Regulated MAS for Sociotechnical Systems

Luis Gustavo Nardin

MINES Saint-Étienne

2023 Workshop-School on Agents, Environments, and Applications

August 31, 2023

Outline

Social Regulation Evolution

Informal Regulations
(e.g., conventions, gossip, ostracism)

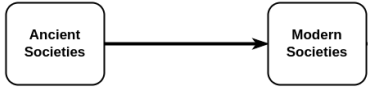
Ancient Societies

- Small groups
- Close kinship
- Long term relationships

Social Regulation Evolution

Informal Regulations
(e.g., conventions, gossip, ostracism)

Formal Regulations
(e.g., laws and legal sanctions)



- Small groups
- Close kinship
- Long term relationships

- Large groups
- Impersonal relationships
- Short term relationships

Social Regulation Evolution

Informal Regulations
(e.g., conventions, gossip, ostracism)

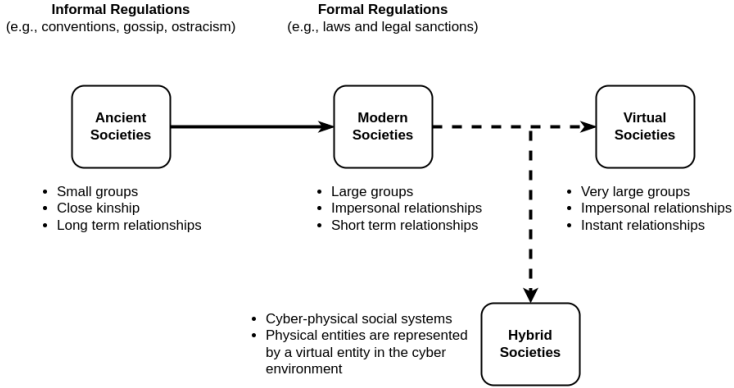
Formal Regulations
(e.g., laws and legal sanctions)

- Small groups
- Close kinship
- Long term relationships

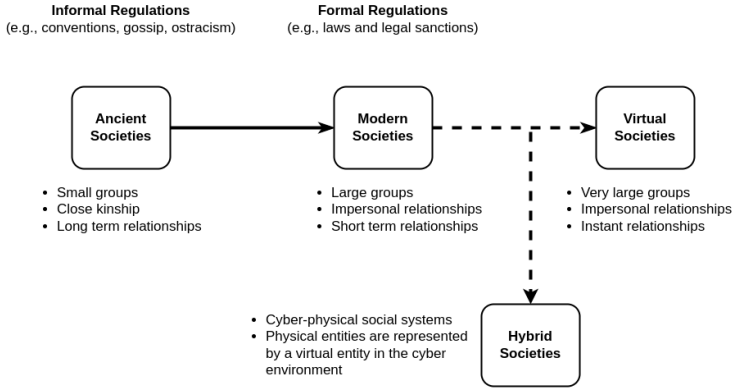
- Large groups
- Impersonal relationships
- Short term relationships

- Very large groups
- Impersonal relationships
- Instant relationships

Social Regulation Evolution



Social Regulation Evolution



How can we regulate these cyber-physical social systems interfering the least with the autonomy of their entities?

Sociotechnical System (STS)

- Human society is invariably becoming more **dependent on** technologies, in particular **information technology**, which is transforming how **human** and **autonomous artificial** agents interact
- Technologies enabling these interactions include technical systems but also operational processes and people who use and interact with these technical systems
- **Sociotechnical systems** (STS) are cyber-physical systems that incorporate **social** and **technical** aspects facilitating the **interactions** of multiple autonomous participants whose interests are at best imperfectly aligned (Sing, 2013)
  - **Publishing system** and **manufacturing industry** are STSs where technology is used in parallel with specific **social** and **organizational strategies** to control production

Sociotechnical System (STS)

- Human society is invariably becoming more **dependent on** technologies, in particular **information technology**, which is transforming how **human** and **autonomous artificial** agents interact
- Technologies enabling these interactions include technical systems but also operational processes and people who use and interact with these technical systems
- **Sociotechnical systems** (STS) are cyber-physical systems that incorporate **social** and **technical** aspects facilitating the **interactions** of multiple autonomous participants whose interests are at best imperfectly aligned (Sing, 2013)
  - **Publishing system** and **manufacturing industry** are STSs where technology is used in parallel with specific **social** and **organizational strategies** to control production

Sociotechnical System (STS)

- Human society is invariably becoming more **dependent on** technologies, in particular **information technology**, which is transforming how **human** and **autonomous artificial** agents interact
- Technologies enabling these interactions include technical systems but also operational processes and people who use and interact with these technical systems
- **Sociotechnical systems** (STS) are cyber-physical systems that incorporate **social** and **technical** aspects facilitating the **interactions** of multiple autonomous participants whose interests are at best imperfectly aligned (Sing, 2013)
  - **Publishing system** and **manufacturing industry** are STSs where technology is used in parallel with specific **social** and **organizational strategies** to control production

Sociotechnical System (STS)

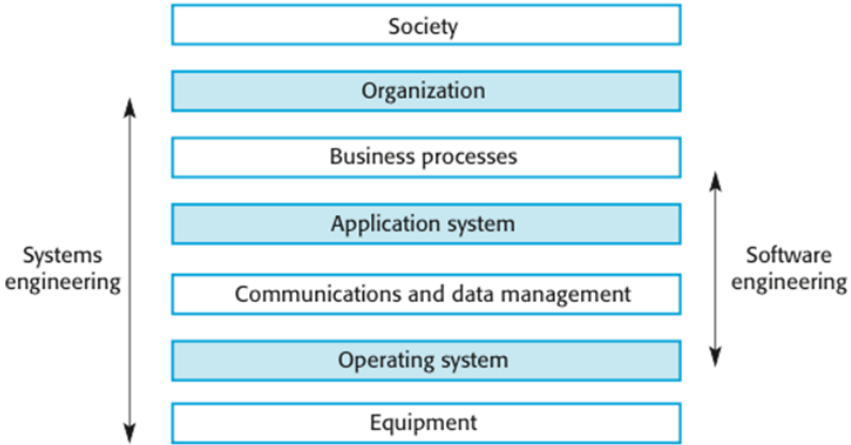


Figure : Sociotechnical system stack

Source: http://csis.pace.edu/~marchese/SE616_New/L10/L10_new.htm

Regulation of Sociotechnical Systems

- The success of STS relies on effective **governance**, which pertains to how the mentioned interactions are controlled (Balke & Villatoro, 2012)
- Governance maintains the equilibrium between the
  - **Autonomy** and **flexibility** of the participants whose interests are at best imperfectly aligned
  - **Predictability** desired in technological systems
- **Normative Multiagent System** has been proposed to govern these systems (Singh, 2013)

Multiagent Systems

A set of autonomous agents interacting with each other within a shared environment, eventually under one to multiple organizations

- **Agents**: autonomous decision-making entities able to react to events while pursuing (pro-actively defined or delegated) goals and directing actions to achieve them
- **Environment**: shared medium providing the surrounding conditions for agents to exist and act
- **Interaction**: motor of dynamics and interoperability in the MAS
- **Organization**: abstractions to declare and make accessible to agents their collective structure and functioning in a shared environment

Multiagent Systems

A set of **autonomous agents** interacting with each other within a shared environment, eventually under one to multiple organizations

- **Agents**: autonomous decision-making entities able to react to events while pursuing (pro-actively defined or delegated) goals and directing actions to achieve them
- **Environment**: shared medium providing the surrounding conditions for agents to exist and act
- **Interaction**: motor of dynamics and interoperability in the MAS
- **Organization**: abstractions to declare and make accessible to agents their collective structure and functioning in a shared environment

Multiagent Systems

A set of autonomous agents interacting with each other within a **shared environment**, eventually under one to multiple organizations

- **Agents**: autonomous decision-making entities able to react to events while pursuing (pro-actively defined or delegated) goals and directing actions to achieve them
- **Environment**: shared medium providing the surrounding conditions for agents to exist and act
- **Interaction**: motor of dynamics and interoperability in the MAS
- **Organization**: abstractions to declare and make accessible to agents their collective structure and functioning in a shared environment

Multiagent Systems

A set of autonomous agents **interacting with each other** within a shared environment, eventually under one to multiple organizations

- **Agents**: autonomous decision-making entities able to react to events while pursuing (pro-actively defined or delegated) goals and directing actions to achieve them
- **Environment**: shared medium providing the surrounding conditions for agents to exist and act
- **Interaction**: motor of dynamics and interoperability in the MAS
- **Organization**: abstractions to declare and make accessible to agents their collective structure and functioning in a shared environment

Multiagent Systems

A set of autonomous agents interacting with each other within a shared environment, eventually under **one to multiple organizations**

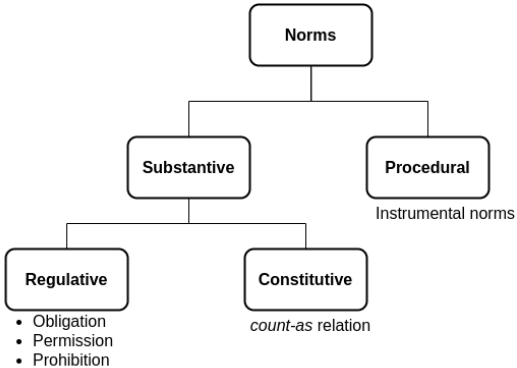
- **Agents**: autonomous decision-making entities able to react to events while pursuing (pro-actively defined or delegated) goals and directing actions to achieve them
- **Environment**: shared medium providing the surrounding conditions for agents to exist and act
- **Interaction**: motor of dynamics and interoperability in the MAS
- **Organization**: abstractions to declare and make accessible to agents their collective structure and functioning in a shared environment

Normative Multiagent Systems

A **Normative Multiagent System** (NMAS) is a multiagent system organized by means of **mechanisms** to represent, communicate, distribute, detect, create, modify, and enforce **norms**, and mechanisms to **deliberate about norms** and **detect norm violation and fulfillment**. (Boella et al., 2008)

Norms

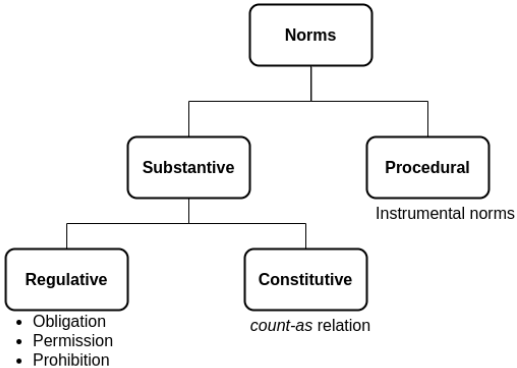
Norms represent the *behavior that each party in a system expects from others* and may be willing to enforce (Hollander & Wu, 2011)



Norms **do not** guarantee that agents will comply with them

Norms

Norms represent the *behavior that each party in a system expects from others* and may be willing to enforce (Hollander & Wu, 2011)



Norms **do not** guarantee that agents will comply with them

Norm Enforcement Mechanisms

Enforcement mechanisms motivate norm compliance

- **Regimentation**
  - Agents are **not allowed to violate** norms
  - Require mechanisms to **block violations**
  - Examples: Prison, Message Filtering
- **Regulation**
  - Agents are **allowed to violate** norms
  - Require mechanisms to **detect and react to violations**
  - Examples: Parking fines, suspension to system access

(Grossi et al, 2007)

Norm Enforcement Mechanisms

Enforcement mechanisms motivate norm compliance

- **Regimentation**
  - Agents are **not allowed to violate** norms
  - Require mechanisms to **block violations**
  - Examples: Prison, Message Filtering
- **Regulation**
  - Agents are **allowed to violate** norms
  - Require mechanisms to **detect and react to violations**
  - Examples: Parking fines, suspension to system access

(Grossi et al, 2007)

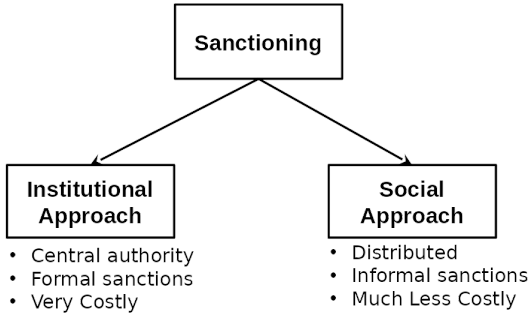
Norm Enforcement Mechanisms

	Observer	Enforcer	Sanctions	Taxonomy
Regimentation	Infrastructure	Infrastructure (mental states)	Impossible violations	Infrastructural control (*white box*)
		Infrastructure (agent actions)		Infrastructural control (*black box*)
Enforcement	Infrastructure entities	Infrastructural entities	Infrastructural sanction	Institutionalization of agents
	Third-party			Infrastructural assisted enforcement
		Social enforcement	Vicarious, Retaliation, Reciprocation	Informal Control
	Second-party	Second-party enforcement	Retaliation, Reciprocation	Promisee-enforced rules
		Infrastructural entities	Infrastructural sanction	Infrastructural assisted enforcement
	First-party	First-party enforcement	Self-sanction	Self-control

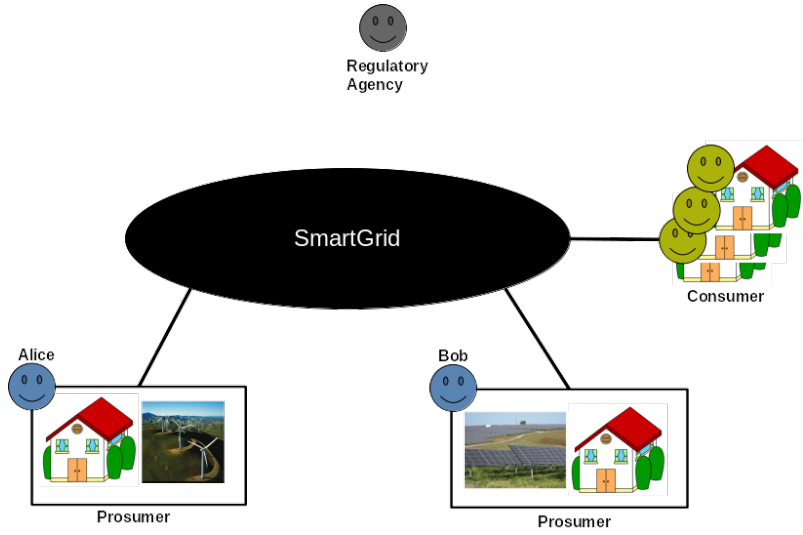
Figure : Norm enforcement taxonomy (Balke, 2009)

Sanctions

- **Sanctions** are negative or positive reactions to potentially any violation of or compliance with an expectation (i.e., a norm)
- Used to **influence agents' decision-making** and steer the system in the preferred direction



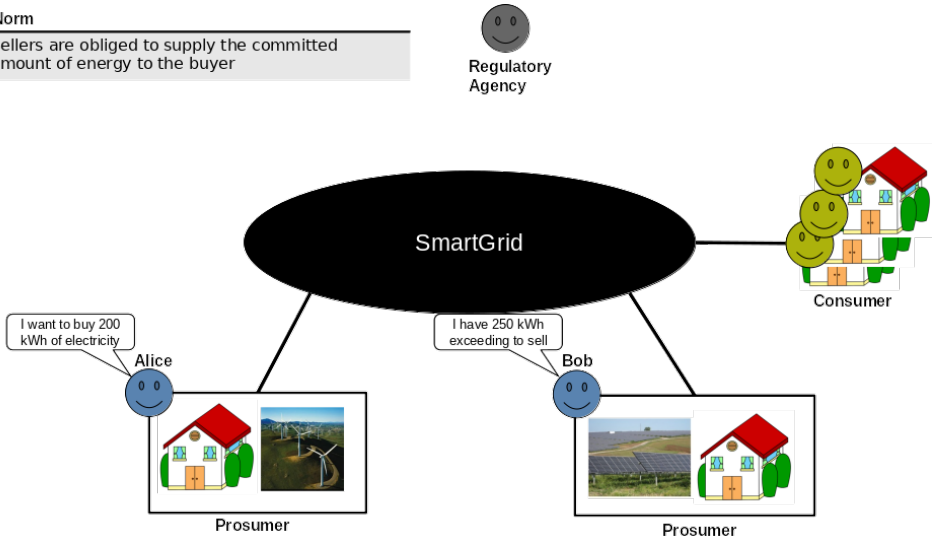
SmartGrid Scenario



SmartGrid Scenario

Norm

Sellers are obliged to supply the committed amount of energy to the buyer

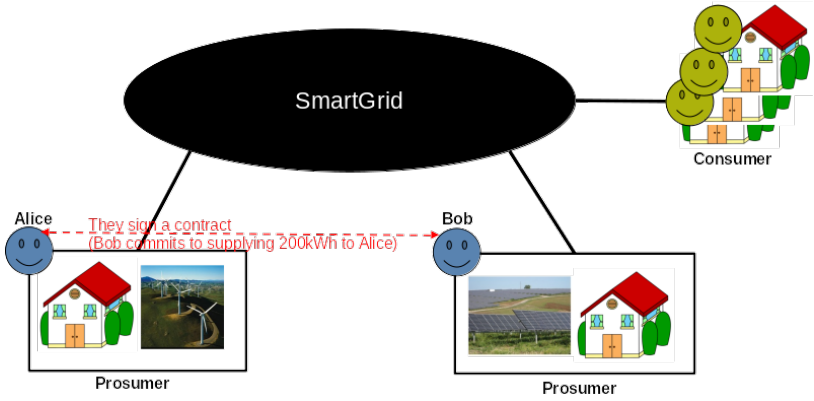


SmartGrid Scenario

Norm

Sellers are obliged to supply the committed amount of energy to the buyer

Regulatory Agency

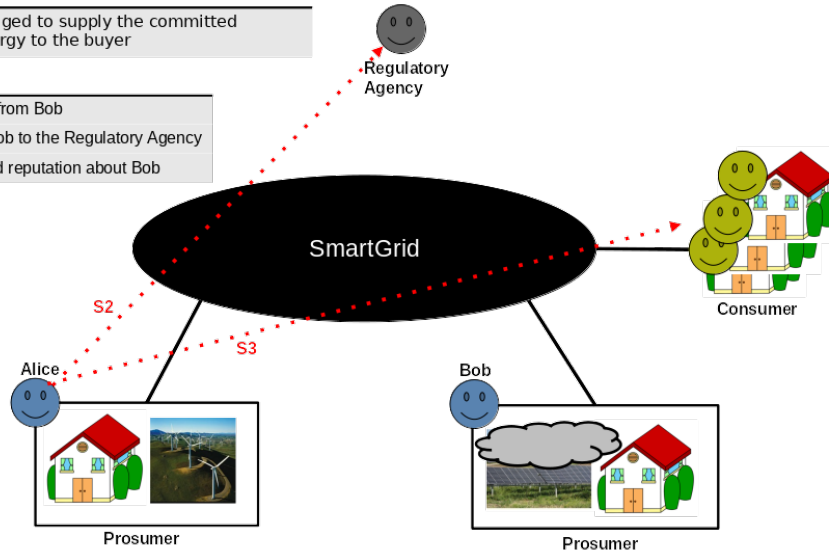


SmartGrid Scenario

Sellers are obliged to supply the committed amount of energy to the buyer

Sanctions

S1	Stop buying from Bob
S2	Denounce Bob to the Regulatory Agency
S3	Spread a bad reputation about Bob



Requirements

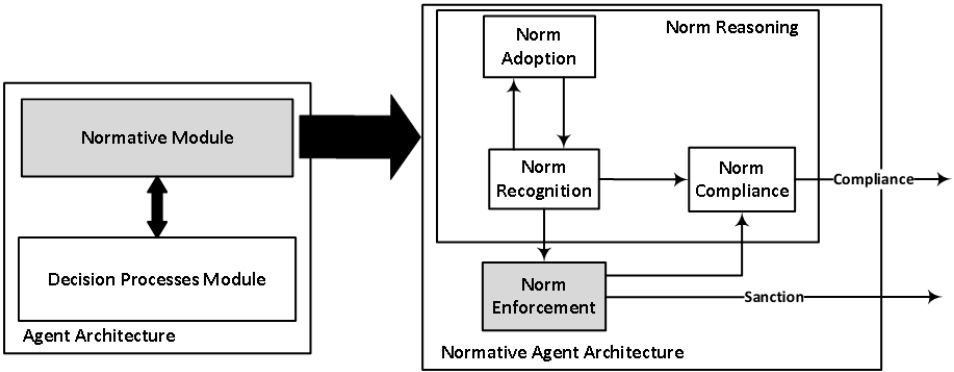
The requirements for an effective enforcement model to be used for regulating sociotechnical systems are

R1 **Multiple categories of sanctions**

R2 Potential **association of multiple sanctions with a norm** violation or compliance

R3 **Reasoning about most adequate sanction(s) to apply** based on several decision factors

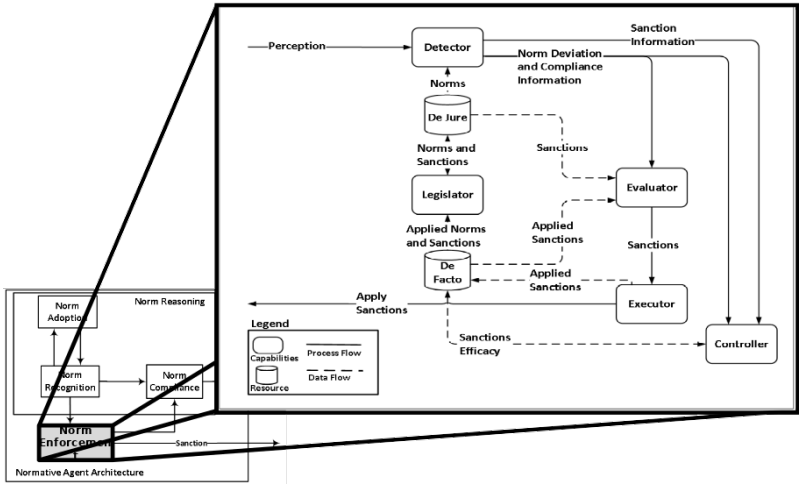
EMILIA: Normative Agent Architecture



Source code available at `https://github.com/gnardin/emilia`

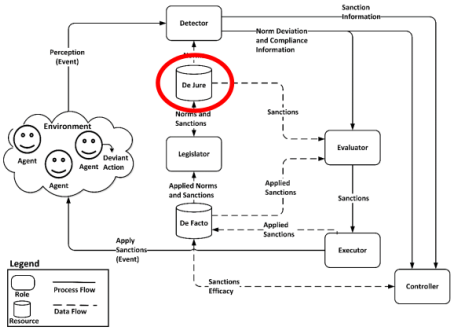
(Nardin et al., 2016a)

Sanctioning Process Model



(Nardin et al., 2016a)

Sanctioning Process Model

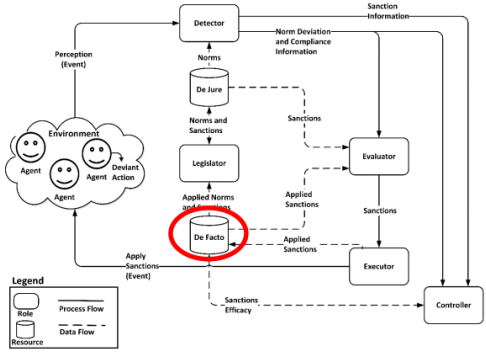


– **De Jure**

- Stores all the norms and sanctions specifications
- Enable the linking between a norm and multiple sanctions
- Relations can be many to many

Norms	Sanctions	
$N_1\langle$**active, role(A) = Prosumer, State,**	**S1**	Denounce violation to the Regulatory Agency
	S2	Spread the negatively updated reputation score to other agent
	S3	Ostracize the agent

Sanctioning Process Model

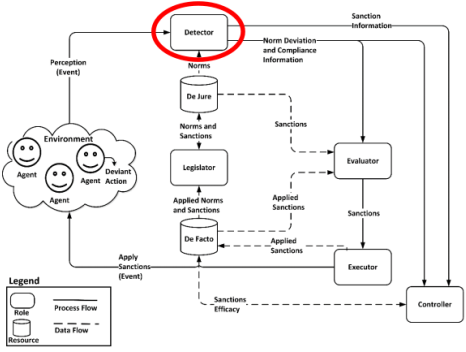


– **De Facto**

- Stores data about sanctioning activities performed or observed
- Stores data about the efficacy of the applied sanctions in promoting compliance

Norm	Sanction	Target	Violations	Efficacy
N1	S1	X	10	2
N1	S2	X	5	4
N1	S3	X	4	0

Sanctioning Process Model

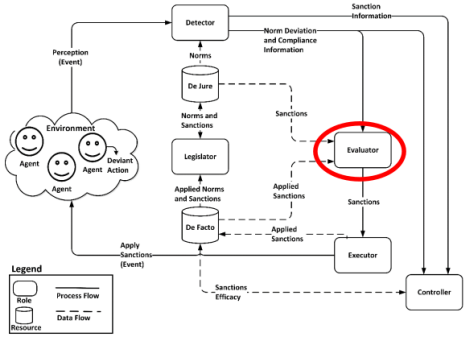


– **Detector**

- Checks whether the content of an observed event is ruled or not by any norm stored in the De Jure
- If matches with norm specification, then the Evaluator and Controller processes are triggered

Event	Norm
$e\langle 01012015, \mathrm{C}, \mathrm{B}, \mathrm{supply}(195\ kWh)\rangle$	$N_1\langle$**active, role(A) = Prosumer, State,**

Sanctioning Process Model



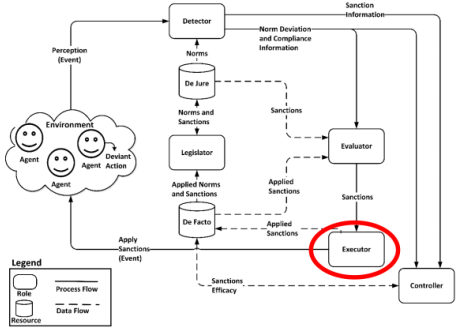
– **Evaluator**

- Evaluates whether the event complies or violates the norm
- Retrieve from the De Jure all the applicable sanctions associated to the norm
- Choose among sanctions the more adequate ones based on Sanctioning Decision Factors (*Factors*)

$$evaluate: NS_e \times LS \times Factors$$

S1	Denounce the violation to the Regulatory Agency
S2	**Spread the negatively updated reputation score to other agent**
S3	Ostracize the Prosumer

Sanctioning Process Model

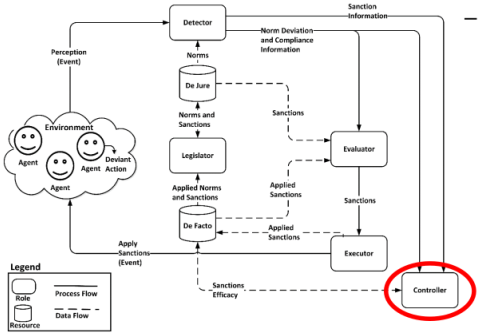


– **Executor**

- Receives the set of sanctions and executes them, if it has the power to execute
- Otherwise, it requests another agent to execute

Sanction	Event
S2 Spread the negatively updated reputation score to other agent	Send to all its neighbors the message *Agent X has a low reputation*

Sanctioning Process Model



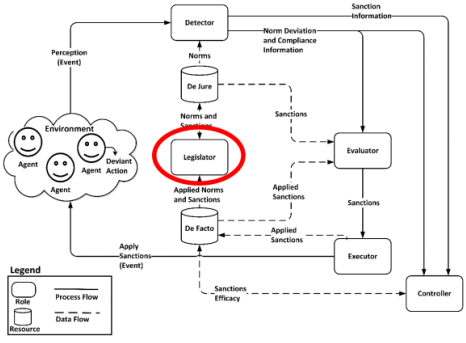
– **Controller**

- Monitors the outcomes of applied sanctions, and stores and reviews the De Facto repository with them

Agent X complied with the norm after had been sanctioned

Norm	Sanction	Target	Violations	Efficacy
N1	S1	X	10	2
N1	**S2**	**X**	**6**	**5**
N1	S3	X	4	0

Sanctioning Process Model



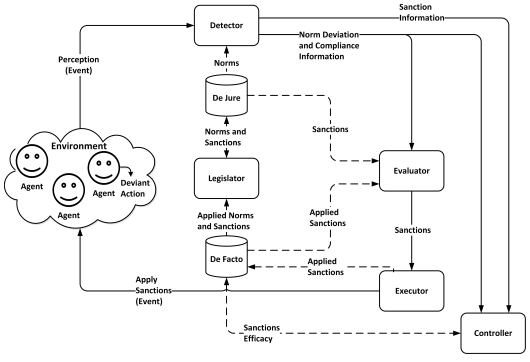
– **Legislator**

- Updates De Jure norms and sanctions specifications based on an assessment of the De Facto repositories

Norms	Sanctions	
$N_1\langle \textbf{active, role(A) = Prosumer, State,}$	**S1**	Denounce violation to the Regulatory Agency
	S2	Spread the negatively updated reputation score to other agent
	S3	~~Ostracize the agent~~

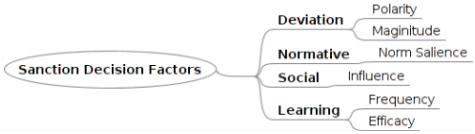
Sanctioning Process Model

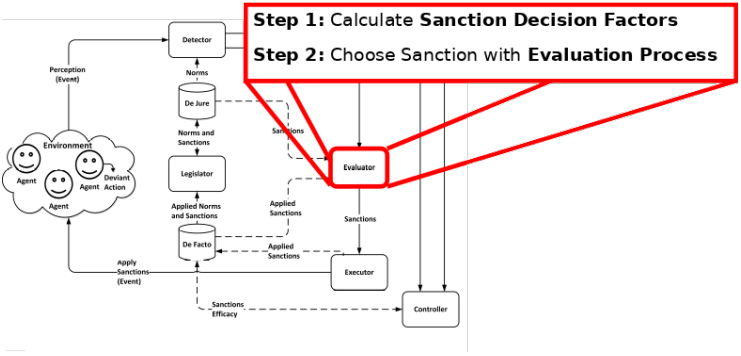
- **De Jure** – stores **norms, sanction, and norm-sanction links**
- **De Facto** – stores information about the **applied sanctions** and other relevant information used to assess the **efficacy of the applied sanctions**



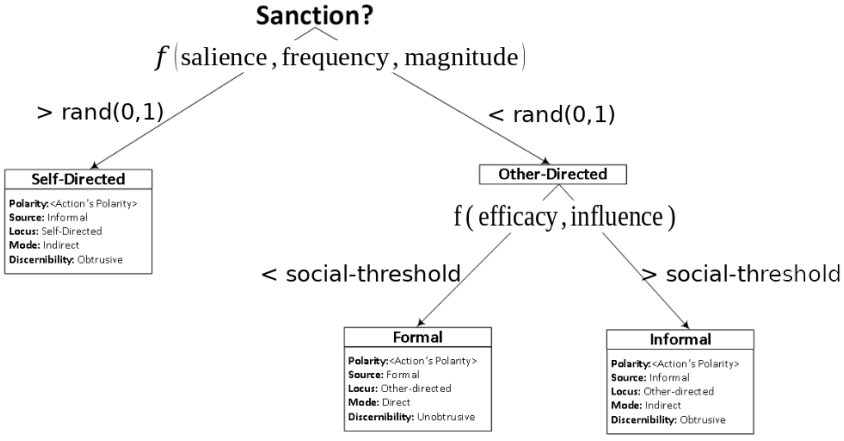
(Nardin et al., 2016b)

Sanctioning Evaluation Model





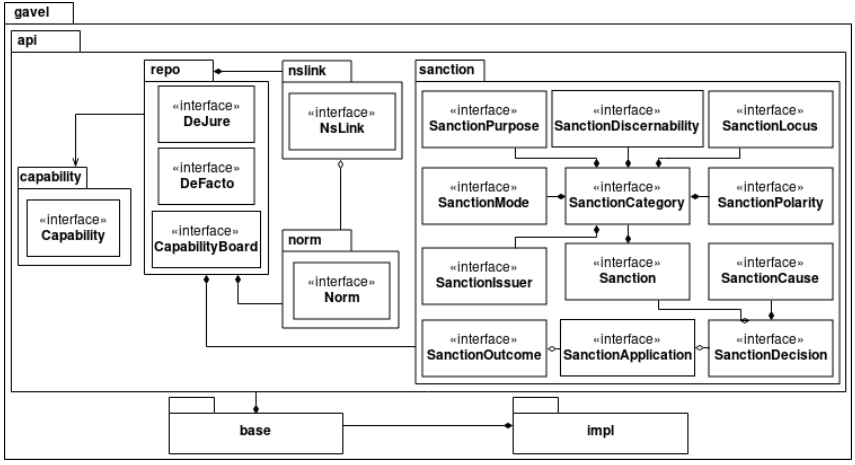
Sanctioning Evaluation Model



Gavel

- **Gavel** is an **adaptive sanctioning enforcement framework** based on the conceptual sanctioning process model.
- The sanctioning process is realised by
  - agents endowed with **special capabilities**; and
  - supporting **data repositories**.

Gavel Architecture



Source code available at `https://github.com/gavelproject/gavel`

(de Lima et al., 2019)

SmartGrid Experiments

- **Research Questions**
  - What is the **effect of a mono-type sanctioning policy** on the level of norm compliance and the use of non-renewable energy **in comparison to a multi-type sanctioning policy?**

- **Hypotheses**
  - **H1** A multi-type sanctioning policy **increases the level of norm compliance** in comparison to a mono-type policy
  - **H2** A multi-type sanctioning policy **decreases the use of non-renewable energy** in comparison to a mono-type policy

SmartGrid Experiments

- Scenario composed of
  - 100 Prosumers
  - 1 Provider
  - 1 Broker
  - 1 Regulatory Agency

Norms

N1	Sellers are obliged to supply the committed amount of energy to the buyer

Sanctions

S1	Denounce the violation to the Regulatory Agency
S2	Spread the negatively updated reputation score to other Prosumers
S3	Ostracize the Prosumer

Policy	**Description**
Base	No Sanction
Formal	Prosumers can use only Formal sanctions (i.e., Denounce)
Informal	Prosumers can use only Informal sanctions (i.e., Reputation spreading)
Hybrid (Formal + Informal)	Prosumers can choose to use between Formal and Informal sanctions (i.e., choose between Denounce or Reputation spreading)

SmartGrid Experiments

Metric	Base	Formal	Informal	Hybrid
Level Compliance	61.79%	**71.94%**	66.26%	68.40
Formal Sanction	0	**6.72**	0	1.16
Informal Sanction	0	0	16.09	**13.67**
Prosumer Sell	2875.60 kWh	**3357.06 kWh**	3125.63 kWh	3228.12 kWh

- It is **worth sanctioning**
- **Formal policy is more efficacious** than Informal and Hybrid policies in maintaining the level of norm compliance
- Formal policy requires **6 times more sanctions** than the Hybrid policy (**more costly**)

Increase Robustness of Production Automation

- Industry 4.0 requires high reconfigurability and flexibility to adapt cyber-physical production systems (CPPS) to changing customer requirements and customized products
- MAS approaches proposed because they support
  - decentralized decision-making
  - flexibility through a modular CPPS control architecture
- MAS approaches success relies partially upon the **agents' accurate and goal-oriented decision-making**
- However, in non-deterministic systems such as process plants, agents are at risk of making inaccurate decisions or not fulfilling their offer due to the non-determinism

(Land, Nardin, Vogel-Heuser, 2023)

Increase Robustness of Production Automation

- Agent-based controlled process plant
- Hybrid combination of logistic and process systems
  - Logistic system consists of conveyor belts, switches, and barcode scanners. Transport bottles to filling stations.
  - Process system represented by filling stations that fill bottles with recipe-specific liquids

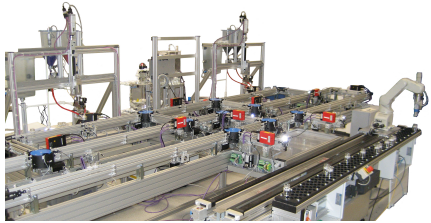


Figure : Production plant consisting of process and logistic parts

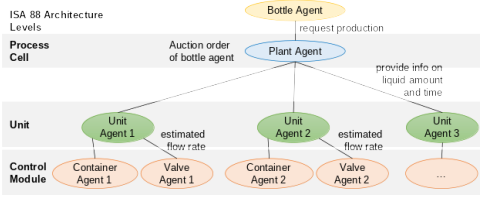


Figure : Multiagent architecture of the process plant based on ISA 88

Increase Robustness of Production Automation

- The filling stations can fill the bottles with different liquids
- Filling process uses a proportioning valve that opens for a specific time window to achieve a specific fill level in the bottle
- Residues of the liquid may adhere to the valve, clogging it over time

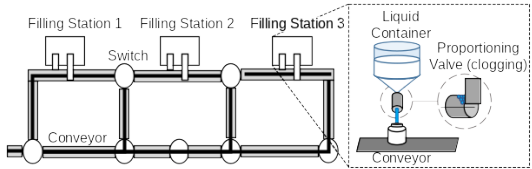


Figure : Process plant with filling stations and clogging of proportioning valve

Increase Robustness of Production Automation

- The norm defines that "*Unit agents are obliged to operate under the cost committed to the plant agent*"
- Cost means that if a unit agent commits to providing 400ml of a specific liquid within two seconds, it must comply with.

	Sanctioning Decision	Sanction
Unit-Agent	1) Update the valve agent's image for that task (liquid and time) based on the polarity and magnitude deviation factors (if the image below a threshold)	Signal to the valve agent to adjust its flow rate estimation
Unit-Agent	2) Update the valve's image for that task (same liquid) based on the polarity and magnitude deviation factors. If image below a threshold and it has not complied with the norm for 2-3 consecutive times	Activate the valve's self-cleaning routine (loss of potential orders during cleaning)
Plant Agent	3) Update the unit agent's image based on the polarity and magnitude deviation factors. If the image below a threshold	Plant agent reduces the likelihood to select unit agent to fulfill the bottle orders
Plant Agent	4) Update the unit agent's image based on the polarity and magnitude deviation factors. If the unit agent's image is below a threshold and it has not complied with the norm for five consecutive times	Plant agent removes unit agent as a possible option for fulfilling bottle orders and generate an alarm: Unit agent requires manual work (long downtime)

Table : Sanctioning decisions and sanctions

Increase Robustness of Production Automation

- Bottle agent enacts the **Detector** capability
- Plant and Unit agents enact the **Evaluator** and **Executor** capabilities
- The **Controller** and **Legislator** as well as the repositories extends the current Production plant MAS architecture

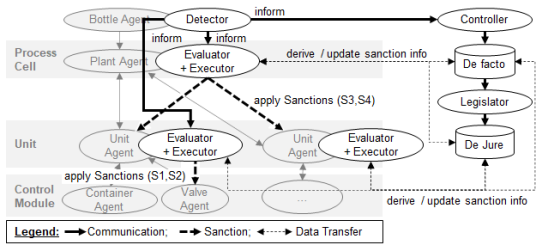


Figure : Sanctioning process integrated into the production plant MAS

Challenges

- **Normative Language**
  - Define a declarative normative language that accounts for sociotechnical system requirements, such as trustworthiness
- **Enforcement Model**
  - Identify main factors impacting the agents' decision-making (probably domain application dependent)
  - Identify psychological models to support the agents' decision-making regarding why, when, how to sanction
- **Applications**
  - Expand the application of the approach to demonstrate its usefulness in different real applications

References

- Balke, T. (2009). A taxonomy for ensuring institutional compliance in utility computing. In Boella, G. et al. (Ed.). *Normative Multi-Agent Systems*. Dagstuhl: Schloss Dagstuhl – Leibniz-Zentrum fuer Informatik.
- Balke, T. & Villatoro, D. (2012) Operationalization of the sanctioning process in utilitarian artificial societies. In Cranefield, S. et al. (Ed.). *Coordination, Organizations, Institutions, and Norms in Agent System VII* (pp. 167–185). Cham: Springer.
- Boella, G., Torre, L. van der, & Verhagen, H. (2008). Introduction to the special issue on normative multiagent systems. *Journal of Autonomous Agents and Multi-Agent Systems*, 17(1), 1–10.
- de Lima, I. C. A., Nardin, L. G., & Sichman, J. S. (2019). Gavel: A sanctioning enforcement framework. In D. Weyns, V. Mascardi, & A. Ricci (Eds.), *Engineering Multi-Agent Systems* (pp. 225–241). Cham: Springer.
- Grossi, D., Aldewereld, H., & Dignum, F. (2007) Ubi Lex, Ibi Poena: Designing norm enforcement in e-institutions. In Noriega, P. et al. (Ed.). *Coordination, Organizations, Institutions, and Norms in Agent Systems II* (pp. 101–114). Cham: Springer.
- Hollander, C. D. & Wu, A. S. (2011). The current state of normative agent-based systems. *Journal of Artificial Societies and Social Simulation*, 14(2).
- Land, K., Nardin, L. G., & Vogel-Heuser, B. (2023). Increasing robustness of agents' decision-making in production automation using sanctioning. In Proceedings of the 2023 IEEE 21st International Conference on Industrial Informatics (INDIN) (pp. 1–6). Pistacaway, NJ: IEEE.
- Nardin, L. G., Andrighetto, G., Conte, R., Székely, Á., Anzola, D., Elsenbroich, C., Lotzmann, U., Neumann, M., Punzo, V., & Troitzsch, K. G. (2016a). Simulating protection rackets: A case study of the Sicilian Mafia. *Journal of Autonomous Agents and Multi-Agent Systems*, 30(6), 1117–1147.
- Nardin, L. G., Balke, T., Ajmeri, N., Kalia, A. A., Sichman, J. S., & Singh, M. P. (2016b). Classifying sanctions and designing a conceptual sanctioning process model for socio-technical systems. *The Knowledge Engineering Review*, 31(2), 142–166.
- Singh, M. P. (2013) Norms as a basis for governing sociotechnical systems. *ACM Transactions on Intelligent Systems and Technology*, ACM Press. 5(1), 21:1–21:23.

Thank you!