

# Lecture 10: Strategy Logics

Eric Pacuit

ILLC, University of Amsterdam

[staff.science.uva.nl/~epacuit](http://staff.science.uva.nl/~epacuit)

[epacuit@science.uva.nl](mailto:epacuit@science.uva.nl)

Lecture Date: April 20, 2006

Caput Logic, Language and Information: Social Software

[staff.science.uva.nl/~epacuit/caputLLI.html](http://staff.science.uva.nl/~epacuit/caputLLI.html)

---

## Introduction

Imagine a set of agents (players, system components) taking actions, simultaneously or in turns, on a common set of states — and thus effecting transitions between these states.

Assume the agents pursue certain goals and in that pursuit they can form **coalitions**

The objective is to develop formal tools for reasoning about coalitions of agents and their ability to achieve specified outcomes in such situations.

---

---

## (Temporal) Logics for Reasoning about Actions

- *Linear Time Temporal Logic (LTL)* [Pnuelli, 1977]: Reasoning about computations:
    - $\diamond\phi$ :  $\phi$  is true some time in the future.
  - *Branching-time Temporal Logic (CTL, CTL\*)* [Clarke and Emerson, 1981, Emerson and Halpern, 1986]: Allows quantification over paths:
    - $\exists\diamond\phi$ : there is a path in which  $\phi$  is eventually true.
  - *Alternating-time Temporal Logic (ATL, ATL\*)* [Alur, Henzinger, Kupferman, 1997]: Selective quantification over paths:
    - $\langle\langle A \rangle\rangle \bigcirc \phi$ : The coalition  $A$  has a joint strategy to ensure that  $\phi$  is true at the next moment.
-

---

## ATL:Syntax

Let  $\Sigma$  be a set of agents,  $\Pi$  a set of propositional variables and  $A \subseteq \Sigma$ .

1.  $p$  where  $p \in \Pi$
  2.  $\neg\phi$
  3.  $\phi \vee \psi$
  4.  $\langle\langle A \rangle\rangle \bigcirc \phi$  meaning ‘The coalition  $A$  can force in the next move an outcome satisfying  $\phi$ ’
  5.  $\langle\langle A \rangle\rangle \Box \phi$  meaning ‘The coalition  $A$  can maintain forever outcomes satisfying  $\phi$ ’
  6.  $\langle\langle A \rangle\rangle \phi U \psi$  meaning ‘The coalition  $A$  can eventually force an outcome satisfying  $\psi$  while meanwhile maintaining the truth of  $\phi$ ’
-

---

## Pauly's Coalitional Logic: Syntax

Given a finite non-empty set of agents  $N$  and a set of atomic propositions  $\Phi_0$ , a formula  $\phi$  can have the following syntactic form

$$\phi := \perp \mid p \mid \neg\phi \mid \phi \vee \phi \mid [C]\phi$$

where  $p \in \Phi_0$  and  $C \subseteq N$ .

$[C]\phi$  is intended to mean “coalition  $C$  can (locally) force  $\phi$  to be true”

---

## Coalitional Logic: Multi-player Game Models

A **Strategic Game Form** is a tuple  $\langle N, \{\Sigma_i \mid i \in N\}, Q, o \rangle$  where

- $N$  is a set of agents
- $\Sigma_i$  is a set of actions
- $Q$  is a set of states
- $o : \prod_{i \in N} \Sigma_i \rightarrow Q$  assigns an outcome to each choice of action.

Let  $\Gamma_Q^N$  be the set of all such strategic game forms.

A **Multi-Player Game Model** is a tuple  $\langle Q, \gamma, \pi \rangle$  where  $Q$  is a set of states and  $\gamma : Q \rightarrow \Gamma_Q^N$  associates strategic games form to each state

$$q \models [C] \phi \text{ iff } \exists \sigma_C \forall \sigma_{N-C}, o(\sigma_C, \sigma_{N-C}) \models \phi$$

---

---

## Effectivity Functions

Let  $G$  be a strategic game.

$$X \in E_G^\alpha(C) \text{ iff } \exists \sigma_C \forall \sigma_{\bar{C}} \quad o(\sigma_C, \sigma_{\bar{C}}) \in X$$

$$X \in E_G^\beta(C) \text{ iff } \forall \sigma_{\bar{C}} \exists \sigma_C \quad o(\sigma_C, \sigma_{\bar{C}}) \in X$$

$$E_G^\alpha \subseteq E_G^\beta$$

---

$$E_G^\beta \not\subseteq E_G^\alpha$$

Player 1 chooses the row, Player 2 chooses the column, Player 3 chooses the table

|     | $l$   | $m$   | $r$   |
|-----|-------|-------|-------|
| $l$ | $s_1$ | $s_2$ | $s_1$ |
| $r$ | $s_2$ | $s_1$ | $s_3$ |

|     | $l$   | $m$   | $r$   |
|-----|-------|-------|-------|
| $l$ | $s_3$ | $s_1$ | $s_2$ |
| $r$ | $s_2$ | $s_3$ | $s_3$ |

$$\{s_2\} \in E_G^\beta(\{2\}) \text{ but } \{s_2\} \notin E_G^\beta(\{2\})$$



---

## Coalitional Logic: Coalition Effectivity Models

An effectivity function is **playable** iff

1. For each  $C \subseteq N$ ,  $\emptyset \notin E(C)$
2. For each  $C \subseteq N$ ,  $Q \in E(C)$
3. If  $X \notin E(N)$ , then  $Q - X \in E(\emptyset)$
4. If  $X \subseteq Y$  and  $X \in E(C)$  then  $Y \in E(C)$
5. for all  $C_1, C_2 \subseteq N$  and  $X_1, X_2 \subseteq Q$ , if  $C_1 \cap C_2 = \emptyset$ ,  $X_1 \in E(C_1)$  and  $X_2 \in E(C_2)$  then  $X_1 \cap X_2 \in E(C_1 \cup C_2)$

**Characterization Theorem:** An  $\alpha$ -effectivity function  $E$  is playable iff it is the effectivity function of some strategic game.

M. Pauly. *Logics for Social Software*. Ph.D. Thesis, ILLC.

---

---

## Coalitional Logic: Coalition Effectivity Models

A **coalitional effectivity model** is a tuple  $\langle Q, E, \pi \rangle$  where  $E : Q \rightarrow (2^N \rightarrow 2^{2^Q})$  assigns a playable effectivity function to each state and  $\pi$  is a valuation function.

$$q \models [C]\phi \text{ iff } \|\phi\| \in E_q(C)$$



---

## Coalitional Logic: Axiomatization

- ( $\perp$ )  $\neg[C]\perp$
- (T)  $[C]T$
- (N)  $(\neg[\emptyset]\neg\phi \rightarrow [N]\phi)$
- (M)  $[C](\phi \wedge \psi) \rightarrow [C]\psi$
- (S)  $([C_1]\phi_1 \wedge [C_2]\phi_2) \rightarrow [C_1 \cup C_2](\phi_1 \wedge \phi_2)$   
provided  $C_1 \cap C_2 = \emptyset$

$$\frac{\phi \quad \phi \rightarrow \psi}{\psi}$$

$$\frac{\phi \leftrightarrow \psi}{[C]\phi \leftrightarrow [C]\psi}$$

---

---

**Theorem** Coalitional Logic is sound and strongly complete with respect to the class of effectivity models.

**Theorem** The complexity of the satisfiability problem of coalitional logic is PSPACE-complete.

M. Pauly. *A Modal Logic for Coalitional Powers in Games*. Journal of Logic and Computation (2002).

---

---

## Coalition Logic is a Fragment of ATL

Define  $[A]\phi$  to be  $\langle\langle A \rangle\rangle \bigcirc \phi$

Multi-player Game Models and Concurrent-game Models only differ in notation

Coalitional Effectivity Models can be used as a semantics for ATL

Goranko and Jamroga. *Comparing Semantics of Logics for Multi-Agent Systems*. See the website.

---

---

## ATL, Semantics I (Concurrent Game Structure)

A **Concurrent Game Structure** is a tuple  $\mathcal{S} = \langle n, Q, \Pi, \pi, d, \delta \rangle$   
with

- A natural number  $n \geq 1$  of players:  $\Sigma = \{1, \dots, n\}$
  - A set  $Q$  of states
  - A finite set  $\Pi$  of atomic propositions
  - For each  $q \in Q$ , a set  $\pi(q) \subseteq \Pi$  of atomic propositions true at  $q$ .
-

---

## ATL, Semantics I (Concurrent Game Structure)

A **Concurrent Game Structure** is a tuple  $S = \langle n, Q, \Pi, \pi, d, \delta \rangle$  with

- For each  $a \in \Sigma$  and each  $q \in Q$ , a natural number  $d_a(q) \geq 1$  of moves available at state  $q$  to player  $a$ . Moves of player  $a$  at  $q$  are identified with numbers  $1, \dots, d_a(q)$ .

A **move vector** at  $q$  is a tuple  $\langle j_1, \dots, j_k \rangle$  such that  $1 \leq j_a \leq d_a(q)$  for each player  $a$ .

Write  $D(q) = \{1, \dots, d_1(q)\} \times \dots \times \{1, \dots, d_n(q)\} \subseteq \mathbb{N}^n$ .

- For each  $q \in Q$  and each move vector  $\langle j_1, \dots, j_n \rangle \in D(q)$ , a state  $\delta(q, j_1, \dots, j_n) \in Q$  that results from the state  $q$  if each player  $a \in \{1, \dots, n\}$  chooses move  $j_a$ .  $\delta$  is called the **transition function**.
-

---

## Notation

- $q'$  is a **successor** of  $q$  iff  $\exists \langle j_1, \dots, j_n \rangle \in D(q)$  such that  $q' = \delta(q, j_1, \dots, j_n)$
  - A **computation** is an infinite sequence  $\lambda = q_0 q_1 \dots$  of states such that for all  $i \geq 0$ ,  $q_{i+1}$  is a successor of  $q_i$
  - If a computation  $\lambda = q_0 q_1 q_2 \dots$  starts at  $q$  (i.e.,  $q_0 = q$ ),  $\lambda$  is called a  $q$ -computation
  - Given a computation  $\lambda$ ,  $\lambda[i]$ ,  $\lambda[0, i]$ , and  $\lambda[i, \infty]$  denote respectively the  $i$ th state of  $\lambda$ , the finite prefix  $q_0 q_1 \dots q_i$  and the infinite suffix  $q_i q_{i+1} \dots$
-



---

## Strategies I

A **strategy** for player  $a \in \Sigma$  is a function  $f_a : Q^+ \rightarrow \mathbb{N}$  that maps every nonempty finite state sequence  $\lambda$  to a natural number such that if the last state is  $q$ ,  $1 \leq f_a(\lambda) \leq d_a(q)$ .

For a state  $q \in Q$  and  $A \subseteq \Sigma$  an  $A$ -move is a tuple  $(\sigma_a)_{a \in A}$  such that  $1 \leq \sigma_a \leq d_a(q)$ . Let  $D_A(q)$  denote the set of  $A$ -moves.

A state  $q'$  is **consistent** with an  $A$ -move  $\sigma \in D_A(q)$  when there is a move vector  $\langle j_1, \dots, j_n \rangle \in D(q)$  such that

1.  $j_a = \sigma_a$  for all  $a \in A$
  2.  $q' = \delta(q, j_1, \dots, j_n)$
-

---

## Strategies II

Let  $\text{out}(\sigma)$  denote the set of state consistent with  $\sigma$

For  $A \subseteq \Sigma$  an  $A$ -strategy is a mapping

$F_A : Q^+ \rightarrow \cup\{D_A(q) \mid q \in Q\}$  such that for all  $\lambda \in Q^*$ , and for all  $q \in Q$ ,  $F_A(\lambda \cdot q) \in D_A(q)$ .

A  $q$ -computation  $\lambda = q_0q_1\dots$  with  $q_0 = q$  is **consistent** with  $F_A$  written  $\lambda \in \text{out}(q, F_A)$  if for all  $i \geq 0$ ,  $q_{i+1}$  is a successor of  $q_i$  and  $q_{i+1} \in \text{out}(F_A(\lambda[0, i]))$ .

---

---

## Definition of Truth

- For  $p \in \Pi$ ,  $q \models p$  iff  $p \in \pi(q)$
  - $q \models \neg\phi$  iff  $q \not\models \phi$
  - $q \models \phi \vee \psi$  iff  $q \models \phi$  or  $q \models \psi$
  - $q \models \langle\langle A \rangle\rangle \bigcirc \phi$  iff there exists an  $A$ -strategy  $F_A$  such that for each computation  $\lambda \in \text{out}(q, F_A)$  we have  $\lambda[1] \models \phi$ ;  
equivalently, iff there exists an  $A$ -move  $\sigma \in D_A(q)$  such that for all  $q' \in \text{out}(\sigma)$ ,  $q' \models \phi$ .
  - $q \models \langle\langle A \rangle\rangle \Box \phi$  iff there exists an  $A$ -strategy  $F_A$  such that for each computation  $\lambda \in \text{out}(q, F_A)$  and all  $i \geq 0$ , we have  $\lambda[i] \models \phi$
  - $q \models \langle\langle A \rangle\rangle \phi U \psi$  iff there exists an  $A$ -strategy  $F_A$  such that for each computation  $\lambda \in \text{out}(q, F_A)$  there exists a position  $i \geq 0$  such that  $\lambda[i] \models \psi$  and for all  $0 \leq j < i$  we have  $\lambda[j] \models \phi$
-

---

## Definition of Truth

- For  $p \in \Pi$ ,  $q \models p$  iff  $p \in \pi(q)$
  - $q \models \neg\phi$  iff  $q \not\models \phi$
  - $q \models \phi \vee \psi$  iff  $q \models \phi$  or  $q \models \psi$
  - $q \models \langle\langle A \rangle\rangle \bigcirc \phi$  iff there exists an  $A$ -strategy  $F_A$  such that for each computation  $\lambda \in \text{out}(q, F_A)$  we have  $\lambda[1] \models \phi$ ;  
equivalently, iff there exists an  $A$ -move  $\sigma \in D_A(q)$  such that for all  $q' \in \text{out}(\sigma)$ ,  $q' \models \phi$ .
  - $q \models \langle\langle A \rangle\rangle \Box \phi$  iff there exists an  $A$ -strategy  $F_A$  such that for each computation  $\lambda \in \text{out}(q, F_A)$  and all  $i \geq 0$ , we have  $\lambda[i] \models \phi$
  - $q \models \langle\langle A \rangle\rangle \phi U \psi$  iff there exists an  $A$ -strategy  $F_A$  such that for each computation  $\lambda \in \text{out}(q, F_A)$  there exists a position  $i \geq 0$  such that  $\lambda[i] \models \psi$  and for all  $0 \leq j < i$  we have  $\lambda[j] \models \phi$
-

---

## Definition of Truth

- For  $p \in \Pi$ ,  $q \models p$  iff  $p \in \pi(q)$
  - $q \models \neg\phi$  iff  $q \not\models \phi$
  - $q \models \phi \vee \psi$  iff  $q \models \phi$  or  $q \models \psi$
  - $q \models \langle\langle A \rangle\rangle \bigcirc \phi$  iff there exists an  $A$ -strategy  $F_A$  such that for each computation  $\lambda \in \text{out}(q, F_A)$  we have  $\lambda[1] \models \phi$ ;  
equivalently, iff there exists an  $A$ -move  $\sigma \in D_A(q)$  such that for all  $q' \in \text{out}(\sigma)$ ,  $q' \models \phi$ .
  - $q \models \langle\langle A \rangle\rangle \Box \phi$  iff there exists an  $A$ -strategy  $F_A$  such that for each computation  $\lambda \in \text{out}(q, F_A)$  and all  $i \geq 0$ , we have  $\lambda[i] \models \phi$
  - $q \models \langle\langle A \rangle\rangle \phi U \psi$  iff there exists an  $A$ -strategy  $F_A$  such that for each computation  $\lambda \in \text{out}(q, F_A)$  there exists a position  $i \geq 0$  such that  $\lambda[i] \models \psi$  and for all  $0 \leq j < i$  we have  $\lambda[j] \models \phi$
-

---

## Definition of Truth

- For  $p \in \Pi$ ,  $q \models p$  iff  $p \in \pi(q)$
  - $q \models \neg\phi$  iff  $q \not\models \phi$
  - $q \models \phi \vee \psi$  iff  $q \models \phi$  or  $q \models \psi$
  - $q \models \langle\langle A \rangle\rangle \bigcirc \phi$  iff there exists an  $A$ -strategy  $F_A$  such that for each computation  $\lambda \in \text{out}(q, F_A)$  we have  $\lambda[1] \models \phi$ ;  
equivalently, iff there exists an  $A$ -move  $\sigma \in D_A(q)$  such that for all  $q' \in \text{out}(\sigma)$ ,  $q' \models \phi$ .
  - $q \models \langle\langle A \rangle\rangle \Box \phi$  iff there exists an  $A$ -strategy  $F_A$  such that for each computation  $\lambda \in \text{out}(q, F_A)$  and all  $i \geq 0$ , we have  $\lambda[i] \models \phi$
  - $q \models \langle\langle A \rangle\rangle \phi U \psi$  iff there exists an  $A$ -strategy  $F_A$  such that for each computation  $\lambda \in \text{out}(q, F_A)$  there exists a position  $i \geq 0$  such that  $\lambda[i] \models \psi$  and for all  $0 \leq j < i$  we have  $\lambda[j] \models \phi$
-

---

## Alternative Semantics for ATL: Alternating Transition Systems

An **alternating transition system** is a tuple  $\langle \Pi, N, Q, \pi, \delta \rangle$  where

- $\Pi$  is a set of atomic propositions
  - $N$  is a set of agents (nonempty and finite)
  - $Q$  is a set of states (nonempty)
  - $\pi : Q \rightarrow 2^\Pi$  is a valuation function
  - $\delta : Q \times N \rightarrow 2^{2^Q}$  is a **transition function**. For each  $i \in N$ ,  $Q_i \in \delta(q, i)$  means  $i$  can force the state to be in  $Q_i$ . We assume that for each that the system is deterministic: given a state  $q$  and the agents choices in  $q$ ,  $Q_1, \dots, Q_n$ ,  $Q_1 \cap \dots \cap Q_n$  is a singleton.
-

---

## Axiomatization

### Coalitional Logic Axioms:

(TAUT) Enough propositional tautologies

$$(\perp) \neg \langle\langle A \rangle\rangle \circ \perp$$

$$(\top) \langle\langle A \rangle\rangle \circ \top$$

$$(\Sigma) \neg \langle\langle \emptyset \rangle\rangle \circ \neg \phi \rightarrow \langle\langle \Sigma \rangle\rangle \circ \phi$$

$$(\text{S}) \langle\langle A_1 \rangle\rangle \circ \phi_1 \wedge \langle\langle A_2 \rangle\rangle \circ \phi_2 \rightarrow \langle\langle A_1 \cup A_2 \rangle\rangle \circ (\phi_1 \wedge \phi_2)$$

---



---

## Axiomatization

### Fixed Point Axioms:

$$(\text{FP}_{\Box}) \langle\langle A \rangle\rangle \Box \phi \leftrightarrow \phi \wedge \langle\langle A \rangle\rangle \Box \phi$$

$$(\text{GFP}_{\Box}) \langle\langle \emptyset \rangle\rangle \Box (\theta \rightarrow (\phi \wedge \langle\langle A \rangle\rangle \Box \theta)) \rightarrow \langle\langle \emptyset \rangle\rangle \Box (\theta \rightarrow \langle\langle A \rangle\rangle \Box \phi)$$

$$(\text{FP}_{\cup}) \langle\langle A \rangle\rangle \phi \cup \psi \leftrightarrow \psi \vee (\phi \wedge \langle\langle A \rangle\rangle \Box \langle\langle A \rangle\rangle \phi \cup \psi)$$

$$(\text{LFP}_{\cup}) \langle\langle \emptyset \rangle\rangle \Box ((\psi \vee (\phi \wedge \langle\langle A \rangle\rangle \Box \theta)) \rightarrow \theta) \rightarrow \langle\langle \emptyset \rangle\rangle \Box (\langle\langle A \rangle\rangle \phi \cup \psi \rightarrow \theta)$$

---

---

## Axiomatization

### Rules:

$$\text{(Modus Ponens)} \quad \frac{\phi \quad \phi \rightarrow \psi}{\psi}$$

$$\text{(\langle\langle A \rangle\rangle\langle\langle \text{O} \rangle\rangle\text{-Monotonicity})} \quad \frac{\phi \rightarrow \psi}{\langle\langle A \rangle\rangle\langle\langle \text{O} \rangle\rangle\phi \rightarrow \langle\langle A \rangle\rangle\langle\langle \text{O} \rangle\rangle\psi}$$

$$\text{(\langle\langle A \rangle\rangle\langle\langle \square \rangle\rangle\text{-Necessitation})} \quad \frac{\phi}{\langle\langle A \rangle\rangle\langle\langle \square \rangle\rangle\phi}$$

---

---

## Results

**Theorem** All of the semantics (concurrent game structures, alternating transitions systems and coalitional effectivity models) are equivalent.

Goranko and Jamroga. *Comparing Semantics of Logics for Multi-Agent Systems*. See the website.

**Theorem** ATL is sound and (weakly) complete.

**Theorem** Given a finite set of players, the satisfiability problem for ATL-formulas over  $N$  with respect to concurrent game structures over  $N$  is EXPTIME-complete.

Goranko and van Drimmelen. *Complete Axiomatization and Decidability of the Alternating-Time Temporal Logic*. Theoretical Computer Science (2005).

---

---

## ATEL

Various attempts have been made in order to introduce imperfect information into these models. The “correct” semantics is still being debated.

---

---

## ATEL

Various attempts have been made in order to introduce imperfect information into these models. The “correct” semantics is still being debated.

Some issues

- Expressions of the form  $\langle\langle C \rangle\rangle \Box p$  presupposes that the group  $C$  is “aware” of  $p$ .
-

---

## ATEL

Various attempts have been made in order to introduce imperfect information into these models. The “correct” semantics is still being debated.

Some issues

- Expressions of the form  $\langle\langle C \rangle\rangle \Box p$  presupposes that the group  $C$  is “aware” of  $p$ .
- By adding knowledge operators, we can express interesting properties:

$$\langle\langle \{1, 2\} \rangle\rangle (\neg K_3 p \ U \ K_2 p)$$

---

---

## ATEL

Various attempts have been made in order to introduce imperfect information into these models. The “correct” semantics is still being debated.

Some issues

- Expressions of the form  $\langle\langle C \rangle\rangle \Box p$  presupposes that the group  $C$  is “aware” of  $p$ .
- By adding knowledge operators, we can express interesting properties:

$$\langle\langle \{1, 2\} \rangle\rangle (\neg K_3 p \ U \ K_2 p)$$

- *de re* vs. *de dicto* knowledge of strategies:
    - agent  $i$  knows there is a winning strategy to achieve  $\phi$
-

---

– there is a winning strategy which  $i$  knows will achieve  $\phi$

---



---

## Where do the coalitions come from?

There is a lot of work on cooperative game theory which attempts to describe *why* agents form a coalition.

One possible way is via aggregation procedures:

M. Pauly. *Axiomatising Judgement Aggregation Procedures in a Minimal Logical Language*. forthcoming in KRA.

M. Pauly. *On the Role of Language in Social Choice Theory*. unpublished.

---

---

## First, A Paradox (Kornhauser and Sager 1993)

*P*: a valid contract was in place

*Q*: the defendant's behavior was such as to breach a contract of that kind

*R*: the court is required to find the defendant liable.

|   | <i>P</i> | <i>Q</i> | $(P \wedge Q) \leftrightarrow R$ | <i>R</i> |
|---|----------|----------|----------------------------------|----------|
| 1 | yes      | yes      | yes                              | yes      |
| 2 | yes      | no       | yes                              | no       |
| 3 | no       | yes      | yes                              | no       |



---

Should we accept  $R$ ?

|   | $P$ | $Q$ | $(P \wedge Q) \leftrightarrow R$ | $R$ |
|---|-----|-----|----------------------------------|-----|
| 1 | yes | yes | yes                              | yes |
| 2 | yes | no  | yes                              | no  |
| 3 | no  | yes | yes                              | no  |

---

Should we accept  $R$ ? No, a simple majority votes no.

|   | $P$ | $Q$ | $(P \wedge Q) \leftrightarrow R$ | $R$ |
|---|-----|-----|----------------------------------|-----|
| 1 | yes | yes | yes                              | yes |
| 2 | yes | no  | yes                              | no  |
| 3 | no  | yes | yes                              | no  |

---

Should we accept  $R$ ? Yes, a majority votes yes for  $P$  and  $Q$  and  $(P \wedge Q) \leftrightarrow R$  is a legal doctrine.

|   | $P$ | $Q$ | $(P \wedge Q) \leftrightarrow R$ | $R$ |
|---|-----|-----|----------------------------------|-----|
| 1 | yes | yes | yes                              | yes |
| 2 | yes | no  | yes                              | no  |
| 3 | no  | yes | yes                              | no  |

---

## List and Pettit Impossibility Result

Suppose there are  $n$  agents and let  $\mathcal{L}$  be a propositional language.

**Personal judgement sets:** a *consistent, complete* and *deductively closed* set of formulas — a maximally consistent set.

**A collective judgement aggregation function:** Let  $\mathbb{M} = \{\Gamma \mid \Gamma \text{ is a maximally consistent set}\}$  then a collective aggregation function is defined as follows:

$$F : \mathbb{M}^n \rightarrow \mathbb{M}$$

---

---

## Some Conditions

**Universal Domain**  $F$  is a total function

**Anonymity** For all  $\vec{\Gamma} \in \mathbb{M}^n$ ,  $F(\Gamma_1, \dots, \Gamma_n) = F(\Gamma_{\pi(1)}, \dots, \Gamma_{\pi(n)})$   
for all permutations  $\pi$

**Systematicity** There exists a function  $f : \{0, 1\}^n \rightarrow \{0, 1\}$  such that for any  $\vec{\Gamma} \in \mathbb{M}^n$ ,  
 $F(\Gamma_1, \dots, \Gamma_n) = \{ \phi \in X \mid f(\delta_1(\phi), \dots, \delta_n(\phi)) = 1 \}$ , where, for each agent  $i$  and each  $\phi \in X$ ,  $\delta_i(\phi) = 1$  if  $\phi \in \Gamma_i$  and  $\delta_i(\phi) = 0$  if  $\phi \notin \Gamma_i$

---

---

**Theorem** [List and Pettit, 2001] There exists no judgement aggregation function generating complete, consistent and deductively closed collective sets of judgements which satisfies Universal Domain, Anonymity and Systematicity.

---



---

## Brief Survey of the Literature

- See [personal.lse.ac.uk/LIST/doctrinalparadox.htm](http://personal.lse.ac.uk/LIST/doctrinalparadox.htm) for a detailed overview of the current state of affairs. Some highlights:
  - Other impossibility results: Pauly and van Hees (2003), van Hees (2004), Gärdenfors (2004), among others
  - List and Pettit, 2005 compare their impossibility result with Arrow's Theorem
-

---

## A Logic for Aggregation Procedures

Let  $\Phi_I$  be the set of **individual formulas** (standard propositional language)

$V_I$  the set of individual valuations

$\Phi_C$  the set of **collective formulas** (standard modal language)

$\Box\phi$ : *The group collectively accepts  $\phi$ .*

---

---

**Conclusion Driven Procedure:**  $\Box\phi \wedge \Box\psi$

**Premise Driven Procedure:**  $\Box(\phi \wedge \psi)$

For majority voting,  $\Box\phi \wedge \Box\psi \rightarrow \Box(\phi \wedge \psi)$  but *not conversely*.

---

---

Let  $\Phi_{\square I} = \{\square\phi \mid \phi \in \Phi_I\}$ , a collective valuation is a function  $v : \Phi_{\square I} \rightarrow \{0, 1\}$

A **decision method** is a function  $D : \{0, 1\}^n \rightarrow \{0, 1\}$

**Dictatorship:** there is a  $d \in N$  such that  $D(x_1, \dots, x_n) = x_d$  for all  $x_1, \dots, x_n \in D$

**Consensus:**  $D(x_1, \dots, x_n) = 1$  iff  $x_i = 1$  for all  $i$

**Majority:**  $D(x_1, \dots, x_n) = 1$  iff  $|\{i \mid x_i = 1\}| > \frac{1}{2}|N|$

---

---

A model  $v \in V_C$  is  **$n$ -systematic** iff there is some  $n$ -ary decision method  $D$  and there are  $n$  individual valuations  $v_1, \dots, v_n \in V_I$  such that for all  $\phi \in \Phi_I$ ,  $v(\Box\phi) = D(v_1(\phi), \dots, v_n(\phi))$

Similarly for the other decision procedures

Connection with simple games.

---

---

$E = \{\Box\phi \rightarrow \Box\psi \mid \phi \rightarrow \psi \in \Phi_I \text{ is a tautology}\}$

**Theorem** (Pauly, 2005)  $V_C(E) = \mathcal{SYS}_n$ , provided  $n \geq 2^{|\Phi_0|}$ .

---

---

KD extends E with the following axiom schemes:

M.  $\Box(\phi \wedge \psi) \rightarrow (\Box\phi \wedge \Box\psi)$

C.  $(\Box\phi \wedge \Box\psi) \rightarrow (\Box\phi \wedge \Box\psi)$

N.  $\Box\top$

D.  $\neg\Box\perp$

**Theorem** (Pauly, 2005)  $V_{\mathcal{C}}(\text{KD}) = \text{CON}_n$ , provided  $n \geq 2^{|\Phi_0|}$ .

---

---

MCY contains  $M$ ,  $C$  and the following scheme:

$$Y. \neg \Box \phi \leftrightarrow \Box \neg \phi$$

**Theorem** (Pauly, 2005)  $V_C(\text{MCY}) = \text{DIC}$ .

---



---

STEM contains E, M and all instances of the following schemes

S.  $[>]\phi \rightarrow \neg[>]\neg\phi$

T.  $([\geq]\phi_1 \wedge \dots \wedge [\geq]\phi_k \wedge [\leq]\psi_1 \wedge \dots \wedge [\leq]\psi_k) \rightarrow \bigwedge_{1 \leq i \leq k} ([=]\phi_i \wedge [=]\psi_i)$   
where  $\forall v \in V_I : |\{i \mid v(\phi_i) = 1\}| = |\{i \mid v(\psi_i) = 1\}|$

**Theorem** (Pauly, 2005)  $V_C(\text{STEM}) = \mathcal{MAJ}$ . (Nontrivial!!)

---

---

Next Week: Class Canceled, sort of — I will speak at the Segerberg  
Celebration

---