Constructive Knowledge: What Agents Can Achieve under Imperfect Information

Wojciech Jamroga[1] — Thomas Ågotnes[2]

[1] *Department of Informatics,*
Clausthal University of Technology, Germany
wjamroga@in.tu-clausthal.de

[2] *Department of Computer Engineering,*
Bergen University College, Norway
tag@hib.no

ABSTRACT. *We propose a non-standard interpretation of Alternating-time Temporal Logic with imperfect information, for which no commonly accepted semantics has been proposed yet. Rather than changing the semantic structures, we generalize the usual interpretation of formulae in* single *states to* sets *of states. We also propose a new epistemic operator for "practical" or "constructive" knowledge, and we show that the new logic (which we call Constructive Strategic Logic) is strictly more expressive than most existing solutions, while it retains the same model checking complexity. Finally, we study properties of constructive knowledge and other operators in this non-standard semantics.*

KEYWORDS: *Alternating-time Temporal Logic, strategic ability, imperfect information, epistemic logic.*

1. Introduction

Modal logics of strategic ability [ALU 97, ALU 02, PAU 00, PAU 02] form one of the fields where logic and game theory can successfully meet. The logics have clear possible worlds semantics, are axiomatizable, and have some interesting computational properties. Moreover, they are underpinned by a clear and intuitively appealing conceptual machinery for modeling and reasoning about systems that involve multiple autonomous agents. The basic notions, used here, originate from temporal logic (i.e., the logic of time and computation) [PRI 67, EME 90, FIS 06], and classical game theory [NEU 44, NAS 50, OSB 94] which emerged in an attempt to give precise meaning to common-sense notions like choices, strategies, or rationality – and to provide formal models of interaction between autonomous entities. Modal logics that embody basic game theory notions – and at the same time build upon branching-time temporal logics, well known and studied in the context of computational systems – seem a good starting point for investigating multi-agent systems.

Alternating-time Temporal Logic (ATL), proposed in [ALU 97] and further developed in [ALU 98, ALU 02], is probably the most important logic of strategic ability that has emerged in recent years. The key elements of ATL are so called *cooperation modalities* $\langle\langle A\rangle\rangle$, one for each possible set of agents A. Informally, the meaning of $\langle\langle A\rangle\rangle\varphi$ is that the group A has a joint strategy to ensure that, no matter what the other agents do, φ will become true. However, ATL considers only agents that possess perfect information about the current state of the world, and such agents seldom exist in reality. On the other hand, imperfect information and knowledge are addressed in epistemic logic in a natural way [HAL 95]. A combination of ATL and epistemic logic, called *Alternating-time Temporal Epistemic Logic* (ATEL), was introduced in [HOE 02, HOE 03] to enable reasoning about agents acting under imperfect information. Still, it has been pointed out in several places [JAM 03, JAM 04, JON 03, ÅGO 04] that the meaning of ATEL formulae can be counterintuitive. Most importantly, an agent's ability to achieve property φ should imply that the agent has enough control and knowledge to *identify* and *execute* a strategy that enforces φ.

EXAMPLE 1. — Let us consider a variant of the example from [SCH 04]. There is a banker b (who knows the code that opens the safe), and a robber r who does not know the code. The banker can also change the code, and he does so from time to time. If a person is in the vault, and types the code correctly, the safe opens. If incorrect code is typed, the vault door closes, jailing the person inside.

Intuitively, there is no feasible plan for r to quickly open the safe whenever he wants to (unless he threatens or corrupts the banker to reveal the code). Reason: whatever the current code is, the vault looks the same to r, and a sensible plan should specify the same choices in indistinguishable situations (otherwise the plan cannot be executed). On the other hand, there *is* a behavior specification (formally: a function from states to actions) that allows r to rob the bank, and it reads as follows: "if you are outside then enter the vault; if you are inside and the code is 00000 then type 00000; if the code is 00001 then type 00001 etc.". Clearly, not every specification like this makes up a strategy that can be executed by the player. Those that do are sometimes called *uniform strategies*, and are required to prescribe the same choices in indistinguishable states.[1] Unfortunately, ATEL accepts all functions from states to actions as a strategies, which does not blend well with the assumption that agents' knowledge is limited.

It should be noted that it is not always enough to restrict strategies to uniform ones. Consider a situation when b has set the code to 23087 and gone for lunch (so he will not change it again for a while), and r is now standing in front of the safe. Obviously,

1. This very much in agreement with game-theoretical treatment of games with imperfect information. A strategy in such games is a function from *information sets* (i.e., sets of indistinguishable states) to actions.

there is a uniform strategy for r that leads to opening the safe, namely: "type 23087, regardless of anything". The robber even knows that such a successful strategy exists. On the other hand, he does not know which strategy it is (because he does not know what the current state is), and thus he does not have the ability to open the safe for sure. □

Reasoning about the collective abilities of teams requires even more sophisticated concepts.

EXAMPLE 2. — Suppose that, instead of a single robber r, a gang of robbers $r_1, ..., r_n$ is operating. If they can discuss their plans before acting, they can share their individual information about the current state of affairs in order to determine the best strategy (which seems to somehow be related to the notion of *distributed knowledge* from epistemic logic). If they have to coordinate on the fly, without communicating, then it is desirable that they all can separately identify the same winning strategy, and they all know that the others can identify this strategy, and they all know that they all know etc. (which looks very much like *common knowledge*). Thus, there seems to be no single notion of collective knowledge that suffices for all possible scenarios involving collective strategic ability. □

EXAMPLE 3. — Let us also consider an industrial company that wants to start production, and looks for a good strategy when and how it should do it. Such a strategy is feasible if it can be carried out by the company (i.e., by its management and employees). However, it does not have to be *prepared* by members of the company themselves. In many cases, a consulting firm is hired to work out the best plan. Then, it is enough that members of the consulting firm can work out a good strategy which can be executed by the management and employees of the industrial company. □

A number of logics were proposed to capture these, and similar, properties [JAM 03, JAM 04, SCH 04, JON 03, OTT 04, HER 06], yet none of them seems the ultimate definitive solution. Most of the solutions agree that only uniform strategies should be taken into account (cf. Example 1). However, in order to identify a successful strategy, the agents must consider not only the possible courses of action starting from the current (actual) state of the system, but also from states that the agents cannot distinguish from the current one. There are many variants here, especially when group epistemics is concerned, as Examples 2 and 3 demonstrate. The agents may have common, mutual, or distributed knowledge[2] about a strategy being successful, or they may be hinted the right strategy by a distinguished member (the "boss"), a subgroup ("headquarters committee") or even another group of agents ("consulting company") etc. In other words, there are many subtle cases in which the (subjectively possible) initial situations should be represented with different sets of states. Some existing solutions treat only some of the cases (albeit often in an elegant way), while the others offer a very general treatment of the problem at the expense of an overblown logical language (which is by no means elegant). *Our aim is to come up with a logic of ability under imperfect information, which is both general and elegant.* By "general", we

2. See Section 2.2 for precise definitions.

mean that it allows to characterize as many meaningful levels of strategic ability as possible (and at least as many as ATOL [JAM 04]). In particular, it should enable the distinction between various readings of knowing a strategy "de re" and "de dicto" for individual as well as collective players. By "elegant", we mean that it allows us to express various levels of ability by *composition* of epistemic operators with strategic operators, instead of assigning a specialized modality to every conceivable combination.

To achieve this, we build our proposal around new epistemic operators for what we call "practical" or "constructive" knowledge. The idea has been inspired by the tradition of *constructivism* which argues that one must find (or "construct") a mathematical object to prove that it exists [TRO 91]. In the same spirit, agents A *constructively know* that $\langle\langle B\rangle\rangle\varphi$ if they can present a strategy for B that guarantees achieving φ. The logic which we propose in this paper has a fairly non-standard semantic interpretation. We use the same semantic structures that were used before for ATEL, ATOL, ATL$_{ir}$ etc.; however, in our semantics formulae are interpreted over *sets of states* rather than single states. This reflects the intuition that the "constructive" ability to enforce φ means that the agents in question have a single strategy that brings about φ for *all* subjectively possible initial situations – and not merely that a successful strategy exists for *each* initial situation (because those could be different strategies for different situations). To do it in a flexible and general way, the type of satisfaction relation in our proposal forces one to specify the set of initial states explicitly. In consequence, we write $M, Q \models \langle\langle A\rangle\rangle\varphi$ to express the fact that A must have a strategy which is successful for all states in a set of states Q.

Semantically, the constructive knowledge operators yield sets of states for which a single evidence (i.e., a successful strategy) should be presented (instead of checking if the required property holds in each of the states separately, like standard epistemic operators do). For example, $M, q \models \mathbb{K}_a\langle\langle a\rangle\rangle\varphi$ holds iff $\langle\langle a\rangle\rangle\varphi$ is satisfied by M, Q, where Q is the set of states which agent a cannot distinguish from q. We point out that the new operators capture the notion of knowing "de re", while the standard epistemic operators refer to knowing "de dicto". We call the resulting logic *Constructive Strategic Logic* (CSL) to emphasize that, in order to prove $M, Q \models \varphi$ true, one must produce "constructive" evidence for all possible cases in Q, rather than "circumstantial" evidence that deals with every case $q \in Q$ separately.

We begin with a short presentation of Alternating-time Temporal Logic and the attempts that have been made to extend ATL to scenarios with imperfect information (Section 2). In Section 3 we present the main contribution of this paper: a new, non-standard semantics for the logic of strategic ability, imperfect information and knowledge. We show that it is strictly more expressive than the existing solutions, with the possible exception of ETSL (Section 4), while it retains the same model checking complexity (Section 5). Then, in Section 6, we study the properties of constructive knowledge itself. It turns out that, when "standard" knowledge is assumed to be S5, constructive knowledge is KD45. Moreover, a simple syntactical restriction is sufficient to guarantee validity of axiom T for constructive knowledge. In Section 7 we show that standard knowledge is definable from constructive knowledge. We also observe that, when we allow a formula to be interpreted in a set of states, several definitions of negation (corresponding to different ways of quantifying over the set) are possible. We introduce and discuss such alternative negations and related operators. Finally, in Section 8 we investigate the relative expressiveness of some of these operators in detail, and we define a normal form for formulae of our language.

Some preliminary results of this research have been reported in [JAM 05b, JAM 06c].

2. What Agents Can Achieve

Alternating-time Temporal Logic ATL [ALU 97, ALU 98, ALU 02] was introduced by Alur, Henzinger and Kupferman in order to capture properties of *open computational systems* (such as computer networks), where different components can act autonomously. Computations in such systems are effected by the components' combined actions. Alternatively, ATL can be seen as a logic for systems involving multiple agents, that allows one to reason about what agents can achieve in game-like scenarios. As ATL does not include imperfect information in its scope, it can be seen as a logic for reasoning about agents who always have complete knowledge about the current state of affairs.

2.1. *ATL: Ability in Perfect Information Games*

ATL can be understood as a generalization of the branching time temporal logic CTL [CLA 81, EME 90], in which path quantifiers are replaced with so called *cooperation modalities*. The formula $\langle\langle A\rangle\rangle\varphi$, where A is a coalition of agents, expresses that A have a collective strategy to enforce φ. ATL formulae include temporal operators: "$\bigcirc$" ("in the next state"), $\Box$ ("always from now on") and $\mathcal{U}$ ("until"). Operator $\Diamond$ ("now or sometime in the future") can be defined as $\Diamond\varphi \equiv \top\,\mathcal{U}\,\varphi$. Similarly to CTL, every occurrence of a temporal operator is immediately preceded by exactly one cooperation modality.[3] The broader language of ATL*, in which no such restriction is imposed, is not discussed in this paper.

Formally, the recursive definition of ATL formulae is:

$$\varphi ::= p \mid \neg\varphi \mid \varphi \wedge \varphi \mid \langle\langle A\rangle\rangle\bigcirc\varphi \mid \langle\langle A\rangle\rangle\Box\varphi \mid \langle\langle A\rangle\rangle\varphi\,\mathcal{U}\,\varphi$$

where A is a set of agents. Example ATL properties are: $\langle\langle jamesbond\rangle\rangle\Diamond\mathsf{win}$ (James Bond has an infallible plan to eventually win), and $\langle\langle jamesbond, bondsgirl\rangle\rangle\mathsf{fun}\,\mathcal{U}\,\mathsf{shot\text{-}at}$ (Bond and his current girlfriend have a collective way of having fun until someone shoots at them).

3. The logic to which such a syntactic restriction applies is sometimes called *"vanilla"* ATL (resp. "vanilla" CTL etc.).

A number of semantics have been defined for ATL, most of them equivalent [GOR 01, GOR 04]. In this paper, we use a variant of *concurrent game structures* (CGSs) as models. A CGS is a tuple $M = \langle \mathbb{A}\mathrm{gt}, St, \Pi, \pi, Act, d, o \rangle$ which includes a nonempty finite set of all agents $\mathbb{A}\mathrm{gt} = \{1, ..., k\}$, a nonempty set of states St, a set of atomic propositions Π, a valuation of propositions $\pi : St \to \mathcal{P}(\Pi)$, and a set of (atomic) actions Act. Function $d : \mathbb{A}\mathrm{gt} \times St \to (\mathcal{P}(Act) \setminus \emptyset)$ defines nonempty sets of actions available to agents at each state, and o is a (deterministic) transition function that assigns the outcome state $q' = o(q, \alpha_1, \ldots, \alpha_k)$ to state q and a tuple of actions $\langle \alpha_1, \ldots, \alpha_k \rangle$, $\alpha_i \in d(i, q)$, that can be executed by $\mathbb{A}\mathrm{gt}$ in q. A *strategy* s_a of agent a is a conditional plan that specifies what a is going to do for every possible situation: $s_a : St \to Act$ such that $s_a(q) \in d(a, q)$. A *collective strategy* S_A for a group of agents A is a tuple of strategies, one per agent from A.

REMARK 4. — This is a deviation from the original semantics of ATL [ALU 97, ALU 98, ALU 02], where strategies assign agents' choices to *sequences* of states, which suggests that agents can by definition recall the whole history of each game. Both types of strategies yield equivalent semantics for "vanilla" ATL, but the choice of one or the other notion of strategy does affect the semantics of the full ATL* and most ATL variants for games with imperfect information [SCH 04]. The main reason why we use "memoryless" strategies here is that model checking strategic abilities of agents with perfect recall and imperfect information is believed to be undecidable (cf. Section 2.10). $\square$

A *path* Λ in model M is an infinite sequence of states that can be effected by subsequent transitions, and refers to a possible course of action (or a possible computation) that may occur in the system; by $\Lambda[i]$, we denote the ith position on path Λ. Function $out(q, S_A)$ returns the set of all paths that may result from agents A executing strategy S_A from state q onward:

$out(q, S_A) = \{\lambda = q_0 q_1 q_2 ... \mid q_0 = q$ and for every $i = 1, 2, ...$ there exists a tuple of agents' decisions $\langle \alpha_1, ..., \alpha_k \rangle$ such that $\alpha_a = S_A(a)(q_{i-1})$ for each $a \in A$, and $\alpha_a \in d(a, q_{i-1})$ for each $a \notin A$, and $o(q_{i-1}, \alpha_1, ..., \alpha_k) = q_i\}$.

Informally speaking, $M, q \models \langle\langle A \rangle\rangle \varphi$ iff there is a collective strategy S_A such that φ holds for every $\Lambda \in out(q, S_A)$. Formally, the semantics of ATL formulae can be given via the following clauses:

$M, q \models p$ iff $p \in \pi(q)$ (for $p \in \Pi$);

$M, q \models \neg\varphi$ iff $M, q \not\models \varphi$;

$M, q \models \varphi \wedge \psi$ iff $M, q \models \varphi$ and $M, q \models \psi$;

$M, q \models \langle\langle A \rangle\rangle \bigcirc \varphi$ iff there is a collective strategy S_A such that, for every $\Lambda \in out(q, S_A)$, we have $M, \Lambda[1] \models \varphi$;

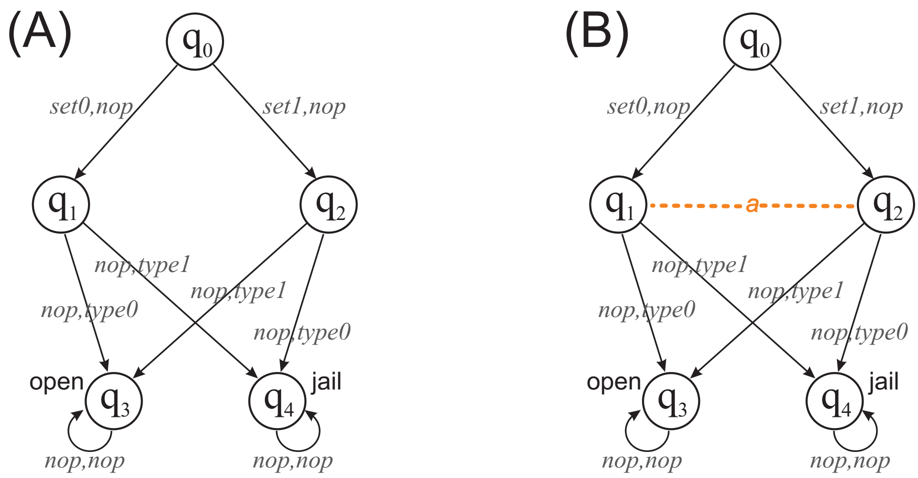

Figure 1. *The banker and the robber: (A) concurrent game structure M_1 for the perfect information case; (B) concurrent epistemic game structure M_2 for the imperfect information case.*

$M, q \models \langle\!\langle A \rangle\!\rangle \Box \varphi$ iff there exists S_A such that, for every $\Lambda \in out(q, S_A)$, we have $M, \Lambda[i]$ for every $i \geq 0$;

$M, q \models \langle\!\langle A \rangle\!\rangle \varphi \mathcal{U} \psi$ iff there exists S_A such that for every $\Lambda \in out(q, S_A)$ there is an $i \geq 0$, for which $M, \Lambda[i] \models \psi$, and $M, \Lambda[j] \models \varphi$ for every $0 \leq j < i$.

EXAMPLE 5. — Consider a simple formalization of the scenario from Example 1, presented in Figure 1A. First, the banker sets the code to either 0 or 1, and walks away. Then, the robber tries to open the safe by typing a number. If the number is correct, the safe opens; otherwise the robber is jailed in the vault. Nodes in the graph represent global states of the system. Transitions are labeled by combinations of actions from b, r, and nop stands for "no operation" or "do nothing" (formally, nop is just another action).

ATL addresses agents with perfect information, so the following naturally holds: $M_1, q_0 \models \langle\!\langle r \rangle\!\rangle \Diamond \mathsf{open}$. The right strategy for the robber is to wait first to see which code is set, and then to type the appropriate number: $s_r(q_0) = nop$, $s_r(q_1) = type0$, and $s_r(q_2) = type1$. □

REMARK 6. — Concurrent game structures model actions as abstract atomic entities, with no underlying structure. This is not necessarily satisfying for everyone's purposes. One may, e.g., want to define actions as state transformations that can occur in the system, like in models of dynamic logic [HAR 00]; STIT models assign actions/choices with even more complicated conceptual structure [BEL 88]. We choose, after [ALU 02, SCH 04, JAM 04, ÅGO 06], to avoid the discussion on the nature of actions, and make the simplifying assumption that actions are identified by unique names. Note that this approach follows closely the tradition of game theory, and the definition of an extensive game form in particular [OSB 94]. □

One of the most appreciated features of ATL is its model checking complexity – linear in the number of transitions in the model and the length of the formula. The model checking problem is, given a formula φ and a model M with a state q, to decide whether $M, q \models \varphi$ or not.

PROPOSITION 7 ([ALU 02]). — *The* ATL *model checking problem is* PTIME*-complete, and can be done in time $\mathbf{O}(ml)$, where m is the number of transitions in the model and l is the length of the formula.*

Note that the complexity is measured, as usual, as a function of the size of the input. Thus, while infinite concurrent game structures make perfect sense in general, they cannot be subjects of model checking unless represented in a finite way.

REMARK 8. — The result in Proposition 7 does not seem so unambiguously optimistic after a closer inspection, i.e., when we measure the size of models in the number of states, actions and agents [JAM 05a, LAR 06, JAM 07], or when we represent systems with so called concurrent programs [HOE 06]. This remark is only meant as a note of warning; such a detailed complexity analysis for the logics of ability under imperfect information (that are the main topic here) is beyond the scope of this paper. □

2.2. *ATL with Epistemic Logic*

ATL is unrealistic in a sense: real-life agents seldom possess complete information about the current state of the world. On the other hand, imperfect information and knowledge are handled in epistemic logic in a natural way. A combination of ATL and epistemic logic, called *Alternating-time Temporal Epistemic Logic* (ATEL), was introduced by van der Hoek and Wooldridge in [HOE 02, HOE 03] to enable reasoning about agents acting under imperfect information.

ATEL enriches the picture with an epistemic component, adding to ATL operators for representing agents' knowledge: $K_a\varphi$ reads as "agent a knows that φ". Additional operators $E_A\varphi$, $C_A\varphi$, and $D_A\varphi$, where A is a set of agents, refer to *mutual knowledge* ("everybody knows"), *common knowledge*, and *distributed knowledge* among the agents from A. Thus, $E_A\varphi$ means that every agent in A knows that φ holds, while $C_A\varphi$ means not only that the agents from A know that φ, but they also know that they know it, and that they know that they know that they know it, etc. The distributed knowledge modality $D_A\varphi$ expresses that if the agents could share their individual information they would be able to recognize that φ.

Models for ATEL extend concurrent game structures with epistemic accessibility relations $\sim_1, ..., \sim_k \subseteq Q \times Q$ (one per agent) for modeling agents' uncertainty.[4] We

4. The relations are assumed to be equivalences.

will call such models *concurrent epistemic game structures* (CEGS) in the rest of the paper. Agent a's epistemic relation is meant to encode a's inability to distinguish between the (global) system states: $q \sim_a q'$ means that, while the system is in state q, agent a cannot determine whether it is in q or q'. Then, the semantics of K_a is defined as:

$M, q \models K_a\varphi$ iff $M, q' \models \varphi$ for every q' such that $q \sim_a q'$.

EXAMPLE 9. — Consider model M_2 from Figure 1B, with the epistemic link between states q_1 and q_2 (we omit the reflexive indistinguishability links from q_0 to q_0, q_1 to q_1 etc. to make the figure easier to read). This time, the scenario is more realistic: the robber does not know the correct code. Thus, one cannot expect him to be able to open the safe. Still, in ATEL, we have that $M_2, q_0 \models \langle\langle r\rangle\rangle\Diamond\mathsf{open}$; the same (non-uniform) strategy as in Example 5 can be used to demonstrate this. Moreover, we have even that $M_2, q_0 \models K_r\langle\langle r\rangle\rangle\Diamond\mathsf{open}$: using knowledge operators does not help, because cooperation modalities are still underpinned by a notion of strategy that does not agree with imperfect information of agents. This is a fundamental problem with ATEL, which we discuss briefly in Section 2.3. □

Relations $\sim^E_A$, $\sim^C_A$ and $\sim^D_A$, used to model group epistemics, are derived from the individual relations of agents from A. First, $\sim^E_A$ is the union of relations $\sim_a$, $a \in A$. Next, $\sim^C_A$ is defined as the transitive closure of $\sim^E_A$. Finally, $\sim^D_A$ is the intersection of all the $\sim_a$, $a \in A$. The semantics of group knowledge can be defined as below (for $\mathcal{K} = C, E, D$):

$M, q \models \mathcal{K}_A\varphi$ iff $M, q' \models \varphi$ for every q' such that $q \sim^{\mathcal{K}}_A q'$.

Note that $K_a \equiv C_{\{a\}} \equiv E_{\{a\}} \equiv D_{\{a\}}$, so individual knowledge operators K_a are actually redundant.

In order to explore the subtleties of collective play, we extend the model from Figure 1B slightly: the pattern is the same, but more complex properties can be demonstrated.

EXAMPLE 10 (GAMBLING ROBOTS). — Two robots (a and b) play a simple card game. The deck consists of Ace, King and Queen (A, K, Q). Normally, it is assumed that A is the best card, K the second best, and Q the worst; so, A beats K and Q, K beats Q, and Q beats no card. At the beginning of the game, the "environment" agent deals a random card to both robots (actions $deal_{AK}, deal_{AQ}, \dots, deal_{QK}$), so that each player can see his own card, but he does not know the card of the other player. Then robot a can choose to exchange his card for the one remaining in the deck (action $exch$), or he can keep the current one ($keep$). At the same time, robot b can change the priorities of the cards to a Rochambeau-like game (that is, A still beats K and K beats Q, but Q becomes better than A), or he can do nothing (nop), i.e. leave the priorities unchanged. If a has a better card than b after that, then a win is scored, otherwise the game ends in a "losing" state.

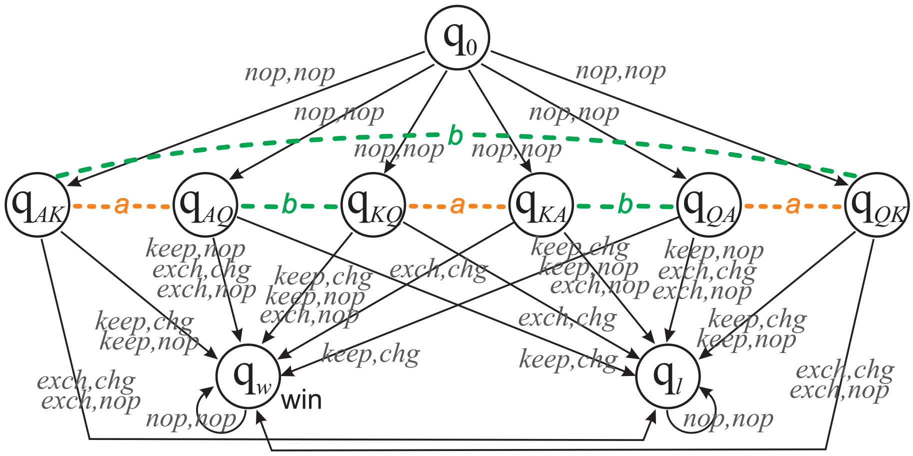

Figure 2. *Gambling Robots game. Nodes represent global states of the system; arrows denote transitions, labeled with combinations of actions from all the agents. Dashed lines indicate states that are indistinguishable for respective agents. Actions of the environment agent are omitted from the picture to make it easier to read. As epistemic relations are by definition reflexive, we omit reflexive epistemic links too.*

A CEGS for the game is shown in Figure 2; we will refer to the model as M_3 throughout the rest of the paper. State q_0 represents the situation before, and states $q_{AK}, \ldots, q_{QK}$ after the cards have been dealt (each $q_{c_1 c_2}$ stands for the situation when a has got card c_1, and b has got card c_2). Actions of the environment are omitted from the figure for the sake of readability. Similarly to the previous example, $M_3, q_0 \models \langle\langle a, b\rangle\rangle\Diamond \mathsf{win}$ (and even $M_3, q_0 \models C_{\{a,b\}}\langle\langle a, b\rangle\rangle\Diamond \mathsf{win}$), but there is no *uniform* strategy to achieve this: in order to win, a must *exchange* his card in state q_{QK}, so he must *exchange* his card in q_{QA} too (if we require uniformity), and playing $exch$ in q_{QA} leads to the losing state. So, again, we have $\langle\langle a, b\rangle\rangle\Diamond \mathsf{win}$, although intuitively $\{a, b\}$ have no feasible way of ensuring a win. □

2.3. *Problems with ATEL*

It has been pointed out in several places that the meaning of ATEL formulae can be counterintuitive [JAM 03, JAM 04, JON 03]. Most importantly, one would expect that an agent's ability to achieve property φ should imply that the agent has enough control and knowledge to *identify* and *execute* a strategy that enforces φ (cf. also [SCH 04]). ATEL adds to ATL the vocabulary of epistemic logic; still, in ATEL the strategic and epistemic layers are combined as if they were independent. They should be – if we do not ask whether the agents in question are able to identify and execute their strategies. They should not if we want to interpret strategies as *executable plans*, about which the agents *know* that they guarantee achieving the goal.

First of all, executable plans should not specify different actions in indistinguishable states. Most (if not all) current approaches to strategic ability under imperfect information [JAM 04, SCH 04, JON 03, OTT 04, HER 06], agree with the postulate from [JAM 03] that only *uniform* strategies should be considered in the semantics of $\langle\!\langle A \rangle\!\rangle$. Formally, strategy s_a is *uniform* iff $q \sim_a q'$ implies that $s_a(q) = s_a(q')$; a collective strategy S_A is uniform iff it consists of only uniform individual strategies. In other words, agents make choices with respect to their *local* (epistemic) states rather than global states of the system. Agents are assumed to know their available actions (i.e., the choices open to them), so they must have the same choices in indistinguishable states. That is, from now on we consider only models in which $q \sim_a q'$ implies $d(a, q) = d(a, q')$.

Second, it was suggested in [JAM 04] that, when reasoning about what an agent can *enforce*, it seems more appropriate to require the agent to *know his winning strategy* rather than to know only *that such a strategy exists*. This problem is closely related to the distinction between knowledge *de re* and knowledge *de dicto*, well known in the philosophy of language [QUI 56], as well as research on the interaction between knowledge and action [MOO 85, MOR 91, WOO 00]. One can naturally distinguish at least four different levels of strategic ability (cf. [JAM 04]):

1) Agent a has a strategy *"de re"* to enforce φ, i.e., he has an executable winning strategy and knows the strategy (he "knows how to play");

2) Agent a has a strategy *"de dicto"* to enforce φ (i.e., he knows only that *some* executable winning strategy is available);

3) Agent a has an executable strategy to enforce φ (but not necessarily even knows about it);

4) Agent a may *happen* to behave in such a way that φ is enforced. However, the behavior can have no executable specification (i.e., there might be no uniform strategy that describes it).

Obviously, (1) $\Rightarrow$ (2) $\Rightarrow$ (3) $\Rightarrow$ (4), but not the other way around. We do think that *all* of these concepts can be useful for reasoning about strategic ability under imperfect information. However, we believe that (1) is particularly important and natural. Unfortunately, ATEL enables to express only ability of type (4), as Example 9 showed. Several variations on "ATL with imperfect information" have been proposed as alternatives, yet none of them seems the ultimate definitive solution. We summarize the most important proposals in the following sections.

2.4. *First Try: ATEL with Uniform Strategies*

The first attempt to cope with these problems was presented in [JAM 03], where it was proposed that only uniform strategies should be used in the semantics of cooperation modalities. "Uniform ATEL" (U-ATEL) captures ability of type (2) and (3): $\langle\!\langle a \rangle\!\rangle\varphi$ says that a has a uniform strategy to achieve φ, and $K_a\langle\!\langle a \rangle\!\rangle\varphi$ denotes having a strategy "de dicto". However, *knowing how to play* still cannot be expressed.

EXAMPLE 11. — Consider model M_2 from Figure 1B, and assume that q_1 is the current state. The robber does have a uniform strategy to open the safe in one step (play *type0* at q_1 and q_2, and *nop* elsewhere), and indeed $M_2, q_1 \models \langle\langle r\rangle\rangle \bigcirc \mathsf{open}$. He also knows that such a strategy is available, and we have $M_2, q_1 \models K_r \langle\langle r\rangle\rangle \bigcirc \mathsf{open}$ (in every state q such that $q_1 \sim_r q$, $M, q \models \langle\langle r\rangle\rangle \bigcirc \mathsf{open}$). Still, the robber does not know how to play in q_1 to achieve open, and this property has no U-ATEL counterpart. Note also that $M_2, q_0 \models \neg\langle\langle r\rangle\rangle \Diamond \mathsf{open} \wedge \langle\langle \emptyset\rangle\rangle \bigcirc \langle\langle r\rangle\rangle \Diamond \mathsf{open}$, and $M_2, q_0 \models \neg K_r \langle\langle r\rangle\rangle \Diamond \mathsf{open} \wedge K_r \langle\langle \emptyset\rangle\rangle \bigcirc K_r \langle\langle r\rangle\rangle \Diamond \mathsf{open}$ (the robber has no strategy to open the safe in q_0, but he can simply wait a moment, and he will magically get one), which suggests that one should be careful when talking about abilities of type (2) and (3).

Likewise, for the gambling robots we have $M_3, q_0 \models \neg\langle\langle a\rangle\rangle \Diamond \mathsf{win}$, and even $M_3, q_0 \models \neg\langle\langle a, b\rangle\rangle \Diamond \mathsf{win}$ (see Section 2.2). On the other hand, $M_3, q_{AK} \models \langle\langle a\rangle\rangle \Diamond \mathsf{win} \wedge K_a \langle\langle a\rangle\rangle \Diamond \mathsf{win}$. □

2.5. *Aggregating Initial States: "Feasible ATEL"*

"Feasible ATEL" [JON 03], which we will sometimes call F-ATEL, is an update of ATEL, in which the "perfect information" cooperation modalities are kept, but the language is extended with new modalities: $\langle\langle A\rangle\rangle^f$, $\langle\langle A\rangle\rangle^f_E$, $\langle\langle A\rangle\rangle^f_C$, $\langle\langle A\rangle\rangle^f_{K_a}$ and $\langle\langle A\rangle\rangle^f_{M_a}$, that represent agents' ability to find a suitable uniform strategy, with the semantics summarized below:

$M, q \models \langle\langle A\rangle\rangle^f \bigcirc \varphi$ iff there is a uniform collective strategy S_A such that, for every $\Lambda \in out(q, S_A)$, we have $M, \Lambda[1] \models \varphi$.
For $\langle\langle A\rangle\rangle^f \Box \varphi$ and $\langle\langle A\rangle\rangle^f \varphi \,\mathcal{U}\, \psi$: analogously;

$M, q \models \langle\langle A\rangle\rangle^f_E \bigcirc \varphi$ iff there is a uniform collective strategy S_A such that, for every q' such that $q \sim^E_A q'$, and for every $\Lambda \in out(q', S_A)$, we have $M, \Lambda[1] \models \varphi$.
For $\langle\langle A\rangle\rangle^f_E \Box \varphi$ and $\langle\langle A\rangle\rangle^f_E \varphi \,\mathcal{U}\, \psi$: analogously;

$M, q \models \langle\langle A\rangle\rangle^f_C \bigcirc \varphi$ iff there is a uniform collective strategy S_A such that, for every q' such that $q \sim^C_A q'$, and for every $\Lambda \in out(q', S_A)$, we have $M, \Lambda[1] \models \varphi$.
For $\langle\langle A\rangle\rangle^f_C \Box \varphi$ and $\langle\langle A\rangle\rangle^f_C \varphi \,\mathcal{U}\, \psi$: analogously;

$M, q \models \langle\langle A\rangle\rangle^f_{K_a} \bigcirc \varphi$ iff there is a uniform collective strategy S_A such that, for every q' such that $q \sim_a q'$, and every $\Lambda \in out(q', S_A)$, we have $M, \Lambda[1] \models \varphi$.
For $\langle\langle A\rangle\rangle^f_{K_a} \Box \varphi$ and $\langle\langle A\rangle\rangle^f_{K_a} \varphi \,\mathcal{U}\, \psi$: analogously;

$M, q \models \langle\langle A\rangle\rangle^f_{M_a} \bigcirc \varphi$ iff there is a uniform collective strategy S_A and state q' with $q \sim_a q'$, such that, for every $\Lambda \in out(q', S_A)$, we have $M, \Lambda[1] \models \varphi$.
For $\langle\langle A\rangle\rangle^f_{M_a} \Box \varphi$ and $\langle\langle A\rangle\rangle^f_{M_a} \varphi \,\mathcal{U}\, \psi$: analogously;

The idea of cooperation modalities with subscripts that indicate the epistemic "mode", in which coalition A can identify their winning strategy, was further developed in the logic of ATOL, which we present in Section 2.6.

We note that "Uniform ATEL" can be seen as a subset of "Feasible ATEL", as the meaning of $\langle\langle A\rangle\rangle\varphi$ proposed in [JAM 03] is, for agents playing memoryless strategies, equivalent to $\langle\langle A\rangle\rangle^f\varphi$ from [JON 03].

2.6. *Going for Expressive Power: ATOL*

Alternating-time Temporal Observational Logic (ATOL), proposed in [JAM 04], follows the same perspective, but it offers a richer language of strategic operators to express subtle differences between various kinds of collective abilities of teams. In this paper, we use the notation proposed in [JAM 05c]. The informal meaning of $\langle\langle A\rangle\rangle_{\mathcal{K}(\Gamma)}\varphi$ is: "group A has a (memoryless uniform) strategy to enforce φ, and agents Γ can identify the strategy as successful for A in the epistemic sense $\mathcal{K}$". For instance, $M, q \models \langle\langle A\rangle\rangle_{D(\Gamma)}\varphi$ iff there is S_A such that, for every q' with $q \sim^D_\Gamma q'$, and every $\Lambda \in out(q', S_A)$, we have that φ is true for Λ.

Formally, let $\mathcal{K} = E, C, D$. The semantics of the enhanced cooperation modalities can be defined as follows:

$M, q \models \langle\langle A\rangle\rangle_{\mathcal{K}(\Gamma)}\bigcirc\varphi$ iff there is a collective memoryless uniform strategy S_A such that, for every q' with $q \sim^{\mathcal{K}}_\Gamma q'$, and every $\Lambda \in out(q', S_A)$, we have that $M, \lambda[1] \models \varphi$.
For $\langle\langle A\rangle\rangle_{\mathcal{K}(\Gamma)}\Box\varphi$ and $\langle\langle A\rangle\rangle_{\mathcal{K}(\Gamma)}\varphi\,\mathcal{U}\,\psi$: analogously.

EXAMPLE 12. — Coming back to our gambling robots, it is easy to see that $M_3, q_0 \models \neg\langle\langle a\rangle\rangle_{K(a)}\Diamond\mathsf{win}$, because, for every a's (uniform) strategy, if it guarantees a win in e.g. state q_{AK} then it fails in q_{AQ} (and similarly for other pairs of indistinguishable states). Let us also observe that $M_3, q_0 \models \neg\langle\langle a, b\rangle\rangle_{E(\{a,b\})}\Diamond\mathsf{win}$: in order to win, a must *exchange* his card in state q_{QK}, so he must *exchange* his card in q_{QA} too (by uniformity), and playing $exch$ in q_{QA} leads to the losing state. On the other hand, $M_3, q_{AQ} \models \langle\langle a, b\rangle\rangle_{E(\{a,b\})}\bigcirc\mathsf{win}$ (a winning strategy: $s_a(q_{AK}) = s_a(q_{AQ}) = s_a(q_{KQ}) = keep$, $s_b(q_{AQ}) = s_b(q_{KQ}) = s_b(q_{AK}) = nop$; q_{AK}, q_{AQ}, q_{QK} are the states that must be considered by a and b in q_{AQ}). Still, $M_3, q_{AK} \models \neg\langle\langle a, b\rangle\rangle_{E(\{a,b\})}\bigcirc\mathsf{win}$.

ATOL allows us to express other ways of identifying a winning strategy too: we have that $M_3, q_{AK} \models \langle\langle a, b\rangle\rangle_{D(\{a,b\})}\bigcirc\mathsf{win} \wedge \langle\langle a, b\rangle\rangle_{K(a)}\bigcirc\mathsf{win}$ (the robots can identify the strategy if they share their views of the world; also, a can be the "boss" who points out the strategy), and $M_3, q_{AQ} \models \neg\langle\langle a, b\rangle\rangle_{C(\{a,b\})}\bigcirc\mathsf{win}$ (despite both a, b knowing the winning strategy, they do not have common knowledge about it). □

ATOL is quite expressive. However, it does not allow for combination of strategic ability and *arbitrary* epistemic modes – the operators $\langle\langle A\rangle\rangle_{\mathcal{K}(\Gamma)}$ are fixed by taking $\mathcal{K} \in \{C, E, D\}$. For example, $\langle\langle A\rangle\rangle_{E_A E_A}\varphi$ is *not* a well formed ATOL formula – although it is easy to give an interpretation of such a formula in a similar manner to the other ATOL operators. Furthermore, the trebly parameterized cooperation modalities are rather baroque.

2.7. *Elegance and Simplicity: ATL$_{ir}$*

Schobbens [SCH 04] approached the problem of combining strategies with uncertainty on a more abstract level. He suggested that it makes sense to talk about agents with perfect as well as imperfect information on one hand, and perfect vs. imperfect recall on the other – and that these two fundamental semantic choices are orthogonal. This gives rise to four different logics of strategic ability: ATL$_{IR}$ (for perfect **I**nformation and perfect **R**ecall, i.e. the original ATL), ATL$_{iR}$ (for imperfect **i**nformation and perfect **R**ecall), etc. As we focus on imperfect information and memoryless strategies in this paper, the logic of ATL$_{ir}$ is most interesting for us.

Informally, $\langle\langle A\rangle\rangle_{ir}\varphi$ holds in M, q iff there is a uniform collective strategy S_A such that, for every agent $a \in A$, state q' with $q \sim_a q'$, and path $\Lambda \in out(q', S_A)$, we have that φ is true for Λ. In other words, there is a strategy such that *everybody in A knows* that executing this strategy will bring about φ. Formally:

$M, q \models \langle\langle A\rangle\rangle_{ir}\bigcirc\varphi$ iff there is a uniform collective strategy S_A such that, for every $a \in A$, q' such that $q \sim_A^E q'$, and path $\Lambda \in out(S_A, q')$, we have $M, \lambda[1] \models \varphi$. For $\langle\langle A\rangle\rangle_{ir}\Box\varphi$ and $\langle\langle A\rangle\rangle_{ir}\varphi\,\mathcal{U}\,\psi$: analogously.

EXAMPLE 13. — For our gambling robots, we get e.g. that: $M_3, q_0 \models \neg\langle\langle a\rangle\rangle_{ir}\Diamond\mathsf{win}$, $M_3, q_0 \models \neg\langle\langle a, b\rangle\rangle_{ir}\Diamond\mathsf{win}$, $M_3, q_{AQ} \models \langle\langle a, b\rangle\rangle_{ir}\bigcirc\mathsf{win}$, and $M_3, q_{AK} \models \neg\langle\langle a, b\rangle\rangle_{ir}\bigcirc\mathsf{win}$. □

Note that $\langle\langle A\rangle\rangle_{ir}\Phi$ is equivalent to the "Feasible ATEL" formula $\langle\langle A\rangle\rangle_E^f\Phi$, and the ATOL formula $\langle\langle A\rangle\rangle_{E(A)}\Phi$. Moreover, it is not possible to express in ATL$_{ir}$ that A have common knowledge about the successful strategy, or that they are able to identify it if they share their information etc. On the other hand, ATL$_{ir}$ stands out among the existing proposals for its simplicity and conceptual clarity, and can be treated as the "core", minimal ATL-based language for ability under imperfect information.

The following proposition sums up some of the results presented in [SCH 04, JAM 06b, JAM 04]:

PROPOSITION 14. — *Model checking "Feasible* ATEL*",* ATL$_{ir}$ *and* ATOL *is* $\mathbf{\Delta_2^P}$*-complete in the number of transitions (and epistemic links) in the model, and the length of the formula.*

In Section 3, we will propose Constructive Strategic Logic (CSL) which strictly subsumes ATOL, while sharing (in our opinion) the elegance of ATL$_{ir}$, and model checking complexity of all of the approaches discussed above. The main idea behind CSL is that we would like to express various levels of ability with combinations of *some* kind of epistemic operators with *some* kind of cooperation modalities. Before we present our proposal, we want to mention two logics that, to a limited extent, have achieved a similar trait. The logics are briefly presented in Sections 2.8 and 2.9.

2.8. *Abilities of Rational Players: ETSL*

Epistemic Temporal Strategic Logic [OTT 04] digs deeper in the repository of game theory, and focuses on the concept of *undominated strategies*. Its variant of the cooperation modalities has a different flavor than the ones from ATL, ATEL, ATOL etc. In a way, $\langle\!\langle A \rangle\!\rangle\varphi$ in ETSL can be summarized as: "if A play *rationally* to achieve φ (meaning: they never play a dominated strategy), they will achieve φ".

ETSL is underpinned by several interesting concepts. Unfortunately, its original semantics from [OTT 04] comes with a plethora of auxiliary functions and definitions (and a couple of omissions), which make it rather hard to read. Moreover, the semantics is defined only for *finite turn-based acyclic* game models, and the satisfaction relation refers not only to models and states (respectively paths), but also to a fixed strategy $S_{\mathbb{A}\mathrm{gt}}$ (assumed to represent the current strategies of all agents). It has been shown in [JAM 06a], that the semantics can be extended to concurrent epistemic game structures, and given in a more compact way. Moreover, for "vanilla" ETSL formulae,[5] it can be given via standard semantic clauses for state formulae.

Let M be a CEGS. First, we define the notion of domination as follows. Let $\Phi \equiv \bigcirc\psi$, $\Box\psi$, or $\psi_1 \mathcal{U} \psi_2$, where ψ, ψ_1, ψ_2 are "vanilla" ETSL formulae. Moreover, let $|\Phi|$ denote the set of paths for which Φ holds; formally, $|\bigcirc\psi| = \{\Lambda \mid M, \Lambda[1] \models \psi\}$, $|\Box\psi| = \{\Lambda \mid \forall_i M, \Lambda[i] \models \psi\}$, and $|\psi_1 \mathcal{U} \psi_2| = \{\Lambda \mid \exists_i (M, \Lambda[i] \models \psi_2 \wedge \forall_{0 \le j < i} M, \Lambda[j] \models \psi_1\}$. Then, *strategy S_A dominates strategy T_A wrt. M, q, and Φ* iff both of the following conditions hold:

1) for every q' with $q \sim_A^E q'$: if $out(q', T_A) \subseteq |\Phi|$ then also $out(q', S_A) \subseteq |\Phi|$;
2) there is q' such that $q \sim_A^E q'$, and $out(q', S_A) \subseteq |\Phi|$, and $out(q', T_A) \not\subseteq |\Phi|$.

Strategy S_A is *undominated* wrt. M, q, Φ iff there is no strategy that dominates S_A wrt M, q, Φ.

Now the semantics of $\langle\!\langle A \rangle\!\rangle$ in ETSL can be expressed entirely in terms of models and their states:

$M, q \models \langle\!\langle A \rangle\!\rangle \bigcirc\varphi$ iff for every strategy S_A, undominated wrt $M, q, \bigcirc\varphi$, and every $\Lambda \in out(q, S_A)$, we have that $M, \Lambda[1] \models \varphi$.
For $\langle\!\langle A \rangle\!\rangle\Box\varphi$ and $\langle\!\langle A \rangle\!\rangle\varphi \mathcal{U} \psi$: analogously.

The relationship between ETSL and Constructive Strategic Logic is briefly discussed in Section 4.4. We conjecture that neither of them subsumes the other, but there are several interesting associations. The most interesting feature of ETSL is perhaps the fact that, by combining standard epistemic operators and its non-standard cooperation modalities, we can capture "knowing how to play" for individual agents (although this does not extend to collective agents), see [JAM 06a] or Section 4.4 for more details.

5. I.e., formulae in which every temporal operator is preceded by exactly one cooperation modality.

2.9. *Explicit Actions: ATEL-A*

In AT(E)L, it is not possible to refer directly to particular actions in the logical language. For example, it is not possible to express the fact that "if agent i chooses action α, then formula φ will necessarily be true in the next moment". ATEL-A [ÅGO 06] allows such expressions by introducing names of actions, in addition to names of agents, inside cooperation modalities. For instance, the above expression can be written as $\langle\langle\alpha_i\rangle\rangle\bigcirc\varphi$. This makes it possible to capture the levels of ability, discussed in Section 2.3, in the limited case of properties that can be achieved in one step:

(4), (3) $\langle\langle i\rangle\rangle\bigcirc\varphi$: agent i may behave in such a way that φ is enforced next. Note that there is no difference between (4) and (3) when we only talk about the next state – then uniformity does not play any role;

(2) $K_i\langle\langle i\rangle\rangle\bigcirc\varphi$: agent i has a strategy "de dicto" to enforce φ next;

(1) $\bigvee_{\alpha\in Act} K_i\langle\langle\alpha_i\rangle\rangle\bigcirc\varphi$: agent i has a strategy "de re" to enforce φ next.

Because of explicit actions, ATEL-A is not directly comparable to the logics considered in this paper, and we will not discuss ATEL-A further.

2.10. *Other Possibilities*

In the original formulation of ATL, agents were assumed to have perfect recall of the game, in the sense that they could base their decisions on *sequences* of states rather than on single states. Variants of ATL for perfect recall and imperfect information have also been considered, cf. ATL$_{iR}$ [SCH 04] and ATEL-R* [JAM 04]. However, as agents seldom have unlimited memory, and logics of strategic ability with imperfect information and perfect recall are believed to have undecidable model checking [ALU 02, SCH 04], we do not investigate this variant of ability here.

Yet another, very recent, proposal [HER 06] approaches the problem of strategic abilities under imperfect information within the framework of STIT (the logic of *seeing to it that*). STIT shares many similarities with ATL, but it comes from a different tradition, and its technical formulation is markedly different from that of ATL. Thus, in order to analyze STIT-based proposals in our new framework, one must first establish the precise relationship between both frameworks, i.e., compare models, semantics, expressive power, pragmatics (e.g., verification issues) etc. Several important results in this respect have already been reported [WöL 04, BRO 06], but there is still much to be done.

3. Constructive Strategic Logic: A New Semantics for Ability and Knowledge

ATOL covers more cases than ATL$_{ir}$ and "Feasible ATEL", and it is not committed to any notion of rationality (unlike ETSL). One major drawback of ATOL is that it vastly increases the number of modal operators necessary to express properties of agents. For team A, a whole family of cooperation modalities $\langle\langle A\rangle\rangle_{\mathcal{K}(\Gamma)}$ is used (instead of a single modality $\langle\langle A\rangle\rangle$ in ATL) to specify who should identify the right strategy for A, in what way etc. It would be much more elegant to modify the semantics of "simple" cooperation modalities $\langle\langle A\rangle\rangle$ and/or epistemic operators, so that they can be composed into sufficiently expressive formulae. The problem with strategic ability under uncertainty is that, when analyzing consequences of their strategies, agents must consider also the outcome paths starting from states other than the current state – namely, from all states that *look the same* as the current state. Thus, a property of a strategy being successful with respect to goal φ is not local to the current state; *the same* strategy must be successful in all "opening" states being considered. In order to capture this feature of strategic ability under imperfect information, we change the type of the satisfaction relation $\models$, and define what it means for a formula φ to be satisfied in a set of states $Q \subseteq St$ of model M. To our best knowledge, nobody has used this kind of semantics yet.

Moreover, we extend the language of ATEL with unary "constructive knowledge" operators $\mathbb{K}_a$, one for each agent a, that yield the set of states, indistinguishable from the current state from a's perspective. Constructive common, mutual, and distributed knowledge are formalized via operators $\mathbb{C}_A, \mathbb{E}_A$, and $\mathbb{D}_A$.

3.1. *Language and Semantics*

The language of Constructive Strategic Logic (CSL) includes atomic propositions, Boolean connectives, strategic formulae, standard epistemic operators, and *constructive knowledge operators* for groups of agents (individual knowledge can be defined as a special case of collective knowledge – see below):

$$\varphi ::= p \mid \neg\varphi \mid \varphi \wedge \varphi \mid \langle\langle A\rangle\rangle\bigcirc\varphi \mid \langle\langle A\rangle\rangle\Box\varphi \mid \langle\langle A\rangle\rangle\varphi\,\mathcal{U}\,\varphi \mid C_A\varphi \mid E_A\varphi \mid D_A\varphi \mid$$
$$\mathbb{C}_A\varphi \mid \mathbb{E}_A\varphi \mid \mathbb{D}_A\varphi.$$

where A is a set of agents.

REMARK 15. — As we will show in Section 7.2, standard knowledge can be defined as a special kind of constructive knowledge, and therefore the standard knowledge operators do *not* have to be included in the language. However, rather than immediately deriving C_A, E_A, D_A from $\mathbb{C}_A, \mathbb{E}_A, \mathbb{D}_A$, we choose to give the semantic clauses for all of them, and only later prove the relationship formally. □

Models are concurrent epistemic game structures again; that is, we interpret the formulae of CSL over exactly the same class of models which was used for ATEL, ATL$_{ir}$, ATOL etc. To recapitulate, a CEGS can be defined as a tuple

$$M = \langle \mathbb{A}\mathrm{gt}, St, \Pi, \pi, Act, d, o, \sim_1, ..., \sim_k\rangle,$$

where:

– $\mathbb{A}\mathrm{gt} = \{1, ..., k\}$ is a finite nonempty set of all agents,

– St is a nonempty set of states,

– Π is a set of atomic propositions,

– $\pi : St \to \mathcal{P}(\Pi)$ is a valuation of propositions,

– Act is a nonempty set of (atomic) actions;

– function $d : \mathbb{A}\mathrm{gt} \times St \to \mathcal{P}(Act)$ defines actions available to an agent in a state; $d(a, q) \neq \emptyset$ for all $a \in \mathbb{A}\mathrm{gt}, q \in St$,

– o is a (deterministic) transition function that assigns an outcome state to each combination of a state and a vector of actions (one action per agent). That is, $o(q, \alpha_1, \ldots, \alpha_k) \in St$ for every $q \in St$ and $\langle \alpha_1, \ldots, \alpha_k \rangle \in d(1, q) \times \cdots \times d(k, q)$;

– $\sim_1, ..., \sim_k \subseteq St \times St$ are epistemic accessibility relations, one per agent. It is assumed that each $\sim_a$ is an equivalence relation, and that $q \sim_a q'$ implies $d(a, q) = d(a, q')$.

Again, a (memoryless) strategy s_a of agent a is a conditional plan represented by function $s_a : St \to Act$ such that $s_a(q) \in d(a, q)$ for every q. A collective strategy S_A is a tuple of strategies, one per agent from A. Strategy s_a is uniform iff $q \sim_a q'$ implies $s_a(q) = s_a(q')$; a collective strategy is uniform iff it consists of only uniform individual strategies. A path Λ is an infinite sequence of states that can be effected by subsequent transitions; by $\Lambda[i]$, we denote the ith position on path Λ. Function $out(q, S_A)$ returns the set of all paths that may result from agents A executing strategy S_A from state q onward (see Section 2.1 for the precise definition). Collective epistemic relations are defined as: $\sim_A^D = \bigcap_{a \in A} \sim_a$, $\sim_A^E = \bigcup_{a \in A} \sim_a$; $\sim_A^C$ is defined as the transitive closure of $\sim_A^E$.

Now we define the notion of a formula φ being satisfied by a (non-empty) set of states Q in a model M, written $M, Q \models \varphi$. We will also write $M, q \models \varphi$ as a shorthand for $M, \{q\} \models \varphi$. Note that it is the latter notion of satisfaction (in single states) that we will ultimately be interested in – but that notion is defined in terms of the (more general) satisfaction in sets of states. Let $\mathrm{img}(q, \mathcal{R})$ be the image of state q with respect to binary relation $\mathcal{R}$, i.e., the set of all states q' such that $q\mathcal{R}q'$. Moreover, we use $out(Q, S_A)$ as a shorthand for $\bigcup_{q \in Q} out(q, S_A)$, and $\mathrm{img}(Q, \mathcal{R})$ as a shorthand for $\bigcup_{q \in Q} \mathrm{img}(q, \mathcal{R})$. The new semantics is given through the following clauses. In the semantics of cooperation modalities, only memoryless uniform strategies are considered.

$M, Q \models p$ iff $p \in \pi(q)$ for every $q \in Q$;

$M, Q \models \neg\varphi$ iff $M, Q \not\models \varphi$;

$M, Q \models \varphi \wedge \psi$ iff $M, Q \models \varphi$ and $M, Q \models \psi$;

$M, Q \models \langle\langle A \rangle\rangle \bigcirc \varphi$ iff there exists S_A such that, for every $\Lambda \in out(Q, S_A)$, we have that $M, \{\Lambda[1]\} \models \varphi$;

$M, Q \models \langle\langle A \rangle\rangle \Box \varphi$ iff there exists S_A such that, for every $\Lambda \in out(Q, S_A)$ and $i \geq 0$, we have $M, \{\Lambda[i]\} \models \varphi$;

$M, Q \models \langle\langle A \rangle\rangle \varphi \mathcal{U} \psi$ iff there exists S_A such that, for every $\Lambda \in out(Q, S_A)$, there is an $i \geq 0$ for which $M, \{\Lambda[i]\} \models \psi$ and $M, \{\Lambda[j]\} \models \varphi$ for every $0 \leq j < i$.

$M, Q \models \mathcal{K}_A \varphi$ iff $M, q \models \varphi$ for every $q \in \text{img}(Q, \sim_A^{\mathcal{K}})$ (where $\mathcal{K} = C, E, D$).

$M, Q \models \hat{\mathcal{K}}_A \varphi$ iff $M, \text{img}(Q, \sim_A^{\mathcal{K}}) \models \varphi$ (where $\hat{\mathcal{K}} = \mathbb{C}, \mathbb{E}, \mathbb{D}$ and $\mathcal{K} = C, E, D$, respectively).

The satisfaction relation $\models$ gives us both the traditional notion of satisfaction in a state, and the more general notion of satisfaction in a set of states. As mentioned above, we are usually interested in the former, but in order to interpret, e.g., an expression such as $\mathbb{C}_A \langle\langle A \rangle\rangle \bigcirc p$ in a single state, we must interpret the subexpression $\langle\langle A \rangle\rangle \bigcirc p$ in a set of states.

Formally, the language includes only operators for representing knowledge of teams. However, individual knowledge operators can be defined in the usual manner as:

$$\begin{aligned} K_a \varphi &\equiv C_{\{a\}} \varphi, \text{ and} \\ \mathbb{K}_a \varphi &\equiv \mathbb{C}_{\{a\}} \varphi. \end{aligned}$$

As a brief example, take the formula $\varphi = \mathbb{K}_a \langle\langle a \rangle\rangle \bigcirc \psi$ where a is an agent. We have that $M, q \models \varphi$ iff there is a strategy S_a for a such that for every $\Lambda \in out(\text{img}(q, \sim_a), S_a)$, $M, \Lambda[1] \models \psi$; in other words iff there is an (executable) strategy for a which is successfull (achieves ψ in the next state) in all the states that a considers to be possible. Or, in the terminology of Section 2.3, a knows a winning strategy – a has a strategy *de re* (for achieving ψ). We will discuss how the logic captures many subtly different properties of ability under imperfect information in more detail in Section 4, after we have clarified a few additional fundamental issues.

We employ the usual definition of the "sometime" operator:

$$\Diamond \varphi \equiv \top \mathcal{U} \varphi$$

We will also use derived propositional connectives. However, the exact meaning of these in the non-standard semantics must be carefully studied, and we will do that in Section 3.2. The CSL concept of validity is discussed in Section 3.3.

A note on notation: as above, we will henceforth use $\mathcal{K}_A$ to denote an arbitrary standard knowledge operator for agents A (i.e., C_A, E_A or D_A), and we use $\hat{\mathcal{K}}_A$ to denote the constructive knowledge operator corresponding to $\mathcal{K}_A$, i.e., $\hat{C}_A = \mathbb{C}_A$, $\hat{E}_A = \mathbb{E}_A$ and $\hat{D}_A = \mathbb{D}_A$. We use $\mathcal{K}, \mathcal{K}', \mathcal{K}_1, \mathcal{K}_2$ etc. to denote arbitrary standard knowledge operators for arbitrary sets of agents, and, again, $\hat{\mathcal{K}}, \hat{\mathcal{K}}', \hat{\mathcal{K}}_1, \hat{\mathcal{K}}_2$ etc. to denote the corresponding constructive modalities.

3.2. *Additional Operators*

In addition to the derived operators introduced in Section 3.1, we use a slightly unusual definition of the Boolean "false" and "true" constants:

$$\begin{array}{rcl} \bot & \equiv & \langle\!\langle\emptyset\rangle\!\rangle(p \wedge \neg p)\,\mathcal{U}\,(p \wedge \neg p), \quad \text{where } p \text{ is an arbitrary primitive proposition,} \\ \top & \equiv & \langle\!\langle\emptyset\rangle\!\rangle(\neg\bot)\,\mathcal{U}\,(\neg\bot) \end{array}$$

and the usual definition of Boolean connectives[6]:

$$\begin{array}{rcl} \varphi_1 \vee \varphi_2 & \equiv & \neg(\neg\varphi_1 \wedge \neg\varphi_2), \\ \varphi_1 \to \varphi_2 & \equiv & \neg\varphi_1 \vee \varphi_2, \text{ and} \\ \varphi_1 \leftrightarrow \varphi_2 & \equiv & (\varphi_1 \to \varphi_2) \wedge (\varphi_2 \to \varphi_2). \end{array}$$

The above Boolean operators have the following semantic characterizations:

PROPOSITION 16. —

1) $M, Q \not\models \bot$ *for all* $Q \subseteq St$, $Q \neq \emptyset$.
2) $M, Q \models \top$ *for all* $Q \subseteq St$, $Q \neq \emptyset$.
3) $M, Q \models \varphi_1 \vee \varphi_2$ *iff* $M, Q \models \varphi_1$ *or* $M, Q \models \varphi_2$.
4) $M, Q \models \varphi_1 \to \varphi_2$ *iff* $M, Q \models \varphi_1$ *implies* $M, Q \models \varphi_2$.
5) $M, Q \models \varphi_1 \leftrightarrow \varphi_2$ *iff we have that* $M, Q \models \varphi_1$ *iff* $M, Q \models \varphi_2$.

Proof.

1) Suppose that $M, Q \models \bot$ for some $Q \neq \emptyset$. Then $M, Q \models \langle\!\langle\emptyset\rangle\!\rangle(p \wedge \neg p)\,\mathcal{U}\,(p \wedge \neg p)$, so for all paths Λ starting from the states in Q we have $M, \Lambda[0] \models p \wedge \neg p$. That is, for all $q \in Q$: $M, q \models p \wedge \neg p$. As Q is nonempty, there is at least one such q. But that means that $p \in \pi(q)$ and $p \notin \pi(q)$, which cannot be the case.

2) Analogous.

3) $M, Q \models \varphi_1 \vee \varphi_2$ iff $M, Q \models \neg(\neg\varphi_1 \wedge \neg\varphi_2)$ iff $M, Q \not\models \neg\varphi_1 \wedge \neg\varphi_2$ iff $M, Q \not\models \neg\varphi_1$ or $M, Q \not\models \neg\varphi_2$ iff $M, Q \models \varphi_1$ or $M, Q \models \varphi_2$.

4), 5) Straightforward from the above.

■

To conclude the analysis of standard connectives in this (rather non-standard) setting, we observe that the $\neg$ operator behaves like classical negation: it obeys the law

6. The reason why we use the above definitions of $\top$ and $\bot$ instead of the more common ones: $\bot \equiv p \wedge \neg p$, $\top \equiv \neg\bot$ is that in the restricted language CSL^-, discussed in Section 6.3, certain formulae are disallowed, namely the ones in which negation (or a sequence of conjunctions, followed by negation) follows a constructive knowledge operator. Defining the Boolean constants the way we do, we make sure that no unraveling of $\top$ or $\bot$ will ever lead to such a formula.

of double negation, the law of excluded middle, and the consistency requirement in every possible context:

PROPOSITION 17. — *We have the following for every M and $Q \subseteq St$:*

1) $M, Q \models \neg\neg\varphi \leftrightarrow \varphi$,

2) $M, Q \models \varphi \vee \neg\varphi$,

3) $M, Q \models \neg(\varphi \wedge \neg\varphi)$.

Proof. Straightforward from Proposition 16 and the semantic definition of $\neg$. ■

It should be noted that there are other possibilities for defining negation, disjunction and implication, corresponding to the different ways of quantifying over the set Q. We discuss the issue in more detail in Section 7.

3.3. *Validity*

We say that a formula is *weakly valid* (or simply *valid*) if it is satisfied individually by *each state* in every model, i.e., if $M, q \models \varphi$ for all models M and states q in M. It is *strongly valid* if it is satisfied by all non-empty *sets* in all models; i.e., if for each M and every non-empty set of states Q it is the case that $M, Q \models \varphi$. We are ultimately interested in the former (see Remark 19 below). The importance of strong validity, on the other hand, lies in the fact that strong validity of $\varphi \leftrightarrow \psi$ makes φ and ψ completely interchangeable (cf. Proposition 20.2). It is not difficult to see that the same is not true for weak validity.

PROPOSITION 18. —

1) Strong validity implies validity.

2) Validity does not imply strong validity.

Proof. (1) Straightforward. (2) We here take the liberty to refer forward to some simple results we haven't proven yet, because it is instructive to point out the distinction between weak and strong validity at this point. By Propositions 16.5 and 44, we have that for any M and set of states Q, $M, Q \models \langle\langle\emptyset\rangle\rangle\varphi\,\mathcal{U}\,\varphi \leftrightarrow \varphi$ iff ($\forall_{q\in Q} M, q \models \varphi$ iff $M, Q \models \varphi$). It follows immediately that $\langle\langle\emptyset\rangle\rangle\varphi\,\mathcal{U}\,\varphi \leftrightarrow \varphi$ is (weakly) valid, for any φ. It follows from Lemma 38.1 that there is a M and a set of states Q and a formula φ such that $M, Q \not\models \varphi$ but $\forall_{q\in Q} M, q \models \varphi$; thus $\langle\langle\emptyset\rangle\rangle\varphi\,\mathcal{U}\,\varphi \leftrightarrow \varphi$ is not strongly valid. ■

REMARK 19. — The term *the logic* is sometimes understood as the set of all valid formulae in the logic. In this sense, we define *the logic of* CSL as the set of all weakly valid formulae of CSL. In a similar way, we say that a formula φ is CSL-*satisfiable* if it is weakly satisfiable in CSL, i.e., there is a model M and a state q such that $M, q \models \varphi$. □

Propositions 16.4 and 16.5 from Section 3.2 have two important consequences. First, the rule of Modus Ponens is correct with respect to this semantics. Second, if

$\varphi_1 \leftrightarrow \varphi_2$ is strongly valid, then formulae φ_1 and φ_1 are completely interchangeable under strong (and hence also weak) validity.

PROPOSITION 20. —

1) If $\varphi_1 \rightarrow \varphi_2$ is strongly (resp. weakly) valid, and φ_1 is strongly (resp. weakly) valid, then φ_2 is strongly (resp. weakly) valid.

2) If $\varphi_1 \leftrightarrow \varphi_2$ is strongly valid, and ψ' is obtained from ψ through replacing an occurrence of φ_1 by φ_2, then $M, Q \models \psi$ iff $M, Q \models \psi'$.

Proof. Straightforward. ■

4. Expressing Agents' Strategic Abilities

In the language of Constructive Strategic Logic, strategic properties of coalitions can be expressed in a flexible and elegant way. To support this claim, we first show that the philosophical discourse on various levels of knowledge and ability, mentioned in Section 2.3, has its formal counterpart in CSL formulae. Then, we present a translation of ATL_{ir}, ATOL and "Feasible ATEL" to CSL, and thus prove that the latter embeds the former ones. We also discuss the relationship between ETSL and CSL. To avoid confusion, we will use the satisfaction sign with subscripts ($\models_{\text{ATOL}}$, $\models_{\text{CSL}}$, $\models_{\text{ETSL}}$ etc.), indicating which semantics is currently referred to.

4.1. *Capturing Levels of Strategic Power*

The reason why we need to interpret formulae over sets of states is that we need non-standard epistemic operators: $M, q \models \mathbb{K}_a\langle\langle a\rangle\rangle\varphi$ expresses the fact that a has a single strategy that enforces φ from *all* states indiscernible from q, instead of stating that φ can be achieved from *every* such state *separately*. Note that the latter property is very much in the spirit of standard epistemic logic, and indeed can be captured with the standard knowledge operator (via $K_a\langle\langle a\rangle\rangle\varphi$). Speaking in more abstract terms:

1) $\mathbb{K}_a\langle\langle a\rangle\rangle\varphi$ refers to agent a having a strategy *"de re"* to enforce φ (i.e. having a successful strategy and knowing the strategy);

2) $K_a\langle\langle a\rangle\rangle\varphi$ refers to agent a having a strategy *"de dicto"* to enforce φ (i.e. knowing only that *some* successful strategy is available);

3) $\langle\langle a\rangle\rangle\varphi$ expresses that agent a has a strategy to enforce φ *from the current state* (but not necessarily even knows about it).

Above, each of the three formulae are informally interpreted in an assumed (single) state q of a model M, i.e., we discuss the meaning of, e.g., $M, q \models \mathbb{K}_a\langle\langle a\rangle\rangle\varphi$. The meaning of this formula in this single state is again defined by interpreting a subformula in a certain set of states. By *strategies* here, we only mean executable (i.e., uniform) strategies. Capturing different ability levels of coalitions is analogous, with various "epistemic modes" of collective recognizing the right strategy.

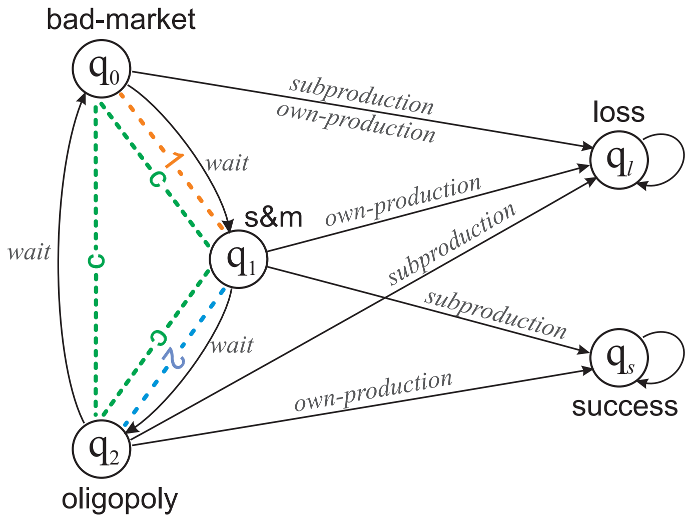

Figure 3. *Simple market example: model M_4*

EXAMPLE 21. — Robot a has no winning strategy in the starting state of the game: $M_3, q_0 \models \neg\langle\langle a\rangle\rangle\Diamond\mathsf{win}$, which implies that it has neither a strategy "de re" nor "de dicto": $M_3, q_0 \models \neg\mathbb{K}_a\langle\langle a\rangle\rangle\Diamond\mathsf{win} \wedge \neg K_a\langle\langle a\rangle\rangle\Diamond\mathsf{win}$. On the other hand, he has a successful strategy in q_{AK} (just play $keep$) and it knows it has one (because another action, $exch$, is bound to win in q_{AQ}); still, the knowledge is not constructive, since a does not know which strategy is the right one in the current situation: $M_3, q_{AK} \models \langle\langle a\rangle\rangle\bigcirc\mathsf{win} \wedge K_a\langle\langle a\rangle\rangle\bigcirc\mathsf{win} \wedge \neg\mathbb{K}_a\langle\langle a\rangle\rangle\bigcirc\mathsf{win}$.

Other properties of the gambling robots, that we discussed in Examples 13 and 12, can be easily expressed in the new logic by combining constructive knowledge with cooperation modalities: $M_3, q_0 \models \neg\mathbb{E}_{\{a,b\}}\langle\langle a,b\rangle\rangle\Diamond\mathsf{win}$, $M_3, q_{AK} \models \mathbb{D}_{\{a,b\}}\langle\langle a,b\rangle\rangle\bigcirc\mathsf{win}\wedge \mathbb{K}_a\langle\langle a,b\rangle\rangle\bigcirc\mathsf{win}\wedge\neg\mathbb{E}_{\{a,b\}}\langle\langle a,b\rangle\rangle\bigcirc\mathsf{win}$, $M_3, q_{AQ} \models \mathbb{E}_{\{a,b\}}\langle\langle a,b\rangle\rangle\bigcirc\mathsf{win}\wedge\neg\mathbb{C}_{\{a,b\}}\langle\langle a,b\rangle\rangle\bigcirc\mathsf{win}$ etc. In fact, it turns out that the new logic is expressive enough to embed most approaches we have discussed. We present an appropriate translation in the next section. □

EXAMPLE 22. — Consider a market model, depicted in Figure 3, which formalizes in a very simple way the scenario from Example 3. The economy is assumed to run in simple cycles: after the moment of bad economy (bad-market), there is always a good time for small and medium enterprises (s&m), after which the market tightens and an oligopoly emerges. At the end, the market gets stale, and we have stagnation and bad economy again.

The company c is the only agent whose actions are represented in the model. The company can wait (action *wait*) or decide to start production: either on its own (*own-production*), or as a subcontractor of a major company (*subproduction*). Both decisions can lead to either loss or success, depending on the current market conditions.

However, the company management cannot recognize the market conditions: bad market, time for small and medium enterprises, and oligopoly market look the same to them, as the epistemic links for c indicate.

The company can call the services of two marketing experts. Expert 1 is a specialist on oligopoly, and can recognize oligopoly conditions (although she cannot distinguish between bad economy and s&m market). Expert 2 can recognize bad economy, but he cannot distinguish between other types of market. The experts' actions have no influence on the actual transitions of the model, and are omitted from the graph in Figure 3. It is easy to see that the company cannot identify a successful strategy on its own: for instance, for the small and medium enterprises period, we have that $M_4, q_1 \models \neg\mathbb{K}_c\langle\langle c\rangle\rangle\Diamond\mathsf{success}$. It is not even enough to call the help of a single expert: $M_4, q_1 \models \neg\mathbb{K}_1\langle\langle c\rangle\rangle\Diamond\mathsf{success} \wedge \neg\mathbb{K}_2\langle\langle c\rangle\rangle\Diamond\mathsf{success}$, or to ask the experts to independently work out a common strategy: $M_4, q_1 \models \neg\mathbb{E}_{\{1,2\}}\langle\langle c\rangle\rangle\Diamond\mathsf{success}$. Still, the experts can propose the right strategy if they join forces and cooperate to find the solution: $M_4, q_1 \models \mathbb{D}_{\{1,2\}}\langle\langle c\rangle\rangle\Diamond\mathsf{success}$.

Note that this is not true any more for bad market, i.e., $M_4, q_0 \models \neg\mathbb{D}_{\{1,2\}}\langle\langle c\rangle\rangle\Diamond\mathsf{success}$, because c is a memoryless agent, and it has no uniform strategy to enforce $\mathsf{success}$ from q_0 at all. However, the experts can suggest a more complex scheme that involves consulting them once again in the future: $M_4, q_0 \models \mathbb{D}_{\{1,2\}}\langle\langle c\rangle\rangle\bigcirc\mathbb{D}_{\{1,2\}}\langle\langle c\rangle\rangle\Diamond\mathsf{success}$. □

For strategic abilities, standard knowledge corresponds to knowing "de dicto", while constructive knowledge captures "knowing how to play". We observe that both kinds of epistemic operators can be combined in a meaningful way. For example, $K_a\mathbb{K}_b\langle\langle b\rangle\rangle\Diamond\mathsf{win}$ says that agent a knows that player b knows how to win. Note that this is substantially different from $\mathbb{K}_a\mathbb{K}_b\langle\langle b\rangle\rangle\Diamond\mathsf{win}$, which says that agent a can identify a strategy which b knows to be winning. Also, when interleaving epistemic operators with strategic operators, we can, e.g., describe an ability to acquire, distribute or maintain *ability*. For instance, $\mathbb{K}_a\langle\langle a\rangle\rangle\Box\mathbb{K}_b\langle\langle b\rangle\rangle\Diamond\mathsf{win}$ means that a knows how to maintain b's (constructive) ability to win, while $K_a\langle\langle a\rangle\rangle\Box\mathbb{K}_b\langle\langle b\rangle\rangle\Diamond\mathsf{win}$ says only that a knows that this is in principle possible, and $\mathbb{K}_a\langle\langle a\rangle\rangle\Box K_b\langle\langle b\rangle\rangle\Diamond\mathsf{win}$ says that a knows how to keep b aware that a winning strategy exists.

4.2. *Expressivity of CSL*

Let $\mathcal{L}$ be the logic of ATL$_{ir}$, ATOL or F-ATEL, and let φ, ψ be formulae of $\mathcal{L}$. Also, let $\mathcal{K} = C, E, D$ and $\hat{\mathcal{K}} = \mathbb{C}, \mathbb{E}, \mathbb{D}$, respectively. Then, let the translation function tr be defined as follows:

$$
\begin{array}{cc}
tr(p) = p & tr(\neg\varphi) = \neg tr(\varphi) \\
tr(\varphi \wedge \psi) = tr(\varphi) \wedge tr(\psi) & tr(\bigcirc\varphi) = \bigcirc tr(\varphi) \\
tr(\Box\varphi) = \Box tr(\varphi) & tr(\varphi\,\mathcal{U}\,\psi) = tr(\varphi)\,\mathcal{U}\,tr(\psi) \\
tr(\langle\langle A\rangle\rangle_{ir}\varphi) = \mathbb{E}_A\langle\langle A\rangle\rangle tr(\varphi) & tr(\langle\langle A\rangle\rangle_{\mathcal{K}(\Gamma)}\varphi) = \hat{\mathcal{K}}_\Gamma\langle\langle A\rangle\rangle tr(\varphi) \\
tr(\langle\langle A\rangle\rangle^f\varphi) = \langle\langle A\rangle\rangle tr(\varphi) & tr(\langle\langle A\rangle\rangle^f_{\mathcal{K}}\varphi) = \hat{\mathcal{K}}_A\langle\langle A\rangle\rangle tr(\varphi) \\
tr(\langle\langle A\rangle\rangle^f_{K_b}\varphi) = \mathbb{K}_b\langle\langle A\rangle\rangle tr(\varphi) & tr(\langle\langle A\rangle\rangle^f_{M_b}\varphi) = \neg K_b\neg\langle\langle A\rangle\rangle tr(\varphi) \\
tr(\mathcal{K}_A\varphi) = \mathcal{K}_A tr(\varphi) &
\end{array}
$$

The following result justifies the translation.

THEOREM 23. — $M, q \models_{\mathcal{L}} \varphi$ *iff* $M, q \models_{\text{CSL}} tr(\varphi)$.
Proof in the Appendix.

COROLLARY 24. — *The translation yields a reduction of* ATL$_{ir}$*,* ATOL *and "Feasible* ATEL*" model checking problems to* CSL *model checking. The time needed for the reduction, and the resulting formula, are linear in the length of the original formula. We summarize the model checking complexity results for* CSL *in Section 5.*

PROPOSITION 25. — *Constructive Strategic Logic is* strictly more expressive *than* ATL$_{ir}$*,* ATOL *etc.*

Proof. It is sufficient to prove that there is a CSL formula φ that has no ATOL equivalent (i.e., there is no ATOL formula which holds in exactly the same models and states as φ). Consider the formula $\varphi \equiv \mathbb{E}_A\mathbb{E}_A\langle\langle A\rangle\rangle\psi$. For most models ($M_3$ from Figure 2 being an example) we have $\sim^D_A \subsetneq \sim^E_A \subsetneq \sim^E_A \circ \sim^E_A \subsetneq \sim^C_A$, so φ is equivalent to neither $\langle\langle A\rangle\rangle_{D(A)}\psi$, $\langle\langle A\rangle\rangle_{E(A)}\psi$, nor $\langle\langle A\rangle\rangle_{C(A)}\psi$. This is of course possible, because $\mathbb{E}_A$ (similarly to E_A) is not a KD45 modality (see Theorem 40 in Section 6.4). ■

Note that the semantics of CSL is based on *exactly the same* class of models as ATEL, ATOL, ATL$_{ir}$ etc. (i.e., on CEGSs). Thus, the above translation can also be used for reduction of validity (resp. satisfiability) problems for ATL$_{ir}$, ATOL and "Feasible ATEL" to weak validity (resp. satisfiability) of CSL. By Theorem 23, we have the following.

COROLLARY 26. — ATL$_{ir}$*,* ATOL *and "Feasible* ATEL*" can be embedded in* CSL*.*

4.3. *Constructive Strategic Logic vs. ATEL*

As we already pointed out in Section 2.3, ATEL only enables expressing ability of type (4): the ATEL formula $\langle\langle A\rangle\rangle\varphi$ says that agents A may *happen* to behave in such a way that φ is enforced (but there might be no executable strategy to enforce it). Thus, ATEL is about a kind of ability different from the "constructive" one we study in this paper. Formally, ATEL differs from CSL in two main ways. First, it does not require uniform strategies. Second, it does not have the constructive knowledge operators.

First, consider non-uniformity. Note that uniform strategies is not a new idea of CSL (see Section 2), and that the differences between the ATEL operators and the uniform variants used by CSL are also shared by all the previously studied logics using uniform strategies. We nevertheless comment briefly on the difference here. First, the "nexttime" fragment of ATEL can be embedded in CSL, as the following proposition shows. It should be remembered that the CEGSs used in ATEL are slightly more general than the ones used in CSL (and the other approaches we have discussed): they do not require that the same actions are available in indistinguishable states. Below we refer to such CEGSs as *uniform* CEGS*s*.

PROPOSITION 27. — *Let φ be an* ATEL *formula that does not include operators* $\Box$, $\mathcal{U}$ *and M be a uniform* CEGS. *Then,* $M, q \models_{\text{ATEL}} \varphi$ *iff* $M, q \models_{\text{CSL}} \varphi$.

Proof. It is sufficient to note that $M, q \models_{\text{ATEL}} \langle\!\langle A \rangle\!\rangle \bigcirc \varphi$ iff $M, q \models_{\text{CSL}} \langle\!\langle A \rangle\!\rangle \bigcirc \varphi$. Thus, we have that the "nexttime" formulae have the same semantics in both logics when interpreted at single states, and ATEL formulae include no "constructive" operators for aggregating sets of states. ∎

REMARK 28. — It is well known that cooperation modalities for strategies of perfect information (e.g., the ones in ATL and ATEL) have the following fixpoint characterizations:

$$\langle\!\langle A \rangle\!\rangle \Box \varphi \quad \leftrightarrow \quad \varphi \wedge \langle\!\langle A \rangle\!\rangle \bigcirc \langle\!\langle A \rangle\!\rangle \Box \varphi, \tag{1}$$

$$\langle\!\langle A \rangle\!\rangle \varphi \,\mathcal{U}\, \psi \quad \leftrightarrow \quad \psi \vee \varphi \wedge \langle\!\langle A \rangle\!\rangle \bigcirc \langle\!\langle A \rangle\!\rangle \varphi \,\mathcal{U}\, \psi. \tag{2}$$

For uniform information strategies, the above formulae are not valid any more (see below). Still, it would be possible to embed the whole ATEL in CSL if we included fixpoint operators in the latter. In that case, the following translation could be used to translate ATL/ATEL modalities to equivalent CSL counterparts:

$$tr(\langle\!\langle A \rangle\!\rangle \Box \varphi) \;=\; \nu \mathbf{Z}.\varphi \wedge \langle\!\langle A \rangle\!\rangle \bigcirc \mathbf{Z},$$

$$tr(\langle\!\langle A \rangle\!\rangle \varphi \,\mathcal{U}\, \psi) \;=\; \mu \mathbf{Z}.\psi \vee \varphi \wedge \langle\!\langle A \rangle\!\rangle \bigcirc \mathbf{Z}.$$

□

We note that due to the uniformity of CSL strategies, the set of ATEL validities is not contained in CSL validities. A counter-example is the formula $\langle\!\langle r \rangle\!\rangle \Box \neg \mathsf{jail} \leftrightarrow \neg \mathsf{jail} \wedge \langle\!\langle r \rangle\!\rangle \bigcirc \langle\!\langle r \rangle\!\rangle \Box \neg \mathsf{jail}$. It is valid in ATEL (it is an instance of the valid scheme that gives a characterization of "always" in terms of "next" in ATL and ATEL). Still, the formula is false in model M_2 and state q_0 from Example 9: the left hand side of the biconditional is false, but the right hand side is true in M_2, q_0.

More importantly, we can show that CSL is more powerful than ATEL when we want to characterize sets of situations in actual systems. First, given a finite model, every ATEL formula has a CSL counterpart (i.e., a CSL formula which holds in exactly the same states). Second, CSL allows for finer-grained specifications than ATEL (in the sense that there are CSL formulae for which there are no ATEL formulae with the same extension). The result is formalized in Propositions 29 and 30.

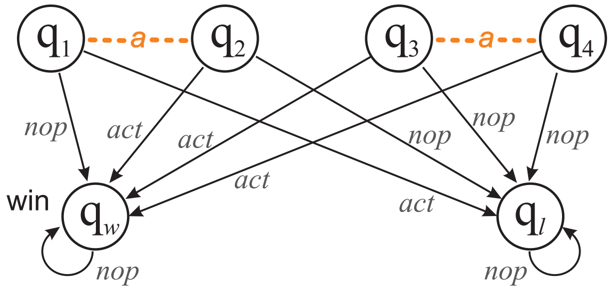

Figure 4. *A model with one agent. From each of the states q_1, q_2, q_3, q_4, the same outcomes can be achieved in one step, albeit through different actions.*

PROPOSITION 29. — *Given a uniform* CEGS*, every* ATEL *formula has a* CSL *counterpart with the same extension (i.e., one which is satisfied in exactly the same states of the model).*

Proof (sketch). For finite models: let M be a model with $|M|$ states, and φ be an ATEL formula. All subformulae $\langle\!\langle A\rangle\!\rangle\Box\psi$ can be equivalently rewritten as $(\psi\wedge\langle\!\langle A\rangle\!\rangle\bigcirc)^{|M|}\psi$, where $|M|$ is the number of states in M. This follows by the property (1) above, and the fact that, after $|M|$ steps, the system is bound to come back to one of the previously visited states, for which a successful action has already been found. Similarly, subformulae $\langle\!\langle A\rangle\!\rangle\psi_1\,\mathcal{U}\,\psi_2$ can be equivalently rewritten as $(\psi_2\vee\psi_1\wedge\langle\!\langle A\rangle\!\rangle\bigcirc)^{|M|}\psi_2$. This way, we get an ATEL formula φ' without $\Box$, $\mathcal{U}$ which holds in exactly the same states as φ. By Proposition 27, φ' has the same extension in ATEL and CSL. ■

PROPOSITION 30. — *Given a uniform* CEGS*, there can be* CSL *formulae that have no* ATEL *counterpart with the same extension (i.e., one which is satisfied in exactly the same states of the model).*

Proof. Consider model M_5 from Figure 4. The formula $\mathbb{K}_a\langle\!\langle a\rangle\!\rangle\bigcirc\mathsf{win}$ holds in q_3 and q_4, but not in q_1 nor q_2. There is no ATEL formula which is true exactly in q_3, q_4: it is easy to see that an ATEL formula is true in q_1 iff it is true in q_2 iff it is true in q_3 iff it is true in q_4. ■

4.4. *Constructive Strategic Logic vs. ETSL*

CSL and ETSL are underpinned by different notions of ability. ETSL can be treated as a logic that describes the outcome of *rational play* under imperfect information,[7]

7. We emphasize that this is a specific notion of rationality (i.e., agents are assumed to *play only undominated strategies*). Game theory proposes several other rationality criteria as well, based e.g. on Nash equilibrium, dominant strategies, or Pareto efficiency. In fact, it is easy to imagine ETSL-like logics based on these notions instead.

in the same way as CSL can be seen as a logic that captures agents' strategic abilities (regardless of whether the agents play rationally or not). Thus, the focus of CSL and ETSL is different, and we suspect that neither logic formally subsumes the other. However, several interesting associations have been already proposed in [JAM 06a].

Let us consider only models with finite state spaces,[8] and formulae $\Phi \equiv \bigcirc\psi, \Box\psi$, or $\psi_1 \,\mathcal{U}\, \psi_2$ where ψ, ψ_1, ψ_2 are "vanilla" ETSL formulae.

PROPOSITION 31 ([JAM 06A]). — *An agent has a strategy "de re" to enforce Φ if, and only if, he knows that his rational play will bring about Φ. Formally:*

$$M, q \models_{\text{ETSL}} K_a\langle\langle a\rangle\rangle\Phi \quad \textit{iff} \quad M, q \models_{\text{CSL}} \mathbb{K}_a\langle\langle a\rangle\rangle\Phi.$$

PROPOSITION 32 ([JAM 06A]). — *If a coalition has common knowledge about how to play, then it has common knowledge that rational play will be successful:*

$$\textit{if} \quad M, q \models_{\text{CSL}} \mathbb{C}_A\langle\langle A\rangle\rangle\Phi \quad \textit{then} \quad M, q \models_{\text{ETSL}} C_A\langle\langle A\rangle\rangle\Phi.$$

The same holds for neither mutual nor distributed knowledge.

PROPOSITION 33 ([JAM 06A]). — *If A have distributed knowledge that rational play will bring about Φ, then they have distributed knowledge how to play to bring about Φ. Formally:*

$$\textit{if} \quad M, q \models_{\text{ETSL}} D_A\langle\langle A\rangle\rangle\Phi \quad \textit{then} \quad M, q \models_{\text{CSL}} \mathbb{D}_A\langle\langle A\rangle\rangle\Phi.$$

The same holds for neither mutual nor common knowledge.

A more definitive study of this issue is beyond the scope of this paper.

5. Verification of Strategic Abilities Through Model Checking

The *model checking* problem asks whether a given formula φ holds in a given model M and state q. We define the *general model checking* problem as the problem that asks whether formula φ holds in model M and *set of states* Q. Let $mctl(\varphi, M)$ be a CTL model checker that returns the set of all states which satisfy φ in M. Below, we sketch an algorithm $mcheck(\varphi, M, Q)$ that returns *true* if $M, Q \models_{\text{CSL}} \varphi$ and *false* otherwise, running in time $\mathbf{\Delta_2^P}$, i.e., in deterministic polynomial time with adaptive queries to an $\mathbf{NP}$ oracle.

Case $\varphi \equiv p$: return(*true*) if $p \in \pi(q)$ for all $q \in Q$, else return(*false*);

Case $\varphi \equiv \neg\psi$: return(*true*) if $mcheck(\psi, M, Q) =$ *false*, else return(*false*);

Case $\varphi \equiv \psi_1 \wedge \psi_2$: return(*true*) if $mcheck(\psi_1, M, Q) =$ *true* and $mcheck(\psi_2, M, Q) =$ *true*, else return(*false*);

8. More generally, we can consider models M such that there exists at least one undominated strategy wrt M, q, Φ.

Case $\varphi \equiv \mathcal{K}_A\psi$**:** Compute $Q' := \mathrm{img}(Q, \sim_A^{\mathcal{K}})$, and then return(*true*) if $mcheck(\psi, M, q) =$ *true* for all $q \in Q'$, else return(*false*);

Case $\varphi \equiv \hat{\mathcal{K}}_A\psi$**:** return($mcheck(\psi, M, \mathrm{img}(Q, \sim_A^{\mathcal{K}}))$);

Case $\varphi \equiv \langle\langle A\rangle\rangle \bigcirc \psi$**:** Run $mcheck(\psi, M, q)$ for every $q \in St$, and label the states in which the answer was *true* with an additional proposition yes (not used elsewhere). Then, guess the strategy of A, and "trim" model M by removing all the transitions inconsistent with the strategy (yielding a sparser model M'). Finally, return(*true*) if $Q \subseteq mctl(\mathsf{A}\bigcirc\mathsf{yes}, M')$, else return(*false*).

Note: subformula ψ is checked in the original model M, and not in M'!

Case $\varphi \equiv \langle\langle A\rangle\rangle \Box \psi$**:** Run $mcheck(\psi, M, q)$ for every $q \in St$, and label the states in which the answer was *true* with an additional proposition yes (not used elsewhere). Then, guess the strategy of A, and "trim" model M by removing all the transitions inconsistent with the strategy (yielding a sparser model M'). Finally, return(*true*) if $Q \subseteq mctl(\mathsf{A}\Box\mathsf{yes}, M')$, else return(*false*). Again, note that ψ is checked in the original model M.

Case $\varphi \equiv \langle\langle A\rangle\rangle \psi_1 \,\mathcal{U}\, \psi_2$**:** analogous.

As model checking CTL can be done in deterministic polynomial time [CLA 86], we get the following.

PROPOSITION 34. — *General model checking for Constructive Strategic Logic is in $\mathbf{\Delta_2^P}$ when the input size is measured with the number of transitions (and epistemic links) in the model, and the length of the formula.*

For the lower bound, we observe that CSL subsumes ATL_{ir}, and model checking ATL_{ir} is $\mathbf{\Delta_2^P}$-complete [SCH 04, JAM 06b]. Thus, we pay no price in terms of complexity for using the more expressive language of CSL:

THEOREM 35. — *General model checking for Constructive Strategic Logic is $\mathbf{\Delta_2^P}$-complete in the number of transitions (and epistemic links) in the model, and the length of the formula.*

6. Constructive Knowledge

Philosophically, constructive knowledge draws inspiration from mathematical constructivism: in order to "constructively know" that φ, agents A must be able to find (or "construct") a mathematical object that supports φ. This is relevant when $\varphi \equiv \langle\langle B\rangle\rangle\psi$ – in that case, the mathematical object in question is a strategy for B which guarantees achieving ψ. The semantic role of *constructive knowledge operators* is to produce sets of states that will appear on the left hand side of the satisfaction relation. In a way, these modalities "aggregate" states into sets, and sets into bigger sets. On the other hand, most of the other operators "split" (or "destroy") sets in the sense that, for evaluating $M, Q \models \varphi$, they require evaluation of subformulae of φ in single states rather than sets of states. Standard epistemic operators (C_A, E_A, D_A) are the most straightforward examples (e.g., evaluating $C_A\psi$ in M, Q "splits" into evaluating ψ in each state from $\mathrm{img}(Q, \sim_A^C)$ separately). Cooperation modalities (combined with temporal operators) are "splitting" in a similar way. Besides the "aggregating" and "splitting" operators, there are also "neutral" ones that do not change the set of reference: namely, conjunction ($\wedge$) and negation ($\neg$). In what follows, we study important properties of these operators in CSL.

6.1. *Properties of Constructive Knowledge*

In the following proposition we list some properties of constructive knowledge (keep in mind that strong validity implies validity).

PROPOSITION 36. — *The following are strongly valid for any $\hat{\mathcal{K}} \in \{\mathbb{C}, \mathbb{D}, \mathbb{E}\}$:*

1) $\hat{\mathcal{K}}_A(\varphi_1 \vee \varphi_2) \leftrightarrow (\hat{\mathcal{K}}_A\varphi_1 \vee \hat{\mathcal{K}}_A\varphi_2)$

2) $\hat{\mathcal{K}}_A\neg\varphi \leftrightarrow \neg\hat{\mathcal{K}}_A\varphi$

3) $\hat{\mathcal{K}}_A(\varphi_1 \wedge \varphi_2) \leftrightarrow (\hat{\mathcal{K}}_A\varphi_1 \wedge \hat{\mathcal{K}}_A\varphi_2)$

4) $\hat{\mathcal{K}}_A(\varphi_1 \rightarrow \varphi_2) \leftrightarrow (\hat{\mathcal{K}}_A\varphi_1 \rightarrow \hat{\mathcal{K}}_A\varphi_2)$

Proof.

1) $M, Q \models \hat{\mathcal{K}}_A(\varphi_1 \vee \varphi_2)$ iff $M, \mathrm{img}(Q, \sim_A^{\hat{\mathcal{K}}}) \models \varphi_1 \vee \varphi_2$ iff $M, \mathrm{img}(Q, \sim_A^{\hat{\mathcal{K}}}) \models \varphi_1$ or $M, \mathrm{img}(Q, \sim_A^{\hat{\mathcal{K}}}) \models \varphi_2$ iff $M, Q \models \hat{\mathcal{K}}_A\varphi_1$ or $M, Q \models \hat{\mathcal{K}}_A\varphi_2$ iff $M, Q \models \hat{\mathcal{K}}_A\varphi_1 \vee \hat{\mathcal{K}}_A\varphi_2$.

2) $M, Q \models \hat{\mathcal{K}}_A\neg\varphi$ iff $M, \mathrm{img}(Q, \sim_A^{\hat{\mathcal{K}}}) \models \neg\varphi$ iff $M, \mathrm{img}(Q, \sim_A^{\hat{\mathcal{K}}}) \not\models \varphi$ iff $M, Q \not\models \hat{\mathcal{K}}_A\varphi$ iff $M, Q \models \neg\hat{\mathcal{K}}_A\varphi$.

3) $M, Q \models \hat{\mathcal{K}}_A(\varphi_1 \wedge \varphi_2)$ iff $M, \mathrm{img}(Q, \sim_A^{\hat{\mathcal{K}}}) \models \varphi_1 \wedge \varphi_2$ iff $M, \mathrm{img}(Q, \sim_A^{\hat{\mathcal{K}}}) \models \varphi_1$ and $M, \mathrm{img}(Q, \sim_A^{\hat{\mathcal{K}}}) \models \varphi_2$ iff $M, Q \models \hat{\mathcal{K}}_A\varphi_1$ and $M, Q \models \hat{\mathcal{K}}_A\varphi_2$ iff $M, Q \models \hat{\mathcal{K}}_A\varphi_1 \wedge \hat{\mathcal{K}}_A\varphi_2$.

4) $M, Q \models \hat{\mathcal{K}}_A(\neg\varphi_1 \vee \varphi_2)$ iff $M, Q \models (\hat{\mathcal{K}}_A\neg\varphi_1) \vee \hat{\mathcal{K}}_A\varphi_2$ iff $M, Q \models (\neg\hat{\mathcal{K}}_A\varphi_1) \vee \hat{\mathcal{K}}_A\varphi_2$ iff $M, Q \models \hat{\mathcal{K}}_A\varphi_1 \rightarrow \hat{\mathcal{K}}_A\varphi_2$.

■

6.2. *Is $\mathbb{K}_a$ an Epistemic Operator?*

We believe that operators $\mathbb{C}_A$, $\mathbb{E}_A$, $\mathbb{D}_A$ and $\mathbb{K}_a$ do capture a special kind of knowledge of agents. An interesting question is: do this notion of knowledge have the properties usually associated with knowledge? In particular, do postulates $\mathbf{K}, \mathbf{D}, \mathbf{T}, \mathbf{4}, \mathbf{5}$ of epistemic logic hold for constructive knowledge? In general, the answer is *no*; particularly, the truth axiom does not hold.

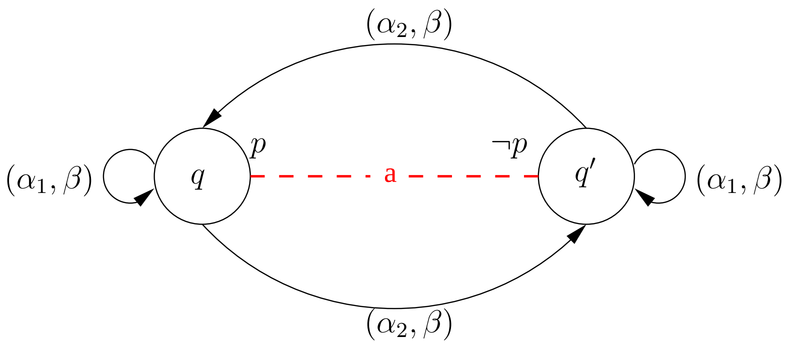

Figure 5. *Model M_6 with two agents a, b, and two states q, q' such that $q \sim_a q'$*

THEOREM 37. — *Below, we list the constructive knowledge versions of some of the S5 properties for individual agents. "Yes" means that the schema is strongly valid; "No" means that it is not even weakly valid (incidentally, none of the properties turns out to be weakly but not strongly valid).*

K	$\mathbb{K}_a(\varphi \to \psi) \to (\mathbb{K}_a\varphi \to \mathbb{K}_a\psi)$	*Yes*
D	$\neg\mathbb{K}_a\bot$	*Yes*
T	$\mathbb{K}_a\varphi \to \varphi$	*No*
4	$\mathbb{K}_a\varphi \to \mathbb{K}_a\mathbb{K}_a\varphi$	*Yes*
4$^+$	$\mathbb{K}_a\varphi \leftrightarrow \mathbb{K}_a\mathbb{K}_a\varphi$	*Yes*
5	$\neg\mathbb{K}_a\varphi \to \mathbb{K}_a\neg\mathbb{K}_a\varphi$	*Yes*
5$^+$	$\neg\mathbb{K}_a\varphi \leftrightarrow \mathbb{K}_a\neg\mathbb{K}_a\varphi$	*Yes*
B	$\varphi \to \mathbb{K}_a\neg\mathbb{K}_a\neg\varphi$	*No*

Before proving Theorem 37, we take a closer look at the relationship between satisfaction by a set of states ($M, Q \models \varphi$), and satisfaction in each of the states ($\forall_{q\in Q} M, q \models \varphi$). The following Lemma shows that the former does not necessarily imply the latter, and that the latter does not necessarily imply the former.

LEMMA 38. —

1) There is a model M, state q, agent a and formula φ such that $M, \mathrm{img}(q, \sim_a) \not\models \varphi$ and for every $q \in \mathrm{img}(q, \sim_a)$, $M, q \models \varphi$.

2) There are M, q, a, φ such that $M, \mathrm{img}(q, \sim_a) \models \varphi$ and $M, q \not\models \varphi$.

Proof. Consider model M_6 from Figure 5.

1) Let $\varphi = \langle\!\langle a \rangle\!\rangle \bigcirc p$. Now $M_6, q \models \varphi$ (a can choose action α_1), and $M_6, q' \models \varphi$ (a can choose action α_2). However, $M_6, \mathrm{img}(q, \sim_a) \not\models \varphi$, because no uniform strategy for a leads to q (in one step) from both q, q'.

2) Let $\varphi = \neg p$. Now $p \notin \pi(q) \cap \pi(q')$, so $M_6, \{q, q'\} \not\models p$, and $M_6, \mathrm{img}(q, \sim_a) \models \varphi$. But $p \in \pi(q)$, so $M_6, q \models p$, and $M_6, q \not\models \varphi$.

■

Proof of Theorem 37.

K: Immediate by Proposition 36.

D: Suppose that $M, Q \models \mathbb{K}_a \bot$ for any $Q \neq \emptyset$. Then $M, \mathrm{img}(Q, \sim_a)) \models \bot$. By reflexivity of $\sim_a$, set $\mathrm{img}(Q, \sim_a)$ is nonempty, which contradicts Proposition 16.1.

T: Let M, q, a, φ be as in Lemma 38.2. $M, q \models \mathbb{K}_a\varphi$, but $M, q \not\models \varphi$, so **T** is not weakly (and hence not strongly) valid.

$\mathbf{4}^+/\mathbf{4}$**:** $M, Q \models \mathbb{K}_a\mathbb{K}_a\varphi$ iff $M, \mathrm{img}(Q, \sim_a) \models \mathbb{K}_a\varphi$ iff $M, \mathrm{img}(\mathrm{img}(Q, \sim_a), \sim_a) \models \varphi$ iff $M, \mathrm{img}(Q, \sim_a) \models \varphi$ (since $\mathrm{img}(\mathrm{img}(Q, \sim_a), \sim_a) = \mathrm{img}(Q, \sim_a)$) iff $M, Q \models \mathbb{K}_a\varphi$.

$\mathbf{5}^+/\mathbf{5}$**:** $M, Q \models \neg\mathbb{K}_a\varphi$ iff $M, Q \not\models \mathbb{K}_a\varphi$ iff, by $\mathbf{4}^+$, $M, Q \not\models \mathbb{K}_a\mathbb{K}_a\varphi$ iff, by Proposition 36, $M, Q \models \mathbb{K}_a\neg\mathbb{K}_a\varphi$.

B: Let M, q, a, φ be as in Lemma 38.1. $M, \mathrm{img}(q, \sim_a) \not\models \varphi$, so $M, q \models \mathbb{K}_a\neg\varphi$. By $\mathbf{4}^+$, $M, q \models \mathbb{K}_a\mathbb{K}_a\neg\varphi$, so $M, q \not\models \neg\mathbb{K}_a\mathbb{K}_a\neg\varphi$, and by Proposition 36 $M, q \not\models \mathbb{K}_a\neg\mathbb{K}_a\neg\varphi$. But $M, q \models \varphi$. Thus, **B** is not weakly (nor strongly) valid.

■

6.3. *In Quest for the Truth Axiom*

We have just showed that, out of the S5 properties, axioms $\mathbf{K}, \mathbf{D}, \mathbf{4}, \mathbf{5}$ (but not $\mathbf{T}$!) hold. However, it also turns out that if we slightly restrict the language, then the corresponding $\mathbf{T}$ axiom becomes strongly valid. Let CSL$^-$ be the subset of CSL in which, between every occurrence of constructive knowledge ($\mathbb{C}_A, \mathbb{E}_A, \mathbb{D}_A$) and negation, there is always at least one operator other than conjunction.[9] Formally, CSL$^-$ formulae are defined by the following grammar (where $\mathcal{K} = C, E, D$ and $\hat{\mathcal{K}} = \mathbb{C}, \mathbb{E}, \mathbb{D}$):

$$\varphi ::= p \mid \neg\varphi \mid \varphi \wedge \varphi \mid \langle\langle A\rangle\rangle \bigcirc \varphi \mid \langle\langle A\rangle\rangle \Box \varphi \mid \langle\langle A\rangle\rangle \varphi \,\mathcal{U}\, \varphi \mid \mathcal{K}_A\varphi \mid \hat{\mathcal{K}}_A\psi,$$
$$\psi ::= p \mid \langle\langle A\rangle\rangle \bigcirc \varphi \mid \langle\langle A\rangle\rangle \Box \varphi \mid \langle\langle A\rangle\rangle \varphi \,\mathcal{U}\, \varphi \mid \mathcal{K}_A\varphi \mid \psi \wedge \psi \mid \hat{\mathcal{K}}_A\psi.$$

THEOREM 39. — *Every* CSL$^-$ *instance of* $\mathbf{T}$ *(i.e.,* $\mathbb{K}_a\psi \to \psi$*) is strongly valid.*

Proof in the Appendix.

9. In particular, the requirement is met when operators $\mathbb{C}_A, \mathbb{E}_A, \mathbb{D}_A$ are never immediately followed by either $\neg$ or $\wedge$.

Thus, the $\mathbf{T}$ axiom holds for CSL$^-$. Note that, by Proposition 36, the meaning of negation or conjunction in the immediate scope of a constructive knowledge operator is the same as if the operator were immediately outside the constructive knowledge operator.[10] In consequence, every formula of the full CSL is equivalent to one in CSL$^-$. Thus, we can restrict our logical language to CSL$^-$ without losing expressive power, and we automatically "get" axiom $\mathbf{T}$. We also observe that, from a more philosophical perspective, it is hard to pinpoint the intuitive meaning of negation immediately following constructive knowledge. Note that, e.g., $\mathbb{K}_a\neg\langle\langle a\rangle\rangle\varphi$ should be read as "a has *constructive* knowledge about being *unable* to achieve φ".[11] It seems thus, first, that the weaker version of the truth axiom in Theorem 39 might be more appropriate for constructive knowledge, and second, that it might be a good idea to consider the logical language of constructive knowledge to be limited to CSL$^-$. In this case, constructive knowledge has the $\mathbf{T}$ property, we do not lose any expressive power, and we leave out only formulae with philosophically unclear reading.

Is then the constructive knowledge in CSL$^-$ S5? First, it must be noted that – even though CSL and CSL$^-$ are expressively equivalent – the extension of the schema $\mathbf{T}$ is *different* in CSL$^-$ (for example, $\mathbb{K}_a\neg p \rightarrow \neg p$ is a CSL instance of $\mathbf{T}$, but even though it is equivalent to the CSL$^-$ formula $\neg\mathbb{K}_a p \rightarrow \neg p$, the latter is *not* a CSL$^-$ instance of $\mathbf{T}$). More importantly, in CSL$^-$ the axiom schemata $\mathbf{K}$ and $\mathbf{5}$, at least written as in Theorem 37, are not valid, but they are not invalid either – they are simply not formulae at all. It does not seem correct to say that an operator has the S5 properties when it cannot even *express* the $\mathbf{K}$ principle or negative introspection. Furthermore, CSL$^-$ lacks the S5 principle of *uniform substitution*.

6.4. *Properties of Collective Constructive Knowledge*

We briefly consider the properties of collective knowledge operators. Theorem 40 should come as no surprise: note that, analogously to standard knowledge, constructive common and distributed knowledge have the same properties as individual knowledge, while mutual knowledge ("everybody knows") differs in that it does not satisfy the introspection axioms $\mathbf{4}$ and $\mathbf{5}$.

THEOREM 40. — *Below, we list some of the S5 properties for collective constructive knowledge operators. We don't state the properties explicitly, but refer to Theorem 37 – axiom* $\mathbf{K}$ *for* $\mathbb{C}_A$ *becomes* $\mathbb{C}_A(\varphi \rightarrow \psi) \rightarrow (\mathbb{C}_A\varphi \rightarrow \mathbb{C}_A\psi)$*, and so on. "Yes" means that the schema is strongly valid; "No" means that it is not even weakly valid (the proof is left for the reader).*

---

10. Which is very much *unlike* the semantics of negation following a standard knowledge operator!
11. $\mathbb{K}_a\langle\langle a\rangle\rangle\neg\varphi$, on the other hand, makes perfect sense: it refers to a's constructive *ability* to prevent φ.

	$\mathbb{C}_A$	$\mathbb{E}_A$	$\mathbb{D}_A$
$\mathbf{K}$	*Yes*	*Yes*	*Yes*
$\mathbf{D}$	*Yes*	*Yes*	*Yes*
$\mathbf{T}$	*No*	*No*	*No*
$\mathbf{4}$	*Yes*	*No*	*Yes*
$\mathbf{4}^+$	*Yes*	*No*	*Yes*
$\mathbf{5}$	*Yes*	*No*	*Yes*
$\mathbf{5}^+$	*Yes*	*No*	*Yes*
$\mathbf{B}$	*No*	*No*	*No*

Note that the proof of Theorem 39 required only that the epistemic relation in question was reflexive. Thus, it can be easily extended to handle collective constructive knowledge.

COROLLARY 41. — *Every* CSL$^-$ *instance of schema* $\mathbf{T}$ *for collective constructive knowledge operators* $\mathbb{C}_A, \mathbb{E}_A, \mathbb{D}_A$ *is strongly valid.*

7. Negation, Localization, and Definability of Knowledge

The semantics of negation presented in Section 3.1 (we call it *weak* negation from now on) yields a very strong notion of disjunction, as Proposition 16 states. Such a strong notion of disjunction makes sense when we talk about agents' abilities, i.e., when used inside a $\mathbb{K}_a$ operator. For example: $M, q \models \mathbb{K}_a(\langle\langle a\rangle\rangle\varphi \vee \langle\langle a\rangle\rangle\psi)$ means in fact that a in q can either identify a plan to achieve φ or to achieve ψ. On the other hand, for a disjunction of simpler formulae, e.g., primitive propositions p and r, a weaker notion seems more intuitive: the disjunction $p \vee r$ should hold in M, Q iff, for any state $q \in Q$, at least one of the disjuncts p and r holds in q (but different disjuncts may hold in different states of Q). This intuition can be captured with a different negation operator $\sim$, which we call "strong" negation. The idea of strong negation can be summarized as: $M, Q \models {\sim}\varphi$ iff $M, q \not\models \varphi$ for every $q \in Q$. However, we will define it in terms of another, more primitive operator that we call *localization*.

As it turns out, the significance of localization goes beyond our discussion on various kinds of negation. Most importantly, localization can be used to define standard knowledge operators from constructive knowledge operators. On the other hand, localization itself proves definable from strategic and temporal operators. In consequence, standard knowledge can be defined in CSL without standard knowledge operators.

7.1. *Local Evaluation of Formulae*

In the semantics of CSL, formulae are interpreted in sets of states; in order for φ to hold in M, Q, the formula must be "globally" satisfied in all states from Q at once (i.e., with single evidence). Another option is to evaluate φ *locally* in particular states from Q. To this end, we introduce a modality that specifies explicitly that the formula must be evaluated for every relevant state separately:

$$M, Q \models \mathsf{loc}\,\varphi \text{ iff } M, q \models \varphi \text{ for every } q \in Q.$$

PROPOSITION 42. — *Below, we investigate some typical axioms with respect to the localization modality. "Yes" means that the scheme is strongly valid, "No" means that the scheme is not strongly valid. Note that all the schemes below are weakly valid, because $M, q \models \mathsf{loc}\,\varphi \leftrightarrow \varphi$ for every individual state q.*

K	$\mathsf{loc}\,(\varphi \rightarrow \psi) \rightarrow (\mathsf{loc}\,\varphi \rightarrow \mathsf{loc}\,\psi)$	*Yes*
D	$\neg\mathsf{loc}\,\bot$	*Yes*
T	$\mathsf{loc}\,\varphi \rightarrow \varphi$	*No*
4	$\mathsf{loc}\,\varphi \rightarrow \mathsf{loc}\,\mathsf{loc}\,\varphi$	*Yes*
$\mathbf{4}^+$	$\mathsf{loc}\,\varphi \leftrightarrow \mathsf{loc}\,\mathsf{loc}\,\varphi$	*Yes*
5	$\neg\mathsf{loc}\,\varphi \rightarrow \mathsf{loc}\,\neg\mathsf{loc}\,\varphi$	*No*
$\mathbf{5}^+$	$\neg\mathsf{loc}\,\varphi \leftrightarrow \mathsf{loc}\,\neg\mathsf{loc}\,\varphi$	*No*
B	$\varphi \rightarrow \mathsf{loc}\,\neg\mathsf{loc}\,\neg\varphi$	*Yes*

Proof in the Appendix.

Thus, localization is weak, but not strong, S5. In particular, S5 properties **T** and **5** do not necessarily hold in some contexts, for example in the immediate scope of a constructive knowledge operator.

PROPOSITION 43. — *Some other localization properties are the following, all strongly valid (proof is left for the reader).*

$\mathsf{loc}\,p \leftrightarrow p,\ p \in \Pi$ $\qquad$ $\mathsf{loc}\,(\varphi \wedge \psi) \leftrightarrow (\mathsf{loc}\,\varphi \wedge \mathsf{loc}\,\psi)$

$\langle\langle A\rangle\rangle \bigcirc \varphi \leftrightarrow \langle\langle A\rangle\rangle \bigcirc \mathsf{loc}\,\varphi$ $\qquad$ $\langle\langle A\rangle\rangle \Box \varphi \leftrightarrow \langle\langle A\rangle\rangle \Box \mathsf{loc}\,\varphi$

$\langle\langle A\rangle\rangle \varphi \mathcal{U}\, \psi \leftrightarrow \langle\langle A\rangle\rangle \mathsf{loc}\,\varphi \mathcal{U}\, \psi \leftrightarrow \langle\langle A\rangle\rangle \varphi \mathcal{U}\, \mathsf{loc}\,\psi$

$\mathsf{loc}\,\mathcal{K}_A\varphi \leftrightarrow \mathcal{K}_A\varphi$ $\qquad$ $\mathcal{K}_A\varphi \leftrightarrow \mathcal{K}_A\mathsf{loc}\,\varphi,\ \mathcal{K} \in \{C, E, D\}$

We will show in the following sections how the loc operator can be used to define standard knowledge and alternative negation operators. This makes the following result very important: it says that localization is definable in the CSL language, from the $\langle\langle\emptyset\rangle\rangle$ and $\mathcal{U}$ operators.

PROPOSITION 44. — *The following formula is strongly valid:*

$$\mathsf{loc}\,\varphi \leftrightarrow \langle\langle\emptyset\rangle\rangle \varphi \mathcal{U}\, \varphi$$

Proof. $M, Q \models \langle\langle\emptyset\rangle\rangle \varphi \mathcal{U}\, \varphi$ iff $\forall_{\lambda \in out(Q,\emptyset)}$ there is an $i \geq 0$ such that $M, \lambda[i] \models \varphi$ and for any j such that $0 \leq j < i$, $M, \lambda[j] \models \varphi$. Since for each $q \in Q$ there is a $\lambda \in out(Q, \emptyset)$ with $\lambda[0] = q$, this implies that $\forall_{q \in Q} M, q \models \varphi$ which is the same as $M, Q \models \mathsf{loc}\ \varphi$. To see that the other direction holds as well, assume that $M, Q \models \mathsf{loc}\ \varphi$ and let $\lambda \in out(Q, \emptyset)$. We must provide a witness for i; take $i = 0$. Now, $M, \lambda[i] \models \varphi$ and there is no j such that $0 \leq j < i$, so $M, Q \models \langle\langle\emptyset\rangle\rangle \varphi \mathcal{U} \varphi$. ∎

7.2. *Defining Standard Knowledge from Constructive Knowledge*

Standard knowledge operators are definable from constructive knowledge and localization:

PROPOSITION 45. — *$\mathcal{K}_A\varphi \leftrightarrow \hat{\mathcal{K}}_A \mathsf{loc}\ \varphi$ is strongly valid for any $\mathcal{K} \in \{C, E, D\}$, $\hat{C} = \mathbb{C}$, $\hat{E} = \mathbb{E}$, $\hat{D} = \mathbb{D}$.*

Proof. $M, Q \models \hat{\mathcal{K}}_A \mathsf{loc}\ \varphi$ iff $M, \mathrm{img}(Q, \sim_A^{\mathcal{K}}) \models \mathsf{loc}\ \varphi$ iff $\forall_{q \in \mathrm{img}(Q, \sim_A^{\mathcal{K}})} M, q \models \varphi$ iff $M, Q \models \mathcal{K}_A\varphi$. ∎

In particular, knowledge of a formula is the same as constructive knowledge of the localization of the formula, i.e. $K_a\varphi \leftrightarrow \mathbb{K}_a \mathsf{loc}\ \varphi$. An important corollary of Propositions 45 and 44 is the following.

THEOREM 46. — *The following is strongly valid:*

$$\mathcal{K}_A\varphi \leftrightarrow \hat{\mathcal{K}}_A \langle\langle\emptyset\rangle\rangle \varphi \mathcal{U} \varphi.$$

Theorem 46 shows that *standard knowledge can be seen as a special case of constructive knowledge*. It follows that the standard knowledge operators are strictly speaking redundant in the CSL language.

7.3. *Non-standard Definitions of Negation*

Negation, as defined in Section 3.1, is "weak" in the sense that it is sufficient for the negation of, e.g., an atomic formula p to hold in a set of states Q that p is false in at least one state from Q. Several other interpretations of negation in a set of states are possible, corresponding to different ways of quantifying over the set. We define *strong negation* as:

$$\sim\varphi \equiv \mathsf{loc}\ \neg\varphi$$

Note that, by Proposition 44, strong negation is definable from weak negation: $\sim\varphi$ can be equivalently defined as $\langle\langle\emptyset\rangle\rangle(\neg\varphi)\mathcal{U}(\neg\varphi)$.

PROPOSITION 47. — *$M, Q \models \sim\varphi$ iff, for every $q \in Q$, we have that $M, q \not\models \varphi$.*

Proof. $M, Q \models \sim\varphi$ iff $M, Q \models \mathsf{loc}\ \neg\varphi$ iff for every $q \in Q$ we have that $M, q \not\models \varphi$. ∎

Strong negation does *not* behave as classical negation: it does not obey the law of double negation, the law of excluded middle, or the consistency requirement under *strong* validity. Nevertheless, it preserves these laws under weak validity.

PROPOSITION 48. —

1) $\sim\sim\varphi \rightarrow \varphi$ *is weakly valid, but not strongly valid.*

2) $\varphi \rightarrow \sim\sim\varphi$ *is weakly valid, but not strongly valid.*

3) $\varphi \vee \sim\varphi$ *is weakly valid, but not strongly valid.*

4) $\neg(\varphi \wedge \sim\varphi)$ *is weakly valid, but not strongly valid.*

Proof.

1), 2) Weak validity is immediate. $M, Q \models \sim\sim\varphi$ iff $M, Q \models \mathsf{loc}\,\neg \sim\varphi$ iff $M, Q \models \mathsf{loc}\,\neg\mathsf{loc}\,\neg\varphi$ iff for every $q \in Q$ we have that $M, q \models \varphi$. Counter-examples for the two implications are found in the two parts of Lemma 38, respectively, by taking $Q = \mathrm{img}(q, \sim_a)$.

3) Weak validity is immediate. As a counter-example to strong validity, take M and φ from Lemma 38.1, and let $Q = \mathrm{img}(q, \sim_a)$. $M, Q \not\models \varphi$, and it is not the case that $M, q' \not\models \varphi$ for every $q' \in Q$.

4) Weak validity: immediate. Strong validity: take $M = M_6$ from Lemma 38, and let $Q = \{q, q'\}$, $\varphi \equiv \neg\langle\langle a \rangle\rangle \bigcirc p$.

■

REMARK 49. — Alternatively, strong negation can be taken as a primary notion: localization is definable from strong negation, and standard knowledge is thus definable from constructive knowledge and strong negation. Formally, the following are strongly valid:

1) $\mathsf{loc}\,\varphi \leftrightarrow \sim\sim\varphi$

2) $\mathcal{K}_A\varphi \leftrightarrow \hat{\mathcal{K}}_A \sim\sim\varphi$.

□

7.3.1. *Boolean Operators Based on Strong Negation*

Recall that connectives like $\vee$ and $\rightarrow$ are defined in terms of weak negation ($\neg$). Similar connectives can be defined for strong negation:

- $\varphi_1 \parallel \varphi_2 \equiv \sim(\sim\varphi_1 \wedge \sim\varphi_2)$,
- $\varphi_1 \rightsquigarrow \varphi_2 \equiv \sim\varphi_1 \parallel \varphi_2$, and
- $\varphi_1 \leftrightsquigarrow \varphi_2 \equiv (\varphi_1 \rightsquigarrow \varphi_2) \wedge \varphi_2 \rightsquigarrow \varphi_1$.

These versions of disjunction, material implication, and material biconditional have the following semantic characterizations:

PROPOSITION 50. —

1) $M, Q \models \varphi_1 \parallel \varphi_2$ *iff, for every* $q \in Q$*, we have* $M, q \models \varphi_1$ *or* $M, q \models \varphi_2$*;*

2) $M, Q \models \varphi_1 \rightsquigarrow \varphi_2$ *iff, for every* $q \in Q$*, we have that* $M, q \models \varphi_1$ *implies* $M, q \models \varphi_2$*;*

3) $M, Q \models \varphi_1 \leftrightsquigarrow \varphi_2$ *iff, for every* $q \in Q$*, we have that* $M, q \models \varphi_1$ *iff* $M, q \models \varphi_2$*.*

Proof.

1) $M, Q \models \sim(\sim\varphi_1 \wedge \sim\varphi_2)$ iff $\forall_{q \in Q} M, q \not\models \sim\varphi_1 \wedge \sim\varphi_2$ iff $\forall_{q \in Q} M, q \not\models \sim\varphi_1$ or $M, q \not\models \sim\varphi_2$ iff $\forall_{q \in Q} M, q \models \varphi_1$ or $M, q \models \varphi_2$.

2) $M, Q \models \varphi_1 \rightsquigarrow \varphi_2$ iff $M, Q \models \sim\varphi_1 \parallel \varphi_2$ iff $\forall_{q \in Q}(M, q \models \sim\varphi_1$ or $M, q \models \varphi_2)$ iff $\forall_{q \in Q}(M, q \not\models \varphi_1$ or $M, q \models \varphi_2)$ iff $\forall_{q \in Q} M, q \models \varphi_1 \rightarrow \varphi_2$.

3) Straightforward.

■

We can also define the strong negation-based versions of Boolean constants "true" and "false", but they coincide with the ones already proposed in Section 3.2.

PROPOSITION 51. — *Let* $\underline{\wr} \equiv p \wedge \sim p$*, and* $\overline{\wr} \equiv \sim \underline{\wr}$*. Then:*

1) $M, Q \not\models \underline{\wr}$ *for all* $Q \subseteq St$*,* $Q \neq \emptyset$*.*

2) $M, Q \models \overline{\wr}$ *for all* $Q \subseteq St$*,* $Q \neq \emptyset$*.*

Proof. Straightforward. ■

7.3.2. *Some Connections Between the Weak and the Strong*

It is immediate from Proposition 50 that, just as strong negation is the localization of weak negation, the operators $\parallel$, $\rightsquigarrow$ and $\leftrightsquigarrow$ defined by strong negation, are the localizations of their counterparts $\vee$, $\rightarrow$, $\leftrightarrow$ defined by weak negation:

PROPOSITION 52. — *The following are strongly valid:*

$$(\varphi_1 \parallel \varphi_2) \leftrightarrow \mathsf{loc}\,(\varphi_1 \vee \varphi_2)$$
$$(\varphi_1 \rightsquigarrow \varphi_2) \leftrightarrow \mathsf{loc}\,(\varphi_1 \rightarrow \varphi_2)$$
$$(\varphi_1 \leftrightsquigarrow \varphi_2) \leftrightarrow \mathsf{loc}\,(\varphi_1 \leftrightarrow \varphi_2)$$

Moreover, for validity (not strong validity), the two negations, the two disjunctions and the two implications coincide:

PROPOSITION 53. — *The following formulae are valid (but not strongly valid):*

1) $\neg\varphi \leftrightarrow \sim\varphi$

2) $(\varphi_1 \vee \varphi_2) \leftrightarrow (\varphi_1 \parallel \varphi_2)$

3) $(\varphi_1 \rightarrow \varphi_2) \leftrightarrow (\varphi_1 \rightsquigarrow \varphi_2)$

Proof. Immediate from Proposition 52, since $M, q \models \psi$ iff $M, q \models \mathsf{loc}\,\psi$, for any (single) state q. ■

The following proposition shows that the notions of strong and weak validity can be seen as dual with respect to the strong and weak versions of the connectives.

PROPOSITION 54. —

1) $\sim\varphi$ *is strongly valid iff* $\neg\varphi$ *is weakly valid.*

2) $\varphi_1 \parallel \varphi_2$ *is strongly valid iff* $\varphi_1 \vee \varphi_2$ *is weakly valid.*

3) $\varphi_1 \rightsquigarrow \varphi_2$ *is strongly valid iff* $\varphi_1 \rightarrow \varphi_2$ *is weakly valid.*

4) $\varphi_1 \leftrightsquigarrow \varphi_2$ *is strongly valid iff* $\varphi_1 \leftrightarrow \varphi_2$ *is weakly valid.*

The laws of negation were stated in Proposition 48 using connectives $\vee$, etc., defined from weak negation. We can now show, however, that the laws of negation do in fact hold for strong negation if we state these laws using the operators defined from strong negation.

PROPOSITION 55. —

1) $\sim\sim\varphi \leftrightsquigarrow \varphi$ *is strongly valid.*

2) $\varphi \parallel \sim\varphi$ *is strongly valid.*

3) $\sim(\varphi\wedge \sim\varphi)$ *is strongly valid.*

Proof. Immediate from Propositions 48 and 54. ■

7.3.3. *Properties of Constructive Knowledge with "Strong" Negation*

In Section 6.2, we discussed the S5 properties of constructive knowledge. These properties can also be stated using strong negation, and derived connectives, instead of weak negation.

THEOREM 56. — *Below, we list constructive knowledge versions of some S5 properties using strong negation. "Yes" means that the schema is strongly valid; "No" means that it is not even weakly valid (again, none of the properties turn out to be weakly but not strongly valid).*

$\widetilde{\mathbf{K}}$	$\mathbb{K}_a(\varphi \rightsquigarrow \psi) \rightsquigarrow (\mathbb{K}_a\varphi \rightsquigarrow \mathbb{K}_a\psi)$	*No*
$\widetilde{\mathbf{D}}$	$\sim\mathbb{K}_a \underline{\wr}$	*Yes*
$\widetilde{\mathbf{T}}$	$\mathbb{K}_a\varphi \rightsquigarrow \varphi$	*No*
$\widetilde{\mathbf{4}}$	$\mathbb{K}_a\varphi \rightsquigarrow \mathbb{K}_a\mathbb{K}_a\varphi$	*Yes*
$\widetilde{\mathbf{4}}^+$	$\mathbb{K}_a\varphi \leftrightsquigarrow \mathbb{K}_a\mathbb{K}_a\varphi$	*Yes*
$\widetilde{\mathbf{5}}$	$\sim\mathbb{K}_a\varphi \rightsquigarrow \mathbb{K}_a\sim\mathbb{K}_a\varphi$	*Yes*
$\widetilde{\mathbf{5}}^+$	$\sim\mathbb{K}_a\varphi \leftrightsquigarrow \mathbb{K}_a\sim\mathbb{K}_a\varphi$	*Yes*
$\widetilde{\mathbf{B}}$	$\varphi \rightsquigarrow \mathbb{K}_a\sim\mathbb{K}_a\sim\varphi$	*Yes*

Proof in the Appendix.

Finally, we point out that if we restrict the language to CSL$^-$, as discussed in Section 6.3, we get the truth axiom $\widetilde{\mathbf{T}}$, i.e., the following variant of Theorem 39.

THEOREM 57. — *Every* CSL$^-$ *instance of schema* $\widetilde{\mathbf{T}}$ *(*$\mathbb{K}_a\varphi \rightsquigarrow \varphi$*) is strongly valid.*

Proof. Note that $\forall_{q\in Q} M, q \models \mathbb{K}_a\varphi \to \varphi$ (by $\mathbf{T}$), which implies that $M, Q \models \mathbb{K}_a\varphi \rightsquigarrow \varphi$ (by Proposition 50). ■

7.3.4. *Other Negations*

We have considered two operators for negation so far. Yet another alternative is: $\angle\varphi \equiv \neg\mathsf{loc}\,\varphi$. The meaning of $\angle$ is characterized with the following proposition.

PROPOSITION 58. — $M, Q \models \angle\varphi$ *iff there exists* $q \in Q$ *such that* $M, q \not\models \varphi$*.*

Proof. $M, Q \models \angle\varphi$ iff $M, Q \models \neg\mathsf{loc}\,\varphi$ iff $M, Q \not\models \mathsf{loc}\,\varphi$ iff there is a $q \in Q$ such that $M, q \not\models \varphi$. ■

8. Normal Forms and Expressiveness

In this section, we investigate expressiveness further, with particular focus on the relationship between localization, weak negation and strong negation. In order to study expressiveness, we will study variants of the language defined in Section 3.1 with other (primary) operators. We have discussed the interpretation of the following operators in sets of states:

$$\neg \quad \wedge \quad \langle\!\langle A\rangle\!\rangle T \quad C_A \quad E_A \quad D_A \quad \mathbb{C}_A \quad \mathbb{E}_A \quad \mathbb{D}_A \quad \mathsf{loc} \quad \sim$$

where T is an ATL temporal connective and A is a set of agents. We use the expression $\mathcal{L}(\neg, \wedge, \langle\!\langle A\rangle\!\rangle T, \mathcal{K}_A, \hat{\mathcal{K}}_A, \mathsf{loc}\ , \sim)$ to denote the language with all the mentioned operators, $\mathcal{L}(\neg, \wedge, \langle\!\langle A\rangle\!\rangle T, \mathcal{K}_A, \hat{\mathcal{K}}_A, \mathsf{loc}\ )$ to denote the language with all operators except strong negation, and so on. The CSL language introduced in Section 3.1 is $\mathcal{L} = \mathcal{L}(\neg, \wedge, \langle\!\langle A\rangle\!\rangle T, \mathcal{K}_A, \hat{\mathcal{K}}_A)$. For simplicity, we sometimes use $\mathcal{L}^*$ for the most extensive language $\mathcal{L}(\neg, \wedge, \langle\!\langle A\rangle\!\rangle T, \mathcal{K}_A, \hat{\mathcal{K}}_A, \mathsf{loc}\ , \sim)$.

We say that two formulae φ and ψ are *equivalent*, if $\varphi \leftrightarrow \psi$ is valid, and that they are *strongly equivalent* if $\varphi \leftrightarrow \psi$ is strongly valid. We say that a language $\mathcal{L}_2$ is *at least as expressive* as a language $\mathcal{L}_1$, if for every $\varphi_1 \in \mathcal{L}_1$ there exist an equivalent $\varphi_2 \in \mathcal{L}_2$. We say that $\mathcal{L}_2$ and $\mathcal{L}_1$ are *expressively equivalent*, if $\mathcal{L}_2$ is at least as expressive as $\mathcal{L}_1$ and $\mathcal{L}_1$ is at least as expressive as $\mathcal{L}_2$.

We will make use of the following definition:

$$Atoms = \Theta \cup \{\langle\!\langle A\rangle\!\rangle T\gamma : \gamma \in \mathcal{L}^*\} \cup \{\sim\gamma : \gamma \in \mathcal{L}^*\}.$$

We begin with defining a *normal form* of our formulae.

8.1. *Constructive Normal Form*

A formula, possibly containing strong negation, is in *constructive normal form* if every subformula starting with a $\hat{\mathcal{K}}_A$ operator is of the form $\hat{\mathcal{K}}_1 \cdots \hat{\mathcal{K}}_k \psi$ where ψ is either a primitive proposition, starts with a cooperation modality, or starts with strong negation. We now show that every $\mathcal{L}^*$ formula is equivalent to one of constructive normal form, and also to a formula of constructive normal form without strong negation.

DEFINITION 59 (CONSTRUCTIVE NORMAL FORM (CSNF)). — *The set of $\mathcal{L}^*$ formulae of* constructive normal form (CSNF) *is defined inductively as follows.*

– *p is of CSNF when $p \in \Theta$,*

– *$\hat{\mathcal{K}}\gamma$ is of CSNF iff γ is of CSNF and either $\gamma \in Atoms$ or $\gamma = \hat{\mathcal{K}}'\chi$,*

– *$\langle\langle G\rangle\rangle T\gamma$ is of CSNF iff γ is of CSNF,*

– *$\neg\gamma$ is of CSNF iff γ is of CSNF,*

– *$\gamma_1 \wedge \gamma_2$ is of CSNF iff both γ_1 and γ_2 are of CSNF,*

– *$\sim\gamma$ is of CSNF iff γ is of CSNF.*

THEOREM 60. — *Every formula in $\mathcal{L}^*$ is strongly equivalent to a formula in constructive normal form.*

Proof in the Appendix.

Thus, any formula of the most general kind we have considered is equivalent to a formula of CSNF. Note that a CSNF formula might contain strong negation. However, we can also get rid of strong negation, as the following result states.

COROLLARY 61. — *Every formula in $\mathcal{L}^*$ is strongly equivalent to a formula in constructive normal form without strong negation.*

Proof in the Appendix.

8.2. *Expressiveness of Strong Negation*

We have shown in Section 7 that standard knowledge, localization and strong negation can be defined with use of weak negation (together with conjunction, constructive knowledge and ATL operators). Thus, $\mathcal{L}(\neg, \wedge, \langle\langle A\rangle\rangle T, \hat{\mathcal{K}})$ is already as expressive as the full $\mathcal{L}^*$. Now we will investigate the other direction: does *weak* negation add expressiveness if we already have strong negation? We show in the following theorem that, in the language $\mathcal{L}$ extended with strong negation, every formula is actually equivalent to one without weak negation.

THEOREM 62. — *Every formula in $\mathcal{L}(\neg, \wedge, \langle\langle A\rangle\rangle T, \hat{\mathcal{K}}, \sim)$ is equivalent to a formula of $\mathcal{L}(\wedge, \langle\langle A\rangle\rangle T, \hat{\mathcal{K}}, \sim)$.*

Proof in the Appendix.

Thus, in particular, the following four languages are expressively equivalent:

$$\mathcal{L}(\neg, \wedge, \langle\!\langle A \rangle\!\rangle T, \hat{\mathcal{K}}) \quad \mathcal{L}(\wedge, \langle\!\langle A \rangle\!\rangle T, \hat{\mathcal{K}}, \sim) \qquad \mathcal{L} \qquad \mathcal{L}^*$$

In consequence, both $\mathcal{L}(\neg, \wedge, \langle\!\langle A \rangle\!\rangle T, \hat{\mathcal{K}})$ and $\mathcal{L}(\wedge, \langle\!\langle A \rangle\!\rangle T, \hat{\mathcal{K}}, \sim)$ are expressively complete with respect to the other operators we have considered. An important difference between $\mathcal{L}(\neg, \wedge, \langle\!\langle A \rangle\!\rangle T, \hat{\mathcal{K}})$ and $\mathcal{L}(\wedge, \langle\!\langle A \rangle\!\rangle T, \hat{\mathcal{K}}, \sim)$ is that strong negation is definable from weak negation and ATL operators by a simple schema ($\langle\!\langle \emptyset \rangle\!\rangle \neg\varphi \, \mathcal{U} \, \neg\varphi$), while this is not the case when we reverse the roles of the negations.

9. Conclusions

In this paper, we propose a non-standard semantics for the modal logic of strategic ability under imperfect information, in which formulae are interpreted over *sets of states* rather than in single states.[12] Moreover, we introduce new epistemic operators for "constructive" knowledge. It turns out that, in this new semantics, simple cooperation modalities $\langle\!\langle A \rangle\!\rangle$ can be combined with "constructive" epistemic operators into sufficiently expressive formulae. Indeed, the new logic *is strictly more expressive than most existing* ATL *versions for imperfect information*, while it retains the same model checking complexity as the least costly of them. The philosophical dimension of constructive knowledge is also natural: the constructive knowledge operators capture the notion of *knowing "de re"*, while the standard epistemic operators refer to *knowing "de dicto"*. Moreover, it turns out that standard (traditional) knowledge is a special case of constructive knowledge. Also, the language of CSL is expressive enough to enable expressing several other interesting operators in a simple way.

Most of the usual S5 properties (with the notable exception of the truth axiom T) hold for constructive knowledge. Furthermore, if we slightly restrict the syntax of CSL, we do not lose expressive power and the schema T becomes a validity.

CSL has novel, meaningful epistemic operators that can be used to capture important properties of the interaction between knowledge, action and ability. In future work, we plan to investigate further the expressivity of CSL, and its relationship with logics like ETSL, ATL_{iR}, ATEL-R*, ATEL-A, and "Uniform STIT". A good case study (together with a more detailed analysis of verification complexity) is essential to determine the applicability of the logic. Also, the (relative) expressive power of various operators in our semantics seems to be worth further study.

We thank anonymous reviewers of JANCL and AAMAS-06 for their helpful remarks. Thomas Ågotnes' work has been supported by the Research Council of Norway under grant 166525/V30. Wojtek Jamroga would also like to thank Jan Broersen, John-Jules Meyer and Wiebe van der Hoek.

12. We emphasize again that we do not propose new models: concurrent epistemic game structures have already been used for several years in ATEL-like logics. What we propose is a new interpretation of formulae.

10. References

[ÅGO 04] ÅGOTNES T., "A note on Syntactic Characterization of Incomplete Information in ATEL", *Procedings of the Workshop on Knowledge and Games, Liverpool*, 2004, p. 34–42.

[ÅGO 06] ÅGOTNES T., "Action and Knowledge in Alternating-time Temporal Logic", *Synthese*, vol. 149, num. 2, 2006, p. 377–409, Section on Knowledge, Rationality and Action.

[ALU 97] ALUR R., HENZINGER T. A., KUPFERMAN O., "Alternating-Time Temporal Logic", *Proceedings of the 38th Annual Symposium on Foundations of Computer Science (FOCS)*, IEEE Computer Society Press, 1997, p. 100–109.

[ALU 98] ALUR R., HENZINGER T. A., KUPFERMAN O., "Alternating-Time Temporal Logic", *Lecture Notes in Computer Science*, vol. 1536, 1998, p. 23–60.

[ALU 02] ALUR R., HENZINGER T. A., KUPFERMAN O., "Alternating-Time Temporal Logic", *Journal of the ACM*, vol. 49, 2002, p. 672–713.

[BEL 88] BELNAP N., PERLOFF M., "Seeing to It That: a Canonical Form for Agentives", *Theoria*, vol. 54, 1988, p. 175–199.

[BRO 06] 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, 2006, p. 559–578.

[CLA 81] CLARKE E., EMERSON E., "Design and Synthesis of Synchronization Skeletons Using Branching Time Temporal Logic", *Proceedings of Logics of Programs Workshop*, vol. 131 of *Lecture Notes in Computer Science*, 1981, p. 52–71.

[CLA 86] CLARKE E., EMERSON E., SISTLA A., "Automatic Verification of Finite-State Concurrent Systems Using Temporal Logic Specifications", *ACM Transactions on Programming Languages and Systems*, vol. 8, num. 2, 1986, p. 244–263.

[EME 90] EMERSON E. A., "Temporal and Modal Logic", VAN LEEUWEN J., Ed., *Handbook of Theoretical Computer Science*, vol. B, p. 995–1072, Elsevier Science Publishers, 1990.

[FIS 06] FISHER M., *Temporal Logics*, Kluwer, 2006.

[GOR 01] GORANKO V., "Coalition Games and Alternating Temporal Logics", VAN BENTHEM J., Ed., *Proceedings of TARK VIII*, Morgan Kaufmann, 2001, p. 259–272.

[GOR 04] GORANKO V., JAMROGA W., "Comparing Semantics of Logics for Multi-agent Systems", *Synthese*, vol. 139, num. 2, 2004, p. 241–280.

[HAL 95] HALPERN J., "Reasoning about Knowledge: a Survey", *Handbook of Logic in Artificial Intelligence and Logic Programming. Vol. 4: Epistemic and Temporal Reasoning*, p. 1–34, Oxford University Press, Oxford, 1995.

[HAR 00] HAREL D., KOZEN D., TIURYN J., *Dynamic Logic*, MIT Press, 2000.

[HER 06] HERZIG A., TROQUARD N., "Knowing How to Play: Uniform Choices in Logics of Agency", *Proceedings of AAMAS'06*, 2006, p. 209–216.

[HOE 02] VAN DER HOEK W., WOOLDRIDGE M., "Tractable Multiagent Planning for Epistemic Goals", CASTELFRANCHI C., JOHNSON W., Eds., *Proceedings of the First International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS-02)*, ACM Press, New York, 2002, p. 1167–1174.

[HOE 03] VAN DER HOEK W., WOOLDRIDGE M., "Cooperation, Knowledge and Time: Alternating-time Temporal Epistemic Logic and its Applications", *Studia Logica*, vol. 75, num. 1, 2003, p. 125–157.

[HOE 06] VAN DER HOEK W., LOMUSCIO A., WOOLDRIDGE M., "On the Complexity of Practical ATL Model Checking", STONE P., WEISS G., Eds., *Proceedings of AAMAS'06*, 2006, p. 201–208.

[JAM 03] JAMROGA W., "Some Remarks on Alternating Temporal Epistemic Logic", DUNIN-KEPLICZ B., VERBRUGGE R., Eds., *Proceedings of Formal Approaches to Multi-Agent Systems (FAMAS 2003)*, 2003, p. 133–140.

[JAM 04] JAMROGA W., VAN DER HOEK W., "Agents that Know how to Play", *Fundamenta Informaticae*, vol. 63, num. 2–3, 2004, p. 185–219, IOS Press : Amsterdam.

[JAM 05a] JAMROGA W., DIX J., "Do Agents Make Model Checking Explode (Computationally)?", PĚCHOUČEK M., PETTA P., VARGA L., Eds., *Proceedings of CEEMAS 2005*, vol. 3690 of *Lecture Notes in Computer Science*, p. 398–407, Springer Verlag, 2005.

[JAM 05b] JAMROGA W., ÅGOTNES T., "Constructive Knowledge: What Agents Can Achieve under Incomplete Information", report num. IfI-05-10, 2005, Clausthal University of Technology.

[JAM 05c] JAMROGA W., VAN DER HOEK W., "Strategic Ability under Uncertainty", report num. IfI-05-06, 2005, Clausthal University of Technology.

[JAM 06a] JAMROGA W., "On the Relationship between Playing Rationally and Knowing how to Play: A Logical Account", FREKSA C., KOHLHASE M., SCHILL K., Eds., *Proceedings of KI 2006*, vol. 4314 of *Lecture Notes in Artificial Intelligence*, p. 419–433, Springer Verlag, 2006.

[JAM 06b] JAMROGA W., DIX J., "Model Checking ATL_{ir} is Indeed Δ_2^P-complete", *Proceedings of EUMAS'06*, 2006.

[JAM 06c] JAMROGA W., ÅGOTNES T., "What Agents Can Achieve under Incomplete Information", *Proceedings of AAMAS'06*, ACM Press, 2006, p. 232–234, Short paper.

[JAM 07] JAMROGA W., DIX J., "Model Checking Abilities of Agents: A Closer Look", *Theory of Computing Systems*, , 2007, To appear.

[JON 03] JONKER G., "Feasible Strategies in Alternating-time Temporal Epistemic Logic", Master thesis, University of Utrecht, 2003.

[LAR 06] LAROUSSINIE F., MARKEY N., OREIBY G., "Expressiveness and Complexity of ATL", report num. LSV-06-03, 2006, CNRS & ENS Cachan, France.

[MOO 85] MOORE R., "A Formal Theory of Knowledge and Action", HOBBS J., MOORE R., Eds., *Formal Theories of the Commonsense World*, Ablex Publishing Corp., 1985.

[MOR 91] MORGENSTERN L., "Knowledge and the Frame Problem", *International Journal of Expert Systems*, vol. 3, num. 4, 1991.

[NAS 50] NASH J., "Equilibrium points in n-person games", *Proceedings of the National Academy of Sciences U.S.A.*, vol. 36, 1950, p. 48–49.

[NEU 44] VON NEUMANN J., MORGENSTERN O., *Theory of Games and Economic Behaviour*, Princeton University Press: Princeton, NJ, 1944.

[OSB 94] OSBORNE M., RUBINSTEIN A., *A Course in Game Theory*, MIT Press, 1994.

[OTT 04] VAN OTTERLOO S., JONKER G., "On Epistemic Temporal Strategic Logic", *Electronic Notes in Theoretical Computer Science*, vol. XX, 2004, p. 35–45, Proceedings of LCMAS'04.

[PAU 00] PAULY M., "Game Logic for Game Theorists", report num. INS-R0017, 2000, CWI.

[PAU 02] PAULY M., "A Modal Logic for Coalitional Power in Games", *Journal of Logic and Computation*, vol. 12, num. 1, 2002, p. 149–166.

[PRI 67] PRIOR A., *Past, Present and Future*, Clarendon Press : Oxford, 1967.

[QUI 56] QUINE W., "Quantifiers and Propositional Attitudes", *Journal of Philosophy*, vol. 53, 1956, p. 177–187.

[SCH 04] SCHOBBENS P. Y., "Alternating-Time Logic with Imperfect Recall", *Electronic Notes in Theoretical Computer Science*, vol. 85, num. 2, 2004, Elsevier.

[TRO 91] TROELSTRA A. S., "History of Constructivism in the Twentieth Century", report num. ML-91-05, 1991, University of Amsterdam.

[WöL 04] WÖLFL S., "Qualitative Action Theory: A Comparison of the Semantics of Alternating-time Temporal Logic and the Kutschera-Belnap Approach to Agency", *Proceedings of JELIA'04*, vol. 3229 of *Lecture Notes in Artificial Intelligence*, p. 70–81, Springer Verlag, 2004.

[WOO 00] WOOLDRIDGE M., *Reasoning about Rational Agents*, MIT Press : Cambridge, Mass, 2000.

Appendix: Some Proofs

Theorem 23

Proof of Theorem 23 (structural induction wrt the structure of φ).

– $M, q \models_{\text{CSL}} tr(p)$ iff $M, q \models_{\text{CSL}} p$ iff $p \in \pi(q)$ iff $M, q \models_{\mathcal{L}} p$.

– $M, q \models_{\text{CSL}} tr(\neg\varphi)$ iff $M, q \not\models_{\text{CSL}} tr(\varphi)$ iff (by induction) $M, q \not\models_{\mathcal{L}} \varphi$ iff $M, q \models_{\mathcal{L}} \neg\varphi$.

– $M, q \models_{\text{CSL}} tr(\varphi \wedge \psi)$ iff $M, q \models_{\text{CSL}} tr(\varphi)$ and $M, q \models_{\text{CSL}} tr(\psi)$ iff (by induction) $M, q \models_{\mathcal{L}} \varphi$ and $M, q \models_{\mathcal{L}} \psi$ iff $M, q \models_{\mathcal{L}} \varphi \wedge \psi$.

– $M, q \models_{\text{CSL}} tr(\langle\!\langle A \rangle\!\rangle_{\mathcal{K}(\Gamma)} \bigcirc \varphi)$ iff $M, q \models_{\text{CSL}} \hat{\mathcal{K}}_\Gamma \langle\!\langle A \rangle\!\rangle \bigcirc tr(\varphi)$ iff $M, \text{img}(q, \sim_\Gamma^{\mathcal{K}}) \models_{\text{CSL}} \langle\!\langle A \rangle\!\rangle \bigcirc tr(\varphi)$ iff $\exists_{S_A} \forall_{\Lambda \in out(\text{img}(q, \sim_\Gamma^{\mathcal{K}}), S_A)} M, \Lambda[1] \models_{\text{CSL}} tr(\varphi)$ iff (by induction) $\exists_{S_A} \forall_{\Lambda \in out(\text{img}(q, \sim_\Gamma^{\mathcal{K}}), S_A)} M, \Lambda[1] \models_{\text{ATOL}} \varphi$ iff $M, q \models_{\text{ATOL}} \langle\!\langle A \rangle\!\rangle_{\mathcal{K}(\Gamma)} \bigcirc \varphi$.

– For $\langle\!\langle A \rangle\!\rangle_{\mathcal{K}(\Gamma)} \Box \varphi$ and $\langle\!\langle A \rangle\!\rangle_{\mathcal{K}(\Gamma)} (\varphi \mathcal{U} \psi)$: analogously. The same for $\langle\!\langle A \rangle\!\rangle_{ir} \varphi$, $\langle\!\langle A \rangle\!\rangle_{\mathcal{K}}^f \varphi$, and $\langle\!\langle A \rangle\!\rangle_{K_b}^f \varphi$.

– $M, q \models_{\text{CSL}} tr(\langle\!\langle A \rangle\!\rangle^f \bigcirc \varphi)$ iff $M, q \models_{\text{CSL}} \langle\!\langle A \rangle\!\rangle \bigcirc tr(\varphi)$ iff $\exists_{S_A} \forall_{\Lambda \in out(q, S_A)} M, \Lambda[1] \models_{\text{CSL}} tr(\varphi)$ iff (by induction) $\exists_{S_A} \forall_{\Lambda \in out(q, S_A)} M, \Lambda[1] \models_{\text{F-ATEL}} \varphi$ iff $M, q \models_{\text{F-ATEL}} \langle\!\langle A \rangle\!\rangle^f \bigcirc \varphi$.

– $M, q \models_{\text{CSL}} tr(\langle\!\langle A \rangle\!\rangle_{M_b}^f \bigcirc \varphi)$ iff $M, q \models_{\text{CSL}} \neg K_b \neg \langle\!\langle A \rangle\!\rangle \bigcirc tr(\varphi)$ iff $\neg\forall_{q' \in \text{img}(q, \sim_b)} \neg M, q' \models_{\text{CSL}} \langle\!\langle A \rangle\!\rangle \bigcirc tr(\varphi)$ iff $\exists_{q' \in \text{img}(q, \sim_b)} \exists_{S_A} \forall_{\Lambda \in out(q', S_A)} M, \Lambda[1] \models_{\text{CSL}} tr(\varphi)$ iff (by induction) $\exists_{S_A} \exists_{q' \in \text{img}(q, \sim_b)} \forall_{\Lambda \in out(q', S_A)} M, \Lambda[1] \models_{\text{F-ATEL}} \varphi$ iff $M, q \models_{\text{F-ATEL}} \langle\!\langle A \rangle\!\rangle_{M_b}^f \bigcirc \varphi$.

– For $\langle\!\langle A \rangle\!\rangle_{M_b}^f \Box \varphi$ and $\langle\!\langle A \rangle\!\rangle_{M_b}^f (\varphi \mathcal{U} \psi)$: analogously.

– $M, q \models_{\text{CSL}} tr(\mathcal{K}_A \varphi)$ iff $M, q \models_{\text{CSL}} \mathcal{K}_A tr(\varphi)$ iff $\forall_{q' \in \text{img}(q, \sim_A^{\mathcal{K}})} M, q' \models_{\text{CSL}} tr(\varphi)$ iff (by induction) $\forall_{q' \in \text{img}(q, \sim_A^{\mathcal{K}})} M, q' \models_{\mathcal{L}} \varphi$ iff $M, q \models_{\mathcal{L}} \mathcal{K}_A \varphi$.

■

Theorem 39

Proof of Theorem 39 (structural induction on the structure of φ). In each case, we will prove that $M, Q \models \mathbb{K}_a \varphi$ implies $M, Q \models \varphi$ for an arbitrary Q. By Proposition 16, we can then conclude that $M, Q \models \mathbb{K}_a \varphi \rightarrow \varphi$.

To simplify the proof, we assume that each φ has been transformed so that no constructive knowledge operator is followed by conjunction (by Proposition 36.3, each subformula $\hat{\mathcal{K}}_A(\psi_1 \wedge \psi_2)$ can be equivalently transformed to $\hat{\mathcal{K}}_A \varphi_1 \wedge \hat{\mathcal{K}}_A \varphi_2$, and we can apply this transformation recursively). Thus, every $\hat{\mathcal{K}}$ in φ is now followed either by some other $\hat{\mathcal{K}}'$, or by $\langle\!\langle A \rangle\!\rangle$, or by a standard knowledge operator $\mathcal{K}''$, or by an atomic proposition p.

Additionally, given Q, we define $Q' = \mathrm{img}(Q, \sim_a)$. Note that $Q \subseteq Q'$ by reflexivity of $\sim_a$. Also, $out(Q, S_A) \subseteq out(Q', S_A)$ by monotonicity of function out wrt Q, and $\mathrm{img}(Q, \sim_A^{\mathcal{K}}) \subseteq \mathrm{img}(Q', \sim_A^{\mathcal{K}})$ by reflexivity of all $\sim_A^{\mathcal{K}}$.

Case $\varphi \equiv p$: Let $M, Q \models \mathbb{K}_a p$. Then $M, Q' \models p$, i.e. $\forall_{q \in Q'} M, q \models p$. So, $\forall_{q \in Q} M, q \models p$, and $M, Q \models p$.

Case $\varphi \equiv \psi_1 \wedge \psi_2$: Let $M, Q \models \mathbb{K}_a(\psi_1 \wedge \psi_2)$. Then $M, Q' \models \psi_1 \wedge \psi_2$, i.e. $M, Q' \models \psi_1$ and $M, Q' \models \psi_2$. So, $M, Q \models \mathbb{K}_a \psi_1$ and $M, Q \models \mathbb{K}_a \psi_2$. By the induction hypothesis, $M, Q \models \psi_1$ and $M, Q \models \psi_2$, and hence $M, Q \models \psi_1 \wedge \psi_2$.

Case $\varphi \equiv \langle\!\langle A \rangle\!\rangle \bigcirc \psi$: Let $M, Q \models \mathbb{K}_a \langle\!\langle A \rangle\!\rangle \bigcirc \psi$. Then $M, Q' \models \langle\!\langle A \rangle\!\rangle \bigcirc \psi$, and so $\exists_{S_A} \forall_{\Lambda \in out(Q', S_A)} M, \Lambda[1] \models \psi$. Thus, $\exists_{S_A} \forall_{\Lambda \in out(Q, S_A)} M, \Lambda[1] \models \psi$, and $M, Q \models \langle\!\langle A \rangle\!\rangle \bigcirc \psi$.

Cases $\varphi \equiv \langle\!\langle A \rangle\!\rangle \Box \psi$ and $\varphi \equiv \langle\!\langle A \rangle\!\rangle \psi_1 \,\mathcal{U}\, \psi_2$: analogous.

Case $\varphi \equiv \mathcal{K}_A \psi$: Let $M, Q \models \mathbb{K}_a \mathcal{K}_A \psi$. Then $M, Q' \models \mathcal{K}_A \psi$, and $\forall_{q \in \mathrm{img}(Q', \sim_A^{\mathcal{K}})} M, q \models \psi$. But then also $\forall_{q \in \mathrm{img}(Q, \sim_A^{\mathcal{K}})} M, q \models \psi$, and $M, Q \models \mathcal{K}_A \psi$.

Before we consider the remaining cases, we define a couple of additional symbols. Let $Q^i = \mathrm{img}(Q^{i-1}, \sim_{A_i}^{\mathcal{K}^i})$, $Q^0 = Q$. That is, $Q^i = \mathrm{img}(\ldots(\mathrm{img}(Q, \sim_{A_1}^{\mathcal{K}^1}), \ldots), \sim_{A_i}^{\mathcal{K}^i})$. Also, let $Q'' = \mathrm{img}(Q^n, \sim_a)$. Note that $Q^n \subseteq Q''$, and $out(Q^n, S_B) \subseteq out(Q'', S_B)$ for any S_B.

Case $\varphi \equiv \hat{\mathcal{K}}_{A_n}^n \ldots \hat{\mathcal{K}}_{A_1}^1 p$: (i.e., φ is a sequence of n possibly different $\hat{\mathcal{K}}$ operators for possibly different coalitions). Let $M, Q \models \mathbb{K}_a \hat{\mathcal{K}}_{A_n}^n \ldots \hat{\mathcal{K}}_{A_1}^1 p$. Then $M, Q'' \models p$, and hence $\forall_{q \in Q''} M, q \models p$. Thus, $\forall_{q \in Q^n} M, q \models p$, so $M, Q^n \models p$, and $M, Q \models \hat{\mathcal{K}}_{A_n}^n \ldots \hat{\mathcal{K}}_{A_1}^1 p$.

Case $\varphi \equiv \hat{\mathcal{K}}_{A_n}^n \ldots \hat{\mathcal{K}}_{A_1}^1 \mathcal{K}_B \psi$: analogous.

Case $\varphi \equiv \hat{\mathcal{K}}_{A_n}^n \ldots \hat{\mathcal{K}}_{A_1}^1 \langle\!\langle B \rangle\!\rangle \bigcirc \psi$: Let $M, Q \models \mathbb{K}_a \hat{\mathcal{K}}_{A_n}^n \ldots \hat{\mathcal{K}}_{A_1}^1 \langle\!\langle B \rangle\!\rangle \bigcirc \psi$. Then $M, Q'' \models \langle\!\langle B \rangle\!\rangle \bigcirc \psi$, and hence $\exists_{S_B} \forall_{\Lambda \in out(Q'', S_B)} M, \Lambda[1] \models \psi$. Thus, $\exists_{S_B} \forall_{\Lambda \in out(Q^n, S_B)} M, \Lambda[1] \models \psi$, so $M, Q^n \models \langle\!\langle B \rangle\!\rangle \bigcirc \psi$, and $M, Q \models \hat{\mathcal{K}}_{A_n}^n \ldots \hat{\mathcal{K}}_{A_1}^1 \langle\!\langle B \rangle\!\rangle \bigcirc \psi$.

Cases $\varphi \equiv \hat{\mathcal{K}}_{A_n}^n \ldots \hat{\mathcal{K}}_{A_1}^1 \langle\!\langle B \rangle\!\rangle \Box \psi$ and $\varphi \equiv \hat{\mathcal{K}}_{A_n}^n \ldots \hat{\mathcal{K}}_{A_1}^1 \langle\!\langle B \rangle\!\rangle \psi_1 \,\mathcal{U}\, \psi_2$: analogous.

■

Proposition 42

Proof of Proposition 42.

K Immediate.

D Suppose that $M, Q \models \mathsf{loc}\ \bot$ for some $Q \neq \emptyset$. Then, there is some q for which $M, q \models \bot$, but this contradicts Proposition 16.1.

T To see that **T** is not strongly valid, let φ and M be as in Lemma 38.1, and take $Q = \mathrm{img}(q, \sim_a)$. $M, Q \models \mathsf{loc}\ \varphi$, but $M, Q \not\models \varphi$.

$\mathbf{4}/\mathbf{4}^+$ $M, Q \models \mathsf{loc}\ \varphi$ iff for every $q \in Q$ we have that $M, q \models \varphi$ iff for every $q \in Q$ we have that $M, q \models \mathsf{loc}\ \varphi$ iff $M, Q \models \mathsf{loc}\ \mathsf{loc}\ \varphi$.

$\mathbf{5}/\mathbf{5}^+$ To see that **5** is not strongly valid, let M be as in Figure 5, $\varphi = p$ and let $Q = \{q, q'\}$. $M, Q \models \neg\mathsf{loc}\ \varphi$ because $M, q' \models \neg\varphi$. However, if it were the case that $M, Q \models \mathsf{loc}\ \neg\mathsf{loc}\ \varphi$, then $M, q \models \neg\mathsf{loc}\ \varphi$ and thus, $M, q \models \neg\varphi$, which is not the case.

B $M, Q \models \varphi$ iff for all $q \in Q$ we have that $M, q \models \varphi$ iff for all $q \in Q$ we have that $M, q \models \neg\mathsf{loc}\ \neg\varphi$ iff $M, Q \models \mathsf{loc}\ \neg\mathsf{loc}\ \neg\varphi$.

■

Theorem 56

Proof of Theorem 56.

$\widetilde{\mathbf{K}}$**:** We construct a counterexample. Let M be a model with states q_1, q_2 and agent a, such that $q_1 \sim_a q_2$, $\pi(q_1) = \{r\}$ and $\pi(q_2) = \{p\}$. Let $\varphi = \neg p$ and $\psi = r$. $p \notin \pi(q_1) \cap \pi(q_2)$, so $M, \mathrm{img}(q_1, \sim_a) \models \varphi$ and $M, q_1 \models \mathbb{K}_a\varphi$. $r \notin \pi(q_1) \cap \pi(q_2)$, so $M, \mathrm{img}(q_1, \sim_a) \not\models \psi$ and $M, q_1 \not\models \mathbb{K}_a\psi$. Thus, $M, q_1 \not\models \mathbb{K}_a\varphi \rightarrow \mathbb{K}_a\psi$ and by Proposition 53: $M, q_1 \not\models \mathbb{K}_a\varphi \rightsquigarrow \mathbb{K}_a\psi$ (*). Since both $M, q_1 \models \varphi \rightarrow \psi$ and $M, q_2 \models \varphi \rightarrow \psi$, by Proposition 50, $M, \mathrm{img}(q_1, \sim_a) \models \varphi \rightsquigarrow \psi$ and thus $M, q_1 \models \mathbb{K}_a(\varphi \rightsquigarrow \psi)$. Together with (*), we get that $M, q_1 \not\models \mathbb{K}_a(\varphi \rightsquigarrow \psi) \rightarrow (\mathbb{K}_a\varphi \rightsquigarrow \mathbb{K}_a\psi)$ and, by Proposition 53, $M, q_1 \not\models \mathbb{K}_a(\varphi \rightsquigarrow \psi) \rightsquigarrow (\mathbb{K}_a\varphi \rightsquigarrow \mathbb{K}_a\psi)$. Thus, $\widetilde{\mathbf{K}}$ is not weakly (and hence not strongly) valid.

$\widetilde{\mathbf{T}}$**:** Let M, q, a, φ be as in Lemma 38.2. $M, q \models \mathbb{K}_a\varphi$ and $M, q \not\models \varphi$, so $M, q \not\models \mathbb{K}_a\varphi \rightsquigarrow \varphi$ by Proposition 53. Thus, $\widetilde{\mathbf{T}}$ is not weakly (and hence not strongly) valid.

$\widetilde{\mathbf{4}}^+/\widetilde{\mathbf{4}}$**:** $M, Q \models \mathbb{K}_a\varphi \leftrightsquigarrow \mathbb{K}_a\mathbb{K}_a\varphi$ iff, by Proposition 50, $\forall_{q \in Q}(M, q \models \mathbb{K}_a\varphi \Leftrightarrow M, q \models \mathbb{K}_a\mathbb{K}_a\varphi)$ iff, by $\mathbf{4}^+$, $\forall_{q \in Q}(M, q \models \mathbb{K}_a\varphi \Leftrightarrow M, q \models \mathbb{K}_a\varphi)$.

$\widetilde{\mathbf{5}}^+/\widetilde{\mathbf{5}}$**:** $M, Q \models \sim \mathbb{K}_a\varphi \leftrightsquigarrow \mathbb{K}_a \sim \mathbb{K}_a\varphi$ iff, by Proposition 50, $\forall_{q \in Q}(M, q \models \sim \mathbb{K}_a\varphi \Leftrightarrow M, q \models \mathbb{K}_a \sim \mathbb{K}_a\varphi)$ iff $\forall_{q \in Q}(M, \mathrm{img}(q, \sim_a) \not\models \varphi \Leftrightarrow M, \mathrm{img}(q, \sim_a) \models \sim \mathbb{K}_a\varphi)$ iff $\forall_{q \in Q}(M, \mathrm{img}(q, \sim_a) \not\models \varphi \Leftrightarrow \forall_{q' \in \mathrm{img}(q, \sim_a)} M, \mathrm{img}(q', \sim_a) \not\models \varphi)$ which is true, since $\mathrm{img}(q', \sim_a) = \mathrm{img}(q, \sim_a)$ for any $q' \in \mathrm{img}(q, \sim_a)$.

$\widetilde{\mathbf{D}}$**:** $M, Q \models {\sim}\mathbb{K}_a \wr$ iff $\forall_{q \in Q} M, q \not\models \mathbb{K}_a \wr$ iff $\forall_{q \in Q} M, \mathrm{img}(q, \sim_a) \not\models \wr$, which is true by Proposition 51.1.

$\widetilde{\mathbf{B}}$**:** $M, Q \models \varphi \rightsquigarrow \mathbb{K}_a \sim \mathbb{K}_a \sim \varphi$ iff $\forall_{q \in Q}(M, q \models \varphi \Rightarrow M, \mathrm{img}(q, \sim_a) \models \sim \mathbb{K}_a \sim \varphi)$ iff $\forall_{q \in Q}(M, q \models \varphi \Rightarrow \forall_{q' \in \mathrm{img}(q, \sim_a)} M, q' \not\models \mathbb{K}_a \sim \varphi)$ iff $\forall_{q \in Q}(M, q \models \varphi \Rightarrow \forall_{q' \in \mathrm{img}(q, \sim_a)} M, \mathrm{img}(q', \sim_a) \not\models \sim \varphi)$ iff $\forall_{q \in Q}(M, q \models \varphi \Rightarrow \forall_{q' \in \mathrm{img}(q, \sim_a)} \exists_{q'' \in \mathrm{img}(q', \sim_a)} M, q'' \models \varphi)$ iff $\forall_{q \in Q}(M, q \models \varphi \Rightarrow \exists_{q' \in \mathrm{img}(q, \sim_a)} M, q' \models \varphi)$. This always holds, by taking $q' = q$.

■

Theorem 60 and Corollary 61 (Constructive Normal Form) and Theorem 62 (Expressiveness of Strong Negation)

In the following we will very often work in the language $\mathcal{L}(\neg, \wedge, \langle\langle A \rangle\rangle T, \hat{\mathcal{K}}_A, \sim)$, and we will henceforth use the shorthand notation $\hat{\mathcal{L}}$ to denote this language, for simplicity.

We use $\mathit{Subf}(\varphi)$ to denote the set of all subformulae of φ (including φ itself). For simplicity, we assume that each subformula of a formula is *unique*, i.e. that there is a unique member of $\mathit{Subf}(\varphi)$ for each occurrence of a subformula in φ[13].

We first present intermediate definitions and results leading up to the main result in Theorem 60. Note that Lemma 65 below gives an alternative (equivalent) definition of constructive normal form.

DEFINITION 63. — *We define the* depth $d_\varphi(\psi)$ *of a subformula* $\psi \in \mathit{Subf}(\varphi)$ *of a formula* $\varphi \in \hat{\mathcal{L}}$ *in the usual way:*

- $d_\varphi(\varphi) = 0$
- $d_\varphi(\hat{\mathcal{K}}\gamma) = d \Rightarrow d_\varphi(\gamma) = d + 1$
- $d_\varphi(\langle\langle G \rangle\rangle T\gamma) = d \Rightarrow d_\varphi(\gamma) = d + 1$
- $d_\varphi(\gamma_1 \wedge \gamma_2) = d \Rightarrow d_\varphi(\gamma_1) = d_\varphi(\gamma_2) = d + 1$
- $d_\varphi(\neg\gamma) = d \Rightarrow d_\varphi(\gamma) = d + 1$
- $d_\varphi(\sim\gamma) = d \Rightarrow d_\varphi(\gamma) = d + 1$

LEMMA 64. — *A formula* $\psi \in \hat{\mathcal{L}}$ *is of CSNF iff every* $\gamma \in \mathit{Subf}(\psi)$ *is of CSNF.*

Proof. The implication to the left is trivial; we prove the one to the right. Assume that ψ is of CSNF. That each $\gamma \in \mathit{Subf}(\psi)$ is of CSNF follows immediately by induction on the depth of γ:

13. This can be achieved by, e.g., adorning the subformulae with unique identifiers, or by taking $\mathit{Subf}(\varphi)$ to be a multiset instead of a set. The only reason for this assumption is to make proofs simpler.

– $d_\psi(\gamma) = 0$: $\gamma = \psi$ is of CSNF

– $d_\psi(\gamma) = d + 1$ ($d \geq 0$). We reason by the possible cases:

- $d_\psi(\hat{\mathcal{K}}\gamma) = d$: by the induction hypothesis $\hat{\mathcal{K}}\gamma$ is of CSNF, and thus γ is of CSNF.

- $d_\psi(\langle\langle G\rangle\rangle T\gamma) = d$: $\langle\langle G\rangle\rangle T\gamma$ is of CSNF; γ is of CSNF.

- $d_\psi(\gamma \wedge \gamma') = d$: $\gamma \wedge \gamma'$ is of CSNF; γ is of CSNF. Similar when $d_\psi(\gamma' \wedge \gamma) = d$.

- $d_\psi(\neg\gamma) = d$: $\neg\gamma$ is of CSNF; γ is of CSNF.

- $d_\psi(\sim\gamma) = d$: $\sim\gamma$ is of CSNF; γ is of CSNF.

■

LEMMA 65. — *A formula $\psi \in \hat{\mathcal{L}}$ is of CSNF iff every $\hat{\mathcal{K}}\gamma \in Subf(\psi)$ is of the form $\hat{\mathcal{K}}\hat{\mathcal{K}}_0 \cdots \hat{\mathcal{K}}_k\alpha$ where $\alpha \in Atoms$, for some $k \geq 0$.*

Proof. For the direction to the right, assume that there is a $\hat{\mathcal{K}}\gamma \in Subf(\psi)$ which is *not* of the form. There are two possibilities: $\gamma = \hat{\mathcal{K}}_0 \cdots \hat{\mathcal{K}}_m\neg\beta$ or $\gamma = \hat{\mathcal{K}}_0 \cdots \hat{\mathcal{K}}_m\beta_1 \wedge \beta_2$ for some $m \geq 0$. In either case, it follows immediately that $\hat{\mathcal{K}}\gamma$ is not of CSNF. By Lemma 64, ψ is not of CSNF.

For the direction to the left, assume that every $\hat{\mathcal{K}}\gamma \in Subf(\psi)$ is of the form. We show that every $\chi \in Subf(\psi)$ is of CSNF by structural induction:

– $\chi = p \in \Theta$: χ is of CSNF.

– $\chi = \hat{\mathcal{K}}\gamma$: by the induction hypothesis, γ is of CSNF. By assumption, $\gamma = \hat{\mathcal{K}}_0 \cdots \hat{\mathcal{K}}_k\alpha$ for some $\alpha \in Atoms$ and some $k \geq 0$. Thus, χ is of CSNF.

– $\chi = \langle\langle A\rangle\rangle T\gamma$: by the induction hypothesis, γ is of CSNF, and thus χ is of CSNF.

– $\chi = \gamma_1 \wedge \gamma_2$: by the induction hypothesis, γ_1 and γ_2 is of CSNF, and thus χ is of CSNF.

– $\chi = \sim\gamma$: by the induction hypothesis, γ is of CSNF and thus χ is of CSNF.

– $\chi = \neg\gamma$: by the induction hypothesis, γ is of CSNF and thus χ is of CSNF.

■

Now that we have established some properties of formulae of CSNF, we go on to define the mapping of a formula to one of CSNF.

DEFINITION 66. — *The value $f(\hat{\mathcal{K}}\psi)$ of the function $f : \{\hat{\mathcal{K}}\psi : \psi \in \hat{\mathcal{L}}\} \to \hat{\mathcal{L}}$ is defined by structural induction over ψ:*

$$\begin{aligned} f(\hat{\mathcal{K}}\psi) &= \hat{\mathcal{K}}\psi \quad \textit{when } \psi \in Atoms \\ f(\hat{\mathcal{K}}\hat{\mathcal{K}}'\gamma) &= \hat{\mathcal{K}}\hat{\mathcal{K}}'\gamma \\ f(\hat{\mathcal{K}}\neg\gamma) &= \neg f(\hat{\mathcal{K}}\gamma) \\ f(\hat{\mathcal{K}}(\gamma_1 \wedge \gamma_2)) &= f(\hat{\mathcal{K}}\gamma_1) \wedge f(\hat{\mathcal{K}}\gamma_2) \end{aligned}$$

LEMMA 67. — *Let $\beta \in \hat{\mathcal{L}}$ be a formula. β is of CSNF iff $f(\hat{\mathcal{K}}\beta)$ is of CSNF for any arbitrary $\hat{\mathcal{K}} \in \{\mathbb{C}_A, \mathbb{D}_A, \mathbb{E}_A : A \subseteq \Sigma\}$.*

Proof. Let $\hat{\mathcal{K}} \in \{\mathbb{C}_A, \mathbb{D}_A, \mathbb{E}_A : A \subseteq \Sigma\}$. The proof is by structural induction over β:

– $\beta = p \in \Theta$: β is of CSNF iff $f(\hat{\mathcal{K}}\beta) = \hat{\mathcal{K}}p$ is of CSNF.

– $\beta = \hat{\mathcal{K}}'\gamma$: β is of CSNF iff $f(\hat{\mathcal{K}}\beta) = \hat{\mathcal{K}}\hat{\mathcal{K}}'\gamma$ is of CSNF.

– $\beta = \langle\langle A\rangle\rangle T\gamma$: β is of CSNF iff $f(\hat{\mathcal{K}}\beta) = \hat{\mathcal{K}}\langle\langle A\rangle\rangle T\gamma$ is of CSNF.

– $\beta = \gamma_1 \wedge \gamma_2$: β is of CSNF iff both γ_1 and γ_2 are of CSNF iff, by the induction hypothesis, both $f(\hat{\mathcal{K}}\gamma_1)$ and $f(\hat{\mathcal{K}}\gamma_2)$ are of CSNF iff $f(\hat{\mathcal{K}}\beta)$ is of CSNF.

– $\beta = \sim\gamma$: β is of CSNF iff $f(\hat{\mathcal{K}}\beta) = \hat{\mathcal{K}}\beta$ is of CSNF.

– $\beta = \neg\gamma$: β is of CSNF iff γ is of CSNF iff, by the induction hypothesis, $f(\hat{\mathcal{K}}\gamma)$ is of CSNF iff $\neg f(\hat{\mathcal{K}}\gamma)$ is of CSNF iff $f(\hat{\mathcal{K}}\neg\gamma)$ is of CSNF.

■

LEMMA 68. — *For any $\psi \in \hat{\mathcal{L}}$,*

$$\hat{\mathcal{K}}\psi \leftrightarrow f(\hat{\mathcal{K}}\psi)$$

is strongly valid for any $\hat{\mathcal{K}} \in \{\mathbb{C}_A, \mathbb{D}_A, \mathbb{E}_A : A \subseteq \Sigma\}$.

Proof. The proof is by structural induction over ψ. When $\psi \in Atoms$ or $\psi = \hat{\mathcal{K}}'\gamma$, $f(\hat{\mathcal{K}}\psi) = \hat{\mathcal{K}}\psi$, and we are done. When $\psi = \neg\gamma$, $M, Q \models \hat{\mathcal{K}}\psi$ iff, by Proposition 36, $M, Q \models \neg\hat{\mathcal{K}}\gamma$ iff $M, Q \not\models \hat{\mathcal{K}}\gamma$ iff, by the induction hypothesis, $M, Q \not\models f(\hat{\mathcal{K}}\gamma)$ iff $M, Q \models \neg f(\hat{\mathcal{K}}\gamma)$ iff $M, Q \models f(\hat{\mathcal{K}}\psi)$. When $\psi = \gamma_1 \wedge \gamma_2$, $M, Q \models \hat{\mathcal{K}}\psi$ iff, by Proposition 36, $M, Q \models \hat{\mathcal{K}}\gamma_1$ and $M, Q \models \hat{\mathcal{K}}\gamma_2$ iff, by the induction hypothesis, $M, Q \models f(\hat{\mathcal{K}}\gamma_1)$ and $M, Q \models f(\hat{\mathcal{K}}\gamma_2)$ iff $M, Q \models f(\hat{\mathcal{K}}\psi)$. ■

DEFINITION 69 (φ_i, X_i, α_i). — *Let $\varphi \in \hat{\mathcal{L}}$ be a formula. Define φ_i, $i \geq 0$:*

– $i = 0$: $\varphi_0 = \varphi$

– $i = j+1$ ($j \geq 0$): Let $X_i = \{\hat{\mathcal{K}}\psi : \hat{\mathcal{K}}\psi \in Subf(\varphi_j), \hat{\mathcal{K}}\psi$ is not of CSNF$\}$. If X_i is empty, let $\varphi_{j+1} = \varphi_j$. Otherwise, select an $\alpha_i \in X_i$ such that $\beta \in X_i$ implies that $d_{\varphi_j}(\beta) \leq d_{\varphi_j}(\alpha_i)$ (several such α_i may exist; select one arbitrarily), and let φ_{j+1} be φ_j with the subformula α_i replaced by $f(\alpha_i)$.

LEMMA 70. — *Let $\varphi \in \hat{\mathcal{L}}$, and let α_i be defined in Def. 69. For each $i \geq 1$, $f(\alpha_i)$ is of CSNF.*

Proof. Let $\alpha_i = \hat{\mathcal{K}}\psi$. We show that for every $\gamma \in Subf(\psi)$, $f(\hat{\mathcal{K}}\gamma)$ is of CSNF by structural induction over γ:

– $\gamma = p \in \Theta$: $f(\hat{\mathcal{K}}\gamma) = \hat{\mathcal{K}}p$ is of CSNF.

– $\gamma = \hat{\mathcal{K}}'\beta$: $f(\hat{\mathcal{K}}\gamma) = \hat{\mathcal{K}}\hat{\mathcal{K}}'\beta$ is of CSNF iff $\hat{\mathcal{K}}'\beta$ is of CSNF. Assume that $\hat{\mathcal{K}}'\beta$ is *not* of CSNF, then $\gamma \in X_i$. Then $d_{\varphi_{i-1}}(\gamma) > d_{\varphi_{i-1}}(\alpha_i)$, but this is a contradiction since there are no $\gamma \in X_i$ with greater depth than α_i. Thus, $f(\hat{\mathcal{K}}\gamma)$ is of CSNF.

– $\gamma = \langle\langle G\rangle\rangle T\beta$: By the induction hypothesis, $f(\hat{\mathcal{K}}\beta)$ is of CSNF; by Lemma 67 β is of CSNF; $\langle\langle G\rangle\rangle T\beta$ is of CSNF; $\hat{\mathcal{K}}\langle\langle G\rangle\rangle T\beta = f(\hat{\mathcal{K}}\gamma)$ is of CSNF.

– $\gamma = \gamma_1 \wedge \gamma_2$: By the induction hypothesis, $f(\hat{\mathcal{K}}\gamma_1)$ and $f(\hat{\mathcal{K}}\gamma_2)$ are of CSNF; $f(\hat{\mathcal{K}}\gamma_1) \wedge f(\hat{\mathcal{K}}\gamma_2)$ is of CSNF; $f(\hat{\mathcal{K}}\gamma)$ is of CSNF.

– $\gamma = \neg\beta$: By the induction hypothesis, $f(\hat{\mathcal{K}}\beta)$ is of CSNF; $\neg f(\hat{\mathcal{K}}\beta)$ is of CSNF; $f(\hat{\mathcal{K}}\gamma)$ is of CSNF.

– $\gamma = \sim\beta$: By the induction hypothesis, $f(\hat{\mathcal{K}}\beta)$ is of CSNF; by Lemma 67 β is of CSNF; γ is of CSNF; $\hat{\mathcal{K}}\gamma = f(\hat{\mathcal{K}}\gamma)$ is of CSNF.

■

LEMMA 71. — *Let $\varphi \in \hat{\mathcal{L}}$, and let φ_i be defined in Def. 69. There is a $p \geq 1$ such that $\varphi_p = \varphi_{p-1}$ and $X_p = \emptyset$. We write $\hat{\varphi} = \varphi_p$ for an arbitrary such p.*

Proof. X_1 is finite. We show that $X_{i+1} \subset X_i$ (proper inclusion) whenever $\varphi_i \neq \varphi_{i-1}$, for any $i \geq 1$. The Lemma follows.

Let $\varphi_i \neq \varphi_{i-1}$. Assume that there is an $\alpha \in X_{i+1}$, $\alpha \notin X_i$. α is not of CSNF, and since $\alpha \in \mathit{Subf}(\varphi_i)$ and $\alpha \notin \mathit{Subf}(\varphi_{i-1})$ the only possibility is that $\alpha \in \mathit{Subf}(f(\alpha_i))$. But by Lemma 70, $f(\alpha_i)$ is of CSNF, and by Lemma 64 α must be of CSNF which is a contradiction. Thus, $X_{i+1} \subseteq X_i$. To see that the inclusion is proper, observe that $\alpha_i \in X_i$ but $\alpha_i \notin X_{i+1}$. ■

Proof of Theorem 60. Let $\varphi'' \in \mathcal{L}^*$, and let φ' be the result of replacing every occurrence of $\mathcal{K}$ in φ'' with the combination $\hat{\mathcal{K}} \sim\sim$, for every $\mathcal{K}$. Let φ be the result of replacing every occurrence of loc in φ' with the combination $\sim\sim$. φ'' and φ are strongly equivalent by Remark 49. Observe that $\varphi \in \hat{\mathcal{L}}$. Let $\hat{\varphi} = \varphi_p$ be defined from φ as in Lemma 71.

First, we argue that $\hat{\varphi}$ is of CSNF. If not, there is a $\hat{\mathcal{K}}\gamma \in \mathit{Subf}(\hat{\varphi})$ where γ is not of the form $\hat{\mathcal{K}}_0 \cdots \hat{\mathcal{K}}_k\alpha$ for $\alpha \in \mathit{Atoms}$ (Lemma 65). Then, $\hat{\mathcal{K}}\gamma$ is not of CSNF, which contradicts the fact that $X_p = \emptyset$. Second, we show that $\hat{\varphi} \leftrightarrow \varphi$ is strongly valid. Let $i \geq 1$. By Lemma 68, $M, Q \models \alpha_i$ iff $M, Q \models f(\alpha_i)$ for any M, Q. It follows immediately that $M, Q \models \varphi_i$ iff $M, Q \models \varphi_{i+1}$. Thus, $M, Q \models \varphi = \varphi_0$ iff $M, Q \models \hat{\varphi} = \varphi_p$. Thus, $\hat{\varphi}$ is of CSNF, and it is equivalent to φ. ■

Proof of Corollary 61. Let $\varphi \in \mathcal{L}^*$. By the theorem, φ is strongly equivalent to a formula $\hat{\varphi}$ which is of CSNF. Now, we recursively replace all subformulae of $\hat{\varphi}$ of the form $\sim\psi$ with $\langle\langle\emptyset\rangle\rangle(\neg\psi)\,\mathcal{U}\,(\neg\psi)$, yielding (by Proposition 44) a strongly equivalent formula φ' without strong negation. We observe that subformulae of CSNF are replaced with subformulae of CSNF, so φ' is of CSNF too. ■

We now go on to present our proof of Theorem 62. Some more notation: when $\varphi \in \hat{\mathcal{L}}$, we use $\widetilde{\varphi}$ to denote the result of replacing each occurrence of $\neg$ in φ with $\sim$. Formally, $\widetilde{p} = p$; $\widetilde{\mathbb{K}_a\psi} = \mathbb{K}_a\widetilde{\psi}$; $\widetilde{\langle\langle G\rangle\rangle T\psi} = \langle\langle G\rangle\rangle T\widetilde{\psi}$; $\widetilde{\psi_1 \wedge \psi_2} = \widetilde{\psi_1} \wedge \widetilde{\psi_2}$; $\widetilde{\neg\psi} = \sim\widetilde{\psi}$; $\widetilde{\sim\psi} = \sim\widetilde{\psi}$.

We begin with defining the notion of *constructive depth* of a subformula – not to be confused with the notion of *depth* in the proof of Theorem 60.

DEFINITION 72 (CONSTRUCTIVE DEPTH). — *Let $\varphi \in \hat{\mathcal{L}}$. The* constructive depth*, or just c-depth, $D_\varphi(\psi)$ in φ of a subformula $\psi \in Subf(\varphi)$ is defined inductively as follows:*

– $D_\varphi(\varphi) = 0$
– $D_\varphi(\hat{\mathcal{K}}\gamma) = D \Rightarrow D_\varphi(\gamma) = D + 1$
– $D_\varphi(\langle\langle G\rangle\rangle T\gamma) = D \Rightarrow D_\varphi(\gamma) = 0$
– $D_\varphi(\gamma_1 \wedge \gamma_2) = D \Rightarrow D_\varphi(\gamma_1) = D_\varphi(\gamma_2) = D$
– $D_\varphi(\neg\gamma) = D \Rightarrow D_\varphi(\gamma) = D$
– $D_\varphi(\sim\gamma) = D \Rightarrow D_\varphi(\gamma) = 0$

If φ has no occurrence of $\neg$ on c-depth D, i.e., if $\neg\psi \in Subf(\varphi)$ implies that $D_\varphi(\neg\psi) \neq D$, we say that φ is *free of $\neg$ on depth D*.

LEMMA 73. — *If a formula $\varphi \in \mathcal{L}(\neg, \wedge, \langle\langle A\rangle\rangle T, \hat{\mathcal{K}}, \sim)$ is free of $\neg$ on all depths > 0, then*

$$\varphi \leftrightarrow \widetilde{\varphi}$$

is valid.

Proof. We show that

$$\psi \leftrightarrow \widetilde{\psi} \text{ is } \begin{cases} \text{valid} & \text{if } D_\varphi(\psi) = 0 \\ \text{strongly valid} & \text{if } D_\varphi(\psi) > 0 \end{cases}$$

for all $\psi \in Subf(\varphi)$ by structural induction.

$\psi = p$**:** immediate ($\widetilde{\psi} = \psi$).

$\psi = \hat{\mathcal{K}}_A\gamma$**:** $D_\varphi(\gamma) > 0$. $M, Q \models \psi$ iff $M, \text{img}(Q, \sim_A^{\mathcal{K}}) \models \gamma$ iff, by the induction hypothesis, $M, \text{img}(Q, \sim_A^{\mathcal{K}}) \models \widetilde{\gamma}$ iff $M, Q \models \mathcal{K}_A\widetilde{\gamma}$ iff $M, Q \models \widetilde{\psi}$.

$\psi = \langle\langle G\rangle\rangle\Box\gamma$**:** $M, Q \models \psi$ iff $\exists_{S_G}\forall_{\Lambda\in out(Q,S_G)}\forall_{j\geq 0}M, \Lambda[j] \models \gamma$ iff, by the induction hypothesis (for γ, where $D_\varphi(\gamma) = 0$), $\exists_{S_G}\forall_{\Lambda\in out(Q,S_G)}\forall_{j\geq 0}M, \Lambda[j] \models \widetilde{\gamma}$ iff $M, Q \models \langle\langle G\rangle\rangle\Box\widetilde{\gamma}$. Similar for the other ATL connectives.

$\psi = \gamma_1 \wedge \gamma_2$**:** First, consider the case that $D_\varphi(\psi) = 0$, in which case $D_\varphi(\gamma_1) = D_\varphi(\gamma_2) = 0$. We must show that $\psi \leftrightarrow \widetilde{\psi}$ is valid. $M, q \models \psi$ iff $M, q \models \gamma_1$ and $M, q \models \gamma_2$ iff, by the induction hypothesis, $M, q \models \widetilde{\gamma_1}$ and $M, q \models \widetilde{\gamma_2}$ iff $M, q \models \widetilde{\gamma_1} \wedge \widetilde{\gamma_2}$. Second, consider the case that $D_\varphi(\psi) > 0$, in which case $D_\varphi(\gamma_1) > 0$ and $D_\varphi(\gamma_2) > 0$. We must show that $\psi \leftrightarrow \widetilde{\psi}$ is strongly valid. $M, Q \models \psi$ iff $M, Q \models \gamma_1$ and $M, Q \models \gamma_2$ iff, by the induction hypothesis, $M, Q \models \widetilde{\gamma_1}$ and $M, Q \models \widetilde{\gamma_2}$ iff $M, Q \models \widetilde{\gamma_1} \wedge \widetilde{\gamma_2}$.

$\psi = \neg\gamma$: By the assumption in the lemma, $D_\varphi(\psi) = 0$. Then also $D_\varphi(\gamma) = 0$. $M, q \models \psi$ iff $M, q \not\models \gamma$ iff, by the induction hypothesis, $M, q \not\models \widetilde{\gamma}$ iff $M, q \models \sim\widetilde{\gamma}$.

$\psi = \sim\gamma$: $M, Q \models \psi$ iff $\forall_{q \in Q} M, q \not\models \gamma$ iff, by the induction hypothesis (for γ, where $D_\varphi(\gamma) = 0$), $\forall_{q \in Q} M, q \not\models \widetilde{\gamma}$ iff $M, Q \models \sim\widetilde{\gamma}$.

■

Proof of Theorem 62. Let $\varphi \in \mathcal{L}(\neg, \wedge, \langle\langle A \rangle\rangle T, \hat{\mathcal{K}}, \sim)$ be a formula, and let $\hat{\varphi}$ be a formula of CSNF equivalent to φ. Note that $\hat{\varphi} \in \mathcal{L}(\neg, \wedge, \langle\langle A \rangle\rangle T, \hat{\mathcal{K}}, \sim)$. We show that

$$D_{\hat{\varphi}}(\psi) > 0 \Rightarrow \psi \in Atoms \text{ or } \psi = \hat{\mathcal{K}}\gamma \tag{3}$$

for every $\psi \in Subf(\hat{\varphi})$ by induction over the depth (not the constructive depth) of ψ, for arbitrary γ and $\hat{\mathcal{K}}$. For the base case, let $\psi = \hat{\varphi}$ and (3) is vacuously true. In the inductive case assume that (3) holds for the parent of ψ. There are three circumstances in which $D_{\hat{\varphi}}(\psi) > 0$. First, $\hat{\mathcal{K}}\psi \in Subf(\hat{\varphi})$. Then, $\psi \in Atoms$ or ψ is of the form $\mathbb{K}_b\gamma$, since $\hat{\varphi}$ is of CSNF. Second, $\neg\psi \in Subf(\hat{\varphi})$ with $D_{\hat{\varphi}}(\psi) = D_{\hat{\varphi}}(\neg\psi)$. By the induction hypothesis, it must be the case that $D_{\hat{\varphi}}(\neg\psi) = 0$, so (3) is vacuously true. Third, $\psi \wedge \psi' \in Subf(\hat{\varphi})$ with $D_{\hat{\varphi}}(\psi) = D_{\hat{\varphi}}(\psi') = D_{\hat{\varphi}}(\psi \wedge \psi')$. By the induction hypothesis, it must be the case that $D_{\hat{\varphi}}(\psi \wedge \psi') = 0$, so (3) is vacuously true. Similarly for the case $\psi' \wedge \psi$. This shows that $\hat{\varphi}$ is free for $\neg$ on all depths > 0, and thus φ is equivalent to $\hat{\varphi}$ which is equivalent to $\widetilde{\hat{\varphi}}$ by Lemma 73 which is without weak negation. ■