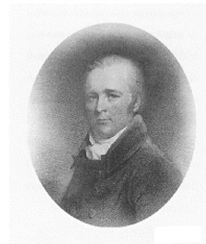
Graphical models and human causal learning

Tom Griffiths
UC Berkeley

Causal induction



Dr. James Currie

"The contagion spread rapidly and before its progress could be arrested, sixteen persons were affected of which two died. Of these sixteen, eight were under my care. On this occasion I used for the first time the affusion of cold water in the manner described by Dr. Wright. It was first tried in two cases ... The effects corresponded exactly with those mentioned by him to have occurred in his own case and thus encouraged the remedy was employed in five other cases. It was repeated daily, and of these seven patients, the whole recovered."

Currie (1798)
Medical Reports on, the Effects of Water, Cold and Warm, as a Remedy in Fevers and Febrile Diseases

	Treated	*Untreated*
Recovered	7	7
Died	0	2

"Does the treatment cause recovery?"

(Currie, 1798)

Causation from contingencies

	C present (c^+)	*C* absent (c^-)
E present (e^+)	*a*	*c*
E absent (e^-)	*b*	*d*

"Does *C* cause *E*?"
(rate on a scale from 0 to 100)

Two models of causal judgment

- Delta-P (Jenkins & Ward, 1965):

[illegible]

- Power PC (Cheng, 1997):

Power [illegible]

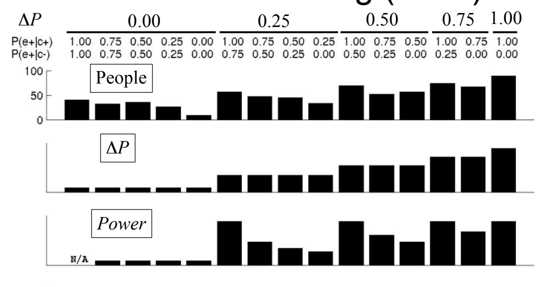
	Treated	Untreated
Recovered	7	7
Died	0	2

$P(e^+|c^+) = 7/7 = 1.00$
$P(e^+|c^-) = 7/9 = 0.78$

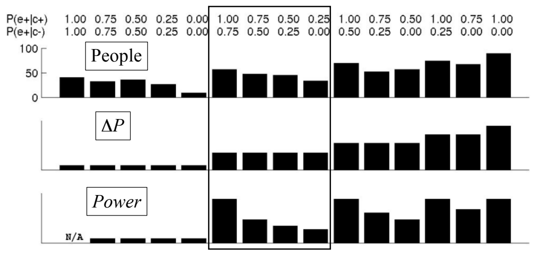
[illegible] $= 1.00 - 0.78 = 0.22$

Power = [illegible] $= 0.22/0.22 = 1.00$

Buehner and Cheng (1997)

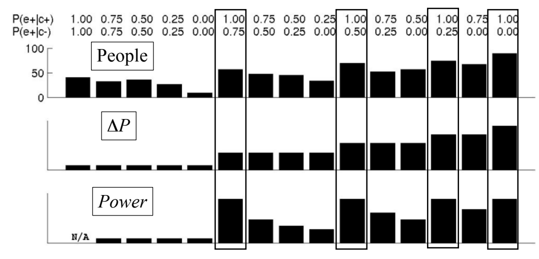


Buehner and Cheng (1997)



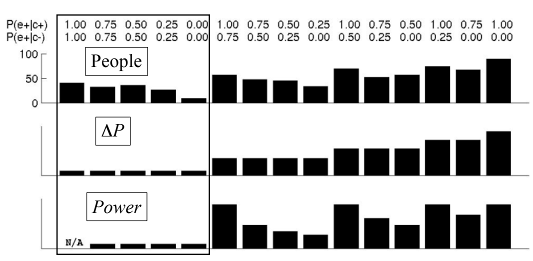
Constant ΔP, changing judgments

Buehner and Cheng (1997)



Constant causal power, changing judgments

Buehner and Cheng (1997)



$\Delta P = 0$, changing judgments

What is the computational problem?

- ΔP and causal power both seem to capture some part of causal induction, and miss something else…
- How can we formulate the problem of causal induction that people are trying to solve?
  - perhaps this way we can discover what's missing
- A first step: a language for talking about causal relationships…

Bayesian networks

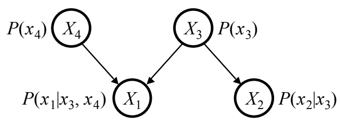
Nodes: variables
Links: dependency

Each node has a conditional probability distribution

Data: observations of $x_1, \ldots, x_4$

Four random variables:

X_1 coin toss produces heads
X_2 pencil levitates
X_3 friend has psychic powers
X_4 friend has two-headed coin



Causal Bayesian networks

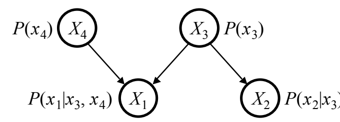
Nodes: variables
Links: causality

Each node has a conditional probability distribution

Data: observations of and interventions on $x_1, \ldots, x_4$

Four random variables:

X_1 coin toss produces heads
X_2 pencil levitates
X_3 friend has psychic powers
X_4 friend has two-headed coin



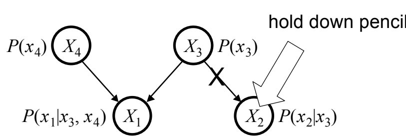
Interventions

Cut all incoming links for the node that we intervene on

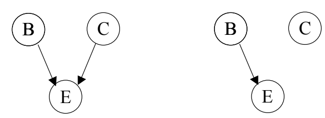
Compute probabilities with "mutilated" Bayes net

Four random variables:

X_1 coin toss produces heads
X_2 pencil levitates
X_3 friend has psychic powers
X_4 friend has two-headed coin

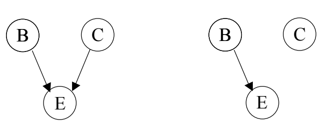


What is the computational problem?



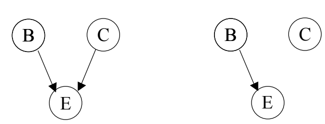
- **Strength:** how strong is a relationship?
- **Structure:** does a relationship exist?

What is the computational problem?



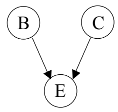
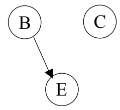
- **Strength:** how strong is a relationship?

What is the computational problem?



- **Strength:** how strong is a relationship?
  – requires defining nature of relationship

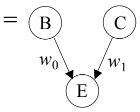
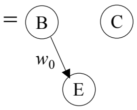
Parameterization

- Structures: $h_1 =$  $h_0 =$ 

- Parameterization: Generic (Bernoulli)

C	B	h_1: P(E = 1 | C, B)	h_0: P(E = 1| C, B)
0	0	p_{00}	p_0
1	0	p_{10}	p_0
0	1	p_{01}	p_1
1	1	p_{11}	p_1

Parameterization

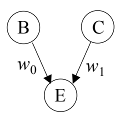
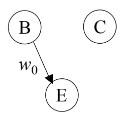
- Structures: $h_1 =$  $h_0 =$ 

w_0, w_1: strength parameters for B, C

- Parameterization: Linear

C	B	h_1: P(E = 1 | C, B)	h_0: P(E = 1| C, B)
0	0	0	0
1	0	w_1	0
0	1	w_0	w_0
1	1	$w_1 + w_0$	w_0

Parameterization

- Structures: $h_1 =$  $h_0 =$ 

w_0, w_1: strength parameters for B, C

- Parameterization: "Noisy-OR"

C	B	h_1: P(E = 1 | C, B)	h_0: P(E = 1| C, B)
0	0	0	0
1	0	w_1	0
0	1	w_0	w_0
1	1	$w_1 + w_0 - w_1 w_0$	w_0

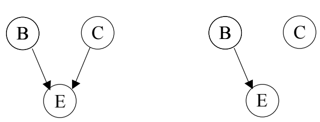
Parameter estimation

- Maximum likelihood estimation:

maximize [illegible]

- Bayesian methods: as for the coinflipping examples we discussed earlier…

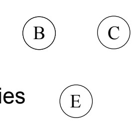
What is the computational problem?



- **Structure:** does a relationship exist?

Approaches to structure learning

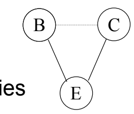
- Constraint-based
  - dependency from statistical tests
  - deduce structure from dependencies



(Pearl, 2000; Spirtes et al., 1993)

Approaches to structure learning

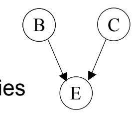
- Constraint-based:
  - dependency from statistical tests
  - deduce structure from dependencies



(Pearl, 2000; Spirtes et al., 1993)

Approaches to structure learning

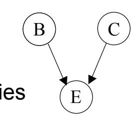
- Constraint-based:
  - dependency from statistical tests
  - deduce structure from dependencies



(Pearl, 2000; Spirtes et al., 1993)

Approaches to structure learning

- Constraint-based:
  - dependency from statistical tests
  - deduce structure from dependencies



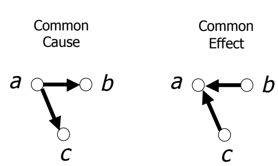
(Pearl, 2000; Spirtes et al., 1993)

Attempts to reduce *inductive* problem to *deductive* problem

Observationally equivalent

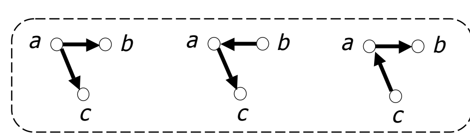
$$P(a)P(b|a) \quad = \quad P(b)P(a|b)$$

Two distinguishable networks

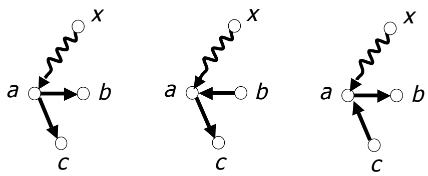


$$P(a)P(b|a)P(c|a) \neq P(a|b,c)P(c)P(b)$$

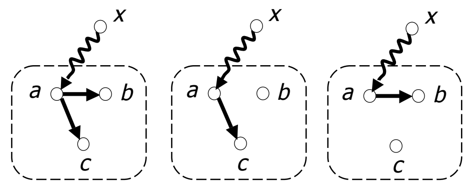
Indistinguishable from observation



Effect of intervention

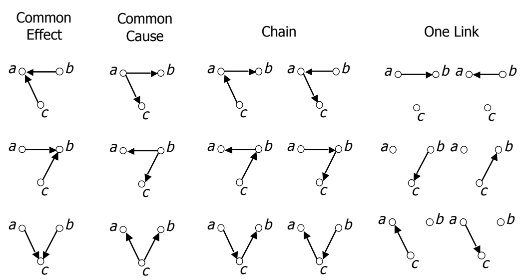


Effect of intervention

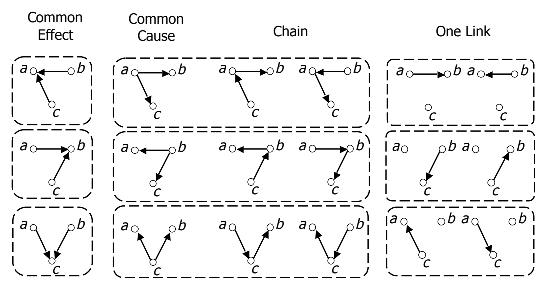


…networks become distinguishable

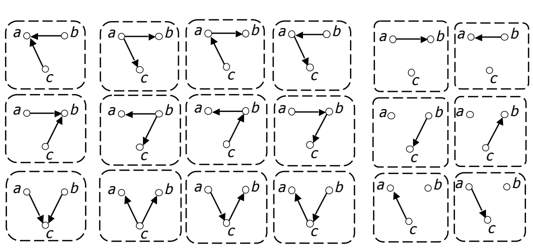
18 causal structures



Equivalence classes

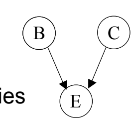


Consequences of intervention (on a well-chosen node)



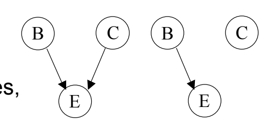
Approaches to structure learning

- Constraint-based:
  - dependency from statistical tests
  - deduce structure from dependencies



(Pearl, 2000; Spirtes et al., 1993)

- Bayesian:
  - compute posterior probability of structures, given observed data

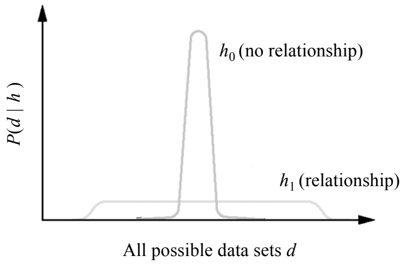


$P(h_1|\text{data})$ $P(h_0|\text{data})$

$P(h|d) \propto P(d|h)\ P(h)$

(Heckerman, 1998; Friedman, 1999)

Bayesian Occam's Razor

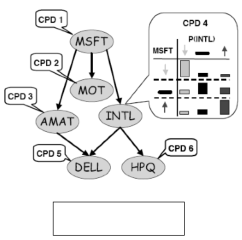


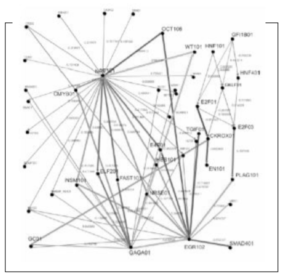
For any model h, [redacted]

Bayesian structure learning

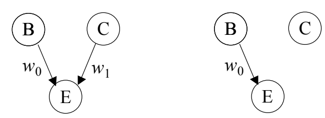
- Option 1: Search for the structure with the highest marginal likelihood
  – Structural EM algorithm (Friedman, 1997)
- Option 2: MCMC over causal structures
  – propose changes to links (implemented in BNT)
  – propose changes to order of variables
  (Friedman & Koller, 2003)

Bayesian structure learning





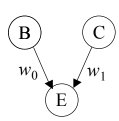
Causal structure vs. causal strength



- **Strength:** how strong is a relationship?
- **Structure:** does a relationship exist?

Causal strength

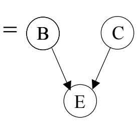
- Assume structure:



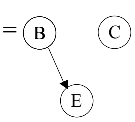
- ΔP and causal power are maximum likelihood estimates of the strength parameter w_1, under different parameterizations for $P(E|B,C)$:
  – linear → ΔP, Noisy-OR → causal power

Causal structure

- Hypotheses: h_1 =



h_0 =



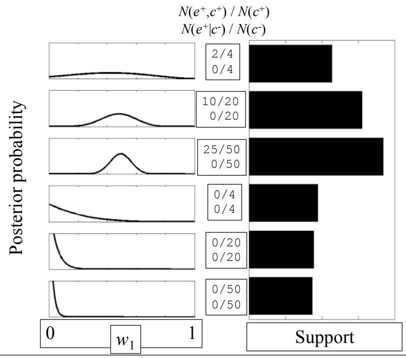
- Bayesian causal inference:

$$\text{support} = \frac{P(d|h_1)}{P(d|h_0)}$$

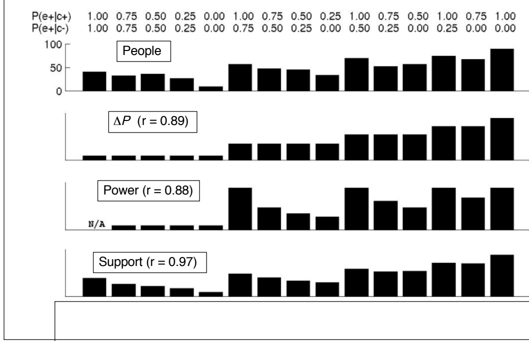
likelihood ratio (Bayes factor) gives evidence in favor of h_1

[redacted]

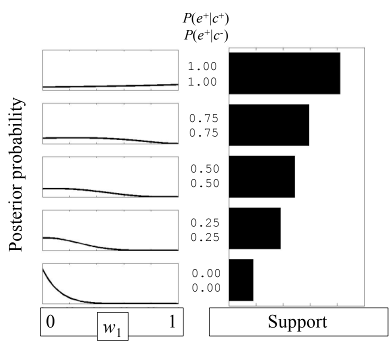
Induction as structure learning



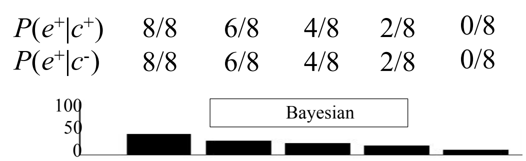
Buehner and Cheng (1997)



Judgments at $\Delta P = 0$



Generativity is essential



- Predictions result from "ceiling effect"
  - ceiling effects only matter if you believe a cause increases the probability of an effect
  - follows from use of Noisy-OR (after Cheng, 1997)

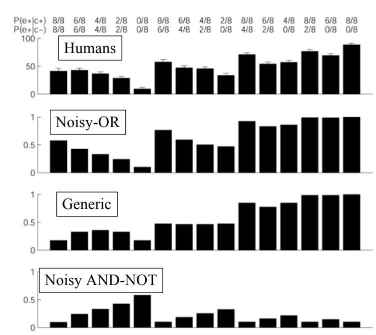
Generativity is essential

Noisy-OR
- causes increase probability of their effects

Generic
- probability differs across conditions

Noisy-AND-NOT
- causes decrease probability of their effects

Generativity is essential



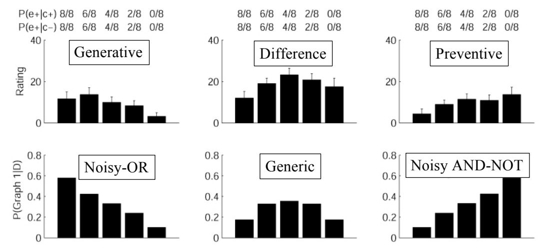
Manipulating functional form

Noisy-OR
- causes increase probability of their effects
- appropriate for *generative* causes

Generic
- probability differs across conditions
- appropriate for assessing *differences*

Noisy-AND-NOT
- causes decrease probability of their effects
- appropriate for *preventive* causes

Manipulating functional form



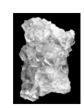
Causal induction from contingencies

- The simplest case of causal learning: a single cause-effect relationship and plentiful data
- Distinction between structure and strength yields different rational models of human causal learning
- Despite simplicity, exhibits complex effects of prior knowledge (in the assumed functional form)

(Griffiths & Tenenbaum, 2005)

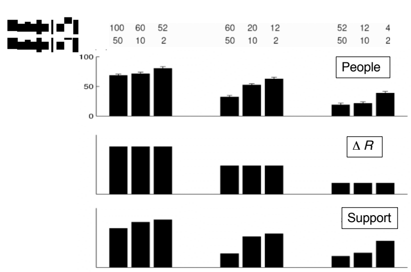
Causal induction with rates

- Causal Bayesian networks can be learned from data other than contingencies
- We can define structure learning and parameter estimation methods for rates

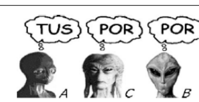


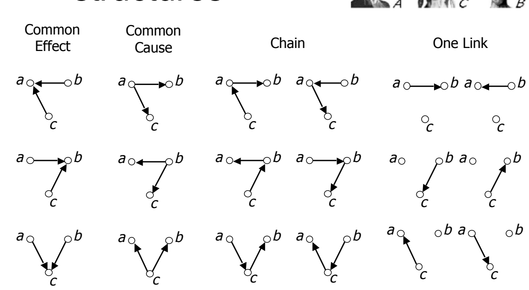
Does the electric field cause the mineral to emit particles?

Causal induction with rates



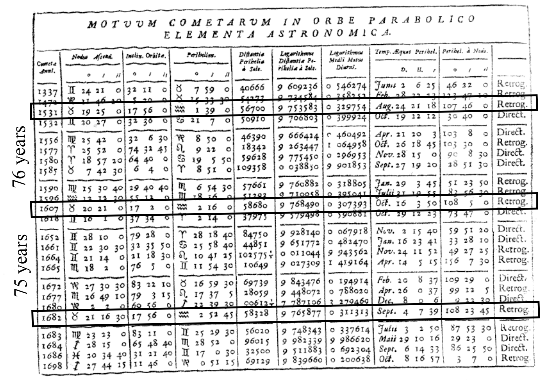
More complex structures





(Steyvers, Tenenbaum, Wagenmakers, & Blum, 2003)

Sir Edmond Halley



MOTUUM COMETARUM IN ORBE PARABOLICO ELEMENTA ASTRONOMICA.

Cometa Anni.	Nodus Ascend. ° ′ ″	Inclin. Orbitæ. ° ′ ″	Perihelium. ° ′ ″	Distantia Perihelia à Sole.	Logarithmus Distantiæ Perihelia à Sole.	Logarithmus Medii Motus Diurni.	Temp. Æquat. Perihel. D. H. ′	Perihel. à Nodo. ° ′ ″	
1337	♊ 24 21 0	32 11 0	♉ 7 59 0	40666	9 609236	0 546274	Junii 2 6 25	46 22 0	Retrog.
1472	♑ [illegible] 46 20	5 20 0	♉ 15 33 30	54273	9 734584	0 358252	Feb. 28 22 23	123 47 10	Retrog.
1531	♉ 19 25 0	17 56 0	♒ 1 39 0	56700	9 753583	0 329754	Aug. 24 21 18	107 46 0	Retrog.
1532	♊ 20 27 0	32 36 0	♋ 21 7 0	50910	9 706803	0 399924	Oct. 19 22 12	30 40 0	Direct.
1556	♍ 25 42 0	32 6 30	♑ 8 50 0	46390	9 666424	0 460492	Apr. 21 20 3	103 8 0	Direct.
1577	♈ 25 52 0	74 32 45	♌ 9 22 0	18342	9 263447	1 064958	Oct. 26 18 45	103 30 0	Retrog.
1580	♈ 18 57 20	64 40 0	♋ 19 5 50	59628	9 775450	0 296953	Nov. 28 15 0	90 8 30	Direct.
1585	♉ 7 42 30	6 4 0	♈ 8 51 0	109358	0 038850	9 901853	Sept. 27 19 20	28 51 30	Direct.
1590	♍ 15 30 40	29 40 40	♏ 6 54 30	57661	9 760882	0 318805	Jan. 29 3 45	51 23 50	Retrog.
1596	♒ 12 12 30	55 12 0	♏ 18 16 0	51293	9 710058	0 395041	Julii 31 19 55	83 56 30	Retrog.
1607	♉ 20 21 0	17 2 0	♒ 2 16 0	58680	9 768490	0 307393	Oct. 16 3 50	108 5 0	Retrog.
1618	♊ 16 1 0	37 34 0	♈ 2 14 0	37975	9 579498	0 590881	Oct. 29 12 23	73 47 0	Direct.
1652	♊ 28 10 0	79 28 0	♈ 28 18 40	84750	9 928140	0 067918	Nov. 2 15 40	59 51 20	Direct.
1661	♊ 22 30 30	32 35 50	♋ 25 58 40	44851	9 651772	0 482470	Jan. 16 23 41	33 28 10	Direct.
1664	♊ 21 14 0	21 18 30	♌ 10 41 25	102575½	0 011044	9 943562	Nov. 24 11 52	49 27 25	Retrog.
1665	♏ 18 2 0	76 5 0	♊ 11 54 30	10649	9 027309	1 419164	Apr. 14 5 15	156 7 30	Retrog.
1672	♑ 27 30 30	83 22 10	♉ 16 59 30	69739	9 843476	0 194914	Feb. 20 8 37	109 29 0	Direct.
1677	♏ 26 49 10	79 3 15	♌ 17 37 5	28059	9 448072	0 788020	Apr. 26 0 37	99 12 5	Retrog.
1680	♑ 2 2 0	60 56 0	♐ 22 39 30	00612½	7 787106	3 279469	Dec. 8 0 6	9 22 30	Direct.
1682	♉ 21 16 30	17 56 0	♒ 2 52 45	58328	9 765877	0 311313	Sept. 4 7 39	108 23 45	Retrog.
1683	♍ 23 23 0	83 11 0	♊ 25 29 30	56020	9 748343	0 337614	Julii 3 2 50	87 53 30	Retrog.
1684	♐ 28 15 0	65 48 40	♏ 28 52 0	96015	9 982339	9 986620	Maii 29 10 16	29 23 0	Direct.
1686	♓ 20 34 40	31 21 40	♊ 17 0 30	32500	9 511883	0 692304	Sept. 6 14 33	86 25 50	Direct.
1698	♐ 27 44 15	11 46 0	♑ 0 51 15	69129	9 839660	0 200638	Oct. 8 16 57	3 7 0	Retrog.

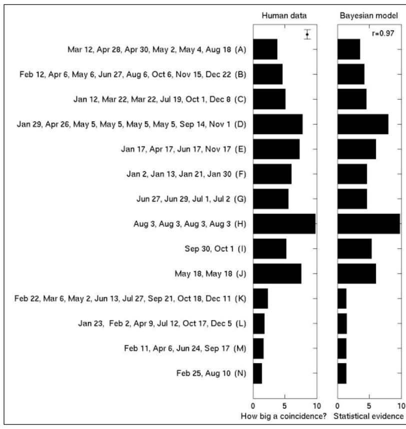
(Halley, 1752)

Detecting coincidences

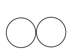
May 14, July 8, August 21, December 25

vs.

August 3, August 3, August 3, August 3



Michotte (1963)



Michotte (1963)

Summary

- Causal induction from contingency data is a well-studied problem, and one where learning structure and strength can provide insights
- Approaching causal induction as learning causal graphical models has the potential to explain many other phenomena:
  - learning from rates
  - learning complex causal structures
  - detecting coincidences
  - perceptual causality
  - inferences about physical causal systems
  - learning and reasoning about dynamic causal systems