What groups *do*, *can* do, and *know* they can do: an analysis in normal modal logics

Jan Broersen* — **Andreas Herzig**** — **Nicolas Troquard*****

**Utrecht University, Department of Information and Computing Science*
PO Box 80.089, 3508 TB Utrecht (The Netherlands)
broersen@cs.uu.nl

***Université de Toulouse, CNRS, Institut de recherche en informatique de Toulouse*
118 route de Narbonne, 31062 Toulouse Cedex 9 (France)
Andreas.Herzig@irit.fr

****University of Liverpool*
Department of Computer Science, Liverpool L69 3BX (UK)
Nicolas.Troquard@liverpool.ac.uk

ABSTRACT. *We investigate a series of logics that allow to reason about agents' actions, abilities, and their knowledge about actions and abilities. These logics include Pauly's Coalition Logic* CL*, Alternating-time Temporal Logic* ATL*, the logic of 'seeing-to-it-that' (*STIT*), and epistemic extensions thereof. While complete axiomatizations of* CL *and* ATL *exist, only the fragment of the* STIT *language without temporal operators and without groups has been axiomatized by Xu (called* Ldm*). We start by recalling a simplification of the* Ldm *that has been proposed in previous work, together with an alternative semantics in terms of standard Kripke models. We extend that semantics to groups via a principle of superadditivity, and give a sound and complete axiomatization that we call* Ldm^G*. We then add a temporal 'next' operator to* Ldm^G*, and again give a sound and complete axiomatization. We show that* Ldm^G *subsumes coalition logic* CL*. Finally, we extend these logics with standard S5 knowledge operators. This enables us to express that agents see to something under uncertainty about the present state or uncertainty about which action is being taken. We focus on the epistemic extension of* $\mathbf{X}$-Ldm^G*, noted* E-$\mathbf{X}$-Ldm^G*. In accordance with established terminology in the planning community, we call this extension of* $\mathbf{X}$-Ldm^G *the conformant* $\mathbf{X}$-Ldm^G*. The conformant* $\mathbf{X}$-Ldm^G *enables us to express that agents are able to perform a uniform strategy. We conclude that in that respect, our epistemic extension of* $\mathbf{X}$-Ldm^G *is better suited than epistemic extensions of* ATL*.*

KEYWORDS: *ATL, CL, STIT, agency, epistemic logic, uniform strategies*

DOI:10.3166/JANCL.19.1–29

1. Introduction

[1]

The *theory of agents and choice in branching time* (Belnap *et al.*, 2001) is maybe the most prominent logic account of agency in the philosophy of action. It is a very rich framework and it appears natural to analyse the notions of agency that are discussed in logics for computer science and logics for social choice theory (Horty, 2001; Troquard, 2007). The aim of this paper is to unify the reasoning about what groups do, can do and know they can do.

The theory of agents and choice in branching time is a family of logics that formalise the linguistic constructions of the form "agent i sees to it that φ holds". (For this reason, we will generally refer to it as *stit theory*.) The term *deliberative stit theories* refers to the particular logic of Chellas STIT operators and deliberative STIT operator. The validities of the logic were axiomatized by Xu, who called his logic Ldm (Xu, 1998; Belnap *et al.*, 2001). We take his logic as a starting point. First we extend it to a group version that we call Ldm^G by adding a principle similar to what is 'superadditivity' in social choice theory (Abdou *et al.*, 1991). In a second step we combine Ldm^G with the logic of the next-time operator **X**. For easy reference, we adopt the name $\mathbf{X}$-Ldm^G for the resulting logic.

Coalition Logic, CL for short, was proposed in (Pauly, 2001; Pauly, 2002) as a logic for reasoning about social procedures characterized by complex strategic interactions between agents, be it in terms of individuals or in terms of groups. Examples of such procedures are fair-division algorithms or voting processes. CL facilitates reasoning about abilities of coalitions in games by extending classical logic with operators $\langle\!\langle J \rangle\!\rangle \mathbf{X}\varphi$ for groups of agents J, reading: "the coalition J has a joint strategy to ensure that φ".[2] We shall show how CL can be naturally embedded in our variant of stit theory. The embedding of CL in $\mathbf{X}$-Ldm^G is an interesting result since it shows how to extend CL with capabilities of reasoning about what a coalition is actually doing (as opposed to what it *could* do).

In social choice theory, in particular since Harsanyi, the interaction between ability models and epistemic models has been a main focus of research. It has been realized that intentionality of action presupposes awareness or knowledge of the means by

1. A preliminary version of the present paper was presented at TARK 2007 (Broersen *et al.*, 2007)
AH: [Dominique is going to place this somewhere.]
2. We have chosen a uniform notation, deviating from the original CL and STIT notation:
– We use $\langle\!\langle J \rangle\!\rangle \mathbf{X}\varphi$ as an alternative notation for Pauly's non-normal operator $[J]\varphi$, because the new syntax highlights both the quantifier combination $\exists$-$\forall$ underlying the semantics, and the temporal aspect.
– We use $[J]\varphi$ as an alternative notation for the STIT operator $[J\, cstit : \varphi]$, thereby emphasizing that this is a normal modal necessity operator.

which effects are ensured. Philosophers refer to this ability of agents as having the *power* to ensure a condition. So, in order to say that an agent 'can' or 'has the power to' ensure a condition, there should not only be an action in the agent's repertoire that ensures the condition, the agent should also know how to choose the action (see (Lorini *et al.*, 2007) for a discussion).

More recently the issue of 'knowing how to act' has come up in the logic ATEL (van der Hoek *et al.*, 2002) which is the epistemic extension of the logic of strategic ability ATL (Alur *et al.*, 2002). The problem is often referred to as the problem of *uniform strategies*. In particular, ATEL does not allow to distinguish the situations where:

1) the agent a knows it has *some* action/choice in its repertoire that ensures φ, while, possibly, it does not know *which* choice ensures φ.
2) the agent a 'knows how to' / 'can' / 'has the power to' ensure φ.

The semantic setup of ATEL, with indistinguishability relations over states is too coarse to distinguish these situations. In the present paper indistinguishability relations range over 'indexes'. These more detailed semantic structures do enable us to distinguish the above two situations.

In this paper we do not reason about series of choices, alias strategies, which is why our starting point is CL instead of ATL. We extend $\mathbf{X}$-Ldm^G with an S5 modal operator for knowledge and show that the resulting complete logic, that we refer to as E-$\mathbf{X}$-Ldm^G, solves the problem of uniform strategies. Furthermore, the epistemic extension enables us to define a notion of 'seeing to it under uncertainty'. In accordance with established terminology in the planning literature, we also call this version of STIT, the 'conformant STIT'.

The paper is organized as follows. In Sections 2, 3 and 4 we respectively recall CL and ATL, their epistemic extensions, and Xu's axiomatization Ldm of the atemporal and individual fragment of deliberative stit theories. In Section 5 we extend his axiomatization to groups (Ldm^G), and in Section 6 we extend it with time ($\mathbf{X}$-Ldm^G) and provide an embedding of CL. In Section 7 we then straightforwardly extend it with knowledge (E-$\mathbf{X}$-Ldm^G). Finally, in Section 8 we discuss what is needed to completely axiomatize STIT-models.

Except in Section 4 where an infinite number of agents is used, AGT is throughout the paper a finite set of agents, and PRP is a countable set of atomic formulas. For all the logical systems that we consider, the standard notions of theoremhood, consistency are defined as usual, as well as validity and satisfiability.

2. Background: Coalition Logic CL and Alternating-time Temporal Logic ATL

The syntax of Coalition Logic is as follows:

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

where p ranges over PRP and J ranges over the subsets of AGT. The other Boolean connectives are defined as usual.

2.1. *Coalition model semantics*

For sets of agents $J \subseteq AGT$, $\overline{J}$ denotes the complement of J w.r.t. AGT, i.e. $\overline{J} = AGT \setminus J$.

DEFINITION 1 (EFFECTIVITY FUNCTION). — *Given a set of states S, an* effectivity function *is a function $E : 2^{AGT} \longrightarrow 2^{2^S}$. An effectivity function is said to be:*

– AGT-maximal *iff for all $Q \subseteq S$, if $S \setminus Q \notin E(\emptyset)$ then $Q \in E(AGT)$;*
– outcome monotonic *iff for all $Q \subseteq Q' \subseteq S$ and for all $J \subseteq AGT$, if $Q \in E(J)$ then $Q' \in E(J)$;*
– superadditive *iff for all Q_1, Q_2, J_1, J_2 such that $J_1 \cap J_2 = \emptyset$, $Q_1 \in E(J_1)$ and $Q_2 \in E(J_2)$ imply that $Q_1 \cap Q_2 \in E(J_1 \cup J_2)$.*

Intuitively, every $Q \in E(J)$ is a possible outcome for which J is effective. That is, J can force the world to be in some state of Q at the next step.

DEFINITION 2 (PLAYABLE EFFECTIVITY FUNCTION). — *An effectivity function E is said to be* playable *iff*

1) $\forall J \subseteq AGT, \emptyset \notin E(J)$; *(Liveness)*
2) $\forall J \subseteq AGT, S \in E(J)$; *(Termination)*
3) E is AGT-maximal;
4) E is outcome-monotonic; and
5) E is superadditive.

These five properties of effectivity functions are independent.

DEFINITION 3. — *An* effectivity structure *is a mapping $E : S \longrightarrow (2^{AGT} \longrightarrow 2^{2^S})$ such that every $E(s)$ is an effectivity function. An effectivity structure is* playable *if it is playable for every effectivity function $E(s)$.*

We write $E_s(J)$ instead of $E(s)(J)$.

DEFINITION 4. — *A* coalition model *is a pair $((S, E), V)$ where:*

– *S is a nonempty set of states;*
– *$E : S \longrightarrow (2^{AGT} \longrightarrow 2^{2^S})$ is a* playable effectivity structure*;*
– *$V : S \longrightarrow 2^{PRP}$ is a valuation function.*

Truth conditions are standard for the Boolean connectives and for atomic formulas. For the modal connective we have:

$$M, s \models \langle\!\langle J \rangle\!\rangle \mathbf{X}\varphi \text{ iff } \{s' \mid M, s' \models \varphi\} \in E_s(J).$$

Coalition Logic is a weak modal logic that does not validate the axiom of normality $\langle\!\langle J \rangle\!\rangle \mathbf{X}(\varphi \to \psi) \to (\langle\!\langle J \rangle\!\rangle \mathbf{X}\varphi \to \langle\!\langle J \rangle\!\rangle \mathbf{X}\psi)$ and hence does not have a standard

Kripke semantics. The above semantics is a so-called neighborhood semantics. For instance, the outcome monotonicity property can be reformulated as reachability of neighborhoods being closed under supersets.

2.2. *Game semantics*

In (Pauly, 2001), Marc Pauly investigates an alternative semantics for CL in terms of game structures. We introduce some notations and results that are going to be useful later.

DEFINITION 5. — *A* strategic game *is a tuple $G = (S, \Sigma, o)$ where S is a nonempty set, Σ associates a nonempty set Σ_i to every agent $i \in AGT$ (a set of choices for i), $o : \prod_{i \in AGT} \Sigma_i \longrightarrow S$ is an outcome function which associates an outcome state in S with every combination of agents' choices (choice profile).*

We write Γ_S for the set of strategic games over the set of states S. For convenience, for every coalition $J \subseteq AGT$, by σ_J we denote a tuple of choices $(\sigma_i)_{i \in J}$ in a strategic game where every $\sigma_i \in \Sigma_i$. We write $\sigma_J \cdot \sigma_{\overline{J}}$ for the concatenation of σ_J and $\sigma_{\overline{J}}$.

It appears that there is a strong link between a coalition model (whose effectivity structure is *playable* by definition) and a strategic game. We define the effectivity function of a strategic game as follows.

DEFINITION 6. — *Given a strategic game $G = (S, \Sigma, o)$, the* effectivity function *$E_G : 2^{AGT} \longrightarrow 2^{2^S}$* of *$G$ is defined as: for every coalition $J \subseteq AGT$, and every $Q \subseteq S$, $Q \in E_G(J)$ iff there is a choice tuple σ_J such that for every $\sigma_{\overline{J}}$ we have $o(\sigma_J \cdot \sigma_{\overline{J}}) \in Q$.*

Pauly then gives the following characterization:

THEOREM 7 (Pauly, 2001). — *An effectivity function E is playable iff it is the effectivity function of some strategic game.*

We now define a new class of models for CL. It merely consists of a set of states, a function associating every state with a strategic game and an evaluation function.

DEFINITION 8. — *A* game model *is a triple $G_M = (S, \gamma, v)$ where:*

- *S is a nonempty set of states;*
- *$\gamma : S \longrightarrow \Gamma_S$ associates every state with a strategic game;*
- *$v : S \longrightarrow 2^{PRP}$.*

For every $s \in S$, we write o_s for the outcome function of the strategic game $\gamma(s)$, and we write $\Sigma_{s,J}$ for the set of choices of coalition J in $\gamma(s)$.

Truth of CL formulas in a game model is as expected. In particular, for a game model $G_M = (S, \gamma, v)$ and a state $s \in S$:

$M_G, s \models \langle\!\langle J \rangle\!\rangle \mathbf{X}\varphi$ iff there is $\sigma_J \in \Sigma_{s,J}$ such that for all $\sigma_{\overline{J}} \in \Sigma_{s,\overline{J}}, M, o_s(\sigma_J \cdot \sigma_{\overline{J}}) \models \varphi$. ▮

The function γ associates a state with a strategic game whilst a playable effectivity structure in a Coalition model associates a state with an effectivity function. Hence, from Theorem 7, it is easy to see that the semantics are equivalent.

2.3. *Axiomatization*

The set of formulas that are valid in coalition models is completely axiomatized by the following principles (Pauly, 2001).

(ProTau)	any sufficient set of propositional logic schemas
$(\bot)$	$\neg⟪J⟫\mathbf{X}\bot$
$(\top)$	$⟪J⟫\mathbf{X}\top$
(N)	$\neg⟪\emptyset⟫\mathbf{X}\neg\varphi \to ⟪AGT⟫\mathbf{X}\varphi$
(M)	$⟪J⟫\mathbf{X}(\varphi \wedge \psi) \to (⟪J⟫\mathbf{X}\varphi \wedge ⟪J⟫\mathbf{X}\psi)$
(S)	$⟪J_1⟫\mathbf{X}\varphi \wedge ⟪J_2⟫\mathbf{X}\psi \to ⟪J_1 \cup J_2⟫\mathbf{X}(\varphi \wedge \psi)$ if $J_1 \cap J_2 = \emptyset$
(MP)	from φ and $\varphi \to \psi$ infer ψ
(RE)	from $\varphi \leftrightarrow \psi$ infer $⟪J⟫\mathbf{X}\varphi \leftrightarrow ⟪J⟫\mathbf{X}\psi$

THEOREM 9 (Pauly, 2001). — *The principles (ProTau), $(\bot)$, $(\top)$, (N), (M), (S), (MP) and (RE) are complete with respect to the class of all coalition models.*

Note that the (N) axiom corresponds to AGT-maximality of the effectivity structures. It says that if a formula is not settled false, the coalition of all agents (AGT) can always coordinate their choices to make it true. The axiom (S) corresponds to superadditivity and says that two disjoint coalitions can combine their efforts to ensure a conjunction of properties. Note that from (S) and $(\bot)$ it follows that $⟪J_1⟫\mathbf{X}\varphi \wedge ⟪J_2⟫\mathbf{X}\neg\varphi$ is inconsistent for disjoint J_1 and J_2. So, two disjoint coalitions cannot ensure opposed facts. This property is known as 'regularity'.

2.4. *ATL as a strategic extension of CL*

While the CL expression $⟪J⟫\mathbf{X}\varphi$ is about the outcome of a single action to be chosen by each agent in J, the logic ATL is about sequences of such actions, alias extensive form. The modal operators **X** ('next'), **U** ('until'), **G** ('henceforth')... of Linear-time Temporal Logic LTL allow to speak about the temporal properties of such strategies.

CL and ATL were proposed independently, and it was shown only later that the former can be viewed as an extension of the latter (Goranko, 2001). Indeed, the ATL-formula $\langle\langle J\rangle\rangle\mathbf{X}\varphi$ says that there exists a joint strategy of J such that when performed by its members, φ will hold at the next state, which boils down to the existence of a joint action of J ensuring φ at the next state. The latter is exactly the reading of the CL-formula $\langle\!\![J]\!\!\rangle \mathbf{X}\varphi$.

Beyond that, the ATL-formula $\langle\langle J\rangle\rangle \varphi \mathbf{U} \psi$ says that there exists a joint strategy of J such that when performed, φ will hold until ψ holds, and $\langle\langle J\rangle\rangle \mathbf{G}\varphi$ says that there exists a joint strategy of J ensuring that φ holds henceforth.

We do not go into the details of semantics and axiomatics of ATL because extensive form strategies are not relevant for the rest of the paper and refer the reader to (Goranko *et al.*, 2006). Our reason to mentioning ATL here is that its epistemic extension has been studied in detail in the agents community, motivating our epistemic extension of CL in Section 7.

3. Background: Epistemic extensions of CL and ATL

The idea of combining a logic for multi-agency with a logic for knowledge naturally stems from game theory (Osborne *et al.*, 1994).

Since the epistemic extension ATEL of Alternating-time Temporal Logic (ATL) had been proposed in (van der Hoek *et al.*, 2002) several refinements have been investigated. In general these logics try to give an account of what game theory calls *uniform strategies*. For an overview see for instance (Jamroga *et al.*, 2004; Jamroga *et al.*, 2007).

The first logic to extend ATL with an epistemic modality was ATEL. Here we prefer to refer to that system as E-ATL for reasons of uniformity of notation. However, we focus on the language fragment where the 'next' is the only temporal operator, and present the logic E-CL. While the latter has not been studied before in the literature, it is the simplest system where the basic features and problems of epistemic extensions of both CL and ATL can be highlighted.

E-CL is the syntactic fusion of the basic epistemic logic S5 and CL. Semantically, for each $i \in AGT$ we add a family of equivalence relations $\sim_i \subseteq S \times S$ to the models. Validity and satisfiability are defined as for CL and S5.

We do not say anything here about the different notions of group knowledge such as distributed knowledge and common knowledge that would be required by a full account. The reason is that the problems can already be highlighted in the individual case. We also do not investigate axiomatizations of E-CL, and instead focus on its limitations in expressiveness.

The problem of representing uniform strategies concerns the disambiguation of the notion of *knowing a strategy*: ATEL is not expressive enough to distinguish the sentence

> "*for all* epistemically indistinguishable states, *there exists* a strategy of J that leads to φ".

from

> *there exists* a strategy σ of the coalition J such that *for all* states epistemically indistinguishable for J, σ leads to φ."

The former is a $\forall$-$\exists$ schema of "knowing a strategy", in philosophy referred to as the *de dicto* reading. It is opposed to the *de re* reading exemplified by the latter sentence, which is a $\exists$-$\forall$ schema.

In Section 7 we shall argue that instead of extending CL and ATL by epistemic modalities, it is more appropriate for modelling purposes to choose STIT-logics as a starting point of such an enterprise.

4. Background: Xu's logic Ldm

The logics of *seeing to it that* are about agents making choices among alternatives in a branching time setting. The starting point for any *stit* theory is that *acting* is the same as *ensuring* the actual world is among a set of possible worlds that satisfy the property being secured by the action. This is the *stit paraphrase thesis* (Belnap *et al.*, 1988). For instance, "agent i closes de door" can be paraphrased as "i sees to it that the door is closed". In a logical language, this is rendered by the expression $[i]\varphi$, reading 'agent i sees to it that φ'. It follows that one of the central axioms for *stit* is the so called 'success axiom': $[i]\varphi \rightarrow \varphi$.

The traditional semantics of *stit* theories is extensively studied in Belnap *et al.* (Belnap *et al.*, 2001). It consists of a branching-time structure (BT) augmented by the set of agents and a choice function (AC). Since in this section we are only interested in axiom systems, we postpone the definition to Section 8.1.

In the sequel we recall the *stit* logic axiomatized in (Belnap *et al.*, 2001, Chap. 17), viz. the (atemporal and individual version of) Ldm, as well as its simplification proposed in (Balbiani *et al.*, 2008b). For historical reasons, Ldm is sometimes referred to as the logic of Chellas STIT.

The language of Ldm is defined by the following BNF:

$$\varphi \ ::= \ p \mid \neg\varphi \mid (\varphi \vee \varphi) \mid [i]\varphi \mid \Box\varphi$$

where p ranges over PRP and i ranges over AGT. $[i]\varphi$ is read 'i sees to it that φ', and $\Box\varphi$ is read 'φ is settled'.[3] In this section, we consider an infinite number of agents. For convenience, AGT is identified with an initial segment of the non-negative integers: $AGT = \{0, 1, 2, \ldots\}$.

Moreover we suppose here that $|AGT| \geq 2$, i.e. there are at least agents 0 and 1.

3. Usually $\Box\varphi$ is noted $Sett : \varphi$, and $[i]\varphi$ is noted $[i\ cstit\colon \varphi]$. Our present notation allows for the dual constructions $\Diamond\varphi$ and $\langle i\rangle\varphi$, abbreviating $\neg\Box\neg\varphi$ and $\neg[i]\neg\varphi$, respectively.

4.1. *Xu's Axiomatics of* **Ldm**

In (Belnap *et al.*, 2001, Chap. 17) Xu provides an axiomatization of BT+AC validities in terms of a family of axiom schemas (AIA_k). These capture a central idea of multi-agent *stit* theories: agents' choices are independent.

S5($\Box$) the axiom schemas of S5 for $\Box$

S5(i) the axiom schemas of S5 for every $[i]$

($\Box\!\to\! i$) $\Box\varphi \to [i]\varphi$

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

The last item is a family of *axiom schemes for independence of agents* that is parameterized by the integer k.[4]

REMARK 10. — As (AIA_{k+1}) implies (AIA_k), the family of schemas can be replaced by the single ($\text{AIA}_{|AGT|-1}$) when AGT is finite. □

Xu's system has the standard inference rules of modus ponens and necessitation for $\Box$. From the latter necessitation rules for every $[i]$ follow by axiom ($\Box\!\to\! i$).

THEOREM 11 (Belnap *et al.*, 2001, Chapter 17). — *An* **Ldm** *formula φ is valid in BT+AC models iff φ is provable from the schemas S5($\Box$), S5(i), ($\Box\!\to\! i$), and (AIA_k) by the rules of modus ponens and $\Box$-necessitation.*

4.2. *An alternative axiomatics of* **Ldm**

In (Balbiani *et al.*, 2008b) an alternative axiomatics is given. To that end, it is first proved that (AIA_k) can be replaced by the family of axiom schemas

(AAIA_k) $\Diamond\varphi \to \langle k\rangle \bigwedge_{0\leq i<k}\langle i\rangle\varphi$ for $k \geq 1$

(AAIA_k) is called the alternative axiom schema for independence of agents. Just as Xu's (AIA_k), (AAIA_k) involves $k+1$ agents.

It can then be seen that the equivalence $\Diamond\varphi \leftrightarrow \langle 1\rangle\langle 0\rangle\varphi$ is provable from (AAIA_1), ($\Box\!\to\! i$) and S5($\Box$). This suggests that $\Box\varphi$ can be viewed as an abbreviation of $[1][0]\varphi$. We can take this as an axiom schema:

4. Xu's original formulation of (AIA_k) resorts to k-ary difference predicates that are part of the language expressing that $i_0, \ldots, i_k$ are all distinct. They are defined from an equality predicate $=$ whose domain is AGT. In consequence Xu's axiomatics has to contain axioms for equality. We here preferred not to introduce equality in order to stay with the same logical language throughout.

Def(□) $\Box\varphi \leftrightarrow [1][0]\varphi$

It then can be proved that under Def(□), axiom (AAIA$_k$) can be replaced by the family of axiom schemas of general permutation:

(GPerm$_k$) $\langle l\rangle\langle m\rangle\varphi \rightarrow \langle n\rangle \bigwedge_{i\leq k, i\neq n}\langle i\rangle\varphi$ for $k \geq 0$

THEOREM 12 (Balbiani *et al.*, 2008b). — *A formula of* **Ldm** *is valid in BT+AC models iff it is provable from S5(i), Def(□), and (GPerm$_k$) by the rules of modus ponens and* $[i]$*-necessitation.*

Note that similar to Xu's axiomatization, if AGT is finite then the single schema (GPerm$_{|AGT|-1}$) is sufficient.

REMARK 13. — If $AGT = \{0, 1\}$ then the BT+AC validities are axiomatized by Def(□), S5(1), S5(2), and $\langle 1\rangle\langle 0\rangle\varphi \leftrightarrow \langle 0\rangle\langle 1\rangle\varphi$. Moreover, the Church-Rosser axiom $\langle 0\rangle[1]\varphi \rightarrow [1]\langle 0\rangle\varphi$. can be proved from S5(1), S5(2) and (GPerm$_1$). Therefore Ldm logic with two agents is a so-called product logic, alias a two-dimensional modal logic (Marx, 1999; Gabbay *et al.*, 2003). Such product logics are characterized by the permutation axiom $\langle 0\rangle\langle 1\rangle\varphi \leftrightarrow \langle 1\rangle\langle 0\rangle\varphi$ together with the Church-Rosser axiom. Hence the logic of the two-agent Ldm is nothing but the product $S5^2 = S5 \otimes S5$. □

4.3. *A simple semantics of* Ldm

All axiom schemas are in the Sahlqvist class (Blackburn *et al.*, 2001), 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 $i \in AGT$ an equivalence relation R_i on W satisfying the *general permutation property*:

- for all $w, v \in W$ and for all $l, m, n \in AGT$, 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 $i \in AGT \setminus \{n\}$;

– V is a mapping from PRP to the set of subsets of W.

When wR_iv then agent i's current choice at w admits v, or allows v, where the verb 'allow' has to be taken in a non-deontic sense. In other words, $R_i(w)$ is the set of outcomes that are possible given i's current choice at w.

We have the usual truth condition for the modal operator:

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

and the usual definitions of validity and satisfiability.

It was shown in (Balbiani *et al.*, 2008b) that the problem of deciding satisfiability of a formula of Ldm is NEXPTIME-complete if AGT contains at least two agents.

5. What groups *do*: the group extension Ldm^G of Xu's Ldm

Xu's axiomatics of Section 4 did not take into account group agency. We now also consider full coalitions instead of just individual agents.

The logic Ldm^G has the following syntax, where p ranges over elements of the set of atomic formulas PRP, and J ranges over the set of subsets of AGT:

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

The intended meaning of the formula $[J]\varphi$ is 'agents J see to it that φ'. In particular, if the empty set of agents ensures that φ then we say that φ is 'settled true', or 'historically necessary'.[5]

Other Boolean connectives are defined by abbreviations, as usual. The abbreviation $\langle J\rangle\varphi =_{def} \neg[J]\neg\varphi$ expresses that the current choices of the members of J allow φ (where 'allow' has to be taken in a non-deontic sense).

5.1. *Axiomatization*

We give the following axiom schemas for Ldm^G.

(ProTau)	any sufficient set of propositional logic schemas	
S5($[J]$)	any sufficient set of S5-schemas, for every $[J]$	
(Mon)	$[J]\varphi \to [J']\varphi$	if $J \subseteq J'$
(Indep)	$[J][\overline{J}]\varphi \to [\emptyset]\varphi$	

Axiom (Indep) characterizes independence of agents: whenever a group J sees to it that the other agents $\overline{J}$ see to it that φ, then that can only be because φ holds trivially, in the sense that φ is settled true (alias historically necessary).

We also assume the standard inference rules of modus ponens, and necessitation for $[\emptyset]$. From the latter necessitation for every $[J]$ follows by the inclusion axiom (Mon).

Note that the converse of (Indep) can be proved from (Mon), S5($\emptyset$) and S5(AGT). Hence, we have $\vdash [\emptyset]\varphi \leftrightarrow [J][\overline{J}]\varphi$.

---

5. The latter terms are from the stit literature. The modal operator of historic necessity $\Box$ there corresponds to our $[\emptyset]$.

5.2. *Kripke semantics*

An Ldm^G-model is a tuple $\mathcal{M} = (W, R, \pi)$, where:

– W is a nonempty set of worlds (alias indexes);
– $R : 2^{AGT} \longrightarrow 2^{(W \times W)}$ associates equivalence relations R_J to every coalition $J \subseteq AGT$ such that:
  - if $J \subseteq J'$ then $R_{J'} \subseteq R_J$
  - $R_\emptyset \subseteq R_J \circ R_{\overline{J}}$
– $\pi : W \longrightarrow 2^{PRP}$ is a valuation function.

As for Ldm, when wR_Jv then group J's current choice at w has v as a possible outcome ('admits' v). $R_{J'} \subseteq R_J$ for $J \subseteq J'$ means that bigger groups have finer choice partitions. $R_\emptyset \subseteq R_J \circ R_{\overline{J}}$ means that historically possible worlds can be attained by combining admission by any group and its complement.

The above conditions on R entail that $R_\emptyset = R_J \circ R_{\overline{J}}$ for every J.

As the R_J are equivalence relations, by (Mon) every R_J partitions $R_\emptyset$. More precisely, every $R_{J_1 \cup J_2}$ is a sub-partition of both R_{J_1} and R_{J_2}, that is, $R_{J_1 \cup J_2}$ partitions the *intersection* classes of R_{J_1} and R_{J_2}. The finest partition is R_{AGT}.

Note that it does *not* follow that the partition R_{AGT} consists just of the intersections of all partitions R_i. More in general it does not follow that the partition of $R_{J_1 \cup J_2}$ is exactly made up of the intersections of the classes for R_{J_1} and R_{J_2}. Again, this is similar to the situation in CL.

The truth conditions are:

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

together with the usual definitions for the other operators.

5.3. *Properties of* Ldm^G

THEOREM 14. — *Ldm^G is sound and complete w.r.t. the class of Ldm^G-models.*

PROOF. — Soundness and completeness follow from Sahlqvist's Theorem: the semantic conditions correspond, via Sahlqvist's correspondence theory, to Ldm^G axioms, cf. (Blackburn *et al.*, 2001, Th. 2.42). ■

In (Schwarzentruber, 2007; Balbiani *et al.*, 2008a) it is shown that satisfiability of formulas of Ldm^G and of $\mathbf{X}$-Ldm^G is decidable, and that its complexity is NEXPTIME-complete if AGT contains more than one agent.

For every J, the (independently) combined choices of J and $\overline{J}$ *cover* the whole space of possible outcomes $R_\emptyset$. The next lemma states that this more generally holds for any two *disjoint* sets of agents.

LEMMA 15. — $\vdash \langle\emptyset\rangle\varphi \to \langle J_1\rangle\langle J_2\rangle\varphi$ *if* $J_1 \cap J_2 = \emptyset$.

PROOF. — By (Indep) we have $\vdash \langle\emptyset\rangle\varphi \to \langle J_1\rangle\langle\overline{J_1}\rangle\varphi$. Then, by hypothesis, $J_1 \cap J_2 = \emptyset$, or equivalently $J_2 \subseteq \overline{J_1}$. Thus, by (Mon), $\vdash \langle\overline{J_1}\rangle\varphi \to \langle J_2\rangle\varphi$. We obtain $\vdash \langle J_1\rangle\langle\overline{J_1}\rangle\varphi \to \langle J_1\rangle\langle J_2\rangle\varphi$ by standard modal principles for $[J_1]$. We conclude that $\vdash \langle\emptyset\rangle\varphi \to \langle J_1\rangle\langle J_2\rangle\varphi$. ■

As the modal operator $\langle\emptyset\rangle$ expresses historic possibility, the formula $\langle\emptyset\rangle[J]\varphi$ can be read 'J can ensure that φ', or 'J has the ability to ensure φ'. The next property highlights that J's ability to ensure φ has to be identified with the other agents $\overline{J}$ allowing J to ensure φ.

LEMMA 16. — $\vdash \langle\emptyset\rangle[J]\varphi \leftrightarrow \langle\overline{J}\rangle[J]\varphi$.

PROOF. — From the left to the right, by (Indep) we have $\vdash \langle\emptyset\rangle[J]\varphi \to \langle\overline{J}\rangle\langle J\rangle[J]\varphi$, and the subformula $\langle J\rangle[J]\varphi$ on the right hand side is equivalent to $[J]\varphi$ by S5(J).

From the right to the left, $\vdash \langle\overline{J}\rangle[J]\varphi \to \langle\emptyset\rangle[J]\varphi$ by (Mon). ■

We now give a theorem of Ldm^G that lifts Xu's axiom (AIA_1) of Section 4 from individuals to coalitions, and that will be instrumental later in the proof of superadditivity in Theorem 23.

LEMMA 17. — $\vdash \langle\emptyset\rangle[J_0]\varphi_0 \wedge \langle\emptyset\rangle[J_1]\varphi_1 \to \langle\emptyset\rangle([J_0]\varphi_0 \wedge [J_1]\varphi_1)$ *for* $J_0 \cap J_1 = \emptyset$.

PROOF. — Suppose $J_0 \cap J_1 = \emptyset$. We establish the following deduction:

1) $\langle\emptyset\rangle[J_0]\varphi_0 \to \langle J_1\rangle\langle J_0\rangle[J_0]\varphi_0$ by Lemma 15
2) $\langle\emptyset\rangle[J_0]\varphi_0 \to \langle J_1\rangle[J_0]\varphi_0$ from 1 by S5($[J_0]$)
3) $\langle\emptyset\rangle[J_0]\varphi_0 \wedge [J_1]\varphi_1 \to \langle J_1\rangle[J_0]\varphi_0 \wedge [J_1][J_1]\varphi_1$ from 2 by S5($[J_1]$)
4) $\langle\emptyset\rangle[J_0]\varphi_0 \wedge [J_1]\varphi_1 \to \langle J_1\rangle([J_0]\varphi_0 \wedge [J_1]\varphi_1)$
from 3 by standard modal principles
5) $\langle\emptyset\rangle(\langle\emptyset\rangle[J_0]\varphi_0 \wedge [J_1]\varphi_1) \to \langle\emptyset\rangle\langle J_1\rangle([J_0]\varphi_0 \wedge [J_1]\varphi_1)$
from 4 by standard modal principles
6) $\langle\emptyset\rangle[J_0]\varphi_0 \wedge \langle\emptyset\rangle[J_1]\varphi_1 \to \langle\emptyset\rangle\langle J_1\rangle([J_0]\varphi_0 \wedge [J_1]\varphi_1)$ from 5 by S5($[\emptyset]$)
7) $\langle\emptyset\rangle[J_0]\varphi_0 \wedge \langle\emptyset\rangle[J_1]\varphi_1 \to \langle\emptyset\rangle([J_0]\varphi_0 \wedge [J_1]\varphi_1)$ from 6 by (Mon) and S5($[\emptyset]$)

■

REMARK 18. — Our logic Ldm^G contains several principles that also hold for product logics (Gabbay *et al.*, 2003). Indeed, for every $i, j \in AGT$ the principles of permutation $[i][j]\varphi \to [j][i]\varphi$ can be proved from Lemma 15, (Mon) and S5($[J]$) and standard modal principles; and the Church-Rosser confluence principle $\langle i\rangle[j]\varphi \to [j]\langle i\rangle\varphi$ can be proved from permutation and S5($[J]$). However, in general Ldm^G is still *weaker* than a product logic.

From a computational perspective it is a blessing that Ldm^G for more than two agents is not a product logic: it is well known that a product of three S5 modalities already yields a not finitely axiomatizable and undecidable logic. □

REMARK 19. — Note that there is an interesting technical link with the definition of common knowledge in so called 'interpreted systems'. An interpreted system is a set of n agents having S5 knowledge only of their own local state. The knowledge of these agents is thus *independent* in the sense that what an agent knows does not depend on the other agents' local states, and thus, the knowledge of other agents. We can make a straightforward connection between interpreted systems and Ldm^G by comparing *historical necessity* to *common knowledge* and *individual choice* to *standard knowledge*. In Ldm^G, the historical necessity modality $[\emptyset]$ is interpreted by the reflexive transitive closure of the accessibility relations of the choices for the individual agents and coalitions. This is because the role of historical necessity is to 'gather' all the worlds that can be the result of agentive processes. Likewise, the notion of common knowledge (C) is semantically defined as the reflexive transitive closure of the S5 classes of the knowledge operators for individual agents. But then, in interpreted systems we have that $C\varphi \leftrightarrow K_i K_j \varphi$ for arbitrary agents $i \neq j$. Note that this is analogous to the situation in Ldm^G. Although properties like these are known for interpreted systems, they are hard to find in published papers. Similar issues are discussed in (Lomuscio *et al.*, 2000). □

6. What groups *can* do: X-Ldm^G as an extension of CL

The Ldm^G-formula $\langle\emptyset\rangle[J]\varphi$ allows to express that it is historically possible that J ensures that φ, or in other words, that J can ensure that φ. But this is just the same as the reading of the CL-formula $\langle\!\langle J\rangle\!\rangle\mathbf{X}\varphi$ given in Section 2. It should therefore be possible to view Ldm^G as an extension of CL. Nevertheless, things are not as straightforward as they might seem. Consider the consistent CL-formula $\langle\!\langle\emptyset\rangle\!\rangle\mathbf{X}p \wedge \langle\!\langle\emptyset\rangle\!\rangle\mathbf{X}\langle\!\langle\emptyset\rangle\!\rangle\mathbf{X}\neg p$: p will be true at the next step, and false in two steps. Its Ldm^G-counterpart is $\langle\emptyset\rangle[\emptyset]p \wedge \langle\emptyset\rangle[\emptyset]\langle\emptyset\rangle[\emptyset]\neg p$, and is inconsistent in Ldm^G (because it collapses to $[\emptyset]p \wedge [\emptyset]\neg p$ in S5).

The reason is that although action and agency are intimately related to time, Ldm^G lacks temporal operators. We now extend Ldm^G by the simplest modal operator of time: the linear 'next' operator $\mathbf{X}$. We call the result $\mathbf{X}$-Ldm^G. We then establish that the resulting logic is an extension of CL by translating $\langle\!\langle J\rangle\!\rangle\mathbf{X}\varphi$ to $\langle\emptyset\rangle[J]\mathbf{X}\varphi$.

The language of $\mathbf{X}$-Ldm^G extends that of $\mathsf{Ldm}^G$by formulas of the form $\mathbf{X}\varphi$, whose intended meaning is 'next time φ'.

6.1. *Axiomatization*

$\mathbf{X}$-Ldm^G has the same axioms as Ldm^G, plus

K($\mathbf{X}$) $\quad \mathbf{X}(\varphi \to \psi) \to (\mathbf{X}\varphi \to \mathbf{X}\psi)$

alt$_1$($\mathbf{X}$) $\quad \neg\mathbf{X}\neg\varphi \to \mathbf{X}\varphi$

We assume the standard inference rules of modus ponens, and necessitation for $\mathbf{X}$ and $[\emptyset]$.

6.2. *Kripke semantics*

An **X**-Ldm^G-model is a tuple $\mathcal{M} = (W, R, F_X, \pi)$, where:

- $\mathcal{M} = (W, R, \pi)$ is an Ldm^G-model;
- $F_X : W \longrightarrow W$ is a partial function.

The truth conditions are as before, plus:

- $\mathcal{M}, w \models \mathbf{X}\varphi$ iff if F_X is defined at w then $\mathcal{M}, F_X(w) \models \varphi$.

6.3. *Properties of X-Ldm^G*

THEOREM 20. — ***X**-Ldm^G is determined by the class of **X**-Ldm^G-models.*

PROOF. — As for Ldm^G, soundness and completeness follow from Sahlqvist's Theorem. ■

REMARK 21. — In **X**-Ldm^G there is no interaction between time and action: the formula $[J]\mathbf{X}p \wedge \mathbf{X}[J]\neg p$ is satisfiable: for example for $J = \{i\}$, i might see to it that the door is closed at time point $t+1$, while at $t+1$ i is not responsible for the door being closed. We will discuss a principle of success preservation in Section 8. □

6.4. *Translating Coalition Logic to X-Ldm^G*

In order to obtain an exact matching with CL we have to add two further constraints on **X**-Ldm^G, which come in terms of the following axioms.

alt$_1$($[AGT]$)	$\langle AGT\rangle\varphi \rightarrow [AGT]\varphi$
D(**X**)	$\mathbf{X}\varphi \rightarrow \neg\mathbf{X}\neg\varphi$

Axiom alt$_1$($[AGT]$) says that if every agent in the 'grand coalition' chooses an action then the system behaves in a deterministic way. It follows from alt$_1$($[AGT]$) and S5($[J]$) that $[AGT]\varphi \leftrightarrow \varphi$. D(**X**) says that time has no end.

We call the resulting logic **X**-Ldm^G + alt$_1$($[AGT]$) + D(**X**).

As expected, a model for **X**-Ldm^G + alt$_1$($[AGT]$) + D(**X**) is a **X**-Ldm^G-model $\mathcal{M} = (W, R, F_X, \pi)$, where:

- $R_{AGT} = Id$;
- $F_X : W \longrightarrow W$ is a total function.

The constraint $R_{AGT} = Id$ means that the choices of the big coalition completely determine the outcome: there are no other agents whose choices would be relevant. It follows that if we want to allow for nondeterminism then we have to include an 'environment', 'nature', or 'god' agent in AGT. See (Broersen *et al.*, 2009) for a discussion.

Now we are ready to give a translation from Coalition Logic to $\mathbf{X}$-Ldm^G:

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

and homomorphic for the other connectives.

REMARK 22. — Note that as the equivalence $\langle\emptyset\rangle[J]\varphi \leftrightarrow \langle\overline{J}\rangle[J]\varphi$ is Ldm^G-valid (cf. Lemma 16), we might as well translate $\langle\!\langle J\rangle\!\rangle\mathbf{X}\varphi$ to $\langle\overline{J}\rangle[J]\mathbf{X}tr(\varphi)$, which reads '$\overline{J}$ allow J to see to it that φ. Actually one may consider that this renders more faithfully the intended reading of $\langle\!\langle J\rangle\!\rangle\mathbf{X}\varphi$ that 'J can ensure that φ *whatever* $\overline{J}$ *does*'. □

THEOREM 23. — *If φ is a theorem of* CL *then $tr(\varphi)$ is a theorem of $\mathbf{X}$-Ldm^G + alt$_1$($[AGT]$) + D($\mathbf{X}$).*

PROOF. — First, the translations of the CL axiom schemas are theorems of $\mathbf{X}$-Ldm^G + alt$_1$($[AGT]$) + D($\mathbf{X}$). The only non-trivial cases are axiom (S) for superadditivity, and axiom (N) for AGT-maximality. We start with the latter:

$tr(\neg\langle\!\langle\emptyset\rangle\!\rangle\mathbf{X}\neg\varphi \rightarrow \langle\!\langle AGT\rangle\!\rangle\mathbf{X}\varphi) = \neg\langle\emptyset\rangle[\emptyset]\mathbf{X}\neg tr(\varphi) \rightarrow \langle\emptyset\rangle[AGT]\mathbf{X}tr(\varphi)$. Since $[AGT]\psi \leftrightarrow \psi$ by Det($[AGT]$) and S5($[AGT]$), and $\langle\emptyset\rangle[\emptyset]\psi \leftrightarrow [\emptyset]\psi$ by S5($[\emptyset]$), the translation of (N) is equivalent to $\neg[\emptyset]\mathbf{X}\neg tr(\varphi) \rightarrow \langle\emptyset\rangle\mathbf{X}tr(\varphi)$. This is equivalent to $\langle\emptyset\rangle\neg\mathbf{X}\neg tr(\varphi) \rightarrow \langle\emptyset\rangle\mathbf{X}tr(\varphi)$ which is proved a theorem from applying $[\emptyset]$-necessitation to alt$_1$($\mathbf{X}$) and applying a variant of the K-axiom.

For the axiom (S) of superadditivity we have:
$tr(\langle\!\langle J_1\rangle\!\rangle\mathbf{X}\varphi \wedge \langle\!\langle J_2\rangle\!\rangle\mathbf{X}\psi \rightarrow \langle\!\langle J_1 \cup J_2\rangle\!\rangle\mathbf{X}(\varphi \wedge \psi)) =$
$\langle\emptyset\rangle[J_1]\mathbf{X}tr(\varphi) \wedge \langle\emptyset\rangle[J_2]\mathbf{X}tr(\psi) \rightarrow \langle\emptyset\rangle[J_1 \cup J_2]\mathbf{X}(tr(\varphi) \wedge tr(\psi))$
We establish the following deduction:

1) $\langle\emptyset\rangle[J_1]\mathbf{X}tr(\varphi) \wedge \langle\emptyset\rangle[J_2]\mathbf{X}tr(\psi) \rightarrow \langle\emptyset\rangle([J_1]\mathbf{X}tr(\varphi) \wedge [J_2]\mathbf{X}tr(\psi))$
by Lemma 17
2) $[J_1]\mathbf{X}tr(\varphi) \wedge [J_2]\mathbf{X}tr(\psi) \rightarrow [J_1 \cup J_2]\mathbf{X}tr(\varphi) \wedge [J_1 \cup J_2]\mathbf{X}tr(\psi)$ by (Mon)
3) $\langle\emptyset\rangle([J_1]\mathbf{X}tr(\varphi) \wedge [J_2]\mathbf{X}tr(\psi)) \rightarrow \langle\emptyset\rangle([J_1 \cup J_2](\mathbf{X}tr(\varphi) \wedge \mathbf{X}tr(\psi))$
from line 2 by standard modal principles
4) $\langle\emptyset\rangle[J_1]\mathbf{X}tr(\varphi) \wedge \langle\emptyset\rangle[J_2]\mathbf{X}tr(\psi) \rightarrow \langle\emptyset\rangle[J_1 \cup J_2]\mathbf{X}(tr(\varphi) \wedge tr(\psi))$
from lines 1 and 3 by standard modal principles for $\mathbf{X}$.

Second, clearly the translation of modus ponens preserves validity. To prove that the translation of CL's (RE) preserves validity suppose $tr(\varphi \leftrightarrow \psi) = tr(\varphi) \leftrightarrow tr(\psi)$ is a theorem of $\mathbf{X}$-Ldm^G. We have to prove that $tr(\langle\!\langle J\rangle\!\rangle\mathbf{X}\varphi \leftrightarrow \langle\!\langle J\rangle\!\rangle\mathbf{X}\psi) = \langle\emptyset\rangle[J]\mathbf{X}tr(\varphi) \leftrightarrow \langle\emptyset\rangle[J]\mathbf{X}tr(\psi)$ is a theorem of $\mathbf{X}$-Ldm^G. This follows from the theoremhood of $tr(\varphi) \rightarrow tr(\psi)$ by standard modal principles. ■

Now we can turn to the proof of satisfiability preservation of the translation and the model construction.

THEOREM 24. — *If φ is* CL*-satisfiable then $tr(\varphi)$ is satisfiable in logic **X**-Ldm^G + alt$_1$($[AGT]$) + D(**X**).*

PROOF. — If φ is CL-satisfiable, then it is satisfiable in a game model. Suppose $M_G = (S, \gamma, v)$ is a game model and $s_0 \in S$ is a state such that $M_G, s_0 \models \varphi$. Let W be the set of all sequences $s_0\sigma_0 \dots s_K\sigma_K$ such that $K \geq 0$ and for $0 \leq k < K$, $s_k \in S$, and $\sigma_k \in \Sigma_{s_k,AGT}$ is such that $o_{s_k}(\sigma_k) = s_{k+1}$ (i.e. σ_k is the choice profile at s_k that leads to s_{k+1}). Define relations $R_J \subseteq W \times W$ as the set of couples

$$(s_0\sigma_0 \dots s_K\sigma_K, s_0\sigma_0 \dots s_K\tau_K)$$

such that $\sigma_K^J = \tau_K^J$, where σ_K^J and τ_K^J are the projections of σ_K and τ_K on the set $\Sigma_{s_K,J}$ of choices for coalition J in the strategic game $\gamma(s_K)$. Define the function $F_X : W \longrightarrow W$ as:

$$F_X(s_0\sigma_0 \dots s_K\sigma_K) = s_0\sigma_0 \dots s_K\sigma_K s_{K+1}\sigma_{K+1}$$

for $s_{K+1} = o_{s_K}(\sigma_K)$ and for some $\sigma_{K+1} \in \Sigma_{s_{K+1},AGT}$. Define $\pi : W \longrightarrow 2^{PRP}$ as:

$$(s_0\sigma_0 \dots s_K\sigma_K) \in \pi(p) \text{ iff } s_K \in v(p)$$

We first prove that (W, R, F_X, π) is a model of **X**-Ldm^G + alt$_1$($[AGT]$) + D(**X**), and then prove that $(W, R, F_X, \pi), s_0\sigma_0 \models tr(\varphi)$.

First, R_J is clearly an equivalence relation, for every $J \subseteq AGT$.

Second, the condition of coalition monotony holds: suppose $J \subseteq J'$ and

$$(s_0\sigma_0 \dots s_K\sigma_K, s_0\sigma_0 \dots s_K\tau_K) \in R_{J'}.$$

Hence $\sigma_K^{J'} = \tau_K^{J'}$. As σ_K^J and τ_K^J are respectively subvectors of $\sigma_K^{J'}$ and $\tau_K^{J'}$, we must also have $\sigma_K^J = \tau_K^J$.

Third, we prove that $R_\emptyset \subseteq R_J \circ R_{\overline{J}}$ for every J. Suppose

$$(s_0\sigma_0 \dots s_K\sigma_K, s_0\sigma_0 \dots s_K\tau_K) \in R_\emptyset.$$

We have $\sigma_K = \sigma_K^J \cdot \sigma_K^{\overline{J}}$ and $\tau_K = \tau_K^J \cdot \tau_K^{\overline{J}}$, for some $\sigma_K^J \in \Sigma_{s_K,J}$, etc. Therefore

$$(s_0\sigma_0 \dots s_K(\sigma_K^J \cdot \sigma_K^{\overline{J}}), s_0\sigma_0 \dots s_K(\sigma_K^J \cdot \tau_K^{\overline{J}})) \in R_J,$$

and

$$(s_0\sigma_0 \dots s_K(\sigma_K^J \cdot \tau_K^{\overline{J}}), s_0\sigma_0 \dots s_K(\tau_K^J \cdot \tau_K^{\overline{J}})) \in R_{\overline{J}}.$$

Fourth, $R_{AGT} = Id$ because $\sigma_K^{AGT} = \sigma_K = \tau_K = \tau_K^{AGT}$.

Fifth, F_X is a total function because $\Sigma_{s_K,AGT}$ is nonempty for every K.

It remains to establish that for all φ and for all $s_0\sigma_0 \dots s_K\sigma_K \in W$ we have

$$M_G, s_K \models \varphi \text{ iff } (W, R, F_X, \pi), s_0\sigma_0 \dots s_K\sigma_K \models tr(\varphi).$$

We prove this by induction on the structure of φ.

For the base case we have: $M_G, s_K \models p$ iff $s_K \in v(p)$ iff $s_0\sigma_0 \dots s_K\sigma_K \in \pi(p)$, for all σ_K. The latter means that $(W, R, F_X, \pi), s_0\sigma_0 \dots s_K\sigma_K \models [\emptyset]p$.

Negation and disjunction are straightforward. We now prove the case where the main connective is $\langle\!\langle J \rangle\!\rangle$**X**.

We have $M_G, s_K \models \langle\!\langle J \rangle\!\rangle \mathbf{X}\varphi$ iff there is a $\sigma_K^J \in \Sigma_{s_K,J}$ such that for all $\sigma_K^{\overline{J}} \in \Sigma_{s_K,\overline{J}}$ we have $G_M, o_{s_K}(\sigma_K^J \cdot \sigma_K^{\overline{J}}) \models \varphi$; Hence, by induction hypothesis

$$(W, R, F_X, \pi), s_0\sigma_0 \ldots s_K\sigma_K o_{s_K}(\sigma_K)\sigma_{K+1} \models tr(\varphi),$$

where $\sigma_K = (\sigma_K^J \cdot \sigma_K^{\overline{J}})$; for the left-to-right direction, $\sigma_{K+1} \in \Sigma_{o_{s_K}(\sigma_K),AGT}$ is specifically such that $F_X(s_0\sigma_0 \ldots s_K\sigma_K) = s_0\sigma_0 \ldots s_K\sigma_K s_{K+1}\sigma_{K+1}$.
This is equivalent to: there is a $\sigma_K^J \in \Sigma_{s_K,J}$ such that

$$(W, R, F_X, \pi), s_0\sigma_0 \ldots s_K\sigma_K \models \mathbf{X}tr(\varphi).$$

for all $\sigma_K^{\overline{J}} \in \Sigma_{s_K,\overline{J}}$.
This is equivalent to: there is a $s_0\sigma_0 \ldots s_K\tau_K \in W$ such that for all $s_0\sigma_0 \ldots s_K\sigma_K$, if $(s_0\sigma_0 \ldots s_K\tau_K, s_0\sigma_0 \ldots s_K\sigma_K) \in R_J$ then

$$(W, R, F_X, \pi), s_0\sigma_0 \ldots s_K\sigma_K \models \mathbf{X}tr(\varphi).$$

Finally, this is the same as:

$$(W, R, F_X, \pi), s_0\sigma_0 \ldots s_K\sigma_K \models \langle\emptyset\rangle[J]\mathbf{X}tr(\varphi).$$

■

COROLLARY 25. — *The formula φ is a theorem of* CL *iff $tr(\varphi)$ is a theorem of* **X**-*Ldm*^{G} + *alt*$_1$*(*$[AGT]$*) + D(***X***).*

PROOF. — The left-to-right direction is Theorem 23. The right-to-left direction follows from Pauly's completeness result for Coalition Logic and Theorem 24. ■

7. What groups *know* they (can) do: E-X-Ldm^G, alias conformant X-Ldm^G

In this section we extend our framework with an S5 knowledge operator. This enables us to express that an agent sees to something although it is uncertain about the present state or the action being taken. The problems with modelling uniformity of strategies already arise with one step choices. We therefore show how, as an extension of **X**-Ldm^G, we can easily obtain a complete system E-**X**-Ldm^G whose semantics distinguishes between uniform and non-uniform strategies.

In the planning community uniform strategies are called *conformant* (Goldman *et al.*, 1996); they ensure a property ('the goal') in spite of uncertainty about the present state. The logic presented here enables us to express this as $K_i[\{i\}]\varphi$ for "agent i knows that it sees to it that φ, without necessarily knowing the present state". To be in accordance with the established terminology in the planning community we may call this combination of the knowledge operator and the agency operator the 'conformant **X**-Ldm^G'.

In (Herzig *et al.*, 2006) we sketched how the problem can be solved in an extension of the *stit* framework. In (Broersen *et al.*, 2006a) we considered a stit-extension of ATL that we called ATL-STIT. The logic system we present here does not have the restricted syntax of the first presented proposal in (Herzig *et al.*, 2006) and, in addition, has a complete and straightforward axiomatization.

E-**X**-Ldm^G has the following BNF:

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

where p ranges over PRP, i over AGT, and J over 2^{AGT} (so, we do not consider group knowledge).

7.1. *Axiomatization*

The axiomatization of E-$\mathbf{X}$-Ldm^G is obtained by adding to $\mathbf{X}$-Ldm^G the principles of the standard epistemic logic S5 for every individual agent i.

7.2. *Kripke semantics*

Models of E-$\mathbf{X}$-Ldm^G are tuples $\mathcal{M} = (W, R, F_X, \sim, \pi)$ where:

- (W, R, F_X, π) is a model of $\mathbf{X}$-Ldm^G;
- $\sim$ is a collection of equivalence relations $\sim_i$ (one for every agent $i \in AGT$).

THEOREM 26. — *E-$\mathbf{X}$-Ldm^G is determined by the class of models of E-$\mathbf{X}$-Ldm^G.*

PROOF. — As for the preceding logics, soundness and completeness follow from Sahlqvist's Theorem. ■

7.3. *Uniform choice in E-X-Ldm^G*

To explain how logic E-$\mathbf{X}$-Ldm^G solves the problem of uniform choice, we consider two scenarios.

EXAMPLE 27. — Ann is in a room. She is blind and cannot distinguish a world where the light is off from a world where the light is on. The light in the room is controlled by a button that activates a timer (as often the case in public buildings). When the button is pushed the light bulb will shine for a determinate time. When the light is on, there is no way to switch it off. Ann can also do nothing (skip). In the actual situation the light is off and Ann is pushing the button. □

Figure 1 models our example. Plain lines correspond to elements of W and dashed lines stand for $\sim_{Ann}$ accessibility. The worlds of the semantics of $\mathbf{X}$-Ldm^G and E-$\mathbf{X}$-Ldm^G have to be seen as state-action pairs. The states are positions before and after execution of an action. In our model there are 6 of these positions, and they result in 8 E-$\mathbf{X}$-Ldm^G-worlds. We thus have the following E-$\mathbf{X}$-Ldm^G-model $\mathcal{M}_1 = \langle W, R, F_X, \sim, \pi\rangle$:

- $W = \{(1, \text{push}), (1, \text{skip}), (2, \text{skip}), (2, \text{push})\} \cup \{(\text{k}, \text{skip}) \mid k \in \{3, 4, 5, 6\}\}$
- $R_\emptyset$ is the smallest equivalence relation containing the set $\{\langle(1, \text{push}), (1, \text{skip})\rangle, \langle(2, \text{skip}), (2, \text{push})\rangle\} \cup \{\langle(\text{k}, \text{skip}) \mid k \in \{3, 4, 5, 6\}\}$
- $R_{Ann} = \{\langle w, w\rangle \mid w \in W\}$

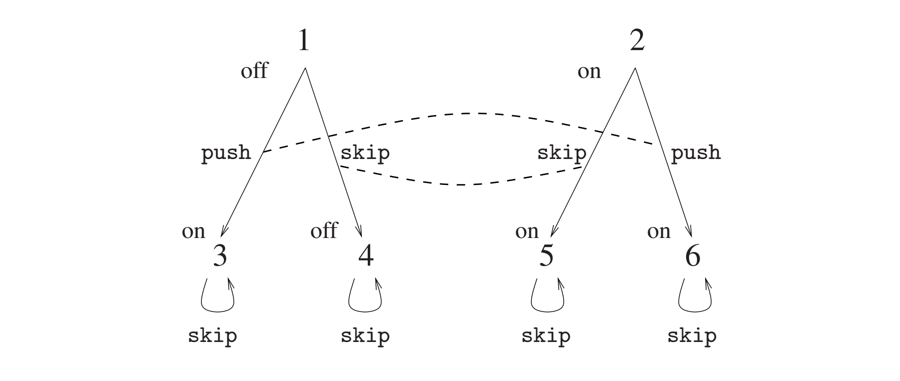

Figure 1. *The E-**X**-Ldm^G-model $\mathcal{M}_1$ for Example 27*

– F_X is defined by:
  $F_X((1, \text{push})) = (3, \text{skip}),\ F_X((1, \text{skip})) = (4, \text{skip}),$
  $F_X((2, \text{skip})) = (5, \text{skip}),\ F_X((2, \text{push})) = (6, \text{skip}),$
  $F_X((\text{k}, \text{skip})) = (\text{k}, \text{skip})$, for $k \in \{3, 4, 5, 6\}$
– $\sim_{Ann}$ is the smallest equivalence relation containing the set
$$\{\langle(1, \text{push}), (2, \text{push})\rangle, \langle(1, \text{skip}), (2, \text{skip})\rangle\}$$
– π is defined by:
  $\pi((1, \text{push})) = \pi((1, \text{skip})) = \text{‘off’}$
  $\pi((2, \text{push})) = \pi((2, \text{skip})) = \text{‘on’},$
  $\pi((3, \text{skip})) = \pi((5, \text{skip})) = \pi((6, \text{skip})) = \text{‘on’},$
  $\pi((4, \text{skip})) = \text{‘off’}$

It is not difficult to check that $\mathcal{M}_1$ is a genuine E-**X**-Ldm^G-model, satisfying also all the constraints we defined for the **X**-Ldm^G-submodels. The reader may have noticed that the model adds detail to the example. In particular, Ann is given the choice between pushing and skipping only once, and "determinate time" is interpreted as *forever*. Of course, the model is a very simple one, with only one agent in the system: $AGT = \{Ann\}$. Ann's actions thus coincide with system actions, and all her choices are deterministic.

The four basic properties we consider are:

– $\varphi_1 = \langle\emptyset\rangle[\{Ann\}]\mathbf{X}\,\text{on}$ ("*One of Ann's choices ensures the light will be on*")
– $\varphi_2 = K_{Ann}\langle\emptyset\rangle[\{Ann\}]\mathbf{X}\,\text{on}$
("*Ann knows one of her choices ensures the light will be on*")
– $\varphi_3 = \langle\emptyset\rangle K_{Ann}[\{Ann\}]\mathbf{X}\,\text{on}$
("*Ann knows she has the power to ensure the light is on*")
– $\varphi_4 = K_{Ann}[\{Ann\}]\mathbf{X}\,\text{on}$ ("*Ann conformantly sees to it that the light is on*")

It is easy to check that in $\mathcal{M}_1$ the first three formulas are true in the first four possible **X**-Ldm^G worlds: $\mathcal{M}_1, w \models \varphi_1 \wedge \varphi_2 \wedge \varphi_3$ for all w in the set

$\{(1,\text{push}),(1,\text{skip}),(2,\text{skip}),(2,\text{push})\}$.

In particular, in the actual world $(1,\text{push})$ the third property holds, saying that Ann has a uniform strategy to ensure the light is on. In the actual world also the fourth property holds ($\mathcal{M}_1,(1,\text{push}) \models \varphi_4$), while in the two worlds where Ann skips, it does not ($\mathcal{M}_1,(1,\text{skip}) \not\models \varphi_4$ and $\mathcal{M}_1,(2,\text{push}) \not\models \varphi_4$).

EXAMPLE 28. — Ann is in a room. She is blind and cannot distinguish a world where the light is off from a world where the light is on. The light in the room is controlled by a switch. In her repertoire of actions, Ann can toggle t or remain passive (skip, s), which correspond to switching the state of the light and maintaining the state of the light, respectively. In the actual situation the light is off and Ann toggles. □

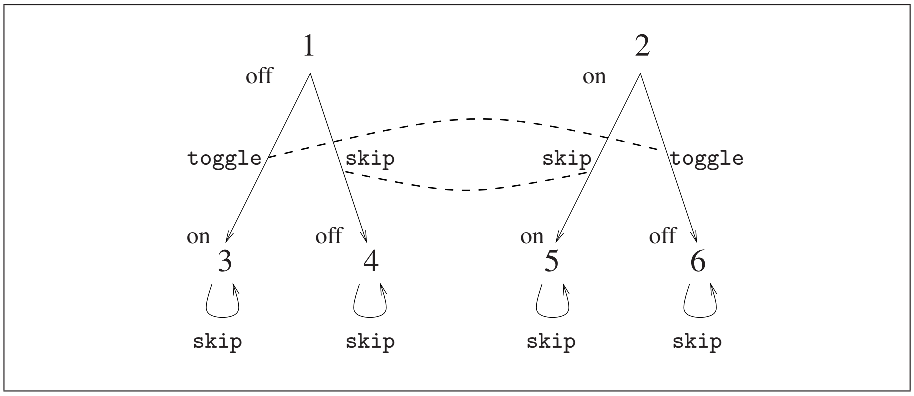

Figure 2. *The E-**X**-Ldm^G-model $\mathcal{M}_2$ for Example 28*

This example is modeled by the E-**X**-Ldm^G-model $\mathcal{M}_2$ that is depicted in Figure 2: $\mathcal{M}_2 = \langle W, R, F_X, \sim, \pi\rangle$, where

- $W = \{(1,t),(1,s),(2,s),(2,t),(3,s),(4,s),(5,s),(6,s)\}$
- $R_\emptyset$ is the smallest equivalence relation containing the set $\{\langle(1,t),(1,s)\rangle,\langle(2,s),(2,t)\rangle,\langle(3,s),(3,s)\rangle,$ $\langle(4,s),(4,s)\rangle,\langle(5,s),(5,s)\rangle,\langle(6,s),(6,s)\rangle\}$
- $R_{Ann} = \{\langle w,w\rangle \mid w \in W\}$
- F_X is defined by:
  $F_X((1,t)) = (3,s),\ F_X((1,s)) = (4,s),$
  $F_X((2,s)) = (5,s),\ F_X((2,t)) = (6,s),$
  $F_X((\text{k},s)) = (\text{k},s),$ for $k \in \{3,4,5,6\}$
- $\sim_{Ann}$ is the smallest equivalence relation containing the set
  $$\{\langle(1,t),(2,t)\rangle,\langle(1,s),(2,s)\rangle\}$$
- π is defined by:
  $\pi((2,t)) = \pi((2,s)) = \pi((3,s)) = \pi((5,s)) =$ 'on',
  $\pi((1,t)) = \pi((1,s)) = \pi((4,s)) = \pi((6,s)) =$ 'off'

Now, in the actual world where the light is off and Ann toggles, the light will actually be on, so the formula $\mathbf{X}on$ holds. Yet, Ann does not conformantly see to it that the light is on, since she does not know that the light is off at the present moment. So, the fourth of the above properties does not hold: $\mathcal{M}_2, (1, t) \not\models \varphi_4$. Also, she does not have a uniform strategy, and indeed the third of the above properties does not hold either: $\mathcal{M}_2, (1, t) \not\models \varphi_3$. The first and the second property do hold in the actual world, since in each state Ann indeed has an action that ensures the light is on and she knows that. But her problem is that the decision to take depends on the state she is in, which is something she does not know: $\mathcal{M}_2, w \models \varphi_1 \wedge \varphi_2$ for all $w \in \{(1, t), (1, s), (2, s), (2, t)\}$.

7.4. *Comparison with* ***ATEL***

Let us compare our approach with the situation in ATEL and CL. For representing uncertainty in ATEL a family of equivalence relations among *states* (one for each agent) is assumed, interpreting a standard normal S5 operator K_i in the language. Since in *stit* uncertainty relations can range over *world-history* pairs[6] (see section 8.1, below), the semantics of our knowledge operators is more fine-grained.

We are going to argue now that the known approaches to the problem of uniform strategies in the literature are unlikely to succeed. Note first that in Example 2 above we might have given different names to the actions. And there is no reason why this renaming should be uniform. In particular, the *left* toggle action can be called '*put the light on*' and the *right* toggle action '*put the light off*'. Obviously, non-uniform renaming of actions should not influence Ann's basic capabilities or her knowledge concerning her capabilities. Our theory satisfies this principle, since changing the names of the actions in the way described, does not in any way change the evaluation of E-**X**-Ldm^G-formulas. In particular, Ann still does not have a uniform strategy: using the new terminology provided by the new action names she now '*cannot distinguish between putting the light on when it is off and putting the light off when it is on*'. However, none of the ATEL-based approaches in the literature satisfies the principle. In these variants and extension of ATEL (see e.g. (Schobbens, 2004)) the following condition is imposed on the models: if one *state* is indistinguishable from another, then any action name appearing for a choice in the first state also appears as an action name for a choice in the second state. It is clear right away that under this restriction, a non-uniform renaming of actions may result in uncertainty relations being eliminated, and thus in a gain in knowledge. In particular, in the renamed version of example 2 above, Ann would always be able to distinguish the two states, and there would be no uncertainty left at all, which directly contradicts the requirement having to express that Ann does *not* know a uniform strategy in this situation.

6. Or 'indexes', as we called them in previous sections.

8. The relation between X-Ldm^G and STIT-models

The semantics of the *stit* operator was extensively studied by Belnap *et al.* (Belnap *et al.*, 2001). It consists of branching-time structures (BT) augmented by the set of agents and a choice function (AC).

We focus here on *discrete* BT+AC models. Such models were introduced in (Broersen *et al.*, 2006b) in order to clarify the relation between STIT on the one hand, and CL and ATL on the other. We show that while discrete BT+AC models validate all the principles of our **X**-Ldm^G of Section 6, there are nevertheless BT+AC validities that are not theorems of **X**-Ldm^G, and explore what is missing.

8.1. *Semantics: discrete BT+AC models*

A *BT structure* is of the form $\langle W, < \rangle$, where W is a nonempty set of moments, and $<$ is a tree-like strict ordering of these moments: for any w_1, w_2 and w_3 in W, if $w_1 < w_3$ and $w_2 < w_3$, then either $w_1 = w_2$ or $w_1 < w_2$ or $w_2 < w_1$. A BT structure is *discrete* if for every $w, u \in W$ such that $w < u$, either there is no $v \in W$ such that $w < v < u$, or there is a $v \in W$ such that $w < v < u$ and there is no $v' \in W$ such that $w < v' < v$.

A maximal set of linearly ordered moments from W is a *history*. When $w \in h$ we say that moment w is *on* the history h. As $<$ is discrete, for every history h and $w \in h$ there is at most one moment $w' \in h$ such that $w < w'$.

We write $Hist$ for the set of all histories. The set $H_w = \{h \mid h \in Hist, w \in h\}$ denotes the set of histories passing through w. An *index* is a pair w/h, consisting of a moment w and a history h from H_w (i.e., a history and a moment in that history).

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

– $\langle W, < \rangle$ is a discrete BT structure;
– $Choice : AGT \times W \rightarrow 2^{2^{Hist}}$ is a function mapping each agent i and each moment w into a partition $Choice_i^w$ of H_w, such that
  - $Choice_i^w \neq \emptyset$;
  - $Q \neq \emptyset$ for every $Q \in Choice_i^w$;
  - for all w and all mappings $s_w : AGT \longrightarrow 2^{H_w}$ such that $s_w(i) \in Choice_i^w$, we have $\bigcap_{i \in AGT} s_w(i) \neq \emptyset$;
– V is valuation function $V : PRP \rightarrow 2^{W \times Hist}$.

The equivalence classes belonging to $Choice_i^w$ can be thought of as possible choices that are available to agent i at w. Given a history $h \in H_w$, $Choice_i^w(h)$ represents the particular choice from $Choice_i^w$ containing h, or in other words, the particular action performed by i at the index w/h.

We say that two histories h_1 and h_2 are *undivided* at w iff there is a w' such that $w < w'$, and $w' \in h_1 \cap h_2$. An important constraint of $BT + AC$ structures is the principle of *no choice between undivided histories*. It forces that if two histories h_1 and h_2 are undivided at w, then $h_2 \in Choice_i^w(h_1)$ for every agent i.

The constraint of nonempty intersection of all possible simultaneous choices of agents (or: strategy profile) is the postulate of *independence of agents*.

This is generalized to groups just in the same way as for CL's game semantics in Section 2.2.

A formula is evaluated with respect to a model and an index:

$$\begin{array}{lll} \mathcal{M}, w/h \models p & \text{iff} & w/h \in V(p), p \in PRP \\ \mathcal{M}, w/h \models \Box\varphi & \text{iff} & \mathcal{M}, w/h' \models \varphi, \forall h' \in H_w \\ \mathcal{M}, w/h \models [i]\varphi & \text{iff} & \mathcal{M}, w/h' \models \varphi, \forall h' \in Choice_i^w(h) \\ \mathcal{M}, w/h \models \mathbf{X}\varphi & \text{iff} & \mathcal{M}, w'/h \models \varphi, \text{ where } w' \text{ is the successor of } w \text{ on } h \end{array}$$

and as usual for the Boolean connectives.

Hence historical necessity (or inevitability) at a moment w in a history is truth in all histories passing through w. According to Chellas, an agent i sees to it that φ in a moment-history pair w/h if φ holds on all histories that agree with i's current choice.

Validity in discrete BT+AC models is defined as truth at every moment-history pair of every discrete BT+AC-model. A formula φ is satisfiable in discrete BT+AC models iff $\neg\varphi$ is not valid in BT+AC models.

8.2. *Incompleteness of X-Ldm^G w.r.t. discrete BT+AC models*

Discrete BT+AC models clearly validate all the principles of $\mathbf{X}$-Ldm^G. Nevertheless, there are BT+AC validities that are not theorems of $\mathbf{X}$-Ldm^G. We here focus on two such principles, the success preservation axiom:

(SuccPresrv) $\quad [J]\mathbf{X}\varphi \to \mathbf{X}[\emptyset]\varphi$

and the 'coalition building axiom':

(CB) $\quad [J_1][J_2]\varphi \to [J_1 \cap J_2]\varphi.$

Both of them are valid in BT+AC models.

Axiom (SuccPresrv) reflects that what an agent does cannot be 'undone' in the sense that any action changes the world irrevocably. (Of course we can think of actions that undo the effects of earlier actions, but that is not the same.) On $\mathbf{X}$-Ldm^G-models, (SuccPresrv) corresponds to the constraint

– $F_X \circ R_\emptyset \subseteq R_J \circ F_X$ (success preservation + no choice between undivided histories) where the partial function F_X is viewed as a relation. Clearly, (SuccPresrv) is not **X**-Ldm^G-valid.

Axiom (CB) is a nice generalization of (Indep) to non-disjoint groups. Roughly speaking, (CB) says that if a group J_1 influences another group J_2 then this is due to J_1's members that are also in J_2. On **X**-Ldm^G-models, (CB) corresponds to

– $R_{J_1 \cap J_2} \subseteq R_{J_1} \circ R_{J_2}$

The following model shows that (CB) is not Ldm^G-valid (Schwarzentruber, 2007, Theorem 16).

EXAMPLE 29. — Let $AGT = \{1, 2, 3\}$, and let $\mathcal{M} = (W, R, \pi)$, be such that $W = \{w, w'\}$, $\pi(w) = \{p\}$ and $\pi(w) = \emptyset$, and

– $R_\emptyset = R_{\{1\}} = R_{\{2\}} = R_{\{3\}} = W \times W$
– $R_{\{1,2\}} = R_{\{1,3\}} = R_{\{2,3\}} = \{(w, w), (w', w')\}$

M is a Ldm^G-model; in particular the constraints $R_{J'} \subseteq R_J$ if $J \subseteq J'$ and $R_\emptyset = R_J \circ R_{\overline{J}}$ are satisfied. But M does not satisfy the above constraint $R_{\{1\}} \subseteq R_{\{1,2\}} \circ R_{\{1,3\}}$. Therefore (CB) is not true in M: $M, w \models [\{1, 2\}][\{1, 3\}]p$, but $M, w \not\models [\{1\}]p$. □

Note that the (CB) axiom can be strengthened to an equivalence

$$[J_1 \cap J_2]\varphi \leftrightarrow [J_1][J_2]\varphi$$

due to axiom (Mon) of **X**-Ldm^G.

8.3. *Non-axiomatizability and undecidability of BT+AC validities*

As we have seen, the axioms (SuccPresrv) and (CB) are valid in BT+AC-models. Thus if we want to axiomatize the latter we have to add these axioms to **X**-Ldm^G. It has been shown in (Herzig *et al.*, 2008, Theorem 23) that this is not enough when there are 3 or more agents: if the set of BT+AC validities was finitely axiomatizable for $n \geq 3$ then $S5^n$ would be finitely axiomatizable, and the latter was proved to be impossible (Gabbay *et al.*, 2003, Theorem 8.2).[7]

THEOREM 30 (Herzig *et al.*, 2008). — *There is no finite axiomatization of BT+AC validities if there are at least 3 agents.*

Moreover, due to undecidability of $S5^n$ for $n \geq 3$ (Venema, 1998, Theorem 8.6), satisfiability in BT+AC models is undecidable (Herzig *et al.*, 2008, Theorem 22):

THEOREM 31 (Herzig *et al.*, 2008). — *The problem of satisfiability in BT+AC models is undecidable if there are at least 3 agents.*

7. A logic is called finitely axiomatizable if there is a finite set of formula schemas from whose instances every theorem is obtained by necessitation and modus ponens.

The main obstacle on the way to a complete axiomatization is the following property of BT+AC-models (as well as by the Alternating Transition Systems ATS of ATL):

– $R_{J_1 \cup J_2} = R_{J_1} \cap R_{J_2}$

This constraint is stronger than that for (CB): it can be shown that the latter entails the former, but not the other way round. Basically, it says that the action repertoire of a group is completely determined by the respective repertoires of its members. While such a constraint can certainly be defended in simple cases of group actions, it may be argued that it is not necessarily so in more complex social situations, where groups may have actions at their disposal that are proper to them, and cannot be attached to individuals.

Note that the fact that BT+AC models satisfy the strong constraint $R_{J_1 \cup J_2} = R_1 \cap R_2$ does not imply that it is *false* that our axiomatization of $\mathbf{X}$-Ldm^G plus the axioms of success preservation and coalition building are complete with respect to BT+AC semantics. While the strong constraint $R_{J_1 \cup J_2} = R_{J_1} \cap R_{J_2}$ is clearly not modally expressible (since intersection is not modally expressible), we might still get completeness, see the literature on Boolean modal logic (Passy *et al.*, 1991).

9. Concluding remarks

We have some brief concluding remarks. The establishment of complete axiomatizations for $\mathbf{X}$-Ldm^G and E-$\mathbf{X}$-Ldm^G opens up interesting perspectives on the use of (semi)-automatic theorem provers for reasoning about properties of games. Such theorem provers could then also be used for conformant planning, through the established link between planning and satisfiability checking (Kautz *et al.*, 1992).

A natural investigation concerns the introduction of group knowledge in the present picture. In particular the integration of *common knowledge* is a worth challenge: some authors, for example Aumann, would say that our system is obsolete without it. It is straightforward to define common knowledge in $\mathbf{X}$-Ldm^G. However, completeness of the resulting logic does not follow immediately, as with standard epistemic logic.

Last but not least, a clear objective is to extend the axiomatizations we gave to the setting with extensive form games. We already studied the semantics for this extension in (Broersen *et al.*, 2006a). The most notable feature of the generalization of the semantics to extensive form games is that evaluation should be defined with respect to state-*strategy* pairs.

Acknowledgements

We would like to thank Philippe Balbiani, Olivier Gasquet and François Schwarzentruber for useful discussions concerning parts of the paper. Thanks are also due to the reviewer of the JANCL.

Andreas Herzig's work is supported by the French Agence Nationale de la Recherche in the framework of the ForTrust[8] project (ANR-06-SETI-006). Nicolas Troquard is supported by the EPSRC grant EP/E061397/1 *Logic for Automated Mechanism Design and Analysis (LAMDA)*.

10. References

Abdou J., Keiding K., *Effectivity Functions in Social Choice Theory*, Kluwer Academic, Dordrecht, 1991.

Alur R., Henzinger T., Kupferman O., "Alternating-time temporal logic", *Journal of the ACM*, vol. 49, pp. 672–713, 2002.

Balbiani P., Gasquet O., Herzig A., Schwarzentruber F., Troquard N., "Coalition games over Kripke semantics", *in* C. Dégremont, L. Keiff, H. Rückert (eds), *Dialogues, Logics and Other Strange Things – Essays in Honour of Shahid Rahman*, College Publications, pp. 11–32, 2008a.

Balbiani P., Herzig A., Troquard N., "Alternative axiomatics and complexity of deliberative STIT theories", *Journal of Philosophical Logic*, vol. 37, num. 4, pp. 387–406, 2008b.

Belnap N., Perloff M., "Seeing to it that: a canonical form for agentives", *Theoria*, vol. 54, pp. 175–199, 1988.

Belnap N., Perloff M., Xu M., *Facing the Future: Agents and Choices in our Indeterminist World*, Oxford, 2001.

Blackburn P., de Rijke M., Venema Y., *Modal Logic*, Cambridge University Press, 2001.

Broersen J., Herzig A., Troquard N., "A STIT-extension of ATL", *Tenth European Conference on Logics in Artificial Intelligence (JELIA'06), Liverpool, UK*, vol. 4160 of *Lecture Notes in Artificial Intelligence*, Springer, pp. 69–81, 2006a.

Broersen J., Herzig A., Troquard N., "Embedding Alternating-time Temporal Logic in strategic STIT logic of agency", *Journal of Logic and Computation*, vol. 16, num. 5, pp. 559-578, 2006b.

Broersen J., Herzig A., Troquard N., "Normal Coalition Logic and its conformant extension", *in* D. Samet (ed.), *Proceedings of the Eleventh Conference Theoretical Aspects of Rationality and Knowledge (TARK)*, Presses universitaires de Louvain, Brussels, Belgium, pp. 42–51, 2007.

Broersen J., Mastop R., Meyer J.-J., Turrini P., "Determining the Environment: A Modal Logic for Closed Interaction", *Synthese, special section of Knowledge, Rationality and Action*, vol. 169, num. 2, pp. 351-369, 2009.

Gabbay D., Kurucz A., Wolter F., Zakharyaschev M., *Many-Dimensional Modal Logics: Theory and Applications*, num. 148 in *Studies in Logic and the Foundations of Mathematics*, Elsevier, North-Holland, 2003.

8. http://www.irit.fr/ForTrust/

Goldman R., Boddy M., “Expressive Planning and Explicit Knowledge”, *Proceedings of the 3rd International Conference on Artificial Intelligence Planning Systems (AIPS-96)*, AAAI Press, pp. 110-117, 1996.

Goranko V., “Coalition games and alternating temporal logics”, *TARK '01: Proceedings of the 8th conference on Theoretical aspects of rationality and knowledge*, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, pp. 259–272, 2001.

Goranko V., van Drimmelen G., “Decidability and Complete Axiomatization of the Alternating-time Temporal Logic”, *Theoretical Computer Science*, vol. 353, num. 1-3, pp. 93–117, 2006.

Herzig A., Schwarzentruber F., “Properties of logics of individual and group agency”, *in* C. Areces, R. Goldblatt (eds), *Advances in Modal Logic (AiML)*, College Publications, Nancy, pp. 133–149, 2008.

Herzig A., Troquard N., “Knowing how to play: uniform choices in logics of agency”, *in* G. Weiss, P. Stone (eds), *5th International Joint Conference on Autonomous Agents and Multi Agent Systems (AAMAS 2006)*, ACM Press, pp. 209–216, 2006.

Horty J., *Agency and Deontic Logic*, Oxford University Press, 2001.

Jamroga W., Ågotnes T., “Constructive Knowledge: What Agents Can Achieve under Imperfect Information”, *Journal of Applied Non-Classical Logics*, vol. 17, num. 4, pp. 423–475, 2007.

Jamroga W., van der Hoek W., “Agents that Know How to Play”, *Fundamenta Informaticae*, 2004.

Kautz H., Selman B., “Planning as satisfiability”, *Proceedings 10th European Conference on AI*, Wiley, pp. 359–363, 1992.

Lomuscio A., van der Meyden R., Ryan M., “Knowledge in multi-agent systems: initial configurations and broadcast”, *ACM Transactions on Computational Logic*, 2000.

Lorini E., Herzig A., Broersen J., Troquard N., “Grounding power on actions and mental attitudes”, *in* R. Verbrugge, B. Dunin-Keplicz (eds), *Proc. Workshop FAMAS'07, England, UK*, 2007.

Marx M., “Complexity of products of modal logics”, *J. of Logic and Computation*, vol. 9, num. 2, pp. 221–238, 1999.

Osborne M., Rubinstein A., *A Course in Game Theory*, The MIT Press, 1994.

Passy S., Tinchev T., “An essay in combinatory dynamic logic”, *Information and Computation*, vol. 93, pp. 263–332, 1991.

Pauly M., Logic for Social Software, PhD thesis, University of Amsterdam, 2001. ILLC Dissertation Series 2001-10.

Pauly M., “A Modal Logic for Coalitional Power in Games.”, *Journal of Logic and Computation*, vol. 12, num. 1, pp. 149-166, 2002.

Schobbens P.-Y., “Alternating-time logic with imperfect recall”, *Electronic Notes in Theoretical Computer Science*, 2004.

Schwarzentruber F., *“Décidabilité et complexité de la logique normale des coalitions”*, Master's thesis, Université Toulouse 3, 2007.

Troquard N., Independent agents in branching time, PhD thesis, University of Toulouse & University of Trento, 2007.

van der Hoek W., Wooldridge M., "Tractable multiagent planning for epistemic goals", *Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2002)*, ACM Press, New York, NY, USA, pp. 1167–1174, 2002.

Venema Y., "Rectangular games", *The Journal of Symbolic Logic*, December, 1998.

Xu M., "Axioms For Deliberative STIT", *Journal of Philosophical Logic*, vol. 27, pp. 505–552, 1998.