KU LEUVEN

# Causal reasoning in a logic with possible causal process semantics

**NMR 2018**

Marc Denecker, Bart Bogaerts, Joost Vennekens

October 29, 2018

# 0 Outline

# 1 Outline

# 1 A common theme of two talks: Construction Knowledge

- Construction knowledge
  - In many domains, certain objects can be constructed in terms of other objects
  - Human experts know how
- Two areas
  - Causal knowledge: the archetypical form?

    the causal process
  - Inductive, recursive and other constructions in mathematics and in logics

    the induction process

# 2 Outline




KU LEUVEN

## 2 Formalizing actual causation

What does it mean to say:

C caused E?

## 2 The counterfactual approach

- Lewis (1973):

  C caused E := "Without C, E would not have been"

- Pearl (2000)

- Halpern & Pearl (2005) (HP)

- Halpern (2016), Fenton-Glynn (2015), Gerstenberg (2015), Vennekens (2011), . . .

- Counterfactual definitions of actual causation in the context of structural equation models.

## 2 Criticisms against counterfactual definitions

- Objections against counterfactual approach

  Regulatory definitions

  Hall (2004), Baumgartner (2013), Bochman & Lifschitz (2015)

- Problematic causal scenarios for all counterfactual definitions
  $\Rightarrow$ Refinements of the definitions

Halpern (2016b) analyzes 6 of these problematic causal scenarios.

## 2 Analysis

Halpern (2016b), scenario A

> *"There are four endogenous binary variables, A, B, C, and S, taking values 1 (on) and 0 (off). Intuitively, A and B are supposed to be alternative causes of C, and S acts as a switch. If $S = 0$, the causal route from A to C is active and that from B to C is dead; and if $S = 1$, the causal route from A to C is dead and the one from B to C is active."*

## 2 Analysis

Halpern (2016b), scenario A

> *"There are four endogenous binary variables, A, B, C, and S, taking values 1 (on) and 0 (off). Intuitively, A and B are supposed to be alternative causes of C, and S acts as a switch. If S = 0, the causal route from A to C is active and that from B to C is dead; and if S = 1, the causal route from A to C is dead and the one from B to C is active."*

$$C := (\neg S \wedge A) \vee (S \wedge B)$$

## 2 Analysis

Halpern (2016b), scenario A

> *"There are four endogenous binary variables, A, B, C, and S, taking values 1 (on) and 0 (off). Intuitively, A and B are supposed to be alternative causes of C, and S acts as a switch. If S = 0, the causal route from A to C is active and that from B to C is dead; and if S = 1, the causal route from A to C is dead and the one from B to C is active."*

$$C := (\neg S \wedge A) \vee (S \wedge B)$$

What is the actual cause of $C$, intuitively?

- when $S$, then $A$
- when $\neg S$, then $B$


KU LEUVEN

## 2 Analysis

Halpern (2016b), scenario A'

> *"But now consider a slightly different story. This time, we view $B$ as the switch, rather than $S$. If $B = 1$, then $C = 1$ if either $A = 1$ or $S = 1$; if $B = 0$, then $C = 1$ only if $A = 1$ and $S = 0$."*

## 2 Analysis

Halpern (2016b), scenario A'

> *"But now consider a slightly different story. This time, we view B as the switch, rather than S. If B = 1, then C = 1 if either A = 1 or S = 1; if B = 0, then C = 1 only if A = 1 and S = 0."*

$$C := (B \wedge (A \vee S)) \vee (\neg B \wedge (A \wedge \neg S))$$

## 2 Analysis

Halpern (2016b), scenario A'

> *"But now consider a slightly different story. This time, we view B as the switch, rather than S. If B = 1, then C = 1 if either A = 1 or S = 1; if B = 0, then C = 1 only if A = 1 and S = 0."*

$$C := (B \land (A \lor S)) \lor (\neg B \land (A \land \neg S))$$

What is the actual cause of $C$, intuitively?

- when $B$, then $A$ or $S$ or both
- when $\neg B$, then $A$ and $\neg S$

## 2 Structural equation models are ambiguous

- The two structural equations are mathematically equivalent!
- In both scenarios, the possible causal worlds are the same.
- Yet, the intuitive answer to actual causation problems is different.
- HP is correct in one of the scenarios.

## 2 Structural equation models are ambiguous

- The two structural equations are mathematically equivalent!
- In both scenarios, the possible causal worlds are the same.
- Yet, the intuitive answer to actual causation problems is different.
- HP is correct in one of the scenarios.

These structural equation models are ambiguous.

## 2 Structural equation models are ambiguous

- It must be the case that some information of these informal scenarios is not expressed by the structural equation model.
- This information does not affect the possible causal worlds.
- This information affects the answer to actual causation problems!

## 2 Structural equation models are ambiguous

- It must be the case that some information of these informal scenarios is not expressed by the structural equation model.
- This information does not affect the possible causal worlds.
- This information affects the answer to actual causation problems!

What kind of information is that? Let's go back to the example.

## 2 The extra information

*"There are four endogenous binary variables, A, B, C, and S, taking values 1 (on) and 0 (off). Intuitively, A and B are supposed to be alternative causes of C, and S acts as a switch. If $S = 0$, the causal route from A to C is active and that from B to C is dead; . . . "*

The extra information:

- separate causal mechanisms
- causes versus switches for causal mechanisms
- causal processes
- causal mechanisms can be alive or dead

dead $\sim$preempted

## 2 Solutions for the ambiguity

- Halpern's solution is a KR methodology :
  "*to add [. . . ] extra variables, which [. . . ] capture the* **mechanism of causality**".

- The approach of this paper:
  - Develop a formal language in which the missing information can be expressed.
  - Regulatory definition(s) of actual causation that exploits the extra information.

# 3 Outline

## 3 The idea

- We see separate causal mechanisms
- Some sets of conditions trigger the causal mechanism
- Other conditions could preempt the causal mechanism if not true.

Information about this strongly influences our idea of actual causation.

# 3 Syntax: Causal theories

**Definition**

A causal theory is a a set of causal mechanisms.

**Definition**

A causal mechanism, or causal law, is an expression of the form

$$A \leftarrow T \,||\, P$$

where $A$ is a literal, $T$ and $S$ sequences of literals

- A literal of $T$ is called a trigger condition of the causal mechanism.
- A literal of $P$ is called a no-preemption condition of the causal mechanism.

# 3 Example

- Scenario A:

$$\left\{ \begin{array}{l} C \leftarrow A \,||\, \neg S \\ C \leftarrow B \,||\, S \end{array} \right\}$$

- Scenario A':

$$\left\{ \begin{array}{l} C \leftarrow A \,||\, B \\ C \leftarrow S \,||\, B \\ C \leftarrow A, S \,||\, \neg B \end{array} \right\}$$

- We made the information explicit that was available in Halperns informal domain description.

## 3 Semantics

A possible causal world semantics is not refined enough.

# 3 Semantics: possible causal processes

- The formal semantics specifies, for a causal theory $\Delta$:
  - the possible causal processes of $\Delta$
  - the possible causal world that each process leads to.
- How to formalize the causal process?
  - a causal process $\sim$ a dependency graph of the causal mechanisms that fire.

## 3 Another example: double preemption

Hall (2004)

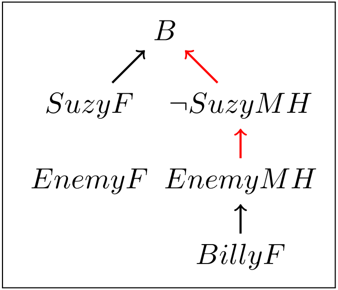
> *Suzy fires a missile (SuzyF) to bomb a target (B); Enemy fires a missile (EnemyF ) to hit Suzy's missile (SuzyMH) and Billy fires a missile (BillyF) to hit Enemy's missile (EnemyMH).*

Theory:

$$\left\{ \begin{array}{l} B \leftarrow SuzyF \,||\, \neg SuzyMH \\ SuzyMH \leftarrow EnemyF \,||\, \neg EnemyMH \\ EnemyMH \leftarrow BillyF \,|| \end{array} \right\}$$

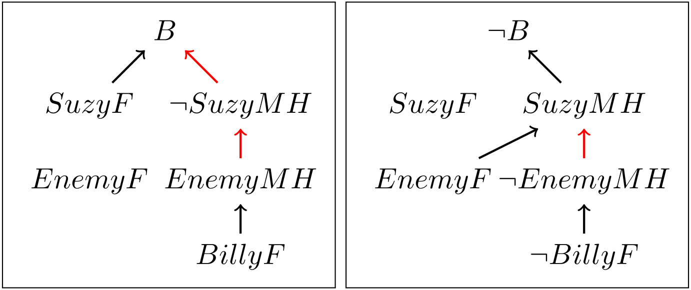
# 3 Another example: double preemption

$$\left\{ \begin{array}{l} B \leftarrow SuzyF \,||\, \neg SuzyMH \\ SuzyMH \leftarrow EnemyF \,||\, \neg EnemyMH \\ EnemyMH \leftarrow BillyF \,|| \end{array} \right\}$$



# 3 Another example: double preemption

$$\left\{\begin{array}{l} B \leftarrow SuzyF \,||\, \neg SuzyMH \\ SuzyMH \leftarrow EnemyF \,||\, \neg EnemyMH \\ EnemyMH \leftarrow BillyF \,|| \end{array}\right\}$$



# 3 Derived concepts and properties

Derived concepts:

- An actual possible causal process induces a unique possible causal world
  - The possible causal process semantics is more refined than the possible world semantics.
- In a possible world, a causal mechanism can be:
  - firing
  - triggered but preempted
  - non-triggered

# 3 Derived concepts and properties

Derived concepts:

- An actual possible causal process induces a unique possible causal world
  - The possible causal process semantics is more refined than the possible world semantics.
- In a possible world, a causal mechanism can be:
  - firing
  - triggered but preempted
  - non-triggered

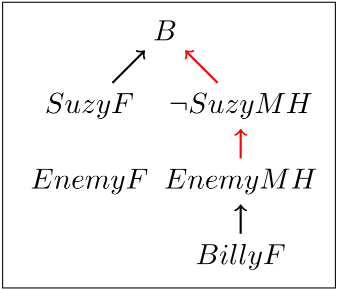
Some derived properties:

- All processes in the same exogeneous state cause the same possible world (confluence property)
- The possible causal worlds of $\Delta$ are the causal worlds of the structural equation model $Completion(\Delta)$

# 4 Outline

# 4 Definitions of actual causation

The possible causal process is a detailed explanation of the world.




- x is an influence of y in possible causal process $\mathcal{P}$
- x is an actual P-cause of y
- x is an actual DP-cause of y



KU LEUVEN

# 5 Outline

## 5 Conclusions

- A study of several sorts of knowledge that are important for actual causation but are not or not well expressed in many causal languages.
- Logic equipped with a possible causal process semantics.
- Some fundamental aspects of causation: the confluence of causal processes and, paradoxically, a theorem explaining why many useful causation problems can be solved without modelling mechanisms and processes.
- A rich and flexible framework for defining several notions of actual causation.

## 5 Future work

- Counterfactual definitions versus Regularity definition: two different sides of the same coin?
- Extending the logic: predicate logic, cyclic causal theories, . . .

## 5 Implementation on-line

http://adams.cs.kuleuven.be/idp/server.html?chapter=intro/11-AC

- An on-line implementation of many of the examples in the paper
- Using the knowledge base system IDP