Synthesizing Optimal Social Laws for Strategical Agents via Bayesian Mechanism Design

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

When rational behavior of the agents and private information are considered, the optimal social law synthesizing problem naturally evolves into a setting which can be handled by the framework of algorithmic mechanism design. We focus on the Bayesian case in this paper, that is, the probability distribution of each agent's cost is known. It is easy to see that in this case our problem closely relates to path/spanning-tree auctions and Myerson's optimal auction mechanism, but the optimization objective is new, that is, we focus on profit maximization instead of payment maximization. By studying this problem: we further extend the logic-based framework of social law optimization problem to the strategic case, and show that it becomes a new problem of algorithmic mechanism design; we find out a mechanism that is incentive compatible, individually rational and maximizes the expected profit for all input cost profiles; however, we can show that this mechanism is computational intractable; so, we finally find out a tractable constant-factor approximation mechanism.

CCS Concepts

•**Computing methodologies → Multi-agent systems;**

Keywords

social laws, logic, normative systems, mechanism design, optimization

1. INTRODUCTION

Social law was initially proposed by Shoham and Tennenholtz [22, 23] as an off-line approach for coordinating multiagent systems, and then extended by a lot of follow-up work, e.g., [26, 27, 1], based on introducing the formal systems of modal and temporal logics. Although the various approaches to social laws proposed in the literature differ on technical details, they all share the same basic intuition that 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 [3]. The purpose is typically to prevent some destructive interactions from taking place, or to facilitate some positive interactions. In this sense, social laws also have much in common with normative systems [17, 7], which is also an important direction in AI research.

Ågotnes and Wooldridge [2] further extended the framework of social laws by enabling expressing multiple differently valued objectives and taking into account the cost of implementing each action restriction, and investigated the problem of optimal social law synthesis defined as to find out the social law that maximizes the difference between the total value of the fulfilled objectives and the total implementation cost. They showed that in general the optimal social law synthesis is intractable, but several tractable subcase can be found out. Moreover, they showed that this problem can be solved by integer programing.

A phenomenon that has attracted a lot of interest in recent years but is generally absent in the original framework of social laws is the information incompleteness and rational behavior of the agents. Such multiagent systems abstract the internal dynamics of many real world distributed systems which can be modeled as non-cooperative games played by a set of agents, all of which always try to act to maximize their individual utilities. This setting brings about some fundamental challenges to social law synthesis: each agent's actual cost on restricting her actions is privately known by herself; each agent decides to obey the social law only if she can obtain a payment that is higher than her cost; and if we decide to obtain these cost values by "asking" the agents, they may misreport it strategically to achieve higher utilities. So, in such a strategic case, obtaining the optimal social law depends crucially on whether we can correctly elicit the private cost information from the agents and take into consideration during social law synthesis.

The above issue naturally falls into the domain of algorithm mechanism design [19, 20, 21], which is recently an active direction in the intersection of computer science and game theory and aims to solve optimization problems in strategic cases just like the one considered in this paper. Basically, we can extend the framework of optimal social laws proposed in [2] to a procurement auction, in which firstly a mechanism (consisting of an *allocation rule* and a *payment rule*) is announced, then every agent reports its cost value, and finally a set of restricted actions (*i.e.,* a social law) is selected and a payment to each agent is decided by applying the allocation rule and payment rule. We aim to find out the mechanism that *reliably* outputs the social law which maximizes the profit defined as the difference between the total value of the fulfilled objectives and the total payment.

*Chongjun Wang is the corresponding author.

Appears in: *Proceedings 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.*

We focus on the Bayesian case in this paper, that is, the probability distribution of each agent's cost is known. In this case our problem closely relates to path auctions [4, 25, 5, 11, 16, 29, 8, 28], spanning-tree auctions [6, 13, 25] and Myerson's optimal auction mechanism [18], but the optimization objective is new, *i.e.*, we focus on *profit maximization* instead of *payment maximization.* The main contributions of this paper can be summarized as follows:

1) We further extend the logic-based framework of social law optimization problem [26, 2] to the strategic case, and show that it becomes a new problem of algorithmic mechanism design – social law auction design;

2) We find out a mechanism that is incentive compatible, individually rational and maximize the expected profit for all possible cost profiles; However, we can show that this mechanism is computational intractable;

3) We then find out a constant-factor approximation mechanism that is incentive compatible, individually rational and computationally tractable.

This paper can be seen as an attempt to introduce the methodology of algorithmic mechanism design into the traditional logic-based approach to artificial intelligence. We have obtained a framework that not only greatly improves the reliability and robustness of social laws, but also discovers some interesting new problems for the further study of algorithmic mechanism design.

The remainder of this paper is structured as follows. We start with some background on social laws and algorithmic mechanism design as well as the formal framework of our work. Next, we present an optimal social law auction and show it is computationally intractable. Then, we present a tractable 2-approximation social law auction. Afterwards, we discuss some related work. Finally, we present some conclusions and introduce some problems for future study.

2. BACKGROUND AND PRELIMINARIES

We give a very brief summary on the formal framework of social laws adopted in [2], and then extend it to the one that will be applied in our work.

2.1 Model of Multiagent Systems

The interaction of the agents can be specified by a weighted Kripke structure, which extend conventional Kripke structure for branching-time logic (see e.g.,[12]) with costs. Formally, a weighted Kripke structure (over a proposition set Φ) is a 7-tuple $K = \langle S, s_0, R, A, \alpha, c, \pi \rangle$ where:

- S is a finite, non-empty set of states;
- $s_0 \in S$ is the initial state;
- $R \subseteq S \times S$ is a total (i.e., for every $s \in S$ there is a $t \in S$ such that $(s, t) \in R$) relation on S, which we refer to as the transition relation;
- $A = \{1, ..., n\}$ is a set of agents;
- $\alpha : R \to A$ labels each transition in R with an agent;
- $c : R \to \mathbb{R}^+$ is a cost function; and
- $\pi : S \to 2^\Phi$ is a valuation function.

Intuitively, a weighted Kripke structure describes the state transitions of a multiagent system where each state $s \in S$ is labeled with a set of propositions $\pi(s)$ that are true in it, each state transition $\tau = (s, t) \in R$ is controlled by an agent $\alpha(\tau)$, and forbidding the transition τ will induce a cost of value $c(\tau)$ to us.

2.2 Computation Tree Logic (CTL)

CTL is a branching-time logic intended for representing the properties of Kripke structures. Since CTL is widely documented in the literature, we only introduce the version of CTL syntax adopted in our work and the informal meaning of the operators. The full version of CTL syntax and semantics can be found in [12].

The syntax of CTL can be defined by the following BNF grammar, where $p \in \Phi$:

$$\varphi ::= p \mid \neg\varphi \mid \varphi_1 \vee \varphi_2 \mid \mathsf{E} \bigcirc \varphi \mid \mathsf{E}(\varphi_1 \mathcal{U} \varphi_2) \mid \mathsf{E}\Box\varphi \qquad (1)$$

The semantics of CTL can be given with respect to the satisfaction relation "$\models$", which holds between pairs of the form K, s (where K is a Kripke structure and s is a state in K), and formula:

- $K, s \models p$ iff $p \in \pi(s)$;
- $K, s \models \neg\varphi$ iff not $K, s \models \varphi$;
- $K, s \models \varphi_1 \vee \varphi_2$ iff $K, s \models \varphi_1$ or $K, s \models \varphi_2$;
- $K, s \models \mathsf{E} \bigcirc \varphi$ iff $\exists \lambda \in paths(s)$: $K, \lambda[1] \models \varphi$;
- $K, s \models \mathsf{E}(\varphi_1 \mathcal{U} \varphi_2)$ iff $\exists \lambda \in paths(s), \exists u \in \mathbb{N}$, s.t. $K, \lambda[u] \models \varphi_2$ and $\forall v, (0 \le v < u)$: $K, \tau[v] \models \varphi_1$;
- $K, s \models \mathsf{E}\Box\varphi$ iff $\exists \lambda \in paths(s)$: $K, \lambda[i] \models \varphi$ for all i.

Note that, $paths(s)$ denotes the set of paths started from state s, and $\lambda[i]$ denotes the $(i + 1)$'s state on path λ. E is a path selection operator intuitively means "there exists a path such that ...", and $\bigcirc, \mathcal{U}$ and $\Box$ are temporal operators intended to capture the notions "next-time", "until" and "always" respectively. For example, $K, s \models \mathsf{E}\varphi_1 \mathcal{U} \varphi_2$ means "there is a path from s such that φ_1 is satisfied in the states along the path until arriving at a state satisfying φ_2.

2.3 Optimal Social Laws

A social law defines, for each possible system transition, whether or not it is legal. Formally, a social law η (w.r.t. a Kripke structure $K = \langle S, s_0, R, A, \alpha, c, \pi \rangle$) is a subset of R, such that $R \setminus \eta$ is a total relation. The latter is a *reasonableness constraint*: it prevents social laws which lead to states with no successor. Let $N(R^a) = \{\eta : (\eta \subseteq R^a)$ *and* $(R \setminus \eta$ is total$)\}$ be the set of social laws over $R^a \subseteq R$. The intended interpretation of a social law η is that $(s, s') \in \eta$ means transition (s, s') is forbidden in the context of η; hence $R \setminus \eta$ denotes the legal transitions of η. The effect of implementing a social law on a Kripke structure is to eliminate from it all transitions that are forbidden according to this social law. If K is a Kripke structure, and η is a social law over K, then $K \dagger \eta$ denotes the Kripke structure obtained from K by deleting transitions forbidden in η. Formally, if $K = \langle S, s_0, R, A, \alpha, c, \pi \rangle$ and $\eta \in N(R)$, then $K \dagger \eta = K'$ is the Kripke structure $K' = \langle S, s_0, R', A, \alpha, c, \pi \rangle$ such that $R' = R \setminus \eta$ and all other components are as in K.

We denote by $\mathcal{K}$ the set of Kripke Structures that may be obtained by implementing some social law on K, i.e.,

$$\mathcal{K} = \{\langle S, s_0, R', A, \alpha, c, \pi\rangle : R' \subseteq R \text{ and } R' \text{ is total }\}. \quad (2)$$

A preference relation of the social law designer can be defined on $\mathcal{K}$ by a valuation function:

$$v : \mathcal{K} \to \mathbb{R}^+ \quad (3)$$

Then the utility of a social law η with respect to a Kripke structure K and valuation function v, which is denoted by $u(\eta, K, v)$, is then the difference between the value brought by the social law and the cost of implementing it, i.e.,

$$u(\eta, K, v) = v(K\dagger\eta) - \sum_{(s,s')\in\eta} c(s, s') \quad (4)$$

So, from the point of view of a designer with valuation function v, the optimal social law $\eta^*(K, v)$ with respect to Kripke structure K and valuation function v will be the one that maximizes the value of the function u:

$$\eta^*(K, v) = \arg\max_{\eta\in N(R)} u(\eta, K, \eta) \quad (5)$$

The OPTIMAL SOCIAL LAW problem is the problem of computing $\eta^*(K, v)$. Ågotnes and Wooldrige [2] showed that every valuation function that doesn't discern between bisimulation equivalent structures (i.e., assigns the same value to such structures) can be compactly and equivalently represented as a feature set: $\mathcal{F} = \{(\varphi_1, a_1), ..., (\varphi_k, a_k)\}$ which defines the following valuation function:

$$v_{\mathcal{F}}(K') = \sum_{(\varphi_i,x_i)\in\mathcal{F},K',s_0\models\varphi_i} a_i \quad (6)$$

and if we define v by this approach, the The OPTIMAL SOCIAL LAW problem is FP^{NP}-complete, and can be solved by integer programming.

2.4 Economic Model

We adopt the above formal framework of social laws, but

1) we focus on the case where the α function of the weighted Kripke structure is a bijection function. So, each transition $(s, s') \in R$ can be uniquely enumerated by the index i of the agent who controls it;

2) For each transition $i \in R$ we reinterpret $c(i)$ as the cost that will be brought to agent i if this transition is forbidden. We assume $c(i)$ to be known only by agent i, but as public information the cost of the agents are independent continuous random variables $\xi_1, \cdots, \xi_n$, where each ξ_i is drawn from the interval $\Xi_i = [0, \omega_i]$ subject to a probability density function f_i.

We use the following usual notations for vectors, *e.g.*, $\xi = (\xi_1, \cdots, \xi_n)$, $\xi_{-i} = (\xi_1, \cdots, \xi_{i-1}, \xi_{i+1}, \cdots, \xi_n)$, and $(\xi_i, \xi_{-i}) = \xi$. We denote by $\Xi = \prod_{i\in A} \Xi_i$ the space of all possible cost profiles. For each $\xi \in \Xi$, we have the joint probability density $f(\xi) = \prod_{i\in A} f_i(\xi_i)$ by the independence assumption.

We run a procurement auction to determine the selected social law: firstly we announce an auction mechanism consisting of an allocation function $\hat{H} : \Xi \to N(R)$ and a payment function $P_i : \Xi \to \mathbb{R}^+$ for each agent i, then collect the cost reports from the agents as their bids and obtain the bid profile ξ, and finally select the social law $\hat{H}(\xi)$ and pay each agent i the amount $P_i(\xi)$. Note that, the allocation function can be equivalently specified as for each agent i and each bid profile ξ, $H_i(\xi) = 1$ (which means agent i is selected) if $i \in \hat{H}(\xi)$, otherwise $H_i(\xi) = 0$.

After a mechanism $\langle H, P\rangle$ is determined, it applies to any instances of agent set A with the real costs randomly drawn from Ξ, the interactions of the agents intrinsically become a Bayesian game $\langle A, (\Xi_i, f_i, B_i, u_i)_{i\in A}\rangle$ where

- Ξ_i denotes not only the cost space but also the action space for agent i, since the actions of each agent are restricted to cost reports, *i.e.*, each $\xi_i \in \Xi_i$ also denotes the action "reporting ξ_i";

- f_i is the probability density function of agent i's cost;

- B_i is the strategy space of agent i containing all the functions of the form $b : \Xi_i \to \Xi_i$, which means agent i can report her cost strategically;

- $u_i : \Xi_i \times \Xi \to \mathbb{R}$ is the utility function of agent i. When agent i's cost value is ξ_i, the bid profile ξ will bring to her the utility:
$$u_i(\xi_i, \xi) = P_i(\xi) - \xi_i \cdot H_i(\xi) \quad (7)$$

Bayesian-Nash Equilibrium (BNE) is a natural solution concept for such games. Intuitively, a strategy profile $(b_1, \cdots, b_n)$ is a BNE means for each agent i, using the strategy $b_i(\xi_i)$ will achieve the highest expected utility for her, no matter which real cost $\xi_i \in \Xi_i$ she takes now. Formally, the expected utility achieved by agent i whose cost value is $\xi_i \in \Xi_i$ via bidding $\xi'_i \in \Xi_i$ is

$$\bar{u}_i(\xi_i, \xi'_i) = \mathbb{E}_{\xi_{-i}\in\Xi_{-i}}[u_i(\xi_i, (\xi'_i, \xi_{-i}))] \quad (8)$$

The strategy profile $(b_1, \cdots, b_n)$ is a BNE if and only if

$$\forall i, \xi_i, b'_i \neq b_i : \bar{u}_i(\xi_i, b(\xi_i)) \geq \bar{u}_i(\xi_i, b'(\xi_i)) \quad (9)$$

Given this solution concept, the task of mechanism design is actually to find out a proper mechanism, so that, there is a BNE and for any instance the outcome of the BNE satisfies the objective of the designer. A mechanism is called Bayesian-Nash Incentive Compatible (BNIC) if and only if truthful bidding, *i.e.*, $b_i(\xi_i) = \xi_i$ for all i and ξ_i, is a BNE. According to the famous *Revelation Principle* [20] we can greatly reduce the search space of mechanisms and only focus on BNIC mechanism without loss of generality.

If agent i bids ξ'_i, and let $h_i(\xi'_i)$ and $p_i(\xi'_i)$ be the expectation of transition i being selected and the expected amount of payment given to agent i respectively, then

$$h_i(\xi'_i) = \int_{\Xi_{-i}} H_i(\xi'_i, \xi_{-i}) f_{-i}(\xi_{-i}) d\xi_{-i} \quad (10)$$

$$p_i(\xi'_i) = \int_{\Xi_{-i}} P_i(\xi'_i, \xi_{-i}) f_{-i}(\xi_{-i}) d\xi_{-i} \quad (11)$$

By equations (8)(7)(10) and (11) we can further obtain

$$\bar{u}_i(\xi_i, \xi'_i) = p_i(\xi'_i) - h_i(\xi'_i)\xi_i \quad (12)$$

So if agent i truthfully bids her cost ξ_i, she will get the expected utility

$$\hat{u}_i(\xi_i) = \bar{u}_i(\xi_i, \xi_i) = p_i(\xi_i) - h_i(\xi_i)\xi_i \quad (13)$$

Implementing a social law by making every agent voluntarily obey the restrictions rather than by "hard constraints" is one of the challenge issues in social law research. In our framework, this requirement can be naturally captured by the game theoretical concept *individually rationality* (IR) which means every agent will get a nonnegative expected utility after the implementation of the social law. Since we focus on BNIC mechanisms, IR is equivalent to nonnegative expected truthful-bidding utilities, *i.e.*,

$$\forall i \in A, \xi_i \in \Xi_i : \hat{u}_i(\xi_i) \geq 0 \tag{14}$$

If all the agents truthfully bid their costs, then for arbitrary real cost profile ξ drawn from Ξ, the bid profile will also be ξ, and the selected social law and the payment to agent i will be $\hat{H}(\xi)$ and $P_i(\xi)$ respectively. For the designer with a feature set $\mathcal{F} = \{(\varphi_1, x_1), ..., (\varphi_k, x_k)\}$, the value of the Kripke structure obtained by implementing the selected social law is $v_{\mathcal{F}}(K \dagger \hat{H}(\xi))$, the total payment is $\sum_{i \in A} P_i(\xi)$ and therefore the achieved profit is

$$\sigma(\xi) = v_{\mathcal{F}}(K \dagger \hat{H}(\xi)) - \sum_{i \in A} P_i(\xi) \tag{15}$$

So, the expected profit of the designer with respect to the cost profile space Ξ is

$$\mathbb{E}_{\xi \in \Xi}[\sigma(\xi)] = \int_{\Xi} \Big(v_{\mathcal{F}}(K \dagger \hat{H}(\xi)) - \sum_{i \in A} P_i(\xi) \Big) f(\xi) d\xi \tag{16}$$

The aim of this paper is to find out the BNIC and IR mechanism $\langle H^*, P^* \rangle$ that maximizes the expected profit.

3. OPTIMAL SOCIAL LAW AUCTIONS

First of all, we can show the requirement of BNIC can be equivalently transformed to some constraints on the allocation function and payment function.

LEMMA 1. *A mechanism $\langle H, P \rangle$ is* BNIC *iff* $\forall\, i \in A,\ \xi_i, \gamma_i \in \Xi_i : \hat{u}_i(\xi_i) - \hat{u}_i(\gamma_i) \geq h_i(\gamma_i)(\gamma_i - \xi_i)$.

PROOF. The following equivalence relations follow trivially from the definition of BNIC and equations (12) and (13): the mechanism $\langle H, P \rangle$ is BNIC iff $\forall i \in A, \xi_i, \gamma_i \in \Xi_i : \bar{u}_i(\xi_i, \xi_i) \geq \bar{u}_i(\xi_i, \gamma_i)$ iff $\forall i \in A, \xi_i, \gamma_i \in \Xi_i : \hat{u}_i(\xi_i) \geq p_i(\gamma_i) - h_i(\gamma_i)\xi_i$ iff $\forall i \in A, \xi_i, \gamma_i \in \Xi_i : \hat{u}_i(\xi_i) - \hat{u}_i(\gamma_i) \geq h_i(\gamma_i)(\gamma_i - \xi_i)$. □

LEMMA 2. *A mechanism $\langle H, P \rangle$ is* BNIC *iff* $\forall \xi_i \in \Xi_i$:

1) $h_i(\xi_i)$ is monotone nonincreasing; and

2) the following equation holds:

$$p_i(\xi_i) = p_i(0) + \xi_i h_i(\xi_i) - \int_0^{\xi_i} h_i(t_i) dt_i \tag{17}$$

.

PROOF. "⇒": By lemma 1, for all i, ξ_i and γ_i, we have both $\hat{u}_i(\xi_i) - \hat{u}_i(\gamma_i) \geq h_i(\gamma_i)(\gamma_i - \xi_i)$ and $\hat{u}_i(\gamma_i) - \hat{u}_i(\xi_i) \geq h_i(\xi_i)(\xi_i - \gamma_i)$. Then we can obtain

$$h_i(\xi_i)(\xi_i - \gamma_i) \leq \hat{u}_i(\gamma_i) - \hat{u}_i(\xi_i) \leq h_i(\gamma_i)(\xi_i - \gamma_i) \tag{18}$$

It follows that $h_i(\xi_i) \leq h_i(\gamma_i)$ iff $\xi_i \geq \gamma_i$. Therefore, $h_i(\xi_i)$ is a monotone nonincreasing function.

To show the correctness of equation (17), we firstly divide the interval $[0, \xi_i]$ into L intervals of length $\delta = \frac{\xi_i}{L}$. Denote by $y^k = (k+1)\delta$ the rightmost end of the kth interval, and by $x^k = k\delta$ its leftmost end. Let $\xi_i = y^k$ and $\gamma_i = x^k$, then

$$\begin{aligned} \sum_{k=0}^{L-1} h_i(y^k)(y^k - x^k) &\leq \sum_{k=0}^{L-1} \hat{u}_i(x^k) - \hat{u}_i(y^k) \\ &\leq \sum_{k=0}^{L-1} h_i(x^k)(y^k - x^k) \end{aligned} \tag{19}$$

Noticing that, $y^k = x^{k+1}$ for all $0 \leq k \leq L-1$, therefore

$$\sum_{k=0}^{L-1} \hat{u}_i(x^k) - \hat{u}_i(y^k) = \hat{u}_i(0) - \hat{u}_i(\xi_i) \tag{20}$$

Both the left part and right part of inequality (19) are Rieman sums. By increasing L, δ gradually approaches 0, both of the left part and right part of inequality (19) converge to $\int_0^{\xi_i} h_i(t_i) dt_i$. Therefore,

$$\hat{u}_i(0) - \hat{u}_i(\xi_i) = \int_0^{\xi_i} h_i(t_i) dt_i \tag{21}$$

Moreover, equation (13) follows

$$\hat{u}_i(0) = p_i(0) \tag{22}$$

Finally, the equation in item 2) of this lemma follows by combining the equations (13)(21)and (22).

"⇐": By equation (13), equation (17) is equivalent to

$$\hat{u}_i(\xi_i) = p_i(0) - \int_0^{\xi_i} h_i(t_i) dt_i \tag{23}$$

So, for all $\gamma_i \in \Xi_i$

$$\hat{u}_i(\xi_i) - \hat{u}_i(\gamma_i) = \int_{\xi_i}^{\gamma_i} h_i(t_i) dt_i \tag{24}$$

Since h_i is monotone nonincreasing,

$$\int_{\xi_i}^{\gamma_i} h_i(t_i) dt_i \geq (\gamma_i - \xi_i) h_i(\gamma_i) \tag{25}$$

then $\hat{u}_i(\xi_i) - \hat{u}_i(\gamma_i) \geq (\gamma_i - \xi_i) h_i(\gamma_i)$. Therefore, by lemma 1, $\langle H, P \rangle$ is BNIC. □

Then, we can transform the expression for expected profit, *i.e.*, equation (16), to a simpler form.

LEMMA 3. *A mechanism $\langle H, P \rangle$ is* BNIC *only if*

$$\begin{aligned} \mathbb{E}_{\xi\ \in \Xi}[\sigma(\xi)] = \int_{\Xi} \Big(& v_{\mathcal{F}}(K \dagger \hat{H}(\xi)) \\ & - \sum_{i \in A} (\xi_i + \frac{F_i(\xi_i)}{f_i(\xi_i)}) H_i(\xi) \Big) f(\xi) d\xi \\ & - \sum_{i \in A} \Big(p_i(0) - \int_0^{\omega_i} h_i(t_i) dt_i \Big) \end{aligned} \tag{26}$$

PROOF. By Lemma 2,

$$p_i(\xi_i) = p_i(0) + \xi_i h_i(\xi_i) - \int_0^{\xi_i} h_i(t_i) dt_i \tag{27}$$

Therefore, by equation (11) we can further obtain

$$\int_{\Xi} P_i(\xi)f(\xi)d\xi = \int_0^{\omega_i} p_i(\xi_i)f(\xi_i)d\xi_i = p_i(0) + \int_0^{\omega_i} \xi_i h_i(\xi_i)f_i(\xi_i)d\xi_i - \int_0^{\omega_i}\int_0^{\xi_i} h_i(t_i)dt_i f(\xi_i)d\xi_i \quad (28)$$

Since

$$\int_0^{\omega_i}\int_0^{\xi_i} h_i(t_i)dt_i f(\xi_i)d\xi_i = \int_0^{\omega_i} h_i(t_i)(1-F_i(t_i))dt_i \quad (29)$$

can be obtained by changing the order of integration. So,

$$\int_{\Xi} P_i(\xi)f(\xi)d\xi = p_i(0) - \int_0^{\omega_i} h_i(\xi_i)d\xi_i + \int_{\Xi}(\xi_i + \frac{F_i(\xi_i)}{f_i(\xi_i)})H_i(\xi)f(\xi)d\xi \quad (30)$$

Finally, equation (26) can be obtained by combining equations (30) and (16). □

Lemma 4. *A mechanism $\langle H, P\rangle$ is* BNIC *and* IR *only if*

$$\sum_{i\in A}\Big(p_i(0) - \int_0^{\omega_i} h_i(t_i)dt_i\Big) \geq 0 \quad (31)$$

Proof. $\langle H, P\rangle$ is BNIC follows equation (23), since we have shown that it is equivalent to equation (17). Then by inequality (14), we have $\forall i \in A : \hat{u}_i(\omega_i) = p_i(0) - \int_0^{\omega_i} h_i(t_i)dt_i \geq 0$. So, inequality (31) trivially holds. □

Note that, in lemma 3, $\xi_i + \frac{F_i(\xi_i)}{f_i(\xi_i)}$ is a famous structure in Myerson's optimal auctions [18]. We denote it as $\lambda(\xi_i)$ and name it as the virtual cost of transition i. We focus on the *regularity* case of the space Ξ, that is, we assume λ is non-decreasing function of ξ_i for every i. Moreover, we let $g_(\eta,\xi) = v_{\mathcal{F}}(K\dagger\eta) - \sum_{i\in A}(\xi_i + \frac{F_i(\xi_i)}{f_i(\xi_i)})\eta_i$ where $\xi \in \Xi$ is the current bid profile, η is a selected social law, and if $i \in \eta$ then $\eta_i = 1$, else $\eta_i = 0$. The above 2 lemmas show that the value of the expected profit can be computed as the difference of two items: the first item is an integral $\int_{\Xi} g(\hat{H}(\xi),\xi)f(\xi)d\xi$ whose value is determined by $\hat{H}(\xi)$ only; and the second item is non-negative. Based on this observation, we can propose the following mechanism, which takes a bid profile ξ as input and output the selected social law $H^*(\xi)$ and the payment vector $P^*(\xi)$ to the agents.

Mechanism $\mathcal{PM}$-$\mathcal{SLA}$

Input: a Kripke structure $K = \langle S, s_0, R, A, \alpha, c, \pi\rangle$, a feature set $\mathcal{F}$, the cost space Ξ_i and probability density function f_i for each agent i, and a bid profile ξ

1. Find out the social law $\eta = \hat{H}^*(\xi) \in N(R)$ that maximizes $g_(\eta,\xi) = v_{\mathcal{F}}(K\dagger\eta) - \sum_{i\in A}\lambda_i(\xi_i)\cdot\eta_i$;
2. Compute the payment to each agent i as $P_i^*(\xi) = H_i^*(\xi)\xi_i + \int_{\xi_i}^{\omega_i} H_i^*(t_i,\xi_{-i})dt_i$;

Output: the allocation and payment $(H^*(\xi), P^*(\xi))$

The intuition behind this mechanism is the allocation function $H^*(\xi)$ maximizes g for every ξ and fortunately is monotone non-increasing (we will show this later), so it maximizes the integration $\int_{\Xi} g(\hat{H}(\xi),\xi)f(\xi)d\xi$; the payment function satisfies the constraints imposed by BNIC and IR, *i.e.*, equations (17) and (31), and makes $\sum_{i\in A}\Big(p_i(0) - \int_0^{\omega_i} h_i(t_i)dt_i\Big)$ get its minimal value 0, and so the expected profit is maximized within all BNIC and IR mechanisms. In the following, we will formally prove these results.

Lemma 5. *The allocation function of $\mathcal{PM}$-$\mathcal{SLA}$ is monotone non-increasing.*

Proof. Let $\hat{H}^*(\xi)$ be the social law that maximizes the function $g(\hat{H}(\xi),\xi) = v_{\mathcal{F}}(K\dagger\hat{H}(\xi)) - \sum_{i\in A}\lambda(\xi_i)H_i(\xi)$, $\hat{H}'(\xi)$ be a social law that $H_k'(\xi) \geq H_k^*(\xi)$, and $\gamma = (\gamma_k, \xi_{-k})$ where $\gamma_k \geq \xi_k$. Suppose $H_k^*(\xi_k,\xi_{-k})$ is a increasing function of ξ_k, that is, when the bid vector changes to γ, the social law $\hat{H}'(\xi)$ becomes the one maximize the objective function. Therefore, $g(\hat{H}'(\xi),\gamma) \geq g(\hat{H}^*(\xi),\gamma)$. So,

$$v_{\mathcal{F}}(K\dagger\hat{H}'(\xi)) - \sum_{i\,\in A}\lambda(\gamma_i)H_i'(\xi) \geq v_{\mathcal{F}}(K\dagger\hat{H}^*(\xi)) - \sum_{i\in A}\lambda(\gamma_i)H_i^*(\xi) \quad (32)$$

Since $\gamma_i = \xi_i$ for all $i \neq k$, inequality (32) is equivalent to

$$g(\hat{H}'(\xi),\xi) + (\lambda(\xi_k) - \lambda(\gamma_k))H_k'(\xi) \geq g(\hat{H}^*(\xi),\xi) + (\lambda(\xi_k) - \lambda(\gamma_k))H_k^*(\xi) \quad (33)$$

By the regularity of the space Ξ, we can further obtain $g(\hat{H}'(\xi),\xi) \geq g(\hat{H}^*(\xi),\xi)$, which contradicts the fact that $\hat{H}^*(\mathbf{x})$ is the social law with highest profit in this case. So, for each k, $H_k^*(\xi_k,\xi_{-k})$ and further $h_k^*(\xi_k)$ is a non-increasing function of ξ_k. □

The intuition behind the above lemma is actually very obvious. Since the higher an agent i bids, the higher virtual cost it has, and the profits of all the social laws that containing transition i will decrease because of increasing total payments, so the likelihood of selecting transition i will decrease. Now we are ready to prove the main results of this section.

Theorem 6. *$\mathcal{PM}$-$\mathcal{SLA}$ is* BNIC *and* IR.

Proof. Since $p_i^*(\xi_i) = \int_{\Xi_{-i}} P_i^*(\xi_i,\xi_{-i})f_{-i}(\xi_{-i})d\xi_{-i} = \xi_i h_i^*(\xi_i) + \int_0^{\omega_i} h_i^*(t_i)dt_i - \int_0^{\xi_i} h_i^*(t_i)dt_i$, we can obtain

$$p_i^*(0) = \int_0^{\omega_i} h_i^*(t_i)dt_i \quad (34)$$

and further $p_i^*(\xi_i) = p_i^*(0) + \xi_i h_i^*(\xi_i) - \int_0^{\xi_i} h_i^*(t_i)dt_i$. Moreover, by lemma 5, h_i^* is non-increasing. So, by lemma 2, PM-NFA is BNIC.

Then, by equations (23) and (34), $\forall\xi_i \in \Xi_i$:

$$\hat{u}_i^*(\xi_i) = \int_0^{\omega_i} h_i^*(t_i)dt_i - \int_0^{\xi_i} h_i^*(t_i)dt_i \geq 0 \quad (35)$$

So, $\mathcal{PM}$-$\mathcal{SLA}$ is also IR. □

Theorem 7. *PM-SLA maximizes the expected profit within all* BNIC *and* IR *mechanisms.*

PROOF. Suppose there is a BNIC and IR mechanism $\mathcal{M}' = (H', P')$ which achieve higher expected profit than the PM-NLA mechanism $\mathcal{M}^* = (H^*, P^*)$. By lemma 2, $\mathcal{M}'$ is IR follows

$$\forall i \in A : p'_i(0) - \int_0^{\omega_i} h'_i(t_i)dt_i \geq 0 \tag{36}$$

Therefore, by lemma 3, equality (34) and inequality (36), $\mathbb{E}_{\xi\in\Xi}[\sigma_{\mathcal{M}'}(\xi)] > \mathbb{E}_{\xi\in\Xi}[\sigma_{\mathcal{M}^*}(\xi)]$ implies

$$\begin{aligned}&\int_{\Xi}\Big(v_{\mathcal{F}}(K\dagger\hat{H}'(\xi)) - \sum_{i\in A}\lambda(\xi_i)H'_i(\xi)\Big)f(\xi)d\xi\\ &> \int_{\Xi}\Big(v_{\mathcal{F}}(K\dagger\hat{H}^*(\xi)) - \sum_{i\in A}\lambda(\xi_i)H^*_i(\xi)\Big)f(\xi)d\xi\end{aligned} \tag{37}$$

which contradicts the fact that $H^*(\xi)$ maximizes $v_{\mathcal{F}}(K\dagger\hat{H}(\xi)) - \sum_{i\in A}\lambda(\xi_i)H_i(\xi)$ with respect to all $\xi \in \Xi$. □

4. COMPUTATIONAL COMPLEXITY

The computational problem of PM-NFA can be specified as given a Kripke structure $K = \langle S, s_0, R, A, \alpha, c, \pi\rangle$, a feature set $\mathcal{F} = \{(\varphi_1, x_1), ..., (\varphi_k, x_k)\}$, the cost value space Ξ_i and the corresponding probability density function f_i for each agent $i \in A$, and the bid profile ξ from the agents, to output the selected social law $\hat{H}^*(\xi)$ and the corresponding payment profile $P^*(\xi)$. Notice that, the two functions H^* and P^* are actually implicitly defined by the mechanism $\mathcal{PM}$-$\mathcal{SLA}$, we are not going to explicitly show the functions but instead compute the values of the two functions for the current bid profile.

The computation of $H^*(\xi)$ is a common optimization problem. The computation of $P^*(\xi)$ seems more complex, because it involves computing the integral $\int_{\xi_i}^{\omega_i} H^*_i(t_i, \xi_{-i})dt_i$ which seems requiring the explicit expression of $H^*_i(t_i, \xi_{-i})$ which is a function of $t_i \in \Xi$. But the following result shows there is a more direct approach to compute $P^*(\xi)$.

LEMMA 8. *For all $i \in A$,*

$$P^*_i(\xi) = \begin{cases} inf\{c_i \in \Xi_i : i \notin \hat{H}^*(c_i, \xi_{-i})\} & i \in \hat{H}^*(\xi)\\ 0 & else\end{cases}$$

PROOF. By lemma 5, for all $i \in A$, $\hat{H}^*(c_i, \xi_{-i})$ is a monotone non-increasing function of c_i, therefore $H^*_i(c_i, \xi_{-i}) = 0$ for all $c_i \in \Xi_i$ or there is a $\theta_i \in \Xi$ such that

$$H^*_i(c_i, \xi_{-i}) = \begin{cases} 1 & 0 \leq c_i \leq \theta_i\\ 0 & \text{else}\end{cases}$$

For the first case, it is trivial that $i \notin \hat{H}^*(\xi)$ and $P^*_i(\xi) = 0$; for the second case, if $\xi_i \leq \theta_i$, then $i \in \hat{H}^*_i(\xi)$ and $P^*(\xi) = \xi_i + \int_{\xi_i}^{\theta_i} 1dt_i = \theta_i = inf\{c_i \in \Xi_i : i \notin \hat{H}^*(c_i, \xi_{-i})\}$, else $i \notin \hat{H}^*(\xi)$ and $P^*_i(\xi) = 0$. □

The above result means the payment to each selected agent is the threshold cost she can bid to keep her in the selected set, given the bids from other agents unchanged. Therefore, firstly we can obtain the following result:

THEOREM 9. *$\mathcal{PM}$-$\mathcal{SLA}$ is dominant strategy incentive compatible (DSIC, i.e., truthful).*

PROOF. By lemma 5, the allocation function of $\mathcal{PM}$-$\mathcal{SLA}$ is monotone, and by lemma 8, the payment to the selected agents are their threshold bid, therefore according to the famous characterization of single-parameter truthful mechanisms [18, 24], $\mathcal{PM}$-$\mathcal{SLA}$ is truthful. □

Moreover, step 2 of mechanism $\mathcal{PM}$-$\mathcal{SLA}$ and be equivalently substituted by the following steps.

> 2.1 For each agent $i \in \eta^*$ find out the social law $\eta = \hat{H}^{-i}(\xi) \in N(R \setminus \{i\})$ that maximizes $g(\eta, \xi)$, and let $P^*_i(\xi)$ be that value that satisfy $\lambda(P^*_i(\xi)) = g(\hat{H}^*(\xi), \xi) + \lambda(\xi_i) - g(\hat{H}^{-i}(\xi), \xi)$;
>
> 2.2 For each agent $i \notin \hat{H}^*(\xi)$ let $P^*_i(\xi) = 0$;

Now, it is obvious that, the solution to the following subproblem is crucial in the computation of $\mathcal{PM}$-$\mathcal{SLA}$.

> PM-SLA-ALLOCATION
>
> **Given**: a Kripke structure $K = \langle S, s_0, R, A, \alpha, c, \pi\rangle$, a feature set $\mathcal{F}$, the cost space Ξ_i and probability density function f_i for each agent i, a bid profile ξ, and an available transition set $R^a \subseteq R$
>
> **Output**: a social law $\eta \in N(R^a)$ that maximize $g_(\eta, \xi) = v_{\mathcal{F}}(K\dagger\eta) - \sum_{i\in A}\lambda_i(\xi_i)\cdot\eta_i$

Finally, we can prove a negative result on the computation of $\mathcal{PM}$-$\mathcal{SLA}$.

THEOREM 10. *The problem of computing $\mathcal{PM}$-$\mathcal{SLA}$ is FP^{NP}-complete.*

PROOF. Firstly, we can show the PM-SLA-ALLOCATION problem is FP^{NP}-complete: the associated decision problem, *i.e.*, "whether there is a social law $\eta \in N(R^a)$ achieving the value of $g(\eta, \xi)$ at least $l \in \mathbb{R}^+$", is trivially in NP, so this problem is in FP^{NP}; Moreover, we can reduce the OPTIMAL SOCIAL LAW problem [2], which is FP^{NP}-complete, to this problem. Then, we can find out that the complexity of computing $\mathcal{PM}$-$\mathcal{SLA}$ mainly involves solving at most $n + 1$ instances of PM-SLA-ALLOCATION, where n is the number of transitions in R^a, so it is also FP^{NP}-complete. □

5. APPROXIMATION

Since the computation of $\mathcal{PM}$-$\mathcal{SLA}$ is intractable, we are going to find out an computationally tractable approximation mechanism. Basically, this approximation mechanism is based on an approximation algorithm for PM-SLA-ALLOCATION: firstly, we formulate an integer program for this problem based on the methodology introduced in [2]. The idea is the value of the objective function $g(\eta, \xi)$, which we aim to maximize, depends not only on whether each transition is selected but also on whether each formula φ_i in the feature set $\mathcal{F} = \{(\varphi_1, a_1), ..., (\varphi_k, a_k)\}$ is satisfied in the current state, which may further depends on whether each sub-formula of φ_i is satisfied in the current state s or the next states $\mathcal{N}(s) = \{s' \in S : \exists(s, s') \in R\}$. Each of the above choice can be captured by a boolean variable, and the combined value assignments to all the boolean variables are constrained by the semantics of CTL.

For each formula φ, its related sub-formulas can be denoted as $cl(\varphi)$, or be called the closure of φ defined as:

$$cl(\varphi) = \{\varphi\} \cup sub(\varphi)$$

where

$$sub(\varphi) = \begin{cases} cl(\psi) \cup cl(\chi) & \text{if } \varphi = \psi \vee \chi \text{ or } \varphi = \mathsf{E}(\psi\mathcal{U}\chi) \\ cl(\psi) & \text{if } \varphi = \neg\psi \text{ or } \varphi = \mathsf{E}\bigcirc\psi \text{ or } \varphi = \mathsf{E}\Box\psi \\ \{\varphi\} & \text{if } \varphi \in \Phi \end{cases}$$

Then for each feature set $\mathcal{F} = \{(\varphi_1, a_1), ..., (\varphi_k, a_k)\}$, we let

$$cl(\mathcal{F}) = cl(\varphi_1) \cup \cdots \cup cl(\varphi_k)$$

Basically, the set $cl(\mathcal{F})$ is the set of all the formulas that may influence the value of the objective function $g(\eta, \xi)$. So we can introduce the following two classes of boolean variables:

- $x^s_\varphi \in \{0, 1\}$ for each $s \in S$ and $\varphi \in cl(\mathcal{F})$;
- $y^s_{s'}$ *(or* y_i*)* $\in \{0, 1\}$ for each $i = (s, s') \in R$.

Intuitively, $x^s_\varphi = 1$ means φ is true in state s, and $y^s_{s'} = 1$ means the transition is selected by the mechanism.

The semantics constraints can be expressed by the following result, which directly follows from CTL semantics:

LEMMA 11. *The following 2 formulas are valid:*

1) $\mathsf{E}(\psi\mathcal{U}\chi) \leftrightarrow (\chi \vee (\psi \wedge \mathsf{E}\bigcirc \mathsf{E}(\psi\mathcal{U}\chi)))$

2) $\mathsf{E}\Box\varphi \leftrightarrow \varphi \wedge \mathsf{E}\bigcirc \mathsf{E}\Box\varphi$

Therefore, we can further obtain the following result.

LEMMA 12.

1) $x^s_p = 1$ *iff* $p \in \pi(x)$*;*

2) $x^s_{\neg\psi} = 1$ *iff* $x^s_\psi = 0$*;*

3) $x^s_{\psi\vee\chi} = 1$ *iff* $x^s_\psi = 1$ *or* $x^s_\chi = 1$*;*

4) $x^s_{\mathsf{E}\bigcirc\psi} = 1$ *iff* $\exists s' \in \mathcal{N}(s) : y^s_{s'} = 0$ *and* $x^{s'}_\psi = 1$*;*

5) $x^s_{\mathsf{E}(\psi\mathcal{U}\chi)} = 1$ *iff at least one of the following hold:*

- $x^s_\chi = 1$*;*
- $x^s_\psi = 1$ *and* $\exists s' \in \mathcal{N}(s) : y^s_{s'} = 0$ *and* $x^{s'}_{\mathsf{E}(\psi\mathcal{U}\chi)} = 1$*;*

6) $x^s_{\mathsf{E}\Box\varphi} = 1$ *iff both of the following hold:*

- $x^s_\varphi = 1$*;*
- $\exists s' \in \mathcal{N}(s) : y^s_{s'} = 0$ *and* $x^{s'}_{\mathsf{E}\Box\varphi} = 1$.

PROOF. Items 1)- 4) follow trivially from the semantics of CTL; and item 5) and item 6) follow from lemma 11. □

Then, for each instance of PM-SLA-ALLOCATION parameterized by $(K, \mathcal{F}, (\Xi_i, f_i)_{i\in A}, R^a, \xi)$, we can give the following integer program (which we refer as $\mathsf{ILP\text{-}PSA}(R^a)$, since we assume the other parameters are clear from the context):

maximize

$$\sum_{(\varphi_i, a_i)\in\mathcal{F}} a_i \cdot x^{s_0}_{\varphi_i} - \sum_{(s,s')\in R} (\xi_i + \frac{F_i(\xi_i)}{f_i(\xi_i)}) y^s_{s'} \quad (38)$$

subject to:

$$x^s_\varphi \in \{0, 1\} \quad \forall s \in S, \varphi \in cl(\mathcal{F}) \quad (39)$$

$$y^s_{s'} \in \{0, 1\}, \quad \forall (s, s') \in R \quad (40)$$

$$y^s_{s'} = 0, \quad \forall (s, s') \in R \setminus R^a \quad (41)$$

$$\sum_{s'\in\mathcal{N}(s)} (1 - y^s_{s'}) \geq 1, \quad \forall s \in S \quad (42)$$

$$x^s_p = 1, \quad \forall s \in S, p \in \pi(s) \quad (43)$$

$$x^s_p = 0, \quad \forall s \in S, p \in (\Phi \cap cl(\mathcal{F})) \setminus \pi(s) \quad (44)$$

$$x^s_{\neg\psi} = 1 - x^s_\psi, \quad \forall s \in S, \neg\psi \in cl(\mathcal{F}) \quad (45)$$

$$x^s_{\psi\vee\chi} \geq x^s_\psi, \quad \forall s \in S, \psi \vee \chi \in cl(\mathcal{F}) \quad (46)$$

$$x^s_{\psi\vee\chi} \geq x^s_\chi, \quad \forall s \in S, \psi \vee \chi \in cl(\mathcal{F}) \quad (47)$$

$$x^s_{\psi\vee\chi} \leq x^s_\psi + x^s_\chi, \quad \forall s \in S, \psi \vee \chi \in cl(\mathcal{F}) \quad (48)$$

$$x^s_{\mathsf{E}\bigcirc\psi} \geq x^{s'}_\psi - y^s_{s'}, \quad \forall s \in S, \mathsf{E}\bigcirc\psi \in cl(\mathcal{F}), s' \in \mathcal{N}(s) \quad (49)$$

$$x^s_{\mathsf{E}\bigcirc\psi} \leq \sum_{s'\in\mathcal{N}(s)} x^{s'}_\psi, \quad \forall s \in S, \mathsf{E}\bigcirc\psi \in cl(\mathcal{F}), s' \in \mathcal{N}(s) \quad (50)$$

$$x^s_{\mathsf{E}(\psi\mathcal{U}\chi)} \geq x^s_\chi, \quad \forall s \in S, \mathsf{E}(\psi\mathcal{U}\chi) \in cl(\mathcal{F}) \quad (51)$$

$$\begin{aligned} x^s_{\mathsf{E}(\psi\mathcal{U}\chi)} &\geq x^s_\psi + x^{s'}_{\mathsf{E}(\psi\mathcal{U}\chi)} - y^s_{s'} - 1, \\ \forall s \in S, &\mathsf{E}(\psi\mathcal{U}\chi) \in cl(\mathcal{F}), s' \in \mathcal{N}(s) \end{aligned} \quad (52)$$

$$x^s_{\mathsf{E}(\psi\mathcal{U}\chi)} \leq x^s_\psi + x^s_\chi, \forall s \in S, \mathsf{E}(\psi\mathcal{U}\chi) \in cl(\mathcal{F}) \quad (53)$$

$$x^s_{\mathsf{E}\Box\psi} \leq x^s_\psi, \quad \forall s \in S, \mathsf{E}\Box\psi \in cl(\mathcal{F}) \quad (54)$$

$$x^s_{\mathsf{E}\Box\psi} \leq \sum_{s'\in\mathcal{N}(s)} x^{s'}_{\mathsf{E}\Box\psi}, \quad \forall s \in S, \mathsf{E}\Box\psi \in cl(\mathcal{F}) \quad (55)$$

$$\begin{aligned} x^s_{\mathsf{E}\Box\psi} &\geq x^s_\psi + x^{s'}_{\mathsf{E}\Box\psi} - y^s_{s'} - 1, \\ \forall s \in S, &\mathsf{E}\Box\psi \in cl(\mathcal{F}), s' \in \mathcal{N}(s) \end{aligned} \quad (56)$$

A solution to $\mathsf{ILP\text{-}PSA}(R^a)$ is an assignment to the boolean variables $(x^s_\varphi)_{s\in S,\varphi\in\mathcal{F}}$, $(y^s_{s'})_{(s,s')\in R}$, which means the social law $\eta = \{(s, s') \in R : y^s_{s'} = 1\}$ is selected. So, in the case we mention η as an solution to $\mathsf{ILP\text{-}PSA}(R^a)$.

THEOREM 13. *A social law* $\eta^* \in N(R^a)$ *is selected in a solution to* $\mathsf{ILP\text{-}PSA}(R^a)$ *iff* $\eta^* = \arg\max_{\eta\in N(R^a)} g(\eta, \xi)$*, i.e.,* η^* *is a solution to an instance of* PM-SLA-ALLOCATION *where the available transition set is* R^a*.*

PROOF. Since (38) is equivalent to $g(\eta, \xi)$, we only have to show (39)- (56) correctly captures all the constraints imposed on the variables: (41) means we only select transitions from R^a; The reasonable constraints of social law are captured by (42); And by lemma 12, (43)-(56) correctly captures the constraints imposed by CTL semantics. □

Since integer programming is intractable, the proposed integer program is still a computationally infeasible solution. But we can obtain a tractable approximation algorithm via relaxing and rounding. We can relax $\mathsf{ILP\text{-}PSA}(R^a)$ to a linear program $\mathsf{LP\text{-}PSA}(R^a)$, which is obtained from $\mathsf{ILP\text{-}PSA}(R^a)$ via replacing (39) and (40) respectively to

- $0 \leq x_\varphi^s \leq 1 \;\; \forall s \in S, \varphi \in cl(\mathcal{F})$ (39′)
- $0 \leq y_{s'}^s \leq 1, \;\; \forall (s,s') \in R$ (40′)

Then, we can design an algorithm for PM-SLA-ALLOCATION:

Algorithm FIND-OSL-APX

Input: an instance of PM-SLA-ALLOCATION $(K, \mathcal{F}, (\Xi_i, f_i)_{i \in A}, R^a, \xi)$, and a threshold $\theta \in [0,1]$

1. Find the solution $(x_\varphi^s)_{s \in S, \varphi \in \mathcal{F}}$, $(y_{s'}^s)_{(s,s') \in R}$ to $\mathsf{LP\text{-}PSA}(R^a)$;
2. Let $\eta^+ = \{(s,s') \in R^a : y_{s'}^s \geq \theta\}$.

Output: the selected social law η^+

Afterward, we can give the following mechanism:

Mechanism $\mathcal{PM\text{-}SLAX}$

Input: a Kripke structure $K = \langle S, s_0, R, A, \alpha, c, \pi \rangle$, a feature set $\mathcal{F}$, the cost space Ξ_i and probability density function f_i for each agent i, and a bid profile ξ

1. $\eta \leftarrow$FIND-OSL-APX(R);
2. For each agent $i \in \eta$: $\eta_{-i} \leftarrow$FIND-OSL-APX$(R \setminus \{i\})$, and let P_i be that value that satisfy $\lambda(P_i) = g(\eta, \xi) + \lambda(\xi_i) - g(\eta^{-i}, \xi)$;
3. For each agent $i \notin \eta$ let $P_i(\xi) = 0$;

Output: the allocation and payment (η, P)

Finally, we can show $\mathcal{PM\text{-}SLAX}$ is a DSIC, IR and computational tractable mechanism, and if choose the threshold $\theta = 0.5$, it produces an approximation ratio equals to 2.

THEOREM 14. *$\mathcal{PM\text{-}SLAX}$ is* DSIC *and* IR.

PROOF. The allocation of $\mathcal{PM\text{-}SLAX}$ is also monotone non-increasing: the idea is when an agent raises her bid, the allocation to it computed by solving $\mathsf{LP\text{-}PSA}(R^a)$ must be non-increasing. Detailed proof can be formulated via reduction to absurdity just like that in the proof of lemma 5. Moreover, the payment to each selected agents is her threshold bid. So, we can conclude $\mathcal{PM\text{-}SLAX}$ is DSIC. IR follows from the fact when an agent is selected, the payment to her is her threshold bid which is higher than her current bid, *i.e.*, her true cost; otherwise the payment to her is 0. □

THEOREM 15. *$\mathcal{PM\text{-}SLAX}$ is tractable 2-approximation mechanism to $\mathcal{PM\text{-}SLA}$ if we choose $\theta = 0.5$.*

PROOF. The computation of $\mathcal{PM\text{-}SLAX}$ is mainly formed by at most $n+1$ calls to the function FIND-OSL-APX, which is tractable since linear programming is tractable, and therefore is tractable. Let OPT be the optimal expected profit (achieved by $\mathcal{PM\text{-}SLA}$), $(x_\varphi^{s+})_{s \in S, \varphi \in \mathcal{F}}$,$(y_i^+)_{i \in A}$ be the result obtained by solving $\mathsf{LP\text{-}PSA}(R^a)$, $(y_i)_{i \in A}$ be the allocation of $\mathcal{PM\text{-}SLAX}$, and $(x_\varphi^s)_{s \in S, \varphi \in \mathcal{F}}$ be the result enforced by $(y_i)_{i \in A}$. We have $\sum_{i \in A} \lambda_i(\xi_i) y_i^+ \geq \frac{1}{\theta} \cdot \sum_{i \in A} \lambda_i(\xi_i) y_i$, since we have $y_i^+ \geq \theta \cdot y_i$ for all i; Moreover, it is easy to show $x_\varphi^s \geq (1-\theta) \cdot x_\varphi^{s+}$ for all $s \in S$ and $\varphi \in cl(\mathcal{F})$, and therefore $\sum_{(\varphi_i, a_i) \in \mathcal{F}} a_i \cdot x_{\varphi_i}^{s_0} \geq (1-\theta) \cdot \sum_{(\varphi_i, a_i) \in \mathcal{F}} a_i \cdot x_{\varphi_i}^{s_0+}$. So, if we choose $\theta = 0.5$, then $\sum_{(\varphi_i, a_i) \in \mathcal{F}} a_i \cdot x_{\varphi_i}^{s_0} - \cdot \sum_{i \in A} \lambda_i(\xi_i) y_i \geq (1-\theta) \cdot \sum_{(\varphi_i, a_i) \in \mathcal{F}} a_i \cdot x_{\varphi_i}^{s_0+} - \theta \sum_{i \in A} \lambda_i(\xi_i) y_i^+ \geq \frac{1}{2} OPT$. □

6. RELATED WORK

The problem considered in this paper is mainly motivated from [2], where the OPTIMAL SOCIAL LAW problem has been formalized and a ILP solution has been proposed. Compared with this work, we non-trivially extend the OPTIMAL SOCIAL LAW problem to the strategic case, find a solution based on algorithmic mechanism design, and explicitly study an approach to solve the discovered computationally intractable problem. Rational behavior of the agents actually has been already considered in some work on social laws, e.g., [3] and so on, but they mainly focus on game theoretical analysis rather than synthesizing which intuitively reduces to mechanism design, as we do in this paper.

On the side of algorithmic mechanism design, our work especially relates to the class of work relates to graphs, e.g., path auctions, spanning-tree auctions, and so on, since social law synthesis is a problem on Kripke structures, which are also intrinsically graphs. Among this class of work, our work is most close to Elkind et al.'s work [11] on path auctions, since it proposes a methodology that solves the payment minimization problem in path auctions in the Bayesian case based on the framework of Myerson's optimal auctions [18]. But our work is different from [11] on mainly two aspects: firstly, we study a different optimization objective, *i.e.*, $(value - payment)$ vs. $payment$; secondly, we further study the computational issue and propose a tractable mechanism with constant-factor approximation guarantee. Actually, profit-maximization has been studied in *competitive auctions* [15, 14, 10, 9], but their settings are totally different, *i.e.*, they focus on buying (or selling) multiple identical items.

7. CONCLUSION AND FUTURE WORK

In this paper we non-trivially extend the OPTIMAL SOCIAL LAW problem to the strategical case, and propose a mechanism, which is truthful, individually rational, computationally tractable, and with constant-approximation guarantee, based on the framework of Bayesian mechanism design. By this work, we not only introduce a meaningful new framework for designing social laws, but also introduce an interesting new problem to algorithm mechanism design research. In our opinion, two aspects of future work are important: firstly, as we currently focus on the Bayesian case, it is also necessary to study the prior-free case; secondly, notice that we assumed each agent can own only one transition. It is interesting to remove this assumption and study the general case where each agent can own multiple transitions.

Acknowledgment

The authors would like to thank the anonymous 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 and Industrialization at Nanjing University.

REFERENCES

[1] T. Ågotnes, W. van der Hoek, and M. Wooldridge. Conservative social laws. In *Proc. of ECAI-12*, pages 49–53, 2012.

[2] T. Ågotnes and M. Wooldridge. Optimal social laws. In *Proceedings of AAMAS-10*, pages 667–674, 2010.

[3] T. Ågotnes, M. Wooldridge, and W. van der Hoek. Normative system games. In *Proceedings of AAMAS-07*, pages 876–883, 2007.

[4] A. Archer and E. Tardos. Frugal path mechanisms. In *Proc. of SODA-02*, pages 991–999, 2002.

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

[6] S. Bikhchandani, S. Vries, J. Schummer, and R. V. Vohra. Linear programming and vickrey auctions. In *Vohra (Eds.), Mathematics of the Internet: E-Auction and Markets*, 2002.

[7] N. Bulling and M. Dastani. Verifying normative behaviour via normative mechanism design. In *Proc. IJCAI-2011*, pages 103–108, 2011.

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

[9] M. C. Cary, A. D. Flaxman, J. D. Hartline, and A. R. Karlin. Auctions for structured procurement. In *Proc. of SODA-08*, pages 304–313, 2008.

[10] N. Chen, N. Gravin, and P. Lu. Optimal competitive auctions. In *Proceedings of STOC-14*, pages 253–262, 2014.

[11] E. Elkind, A. Sahai, and K. Steiglitz. Frugality in path auctions. In *Proc. of SODA-04*, pages 701–709, 2004.

[12] E. A. Emerson. *Handbook of Theoretical Computer science*, chapter Temporal and modal logic, pages 997–1072. 1990.

[13] R. Garg, V. Kumar, A. Rudra, and A. Verma. Coalitional games on graphs: core structure, substitutes and frugality. *Technical Report TR-02-60, UTCS*, 2002.

[14] A. V. Goldberg, J. D. Hartline, A. R. Karlin, M. Saks, and A. Wright. Competitive auctions. *Games and Economic Behavior*, 55(2):242–269, 2006.

[15] A. V. Goldberg, J. D. Hartline, and A. Wright. Competitive auctions and digital goods. In *Proceedings of SODA-01*, pages 735–744, 2001.

[16] A. R. Karlin and D. Kempe. Beyond VCG: Frugality of truthful mechanisms. In *Proc. of FOCS-05*, pages 615–626, 2005.

[17] J. Morales, M. Lopez-Sanchez, M. Wooldridge, and W. Vasconcelos. Automated synthesis of normative systems. In *Proceedings of AAMAS-13*, pages 484–490, 2013.

[18] R. B. Myerson. Optimal auction design. *Mathematics of Operations Research*, 6(1):58–73, 1981.

[19] N. Nisan and A. Ronen. Algorithmic mechanism design. In *Proceedings of the thirty-first annual ACM symposium on Theory of computing*, pages 129–140. ACM, 1999.

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

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

[22] Y. Shoham and M. Tennenholtz. On the synthesis of useful social laws for artificial agent societies. In *Proceedings of AAAI-92*, pages 276–281, 1992.

[23] 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.

[24] Y. Singer. Budget feasible mechanisms. In *Proc. of FOCS-10*, pages 765–774, 2010.

[25] K. Talwar. The price of truth: Frugality in truthful mechanisms. In *Proc. of STACS-03*, pages 608–619, 2003.

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

[27] J. Wu, C. Wang, and J. Xie. A framework for coalitional normative systems. In *Proceedings of AAMAS-11*, 2011.

[28] L. Zhang, H. Chen, J. Wu, C. Wang, and J. Xie. False-name-proof mechanisms for path auctions in social networks. In *Proceedings of ECAI-16*, pages 1485–1492, 2016.

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