ESSAI 2024 course: Logic-based specification and verification of multi-agent systems Lecture 4.2: Specification and Verification of Homogeneous Dynamic Multi-agent Systems

#### Valentin Goranko Stockholm University



2nd European Summer School on Artificial Intelligence ESSAI 2024 Athens, July 15-19, 2024



Based on:

Riccardo De Masellis and Valentin Goranko: Logic-based Specification and Verification of Homogeneous Dynamic Multi-agent Systems, J. of Autonomous Agents and Multiagent Systems, vol. 34 (2), 2020.



## Introduction: Homogeneous Dynamic Multi-agent Systems (HDMAS)

We consider multi-agent systems which evolve by means of discrete transitions from state to state, and have two special features:

#### Homogeneous:

- ▶ all agents have the same type (protocol). In particular:
  - all agents have the same available actions at each system state;
  - the effect of any action does not depend on which agent performs it, but only on how many agents perform it.
- thus, transitions are completely determined by how many agents perform each possible action.

#### Dynamic:

- unbounded (but always finite) number of agents in the system.
- at every round, each agent may be 'active', by taking a 'real' action, or 'inactive' by performing action 'idle'.
  So, agents can 'join' and 'leave' the system dynamically.



All agents in a HDMAS are split into

- controllable (by the leader/supervisor), and
- uncontrollable (regarded as environment or adversaries).

All controllable and uncontrollable agents have the same type.

However, the controllable agents follow a prescribed strategy, whereas the uncontrollable ones are unconstrained.



# Homogeneous Dynamic Multi-agent Systems: some examples

- sensor/computer/social networks;
- election systems and voting procedures;
- more abstractly, systems of strategic dynamic resource allocation, such as Colonel Blotto games.



Assuming perfect coordination between the controllable agents, an HDMAS can be regarded as a concurrent game between two 'super-players': Proponent and Opponent.

Proponent has a (temporalised) qualitative objective, for instance:

- to eventually reach a desired goal state in the system, or

to keep the system in a safe region, until a goal state is reached.
Opponent tries to prevent the achievement of that objective.

A typical property to specify and verify in an HDMAS:

Proponent has a joint strategy for a coalition of M controllable agents acting against (Opponent represented by) at most N non-controllable agents, to ensure that Proponent's objective is satisfied on every play enabled by that strategy.



### HDMAS: technical preliminaries

- $Ag = \{ag_1, ag_2, \ldots\}$ : unbounded universe of agents.
- Act = {act<sub>1</sub>,..., act<sub>n</sub>}: a finite set of (names for) actions. Act<sup>+</sup> = Act ∪ {ε}, where ε is a specific 'idle' action.
- X = {x<sub>1</sub>,...,x<sub>n</sub>} and X<sup>+</sup> = X ∪ {x<sub>ε</sub>}: a fixed set of variables, called action counters, respectively associated with the actions in Act<sup>+</sup>.
- action profile: a tuple of actions in  $Act^+$ , one for each agent in Ag.
- An action distribution for an action profile σ is a function act<sub>σ</sub> : X → N, where, for i = 1, ..., n, act<sub>σ</sub>(x<sub>i</sub>) is the number of agents taking action act<sub>i</sub> in σ. (The idle action ε is not counted.)
- A (transition) guard is an open formula of Presburger arithmetic (PrA) over the variables in X.
- ► Satisfaction of a transition guard g by an action distribution act, denoted act |= g, is defined by the standard semantics of PrA.



### HDMAS models: technical definition

An **HDMAS** structure for a set of agents *Ag* and a set of actions *Act*<sup>+</sup>:

 $\mathcal{S} = \langle Ag, Act^+, S, \delta \rangle$ 

► *S* is a (usually finite) set of **states**.

•  $\delta: S \times S \rightarrow G$  is a transition guards function.

 $\delta$  labels all possible transitions between states with guards that determine, for every possible action distribution, a unique transition to a successor state.

NB: the transition caused by any action profile only depends on its action distribution.

Hereafter, we use action distributions instead of action profiles.

HDMAS-based model: HDMAS structure + labelling  $\lambda$  of states with sets of atomic propositions in a fixed set  $\Phi$ .



#### An example of HDMAS



The model involves 6 states and 2 atomic propositions (p and q). Besides  $\varepsilon$ , there are 3 'real' actions, with respective counters  $x_1, x_2, x_3$ . The guards:

$$\begin{array}{ll} g_1 := & (x_1 \geq 2x_2) \land (x_3 \leq 3) & g_4 := & x_1 > 5 \land (3x_2 < x_1 + 2x_3) \\ g_2 := & (x_1 + x_2 + x_3 \leq 10) \land (x_3 > 3) & g_5 := & x_1 = x_1 \\ g_3 := & (x_1 > 5) \land (x_3 > x_1) & g_6 := & x_1 + 2x_2 \geq x_3 \\ & g_7 := & x_2 = x_3 \end{array}$$

 $\mathcal{L}_{HDMAS}$  is a logic (extending ATL) for specifying strategic properties in HDMAS, such as existence of a strategy for the controllable agents that guarantees satisfaction of Proponent's objective against any behavour of the uncontrollable agents.



The vocabulary of  $\mathcal{L}_{HDMAS}$  contains:

- ► a fixed set of atomic propositions  $\Phi = \{p_1, p_2, ....\}$ ,
- ► a set of two special variables Y = {y<sub>1</sub>, y<sub>2</sub>} ranging over N, called agent counters, representing respectively the current numbers of controllable and uncontrollable agents,
- ▶ a set of auxiliary agent-counting parameters  $Z = \{z_1, z_2, ...\}$ .

The set of **terms**:  $T = \mathbb{N} \cup Y \cup Z$ .

(Natural numbers will be identified with their numerals.)



Two sorts of formulae:

Path formulae:  $\chi ::= X \varphi | G \varphi | \varphi U \varphi$ , where  $\varphi$  are state formulae.

#### State formulae:

$$\varphi ::= \top \mid p \mid \neg \varphi \mid (\varphi \land \varphi) \mid (\varphi \lor \varphi) \mid \forall y \varphi \mid \exists y \varphi \mid \langle \langle t_1, t_2 \rangle \rangle \chi$$

where  $p \in \Phi$ ,  $y \in Y$ ,  $t_1 \in T \setminus \{y_2\}$ ,  $t_2 \in T \setminus \{y_1\}$ , and  $\chi$  is a path formula.

In the formula  $\langle\!\langle t_1, t_2 \rangle\!\rangle \chi$ ,  $t_1$  denotes the number of controllable agents, and  $t_2$  – the number of uncontrollable agents.

**Positive polarity constraint** for the formulae  $\forall y\varphi$  and  $\exists y\varphi$ : all free occurrences of y in  $\varphi$  must have a positive polarity (i.e., must be in the scope of an even number of negations).



#### ► (((7,5)) X p:

*"7 controllable agents have a joint action ensuring against 5 uncontrollable agents that any outcome state satisfies p" .* 

 $\blacktriangleright \forall y_2 \langle\!\langle 7, y_2 \rangle\!\rangle \, \mathsf{G} \, \neg p:$ 

"for any number  $y_2$ , 7 controllable agents have a joint strategy to ensure against  $y_2$  uncontrollable agents that any outcome play never reaches a *p*-state."

 $\blacktriangleright \forall y_2 \exists y_1 \langle \langle y_1, y_2 \rangle \rangle q \cup \neg p:$ 

"For any number  $(y_2)$  of uncontrollable agents there is a number  $(y_1)$  of controllable agents who have a joint strategy to ensure that any outcome play will stay within a q-region until it eventually reaches a non-p-state."



**Assignment**: a function  $\theta : T \to \mathbb{N}$ , where  $\theta(i) = i$  for  $i \in \mathbb{N}$ . Let  $\mathcal{M}$  be a HDMAS, *s* be a state and  $\theta$  an assignment in it. The satisfaction relation  $\models$  is inductively defined as follows:

• 
$$\mathcal{M}, s, \theta \models p$$
 iff  $p \in \lambda(s)$ ,

- ▶ The semantics of  $\top$ ,  $\land$ ,  $\lor$ ,  $\neg$ ,  $\forall$ ,  $\exists$ : as in classical logic.
- $\mathcal{M}, s, \theta \models \langle\!\langle t_1, t_2 \rangle\!\rangle \chi$  iff

there exists a joint strategy  $\sigma$  for a coalition of  $\theta(t_1)$  controllable agents such that every play enabled by  $\sigma$ , against (at most)  $\theta(t_2)$ uncontrollable agents, satisfies the temporal objective  $\chi$ .



### Monotonicity properties and quantifier elimination equivalences

#### $\langle\!\langle t_1, t_2 \rangle\!\rangle \chi$ is monotone with respect to $t_1$ and anti-monotone with respect to $t_2$ .

That means:

- If  $\mathcal{M}, s, \theta \models \langle\!\langle t_1, t_2 \rangle\!\rangle \chi$  and  $C > t_1$  then  $\mathcal{M}, s, \theta \models \langle\!\langle C, t_2 \rangle\!\rangle \chi$ ;
- ► If  $\mathcal{M}, s, \theta \models \langle\!\langle t_1, t_2 \rangle\!\rangle \chi$  and  $N < t_2$  then  $\mathcal{M}, s, \theta \models \langle\!\langle t_2, N \rangle\!\rangle \chi$ .

Based on these, some quantifier elimination equivalences are valid, e.g.:

$$\blacktriangleright \forall y_1 \langle\!\langle y_1, t \rangle\!\rangle \chi \equiv \langle\!\langle 0, t \rangle\!\rangle \chi [0/y_1],$$

- $\blacktriangleright \exists y_2 \langle\!\langle t, y_2 \rangle\!\rangle \chi \equiv \langle\!\langle t, 0 \rangle\!\rangle \chi [0/y_2],$
- $\forall y_2 \forall y_1 \langle\!\langle y_1, y_2 \rangle\!\rangle \chi \equiv \forall y_2 \langle\!\langle 0, y_2 \rangle\!\rangle \chi [0/y_1],$
- ►  $\exists y_1 \exists y_2 \langle\!\langle y_1, y_2 \rangle\!\rangle \chi \equiv \exists y_1 \langle\!\langle y_1, 0 \rangle\!\rangle \chi [0/y_2], \text{ etc.}$



The full language  $\mathcal{L}_{HDMAS}$  is not suitable for algorithmic model checking, because of the unconstrained nesting of strategic operators and quantification over agent counters.

For model checking we transform  $\mathcal{L}_{HDMAS}$ -formulae to normal form, by imposing syntactic restrictions on the patterns of quantification.

Informally, the formulae in  $\mathcal{L}_{HDMAS}^{NF}$  are defined by modifying the recursive definition of state formulae of  $\mathcal{L}_{HDMAS}$ , where the clauses  $\forall y \varphi$  and  $\exists y \varphi$  are replaced with the following, where  $\chi$  is a temporal objective:

 $\exists y_1 \langle\!\langle y_1, t_2 \rangle\!\rangle \chi \mid \forall y_2 \exists y_1 \langle\!\langle y_1, y_2 \rangle\!\rangle \chi \mid \forall y_2 \langle\!\langle t_1, y_2 \rangle\!\rangle \chi \mid \exists y_1 \forall y_2 \langle\!\langle y_1, y_2 \rangle\!\rangle \chi$ 

 $\mathcal{L}_{HDMAS}^{NF}$ : the fragment of  $\mathcal{L}_{HDMAS}$ -formulae in normal form.



Normal forms restrict the language syntactically, but not its expressiveness.

- A key technical result: a recursive procedure NF, converting every  $\mathcal{L}_{HDMAS}$ -formula  $\varphi$  into NF( $\varphi$ )  $\in \mathcal{L}_{HDMAS}^{NF}$ , such that:
  - 1.  $NF(\varphi) \equiv_{fin} \varphi$ . ( $\equiv_{fin}$  is equivalence on all finite HDMAS models)
  - 2. If  $\varphi \in \mathcal{L}_{HDMAS}^{NF}$  then  $NF(\varphi) = \varphi$ .
  - 3. NF( $\varphi$ ) can be computed effectively and has length linearly bounded above by the length of  $\varphi$ .

Thus, normal forms restrict the language syntactically, but do not reduce its expressiveness over finite models.



#### Fixpoint equivalences for formulae in normal form

The strategic operators for formulae in  $\mathcal{L}_{HDMAS}^{NF}$  satisfy fixpoint equivalences over finite models, listed in the theorem below.

**Theorem.** For every terms t, t', t'' the following equivalences hold, where the formulae on the left are in  $\mathcal{L}_{HDMAS}^{NF}$ .

$$\blacktriangleright \langle\!\langle t', t'' \rangle\!\rangle \,\mathsf{G}\,\varphi \equiv \varphi \wedge \langle\!\langle t', t'' \rangle\!\rangle \,\mathsf{X}\,\langle\!\langle t', t'' \rangle\!\rangle \,\mathsf{G}\,\varphi$$

$$\blacktriangleright \quad \langle\!\langle t',t''\rangle\!\rangle \,\psi \,\mathsf{U}\,\varphi \equiv \varphi \lor (\psi \land \langle\!\langle t',t''\rangle\!\rangle \,\mathsf{X}\,\langle\!\langle t',t''\rangle\!\rangle \,\psi \,\mathsf{U}\,\varphi)$$

$$\blacktriangleright \exists y_1 \langle\!\langle y_1, t \rangle\!\rangle \, \mathsf{G} \, \varphi \equiv_{\mathsf{fin}} \varphi \land \exists y_1 \langle\!\langle y_1, t \rangle\!\rangle \, \mathsf{X} \, \exists y_1 \langle\!\langle y_1, t \rangle\!\rangle \, \mathsf{G} \, \varphi$$

 $\blacktriangleright \quad \forall y_2 \langle\!\langle t, y_2 \rangle\!\rangle \, \mathsf{G} \, \varphi \equiv_{\mathsf{fin}} \varphi \land \forall y_2 \langle\!\langle t, y_2 \rangle\!\rangle \, \mathsf{X} \, \forall y_2 \langle\!\langle t, y_2 \rangle\!\rangle \, \mathsf{G} \, \varphi$ 

$$\blacktriangleright \exists y_1 \langle\!\langle y_1, t \rangle\!\rangle \, \psi \, \mathsf{U} \, \varphi \equiv_{\mathsf{fin}} \varphi \lor (\psi \land \exists y_1 \langle\!\langle y_1, t \rangle\!\rangle \, \mathsf{X} \, \exists y_1 \langle\!\langle y_1, t \rangle\!\rangle \, \psi \, \mathsf{U} \, \varphi)$$

- $\blacktriangleright \forall y_2 \langle\!\langle t, y_2 \rangle\!\rangle \psi \, \mathsf{U} \, \varphi \equiv_{\mathsf{fin}} \varphi \lor (\psi \land \forall y_2 \langle\!\langle t, y_2 \rangle\!\rangle \, \mathsf{X} \, \forall y_2 \langle\!\langle t, y_2 \rangle\!\rangle \, \psi \, \mathsf{U} \, \varphi)$
- $\blacktriangleright \ \forall y_2 \exists y_1 \langle\!\langle y_1, y_2 \rangle\!\rangle \operatorname{\mathsf{G}} \varphi \equiv_{\operatorname{\mathsf{fin}}} \varphi \land \forall y_2 \exists y_1 \langle\!\langle y_1, y_2 \rangle\!\rangle \operatorname{\mathsf{X}} \forall y_2 \exists y_1 \langle\!\langle y_1, y_2 \rangle\!\rangle \operatorname{\mathsf{G}} \varphi.$
- $\blacktriangleright \exists y_1 \forall y_2 \langle\!\langle y_1, y_2 \rangle\!\rangle \operatorname{\mathsf{G}} \varphi \equiv_{\operatorname{\mathsf{fin}}} \varphi \land \exists y_1 \forall y_2 \langle\!\langle y_1, y_2 \rangle\!\rangle \operatorname{\mathsf{X}} \exists y_1 \forall y_2 \langle\!\langle y_1, y_2 \rangle\!\rangle \operatorname{\mathsf{G}} \varphi.$
- $\blacktriangleright \forall y_2 \exists y_1 \langle\!\langle y_1, y_2 \rangle\!\rangle \psi \, \mathsf{U} \, \varphi \equiv_{\mathsf{fin}} \varphi \lor (\psi \land \forall y_2 \exists y_1 \langle\!\langle y_1, y_2 \rangle\!\rangle \, \mathsf{X} \, \forall y_2 \exists y_1 \langle\!\langle y_1, y_2 \rangle\!\rangle \, \psi \, \mathsf{U} \, \varphi).$
- $\blacktriangleright \exists y_1 \forall y_2 \langle\!\langle y_1, y_2 \rangle\!\rangle \psi \, \mathsf{U} \, \varphi \equiv_{\mathsf{fin}} \varphi \lor (\psi \land \exists y_1 \forall y_2 \langle\!\langle y_1, y_2 \rangle\!\rangle \, \mathsf{X} \, \exists y_1 \forall y_2 \langle\!\langle y_1, y_2 \rangle\!\rangle \, \psi \, \mathsf{U} \, \varphi).$



# Model checking of formulae in $\mathcal{L}_{HDMAS}^{NF}$

Given a state formula  $\varphi$  of  $\mathcal{L}_{HDMAS}$ , a HDMAS model  $\mathcal{M}$ , a state  $s \in \mathcal{M}$ , and an assignment  $\theta$  in  $\mathcal{M}$ :

- ► the local model checking problem for HDMAS is the problem of deciding whether M, s, θ ⊨ φ,
- the global model checking problem is the problem of computing the state extension of φ in M for θ, formally defined as:

$$\llbracket \varphi \rrbracket_{\mathcal{M}}^{\theta} = \{ s \in S \mid \mathcal{M}, s, \theta \models \varphi \}.$$

Recall: the transitions in HDMAS models are represented symbolically, in terms of the guards.

So, an explicit representation of the transition graph is generally infinite.

That is why, the model checking of HDMAS models is done symbolically, by reduction to Presburger arithmetic (PrA).



# Algorithm for global model checking in $\mathcal{L}_{HDMAS}^{NF}$ : the core sub-procedure PREIMG

For a set of states  $Q \subseteq S$  and integers  $C, N \in \mathbb{N}$ , the

(C, N)-controllable pre-image of Q is the set of states from which C controllable agents have a joint action, which, when played against *any* joint action of N uncontrollable agents produces an outcome state in Q.

The procedure PREIMG returns the (C, N)-controllable pre-image of Q.

PREIMG is extended to compute, for any terms  $t_1$ ,  $t_2$ , the  $(t_1, t_2)$ -controllable pre-image of Q by means of a PrA-formula with  $t_1$ ,  $t_2$  as parameters.



# Algorithm for global model checking in $\mathcal{L}_{HDMAS}^{NF}$ : informal description

1. For  $\varphi = \langle\!\langle t_1, t_2 \rangle\!\rangle X \psi$ , PREIMG applied to  $Q = \llbracket \psi \rrbracket_{\mathcal{M}}^{\theta}$ , computes the state extension of  $\varphi$ , as a PrA-formula parameterised with  $t_1, t_2$ .

2. Then the procedure is readily extended to all quantified extensions of  $\langle\!\langle t_1, t_2 \rangle\!\rangle X \psi$ , by adding the respective quantification to the result.

3. Lastly, for the long-term temporal objectives model checking is done by fixpoint unfolding iterations (like in model checking of CTL or ATL). Every iteration stage produces again a PrA-formula.

Stabilisation and reaching the fixpoint is detected by checking equivalence of the PrA-formulae produced at the successive iterations.

So, in all cases, computing the state extension of a  $\mathcal{L}_{HDMAS}^{NF}$ -formula is reduced to computing a PrA-formula.

Thus, the local model checking problem for  $\mathcal{L}_{HDMAS}^{NF}$  is reduced to checking the truth of PrA-formulae.



# Local model checking in $\mathcal{L}_{HDMAS}^{NF}$ : example 1

 $g_3 := (x_1 > 5) \land (x_3 > x_1)$  $g_4 := x_1 > 5 \land (3x_2 - 2x_3 < x_1)$ 

 $g_5 := \dots$ 



## Global model checking problem for $\mathcal{L}_{HDMAS}^{NF}$ : example 2

Computing  $\llbracket \varphi \rrbracket_{\mathcal{M}}$  for  $\varphi = \exists y_1 \forall y_2 \langle \langle y_1, y_2 \rangle \rangle X (p \lor q)$  in the given model  $\mathcal{M}$ .



1. Compute  $[\![p \lor q]\!]_{\mathcal{M}} = \{s_2, s_3, s_4, s_5, s_6\}.$ 

2. For each  $s \in \mathcal{M}$  check the truth of  $\exists y_1 \forall y_2 \Pr{F(\mathcal{M}, s, y_1, y_2, \llbracket p \lor q \rrbracket_{\mathcal{M}})}$ .

- 11 uncontrollable agents can keep the system in  $s_1$  by all performing  $act_3$ ; so,  $\exists y_1 \forall y_2 \Pr F(\mathcal{M}, s_1, y_1, y_2, \llbracket p \lor q \rrbracket_{\mathcal{M}})$  is false, hence  $s_1$  is not in the  $\exists y_1 \forall y_2(y_1, y_2)$ - controllable pre-image of  $\llbracket p \lor q \rrbracket_{\mathcal{M}}$ .

- All outgoing transitions from  $s_2$  lead to states in  $\llbracket p \lor q \rrbracket_{\mathcal{M}}$ ; hence  $\exists y_1 \forall y_2 \Pr F(\mathcal{M}, s_2, y_1, y_2, \llbracket p \lor q \rrbracket_{\mathcal{M}})$  is true, so  $s_2$  is in the  $\exists y_1 \forall y_2(y_1, y_2)$ controllable pre-image of  $\llbracket p \lor q \rrbracket_{\mathcal{M}}$ .

- Checking all other states likewise produces the final result:  $\llbracket \varphi \rrbracket_{\mathcal{M}} = \{s_2, s_4, s_5, s_6\}.$ 



## Global model checking problem for $\mathcal{L}_{HDMAS}^{NF}$ : example 3

Computing  $\llbracket \psi \rrbracket_{\mathcal{M}}$  for  $\psi = \langle\!\langle 7, 4 \rangle\!\rangle X (\forall y_2 \exists y_1 \langle\!\langle y_1, y_2 \rangle\!\rangle G p)$  in the model  $\mathcal{M}$ .



1. Initialize  $Z \leftarrow \{s_2, s_3, s_4\}$  and  $W \leftarrow S = \{s_1, \dots, s_6\}$ . A while-loop computing the fixpoint:  $-W \leftarrow \{s_2, s_3, s_4\};$   $-\operatorname{PREIMG}(\mathcal{M}, y_1, y_2, \{s_2, s_3, s_4\}, \theta, \forall y_2 \exists y_1) = \{s_2, s_4, s_5\};$   $-Z \leftarrow \{s_2, s_4, s_5\} \cap \{s_2, s_3, s_4\} = \{s_2, s_4\}.$ Next round:  $Z \leftarrow \dots \{s_2, s_4\}.$ Now the fixpoint is reached.

So,  $\llbracket \forall y_2 \exists y_1 \langle \langle y_1, y_2 \rangle \rangle$  G  $\rho \rrbracket_{\mathcal{M}} = \{s_2, s_4\}.$ 

Lastly, for computing the outer Next-formula the algo calls the  $\ensuremath{\mathrm{PREIMG}}$  procedure.

For each  $s \in S$  the truth of formula  $PrF(\mathcal{M}, s, 7, 4, \{s_2, s_4\})$  is called.

The final result is  $\llbracket \psi \rrbracket_{\mathcal{M}} = \{s_4, s_5\}.$ 



Complexity of model checking of  $\mathcal{L}_{HDMAS}^{NF}$ - formulae: by using results (by Hasse and others) on complexity of model checking of PrA-formulas.

- Ranges from Σ<sub>3</sub><sup>EXP</sup> in the general case, to NP-complete when the number of controllable or uncontrollable agents is fixed or bounded.
- ▶ When the number of actions is bounded, too, it is **P**-complete.



Future works (in some possible futures) include:

- Allowing several types of agents.
- Allowing several coalitions of controlled agents.
- Extending the language, e.g. by relaxing some syntactic restrictions.
- Possible applications include:
  - solving games of the type of generalised Colonel Blotto games by model checking  $\mathcal{L}_{\text{HDMAS}}^{\text{NF}}\text{-}\text{formulae}.$
  - design and verification of sensor networks and voting procedures.
  - etc.

#### **END OF LECTURE 4.2**

