Leveraging both Successes and Failures in Case-Based Reasoning for Optimal Solutions

Fateh Boulmaiz(✉), Patrick Reignier, and Stephane Ploix

Université Grenoble Alpes, INP Grenoble, LIG, 38000 Grenoble, France
{fateh.boulmaiz,patrick.reignier,stephane.ploix}@univ-genoble-alpes.fr

Abstract. Usually, existing works on adaptation in case-based reasoning assume that the case base holds only successful cases, i.e., cases having solutions believed to be appropriate for the corresponding problems. However, in practice, the case base could hold failed cases, resulting from an earlier adaptation process but discarded by the revision process. Not considering failed cases would be missing an interesting opportunity to learn more knowledge for improving the adaptation process. This paper proposes a novel approach to the adaptation process in the case-based reasoning paradigm, based on an improved barycentric approach by considering the failed cases. The experiment performed on real data demonstrates the benefit of the method considering the failed cases in the adaptation process compared to the classical ones that ignore them, thus, improving the performance of the case-based reasoning system.

Keywords: Case-based reasoning · adaptation · successful case · failed case

1 Introduction

Case-based reasoning (CBR) is undoubtedly the most intuitive artificial intelligence approach for problem-solving, as it emulates human behavior. A CBR system searches through its memory, which is composed of a base of previously solved cases known as source cases, to find cases that exhibit similar problems to the target problem for which a solution is sought. It then adapts their solutions, if necessary, to solve the target problem. The target solution is thoroughly reviewed to ensure its suitability for resolving the target problem, and subsequently, the case base is updated with the new resolution experiment for the target case. Each step of the reasoning process is supported by a knowledge acquisition process required for that particular step.

Adaptation, one of the four key stages in the reasoning process, holds great significance as the quality of the solution heavily relies on its performance. The primary objective of adaptation is to tailor the solutions of similar source cases to meet the specific requirements of the target problem. This step is particularly

N. T. Nguyen et al. (Eds.): ACIIDS 2023, LNAI 13995, pp. 31–44, 2023.
https://doi.org/10.1007/978-981-99-5834-4_3

crucial because the source problems usually do not align perfectly with the target problem. Without successful adaptation, the CBR system cannot generate an appropriate solution for the target problem. The importance of adaptation has been recognized since the early days of CBR systems, leading to numerous studies that explore different approaches for acquiring adaptation knowledge to enhance its performance. According to [13], two distinct approaches to adaptation knowledge acquisition can be distinguished: knowledge-light approaches, which leverage existing knowledge within the system without requiring additional acquisition [11], and knowledge-intensive approaches, which rely on external knowledge sources, such as knowledge obtained from experts or users [5,6].

Existing adaptation approaches primarily concentrate on successful cases (referred to as $\mathbb{C}^+$) that provide relevant solutions to the corresponding problems. The definition of success is subjective and varies depending on the application domain. For example, in the context of a CBR application for an energy management system in a building, a successful case would involve achieving user comfort while minimizing energy expenditure. However, there are also cases that fail to meet the desired criteria. These failed cases (referred to as $\mathbb{C}^-$) have solutions that are deemed unsatisfactory and are typically rejected during the validation phase of the adaptation process. Additionally, the adaptation process often requires acquiring domain-specific knowledge to generate adaptation rules. This knowledge acquisition process is complex and challenging due to its strong dependence on the specific application domain, making it difficult to comprehend and grasp.

In spite of the abundance of research studies and the increased interest in the issue of adaptation, there are few works that specifically address the challenge of proposing a domain-independent adaptation approach. Moreover, there is limited research that considers adaptation from the perspective of solution quality, which encompasses both failed and successful cases. Surprisingly, these cases, which could potentially provide valuable knowledge, are rarely employed by CBR systems. This work introduces a fresh viewpoint on the adaptation process within the CBR paradigm, presenting a fully domain-independent approach that incorporates both successful and failed cases. The study proposes a novel method for acquiring adaptation knowledge, drawing inspiration from research on planning the path of a robot navigating through an unfamiliar and hazardous environment, including obstacles. The uniqueness of this approach lies in applying artificial forces to the proposed solution, aiming to distance itself from failed source solutions while gravitating towards successful ones.

The structure of this paper is organized as follows: Sect. 2 provides an overview of the motivation and background for this work. Section 3 elaborates on the contribution made towards leveraging failed and successful cases for a novel adaptation approach. The evaluation of the proposed approach is presented and discussed in Sect. 4. Finally, Sect. 5 concludes this work, highlights its key findings, and outlines future research directions.

2 Illustrative Example and Preliminary Concepts

A CBR-based energy management system (EMS) in a building serves as a representative case study within the scope of this research. The objective of an EMS is to meet user preferences for thermal comfort, air quality, and other factors, while minimizing energy consumption in the building. Undoubtedly, a building is a complex system influenced by various factors, including climate, building materials, geographical location, energy rates, and the occupants themselves, making it challenging to identify dependencies [3]. Earlier studies [9] have already highlighted the benefits of acquiring adaptation knowledge to enhance the performance of a CBR-based EMS. Additionally, the increasing awareness of environmental concerns has prompted numerous studies to explore the relationship between building energy consumption and occupant comfort, resulting in the establishment of standards [1,2,7] for evaluating user comfort. These standards provide a framework for assessing the quality of the target solution proposed by the adaptation process during the revision phase. Consequently, the solution can be classified as either a successful case ($\mathbb{C}^+$) or a failed case ($\mathbb{C}^-$).

In the CBR-based EMS described in [3], the primary goal is to raise the user's awareness of the impact of their actions on the energy efficiency of the building. To achieve this, the system assists the user by providing recommendations on a set of actions aimed at reducing energy wastage while taking their comfort into account. Each case within the system represents a specific energy management scenario for a building over a single day. The actions stored in the case base correspond to the actions actually performed by the building occupant; however, there is no guarantee that these actions yield satisfactory outcomes for the occupant. To address this, the system incorporates a function to evaluate the effectiveness of the actions stored in the case base, enabling the appropriate labeling ($\mathbb{C}^-$ or $\mathbb{C}^+$) of the corresponding cases.

2.1 Key Concepts and Notations Related to the CBR Paradigm

Each past experience of a CBR system, which forms the foundation for solving new problems, is stored in a structure called a source case ($\mathbb{C}^{sr}$), and the collection of source cases constitutes a case base ($\mathbb{CB}$). Below is a concise introduction to essential concepts within the CBR paradigm, necessary for comprehending our approach.

Case Organization. Consider three sets, $\mathbb{C}$, $\mathbb{A}$, and $\mathbb{E}$, which are mutually disjoint. A case is defined as a triplet $(\mathscr{C}, \mathcal{A}, \mathcal{E})$ where:

- $\mathscr{C}$ belongs to the context domain $\mathbb{C}$ and represents the fixed elements of the problem that cannot be controlled. For instance, in a CBR-based medical diagnostic system, $\mathscr{C}$ can represent physiological indicators of the patient such as heart rate, respiratory rate, etc.

- $\mathcal{A}$ belongs to the action domain $\mathbb{A}$, representing elements that can be controlled to achieve desired outcomes. It represents the suggested solution for the system. In a medical diagnostic system, this could entail the names of recommended medications and their corresponding administration protocols.
- $\mathcal{E}$ belongs to the effect domain $\mathbb{E}$, which characterizes the system's outcome resulting from action $\mathcal{A}$ in context $\mathscr{C}$. In a medical diagnostic system, $\mathcal{E}$ can denote the patient's post-treatment clinical observations or test results.

A target context, denoted as $\mathscr{C}^{tg}$, represents a specific context for which the CBR system aims to determine appropriate actions $\mathcal{A}^{tg}$ in order to produce desired effects $\mathcal{E}^{tg}$ and ultimately generate a target case $\mathbb{C}^{tg}$. The resolution of a problem within the CBR paradigm can be formally described by Eq. (1).

$$\begin{gathered}\text{CBR system: } (\mathbb{CB}, \mathscr{C}^{tg}) \longmapsto \mathcal{A}^{tg} \\ \mathbb{C}^{tg} \stackrel{\text{def}}{=\!=} (\mathscr{C}^{tg}, \mathcal{A}^{tg}, \mathcal{E}^{tg})\end{gathered} \tag{1}$$

The Retrieval and Adaptation Processes. While this paper does not delve into a comprehensive exposition of the reasoning process, it is crucial to acknowledge the inherent relationship between adaptation and knowledge retrieval. Consequently, it is often imperative to present the adaptation process alongside the retrieval process.

- *retrieval stage*: In the retrieval process, the goal is to find source cases that exhibit a context similar to the target context, using a threshold $\Theta_{\mathbb{C}^{tg}}$ to measure the distance between their context variables. Precisely, the process involves locating cases where the context distance from the target context is below $\Theta_{\mathbb{C}^{tg}}$. The retrieval function's profile is outlined by Eq. (2).

$$\text{Retrieval process:} \quad \mathscr{C}^{tg} \longmapsto \{\forall\ \mathbb{C}^{sr} \in \mathbb{CB}, D^{ct}(\mathbb{C}^{tg}, \mathbb{C}^{sr}) \leq \Theta_{\mathbb{C}^{tg}}\} = \mathbb{S}^{\mathbb{C}^{tg}} \tag{2}$$

  where D^{ct} represents the distance between the target context variables $\mathscr{C}^{tg}$ of the target case $\mathbb{C}^{tg}$ and the context variables $\mathscr{C}^{sr}$ of the source case $\mathbb{C}^{sr}$.
  There are no limitations on the choice of distance metric in order to accommodate the context variables. This flexibility allows for the utilization of various distance measures based on the nature of the context variables. For instance, in a CBR-based EMS, the Manhattan metric can be employed to calculate the contextual distance. This is particularly suitable when the context variables involved are real-valued.
- *adaptation stage*: Given that the source contexts often differ from the target context, it becomes necessary to establish a function that can modify the source actions in order to meet the requirements of the target context. The characteristics of this adaptation function can be described by Formula (3).

$$\begin{aligned}\text{Adaptation process: } &\forall\ \mathbb{C}^{sr} \stackrel{\text{def}}{=\!=} (\mathscr{C}^{sr}, \mathcal{A}^{sr}, \mathcal{E}^{sr}) \in \mathbb{S}^{\mathbb{C}^{tg}}, \\ &(\{(\mathscr{C}^{sr}, \mathcal{A}^{sr}, \mathcal{E}^{sr})\}, \mathscr{C}^{tg}) \longmapsto \mathcal{A}^{tg}\end{aligned} \tag{3}$$

where the set $\mathbb{S}^{\mathbb{C}^{tg}}$ refers to the collection of source cases that are considered similar based on the definition provided by Eq. (2).

It is important to note that Eq. (3) does not place any limitations on the number of similar cases that can be considered during the adaptation process. As a result, we are dealing with a form of adaptation that involves combining solutions from multiple source cases to generate a target solution. This type of adaptation is known as compositional adaptation, with single case adaptation being a special case of it. Indeed, the experimental findings from [12] demonstrate that relying solely on a single case often produces less accurate outcomes. This can be attributed to the fact that, in many cases, only a portion of the problem exhibited in the similar source case is relevant to the target problem. Consequently, the process of adaptation becomes complex and, at times, even impossible.

2.2 Collision Avoidance Navigation

The primary objective of studying robot path planning is to address the movement of an autonomous robot within an unfamiliar environment. This involves guiding the robot from its starting point to a designated target position, while placing emphasis not only on finding the most efficient path but also on ensuring the utmost safety. The aim is to calculate a path that optimizes both efficiency and safety by effectively avoiding any obstacles that may arise along the trajectory leading to the target.

Numerous strategies have been proposed to address this challenge, with the Artificial Potential Field (APF) approach, originally introduced in [8], being widely utilized for robot guidance. The APF approach effectively handles the real-world environment in which a robot operates, taking into account both the desired objectives and the obstacles that need to be avoided during movement. The fundamental concept behind this approach is to treat the robot as a point moving within a two-dimensional space (in a basic scenario), influenced by a field created by the targets to be reached and the obstacles to be avoided. Consequently, the robot experiences two types of forces: an attractive force $\mathbb{F}^A$ generated by the targets, and a repulsive force $\mathbb{F}^R$ generated by the obstacles, which collectively determine the robot's movement.

While the repulsive forces exerted on the robot are stronger when it is closer to obstacles and weaker as the distance increases, the attractive forces acting on the robot are directly proportional to the distance between the robot and its target. By combining these forces, denoted as $\overrightarrow{\mathbb{F}} = \overrightarrow{\mathbb{F}^A} + \overrightarrow{\mathbb{F}^R}$, the robot's movement direction and speed can be determined while avoiding collisions with obstacles. Figure 1 illustrates the basic principle of this method, specifically designed for a robot moving in a two-dimensional environment, for the purpose of simplification.

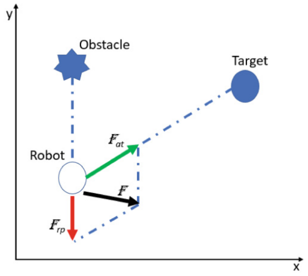

Fig. 1. Robotics's Artificial potential field.

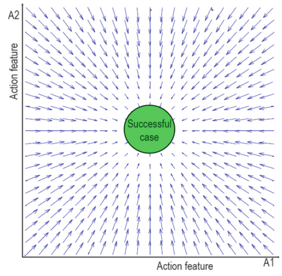

Fig. 2. CBR attractive force.

3 Adaptation Through Failed and Successful Cases

3.1 Problem Statement

The adaptation problem, taking into account both failed and successful cases, can be expressed through the following problem statement. Considering the following observations:

– the case base $\mathbb{CB}$ is partitioned into two subsets: failed cases denoted as $\mathbb{CB}^-$, and successful cases denoted as $\mathbb{CB}^+$. Therefore, $\mathbb{CB} = \mathbb{CB}^- \cup \mathbb{CB}^+$.
– through language misuse, we use the term "target case" to refer to the elements within a specific context for which we are seeking a solution. However, the structure of this case is not fully defined, particularly regarding the elements representing the actions and their subsequent effects, which remain unknown.

The objective of finding a solution for a target case that is currently under construction involves inferring a set of target actions from source cases that share a similar context. These target actions are aimed at satisfying the specific context of the target case. This process leads to the identification and definition of the target effects. Ultimately, the goal is to construct a comprehensive case that encompasses the three parties: context, actions, and effects.

When dealing with similar source cases, which encompass both successful (member of $\mathbb{CB}^+$) and failed cases (member of $\mathbb{CB}^-$), it is important to handle them differently based on their outcome (failure or success) and their level of similarity to the target case. To address this, the proposed method should incorporate mechanisms that guide the approach towards successful solutions for similar source cases while moving away from failed ones. It should also consider that the closer a source case is to the target case, the greater impact its solution will have on the desired target solution.

3.2 Principle

Our approach to addressing failed cases in the adaptation process draws inspiration from navigation algorithms employed in the programming of autonomous robots. Specifically, we adopt the concept of artificial potential field discussed in Sect. 2.2.

Prior to delving into the specifics of our approach in the following section, we have established a set of assumptions to guarantee the effective integration of an artificial potential field-like concept within the framework of this study:

– although the study does not cover the labeling process, it is assumed that prior experiences (referred to as source cases) have already been labeled as either successful or failed cases. Additionally, it is presumed that the CBR system is equipped with a quality function QF, which evaluates the effectiveness of the actions taken within a given context. Higher scores indicate better performance. Consequently, this implicitly establishes a threshold value $\mathcal{TS}^{\mathcal{E}_i}$ for each effect feature $\mathcal{E}_i$, as defined by Eq. (4).

$$\forall\ \mathbb{C} \stackrel{\text{def}}{=} (\mathscr{C}, \mathcal{A}, \mathcal{E}) \in \mathbb{CB},\ \ \mathrm{QF} : \mathcal{E} \longmapsto \mathbb{R}$$
$$\mathrm{LF}(\mathbb{C}) = \begin{cases} \mathbb{C}^{+} & \text{if } \mathrm{QF}(\mathcal{E}) \geq \mathcal{TS}^{\mathcal{E}}_{s}\ ,\ \forall\ \mathcal{E} \in \mathbb{E} \\ \mathbb{C}^{-} & \text{otherwise.} \end{cases} \tag{4}$$

With LF – the labeling function, $\mathcal{E}$ – an effect variable of the case $\mathbb{C}$.

– classical CBR methods typically retrieve a fixed number of neighboring cases from the case base $\mathbb{CB}$, without considering the optimal number of similar cases specific to the target case. This approach, resembling KNN, gives rise to certain issues. Not all target cases necessarily possess the same number of similar neighbors; some may have more while others may have fewer. Additionally, handling situations where there are significantly more source cases equidistant from a target case than the predefined number becomes challenging. In this study, we assume the presence of a retrieval approach that adjusts the number of source cases based on their similarity to the target case $\mathbb{C}^{tg}$. This adjustment is achieved by dynamically defining a similarity threshold $\varphi^{\mathbb{C}^{tg}}$ for the context distance between $\mathbb{C}^{tg}$ and its neighboring source cases. For example, a method proposed in [3] offers a technique to determine this threshold by combining statistical analysis and a genetic algorithm.

The main concept behind the approach proposed in this work involves associating the types of source cases available in the case base, namely successful and failed cases, with the types of objects involved in the domain of robot movement, specifically targets and obstacles. As a result, failed cases are interpreted as obstacles, while successful cases are treated as targets. In this framework, cases $\mathbb{C}^{+} \in \mathbb{S}^{\mathbb{C}^{tg}}$ that exhibit positive outcomes generate an attractive force $\mathbb{F}^{A}$, which draws the target solution towards them. On the other hand, failed cases $\mathbb{C}^{-} \in \mathbb{S}^{\mathbb{C}^{tg}}$ produce a repulsive force $\mathbb{F}^{R}$ pushing the solution away from them.

The source, both successful and failed, cases are utilized to create a CBR potential field that represents the characteristics of the desired solution. Similar to the approach in the robotic potential field method, the CBR potential field consists of two fields. For exemple, in the case of the attractive potential field, a force of attraction is generated from the target solution towards the source solutions of the successful cases. This is achieved by configuring the latter in a way that enables pulling the target solution closer to the solutions of these cases.

To help explain this idea, let's imagine a system that incorporates domain knowledge with just two action variables. In this system, we can visualize the attractive potential field created by any successful case as shown in Fig. 2. In this figure, at every point in the context space representing the target context, the force vectors point towards the successful source case. On the other hand, the repulsive potential field generates a pushback force from the failed case towards the target solution. This force helps move the target solution away from the solutions associated with these failed cases. Figure 3 provides a visual representation of the repulsive force in a configuration similar to the example that demonstrates the CBR attractive force.

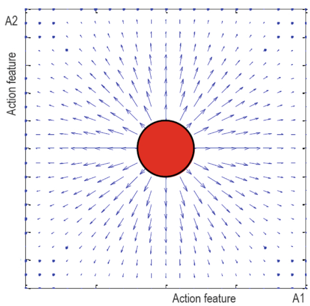

Fig. 3. Repulsive force in CBR context.

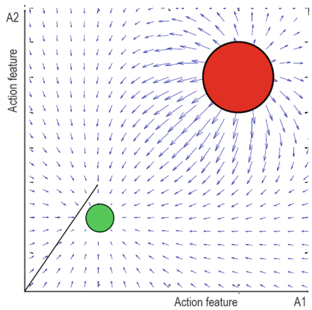

Fig. 4. Total potential force in CBR.

In the end, the positioning of the target solution within the solution space (actions) is established by combining the attractive and repulsive forces exerted by neighboring successful and failed cases, respectively. In the case where there are only two similar cases, one being a successful case and the other a failed case, the overall potential field takes the form depicted in Fig. 4.

3.3 Local Prediction of the Target Solution

While we draw inspiration from the potential artificial field method, applying it in the context of this work, as it is typically done in the robotics field, is not suitable for determining the target solution due to several reasons:

– in the realm of robotics, the overall force that a robot can exert is determined solely by the distance between the robot and the goal or obstacles it encounters. However, in the context of case-based reasoning, the strength of the attractive and repulsive forces is not solely dependent on the distance between the target context and the surrounding similar cases. It also takes into account the performance or quality of the neighboring similar cases.

- within the realm of robotics, the repulsive force in proximity to an obstacle is strongest, gradually diminishing as the distance from the obstacle increases, unlike the attractive force. In the context of case-based reasoning applications, the magnitude of both forces should be directly proportional to the performance of the source solutions, but inversely proportional to the distance between the source similar contexts and the target context.
- typically, robotic applications involve a single goal to be reached. However, in the scenario of a multi-goal environment, the objective is to find a path that sequentially traverses all these goals while optimizing specific criteria. In the context of CBR systems, the goal is to leverage the knowledge from neighboring source cases to infer the desired solution for the target case.
- while the objective of the potential artificial field in robotics is to identify a secure path towards the goal, its purpose within the context of CBR is to acquire fresh knowledge that facilitate the adaptation process in constructing the target solution. In other words, its role is to guide the reasoning process towards the most valuable solutions, which are typically the closest and best-performing cases, while steering away from unfavorable cases that are either farthest away or exhibit poor performance.

Table 1. Summery of results on synthetic dataset.

Approach	Test Set																	
	P_1			P_2			P_3			P_4			P_5			GLOBAL		
	Metrics			Metrics			Metrics			Metrics			Metrics			Metrics		
	IRP (%)	*EII(%)*	*EER(%)*	*IRP*	*EII*	*EER*	*IRP*	*EII*	*EER*	*IRP*	*EII*	*EER*	*IRP*	*EII*	*EER*	*IRP*	*EII*	*EER*
CBR_S	16.73	59.13	59.13	17.85	48.57	48.57	19.53	60.12	60.12	20.48	56.07	56.07	18.79	64.48	64.48	18.68	57.67	57.67
CBR_B	18.27	57.51	57.51	15.36	63.90	63.90	22.85	59.69	59.69	24.23	65.52	65.52	21.10	662.71	62.71	20.36	61.87	61.87
CBR_P	22.62	42.26	57.10	18.54	48.85	63.71	20.14	50.21	60.10	22.48	52.92	70.19	23.47	39.86	60.09	21.45	46;82	62,24
CBR_R	−2.56	32.18	49.75	9.12	29.89	51.19	14.71	43.07	64.24	17.45	39.52	57.74	12.04	41.26	62.84	10.15	37.18	57.15
$APF - CBR$	**34.68**	**100**	**100**	**28.85**	**99.76**	**99.76**	**33.91**	**100**	**100**	**31.27**	**100**	**100**	**38.73**	**99.88**	**99.88**	**33.49**	**99.92**	**99.92**

To effectively incorporate the specificities of the CBR adaptation process, it becomes necessary to modify the artificial potential field approach. In our proposed approach, the target solution (actions), denoted as $\mathcal{A}^{tg}$, is determined by the vector sum of all attractive forces ($\mathbb{F}^{A}_{\mathbb{C}^+}, \forall\ \mathbb{C}^+ \in \mathbb{S}^{\mathbb{C}^{tg}}$) and all repulsive forces ($\mathbb{F}^{R}_{\mathbb{C}^-}, \forall\ \mathbb{C}^- \in \mathbb{S}^{\mathbb{C}^{tg}}$), as defined in Eq. (5).

$$\forall\ \mathbb{C}^+, \mathbb{C}^- \in \mathbb{S}^{\mathbb{C}^{tg}}, \sum_{\mathbb{C}^+} \mathbb{F}^{A}_{\mathbb{C}^+} \overrightarrow{\mathcal{A}_{tg} A_{C_i+}} + \sum_{\mathbb{C}^-} \mathbb{F}^{R}_{\mathbb{C}^-} \overrightarrow{\mathcal{A}_{tg} A_{\mathbb{C}^-}} = 0 \tag{5}$$

As previously stated, the strength of the repulsion and attraction forces is influenced by both the distance between the target context and the context of the similar source case, as well as the performance of the source case. Eq. (5) provides a metric $\mathbb{F}_{\mathbb{C}}$, which determines the magnitude and direction of the corresponding force associated with case $\mathbb{C}$. To estimate the value of this force, we introduce Eq. (6).

$$\forall\ \mathbb{C} \in \mathbb{S}^{\mathbb{C}^{tg}}, \mathbb{F}_{\mathbb{C}} = \begin{cases} \left(1 - \frac{D^{ct}(\mathbb{C}^{tg}, \mathbb{C}^{sr})}{\Theta_{\mathbb{C}^{tg}}}\right) \times (\mathrm{QF}_{\mathbb{C}} - \mathcal{TS}) & \text{if } \mathrm{QF}_{\mathbb{C}} \neq \mathcal{TS} \\ 1 - \frac{D^{ct}(\mathbb{C}^{tg}, \mathbb{C}^{sr})}{\Theta_{\mathbb{C}^{tg}}} & \text{else} \end{cases} \tag{6}$$

where $D^{ct}(\mathbb{C}^{tg}, \mathbb{C}^{sr})$ indicate the context distance between $\mathbb{C}^{tg}$ and its neighboring case $\mathbb{C}^{sr}$, $\Theta_{\mathbb{C}^{tg}}$ represent the context distance threshold, $\mathcal{TS}$ represent the performance threshold, and $\mathrm{QF}_{\mathbb{C}}$ denote the performance of $\mathbb{C}$.

The Eq. (6) demonstrates that regardless of the force's nature, its strength gradually diminishes as the contextual distance increases, until it reaches zero when the contextual distance reaches the similarity threshold $\Theta_{\mathbb{C}^{tg}}$. In addition to determining the force's strength, the term $\mathrm{QF}_{\mathbb{C}} - \mathcal{TS}$ specifies the type of force. If $\mathrm{QF}_{\mathbb{C}} \geq \mathcal{TS}$, then $\mathbb{F}_{\mathbb{C}} \geq 0$, indicating an attractive force. Conversely, if $\mathrm{QF}_{\mathbb{C}} < \mathcal{TS}$, the force is repulsive. Therefore, it is necessary for the proposed actions $\mathcal{A}^{tg}$ to adhere to the following conditions:

$$\forall\ \mathbb{C} \stackrel{\text{def}}{=} (\mathscr{C}, \mathcal{A}, \mathcal{E}) \in \mathbb{S}^{\mathbb{C}^{tg}},\ \mathcal{A}^{tg} = \frac{1}{\sum_{\mathbb{C}} \mathbb{F}_{\mathbb{C}}} \sum_{\mathbb{C}} \mathbb{F}_{\mathbb{C}} \mathcal{A} \tag{7}$$

4 Evaluation

In this section, we provide a practical evaluation of the proposed approach, referred to as APF-CBR hereafter. The evaluation aims to accomplish two main objectives. Firstly, investigate the influence of considering both failed and successful cases on enhancing the effectiveness of the CBR system. Secondly, evaluate the efficacy of the APF-CBR approach in comparison to existing adaptation approaches.

4.1 Experimental Design

As indicated in Sect. 2, the APF-CBR approach is applied within an EMS, aiming to raise the user's consciousness about the consequences of their actions on energy consumption in a building. More specifically, the EMS suggests a set of measures to the occupant that enhance comfort while simultaneously reducing energy consumption.

To assess the effectiveness of the APF-CBR approach, we carried out an experiment utilizing semi-synthetic data derived from actual data presented in [4]. The data base used for this experiment consisted of a total of 15,948 cases, with each case comprising three types of variables: effect variables, action variables, and context variables. The effect variables in our cases represent the temperature and air quality within the building. The action variables are used to model the opening of the door and window, while the context variables capture the weather conditions. We utilized a 24-value vector to represent each variable, corresponding to a single day. We utilized a 5-fold cross-validation approach to assess the variables in our study. To evaluate our model's performance, we employed a 5-fold cross-validation approach. Initially, we randomly divided the original case base into five subsets of equal size, namely P_1, P_2, P_3, P_4, and P_5. During each iteration of the cross-validation process, one subset was selected as the test set, denoted as $\mathbb{CB}_T$, consisting of target cases. The remaining four subsets served as the learning set, denoted as $\mathbb{CB}_L$, which comprised the source cases. This process was repeated five times, with each subset being used once as the test set. The final values of the metrics adopted in the evaluation correspond to the average of the values obtained in the five iterations

To assess the effectiveness of the proposed actions, we employed functions that measured the level of user dissatisfaction with temperature ($\mathrm{QF}_{\mathbb{C}}^{\mathcal{E}_T}$) and air quality ($\mathrm{QF}_{\mathbb{C}}^{\mathcal{E}_{CO_2}}$), as depicted by Formula (8). To simulate the consequences resulting from the implementation of the suggested actions, we constructed a physical model of the building that was utilized in the experiment.

$$\mathrm{QF}_{\mathbb{C}}^{\mathcal{E}_T}(h) = \begin{cases} 0 & \text{if } \mathcal{E}_T(h) \in [21, 23] \\ \frac{\mathcal{E}_T(h)-23}{26-23} & \text{if } \mathcal{E}_T(h) > 23 \\ \frac{21-\mathcal{E}_T(h)}{21-18} & \text{if } \mathcal{E}_T(h) < 21 \end{cases}, \quad \mathrm{QF}_{\mathbb{C}}^{\mathcal{E}_{CO_2}}(h) = \begin{cases} 0 & \text{if } \mathcal{E}_{CO_2}(h) \leq 500 \\ \frac{\mathcal{E}_{CO_2}(h)-500}{1500-1000} & \text{if } \mathcal{E}_{CO_2}(h) > 500 \end{cases} \tag{8}$$

4.2 Baselines and Metrics

The evaluation process includes several baselines:

1. the $\mathrm{CBR_S}$ approach, which is discussed in [4], utilizes both failed and successful cases. However, it lacks an adaptation process, as it simply involves taking a vote among the solutions of similar cases and selecting the solution with the best performance (maximizing the quality function) to be directly applied to the target case. This baseline choice aims to assess the importance of incorporating multiple source cases in establishing an adaptation process.
2. one approach, referred to as $\mathrm{CBR_B}$, employs a standard barycentric method to merge solutions from both successful and failed similar cases, without the use of artificial forces. The primary objective of this approach is to assess the effectiveness of artificial forces in enhancing the reasoning process.
3. to illustrate the advantages of considering both negative and positive cases over only positive cases, we tested a modified variant of our approach called $\mathrm{CBR_P}$. This variant exclusively considers positive cases, relying solely on attractive forces. The objective behind this modification was to highlight the benefits of incorporating both negative and positive cases in comparison to exclusively focusing on positive cases.
4. as an additional baseline, the method presented in [10] (denoted as $\mathrm{CBR_R}$) is employed. $\mathrm{CBR_R}$ utilizes a K-Nearest Neighbors (KNN) algorithm to identify source cases that are similar to the target case. From these similar cases, a generalized case is created. Moreover, the similar cases are utilized to train a linear regression model. This regression model is then employed to predict the solution for the target case based on the generalized case.

It's important to note that in the experiment, the comparison of all the approaches being tested is based on the actions performed by the user without any assistance. This evaluation is conducted using three specific metrics:

– *Improvement Ratio of Performance (IRP)*: The IRP metric evaluates the performance enhancement achieved by each tested approach. It is determined by comparing the average of the global satisfaction $\mathrm{QF}_{\mathbb{C}}^{P}$, of the proposed actions with the corresponding value $\mathrm{QF}_{\mathbb{C}}^{U}$, of the actions performed by the user without assistance for each test case $\mathbb{C}^t$.

$$IRP_{\mathbb{C}^t} = \frac{\mathrm{QF}_{\mathbb{C}}^{P} - \mathrm{QF}_{\mathbb{C}}^{U}}{\mathrm{QF}_{\mathbb{C}}^{U}} \tag{9}$$

– *Effectiveness Improvement Index (EII)* is calculated as the average ratio of the number of test cases that show performance improvement when the actions recommended by this approach are applied, to the total number of test cases.

$$EII = \frac{\beta^+}{\beta} \tag{10}$$

With $\beta = |\mathbb{CB}_T|$ – the set of test cases, $\beta^+ = \{\mathbb{C} \in \mathbb{CB}_T \ / \ IRP_{\mathbb{C}^t} > 0\}$

– *Effective Enhancement Ratio (EER)* refers to the average ratio between the number of test cases that show improved performance when the recommended actions from the approach are applied, and the total number of test cases for which the approach successfully suggests a solution (whether it improves or degrades performance compared to the user's actions).

4.3 Results

Regardless of the adaptation approach used in a CBR system, the retrieval process plays a significant role in determining its performance. Although this paper does not delve into analyzing the retrieval process, we adopt the methodology proposed in [4] to assess similarity and identify similar source cases within the training set. After applying this approach, it is observed that each target case from the test set has at least one similar source case from the training set.

Table 1 provides a summary of the results obtained through 5-fold cross-validation, comparing our approach to the four baselines. Notable observations from this experiment are:

– while the EER metric aligns with the EII value for the $\mathrm{CBR_S}$, $\mathrm{CBR_B}$, and APF-CBR approaches, the $\mathrm{CBR_P}$ and $\mathrm{CBR_R}$ approaches exhibit a lower EII value compared to EER. This disparity can be attributed to the fact that the first three approaches are capable of producing a solution even when provided with a collection of exclusively failed cases. On the other hand, the $\mathrm{CBR_P}$ and $\mathrm{CBR_R}$ approaches do not possess this capability.
– irrespective of the test set employed, our APF-CBR approach consistently outperforms all other baseline methods, demonstrating superior performance in terms of EII and EER.

- the quality of the adaptation process is significantly influenced by the number of similar source cases. When employing a compositional adaptation approach, the IRP tends to be superior compared to using a single similar case, as exemplified by the comparison between APF-CBR (compositional approach) and CBR_S (single similar case).
- the inclusion of attraction and repulsion forces significantly impacts the results of the adaptation process. Utilizing these forces, our APF-CBR approach surpasses the CBR_B baseline, which does not incorporate them, even when considering an equal number of similar cases. APF-CBR demonstrates a superior performance, being 1.64 times more effective than CBR_B in terms of enhancing case performance (global IRP = 33.49% compared to 20.36%). Furthermore, APF-CBR is 1.61 times more efficient in terms of the number of cases for which it successfully finds a solution. It enhances the performance of user-proposed solutions without assistance in 99.92% of cases, as opposed to 61.87% for CBR_B.
- the performance of a Case-Based Reasoning (CBR) system is significantly influenced by the utilization of failed cases. By incorporating both successful and failed cases, the system enhances the outcomes of the reasoning process. When comparing the performance of three different approaches-APF-CBR, CBR_P, and CBR_R, the EER results demonstrate that the APF-CBR approach surpasses the other baselines. The APF-CBR approach exhibits more than three times greater efficiency than CBR_R and more than 1.5 times greater efficiency than CBR_P in enhancing the performance (PER).

5 Conclusion

This paper introduces a novel method for improving the adaptation process in the Case-Based Reasoning paradigm. Instead of relying solely on successful source cases, we consider both failed and successful cases. We draw inspiration from studies on planning safe paths for robots in unknown environments. Our approach involves generating attraction and repulsion forces from successful and failed cases, respectively, to guide reasoning towards the best solutions and away from the failed ones. Experimental results in an EMS context demonstrate a significant ienhancement in system performance by considering both successful and failed cases. We have developed and evaluated an approach that incorporates the entire set of successful and failed similar cases. Further evaluation could explore the impact of the number of neighboring successful and failed cases, focusing on the top-performing cases and the worst-performing cases. Additionally, future research could investigate the potential influence of failed cases on the domain ontology. Understanding this impact could provide insights into how the domain ontology could be refined or modified to prevent the recurrence of negative cases in the future.

References

1. Ashrae, (ed.). ASHRAE Standard Thermal Environmental Conditions for Human Occupancy. American Society of Heating, Refrigerating and Air-Conditioning Engineers., Atlanta, USA (1992)
2. Ashrae, (ed.). Indoor air quality guide: best practices for design, construction, and commissioning. American Society of Heating, Refrigerating and Air-Conditioning Engineers., Atlanta, USA (2009)
3. Boulmaiz, F., Alyafi, A.A., Ploix, S., Reignier, P.: Optimizing occupant actions to enhance his comfort while reducing energy demand in buildings. In: 11th IEEE IDAACS (2021)
4. Boulmaiz, F., Reignier, P., Ploix, S.: An occupant-centered approach to improve both his comfort and the energy efficiency of the building. Knowl.-Based Syst. **249**, 108970 (2022)
5. Díaz-Agudo, B., González-Calero, P.A.: An architecture for knowledge intensive CBR systems. In: Blanzieri, E., Portinale, L. (eds.) EWCBR 2000. LNCS, vol. 1898, pp. 37–48. Springer, Heidelberg (2000). https://doi.org/10.1007/3-540-44527-7_5
6. Govedarova, N., Stoyanov, S. and Popchev, I.: An ontology based CBR architecture for knowledge management in bulchino catalogue. In: CompSysTech (2008)
7. CSA Group. Z412–17 Office ergonomics - an application standard for workplace ergonomics (2017)
8. Khatib, O.: Real-time obstacle avoidance for manipulators and mobile robots. In: Proceedings of IEEE International Conference on Robotics and Automation (1985)
9. Minor, M., Marx, L.: Case-based reasoning for inert systems in building energy management. In: Aha, D.W., Lieber, J. (eds.) ICCBR 2017. LNCS (LNAI), vol. 10339, pp. 200–211. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-61030-6_14
10. Patterson, D., Rooney, N., Galushka, M.: A regression based adaptation strategy for case-based reasoning, In: AAAI/IAAI (2002)
11. Petrovic, S., Khussainova, G., Jagannathan, R.: Knowledge-light adaptation approaches in case-based reasoning for radiotherapy treatment planning. Artif. Intell. Med. **68**, 17–28 (2016)
12. Sizov, G., Öztürk, P., Marsi, E.: Compositional adaptation of explanations in textual case-based reasoning. In: Goel, A., Díaz-Agudo, M.B., Roth-Berghofer, T. (eds.) ICCBR 2016. LNCS (LNAI), vol. 9969, pp. 387–401. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-47096-2_26
13. Wilke, W., Vollrath, I., Althoff, K.D., Bergmann, R.: A framework for learning adaptation knowledge based on knowledge light approaches. In: Fifth German Workshop on Case-BasedReasoning, pp. 235–242 (1997)