Norm Identification in Jason using a Bayesian Approach

Guilherme Krzisch Felipe Meneguzzi

School of Computer Science
Pontifical Catholic University of Rio Grande do Sul
Porto Alegre, Brazil

guilhermekrz@gmail.com,felipe.meneguzzi@pucrs.br

May 8, 2017

Overview

1. Introduction
2. Background
3. Norm Detecting System
4. Experiments and Results
5. Conclusion and Future Work

Introduction

- Open Multi-Agent Systems
- Normative Environments
- How to identify existing norms, in order to avoid sanctions?
  - E.g. when there are no explicit norms or if there is no trust between agents
  - Using an existing Bayesian Identification approach
- We propose a multi-agent system testbed to validate the approach

Background - Norms

Definition 1 (Norm)

A norm exists in a given social setting to the extent that individuals usually act in a certain way and are often punished when seen not to be acting in this way.

- Regulate conflicts
- Coordinate behavior
- Central vs No central authority
- Examples:
  - Driving on the left/right side of the road
  - Dueling
  - Smoke in public

Background - Norms

Five phases of norm development:

1. Creation
2. **Identification** - How a new agent entering the society can infer the norms created and currently being enforced in the system
3. Spreading
4. Enforcement
5. Emergence

Background - Norm identification

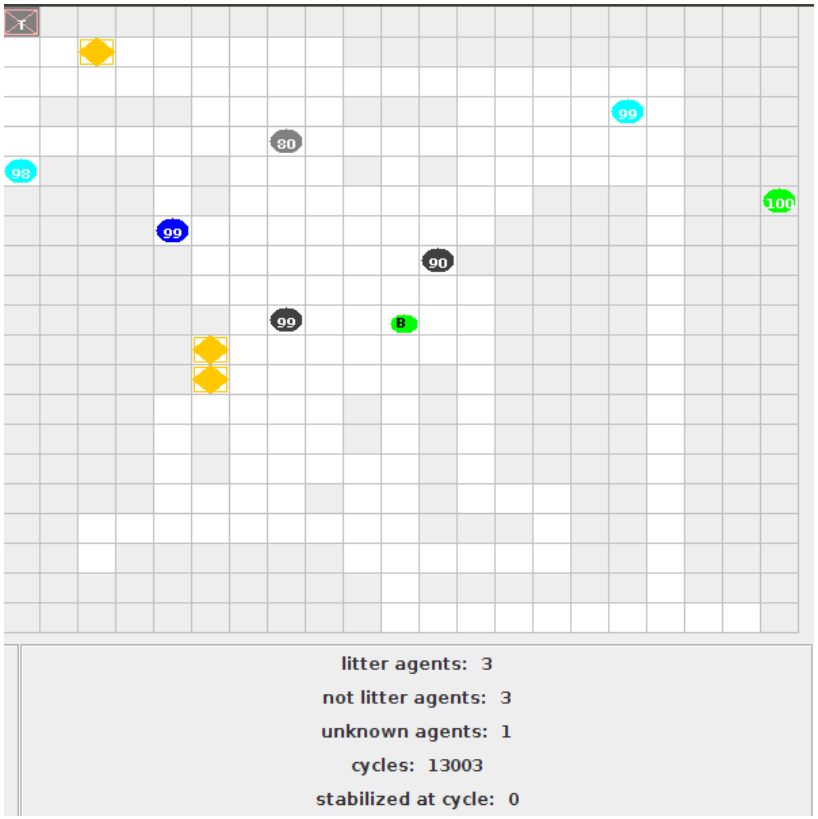
- We use a Bayesian approach from *Cranefield et al.* [1]
- Norms are defined in a subset of Linear Temporal Logic
  - *eventually*(δ), *never*(δ), *next*(γ, δ), *not_next*(γ, δ), *eventually*(γ, δ) and *never*(γ, δ)
- Weight norm hypothesis according to a set of observations over a model of the system (state-space graph)
- Alternative interpretation of the Bayes Theorem, which uses odds against a null hypothesis:

$$O(H_1 : H_2|D) = \frac{p(H_1|D)}{p(H_2|D)} = \frac{p(H_1)p(D|H_1)/p(D)}{p(H_2)p(D|H_2)/p(D)} = O(H_1 : H_2)\frac{p(D|H_1)}{p(D|H_2)}$$

- We select the top norm hypothesis

Norm Detecting System - Overview

- Jason + CArtAgO
- Park environment
- *Non littering agents* and *littering agents*, with an associated score
- There is a norm when the majority of the agents is from the same type



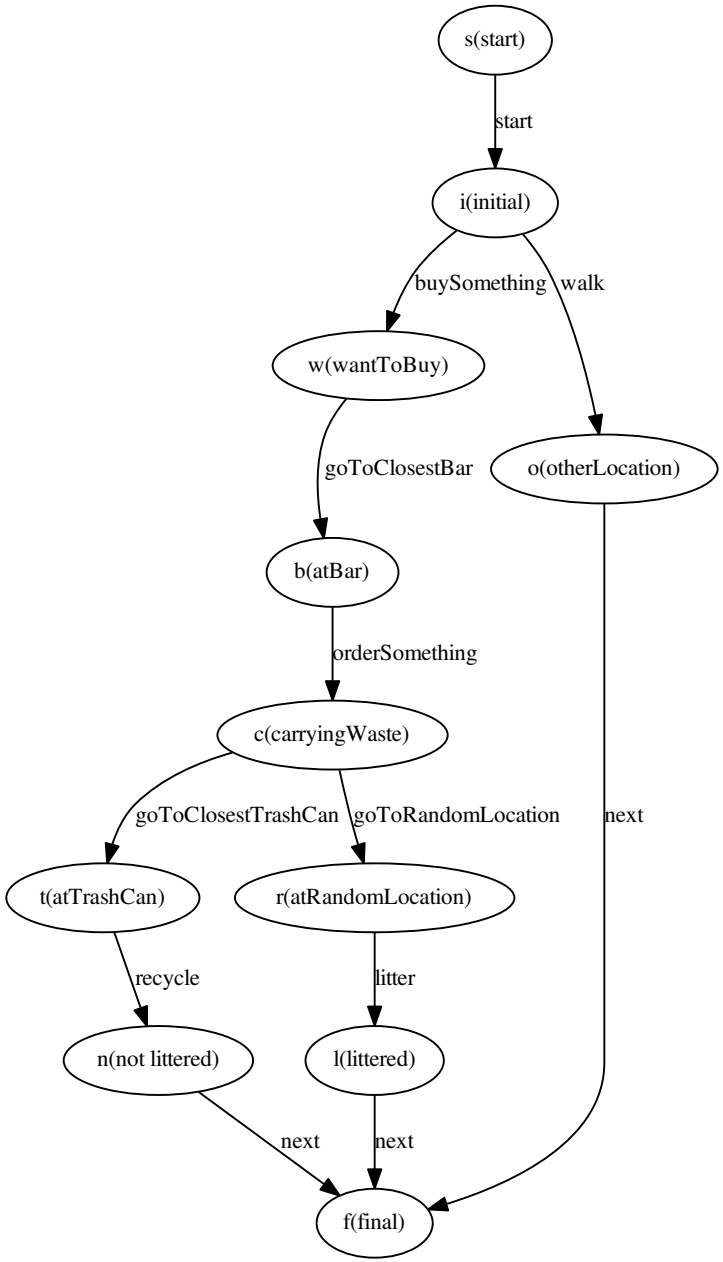
Norm Detecting System - State-space graph

Norm Detecting System - Algorithm

Algorithm 1 Norm Inference Procedure

1: **procedure** NormInferenceProcedure(stateSpaceGraph)
2: *observations* ← collect a set of observations
3: *normHypotheses* ← *normIdentificationAlgorithm*(*stateSpaceGraph*, *observations*)
4: *topNormHypotheses* ← retrieve top hypotheses from normHypotheses
5: *filtered* ← filter relevant topNormHypotheses
6: **for** normHypothesis in filtered **do**
7: check if normHypothesis indicates a litter or not litter norm
8: **end for**
9: return most probable norm based on the number of each norm indications
10: **end procedure**

Experiments and Results - Not Litter Example



Observation examples:

- i, o, f
- i, w, b, c
- i, w, b, c, t, n, f

Top norm hypothesis:

1. **('c', 'next', 't')***
2. **('c', 'not next', 'r')***
3. **('c', 'eventually', 'n')***
4. ('l', 'not next', 'f')
5. **('c', 'never', 'r')***
6. ('r', 'never', 'l')
7. ('r', 'eventually', 'i')
8. ('l', 'eventually', 's')
9. ('l', 'eventually', 'n')
10. ('r', 'eventually', 'c')

* Indicate a *not litter* norm

Experiments and Results - 10000 execution cycles

Percentage of litter agents	# of agents	Inferred norm
100% to 90%	6	litter
	50	litter
	100	litter
85% to 10%	6	none
	50	none
	100	none
5% to 0%	6	not litter
	50	not litter
	100	not litter

Table: Inferred norms for an increasingly number of agents, and the percentage of *littering* and *non littering* agents

Conclusion and Future Work

- Contributions
  - Norm inference testbed that can be used for experiments of norm-identification algorithms
  - Further experiments to validate the bayesian norm-identification approach by Cranefield *et al.*, confirming their positive result in a multi-agent setting
- Future Work
  - Automatically generate the state-space graph of the plan library built in Jason
  - Investigate different ways of combining the top norm hypotheses
  - Run more experiments

References

Cranefield, S.; Savarimuthu, T.; Meneguzzi, F.; and Oren, N. (2015)
A Bayesian approach to norm identification
Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems 1743–1744.

Savarimuthu, B. T. R.; Purvis, M.; Purvis, M.; and Cranefield, S. (2008)
Social norm emergence in virtual agent societies
Declarative Agent Languages and Technologies VI. Springer 1828.

Questions?

Norm Identification in Jason using a Bayesian Approach

Guilherme Krzisch Felipe Meneguzzi

School of Computer Science
Pontifical Catholic University of Rio Grande do Sul
Porto Alegre, Brazil

guilhermekrz@gmail.com,felipe.meneguzzi@pucrs.br

May 8, 2017