Causation in Semantics and Grammatical Structure

Week 10: Approaches to modeling causation

Prerna Nadathur

December 12, 2019

Event semantics: recap

In "The Individuation of Events", Davidson (1969) is trying to do two things:

Event semantics: recap

In "The Individuation of Events", Davidson (1969) is trying to do two things:

1. provide arguments in support of introducing events as elements in a semantic interpretation system (and thus as a conceptual object)

Event semantics: recap

In "The Individuation of Events", Davidson (1969) is trying to do two things:

1. provide arguments in support of introducing events as elements in a semantic interpretation system (and thus as a conceptual object)
2. put events on a par with other conceptual/semantic objects, by providing a means of understanding how to identify and differentiate events

Event semantics: recap

In "The Individuation of Events", Davidson (1969) is trying to do two things:

1. provide arguments in support of introducing events as elements in a semantic interpretation system (and thus as a conceptual object)
2. put events on a par with other conceptual/semantic objects, by providing a means of understanding how to identify and differentiate events

Davidson (1969):

- builds on Davidson (1967)

Event semantics: recap

In "The Individuation of Events", Davidson (1969) is trying to do two things:

1. provide arguments in support of introducing events as elements in a semantic interpretation system (and thus as a conceptual object)
2. put events on a par with other conceptual/semantic objects, by providing a means of understanding how to identify and differentiate events

Davidson (1969):

- builds on Davidson (1967)
- . . . but goes beyond the earlier paper, in connecting events with notions of causation

The logical form of action sentences (1967)

(1) Brutus stabbed Caesar with a knife in the forum.

a. **entails:** Brutus stabbed Caesar with a knife.

b. **entails:** Brutus stabbed Caesar in the forum.

c. **entails:** Brutus stabbed Caesar.

The logical form of action sentences (1967)

(1) Brutus stabbed Caesar with a knife in the forum.

a. **entails:** Brutus stabbed Caesar with a knife.

b. **entails:** Brutus stabbed Caesar in the forum.

c. **entails:** Brutus stabbed Caesar.

Prior to Davidson:

- (1) is represented by the verb STAB, with a set of arguments:

(2) (1) $\equiv$ STAB(B, C, *with a knife*, *in the forum*)

The logical form of action sentences (1967)

(1) Brutus stabbed Caesar with a knife in the forum.

a. **entails:** Brutus stabbed Caesar with a knife.

b. **entails:** Brutus stabbed Caesar in the forum.

c. **entails:** Brutus stabbed Caesar.

Prior to Davidson:

- (1) is represented by the verb STAB, with a set of arguments:

  (2) (1) $\equiv$ STAB(B, C, *with a knife*, *in the forum*)

- **problem 1:** not all of the NPs and PPs in (1) are *required* arguments of the verb *stab*

The logical form of action sentences (1967)

(1) Brutus stabbed Caesar with a knife in the forum.

a. **entails:** Brutus stabbed Caesar with a knife.

b. **entails:** Brutus stabbed Caesar in the forum.

c. **entails:** Brutus stabbed Caesar.

Prior to Davidson:

- (1) is represented by the verb STAB, with a set of arguments:

  (2) (1) $\equiv$ STAB(B, C, *with a knife*, *in the forum*)

- **problem 1:** not all of the NPs and PPs in (1) are *required* arguments of the verb *stab*
- **problem 2:** how can we explain the entailment relationships in (1) in a natural way?

The logical form of action sentences (1967)

(1) Brutus stabbed Caesar with a knife in the forum.

a. **entails:** Brutus stabbed Caesar with a knife.

b. **entails:** Brutus stabbed Caesar in the forum.

c. **entails:** Brutus stabbed Caesar.

Prior to Davidson:

- (1) is represented by the verb STAB, with a set of arguments:

  (2) (1) $\equiv$ STAB(*B*, *C*, *with a knife*, *in the forum*)

- **problem 1:** not all of the NPs and PPs in (1) are *required* arguments of the verb *stab*
- **problem 2:** how can we explain the entailment relationships in (1) in a natural way?
  - (that is, without simply stipulating a number of meaning postulates relating various different *stab* verbs)

The logical form of action sentences

Davidson's solution: action sentences describe events

The logical form of action sentences

Davidson's solution: action sentences describe events

- *stab* has certain 'core' arguments (AGENT, PATIENT)

The logical form of action sentences

Davidson's solution: action sentences describe events

- *stab* has certain 'core' arguments (AGENT, PATIENT)
- *with a knife* and *in the forum* are **adjuncts** which specify further information about the stabbing **event**

The logical form of action sentences

Davidson's solution: action sentences describe events

- *stab* has certain 'core' arguments (AGENT, PATIENT)
- *with a knife* and *in the forum* are **adjuncts** which specify further information about the stabbing **event**
- we can add this information with conjuncts relating the main event to additional participants (using semantic roles)

The logical form of action sentences

Davidson's solution: action sentences describe events

- *stab* has certain 'core' arguments (AGENT, PATIENT)
- *with a knife* and *in the forum* are **adjuncts** which specify further information about the stabbing **event**
- we can add this information with conjuncts relating the main event to additional participants (using semantic roles)

(3) Brutus stabbed Caesar with a knife in the forum.
$\equiv \exists e :$ STAB(e, B, C) & INST$(e, \textit{knife})$ & LOC$(e, \textit{forum})$

a. **entails:** $\exists e :$ STAB(e, B, C) & INST$(e, \textit{knife})$

b. **entails:** $\exists e :$ STAB(e, B, C) & LOC$(e, \textit{forum})$

c. **entails:** $\exists e :$ STAB(e, B, C)

The logical form of action sentences

Davidson's solution: action sentences describe events

- *stab* has certain 'core' arguments (AGENT, PATIENT)
- *with a knife* and *in the forum* are **adjuncts** which specify further information about the stabbing **event**
- we can add this information with conjuncts relating the main event to additional participants (using semantic roles)

(3) Brutus stabbed Caesar with a knife in the forum.
$\equiv \exists e :$ STAB(e, B, C) & INST$(e, \textit{knife})$ & LOC$(e, \textit{forum})$

a. **entails:** $\exists e :$ STAB(e, B, C) & INST$(e, \textit{knife})$
b. **entails:** $\exists e :$ STAB(e, B, C) & LOC$(e, \textit{forum})$
c. **entails:** $\exists e :$ STAB(e, B, C)

- the entailment relationships in (1) follow using basic logical principles

Events in semantics

What are the basic objects in our semantic interpretation system?

- individuals (*Brutus*, *Caesar*, . . .)
- truth values (T, F)
- (functions from one type to another)

Events in semantics

What are the basic objects in our semantic interpretation system?

- individuals (*Brutus*, *Caesar*, . . .)
- truth values (T, F)
- (functions from one type to another)

Events are like individuals, but abstract:

- are they referenced directly?
  - maybe by certain kinds of nominal: *the killing of Caesar*, *the destruction of Carthage*

Events in semantics

What are the basic objects in our semantic interpretation system?

- individuals (*Brutus*, *Caesar*, . . .)
- truth values (T, F)
- (functions from one type to another)

Events are like individuals, but abstract:

- are they referenced directly?
  - maybe by certain kinds of nominal: *the killing of Caesar*, *the destruction of Carthage*
- Davidson's (1969) point: having events as a basic object isn't any stranger than having individuals

The individuation of events

We have intuitions about when two individuals are identical:

- do we have the same sense for events?

The individuation of events

We have intuitions about when two individuals are identical:

- do we have the same sense for events?
- **no entity without identity:**

The individuation of events

We have intuitions about when two individuals are identical:

- do we have the same sense for events?
- **no entity without identity:**
  - if events are real objects, we should 'know them when we see them'

The individuation of events

We have intuitions about when two individuals are identical:

- do we have the same sense for events?
- **no entity without identity:**
  - if events are real objects, we should 'know them when we see them'
  - even if we can't define events directly, we have ways of telling when events are the same or different

The individuation of events

We have intuitions about when two individuals are identical:

- do we have the same sense for events?
- **no entity without identity:**
  - if events are real objects, we should 'know them when we see them'
  - even if we can't define events directly, we have ways of telling when events are the same or different
  - just like with individuals, the descriptions may not be the same:

(4) The morning star is the evening star.

(5) a. Brutus stabbed Caesar.

b. Caesar was killed.

The individuation of events

We have intuitions about when two individuals are identical:

- do we have the same sense for events?
- **no entity without identity:**
  - if events are real objects, we should 'know them when we see them'
  - even if we can't define events directly, we have ways of telling when events are the same or different
  - just like with individuals, the descriptions may not be the same:

    (4) The morning star is the evening star.

    (5) a. Brutus stabbed Caesar.

    b. Caesar was killed.

So, if we can specify identity conditions for events, this eliminates some arguments against including them in our semantics

- for us, the identity conditions themselves are more interesting!

The individuation of events

Some problems for intuitions about when two events are the same:

The individuation of events

Some problems for intuitions about when two events are the same:

- . . . if they occur at the same time

  (6) Jones buttered the toast at midnight.

  (7) The clock struck at midnight.

The individuation of events

Some problems for intuitions about when two events are the same:

- . . . if they occur at the same time

  (6) Jones buttered the toast at midnight.

  (7) The clock struck at midnight.

- . . . the same time AND place

  (8) Jones buttered the toast at midnight in the bathroom.

  (9) The clock struck at midnight in the bathroom.

The individuation of events

Some problems for intuitions about when two events are the same:

- . . . if they occur at the same time

  (6) Jones buttered the toast at midnight.

  (7) The clock struck at midnight.

- . . . the same time AND place

  (8) Jones buttered the toast at midnight in the bathroom.

  (9) The clock struck at midnight in the bathroom.

- . . . same time, same place, AND involve the same participants

  (10) Jones buttered the toast at midnight in the bathroom by candlelight.

  (11) Jones heated the toast at midnight in the bathroom with a candle.

The individuation of events

Some problems for intuitions about when two events are the same:

- . . . if they occur at the same time

  (6) Jones buttered the toast at midnight.

  (7) The clock struck at midnight.

- . . . the same time AND place

  (8) Jones buttered the toast at midnight in the bathroom.

  (9) The clock struck at midnight in the bathroom.

- . . . same time, same place, AND involve the same participants

  (10) Jones buttered the toast at midnight in the bathroom by candlelight.

  (11) Jones heated the toast at midnight in the bathroom with a candle.

  - are the heating and buttering events the same?

The individuation of events

Identity condition for events
Events, e_1 and e_2, are identical if and only if they have the same **causes** and **effects**:

$$e_1 = e_2 \text{ if and only if}$$
$$\forall z.(z \text{ caused } e_1 \leftrightarrow z \text{ caused } e_2) \;\&\; \forall z.(e_1 \text{ caused } z \leftrightarrow e_2 \text{ caused } z)$$

The individuation of events

Identity condition for events
Events, e_1 and e_2, are identical if and only if they have the same **causes** and **effects**:

$$e_1 = e_2 \text{ if and only if}$$
$$\forall z.(z \text{ caused } e_1 \leftrightarrow z \text{ caused } e_2) \;\&\; \forall z.(e_1 \text{ caused } z \leftrightarrow e_2 \text{ caused } z)$$

- causal relationships define/identify events (not a particular description)

The individuation of events

Identity condition for events
Events, e_1 and e_2, are identical if and only if they have the same **causes** and **effects**:

$$e_1 = e_2 \text{ if and only if}$$
$$\forall z.(z \text{ caused } e_1 \leftrightarrow z \text{ caused } e_2) \;\&\; \forall z.(e_1 \text{ caused } z \leftrightarrow e_2 \text{ caused } z)$$

- causal relationships define/identify events (not a particular description)

  (12) a. Brutus stabbed Caesar
  b. Brutus killed Caesar
  c. Brutus betrayed Caesar

- (12)a-c can all describe the same event, provided:

The individuation of events

Identity condition for events
Events, e_1 and e_2, are identical if and only if they have the same **causes** and **effects**:

$$e_1 = e_2 \text{ if and only if}$$
$$\forall z.(z \text{ caused } e_1 \leftrightarrow z \text{ caused } e_2) \;\&\; \forall z.(e_1 \text{ caused } z \leftrightarrow e_2 \text{ caused } z)$$

- causal relationships define/identify events (not a particular description)

  (12) a. Brutus stabbed Caesar
  b. Brutus killed Caesar
  c. Brutus betrayed Caesar

- (12)a-c can all describe the same event, provided:
  - they have the same causes and effects

The individuation of events

Identity condition for events
Events, e_1 and e_2, are identical if and only if they have the same **causes** and **effects**:

$$e_1 = e_2 \text{ if and only if}$$
$$\forall z.(z \text{ caused } e_1 \leftrightarrow z \text{ caused } e_2) \;\&\; \forall z.(e_1 \text{ caused } z \leftrightarrow e_2 \text{ caused } z)$$

- causal relationships define/identify events (not a particular description)

  (12) a. Brutus stabbed Caesar
  b. Brutus killed Caesar
  c. Brutus betrayed Caesar

- (12)a-c can all describe the same event, provided:
  - they have the same causes and effects
  - as a result of which they must also involve the same participants, and take place at the same time and place

Events and causation

Davidson's idea: we perceive and distinguish events in terms of causal relationships

> "Events are identical if and only if they have exactly the same causes and effects. Events have a unique position in the framework of causal relations . . . in somewhat the same way objects have a unique framework in the spatial framework of objects." Davidson (1969, p.306)

Events and causation

Davidson's idea: we perceive and distinguish events in terms of causal relationships

> "Events are identical if and only if they have exactly the same causes and effects. Events have a unique position in the framework of causal relations . . . in somewhat the same way objects have a unique framework in the spatial framework of objects." Davidson (1969, p.306)

- both events and objects exist as part of collections of alike things

Events and causation

Davidson's idea: we perceive and distinguish events in terms of causal relationships

> "Events are identical if and only if they have exactly the same causes and effects. Events have a unique position in the framework of causal relations . . . in somewhat the same way objects have a unique framework in the spatial framework of objects." Davidson (1969, p.306)

- both events and objects exist as part of collections of alike things
- these things relate to one another

Events and causation

Davidson's idea: we perceive and distinguish events in terms of causal relationships

> "Events are identical if and only if they have exactly the same causes and effects. Events have a unique position in the framework of causal relations . . . in somewhat the same way objects have a unique framework in the spatial framework of objects." Davidson (1969, p.306)

- both events and objects exist as part of collections of alike things
- these things relate to one another
- concrete objects have spatial (concrete) relationships

Events and causation

Davidson's idea: we perceive and distinguish events in terms of causal relationships

> "Events are identical if and only if they have exactly the same causes and effects. Events have a unique position in the framework of causal relations . . . in somewhat the same way objects have a unique framework in the spatial framework of objects." Davidson (1969, p.306)

- both events and objects exist as part of collections of alike things
- these things relate to one another
- concrete objects have spatial (concrete) relationships
- events have causal relationships

Events and causation

Davidson's idea: we perceive and distinguish events in terms of causal relationships

> "Events are identical if and only if they have exactly the same causes and effects. Events have a unique position in the framework of causal relations . . . in somewhat the same way objects have a unique framework in the spatial framework of objects." Davidson (1969, p.306)

- both events and objects exist as part of collections of alike things
- these things relate to one another
- concrete objects have spatial (concrete) relationships
- events have causal relationships

Croft 1991: pushes on this idea by proposing to account for internal event structure in terms of causal relationships

Approaches to modeling causation

This gives us a new way of thinking about causation:

Approaches to modeling causation

This gives us a new way of thinking about causation:

- the analogy between concrete 'objects' and events is suggestive

Approaches to modeling causation

This gives us a new way of thinking about causation:

- the analogy between concrete 'objects' and events is suggestive
- spatial (and temporal) relationships have to do with the way human cognition works:

Approaches to modeling causation

This gives us a new way of thinking about causation:

- the analogy between concrete 'objects' and events is suggestive
- spatial (and temporal) relationships have to do with the way human cognition works:
  - they are *forms of experience* (Kant 1781), or ways in which we perceive the world

Approaches to modeling causation

This gives us a new way of thinking about causation:

- the analogy between concrete 'objects' and events is suggestive
- spatial (and temporal) relationships have to do with the way human cognition works:
  - they are *forms of experience* (Kant 1781), or ways in which we perceive the world
- causal relationships are also about cognition, and also constitute a way in which we perceive the world

Approaches to modeling causation

This gives us a new way of thinking about causation:

- the analogy between concrete 'objects' and events is suggestive
- spatial (and temporal) relationships have to do with the way human cognition works:
  - they are *forms of experience* (Kant 1781), or ways in which we perceive the world
- causal relationships are also about cognition, and also constitute a way in which we perceive the world
  - so, just like we have models for geometry that explain the way we perceive spatial relations

Approaches to modeling causation

This gives us a new way of thinking about causation:

- the analogy between concrete 'objects' and events is suggestive
- spatial (and temporal) relationships have to do with the way human cognition works:
  - they are *forms of experience* (Kant 1781), or ways in which we perceive the world
- causal relationships are also about cognition, and also constitute a way in which we perceive the world
  - so, just like we have models for geometry that explain the way we perceive spatial relations
  - we can understand something more about causation-as-cognition by building a model of relationships between events and seeing whether it predicts aspects of our understanding

Approaches to modeling causation

This gives us a new way of thinking about causation:

- the analogy between concrete 'objects' and events is suggestive
- spatial (and temporal) relationships have to do with the way human cognition works:
  - they are *forms of experience* (Kant 1781), or ways in which we perceive the world
- causal relationships are also about cognition, and also constitute a way in which we perceive the world
  - so, just like we have models for geometry that explain the way we perceive spatial relations
  - we can understand something more about causation-as-cognition by building a model of relationships between events and seeing whether it predicts aspects of our understanding
  - linguistic judgements are one way of 'testing' models

Language and causal theories

- philosophers have spent a lot of time arguing about what it means for one thing to *cause* another:

> "...we may define a cause to be *an object followed by another, and where all the objects, similar to the first, are followed by objects similar to the second.* Or, in other words, *where, if the first object had not been, the second never had existed.*" Hume, 1748

Language and causal theories

- philosophers have spent a lot of time arguing about what it means for one thing to *cause* another:

> "...we may define a cause to be *an object followed by another, and where all the objects, similar to the first, are followed by objects similar to the second.* Or, in other words, *where, if the first object had not been, the second never had existed.*" Hume, 1748

- Challenge: how we should get access to the facts?

Language and causal theories

- philosophers have spent a lot of time arguing about what it means for one thing to *cause* another:

> "...we may define a cause to be *an object followed by another, and where all the objects, similar to the first, are followed by objects similar to the second.* Or, in other words, *where, if the first object had not been, the second never had existed.*" Hume, 1748

- Challenge: how we should get access to the facts?
- **Now:** we can use language – we want to build models to 'fit' two kinds of linguistic data:

Language and causal theories

- philosophers have spent a lot of time arguing about what it means for one thing to *cause* another:

> "...we may define a cause to be *an object followed by another, and where all the objects, similar to the first, are followed by objects similar to the second.* Or, in other words, *where, if the first object had not been, the second never had existed.*" Hume, 1748

- Challenge: how we should get access to the facts?
- **Now:** we can use language – we want to build models to 'fit' two kinds of linguistic data:
  1. variations in linguistic marking for different features of causal situations (up to now)

Language and causal theories

- philosophers have spent a lot of time arguing about what it means for one thing to *cause* another:

> "...we may define a cause to be *an object followed by another, and where all the objects, similar to the first, are followed by objects similar to the second.* Or, in other words, *where, if the first object had not been, the second never had existed.*" Hume, 1748

- Challenge: how we should get access to the facts?
- **Now:** we can use language – we want to build models to 'fit' two kinds of linguistic data:
  1. variations in linguistic marking for different features of causal situations (up to now)
  2. intuitions about when *cause* (or other causal language) is appropriate

Language and causal theories

Copley & Wolff (2014):

- given a link between verbs, events, and causation, it must be the case that language structure and linguistic judgements reflect facts about (our conceptualization of) causation

Language and causal theories

Copley & Wolff (2014):

- given a link between verbs, events, and causation, it must be the case that language structure and linguistic judgements reflect facts about (our conceptualization of) causation
- so:
  - if linguistic structure reflects our understanding of causation

Language and causal theories

Copley & Wolff (2014):

- given a link between verbs, events, and causation, it must be the case that language structure and linguistic judgements reflect facts about (our conceptualization of) causation
- so:
  - if linguistic structure reflects our understanding of causation
  - and causation is something like a form of cognitive experience

Language and causal theories

Copley & Wolff (2014):

- given a link between verbs, events, and causation, it must be the case that language structure and linguistic judgements reflect facts about (our conceptualization of) causation
- so:
  - if linguistic structure reflects our understanding of causation
  - and causation is something like a form of cognitive experience
  - then, a model of causal relationships that predicts our linguistic judgements models what we 'know' about causation

Language and causal theories

Copley & Wolff (2014):

- given a link between verbs, events, and causation, it must be the case that language structure and linguistic judgements reflect facts about (our conceptualization of) causation
- so:
  - if linguistic structure reflects our understanding of causation
  - and causation is something like a form of cognitive experience
  - then, a model of causal relationships that predicts our linguistic judgements models what we 'know' about causation
  - . . . which, from this perspective, is all that there really IS to causation in the first place

Language and causal theories

Building a model isn't trivial:

Language and causal theories

Building a model isn't trivial:

- the data are complex (as we've seen), intuitions can be subtle

Language and causal theories

Building a model isn't trivial:

- the data are complex (as we've seen), intuitions can be subtle
- Copley & Wolff: linguists, philosophers can help each other

Language and causal theories

Building a model isn't trivial:

- the data are complex (as we've seen), intuitions can be subtle
- Copley & Wolff: linguists, philosophers can help each other
  - a good place for linguists to start is with existing theories (developed in philosophy, psych, comp sci . . .)

Language and causal theories

Building a model isn't trivial:

- the data are complex (as we've seen), intuitions can be subtle
- Copley & Wolff: linguists, philosophers can help each other
  - a good place for linguists to start is with existing theories (developed in philosophy, psych, comp sci . . .)
  - philosophers can refine their theories by seeing what is needed to explain language data

Language and causal theories

Building a model isn't trivial:

- the data are complex (as we've seen), intuitions can be subtle
- Copley & Wolff: linguists, philosophers can help each other
  - a good place for linguists to start is with existing theories (developed in philosophy, psych, comp sci . . .)
  - philosophers can refine their theories by seeing what is needed to explain language data

Today: major classes of causal theory

- **dependency theories:** pseudo-logical, abstract relationship between cause and effect
  - **focus:** Lewis's counterfactual theory of causation
  - next two weeks: network models and applications
- **production theories:** something is produced and/or transmitted from a cause to an effect
  - later: back to force dynamics and its applications

Dependency theories of causation

For Copley & Wolff, a **dependency theory** of causation involves the idea that there is a (pseudo-)logical, abstract relationship between a cause and an effect:

Dependency theories of causation

For Copley & Wolff, a **dependency theory** of causation involves the idea that there is a (pseudo-)logical, abstract relationship between a cause and an effect:

1. logical dependency theories define this relationship in the terms of formal logic

Dependency theories of causation

For Copley & Wolff, a **dependency theory** of causation involves the idea that there is a (pseudo-)logical, abstract relationship between a cause and an effect:

1. logical dependency theories define this relationship in the terms of formal logic
2. counterfactual dependency theories (following Lewis 1973) use formal logic as a base, but add additional stipulations

Dependency theories of causation

For Copley & Wolff, a **dependency theory** of causation involves the idea that there is a (pseudo-)logical, abstract relationship between a cause and an effect:

1. logical dependency theories define this relationship in the terms of formal logic
2. counterfactual dependency theories (following Lewis 1973) use formal logic as a base, but add additional stipulations
3. probabilistic or statistical dependency theories define the cause-effect relationship in terms of conditional probabilities

Dependency theories of causation

For Copley & Wolff, a **dependency theory** of causation involves the idea that there is a (pseudo-)logical, abstract relationship between a cause and an effect:

1. logical dependency theories define this relationship in the terms of formal logic
2. counterfactual dependency theories (following Lewis 1973) use formal logic as a base, but add additional stipulations
3. probabilistic or statistical dependency theories define the cause-effect relationship in terms of conditional probabilities
4. network dependency theories (often) take a causal link to be abstract, but define (linguistically-relevant) causal relationships in terms of network configurations

Dependency theories of causation

1. Defining causal relationships in terms of classical logic

Dependency theories of causation

1. Defining causal relationships in terms of classical logic

Using formal logic, we have three basic options for defining causal dependence: **necessity, sufficiency**, or a combination

Dependency theories of causation

1. Defining causal relationships in terms of classical logic

Using formal logic, we have three basic options for defining causal dependence: **necessity, sufficiency**, or a combination

(13) **Necessity**: A is logically necessary for B iff $\neg A \rightarrow \neg B$
. . . knowing that A fails, you can be sure that B fails

- *Jakub is male* is logically necessary for *Jakub is a bachelor*

Dependency theories of causation

1. Defining causal relationships in terms of classical logic

Using formal logic, we have three basic options for defining causal dependence: **necessity, sufficiency**, or a combination

(13) **Necessity**: A is logically necessary for B iff $\neg A \rightarrow \neg B$
. . . knowing that A fails, you can be sure that B fails

- *Jakub is male* is logically necessary for *Jakub is a bachelor*

(14) **Sufficiency:** A is logically sufficient for B iff $A \rightarrow B$
. . . knowing that A, you can be sure that B

- *Socrates is a man* is logically sufficient for *Socrates is mortal*

Dependency theories of causation

1. Defining causal relationships in terms of classical logic

Using formal logic, we have three basic options for defining causal dependence: **necessity, sufficiency**, or a combination

(13) **Necessity**: A is logically necessary for B iff $\neg A \rightarrow \neg B$
. . . knowing that A fails, you can be sure that B fails

- *Jakub is male* is logically necessary for *Jakub is a bachelor*

(14) **Sufficiency:** A is logically sufficient for B iff $A \rightarrow B$
. . . knowing that A, you can be sure that B

- *Socrates is a man* is logically sufficient for *Socrates is mortal*

Note that neither of these relationships inherently invokes intuitions about causation

Dependency theories: logical dependency

Late pre-emption scenarios show that causal dependency **can't just be about logical necessity:**

(15) **A case of late pre-emption** (Hall 2004). Billy and Suzy both throw stones at a bottle on the wall. Their throws are on target. Suzy throws first, so her stone reaches the bottle before Billy's, and the bottle breaks.

Dependency theories: logical dependency

Late pre-emption scenarios show that causal dependency **can't just be about logical necessity:**

(15) **A case of late pre-emption** (Hall 2004). Billy and Suzy both throw stones at a bottle on the wall. Their throws are on target. Suzy throws first, so her stone reaches the bottle before Billy's, and the bottle breaks.

a. Billy's throw is *pre-empted* by Suzy's throw

Dependency theories: logical dependency

Late pre-emption scenarios show that causal dependency **can't just be about logical necessity:**

(15) **A case of late pre-emption** (Hall 2004). Billy and Suzy both throw stones at a bottle on the wall. Their throws are on target. Suzy throws first, so her stone reaches the bottle before Billy's, and the bottle breaks.

a. Billy's throw is *pre-empted* by Suzy's throw

b. **intuition:** Suzy caused the bottle to break

Dependency theories: logical dependency

Late pre-emption scenarios show that causal dependency **can't just be about logical necessity:**

(15) **A case of late pre-emption** (Hall 2004). Billy and Suzy both throw stones at a bottle on the wall. Their throws are on target. Suzy throws first, so her stone reaches the bottle before Billy's, and the bottle breaks.

a. Billy's throw is *pre-empted* by Suzy's throw

b. **intuition:** Suzy caused the bottle to break

c. **problem:** Suzy's throw was not *necessary*: if she hadn't thrown, Billy's stone would have hit and broken the bottle

Dependency theories: logical dependency

Late pre-emption scenarios show that causal dependency **can't just be about logical necessity:**

(15) **A case of late pre-emption** (Hall 2004). Billy and Suzy both throw stones at a bottle on the wall. Their throws are on target. Suzy throws first, so her stone reaches the bottle before Billy's, and the bottle breaks.

a. Billy's throw is *pre-empted* by Suzy's throw

b. **intuition:** Suzy caused the bottle to break

c. **problem:** Suzy's throw was not *necessary*: if she hadn't thrown, Billy's stone would have hit and broken the bottle

- to define causation, we'd like to identify a set of conditions that predict all and only those cases where causation is intuitively present

Dependency theories: logical dependency

Late pre-emption scenarios show that causal dependency **can't just be about logical necessity:**

(15) **A case of late pre-emption** (Hall 2004). Billy and Suzy both throw stones at a bottle on the wall. Their throws are on target. Suzy throws first, so her stone reaches the bottle before Billy's, and the bottle breaks.

a. Billy's throw is *pre-empted* by Suzy's throw

b. **intuition:** Suzy caused the bottle to break

c. **problem:** Suzy's throw was not *necessary*: if she hadn't thrown, Billy's stone would have hit and broken the bottle

- to define causation, we'd like to identify a set of conditions that predict all and only those cases where causation is intuitively present
- so those conditions should (collectively) be both necessary and sufficient for a causal relationship

Dependency theories: logical dependency

Late pre-emption scenarios show that causal dependency **can't just be about logical necessity:**

(15) **A case of late pre-emption** (Hall 2004). Billy and Suzy both throw stones at a bottle on the wall. Their throws are on target. Suzy throws first, so her stone reaches the bottle before Billy's, and the bottle breaks.

a. Billy's throw is *pre-empted* by Suzy's throw

b. **intuition:** Suzy caused the bottle to break

c. **problem:** Suzy's throw was not *necessary*: if she hadn't thrown, Billy's stone would have hit and broken the bottle

- to define causation, we'd like to identify a set of conditions that predict all and only those cases where causation is intuitively present
- so those conditions should (collectively) be both necessary and sufficient for a causal relationship
- (15) shows that we can judge causation to be present without logical necessity holding between cause and effect

Dependency theories: logical dependency

Late pre-emption scenarios show that causal dependency **can't just be about logical necessity:**

(15) **A case of late pre-emption** (Hall 2004). Billy and Suzy both throw stones at a bottle on the wall. Their throws are on target. Suzy throws first, so her stone reaches the bottle before Billy's, and the bottle breaks.

a. Billy's throw is *pre-empted* by Suzy's throw

b. **intuition:** Suzy caused the bottle to break

c. **problem:** Suzy's throw was not *necessary*: if she hadn't thrown, Billy's stone would have hit and broken the bottle

- to define causation, we'd like to identify a set of conditions that predict all and only those cases where causation is intuitively present
- so those conditions should (collectively) be both necessary and sufficient for a causal relationship
- (15) shows that we can judge causation to be present without logical necessity holding between cause and effect
- so, **logical necessity of a cause for an effect is not a necessary condition for causation**

Dependency theories: logical dependency

The same scenario shows that causal dependency **can't just be about logical sufficiency:**

(16) **A case of late pre-emption** (Hall 2004). Billy and Suzy both throw stones at a bottle on the wall. Their throws are on target. Suzy throws first, so her stone reaches the bottle before Billy's, and the bottle breaks.

a. Billy's throw is *pre-empted* by Suzy's throw

Dependency theories: logical dependency

The same scenario shows that causal dependency **can't just be about logical sufficiency:**

(16) **A case of late pre-emption** (Hall 2004). Billy and Suzy both throw stones at a bottle on the wall. Their throws are on target. Suzy throws first, so her stone reaches the bottle before Billy's, and the bottle breaks.

a. Billy's throw is *pre-empted* by Suzy's throw

b. **intuition:** Billy did not cause the bottle to break

Dependency theories: logical dependency

The same scenario shows that causal dependency **can't just be about logical sufficiency:**

(16) **A case of late pre-emption** (Hall 2004). Billy and Suzy both throw stones at a bottle on the wall. Their throws are on target. Suzy throws first, so her stone reaches the bottle before Billy's, and the bottle breaks.

a. Billy's throw is *pre-empted* by Suzy's throw

b. **intuition:** Billy did not cause the bottle to break

c. **problem:** Billy's throw was sufficient for the bottle to break: if Suzy hadn't thrown, Billy would have broken the bottle

Dependency theories: logical dependency

The same scenario shows that causal dependency **can't just be about logical sufficiency:**

(16) **A case of late pre-emption** (Hall 2004). Billy and Suzy both throw stones at a bottle on the wall. Their throws are on target. Suzy throws first, so her stone reaches the bottle before Billy's, and the bottle breaks.

a. Billy's throw is *pre-empted* by Suzy's throw

b. **intuition:** Billy did not cause the bottle to break

c. **problem:** Billy's throw was sufficient for the bottle to break: if Suzy hadn't thrown, Billy would have broken the bottle

- since (16) shows that we fail to identify a causal dependency in the presence of logical sufficiency

Dependency theories: logical dependency

The same scenario shows that causal dependency **can't just be about logical sufficiency:**

(16) **A case of late pre-emption** (Hall 2004). Billy and Suzy both throw stones at a bottle on the wall. Their throws are on target. Suzy throws first, so her stone reaches the bottle before Billy's, and the bottle breaks.

a. Billy's throw is *pre-empted* by Suzy's throw

b. **intuition:** Billy did not cause the bottle to break

c. **problem:** Billy's throw was sufficient for the bottle to break: if Suzy hadn't thrown, Billy would have broken the bottle

- since (16) shows that we fail to identify a causal dependency in the presence of logical sufficiency
- . . . it shows that **logical sufficiency of a cause for an effect is not a sufficient condition for causation**

Dependency theories: logical dependency

The same scenario shows that causal dependency **can't just be about logical sufficiency:**

(16) **A case of late pre-emption** (Hall 2004). Billy and Suzy both throw stones at a bottle on the wall. Their throws are on target. Suzy throws first, so her stone reaches the bottle before Billy's, and the bottle breaks.

a. Billy's throw is *pre-empted* by Suzy's throw

b. **intuition:** Billy did not cause the bottle to break

c. **problem:** Billy's throw was sufficient for the bottle to break: if Suzy hadn't thrown, Billy would have broken the bottle

- since (16) shows that we fail to identify a causal dependency in the presence of logical sufficiency
- . . . it shows that **logical sufficiency of a cause for an effect is not a sufficient condition for causation**
- combinations of necessity and sufficiency run into similar problems

Dependency theories: logical dependency

The same scenario shows that causal dependency **can't just be about logical sufficiency:**

(16) **A case of late pre-emption** (Hall 2004). Billy and Suzy both throw stones at a bottle on the wall. Their throws are on target. Suzy throws first, so her stone reaches the bottle before Billy's, and the bottle breaks.

a. Billy's throw is *pre-empted* by Suzy's throw

b. **intuition:** Billy did not cause the bottle to break

c. **problem:** Billy's throw was sufficient for the bottle to break: if Suzy hadn't thrown, Billy would have broken the bottle

- since (16) shows that we fail to identify a causal dependency in the presence of logical sufficiency
- . . . it shows that **logical sufficiency of a cause for an effect is not a sufficient condition for causation**
- combinations of necessity and sufficiency run into similar problems

In (16), we have to reason about alternatives to reality: this suggests the relevance of **counterfactuals** to causation

Dependency theories of causation

2. Defining causal relationships in terms of counterfactuals

Dependency theories of causation

2. Defining causal relationships in terms of counterfactuals

The most prevalent linguistic theory of causation is David Lewis's (1973, 2000) **counterfactual** theory of causation (cf. Shibatani 1976, Neeleman & van de Koot 2012):

Dependency theories of causation

2. Defining causal relationships in terms of counterfactuals

The most prevalent linguistic theory of causation is David Lewis's (1973, 2000) **counterfactual** theory of causation (cf. Shibatani 1976, Neeleman & van de Koot 2012):

- **idea:** capitalize on the intuition that statements like (17) are closely correlated with counterfactual claims

  (17) The recession caused Jerry to lose his job.
  Counterfactual: If the recession had not happened, Jerry would not have lost his job.

Dependency theories of causation

2. Defining causal relationships in terms of counterfactuals

The most prevalent linguistic theory of causation is David Lewis's (1973, 2000) **counterfactual** theory of causation (cf. Shibatani 1976, Neeleman & van de Koot 2012):

- **idea:** capitalize on the intuition that statements like (17) are closely correlated with counterfactual claims

  (17) The recession caused Jerry to lose his job.
  Counterfactual: If the recession had not happened, Jerry would not have lost his job.

- the counterfactual relationship amounts to logical necessity

Dependency theories of causation

2. Defining causal relationships in terms of counterfactuals

The most prevalent linguistic theory of causation is David Lewis's (1973, 2000) **counterfactual** theory of causation (cf. Shibatani 1976, Neeleman & van de Koot 2012):

- **idea:** capitalize on the intuition that statements like (17) are closely correlated with counterfactual claims

  (17) The recession caused Jerry to lose his job.
  Counterfactual: If the recession had not happened, Jerry would not have lost his job.

- the counterfactual relationship amounts to logical necessity
- Lewis avoids the late pre-emption problem by using the counterfactual as a *diagnosis*, not a complete definition

Dependency theories of causation

2. Defining causal relationships in terms of counterfactuals

The most prevalent linguistic theory of causation is David Lewis's (1973, 2000) **counterfactual** theory of causation (cf. Shibatani 1976, Neeleman & van de Koot 2012):

- **idea:** capitalize on the intuition that statements like (17) are closely correlated with counterfactual claims

  (17) The recession caused Jerry to lose his job.
  Counterfactual: If the recession had not happened, Jerry would not have lost his job.

- the counterfactual relationship amounts to logical necessity
- Lewis avoids the late pre-emption problem by using the counterfactual as a *diagnosis*, not a complete definition
  - the distinction he makes between *causal dependence* and *nomic dependence* is a distinction between causal dependence and purely logical dependence

Lewis's counterfactual theory of causation

Prior to Lewis, philosophical theories of causation were focused on **regularities:** "...where all the objects, similar to the first, are followed by objects similar to the second ..."

Lewis's counterfactual theory of causation

Prior to Lewis, philosophical theories of causation were focused on **regularities:** "...where all the objects, similar to the first, are followed by objects similar to the second ..."

- Lewis's objections: regularity theories don't properly distinguish causes and effects (A), epiphenomena (B), and pre-empted causes (C)
  - (A) barometer readings and pressure change at the same time, so how can a regularity theory tell us which is the cause and which the effect
  - (B) changes in pressure change barometer readings and often cause rain, but there's no real causal link between rain and barometer readings
  - (C) similar to above: but Billy throws his stone only if Suzy misses (he didn't cause the bottle to break, but he would have if he had thrown)

Lewis's counterfactual theory of causation

Define causation in terms of counterfactuality, not regularity

- Hume: " . . . where, if the first object had not been, the second never had existed."
- Lewis, p.557: "We think of a cause as something that makes a difference, and the difference it makes must be a difference from what would have happened without it."

Lewis's counterfactual theory of causation

Define causation in terms of counterfactuality, not regularity

- Hume: " . . . where, if the first object had not been, the second never had existed."
- Lewis, p.557: "We think of a cause as something that makes a difference, and the difference it makes must be a difference from what would have happened without it."

The relevant notions for Lewis's theory:

Lewis's counterfactual theory of causation

Define causation in terms of counterfactuality, not regularity

- Hume: " . . . where, if the first object had not been, the second never had existed."
- Lewis, p.557: "We think of a cause as something that makes a difference, and the difference it makes must be a difference from what would have happened without it."

The relevant notions for Lewis's theory:

- possible worlds, which can differ with respect to historical facts and causal laws

Lewis's counterfactual theory of causation

Define causation in terms of counterfactuality, not regularity

- Hume: " . . . where, if the first object had not been, the second never had existed."
- Lewis, p.557: "We think of a cause as something that makes a difference, and the difference it makes must be a difference from what would have happened without it."

The relevant notions for Lewis's theory:

- possible worlds, which can differ with respect to historical facts and causal laws
- a **similarity relationship**, by which we can judge worlds as closer or farther away from one another
  - a world in which Suzy misses and Billy's stone breaks the bottle is 'closer to' the actual world than one in which Suzy's stone evaporates mid-flight

Lewis's counterfactual theory of causation

Define causation in terms of counterfactuality, not regularity

- Hume: " . . . where, if the first object had not been, the second never had existed."
- Lewis, p.557: "We think of a cause as something that makes a difference, and the difference it makes must be a difference from what would have happened without it."

The relevant notions for Lewis's theory:

- possible worlds, which can differ with respect to historical facts and causal laws
- a **similarity relationship**, by which we can judge worlds as closer or farther away from one another
  - a world in which Suzy misses and Billy's stone breaks the bottle is 'closer to' the actual world than one in which Suzy's stone evaporates mid-flight
- we judge the truth of a counterfactual by whether or not the most similar worlds (with some fact changed) support it

Lewis's counterfactual theory of causation

Connecting the evaluation of counterfactuals to causation:

Lewis's counterfactual theory of causation

Connecting the evaluation of counterfactuals to causation:

- the counterfactual dependence of C on A, written $A \mathbin{\Box\!\!\rightarrow} C$:

Lewis's counterfactual theory of causation

Connecting the evaluation of counterfactuals to causation:

- the counterfactual dependence of C on A, written $A \mathbin{\Box\!\!\rightarrow} C$:
  - if A holds in the evaluation world, then we have to go 'farther' to get to $\neg C$ worlds than we do to get to C worlds

Lewis's counterfactual theory of causation

Connecting the evaluation of counterfactuals to causation:

- the counterfactual dependence of C on A, written $A \boxed{\rightarrow} C$:
  - if A holds in the evaluation world, then we have to go 'farther' to get to $\neg C$ worlds than we do to get to C worlds
  - **NB:** we may already have to move away from the actual world to find an A world – similarity plays a role here as well

Lewis's counterfactual theory of causation

Connecting the evaluation of counterfactuals to causation:

- the counterfactual dependence of C on A, written $A \mathbin{\Box\!\!\rightarrow} C$:
  - if A holds in the evaluation world, then we have to go 'farther' to get to $\neg C$ worlds than we do to get to C worlds
  - **NB:** we may already have to move away from the actual world to find an A world – similarity plays a role here as well

Some problems:

Lewis's counterfactual theory of causation

Connecting the evaluation of counterfactuals to causation:

- the counterfactual dependence of C on A, written $A \mathbin{\Box\!\!\rightarrow} C$:
  - if A holds in the evaluation world, then we have to go 'farther' to get to $\neg C$ worlds than we do to get to C worlds
  - **NB:** we may already have to move away from the actual world to find an A world – similarity plays a role here as well

Some problems:

- what counts as 'closer': a world where Suzy misses entirely, or a world where Suzy hits the bottle but it doesn't break?

Lewis's counterfactual theory of causation

Connecting the evaluation of counterfactuals to causation:

- the counterfactual dependence of C on A, written $A \Box\!\!\rightarrow C$:
  - if A holds in the evaluation world, then we have to go 'farther' to get to $\neg C$ worlds than we do to get to C worlds
  - **NB:** we may already have to move away from the actual world to find an A world – similarity plays a role here as well

Some problems:

- what counts as 'closer': a world where Suzy misses entirely, or a world where Suzy hits the bottle but it doesn't break?
- we have intuitions about this, but there's a circularity: our intuitions seem to be about what preserves causal laws

Lewis's counterfactual theory of causation

Connecting the evaluation of counterfactuals to causation:

- the counterfactual dependence of C on A, written $A \mathbin{\Box\!\!\rightarrow} C$:
  - if A holds in the evaluation world, then we have to go 'farther' to get to $\neg C$ worlds than we do to get to C worlds
  - **NB:** we may already have to move away from the actual world to find an A world – similarity plays a role here as well

Some problems:

- what counts as 'closer': a world where Suzy misses entirely, or a world where Suzy hits the bottle but it doesn't break?
- we have intuitions about this, but there's a circularity: our intuitions seem to be about what preserves causal laws
- (more recently, linguists have analyzed counterfactuals in terms of causation instead of the other way around)

Lewis's counterfactual theory of causation

How does Lewis avoid the problems of regularity and necessity?

Lewis's counterfactual theory of causation

How does Lewis avoid the problems of regularity and necessity?

- **Claim:** causation is **transitive**, but counterfactual dependence ($\Box\!\!\rightarrow$) is not!

(18) **causation:**

a. the recession caused Jerry's company to lose money

b. losing money caused Jerry's company to fire Jerry

c. $\rightarrow$ the recession caused Jerry's company to fire him

(19) **counterfactuality:**

a. if Hoover had been born in Russia, he would have been a communist

b. if Hoover had been a communist, he would have been a traitor in the eyes of the US government

c. $\not\rightarrow$ if Hoover had been born in Russia, he would have been a traitor ...

Lewis's counterfactual theory of causation

How does Lewis avoid the problems of regularity and necessity?

- **Claim:** causation is **transitive**, but counterfactual dependence ($\Box\!\!\rightarrow$) is not!

  (18) **causation:**

  a. the recession caused Jerry's company to lose money

  b. losing money caused Jerry's company to fire Jerry

  c. $\rightarrow$ the recession caused Jerry's company to fire him

  (19) **counterfactuality:**

  a. if Hoover had been born in Russia, he would have been a communist

  b. if Hoover had been a communist, he would have been a traitor in the eyes of the US government

  c. $\not\rightarrow$ if Hoover had been born in Russia, he would have been a traitor ...

- counterfactual dependence entails causation, but not vice versa

Lewis's counterfactual theory of causation

How does Lewis avoid the problems of regularity and necessity?

- **Claim:** causation is **transitive**, but counterfactual dependence ($\Box\!\!\rightarrow$) is not!

  (18) **causation:**

  a. the recession caused Jerry's company to lose money

  b. losing money caused Jerry's company to fire Jerry

  c. $\rightarrow$ the recession caused Jerry's company to fire him

  (19) **counterfactuality:**

  a. if Hoover had been born in Russia, he would have been a communist

  b. if Hoover had been a communist, he would have been a traitor in the eyes of the US government

  c. $\not\rightarrow$ if Hoover had been born in Russia, he would have been a traitor ...

- counterfactual dependence entails causation, but not vice versa
- counterfactuality is a sufficient but not a necessary condition

Lewis's counterfactual theory of causation

An example involving **early pre-emption**:

(19) Billy and Suzy are both ready with stones to throw at a bottle on the wall. Suzy throws her stone, breaking the bottle, and Billy stands by as backup – he'll throw his stone if she doesn't throw hers (or if she misses)

a. Again, Billy's throw is pre-empted

b. **intuition:** Suzy caused the breaking

c. **problem:** there's no counterfactual dependence between Suzy's throw and the bottle breaking

Lewis's counterfactual theory of causation

An example involving **early pre-emption**:

(19) Billy and Suzy are both ready with stones to throw at a bottle on the wall. Suzy throws her stone, breaking the bottle, and Billy stands by as backup – he'll throw his stone if she doesn't throw hers (or if she misses)

a. Again, Billy's throw is pre-empted

b. **intuition:** Suzy caused the breaking

c. **problem:** there's no counterfactual dependence between Suzy's throw and the bottle breaking

Lewis's solution: **stepwise counterfactual dependency**

Lewis's counterfactual theory of causation

An example involving **early pre-emption**:

(19) Billy and Suzy are both ready with stones to throw at a bottle on the wall. Suzy throws her stone, breaking the bottle, and Billy stands by as backup – he'll throw his stone if she doesn't throw hers (or if she misses)

a. Again, Billy's throw is pre-empted

b. **intuition:** Suzy caused the breaking

c. **problem:** there's no counterfactual dependence between Suzy's throw and the bottle breaking

Lewis's solution: **stepwise counterfactual dependency**

- we need to consider causation in terms of causal chains (familiar!)

Lewis's counterfactual theory of causation

An example involving **early pre-emption**:

(19) Billy and Suzy are both ready with stones to throw at a bottle on the wall. Suzy throws her stone, breaking the bottle, and Billy stands by as backup – he'll throw his stone if she doesn't throw hers (or if she misses)

a. Again, Billy's throw is pre-empted

b. **intuition:** Suzy caused the breaking

c. **problem:** there's no counterfactual dependence between Suzy's throw and the bottle breaking

Lewis's solution: **stepwise counterfactual dependency**

- we need to consider causation in terms of causal chains (familiar!)
- Suzy's throw $\Box\!\!\rightarrow$ Suzy's stone flies through the air $\rightarrow$ **cause**
- Suzy's stone flies $\Box\!\!\rightarrow$ the bottle breaks $\rightarrow$ **cause**

Lewis's counterfactual theory of causation

An example involving **early pre-emption**:

(19) Billy and Suzy are both ready with stones to throw at a bottle on the wall. Suzy throws her stone, breaking the bottle, and Billy stands by as backup – he'll throw his stone if she doesn't throw hers (or if she misses)

a. Again, Billy's throw is pre-empted

b. **intuition:** Suzy caused the breaking

c. **problem:** there's no counterfactual dependence between Suzy's throw and the bottle breaking

Lewis's solution: **stepwise counterfactual dependency**

- we need to consider causation in terms of causal chains (familiar!)
- Suzy's throw $\Box\!\!\rightarrow$ Suzy's stone flies through the air $\rightarrow$ **cause**
- Suzy's stone flies $\Box\!\!\rightarrow$ the bottle breaks $\rightarrow$ **cause**
- the transitivity of causation then gives us causal dependence between Suzy throwing and the bottle breaking, *even though we don't have counterfactual dependence*

Lewis's counterfactual theory of causation

The problem of **effects** and spurious causes (p.565):

- Suppose: c **causes** e, but not the reverse
- Suppose: given the laws and circumstances, c **could not but cause** e
- problem: it follows that if e had not occurred, c would not have either $?\rightarrow$ **cause**

Lewis's counterfactual theory of causation

The problem of **effects** and spurious causes (p.565):

- Suppose: c **causes** e, but not the reverse
- Suppose: given the laws and circumstances, c **could not but cause** e
- problem: it follows that if e had not occurred, c would not have either $?\rightarrow$ **cause**

The problem of **epiphenomena** and spurious causes (p.566):

- Suppose: c **causes** e **and** f (subsequently)
- Suppose: given the laws and circumstances
  - (i) c **could not but cause** e
  - (ii) f **could not have been caused by anything but** c
- problem: if e had not occurred, we know that c could not have occurred, and thus f also does not hold
  ...which makes it seem like e **causes** f
- (e is epiphenomenal to the causal relationship between c and e)

Lewis's counterfactual theory of causation

Lewis's solution to the effect and epiphenomena problems (p.566):

> *"The proper solution to both problems, I think, is to flatly deny the counterfactuals that cause the trouble."*

Lewis's counterfactual theory of causation

Lewis's solution to the effect and epiphenomena problems (p.566):

> *"The proper solution to both problems, I think, is to flatly deny the counterfactuals that cause the trouble."*

- **Claim:** if e had been absent, it is not that c would have been absent (and f as well)
- this comes from the similarity relationship:

Lewis's counterfactual theory of causation

Lewis's solution to the effect and epiphenomena problems (p.566):

> *"The proper solution to both problems, I think, is to flatly deny the counterfactuals that cause the trouble."*

- **Claim:** if e had been absent, it is not that c would have been absent (and f as well)
- this comes from the similarity relationship:
  - worlds where c occurs and fails to cause e are *closer* than worlds where c also does not occur

Lewis's counterfactual theory of causation

Lewis's solution to the effect and epiphenomena problems (p.566):

> *"The proper solution to both problems, I think, is to flatly deny the counterfactuals that cause the trouble."*

- **Claim:** if e had been absent, it is not that c would have been absent (and f as well)
- this comes from the similarity relationship:
  - worlds where c occurs and fails to cause e are *closer* than worlds where c also does not occur
  - we want to preserve the facts of the world as much as possible in evaluating counterfactuals

> *"To get rid of an actual event e with the least over-all departure from actuality, it will normally be best not to diverge at all from the actual course of events until just before the time of e."*

Problem? doesn't this rely on intuitions about facts vs. causal laws?

Lewis's theory of causation: problems

Late pre-emption is still an issue:

(15) **A case of late pre-emption** (Hall 2004). Billy and Suzy both throw stones at a bottle on the wall. Their throws are on target. Suzy throws first, so her stone reaches the bottle before Billy's, and the bottle breaks.

Lewis's theory of causation: problems

Late pre-emption is still an issue:

(15) **A case of late pre-emption** (Hall 2004). Billy and Suzy both throw stones at a bottle on the wall. Their throws are on target. Suzy throws first, so her stone reaches the bottle before Billy's, and the bottle breaks.

- there's no stepwise counterfactual dependency between Suzy and the bottle breaking (because Billy throws regardless of Suzy)

Lewis's theory of causation: problems

Late pre-emption is still an issue:

(15) **A case of late pre-emption** (Hall 2004). Billy and Suzy both throw stones at a bottle on the wall. Their throws are on target. Suzy throws first, so her stone reaches the bottle before Billy's, and the bottle breaks.

- there's no stepwise counterfactual dependency between Suzy and the bottle breaking (because Billy throws regardless of Suzy)
- counterfactual dependence is still not necessary for causation

Lewis's theory of causation: problems

Late pre-emption is still an issue:

(15) **A case of late pre-emption** (Hall 2004). Billy and Suzy both throw stones at a bottle on the wall. Their throws are on target. Suzy throws first, so her stone reaches the bottle before Billy's, and the bottle breaks.

- there's no stepwise counterfactual dependency between Suzy and the bottle breaking (because Billy throws regardless of Suzy)
- counterfactual dependence is still not necessary for causation

Counterfactual dependence is also, contra Lewis, not **sufficient** for causation (i.e., $\Box\!\!\rightarrow$ does NOT imply causation)

(20) **The assassin scenario** (Hall 2000). An assassin places a bomb under your desk, causing you to find it, which causes you to remove it. Removing it causes (ensures) your survival.

Lewis's theory of causation: problems

Late pre-emption is still an issue:

(15) **A case of late pre-emption** (Hall 2004). Billy and Suzy both throw stones at a bottle on the wall. Their throws are on target. Suzy throws first, so her stone reaches the bottle before Billy's, and the bottle breaks.

- there's no stepwise counterfactual dependency between Suzy and the bottle breaking (because Billy throws regardless of Suzy)
- counterfactual dependence is still not necessary for causation

Counterfactual dependence is also, contra Lewis, not **sufficient** for causation (i.e., $\Box\!\!\rightarrow$ does NOT imply causation)

(20) **The assassin scenario** (Hall 2000). An assassin places a bomb under your desk, causing you to find it, which causes you to remove it. Removing it causes (ensures) your survival.

- there IS stepwise counterfactual dependence along the whole chain

Lewis's theory of causation: problems

Late pre-emption is still an issue:

(15) **A case of late pre-emption** (Hall 2004). Billy and Suzy both throw stones at a bottle on the wall. Their throws are on target. Suzy throws first, so her stone reaches the bottle before Billy's, and the bottle breaks.

- there's no stepwise counterfactual dependency between Suzy and the bottle breaking (because Billy throws regardless of Suzy)
- counterfactual dependence is still not necessary for causation

Counterfactual dependence is also, contra Lewis, not **sufficient** for causation (i.e., $\Box\!\!\rightarrow$ does NOT imply causation)

(20) **The assassin scenario** (Hall 2000). An assassin places a bomb under your desk, causing you to find it, which causes you to remove it. Removing it causes (ensures) your survival.

- there IS stepwise counterfactual dependence along the whole chain
- **intuition:** it is not the case that the assassin caused your survival

Lewis's theory of causation: problems

Late pre-emption is still an issue:

(15) **A case of late pre-emption** (Hall 2004). Billy and Suzy both throw stones at a bottle on the wall. Their throws are on target. Suzy throws first, so her stone reaches the bottle before Billy's, and the bottle breaks.

- there's no stepwise counterfactual dependency between Suzy and the bottle breaking (because Billy throws regardless of Suzy)
- counterfactual dependence is still not necessary for causation

Counterfactual dependence is also, contra Lewis, not **sufficient** for causation (i.e., $\Box\!\!\rightarrow$ does NOT imply causation)

(20) **The assassin scenario** (Hall 2000). An assassin places a bomb under your desk, causing you to find it, which causes you to remove it. Removing it causes (ensures) your survival.

- there IS stepwise counterfactual dependence along the whole chain
- **intuition:** it is not the case that the assassin caused your survival
- **more broadly:** causation may not be transitive either

Dependency theories of causation

Upshot: intuitions about a connection between causation and counterfactuals are important, but there's more going on!

Dependency theories of causation

Upshot: intuitions about a connection between causation and counterfactuals are important, but there's more going on!

3. Defining causal relationships in terms of probability

Dependency theories of causation

Upshot: intuitions about a connection between causation and counterfactuals are important, but there's more going on!

3. Defining causal relationships in terms of probability

- **main idea:** certain events *raise* the probability of others
  - event C causes (or is causally involved) in event E if C's occurrence makes E more likely
  - $\Pr(E|C) > \Pr(E|\neg C)$ (Cheng & Novick 1992)

(21) You're more likely to get cancer if you smoke than if you do not smoke

a. $\Pr(\text{cancer} \mid \text{smoker}) > \Pr(\text{cancer} \mid \text{non-smoker})$

b. Smoking **causes** cancer

Dependency theories of causation

Upshot: intuitions about a connection between causation and counterfactuals are important, but there's more going on!

3. Defining causal relationships in terms of probability

- **main idea:** certain events *raise* the probability of others
  - event C causes (or is causally involved) in event E if C's occurrence makes E more likely
  - $\Pr(E|C) > \Pr(E|\neg C)$ (Cheng & Novick 1992)

(21) You're more likely to get cancer if you smoke than if you do not smoke

a. $\Pr(\text{cancer} \mid \text{smoker}) > \Pr(\text{cancer} \mid \text{non-smoker})$

b. Smoking **causes** cancer

- probability theories build on Hume's (first) regularity theory

Probability raising theories of causation

Correlation is not causation!

- events can be probabilistically correlated because C is a cause of E, because E is a cause of C, or because both C and E are caused by some other event, F
- these are the *effect* and *epiphenomena* problems, again

Probability raising theories of causation

Correlation is not causation!

- events can be probabilistically correlated because C is a cause of E, because E is a cause of C, or because both C and E are caused by some other event, F
- these are the *effect* and *epiphenomena* problems, again
- probability-raising is not **sufficient** for causation

Probability raising theories of causation

Correlation is not causation!

- events can be probabilistically correlated because C is a cause of E, because E is a cause of C, or because both C and E are caused by some other event, F
- these are the *effect* and *epiphenomena* problems, again
- probability-raising is not **sufficient** for causation

Probability raising is also not **necessary** for causation:

- early pre-emption:
  - suppose Suzy doesn't have great aim, and only has a 25% chance of hitting the bottle
  - Billy has better aim, with a 70% chance

Probability raising theories of causation

Correlation is not causation!

- events can be probabilistically correlated because C is a cause of E, because E is a cause of C, or because both C and E are caused by some other event, F
- these are the *effect* and *epiphenomena* problems, again
- probability-raising is not **sufficient** for causation

Probability raising is also not **necessary** for causation:

- early pre-emption:
  - suppose Suzy doesn't have great aim, and only has a 25% chance of hitting the bottle
  - Billy has better aim, with a 70% chance
  - if Suzy doesn't throw, Billy will throw, increasing the likelihood that the bottle will break

Probability raising theories of causation

Correlation is not causation!

- events can be probabilistically correlated because C is a cause of E, because E is a cause of C, or because both C and E are caused by some other event, F
- these are the *effect* and *epiphenomena* problems, again
- probability-raising is not **sufficient** for causation

Probability raising is also not **necessary** for causation:

- early pre-emption:
  - suppose Suzy doesn't have great aim, and only has a 25% chance of hitting the bottle
  - Billy has better aim, with a 70% chance
  - if Suzy doesn't throw, Billy will throw, increasing the likelihood that the bottle will break
  - so Suzy's throw actually REDUCES the likelihood that the bottle breaks, but we still feel that Suzy caused the breaking

Dependency theories of causation

4. Causal network models

Dependency theories of causation

4. Causal network models

The basic ideas:

- events are connected in a network, represented as a **directed acyclic graph**
- 'arrows' just indicate a causal link, which is unanalyzed
- probability distributions or **structural equations** tell us how the various events determine one another
- only 'upstream' events can influence downstream ones

Dependency theories of causation

4. Causal network models

The basic ideas:

- events are connected in a network, represented as a **directed acyclic graph**
- 'arrows' just indicate a causal link, which is unanalyzed
- probability distributions or **structural equations** tell us how the various events determine one another
- only 'upstream' events can influence downstream ones

Some positives:

- having a network instead of just a chain lets us make sense of complex relationships (epiphenomena, late pre-emption)
- Lewis's intuitions about preserving facts up until the time of an effect can be related to 'intervention' tests

Causal network models

The late pre-emption example in a network model:

SH := ST
BH := BT
BS := SH v BH
BH := BT and ∼S

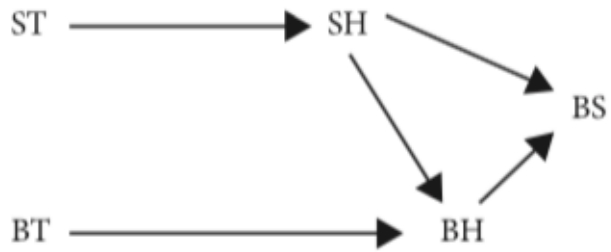

FIG. 2.2 An example of late pre-emption in terms of a direct graph (Hitchcock 2010)

Causal network models

The late pre-emption example in a network model:

SH := ST
BH := BT
BS := SH v BH
BH := BT and ∼S

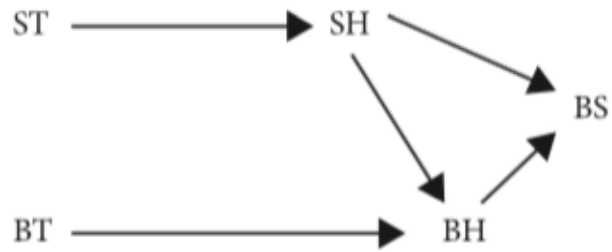

FIG. 2.2 An example of late pre-emption in terms of a direct graph (Hitchcock 2010)

- the network model allows us to include the influence that Suzy's hit (SH) has on Billy's hit (BH)

Causal network models

The late pre-emption example in a network model:

SH := ST
BH := BT
BS := SH v BH
BH := BT and ∼S

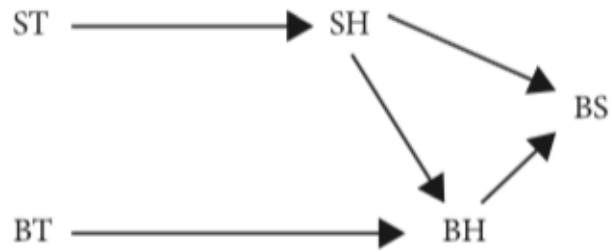

FIG. 2.2 An example of late pre-emption in terms of a direct graph (Hitchcock 2010)

- the network model allows us to include the influence that Suzy's hit (SH) has on Billy's hit (BH)
- Suzy's throw (ST) only influences Suzy's hit (SH), but the value of SH determines the value of BH (assuming we don't change the facts about whether or not Billy throws)

Causal network models

The late pre-emption example in a network model:

SH := ST
BH := BT
BS := SH v BH
BH := BT and ∼S

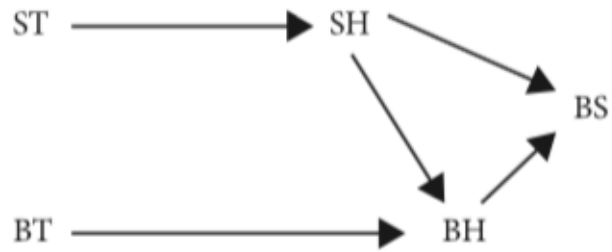

FIG. 2.2 An example of late pre-emption in terms of a direct graph (Hitchcock 2010)

- the network model allows us to include the influence that Suzy's hit (SH) has on Billy's hit (BH)
- Suzy's throw (ST) only influences Suzy's hit (SH), but the value of SH determines the value of BH (assuming we don't change the facts about whether or not Billy throws)
- this explains why we think Suzy's throw is the cause and Billy's throw is not: there's an asymmetry in the arrows in the model

Causal network models

The late pre-emption example in a network model:

SH := ST
BH := BT
BS := SH v BH
BH := BT and ∼S

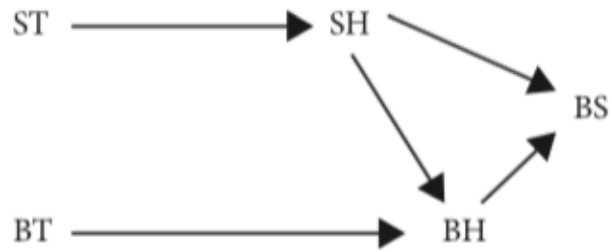

FIG. 2.2 An example of late pre-emption in terms of a direct graph (Hitchcock 2010)

- the network model allows us to include the influence that Suzy's hit (SH) has on Billy's hit (BH)
- Suzy's throw (ST) only influences Suzy's hit (SH), but the value of SH determines the value of BH (assuming we don't change the facts about whether or not Billy throws)
- this explains why we think Suzy's throw is the cause and Billy's throw is not: there's an asymmetry in the arrows in the model

(We'll spend more time on network models in the next two weeks)

Production theories of causation

Production theories are motivated by the intuition "that something more than correlation or regularity is involved in causation." (Copley & Wolff, p.23)

Production theories of causation

Production theories are motivated by the intuition "that something more than correlation or regularity is involved in causation." (Copley & Wolff, p.23)

- Back to Hume:
  - causation is often associated with (physical or mental) force, and (kinetic or potential) energy
  - for Hume, these associations are the way in which we perceive the thing that we call causation
  - production theories reverse this idea: causation is a name for certain kinds of transfers of force and energy (correlation and regularity are epiphenomena, p.25)

Production theories of causation

Production theories are motivated by the intuition "that something more than correlation or regularity is involved in causation." (Copley & Wolff, p.23)

- Back to Hume:
  - causation is often associated with (physical or mental) force, and (kinetic or potential) energy
  - for Hume, these associations are the way in which we perceive the thing that we call causation
  - production theories reverse this idea: causation is a name for certain kinds of transfers of force and energy (correlation and regularity are epiphenomena, p.25)
- properties that a production theory might explain:
  - temporal orderings between cause and effect (cf. Shibatani 1976): dependency theories alone cannot explain the relevance of temporal order
  - physical links: why is it difficult for us to accept "spooky action at a distance"?

Production theories of causation

Production theories are motivated by the intuition "that something more than correlation or regularity is involved in causation." (Copley & Wolff, p.23)

- Back to Hume:
  - causation is often associated with (physical or mental) force, and (kinetic or potential) energy
  - for Hume, these associations are the way in which we perceive the thing that we call causation
  - production theories reverse this idea: causation is a name for certain kinds of transfers of force and energy (correlation and regularity are epiphenomena, p.25)
- properties that a production theory might explain:
  - temporal orderings between cause and effect (cf. Shibatani 1976): dependency theories alone cannot explain the relevance of temporal order
  - physical links: why is it difficult for us to accept "spooky action at a distance"?
- **upshot**: causal links have real, tangible correlates

Production theories of causation

1. Transmission theories

- **mark transmission** (Salmon 1984): causal processes propagate some kind of modification, or mark – something that starts in one place and has a reflection or realization elsewhere at the end of the process
  - example: shining a light through a red filter causes a red mark to appear on a screen

Production theories of causation

1. Transmission theories

- **mark transmission** (Salmon 1984): causal processes propagate some kind of modification, or mark – something that starts in one place and has a reflection or realization elsewhere at the end of the process
  - example: shining a light through a red filter causes a red mark to appear on a screen
- **transmission of conserved properties** (Dowe 2000, Kistler 2006):
  - physical laws tells us that heat and momentum are conserved quantities
  - causal processes transmit these properties from one object to another, in keeping with the conservation laws
  - example: a billiard ball can only start to move if it gets an input of kinetic energy from something else (a causer)

Production theories of causation

2. Force dynamics

Basic ideas:

- (physical) objects are associated with forces: inertia, momentum, dispositions
- causation/causal processes have to do with the convergence of forces from various objects in an interaction
- to evaluate causation, we look at the difference between starting and ending force vectors
- Wolff (2003): the 'tendency' of the causer, causee, etc

Production theories of causation

2. Force dynamics

Basic ideas:

- (physical) objects are associated with forces: inertia, momentum, dispositions
- causation/causal processes have to do with the convergence of forces from various objects in an interaction
- to evaluate causation, we look at the difference between starting and ending force vectors
- Wolff (2003): the 'tendency' of the causer, causee, etc

(We'll look more closely at modern versions of force dynamics and its applications in January)

Production theories of causation

Positives for production theories:

Production theories of causation

Positives for production theories:

- by looking at 'resultant' forces, we can explain the similarity between causing and prevention, causing and enabling (causation by omission), causation and maintenance (N&vdK)

Production theories of causation

Positives for production theories:

- by looking at 'resultant' forces, we can explain the similarity between causing and prevention, causing and enabling (causation by omission), causation and maintenance (N&vdK)
- we can do many of the same things as network models:
  - pre-emption: Suzy is the cause of the bottle-breaking (in both pre-emption scenarios), and Billy is not, because Suzy's force/energy is *actually* transmitted to the bottle

Production theories of causation

Positives for production theories:

- by looking at 'resultant' forces, we can explain the similarity between causing and prevention, causing and enabling (causation by omission), causation and maintenance (N&vdK)
- we can do many of the same things as network models:
  - pre-emption: Suzy is the cause of the bottle-breaking (in both pre-emption scenarios), and Billy is not, because Suzy's force/energy is *actually* transmitted to the bottle

But challenges remain:

- we often (correctly) infer causation in cases where we can't 'track' the transmitted quantity
- how can you track the transmission of mental (as opposed to physical) influence?

(25) **Persuasion.** John talked the children into dancing.

a. **intuition:** John caused the dancing

b. **problem:** no physical force or momentum transmitted

Language and causation

If the way we talk about events is shaped by our causal cognition:

Language and causation

If the way we talk about events is shaped by our causal cognition:

- it's not just that a good theory or model of causation should explain or predict linguistic judgements

Language and causation

If the way we talk about events is shaped by our causal cognition:

- it's not just that a good theory or model of causation should explain or predict linguistic judgements
- linguistic phenomena are a way to investigate the cognitive parameters of causation

Language and causation

If the way we talk about events is shaped by our causal cognition:

- it's not just that a good theory or model of causation should explain or predict linguistic judgements
- linguistic phenomena are a way to investigate the cognitive parameters of causation
- e.g., if linguistic phenomena distinguish between cases of early and late pre-emption, it's clear that the way we model (mentally represent) causation must also be able to distinguish these things

Language and causation

Our cognitive model cares about:

Language and causation

Our cognitive model cares about:

- **agentive** causation ('less marked', DeLancey 1984)
  - by looking at the edge cases, we can narrow in on what the underlyingly relevant notion is (C&W: **dispositions**)
  - these kinds of properties might be well represented in a force dynamics model

Language and causation

Our cognitive model cares about:

- **agentive** causation ('less marked', DeLancey 1984)
  - by looking at the edge cases, we can narrow in on what the underlyingly relevant notion is (C&W: **dispositions**)
  - these kinds of properties might be well represented in a force dynamics model
  - *question:* do we distinguish between latent properties (e.g., potential energy) and realized properties (e.g., momentum)?

Language and causation

Our cognitive model cares about:

- **agentive** causation ('less marked', DeLancey 1984)
  - by looking at the edge cases, we can narrow in on what the underlyingly relevant notion is (C&W: **dispositions**)
  - these kinds of properties might be well represented in a force dynamics model
  - *question:* do we distinguish between latent properties (e.g., potential energy) and realized properties (e.g., momentum)?
- a related question: **defeasibility**

(22) a. Billy insulted Suzy, but she wasn't offended.

b. The low salary offer insulted Suzy, #but she wasn't offended

Language and causation

Our cognitive model cares about:

- **agentive** causation ('less marked', DeLancey 1984)
  - by looking at the edge cases, we can narrow in on what the underlyingly relevant notion is (C&W: **dispositions**)
  - these kinds of properties might be well represented in a force dynamics model
  - *question:* do we distinguish between latent properties (e.g., potential energy) and realized properties (e.g., momentum)?
- a related question: **defeasibility**

(22) a. Billy insulted Suzy, but she wasn't offended.

b. The low salary offer insulted Suzy, #but she wasn't offended

- if *insult* is a causal verb, what kind of theory best accounts for the difference in (22)?

Language and causation

Our cognitive model cares about:

- **agentive** causation ('less marked', DeLancey 1984)
  - by looking at the edge cases, we can narrow in on what the underlyingly relevant notion is (C&W: **dispositions**)
  - these kinds of properties might be well represented in a force dynamics model
  - *question:* do we distinguish between latent properties (e.g., potential energy) and realized properties (e.g., momentum)?
- a related question: **defeasibility**

(22) a. Billy insulted Suzy, but she wasn't offended.

b. The low salary offer insulted Suzy, #but she wasn't offended

  - if *insult* is a causal verb, what kind of theory best accounts for the difference in (22)?
- direct vs. indirect causal chains

Language and causation

Our cognitive model cares about:

- **agentive** causation ('less marked', DeLancey 1984)
  - by looking at the edge cases, we can narrow in on what the underlyingly relevant notion is (C&W: **dispositions**)
  - these kinds of properties might be well represented in a force dynamics model
  - *question:* do we distinguish between latent properties (e.g., potential energy) and realized properties (e.g., momentum)?
- a related question: **defeasibility**

(22) a. Billy insulted Suzy, but she wasn't offended.

b. The low salary offer insulted Suzy, #but she wasn't offended

  - if *insult* is a causal verb, what kind of theory best accounts for the difference in (22)?
- direct vs. indirect causal chains
  - what counts as a link in a network vs. a probability-raising model vs. the force dynamics?

Overview and road map

Today:

Overview and road map

Today:

- in a Davidsonian event-based semantics, basic sentential descriptions invoke our causal understanding of the world

Overview and road map

Today:

- in a Davidsonian event-based semantics, basic sentential descriptions invoke our causal understanding of the world
- part of what has made the data so far difficult is that it's hard to get a handle on what causation *is*

Overview and road map

Today:

- in a Davidsonian event-based semantics, basic sentential descriptions invoke our causal understanding of the world
- part of what has made the data so far difficult is that it's hard to get a handle on what causation *is*
- instead of taking causation as a black box, and then trying to encode features like $\pm$ VOL in our linguistic/lexical representations

Overview and road map

Today:

- in a Davidsonian event-based semantics, basic sentential descriptions invoke our causal understanding of the world
- part of what has made the data so far difficult is that it's hard to get a handle on what causation *is*
- instead of taking causation as a black box, and then trying to encode features like $\pm$ VOL in our linguistic/lexical representations
- . . . the notion of causation is *constituted* by the features that language cares about

Overview and road map

Today:

- in a Davidsonian event-based semantics, basic sentential descriptions invoke our causal understanding of the world
- part of what has made the data so far difficult is that it's hard to get a handle on what causation *is*
- instead of taking causation as a black box, and then trying to encode features like $\pm$ VOL in our linguistic/lexical representations
- . . . the notion of causation is *constituted* by the features that language cares about
- so, by building a model which predicts linguistic data, we find out how we represent causation
- **caveat:** this may or may not reflect reality (e.g. quantum mechanics)

Overview and road map

Today:

- in a Davidsonian event-based semantics, basic sentential descriptions invoke our causal understanding of the world
- part of what has made the data so far difficult is that it's hard to get a handle on what causation *is*
- instead of taking causation as a black box, and then trying to encode features like $\pm$ VOL in our linguistic/lexical representations
- . . . the notion of causation is *constituted* by the features that language cares about
- so, by building a model which predicts linguistic data, we find out how we represent causation
- **caveat:** this may or may not reflect reality (e.g. quantum mechanics)
- different models may be good at different things

Overview and road map

Looking ahead:

Overview and road map

Looking ahead:

- next week: we'll take a look at how network models are formally defined
  - . . . and then see some applications of these models in explaining periphrastic and lexical causatives, including 'directness' and 'agency' contrasts

Overview and road map

Looking ahead:

- next week: we'll take a look at how network models are formally defined
  - . . . and then see some applications of these models in explaining periphrastic and lexical causatives, including 'directness' and 'agency' contrasts
- in January: we'll revisit the force dynamics
  - . . . and look at how cases of defeasible causation are treated with forces

Overview and road map

Looking ahead:

- next week: we'll take a look at how network models are formally defined
  - . . . and then see some applications of these models in explaining periphrastic and lexical causatives, including 'directness' and 'agency' contrasts
- in January: we'll revisit the force dynamics
  - . . . and look at how cases of defeasible causation are treated with forces
- finally: we'll look outside of overtly causal language, and see how causal modeling approaches can help with other lexical semantic questions (counterfactuals, modals, implicatives)

Overview and road map

If you're interested in doing the AP, now is a good time to start thinking about topics!

- a good empirical topic would be to gather some data on a causal construction (in German or another language)
- use the data to argue for or against a particular kind of model
- or to extend one of the applications we look at