Epistemic Democracy with defensible premises

by Franz Dietrich & Kai Spiekermann

New Developments in Judgement Aggregation and Voting Theory

Workshop
Freudenstadt, Schwarzwald
September 2011

The talk is based on our two working papers:
'Epistemic democracy with defensible premises', October 2010
'Independent Opinions?', October 2010

Background

- Popular defence of democracy in social epistemology: crowds can be ‘wise’, even if single people are ‘not so wise’
- The argument has been formalised in the classic Condorcet Jury Theorem

Background (cont.)

The Condorcet Jury Theorem's (CJT) remarkable history:

- goes back to Nicolas de Caritat, Marquis de Condorcet, 1785, french enlightenment period, just before the revolution
- first proved formally by Laplace in 1812
- then long forgotten
- finally rediscovered by Duncan Black (Black 1958, Grofman & Feld 1988)
- today very popular

The classical Condorcet Jury Theorem (informally)

Premise 1:	voters are 'independent'
Premise 2:	voters are 'competent'
Conclusion 1 (non-asymptotic):	larger groups perform better (plausible!)
Conclusion 2 (asymptotic):	huge groups are infallible (implausible!)

Epistemic Democracy

Note:

- This paper (and the CJT) pursue an epistemic goal
  - epistemic vs. procedural democracy

Institutional design (from an epistemic perspective)

Roughly, institutional design operates at two levels:

(1) designing the environment in which people form their opinions, ideally ensuring that opinions are

  (1a) independent

  (1b) competent (i.e., ‘often true’ in a suitable sense)

(2) designing the voting/aggregation rule used to merge the opinions once they are formed.

Goals for today

- The literature on the CJT focuses on (2), taking (1a) and (1b) for granted.
- This talk addresses both parts.

- **First part of talk (first paper):** A new jury theorem with more defensible 'independence' and 'competence' premises
- **Second part of talk (second paper):** What kind of causal environment promote independent opinions – and in what sense?

Outline

Part 1

- *The classical Condorcet Jury Theorem recapitulated*
- Common causes and the failure of Classical Independence
- The need to revise the classical competence assumption
- A new jury theorem
- The merits of deliberation

Part 2

[...]

Model ingredients

- Group of individuals: $i = 1, 2, ...$
  - e.g., group of jurors in a jury trial
- In total n individuals (n odd to avoid ties)
- Majority vote between two alternatives, labelled 0 and 1.
  - e.g., 'guilty' or 'not guilty' in a jury trial.
- One of the alternatives is factually 'correct', 'right' or 'better'.
  - called the *state (of the world)*
  - denoted x, generated by a random variable $\mathbf{x}$ (in bold!)
- R_i is the event that voter i votes correctly, i.e., for the state $\mathbf{x}$.

The classical jury theorem

Classical Independence Condition: Given any state of the world x in $\{0, 1\}$, the events of correct voting $R_1, R_2, \ldots$ are independent.

Classical Competence Condition: Given any state of the world x in $\{0, 1\}$, the probability of correct voting $\Pr(R_i|x)$ exceeds $\frac{1}{2}$ and does not depend on the voter i.

Condorcet Jury Theorem: Under these conditions, as the group size increases, the probability that a majority votes correctly (i) increases and (ii) converges to one.

What went wrong?

- The independence premise is unrealistic!
- Strategy: revise the premises, obtain a more realistic asymptotic conclusion.

Brief preview at the jury theorem

	classical theorem	new theorem
Premise 1:	‘independence’	‘conditional independence’
Premise 2:	‘competence’	‘competence more often than incompetence’
Conclusion 1:	‘the larger the better’	‘the larger the better’
Conclusion 2:	‘huge groups **infallible**’	‘huge groups **fallible**’

Outline

Part 1

- The classical Condorcet Jury Theorem recapitulated
- *Common causes and the failure of Classical Independence*
- The need to revise the classical competence assumption
- A new jury theorem
- The merits of deliberation

Part 2

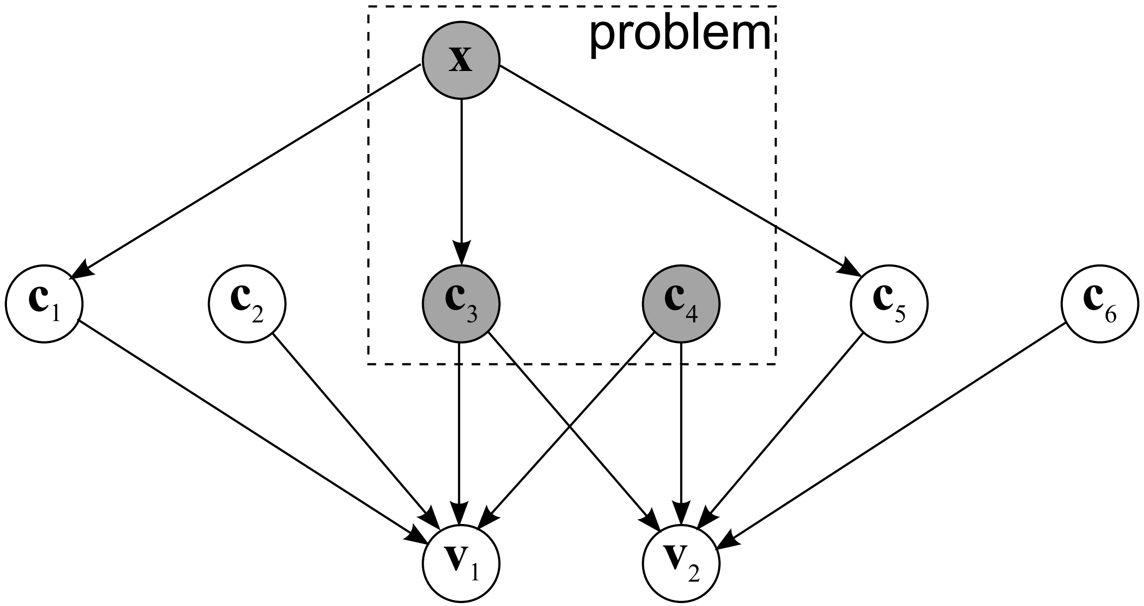
[...]

Common causes

- The standard critique of Classical Independence: it may fail since voters can influence each other.
- Our critique: it may fail even if voters are isolated from each other.
  - Because of common causes
  - We draw on the well-established theory of causal networks and Reichenbach's influential *common cause principle*.
- Common causes for economic advisors in 2007 before the economic crisis broke out:
  - shared theoretical assumptions about the economy.
  - shared evidence (e.g., apparently safe balance sheets of banks)
  - shared exposure to room temperature
- Common causes push all into the same (possibly wrong) direction!

A causal network to illustrate common causes

- A causal network is a directed acyclic graph representing causal effects between variables/phenomena.
- This causal network contains the votes (only the first two votes are shown), the state $\mathbf{x}$, and other causes of votes $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_6$
- Some causes are *common* (see box), others *private*.
- Some are *evidential* (i.e., related to $\mathbf{x}$), others *non-evidential*.



A new independence condition

- We 'conditionalise away' all dependence between voters by conditionalising
  - not just on the state of the world (as in the classical model)
  - but on all circumstances, conceptualized as the common causes of votes.
- So, we conditionalise on what we call the group's *decision problem* (following Dietrich 2008).
- Formally, the decision problem is a random variable $\boldsymbol{\pi}$ taking values in some (arbitrarily complex) space.

New Independence Condition: Given the decision problem $\boldsymbol{\pi}$, the correct voting events $R_1, R_2, \ldots$ are independent.

Outline

Part 1

- The classical Condorcet Jury Theorem recapitulated
- Common causes and the failure of Classical Independence
- *The need to revise the classical competence assumption*
- A new jury theorem
- The merits of deliberation

Part 2

[...]

The need to revise the competence assumption

- Classical Competence: For each state x, $\mathsf{Pr}(R_i|x)$ exceeds $1/2$ (and is the same for each voter i).
- Plausible!
- But one can't fruitfully combine this *state*-conditional notion of competence with our *problem*-conditional notion of independence (rather than with the unrealistic state-conditional independence).
- This wouldn't deliver the desired conclusion!
  - Recall we look for plausible premises implying that larger groups perform better, i.e., that 'crowds are wise'.

Example of larger groups performing worse

Let our economists face only two types of economic problems:

- *easy* problems, on which each expert is right with 99% probability.
- *difficult* problems, on which each expert is right with 49% probability
  - Presumably, the problem of predicting whether the 2008 banking crisis would trigger a major recession in 2009 was difficult.

Example of larger groups performing worse

- Formally: $\text{Pr}(R_i|\pi) = \begin{cases} 0.99 & \text{for every easy problem } \pi \\ 0.49 & \text{for every difficult problem } \pi \end{cases}$
- Suppose each problem type occurs with probability $\frac{1}{2}$.
- Each voter i is *unconditionally* competent:

$$\text{Pr}(R_i) = \frac{1}{2} \times 0.99 + \frac{1}{2} \times 0.49 = 0.74 > \frac{1}{2}.$$

- Each voter i is also *state*-conditionally competent:[1]

$$\text{Pr}(R_i|x) > \frac{1}{2} \text{ for each state } x$$

[1] Under mild extra conditions (essentially, there shouldn't be a too high correlation between problem type and state).

Example of larger groups performing worse

- So, Classical Competence holds.
- Yet large groups are much worse:

$$\text{Pr}(M_n) = \frac{1}{2} \times \text{Pr}(M_n|\boldsymbol{\pi} \text{ is easy}) + \frac{1}{2} \times \text{Pr}(M_n|\boldsymbol{\pi} \text{ is difficult})$$

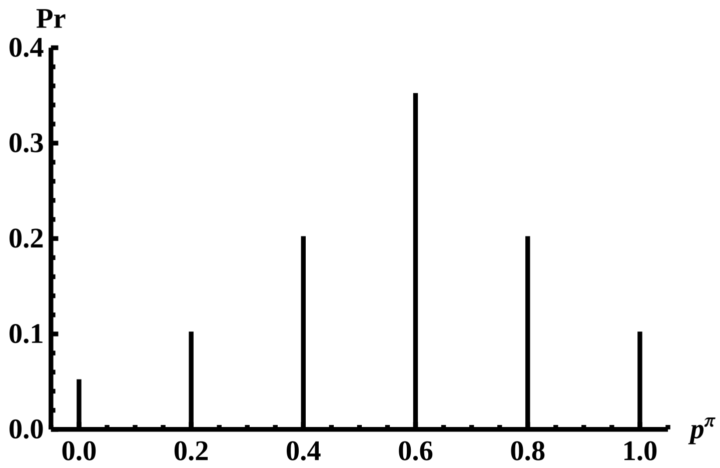
(... assuming New Independence)

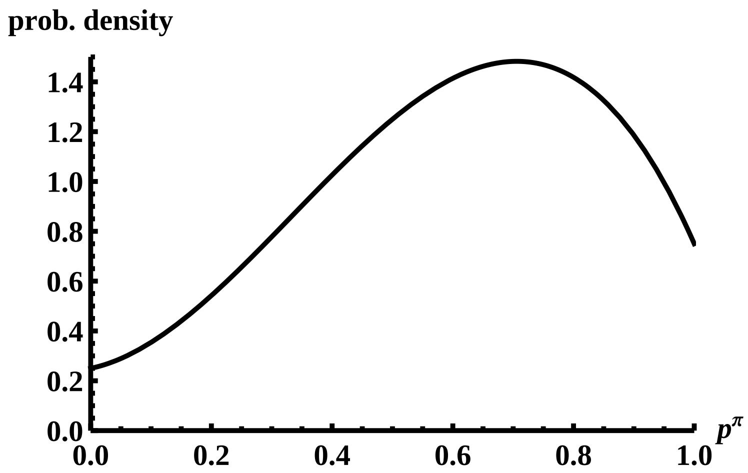
$$\approx \begin{cases} \frac{1}{2} \times 1 + \frac{1}{2} \times \frac{1}{2} = \frac{3}{4} & \text{for small } n \\ \frac{1}{2} \times 1 + \frac{1}{2} \times 0 = \frac{1}{2} & \text{for very large } n. \end{cases}$$

A problem-specific notion of competence (1)

- A voter i's *problem-specific competence* is the probability of voting correctly conditional on the problem, $p_i^\pi = \mathrm{Pr}(R_i|\boldsymbol{\pi})$.
  - It is high if the problem is 'easy' and low if the problem is 'difficult'.
- Since the problem is a random variable, so is propblem-specific competence.
- So, problem-specific competence has a distribution.

Examples:





The new competence assumption

New Competence

- (informally) Competence is more often high than low
  –> so the distribution of problem-specific competence is right-skewed
- (formally) Problem-specific competence p_i^π is more likely to be high than low – that is, is $\frac{1}{2} + \epsilon$ with at least as much probability as it is $\frac{1}{2} - \epsilon$, for all $\epsilon > 0$ – and is the same for all voters i – that is, $p_i^\pi \equiv p^\pi$.[2]

[2]The clause 'that is, is $\frac{1}{2} + \epsilon$ ... for all $\epsilon > 0$' is stated for the case that p^π has a discrete distribution (as in figure 4 but not as in figure 5). The general statement is as follows: 'that is, belongs to $\left[\frac{1}{2} + \epsilon, \frac{1}{2} + \epsilon'\right]$ with at least as much probability as it belongs to $\left[\frac{1}{2} - \epsilon', \frac{1}{2} - \epsilon\right]$, for all $\epsilon' \geq \epsilon > 0$'. The reason is, roughly, that a continuous distribution is given not by the probabilities of single points (these are all zero) but by the probabilities of intervals.

Outline

Part 1

- The classical Condorcet Jury Theorem recapitulated
- Common causes and the failure of Classical Independence
- The need to revise the classical competence assumption
- *A new jury theorem*
- The merits of deliberation

Part 2

[...]

The new jury theorem (1)

New Independence Condition (recall): Given the decision problem $\boldsymbol{\pi}$, the correct voting events $R_1, R_2, \ldots$ are independent.

New Competence Condition (recall): Problem-specific competence is more likely to be 'high' than 'low' and does not depend on the voter.

New Jury Theorem. Under the new conditions, as the group size increases, the probability that a majority votes correctly (i) increases, and (ii) converges to a value below one if not all problems are 'easy', i.e., if $\mathsf{Pr}\left(p^{\boldsymbol{\pi}} > \frac{1}{2}\right) \neq 1$ (and to one otherwise).

The exact limiting group performance

As the proof shows:

- The value to which the probability converges is $\mathsf{Pr}\left(p^{\pi} > \frac{1}{2}\right) + \frac{1}{2}\mathsf{Pr}\left(p^{\pi} = \frac{1}{2}\right)$, the probability that the problem is easy plus half of the probability that it is on the boundary between easy and difficult.

There are counterexamples to the premises!

- Our earlier example violates New Competence.
- Because competence is *less* likely to be 0.51 than 0.49:

$$\mathsf{Pr}(p_i^{\pi} = 0.51) = 0 < \frac{1}{2} = \mathsf{Pr}(p_i^{\pi} = 0.49).$$

- That's why larger groups could perform worse here!

Recovering the classical CJT as a special case

- Our model is very flexible since the problem variable $\boldsymbol{\pi}$ can be specified arbitrarily according to one's needs.
- A very simple specification of $\boldsymbol{\pi}$ yields the classical CJT.
  - This specification goes against the spirit of our analysis, but is mathematically meaningful.
- Formally, if we choose $\boldsymbol{\pi}$ be identical with the state of the world, then:
  - New Independence $\Leftrightarrow$ Classical Independence
  - New Competence $\Leftarrow$ Classical Competence[3]
  - New conclusions $\Leftrightarrow$ classical conclusions (when we have Classical Competence, i.e., when $\Pr\left(p^{\mathbf{x}} > \frac{1}{2}\right) = 1$).
- In fact, this strengthens the CJT by using logically weaker premises.

[3]Classical Competence is the special case of New Competence in which the distribution of problem/state-specific competence is fully concentrated on the right-half interval $(1/2, 1]$.

Recovering another variant of the classical CJT

- There is not 'one' classical CJT but different related variants.
- To recover the simplest of all variants, suppose the problem $\boldsymbol{\pi}$ takes only *one* value
- ... so that conditionalizing on $\boldsymbol{\pi}$ is as much as not conditionalizing at all!
- Our two premises then reduce to the following premises:
  - the events $R_1, R_2, \ldots$ are (unconditionally) independent;
  - unconditional competence, $\mathsf{Pr}(R_i)$, is at least $\frac{1}{2}$ and is the same across voters.
- Our conclusions are equivalent to the classical conclusions: majority competence increases in group size and converges to one (or to $1/2$ if $\mathsf{Pr}(R_i) = 1/2$).

Outline

Part 1

- The classical Condorcet Jury Theorem recapitulated
- Common causes and the failure of Classical Independence
- The need to revise the classical competence assumption
- A new jury theorem
- *The merits of deliberation*

Part 2

[...]

The merits of deliberation

- Education and deliberation rehabilitated:
  - The classical framework makes them appear unnecessary (and partly counter-productive as deliberation threatens Classical Independence).
  - In our framework, they can improve group performance by
    - making more problems 'easy'
    - i.e., right-shifting the distribution of problem-specific competence
    - hence, increasing the limiting group performance

Outline

Part 1

[...]

Part 2

Goal: Causal foundations

Four types of probabilistic independence

Theorem

Background (cont.)

Thinking about opinion independence reveals a systematic difference between individual and social epistemology:

- individual epistemology recommends dependent opinions in the form of positive correlation with experts,
- social epistemology recommends independent opinions (and other things)
  - tries to avoid pathologies of social opinion formation, such as informational cascades, biases and the influence of opinion leaders

Background

- But what means **independence** of opinions?
  - probabilistic vs. causal independence
- Goal:
  - distinguish 4 notions of probabilistic opinion independence
  - identify their **causal foundations**, i.e., the causal environments that deliver each of them.
- Two of these notions will be the above 'Classical' and 'New' Independence.

Arbitrary opinions

- So far, opinions (votes) were binary; e.g.:
  - Is the defendant in a court trial guilty or innocent?
  - Will global warming continue or not?
- But from now on, opinions are arbitrary, e.g.:
  - sets of believed propositions (belief sets or judgment sets),
  - numerical estimates (say, of the height of a mountain),
  - degrees of belief (probabilities)
  - ...
- Formally, there is an arbitrary set O of possible opinions.

An arbitrary state

- Exactly one opinion in O is 'correct' ('right' or 'best').
- Which opinion is correct is determined by an external fact, called the *state (of the world)* and denoted x.
  - e.g., the opinion 'the defendant is guilty' is true just in case the defendant has committed the crime in question.

An arbitrary state (cont)

- One might formally identify the state with the opinion thereby made correct
  - So that states and opinions would be the same kind of object.
  - Such an identification is implicitly made in the literature.
- What's formally convenient isn't always conceptually plausible!
- So, we allow to keep opinions and facts ('truth-makers') apart.

- Formally, let $\mathbf{x}$, $\mathbf{o}_1$, $\mathbf{o}_2$, ... be random variables generating the state of the world, 1's opinion, 2's opinion, ...

Outline

Part 1

[...]

Part 2

Goal: Causal foundations

Four types of probabilistic independence

Theorem

Independence: first version

Simplest independence notion one may come up with:

Unconditional Independence (UI). The opinions $\mathbf{o}_1$, $\mathbf{o}_2$, ... are unconditionally independent.

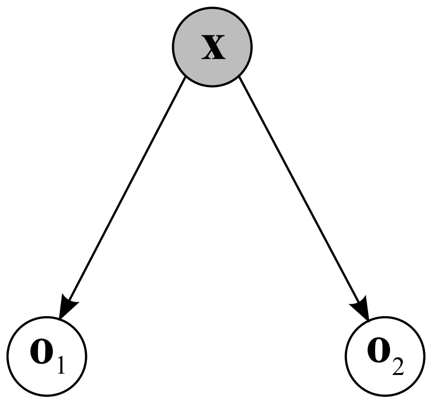
- Objection to UI: since each opinion is (hopefully) correlated to the truth, how could they be mutually independent?
- Surprisingly, UI is less absurd than one might have thought:
  - Interesting causal setups support UI.
  - Details later. For now, just notice that the above objection is not fully convincing, since correlation isn't transitive!

Independence: second version

In response to the above objection against UI, let's conditionalise on the state:

State-Conditional Independence (SI). The opinions $\mathbf{o}_1$, $\mathbf{o}_2$, ... are independent conditional on the state $\mathbf{x}$.

- SI underlies Condorcet's classical jury theorem.
- SI can be motivated by a causal network (all our plots of causal networks show only the first two opinions):



Independence: second version (cont.)

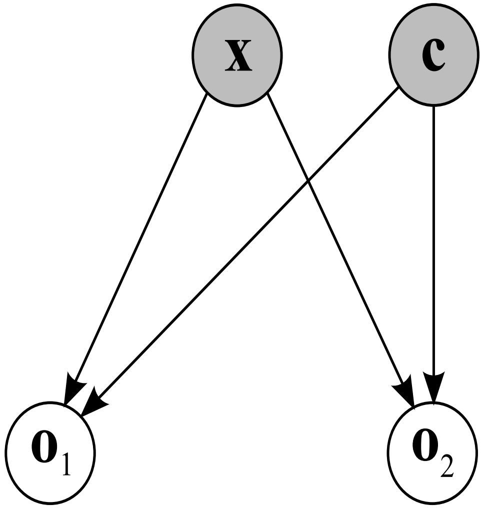
- Assuming the causal environment of Fig. 1, SI *must* hold.
- Informal reason: conditionalising on $\mathbf{x}$ blocks the information flow between opinions.
- Formal reason: theory of Bayesian networks, and more specifically, Reichenbach's (1956) famous *Common Cause Principle*:

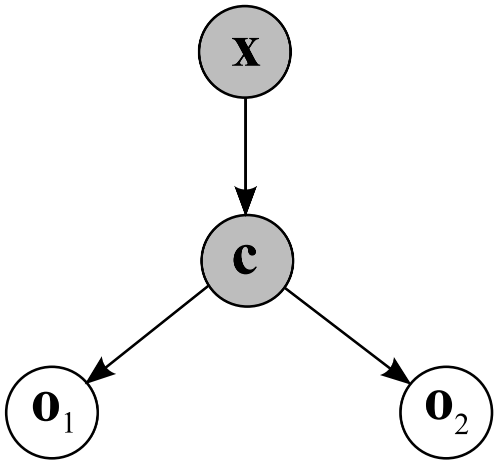
Common Cause Principle (stated informally). Phenomena which do not causally affect each other

- can only be probabilistically dependent if they have (one or more) common causes,
- are probabilistically independent conditional on these common causes.

Independence: second version (cont.)

- However, the Common Cause Principle can be turned against SI once we consider other causal networks with other common causes than $\mathbf{x}$.
- Examples:





Independence: third version

- The possibility of multiple common causes suggests adopting an independence condition that conditionalises on ('controls for') all common causes.
- To define it, we need to extend the framework slightly.
- In addition to the random variables $\mathbf{x}$, $\mathbf{o}_1$, $\mathbf{o}_2$, ..., we consider an arbitrary number of additional random variables (phenomena).
- Consider a causal network (i.e., a *directed acyclic graph*) over all the variables.
- A variable $\mathbf{a}$ is a *cause* of another $\mathbf{b}$ (and $\mathbf{b}$ an *effect* of $\mathbf{a}$) if there is a directed path from $\mathbf{a}$ to $\mathbf{b}$.

Independence: third version

- $\mathbf{a}$ is a *common cause* (*effect*) of some variables if it is a cause (effect) of each of them.
- $\mathbf{a}$ is a *private cause* of an opinion it it's a cause of this opinion and of no other opinion.
- Let $\boldsymbol{\chi}$ (Greek 'chi') be the family of all common causes of votes.

Common-Cause-Conditional Independence (CI). The opinions $\mathbf{o}_1, \mathbf{o}_2, \ldots$ are independent conditional on the common causes $\boldsymbol{\chi}$.

Independence: third version (cont.)

Nice about CI:

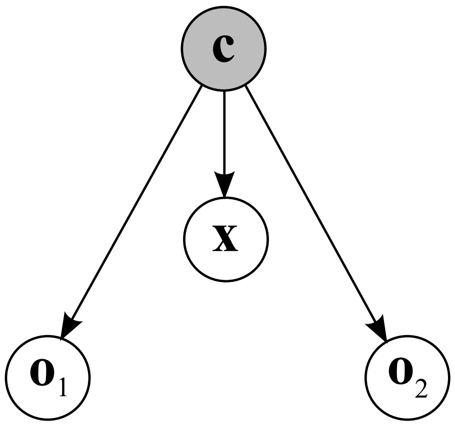
- CI is backed by the Common Cause Principle and more generally by probabilistic theories of causality.
- CI is guaranteed to hold as long as the opinions do not causally affect each other.

Flaw of CI (and of UI):

- CI doesn't lend itself easily to arguments for the 'wisdom of crowds'.
- CI isn't a suitable premise for a jury theorem.
- Why? Next slide.

Independence: third version (cont.)

- CI may fail to conditionalise on the state $\mathbf{x}$.
  - Because $\mathbf{x}$ need not be a common cause; e.g., in a murder case the jurors might learn that the defendant has bought cyanide ($\mathbf{c}$).



- What's so important about conditionalising on $\mathbf{x}$?
- Next slide!

Independence: third version (cont.)

- What matters ultimately isn't independence of *opinions* but independence of the *events of correct opinions*, i.e., the events that opinions match $\mathbf{x}$
  - because we typically want to argue that a group whose members are independently more likely *to get it right* will quite probably get it right in majority.
- Fortunately, independence of opinions implies independence of the events of correct opinions, *provided that we conditionalise on* $\mathbf{x}$.

Independence: fourth version

- In response: let's conditionalise on all common causes *plus* the state.
- As in Part 1 we conceptualize the group's *decision problem* as being composed of two things:
  - the fact to find out about, conceptualized as the state $\mathbf{x}$,
  - the circumstances in which people search, conceptualized as the common causes influencing the opinions.
- So, what we need to conditionalise on is the decision problem.

Independence: fourth version (cont.)

- Formally, we write $\boldsymbol{\pi}$ for the decision problem defined as a family containing the state $\mathbf{x}$ and all common causes.
  - $\boldsymbol{\pi}$ reduces $\boldsymbol{\chi}$ if $\mathbf{x}$ is a common cause
  - Note: $\boldsymbol{\pi}$ was a primitive in Part 1's model, but not it's defined from the network

Problem-Conditional Independence (PI). The opinions $\mathbf{o}_1$, $\mathbf{o}_2$, ... are independent conditional on the problem $\boldsymbol{\pi}$.

Outline

Part 1

[...]

Part 2

Goal: Causal foundations

Four types of probabilistic independence

Theorem

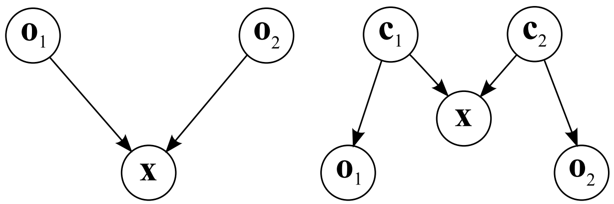
Theorem

- We now give precise sufficient (and in fact essentially necessary) conditions on causal environment for each independence condition to hold.
- To infer probabilistic features from causal interconnections, we must of course assume that probabilities are **compatible** with the causal network
  - (in the standard sense of the *Parental Markov Condition*: any variable in the network is independent of its non-effects conditional on its direct causes).

Theorem (cont.)

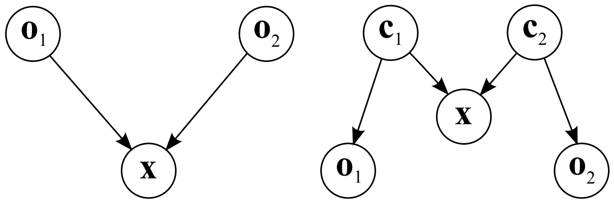
Theorem 1. Suppose probabilities are compatible with the causal network, and no opinion is a cause of any other opinion. Then:
(a) Common-Cause-Conditional Independence holds;
(b) Problem-Conditional Independence holds if the state is not a common effect of any opinions or private causes thereof.

Part (a): an instance of the Common Cause Principle and as such should come without surprise to specialists.
Part (b): PI 'often' holds, though there are counterexamples:



(6a) (6b)

Theorem (cont.)



(6a) (6b)

- 6a: the opinions *influence* the state.
  - e.g., the prediction of a bank run might cause the bank run.
  - 6a violates a frequent assumptions in social epistemology: that an **external** fact determines what opinion is correct. 'Self-fulfilling prophecies' are ruled out!
- 6b: the state is a common effect of private causes of opinions.
  - e.g., an intelligence agency wants to find out about whether certain subjects will attend a conspiracy meeting ($\mathbf{x}$), and each member i of the agency observes a different subject prior to the potential meeting ($\mathbf{c}_i$)

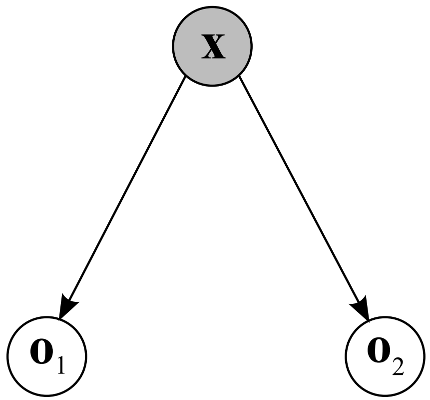
Theorem (cont.)

Let's turn to our remaining two independence conditions:

Corollary. Suppose probabilities are compatible with the causal network, and no opinion is a cause of any other opinion. Then:

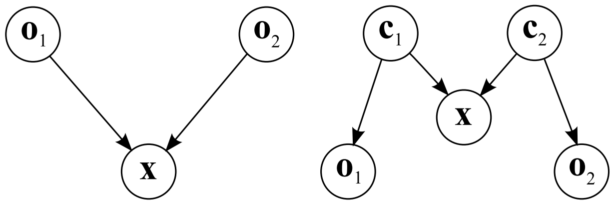
- State-Conditional Independence holds if only the state is a common cause;
- Unconditional Independence holds if there are no common causes at all.

- The conditions supporting SI and UI are rather strong.
- The paradigmatic causal network supporting SI is our first, very simple network:



Theorem (cont.)

- The following (earlier-used) causal network supports UI, since there are *no* common causes:



(6a) (6b)

Conclusion

- To make a notion of independence realistic, the conditionalisation should include the common causes.
- To make it suitable for jury theorems (or 'wisdom of crowds' arguments), the conditionalisation should include the state of the world.

	more realistic	less realistic
lends itself to jury theorems	PI	SI
doesn't	CI	UI

- A jury theorem based on SI: classical CJT
- A jury theorem based on PI: see Part 1 of the talk

Conclusion (cont.)

Future challenges:

- Develop the causal approach!
- Develop jury theorems for the aggregation of *non-binary* opinions, such as judgment sets or degrees of belief.

Conclusion (cont)

- The classical CJT's implausible independence assumption must be held responsible for its implausible conclusion of 'infallible crowds'.
- The goal hasn't been to banish the 'Condorcetian programme' of a formal epistemic justification for democracy
  ... but to place it on better premises
  ... and to vindicate the epistemic merits of deliberation.