Logics for Agency and Multi-Agent systems

What agents do, what agents know they (can) do

Jan Broersen, Andreas Herzig, *Nicolas Troquard*

ESSLLI'07, Dublin, Wednesday, August 15, 2007

Course overview

- Monsday: introduction to modal logic
- Tuesday: what agent can do
  - logic of ability: Coalition Logic
- Wednesday: what agents do; what agents know they (can) do
  - logic of agency: STIT
- Thursday: what agents want
  - logic of intention: Cohen and Levesque
- Friday: what agents can plan
  - ATL, strategic STIT

Today outline

1. Logical frameworks for MAS and agency
2. Time and Choice aspects of $BT + AC$ models
3. Reasoning about individual choice
4. Reasoning about coalitional choice
5. Application to uniform choices

Logics for social choice and MAS

- Social software [Parikh 2001]: constructing and verifying social procedures
  - Traditional problems of Game Theory (Prisoner's Dilemma, Battle of the sexes...)
  - Preference aggregation (voting procedures...)
  - Judgement aggregation (possibly inconsistent 'multi-expert' systems...)
  - Auction procedures (Ebay...)
  - Fair division problems (cake cutting problem, social welfare...)
- Logics for social software: Coalition Logic [Pauly 2001] $\Longrightarrow$ **Tuesday**
- Logics of cooperation in Computer Science: ATL [Alur, Henzinger, Kupferman 1997] $\Longrightarrow$ **Friday**

Philosophy of action

- *Ontological view*: Davidson, Thomson, Anscombe...
- *Modal view*: Kanger, Chellas, von Kutschera, Belnap and Perloff, Segerberg...

The ontological view

- An action (token) is a particular event
  - + courses over time
  - + participant/actor
  - + intention
  - + ...
- Plural action: John and Mary going upstairs
  $\Longrightarrow$ individual intentions
- Collective action: John and Mary lifting the table
  $\Longrightarrow$ coordination and collective intention

The modal view

- Dynamic Logic [Pratt 1976]
  - Logics of programs: actions are given explicitly
  - Originally designed to explain program verification and Floyd-Hoare logic [Hoare 1969]
- Logics of agency
  - Actions are abstract, and identified with what they cause
  - First semantics [Chellas 1969]

$\Longrightarrow$ Two main families:

  - bringing it about: [Kanger 1972], [Pörn 1977]
  - seeing to it that: [Belnap, Perloff 1988], [Horty, Belnap 1995]...

Pörn's account

- $D_a\varphi$ is true at a world w if φ is true at every hypothetical situation where agent a *"does at least as much as he does in w"*
- $D'_a\varphi$ is true in w if $\neg\varphi$ is true in every hypothetical situation w' such that *"the opposite of everything a does in w is the case in w'"*
- Combination of two normal operators in a non-normal modality:
  - $D_a\varphi$ reading "it is necessary for something a does that φ"
  - $D'_a\varphi$ reading "but for a's action, it would not be the case that φ"
  - $E_a\varphi \triangleq D_a\varphi \wedge \neg D'_a\neg\varphi$ reads "agent a brings it about that φ".

A controversial framework

- (Sometimes adapted and) used a lot for modeling institutionalized power and law [Jones, Sergot 1996], [Royakkers 2000], [Carmo, Pacheco 2001], but...
- “one problem with the proposed semantics is that ‘doing at least as much as’ he does in [a world], and the notion of an agent doing ‘the opposite’ of everything he does in [a world], are of dubious intelligibility without substantial further elucidation, and Pörn offers none.” [Horgan 1979]
- “the intuitive significance of this semantics is not altogether clear” [Segerberg 1992]

Belnap and Perloff's STIT theory

Problem: distinguish between sentences which involve agency and those which do not.

The Moby Dick example [Belnap and Perloff 1988]:

- Is "Ishmael sails on board the Pequod" agentive for Ishmael?
  - try to uncover general principles for deciding agentivity of sentences
- An agentive sentence must emphasize a sort of causality and responsability of an agent for the truth of a state of affairs.
  - For Ishmael being agentive for sailing on the Pequod, there should be a choice by Ishmael which permitted it.
  - E.g. he chose deliberatively to engage on the Pequod to break out of his depressive cycle.

STIT paraphrase thesis

Definition (STIT paraphrase thesis [Belnap and Perloff 1988])

The sentence φ marks the agentiveness of agent *a* just in case φ may be usefully paraphrased as "*a* sees to it that φ".

This way, deciding whether the sentence

"*Ishmael* sails on board the Pequod"

is agentive for Ishmael, is deciding whether it is equivalent to

"*Ishmael* sees to it that *Ishmael* sails on board the Pequod"

STIT logics

Also called “the theory of agents and choices in branching time”.

- John puts the cube on the table

  $\Longrightarrow$ John sees to it that the cube is on the table.
- Notation:
  - original: [John *stit* : *cubeOnTable*]
  - here: [John]*cubeOnTable*
- Several logics
  - Achievement stit [Belnap and Perloff 1988]
  - Deliberative stit, Chellas stit [Horty and Belnap 1995] (also [von Kutschera 1986])
  - Strategic stit [Horty 2001], [Belnap et al. 2001]
  - ...

Content

We are interested in the formal framework behind this linguistic agenda.

- We present STIT models that might seem quite odd to the logician in Computer Science.
- We introduce a more CS friendly presentation $\Longrightarrow$ Kripke models
- We present applications of STIT to new trends of research at the intersection of CS and Game Theory
  - Group choices $\Longrightarrow$ as Coalition Logic
  - (Epistemic) uniform choices $\Longrightarrow$ problem of "knowing how to play"

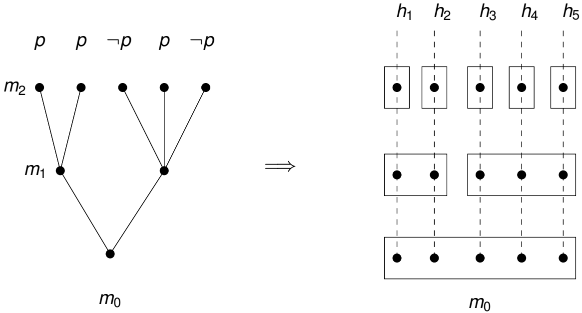
Today outline

“I shall finish this lecture”

- Controversial evaluation of future-tense sentences in branching time
  - [Thomason 1984] talks about a truth value gap
- [Prior 1967] opposes:
  - Peircean temporal logic
  - Ockhamist/actualist temporal logic
- “I shall finish this lecture” is true whenever “I finish this lecture” is true at some moment on the *actual* history
  - Truth in a tree-structure should in general be relative to moment/history pairs

Ockhamist branching time temporal logic (BT)

BT structure $\langle Mom, < \rangle$:



- History = maximally $<$-ordered set of moments
- $Hist$ = set of all histories
- H_m = set of histories passing through the moment m
- Explode moments into contexts (moment/history pairs)
  - $m_0/h_3 \not\models \mathbf{F}p$
  - $m_0/h_1 \models \mathbf{F}p$

$BT + AC$ models

A *$BT + AC$ model* is a tuple $\mathcal{M} = \langle Mom, <, Choice, v \rangle$, where:

- $\langle Mom, < \rangle$ is a BT structure;
- $Choice : \mathcal{A}gt \times Mom \to 2^{2^{Hist}}$ is a function mapping each agent and each moment m into a partition of H_m;
- v is valuation function $v : \mathcal{A}tm \to 2^{Mom \times Hist}$.

Choice

Choice is the most fundamental primitive of $BT + AC$ structures.

- $Choice_a^m$ = repertoire of choices for agent a at moment m
- For $h \in H_m$: $Choice_a^m(h)$ = the particular choice of a at context m/h.

We need to further specify it.

Definition

A function s_m from $\mathcal{A}gt$ into 2^{H_m} such that for each $m \in Mom$ and $a \in \mathcal{A}gt$, $s_m(a) \in Choice_a^m$ is a *selection function*.

For a given m, $Select_m$ is the set of all selection functions s_m.

Choice: independence and groups

Assumption (independence of agents)

For every $s_m \in \mathit{Select}_m$, $\bigcap_{a \in \mathcal{A}gt} s_m(a) \neq \emptyset$.

- More assumptions: Liveness, No choice between undivided histories...

Generalizing to coalitions: $\mathit{Choice} : 2^{\mathcal{A}gt} \times \mathit{Mom} \to 2^{2^{\mathit{Hist}}}$.

Definition (choice function for a group of agents)

For $A \subseteq \mathcal{A}gt$,

$$\mathit{Choice}^m_A = \{\bigcap_{a \in A} s_m(a) | s_m \in \mathit{Select}_m\}$$

Truth values

A formula is evaluated with respect to a model and a context:

$$
\begin{array}{lcl}
\mathcal{M}, m/h \models p & \iff & m/h \in v(p), p \in \mathcal{A}tm. \\
\mathcal{M}, m/h \models \neg\varphi & \iff & \mathcal{M}, m/h \not\models \varphi \\
\mathcal{M}, m/h \models \varphi \vee \psi & \iff & \mathcal{M}, m/h \models \varphi \text{ or } \mathcal{M}, m/h \models \psi \\
\\
\mathcal{M}, m/h \models \mathbf{F}\varphi & \iff & \exists m',\ m < m',\ \mathcal{M}, m'/h \models \varphi
\end{array}
$$

Truth values

- Historical necessity: $\Box\varphi$ = "whatever happens, φ is true at the current moment"

$$\mathcal{M}, m/h \models \Box\varphi \iff \mathcal{M}, m/h' \models \varphi,\ \forall h' \in H_m$$

- Chellas's stit: $[A]\varphi$ = "A sees to it that φ", "the alternative that is presently and actually chosen by A guarantees that φ is true".

$$\mathcal{M}, m/h \models [A]\varphi \iff \mathcal{M}, m/h' \models \varphi,\ \forall h' \in Choice_A^m(h)$$

- $\Diamond\varphi$ = "φ is historically possible"
- $\Diamond[\{Ann, Bill\}]\mathbf{F}lightOn$ = "Ann and Bill can ensure that the light is eventually on"

A discrete-deterministic STIT

Hypothesis (discreteness)

Given a moment m_1, there exists a successor moment m_2 such that $m_1 < m_2$ and there is no moment m_3 such that $m_1 < m_3 < m_2$.

$m/h \models \mathbf{X}\varphi$ iff φ is true at the moment immediately after m on h

Hypothesis (determinism)

$\forall m \in Mom, \exists m' \in Mom\,(m < m'$ *and* $\forall h \in m', \; Choice^{m}_{\mathcal{A}gt}(h) = H_{m'})$

Translation of Coalition Logic to discrete-deterministic STIT

$$
\begin{array}{lcl}
tr(p) & = & \Box p, \text{ for } p \in \mathcal{A}tm \\
tr(\neg\varphi) & = & \neg tr(\varphi) \\
tr(\varphi \vee \psi) & = & tr(\varphi) \vee tr(\psi) \\
tr(\langle\![J]\!\rangle \mathbf{X}\varphi) & = & \Diamond [J]\mathbf{X} tr(\varphi)
\end{array}
$$

In STIT terminology

"the coalition J is able to ensure φ"

can be paraphrased by

"it is historically possible that J sees to it that next φ"

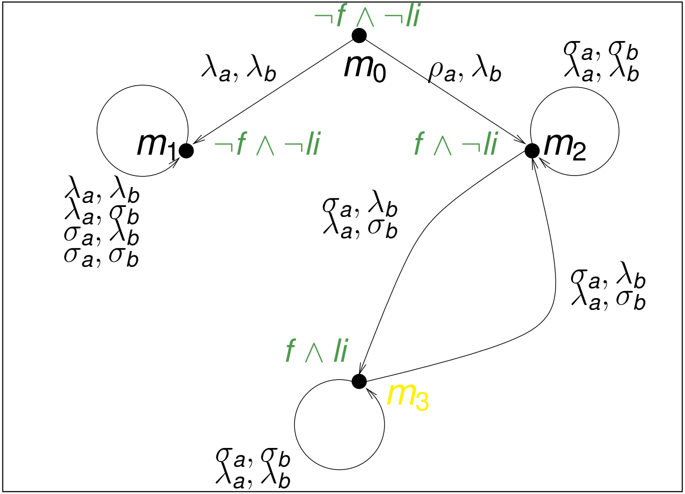
Theorem ([Broersen, Herzig, Troquard 2006])

tr is a correct embedding of CL into discrete-deterministic STIT.

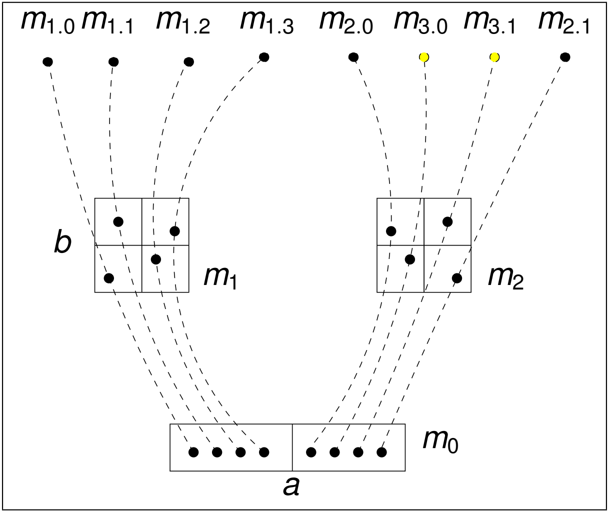
Example: Ann and Bill switch the light

- Four states: m_0, m_1, m_2, m_3
- li = light is on (at m_3)
- f = lamp is functioning (at m_2 and m_3)
- At moment m_0, agent a has the choice between *repairing* a broken lamp (ρ_a) or *remaining passive* (λ_a). Agent b has the vacuous choice of *remaining passive* (λ_b).
- If a chooses not to repair, the system reaches m_1. If a chooses to repair, the system reaches m_2.
- In m_1, m_2 and m_3 both a and b can choose to *toggle* a light switch (τ_a and τ_b) or *not toggle* (λ_a and λ_b).
- If a repairs at m_0 then a and b 'play toggling' between m_2 and m_3

Game model


$\neg f \wedge \neg li$
m_0
λ_a, λ_b
ρ_a, λ_b
σ_a, σ_b
λ_a, λ_b
m_1
$\neg f \wedge \neg li$
$f \wedge \neg li$
m_2
λ_a, λ_b
λ_a, σ_b
σ_a, λ_b
σ_a, σ_b
σ_a, λ_b
λ_a, σ_b
σ_a, λ_b
λ_a, σ_b
$f \wedge li$
m_3
σ_a, σ_b
λ_a, λ_b

Corresponding STIT model



So far so good

What is this fragment of discrete-deterministic STIT?

- Axiomatics
- Decidability
- Complexity

Let's see!

Today outline

Reminder of $S5$ modal logic [Lewis, Langford 1932]

- $S5$ is characterized by equivalence frames (reflexive, transitive, and symmetrical).
- Axiomatics: (K, T, 4, B), (K, D, T, 4, 5)...

K $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$
T $\Box\varphi \to \varphi$
4 $\Box\varphi \to \Box\Box\varphi$
5 $\Diamond\varphi \to \Box\Diamond\varphi$
B $\varphi \to \Box\Diamond\varphi$
D $\Box\varphi \to \Diamond\varphi$

Lemma

$A_1 A_2 \ldots A_k \varphi \leftrightarrow A_k \varphi$, $A_i \in \{\Box, \Diamond\}$.

Xu's *Ldm* axiomatics of *individual* Chellas stit

Convenient notation:

- $[i]\varphi$ instead of $[\{i\}]\varphi$

S5($\Box$)	axiom schemas of S5 for $\Box$
S5($[i]$)	axiom schemas of S5 for every $[i]$
($\Box \to [i]$)	$\Box\varphi \to [i]\varphi$
(AIA_k)	$(\Diamond[0]\varphi_0 \wedge \ldots \wedge \Diamond[k]\varphi_k) \to \Diamond([0]\varphi_0 \wedge \ldots \wedge [k]\varphi_k)$

Theorem ([Xu 1994])

Ldm is sound and complete w.r.t. BT + AC models.

A convenient truth

Clearly via the semantics and the completeness theorem:

$$\vdash [1][0]\varphi \rightarrow \Box\varphi$$

- Advanced (?) problem: derive it from *Ldm*.
  We do not know the solution.

The other way round holds too!

- Simple exercise: derive it from *Ldm*.

Then

- $\vdash \Box\varphi \leftrightarrow [1][0]\varphi$
- we can get rid off the $\Box$ operator!

A convenient truth

Clearly via the semantics and the completeness theorem:

$$\vdash [1][0]\varphi \rightarrow \Box\varphi$$

- Advanced (?) problem: derive it from *Ldm*.
  We do not know the solution.

The other way round holds too!

- Simple exercise: derive it from *Ldm*.

Then

- $\vdash \Box\varphi \leftrightarrow [1][0]\varphi$
- we can get rid off the $\Box$ operator!

Alternative *Ldm*

- Independence of agents in *Ldm*: (AIA_k)
  $\Diamond[0]\varphi_0 \wedge \ldots \Diamond[k]\varphi_k \rightarrow \Diamond([0]\varphi_0 \ldots [k]\varphi_k)$
- Alternative axiomatization of *Ldm*
  [Balbiani, Herzig, Troquard 2007]:

S5(i)	enough S5-theorems, for every [i]
Def($\Box$)	$\Box\varphi \leftrightarrow [1][0]\varphi$
(GPerm_k)	$\langle l\rangle\langle m\rangle\varphi \rightarrow \langle n\rangle \bigwedge_{i\in Agt\setminus\{n\}}\langle i\rangle\varphi$

- (GPerm_k) captures independence of agents

Alternative semantics

All axiom schemes are in the Sahlqvist class, and therefore have a standard possible worlds semantics.

Kripke models are of the form $M = \langle W, R, V \rangle$, where

- W is a nonempty set of possible worlds;
- R is a mapping associating to every $i \in \mathcal{A}gt$ an equivalence relation R_i on W;
- V is a mapping from $\mathcal{A}tm$ to the set of subsets of W.

We impose that R satisfies the general permutation property.

Alternative semantics (ctd.)

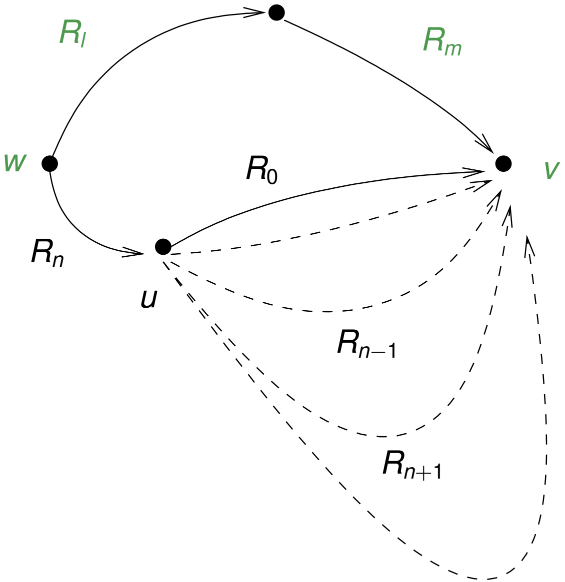
Definition (general permutation property)

R satisfies the *general permutation property* iff:
for all $w, v \in W$ and for all $l, m, n \in \mathcal{A}gt$, if $\langle w, v\rangle \in R_l \circ R_m$ then there is $u \in W$ such that: $\langle w, u\rangle \in R_n$ and $\langle u, v\rangle \in R_i$ for every $i \in \mathcal{A}gt \setminus \{n\}$.

We have the usual truth condition:

$$M, w \models [i]\varphi \text{ iff } M, u \models \varphi \text{ for every } u \text{ such that } \langle w, u\rangle \in R_i$$

Alternative semantics (illustration)



Link with product logic and complexity

If $\mathcal{A}gt = \{0, 1\}$ then the validities are axiomatized by:

- Def($\Box$): $\Box\varphi \leftrightarrow [1][0]\varphi$
- S5(0)
- S5(1)
- (GPerm_1), two instances:
  - $\langle 1\rangle\langle 0\rangle\varphi \rightarrow \langle 0\rangle\langle 1\rangle\varphi$
  - $\langle 0\rangle\langle 1\rangle\varphi \rightarrow \langle 1\rangle\langle 0\rangle\varphi$

Moreover,

- the permutation axiom $\langle 1\rangle\langle 0\rangle\varphi \leftrightarrow \langle 0\rangle\langle 1\rangle\varphi$
- Church-Rosser axioms $\langle 0\rangle[1]\varphi \rightarrow [1]\langle 0\rangle\varphi$, $\langle 1\rangle[0]\varphi \rightarrow [0]\langle 1\rangle\varphi$

can be proved.

Proof of Church-Rosser

1. $\langle 0\rangle\langle 1\rangle[1]\varphi \to \langle 1\rangle\langle 0\rangle[1]\varphi$ (GPerm$_1$)
2. $\langle 0\rangle[1]\varphi \to \langle 1\rangle\langle 0\rangle[1]\varphi$ (S5(1))

3. $\langle 1\rangle\langle 0\rangle[1]\varphi \to \langle 0\rangle\langle 1\rangle[1]\varphi$ (GPerm$_1$)
4. $[1]\langle 1\rangle\langle 0\rangle[1]\varphi \to [1]\langle 0\rangle\langle 1\rangle[1]\varphi$ (K(1))
5. $\langle 1\rangle\langle 0\rangle[1]\varphi \to [1]\langle 0\rangle\langle 1\rangle[1]\varphi$ (S5(1))
6. $\langle 1\rangle\langle 0\rangle[1]\varphi \to [1]\langle 0\rangle[1]\varphi$ (S5(1))
7. $\langle 1\rangle\langle 0\rangle[1]\varphi \to [1]\langle 0\rangle\varphi$ (S5(1))

8. $\langle 0\rangle[1]\varphi \to [1]\langle 0\rangle\varphi$ (From 3 and 7)

Link with product logic and complexity

- Hence the logic of the two-agent *Ldm* is nothing but the product $S5^2 = S5 \otimes S5$ [Marx 1999], [Gabbay et al. 2003] (cf. **Monday**).
- NEXPTIME-complete.

Fortunately, adding more agents does not lead to a more complex logic:

Theorem ([Balbiani, Herzig, Troquard 2007])

(Full) Ldm is NEXPTIME-complete.

Link with product logic and complexity

- Hence the logic of the two-agent *Ldm* is nothing but the product $S5^2 = S5 \otimes S5$ [Marx 1999], [Gabbay et al. 2003] (cf. **Monday**).
- NEXPTIME-complete.

Fortunately, adding more agents does not lead to a more complex logic:

Theorem ([Balbiani, Herzig, Troquard 2007])

(Full) Ldm is NEXPTIME-complete.

Today outline

Group STIT (language)

Syntax is as follows:

$$\varphi ::= p \mid \neg\varphi \mid \varphi \vee \varphi \mid [J]\varphi$$

- $[J]\varphi$ reads "group J chooses/forces φ"
- $\langle J\rangle\varphi$ reads "group J allows φ"

- $[\emptyset]\varphi \approx \Box\varphi$: "nature forces φ" = outcome settledness
- $\langle\emptyset\rangle\varphi \approx \Diamond\varphi$: "nature allows φ" = outcome possibility

Group STIT (Semantics)

A $\mathcal{G}$STIT-model is a tuple $\mathcal{M} = (W, R, V)$ where:

- W is a set of contexts;
- R is a collection of equivalence relations R_J (one for every coalition $J \subseteq \mathcal{A}gt$):

  $\Longrightarrow$ choice alternatives
  - $R_{J_1 \cup J_2} \subseteq R_{J_1}$ (more members $\Rightarrow$ tighter choices)
  - $R_\emptyset \subseteq R_J \circ R_{\overline{J}}$ (nature allows $\Rightarrow$ independent groups allow to allow)
- $V : \mathcal{A}tm \longrightarrow 2^W$ is a valuation function.

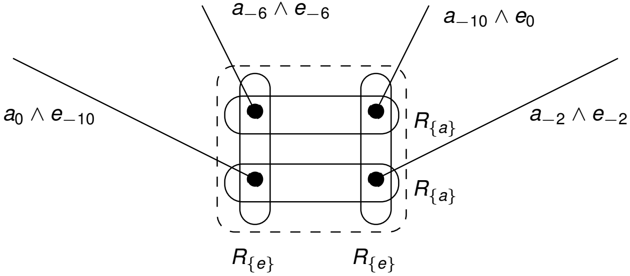
Truth conditions:

- $\mathcal{M}, w \models p$ iff $w \in V(p)$
- $\mathcal{M}, w \models [J]\varphi$ iff for all $u \in R_J(w)$, $\mathcal{M}, u \models \varphi$

$\mathcal{G}$STIT-models and strategic games

	defect_e	silent_e
defect_a	$(-6, -6)$	$(-10, 0)$
silent_a	$(0, -10)$	$(-2, -2)$

$\approx$



Group STIT (Axiomatics)

S5($[J]$)	S5 axioms for every $[J]$
(Mon)	$[J_1]\varphi \rightarrow [J_1 \cup J_2]\varphi$
Elim($[\emptyset]$)	$\langle\emptyset\rangle\varphi \rightarrow \langle J\rangle\langle\overline{J}\rangle\varphi$

- Elim($[\emptyset]$) is $[J][\overline{J}]\varphi \rightarrow \Box\varphi$
  $\Longrightarrow$ If a coalition can make the complementary (and thus independent) coalition do φ then φ is settled. (Remember $\Box \approx [\emptyset]$.)

Theorem (corollary of [Sahlqvist 1975])

$\mathcal{G}$STIT is sound and complete w.r.t. the class of $\mathcal{G}$STIT-models.

Is there a link with Coalition Logic?

Coalition Logic language:

$$\varphi ::= p \mid \neg\varphi \mid \varphi \vee \varphi \mid \langle\!\lbrack J \rbrack\!\rangle \mathbf{X}\varphi$$

- $\langle\!\lbrack J \rbrack\!\rangle \mathbf{X}\varphi$ reads "J are able to enforce φ *whatever other agents do*"
- $\exists$ a strategy of J, s.t. $\forall$ outcome φ

From Coalition Logic to $\mathcal{G}$STIT (tentative)

We define tr_0 a function from CL formulas to $\mathcal{G}$STIT formulas s.t.:

$$\begin{array}{lcl} tr_0(p) & = & p \\ tr_0(\langle\!\![J]\!\!\rangle \mathbf{X}\varphi) & = & \langle \emptyset \rangle [J] tr_0(\varphi) \end{array}$$

and homomorphic for classical connectives.

Problem.

- $\langle\!\![\emptyset]\!\!\rangle \mathbf{X}\varphi \wedge \langle\!\![\emptyset]\!\!\rangle \mathbf{X} \langle\!\![\emptyset]\!\!\rangle \mathbf{X} \neg\varphi$ is CL-consistent
- by tr_0 it translates to: $\langle \emptyset \rangle [\emptyset]\varphi \wedge \langle \emptyset \rangle [\emptyset] \langle \emptyset \rangle [\emptyset] \neg\varphi$
- and collapses to the inconsistent formula $[\emptyset]\varphi \wedge [\emptyset]\neg\varphi$
- We need a way to jump from moment to moment

From Coalition Logic to $\mathcal{G}$STIT (tentative)

We define tr_0 a function from CL formulas to $\mathcal{G}$STIT formulas s.t.:

$$\begin{array}{lcl} tr_0(p) & = & p \\ tr_0(\langle\!\![J]\!\!\rangle \mathbf{X}\varphi) & = & \langle\emptyset\rangle[J]tr_0(\varphi) \end{array}$$

and homomorphic for classical connectives.

Problem.

- $\langle\!\![\emptyset]\!\!\rangle\mathbf{X}\varphi \wedge \langle\!\![\emptyset]\!\!\rangle\mathbf{X}\langle\!\![\emptyset]\!\!\rangle\mathbf{X}\neg\varphi$ is CL-consistent
- by tr_0 it translates to: $\langle\emptyset\rangle[\emptyset]\varphi \wedge \langle\emptyset\rangle[\emptyset]\langle\emptyset\rangle[\emptyset]\neg\varphi$
- and collapses to the inconsistent formula $[\emptyset]\varphi \wedge [\emptyset]\neg\varphi$
- We need a way to jump from moment to moment

Normal Simulation of Coalition Logic (semantics)

Language.

$$\varphi ::= p \mid \neg\varphi \mid \varphi \vee \varphi \mid \mathbf{X}\varphi \mid [J]\varphi$$

An NCL-model is a tuple $\mathcal{M} = (W, R, F_X, V)$ where:

- (W, R, V) is a $\mathcal{G}$STIT model, further constrained by $R_{\mathcal{A}gt} = Id$;
- $F_X : W \longrightarrow W$ is a total function;

Extra operator:

- $\mathcal{M}, w \models \mathbf{X}\varphi$ iff $\mathcal{M}, F_X(w) \models \varphi$

Normal Simulation of Coalition Logic (axiomatics)

S5($[J]$)	enough S5-schemas, for every $[J]$
(Mon)	$[J_1]\varphi \to [J_1 \cup J_2]\varphi$
Elim($[\emptyset]$)	$\langle\emptyset\rangle\varphi \to \langle J\rangle\langle\overline{J}\rangle\varphi$
Triv($[\mathcal{A}gt]$)	$\varphi \to [\mathcal{A}gt]\varphi$
K($\mathbf{X}$)	$\mathbf{X}(\varphi \to \psi) \to (\mathbf{X}\varphi \to \mathbf{X}\psi)$
D($\mathbf{X}$)	$\mathbf{X}\varphi \to \neg\mathbf{X}\neg\varphi$
Det($\mathbf{X}$)	$\neg\mathbf{X}\neg\varphi \to \mathbf{X}\varphi$

Theorem ([Broersen, Herzig, Troquard 2007], corollary of [Sahlqvist 1975])

NCL is sound and complete w.r.t. the class of NCL-models.

From Coalition Logic to NCL

We now define *tr* the translation from CL formulas to NCL formulas s.t.:

$$\begin{array}{rcl} tr(p) & = & p \\ tr(\langle\!\![J]\!\!\rangle \mathbf{X}\varphi) & = & \langle\emptyset\rangle[J]\mathbf{X}tr(\varphi) \end{array}$$

and homomorphic for classical connectives.

Is it better than with $\mathcal{G}$STIT?

- $tr(\langle\!\![\emptyset]\!\!\rangle \mathbf{X}\varphi \wedge \langle\!\![\emptyset]\!\!\rangle \mathbf{X}\langle\!\![\emptyset]\!\!\rangle \mathbf{X}\neg\varphi) = \langle\emptyset\rangle[\emptyset]\mathbf{X}\varphi \wedge \langle\emptyset\rangle[\emptyset]\mathbf{X}\langle\emptyset\rangle[\emptyset]\mathbf{X}\neg\varphi$
- just entails the consistent formula $[\emptyset]\mathbf{X}\varphi \wedge [\emptyset]\mathbf{X}[\emptyset]\mathbf{X}\neg\varphi$

Theorem (correct embedding)

φ is a theorem of CL iff $tr(\varphi)$ is a theorem of NCL.

What do we have?

- A logic for reasoning about coalitions
- Stems from logic in philosophy of action
- More expressive than Coalition Logic
- But NEXPTIME-complete...
  [Schwarzentruber et al. forthcoming]

- So for what would NCL be particularly suitable, hey?

What do we have?

- A logic for reasoning about coalitions
- Stems from logic in philosophy of action
- More expressive than Coalition Logic
- But NEXPTIME-complete...
  [Schwarzentruber et al. forthcoming]

- So for what would NCL be particularly suitable, hey?

Today outline

CL models vs. NCL-models

Coalition Logic

- Neighborhood models
- Game models
- Idea: associate a strategic game to every state

In NCL, *contexts*

- are ‘part’ of the strategic game,
- and model
  - physical description of the world
  - current choice/commitment of agents

Helpful modeling power! Look!

CL models vs. NCL-models

Coalition Logic

- Neighborhood models
- Game models
- Idea: associate a strategic game to every state

In NCL, *contexts*

- are ‘part’ of the strategic game,
- and model
  - physical description of the world
  - current choice/commitment of agents

Helpful modeling power! Look!

CL models vs. NCL-models

Coalition Logic

- Neighborhood models
- Game models
- Idea: associate a strategic game to every state

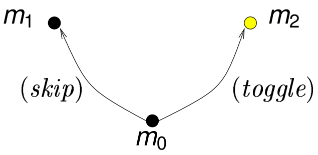
In NCL, *contexts*

- are ‘part’ of the strategic game,
- and model
    - physical description of the world
    - current choice/commitment of agents

Helpful modeling power! Look!

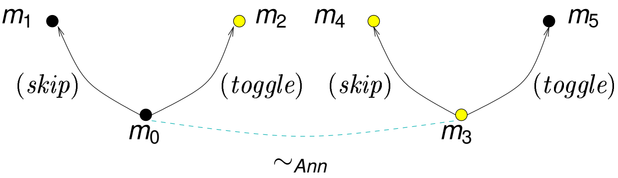
Ann toggles

- At m_0, the light is off: $m_0 \models \neg li$
- Ann can *toggle* or *skip*
- $m_0 \models \langle\!\![Ann]\!\!\rangle \mathbf{X} li$
  at m_0, "Ann is able to achieve *li*"



Poor blind Ann – a CL account

- As previously, the light is off: $m_0 \models \neg li$
- Ann is blind and cannot distinguish a world where the light is on from a world where the light is off
- $m_0 \models K_{Ann}\langle\!\langle Ann \rangle\!\rangle \mathbf{X} li$
  at m_0, "Ann knows she is able to achieve li"



Adding knowledge

A logical language of action and knowledge must be able to distinguish the following scenarii:

1. the agent *a* knows it has a particular action/choice in its repertoire that ensures φ, possibly without knowing which choice to make to ensure φ.
2. the agent *a* 'knows how to' / 'can' / 'has the power to' ensure φ.

Two readings of "having a strategy"

- $tr(K_J \langle\!\langle J \rangle\!\rangle \mathbf{X}\varphi) = K_J\langle\emptyset\rangle[J]\mathbf{X}\varphi$ **(de dicto)**
  Group J knows (K) there is ($\exists$) a choice s.t. for all ($\forall$) possible outcomes φ
  - Alternating-time *Epistemic* Temporal Logic ATEL [Wooldridge, van der Hoek 2002]
- We want: $\langle\emptyset\rangle K_J[J]\mathbf{X}\varphi$ **(de re)**
  There is a choice ($\exists$), s.t. group J knows (K) that for all ($\forall$) possible outcomes φ
  - ATEL does not deal with *de re* strategies [Jamroga 2003], [Schobbens 2004]
  - Several corrections [Schobbens 2004], [Jamroga, van der Hoek 2004], [Jamroga, Ågotnes 2006, 2007]
  - First semantics with STIT [Herzig, Troquard 2006]

Two readings of "having a strategy"

- $tr(K_J \langle\!\langle J \rangle\!\rangle \mathbf{X}\varphi) = K_J \langle\emptyset\rangle [J]\mathbf{X}\varphi$ **(de dicto)**
  Group J knows (K) there is ($\exists$) a choice s.t. for all ($\forall$) possible outcomes φ
  - Alternating-time *Epistemic* Temporal Logic ATEL [Wooldridge, van der Hoek 2002]
- We want: $\langle\emptyset\rangle K_J [J]\mathbf{X}\varphi$ **(de re)**
  There is a choice ($\exists$), s.t. group J knows (K) that for all ($\forall$) possible outcomes φ
  - ATEL does not deal with *de re* strategies [Jamroga 2003], [Schobbens 2004]
  - Several corrections [Schobbens 2004], [Jamroga, van der Hoek 2004], [Jamroga, Ågotnes 2006, 2007]
  - First semantics with STIT [Herzig, Troquard 2006]

Epistemic NCL

Language.

$$\varphi ::= p \mid \neg\varphi \mid \varphi \vee \varphi \mid \mathbf{X}\varphi \mid [J]\varphi \mid K_i\varphi$$

ENCL-models are tuples $\mathcal{M} = (W, R, F_X, \sim, V)$ where:

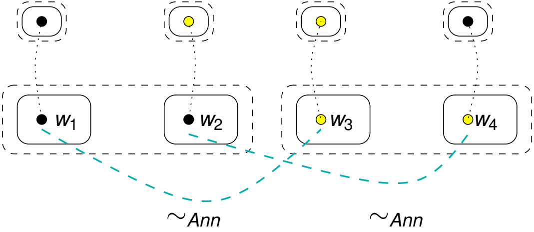
- (W, R, F_X, V) is an NCL-model.
- $\sim$ is a collection of equivalence relations $\sim_i$ (one for every agent $i \in \mathcal{A}gt$).

Extra operators:

- $\mathcal{M}, w \models K_i\varphi$ iff for all $u \sim_i w$, $\mathcal{M}, u \models \varphi$

Every K_i is axiomatized as a standard epistemic modality. [Hintikka 1962]

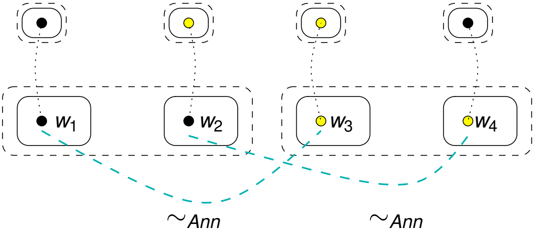
Poor blind Ann again



Epistemic relations are over contexts instead of ‘pure’ states.

- $w_i \models K_{Ann}\langle\emptyset\rangle[Ann]\mathbf{X}\varphi$
  Ann knows she has an action that leads to a lighten moment.
- $w_i \not\models \langle\emptyset\rangle K_{Ann}[Ann]\mathbf{X}\varphi$
  Ann does **not** know how to achieve it.

Poor blind Ann again



Epistemic relations are over contexts instead of ‘pure’ states.

- $w_i \models K_{Ann}\langle\emptyset\rangle[Ann]\mathbf{X}\varphi$
  Ann knows she has an action that leads to a lighten moment.
- $w_i \not\models \langle\emptyset\rangle K_{Ann}[Ann]\mathbf{X}\varphi$
  Ann does **not** know how to achieve it.

Outlook for the remaining courses

Thursday (tomorrow): What agents want

- logic of intention: Cohen and Levesque

Friday: What agents can plan

- ATL, strategic STIT