A traffic flow cellular automaton model to considering drivers' learning and forgetting behaviour*

Ding Jian-Xun(丁建勋), Huang Hai-Jun(黄海军)†, and Tian Qiong(田 琼)

School of Economics and Management, Beihang University, Beijing 100191, *China*

(Received 26 July 2010; revised manuscript received 30 August 2010)

It is known that the commonly used NaSch cellular automaton (CA) model and its modifications can help explain the internal causes of the macro phenomena of traffic flow. However, the randomization probability of vehicle velocity used in these models is assumed to be an exogenous constant or a conditional constant, which cannot reflect the learning and forgetting behaviour of drivers with historical experiences. This paper further modifies the NaSch model by enabling the randomization probability to be adjusted on the bases of drivers' memory. The Markov properties of this modified model are discussed. Analytical and simulation results show that the traffic fundamental diagrams can be indeed improved when considering drivers' intelligent behaviour. Some new features of traffic are revealed by differently combining the model parameters representing learning and forgetting behaviour.

Keywords: cellular automaton model, learning and forgetting behaviour, Markov property

PACS: 89.40.–a, 45.70.Vn, 64.60.Ak **DOI:** 10.1088/1674-1056/20/2/028901

1. Introduction

Since the 1990s, there has been an increased awareness of cellular automata (CA) models. CA models have been used in various dynamic systems in physics, such as the biological self-reproduction system, the thermodynamic system and other evolution systems. These models arise from the physical discipline of statistical mechanics and have an ultimate goal of reproducing natural macroscopic behaviour based on a simple description of microscopic interactions. This has led to the use of alternative technologies in the traffic flow field. There are other different approaches towards the problem, such as car-following theory, hydrodynamic fluid theory and other approaches that use macroscopic models.[1−3]

Most researchers attempt to deal with the prediction and modeling of traffic flow using CA models (see Refs. [4] and [5]). This is of interest for both applications (i.e., road system with ramps in Ref. [6] and intersections in Ref. [7]) and theoretical investigations because CA models are capable of modeling interaction between individual vehicles and capturing the macroscopic properties of traffic flow based on a minimal set of rules. Following Wolfram's Rule 184 (CA-184), Nagel and Schreckenberg (NaSch) proposed a minimal CA model of traffic flow that is able to reproduce basic phenomena encountered in real traffic, such as soliton waves, phantom jams, stop-and-go waves and synchronized flow patterns.[8] This model is called the NaSch model in the literature.

For an idea of what has recently been done in this field, in magnitude if not in specifics, we briefly look to the past. Numerous studies have focused on the acceleration and deceleration rules of vehicles, and the randomization rule of vehicle velocity.[9−11] These rules were added into the original NaSch model. With these extensions, the CA models are becoming increasingly more effective in simulating traffic phenomena.

With regard to whether all vehicles' characteristics are homogeneous, we can classify previous works into two groups. In the first group, traffic is assumed to be homogeneous, namely, all of the drivers' characteristics are identical, for example, the CA-184 model and the DFI-TCA model proposed by Fukui and Ishibashi.[9] These two models only deal with the deterministic behaviour of the system and are essentially all congested patterns stemming from the initial conditions. The stochastic CA models allow for the spontaneous emergence of phantom jams, like the NaSch model, which is also within this group. In contrast to the fixed randomization probability (P) of

*Project supported by the National Natural Science Foundation of China (Grant No. 70821061) and the National Basic Research Program of China (Grant No. 2006CB705503).

†Corresponding author. E-mail: haijunhuang@buaa.edu.cn

© 2011 Chinese Physical Society and IOP Publishing Ltd

vehicle velocity used in the NaSch model, Barlovic *et al.* proposed a randomization model in which P is a function of the vehicle velocity.[12] Recent studies by Jiang and Wu have suggested that the sensitivity of a driver is affected by the stopped time of the vehicle, i.e., P should be an increasing function of the stopped time.[13] Simulation by these stochastic models can reproduce the metastable state of high flow and all other three phases observed in real traffic (i.e., free-flow, synchronized and wide jams). The cause of the randomization rule still remains unexplained and it is then criticized for having no realistic meaning. However, it has been recognized that, in the real world, traffic flow is normally a mix of several types of vehicles, such as cars, buses and trucks. Different types of vehicles may produce heterogeneous traffic flow phenomena.[14−19] Relevant studies have demonstrated that mixed traffic flow can lead to platoon formations, even at low densities in single-lane traffic flow.

However, the studies in Refs. [13]–[19] only focus on the external differences of drivers. Until now, there has been less investigation of internal diversity among drivers. In literature, no works have been reported to prove that different drivers with different behaviour characters related to psychology (e.g., the learning, forgetting and other intelligent behaviour) could also lead to platoon formations. Considerable works are devoted to modifying the rules associated with the randomization probability of velocity, let P relate to the external driving conditions, not to the evolution randomization alone. Obviously, if the randomization is regarded as an overreaction or careless driving behaviour, an unchangeable value of P may sound unrealistic. Such behaviour as overreaction or careless driving generally occurs with a time-varying P.

Hence, it may be descriptively unrealistic to assume that the randomization probability of vehicle velocity is constant over time. This assumption should be replaced by a more natural rule, inspired by the idea of a reinforcement learning scheme.[20,21] In this paper, we will formulate a historical randomization behaviour-dependent CA model for traffic flow.

The paper is organized as follows. The basic NaSch CA model and our modified one are presented in Section 2. The Markov properties of our modified model are discussed in Section 3. Special attention is paid to the newly introduced parameters. Section 4 presents the simulation results generated by our modified model and the basic NaSch model. Comparisons between the results obtained using these two models are made. Section 5 concludes the paper.

2. The CA models

For the sake of completeness, we first briefly recall the basic NaSch model and then present our modifications of it.

2.1. The basic NaSch CA model

The basic NaSch model is subjected to single-lane traffic. The lane is subdivided into discrete cells and the physical length of each cell is generally 7.5 m. A cell can be either empty or occupied by exactly one vehicle. All vehicles move forward according to simple rules. The state of the vehicle n is determined by its speed v_n and the position x_n. The value of v_n lies between 0 and $v_{\max}$. Here, the maximum speed is given and fixed. The model is composed of velocity and position updating rules as follows (parallel dynamics): Rule 1: acceleration, $v_n(t+1) \to \min(v_n(t)+1, v_{\max})$; Rule 2: deceleration for avoiding accidents, $v_n(t+1) \to \min(v_n(t+1), d_n)$; Rule 3: velocity randomization with a certain probability P_0, $v_n(t+1) \to \max(v_n(t+1)-1, 0)$; Rule 4: movement, $x_n(t+1) \to x_n(t) + v_n(t+1)$.

In the above model, t denotes the time step of the simulation, d_n is the number of empty cells between a vehicle and its leading vehicle $n-1$, $d_n = x_{n-1}(t) - x_n(t) - 1$. The model states that every driver always accelerates if the vehicle has not reached $v_{\max}$ but brakes to avoid accidents. The probability P_0 plays a role in reflecting the drivers' randomization of finally determining their velocities. It is clear that, in the NaSch model, the randomization rule captures natural speed fluctuations due to human behaviour or varying external conditions. This rule introduces the overreactions of drivers when braking and then leads to the formation of spontaneously emerging jams.

2.2. Our modified CA model

In this subsection, we show that the human desire for self-adjustable randomization behaviour plays an important role in driving vehicles on the road. This consideration is added into the basic NaSch model and causes the randomization behaviour to be closely consistent with driving psychology.

In contrast to the original NaSch model, in the so-called learning-randomization model proposed in this paper, we let the randomization parameter depend on the driver's previous randomization behaviour, i.e., $P(t+1) = f(P(t))$, where $P(t)$ will increase or decrease according to traffic conditions. Then, Rule 3 in Subsection 2.1 is supplemented by

$$P(t+1) = \begin{cases} (1-\alpha)P(t), & \text{if a randominization occurs at } t, \\ \min\{P(t) + \beta(1-P(t)), P_0\}, & \text{otherwise,} \end{cases} \tag{1}$$

where P_0 is a constant. Equation (1) identifies two types of velocity randomization behaviour, namely learning and forgetting. In the learning situation, the randomization occurs at t and continues to do so at $t+1$, the probability is then given by

$$P(t+1) = (1-\alpha)P(t), \tag{2}$$

where α is the learning parameter, $0 \leqslant \alpha \leqslant 1$. When $\alpha = 0$, we have $P(t+1) = P(t)$. This means that the randomization probability remains unchanged from t to $t+1$. When $\alpha = 1$, $P(t+1) = 0$ holds. This implies that perfect learning has been realized, i.e., the driver completely accepts the velocity determined by Rules 1 and 2. It is clear that $P(t+1) \leqslant P(t)$ always holds in the evolution process. Thus, learning can stop the decrease of velocity by reducing the randomization probability.

In the forgetting situation, we assume that randomization does not occur at t (i.e., drivers do not care if there is randomization in the last time step) but it may exist at $t+1$. The randomization probability at $t+1$ is given by

$$P(t+1) = \min\{P(t) + \beta(1-P(t)), P_0\}, \tag{3}$$

where β is a forgetting parameter, $0 \leqslant \beta \leqslant 1$. When $\beta = 0$, we have $P(t+1) = \min\{P(t), P_0\}$, which assures $P(t+1) \leqslant P(t) \leqslant P_0$ in the evolution process. Hence, if P_0 is the initial probability and $P(1) = P_0$ is defined, our modified model becomes the basic NaSch model when $\alpha = 0$, i.e., $P(t) = P_0$ for all t, regardless of learning or forgetting situation. When $\beta = 1$, we have $P(t+1) = P_0$ for $P_0 < 1$. This states that the probability at $t+1$ will increase in comparison with that at t if $P(t) < P_0$. Thus, forgetting may decrease the drivers' velocities by increasing the randomization probability in the evolution process.

In the next section, we will further discuss the properties of our modified model when α and β take some specific values.

3. The Markov properties of our model

We can regard the series of evolutionary randomization behaviour as a stochastic process, i.e., $X = \{X_t, t \in T^+\}, X_t \in E, T^+ \in \{1, 2, \ldots\}$, the state space $E = \{0, 1\}$ where 0 means the randomization occurs and 1 not. As shown in Fig. 1, there are two states and four transition probabilities $\{P_{i,j}(t+1) : i, j = 0, 1\}$ from time step t to $t+1$, as follows:

$$\begin{aligned} P(t+1) &= \begin{bmatrix} P_{0,0}(t+1) & P_{0,1}(t+1) \\ P_{1,0}(t+1) & P_{1,1}(t+1) \end{bmatrix} \\ &= \begin{bmatrix} P'(t) & 1-P'(t) \\ P''(t) & 1-P''(t) \end{bmatrix}, \end{aligned} \tag{4}$$

where the transition probability $P_{i,j}(t+1)$ is the conditional probability of being in state j "tomorrow" given that we are in state i "today". It follows that

$$\begin{aligned} P_{0,0}(t+1) &= P'(t+1) = P(X_{t+1} = 0 | X_0, X_1, \ldots, X_t) \\ &= P(X_{t+1} = 0 | X_t = 0) = (1-\alpha)P(t), \end{aligned} \tag{5}$$

$$\begin{aligned} P_{1,0}(t+1) &= P''(t+1) = P(X_{t+1} = 0 | X_0, X_1, \ldots, X_t) \\ &= P(X_{t+1} = 0 | X_t = 1) \\ &= \min\{P(t) + \beta(1-P(t)), P_0\}. \end{aligned} \tag{6}$$

Observing Eqs. (5) and (6), we easily find that given $(X_0, \ldots, X_t)$, the conditional distribution of X_{t+1} depends only on X_t. This phenomenon is called the memoryless property, also known as the Markov property. Thus, the stochastic process studied in this paper is a two-state Markov chain.

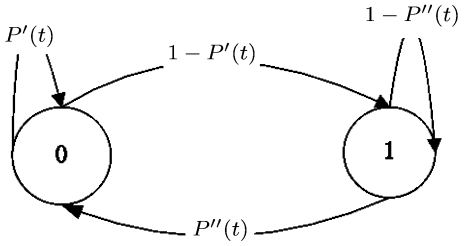


Fig. 1. Transition probabilities of the states on space E.

The purpose of the update chain above is to increase an action probability that produces a favourable response (i.e, to be confronted with no randomization behaviour). Conversely, for an unfavourable response (i.e, the randomization behaviour occurs), the corresponding action probability decreases. This simple update principle is based on the reinforcement learning scheme.

The initial distributions of the transition probabilities are as follows:

$$\begin{cases} P(X_0 = i) = P_i(0), P_i(0) \geqslant 0, \\ \sum\limits_{i \in E} P_i(0) = 1, i \in E, \\ [P_0(0), P_1(0)] = [P_0, 1 - P_0]\,. \end{cases} \tag{7}$$

Note that $P_0 = 0$ implies that no randomization will occur while driving and $P_0 = 1$ means that randomization will definitely occur. Therefore, further investigations should be conducted of $P_0 \in (0, 1)$ together with α and $\beta \in [0, 1]$. Different driving strategies can be simulated by taking α and β as different values. Considering this, we shall study the following transition graphs of the states underlying the evolutionary schemes.

3.1. Scheme without learning and forgetting (the basic NaSch model, $\alpha = 0$)

Figure 2(a) demonstrates that the basic NaSch model is reached as there is no internal or external driving force to change $P(t)$. Thus, the evolution of the randomization probability can be regarded as a Markov transition process with a stationary transition matrix

$$P(t) = \begin{bmatrix} P_0 & 1 - P_0 \\ P_0 & 1 - P_0 \end{bmatrix}, \quad t = 1, 2, \ldots. \tag{8}$$

Clearly, the transition matrix $P(t)$ is irreducible for all possible communications $i \leftrightarrow j$ of states, $i, j \in (0, 1)$. Here, we say that state i communicates with state j if j is from i ($i \to j$) with a positive probability (i.e., $P_{i,j}(t) > 0$) at time step t. In addition, such a Markov chain is said to be aperiodic (irreducibility is a property that all states of the Markov chain can be reached from others). Because $P_{i,i}(t) > 0$ for all t and all states i, thus the chain is aperiodic. Therefore, $X = \{X_t, t \in T^+\}$ is an irreducible and aperiodic Markov chain with state space $E = \{0, 1\}$ and transition matrix $P(t)$.

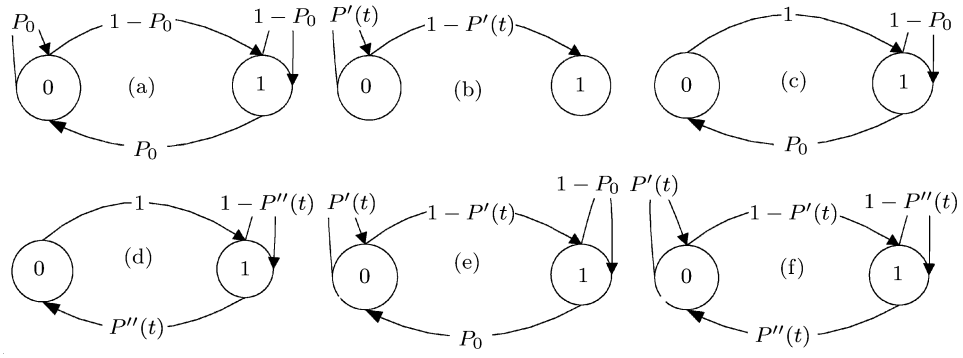


Fig. 2. Transition probabilities of the states: (a) $\alpha = 0$; (b) $\alpha = 1$ and $\beta \geqslant P_0$; (c) $\alpha = 1$ and $\beta = 0$; (d) $\alpha = 1$ and $0 < \beta < P_0$; (e) $\alpha \in (0, 1)$ and $\beta \geqslant P_0$; (f) $\alpha \in (0, 1)$ and $\beta < P_0$.

Now we turn to examine the long-term behaviour of the Markov chain under this scheme. Firstly, we introduce a definition of the hitting time as $T_{i,j} = \min\{t \geqslant 1 : x_t = j\}$ with the convention that $T_{i,j} = \infty$ if the Markov chain never visits state j. Then, we define the mean hitting time as $\tau_{i,j} = E[T_{i,j}]$, i.e., the expected time taken from state i to state j. When $j = i$, $\tau_{i,j}$ is called the mean return time of state i. Obviously, $P(T_{i,j} < \infty) = 1$ and $E[T_{i,j}] < \infty$ hold.

For a transition matrix P, if a row vector $\pi = \{\pi_1, \ldots, \pi_k\}$ satisfies $\pi_i \geqslant 0$, $\sum_{i=1}^{k} \pi_i = 1$, and $\sum_{i=1}^{k} \pi_i P_{i,j} = \pi_j$ for $j = (1, 2, \ldots, k)$, it is called a stationary distribution of the Markov chain. It is known that for an irreducible and aperiodic Markov chain, there is exactly stationary distribution. We can then find the mean return time of state i by $\pi_i = \lim\limits_{t \to \infty} P_{i,i}(t) = 1/\tau_{i,i}$. In the NaSch scheme, the stationary distribution is $\pi = \{p_0, 1 - p_0\}$. Therefore,

the limiting probability of the Markov process is

$$P^{\infty} = \lim_{t\to\infty} P^t = \begin{bmatrix} P_0 & 1-P_0 \\ P_0 & 1-P_0 \end{bmatrix}, \tag{9}$$

and the mean return time of state 0 is

$$\tau_{0,0} = E[T_{0,0}] = \frac{1}{P_0}. \tag{10}$$

3.2. Scheme of perfect learning and non-forgetting ($\alpha = 1$, $\beta = 0$)

As shown in Fig. 2(b), if the previous action is not a randomization, the randomization probability remains unchanged in the next time step. The evolution of the randomization probability can be regarded as a homogeneous Markov transition process, with the following stationary transition matrix

$$P(t+1) = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}, \quad t = 1, 2, \ldots. \tag{11}$$

Obviously, this Markov chain, which models the learning automaton, has only one absorbing state (i.e., state 1, hence there is no randomization at all after one randomization occurs). The stationary distribution is $\pi = [0, 1]$.

3.3. Scheme of perfect learning and rapid forgetting ($\alpha = 1$, $\beta \geqslant P_0$)

As shown in Fig. 2(c), under the condition $\alpha = 1$ and $\beta \geqslant P_0$, we find that a randomization occurs and the probability of randomization decreases abruptly to 0 in the next time step. Furthermore, once the previous action is not a randomization, the probability of randomization climbs abruptly up to P_0. Therefore, it is a scheme of perfect learning and rapid forgetting.

The evolution of the randomization probability can be regarded as a homogeneous Markov transition process with the following stationary transition matrix:

$$P(t+1) = \begin{bmatrix} 0 & 1 \\ P_0 & 1-P_0 \end{bmatrix}, \quad t = 1, 2, \ldots. \tag{12}$$

The corresponding stationary distribution is $\pi = \{P_0/(1+P_0),\ 1/(1+P_0)\}$. The limiting probability of the Markov process is

$$P^{\infty} = \lim_{t\to\infty} P^t = \begin{bmatrix} \dfrac{P_0}{1+P_0} & 1-\dfrac{P_0}{1+P_0} \\ \dfrac{P_0}{1+P_0} & 1-\dfrac{P_0}{1+P_0} \end{bmatrix}, \tag{13}$$

and the mean return time of state 0 is

$$\tau_{0,0} = E[T_{0,0}] = \frac{1+P_0}{P_0}. \tag{14}$$

It can be found that the mean return time of state 0 increases as P_0 goes down. Since $1/P_0 < (1+P_0)/P_0$, this scheme has a longer time interval between two consecutive randomization behaviours than that of the NaSch model.

3.4. Scheme of perfect learning and slow forgetting ($\alpha = 1$, $0 < \beta < P_0$)

As shown in Fig. 2(d), if randomization occurs, the probability of randomization abruptly decreases to 0 in the next time step. Furthermore, if the previous action is not a randomization, the probability of randomization slowly increases to P_0. The evolution of the randomization probability can be regarded as an inhomogeneous Markov transition process, with the following time-varying transition matrix:

$$P(t+1) = \begin{bmatrix} 0 & 1 \\ P''(t) & 1-P''(t) \end{bmatrix}, \quad t = 1, 2, \ldots. \tag{15}$$

The above transition matrix $P(t+1)$ is irreducible and the Markov chain is aperiodic. For any irreducible, aperiodic Markov chain with the transition matrix $P(t+1)$, we have $P(T_{i,j} < \infty) = 1$. Moreover, the mean hitting time from state i to state j is finite, i.e., $E[T_{i,j}] < \infty$. Hence, the state i is positively recurrent.

The mean return time of state 1, $\tau_{1,1}$, is finite and calculated as follows:

$$\begin{cases} f_{1,1}^{(1)} = 1 - P''(1) = 1 - \min\{P_0 + \beta(1-P_0), P_0\}, \\ f_{1,1}^{(t)} = \begin{cases} P''(1), & t = 2, \\ 0, & t = 3, 4, \ldots, \end{cases} \end{cases} \tag{16}$$

$$\begin{aligned} \tau_{1,1} = E[T_{1,1}] &= \sum_{t=1}^{\infty} t\ f_{1,1}^{(t)} \\ &= 1 + \min\{P_0 + \beta(1-P_0), P_0\} = 1 + P_0. \end{aligned} \tag{17}$$

In the long-run case, the stationary distribution π and the mean return time of state 0 are given as

$$\pi = \left[1 - \frac{1}{1+P_0}, \frac{1}{1+P_0}\right], \tag{18}$$

$$\tau_{0,0} = \frac{1}{\pi} = \frac{1+P_0}{P_0}. \tag{19}$$

Equation (18) shows that the mean return time of state 0 is dependent on P_0 and independent of (α, β).

Since $1/P_0 < (1+P_0)/P_0$, the scheme discussed in this subsection has a longer time interval between two consecutive randomization behaviours than that of the NaSch model.

3.5. Scheme of partial learning and rapid forgetting ($\alpha \in (0,1)$, $\beta \geqslant P_0$)

Figure 2(e) shows that once a randomization occurs, the probability of randomization decreases in the next time step. Furthermore, once the previous action is not a randomization, the probability of randomization rapidly increases to P_0 in the next time step. Therefore, this is a scheme of partial learning and rapid forgetting.

The evolution of the randomization probability can be represented as an inhomogeneous Markov transition process, with the following time-varying transition matrix:

$$P(t+1) = \begin{bmatrix} P'(t) & 1-P'(t) \\ P_0 & 1-P_0 \end{bmatrix}, \quad t = 1, 2, \ldots. \tag{20}$$

As with all of the transition probabilities $P_{i,j}(t) > 0$ for all t, the transition matrix, P is then irreducible and the Markov chain is aperiodic. For any irreducible, aperiodic Markov chain on the state space $E = \{0, 1\}$, we have $P(T_{i,j} < \infty) = 1$. Moreover, the mean hitting time from state i to state j is finite, i.e., $E[T_{i,j}] < \infty$. Thus, all of the states are positively recurrent.

The mean return time of state 0, $\tau_{0,0}$, is finite and calculated as follows:

$$\begin{cases} f_{0,0}^{(1)} = P'(1), \\ f_{0,0}^{(t)} = P_0 \left(1-P_0\right)^{t-2} \left(1-P'(1)\right), t = 2, 3, \ldots, \end{cases} \tag{21}$$

$$\begin{aligned} \tau_{0,0} = E\left[T_{0,0}\right] &= \sum_{t=1}^{\infty} t f_{0,0}^{(t)} \\ &= P'(1) + \sum_{t=2}^{\infty} t\ P_0 \left(1-P_0\right)^{t-2} \left(1-P'(1)\right) \\ &= \frac{1+P_0}{P_0} - (1-\alpha) = \frac{1}{P_0} + \alpha. \end{aligned} \tag{22}$$

From Eq. (22), we find that the value of $\tau_{0,0}$ is between those values given by the NaSch scheme and the scheme of perfect learning and rapid forgetting. In addition, the value of $\tau_{0,0}$ increases with α. When $\alpha = 1$, the scheme reduces to a scheme of perfect learning and rapid forgetting. When $\alpha = 0$, it becomes the NaSch scheme. The stationary distribution π is given as

$$\pi = \left[\frac{P_0}{1+\alpha P_0}, 1 - \frac{P_0}{1+\alpha P_0}\right]. \tag{23}$$

3.6. Scheme with partial learning and slow forgetting ($\alpha \in (0,1)$, $\beta < P_0$)

Figure 2(f) shows that once a randomization occurs, the probability of randomization decreases in the next time step. If the previous action is not a randomization, the probability of randomization slowly increases to P_0 in the next time step. Hence, it is a scheme of partial learning and slow forgetting.

The evolution of randomization probability can be considered as inhomogeneous Markov transition process, with the following transition matrix:

$$P(t+1) = \begin{bmatrix} P'(t) & 1-P'(t) \\ P''(t) & 1-P''(t) \end{bmatrix}, \quad t = 1, 2, \ldots. \tag{24}$$

From Fig. 2(f), we can find that $P_{i,j}(t) > 0$ at any time step t, the transition matrix $P(t+1)$ is said to be irreducible and the Markov chain is aperiodic. For any irreducible, aperiodic Markov chain on state space $E = \{0, 1\}$ and transition matrix P, we have $P(T_{i,j} < \infty) = 1$. Moreover, the mean hitting time from state i to state j is finite, i.e., $E[T_{i,j}] < \infty$. This is to say that state i is positively recurrent.

Since the transition matrix $P(t+1)$ consists of four time-variants under this scheme, it is difficult to obtain the mean return time of all of the states and the stationary distribution of the Markov chain. We can only perform a qualitative analysis on the Markov chain under this scheme. The role of α is to make $P(t)$ decline, while β increases. The higher the α-value, the more $P(t)$ decreases. The smaller the β-value, the slower $P(t)$ increases. In addition, the value of the mean return time of state 0 stays between the values from the NaSch scheme and the scheme of partial learning and slow forgetting.

3.7. Mean return time and flux

Let ρ_c be the critical density that leads to traffic jam. For $\rho \ll \rho_c$, the speed of a vehicle fluctuates between $v_{\max}$ and $v_{\max} - 1$ due to the randomization rule (NaSch model). We can compute the space-mean speed for a sufficiently low density by means of a weighted average. Thus, we have $\bar{v} = \bar{P}(t)(v_{\max} - 1) + (1 - \bar{P}(t))v_{\max} = v_{\max} - \bar{P}(t)$, where $\bar{P}(t)$ is the average probability for slowing down

the speed $v_{\max}$ to $(v_{\max}-1)$. The corresponding flux is $J_{\text{free}} = \rho(v_{\max} - \bar{P}(t))$. On the other hand, in the congested region (here, for simplicity, let $v_{\max} = 1$), $d_n(t) \leqslant \min[v_n(t)+1, v_{\max}]$ and $v_n(t) = d_n(t)$, the average distance-headway is $(\rho^{-1}-1)$. The exact expression of the flux becomes $J_{\text{jam}} = \rho v_n = (1-\bar{P}(t))(1-\rho)$, $v_n = (1-\bar{P}(t))(\rho^{-1}-1) + \bar{P}(t) \times 0 = (1-\bar{P}(t))(\rho^{-1}-1)$.

We derive the flux based on the investigated dynamic randomization behaviour. Suppose that the mean return time of state 0, $\tau_{0,0}$, is synonymous with the average time T_w required by the n-th vehicle for randomly slowing down its velocity v_n to $v_n - 1$ after the latest randomization behaviour. The T_w is determined by the randomization probability $P(t)$ through $T_w = \tau_{0,0}$. Let the mean hitting time from state 0 to state 1, $\tau_{0,1}$, be synonymous with the average time T'_w required by the n-th vehicle for randomly accelerating to the velocity $v_n + 1$ after the latest randomization behaviour. The T'_w is determined by the randomization probability $P(t)$ through $T'_w = \tau_{0,1}$. Then, we obtain an expression of the flux as follows:

$$\begin{cases} J_{\text{free}} = \rho(v_{\max} - \pi_0) = \rho\left(v_{\max} - \dfrac{1}{\tau_{0,0}}\right), & \text{(sufficiently low density)}, \\ J_{\text{jam}} = (1-\pi_0)(1-\rho) = \left(1 - \dfrac{1}{\tau_{0,0}}\right)(1-\rho), & \text{(highly congested region)}. \end{cases} \tag{25}$$

Note that in the above discussion, π_0 represents the stationary distribution and a limiting probability of state 0 (i.e., $\pi_0 = \lim\limits_{t\to\infty} P_{0,0}(t)$). Obviously, equation (25) is a good expression of the flux, with a consideration of the vehicles' dynamic randomization behaviour.

Equation (25) can be extended to the schemes discussed previously. For simplicity, we only provide the results that correspond to a sufficiently low density region as follows:

$$J_{\text{free}} = \begin{cases} \rho(v_{\max} - P_0), & \text{for } \alpha{=}0, \\ \rho\left(v_{\max} - \dfrac{P_0}{1+P_0}\right), & \text{for } \alpha = 1,\ \beta > 0, \\ \rho v_{\max}, & \text{for } \alpha = 1,\ \beta = 0, \\ \rho\left(v_{\max} - \dfrac{P_0}{1+\alpha P_0}\right), & \text{for } \alpha \in (0,1),\ \beta \geqslant p_0. \end{cases} \tag{26}$$

In the next section, these results are verified well by our simulation outputs.

4. Simulation results and discussion

4.1. Description of the simulation procedure

In our simulation, a system containing 1000 cells is studied under the periodic boundary condition (i.e., $L = 1000$). The length of each cell is 7.5 m, so the system is equivalent to a ring road of 7.5 km long. One iteration time step is 1 s, $v_{\max} = 5$ cells per time-step (i.e., 135 km/h), and $P_0 = 0.5$. For simplicity, only one type of vehicle is considered, i.e., all vehicles have the same maximum velocity.

At the beginning of simulation, all of the vehicles are randomly distributed on the road. Clearly, the global density remains constant during the entire period. The global density is given by

$$\rho = \frac{N}{L}, \tag{27}$$

where N denotes the number of vehicles.

In each simulation, $T = 11000$ time steps are carried out and the results of the last 1000 time steps are recorded for conducting statistical analyses. The average velocity v and the flux q are computed as follows ($t_0 = 10000$):

$$v = \frac{1}{T}\sum_{t=t_0+1}^{t=T}\left(\frac{1}{N}\sum_{i=1}^{N} v_i(t)\right), \tag{28}$$

$$q = \rho v. \tag{29}$$

The simulation results are depicted in the v–ρ and q–ρ fundamental diagrams, each with respect to a set of α and β values. Each curve in the fundamental diagrams is produced by 20 simulations through changing N from 0 to 1000. In the following subsections, we will analyse these diagrams in detail.

4.2. Analyses of the fundamental diagrams

A universal fact can be found from these diagrams: either v–ρ or q–ρ curve is divided into two segments that correspond to free-flow and congested flow regions, respectively. In the free-flow region, the velocity remains constant while the flux increases with density linearly; in the congested flow region, both the velocity and the flux decrease with density. This is certainly true, as the commonly used fundamental diagram should demonstrate. However, the maximal values to which velocity and flux can reach are different against different α and β values, as explained below.

Figure 3 is produced by the basic NaSch model in which α is 0. It can be seen that there is a critical density, $\rho_c \approx 0.08$. The velocity remains a maximal value 4.5 when $\rho \leqslant \rho_c$ and the flux reaches a maximal value 0.332 when $\rho = \rho_c$. Figure 3 also shows that the fundamental diagram is not sensitive to the β value, provided $\alpha = 0$. This coincides with the analysis made previously.

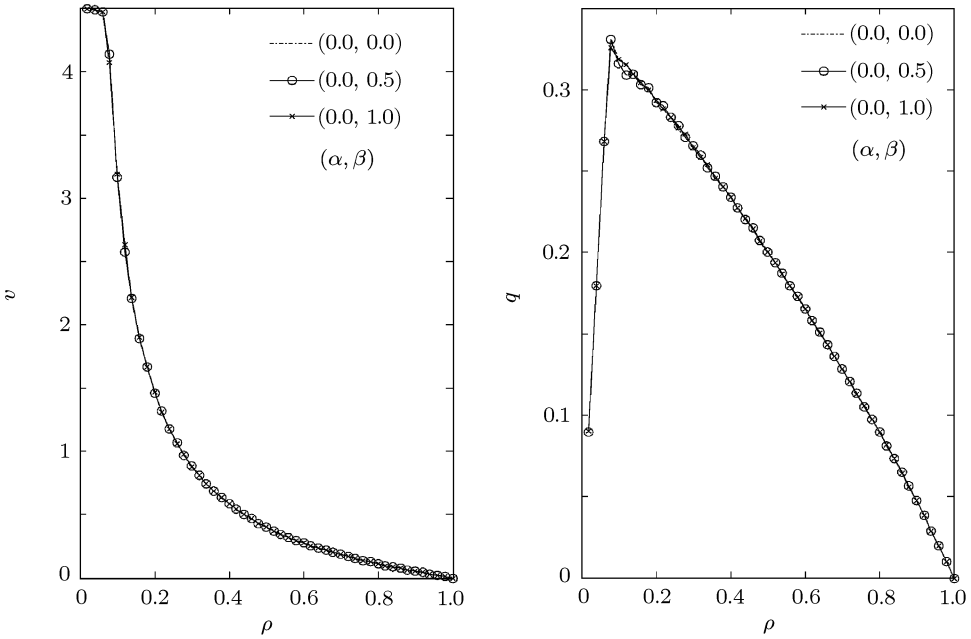


Fig. 3. Fundamental diagrams under the NaSch scheme ($\alpha = 0$).

Figure 4 depicts the fundamental diagrams under schemes of perfect learning ($\alpha = 1$) together with the non-forgetting ($\beta = 0$), rapid forgetting ($\beta \geqslant P_0$) and slow forgetting schemes ($\beta < P_0$). For comparison, the fundamental diagram under the NaSch scheme is also given in this figure.

Figure 4 shows that in comparison with the basic NaSch scheme, the critical density for remaining the maximal velocity or realizing the maximal flux is increased under the perfect learning scheme. The velocity and flux under the perfect learning scheme are always larger than that under the basic NaSch scheme. Particularly, when $\beta = 0$ (i.e, there is no forgetting behaviour), the fundamental diagrams are on top of others. When $\beta \geqslant P_0$ (i.e., rapidly forgetting), the curves v–ρ overlap basically, so do the curves q–ρ. When $\beta < P_0$ (i.e., slowly forgetting), the fundamental diagrams move up with decreasing β-value. These findings are consistent with the analyses made in Section 3.

Figure 5 depicts the fundamental diagrams under schemes of partial learning (i.e., $0 < \alpha < 1$) together with rapid forgetting (i.e., $\beta \geqslant P_0$). It can be seen that there is no big difference between the curves v–ρ given by the basic NaSch model and our modified model. However, the flux by the modified model is significantly larger than that by the basic NaSch model. For a given β-value, the curve q–ρ under a scheme with a larger α is on top of the curve under the scheme with a lower larger α. For a given α-value, either curves v–ρ or q–ρ basically overlap, regardless of the β-value. The critical density for generating the maximal flux increases when α takes larger value.

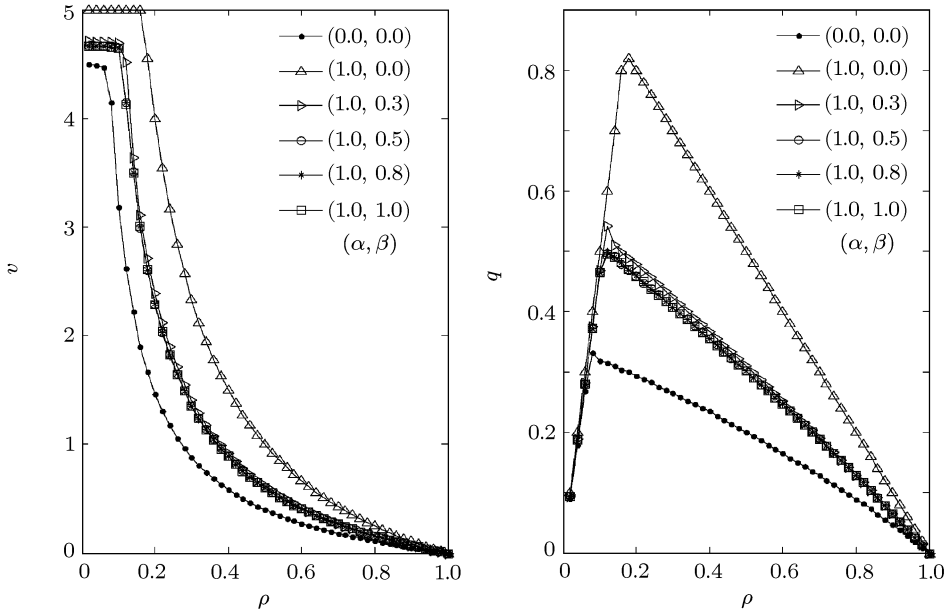


Fig. 4. Fundamental diagrams under schemes of perfect learning ($\alpha = 1$).

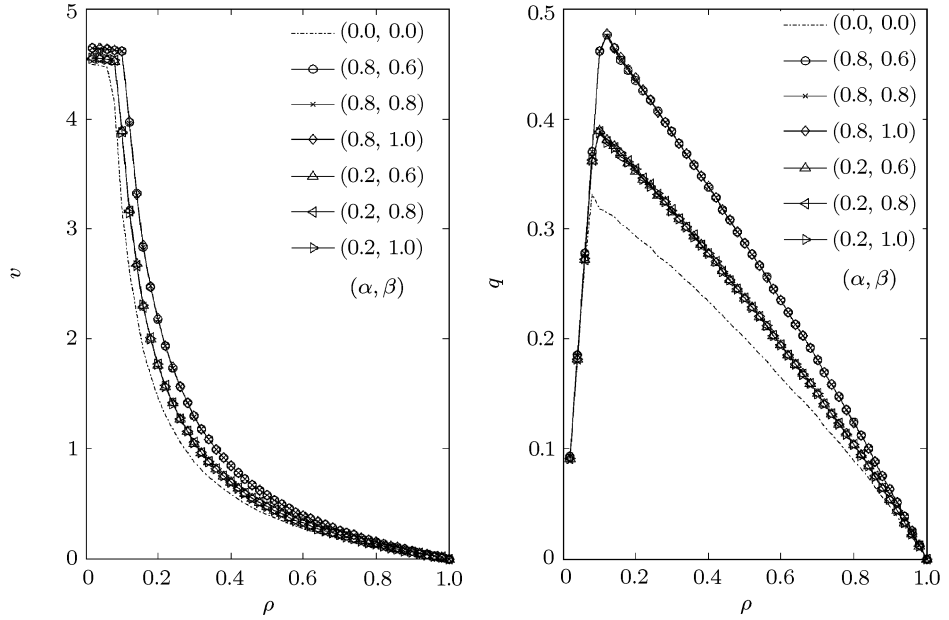


Fig. 5. Fundamental diagrams under schemes of partial learning and rapid forgetting.

Figure 6 compares the fundamental diagrams obtained under the schemes of partial learning and slow forgetting (i.e., $\beta < P_0$) in varying degrees. It can be seen that the schemes with $\beta = 0$ produce the the highest velocity and flux regardless of whether α takes 0.2 or 0.8 ($\alpha = 0$ is removed in this comparison). If $\beta \neq 0$, curves v–ρ and q–ρ are almost solely dependent upon the α-value, and a larger α-value leads to a higher velocity and flux.

In summary, the above figures show that the maximal flux given by the basic NaSch model is always less than that given by our improved model proposed in this paper. In the free-flow regions (i.e., the density is less than the critical value), the q–ρ fundamental diagrams overlap no matter what types of learning and forgetting behaviour are assumed. In the congested-flow regions, fundamental diagrams are sensitive to the combination of parameters representing learning and forgetting behaviour. Simulation results verify the analytical investigation made in this paper.

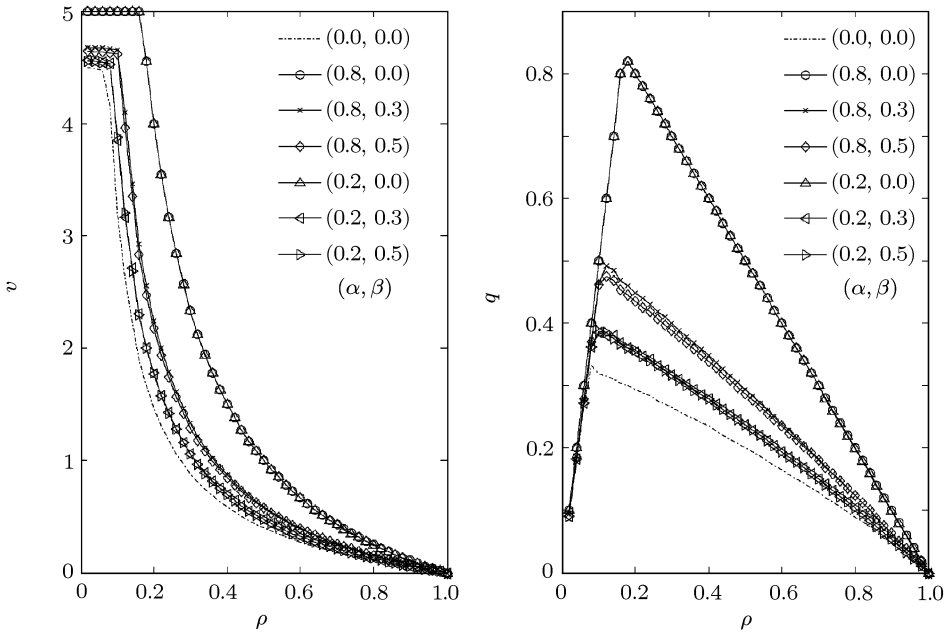


Fig. 6. Fundamental diagrams under schemes of partly learning and slowly forgetting.

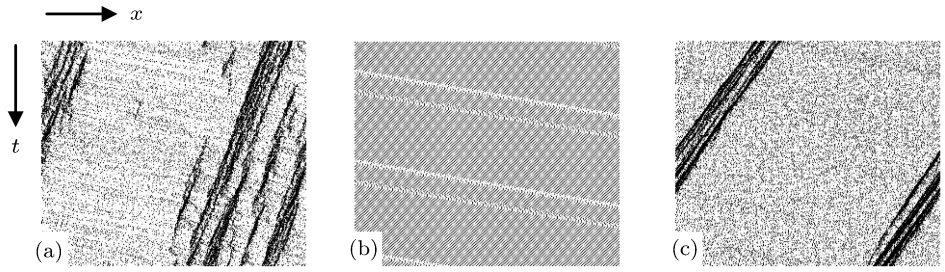


Fig. 7. Spacetime plot of car trajectories against different sets of parameters: (a) (α, β)=(0, 0); (b) (α, β)=(1, 0); (c) (α, β)=(0.8, 0.1).

We now use the spacetime plot of car trajectories to further verify our findings. Figure 7 depicts the spatio-temporal pattern of the traffic flow produced under three schemes, each corresponding to a set of model parameters ($P_0 = 0.5$ and ρ=0.16). When (α, β)=(0, 0), cars sometimes stop and form a queue, so 'stop-and-go' waves can be observed. Thus, the basic NaSch model gives the lowest flux. When (α, β)=(1, 0), cars move forward with larger gaps, so there is no traffic jam. As a result, the scheme of perfect learning generates the highest flux (see Fig. 4). When (α, β)=(0.8, 0.1), spontaneous jams occur but with much less frequency. Hence, the learning behaviour of drivers can indeed improve traffic.

5. Conclusions

In this paper, we proposed a modified CA model based on the basic NaSch model by considering the driver's intelligent behaviour (i.e., learning and forgetting behaviour). With the use of stochastic process theory, sensitivity analyses of the parameters representing learning and forgetting behaviour were extensively made. Analytical and simulation results support the following findings.

(i) Our modified model may reduce to the basic NaSch model, which depends on the value of α. If $\alpha > 0$, our model can improve the fundamental diagrams through increasing the critical density and then the maximal flux.

(ii) Changing the two parameters that reflect the learning and forgetting behaviour respectively, allows us to simulate different driving strategies and then explore new traffic features. The mean hitting time from one state to another state and the mean return time of a state are introduced to measure the characteristics of time-varying randomization behaviour. This enables

us to obtain the exact expression of flux in the free-flow region. This exact expression was reproduced by simulations. The velocity and flux in the congested-flow region were generated by simulations. Simulation results show that the self-adjustment behaviour based on reinforcement learning can increase the traffic flux a great deal in the congested-flow region.

Note that, in this paper we do not modify other rules of the CA model to investigate the phase transitions from free flow to synchronized flow to wide moving jam. However, phase transition is an important concept and worth studying as the complex spatiotemporal behaviour of real traffic is due to the transitions among the three traffic phases. Our on-going work is to further improve the CA model for reproducing various phase transitions. Certainly, this should be done by keeping the reinforcement learning of drivers in the model. In addition, we also intend to apply our model to the two-lane traffic where more interesting behaviours are open to study.

References

[1] Lighthill M J and Whitham G B 1955 *Proc. Royal Society* Ser. A **229** 281

[2] Gazis D C, Herman R and Rothery R W 1961 *Oper. Res.* **9** 545

[3] Kuznetsov A V and Avramenko A A 2009 *Math. Biosciences* **218** 142

[4] Qian Y S, Shi P J, Zeng Q, Ma C X, Lin F, Sun P and Yin X T 2009 *Chin. Phys.* B **18** 4037

[5] Qian Y S, Shi P J, Zeng Q, Ma C X, Lin F, Sun P and Wang H L 2010 *Chin. Phys.* B **19** 048201

[6] Belitsky V, Marić N and Schütz G M 2007 *J. Phys.* A **40** 11221

[7] Foulaadvand M E and Belbasi S 2007 *J. Phys.* A **40** 8289

[8] Nagel K and Schreckenberg M 1992 *J. Physique* I (France) **2** 2221

[9] Fukui M and Ishibashi Y 1996 *J. Phys. Soc. Jpn.* **65** 1868

[10] Hu S X, Gao K, Wang B H and Lu Y F 2008 *Chin. Phys.* B **17** 1863

[11] Li X G, Gao Z Y, Jia B and Jiang R 2009*Physica* A **388** 2051

[12] Barlovic R, Santen L, Schadschneider A and Schreckenberg M 1998 *Euro. Phys. J.* B **5** 793

[13] Jiang R and Wu Q S 2006 *Physica* A **364** 493

[14] Krug J and Ferrari J A 1996 *J. Phys.* A **29** 465

[15] Evans M R 1997 *J. Phys.* A **30** 5669

[16] Ktitarev D V, Chowdhury D and Wolf D E 1997 *J. Phys.* A **30** 221

[17] Nagel K, Wolf D F and Wargner P 1998 *Phys. Rev.* E **58** 1425

[18] Knospe W, Santen L and Schadschneider A 1999 *Physica* A **265** 614

[19] Treiber M and Helbing D 1999 *J. Phys.* A **32** 17

[20] Halici U 2001 *Bio. System* **63** 21

[21] Tsoularis A and Wallace J 2005 *Math. Biosciences* **195** 76