Cognitive Diagnosis-based Personalized Exercise Group Assembly via A Multi-Objective Evolutionary Algorithm

Journal:	*IEEE Transactions on Emerging Topics in Computational Intelligence*
Manuscript ID	TETCI-2022-0141
Manuscript Type:	Full Paper
Date Submitted by the Author:	07-Apr-2022
Complete List of Authors:	Yang, Shangshang; Anhui University, School of Artificial Intelligence Wei, Haoyu; Anhui University, School of Artificial Intelligence Ma, Haiping; Anhui University, Institutes of Physical Science and Information Technology Tian, Ye; Anhui University, School of Computer Sciences and Technolgoy Zhang, Xingyi; Anhui University, School of Computer Sciences and Technology Cao, Yunbo; Tencent Technology (Beijing) Co., Ltd, The Center of Natural Language Processing Jin, Yaochu; Bielefeld University, Faculty of Technology; University of Surrey, Department of Computer Science
Keywords:	Exercise group recommendation, assemble exercises, personalized exercise group assembly, multi-objective problem, cognitive diagnosis, co-evolutionary algorithm

Cognitive Diagnosis-based Personalized Exercise Group Assembly via A Multi-Objective Evolutionary Algorithm

Shangshang Yang, Haoyu Wei, Haiping Ma, Ye Tian, Xingyi Zhang, *Senior Member, IEEE*, Yunbo Cao, and Yaochu Jin, *Fellow, IEEE*

***Abstract*—Exercise group recommendation plays an important role in many intelligent education tasks. However, existing approaches make recommendations based on the intrinsic features of exercises without considering students' learning abilities, or make selections from several pre-built exercise groups at the expense of flexibility. Furthermore, although many cognitive diagnosis approaches have successfully revealed students' abilities, how to leverage the diagnosis results for exercise group recommendation is hardly explored. To flexibly recommend suitable exercise groups to students, this paper proposes to assemble personalized exercises as a group based on students' abilities, called Personalized Exercise Group Assembly (PEGA). To solve the PEGA task, we first formulate it as a constrained multi-objective problem (CMOP), where three objectives are designed for enabling the assembled exercises to enhance students' abilities on both weak and new knowledge concepts. Then, we devise an extended neural cognitive diagnosis model to learn student's ability/proficiency on all knowledge concepts to compute the weakness consolidation objective. Besides, we propose a dual-encoding and dual-population based co-evolutionary algorithm to tackle the CMOP, where the main population with binary encoding is used to search which exercises are selected, and the auxiliary population with integer encoding is responsible for accelerating the convergence of the main population via guiding offspring generation of the main population. Experiments on two popular datasets demonstrate the effectiveness of exercises assembled by the proposed algorithm compared to state-of-the-art exercise group recommendation approaches, where our assembled exercises can enhance students' proficiency on both poorly mastered and new knowledge concepts.**

***Index Terms*—Exercise group recommendation, assemble exercises, personalized exercise group assembly, multi-objective problem, cognitive diagnosis, co-evolutionary algorithm**

Manuscript received –. This work was supported in part by the National Natural Science Foundation of China under Grant (U21A20512), in part by the National Natural Science Foundation of China (62107001), in part by the National Key Research and Development Project under Grant (2018AAA0100105), in part by the Anhui Provincial Natural Science Foundation (2108085QF272), in part by Key Program of Natural Science Project of Educational Commission of Anhui Province (KJ2020A0036), in part by The University Synergy Innovation Program of Anhui Province (GXXT-2021-004, GXXT-2020-013), in part by CCF-Tencent Open Fund (RAGR20200121), and in part by the Natural Science Foundation of Qinghai Province under Grant (2020-ZJ-913). *(Corresponding authors: Haiping Ma and Xingyi Zhang.)*

S. Yang, H. Wei and X. Zhang are with the Key Laboratory of Intelligent Computing and Signal Processing of Ministry of Education, School of Artificial Intelligence, Anhui University, Hefei 230039, China (email: yangshang0308@gmail.com;1020179721@qq.com; xyzhanghust@gmail.com).

H. Ma and Y. Tian are with the Key Laboratory of Intelligent Computing and Signal Processing of Ministry of Education, Institutes of Physical Science and Information Technology, Anhui University, Hefei 230601, China, and also with Institute of Artificial Intelligence, Hefei Comprehensive National Science Center, Hefei 231299, China (email: hpma@ahu.edu.cn;field910921@gmail.com).

Y. Cao is with Tencent Technology (Beijing) Co., Ltd, Beijing 100080, China (email: yunbocao@tencent.com).

Y. Jin is with the Faculty of Technology, Bielefeld Unversity, Bielefeld 33619, Germany (email:yaochu.jin@uni-bielefeld.de).

I. Introduction

Exercise group recommendation adaptively selects m exercises from a large exercise pool as a group to each student for improving student's ability/proficiency on the studied knowledge concepts and new knowledge concepts as much as possible, which plays an important role in many adaptive education tasks, such as adaptive homework assignment [1] and adaptive test paper generation [2]. Compared to traditional unified exercising mode that assigns a unified exercise group to all students in the same class, there are two advantages held by the exercise group recommendation. On the one hand, the exercise group recommendation can improve students' learning efficiencies and decrease their learning burdens via reducing some redundant exercises in their recommended exercise groups. One the other hand, the exercise group for each student provided by the exercise group recommendation enables each student to learn and practice on their own pace according to the individual needs and learning ability, which can simulate students' interests in learning [3] to some extent.

Up to now, there have been some exercise group recommendation approaches proposed to recommend suitable exercise groups to students as much as possible, where these existing approaches can be mainly classified into two categories. The first type of approaches [1], [2] commonly recommend a series of exercises as a group to a student mainly based on the intrinsic characteristics of exercises (e.g, exercise difficulty and distinction) without considering students' personalized learning abilities, which leads to the fact that the recommended exercise groups are usually of little avail for the students with strong learning abilities but a bit difficult for the students with weak knowledge proficiency to complete. The

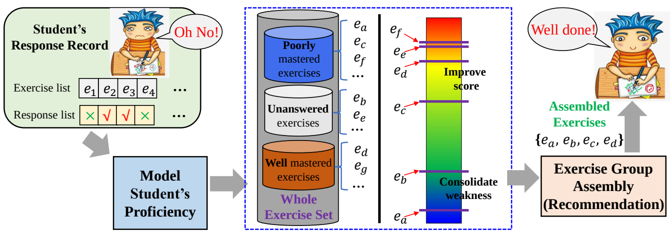


Fig. 1. The illustration of personalized exercise group assembly.

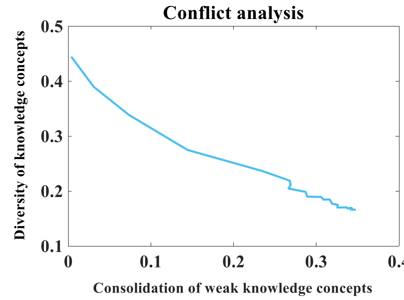


Fig. 2. Illustration of conflict analysis for the first and third objectives.

approaches adopted by recent several famous adaptive learning systems are the second type. These adaptive learning systems commonly provide a dataset for homework assignment or test paper generation consisting of multiple exercise groups with different difficulty levels (e.g., easy, medium, difficult), where the exercise groups are constructed by education experts in advance. Then a simple rule is made to match the students with the homework/test paper according to students' integrated ability levels. However, the selected exercise group is not suitable for each student due to the poor flexibility of the pre-built datasets.

Furthermore, with the rise of online education platforms [4], [5] and computer supported education environment [6], [7], abundant learning materials (e.g., exercises) and adequate students' exercise data can be collected from these learning systems for diagnosing students' abilities. Correspondingly, many cognitive diagnosis (CD) models [8]–[10] have been proposed to exploit these exercising data to diagnose students' learning ability or knowledge proficiency. In spite of the easy access to the diagnosis results of students, how to leverage the diagnosis results for exercise group recommendation has been hardly explored.

To achieve the flexible exercise group recommendation suitable to each student, we propose to assemble personalized exercises as a group with respect to a student's learning ability, which is formally called Personal Exercise Group Assembly (PEGA) in this paper. Fig. 1 summarizes the main procedure of the PEGA. The PEGA first employs a CD model to diagnose a student's ability/proficiency on all knowledge concepts based on the student's response record. Then, an exercise group assembly (recommendation) strategy is used to assemble a group of personalized exercises to the student according to the diagnosed results. As can be seen, the proposed PEGA task is a special case of exercise group recommendation, but there exist two important differences when assembling/recommending a group of exercises in the PEGA. The first point is each student's personalized proficiency, which has a high influence on assembling an exercise group to each student. The second point is the exercise group recommended to each student, which is built via assembling different exercise flexibly instead of selecting an exercise group from multiple pre-built exercise groups.

To tackle the PEGA task well, we formulate the task as a constrained multi-objective optimization problem (CMOP) with three objectives: the weakness consolidation objective, the new knowledge exploration objective, and the knowledge diversity objective. However, in addition to the conflict between proposed three objectives as shown in Fig. 2, there exist two challenges to be overcome when solving the CMOP. The first challenge is the difficulty of learning accurate students' proficiency on all concepts for the computation of weakness consolidation objective. Learning accurate diagnosis results will provide accurate guidance for exercise group assembly, but existing CD model cannot learn the good diagnosis results on the knowledge concepts that the students have not done the associated exercises. The second challenge is the fact that the formulated CMOP is a constrained sparse large-scale multi-objective optimization problem (MOP). As demonstrated in [11], although several multi-objective evolutionary algorithms (MOEAs) have been proposed for constrained large-scale MOPs or sparse large-scale MOPs, there is no MOEA devised for constrained sparse large-scale MOPs.

To overcome the above challenges, we suggest an effective CD model for obtaining student's proficiency and proposes a co-evolutionary algorithm to tackle such a complex CMOP. Specifically, the contributions of this paper are summarized as follows:

1) We are the first to propose the concept of personalized exercise group assembly and devise three objectives to ensure the quality of the assembled exercise group, including weakness consolidation, new knowledge exploration, and knowledge diversity. The first objective is to make students' weak knowledge concepts benefit from the assembled exercise group as much as possible, the second objective is responsible for maintaining more new knowledge concepts in the assembled exercise group, and the last objective is used to ensure the diversity of included knowledge concepts. Equipped with the above objectives, we finally formulate the PEGA task as a multi-objective optimization problem with four constraints, where the first two constraints are caused by the intrinsic nature of the problem and the other two constraints are used to make the student maintain high learning enthusiasm.
2) To address two issues in the formulated CMOP, we

first devise an extended neural cognitive diagnosis model to fuse more latent representations via adding extra hidden layers to NeuralCD model [10], termed ENCD. Based on the proposed ENCD, the student's proficiency on all knowledge concepts can be well diagnosed for the computation of the first objective (i.e., weakness consolidation) in the CMOP. Then, a **co-e**volutionary algorithm based on **d**ual-en**c**oding and **d**ual-**p**opulation (CoE-DCDP for short) is proposed to tackle the formulated CMOP. To be specific, a main population encoded by binary vectors is used to search which exercises are selected, while an auxiliary population with integer encoding is responsible for accelerating the convergence of the main population via guiding the offspring generation of the main population.

3) Experimental results on two widely used datasets, namely ASSISTments and JunYi, demonstrate the effectiveness of exercises assembled by the proposed CoE-DCDP compared with state-of-the-art exercise group recommendation approaches. Besides, it is verified that our assembled exercise groups can enhance students' proficiency on not only poorly mastered knowledge concepts but also new never saw knowledge concepts. Moreover, the effectiveness of the proposed cognitive diagnosis model ENCD and the proposed co-evolutionary algorithm CoE-DCDP by comparing with state-of-the-art CD approaches and state-of-the-art MOEAs, where the ENCD can provide better prediction results than its competitors and the employed auxiliary population is validated to accelerate and improve the convergence of the main population indeed significantly.

The rest of the paper is organized as follows. Section II will present the background and the related work, and Section III gives the mathematical formulation of the PEGA task and the proposed extended neural CD model. The proposed approach is elaborated in Section IV, followed by the empirical evaluations presented in Section V. Finally, conclusions are drawn in Section VI.

II. Background and Related Work

In this section, we first present the background about the proposed PEGA task, then the related work on cognitive diagnosis and the related work on large-scale evolutionary multi-objective optimization are introduced consecutively, which are crucial techniques for the proposed approach to solve the CMOP formulated for the PEGA task.

A. The PEGA Task

The more relevant tasks to our work are homework assignment [1] and test paper generation [2]. The traditional approaches mainly focus on the exercise features, e.g., difficulty, distinction and do not consider customized homework or test with respect to student abilities. In recent several adaptive learning systems, the homework or test papers with different difficulty levels are constructed by education experts. A simple rule is made to match the students with the homework/test paper according to their integrated ability levels. Different from these existing work, we attempt to auto-generate a homework or test paper for each student according to their detailed knowledge proficiency.

Personalize exercise recommendation with respect to students' learning ability is also related to our work. The main difference among them is that the former just consider the benefit of single exercise while the latter consider the portfolio benefits of m exercises. In addition, these traditional exercise recommendation approaches mainly focus on make decision following one goal. For example, some existing studies [3], [12]–[14] try to discover the weakness of students and then make decision of recommending their non-mastered exercises (i.e., the goal is to address the weakness of students); Some others [15]–[18] attempt to obtain students' learning abilities and the difficulty of exercises or knowledge concepts and then make decisions of recommending the exercises whose difficulty is suitable for the students (i.e., the goal is to maintain the learning enthusiasm of students).

In recent years, some researchers have pointed out that making recommending decisions of simply considering one goal is reasonable but ignores the long-term learning needs of students in practice [17], [19]. For example, it is not optimal that the exercise recommender system would tend to keep pushing the student to practice non-mastered exercises including even the ones that the student cannot deal with. Almost these studies firstly design the task-oriented optimization objectives and then linearly merging these objectives with manually-designed weight as a single objective optimization problem. For example, Huang *et al.* [17] proposed a novel deep reinforcement learning framework for recommending exercise learning path, which linearly merges three objectives (i.e., Review & Explore, Smoothness of difficulty level and Engagement) as the reward function. Li *et al.* [15] defined five objectives (e.g., the difference of students' learning ability and difficulty of learning materials) for extracurricular learning materials recommendation and linearly merging these as an optimization objective, and then exploited an evolutionary algorithm named FBPSOS to search the suitable learning materials for each student. However, how to make strategies and design optimization objectives for adaptively recommending m exercises from a larger exercise bank is less explored. Moreover, these single-objective optimization methods are not suitable for such complex combinatorial optimization problem.

B. Cognitive Diagnosis (CD)

CD refers to the diagnosis and evaluation of individual knowledge structure, processing skills or cognitive process (collectively referred to as attribute). Specifically, in intelligent education systems, the purpose of CD is to

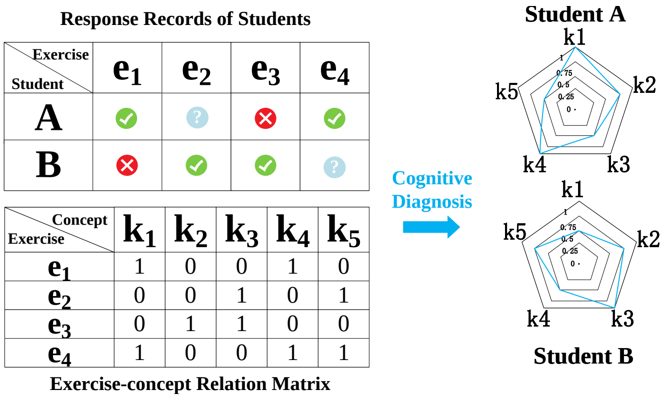


Fig. 3. Illustration of cognitive diagnosis in intelligent education systems.

diagnose the students' proficiency on specific knowledge concepts according to the students' response records. As shown in Fig. 3, students $\{A, B\}$ have practiced a series of exercises and the response matrix is shown in upper left, where check and cross represent the right and wrong response respectively, question mark means the student does not answer this exercise; the exercise-concept relational matrix is shown in bottom left; then, the students' proficiency on the corresponding knowledge concepts is obtained via the CD as shown in the right.

In the context of educational psychology and data mining, many cognitive diagnostic models have been proposed to discover students' knowledge proficiency and predict students' exercising performance. These existing CD models can be generally grouped into two genres. The first genres of approaches usually portray the student learning ability by a hidden vector, where the hidden vector is a one-dimensional ability variable. Among these, Item response theory (IRT) [20], Multidimensional IRT (MIRT) [21] and Matrix Factorization (MF) [22] are the typical representatives of the first genres. The second genres of approaches portray the student learning ability by a knowledge proficiency vector (each dimension corresponds to a specific knowledge concept). Due to the better interpretability provided by the proficiency levels over specific knowledge concepts than the first genres, the recent work mainly focus on the second genres, where the typical representatives include Neural Cognitive Diagnosis Model (NeuralCDM) [10] and Deep Item Response Theory (DIRT) [23].

However, these CD approaches cannot discover the student's learning ability over knowledge concepts that are not related to his/her response records. Although there have proposed Relation map driven Cognitive Diagnosis (RCD) [24] and Concept Aggregation based Cognitive Diagnosis (CDGK) [25] attempted to address this issue by incorporating the prior relations between knowledge concepts (i.e., the prerequisite relation), the annotation of the relations is labor-intensive and costly, which is still difficult for most online learning systems to obtain this information. Therefore, it is necessary to design an effective and efficient CD model, which can diagnose the student's proficiency on all knowledge concepts merely based on the student's response logs. Then the diagnosis results can be used to compute our proposed first objective for the PEGA task.

C. Large-scale Evolutionary Multi-objective Optimization

Since Antonio and Coello first proposed the concept of large-scale multi-objective optimization problems (LSMOPs) in [26], there have been a lot of large-scale MOEAs proposed to solve various types of LSMOPs [11], [27]–[32], including traditional LSMOPs, sparse LSMOPs, constrained LSMOPs, and so on. Specifically, traditional LSMOPs refer to the problems whose number of decision variables is greater than 100, and the representative MOEAs to solve traditional LSMOPs include CSO [33], MOEA/DVA [34], and LSMOF [35]. Compared to traditional LSMOPs, sparse LSMOPs are the special case of traditional LSMOPs, where only a few decision variables of Pareto-optimal solutions are valid (i.e., not equal to 0). The representative MOEAs to solve sparse LSMOPs include SparseEA [36], MOEA/PSL [37], and PM-MOEA [38]. Similarly, constrained LSMOPs refer to traditional LSMOPs equipped with some constraints, which are usually caused by the environment of real-world and make common MOEAs or large-scale MOEAs fail to solve the LSMOPs. The representative MOEAs to solve sparse LSMOPs include SATVaEA [39] for software product configuration, EMRG [40] and CCMO [41] for various vehicle routing problems.

In addition, there are also many large-scale MOEAs proposed to develop common recommendation systems, such as PMOEA [42], NNIA-RS [43], and MOEA-Probs [44]. Note that most of these MOEAs mainly focus on how to build a reasonable and effective MOP for their recommendation tasks, and their formulated MOPs are still solved under classic evolutionary frameworks proposed 15 years ago (e.g., NSGA-II [45] and MOEA/D [46]), which cannot effectively solve sparse LSMOPs or constrained LSMOPs as demonstrated in [36], [37], [41].

However, in this paper, our formulated MOP for the PEGA task is actually a constrained sparse large-scale MOP due to the task's intrinsic characteristics. As stated in [11], there is no MOEA designed for such the kind of problem, which is very difficult for existing MOEAs to solve well. Therefore, it is urgent for us to design an effective MOEA to solve this constrained sparse LSMOP.

III. Mathematical Formulation of PEGA

A. Problem Statement

Suppose there are N students, M exercises, K knowledge concepts provided by an online education system, which can be represented as $\mathcal{S} = \{s_1, s_2, ..., s_N\}$, $\mathcal{E} = \{e_1, e_2, ..., e_M\}$, and $\mathcal{C} = \{c_1, c_2, ..., c_K\}$, respectively. In addition, there is an exercise-concept relation matrix

$Q = (Q_{ij} \in \{0,1\})^{M\times K}$ measuring the relationship between exercise and knowledge concepts, where $Q_{ij} = 1$ means the exercise e_i is related to knowledge concept c_j, and $Q_{ij} = 0$ otherwise. The $\mathcal{R}$ is used to represent the students' exercise response logs and it can be denoted by a set of triple (s_i, e_j, r_{ij}), where r_{ij} is the response score of student s_i got on exercise e_j, $s_i \in \mathcal{S}$, $e_j \in \mathcal{E}$, and $r_{ij} \in \{0,1\}$. Here $r_{ij} = 1$ indicates student s_i answered correctly on exercise e_j and $r_{ij} = 0$ otherwise.

The personalized exercise group assembly aims to select m exercises from the exercise bank $\mathcal{E}$ as a group to each student, then improve each student's knowledge proficiency on all knowledge concepts based on the m exercises recommended by an algorithm. The recommended exercise group consisting of m exercises for student s_i can be denoted by

$$U_i = \{u_i^1, u_i^2, \cdots, u_i^j, \cdots, u_i^m\},\ 1 \le m \le M, \tag{1}$$

where $u_i^j \in \mathcal{E}$ and it can be also represented by the following binary vector as

$$\mathbf{X}_i = (x_i^0, x_i^1, \cdots, x_i^M) \in \{0,1\}^M, \tag{2}$$

where $x_i = 1$ means the exercise e_i is selected and $x_i = 0$ indicates the exercise e_i is not selected.

In addition, there is an important vector for recording the numbers of exercises in U_i (i.e., $\mathbf{X}_i$) related to knowledge concepts, which is defined as

$$\mathbf{n}_i = (n_i^1, n_i^2, \cdots, n_i^K) \tag{3}$$

where each bit $\mathbf{n}_i^k$ ranges from 0 to m and indicates the number of exercises in U_i related to knowledge concept c_k, and it can be computed by

$$\begin{aligned} &\mathbf{n}_{temp} = Q[U_i, :] \text{ or } Q[\mathbf{X}_i, :] \\ &\mathbf{n}_i = row\_sum(\mathbf{n}_{temp}) \end{aligned}, \tag{4}$$

where $Q[U_i, :]$ and $Q[\mathbf{X}_i, :]$ are the operations that extract the rows of the matrix Q corresponding to exercises included in U_i (i.e., $\mathbf{X}_i$), and $row\_sum(\cdot)$ is the sum operation to the rows.

B. Formulated Multi-Objective Optimization Problem

In order to make the recommended exercise group U_i can help student s_i enhance student's proficiency on knowledge concepts, especially on these knowledge concepts that student master poorly, the distribution of exercises in U_i related to knowledge concepts should be highly correlated to that of the student s_i's proficiency on the knowledge concepts. Specifically, the poorer a knowledge concept student masters, the more exercises containing this knowledge concept should be recommended, and fewer exercises should be recommended, whose containing knowledge concepts have been well mastered by student s_i.

For this aim, we employ *Pearson Correlation Coefficient* [47] to measure the above correlation as follows:

$$O_1(\mathbf{n}_i, \alpha_i) = \frac{1}{(K-1)} \cdot \sum_{k=1}^{K} \frac{(\alpha_{i,k} - \bar{\alpha_i})}{\sigma\alpha_i} \cdot \frac{(\mathbf{n}_{i,k} - \bar{\mathbf{n}_i})}{\sigma\mathbf{n}_i}, \tag{5}$$

where $\alpha_i \in R^{1\times K}$ is a real vector with length of K representing for the knowledge proficiency of student s_i. $\bar{\alpha_i}$ and $\sigma\alpha_i$ are the mean and standard deviation of α_i, respectively, and it is the same case for $\mathbf{n}_i$. The smaller value of $O_1(\mathbf{n}_i, \alpha_i)$ means the higher correlation, which further indicates that the recommended exercises in U_i contain more knowledge concepts the student masters poorly. Thus we take $O_1(\mathbf{n}_i, \alpha_i)$ as our ***weakness consolidation*** objective (O_1).

In addition to consolidating student's weakness, the recommended exercise group U_i also should provide some exercises, which contain new knowledge concepts that the student never saw as many as possible. To this end, we propose another objective: ***new knowledge exploration*** (O_2), which can be represented by

$$O_2(\mathbf{n}_i) = \frac{sum(\mathbf{n}_i \ne 0) - sum(\ (\mathbf{n}_i \ne 0) \odot \mu_i)}{K - sum(\mu_i\ )}, \tag{6}$$

where $\odot$ is the element-wise product, $sum(\cdot)$ is a sum operation, $(\mathbf{n}_i \ne 0) \in \{0,1\}^{1\times K}$ is based on a logical operation and outputs a binary vector indicating whether knowledge concepts are contained in U_i, and $\mu_i \in \{0,1\}^{1\times K}$ represents whether student s_i has done the corresponding knowledge concepts. The bigger value of $O_2(\mathbf{n}_i)$ means the more new knowledge concepts are included in U_i, which can promote the student to explore new knowledge.

For further improving the diversity of knowledge concepts included in U_i, a ***knowledge diversity*** objective (O_3) is proposed, whose formal definition is as follow:

$$O_3(\mathbf{n}_i) = \frac{\sum_{k=1}^{K} min\,(\mathbf{n}_{i,k}, 1)}{K}. \tag{7}$$

Here, $O_3(\mathbf{n}_i)$ holds bigger value means than better diversity of contained knowledge concepts.

Equipped with the above-introduced objectives, in this paper we formulate the personalized exercise group assembly as a multi-objective optimization problem, which is as follows:

$$\min_{\mathbf{X}_i} F(\mathbf{X}_i, s_i) = \begin{cases} f_1(\mathbf{X}_i, s_i) = O_1(\mathbf{n}_i, \alpha_i) \\ f_2(\mathbf{X}_i) = -O_2(\mathbf{n}_i) \\ f_3(\mathbf{X}_i) = -O_3(\mathbf{n}_i) \end{cases}, \tag{8}$$

where $\mathbf{n}_i$ in computed from $\mathbf{X}_i$ based on (4), $\alpha_i = CD(s_i)$, and thus there is a necessary step to get student s_i's knowledge proficiency via a CD model $CD(\cdot)$.

In addition, due to the intrinsic characteristics of the personalized exercise group assembly, there are four constraints to the formulated problem in (8), which are as follows:

$$\begin{aligned} &\min_{\mathbf{X}_i} F(\mathbf{X}_i, s_i) = (f_1(\mathbf{X}_i, s_i), f_2(\mathbf{X}_i), f_3(\mathbf{X}_i)) \\ &\text{subject to } C_1 : sum(\mathbf{X}_i) = m \\ &\qquad C_2 : sum(\mathbf{X}_i \odot \nu_i) = 0 \\ &\qquad C_3 : l \le x_i^j * y_i^j \le h, \text{for all } x_i^j = 1 \\ &\qquad C_4 : L \le (\frac{\sum_{j=1}^{M} x_i^j * y_i^j}{m}) \le H \end{aligned}, \tag{9}$$

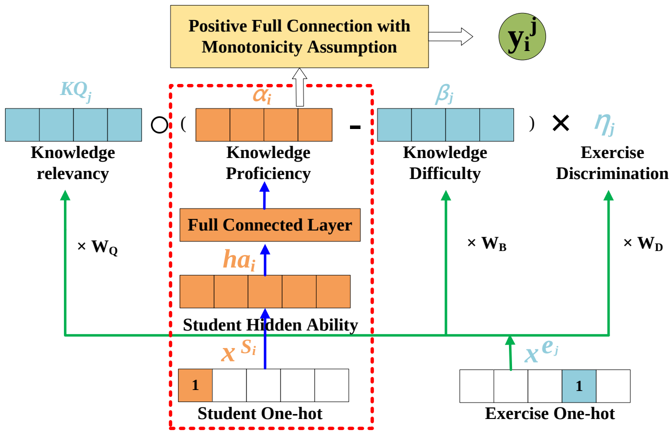


Fig. 4. The proposed extended neural cognitive diagnosis model.

where $\nu_i \in \{0,1\}^{1\times M}$ indicates whether the student s_i has done the corresponding exercise, and y_i^j is the probability of student s_i answering correctly on exercise e_j. The first constraint C_1 ensures the number of selected exercises in U_i equaling to m, the second constraint C_2 enable the exercises that have been done by student s_i not to be selected, the third constraint C_3 make each selected exercise not be too easy or too difficult, and the final constraint C_4 is used to guarantee the overall correct rate of student s_i on the recommended exercise group U_i to be a high level. Note that the first two constraints are basic constrained conditions caused by the innate character of this problem, and other two constraints aim to make the student maintain a high interest in learning.

To solve the CMOP presented in (9) for one student, there are two key challenges need to be overcome. The first is to diagnose an accurate knowledge proficiency α_i for the student based on a CD model, which is necessary for computing the $f_1(\cdot)$ value in (8), but existing CD models cannot provide the diagnosed knowledge proficiency on the knowledge concepts that are not correlated to the student's response records. The second is to design an effective algorithm to solve such a difficult constrained multi-objective problem, where common algorithms fail to solve the constraints in (9). In the following subsection, we will present the proposed extended neural CD model for the first challenge, and we will introduce a co-evolutionary algorithm to overcome the second challenge in Section IV.

C. Proposed Extended Neural CD

Due to the sparse interactions between a student and knowledge concepts, existing CD approaches cannot diagnose the student's proficiency on all knowledge concepts, and the diagnosed proficiency will lead to the inaccurate computation of (8). To address this issue, we design a new CD model by extending the NeuralCD [10], which is named extended neural CD (termed ENCD).

Instead of directly modeling α_i via a simple embedding layer that multiplies the student's one-hot vector $\mathbf{x}^{s_i} \in \{0,1\}^{1\times N}$ with a trainable matrix, as shown in the red dotted box of Fig. 4, we employ two fully connected (FC) layers to capture the complex latent representation for the student's knowledge proficiency, which can be defined as follows:

$$\begin{aligned} \mathbf{ha}_i &= FC_1(\mathbf{x}^{s_i}),\ \mathbf{ha}_i \in R^{1\times d} \\ \alpha_i &= FC_2(\mathbf{ha}_i),\ \alpha_i \in R^{1\times K} \end{aligned}, \tag{10}$$

where d is the embedding size, $\mathbf{ha}_i$ is the student's hidden ability representation vector, and α_i is the student's knowledge proficiency on all knowledge concepts.

Except for the computation of student's knowledge proficiency, the other parts of the proposed ENCD keep same with that of NeuralCD. To be specific, the knowledge relevancy KQ_j, knowledge difficult β_j, and exercise discrimination η_j can be obtained by

$$\begin{aligned} KQ_j &= \sigma(\mathbf{x}^{e_j} \times \mathbf{W_Q}), \mathbf{W_Q}^{M\times d} \\ \beta_j &= \sigma(\mathbf{x}^{e_j} \times \mathbf{W_B}), \mathbf{W_B}^{M\times d} \\ \eta_j &= \sigma(\mathbf{x}^{e_j} \times \mathbf{W_D}), \mathbf{W_D}^{M\times 1} \end{aligned}, \tag{11}$$

where $\mathbf{W_Q}$, $\mathbf{W_B}$, and $\mathbf{W_D}$ are trainable matrices, $\mathbf{x}^{e_j} \in \{0,1\}^{1\times M}$ denotes the exercise's one hot vector, $\sigma(\cdot)$ is the Sigmoid [48] activation function.

With the above obtained elements, the probability of student s_i answering correctly on exercise e_j is computed by the final diagnosis layer as

$$\begin{aligned} temp &= KQ_j \odot (\alpha_i - \beta_j) \times \eta_j \\ y_i^j &= FC_5(FC_4(FC_3(temp))) \end{aligned}, \tag{12}$$

where the first layer is used to get the aggregation results, and three FC layers are used to get the final output y_i^j.

To train the proposed ENCD, the Cross-Entropy loss function [49] is adopted:

$$\text{loss} = -\sum_i \sum_j \left(r_{ij} \log y_i^j + (1 - r_{ij}) \log \left(1 - y_i^j\right)\right). \tag{13}$$

Same as the NeuralCD, the training process is completed with a standard optimizer. After training, the value of α_i is the diagnosis result for each student, which denotes the student's proficiency on all knowledge concepts.

IV. The Proposed CoE-DCDP

In this section, we first present the overall framework of the proposed CoE-DCDP for solving the formulated CMOP. Then, some important parts of the proposed CoE-DCDP will be elaborated. Finally, other related techniques will be briefly introduced.

A. Overall Framework of CoE-DCDP

After training the proposed ENCD (denoted by $ENCD(\cdot)$) based on students' response logs R, the proposed CoE-DCDP receives one student each time as the input and outputs one group of exercises (i.e.,

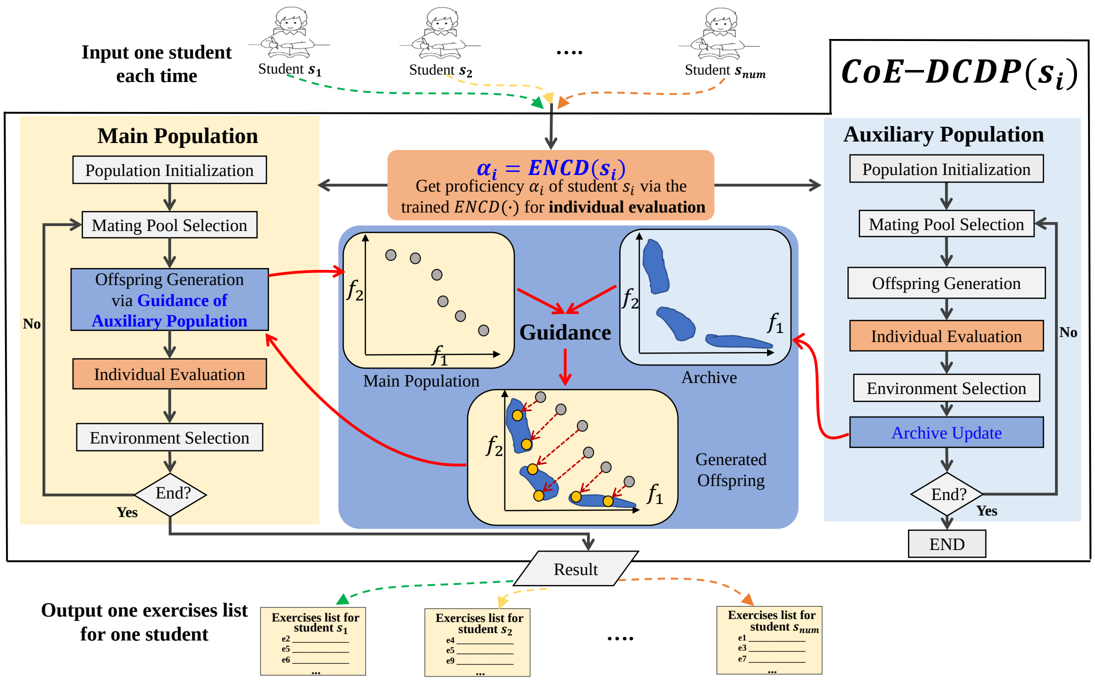


Fig. 5. The overall procedure of the proposed CoE-DCDP for each student, where $ENCD(\cdot)$ is trained based on all students' response logs.

exercises list) for the student. Fig. 5 presents the overall framework of the proposed CoE-DCDP containing two populations: a main population and an auxiliary population, where the main population is equipped by exercise binary encoding for fine search (i.e., exploitation), and the auxiliary population is equipped with knowledge concept integer encoding for rough search (i.e., exploration).

First, the student's knowledge proficiency α_i is diagnosed via the trained ENCD, where the obtained α_i is used for individual evaluation in two populations. Second, the main population and the auxiliary population start to evolve, where two populations are initialized by different initialization strategies. Third, a binary tournament selection is employed to mate individuals for both the main population and the auxiliary population, respectively. Next, the auxiliary population will execute offspring generation, individual evaluation, environment selection, and archive update, while the main population is blocked to wait for the archive updated by the auxiliary population, where the Pareto-optimal individuals are kept. Fifth, the mated parent individuals in the main population generate offspring individuals based on the guidance of the individuals in the archive, and each offspring individual needs to be evaluated. Finally, the environment selection will be applied to two populations to maintain potentially good individuals. The second to the sixth step will repeat until the evolution termination criterion is satisfied, after which one individual will be output as the exercises list for the input student. For details, Algorithm 1 demonstrates main steps of the proposed CoE-DCDP.

Algorithm 1: Main Steps of CoE-DCDP

Input: Gen: Maximum number of generations; Pop: Population size; s_i: A student's one-hot vector;
Output: $\mathbf{P}^i_m$: One individual in the main population;
$\alpha_i \leftarrow ENCD(s_i)$; % `Get student's proficiency via trained ENCD`
$\mathbf{P}_m, \mathbf{P}_a, A \leftarrow$ Initialization$_1(Pop)$, Initialization$_2(Pop)$, $\emptyset$; % `Initialize two populations and set the archive`
for $g = 1$ to Gen **do**
  Update $Ratio$ via a parameter adaption strategy;
  $\mathbf{P}'_m, \mathbf{P}'_a \leftarrow$ Select Pop parents from $\mathbf{P}_m$ and $\mathbf{P}_a$, respectively; % `Mating pool selection`
  ———**Auxiliary population**———
  $\mathbf{Off}_a \leftarrow$ Genetic Operator($\mathbf{P}'_a$);
  Evaluate individuals in $\mathbf{Off}_a$ according to (8) and α_i; $\mathbf{P}_a \leftarrow$ EnvironmentSelection($\mathbf{P}_a \cup \mathbf{Off}_a$);
  $A \leftarrow$ ArchiveUpdate($A \cup \mathbf{P}_a$);
  ———**Main population**———
  $\mathbf{Off}_m \leftarrow$ Guidance-based Genetic Operator($\mathbf{P}'_m, A, Ratio, Pop$); % `Algorithm 2`
  Evaluate individuals in $\mathbf{Off}_M$ according to (8) and α_i;
  $\mathbf{P}_m \leftarrow$ EnvironmentSelection($\mathbf{P}_m \cup \mathbf{Off}_m$);
end for
Select one individual $\mathbf{P}^i_m$ from $\mathbf{P}_m$ as the output
return $\mathbf{P}^i_m$;

B. Encoding Strategy

There are two types of encoding respectively used for two populations, including the exercise binary encoding and the knowledge concept integer encoding. The exercise binary encoding employs a binary vector $\mathbf{X}_i$ in (2), while an integer vector $\mathbf{n}_i$ in (3) is used for the knowledge concept integer encoding.

For better observation, Fig. 6 presents an illustrative example of two types of encoding. As can be seen, the recommended $m = 5$ exercises $U_i = \{e_1, e_2, e_6, e_8.e_{10}\}$ can be represented by its corresponding exercise binary

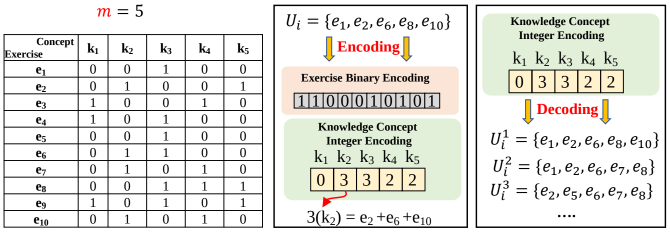


Fig. 6. An illustrative example of proposed encoding strategy.

encoding $(1, 1, 0, 0, 0, 1, 0, 1, 0, 1)$ and knowledge concept integer encoding $(0, 3, 3, 2, 2)$, where the second bit of integer encoding equaling to 3 means that there are 3 exercises related to the second concept k_2 including e_2, e_6, and e_{10}. In contrast to the one-to-one correspondence between the exercise binary encoding and the exercises set U_i, one knowledge concept integer encoding can be decoded into multiple exercises sets, which enables the population encoded by the knowledge concept integer easily to get trapped from local optimal and explore more new areas.

As a result, the knowledge concept integer encoding is used for the auxiliary population to perform a rough search, which is responsible for searching for the optimal distribution of exercises on knowledge concepts and thus provides fast convergence, while the exercise binary encoding is used for the main population to perform a fine search, which is responsible for determining which exercises are selected as the recommendation and thus provides good diversity.

It is worth noting that the CMOP in (9) will degenerate into the following problem for the auxiliary population:

$$\begin{aligned} &\min_{\mathbf{n}_i} F(\mathbf{n}_i, s_i) = (f_1(\mathbf{n}_i, s_i), f_2(\mathbf{n}_i), f_3(\mathbf{n}_i)) \\ &\text{subject to } C_1 : sum(\mathbf{n}_i) = m \end{aligned}, \tag{14}$$

where only the constraint C_1 is available due to the unique properties of the integer encoding.

C. Guidance-based Offspring Generation

As mentioned above, the characteristics of the knowledge concept integer encoding enable the auxiliary population with faster convergence speed but worse diversity than the main population. To make the best of two populations, we propose a guidance-based genetic operator for the main population to accelerate its convergence as well as hold high diversity. Specifically, individuals in the archive A of the auxiliary population will be employed to guide the crossover and mutation of individuals in the main population, where each offspring individual has to make the distribution of exercises over knowledge concepts be consistent with the corresponding guidance individual that selected the from archive.

Given an offspring individual (denoted by I_{off}) and a guidance individual (denoted by $I_{guidance}$), their knowledge concept integer encoding are Int-En(I_{off}) and Int-En$(I_{guidance})$, where Int-En(I_{off}) is obtained based on (3). Then the difference in the number of exercises over knowledge concepts between I_{off} and $I_{guidance}$ can be computed as follows:

$$\begin{aligned} \textbf{Diff} &= \text{Int-En}(I_{guidance}) - \text{Int-En}(I_{off}) \\ &= (Diff_1, Diff_2, \cdots, Diff_k) \end{aligned}, \tag{15}$$

where $Diff_1$ represents the difference in the number of exercises over knowledge concept k_1 between I_{off} and $I_{guidance}$, and further means there should be $Diff_1$ more exercises that need to be included in I_{off}.

Suppose there are $N_{kj} = sum(Q[:, j])$ exercises related to knowledge concept k_j in exercise set $\mathcal{E}$, I_{off} selects N_{kj}^{off} exercises on knowledge concept k_j, and $I_{guidance}$ holds N_{k1}^{gui} exercises on knowledge concept k_j. Then we propose the following two probabilities of flipping as

$$\begin{aligned} Pof_{kj}^1 &= \frac{p_c * N_{kj} + Diff_1}{2 * N_{kj}^{off}} \\ Pof_{kj}^0 &= \frac{p_c * N_{kj} - Diff_1}{2 * (N_{kj} - N_{kj}^{off})} \end{aligned}, \tag{16}$$

where p_c is the crossover probability commonly set to 0.5, Pof_{kj}^1 is the probability of flipping zeros related to k_j in I_{off} to ones, and Pof_{kj}^0 refers to the probability of flipping ones related to k_j in I_{off} to zeros. As can be observed from Pof_{kj}^1, the big value of $Diff_j$ and small value of N_{kj}^{off} will lead to more zeros flipped to ones, which means that more exercises on concept k_j will be selected in I_{off}, and it is the same case for Pof_{kj}^0.

Next, two random vectors $R0^{1\times(N_{ki}-N_{kj}^{off})}$ and $R1^{1\times N_{kj}^{off}}$ are generated to determine which zeros or ones related to k_j need to be flipped by

$$Ff_{kj}^1 = R0 < Pof_{kj}^1, \ Ff_{kj}^0 = R1 < Pof_{kj}^0, \tag{17}$$

where $Ff_{kj}^1 \in \{0, 1\}^{1\times(N_{ki}-N_{kj}^{off})}$ and $Ff_{kj}^0 \in \{0, 1\}^{1\times N_{kj}^{off}}$ are two flipping flags, and the flag value equaling to 1 means executing the flipping.

Based on the above tailored flipping operation, Algorithm 2 lists the detailed procedure of the proposed guidance-based genetic operator. First, two individuals I_{off} and I_{off1} are selected from mated parent population (Line 3). Then, a randomly generated number $rand$ is compared to $Ratio$ that is the ratio of offspring individuals generated with the guidance of A (Line 4). If $rand > Ratio$, the widely used single-point crossover and bit-wise mutation [36] will be applied to I_{off} and I_{off1}, where a new offspring individual (also denoted by I_{off}) will be generated (Line 5), otherwise a guidance individual $I_{guidance}$ will be randomly selected from archive A (Line 7). Third, the **Diff** between I_{off} and $I_{guidance}$ is computed based on (15) (Line 9). Next, the flipping operation in (17) is executed to I_{off} according to two probabilities Pof_{kj}^1 and Pof_{kj}^0 obtained from (16) (Lines 10-11), and the unflipped bits related to k_j in I_{off} will be randomly flipped to keep consistent with $I_{guidance}$, i.e., $Diff_j = 0$, (Line 12). Note that the fourth step will repeat for all knowledge concepts. Fifth, the generated

Algorithm 2: Guidance-based Genetic Operator()

Input: $\mathbf{P}'_m$:Mated parent population, A: Archive, $Ratio$: Ratio of offspring generated with the guidance of A; Pop: Population size
Output: Off: Offspring;
$Off \leftarrow \emptyset$;
for $i = 1$ to Pop **do**
  $I_{off}, I_{off1} \leftarrow \mathbf{P}'_m[i,:], \mathbf{P}'_m[i+1,:]$;
  if rand $> Ratio$ **then**
    $I_{off} \leftarrow$ Apply single-point crossover and bitwise mutation to I_{off} and I_{off1};
  else
    $I_{guidance} \leftarrow A$[random,:]; % Random selection
    for $j = 1$ to K **do**
      Diff $\leftarrow$Compute the difference via (15);
      Pof^1_{kj}, $Pof^0_{kj} \leftarrow$ Get two probabilities by (16);
      ———**Crossover**———
      Flip zeros and ones according to (17);
      ———**Mutation**———
      Randomly flip rest bits to make $Diff_j$ to be zero;
    end for
  end if
  $I_{off} \leftarrow$Repair constraints of I_{off} by local search;
  $Off \leftarrow Off \cup I_{off}$;
end for
return Off;

I_{off} (from the second step or fourth step) has to repair its constrains by a local search technique, which will be presented in Section IV-E3. Finally, the adjusted I_{off} will be added to offspring population Off.

For better understanding, Fig. 7 gives an illustrative example of the guidance-based genetic operator. Given an individual $I_{off} = (1, 0, 1, 1, 0, 0, 1, 0, 1, 0)$ and a guidance individual $I_{guidance} = (1, 3, 1)$, the knowledge concept integer encoding of I_{off} is obtained as Int-En$(I_{off}) = (3, 1, 2)$. The concept k_1 is related to 4 exercises: e_3, e_4, e_5, and e_9, where corresponding decision variables $(1, 1, 0, 1)$ are extracted. The next thing is to compute the two probabilities of flipping $Pof^1_{k1} = 0$ and $Pof^0_{k1} = \frac{5}{6}$, randomly generate two vectors $R0 = (0.3)$ and $R1 = (0.6, 0.2, 0.5)$, and then obtain two flipping flags $Ff^1_{k1} = (0)$ and $Ff^0_{k1} = (1, 0, 1)$. With the two flags, the corresponding bits of extracted variables $(1, 1, 0, 1)$ are flipped, where new variables $(0, 1, 0, 0)$ are formed. The above procedure is repeated for concept k_2 and k_3, where all newly formed variables $(0, 1, 0, 0)$, $(1, 1, 1)$, and $(1, 0, 0, 0)$ will be merged as a new individual $(1, 1, 0, 1, 0, 0, 1, 0, 0, 1)$. Note that the bits of e_2, e_3, e_9, and e_{10} are flipped.

D. Recommendation Strategy

Due to multiple non-dominated individuals (solutions) found by the MOEA in one run, it is necessary to select one suitable solution for final recommendation. To this end, the solution with the best value of weaknesses consolidation will be finally recommended.

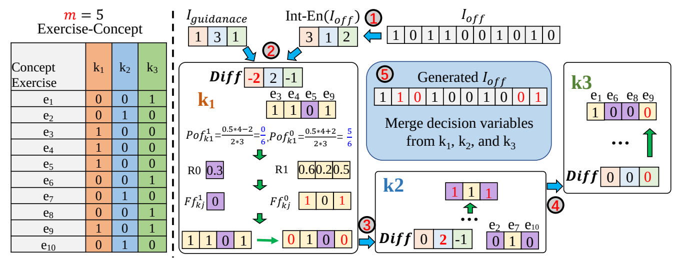


Fig. 7. An illustrative example of guidance-based genetic operator.

E. Other Techniques

In the proposed CoE-DCDP, the mating pool selection and the environment selection are same as that of SPEA2 [50], which demonstrates pretty good effectiveness on the MOPs. In addition, there are three other techniques, including initialization strategy for the auxiliary population, parameter adaption for $Ratio$, and constrains repair technique.

1) Initialization Strategy: For better initial diversity and convergence of the auxiliary population, more exercises for knowledge concepts that the student masters poorly (i.e., weak concepts) should be selected. To achieve this, each individual has to compute the average number of exercises on weak concepts as

$$AvgNum_{weak} = \text{round}\ (\frac{m * U(\cdot)}{\frac{K}{2}})\ , \tag{18}$$

where $U(\cdot)$ is a uniform distribution between 0.5 and 1, and $round(\cdot)$ is the round operation. Then the average number of exercises on strong concepts can be also inferred. All individuals in the auxiliary population are initialized based on the above idea.

2) Parameter Adaptation for Ratio: To achiever a good balance exploration and exploitation, a parameter adaptation strategy for $Ratio$ is designed as

$$Ratio = 0.5 * (1 - \cos((1 - \frac{g}{Gen}) * \pi)), \tag{19}$$

where g is the current generation and Gen is the maximal number of generations. By doing so, the main population depends more on the guidance of auxiliary population in the early evolution stage for fast convergence, while in the late stage the main population will gradually take the advantage of the binary encoding to get trapped from local optimal for good diversity and convergence.

3) Local Search-based Constraints Repair: Due to the simple rules that can deal with constraints C_2 and C_3, there are merely two sub-techniques for repairing constraints C_1 and C_4.

Repair C_1. If the number of selected exercises is more than m, randomly flip extra ones to zeros, otherwise, randomly flip some zeros to ones to satisfy the constraint.

Repair C_4. If the overall correct rate on selected exercises is lower than L, more exercises that are easy to answer will be used to replace these selected exercises that are difficult to answer, otherwise, more difficult exercises are used to replace these selected easy exercises.

TABLE I
DATASET STATISTICS.

Dataset	# Students	# Concepts	# Exercises	# Response logs
ASSISTments	4163	123	17746	324572
JunYi	1000	39	712	202945

The difficulty degree of an exercise is judged according to the probability predicted by the proposed ENCD.

V. EXPERIMENTAL STUDIES

A. Experimental Settings

*1) **Benchmark Datasets**:*

Two widely used datasets ASSISTments[1] and JunYi [51] are used to verity the effectiveness of the proposed CoE-DCDP, whose statistics are summarized in Table I and descriptions are presented as follows:

- ASSISTments is an open available dataset collected and created in 2004 by the ASSISTments online tutoring service system.
- JunYi is an open available dataset gathered from Junyi Academy, which is a Taiwan's education platform built in 2012.

*2) **Performance Metrics** :*

The quality of the recommended exercise group is determined by whether more exercises are selected for student's real weak knowledge concepts and fewer exercises are selected for student's real strong knowledge concepts. However, there is no ground truth due to unavailability of the student's real ability/proficiency, which is different from traditional goods recommendation and movie recommendation, thus traditional metrics such as precision and recall cannot be employed.

To address the issue, we execute a simulation experiment to build two response datasets for ASSISTments and JunYi. Each response dataset is composed of some newly added students and their response logs, where the newly added student's real ability/ proficiency is randomly generated in advance. Specifically, the NeuralCD [10] model is first employed to be trained based on response logs in ASSISTments or JunYi, where there are two types of parameters in NeuralCD: student's ability parameters $Para_a$ and exercise parameters $Para_e$. To generate the response log of a new student s_{new}, then a randomly generated vector is used to replace $Para_a$, which is regarded as s_{new}'s real ability/proficiency. In the following, the NeuralCD with replaced $Para_a$ is used to predict s_{new}'s answer to each exercise e_j of $\mathcal{E}$ in ASSISTments or JunYi, where the output is greater than 0.5 means the answer is correct (i.e., $r_{new\ j} = 1$), otherwise $r_{new\ j} = 0$. Finally, student s_{new}'s response log can be built as $(s_{new}, e_j, r_{new\ j})$. The second to the fourth step will repeat for Num_{add} times to add Num_{add} students and their response logs to original ASSISTments or JunYi. By doing so, the students' proficiency generated in advance can be seen as student's real proficiency (i.e., the ground truth), where Num_{add} students have their own ground truth that can be for algorithm comparison.

Note that the students' real proficiency on each knowledge concept is measured by the value of corresponding bit in the student's ability parameters, where big value stands for good proficiency. The better the student's real proficiency on one knowledge concept is, the fewer the number of exercises related to this knowledge concept should be recommended. Thus, the lower correlation between the student's real proficiency and the distribution of exercises in the recommended exercise group over knowledge concepts is, the better the quality of the recommended exercise group are. To measure such the correlation, three widely adopted metrics are employed, which are as follows:

- *Pearson Correlation Coefficients* [47].
- *Spearman Rank Correlation Coefficients* [52].
- *Kendall'τ Rank Correlation Coefficients* [53].

In addition, there are other metrics that will be briefly introduced in the corresponding sections.

*3) **Compared Approaches** :*

There are three group of state-of-the-art comparison approaches for validation, including exercise group recommendation approaches, MOEAs, and CD approaches.

To verify the effectiveness of the proposed CoE-DCDP, state-of-the-art exercise group recommendation approaches are used to comparing their recommended exercise groups, which are as

- CF: CF [54] is a traditional collaborative filtering recommendation algorithm, which will recommend top-m exercises to the student according to the probability of answering correctly.
- SGL: SGL [55] is a recently published recommendation algorithm based on the graph self-supervised learning, which employs a self-discriminator to strengthen the node representation learning for striking the balance between recommendation accuracy and recommendation diversity.
- MOEA-Probs-ENCD: MOEA-Probs [44] is a traditional MOEA for recommendation. Here only its objectives are replaced by our proposed objectives in (8) that is computed based on ENCD.
- MOEA-Probs-NCD: Same as MOEA-Probs-ENCD, but the objectives in (8) that is computed based on NeuralCD.
- SparseEA-ENCD: SparseEA [36] is an effective MOEA for sparse LSMOPs. Here we directly employ SparseEA to solve the CMOP in (9) that is based on ENCD.
- SparseEA-NCD: Same as SparseEA-ENCD with the objectives computed based on NeuralCD.
- CCMO-ENCD: CCMO [41] is an effective MOEA for constrained LSMOPs under co-evolutionary framework. Here we directly employ CCMO to solve the CMOP in (9) that is based on ENCD.

[1]https://www.etrialstestbed.org/resources/featured-studies/dataset-papers

TABLE II
THE COMPARISON OF BETWEEN THE PROPOSED COE-DCDP AND COMPARED APPROACHES IN TERMS OF THREE CORRELATION METRICS: PEARSON, SPEARMAN, AND KENDALL ON BOTH ASSISTMENTS AND JUNYI, AVERAGED ON 100 STUDENTS. THE LOWER THE CORRELATION IS, THE BETTER THE RESULT IS.

Dataset	Metric	CF	SGL	MOEA-Probs -NCD	SparseEA -NCD	CCMO -NCD	CoE-DCDP -NCD	MOEA-Probs -ENCD	SparseEA -ENCD	CCMO -ENCD	CoE-DCDP
ASSISTments	**Pearson**	-0.0273	-0.0172	-0.0011	-0.0865	-0.0336	-0.1792	-0.1716	-0.2518	-0.1508	-0.5589
	Spearman	0.0186	-0.0496	-0.0118	-0.0793	-0.0448	-0.1698	-0.1628	-0.1949	-0.1187	-0.4771
	Kendall	0.0165	-0.0254	-0.0105	-0.0601	-0.0322	-0.1605	-0.1237	-0.1472	-0.0887	-0.3721
JunYi	**Pearson**	-0.0581	-0.0837	-0.0491	-0.0201	-0.0084	-0.0981	-0.2235	-0.2121	-0.1486	-0.2396
	Spearman	-0.0241	-0.0628	-0.0649	-0.0175	-0.0411	-0.0116	-0.1975	-0.1803	-0.1242	-0.2132
	Kendall	-0.0446	-0.0601	-0.0744	-0.0312	-0.0450	-0.0325	-0.1801	-0.1637	-0.1091	-0.1867

- CCMO-NCD: Same as CCMO-ENCD with the objectives computed based on NeuralCD.
- CoE-DCDP-NCD: The variant of CoE-DCDP, whose objectives computed based on NeuralCD.

To verify the effectiveness of the proposed co-evolutionary algorithm, state-of-the-art MOEAs are used to compare their searched Pareto fronts, which contain two representative EAs: NSGA-II [45] and MOEA/D [46] and four large-scale MOEAs: SparseEA [36], LSMOF [35], CCMO [41], and MOEA/PSL [37].

To verify the effectiveness of the proposed ENCD, state-of-the-art CD approaches are used to compare their prediction results, which consist of DINA [9], MIRT [21], and NeuralCD [10].

*4) **Parameter Settings** :*

Dataset and Model Settings. Only the students whose response records' length is more than 15 will be used for experiments, and all remained students in each dataset are randomly split into 70%/10%/20% for training/validation/testing. Num_{add} is set to 100, which leads to 100 randomly generated students added to each dataset. For the model ENCD, d is set to 20 and the dimensions of final three FC layer are set to 512, 256, and 1, respectively.

Algorithm Parameter Settings. The proposed CoE-DCDP will be executed num times for num students, here num is set to 100. The number of exercises m assembled in one exercise group is set to 100 for ASSISTments and set to 50 for JunYi. Population size Pop is set to 100, the maximal number of generations Gen is set to 400 and 200 for ASSISTments and JunYi, respectively.

For fair comparison, except for the number of evaluations ($Pop \times Gen$), other parameter settings for all compared approaches are same as those in their original papers to hold their best performances.

B. Effectiveness of Exercises Assembled by Proposed CoE-DCDP

To demonstrate the effectiveness of the exercises assembled by the proposed CoE-DCDP, Table II compares the proposed CoE-DCDP with compared approaches in terms of three types of correlations between the student's real ability and the exercises in their recommended exercise groups, averaged over 100 students. As can be observed, the correlation coefficients achieved by the proposed CoE-DCDP are significantly lower than that achieved by other approaches, which indicates that our assembled exercises are more suitable for students to enhance their poorly mastered knowledge concepts.

TABLE III
THE TESTING PERFORMANCE COMPARISON OF THE PROPOSED ENCD AND OTHER CD APPROACHES IN TERMS OF ANSWER PREDICTION.

Dataset	ASSISTments			JunYi		
Model	Accuracy	RMSE	AUC	Accuracy	RMSE	AUC
DINA	0.669	0.477	0.717	0.676	0.461	0.752
MIRT	0.676	0.504	0.710	0.743	0.418	0.777
NeuralCD	0.719	0.439	0.749	0.748	0.408	0.797
ENCD	**0.732**	**0.426**	**0.776**	**0.778**	**0.392**	**0.820**

For deep insight to the results, we take the results of num=100 students on ASSISTments dataset as an illustrate examples in Fig. 8. First, all knowledge concepts are sorted according to the students' real proficiency/ability, where the weak knowledge concept are located in the left of the horizontal axis. Then, we plot the distribution of exercises assembled by the proposed CoE-DCDP and compared approaches, where the vertical axis represents the proportion of recommended exercises for the corresponding knowledge concepts. As can be seen, compared to other nine comparison approaches, the proposed CoE-DCDP has a better distribution of exercises over knowledge concepts, the proposed CoE-DCDP recommend more exercises for the student's weaker knowledge concepts and fewer exercises for the student's stronger knowledge concepts, which is in line with the goal of strengthening students' weaknesses. Therefore, it can be proved that our proposed approach can achieve the goal of consolidating the concept of weak knowledge of students.

For further investigating the competitiveness of the proposed CoE-DCDP, there are two more metrics used for comparison including the novelty and the diversity. Specifically, the novelty metric, computed by (6), measures the coverage of new knowledge concepts included in the recommended exercise group, where high novelty coverage rates are usually preferred that

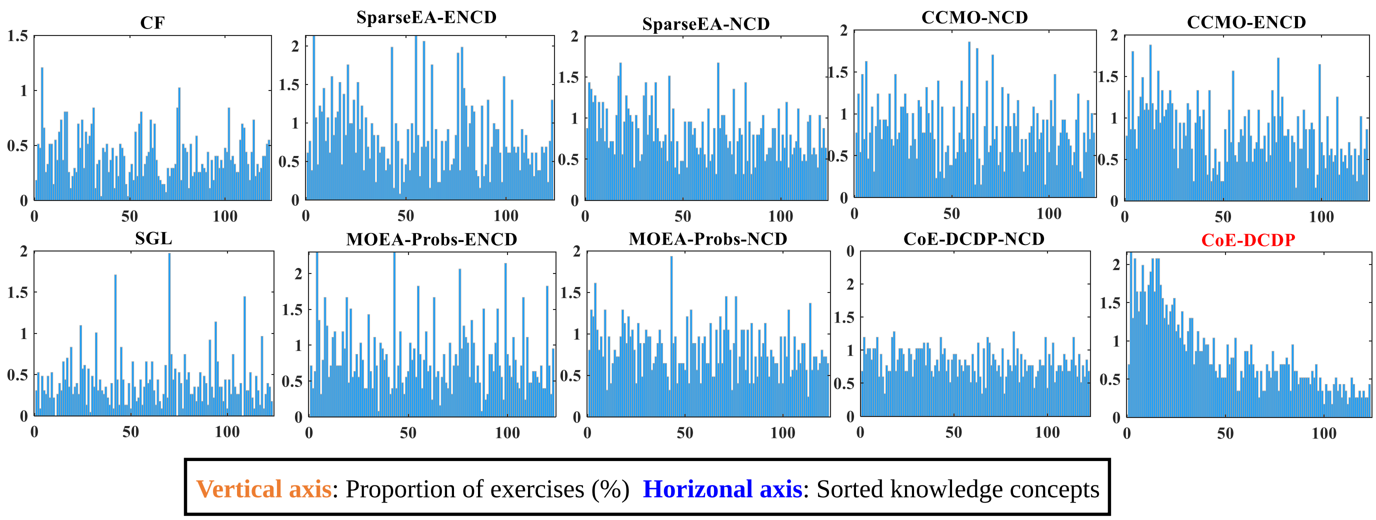
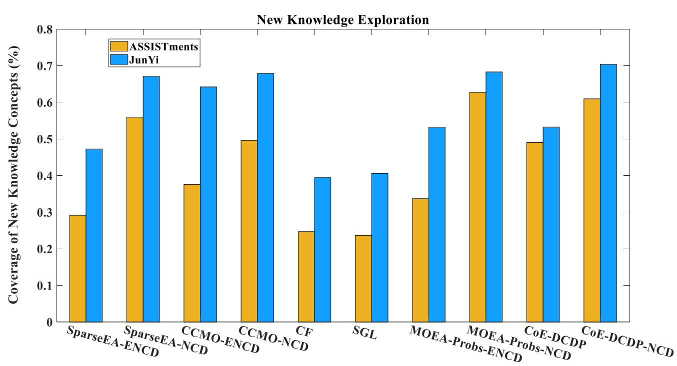
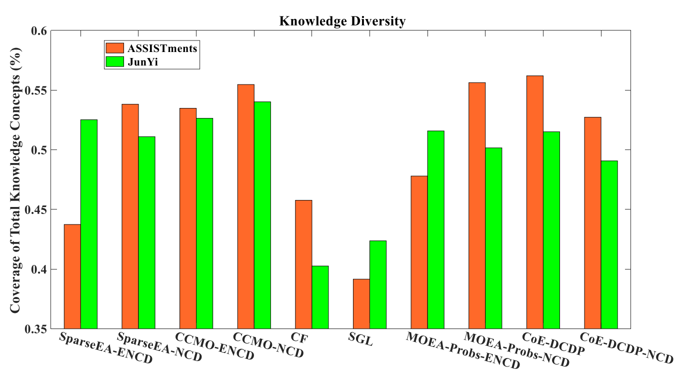


Fig. 8. The comparison of the proposed CoE-DCDP and other compared approaches in terms of the distribution of exercises on all knowledge concepts in their recommended exercise groups.



(a) The novelty comparison.



(b) The diversity comparison.

Fig. 9. The comparison between the proposed CoE-DCDP and comparison approaches in terms of the novelty and diversity metrics.

means more new knowledge concepts are introduced, while the diversity metric, computed by (7), measures the coverage of total knowledge concepts included in the recommended exercise group, where high diversity coverage rates are usually preferred. Fig. 9 shows the comparison of the proposed CoE-DCDP and compared approaches in terms of the novelty and diversity, it can be seen that our approach holds competitively high novelty coverage rates and diversity coverage rates to compared approaches.

In summary, the proposed CoE-DCDP is demonstrated to be superior over or competitive to state-of-the-art exercise group recommendation approaches in terms of correlation, novelty, and diversity of the recommended exercise group.

C. *Effectiveness of The Proposed ENCD*

As can be seen from Table II, all NeuralCD-based approaches including MOEA-Probs-NCD, SparseEA-NCD, CCMO-NCD, and CoE-DCDP-NCD are worse than their own variant approaches, i.e., MOEA-Probs-ENCD, SparseEA-ENCD, CCMO-ENCD, and CoE-DCDP, which are based on the proposed ENCD. The above observation proves the effectiveness of the proposed ENCD to some extent, and the same conclusion can be also obtained from Figs. 8 and 9.

To further investigate the effectiveness of ENCD on cognitive diagnosis, we compare the proposed ENCD with other three state-of-the-art CD approaches, including DINA, MIRT, and NeuralCD. Table III summarizes their testing performances on predicting student's answer to exercises in terms of three metrics [10]: accuracy, RMSE (root mean square error), and AUC (area under curve). It can be seen from Table III that the proposed ENCD outperforms than DINA, MIRT, and NeuralCD in terms of all indicators on both two datasets, which further demonstrates the effectiveness of our proposed ENCD.

To sum up, the suggested ENCD is effective in not only assisting our approach for the PEGA task but also diagnosing student's proficiency.

D. *Effectiveness of The Proposed CoE-DCDP*

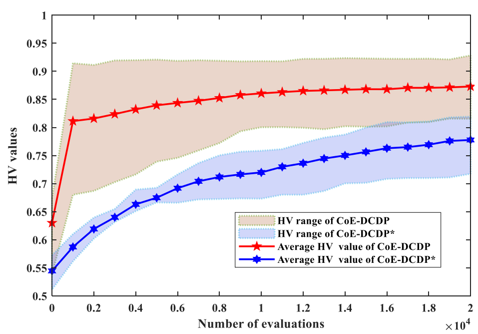
As can be also found in Table II, despite the fact that same CD models (i.e., NeurealCD and ENCD) are

TABLE IV
STATISTICAL RESULTS OF AVERAGE HV OBTAINED BY NSGA-II, MOEA/D, SPARSEEA, LSMOF, MOEA/PSL, CCMO, AND THE PROPOSED COE-DCDP ON THE ASSISTMENTS AND JUNYI. BEST RESULT IN EACH ROW IS HIGHLIGHTED.

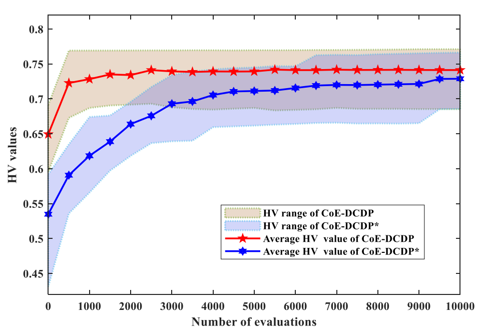
Dataset	#Exercise (m)	NSGA-II	MOEA/D	SparseEA	LSMOF	CCMO	MOEA/PSL	CoE-DCDP
ASSISTments	50	0.4057	0.4591	0.5592	0.4475	0.4239	0.4817	0.5860
	100	0.5015	0.6042	0.6501	0.6197	0.5684	0.6258	0.8817
JunYi	30	0.5716	0.5988	0.7482	0.6717	0.6446	0.6519	0.7570
	50	0.7145	0.7242	0.8121	0.7746	0.7287	0.7894	0.8246

TABLE V
STATISTICAL RESULTS OF AVERAGE HV OBTAINED BY THE PROPOSED COE-DCDP AND ITS VARIANT ON TWO DATASETS, AVERAGED ON 100 STUDENTS. BEST RESULT IN EACH ROW IS HIGHLIGHTED.

Dataset	# Exercise (m)	CoE-DCDP*	CoE-DCDP
ASSISTments	50	0.7993	0.8247
	100	0.7445	0.8732
JunYi	30	0.7367	0.7570
	50	0.5103	0.5757



(a) ASSISTments dataset (m=100)



(b) JunYi dataset (m=50)

Fig. 10. Effectiveness verification of the auxiliary population in the proposed approach on two datasets.

used for MOEAs, the performance of CoE-DCDP and CoE-DCDP-NCD are significantly better than SparseEA-variants (SparseEA-NCD and SparseEA-ENCD), CCMO-variants (CCMO-NCD and CCMO-ENCD), and MOEA-Probs-variants (MOEA-Probs-NCD and MOEA-Probs-ENCD), which demonstrates that the proposed CoE-DCDP is a more effective evolutionary algorithm than SparseEA and MOEA-Probs for solving the formulated CMOP.

For a deeper insight into the effectiveness of the proposed CoE-DCDP, Table IV shows the comparison of the proposed CoE-DCDP and the compared state-of-the-art MOEAs in terms of average hypervolume (HV) [56] values on ASSTISments and Junyi, where the HV metric measures convergence and diversity of a population and a large HV value indicates a good convergence and diversity. Note that two settings for the number of recommended exercises m is used for fair comparison, where m is set to 50 and 100 for ASSISTments and m is set to 30 and 50 for JunYi. As can be observed from Table IV, the proposed algorithm is superior to all comparison algorithms under different numbers of recommended exercises on both ASSISTments and Junyi datasets. The effectiveness of the proposed CoE-DCDP is mainly attributed to the designed co-evolutionary framework that employs two populations equipped with two types of encoding, where the auxiliary population can accelerate the convergence of the main population to hold a better convergence and diversity than a single main population.

To demonstrate this, a CoE-DCDP's variant (denoted by CoE-DCDP*) is used for the subsequent comparison, where the CoE-DCDP* has only a main population without the help of the auxiliary population. For fast validation, the number of generations are set to 200 and 100 for ASSISTments and JunYi. Table V compares the HV values achieved by the CoE-DCDP and CoE-DCDP* under two exercise number settings on both two datasets, it is obvious that HV values obtained by the proposed CoE-DCDP are better than that obtained by CoE-DCDP*, which indicates that better final convergence and diversity can be achieved with the help of the auxiliary population.

To further investigate what specific effects the auxiliary population poses on the proposed CoE-DCDP, Fig. 10 depicts the convergence profiles of HV obtained by the two approaches on two datasets. Here only one setting is used to obtain the results of 100 students, where m is set to 100 and 50 for ASSISTments and JunYi. It is worth noting that the red and blue lines are HV convergence profiles averaged on 100 students (100 runs for each algorithm), and the light red and light blue area denote the region of HV values. Two important facts can be observed: first, the proposed CoE-DCDP demonstrates high convergence speed at the early evolution stage, which proves that the auxiliary population indeed help the main population to accelerate its convergence. Second, the HV values achieved by the CoE-DCDP are not only better but also more stable than that achieved by the CoE-DCDP*, which indicates that our devised co-evolutionary framework is more robust to different students with different abilities. Therefore, the auxiliary population can significantly accelerate and improve the convergence of diversity of the proposed CoE-DCDP.

To summarize, not only the proposed CoE-DCDP is superior to state-of-the-art MOEAs, but also the employed auxiliary population are effective is improving the performance of CoE-DCDP.

VI. Conclusion and Future Work

In this paper, we proposed the task of personalized exercise group assembly to flexibly recommend a suitable exercise group to a student with respect to her/his learning ability or knowledge proficiency. To tackle the task well, we designed three objectives and thus formulated the task as a constrained multi-objective optimization problem. Then, an extended neural CD model was proposed to diagnose student's proficiency on all knowledge concept for the computation of the first objective, and a dual-encoding and dual-population based co-evolutionary algorithm was also proposed to solve the formulated constrained sparse large-scale MOP. Experimental results verified the effectiveness of exercises assembled by the proposed approach. In the future, we will be devoted to problem formulation and optimization approach design for various exercise group recommendation situations in the actual learning process. In addition, we would like to extend our approach and further propose more generic approaches to solve other constrained sparse large-scale multi-objective optimization problems.

References

[1] P. Kotas and J. Finck, "Collaborative learning and other successful strategies for on-line homework," in *Proceedings of the 2002 International Conference on Computers in Education*. IEEE, 2002, pp. 1068–1069.
[2] Z. Lou, "An intelligent teaching test paper generation system based on ant colony hybrid genetic algorithms," in *Proceedings of the 2021/13th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA)*. IEEE, 2021, pp. 467–470.
[3] Z. Li, H. Hu, Z. Xia, J. Zhang, X. Li, J. Shi, H. Li, and X. Li, "Exercise recommendation algorithm based on improved collaborative filtering," in *Proceedings of the 2021 International Conference on Advanced Learning Technologies (ICALT)*. IEEE, 2021, pp. 47–49.
[4] B. Liu, H. Chen, and H. Junmei, "Research on the platform of online education platform based on cloud computing," in *Proceedings of the 2019/14th International Conference on Computer Science & Education (ICCSE)*. IEEE, 2019, pp. 22–25.
[5] X.-d. Chen, J. Lu, M. Gong, B.-J. Guo, and Y.-p. Xu, "Design and implementation of decentralized online education platform," in *Proceedings of the 2020/5th International Conference on Mechanical, Control and Computer Engineering (ICMCCE)*. IEEE, 2020, pp. 970–974.
[6] J. A. Muñoz-Cristóbal, I. M. Jorrín-Abellán, J. I. Asensio-Pérez, A. Martínez-Monés, L. P. Prieto, and Y. Dimitriadis, "Supporting teacher orchestration in ubiquitous learning environments: A study in primary education," *IEEE Transactions on Learning Technologies*, vol. 8, no. 1, pp. 83–97, 2015.
[7] R. Ford, J. Bondzie, and P. Kitcho, "Java applets for microelectronics education," *IEEE Transactions on Education*, vol. 44, no. 2, pp. 10 pp.–, 2001.
[8] Liu, Qi, Runze, Chen, Enhong, Guandong, Yu, Zhigang, and Guoping, "Fuzzy cognitive diagnosis for modelling examinee performance," *ACM Transactions on Intelligent Systems and Technology*, vol. 9, no. 4, pp. 296–296, 2018.
[9] D. L. Torre and J., "Dina model and parameter estimation: A didactic," *Journal of Educational and Behavioral Statistics*, vol. 34, no. 1, pp. 115–130, 2009.
[10] F. Wang, Q. Liu, E. Chen, Z. Huang, Y. Chen, Y. Yin, Z. Huang, and S. Wang, "Neural cognitive diagnosis for intelligent education systems," in *Proceedings of the 2020 AAAI Conference on Artificial Intelligence*, vol. 34, no. 04, 2020, pp. 6153–6161.
[11] Y. Tian, L. Si, X. Zhang, R. Cheng, C. He, K. C. Tan, and Y. Jin, "Evolutionary large-scale multi-objective optimization: A survey," *ACM Computing Surveys (CSUR)*, vol. 54, no. 8, pp. 1–34, 2021.
[12] Y. Huo, J. Xiao, and L. M. Ni, "Towards personalized learning through class contextual factors-based exercise recommendation," in *Proceedings of the 2018/24th IEEE International Conference on Parallel and Distributed Systems (ICPADS)*, 2018.
[13] S. Zhang, J. Cai, B. Zhuge, L. Dong, and X. Jiang, "Accurate exercise recommendation based on multidimension -al feature analysis," *Journal of Physics Conference Series*, vol. 1544, p. 012048, 2020.
[14] Z. Li, H. Hu, Z. Xia, J. Zhang, X. Li, Z. Wang, X. Huang, S. Zeng, and B. Qiu, "Exercise recommendation method based on machine learning," in *Proceedings of the 2021 International Conference on Advanced Learning Technologies (ICALT)*. IEEE, 2021, pp. 50–52.
[15] L. I. Quan, X. Liu, X. U. Xinhua, and S. Lin, "Personalized test question recommendation method based on unified probalilistic matrix factorization," *Journal of Computer Applications*, 2018.
[16] Z. Wu, M. Li, Y. Tang, and Q. Liang, "Exercise recommendation based on knowledge concept prediction," *Knowledge-Based Systems*, vol. 210, p. 106481, 2020.
[17] Z. Huang, Q. Liu, C. Zhai, Y. Yin, E. Chen, W. Gao, and G. Hu, "Exploring multi-objective exercise recommendations in online education systems," in *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, 2019, pp. 1261–1270.
[18] Y. Cheng, M. Li, H. Chen, Y. Cai, H. Sun, H. Zou, and G. Zhang, "Exercise recommendation method combining neuralcd and neumf models," in *Proceedings of the 2021/7th Annual International Conference on Network and Information Systems for Computers (ICNISC)*. IEEE, 2021, pp. 646–651.
[19] K. Nagatani, Q. Zhang, M. Sato, Y.-Y. Chen, F. Chen, and T. Ohkuma, "Augmenting knowledge tracing by considering forgetting behavior," in *Proceedings of the 2019 World Wide Web Conference*, 2019, pp. 3101–3107.
[20] S. E. Embretson and S. P. Reise, "Item response theory for psychologists," in *Proceedings of the 2020 Health Outcome Methodology Symposium*, 2000.
[21] M. D. Reckase, "Multidimensional item response theory models," in *Multidimensional Item Response Theory*. Springer, 2009, pp. 79–112.
[22] A. Toscher and M. Jahrer, "Collaborative filtering applied to educational data mining," *KDD cup*, 2010.
[23] S. Cheng, Q. Liu, E. Chen, Z. Huang, Z. Huang, Y. Chen, H. Ma, and G. Hu, "Dirt: Deep learning enhanced item response theory for cognitive diagnosis," in *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, 2019, pp. 2397–2400.
[24] W. Gao, Q. Liu, Z. Huang, Y. Yin, H. Bi, M.-C. Wang, J. Ma, S. Wang, and Y. Su, "Rcd: Relation map driven cognitive diagnosis for intelligent education systems," in *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2021, pp. 501–510.
[25] X. Wang, C. Huang, J. Cai, and L. Chen, "Using knowledge concept aggregation towards accurate cognitive diagnosis," in *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*, 2021, pp. 2010–2019.
[26] L. M. Antonio and C. A. C. Coello, "Use of cooperative coevolution for solving large scale multiobjective optimization problems," in *Proceedings of the 2013 IEEE Congress on Evolutionary Computation*. IEEE, 2013, pp. 2758–2765.
[27] Y. Tian, S. Yang, L. Zhang, F. Duan, and X. Zhang, "A surrogate-assisted multiobjective evolutionary algorithm for large-scale task-oriented pattern mining," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 3, no. 2, pp. 106–116, 2019.

[28] Y. Xue, Y. Tang, X. Xu, J. Liang, and F. Neri, "Multi-objective feature selection with missing data in classification," *IEEE Transactions on Emerging Topics in Computational Intelligence*, 2021.
[29] Y. Mei, S. Nguyen, B. Xue, and M. Zhang, "An efficient feature selection algorithm for evolving job shop scheduling rules with genetic programming," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 1, no. 5, pp. 339–353, 2017.
[30] U. Ahmed, J. C.-W. Lin, G. Srivastava, R. Yasin, and Y. Djenouri, "An evolutionary model to mine high expected utility patterns from uncertain databases," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 5, no. 1, pp. 19–28, 2020.
[31] S. Yang, Y. Tian, C. He, X. Zhang, K. C. Tan, and Y. Jin, "A gradient-guided evolutionary approach to training deep neural networks," *IEEE Transactions on Neural Networks and Learning Systems*, 2021.
[32] Y. Tian, X. Su, Y. Su, and X. Zhang, "Emodmi: A multi-objective optimization based method to identify disease modules," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 5, no. 4, pp. 570–582, 2020.
[33] Y. Tian, X. Zheng, X. Zhang, and Y. Jin, "Efficient large-scale multi-objective optimization based on a competitive swarm optimizer," *IEEE Transactions on Cybernetics*, vol. 50, no. 8, pp. 3696–3708, 2020.
[34] X. Ma, F. Liu, Y. Qi, X. Wang, L. Li, L. Jiao, M. Yin, and M. Gong, "A multiobjective evolutionary algorithm based on decision variable analyses for multiobjective optimization problems with large-scale variables," *IEEE Transactions on Evolutionary Computation*, vol. 20, no. 2, pp. 275–298, 2016.
[35] C. He, L. Li, Y. Tian, X. Zhang, R. Cheng, Y. Jin, and X. Yao, "Accelerating large-scale multiobjective optimization via problem reformulation," *IEEE Transactions on Evolutionary Computation*, vol. 23, no. 6, pp. 949–961, 2019.
[36] Y. Tian, X. Zhang, C. Wang, and Y. Jin, "An evolutionary algorithm for large-scale sparse multiobjective optimization problems," *IEEE Transactions on Evolutionary Computation*, vol. 24, no. 2, pp. 380–393, 2020.
[37] Y. Tian, C. Lu, X. Zhang, K. C. Tan, and Y. Jin, "Solving large-scale multiobjective optimization problems with sparse optimal solutions via unsupervised neural networks," *IEEE Transactions on Cybernetics*, vol. 51, no. 6, pp. 3115–3128, 2020.
[38] Y. Tian, C. Lu, X. Zhang, F. Cheng, and Y. Jin, "A pattern mining based evolutionary algorithm for large-scale sparse multi-objective optimization problems," *IEEE Transactions on Cybernetics*, vol. PP, no. 99, 2020.
[39] Y. Xiang, Y. Zhou, Z. Zheng, and M. Li, "Configuring software product lines by combining many-objective optimization and sat solvers," *ACM Transactions on Software Engineering and Methodology (TOSEM)*, vol. 26, no. 4, pp. 1–46, 2018.
[40] J. Xiao, T. Zhang, J. Du, and X. Zhang, "An evolutionary multiobjective route grouping-based heuristic algorithm for large-scale capacitated vehicle routing problems," *IEEE Transactions on Cybernetics*, vol. 51, no. 8, pp. 4173–4186, 2021.
[41] Y. Tian, T. Zhang, J. Xiao, X. Zhang, and Y. Jin, "A coevolutionary framework for constrained multiobjective optimization problems," *IEEE Transactions on Evolutionary Computation*, vol. 25, no. 1, pp. 102–116, 2020.
[42] L. Cui, P. Ou, X. Fu, Z. Wen, and N. Lu, "A novel multi-objective evolutionary algorithm for recommendation systems," *Journal of Parallel and Distributed Computing*, vol. 103, pp. 53–63, 2017.
[43] B. Geng, L. Li, L. Jiao, M. Gong, Q. Cai, and Y. Wu, "Nnia-rs: A multi-objective optimization based recommender system," *Physica A: Statistical Mechanics and its Applications*, vol. 424, pp. 383–397, 2015.
[44] Y. Zuo, M. Gong, J. Zeng, L. Ma, and L. Jiao, "Personalized recommendation based on evolutionary multi-objective optimization [research frontier]," *IEEE Computational Intelligence Magazine*, vol. 10, no. 1, pp. 52–62, 2015.
[45] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, 2002.
[46] Q. Zhang and H. Li, "MOEA/D: A multiobjective evolutionary algorithm based on decomposition," *IEEE Transactions on Evolutionary Computation*, vol. 11, no. 6, pp. 712–731, 2007.
[47] J. Benesty, J. Chen, and Y. Huang, "On the importance of the pearson correlation coefficient in noise reduction," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 16, no. 4, pp. 757–765, 2008.
[48] G. Cybenko, "Approximation by superpositions of a sigmoidal function," *Mathematics of control, signals and systems*, vol. 2, no. 4, pp. 303–314, 1989.
[49] P. Vincent, H. Larochelle, Y. Bengio, and P.-A. Manzagol, "Extracting and composing robust features with denoising autoencoders," in *Proceedings of the 25th International Conference on Machine Learning*, 2008, pp. 1096–1103.
[50] E. Zitzler, M. Laumanns, and L. Thiele, "Spea2: Improving the strength pareto evolutionary algorithm," *TIK-report*, vol. 103, 2001.
[51] C. Pojen, H. Mingen, and T. Tzuyang, "Junyi academy online learning activity dataset: a large-scale public online learning activity dataset from elementary to senior high school students (2020)," *Dataset available from https://www. kaggle. com/junyiacademy/learning-activity-public-dataset-by-junyi-academy*.
[52] C. Croux and C. Dehon, "Influence functions of the spearman and kendall correlation measures," *Statistical Methods & Applications*, vol. 19, no. 4, pp. 497–515, 2010.
[53] P. K. Sen, "Estimates of the regression coefficient based on kendall's tau," *Journal of the American Statistical Association*, vol. 63, no. 324, pp. 1379–1389, 1968.
[54] M. Nilashi, K. Bagherifard, O. Ibrahim, H. Alizadeh, L. A. Nojeem, and N. Roozegar, "Collaborative filtering recommender systems," *Research Journal of Applied Sciences Engineering and Technology*, vol. 5, no. 16, pp. 4168–4182, 2013.
[55] J. Wu, X. Wang, F. Feng, X. He, L. Chen, J. Lian, and X. Xie, "Self-supervised graph learning for recommendation," in *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2021, pp. 726–735.
[56] L. While, P. Hingston, L. Barone, and S. Huband, "A faster algorithm for calculating hypervolume," *IEEE Transactions on Evolutionary Computation*, vol. 10, no. 1, pp. 29–38, 2006.