Mechanism Design for Social Law Synthesis under Incomplete Information

(Extended Abstract)

Jun Wu, Lei Zhang, Chongjun Wang*, and Junyuan Xie
State Key Laboratory for Novel Software Technology at Nanjing University
163 Xianlin AVE, Nanjing, China
{wujun,zhangl,chjwang,jyxie}@nju.edu.cn

ABSTRACT

For the social law synthesis problem, when the agents are rational in the sense of game theory and hold some information we need as private information, it naturally evolves into a setting that is perfectly addressed by the framework of algorithmic mechanism design. In this strategic setting, we are not only required to find out the feasible social law for the objective, but also required to formulate the right payment to the agents to induce incentive compatibility and individual rationality. We design a mechanism for this setting, prove that it satisfies all the required formal properties, and characterize the conditions for the existence of feasible mechanisms. Moreover, we show that the upper-bound of the total payment of the proposed mechanism is high.

Keywords

social laws, logic, mechanism design, optimization

1. INTRODUCTION

Social law was extensively studied in the past decades as an off-line approach for coordinating multiagent systems by e.g., [16, 17, 19, 21, 1]. In general, a social law is a set of restrictions on the available actions of agents. By imposing these restrictions, it is hoped that some desirable objectives will emerge [2]. It is easy to see that the strategic setting brings about some fundamental challenges to social law synthesis: firstly, the agents are autonomous and self-interested. They choose to comply with the social law if and only if it is profitable. So, it is necessary to take the *gains* and *costs* of every agent into consideration in the implementation of a social law, rather than naively treat it as hard constraints. Basically, we can try to pay each agent a proper amount which guarantees a positive profit for it. Unfortunately, as the gains and costs of an agent may be its private information, we are not sure about how much we should pay each agent; secondly, modifying the original system as little as possible, *i.e., minimality* [9, 10], is sometimes a desired property of social laws. In the seminal papers, minimality is measured by the total number of available actions deleted by the social law. More generally, it can be defined as the total cost of the deleted actions. Moreover, we can also formalize some other optimality concepts related to the gains and costs information of the agents, but when this information is privately held by the agents, we are not sure about which of the social laws is the best.

The above facts mean some private information of the agents should be properly elicited and taken into consideration during social law synthesis. This issue can be naturally handled by the framework of algorithmic mechanism design [13, 14, 15]. Our approach is to extend the logic-based framework of social laws [19] by adding game theoretical components that capture the utilities of the agents, then we can propose a social law auction based on Vickrey-Clarke-Groves (VCG) mechanisms [20, 7, 11]. This paper can be seen as an attempt to introduce the methodology of algorithmic mechanism design into the traditional logic-based approach to AI.

2. FORMAL FRAMEWORK

The interactions of a multiagent system Ag with k agents, a state space Q in which each state q is labeled by a set of propositions $\pi(q) \subseteq \Pi$ can be modeled by a Concurrent Game Structure (CGS) $S = \langle k, Q, \Pi, \pi, \varepsilon, \delta \rangle$, in which for each state $q \in Q$ the available actions for each agent i is specified as a non-empty set $\varepsilon_i(q)$ and a state $\delta(q, j_1, ..., j_k) \in Q$ will be the next state if every agent i chooses action $j_i \in \varepsilon_i(q)$. We adopt Alternating-time Temporal Logic (ATL) for specifying and verifying CGS. The language of ATL is generated by BNF:

$$\varphi ::= p|\neg\varphi|\varphi_1 \vee \varphi_2|\langle\langle A\rangle\rangle\bigcirc\varphi|\langle\langle A\rangle\rangle \Box \varphi|\langle\langle A\rangle\rangle\varphi_1\mathcal{U}\varphi_2,$$

where $p \in \Pi$, and $A \subseteq Ag = \{1, ..., k\}$ is an agent coalition. A social law for S is a behavioral constraint η defined as a function $\forall i \in Ag, q \in Q : \eta_i(q) \subset \varepsilon_i(q)$. The new structure obtained by implementing η on S, denoted $S\dagger\eta$, is the structure $S' = \langle k, Q, \Pi, \pi, \varepsilon', \delta' \rangle$ where $\forall i \in Ag, q \in Q : \varepsilon'_i(q) = \varepsilon_i(q) \setminus \eta_i(q)$, and δ' is obtained from δ by restricting the action profiles to $\varepsilon'_1(q) \times ... \times \varepsilon'_k(q)$ in every q. We denote the set of all the possible social laws for S as $\mathcal{SL}_S$, and denote the set of all the possible CGS that can be obtained from S by implementing a social law as $\mathcal{W}_S = \{S' \mid \exists\eta \in \mathcal{SL}_S : S\dagger\eta = S'\}$. The underlying economic model can be specified as follows.

With respect to each agent, a restriction on its available actions can result in a *cost* [9, 10], and the structural properties of the obtained new CGS can result in a

*Chongjun Wang is the corresponding author.

Appears in: *Proc. of the 16th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2017), S. Das, E. Durfee, K. Larson, M. Winikoff (eds.), May 8–12, 2017, São Paulo, Brazil.*

change in its *structural valuation*, which manifests the agent's preference on different system properties [2, 5]. The cost of a social law $\eta \in \mathcal{SL}$ to an agent i with unit cost c_i is $\sigma_i(c_i, \eta) = c_i \cdot \sum_{q \in Q} |\eta_i(q)|$, where $\sum_{q \in Q} |\eta_i(q)|$ is the number of i's actions restricted by the social law η. The *structural valuations* of an agent i on a CGS can be defined based on a fixed *feature set* $\mathbb{F}_i$ of state-formula pairs of the form $\langle q, \varphi \rangle$, which is called an *objective* expressing its aimed properties with respect to some of the states, and an agent's structural valuation on a social law equals its structural value gain (which sometimes can be a negative number) on obtaining the new CGS. We assume each agent $i \in Ag$ has $N_i \in \mathbb{N}$ different *structural types*, denoted as the set $\{1, ..., N_i\}$, and each type $j \in \{1, ..., N_i\}$ is specified by a list of the form $\langle (q_i^j[1], \varphi_i^j[1], x_i^j[1]), ..., (q_i^j[l_i^j], \varphi_i^j[l_i^j], x_i^j[l_i^j]) \rangle$ where $l_i^j \leq |\mathbb{F}_i|+1$, $\forall k \in \{1, ..., l_i^j - 1\} : (q_i^j[k], \varphi_i^j[k]) \in \mathbb{F}_i$, $x_i^j[k] \in R^+$; and each tuple in $\mathbb{F}_i$ can appear at most once in the list. Moreover, we assume $\varphi_i^j[l_i^j] = \top$, $x_i^j[l_i^j] = 0$, $q_i^j[l_i^j]$ is an arbitrary state; and $x_i^j[1] \geq ... \geq x_i^j[l_i^j]$. Given a CGS S, if agent i's structural valuation type is j, then for each CGS $G \in \mathcal{W}_S$, its structural value to agent i is $e_i(j, G) = x_i^j[k]$ where k is the smallest index in $\{1, ..., l_i^j\}$ satisfying $G, q_i^j[k] \models \varphi_i^j[k]$; for each social law $\eta \in \mathcal{SL}_S$, its structural value to agent i is $e_i(j, \eta) = e_i(j, S \dagger \eta) - e_i(j, S)$. Then the value of a social law $\eta \in \mathcal{SL}_S$ to an agent i with unit cost c_i and structural type j_i is $v_i(c_i, j_i, \eta) = e_i(j_i, S \dagger \eta) - e_i(j_i, S) - c_i \cdot \sum_{q \in Q} |\eta_i(q)|$, i.e., the structural value minus the cost.

A *type* of agent i can be denoted as $\theta_i = \langle c_i, j_i \rangle$ where $c_i \in \mathbb{R}^+$ and $j_i \in \{1, ..., N_i\}$. The type space of agent i is $\Theta_i = \mathbb{R}^+ \times \{1, ..., N_i\}$. While each agent's actual type is assumed to be their private information, the type space of each agent is assumed to be public information. A social law auction mechanism is a tuple of allocation function and payment function $\langle a, t \rangle$, where $a : \Theta_1 \times ... \times \Theta_k \to \mathcal{SL}_S$, $t_i : \Theta_1 \times ... \times \Theta_k \to \mathbb{R}$, and Θ_i is the type space of agent i. A social law auction proceeds in 3 steps: firstly it is announced and asks every agent to bid their type simultaneously; then each agent chooses a type from its type space and submits it at the same time with the others; and finally according to the collected bids, a social law is selected and the amount of payment to each agent is determined.

A selected social law η can be mapped to a set of relevant agents ξ_η. Agent $i \in \xi_\eta$, if and only if at least one of the following two conditions hold: 1) $\exists q \in Q : \eta_i(q) \neq \varnothing$; 2) $\exists \langle q, \varphi \rangle \in \mathbb{F}_i : S, q \models \neg\varphi$ *iff* $S \dagger \eta, q \models \varphi$. So, a relevant agent is actually an agent whose value (= *structural value* − *cost*) is changed by the social law. We require the mechanisms to be *normalized*, that is, if an agent is not in ξ_η, then the payment to it is 0. Obviously, each agent outside the relevant set will get a 0 utility, and each agent i in the relevant set will finally get a utility $u_i(\theta_i^*, \theta_i, \theta_{-i}) = e_i(j_i^*, S \dagger \eta) - e_i(j_i^*, S) - c_i^* \cdot \sum_{q \in Q} |\eta_i(q)| + t_i(\theta_i, \theta_{-i})$, where $\theta_i^* = \langle c_i^*, j_i^* \rangle \in \Theta_i$ is its true type, $\theta_i \in \Theta_i$ is its bidded type and $\theta_{-i} \in \Theta_{-i}$ is the bid vector of all the other agents.

We say a social law auction mechanism $\langle a, t \rangle$ is *feasible* for an objective $\omega = \langle q, \varphi \rangle$ if all of the following items hold: 1) (*Effective*) $\forall \vec{\theta} \in \Theta_1 \times ... \times \Theta_k : S \dagger a(\vec{\theta}), q \models \varphi$, *i.e.*, always allocates an effective social law fulfilling the objective; 2) (*Incentive Compatible*) $\forall i$, $\forall \theta_i, \theta_i' \in \Theta_i$, $\forall \theta_{-i} \in \Theta_{-i}$: $u_i(\theta_i, \theta_i, \theta_{-i}) \geq u_i(\theta_i, \theta_i', \theta_{-i})$, *i.e.*, to bid truthfully is the dominant strategy; 3) (*Individual Rationality*) $\forall i$, $\forall (\theta_i, \theta_{-i}) \in \Theta$: $u_i(\theta_i, \theta_i, \theta_{-i}) \geq 0$, *i.e.*, every agent will get a non-negative utility.

3. MECHANISM DESIGN

Given a CGS S, an objective $\omega = \langle q, \varphi \rangle$, and a bid vector $\langle c_1, j_1 \rangle, ..., \langle c_k, j_k \rangle$, we firstly find out the set $\mathcal{F}_{S,\omega} \subseteq \mathcal{SL}_S$ of all the effective social laws; then allocate the social law in $\mathcal{F}_{S,\omega}$ with the maximal social welfare, and pay each agent i in the relevant set ξ_η the amount of the social welfare gain of all the other agents brought about by it, assuming the bids from all the agents are their true types.

Given the type vector $\vec{\theta}$, the social welfare of a set of agents $A \subseteq Ag$ with respect to a social law $\eta \in \mathcal{SL}_S$ is $SW_{\vec{\theta}}(A, \eta) = \sum_{i \in A} [e_i(j_i, S \dagger \eta) - e_i(j_i, S) - c_i \cdot \sum_{q \in Q} |\eta_i(q)|]$. After obtaining $\mathcal{F}_{S,\omega}$ (which has already been studied by the literature on social law synthesis, *e.g.*, [19, 21]), for any bid profile $\vec{\theta} \in \Theta_1 \times ... \times \Theta_k$, the allocation can be computed as:

$$a(\vec{\theta}) = \arg \max_{\eta \in \mathcal{F}_{S,\omega}} SW_{\vec{\theta}}(Ag, \eta) \tag{1}$$

For an arbitrary agent $i \in Ag$ in the relevant set $\xi_{a(\vec{\theta})}$, we can compute the payment as

$$t_i(\vec{\theta}) = SW_{\vec{\theta}}(\bar{i}, a(\vec{\theta})) - \max_{\eta \in \mathcal{F}_{S,\omega}} [v_i(+\infty, j_i^\infty, \eta) + SW_{\theta_{-i}}(\bar{i}, \eta)] \tag{2}$$

We call the mechanism composed by the allocation function defined by equation 1 and the payment function defined by equation 2 as the *VCG Social Law Auction Mechanism* (VCG-SLA). Moreover, the following results can be proved.

THEOREM 1. *Given a* CGS S*,* VCG-SLA *is feasible for the objective* ω *iff* $\mathcal{F}_{S,\omega}$ *is nonempty and monopoly-free.*

Note that, *monopoly-free* means for every agent i there is always a social law according to which i is not in the relevant set, *i.e.*, $\forall i \in Ag, \exists \eta \in \mathcal{F}_{S,\omega} : i \notin \xi_\eta$.

COROLLARY 2. *There exists a feasible social law mechanism if and only if* VCG-SLA *is feasible.*

Therefore, the proposed mechanism has soundly and completely solved the social law synthesis problem for rational agents defined in this paper. Finally, the high total payment of VCG-SLA can be manifested by the following result.

PROPOSITION 3. *There are instances where the best social welfare* $V \geq 0$ *and the total amount of payment is* $(|\xi_{a(\vec{\theta})}| - 1)V$*, and instances where the second best social welfare* $V' \leq 0$ *and the total amount of payment is* $-\Theta(|\xi_{a(\vec{\theta})}|)V'$*.*

So the proposed setting potentially relates to and introduces an interesting new problem to the research of frugal mechanisms [3, 18, 4, 8, 12, 22, 6]. Moreover, it is also interesting to further study the general case where the optimization objective is not simply social welfare, the cost of each action is different and the costs of sets of actions are not simply the sum of the cost of the individual actions; Finally, since the computation of the proposed VCG social law mechanism is intractable, it is interesting to find out tractable approximation mechanisms with good properties.

Acknowledgment

The authors are grateful to the reviewers for their helpful comments. This paper is supported by the National Key Research and Development Program of China (Grant No. 2016YFB1001102), the National Natural Science Foundation of China (Grant No. 61375069, 61403156, 61502227), and the Collaborative Innovation Center of Novel Software Technology & Industrialization at Nanjing University.

REFERENCES

[1] T. Ågotnes, W. van der Hoek, and M. Wooldridge. Conservative social laws. In *Proc. of 20th European Conference on Artificial Intelligence (ECAI-12)*, pages 49–53, 2012.

[2] T. Ågotnes, M. Wooldridge, and W. van der Hoek. Normative system games. In *Proc. of 6th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-07)*, pages 876–883, 2007.

[3] A. Archer and E. Tardos. Frugal path mechanisms. In *Proc. of 13th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA-02)*, pages 991–999, 2002.

[4] A. Archer and E. Tardos. Frugal path mechanisms. *ACM Transactions on Algorithms*, 3(1):1–19, 2004.

[5] N. Bulling and M. Dastani. Verifying normative behaviour via normative mechanism design. In *Proc. of 22nd International Joint Conference on Artificial Intelligence (IJCAI-11)*, pages 103–108, 2011.

[6] G. Calinescu. Bounding the payment of approximate truthful mechanisms. *Theoretical Computer Science*, 562(C):419–435, 2015.

[7] E. H. Clarke. Multipart pricing of public goods. *Public Choice*, 11(11):17–33, 1971.

[8] E. Elkind, A. Sahai, and K. Steiglitz. Frugality in path auctions. In *Proc. of 15th ACM-SIAM Symposium on Discrete Algorithms (SODA-04)*, pages 701–709, 2004.

[9] D. Fitoussi and M. Tennenholtz. Minimal social laws. In *Proc. of 15th AAAI Conference on Artificial Intelligence (AAAI-98)*, pages 26–31, 1998.

[10] D. Fitoussi and M. Tennenholtz. Choosing social laws for multi-agent systems: Minimality and simplicity. *Artificial Intelligence*, 119:61–101, 2000.

[11] T. Groves. Incentives in teams. *Econometrica*, 41:617–631, 1973.

[12] A. R. Karlin and D. Kempe. Beyond VCG: Frugality of truthful mechanisms. In *Proc. of 46th Annual Symposium on Foundations of Computer Science (FOCS-05)*, pages 615–626, 2005.

[13] N. Nisan and A. Ronen. Algorithmic mechanism design. In *Proc. of 31st annual ACM symposium on Theory of computing (STOC-99)*, pages 129–140. ACM, 1999.

[14] N. Nisan and A. Ronen. Algorithmic mechanism design. *Games and Economic Behavior*, 35:129–140, 2001.

[15] N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani. *Algorithmic game theory*, volume 1. Cambridge University Press Cambridge, 2007.

[16] Y. Shoham and M. Tennenholtz. On the synthesis of useful social laws for artificial agent societies. In *Proc. of 9th AAAI Conference on Artificial Intelligence (AAAI-92)*, pages 276–281, 1992.

[17] Y. Shoham and M. Tennenholtz. On social laws for artificial agent societies: Off-line design. In *P.E. Agre and S.J.Rosenschein (Eds.),Computational Theories of Interaction and Agency*, pages 597–618, MIT Press,Cambridge, MA, 1996.

[18] K. Talwar. The price of truth: Frugality in truthful mechanisms. In *Proc. of 20th Annual Symposium on Theoretical Aspects of Computer Science (STACS-03)*, pages 608–619, 2003.

[19] W. van der Hoek, M. Roberts, and M. Wooldridge. Social laws in alternating time: Effectiveness feasibility, and synthesis. *Synthese*, 156:1–19, 2007.

[20] W. Vickrey. Counterspeculation, auctions, and competitive sealed tenders. *Journal of Finance*, 16(1):8–37, 1961.

[21] J. Wu, C. Wang, and J. Xie. A framework for coalitional normative systems. In *Proc. of 10th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-11)*, 2011.

[22] D. Zhao, H. Ma, and L. Liu. Frugal online incentive mechanisms for crowdsourcing tasks truthfully. *CoRR abs/1404.2399*, 2014.