Chapter 7

Interacting Vehicles: “Rules of the Game”

In previous chapters, we introduced an intelligent control method for autonomous navigation and path planning. The decision system mainly uses local information, and as a result, the actions learned by the intelligent controller are not globally optimal; the vehicles can survive, but may not be able to reach some of their goals. To overcome this problem, we visualize the interaction between vehicles as sequences of games played between pairs of automata. Every game corresponds to a “state” of the physical environment (the highway and the vehicles) as described in Section 7.3. By evaluating these games, it is possible to design new decision rules, and analyze the interactions by predicting the behavior and the position of each vehicle. Some of the additional modules in Chapter 5 are actually the result of the approach described in detail here.

7.1 Interconnected Automata and Games

Automata are, by design, “simple gadgets for doing simple things.” As a control tool, a learning automaton has limited applications since increasing the number of actions results in a decrease in the speed of responses [Norman68]. The full potential of the stochastic automaton is realized when multiple automata interact. For example, the separation of lateral and longitudinal actions by using two automata leads to faster convergence, while bringing the issues of interaction and coordination to the design problem. As stated in [Narendra89], the first question that needs to be answered is whether the automata can be interconnected in useful ways for desired group behavior for modeling or controlling complex systems. Narendra and Thathachar also ask if an individual automata can be designed to function as a distributed yet coordinated intelligent control system. The application of learning automata described in previous chapters partially answers that question. This chapter introduces the idea and methodology behind the design of automata as interacting agents.

Although the interaction between automata in our application is slightly different than the definitions given in the literature, we will briefly introduce the basic ideas and previous results on interacting automata. A detailed treatment of the subject can be found in [Narendra89, Lakshmivarahan81].

Similar to other definitions in game theory, two or more automata are assumed to play a game formulated in terms of automata actions and reward-penalty responses. At every step, automata send an action to the environment which in turn evaluates the actions and sends out a feedback. This feedback is called the outcome or *payoff* as in the general game theory. The probability vector of actions defines the mixed strategy of a player/automaton; each element of the probability vector corresponds to a specific action, or *pure strategy.*

Standard game definitions such as *equilibrium point*(*s*), *dominant strategy,* and *pareto optimality*, are defined similarly. In our application, we are interested in equilibrium points where no player has a positive reason to change its strategy, assuming that none of the other players is going to change strategies. Such equilibrium points are also called *Nash equilibria.* All *N*-player games with finite mixed strategy sets have at least one equilibrium. In a two-person zero-sum game, there may be more than one equilibrium point with the same payoff. This is not necessarily true for two-person non-zero sum games.

All descriptions above are defined on a *game matrix*. A game matrix is used to show the payoff structure of a game. For example, for a two-player game where each player has two actions (pure strategies), the game matrix D can be shown as:

$$D = \begin{bmatrix} \left(d_{11}^1, d_{11}^2\right) & \left(d_{12}^1, d_{12}^2\right) \\ \left(d_{21}^1, d_{21}^2\right) & \left(d_{22}^1, d_{22}^2\right) \end{bmatrix}$$

where d_{ij}^k is the payoff[1] to player k when the players play pure strategies i and j. For a two-player zero-sum game, the equality $d_{ij}^1 = -d_{ij}^2$ is true. There were several learning schemes designed for zero-sum games; the solution to the problem (of obtaining the pareto optimal[2] or equlibrium point outcome) is given in [Lakshmivarahan82]. It is known that linear reward-inaction L_{R-I} and ε-optimal linear reward-penalty $L_{R-\varepsilon P}$ (with b << a) schemes guarantee that the automata reach the best possible solution, *i.e.*, the expected penalty reaches its minimum value. The results are also valid for identical payoff games [Lakshmivarahan81], and may be extended to *N*-player games [Narendra89].

For non-zero sum games, however, unique rational solutions are difficult to define. While considering non-zero sum automata games, it is not possible to evaluate the performance of the automata using game theoretic arguments. "The role of the learning schemes must be reevaluated in such situations" [Narendra89]. Unfortunately, relatively little work exists in this area. Here, we will attempt to provide some insight on automata interaction of several vehicles by defining games of automata and vehicles. The results obtained for non-zero sum games can be extended to *N*-player non-zero sum games to show that the mixed strategies will converge to the equilibrium point(s). In the case of multiple equilibrium points, the convergence depends on the initial conditions of the probability vectors.

It is also important to note that, in an automata game, the players are not aware of the mixed strategy used by the other player(s) nor its previous actions. In fact, the players do not even have the knowledge on the distribution of random payoff structure as a function of the strategy combinations. The interaction medium between automata is the payoff function.

---

[1] This is still the probability of receiving penalty, and therefore a player wants to *minimize* its payoff.

[2] In an *N*-person game, an outcome is said to be pareto optimal if there is no other outcome in which all players simultaneously do better, *i.e.*, receive less penalty from the environment.

7.2 Interacting Automata and Vehicles

As described in Chapters 4 and 5, the planning layer of an autonomous vehicle includes two automata for lateral and longitudinal actions. Furthermore, the actions of each automaton, sooner or later, affects the physical environment, thus creating another level of interactions although indirect. In this section, we will describe the interaction between automata in the same vehicle, and the automata in different vehicles. Furthermore, the nature of the games "played" by learning automata is defined.

7.2.1 Interacting Automata in an Autonomous Vehicle

Each vehicle's intelligent path controller consists of two automata. The interaction between these two automata can be visualized as in Figure 7.1. As described in Section 5.3.3 (Figure 5.19), lateral and longitudinal automata synchronously update their action probabilities based on the response(s) of the environment. Furthermore, the value of the current longitudinal action changes the environment response to the lateral automaton. In other words, the lateral environment[3] is determined by the current longitudinal actions. This is different than an automata game since the interaction is not via the environment, but the automata directly interact. The idea of interacting automata was first introduced in [Wheeler85]. The resulting configurations can be viewed as games of automata with a particular payoff structure. Some of the results mentioned in Section 7.1 can then be applied to such interconnected automata. The model given in Figure 7.1 is an example of interconnected automata with a changing environment.

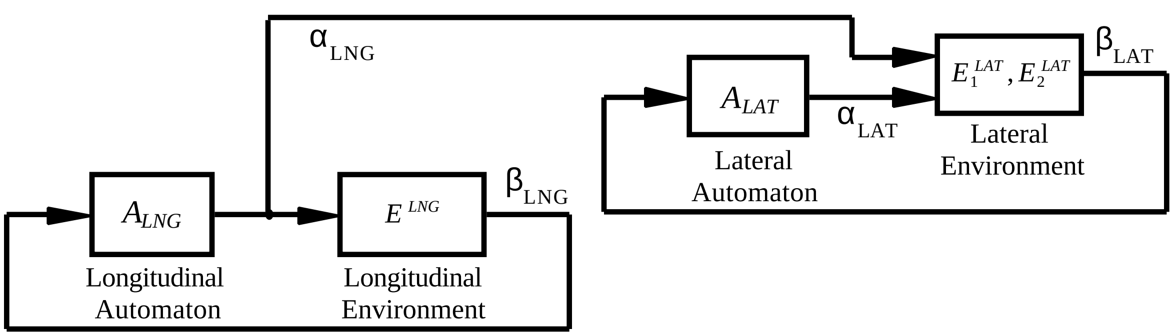


Figure 7.1. The longitudinal automaton determines the lateral automaton's environment (Adapted from [Narendra89]).

We know that lateral automaton A_{LAT} can operate in two stationary environments E_i^{LAT} (i = 1, 2). The difference between these two automaton environments is the response of the headway module (see Section 5.3.3). In some situations, the choice of longitudinal action α_{LONG} affects the response of the lateral environment. All other environment changes are due to changes in the physical environment, and we visualize those changes as switching environments described in Section 7.3. Due to interconnections, one would expect the longitudinal automaton A_{LNG} to

[3] The word 'environment' here describes the automata environment, not the physical one.

converge to its best action using absolutely expedient schemes. Lateral automaton A_{LAT} in turn would converge to the best action in the environment determined by A_{LNG}.

Let us assume that an automated vehicle finds itself in the rightmost lane of a two-lane highway, say by merging from an exit. Consider the situation in the first few seconds where the automata environment stays stationary. With a relatively fast update rate, this assumption is possible. The lateral action SR will receive a penalty until the vehicle shifts lane. Assuming the vehicle is in its desired lane and speed range, the environment response for the longitudinal and lateral actions depends on the output of the headway and left side sensor modules as given in Table 7.1. Let us also assume that the probability that a sensed vehicle is moving away from the autonomous vehicle is 1/3, and we know the probability of a vehicle being in a particular sensor range[4] as given in Table 7.2.

Range / Action	**A**	***B***	***C***	***D***	**E**	**none**
ACC	1	1	1/3	0	0	0
DEC	0	0	0	0	0	0
SM	1	1/3	0	0	0	0
SL	0	0	0	1	1/3	0
SR	1	1	1	1	1	1
SiL [5]	1	1/3	1/3	0	0	0
	0	0	1/3	0	0	0

Table 7.1. Assumed probabilities of penalty for each action based on the front and side sensors (see Figure 5.1 for range definitions).

.

Probability	**A**	**B**	**C**	**D**	**E**	**none**
(a)	1/10	1/10	1/10	1/10	1/10	_

Table 7.2. Probability of sensing a vehicle in the sensor range.

Based on the numbers given in Tables 7.1 and 7.2, the probabilities of receiving penalty for both automata can be calculated. For example, for longitudinal action ACC, the payoff is $\frac{1}{10}\ 1+\frac{1}{10}\ \frac{1}{3}+\frac{1}{10}\ \frac{1}{3}+\frac{1}{10}\ 0+\frac{1}{10}\ 0+\frac{1}{2}\ 0=\frac{7}{10}$. Therefore, the game matrix for longitudinal and lateral automata is:

[4] Note that this is not required for the actual application, it is just an assumption to illustrate the effect of changing environment for interconnected automata.

[5] The second row indicates the probability of penalty when the longitudinal action is DEC.

$$
\begin{array}{c} \\ ACC \\ DEC \\ SM \end{array}
\begin{array}{c}
\begin{array}{ccc} \quad SL \quad & \quad SR \quad & \quad SiL \quad \end{array} \\
\left[\begin{array}{ccc}
\left(\frac{7}{30},\frac{4}{30}\right) & \left(\frac{7}{30},1\right) & \left(\frac{7}{30},\frac{5}{30}\right) \\
\left(0,\frac{4}{30}\right) & \left(0,1\right) & \left(0,\frac{1}{30}\right) \\
\left(\frac{4}{30},\frac{4}{30}\right) & \left(\frac{4}{30},1\right) & \left(\frac{4}{30},\frac{5}{30}\right)
\end{array}\right]
\end{array}
$$

Entries in the first and third row correspond to the first environment for the lateral automaton while the entries of the second row to the second environment. The difference between the two environment is the longitudinal action DEC: when the longitudinal automaton chooses DEC, lateral environment A_{LAT} is switched to second environment which changes the response to the lateral action SiL. If the vehicle is slowing down, lane shifts are discouraged (Figure 5.19). If the automata were not connected, absolutely expedient algorithms would be expected to converge to actions DEC and SL (without the connection, the lateral action SL is optimal since the penalty response from the front sensor in not suppressed.) When the automata are connected as described, the optimal actions are DEC and SiL. Based on the probabilities given in Tables 7.1 and 7.2, this solution is pareto optimal and an equilibrium point for this game. In a situation where the probability of sensing a vehicle in region C is much larger than others, the payoff of the lateral action SL is always less than SiL, and the action pair (DEC, SL) becomes the pareto optimal solution. The two-automata game will converge to this new optimal solution in this case.

Note that the situation above is very specific when we consider all sensors: the vehicles must be cruising at their desired speed range and lane, the pinch sensor must not send a penalty response, etc. For all other situations, the two automata may be considered 'unconnected.' Using the algorithms given in Chapter 6, the automata will converge to their best (optimal) actions separately. The interaction between automata is via the physical environment, and for the duration of a specific "game" we consider it to be stationary, resulting in a stationary automata environment. Of course, the solution of such a 'disjoint game' will be an equilibrium point (and a pareto optimal solution) due to the convergence characteristics of the reinforcement schemes.

7.2.2 Interacting Vehicles

While two automata in each autonomous vehicle form an interconnected automata pair that is guaranteed to reach the optimal solution for a stationary (physical) environment, interaction between autonomous vehicles creates another level of interaction via the physical environment. As seen in Figure 7.2, automata actions from other autonomous vehicles changes the physical environment which in turn affects the feedback responses sent to an automaton. This type of interaction is indirect, and therefore cannot be formulated using a game matrix. Furthermore, the fact that the game matrix needs to be time variant when considering multiple interacting vehicles complicates the matter.

Instead, we visualize the automata environments resulting from the ever-changing physical environment as a switching environment. Based on a certain change in the physical world, the automata environment is changing from one state to another. Every automata environment state includes a different set of feedback responses to lateral and longitudinal automata. All these different states of the environment are assumed to be stationary since the automata converge to the optimal actions before another change takes place. As discussed previously, the automata pair in every vehicle is absolutely expedient (or optimal when using linear schemes) in each stationary environment, and therefore, the autonomous vehicles are guaranteed to reach an intelligent decision. Once a decision is made and sent to the regulation layer, corresponding actions are fired and the physical environment changes, forcing the automata environment to switch to another state.

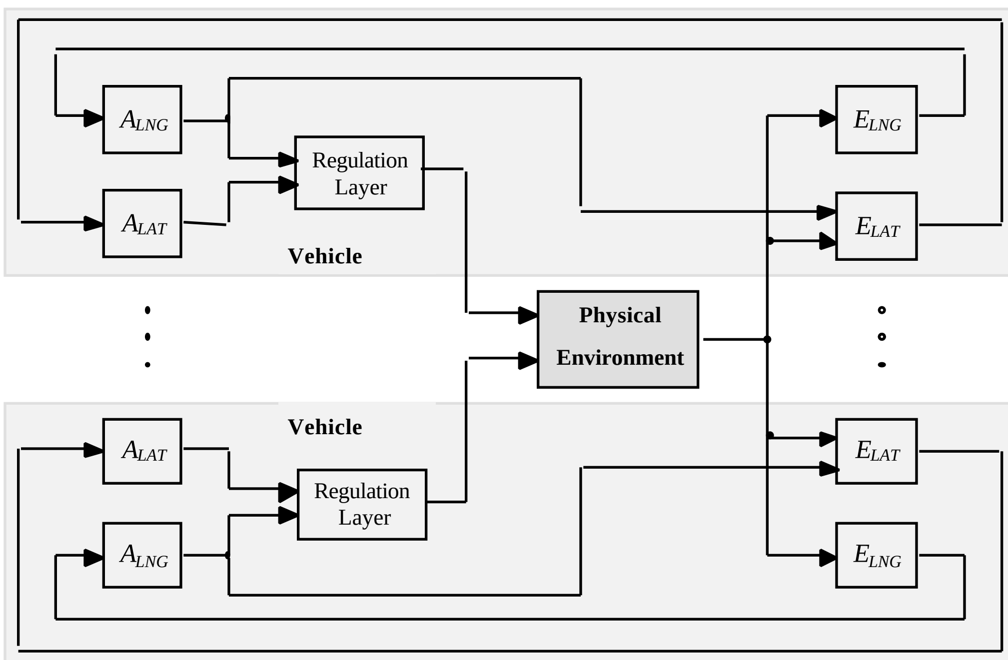


Figure 7.2. Multiple vehicles interacting through the physical environment.

It is important to realize that lateral or longitudinal actions need not be fired for the automata environment to switch from one state to another. For example, if two vehicles cruising at their desired speed are in adjacent lanes, and if their idle actions SiL and SM are optimal, the physical environment may change due to the speed difference between the vehicles. The moment that one of the vehicles clears the other's side sensor detection area, the lateral automata environment for both vehicles change. Similarly, when ideal actions are fired, the physical and automata environments may not change. The interactions between the actions and the physical environment, and the physical and automata environment, are fairly complicated. In the next section, we will introduce a representation scheme that will facilitate analysis of the changes in the physical environment in relation to the automata environment.

7.3 States of the Switching Automata Environment

Although illustrating vehicle interactions as automata games for every (stationary) instance of the automata environment is not feasible, it may be possible to define a similar matrix for all the actions of autonomous vehicles. Consider a situation where two autonomous vehicles interact via their sensors, and communication (or signaling) devices. The physical presence of one vehicle is affecting the automata environment of the other via range sensors. Note that the vehicles are not actually aware of the presence of others: the automata in each vehicle are simply trying to find the best action to take given a set of feedback responses from the modules. For example, consider two vehicles cruising at their desired speeds and occupying the spot next to each other without any other vehicles in the sensor ranges (Figure 7.3a). Let us also assume that the vehicles have no lane preferences, resulting in a constant reward from the lane detection module. In this case, the penalty-reward structure for both vehicles is given in Figure 7.3c. The matrix in this table gives the conditions in a particular automata environment resulting from physical location, current conditions, and predefined vehicle parameters. For example, lateral action SR for vehicle 1 receives penalty from the right sensor, and the combined response is a penalty. This penalty response is shown as a shaded lower triangle in the matrix; reward responses are shown by white triangles. Combined feedback responses are evaluated using the decision structure given in Section 5.4.1. Again, note that the entries in this matrix are the environment responses, not the probabilities of receiving a penalty.

As seen from the Figure 7.3c, the only intelligent choice for vehicle 2 is the idle action (SiL and SM) while vehicle 1 may shift the left lane or cruise on its current lane while keeping its speed. Let us assume that vehicle 1 stays in its current lane, and it is moving faster than vehicle 2. After a short period of time, the physical situation shown in Figure 7.3b is reached; vehicle 1 moves out of the left sensor range of vehicle 2. In the mean time, the idle actions may be fired repeatedly in both vehicle. This new situation affects the automata environment as shown in Figure 7.3d. The vehicles now have more choices: vehicle 2 may also shift left while it is possible for vehicle 1 to shift right, left or stay in the middle lane. There are six pairs of possible (vehicle) actions; five of these force the automata environment to switch to another state. The idle action pairs (SiL and SM) do not affect the physical environment instantaneously; in fact as far as the two vehicles are concerned, these actions do not change the automata environment at all. For our analysis of vehicle interactions, all the physical situations where a vehicle is outside of another vehicle's sensor range lead to the same automata environment since the reward-penalty structure does not change.

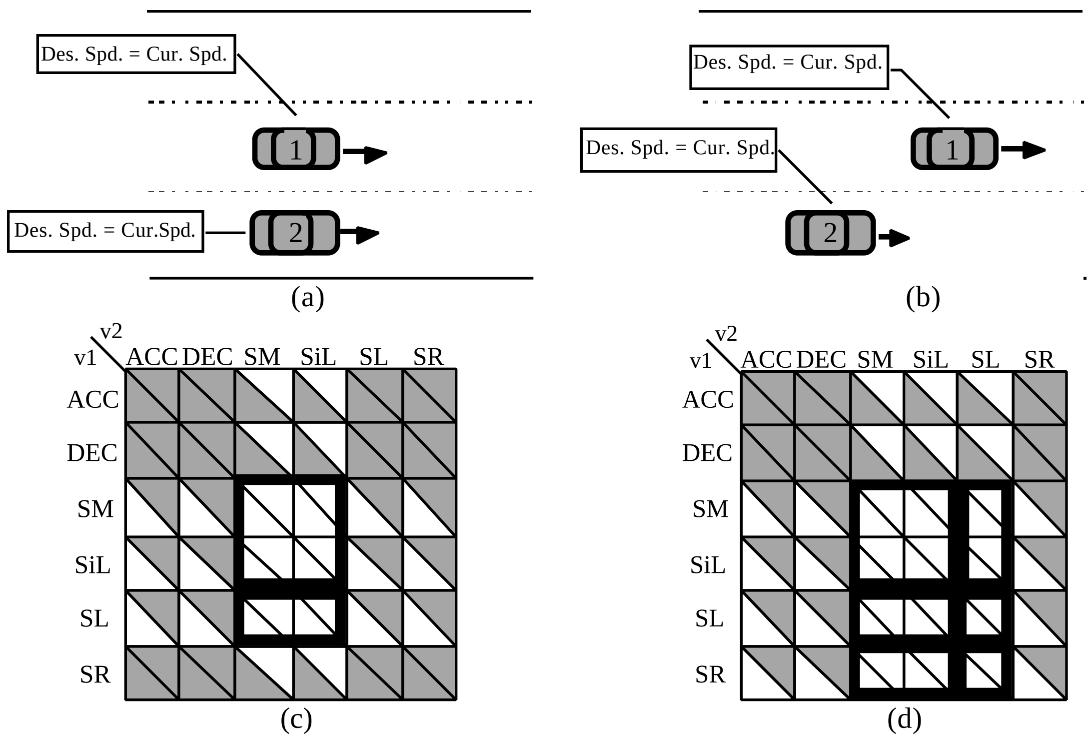


Figure 7.3. Situations for two interacting vehicles (a, b), and combined environment responses (c, d) for situations (a) and (b) respectively (shaded triangles indicate a penalty; upper triangles are associated with vehicle 2; optimal action pairs are indicated with black borders).

Consider a situation with three vehicles as shown in Figure 7.4a. Vehicle 1 and 2 are autonomous; vehicle 3 is not automated and cannot sense nor communicate. It is just an obstacle as far as the 'intelligent' vehicles are considered. Vehicle velocities are given as $V_1 = V_3 > V_2$. Vehicle 2 has no lane preference while vehicle 1 wants to shift to the middle lane. However, vehicle 1 cannot shift immediately to the middle lane since vehicle 2 is in the side sensor range (Figure 7.4a); the automata environment for this situation is given in Figure 7.4d (The actions SM and SiL are combined as a single action *IDLE*. If a lateral action other than SiL is chosen, the row/column for combined action IDLE refers to the lateral idle action, and vice versa. If both SiL and SM are chosen the table shows the OR-ed response).

Due to velocity differences, vehicle 2 drifts away from vehicle 1's sensor range (Figure 7.4b), and the automata environment switches (Figure 7.4e). In the mean time, the idle actions are fired repeatedly. With the new environment, the number of possible actions for vehicles 1 and 2 increases, and lateral action SL becomes the optimal solution for vehicle 1. As a result, vehicle 1 changes lane (Figure 7.4c) which in turn causes another automata environment change (Figure 7.4f).

Using the same reasoning, we can establish which automata environment corresponds to what physical situation-vehicle condition pairs. Since the automata in each vehicle uses optimal or absolutely expedient algorithms, convergence to the optimal solution is guaranteed for all these situations, provided that the automata have enough time to learn. Using this method, it is then possible to predict how the vehicle will react to a specific physical situation. This will enable us to define *highway scenarios* as described in the next section, and find solutions for intelligent path planning.

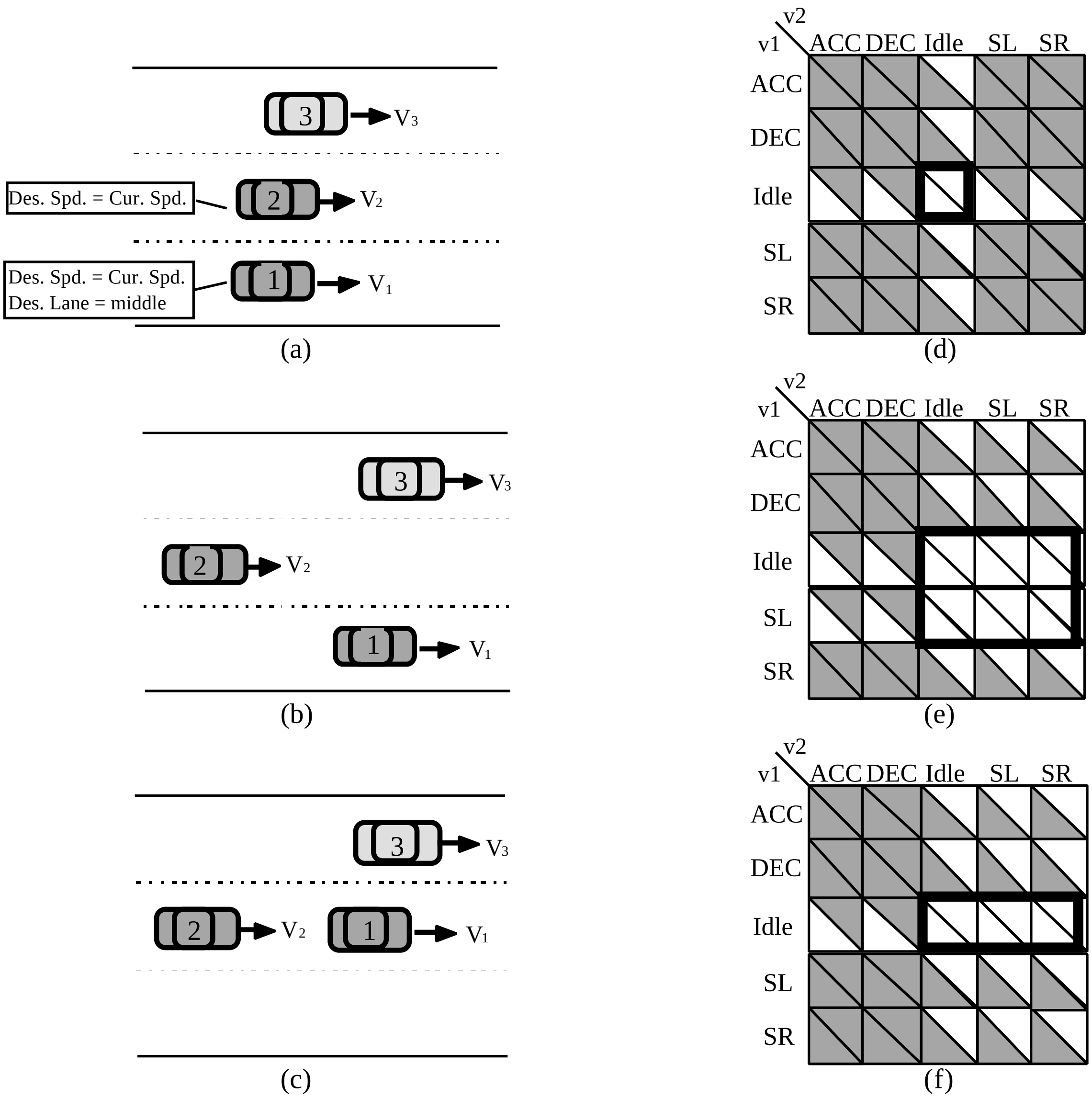


Figure 7.4. Changes in the physical (left) and automata environments (right): vehicle 1 shifts to the middle lane.

7.4 Highway Scenarios as State Transitions

Autonomous vehicles described in previous chapters are able to avoid collisions by making decisions based on local information available to them. However, the resulting vehicle path may not be the best solution toward the problem of congestion. Furthermore, some of the vehicle paths may conflict, and prevent the vehicles from reaching their desired goals. The problem exists for both hierarchical control and autonomous vehicle approaches. However, the second approach is a decentralized control method in nature, and finding a solution may be much more difficult than finding a globally optimal path strategy with a hierarchical architecture that has all the information about the highway situation.

This problem with the autonomous vehicle approach has not yet been answered by the previous research efforts, while the hierarchical control structure inherently possesses the methodology to solve the problem. We visualize a possible situation with multiple interacting vehicles as a sequence of environment states. For all states of the physical environment ¾ which includes the positions of the vehicles and current parameters defining their behavior ¾ a corresponding automata environment can be defined. The automata environment is analyzed to predict possible physical environment changes. These changes will be illustrated as state transitions. State diagrams formed using possible environment state transitions can then be used for analysis as well as design purposes.

Consider two vehicles sharing a 3-lane highway. The possible physical situations are given in Figure 7.5. Besides the relative lateral positions, we assumed that only three possibilities exist for relative longitudinal positions. The distinguishing factor between these positions is whether a vehicle is in the side sensor range or not. Also note that for each state given in Figure 7.5, there is a reciprocal state with switched vehicle positions, denoted by an asterisk (*e.g.*, B1*; therefore, total number of states is 24). Figure 7.5 is basically a list of situations that are interesting for the analysis of vehicle interactions; there may be more situations. Appendix D.1 shows the relative positions of the two vehicles in a larger physical environment. Situations where the vehicles fall outside of the sensor range are combined into a single ‘state,’ and further simplification gives the twelve situations given in this section.

Note that for any physical environment state in Figure 7.5, there may be multiple corresponding automata environments due to several factors such as desired speed, desired lane as indicated in Figure 7.3 and 7.4.

To analyze the behavior of autonomous vehicles and the conflicts resulting from their interactions, we define *highway scenarios*. A *scenario* is a specific situation with physical locations of vehicles, their sensor outputs, and internal parameters such as desired lanes and desired speeds. Once we know the automata environment at the beginning of a scenario, we can predict the (state) changes in the physical environment. Then, all possible changes are combined to form a state diagram showing the progress of the physical environment. The transitions from one state are the direct results of the automata environment given by the matrices such as those in Table 7.4.

To illustrate the idea, let us consider the situation A1 in Figure 7.5. Assume the probabilities of possible vehicle actions are equal. For example, vehicle 1 may shift right, move ahead of the vehicle 2, slow down or do nothing while vehicle 2 keeps its speed and lane the same. These vehicle actions will cause transitions from A1 to A3, B1, C1 and A1 respectively. All possible transitions between defined states are shown in Figure 7.6. The links indicated with ‘*’ show a transition to a reciprocal state, *e.g.*, D1 → B1*, B2 → D2* or B3* → D1*.

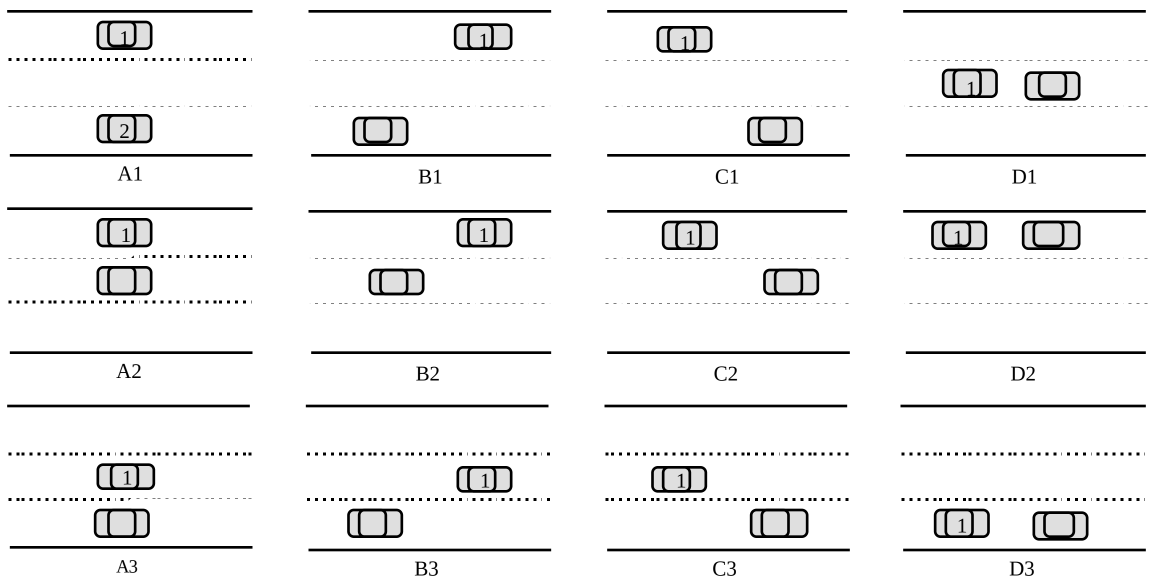


Figure 7.5. Possible physical environment states for 2 vehicles in 3-lane highway.

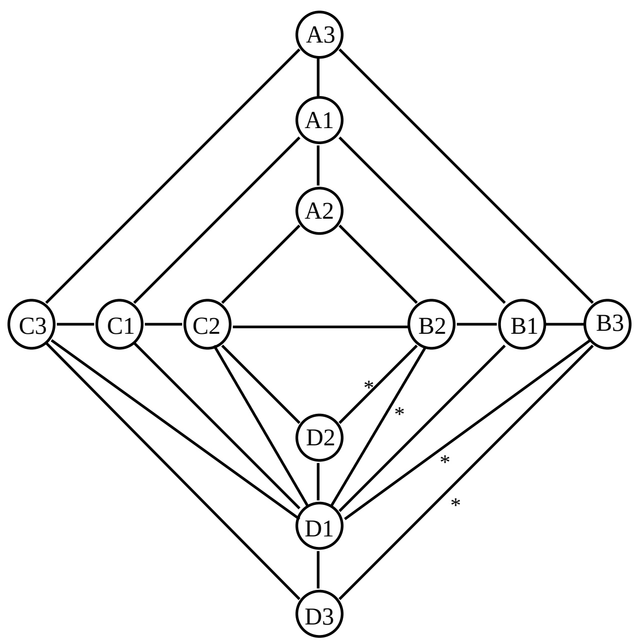


Figure 7.6. State transition diagram for two vehicles on a 3-lane highway; reflexive transitions are not shown (see also Figure 7.5).

Assuming equal probability for all vehicle actions in state A1, we can define a transition matrix as follows. At state A1, vehicle 1 can shift right, move ahead, stay back or keep its speed and lane. Similarly, vehicle 2 can shift left, move ahead, stay back or keep its speed and lane. We assume that vehicles do not take actions simultaneously, *i.e.*, only one vehicle can fire an action at a given time. Thus, there are eight possible transitions with equal probability (Table 7.3).

Actions	**Transition to**
Vehicle 2 idle, Vehicle 1 move ahead	B1
Vehicle 2 idle, Vehicle 1 slips back	C1
Vehicle 2 idle, Vehicle 1 idle	A1
Vehicle 2 idle, Vehicle 1 shifts right	A3
Vehicle 1 idle, Vehicle 2 move ahead	C1
Vehicle 1 idle, Vehicle 2 slips back	B1
Vehicle 1 idle, Vehicle 2 idle	A1
Vehicle 1 idle, Vehicle 2 shifts left	A2

Table 7.3. Action-transition pairs for state A1.

Therefore, the probabilities of transition from state A1 to states B1, C1, A2 and A3 are 0.250, 0.250, 0.125, and 0.125 respectively. The environment stays at state A1 with probability of 0.250. All other transitions for state A1 are not possible. Similarly, all probabilities of transition for other states can be written, and the transition matrix *T* becomes:

		A			*B*			*C*			*D*			*A**			*B**			*C**			*D**		*K*
*A*1	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{4}$	0	0	$\frac{1}{4}$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
*A*2	$\frac{1}{9}$	$\frac{2}{9}$	0	0	$\frac{2}{9}$	0	0	$\frac{2}{9}$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	$\frac{2}{9}$
*A*3	$\frac{1}{9}$	0	$\frac{2}{9}$	0	0	$\frac{2}{9}$	0	0	$\frac{2}{9}$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	$\frac{2}{9}$
*B*1	$\frac{1}{4}$	0	0	$\frac{1}{2}$	$\frac{1}{8}$	$\frac{1}{8}$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
*B*2	0	$\frac{2}{9}$	0	$\frac{1}{9}$	$\frac{4}{9}$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	$\frac{1}{9}$	$\frac{1}{9}$	0	0
*B*3	0	0	$\frac{2}{9}$	$\frac{1}{9}$	0	$\frac{4}{9}$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	$\frac{1}{9}$	0	$\frac{1}{9}$	0
*C*1	$\frac{1}{4}$	0	0	0	0	0	$\frac{1}{2}$	$\frac{1}{8}$	$\frac{1}{8}$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
*C*2	0	$\frac{2}{9}$	0	0	0	0	$\frac{1}{9}$	$\frac{4}{9}$	0	$\frac{1}{9}$	$\frac{1}{9}$	0	0	0	0	0	0	0	0	0	0	0	0	0	0
*C*3	0	0	$\frac{2}{9}$	0	0	0	$\frac{1}{9}$	0	$\frac{4}{9}$	$\frac{1}{9}$	0	$\frac{1}{9}$	0	0	0	0	0	0	0	0	0	0	0	0	0
*D*1	0	0	0	0	0	0	0	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{2}{5}$	0	0	0	0	0	$\frac{1}{10}$	$\frac{1}{10}$	0	0	0	0	0	0	0	$\frac{1}{5}$
*D*2	0	0	0	0	0	0	0	$\frac{1}{8}$	0	0	$\frac{1}{2}$	0	0	0	0	$\frac{1}{8}$	0	0	0	0	0	0	0	0	$\frac{1}{4}$
*D*3	0	0	0	0	0	0	0	0	$\frac{1}{8}$	0	0	$\frac{1}{2}$	0	0	0	0	0	$\frac{1}{8}$	0	0	0	0	0	0	$\frac{1}{4}$
*A*1*	0	0	0	0	0	0	0	0	0	0	0	0	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{4}$	0	0	$\frac{1}{4}$	0	0	0	0	0	0
*A*2*	0	0	0	0	0	0	0	0	0	0	0	0	$\frac{1}{9}$	$\frac{2}{9}$	0	0	$\frac{2}{9}$	0	0	$\frac{2}{9}$	0	0	0	0	$\frac{2}{9}$
*A*3*	0	0	0	0	0	0	0	0	0	0	0	0	$\frac{1}{9}$	0	$\frac{2}{9}$	0	0	$\frac{2}{9}$	0	0	$\frac{2}{9}$	0	0	0	$\frac{2}{9}$
*B*1*	0	0	0	0	0	0	0	0	0	0	0	0	$\frac{1}{4}$	0	0	$\frac{1}{2}$	$\frac{1}{8}$	$\frac{1}{8}$	0	0	0	0	0	0	0
*B*2*	0	0	0	0	0	0	0	0	0	$\frac{1}{9}$	$\frac{1}{9}$	0	0	$\frac{2}{9}$	0	$\frac{1}{9}$	$\frac{4}{9}$	0	0	0	0	0	0	0	0
*B*3*	0	0	0	0	0	0	0	0	0	$\frac{1}{9}$	0	$\frac{1}{9}$	0	0	$\frac{2}{9}$	$\frac{1}{9}$	0	$\frac{4}{9}$	0	0	0	0	0	0	0
*C*1*	0	0	0	0	0	0	0	0	0	0	0	0	$\frac{1}{4}$	0	0	0	0	0	$\frac{1}{2}$	$\frac{1}{8}$	$\frac{1}{8}$	0	0	0	0
*C*2*	0	0	0	0	0	0	0	0	0	0	0	0	0	$\frac{2}{9}$	0	0	0	0	$\frac{1}{9}$	$\frac{4}{9}$	0	$\frac{1}{9}$	$\frac{1}{9}$	0	0
*C*3*	0	0	0	0	0	0	0	0	0	0	0	0	0	0	$\frac{2}{9}$	0	0	0	$\frac{1}{9}$	0	$\frac{4}{9}$	$\frac{1}{9}$	0	$\frac{1}{9}$	0
*D*1*	0	0	0	0	$\frac{1}{10}$	$\frac{1}{10}$	0	0	0	0	0	0	0	0	0	0	0	0	0	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{2}{5}$	0	0	$\frac{1}{5}$
*D*2*	0	0	0	0	$\frac{1}{8}$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	$\frac{1}{8}$	0	0	$\frac{1}{2}$	0	$\frac{1}{4}$
*D*3*	0	0	0	0	0	$\frac{1}{8}$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	$\frac{1}{8}$	0	0	$\frac{1}{2}$	$\frac{1}{4}$
K	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

Each row of the matrix T shows the probabilities of transitions from a state, and sums up to 1. Such a matrix is called a *stochastic matrix*. As seen above, it also includes the reciprocal states obtained by changing the locations of the vehicles in states shown in Figure 7.5. The last row and column of the matrix corresponds to a collision state, denoted by K.

It is also possible to consider simultaneous vehicle actions. For example, if vehicle 1 shifts to the middle lane while vehicle 2 move ahead in state A1, a transition to state C3 will occur. The probability of simultaneous actions is relatively small, and also, even if this is the case, the transition may be thought of as a sequence of to separate transitions from state A1 to C1 or A3, and then to state C3. Thus, we assumed that only one vehicle may take an action at one time.

For the scenario above, we may assume that the probability of transitions are constant for all steps. If the probabilities of finding the environment at a given state are known, the row vector which consists of these probabilities can then be multiplied with the transition matrix to evaluate the probability distribution between states at the next time step. Note that the use of the transition matrix suggests that probability of the environment being in one state at time n depends only on the probability distribution at the previous step n-1. Therefore, the state transitions of the physical environment given above constitute a Markov Chain.

Let us assume that the environment is at state A1 initially (n = 0). The row vector defining the probability distribution for the states is then:

$$s(0) = \begin{bmatrix} 1 & 0 & \quad & 0 \end{bmatrix}_{1x25}$$

The probabilities of being at a specific state is then calculated by:

$$s(0)\ \ T = s(1) = \begin{matrix} A1 & A2 & A3 & B1 & B2 & B3 & C1 & C2 & C3 & \quad & K \\ [0250 & 0125 & 0125 & 0250 & 0 & 0 & 0250 & 0 & 0 & \quad & 0] \end{matrix}$$

For consecutive steps, the probability of being at state K increases. At n = 20, we have:

$$s = \begin{bmatrix} 23 & 15 & 15 & 23 & 18 & 18 & 23 & 18 & 18 & 11 & 7 & 7 & 7 & 6 & 6 & 7 & 8 & 8 & 7 & 8 & 8 & 10 & 7 & 7 & 716 \end{bmatrix} .10^{-3}$$

The state K for this Markov chain is an *absorbing state*[6]. Then, we conclude that two vehicles randomly choosing one of the possible actions in a three-lane highway will collide sooner or later, if, of course, our probabilistic model is correct.

7.4.1 Scenario 1: Two Vehicles on a Three-Lane Highway

Now, let us consider the situation A1 in Figure 7.5, now with two intelligent vehicles equipped with the sensor and teacher modules, except for the flag structures discussed in Chapter 5. The speeds of the vehicles, and their lateral positions are the same. Vehicle 1 wants to shift to lane 3, and vehicle 2 to lane 1 (Figure 7.7). Since they are initially traveling at the same speed, with the

---

[6] A subset C of the state space for a Markov process is said to be closed if for every state s_i C and s_j C, and the transition probability t_{ij} is zero. If the subset C consists of a single state, that state is called an absorbing state. Only state K has that probability in our example.

basic sensors and modules, there are no decisions/actions that would lead to a goal state (or states). Starting at state A1, possible transitions are to states A1, A2, and A3. For transitions to states A2 or A3, one of the vehicles must fill its memory vector with a lane shifting action much faster than the other. If one of the pinch modules detects the other vehicle signaling for lane shift, the state A1 will be permanent, and idle action SM will fire repeatedly.

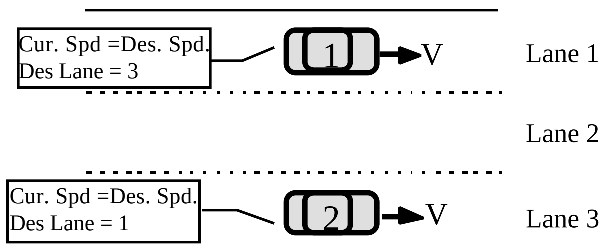


Figure 7.7. Scenario 1: Two vehicles with conflicting desired paths.

If the vehicles were to change their speed, the transitions shown in Table 7.4 are possible with the current internal parameters. Figure 7.8 shows the corresponding state diagram where arrows indicate the direction of transition. Two possible chains with the solution states as absorbing states B1* and C1* are distinguished with solid and dashed lines.

From	To
Chain 1	
A1	A1,A2, A3
A2	A2
A3	A3
Chain 2	
B1	B1,B2, B3
B2	B2,D1*, D2*
B3	B3,D1*, D3*
D1*	D1*,C2*, C3*
D2*	D2*,C2*
D3*	D3*,C3*
C2*	C1*, C2*
C3*	C1*,C3*
C1*	C1*
Chain 3	
C1	C1, C2, C3
C2	C2, D1, D2
C3	C3,D1,D3
D1	D1,B2*, B3*
D2	D2,B2
D3	D3,B3
B2*	B1*, B2*
B3*	B1*,B3*
B1*	B1*

Table 7.4. Possible transitions for Scenario 1.

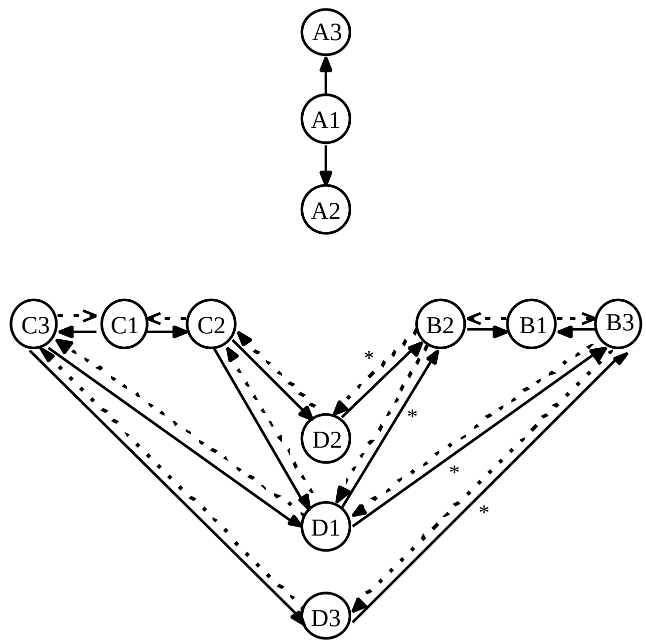


Figure 7.8. Possible chains for Scenario 1 (reflexive transitions are not shown; two chains are distinguished with dashed and solid lines).

From Table 7.4 and Figure 7.8, it is obvious that if the initial state is A1, the vehicles will not be able to reach their desired lanes under the current circumstances. The final situation is then A1, A2 or A3. The states C1* and B1* represent the goal state for this situation and are reachable from all other states in their respective chains. Furthermore, these two states are absorbing states for chains 2 and 3 respectively. Given the vehicle's desire to shift lanes, the physical environment is guaranteed to switch to states C1* or B1* starting at any state in chain 2 or 3. Therefore, all we need to do is to force the physical environment to switch from A1, A2 or A3 to any of the states in chain 2 or 3. Possible transitions are from Ai to Bi or Ci where i = 1,2,3. All these transitions require a speed change for at least one of the vehicles.

Consider the situation A1 with the automata environment given in Figure 7.9a. If the penalty-reward structure can be changed to one of the matrices shown in Figures 7.9b-7.9e, or any suitable combination of longitudinal actions, the physical environment will then switch to state B1 or C1 depending on the chosen action(s). If at least one of the vehicles can be forced to change its speed by changing its automata environment, the physical environment will switch after some time, leading to a goal state. The same reasoning is true for states A2 and A3.

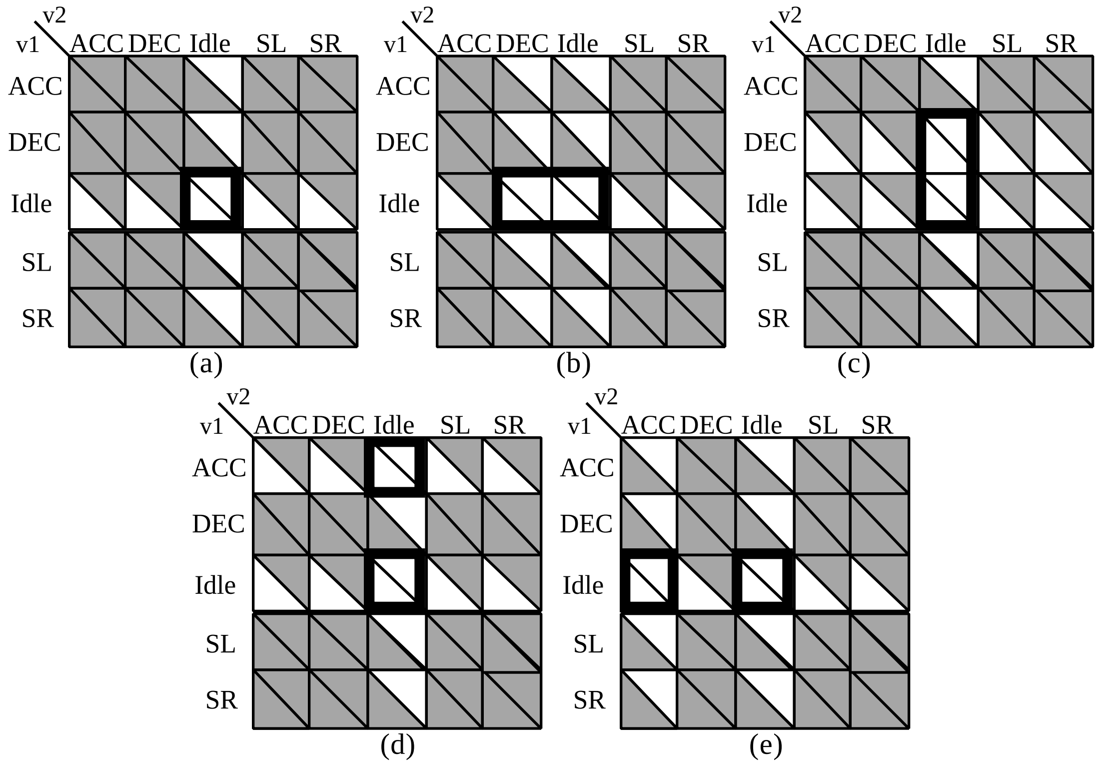


Figure 7.9. Possible penalty-reward structures (b-e) to force physical environment to switch to states B1 or C1 from current state A1.

Therefore, in order to introduce a change to the automata environment, the flag structure given in Section 5.3.1 is designed. If a vehicle cannot shift to its desired lane in a predefined time interval, a flag is set. This flag changes the vehicles’ desired speed, usually to a value smaller than the current speed. The learning automata environment then changes as shown in Figure 7.9b or 7.9c; the transition in physical environment state comes later, attaching the states Ai to chain 2 or 3 (Figure 7.8). Consecutive state transition are automatic for this situation.

7.4.2 Scenario 2: Three Vehicles on a Three-Lane Highway

In this section, we consider three intelligent vehicles in a situation similar to the one given in previous section. As seen in Figure 7.10, the speed and lateral positions of the vehicles are the same, and their desired lane parameters create a conflict. Similar to scenario 1, the solution to the problem lies in changing the relative speeds of the vehicles. Again, the lane flag is used to decrease the speeds of vehicles 1 and 2 to a smaller value than that of vehicle 3.

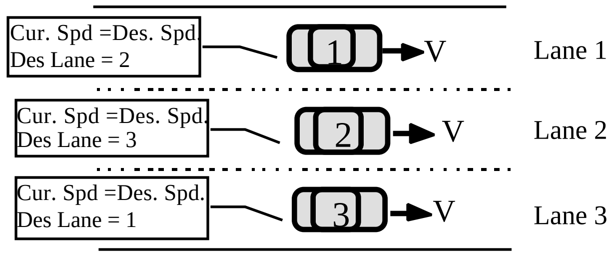


Figure 7.10. Scenario 2: Three vehicles with conflicting desired paths.

All possible environment states for three vehicles on a three-lane highway are given in Appendix D.2. Again, assuming that only one vehicle take can take an action at a time, the state transitions leading to a solution would be similar to the one given in Figure 7.11. This transition diagram only shows a solution when vehicles 1 and 2 slow down. A few other solutions are also possible if different speed adjustment are considered.

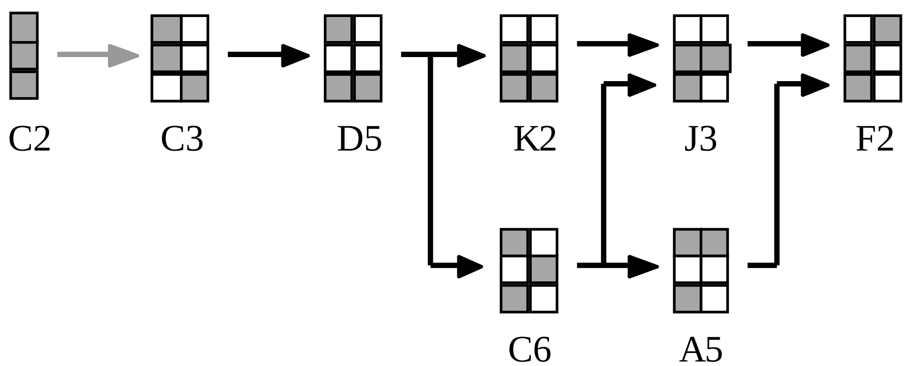


Figure 7.11. A possible chain for Scenario 2: lane flag forces vehicles 1 and 2 to slow down.

All the transitions except the first one are automatic under current circumstances. For the first transition, on the other hand, the lane flag needs to be set in at least one vehicle (if it is vehicle 3), breaking the symmetry. The problem and the solution for this case is similar to the two-vehicle situation given in the previous section. This is not a coincidence; it is due to the *superposition* of two two-vehicle situations.

The term 'superposition' indicates that a three-vehicle situation given here can be treated as three separate two-vehicle interactions. In terms of the two-vehicle states, the state transition diagram above can be written as separate transition diagrams using the two-vehicle situations previously described, as shown in Figure 7.12.

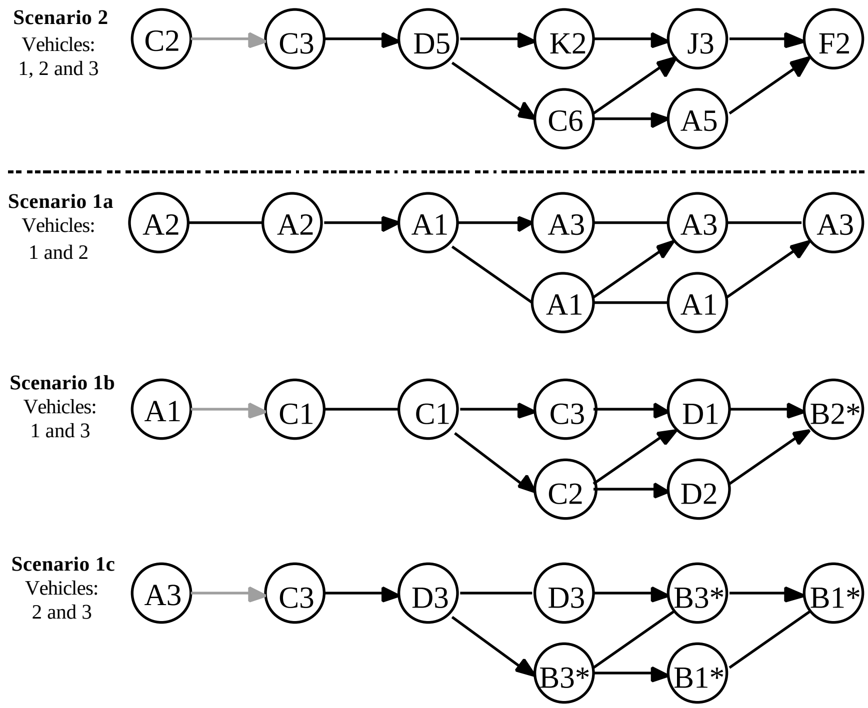


Figure 7.12. Three-vehicle transition diagram can be written as three separate two-vehicle transition diagrams using the definitions in Figure 7.8.

As seen from Figure 7.12, the three-vehicle situation is nothing more than three asynchronous Markov chains representing two-vehicle situations. In Figure 7.12, the corresponding two-vehicle states are aligned with the three-vehicle environment states. Transitions that need to be forced by the lane flag are shown in gray, and they are (and must be) between corresponding states in both three-vehicle and two-vehicle transition diagrams. The two-vehicle scenario including vehicles 1 and 2 is automatic, *i.e.*, there are no conflicts. The other two scenarios both have a synchronous 'forced' transition.

Similarly, for the three-vehicle scenario of Section 5.3.1, it is possible to view the situation as three two-vehicle transition diagrams. Figure 7.13 shows the corresponding two- and three-vehicle chains. Again, the transition breaking the symmetry occurs at the between the corresponding states of three- and two-vehicle scenarios. When three-vehicle situation is forced from state K2 to state L2, two-vehicle scenario of vehicles 1 and 2 is forced from state A3 (corresponding to state K2) to state B3* (corresponding to state L2). Two-vehicle scenario including vehicle 1 and 3 is automatic, the one between 2 and 3 is not significant, *i.e.*, there are no conflicts, nor transitions. The situation between vehicle 1 and 2 is the "defining" scenario.

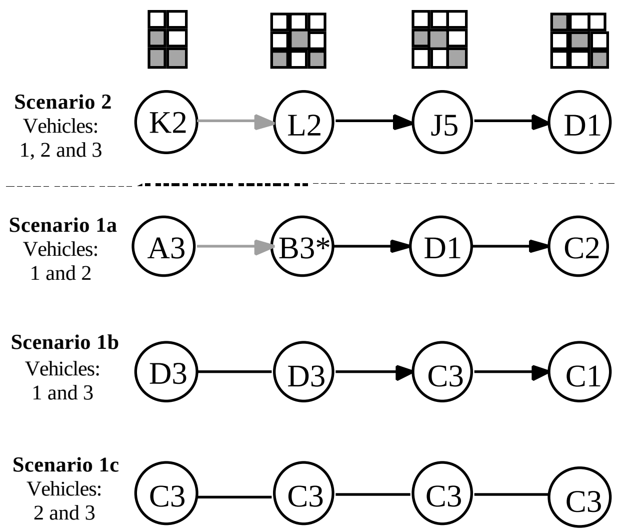


Figure 7.13. Three-vehicle transition diagram is equivalent to two separate two-vehicle diagrams for the example in Section 5.3.1 (See Figure 5.17).

Analysis of the example situations given in this section indicates that it is possible to define complex situations of multiple interacting vehicles as a group of many (conflicting and non-conflicting) two-vehicle situations. A complex scenario is nothing more than a superposition of multiple two-vehicle scenarios. When the state transitions of significant two-vehicle scenarios corresponding to the transition diagram of the N-vehicle ($N > 2$) scenario are studied, it is seen that the key transition that breaks the symmetry in N-vehicle situation is 'synchronous' with at least one of the two-vehicle symmetry-breaking transitions[7].

7.4.3 Scenario 3: Four Vehicles on a Three-Lane Highway

As we discussed in the previous section, a scenario of multiple vehicles can be treated as multiple asynchronous two-vehicle transition diagrams. In the previous example, interaction between vehicle 2 and 3 was not significant since their relative positions and speeds, as well as other parameters, did not change during the time interval of interest. Similarly, the example with four vehicles given in this section will be equivalent to a single two-vehicle situation due to relative position and speed of vehicles. This example also emphasizes the need for a speed flag as described in Section 5.3.2.

Consider the situation given in Figure 7.14. Vehicles 3 and 4 are traveling at their desired speed and lane, unaware of the fact that vehicles 1 and 2 are approaching. Assume that speeds of vehicles 1 and 2 are equal to each other and greater than those of vehicles 3 and 4 while vehicles 3 and 4 are traveling at the same lateral position with equal speeds. This situation is the most complicated among similar situations. For vehicles 1 and 2 to keep their desired speed, it is necessary that they shift to the middle lane and pass vehicles 3 and 4. Although they will be able

[7] The term 'synchronous' here means that corresponding transitions occur between the corresponding states in N- and two-vehicle scenarios.

to slow down to avoid collisions, the fact that they do not have a lane preference will prevent them from shifting to the middle lane.

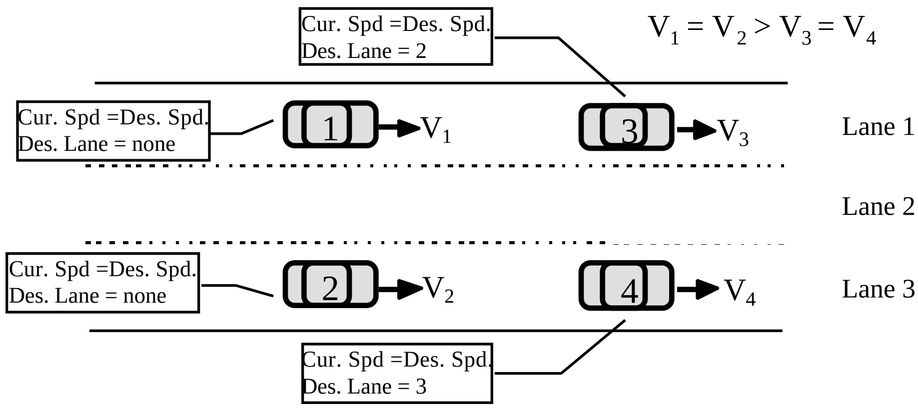


Figure 7.14. Scenario 3: Four vehicles with conflicting desired speeds.

For our analysis of the situation, we only consider vehicles 1 and 2 while defining the environment states since interactions of vehicles 3 and 4 with others are unaffected by any possible actions vehicles 1 and 2 may take. Therefore, by analyzing the interactions between the first two vehicles, we must be able to find a solution to this conflict. The states in which we are interested are listed in Appendix D.3. Initially the environment is in state B2. Since vehicles 1 and 2 are traveling at the same speed and faster than vehicles 3 and 4, the transition to state D1 is automatic (Figure 7.15).

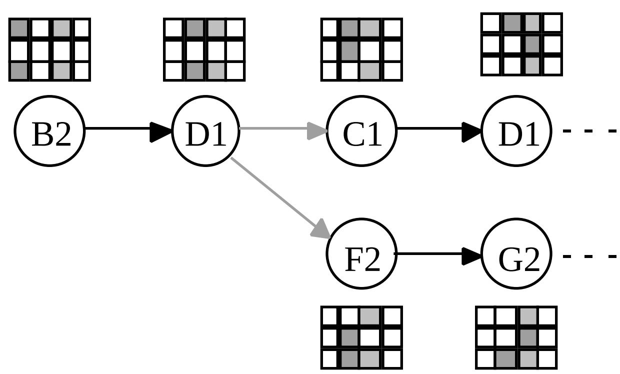
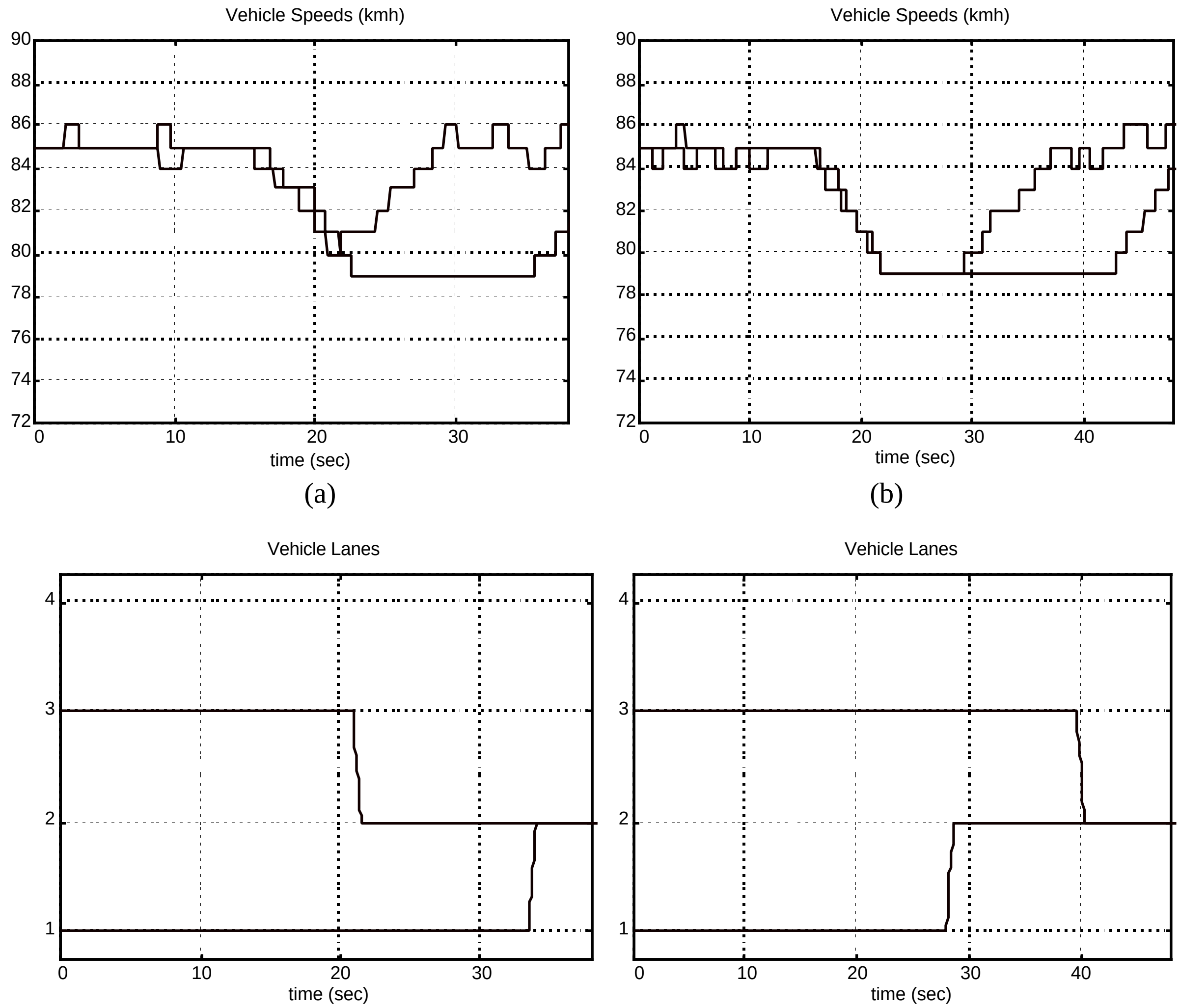


Figure 7.15. Two possible solutions to a situation with 4 vehicles.

At state D1, two transitions that will solve the conflict, but need to be forced, are transitions to form state D1 to states C1 and F2. From these states, the chain will move to goal states C3, E4, or G3 (see Appendix D.3). Then, vehicles 1 and 2 will increase their speed and pass other vehicles. Since the desired lane for vehicles 1 and 2 are not set, another method needs to be found to force the environment state to switch to one of the chains leading to a goal state.

The solution to the problem is what we defined as speed flag in Section 5.3.2. Under current circumstances, vehicles 1 and 2 both fire their idle actions repetitively at state D1. With the addition of the speed flag, both vehicles decide to shift to another lane after a predefined time interval. Since shifting to the middle lane becomes the only possible lateral action, one of the states C1 or F2 is reached.

If both vehicles start filling their memory vectors with conflicting actions, then the pinch modules will start returning a penalty for these actions. At this point one of the vehicles will decide against shifting to the middle lane, and the other will be the first to shift as described in Section 5.2.2, therefore changing environment state to either C1 or F2. Figure 7.16 includes the *mpeg* movies of the two separate runs for this highway scenario where vehicle 1 and 2 mutually find a solution through their pinch modules. The time interval for the speed flag is chosen as 6 sec. In the first simulation (Figures 7.16a, c, e), vehicle 1 shifts to the middle lane immediately after the speed flag is set, returning a penalty response to lateral action SiL. In the second example (Figures 7.16b, d, f), although the flag is again set after 6 seconds, vehicles 1 and 2 need some time to ‘negatiate’ the shift via their pinch modules. Since the last deviation from the desired speed of 85kmh is approximately the same for both vehicles (Figure 7.16b), they attempt to shift to the middle lane at approximately the same time, leading to a pinch condition. Therefore, reaching the solution takes slightly more time in the second simulation.



(a) (b) (c) (d)

ch7m1.mpg
(0.065 Mb; every frame represents one fifth of a second)
(e)

ch7m2.mpg
(0.177 Mb; every frame represents one fifth of a second)
(f)

Figure 7.16. Speeds (a, b) and positions (c, d) of vehicles 1 and 2 for scenario 3: (left) vehicle 1 shifts first; (right) vehicle 2 shifts first. *Mpeg* movies of the simulations are accessible via icons (e, f).

Figure 7.15 shows only two of the possible transitions from state D1; there are two other possibilities ¾ switching to state B3 or to state C4. These two possibilities become a solution to the problem at hand when both vehicles cannot shift lanes because of their pinch module's respective penalty outputs. The probability of occurrence for this situation is very small; however, it is not zero. If this is the case, the vehicles will not be able to change to the middle lane defined now as the desired lane by the speed flag interference. Then, the lane flag will be set (see Section 5.3.1) after a time period, forcing the vehicles to slow down. Since the desired speeds set by the lane flag are different, the symmetry will be broken, and one of the goal states will be reached via state B3 or C4.

7.5 Discussion

This chapter introduced the treatment of path decisions for autonomous vehicles as interacting automata, and then at a higher level, as interacting vehicles. In each intelligent vehicle, lateral and longitudinal automata create an interacting system which can be visualized as an automata game. In our application, this game is usually 'disjoint' *i.e.*, the mixed strategies of automata do not depend on other's strategy. Therefore, the optimal action pair is reached. For the case where the longitudinal automaton's actions affect the lateral automaton's environment, the interacting system again reaches the equilibrium point which happens to be the optimal solution for a given environment.

The visualization of multiple automata interactions as games is based on the assumption that the physical (and therefore, the automata) environment is "stationary." Assuming that the physical environment does not change for a relatively long period of time, and provided that the learning rates for all automata are fast enough, it is possible to assign a stationary automata environment to any physical environment consisting of locations and internal parameters of vehicles in which we are interested for a given situation. Changes of the physical environment can then be treated as a Markov chain whose transitions are direct results of the automata environment evaluated for each state of the physical environment. Situations involving multiple vehicles can be analyzed by investigating the transition diagrams resulting from a specific situation, as described in Section 7.4.

Despite the fact that we used automata environments to evaluate a physical environment's state transitions, the method can also be used for other decision mechanisms. As long as we can define a formal way of describing the decision/control procedure resulting from a specific environment condition, the transition diagrams for physical environment states can be created.

The treatment of the highway scenarios is based on 3x3 or 3x4 location matrices. Larger matrices can also be used, however those may be redundant for our analysis purposes of the situations given in Section 7.4. For a larger number of vehicles or different sensor definitions, larger location matrices may be needed, but again, it is possible to represent multi-vehicle scenarios as a simultaneous two-vehicle interactions as described in Section 7.4.2.