Logics of Rational Agency

Lecture 1

Eric Pacuit

Tilburg Institute for Logic and Philosophy of Science
Tilburg Univeristy
ai.stanford.edu/~epacuit

July 27, 2009

Lecture 1: Introduction, Motivation and Background

Lecture 2: Basic Ingredients for a Logic of Rational Agency

Lecture 3: Logics of Rational Agency and Social Interaction, Part I

Lecture 4: Logics of Rational Agency and Social Interaction, Part II

Lecture 5: Conclusions and General Issues

Lecture 1: Introduction, Motivation and Background

Lecture 2: Basic Ingredients for a Logic of Rational Agency

Lecture 3: Logics of Rational Agency and Social Interaction, Part I

Lecture 4: Logics of Rational Agency and Social Interaction, Part II

Lecture 5: Conclusions and General Issues

Lecture 1: Introduction, Motivation and Background

Lecture 2: Basic Ingredients for a Logic of Rational Agency

Lecture 3: Logics of Rational Agency and Social Interaction, Part I

Lecture 4: Logics of Rational Agency and Social Interaction, Part II

Lecture 5: Conclusions and General Issues

Lecture 1: Introduction, Motivation and Background

Lecture 2: Basic Ingredients for a Logic of Rational Agency

Lecture 3: Logics of Rational Agency and Social Interaction, Part I

Lecture 4: Logics of Rational Agency and Social Interaction, Part II

Lecture 5: Conclusions and General Issues

Lecture 1: Introduction, Motivation and Background

Lecture 2: Basic Ingredients for a Logic of Rational Agency

Lecture 3: Logics of Rational Agency and Social Interaction, Part I

Lecture 4: Logics of Rational Agency and Social Interaction, Part II

Lecture 5: Conclusions and General Issues

Lecture 1: Introduction, Motivation and Background

Lecture 2: Basic Ingredients for a Logic of Rational Agency

Lecture 3: Logics of Rational Agency and Social Interaction, Part I

Lecture 4: Logics of Rational Agency and Social Interaction, Part II

Lecture 5: Conclusions and General Issues

Course Website

`http://ai.stanford.edu/~epacuit/classes/esslli/log-ratagency.html`

Reading Material

✓ Pointers to literature on the website

Concerning Modal Logic

✓ *Modal Logic* by P. Blackburn, M. de Rijke and Y. Venema.

✓ *Modal Logic for Open Minds* by Johan van Benthem (published soon)

We are interested in reasoning about rational agents interacting in *social* situations.

We are interested in reasoning about rational agents interacting in *social* situations.

- Philosophy (social philosophy, epistemology)
- Game Theory
- Social Choice Theory
- AI (multiagent systems)

We are interested in reasoning about rational agents interacting in *social* situations.

What is a rational agent?

- maximize expected utility (instrumentally rational)
- react to observations
- revise beliefs when learning a *surprising* piece of information
- understand higher-order information
- plans for the future
- ????

We are interested in reasoning about rational agents interacting in *social* situations.

There is a jungle of formal systems!

- logics of informational attitudes (knowledge, beliefs, certainty)
- logics of action & agency
- temporal logics/dynamic logics
- logics of motivational attitudes (preferences, intentions)

(*Not to mention various game-theoretic/social choice models and logical languages for reasoning about them*)

We are interested in reasoning about rational agents interacting in *social* situations.

There is a jungle of formal systems!

- How do we compare different logical systems studying the same phenomena?
- How *complex* is it to reason about rational agents?
- (How) should we *merge* the various logical systems?
- What do the logical frameworks contribute to the discussion on rational agency?

and logical languages for reasoning about them)

We are interested in reasoning about rational agents interacting in *social* situations.

- playing a card game
- having a conversation
- executing a *social procedure*
- ....

Goal: incorporate/extend existing game-theoretic/social choice analyses

Formally, a game is described by its strategy sets and payoff functions.

Formally, a game is described by its strategy sets and payoff functions. But in real life, may other parameters are relevant; there is a lot more going on. Situations that substantively are vastly different may nevertheless correspond to precisely the same strategic game.

Formally, a game is described by its strategy sets and payoff functions. But in real life, may other parameters are relevant; there is a lot more going on. Situations that substantively are vastly different may nevertheless correspond to precisely the same strategic game. For example, in a parliamentary democracy with three parties, the winning coalitions are the same whether the parties hold a third of the seats, or, say, 49%, 39%, and 12 % respectively.

Formally, a game is described by its strategy sets and payoff functions. But in real life, may other parameters are relevant; there is a lot more going on. Situations that substantively are vastly different may nevertheless correspond to precisely the same strategic game. For example, in a parliamentary democracy with three parties, the winning coalitions are the same whether the parties hold a third of the seats, or, say, 49%, 39%, and 12 % respectively. But the political situations are quite different.

Formally, a game is described by its strategy sets and payoff functions. But in real life, may other parameters are relevant; there is a lot more going on. Situations that substantively are vastly different may nevertheless correspond to precisely the same strategic game. For example, in a parliamentary democracy with three parties, the winning coalitions are the same whether the parties hold a third of the seats, or, say, 49%, 39%, and 12 % respectively. But the political situations are quite different. The difference lies in the attitudes of the players, in their expectations about each other, in custom, and in history, though the rules of the game do not distinguish between the two situations.

R. Aumann and J. H. Dreze. *Rational Expectation in Games*. American Economic Review (2008).

Adjusted Winner

Adjusted winner (AW) is an algorithm for dividing n divisible goods among two people (invented by Steven Brams and Alan Taylor).

For more information see

- *Fair Division: From cake-cutting to dispute resolution* by Brams and Taylor, 1998
- *The Win-Win Solution* by Brams and Taylor, 2000
- www.nyu.edu/projects/adjustedwinner

Adjusted Winner: Example

Adjusted Winner: Example

Suppose Ann and Bob are dividing three goods: A, B, and C.

Adjusted Winner: Example

Suppose Ann and Bob are dividing three goods: A, B, and C.

Step 1. Both Ann and Bob divide 100 points among the three goods.

Adjusted Winner: Example

Suppose Ann and Bob are dividing three goods: A, B, and C.

Step 1. Both Ann and Bob divide 100 points among the three goods.

Item	Ann	Bob
A	5	4
B	65	46
C	30	50
Total	100	100

Adjusted Winner: Example

Suppose Ann and Bob are dividing three goods: A, B, and C.

Step 2. The agent who assigns the most points receives the item.

Adjusted Winner: Example

Suppose Ann and Bob are dividing three goods: A, B, and C.

Step 2. The agent who assigns the most points receives the item.

Item	Ann	Bob
A	5	4
B	65	46
C	30	50
Total	100	100

Adjusted Winner: Example

Suppose Ann and Bob are dividing three goods: A, B, and C.

Step 2. The agent who assigns the most points receives the item.

Item	Ann	Bob
A	5	0
B	65	0
C	0	50
Total	70	50

Adjusted Winner: Example

Suppose Ann and Bob are dividing three goods: A, B, and C.

Step 3. Equitability adjustment:

Adjusted Winner: Example

Suppose Ann and Bob are dividing three goods: A, B, and C.

Step 3. Equitability adjustment:
Notice that $65/46 \geq 5/4 \geq 1 \geq 30/50$

Item	Ann	Bob
A	5	4
B	65	46
C	30	50
Total	100	100

Adjusted Winner: Example

Suppose Ann and Bob are dividing three goods: A, B, and C.

Step 3. Equitability adjustment:

Adjusted Winner: Example

Suppose Ann and Bob are dividing three goods: A, B, and C.

Step 3. Equitability adjustment:

Give A to Bob (the item whose ratio is closest to 1)

Adjusted Winner: Example

Suppose Ann and Bob are dividing three goods: A, B, and C.

Step 3. Equitability adjustment:

Give A to Bob (the item whose ratio is closest to 1)

Item	Ann	Bob
A	5	0
B	65	0
C	0	50
Total	70	50

Adjusted Winner: Example

Suppose Ann and Bob are dividing three goods: A, B, and C.

Step 3. Equitability adjustment:

Give A to Bob (the item whose ratio is closest to 1)

Item	Ann	Bob
A	0	4
B	65	0
C	0	50
Total	65	54

Adjusted Winner: Example

Suppose Ann and Bob are dividing three goods: A, B, and C.

Step 3. Equitability adjustment:

Still not equal, so give (some of) B to Bob: $65p = 100 - 46p$.

Item	Ann	Bob
A	0	4
B	65	0
C	0	50
Total	65	54

Adjusted Winner: Example

Suppose Ann and Bob are dividing three goods: A, B, and C.

Step 3. Equitability adjustment:

yielding $p = 100/111 = 0.9009$

Item	Ann	Bob
A	0	4
B	65	0
C	0	50
Total	65	54

Adjusted Winner: Example

Suppose Ann and Bob are dividing three goods: A, B, and C.

Step 3. Equitability adjustment:

yielding $p = 100/111 = 0.9009$

Item	Ann	Bob
A	0	4
B	58.559	4.559
C	0	50
Total	58.559	58.559

Adjusted Winner: Formal Definition

Suppose that $G_1, \ldots, G_n$ is a fixed set of goods.

Adjusted Winner: Formal Definition

Suppose that $G_1, \ldots, G_n$ is a fixed set of goods.

A **valuation** of these goods is a vector of natural numbers $\langle a_1, \ldots, a_n \rangle$ whose sum is 100.

Let $\alpha, \alpha', \alpha'', \ldots$ denote possible valuations for Ann and $\beta, \beta', \beta'', \ldots$ denote possible valuations for Bob.

Adjusted Winner: Formal Definition

Suppose that $G_1, \ldots, G_n$ is a fixed set of goods.

Adjusted Winner: Formal Definition

Suppose that $G_1, \ldots, G_n$ is a fixed set of goods.

An **allocation** is a vector of n real numbers where each component is between 0 and 1 (inclusive). An allocation $\sigma = \langle s_1, \ldots, s_n \rangle$ is interpreted as follows.

For each $i = 1, \ldots, n$, s_i is the proportion of G_i given to Ann.

Thus if there are three goods, then $\langle 1, 0.5, 0 \rangle$ means, "Give all of item 1 and half of item 2 to Ann and all of item 3 and half of item 2 to Bob."

Fairness

- **Proportional** if both Ann and Bob receive at least 50% of their valuation: $\sum_{i=1}^{n} s_i a_i \geq 50$ and $\sum_{i=1}^{n}(1-s_i)b_i \geq 50$

Fairness

- **Proportional** if both Ann and Bob receive at least 50% of their valuation: $\sum_{i=1}^{n} s_i a_i \geq 50$ and $\sum_{i=1}^{n}(1 - s_i)b_i \geq 50$
- **Envy-Free** if no party is willing to give up its allocation in exchange for the other player's allocation: $\sum_{i=1}^{n} s_i a_i \geq \sum_{i=1}^{n}(1 - s_i)a_i$ and $\sum_{i=1}^{n}(1 - s_i)b_i \geq \sum_{i=1}^{n} s_i b_i$

Fairness

- **Proportional** if both Ann and Bob receive at least 50% of their valuation: $\sum_{i=1}^{n} s_i a_i \geq 50$ and $\sum_{i=1}^{n}(1-s_i)b_i \geq 50$
- **Envy-Free** if no party is willing to give up its allocation in exchange for the other player's allocation: $\sum_{i=1}^{n} s_i a_i \geq \sum_{i=1}^{n}(1-s_i)a_i$ and $\sum_{i=1}^{n}(1-s_i)b_i \geq \sum_{i=1}^{n} s_i b_i$
- **Equitable** if both players receive the same total number of points: $\sum_{i=1}^{n} s_i a_i = \sum_{i=1}^{n}(1-s_i)b_i$

Fairness

- **Proportional** if both Ann and Bob receive at least 50% of their valuation: $\sum_{i=1}^n s_i a_i \geq 50$ and $\sum_{i=1}^n (1 - s_i) b_i \geq 50$
- **Envy-Free** if no party is willing to give up its allocation in exchange for the other player's allocation: $\sum_{i=1}^n s_i a_i \geq \sum_{i=1}^n (1 - s_i) a_i$ and $\sum_{i=1}^n (1 - s_i) b_i \geq \sum_{i=1}^n s_i b_i$
- **Equitable** if both players receive the same total number of points: $\sum_{i=1}^n s_i a_i = \sum_{i=1}^n (1 - s_i) b_i$
- **Efficient** if there is no other allocation that is strictly better for one party without being worse for another party: for each allocation $\sigma' = \langle s_1', \ldots, s_n' \rangle$ if $\sum_{i=1}^n a_i s_i' > \sum_{i=1}^n a_i s_i$, then $\sum_{i=1}^n (1 - s_i') b_i < \sum_{i=1}^n (1 - s_i) b_i$. (Similarly for Bob)

Easy Observations

- For two-party disputes, proportionality and envy-freeness are equivalent.
- *AW* only produces equitable allocations (equitability is essentially built in to the procedure).
- *AW* produces allocations σ that in which at most one good is split.

Adjusted Winner is Fair

Theorem (Brams and Taylor) *AW produces allocations that are efficient, equitable and envy-free (with respect to the announced valuations)*

Some Questions

- Can we make use of geometric intuitions?
- Is *AW* a "continuous" function?
- It seems that the more the agents' utilities differ, the more points *AW* gives to each agent.
- The agents' utility functions are assumed to be linear, what about non-linear utility functions?
- Can an agent benefit by making use of information about the other agent's valuation?

Some Questions

- Can we make use of geometric intuitions? Yes!
- Is *AW* a "continuous" function? Yes and No
- It seems that the more the agents' utilities differ, the more points *AW* gives to each agent. Yes, we can prove this.
- The agents' utility functions are assumed to be linear, what about non-linear utility functions? The nonlinear situation may be interesting.
- Can an agent benefit by making use of information about the other agent's valuation? Yes, but in most cases it is not a "safe" strategy.

EP, R. Parikh and S. Salame. *Some Results on Adjusted Winner*. Proceedings of Indian Conference on Logic and Interaction, 2008.

Strategizing

Can the agents improve their allocation by misrepresenting their preferences?

Strategizing

Can the agents improve their allocation by misrepresenting their preferences?

Yes

Strategizing

Can the agents improve their allocation by misrepresenting their preferences?

Yes

However, while honesty may not always be the best policy it is the only **safe** *one, i.e., the only one which will guarantee 50%.*

Strategizing

Item	Ann	Bob
Matisse	75	25
Picasso	25	75

Ann will get the Matisse and Bob will get the Picasso and each gets 75 of his or her points.

Strategizing: Example

Suppose Ann knows Bob's preferences, but Bob does not know Ann's.

Item	Ann	Bob
M	75	25
P	25	75

Item	Ann	Bob
M	26	25
P	74	75

So Ann will get M plus a portion of P.

According to Ann's announced allocation, she receives 50 points

According to Ann's actual allocation, she receives $75 + 0.33 * 25 = 83.33$ points.

Strategizing: A Theorem

Theorem (Brams and Taylor) *Assume there are two goods, G_1 and G_2, all true and announced values are restricted to integers, and suppose Bob's announced valuation of G_1 is x, where $x \geq 50$. Assume Ann's true valuation of G_1 is b. Then her optimal announced valuation of G_1 is:*

$$\begin{cases} x+1 & \text{if } b > x \\ x & \text{if } b = x \\ x-1 & \text{if } b < x \end{cases}$$

Strategizing: Example

Suppose *both* players know each other's preferences but neither knows that the other knows their own preference.

Item	Ann	Bob
M	75	25
P	25	75

Item	Ann	Bob
M	26	74
P	74	26

Each will get 74 of his or her announced points, but each one is really getting only 25 of his or her *true* points.

Strategizing: Example

Suppose *both* players know each other's preferences. Moreover, Ann knows that Bob knows her preference and Bob doesn't know that Ann knows.

Item	Ann	Bob
M	26	74
P	74	26

Item	Ann	Bob
M	73	74
P	27	26

What happens as the level of knowledge increases?

What about a logical analysis?

What about a logical analysis?

- Which aspects of social situations should we focus on? Knowledge, Beliefs, Group Knowledge, Preferences, Desires, Ability, Actions, Intentions, Goals, Obligations, etc.

What about a logical analysis?

- Which aspects of social situations should we focus on? Knowledge, Beliefs, Group Knowledge, Preferences, Desires, Ability, Actions, Intentions, Goals, Obligations, etc.
- One grand system, or many smaller systems that loosely "fit" together?

What about a logical analysis?

- Which aspects of social situations should we focus on? Knowledge, Beliefs, Group Knowledge, Preferences, Desires, Ability, Actions, Intentions, Goals, Obligations, etc.
- One grand system, or many smaller systems that loosely "fit" together?
- *Combining* systems is hard! (conceptually and technically)

What about a logical analysis?

- Which aspects of social situations should we focus on? Knowledge, Beliefs, Group Knowledge, Preferences, Desires, Ability, Actions, Intentions, Goals, Obligations, etc.
- One grand system, or many smaller systems that loosely "fit" together?
- *Combining* systems is hard! (conceptually and technically)
- Logics *of* rational agents in social situations.

  vs.

  Logics *about* rational agents in social situations.

What about a logical analysis?

- Which aspects of social situations should we focus on? Knowledge, Beliefs, Group Knowledge, Preferences, Desires, Ability, Actions, Intentions, Goals, Obligations, etc.
- One grand system, or many smaller systems that loosely "fit" together?
- *Combining* systems is hard! (conceptually and technically)
- Logics *of* rational agents in social situations.

  vs.

  Logics *about* rational agents in social situations.
- Normative vs. Descriptive

The point of view of this model is not normative; it is not meant to advise the players what to do. The players do whatever they do; their strategies are taken as given.

The point of view of this model is not normative; it is not meant to advise the players what to do. The players do whatever they do; their strategies are taken as given. Neither is it meant as a description of what human beings actually do in interactive situations.

The point of view of this model is not normative; it is not meant to advise the players what to do. The players do whatever they do; their strategies are taken as given. Neither is it meant as a description of what human beings actually do in interactive situations. The most appropriate term is perhaps "analytic"; it asks, what are the implications *of rationality in interactive situations? Where does it lead?*

The point of view of this model is not normative; it is not meant to advise the players what to do. The players do whatever they do; their strategies are taken as given. Neither is it meant as a description of what human beings actually do in interactive situations. The most appropriate term is perhaps "analytic"; it asks, what are the implications *of rationality in interactive situations? Where does it lead? This question may be as important as, or even more important than, more direct "tests" of the relevance of the rationality hypothesis.*

R. Aumann. *Irrationality in Game Theory*. 1992.

Time for some details.

What is a modal?

A modal qualifies the truth of a judgement.

What is a modal?

A modal qualifies the truth of a judgement.

John ________ happy.

What is a modal?

A modal qualifies the truth of a judgement.

John _______ happy.

- is necessarily
- is possibly

What is a modal?

A modal qualifies the truth of a judgement.

John _______ happy.

- is necessarily
- is possibly
- is known/believed/certain (by Ann) to be

What is a modal?

A modal qualifies the truth of a judgement.

John _______ happy.

- is necessarily
- is possibly
- is known/believed/certain (by Ann) to be
- is permitted to be
- is obliged to be

What is a modal?

A modal qualifies the truth of a judgement.

John _______ happy.

- is necessarily
- is possibly
- is known/believed/certain (by Ann) to be
- is permitted to be
- is obliged to be
- is now
- will be

What is a modal?

A modal qualifies the truth of a judgement.

John ______ happy.

- is necessarily
- is possibly
- is known/believed/certain (by Ann) to be
- is permitted to be
- is obliged to be
- is now
- will be
- can do something to ensure that he is

What is a modal?

A modal qualifies the truth of a judgement.

John _______ happy.

- is necessarily
- is possibly
- is known/believed/certain (by Ann) to be
- is permitted to be
- is obliged to be
- is now
- will be
- can do something to ensure that he is
- ...

The Basic Modal Language

A wff of Modal Logic is defined *inductively*:

1. Any atomic propositional variable is a wff
2. If P and Q are wff, then so are $\neg P$, $P \wedge Q$, $P \vee Q$ and $P \rightarrow Q$
3. If P is a wff, then so is $\Box P$ and $\Diamond P$

The Basic Modal Language

A wff of Modal Logic is defined *inductively*:

1. Any atomic propositional variable is a wff
2. If P and Q are wff, then so are $\neg P$, $P \wedge Q$, $P \vee Q$ and $P \rightarrow Q$
3. If P is a wff, then so is $\Box P$ and $\Diamond P$

Boolean Logic

The Basic Modal Language

A wff of Modal Logic is defined *inductively*:

1. Any atomic propositional variable is a wff
2. If P and Q are wff, then so are $\neg P$, $P \wedge Q$, $P \vee Q$ and $P \rightarrow Q$
3. If P is a wff, then so is $\Box P$ and $\Diamond P$

Unary operator

The Basic Modal Language

A wff of Modal Logic is defined *inductively*:

1. Any atomic propositional variable is a wff
2. If P and Q are wff, then so are $\neg P$, $P \wedge Q$, $P \vee Q$ and $P \rightarrow Q$
3. If P is a wff, then so is $\Box P$ and $\Diamond P$

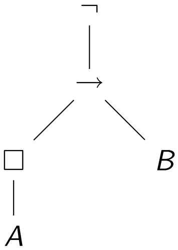
Eg., $\Box(P \rightarrow \Diamond Q) \vee \Box\Diamond\neg R$

Modal Formulas

$\boxed{\neg(\Box A \rightarrow B)}$ $\boxed{\neg\Box(A \rightarrow B)}$ $\boxed{(\neg\Box A \rightarrow B)}$

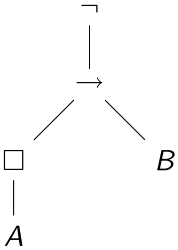
Modal Formulas

$\neg(\Box A \rightarrow B)$ $\neg\Box(A \rightarrow B)$ $(\neg\Box A \rightarrow B)$

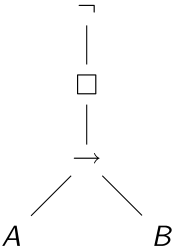


Modal Formulas

$\neg(\Box A \to B)$



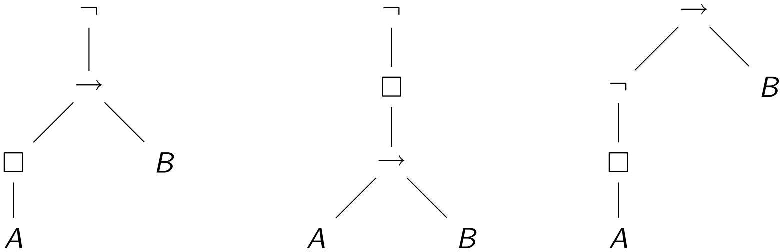
$\neg\Box(A \to B)$



$(\neg\Box A \to B)$

Modal Formulas

$\neg(\Box A \to B)$ $\neg\Box(A \to B)$ $(\neg\Box A \to B)$



One Language, Many Interpretations

Alethic

$\Box A$: A is necessary

$\Diamond A$: A is possible

One Language, Many Interpretations

Alethic
$\Box A$: A is necessary
$\Diamond A$: A is possible

Deontic
$\Box A$: A is obligatory
$\Diamond A$: A is permitted
(OA, PA)

One Language, Many Interpretations

Alethic
$\Box A$: A is necessary
$\Diamond A$: A is possible

Deontic
$\Box A$: A is obligatory
$\Diamond A$: A is permitted
(OA, PA)

Epistemic
$\Box A$: Ann knows φ
$\Diamond A$: it is consistent with Ann's information that φ
(KA, LA)

Valid?

$$\Box P \leftrightarrow \neg\Diamond\neg P$$

Valid?

$\Box P \leftrightarrow \neg\Diamond\neg P$
P is necessary/obligatory iff $\neg P$ *is not possible/permitted*

Valid?

$\Box P \leftrightarrow \neg\Diamond\neg P$
P is necessary/obligatory iff $\neg P$ is not possible/permitted

$\Box P \rightarrow P$

Valid?

$\Box P \leftrightarrow \neg\Diamond\neg P$
P is necessary/obligatory iff $\neg P$ is not possible/permitted

$\Box P \rightarrow P$
If P is necessary/known/obligatory then P is true

Valid?

$\Box P \leftrightarrow \neg\Diamond\neg P$

P is necessary/obligatory iff $\neg P$ *is not possible/permitted*

$\Box P \rightarrow P$

If P is necessary/known/obligatory then P is true

$\Box P \rightarrow \Box\Box P$

Valid?

$\Box P \leftrightarrow \neg\Diamond\neg P$

P is necessary/obligatory iff $\neg P$ *is not possible/permitted*

$\Box P \rightarrow P$

If P is necessary/known/obligatory then P is true

$\Box P \rightarrow \Box\Box P$

If Ann knows P then Ann knows that she knows P

Valid?

$\Box P \leftrightarrow \neg\Diamond\neg P$
P is necessary/obligatory iff $\neg P$ is not possible/permitted

$\Box P \rightarrow P$
If P is necessary/known/obligatory then P is true

$\Box P \rightarrow \Box\Box P$
If Ann knows P then Ann knows that she knows P

$\Diamond P \rightarrow \Box\Diamond P$,

Valid?

$\Box P \leftrightarrow \neg\Diamond\neg P$
P is necessary/obligatory iff $\neg P$ is not possible/permitted

$\Box P \rightarrow P$
If P is necessary/known/obligatory then P is true

$\Box P \rightarrow \Box\Box P$
If Ann knows P then Ann knows that she knows P

$\Diamond P \rightarrow \Box\Diamond P$, $\Box\Diamond P \rightarrow \Diamond\Box P$, etc.

Can we give find a natural *semantics* for the basic modal language?

Kripke Structures

Kripke Structures

The main idea:

- ‘It is sunny outside’ is currently true

Kripke Structures

The main idea:

- 'It is sunny outside' is currently true, but it is not necessary (for example, if we were currently in Ohio).

Kripke Structures

The main idea:

- ‘It is sunny outside’ is currently true, but it is not necessary (for example, if we were currently in Ohio).

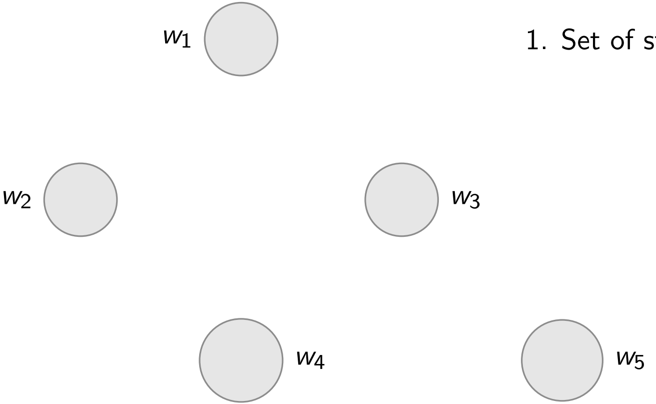
- We say P is **necessary** provided P is true in all (relevant) situations (states, worlds, possibilities).

Kripke Structures

The main idea:

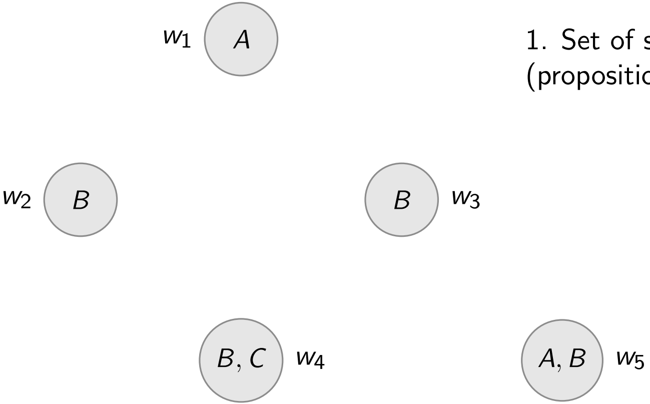
- 'It is sunny outside' is currently true, but it is not necessary (for example, if we were currently in Ohio).
- We say P is **necessary** provided P is true in all (relevant) situations (states, worlds, possibilities).
- A **Kripke structure** is
  1. A set of states, or worlds (each world specifies the truth value of all propositional variables)
  2. A **relation** on the set of states (specifying the "relevant situations")

A Kripke Structure



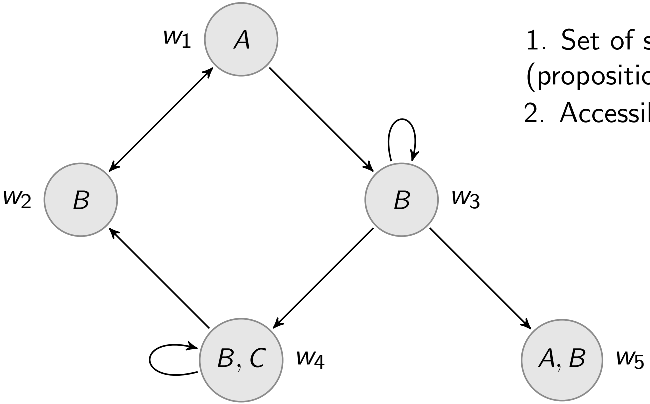
1. Set of states

A Kripke Structure



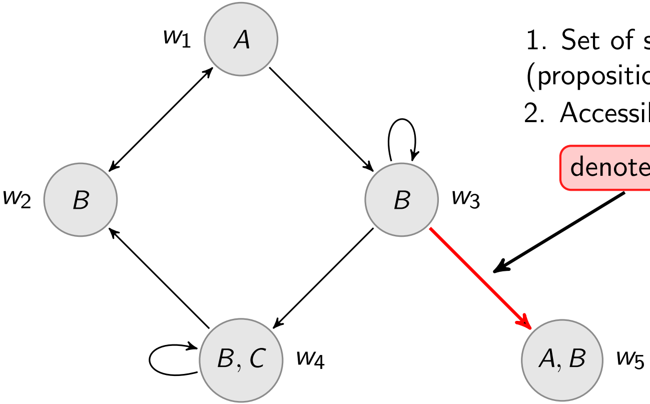
1. Set of states (propositional valuations)

A Kripke Structure



1. Set of states (propositional valuations)
2. Accessibility relation

A Kripke Structure



1. Set of states (propositional valuations)
2. Accessibility relation

denoted $w_3 R w_5$

Truth of Modal Formulas

We interpret formulas at states in a Kripke structure: $w \models P$ means P is true at state w.

Truth of Modal Formulas

We interpret formulas at states in a Kripke structure: $w \models P$ means P is true at state w.

We write wRv is v is accessible from state w.

Truth of Modal Formulas

We interpret formulas at states in a Kripke structure: $w \models P$ means P is true at state w.

We write wRv is v is accessible from state w.

1. $\Box P$ is true at state w iff P is true in all accessible worlds.

Truth of Modal Formulas

We interpret formulas at states in a Kripke structure: $w \models P$ means P is true at state w.

We write wRv is v is accessible from state w.

1. $\Box P$ is true at state w iff P is true in all accessible worlds.
   $w \models \Box P$ iff for all v, if wRv then $v \models P$

Truth of Modal Formulas

We interpret formulas at states in a Kripke structure: $w \models P$ means P is true at state w.

We write wRv is v is accessible from state w.

1. $\Box P$ is true at state w iff P is true in all accessible worlds.
   $w \models \Box P$ iff for all v, if wRv then $v \models P$

2. $\Diamond P$ is true at state w iff P is true at some accessible world.

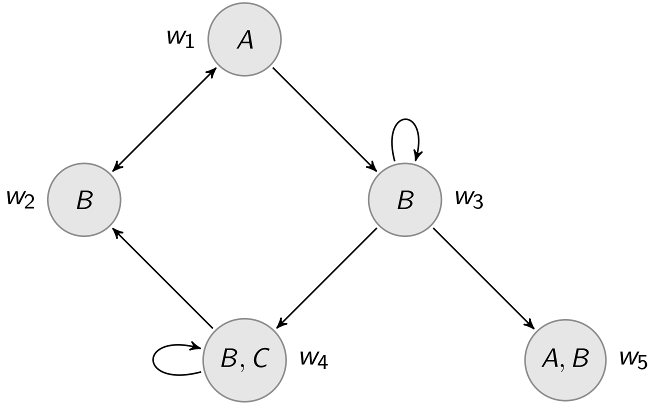
Truth of Modal Formulas

We interpret formulas at states in a Kripke structure: $w \models P$ means P is true at state w.

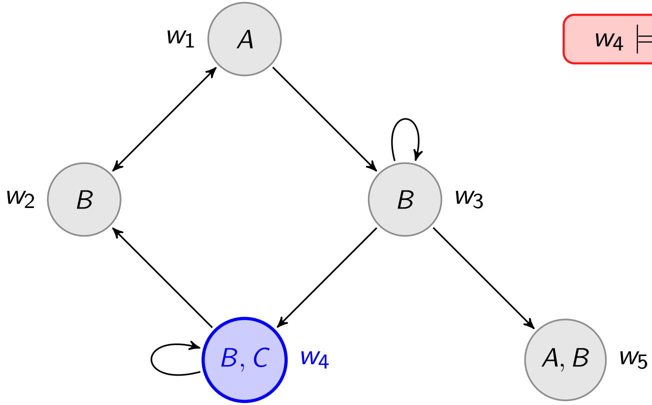
We write wRv is v is accessible from state w.

1. $\Box P$ is true at state w iff P is true in all accessible worlds.
   $w \models \Box P$ iff for all v, if wRv then $v \models P$
2. $\Diamond P$ is true at state w iff P is true at some accessible world.
   $w \models \Diamond P$ iff there exists v such that wRv and $v \models P$.

Example

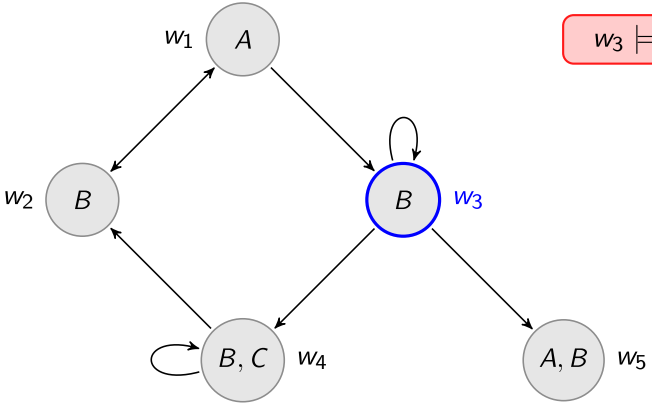


Example



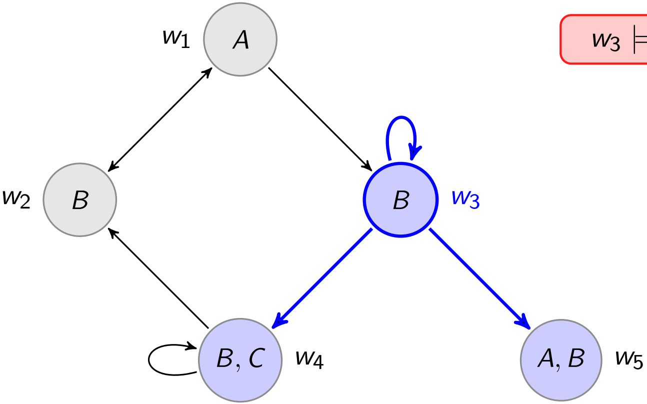
$w_4 \models B \wedge C$

Example



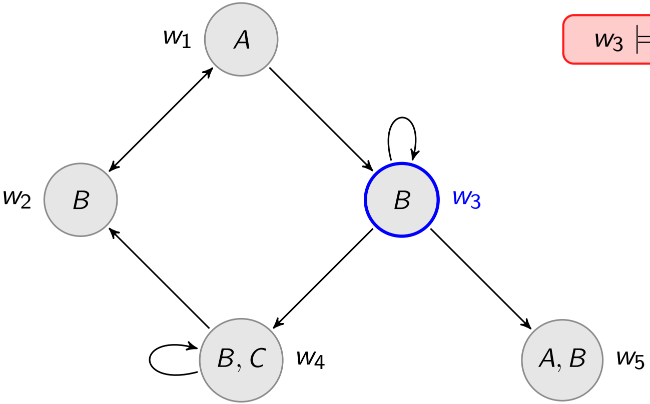
$w_3 \models \Box B$

Example



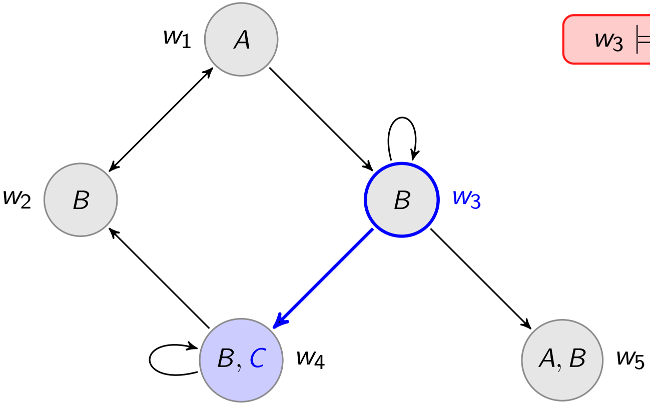
$w_3 \models \Box B$

Example



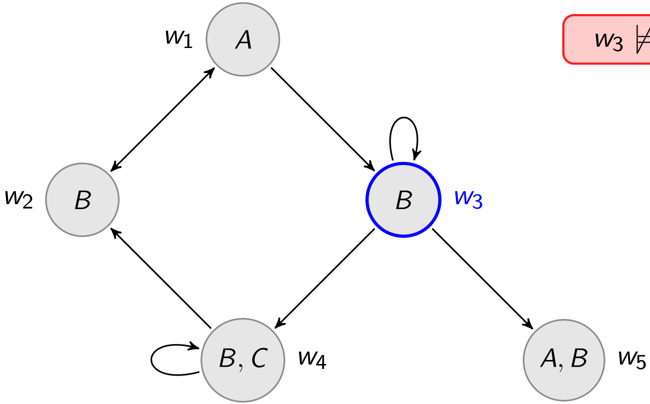
$w_3 \models \Diamond C$

Example



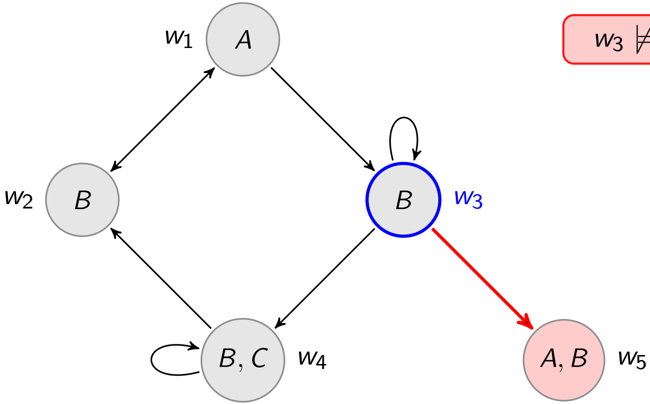
$w_3 \models \Diamond C$

Example



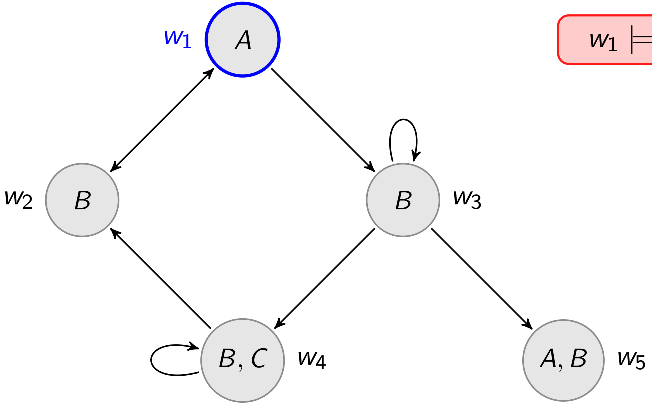
$w_3 \not\models \Box C$

Example



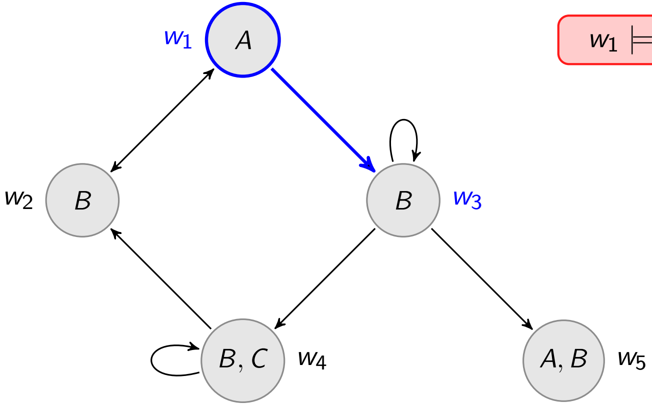
$w_3 \not\models \Box C$

Example



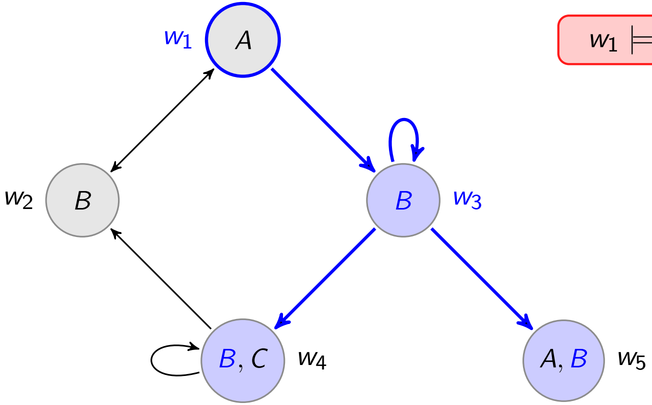
$w_1 \models \Diamond\Box B$

Example



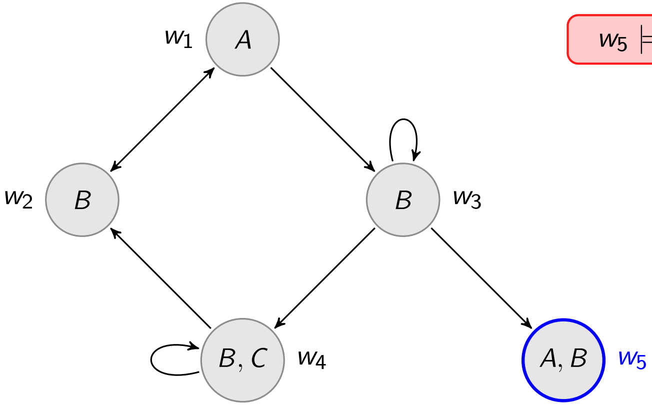
$w_1 \models \Diamond\Box B$

Example



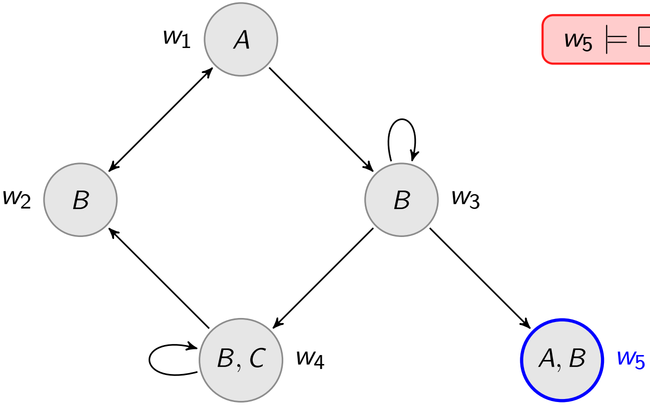
$w_1 \models \Diamond\Box B$

Example



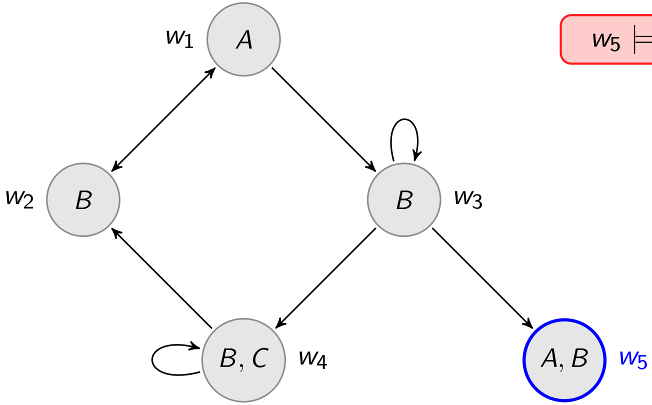
$w_5 \models \Box C$

Example

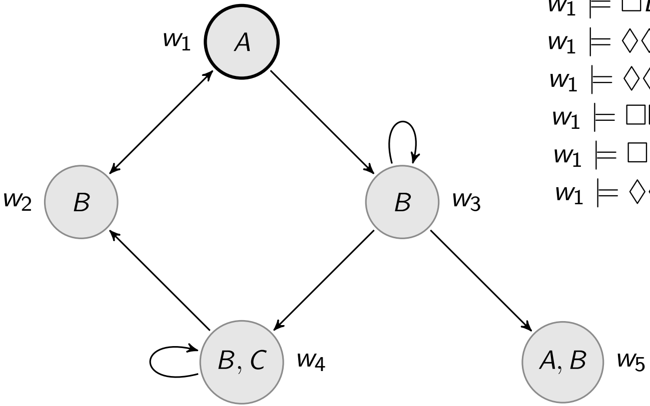


$w_5 \models \Box(B \wedge \neg B)$

Example



$w_5 \models \neg\Diamond B$



$w_1 \models \Box B \wedge B?$

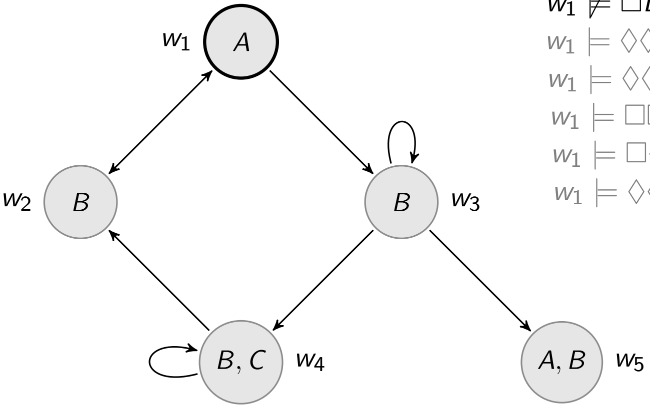
$w_1 \models \Diamond\Diamond B?$

$w_1 \models \Diamond\Diamond\Diamond B?$

$w_1 \models \Box\Box B?$

$w_1 \models \Box\Diamond C?$

$w_1 \models \Diamond\Diamond C?$



$w_1 \not\models \Box B \wedge B$

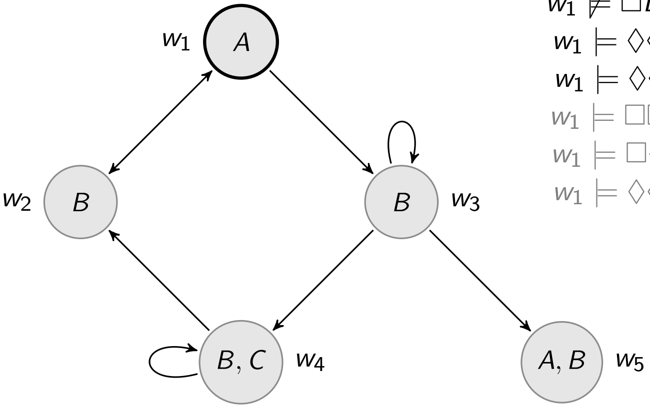
$w_1 \models \Diamond\Diamond B$?

$w_1 \models \Diamond\Diamond\Diamond B$?

$w_1 \models \Box\Box B$?

$w_1 \models \Box\Diamond C$?

$w_1 \models \Diamond\Diamond C$?



$w_1 \not\models \Box B \wedge B$

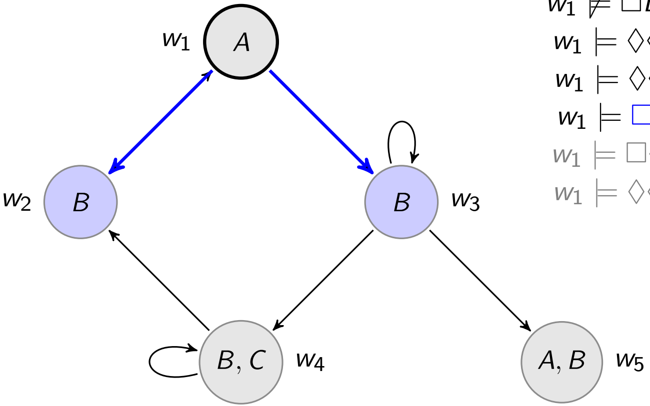
$w_1 \models \Diamond\Diamond B$

$w_1 \models \Diamond\Diamond\Diamond B$

$w_1 \models \Box\Box B?$

$w_1 \models \Box\Diamond C?$

$w_1 \models \Diamond\Diamond C?$



$w_1 \not\models \Box B \wedge B$

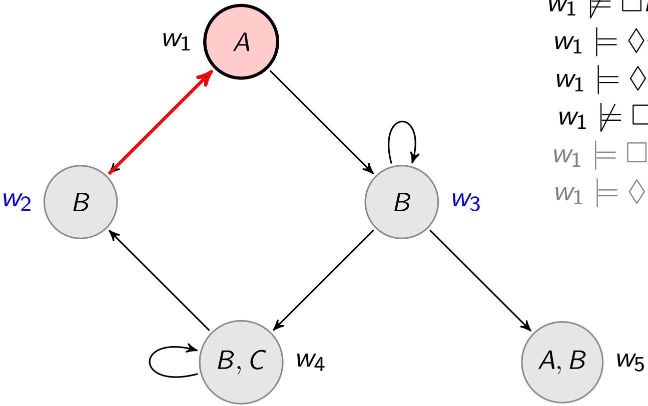
$w_1 \models \Diamond\Diamond B$

$w_1 \models \Diamond\Diamond\Diamond B$

$w_1 \models \Box\Box B$

$w_1 \models \Box\Diamond C?$

$w_1 \models \Diamond\Diamond C?$



$w_1 \not\models \Box B \wedge B$

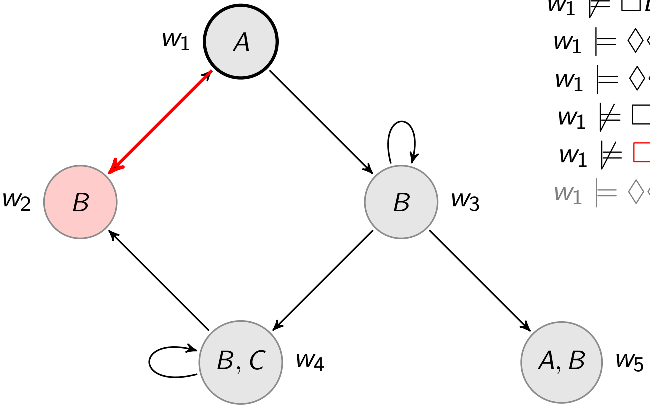
$w_1 \models \Diamond\Diamond B$

$w_1 \models \Diamond\Diamond\Diamond B$

$w_1 \not\models \Box\Box B$

$w_1 \models \Box\Diamond C?$

$w_1 \models \Diamond\Diamond C?$



$w_1 \not\models \Box B \wedge B$

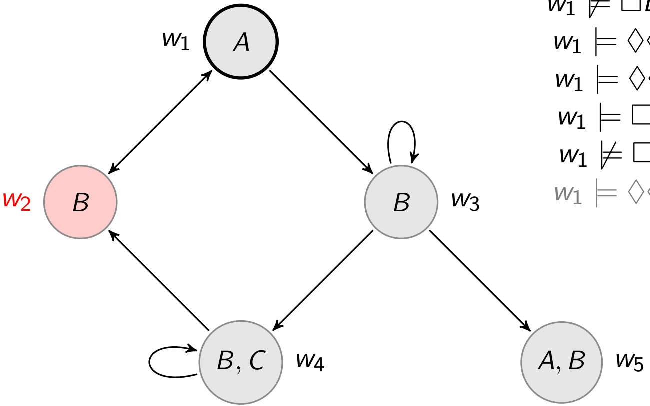
$w_1 \models \Diamond\Diamond B$

$w_1 \models \Diamond\Diamond\Diamond B$

$w_1 \not\models \Box\Box B$

$w_1 \not\models \Box\Diamond C$

$w_1 \models \Diamond\Diamond C?$



$w_1 \not\models \Box B \wedge B$

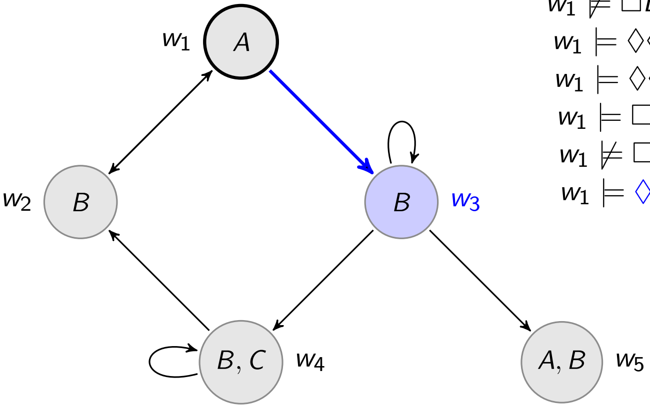
$w_1 \models \Diamond\Diamond B$

$w_1 \models \Diamond\Diamond\Diamond B$

$w_1 \models \Box\Box B$

$w_1 \not\models \Box\Diamond C$

$w_1 \models \Diamond\Diamond C?$



$w_1 \not\models \Box B \wedge B$

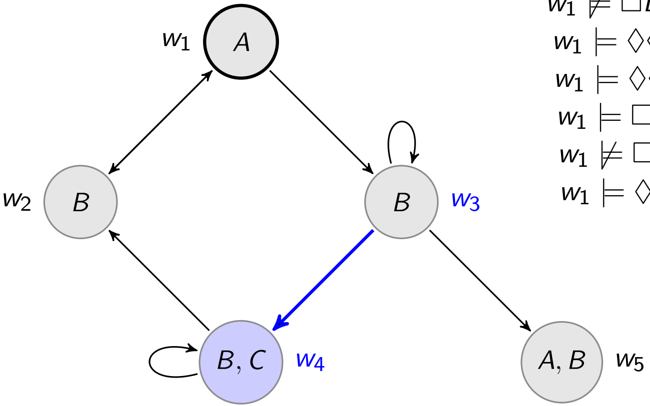
$w_1 \models \Diamond\Diamond B$

$w_1 \models \Diamond\Diamond\Diamond B$

$w_1 \models \Box\Box B$

$w_1 \not\models \Box\Diamond C$

$w_1 \models \Diamond\Diamond C$



$w_1 \not\models \Box B \wedge B$

$w_1 \models \Diamond\Diamond B$

$w_1 \models \Diamond\Diamond\Diamond B$

$w_1 \models \Box\Box B$

$w_1 \not\models \Box\Diamond C$

$w_1 \models \Diamond\Diamond C$

Some Facts

- $\Box P \vee \neg\Box P$ is always true (i.e., true at any state in any Kripke structure), but what about $\Box P \vee \Box\neg P$?

Some Facts

- $\Box P \vee \neg\Box P$ is always true (i.e., true at any state in any Kripke structure), but what about $\Box P \vee \Box\neg P$?

- $\Box P \wedge \Box Q \rightarrow \Box(P \wedge Q)$ is true at any state in any Kripke structure.

Some Facts

- $\Box P \vee \neg\Box P$ is always true (i.e., true at any state in any Kripke structure), but what about $\Box P \vee \Box\neg P$?

- $\Box P \wedge \Box Q \rightarrow \Box(P \wedge Q)$ is true at any state in any Kripke structure. What about $\Box(P \vee Q) \rightarrow \Box P \vee \Box Q$?

Some Facts

- $\Box P \vee \neg\Box P$ is always true (i.e., true at any state in any Kripke structure), but what about $\Box P \vee \Box\neg P$?

- $\Box P \wedge \Box Q \rightarrow \Box(P \wedge Q)$ is true at any state in any Kripke structure. What about $\Box(P \vee Q) \rightarrow \Box P \vee \Box Q$?

- $\Box P \leftrightarrow \neg\Diamond\neg P$ is true at any state in any Kripke structure.

But, we are not always interested in all Kripke structures.

But, we are not always interested in all Kripke structures.

For example, consider the epistemic interpretation: A state v is accessible from w (wRv) provided "given the agents information, w and v are indistinguishable".

But, we are not always interested in all Kripke structures.

For example, consider the epistemic interpretation: A state v is accessible from w (wRv) provided "given the agents information, w and v are indistinguishable". What are natural properties?

But, we are not always interested in all Kripke structures.

For example, consider the epistemic interpretation: A state v is accessible from w (wRv) provided "given the agents information, w and v are indistinguishable". What are natural properties?

Eg., for each state w, w is accessible from itself (R is a reflexive relation).

But, we are not always interested in all Kripke structures.

For example, consider the epistemic interpretation: A state v is accessible from w (wRv) provided "given the agents information, w and v are indistinguishable". What are natural properties?

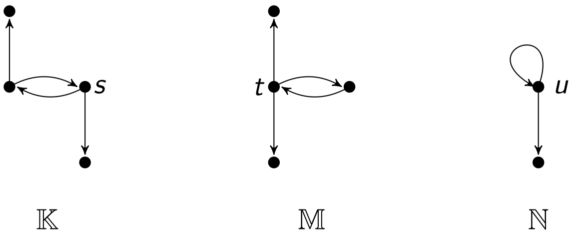
Eg., for each state w, w is accessible from itself (R is a reflexive relation).

Some Facts

- $\Box P \rightarrow P$ is true at any state in any Kripke structure where each state is accessible from itself.
- $\Box P \rightarrow \Diamond P$ is true at any state in any Kripke structure where each state has at least one accessible world.

Something to think about....

Which pair of states cannot be distinguished by a modal formula?
What about a first order formula?

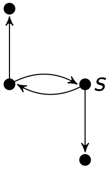


Slogan 1: Modal languages are simple yet expressive languages for talking about relational structures.

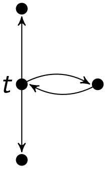
Slogan 2: Modal languages provide an internal, local perspective on relational structures.

P. Blackburn, M. de Rijke and Y. Venema. *Modal Logic*. 2001.

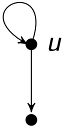
P. Blackburn and J. van Benthem. *Modal Logic: A Semantics Perspective*. Handbook of Modal Logic (2007).



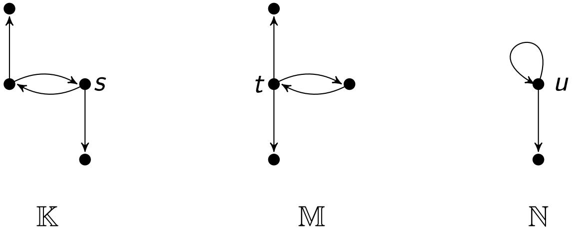
$\mathbb{K}$



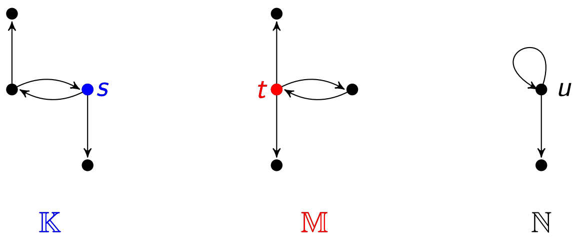
$\mathbb{M}$



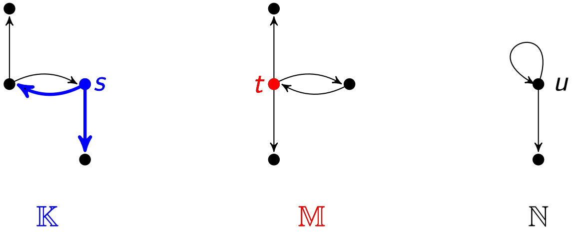
$\mathbb{N}$



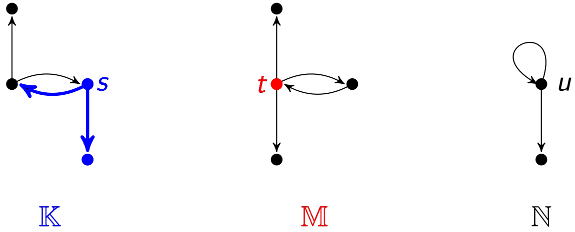
$$\Box(\Box\bot \vee \Diamond\Box\bot)$$



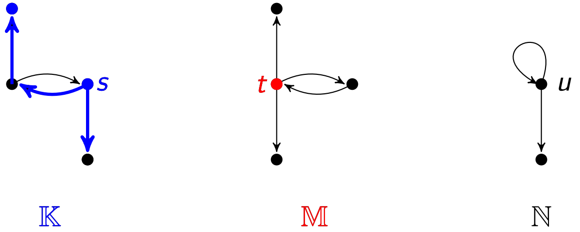
$$\Box(\Box\bot \vee \Diamond\Box\bot)$$



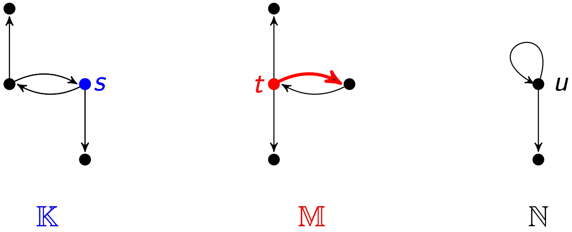
$$\Box(\Box\bot \vee \Diamond\Box\bot)$$



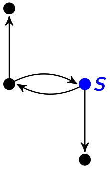
$$\Box(\Box\bot \vee \Diamond\Box\bot)$$



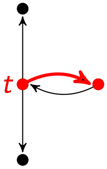
$$\Box(\Box\bot \vee \Diamond\Box\bot)$$



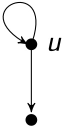
$$\Box(\Box\bot \vee \Diamond\Box\bot)$$



$\mathbb{K}$

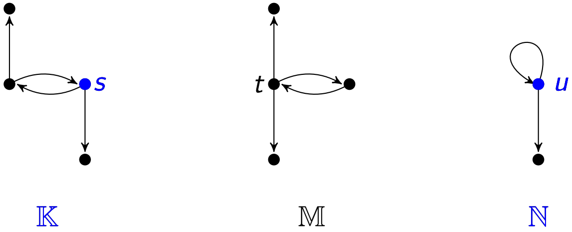


$\mathbb{M}$



$\mathbb{N}$

$$\Box(\Box\bot \vee \Diamond\Box\bot)$$


s
t
u
$\mathbb{K}$
$\mathbb{M}$
$\mathbb{N}$

Notation

A Kripke frame is a pair $\langle W, R\rangle$ where $R \subseteq W \times W$.

Notation

A Kripke frame is a pair $\langle W, R\rangle$ where $R \subseteq W \times W$.

Let $\mathbb{F} = \langle W, R\rangle$ be a Kripke frame and $\mathbb{M} = \langle W, R, V\rangle$ a model based on $\mathbb{M}$.

Notation

A Kripke frame is a pair $\langle W, R\rangle$ where $R \subseteq W \times W$.

Let $\mathbb{F} = \langle W, R\rangle$ be a Kripke frame and $\mathbb{M} = \langle W, R, V\rangle$ a model based on $\mathbb{M}$.

φ is satisfiable in $\mathbb{M}$ if there exists $w \in W$ such that $\mathbb{M}, w \models \varphi$

Notation

A Kripke frame is a pair $\langle W, R\rangle$ where $R \subseteq W \times W$.

Let $\mathbb{F} = \langle W, R\rangle$ be a Kripke frame and $\mathbb{M} = \langle W, R, V\rangle$ a model based on $\mathbb{M}$.

φ is satisfiable in $\mathbb{M}$ if there exists $w \in W$ such that $\mathbb{M}, w \models \varphi$

φ is valid in $\mathbb{M}$ ($\mathbb{M} \models \varphi$) if $\forall w \in W, \quad \mathbb{M}, w \models \varphi$

Notation

A Kripke frame is a pair $\langle W, R \rangle$ where $R \subseteq W \times W$.

Let $\mathbb{F} = \langle W, R \rangle$ be a Kripke frame and $\mathbb{M} = \langle W, R, V \rangle$ a model based on $\mathbb{M}$.

φ is satisfiable in $\mathbb{M}$ if there exists $w \in W$ such that $\mathbb{M}, w \models \varphi$

φ is valid in $\mathbb{M}$ ($\mathbb{M} \models \varphi$) if $\forall w \in W, \quad \mathbb{M}, w \models \varphi$

φ is valid on a frame $\mathbb{F}$ ($\mathbb{F} \models \varphi$) if for all models $\mathbb{M}$ based on $\mathbb{F}$, $\mathbb{M} \models \varphi$.

Definable Properties

A modal formula φ defines a class of frames K provided

$$\mathbb{F} \in \mathsf{K} \text{ iff } \mathbb{F} \models \varphi$$

- $\Box\varphi \rightarrow \Box\Box\varphi$ defines the class of transitive frames.
- $\varphi \leftrightarrow \Box\varphi$ defines the class of frames consisting of isolated reflexive points ($\forall x \in W,\ xRy \rightarrow x = y$).
- $\Box(\Box\varphi \rightarrow \varphi)$ defines the class of secondary-reflexive frames ($\forall w, v \in W,\ \text{ if } wRv \text{ then } vRv$).

Definable Properties

A modal formula φ defines a class of frames K provided

$$\mathbb{F} \in \mathsf{K} \text{ iff } \mathbb{F} \models \varphi$$

- ✓ $\Box\varphi \rightarrow \Box\Box\varphi$ defines the class of transitive frames.
- $\varphi \leftrightarrow \Box\varphi$ defines the class of frames consisting of isolated reflexive points ($\forall x \in W,\ xRy \rightarrow x = y$).
- $\Box(\Box\varphi \rightarrow \varphi)$ defines the class of secondary-reflexive frames ($\forall w, v \in W,$ if wRv then vRv).

Definable Properties

A modal formula φ defines a class of frames K provided

$$\mathbb{F} \in \mathsf{K} \text{ iff } \mathbb{F} \models \varphi$$

- ✓ $\Box\varphi \to \Box\Box\varphi$ defines the class of transitive frames.
- ✓ $\varphi \leftrightarrow \Box\varphi$ defines the class of frames consisting of isolated reflexive points ($\forall x \in W,\ xRy \to x = y$).
- $\Box(\Box\varphi \to \varphi)$ defines the class of secondary-reflexive frames ($\forall w, v \in W,\ $ if wRv then vRv).

Definable Properties

A modal formula φ defines a class of frames K provided

$$\mathbb{F} \in \mathsf{K} \text{ iff } \mathbb{F} \models \varphi$$

- ✓ $\Box\varphi \to \Box\Box\varphi$ defines the class of transitive frames.
- ✓ $\varphi \leftrightarrow \Box\varphi$ defines the class of frames consisting of isolated reflexive points ($\forall x \in W,\ xRy \to x = y$).
- ✓ $\Box(\Box\varphi \to \varphi)$ defines the class of secondary-reflexive frames ($\forall w, v \in W,$ if wRv then vRv).

Definable Properties

A modal formula φ defines a class of frames K provided

$$\mathbb{F} \in \mathsf{K} \text{ iff } \mathbb{F} \models \varphi$$

- ✓ $\Box\varphi \rightarrow \Box\Box\varphi$ defines the class of transitive frames.
- ✓ $\varphi \leftrightarrow \Box\varphi$ defines the class of frames consisting of isolated reflexive points ($\forall x \in W,\ xRy \rightarrow x = y$).
- ✓ $\Box(\Box\varphi \rightarrow \varphi)$ defines the class of secondary-reflexive frames ($\forall w, v \in W$, if wRv then vRv).

Some modal formulas correspond to genuine second-order properties: Löb ($\Box(\Box\varphi \rightarrow \varphi) \rightarrow \Box\varphi$), *McKinsey* ($\Box\Diamond\varphi \rightarrow \Diamond\Box\varphi$)

Definable Properties

A modal formula φ defines a class of frames K provided

$$\mathbb{F} \in \mathsf{K} \text{ iff } \mathbb{F} \models \varphi$$

- ✓ $\Box\varphi \rightarrow \Box\Box\varphi$ defines the class of transitive frames.
- ✓ $\varphi \leftrightarrow \Box\varphi$ defines the class of frames consisting of isolated reflexive points ($\forall x \in W,\ xRy \rightarrow x = y$).
- ✓ $\Box(\Box\varphi \rightarrow \varphi)$ defines the class of secondary-reflexive frames ($\forall w, v \in W,$ if wRv then vRv).

The Sahlqvist Theorem gives an algorithm *for finding a first-order correspondant for certain modal formulas.*

Slogan 3: Modal logics are not isolated formal systems.

The Standard Translation

$$st_x : \mathcal{L} \rightarrow \mathcal{L}_1$$

The Standard Translation

First-order language with an appropriate signature

$$st_x : \mathcal{L} \rightarrow \mathcal{L}_1$$

The Standard Translation

$st_x : \mathcal{L} \to \mathcal{L}_1$

$$\begin{array}{lcl} st_x(p) & = & Px \\ st_x(\neg\varphi) & = & \neg st_x(\varphi) \\ st_x(\varphi \wedge \psi) & = & st_x(\varphi) \wedge st_x(\psi) \end{array}$$

The Standard Translation

$$st_x : \mathcal{L} \rightarrow \mathcal{L}_1$$

$$\begin{array}{lcl} st_x(p) & = & Px \\ st_x(\neg\varphi) & = & \neg st_x(\varphi) \\ st_x(\varphi \wedge \psi) & = & st_x(\varphi) \wedge st_x(\psi) \\ st_x(\Box\varphi) & = & \forall y(xRy \rightarrow st_y(\varphi)) \end{array}$$

The Standard Translation

$st_x : \mathcal{L} \to \mathcal{L}_1$

$$
\begin{array}{lcl}
st_x(p) & = & Px \\
st_x(\neg\varphi) & = & \neg st_x(\varphi) \\
st_x(\varphi \wedge \psi) & = & st_x(\varphi) \wedge st_x(\psi) \\
st_x(\Box\varphi) & = & \forall y(xRy \to st_y(\varphi)) \\
st_x(\Diamond\varphi) & = & \exists y(xRy \wedge st_y(\varphi))
\end{array}
$$

The Standard Translation

$st_x : \mathcal{L} \to \mathcal{L}_1$

$$
\begin{array}{lcl}
st_x(p) & = & Px \\
st_x(\neg\varphi) & = & \neg st_x(\varphi) \\
st_x(\varphi \wedge \psi) & = & st_x(\varphi) \wedge st_x(\psi) \\
st_x(\Box\varphi) & = & \forall y(xRy \to st_y(\varphi)) \\
st_x(\Diamond\varphi) & = & \exists y(xRy \wedge st_y(\varphi))
\end{array}
$$

Fact: Modal logic falls in the two-variable fragment of $\mathcal{L}_1$.

The Standard Translation

$$st_x : \mathcal{L} \to \mathcal{L}_1$$

$$\begin{array}{lcl}
st_x(p) & = & Px \\
st_x(\neg\varphi) & = & \neg st_x(\varphi) \\
st_x(\varphi \wedge \psi) & = & st_x(\varphi) \wedge st_x(\psi) \\
st_x(\Box\varphi) & = & \forall y(xRy \to st_y(\varphi)) \\
st_x(\Diamond\varphi) & = & \exists y(xRy \wedge st_y(\varphi))
\end{array}$$

Lemma For each $w \in W$, $\mathbb{M}, w \models \varphi$ iff $\mathbb{M} \Vdash st_x(\varphi)[x/w]$.

What can we say with modal logic? What about in comparison with first-order logic?

Disjoint Union

Definition Let $\mathbb{M}_1 = \langle W_1, R_1, V_1 \rangle$ and $\mathbb{M}_2 = \langle W_2, R_2, V_2 \rangle$. The disjoint union is the structure $\mathbb{M}_1 \uplus \mathbb{M}_2 = \langle W, R, V \rangle$ where

- $W = W_1 \cup W_2$
- $R = R_1 \cup R_2$
- for all $p \in \mathsf{At}$, $V(p) = V_1(p) \cup V_2(p)$

Disjoint Union

Definition Let $\mathbb{M}_1 = \langle W_1, R_1, V_1 \rangle$ and $\mathbb{M}_2 = \langle W_2, R_2, V_2 \rangle$. The disjoint union is the structure $\mathbb{M}_1 \uplus \mathbb{M}_2 = \langle W, R, V \rangle$ where

- $W = W_1 \cup W_2$
- $R = R_1 \cup R_2$
- for all $p \in \mathsf{At}$, $V(p) = V_1(p) \cup V_2(p)$

Lemma For each collection of Kripke structures $\{\mathbb{M}_i \mid i \in I\}$, for each $w \in W_i$, $\mathbb{M}_i, w \models \varphi$ iff $\biguplus_{i \in I} \mathbb{M}_i, w \models \varphi$

Disjoint Union

Definition Let $\mathbb{M}_1 = \langle W_1, R_1, V_1 \rangle$ and $\mathbb{M}_2 = \langle W_2, R_2, V_2 \rangle$. The disjoint union is the structure $\mathbb{M}_1 \uplus \mathbb{M}_2 = \langle W, R, V \rangle$ where

- $W = W_1 \cup W_2$
- $R = R_1 \cup R_2$
- for all $p \in \mathsf{At}$, $V(p) = V_1(p) \cup V_2(p)$

Fact The universal modality is not definable in the basic modal language.

Generated Submodel

Definition $\mathbb{M}' = \langle W', R', V' \rangle$ is a generated submodel of $\mathbb{M} = \langle W, R, V \rangle$ provided

- $W' \subseteq W$ is R-closed:
  for each $w' \in W$ and $v \in W$, if wRv then $v \in W'$.
- $R' = R \cap W' \times W'$
- for all $p \in \mathrm{At}$, $V'(p) = V(p) \cap W'$

Generated Submodel

Definition $\mathbb{M}' = \langle W', R', V' \rangle$ is a generated submodel of $\mathbb{M} = \langle W, R, V \rangle$ provided

- $W' \subseteq W$ is R-closed:
  for each $w' \in W$ and $v \in W$, if wRv then $v \in W'$.
- $R' = R \cap W' \times W'$
- for all $p \in \mathsf{At}$, $V'(p) = V(p) \cap W'$

Lemma If $\mathbb{M}'$ is a generated submodel of $\mathbb{M}$ then for each $w \in W'$, $\mathbb{M}', w \models \varphi$ iff $\mathbb{M}, w \models \varphi$

Generated Submodel

Definition $\mathbb{M}' = \langle W', R', V' \rangle$ is a generated submodel of $\mathbb{M} = \langle W, R, V \rangle$ provided

- $W' \subseteq W$ is R-closed:
  for each $w' \in W$ and $v \in W$, if wRv then $v \in W'$.
- $R' = R \cap W' \times W'$
- for all $p \in \mathrm{At}$, $V'(p) = V(p) \cap W'$

Fact The backwards looking modality is not definable in the basic modal language.

Bounded Morphism

Definition A bounded morphism between models $\mathbb{M} = \langle W, R, V\rangle$ and $\mathbb{M}' = \langle W', R', V'\rangle$ is a function f with domain W and range W' such that:

Atomic harmony: for each $p \in \mathrm{At}$, $w \in V(p)$ iff $f(w) \in V'(p)$

Morphism: if wRv then $f(w)Rf(v)$

Zag: if $f(w)R'v'$ then $\exists v \in W$ such that $f(v) = v'$ and wRv

Bounded Morphism

Definition A bounded morphism between models $\mathbb{M} = \langle W, R, V \rangle$ and $\mathbb{M}' = \langle W', R', V' \rangle$ is a function f with domain W and range W' such that:

Atomic harmony: for each $p \in \mathrm{At}$, $w \in V(p)$ iff $f(w) \in V'(p)$

Morphism: if wRv then $f(w)Rf(v)$

Zag: if $f(w)R'v'$ then $\exists v \in W$ such that $f(v) = v'$ and wRv

Lemma If $\mathbb{M}'$ is a bounded morphic image of $\mathbb{M}$ then for each $w \in W$, $\mathbb{M}, w \models \varphi$ iff $\mathbb{M}', f(w) \models \varphi$

Bounded Morphism

Definition A bounded morphism between models $\mathbb{M} = \langle W, R, V\rangle$ and $\mathbb{M}' = \langle W', R', V'\rangle$ is a function f with domain W and range W' such that:

Atomic harmony: for each $p \in \mathrm{At}$, $w \in V(p)$ iff $f(w) \in V'(p)$

Morphism: if wRv then $f(w)Rf(v)$

Zag: if $f(w)R'v'$ then $\exists v \in W$ such that $f(v) = v'$ and wRv

Fact Counting modalities are not definable in the basic modal language (eg., $\Diamond_1\varphi$ iff φ is true in more than 1 accessible world).

Bisimulation

A bisimulation between $\mathbb{M} = \langle W, R, V\rangle$and $\mathbb{M}' = \langle W', R', V'\rangle$ is a non-empty binary relation $Z \subseteq W \times W'$ such that whenever wZw':

Atomic harmony: for each $p \in \mathsf{At}$, $w \in V(p)$ iff $w' \in V'(p)$

Zig: if wRv, then $\exists v' \in W'$ such that vZv' and $w'R'v'$

Zag: if $w'R'v'$ then $\exists v \in W$ such that vZv' and wRv

Bisimulation

A bisimulation between $\mathbb{M} = \langle W, R, V\rangle$and $\mathbb{M}' = \langle W', R', V'\rangle$ is a non-empty binary relation $Z \subseteq W \times W'$ such that whenever wZw':

Atomic harmony: for each $p \in \mathrm{At}$, $w \in V(p)$ iff $w' \in V'(p)$

Zig: if wRv, then $\exists v' \in W'$ such that vZv' and $w'R'v'$

Zag: if $w'R'v'$ then $\exists v \in W$ such that vZv' and wRv

We write $\mathbb{M}, w \,\underline{\leftrightarrow}\, \mathbb{M}', w'$ *if there is a* Z *such that* wZw'*.*

Bisimulation

A bisimulation between $\mathbb{M} = \langle W, R, V\rangle$and $\mathbb{M}' = \langle W', R', V'\rangle$ is a non-empty binary relation $Z \subseteq W \times W'$ such that whenever wZw':

Atomic harmony: for each $p \in \text{At}$, $w \in V(p)$ iff $w' \in V'(p)$

Zig: if wRv, then $\exists v' \in W'$ such that vZv' and $w'R'v'$

Zag: if $w'R'v'$ then $\exists v \in W$ such that vZv' and wRv

We write $\mathbb{M}, w \leftrightsquigarrow \mathbb{M}', w'$ *iff* $\forall \varphi \in \mathcal{L}$, $\mathbb{M}, w \models \varphi$ *iff* $\mathbb{M}', w' \models \varphi$.

Bisimulation

A bisimulation between $\mathbb{M} = \langle W, R, V \rangle$and $\mathbb{M}' = \langle W', R', V' \rangle$ is a non-empty binary relation $Z \subseteq W \times W'$ such that whenever wZw':

Atomic harmony: for each $p \in \mathsf{At}$, $w \in V(p)$ iff $w' \in V'(p)$

Zig: if wRv, then $\exists v' \in W'$ such that vZv' and $w'R'v'$

Zag: if $w'R'v'$ then $\exists v \in W$ such that vZv' and wRv

Lemma If $\mathbb{M}, w \underline{\leftrightarrow} \mathbb{M}', w'$ then $\mathbb{M}, w \leftrightsquigarrow \mathbb{M}', w'$.

Bisimulation

A bisimulation between $\mathbb{M} = \langle W, R, V\rangle$and $\mathbb{M}' = \langle W', R', V'\rangle$ is a non-empty binary relation $Z \subseteq W \times W'$ such that whenever wZw':

Atomic harmony: for each $p \in \mathsf{At}$, $w \in V(p)$ iff $w' \in V'(p)$

Zig: if wRv, then $\exists v' \in W'$ such that vZv' and $w'R'v'$

Zag: if $w'R'v'$ then $\exists v \in W$ such that vZv' and wRv

Lemma On finite frames, if $\mathbb{M}, w \leftrightsquigarrow \mathbb{M}', w'$ then $\mathbb{M}, w \underline{\leftrightarrow} \mathbb{M}', w'$.

The Van Benthem Characterization Theorem

Modal logic is the bisimulation invariant fragment of first-order logic.

The Van Benthem Characterization Theorem

For any first-order formula $\varphi(x)$, TFAE:

1. $\varphi(x)$ is invariant for bisimulation
2. $\varphi(x)$ is equivalent to the standard translation of a basic modal formula.

Logics of Rational Agency

Basic Ingredients

- What are the basic building blocks? (the nature of time (continuous or discrete/branching or linear), how (primitive) *events* or *actions* are represented, how *causal* relationships are represented and what constitutes a *state of affairs*.)

Basic Ingredients

- What are the basic building blocks? (the nature of time (continuous or discrete/branching or linear), how (primitive) *events* or *actions* are represented, how *causal* relationships are represented and what constitutes a *state of affairs*.)
- Single agent vs. many agents.

Basic Ingredients

- What are the basic building blocks? (the nature of time (continuous or discrete/branching or linear), how (primitive) *events* or *actions* are represented, how *causal* relationships are represented and what constitutes a *state of affairs*.)
- Single agent vs. many agents.
- What the the primitive operators?
  - Informational attitudes
  - Motivational attitudes
  - Normative attitudes

Basic Ingredients

- What are the basic building blocks? (the nature of time (continuous or discrete/branching or linear), how (primitive) *events* or *actions* are represented, how *causal* relationships are represented and what constitutes a *state of affairs*.)
- Single agent vs. many agents.
- What the the primitive operators?
  - Informational attitudes
  - Motivational attitudes
  - Normative attitudes
- Static vs. dynamic

✓ informational attitudes (eg., knowledge, belief, certainty)
✓ time, actions and ability
✓ motivational attitudes (eg., preferences)
✓ group notions (eg., common knowledge and coalitional ability)
✓ normative attitudes (eg., obligations)

End of lecture 1.