A Framework for Monitoring Agent-Based Normative Systems *

Sanjay Modgil, Noura Faci, Felipe Meneguzzi, Nir Oren, Simon Miles, Michael Luck

Agents and Intelligent Systems Group
Dept Computer Science
King’s College London

* Research Funded by European Commission 6th framework CONTRACT project (www.ist-contract.org)
Outline of Talk

- Requirements for monitoring
- Overview of monitoring framework
- Representation and processing of norms for monitoring
- Agreements on what observations count as norm statuses and trusted observers
- Implementation and use case
- Conclusions and future work
Requirements for monitoring normative systems

- Norms (obligations, prohibitions, permissions) used to regulate and coordinate agent behaviours

- Autonomous agents may violate norms
  - enforcement mechanisms (sanctions) required to motivate compliance
  - monitoring of agent behaviours to determine compliance / violation

- Example obligation on purchaser P of goods G from a supplier S:

  If P is notified by S that goods G are in stock, then unless S is declared bankrupt, either P must
  • cancel the order within 7 days of receipt of notification,
  or;
  • accept the order and pay within 7 days of receipt of notification.
Requirements for monitoring normative systems

- Norms (obligations, prohibitions, permissions) used to regulate and coordinate agent behaviours

- Autonomous agents may violate norms
  - enforcement mechanisms (sanctions) required to motivate compliance
  - monitoring of agent behaviours to determine compliance / violation
General Framework for Monitoring Normative Systems

Monitoring Framework

- Manager
  - norm status reports
  - with explanations
  - inform agents and impose sanctions

- Monitor
  - observations

- Norms = N1, N2, ...
  - system neutral representation of individual norms for monitoring
  - abstract model of norms

Normative System
An abstract model of norms for monitoring

- Norms describe complex behaviours and world states brought about by (groups of) agents, e.g., consider following obligation:

  *If P is notified by S that goods G are in stock, then unless S is declared bankrupt, either P must:

  • cancel the order within 7 days of receipt of notification, or;
  • accept the order and pay within 7 days of receipt of notification.*

  \[(Type, Target, Activation, Condition, Expiration)\]

  *Type* ∈ \{obligation, permission, prohibition\},
  *Target* = agents whose behaviour is governed by the norm \(P\)
An abstract model of norms for monitoring

- Norms describe complex behaviours and world states brought about by (groups of) agents

*If P is notified by S that goods G are in stock, then unless S is declared bankrupt, either P must:

• cancel the order within 7 days of receipt of notification,
• accept the order and pay within 7 days of receipt of notification.*

*Activation* = conditions under which norm applies to *Target*
An abstract model of norms for monitoring

- Norms describe complex behaviours and world states brought about by (groups of) agents

  If $P$ is notified by $S$ that goods $G$ are in stock, then unless $S$ is declared bankrupt, either $P$ must:
  
  - cancel the order within 7 days of receipt of notification, or;
  - accept the order and pay within 7 days of receipt of notification.

**Activation** = conditions under which norm applies to **Target**

**Condition** = state that must be (obligation) may be (permission) or must not be (prohibition) realised by **Target**
An abstract model of norms for monitoring

- Norms describe complex behaviours and world states brought about by (groups of) agents

If P is notified by S that goods G are in stock, then unless S is declared bankrupt, either P must:

- cancel the order within 7 days of receipt of notification, or;
- accept the order and pay within 7 days of receipt of notification.

**Activation** = conditions under which norm applies to Target

**Condition** = state that must be (obligation) may be (permission) or must not be (prohibition) realised by Target

**Expiration** = conditions under which norm no longer applies
Representing norms for Monitoring

- Norms in many normative systems conform to an abstract model
  - norms can be mapped to Augmented Transition Networks (ATNs) that reference elements of the model and are used for monitoring status of norms.

ATN-1
ATN-2
Monitor
Manager

status reports on norms

Norms = N1, N2, ...

(\textit{Type, Target, Activation, Condition, Expiration})

Manager

Monitor

status reports on norms
Processing ATN representations of norms

- ATNs are directed labelled graphs that transition from node to node based on satisfaction of connecting arcs’ labels.

- DNF representations of Activation = \((\alpha_1 \land ... \land \alpha_n) \lor (\beta_1 \land ... \land \beta_m) \lor ...\) and Condition = \((\gamma_1 \land ... \land \gamma_n) \lor (\delta_1 \land ... \land \delta_m) \lor ...\) map to labels of ATN.

- If Expiration does not hold, and:
  - ATN in S1 (norm not active) and at least one arc label satisfied then transition arc to activation state S2.
  - ATN in S2 and at least one arc label satisfied then transition arc to fulfillment state S3 (if obligation/permission) or violation state S3 (if prohibition).
  - Obligation ATN in S2 and time window elapses then obligation violated.
Processing ATN representations of norms
If \textit{NormExpiration} does not hold, and:

- \( P \) notified by \( S \) that \( G \) in stock at time \( T \) then transition to \( S_2 \) and norm activated (in force w.r.t. \( P \))

- If \( P \) cancels at \( T_1 \) (\( T_1 \leq T + 7 \)) or
  
  \( P \) accepts at \( T_2 \) (\( T_2 \leq T + 7 \)) and pays at \( T_3 \) (\( T_3 \leq T + 7 \))

  then transition corresponding arc to fulfillment node \( S_3 \)

- If current time \( T' \) greater than time window \( T + 7 \) and ATN in \( S_2 \) then obligation violated
Processing ATN representations of maintenance norms

- 3 node ATNs for achievement norms
- 4 node ATNs for maintenance norms that may toggle between violated and not violated (e.g. obligation to drive on left)

Diagram:

- S1: Agent starts driving
- S2: driving on right
- S3: not violated
- S4: violated

Transitions:
- From S1 to S2: driving on right
- From S2 to S3: driving on left
- From S3 to S4: driving on left
- From S4 to S2: driving on right
Matching observations with ATN labels

- Monitors match observations of world states and agent behaviours with ATN labels => transition ATNs across corresponding arcs

Diagram:
- Monitoring Framework
- Manager
- Monitor
- Observers
- Environment
- Norms = N1, N2, ...
- Normative System
Counts as agreements and trusted observers

- Choice of observations (brute facts) that \textit{count as Activation, Condition, Expiration and} choice of observers responsible for observations
  
  \[ \Rightarrow \] important for motivating agent participation in normative systems since these choices impact on agents’ confidence that sanctions applied appropriately

- What observed brute facts should count as \textit{P paying S} and who should make and report these observations?
Counts as agreements and trusted observers

- Observed message sent from $P$ to $S$ notifying $S$ of payment - sent message counts as $P$ pays $S$
  
  => $P$ may send message without actually paying and so avoid sanction
  => $S$ not motivated to participate

- $S$ sends notification message to monitor (i.e., $S$ acts as observer) that money deposited in bank - money in bank counts as $P$ pays $S$
  
  => money may be in bank but $S$ does not notify and so sanction inappropriately imposed on $P$ (resulting in some gain for $S$)
  => $P$ not motivated to participate
Counts as agreements and trusted observers

- Agents agree on what brute facts count as Activation, Condition, and Expiration and agree on observers responsible for observing and relaying facts to monitor.

Agents entrust observer to observe for and report $X$ to monitor

- e.g., $S$ and $P$ agree to trust bank

$X$ counts as $Y$

agreed by agents

ATN arcs

e.g. $S$ and $P$ agree that money in bank counts as $S$ pays $P$
Implementation and Use Case *

- Jason implementation of interacting aerospace agents governed by electronic contract

- Examples of contract normative clauses:
  - Service site obliged to repair engine for airline operator within 7 days
  - Service site prohibited from sourcing parts from part manufacturer

- JAVA implemented monitor processes observed messages exchanged between aerospace agents, together with ATN representations of norms

  Norm status reports visualised in GUI proxy for manager

Conclusions and Future Work

- General framework for monitoring individual norms in underlying normative systems
  - Monitors process ATN representations of norms labelled by observed facts agreed by system agents as counting as statuses of norms
  - Facts relayed to monitors by observers explicitly entrusted by agents
  - Monitors create status reports on norms

- Implementation of monitor validated on electronic contract

- Future Work:
  - predictive monitoring whereby recognition of danger states (encoded as additional nodes in ATNs) signals danger of normative violations
  - enhanced explanations of normative violations