CAUSAL EXPLANATORY POWER, BLAMEWORTHINESS, AND NEGLIGENCE

BENJAMIN EVA (KONSTANZ) & REUBEN STERN (LMU MUNICH)
JUNE 2018 – TURIN, ITALY

BLAMEWORTHINESS AND CAUSAL EXPLANATORY POWER

NOT SO FAST!

LIQUEFIED ASBESTOS in handy pressurized cans for spraying on heating pipes, water pipes, above furnaces and around hot-air registers.

CAUSAL EXPLANATORY POWER – STEP 1

- Use a graphical causal model to represent the causal system in which the *explanans* and *explanandum* are embedded, where the probability distribution encodes what would have been reasonable to expect about these events, given what we believe now about the causal system at play.

CAUSAL EXPLANATORY POWER – STEP 2

- Update on all your background knowledge (excluding only the *explanans* and the *explanandum*) regarding the causal system by intervening to set the relevant variables to their known values.

CAUSAL EXPLANATORY POWER – STEP 3

- Use the updated probability distribution to calculate the explanatory power that the intervention to bring about the *explanans* exerts over the *explanandum* by deploying Schupbach and Sprenger's (2011) measure of statistical relevance, ε.

CAUSAL EXPLANATORY POWER & ASBESTOS

- Our framework for assessing causal explanatory power says that the asbestos guy's actions powerfully causally explain the bad outcomes (because the intervention to install asbestos in S's home is highly statistically relevant to whether S gets lung cancer in the updated probability distribution).

CAPTURING BLAMEWORTHINESS

- What are the minimal changes that must be made to this framework in order to render it suitable for measuring blameworthiness?

REVISING THE PROCEDURE

- Use a graphical causal model to represent the causal system in which the **agent's *action, a,* and *outcome, o***, are embedded, where the probability distribution encodes what **the agent should have expected**.

- Update on anything that **the agent should have known** about the causal system (excluding only the action and the outcome) by intervening to set the relevant variables to their known values.

- Calculate the power that $do(a)$ exerts over o using Schupbach and Sprenger's (2011) measure, ε, relative to the updated distribution.
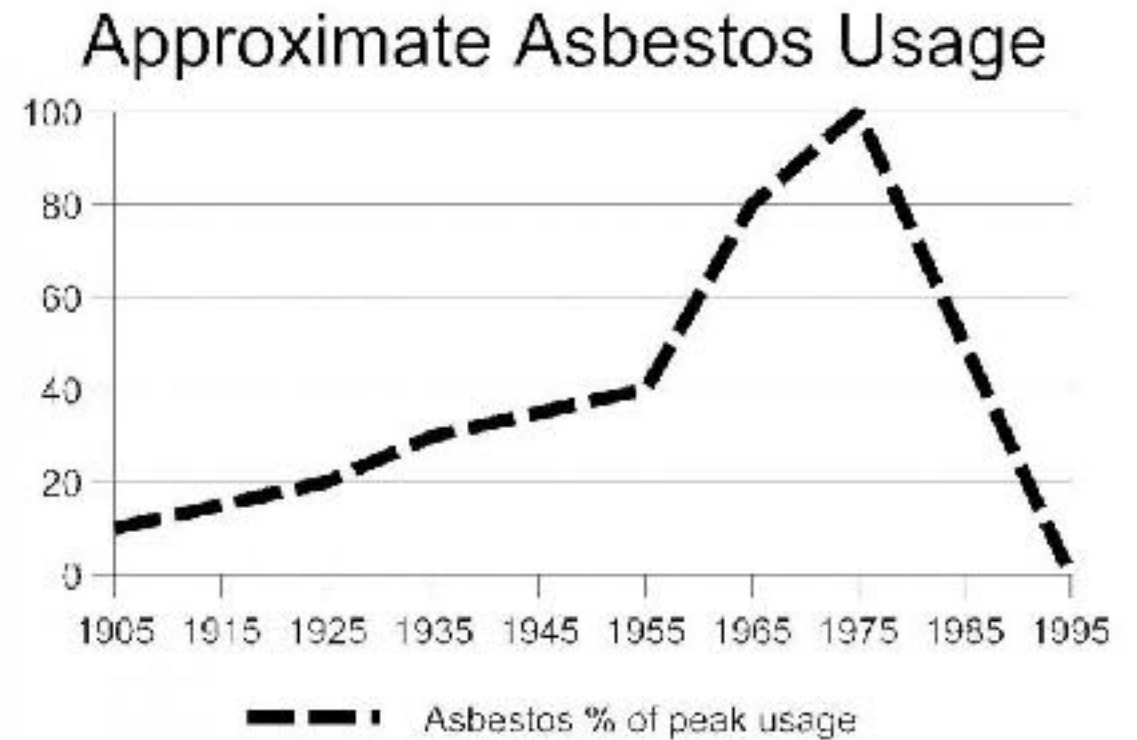
WHAT THE AGENT SHOULD HAVE KNOWN

- Depends on when the agent acted – e.g., 1970 vs. 2018.
- Depends on the agent's capacities – e.g., kids vs. adults.
- Depends on the agent's societal role – e.g., construction workers vs. government officials.

PROBLEM SOLVED?

ASBESTOS IN 1985

Approximate Asbestos Usage

Asbestos % of peak usage

EQUALLY BLAMEWORTHY?

- When the revised procedure is used to assess the blameworthiness of the 1985 guys, they come out as equally blameworthy.

- But surely the one who knew what he should have known about the harms of asbestos is more blameworthy.

NEGLIGENCE VS. RECKLESSNESS

- Let us say that the agent who knows that asbestos is harmful is *reckless*.
- Let us say that who doesn't know what she should know – that asbestos is harmful – is considered *negligent*.
- Negligent agents aren't as blameworthy as reckless agents.
- In order to account for this, we must *discount* the extent to which negligent individuals should be blamed.

INCORPORATING NEGLIGENCE

- Use the earlier procedure to establish how *e-blameworthy* the individual is relative to what **she should have known and expected**.

- Use a slightly revised procedure to establish how *i-blameworthy* the individual is relative to what **she actually knew and expected**.

- Use the difference between these amounts to somehow discount the amount of blame that is output in the first step of this procedure.

HOW TO DISCOUNT FOR NEGLIGENCE

- We need some discount function f that takes as an argument the difference between the agent's e-blameworthiness and i-blameworthiness and returns a discount term.
- The agent's overall blameworthiness is then calculated by subtracting the discount term from the agent's e-blameworthiness.

THE SHAPE OF THE DISCOUNT FUNCTION

- Identifying the exact form of *f* is a substantive philosophical problem.
  - Should the output of *f* be 0 when the difference is 0? (Yes.)
  - Does the value of *f* vary linearly with the size of the difference between *e*-blameworthiness and *i*-blameworthiness?
  - What are the upper and lower bounds of *f*?
  - Can the output of *f* be negative?

PROBLEM SOLVED!

WHAT ABOUT MOTIVE?

PROBABILITIES VS. UTILITIES

- It is plausible that the extent to which an agent should be blamed depends not only her beliefs (or probabilities), but also her desires (or utilities).

  - Probabilistic treatments of explanatory power can help to capture the first kind of consideration, but not the second, and therefore may need to be supplemented.

  - But it's also plausible that the extent to which the agent should be blamed for the *outcome* does not depend on utilities.

IN SUM

- Causal explanatory power and blameworthiness are related, but not the same thing.
- Unlike the extent to which an agent's action causally explains some outcome, the extent to which the agent is blameworthy for the outcome depends on what the agent should have anticipated and did anticipate.
- When it comes to developing a rigorous degreed notion of blameworthiness, the future is promising.

THE RECIPE FOR DETERMINING BLAMEWORTHINESS

We can plausibly assess the blameworthiness of an agent for some outcome if we have the following ingredients:

1. A description of the epistemic state that the agent should have occupied.
2. A description of the agent's actual epistemic state at the time of her action.
3. A specification of the discount function f.

OPEN QUESTIONS

1. When assessing blameworthiness, what considerations are relevant to the determination of the epistemic state that the agent should have occupied?

2. Are there multiple epistemic states that the agent could have permissibly occupied?

3. If so, how can our framework accommodate this?

4. How should we axiomatize the discount function f?

5. Other difficult cases?

THANKS!