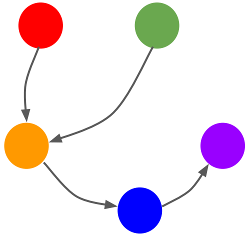
Enhancing Human-Robot Spatial Interaction through Causal Inference

Luca Castri,
lcastri@lincoln.ac.uk
PhD student
University of Lincoln

Website: ***https://darko-project.eu***

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101017274

Outline

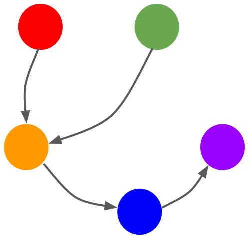


Motivation

Causal Inference overview

Causal Inference in Human-Robot Spatial Interaction

Outline



Motivation

Causal Inference overview

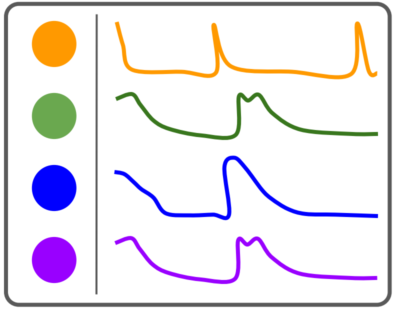
Causal Inference in Human-Robot Spatial Interaction

Motivation – what is Causal Inference?

- What is it?

  Science that studies the cause-and-effect relationship between events
  [Pearl, J., & Mackenzie, D. (2019). The book of why]

- It is divided into two main areas:

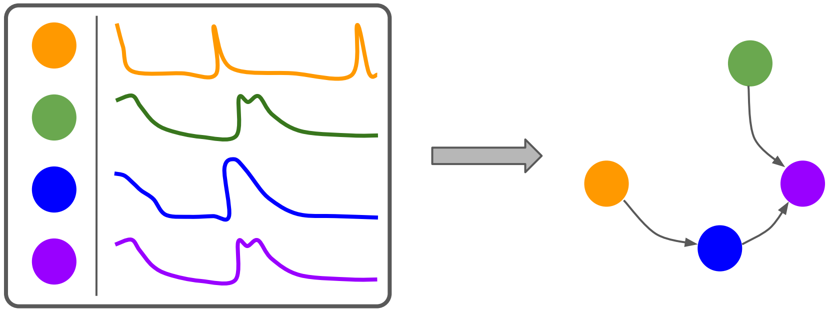


Motivation – what is Causal Inference?

- What is it?

  Science that studies the cause-and-effect relationship between events
  [Pearl, J., & Mackenzie, D. (2019). The book of why]

- It is divided into two main areas:



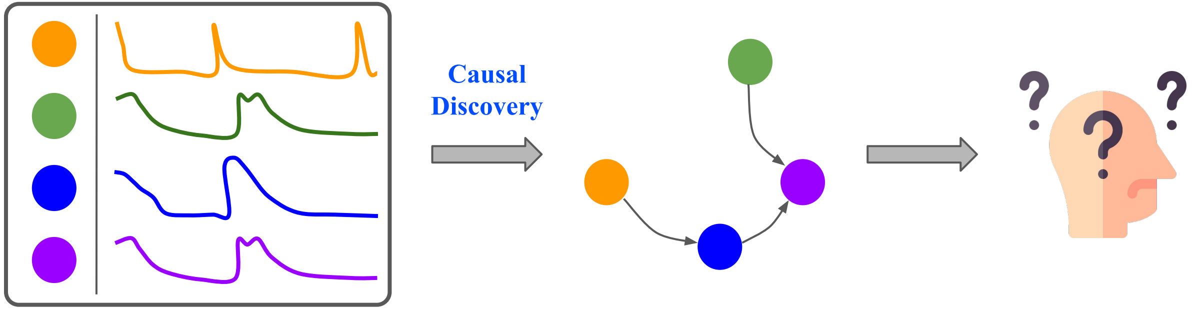
Causal Discovery
starting from a set of variables (events) aims to reconstruct the cause-effect model underlying them

Motivation – what is Causal Inference?

- What is it?

  Science that studies the cause-and-effect relationship between events
  [Pearl, J., & Mackenzie, D. (2019). The book of why]

- It is divided into two main areas:



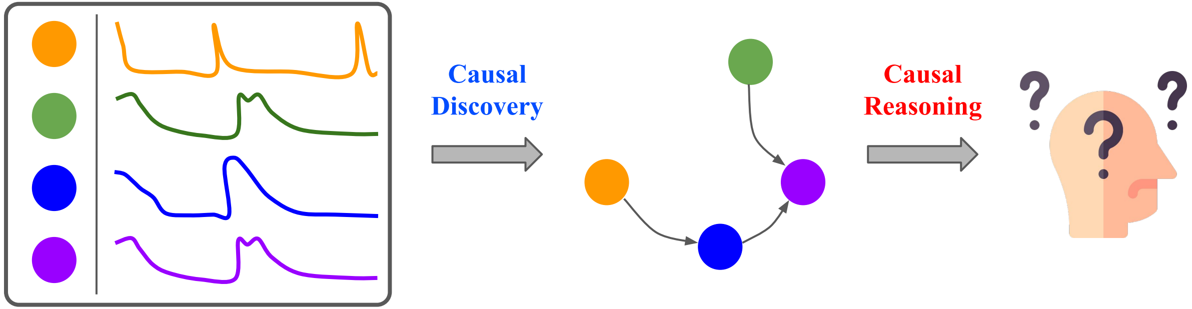
Causal Reasoning
reason on the causal model structure and on the cause-effect strength to predict how the observed system evolves

Motivation – what is Causal Inference?

- What is it?

  Science that studies the cause-and-effect relationship between events
  [Pearl, J., & Mackenzie, D. (2019). The book of why]

- It is divided into two main areas:

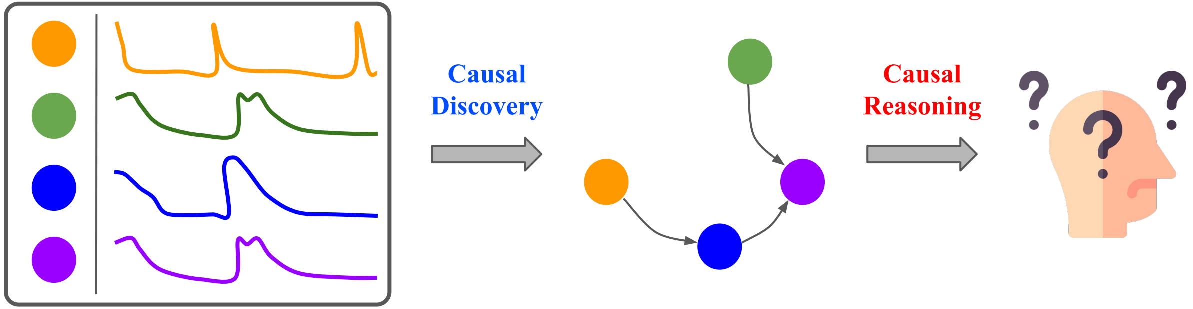


Motivation – what is Causal Inference?

- What is it?

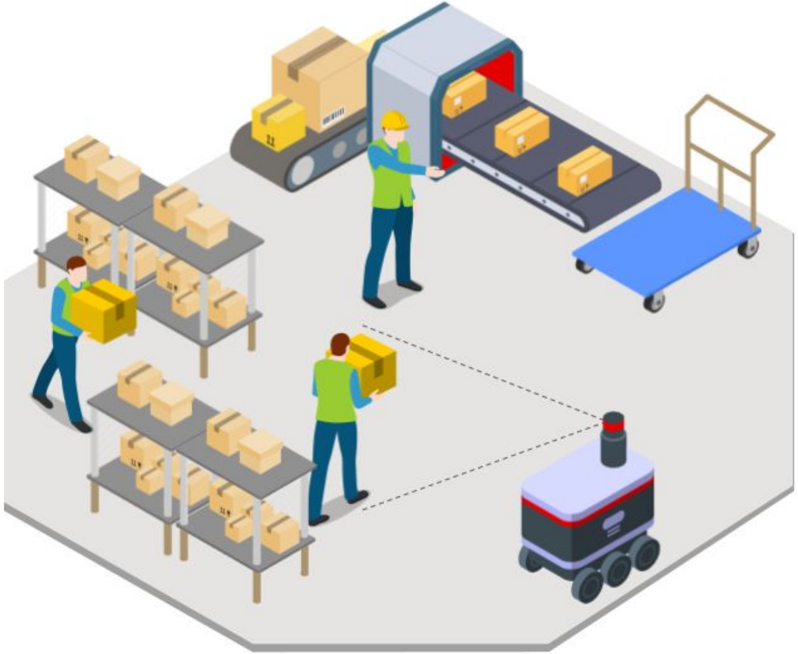
  Science that studies the cause-and-effect relationship between events
  [Pearl, J., & Mackenzie, D. (2019). The book of why]

- It is divided into two main areas:



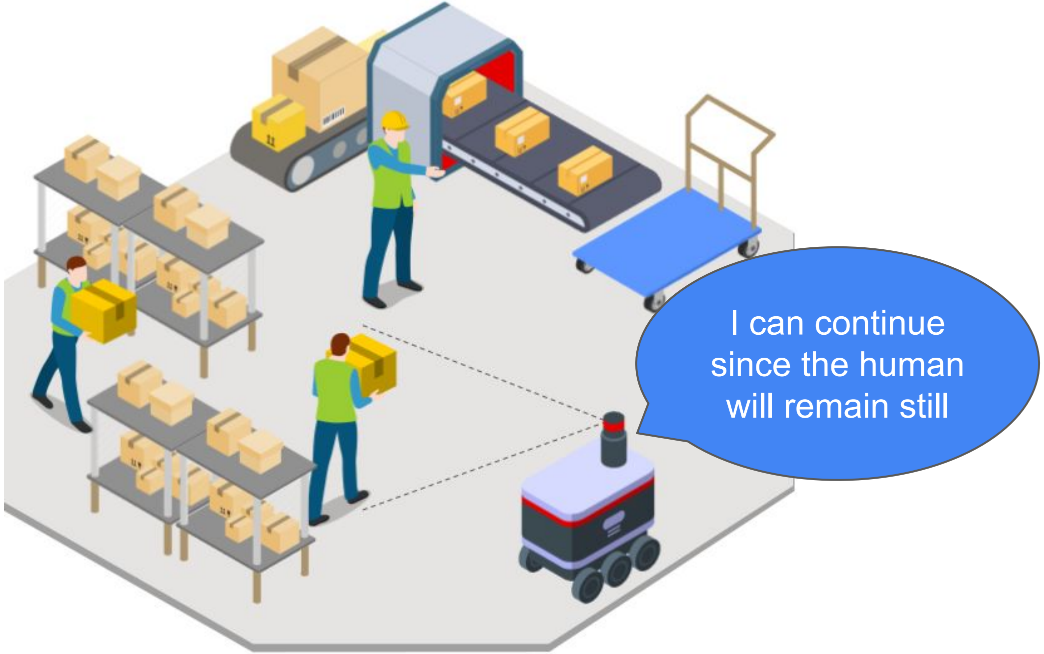
Can robots benefit from Causal Inference ?

Motivation – robotics scenario



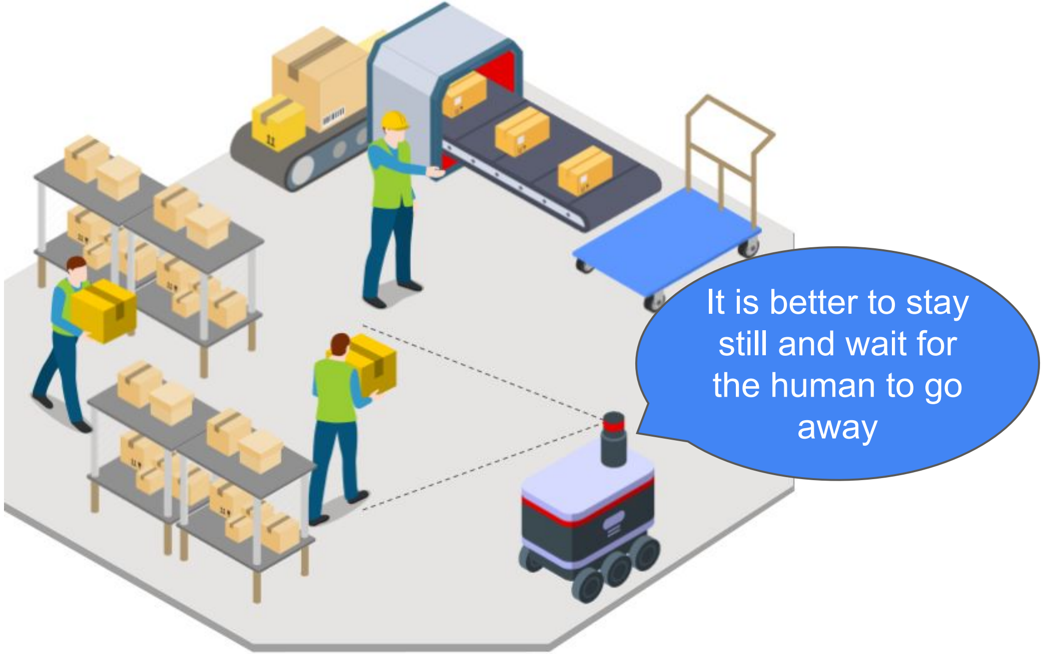
Dynamic Agile Production Robots That Learn and Optimise Knowledge and Operations

Motivation – robotics scenario



DARKO
Dynamic Agile Production Robots That Learn
and Optimise Knowledge and Operations

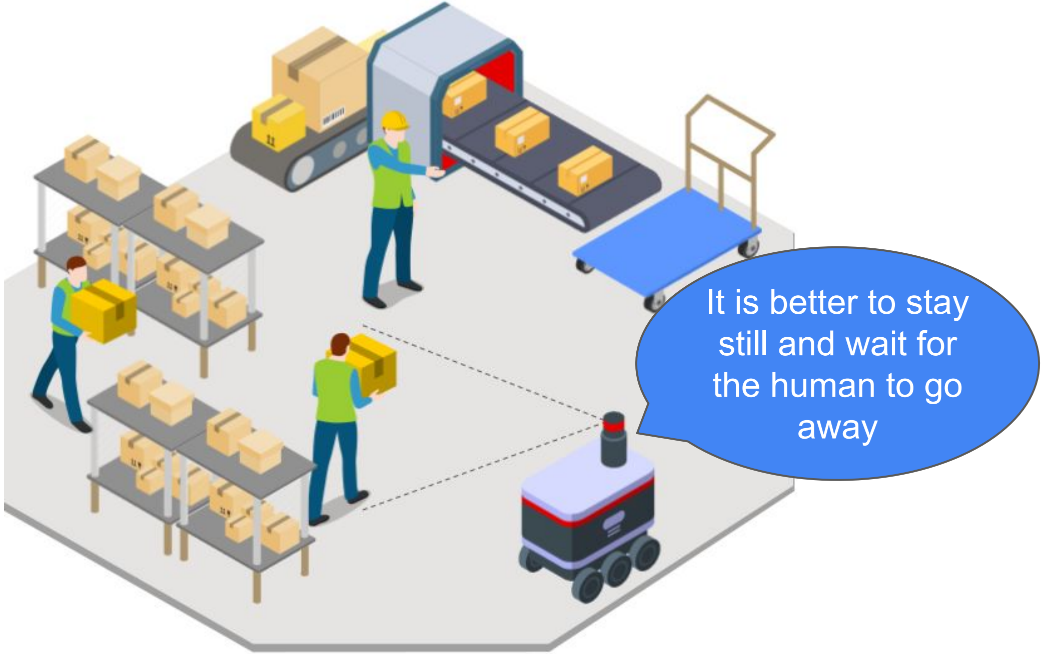
Motivation – robotics scenario



Motivation – robotics scenario

Discovering the causal model will enable the robot to reason on it and to answer questions like:

- “what happens if I go this way?”
- “what would have happened if I remained still instead of moving?”



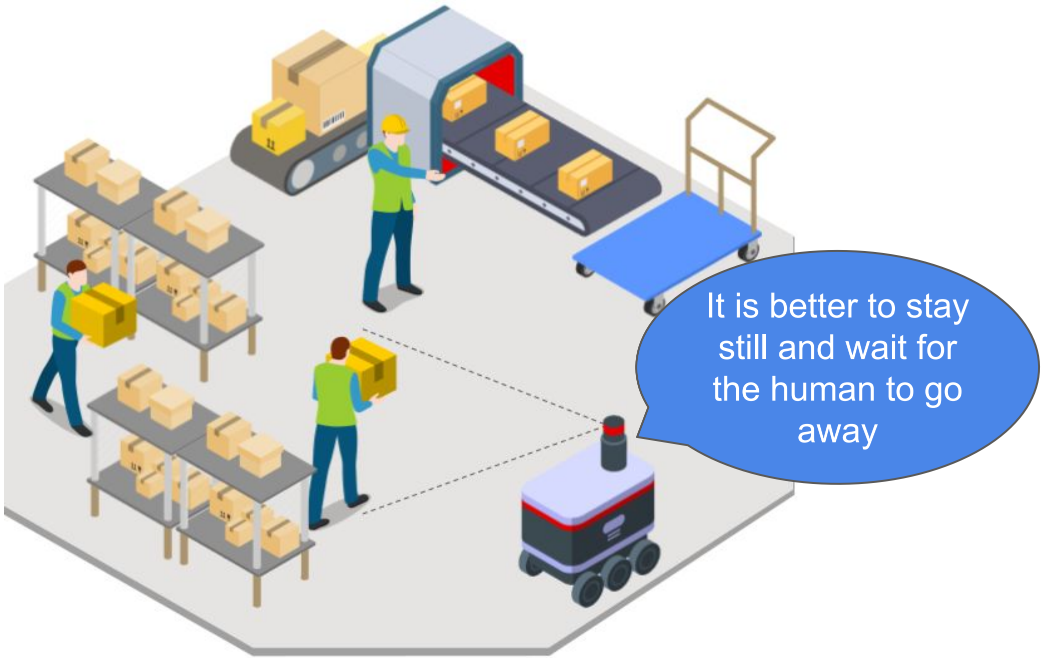
DARKO
Dynamic Agile Production Robots That Learn and Optimise Knowledge and Operations

Motivation – robotics scenario

Discovering the causal model will enable the robot to reason on it and to answer questions like:

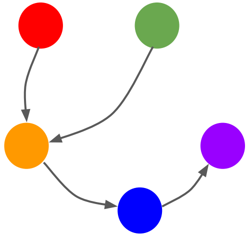
- "what happens if I go this way?"
- "what would have happened if I remained still instead of moving?"

take the best choice among possible HRSIs



Dynamic Agile Production Robots That Learn and Optimise Knowledge and Operations

Outline



Motivation

Causal Inference overview

Causal Inference in Human-Robot Spatial Interaction

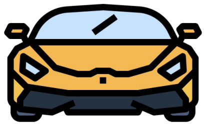
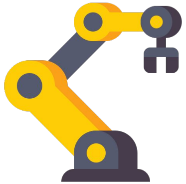
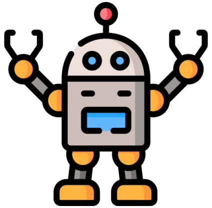
Causal Inference overview – why is it important?

Science that studies the cause-and-effect relationship between events

- **Cause**: I never brush my teeth. **Effect**: I have 5 cavities.
- **Cause**: I've smoked cigarettes daily for 20 years. **Effect**: I have lung cancer.

Humans reason causally

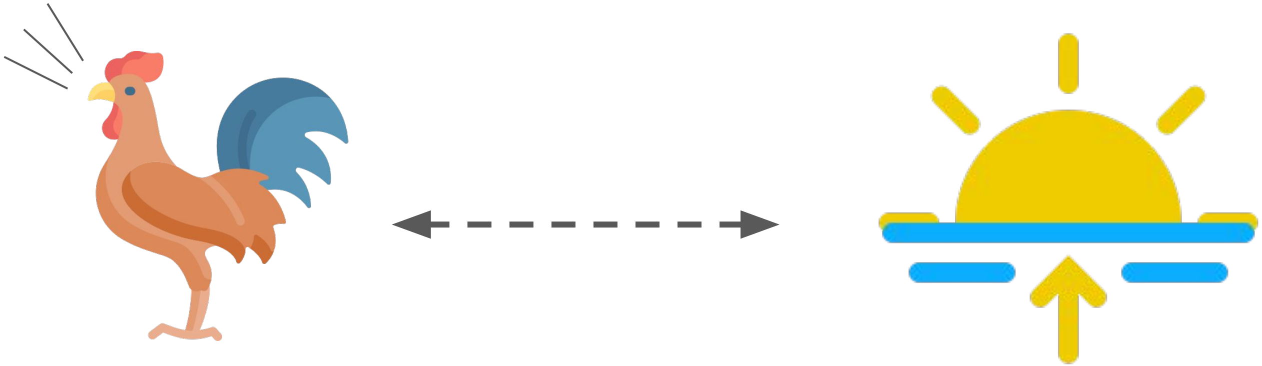
- What about machines? They simply analyse data


- Performing NON-causal analysis on data can lead to wrong relationships

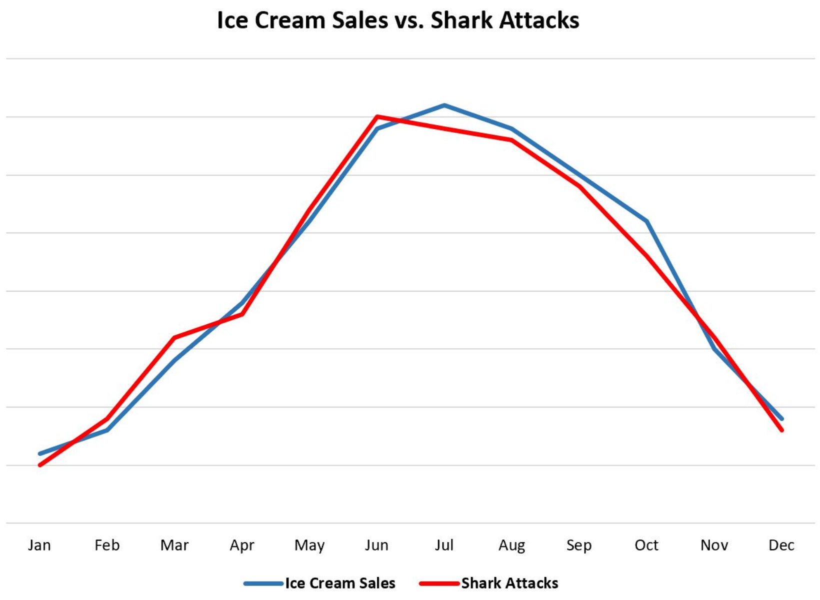
Causal Inference overview – why is it important?

The cockcrow is **strongly** <u>**associated**</u> with sunrise



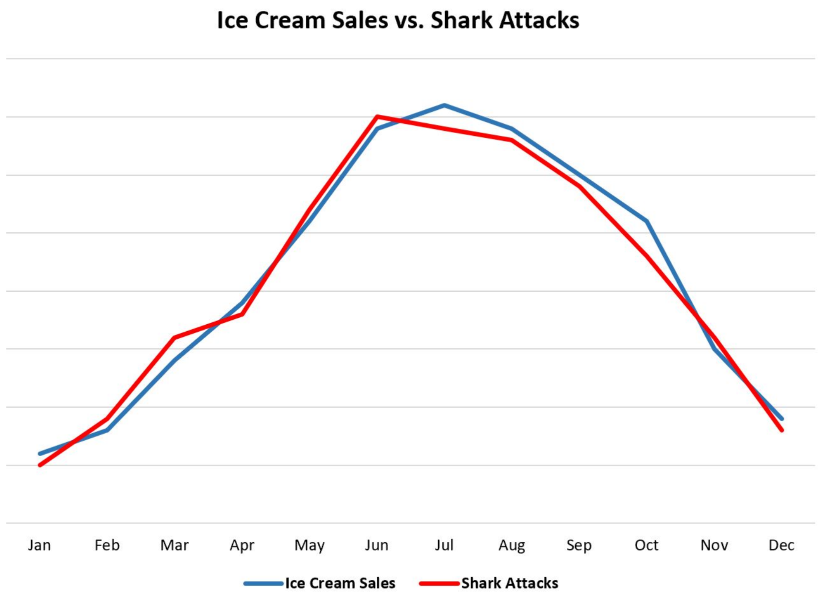
Causal Inference overview – why is it important?

The ice cream sales is **strongly <u>associated</u>** with shark attacks

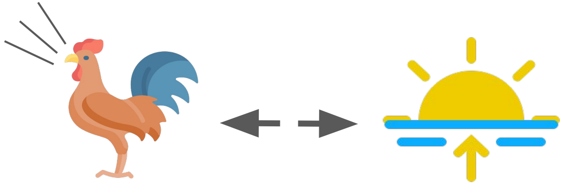


Causal Inference overview – why is it important?

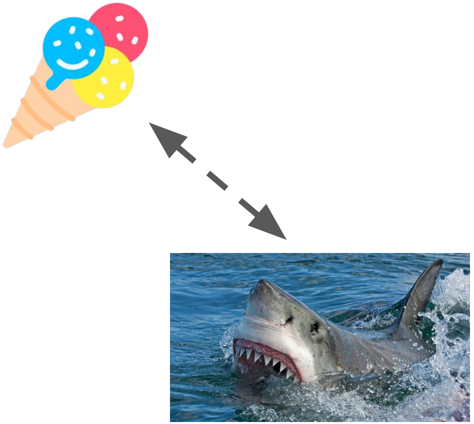
The ice cream sales is **strongly <u>associated</u>** with shark attacks



Causal Inference overview – why is it important?



Correlation
is not
Causation

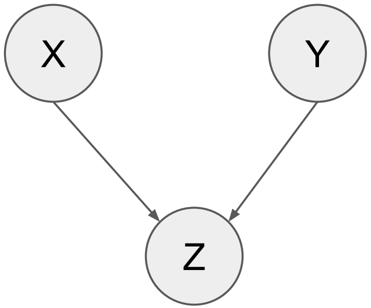


Causal Inference overview – causal discovery

- What are we trying to discover? A **causal** graph

  data → causal graph

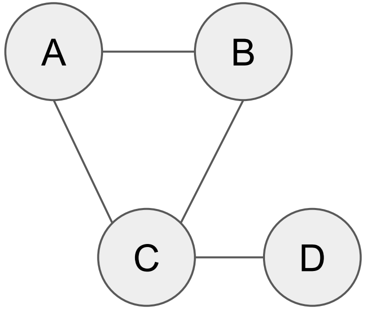
	X	Y	Z
0	20.000000	100.0	340.000000
1	20.204082	100.0	340.408163
2	20.408163	100.0	340.816327
3	20.612245	100.0	341.224490
4	20.816327	100.0	341.632653
...	...	...	...
2495	29.183673	300.0	958.367347
2496	29.387755	300.0	958.775510
2497	29.591837	300.0	959.183673
2498	29.795918	300.0	959.591837
2499	30.000000	300.0	960.000000



Causal Inference overview – causal discovery

- What are we trying to discover? A **causal** graph
  Let's start from…

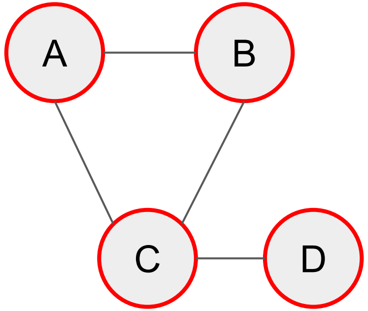
Undirected Graph



Causal Inference overview – causal discovery

- What are we trying to discover? A **causal** graph
  Let's start from…

Undirected Graph

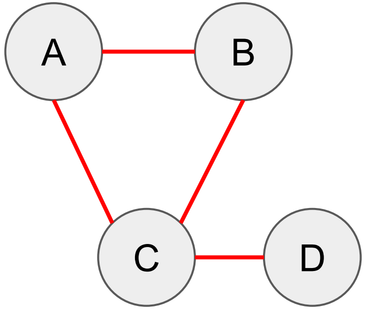


nodes

Causal Inference overview – causal discovery

- What are we trying to discover? A **causal** graph
  Let's start from…

Undirected Graph

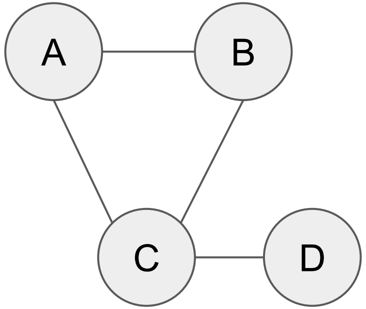


edges

Causal Inference overview – causal discovery

- What are we trying to discover? A **causal** graph
  Let's start from…

Undirected Graph

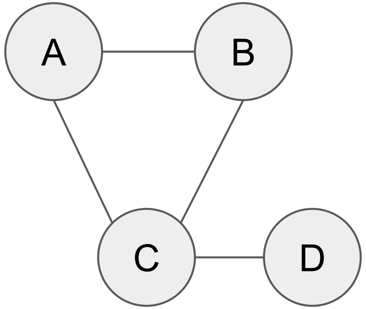


We don't want this

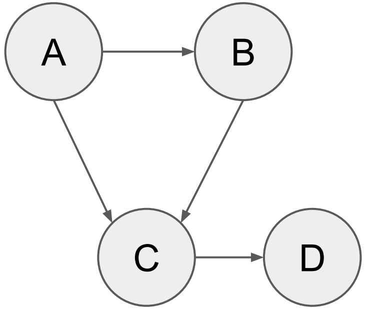
Causal Inference overview – causal discovery

- What are we trying to discover? A **causal** graph

Undirected Graph



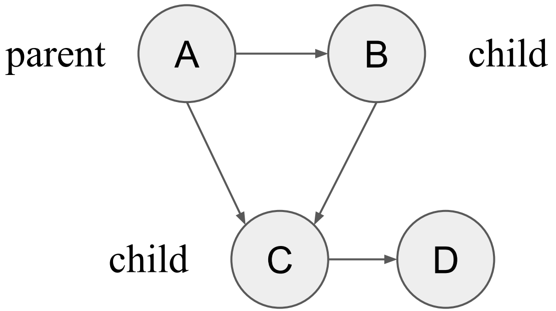
Directed Graph



Causal Inference overview – causal discovery

- What are we trying to discover? A **causal** graph

Directed Graph

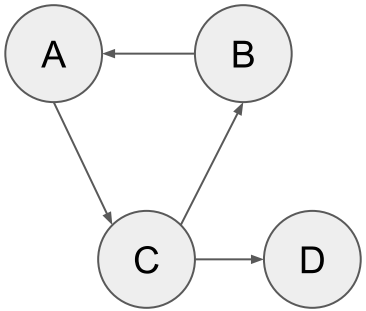


Causal Inference overview – causal discovery

- What are we trying to discover? A **causal** graph

We do not want cycles

Directed Graph



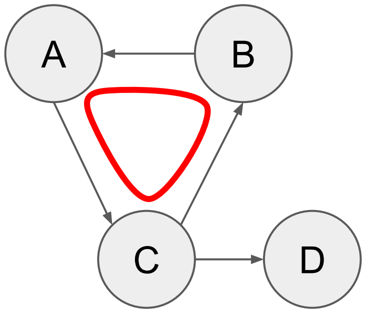
Causal Inference overview – causal discovery

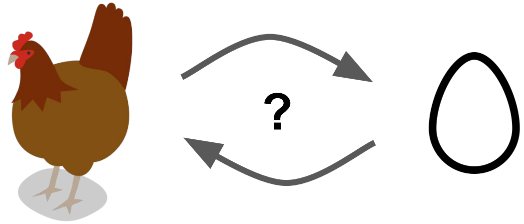
- What are we trying to discover? A **causal** graph

We do not want cycles

We need to guarantee the **acyclicity** assumption otherwise we can not distinguish between cause and effect

Directed Graph

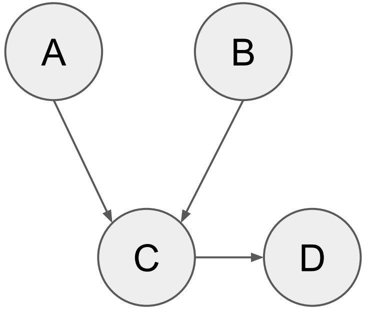




Causal Inference overview – causal discovery

- What are we trying to discover? A **causal** graph

Directed Acyclic Graph (**DAG**)



- **Why "Direct"?**

  We need oriented edges otherwise they do not represent cause-and-effect relationships

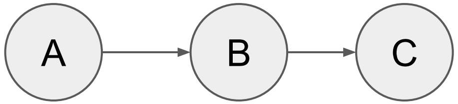
- **Why "Acyclic"?**

  We need to guarantee the **acyclicity** assumption otherwise we can not distinguish between cause and effect

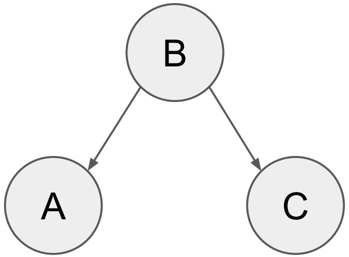
Causal Inference overview – causal discovery

- DAG configurations

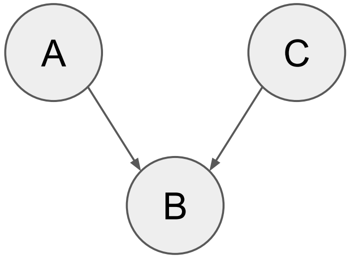
Chain



Fork



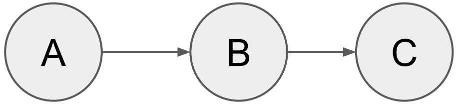
Collider



Causal Inference overview – causal discovery

- DAG configurations

Chain



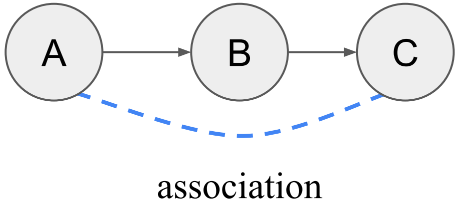
A is a direct cause of B
B is a direct cause of C

What about A and C?

Causal Inference overview – causal discovery

- DAG configurations

Chain



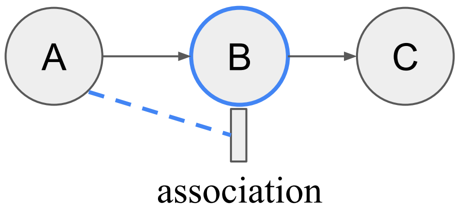
A is a direct cause of B
B is a direct cause of C

What about A and C?

They are associated (statistical dependent) through B

- DAG configurations

Chain



A is a direct cause of B
B is a direct cause of C

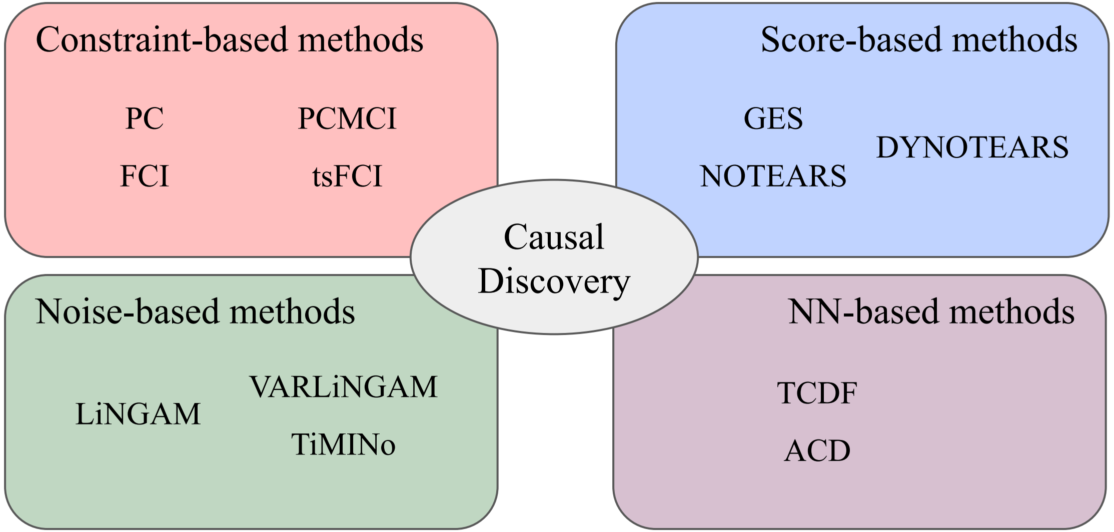
What about A and C?
They are associated (statistical dependent) through B

If we condition on B ⇒ A and C are conditionally independent

By conditioning on B, we are creating a **blocked** path. If we do not condition on B, the path from A to C is **unblocked** and the association is free to flow.

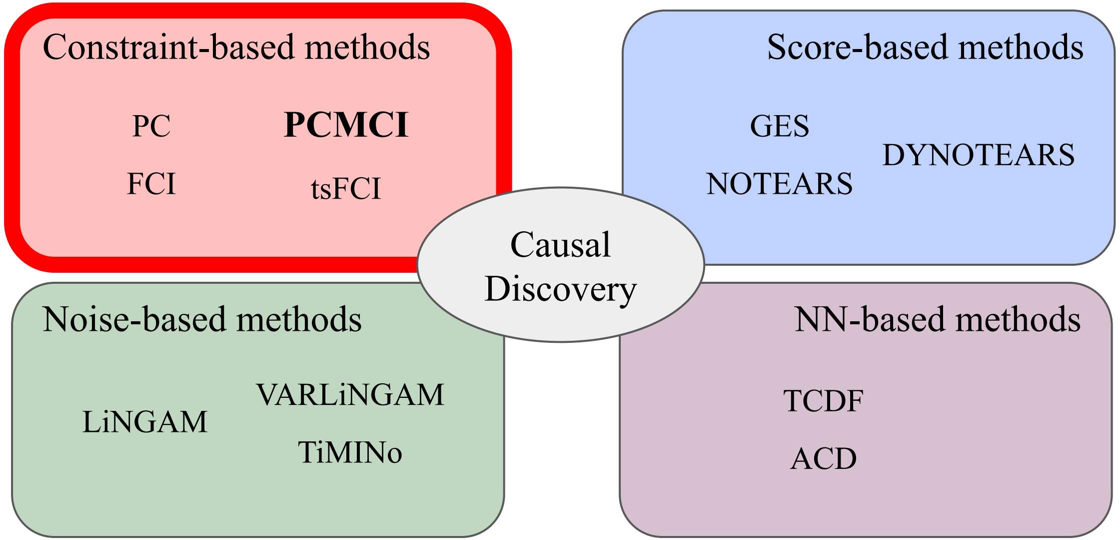
Causal Inference overview – causal discovery

- How can we discover a causal graph from observational data?



Causal Inference overview – causal discovery

- How can we discover a causal graph from observational data?



Causal Inference overview – causal discovery

- PCMCI algorithm

It consists of two main steps:

- **PC algorithm**
  retrieves the causal model structure by considering ONLY lagged dependencies as possible causal relationships between variables

- **MCI test**
  validates the structure found at the previous step by performing a false positive rate optimisation control

$$X^i_{t-\tau} \perp\!\!\!\perp X^j_t \,|\, \tilde{P}(X^i_{t-\tau}), \tilde{P}(X^j_t)$$

Key parameter: τ maximum time delay

Causal Inference overview – causal discovery

- PCMCI algorithm

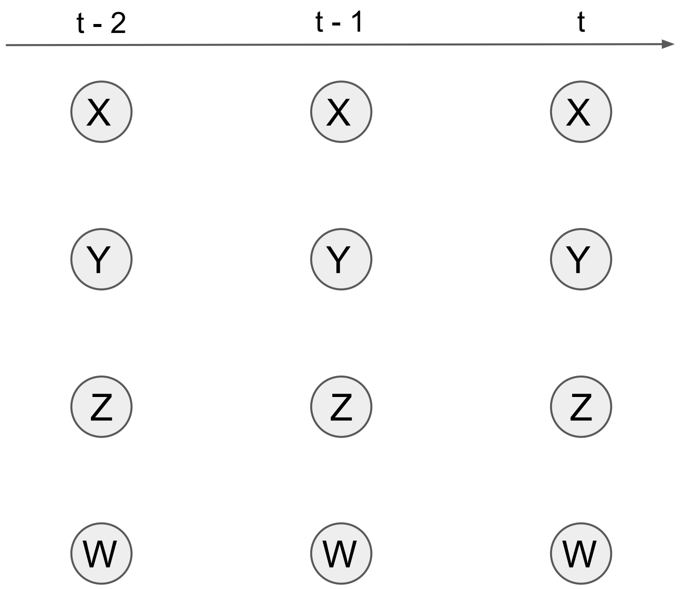
It consists of two main steps:

- **PC algorithm**
  retrieves the causal model structure by considering ONLY lagged dependencies as possible causal relationships between variables

- **MCI test**
  validates the structure found at the previous step by performing a false positive rate optimisation control

$$X_{t-\tau}^{i} \perp\!\!\!\perp X_{t}^{j} | \tilde{P}(X_{t-\tau}^{i}), \tilde{P}(X_{t}^{j})$$

Key parameter: τ maximum time delay



Causal Inference overview – causal discovery

- PCMCI algorithm

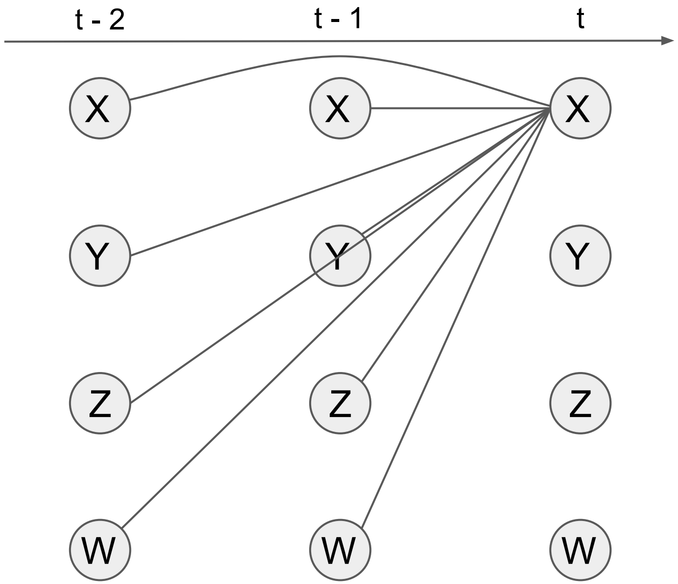
It consists of two main steps:

- **PC algorithm**
  retrieves the causal model structure by considering ONLY lagged dependencies as possible causal relationships between variables

- **MCI test**
  validates the structure found at the previous step by performing a false positive rate optimisation control

$$X^i_{t-\tau} \perp\!\!\!\perp X^j_t \,|\, \tilde{P}(X^i_{t-\tau}), \tilde{P}(X^j_t)$$

Key parameter: τ maximum time delay



Causal Inference overview – causal discovery

- PCMCI algorithm

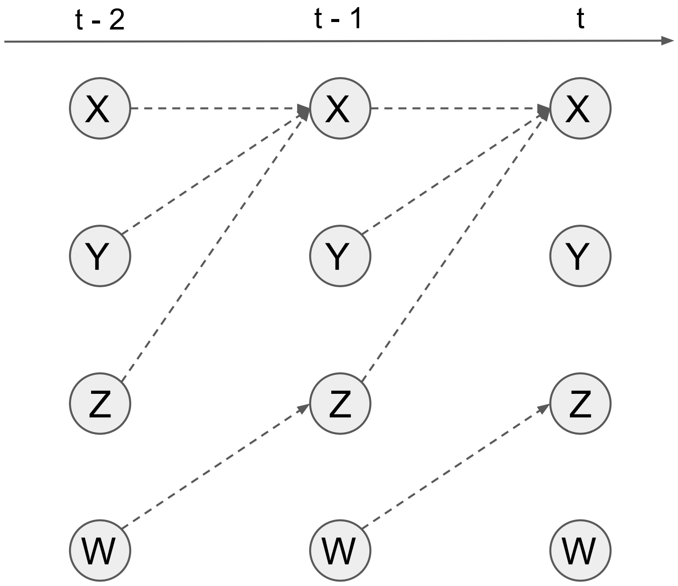
It consists of two main steps:

- **PC algorithm**
  retrieves the causal model structure by considering ONLY lagged dependencies as possible causal relationships between variables

- **MCI test**
  validates the structure found at the previous step by performing a false positive rate optimisation control

$$X^i_{t-\tau} \perp\!\!\!\perp X^j_t | \tilde{P}(X^i_{t-\tau}), \tilde{P}(X^j_t)$$

Key parameter: τ maximum time delay



Causal Inference overview – causal discovery

- PCMCI algorithm

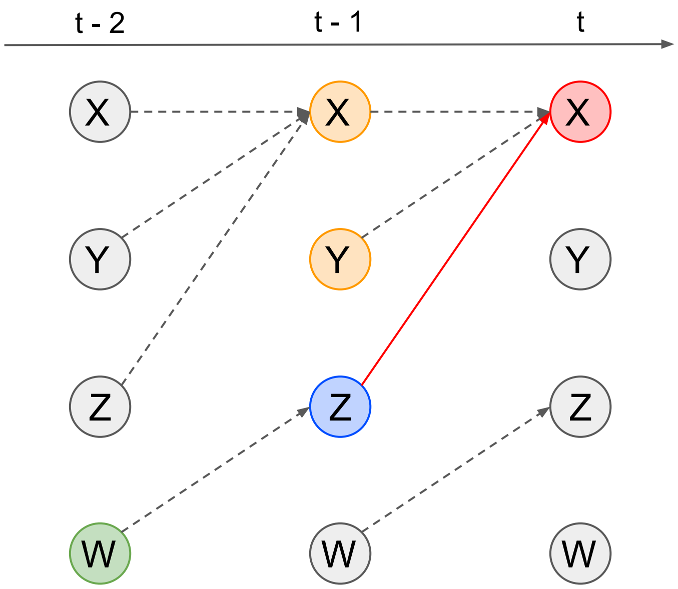
It consists of two main steps:

- **PC algorithm**
  retrieves the causal model structure by considering ONLY lagged dependencies as possible causal relationships between variables

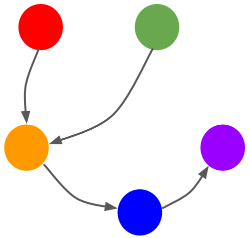
- **MCI test**
  validates the structure found at the previous step by performing a false positive rate optimisation control

$$X^i_{t-\tau} \perp\!\!\!\perp X^j_t \,|\, \tilde{P}(X^i_{t-\tau}), \tilde{P}(X^j_t)$$

Key parameter: τ maximum time delay



Outline

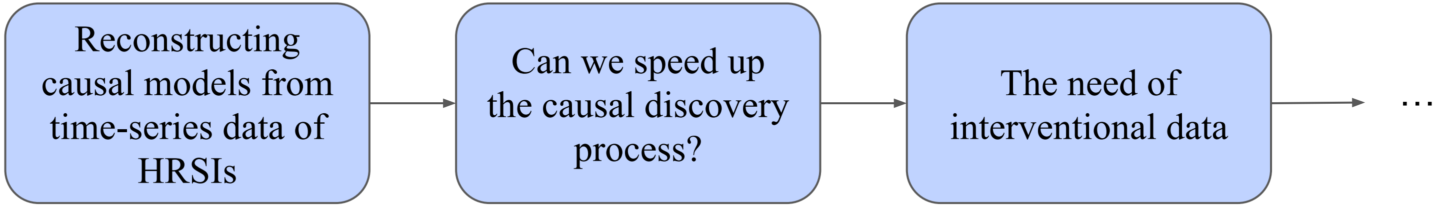


Motivation

Causal Inference overview

Causal Inference in Human-Robot Spatial Interaction

Causal Inference in Human-Robot Spatial Interaction



Causal Inference in Human-Robot Spatial Interaction

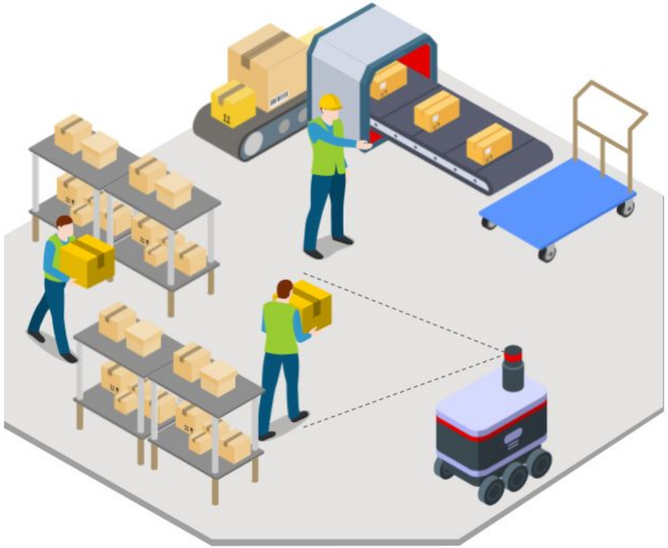
Reconstructing causal models from time-series data of HRSIs → Can we speed up the causal discovery process? → The need of interventional data → …

Causal Inference in Human-Robot Spatial Interaction

Reconstructing causal models from time-series data of HRSIs

Aim

enable the robot to understand human behaviours by discovering the cause-effect relationship between events during a Human-Robot Spatial Interaction (HRSI)



DARKO
Dynamic Agile Production Robots That Learn and Optimise Knowledge and Operations

Causal Inference in Human-Robot Spatial Interaction

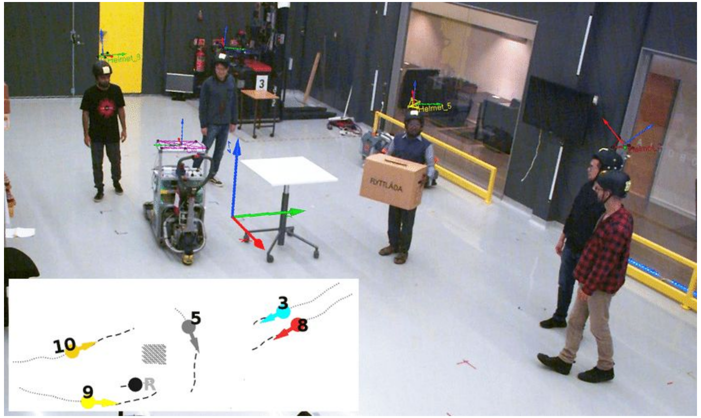
Reconstructing causal models from time-series data of HRSIs

Single-agent scenario
system variables:

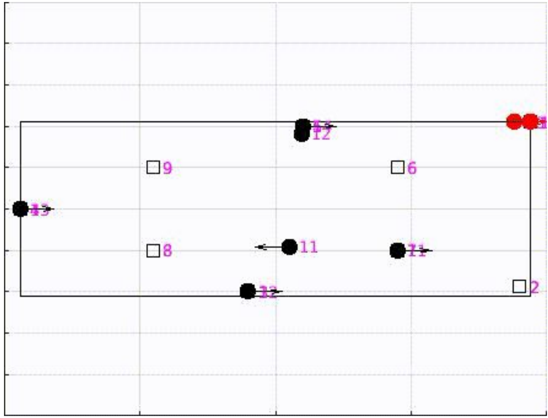
$$\theta_g, d_g, v$$

expected cause-effect relationships:

$$\theta_g = f(\theta_g, d_g)$$
$$d_g = f(d_g, \theta_g, v)$$
$$v = f(v, \theta_g)$$



THÖR Dataset [A. Rudenko et al. 2020]



Multi-agent scenario
system variables:

$$d_g, v, risk$$

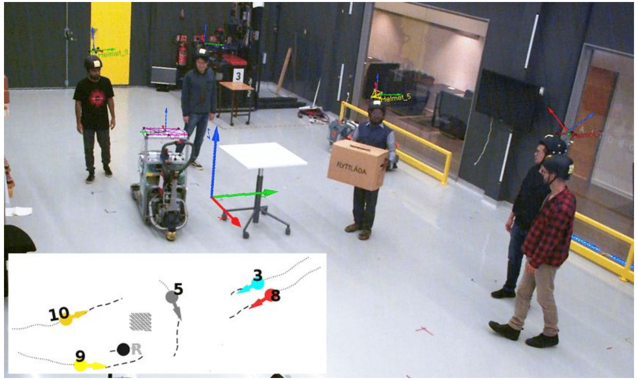
expected cause-effect relationships:

$$d_g = f(d_g, v)$$
$$v = f(v, d_g, risk)$$
$$risk = f(risk, v)$$

Causal Inference in Human-Robot Spatial Interaction

Reconstructing causal models from time-series data of HRSIs

THÖR
warehouse-like environment



ATC
shopping centre

SCENARIO			
DATASET		Single-agent	Multi-agent
	THÖR	X	X
	ATC	X	

Different scenario observed from the same dataset
➡ different causal models

Same scenario observed from different datasets
➡ same causal models

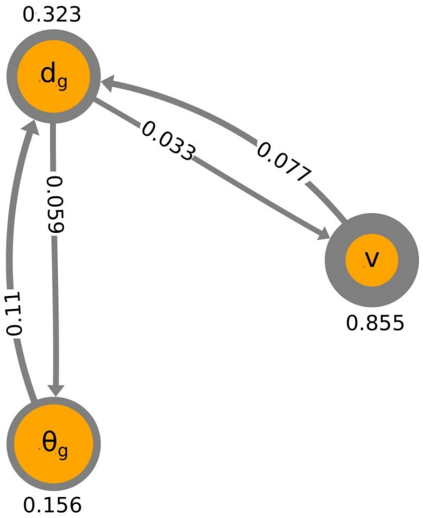
Different scenario observed from different datasets
➡ different causal models

Causal Inference in Human-Robot Spatial Interaction

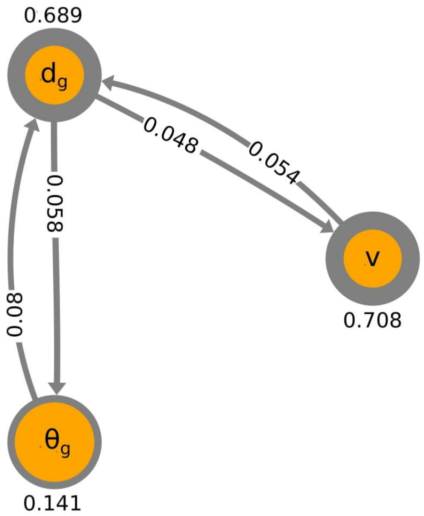
Reconstructing causal models from time-series data of HRSIs

$\tau = 1$

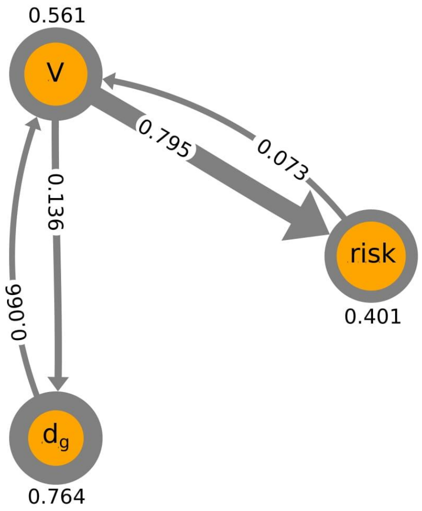
Single-agent THÖR



Single-agent ATC



Multi-agent THÖR



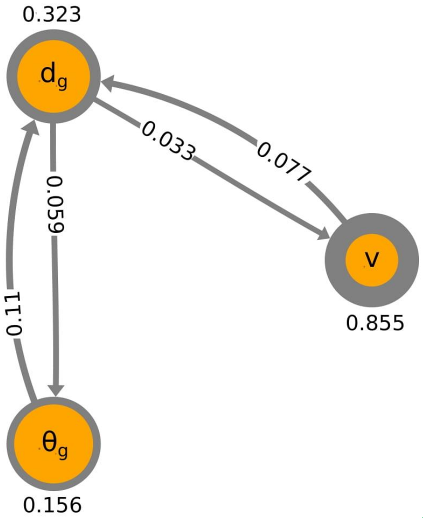
Causal Inference in Human-Robot Spatial Interaction

Reconstructing causal models from time-series data of HRSIs

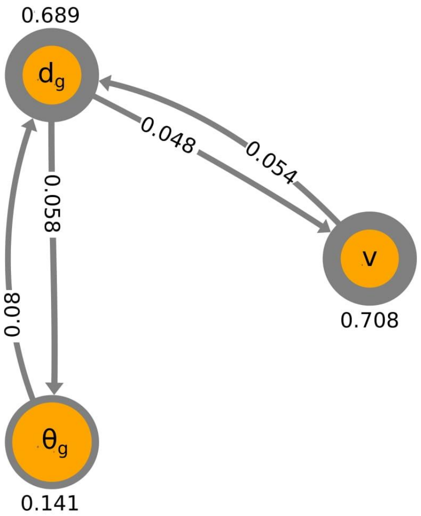
Different scenario observed from the same dataset
➡ different causal models

$\tau = 1$

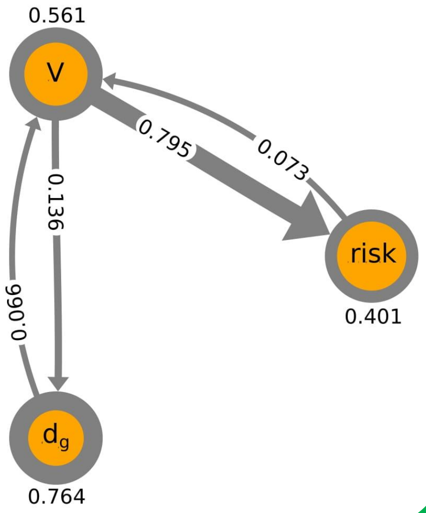
Single-agent
THÖR



Single-agent
ATC



Multi-agent
THÖR



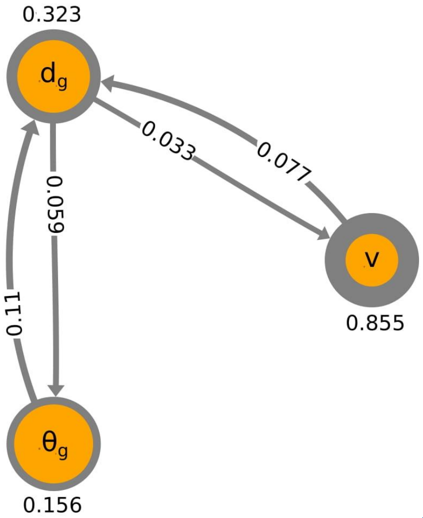
Causal Inference in Human-Robot Spatial Interaction

Reconstructing causal models from time-series data of HRSIs

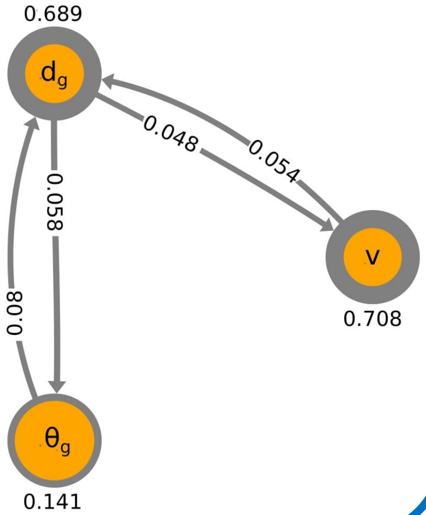
Same scenario observed from different datasets
➡ same causal models

τ = 1

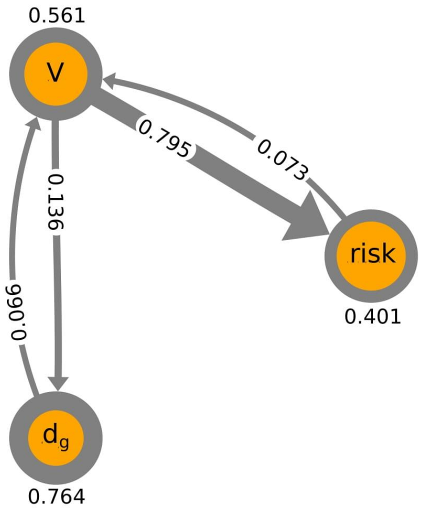
Single-agent THÖR



Single-agent ATC



Multi-agent THÖR



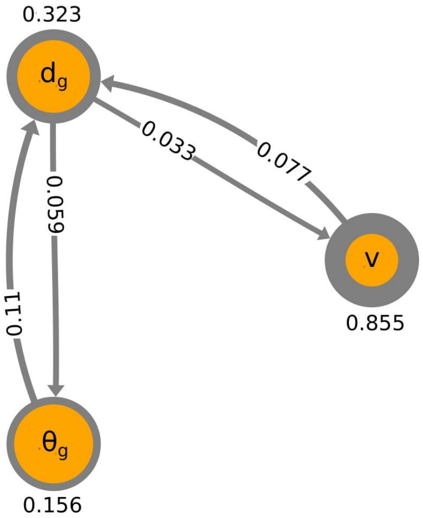
Causal Inference in Human-Robot Spatial Interaction

Reconstructing causal models from time-series data of HRSIs

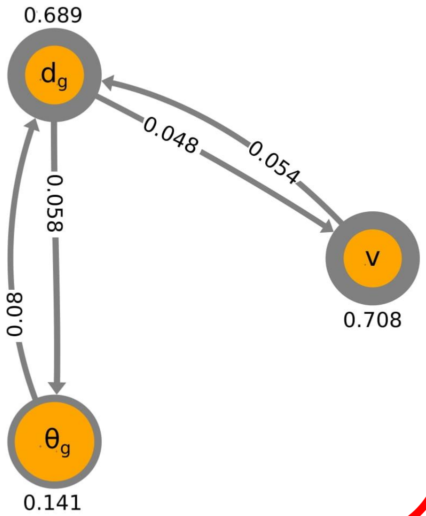
Same scenario observed from different datasets
➡ same causal models

$\tau = 1$

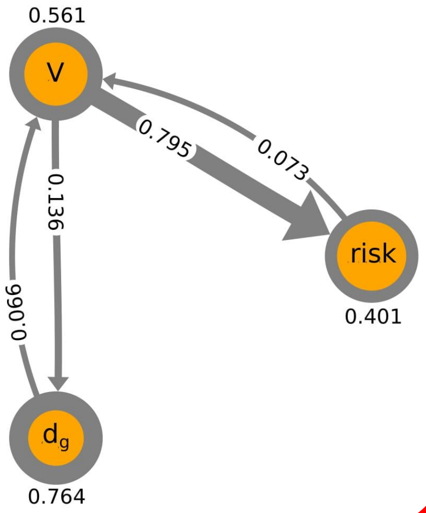
Single-agent THÖR



Single-agent ATC

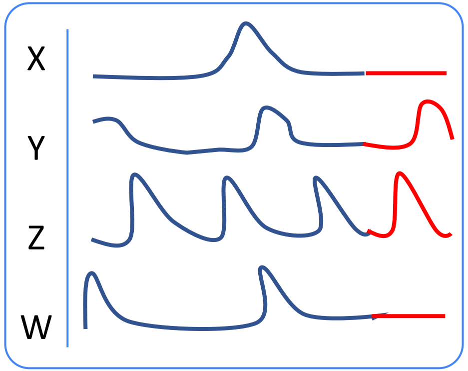


Multi-agent THÖR



Gaussian Process Regressor:

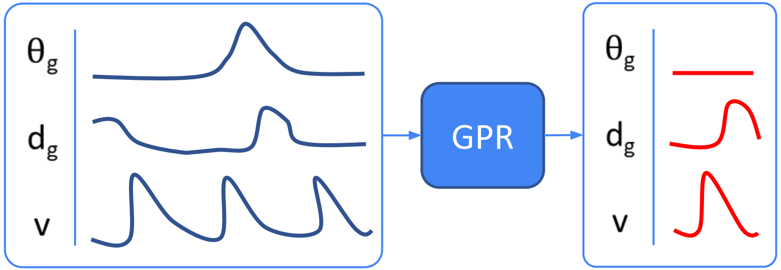
- supervised learning method designed to solve regression and probabilistic classification problems
- widely used for time-series prediction
- embedding the causal structure in the GPR ⇨ **Causal GPR**



Causal Inference in Human-Robot Spatial Interaction

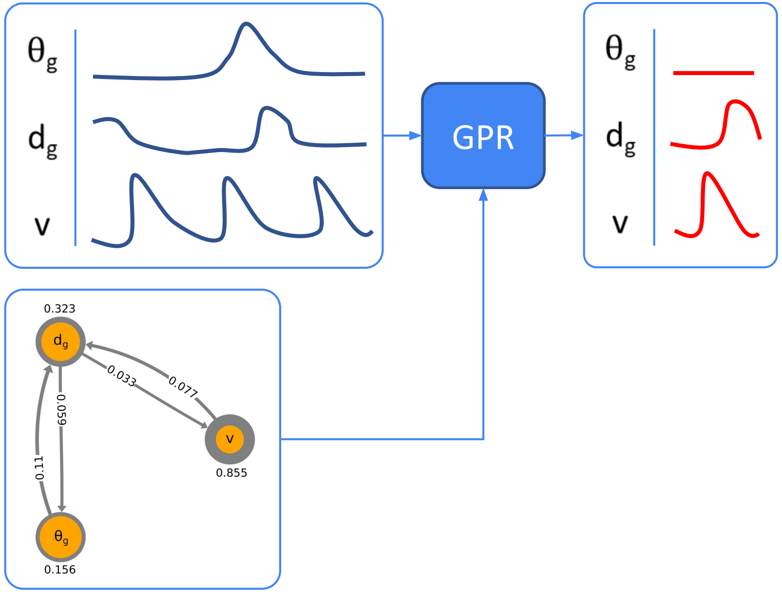
Reconstructing causal models from time-series data of HRSIs

Non-causal GPR approach



VS

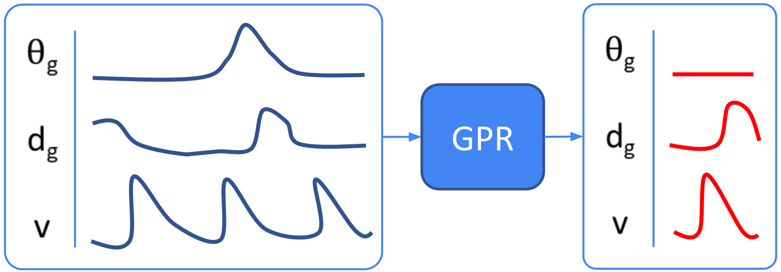
Causal GPR approach



Causal Inference in Human-Robot Spatial Interaction

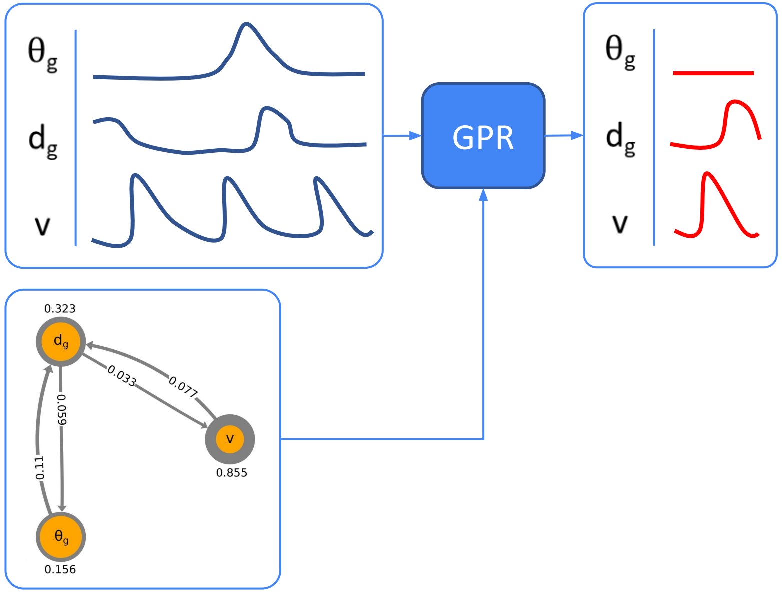
Reconstructing causal models from time-series data of HRSIs

Non-causal GPR approach



VS

Causal GPR approach

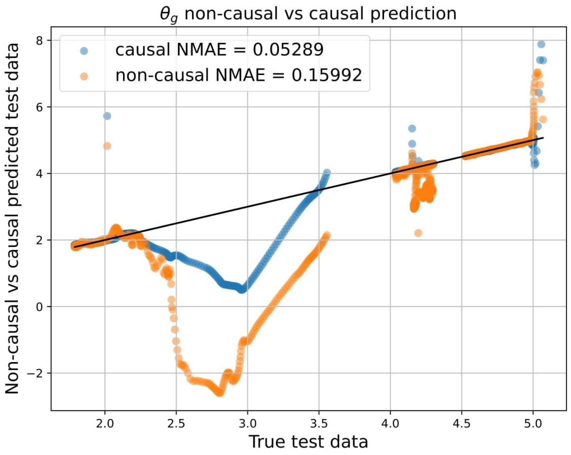


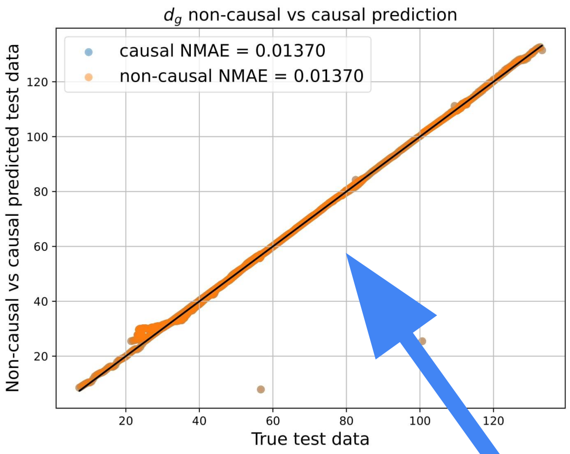
Evaluation metric

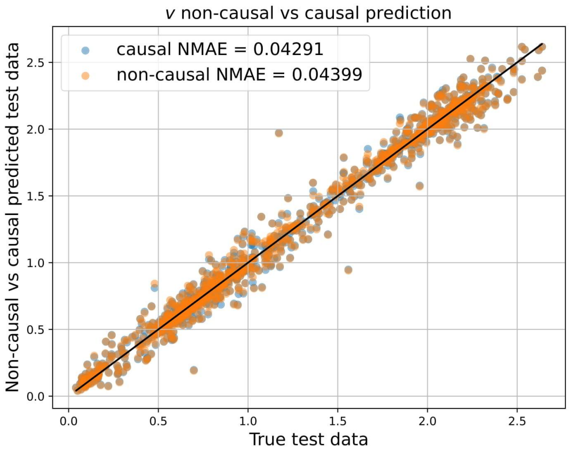
$$NMAE(y, \hat{y}) = \frac{\sum_{i=1}^{n} \frac{|y_i - \hat{y}_i|}{n}}{\frac{1}{n}\sum_{i=1}^{n} y_i}$$

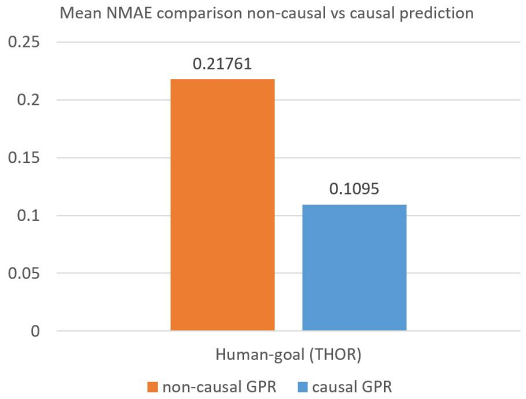
Causal Inference in Human-Robot Spatial Interaction

Reconstructing causal models from time-series data of HRSIs









$d_g(t) = f(\theta_g, d_g, v)(t-1)$
→ No difference between non-causal and causal GPR

Mean across the NMAE for each variable

Reconstructing causal models from time-series data of HRSIs

- Non-causal vs causal GPR comparison for the scenarios:
  - Single-agent (THÖR)
  - Single-agent (ATC)
  - Multi-agent (THÖR)

Mean NMAE	**Single-agent**		**Multi-agent**
	THÖR	ATC	THÖR
Non-causal GPR	0.21761	1.61692	0.37849
Causal GPR	**0.1095**	**1.54552**	**0.36453**

Causal Inference in Human-Robot Spatial Interaction

Reconstructing causal models from time-series data of HRSIs

- Non-causal vs causal GPR comparison for the scenarios:
  - Single-agent (THÖR)
  - Single-agent (ATC)
  - Multi-agent (THÖR)

Mean NMAE	**Single-agent**		**Multi-agent**
	THÖR	ATC	THÖR
Non-causal GPR	0.21761	1.61692	0.37849
Causal GPR	**0.1095**	**1.54552**	**0.36453**

~ -50%
prediction error

~ -4%
prediction error

Causal Inference in Human-Robot Spatial Interaction

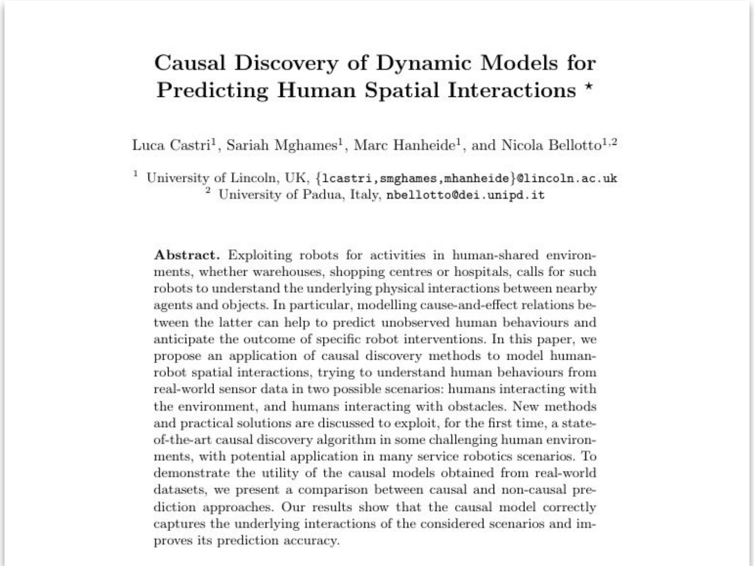
Reconstructing causal models from time-series data of HRSIs

Summing up

- First application of a causal discovery method to real-world sensor data for modelling HRSI
- New causal models from HRSI

Main limitation

- The PCMCI causal discovery is extremely demanding in terms of computational cost and hardware resources



Causal Discovery of Dynamic Models for Predicting Human Spatial Interactions *

Luca Castri[1], Sariah Mghames[1], Marc Hanheide[1], and Nicola Bellotto[1,2]

[1] University of Lincoln, UK, {lcastri,smghames,mhanheide}@lincoln.ac.uk
[2] University of Padua, Italy, nbellotto@dei.unipd.it

Abstract. Exploiting robots for activities in human-shared environments, whether warehouses, shopping centres or hospitals, calls for such robots to understand the underlying physical interactions between nearby agents and objects. In particular, modelling cause-and-effect relations between the latter can help to predict unobserved human behaviours and anticipate the outcome of specific robot interventions. In this paper, we propose an application of causal discovery methods to model human-robot spatial interactions, trying to understand human behaviours from real-world sensor data in two possible scenarios: humans interacting with the environment, and humans interacting with obstacles. New methods and practical solutions are discussed to exploit, for the first time, a state-of-the-art causal discovery algorithm in some challenging human environments, with potential application in many service robotics scenarios. To demonstrate the utility of the causal models obtained from real-world datasets, we present a comparison between causal and non-causal prediction approaches. Our results show that the causal model correctly captures the underlying interactions of the considered scenarios and improves its prediction accuracy.

L. Castri, S. Mghames, M. Hanheide, and N. Bellotto
“Causal discovery of dynamic models for predicting human spatial interactions,”
in International Conference on Social Robotics (ICSR), 2022.

Causal Inference in Human-Robot Spatial Interaction

Reconstructing causal models from time-series data of HRSIs → Can we speed up the causal discovery process? → The need of interventional data → …

Causal Inference in Human-Robot Spatial Interaction

Can we speed up the causal discovery process?

Motivation

Causal analysis of complex and dynamical systems is extremely demanding in terms of time and hardware resources, making it a challenge for autonomous robotics with limited hardware resources and real-time requirements.

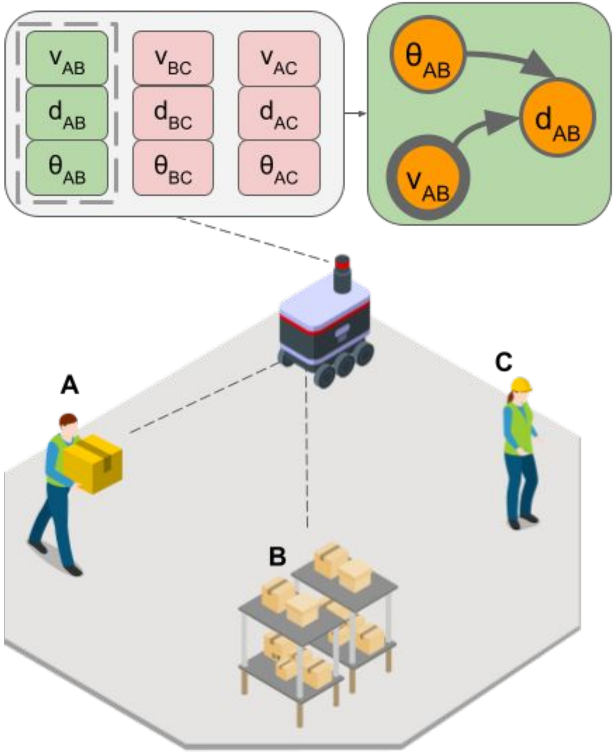
None of the state-of-the-art approaches extracts both the important features representing the system and the causal association between them, while at the same time taking into account the execution time and the computational cost for completing the task.

DARKO
Dynamic Agile Production Robots That Learn and Optimise Knowledge and Operations

Causal Inference in Human-Robot Spatial Interaction

Can we speed up the causal discovery process?

Are all the observable variables useful to understand the observed scenario?



DARKO
Dynamic Agile Production Robots That Learn and Optimise Knowledge and Operations

Causal Inference in Human-Robot Spatial Interaction

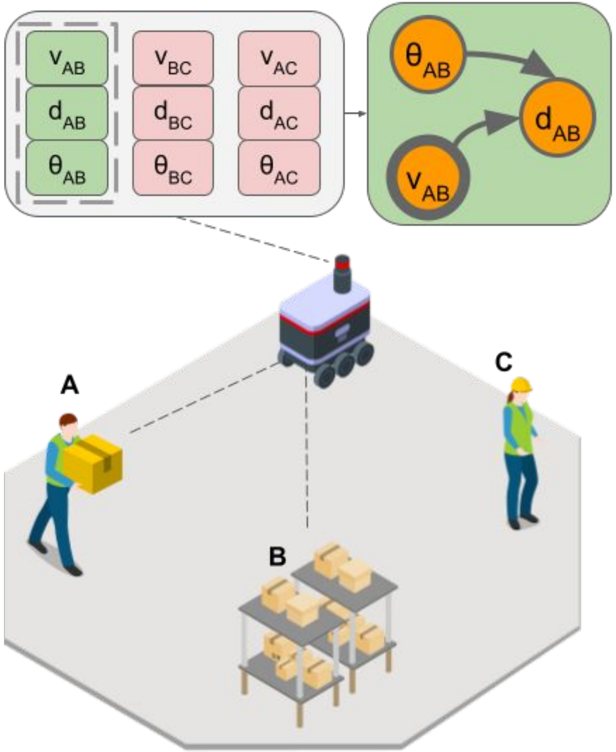
Can we speed up the causal discovery process?

Aim

Create an all-in-one algorithm able to:

- select the most meaningful features from a prefixed set of variables
- build a causal model from such selection

in order to enhance speed and accuracy of the causal discovery and make it more efficient and feasible for robotics applications.



DARKO
Dynamic Agile Production Robots That Learn and Optimise Knowledge and Operations

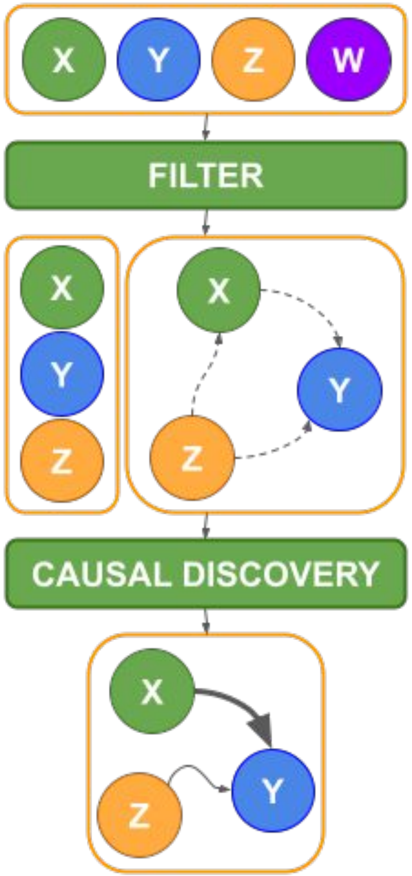
Causal Inference in Human-Robot Spatial Interaction

Can we speed up the causal discovery process?

F-PCMCI steps:

- takes in input a prefixed set of variables
- the Transfer Entropy-based filter analyses and removes irrelevant variables (e.g., constants or isolated ones). The reduced variable set is used to create a hypothetical causal model
- the latter needs to be validated by a proper causal analysis, which is performed by the PCMCI causal discovery algorithm

This strategy enables **faster** and **more accurate** causal discovery



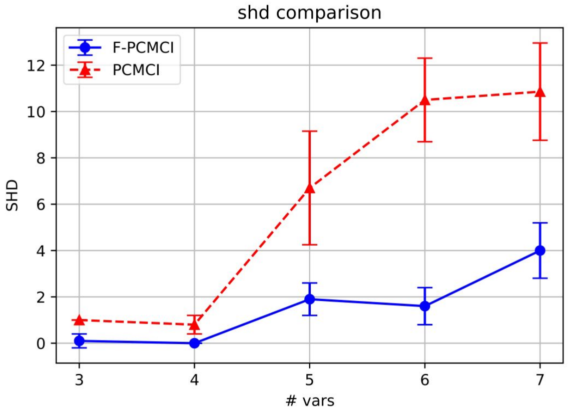
PCMCI vs F-PCMCI

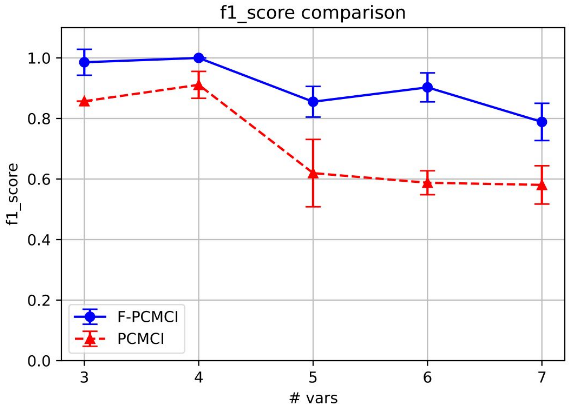
The correctness of our approach was evaluated based on toy problems with known ground-truth causal models. Various types of dependencies:

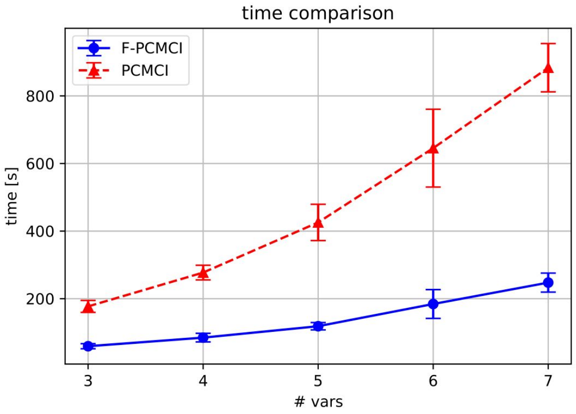
- linear and non-linear cross- and auto-dependency;
- noise-only equations;
- independent and dependent equations;
- different time-lag dependencies.

The analysis was carried out considering a number of system variables varying between 3 and 7. For each configuration, we performed 10 run tests with random system coefficients, using as evaluation metrics the mean over all the tests of:

- Structural Hamming Distance SHD;
- F1-score;
- execution time (in secs).



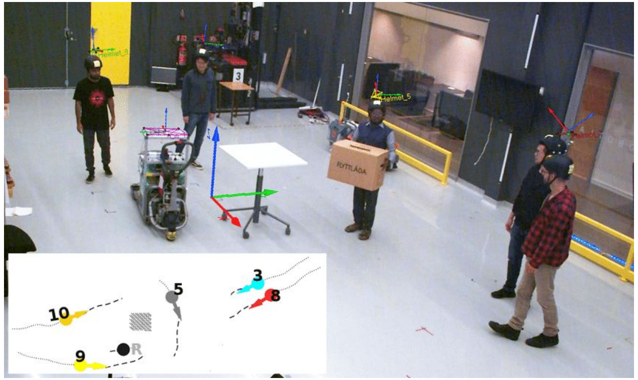




Causal Inference in Human-Robot Spatial Interaction

Can we speed up the causal discovery process?

Modeling Real-world Human Spatial Interactions



We used our approach to model and predict spatial

interactions. This application involves three steps:

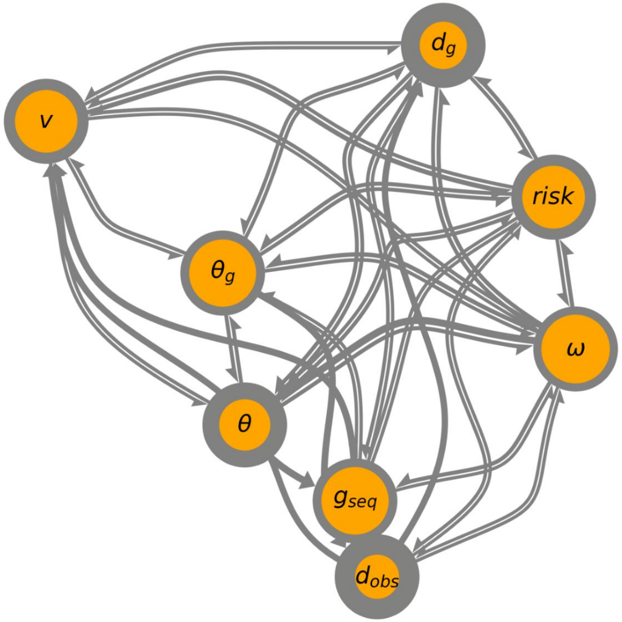
- extracting time-series of sensor data from human spatial interaction scenarios using the THÖR dataset;
- reconstruct the causal model using F-PCMCI;
- embedding the causal model in a LSTM-based prediction system.

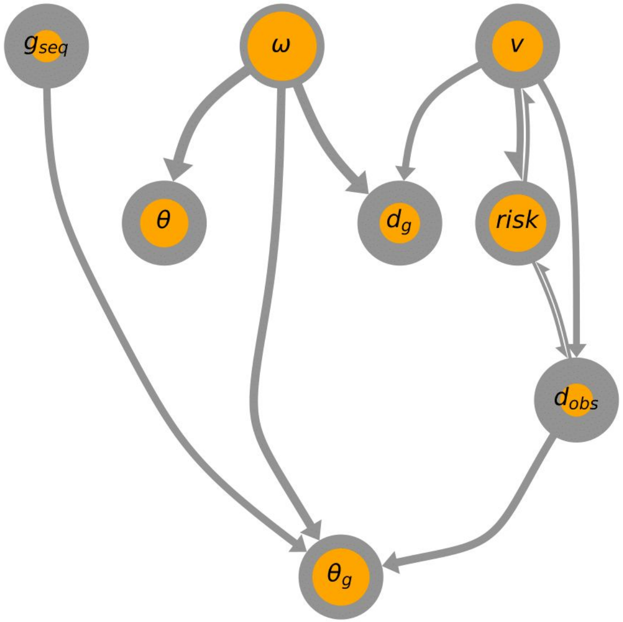
In order to represent human spatial interactions, for each agent we considered 8 variables, which were then used in the causal analysis.

$\mathbf{d_g}$	distance to goal	$\boldsymbol{\theta_g}$	angle to goal
$\mathbf{v}$	velocity	$\boldsymbol{\omega}$	angular velocity
risk	collision risk	$\mathbf{g_{seq}}$	goal position sequence
$\boldsymbol{\theta}$	orientation	$\mathbf{d_{obs}}$	distance to closest obstacle

Can we speed up the causal discovery process?

Modeling Real-world Human Spatial Interactions



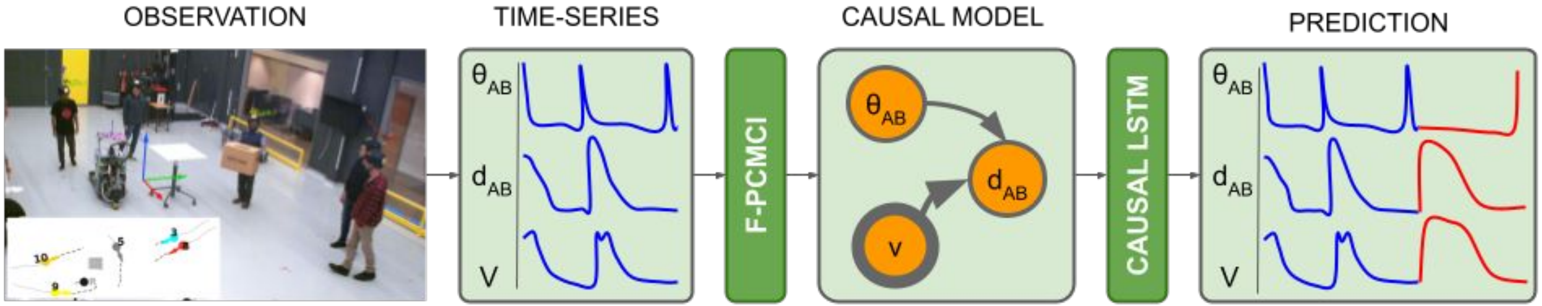


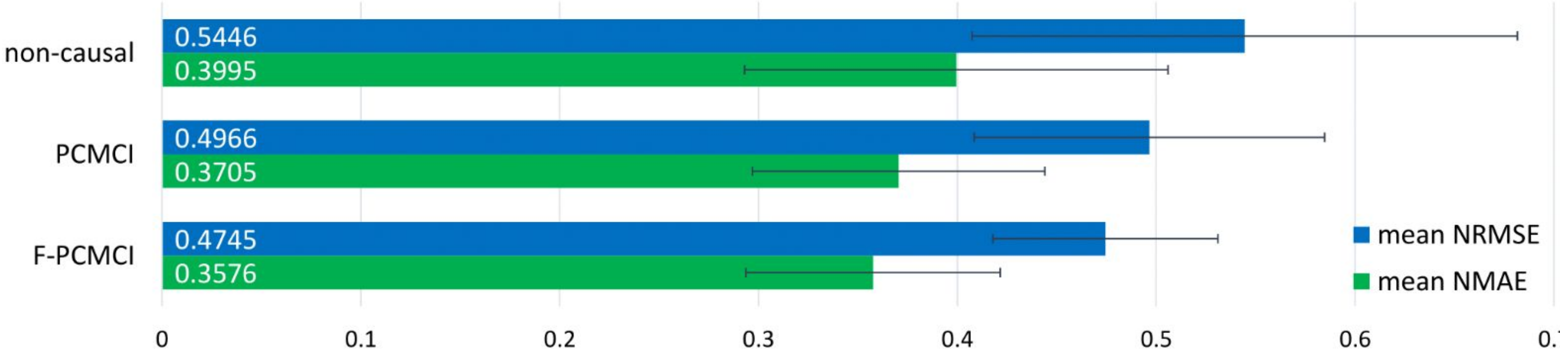
PCMCI execution time 79’45”	**F-PCMCI execution time 17’33”**

Causal Inference in Human-Robot Spatial Interaction

Can we speed up the causal discovery process?

Modeling Real-world Human Spatial Interactions





Can we speed up the causal discovery process?

Summing up

- We extended and improved a state-of-the-art causal discovery algorithm, PCMCI, embedding an additional feature-selection module based on transfer entropy

- **F-PCMCI**
  https://github.com/lcastri/fpcmci
  `pip install fpcmci`

Main limitation

- We are not exploiting the full power of causal inference: **the intervention**

Proceedings of Machine Learning Research vol 213:1–16, 2023 2nd Conference on Causal Learning and Reasoning

Enhancing Causal Discovery from Robot Sensor Data in Dynamic Scenarios

Luca Castri LCASTRI@LINCOLN.AC.UK
Sariah Mghames SMGHAMES@LINCOLN.AC.UK
Marc Hanheide MHANHEIDE@LINCOLN.AC.UK
University of Lincoln, UK

Nicola Bellotto NBELLOTTO@DEI.UNIPD.IT
University of Padua, Italy

Editors: Mihaela van der Schaar, Dominik Janzing and Cheng Zhang

Abstract

Identifying the main features and learning the causal relationships of a dynamic system from time-series of sensor data are key problems in many real-world robot applications. In this paper, we propose an extension of a state-of-the-art causal discovery method, PCMCI, embedding an additional feature-selection module based on transfer entropy. Starting from a prefixed set of variables, the new algorithm reconstructs the causal model of the observed system by considering only its main features and neglecting those deemed unnecessary for understanding the evolution of the system. We first validate the method on a toy problem and on synthetic data of brain network, for which the ground-truth models are available, and then on a real-world robotics scenario using a large-scale time-series dataset of human trajectories. The experiments demonstrate that our solution outperforms the previous state-of-the-art technique in terms of accuracy and computational efficiency, allowing better and faster causal discovery of meaningful models from robot sensor data.

Keywords: causal discovery, feature selection, time-series, transfer entropy, causal robotics.

L. Castri, S. Mghames, M. Hanheide, and N. Bellotto, "Enhancing causal discovery from robot sensor data in dynamic scenarios," in Conference on Causal Learning and Reasoning, 2023.

Causal Inference in Human-Robot Spatial Interaction

Reconstructing causal models from time-series data of HRSIs → Can we speed up the causal discovery process? → The need of interventional data → …

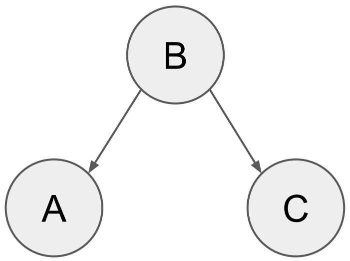
Motivation

a primary limitation of the current state-of-the-art algorithms is that they can only handle **observational data**.

The latter are often insufficient to retrieve the correct causal model in complex scenarios where it is impossible to account for all the variables responsible for the system's evolution. In such cases, data from experiments, i.e. **interventional data**, are needed to eliminate spurious links and enhance the quality of the causal model

Example

Hidden confounder scenario



The need of interventional data

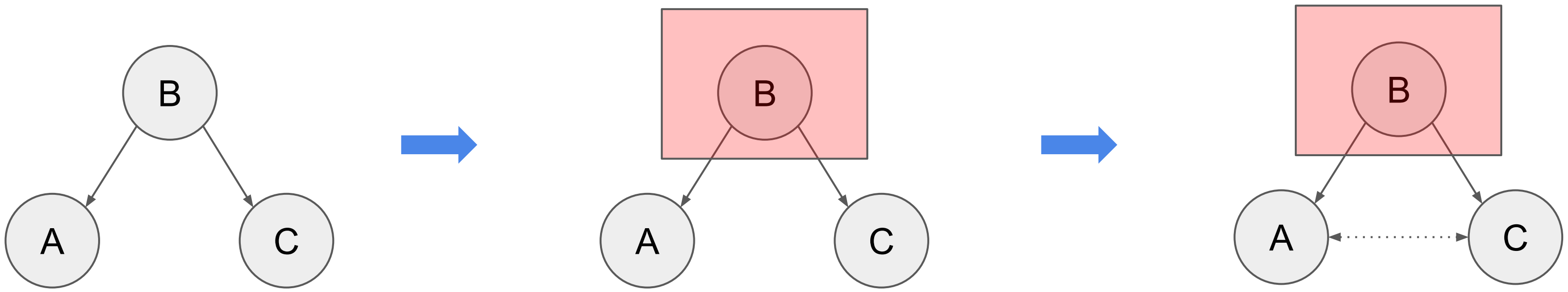
Motivation

a primary limitation of the current state-of-the-art algorithms is that they can only handle **observational data**.

The latter are often insufficient to retrieve the correct causal model in complex scenarios where it is impossible to account for all the variables responsible for the system's evolution. In such cases, data from experiments, i.e. **interventional data**, are needed to eliminate spurious links and enhance the quality of the causal model

Example

Hidden confounder scenario



The need of interventional data

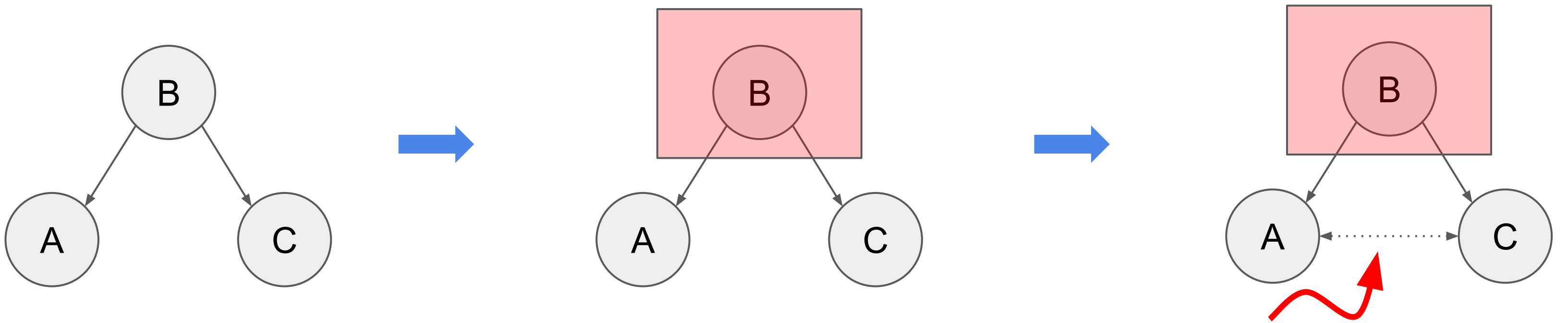
Motivation

a primary limitation of the current state-of-the-art algorithms is that they can only handle **observational data**.

The latter are often insufficient to retrieve the correct causal model in complex scenarios where it is impossible to account for all the variables responsible for the system's evolution. In such cases, data from experiments, i.e. **interventional data**, are needed to eliminate spurious links and enhance the quality of the causal model

Example

Hidden confounder scenario



Interventional data are needed to remove spurious links

The need of interventional data

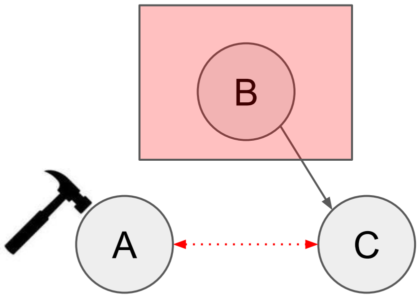
Motivation

a primary limitation of the current state-of-the-art algorithms is that they can only handle **observational data**.

The latter are often insufficient to retrieve the correct causal model in complex scenarios where it is impossible to account for all the variables responsible for the system's evolution. In such cases, data from experiments, i.e. **interventional data**, are needed to eliminate spurious links and enhance the quality of the causal model

Example

Hidden confounder scenario



Performing an intervention on A, i.e. forcing its value:

- breaks the input links to A
- we can study the effect on C
- If varying A does not lead any change on C ➡ A-C link is spurious

Thank you

questions?

Luca Castri,
lcastri@lincoln.ac.uk
PhD student
University of Lincoln

Website: ***https://darko-project.eu***
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101017274