A Reinforcement Learning for Freight Train with Augmented Collective Motions in a Marshaling

Yoichi Hirashima

***Abstract*— This paper proposes a new reinforcement learning method to solve a train marshaling problem for assembling an outgoing train. In the problem, the arrangement of incoming freight cars is assumed to be random. Then, the cars are rearranged to the desired layout in order to assemble an outgoing train. 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. Then, a rearrange operation is conducted by using several sub-tracks and an outgoing train is assembled in the main track. When a rearrangement operation is conducted in the proposed method, several cars located on different sub-tracks are collected by a locomotive so that the transfer distance of locomotive is reduced. In order to rearrange cars by the desirable order, cars are removed from a sub-track to one of other sub-tracks. Each marshaling plan that consists of series of removal and rearrangement operations are generated by a reinforcement learning system based on the transfer distance of a locomotive. The total transfer distance of the locomotive required to assemble an outgoing train can be minimized by the proposed method.**

***Index Terms*— Collective action, Freight train, Marshaling, Q-Learning, Scheduling, Container transfer problem**

I. Introduction

Railway freight transportation has smaller environmental load as compared to the transportation by tracks. However, modal shifts are required for area that has no railway in order to link different modes of transportation including rail. A freight train consists of several freight cars, and each car has its own destination. The train driven by a locomotive travels several destinations decoupling corresponding freight cars at each freight station. 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 carried by non-rail methods. Containers carried into the station are loaded on freight cars and the initial layout of freight cars is thus random. 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 initially placed on sub-tracks are 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]-[4]. However, in these methods, the marshaling is based on the layout generated by the classification process that assigns each car into a certain sub-track. This causes restrictions on the number of sub-tracks and the number of cars in a sub-track. On the other hands, in our research group, several methods that can generate marshaling plans from the random arrangement of cars in sub-tracks have been proposed [5],[6]. Since, in these methods, each rearrange operation is conducted for cars in the same sub-track, the marshaling plan can be improved by using method proposed in [7] considering collective movements in a rearrange operation. A rearrange operation in the method in [7] is conducted after collecting cars in the same group from several sub-tracks by the desirable order. The method is effective for reducing the total transfer distance of locomotive.

In this paper, the collective movements is augmented and a new learning algorithm for improving marshaling plan of freight cars in a train is proposed. In the proposed method, the initial arrangement of freight cars in sub-tracks are assumed to be random, and thus any arrangements as the result of classification of incoming freight cars can be managed. Then, marshaling plans based on the transfer distance of locomotive are obtained by a reinforcement learning method [8]. Freight cars to be rearranged are collected from several sub-tracks, coupled to each other and moved into the main track by a locomotive. Simultaneously, the optimal sequence of car-movements as well as the number of freight cars that can achieve the desired layout of outgoing train is obtained by autonomous learning. In order to show the effectiveness of the proposed method, computer simulations are conducted for several methods.

II. Problem description

A. Freight Yard

A freight yard is assumed to have 1 main track and *m* sub-

Manuscript received January 7, 2017.
Y. Hirashima is with Osaka Institute of Technology, 1-79-1 Kitayama, Hirakata, Osaka, 573-0196 Japan (phone, fax: +81-72-866-5187; e-mail: yoichi.hirashima@oit.ac.jp).

tracks. Define k as the number of freight cars carried in and assigned to sub-tracks in a classification stage. Then, they are moved to the main track by the desirable order based on their destination in a marshaling stage. In the yard, a locomotive moves freight cars, and 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 ith car is recognized by an unique symbol c_i, (i=1, ⋯, k). Fig. 1 shows the outline of freight yard in the case k=30, m=n=6. In the figure, track Tm denotes the main track, and other tracks [1]-[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.

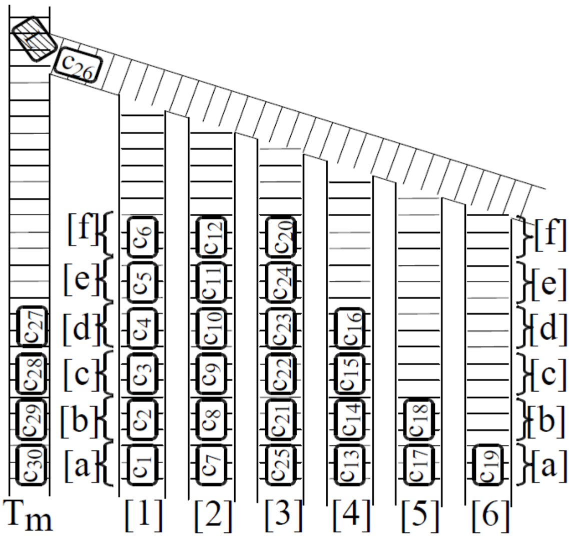


Fig. 1: Freight yard

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 $k \leq nm\text{-}(n\text{-}1)$ 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 mn integers, and the position number for cars at the main track is 0.

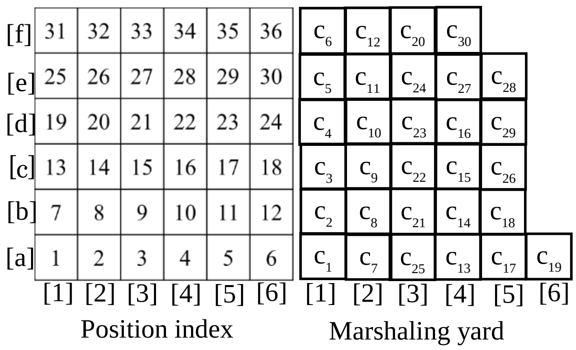


Fig. 2 Example of position index and yard arrangement

Fig. 2 shows an example of position index for k=30, m=n=6 and the layout of cars for Fig. 1-(b). In Fig. 2, the position ``[a][1]" that is located at row ``[a]" in the sub-track ``[1]" has the position number 1. 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.

B. 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 arrangements 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, making a group corresponding to each destination, the order of groups lined in the main track is predetermined by destinations. This feature yields several desirable layouts in the main track.

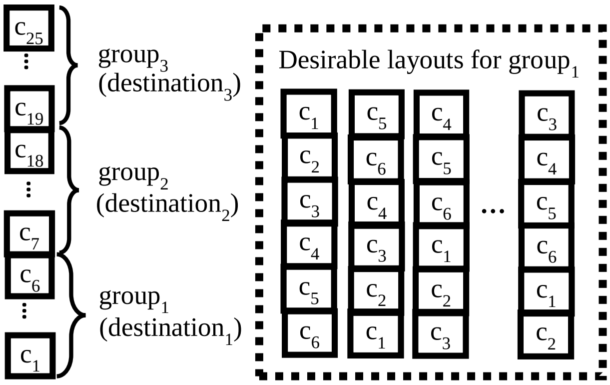
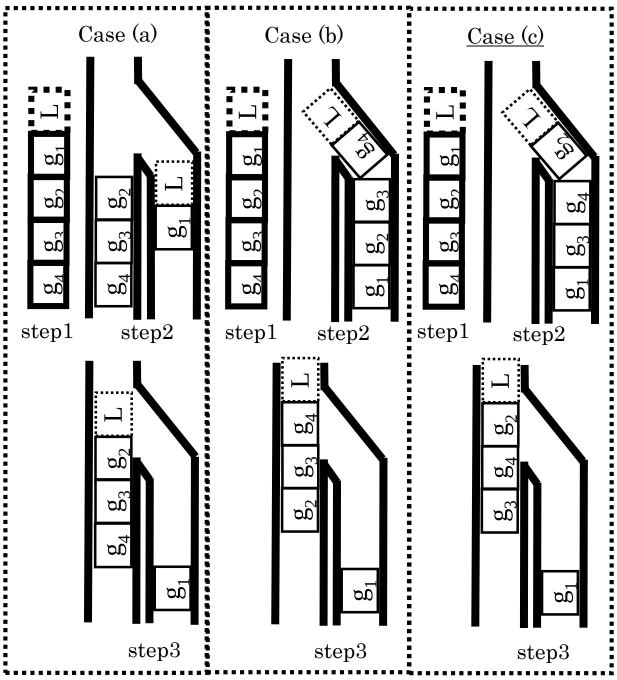


Fig 3: Groups and arrangements of freight cars

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, ⋯, c_6 to the destination1 make group1, c_7, ⋯, c_{18} to the destination2 make group2, c_{19}, ⋯, c_{25} to the destination3 make group3. Groups1,2,3 are lined by ascending order in the main track, which make a desirable layout. Also, in the figure, examples of layout in group1 are in the dashed square.

The layout of groups lined by the reverse order does 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



ISBN: 978-988-14047-3-2
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

from both ends of the train are regarded as desired layouts. By defining r as the number of groups, the total number of layouts of group is 2^{r-1}.

Fig. 4 depicts examples of material handling operation for extended layout of groups at the destination of group1. In the figure, step 1 shows the layout of the incoming train. In case (a), cars in group1 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 group1 are carried in a sub-track, and group1 is separated at the sub-track. In the cases, group1 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.

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 2 cases of marshaling including a removal and rearrangements, 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, and the initial layout of cars in sub-tracks followed by 2 cases in 2 columns are depicted for m=4, n=5, k=9. In both cases, c_1, c_2, c_3, c_4 are in group1, c_5, c_6, c_7, c_8 are in group2, and group1 must be rearranged first to the main track. In each group, any layouts of cars can be acceptable. In case1, c_2 in step 1, c_3 in step 3 and step 4 are applied the direct rearrangement. Also, in case2, steps 2 through 4 are direct rearrangements. Step2 in case1 and step1 in case2 are removals. Step 4 in case1, 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 steps 2 and 3 in case 2, cars in group1 and group2 are collected in sub-tracks to rearrange directly, since the arrangement of c_3, c_2, c_1, c_4, c_5 can satisfy the desired layout of groups in the main track. Whereas, at steps 1,3,4 in case 1, 3 direct rearrangements are conducted separately.

In addition, Fig.6 shows an example of augmented collective movements. In case3 in the figure, the transfer distance of locomotive is reduced in steps 1 and 4 because the freight cars are collected from c_9 through c_3 without a decoupling operation, and because a removal is included in the collective movements. Whereas a decoupling operation is conducted in each step in case1 and two decoupling in steps 1, 4 in case2.

C. Marshaling process

A marshaling process consists of following 6 operations:

(I) selections: group-layout in the main track,

(II) rearrangement of freight cars to the main track, when they can be moved directly,

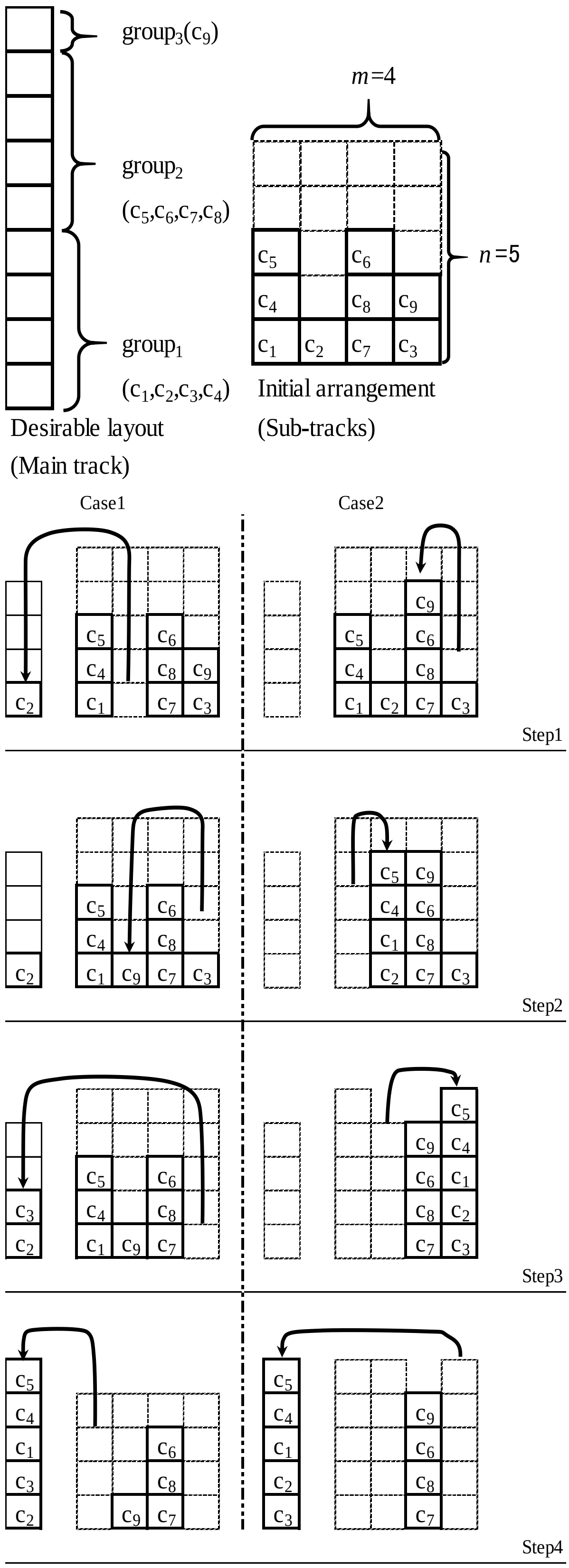


Fig. 5 Example of marshaling

(III) selection: a freight car to be rearranged into the main track,

(IV) selection: removal destinations of the cars in front of the car selected in (IV),

(V) selection: the number of cars to be moved in (V),

(VI) removal of the cars determined in (VI) to the selected sub-track in (V).

After operations (I),(II) are finished, (III)-(VII) are repeated until one of desirable layouts is achieved in the main track, and a series of operations from the initial one in (I) to the final one in (III) achieving the desirable layout is defined as a trial.

Now, define Go as the desired layout and h_1 as the number of candidates of Go. Each candidate in operation (I) is represented by u_{j1} ($1\leq j_1\leq h_1$, $h_1=2^{r-1}$)

In the operation (IV), each group has the predetermined position *in* the main track. Then, the car to be rearranged is defined as c_T, and candidates of c_T is determined by the number of freight cars that have already rearranged to the main track and the group layout in Go. h_2 is defined as the number of freight cars in group to be rearranged, and u_{j3} ($h_1+1\leq j_3\leq h_1+h_2$) as candidates of c_T.

In the operation (V), the removal destination of cars located on the car to be rearranged is defined as T_R. Then, defining u_{j4} ($h_1+h_2+1\leq j_4\leq h_1+h_2+m-1$) as candidates of T_R, excluding the sub-track that has the car to be removed, and the number of candidates is m-1.

In the operation (VI), defining n_p as the number of removal cars required to rearrange c_T, and defining n_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_{j5} ($1\leq u_{j5}\leq\min\{n_p,n_q\}$, $h_1+h_2+m\leq j_5\leq h_1+h_2+m+\min\{n_p, n_q\}$). (III)-(VI) are repeated until all the cars are rearranged into the main track.

In the operation (II), collective movements are generated by the following operations (IIa-IId):

(IIa) Determine the cars that can be rearranged directly to the main-track, and the sub-tracks including those cars. Also, determine if such sub-tracks can be yielded by a single removal operation, and define h_4 the number of candidates of removal-destination for the removal cars.

(IIb) Define T_C as the sub-track including cars to be rearranged directly, and h_3 as the number of candidates of T_C.

(IIc) If $h_3>1$, select a T_C as the tail of the collected cars,
if $h_3=1$, rearrange T_C without collective movements,
if ($h_3=0$, $h_4<2$), operation (II) is finished,
and if ($h_3=0$, $h_4>1$), select a removal operation,

(IId) Conduct rearrange or removal and go to (IIa).

Then, defining u_{j2} ($h_1+h_2+m+\min\{n_p, n_q\}\leq j_2\leq h_1+h_2+h_3+h_4+m+\min\{n_p, n_q\}$) as candidates

D. 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) restarts in reverse direction to the target car to be moved, (E3) joints them, (E4) pulls out them to the joint track, (E5) restarts 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, and the distance of the unit transition D is calculated by $D= D_1+ D_2+ D_3 +D_4$. Also, define the unit distance of a movement for cars in each sub-track as D_{minv}, the length of joint track between adjacent sub-tracks, or, sub-track and main track as D_{minh}. The location of the locomotive at the end of above process is the start location of the next movement process of the selected car. The initial position of the locomotive is located on the joint track nearest to the main track.

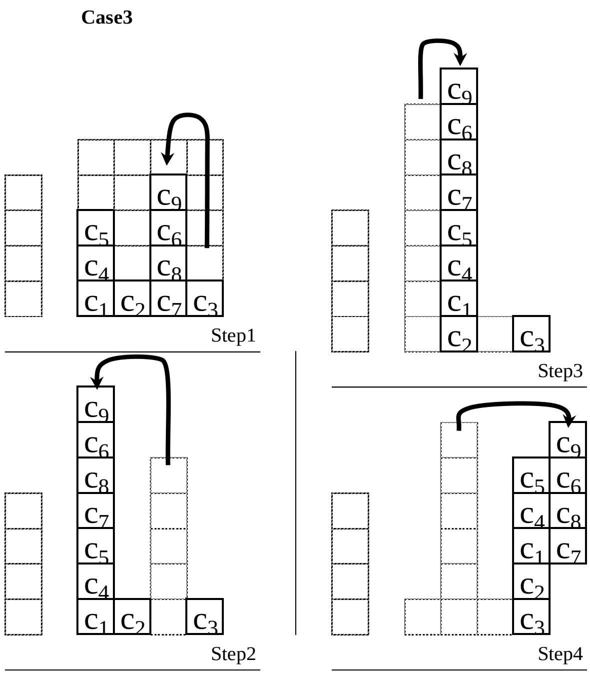


Fig. 6 Augmented collective movements

Fig. 7 shows an example of transfer distance. In the figure, $m=n=6$, $D_{minv}=D_{minh}=1$, $k=18$, (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 D1+D2=12, D1=5,D2=7, whereas it moves from 24 to 16 with 2 freight cars, and the transfer distance is D3+D4=13 (D3=7,D4=6).

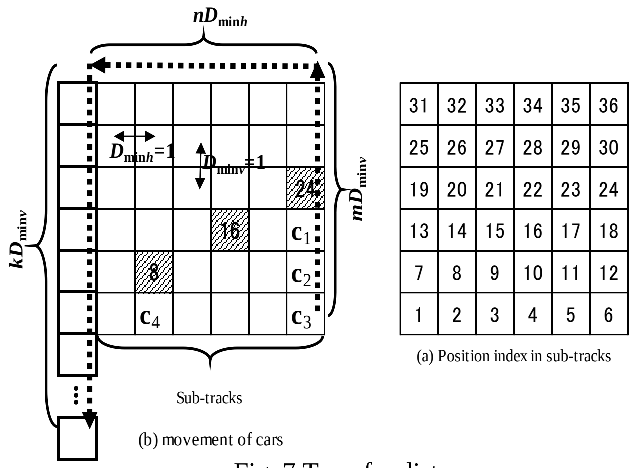


Fig. 7 Transfer distance

III. Learning Algorithm

Defining Q_1 as an evaluation value for Go, $Q_1(G$o$)$ is updated by the following rule when one of desired layout is achieved in the main track:

$$Q_1(G_0) \leftarrow \max\left\{Q_1(G_0), (1-\alpha)Q_1(G_0) + \alpha R\prod_{i=1}^{l}\gamma_i\right\} \quad (1)$$

where l denotes the total movement counts required to achieve the desired layout， α is learning rate, γ_i 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 $\boldsymbol{s}(t)$ as the state at time t, Tc 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 $\boldsymbol{s}$' as the state that follows $\boldsymbol{s}$. In the direct rearrangement, $Q_{2\mathrm{a}}$ is defined as evaluation values for $(\boldsymbol{s}_1,u_{j2\mathrm{a}})$, where $\boldsymbol{s}_1=[\boldsymbol{s}, G\mathrm{o}]$, $\boldsymbol{s}_2=[\boldsymbol{s}_1, T_{\mathrm{C}}]$, and $Q_{2\mathrm{b}}$ is defined as evaluation values for $(\boldsymbol{s}_1,u_{j2\mathrm{b}})$. Then, $Q_{2\mathrm{a}}(\boldsymbol{s}_1,T_{\mathrm{C}})$ is updated by the following rule:

$$Q_{2a}(\boldsymbol{s}_1,T_{\mathrm{C}}) \leftarrow (1-\alpha)Q_{2a}(\boldsymbol{s}_1,T_{\mathrm{C}})+\alpha V_1$$

$$V_1=\begin{cases}\gamma\mathrm{R} & \text{(all cars moved into the main track)}\\ \gamma\max\limits_{u_{j3}} Q_3(\boldsymbol{s}_3,u_{j_3}) & \text{(collective movement is finished)}\\ \gamma\max\limits_{u_{j2a}} Q_{2a}(\boldsymbol{s}_1,u_{j_{2a}}) & \text{(collective movement is continued)}\end{cases} \tag{2}$$

In removal operations, define p_{M} as the number of cars moved. Q_3,Q_4 and Q_5 are defined as evaluation values for $(\boldsymbol{s}_1,u_{j3})$,$(\boldsymbol{s}_3,u_{j4})$,$(\boldsymbol{s}_4,u_{j5})$ respectively, where $\boldsymbol{s}_3=\boldsymbol{s}_1$, $\boldsymbol{s}_4=[\boldsymbol{s}_3, \mathrm{c}_T]$, $\boldsymbol{s}_5=[\boldsymbol{s}_4, T_{\mathrm{R}}]$. $Q_{2\mathrm{b}}(\boldsymbol{s}_4,T_{\mathrm{R}})$,$Q_3(\boldsymbol{s}_3,\mathrm{c}_T)$ $Q_4(\boldsymbol{s}_4, T_{\mathrm{R}})$, and $Q_5(\boldsymbol{s}_5,p_{\mathrm{M}})$ are updated by following rules:

$$Q_{2b}(\boldsymbol{s}_4,T_{\mathrm{R}}) \leftarrow (1-\alpha)Q_{2b}(\boldsymbol{s}_4,T_{\mathrm{R}})+\alpha V_2$$

$$V_2=\gamma\max_{u_{j2a}} Q_{2a}(\boldsymbol{s}_1,u_{j_{2a}})$$

$$Q_3(\boldsymbol{s}_3,\mathrm{c}_T) \leftarrow \max_{u_{j4}} Q_4(\boldsymbol{s}_4,u_{j4}) \tag{3}$$

$$Q_4(\boldsymbol{s}_4,T_{\mathrm{R}}) \leftarrow \max_{u_{j5}} Q_5(\boldsymbol{s}_5,u_{j5}) \tag{4}$$

c6											
c4					c26						c10
c27			c14		c24	c15					c5
c30			c13		c19	c7		c1			c16
c2	c29	c8	c25	c3	c21	c35		c33	c20		c11
c31	c28	c9	c23	c36	c17	c34		c22	c12	c18	c32

Fig. 8 Initial arrangement of cars in sub-tracks

$$Q_5(\boldsymbol{s}_5,p_{\mathrm{M}}) \leftarrow \begin{cases}(1-\alpha)Q_5(\boldsymbol{s}_5,p_{\mathrm{M}})+\alpha[\gamma V_3],\\ V_3=\max\limits_{u_{j2a}} Q_{2\mathrm{a}}(\boldsymbol{s}_1',u_{j2a}) \quad (u \text{ is a rearrangement})\\ (1\text{-}\alpha)Q_5(\boldsymbol{s}_5,p_{\mathrm{M}})+\alpha[\mathrm{R}+\gamma V_4],\\ V_4=\max\limits_{u_{j4}} Q_4(\boldsymbol{s}_4',u_{j4}) \quad (u \text{ is a removal})\end{cases} \tag{5}$$

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 locomotive and calculated by the following equation:

$$\gamma=\delta\frac{D_{\max}-\beta D}{D_{\max}},(0<\beta<1,0<\delta<1), \tag{6}$$

where $D_{\max}$ is the maximum value of D.

Propagating Q-values by using eqs.(1)-(6), Q-values are discounted according to the transfer distance of locomotive. In other words, by selecting the removal destination that has the largest Q-value, the transfer distance of locomotive can be reduced.

In the learning stages, each u_{j}, $(1\leq j \leq h_1+h_2+h_3+h_4+m+\min\{n_{\mathrm{p}}, n_{\mathrm{q}}\})$ is selected by the soft-max action selection method [8] In the addressed problem, Q_2, Q_3, Q_4, Q_5 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, $Q_{2\mathrm{a}}$, $Q_{2\mathrm{b}}$, Q_3, Q_4, Q_5, Q_6 are normalized, and probability P for selection of each candidate is calculated by

$$\tilde{Q}_i(\boldsymbol{s}_i,u_{j_i})=\frac{Q_i(\boldsymbol{s}_i,u_{j_i})-\min\limits_{u_{ji}} Q_i(\boldsymbol{s}_i,u_{j_i})}{\max\limits_{u_{ji}} Q_i(\boldsymbol{s}_i,u_{j_i})-\min\limits_{u_{ji}} Q_i(\boldsymbol{s}_i,u_{j_i})}, \tag{7}$$

$$P(\boldsymbol{s}_i,u_{j_i})=\frac{\exp(\tilde{Q}_i(\boldsymbol{s}_i,u_{j_i})/\xi)}{\sum\limits_{u\in u_{ji}}\exp(\tilde{Q}_i(\boldsymbol{s}_i,u)/\xi)}, \tag{8}$$

$(\mathrm{i}=2,3,4,5)$

$$P(u_{j_1})=\frac{\exp(\tilde{Q}_1(u_{j_1})/\xi)}{\sum\limits_{u\in u_{j_1}}\exp(\tilde{Q}_1(u)/\xi)},$$

wherer ξ is a thermo constant.

IV. Computer Simulations

Computer simulations are conducted for m=12,n=6,k=36 and learning performances of following 2 methods are compared:

(A) Proposed method considering with augmented collective movements in the direct rearrangement,
(B) Proposed method with conventional collective movements [7]
(C) Proposed method without collective movements in the direct rearrangement [6].

The initial arrangement of freight cars in sub-tracks train is described in Fig.8. The original group-layout is group_1, group_2, group_3, group_4, the order that is depicted in Fig.9. Also, 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, c_{28}, $\cdots$,c_{36} are in group_4. Other parameters are set as α =0.9, β =0.2, δ =0.9, R=1.0, ξ =0.1. Both methods consider extended layout of groups and derive desirable layouts autonomously in order to reduce the total transfer distance of the locomotive. The locomotive and freight cars assumed to have the same length, and $D_{\mathrm{minv}}=D_{\mathrm{min}h}$=20m.

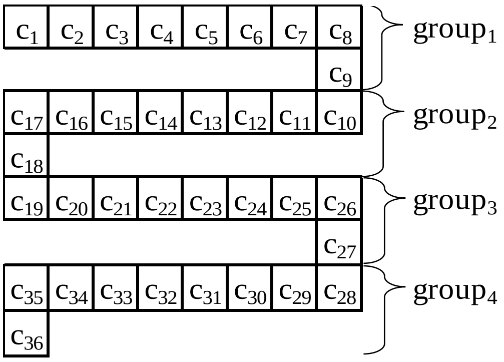


Fig. 9 Original group-layout in the main track

The results are shown in Fig.10. In the figure, horizontal axis expresses the number of trials and the vertical axis expresses the minimum processing time found in the past trials. A trail starts with an initial arrangement in sub-tracks, and finished when all the freight cars are rearranged into the main track. Each result is averaged over 20 independent simulations. In Fig.10, although the learning performance of method (A) is under preparing, (A) includes all the collective movements in (B) so that the marshaling plan can be improved as compared to that in (B). Also, the learning performance of method (B) is better than that of (C), because (B) collects freight cars in sub-tracks for each direct rearrangement. Simulation results obtained by (A) will be shown at the presentation.

Table 1 Total transfer distance

	Transfer distance [m]		
	best	average	worst
Method(B)	15280	15480	15800
Method(C)	24980	25600	26720

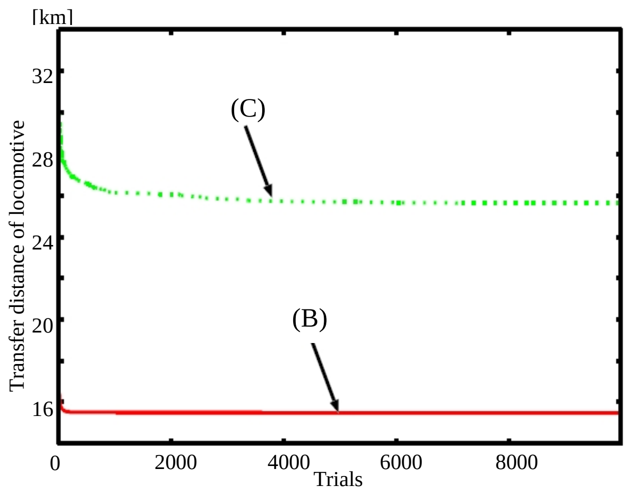


Fig. 10 Performance comparison

V. Conclusions

A new scheduling method has been proposed in order to rearrange and line cars in the desirable order onto the main track considering augmented collective movements of the locomotive in the direct rearrangement of freight cars.

The learning algorithm of the proposed method is derived based on the reinforcement learning that evaluates the total transfer distance of locomotive in a marshaling. In order to reduce the total transfer distance of locomotive in a marshaling, the proposed method obtain group-layout of desirable arrangement of outgoing train and collecting order of freight cars in sub-tracks, so that the total transfer distance in a marshaling plan derived by the proposed method has been reduced by about 40% as compared to the conventional method in computer simulations. In addition, the arrangement of freight cars in the main track, the rearrange order of cars, the position of each removal car, the number of cars to be removed, and the group layout in the outgoing train has been obtained simultaneously so that the learning performance of the proposed method has been improved.

References

[1] C. F. Daganzo, R. G. Dowling, R. W. Hall: *Railroad classification yard throughput: The case of multistage triangular sorting*, Transportation Research, Part A, No.17, Vol. 2 (1983), pp. 95-106.

[2] 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 (2000) , pp. 137-148.

[3] R. Jacob, P. Marton, J. Maue and M. Nunkesser: *Multistage Methods for Freight Train Classification*, Proceedings of 7th Workshop on Algorithmic Approaches for Transportation Modeling, Optimization, and Systems (2007) , pp. 158-174.

[4] T. Nonner and A. Souza: *Optimal Algorithms for Train Shunting and Relaxed List Update Problems*, 12th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems (2012), pp. 97-107.

[5] Y. Hirashima: *An Intelligent Train Marshaling Based on the Processing Time Considering Group Layout of Freight Cars* in IAENG Transactions on Electrical Engineering Volume 1, World Scientific (2013), pp.229-243.

[6] Y. Hirashima: *A Reinforcement Learning System for Generating Train Marshaling Plan of Freight Cars Based on the Processing Time Considering Group Layout*, Proceedings of The International MultiConference of Engineers and Computer Scientists 2013, (2013) , pp.30-35.

[7] Y. Hirashima: *A Reinforcement Learning for Marshaling of Freight Train Considering Collective Motions*, Proceedings of The International MultiConference of Engineers and Computer Scientists 2016, (2016) , pp. pp19-24.

[8] R. Sutton and A. Barto: *Reinforcement Learning*, MIT Press (1999).