A very gentle introduction to

Causality

“The Art and Science of Cause and Effect”

Outline

1. History of causality & why causality matters
2. How to approach causal inference
3. Applications; connection to Natural Language Understanding

In J. Pearl, Causality: Models, Reasoning, and Inference, New York: Cambridge University Press, pp. 401-428, 2009.

EPILOGUE

The Art and Science of Cause and Effect

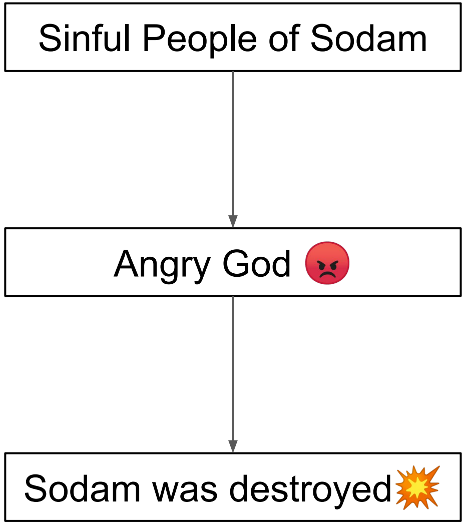
A public lecture delivered November 1996 as part of the UCLA Faculty Research Lectureship Program

Outline

1. History of causality & why causality matters
2. How to approach causal inference
3. Applications; connection to Natural Language Understanding

A Tale of Cause and Effect

“The Flight of Lot and His Family from Sodom” by Peter Paul Rubens, ~1614



A Tale of Cause and Effect

"The Flight of Lot and His Family from Sodom" by Peter Paul Rubens, ~1614

Sinful People of Sodam

$\downarrow$

Sodam was destroyed💥

As a modern statistician will tell you...
Correlation ≠ Causation

Causal Relations are Difficult to Examine

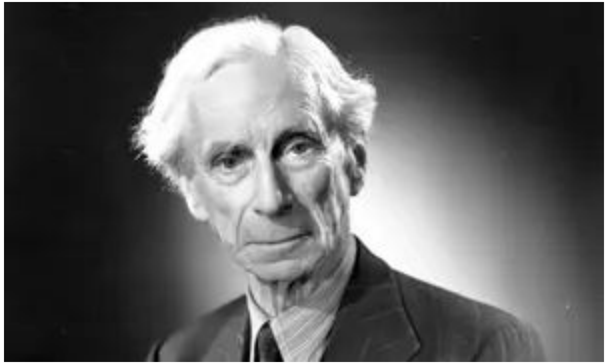
What kind of **empirical evidence** is sufficient to prove causal relation?

But wait… Do we even need Causality?

Statisticians from 100 years ago started to wonder...

"All philosophers imagine that causation is one of the fundamental axioms of science, yet oddly enough, ***in advanced sciences, the word 'cause' never occurs****. . . .*

The law of causality, I believe, is a relic of bygone age, surviving, like the monarchy, only because it is erroneously supposed to do no harm."



Philosopher Bertrand Russell

The Invention of Correlation

- Early 20th century statisticians don't see why causation need to be studied as an independent concept beyond correlation.

- Quote from Pearson --

  "Beyond such discarded fundamentals as 'matter' and 'force' lies still another fetish amidst the inscrutable arcana of modern science, namely, the category of cause and effect."

Karl Pearson (1857-1936)
English Mathematician

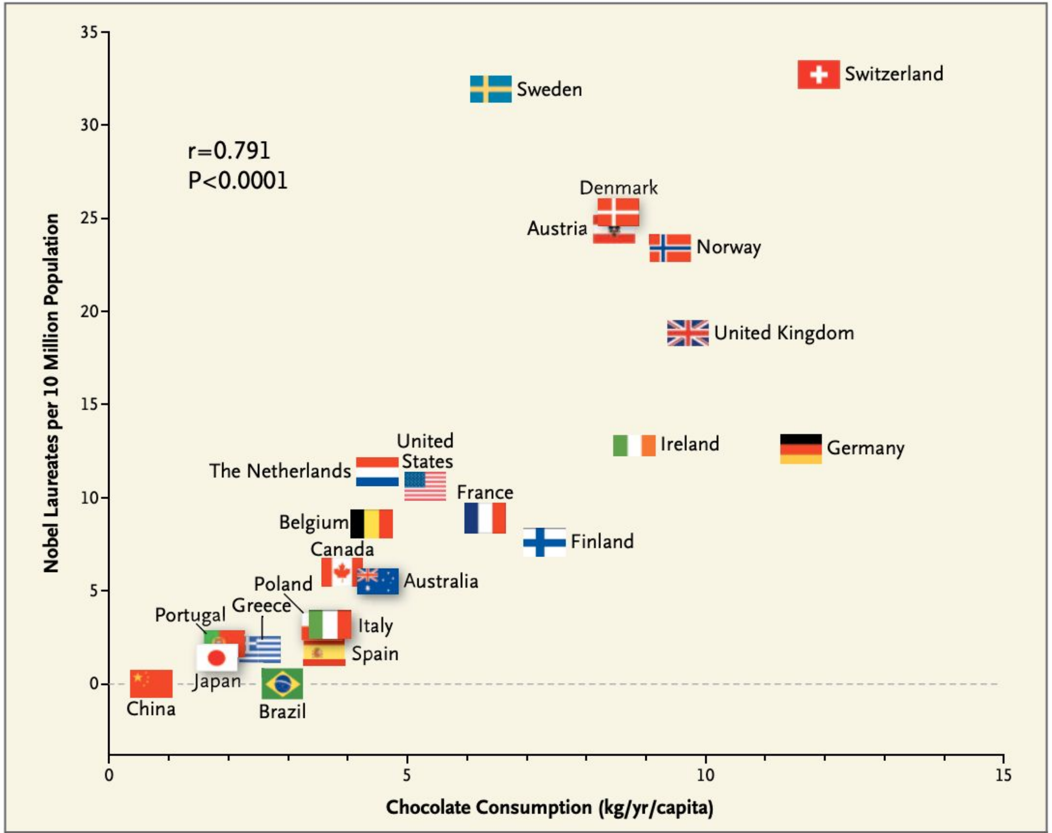


Figure 1. Correlation between Countries' Annual Per Capita Chocolate Consumption and the Number of Nobel Laureates per 10 Million Population.

F. H. Messerli: Chocolate Consumption, Cognitive Function, and Nobel Laureates, N Engl J Med 2012

TIME

FOOD & DRINK

Secret to Winning a Nobel Prize? Eat More Chocolate

By Olivia B. Waxman @OBWax | Oct. 12, 2012

Getty Imag

As the Nobel Prizes are being awarded this week, one U.S. scientist asks: could eating chocolate have anything to do with becoming a laureate?

Why would the sweet treat be linked to winning the most prestigious intellectual award, you ask? In a "note" published in the *New England Journal of Medicine*, Dr. Franz H. Messerli, a cardiologist at St. Luke's–Roosevelt Hospital in New York City, writes that cocoa contains flavanols, plant-based compounds that previous studies have linked to the slowing or reversing of age-related cognitive decline. (You can also get flavanols in green tea, red wine and

Outline

1. History of causality & why causality matters
2. How to approach causal inference
3. Applications; connection to Natural Language Understanding

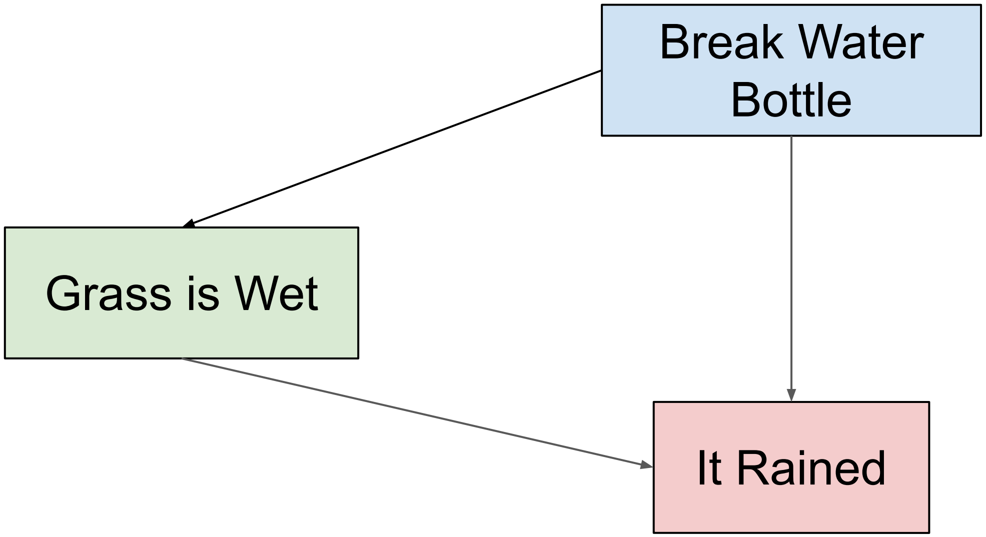
Causal Inference - Example

Rule #1: If the grass is wet, then it rained

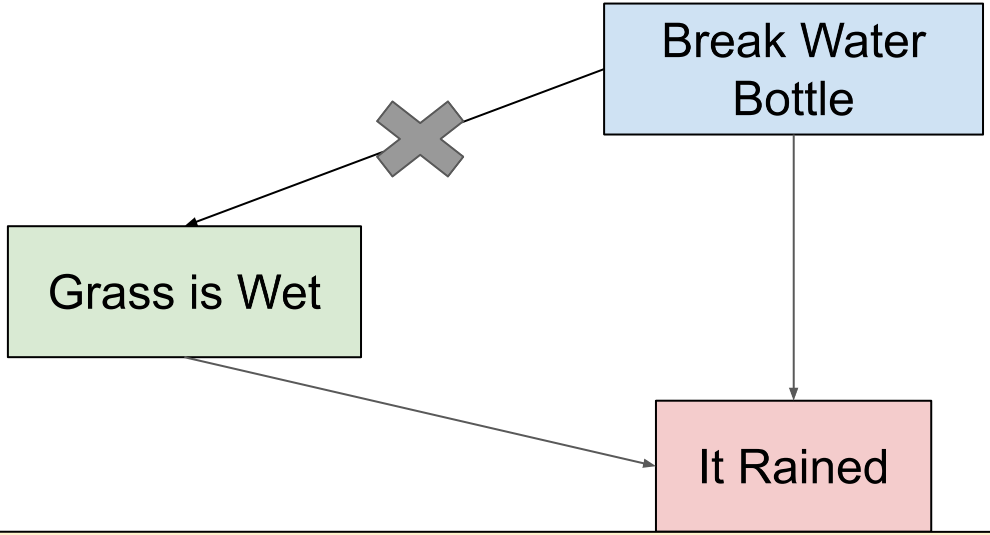
Rule #2: If we break this bottle, the glass will get wet

Causal Graph

A **causal graph** is a Bayesian network where the parents of each vertex are its direct causes.

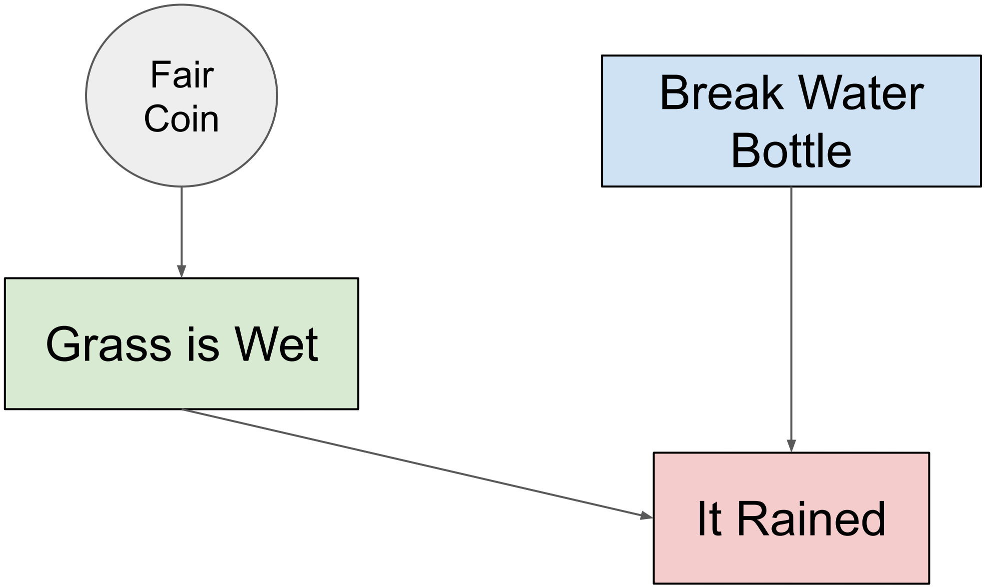


Causal Inference



When you observed Grass is Wet, the distribution of Break Water Bottle also changes due to the observation or added condition.
To examine causality of It Rained, we have to "**break the link**" somehow...

Intervention + Randomization



Think of it as a control experiment, where you have two groups of subject -- **Treatment** and **Control**. Your goal is to know the difference in experiment result.

What is different here?

Prediction: Predict Y after **observing** X = x

P(Y | X=x)

Causation: Predict Y after **setting** X = x.

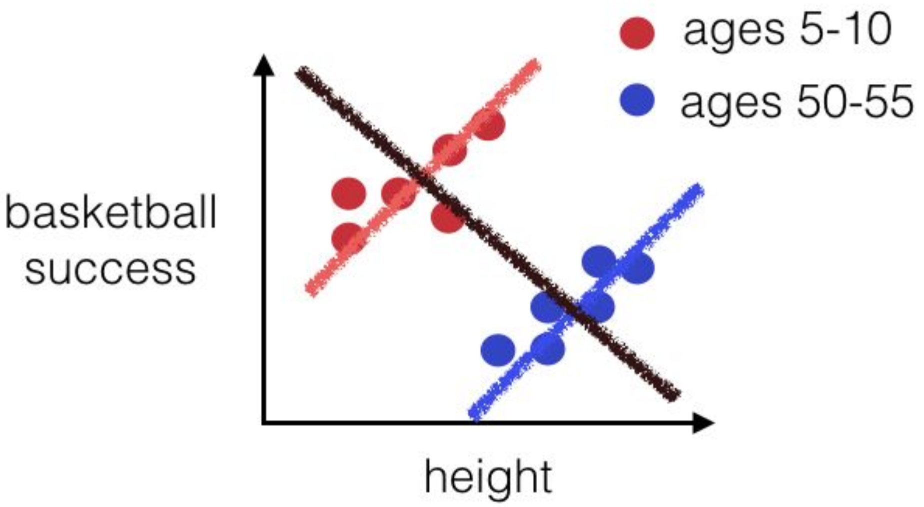
P(Y | set X=x)

Now we can write the phrase "*Correlation is not equal to Causation*" mathematically as …

$$P(Y \mid X{=}x) \neq P(Y \mid \text{set } X{=}x)$$

Example -- Simpson's Paradox

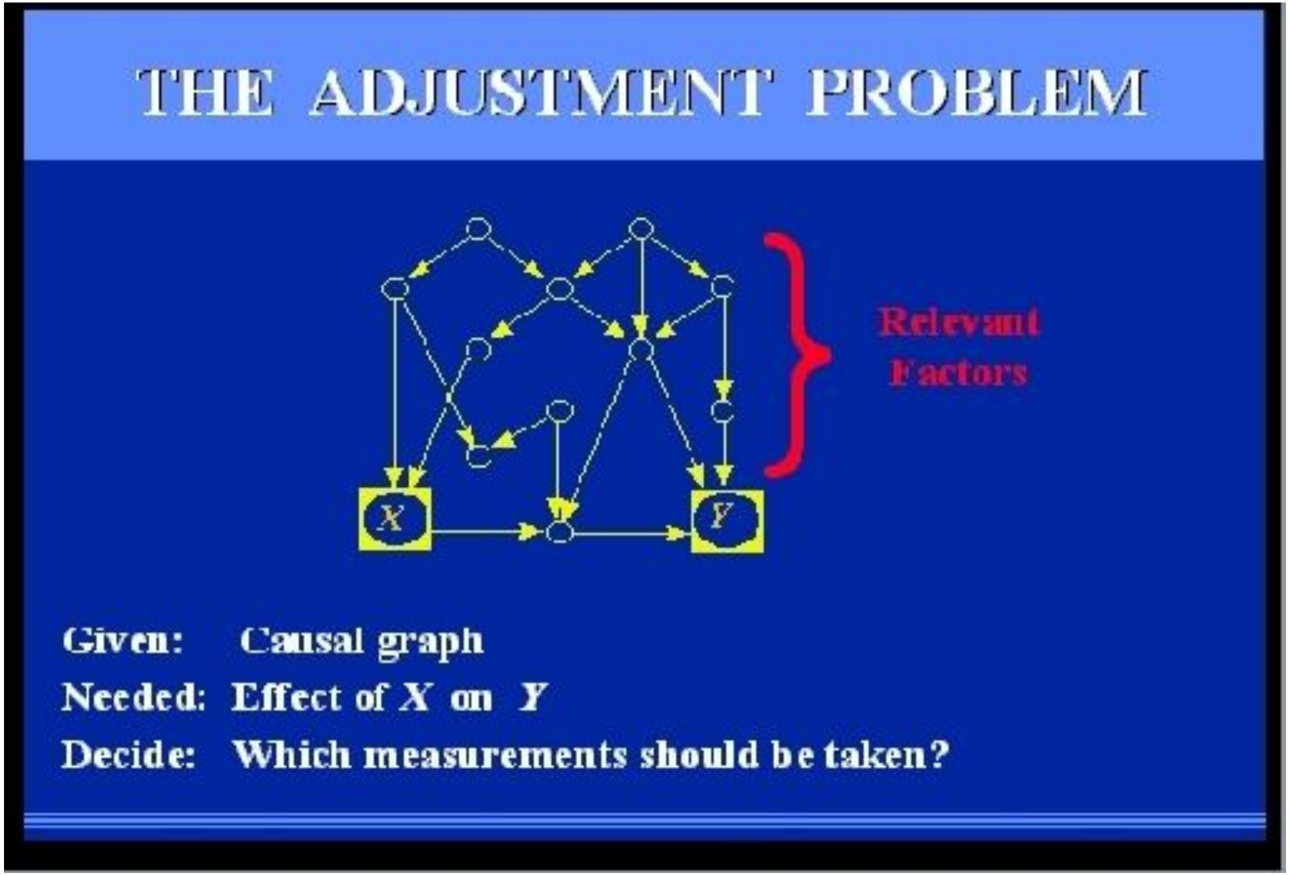
The statistical relationship between two variables may be REVERSED by including additional factors in the analysis



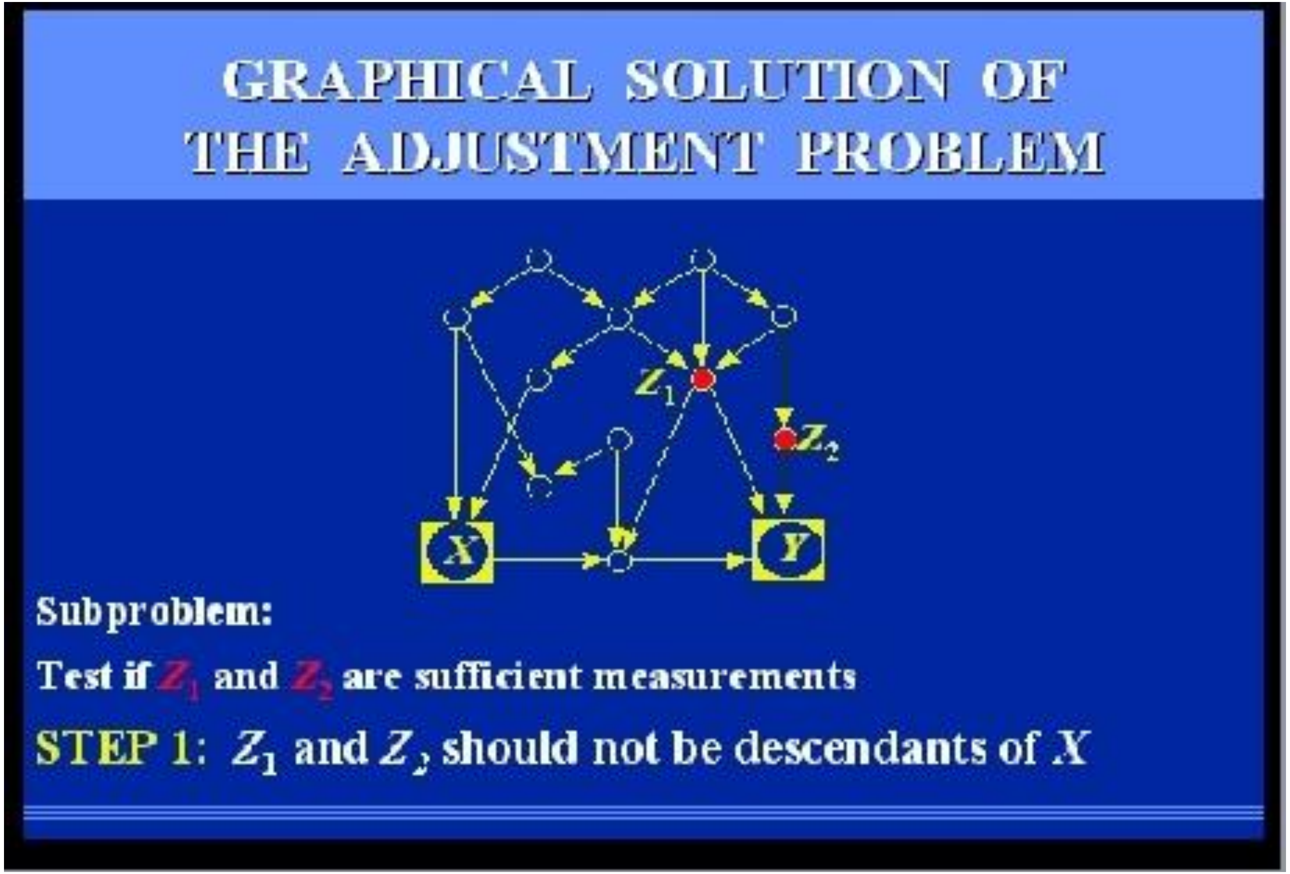
(Ironically, the paradox is first discovered by none other than Karl Pearson himself…)

https://towardsdatascience.com/simpsons-paradox-d2f4d8f08d42

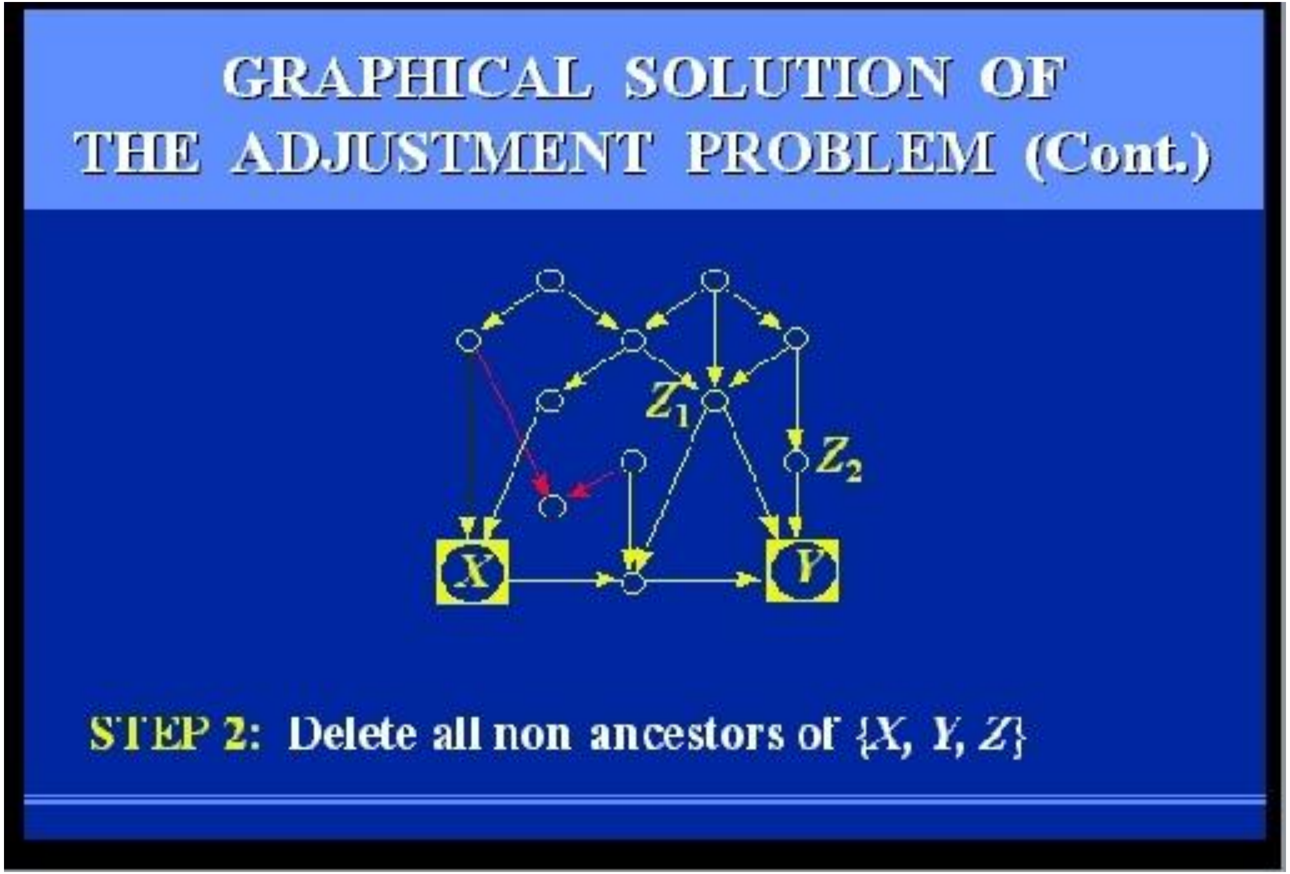
*If you have taken CIS520, This might remind you of d-separation.



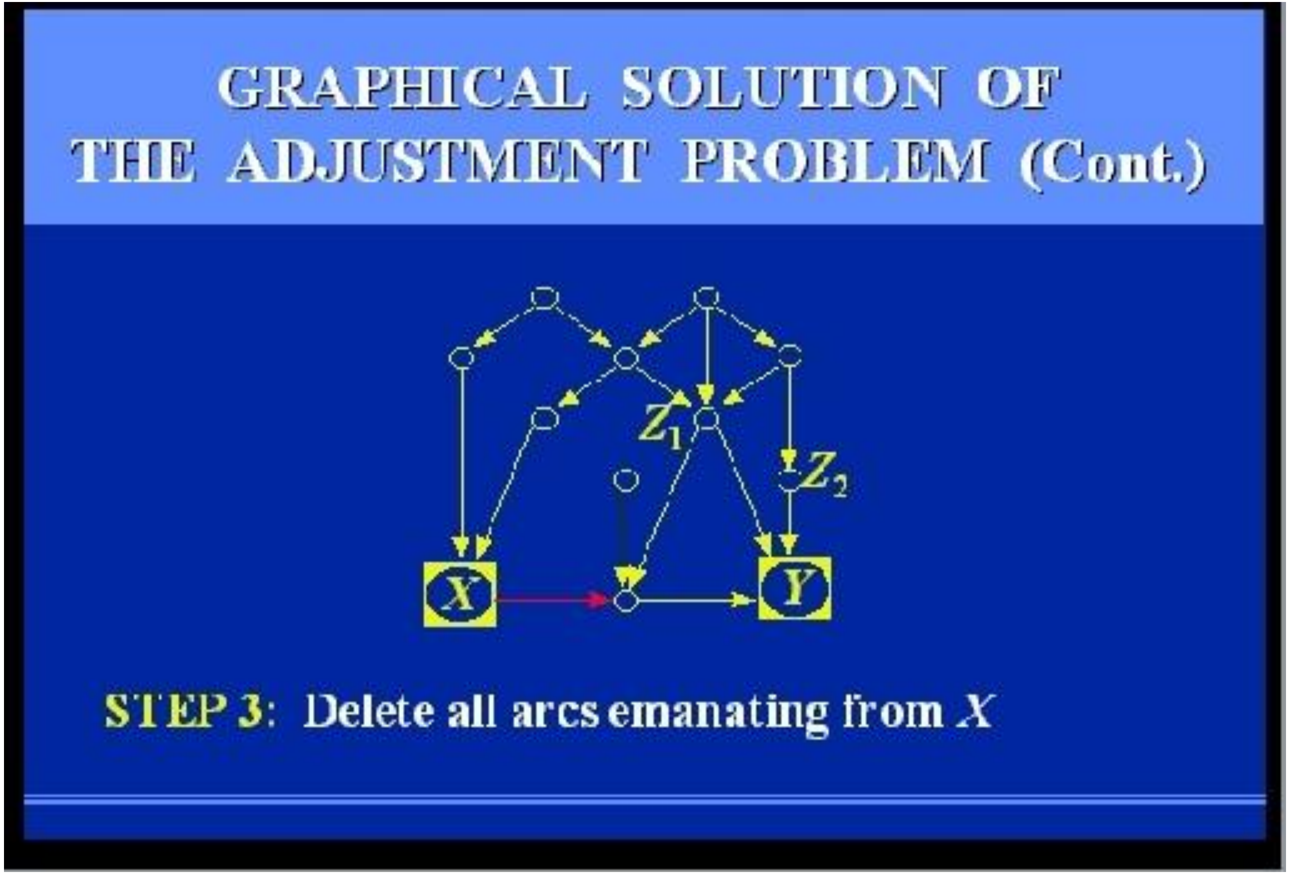
From Judea Pearl's original lecture slides...http://singapore.cs.ucla.edu/LECTURE/lecture_sec3.htm



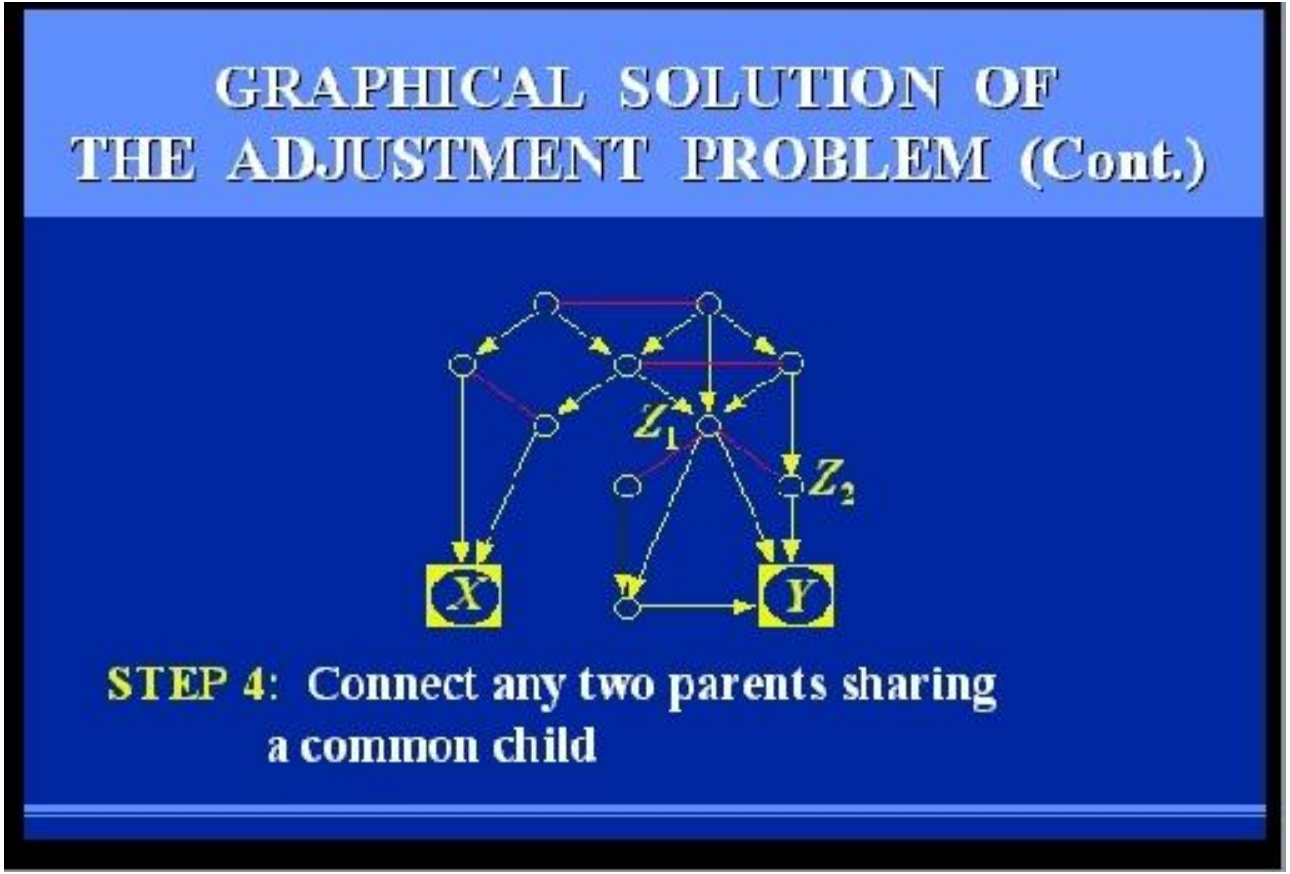
From Judea Pearl's original lecture slides...http://singapore.cs.ucla.edu/LECTURE/lecture_sec3.htm



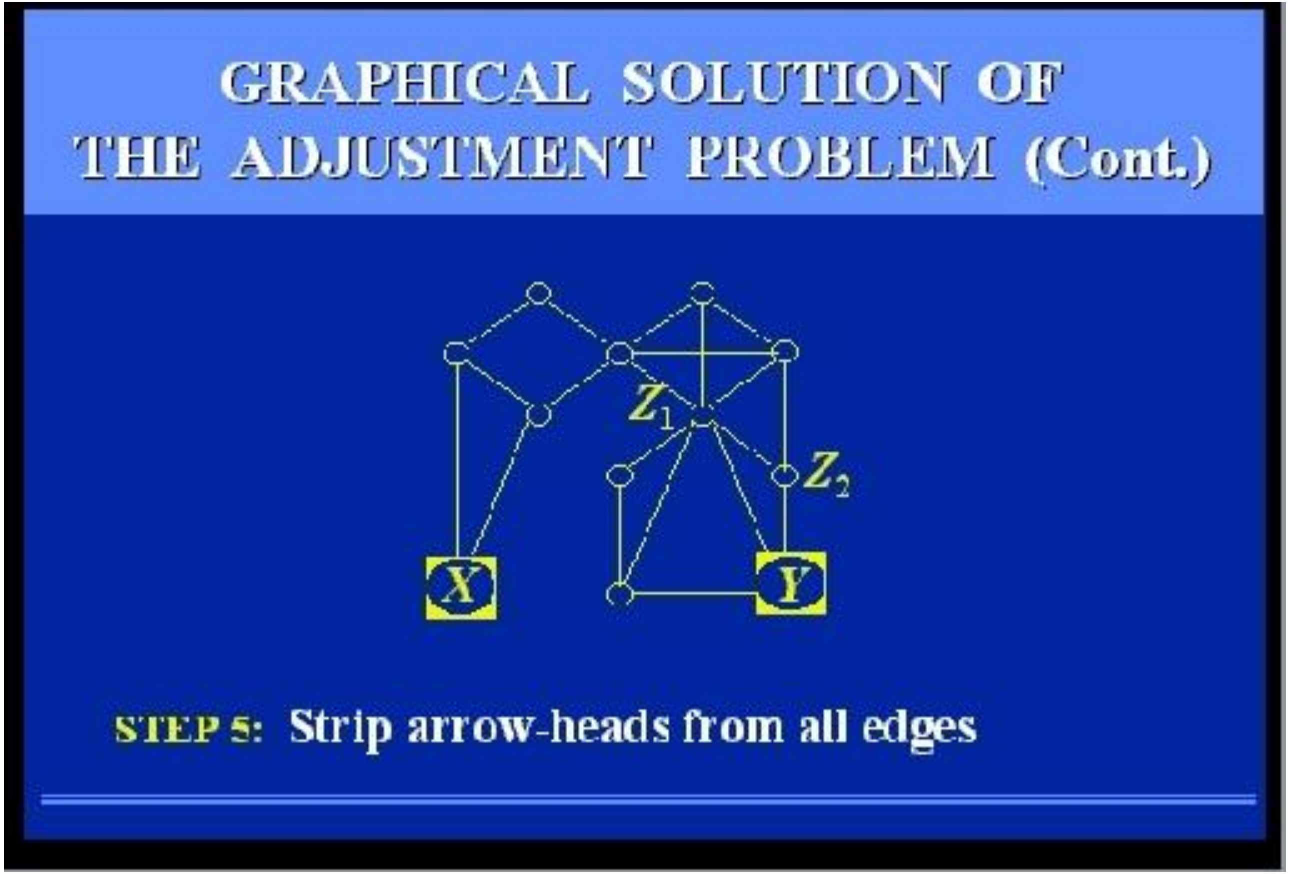
From Judea Pearl's original lecture slides...http://singapore.cs.ucla.edu/LECTURE/lecture_sec3.htm



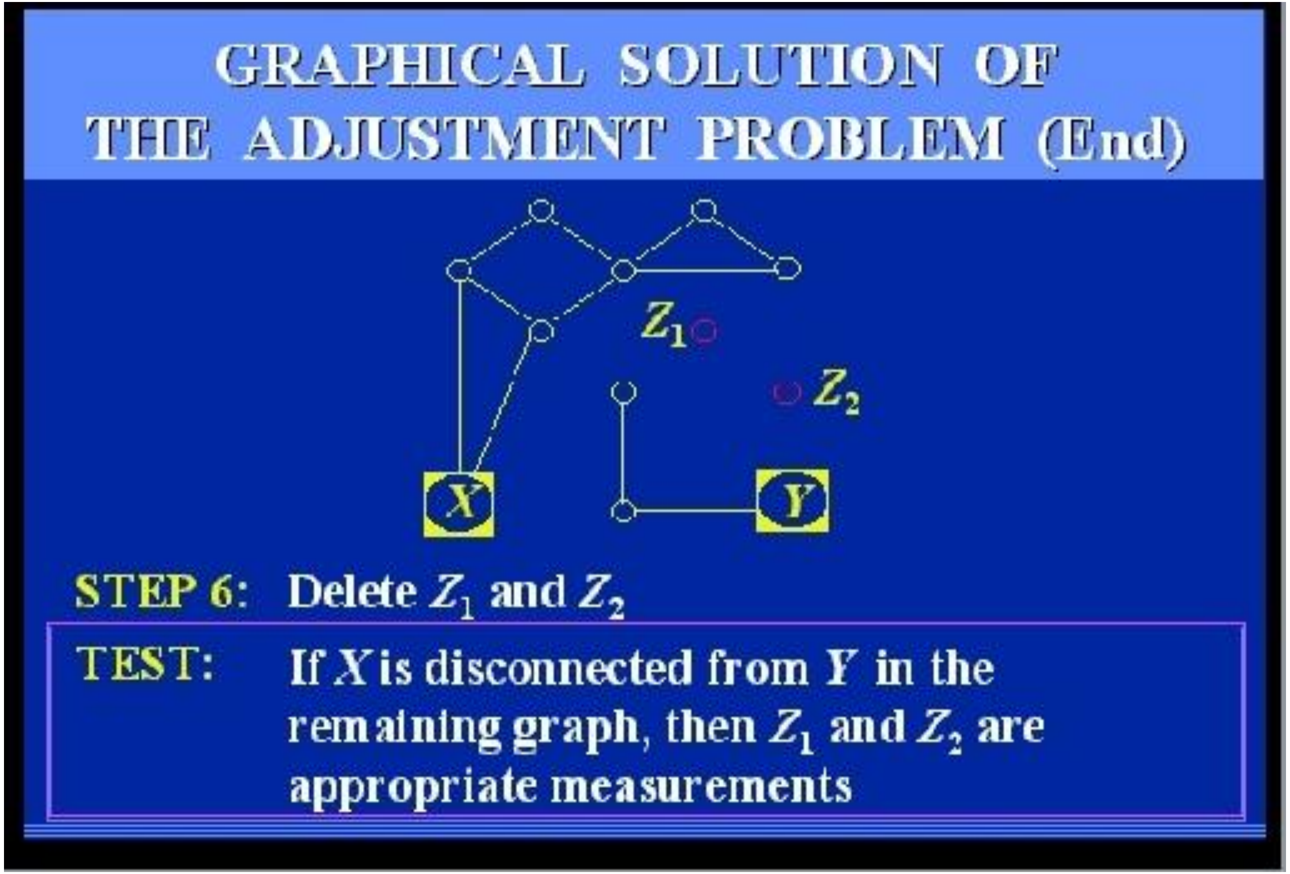
From Judea Pearl's original lecture slides...http://singapore.cs.ucla.edu/LECTURE/lecture_sec3.htm



From Judea Pearl’s original lecture slides...http://singapore.cs.ucla.edu/LECTURE/lecture_sec3.htm



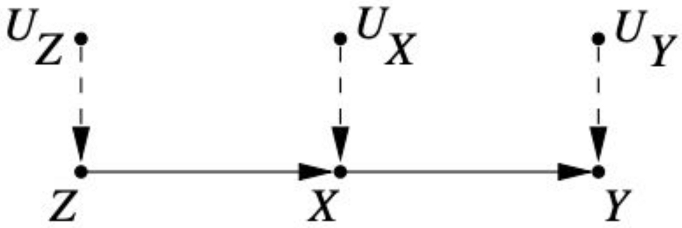
From Judea Pearl’s original lecture slides...http://singapore.cs.ucla.edu/LECTURE/lecture_sec3.htm



From Judea Pearl’s original lecture slides...http://singapore.cs.ucla.edu/LECTURE/lecture_sec3.htm

Structural Equation Modeling (SEM)

Given X, Y, Z and background factors (either observed or unobserved) U_x, U_y, U_z



The causal graph can be represented by

$$X = f_x(Z, U_x)$$
$$Y = f_y(X, U_y)$$
$$Z = f_z(U_z)$$

Intervention by setting value to a RV.
e.g. Fixing X=x

Outline

1. History of causality & why causality matters
2. How to approach causal inference
3. Applications; connection to Natural Language Understanding

Applications

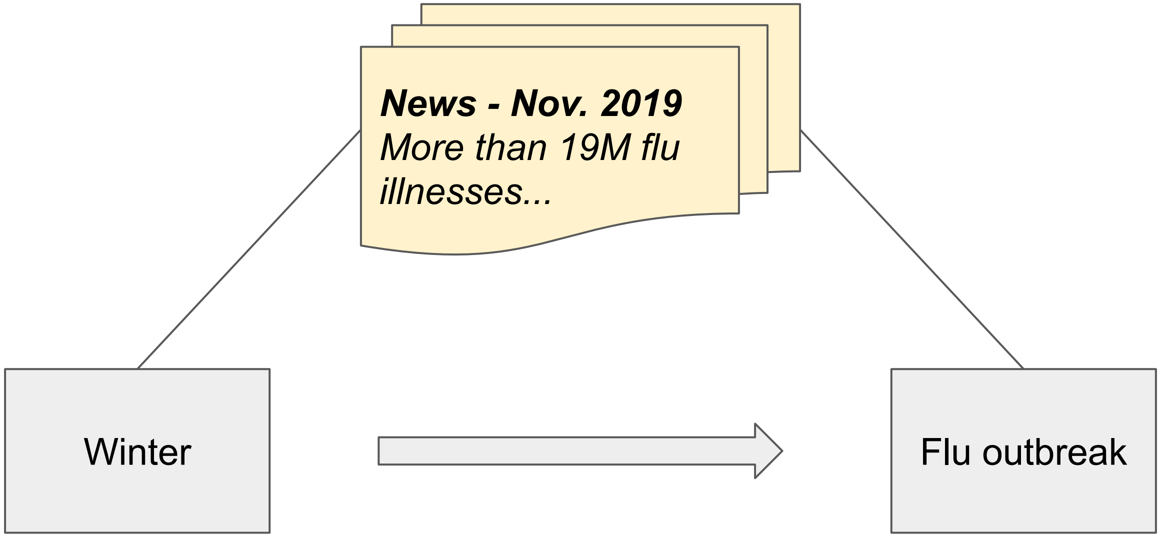
- Public Health, Medical Applications
  - e.g. Examining the cause of disease or the effect of medicine

- Social-Economical
  - “If we publish this policy, will we get the expected effect we want?”

- ...

In the World of Natural Language...

NL encodes information that either
(1) could serve as indicators for variables/events we want to study causality on.
(2) describe concepts (e.g. events) that could be "grounded" in the real world



EMNLP' 18

Framing and Agenda-setting in Russian News: a Computational Analysis of Intricate Political Strategies

Anjalie Field[♠] Doron Kliger[♦] Shuly Wintner[♦] Jennifer Pan[♥] Dan Jurafsky[♥] Yulia Tsvetkov[♠]
[♠]Carnegie Mellon University [♦]University of Haifa [♥]Stanford University
{anjalief, ytsvetko}@cs.cmu.edu,
{kliger@econ, shuly@cs}.haifa.ac.il, {jp1, jurafsky}@stanford.edu

They tested whether there's causal relation between
(1) The downturn of Russia's economy
(2) Increased state media coverage of U.S. (to deviate public attention)

In this work, NLP is used for building indicators for (2)
The natural of the problem is still social-economical.

RESEARCH Open Access

Extracting health-related causality from twitter messages using natural language processing

Son Doan*, Elly W. Yang, Sameer S. Tilak, Peter W. Li, Daniel S. Zisook and Manabu Torii

From The Sixth IEEE International Conference on Healthcare Informatics (ICHI 2018)
New York, NY, USA. 4-7 June 2018

They used pattern matching to get patient’s self-identified <Cause, Symptoms> pairs.

For example, “My insomnia was caused by stress.”

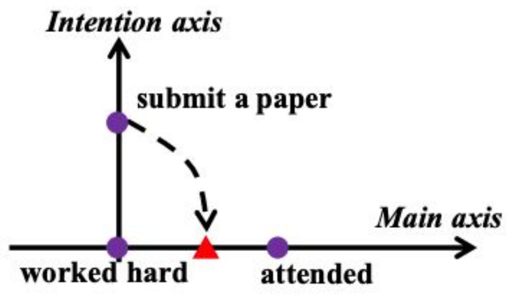
This can potentially be used to build/infer a rough causal graph

Temporality vs. Causality

- Temporality doesn’t imply causality, but one can be used to reason about the other...

- **Best of two worlds**: Natural Language
  - In NL, you have many context information to reason about both temporality and causality
  - **Temporality**:
  - **Causality**: “if” conjunctions; lexical patterns “<A> causes <b>”; Global statistics...
  - Maybe joint reasoning on the two?

“I worked hard to submit a paper . . . I attended the conference.”



“A Multi-Axis Annotation Scheme for Event Temporal Relations”, Ning et. al. 2018

Joint Reasoning for Temporal and Causal Relations

Qiang Ning,[1] **Zhili Feng,**[2] **Hao Wu,**[3] **Dan Roth**[1,3]
Department of Computer Science
[1]University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
[2]University of Wisconsin-Madison, Madison, WI 53706, USA
[3]University of Pennsylvania, Philadelphia, PA 19104, USA
qning2@illinois.edu, zfeng49@cs.wisc.edu, {haowu4,danroth}@seas.upenn.edu

Ex 1: Temporal relation dictated by causal relation.
More than 10 people ***(e1:died)*** on their way to the nearest hospital, police said. A suicide car bomb ***(e2:exploded)*** on Friday in the middle of a group of men playing volleyball in northwest Pakistan.
*Since **e2:exploded** is the reason of **e1:died**, the temporal relation is thus e2 being before e1.*

Ex 2: Causal relation dictated by temporal relation.
Mir-Hossein Moussavi ***(e3:raged)*** after government's efforts to ***(e4:stifle)*** protesters.
*Since **e3:raged** is temporally after **e4:stifle**, e4 should be the cause of e3.*

KORDING LAB

HOME ABOUT PUBLICA

Principal Investigator

Konrad Kording

http://kordinglab.com/

Interested in causal inference from a ML perspective;
A lot of cool work (and people) from their lab!

Questions?

What causes these questions? 👀