Norm Creation in Proposition Control Games*

Xin Sun
xin.sun@uni.lu

Livio Robaldo
livio.robaldo@uni.lu

Faculty of Science, Technology and Communication, University of Luxembourg

Abstract

In this paper we study how to create norms in games. We consider norms as deontic statements which are used to describe prescriptions made by certain randomized signals of which the usage is to guide agents' behavior. Such randomized signals are implemented from correlated equilibria in games. Our proposal belongs to the offline norm creation approach [ST96]. Agents' compliance and computational tractability have been identified as interesting problems to cope with in the offline norm creation approach. In this paper, four types of norms are studied: utilitarian norms, egalitarian norms, Nash product norms and opportunity-balanced norms. We show that in our framework all four types of norms can be created in polynomial time and all rational agents will comply with those norms.

1 Introduction

Norms, or social laws, have been extensively studied as a mechanism for coordinating interactions within multi-agent systems. Two general approaches to the design of norms have been investigated in the literature: offline approach [ST96, vdHRW07, ÅW10, ÅvdHW12] and online approach [ST97, SA07, MLE11, MLR+15]. Online approach aims to equip agents with the ability to dynamically coordinate their activities, for example by explicitly reasoning about coordination at run-time or learning from the interaction with other agents. In contrast, offline techniques aim at developing a coordination regime at design-time, and build this regime into a system for use at run-time. There are arguments in favor of both approaches: the online approach is potentially more flexible, and may be more robust against unanticipated events, while the offline approach benefits from offline reasoning about coordination, thereby reducing the run-time decision-making burden on agents. This paper belongs to the offline approach.

One of the most popular offline approaches is the social law paradigm, originally introduced by Shohan and Tennenholtz [ST96]. Social laws are understood as a set of rules imposed upon a multi-agent system with the goal of ensuring some desirable states. Social laws work by constraining the behavior of the agents in the system—by forbidding agents from performing certain actions in certain states.

There are two disadvantages of the social law paradigm. Firstly, the computational complexity of the creation of social laws that will effectively coordinate a multi-agent system is intractable. In Shohan and Tennenholtz [ST96], the design of social law is to find suitable restrictions for the system, which is NP complete. In van der Hoek *et al* [vdHRW07], the design of social laws is framed as model checking problem of alternating-time temporal logic. It is again NP complete. Ågotnes and Woolridge [ÅW10] extend the model by taking into account both the implementation costs of social laws

---

and multiple (possibly conflicting) design objectives with different priorities. In this setting, the design of social laws becomes an optimization problem and is FP^{NP} complete. Secondly, there is no guarantee whether agents will comply with the created social laws. There might be conflicts between the interest of agents and the designer. The complexity of the generation of social laws such that every agent will choose to comply is even more intractable.

In this paper, following ideas from the game theoretical analysis of norms [Gin10], we develop a different offline paradigm, such that the complexity of the norm creation problem is tractable and every agent will comply with the created norms. The main features of our framework are the following:

1. Instead of relational structures, we use a kind of succinct game called proposition control game to represent multi-agent system.

2. Instead of constraints, norms in our framework are viewed as randomized signals, like traffic lights, to guide agents' behavior. We generate randomized signals by computing correlated equilibrium of proposition control games. Norms are descriptions of those randomized signals using deontic statements.

3. Four kinds of norms are created: utilitarian norms, egalitarian norms, Nash product norms and opportunity-balanced norms.

    (a) Utilitarian norms are created from those correlated equilibria which maximize the sum of the expect utility of each agents.

    (b) Egalitarian norms are created from those correlated equilibria which maximize the expected utility of the weakest agent.

    (c) Nash product norms are created from those correlated equilibria which maximize the product of the expected utility of all agent.

    (d) Opportunity-balanced norms are created from those correlated equilibria which are computed by taking the average of correlated equilibrium which maximize the expect utility of every single agent.

The procedure of norm creation in this paper is as follows: At first a proposition control game is given. Then we compute a correlated equilibrium of the given game. The resulting correlated equilibrium is a probability distribution over agents' strategy profiles. We then transform the probability distribution to randomized signals to guide agents' behavior. Finally we use deontic statements to describe the prescriptions given by those randomized signals. Those deontic statements are norms.

The structure of the rest of this paper is as follows. We briefly review proposition control game in Section 2. Then in Section 3 we explain in details our framework and show how norms can be created in polynomial time. In Section 4 we discuss related work. We summarize this paper in Section 5 with suggestions of future work.

2 Proposition control game

Proposition control game is a variant of Boolean game. Boolean game is super succinct in the sense that agents' strategy and utility function are represented implicitly. Such succinctness is reached with a cost: many decision problems in Boolean games are intractable. For example deciding whether there is a pure strategy Nash equilibrium in a given Boolean game is Σ_2^P hard [BLLZ06]. To find a balance between succinctness and tractability, we introduce proposition control games.

In a proposition control game, the strategies available to each agent consist in assigning a truth value to each variable he can control. The goal of each agent is represented by a set of weighted formulas. Formally, let $\mathbb{P} = \{p_0, p_1, \ldots\}$ be a finite set of propositional variables and let $L_{\mathbb{P}}$ be the propositional language built from $\mathbb{P}$. $2^{\mathbb{P}}$ is the set of all valuations for $\mathbb{P}$, with the usual convention that for $V \in 2^{\mathbb{P}}$ and $p \in V$, V gives the value true to p if $p \in V$ and false otherwise. Let $X \subseteq \mathbb{P}$, 2^X is the set of X-valuations. A partial valuation (for $\mathbb{P}$) is an X-valuation for some $X \subseteq \mathbb{P}$. Partial valuations are denoted by listing all variables of X, with a " $+$ " symbol when the variable is set to be true and a " $-$ " symbol when the variable is set to be false: for instance, let $X = \{p, q, r\}$, then the X-valuation $V = \{p, r\}$ is denoted $\{+p, -q, +r\}$. If $\{\mathbb{P}_1, \ldots, \mathbb{P}_n\}$ is a partition of $\mathbb{P}$ and $V_1, \ldots, V_n$ are partial valuations, where $V_i \in 2^{\mathbb{P}_i}$, $(V_1, \ldots, V_n)$ denotes the valuation $V_1 \cup \ldots \cup V_n$.

Definition 1 (proposition control game) *A proposition control game is a tuple* $(Agent, \mathbb{P}, \pi, S_1, \ldots, S_n, Goal)$, *where*

1. $Agent = \{1, \ldots, n\}$ *is a set of agents.*

2. $\mathbb{P}$ *is a finite set of propositional variables.*

3. $\pi : Agent \mapsto 2^{\mathbb{P}}$ *is a control assignment function such that* $\{\pi(1), \ldots, \pi(n)\}$ *forms a partition of* $\mathbb{P}$.

4. *For each agent* i, $S_i \subseteq 2^{\pi(i)}$ *is his strategy set.*

5. $Goal = \{Goal_1, \ldots, Goal_n\}$ *is a set of sets of weighted formulas of* $L_{\mathbb{P}}$. *Each* $Goal_i$ *is a finite set* $\{\langle x_1, m_1 \rangle, \ldots, \langle x_k, m_k \rangle\}$ *where* $x_j \in L_{\mathbb{P}}$ *and* m_j *is a real number representing the weight of* x_j.

A strategy for agent i is a $\pi(i)$−valuation. Note that since $\{\pi(1), \ldots, \pi(n)\}$ forms a partition of $\mathbb{P}$, a strategy profile $s = (s_1, \ldots, s_n)$ is a valuation for $\mathbb{P}$. Agents' utilities in proposition control games are induced by their goals. For every agent i and every strategy profile s, $u_i(s) = \Sigma\{m_j : \langle \phi_j, m_j \rangle \in Goal_i, s \vDash \phi_j\}$. Agent's preference over strategy profiles is induced by his utility function naturally: $s \leq_i s'$ iff $u_i(s) \leq u_i(s')$. Let $s = (s_1, \ldots, s_n)$ be a strategy profile, we use s_{-i} to denote the projection of s on $Agent - \{i\}$: $s_{-i} = (s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_n)$ and s_i to denote the projection of s on i's strategy.

The main difference between proposition control game and Boolean game is the representation of agents' strategy set. In a proposition control game, an agent's strategy set is a **subset** of the power set of the propositional variables he can control, while in a Boolean game the strategy set of each agent is the **entire power set** of the propositional variables he can control. Such difference is the reason for proposition control game being computationally easier than Boolean game. [1] Moreover, reality supports the restriction we made on proposition control game. In real life, it is common that we can set something to be true, but not able to set it to be false, or vice versa. For example, the president of USA is able to start a war ($+w$), but has no ability to stop a war ($-w$). That is to say, $+w$ is in the president's strategy set while $-w$ is not. As another example, we are able to not hit the bull's eye ($-b$) when we are throwing a dart, but not able to hit the bull's eye ($+b$), which means $-b$ is in our strategy set while $+b$ is not. For the sake of tractability and being realistic, we sacrifice the super-succinctness of Boolean game and use proposition control game instead.

Example 1 *Let* $G = (Agent, \mathbb{P}, \pi, S_1, S_2, Goal)$ *where* $Agent = \{1, 2\}$, $\mathbb{P} = \{p, q, r, s\}$, $\pi(1) = \{p, r\}$, $\pi(2) = \{q, s\}$, $S_1 = \{\{p, r\}, \{p\}, \{r\}\}$, $S_2 = \{\{q, s\}, \{q\}, \{s\}\}$, $Goal_1 = \{\langle p \leftrightarrow q, 1 \rangle, \langle s, 2 \rangle\}$, $Goal_2 = \{\langle p \wedge q, 2 \rangle, \langle \neg s, 1 \rangle\}$. *This proposition control game is depicted as follows:*

	$+q, +s$	$+q, -s$	$-q, +s$
$+p, +r$	$(3, 2)$	$(1, 3)$	$(2, 0)$
$-p, +r$	$(2, 0)$	$(0, 1)$	$(3, 0)$
$+p, -r$	$(3, 2)$	$(1, 3)$	$(2, 0)$

3 Norms and correlated equilibrium

In game theory, correlated equilibrium is a solution concept that is more general than the well known (mixed-strategy) Nash equilibrium. But it is insufficiently studied within the normative multi-agent system community, if not completely ignored. Correlated equilibrium is first discussed in Aumann [Aum74]. A correlated equilibrium is a probability distribution over agents' strategy profiles, which can be understood as a public randomized signal, such that each agent can choose his action according to the recommendation of the signal. If no agent would want to deviate from the recommended strategy (assuming the others don't deviate), the probability distribution is called a correlated equilibrium.

Definition 2 (correlated equilibrium) *Given a proposition control game* $(Agent, \mathbb{P}, \pi, S_1, \ldots, S_n, Goal)$, *a correlated equilibrium is a probability distribution* τ *over* $S = S_1 \times \ldots \times S_n$ *such that for every agent* $i \in Agent$, *for every pair of strategies* $s_i, s_i' \in S_i$, *we have:*

$$\sum_{s_{-i} \in S_{-i}} \tau(s_i, s_{-i}) \times u_i(s_i, s_{-i}) \geq \sum_{s_{-i} \in S_{-i}} \tau(s_i, s_{-i}) \times u_i(s_i', s_{-i})$$

where $S_{-i} = S_1 \times \ldots \times S_{i-1} \times S_{i+1} \times \ldots \times S_n$.

---

[1]As a side effect, proposition control game is not as succinct as Boolean game because in Boolean games strategy sets are represented implicitly. However, the succinctness of proposition control game is still better than normal form game. In a normal form game, both strategy set and utility function are represented explicitly, while in propositional control game the utility function is compactly represented by propositional formulas. Every proposition control game can be transformed into a normal form game in polynomial time. Moreover, we can improve the succinctness of proposition control game without damaging the computational tractability by representing strategy set as a propositional formula of which the true valuation can be effectively computed. Similar treatment has been studied by Bonzon [Bon07] for Boolean game. We leave such task for proposition control game as future work.

Intuitively, in a correlated equilibrium the expected utility of an agent following the randomized signal produced by the correlated equilibrium ($\sum_{s_{-i} \in S_{-i}} \tau(s_i, s_{-i}) \times u_i(s_i, s_{-i})$) is larger than his expected utility of deviating from the signal ($\sum_{s_{-i} \in S_{-i}} \tau(s_i, s_{-i}) \times u_i(s'_i, s_{-i})$). An immediate positive consequence of this properties is that all those norms created from correlated equilibrium will be obeyed by rational agents because they cannot be better-off by violating those norms (deviating from the randomized signal).

An equivalent characterization of correlated equilibrium is

$$\sum_{s_{-i} \in S_{-i}} \tau(s_i, s_{-i}) \times (u_i(s_i, s_{-i}) - u_i(s'_i, s_{-i})) \geq 0.$$

The difference between $u_i(s'_i, s_{-i})$ and $u_i(s_i, s_{-i})$ can be understood as a measure of the *regret* of agent i choosing s'_i instead of s_i, when other agents chooses s_{-i}. The regret grows when $(u_i(s'_i, s_{-i}) - u_i(s_i, s_{-i}))$ grows. When $(u_i(s_i, s_{-i}) - u_i(s'_i, s_{-i}))$ is positive, agent i has no regret for choosing s_i compared to s'_i, when other agents choosing s_{-i}. Therefore this characterization of correlated equilibrium shows that there is no expected regret for each agent in a correlated equilibrium, which is why all agents are willing to follow the signal produced by a correlated equilibrium.

Example 2 *Let $G = (Agent, \mathbb{P}, \pi, S_1, S_2, Goal)$ where $Agent = \{1, 2\}$, $\mathbb{P} = \{p, q\}$, $\pi(1) = \{p\}$, $\pi(2) = \{q\}$, $S_1 = \{\{p\}, \emptyset\}$, $S_2 = \{\{q\}, \emptyset\}$, $Goal_1 = \{\langle p \wedge q, 4\rangle, \langle p \wedge \neg q, 1\rangle, \langle \neg p \wedge q, 5\rangle, \langle \neg p \wedge \neg q, 0\rangle\}$, $Goal_2 = \{\langle p \wedge q, 4\rangle, \langle p \wedge \neg q, 5\rangle, \langle \neg p \wedge q, 1\rangle, \langle \neg p \wedge \neg q, 0\rangle\}$. This game is called chicken game in the game theory literature. A realistic scenario for this game is two driver meet at a single lane bridge from opposite directions. The first to swerve away (playing $+p$ or $+q$) yields the bridge to the other. If neither player swerves, the result is a costly deadlock in the middle of the bridge, or a potentially fatal head-on collision. It is depicted by a payoff matrix as follows:*

	$+q,$	$-q$
$+p$	$(4, 4)$	$(1, 5)$
$-p$	$(5, 1)$	$(0, 0)$

This game has two pure Nash equilibria: pure strategy profile $(-p, +q)$ with expected utility vector $(5, 1)$ and $(+p, -q)$ with expected utility vector $(1, 5)$. It has one mixed Nash equilibrium: the mixed strategy profile $(1/2, 1/2)$, which assign probability $1/2$ to $+p$ and $+q$, with expected utility vector $(2.5, 2.5)$.

τ is a correlated equilibrium in this game if the following is satisfied:

- $\tau(+p, +q) \times 4 + \tau(+p, -q) \times 1 \geq \tau(+p, +q) \times 5 + \tau(+p, -q) \times 0$

- $\tau(-p, +q) \times 5 + \tau(-p, -q) \times 0 \geq \tau(-p, +q) \times 4 + \tau(-p, -q) \times 1$

- $\tau(+p, +q) \times 4 + \tau(-p, +q) \times 1 \geq \tau(+p, +q) \times 5 + \tau(-p, +q) \times 0$

- $\tau(+p, -q) \times 5 + \tau(-p, -q) \times 0 \geq \tau(+p, -q) \times 4 + \tau(-p, -q) \times 1$

Therefore $\tau_1(+p, +q) = 0.2, \tau_1(+p, -q) = 0.3, \tau_1(-p, +q) = 0.4, \tau_1(-p, -q) = 0.1$ is a correlated equilibrium with expected utility $(3.1, 2.7)$. $\tau_2(+p, +q) = \tau_2(+p, -q) = \tau_2(-p, +q) = 1/3$ is another correlated equilibrium with expected utility $(10/3, 10/3)$.

One advantage of correlated equilibrium from a computational perspective is that they are computationally less expensive than Nash equilibrium.[2] This can be captured by the fact that computing a correlated equilibrium only requires solving a linear program whereas solving a Nash equilibrium requires solving a mixed integer programming problem. It is known from the algorithmic game theory literature that correlated equilibrium of normal form games can be computed in polynomial time [PR08], while computing Nash equilibrium is PPDA complete [DGP09].

Although we view norms are created from correlated equilibrium, this does not mean each correlated equilibrium should be implemented as norms. Instead, we are only interested in those correlated equilibria which satisfy certain desired properties, for example maximize social welfare, protect the weakest member of the society, or ensure fairness.

In the theory pf multiagent resource allocation [CDE$^+$06], three types of social welfare are widely used.

- Utilitarian social welfare is the sum of the utilities of all agents.

---

[2]Gintis [Gin10] presents more conceptual justifications for viewing norms as correlated equilibria instead of Nash equilibria.

- Egalitarian social welfare is the utility of the worst off agents.

- Nash product social welfare is the product of the utilities of all agents.

Correspondingly we introduce utilitarian/egalitarian/Nash product correlated equilibrium. Moreover, we introduce opportunity-balanced correlated equilibrium, which in certain sense ensures fairness. In this paper we only use those types of correlated equilibrium to create norms.

3.1 Utilitarian correlated equilibrium

For each agent i, his expected utility in a correlated equilibrium τ is calculated as follows:

$$EU_i(\tau) = \sum_{s \in S} \tau(s) \times u_i(s)$$

Given a proposition control game $(Agent, \mathbb{P}, \pi, S_1, \ldots, S_n, Goal)$ and a correlated equilibrium τ. The utilitarian social welfare induced by τ is

$$SW^u(\tau) = \sum_{i \in Agent} EU_i(\tau)$$

Utilitarian social welfare sums up the agents' expected utilities in a given correlated equilibrium, thus providing a useful measure of the overall benefit of the society.

Definition 3 (utilitarian correlated equilibrium) *Given a proposition control game, and Θ the set of all correlated equilibrium of this game. $\tau \in \Theta$ is a utilitarian correlated equilibrium if τ maximize $SW^u(\tau)$, i.e. for all $\tau' \in \Theta$, $SW^u(\tau) \geq SW^u(\tau')$.*

A utilitarian correlated equilibrium can be computed using linear programming with maximizing $SW^u(\tau)$ as the objective function and requirements of correlated equilibrium as constrains. Utilitarian norms are deontic statements which describes those randomized signals implemented from utilitarian correlated equilibrium. For example, in the chicken game utilitarian correlated equilibrium are solutions of the following linear program:

$$\max \tau(+p, +q) \times (4 + 4) + \tau(-p, +q) \times (5 + 1) + \tau(+p, -q) \times (1 + 5) + \tau(-p, -q) \times (0 + 0)$$

subject to

- $\tau(+p, +q) \times 4 + \tau(+p, -q) \times 1 \geq \tau(+p, +q) \times 5 + \tau(+p, -q) \times 0$

- $\tau(-p, +q) \times 5 + \tau(-p, -q) \times 0 \geq \tau(-p, +q) \times 4 + \tau(-p, -q) \times 1$

- $\tau(+p, +q) \times 4 + \tau(-p, +q) \times 1 \geq \tau(+p, +q) \times 5 + \tau(-p, +q) \times 0$

- $\tau(+p, -q) \times 5 + \tau(-p, -q) \times 0 \geq \tau(+p, -q) \times 4 + \tau(-p, -q) \times 1$

- $\tau(+p, +q) + \tau(-p, +q) + \tau(+p, -q) + \tau(-p, -q) = 1$

- $\tau(+p, +q), \tau(-p, +q), \tau(+p, -q), \tau(-p, -q) \geq 0$

Let $\tau(+p, +q)$ be x_1, $\tau(-p, +q)$ be x_2, $\tau(+p, -q)$ be x_3 and $\tau(-p, -q)$ be x_4. This linear programme is transformed to:

$$\max 8x_1 + 6x_2 + 6x_3$$

subject to

- $x_1 - x_3 \leq 0$

- $-x_2 + x_4 \leq 0$

- $x_1 - x_2 \leq 0$

- $-x_3 + x_4 \leq 0$

- $x_1 + x_2 + x_3 + x_4 = 1$

- $x_1, x_2, x_3, x_4 \geq 0$

Using techniques of linear programming, we find a solution for the linear program with $x_1 = x_2 = x_3 = \frac{1}{3}$. Such a utilitarian correlated equilibrium can be transformed to randomized signals. For example we can use a random number generator to generate a real number between 0 and 1. If the number is less than $\frac{1}{3}$, then we show a red signal to agent 1 and a green signal to agent 2. If the number is larger than $\frac{2}{3}$, then we show a green signal to agent 1 and a red signal to agent 2. Otherwise we show a red signal to both agents. Then we create the following two norms for agent 1:

- if the signal you see is red, then p is obligatory.

- if the signal you you see is green, then $\neg p$ is obligatory.

And we generate the following two norms for agent 2:

- if the signal you see is red, then q is obligatory.

- if the signal you you see is green, then $\neg q$ is obligatory.

Note that to ensure that each agent to comply with the norms, we should make the probability distribution of signals public meanwhile keep signals private for each agent. When one agent sees both agents' signals, it is possible that he want to violate the norms. For example, when agent 1 see a red signal for himself as well as a red signal for agent 2, then he has an incentive to choose $-p$, which is a violation of his norm. But if agent 1 only sees his own red signal, then he knows that the probability of the signal for agent 2 being red is $\frac{1}{2}$. Therefore if he choose $-p$, his expected utility will be $\frac{5}{2}$, which is no larger than his expected utility of choosing $+p$. The incentive of violation disappears. Such treatment of publicity and privacy is also required for implementing other types of norms in order to ensure agent's compliance.

Now consider a different example where one ambulance meet a bus in a crossroad.

Example 3 *(ambulance game) Let $G = (Agent, \mathbb{P}, \pi, S_1, S_2, Goal)$ where $Agent = \{1,2\}$, $\mathbb{P} = \{p,q\}$, $\pi(1) = \{p\}$, $\pi(2) = \{q\}$, $S_1 = \{\{p\}, \emptyset\}$, $S_2 = \{\{q\}, \emptyset\}$, $Goal_1 = \{\langle p, -10 \rangle, \langle \neg p \wedge q, 10 \rangle, \langle \neg p \wedge \neg q, -20 \rangle\}$, $Goal_2 = \{\langle q, 0 \rangle, \langle p \wedge \neg q, 5 \rangle, \langle \neg p \wedge \neg q, -10 \rangle\}$. It is depicted as follows:*

	$+q,$	$-q$
$+p$	$(-10, 0)$	$(-10, 5)$
$-p$	$(10, 0)$	$(-20, -10)$

Here $+p$ means the ambulance stop while $+q$ means the bus stop. τ is a correlated equilibrium in this game if the following is satisfied:

- $\tau(+p, +q) \times (-10) + \tau(+p, -q) \times (-10) \geq \tau(+p, +q) \times 10 + \tau(+p, -q) \times (-20)$

- $\tau(-p, +q) \times 10 + \tau(-p, -q) \times (-20) \geq \tau(-p, +q) \times (-10) + \tau(-p, -q) \times (-10)$

- $\tau(+p, +q) \times 0 + \tau(-p, +q) \times 0 \geq \tau(+p, +q) \times 5 + \tau(-p, +q) \times (-10)$

- $\tau(+p, -q) \times 5 + \tau(-p, -q) \times (-10) \geq \tau(+p, -q) \times 0 + \tau(-p, -q) \times 0$

Let $\tau(+p, +q)$ be x_1, $\tau(-p, +q)$ be x_2, $\tau(+p, -q)$ be x_3 and $\tau(-p, -q)$ be x_4. The linear program we need to calculate correlated equilibria is:

$$max \ z = -10x_1 + 10x_2 - 5x_3 - 30x_4$$

subject to

- $20x_1 - 10x_3 \leq 0$

- $-20x_2 + 10x_4 \leq 0$

- $5x_1 - 10x_2 \leq 0$

- $-5x_3 + 10x_4 \leq 0$

- $x_1 + x_2 + x_3 + x_4 = 1$

- $x_1, x_2, x_3, x_4 \geq 0$

Using techniques of linear programming, we find a solution for the linear program with $x_2 = 1, x_1 = x_3 = x_4 = 0$. Such a utilitarian correlated equilibrium can be implemented to a norm:

- $\neg p \wedge q$ *is obligatory.*

The norm created in the ambulance game explains real traffic rule quite well.

Theorem 1 *Utilitarian correlated equilibria can be computed in polynomial time.*

Proof: A utilitarian correlated equilibrium can be computed using linear programming with the objective function maximizing the sum of the expected utility of all agents. A linear program can be solved in polynomial time using standard techniques like ellipsoid algorithms Alevras and Padberg [AP01]. $\qquad \square$

3.2 Egalitarian correlated equilibrium

The concept of egalitarian social welfare is inspired by the work of Rawls [Raw71]. Intuitively, egalitarian social welfare is measured by the situation of the weakest member of society. Given a proposition control game $(Agent, \mathbb{P}, \pi, S_1, \ldots, S_n, Goal)$ and a correlated equilibrium τ. The egalitarian social welfare induced by τ is

$$SW^e(\tau) = min\{EU_i(\tau) : i \in Agent\}$$

Egalitarian social welfare offers a measure of fairness in cases where the minimum needs of all agents has to be satisfied.

Definition 4 (egalitarian correlated equilibrium) *Given a proposition control game, and Θ the set of all correlated equilibrium of this game. $\tau \in \Theta$ is a egalitarian correlated equilibrium if τ maximize $SW^e(\tau)$, i.e. for all $\tau' \in \Theta$, $SW^e(\tau) \geq SW^e(\tau')$.*

Egalitarian norms are deontic statements of the prescription of those randomized signals transformed from egalitarian correlated equilibria. Computing egalitarian correlated equilibrium seems to be more difficult than computing utilitarian correlated equilibrium at the first sight because maximizing egalitarian social welfare is not a linear function. However, the good news is, we can still use linear programming to compute egalitarian correlated equilibrium efficiently, as the following theorem shows.

Theorem 2 *Egalitarian correlated equilibrium can be computed in polynomial time.*

Proof: We use linear programming to compute egalitarian correlated equilibrium. The variables of the linear programming are random variables $\tau(s), \tau(s'), \ldots$, representing the probability assigned to each strategy profile of the game. The objective function of the linear programming is to maximize the egalitarian social welfare, which is represented by a variable z. There are four groups of constrains in this linear program.

1. $\tau(s) \geq 0$, for all $s \in S_1 \times \ldots \times S_n$.

2.
$$\sum_{s \in S_1 \times \ldots \times S_n} \tau(s) = 1$$

These two groups of constrains together ensures that the function τ is a probability distribution over strategy profiles.

3. For all i,, for all s_i and $s_i' \in S_i$,

$$\sum_{s_{-i} \in S_{-i}} \tau(s_i, s_{-i}) \times u_i(s_i, s_{-i}) \geq$$

$$\sum_{s_{-i} \in S_{-i}} \tau(s_i, s_{-i}) \times u_i(s_i', s_{-i})$$

The third group of constrains say that the solution of this linear programming must be a correlated equilibrium.

4. For all i,

$$\sum_{s \in S_1 \times \ldots \times S_n} u_i(s)\tau(s) \geq z$$

The fourth group of constrains ensures that z is no larger than the expected utility of any agent in this correlated equilibrium.

These four groups of constrains together ensures that the solution of this linear programming indeed maximizes the egalitarian social welfare. It can be verified that the size of this linear program is polynomial to the size of the give proposition control game. Since a linear program can be solved in polynomial time, we know that egalitarian correlated equilibrium can be computed in polynomial time. □

3.3 Nash product correlated equilibrium

Given a proposition control game $(Agent, \mathbb{P}, \pi, S_1, \ldots, S_n, Goal)$ in which every goal of every agent is non-negative and a correlated equilibrium τ. The Nash product social welfare induced by τ is

$$SW^n(\tau) = \prod_{i \in Agent} EU_i(\tau)$$

Nash product social welfare can be regarded as a compromise between the utilitarian and the egalitarian social welfare. On the one hand, just as utilitarian social welfare, the Nash product increases with single increasing of individual utilities. On the other hand, just as egalitarian social welfare, the Nash product reaches its maximum when the utilities distributed equally over all agents.

Definition 5 (Nash product correlated equilibrium) *Given a proposition control game, and Θ the set of all correlated equilibrium of this game. $\tau \in \Theta$ is a Nash product correlated equilibrium if τ maximize $SW^n(\tau)$, i.e. for all $\tau' \in \Theta$, $SW^n(\tau) \geq SW^n(\tau')$.*

Nash product norms are deontic statements of the prescription of those randomized signals transformed from Nash product correlated equilibrium. The following theorem shows that the computation of Nash product correlated equilibrium is tractable.

Theorem 3 *Nash product correlated equilibrium can be computed in polynomial time.*

Proof: We use convex optimization to compute Nash product correlated equilibrium.[3] Suppose $s^1, \ldots, s^k$ are all strategy profiles in the given game. We use $\tau(s^1), \ldots, \tau(s^k)$ to represent the probability assigned to each strategy profile of the game. To compute Nash product correlated equilibrium we need to maximize the Nash product social welfare:

$$\prod_{i \in Agent} \sum_{s \in S} \tau(s) \times u_i(s) =$$

$$\prod_{i \in Agent}(\tau(s^1) \times u_i(s^1) + \ldots + \tau(s^k) \times u_i(s^k)) =$$

$$(\tau(s^1) \times u_1(s^1) + \ldots + \tau(s^k) \times u_1(s^k)) \times \ldots \times (\tau(s^1) \times u_n(s^1) + \ldots + \tau(s^k) \times u_n(s^k)) =$$

$$(\tau(s^1))^n \times \prod_{i \in Agent} u_i(s^1) + \ldots + (\tau(s^k))^n \times \prod_{i \in Agent} u_i(s^k)$$

Note that the above function is a convex function because every goal of every agent is non-negative, which ensures that $u_i(s^j)$ is non-negative. The above function is maximized iff the following function is minimized

$$-\log((\tau(s^1))^n \times \prod_{i \in Agent} u_i(s^1) + \ldots + (\tau(s^k))^n \times \prod_{i \in Agent} u_i(s^k)) \qquad (\divideontimes)$$

Note that $(\divideontimes)$ is also a convex function because the function $f(x) = -\log x$ is a convex function and the composition of two convex functions is again a convex function. Now we take minimizing $(\divideontimes)$ as our objective function and built a convex optimization problem by imposing four groups of constrains the same as in the proof of Theorem 2. Such a convex optimization problem can be solved using standard techniques from convex optimization, say interior-point methods, in polynomial time. □

---

[3]A comprehensive introduction of convex optimization can be found in Boyd and Vandenberghe [BV04].

3.4 Opportunity-balanced correlated equilibrium

A fourth kind of correlated equilibrium of interest to us is what we call opportunity-balanced correlated equilibrium. Intuitively, an opportunity-balanced correlated equilibrium is the average of those correlated equilibria which maximize each single agent's expected utility.

Definition 6 (opportunity-balanced correlated equilibrium) *Given a proposition control game* $(Agent, \mathbb{P}, \pi, S_1, \ldots, S_n, Goal)$, *an opportunity-balanced correlated equilibrium is a correlated equilibrium* τ *such that* $\tau(s) = (\tau_i(s) + \ldots + \tau_n(s))/n$, *where* τ_i *is a correlated equilibrium which maximize the expected utility of agent* i.

The set of all correlated equilibria of a game is a convex set because correlated equilibrium is defined using linear constrains. Therefore the average of finite correlated equilibria is again a correlated equilibrium.

Example 4 *(opportunity-balanced correlated equilibrium for the chicken game) We first solve two linear programs with objective function max* $4x_1 + 5x_2 + x_3$ *and* $4x_1 + x_2 + 5x_3$ *respectively. Solving the first linear program give us* τ_1 *with* $\tau_1(x_2) = 1, \tau_1(x_1) = \tau_1(x_3) = \tau_1(x_4) = 0$. *Solving the second linear program give us* τ_2 *with* $\tau_2(x_3) = 1, \tau_2(x_1) = \tau_2(x_2) = \tau_2(x_4) = 0$. *Then we calculate the opportunity-balanced correlated equilibrium* τ *by taking the average of* τ_1 *and* τ_2: $\tau(x_2) = \tau(x_3) = \frac{1}{2}, \tau(x_1) = \tau(x_4) = 0$.

Example 5 *(opportunity-balanced correlated equilibrium for the ambulance game) We first solve two linear programs with objective function max* $-10x_1 + 10x_2 - 10x_3 - 20x_4$ *and* $5x_3 - 10x_4$ *respectively. Solving the first linear program give us* τ_1 *with* $\tau_1(x_2) = 1, \tau_1(x_1) = \tau_1(x_3) = \tau_1(x_4) = 0$. *Solving the second linear program give us* τ_2 *with* $\tau_2(x_3) = 1, \tau_2(x_1) = \tau_2(x_2) = \tau_2(x_4) = 0$. *Then we calculate the opportunity-balanced correlated equilibrium* τ *by taking the average of* τ_1 *and* τ_2: $\tau(x_2) = \tau(x_3) = \frac{1}{2}, \tau(x_1) = \tau(x_4) = 0$.

Opportunity-balanced norms are deontic statements of the prescription of those randomized signals transformed from opportunity-balanced correlated equilibria. The following theorem shows that the computation of opportunity-balanced correlated equilibria is tractable.

Theorem 4 *Opportunity-balanced correlated equilibria can be computed in polynomial time.*

Proof: Given an arbitrary proposition control game with n agents. An opportunity-balanced correlated equilibrium can be computed as follows:

1. For each agent, use linear programming to compute a correlated equilibrium which maximizes the agent's expected utility.

2. Take the average of the n correlated equilibria from the previous step.

Since each linear program can be solved in polynomial time, an opportunity-balanced correlated equilibrium can be computed in polynomial time. □

4 Related work

Except the literature mentioned in the introductory section, the generation of norms is also studied in sociology and philosophy. In Binmore [Bin05] social norms are tools to solve the equilibrium selection in a coordination game with multiple Nash equilibria. His approach is based on evolutionary game theory as he believes that the interaction of agent's strategies are infinitely repeated games. While all equilibria in a game are not pareto optimal, norms arise to solve the equilibrium selection problem. Other evolutionary game theoretical approach of norm emergence can be found in Alexander [Ale07] and Skyrms [Sky14].

Gintis [Gin10] present his work as a complementary step for the work of Binmore. Instead of Nash equilibrium, Ginitis propose to use correlated equilibrium to generate norms. He considered norms as choreographers who send signals to agents to guide agents' behavior. A rational agent would follow the suggestion of choreographers if he believes that others agents are following it, too. This paper is a further development of Gintis' original work from a computational perspective.

We understand creation and emergence as two different methods of norm generation. It seems concepts from evolutionary game theory, like evolutionary stable strategy and population dynamics, are more suitable for norm emergence, while concepts from traditional game theory, like Nash/correlated equilibrium, are more suitable for norm creation.

5 Summary and future work

In this paper we study how to create norms in proposition control games. We consider norms as deontic statements which are used to describe the prescriptions given by certain randomized signals of which the usage is to guide agents' behavior. Such randomized signals are implemented from correlated equilibrium in games. Four types of norms are studied: utilitarian norms, egalitarian norms, Nash product norms and opportunity-balanced norms. All these norms can be created in polynomial time in our framework. Moreover, since all these norms created from correlated equilibrium, it is to each agent's interest to comply with these norms.

An avenue of future work is to investigate the limitation of norms. In those games where there is a single pure Nash equilibrium, like the well-known prisoner's dilemma, the only norm which will be created using our method is the norm indicating the Nash equilibrium is obligatory, because the Nash equilibrium is also the unique correlated equilibrium. Therefore in such situations norms are not helpful in avoiding dominated outcomes. What we need in those situations is a stronger mechanism, for example laws, to enforce punishment to help agents to avoid undesirable outcomes.

References

[Ale07]   J. McKenzie Alexander. *The Structural Evolution of Morality*. Cambridge University Press, 2007.

[AP01]   Dimitris Alevras and Manfred W. Padberg. *Linear Optimization and Extensions: Problems and Solutions*. Springer, 2001.

[Aum74]   Robert Aumann. Subjectivity and correlation in randomized strategies. *Journal of Mathematical Economics*, 1:67–96, 1974.

[ÅvdHW12]   Thomas Ågotnes, Wiebe van der Hoek, and Michael Wooldridge. Conservative social laws. In Luc De Raedt, Christian Bessière, Didier Dubois, Patrick Doherty, Paolo Frasconi, Fredrik Heintz, and Peter J. F. Lucas, editors, *ECAI 2012 - 20th European Conference on Artificial Intelligence. Including Prestigious Applications of Artificial Intelligence (PAIS-2012) System Demonstrations Track, Montpellier, France, August 27-31 , 2012*, volume 242 of *Frontiers in Artificial Intelligence and Applications*, pages 49–54. IOS Press, 2012.

[ÅW10]   Thomas Ågotnes and Michael Wooldridge. Optimal social laws. In Wiebe van der Hoek, Gal A. Kaminka, Yves Lespérance, Michael Luck, and Sandip Sen, editors, *9th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2010), Toronto, Canada, May 10-14, 2010, Volume 1-3*, pages 667–674. IFAAMAS, 2010.

[Bin05]   Ken Binmore. *Natural Justice*. Oxford University Press, 2005.

[BLLZ06]   Elise Bonzon, Marie-Christine Lagasquie-Schiex, Jérôme Lang, and Bruno Zanuttini. Boolean games revisited. In Gerhard Brewka, Silvia Coradeschi, Anna Perini, and Paolo Traverso, editors, *ECAI 2006, 17th European Conference on Artificial Intelligence, August 29 - September 1, 2006, Riva del Garda, Italy, Including Prestigious Applications of Intelligent Systems (PAIS 2006), Proceedings*, volume 141 of *Frontiers in Artificial Intelligence and Applications*, pages 265–269. IOS Press, 2006.

[Bon07]   Elise Bonzon. *Modélisation des interactions entre agents rationnels : les jeux boolens*. PhD thesis, Universit Paul Sabatier, 2007.

[BV04]   Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[CDE⁺06]   Yann Chevaleyre, Paul E. Dunne, Ulle Endriss, Jérôme Lang, Michel Lemaître, Nicolas Maudet, Julian A. Padget, Steve Phelps, Juan A. Rodríguez-Aguilar, and Paulo Sousa. Issues in multiagent resource allocation. *Informatica (Slovenia)*, 30(1):3–31, 2006.

[DGP09]   Constantinos Daskalakis, Paul W. Goldberg, and Christos H. Papadimitriou. The complexity of computing a nash equilibrium. *Commun. ACM*, 52(2):89–97, 2009.

[Gin10]   Herbert Gintis. Social norms as choreography. *politics, philosophy & economics*, 9(3):251–264, 2010.

[MLE11]   Javier Morales, Maite López-Sánchez, and Marc Esteva. Using experience to generate new regulations. In Toby Walsh, editor, *IJCAI 2011, Proceedings of the 22nd International Joint Conference on Artificial Intelligence, Barcelona, Catalonia, Spain, July 16-22, 2011*, pages 307–312. IJCAI/AAAI, 2011.

[MLR⁺15] Javier Morales, Maite López-Sánchez, Juan Antonio Rodríguez-Aguilar, Michael Wooldridge, and Wamberto Weber Vasconcelos. Synthesising liberal normative systems. In Gerhard Weiss, Pinar Yolum, Rafael H. Bordini, and Edith Elkind, editors, *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems, AAMAS 2015, Istanbul, Turkey, May 4-8, 2015*, pages 433–441. ACM, 2015.

[PR08] Christos H. Papadimitriou and Tim Roughgarden. Computing correlated equilibria in multi-player games. *J. ACM*, 55(3), 2008.

[Raw71] John Rawls. *A Theory of Justice*. Oxford University Press, 1971.

[SA07] Sandip Sen and Stéphane Airiau. Emergence of norms through social learning. In Manuela M. Veloso, editor, *IJCAI 2007, Proceedings of the 20th International Joint Conference on Artificial Intelligence, Hyderabad, India, January 6-12, 2007*, pages 1507–1512, 2007.

[Sky14] Brian Skyrms. *Evolution of the Social Contract*. Cambridge University Press, 2014.

[ST96] Yoav Shoham and Moshe Tennenholtz. On social laws for artificial agent societies: Off-line design. *Artif. Intell.*, 73(1-2):231–252, 1996.

[ST97] Yoav Shoham and Moshe Tennenholtz. On the emergence of social conventions: Modeling, analysis, and simulations. *Artif. Intell.*, 94(1-2):139–166, 1997.

[vdHRW07] Wiebe van der Hoek, Mark Roberts, and Michael Wooldridge. Social laws in alternating time: effectiveness, feasibility, and synthesis. *Synthese*, 156(1):1–19, 2007.