

Epistemic Games in Modal Logic: Joint Actions, Knowledge and Preferences all together

Emiliano Lorini, François Schwarzentruber, Andreas Herzig
IRIT, Toulouse, France

Abstract

We present in this work a sound and complete modal logic called \mathcal{EDLA} (*Epistemic Dynamic Logic of Agency*) integrating the concepts of joint action, preference and knowledge and enabling to reason about epistemic games in strategic form. We provide complexity results for \mathcal{EDLA} . In the second part of the paper, we study in \mathcal{EDLA} the epistemic and rationality conditions of some classical solution concepts like Nash equilibrium and Iterated Deletion of Strictly Dominated Strategies (IDSDS). In the last part of the paper we combine \mathcal{EDLA} with Dynamic Epistemic Logic (DEL) in order to model epistemic game dynamics.

1 Introduction

We present a modal logic integrating the concepts of joint action, preference and knowledge. Our logic supports reasoning about epistemic games in strategic form in which agents decide what to do according to some general principles of rationality while being uncertain about several aspects of the interaction such as other agents' choices, other agents' preferences, etc. Although epistemic games have been extensively studied in economics in the so-called area of interactive epistemology (see, e.g., [1, 10, 9, 3, 11]) and there have been some analysis of epistemic games in modal logic (see, e.g., [5, 12, 9, 21]), no modal approach to epistemic games in strategic form has been proposed up to now which addresses all the following issues at the same time: to provide a *formal language*, and a corresponding *formal semantics*, which is sufficiently general to express solution concepts like Nash Equilibrium or Iterated Deletion of Strictly Dominated Strategies (IDSDS) and to deduce formally the epistemic and rationality conditions on which such solution concepts are based; to prove its *soundness* and *completeness*; to study its computational properties like *decidability* and *complexity*. In this paper, we try to fill this gap by proposing a sound and complete modal logic for epistemic games interpreted on a Kripke-style semantics. We also provide complexity results for our logic.

The remainder of the paper is organized as follows. In Section 2 we present our modal logic of joint actions, preference and knowledge called \mathcal{EDLA} (*Epistemic Dynamic Logic of Agency*). Section 3 is devoted to the analysis in \mathcal{EDLA} of the epistemic conditions of Nash equilibrium and IDSDS. In Section 4 we make \mathcal{EDLA} dynamic by extending it with constructions of Dynamic Epistemic Logic (DEL) [13], and we show that this dynamic version of \mathcal{EDLA} allows to express IDSDS in a more compact way than in the static \mathcal{EDLA} . In Section 5 we show how our logical framework can be applied to the analysis of strategic interaction with imperfect information about the

game structure. Finally, in Section 6, we compare our approach with some existing approaches to epistemic games in modal logic.

2 A logic of joint actions, knowledge and preferences

The logic \mathcal{EDLA} (*Epistemic Dynamic Logic of Agency*) is an extension of the logic \mathcal{DLA} (*Dynamic Logic of Agency*) with modal operators for preference and knowledge modalities. \mathcal{DLA} itself, which was presented in [15, 18], extends dynamic logic by a modal operator of historic possibility quantifying over possible joint actions of all agents. This operator is borrowed from STIT theory [4]. In [15, 18] the relationships between \mathcal{DLA} and Coalition Logic (CL), and \mathcal{DLA} and STIT have been studied. We will come back to this point in Section 2.3.

2.1 Syntax

The syntactic primitives of \mathcal{EDLA} are the finite set of agents Agt , the set of atomic formulas Atm and a nonempty finite set of atomic action names $Act = \{a_1, a_2, \dots, a_{|Act|}\}$.

The language $\mathcal{L}_{\mathcal{EDLA}}$ of the logic \mathcal{EDLA} is given by the following BNF:

$$\varphi ::= p \mid \perp \mid \neg\varphi \mid \varphi \vee \varphi \mid \langle i:a \rangle \varphi \mid \diamond\varphi \mid K_i\varphi \mid [\text{good}]_i\varphi$$

where p ranges over Atm , a ranges over Act , and i ranges over Agt .

It is supposed that every agent performs exactly one action at a time, that all actions are independent, that actions of different agents are performed in parallel and lead to a unique successor state. Therefore the formula $\langle i:a \rangle \varphi$ reads “ i performs action a and φ holds afterwards”, and $\langle i:a \rangle \top$ reads “ i performs a ”. Note that this is slightly different from the standard PDL reading “*there is a possible execution of action a after which φ holds*”, which takes into account that there could be different executions of the same action leading to different successor states. $\langle i:a \rangle \top \wedge \langle j:b \rangle \top$ means that i and j respectively perform a and b in parallel.

The operator \diamond quantifies over possible joint actions of all agents, that is, over the strategy profiles of the current game (the terms “joint actions of all agents” and “strategy profiles” are supposed here to be synonymous). $\diamond\varphi$ reads “ φ holds for some alternative strategy profile of the current game”, or simply “ φ is possibly true”.

The classical Boolean connectives \wedge , \rightarrow , \leftrightarrow and \top (tautology) are defined from \perp , \vee and \neg in the usual manner. Moreover, $[i:a]\varphi$ abbreviates $\neg\langle i:a \rangle\neg\varphi$, $\square\varphi$ abbreviates $\neg\diamond\neg\varphi$ and $\widehat{K}_i\varphi$ abbreviates $\neg K_i\neg\varphi$. $\square\varphi$ means “ φ is necessarily true”. Therefore $[i:a]\perp$ reads “ i does not perform action a ”, and $[i:a]\varphi$ reads “if i performs a then φ holds afterwards. The following abbreviations are convenient to speak about joint actions. Sets of agents are called *coalitions*, noted C_1, C_2, \dots . To every agent $i \in Agt$ we associate the set of all possible ordered pairs $i:a$, that is, $Act_i = \{i:a \mid a \in Act\}$. Besides, we note Δ the set of all joint actions of all agents (alias strategy profiles), that is, $\Delta = \prod_{i \in Agt} Act_i$.

Elements in Δ are n -tuples noted $\alpha, \beta, \gamma, \delta, \dots$. Given $\delta \in \Delta$, we note δ_i the element in δ corresponding to agent i . Finally, we note $\delta_C = (\delta_i)_{i \in C}$ the tuple which consists of the vector of all δ_i for $i \in C$. Therefore $\delta_{Agt} = \delta$. Moreover, we write $\delta_{-i} = \delta_{Agt \setminus \{i\}}$.

The following abbreviation will be useful to axiomatize \mathcal{EDLA} . For every $\delta \in \Delta$ and $C \subseteq Agt$: $\langle \delta_C \rangle \varphi \stackrel{\text{def}}{=} \bigwedge_{j \in C} \langle \delta_j \rangle \varphi$. $\langle \delta_C \rangle \varphi$ reads “the joint action δ_C is going to be performed by coalition C and φ will be true afterwards”. For example, $\langle i:a, j:b \rangle \varphi$

abbreviates $\langle i:a \rangle \varphi \wedge \langle j:b \rangle \varphi$, and stands for “the joint action $\langle i:a, j:b \rangle$ is going to be performed, and φ will be true afterwards”. As usual $[\delta_C] \varphi \stackrel{\text{def}}{=} \neg \langle \delta_C \rangle \neg \varphi$.

Construction $K_i \varphi$ is read as usual “agent i knows that φ ”, whereas the construction $[\text{good}]_i \varphi$ is read “ φ is true in all worlds which are for agent i at least as good as the current one concerning the strategy profile that is chosen”. We define $\langle \text{good} \rangle_i \varphi$ as an abbreviation of $\neg [\text{good}]_i \neg \varphi$. Operators $[\text{good}]_i$ are used in \mathcal{EDLA} to define agents’ preference orderings over the strategy profiles of the current game. Similar operators are studied in [7, 6]. We use $EK_C \varphi$ as an abbreviation of $\bigwedge_{i \in C} K_i \varphi$, i.e. every agent in C knows φ . Then we define by induction $EK_C^k \varphi$ for all $k \in \mathbf{N}$: $EK_C^0 \varphi \stackrel{\text{def}}{=} \varphi$ and for all $k \geq 1$, $EK_C^k \varphi \stackrel{\text{def}}{=} EK_C(EK_C^{k-1} \varphi)$. We define for all $n \in \mathbf{N}$, $MK_C^n \varphi$ as an abbreviation of $\bigwedge_{1 \leq k \leq n} EK_C^k \varphi$. $MK_C^n \varphi$ expresses C ’s mutual knowledge that φ to n iterations, i.e. everyone in C knows φ , everyone in C knows that everyone in C knows φ , and so on until level n .

2.2 Semantics

Frames are tuples $F = \langle W, R, \sim, E, \preceq \rangle$ where:

- W is a nonempty set of possible worlds or states;
- $R : \text{Agt} \times \text{Act} \rightarrow W \times W$ maps every agent-action pair $i:a$ to a transition relation $R_{i:a} \subseteq W \times W$ between possible worlds;
- \sim is an equivalence relation on W ;
- $E : \text{Agt} \rightarrow W \times W$ maps to every agent i an equivalence relation E_i on W ;
- $\preceq : \text{Agt} \rightarrow W \times W$ maps to every agent i a reflexive and transitive relation \preceq_i on W .

It is convenient to use $R_{\delta_C} = \bigcap_{i \in C} R_{\delta_i}$, and $R_{\delta_C}(w) = \{w' \in W \mid w R_{\delta_C} w'\}$. If $R_{i:a}(w) \neq \emptyset$ then i performs a at w . More generally, if $R_{\delta_C}(w) \neq \emptyset$ then coalition C performs joint action δ_C at w . If $w' \in R_{\delta_C}(w)$ then world w' results from the performance of joint action δ_C by coalition C at w .

If $w' \sim w$ then w and w' correspond to alternative strategy profiles of the same game. For short, we say that w' is alternative to w . Given a world w , we use the notation $\sim(w) = \{w' \mid w' \sim w\}$ to denote the equivalence class made up of those worlds which correspond to alternative strategy profiles of the game of which w is part. Consider e.g. $\text{Agt} = \{1, 2\}$ and $\text{Act} = \{a, b\}$. In the frame in Fig. 1 we have $w_1 \sim w_2$. This means that the strategy profile performed at w_1 (viz. $\langle 1:c, 1:c \rangle$) and the one performed at w_2 (viz. $\langle 1:c, 1:d \rangle$) are alternative strategy profiles of the same game defined by the equivalence class $\sim(w_1) = \{w_1, w_2, w_3, w_4\}$. For every $C \subseteq \text{Agt}$, if there exists $w' \in \sim(w)$ such that C performs δ_C at w' then we say that δ_C is *possible* at w (or δ_C can be performed at w).

$w E_i w'$ means that, for agent i , world w' is (epistemically) possible at w , whilst $w \preceq_i w'$ means that for agent i , world w' is at least as good as world w . We write $w =_i w'$ iff $w \preceq_i w'$ and $w' \preceq_i w$, and $w <_i w'$ iff $w \preceq_i w'$ and not $w' \preceq_i w$.

Frames have to satisfy the semantic constraints S1-S9 in Table 1 in order to be \mathcal{EDLA} -frames. According to Constraint S1, for every world w there exists exactly one successor of w , viz. the world resulting from the execution of the strategy profile associated to w . According to the Constraints S2 and S3, every world is associated to

For every $w, v, v' \in W$ and $\delta, \delta' \in \Delta$ and $a \in Act$ and $i \in Agt$, we have:

- (S1) If $v \in R_\delta(w)$ and $v' \in R_{\delta'}(w)$ then $v = v'$;
- (S2) $\bigcup_{\delta \in \Delta} R_\delta(w) \neq \emptyset$;
- (S3) If $\delta \neq \delta'$ then $R_\delta(w) = \emptyset$ or $R_{\delta'}(w) = \emptyset$;
- (S4) If for every $i \in Agt$ there is v_i such that $v_i \sim w$ and $R_{\delta_i}(v_i) \neq \emptyset$ then there is a v such that $v \sim w$ and $R_\delta(v) \neq \emptyset$;
- (S5) If $w \sim v$ and $R_\delta(w) \neq \emptyset$ and $R_\delta(v) \neq \emptyset$, then $w = v$;
- (S6) If wE_iv then $R_{i:a}(w) \neq \emptyset$ iff $R_{i:a}(v) \neq \emptyset$;
- (S7) If $w \preceq_i v$ then $w \sim v$;
- (S8) If $w \sim v$ and $w \sim v'$ then $v \preceq_i v'$ or $v' \preceq_i v$;
- (S9) If wE_iw' then $w \sim w'$.

Table 1: Semantic constraints over \mathcal{EDLA} -frames

exactly one joint action of all agents (alias strategy profile). Note that constraints S1 and S2 together ensure that for every world w there is exactly one next (future) world. According to the Constraint S4, if every individual action in a joint action δ is possible at w , then their simultaneous occurrence is also possible at w . We suppose determinism for the joint actions of all agents: different worlds in an equivalence class $\sim(w)$ correspond to the occurrences of *different* strategy profiles (Constraint S5). Constraint S6 just says that agents know what they are doing. This is a standard assumption in interactive epistemology and epistemic analysis of games (see [9] for instance). We also have two constraints over the relations \preceq_i . We suppose that a world w' is for agent i at least as good as w only if w' is a world which is possible at w , i.e. only if w' and w correspond to alternative strategy profiles of the same game (Constraint S7). Furthermore, we suppose that every agent has a complete preference ordering over the strategy profiles of the current game (Constraint S8).¹ Finally, we suppose perfect information about the specification of the game, including the players' strategy sets (or action repertoires) and the players' preference ordering over strategy profiles. This assumption is formally expressed by Constraint S9: if world w' is epistemically possible for agent i at w , then w and w' correspond to alternative strategy profiles of the same game. Perfect information about the structure of the game is a standard assumption in game theory. In Section 5, this assumption will be relaxed in order to deal with realistic situations in which an agent might be uncertain about his own utility and other agents' utilities associated to a certain strategy profile, as well as about his own action repertoire and other agents' action repertoires.

A frame F is a \mathcal{EDLA} -frame if F satisfies constraints S1-S9.

A \mathcal{EDLA} -model is a couple $M = \langle F, \pi \rangle$ where F is a \mathcal{EDLA} -frame (satisfying constraints S1-S9) and $\pi : Atm \rightarrow 2^W$ is a valuation function.

Truth conditions for atomic formulas and the Boolean operators are standard. The truth conditions for the modal operators are:

- $M, w \models \langle i:a \rangle \varphi$ iff $M, w' \models \varphi$ for some $w' \in R_{i:a}(w)$;
- $M, w \models \Diamond \varphi$ iff $M, w' \models \varphi$ for some $w' \in \sim(w)$;
- $M, w \models K_i \varphi$ iff $M, w' \models \varphi$ for all w' such that wE_iw' ;
- $M, w \models [\text{good}]_i \varphi$ iff $M, w' \models \varphi$ for all w' such that $w \preceq_i w'$.

¹Note that, given the properties of \sim , S7 and S8 can be replaced by: $\sim = \preceq_i \cup (\preceq_i)^{-1}$.

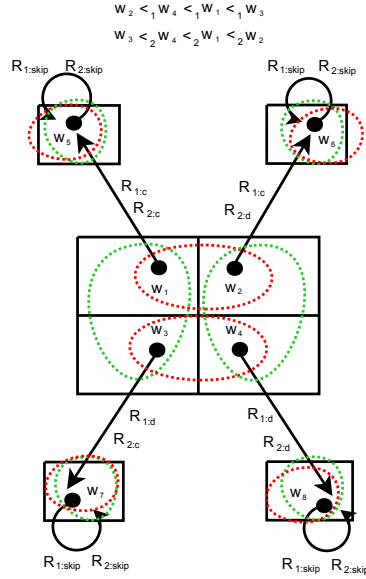


Figure 1: The equivalence class $\{w_1, w_2, w_3, w_4\}$ represents the classical Prisoner's Dilemma game [19]. At w_1 agent 1 and agent 2 cooperate, at w_4 they both defect, at w_2 1 cooperates and 2 defects, and at w_3 1 defects and 2 cooperates. Red circles are epistemic possibility relations for 1, green circles are epistemic possibility relations for 2 (1 and 2 know what they are going to do but are uncertain about the other's action).

A formula φ is *true in a \mathcal{EDLA} -model M* iff $M, w \models \varphi$ for every world w in M . φ is *\mathcal{EDLA} -valid* (noted $\models \varphi$) iff φ is true in all \mathcal{EDLA} -models. φ is *\mathcal{EDLA} -satisfiable* iff $\neg\varphi$ is not \mathcal{EDLA} -valid.

2.3 Axiomatization

We call \mathcal{EDLA} the logic that is axiomatized by the principles given in Table 2. Note that Axiom **Indep** is the \mathcal{EDLA} counterpart of the so-called *axiom of independence of agents* of STIT logic [4]. This axiom allows to express the basic game theoretic assumption that the set of strategy profiles of a game in strategic form is the cartesian product of the sets of individual actions for the agents in Agt .

We write $\vdash_{\mathcal{EDLA}} \varphi$ if φ is a theorem of \mathcal{EDLA} . We can define in \mathcal{EDLA} an operator *next* of linear temporal logic: $X\varphi \stackrel{\text{def}}{=} \bigvee_{\delta \in \Delta} \langle \delta \rangle \varphi$. Due to Axioms **Active** and **Alt**_[δ], X obeys the standard validity $X\varphi \leftrightarrow \neg X\neg\varphi$.

Theorem 1. *\mathcal{EDLA} is determined by the class of \mathcal{EDLA} -frames.*

Theorem 2. *The satisfiability problem of \mathcal{EDLA} is PSPACE-complete².*

It has been proved in [18] that the basic fragment of \mathcal{EDLA} without preference and knowledge modalities embeds Coalition Logic (CL) [20]. In particular, CL cooperation

²You can find a basic implementation of a \mathcal{EDLA} -prover for two agents and two actions at <http://www.irit.fr/~Francois.Schwarzentruber/publications/tark2009/>

(CPL)	All principles of classical propositional logic
(S5 \Box)	All principles of modal logic S5 for \Box
(K $_{[i:a]}$)	All principles of modal logic K for every $[i:a]$
(S5 $_{K_i}$)	All principles of modal logic S5 for every K_i
(S4 $_{[\text{good}]_i}$)	All principles of modal logic S4 for every $[\text{good}]_i$
(Alt $_{[\delta]}$)	$\langle \delta \rangle \varphi \rightarrow [\delta'] \varphi$
(Active)	$\bigvee_{\delta \in \Delta} \langle \delta \rangle \top$
(Single)	$\langle \delta \rangle \top \rightarrow [\delta'] \perp$ if $\delta \neq \delta'$
(Indep)	$\left(\bigwedge_{i \in \text{Agt}} \diamond \langle \delta_i \rangle \top \right) \rightarrow \diamond \langle \delta \rangle \top$
(JointDet)	$\langle \delta \rangle \top \wedge \varphi \rightarrow \Box \langle \delta \rangle \top \rightarrow \varphi$
(Aware)	$\langle i:a \rangle \top \rightarrow K_i \langle i:a \rangle \top$
(Incl $_{[\text{good}]_i, \Box}$)	$\Box \varphi \rightarrow [\text{good}]_i \varphi$
(PrefConnect)	$\diamond \varphi \wedge \diamond \psi \rightarrow \diamond (\varphi \wedge \langle \text{good} \rangle_i \psi) \vee \diamond (\psi \wedge \langle \text{good} \rangle_i \varphi)$
(PerfectInfo)	$\Box \varphi \rightarrow K_i \varphi$

Table 2: Axiomatization of \mathcal{EDLA}

modalities of the form $[C]$ can be reconstructed in our logic as follows.

$$tr([C] \varphi) = \bigvee_{\delta \in \Delta} (\diamond \langle \delta_C \rangle \top \wedge \Box (\langle \delta_C \rangle \top \rightarrow \varphi))$$

That is, the CL expression “coalition C can enforce an outcome state satisfying φ ” (noted $[C] \varphi$) is translated in our logic as “there exists a joint action δ_C of the agents in C such that the agents in C can perform δ_C , and necessarily if the agents in C perform δ_C then φ will be true in the next state, no matter what the agents outside C do”.

It has also been shown in [15] that a slightly different variant of the logic presented in this paper embeds Chellas’ STIT logic with agents and groups [17], under the hypothesis that the number of agents’ choices is bounded. STIT logic has formulas of the form $[C \text{ cstit}:\varphi]$ that are read “group C sees to it that φ ”. To obtain this embedding, it is sufficient to remove from \mathcal{EDLA} the Axiom of joint determinism **JointDet** and to add an Axiom of the form $\langle \delta \rangle \diamond \varphi \rightarrow \diamond \langle \delta \rangle \varphi$ which allows to capture the so-called semantic property *no choice between undivided histories* on STIT frames (see [4]). The translation of STIT modalities of the form $[C \text{ cstit}:\varphi]$ into our logic would be the following:

$$tr([C \text{ cstit}:\varphi]) = \bigvee_{\delta \in \Delta} (\langle \delta_C \rangle \top \wedge \Box (\langle \delta_C \rangle \top \rightarrow \varphi))$$

That is, the STIT expression “group C sees to it that φ ” is translated into \mathcal{DLA} as “there exists a joint action δ_C of the agents in C such that the agents in C perform δ_C , and necessarily if the agents in C perform δ_C then φ will be true, no matter what the agents outside C do”.

3 A logical account of epistemic games

3.1 Best Response and Nash Equilibrium

The modal operators $[\text{good}]_i$ and \Box allow to capture in \mathcal{EDLA} a notion of comparative goodness over formulas of the kind “ φ is for agent i at least as good as ψ ”, noted $\psi \leq_i \varphi$:

$$\psi \leq_i \varphi \stackrel{\text{def}}{=} \Box (\psi \rightarrow \langle \text{good} \rangle_i \varphi).$$

$\psi \leq_i \varphi$ is a total preorder: the formulas $\psi \leq_i \psi$, $(\varphi_1 \leq_i \varphi_2) \wedge (\varphi_2 \leq_i \varphi_3) \rightarrow (\varphi_1 \leq_i \varphi_3)$ and $(\varphi_1 \leq_i \varphi_2) \vee (\varphi_2 \leq_i \varphi_1)$ are valid in \mathcal{EDLA} . More details on this notion of comparative goodness can be found in [6]. We note $\psi <_i \varphi \stackrel{\text{def}}{=} (\psi \leq_i \varphi) \wedge \neg(\varphi \leq_i \psi)$ and $\delta \leq_i \delta' \stackrel{\text{def}}{=}} \langle \delta \rangle \top \leq_i \langle \delta' \rangle \top$. Finally we note $\delta <_i \delta' \stackrel{\text{def}}{=} (\delta \leq_i \delta') \wedge \neg(\delta' \leq_i \delta)$.

Some basic concepts of game theory can be expressed in \mathcal{EDLA} in terms of comparative goodness. We first consider *best response*. Agent i 's action a is said to be a best response to the other agents' joint action δ_{-i} , noted $\text{BR}(i:a, \delta_{-i})$, if and only if i cannot improve his utility by deciding to do something different from a while the others choose the joint action δ_{-i} , that is:

$$\text{BR}(i:a, \delta_{-i}) \stackrel{\text{def}}{=} \bigwedge_{b \in \text{Act}} (\langle i:b \rangle \perp \wedge \langle \delta_{-i} \rangle \top) \leq_i (\langle i:a \rangle \top \wedge \langle \delta_{-i} \rangle \top).$$

Note that the following equivalence is valid in \mathcal{EDLA} :

$$\text{BR}(i:a, \delta_{-i}) \leftrightarrow (\langle i:a \rangle \perp \wedge \langle \delta_{-i} \rangle \top) \leq_i (\langle i:a \rangle \top \wedge \langle \delta_{-i} \rangle \top).$$

Given a certain strategic game, the strategy profile (or joint action) δ is said to be a *Nash equilibrium* if and only if for every agent $i \in \text{Agt}$, i 's action δ_i is a best response to the other agents' joint action δ_{-i} :

$$\text{Nash}(\delta) \stackrel{\text{def}}{=} \bigwedge_{i \in \text{Agt}} \text{BR}(\delta_i, \delta_{-i}).$$

From Axiom **PerfectInfo** and S5 for \Box , the following theorems are provable expressing perfect information about the structure of the game, and in particular, perfect information about the players' preferences ordering over strategy profiles, perfect information about the existence of a Nash equilibrium, and perfect information about the players' repertoires: $(\psi \leq_i \varphi) \leftrightarrow \text{MK}_{\text{Agt}}^n(\psi \leq_i \varphi)$, $\text{Nash}(\delta) \leftrightarrow \text{MK}_{\text{Agt}}^n \text{Nash}(\delta)$ and $\langle \delta_i \rangle \top \leftrightarrow \text{MK}_{\text{Agt}}^n \langle \delta_i \rangle \top$, for every $n \in \mathbf{N}$.

3.2 Epistemic rationality

The following \mathcal{EDLA} formula characterizes a notion of rationality which is commonly supposed in epistemic analysis of games (see, e.g., [3, 5]):

$$\bigwedge_{a,b \in \text{Act}} (\langle i:a \rangle \top \rightarrow \bigvee_{\delta \in \Delta} (\widehat{\text{K}}_i \langle \delta_{-i} \rangle \top \wedge (\langle \delta_{-i}, i:b \rangle \leq_i \langle \delta_{-i}, i:a \rangle))).$$

This means that an agent i is rational if and only if, if he chooses a particular action a then for every alternative action b , there exists a joint action δ_{-i} of the other agents that he considers possible such that, playing a while the others play δ_{-i} is for i at least as good as playing b while the others play δ_{-i} . As \mathcal{EDLA} $\delta \leq_i \delta'$ and $\text{K}_i(\delta \leq_i \delta')$ are equivalent, the previous definition of rationality can be rewritten in the following equivalent form:

$$\text{Rat}_i \stackrel{\text{def}}{=} \bigwedge_{a,b \in \text{Act}} (\langle i:a \rangle \top \rightarrow \bigvee_{\delta \in \Delta} (\widehat{\text{K}}_i \langle \delta_{-i} \rangle \top \wedge \text{K}_i(\langle \delta_{-i}, i:b \rangle \leq_i \langle \delta_{-i}, i:a \rangle))).$$

Theorem 3. For all $i \in \text{Agt}$, we have:

$$(3a) \quad \vdash_{\mathcal{EDLA}} \text{Rat}_i \leftrightarrow \text{K}_i \text{Rat}_i$$

$$(3b) \quad \vdash_{\mathcal{EDLA}} \neg \text{Rat}_i \leftrightarrow \text{K}_i \neg \text{Rat}_i$$

Theorem 3 highlights that the concepts of rationality and irrationality are introspective. The following theorem specifies some sufficient epistemic conditions for guaranteeing that the chosen strategy profile is a Nash equilibrium: if all agents are rational and every agent knows the choices of the other agents, then the selected strategy profile is a Nash equilibrium. A similar theorem has been stated for the first time in [1, 11].

Theorem 4. For all $n \in \mathbb{N}$, for all $\delta \in \Delta$, we have:

$$\vdash_{\mathcal{EDLA}} \left(\left(\bigwedge_{i \in \text{Agt}} \text{Rat}_i \right) \wedge \bigwedge_{i \in \text{Agt}} \text{K}_i \langle \delta_{-i} \rangle \top \right) \rightarrow \text{Nash}(\delta)$$

3.3 Iterated deletion of strictly dominated strategies

A strategy a for agent i is a *strictly dominated strategy*, noted $\text{SD}^{\leq 0}(i:a)$, if and only if, if a can be performed then there is another strategy b such that, no matter what joint action δ_{-i} the other agents choose, playing b is for i strictly better than playing a :

$$\text{SD}^{\leq 0}(i:a) \stackrel{\text{def}}{=} \langle i:a \rangle \top \rightarrow \bigvee_{b \in \text{Act}} \left(\langle i:b \rangle \top \wedge \bigwedge_{\delta \in \Delta} (\langle \delta_{-i} \rangle \rightarrow (\langle \delta_{-i}, i:a \rangle <_i \langle \delta_{-i}, i:b \rangle)) \right).$$

An example of strictly dominated strategy is cooperation in the Prisoners Dilemma (PD) game: whether ones opponent chooses to cooperate or defect, defection yields a higher payoff than cooperation. Therefore, a rational player will never play a dominated strategy: whatever he believes his opponents will do, he will achieve a better outcome if he plays a dominating strategy. So when trying to predict the behavior of rational players, we can rule out all strictly dominated strategies. The so-called Iterated Deletion of Strictly Dominated Strategies (IDSDS) (or iterated strict dominance) [19] is a procedure that starts with the original game and, at each step, for every player i removes from the game all i 's strictly dominated strategies, thereby generating a subgame of the original game, and that repeats this process again and again. IDSDS can be inductively characterized in our logic \mathcal{EDLA} by defining a concept of strict dominance in the subgame of depth at most n , noted $\text{SD}^{\leq n}(i:a)$. For every $n \geq 1$:

$$\text{SD}^{\leq n}(i:a) \stackrel{\text{def}}{=} \neg \text{SD}^{\leq n-1}(i:a) \rightarrow \bigvee_{b \in \text{Act}} \left(\neg \text{SD}^{\leq n-1}(i:b) \wedge \bigwedge_{\delta \in \Delta} (\neg \text{SD}^{\leq n-1}(\delta_{-i}) \rightarrow (\langle \delta_{-i}, i:a \rangle <_i \langle \delta_{-i}, i:b \rangle)) \right).$$

where $\text{SD}^{\leq k}(\delta_C)$ is an abbreviation of $\bigvee_{i \in C} \text{SD}^{\leq k}(\delta_i)$ for every $k \geq 0$ and for every δ_C . According to this definition, a is a strictly dominated strategy for agent i in a subgame of depth at most n , noted $\text{SD}^{\leq n}(i:a)$, if and only if, if a is not strictly dominated for i in all subgames of depth $k < n$ then there is another strategy b such that b is not strictly dominated for i in all subgames of depth $k < n$ and, no matter what joint action δ_{-i} the other agents choose, if the elements in δ_{-i} are not dominated in all subgames of depth $k < n$ then playing b is for i strictly better than playing a . In other terms $\text{SD}^{\leq n}(i:a)$ means that strategy $i:a$ does not survive after n rounds of IDSDS. On the contrary, $\neg \text{SD}^{\leq n}(i:a)$ means that strategy $i:a$ survives IDSDS until the subgame of depth $n+1$.

It has been shown that common knowledge of rationality implies that players choose strategies which survive IDSDS ([10, 9, 3, 11]). This latter principle can be derived in our logic \mathcal{EDLA} . According to the following Theorem 5, if there is mutual knowledge of rationality among the players to n levels and the agents play the strategy profile δ then, for every agent i , δ_i survives IDSDS until the subgame of depth $n+1$.

Theorem 5. For all $\delta \in \Delta$, $\vdash_{\mathcal{EDLA}} \left(\left(\text{MK}_{\text{Agt}}^n \bigwedge_{i \in \text{Agt}} \text{Rat}_i \right) \wedge \langle \delta \rangle \top \right) \rightarrow \neg \text{SD}^{\leq n}(\delta)$

4 Game transformation

We provide in this section an alternative and more compact characterization of the procedure IDSDS in our logic \mathcal{EDLA} . To this aim, we introduce special events whose effect is to delete a strictly dominated strategy from the current game. These events are similar to announcements in Dynamic Epistemic Logic (DEL) [13].

\mathcal{L}_{AN} is the set of *announcable formulas* and is defined by the following BNF:

$$\chi ::= \Box\psi \rightarrow [i:a] \perp \mid \chi \wedge \chi$$

where $\psi \in \mathcal{L}_{\mathcal{EDLA}}$, $i \in \text{Agt}$ and $a \in \text{Act}$. Thus, announcable formulas are of the form ‘if property ψ (globally) holds in the current game, then action a should not be performed by agent i ’.

We extend the \mathcal{EDLA} language with announcements operators of the form $[\chi!]$ with $\chi \in \mathcal{L}_{AN}$. We call \mathcal{EDLA}^{AN} the extended logic. The truth condition for $[\chi!]\varphi$ is:

$$M, w \models [\chi!]\varphi \text{ iff } M, w \models \chi \text{ implies } M^X, w \models \varphi$$

with:

$$M^X = \langle W^X, R^X, \sim^X, E^X, \preceq^X, \pi^X \rangle$$

$$W^X = \|\chi\|_M$$

$$R_{i:a}^X = [R_{i:a} \cap (W^X \times W^X)] \cup \{(w, w) \mid R_{i:a}(w) \neq \emptyset \ \& \ w \in \|\chi\|_M \ \& \ R_{i:a}(w) \notin \|\chi\|_M\}$$

$$\sim^X = \sim \cap (W^X \times W^X)$$

$$E_i^X = E_i \cap (W^X \times W^X)$$

$$\preceq_i^X = \preceq_i \cap (W^X \times W^X)$$

$$\pi^X(p) = \pi(p) \cap W^X$$

Thus, an event $\chi!$ removes from every game $\sim(w)$ in the model M all worlds in which χ is false. The same operation is applied for the agents’ epistemic accessibility relations E_i and preference orderings \preceq_i . This ensures that the constraints S7 and S9 on \mathcal{EDLA} frames will be preserved after the model transformation. Moreover, if v is the world which results from the execution of a by i at w , and w and v are not removed from the model by the event $\chi!$ then, after the occurrence of $\chi!$, v is still the world which results from the execution of a by i at w ; if v is the world which results from the execution of a by i at w , w is not removed from the model by the event $\chi!$ and v is removed then, after the occurrence of $\chi!$, we impose that w itself is the world which results from the execution of a by i at w . This ensures that the constraint S2 on \mathcal{EDLA} frames will be preserved after the model transformation.

Theorem 6. *If M is a \mathcal{EDLA} model then M^X is a \mathcal{EDLA} model.*

We have reduction axioms for $\chi!$ which guarantee the completeness of \mathcal{EDLA}^{AN} .

Theorem 7. *The following schemata are valid in \mathcal{EDLA}^{AN} .*

$$\mathbf{R1.} \quad [\chi!]p \leftrightarrow (\chi \rightarrow p)$$

$$\mathbf{R2.} \quad [\chi!]\neg\varphi \leftrightarrow (\chi \rightarrow \neg[\chi!]\varphi)$$

$$\mathbf{R3.} \quad [\chi!](\varphi_1 \wedge \varphi_2) \leftrightarrow ([\chi!]\varphi_1 \wedge [\chi!]\varphi_2)$$

$$\mathbf{R4.} \quad [\chi!]\Box\varphi \leftrightarrow (\chi \rightarrow \Box[\chi!]\varphi)$$

$$\mathbf{R5.} \quad [\chi!]K_i\varphi \leftrightarrow (\chi \rightarrow K_i[\chi!]\varphi)$$

$$\mathbf{R6.} \quad [\chi!][\text{good}]_i\varphi \leftrightarrow (\chi \rightarrow [\text{good}]_i[\chi!]\varphi)$$

$$\mathbf{R7.} \quad [\chi!][i:a]\varphi \leftrightarrow (\neg\chi \vee ([i:a]\chi \wedge [i:a][\chi!]\varphi) \vee (\neg[i:a]\chi \wedge [\chi!]\varphi))$$

$$\mathbf{R8.} \quad \text{Rule of replacement of proved equivalence}$$

Theorem 8. *The logic \mathcal{EDLA}^{AN} is completely axiomatized by the axioms and inference rules of \mathcal{EDLA} together with the schemata of Theorem 7.*

Now, consider the following formula:

$$\chi_{SD} \stackrel{\text{def}}{=} \bigwedge_{i \in \text{Agt}, a \in \text{Act}} (\Box SD^{\leq 0}(i:a) \rightarrow [i:a] \perp).$$

The effect of $\chi_{SD}!$ is to delete from every game $\sim(w)$ in the model M all worlds in which a strictly dominated strategy is played by some agent. Note that χ_{SD} is equivalent to $\bigwedge_{i \in \text{Agt}, a \in \text{Act}} (SD^{\leq 0}(i:a) \rightarrow [i:a] \perp)$.

As the following Theorem 9 highlights, the procedure IDSDS that we have characterized in Section 3.3 in the static \mathcal{EDLA} can be characterized in a more compact way in \mathcal{EDLA}^{AN} . Suppose δ is the selected strategy profile. Then, for every agent i , δ_i survives IDSDS until the subgame of depth $n+1$ if and only if, the event $\chi_{SD}!$ can occur $n+1$ times in sequence.

Theorem 9. *For all $\delta \in \Delta$, for all $n \geq 0$,*

$$\vdash_{\mathcal{EDLA}^{AN}} \langle \delta \rangle \top \rightarrow \left(\neg SD^{\leq n}(\delta) \leftrightarrow \langle \chi_{SD}! \rangle^{n+1} \top \right).$$

Finally, here is a reformulation of Theorem 5 in \mathcal{EDLA}^{AN} :

Theorem 10. *For all $n \geq 0$, $\vdash_{\mathcal{EDLA}^{AN}} \left(\text{MK}_{\text{Agt}}^n \bigwedge_{i \in \text{Agt}} \text{Rat}_i \right) \rightarrow \langle \chi_{SD}! \rangle^{n+1} \top$.*

5 Concluding remarks: imperfect information

Let us consider a more general class of games which includes strategic games with imperfect information about the game structure. Apart from few exceptions (see, e.g., [14]), these games have been rarely explored. Indeed, most work in game theory assumed that players have common knowledge of all relevant aspects of the game. We are interested in verifying whether the results obtained in Sections 3.2 and 3.3 can be generalized to this kind of games, that is:

1. Are rationality of every player and every agent's knowledge about other agents' choices still sufficient to ensure that the selected strategy profile is a Nash equilibrium in a strategic game with imperfect information about the game structure?
2. Is mutual knowledge of rationality among the players still sufficient to ensure that the selected strategy profile survives iterated deletion of dominated strategies in a strategic game with imperfect information about the game structure?

To answer these questions, we have to remove Axiom **PerfectInfo** from \mathcal{EDLA} and the corresponding semantic constraint S9 from the definition of \mathcal{EDLA} frames expressing the hypothesis of perfect information about the game structure. We call \mathcal{EDLA}^* the resulting logic and \mathcal{EDLA}^* -frames the resulting class of frames. Then we have to check whether Theorems 4 and 5 given in Sections 3.2 and 3.3 are still derivable in \mathcal{EDLA}^* .

We have a positive answer to the previous first question. Indeed, the formula

$$\left(\left(\bigwedge_{i \in \text{Agt}} \text{Rat}_i \right) \wedge \bigwedge_{i \in \text{Agt}} \text{K}_i \langle \delta_{-i} \rangle \top \right) \rightarrow \text{Nash}(\delta)$$

is derivable in \mathcal{EDLA}^* . But we have a negative answer to the second question. Indeed, the following formula is invalid in \mathcal{EDLA}^* for every $\delta \in \Delta$ and for every $n \in \mathbf{N}$ such that $n > 0$ (see Section A.11 in the Annex for a proof):

$$\left(\left(\text{MK}_{\text{Agt}}^n \bigwedge_{i \in \text{Agt}} \text{Rat}_i \right) \wedge \langle \delta \rangle \top \right) \rightarrow \neg SD^{\leq n}(\delta).$$

6 Related works

Although several modal logics of games in strategic forms have been proposed (see, e.g., [16, 22]), few modal logics exist which support reasoning about epistemic (strategic) games. Among them we should mention [12, 21, 9].

De Bruin [12] has developed a very rich logical framework which enables to reason about the epistemic aspects of strategic games and of extensive games. His system deals with several game-theoretic concepts like the concepts of knowledge, rationality, Nash equilibrium, iterated strict dominance, backward induction. Nevertheless, de Bruin's approach differs on many points from our approach. First of all, our logical approach to epistemic games is *minimalistic* since it relies on few primitive concepts and corresponding modal operators for knowledge, action, historical necessity and preference. All other notions such as Nash equilibrium, rationality, iterated strict dominance are defined by means of these four primitive concepts. On the contrary, in de Bruin's logic all those notions are atomic propositions managed by a *ad hoc* axiomatization (see, e.g., [12, pp. 61,65] where special propositions for rationality and iterated strict dominance are introduced). Secondly, we provide a semantics and a complete axiomatics for our logic of epistemic games. De Bruin's approach is purely syntactic: no model-theoretic analysis of games is proposed nor completeness result for the proposed logic is given. Finally, de Bruin does not provide any complexity results about his logic while we prove that the satisfiability problem of a formula in our logic is PSPACE-complete.

Roy [21] has recently proposed a modal logic integrating preferences, knowledge and intentions. In his approach every world in a model is associated to a nominal which directly refers to a strategy profile in a strategic game. This approach is however limited in expressing formally the structure of a strategic game. In particular, in Roy's logic there is no principle like the \mathcal{EDLA} Axiom **Indep** explaining how possible actions δ_i of individual agents are combined to form a strategy profile δ of the current game. Another limitation of Roy's approach is that it does not allow to express the concept of (weak) rationality that we have been able to define in Section 3.2 (see [21, pp. 101]). As discussed in the previous sections this is a crucial concept in interactive epistemology since it is used for giving epistemic justifications of several solution concepts like Nash equilibrium and IDSDS (see Theorems 4 and 5).

Bonanno [9] integrates modal operators for belief, common belief with constructions expressing agents' preferences over individual actions and strategy profiles, and apply them to the semantic characterization of solution concepts like Iterated Deletion of Strictly Dominated Strategies (IDSDS) and Iterated Deletion of Inferior Profiles (IDIP). As in [21], in Bonanno's logic every world in a model corresponds to a strategy profile of the current game. Although this logic allows to express the concept of weak rationality, it is not sufficiently general to enable to express in the object language solution concepts like Nash equilibrium and IDSDS (note that the latter is defined by Bonanno only in the metalanguage).

It is to be noted that, differently from \mathcal{EDLA} , most modal logics of epistemic games in strategic form (included Roy's logic and Bonanno's logic) postulate a one-to-one correspondence between models and games (i.e. every model of the logic corresponds to a unique strategic game, and worlds in the model are all strategy profiles of this game). Such an assumption is quite restrictive since it prevents from analyzing in the logic games with imperfect information about the game structure in which an agent can imagine alternative games. As shown in Section 5, this is something we can do in our logical framework by removing Axiom **PerfectInfo** from \mathcal{EDLA} .

Before concluding this section about related works it is to be noted that the ap-

proach to game dynamics based on Dynamic Epistemic Logic (DEL) we proposed in Section 4 is inspired by [5] in which strategic equilibrium is defined by fixed-points of operations of repeated announcement of suitable epistemic statements and rationality assertions. However, the analysis of epistemic games proposed in [5] is mainly semantical and the author does not provide a full-fledged modal language for epistemic games which allows to express in the object language solution concepts like Nash Equilibrium or IDSDS, and the concept of rationality. Moreover, van Benthem’s analysis does not include any completeness result for the proposed framework and there is no proposal of reduction axioms for a combination of DEL with a static logic of epistemic games. On the contrary, these two aspects are central in our analysis.

References

- [1] R. J. Aumann and A. Brandenburger. Epistemic conditions for Nash equilibrium. *Econometrica*, 63:1161–1180, 2008.
- [2] P. Balbiani, O. Gasquet, A. Herzig, F. Schwarzentruber, and N. Troquard. Coalition games over Kripke semantics. In *Festschrift in Honour of Shahid Rahman*, pages 1–12. College Publications, 2008.
- [3] P. Battigalli and G. Bonanno. Recent results on belief, knowledge and the epistemic foundations of game theory. *Research in Economics*, 53:149–225, 1999.
- [4] N. Belnap, M. Perloff, and M. Xu. *Facing the future: agents and choices in our indeterminist world*. Oxford University Press, 2001.
- [5] J. van Benthem. Rational dynamics and epistemic logic in games. *International Game Theory Review*, 9(1):13–45, 2007.
- [6] J. van Benthem, P. Girard, and O. Roy. Everything else being equal: a modal logic for *ceteris paribus* preferences. *Journal of Philosophical Logic*, 38(1):83–125, 2009.
- [7] J. van Benthem and F. Liu. Dynamic logic of preference upgrade. *Journal of Applied Non-Classical Logics*, 17(2):157–182, 2007.
- [8] P. Blackburn, M. de Rijke, and Y. Venema. *Modal Logic*. Cambridge University Press, 2001.
- [9] G. Bonanno. A syntactic approach to rationality in games with ordinal payoffs. In *Proc. of LOFT 2008*, pages 59–86, 2008.
- [10] A. Brandenburger and E. Dekel. Rationalizability and correlated equilibria. *Econometrica*, 55:1391–1402, 1987.
- [11] A. Brandenburger. Knowledge and equilibrium in games. *Journal of Economic Perspectives*, 6:83–101, 1992.
- [12] B. de Bruin. *Explaining games: on the logic of game theoretic explanations*. PhD thesis, University of Amsterdam, 2004.
- [13] H. P. van Ditmarsch, W. van der Hoek, and B. Kooi. *Dynamic Epistemic Logic*. Kluwer Academic Publishers, 2007.
- [14] J. Y. Halpern and L. Rego. Generalized solution concepts in games with possibly unaware players. In *Proc. of TARK’07*, pages 253–262, 2007.
- [15] A. Herzig and E. Lorini. A dynamic logic of agency I: STIT, abilities and powers. Technical Report IRIT/RR2009-4FR, Institut de Recherche en Informatique de Toulouse, 2009.
- [16] W. van der Hoek, W. Jamroga, and M. Wooldridge. A logic for strategic reasoning. In *Proc. of AAMAS 2005*, pages 157–164, ACM Press, 2005.
- [17] J. F. Horty. *Agency and Deontic Logic*. Oxford University Press, Oxford, 2001.
- [18] E. Lorini. A dynamic logic of agency II: deterministic DLA, coalition logic, and game theory. Technical Report IRIT/RR2009-5FR, Institut de Recherche en Informatique de Toulouse, 2009.
- [19] M. J. Osborne and A. Rubinstein. *A course in game theory*. MIT Press, 1994.
- [20] M. Pauly. A modal logic for coalitional power in games. *Journal of Logic and Computation*, 12(1):149–166, 2002.
- [21] O. Roy. *Thinking before acting: intentions, logic, rational choice*. PhD thesis, University of Amsterdam, 2008.
- [22] N. Troquard, W. van der Hoek, and M. Wooldridge. A logic of games and propositional control. In *Proc. of AAMAS’09*, pages 961–968, ACM Press, 2009.

A ANNEX: Proofs of some theorems

We add this long annex since it may be useful for evaluating our work. Indeed, some of the proofs are not straightforward. However, if our submission will be accepted, the annex can be partially removed from the final version without affecting the contribution of the paper.

A.1 Proof of Theorem 1

\mathcal{EDLA} is determined by the class of \mathcal{EDLA} -frames.

Proof. All axioms of \mathcal{EDLA} are in the Sahlqvist class. Using the Sahlqvist algorithm it is routine to prove that the Axioms **Alt**_[δ], **Active**, **Single**, **Indep**, **JointDet**, **Aware**, **Incl**_[good], \square , **PrefConnect** and **PerfectInfo** of \mathcal{EDLA} respectively correspond to the constraints S1, S2, S3, S4, S5, S6, S7, S8 and S9. Completeness of \mathcal{EDLA} then follows from Sahlqvist's completeness theorem, cf. [8, Th. 2.42]. \square

A.2 Proof of Theorem 2

We give here a sketch of proof. The Theorem 2 is implied by two facts:

Proposition 1. *The satisfiability problem of \mathcal{EDLA} is PSPACE-hard.*

Proof. Consider B the logic of the class of infinite binary trees. The logic B is PSPACE-complete (we leave to the reader the adaptation of the proof given in [8]). Then the idea is that we can simulate binary trees with one agent and two actions. Let φ a B -formula. We define a translation tr from the language of B to the language of \mathcal{EDLA} by:

$$tr(\varphi) = \square[\bigwedge_{i \in md(\varphi)} (X\square)^i [atmost_{a_1 a_2} foragent1 \wedge justaction_{a_1} foragent1] \wedge tr_1(\varphi)]$$

where:

- $atmost_{a_1 a_2} foragent1 = \bigwedge_{j \in \{3, \dots, n\}} [1:a_j] \perp$;
- $justaction_{a_1} foragent1 = \bigwedge_{i \in Agt, i \neq 1} \bigwedge_{j \in \{2, \dots, n\}} [i:a_j] \perp$;
- $tr_1(\diamond\psi) = \diamond X tr_1(\psi)$;
- $X\psi = \bigvee_{a \in Act} \langle 1:a \rangle \psi$.

We can prove that φ is B-sat iff $tr(\varphi)$ is \mathcal{EDLA} -sat. \square

Proposition 2. *The satisfiability problem of \mathcal{EDLA} is PSPACE.*

Proof. • First, the satisfiability problem of the fragment of \mathcal{EDLA} without time (we only consider formulas φ where all subformulas of the form $[i:a]\psi$ implies $\psi = \perp$), is NP-complete. Indeed, we can prove that if a formula is satisfiable it is in a model where the size is bounded by $card(Act)^{card(Agt)}$;

- Finally we can prove that there exists a deep first search algorithm running in PSPACE. The idea is the same than in [2]. \square

A.3 Proof of Theorem 3a

For all $i \in \text{Agt}$, we have: $\vdash_{\mathcal{EDLA}} \text{Rat}_i \leftrightarrow \text{K}_i \text{Rat}_i$

Proof. Rat_i is equivalent to

$$\bigwedge_{a,b \in \text{Act}} (\langle i:a \rangle \top \rightarrow \bigvee_{\beta \in \Delta} (\widehat{\text{K}}_i \langle \beta_{-i} \rangle \top \wedge (\langle \beta_{-i}, i:b \rangle \leq_i \langle \beta_{-i}, i:a \rangle)))$$

which is equivalent to

$$\bigwedge_{a,b \in \text{Act}} (\text{K}_i [i:a] \perp \vee \bigvee_{\beta \in \Delta} \text{K}_i (\widehat{\text{K}}_i \langle \beta_{-i} \rangle \top \wedge (\langle \beta_{-i}, i:b \rangle \leq_i \langle \beta_{-i}, i:a \rangle))),$$

by classical principles of propositional logic, Axiom 5 for K_i , and the two \mathcal{EDLA} theorems $(\beta' \leq_i \beta) \leftrightarrow \text{K}_i(\beta' \leq_i \beta)$ and $[i:a] \perp \leftrightarrow \text{K}_i [i:a] \perp$.

The latter implies

$$\bigwedge_{a,b \in \text{Act}} \text{K}_i ([i:a] \perp \vee \bigvee_{\beta \in \Delta} (\widehat{\text{K}}_i \langle \beta_{-i} \rangle \top \wedge (\langle \beta_{-i}, i:b \rangle \leq_i \langle \beta_{-i}, i:a \rangle))),$$

by standard principles of normal modal logic. By standard principles of normal modal logic, the latter is equivalent to

$$\text{K}_i \bigwedge_{a,b \in \text{Act}} ([i:a] \perp \vee \bigvee_{\beta \in \Delta} (\widehat{\text{K}}_i \langle \beta_{-i} \rangle \top \wedge (\langle \beta_{-i}, i:b \rangle \leq_i \langle \beta_{-i}, i:a \rangle))).$$

This is equivalent to $\text{K}_i \text{Rat}_i$. \square

A.4 Proof of Theorem 4

For all $n \in \mathbf{N}$, for all $\delta \in \Delta$, we have:

$$\vdash_{\mathcal{EDLA}} \left(\left(\bigwedge_{i \in \text{Agt}} \text{Rat}_i \right) \wedge \bigwedge_{i \in \text{Agt}} \text{K}_i \langle \delta_{-i} \rangle \top \right) \rightarrow \text{Nash}(\delta)$$

Proof. Let us take a \mathcal{EDLA} -model M and a world w such that $M, w \models (\bigwedge_{i \in \text{Agt}} \text{Rat}_i \wedge \bigwedge_{i \in \text{Agt}} \text{K}_i \langle \delta_{-i} \rangle \top)$. Now, let us prove that $M, w \models \text{Nash}(\delta)$.

Let $i \in \text{Agt}$ and let us prove that $M, w \models \text{BR}(\delta_i, \delta_{-i})$.

More precisely, we have to prove that $M, w \models ([\delta_i] \perp \wedge \langle \delta_{-i} \rangle \top) \leq_i (\langle \delta_i \rangle \top \wedge \langle \delta_{-i} \rangle \top)$.

Let $j \neq i$. We have $\text{K}_j \langle -j \rangle \top$, so $M, w \models \langle \delta_i \rangle \top$ by Axiom T for K_j . As $M, w \models \text{Rat}_i$, we then have $M, w \models \bigwedge_{b \in \text{Act}} \bigvee_{\beta \in \Delta} (\widehat{\text{K}}_i \langle \beta_{-i} \rangle \top \wedge (\langle \beta_{-i}, i:b \rangle \leq_i \langle \beta_{-i}, \delta_i \rangle))$. That is to say for all $b \in \text{Act}$, there exists $\beta \in \Delta$ such that $M, w \models \widehat{\text{K}}_i \langle \beta_{-i} \rangle \top \wedge (\langle \beta_{-i}, i:b \rangle \leq_i \langle \beta_{-i}, \delta_i \rangle)$.

But, we have $\text{K}_i \langle \delta_{-i} \rangle \top$. So for all $b \in \text{Act}$, $\beta_{-i} = \delta_{-i}$.

So we have $M, w \models \bigwedge_{b \in \text{Act}} (\langle \beta_{-i}, i:b \rangle \leq_i \langle \beta_{-i}, \delta_i \rangle)$. This is equivalent to $M, w \models ([\delta_i] \perp \wedge \langle \delta_{-i} \rangle \top) \leq_i (\langle \delta_i \rangle \top \wedge \langle \delta_{-i} \rangle \top)$. \square

A.5 Proof of Theorem 5

For all $\delta \in \Delta$, $\vdash_{\mathcal{EDLA}} \left(\left(\text{MK}_{\text{Agt}}^n \bigwedge_{i \in \text{Agt}} \text{Rat}_i \right) \wedge \langle \delta \rangle \top \right) \rightarrow \neg \text{SD}^{\leq n}(\delta)$

Proof. We are going to prove the theorem by induction on n .

- Let us begin to prove the theorem for $n = 0$. Let us take a \mathcal{EDLA} -model M and a world w such that $M, w \models \bigwedge_{i \in \text{Agt}} \text{Rat}_i \wedge \langle \delta \rangle \top$. By definition of Rat_i , we have:

$$M, w \models \bigwedge_{i \in \text{Agt}} \bigwedge_{b \in \text{Act}} \left(\bigvee_{\beta \in \Delta} (\widehat{\text{K}}_i \langle \beta_{-i} \rangle \top \wedge (\langle \beta_{-i}, i:b \rangle \leq_i \langle \beta_{-i}, i : \delta_i \rangle)) \right)$$

This implies:

$$M, w \models \bigwedge_{i \in \text{Agt}} \bigwedge_{b \in \text{Act}} \left(\bigvee_{\beta \in \Delta} (\diamond \langle \beta_{-i} \rangle \top \wedge \langle \beta_{-i}, i:b \rangle \leq_i \langle \beta_{-i}, i : \delta_i \rangle) \right), \text{ by Axiom PerfectInfo.}$$

Furthermore, we have $M, w \models \bigwedge_{i \in \text{Agt}} \langle \delta_i \rangle \top$. So, $M, w \models \bigwedge_{i \in \text{Agt}} \neg \text{SD}^{\leq 0}(\delta_i)$ and $M, w \models \neg \text{SD}^{\leq 0}(\delta)$.

So

$$\vdash_{\mathcal{EDLA}} \left(\bigwedge_{i \in \text{Agt}} \text{Rat}_i \wedge \langle \delta \rangle \top \right) \rightarrow \neg \text{SD}^{\leq 0}(\delta).$$

- Now, let $n \in \mathbf{N}$ and let us prove that if the theorem 5 is true for all $k \leq n$ then it is true for $n+1$. Let us take a \mathcal{EDLA} -model M and a world w such that $M, w \models (\text{MK}_{\text{Agt}}^{n+1} \bigwedge_{i \in \text{Agt}} \text{Rat}_i) \wedge \langle \delta \rangle \top$. We have to prove $M, w \models \neg \text{SD}^{\leq n+1}(\delta)$. That is to say, we have to prove that for all $i \in \text{Agt}$, $M, w \models \neg \text{SD}^{\leq n+1}(\delta_i)$.

$$\neg \text{SD}^{\leq n+1}(\delta_i) = \neg \text{SD}^{\leq n}(\delta_i) \wedge \bigwedge_{b \in \text{Act}} \left(\neg \text{SD}^{\leq n}(i:b) \rightarrow \bigvee_{\beta \in \Delta} \left(\neg \text{SD}^{\leq n}(\beta_{-i}) \wedge (\langle \beta_{-i}, i:b \rangle \leq_i \langle \beta_{-i}, \delta_i \rangle) \right) \right)$$

First, as $M, w \models (\text{MK}_{\text{Agt}}^{n+1} \bigwedge_{i \in \text{Agt}} \text{Rat}_i) \wedge \langle \delta \rangle \top$ we also have $M, w \models (\text{MK}_{\text{Agt}}^n \bigwedge_{i \in \text{Agt}} \text{Rat}_i) \wedge \langle \delta \rangle \top$. So by applying Theorem 5 for n we have $M, w \models \neg \text{SD}^{\leq n}(\delta_i)$.

It remains to be proven $M, w \models \bigwedge_{b \in \text{Act}} (\neg \text{SD}^{\leq n}(i:b) \rightarrow \bigvee_{\beta \in \Delta} (\neg \text{SD}^{\leq n}(\beta_{-i}) \wedge (\langle \beta_{-i}, i:b \rangle \leq_i \langle \beta_{-i}, \delta_i \rangle)))$.

In fact, we are going to prove something less strong:

$$M, w \models \bigwedge_{b \in \text{Act}} \bigvee_{\beta \in \Delta} (\neg \text{SD}^{\leq n}(\beta_{-i}) \wedge (\langle \beta_{-i}, i:b \rangle \leq_i \langle \beta_{-i}, i:\delta_i \rangle)).$$

But, we have $M, w \models \text{Rat}_i \wedge \langle \delta \rangle \top$. So, $M, w \models \bigwedge_{b \in \text{Act}} \bigvee_{\beta \in \Delta} (\widehat{\text{K}}_i \langle \beta_{-i} \rangle \top \wedge (\langle \beta_{-i}, i:b \rangle \leq_i \langle \beta_{-i}, \delta_i \rangle))$.

The only thing which remains to be proven is that we have ' $\neg \text{SD}^{\leq n}(\beta_{-i})$ '.

But for all $b \in \text{Act}$, there exists $\beta \in \Delta$ such that $M, w \models \widehat{\text{K}}_i \langle \beta_{-i} \rangle \top \wedge (\langle \beta_{-i}, i:b \rangle \leq_i \langle \beta_{-i}, \delta_i \rangle)$.

For all $b \in \text{Act}$, there exists a world u such that $w E_i u$ and $M, u \models \langle \beta_{-i} \rangle \top$. As $M, w \models \text{MK}_{\text{Agt}}^{n+1} \bigwedge_{i \in \text{Agt}} \text{Rat}_i$, we have for all $k \leq n$, $M, u \models \text{MK}_{\text{Agt}}^k \bigwedge_{i \in \text{Agt}} \text{Rat}_i$.

The Theorem 5 is supposed to be true by induction for n so $M, u \models \neg \text{SD}^{\leq n}(\beta_{-i})$. But as $\models \neg \text{SD}^{\leq n}(\beta_{-i}) \leftrightarrow \Box \neg \text{SD}^{\leq n}(\beta_{-i})$, and as $E_i \subseteq \sim$ (constraint S9), we have $M, w \models \neg \text{SD}^{\leq n}(\beta_{-i})$, this for all $b \in \text{Act}$.

□

A.6 Proof of Theorem 6

If M is a \mathcal{EDLA} model then M^X is a \mathcal{EDLA} model.

Proof. It is just a routine to verify that \sim^X and every E_i^X are equivalence relations, every \preceq_i^X is reflexive and transitive, and the model M^X satisfies the constraints S7, S8 and S9. The constraints S1, S2, S3, S5 on M^X follow straightforwardly from the definition of $R_{i;a}^X$.

Let us prove that M^X satisfies constraints S4 and S6.

We first prove that M^x satisfies constraint S4. We introduce the following useful notation. Suppose $\chi_1, \chi_2 \in \mathcal{L}_{\mathcal{AN}}$. Then, $\chi_2 \rightsquigarrow \chi_3$ iff there is $\chi_3 \in \mathcal{L}_{\mathcal{AN}}$ such that $\chi_1 = \chi_2 \wedge \chi_3$.

Now, suppose for every $i \in \text{Agt}$ there is v_i such that $v_i \sim^x w$ and $R_{\delta_i}^x(v_i) \neq \emptyset$. It follows that for every $i \in \text{Agt}$ there is v_i such that $v_i \sim w$ and $R_{\delta_i}(v_i) \neq \emptyset$. The latter implies that there is v such that $v \sim w$ and $R_\delta(v) \neq \emptyset$ (by the semantic constraint S4). Now, suppose for all v' if $v' \sim^x w$ then $R_\delta^x(v') = \emptyset$. It follows that: there is $i \in \text{Agt}$ and $\psi \in \mathcal{L}_{\mathcal{EDLA}}$ such that $\Box\psi \rightarrow [\delta_i] \perp \rightsquigarrow \chi$ and $M, v \models \Box\psi$. The latter implies that there is $i \in \text{Agt}$ and $\psi \in \mathcal{L}_{\mathcal{EDLA}}$ such that $\Box\psi \rightarrow [\delta_i] \perp \rightsquigarrow \chi$ and for all $v' \sim w$, $M, v' \models \Box\psi$. We conclude that there is no $v_i \sim^x w$ such that $R_{\delta_i}^x(v_i) \neq \emptyset$ which leads to a contradiction.

We now consider constraint S6. Suppose $wE_i^x v$ and $R_{i:a}^x(w) \neq \emptyset$. It follows that $wE_i v$ and $R_{i:a}(w) \neq \emptyset$ which implies $R_{i:a}(v) \neq \emptyset$, because M satisfies constraint S6. The latter implies $R_{i:a}^x(v) \neq \emptyset$. Now, suppose $wE_i^x v$ and $R_{i:a}^x(v) \neq \emptyset$. It follows that $wE_i v$ and $R_{i:a}(v) \neq \emptyset$ which implies $R_{i:a}(w) \neq \emptyset$, because M satisfies constraint S6. The latter implies $R_{i:a}^x(w) \neq \emptyset$. □

A.7 Proof of Theorem 7

Proof. The proofs of reduction axioms **R1-R6** and of rule **R8** go as in DEL (see [13]). We here prove reduction axiom **R7**.

CASE 1. $\neg\chi$ and $[\chi!] \perp$ are equivalent in \mathcal{EDLA}^{AN} . Therefore,

$$(A) \quad \neg\chi \rightarrow ([\chi!] [i:a] \varphi \leftrightarrow \neg\chi)$$

is valid in \mathcal{EDLA}^{AN} .

CASE 2. Let us suppose $M, w \models [i:a] \chi \wedge \chi$.

$$M, w \models [\chi!] [i:a] \varphi$$

$$\text{IFF if } M, w \models \chi \text{ then } M^x, w \models [i:a] \varphi$$

$$\text{IFF if } M, w \models \chi \text{ then, if } v \in R_{i:a}^x(w) \text{ then } M^x, v \models \varphi$$

$$\text{IFF if } M, w \models \chi \text{ then, if } v \in R_{i:a}(w) \text{ then } M^x, v \models \varphi$$

(because $M, w \models [i:a] \chi \wedge \chi$ implies $R_{i:a}^x(w) = R_{i:a}(w)$),

$$\text{IFF if } M, w \models \chi \text{ then, if } v \in R_{i:a}(w) \text{ then } M^x, v \models \varphi \text{ and } M, v \models \chi$$

(by $M, w \models [i:a] \chi$),

$$\text{IFF if } M, w \models \chi \text{ then, if } v \in R_{i:a}(w) \text{ then } M, v \models [\chi!] \varphi \text{ and } M, v \models \chi$$

$$\text{IFF if } M, w \models \chi \text{ then } M, w \models [i:a] ([\chi!] \varphi \wedge \chi)$$

$$\text{IFF if } M, w \models \chi \rightarrow ([i:a] ([\chi!] \varphi \wedge \chi))$$

$$\text{IFF if } M, w \models \chi \rightarrow [i:a] [\chi!] \varphi$$

(by the hypothesis $M, w \models [i:a] \chi$).

This proves that $([i:a] \chi \wedge \chi) \rightarrow ([\chi!] [i:a] \varphi \leftrightarrow (\chi \rightarrow [i:a] [\chi!] \varphi))$ is valid in \mathcal{EDLA}^{AN} .

It follows that

$$(B) \quad ([i:a] \chi \wedge \chi) \rightarrow ([\chi!] [i:a] \varphi \leftrightarrow [i:a] [\chi!] \varphi)$$

is valid in \mathcal{EDLA}^{AN} too.

CASE 3. Let us suppose $M, w \models \neg[i:a] \chi \wedge \chi$ which is equivalent to $M, w \models \langle i:a \rangle \neg\chi \wedge \chi$.

$$M, w \models [\chi!] [i:a] \varphi$$

$$\text{IFF if } M, w \models \chi \text{ then } M^x, w \models [i:a] \varphi$$

$$\text{IFF if } M, w \models \chi \text{ then, if } v \in R_{i:a}^x(w) \text{ then } M^x, v \models \varphi$$

IFF if $M, w \models \chi$ then $M^x, w \models \varphi$
 (because $M, w \models \langle i:a \rangle \neg \chi \wedge \chi$ implies $R_{i:a}^x(w) = \{w\}$),
 IFF if $M, w \models \chi$ then $M^x, w \models \varphi$ and $M, w \models \chi$
 (by the hypothesis $M, w \models \chi$),
 IFF if $M, w \models \chi$ then $M^x, w \models [\chi!]\varphi$
 IFF if $M, w \models \chi \rightarrow [\chi!]\varphi$
 IFF if $M, w \models [\chi!]\varphi$
 (by the hypothesis $M, w \models \chi$). Therefore

(C) $(\neg [i:a] \chi \wedge \chi) \rightarrow ([\chi!] [i:a] \varphi \leftrightarrow [\chi!]\varphi)$

is valid in \mathcal{EDLA}^{AN} .

From the previous three \mathcal{EDLA}^{AN} validities A, B and C it follows that:

$[\chi!] [i:a] \varphi \leftrightarrow (\neg \chi \vee ([i:a] \chi \wedge [i:a] [\chi!]\varphi) \vee (\neg [i:a] \chi \wedge [\chi!]\varphi))$
 is valid in \mathcal{EDLA}^{AN} .

□

A.8 Proof of Theorem 8

The logic \mathcal{EDLA}^{AN} is completely axiomatized by the axioms and inference rules of \mathcal{EDLA} together with the schemata of Theorem 7.

Proof. By means of the principles **R1-R8** in Theorem 7, it is straightforward to prove that for every \mathcal{EDLA}^{AN} formula there is an equivalent \mathcal{EDLA} formula. In fact, each reduction axiom **R2-R7**, when applied from the left to the right by means of the rule of replacement of proved equivalence **R8**, yields a simpler formula, where 'simpler' roughly speaking means that the dynamic operator is pushed inwards. Once the dynamic operator attains an atom it is eliminated by the equivalence **R1**. Hence, the completeness of \mathcal{EDLA}^{AN} is a straightforward consequence of Theorem 1. □

A.9 Proof of Theorem 9

For all $\delta \in \Delta$, for all $n \geq 0$,

$$\vdash_{\mathcal{EDLA}^{AN}} \langle \delta \rangle \top \rightarrow (\neg \text{SD}^{\leq n}(\delta) \leftrightarrow \langle \chi_{\text{SD}}! \rangle^{n+1} \top).$$

Proof. We are going first to prove the theorem by induction.

Let us begin to prove the case $n = 0$. Let M, w be a \mathcal{EDLA} -pointed-model such that $M, w \models \langle \delta \rangle \top$. $M, w \models \neg \text{SD}^{\leq 0}(\delta)$ means that for all $i \in \text{Agt}$, we have $M, w \models \neg \text{SD}^{\leq 0}(\delta_i)$. It is equivalent to: for all $i \in \text{Agt}$, for all $a \in \text{Act}$, $M, w \models \text{SD}^{\leq 0}(i:a) \rightarrow [i:a] \perp$ (indeed, if $a = \delta_i$, we have $M, w \models \neg \text{SD}^{\leq 0}(\delta_i)$ and if $a \neq \delta_i$, we have $M, w \models [i:a] \perp$). So it is equivalent to $M, w \models \chi_{\text{SD}}$ which is equivalent to $M, w \models \langle \chi_{\text{SD}}! \rangle \top$.

Now, we suppose the theorem true for $n - 1$. We suppose that

$$\vdash_{\mathcal{EDLA}^{AN}} \langle \delta \rangle \top \rightarrow (\neg \text{SD}^{\leq n-1}(\delta) \leftrightarrow \langle \chi_{\text{SD}}! \rangle^n \top) \quad (*)$$

Lemma 1. *Let M, w a \mathcal{EDLA} -pointed model. Let $i \in \text{Agt}$. There exists $a \in \text{Act}$ such that $M \models \neg \text{SD}^{\leq 0}(i:a)$.*

Proof. By contradiction. Suppose for all $a \in Act$, we have $M, w \models SD^{\leq 0}(i:a)$. Let $\beta \in \Delta$. Let $a_1 \in Act$ such that $M, w \models \diamond \langle i:a_1 \rangle \top$. We have $M, w \models SD^{\leq 0}(i:a_1)$. We recall the definition of $SD^{\leq 0}(i:a_1)$:

$$SD^{\leq 0}(i:a_1) \stackrel{\text{def}}{=} \langle i:a_1 \rangle \top \rightarrow \bigvee_{a_2 \in Act} (\diamond \langle i:a_2 \rangle \top \wedge \bigwedge_{\delta \in \Delta} (\diamond \langle \delta_{-i} \rangle \rightarrow (\langle \delta_{-i}, i:a_1 \rangle <_i \langle \delta_{-i}, i:a_2 \rangle))).$$

By definition of $SD^{\leq 0}(i:a_1)$, there exists $a_2 \in Act$ such that $\langle \delta_{-i}, i:a_1 \rangle <_i \langle \delta_{-i}, i:a_2 \rangle$.

We have $M, w \models SD^{\leq 0}(i:a_2)$. So we can find a_3 such that $M, w \models \langle \delta_{-i}, i:a_2 \rangle <_i \langle \delta_{-i}, i:a_3 \rangle$. We continue the process and we define a sequence of actions a_1, a_2, a_3, \dots such that for all $j \geq 1$, $M, w \models \langle \delta_{-i}, i:a_j \rangle <_i \langle \delta_{-i}, i:a_{j+1} \rangle$. But Act is finite, so there exists $k > 1$ such that $a_1 = a_k$. By transitivity of $<_i$, we have $M, w \models \langle \delta_{-i}, i:a_1 \rangle <_i \langle \delta_{-i}, i:a_k \rangle$. This is not possible. \square

Lemma 2. For all $i \in Agt$, for all $a \in Act$,

- $$M, w \models \neg SD^{\leq 0}(i:a) \text{ iff } M^{\chi_{SD}}, w \models \diamond \langle i:a \rangle \top.$$

- for all $n \geq 0$, we have:

$$M, w \models \neg SD^{\leq n+1}(i:a) \text{ iff } M^{\chi_{SD}}, w \models \neg SD^{\leq n}(i:a).$$

Proof. • Consider a pointed-model M, w such that $M, w \models \neg SD^{\leq 0}(i:a)$. Thus, by definition of $\neg SD^{\leq 0}(i:a)$, we have $M, w \models \diamond \langle i:a \rangle \top$. According to Lemma 1 we have for all $j \neq i$ the existence of β_j such that $M, w \models \neg SD^{\leq 0}(\beta_j)$. We define δ as $\delta_i = i:a$ and $\delta_j = \beta_j$ for all $j \neq i$. So, by the semantic constraint S4, there exists a point u where $M, u \models \langle \delta \rangle \top$. The world u is not removed by the event $\chi_{SD}!$. Thus, we have $M^{\chi_{SD}}, u \models \langle i:a \rangle \top$.

If $M, w \models SD^{\leq 0}(i:a)$, then all worlds w in which $M, w \models \langle i:a \rangle \top$ are removed because χ_{SD} is false in w . So $M^{\chi_{SD}}, w \not\models \diamond \langle i:a \rangle \top$.

- The second point is the induction case. You can read the first case as the initial case of induction by defining $\neg SD^{\leq -1}(i:a) \stackrel{\text{def}}{=} \diamond \langle i:a \rangle \top$.

Let $n \in \mathbf{N}$. Suppose that we have $M \models \neg SD^{\leq n}(i:a) \text{ iff } M^{\chi_{SD}} \models \neg SD^{\leq n-1}(i:a)$.

We leave to the reader checking that the latter implies $M \models \neg SD^{\leq n+1}(i:a) \text{ iff } M^{\chi_{SD}} \models \neg SD^{\leq n}(i:a)$. \square

Let M, w be a \mathcal{EDCA} -pointed-model such that $M, w \models \langle \delta \rangle \top$. If $M, w \models \neg SD^{\leq n}(\delta)$ then $M, w \models \neg SD^{\leq 0}(\delta)$. So, $M, w \models \langle \chi_{SD}! \rangle \top$, and w remains in $M^{\chi_{SD}}$.

As $M, w \models \neg SD^{\leq n}(\delta)$, the lemma gives us $M^{\chi_{SD}}, w \models \neg SD^{\leq n-1}(\delta)$. Now we are going to apply the induction hypothesis (*). We obtain $M^{\chi_{SD}}, w \models \langle \chi_{SD}! \rangle^n \top$. So $M, w \models \langle \chi_{SD}! \rangle^{n+1} \top$.

If $M, w \models \langle \chi_{SD}! \rangle^{n+1} \top$, we have also $M, w \models \langle \chi_{SD}! \rangle \top$. So w remains in $M^{\chi_{SD}}$. By applying Lemma 2 and induction hypothesis (*), we obtain $M, w \models \neg SD^{\leq n-1}(\delta)$. Finally, $M, w \models \neg SD^{\leq n}(\delta)$. \square

A.10 Proof of Theorem 10.

For all $n \geq 0$, $\vdash_{\mathcal{EDLA}^N} \left(\text{MK}_{\text{Agt}}^n \bigwedge_{i \in \text{Agt}} \text{Rat}_i \right) \rightarrow \langle \chi_{\text{SD}}! \rangle^{n+1} \top$.

Proof. By Theorem 5 and 9. Indeed, let M, w a \mathcal{EDLA} -pointed-model such that $M, w \models \text{MK}_{\text{Agt}}^n \bigwedge_{i \in \text{Agt}} \text{Rat}_i$. There exists $\delta \in \Delta$ such that $M, w \models \langle \delta \rangle \top$. Theorem 5 gives $M, w \models \neg \text{SD}^{\leq n}(\delta)$. Theorem 9 gives $M, w \models \neg \text{SD}^{\leq n}(\delta) \leftrightarrow \langle \chi_{\text{SD}}! \rangle^{n+1} \top$. So, $M, w \models \langle \chi_{\text{SD}}! \rangle^{n+1} \top$. \square

A.11 Note on imperfect information about game structure

The following formula is invalid in \mathcal{EDLA}^* for every $\delta \in \Delta$ and for every $n \in \mathbf{N}$ such that $n > 0$:

$$\left(\left(\text{MK}_{\text{Agt}}^n \bigwedge_{i \in \text{Agt}} \text{Rat}_i \right) \wedge \langle \delta \rangle \top \right) \rightarrow \neg \text{SD}^{\leq n}(\delta).$$

Proof. We suppose $\text{Agt} = \{1, 2\}$ and we exhibit in Figure 2 a \mathcal{EDLA}^* -model M and a world w_1 in M in which for all n , $(\text{MK}_{\{1,2\}}^n \bigwedge_{i \in \{1,2\}} \text{Rat}_i) \wedge \langle 1:\text{main} \rangle \top \wedge \text{SD}^{\leq 1}(1:\text{main})$ is true. \square

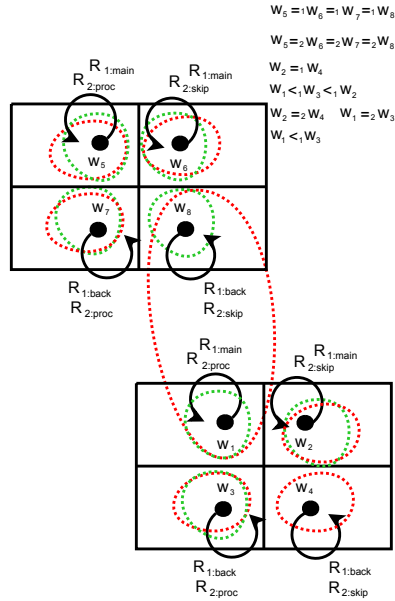


Figure 2: Alarm Game. Again red circles represent epistemic possibility relations for agent 1 whereas green circles represent epistemic possibility relations for agent 2. The two equivalence classes $\sim(w_1) = \{w_1, w_2, w_3, w_4\}$ and $\sim(w_5) = \{w_5, w_6, w_7, w_8\}$ correspond to two different games where agents have different preference ordering over strategy profiles.

SCENARIO DESCRIPTION. We call Alarm Game the scenario represented by the model in Fig. 2. Agent 1 is a thief who intends to burgle agent 2’s apartment. Agent 1 can enter the apartment either by the main door or by the back door (action 1:*main* or action 1:*back*). Agent 2 has two actions available. Either he does nothing (action 2:*skip*) or he follows a security procedure (action 2:*proc*) which consists in locking the two doors and in activating a surveillance camera on the main door. Entering the apartment by the main door when agent 2 does nothing (i.e. the strategy profile $\langle 1:main, 2:skip \rangle$ executed at world w_2) and entering by the back door when agent 2 does nothing (i.e. the strategy profile $\langle 1:back, 2:skip \rangle$ executed at world w_4) are for agent 1 the best situations and are for him equally preferable. Indeed, in both cases agent 1 will successfully enter and burgle the apartment. On the contrary, trying to enter the apartment by the back door when 2 follows the security procedure (i.e. the strategy profile $\langle 1:back, 2:proc \rangle$ executed at world w_3) is for 1 strictly better than trying to enter by the main door when 2 follows the security procedure (i.e. the strategy profile $\langle 1:main, 2:proc \rangle$ executed at world w_1). Indeed, in the former case agent 1 will be simply unable to burgle the apartment, in the latter case not only he will be unable to burgle the apartment but also he will disclose his identity. The two possible situations in which agent 1 does not succeed in burgling the apartment (worlds w_1 and w_3) are equally preferable for agent 2 and are for 2 strictly better than the situations in which agent 1 successfully burgles the apartment (worlds w_2 and w_4).

At world w_1 agent 1 enters by the main door while agent 2 follows the security procedure. This is the only world in the model M in which agent 1 has some uncertainty. Indeed, in this world agent 1 can imagine the alternative game defined by the equivalence class $\sim(w_5) = \{w_5, w_6, w_7, w_8\}$ in which he enters by the back door while agent 2 does nothing (world w_8). We suppose that in such a game, even if agent 2 follows the security procedure, agent 1 will succeed in burgling his apartment. This is the reason why the four strategy profiles $\langle 1:main, 2:skip \rangle$, $\langle 1:back, 2:skip \rangle$, $\langle 1:main, 2:proc \rangle$ and $\langle 1:back, 2:proc \rangle$ are equally preferable for the two agents.