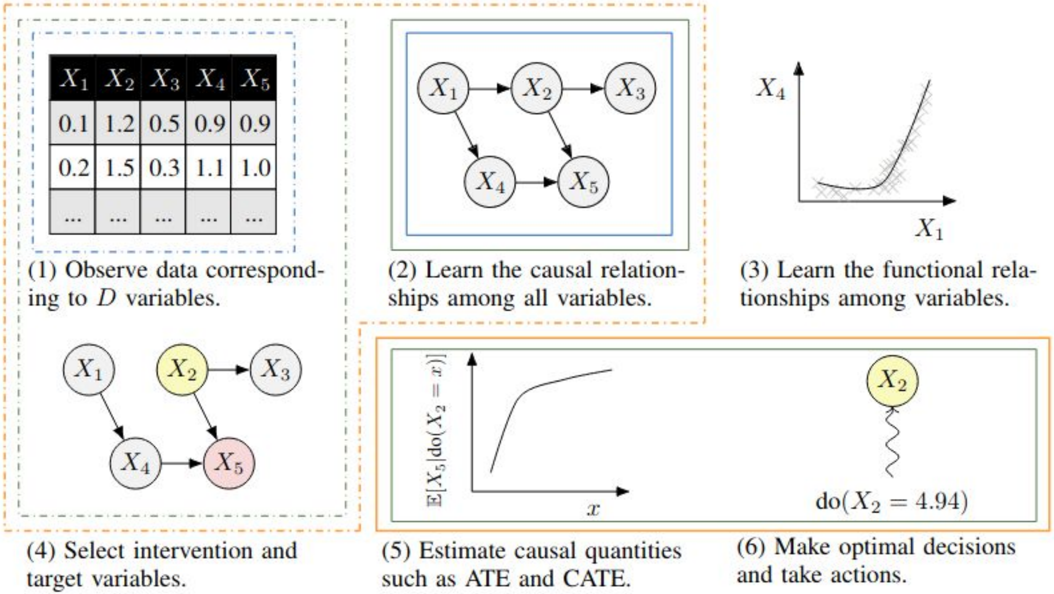


Figure 1: An overview of the deep end-to-end causal inference pipeline compared to traditional causal discovery and causal inference. The dashed line boxes show the inputs and the solid line boxes show the outputs. In causal discovery, a user provides observational data (1) as input. The output is the causal relationship (2) which are DAGs or partial DAGs. In causal inference, the user needs to provide both the data (1) and the causal graph (2) as input and provide a causal question by specifying treatment and effect (4), a model is learned and outputs the causal quantities (5) which helps decision making (6). In this work, we aim to answer causal questions end-to-end. DECI allows the user to provide the observational data only and specify any causal questions and output both the discovered causal relationship (2) and the causal quantities (5) that helps decision making (6).

Causal Discovery in DECI

- Models relationships among variables $x_1,\ldots,x_N$ and a causal graph G using joint probability $p_\theta(x_1,\ldots,x_N,G) = p(G)\prod_{n=1}^{N} p_\theta(x_n|G)$

  Where p(G) represents a prior over graphs and $p_\theta(x_n|G)$ is the likelihood of observing x_n given the graph G and parameters θ.

- The graph prior p(G) encourages the graph structure to be a DAG using a penalty function on the adjacency matrix A of G

Prior over Graphs. The graph prior $p(G)$ should characterize the graph as a DAG. We implement this by leveraging the continuous DAG penalty from Zheng et al. [73],

$$h(G) = \operatorname{tr}\left(e^{G \odot G}\right) - D, \quad (5)$$

which is non-negative and zero only if G is a DAG. We then implement the prior as

$$p(G) \propto \exp\left(-\lambda_s \|G\|_F^2 - \rho\, h(G)^2 - \alpha\, h(G)\right), \quad (6)$$

Causal Discovery in DECI

FACULTY OF INFORMATION TECHNOLOGY CTU IN PRAGUE
DATA SCIENCE LABORATORY

- Given a graph *G*, the likelihood for a single observation x_n is factorized autoregresively assuming an additive noise model

$$p_\theta(x_n|G) = \prod_{i=1}^{D} p_{z_i}(x_i - f_i(x_{\text{pa}(i,G)}, \theta))$$

- $x_{pa}(i,G)$ are parent variables of x_i in G
- f_i is a function specifying the causal mechanism from $x_{pa}(i,G)$ to x_i parametrized by θ
- p_{zi} is the distribution of additive noise for variable x_i
- additive noise model

$$x_i = f_i(x_{\text{pa}(i,G)}, \theta) + z_i$$

Causal Discovery in DECI

- True posterior $p_\theta(G|x_1,\ldots,x_n)$ is intractable
- Deci approximates it with variational distribution $q_\phi(G)$ and maximizes the Evidence Lower Bound (ELBO) to learn θ and ϕ

$$\mathrm{ELBO}(\theta, \phi) = \mathbb{E}_{q_\phi(G)}\left[\log p(G) + \sum_{n=1}^{N} \log p_\theta(x_n|G)\right] + H(q_\phi)$$

- $H(q_\phi)$ is the entropy of $q_\phi(G)$, encouraging exploration of different graph structures

Estimating ATE in DECI

- After DECI has been trained, it can simulate interventions on the treatment variables
- DECI estimates these expectations by generating samples from the interventional distributions and calculating the mean outcome for both treated and untreated scenarios:

  - Generate samples x_Y^a from $p(x_Y|do(X_T = a))$ and calculate $\mathbb{E}[x_Y|do(X_T = a)]$ as the mean of x_Y^a.
  - Generate samples x_Y^b from $p(x_Y|do(X_T = b))$ and calculate $\mathbb{E}[x_Y|do(X_T = b)]$ as the mean of x_Y^b.

- Then

$$\text{ATE} = \mathbb{E}[x_Y|do(X_T = a)] - \mathbb{E}[x_Y|do(X_T = b)]$$

Towards Causal Representation Learning

- Bengio et al., 2021
- Causal inference can help address important challenges in machine learning such as generalization, transfer learning, and data efficiency.
- Causal representation learning is a crucial problem for artificial intelligence and could unlock new capabilities in learning from data.
- Incorporating causality into machine learning models requires careful consideration of assumptions, limitations, and trade-offs.
- Combining causal inference techniques with machine learning models to improve generalization and transfer learning.
- Developing algorithms for causal representation learning that can handle complex data types such as images, audio, and video.
- Incorporating causality into reinforcement learning algorithms to improve the performance of agents in complex environments.

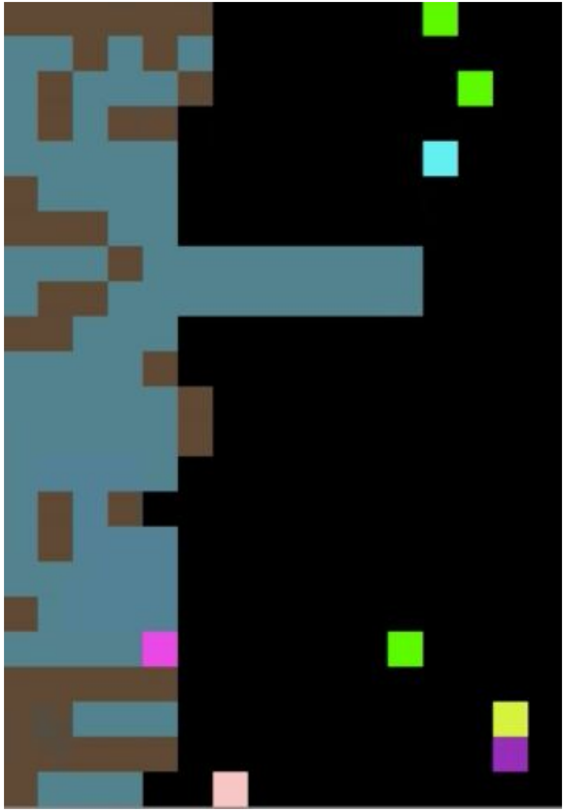
Social Influence in MARL

- Jacques et al.: Social Influence as Intrinsic Motivation for Multi-Agent Deep Reinforcement Learning
- mechanism for achieving coordination and communication
- rewarding agents for having causal influence over other agents' actions
- causal influence is assessed using **counterfactual reasoning**
- At each timestep, an agent simulates alternate actions that it could have taken, and computes their effect on the behavior of other agents.
- Actions that lead to bigger changes in other agents' behavior are considered influential and are rewarded
- influence leads to enhanced coordination and communication in challenging social dilemma environments, dramatically increasing the learning curves of the deep RL agent

Sequential Social Dilemmas

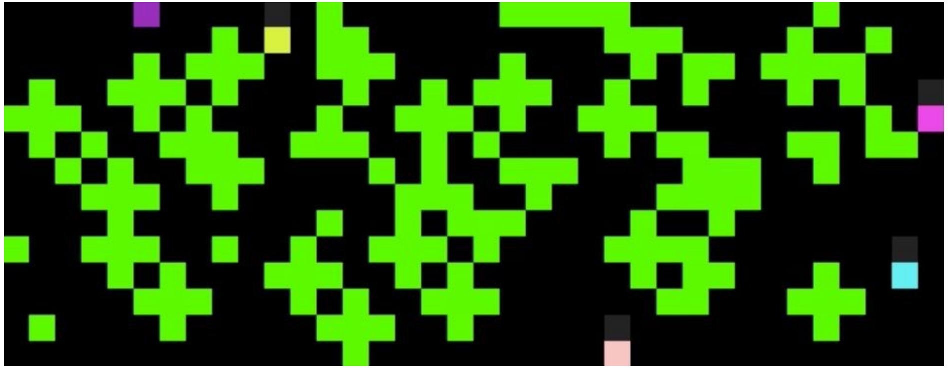
- Can be thought of as analogous to spatially and temporally extended Prisoner's Dilemma-like games.
- The reward structure poses a dilemma because individual short-term optimal strategies lead to poor long-term outcomes for the group

Cleanup



- A public goods dilemma in which
- agents get a reward for consuming apples, but must use a cleaning beam to clean a river in order for apples to grow.
- While an agent is cleaning the river, other agents can exploit it by consuming the apples that appear.

Harvest



- A tragedy-of-the-commons dilemma
- apples regrow at a rate that depends on the amount of nearby apples.
- If individual agents employ an exploitative strategy by greedily consuming too many apples, the collective reward of all agents is reduced.

Multi-agent Reinforecement Learning

Definition 2.2 (Dec-POMDP) *Decentralised Partially Observable Multi-Agent Markov Decision process is It is a 7-tuple* $\{S, \{A_i\}, T, R, \{\Omega_i\}, O, \gamma\}$ *, where* S *are states,* $\{A_i\}$ *is the joint action set,* $T = P(s'|s, a)$ *is the set of conditional transition probabilities between states,* R *is the reward function,* $\{\Omega_i\}$ *is the joint observation set,* $O(s', a, o) = P(o|s', a)$ *gives the conditional observation distribution, and* $\gamma \in [0, 1]$ *is the discount factor.*

Intrinsic motivation

- To stimulate agents to learn cooperative behavior introduce
- category of reward functions that allow learning of desired behaviors in a wide range of environments and tasks, sometimes even in the absence of environmental rewards.
- Social influence intrinsic motivation gives an agent k additional reward when it has causal influence on the actions of other agents.
- It adds a causal influence reward c_k^t to the agent's immediate environmental (extrinsic) reward e_t^k at time t:

$$r_t^k = \alpha e_t^k + \beta c_t^k.$$

Evaluation of social influence

FACULTY OF INFORMATION TECHNOLOGY CTU IN PRAGUE | DATA SCIENCE LABORATORY

To evaluate the causal influence of agent k on agent j at time t, agent j should be able to condition its action a_t^j on a_t^k, agent's k action at time t. Therefore, a_j can quantify the probability of the next step action as

$$p(a_t^j | a_k^t, s_t^j).$$

Then we can we can replace a_t^k by $\tilde{a}_t^k$, the counterfactual action, and compute a new next step probability

$$p(a_t^j | \tilde{a}_k^t, s_t^j).$$

Evaluation of social influence

By averaging the policy distribution from a sampling of several counterfactual actions, we would obtain the marginal policy of agent j:

$$p(a_t^j|s_t^j) = \sum_{\tilde{a}_t^k} p(a_t^j|\tilde{a}_t^k, s_t^j)p(\tilde{a}_t^k, s_t^j),$$

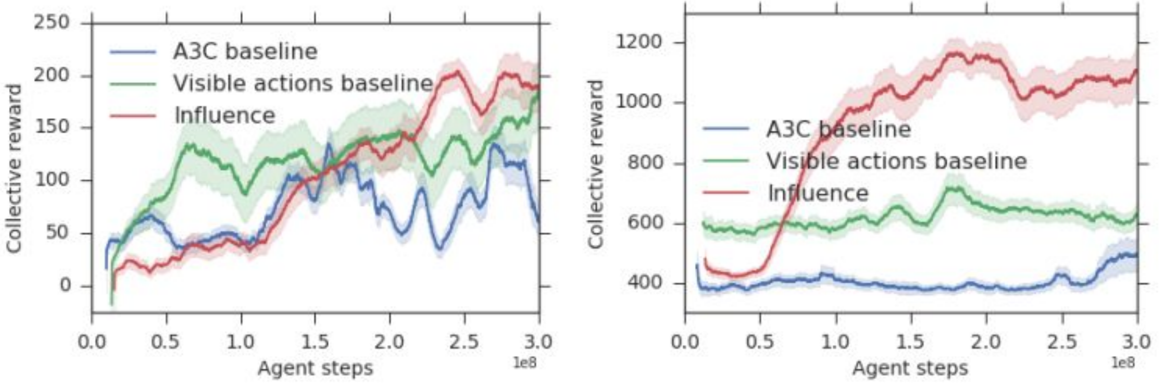
i.e. agent's j policy if it did not take into account actions of agent k.

The difference between agent's j marginal policy and the conditional policy of agent j after observing agent's k action is a degree of how agent k is causually influencing agent j.

Therefore, the overall causal influence of agent k on all other agents is given by:

$$\begin{aligned} c_t^k &= \sum_{j=0, j\neq k}^{N} \left[D_{KL} \left[p\left(a_t^j \mid a_t^k, s_t^j\right) \| \sum_{\tilde{a}_t^k} p\left(a_t^j \mid \tilde{a}_t^k, s_t^j\right) p\left(\tilde{a}_t^k \mid s_t^j\right) \right] \right] \\ &= \sum_{j=0, j\neq k}^{N} \left[D_{KL} \left[p\left(a_t^j \mid a_t^k, s_t^j\right) \| p\left(a_t^j \mid s_t^j\right) \right] \right], \end{aligned} \tag{4.1}$$

Effect of social influence



(a) *Cleanup* (b) *Harvest*

Figure 1: Total collective reward obtained in Experiment 1. Agents trained with influence (red) significantly outperform the baseline and ablated agents. In Harvest, the influence reward is essential to achieve any meaningful learning.

Social influence

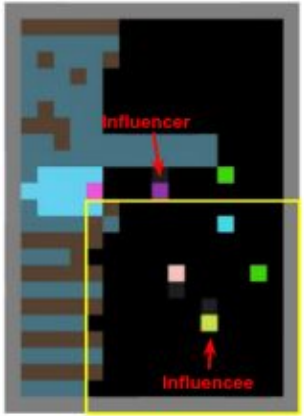


Figure 2: A moment of high influence when the purple influencer signals the presence of an apple (green tiles) outside the yellow influencee's field-of-view (yellow outlined box).

Figure 2 shows a moment of high influence between the influencer and the yellow influencee. The influencer has chosen to move towards an apple that is outside of the ego-centric field-of-view of the yellow agent. Because the influencer only moves when apples are available, this signals to the yellow agent that an apple must be present above it which it cannot see. This changes the yellow agent's distribution over its planned action, $p(a_t^j | a_t^k, s_t^j)$, and allows the purple agent to gain influence. A similar moment occurs when the influencer signals to an agent that has been cleaning the river that no apples have appeared by staying still (see Figure 14 in the Supplementary Material).

- Agents continue to move and explore randomly while waiting for apples to spawn,
- The **influencer** only traverses the map when it is pursuing an apple, then stops. The rest of the time it stays still.
- The **influencer** agent learned to use its own actions as a binary code which signals the presence or absence of apples in the environment

Model of Other Agents

- Computing the causal influence reward requires knowing the probability of another agent’s action given a counterfactual,
- Requires a centralized training approach in which agents could access other agents’ policy network
- To relax this unrealistic assumption we equip each agent with its own internal Model of Other Agents (MOA).
- The MOA is trained to predict all other agents’ next actions given their previous actions, and the agent’s egocentric view of the state: $p(at+1|at,sk\ t\ )$.

Model of other agents

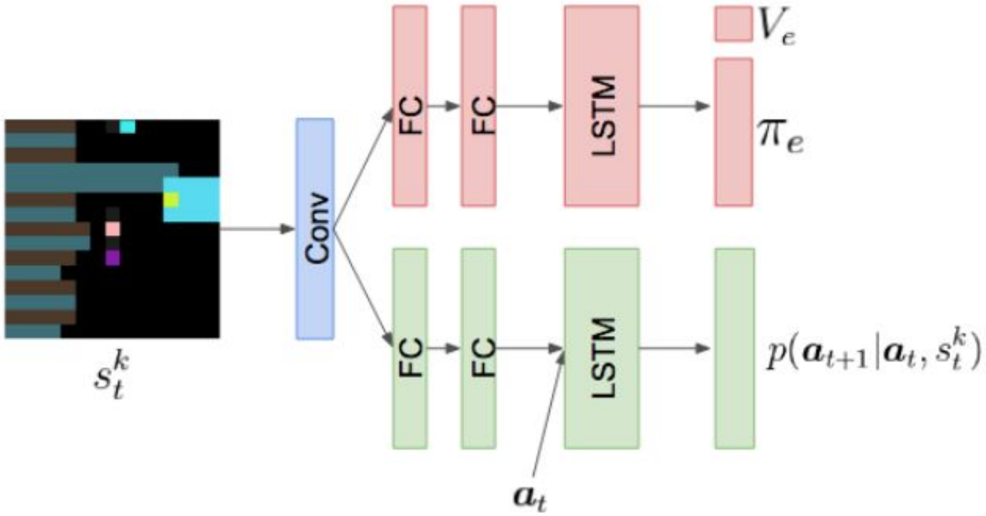


Figure 6: The Model of Other Agents (MOA) architecture learns both an RL policy π_e, and a supervised model that predicts the actions of other agents, $\boldsymbol{a}_{t+1}$. The supervised model is used for internally computing the influence reward.

Causal Diagram

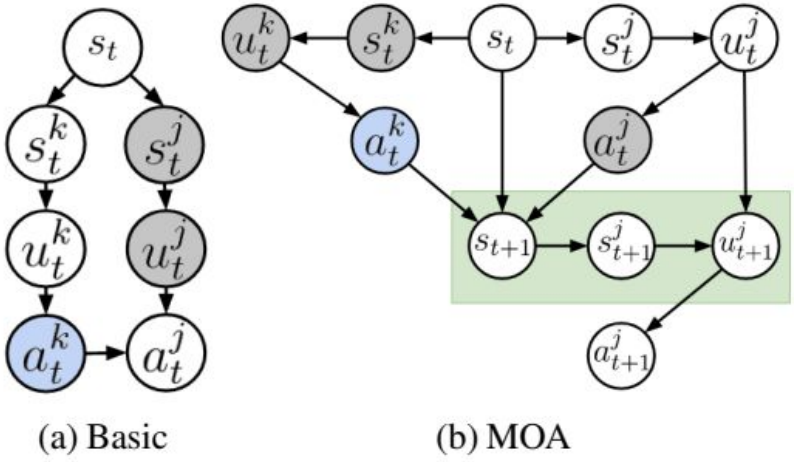


Figure 8: Causal diagrams of agent k's effect on j's action. Shaded nodes are conditioned on, and we intervene on a_t^k (blue node) by replacing it with counterfactuals. Nodes with a green background must be modeled using the MOA module. Note that there is no backdoor path between a_t^k and s_t in the MOA case, since it would require traversing a collider that is not in the conditioning set.

Causal Explanations for Sequential Decision-Making In Multi-Agent Systems

FACULTY OF INFORMATION TECHNOLOGY CTU IN PRAGUE
DATA SCIENCE LABORATORY

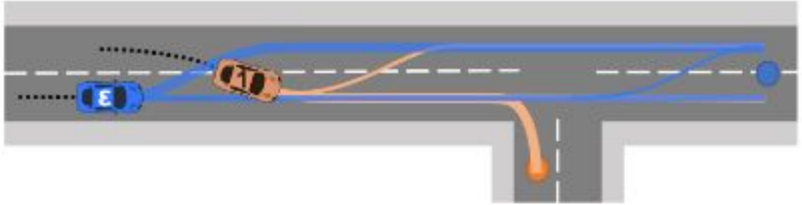
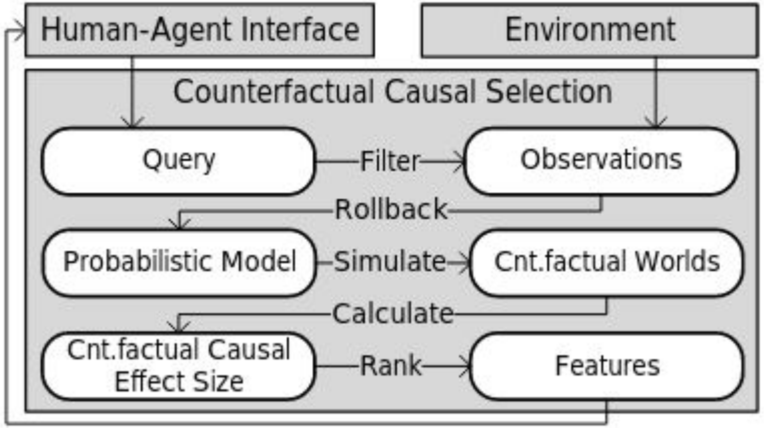


Figure 1: The autonomous vehicle (ε) is heading to the blue goal. It decided to change lanes after the other vehicle (1) cut in front of it and began to slow down. A passenger asks: *Why did you change lanes?* "To decrease the time to reach the goal." [teleological] *Why was changing lanes faster?* "Because the other vehicle is slower than us and is decelerating." [mechanistic] – Actual explanations by CEMA with explanation types in brackets. Blue/orange lines illustrate forward simulations using the probabilistic forward model.



F might include a discretized summary of actions, such as average acceleration or distance to the leading vehicle

Causal Explanations for Sequential Decision-Making In Multi-Agent Systems

Table 1: Binary features $\mathcal{F}$ to describe the fundamental motions and high-level actions of vehicles (including ego). For continuous values, the mean value is calculated along the length of the trajectory and thresholded with small value δ.

Feature	Calculation	Explanation
Acceleration	$a^i > \delta_a$	Accelerate
	$a^i < -\delta_a$	Decelerate
	$a^i \in [-\delta_a, \delta_a]$	Maintain velocity
Relative speed	$v^i - v^\varepsilon > \delta_v$	Faster than ego
	$v^i - v^\varepsilon < -\delta_v$	Slower than ego
	$v^i - v^\varepsilon \in [-\delta_v, \delta_v]$	Same speed as ego
Stop	$v^i \in [0, \delta_s]$	Does it stop
Maneuver	One-hot encode	Longest maneuver
Macro Action	One-hot encode	Longest macro action

Reward components R are

- longitudinal and lateral acceleration
- presence of collisions
- time to reach a destination
- goal completion

Causal Reasoning and Large Language Models

FACULTY OF INFORMATION TECHNOLOGY CTU IN PRAGUE
DATA SCIENCE LABORATORY

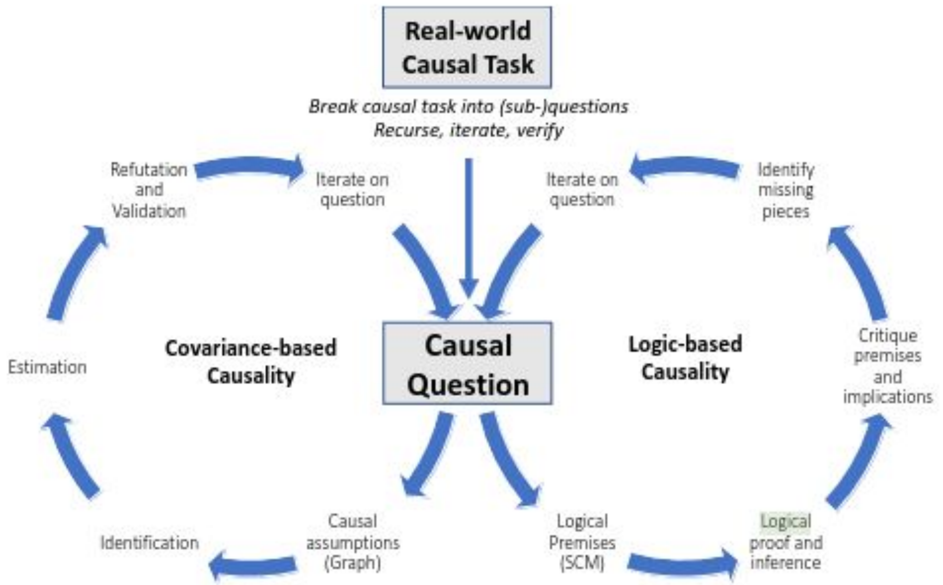
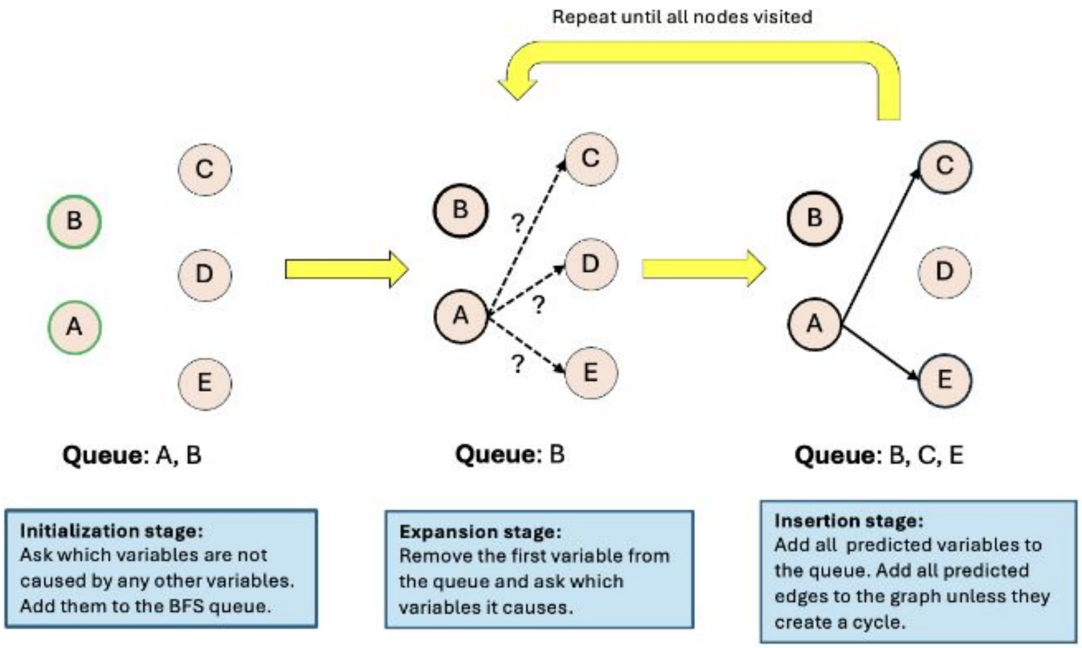


Figure 1: When tackling real-world causal tasks, people strategically alternate between logical- and covariance-based causal reasoning as they formulate (sub-)questions, iterate, and verify their premises and implications. Now, LLMs may have the capability to automate or assist with every step of this process and seamlessly transition bteween covariance- and logic-based causality.

Causal Reasoning and Large Language Models

- LLMs enable knowledge-based causal discovery, and achieve competitive performance in determining pairwise causal relationships between variables, across datasets from multiple domains, including medicine and climate science.
- Extending knowledge-based causal discovery to full graph discovery poses additional chalenges, such as distinguishing between direct and indirect causes.
- LLMs capture and apply common sense and domain knowledge enables substantial improvements in counterfactual reasoning and actual causality tasks, making them valuable in real-world applications
-

Efficient Causal Graph Discovery Using Large Language Models



Algorithm 1 BFS with LLMs

Require: LM p_θ, descriptions of variables X, initial variable selector $I()$, expansion generator $E()$, cycle checker $CheckCycle()$

$G \leftarrow \{\}$ ▷ Create an empty graph to store the result.
$frontier, visited \leftarrow I(p_\theta, X)$ ▷ With initialization prompt.
while $frontier$ is not empty **do**
  $toVisit \leftarrow frontier[0]$
  $frontier.remove(toVisit)$
  $visited.add(toVisit)$
  for $node$ in $E(p_\theta, G)$ **do** ▷ Expand with expansion prompt.
    if not $CheckCycle(G, toVisit, node)$ **then**
      ▷ Check if adding $toVisit \rightarrow node$ will create cycle.
      $G.add((toVisit, node))$
    end if
    if $node$ not in $frontier \cup visited$ **then**
      $frontier.add(node)$
    end if
  end for
end while
return G

```
You are a helpful assistant to a
neuropathic pain diagnosis expert.
The following factors are key
variables related to neuropathic pain
diagnosis which have various causal
effects on each other. Our goal is
to construct a causal graph between
these variables.

<A>: Description of variable A
<B>: Description of variable B
...

Now you are going to use the data to
construct a causal graph. You will
start with identifying the variable(s)
that are unaffected by any other
variables.
Think step by step. Then, provide your
final answer (variable names only)
within the tags <Answer>...</Answer>.
```

Initialization Stage

```
Given <Independent Variables> is(are)
not affected by any other variable and
the following causal relationships:

A causes B, C, D
C causes D, E
...

Select the variables that are caused by
<Currently Visited Node>.
Think step by step. Then, provide your
final answer (variable names only)
within the tags <Answer>...</Answer>.
```

Expansion Stage