Decidable Verification of Multi-agent Systems with Bounded Private Actions

Extended Abstract

Francesco Belardinelli
Laboratoire IBISC, UEVE
Evry, France
belardinelli@ibisc.fr

Alessio Lomuscio
Imperial College London
London, UK
a.lomuscio@imperial.ac.uk

Aniello Murano
Università degli Studi di Napoli
Naples, Italy
murano@na.infn.it

Sasha Rubin
Università degli Studi di Napoli
Naples, Italy
rubin@unina.it

ACM Reference Format:
Francesco Belardinelli, Alessio Lomuscio, Aniello Murano, and Sasha Rubin. 2018. Decidable Verification of Multi-agent Systems with Bounded Private Actions. In *Proc. of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2018), Stockholm, Sweden, July 10–15, 2018,* IFAAMAS, 3 pages.

Introduction. A key feature of the frontier between decidability and undecidability in the verification of multi-agent systems (MAS) with respect to strategy-based specifications is the assumptions made on the agents' information. For instance, the complexity of model checking MAS against ATL is markedly different for agents with complete information compared with agents with incomplete information, i.e., PTIME-complete [1] and undecidable [7], respectively, assuming agents with perfect recall. It follows that the verification problem for agents with incomplete information and perfect recall remains undecidable in any formalism stronger than ATL, such as Strategy Logic and most of its variants [4–6, 12]. It therefore remains of importance to identify expressive fragments whose model checking problem is decidable.

A way of achieving this consists in identifying classes of MAS, still endowed with perfect recall and incomplete information, for which model checking is decidable. For example, such a result is proved in [2, 3] for MAS where all communication is via public actions. In this paper we further pursue this line and show that the work in [2] can be generalised much further. Specifically, we show that: i) model checking remains decidable even if non-public actions are permitted a bounded number of times along any execution; ii) the temporal language underlying epistemic SL can be considerably extended from LTL to LDL [13] at no extra computational cost. By doing so, we show decidability for a much larger class of MAS against a considerably expressive specification language.

Interpreted Systems (IS) with Explicit Public Actions. Interpreted systems [8] are a formal setting for multi-agent systems (MAS), where each agent a is defined by its set of local states L_a, set of actions act_a, local protocol function $P_a : L_a \to \mathbb{P}(act_a)\backslash\{\emptyset\}$ (specifying available actions), and a local transition function τ_a : $L_a \times \mathcal{J}act \to L_a$ where $\mathcal{J}act = \prod_a act_a$ is the set of joint actions. This induces a transition system with state set $S = \prod_a L_a$, initial states in $S_0 \subseteq S$, and transition function $\tau : S \times \mathcal{J}act \to S$. An *interpreted system* S is such a transition system with a valuation function $\pi : AP \to \mathbb{P}(S)$, where AP is a set of atomic predicates.

A *run* (resp. *history*) is an infinite (resp. finite) sequence r of global states such that $r(0) \in S_0$ and for every $n < |r|$ there exists a joint action $J \in \mathcal{J}act$ such that i) all agent actions are allowed by the respective individual protocols, and ii) $\tau(r(n), J) = r(n+1)$. For $a \in Ag$ and $n < |r|$, let $r_a(n)$ be the local state of agent a in the nth global state of r. The set of all histories is denoted by $Hist$.

We now define a variant of IS in which some actions are public.

Definition 1. *An* interpreted system with explicit public actions *is an IS such that, for every $a \in Ag$, there exists a set $pb\_act_a \subseteq act_a$ of* public actions*, and a set L_pr_a of* private components *such that:*
i) $L_a = L\_pr_a \times \mathcal{J}act_\epsilon$ where $\mathcal{J}act_\epsilon = \prod_{a \in Ag}(pb\_act_a \cup \{\epsilon\})$.
ii) The local transition function τ_a satisfies that $\tau_a((p, v), J) = (p', J')$ implies that for all $a \in Ag$, if $J_a \in pb\_act_a$ then $J'_a = J_a$, else $J'_a = \epsilon$.
iii) The initial global states are of the form $(l_a, (\epsilon, \cdots \epsilon))_{a \in Ag}$.

That is, a local state is a pair (p, v) where $p \in L\_pr_a$ is the private local state proper, and $v \in \mathcal{J}act_\epsilon$ represents the last joint action performed in the system, with public actions only being visible and non-public actions represented by ϵ.

Any system conforming to Def. 1 is an IS. Also, any IS is isomorphic to some system conforming to Def. 1 (set $pb\_act_a = \emptyset$ for all a). Thus, for convenience we will call systems conforming to Def. 1 simply IS. Also if $act_a = pb\_act_a$ for all $a \in Ag$, we obtain systems with public actions only. These are closely related to the recording contexts in [8], game structures with public actions only [3], and MAS with broadcasting environment [10].

Definition 2 (*BMNPA*). *A joint action $J \in \mathcal{J}act$ is called* public *if $J(a) \in pb\_act_a$ for all $a \in Ag$. Let Δ denote the set of all public joint actions.*

An interpreted system has only public actions after time *b if for every history $h \in Hist$, joint action $J \in \mathcal{J}act$, and $n \geq b$, if $\tau(h(n), J) = h(n+1)$ then J is a public joint action. An interpreted system that has only public actions after time b, for some bound b, is said to have* boundedly-many non-public actions*; the class of such systems is denoted by BMNPA.*

Proc. of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2018), M. Dastani, G. Sukthankar, E. André, S. Koenig (eds.), July 10–15, 2018, Stockholm, Sweden.

If an action is not in pb_act_a it may still be observed by all agents, e.g., by being recorded in their private components. Thus, pb_act_a should not be considered as the set of all public actions of agent a, but only those explicitly identified as such.

Epistemic Dynamic Strategy Logic We now define EDSL, an extension of Strategy Logic [12] in two directions: its temporal component is based on LDL [13] rather than LTL, and it adds an epistemic dimension.

Syntax. Fix a finite set AP of *atomic propositions (atoms)*, a finite set Ag of *agents* , and a finite set Var of strategy variables $x_0, x_1, \ldots$

Definition 3 (EDSL). *The* EDSL formulas φ *and* EDSL expressions ρ *over AP, Ag, and Var are built according to the following grammar:*

$$\varphi ::= p \mid \neg\varphi \mid \varphi \wedge \varphi \mid \langle\rho\rangle\varphi \mid \langle\!\langle x\rangle\!\rangle\varphi \mid (x,a)\varphi \mid C_A\varphi \mid D_A\varphi$$
$$\rho ::= \phi \mid \varphi? \mid \rho + \rho \mid \rho;\rho \mid \rho^*$$

where $p \in AP$, $x \in Var$, $a \in Ag$, $A \subseteq Ag$, *and* ϕ *is a propositional formula (i.e., a Boolean combination over AP).*

Semantics. We interpret the logic EDSL on interpreted systems. To interpret the epistemic operators we introduce an indistinguishability relation $\sim_a$ on S, for every agent $a \in Ag$, such that $s \sim_a s'$ iff $s(a) = s'(a)$ [8]. We extend $\sim_a$ to histories as follows: for $h, h' \in Hist$ define $h \equiv_a h'$ if $|h| = |h'|$ and $h(i) \sim_a h'(i)$ for all $i \leq |h|$. Let $\sim_A^C$ be $(\cup_{a\in A} \sim_a)^*$, where $*$ denotes the reflexive and transitive closure (w.r.t. relation composition), and its extension to histories $\equiv_A^C$. Also, let $\sim_A^D$ be $\cap_{a\in A} \sim_a$, and let $\equiv_A^D$ be its extension to histories.

A *strategy* is a function of the form $\sigma : Hist \to Act$, and let Str denote the set of all strategies. Strategy σ is *coherent for a* if action $\sigma(h)$ is available to a in local state $last(h)(a)$; it is *uniform for a* if $h \sim_a h'$ implies $\sigma_a(h) = \sigma_a(h')$, that is, in indistinguishable states a is bound to play the same action [9]. An *assignment* is a function $\chi : Var \cup Ag \to \text{Str}$ such that for every $a \in Ag$ the strategy $\chi(a)$ is coherent and uniform for a. For $x \in Var$ and $\sigma \in \text{Str}$, the *variant* χ^x_σ is the assignment that maps x to σ and coincides with χ on all other variables and agents. Similarly, if $a \in Ag$ and σ is coherent and uniform for a, then the *variant* χ^a_σ is the assignment that maps a to $\sigma \in \text{Str}$ and coincides with χ on all other variables and agents. An assignment χ is *φ-compatible* if, for every $x \in Var$, the strategy $\chi(x)$ is coherent and uniform for every agent in $shr(x,\varphi) = \{a \in Ag \mid (x,a)\psi \text{ is a subformula of } \varphi\}$. Here $shr(x,\varphi)$ represents the set of agents using strategy x in evaluating formula φ. We write $\text{out}(h,\chi)$ for the set of histories h' *generated by* χ, i.e., $h' \in \text{out}(h,\chi)$ iff h is a prefix of h' and for every i with $|h| \leq i < |h'|$, $h'_{i+1} = \tau(h'_i, \prod_{a\in Ag} \chi(a)(h'_{\leq i}))$.

Definition 4 (Satisfaction). *Define the satisfaction relation* $(\mathrm{S}, h, \chi) \models \varphi$, *where* $h \in Hist$, φ *is an EDSL-formula, and* χ *is a* φ*-compatible assignment as follows:*[1]

$(\mathrm{S}, h, \chi) \models p$	*iff*	$last(h) \in \pi(p)$, *for* $p \in AP$
$(\mathrm{S}, h, \chi) \models \neg\varphi_1$	*iff*	*it is not the case that* $(\mathrm{S}, h, \chi) \models \varphi_1$
$(\mathrm{S}, h, \chi) \models \varphi_1 \wedge \varphi_2$	*iff*	$(\mathrm{S}, h, \chi) \models \varphi_i$ *for* $i \in \{1,2\}$
$(\mathrm{S}, h, \chi) \models \langle\!\langle x\rangle\!\rangle\varphi_1$	*iff*	*there exists a* φ_1*-compatible variant* χ^x_σ *such that* $(\mathrm{S}, h, \chi^x_\sigma) \models \varphi_1$
$(\mathrm{S}, h, \chi) \models (x,a)\varphi_1$	*iff*	$(\mathrm{S}, h, \chi^a_{\chi(x)}) \models \varphi_1$
$(\mathrm{S}, h, \chi) \models C_A\varphi_1$	*iff*	*for every history* $h' \in Hist$, $h' \equiv^C_A h$ *implies* $(\mathrm{S}, h', \chi) \models \varphi_1$
$(\mathrm{S}, h, \chi) \models D_A\varphi_1$	*iff*	*for every history* $h' \in Hist$, $h' \equiv^D_A h$ *implies* $(\mathrm{S}, h', \chi) \models \varphi_1$
$(\mathrm{S}, h, \chi) \models \langle\rho\rangle\varphi_1$	*iff*	*for some* $h' \in \text{out}(h,\chi)$, $(h,h') \in R(\rho,\chi)$ *and* $(\mathrm{S}, h', \chi) \models \varphi_1$

where $R(\rho,\chi) \subseteq Hist \times Hist$ *(which also depends on* S*, although we supress it) is defined as follows:*

$R(\phi,\chi)$	=	$\{(h,h') \mid \lvert h'\rvert = \lvert h\rvert + 1, (\mathrm{S}, h, \chi) \models \phi\}$
$R(\varphi?,\chi)$	=	$\{(h,h) \mid (\mathrm{S}, h, \chi) \models \varphi\}$
$R(\rho_1 + \rho_2,\chi)$	=	$R(\rho_1,\chi) \cup R(\rho_2,\chi)$
$R(\rho_1;\rho_2,\chi)$	=	$R(\rho_1,\chi) \circ R(\rho_2,\chi)$
$R(\rho^*,\chi)$	=	$R^*(\rho,\chi)$

where $R \circ R'$ *is the sequential composition of relations* R *and* R', *and* R^* *is the reflexive and transitive closure of* R *under the sequential composition operation.*

Write $\mathrm{S} \models \varphi$ to mean that $(\mathrm{S}, s, \chi) \models \varphi$ for every initial state $s \in S_0$ (note that states are histories of length 1) and every assignment χ. Given an interpreted system S and an EDSL formula φ, the model checking problem concerns establishing whether $\mathrm{S} \models \varphi$.

Theorem 1. *Model checking IS against EDSL is undecidable.*

We can now state the main result of this abstract. The theorem below shows that verifying *BMNPA* against EDSL is decidable. This is in contrast with general undecidability results of the verification and synthesis of MAS for agents with perfect recall and incomplete information for much weaker logics [1, 11]. Define $\exp(k,n)$ by recursion as follows: $\exp(1,n) = 2^n$ for all $n \in \mathbb{N}$, and $\exp(k+1,n) = \exp(1, exp(k,n))$ for $k \geq 1$.

Theorem 2. *Model checking BMNPA against EDSL is decidable: model checking interpreted systems of size n against formulas of size k is in* $\exp(2^{2^{O(nk)}}, O(n))$*-time and is* $exp(k, O(1))$*-space hard.*

Discussion. The results above show that the *BMNPA* class of MAS admits decidable verification. In common with existing proposals [3], a key ingredient of *BMNPA* is infinite runs generated by public actions, but, differently from the existing state of the art, *BMNPA* admit bounded rounds of non-public communication. This permits the modelling of properties such as collusion and, generally, properties resulting from non-public, or hidden, actions. This is relevant in a number of contexts including when studying unreliable communication channels where some of the actions, or messages, may not reach all its intended recipients.

In the future we intend to study applications of *BMNPA*. We believe that this will be facilitated by the set-up based on interpreted systems, rather than more abstract models such as CGS as in [3], and by the considerable expressive power offered by the specification language here used. We also intend to try and identify a noteworthy subclass of *BMNPA* for which the verification problem admits a lower complexity.

Acknowledgements. The authors acknowledge support of: ANR JCJC Project SVeDaS (ANR-16-CE40-0021); the Royal Academy of Engineering and the Royal Society; INDAM GNCS 2018 Project "Metodi Formali per la Verifica e la Sintesi di Sistemi Discreti e Ibridi".

[1] It can be shown by induction that satisfaction is well-defined: In expressions $(S, h, \chi') \models \varphi'$, on the right-hand sides assignment χ' is always φ'-compatible.

REFERENCES

[1] R. Alur, T. A. Henzinger, and O. Kupferman. 2002. Alternating-Time Temporal Logic. *J. ACM* 49, 5 (2002), 672–713.

[2] F. Belardinelli, A. Lomuscio, A. Murano, and S. Rubin. 2017. Verification of Broadcasting Multi-Agent Systems against an Epistemic Strategy Logic. In *IJCAI'17*. 91–97.

[3] F. Belardinelli, A. Lomuscio, A. Murano, and S. Rubin. 2017. Verification of Multi-agent Systems with Imperfect Information and Public Actions. In *AAMAS'17*. 1268–1276.

[4] R. Berthon, B. Maubert, A. Murano, S. Rubin, and M.Y. Vardi. 2017. Strategy logic with imperfect information. In *LICS'17*. 1–12.

[5] P. Cermák, A. Lomuscio, F. Mogavero, and A. Murano. 2014. MCMAS-SLK: A Model Checker for the Verification of Strategy Logic Specifications. In *CAV'14 (LNCS)*, Vol. 8559. 525–532.

[6] P. Cermák, A. Lomuscio, F. Mogavero, and A. Murano. 2015. Verifying and Synthesising Multi-Agent Systems against One-Goal Strategy Logic Specifications. In *AAAI '15*. 2038–2044.

[7] C. Dima and F.L. Tiplea. 2011. Model-checking ATL under Imperfect Information and Perfect Recall Semantics is Undecidable. *CoRR* abs/1102.4225 (2011).

[8] R. Fagin, J.Y. Halpern, Y. Moses, and M.Y. Vardi. 1995. *Reasoning about Knowledge*. MIT.

[9] W. Jamroga and W. van der Hoek. 2004. Agents that Know How to Play. *Fund. Inf.* 62 (2004), 1–35.

[10] A. Lomuscio, R. van der Meyden, and M. Ryan. 2000. Knowledge in Multi-Agent Systems: Initial Configurations and Broadcast. *TOCL* 1, 2 (2000), 246–282.

[11] R. van der Meyden and T. Wilke. 2005. Synthesis of Distributed Systems from Knowledge-Based Specifications. In *CONCUR'05*. 562–576.

[12] F. Mogavero, A. Murano, G. Perelli, and M.Y. Vardi. 2014. Reasoning About Strategies: On the Model-Checking Problem. *TOCL* 15, 4 (2014), 34:1–34:47.

[13] M. Y. Vardi. 2011. The rise and fall of LTL. In *GandALF 2011, EPTCS 54*.