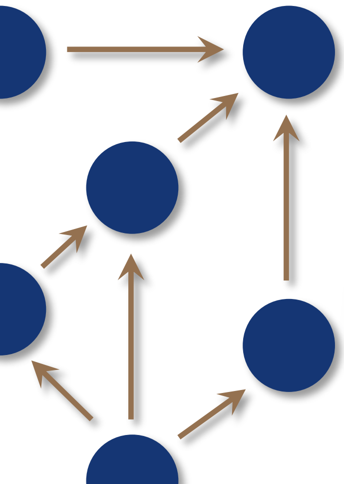
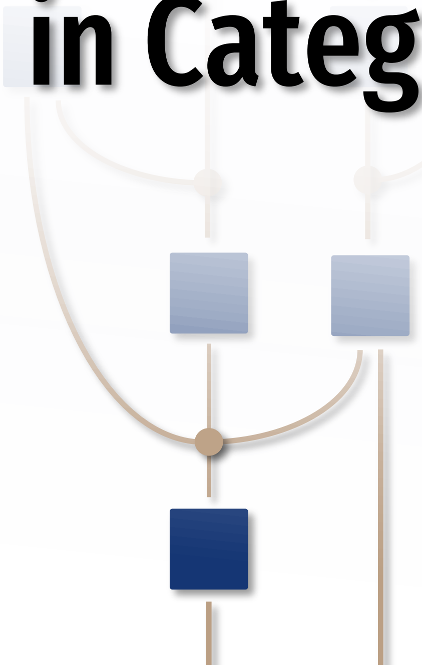
The d-separation criterion in Categorical Probability

Andreas Klingler

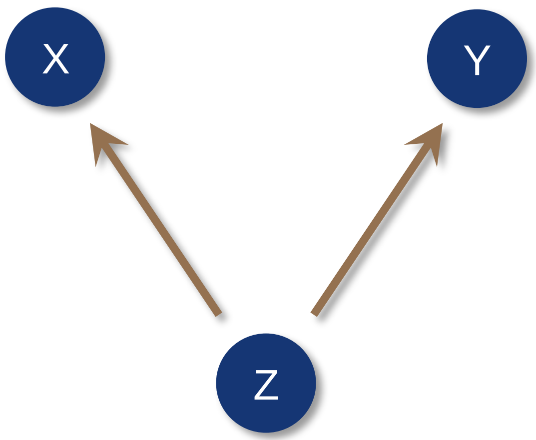
joint with Tobias Fritz

arXiv:2207.05740



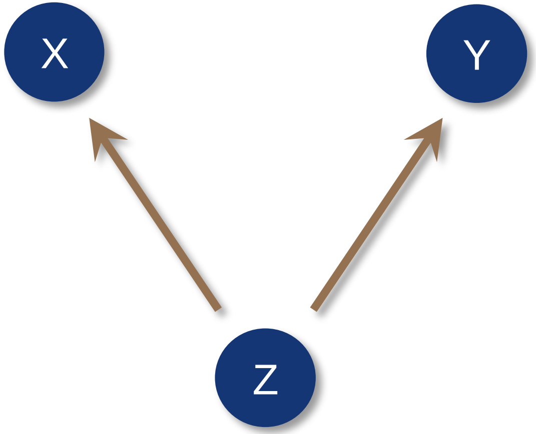


Bayesian networks



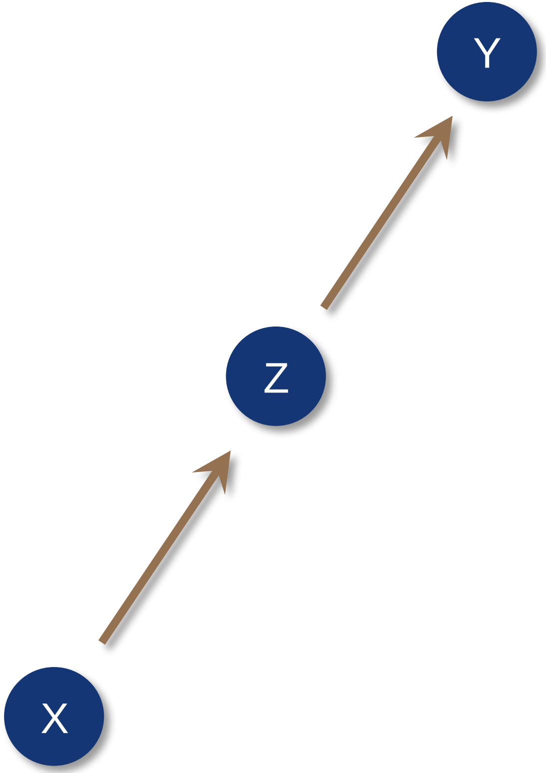
$$P(XYZ) = P(X|Z) \cdot P(Y|Z) \cdot P(Z)$$

How to test a causal structure?



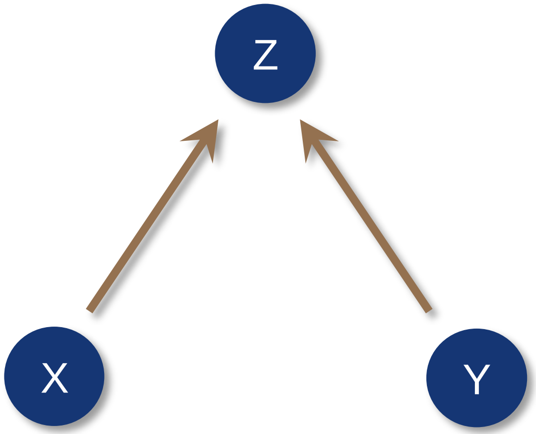
$$X \leftarrow Z \rightarrow Y$$

$$X \perp Y \mid Z$$



$$X \rightarrow Z \rightarrow Y$$

$$X \perp Y \mid Z$$



$X \to Z \leftarrow Y$

~~$X \perp Y \mid Z$~~

The d-separation criterion

P is compatible with a causal structure

$\Updownarrow$

d-separation in the causal structure

$\Longrightarrow$ conditional independence

The d-separation criterion

- Discrete random variables

  J. Pearl. Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. 1988.

  T. Verma and J. Pearl. Causal networks: Semantics and expressiveness. Machine intelligence and pattern recognition. 1990.

- Continuous and mixed random variables

  S. L. Lauritzen et al. Independence properties of directed Markov fields. Networks, 1990.

  P. Forré and J. M. Mooij. Markov properties for graphical models with cycles and latent variables. 2017.

d-separation

based on string diagrams

as causal models

Classical approach	Categorical approach
Specific proof for each probability class	Independent of probability class
Proof via measure theory	Diagrammatic reasoning
Probability distributions	Stochastic maps
d-separation on DAGs	Connectedness of string diagrams

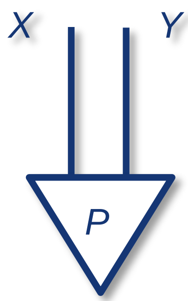
1. Categorical Probability: Markov categories
   Generalized causal models
2. Conditional independence
   Categorical d-separation
3. The d-separation criterion

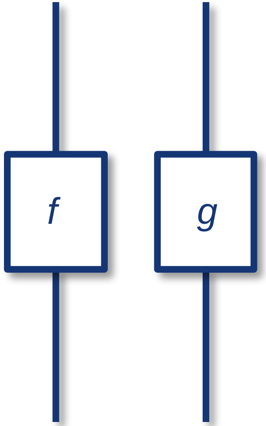
Markov categories – modeling information flow

Symmetric monoidal category

Objects: sample spaces

Morphisms: Stochastic maps





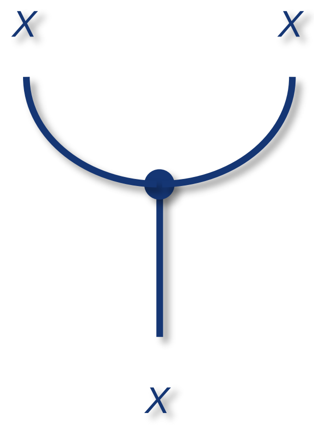
T. Fritz. A synthetic approach to Markov kernels, conditional independence and theorems on sufficient statistics. Adv. Math. 2020

Markov categories – modeling information flow

Symmetric monoidal category with a comonoid structure

Objects: sample spaces

Morphisms: Stochastic maps



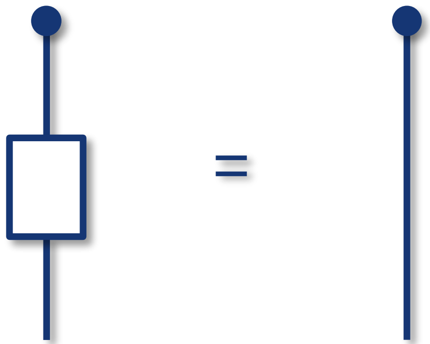
T. Fritz. A synthetic approach to Markov kernels, conditional independence and theorems on sufficient statistics. Adv. Math. 2020

Markov categories – modeling information flow

Symmetric monoidal category with a comonoid structure

Objects: sample spaces

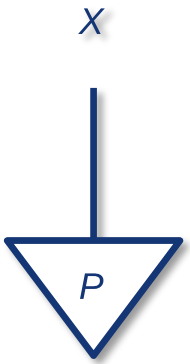
Morphisms: Stochastic maps



T. Fritz. A synthetic approach to Markov kernels, conditional independence and theorems on sufficient statistics. Adv. Math. 2020

Examples -- FinStoch

Objects: finite sets

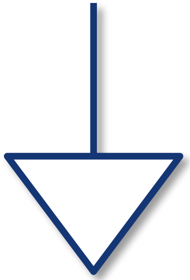


$$P: \quad X \to [0,1]$$
$$x \mapsto p_x$$

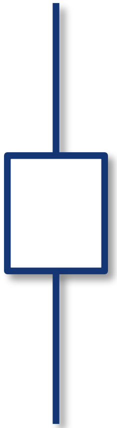
$$f: \quad X \times Y \to [0,1]$$
$$(x, y) \mapsto f(y|x)$$

Examples -- Gauss

Objects: $\mathbb{R}^n$



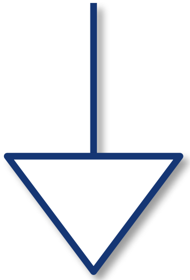
Gaussian vector



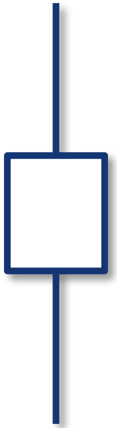
$$x \mapsto Ax + \xi$$

Examples -- Stoch

Objects: measurable spaces (X, Σ_X)



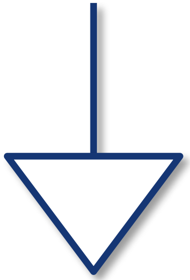
$P: \Sigma_X \to [0,1]$



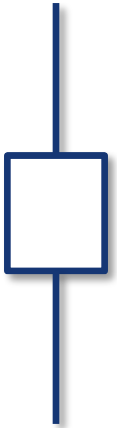
Measurable Markov kernels

Examples -- BorelStoch

Objects: standard Borel spaces $(X, \mathcal{B}(X))$



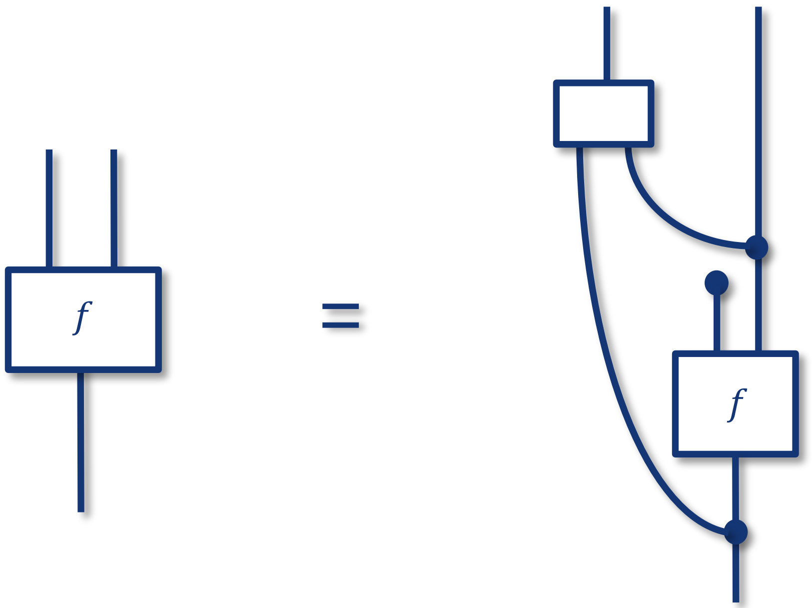
$P: \mathcal{B}(X) \to [0,1]$



Measurable Markov kernels

Markov categories with conditionals

A Markov category has **conditionals** if every morphism f



T. Fritz. A synthetic approach to Markov kernels, conditional independence and theorems on sufficient statistics. Adv. Math. 2020
K. Cho, B. Jacobs. Disintegration and Bayesian inversion via string diagrams. Math. Struct. Comp. Sci. 2019

Existence of conditionals

FinStoch	✓	**BorelStoch**	✓
Gauss	✓	**Stoch**	✗

T. Fritz. A synthetic approach to Markov kernels, conditional independence and theorems on sufficient statistics. Adv. Math. 2020
V. Bogachev, I. Malofeev. Kantorovich problems and conditional measures depending on a parameter. J. Math. Anal. Appl. 2020
J. L. Doob. Stochastic processes. 1953

Concrete Markov cateogories:

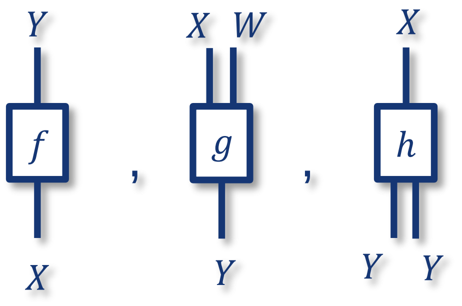
FinStoch

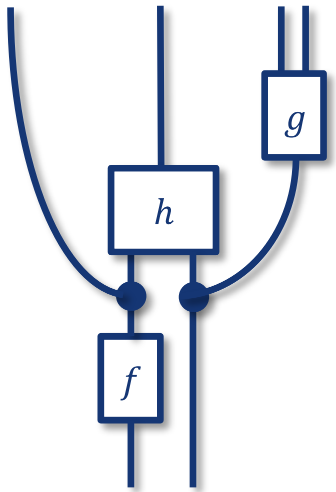
Gauss

BorelStoch

Stoch

Free Markov categories:



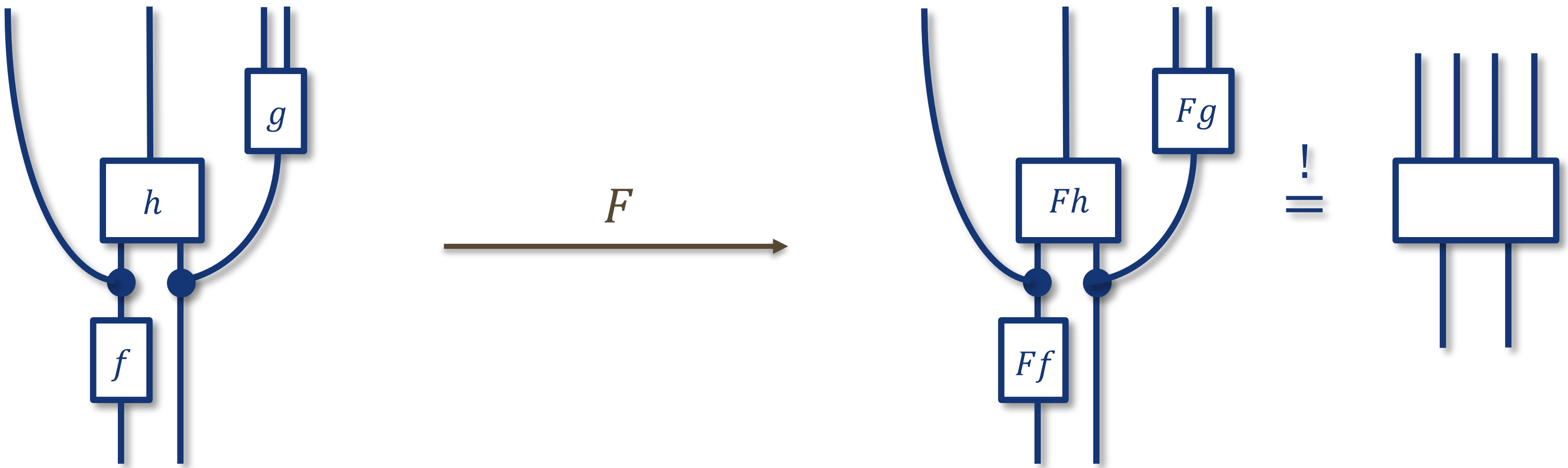


T. Fritz, W. Liang. Free gs-monoidal categories and free Markov categories. 2022

Compatibility

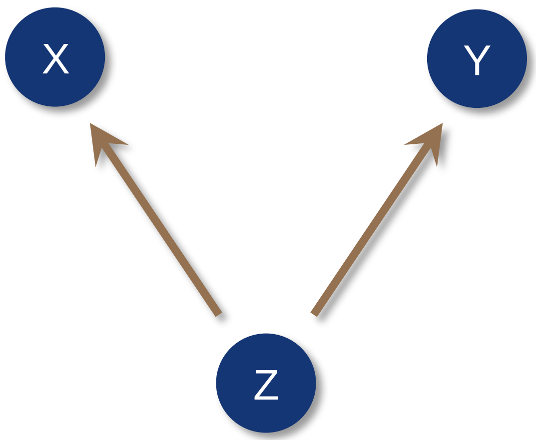
Free Markov category

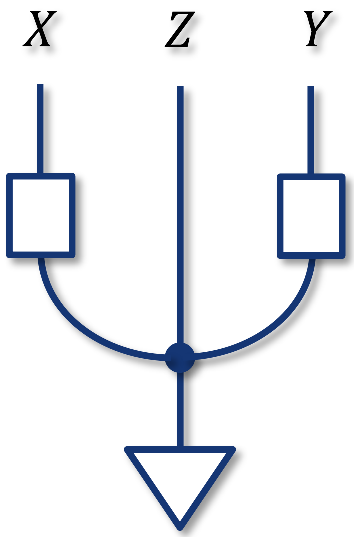
Concrete Markov category



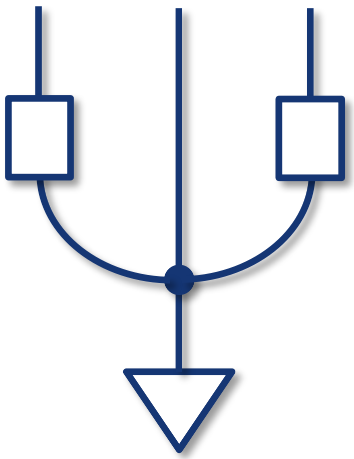
T. Fritz, W. Liang. Free gs-monoidal categories and free Markov categories. 2022

Examples: DAG models



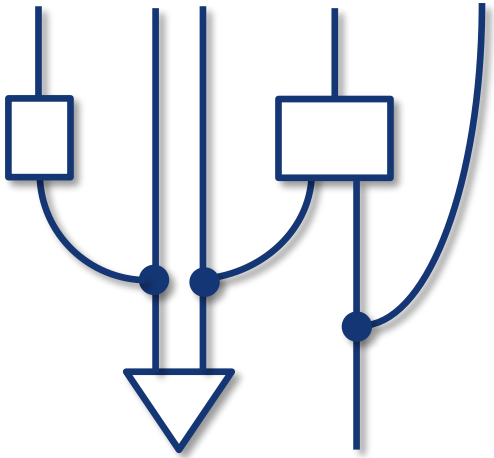


Examples: DAG models



- Every box has one output
- No global input wires
- Every wire is a global output

Generalized causal models



- ~~Every box has one output~~
- ~~No global input wires~~
- Every wire is a global output

1

Categorical Probability Markov categories

Generalized causal models

2

Conditional independence

Categorical d-separation

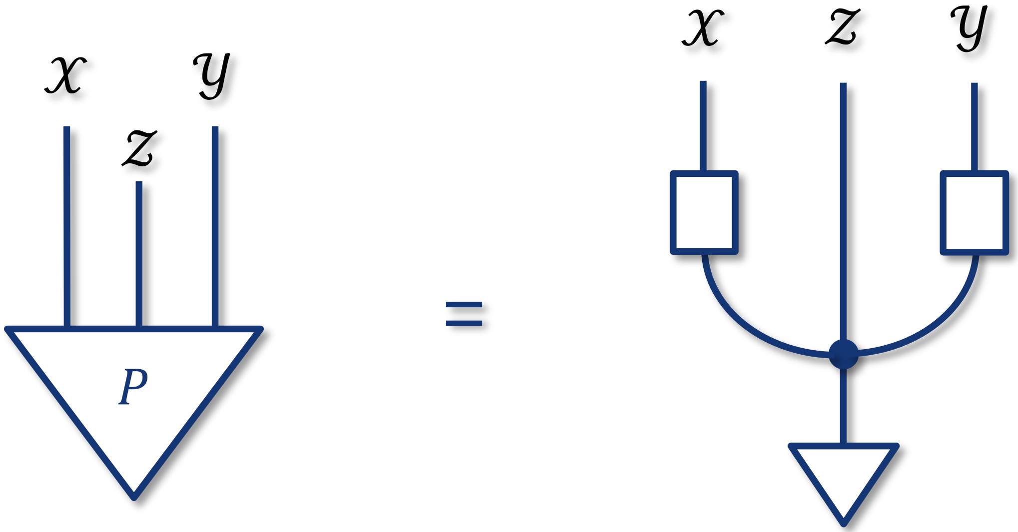
3

The d-separation criterion

Conditional independence (for probabilities)

$\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ sets of wires.

P shows **conditional independence** $\mathcal{X} \perp \mathcal{Y} \mid \mathcal{Z}$, if



d-separation

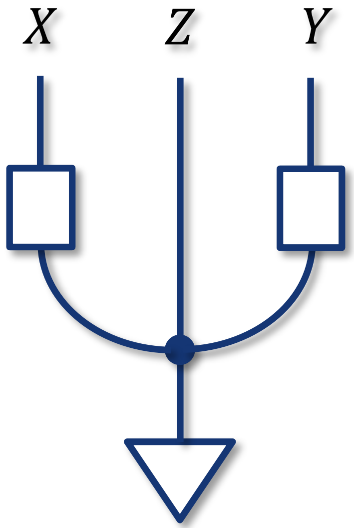
$\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ disjoint sets of wires.

$\mathcal{Z}$ **d-separates** $\mathcal{X}$ and $\mathcal{Y}$

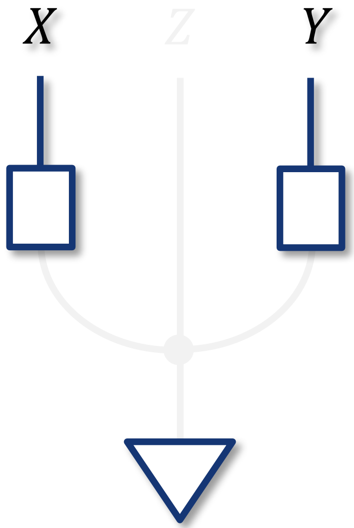
if

$\mathcal{X}$ and $\mathcal{Y}$ **disconnected** when removing $\mathcal{Z}$

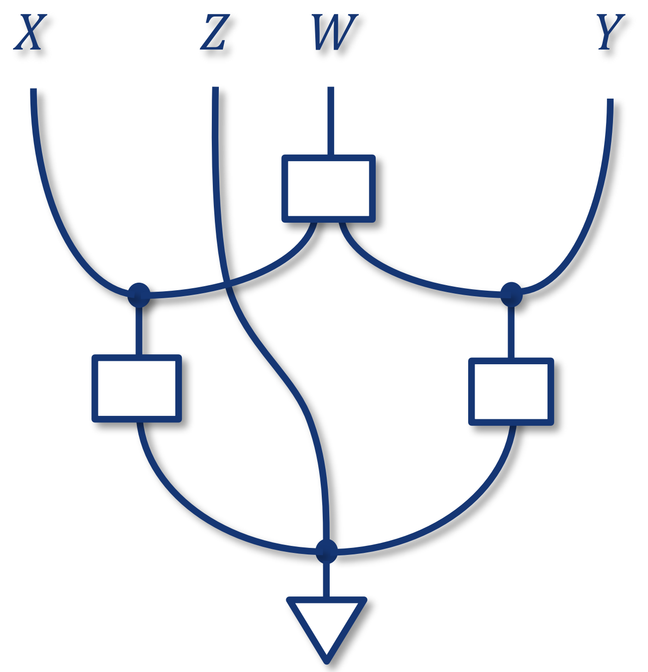
Z **d-separates** X and Y



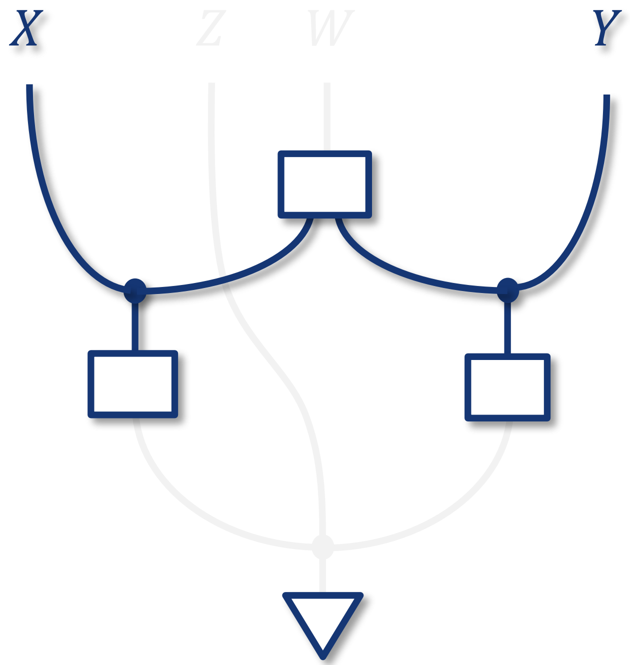
Z **d-separates** X and Y



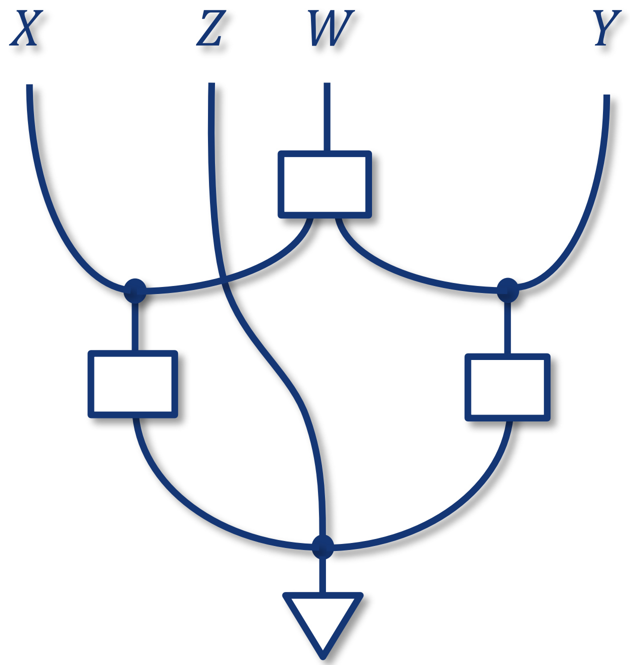
Z, W do not **d-separate** X and Y



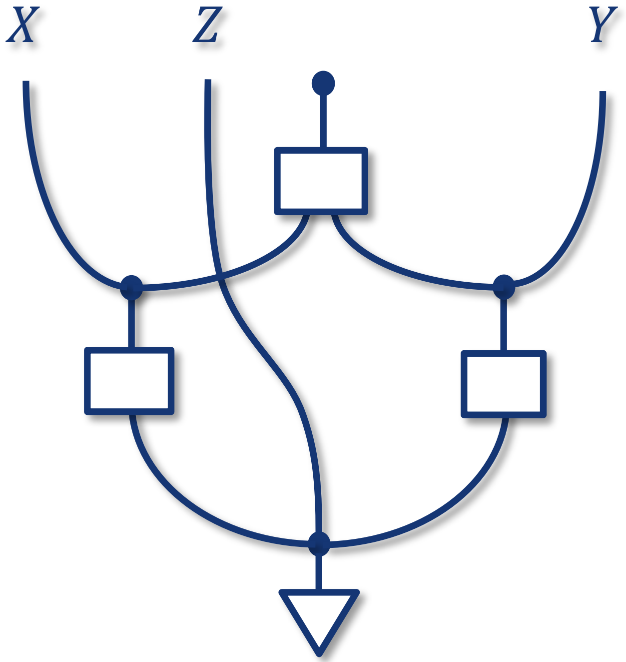
Z, W do not **d-separate** X and Y



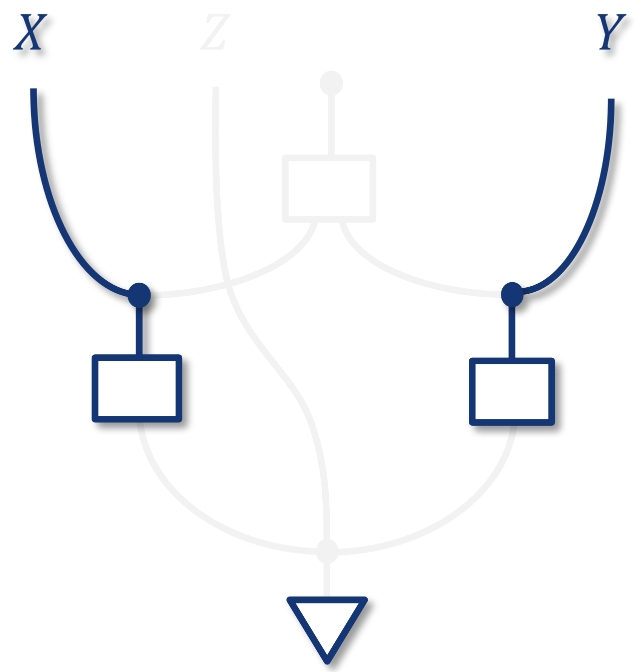
Z **d-separates** X and Y



Z **d-separates** X and Y



Z **d-separates** X and Y



1

Categorical Probability Markov categories

Generalized causal models

2

Conditional independence

Categorical d-separation

3

The d-separation criterion

Soundness of the d-Separation criterion

C a Markov category with conditionals

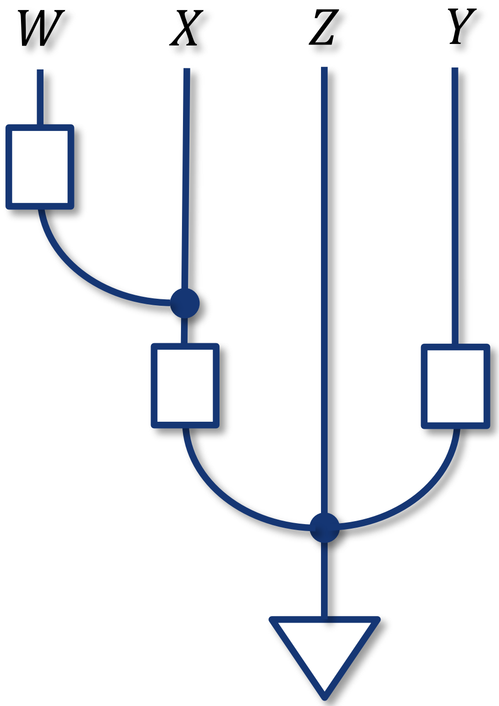
P is compatible with a causal structure

⇓

d-separation in the string diagram

$\Longrightarrow$ conditional independence

Example

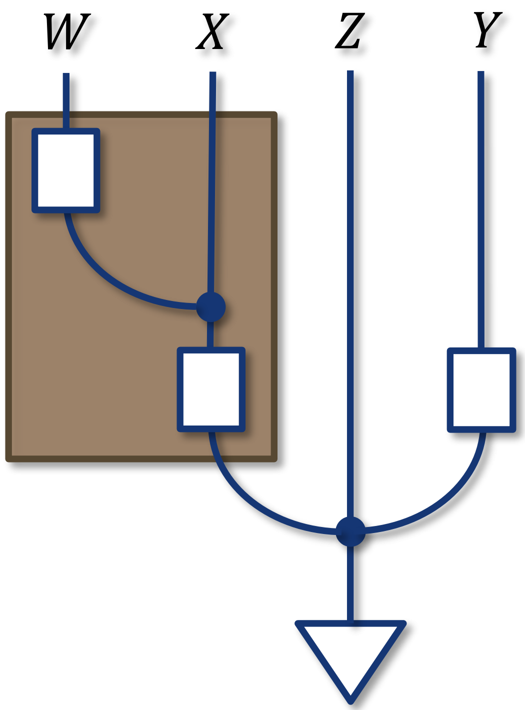


Z, W d-separates X and Y

⇓ ?

$X \perp Y \mid Z, W$

Example

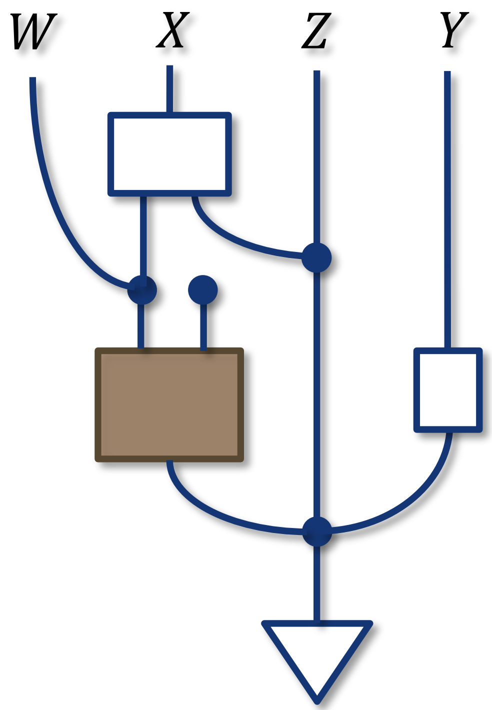


Z, W d-separates X and Y

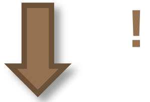
⇩ ?

$$X \perp Y \mid Z, W$$

Example



Z, W d-separates X and Y



$$X \perp Y \mid Z, W$$

The reverse direction

C a Markov category with conditionals

P is compatible with a causal structure

⇑

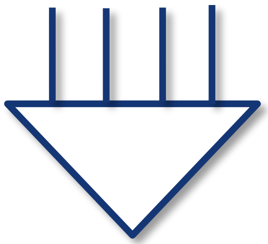
Every d-separation in the string diagram

implies conditional independence

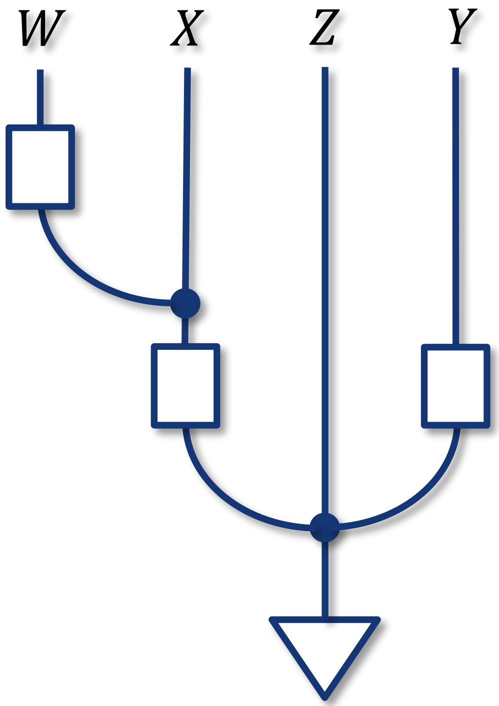
Global Markov property

Example

P satisfies global Markov property



$\overset{?}{=}$

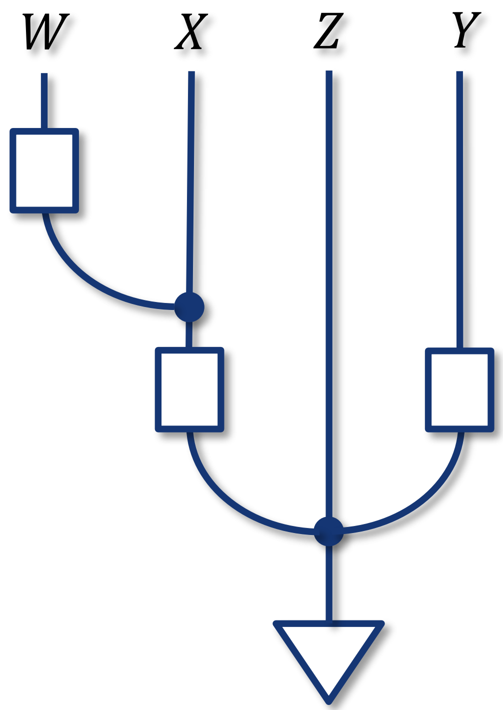


Example

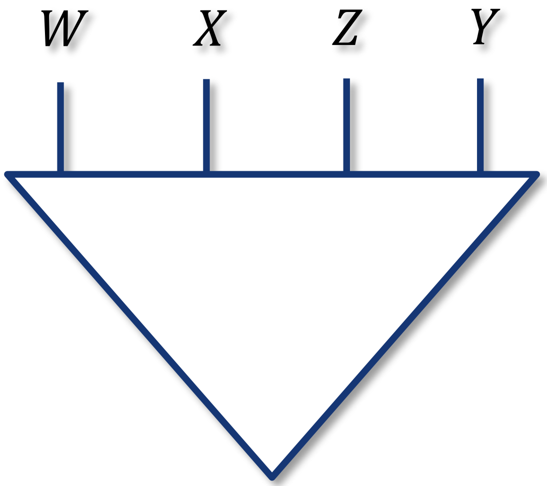
P satisfies global Markov property

$$W \perp YZ \mid X$$

$$X \perp Y \mid Z$$



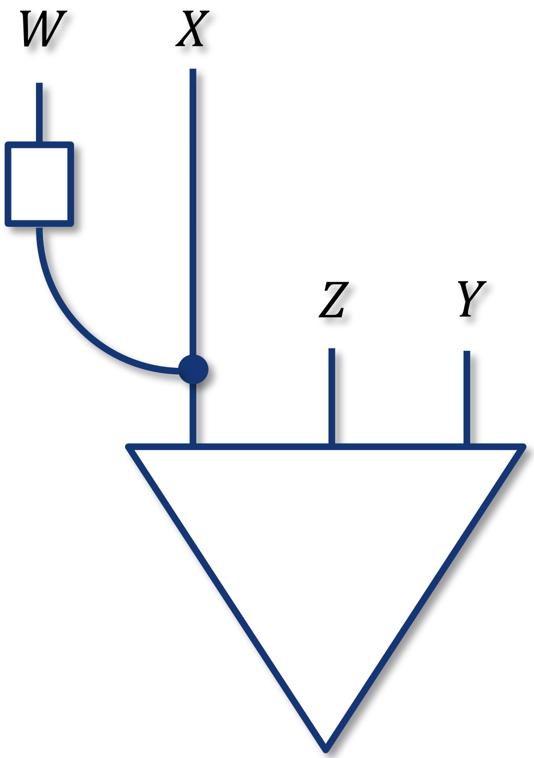
Example



$$W \perp YZ \mid X$$

$$X \perp Y \mid Z$$

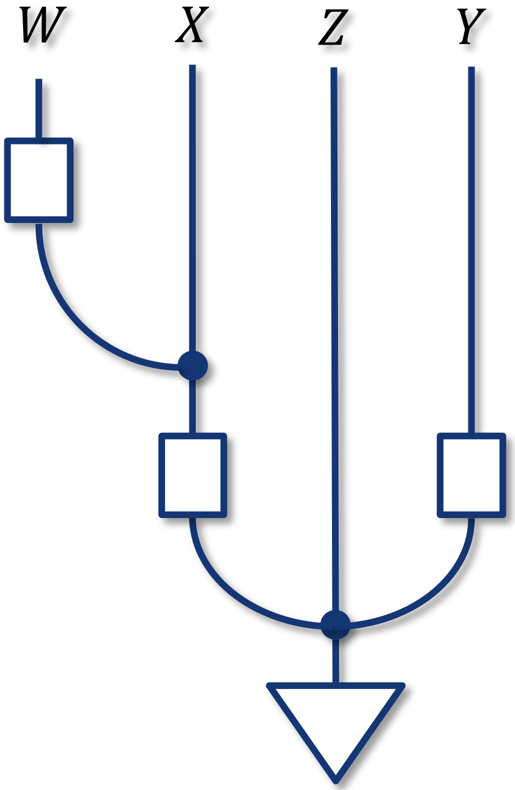
Example



$W \perp YZ \mid X$

$X \perp Y \mid Z$

Example

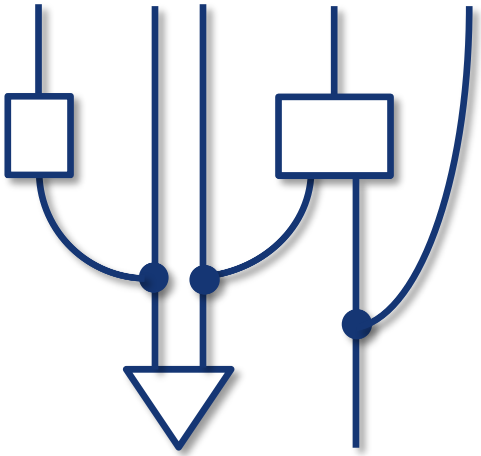


$W \perp YZ \mid X$

$X \perp Y \mid Z$

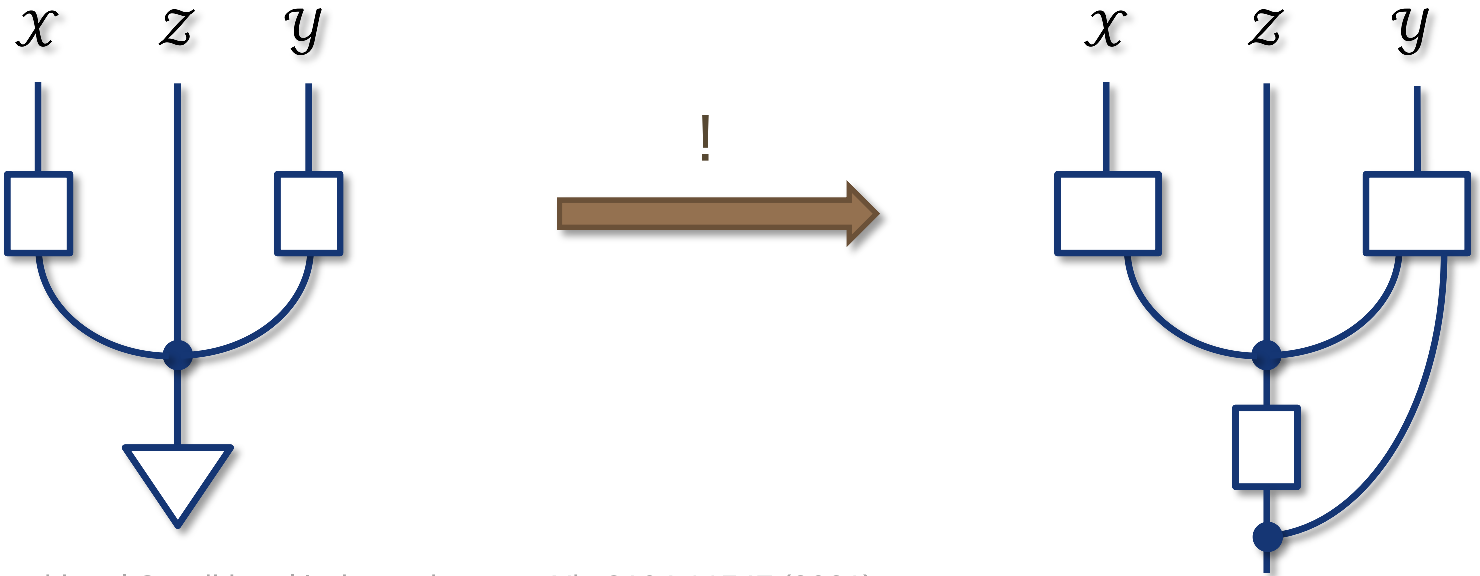
Things I have not talked about ...

The d-separation criterion for Markov kernels



Things I have not talked about ...

$$\mathcal{X} \perp \mathcal{Y} \mid \mathcal{Z}$$



P.Forré .Transitional Conditional Independence. arXiv:2104.11547 (2021)

Summary:

- d-separation criterion on **generalized causal models**
- Categorical d-separation $\leftrightarrow$ **disconnectedness** of string diagrams
- Proof of the d-separation criterion for **BorelStoch**, **Gauss**, **FinStoch**

Open questions:

- Extending the d-separation criterion to **symmetric** causal models
- Applying d-separation to **"exotic"** Markov categories
- **Disconnectedness** of string diagrams $\leftrightarrow$ **conditional independence**:

  Are string diagrams a **more natural** representation of causal models than DAGs?

Andreas Klingler

andreas.klingler@uibk.ac.at

Tobias Fritz

tobias.fritz@uibk.ac.at

arXiv:2207.05740