Lecture 35: More Common Knowledge

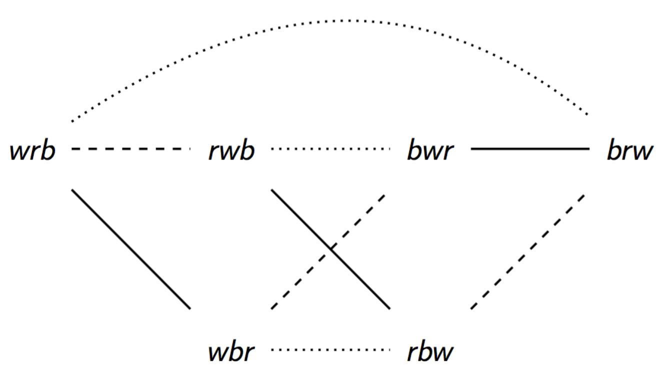
CS 181O
Spring 2016
Kim Bruce

Some slides based on those of Christina Unger

New situation

- Three cards: red, white, blue, and three players: Alice, Bob, and Carol. Deal the three cards out to three players, face down. Each looks at her own card.

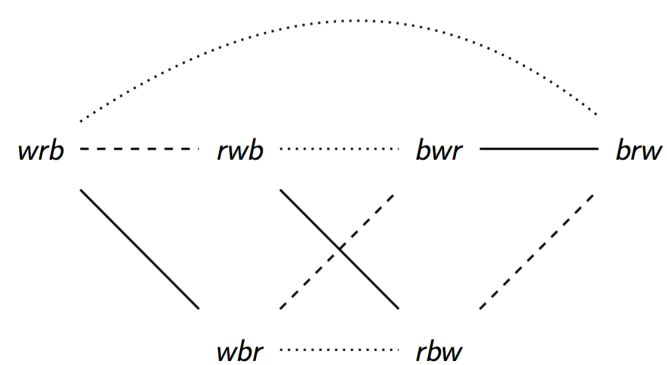
Possible Worlds



Solid = Alice
Small dotted = Bob
Big dotted = Carol

Lines represent accessibility relation for each player

Possible Worlds



What if Alice says her card is red? How does this change possible worlds?
What if Alice says her card is not white

Propositional Dynamic Logic

- Originated in theoretical CS to prove programs correct.
- Added infinite number of modal operators associated with (non-deterministic) programs.
  - $[\pi]\phi$ true iff after every execution of π, ϕ will be true.
  - $<\pi>\phi$ true iff after some execution of π, ϕ will be true.

Programs

- Programs built out of primitive statements a,b, ... using
  - $\phi \cup \psi$ (for non-determinism),
  - $\phi ; \psi$ (for composition)
  - ϕ^* (for iteration)
  - $\phi?$ (for test): if ϕ true then continue, else fail

Models

- ... are of the form
  $\mathcal{M} = (W, \{R^\pi \mid \pi \text{ is a program }\}, V)$ where
  W is set of worlds, each $R^\pi \subseteq W \times W$ and V is valuation on each world.
  - $R^{[[\phi?]]} = \{(w,w) \mid \mathcal{M},w \vDash \phi\}$
  - $R^{\pi 1 \cup \pi 2} = R^{\pi 1} \cup R^{\pi 2}$,
  - $R^{\pi 1;\pi 2} = R^{\pi 1} \circ R^{\pi 2}$ (= $\{(x,y) \mid \exists z\ (R^{\pi 1}(x,z) \wedge R^{\pi 2}(z,y)\}$),
  - $R^{(\pi *)} = (R^\pi)^*$, the reflexive transitive closure of R^π
- Start w/ transitions on primitive programs & compose meanings as above.

Expressiveness

- Can encode regular programs:
  - $(p? ; a) \cup (\neg p? ; b) \approx$ if p then a else b
  - $a; (\neg p?; a)*$ ≈ repeat a until p
  - $(p?; a)*; \neg p?$ ≈ while p do a
- Theorem. PDL has the finite model property and is decidable. Its satisfiability problem is EXPTIME-complete.

Epistemic Dynamic Logic

- Let a be accessibility relation for an agent, ϕ is statement that can be true or false, while α is an accessibility relation.
- $\phi ::= \top, p, \neg\phi, \phi_1 \wedge \phi_2, [\alpha]\phi, [!\phi_1]\phi_2$
- $\alpha ::= a \mid ?\phi \mid \alpha_1; \alpha_2 \mid \alpha_1 \cup \alpha_2 \mid \alpha^*$
- Interpret in model $\mathcal{M} = \langle W,V,R \rangle$ where W is possible worlds, V assigns values to prop letters, and R assigns each agent an equivalence relation to show accessible worlds

Epistemic Dynamic Logic

- Let P be set of proposition letters and A set of agents.
- Interpret formulas over A, P in model
  $\mathcal{M} = \langle W,V,R \rangle$
  where W is possible worlds,
  $V{:}W \rightarrow t$ assigns truth values to prop letters, &
  R assigns each agent an equivalence relation to relate mutually accessible worlds
  - Assume accessibility is reflexive, symmetric, transitive

Interpreting formulas

- Formulas interpreted as set of worlds true of.
- Interpret $\top, p, \neg\phi, \phi_1 \wedge \phi_2$ as usual.
- $[[\ [\alpha]\phi\ ]]^{\mathcal{M}} = \{w \in W \mid \forall v, \text{ if } (w,v) \in [[\alpha]]^{\mathcal{M}} \text{ then } v \in [[\phi\ ]]^{\mathcal{M}}\}$, I.e. ϕ true in all accessible worlds for $[[\alpha]]^{\mathcal{M}}$, so "α knows ϕ"
- $[[\ [!\phi_1]\phi_2]]^{\mathcal{M}} = (W - [[\phi_1]]^{\mathcal{M}}) \cup\ [[\phi_2]]^{\mathcal{M}|\phi_1}$
  where $\mathcal{M}|\ \phi_1 = (W', V', R')$ with $W' = [[\phi_1]]^{\mathcal{M}}$,
  V' = V restricted to W', &
  R'(a) = R(a) restricted to W' for each a.

Example Formulas

- $[a]p$: a "knows" p (because true in all of a's accessible worlds)
- $\neg[a]\neg p$: p is compatible with what a knows (some accessible world makes it true)
- $[(a \cup b)^*]p$: It is common knowledge for a, b that p
- $[!\phi_1]\phi_2$ says after announcement of ϕ_1, ϕ_2 will be true. Effect is to restrict model to only those worlds in which ϕ_1 is true.
  - $M,w \models [!\phi]\psi$ iff ($M,w \models \phi$ implies $M|\phi,w \models \psi$)

Properties of $!\phi$

- $[!\phi]\, p \Leftrightarrow \phi \rightarrow p$
- $[!\phi]\, \neg\psi \Leftrightarrow \phi \rightarrow \neg\, [!\phi]\, \psi$
- $[!\phi]\, (\psi_1 \wedge \psi_2) \Leftrightarrow [!\phi]\, \psi_1 \wedge [!\phi]\, \psi_2$

- Go back to card model and see effect of public announcement.

Application: Presuppositions

- A presupposition of an utterance is an implicit assumption about the world or a background belief shared by speaker & hearer in a discourse.
- "Shall we do it again?"
  - Presupposition: we have done it before.
- "Jan washed her car."
  - Presupposition: 'Jan' has a car (& Jan is female)

More Presuppositions

- The king of France is bald
  - There is a unique king of France
- Mary knows Bob is at the store
  - Bob is at the store
- Mary stopped beating her dog.
  - Mary has a dog and formerly beat it.
- If I were superman, then I could fly.
  - I'm not superman

Entailment vs Presupposition

- Jan has a red car $\Rightarrow$ Jan has a car
- Jan washed her car $\Rightarrow$ Jan has a car
- Tests for distinction:
  - Presuppositions are preserved under negation
    - Jan did not wash her car $\Rightarrow$ Jan has a car
    - Jan does not have a red car $\not\Rightarrow$ Jan has a car

In Logic

- Jan washed her car:
  - $\exists$c. (car(c) $\wedge$ owns(jan,c) $\wedge$ female(jan) $\wedge$ washed(jan,c))
- Jan did not wash her car:
  - $\exists$c. (car(c) $\wedge$ owns(jan,c) $\wedge$ $\wedge$ female(jan) $\neg$washed(jan,c))
  - This is *not* the negation of the first!
- Strawson:
  - Let ϕ and ψ be propositions: ϕ presupposes ψ iff ψ is a precondition of the truth or falsity of ϕ. I.e., if ϕ is either true or false then ψ is true.

Test for Presuppositions

- Negation ✔ *Check on earlier examples*
- Question: Make statement into question.
  - Did Jan wash her car?
  - Is the king of France bald?
  - Compare: Does Jan have a red car?
- Conditional: Make it an antecedent:
  - If Jan washed her car then it is clean
    - Presupposition Jan has a car remains
    - Compare: If Jan has a red car then it is clear
    - Entailment of Jan has a car disappears

Common Knowledge

- Suppose have update: $![(i \cup j)^*]\phi$
  - It is common knowledge between i and j that ϕ.
  - If ϕ is already common knowledge then no impact
  - If not common knowledge then remove all worlds!
    - $[[\ [(i \cup j)^*]\phi\ ]] = \{w \mid \forall v \text{ s.t. } (w,v) \in (i \cup j)^*,\ v \in [[\phi]]\}$

Common Knowledge & Presuppositions

- Presupposition is common knowledge between speaker and listener.
  - Jan washed her car
  - $\exists$c. $[(i \cup j)^*]$(car(c) $\wedge$ owns(jan,c) $\wedge$ female(jan)) $\wedge$ washed(jan,c))
  - First part common knowledge, last not!

Questions?