Tutorial "Action Sentences"

Jan Broersen and Andreas Herzig
broersen@uu.nl, herzig@irit.fr

Semantics & Philosophy in Europe (SPE4)
Bochum, September 27, 2011

Overview

1. An introduction to STIT theory and related logics
   Andreas Herzig

2. Extensions: strategic STIT; STIT with deontic and epistemic modalities
   Jan Broersen

1. An introduction to STIT theory and related logics

Andreas Herzig
(joint work with Emiliano Lorini and Nicolas Troquard)

slides at http://www.irit.fr/~Andreas.Herzig

Introduction and motivation

Philosophy of action

- *Ontological view*: Davidson, Thomson, Anscombe,. . .
- *Modal view*: Kanger, Chellas, von Kutschera, Belnap et col., 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 Floyd-Hoare logic of program verification [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]...

- link? $\Longrightarrow$ my talk on Thursday

Logics of agency: Pörn's account

- $D_i\varphi$ true at world w if φ is true at every hypothetical situation where agent i *"does at least as much as he does in w"*
- $D'_i\varphi$ true at w if $\neg\varphi$ is true in every hypothetical situation w' such that *"the opposite of everything i does in w is the case in w'"*
- Combination of two normal operators in a non-normal modality:
  - $D_i\varphi$ = "it is necessary for something i does that φ"
  - $D'_i\varphi$ = "but for i's action, it would not be the case that φ"
  - $E_i\varphi \overset{\text{def}}{=} D_i\varphi \wedge \neg D'_i\varphi$ = "agent i brings it about that φ"

A controversial framework

- Used a lot for modeling institutionalized power and law [Jones and Sergot 1996], [Royakkers 2000], [Carmo and 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 everthing he does in [a world], are of dubious intelligibility without substantial further elucidation, and Pörn offers none.” [Horgan 1979]
- “the most detailed working out of the modal logic of agency as based on binary relational semantics”; “the semantics proper is problematic” [Belnap 1991]
- “the intuitive significance of this semantics is not altogether clear” [Segerberg 1992]

Logics of agency: Belnap and Perloff's STIT theory

How can we distinguish between sentences which involve agency and those which do not?

- Moby Dick example [Belnap and Perloff 1988]:
  - Is "Ishmael sails on board the Pequod" agentive for Ishmael?

Criterion:

- 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".

$\Longrightarrow$ 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 theory

alias "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 $\texttt{stit}$: *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]
  - ...

From individual to collective agency [Horty 2001]

- $[\{\text{John}, \text{Mary}\}]\mathit{cubeOnTable}$
  - monotony:

    $$[\text{John}]\mathit{cubeOnTable} \to [\text{John}, \text{Mary}]\mathit{cubeOnTable}$$

  - independence:

    $$[\text{John}]\mathit{cubeOnTable} \wedge [\text{Mary}]\mathit{ballOnTable} \to [\text{John}, \text{Mary}](\mathit{cubeOnTable} \wedge \mathit{ballOnTable})$$

  - ...

‘Do’ and ‘can’: agency, ability, and opportunity [von Wright, Kenny]

- agency: John puts the cube on the table
- ability: Obelix can put the cube on the table (although there is no cube around right now; repeatable under normal conditions)
- opportunity: Asterix can put the cube on the table (right now, and because he has just drunk magic potion)
  - reducible to possibility of agency

The rest of Part I in a nutshell

Formal machinery behind this linguistic agenda?

- coalition logic
  - opportunity
- 'Chellas STIT' logic
  - opportunity + agency
  - embeds coalition logic (version with discrete time)
  - has much more logical principles than Pörn's logic: normal modal operators; agents' independence

Logical investigation:

- models?
- axioms?
- mathematical properties?
- applications?

Outline

Introduction and motivation

Coalition Logic

STIT: branching time and agents' choices

STIT: reasoning about individual choice

STIT: reasoning about coalitional choice

Epistemic extension of STIT and the concept of uniform strategy

Coalition Logic

Coalition logic [Pauly 2001, 2002]

- aim: model procedures involving the interactions between multiple agents ('social software')
  - voting procedures
  - multiagent systems
  - ...

- allows to express what a coalition of agents can ensure by doing a joint action (in a given state)

Coalition Logic: main operator

$$\langle\!\![J]\!\!\rangle\varphi \quad = \quad \text{“agents in } J \text{ can act together to ensure that } \varphi \text{ is true in the next state (whatever the other agents do)”}$$

$$= \quad \text{“}\exists \text{ a joint action of } J \text{ such that } \forall \text{ joint actions of the agents outside } J, \ \varphi \text{ holds in the next state”}$$

- case $J = \emptyset$: “φ is necessarily true at the next time point”

Coalition logic (CL): language

- $\texttt{AGT} = \{i, j, \dots\}$: set of agents (finite)
  - sets of agents ('coalitions') $J, J', \dots \subseteq \texttt{AGT}$
- $\texttt{PRP} = \{p, q, \dots\}$: set of propositional variables (countable)
- grammar:

$$\varphi ::= p \mid \varphi \land \varphi \mid \neg\varphi \mid \langle\!\![J]\!\!\rangle\varphi$$

$\Rightarrow$ write $\langle\!\![i]\!\!\rangle\varphi$ instead of $\langle\!\![\{i\}]\!\!\rangle\varphi$,
write $\langle\!\![i, j]\!\!\rangle\varphi$ instead of $\langle\!\![\{i, j\}]\!\!\rangle\varphi$

Example: coordinated attack

- Two agents 1 and 2 are trying to move an attack against a common enemy. The enemy will be defeated iff 1 and 2 move a coordinated attack (both 1 and 2 attack the enemy).
  - 1 has two actions available: *attack* and *wait*
  - 2 has two actions available: *attack* and *wait*

In formulas:
$\langle\!\langle 1,2 \rangle\!\rangle \mathit{defeatEnemy} \wedge \neg \langle\!\langle 1 \rangle\!\rangle \mathit{defeatEnemy} \wedge \neg \langle\!\langle 2 \rangle\!\rangle \mathit{defeatEnemy}$

mechanism design: Bonn vs. Berlin [Pauly 2001]

After the "German Question" had been solved, on June 20th, 1991, the German parliament was faced with the Berlin question: Should the German parliament and the seat of government move to Berlin or stay in Bonn? In this historic debate, the parliament was very divided and this division ran through all parties. About 100 speeches were made while another 100 speeches were placed on record. [...]

In its debate, the German parliament considered 5 different motions, the 3 central ones being:

p_1 = *parliament and government move to Berlin*

p_2 = *parliament moves to Berlin but government remains in Bonn*

p_3 = *both parliament and government remain in Bonn*

[...] Since there were more than 2 motions up for vote, the

parliamentary council of elders first had to decide on a voting procedure.

Bonn vs. Berlin (2) [Pauly 2001]

- initially: $\neg p_1 \wedge \neg p_2 \wedge \neg p_3$
- aim: design a procedure warranting a democratic decision
  - effective: no eternal voting
  - fair: moves are brought about by majorities
  - ...
- basic ideas:
  - use majority voting
  - eliminate step by step some alternative(s) until only a single remains

Bonn vs. Berlin (3) [Pauly 2001]

- a list of requirements
  1. alternatives which have been eliminated remain eliminated:
     $$\neg p \to \langle\![\emptyset]\!\rangle \neg p \text{ for all atomic } p$$
  2. at every stage at least one alternative has to be eliminated:
     $$\bigwedge_{\delta \in Sit}(\delta \to \langle\![\emptyset]\!\rangle \neg\delta)$$
     - where
       $Sit = \{l_1 \wedge l_2 \wedge l_3 : l_i = p_i \text{ or } l_i = \neg p_i\} \setminus \{\neg p_1 \wedge \neg p_2 \wedge \neg p_3\}$
  3. each vote must be a democratic majority vote:
     $$\langle\![\mathtt{AGT}]\!\rangle\varphi \leftrightarrow \langle\![J]\!\rangle\varphi \text{ for every } J \text{ with } |J| > \tfrac{1}{2} \times |\mathtt{AGT}|$$
     - supposing $|\mathtt{AGT}|$ is odd
  4. each vote between two alternatives only:
     $$(\langle\![\mathtt{AGT}]\!\rangle(\varphi \wedge \psi) \wedge \langle\![\mathtt{AGT}]\!\rangle(\neg\varphi \wedge \psi)) \to \langle\![\emptyset]\!\rangle\psi$$
     - if two different outcomes then these are the only ones

Bonn vs. Berlin (4) [Pauly 2001]

- are the requirements consistent?
- is there a model M where
  - requirements (1)-(4) are valid, and
  - there is a state where $p_1 \wedge p_2 \wedge p_3$ holds?
- solution 1 (adopted by council of elders)
  1. majority vote whether p_2 or $\neg p_2$
  2. if $\neg p_2$ then majority vote whether p_1 or p_3

  $\Rightarrow$ not all 3 alternatives considered equally
  - p_2 stands alone against p_1 and p_3
- solution 2
  1. ...
  2. ...
  3. ...
  4. ...

  satisfies requirement "eliminate only one option per step"

$$(p \wedge q) \rightarrow \langle\!\langle \emptyset \rangle\!\rangle (p \vee q) \quad \text{for atomic } p \neq q$$

Effectivity functions

- S: a set of states
- $e : 2^{\mathtt{AGT}} \longrightarrow 2^{2^S}$: effectivity function
- When $Q \in e(J)$ then J is *effective* for Q: J can force the world to be among the states Q at the next step
  - When $Q \in e(\emptyset)$ then at the next step the world is necessarily among the states in Q

Example: coordinated attack (cont.)

- $S = \{s_{Defeat}, s_{Undefeat}\}$
- $e(\emptyset) = e(\{1\}) = e(\{2\}) = \{S\}$
- $e(\{1,2\}) = \{\{s_{Defeat}, s_{Undefeat}\}, \{s_{Defeat}\}, \{s_{Undefeat}\}\}$

Playable effectivity function

e is **playable** iff:

1. $\emptyset \notin e(J)$
2. $S \in e(J)$
3. if $Q_1 \in e(J)$ then $Q_1 \cup Q_2 \in e(J)$
   (*outcome monotonicity*)
4. if $S \setminus Q \notin e(\emptyset)$ then $Q \in e(\mathtt{AGT})$, for all $Q \subseteq S$
   ($\mathtt{AGT}$-*maximality*)
5. for $J \cap I = \emptyset$,
   if $Q_1 \in e(J)$ and $Q_2 \in e(I)$ then $Q_1 \cap Q_2 \in e(J \cup I)$
   (*superadditivity*)

CL models

A CL-model is a tuple $M = ((S, E), V)$ where:

- S is a set of states;
- $E : S \longrightarrow (2^{\texttt{AGT}} \longrightarrow 2^{2^S})$ such that every $E(s)$ is a playable effectivity function;
- $V : \texttt{PRP} \longrightarrow 2^S$ is a valuation function.

Truth conditions

$$\begin{array}{lll} M, s \models p & \text{iff} & p \in V(s) \\ M, s \models \neg\varphi & \text{iff} & M, s \not\models \varphi \\ M, s \models \varphi \wedge \psi & \text{iff} & M, s \models \varphi \text{ and } M, s \models \psi \\ M, s \models \langle\!\langle J \rangle\!\rangle\varphi & \text{iff} & \{s' \ : \ M, s' \models \varphi\} \in E_s(J) \end{array}$$

validity and satisfiability defined as usual:

- φ is valid iff $M, s \models \varphi$ for every model M and every s in M

A complete axiomatization of CL

(RE)	If $\varphi \leftrightarrow \psi$ then $\langle\![J]\!\rangle\varphi \leftrightarrow \langle\![J]\!\rangle\psi$
(M)	$\langle\![J]\!\rangle(\varphi \wedge \psi) \rightarrow (\langle\![J]\!\rangle\varphi \wedge \langle\![J]\!\rangle\psi)$
($\bot$)	$\neg\langle\![J]\!\rangle\bot$
($\top$)	$\langle\![J]\!\rangle\top$
($\mathtt{AGT}$)	$\neg\langle\![\emptyset]\!\rangle\varphi \rightarrow \langle\![\mathtt{AGT}]\!\rangle\neg\varphi$
(S)	$(\langle\![J]\!\rangle\varphi \wedge \langle\![I]\!\rangle\psi) \rightarrow \langle\![J \cup I]\!\rangle(\varphi \wedge \psi)$ if $J \cap I = \emptyset$

Properties of CL

- $\langle\!\![\emptyset]\!\!\rangle$ is a normal 'necessary' operator:
  - $\vdash \langle\!\![\emptyset]\!\!\rangle(\varphi \wedge \psi) \leftrightarrow (\langle\!\![\emptyset]\!\!\rangle\varphi \wedge \langle\!\![\emptyset]\!\!\rangle\psi)$
- $\langle\!\![\mathtt{AGT}]\!\!\rangle$ is a normal 'possible' operator:
  - $\vdash \langle\!\![\mathtt{AGT}]\!\!\rangle(\varphi \vee \psi) \leftrightarrow (\langle\!\![\mathtt{AGT}]\!\!\rangle\varphi \vee \langle\!\![\mathtt{AGT}]\!\!\rangle\psi)$
- $\langle\!\![\emptyset]\!\!\rangle$ and $\langle\!\![\mathtt{AGT}]\!\!\rangle$ are dual modal operators:
  - $\vdash \langle\!\![\emptyset]\!\!\rangle\varphi \leftrightarrow \neg\langle\!\![\mathtt{AGT}]\!\!\rangle\neg\varphi$
- Two disjoint coalitions cannot bring about conflicting effects:
  - if $I \cap J = \emptyset$ then $\vdash (\langle\!\![I]\!\!\rangle\varphi \wedge \langle\!\![J]\!\!\rangle\neg\varphi) \rightarrow \bot$
- Complexity of the satisfiability problem and of the validity problem: PSPACE complete

Alternating-time Temporal Logic ATL
[Alur, Henzinger and Kupferman, 2002]

- temporal extension of Coalition Logic [Goranko 2003]
- theoretical computer science: reasoning about distributed systems
- multi-agent systems: reasoning about dynamic systems with multiple agents

ATL: main operators

$$\langle\langle J\rangle\rangle \mathtt{X}\varphi \quad = \quad \text{“}J \text{ can act together to ensure } \varphi \text{ in the next state”}$$
$$= \quad \text{“}\exists \text{ joint action of } J \text{ such that } \forall \text{ outcome, } \varphi \text{ holds in the next state”}$$

$$\langle\langle J\rangle\rangle \mathtt{G}\, \varphi \quad = \quad \text{“}J \text{ can act to ensure that } \varphi \text{ is true henceforth”}$$

$$\langle\langle J\rangle\rangle (\varphi \mathtt{U} \psi) \quad = \quad \text{“}J \text{ can act together to ensure that } \varphi \text{ remains true until } \psi \text{ is true”}$$

- Complexity of the satisfiability problem: EXPTIME complete

Relation between ATL and CL [Goranko 2003]

$$\langle\!\![J]\!\!\rangle \varphi \approx \langle\langle J \rangle\rangle \mathtt{X} \varphi$$

STIT: branching time and agents' choices

“I shall finish this lecture”

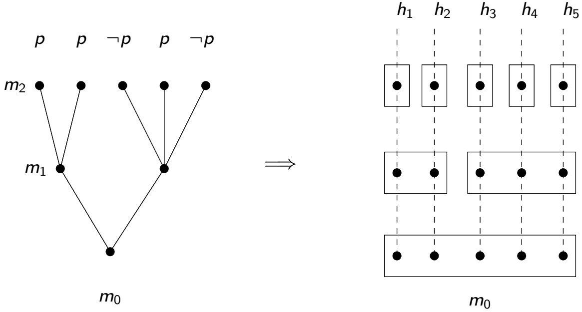
- branching time (BT) structure ($Mom, <$):

  $$\begin{array}{rcl} Mom & = & \text{set of moments} \\ < & = & \text{temporal precedence} \end{array}$$

- controversial evaluation of future-tense sentences in branching time
  - “truth value gap” [Thomason 1984]
- “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/actualist temporal logic [Prior 1967]

Ockhamist branching time temporal logic (BT)

BT structure $(Mom, <)$:



- 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)
  - $m_0/h_3 \not\models \mathtt{F}\ p$
  - $m_0/h_1 \models \mathtt{F}\ p$

Agents' choices (AC)

idea [Kracht et al.]:

- usually an agent is not able to select one possible future to become the unique actual future
- but by his action he can make sure that certain futures, which before his action are possible, are no longer possible after his action

$Choice_i^m$ = repertoire of choices for agent i at moment m

$BT{+}AC$ models

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

- $\langle Mom, < \rangle$ is a branching time structure;
- $Choice : \texttt{AGT} \times Mom \to 2^{2^{Hist}}$ is a function mapping each agent and each moment m into a partition of H_m;
- v is valuation function $v : \texttt{PRP} \to 2^{Mom \times Hist}$.

Choice: independence

Definition

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

$Select_m$ = the set of all selection functions s_m, for a given m

Assumption (independence of agents)

For every $s_m \in Select_m$, $\bigcap_{i \in \mathtt{AGT}} s_m(i) \neq \emptyset$

- more assumptions:
  - liveness
  - no choice between undivided histories
  - ...

Choice: groups

Generalize choice function from individuals to coalitions:

$$Choice : 2^{\texttt{AGT}} \times Mom \rightarrow 2^{2^{Hist}}$$

Definition (choice function for a group of agents)

For $J \subseteq \texttt{AGT}$,

$$Choice_J^m = \{\bigcap_{i \in J} s_m(i) \ : \ s_m \in Select_m\}$$

N.B.: $Choice_{\emptyset}^m = \{H_m\}$

Truth conditions

A formula is evaluated at a moment/history pair:

$$
\begin{array}{lll}
\mathcal{M}, m/h \models p & \text{iff} & m/h \in v(p), p \in \texttt{PRP} \\
\mathcal{M}, m/h \models \neg\varphi & \text{iff} & \mathcal{M}, m/h \not\models \varphi \\
\mathcal{M}, m/h \models \varphi \wedge \psi & \text{iff} & \mathcal{M}, m/h \models \varphi \text{ and } \mathcal{M}, m/h \models \psi \\
 & & \\
\mathcal{M}, m/h \models \texttt{F}\,\varphi & \text{iff} & \exists m',\ m < m',\ \mathcal{M}, m'/h \models \varphi \\
 & & \\
\mathcal{M}, m/h \models \Box\varphi & \text{iff} & \text{for all } h' \in H_m,\ \mathcal{M}, m/h' \models \varphi \\
\mathcal{M}, m/h \models \Diamond\varphi & \text{iff} & \text{for some } h' \in H_m,\ \mathcal{M}, m/h' \models \varphi
\end{array}
$$

- $\Box\varphi$ = "whatever happens, φ is true at the current moment" (historical necessity)
- $\Diamond\varphi = \neg\Box\neg\varphi$ = "φ is historically possible"
  - $\Diamond[\{Jack, Mary\}]\texttt{F}\ cubeOnTable$ = "Jack and Mary can ensure that the cube is eventually on the table"

Truth conditions: the Chellas stit

for $h \in H_m$:

$$
\begin{aligned}
Choice_i^m(h) &= \{h' \ : \ \text{there is } Q \in Choice_i^m \text{ s.th. } h, h' \in Choice_i^m\} \\
&= \text{the particular choice of } i \text{ at context } m/h
\end{aligned}
$$

$$\mathcal{M}, m/h \models [J]\varphi \quad \text{iff} \quad \mathcal{M}, m/h' \models \varphi, \ \forall h' \in Choice_J^m(h)$$

$[J]\varphi$ = "the alternative that is presently and actually chosen by J guarantees that φ is true"
= "J sees to it that φ"

The link with Coalition Logic: 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 $m_1 < m_3 < m_2$.

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

Hypothesis (determinism)

$\forall m \in Mom, \ \exists m' \in W \ (m < m' \ and \ \forall h \in m', \ Choice^{m}_{\mathtt{AGT}}(h) = H_{m'})$

Translation of Coalition Logic to discrete-deterministic STIT

$$
\begin{array}{lcl}
tr(p) & = & \Box p, \text{ for } p \in \mathtt{PRP} \\
tr(\neg\varphi) & = & \neg tr(\varphi) \\
tr(\varphi \wedge \psi) & = & tr(\varphi) \wedge tr(\psi) \\
tr(\langle\![J]\!\rangle\varphi) & = & \Diamond [J] \mathtt{X}\, tr(\varphi)
\end{array}
$$

In STIT terminology

"coalition J is able to ensure φ"

can be paraphrased by

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

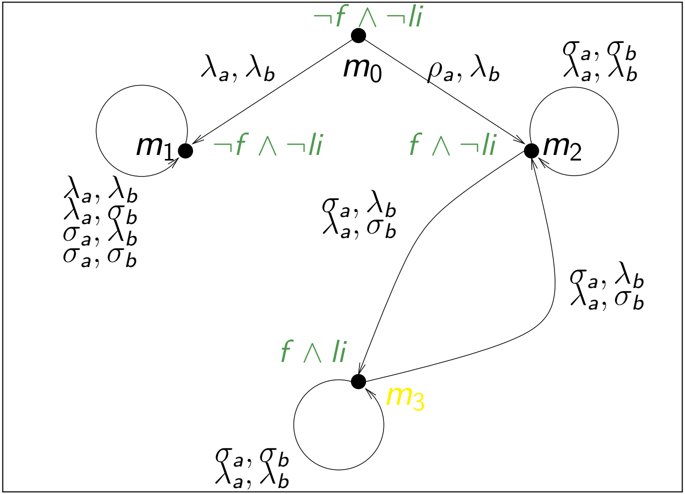
Theorem ([Broersen, Herzig, Troquard 2006])

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

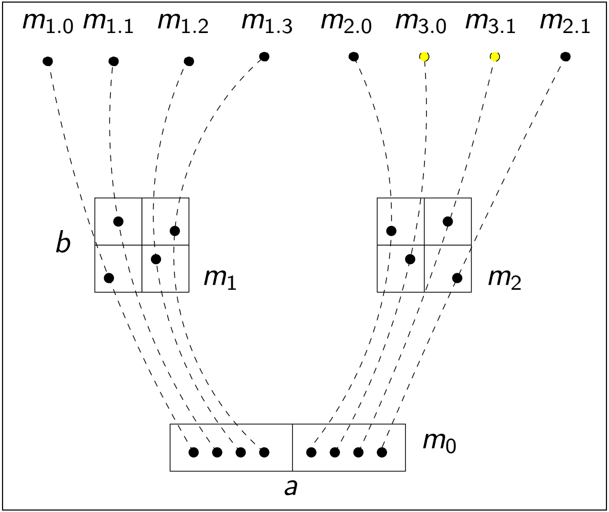
Example: Ann and Bill switch the light

- 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

Coalition logic model (game model)



Corresponding STIT model



So far so good

What is the logic of $BT{+}AC$ structures?

- axiomatics
- decidability
- complexity

the message is going to be:
reasoning about coalitional action is more difficult
than reasoning about individual action

Reasoning about individual choice

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

- $S5$ is characterized by equivalence frames (reflexive, transitive, and symmetrical)
- axiomatics: (K+T+4+B), or (K+T+4+5), or ...
  - 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 $\varphi \to \Box\Diamond\varphi$

Property

$\vdash \triangle_1\triangle_2 \ldots \triangle_k\varphi \leftrightarrow \triangle_k\varphi$, *for* $\triangle_i \in \{\Box, \Diamond\}$.

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

N.B.: no temporal operators

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)	$(\Diamond[0]\varphi_0 \wedge \ldots \wedge \Diamond[k]\varphi_k) \to \Diamond([0]\varphi_0 \wedge \ldots \wedge [k]\varphi_k)$

(AIA_k): family of Axiom schemas for the Independence of Agents

Theorem ([Xu 1994])

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

other completeness results:

- achievement STIT (single agent) [Xu 1995]
- STIT with restricted quantification over histories [Ciuni and Zanardo 2010]

A convenient truth

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

- valid, and hence provable due to the completeness theorem

A convenient truth

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

- valid, and hence provable due to the completeness theorem
  - problem: derive it from *Ldm* (we do not know the solution)
- holds the other way round, too: $\vdash \Box\varphi \leftrightarrow [1][0]\varphi$
  - $\Box\varphi \to \Box\Box\varphi$ (by S5($\Box$))
  - $\Box\Box\varphi \to [1][0]\varphi$ (by ($\Box\!\to\![i]$))
- allows to eliminate the $\Box$ operator
- consequence, for $l \neq m$ and $n \neq i$:

$$\vdash [l][m]\varphi \to [n][i]\varphi$$

  - to be generalised $\Rightarrow$ simpler axiomatisation

Alternative *Ldm*

- independence of agents in *Ldm*: (AIA_k)
  $(\Diamond[0]\varphi_0 \wedge \ldots \Diamond[k]\varphi_k) \to \Diamond([0]\varphi_0 \ldots [k]\varphi_k)$

- alternative axiomatization of *Ldm*
  [Balbiani, Herzig, Troquard 2007]:

$S5(i)$	axiom schemas of S5 for every $[i]$
$Def(\Box)$	$\Box\varphi \leftrightarrow [1][0]\varphi$
$(GPerm_k)$	$\langle l\rangle\langle m\rangle\varphi \to \langle n\rangle \bigwedge_{i\in \mathtt{AGT}\setminus\{n\}}\langle i\rangle\varphi$

Alternative semantics

All axiom schemes are in the Sahlqvist class $\Rightarrow$ standard possible worlds semantics!

Kripke models are of the form $M = (W, R, V)$, where

- W = nonempty set of possible worlds;
- R is a mapping associating to every $i \in \texttt{AGT}$ an equivalence relation R_i on W;
- $V : \texttt{PRP} \longrightarrow 2^W$.

R must satisfy the general permutation property ...

Alternative semantics (ctd.)

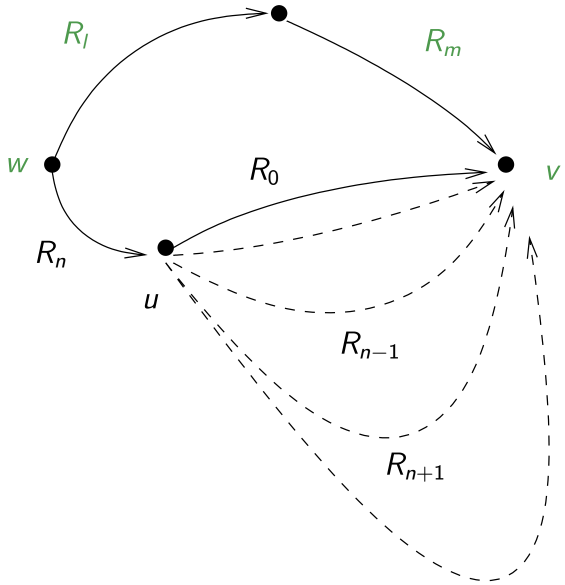
Definition (general permutation property)

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

standard modal truth condition:

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

Alternative semantics (illustration)



The link with product logics

If $\mathtt{AGT} = \{0, 1\}$ then the validities are axiomatized by:

- Def($\Box$): $\Box\varphi \leftrightarrow [1][0]\varphi$
- S5(0)
- S5(1)
- two instances of (GPerm$_1$):
  - $\langle 1\rangle\langle 0\rangle \to \langle 0\rangle\langle 1\rangle\varphi$
  - $\langle 0\rangle\langle 1\rangle \to \langle 1\rangle\langle 0\rangle\varphi$

provable:

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

The link with product logics: 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 9)

The link with product logics: complexity result

- product logic = many-dimensional logic: validates permutation axiom and Church-Rosser axioms
- logic of the two-agent Chellas STIT ≈ product $S5^2 = S5 \otimes S5$ [Marx 1999, Gabbay et al. 2003]
- complexity of $S5 \otimes S5$ satisfiability: NEXPTIME-complete

The link with product logics: complexity result

- product logic = many-dimensional logic: validates permutation axiom and Church-Rosser axioms
- logic of the two-agent Chellas STIT ≈ product $S5^2 = S5 \otimes S5$ [Marx 1999, Gabbay et al. 2003]
- complexity of $S5 \otimes S5$ satisfiability: NEXPTIME-complete

Adding more agents does not lead to a more complex logic:

Theorem ([Balbiani, Herzig, Troquard 2007])

Deciding satisfiability of a formula of individual STIT logic (without temporal operators) is NEXPTIME-complete.

Reasoning about coalitional choice

STIT logic with agents and coalitions

grammar (no temporal operators):

$$\varphi ::= p \mid \varphi \wedge \varphi \mid \neg\varphi \mid [J]\varphi \mid \Box\varphi$$

⇒ $[J]\varphi$ = "coalition J sees to it that φ no matter what the other agents in $\mathtt{AGT} \setminus J$ do".

⇒ $\Box\varphi$ = "φ is necessarily true" (historic necessity)

- $\Box\varphi$ same as $[\emptyset]\varphi$
- $\Diamond[J]\varphi$ = "J can see to it that φ whatever the other agents in $\mathtt{AGT} \setminus J$ do".

STIT models

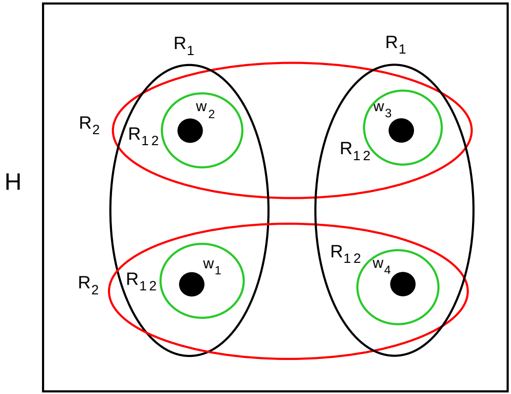
A STIT-model is a tuple $M = (W, \{R_J\}_{J \subseteq \texttt{AGT}}, H, V)$ where:

- W is a non-empty set of possible worlds or states;
- For all $J \subseteq \texttt{AGT}$, R_J is an equivalence relation over W:
  1. **Reflexive**: $(w, w) \in R_J$;
  2. **Transitive**: if $(w, v) \in R_J$ and $(v, u) \in R_J$ then $(w, u) \in R_J$;
  3. **Symmetric**: if $(w, v) \in R_J$ then $(v, w) \in R_J$;
- H is an equivalence relation over W;
- $V : \texttt{PRP} \longrightarrow 2^W$ is a valuation function.

$\Rightarrow R_J(w) = \{v \in W \ : \ (w, v) \in R_J\}$ is the set of outcomes that are forced by the action chosen by coalition J at w
$\Rightarrow H(w) = \{v \in W \ : \ (w, v) \in H\}$ is the set of all possible outcomes at w

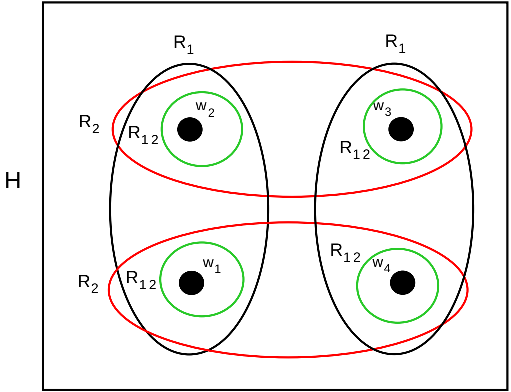
Constraints on STIT models

Action Outcomes: $R_{\{i\}} \subseteq H = R_{\emptyset}$



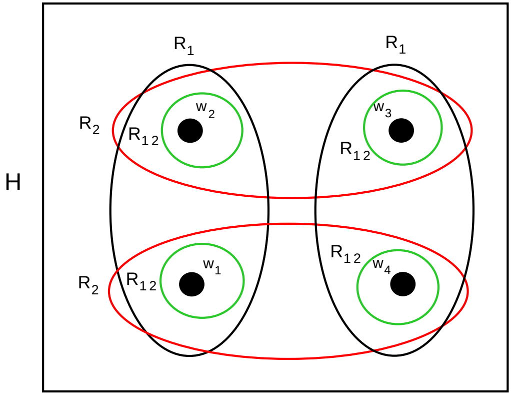
Constraints on STIT models (cont.)

Independence of Agents:
for all $w \in W$, for all $\langle w_i \rangle_{i \in \mathtt{AGT}} \in H(w)^n$, $\bigcap_{i \in \mathtt{AGT}} R_{\{i\}}(w_i) \neq \emptyset$



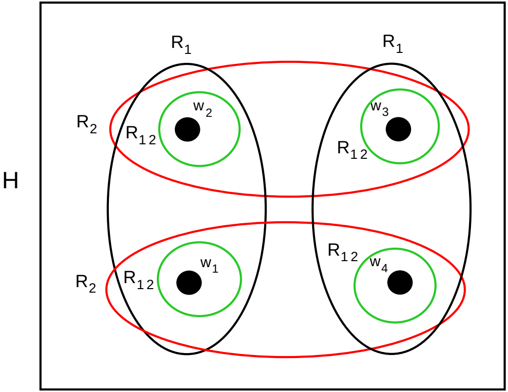
Constraints on STIT models (cont.)

Group Action: $R_J = \bigcap_{i \in J} R_{\{i\}}$



Constraints on STIT models (cont.)

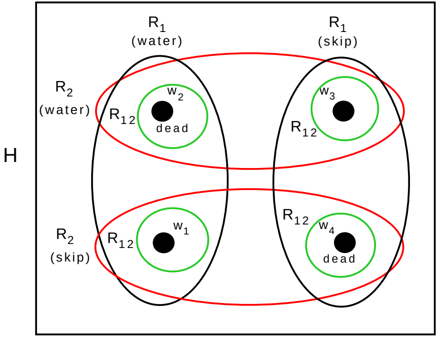
$\texttt{AGT}$-**Determinism**: $R_{\texttt{AGT}} = id_W$



Truth conditions

$$\begin{array}{lll}
M, w \models p & \text{iff} & w \in V(p) \\
M, w \models \neg\varphi & \text{iff} & M, w \not\models \varphi \\
M, w \models \varphi \land \psi & \text{iff} & M, w \models \varphi \text{ and } M, w \models \psi \\
M, w \models [J]\varphi & \text{iff} & M, v \models \varphi \text{ for all } v \text{ s.th. } (w, v) \in R_J \\
M, w \models \Box\varphi & \text{iff} & M, v \models \varphi \text{ for all } v \text{ s.th. } (w, v) \in H
\end{array}$$

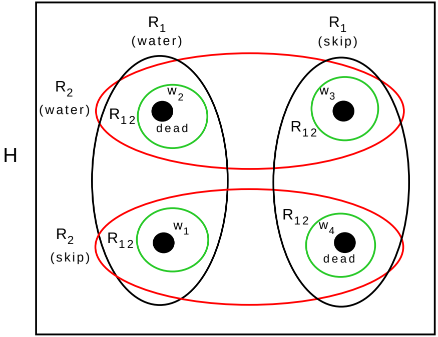
Example 1: two agents taking care of a plant



Two agents $\mathtt{AGT} = \{1, 2\}$ have to take care of a plant. If both of them water the plant (resp. do nothing) the plant will die.

- $[\{1, 2\}]$*alive* is true at w_1 and w_3.

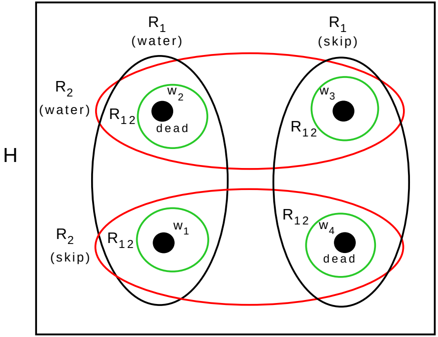
Example 1: two agents taking care of a plant



Two agents $\texttt{AGT} = \{1, 2\}$ have to take care of a plant. If both of them water the plant (resp. do nothing) the plant will die.

- $\Diamond[\{1, 2\}]\mathit{alive}$ is true at w_1, w_2, w_3, w_4.

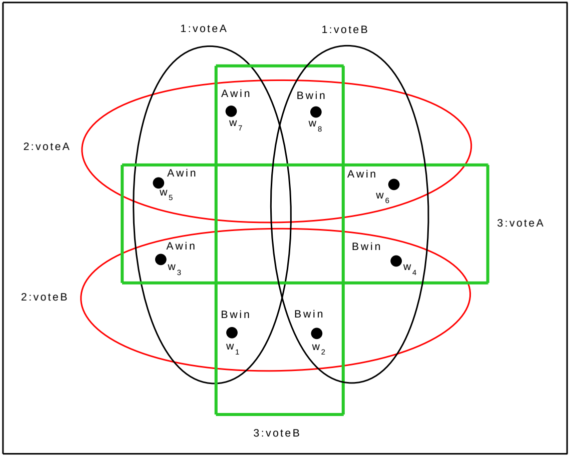
Example 1: two agents taking care of a plant



Two agents $\mathtt{AGT} = \{1, 2\}$ have to take care of a plant. If both of them water the plant (resp. do nothing) the plant will die.

- $\neg\Diamond[1]$*alive* and $\neg\Diamond[2]$*alive* are true at w_1, w_2, w_3, w_4.

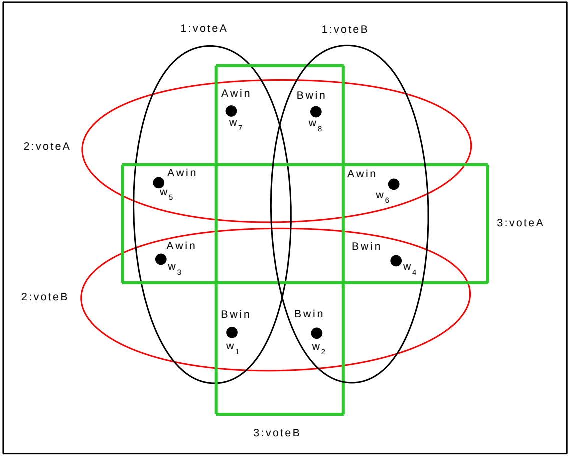
Example 2: voting for a candidate



A and B are the two candidates and 1, 2, 3 are the three voters.

- $[\{1,2\}]Awin$ holds at w_5 and w_7.

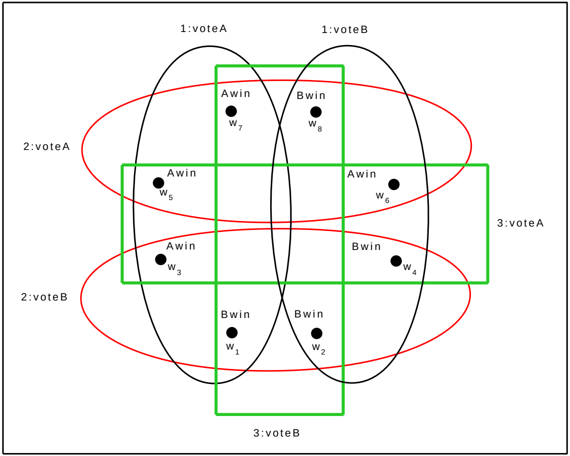
Example 2: voting for a candidate



A and B are the two candidates and $1, 2, 3$ are the three voters.

- $\Diamond[\{1, 2\}]Awin$ holds at w_1-w_8.

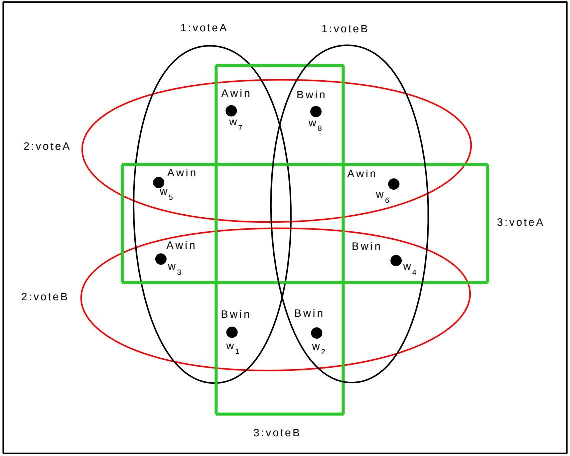
Example 2: voting for a candidate



A and B are the two candidates and 1, 2, 3 are the three voters.

- $\neg\Diamond[\{1\}]Awin \wedge \neg\Diamond[\{1\}]Bwin$ holds at w_1-w_8.

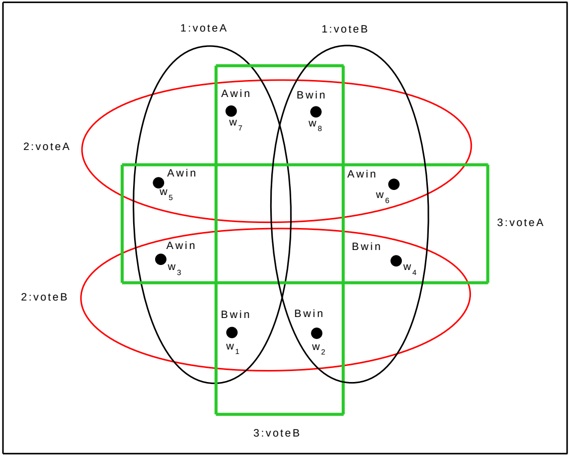
Example 2: voting for a candidate



A and B are the two candidates and $1, 2, 3$ are the three voters.

- $\bigwedge_{i \in \{1,2,3\}} (\neg \Diamond [\{i\}] \mathit{1win} \wedge \neg \Diamond [\{i\}] \mathit{2win})$ holds at w_1-w_8.

Example 2: voting for a candidate



A and B are the two candidates and 1, 2, 3 are the three voters.

- $\bigwedge_{J\subseteq\{1,2,3\},|J|\geq 2}(\Diamond[J]Awin \wedge \Diamond[J]Bwin)$ holds at w_1-w_8.

Some STIT validities

$\Box$ and $[J]$ are S5 modalities

1. $\models ([J]\varphi \land [J]\psi) \to [J](\varphi \land \psi)$
2. $\models [J]\varphi \to \varphi$
3. $\models [J]\varphi \to [J][J]\varphi$
4. $\models \langle J \rangle\varphi \to [J]\langle J \rangle\varphi$
5. $\models (\Box\varphi \land \Box\psi) \to \Box(\varphi \land \psi)$
6. $\models \Box\varphi \to \varphi$
7. $\models \Box\varphi \to \Box\Box\varphi$
8. $\models \Diamond\varphi \to \Box\Diamond\varphi$

Some STIT validities (cont.)

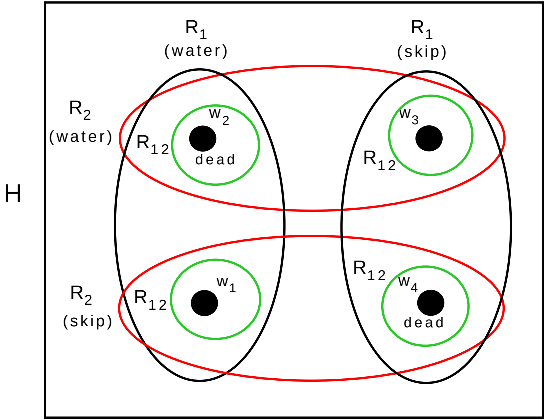
1. $\models \Box\varphi \rightarrow [J]\varphi$
2. $\models \Box\varphi \rightarrow \Box[J]\varphi$
3. $\models \langle \mathtt{AGT} \rangle\varphi \leftrightarrow [\mathtt{AGT}]\varphi$
4. $\models [J]\varphi \rightarrow [J \cup J']\varphi$
5. $\models \Diamond[J]\varphi \rightarrow \Diamond[J \cup J']\varphi$ (cf. coalition monotony in CL)
6. $\models [J]\varphi \rightarrow \Diamond[J]\varphi$ ('do' implies 'can')
7. $\models (\Diamond[J]\varphi \wedge \Diamond[I]\psi) \rightarrow \Diamond[J \cup I](\varphi \wedge \psi)$ if $I \cap J = \emptyset$ (cf. superadditivity in CL)

Discussion: the notion of responsibility

$\varphi \wedge \neg[\mathtt{AGT} \setminus J]\varphi$ expresses a basic notion of **responsibility** of the form “coalition J *could have prevented* a certain state of affairs φ to be true now”.

$$\mathtt{Resp}_J\varphi \stackrel{\text{def}}{=} \varphi \wedge \neg[\mathtt{AGT} \setminus J]\varphi$$

Discussion: the notion of responsibility (cont.)



- $\mathtt{Resp}_1 \neg alive$ and $\mathtt{Resp}_2 \neg alive$ are true at w_2 and w_4.
- $\mathtt{Resp}_{\{1,2\}} \neg alive$ and $\mathtt{Resp}_{\{1,2\}} \neg alive$ are true at w_1 and w_4.

Mathematical properties of STIT

Definition

A logic $\mathcal{L}$ is *finitely axiomatizable* if there is a finite set Ax of formula schemas such that $\varphi \in \mathcal{L}$ iff there is a sequence $(\varphi_1, \ldots, \varphi_n)$ of formulas such that $\varphi_n = \varphi$ and for $1 \leq i \leq n$:

- φ_i is a tautology of propositional calculus or an instance of an axiom in Ax or,
- φ_i is obtained by necessitation from φ_j with $j < i$ or,
- φ_i is obtained by modus ponens from φ_j and φ_k with $j, k < i$.

Theorem ([Herzig & Schwarzentruber, 2008])

STIT with agents and coalitions undecidable and not finitely axiomatizable if there are at least 3 agents

A decidable fragment of STIT

The following fragment of STIT called $\mathsf{DF}^{\mathsf{STIT}}$ is decidable:

$$
\begin{array}{lcll}
\chi & ::= & \bot \mid p \mid \chi \wedge \chi \mid \neg\chi & \text{(propositional formulas)} \\
\psi & ::= & [J]\chi \mid \psi \wedge \psi & \text{("see-to-it" formulas)} \\
\varphi & ::= & \chi \mid \psi \mid \varphi \wedge \varphi \mid \neg\varphi \mid \Diamond\psi & \text{("see-to-it" and "can" formulas)}
\end{array}
$$

For instance, $[\{1\}][\{1,2\}]p$ is in STIT but is not in $\mathsf{DF}^{\mathsf{STIT}}$

Theorem ([Lorini & Schwarzentruber 2009])

The satisfiability problem of DF^{STIT} is NP-complete.

other decidable fragments:
[Schwarzentruber, to appear in Studia Logica]

Epistemic extension of STIT and the concept of uniform strategy

A STIT extension with knowledge

We add modalities for knowledge and a temporal operator to the STIT language:

⇒ $\mathtt{K}_i\varphi$: agent i knows that φ is true

⇒ $\mathtt{X}\varphi$: φ will be true in the next state

Epistemic STIT models

An epistemic STIT-model with discrete time is a tuple $M = (W, \{R_J\}_{J \subseteq \texttt{AGT}}, H, \{E_i\}_{i \in \texttt{AGT}}, F_X, V)$ where:

- W, R_J, H and V are defined as in STIT models;
- For every $i \in \texttt{AGT}$, E_i is an equivalence (epistemic) relation over W;
- $F_X : W \longrightarrow W$.

$\Rightarrow$ $E_i(w) = \{v \ : \ (w, v) \in E_i\}$ are the worlds that agent i considers possible at w

$\Rightarrow$ $F_X(w)$ is the successor of world w

Truth conditions for knowledge

$$M, w \models \mathtt{K}_i\varphi \quad \text{iff} \quad M, v \models \varphi \text{ for all } (w, v) \in E_i$$

$$M, w \models \mathtt{X}\varphi \quad \text{iff} \quad M, F_X(w) \models \varphi$$

Interactions between knowledge and historic necessity: discussion

$\Rightarrow$ *Semantic constraint* S1: For every $w \in W$, $E_i(w) \subseteq H(w)$
$\Rightarrow$ *Corresponding axiom* **PerfectInfo**: $\Box\varphi \rightarrow \mathtt{K}_i\varphi$

Perfect information about the situation of interaction

Interactions between knowledge and historic necessity: discussion (cont.)

⇒ *Semantic constraint* S2: For every $w \in W$, $E_i(w) \subseteq R_i(w)$
⇒ *Corresponding axiom* **PerfectInfo**+**ActAware**: $[i]\varphi \rightarrow \mathtt{K}_i\varphi$

Perfect information about the structure of interaction + agents' knowledge about their current choice (the only uncertainty is about choices of others)

Theorem
If S2 then S1.

Interactions between knowledge and historic necessity: discussion (cont.)

More realistic principle:

⇒ *Semantic constraint (**KChoiceDet**)* S3: For every $w \in W$, if $v \in H(w)$ and $u \in E_i(v)$ then $u \in E_i(w)$
⇒ *Corresponding axiom*: $\texttt{K}_i\varphi \rightarrow [i]\texttt{K}_i\varphi$

Knowledge is choice determinate

The concept of uniform strategy

For an agent i to have the **power of** ensuring φ, i must have both: the objective capability to achieve φ and, the discretion (awareness) over his capability [Castelfranchi 2003, Barnes 1988]

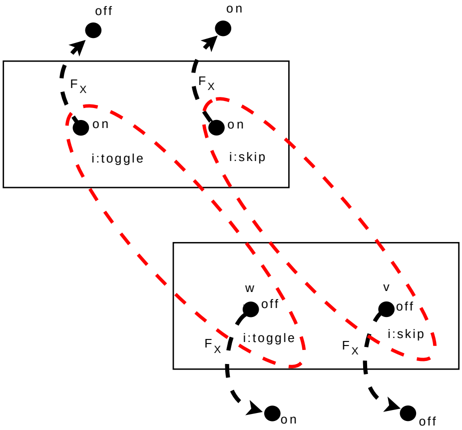
The concept of uniform strategy (cont.)

- *DE DICTO* sentence: "*i* knows that there is some action such that if he chooses it, he will ensure φ in the next state"
- *DE RE* sentence: "there is some action such that if agent *i* chooses it, he knows that he will ensure φ in the next state"

$\Rightarrow$ DE DICTO: $\mathtt{K}_i \Diamond [i] \mathtt{X} \varphi$
$\Rightarrow$ DE RE: $\Diamond \mathtt{K}_i [i] \mathtt{X} \varphi$

Only formula $\Diamond \mathtt{K}_i [i] \mathtt{X} \varphi$ captures a proper concept of agent *i*'s *power of* ensuring φ or agent *i*'s *uniform strategy* over φ

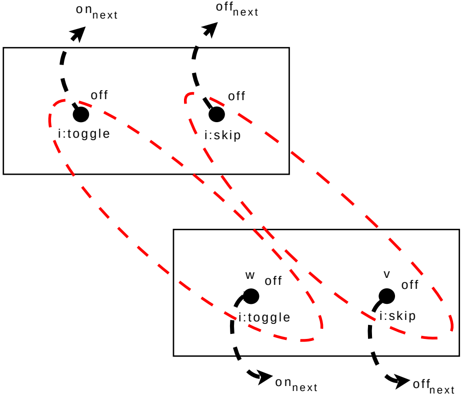
Example: an agent trying to switch on a light



Agent i can either *toggle* or *skip* (do nothing) but he is uncertain about the current state of the light (*on* or *off*)

- $\mathtt{K}_i\Diamond[i]\mathtt{X}on$ is true at w and v.
- $\Diamond\mathtt{K}_i[i]\mathtt{X}on$ is false at w and v.

Example: an agent trying to switch on a light (cont.)



Agent i can either *toggle* or *skip* (do nothing) and he knows the current state of the light

- $\mathtt{K}_i \Diamond [i] \mathtt{X} \mathit{on}$ is true at w and v.
- $\Diamond \mathtt{K}_i [i] \mathtt{X} \mathit{on}$ is true at w and v.

References

- Alur, R., Henzinger, T., Kupferman, O. (2002). *Alternating-time temporal logic*. Journal of the ACM, 49:672-713.
- Belnap, N. Perloff, M., Xu, M. (2001). *Facing the future: agents and choices in our indeterminist world*. Oxford University Press.
- Broersen, J. Herzig, A., Troquard, N. (2006). *Embedding Alternating-time Temporal Logic in strategic STIT logic of agency*. Journal of Logic and Computation, 16(5):559-578.
- Ciuni, R. and Zanardo, A. (2010). *Completeness of a Branching-Time Logic with Possible Choices*. Studia Logica 96(3)
- Goranko, V. (2001). *Coalition games and alternating temporal logics*. In Proceedings of the 8th conference on Theoretical Aspects of Rationality and Knowledge (TARK'01), Morgan Kaufmann Publishers Inc, p. 259-272.
- Goranko, V., Jamroga, W. (2004). *Comparing Semantics for Logics of Multi-agent Systems*. Synthese, 139(2):241-280.

References

- Herzig, A., Schwarzentruber, F. (2008). *Properties of logics of individual and group agency*. Proceedings of Advances in Modal Logic 2008 (AiML 2008).
- Horty, J. F., Belnap, N. (1995). *The deliberative STIT: A study of action, omission, and obligation*. Journal of Philosophical Logic, 24(6):583-644.
- van der Hoek, W., Wooldridge, M. (2003). *Cooperation, Knowledge, and Time: Alternating-Time Temporal Epistemic Logic and its Applications*. Studia Logica, 75(1):125-157.
- Horty, J. F. (2001). *Agency and Deontic Logic*, Oxford University Press.
- Lorini, E., Schwarzentruber, F. (2011). *A Logic for Reasoning about Counterfactual Emotions*. Artificial Intelligence, 175(3-4): 814-847.
- Pauly, M. (2002). *A modal logic for coalitional power in games*. Journal of Logic and Computation, 12(1):149-166.
- Pauly, M. (2001). *Logic for Social Software*. PhD thesis, University of Amsterdam, The Netherlands.
- Schwarzentruber, F. (to appear). *Complexity Results of STIT Fragments*. Studia Logica.
- Xu, M. (1995). *On the basic logic of STIT with a single agent*. Journal of Symbolic Logic 60, 459-483.