TECHNISCHE UNIVERSITÄT DARMSTADT

On the Tractability of Inference for the Spectrum of Causal Models

Matej Zečević

Eurandom

Workshop Centre in the area of Stochastics

YES Causal Inference, 15th March 2023, Eindhoven

TECHNISCHE UNIVERSITÄT DARMSTADT

Not All Causal Inference is the Same

~~On the Tractability of Inference for the Spectrum of Causal Models~~

Matej Zečević

Eurandom
Workshop Centre in the area of Stochastics

YES Causal Inference, 15th March 2023, Eindhoven

Causal Inference

modelling assumptions **outside** **the data**
identification & estimation
graph learning
etc.

Causal Inference **in/for/with AI (implied ML)**

get these machines finally ‘smart’

attributions/explanations

robust/invariant predictions

reasoning

etc.

TECHNISCHE UNIVERSITÄT DARMSTADT

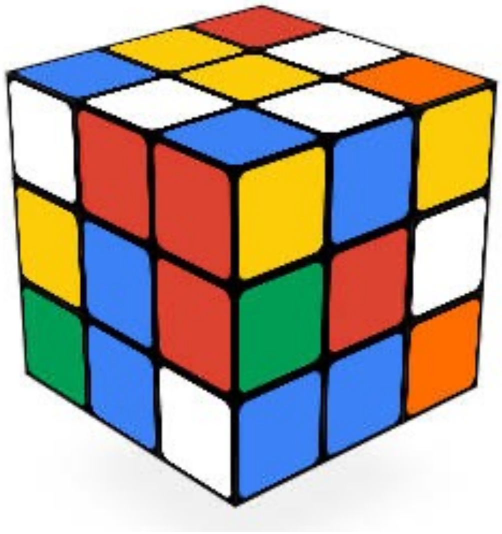
Causal **Inference** in/for/with AI (implied ML)

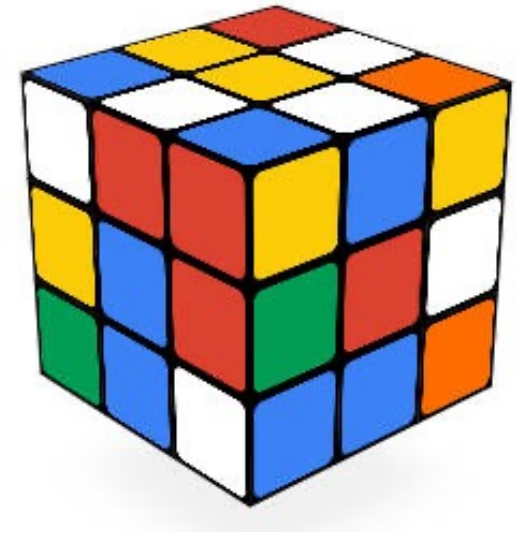
get the answers we want

get them **quick**

TECHNISCHE UNIVERSITÄT DARMSTADT

3 x 3 x 3 cube

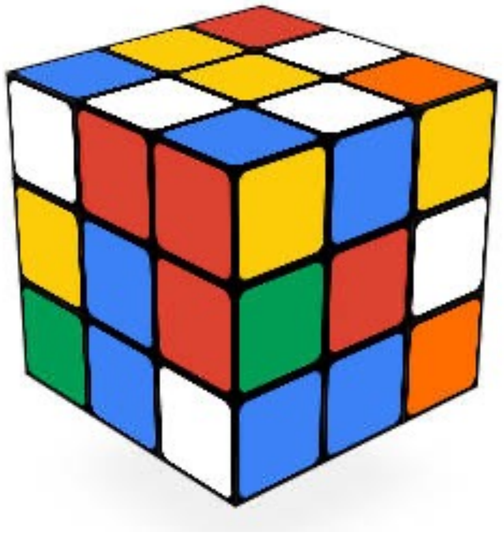




3 x 3 x 3 cube

43,252,003,274,489,856,000 potential positions

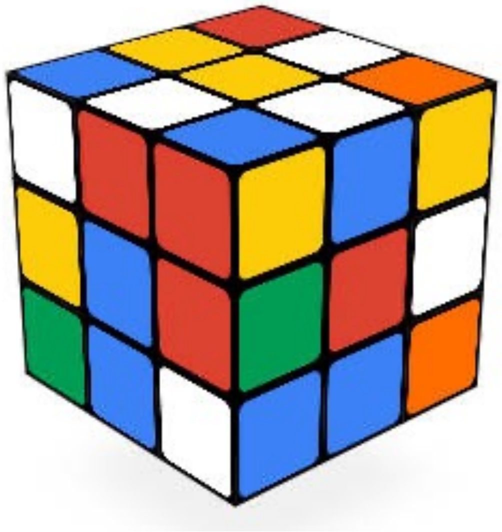
TECHNISCHE UNIVERSITÄT DARMSTADT



3 x 3 x 3 cube

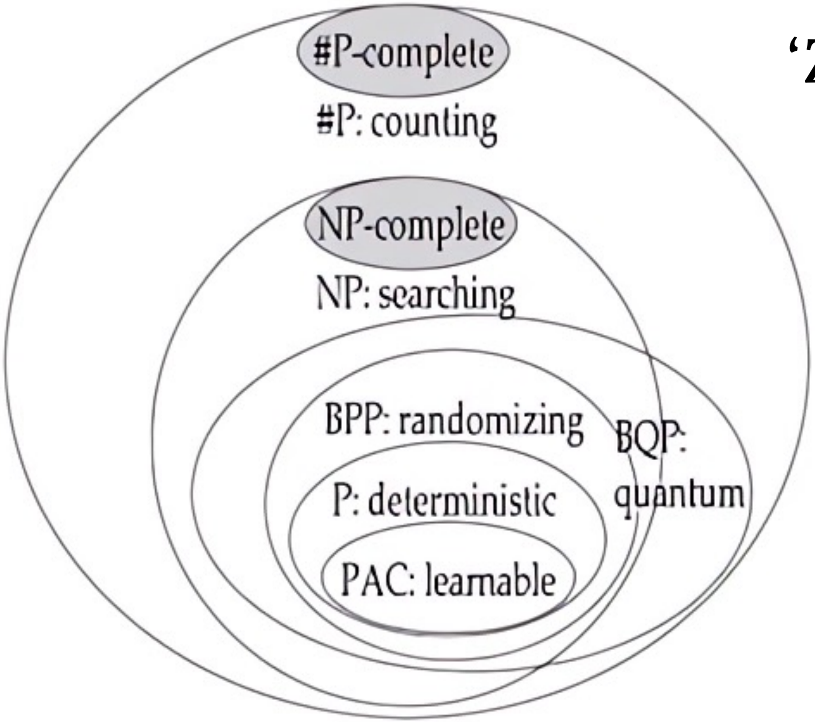
43,252,003,274,489,856,000 potential positions

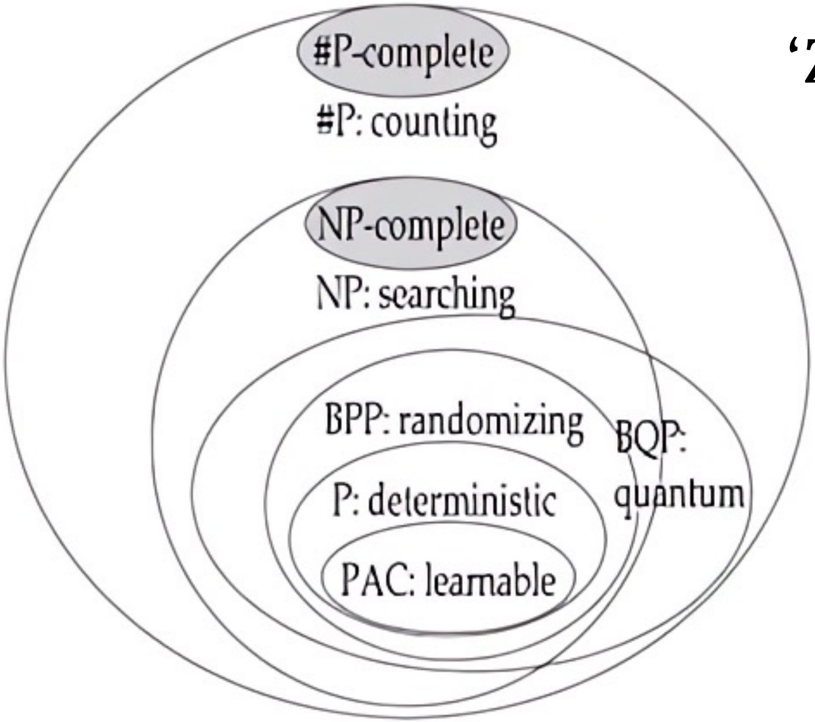
max. 20 moves from any position necessary



NP-hard

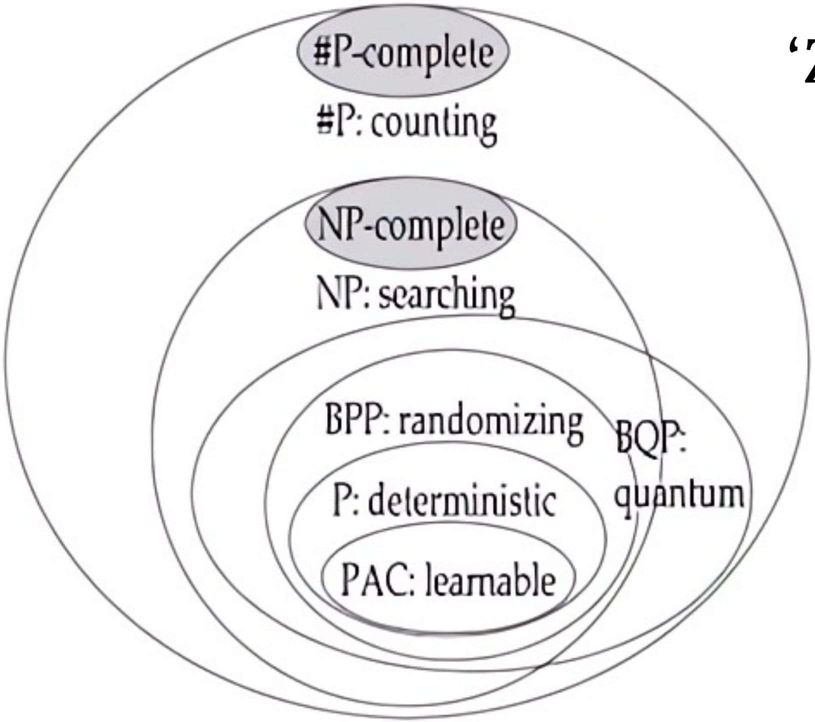
‘Zoo’ of Complexity Classes





‘Zoo’ of Complexity Classes

NP-hard,
#P-hard, etc.



‘Zoo’ of Complexity Classes

NP-hard,
#P-hard, etc.

“<problem> is
<complexity>-hard”
= <problem> solves all in that class

This talk is brief and not too formal…

This talk is brief and not too formal…

…since doing that would be: **Very-hard**

sorry that pun was necessary

TECHNISCHE UNIVERSITÄT DARMSTADT

For technical details and all the references,
please consider the preliminary paper:

https://arxiv.org/pdf/**2110.12052**.pdf

This talk is brief and not too formal…

…since doing that would be: **Very-hard**

sorry that pun was necessary

TECHNISCHE
UNIVERSITÄT
DARMSTADT

Justifying to your Boss *the wrong way*

You to advisor:
"Cannot solve this, I'm too dumb."

Justifying to your Boss *the right way*

You to advisor:
"I cannot solve this,
but neither can these people."

TECHNISCHE UNIVERSITÄT DARMSTADT

In this sense, being **NP-hard** is a good thing ;-)

Tractability

Definition. A class of queries Q is tractable on a family of probabilistic models M
iff for any query $q \in Q$ and model $m \in M$
exactly computing $q(m)$ runs in time $\mathbf{O(poly(|m|))}$.

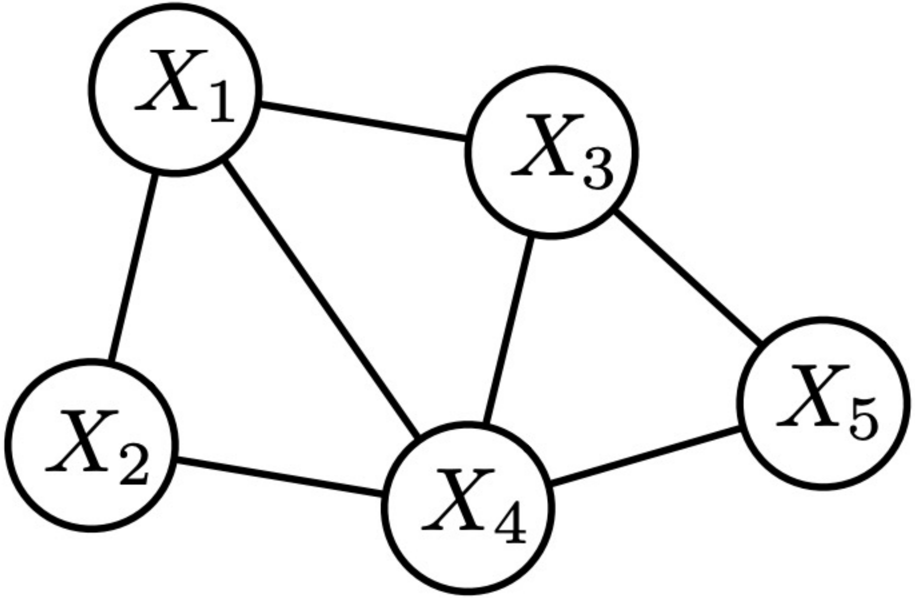
Tractability

Definition. A class of queries Q is tractable on a family of probabilistic models M iff for any query $q \in Q$ and model $m \in M$ **exactly** computing $q(m)$ runs in time $\mathbf{O(poly(|m|))}$.

cubic $O(x^3)$, high-degree $O(x^{2023})$, linear $O(x)$

Known results for PGM incl. Bayesian Nets (1990s):

exact marginal inference is #P-hard
approximate is NP-hard



What is the **landscape** of tractability of inference in causal models?

TECHNISCHE UNIVERSITÄT DARMSTADT

What is the **landscape** of tractability of inference
in causal models?

In the cases where it is bad…
what kind of design choices for **tradeoffs** are available?

TECHNISCHE
UNIVERSITÄT
DARMSTADT

Overview

I Are There Even Different Types of Causal Models?

Overview

1 Are There Even Different Types of Causal Models?

2 Inference in Non-Causal (i),

TECHNISCHE UNIVERSITÄT DARMSTADT

Overview

1 Are There Even Different Types of Causal Models?

2 Inference in Non-Causal (i),
Partially and Structural Causal Models (ii)

TECHNISCHE
UNIVERSITÄT
DARMSTADT

Overview

1 Are There Even Different Types of Causal Models?

2 Inference in Non-Causal (i),
Partially and Structural Causal Models (ii)

3 Summarizing the Key Differences

TECHNISCHE
UNIVERSITÄT
DARMSTADT

Overview

1 Are There Even Different Types of Causal Models?

2 Inference in Non-Causal (i),
Partially and Structural Causal Models (ii)

3 Summarizing the Key Differences

4 Bonus: Speeding Up Mechanism Inference

TL;DR

AI requires the ‘inference’ in causal inference
to be **effectively computable**.
Classical causal models might **not** do the trick.

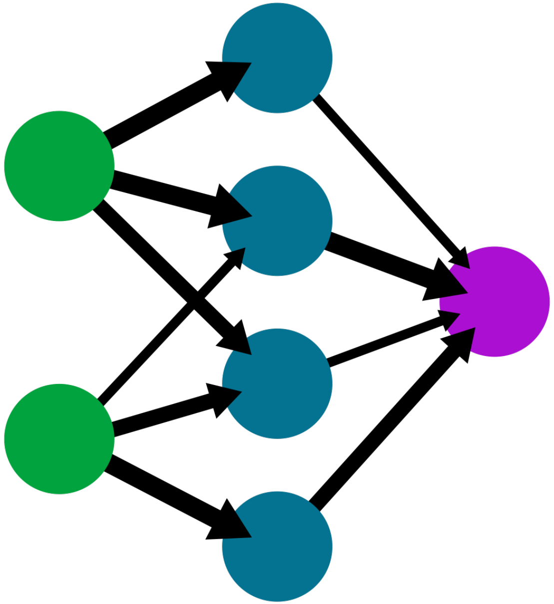
TL;DR

AI requires the ‘inference’ in causal inference
to be **effectively computable**.
Classical causal models might **not** do the trick.

…but we will do something about it! ;-)

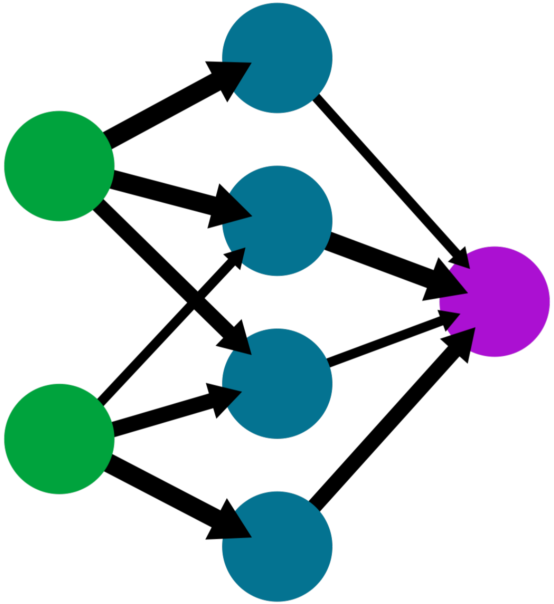
I | Are There Even Different Types of Causal Models?

Neural Nets

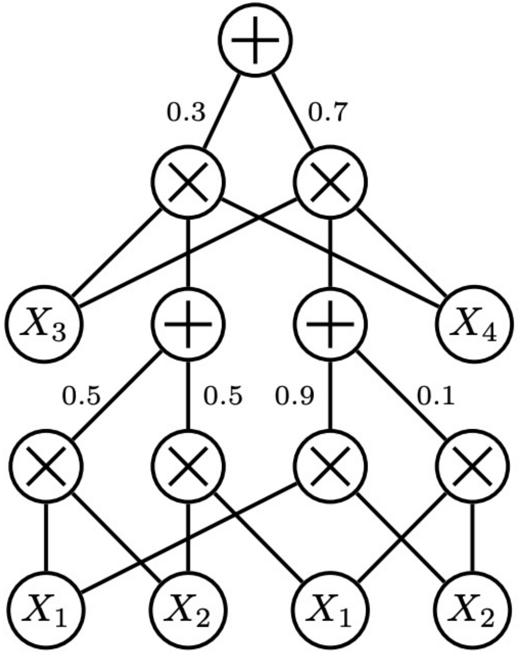


TECHNISCHE
UNIVERSITÄT
DARMSTADT

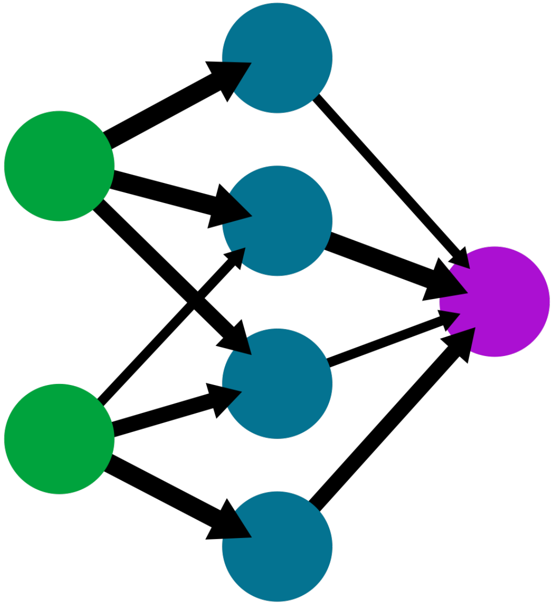
Neural Nets



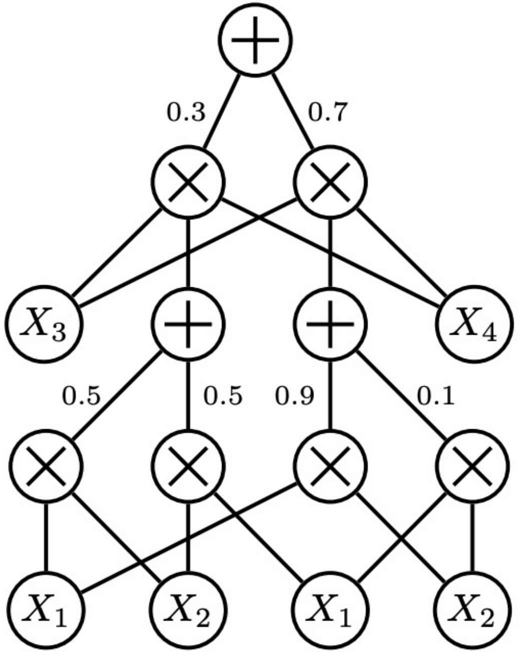
Sum-Product Nets



Neural Nets



Sum-Product Nets

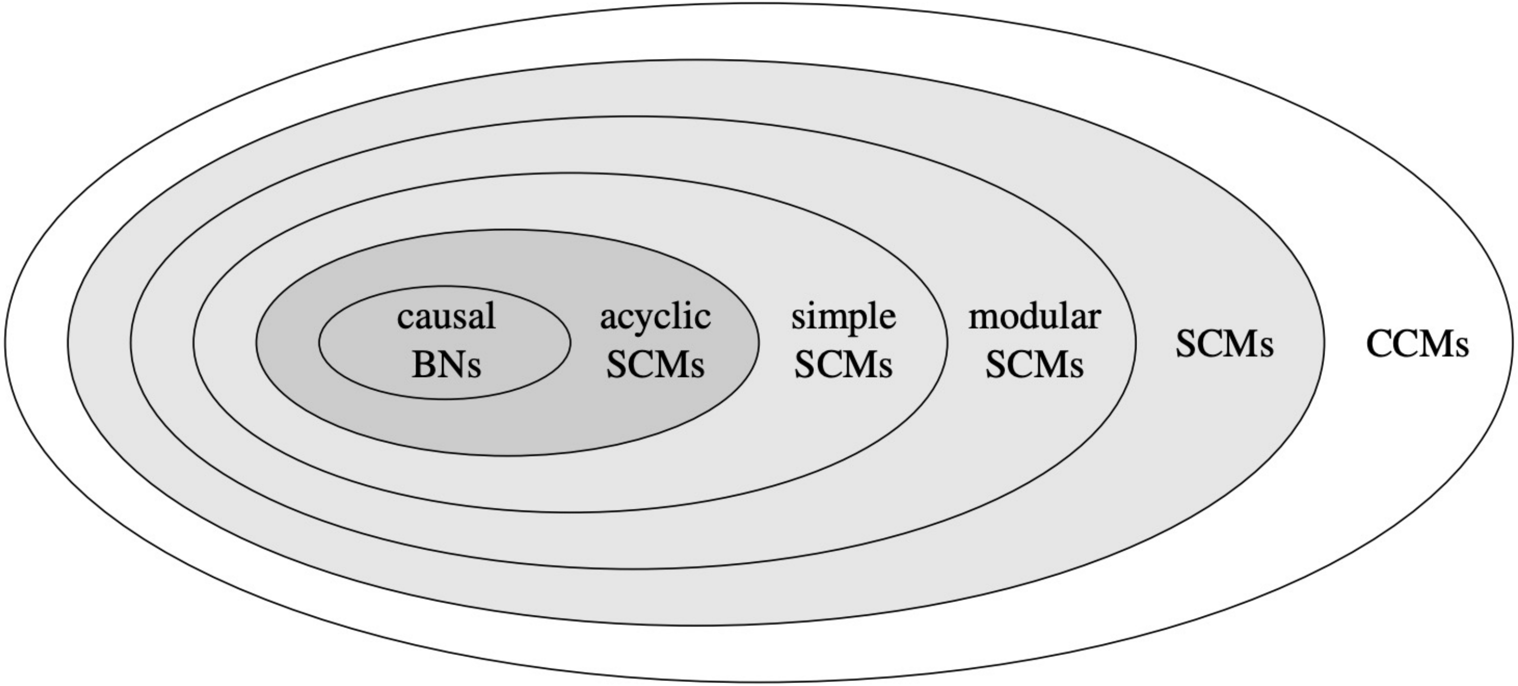


many more **non-causal** models...

Structural Causal Model (SCM)

$$\mathcal{M} = \langle \mathbf{V}, \mathbf{U}, \mathcal{F}, P(\mathbf{u}) \rangle$$

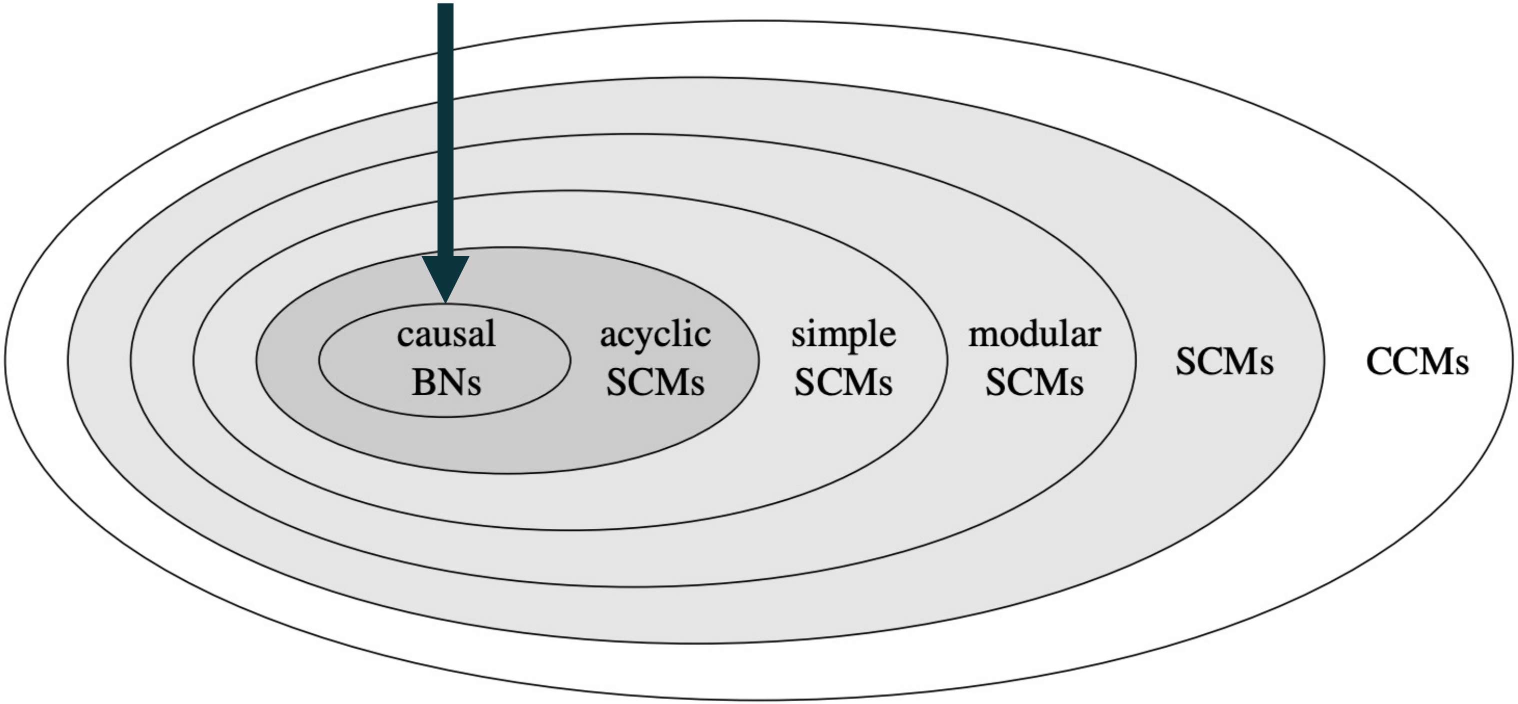
$$\mathcal{M} = \begin{cases} \mathbf{V} = \{X, M, Y\} \\ \mathbf{U} = \{U_{XY}, U_X, U_M, U_Y\} \\ \mathcal{F} = \begin{cases} f_X(U_X, U_{XY}) \\ f_M(X, U_M) \\ f_Y(M, U_Y, U_{XY}) \end{cases} \\ P(\mathbf{U}) \end{cases}$$



TECHNISCHE UNIVERSITÄT DARMSTADT

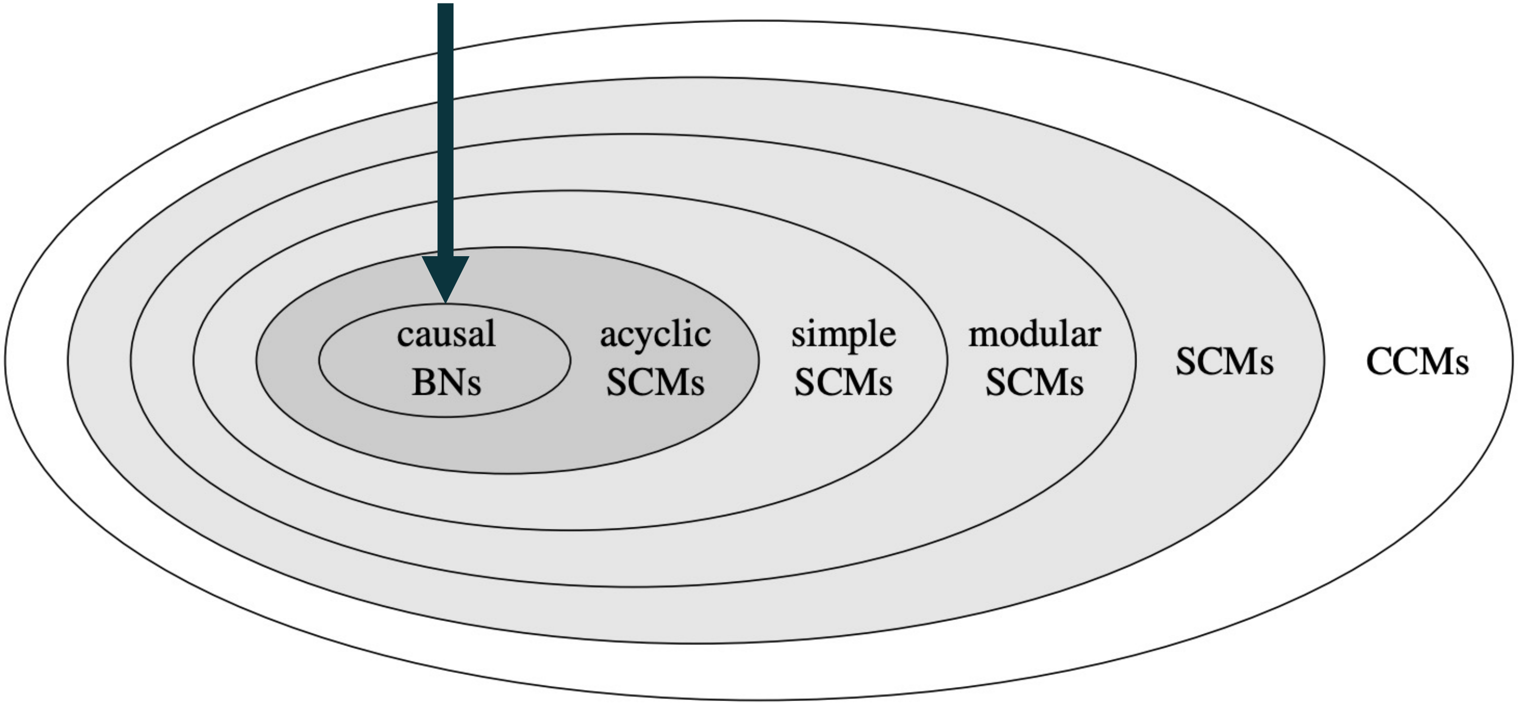
cannot define them

don’t even ‘speak’ counterfactuals \(o.o)/



cannot define them

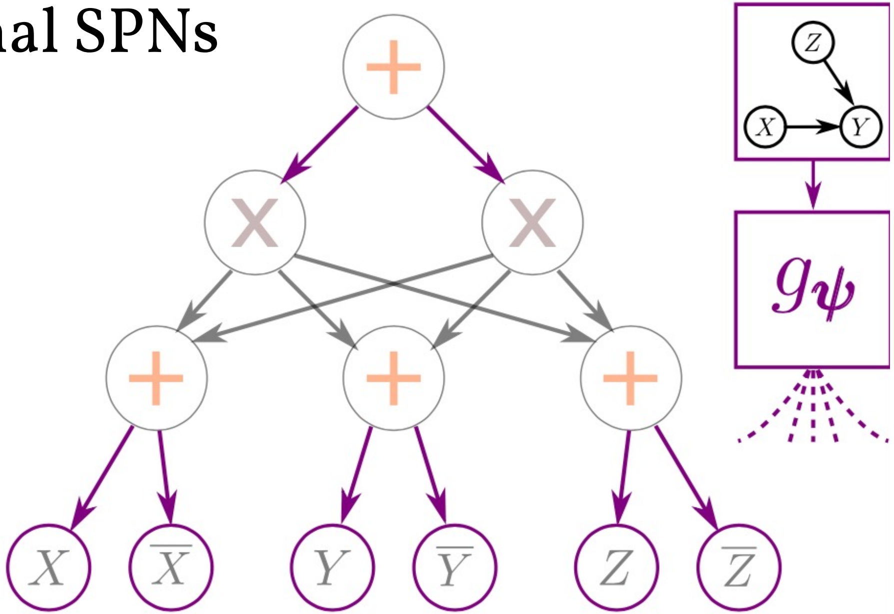
don’t even ‘speak’ counterfactuals \(o.o)/



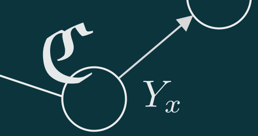
…well, CBN not a proper causal model then??

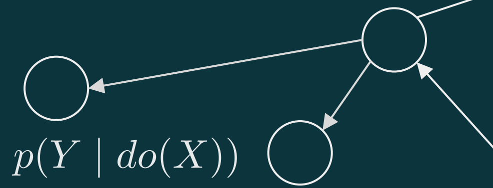
Other **Partially** Causal Models include..

Interventional SPNs

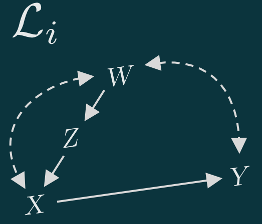


TECHNISCHE
UNIVERSITÄT
DARMSTADT





2(i) Inference in Non-Causal / Correlation-based Models



TECHNISCHE
UNIVERSITÄT
DARMSTADT

Query

do-Calculus

Estimand

Data + **Model**

Estimate

TECHNISCHE UNIVERSITÄT DARMSTADT

Query

$\downarrow$ do-Calculus

Estimand

$\downarrow$ Data + **Model**

Estimate

Example

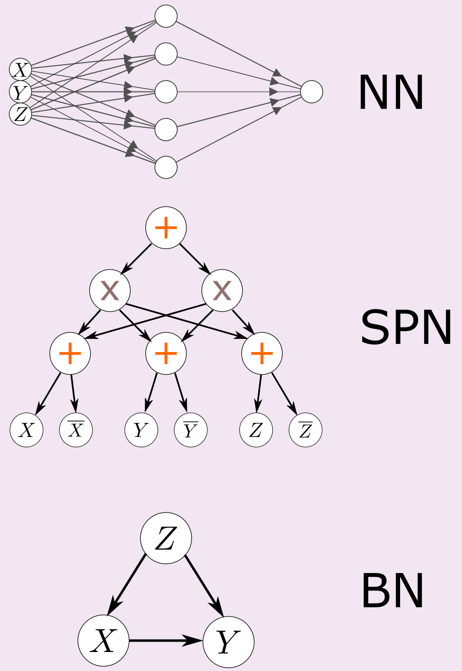
$$p(y|do(x))$$

$$\sum_z p(y|x,z)p(z)$$

$$\boxed{\sum_z \hat{p}(y|x,z)\hat{p}(z)}$$

	Example
Query	$p(y \vert do(x))$
↓ *do*-Calculus	
Estimand	$\sum_z p(y \vert x, z)p(z)$
↓ Data + **Model**	
Estimate	$\boxed{\sum_z \hat{p}(y \vert x, z)\hat{p}(z)}$

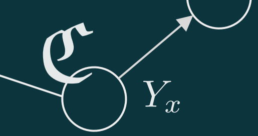
Example Models

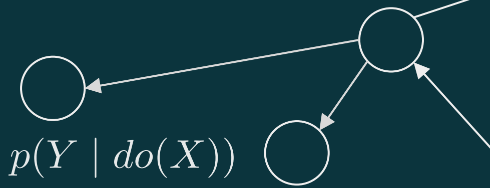


Corollary. Let Q be an identifiable causal query, $|Q|$ be the number of terms in the estimand and R be the number of edges in the DAG of SPN S. If $|Q| < R$, then Q is tractable.

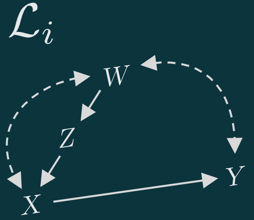
Corollary. Let Q be an identifiable causal query, $|Q|$ be the number of terms in the estimand and R be the number of edges in the DAG of SPN S. If $|Q| < R$, then Q is tractable.

For SPN Q is linear-time tractable.

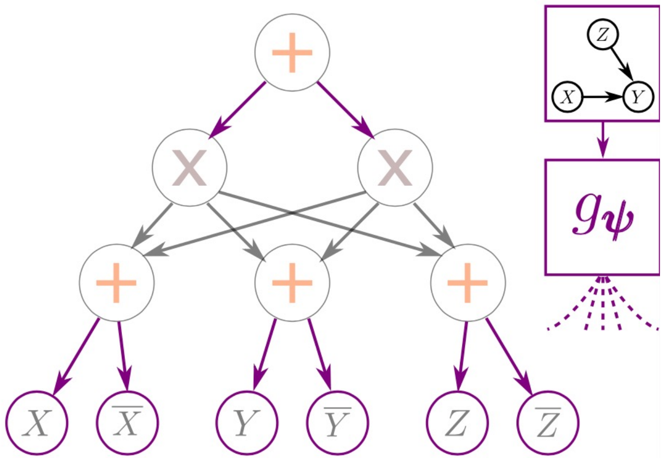




2(ii) Inference in (parameterized) Partially/ Structural Causal Models

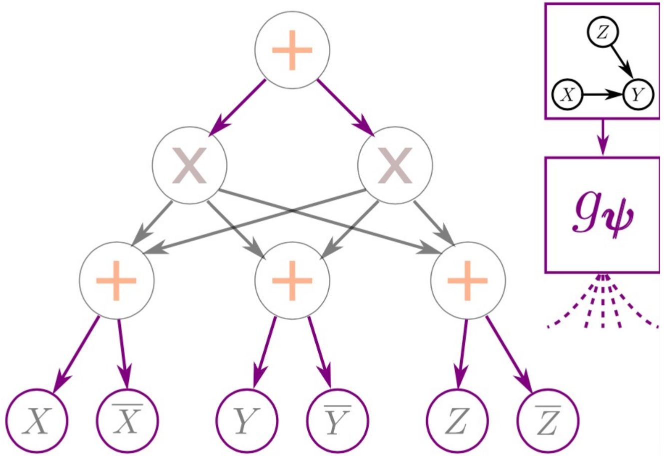


TECHNISCHE
UNIVERSITÄT
DARMSTADT

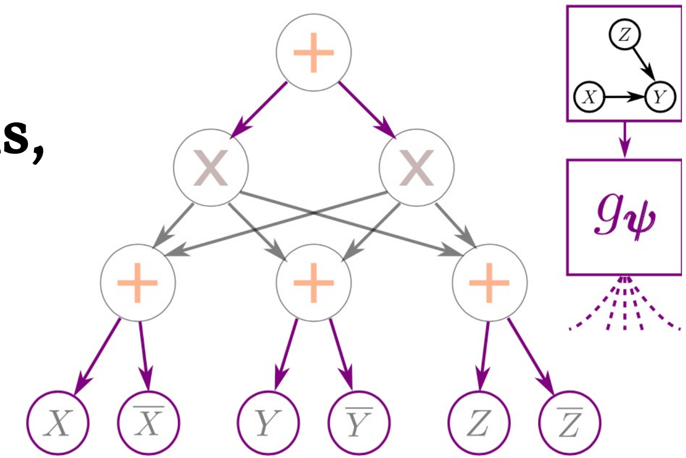


Proposition. Let $\{Q_i\}$ be a set of queries and R as before for iSPN I.
Any $\{Q_k\}$ is linear-time tractable.

no inherent identification, data assumption necessary



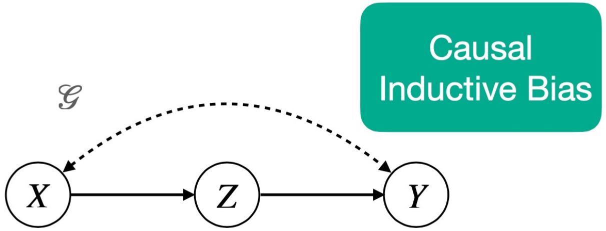
Proposition. Let $\{Q_i\}$ be a set of queries and R as before for iSPN I.
Any $\{Q_k\}$ is linear-time tractable.



arbitrary many distributions, efficient exact inference

Proposition. Let $\{Q_i\}$ be a set of queries and R as before for iSPN I.
Any $\{Q_k\}$ is linear-time tractable.

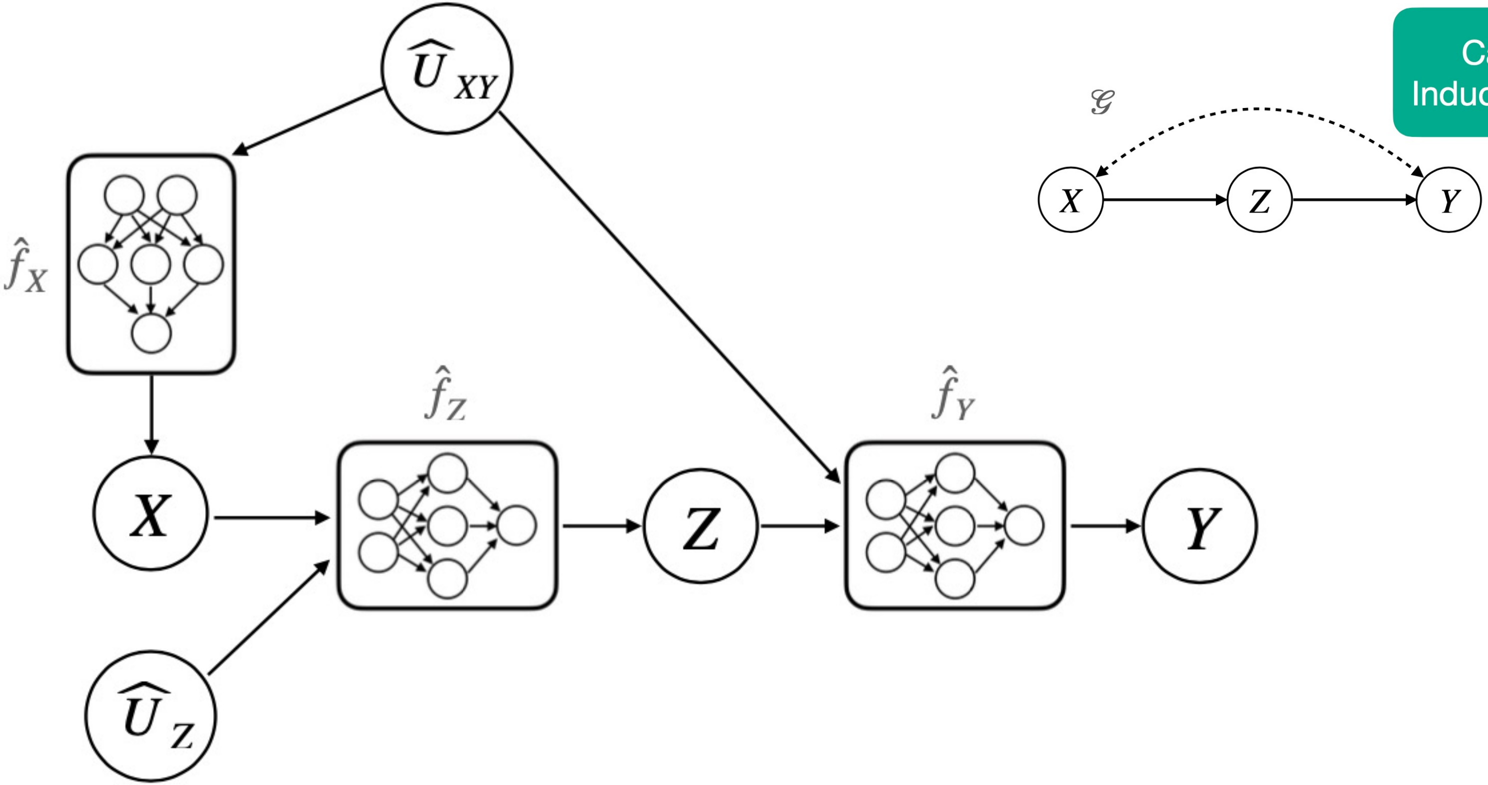
= Neural Nets + Causal Graph



TECHNISCHE UNIVERSITÄT DARMSTADT

Neural Causal Model = Neural Nets + Causal Graph

= parameterized SCM with neural nets mechanisms



Causal Inductive Bias

Theorem. Causal (marginal) inference in **any** parameterized SCM is NP-hard.

TECHNISCHE
UNIVERSITÄT
DARMSTADT

Theorem. Causal (marginal) inference in **any** parameterized SCM is NP-hard.

Proof.
Reduction to **3-SAT**, …

$$(x_1 \vee x_2 \vee x_3) \wedge (x_4 \vee x_5 \vee x_6)$$

exists x_i-tuple s.t. true?

TECHNISCHE UNIVERSITÄT DARMSTADT

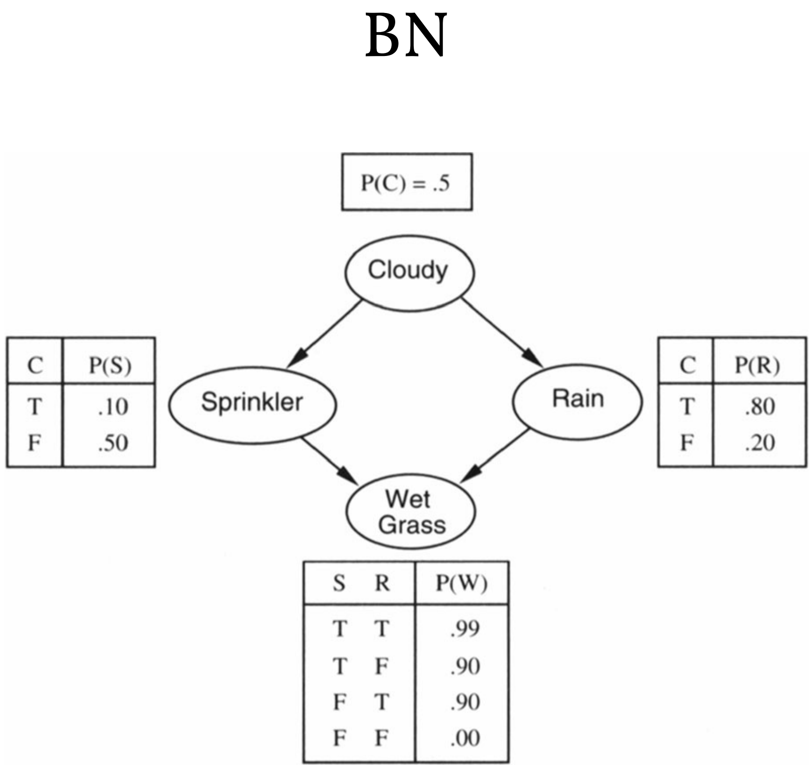
Theorem. Causal (marginal) inference
in **any** parameterized SCM is NP-hard.

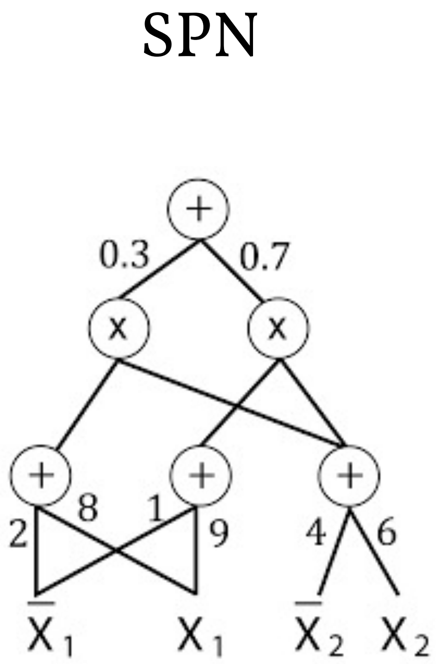
Proof.
Reduction to 3-SAT, so that solving causal inference amounts to solving 3-SAT. Since 3-SAT is NP-hard, we conclude that causal inference is NP-hard.

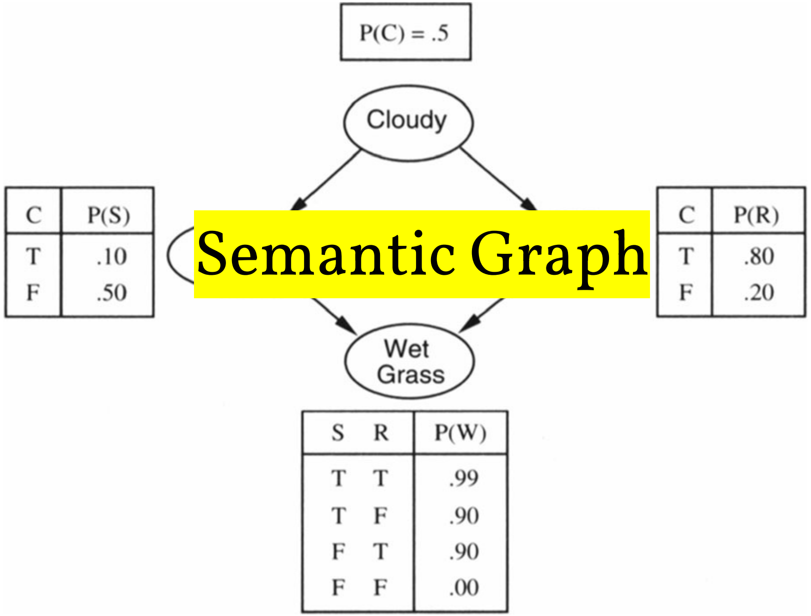
TECHNISCHE UNIVERSITÄT DARMSTADT

Theorem. Causal (marginal) inference in **any** parameterized SCM is NP-hard.

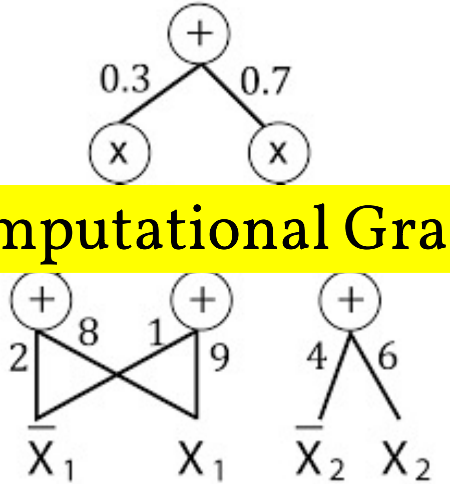
Historcal Issue: Inherited from BNs.


BN
P(C) = .5
Cloudy
Sprinkler
Rain
Wet Grass
C P(S)
T .10
F .50
C P(R)
T .80
F .20
S R P(W)
T T .99
T F .90
F T .90
F F .00


SPN
0.3
0.7
2
8
1
9
4
6
$\bar{X}_1$
X_1
$\bar{X}_2$
X_2



Semantic Graph



Computational Graph

Open Question:

Can we have the best of **both** worlds?

TECHNISCHE UNIVERSITÄT DARMSTADT

3 Summarizing the Key Differences

TECHNISCHE
UNIVERSITÄT
DARMSTADT

Little Taxonomy:

Model Family	PCH	Identification	Mechanism Inference	Marginal Inference

e.g. Neural Net

Little Taxonomy:

Model Family	PCH	Identification	Mechanism Inference	Marginal Inference

e.g. Neural Net

observational (L_1),
interventional (L_2),
counterfactual (L_3)

Little Taxonomy:

Model Family	PCH	Identification	Mechanism Inference	Marginal Inference

e.g. Neural Net

produce e.g.
L_2 with L_1?

observational (L_1),
interventional (L_2),
counterfactual (L_3)

Little Taxonomy:

Model Family	PCH	Identification	Mechanism Inference	Marginal Inference

e.g. Neural Net

observational (L_1),
interventional (L_2),
counterfactual (L_3)

produce e.g.
L_2 with L_1?

complexity of
$P(V_i \mid do(pa_i))$

TECHNISCHE UNIVERSITÄT DARMSTADT

Little Taxonomy:

Model Family	PCH	Identification	Mechanism Inference	Marginal Inference

e.g. Neural Net

produce e.g.
L_2 with L_1?

complexity
of $P(V_i)$

observational (L_1),
interventional (L_2),
counterfactual (L_3)

complexity of
$P(V_i \mid do(pa_i))$

TECHNISCHE
UNIVERSITÄT
DARMSTADT

Non-causals

Model Family	PCH	Identification	Mechanism Inference	Marginal Inference
OLS, CNN, GAN	$\mathcal{L}_1$	✗	-	polynomial

Best one but not expressivity of course

Model Family	PCH	Identification	Mechanism Inference	Marginal Inference
OLS, CNN, GAN	$\mathcal{L}_1$	✗	-	polynomial
SPN	$\mathcal{L}_1$	✗	-	linear (Cor. $\boxed{1}$)

TECHNISCHE UNIVERSITÄT DARMSTADT

Partially Causal

Model Family	PCH	Identification	Mechanism Inference	Marginal Inference
OLS, CNN, GAN	$\mathcal{L}_1$	✗	-	polynomial
SPN	$\mathcal{L}_1$	✗	-	linear (Cor. 1)
CausalVAE, iVGAE, CausalGAN	$\mathcal{L}_2$	✗	-	polynomial

TECHNISCHE UNIVERSITÄT DARMSTADT

Partially Causal with SPN

Model Family	PCH	Identification	Mechanism Inference	Marginal Inference
OLS, CNN, GAN	$\mathcal{L}_1$	✗	-	polynomial
SPN	$\mathcal{L}_1$	✗	-	linear (Cor. 1)
CausalVAE, iVGAE, CausalGAN	$\mathcal{L}_2$	✗	-	polynomial
iSPN	$\mathcal{L}_2$	✗	-	linear (Prop. 3)

TECHNISCHE
UNIVERSITÄT
DARMSTADT

SCMs

Model Family	PCH	Identification	Mechanism Inference	Marginal Inference
OLS, CNN, GAN	$\mathcal{L}_1$	✗	-	polynomial
SPN	$\mathcal{L}_1$	✗	-	linear (Cor. 1)
CausalVAE, iVGAE, CausalGAN	$\mathcal{L}_2$	✗	-	polynomial
iSPN	$\mathcal{L}_2$	✗	-	linear (Prop. 3)
NCM, DeepSCM	$\mathcal{L}_3$	✔	polynomial (Cor. 2)	intractable (Thm. 1)

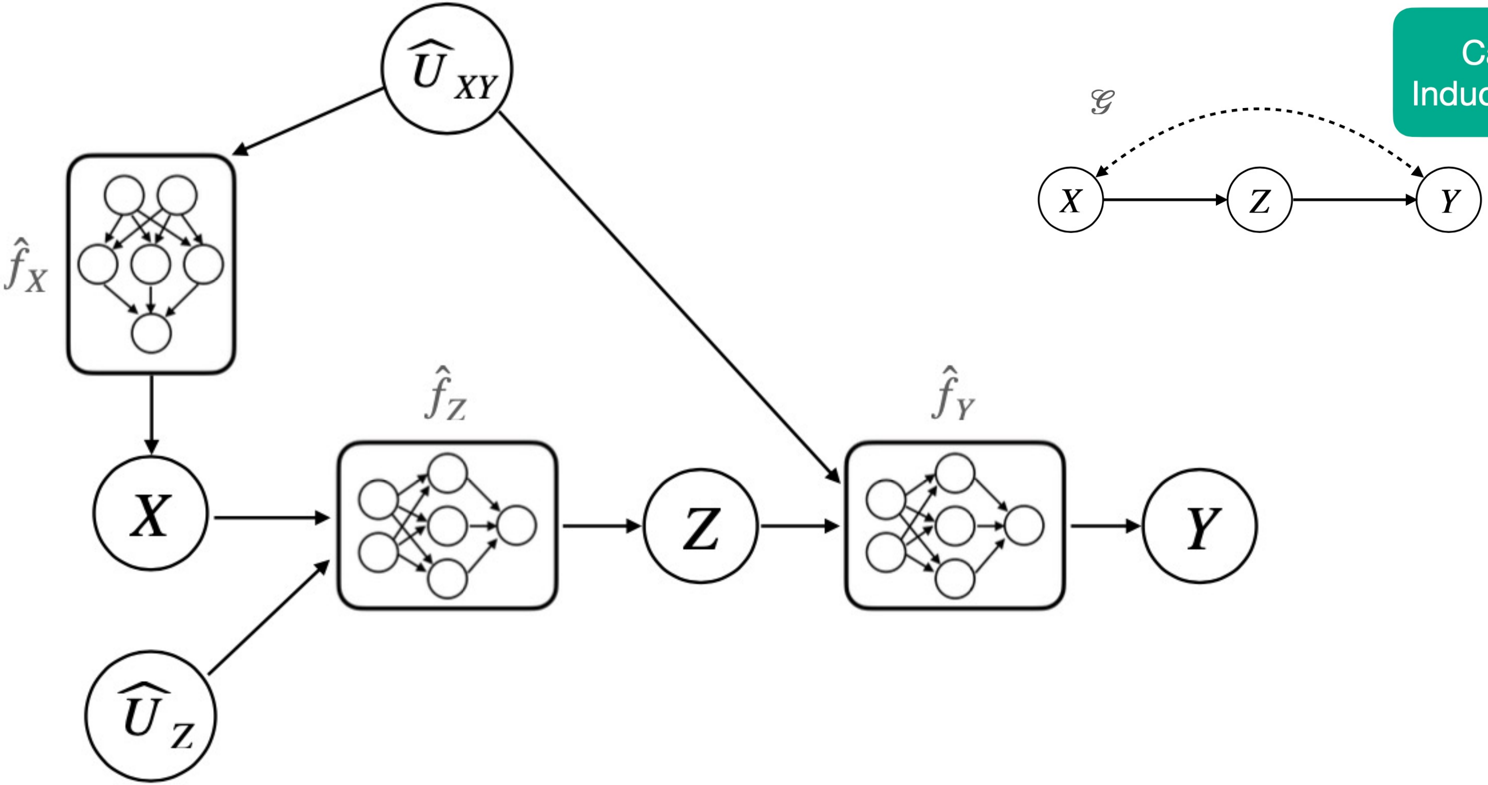
Open: SCMs with linear-time mechanisms?

Model Family	PCH	Identification	Mechanism Inference	Marginal Inference
OLS, CNN, GAN	$\mathcal{L}_1$	✗	-	polynomial
SPN	$\mathcal{L}_1$	✗	-	linear (Cor. 1)
CausalVAE, iVGAE, CausalGAN	$\mathcal{L}_2$	✗	-	polynomial
iSPN	$\mathcal{L}_2$	✗	-	linear (Prop. 3)
NCM, DeepSCM	$\mathcal{L}_3$	✔	polynomial (Cor. 2)	intractable (Thm. 1)
???	$\mathcal{L}_3$	✔	linear ?	intractable (Thm. 1)

TECHNISCHE UNIVERSITÄT DARMSTADT

4 Bonus: Speeding Up Mechanism Inference

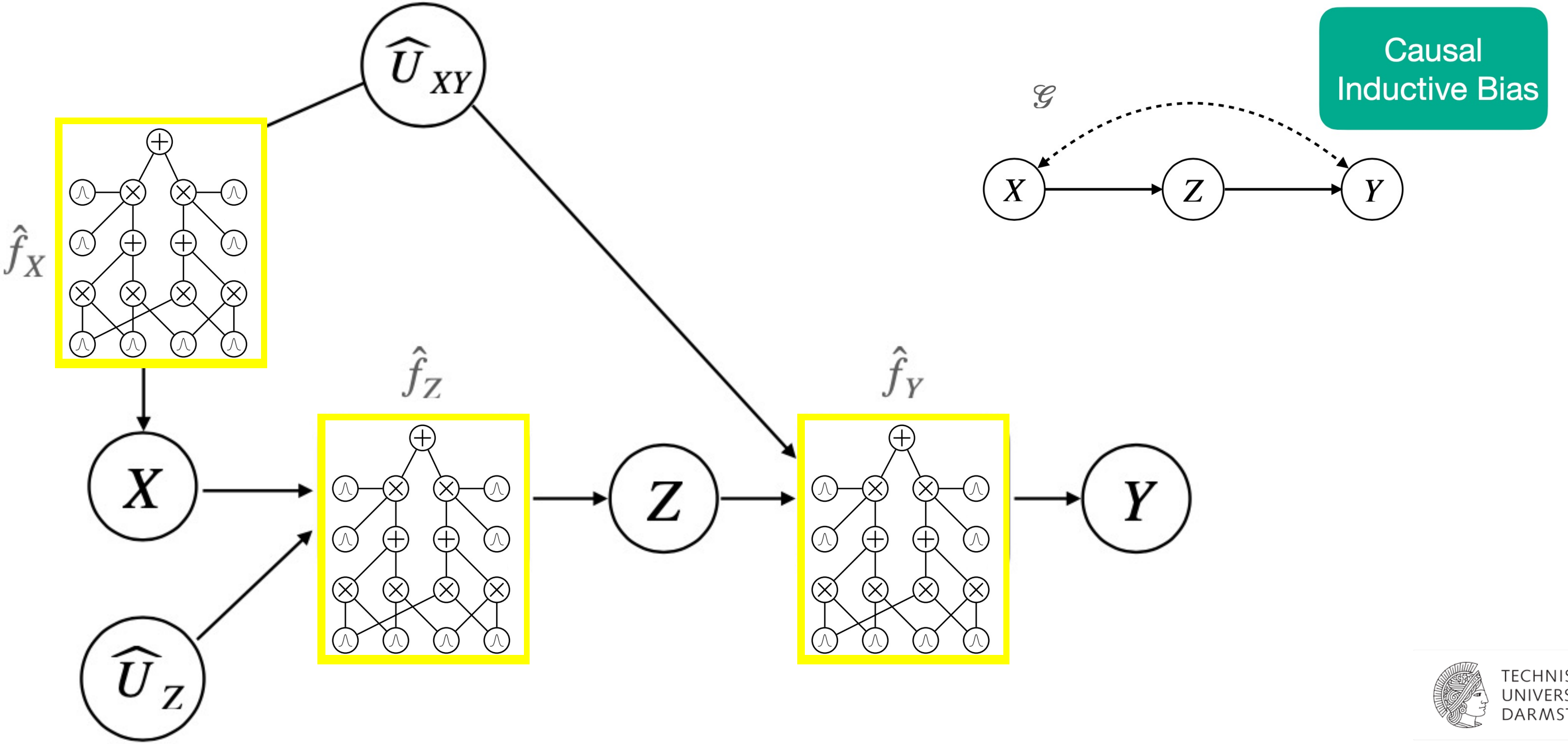
Neural Causal Model = Neural Nets + Causal Graph
= parameterized SCM with neural nets mechanisms



Causal Inductive Bias

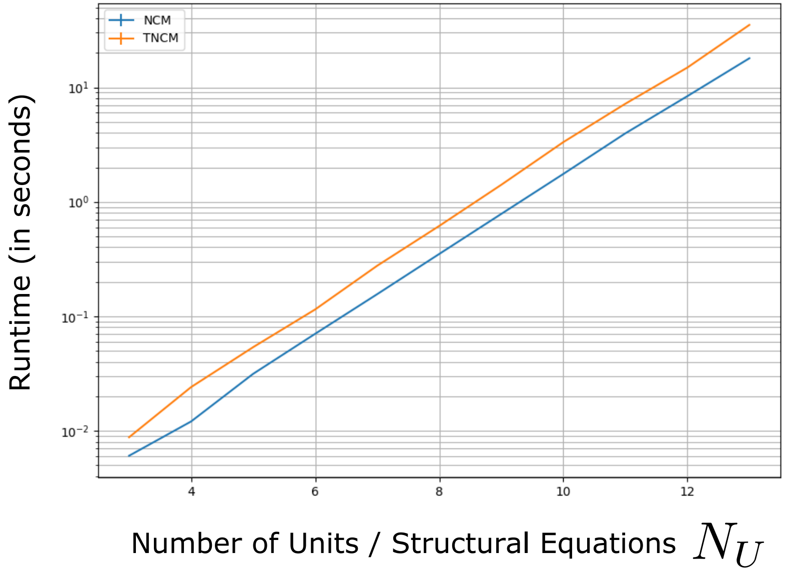
Circuit Causal Model = Sum-Product Nets + Causal Graph

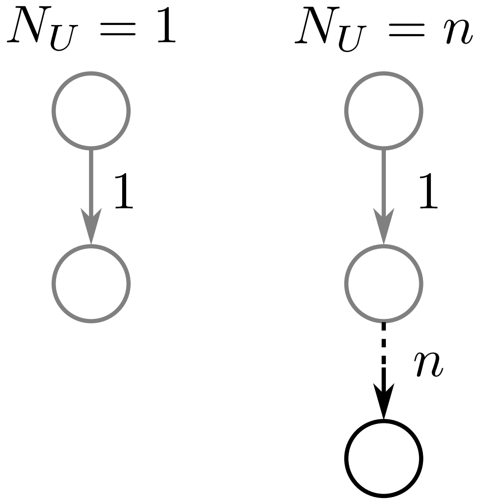
= parameterized SCM with SPN mechanisms



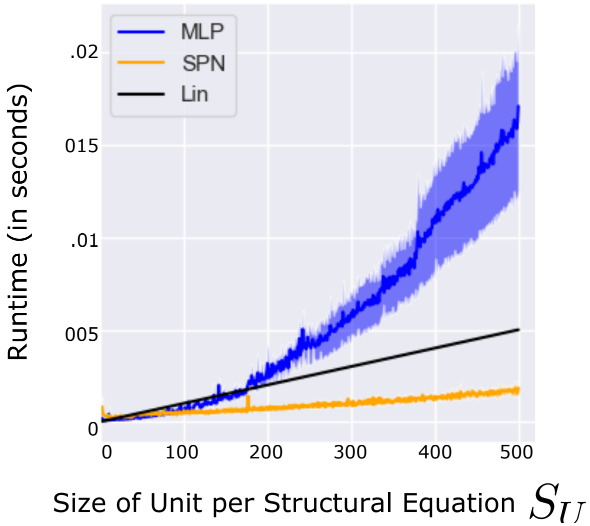
(a)

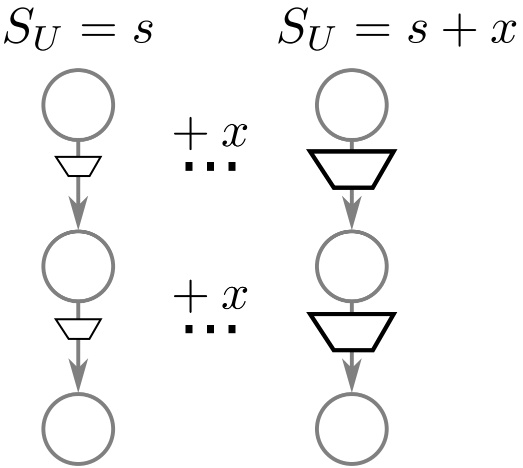
Marginal inference:
Both NCM and CCM
are intractable





(b) Mechanism inference:
Only CCM is
linear-time tractable





Answered: SCMs with linear-time mechanisms, yes!

Model Family	PCH	Identification	Mechanism Inference	Marginal Inference
OLS, CNN, GAN	$\mathcal{L}_1$	✗	-	polynomial
SPN	$\mathcal{L}_1$	✗	-	linear (Cor. 1)
CausalVAE, iVGAE, CausalGAN	$\mathcal{L}_2$	✗	-	polynomial
iSPN	$\mathcal{L}_2$	✗	-	linear (Prop. 3)
NCM, DeepSCM	$\mathcal{L}_3$	✔	polynomial (Cor. 2)	intractable (Thm. 1)
CCM	$\mathcal{L}_3$	✔	linear (Cor. 2)	intractable (Thm. 1)

Z | Final Remarks

It is clear: we want AI with causal reasoning capabilities because interacting with **change** is fundamental in nature

It is clear: we want AI with causal reasoning capabilities because interacting with **change** is fundamental in nature

It is also clear: we don't want to wait aeons and waste endless resources to get an answers from our AI

It is clear: we want AI with causal reasoning capabilities because interacting with **change** is fundamental in nature

It is also clear: we don’t want to wait aeons and waste endless resources to get an answers from our AI

Trade-off ?
“**how much**” causal vs. **how efficient** the acquisition
Future: Partially Causal Models?

Prof. Dr. Kristian Kersting, FEurAI, FELLIS

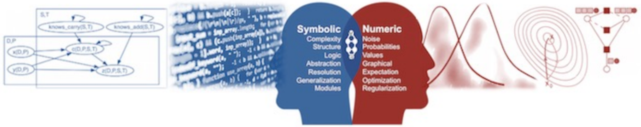
Computer Science Department and Centre for Cognitive Science, TU Darmstadt

Altes Hauptgebäude, Room 074, Hochschulstrasse 1, 64289 Darmstadt, Germany

+49-6151-16-24411 kersting (at) cs (dot) tu-darmstadt (dot) de

Informatik

TECHNISCHE UNIVERSITÄT DARMSTADT



Devendra Singh Dhami

Machine Learning Group, Computer Science Department, TU Darmstadt.

Hochschulstrasse 1, Room S1|66, 64289 Darmstadt, Germany

+49 1523 79 19263 devendra (dot) dhami (at) cs (dot) tu-darmstadt (dot) de

hessian.AI

Matej Zečević

Machine Learning Group, Computer Science Department, TU Darmstadt.

Hochschulstrasse 1, Room S1|03 066, 64289 Darmstadt, Germany.

matej (dot) zecevic (at) cs (dot) tu-darmstadt (dot) de

CAUSE

TECHNISCHE UNIVERSITÄT DARMSTADT

Thank you!

Questions?

matej.zecevic@tu-darmstadt.de | https://www.matej-zecevic.de

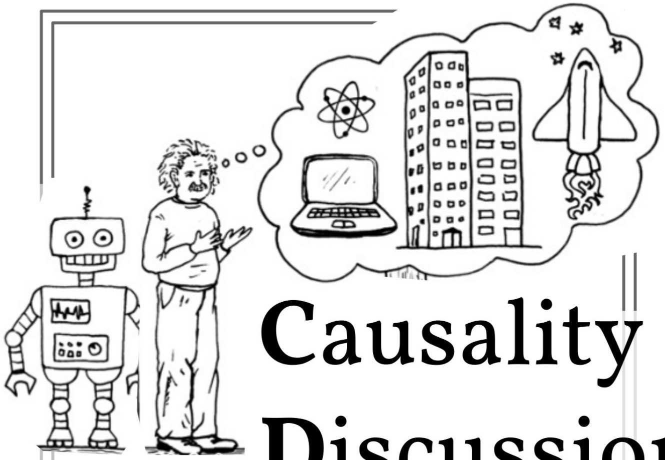
https://arxiv.org/pdf/2110.12052.pdf

Announcements

Want to discuss more *following* the "YES Causal Inference" Workshop ?

TECHNISCHE UNIVERSITÄT DARMSTADT

Every Week with Paper Authors → Discuss LIVE



Causality
Discussion
Group

465 members
on Slack

Join the community
via
discuss.causality.link

TECHNISCHE
UNIVERSITÄT
DARMSTADT

Past Sessions: [Password: Causa1ity, Direct Access Link]

- Session 01.03.2023 | **Deep Counterfactual Estimation with Categorical Background Variables** | Discussant: Edward De Brouwer
- Session 22.02.2023 | **Information-Theoretic Causal Discovery and Intervention Detection over Multiple Environments** | Discussant: Osman Ali Mian
- Session 08.02.2023 | **CLEAR: Generative Counterfactual Explanations on Graphs** | Discussants: Jing Ma, Ruocheng Guo
- Session 01.02.2023 | **Causal Transformer for Estimating Counterfactual Outcomes** | Discussant: Valentyn Melnychuk
- Session 25.01.2023 | **Abstracting Causal Models** | Discussant: Sander Beckers
- Session 18.01.2023 | **Desiderata for Representation Learning: A Causal Perspective** | Discussant: Yixin Wang
- Session 11.01.2023 | **Causal Feature Selection via Orthogonal Search** | Discussant: Ashkan Soleymani
- Session 14.11.2022 | **Rewind 2022** | Final session of 2022 to simply rewind on what we experienced throughout the year
- Session 07.12.2022 | **Causal Inference Through the Structural Causal Marginal Problem** | Discussant: Luigi Gresele
- Session 30.11.2022 | **Selecting Data Augmentation for Simulating Interventions** | Discussant: Maximilian Ilse
- Session 23.11.2022 | **On Disentangled Representations Learned from Correlated Data** | Discussant: Frederik Träuble
- Session 16.11.2022 | **Causal Curiosity: RL Agents Discovering Self-supervised Experiments for Causal Repr. Learning** | Discussant: Sumedh Sontakke
- Session 09.11.2022 | **Causal Machine Learning: A Survey and Open Problems** | Discussants: Jean Kaddour, Aengus Lynch
- Session 02.11.2022 | **A Critical Look at the Consistency of Causal Estimation with Deep Latent Variable Models** | Discussant: Severi Rissanen
- Session 26.10.2022 | **Nonlinear Invariant Risk Minimization: A Causal Approach** | Discussant: Chaochao Lu
- Session 19.10.2022 | **CausalVAE: Disentangled Representation Learning via Neural Structural Causal Models** | Discussant: Mengyue Yang
- Session 12.10.2022 | **Weakly Supervised Causal Representation Learning** | Discussant: Johann Brehmer
- Session 05.10.2022 | **Towards Causal Representation Learning** | Discussant: Anirudh Goyal
- Session 21.09.2022 | **Selection Collider Bias in Large Language Models** | Discussant: Emily McMilin
- Session 14.09.2022 | **The Causal-Neural Connection: Expressiveness, Learnability, and Inference** | Discussants: Kai-Zhan Lee, Kevin Xia
- Session 07.09.2022 | **Self-Supervised Learning with Data Augmentations Provably Isolates Content from Style** | Discussant: Julius von Kügelgen

⋮

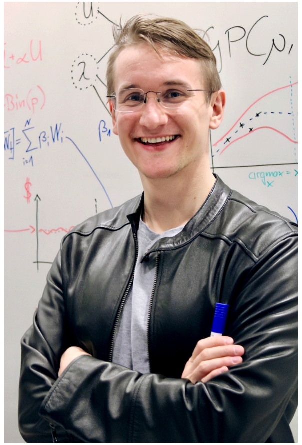
30 Sessions Completed and All Recorded

TECHNISCHE UNIVERSITÄT DARMSTADT

Are you also overwhelmed by all the different flavours of causality and all its different protagonists?

TECHNISCHE UNIVERSITÄT DARMSTADT

Genealogy of Causality



with
Jakob Zeitler
(meet him here today!)

Name	Institution	Superv		
UCLA				
Judea Pearl	UCLA	?	US	Rutgers, Tech
Wesley Salmon	UCLA	Hans Reichenbach	US	?
Hans Reichenbach	UCLA	Paul Hensel, Max Noeth	US	Berlin, Istanbu
John Hopkins				
Ilya Shpitser	John Hopkins		US	UCLA, Judea
Oregon State University				
Karthika Mohan	Oregon State University	Judea Pearl	US	
CMU				
Kun Zhang	CMU		Pittsburgh, US	MPI Tübingen
Clark Glymour	CMU	Wesley Salmon	Pittsburgh, US	
Peter Spirtes	CMU		Pittsburgh, US	
ETH Zürich				
Peter Bühlmann	ETH		Zürich	?
Marloes Maathuis	ETH		Zürich	?
Nicolai Meinshausen	ETH			
LMU Munich				
Stephan Hartmann	LMU		Munich, Germany	
MPI Tübingen				
Bernhard Schölkopf	MPI Tübingen	Vladimir Vapnik	Tübingen, Germ	TU Berlin
Ulrike von Luxburg	MPI Tübingen		Tübingen, Germany	
Michel Besserve				

To Extend This!

Name	Institution	Supervisor	Location	Previous Pos
UCLA				
Judea Pearl	UCLA	?	US	Rutgers, Tech
Wesley Salmon	UCLA	Hans Reichenbach	US	?
Hans Reichenbach	UCLA	Paul Hensel, Max Noeth	US	Berlin, Istanbu
John Hopkins				
Ilya Shpitser	John Hopkins		US	UCLA, Judea
Oregon State University				
Karthika Mohan	Oregon State University	Judea Pearl	US	
CMU				
Kun Zhang	CMU		Pittsburgh, US	MPI Tübingen
Clark Glymour	CMU	Wesley Salmon	Pittsburgh, US	
Peter Spirtes	CMU		Pittsburgh, US	
ETH Zürich				
Peter Bühlmann	ETH		Zürich	?
Marloes Maathuis	ETH		Zürich	?
Nicolai Meinshausen	ETH			
LMU Munich				
Stephan Hartmann	LMU		Munich, Germany	
MPI Tübingen				
Bernhard Schölkopf	MPI Tübingen	Vladimir Vapnik	Tübingen, Germ	TU Berlin
Ulrike von Luxburg	MPI Tübingen		Tübingen, Germany	
Michel Besserve				