Complex Loop of Norm Emergence

Giulia Andrighetto
Laboratory of Agent Based Social Simulation,
Institute of Cognitive Science and Technology,
CNR, Rome, Italy

Outline

- State of the art
- A social cognitive view of norms
- Normative MAgent architecture:
  - EMIL-A
- Simulation model and results
- Conclusions

State of the art

Conventionalistic view

- How do norms emerge, spread and change over time?
- Norms as conventions
- Norms emerge from interaction among agents, driven by non normative internal mechanisms

- Ullman-Margalit, E. The Emergence of Norms. Clarendon Press, Oxford, 1977.
- Axelrod, R, An evolutionary approach to norms. The American Political Science Review, 4(80):1095-1111, 1986.
- Bicchieri, C. (2006). The Grammar of Society: the Nature and Dynamics of Social Norms, Cambridge University Press.
- Young, H. P. (1998). Individual Strategy and Social Structure: An Evolutionary Theory of Institutions. Princeton University Press.
- Sen, S. and Airiau, S. (2007) Emergence of norms through social learning. In Proceedings of the Twentieth International Joint Conference on Artificial Intelligence.

Imperativistic view

- Why do agents comply with norms and how is it possible that norms operate upon autonomous intelligent agents?
- Cognitively rich agents*:* agents decide upon norms
- Investigation of the effects of norms, i.e. a *functional analysis.*

- Conte, R. and Castelfranchi, C. (1995) Cognitive and social action. University College of London Press, London.
- Jones, A. and M. Sergot. (1996) A formal characterization of institutionalized power. Logic Journal of the IGPL 4(3):429-445
- Dignum, F. (1999) Autonomous agents with norms. Artificial Intelligence and Law, 7(1):69-79.
- Shoham, Y. and M. Tennenholtz (1992) On the synthesis of useful social laws in artificial societies. In Proceedings of the 10th National Conference on Artificial Intel ligence, number 276-282. Kaufmann: San Mateo, California.
- Van der Torre, L. and Tan, Y. (1999) Contrary-to-duty reasoning with preference-based dyadic obligations. Annals of Mathematics and Artificial Intel ligence, pages 1239-1246.

Two Current Views

- Conventions (mainly bottom-up)
- Legal norms (mainly top-down)
- Open questions
  - As to conventions:
    - What about *mandatory* social norms?
  - As to legal norms
    - How do they evolve?
  - As to both
    - What about a unifying view?

Main lessons from state of the art

Two traditions

- Either norm emergence
  - As epiphenomena
  - Or non-normative internal mechanisms
    - Moral dispositions
    - Social learning (imitation)
- Or norm-based reasoning and decisions

No integration!

Norms

A ***norm*** "is a *prescribed* guide for conduct which is generally complied with by the members of society" (Ullman-Margalit, 1977).

"first-come, first-served norm"

For a norm-based behavior to take place:

- N-beliefs: beliefs that a given behaviour, in a given context, for a given set of agents, is either forbidden, obligatory, or permitted (Kelsen 1979, von Wright 1963 Conte and Castelfranchi 1999; 2006).
- N-goals: goals to (not) achieve/accomplish obligatory/forbidden/permitted actions (Conte 2009).

- von Wright, G. H. (1963). Norm and Action. A Logical Inquiry. Routledge and Kegan Paul, London.
- Conte, R. and Castelfranchi, C. (1999). From conventions to prescriptions. towards a unified theory of norms. AI and Law, 7:323–340.
- Conte, R. and Castelfranchi, C. (2006). The mental path of norms. Ratio Juris, 19(4):501- 517.
- Kelsen, H. (1979). General Theory of Norms. Hardcover.
- Conte, R: (2009). Rational, goal governed agents. Encyclopedia of Complexity systems, Science, Springer
- Ullman-Margalit, E. (1977) The Emergence of Norms. Clarendon Press, Oxford.

Complex loop of Norm emergence

Gradual and complex dynamics by which the macro-social effect, in our case a specific norm,

- emerges in the society *while*
- immerging (Castelfranchi, 1998; Conte et al., 2007) in the minds of the agents producing it,generating a number of intermediate loops.

Local loops of Norm Emergence

- *Local loops*
  - partial or initial observable macroscopic effects of local behaviours occur;
  - retroact on (a subset of) the observers' minds, modifying them (producing new internal states, emotions, normative goals, normative beliefs, etc.);
  - agents start to behave accordingly with their mental states;
  - agents communicate internal states to one another, thus activating a process of normative influencing (see Conte and Dignum, 2001);
  - these normative beliefs spread through agents' minds;
  - behaviours progressively conform to spreading states;
  - initial macroscopic effects get reinforced/weakened depending on the type of mental states spreading.

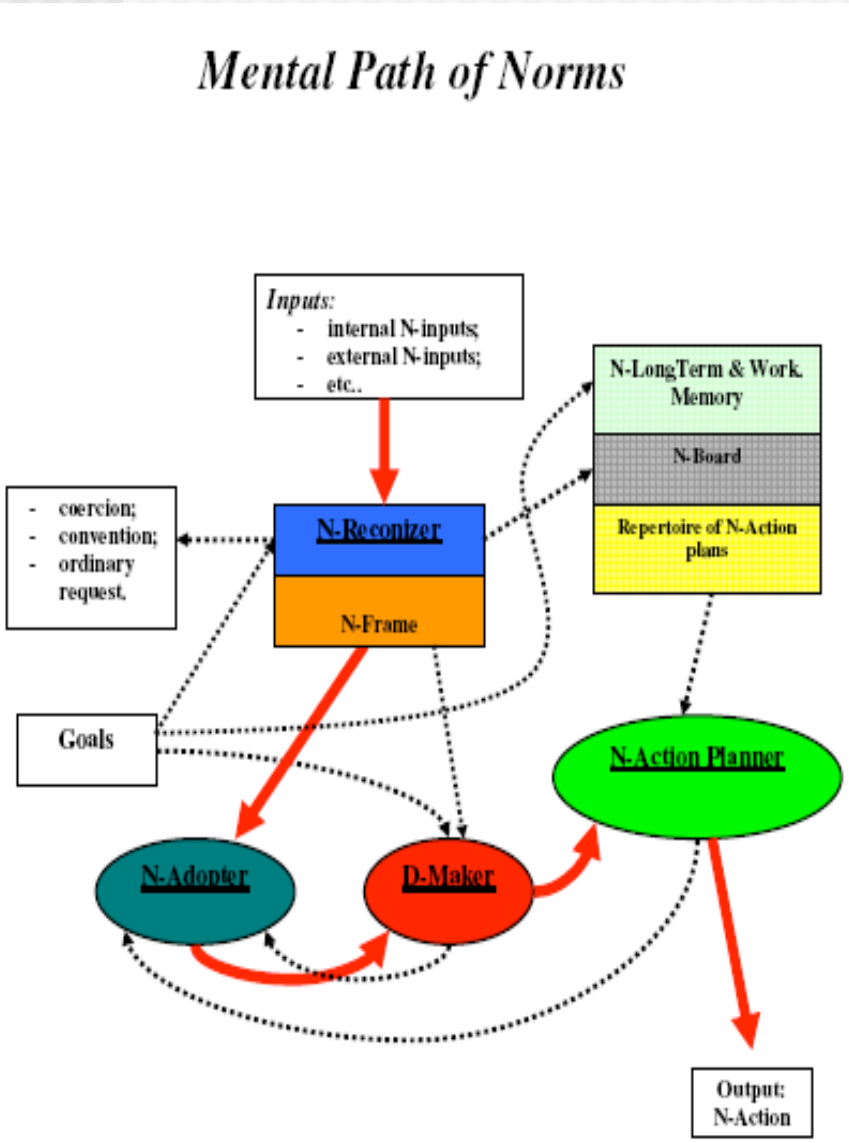
Conte R. and Dignum F. (2001). From social monitoring to normative influence. Jasss - the Journal of Artificial Societies and Social Simulation, 4(2). http://jasss.soc.surrey.ac.uk/4/2/7.html

Cognitive mediators of norms

Social phenomena are due to the agents' behaviors, but… the agents' behaviors are due the the *mental mechanisms* controlling and (re)producing them.

- How the norm should ***work through*** the minds of the agents? How is it *represented*?
- Which are the *proximate mechanisms* underlying the normative behavior?
- What does it mean to *conform* to a norm from a mental - not just a behavioral - point of view? What does it mean to *obey*?

EMIL A



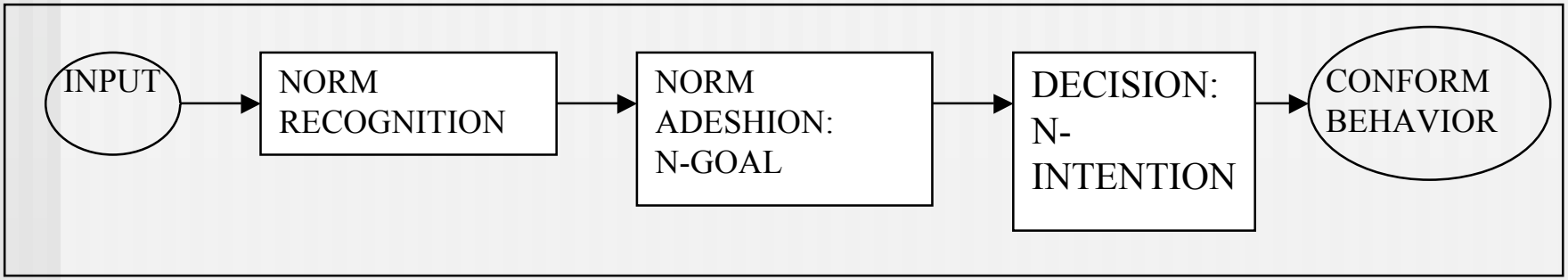
- *tick red arrows represent the standard information flow*
- *dotted black arrows represent alternative directions of the information flow.*

- is able to *recognize* N, tell what is a N and what is not and form a N-bel corresponding to N;
- is able to *assess* whether it is concerned by N;
- *accepts* N, forms a N-goal corresponding to N;
- *decides to comply* with N or not (*intention*);
- is able to *re-issue* N, to prescribe it to other fellows subject to N, and
- is able to observe, *monitor* their behaviors wrt N and react in a positive or negative way to them.

Andrighetto, G., Campennì, M, Conte, R., Paolucci, M. (2007). On the Immergence of Norms: a Normative Agent Architecture. In Proceedings of AAAI Symposium, Social and Organizational Aspects of Intelligence, Washington DC.

Conformity as routine behavior 1/2

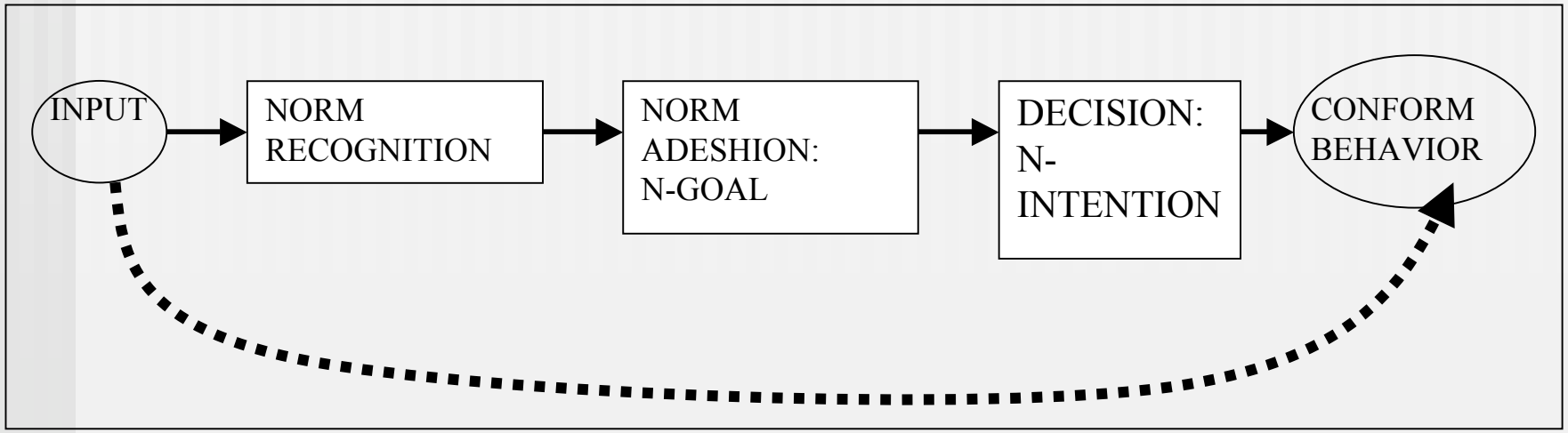
Our quite rich cognitive characterization of the representations and processes underlying a behavior obedient to a norm



…. **shouldn't** however give the idea of behavioral conformity as always based on such a complex 'reasoning' and 'deliberation'.

Conformity as routine behavior 2/2

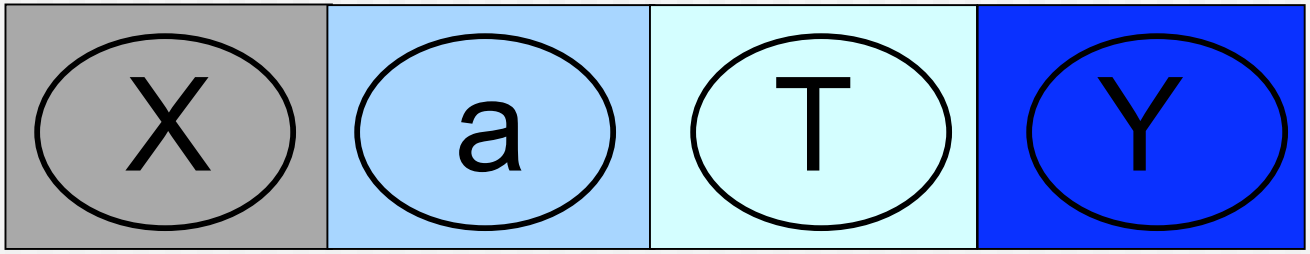
Norm conformity and obedience become a *habit*, an *automatism*, a *routine* behavior.



But before, norms must be acquired (*immergence*)

NORM RECOGNITION

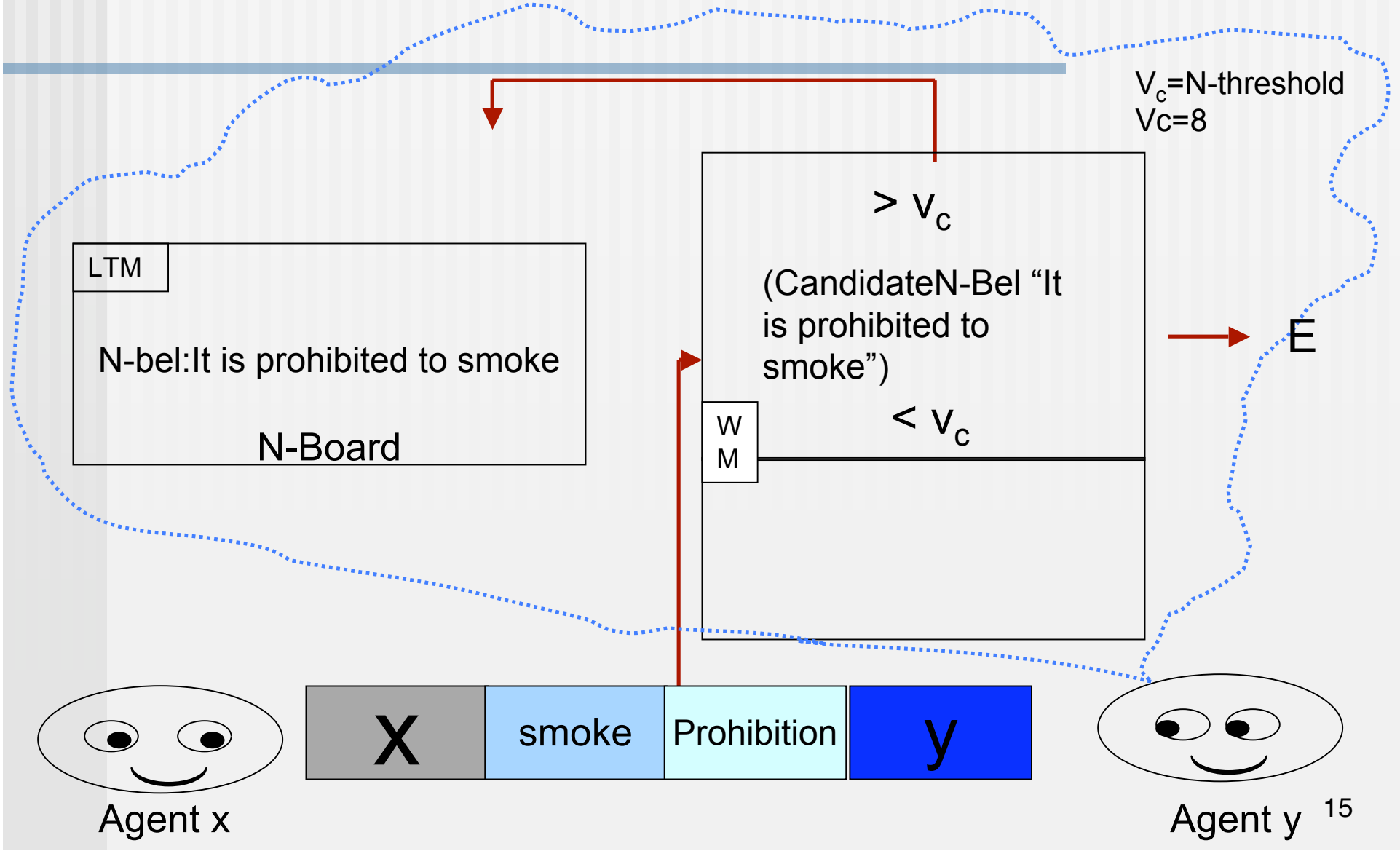
INPUT



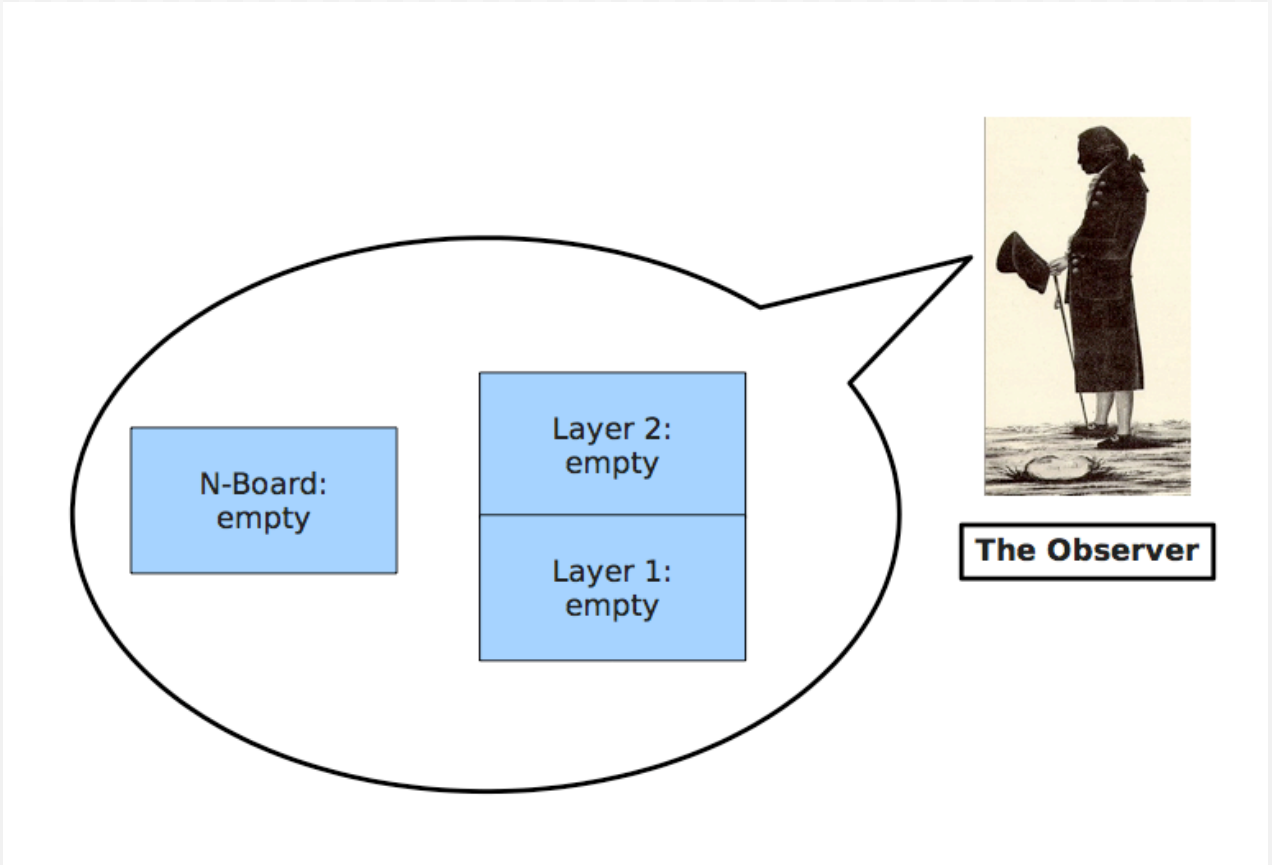
Each input is presented as an ordered vector

- Source (x);
- Action transmitted (a) (potential norm)
- Type of input:
  - Behaviors
  - Messages: assertions (A), behaviours (B), requests (R), deontics (D), evaluations (V), sanctions (S);
- Observer (y);

N-RECOGNITION MODULE



NORM RECOGNIZER AT WORK 1/6



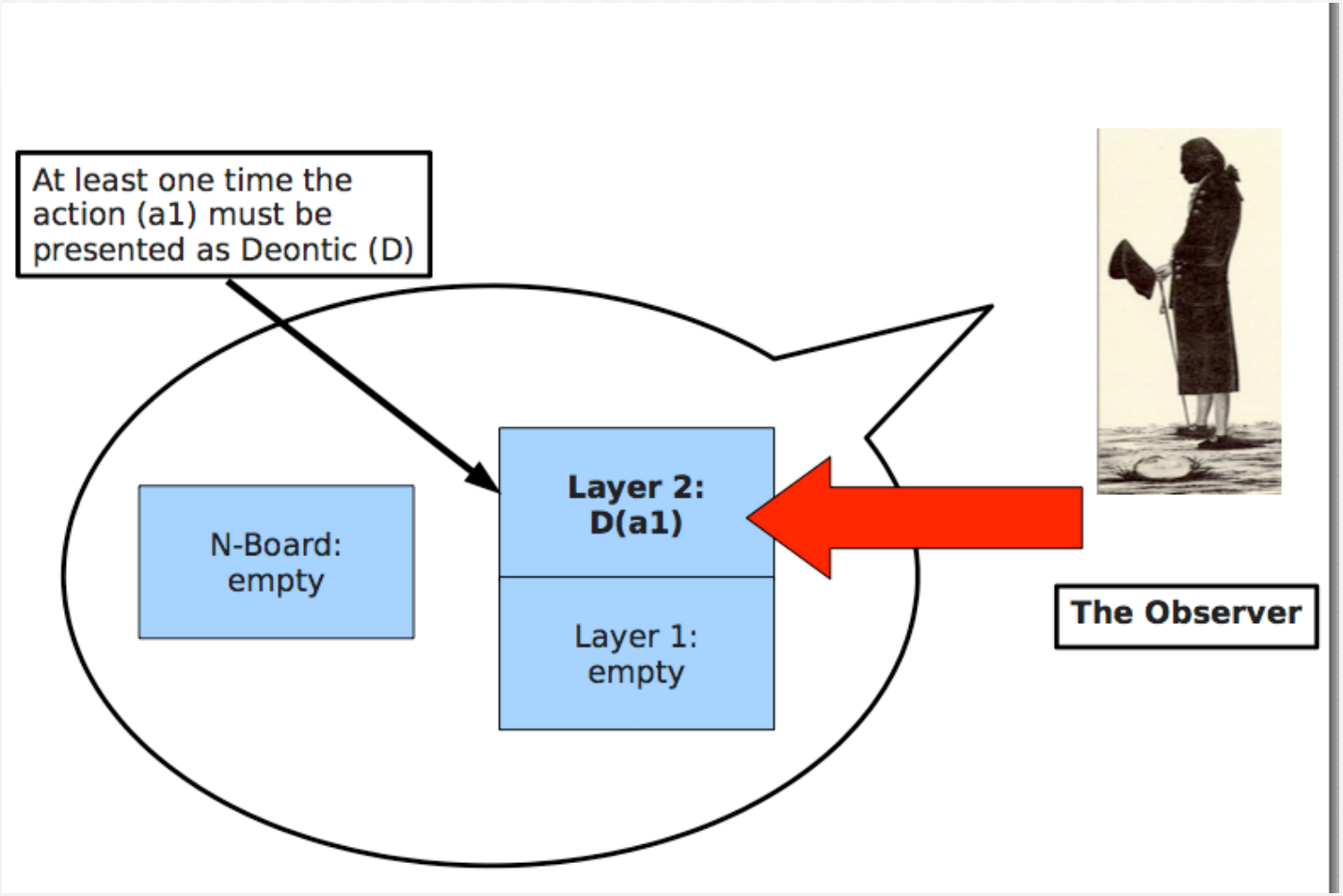
NORM RECOGNIZER AT WORK 2/6

Agent 2 | a1 | D(a1) | Agent 1

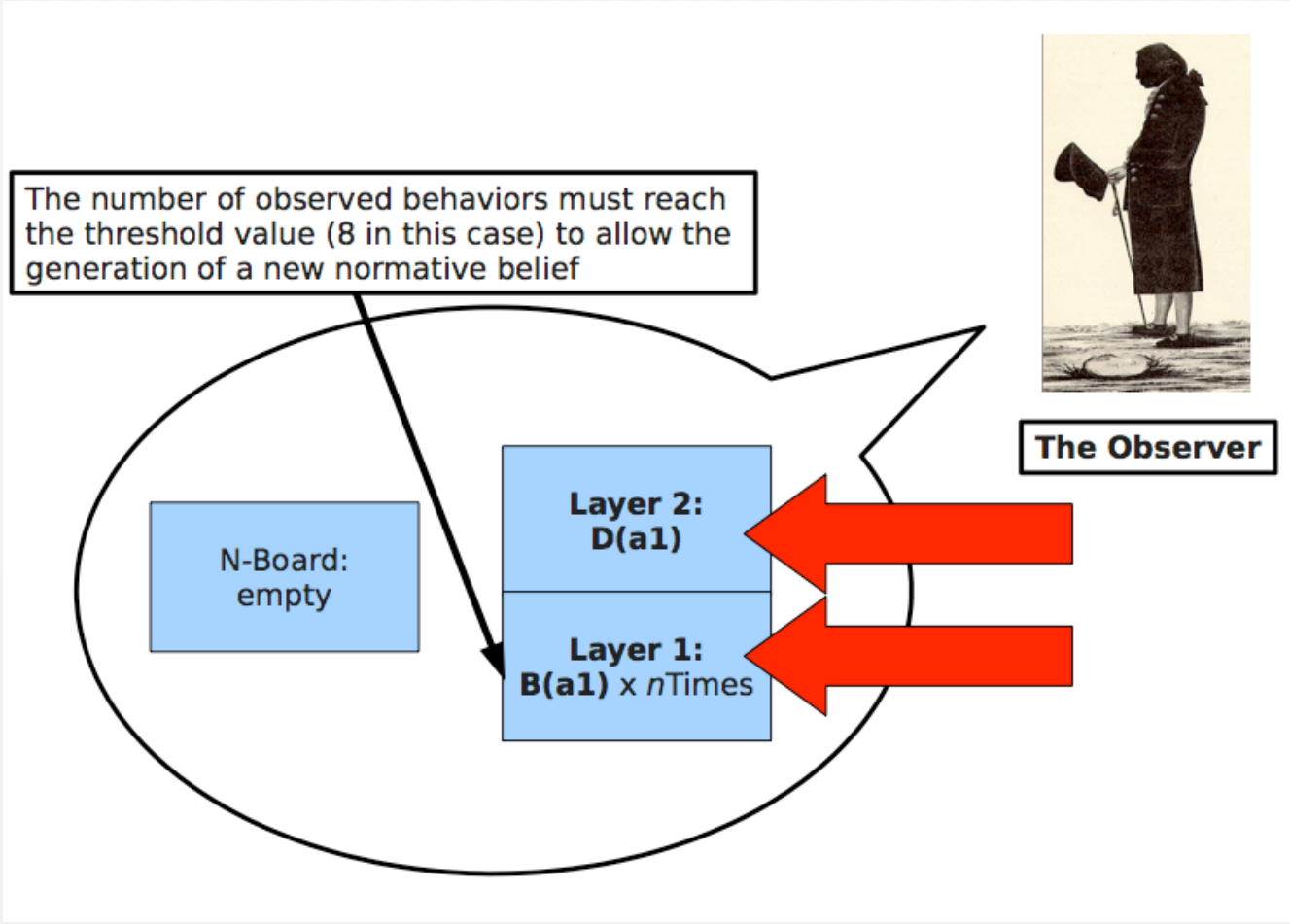
The Sender

The Observer

NORM RECOGNIZER AT WORK 3/6



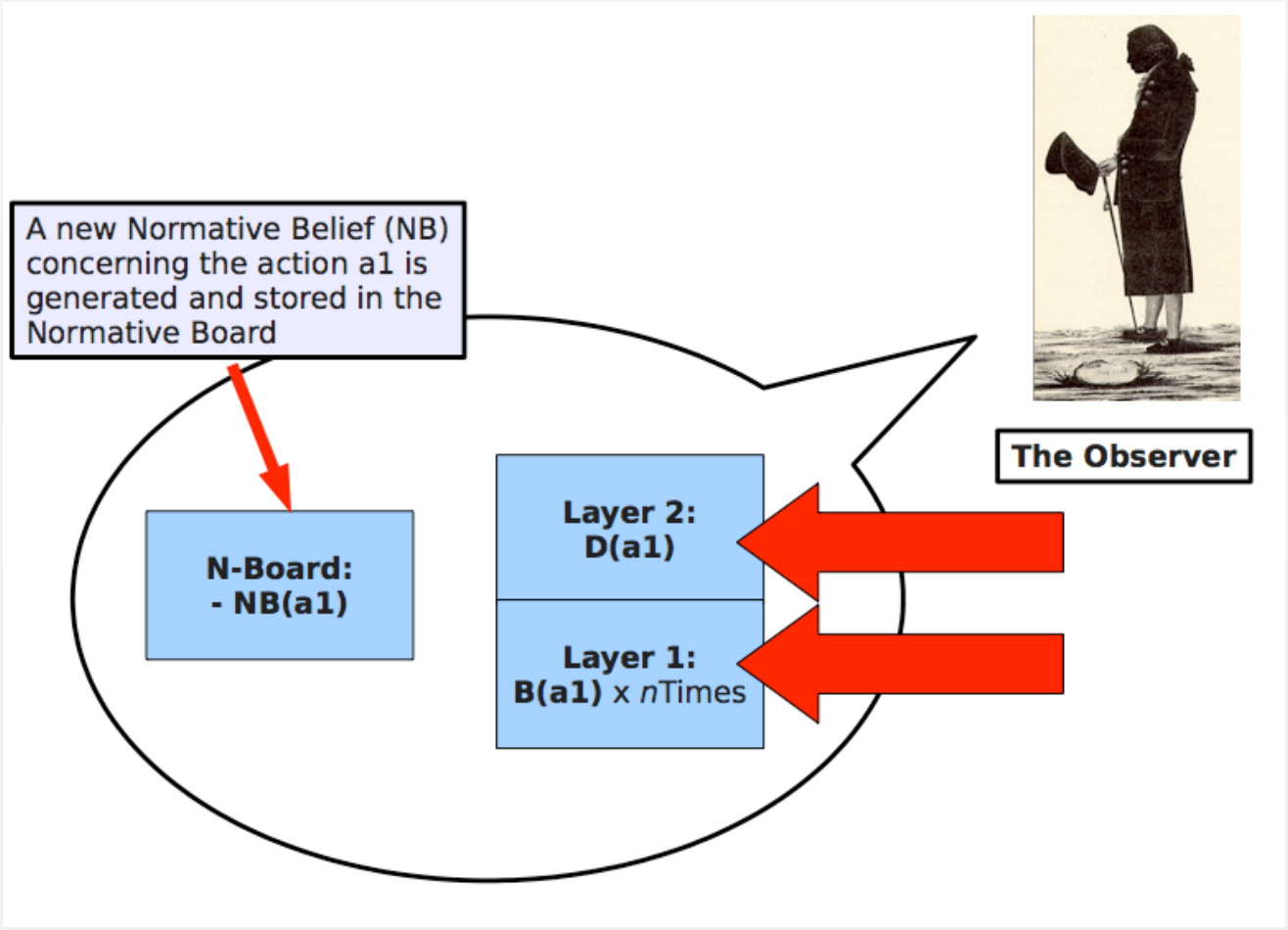
NORM RECOGNIZER AT WORK 4/6



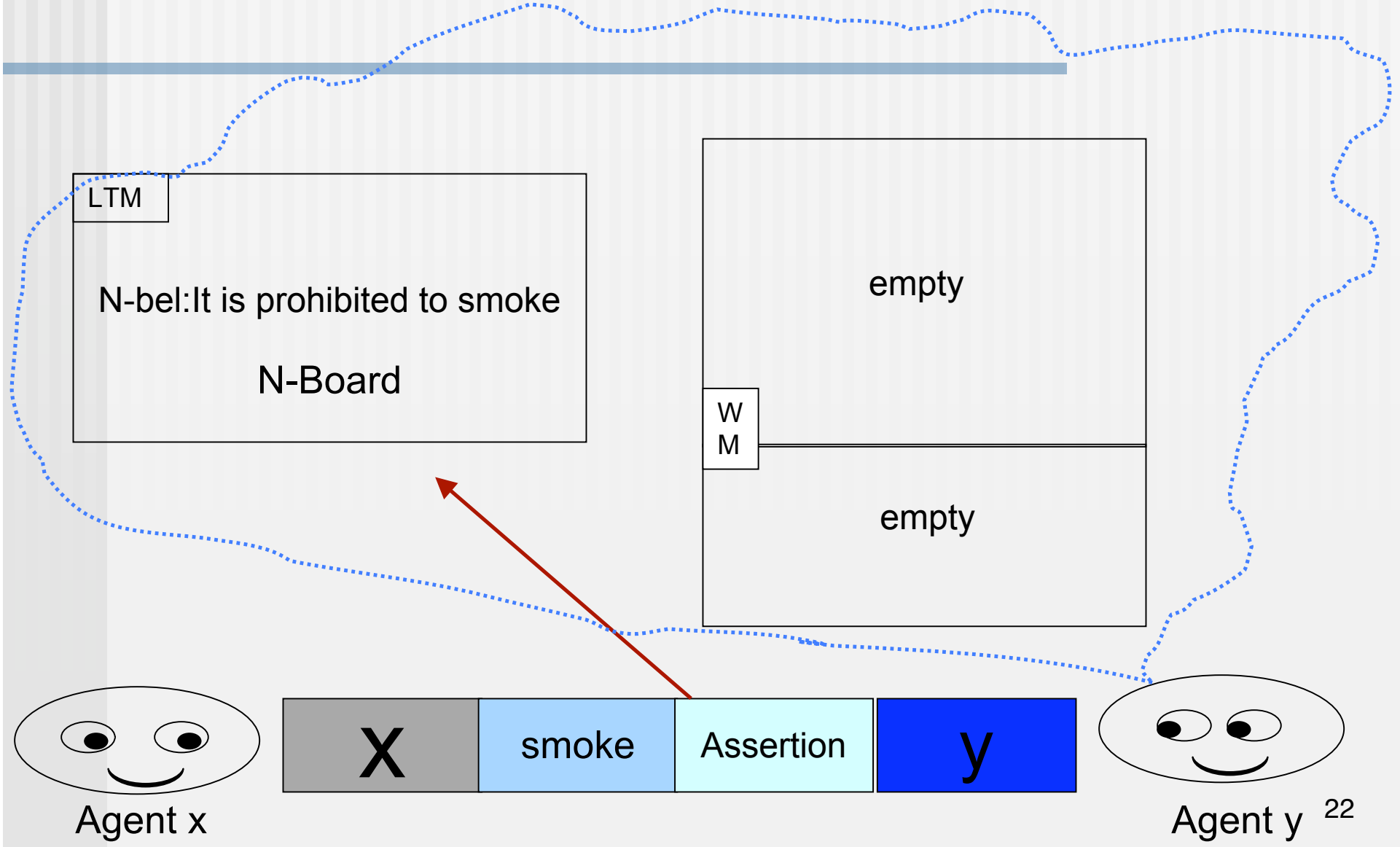
NORM RECOGNIZER AT WORK 5/6

The Observer

NORM RECOGNIZER AT WORK 6/6



N-RECOGNITION MODULE



A SIMULATION STUDY

NORM-RECOGNIZERS VS SOCIAL CONFORMERS

- What are observable effects of norm recognition?
- Implement different populations (Andrighetto et al., 2008, Campennì et al., 2008):
  - Social conformers follow actions most frequently done in observation window (parameter)
  - Norm recognizers take input from others, form beliefs and act based on those.

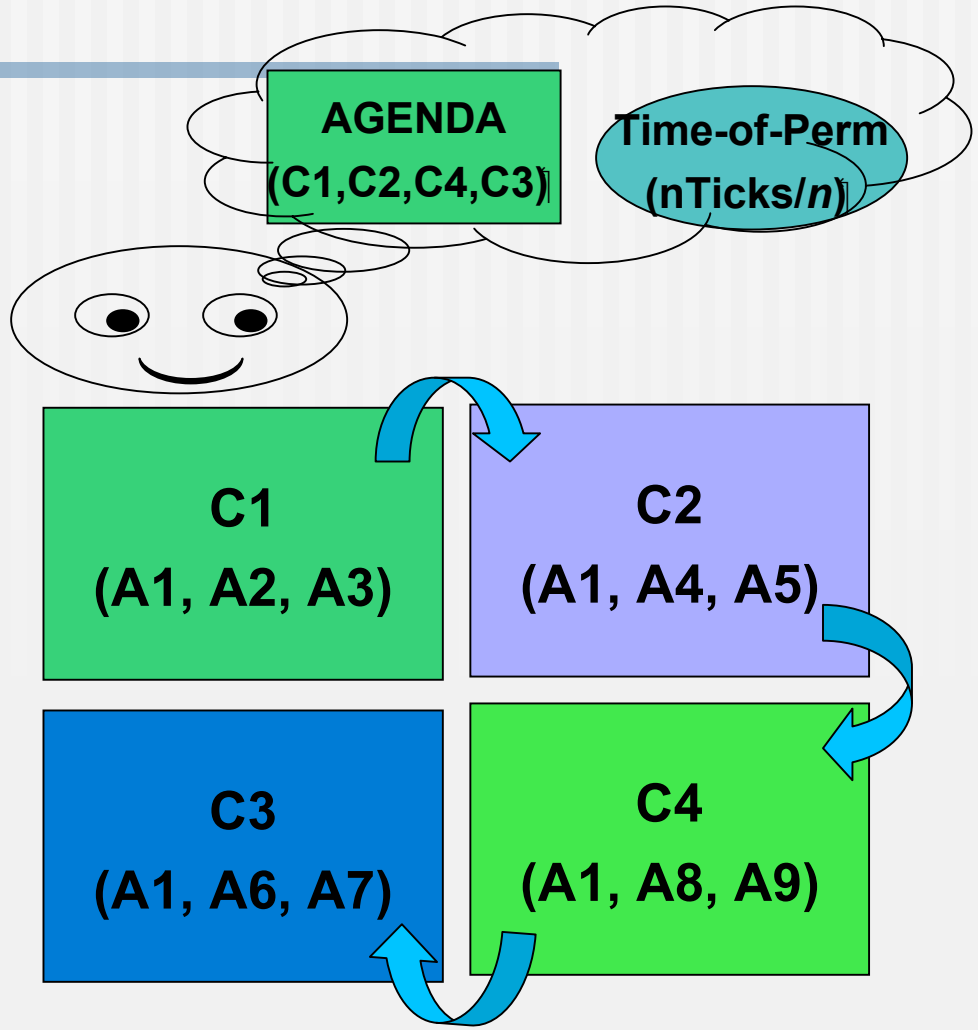
Andrighetto, G., Campennì, M., Conte, R., Cecconi, F (2008). How Agents Find out Norms: A Simulation Based Model of Norm Innovation. In 3rd International Workshop on Normative Multiagent Systems (NorMAS 2008) 15-16 July, 2008, Luxembourg.

Campennì, M., Andrighetto, G., Cecconi, F., Conte, R. (2008). Normal = Normative? The Role of Intelligent Agents in Norm Innovation. In The Fifth Conference of the European Social Brescia, September 1-5, 2008.

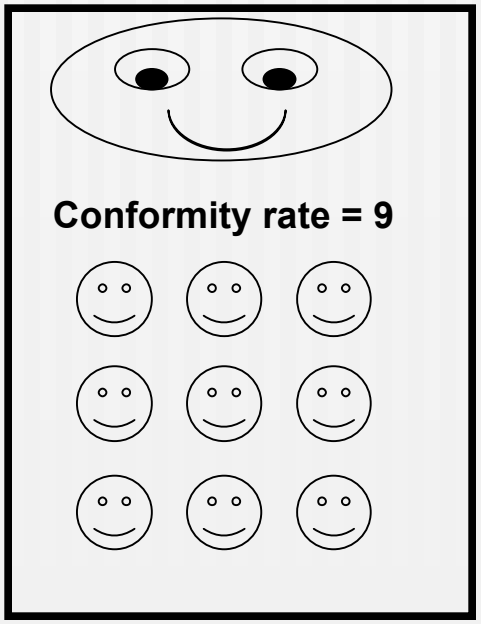
AGENT AND WORLD

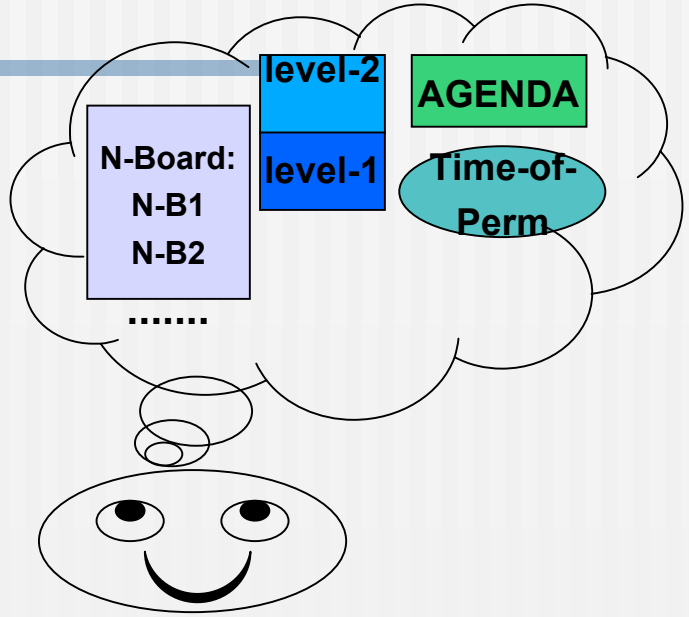
4 contexts:

- following its agenda and time of permanence, each agent moves among contexts;
- in each context, agents can produce 1 out of **3 actions;**
- 1 action is the same for all of the contexts.



Social Conformers & Norm Recognizers





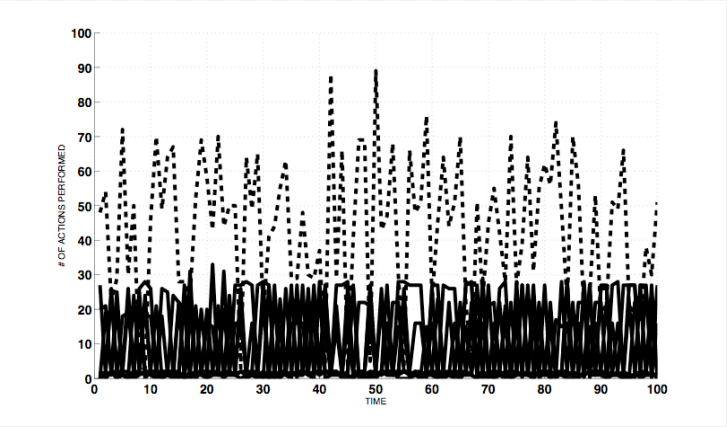
- Each observes other agents in same context
- Local rule: according to conformity rate, imitates most frequent action

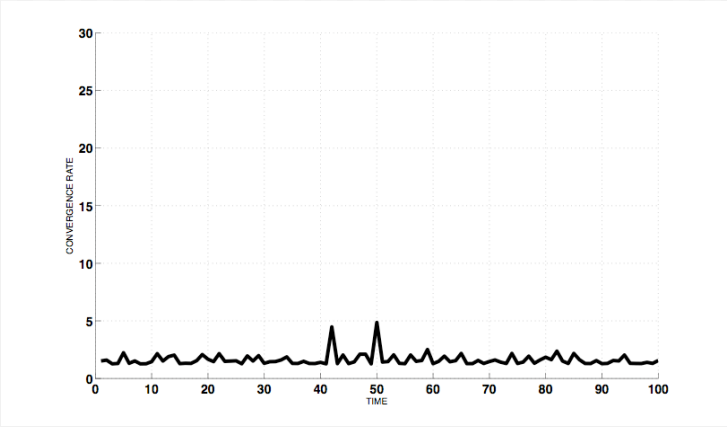
Each is provided with:

- Normative Board;
- Double-layer architecture;
- Agenda: individual time of permanence (in contexts);
- New normative beliefs contribute to choose action;

SIMULATIONS' RESULTS

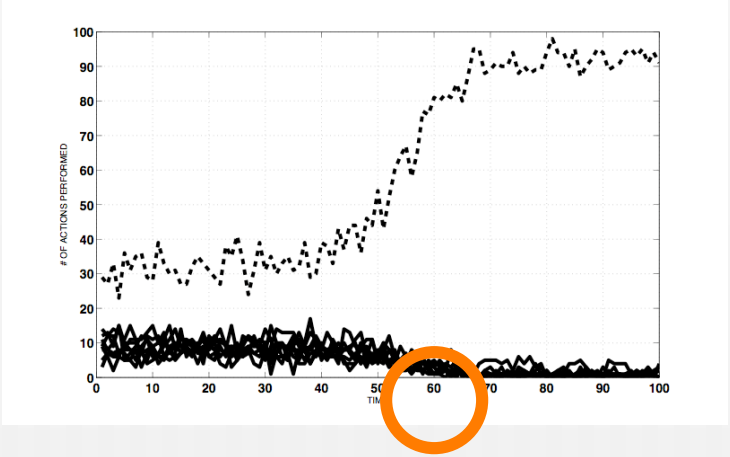
Findings: Social Conformers (Andrighetto et al. 2008; Campennì et al. 2008)

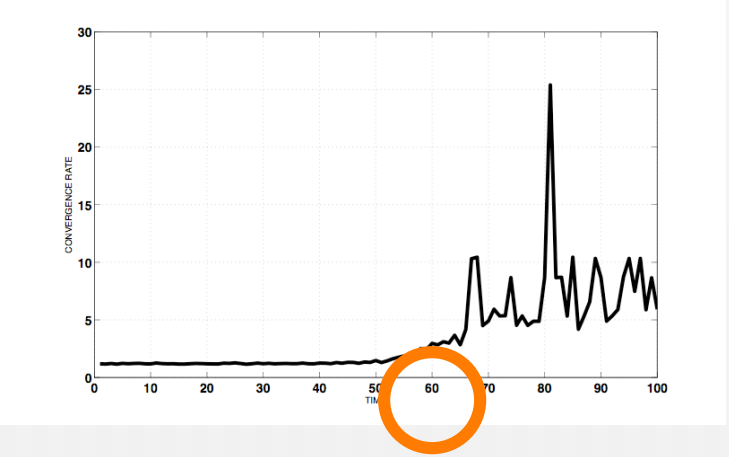




Social conformers do not converge on one action

Findings: Norm Recognizers (Andrighetto et al 2008; Campennì et al. 2008)

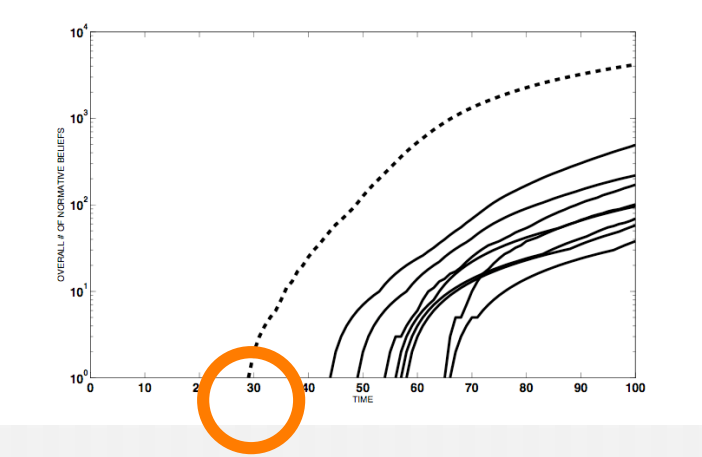




- After 60th tick, something emerges… What is it?

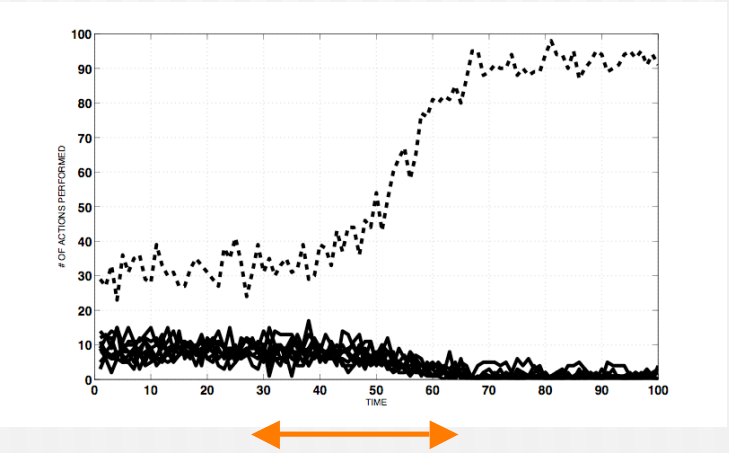
Lets look into the agents' minds…

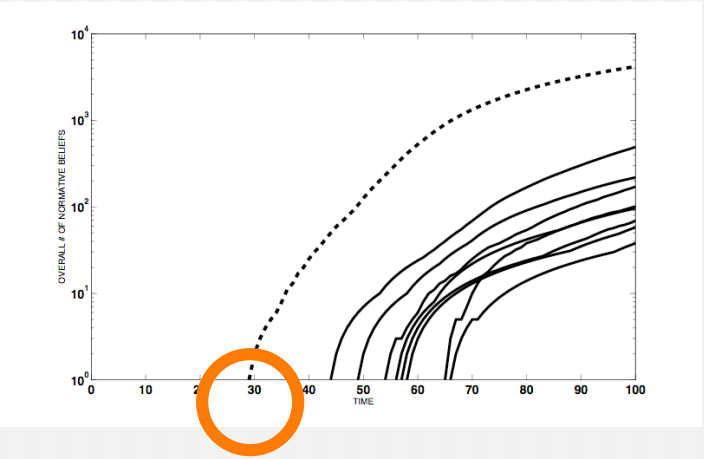
IMMERGENCE



- At 30th tick a normative belief starts to spread
- What has happened in the interval?
  - Other normative beliefs got formed, although earlier is more frequent
  - Immergence is earlier: it takes time for effect to emerge

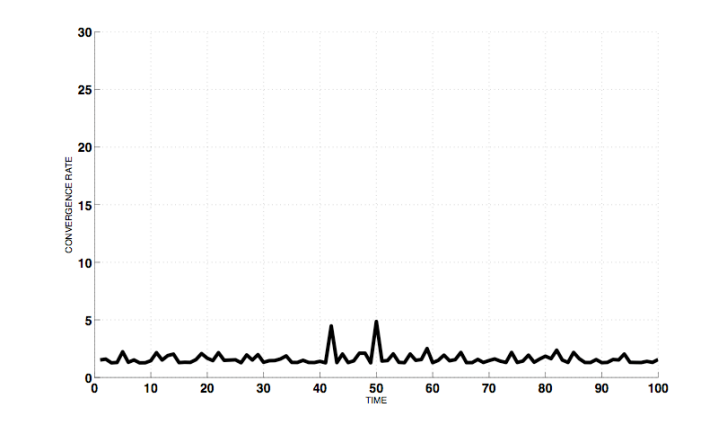
Norm latency

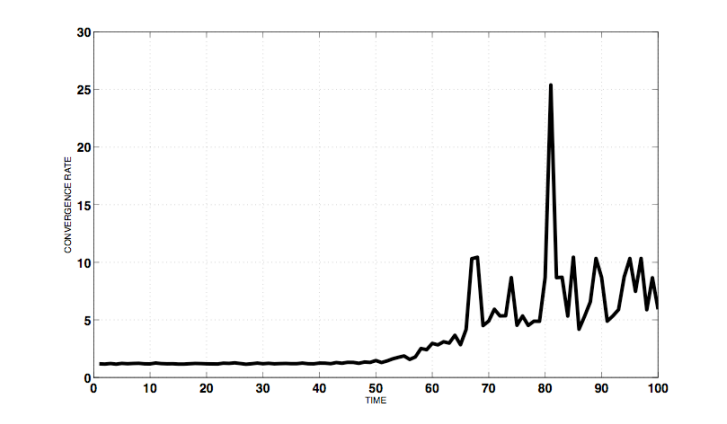




- Time interval between N-bels appearance and convergence on corresponding action.
- Actually, a complex loop
  - from N-Belx to N-actionx
  - from N-actionx to N-bely
  - from N-bely to N-actiony
  - Etc.

TO SUM UP

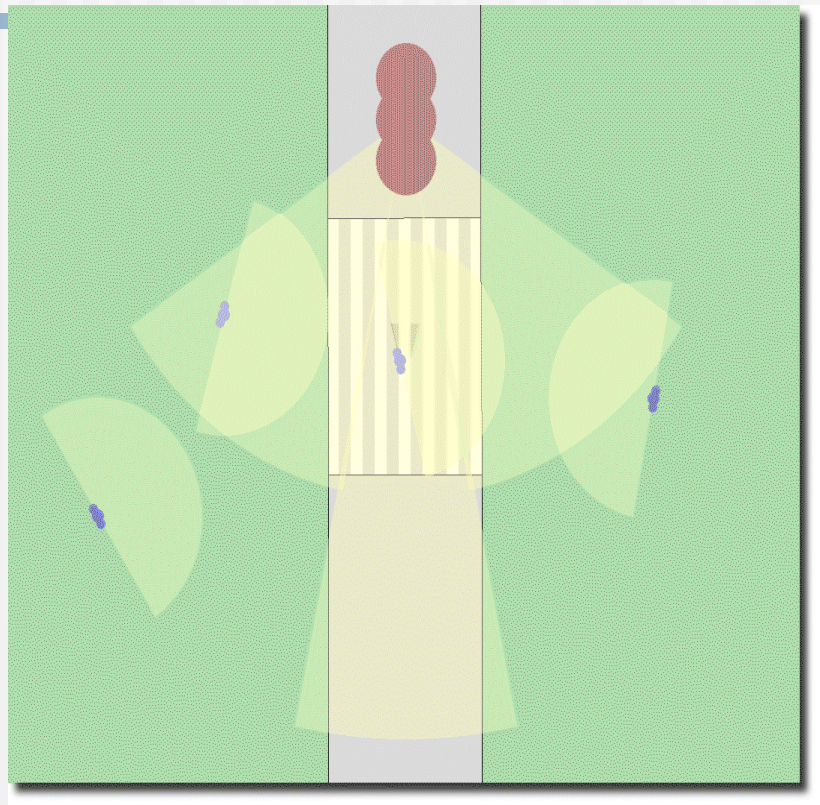




- In a multi-scenario world, unlike social conformers, agents endowed with a rich cognitive architecture and able to influence each other converge on one single norm
- Norms immerge in the minds before emerging in behavior.
- Norms have a latency time

TRAFFIC SCENARIO

- One-way road with cars moving from North to South
- Two meadows with children moving between East and West
- Car drivers and children learn how to behave reasonably in this scenario and internalise emerging norms



Lotzmann et al. 2008

Lotzmann, Ulf; Moring, Michael; Troitzsch, Klaus G. (2008): Simulating Norm Formation in a Traffic Scenario. In Proceedings"The Fifth Conference of the European Social Simulation Association, September 1-5, 2008". Brescia

- For both types of agents, the deviation from the allowed duration leads to a *penalty* when more time was required and accordingly a *gratification* when the target was reached early.
- Due to the interaction between agents, occasional *collisions* are likely to happen. Such an event, when occurring between a car and pedestrian, is classified as *undesirable*.
- Observations of a collision provoke other agents to issue *sanctions* against the blameable agents.
- The strength of the sanction is determined by various factors:
  - ✓ *environmental situation* (e.g. the road section in which the collision occurred)
  - ✓ *normative beliefs* of the valuating agent (e.g. a collision on a crosswalk might result in a harder sanction than on the rest of the road).
- Sanctions lead to a temporary stop of motion for the involved agents. To avoid sanctions is a *competing goal*: reaching the target point or end of the road, respectively, in due time).

TRAFFIC SCENARIO RUNNING

- the first (simulated) minute (20 pedestrians, random cars
- several (simulated) minutes later (again 20 pedestrians, random cars)
- the same, some pedestrians have not learnt to use the crossing

THANK YOU FOR YOUR ATTENTION

References and online simulations can be found on http://labss.istc.cnr.it/