

Contents

Acknowledgments	v
Abstract	vii
Resumen	ix
Resum	xi
Contents	xvii
List of Figures	xxi
List of Tables	xxii
1 Introduction	1
1.1 Motivation	2
1.2 Objectives	5
1.3 Contributions	6
1.4 Document Structure	7
2 State of the Art	9
2.1 Introduction	9
2.2 Approaches to Norms	11
2.2.1 Sociological Approach to Norms	11
2.2.2 Philosophical Approach to Norms	13
2.2.3 Legal Approaches	14
2.2.4 Artificial Intelligence & Law	14
2.3 Norm Definition in Multi-Agent Systems	15
2.3.1 Norm Definition	15
2.3.2 Normative Multi-Agent Systems	18

2.4	Norm Representation	21
2.4.1	Deontic Logic: Logic of Norms	21
2.4.2	Input/Output Logic	25
2.4.3	Commitments	25
2.4.4	Social Law	26
2.4.5	Normative Positions	27
2.4.6	Power in Normative Systems	28
2.4.7	Norms and Time Considerations	29
2.4.8	Open Issues for a Logic of Normative Systems	29
2.5	Norm Implementation	30
2.5.1	Normative Language	30
2.5.2	Operational Norms	32
2.5.3	Implementation Mechanisms	36
2.5.4	Open Issues for Implementing Normative Multi-Agent Systems	42
2.6	Norm Reasoning	42
2.6.1	Norm Decision Making Systems	43
2.6.2	Norm Diagnosis Systems	49
2.6.3	Open Issues for Normative Reasoning	51
2.7	Norm Creation Process	54
2.7.1	Top-Down Approach	55
2.7.2	Bottom-Up Approach: Dynamic Emergence	57
2.7.3	Open Issues for the Emergence of Norms	60
2.8	Conclusions	61
2.8.1	Specification of Normative Systems	62
2.8.2	Individual Normative Reasoning	63
2.8.3	Implementation of Norms	63
2.8.4	Software Tools for Normative Multi-Agent Systems . . .	64
3	Normative Definitions	67
3.1	Introduction	68
3.2	Deontic Norms	70
3.2.1	Deontic Norm Definition	71
3.2.2	Deontic Instance Definition	72
3.3	Constitutive Norms	72
3.3.1	Constitutive Norm Definition	73
3.3.2	Constitutive Instance Definition	74
3.4	Conclusions	75

4	The n-BDI Architecture	77
4.1	Illustrative Scenario	78
4.2	Normative Multi-context Graded BDI Architecture	80
4.2.1	Mental Contexts	81
4.2.2	Functional Contexts	83
4.2.3	Normative Contexts	84
4.2.4	Reasoning Process in a n-BDI Agent	85
4.3	Norm Acquisition Context (NAC)	90
4.3.1	NAC Language	93
4.3.2	Norm Dynamics	95
4.4	Norm Compliance Context (NCC)	99
4.4.1	NCC Language	99
4.4.2	Instance Dynamics	101
4.5	Acquiring Norms: Evaluation	105
4.6	Contributions	107
4.7	Conclusions	111
5	Reasoning About Deontic Norms	113
5.1	Introduction	113
5.2	Norm-based Expansion for Deontic Norms	114
5.2.1	Obligation Internalization	115
5.2.2	Prohibition Internalization	117
5.2.3	Permission Internalization	117
5.3	Determining the Willingness to Norm Compliance	118
5.3.1	$\theta_{interest}$	119
5.3.2	$\theta_{expectation}$	120
5.3.3	$\theta_{emotion}$	122
5.4	Evaluation	127
5.4.1	Simulation Description	127
5.4.2	Results	132
5.4.3	Discussion	140
5.5	Contributions	142
5.6	Conclusions	143
6	Reasoning About Constitutive Norms	145
6.1	Introduction	145
6.2	Norm-based Expansion for Constitutive Norms	147
6.3	Case Study	149
6.3.1	Initial Situation	149
6.3.2	Normative Reasoning Process	149

6.4	Evaluation	152
6.4.1	Agent Implementation	154
6.4.2	Metrics	155
6.4.3	Results	156
6.5	Contributions	158
6.6	Conclusions	160
7	Coherence-based Contraction	161
7.1	Introduction	161
7.2	Coherence Theory	162
7.2.1	Deductive Coherence	164
7.3	Coherence for Multi-context Graded BDI Agents	165
7.3.1	Formalization of Deductive Coherence	165
7.3.2	Building the Coherence Graph	166
7.4	Coherence for n-BDI Agents	167
7.4.1	Coherence for the BC: Explanatory Constraints	169
7.4.2	Coherence for the NCC: Normative Constraints	170
7.4.3	Coherence for the DC: Deliberative Constraints	171
7.4.4	Coherence Between Contexts: Normative Bridge Rules	171
7.4.5	Coherence Maximization	176
7.5	Case Study	176
7.6	Contributions	178
7.7	Conclusions	180
8	Case Study	181
8.1	Introduction	181
8.1.1	Fire-Rescue Scenario Modelling	182
8.2	Non-Normative Fireman	183
8.3	Norm-Constrained Fireman	183
8.4	n-BDI Fireman	185
8.5	Experimental Description	188
8.5.1	Metrics	188
8.5.2	Experiment Results	189
8.6	Conclusions	194
9	MaNEA: A Distributed Architecture for Enforcing Norms in Open MAS	195
9.1	Introduction	196
9.2	Related Work	196
9.2.1	Infrastructural Observability	197

9.2.2	Requirements for Norm Enforcing Architectures	198
9.3	The Magentix2 Platform	199
9.3.1	Tracing Service Support	201
9.3.2	Organization Management System (OMS)	203
9.4	Norm-Enforcing Architecture: MaNEA	208
9.4.1	Norm Manager	208
9.4.2	Norm Enforcer	209
9.5	Implementation of the n-BDI Architecture	216
9.5.1	Jason	216
9.5.2	Implementing the n-BDI Architecture in Magentix2 using Jason	219
9.6	Case Study	226
9.7	Evaluation	228
9.7.1	Theoretical Results	231
9.7.2	Experimental Results	238
9.8	Contributions	245
9.9	Conclusions	246
10	Conclusions	247
10.1	Contributions	247
10.2	Future Work	248
10.3	Related Publications	250
10.3.1	Publications in Journals	250
10.3.2	Publications in Conferences	251
10.3.3	Book Chapters	253
	Bibliography	280

List of Figures

2.1	Operational interpretation of norms	37
2.2	Norm emergence	58
4.1	Perception phase in the n-BDI architecture	86
4.2	Decision-making phase in the n-BDI architecture.	89
4.3	Representation of norms and instances in the n-BDI architecture	91
4.4	Relative error with respect to the agent accuracy	108
5.1	Norm-based expansion for deontic norms in the n-BDI archi- tecture	116
5.2	Desire distribution of a randomly generated agent	129
5.3	Explanatory relationship graph	130
5.4	Percentage of instances that belong to each willingness category on average when $\delta_{compliance}$ takes value 0.1	133
5.5	Percentage of instances that belong to each willingness category on average when $\delta_{compliance}$ takes value 0.05	135
5.6	Percentage of instances that belong to each willingness category on average when $\delta_{compliance}$ takes value 0.2	135
5.7	Percentage of instances that belong to each willingness category on average when $\rho_{NAC} \in [0.0, 0.25]$	136
5.8	Percentage of instances that belong to each willingness category on average when $\rho_{NAC} \in [0.25, 0.5]$	137
5.9	Percentage of instances that belong to each willingness category on average when $\rho_{NAC} \in [0.5, 0.75]$	137
5.10	Percentage of instances that belong to each willingness category on average when $\rho_{NAC} \in [0.75, 0.1]$	138
5.11	Percentage of instances that belong to each willingness category on average when the number of goals is 10	139
5.12	Percentage of instances that belong to each willingness category on average when the number of goals is 50	139

5.13	Percentage of instances that belong to each willingness category on average when the number of goals is 100	140
5.14	Percentage of instances that belong to each willingness category on average when the number of explanatory relationships is 10	141
5.15	Percentage of instances that belong to each willingness category on average when the number of explanatory relationships is 20	141
5.16	Percentage of instances that belong to each willingness category on average when the number of explanatory relationships is 40	142
6.1	Norm-based expansion for constitutive norms in the n-BDI architecture	147
6.2	MCC with respect to the internalization threshold	157
6.3	MCC with respect to the observation threshold	157
7.1	Coherence for normative reasoning	168
7.2	Coherence graph of the case study	179
8.1	Example of a grid that models a building in flames	182
9.1	The Jason reasoning cycle	217
9.2	Norm activation	227
9.3	Observation of behaviours	227
9.4	Norm expiration	227
9.5	Messages exchanged in Cardoso & Oliveira' approach when one norm is controlled	232
9.6	Messages exchanged in Modgil et al. approach when one norm is controlled	234
9.7	Messages exchanged in MaNEA when one norm is controlled	237
9.8	Performance of MaNEA, Cardoso & Oliveira and Modgil et al. frameworks with respect to the number of iterations	240
9.9	Performance of MaNEA, Cardoso & Oliveira and Modgil et al. frameworks with respect to the number of actions	241
9.10	Performance of MaNEA, Cardoso & Oliveira and Modgil et al. frameworks with respect to the number of norms	242
9.11	Performance of MaNEA, Cardoso & Oliveira and Modgil et al. frameworks with respect to the number of instantiations	243
9.12	Performance of MaNEA, Cardoso & Oliveira and Modgil et al. frameworks with respect to the number of agents	244
9.13	Performance of MaNEA, Cardoso & Oliveira and Modgil et al. frameworks with respect to the number of roles	245

List of Tables

2.1	Levels in the development of NMAS	20
2.2	Comparison among languages for specifying norms	33
2.3	Summary of proposals on norm-autonomous agents	53
4.1	Operational rules of the NAC Language	95
4.2	Operational rules of the NCC Language	100
4.3	Parameters used in the norm recognition experiment	106
4.4	95% confidence interval for the relative error made by agents .	107
5.1	Parameters used in the simulations	128
6.1	Parameters used in the norm expansion experiment	153
6.2	95% confidence interval for the Sensitivity, the Specificity and the MCC achieved for each type of agent.	158
8.1	95% confidence interval for the victim survival percentage, the fireman survival percentage and the success that each imple- mentation achieves in all the simulations.	190
8.2	95% confidence interval for the victim survival percent- age, the fireman survival percentage and the success when <code>riskThreshold</code> and <code>internalizationThreshold</code> vary within the [0,0.33) interval.	192
8.3	95% confidence interval for the victim survival percent- age, the fireman survival percentage and the success when <code>riskThreshold</code> and <code>internalizationThreshold</code> vary within the [0.33,0.66) interval.	192
8.4	95% confidence interval for the victim survival percent- age, the fireman survival percentage and the success when <code>riskThreshold</code> and <code>internalizationThreshold</code> vary within the [0.66,1] interval.	193

9.1	Summary of distributed proposals on infrastructural enforcement	200
9.2	Parameters used in the experiments	239