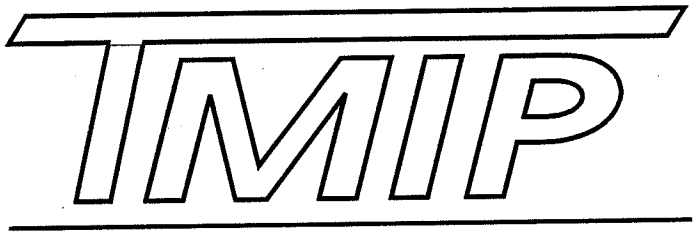
Los Alamos National Laboratory

TRANSIMS REPORT SERIES

Issues in the Representation of Traffic Using Multiresolution Cellular Automata

Christopher L. Barrett, Steven Eubank,
Kai Nagel, Steen Rasmussen,
Jason Riordan, Murray Wolinsky



Travel Model Improvement Program

Department of Transportation
Federal Highway Administration
Bureau of Transportation Statistics
Federal Transit Administration
Assistant Secretary for Policy Analysis

Environmental Protection Agency

U.S.Department of Transportation

U.S.Environmental Protection Agency

LA-UR- 95:2658

Title: ISSUES IN THE REPRESENTATION OF TRAFFIC USING MULTI-RESOLUTION CELLULAR AUTOMATA

Author(s): Christopher L. Barrett, Steven Eubank, Kai Nagel, Steen Rasmussen, Jason Riordan, Murray Wolinsky

Submitted to: INFORMAL Distribution

Los Alamos
NATIONAL LABORATORY

Form No. 836 R5
ST 2629 10/91

Issues in the Representation of Traffic Using Multi-Resolution Cellular Automata

Christopher L. Barrett, Steven Eubank, Kai Nagel,
Steen Rasmussen, Jason Riordan, and Murray Wolinsky

TSA Division, Los Alamos National Laboratory
Los Alamos, New Mexico 87545

1.0 Introduction

A cellular automata (CA) approach to traffic simulation is potentially useful in order to achieve a very high computational rate in mircoscopic simulation (e.g. on the level of the individual vehicles), and to facilitate distributed computing.

There are many obvious reasons for this, for example

- the rule set that describes the update of each vehicle is very small;
- different pieces of roadway are represented in identical or nearly-identical ways;
- the update schedule, being completely parallel, is extremely simple.

And there is one additional, more subtle reason for the high speed of the CA: much of the behavioral model of the vehicles is computed implicitly. That is, behavioral properties are arranges to occur during an update, but are not computed explicitly and do not exist in the code. Driver reaction times, for example, are the times that vehicles take from the occurrence of an event until the event is responded to. These times are represented in the CA by arranging an interaction among the spatial scaling of the CA, the acceleration logic, and the update scheduling method; they do not appear in any explicit way in the driving logic rules. We will demonstrate the implicit incorporation of these terms later in this paper.

The implication of this is that if we want to translate the basic CA approach to a more general setting, we first have to know what it involves implicitly, in addition to the explicit rule sets, scheduling algorithms, etc... We may, for instance, want to create a new CA that has a different lattice structure but produces the same dynamics as the original CA. But in writing down the update rules for the new CA, we cannot just translate the old rules to take account of the new lattice structure. Instead, we have to consider all that is implemented explicitly and implicitly in the old CA, and implement it either explicitly or implicitly in the new CA.

In other words, the implicit characteristics of a CA approach make it necessary to "retune" the simulation if anything is changed. The beneficial aspect of using implicit computation, however, is that it minimizes the computation and thereby increases speed.

If we proceed to use a multi-resolution cellular automata approach to representing traffic in simulation, we must consider this general issue of implicit characteristics, which encapsulates a number of particular issues like driver reaction time delays. This paper is intended to motivate these issues, and introduce some approaches to creating multi-resolution cellular automata with this issue in mind.

One natural outgrowth of this issue is what to do about connecting segments of different resolutions. We investigate this issue here also, and demonstrate some of the difficulties involved. It will turn out that the most important aspect of this issue is to connect segments of different resolutions at intersections only, since these connections introduce areas of low throughput, which are present in intersections anyway.

The rest of this paper is organized as follows:

In Section 2 we examine the Nagel-Schreckenberg CA traffic rules and point out some of the emergent properties that these rules are capable of producing. Then we show how one can calibrate the CA representation with reality, using experimental data from traffic systems, and discuss the effectiveness of these calibration arguments in urban traffic regimes. We are then in a position to answer the question: What exactly is the purpose of the CA with regard to the types of real-world phenomena that it is capable of reproducing? We close this section by elucidating some features of the CA computation that are done implicitly.

In Section 3, we address how the resolution of the CA can be changed, pointing out two distinct approaches that one could take towards this goal. The first approach, which we call Refinement, is capable of producing dynamics that reduces to the original dynamics under certain parameter choices. The second approach, which we call Subdivision, involves many other intricate issues and raises unique questions of its own. We also discuss some of the issues raised by having simulations of different resolutions in the same road network.

In Section 4, we invoke universality and minimalism arguments to demonstrate that the CA should be thought of as a model of a driving model. We present here a stochastic differential equation approach that serves, in part, to illustrate this point, and also to show the difficulties involved in moving to smaller scales.

Finally, in Section 5, we summarize our conclusions in a practical manner.

2.0 Overview of the fixed-resolution CA representation

2.1 The CA is defined by simple, local rules

The basic computational model (Nagel and Schreckenberg, 1992) is defined on a one-dimensional array of sites. This could, for example, be an edge in a road network. Each site is either occupied by a single vehicle, or empty.

Each vehicle has an integer velocity v (in units of lattice spacings per time step). This velocity can take values between 0 and v_{max}, inclusive.

The position on the grid (in units of lattice spacings) of the front bumper of a vehicle is denoted by x, and the number of empty sites in front of a vehicle (also in units of lattice spacings) by gap. At any time t we have

$$gap(t) = (x_{ahead}(t) - 1) - x(t) ,$$

where x_{ahead} denotes the front-bumper position of the vehicle immediately ahead, and the subtracted 1 accounts for the space this vehicle occupies.

For any configuration of the system, one update consists of the following four consecutive steps, which are performed in parallel for all vehicles:

- ***Acceleration:*** If the velocity v is small compared to gap, the speed is increased by one, up to a maximum speed v_{max}.
- ***Interaction:*** If the next vehicle is too close (if the velocity is large compared to the gap), the speed is reduced to gap.
- ***Randomization:*** With probability p, the velocity of each vehicle is decreased by one, down to a minimum speed of zero.[1]
- ***Motion:*** Each vehicle is advanced v sites.

Technically, this amounts to the following:

FOR all vehicles DO IN PARALLEL
  $gap(t) \leftarrow (x_{ahead}(t) - 1) - x(t)$
  IF ($v(t) \leq gap(t) - 1$) THEN *; Acceleration*
    $v(t+1) \leftarrow min[v(t) + 1, v_{max}]$
  ELSE IF ($v(t) \geq gap(t) + 1$) THEN *; Interaction*

1. This randomness actually condenses four different behavioral patterns into one computational rule: fluctuations at maximum speed (p_{free}), retarded acceleration (p_{accel}), over-reactions at braking (p_{brake}), and fluctuations in car following (p_{follow}). These can be distinguished in the CA for greater realism, without a serious slowdown in computational speed.

```
        v(t+1) ← gap(t)
    ENDIF
    WITH PROBABILITY p DO                    ; Randomization
        v(t+1) ← max[0, v(t+1) - 1]
    x(t+1) ← x(t) + v(t+1)                   ; Motion
ENDFOR
```

We use $\leftarrow$ to denote assignment. Two different time levels t and $t+1$ are used to indicate that only old information from time t is used for each individual car's update.

Note that, because of integer arithmetic, the expression $v(t) \leq gap(t) - 1$ from the first IF condition is equivalent to $v(t) < gap(t)$. Similarly, the expression $v(t) \geq gap(t) + 1$ from the ELSE IF condition is equivalent to $v(t) > gap(t)$.

Since this model operates on discrete space and discrete time, has only a small number of discrete states per site, and has a completely local and parallel update, this is formally a cellular automaton.

2.2 Nontrivial and realistic behavior emerges from these rules

Already this simple model shows nontrivial and realistic *emergent* behavior, including a realistic fundamental diagram, and jam behavior that has the same qualitative features as that found in real traffic systems. More specifically:

- Space-time plots show the characteristic start-stop waves seen in real traffic (Nagel and Schreckenberg, 1992; Nagel, 1994b).
- Short-time averages, taken from simulation in the same way that they are taken in reality, result in a data cloud pattern typical for real-world fundamental diagrams (Nagel and Schreckenberg, 1992; Nagel, 1994b).
- In agreement with others (Bando et al, 1994; Bando et al, 1995; Kerner and Konhäuser, 1993; Kerner and Konhäuser, 1994; Montroll, 1962), traffic flow above a certain density is no longer completely laminar, but has jammed portions interdispersed in laminar parts (Nagel, 1994a; Nagel, 1994b; Nagel and Paczuski, 1995; Nagel, 1995). It is these jammed portions which are recognized as start-stop waves by drivers.
- It is possible to quantitatively characterize the size distribution of traffic jams as a function of average overall density. It turns out that, exactly at average maximum flow, the expectation value for the life-time of a traffic jam diverges, i.e. becomes infinite (Nagel and Paczuski, 1995).
- In addition, there are large fluctuations of average travel time at maximum flow (Nagel and Rasmussen, 1994), which have been confirmed using a much more realistic simulation setup (Wolinsky et al., 1995)..

2.3 Calibrating the CA to reality

With the rules above, the variables of the CA (i.e. positions, velocities) are all in units of lattice spacings. It is relatively easy to determine what this unit must correspond to in reality.

2.3.1 Jam density

The CA represents a lane as an array of boxes, each of equal length L, where L is the average space a vehicle occupies in a jam. If we assume that a vehicle's front bumper is at one edge of the box, then the vehicle extends a length l into the box, and is followed by some empty space to the next vehicle, gap_{jam}, and $L = l + gap_{jam}$.

As a consequence, if one knows the average vehicle density at a complete standstill, ρ_{jam} [1], then the CA box size L is given by

$$L = \frac{1}{\rho_{jam}}.$$

Note that this quantity is thus given by the real world, and *cannot be adjusted at convenience.*

2.3.2 Maximum speed

There is one constant in the above CA rules that must be determined-- the maximum vehicle speed v_{max}. It should be noted that this value does not represent the maximal mechanical speed of each car, but rather a maximum speed as attainable in reality (i.e. as restricted by speed limits or by preferences of the driver). We can determine this constant by calibrating to an important feature of the flow-density fundamental diagram; namely, the density of maximum flow.

The parameter v_{max} roughly determines the density at which maximum flow occurs, and so we can use this density to set v_{max}. As a rule of thumb, *increasing the maximum speed value shifts the density of maximum flow to lower densities.*

1. This is also denoted ρ_j, k_j, or k_{jam} in the literature.

The following table gives some approximate values for the above rules with $p = 0.5$ (p is the random deceleration probability).

TABLE 1. Approximate densities at which maximum flow occurs under varying values of v_{max}, with $p = 0.5$.

v_{max}	maximum flow occurs at density
1	$\rho = (0.5 \cdot \rho_{jam})$
3	$\rho \approx (0.2 \cdot \rho_{jam})$
5	$\rho \approx (0.08 \cdot \rho_{jam})$
∞	$\rho = 0$ (!)

$v_{max} = 5$ is approximately realistic for traffic which consists entirely of fast cars (i.e. freeway traffic without trucks).

But this is not the entire picture-- in a mixture of cars of varying v_{max} at maximum flow, the slow cars dominate the dynamics (in both reality and simulation). This means that the v_{max} *of the slow cars* determines the density of maximum flow. If only 10% of the care are "slow", with $v_{max} = 3$, the maximum flow occurs at a much higher density than in uniformly "fast" traffic (Rickert, 1994). This is a realistic result.

2.3.3 Time step

Nowhere in this kind of simulation will one find a $v \cdot \Delta t$ calculation-- there are no explicit units other than number of cells. That is, time is entirely implicit. The implicit time scale is derived from some knowledge of the traffic being simulated (e.g. length of cars, or achievable maximum speed), as follows.

If L is the real length that corresponds to the CA box size, since $v_{max} - p$ is the average number of boxes a car goes at maximum velocity per update[1], then $L \cdot (v_{max} - p)$ is the real distance a vehicle goes per update, and therefore the time step

1. Imagine a vehicle driving at v_{max} with no need to slow down due to the *Interaction* rule. This is the case if the vehicle in front is far enough ahead. Since the vehicle is already at maximum velocity, it is not subject to the *Acceleration* rule; it is only subject to the *Randomization* step. As a result, the velocity after the update is v_{max} with probability $1 - p$, and $v_{max} - 1$ with probability p.

   If the vehicle's velocity is $v_{max} - 1$ after the update, however, it immediately accelerates back to v_{max} in the *Acceleration* step of the next update. It is again subject to the *Randomization* step, and so the velocity after the second update (and all updates thereafter) is again v_{max} with probability $1 - p$, and $v_{max} - 1$ with probability p.

$$\tau = \frac{L \cdot (v_{max} - p)}{v_{max}^{real}},$$

where v_{max}^{real} is the reality maximum velocity, e.g. given in km/h.

For a homogenous system of fast cars, an appropriate value to use for v_{max}^{real} is 120 km/h, or 75 mph. Using the values $L = 7.5m$, $v_{max} = 5$ and $p = 0.5$, as above, we get a time step of 1.0125 seconds, which we approximate as 1 second for the rest of this paper.

This inherent 1 second scale allows implicit representation of driver/vehicle reaction time delays and minimum following times. These will be discussed below.

It is instructive to look at the units of the time step τ. L is in units of meters. v_{max} and p are CA variables, and are in units of lattice spacings per update, or "cells" per update. v_{max}^{real} is in units of meters per second. The result is that τ has units of cell seconds per update.

2.3.4 Calibration and urban networks

In any network of roads and intersections, the network flow is bottlenecked by areas of low maximum flow interspersed in areas of higher maximum flow.

For intersections, however, the flow in any given direction is practically always lower than that of the adjacent segments. This is because of red light phases, for example, or traffic building up behind a vehicle slowing down to turn (Axhausen, personal communication).

In an urban network, then, the network flow will be determined by the intersections. To obtain the correct maximum flow over the network, *a correct calibration of the intersection fundamental diagrams is much more important than a correct calibration for the segments.*

One way to partially overcome this problem is to make intersections 2-level. That is, to remove all impediments in the "straight" directions. A bottleneck remains, however, for the "turning" directions (Nagatani, 1993). Another option is to add enough lanes for each direction through the intersection.

2.4 Minimalism and calibration

The CA approach described above is intended to be a minimal model with respect to an accurate fundamental diagram and realistic jam dynamics. That is, it is designed to be the simplest possible model that exhibits these emergent properties. We can imagine that other models, like an event-driven simulation with complex objects, also exhibit these properties, but with a much greater strain on computational resources.

Considering the calibration arguments above, however, one finds that the CA must be restricted exactly to the computed calibrations (e.g. L, τ) if it is to remain a model of real-world traffic. Hence, it is not possible to use the CA "as is" to resolve a simulation at a finer length-scale detail than the grid spacing L, or at a finer time-scale detail than the time step τ.

In fact, above we have calibrated the CA using a macroscopic property: namely, the fundamental diagram. We did not calibrate to any sort of micro-statistical property (i.e. any property at the length scale of the lattice spacing, or at the time scale of a time step). For this reason, we could expect the micro-statistics of the simulated system to be significantly different compared to the micro-statistics of a real traffic system, while the macroscopic properties exhibit a close correspondence.

Another factor that has a similar effect is the sheer discontinuity of the CA rules. All actions occur at order L in length and order τ in time. Vehicles in the real world, however, certainly trace out continuous paths in space-time. To see somewhat continuous paths in simulation, then, we have to look at scales of higher orders of magnitude.

As a consequence of all this, it is clear that *the CA approach is incapable of answering questions about small-scale dynamics*. The trade-off, however, is a very fast, optimized simulation, perfectly capable of reproducing the large-scale dynamics of interest.

One may, however, need to discern the small-scale dynamics in simulation, in which case there are two options. The first is to create a full-blown event-driven simulation with high-fidelity objects, regardless of the computer resources required. The second approach is to create a minimal model that resolves detail at all length-scales equal to and greater than the desired level of resolution-- to create, for instance, some generalization of the CA that can be calibrated to a smaller length and/or time scale so that finer detail can be resolved.

The issues raised by the latter approach are the topic of the rest of this paper.

2.5 Some things aren't so obvious in the CA

If an attempt is made to build a higher resolution model, then it immediately becomes apparent that the CA contains some implicit features which are not obvious. These include

- driver reaction times, which are the times a driver/vehicle combination requires to respond to a situation;
- following times, or the preferred times for a vehicle to follow behind another vehicle, which can easily be converted into preferred following distances;
- jam spacings, or the amount of space between vehicles in standstill jam situations.

2.5.1 Time delays in the update scheduling

The rule in the CA for car following and for braking essentially is

$$v(t+1) \leftarrow gap(t),$$

which is the *Interaction* rule with the assumption that the IF condition is fulfilled.

Translating from CA units to real units, using the calibration above, gives

$$v^{real}\left(t^{real} + 1s\right) \leftarrow \frac{gap^{real}\left(t^{real}\right)}{1s},$$

In words this means that a driver sees a certain distance to a car ahead at time t^{real}, and he tries to adjust the velocity such that at time $t^{real} + 1s$ he has reached velocity $\frac{gap^{real}\left(t^{real}\right)}{1s}$. In a more general way, this can be written as

$$v^{real}\left(t^{real} + T_d\right) \leftarrow \frac{gap^{real}\left(t^{real}\right)}{T_f},$$

where T_d is the delay time the driver/vehicle combination needs to observe the situation, determine the action, and accomplish the action. Note that the environment at time t^{real} produces actions only after a time $t^{real} + T_d$. T_f is a following time. If a vehicle drives at constant speed according to the above rule, then at time $t^{real} + T_f$ its front bumper is where the rear bumper of the car ahead was at time t^{real}.

Assuming identical braking characteristics, it is fairly straightforward to see that for accident free driving, $T_f > T_d$. In reality, T_f and T_d are, at least to the order of magnitude, comparable times.

2.5.2 Implicit following distances in the cell definition

In this section, we will consider all variables to be in real units, rather than CA units. Thus, $gap(t)$ and $x(t)$ are in units of meters, rather than lattice spacings.

In a standstill jam, there is still some distance between a vehicle's front bumper and the rear bumper of the vehicle ahead. Yet, the CA sets $gap = 0$ in this case and assumes that there is some additional separating space, gap_{jam}, since the box is longer than the car.

Along with the vehicle length, this can be made explicit by decomposing the front-bumper-to-front-bumper distance, $x_{ahead}(t) - x(t) \equiv \Delta x(t)$. We decompose this into gap_{jam}, the standstill-jam value (i.e. the minimum separation between vehicles); the variable remaining space, represented by $gap_{additional}(t)$; and the length of the vehicle ahead, l_{ahead}. Then

$$\Delta x(t) = gap_{additional}(t) + gap_{jam} + l_{ahead} .$$

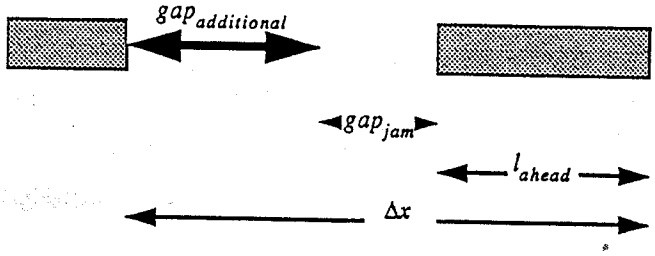


FIGURE 1. Decomposition of the front-bumper-to-front-bumper spacing

We note that the $gap(t)$ of the CA rules corresponds to $gap_{additional}(t)$ (but in the CA units of lattice spacings). Then we have, in a jam situation, $gap = 0$. Note that $gap_{additional}(t)$ denotes the distance a vehicle has *available* in front of it to move into, whereas the summed value $gap_{additional}(t) + gap_{jam}$ denotes the *actual* space ahead of a vehicle.

We can use these two expressions to create a gap computation step that makes these values explicit. The new step is, quite simply,

$$gap(t) \leftarrow \Delta x(t) - (gap_{jam} + l_{ahead}) .$$

2.5.3 The CA rules themselves are not an entire description of the dynamics

The considerations of this section demonstrate that we cannot consider the CA rules themselves to be a full description of the simulation dynamics. Rather, we must think of the CA as a package consisting of the update rules, a scheduler (here, a parallel scheduler), and a (here discrete) representation of states. Then we can make explicit the implicit characteristics that result from the interplay of these three components.

For instance, the two-time-rule of (2.5.1), with T_d and T_f, is implemented in the original CA by an implicit interplay between the scheduler and rule set. T_d is implemented by a parallel scheduler, which insures that the next time the state is accessed after an update is exactly one time step later. In other words, it is the scheduler that insures that by changing the value of v at a site, the update is effectively setting $v(t+T_d)$. T_f is implemented by equating directly v and gap, which actually have different units, so that the units are implicitly rectified by a conversion factor. Since v is in units of lattice spacing per time step, and gap is in units of lattice spacing, this conversion factor is one time step, which is again given by the calibration above.

If a different scheduler is used (e.g. a random-order sequential update), the dynamics is completely different (Nagel, thesis) since the *Interaction* rule is no longer implemented as in (2.5.1).

This is an example of the general issue of differences between continuous mathematics and discrete computation. We often think of a discretization as a model of a continuous

system, and try to minimize any effect of the discretization (i.e. by minimizing "error terms" in a numerical integration of a differential equation). But we are in a different scenario here, since we are beginning with a discrete model and attempting to find a continuous model that reduces to the discrete model under particular parameter choices. And in this endeavor, we have to explicitly model the effects of the discretization. An analogy would be that we have to determine the "error terms" of a numerical integration, and place them explicitly into a continuum representation.

3.0 Changing the resolution of the CA

3.1 Why change it?

As pointed out in the last section, we may need to rectify a simulation to a level of detail too small for the original CA, and thus may need to develop a discrete model capable of producing realistic dynamics at smaller length and time scales. In particular, we may look to a more detailed CA

- to rectify the dynamics at the level of car-following, so that one can observe a realistic time-series in the objects states, perhaps to detect problem situations involving two cars (e.g. situations that could more easily cause an accident), or to use this information for another modeling purpose (e.g. emissions and air quality modeling).
- to model more complex vehicles whose characteristics are not compatible with the CA granularity; for example, vehicle lengths which are not a multiple of the calibrated L; or maximal velocities that are not a multiple of $\frac{L}{\tau}$.
- to model more refined spatial structures whose characteristics are not compatible with the CA; for example, roadway segments that are not simple multiples of L.

Also, there is no a priori reason why a more coarse approach couldn't be found that rectifies the dynamics at larger length and time scales. The obvious reason to do this would be the increase in computational speed for areas where the small-scale dynamics can be ignored. What makes this approach more difficult is that we must allow for the possibility that there is more than one car in a site. For this reason, this approach has intricacies and problems of its own, and we do not address it here.

3.2 Approaches to the refining the spatial resolution of the CA

There are basically two ways to refine the spatial resolution of the existing CA:

- ***Refinement:*** The new grid has a site length that is an integer quotient of the length of the sites on the original grid. That is,

$$L_{new} = \frac{1}{n} \cdot L,$$

where n is an integer.

- *Subdivision:* The new grid has a site length that is less than that of the original grid, though not necessarily an integer quotient.

We will find that each of these approaches has its own intricacies and raises its own issues.

3.3 Refining the spatial resolution with no change in behavior

3.3.1 Rewriting the CA rules for a refinement of the original grid

It is straightforward to rewrite the original traffic CA for a refinement of the original grid. Consider the rules of (2.1), taking $v_{max} = 5$.

Now consider two vehicles on the CA grid ($L = 7.5m$, $\tau = 1s$) at two consecutive time steps. Assume that the vehicle on the left (shown in black) has an initial velocity of 5 in CA units, or $5 \cdot \frac{L}{\tau}$ in reality units. Also assume that the vehicle on the right (shown in gray) has an initial velocity of zero.

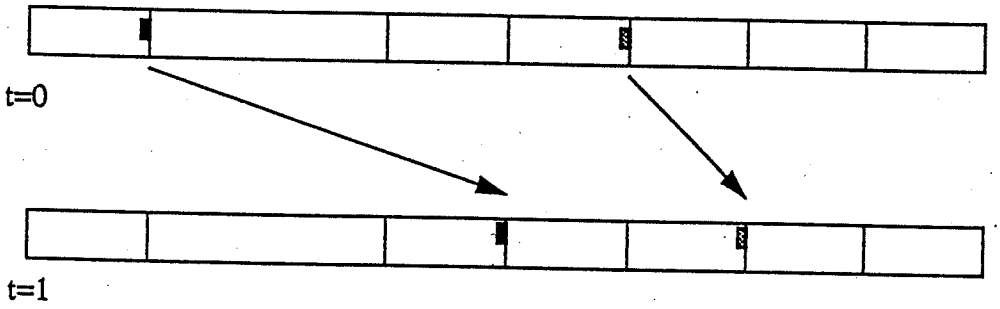


FIGURE 2. Two vehicles at two successive moments on the original 7.5 meter grid

Look at the rules for the black car on the left (and temporarily ignore the vehicle on the right). The *gap* is computed to be 3. The *Acceleration* condition is not met. The *Interaction* condition, however, is met, and $v(t+1)$ takes the value 3 (from *gap*). We'll assume no effect from the *Randomization* step.

The rules for this case give a final velocity of 3 and a motion of 3 cells.

We want to examine the motion of the same vehicles on a finer grid. The figure below shows the original grid but divided into pieces of one-fifth the length. It is obvious that the situation above can be portrayed on the finer grid without difficulty.

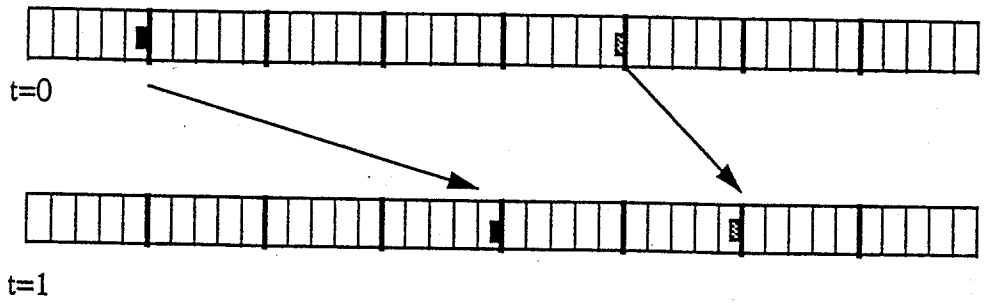


FIGURE 3. The same two vehicles at the same times on a 1.5 meter refinement of the original grid

The new rules are obtained from the old ones by multiplying all length and velocity constants by 5:

FOR all vehicles DO IN PARALLEL

$gap(t) \leftarrow (x_{ahead}(t) - 5) - x(t)$

IF ($v(t) \leq gap(t) - 5$) THEN ; *Acceleration*

$v(t+1) \leftarrow max\left[v(t) + 5, 5 \cdot (v_{max})^{old}\right]$

ELSE IF ($v(t) \geq gap(t) + 5$) THEN ; *Interaction*

$v(t+1) \leftarrow gap(t)$

ENDIF

WITH PROBABILITY p DO ; *Randomization*

$v(t+1) \leftarrow max[0, v(t+1) - 5]$

$x(t+1) \leftarrow x(t) + v(t+1)$; *Motion*

ENDFOR

Recognize that the gap is measured in the smaller units, so that the gap for the vehicle on the left, originally 3, is now 15. Similarly, velocity is measured in cells per time step, so that the initial velocity of the vehicle on the left, originally 5, is now 25. The final velocity, by the rules above, is 15 and there is a motion of 15 cells.

It is evident that this procedure can be carried out for any integral refinement of the original grid and that, if corresponding initial conditions are employed, then identical results will be obtained. It should also be obvious that nothing is accomplished by this refinement except the (mis-)use of additional storage and time to compute exactly the same dynamics.

However, the finer grid opens up new representational possibilities and suggests extensions of the original model. One such extension will be considered shortly. Before doing so it is convenient to look at how to connect grids with differing spacing.

3.3.2 Coupling between mixed-resolution grids

We can imagine very simple grid couplers which simply transfer vehicles from an original grid to a finer version and from the finer version back to the original grid.

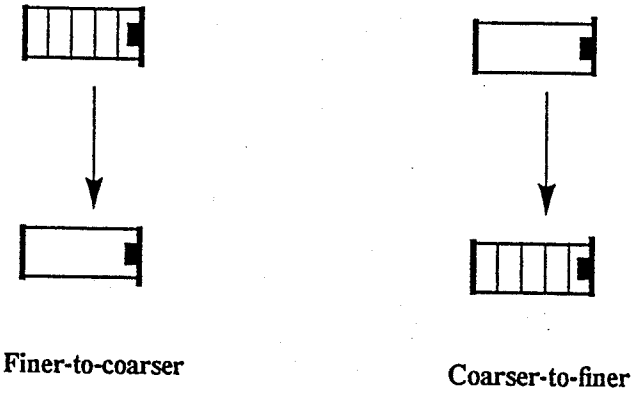


FIGURE 4. The naive grid couplers

Conceptually (though perhaps not practically) these couplers can be used as depicted below. Shown is a mixed-resolution grid consisting of a section of 15 cells of fine resolution, followed by 3 cells of coarse resolution and then 10 additional cells of fine resolution. A vehicle, moving with fixed speed of 2 cells per time step (coarse resolution) is depicted.

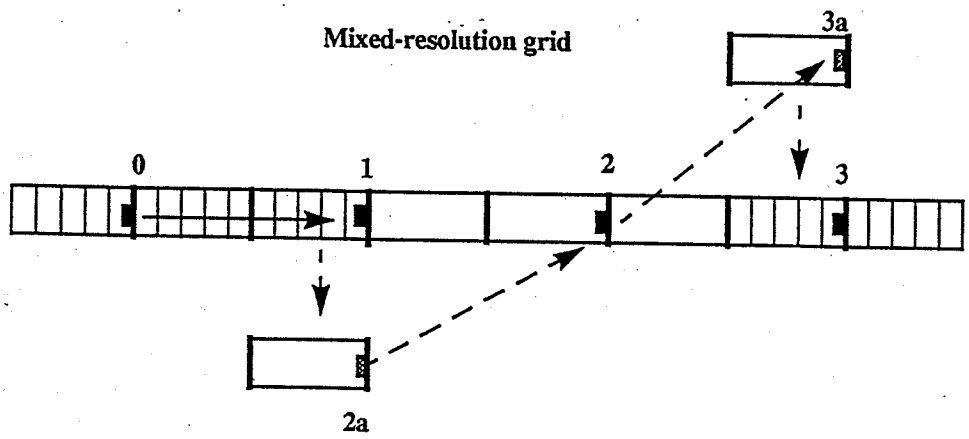


FIGURE 5. Coupling grids of different resolutions

From time 0 to time 1 the vehicle is on the fine grid. Between time 1 and time 2 it must move from the fine grid to the coarse grid. One mechanism for doing so places a coupler before the coarse grid segment and transfers the vehicle instantaneously to the coarse grid, then moves the vehicle on the coarse grid to the position marked 2. Between times 2 and 3 the vehicle must move off the coarse grid and back to the fine grid. One can regard this as motion on the coarse grid, followed by a coupling back to the fine grid.

Although the details of this transfer process require some care, it is clear that there are no conceptual problems involved. The transfers can be done smoothly with no change to vehicle dynamics. However, all this nice behavior breaks down immediately when we try to enhance the model.

3.4 A new dynamics is introduced by accessing the new states

By accessing the new states (i.e. new positions, velocities) that aren't present in the 7.5 meter representation, the dynamics of the 1.5 meter grid exhibits some different characteristics.

Consider the 1.5 meter grid, but with the initial condition (as above) where all cars are aligned at position multiples of 7.5 meters. With the rules as they stand now, we have not gained anything in representation in this case, since the dynamics is exactly equivalent to that of the 7.5 meter grid.

There are many ways to access the new states. Effectively, the new granularity makes possible a finer representation of all length variables in the simulation, and hence all velocity and acceleration variables also. One obvious way to take advantage of this new granularity, then, is to set up the initial conditions at the new, finer scale.

An instructive example, however, would be to take a minimal advantage of the new state representation, and look at the manifest effects of such a minimal change. This is what is done in the next section, where we introduce a distribution of vehicle lengths (where the lengths are in multiples of 1.5 meters) to the 1.5 meter refined grid.

3.4.1 How only making different length vehicles changes things

A simple approach to allowing varying vehicle lengths begins by parameterizing the gap computation equation. The number 5 in the existing assignment,

$$gap\,(t) \leftarrow (x_{ahead}\,(t) - 5) - x\,(t)\,,$$

is made variable as in (2.5.2), and hence we have the new gap computation equation

$$gap\,(t) \leftarrow [x_{ahead}\,(t) - (gap_{jam} + l_{ahead})] - x\,(t)\,.$$

The 1.5 meter CA, as it stands now, sets

$$gap_{jam} + l_{ahead} = 5\,,$$

and does not distinguish differences in two values, as long as the sum is 5.

If we change gap_{jam} and l_{ahead} so that their sum remains equal to 5, we recover the exact same dynamics. But this is not the natural thing to do. Instead, one is tempted to do

something like set $gap_{jam} = 1$ regardless of the vehicle and then allow l_{ahead} to vary from vehicle to vehicle.

Do we need to change any of the rules in our 1.5 meter rule set to account for these different vehicle lengths? More specifically, since we changed the hard-wired vehicle length of 5 lattice spacings, do we need to modify any of the other 5's that are floating around in our 1.5 meter rules? This turns out to be a surprisingly intricate issue and the answer depends on the precise form we use to write the rules as well as on what we decide is the "intent" of the rules.

We must certainly change the Interaction rule, however, if we are to keep collisionless dynamics. This rule immediately leads to velocities which are no longer multiples of 5 units. Of itself, that is not problematic; rather, it is commendable, for it is what allows us to access states that were previously unaccessible. However, the inequality in this rule no longer catches all possible collisions. The simplest change is to rewrite it as

ELSE IF ($v(t) \geq gap(t) + 1$) THEN ; *Interaction*
    $v(t+1) \leftarrow gap(t)$
ENDIF.

Now if the sum of gap_{jam} and l_{ahead} is set equal to 5 and if the same initial conditions are kept, then the new model agrees with the previous one. Otherwise it no longer necessarily does.

It is clear that if vehicles remain 5 units long, the jam densities for the new rule set are less than those original, since at jams vehicles will be spaced gap_{jam} units apart rather than being allowed to go bumper-to-bumper. Of course, the rectification of this problem is in the fact that the effect of gap_{jam} is modeled *implicitly* in the original CA, while we make it *explicit* in the 1.5 meter refinement.

3.4.2 The new coupling between mixed-resolution grids

It is also clear that the naive method for transferring vehicles from the fine grid to the initial grid breaks down. The coupler is no longer guaranteed that a vehicle arrives in precisely the correct cell to be transferred. By itself this may appear to be only a nuisance. But two different vehicles might arrive on the fine side of the coupler during the same time step and both need to be transferred simultaneously. Since vehicles can not be superimposed, this is disastrous.

The breakdown of the naive coupling mechanism is intimately tied to the fact that the dynamics with varying vehicle lengths is no longer the same as the original dynamics.

3.5 Effects induced by changes of the model in simulated traffic

3.5.1 Introduction

Assume two different models, A and B. Since maximum throughput is an emergent quantity of the models (i.e. a quantity which can only be adjusted indirectly, as opposed to, e.g. maximum speed), models A and B will have different maximum throughput, if even by a tiny amount. As a result, there can be a calculation induced traffic jam at the position where the two models connect.

Such effects are well known, e.g. from nested-grid meteorological models. Yet, they look unrealistic; for that reason alone one should connect different resolutions only at positions where the formation of jams would make sense anyway, e.g. at intersections, or at a decrease of the speed limit.

3.5.2 Coarse to fine grain spatial resolution

In this case, additional problems arise. For instance, after a change to higher resolution, the vehicles do not automatically explore every state available to them. Rather, they begin on the states that were present in the coarse resolution.

As an example, consider a change from a grid spacing of 7.5 meters to one of 1.5 meters. As the vehicles enter the 1.5 meter grid, they first occupy the positions that were present in the old grid; namely, those of multiples of 7.5 meters. It will take some time before the system accesses the other newly available states, and we call this time the *adaption time.*

We can use this time to form an equivalent length, the *adaption length*, which is the average length a vehicle must travel before it is adapted to the finer grid.

The point of all this is that there is some finite length and time that the vehicle must travel in a higher-resolution simulation to experience the effects of that resolution. If sensors are placed in a high-resolution area but before this length and time is reached, then having the high-resolution area there will have no effect. That is, *going to high resolution just before sensors will result in low resolution traffic patterns at the sensor location.*

Further trouble can arise when we have cars with different lengths in the simulation. That means that we probably have cars that are longer than one coarse box, which are "stuffed" into only one coarse box because of rounding. When, say, doubling the resolution, these vehicles can expand into 3 boxes. If you have several cars of this type in a queue, this might become a problem because in the higher resolution they occupy much more space than in the lower resolution.

3.5.3 Fine to coarse grain spatial resolution

The analog is true for fine to coarse resolution: compact cars, which do not occupy much space in the high resolution model, have to expand to full box sizes for the coarse grain spatial resolution.

3.5.4 Recommendation

For all these reasons, we recommend to connect different resolutions only at intersections. Intersections cause jams anyway; and the intersection object should also anyway have a mechanism of how to deal with vehicles of different lengths.

3.6 Subdivision

3.6.1 Versus refinement

Refining a grid, as above, splits the original grid into an integral number of smaller pieces. In doing so, the representational capability of the grid increases without loss: new things can be represented on the finer grid that couldn't be represented on the original grid and anything that could be represented previously can still be represented. However, grid refinement leads to a severely limited number of additional possible cell lengths; namely, only those that are an integer quotient of the original CA spacing.

This sections discusses the issues raised if we decide to represent traffic flows using a grid spacing that is not an integer quotient. That is, issues raised by considering a general subdivision.

3.6.2 The 3.0 meter grid problem

A cell size of 3.0 meters does not represent a refinement of the original grid, but a subdivision. As we will immediately see, the difference is not merely semantic.

Let us then take cells of 3.0 meters. We expect to have increased our representational powers by going to a smaller spacing: we believe we should be able to represent new things. And, of course, we can represent more things than before, on average. But we also lose the ability to represent certain things. In particular, we can no longer represent vehicles uniformly spaced at 7.5 meters!

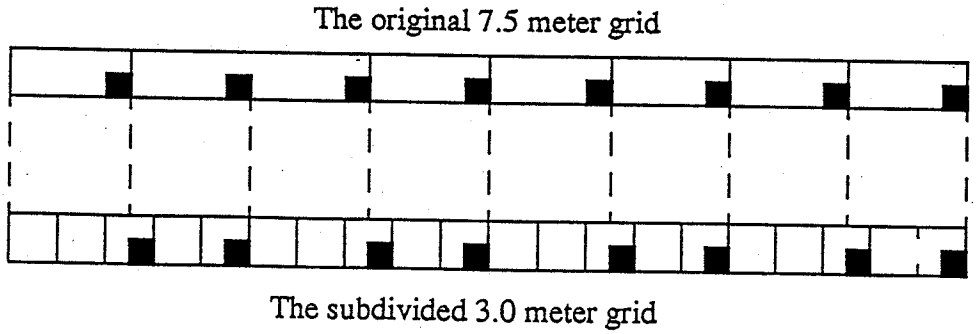


FIGURE 6. An unsuccessful attempt to place vehicles on the 3.0 meter grid using their positions on the 7.5 meter grid

On average, on the 3.0 meter grid, vehicles are spaced 7.5 meters apart, but half are 6.0 meters apart and the other half are 9.0 meters apart.

3.6.3 "Enhancing" the 3.0 meter grid fails to solve the problem

There are two general strategies for dealing with this problem. The first strategy insists that a grid with smaller spacing should have greater representational power than a grid with larger spacing; therefore the 3.0 meter grid must be capable of representing anything that the 7.5 meter grid does. This strategy leads to an "enhanced" 3.0 meter grid on which the original spacing can be represented, by adding a variable to the cell state that indicates whether the vehicle is at the left or right of the cell. It soon becomes obvious that this solution is a Pyrrhic victory. That is, this strategy only leads to a clumsy emulation of the 1.5 meter grid refinement. If one can use the 1.5 meter spacing one should do so. If there is a reason to resolve to 3.0 meters but not to 1.5 meters, then the "enhancement" strategy fails.

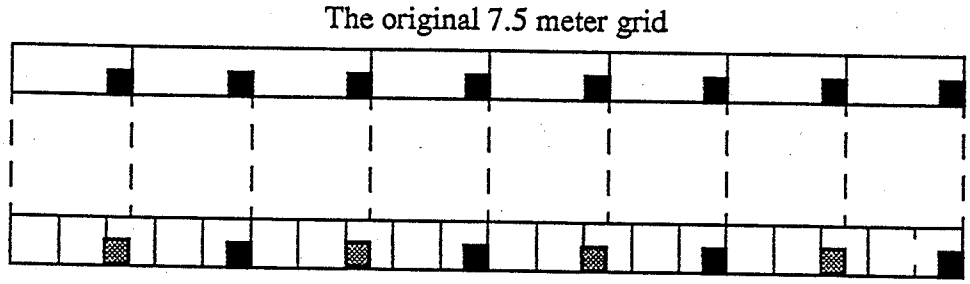


FIGURE 7. The (ultimately) unsuccessful enhancement strategy

So we must turn to the second strategy.

3.6.4 Accepting the limits of the 3.0 meter grid leads to a search for a continuum model

If we are persuaded by the argument above, then we must accept the limitations of the 3.0 meter grid, and in particular, its inability to represent things such as a row of vehicles uniformly spaced at 7.5 meters. A simple argument now leads directly to the need for a continuum representation.

Since we do not want a change in grid spacing to change the predictions of our traffic model, we would like the predictions of the 3.0 meter grid model to agree with those of the original grid whenever possible. In general we want the predictions of any subdivision of a grid to agree with those of the original grid whenever possible. But it is already clear that agreement must be found only in some statistical sense: the best we might hope for is that the underlying probability distributions of different outcomes are unchanged for grids of different spacings.

We must therefore look at the nature of the probability distributions for different outcomes as a function of grid spacing. If we know that these distributions are of a particular form, characterized by certain parameters, then changes to grid spacing must preserve the parameter settings of the original grid.

Generally, however, we do not know the underlying probability distributions; but we can specify them by their means and higher-order moments. We want the probability distribution for a grid subdivision to preserve all moments of the original distribution. But, in general, each new moment represents additional information and requires a new parameter to determine it.

So, in general, we can't match the new distribution to the old one if we only have a finite number of parameters to twiddle.

There is really only one way out: we need to know the form of the probability distribution so that we know how many parameters are required. But, there is an approximation to this situation which is more realistic. Namely, if our CA models are derivable from a continuum model, then we can determine the probability distribution in the limit and we can expect our CA's to approximate this distribution more closely as the lattice spacing is changed. The subdivisions will not, in all likelihood, give the same probability distribution as the original grid, but any differences will be understandable and controllable.

3.6.5 Subdivision presents a simpler, but not simple, arena for studying coarsening

The general problem of subdivision is difficult but not especially interesting. If we want a smaller grid, we can confine ourselves to a refinement and avoid studying subdivisions. However, the problem of *coarsening* (going to larger grids for increased performance) is interesting. Although coarsening involves many concerns beyond those we've touched upon, subdivision provides a fertile domain for studying coarsening. Specifically, since a general subdivision can be viewed as a grid refinement followed by a coarsening, then if we knew how to subdivide, we would have a good clue as how to coarsen.

4.0 The CA should be thought of as a model of a driving model

4.1 Minimalism and universality

When judging the realism of a model, one should keep in mind that the CA is intended to be a *minimal* microscopic description, with regard to jam dynamics and a fundamental diagram.

4.1.1 Universal exponents and jams

The theory of critical phenomena can be used to describe the behavior of traffic jams in the CA model near average maximum flow-- exactly where the most interesting jam behavior occurs.

Using this theory, it can be shown that the traffic jams follow certain critical exponents. Often in critical phenomena, it rungs out that critical exponents are *universal*, i.e. do not depend on the exact microscopic characteristics of the model. Since we also have a theory for the critical exponents we find, and this theory makes use of only the most basic properties of traffic flow (such as the conservation of cars and "drive slower when traffic is dense"), we conjecture that all car following models fall into this common universality class and are therefore *equivalent* in this respect.

It should be noted that such arguments are only valid in the thermodynamic limit, i.e. for infinitely large system sizes and for infinitely long time averages. Usually, finite size behavior gets more and more distorted in some way with smaller scales, and one would have to check for which scales or which questions that actually matter.

4.1.2 Only a few characteristics have an effect on the fundamental diagram

There are only a few characteristics of the CA model which have an effect on the fundamental diagram:

- As mentioned in (2.3.2), v_{max} controls the density of maximum flow. More technically, it controls the slope of the flow-density fundamental diagram in the limit $\rho \to 0$ (when the calibration of the time step is kept fixed). This is because the velocity in this limit is determined solely by the maximum velocity of the vehicles (since interactions are effectively nonexistent).
- The level of maximum flow is determined by acceleration at low speeds; higher acceleration produces a higher maximum flow. In the CA, the acceleration at low speeds in determined by the slow-down probability p, or, more specifically, the effect of this slow down probability when the vehicle is accelerating (p_{accel} in the context of the footnote of (2.1))..
- Time-to-breakdown, the time that a homogenous system above the critical density takes to break into distinct laminar and jammed regions, is determined by fluctuations at maximum speed. In the CA, this is also determined by the p (or, more precisely, p_{free}).

4.1.3 The intent of the CA is to capture these essential features

The intent of the CA approach is to capture the features that are essential in realistic jam dynamics and in producing an accurate fundamental diagram in a minimal way, in that no simpler microscopic model does the same.

In this respect, the CA can be thought of as a *model of driving models*. Other driving models may contain these features and more, and may reduce to the CA model in special cases.

In the next section we give an example of a stochastic differential equation approach that reduces to the CA under appropriate parameter values. Also, in the neighborhood of these parameter values, the stochastic differential equation below produces CA-like dynamics that are in accord with the considerations of the two sections above. In this sense, the CA models the stochastic differential equation model.

4.2 A stochastic differential equation generalization of the CA

In a certain sense, to be explicated below, the original CA model is equivalent to the system of stochastic differential equations (SDEs)

$$\begin{aligned} dx(t) &= v(t)\,dt \\ dv(t) &= a[v(t), gap(t)]\,dt + b\,dW \end{aligned}$$

where dW represents white noise, and a and b are given by

$$a[v(t), gap(t)] = \begin{cases} -p & \text{if } \left(\left(v(t) < \frac{gap(t)}{T}\right) \wedge (v(t) = v_{max})\right) \\ 1-p & \text{if } \left(\left(v(t) < \frac{gap(t)}{T}\right) \wedge (v(t) < v_{max})\right) \\ \frac{1}{T}\left(\frac{gap}{T} - v\right) - p & \text{if } \left(v(t) > \frac{gap(t)}{T}\right) \\ 0 & \text{otherwise} \end{cases}$$

$$b = \sqrt{p(1-p)}\,.$$

The discrete-time approximation to the SDEs, using an Euler discretization, is given by

$$\begin{aligned} v(t+\tau) &= v(t) + \tau a(v(t), gap(t)) + b\eta(t)\sqrt{\tau} \\ x(t+\tau) &= x(t) + \tau v(t) \end{aligned}$$

where $\eta(t)$ is any random process with mean zero and unit variance. We will find that, with $\tau = 1$, and with the appropriate choice of the process η, these equations become identical with the CA model.

The above equations for dx are intuitively plausible. The equations for dv require some explanation, which we proceed to supply. It is convenient to start by looking at the *Randomization* step of the CA model.

4.2.1 Modeling the *Randomization* step

This step has interesting implications for the continuous-time limit. It either preserves the velocity at time t, with probability $1-p$, or lowers it by one unit, with probability p (where we have temporarily neglected the boundary condition that the velocity cannot become negative).

Consider an ensemble of vehicles, all of which have the same initial velocity and with are subject only to the *Randomization* step. There are two effects of the randomization. The first effect is to add noise to the distribution: the initially sharp velocity distribution is blurred; equivalently, the velocity diffuses. The second effect is that the mean of the velocity distribution decreases by p velocity units per time step: the velocity distribution drifts downward.

If we separate out the effect on the mean (the drift term) from the noise effect (the diffusion term) we can imagine carrying out the randomization step in two parts. We first prop-

agate the velocity v at time t to the velocity $v-p$ at time $t+1$. Next we add noise; $+p$ (positive noise bringing the velocity back up to v) with probability $1-p$, and $-(1-p)$ (negative noise bringing the velocity down to $v-1$) with probability p.

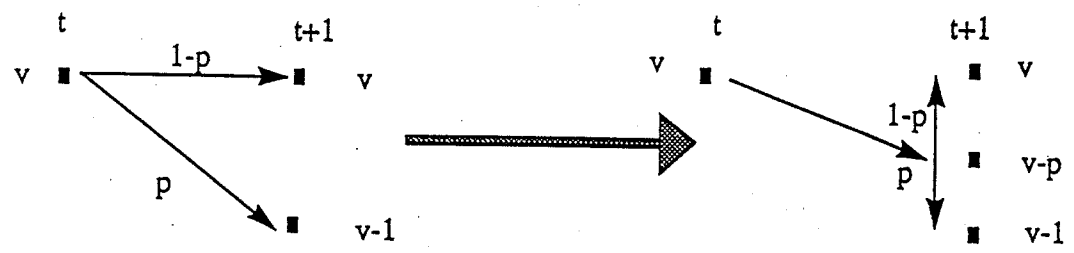


FIGURE 8. Drift and diffusion reformulation of CA *Randomization* step

Note that the added noise has two key properties: it has zero mean and variance $\sqrt{p(1-p)}$. The vanishing of this variance at $p = 0$ and $p = 1$ is a result of the fact that these two values represent pure determinism.

A noise process with this mean and variance, and with a downward drift of $-p$, is modeled directly by the stochastic differential equation

$$dv = (-p)\,dt + \sqrt{p(1-p)}\,dW.$$

The first term represents the drift term: there is a loss (on average) of p units of velocity per unit of time. The second term is the noise term. We imagine a Wiener, or Brownian motion, process, W, corresponding to a random walk with zero mean and unit increase in variance per unit time. We take an increment of that random walk, dW, to get the second term.

This model has all of the properties of the original *Randomization* step with two exceptions. First, it does not yet take into account the boundary condition at zero. Second, the velocity noise term is not restricted to amounts of p for positive noise and $-(1-p)$ for negative (downward) noise. We will return to these matters later.

We should point out that the equation above leads to a discrete-time model of the form

$$\Delta v = -p\tau + \sqrt{p(1-p)}\,\eta(t)\,\sqrt{\tau}$$

where τ is the time-step interval and $\eta(t)$ is a random number generated at time t in any manner which realizes the properties of independence (no correlation between numbers generated at different times), zero-mean, and unit variance. An example of an appropriate random number generator is

$$\eta(t) = \xi(t) = \begin{cases} \dfrac{p}{\sqrt{p(1-p)}} & \text{with probability } 1-p \\ \dfrac{-(1-p)}{\sqrt{p(1-p)}} & \text{with probability } p \end{cases}$$

This choice of random number generator fulfills the second requirement above for the velocity noise (namely, that the noise be restricted to a particular value in the positive direction, and a particular value in the negative direction). Another example of an appropriate random number generator, but one that does not fill the above requirement for the velocity noise, is one which generates $\pm\sqrt{2}$ with equal probabilities.

At this stage we have accomplished several things. We have derived the form of the noise term for both the SDE and its associated discrete approximation, and we have derived the value of $b = \sqrt{p(1-p)}$ used above.

4.2.2 Modeling the *Acceleration* step

We will examine the applicability condition and the action item of the Acceleration rule separately. It is convenient to rewrite the applicability condition, $v(t) \leq gap(t) - 1$, as

$$v(t) < gap(t) .$$

The simplest approach to translating this term to the continuum limit is to introduce a parameter T with dimension of time and the value unity and rewrite the above condition as

$$v(t)T < gap(t) ,$$

which is now dimensionally correct. It is also intuitively plausible, since it represents a driver who is attempting to maintain a following time of T behind the vehicle in front, but is travelling too slowly. In fact, it is equivalent to the following time T_f of (2.5.1).

The action item,

$$v(t+T) \leftarrow max[v(t) + 1, v_{max}] ,$$

represents an acceleration of one unit of velocity per unit of time, assuming that the vehicle is not at its maximum velocity. So our continuum representation of this step is simply

IF $(v(t)T < gap(t))$ THEN

$$a(t) \leftarrow \begin{cases} 0 & \text{if } v = v_{max} \\ 1 & \text{if } (v < v_{max}) \end{cases}$$

ELSE ...

which is represented in the expression for acceleration given previously, corrected by adding $-p$ from the *Randomization* step.

4.2.3 Modeling the *Interaction* step

The applicability condition is corrected an interpreted exactly as above, but applies to the complementary possibility, namely the vehicle is traveling too fast. The action item, $v \rightarrow gap$, is a slowing down so that the vehicle arrives at the end of the present gap T units of time later. It is easy to derive the acceleration which is required to do so; however, the equations which result do not become identical to those of the CA. There is an alternative interpretation, though, for which exact equivalence does follow.

We write the action item as

$$v(t+T) \leftarrow \frac{gap(t)}{T}$$

and interpret it as providing an updated velocity of $\frac{gap}{T}$ after an elapsed time T. The acceleration required to change an initial velocity v into the velocity $\frac{gap}{T}$ after time T is simply

$$\frac{1}{T}\left(\frac{gap}{T} - v\right),$$

so that the continuum representation of this step is simply

ELSE IF $(v(t)T > gap(t))$ THEN

$$a(t) \leftarrow \frac{1}{T}\left(\frac{gap(t)}{T} - v(t)\right)$$

ENDIF

which is again corrected in the expression above for the acceleration because of the *Randomization* step.

4.2.4 How the SDE and CA agree, and how they don't

One finds that upon substituting $\tau = 1$ in the discrete-time approximation to the SDE it become identical to the original CA. However, there are two minor quibbles with this equivalence and one major flaw.

The minor quibbles have to do with the noise. There is nothing in the formulation which requires the noise to be of the proper amplitude to maintain the velocity discretization. However, if one chooses a specific model for the random noise process $\eta(t)$, namely the $\xi(t)$ process given above, then this model, for $\tau = 1$, does so. It should be observed that this specific requirement goes beyond the SDE. A second quibble with the exact equivalence has to do with the boundary condition on the velocity at zero. It is not correctly encapsulated by the SDEs and should be imposed as an additional boundary condition.

Also, to exactly reproduce the CA with $\tau = 1$, we must also impose the initial condition that all velocities and positions start out with discrete, integer values.

However, the major flaw of the model is that it agrees exactly with the CA! This is bad because the equivalence depends crucially on selecting $\tau = 1$. For any other time step, there will not be agreement. For significantly different time steps, even the large-scale, macroscopic behavior will be significantly different! This is because the implicit control terms, some of which were outlined in (2.5), *have not been incorporated*, but are incorporated implicitly when $\tau = 1$, and are incorporated implicitly in an approximate way when $\tau \approx 1$.

5.0 Conclusions, observations, and rules of thumb

5.1 Control terms that are implicit in the original CA must be made explicit in other models

There are a number of implicit rules in the CA, which result from the interplay of the rule set, the scheduler, and the state space representation. It is another research project to elucidate these terms, and incorporate them into some of the alternative modeling approaches outlined in this paper (e.g. into a variable time step model, or into a stochastic differential equation model).

5.2 As the scale gets small, a continuum model becomes more efficient

Our current estimates say that the CA is approximately a factor of 20 faster than car following models which are also time-stepped based on a 1 second basis but which have continuous spatial resolution. Enhancing the resolution of the CA leads to slower computing speed, since the state space expands and previously implicit control terms must be made explicit, and eventually a continuous model will both become faster and more straightforward.

5.3 Going to larger scales is not addressed here

In this paper, we only address box sizes of $L = \frac{1}{\rho_{jam}}$ and smaller. Going in the other direction, i.e. box sixes larger than this means that one has to deal with multiple cars in one box. When one goes far enough in this direction (box sizes of 150 meters or more), one would use a fluid-dynamical approximation as is indeed used by several models. This is very different from what has been said in this paper, and we recommend not implementing this currently.

5.4 Connect multiple-scale segments at nodes only

As explained in the text, connecting segments with different resolutions may lead to inconsistencies. To circumvent this problem, we recommend to change resolution at intersections only. Intersections should be designed in a way that they can absorb and emit vehicles from and to models of arbitrary resolution.

5.5 Intersection buffers have similar problems of mediating between resolutions

Note that intersection objects will ultimately face the same problems of mediating between resolutions as mentioned in the text. However, instead of designing connectors between all kind of resolutions, which would be of order $O(N^2)$ connectors if we had $O(N)$ different resolutions, connecting at intersections means that we only have to design connectors between the intersection resolution and all other resolutions, which means only $O(N)$ connectors.

5.6 Intersections act as spatial buffers

We also explained in the text that the box size of the CA is given by $\frac{1}{\rho_{jam}}$, where ρ_{jam} is the density of vehicles in a jam (or fractions of this for higher resolution CAs). It is therefore given externally and cannot be used to adjust if link lengths are not multiples of the box size.

We recommend to include additional spatial "slack" into the design of intersections to deal with these problems.

5.7 In urban contexts, the intersection calibration is most important

Remember for calibration that in an urban context, intersections actually have the dominant influence on throughput and are therefore much more important than calibration of the CA for homogenous freeways.

5.8 Pick your questions carefully

The CA design has been chosen for high computational speed while retaining a microscopic resolution. It is clear that this involves compromises with regard to fidelity. As a consequence, the CA version of the TRANSIMS microsimulation will mostly be useful for questions which are as large scale (network size, number of travelers, necessary computing time) as we can handle. Questions where we don't reach these limits might benefit from higher fidelity microsimulation.

REFERENCES

(Bando et al., 1994) M. Bando, K. Hasebe, A. Nakayama, A. Shibata, and Y. Sugiyama, Structure stability of congestion in traffic dynamics, Japan Journal of Industrial and Applied Mathematics 11(2), 203 (1994).

(Bando et al., 1995) M. Bando, K. Hasebe, A. Nakayama, A. Shibata, and Y. Sugiyama, Dynamical model of traffic congestion and numerical simulation, Phys. Rev. E 51(N2), 1035 (1995).

(Benz, 1993) The microscopic traffic simulator AS (Autobahn Simulator), in: *Proc. 1993 European Simulation Multiconference*, edited by A. Pave (1993), p. 486.

(Daganzo, 1993) C. F. Daganzo, The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory, California PATH Report No. UCB-ITS-PRR-93-7, 1993.

(Kerner and Konhäuser, 1993) B. S. Kerner and P. Konhäuser, Cluster effects in initially homogenous traffic flow, Phys. Rev. E 48(4), R2335 (1993).

(Kerner and Konhäuser, 1994) B. S. Kerner and P. Konhäuser, Structure and parameters of clusters in traffic flow, Phys. Rev. E 50(1), 54 (1994).

(Mahmassani et al., 1992) H. S. Mahmassani, T. Hu, R. Jaykrishnan, Dynamic traffic assignment and simulation for advanced network informatics (DYNASMART), *Proceedings of the 2nd International Capri Seminar on Urban Traffic Networks*, Capri, Italy, July 1992.

(Montroll, 1962) E. W. Montroll, Theory and observations of thge dynamics and statistics of traffic on an open road, *Proceedings* (1962), p. 231.

(Nagatani, 1993) T. Nagatani, Jamming transition in the traffic-flow model with two-level crossings, Phys. Rev. E. 48(5), 2533 (1993).

(Nagel and Schreckenberg, 1992) K. Nagel and M. Schreckenvberg, A cellular automaton model for freeway traffic, J. Phys. I France 2, 2221 (1992).

(Nagel and Herrmann, 1993) K. Nagel and H. J. Herrmann, Deterministic model for traffic jams, Physica A 199, 254 (1993).

(Nagel, 1994a) K. Nagel, Life-times of simulated traffic jams, Int. J. Mod. Physics C 5(3), 567 (1994).

(Nagel, 1994b) K. Nagel, High-speed microsimulation of traffic flow, Ph.D. thesis, University of Cologne, 1994.

(Nagel and Rasmussen, 1994) K. Nagel and S. Rasmussen, Traffic at the edge of chaos, in *Artificial Life IV: Proceedings of the Fourth International Workshop on the Synthesis and Simulation of Living Systems*, edited by R. A. Brooks and P. Maes (MIT Press, Cambridge, MA, 1994), p. 222.

(Nagel and Paczuski, 1995) K. Nagel and M. Paczuski, Emergent Traffic Jams, Phys. Rev. E 51, 2909 (1995).

(Nagel, 1995) K. Nagel, Particle hopping models and traffic flow theory, in preparation (1995).

(Rickert, 1994) M. Rickert, Simulationen zweispurigen Autobahnverkehrs mit Zellularautomaten, Master Thesis, University of Cologne, 1994.

(Schwerdtfeger, 1987) T. Schwerdtfeger, Makrosckopisches Simulationsmodell für Schnellstraßennetze mit Berücksichtigung von Einzelfahrzeugen (DYNEMO), Ph.D. thesis, TH Karlsruhe, 1987.

(Stauffer, 1991) D. Stauffer, Computer simulations of cellular automata, J. Phys. AS 24, 909 (1991).

(Wolfram, 1986) S. Wolfram, *Theory and Applications of Cellular Automata*, (World Scientific, SIngapore, 1986).

(Wolinsky el al., 1995) M. Wolinsky, M. Stein, and C. Barrett, IVHS Incident Analysis and Detection Test Bed, Los Alamos National Labortaory Technical Reoprt, LA-UR-95-1662, 1995.

(Wylie et al., 1992) B. J. N. Wylie, D. McArthur, M. D. Brown, PARAMICS parallelisation schemes, Report EPCC-PARAMICS-CT.10, Edingburgh Parallel Computing Centre, University of Edingburgh, 1992.