A Norms-Based Probabilistic Decision-Making Model for Autonomic Traffic Networks

Maksims Fiosins, Jörg P. Müller and Michaela Huhn

Abstract We propose a norms-based agent-oriented model of decision-making of semi-autonomous vehicles in urban traffic scenarios. Computational norms are used to represent the driving rules and conventions that influence the distributed decision-making process of the vehicles. As norms restrict admissible behaviour of the agents, we propose to represent them as constraints, and we express the agents' individual and group decision-making in terms of distributed constraint optimization problems. The uncertain nature of the driving environment is reflected in our model through probabilistic constraints – collective norm compliance is considered as a stochastic distributed constraint optimization problem. In this paper, we introduce the basic conceptual and algorithmic ingredients of our model, including the norms provisioning and enforcement mechanisms (where electronic institutions are used), the norms semantics, as well as methods of the agents' cooperative decision-making. For motivation and illustration of our approach, we study a cooperative multi-lane highway driving scenario; we propose a formal model, and illustrate our approach by a small example.

Key words: cooperative traffic management; multi-agent decision-making; computational norms and institutions; probabilistic distributed constraint optimization; resampling

1 Introduction

The growing complexity of traffic management systems (TMS) in conjunction with new technological trends such as the increasing availability and growing amount of

Clausthal University of Technology,
Department of Informatics,
Julius-Albert Str. 4, D-38678, Clausthal-Zellerfeld, Germany
e-mail: {Maksims.Fiosins, joerg.mueller, michaela.huhn}@tu-clausthal.de

real-time traffic data, intelligence and autonomy of vehicular assistance functions (and indeed: of vehicles), and the capability of Car-to-X (C2X) communication, create new challenges for future cooperative traffic systems. As an example, the integration of car navigation with intelligent assistance functions and car-to-car communication enables software-based driving assistants not only assist the driver but make decisions, take actions, and communicate with other vehicles and traffic control devices autonomously, i.e., without explicit human command. Current examples of this development are BMWs cross-traffic assistant and traffic light assistant. In the future, we shall see much more advanced such services with a higher degree of autonomy. Reliability and efficiency of such systems will have crucial impact (positive or negative) on our society [1].

Architectural approaches towards modeling and controlling such future cooperative traffic management systems (CTMS) must (i) support system scalability and reconfigurability, (ii) provide adaptiveness to dynamic and stochastic environments, and (iii) enable a decentralized modeling and coordination approach which allows keeping local structures and decision models simple, and does not require complete models of the environment. We claim that multi-agent systems (MAS) [2] are a promising architectural approach to , as it provides appropriate paradigms and methods to model concepts such as autonomy, interaction, and adaptation.

One of the key questions in MAS research and design is how to control the behavior of agents while preserving their autonomy [3]. There are two main approaches to answer this question: The first is to define dedicated services, which coordinate the agents' behavior in terms of action synchronization or resource access planning. An example for this approach is the work on automated intersection control by Dresner and Stone [4]. The second, more decentralized approach employs indirect organizational and social control concepts such as computational norms [5], including permission, obligation or prohibition of states or actions, and provides monitoring, incentives, and penalizing mechanisms to organize and control agent behavior. While both approaches have pros and cons, our research mostly concentrates on the latter aspect, because we argue that direct, centralized control is often unfeasible in large MAS.

In our research we start from concepts and methods from normative multiagent systems [6], which we extend by a distributed constraints semantics and by the ability to deal with uncertainty. Computational norms can be conceived as rules, which define appropriate (or unacceptable) states or actions in a given environment (such as maximum speed, minimum distance, or priority at intersections). Compliance to norms is about to effect safe, stable, and efficient functioning of the overall system. Norms lifecycle support mechanisms such as electronic institutions [7] support (i) norms creation, maintenance, and evolution; (ii) provisioning of norms to the agents, and (iii) norms monitoring and enforcement.

Constraints have been proposed as an operational semantics of norms [8], allowing to detect norms violations and optimizing decisions under a given set of norms, based on an established computational framework. We reflect the distributed nature of decision-making situations in traffic scenarios by using distributed constraint satisfaction / optimization (DCSP/DCOP) [9]. In a DCSP/DCOP, each agent controls a subset of the variables and has only local knowledge about constraints of the DCSP or DCOP [10].

In addition, environment uncertainty requires probabilistic decision-making models [11]: Norms are subject to changes, the reasons of which are unknown to the agents. We propose probabilistic distributed constraint satisfaction,to allow agents to analyze the uncertainty of the environment and make appropriate decisions.

This introductory paper sets out the conceptual and algorithmic pillars of our approach. Starting from a simple application scenario (cooperative multi-lane highway driving), we propose a novel agent-based coordinated decision-making model for autonomic CTMS. We introduce computational norms, describe their semantics in terms of constraints, and formalize deciding norms compliance as a stochastic DCOP. We outline a generic multi-agent architecture for norms provisioning.

The paper is organized as follows. Section 2 presents the motivating scenario. Section 3 describes the overall MAS architecture and models, which support norm provision and interpretation by the agents. In Section 4, we formalize the norms semantics in terms of stochastic constraints. Section 5 presents and discusses an illustrative example. Section 6 contains conclusions and points to future directions.

2 Norms in a Traffic Scenario

In this section we present a simple traffic scenario that is regulated by norms. We consider autonomous vehicles (AV) driving one-way on a multi-lane highway (see Fig. 1). Each AV has its individual goals (destination and preferred arrival time) and each AV respects the physical laws for safe driving; it also knows the traffic rules that apply in a situation. Here we distinguish between two types of rules: (i) basic safety-related rules, which the agent will not willingly break, because doing so would violate the physical integrity of itself or others (e.g., entering a motorway in the wrong direction)[1], and (ii) efficiency-related rules (like speed limits) that mainly serve to optimizing traffic flows with respect to superordinate targets like maximizing throughput, or minimizing the overall time in traffic congestions or environmental pollution (noise or emissions).

The notion of norms serves us as a conceptual means to express these rules: Norms are behavioral guidelines provided by so-called electronic institutions [7] in order to enforce safety rules and encourage efficiency rules. An institution will offer positive incentives for norm-compliant behavior, and sanctions in case of detected norm violation. For safety norms, we can imagine that an infinitely high penalty will be issued to the violating agent. For efficiency norms, an agent may decide freely whether they fits to its goals, as a norm violation effects neither its physical safety nor results in a capital offense.

Norms refer to the externally observable dynamic state of an AV that consists of a list of parameters like its current location on a certain lane, its speed, and its

[1] If a safety level is needed to guarantee agents will not coincidentally break rules, even stronger measures (like physical precautions) need to be taken. This problem is not addressed in this paper.

distance to neighboring vehicles etc. Norms may have a restricted *scope*, e.g. they may only apply for some kind of vehicles in a certain lane section. We assume that norms are published to all traffic participants by the institution.

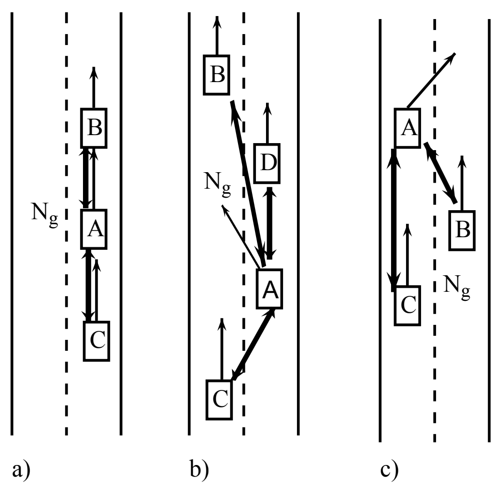

Fig. 1 Examples of group norms for cooperative multi-lane driving in the viewpoint of vehicle A. a) illustrates platooning; b) lane change of a faster vehicle; c) giving way by a slower vehicle

Examples for norms applying to individual vehicles are:

- *Maximum speed limit:* Such norms recommend an upper speed limit at certain sections of a lane.
- *Minimum speed limit:* These norms are activated, for instance at up-hill sections on certain lanes. They aim at preventing slow vehicles from occupying the lanes that shall be scheduled for the faster ones.
- *Stop at red light:* Such norms prohibit crossing against red light. Compliance of this norm is usually controlled for each traffic light.
- *Prohibition of lane change:* Such norms discourage or prohibit lane changes in heavy-traffic, dangerous road sections, or close to exits.

Group norms require the cooperation of several vehicles:

- *Optimal platooning distance*: This norm specifies an interval for the distance that is recommended between vehicles driving in a convoy on the same lane.
- *Polite lane change by faster vehicles*: Vehicles that aim to drive significantly faster than the vehicle ahead shall change one lane left, if the safety distances to vehicles on the new lane are respected when taking their speed and the possible acceleration of the vehicle itself into account.
- *Polite lane change by slower vehicles*: Vehicles that aim to drive significantly slower than the vehicle behind shall change one lane right. Again, the safety distances to vehicles driving on the right lane have to be considered.

3 Norms-Based Multi-Agent Architecture

In this section, we describe an architecture for indirect norms-based control of semi-autonomous vehicles. Our conceptual architecture for norms-based MAS control consists of the controlled MAS (below) and of the electronic institutions institution (system controller, above). In this paper, our focus is on the agents' viewpoint towards norms rather than on the methods of the norms provision, norms/system monitoring and efficiency analysis from the institution point of view — the system controller is considered a "black box" for the agents. Thus, we only provide a brief sketch of the institution side: An institution has three primary tasks: (i) exert indirect control over the system by dynamically providing norms to components (agents); (ii) observe agents and obtain a model by mining their behavior; and (iii) calculate values of system metrics to evaluate efficiency, reliability and controllability of system operation. Based on the metrics and the system model, decisions are made, e.g., about provision of new norms.

Our institutional model is based on the AAOL (Autonomous Agents in Organized Localities) modeling methodology. AAOL provides a metamodel for institution-based multiagent system, which in particular takes the distributedness of institutions and their range of influence into account. In particular, in AAOL, institutions can be associated with a physical or virtual space, a so-called *locality*. The locality of an institution determines the outreach of an institution in terms of the validity of norms as well as the power of norms enforcement. Using localities, we can describe a large range of control regimes, ranging from purely centralized over regional to decentralized. The AAOL model has been described in detail in [12]. In this paper, we extend AAOL by models and methods for constraint-based probabilistic decision-making.

Within a locality, norms act as (hard or soft) constraints to the agents' behavior (while the agent acts within the borders of the locality). We assume that initially each agent acts according to its own interests. Norms restrict this behavior by providing possible sanctions in order to avoid system failures and ineffective system operation due to egoistic behavior of some agents. Within a locality, we assume that norms refer to a system state which will be formalized as a tuple of values each expressing a parameter of a (sub)system as it may be observed by the institution at a certain point in time. A norm consists of (1) a *context* that indicates at which states of the system the norm shall be respected (the locality plus further pre-conditions), (2) a *normative* predicate that specifies the system states preferred by the norm and (3) an *incentive* which is expressed as a function that assigns a positive or negative reward in case the system state complies with or violates the norm. A norm is called *applicable* on an agent, if at least one variable that is addressed in the normative predicate is under the control of that agent. I.e., norms are expressed in terms of (projections of) system states as they are seen from the viewpoint of an institution.

An agent is assumed to be capable to correctly interpret a norm N; i.e., it is able to decide whether the pre-condition applies in the current situation, and it may take the consequences of sanctioning into account for evaluating and choosing its plans. According to its current state and information available, each agent creates a set of alternative plans, each of them containing a sequence of future actions. If egoistic behavior of the agents is supposed, each agent should evaluate the plans according to its preferences and select an optimal plan with respect to the rewards it can obtain by using the plan. Norms act as an additional plan evaluation criteria that force agents to take possible sanctions (i.e., negative rewards) into account.

We consider the two types of norms identified in Section 2:

- *Single-agent norms*, which restrict plans of single agents in certain situations;
- *Group norms*, which restrict joint plans of multiple agents. Note that N_g norms are not symmetric in general: they may provide different sanctions to the participating agents.

Each norm is associated with a sanction, i.e. a function that calculates a penalty (negative reward) to be paid by the agent if it violates the norm and has a certain state in the moment of the norm monitoring (e.g. *if* the speed limit is 50 km/h and at the moment of monitoring the speed was 73 km/h *then* the reward is -30). So each plan should be evaluated according to payoff and possible penalties and an optimal plan should be selected. In Section 4 we specify an overall reward function for that purpose.

However, a major complicating factor in this process is uncertainty, which appears in three guises: **uncertainty in plan execution**, **uncertainty in other agents' plans**, and **uncertainty in norms**. So a model is required that enables probabilistic situation forecasting [13].

Given these different flavors of uncertainty, our goal is to make probabilistic estimation of sanctions from single-agent norms for each individual plan selection and from group norms for each multiagent plan. As a method for deciding/detecting cooperative norm compliance, we propose Stochastic Distributed Constraint Optimization (SDCOP). It enables optimization of certain system-wide characteristics (in our case, overall reward) under a set of conditions for the entire system. As an optimization approach, resampling can be effective [14].

We propose that each agent be equipped by a "Probabilistic Constraint Optimizer" (PCO) module, which implements the mentioned SDCOP algorithm and detects an optimal plan for the agent from a given plan library under a given set of norms. For each plan in the plan library the PCO module forecasts the situations and estimates sanctions from the available norms. The agent can communicate with other agents in order to make joint plans and estimate sanctions from group norms. Then an optimal plan (with respect to a defined objective function, see Section 4) is selected for execution.

4 Formal constraint-based model of normative regulation

In this section we describe a formal model of norms-based decision-making of the vehicle agents. The norms are represented as soft and hard constraints over the system configurations as described in the previous sections.

We define a set of discrete system configurations $Conf = \{\langle \vec{c}_1(t), \vec{c}_2(t), \ldots, \vec{c}_m(t)\rangle \mid t \in \mathbb{N}\}$ where t denotes the discrete time. A global configuration $\mathbf{c}$ represents the system state to which the norms relate. It is composed of local configurations representing the current state of a component, e.g. a vehicle or a traffic light. Each subconfiguration $\vec{c}(t)$ is a vector $\vec{c}(t) = \langle v_1(t), v_2(t), \ldots, v_k(t)\rangle$, where each parameter describes a certain characteristics of the component (for example, $v_1(t)$ may describe the current speed of a vehicle, $v_2(t)$ its distance to the next intersection, v_3 its position etc.). Some of the parameters may be undefined, denoted by "$\perp$", at specific time steps.

The considered MAS consists of a set of n agents $AG = \{ag^1, ag^2, \ldots, ag^n\}$. Each agent ag^j has a set of the internal states $S^j = \{\langle s_1^j(t), s_2^j(t), \ldots, s_{n^j}^j(t)\rangle \mid t \in \mathbb{N}\}$, representing its internal information about itself and its environment. An internal state $\vec{s}^j(t) \in S^j$ can be translated to a system configuration using a translation function $U^j : S^j \to Conf$. We assume U^j to be a *correct*, but possibly partial view of agent j of the system configuration; i.e., if two agents i and j map their local states $\vec{s}^i(t)$ and $\vec{s}^j(t)$ to the system state, and for some parameter $v_k(t)$ within some subconfiguration $c(t)$ both mappings yield a value (i.e. the projections $U^i(\vec{s}^i(t))|_{c,k} \neq \perp \neq U^j(\vec{s}^j(t))|_{c,k}$), then $U^i(\vec{s}^i(t))|_{c,k} = U^j(\vec{s}^j(t))|_{c,k}$. A correct mapping ensures that any two agents (and the institution) coincide on their common view of the system.

An agent ag^j has a finite set of available actions Act^j. An action execution forces deterministic state change; the actions from Act^j are used to modify the parameters under the control of the agent, such as its speed. We further assume a set of oracle actions $\omega \in \Omega^j$ that we use to describe what change the agent observes in its environment as far as it is aware of it like e.g. the traffic light will be red in the next state and the vehicle ahead will be on position x. Formally, there is a deterministic transition function $T^j : S^j \times Act^j \times \Omega^j \to S^j$, which for each state-action pair returns the next state.

Each agent is able to generate a set of alternative plans, whenever needed. In this paper, we assume a planning capability of the agent to be given, so we do not consider details. For simplicity we choose a fixed planning horizon $T \in \mathbb{N}$ that is common to all the agents in the system. A plan $plan^j \in Plans^j$ is an ordered sequence of actions $plan^j = ((a_1, \omega_1), (a_2, \omega_2), \ldots, (a_T, \omega_T))$, $a_k \in Act^j, \omega_k \in \Omega^j$. An execution of the plan $plan^j$ at time t means that the agent selects the pair (a_1, ω_1) at time t, action (a_2, ω_2) at time $t+1$, …, action (a_T, ω_T) at time $t+T-1$. If the state $\vec{s}^j(t) \in S^j$ at the moment t is known and the abovementioned deterministic state transition schema is used, an execution of $plan^j$ generates a defined sequence of the states $\vec{s}^j(t+1), \vec{s}^j(t+2), \ldots, \vec{s}^j(t+T)$, where $\vec{s}^j(t+k) = T^j(\vec{s}^j(t+k-1), a_k, \omega_k)$, $k = 1, 2 \ldots, T$.

The set of norms $Norm$ is finite and fixed. A norm $n \in Norm$ is a tuple $\langle cond, pred, reward\rangle$ where $cond : Conf \to \mathbb{B}$ specifies the *enabling condition* of the norm, $pred : Conf \to \mathbb{B}$ the normative predicate describing the norm compliant states and $reward : Conf \times AG \to (\mathbb{R} \cup \{-\infty\})$ is a reward function, that formalizes the rewards (positive values) and sanctions (negative values) that are imposed on the agents when norm compliance is monitored. Within $Norm$ we distinguish single-agent norms N_a and group norms N_g: A single-agent norm n_a refers only to a single agent, meaning that the set of components on which subconfigurations $cond_a$ and $pred_a$ truly depend, contains exactly one agent[2]. We note that, however, the state of non-agent components may be relevant in a single agent norm, e.g. the norm may state that a vehicle (agent) must stop in case of a red traffic light ahead (component with property $v_{light} = red$). In a group norm n_g, the enabling condition and the normative predicate refer to the status of more than one agent.

In addition we require that at most those agents, on which the norm truly depends, may receive a non-zero reward by the *reward* function. The value $-\infty$ is used in a reward assignment for safety norms that must not be violated.

In order to reason about norms, an agent must be able to interpret a norm, by mapping its internal state to a system configuration and evaluating the enabling condition, the normative predicate and the reward function for itself. This implies that an agent's perception, its internal state and the translation function U^j need to be sufficiently complete, i.e. the agent must be aware of those parameters of a system configuration that are relevant to a norm. As translation functions are correct we can be sure, that all agents that are aware of a certain parameter coincide on its value.

As we described before, a norm aims to restricts an agent's behavior, however an agent can in principle violate it. However, due to its egoistic but rational behavior the agent will take possible sanctions imposed in case of norm violation into account when evaluating its plans.

Formally, the sanctions for agent j imposed by a norm $n \in Norm$ can be calculated for each state $\vec{s}^j(t+k)$ in the state sequence produced by the plan $plan^j$, by $reward(plan^j) = \sum_{k=0}^{T} reward(U^j(\vec{s}^j(t+k)), j)$. We note that via the sequence of oracle actions the effect of future behavior of other components and in particular other agents are taken into account as far as it is represented in the internal state of agent j. However, the prediction is individual for agent j.

A *joint* plan means that a group of agents shares each others plans. Here we will model joint plans as a coincidence of the action sequence of one agent with the oracle sequences on its behavior by the other group members regarding the effects on those parameters that are under the control of the group of agents: I.e. let $A \subseteq AG$ be a group of agents. Then $plan^A = \{plan^j \mid plan^j \in Plans^j \text{ for each } j \in A\}$ is called a *joint plan* of A at time t_0 iff for all $i \in A$ the following holds: If $v^{(i)}$ is the mth parameter in the subconfiguration $\vec{c}$ that represents agent i on the system level and any agent $j \in A$ that is aware of agent i's property $v^{(i)}$, i.e. $U^j(\vec{s}^j(t))$ yields a value for $v^{(i)}$, then for the two state sequences $\vec{s}^i(t_0+1), \vec{s}^i(t_0+2), \ldots, \vec{s}^i(t_0+T)$ generated from $plan^i$ and $\vec{s}^j(t_0+1), \vec{s}^j(t_0+2), \ldots, \vec{s}^j(t_0+T)$ generated from $plan^j$ we have

$$v^{(i)} = U^i(\vec{s}^i(t_0+k))|_{c,m} = U^j(\vec{s}^j(t_0+k))|_{c,m} \text{ for } k = 1, \ldots T,$$

[2] We say a predicate or function p *truly depends on* a component if two system configurations s_1,s_2 exist which differ only by this component, i.e. $s_1 = \langle c_1, c_2, \ldots, c_i, \ldots, c_k \rangle$ and $s_2 = \langle c_1, c_2, \ldots, c_i', \ldots, c_k \rangle$, and $p(s_1) \neq p(s_2)$.

i.e. agent j's prediction on the future of agent i coincides with i's plan. Thus the notion of a joint plan formalizes that the agents somehow share their plans. In case of a joint plan, each agent of the group may evaluate the norms individually with respect to its individual rewards. However, it can be sure that the others will behave accordingly. An agent evaluates a $plan^j \in Plans^j$ by summing up the sanctions $reward(plan^j)$ it causes over all applicable norms.

However, in general not every norm violation is monitored and sanctioned. Thus we assume a agent-local likelihood information that may result from its former experience. In order to model it, we introduce an experience function $F^j(n) : S^T \to P(\mathbb{R} \cup \{-\infty\})$, which is associated by agent j with the each norm $n \in Norm$. $F^j(n)$ returns for a sequence of states a probability distribution of the rewards. A sanction or incentive of a plan $plan^j$ of the agent j caused by the norm n is a random variable $X^j_{plan,n}$ with a distribution $F^j(plan^j, n)$;

Then the reward from all norms is also a random variable defined as

$$R^j_{norm}(plan^j) = \sum_{n \in Norm} X^j_{plan^j,n}$$

Each agent $ag^j \in AG$ further has an individual reward function $R^j_{goal}(plan^j)$, which calculates the usability of a plan $plan^j$. The reward function lets us compare the result of different plans relative to the agent goals (for example, the agent wants to reach its destination as quickly as possible; in this case $R^j_{goal}(plan^j)$ may be measured as costs of time).

A total reward $R^j(plan^j)$ of the plan $plan^j$ is calculated as

$$R^j(plan^j) = R^j_{norm}(plan^j) + R^j_{goal}(plan^j).$$

The final reward is a random variable. As a criterion for an optimal plan selection the agent takes into account its expectation $E[R^j(plan^j)]$ and variance $Var[R^j(plan^j)]$

$$Eff^j(plan^j) = \alpha^j E[R^j(plan^j)] + \beta^j \sqrt{Var[R^j(plan^j)]}.$$

Cooperative planning is a process of a maximizing the final reward over a group of agents $A \subseteq AG$ in the system. The goal of the group is to select a joint plan $\{plan^{*1}, plan^{*2}, \ldots, plan^{*n}\} \in Plans^A$ such that the total expected efficiency of the system is maximized:

$$Eff = \max_{Plans^A} \sum_{ag^j \in A} Eff^j(plan^A|^j)$$

The formulated problem is a problem of stochastic distributed constraint optimization.

5 Example

We illustrate our approach by using a simple example. Let us consider a road segment, where three vehicles ag^1, ag^2 and ag^3 are situated. The planning horizon of all the vehicles is $T = 4$.

We use a variant of the Nagel-Schreckenberg traffic model, which means that the road is split to cells and each vehicle occupies exactly one cell. The state of each vehicle j is described by a vector $S^j = \langle s_1^j(t), s_2^j(t), s_3^j(t)\rangle$, where $s_1^j(t)$ denotes the position of the vehicle (number of cells from beginning of the road), $s_2^j(t)$ denotes the lane and $s_3^j(t)$ denotes speed. We also assume that a vehicle can accelerate/decelerate by one cell per time unit. The movement of a vehicle is possible forward (F) or forward with lane change left (L) or right (R); a vehicle can move to one lane left or right. The relative speeds of the vehicles are expressed by the number of cells which they pass per time unit. We further assume that a vehicle can accelerate (A) or decelerate (D) by one cell per time unit or not change its speed (N). So the set of actions Act^j for each agent j consists of nine actions: $Act^j = \{FA, FD, FN, LA, LD, LN, RA, RD, RN\}$. We finally assume that the subconfigurations $c_i(t)$ of the system configuration $Conf$ are equal to the agent states S^j, i.e. the function U^j is an identity function.

The initial states of the vehicles are $S^1 = \langle 1,2,3\rangle$, $S^2 = \langle 5,2,2\rangle$ and $S^3 = \langle 6,1,1\rangle$. There are the following norms, which are enabled ($cond = true$):

- n_1: Maximum speed limit for all lanes is 4. The norm compliant states are $pred : s_3^j \leq 4$ and the reward function of the state $reward(s_3^j) = -(s_3^j - 4) * 10$
- n_2: Minimum following distance for all lanes is 1. The norm compliant states are $pred : \forall k \neq j : (s_2^j = s_2^k) \& (s_2^k > s_2^j) \rightarrow s_1^k - s_1^j > 1$ and the reward function of the first vehicle $reward(s_1^j, s_1^k) = -20$ and of the second vehicle $reward(s_1^k, s_1^j) = -1$
- n_3: Safety distance for the lane change is 1. The norm compliant states are $pred : \forall k \neq j : (s_2^j(t-1) - s_2^k(t-1) = 1) \& (s_2^j(t) = s_2^k(t)) \& (s_2^j > s_2^k) \rightarrow s_1^j(t) - s_1^k(t) > 1$ and the reward function of the first vehicle $reward(s_1^j, s_1^k) = -20$ and of the second vehicle $reward(s_1^k, s_1^j) = -1$

The agents consider the following alternative plans: $plan_1^1$ = {FN, FN, FN, FN}; $plan_2^1$ = {FA, FN, FD, FN}; $plan_1^2$ = {FN, FN, RN, FN}; $plan_2^2$ = {FN, RN, FN, FN}; $plan_1^3$ = {FN, FN, FN, FN}. Tables 1, 2 and 3 list the estimated sanctions from the norms for the vehicles ag^1, ag^2 and ag^3 correspondingly.

Table 1 Sanctions for the plans of the vehicle ag^1

Plan	Norm n_1	Norm n_2		Norm n_3		Reward
		$plan_1^2$, $plan_1^3$	$plan_2^2$, $plan_1^3$	$plan_1^2$, $plan_1^3$	$plan_2^2$, $plan_1^3$	
$plan_1^1$	0	0	0	0	0	10
$plan_2^1$	-10, $p = 0.1$	-20, $p = 0.2$	0	0	0	20

Table 2 Sanctions for the plans of the vehicle ag^2

Plan	Norm n_1	Norm n_2		Norm n_3		Reward
		$plan^1_1$, $plan^3_1$	$plan^1_2$, $plan^3_1$	$plan^1_1$, $plan^3_1$	$plan^1_2$, $plan^3_1$	
$plan^2_1$	0	0	-1, $p = 0.2$	0	0	10
$plan^2_2$	0	0	0	-30, $p = 0.3$	-30, $p = 0.3$	10

Table 3 Sanctions for the plan of the vehicle ag^3 (reward = 10)

Norm n_1	0			
	$plan^1_1$, $plan^2_1$	$plan^1_1$, $plan^2_2$	$plan^1_2$, $plan^2_1$	$plan^1_2$, $plan^2_2$
Norm n_2	0	0	0	0
Norm n_3	0	-1, $p = 0.3$	0	-1, $p = 0.3$

Now any distributed constraint optimization algorithm [9] can be applied to find optimal combinations of the agent plans. In our simple case there are only 4 possible combinations of plans, which are shown in Table 4.

Table 4 Summarized sanctions for possible combinations of the vehicle plans

Plans	ag^1	ag^2	ag^3	Sum	Eff
$plan^1_1$, $plan^2_1$, $plan^3_1$	10	10	10	30	30
$plan^1_1$, $plan^2_2$, $plan^3_1$	10	10, $p = 0.7$ -20, $p = 0.3$	10, $p = 0.7$ 9, $p = 0.3$	30, $p = 0.7$ -1, $p = 0.3$	20.7
$plan^1_2$, $plan^2_1$, $plan^3_1$	-10, $p = 0.02$ 0, $p = 0.18$ 10, $p = 0.18$ 20, $p = 0.72$	10, $p = 0.8$ 9, $p = 0.2$	10	9, $p = 0.02$ 19, $p = 0.18$ 30, $p = 0.18$ 40, $p = 0.72$	37.8
$plan^1_2$, $plan^2_2$, $plan^3_1$	20, $p = 0.9$ 10, $p = 0.1$	10, $p = 0.7$ -20, $p = 0.3$	10, $p = 0.7$ 9, $p = 0.3$	40, $p = 0.63$ 30, $p = 0.07$ 9, $p = 0.27$ -1, $p = 0.03$	29.7

We see that the combination of plans $plan^1_2$, $plan^2_1$, $plan^3_1$ corresponds to maximal system efficiency 37.8 and will be selected in the considered situation.

6 Conclusion

In this paper, we proposed a constraint-based decision-making model for vehicles in cooperative traffic management. The model supports indirect regulation of the vehicles by a (centralized or federated) authority while preserving and respecting the autonomy of traffic participants. We illustrated our approach by a simple use case scenario and provided a formalism based on stochastic distributed constraint optimization (SDOP). While this paper has outlined the conceptual and algorithmic cornerstones of our approach, numerous future activities are on our research agenda. The next steps will be to provide a detailed description of corresponding optimization algorithms including their implementation and evaluation both in terms of computational complexity / tractability and with real-world traffic data obtained from the PLANETS research project [1]. Longer term issues relate to the study of more expressive norms semantics (in particular by using temporal logic languages), the consideration of more elaborate methods for norms design, norms emergence, and norms efficiency evaluation.

References

1. J. Görmer, J. Ehmke, M. Fiosins, D. Schmidt, H. Schumacher, and H. Tchouankem, "Decision support for dynamic city traffic management using vehicular communication," in *Proc. of 1st Int. Conf. on Simulation and Modeling Methodologies, Technologies and Applications (SIMULTECH)*. SciTePress.
2. J. P. Müller, "The design of autonomous agents - a layered approach," *LNAI*, vol. 1177, 1996.
3. R. Conte, C. Castelfranchi, and F. Dignum, "Autonomous norm-acceptance," *Intelligent agents V: Agent theories, architectures, and languages*, pp. 99–112, 1999.
4. K. Dresner and P. Stone, "Multiagent traffic management: A reservation-based intersection control mechanism," in *3rd International Joint Conference on Autonomous Agents and Multiagent Systems*, July 2004, pp. 530–537.
5. O. Boissier, J. Padget, V. Dignum, G. Lindemann, E. Matson, S. Ossowski, J. Sichman, and J. e. Vazquez-Salceda, "Coordination, organizations, institutions and norms in multi-agent systems," *LNCS*, vol. 3913, 2006.
6. G. Boella, L. W. N. van der Torre, and H. Verhagen, "Introduction to normative multiagent systems," in *Normative Multi-agent Systems*, ser. Dagstuhl Seminar, vol. DBLP:conf/dagstuhl/2007P7122, 2007.
7. M. Esteba, J. Rodriguez-Aguilar, C. Sierra, P. Garcia, and J. L. Arcos, "On the formal specification of electronic institutions," *Agent-mediated Electronic Commerce IV, LNCS*, vol. 1991, pp. 126–147, 2001.
8. N. Lacey and H. Hexmoor, "A constraint-based approach to multiagent planning," in *Proc. of 13th Midwest AI and Cognitive Science Conference*, 2002, pp. 1–6.
9. M. Yokoo, *Distributed Constraint Satisfaction: Foundations of Cooperation in Multi-agent Systems (Springer Series on Agent Technology)*. Springer, 2001.
10. E. Bowring, M. Tambe, and M. Yokoo, "Multiply-constrained distributed constraint optimization," in *Proc. of the 5-th Int. joint Conf. on Autonomous agents and multiagent systems (AAMAS'06)*, 2006, pp. 1413–1420.
11. M. Fiosins, J. Fiosina, J. P. Müller, and J. Görmer, "Reconciling strategic and tactical decision making in agent-oriented simulation of vehicles in urban traffic," in *Proc. of the 4th Int. ICST Conf. Simulation Tools and Techniques*. ACM Digital Library, 2011, pp. 144–151.
12. M. Huhn, J. P. Müller, J. Görmer, G. Homoceanu, N.-T. Le, L. Märtin, C. Mumme, C. Schulz, N. Pinkwart, and C. Müller-Schloer, "Autonomous agents in organized localities regulated by institutions," in *Proc.of IEEE DEST 2011*, 2011, pp. 54–61.
13. J. Fiosina and M. Fiosins, "Cooperative kernel-based forecasting in decentralized multi-agent systems for urban traffic networks," in *Proc. of Ubiquitous Data Mining (UDM) Workshop at the 20th European Conf. on Artif. Intelligence*, 2012, pp. 3–7.
14. M. Fiosins, J. Fiosina, and J. P. Müller, "Change point analysis for intelligent agents in city traffic," in *Agents and Data Mining Interaction, LNAI*, L. Cao, A. Bazzan, A. Symeonidis, V. Gorodetsky, G. Weiss, and P. Yu, Eds., 2012, vol. 7103, pp. 195–210.