Scientific report for the COST action

Emiliano Lorini

Institut de Recherche en Informatique de Toulouse (IRIT), CNRS, France

During my stay at IIIA-CSIC in Barcelona I had an intense collaboration with Pablo Noriega, Carles Sierra, Pilar Dellunde, Marc Esteva and Jordi Sabater-Mir. We have discussed on many different topics related with the theory of institutions, speech act theory and game theory.

In the first part of my stay I have presented my general approach to institutions based on the notion of acceptance [4, 5]. I have discussed with Pablo Noriega, Carles Sierra, Pilar Dellunde, Marc Esteva about the integration of this model with a BDI architecture of an autonomous agent, in such a way that the extended model could explain not only how norms are established through agreement by the agents in a population, but also how norms are internalized by an agent and how they affect his decisions. We have also investigated the relationship between speech acts (orders, promises, delegation, etc.) and institutions:

- how speech acts are responsible for norm change (e.g. how order and delegation create obligations, how promises create commitments, etc.)
- how certain rules of institutions specify the sequences of speech acts that are executable, and the conditions under which a given speech act is executable (e.g. during an auction, a person can make a bid only if she is registered in the auction).

In the second part of my stay at IIIA-CSIC, I gave a seminar on game theory and on the logical representation of individual preferences and group preferences with title "A Logical Account of Social Rationality in Strategic Games".

Moreover, together with Pablo Noriega and Carles Sierra, I have put the basis for a joint the work on the logical representation of institutions. The idea of this work is to provide a simple logical formalism which allows to formalize institutions and their dynamic aspects, and which directly maps into the model of institutions proposed by the IIIA-CSIC group [1].

The following are the main features of the notion of institution that we have discussed and that the logical formalism should be able to represent:

1. An institution can be seen as a set of *scenes* (or contexts) *plus* a set of *variables* that are instantiated with certain values. Variables are used to specify the current state of institution. For example, in the case of an auction, a variable is used to specify the current bid of an agent playing in the auction.

2. A scene is a set of rules that specifies the *actions* that can be performed (i.e. are allowed) in a given situation, and the effects that an action will produce when it is performed in a given situation. For example, in the case of an auction, there might be a rule saying that one can make a bid only if he is registered in the auction. There might be another rule saying that if a person bids a certain amount of euros for a given product and nobody else makes a higher bid, then the person will be the buyer of the product.

3. An action is nothing else than the process of changing the value of a given variable. Therefore, a first kind of institutional dynamics consists in changing the values of the variables specifying the current state of institution. For example, in the case of an auction, a person might bid 300 euros for a given product (i.e. set to 300 the value of the variable *Bid*).

4. Another kind of institutional dynamics consists in changing the rules of a scene in a given institution. For example, in the case of an auction, one might cancel the rule saying that a person can make a bid only if she is registered in the auction. After this rule has been cancel, it will possible to make a bid without being registered in the auction.

In the following section I am going to sketch the main ingredients of the logical formalism that we have elaborated during my stay at IIIA-CSIC and that we intend to use as a specification language for institutions.

1 DLI: towards a dynamic logic of institutions

This section introduces the main features of the syntax and the semantics of the logic DLI: a dynamic logic of institutions.

1.1 Language

Let $\mathbb{V} = \{\mathbf{x}, \mathbf{y}, \ldots\}$ be a non-empty finite set of variables. We suppose that each variable $\mathbf{x} \in \mathbb{V}$ takes values from a non-empty finite set of variable assignments $Val_{\mathbf{x}}$. For each set $Val_{\mathbf{x}}$, $Inst_{\mathbf{x}}$ is the corresponding set of all possible instantiations of variable $\mathbf{x}$. For example, suppose that $Val_{\mathbf{x}} = \{v_1, \ldots, v_r\}$ then $Inst_{\mathbf{x}} = \{\mathbf{x}{=}v_1, \ldots, \mathbf{x}{=}v_r\}$. We denote by *Inst* the set of all possible instantiations of all variables, that is: $Inst = \bigcup_{\mathbf{x} \in \mathbb{V}} Inst_{\mathbf{x}}$.

$At = \{\mathbf{x}{\leftarrow}v | \mathbf{x} \in \mathbb{V} \text{ and } v \in Val_{\mathbf{x}}\}$ is the set of variable assignments. For instance $\mathbf{x}{\leftarrow}v$ is the event of setting to v the value of the variable $\mathbf{x}$. Variable assignment is a generalization of the concept of assignment in the sense of [3, 2] where variables are propositional, i.e. they can have only two values (1 for true or 0 for false).

The language of DLI is made up of events α and formulas φ and is defined by the following BNF:

$$\alpha \quad ::= \quad a \mid \alpha;\alpha \mid \alpha \cup \alpha$$
$$\varphi \quad ::= \quad \mathbf{x}{=}v \mid a \rightsquigarrow b \mid \mathrm{Pre}(a,\mathbf{x}{=}v) \mid \top \mid \bot \mid \neg\varphi \mid \varphi \vee \varphi \mid [\alpha]\varphi$$

where a, b ranges over At, $\mathbf{x}{=}v$ ranges over $Inst$. The set of all formulas φ is noted Fml, and the set of all events α is noted Evt.

The set of events Evt includes variable assignments, sequences of events (;) and undetermistic choice between events ($\cup$). The set of formulas Fml includes atomic formulas $\mathbf{x}{=}v$ that has to be read "the value of variable x is v", and also special constructions $a \rightsquigarrow b$ and $\mathrm{Pre}(a,\mathbf{x}{=}v)$. $a \rightsquigarrow b$ means that "the variable assigment a triggers the variable assignment b", that is, "if the variable assignment a occurs then the variable assignment b will also occur". For instance, $(\mathbf{x}{\leftarrow}v) \rightsquigarrow (\mathbf{y}{\leftarrow}u)$ means that if the value of variable $\mathbf{x}$ is set to v then the value of variable $\mathbf{y}$ will be set to u. $\mathrm{Pre}(a,\mathbf{x}{=}v)$ means that "$\mathbf{x}{=}v$ is a necessary precondition for the occurrence of the variable assignment a", that is, "a can occur (is allowed) only if the the value of variable $\mathbf{x}$ is v". For instance, $\mathrm{Pre}(\mathbf{x}{\leftarrow}v,\mathbf{y}{=}u)$ means that the value of $\mathbf{x}$ can be set to v only if the value of $\mathbf{y}$ is u.

Here a couple of examples that better illustrate the meaning of these formal construnctions.

Example 1 *The formula* $(\mathit{UrbanCenterSpeed}{\leftarrow}\mathit{80}) \rightsquigarrow (\mathit{SpeedLimit}{\leftarrow}\mathit{exceded})$ *means that if a car runs at 80 km/h in a urban center then the driver will excede the speed limit and will make a traffic violation.*

Example 2 *The formula* $\mathrm{Pre}(\mathit{Speed}{\leftarrow}\mathit{130},\mathit{Zone}{=}\mathit{highway})$ *means that a driver can run at 130 km/h only if he is driving in a highway.*

Finally, $[\alpha]\varphi$ are dynamic operators describing the effect of a given event α. $[\alpha]\varphi$ means "after the occurrence of the event α, φ holds".

1.2 Scences and DLI models

The following are the formal definitions of a scene and of a DLI model that are used to interpret the formulas given in Section 1.

Definition 1 *A scene is a tuple* $S = (\mathcal{D}, \mathcal{P})$ *where:*

- $\mathcal{D} \subseteq At \times At$ *is a dependence relation between variable assignments such that:*
  - *if* $(\mathbf{x}{\leftarrow}v, \mathbf{y}{\leftarrow}u) \in \mathcal{D}$ *and* $(\mathbf{x}{\leftarrow}v, \mathbf{y}{\leftarrow}t) \in \mathcal{D}$ *then* $u = t$,
  - $(\mathbf{x}{\leftarrow}v, \mathbf{x}{\leftarrow}v) \in \mathcal{D}$;
- $\mathcal{P} : At \longrightarrow 2^{Inst}$ *is a precondition function for variable assignments.*

The relation $\mathcal{D}$ specifies the dependence between assignments. $(\mathbf{x}{\leftarrow}v, \mathbf{y}{\leftarrow}u) \in \mathcal{D}$ means that the variable assignment $\mathbf{y}{\leftarrow}u$ depends on the variable assignment $\mathbf{x}{\leftarrow}v$, that is, if the value of variable $\mathbf{x}$ is set to v then the value of variable $\mathbf{y}$ will be set to u. The function $\mathcal{P}$ specifies the preconditions of a given variable assignment. $\mathbf{x}{=}v \in \mathcal{P}(a)$ means that the variable assignment a can occur (is allowed) only if $\mathbf{x}{=}v$ holds.

Definition 2 *A model of an institution (DLI model) is a tuple* $M = (S, (\pi_{\mathbf{x}} | \mathbf{x} \in \mathbb{V}))$ *where:*

- S *is a scene;*
- *for every* $\mathbf{x} \in \mathbb{V}$, $\pi_{\mathbf{x}} \in Val_{\mathbf{x}}$.

1.3 Truth conditions

Given a DLI model, the events $\mathbf{x}{\leftarrow}v$ are going to be interpreted as valuation modifiers: they change the value of a given variable. In particular, $\mathbf{x}{\leftarrow}v$ sets the value of the variable $\mathbf{x}$ to v.

The truth conditions are the usual ones for $\top$, $\bot$, negation and disjunction, plus:

$$\begin{array}{lll} M \models \mathbf{x}{=}v & \text{iff} & \pi_{\mathbf{x}} = v \\ M \models a \rightsquigarrow b & \text{iff} & (a, b) \in \mathcal{D} \\ M \models \mathrm{Pre}(a, \mathbf{x}{=}v) & \text{iff} & \mathbf{x}{=}v \in \mathcal{P}(a) \\ M \models [\mathbf{x}{\leftarrow}v]\varphi & \text{iff} & M^{\mathbf{x}\leftarrow v} \models \varphi \\ M \models [\alpha; \beta]\varphi & \text{iff} & (M^{\alpha})^{\beta} \models \varphi \\ M \models [\alpha \cup \beta]\varphi & \text{iff} & M^{\alpha} \models \varphi \text{ or } M^{\beta} \models \varphi \end{array}$$

where the model $M^{\mathbf{x}\leftarrow v} = (\mathcal{D}^{\mathbf{x}\leftarrow v}, \mathcal{P}^{\mathbf{x}\leftarrow v}, (\pi_{\mathbf{x}}^{\mathbf{x}\leftarrow v} | \mathbf{x} \in \mathbb{V}))$ is defined as follows.

$$\begin{array}{lll} \mathcal{D}^{\mathbf{x}\leftarrow v} & = & \mathcal{D} \\ \mathcal{P}^{\mathbf{x}\leftarrow v} & = & \mathcal{P} \\ \pi_{\mathbf{y}}^{\mathbf{x}\leftarrow v} & = & \begin{cases} \pi_{\mathbf{y}} & \text{if there is no } u \in Val_{\mathbf{y}} \text{ such that } (\mathbf{x}{\leftarrow}v, \mathbf{y}{\leftarrow}u) \in \mathcal{D} \\ u & \text{if } (\mathbf{x}{\leftarrow}v, \mathbf{y}{\leftarrow}u) \in \mathcal{D} \end{cases} \end{array}$$

Validity and satisfiability are defined as usual.

References

[1] J. L. Arcos, M. Esteva, P. Noriega, J. A. Rodriguez-Aguilar, and C. Sierra. Engineering open evironments with electronic institutions. *Engineering applications of artificial intelligence*, 18(2):191–204, 2005.

[2] J. van Benthem, J. van Eijck, and B. Kooi. Logics of communication and change. *Information and Computation*, 204(204):1620–1662, 2006.

[3] H. P. van Ditmarsch, W. van der Hoek, and B. Kooi. Dynamic epistemic logic with assignment. In *Proceedings of Fourth International Joint Conference on Autonomous Agents and Mult-Agent Systems (AAMAS'05)*, pages 141–148. ACM Press, 2005.

[4] E. Lorini and D. Longin. A logical account of institutions: from acceptances to norms via legislators. In G. Brewka and J. Lang, editors, *Proceedings of the Eleventh International Conference on Principles on Principles of Knowledge Representation and Reasoning (KR 2008)*, pages 38–48. AAAI Press, 2008.

[5] E. Lorini, D. Longin, B. Gaudou, and A. Herzig. The logic of acceptance: grounding institutions on agents' attitudes. *Journal of Logic and Computation*, 19(6):901–940, 2009.