Proceedings of the International MultiConference of Engineers and Computer Scientists 2011 Vol I,
IMECS 2011, March 16 - 18, 2011, Hong Kong

A New Reinforcement Learning Method for Train Marshaling Based on the Transfer Distance of Locomotive

Yoichi Hirashima

***Abstract*—In this paper a new reinforcement learning system for generating marshaling plan of freight cars in a train is designed. In the proposed method, the total transfer distance of a locomotive is minimized to obtain the desired layout of freight cars for an outbound train. The order of movements of freight cars, the position for each removed car, the layout of cars in a train and the number of cars to be moved are simultaneously optimized to achieve minimization of the total transfer distance of a locomotive. 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 out bound train. A layout and movements of freight cars are used to describe a state of marshaling yard, and the state transitions are defined based on the Markov Decision Process (MDP). Q-Learning is applied to reflect the transfer distance of the locomotive that are used to achieve one of the desired layouts in the main track. After adequate autonomous learning, the optimum schedule can be obtained by selecting a series of movements of freight cars that has the best evaluation.**

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

I. INTRODUCTION

TRAIN 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 disjointing corresponding freight cars at each freight station. In addition, since freight trains can transport goods only between railway stations, modal shifts are required for delivering them to area that has no railway. In intermodal transports from the road and the rail, containers carried into the station are loaded on freight cars in the arriving order. The initial layout of freight cars is thus random. For efficient shift, the desirable layout should be determined considering destination of container. Then, freight cars must be rearranged before jointing 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].

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

In this paper, a new scheduling method is proposed in order to rearrange and line freight cars by the desirable order onto the main track. In the proposed method, the focus is centered on to reduce the total transfer distance of a locomotive required to achieve desirable layout on the main track. The optimal layout of freight cars in the main track is derived based on the destination of freight cars. This yields several desirable layouts of freight cars in the main track, and the optimal layout that can achieve the smallest transfer distance of the locomotive is obtained by autonomous learning. Simultaneously, the optimal sequence of car-movements as well as the number of freight cars that can achieve the desired layout is obtained by autonomous learning. Also, the feature is considered in the learning algorithm, so that, at each arrangement on sub track, an evaluation value represents the smallest transfer distance of the locomotive 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 transfer distance of the locomotive required 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

The yard consist of 1 main track and m sub tracks. Define k as the number of freight cars placed on the sub tracks, and they are carried 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

ISBN: 978-988-18210-3-4
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

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 ith car is recognized by an unique symbol c_i $(i = 1, \cdots, k)$. Fig.1 shows the outline of freight yard in the case $k = 30, m = n = 6$. In the figure,

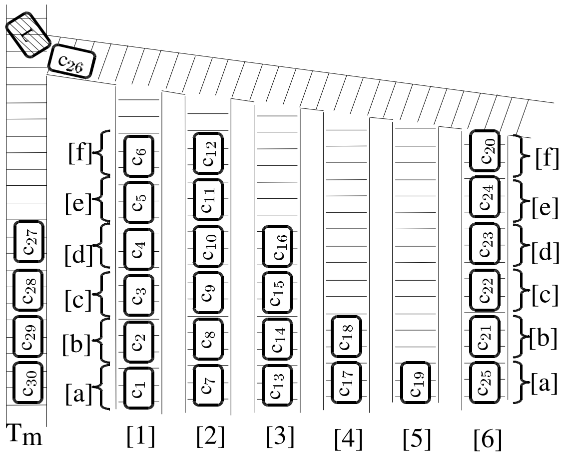

Fig. 1. Freight yard

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. In the figure, 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 main 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.

Position index

	[1]	[2]	[3]	[4]	[5]	[6]
[f]	31	32	33	34	35	36
[e]	25	26	27	28	29	30
[d]	19	20	21	22	23	24
[c]	13	14	15	16	17	18
[b]	7	8	9	10	11	12
[a]	1	2	3	4	5	6

Yard layout

	[1]	[2]	[3]	[4]	[5]	[6]
[f]	c_6	c_{12}				c_{20}
[e]	c_5	c_{11}				c_{24}
[d]	c_4	c_{10}	c_{16}			c_{23}
[c]	c_3	c_9	c_{15}			c_{22}
[b]	c_2	c_8	c_{14}	c_{18}		c_{21}
[a]	c_1	c_7	c_{13}	c_{17}	c_{19}	c_{25}

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 car in sub tracks, cars are placed from the row "[a]" in every track, and newly placed car is jointed 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 $x_i (1 \leq x_i \leq n \cdot m, i = 1, \cdots, k)$ as the position number of the car c_i, and $\boldsymbol{s} = [x_1, \cdots, x_k]$ as the state vector of the sub tracks. For example, in Fig.2, the state is represented by $\boldsymbol{s} = [1, 7, 13, 19, 25, 31, 2, 8, 14, 20, 26, 32, 3, 9, 15, 21, 4, 10, 5, 36, 12, 18, 24, 30, 6, 0, 0, 0, 0, 0]$. A trial of the rearrange process starts with the initial layout, rearranging freight cars according to the desirable layout in the main track, and finishs 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, $\cdots$, c_6 to the destination$_1$ make group$_1$, c_7, $\cdots$, c_{18} to the destination$_2$ make group$_2$, c_{19}, $\cdots$, c_{25} to the destination$_3$ make group$_3$, and c_{26}, $\cdots$, c_{30} to the destination$_4$ make group$_4$. Groups$_{1,2,3,4}$ are lined by ascending order in the main track, which make a desirable layout. In the figure, examples of layout in group$_1$ are in the dashed square.

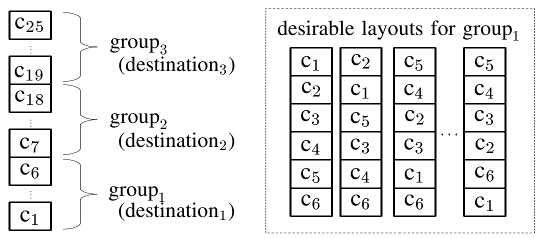

Fig. 3. Example of groups

IV. Direct Rearrangement

When rearranging car that has no car to be removed on it is exist, its rearrangement precede 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 counts of cars has no correlation with rearrangement orders 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

ISBN: 978-988-18210-3-4
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

and when several cars located adjacent positions in the same sub track satisfy the layout of group in main track, they are jointed and applied direct rearrangement.

Fig.4 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 $m = n = 4, k = 9$. $\mathrm{c}_1, \mathrm{c}_2, \mathrm{c}_3, \mathrm{c}_4$ are in group_1, $\mathrm{c}_5, \mathrm{c}_6, \mathrm{c}_7, \mathrm{c}_8$ are in group_2, and group_1 must be rearranged first to the main track. In each group, any layouts of cars can be acceptable. In both cases, c_2 in step1 and c_3 in step3 are applied the direct rearrangement. Also, in step4, 3 cars $\mathrm{c}_1, \mathrm{c}_4, \mathrm{c}_5$ located ajacent positions are jointed and moved to the main track by a direct rearrangement operation. In addition, at step5 in Case2, cars in group_2 and group_3 are moved by a direct rearrangement, since the positions of $\mathrm{c}_7, \mathrm{c}_8, \mathrm{c}_6, \mathrm{c}_9$ are satisfied the desired layout of groups in the main track.

In Case1 of the example, the rearrangement order of cars that require no removal is $\mathrm{c}_1, \mathrm{c}_2, \mathrm{c}_3, \mathrm{c}_4$, and in Case2, the order is $\mathrm{c}_3, \mathrm{c}_2, \mathrm{c}_1, \mathrm{c}_4$. Although 2 cases have different orders of rearrangement, the arrangements of cars in sub tracks and the numbers of movements of cars have no difference.

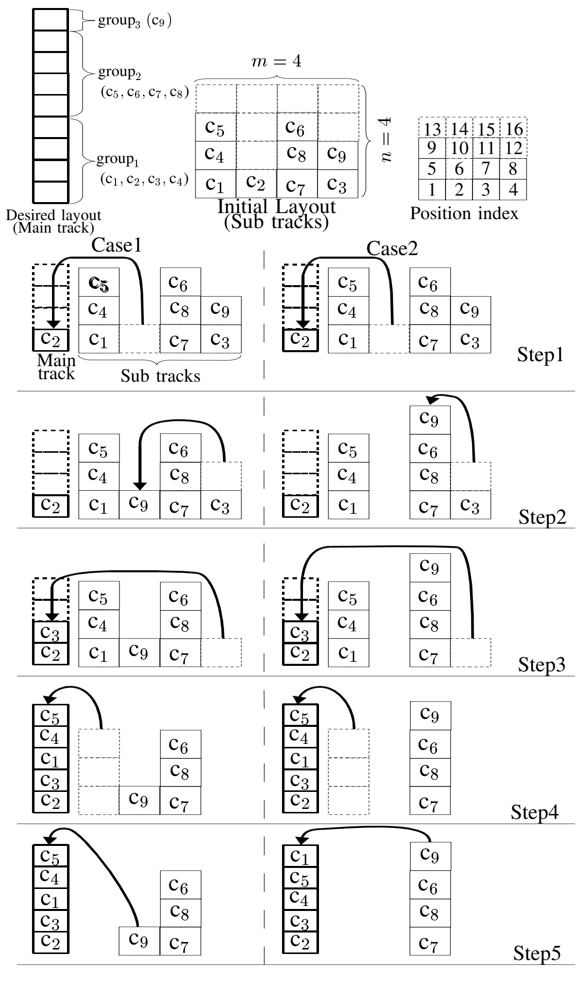

Fig. 4. Direct rearrangements

V. REARRANGEMENT PROCESS

The rearrangement process for cars consists of following 4 operations :

(1) rearrangement for all the cars that can apply the direct rearrangement into the main track,
(2) selection of a freight car to be rearranged into the main track,
(3) selection of a removal destination of cars located between the locomotive and the freight car selected in (2),
(4) removal of the cars to the selected sub track,
(5) rearrangement of the selected car to the main 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.

In the operation (2), 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 that is determined uniquely by the desired layout of groups in the main track and the number of rearranged cars.

Now, define r as the number of groups, g_l as the number of freight cars in $\mathrm{group}_l (1 \leq l \leq r)$, and $u_{j_1} (1 \leq j_1 \leq g_l)$ as candidates of c_T.

In the operation (3), the removal destination of car located on the car to be rearranged is defined as r_M. Then, defining $u_{j_2} (g_l + 1 \leq j_2 \leq g_l + m - 1)$ as candidates of r_M, excluding the sub track that has the car to be removed, and the number of candidates is $m - 1$.

In the operation (4), defining p_{s} as the number of removal cars required to rearrange c_T, and defining p_{d} as the number of removal cars that can be located on the sub track selected in the operation (3), the candidate numbers of cars to be moved are determined by $u_{j_3}, 2m \leq j_3 \leq 2m + \min\{p_{\mathrm{s}}, p_{\mathrm{d}}\} - 1$.

In both cases of Fig.4, the direct rearrangement is conducted for c_2 at step1, and the selection of c_T conducted at step2, candidates are $u_1 = [1], u_2 = [4]$, that is, sub tracks where cars in group_1 are located at the top. u_3, u_4 are excluded from candidates. Then, $u_2 = [4]$ is selected as c_T. Candidates for the location of c_T are $u_5 = [1], u_6 = [2], u_7 = [3]$, sub tracks [1],[2], and [3]. In Case1, $u_6 = [2]$ is selected as c_M, and in Case2, $u_7 = [3]$ is selected. After direct rearrangements of c_3 at step3 and $\mathrm{c}_1, \mathrm{c}_4, \mathrm{c}_5$ at step4, the marshaling process is finished at step5 in Case2, whereas Case1 requires one more step in order to finish the process. Therefore, the layout of cars and groups in the main track, the number of cars to be moved, the location the car to be rearranged and the order of rearrangement affect the total movement counts of cars as well as the total transfer distance of locomotive.

A. *Transfer distance of locomotive*

A locomotive starts without freight cars, directs to the target car to be moved, and locates it at the corresponding destination. The distance D where the locomotive travels from the start location to the destination of the target car is defined as the transfer distance of the locomotive. Then, 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.5

ISBN: 978-988-18210-3-4
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

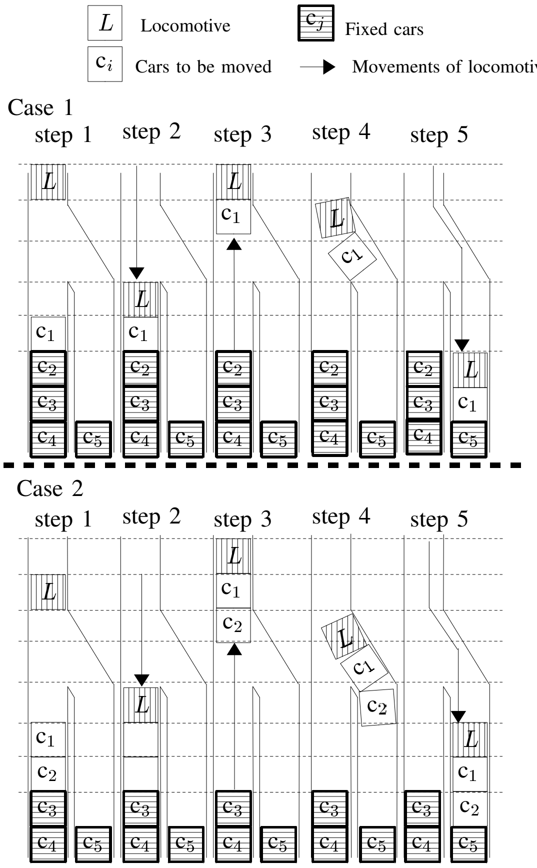

Fig. 5. Transfer distance of locomotive

shows an example of the transfer distance of a locomotive. In the figure, L is the locomotive, $c_1 - c_5$ are freight cars. Cars with hatching are to be moved. In Case1, a freight car is removed to the ajascent sub track, whereas, in Case2, 2 cars are removed. The transfer distances of the locomotive in Cases1,2 are the same from step1 through step2, and from step3 through step 5. While, from step2 through step3, the transfer distance of the locomotive in Case2 is larger than that in Case1. Thus, the number of cars to be moved affects the transfer distance of the locomotive. Also, the transfer distance is affected by the arrangment of cars in sub tracks, the order of cars to be moved, and the destination of moved cars. Thus, the transfer distance must be considered in each selection in the marshaling process in order to reduce the total transfer distance of the locomotive.

Define the unit distance of a movement for cars in each subtrack as $D_{\min_v}$, the length of transition track between adjacent subtracks, or, subtrack and main track as $D_{\min_h}$. Then, the transfer distance of the locomotive is D, and the maximum of D is $D_{\max} = 2(mD_{\min_v} + nD_{\min_h} + kD_{\min_v})$.

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 12, whereas it moves from 24 to 16 with 2 freight cars, and the transfer distance is 13, $D = 25$ and $D_{\max} = 60$.

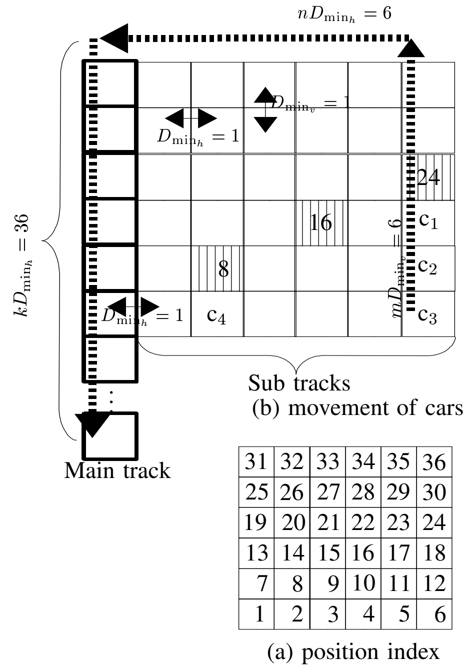

Fig. 6. Calculation of transfer distance

VI. LEARNING ALGORITHM

Define $s(t)$ as the state at time t, $r_{\rm M}$ as the sub track selected as the destination for the removed car, $p_{\rm M}$ as the number of removed cars, q as the movement counts of freight cars by direct rearrangement, and s' as the state that follows s. Also, Q_1, Q_2, Q_3 are defined as evaluation values for (s_1, u_{j_1}), (s_2, u_{j_2}), (s_3, u_{j_3}), respectively, where $s_1 = s, s_2 = [s, c_T]$, $s_3 = [s, c_T, r_{\rm M}]$. $Q_1(s_1, u_{j_1})$, $Q_2(s_2, u_{j_2})$ and $Q_3(s_3, u_{j_3})$ are updated by following rules:

$$Q_1 \quad (s_1, c_T) \leftarrow \max_{u_{j_2}} Q_2(s_2, u_{j_2}), \tag{1}$$

$$Q_2 \quad (s_2, r_{\rm M}) \leftarrow \max_{u_{j_3}} Q_3(s_3, u_{j_3}), \tag{2}$$

$$Q_3 \quad (s_3, p_{\rm M}) \leftarrow \tag{3}$$

$$\begin{cases} (1-\alpha)Q_3(s_3, p_{\rm M}) + \alpha[R + \gamma^{q+1} V_1] \\ \quad \text{(next action is rearrangement)} \\ (1-\alpha)Q_3(s_3, p_{\rm M}) + \alpha[R + \gamma V_2] \\ \quad \text{(next action is removal)} \end{cases}$$

$$V_1 = \max_{u_{j_1}} Q_1(s'_1, u_{j_1}),$$

$$V_2 = \max_{u_{j_2}} Q_2(s'_2, u_{j_2})$$

where α is the learning rate, R is the reward that is given when one of desirable layout is achieved, and γ is the discount factor that is used to reflect the transfer distance of the locomotive and calculated by the following equation.

$$\gamma = \delta \frac{D_{\max} - \beta D}{D_{\max}},\ 0 < \beta < 1, 0 < \delta < 1 \tag{4}$$

Propagating Q-values by using eqs.(1)-(4), Q-values are discounted according to the number of removals of cars. In other words, by selecting the removal destination that has the largest Q-value, the transfer distance of the locomotive can be reduced.

In the learning stages, each $u_j\,(1 \le j \le 2m + \min\{p_{\rm s}, p_{\rm d}\} - 1)$ is selected by the soft-max action selection method[9]. Probability P for selection of each candidate is

ISBN: 978-988-18210-3-4
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

calculated by

$$\tilde{Q}_i(\boldsymbol{s}, u_{j_i}) = \frac{Q_i(\boldsymbol{s}, u_{j_i}) - \min_u Q_i(\boldsymbol{s}, u_{j_i})}{\max_u Q_i(\boldsymbol{s}, u_{j_i}) - \min_u Q_i(\boldsymbol{s}, u_{j_i})} \quad (5)$$

$$P(\boldsymbol{s}_i, u_{j_i}) = \frac{\exp(\tilde{Q}_i(\boldsymbol{s}_{j_i}, u_{j_i})/T)}{\sum_{u \in u_{j_i}} \exp(\tilde{Q}_i(\boldsymbol{s}_i, u)/T)}, \quad (6)$$

$$(i = 1, 2, 3).$$

In the addressed problem, Q_1, Q_2, Q_3 become smaller when the number of discounts becomes larger. Then, for complex problems, the difference between probabilities in candidate selection remain small at the initial state and large at final state before achieving desired layout, even after repetitive learning. In this case, obtained evaluation does not contribute to selections in initial stage of marshaling process, and search movements to reduce the transfer distance of locomotive is spoiled in final stage. To conquer this drawback, Q_1, Q_2, Q_3 are normalized by eq.(5), and the thermo constant T is switched from T_1 to T_2 ($T_1 > T_2$) when the following condition is satisfied:

$$[\text{The count of } Q_i(\boldsymbol{s}_{j_i}, u_{j_i})] > \eta, \quad \text{s.t. } Q_i(\boldsymbol{s}_{j_i}, u_{j_i}) > 0, \quad 0 < \eta \leq [\text{the number of candidates for } u_{j_i}] \quad (7)$$

where η is the threshold to judge the progress of learning.

The proposed learning algorithm can be summarized as follows:

1) Initialize all the Q-values as 0
2) ⓐ When no cars are placed on candidates of c_T, all of them are rearranged
   ⓑ Update corresponding $Q_3(\boldsymbol{s}_3, p_{\mathrm{M}})$ by eq.(3)
   ⓒ Store $\boldsymbol{s}_1, c_T$
3) If no cars are in sub tracks, go to 9, otherwise go to 4
4) ⓐ Determine c_T among the candidates by roulette selection (probabilities are calculated by eq. (6)),
   ⓑ Put reward as R = 0,
   ⓒ Update the corresponding $Q_3(\boldsymbol{s}_3, p_{\mathrm{M}})$ by eq.(3)
   ⓓ Store $\boldsymbol{s}_1, c_T$
5) ⓐ Determine r_{M}(probability for the selection is calculated by eq.(6))
   ⓑ Update corresponding $Q_2(\boldsymbol{s}_2, r_{\mathrm{M}})$ by eq.(2),
   ⓒ store $\boldsymbol{s}_2, r_{\mathrm{M}}$
6) ⓐ Determine p_{M}(probability for the selection is calculated by eq.(6))
   ⓑ Update corresponding $Q_3(\boldsymbol{s}_3, p_{\mathrm{M}})$ by eq.(3)
   ⓒ Store $\boldsymbol{s}_3, p_{\mathrm{M}}$
7) Remove p_{M} cars and place at r_{M}
8) Go to 2
9) Receive the reward R, update $Q_1(\boldsymbol{s}_1, c_T)$ by eq.(1)

Also, flowchart of the proposed learning algorithm is shown in Fig.7.

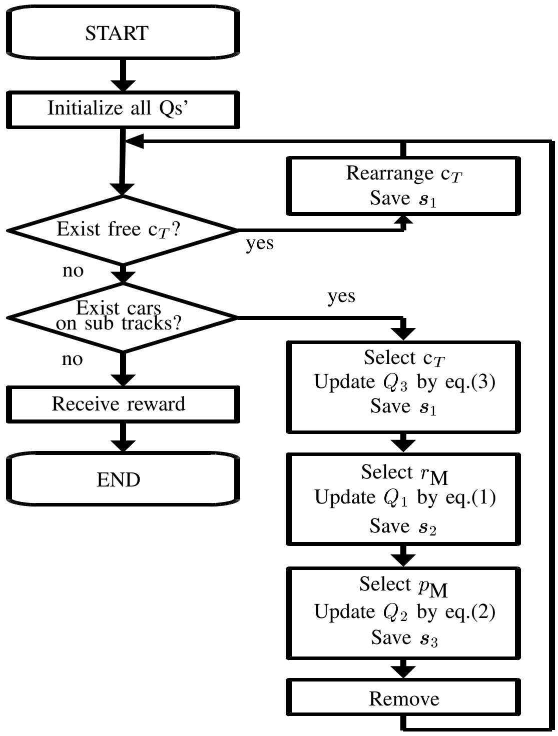

Fig. 7. Flowchart of the learning algorithm

VII. COMPUTER SIMULATIONS

Computer simulations are conducted for $m = 12, n = 6, k = 36$ and learning performances of following 5 methods are compared:

(A) proposed method that evaluates the transfer distance of the locomotive, considers the number of cars to be moved, and uses 2 thermo constants T_1, T_2 with normalized evaluation values,
(B) a method that the number of cars to be moved is 1, and uses 2 thermo constants with normalized evaluation values,
(C) a method that evaluates the number of movements of freight cars, considers the number of cars to be moved, and uses 2 thermo constants with normalized evaluation values,
(D) a method that evaluates the transfer distance of the locomotive, considers the number of cars to be moved, and uses 1 thermo constant T_1,
(E) the same method as (D) that the thermo constant is T_2.

The desirable layout of groups in the main track is depicted in Fig.8, and the initial arrangement of cars in sub tracks is described in Fig.9. In this case, the rearrantement order of groups is group$_1$, group$_2$, group$_3$, group$_4$. Cars $c_1, \cdots, c_9$ are in group$_1$, $c_{10}, \cdots, c_{18}$ are in group$_2$, $c_{19}, \cdots, c_{27}$ are in group$_3$, and $c_{28}, \cdots, c_{36}$ are in group$_4$. Other parameters are set as $\alpha = 0.9, \beta = 0.2, \delta = 0.9, R = 1.0, \eta = 0.95, T_1 = 0.1, T_2 = 0.05$. In method (C), the discount factor γ is assumed to be constant, and set as $\gamma = 0.9$ instead of calculationg by eq.(4).

Figs.10,11 show the results. In Figs.10,11, horizontal axis expresses the number of trials and the vertical axis expresses the minimum transfer distance of locomotive to achieve a desirable layout found in the past trials. Vertical lines in Fig.10 indicate dispersions at the corresponding data points. Each result is averaged over 20 independent simulations. In Fig.10, as the number of trials increases, the transfer distance of locomotive reduces, and method (A) derives solutions

TABLE I
TOTAL TRANSFER DISTANCES OF THE LOCOMOTIVE

methos	transfer distances best	average	worst
method (A)	981	1013.60	1040
method (B)	1892	1929.0	1954
method (C)	1002	1035.75	1078
method (D)	999	1026.75	1051
method (E)	984	1023.35	1049

that require smaller distance of movements of locomotive as compared to method (B). The total transfer distance can be reduced by method (A), because method (A) learns the number of cars to be moved, in addition to the solutions derived by method (B). In Fig.11, the learning performance of method (A) is better than that of methods (D),(E), because normalized evaluation and switching thermo constants in method (A) is effective for reducing the transfer distance of the locomotive. In method (C), the learning algorithm evaluates the number of movements of freight cars, and is not effective to reduce the total transfer distance of locomotive. Total transfer distances of the locomotive at 1×10^6th trial are described in table.I for each method.

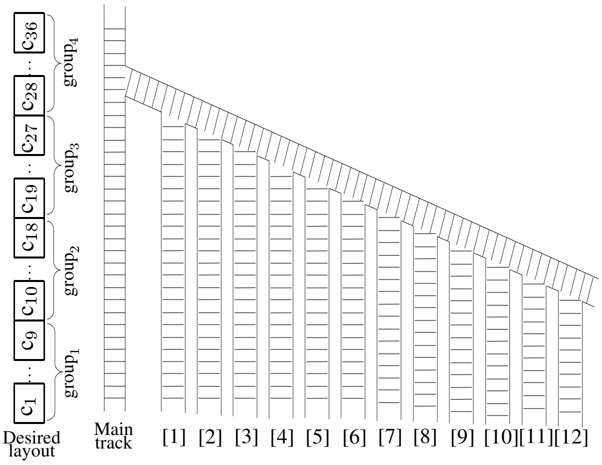

Fig. 8. Yard setting

				c_{19}	c_{32}	c_{31}					
c_{10}		c_{34}	c_{28}	c_{16}	c_4	c_5		c_{25}	c_{29}		c_{36}
c_2	c_{13}	c_8	c_{11}	c_{14}	c_{17}	c_{20}	c_{23}	c_{26}	c_1	c_{33}	c_{35}
c_3	c_6	c_9	c_{12}	c_{15}	c_{18}	c_{21}	c_{24}	c_{27}	c_{30}	c_{22}	c_7

Fig. 9. Initial layout

VIII. CONCLUSIONS

A new scheduling method has been proposed in order to rearrange and line cars in the desirable order onto the main track. The learning algorithm of the proposed method is derived based on the reinforcement learning, considering the total transfer distance of locomotive. In order to reduce the tansfer distance of locomotive, the proposed method learns the number of cars to be moved, as well as the layout of main track, the rearrangement order of cars, and the removal destination of cars, simultaneously. In computer

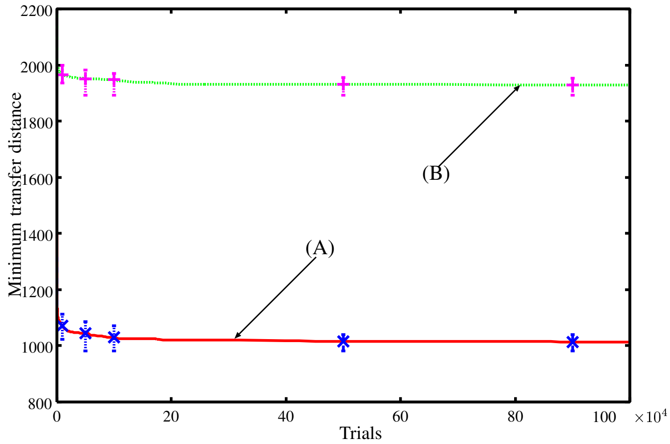

Fig. 10. Minimum transfer distances

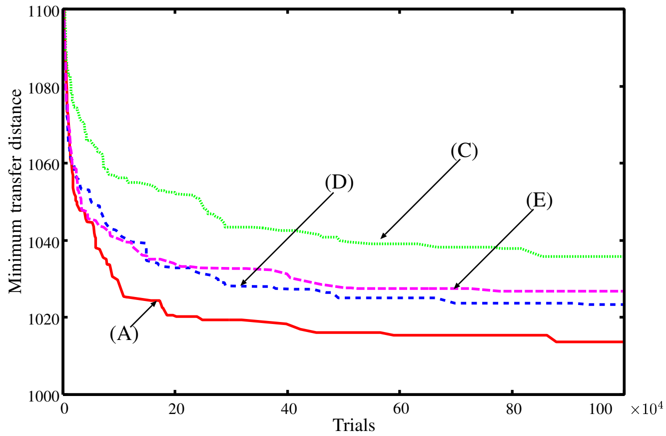

Fig. 11. Comparison of learning performances

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.

ISBN: 978-988-18210-3-4
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)