A Computational Framework for Non-Monotonic Agency, Institutionalised Power and Multi-Agent Systems

Guido Governatori* Antonino Rotolo†
* *School of Information Technology and Electrical Engineering*
The University of Queensland, Brisbane, Australia
guido@itee.uq.edu.au
† *CIRSFID, University of Bologna, Bologna, Italy*
rotolo@cirfid.unibo.it

In [3, 4] the present authors and colleagues have provided a model of organisations of agents in terms of policy-based normative systems. An organisation is thus characterised by specifying the normative positions relevant to design its structure. This abstract summarises further results in this direction[1]. In particular, the research provides a computational framework based on Defeasible Logic to capture some theoretical intuitions developed in [3, 4]. The focus is on the idea of institutionalised power, as represented by the counts-as link, and on two aspects of agency: the notions of attempt and of personal and direct action to realise states of affairs. For the first issue, we adopted the approach of [3, 4], inspired, with some differences, by [5, 6]. For the second aspect, the reference is to recent developments of the Kanger-Lindahl-Pörn theory of action [7, 8, 6]. The model embeds also the modal operator *proc*, which represents the act of proclaiming to capture some minimal properties of all speech acts that are intended to modify the institutional world [3, 4]. The goal of the research is computational insofar as it proposes an efficient and flexible formalism to deal with counts-as conditionals and their interplay with agency. In fact, [4, 3], but also [5, 6], can hardly be used directly for implementation since they are based on conditional logics. On the other hand, this research aims also to provide some hints to define a non-monotonic theory of agency.

Our logic of institutional agency is based on the framework for Defeasible Logic developed in [1, 2]. The idea is to extend this framework to account for: (1) the underlying logical structure of any propositional base describing an institution; (2) the logic behavior of the modal operators of institutional agency; and (3) the relationships among such operators. The logic of agency should specify how modalities can be introduced and manipulated. To comply with this, in the setting provided by Defeasible Logic we have to set 1) the rules describing the logical inferences within the institution and 2) the rules to introduce the modal operators of agency E_i (*the agent i brings it about*), and H_i (*the agent i attempts*). Accordingly, two types of rules (strict, defeasible, and defeaters) are considered: rules for the notion

*This research was partially supported by the University of Queensland under the grant UQRSF ITEE-03/2002001335.

[1]Motivation and details of the research are provided in the full version of this paper, which is available at `http://www.itee.uq.edu.au/~guido/Papers/deontic/Defeasible-jurix03.pdf`.

Guido Governatori and Antonino Rotolo, 'A Computational Framework for Non-Monotonic Agency, Institutionalised Power and Multi-Agent Systems' in D. Bourcier (ed.), *Legal Knowledge and Information Systems. Jurix 2003: The Sixteenth Annual Conference*. Amsterdam: IOS Press, 2003, pp. 151-152.

of *counts-as* and rules for the notion of *results-in*. The language of Defeasible Logic is extended with a set of action symbols; $\alpha_i, \beta_i, \gamma_i$, which denote atomic actions, may occur in the rules. The meaning of an action symbol α_i is that the action corresponding to it has been performed by the agent i. In this perspective, a defeasible institutional action theory is a structure $I = (A, F, R^c, \{R^i\}_{i \in A}, >)$ where, A is a finite set of agents, F is a set of facts, R^c is a set of counts-as rules (i.e., $\rightarrow_c$, $\Rightarrow_c$, $\leadsto_c$), $\{R^i\}_{i \in A}$ is a family of sets of results-in rules (i.e., $\rightarrow^i$, $\Rightarrow^i$, $\leadsto^i$ for each agent $i \in A$), and $>$, the superiority relation, is a binary relation over the set of rules (i.e., $> \subseteq (R^c \cup R^A)^2$), where $R^A = \bigcup_{i \in A} R^i$. The intuition is that, given an institution I, F consists of the description of the raw institutional facts, either in form of states of affairs and actions that have been performed. R^c describes the basic inference mechanism internal to an institution, while R^A encodes the transitions from state to state occurring as the results of actions performed by agents within the organisation. The crucial notion in this model is that of derivability. Counts-as rules and conditions for derivation involving them produce the institutional facts that hold within the institution. However, the machinery requires also to account for action transitions that are recognised by the organisation. Accordingly, the introduction of the modalities of agency within the institution is obtained as the output of the general definition of derivability that involves counts-as as well as results-in rules. Suppose the agent i is acting in the context of an auction[2]:

$$r_1: \ \mathbf{raises\_hand}_i, \ auction\_begun \Rightarrow_c \mathbf{bids}_i \qquad\qquad r_2: \ \mathbf{bids}_i \Rightarrow^i offer$$

If no attack is possible, i's fulfilment of the conditions of r_1 produces *offer*: i's action of raising one hand has the result that i has made an offer, and so the derivation of *offer* permits to introduce $E_i(offer)$. But the framework can deal with more complex scenarios where, e.g., i acts on behalf of the agent j. In this perspective, we may have rules such as the following:

$$r_3: \ \mathbf{bids}_i, \ proc_i(E_j offer) \Rightarrow^j offer \qquad\qquad r_4: \ proc_i(E_j offer), \mathbf{raises\_hand}_i \Rightarrow_c \mathbf{bids}_j$$

Rule r_3 means that the fact that i makes a bid and proclaims that j makes the offer may permit to introduce $E_j offer$, while r_4 expresses that i's proclamation that j makes an offer counts as j's action of bidding.

References

[1] G. Antoniou, D. Billington, G. Governatori, and M. J. Maher. A Flexible Framework for Defeasible Logics. In *AAAI-2000*, pages 401–405. AAAI Press, Menlo Park, 2000.

[2] G. Antoniou, D. Billington, G. Governatori, and M. J. Maher. Representation Results for Defeasible Logic. *ACM Transactions on Computational Logic*, 2(2):255–287, 2001.

[3] J. Gelati, G. Governatori, A. Rotolo, and G. Sartor. Declarative Power, Representation, and Mandate: A Formal Analysis. In *Proceedings of Jurix'02*, pages 41–52. IOS Press, Amsterdam, 2002.

[4] G. Governatori, J. Gelati, A. Rotolo, and G. Sartor. Actions, Institutions, Powers. Preliminary Notes. In *Proceedings of RASTA'02*, pages 131–147. University of Hamburg, Hamburg, 2002.

[5] A. Jones and M. Sergot. A Formal Characterisation of Institutionalised Power. *Journal of IGPL*, 3:427–443, 1996.

[6] A. Jones. A Logical Framework. In J. Pitt, editor, *Open Agent Societies: Normative Specifications in Multi-Agent Systems*. Wiley, Chichester, 2003.

[7] F. Santos and J. Carmo. Indirect Action, Influence and Responsibility. In M. Brown and J. Carmo, editors, *Deontic Logic, Agency and Normative Systems*. Springer, Berlin, 1996.

[8] F. Santos, A. Jones, and J. Carmo. Action Concepts for Describing Organised Interaction. In *13th HICSS*. IEEE Computer Society Press, Los Alamitos, 1997.

[2] Bold type expressions correspond to action symbols, the italicised ones to state of affairs.