Logic and Artificial Intelligence

Lecture 21

Eric Pacuit

Currently Visiting the Center for Formal Epistemology, CMU

Center for Logic and Philosophy of Science
Tilburg University

ai.stanford.edu/~epacuit
e.j.pacuit@uvt.nl

November 16, 2011
Recap: Logics of Action and Ability

- $F\varphi$: $\varphi$ is true at some moment in the future
Recap: Logics of Action and Ability

- $F\varphi$: $\varphi$ is true at some moment in the future
- $\exists F\varphi$: there is a history where $\varphi$ is true some moment in the future
Recap: Logics of Action and Ability

- $F\varphi$: $\varphi$ is true at some moment in the future
- $\exists F\varphi$: there is a history where $\varphi$ is true some moment in the future
- $[\alpha]\varphi$: after doing action $\alpha$, $\varphi$ is true
Recap: Logics of Action and Ability

- $F\varphi$: $\varphi$ is true at some moment in the future
- $\exists F\varphi$: there is a history where $\varphi$ is true some moment in the future
- $[\alpha]\varphi$: after doing action $\alpha$, $\varphi$ is true
- $[\delta \varphi]\psi$: after bringing about $\varphi$, $\psi$ is true
Recap: Logics of Action and Ability

- $F\varphi$: $\varphi$ is true at some moment in the future
- $\exists F\varphi$: there is a history where $\varphi$ is true some moment in the future
- $[\alpha]\varphi$: after doing action $\alpha$, $\varphi$ is true
- $[\delta \varphi] \psi$: after bringing about $\varphi$, $\psi$ is true
- $[i stit] \varphi$: the agent can “see to it that” $\varphi$ is true
Recap: Logics of Action and Ability

- $F\varphi$: $\varphi$ is true at some moment in the future.
- $\exists F\varphi$: there is a history where $\varphi$ is true some moment in the future.
- $[\alpha]\varphi$: after doing action $\alpha$, $\varphi$ is true.
- $[\delta\varphi]\psi$: after bringing about $\varphi$, $\psi$ is true.
- $[i\ stit]\varphi$: the agent can “see to it that” $\varphi$ is true.
- $\diamond[i\ stit]\varphi$: the agent has the ability to bring about $\varphi$. 
Group/Collective actions and abilities: for $J \subseteq N$, $[J]\varphi$ means “the group can make $\varphi$ true...”
Group Actions: Example
Group Actions: Example

Suppose that there are two agents: a server ($s$) and a client ($c$). The client asks to set the value of $x$ and the server can either grant or deny the request. Assume the agents make simultaneous moves.
Group Actions: Example

Suppose that there are two agents: a server ($s$) and a client ($c$). The client asks to set the value of $x$ and the server can either grant or deny the request. Assume the agents make simultaneous moves.

\[
\begin{array}{c|c|c}
\text{set1} & \text{deny} & \text{grant} \\
\hline
\text{set2} & & \\
\end{array}
\]
Suppose that there are two agents: a server \((s)\) and a client \((c)\). The client asks to set the value of \(x\) and the server can either grant or deny the request. Assume the agents make simultaneous moves.

<table>
<thead>
<tr>
<th></th>
<th>deny</th>
<th>grant</th>
</tr>
</thead>
<tbody>
<tr>
<td>set1</td>
<td></td>
<td>(q_0 \Rightarrow q_0, q_1 \Rightarrow q_0)</td>
</tr>
<tr>
<td>set2</td>
<td></td>
<td>(q_0 \Rightarrow q_1, q_1 \Rightarrow q_1)</td>
</tr>
</tbody>
</table>
Example: Suppose that there are two agents: a server (s) and a client (c). The client asks to set the value of \( x \) and the server can either grant or deny the request. Assume the agents make simultaneous moves.

<table>
<thead>
<tr>
<th>set1</th>
<th>deny</th>
<th>grant</th>
</tr>
</thead>
<tbody>
<tr>
<td>( q \rightarrow q )</td>
<td>( q_0 \rightarrow q_0, q_1 \rightarrow q_0 )</td>
<td></td>
</tr>
<tr>
<td>( q \rightarrow q )</td>
<td>( q_0 \rightarrow q_1, q_1 \rightarrow q_1 )</td>
<td></td>
</tr>
</tbody>
</table>
Strategy Logics

- **Coalitional Logic**: Reasoning about (local) group power.

  \([C]\varphi\): coalition \(C\) has a **joint action** to bring about \(\varphi\).

Strategy Logics

- **Coalitional Logic**: Reasoning about (local) group power.
  
  \([C]\varphi\): coalition \(C\) has a **joint action** to bring about \(\varphi\).


- **Alternating-time Temporal Logic**: Reasoning about (local and global) group power:

  \(\langle\langle A\rangle\rangle G \varphi\): The coalition \(A\) has a **joint action** to ensure that \(\varphi\) will remain true.

Multi-agent Transition Systems

\[ (P_{x=1} \rightarrow [s]P_{x=1}) \land (P_{x=2} \rightarrow [s]P_{x=2}) \]

\[ q_0 \quad x = 1 \]

\[ (\text{set2, grant}) \]

\[ q_1 \quad x = 2 \]

\[ (\text{set1, grant}) \]

\[ q_0 \quad (\text{* , deny}) \]

\[ q_1 \quad (\text{* , deny}) \]
Multi-agent Transition Systems

\[
P_{x=1} \rightarrow \neg[s]P_{x=2}
\]
Multi-agent Transition Systems

\[ P_{x=1} \rightarrow [s, c]P_{x=2} \]

\[ \langle *, deny \rangle \]

\[ \langle set2, grant \rangle \quad \langle set1, grant \rangle \]

\[ q_0 \quad x = 1 \]

\[ q_1 \quad x = 2 \]
∃ “something an agent/a group can do” such that ∀ “actions of the other players/nature”...


∃ “something an agent/a group can do” such that ∀ “actions of the other players/nature”...


∀ “(joint) actions of the other players”, ∃ “something the agent/coalition can do”...

Coalitional Logic

Effectivity Functions

Let $N$ be a (finite) set of agents and $W$ a set of worlds.
Effectivity Functions

Let $N$ be a (finite) set of agents and $W$ a set of worlds.

For each $J \subseteq N$, the **effectivity function** for $J$ is $e : \wp(N) \to \wp(\wp(W))$.

$X \in e(J)$ means $X$ is a set of possible outcomes for which $J$ is effective (or $J$ can force the world to be in some state of $X$ at the next step).
e is **playable** (for a set of states \( \mathcal{W} \)) iff

1. \( \emptyset \not\in e(\mathcal{J}) \) (Liveness)
2. \( \mathcal{W} \in e(\mathcal{J}) \) (Safety)
3. if \( \mathcal{W} - X \not\in e(\emptyset) \) then \( X \in e(\mathcal{N}) \) (\( N \)-maximality)
4. if \( X \in e(\mathcal{J}) \) and \( X \subseteq Y \) then \( Y \in e(\mathcal{J}) \) (Outcome monotonicity)
5. If \( J \cap I = \emptyset \), then if \( X_1 \in e(J) \) and \( X_2 \in e(I) \), then \( X_1 \cap X_2 \in e(J \cup I) \) (Super-additivity)
Playable Effectivity Functions

$e$ is the effectivity function of some strategic game provided $X \in e(J)$ if there is a joint strategy for $J$ such that no matter what strategy is chosen by the agents $N - J$, the outcome of the game is in $X$.

**Theorem** (Pauly). An effectivity function $e$ is playable iff it is the effectivity function of some strategic game.
Playable Effectivity Functions

\( e \) is the effectivity function of some strategic game provided \( X \in e(J) \) if there is a joint strategy for \( J \) such that no matter what strategy is chosen by the agents \( N - J \), the outcome of the game is in \( X \).

**Theorem** (Pauly). An effectivity function \( e \) is playable iff it is the effectivity function of some strategic game.

See, also

A coalitional logic model is a tuple $\mathcal{M} = \langle W, E, V \rangle$ where $W$ is a set of states, $E : W \rightarrow (\wp(N) \rightarrow \wp(\wp(W)))$ assigns to each state a playable effectivity function, and $V : \text{At} \rightarrow \wp(W)$ is a valuation function.

$\mathcal{M}, w \models [J]\varphi$ iff $\llbracket \varphi \rrbracket_{\mathcal{M}} = \{ w | \mathcal{M}, w \models \varphi \} \in E(w)(J)$
Coalitional Logic: Axiomatics

- $\neg [J] \bot$
- $[J] \top$
- $(\neg [\emptyset] \neg \phi) \rightarrow [N] \phi$
- $[J] (\phi \land \psi) \rightarrow ([J] \phi \land [J] \psi)$
- $([J_1] \varphi_1 \land [J_2] \varphi_2) \rightarrow [J_1 \cup J_2] (\varphi_1 \land \varphi_2)$, where $J_1 \cap J_2 = \emptyset$
Logics of preference...
Preference (Modal) Logics

\( x, y \) objects

\( x \succeq y: x \) is at least as good as \( y \)
Preference (Modal) Logics

\( x, y \) objects

\( x \succeq y \): \( x \) is at least as good as \( y \)

1. \( x \succeq y \) and \( y \not\succeq x \) (\( x \succ y \))
2. \( x \not\succeq y \) and \( y \succeq x \) (\( y \succ x \))
3. \( x \succeq y \) and \( y \succeq x \) (\( x \sim y \))
4. \( x \not\succeq y \) and \( y \not\succeq x \) (\( x \perp y \))
Preference (Modal) Logics

$x, y$ objects

$x \succeq y$: $x$ is at least as good as $y$

1. $x \succeq y$ and $y \not\succeq x$ ($x \succ y$)
2. $x \not\succeq y$ and $y \succeq x$ ($y \succ x$)
3. $x \succeq y$ and $y \succeq x$ ($x \sim y$)
4. $x \not\succeq y$ and $y \not\succeq x$ ($x \perp y$)

Properties: transitivity, connectedness, etc.
Preference (Modal) Logics

Modal betterness model \( \mathcal{M} = \langle W, \succeq, V \rangle \)
Preference (Modal) Logics

Modal betterness model $\mathcal{M} = \langle W, \succeq, V \rangle$

Preference Modalities $\langle \succeq \rangle \varphi$: “there is a world at least as good (as the current world) satisfying $\varphi$”

$\mathcal{M}, w \models \langle \succeq \rangle \varphi$ iff there is a $v \succeq w$ such that $\mathcal{M}, v \models \varphi$
Preference (Modal) Logics

Modal betterness model $\mathcal{M} = \langle W, \succeq, V \rangle$

Preference Modalities $\langle \succeq \rangle \varphi$: “there is a world at least as good (as the current world) satisfying $\varphi$”

$\mathcal{M}, w \models \langle \succeq \rangle \varphi$ iff there is a $v \succeq w$ such that $\mathcal{M}, v \models \varphi$

$\mathcal{M}, w \models \langle \succ \rangle \varphi$ iff there is $v \succeq w$ and $w \not\succeq v$ such that $\mathcal{M}, v \models \varphi$
Preference (Modal) Logics

1. $⟨\neg⟩\varphi \rightarrow ⟨\geq⟩\varphi$
2. $⟨\geq⟩⟨\neg⟩\varphi \rightarrow ⟨\neg⟩\varphi$
3. $\varphi \land ⟨\geq⟩\psi \rightarrow (⟨\neg⟩\psi \lor ⟨\geq⟩(\psi \land ⟨\geq⟩\varphi))$
4. $⟨\neg⟩⟨\geq⟩\varphi \rightarrow ⟨\neg⟩\varphi$

**Theorem** The above logic (with Necessitation and Modus Ponens) is sound and complete with respect to the class of preference models.

Preference Modalities

\( \varphi \geq \psi \): the state of affairs \( \varphi \) is at least as good as \( \psi \) (ceteris paribus)

Preference Modalities

$\varphi \geq \psi$: the state of affairs $\varphi$ is at least as good as $\psi$ (ceteris paribus)


$\langle \Gamma \rangle \leq \varphi$: $\varphi$ is true in “better” world, *all things being equal*.

All Things Being Equal...

- With boots (\(b\)), I prefer my raincoat (\(r\)) over my umbrella (\(u\)).
- Without boots (\(\neg b\)), I also prefer my raincoat (\(r\)) over my umbrella (\(u\)).
- But I do prefer an umbrella and boots over a raincoat and no boots.
All Things Being Equal...

With boots ($b$), I prefer my raincoat ($r$) over my umbrella ($u$).

But I do prefer an umbrella and boots over a raincoat and no boots.

With boots ($b$), I prefer my raincoat ($r$) over my umbrella ($u$).
All Things Being Equal...

- With boots ($b$), I prefer my raincoat ($r$) over my umbrella ($u$).
- Without boots ($\neg b$), I also prefer my raincoat ($r$) over my umbrella ($u$).
All Things Being Equal...

- With boots ($b$), I prefer my raincoat ($r$) over my umbrella ($u$)
- Without boots ($\neg b$), I also prefer my raincoat ($r$) over my umbrella ($u$)
- But I do prefer an umbrella and boots over a raincoat and no boots
All things being equal, I prefer my raincoat over my umbrella.
All Things Being Equal...

Let $\Gamma$ be a set of (preference) formulas. Write $w \equiv_{\Gamma} v$ if for all $\varphi \in \Gamma$, $w \models \varphi$ iff $v \models \varphi$.

1. $M, w \models \langle \Gamma \rangle \varphi$ iff there is a $v \in W$ such that $w \equiv_{\Gamma} v$ and $M, v \models \varphi$.

2. $M, w \models \langle \Gamma \rangle \leq \varphi$ iff there is a $v \in W$ such that $w(\equiv_{\Gamma} \cap \leq) v$ and $M, v \models \varphi$.

3. $M, w \models \langle \Gamma \rangle < \varphi$ iff there is a $v \in W$ such that $w(\equiv_{\Gamma} \cap <) v$ and $M, v \models \varphi$.

Key Principles:

$\langle \Gamma' \rangle \varphi \rightarrow \langle \Gamma \rangle \varphi$ if $\Gamma \subseteq \Gamma'$

$\pm \varphi \land \langle \Gamma \rangle \left( \alpha \land \pm \varphi \right) \rightarrow \langle \Gamma \cup \{ \varphi \} \rangle \alpha$
All Things Being Equal...

Let $\Gamma$ be a set of (preference) formulas. Write $w \equiv_{\Gamma} v$ if for all $\varphi \in \Gamma$, $w \models \varphi$ iff $v \models \varphi$.

1. $\mathcal{M}, w \models \langle \Gamma \rangle \varphi$ iff there is a $v \in W$ such that $w \equiv_{\Gamma} v$ and $\mathcal{M}, v \models \varphi$.

2. $\mathcal{M}, w \models \langle \Gamma \rangle \leq \varphi$ iff there is a $v \in W$ such that $w(\equiv_{\Gamma} \cap \leq) v$ and $\mathcal{M}, v \models \varphi$.

3. $\mathcal{M}, w \models \langle \Gamma \rangle < \varphi$ iff there is a $v \in W$ such that $w(\equiv_{\Gamma} \cap <) v$ and $\mathcal{M}, v \models \varphi$.

Key Principles:

- $\langle \Gamma' \rangle \varphi \rightarrow \langle \Gamma \rangle \varphi$ if $\Gamma \subseteq \Gamma'$
- $\pm \varphi \land \langle \Gamma \rangle (\alpha \land \pm \varphi) \rightarrow \langle \Gamma \cup \{\varphi\} \rangle \alpha$
All Things Being Equal...

Let $\Gamma$ be a set of (preference) formulas. Write $w \equiv_\Gamma v$ if for all $\varphi \in \Gamma$, $w \models \varphi$ iff $v \models \varphi$.

1. $\mathcal{M}, w \models \langle \Gamma \rangle \varphi$ iff there is a $v \in \mathcal{W}$ such that $w \equiv_\Gamma v$ and $\mathcal{M}, v \models \varphi$.
2. $\mathcal{M}, w \models \langle \Gamma \rangle \leq \varphi$ iff there is a $v \in \mathcal{W}$ such that $w(\equiv_\Gamma \cap \leq) v$ and $\mathcal{M}, v \models \varphi$.
3. $\mathcal{M}, w \models \langle \Gamma \rangle < \varphi$ iff there is a $v \in \mathcal{W}$ such that $w(\equiv_\Gamma \cap <) v$ and $\mathcal{M}, v \models \varphi$.

Key Principles:

- $\langle \Gamma' \rangle \leq \varphi \rightarrow \langle \Gamma \rangle \leq \varphi$ if $\Gamma \subseteq \Gamma'$
- $\pm \varphi \land \langle \Gamma \rangle \leq (\alpha \land \pm \varphi) \rightarrow \langle \Gamma \cup \{\varphi\} \rangle \leq \alpha$
Preference Lifting, I

Given a preference ordering \( \preceq \) over a set of objects \( X \), we want to \textbf{lift} this to an ordering \( \widehat{\preceq} \) over \( \wp(X) \).

Given \( \preceq \), what reasonable properties can we infer about \( \widehat{\preceq} \)?

---

Preference Lifting, II

- You know that $x \prec y \prec z$
  Can you infer that $\{x, y\} \hat{\prec} \{z\}$?
You know that $x ≺ y ≺ z$
Can you infer that $\{x, y\} ≺ \{z\}$?

You know that $x ≺ y ≺ z$
Can you infer anything about $\{y\}$ and $\{x, z\}$?
Preference Lifting, II

- You know that $x \prec y \prec z$
  Can you infer that $\{x, y\} \hat{\prec} \{z\}$?

- You know that $x \prec y \prec z$
  Can you infer anything about $\{y\}$ and $\{x, z\}$?

- You know that $w \prec x \prec y \prec z$
  Can you infer that $\{w, x, y\} \hat{\preceq} \{w, y, z\}$?
Preference Lifting, II

▶ You know that $x \prec y \prec z$
  Can you infer that $\{x, y\} \sim \{z\}$?

▶ You know that $x \prec y \prec z$
  Can you infer anything about $\{y\}$ and $\{x, z\}$?

▶ You know that $w \prec x \prec y \prec z$
  Can you infer that $\{w, x, y\} \lesssim \{w, y, z\}$?

▶ You know that $w \prec x \prec y \prec z$
  Can you infer that $\{w, x\} \sim \{y, z\}$?
Preference Lifting, III

There are different interpretations of $X \hat{\leq} Y$:

- You will get one of the elements, but cannot control which.
- You can choose one of the elements.
- You will get the full set.