# Modelling the Impact of Role Specialisation on Cooperative Behaviour in Historic Trader Scenarios

Christopher K. Frantz[1], Martin K. Purvis[1], Bastin Tony Roy Savarimuthu[1], and Mariusz Nowostawski[2]

[1] Department of Information Science, University of Otago, New Zealand
{christopher.frantz,martin.purvis,tony.savarimuthu}@otago.ac.nz
[2] Faculty of Computer Science and Media Technology, Gjøvik University College, Norway
mariusz.nowostawski@hig.no

**Abstract.** We analyse two well-established historic trader scenarios from the area of comparative economics known as the Maghribi Traders Coalition and the contemporary Genoese traders, which contrast the otherwise comparable individualistic Genoese and collectivistic North-African trader societies by the institutions they used to sustain cooperative behaviour. We employ agent-based modelling to test a previously unexplored aspect, namely whether a unified role structure (unifying the contrasting investor and merchant perspectives – something that could have characterised one of the two communities in question, the Maghribis) could have been a contributing factor to sustain cooperation for the collective group of Maghribi Traders. To model the emerging institutions, we utilise a continuous notion of deontics that supports the adoption of norms from an experiential perspective. Our simulation results support the idea that experiencing economic transactions from different perspectives increases the convergence performance towards stable behaviour, and supports the enforcement of cooperation by informal means, such as norms, based on their stronger normative alignment.

**Keywords:** Institutions, Role Specialisation, Maghribi Traders Coalition, Genoese Traders, Norms, Rules, Dynamic Deontics, Social Simulation, Multi-Agent Systems

## 1 Introduction

In recent decades the importance of institutions as a fundamental element to determine economic success has gained strong reflection in economic literature. Notable works include North's seminal work [13], but also more recent efforts such as Robinson and Acemoglu's [1] and Greif's [9].

Recent achievements modelling norm emergence using multi-agent systems ([18], [2], [11]) demonstrate the suitability to represent those subtle social coordination mechanisms in silico. We build on that effort and show how perspective-taking can drive differing norm understandings without prescribing specific norms ex ante.

To realise this, we capitalise on the previously introduced Dynamic Deontics [6] that relax the otherwise rigid categorisation of prescriptions into *may*'s, *must*'s and *must not*'s.

In the next section (Section 2), we introduce Dynamic Deontics in detail, before presenting the trader scenarios and their historical context (Section 3) in order to derive a simulation model that allows us to generate behavioural norms based on reinforcement learning (Section 4). The final Section 5 discusses the simulation results.

## 2 Dynamic Deontics

A conceptual foundation of this work is the notion of Dynamic Deontics introduced by previous work [6]. Conventionally, the notion of discrete, interdefinable deontics (often represented using the deontic primitives *must*, *must not*, and *may*) based on deontic logic [19] is appealing and offers a clear interpretation of associated prescriptions. However, to model the emergence of norms and institutions (especially when we cannot make presumptions about pre-existing norms) as well as the dynamics associated with this, we use a continuous notion of deontics. This enables agents to operate along a *deontic range* spanning from the extreme of proscriptions (or prohibitions) via permissions to prescriptions (or obligations), as illustrated in Figure 1.

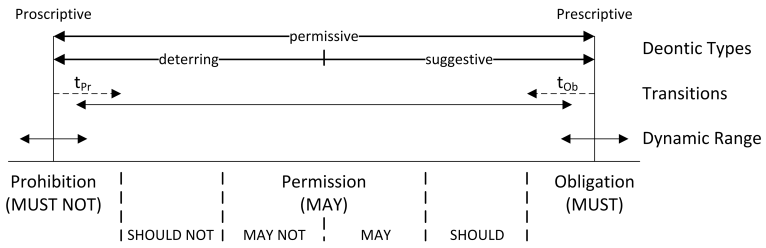



Fig. 1: Dynamic Deontic Range

Note that this approach is not related to dynamic deontic logic [12], which attempts to resolve the ambiguities of 'ought-to-do' and 'ought-to-be' in standard deontic logic. Our approach does not entail a refinement of deontic logic, but uses the term 'dynamic' to describe the expanding and contracting boundaries of the deontic range.

In the remainder of this section we will briefly explore the central characteristics of Dynamic Deontics and discuss possible operationalisation approaches.

### 2.1 Characteristics

**Continuous Notion of Deontics** – Concrete deontics associated with a given action or outcome can be allocated along the deontic range and shift continuously, including moderate movements along the range, varying across short time frames. However, norms (and institutions in general), can potentially likewise shift from one deontic extreme to the other. An example of this is the societal attitude towards homosexuality, which in the past decades experienced a considerable shift from (former) legal prohibition towards increasing acceptance. A similar example is the societal attitude towards spanking of children as a punitive measure. To make the continuous notion more compatible with customary linguistic usage, we can allocate terms along this range to express

the varying extent of normative prescriptiveness. For this work here, we operationalise common norm understandings and employ the terms *must not*, *should not*, *may not*, *may*, *should* and *must*, which we allocate in *deontic compartments* of equal size along the deontic range. However, the choice of deontic terms and their number is flexible. Likewise the range and compartment may not be necessarily conceived symmetric respectively equally sized. However, in this context, we concentrate on the core idea of compartmentalising different institutions in order to simplify their interpretation.

**Stability** – Adopting a continuous understanding implicitly suggests the situational shift of norms. However, a core characteristic of institutions (and thus norms) is their stability [15]. Once successfully established, they exhibit 'stickiness' and change resistance which opens up an arena for potential norm violations. If a model exhibits stability properties itself, we can observe the emerging stabilisation towards the outer extremes without explicit operationalisation (*Modelling Variant 1*). However, depending on the model objectives and underlying assumptions about institutional change (see [10]) this characteristic can be represented using the metaphor of hysteresis. Stability can be modelled using discrete tolerance zones around the deontic extremes (denoted as $t_{Pr}$ and $t_{Ob}$ in Figure 1), in which institutions, if penetrating those extremal deontic compartments for sufficiently long time, become engrained and stable (*Modelling Variant 2*). If in such state, institutions likely require strong reinforcement to give up this stability and shift back into an adaptive state, which allows their resumed movement along the deontic scale. Translating this into simulation models, it can be operationalised by counting simulation rounds for which a particular statement remains in the extremal ends of the deontic range, with thresholds for their establishment and dissolution. An alternative, continuous operationalisation could establish increasing levels of friction along the range towards the deontic extremes (*Modelling Variant 3*), avoiding the discrete tolerance zones at the outer ends.

**Dynamic Deontic Range** – A final important aspect is the dynamic nature of the deontic range. Individuals experience the world subjectively and absorb feedback in a varying fashion. As an example imagine individuals moving in culturally diverse environments in which different, potentially conflicting, influences coexist. An individual might be inclined to incorporate those different aspects and develop a wide deontic range with greater degrees of tolerance ('openness'), i.e. a wider range between his deontic extremes, his no-go's. This can be seen in contrast to individuals that adopted existing rigid political or cultural rulesets, and act within rather narrow boundaries of permissiveness, but in case of uncertainty quickly resort to stable internalised rules. We thus suggest that the deontic range changes throughout an agent's lifetime, widening with experience, but also narrowing if experiential stimuli cease.

### 2.2 Operationalisation

The Dynamic Deontics concept itself does not prescribe a specific operationalisation, but for our purpose, here of modelling the establishment of institutions (and norms in particular), we adopt an experiential perspective and do not rely on predefined norms.

For this paper we operationalise those using reinforcement learning (RL, specifically Q-Learning [20]), and use the mean of a sliding window across the highest and

lowest Q-values as the deontic range boundaries. The middle point of the scale (normative centre) is the mean of the upper and lower boundary values, which depend on the situational Q-values. The discounting characteristics associated with Q-Learning reflect the dynamic adaptation, i.e. expansion or retraction, of the deontic range over time. The operationalisation of stability follows the *Modelling Variant 2* introduced in Subsection 2.1; stability characteristics are represented by the time range of Q-values within the tolerance zones around the deontic extremes (e.g. number of rounds), the values of which are specified as part of the simulation parameter set.

Doing so, at any time of simulation runtime, an agent's Q-values can be resolved to deontic terms associated with the respective compartment along the deontic scale, such as *may not*, *should not* or *must not*.

Even if the norm assignment to compartments across the deontic range is not accurate, the intuition of the individualised norm understanding and its varying strength is retraceable. This categorisation simplifies the interpretation of the differentiated norm understandings, especially given that the core interest is not centred around a precisely accurate representation of what an agent 'thinks', but instead to provide a situational understanding of the overall normative landscape.

Important to note at this stage is that the chosen operationalisation adopts a consequentialist perspective in opposition to the traditional deontic perspective, in which individuals evaluate their norm compliance behaviour based on given norms or rules. In this case, individuals need to learn which behaviour provides them with the best outcome, shaping behavioural norms from experience. In this context the role of the Dynamic Deontics is not to prescribe rules. Instead this operationalisation extracts deontic values from the existing RL instances maintained by each agent in order to derive the agents' understanding of normative behaviour.

## 3 Historic Trader Scenarios

We use Dynamic Deontics for the exploration of scenarios from the area of comparative economics. A core topic in this area is to analyse the impact of institutions on the development of economies, thus asking the question why some, often more closed, societies could rely on informal mechanisms to assure cooperation, while societies made up of weaker social ties had to rely on legal instruments, i.e. formal institutions, to bind agents to their commitments.

An interesting example is Avner Greif's comparison of long-distance trading in 12th century Genoa, which is considered an important early historical example for the systematic use of formal institutional mechanisms. He contrasts this with what he called the *Maghribi Traders Coalition*, a contemporary homogeneous group of traders that were unified by their cultural background and belief, and operated along the North African coast.

Greif's work [9] combines rational choice theory with game-theoretical analysis to show that Maghribi traders could sustain cooperation, among other aspects, based on the high cost that was associated with non-cooperation. As a central characteristic of their group, Greif elaborates on the information transmission mechanisms employed by Maghribi traders. They maintained communicative ties by frequent exchange of letters

among associates by means of which they shared market information and coordinated agency services for each other. However, they used such medium not only to manage their business operations (which usually extended to remote ports across the Southern part of the Mediterranean basin), but likewise to share information about fellow traders, suggesting a fast spread of information,[3] should a trader attempt to misreport profits when acting as an agent for a remote associate.

Maghribis, named by their geographical descent ('West') within the Arabic world , were in fact traders from Jewish communities in nowadays Tunisia that were united by their operation in a culturally contrasting Islamic environment, making it hard for outsiders to enter their group. Accordingly, likewise high exit costs were associated with defecting from cooperation.

Genoa, on the other hand, was different in its structure. It operated in opposition to other influential city states, such as Venice, and relied on a constant influx of foreigners to sustain its development. Consequently, binding features were limited, which facilitated easy defection from business commitment, given the limited effect of informal enforcement.[4] Consequently Genoa had to rely on formal institutions, such as commercial courts and associated legal instruments, to sustain cooperation. Trade operations in Genoa's open society thus neatly contrast with the kinds of interactions in the closed Maghribian trader community.

For our simulation model, we adopt the comparative nature of the scenario, but concentrate our focus on an aspect that has been mentioned in historical commentaries but not explored in previous analyses: individuals involved in Genoese long-distance trade were stratified into different roles, a characteristic that is reflected in the dominantly used institutional mechanism, the 'commenda', namely

- investors ('commendatores') that supplied funds and goods for travels to remote trade locations, and
- actual merchants ('tractatores') that ran the actual operation, thus bearing the laborious share of the agreement.[5]

Given Genoa's central role as trading port, long-distance trade was seen as an investment opportunity that attracted rich citizens as well as foreigners, who often did not have any trade experience themselves or ceded that part of their enterprise to a third party (see van Doosselaere [17] for an overview of the structure in commenda relationships). The actual merchants, however, were often opportunists themselves, or workless artisans that saw the adventure of long-distance trade as a promising temporary job opportunity.

The Maghribi traders, in contrast had cultivated a rigorous apprenticeship system, in which young aspiring traders operated under the supervision of an often unrelated experienced trader. In this process apprentices would be increasingly embedded in the trade operations (and information transmission aspects) and so could eventually establish themselves as full traders – a process which could last more than a decade [8].

[3] Goldberg [7] allocates the fraction of communication dedicated to such gossip at around 20 percent.

[4] At that time Genoa had more than 30,000 citizens [9].

[5] In this text we use 'trader' to capture both roles and use 'investor' and 'merchant' to address the respective specialised roles.

A second characteristic was the unification of the investor and merchant roles. While senior traders tended to concentrate on the investment aspect, as part of the trader coalition, the reciprocity-based informal rule system still required them to process agent services for other traders (or at least store their goods at no expense). So the clear role differentiation as found with the Genoese did not exist in the Maghribi case.

Based on the available information we hypothesise that, notwithstanding the core differences between open vs. closed societies, the role stratification in the Genoese trader community and their unified character in the Maghribian case could have been an important difference that might have driven cooperation based on informal mechanisms. We postulate that Maghribi cooperation was largely facilitated by the mutual interest to sanction violators, and more so, by the desire of the potential violators not to be detected, knowing that they themselves, when acting as an 'investor', could be cheated if delegating their goods-handling to fellow traders. So even in the attempt of cheating, they still had an incentive to sustain cooperation to suppress cheating by others. In the Genoese trader community, opportunistic merchants could not expect to undertake multiple journeys with the same investor, and were, unless affiliated with a family firm, hardly ever in the position to take up the investor role. Given this role separation, merchants did not have any incentive to avoid non-cooperative behaviour unless he could exercise control by formal means or private-order enforcement (e.g. retaliation against family members).

## 4 Model

To test the hypothesis that the Maghribis' integrated role understanding could have been fundamental to drive cooperative behaviour, we have developed an agent-based simulation model that captures the essence of the aforementioned scenarios. We start with a basic scenario that employs the commission-based trading metaphor. During each round traders can randomly choose fellow traders to whom they wish to send goods and expect profit in return. Instead of modelling the entire trade interaction in detail, we concentrate on the essential decisions, namely whether goods-receiving agents cooperate or withhold profits. As part of this scenario, each round an agent (Investor 'Inv') chooses a random trade partner (Merchant 'Mer') before sending him goods. The receiver can then decide whether to trade fair and return realized profits, or to cheat, and withhold profits. The investing party (Inv) then reacts to the merchant's behaviour using the reaction he considers suitable based on his experience. For this purpose Inv has selected reactions at its disposal. For given actions with corresponding reactions, we specify the effects in terms of payoffs for individual action-reaction combinations as shown in Table 1. To memorise the respective feedback, agents use reinforcement learning from which we can derive their respective norm understanding using the Dynamic Deontics operationalisation introduced in Section 2.

Central to this is the integrated nature of the memory structure (used to internalise feedback from actions and reactions) and the operationalisation of Dynamic Deontics as part of the simulation infrastructure (see also [6]). To represent the experiential aspect, reinforcement is associated with action-reaction pairs (e.g. [TRADE FAIR, PAY COMMISSION]), since feedback information entails the combination of action and reaction,

Table 1: Action Reaction Feedback Combinations

| Action-Reaction Combinations | | Utility from Actions | |
|---|---|---|---|
| **Action (Mer)** | **Reaction (Inv)** | **for Mer** | **for Inv** |
| TRADE FAIR | FIRE | -2 | -1 |
| TRADE FAIR | RETALIATE FAMILY | -3 | -1 |
| TRADE FAIR | PAY COMMISSION | 1 | 1 |
| WITHHOLD PROFIT | FIRE | -1 | 0 |
| WITHHOLD PROFIT | RETALIATE FAMILY | -3 | 1 |
| WITHHOLD PROFIT | PAY COMMISSION | 2 | -2 |

independent of whether they act as investors or merchants. Further, choosing action-reaction combinations allows the use of this memory structure independent of the role stratification, i.e. the RL instance can be used to store experience from a merchant's perspective ('What reaction followed my action?') as well as investor's perspective ('What reaction did I choose to address a given action?'). We use the integrated RL memory instance as a mechanism to unify all memory entries by the action the statement describes. As a consequence, for each action, the deontic associated with that action needs to be derived from all potential consequences (i.e. reactions) an individual has faced, e.g. the action 'TRADE FAIR' may have been usually reciprocated with 'PAY COMMISSION', but potentially also with 'FIRE' and 'RETALIATE FAMILY' at different times.

The intuitive approach to derive the deontic to be associated with the given action is to calculate the sum of all individual Q-values. This offers an integrated picture of the individual's experience. However, as actions hardly co-occur, this representation may not be faithful to the individual's perception and neither reflect an individual's fear of uncertainty, a central driver for the establishment of institutions. Instead of adding the Q-values, we thus choose the most extreme Q-value, representing an individual's expected greatest gain or pain.

To operationalise this, we derive the deontic from the Q-value with the greatest deviation (extremal) from the centre of the deontic range ($c_{deonticRange}$) towards the direction indicated by the sum of all Q-values (*deontic bias*). With *stmt* representing individual statements and $d(stmt_{l,i})$ as the deontic value for the $i$th statement on nesting level $l$, we can say

$$extremeDeontic(stmt_l) := \begin{cases} max(d(stmt_{(l+1),0}), \ldots, d(stmt_{(l+1),count_{(l+1)}})), & \text{if } (\sum_{i=0}^{count_{(l+1)}} d(stmt_{(l+1),i})) \\ & > c_{deonticRange} \\ min(d(stmt_{(l+1),0}), \ldots, d(stmt_{(l+1),count_{(l+1)}})), & \text{otherwise} \end{cases}$$

The extreme deontic is applied unless the sum of the Q-values is located at the deontic range centre $c_{deonticRange}$, in which case the Q-values associated with action-reaction pairs cancel each other out. In that case, the deontic range centre itself describes the action's deontic (which, under the assumption of a symmetric deontic range, resolves to *may*), i.e.

$$d(stmt_l) := \begin{cases} c_{deonticRange}, & \text{if } (\sum_{i=0}^{count_{(l+1)}} d(stmt_{(l+1),i})) = c_{deonticRange} \\ extremeDeontic(stmt_l), & \text{otherwise} \end{cases}$$

To illustrate our mapping from RL to the deontic range value associated with an action, we show in Figure 2 how deontic terms are derived from a situational deontic range for a given agent ranging from around -30 to 20.1. Based on the deontic range and the reinforcement values, the figure displays the different Q-values associated with various reactions (e.g. retaliate against family, pay commission, fire) grouped by the action 'WITHHOLD PROFIT' (represented in the nADICO syntax [5]) and derives the action's deontic term using the aforementioned principle.

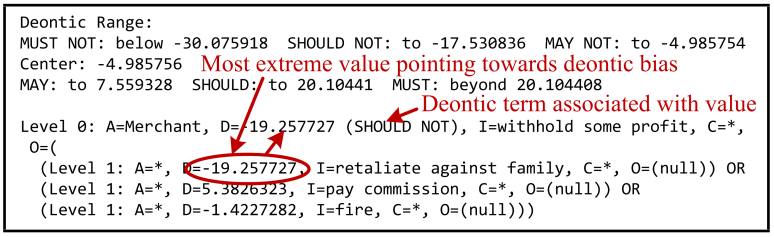



Fig. 2: Example for Deriving Deontic Term from Situational Deontic Range

Before discussing the entire agent execution cycle, we briefly discuss the intuitions associated with the value choices (shown in Table 1): Being fired after trading fair has a negative impact on the merchant who operated truthfully. For the sanctioning investor, however, this is likewise of negative impact, given that he sanctions a compliant merchant. Similarly, retaliation against a compliant merchant's family is counterintuitive. Private-order enforcement of contractual obligations was very well present in medieval Genoa [17], given that pursuing the legal track was cumbersome and time-consuming [4], making private-order enforcement against cheating merchants a realistic option. Paying the commission to a compliant merchant is considered the regular outcome if cooperation should be sustained. The lower payoffs associated with this imply that both parties had the general expectation that the commitments associated with their trade interaction (i.e. being paid for fair trading) would be honoured. For the negative case of a non-compliant merchant, payoffs are amended. Withholding profit and being fired as a consequence has mild negative feedback for the merchant (who indeed cheated) and neutral feedback for the investor (who identified and fired a cheater). Retaliation against family has a strongly negative feedback, as it possibly is the greatest threat associated with non-compliance. However, given the elicited satisfaction for the sanctioner, we associate a mild positive feedback for this reaction.[6] As a final aspect, paying commission to an unloyal merchant has a negative effect for the investor, and is highly rewarding for the cheater. Note that this work is based on historical scenarios, which constrains an authentic representation. Nevertheless, instead of putting the

[6] Neuro-scientific findings [3] support the idea that performing punishments can elicit feelings of reward, especially if they are considered 'deserved'.

emphasis on precision, we rather seek to improve the understanding of an otherwise unexplored aspect of the scenarios based on available information.

Given this overview on the infrastructural aspects, we can return to the discussion of the scenario. Our model of the trader scenario allows the representation of a characteristic that sets apart different society types, using Simmel's social circles [16] as a metaphor. Following this understanding, in more homogeneous societies (in line with North, Wallis and Weingast's primitive societies [14]) we can find a lower extent of role specialisations. Thus roles in such societies are of more general nature, allowing members to develop a more unified understanding of roles and overlapping social circles, rather than a differentiated and stratified role experience (which we postulate for the more individualistic Genoese society). This drives our hypothesis that a more integrated role understanding of traders in the Maghribi society (i.e. taking the perspective of both investor and merchant at different times) could have been a contributor to the more compliant behaviour without need for formal institutions.

The basic execution cycle is shown in Algorithm 1. It does not differentiate between different roles for investors and merchants. We thus interpret it as a representation of the Maghribi trading behaviour (denoted as 'Maghribi version'). Note that we include the choice to activate norm enforcement. If choosing to exploit, an agent chooses an action based on the Q-values associated with it. In this context norm enforcement refers to the sanctioning of other merchants' actions, using the memorised action-reaction combination associated with the highest Q-value, or if not existent, a randomly chosen reaction. The Genoese variant of the algorithm (see Algorithm 2) introduces the role specialisation discussed in Section 3. Agents are thus instantiated as either investors or merchants and take actions only related to their respective role. If acting as merchants, they engage in exploration and exploitation of actions (with a bias towards exploitation, see Table 2). If norm enforcement is activated, investors can sanction unrelated merchants' behaviour in addition to applying sanctions according to their action-reaction Q-values.

We test both scenarios using the same parameter set shown in Table 2. The different scenarios sketched here allow us to specify four possible configurations:

- Scenario 1 – Role Unification w/o Norm Enforcement
- Scenario 2 – Role Unification with Norm Enforcement
- Scenario 3 – Role Specialisation w/o Norm Enforcement
- Scenario 4 – Role Specialisation with Norm Enforcement

We ran each scenario for 20,000 rounds. The high number of rounds was chosen to allow the stabilisation of changing norm understandings in the given simulation. The simulation outcomes are discussed in the following section.

## 5 Simulation Results

As explained previously, throughout the simulation runtime agents develop a normative understanding of the different actions aligned with the deontic compartment in their respective deontic range (e.g. *must not*, *should not*, *may not*, *may*, *should*, *must*, etc.). We can thus show the progression in the developing norm understanding using time-series

**Algorithm 1:** Agent Execution Cycle – Maghribi version

```
Decide whether to explore or exploit in this round;
if exploring then
    Pick random action from action pool;
else
    Pick action with highest Q-value from action pool;
    if norm enforcement activated then
        Sanction action taken by randomly chosen agent using sanction with highest
            Q-value;
        Memorize feedback from sanction choice;
end
Execute picked action and apply to randomly chosen agent;
Memorize reaction and make action-reaction combination (with valence representation of
    feedback) visible to other agents;
Update deontic range;
Check Q-values for stability (shifts from/to obligation or prohibition norms);
Apply discount factor to all memory entries;
```

**Algorithm 2:** Agent Execution Cycle – Genoese version

```
During setup: Assign either investor or merchant role;
Decide whether to explore or exploit in this round;
if exploring then
    if is merchant then
        Pick random action from action pool;
else
    if is merchant then
        Pick action with highest Q-value from action pool;
    if is investor & norm enforcement activated then
        Sanction action taken by randomly chosen agent using sanction with highest
            Q-value;
        Memorize feedback from sanction choice;
end
if is merchant then
    Execute picked action and apply to randomly chosen agent;
    Memorize reaction and make action-reaction combination (with valence
        representation of feedback) visible to other agents;
Update deontic range;
Check Q-values for stability (shifts from/to obligation or prohibition norms);
Apply discount factor to all memory entries;
```

diagrams in which the different understandings for a particular action accumulate to 100 percent (i.e. each agent has a normative attitude towards an action). Combining both actions in one diagram thus provides us with a macro-view of the normative landscape. Given our interest in the developing normative understanding, we concentrate on this aspect in our analysis. Given the vast number of possible combinations of action and

Table 2: Simulation Parameters

| **Parameter** | **Value** |
|---|---|
| Number of agents | 100 |
| Tolerance zone around extreme deontics ($t_{Pr}$, $t_{Ob}$) | 0.05 of deontic range amplitude |
| Norm establishment threshold | 100 rounds |
| Norm destruction threshold | 200 rounds |
| Deontic range history length | 100 rounds |
| Memory discount factor | 0.99 |
| Exploration probability | 0.1 |

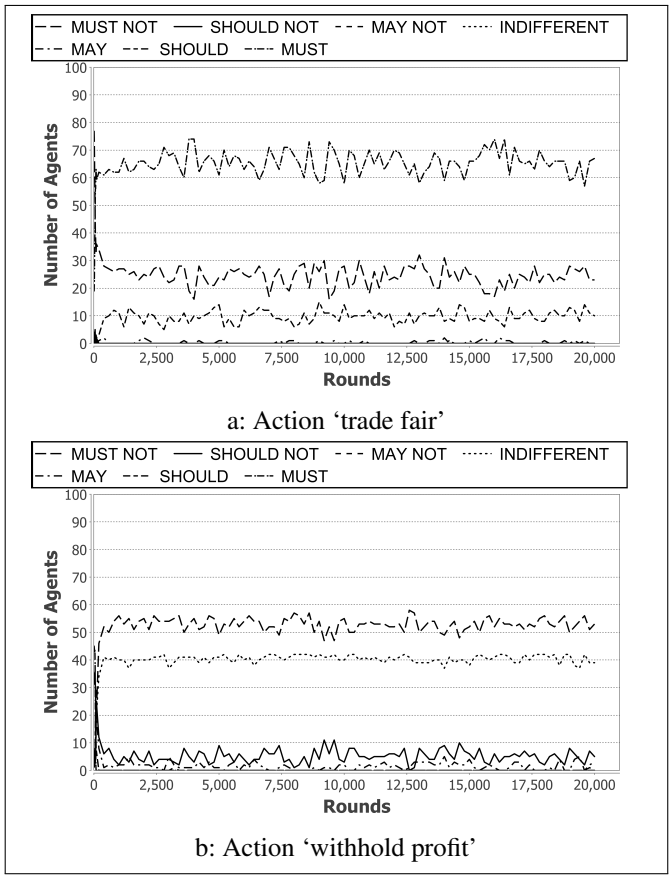
deontics, we highlight the essential findings for each scenario. To do this, we show a representative simulation run for each scenario and interpret the displayed dynamics.

### 5.1 Role Unification without Norm Enforcement

For the first scenario (Maghribi-like), individuals adopt both roles, investor and merchant, throughout the simulation runtime, but do not engage in norm enforcement (i.e. sanctioning of merchants in observed trade interaction with another investor). Instead, agents operate purely based on experiential learning from feedback they receive for chosen actions (and the reaction chosen by their counterpart).

Looking at the simulation results for this configuration (Figure 3), we can observe that most agents quickly converge to the understanding that they can act selfishly and cheat repeatedly. They mostly act in a compliant manner, which is driven by the integrated roles in which they act. If situationally acting as merchant, cheating is a beneficial option. When acting as an investor, in contrast, cheating is not desirable. However, as investors they can likewise exploit their agent, e.g. by firing him despite compliant behaviour. But by integrating the different perspectives, over time up to 70 percent of all agents converge to the understanding that they *must* trade fair, mirrored by around 20 to 30 percent that think they *must not* trade fair. The remainder (less than 10 percent) believe they *should* trade fair.

It is important to understand that both actions cannot simply be assumed complementary and mirror each other. Firstly, the evaluation relies on the reinforcement (i.e. continuous experience) of the different actions in combination with reactions chosen by the counterpart, which may vary for different actions. Secondly, the norm understanding provided here is derived from the Q-values of individual agents, but that does not necessarily reflect their situational choice as the choice of actions is based on the individual Q-values, not aggregated ones from which we derived the overall perspective. This way agents can maintain in principle conflicting norms (e.g. based on negative reinforcement for individual actions), but solely base their choice on the highest Q-value, which allows them to overrule the extracted normative understanding.



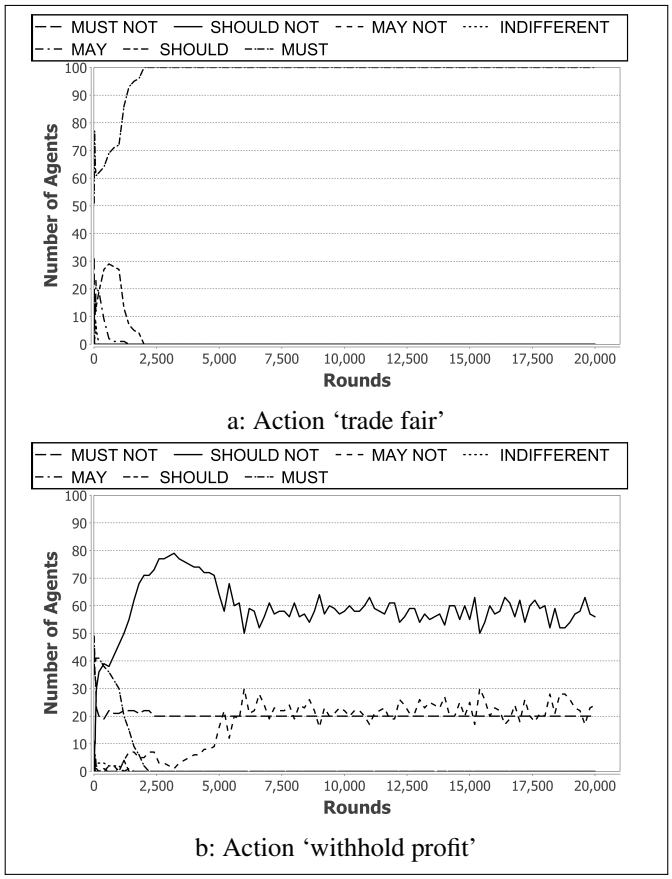

a: Action 'trade fair'

b: Action 'withhold profit'

Fig. 3: Role Unification, No Norm Enforcement

### 5.2 Role Unification with Norm Enforcement

Another outcome can be observed when including norm enforcement (Figure 4) in the Maghribian scenario. This configuration is the closest match to the institutional setup in the real Maghribian society. Individuals acted in role unity and are aware of constant norm enforcement (see Section 3). In this simulation model norm enforcement introduces a bias towards the investor role. Agents judge other agent's behaviour from the investor perspective, i.e. interpret it as if they had been subject of that action, and reinforce their reaction choice. As a consequence of this, agents acting as situational merchants need to expect multiple reactions to their action, an aspect that considers the uncertainty about both occurrence of consequences ('Will I be sanctioned?') and chosen measure ('What will the sanction be?') associated with norms, as opposed to precisely prescribed consequences in the context of laws or rules. The result of this shift (Figure 4) is a societal perspective on compliant behaviour. After initial low measures for compliant behaviour (*must* trade fair), the norm enforcement (once sufficiently




a: Action ‘trade fair’

b: Action ‘withhold profit’

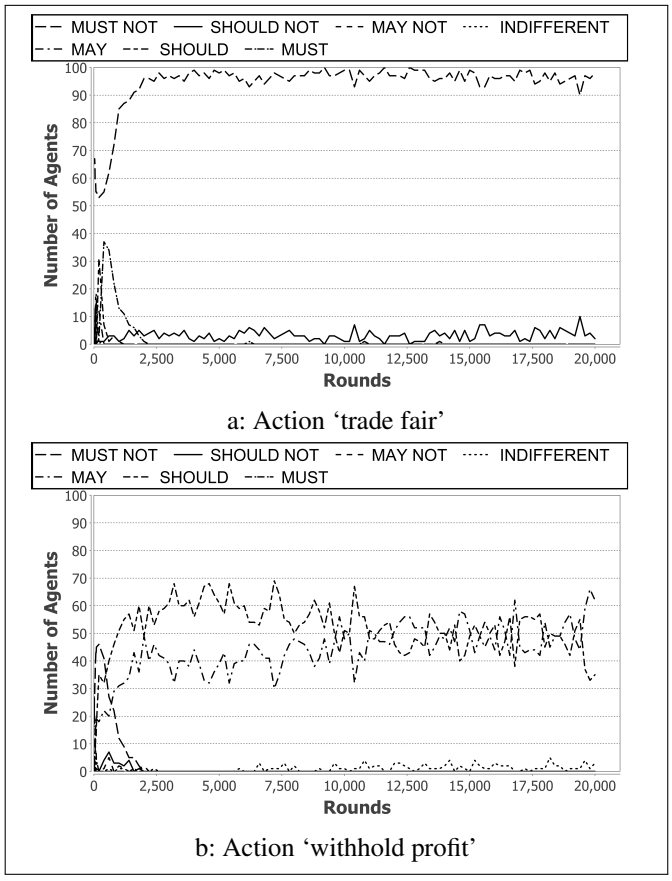
Fig. 4: Role Unification, Norm Enforcement

explored and settled in individual agents) leads to a fully cooperative trader society. In parallel, agents adjust their understanding of withholding profit and arrive at a majority of agents that think they *should not* cheat (around 60 percent). Complementing this, a stable fraction of 20 percent persist that they *must not* cheat. The reason for the lower convergence towards extreme values is the lesser reinforcement of the action ‘withhold profit’, because agents more strongly reinforce fair trading as opposed to cheating. This aspect is an artefact caused by the operationalisation using the discounting mechanism of reinforcement learning.

### 5.3 Role Specialisation without Norm Enforcement

Introducing role specialisation requires further considerations in order to maintain comparability of simulation results. In the Maghribi case each agent could act as investor and merchant, enabling each individual to act as a merchant (and thus each trade compliantly or cheat). Simply separating the roles would render us with 50 active merchants

and 50 purely reactive investors as opposed to the Maghribian case where each individual could act as a merchant. To reflect the effect of role stratification and establish comparable outcomes, for the role-specialised Genoese scenario we double the number of agents to maintain the same number of acting merchants. All other parameters remain unchanged, and so we increase the number of agents to 200 for all remaining simulations.




a: Action 'trade fair'

b: Action 'withhold profit'

Fig. 5: Role Specialisation, No Norm Enforcement

Analysing the simulation outcomes for this configuration (Figure 5), we observe that traders nearly fully converge to the understanding that they *must not* trade fair, framed with around 5 percent that retain a weaker normative understanding and believe that they *should not* trade fair. Regarding the action 'withhold profit', agents provide a more divided view; agents are largely equally divided (but shifting over time) between *may* and *should* withhold profit. The general trend points towards a stronger dominance of the weaker *may* withhold profit.

The scenario described here is closer related to the actual, historical Genoese society. We can clearly see that role specialisation could not have sustained cooperative behaviour without the introduction of formal mechanisms that afforded legal commitments of participating individuals. Given their individualised roles, individuals would never be able to perform the perspective taking as done implicitly in the context of role unification (i.e. both roles fulfilled by same trader at different times). But given the exploratory possibilities of our simulation setup, we complement our simulation runs by exploring how norm enforcement would impact our modelled Genoese case.

### 5.4 Role Specialisation with Norm Enforcement

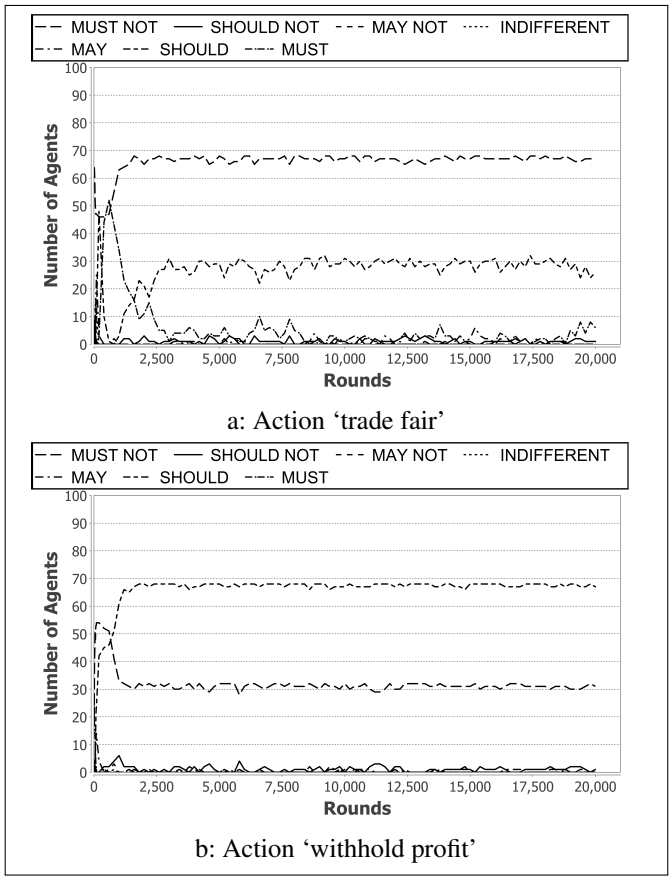
The activation of norm enforcement in addition to role specialisation adds an artificial aspect in the sense that it ignores the fact that Genoa was an open society, in which constant influx of new merchants and investors limited the effect of society-internal normative enforcement (although private-order enforcement is indeed documented as sanctioning mechanism [9]). However, in the scenarios presented here, the number of agents is constant. Neither do we model trader generations nor an open society. However, introducing norm enforcement allows us to explore the hypothetical case of norm enforcement in a closed society with role specialisation.

The results for this configuration (Figure 6) show that norm enforcement by investors could indeed have an impact on the normative attitude of merchants towards compliant trading. Around 30 percent of merchants share the view that merchants *should* trade fair. The dominant attitude, however, remains that merchants *must not* trade fair (nearly 70 percent). In this case, this distribution is contrasted with 70 percent of traders who think they *should* withhold profit and 30 percent maintaining that they *must not* withhold profit as a reaction to norm enforcement they experienced.

It is important to stress again, that this scenario ignores further characteristics of that society, but the isolated perspective on role stratification supports the presumption that it could have played an important role in preventing the society from maintaining cooperative behaviour based on informal means.

## 6 Discussion, Conclusions and Outlook

This work addresses important scenarios from the area of comparative economics, the Maghribi Traders Coalition and its Genoese counterpart, both of which are some of the earliest well-documented historical examples for long-distance trade by in/formal means. Particular focus lies on a specific previously documented but unexplored aspect, namely the question whether the role specialisation in the Genoese society could have made the difference in driving the society towards stronger reliance on formal institutional mechanisms to assure compliance. Their historical counterpart, a North African trader collective – called 'Maghribis' – could maintain cooperation based on informal means but shared a unified role understanding. We model and explore differing outcomes for the characteristics of role specialisation and norm enforcement in otherwise unchanged scenarios. Our findings support the hypothesis that the normative understanding of the individualistic Genoese society, at least in part, drifted apart over time based on the specialisation of individuals.




a: Action ‘trade fair’

b: Action ‘withhold profit’

Fig. 6: Role Specialisation, Norm Enforcement

The experiments described here bear further interesting findings: Looking at the model, even in the informally regulated society norm enforcement remains the important driver for fully compliant behaviour. However, even without norm enforcement, around 70 to 80 percent (*must* and *should* trade fair) of traders act compliantly. Norm enforcement in the Maghribi society initially produces diverse compliance levels (nearly 100 percent for *may*, *should* and *must*), which fully converge to the prescription (*must*) to trade fair. For role specialisation we cannot observe such behaviour. The individualistic specialised perspective drives selfish behaviour. The hypothetical case of introducing norm enforcement drives a more diverse understanding with a significant minority of around 30 percent internalising the understanding that compliant behaviour is desirable (*should*). This leaves to suggest that even in specialised societies, normative influence still proves to be supportive for achieving a socially desirable outcome. However, the sketched simulation models an idealised social representation. The simulation scenario focuses on the essential representation of the social features of interest, but

omits specific societal characteristics (open vs. closed society) and the consideration of possible psychological components. This includes a limitation to the fixed representation of utilities as well the lacking consideration of situational, rather than randomised, choice of compliance behaviour. This is hardened by the challenge to find more grounded data on the historic societies. But even though this analysis specifically focuses on the Maghribi Trader Coalition and their Genoese counterpart, the results bear general value in that they support the hypothesis that role specialisation in combination with the assumption of the selfish individual potentially supports antisocial behaviour, insofar as individuals do not have mutual awareness about their individual preferences. This challenges the ability to regulate behaviour in a normative fashion driving increasing formalisation of institutions in open specialised societies.

Beyond the simulation outcomes it is worthwhile to discuss the concept of Dynamic Deontics used to operationalise the simulation model. It is important to reemphasise that the operationalisation showcased here adopts a purely consequentialist perspective and does not explicitly preimpose normative statements or rules, but agents experience feedback from both their actions and their social environment in a greenfield approach. Dynamic Deontics allow us to represent the society's normative understanding on an individual level (see memory sequences shown in Figure 2) as well as collective level (see time-series charts shown in Section 5). By mapping normative understandings onto deontic terms the salience of established norms becomes accessible, which allows us to follow the dynamics in which norms emerge and stabilise.

The operationalisation presented here does not exploit the full capabilities of the Dynamic Deontics concept. Agents can in principle develop independent normative understandings for individual actions (see the operationalisation in Subsection 2.2). However, the current action representation is too simplistic to capture different situational contexts. To allow a more comprehensive application, we intend to introduce a more complex action representation that incorporates context, with the use of statements in the nADICO syntax [5] as a starting point (which is briefly highlighted in Figure 2). Further aspects that require future exploration (and highlighted previously [6]) include the allocation of deontic terms along the deontic scale, but likewise the assumption of symmetry of deontic compartments. Those refinements will naturally rely on empirical input based on user studies to establish the necessary grounding. We are further investigating mechanisms that allow the meaningful aggregation of individual normative understandings beyond the simplified conflation of individual deontic compartments in time-series.

Concluding, we believe that Dynamic Deontics is an intuitively accessible concept that offers the potential to incorporate the representation of different mindsets, such as cultural or social backgrounds (e.g. by different experiences and deontic scale widths), as well as preimposed norms, which may potentially change over time. Moreover, the inclusion of different contexts and experiences may pave the path towards a representation of morality (here: the agent's ability to infer what is 'good' or 'bad' based on the developed deontic scale derived from contextual experience) within individual agents.

## References


1. D. Acemoglu and J. Robinson. *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*. Crown Business, New York (NY), 2012.
2. G. Andrighetto, D. Villatoro, and R. Conte. Norm internalization in artificial societies. *AI Communications*, 23(4):325–339, Dec. 2010.
3. D. J.-F. de Quervain, U. Fischbacher, V. Treyer, M. Schellhammer, U. Schnyder, A. Buck, and E. Fehr. The Neural Basis of Altruistic Punishment. *Science*, 305(5688):1254–1258, 2005.
4. S. A. Epstein. Secrecy and Genoese commercial practices. *Journal of Medieval History*, 20(4):313–325, 1994.
5. C. Frantz, M. K. Purvis, M. Nowostawski, and B. T. R. Savarimuthu. nADICO: A Nested Grammar of Institutions. In G. Boella, E. Elkind, B. T. R. Savarimuthu, F. Dignum, and M. K. Purvis, editors, *PRIMA 2013: Principles and Practice of Multi-Agent Systems*, volume 8291 of *Lecture Notes in Artificial Intelligence*, pages 429–436, Berlin, 2013. Springer.
6. C. Frantz, M. K. Purvis, M. Nowostawski, and B. T. R. Savarimuthu. Modelling Institutions using Dynamic Deontics. In T. Balke, F. Dignum, M. B. van Riemsdijk, and A. K. Chopra, editors, *Coordination, Organizations, Institutions and Norms in Agent Systems IX*, volume 8386 of *Lecture Notes in Artificial Intelligence*, pages 211–233, Berlin, 2014. Springer.
7. J. L. Goldberg. Choosing and Enforcing Business Relationships in the Eleventh Century Mediterranean: Reassessing the 'Maghribī Traders'. *Past & Present*, 216(1):3–40, August 2012.
8. J. L. Goldberg. *Trade and Institutions in the Medieval Mediterranean: The Geniza Merchants and their Business World*. Cambridge University Press, New York (NY), 2012.
9. A. Greif. *Institutions and the Path to the Modern Economy: Lessons from Medieval Trade*. Cambridge University Press, New York (NY), 2006.
10. C. Kingston and G. Caballero. Comparing theories of institutional change. *Journal of Institutional Economics*, 5(2):151–180, 8 2009.
11. S. Mahmoud, N. Griffiths, J. Keppens, and M. Luck. Efficient Norm Emergence through Experiential Dynamic Punishment. In *ECAI'12*, pages 576–581, 2012.
12. J.-J. C. Meyer. A Different Approach to Deontic Logic: Deontic Logic Viewed as a Variant of Dynamic Logic. *Notre Dame Journal of Formal Logic*, 29(1):109–136, 1988.
13. D. C. North. *Institutions, Institutional Change, and Economic Performance*. Cambridge University Press, New York (NY), 1990.
14. D. C. North, J. J. Wallis, and B. R. Weingast. *Violence and Social Orders: A Conceptual Framework for Interpreting Recorded Human History*. Cambridge University Press, New York (NY), 2009.
15. W. R. Scott. Approaching Adulthood: The Maturing of Institutional Theory. *Theory and Society*, 37(5):427–442, 2008.
16. G. Simmel. *The Sociology of Georg Simmel*. Free Press, New York (NY), 1964. Originally published: 1908.
17. Q. van Doosselaere. *Commercial Agreements and Social Dynamics in Medieval Genoa*. Cambridge University Press, Cambridge (UK), 2009.
18. D. Villatoro, G. Andrighetto, J. Sabater-Mir, and R. Conte. Dynamic Sanctioning for Robust and Cost-efficient Norm Compliance. In T. Walsh, editor, *Proceedings of the 22nd International Joint Conference on Artificial Intelligence*, volume 1 of *IJCAI'11*, pages 414–419. AAAI Press, 2011.
19. G. H. von Wright. *Norm and Action: A Logical Enquiry*. Routledge & Kegan Paul, London, 1963.
20. C. Watkins. *Learning from Delayed Rewards*. PhD thesis, Cambridge University, Cambridge (UK), 1989.