

# A Reinforcement Learning System for Generating Train Marshaling Plan of Freight Cars Based on the Processing Time Considering Group Layout

Yoichi Hirashima

**Abstract**—This paper proposes a new marshaling method for assembling an outgoing train considering the layout of incoming freight cars. In the proposed method, each set of freight cars that have the same destination make a group, and the desirable group layout constitutes the best outgoing train. The incoming freight cars are classified into several “sub-tracks” searching better assignment in order to reduce the total processing time. Classifications and marshaling plans based on the processing time are obtained by a reinforcement learning system. In order to evaluate the processing time, the total transfer distance of a locomotive and the total movement counts of freight cars are simultaneously considered. Moreover, by grouping of freight cars, candidates of classification of incoming train and the desired arrangement of the outgoing train is extended. This feature is considered in the learning algorithm, so that the total processing time is reduced. Then, the classification of incoming cars, the order of movements of freight cars, the position for each removed car, the layout of groups in a train, the arrangement of cars in a group and the number of cars to be moved are simultaneously optimized to achieve minimization of the total processing time for obtaining the desired arrangement of freight cars for an outgoing train. Initially, freight cars are located in a freight yard by the random layout, and they are moved and lined into a main track in a certain desired order in order to assemble an outgoing train. Learning algorithm in the proposed method is based on the Q-Learning, and, after adequate autonomous learning, the optimum marshaling plan can be obtained by selecting a series of movements of freight cars each of which has the best evaluation.

**Index Terms**—Scheduling, Freight train, Marshaling, Q-Learning, Container Transfer Problem

## I. INTRODUCTION

**T**RAIN marshaling operation at freight yard is required to joint several rail transports, or different modes of transportation including rail. Transporting goods are carried in containers, each of which is loaded on a freight car. A freight train is consists of several freight cars, and each car has its own destination. Thus, the train driven by a locomotive travels several destinations decoupling corresponding freight cars at each freight station. In addition, since freight trains can transport goods only between railway stations, modal shifts are required for area that has no railway. In intermodal transports including rail, containers carried into the station are loaded on freight cars and located at the freight yard in the arriving order. The initial layout of freight cars in the yard is determined by considering both arrangement of incoming train and the arriving order of the containers. For efficient shift in assembling outgoing train, freight cars must be rearranged before coupling to the freight train. In

general, the rearrangement process is conducted in a freight yard that consists of a main-track and several sub-tracks. Freight cars are initially placed on sub-tracks, rearranged, and lined into the main track. This series of operation is called marshaling, and several methods to solve the marshaling problem have been proposed [1], [2]. Also, many similar problems are treated by mathematical programming and genetic algorithm[3], [4], [5], [6], and some analyses are conducted for computational complexities [6], [7]. However, these methods do not consider the processing time for each transfer movement of freight car that is moved by a locomotive.

In this paper a new method for generating marshaling plan of freight cars in a train is proposed. In the proposed method, the incoming freight cars are classified into several sub-tracks searching better assignment in order to reduce the total processing time. Then, classifications and marshaling plans based on the processing time are obtained by a reinforcement learning system. A movement of a freight car consists of 4 elements: 1. moving a locomotive to the car to be transferred, 2. coupling cars with the locomotive, 3. transferring cars to their new position by the locomotive, and 4. decoupling the cars from the locomotive. The processing times for elements 1. and 3. are determined by the transfer distance of the locomotive, the weight of the train, and the performance of the locomotive. The total processing time for elements 1. and 3. is determined by the number of movements of freight cars. Thus, the transfer distance of the locomotive and the number of movements of freight cars are simultaneously considered, and used to evaluate and minimize the processing time of marshaling for obtaining the desired layout of freight cars for an outgoing train. The total processing time of marshaling is considered by using a weighted cost of a transfer distance of the locomotive and the number of movements of freight cars. Then, the order of movements of freight cars, the position for each removed car, the arrangement of cars in a train and the number of cars to be moved are simultaneously optimized to achieve minimization of the total processing time. The *original* desired arrangement of freight cars in the main track is derived based on the destination of freight cars. In the proposed method, by grouping freight cars that have the same destination, several desirable positions for each freight car in a group are generated from the original one, and the optimal group-layout that can achieve the smallest processing time of marshaling is obtained by autonomous learning. Simultaneously, the desirable classification of incoming cars as well as, the optimal sequence of car-movements, the number of freight cars that can achieve the desired layout of outgoing train is obtained by autonomous learning. Also,

Faculty of Information Science and Technology, Osaka Institute of Technology, 1-79-1, Kita-yama, Hirakata City, Osaka, 573-0196, Japan. Tel/Fax: +86-72-866-5187 Email: hirash-y@is.oit.ac.jp

the feature is considered in the learning algorithm, so that, at each arrangement on sub-track, the corresponding evaluation value reflects the smallest processing time of marshaling to achieve the best layout on the main track. The learning algorithm is derived based on the Q-Learning [8], which is known as one of the well established realization algorithm of the reinforcement learning.

In the learning algorithm, the state is defined by using a layout of freight cars, the car to be moved, the number of cars to be moved, and the destination of the removed car. An evaluation value called Q-value is assigned to each state, and the evaluation value is calculated by several update rules based on the Q-Learning algorithm. In the learning process, a Q-value in a certain update rule is referred from another update rule, in accordance with the state transition. Then, the Q-value is discounted according to the transfer distance of the locomotive. Consequently, Q-values at each state represent the total processing time of marshaling to achieve the best layout from the state. Moreover, in the proposed method, only referred Q-values are stored by using table look-up technique, and the table is dynamically constructed by binary tree in order to obtain the best solution with feasible memory space. In order to show effectiveness of the proposed method, computer simulations are conducted for several methods.

## II. PROBLEM DESCRIPTION

A freight yard is assumed to have 1 main track and  $m$  sub-tracks. Define  $k$  as the number of freight cars carried in and placed on the sub-tracks. Then, they are moved to the main track by the desirable order based on their destination. In the yard, a locomotive moves freight cars from sub-track to sub-track or from sub-track to main track. The movement of freight cars from sub-track to sub-track is called removal, and the car-movement from sub-track to main track is called rearrangement. For simplicity, the maximum number of freight cars that each sub-track can have is assumed to be  $n$ , the  $i$ th car is recognized by a unique symbol  $c_i$  ( $i = 1, \dots, k$ ). Fig.1 shows the outline of freight yard in the case  $k = 30, m = n = 6$ . In the figure, track  $T_m$  denotes the main track, and other tracks [1], [2], [3], [4], [5], [6] are sub-tracks. The main track is linked with sub-tracks by a joint track, which is used for moving cars between sub-tracks, or for moving them from a sub-track to the main track. Fig.1-(a) depicts an example of classification and Fig.1-(b) is an example of rearrangement. In Fig.1-(a), after cars  $c_1$  through  $c_{12}$  and  $c_{20}$  through  $c_{26}$  are classified into sub-tracks [1] [2] [3],  $c_{19}$  is placed on the sub-track [6]. Then,  $c_{26}$  through  $c_{30}$  carried by trucks are placed on sub-track [6] by the arriving order. In Fig.1-(b), freight cars are moved from sub-tracks, and lined in the main track by the descending order, that is, rearrangement starts with  $c_{30}$  and finishes with  $c_1$ . When the locomotive  $L$  moves a certain car, other cars locating between the locomotive and the car to be moved must be removed to other sub-tracks. This operation is called removal. Then, if  $k \leq n \cdot m - (n - 1)$  is satisfied for keeping adequate space to conduct removal process, every car can be rearranged to the main track.

In each sub-track, positions of cars are defined by  $n$  rows. Every position has unique position number represented by  $m \cdot n$  integers, and the position number for cars at the main

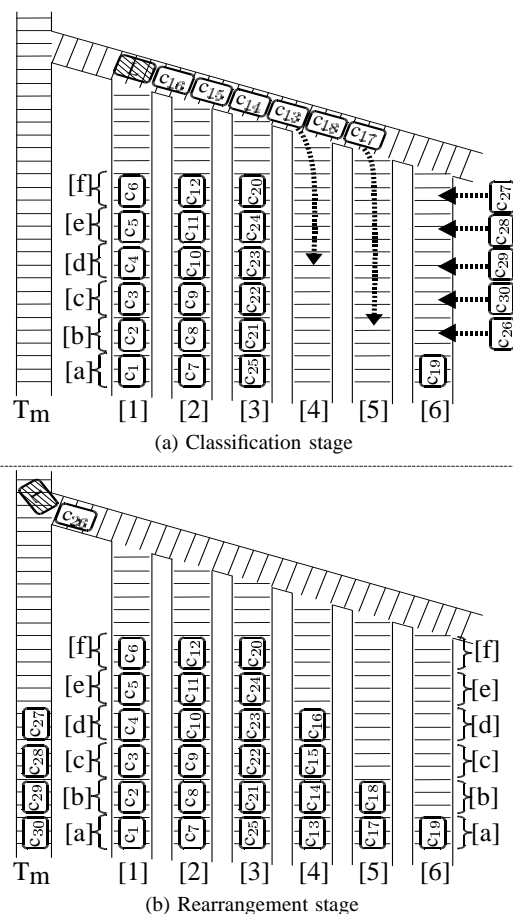


Fig. 1. Freight yard

track is 0. Fig.2 shows an example of position index for  $k = 30, m = n = 6$  and the layout of cars for fig.1-(b).

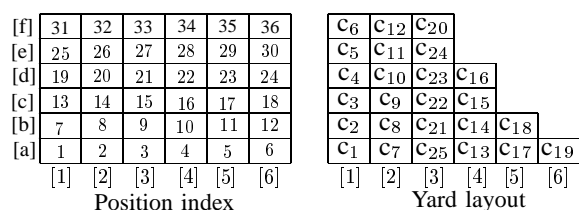


Fig. 2. Example of position index and yard state

In Fig.2, the position “[a][1]” that is located at row “[a]” in the sub-track “[1]” has the position number 1, and the position “[f][6]” has the position number 36. For unified representation of layout of cars in sub-tracks, the first car is placed at the row “[a]” in every track, and a newly placed car is coupled with the adjacent freight car. In the figure, in order to rearrange  $c_{25}$ , cars  $c_{24}, c_{23}, c_{22}, c_{21}$  and  $c_{20}$  have to be removed to other sub-tracks. Then, since  $k \leq n \cdot m - (n - 1)$  is satisfied,  $c_{25}$  can be moved even when all the other cars are placed in sub-tracks.

In the freight yard, define  $x_i (1 \leq x_i \leq n \cdot m, i = 1, \dots, k)$  as the position number of the car  $c_i$ , and  $s = [x_1, \dots, x_k]$  as the state vector of the sub-tracks. For example, in Fig.2, the state is represented by  $s = [1, 7, 13, 19, 25, 31, 2, 8, 14, 20, 26, 32, 4, 10, 16, 22, 5, 11, 6, 33, 9, 15, 21, 27, 3, 0, 0, 0, 0, 0]$ . A trial of the rearrange process starts with the

initial layout, rearranging freight cars according to the desirable layout in the main track, and finishes when all the cars are rearranged to the main track.

### III. DESIRED LAYOUT IN THE MAIN TRACK

In the main track, freight cars that have the same destination are placed at the neighboring positions. In this case, removal operations of these cars are not required at the destination regardless of layouts of these cars. In order to consider this feature in the desired layout in the main track, a group is organized by cars that have the same destination, and these cars can be placed at any positions in the group. Then, for each destination, make a corresponding group, and the order of groups lined in the main track is predetermined by destinations. This feature yields several desirable layouts in the main track.

Fig.3 depicts examples of desirable layouts of cars and the desired layout of groups in the main track. In the figure, freight cars  $c_1, \dots, c_6$  to the destination<sub>1</sub> make group<sub>1</sub>,  $c_7, \dots, c_{18}$  to the destination<sub>2</sub> make group<sub>2</sub>,  $c_{19}$  and  $\dots, c_{25}$  to the destination<sub>3</sub> make group<sub>3</sub>. Groups<sub>1,2,3</sub> are lined by ascending order in the main track, which make a desirable layout. Also, in the figure, examples of layout in group<sub>1</sub> are in the dashed square.

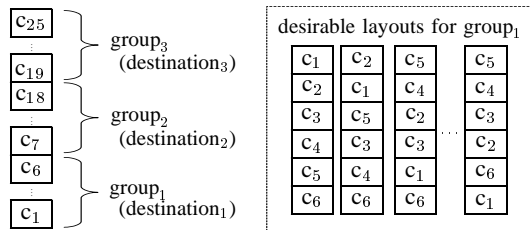


Fig. 3. Example of groups

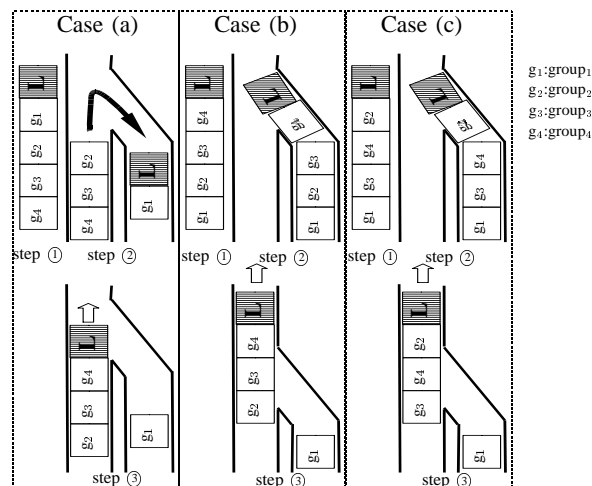


Fig. 4. Group layouts

The layout of groups lined by the reverse order do not yield additional removal actions at the destination of each group. Thus, in the proposed method, the layout lined groups by the reverse order and the layout lined by ascending order from both ends of the train are regarded as desired layouts. Fig.4 depicts examples of material handling operation for

extended layout of groups at the destination of group<sub>1</sub>. In the figure, step ① shows the layout of the incoming train. In case (a), cars in group<sub>1</sub> are separated at the main track, and moved to a sub-track by the locomotive L at step ②. In cases (b),(c), cars in group<sub>1</sub> are carried in a sub-track, and group<sub>1</sub> is separated at the sub-track. In the cases, group<sub>1</sub> can be located without any removal actions for cars in each group. Thus, these layouts of groups are regarded as candidate for desired one in the learning process of the proposed method.

### IV. DIRECT REARRANGEMENT

When there exists a rearranging car that has no car to be removed on it, its rearrangement precedes any removals. In the case that several cars can be rearranged without a removal, rearrangements are repeated until all the candidates for rearrangement requires at least one removal. If several candidates for rearrangement require no removal, the order of selection is random, because any orders satisfy the desirable layout of groups in the main track. In this case, the arrangement of cars in sub-tracks obtained after rearrangements is unique, so that the movement count of cars has no correlation with rearrangement order of cars that require no removal. This operation is called direct rearrangement. When a car in a certain sub-track can be rearrange directly to the main track and when several cars located adjacent positions in the same sub-track satisfy the layout of group in the main track, they are coupled and applied direct rearrangement.

Fig.5 shows an example of arrangement in sub-tracks existing candidates for rearranging cars that require no removal. At the top of figure, from the left side, a desired layout of cars and groups, the initial layout of cars in sub-tracks, and the position index in sub-tracks are depicted for  $m = n = 4, k = 9$ .  $c_1, c_2, c_3, c_4$  are in group<sub>1</sub>,  $c_5, c_6, c_7, c_8$  are in group<sub>2</sub>, and group<sub>1</sub> must be rearranged first to the main track. In each group, any layouts of cars can be acceptable. In both cases,  $c_2$  in step 1 and  $c_3$  in step 3 are applied the direct rearrangement. Also, in step 4, 3 cars  $c_1, c_4, c_5$  located adjacent positions are coupled with each other and moved to the main track by a direct rearrangement operation. In addition, at step 5 in case 2, cars in group<sub>2</sub> and group<sub>3</sub> are moved by a direct rearrangement, since the positions of  $c_7, c_8, c_6, c_9$  are satisfied the desired layout of groups in the main track. Whereas, at step 5, case 1 includes 2 direct rearrangements separately for group<sub>2</sub> and group<sub>3</sub>.

### V. MARSHALING PROCESS

A marshaling process consists of following 7 operations:

- (I) selection of a layout of groups in the main track,
- (II) classification of the incoming freight cars into sub-tracks,
- (III) direct rearrangement,
- (IV) selection of a freight car to be rearranged into the main track,
- (V) selection of a removal destinations of the cars in front of the car selected in (I),
- (VI) selection of the number of cars to be moved,
- (VII) removal of the cars to the selected sub-track,

These operations are repeated until one of desirable layouts is achieved in the main track, and a series of operations from the initial state to the desirable layout is defined as a trial.

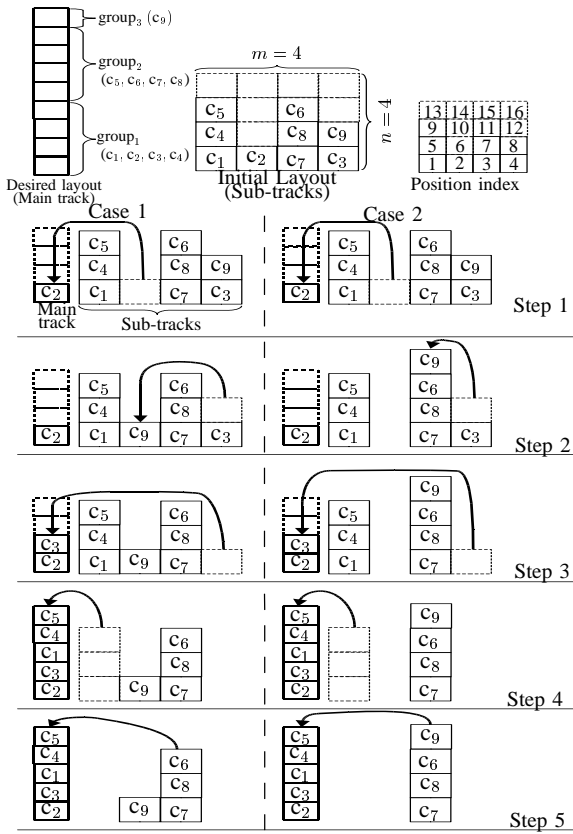


Fig. 5. Direct rearrangements

Now, define  $h$  as the number of candidates of the desired layout of groups. Each candidate in operation (I) is represented by  $u_{j_1}$  ( $1 \leq j_1 \leq h$ ).

In the operation (II), a destination for each car is determined from the tail of the train. The destination is defined as  $T_C$ , and candidates of  $T_C$  are defined as  $u_{j_2}$  ( $h+1 \leq j_2 \leq h+m$ ).  $u_{j_2}$  are sub-tracks each of which have the car belonging to the group moved to the main track before the car of the tail of the incoming train. When there is no such sub-track,  $T_C$  is selected from  $m$  sub-tracks. Then, the number of groups classified to  $T_C$  is determined. Candidates are groups that satisfy the movement-order selected in (I), and are defined as  $u_{j_3}$  ( $h+m+1 \leq j_3 \leq h+m+v$ ), where  $v$  is the number of the candidates.

In the operation (IV), each group has the predetermined position in the main track. The car to be rearranged is defined as  $c_T$ , and candidates of  $c_T$  can be determined by excluding freight cars that have already rearranged to the main track. These candidates must belong to the same group.

Also, define  $r$  as the number of groups,  $g_l$  as the number of freight cars in group  $l$  ( $1 \leq l \leq r$ ), and  $u_{j_4}$  ( $h+m+v+1 \leq j_4 \leq h+m+v+g_l$ ) as candidates of  $c_T$ .

In the operation (V), the removal destination of cars located on the car to be rearranged is defined as  $c_M$ . Then, defining  $u_{j_5}$  ( $h+m+v+g_l+1 \leq j_5 \leq h+m+v+g_l+m-1$ ) as candidates of  $c_M$ , excluding the sub-track that has the car to be removed, and the number of candidates is  $m-1$ .

In the operation (VI), defining  $p$  as the number of removal cars required to rearrange  $c_T$ , and defining  $q$  as the number of removal cars that can be located the sub-track selected in the operation (V), the candidate numbers of cars to be moved are

determined by  $u_{j_6}$  ( $1 \leq u_{j_6} \leq \min\{p, q\}$ ,  $h+2m+v+g_l \leq j_6 \leq h+2m+v+g_l+\min\{p, q\}$ ).

In both cases of Fig.5, the direct rearrangement is conducted for  $c_2$  at step 1, and the selection of  $c_T$  conducted at step 2, candidates are  $u_{h+m+v+1} = [1]$ ,  $u_{h+m+v+2} = [4]$ , that is, sub-tracks where cars in group1 are located at the top.  $u_{h+m+v+3}$ ,  $u_{h+m+v+4}$  are excluded from candidates. Then,  $u_{h+m+v+2} = [4]$  is selected as  $c_T$ . Candidates for the location of  $c_9$  are  $u_{h+m+v+5} = [1]$ ,  $u_{h+m+v+6} = [2]$ ,  $u_{h+m+v+7} = [3]$ , sub-tracks [1],[2], and [3]. In case 1,  $u_{h+m+v+6} = [2]$  is selected as  $c_M$ , and in case 2,  $u_{h+m+v+7} = [3]$  is selected. After direct rearrangements of  $c_3$  at step 3 and  $c_1, c_4, c_5$  at step 4, the marshaling process is finished at step 5 in case 2. Then, total step counts of marshaling process for case 2 is 5, whereas 6 for case 1.

## VI. PROCESSING TIME FOR A MOVEMENT OF LOCOMOTIVE

### A. Transfer distance of locomotive

When a locomotive transfers freight cars, the process of the unit transition is as follows: (E1). starts without freight cars, and reaches to the joint track, (E2) restart in reverse direction to the target car to be moved, (E3). joints them, (E4) pull out them to the joint track, (E5) restart in reverse direction, and transfers them to the indicated location, and (E6) disjoints them from the locomotive. Then, the transfer distance of locomotive in (E1), (E2), (E4) and (E5) is defined as  $D_1$ ,  $D_2$ ,  $D_3$  and  $D_4$ , respectively. Also, define the unit distance of a movement for cars in each sub-track as  $D_{\min_v}$ , the length of joint track between adjacent sub-tracks, or, sub-track and main track as  $D_{\min_h}$ . The location of the locomotive at the end of above process is the start location of the next movement process of the selected car. Also, the initial position of the locomotive is located on the joint track nearest to the main track.

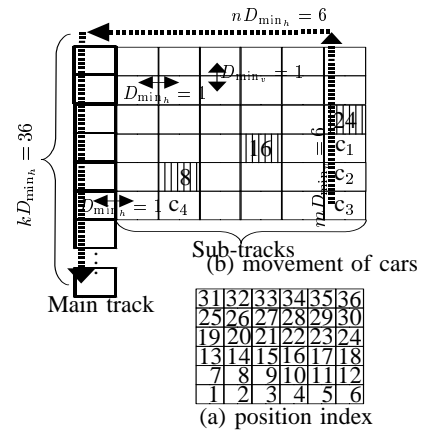


Fig. 6. Calculation of transfer distance

Fig.6 shows an example of transfer distance. In the figure,  $m = n = 6$ ,  $D_{\min_v} = D_{\min_h} = 1$ ,  $k = 18$ , (a) is position index, and (b) depicts movements of locomotive and freight car. Also, the locomotive starts from position 8, the target is located on the position 18, the destination of the target is 4, and the number of cars to be moved is 2. Since the locomotive moves without freight cars from 8 to 24, the transfer distance is  $D_1 + D_2 = 12$  ( $D_1 = 5, D_2 = 7$ ),

whereas it moves from 24 to 16 with 2 freight cars, and the transfer distance is  $D_3 + D_4 = 13$  ( $D_3 = 7, D_4 = 6$ ).

### B. Processing time for the unit transition

In the process of the unit transition, the each time for (E3) and (E6) is assumed to be the constant  $t_E$ .

The processing times for elements (E1), (E2), (E4) and (E5) are determined by the transfer distance of the locomotive  $D_i$  ( $i = 1, 2, 3, 4$ ), the weight of the freight cars  $W$  moved in the process, and the performance of the locomotive. Then, the time each for (E1), (E2), (E4) and (E5) is assumed to be obtained by the function  $f(D_i, W)$  derived considering dynamics of the locomotive, limitation of the velocity, and control rules. Thus, the processing time for the unit transition  $t_U$  is calculated by  $t_U = t_E + \sum_{i=1}^2 f(D_i, 0) + \sum_{j=4}^5 f(D_j, W)$ . The maximum value of  $t_U$  is define as  $t_{\max}$  and is calculated by

$$t_{\max} = t_E + f(kD_{\min_v}, 0) + f(mD_{\min_h}, 0) + f(mD_{\min_h} + n, W) + f(kD_{\min_v}, W) \quad (1)$$

## VII. LEARNING ALGORITHM

Defining  $G_o$  as the desired layout selected among  $u_{j_1}$ ,  $Q_1(G_o)$  is updated by the following rule when one of desired layout is achieved in the main track:

$$Q_1(G_o) \leftarrow \max \left\{ \begin{array}{l} Q_1(G_o), \\ (1 - \alpha)Q_1(G_o) + \alpha R \prod_{i=1}^l \gamma_i \end{array} \right\} \quad (2)$$

where  $l$  denotes the total movement counts required to achieve the desired layout,  $\alpha$  is learning rate,  $\gamma$  is discount factor calculated for each movement,  $R$  is reward that is given only when one of desired layout is achieved in the main track.

Define  $s(t)$  as the state at time  $t$ ,  $T_c$  as the sub-track selected as the destination for the removed car,  $p_c$  as the number of classified groups,  $q_M$  as the movement counts of freight cars by direct rearrangement, and  $s'$  as the state that follows  $s$ . In the classification stage,  $Q_2, Q_3$  are defined as evaluation values for  $(s_1, u_{j_2}), (s_2, u_{j_3})$  respectively, where  $s_1 = [s, G_o], s_2 = [s_1, T_c]$ .  $Q_2(s_1, T_c)$  and  $Q_3(s_2, p_c)$  are updated by following rules:

$$Q_2(s_1, T_c) \leftarrow \max_{u_{j_3}} Q_3(s_1, u_{j_3}) \quad (3)$$

$$Q_3(s_2, p_c) \leftarrow (1 - \alpha)Q_3(s_2, p_c) + \alpha V_1$$

$$V_1 = \begin{cases} R \prod_{i=1}^l \gamma_i \\ \text{(all cars assigned)} \\ \gamma \max_{u_{j_2}} Q_2(s_1, u_{j_2}) \\ \text{(otherwise)} \end{cases} \quad (4)$$

In the rearrangement stage, define  $q$  as the number of direct movements conducted sequentially,  $p_M$  as the number of cars moved.  $Q_4, Q_5$  and  $Q_6$  are defined as evaluation values for  $(s_1, u_{j_4}), (s_3, u_{j_5}), (s_4, u_{j_6})$  respectively, where  $s_3 = s_1, s_4 = [s_3, c_T], s_5 = [s_4, c_M]$ .  $Q_4(s_3, c_T), Q_5(s_4, c_M)$  and

$Q_6(s_5, p_M)$  are updated by following rules:

$$Q_4(s_3, c_T) \leftarrow \max_{u_{j_5}} Q_5(s_4, u_{j_5}), \quad (5)$$

$$Q_5(s_4, c_M) \leftarrow \max_{u_{j_6}} Q_6(s_5, u_{j_6}), \quad (6)$$

$$Q_6(s_5, p_M) \leftarrow \begin{cases} (1 - \alpha)Q_6(s_5, p_M) + \alpha \left[ R + V_2 \prod_{i=1}^{q+1} \gamma_i \right] \\ \text{(u is a rearrangement)} \\ (1 - \alpha)Q_6(s_5, p_M) + \alpha [R + \gamma V_3] \\ \text{(u is a removal)} \end{cases} \quad (7)$$

$$V_2 = \max_{u_{j_4}} Q_4(s'_3, u_{j_4}), V_3 = \max_{u_{j_5}} Q_5(s'_4, u_{j_5})$$

where  $\alpha$  is the learning rate,  $R$  is the reward that is given when one of desirable layout is achieved, and  $\gamma$  is the discount factor that is used to reflect the processing time of the marshaling and calculated by the following equation.

$$\gamma = \delta \frac{t_{\max} - \beta t_U}{t_{\max}}, 0 < \beta < 1, 0 < \delta < 1 \quad (8)$$

Propagating Q-values by using eqs.(5)-(8), Q-values are discounted according to the processing time of marshaling. In other words, by selecting the removal destination that has the largest Q-value, the processing time of the marshaling can be reduced.

In the learning stages, each  $u_j$  ( $1 \leq j \leq h + 2m + v + g_l + \min\{p, q\}$ ) is selected by the soft-max action selection method[9]. Probability  $P$  for selection of each candidate is calculated by

$$\tilde{Q}_i(s_{i-1}, u_{j_i}) = \frac{Q_i(s_{i-1}, u_{j_i}) - \min_u Q_i(s_{i-1}, u_{j_i})}{\max_u Q_i(s_{i-1}, u_{j_i}) - \min_u Q_i(s_{i-1}, u_{j_i})} \quad (9)$$

$$P(s_{i-1}, u_{j_i}) = \frac{\exp(\tilde{Q}_i(s_{i-1}, u_{j_i})/\xi)}{\sum_{u \in u_{j_i}} \exp(\tilde{Q}_i(s_{i-1}, u)/\xi)}, \quad (10)$$

$$(i = 2, 3, 4, 5, 6).$$

$$P(u_{j_1}) = \frac{\exp(Q_1(u_{j_1})/\xi)}{\sum_{u \in u_{j_1}} \exp(Q_1(u)/\xi)}$$

## VIII. COMPUTER SIMULATIONS

Computer simulations are conducted for  $m = 12, n = 6, k = 36$  and learning performances of following 3 methods are compared:

- (A) proposed method that evaluates the processing time of the marshaling operation, considering the layout of groups and classification of incoming cars,
- (B) a method that evaluates the processing time considering the layout of groups, and the classification is fixed[10],
- (C) a method that evaluates the processing time, has the fixed layout of groups, and has the fixed classification[10].

The initial arrangement of incoming train is described in Fig.7. The original rearrangement order of groups is group<sub>1</sub>, group<sub>2</sub>, group<sub>3</sub>, group<sub>4</sub>. Cars  $c_1, \dots, c_9$  are in

TABLE I  
TOTAL PROCESSING TIME

| methods    | processing time (sec.) |          |          |
|------------|------------------------|----------|----------|
|            | best                   | average  | worst    |
| method (A) | 5876.332               | 6157.908 | 6294.066 |
| method (B) | 6437.393               | 6785.930 | 6919.313 |
| method (C) | 7857.711               | 7973.113 | 8003.003 |

group<sub>1</sub>, c<sub>10</sub>, ..., c<sub>18</sub> are in group<sub>2</sub>, c<sub>19</sub>, ..., c<sub>27</sub> are in group<sub>3</sub>, and c<sub>28</sub>, ..., c<sub>36</sub> are in group<sub>4</sub>. Other parameters are set as  $\alpha = 0.9, \beta = 0.2, \delta = 0.9, R = 1.0, \eta = 0.95, \xi_1 = 0.1, \xi_2 = 0.05$ . Method (C) accepts only the original rearrangement order of groups, whereas other methods consider extended layout of groups. In methods (B),(C), the classification generates the fixed layout depicted in Fig.8.

The locomotive assumed to accelerate and decelerate the train with the constant force  $100 \times 10^3 \text{N}$ , and to be  $100 \times 10^3 \text{kg}$  in weight. Also, all the freight cars have the same weight,  $10 \times 10^3 \text{kg}$ . The locomotive and freight cars assumed to have the same length, and  $D_{\min_v} = D_{\min_h} = 20 \text{m}$ . The velocity of the locomotive is limited to no more than  $10 \text{m/s}$ . Then, the locomotive accelerates the train until the velocity reaches  $10 \text{m/s}$ , keeps the velocity, and decelerates until the train stops within the indicated distance. When the velocity does not reach  $10 \text{m/s}$  at the half way point, the locomotive starts to decelerate immediately.

The results are shown in Fig.9. In the figure, horizontal axis expresses the number of trials and the vertical axis expresses the minimum processing time to achieve a desirable layout found in the past trials. Each result is averaged over 20 independent simulations. In Fig.9, the learning performance of method (A) is better than that of method (B), because solutions derived by method (A) considers the classification of groups effectively for reducing the total processing time. Since the group layout for the outbound train and classification are fixed, method (C) is not effective to reduce the total processing time as compared to methods (A),(B). For method (A), since no freight cars are carried in the yard by trucks, the solution is so simple that the best marshaling plan includes no removals. In case the result of classification requires removals in order to achieve the desired layout of groups in the main track, method (A) continues to derive the marshaling plan that can reduce the total processing time is improved as shown by method (B). Total transfer distances of the locomotive at  $1.5 \times 10^6$ th trial are described in table.I for each method.

|                 |                 |                 |                 |                 |                 |                 |                 |                 |  |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|--|
| tail            |                 |                 |                 |                 |                 |                 |                 |                 |  |
| c <sub>36</sub> | c <sub>35</sub> | c <sub>7</sub>  | c <sub>33</sub> | c <sub>22</sub> | c <sub>29</sub> | c <sub>1</sub>  | c <sub>30</sub> | c <sub>25</sub> |  |
| c <sub>32</sub> | c <sub>21</sub> | c <sub>20</sub> | c <sub>5</sub>  | c <sub>31</sub> | c <sub>24</sub> | c <sub>23</sub> | c <sub>27</sub> | c <sub>26</sub> |  |
| c <sub>4</sub>  | c <sub>17</sub> | c <sub>18</sub> | c <sub>19</sub> | c <sub>16</sub> | c <sub>14</sub> | c <sub>15</sub> | c <sub>28</sub> | c <sub>11</sub> |  |
| c <sub>3</sub>  | c <sub>2</sub>  | c <sub>10</sub> | c <sub>6</sub>  | c <sub>13</sub> | c <sub>9</sub>  | c <sub>8</sub>  | c <sub>34</sub> | c <sub>12</sub> |  |
| head            |                 |                 |                 |                 |                 |                 |                 |                 |  |

Fig. 7. Arrangement of incoming train

|                 |                 |                 |                 |                 |                 |                 |                 |                 |                 |                 |                 |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| c <sub>10</sub> | c <sub>9</sub>  | c <sub>12</sub> | c <sub>15</sub> | c <sub>19</sub> | c <sub>4</sub>  | c <sub>20</sub> | c <sub>24</sub> | c <sub>26</sub> | c <sub>1</sub>  | c <sub>33</sub> | c <sub>36</sub> |
| c <sub>2</sub>  | c <sub>13</sub> | c <sub>34</sub> | c <sub>28</sub> | c <sub>16</sub> | c <sub>17</sub> | c <sub>21</sub> | c <sub>31</sub> | c <sub>27</sub> | c <sub>30</sub> | c <sub>22</sub> | c <sub>35</sub> |
| c <sub>3</sub>  | c <sub>6</sub>  | c <sub>8</sub>  | c <sub>11</sub> | c <sub>14</sub> | c <sub>18</sub> | c <sub>32</sub> | c <sub>5</sub>  | c <sub>23</sub> | c <sub>25</sub> | c <sub>29</sub> | c <sub>7</sub>  |

Fig. 8. Classification for methods (B), (C)

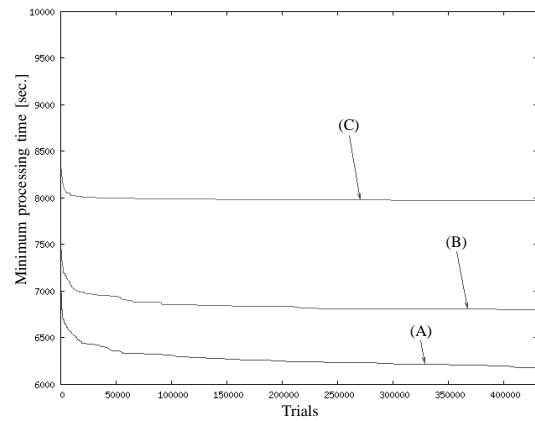


Fig. 9. Comparison of learning performances

## IX. CONCLUSIONS

A new scheduling method has been proposed in order to rearrange and line cars in the desirable order onto the main track considering classifications for incoming freight cars. The learning algorithm of the proposed method is derived based on the reinforcement learning, considering the total processing time of marshaling. In order to reduce the total processing time of marshaling, the proposed method learns the classification of incoming cars and the layout of groups, as well as the arrangement of freight cars in each group, the rearrangement order of cars, the number of cars to be moved and the removal destination of cars, simultaneously. In computer simulations, learning performance of the proposed method has been improved by using normalized evaluation and switching thermo constants in accordance with the progress of learning.

## REFERENCES

- [1] U. Blasum, M. R. Bussieck, W. Hochstättler, C. Moll, H.-H. Scheel, and T. Winter, "Scheduling trams in the morning," *Mathematical Methods of Operations Research*, vol. 49, no. 1, pp. 137–148, 2000.
- [2] R. Jacob, P. Marton, J. Maue, and M. Nunkesser, "Multistage methods for freight train classification," in *Proceedings of 7th Workshop on Algorithmic Approaches for Transportation Modeling, Optimization, and Systems*, 2007, pp. 158–174.
- [3] L. Kroon, R. Lentink, and A. Schrijver, "Shunting of passenger train units: an integrated approach," *Transportation Science*, vol. 42, pp. 436–449, 2008.
- [4] N. TOMII and Z. L. Jian, "Depot shunting scheduling with combining genetic algorithm and pert," in *Proceedings of 7th International Conference on Computer Aided Design, Manufacture and Operation in the Railway and Other Advanced Mass Transit Systems*, 2000, pp. 437–446.
- [5] S. He, R. Song, and S. Chaudhry, "Fuzzy dispatching model and genetic algorithms for railyards operations," *European Journal of Operational Research*, vol. 124, no. 2, pp. 307–331, 2000.
- [6] E. Dahlhaus, F. Manne, M. Miller, and J. Ryan, "Algorithms for combinatorial problems related to train marshalling," in *Proceedings of the 11th Australasian Workshop on Combinatorial algorithms*, 2000, pp. 7–16.
- [7] C. Eggermont, C. A. J. Hurkens, M. Modelski, and G. J. Woeginger, "The hardness of train rearrangements," *Operations Research Letters*, vol. 37, pp. 80–82, 2009.
- [8] C. J. C. H. Watkins and P. Dayan, "Q-learning," *Machine Learning*, vol. 8, pp. 279–292, 1992.
- [9] R. Sutton and A. Barto, *Reinforcement Learning*. MIT Press, 1999.
- [10] Y. Hirashima, "A reinforcement learning for train marshaling based on the processing time considering group layout of freight cars," *Lecture Notes in Engineering and Computer Science: Proceedings of The International MultiConference of Engineers and Computer Scientists 2012, IMECS 2012*, pp. 47–52, 2012.