Causal Discovery of Linear Cyclic Models from Multiple Experimental Data Sets with Overlapping Variables

Antti Hyttinen[1], Frederick Eberhardt[2] and Patrik O. Hoyer[1]

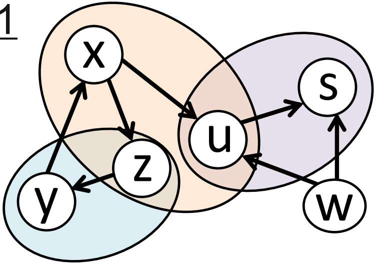
[1] Helsinki Institute for Information Technology (HIIT)
[1] Dept. of Computer Science, University of Helsinki
[2] Dept. of Philosophy, Carnegie Mellon University

UAI 2012
17.8.2012

General Learning Setting

Experiment 1
intervened: x
observed: z, u

Experiment 2
intervened: y
observed: z



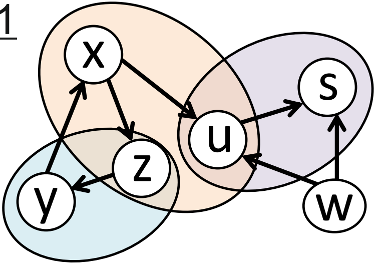
Experiment 3
intervened: -
observed: u, s

- **Multiple experimental data sets** from randomized controlled trials (Cooper and Yoo 1999, Murphy et al. 2001-2009)

General Learning Setting

Experiment 1
intervened: x
observed: z, u

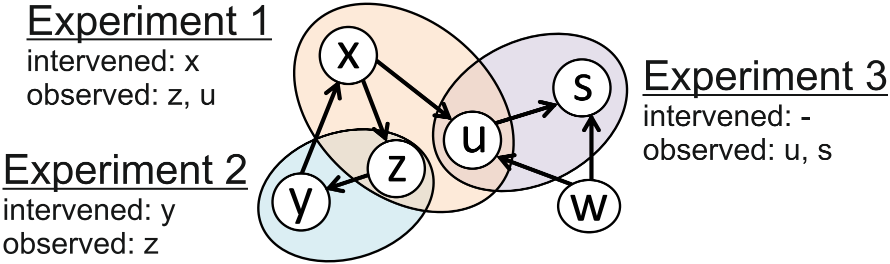
Experiment 2
intervened: y
observed: z



Experiment 3
intervened: -
observed: u, s

- **Multiple experimental data sets** from randomized controlled trials (Cooper and Yoo 1999, Murphy et al. 2001-2009)
- Multiple partially **overlapping data sets** (Tillman et al. 2009-2011, Tsamardinos, Triantafillou et al. 2010-2012)

General Learning Setting

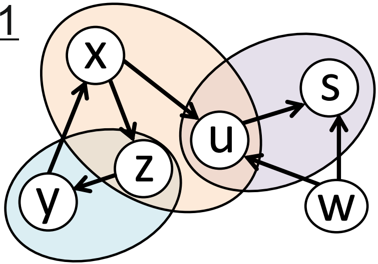


- **Multiple experimental data sets** from randomized controlled trials (Cooper and Yoo 1999, Murphy et al. 2001-2009)
- Multiple partially **overlapping data sets** (Tillman et al. 2009-2011, Tsamardinos, Triantafillou et al. 2010-2012)
- **No joint causal sufficiency**: there may be latent confounders not measured in any of the data sets

General Learning Setting

Experiment 1
intervened: x
observed: z, u

Experiment 2
intervened: y
observed: z

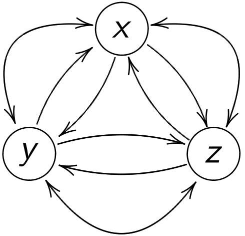


Experiment 3
intervened: -
observed: u, s

- **Multiple experimental data sets** from randomized controlled trials (Cooper and Yoo 1999, Murphy et al. 2001-2009)
- Multiple partially **overlapping data sets** (Tillman et al. 2009-2011, Tsamardinos, Triantafillou et al. 2010-2012)
- **No joint causal sufficiency**: there may be latent confounders not measured in any of the data sets
- The underlying structure may be **cyclic**
- Causal relations restricted to be **linear**

Background: Linear Cyclic Model with Latent Variables

- Over the joint set of variables in the overlapping experiments



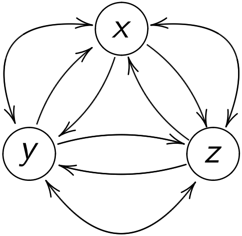
$$\begin{aligned} x &:= b(y \to x)y + b(z \to x)z + e_x \\ y &:= b(x \to y)x + b(z \to y)z + e_y \\ z &:= b(x \to w)x + b(y \to z)y + e_z \end{aligned}$$

$$\mathbf{x} := \mathbf{B}\mathbf{x} + \mathbf{e}, \quad \mathbf{e} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{\mathbf{e}})$$

$$\mathbf{B} = \begin{pmatrix} 0 & b(y \to x) & b(z \to x) \\ b(x \to y) & 0 & b(z \to y) \\ b(x \to z) & b(y \to z) & 0 \end{pmatrix} \quad \boldsymbol{\Sigma}_e = \begin{pmatrix} \sigma_x^2 & \sigma_{xy} & \sigma_{xz} \\ \sigma_{xy} & \sigma_y^2 & \sigma_{yz} \\ \sigma_{xz} & \sigma_{yz} & \sigma_z^2 \end{pmatrix}$$

Background: Linear Cyclic Model with Latent Variables

- Over the joint set of variables in the overlapping experiments



$$\begin{aligned} x &:= b(y \to x)y + b(z \to x)z + e_x \\ y &:= b(x \to y)x + b(z \to y)z + e_y \\ z &:= b(x \to w)x + b(y \to z)y + e_z \end{aligned}$$

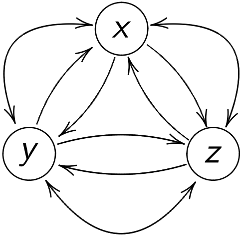
$$\mathbf{x} := \mathbf{B}\mathbf{x} + \mathbf{e}, \quad \mathbf{e} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_\mathbf{e})$$

$$\mathbf{B} = \begin{pmatrix} 0 & b(y \to x) & b(z \to x) \\ b(x \to y) & 0 & b(z \to y) \\ b(x \to z) & b(y \to z) & 0 \end{pmatrix} \quad \boldsymbol{\Sigma}_e = \begin{pmatrix} \sigma_x^2 & \sigma_{xy} & \sigma_{xz} \\ \sigma_{xy} & \sigma_y^2 & \sigma_{yz} \\ \sigma_{xz} & \sigma_{yz} & \sigma_z^2 \end{pmatrix}$$

- Latent confounding represented by correlated disturbances

Background: Linear Cyclic Model with Latent Variables

- Over the joint set of variables in the overlapping experiments



$$\begin{aligned} x &:= b(y \to x)y + b(z \to x)z + e_x \\ y &:= b(x \to y)x + b(z \to y)z + e_y \\ z &:= b(x \to w)x + b(y \to z)y + e_z \end{aligned}$$

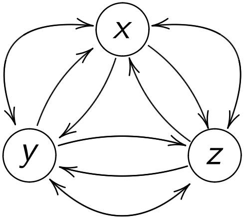
$$\mathbf{x} := \mathbf{B}\mathbf{x} + \mathbf{e}, \quad \mathbf{e} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_\mathbf{e})$$

$$\mathbf{B} = \begin{pmatrix} 0 & b(y \to x) & b(z \to x) \\ b(x \to y) & 0 & b(z \to y) \\ b(x \to z) & b(y \to z) & 0 \end{pmatrix} \quad \boldsymbol{\Sigma}_e = \begin{pmatrix} \sigma_x^2 & \sigma_{xy} & \sigma_{xz} \\ \sigma_{xy} & \sigma_y^2 & \sigma_{yz} \\ \sigma_{xz} & \sigma_{yz} & \sigma_z^2 \end{pmatrix}$$

- Latent confounding represented by correlated disturbances
- Generated distribution

$$\begin{aligned} \mathbf{x} &= (\mathbf{I} - \mathbf{B})^{-1}\mathbf{e} \\ \mathbf{x} &\sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_\mathbf{x}), \text{ where } \boldsymbol{\Sigma}_\mathbf{x} = (\mathbf{I} - \mathbf{B})^{-1}\boldsymbol{\Sigma}_\mathbf{e}(\mathbf{I} - \mathbf{B})^{-T} \end{aligned}$$

Problem Formalization



$$\mathbf{x} \;:=\; \mathbf{B}\mathbf{x} + \mathbf{e}, \quad \mathbf{e} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{\mathbf{e}})$$

$\Rightarrow$

Experiment 1

x	y	z
$-$**0.3**	-0.7	?
⋮	⋮	⋮
int.	obs.	unobs.

Experiment 2

x	y	**z**
?	-0.1	**0.2**
⋮	⋮	⋮
unobs.	obs.	**int.**

$\Rightarrow \mathbf{B} = ?$

- Given multiple data sets
  - in which some variables are surgically intervened on, some are (passively) observed and some are unobserved
  - generated by a (manipulated) linear cyclic model with latent variables $(\mathbf{B}, \boldsymbol{\Sigma}_{\mathbf{e}})$ over the joint set of variables in the experiments
- Identify as many causal relations among the joint set of variables in the experiments, i.e. entries in $\mathbf{B}$, such as $b(x \rightarrow y)$, as possible.

Background: Experimental Effects

- The sum-product of coefficients on directed paths from an intervened variable x to an observed variable y is called an **experimental effect**:

intervened variable observed variable

$$t(x \rightsquigarrow y \,||\, \mathcal{J})$$

total sum-prod. paths all intervened variables

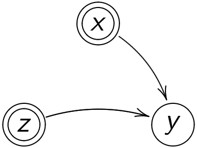
Background: Experimental Effects

- The sum-product of coefficients on directed paths from an intervened variable x to an observed variable y is called an **experimental effect**:

intervened variable observed variable

$$t(\,x \rightsquigarrow y \,||\, \mathcal{J}\,)$$

total sum-prod. paths all intervened variables



$$t(x \rightsquigarrow y || \mathcal{V} \setminus y) \quad = \quad b(x \to y)$$

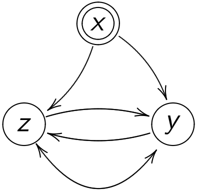
Background: Experimental Effects

- The sum-product of coefficients on directed paths from an intervened variable x to an observed variable y is called an **experimental effect**:

intervened variable observed variable

$$t(x \rightsquigarrow y \,||\, \mathcal{J})$$

total sum-prod. paths all intervened variables



$$
\begin{aligned}
t(x \rightsquigarrow y || \mathcal{V} \setminus y) &= b(x \rightarrow y) \\
t(x \rightsquigarrow y || x) &= (b(x \rightarrow y) + b(x \rightarrow z) b(z \rightarrow y)) \\
&\quad \cdot (1 + b(y \rightarrow z) b(z \rightarrow y) + \cdots) \\
&= \frac{b(x \rightarrow y) + b(x \rightarrow z) b(z \rightarrow y)}{1 - b(y \rightarrow z) b(z \rightarrow y)}
\end{aligned}
$$

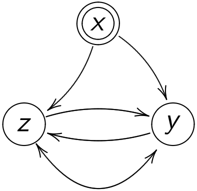
Background: Experimental Effects

- The sum-product of coefficients on directed paths from an intervened variable x to an observed variable y is called an **experimental effect**:

intervened variable　observed variable

$$t(\, x \rightsquigarrow y \,||\, \mathcal{J}\,)$$

total sum-prod.　paths　all intervened variables



$$
\begin{aligned}
t(x \rightsquigarrow y || \mathcal{V} \setminus y) &= b(x \to y) \\
t(x \rightsquigarrow y || x) &= (b(x \to y) + b(x \to z) b(z \to y)) \\
&\quad \cdot (1 + b(y \to z) b(z \to y) + \cdots) \\
&= \frac{b(x \to y) + b(x \to z) b(z \to y)}{1 - b(y \to z) b(z \to y)}
\end{aligned}
$$

- From an experiment where x and w are intervened on and y is observed the following exp. effects can be estimated by regression

$$t(x \rightsquigarrow y || x, w) = 0.227 \quad t(w \rightsquigarrow y || x, w) = -0.345$$

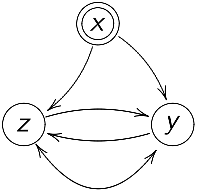
Background: Experimental Effects

- The sum-product of coefficients on directed paths from an intervened variable x to an observed variable y is called an **experimental effect**:

intervened variable   observed variable

$$t(\, x \rightsquigarrow y \,\|\, \mathcal{J}\,)$$

total sum-prod.   paths   all intervened variables



$$\begin{aligned} t(x \rightsquigarrow y \| \mathcal{V} \setminus y) &= b(x \to y) \\ t(x \rightsquigarrow y \| x) &= (b(x \to y) + b(x \to z)b(z \to y)) \\ &\quad \cdot (1 + b(y \to z)b(z \to y) + \cdots) \\ &= \frac{b(x \to y) + b(x \to z)b(z \to y)}{1 - b(y \to z)b(z \to y)} \end{aligned}$$

- From an experiment where x and w are intervened on and y is observed the following exp. effects can be estimated by regression

$$t(x \rightsquigarrow y \| x, w) = 0.227 \quad t(w \rightsquigarrow y \| x, w) = -0.345$$

- When can we infer coeffs $b(x \to y)$ from the observed exp. effects?

Identifiability from Overlapping Data Sets (Result 1)

- We can extend the identifiability results given by Eberhardt et al. (2010) for experimental data sets sharing the same variables:

Identifiability from Overlapping Data Sets (Result 1)

- We can extend the identifiability results given by Eberhardt et al. (2010) for experimental data sets sharing the same variables:
- A sufficient condition for fully identifying $\mathbf{B}$ is **the pair condition**: for each ordered pair of variables (u, v) there is an experiment such that u is intervened on and v is (passively) observed.

	x	y	z
x	$\cdot$	$\checkmark$	$\checkmark$
y	$\checkmark$	$\cdot$	$\checkmark$
z	$\checkmark$	$\checkmark$	$\cdot$

Identifiability from Overlapping Data Sets (Result 1)

- We can extend the identifiability results given by Eberhardt et al. (2010) for experimental data sets sharing the same variables:
- A sufficient condition for fully identifying **B** is **the pair condition**: for each ordered pair of variables (u, v) there is an experiment such that u is intervened on and v is (passively) observed.

	x	y	z
x	$\cdot$	✓	✓
y	✓	$\cdot$	✓
z	✓	✓	$\cdot$

- Also **Worst Case Necessary**: If the pair condition is not satisfied for some ordered pair of variables, there exist distinct **B**s that have identical distributions for the given experiments.

Identifiability from Overlapping Data Sets (Result 1)

- We can extend the identifiability results given by Eberhardt et al. (2010) for experimental data sets sharing the same variables:
- A sufficient condition for fully identifying **B** is **the pair condition**: for each ordered pair of variables (u, v) there is an experiment such that u is intervened on and v is (passively) observed.

	x	y	z
x	$\cdot$	$\checkmark$	$\checkmark$
y	$\checkmark$	$\cdot$	$\checkmark$
z	$\checkmark$	$\checkmark$	$\cdot$

- Also **Worst Case Necessary**: If the pair condition is not satisfied for some ordered pair of variables, there exist distinct **B**s that have identical distributions for the given experiments.
- Their learning algorithm is still **Complete**: Without further assumptions on the model space, the algorithm extracts all available information about **B**.

Identifiability from Overlapping Data Sets (Result 1)

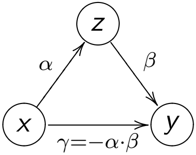
- We can extend the identifiability results given by Eberhardt et al. (2010) for experimental data sets sharing the same variables:
- A sufficient condition for fully identifying **B** is **the pair condition**: for each ordered pair of variables (u, v) there is an experiment such that u is intervened on and v is (passively) observed.

	x	y	z
x	$\cdot$	✓	✓
y	✓	$\cdot$	✓
z	✓	✓	$\cdot$

- Also **Worst Case Necessary**: If the pair condition is not satisfied for some ordered pair of variables, there exist distinct **B**s that have identical distributions for the given experiments.
- Their learning algorithm is still **Complete**: Without further assumptions on the model space, the algorithm extracts all available information about **B**.
- Usually, the overlapping experiments will not satisfy the pair condition and very little will be learned, so further assumptions are needed.

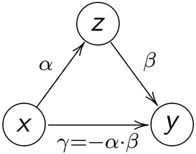
Representing Faithfulness Constraints with Exp. Effects (Result 2)

- All observed independencies structural
- No exactly canceling paths

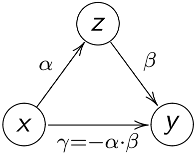


Representing Faithfulness Constraints with Exp. Effects (Result 2)

- All observed independencies structural
- No exactly canceling paths
- If we observe $x \perp\!\!\!\perp y$:

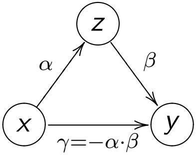


Representing Faithfulness Constraints with Exp. Effects (Result 2)



- All observed independencies structural
- No exactly canceling paths
- If we observe $x \perp\!\!\!\perp y$:
  - Directed paths
    $t(x \rightsquigarrow y || x) = 0, \quad t(y \rightsquigarrow x || y) = 0$
  - Paths through specific variables
    $t(x \rightsquigarrow z || x) t(z \rightsquigarrow y || z) = 0, \quad t(y \rightsquigarrow z || y) t(z \rightsquigarrow x || z) = 0$
  - Paths to each variable from a possible confounder
    $t(z \rightsquigarrow x || z) t(z \rightsquigarrow y || z) = 0$

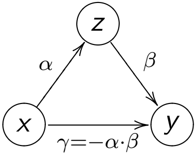
Representing Faithfulness Constraints with Exp. Effects (Result 2)



- All observed independencies structural
- No exactly canceling paths
- If we observe $x \perp\!\!\!\perp y$:
  - Directed paths
    $t(x \rightsquigarrow y || x) = 0, \quad t(y \rightsquigarrow x || y) = 0$
  - Paths through specific variables
    $t(x \rightsquigarrow z || x) t(z \rightsquigarrow y || z) = 0, \quad t(y \rightsquigarrow z || y) t(z \rightsquigarrow x || z) = 0$
  - Paths to each variable from a possible confounder
    $t(z \rightsquigarrow x || z) t(z \rightsquigarrow y || z) = 0$
- These apply for all supersets of the intervention sets. For example:

$$\underbrace{t(x \rightsquigarrow y || x)}_{\text{paths from } x \text{ to } y} = 0 \quad \Rightarrow \quad \underbrace{t(x \rightsquigarrow y || x, z)}_{\text{not through } z} = 0$$

Representing Faithfulness Constraints with Exp. Effects (Result 2)



- All observed independencies structural
- No exactly canceling paths
- If we observe $x \perp\!\!\!\perp y$:
  - Directed paths
    $t(x \rightsquigarrow y || x) = 0, \quad t(y \rightsquigarrow x || y) = 0$
  - Paths through specific variables
    $t(x \rightsquigarrow z || x) t(z \rightsquigarrow y || z) = 0, \quad t(y \rightsquigarrow z || y) t(z \rightsquigarrow x || z) = 0$
  - Paths to each variable from a possible confounder
    $t(z \rightsquigarrow x || z) t(z \rightsquigarrow y || z) = 0$
- These apply for all supersets of the intervention sets. For example:

$$\underbrace{t(x \rightsquigarrow y || x)}_{\text{paths from } x \text{ to } y} = 0 \quad \Rightarrow \quad \underbrace{t(x \rightsquigarrow y || x, z)}_{\text{not through } z} = 0$$

- Other rules extensions to cyclic models from Claassen et al. (2011)

Problem

- Some 'observed' experimental effects.

$$t(x \rightsquigarrow z || x, w) = 0.654, \quad t(w \rightsquigarrow y || x, w) = -0.345, \quad \ldots$$

- Some equations implied by faithfulness & detected independencies:

$$t(x \rightsquigarrow y || x, w) = 0, \quad t(x \rightsquigarrow z || x, w) t(z \rightsquigarrow y || x, z, w) = 0, \quad \ldots$$

- Output the known coefficients $b(x \rightarrow y)$

Problem Solution: Linear Inference -algorithm (Result 3)

- Some 'observed' experimental effects.

$$t(x \rightsquigarrow z || x, w) = 0.654, \quad t(w \rightsquigarrow y || x, w) = -0.345, \quad \ldots$$

- Some equations implied by faithfulness & detected independencies:

$$t(x \rightsquigarrow y || x, w) = 0, \quad t(x \rightsquigarrow z || x, w) t(z \rightsquigarrow y || x, z, w) = 0, \quad \ldots$$

- Output the known coefficients $b(x \to y)$

Problem Solution: Linear Inference -algorithm (Result 3)

- **Overparametrize:** Rather than only $b(x \to y) = t(x \rightsquigarrow y || \mathcal{V} \setminus y)$, consider all exp. effects $t(x \rightsquigarrow y || \mathcal{J})$ as unknown model parameters.

- Some 'observed' experimental effects.

$$t(x \rightsquigarrow z || x, w) = 0.654, \quad t(w \rightsquigarrow y || x, w) = -0.345, \quad \ldots$$

- Some equations implied by faithfulness & detected independencies:

$$t(x \rightsquigarrow y || x, w) = 0, \quad t(x \rightsquigarrow z || x, w) t(z \rightsquigarrow y || x, z, w) = 0, \quad \ldots$$

- Output the known coefficients $b(x \to y)$

Problem Solution: Linear Inference -algorithm (Result 3)

- **Overparametrize:** Rather than only $b(x \to y) = t(x \rightsquigarrow y || \mathcal{V} \setminus y)$, consider all exp. effects $t(x \rightsquigarrow y || \mathcal{J})$ as unknown model parameters.
- Generate all **constraint** equations relating exp. effects.

$$t(x \rightsquigarrow y || x, w) \quad = \quad t(x \rightsquigarrow y || x, z, w) + t(x \rightsquigarrow z || x, w) t(z \rightsquigarrow y || x, z, w)$$

- Some 'observed' experimental effects.

$$t(x \rightsquigarrow z || x, w) = 0.654, \quad t(w \rightsquigarrow y || x, w) = -0.345, \quad \ldots$$

- Some equations implied by faithfulness & detected independencies:

$$t(x \rightsquigarrow y || x, w) = 0, \quad t(x \rightsquigarrow z || x, w) t(z \rightsquigarrow y || x, z, w) = 0, \quad \ldots$$

- Output the known coefficients $b(x \to y)$

Problem Solution: Linear Inference -algorithm (Result 3)

- **Overparametrize:** Rather than only $b(x \to y) = t(x \rightsquigarrow y || \mathcal{V} \setminus y)$, consider all exp. effects $t(x \rightsquigarrow y || \mathcal{J})$ as unknown model parameters.
- Generate all **constraint** equations relating exp. effects.

$$t(x \rightsquigarrow y || x, w) \quad = \quad t(x \rightsquigarrow y || x, z, w) + t(x \rightsquigarrow z || x, w) t(z \rightsquigarrow y || x, z, w)$$

- Some 'observed' experimental effects. Mark them as known.

$$t(x \rightsquigarrow z || x, w) = 0.654, \quad t(w \rightsquigarrow y || x, w) = -0.345, \quad \ldots$$

- Some equations implied by faithfulness & detected independencies:

$$t(x \rightsquigarrow y || x, w) = 0, \quad t(x \rightsquigarrow z || x, w) t(z \rightsquigarrow y || x, z, w) = 0, \quad \ldots$$

- Output the known coefficients $b(x \to y)$

Problem Solution: Linear Inference -algorithm (Result 3)

- **Overparametrize:** Rather than only $b(x \to y) = t(x \rightsquigarrow y || \mathcal{V} \setminus y)$, consider all exp. effects $t(x \rightsquigarrow y || \mathcal{J})$ as unknown model parameters.
- Generate all **constraint** equations relating exp. effects.

$$t(x \rightsquigarrow y || x, w) \;=\; t(x \rightsquigarrow y || x, z, w) + t(x \rightsquigarrow z || x, w) t(z \rightsquigarrow y || x, z, w)$$

- Some 'observed' experimental effects. Mark them as known.

$$t(x \rightsquigarrow z || x, w) = 0.654, \quad t(w \rightsquigarrow y || x, w) = -0.345, \quad \ldots$$

- Some equations implied by faithfulness & detected independencies:

$$t(x \rightsquigarrow y || x, w) = 0, \quad t(x \rightsquigarrow z || x, w) t(z \rightsquigarrow y || x, z, w) = 0, \quad \ldots$$

- Collect all equations that are **linear** on the unknowns, and determine the uniquely identified unknowns of the linear system.

- Output the known coefficients $b(x \to y)$

Problem Solution: Linear Inference -algorithm (Result 3)

- **Overparametrize:** Rather than only $b(x \to y) = t(x \rightsquigarrow y || \mathcal{V} \setminus y)$, consider all exp. effects $t(x \rightsquigarrow y || \mathcal{J})$ as unknown model parameters.
- Generate all **constraint** equations relating exp. effects.

$$t(x \rightsquigarrow y || x, w) \quad = \quad t(x \rightsquigarrow y || x, z, w) + t(x \rightsquigarrow z || x, w) t(z \rightsquigarrow y || x, z, w)$$

- Some 'observed' experimental effects. Mark them as known.

$$t(x \rightsquigarrow z || x, w) = 0.654, \quad t(w \rightsquigarrow y || x, w) = -0.345, \quad \ldots$$

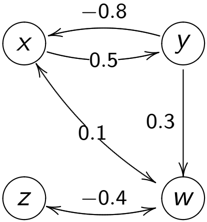
- Some equations implied by faithfulness & detected independencies:

$$t(x \rightsquigarrow y || x, w) = 0, \quad t(x \rightsquigarrow z || x, w) t(z \rightsquigarrow y || x, z, w) = 0, \quad \ldots$$

- Collect all equations that are **linear** on the unknowns, and determine the uniquely identified unknowns of the linear system.
- Rerun: more equations might become linear!
- Output the known coefficients $b(x \to y)$

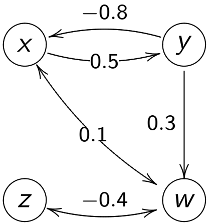
Toy Example

True Model



Toy Example

True Model



Experiment 1

$\mathbf{x}$	y	w
$-\mathbf{0.2}$	0.8	1.2
$\vdots$	$\vdots$	$\vdots$

$t(x \rightsquigarrow y || x) = 0.5$

$t(x \rightsquigarrow w || x) = 0.15$

$x \perp\!\!\!\perp w | y \ \ || x$

Experiment 2

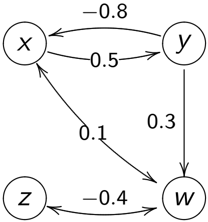
y	z	w
-0.3	-0.7	1.4
$\vdots$	$\vdots$	$\vdots$

$y \perp\!\!\!\perp z$

$y \not\perp\!\!\!\perp z | w$

Toy Example

True Model



Experiment 1

x	y	w
$-\mathbf{0.2}$	0.8	1.2
⋮	⋮	⋮

$t(x \rightsquigarrow y || x) = 0.5$

$t(x \rightsquigarrow w || x) = 0.15$

$x \perp\!\!\!\perp w | y \quad || x$

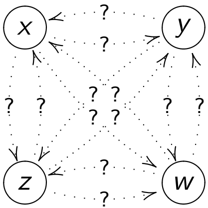
Experiment 2

y	z	w
-0.3	-0.7	1.4
⋮	⋮	⋮

$y \perp\!\!\!\perp z$

$y \not\!\perp\!\!\!\perp z | w$

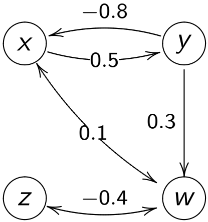
N.a. faithfulness:



(double-headed edges not shown here)

Toy Example

True Model



Experiment 1

$\mathbf{x}$	y	w
$-\mathbf{0.2}$	0.8	1.2
⋮	⋮	⋮

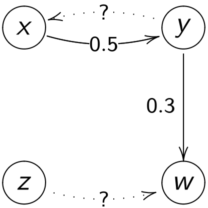
$$t(x \rightsquigarrow y || x) = 0.5$$
$$t(x \rightsquigarrow w || x) = 0.15$$
$$x \perp\!\!\!\perp w | y \quad || x$$

Experiment 2

y	z	w
−0.3	−0.7	1.4
⋮	⋮	⋮

$$y \perp\!\!\!\perp z$$
$$y \not\!\perp\!\!\!\perp z | w$$

Linear Inference:

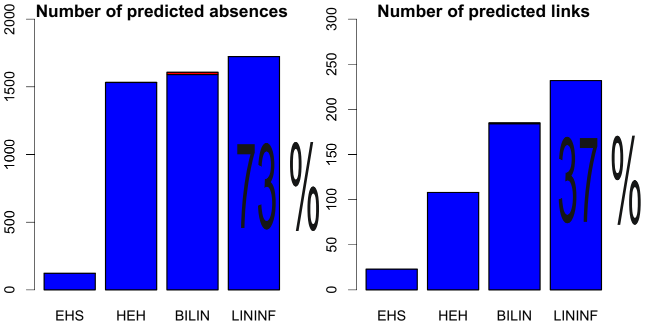


(double-headed edges not shown here)

Simulations

- 100 random models with 6 obs. variables, 5 random exp. settings
- 624 present links (20%), 2376 absences of links

Infinite sample limit
No prediction error
Independence oracle

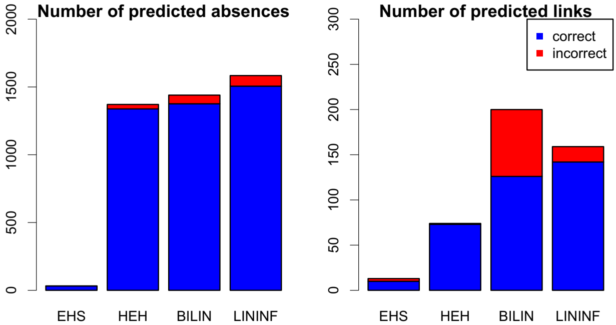


- EHS: Algorithm shown complete when NOT assuming faithfulness.
- HEH: Utilizing faithfulness for data sets sharing the same variables.
- BILIN: Another method in the article, cut out of the presentation.
- LININF: Linear Inference -algorithm just presented.

Simulations

- 100 random models with 6 obs. variables, 5 random exp. settings
- 624 present links (20%), 2376 absences of links

1000 samples/exp.

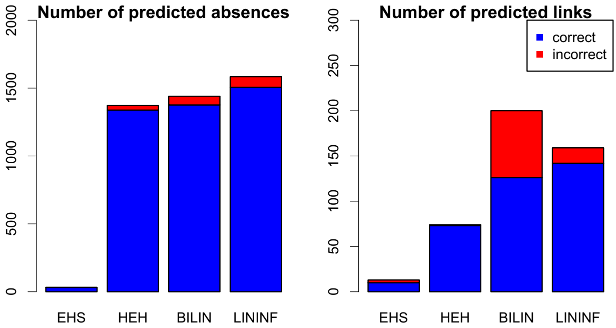


- EHS: Algorithm shown complete when NOT assuming faithfulness.
- HEH: Utilizing faithfulness for data sets sharing the same variables.
- BILIN: Another method in the article, cut out of the presentation.
- LININF: Linear Inference -algorithm just presented.

Simulations

- 100 random models with 6 obs. variables, 5 random exp. settings
- 624 present links (20%), 2376 absences of links

1000 samples/exp.

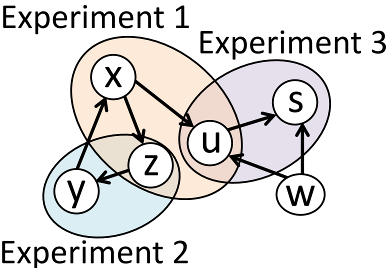


- EHS: Algorithm shown complete when NOT assuming faithfulness.
- HEH: Utilizing faithfulness for data sets sharing the same variables.
- BILIN: Another method in the article, cut out of the presentation.
- LININF: Linear Inference -algorithm just presented.
- **Which of the unidentified links are inherently undetermined?**

Conclusion

General learning setting:

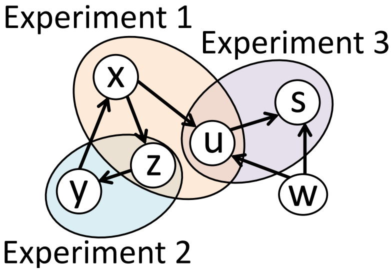
- Multiple completely or partially overlapping data sets
- Experimental or non-experimental data
- Cycles, No joint causal sufficiency, Linear causal relations



Conclusion

General learning setting:

- Multiple completely or partially overlapping data sets
- Experimental or non-experimental data
- Cycles, No joint causal sufficiency, Linear causal relations



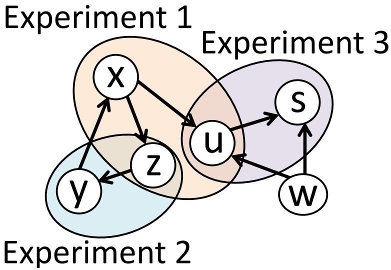
Contributions:

- Completeness and Identifiability results without faithfulness
- Faithfulness constraints decoded with exp. effects
- Linear Inference -algorithm by **overparametrizing**, generalizing previous **constraint** equations, and enforcing only the **linear** equations

Conclusion

General learning setting:

- Multiple completely or partially overlapping data sets
- Experimental or non-experimental data
- Cycles, No joint causal sufficiency, Linear causal relations



Contributions:

- Completeness and Identifiability results without faithfulness
- Faithfulness constraints decoded with exp. effects
- Linear Inference -algorithm by **overparametrizing**, generalizing previous **constraint** equations, and enforcing only the **linear** equations

Thank you!