Proceeding of the 2004 American Control Conference
Boston, Massachusetts June 30 - July 2, 2004

Automatic Core Design Using Reinforcement Learning

Yoko Kobayashi
TEPCO SYSTEMS CORPORATION,
6-19-15 Shinbashi Minato-ku Tokyo, Japan, kobayashi-youko@tepsys.co.jp

Eitaro Aiyoshi
Keio University,
3-14-1 Hiyoshi Kouhoku-ku Yokohama Kanagawa, Japan

Abstract— *This paper deals with the application of multi-agents algorithm to the core design tool in a nuclear industry. We develop an original solution algorithm for the automatic core design of boiling water reactor using multi-agents and reinforcement learning. The characteristics of this algorithm are that the coupling structure and the coupling operation suitable for the assigned problem are assumed, and an optimal solution is obtained by mutual interference in multi state transitions using multi-agents. We have already proposed an integrated optimization algorithm using a two-stage genetic algorithm for the automatic core design. The objective of this approach is to improve the convergence performance of the optimization in the automatic core design. We compared the results of the proposed technique using multi-agents algorithm with the two-stage genetic algorithm that had been proposed before. The proposed technique is shown to be effective in reducing the iteration numbers in the search process.*

1. Introduction

A boiling water reactor (BWR) is one of the commercial nuclear power reactors. The reactor core of a BWR is arranged as an upright cylinder containing a large number of fuel assemblies. The coolant flows upward through the core. Control rods occupy alternate space between fuel assemblies and can be withdrawn into the guide tubes. Positive core reactivity control is maintained by using movable control rods interspersed throughout the core.

When a BWR is shut down between successive operation cycles refueling or reloading is performed. This refueling or reloading is called "loading pattern (LP)". The number of control rods inserted at many particular time and depth to which they are inserted is called "control rod pattern (CRP)". In the actual core design of a BWR, it is necessary to decide these two optimal patterns at the same time. The core design is a NP-complete combinatorial optimization problem. Therefore, although there is a strong desire for to save labor, automatic optimization of the core design of a BWR has been assumed to be a very difficult combinatorial optimization problem.

For the last several years, we have been using genetic algorithms (GA) without transcendental information to optimize the core design of BWR. We were able to achieve the automatic optimization of the complex core design of a BWR by the two-stage optimization algorithm using GA. In this paper, we further proposed a new original algorithm for combinatorial optimization using multi-agents and reinforcement learning. We name it as multi-agents algorithm (MAA). In order to improve the convergence performance of the core design optimization of BWR, we introduce this new algorithm to the first stage of the two-stage GA previously developed.

2. Optimization Problem of Core Design

Since the LP of a BWR usually has one-eight symmetry property, the optimization of LP and CRP is carried out generally on the octant core as shown in Fig.1. (In this example, the numbers of fuel assemblies are 115.) The numbers of the loading position of each fuel are applied as shown in Fig.1. The fuel assembly placed in loading position l is defined as x_l and this list is arranged as $\boldsymbol{x} = (x_1, \cdots, x_l, \cdots, x_L)$, where $x_l \neq x_{l'}$, $l \neq l'$, $x_l, x_{l'} \in \{1, \cdots, L\}$. L is the number of fuel assembly. Then, if the insertion depth of the CR position n ($n = 1, \cdots, N_c$) at each burn-up step t is defined as y_n , the list is expressed as $y(t) = (y_1(t), \cdots, y_n(t), \cdots, y_{N_c}(t))$ (see Fig.2). There are N_t burn-up steps including both BOC (beginning of cycle) and EOC (end of cycle), and defined as $t = (1, \cdots, N_t)$. The time change of control rod pattern list $\boldsymbol{y}$ at all the burn-up points can be written as $\boldsymbol{y}(\bullet) = (y(1), \cdots, y(N_t))$.

LP evaluation of the list $\boldsymbol{x}$ is performed with a three-dimensional diffusion code coupled with neutronic and thermal hydraulic models. The core performance calculation based on CRP outputs some parameters, which are the two limiting value of FLCPR (fraction of limiting critical power ratio) and FLPD (fraction of limiting power density) and k_{eff} (reactor eigenvalue) and relative nodal power distribution. These output parameters are shown as a function of LP list $\boldsymbol{x}$ and CRP list $\boldsymbol{y}$. Because $\boldsymbol{y}$ is expressed as a function of t , they can be respectively expressed as $e_f(\boldsymbol{x}, \boldsymbol{y}(t))$, $r_{lm}(\boldsymbol{x}, \boldsymbol{y}(t))$, $flcpr_l(\boldsymbol{x}, \boldsymbol{y}(t))$, $flpd_{lm}(\boldsymbol{x}, \boldsymbol{y}(t))$. Subscript l indicates each bundle in octant core ($l = 1, \cdots, L$) , and subscript m indicates the axial node of fuel ($m = 1, \cdots, 24$). Because the target values and the upper limit values of these parameters are expressed as a function of t , they are defined as:

$\overline{e_f}(t)$: the target value of k_{eff} ,

$\overline{r}(t)$: the upper limit value of relative nodal power,

$\overline{flcpr}(t)$: the upper limit value of FLCPR,

$\overline{flpd}(t)$: the upper limit value of FLPD.

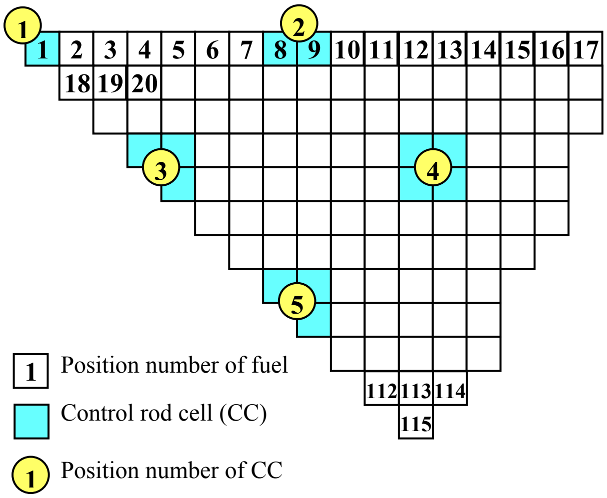

Fig. 1 Position numbers of fuel and locations of control rods (example of 5CC)

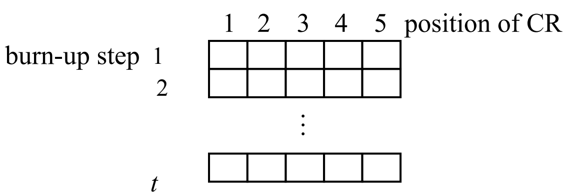

Fig. 2 Sample coding of chromosome in second stage

In the master-servant relationship between LP and CRP, it is important that the set of CRP $\boldsymbol{y}$ searched later agrees with the target value of $\overline{e_f}(t)$ for the set of LP searched first. Such a pattern $\boldsymbol{x}$ is called a controllable LP, and the cluster is expressed as Eq. (1) using tolerance ε_{LP} :

$$X = \{\boldsymbol{x} \mid |e_f(t) - \overline{e_f}(t)| < \varepsilon_{LP}, t = 1, \cdots, N_t \text{ for} \exists \boldsymbol{y}(\bullet) \tag{1}$$

where $\boldsymbol{y}(\bullet) = (y(1), \cdots, y(N_t))$. As the control rod position, constrained set Y concerning $\boldsymbol{y}(\bullet)$ is given by:

$$Y = \{\boldsymbol{y}(\bullet) \mid y_n(t) \in \{pos_{in}, \cdots, pos_{out}, pos_{all}\}, t = 1, \cdots, N_t\}, \tag{2}$$

where pos_{in} , pos_{out} , and pos_{all} are the limitation of insertion of control rods, the limitation of drawing, and all rods out. In addition, there is a core flow $flow(t)$ which can be adjusted at each burn-up point as the control parameter. The upper limit value of flow at EOC is defined as $\overline{flow_{EOC}^{upper}}$. From these, the objective function and the penalty function are shown as the following at every burn-up step t :

$$g_0(\boldsymbol{x}, \boldsymbol{y}(\bullet)) = \max_t \{|e_f(t)) - \overline{e_f}(t)|\}, \tag{3}$$

$$g_0(\boldsymbol{x}, \overline{\boldsymbol{y}}(\bullet)) < \varepsilon_{LP}, \tag{4}$$

$$g_1(\boldsymbol{x}, \overline{\boldsymbol{y}}(\bullet)) = \max_t \{\max_{(m,l)} \{\max\{r_{lm}(t)) - \overline{r}(t), 0\}\}\}, \tag{5}$$

$$g_2(\boldsymbol{x}, \overline{\boldsymbol{y}}(\bullet)) = \max_t \{\max_l \{\max \{flcpr_l((t)) - \overline{flcpr}(t), 0\}\}\}, \tag{6}$$

$$g_3(\boldsymbol{x}, \overline{\boldsymbol{y}}(\bullet)) = \max_t \{\max_{(m,l)} \{\max \{flpd_{lm}((t)) - \overline{flpd}(t), 0\}\}\}, \tag{7}$$

$$g_4(\boldsymbol{x}, \overline{\boldsymbol{y}}(\bullet)) = \sum_{t=2}^{N_t} \sum_{k=1}^{N_c} \{y_k(t) - y_k(t-1)\}, \tag{8}$$

$$P(\boldsymbol{x}, \overline{\boldsymbol{y}}(\bullet)) = \max\{flow(\boldsymbol{x}, \boldsymbol{y}(N_t)) - \overline{flow_{EOC}^{upper}}, 0\}, \tag{9}$$

where $\overline{\boldsymbol{y}}(\bullet)$ is given by:

$$\overline{\boldsymbol{y}}(\bullet) = \arg \min_{\boldsymbol{y}(\bullet)} g_0(\boldsymbol{x}, \boldsymbol{y}(\bullet)). \tag{10}$$

Eq. (8) is the objective that reduces the fluctuation of the control rod positions as much as possible. From these, the problem of core design is formulated as:

$$\min_{\boldsymbol{x}} d_1 g_1(\boldsymbol{x}, \overline{\boldsymbol{y}}(\bullet)) + d_2 g_2(\boldsymbol{x}, \overline{\boldsymbol{y}}(\bullet)) + d_3 g_3(\boldsymbol{x}, \overline{\boldsymbol{y}}(\bullet)) + d_4 g_4(\boldsymbol{x}, \overline{\boldsymbol{y}}(\bullet)) + d_5 P(\boldsymbol{x}, \overline{\boldsymbol{y}}(\bullet)), \tag{11}$$

$$\text{subject to } \boldsymbol{x} \in \boldsymbol{X}, \tag{12}$$

$$g_0(\boldsymbol{x}, \overline{\boldsymbol{y}}(\bullet)) < \varepsilon_{LP}, \tag{13}$$

where

$$\overline{\boldsymbol{y}}(\bullet) = \arg \min_{\boldsymbol{y}(\bullet) \in Y} g_0(\boldsymbol{x}, \boldsymbol{y}(\bullet)), \tag{14}$$

$$\boldsymbol{x} = (x_1, \cdots, x_l \cdots, x_{115}), \ x_l \neq x_{l'}, \ l \neq l', \quad x_l, x_{l'} \in \{1, \cdots, 115\}, \tag{15}$$

$$\boldsymbol{y}(t) = (y_1(t), \cdots, y_n(t), \cdots, y_{N_c}(t)), \quad t \in \{t_1, \cdots, t_{N_t}\}, \tag{16}$$

$$y_n(t) \in \{pos_{in}, \cdots, pos_{out}, pos_{all}\}. \tag{17}$$

The values of d_1 to d_5 in Eq. (11) are determined by some trial-and-error based on the empirical rule obtained by the reload core design of a real plant. In the optimization of the first stage, $\boldsymbol{x}$ is renewed under the constraint of the controllability for $\boldsymbol{x}$ of Eq. (12). In the optimization of the second stage, $\overline{\boldsymbol{y}}(\bullet)$ which is subordinate for $\boldsymbol{x}$ is required, after Eq. (14) is executed.

3. Our Previous Approaches

The optimization problem of LP and CRP formulated in the previous section is a two-stage complex combinatorial problem. Therefore, the application of an efficient and convenient search algorithm is desired. In our previous approaches, we use GA as such algorithm. Although comparatively many optimal solutions exist for such a complex combinatorial optimization problem, the reason for using GA is the simplicity and fast in comparison with other heuristic search algorithm. GA is a multi-point search method, where a lot of solutions are explored concurrently is the solution space. GA has the potential to find a good global solution in conditions with multiple variables, and it is more likely to obtain a solution that is better than one an expert finds stochastically. For these reasons, GA was chosen as the search method, and we aimed at improvement of efficiency in solving this problem. In this section, we present our modification of GA designed to enhance its performance.

3.1 Improved GA

First of all, we developed an improved GA [1] for a complex combination optimization problem. The features of improved GA that we developed before are described as follows:

1. Performance of the GA is improved by the execution of the deterministic operator,
2. Convergence efficiency is raised by the adoption of an elite strategy that utilizes the fact that the LP problem is a two-objective problem,
3. Convergence efficiency is raised further by self-reproduction done every several generation, and
4. Convergence performance is improved by using the initial value dependence.

This algorithm provides both good convergence and global searching ability.

3.2 Integrative two-stage GA

Next, we developed an integrative two-stage optimization technique [2] of dynamic stage and static stage based on GA. In the first stage optimization, the improved GA is applied and the LP is optimized by the improved GA. The fitness function in the first stage is defined as:

$$G(\boldsymbol{x}, \boldsymbol{y}) = \left(\frac{1}{1 + \exp(g(\boldsymbol{x}, \bar{\boldsymbol{y}}(\bullet)))} \right)^3, \quad (18)$$

where :

$$\begin{aligned} g(\boldsymbol{x}, \bar{\boldsymbol{y}}(\bullet)) = & d_1 g_1(\boldsymbol{x}, \bar{\boldsymbol{y}}(\bullet)) + d_2 g_2(\boldsymbol{x}, \bar{\boldsymbol{y}}(\bullet)) \\ & + d_3 g_3(\boldsymbol{x}, \bar{\boldsymbol{y}}(\bullet)) + d_4 g_4(\boldsymbol{x}, \bar{\boldsymbol{y}}(\bullet)) \quad (19) \\ & + d_5 P(\boldsymbol{x}, \bar{\boldsymbol{y}}(\bullet)). \end{aligned}$$

In the second stage, for each individual of the LP, the CRP $\bar{\boldsymbol{y}}(\bullet)$ of Eq. (14) is optimized by a powerful heuristic if-then rule. In the search process for the CRP, the search space of $\bar{\boldsymbol{y}}(\bullet)$ is constrained to the one that satisfies Eq. (13); in addition, the search space of the epistatic LP is also constrained. Thus, the search efficiency is improved by introducing the if-then rules in the search process of $\bar{\boldsymbol{y}}(\bullet)$. The fitness function in the second stage is defined as:

$$G_0(\boldsymbol{x}, \boldsymbol{y}(\bullet)) = \left(\frac{1}{1 + \exp(g_0(\boldsymbol{x}, \boldsymbol{y}(\bullet)))} \right)^3. \quad (20)$$

4. Concept of Multi-Agents Algorithm

We present the basic concept of the proposed algorithm in this section. First, consider an agent space that has several agents in it, as shown in Fig.3. Each agent has a state. In our MAA, the optimal solution in the combinatorial optimization problem is obtained by exchanges of information using the coupling relations among multiple agents. The concept of our MAA is composed of the following five items:

1. The coupling structure among agents,
2. The dimensions and neighborhood of the coupling structure,
3. The type of coupling operation,
4. The acceptable threshold of the minimization function value, and
5. The selection of the firing elements by the operation.

A state transition of the agents' element is called a "firing." We will explain the five elements according to the above criteria.

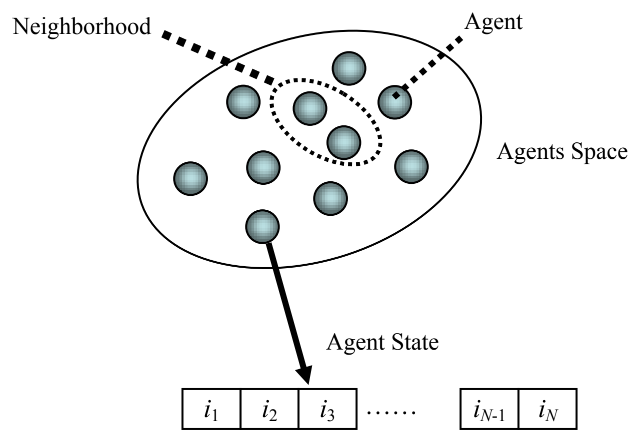

Fig.3 Agents Space and Agent State

4.1 Coupling Structure among Agents

A structure in a specific neighborhood that can be defined for any agent in the space among agents is called the "coupling structure." An agent's state is changed by exchanging information according to specific rules in the neighborhood of the agent. A specific coupling structure among agents in a specific neighborhood is defined, and the coupling structure is variable when one partner of the information exchange is changed according to a specific rule. There are three kinds of coupling structure as examples of the variable structure: (1) variable coupling structure, (2) probable variable coupling structure, (3) random variable coupling structure.

4.2 Dimension and Neighborhood of coupling Structure

In the coupling structure, the spread of a coupling is expressed by the concept of dimension. Let us consider a combinatorial minimization problem in which all elements of variable $\boldsymbol{x}$ are the integer value $\{1, \cdots, N\}$ of the permutation type. If these multiple individuals are associated with a set of M numbers multi-agents, this optimization problem can be expressed as the following problem which solves for multiple states:

$$\min_{\boldsymbol{x}^m} E(\boldsymbol{x}^m), \tag{21}$$

$$\text{subj.to} \quad x_i^m, x_{i'}^m \in \{1, \cdots, N\}, \; m = 1, \cdots, M, \tag{22}$$

$$\text{where} \quad x_i^m \neq x_{i'}^m, \; i \neq j. \tag{23}$$

At this time, in the case of a one-dimensional coupling structure, the state of agent $\boldsymbol{x}^m$ in problem (21) is expressed by $\boldsymbol{x}(k)$. In the case of a two-dimensional coupling structure, it is expressed by $\boldsymbol{x}(k_1, k_2)$. Here, k, k_1 and k_2 are integer values that express the structure of the agent set.

4.3 Coupling Operation

Some coupling operations are defined by the distance between agents in the neighborhood. In this paper, the distance between two agents $\boldsymbol{x}$ and $\boldsymbol{y}$ is defined as follows:

$$\| \boldsymbol{x} - \boldsymbol{y} \| = \sum_{i=1}^{N_e} | x_i - y_i | , \tag{24}$$

$$| x_i - y_i | = \begin{cases} 0, & \text{if } x_i = y_i \\ 1, & \text{if } x_i \neq y_i \end{cases} \tag{25}$$

where N_e is the total number of element in an agent. From this definition, the coupling operations can be classified into the following three types:

Forward coupling operation
An operation in which the specific agent moves toward all agents' states in the neighborhood,

Backward coupling operation
An operation in which the specific agent goes away from all agents' states in the neighborhood,

Neutral coupling operation
An operation in which the specific agent keeps a constant distance (h) from all agents' states in the neighborhood.

In addition, an operation that adopts one side of the neighborhood as a neighborhood is called a "convection coupling operation," and the operation that adopts both sides of the neighborhood is called a "diffusion coupling operation." For example, in the case of one-dimensional coupling structure, the coupling operations between $\boldsymbol{x}(k \pm 1)$ and $\boldsymbol{x}(k)$ can be classified into the following six types.

Forward convection coupling operation

$$\min_{\boldsymbol{x}(k)} \| \boldsymbol{x}(k+1) - \boldsymbol{x}(k) \| \tag{26}$$

Backward convection coupling operation

$$\max_{\boldsymbol{x}(k)} \| \boldsymbol{x}(k+1) - \boldsymbol{x}(k) \| \tag{27}$$

Neutral convection coupling operation

$$\min_{\boldsymbol{x}(k)} \{ \| \boldsymbol{x}(k+1) - \boldsymbol{x}(k) \| - h \}^2 \tag{28}$$

Forward diffusion coupling operation

$$\min_{\boldsymbol{x}(k)} \{ \| \boldsymbol{x}(k+1) - \boldsymbol{x}(k) \| + \| \boldsymbol{x}(k-1) - \boldsymbol{x}(k) \| \} \tag{29}$$

Backward diffusion coupling operation

$$\max_{\boldsymbol{x}(k)} \{ \| \boldsymbol{x}(k+1) - \boldsymbol{x}(k) \| + \| \boldsymbol{x}(k-1) - \boldsymbol{x}(k) \| \} \tag{30}$$

Neutral diffusion coupling operation

$$\min_{\boldsymbol{x}(k)} \{ \| \boldsymbol{x}(k+1) - \boldsymbol{x}(k) \| + \| \boldsymbol{x}(k-1) - \boldsymbol{x}(k) \| - h \}^2 \tag{31}$$

Multi-agents should minimize the function E during the same time as the state transition while interfering with the coupling operations each other. Various synthesis methods for the minimization operation and the coupling operation are also considered. However, in this paper, the algorithm in which the minimization operation of the function was maintained only by the coupling operation was adopted.

4.4 Threshold Acceptance of Minimization Function Value

In the coupling operation, the ability to decrease the minimization function is not provided. Therefore, the minimization of the function depends on a method that allows a wide range of deterioration of the function value. Let $\boldsymbol{x}(k, t)$ be the state of an agent in the current generation and let $\boldsymbol{x}(k, t+1)$ be the state of the agent in the next generation. If the difference in the minimization function value between both is expressed as $\Delta E(\boldsymbol{x}(k;t)) = E(\boldsymbol{x}(k;t+1)) - E(\boldsymbol{x}(k;t))$, and the threshold of the threshold acceptance is set to $T(\geq 0)$, the formula that judges whether a new state is permitted will be given as follows:

$$\boldsymbol{x}(k;t+1) := \begin{cases} \boldsymbol{x}(k;\, t+1), & \text{if } \Delta E(\boldsymbol{x}(k;t)) \leq T \\ \boldsymbol{x}(k;\, t), & \text{if } \Delta E(\boldsymbol{x}(k;\, t)) > T \end{cases} \tag{32}$$

4.5 Selection of Firing Element

As for the state transition in the combinatorial state space, it is necessary to adopt either an asynchronous system or a linking system. The asynchronous system is a state transition that fires only a single element and changes. The linking system is a state transition that synchronizes and fires a small number of elements. In particular, the state transitions of the linking system in the constrained problem are adopted for the state transition that satisfies constraints.
In this case, there are two methods to select the firing element:

(a) A rule selection system that selects an firing element regularly according to element subscript numerical order, and a random selection system in which it is uniformly selected completely at random,
(b) A selection element fixed system that selects the firing element of the fixed number, and a selection element gradual increase system, which increases and selects the number of selection elements.

In addition, we introduced the concept of reinforcement learning (RL) to the selection of firing elements. An adaptation of the standard one-step Q-learning [3] was used:

$$Q(x_n, x_{n+1}, p) = Q(x_n, x_{n+1}, p) + \\ + \alpha \, [r(n) + \gamma \max_a Q(x_{n+1}, a, p) - Q(x_n, x_{n+1}, p)] \quad (33)$$

where $Q(x_n, x_{n+1}, p)$ reflects the desirability of choosing fuel assembly x_n at position p and fuel assembly x_{n+1} at position $p+1$. a denotes all actions available in x_{n+1}. α and γ are learning and discounting parameters respectively. The reward $r(n)$ is a weighted combination of local and global rewards:

$$r(n) = \varphi \, r_{local}(n) + (1-\varphi) \, r_{global}(n) \quad (34)$$

$$r_{local}(n) = G_k / G_{ave} \quad (35)$$

$$r_{global}(n) = G_k / G_{top} \quad (36)$$

where r_{local} is the local reward, r_{global} is the global reward, φ is a weighting factor, G_k is the fitness of agent k, G_{ave} is the average fitness of all agent at the current generation, and G_{top} is the best fitness until previous generation. We chose a simple $\varepsilon - greedy$ proportional policy.

5. Application to The Core Design of A BWR

We applied the MAA proposed in this paper to the actual core design of a BWR plant with 115 fuel assemblies and 1356 MWe, and the performance was compared with the integrated two-stage optimization technique using improved GA that had been proposed before. This MAA was introduced into the first stage of this integrated two-stage optimization technique. Table 1 shows the parameter of LP optimization in the first stage. $\bar{s}$ is upper generation numbers of LP, $\bar{t}$ is upper generation numbers of CRP, and Sc is interval generations of self-reproduction. Table 2 shows the parameters of CRP optimization in the second stage. Target values in the objective functions are shown in Table 3. The parameters were used 0.1 for α, 0.5 for ε, φ, and 0.9 for γ. In the integrative two-stage optimization using the MAA, we used the value of 30 as the number of agents (M).

Table 1 Parameters of LP optimization in the first stage

M	$\bar{s}$	$\bar{t}$	Sc
30	100	2	10

Table 2 Parameters of CRP optimization in the second stage

Nc	Nt	Pos_{in}	Pos_{out}	Pos_{all}
5	6	60	100	200

Table 3 Target values

$\overline{e_f}(t_1)$	$\overline{e_f}(t_2)$	$\overline{e_f}(t_3)$	$\overline{e_f}(t_4)$	$\overline{e_f}(t_5)$	$\overline{e_f}(t_6)$
0.9980	0.9974	0.9968	0.9963	0.9962	0.9986
$\overline{r}(t)$	$\overline{flcpr}(t)$	$\overline{flpd}(t)$	$\overline{flow_{EOC}^{upper}}$	ε_{LP}	
1.70	0.95	0.95	111%	0.0003	

In applying this multi-agent technique to an actual combinatorial optimization problem, only the single coupling operation is not performed but convergent ability improves by the combination of several coupling operations corresponding to the problem to be solved. In out pretest, the best convergence performance was obtained in the combination of a coupling operation of executing the backward convection coupling calculation of one dimension during nine generations after the forward convection coupling calculation of one dimension during one generation. The following three cases were compared as an agent in the neighborhood.

- Case 1 Random variable coupling
- Case 2 Probable variable coupling by the roulette strategy
- Case 3 Variable coupling by the competition rule

In the case of the backward convection coupling calculation, when each case was adjusted to four as the number of agent's firing elements, the best result was obtained. In the case of the forward convection coupling calculation, the method in which all elements states of the current agent change to all elements states of the neighborhood agent was adopted. In addition, the minimization problem was converted to a maximization problem by using the fitness functions of the Eq. (18) and the Eq. (20) respectively in the first stage and the second stage instead of the minimization function of the Eq. (21).

The results of applying the MAA and conventional two-stage GA to the optimization of the actual core design of a BWR are shown in Table 4. This table shows the generation numbers of the average value, the best value, the worst value, and the standard deviation until reaching the optimum LP in ten trials with a different initial value in each case. The initial values are decided randomly using random numbers. The convergence performance is improved in three cases using the MAA compared with conventional two-stage GA, as shown in Table 4. Moreover, of cases 1-3, the best convergence performance was obtained in case 2 using a probable variable coupling by the roulette strategy.

The transition of entropy, numbers of renewal agents, and fitness function with generation renewal in a case near the average value of case 2 are shown in Fig.4. Entropy [4] is an index of the agent's diversity. As can be seen, the repetition of the process in which a local search is performed concentrating on the best agent by the forward convection coupling, while the agent's diversity is maintained during nine generations by the backward convection coupling and a global search is performed in part, plays the role of keeping the balance between a global search and a local search.

Fig.5 shows an example of optimal LP and Fig.6 shows an example of optimal CRP.

Table 4 Numbers of generations for optimum LP and CRP

	Two-stage GA	Multi-agents + GA		
		Case 1	Case 2	Case 3
Generation Numbers				
Average	48	40	33	42
Worst	74	72	60	68
Best	21	15	13	22
Sigma (%)	5.8	5.5	4.5	4.5

(M = 30, 10 trials / case)

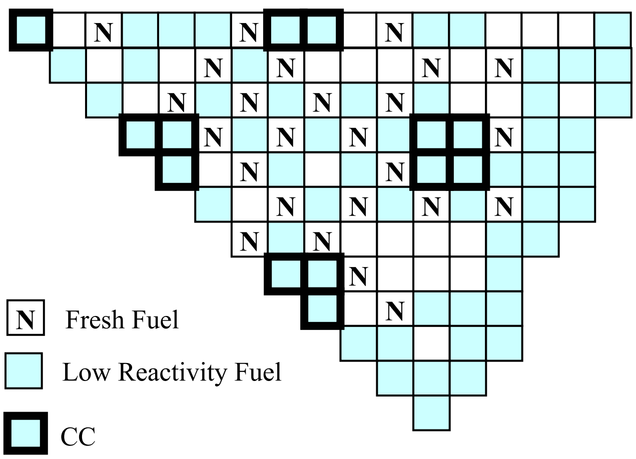

Fig. 5 Example of optimal loading pattern (case of 5CC)

6. Conclusion

In this paper, MAA was applied to the optimization of the automatic core design of a BWR, and the performances were compared with conventional two-stage GA. The convergence performance was considerably better than conventional two-stage GA alone. The reason is that A good solution is searched early by combining the coupling operation using the multi agents. As a result, the design time has decreased by about 30% in the complex automatic core design of a BWR.

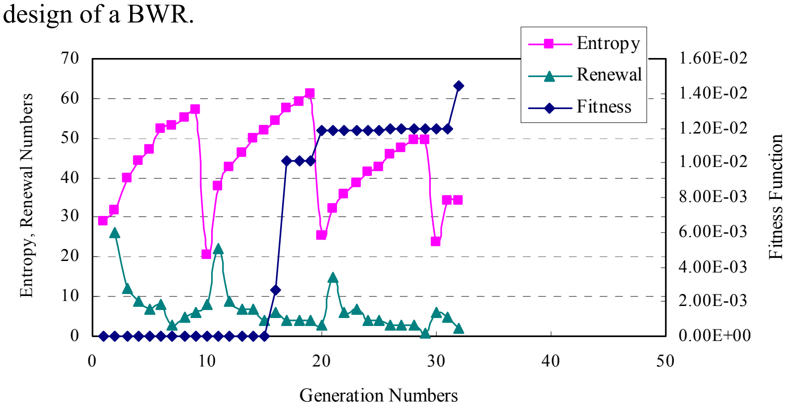

Fig. 4 Transition of entropy, numbers of renewal agents and fitness Function

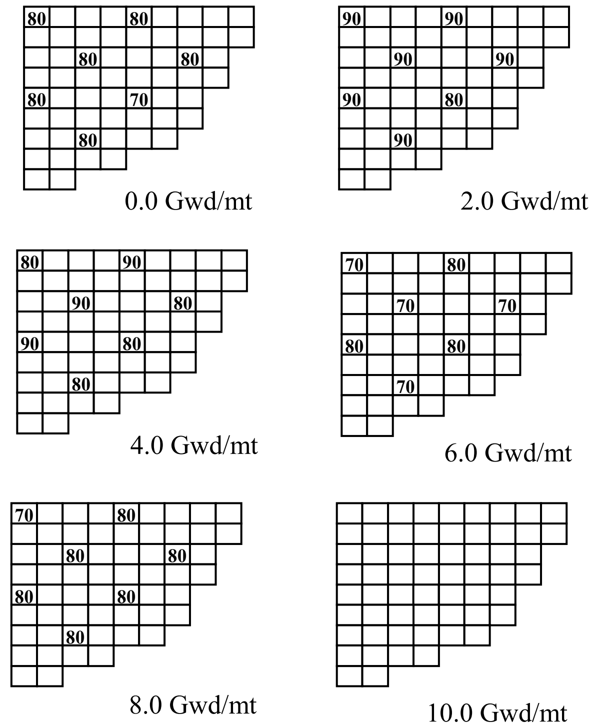

Fig. 6 Examples of optimal control rod pattern

References

[1] Y. Kobayashi, E. Aiyoshi, "Optimization of Boiling Water Reactor Loading Pattern Using Improved Genetic Algorithm," *Proceeding of American Nuclear Society International Topical Meeting on Nuclear Plant Instrumentation Control and Human-Machine Interface Technologies*, Washington DC, November, 134-156, 2000.

[2] Y. Kobayashi, E. Aiyoshi, "Optimization of Boiling Water Reactor Loading Pattern Using Genetic Algorithm," *Nuclear Science and Engineering*, 142,119-139, 2002.

[3] C. Watkins, P. Dayan, "Q-learning," *Machine learning*, 8, 279-292, 1992.

[4] N. Mori, J. Yoshida, H. Tamaki, H. Kita, and Y. Nishikawa, "A Thermodynamical Selection Rule for the Genetic Algorithm," *Proceedings of IEEE ICEC'95*, 188-192, 1995.