Abstract. This contribution presents the use of Fuzzy Case-Based Reasoning methods for modelling student’s Historical Text Comprehension (F-CBR-DHTC). The proposed fuzzy-Case-Based Reasoning synergism handles the uncertainty in the acquisition of expert’s knowledge regarding student’s observable behaviour during Historical Text Comprehension (HTC). It integrates the right balance between expert’s knowledge described in the form of fuzzy sets and previous experience-cases. A fuzzy k-nn retrieval and similarity-measuring algorithm helps the student model to handle case adaptation by facilitating the retrieval process from the case base. The F-CBR-DHTC student model assesses the student’s cognitive profile and profile descriptor, which describe the student’s learning difficulties, and reflects assessment results back to the student. This open student model helps students to better understand their beliefs and instructors to target their teaching strategies to student groups with specific learning difficulties. Preliminary experiment with human experts indicates that the model is effective in handling uncertainty in student diagnosis.

1 Introduction

A main issue in Artificial Intelligence (AI) is to imitate human intelligence in resolving problems in real world domains. In an Intelligent Tutoring System (ITS), student diagnosis process imitates the human expert’s process of inferring the student’s internal characteristics from his observable behaviour [24]. In the domain of history, this computational diagnostic process imitates a human expert’s ability to estimate how the students comprehend the historical text and what are their learning difficulties. An attempt towards this direction is our previous work concerning the Diagnosis of students’ cognitive profiles of Historical Text Comprehension student model (DHTC), based on MOCOHN, a pencil-and-paper diagnosis model [22].

Results of the diagnostic process are beneficial in guiding individualised learning in history. Recently, there is a growing interest in opening the student model to the learner, encouraging him to reflect on his beliefs and on the learning process [4], [11]. Systems are being built to give the learner greater responsibility and control over learning. Instructors find that these open student models are a useful way of helping them to recognise student difficulties, enabling them to respond to specific student populations or individuals in appropriate ways [4].
Such a student model demands knowledge acquisition from a human expert. The main obstacle in this process is the uncertainty derived not only from the knowledge communication among the developer, the human expert and the system, but also from inaccuracy of the information captured and of approximation involved in all process steps [18]. The success of this process depends on the method adopted with regards to AI perspective.

CBR is claimed to be a paradigm that is more akin to the human way of solving complex diagnostic problems in domains like medicine, law or history. Much of the research in CBR deals with a variety of problem solving techniques which offer a platform for an efficient imitation of the expert’s reasoning by organising empirical knowledge of previous experience in a case base. A human expert solves a diagnostic problem using rules derived from his previous experience-cases, whereas a novice requires complete and concrete rules. CBR integrates the right balance between hard to acquire expert knowledge and more easily acquired knowledge in the form of cases. So, in the building of an ITS, CBR helps more easily than other methods to overcome problems of knowledge acquisition from the expert. CBR has been proposed for a variety of diagnostic applications such as PROTOS for clinical audiology, CASEY for heart failure, CASCADE for causes of crashes or PAKAR for possible causes for building defect, CHEF for case-based planning [13]. CBR has been used in educational systems such as in CELIA, for modelling the memory and reasoning capabilities of a novice, in Engines for Education for case based coach [19], for student modelling [20], in Tutoring and Help Systems [5][26] and in SYIM for distance education [23].

To overcome complex problems like uncertainty developers recently build more hybrid case- and knowledge-based systems than pure CBR systems. Fuzzy logic is designed to operate with linguistics expressions and express imprecision and subjectivity in human thinking. The uncertainty involved in linguistic expressions human experts employ, motivates the utilization of fuzzy logic in connection with CBR. Fuzzy logic contributes to CBR in overcoming problems of managing the uncertainty and problems concerning case adaptation by improving performance of case retrieval [10]. In research community the interests of fuzzy logic and CBR for diagnosis recently intersect [8][17][9]. Fuzzy logic is widely used in student modelling where variables are continuous, imprecise, or ambiguous. Fuzzy-based techniques have been used in educational systems for flexible case-based querying [6].

The objective of this work is to present the Fuzzy-CBR method that implements student diagnosis of students’ cognitive profiles for HTC and of their learning difficulties. In section 2, the DHTC model is outlined and problems concerning uncertainty in knowledge acquisition are discussed. In section 3, an overview of the design methodology of the F-CBR-DHTC model with details is presented. In section 4, the student model open to the student and to the instructor is analysed. In section 5, preliminary experiment and formative evaluation regarding diagnosis accuracy by human experts are presented. In the last section 6, conclusion and some short-term perspectives are discussed.

2 DHTC: Student Diagnosis of HTC

2.1 Student’s Cognitive Profiles for HTC

Comprehension of historical text is a special kind of the complex and interactive cognitive process of comprehension [7][22]. The general hypothesis is that the reader
utilises certain fundamental cognitive categories for establishing and organising the meaning of the text. Comprehension is viewed as attribution of meaning to causal connections between events in the text. Comprehension of historical text is associated with causal connections and arguments made by the reader, which are based on the three cognitive categories action, state and event.

The DHTC student diagnosis model of cognitive profiles for HTC is based on MOCOHN, a pencil-and-paper diagnosis model, and on experimental results [22]. The experimental research conducted, used a historical text, question-pairs and alternative answers. Question-pairs asked the student to use the alternative answers in order to express his position against certain historical issues and support it selecting a justification. The student diagnosis model composes the student’s arguments by combining students’ positions and justifications. An argument is considered complete in case both position and corresponding justification are right. Then, taking into account the number of complete arguments, the model results in the formulation of the students’ cognitive models and cognitive profiles for HTC.

Cognitive models, which reflect the students’ levels of historical thought, concern the recognition of the three general cognitive categories: event, state and action. These categories are represented by their instances in the historical text. When a student composes a complete argument it means that he recognizes the corresponding cognitive category.

Table 1. Cognitive models and the corresponding cognitive profiles of DHTC

<table>
<thead>
<tr>
<th>Cognitive models</th>
<th>Number and types of recognized cognitive categories</th>
<th>Cognitive profiles</th>
</tr>
</thead>
<tbody>
<tr>
<td>NHT</td>
<td>-</td>
<td>Very low</td>
</tr>
<tr>
<td>TAHTnx</td>
<td>TAHT1 1 event or 1 state or 1 action</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>TAHT1x more than 1 events or more than 1 states or more than 1 actions</td>
<td>Nearly Low</td>
</tr>
<tr>
<td></td>
<td>TAHT2 (1 event and 1 state) or (1 event and 1 action) or (1 state and 1 action)</td>
<td>Below Intermediate</td>
</tr>
<tr>
<td></td>
<td>TAHT2x (more than 1 events and states) or (more than 1 states and actions) or (more than 1 events and actions)</td>
<td>Above Intermediate</td>
</tr>
<tr>
<td></td>
<td>TAHT3 1 event and 1 state and 1 action</td>
<td>Nearly High</td>
</tr>
<tr>
<td></td>
<td>TAHT3x more than 1 events, states and actions</td>
<td>High</td>
</tr>
<tr>
<td>HT</td>
<td>all events, states and actions</td>
<td>Very High</td>
</tr>
</tbody>
</table>

Table 1 depicts the cognitive models and the cognitive profiles of DHTC. The general categories of cognitive models considered are: Historical Thought (HT), Towards Acquiring Historical Thought (TAHTnx) and Non-Historical Thought (NHT). TAHTnx cognitive models are categorized in more detail according to the number n of recognized by the student categories and to the number x of instances. TAHT1 means that the student recognizes 1 instance of a cognitive category, whereas TAHT1x means that the student recognizes x instances of a cognitive category, where x>1. The same stands for TAHT2, TAHT2x, TAHT3 and TAHT3x. The number n of recognized categories and the number x of recognized instances of every cognitive category formulate the cognitive profiles and give a first level of classification based on quantitative characteristics.
The cognitive profiles determine the degree to which the students face learning difficulties. Students with Very Low profile seem to have serious difficulties in thinking historically. Students characterised by terms like Low, Nearly Low, Below Intermediate, Above Intermediate, Nearly High and High, seem to have some specific difficulties in thinking historically. Students with Very High profile seem not to have learning difficulties. The learning difficulties can guide the instructor to put down the learning goals and can help him to construct the appropriate for a student learning strategies.

2.2 Handling the Problems of Uncertainty in Knowledge Acquisition

DHTC uses quantitative criteria, which describe the number of successfully recognized cognitive categories. For example, the characterization Low of Table 1 means that the student recognizes one cognitive category but it is not clear, which is that category and what are the student’s difficulties concerning the other two categories, which the student does not recognize. The adoption of qualitative characteristics, which describe the cognitive profiles, enhances DHTC with the ability to trace students’ learning difficulties. The qualitative characteristics of the arguments may be different for students with the same cognitive profile, as they may have different position - justification pairs. The characteristics are described by two attributes: completeness and quality, which are the uncertainty factors. Table 2 demonstrates all possible combinations of position-justification pairs and the corresponding argument completeness. For example, if the position is characterized right and the justification wrong then the argument is incomplete. Possible values of the argument completeness are: complete, almost complete, intermediate, nearly incomplete and incomplete.

Table 2: Argument completeness for position - justification combinations

<table>
<thead>
<tr>
<th>Student responses</th>
<th>Argument completeness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Position</td>
<td>Justification</td>
</tr>
<tr>
<td>right</td>
<td>right</td>
</tr>
<tr>
<td>right</td>
<td>mediocre</td>
</tr>
<tr>
<td>right</td>
<td>wrong</td>
</tr>
<tr>
<td>mediocre</td>
<td>right</td>
</tr>
<tr>
<td>mediocre</td>
<td>mediocre</td>
</tr>
<tr>
<td>mediocre</td>
<td>wrong</td>
</tr>
<tr>
<td>wrong</td>
<td>right</td>
</tr>
<tr>
<td>wrong</td>
<td>mediocre</td>
</tr>
<tr>
<td>wrong</td>
<td>wrong</td>
</tr>
</tbody>
</table>

During comprehension of historical text, the recognition of the cognitive category action demands superior quality process than that of state, whereas the recognition of the cognitive category event demands inferior quality process than that of state [7]. Table 3 demonstrates the argument quality values concerning the cognitive categories. Possible values of the argument quality are: superior for the action, medium for the state and inferior for the event.
Table 3: Argument quality values concerning the cognitive categories

<table>
<thead>
<tr>
<th>Cognitive category</th>
<th>Argument quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>action</td>
<td>superior</td>
</tr>
<tr>
<td>state</td>
<td>medium</td>
</tr>
<tr>
<td>event</td>
<td>inferior</td>
</tr>
</tbody>
</table>

The **profile descriptor** describes the learner’s cognitive profile by exploiting information carried by the arguments qualitative characteristics: the **completeness** and the **quality**. For example, a **Nearly Low** cognitive profile of a learner, during comprehension of a historical text with 5 fragments and 5 question-pairs, can be described by the following profile descriptor: “The learner gives one complete and inferior argument, one nearly incomplete and superior argument, one nearly incomplete and superior argument, one incomplete and superior argument and one incomplete argument of medium quality”.

Figure 1 depicts a part of the historical text concerning different concepts in a historical text concerning the outbreak of French Revolution a question-pair and the corresponding alternative answers. It also indicates the characterizations of the answers, which are not visible to the learner. The answers a1 to a3 are alternative answers to question 3a concerning the position, whereas the answers b1 to b5 are alternative answers to question 3b concerning the corresponding justification.

**Figure 1:** A screenshot depicting a fragment of the historical text concerning the cognitive category event, with a question-pair and alternative answers.

For example, if a student selects the answers a3 and b1 this constitutes a **complete argument of inferior quality** and indicates the recognition of one instance of the cognitive category (in this example the category event). If a student selects the answers a2 and b4 this constitutes a **nearly incomplete argument of inferior quality** and indicates no
recognition of the cognitive category (event). Based on student’s answers: *position-justification* pairs the model infers the corresponding *arguments*. Based on student’s arguments the model formulates the student’s *cognitive profile* and the *profile descriptor*.

3 F-CBR-DHTC: Employing Fuzzy- CBR in DHTC

3.1 Overview of the F-CBR-DHTC System

The F-CBR-DHTC system is a hybrid Fuzzy-CBR system, which handles the complexity of diagnosis of student’s cognitive profile and profile descriptor. The system presents to the student a historical text with question-pairs and alternative answers and asks him to select one answer for every question. Student’s observable behavior is defined by his responses, which are position-justification pair values. The F-CBR-DHTC system, Figure 2, solves the diagnostic problem in two stages: (1) the *Fuzzy inference stage*, and (2) the *Fuzzy-CBR inference stage*. The *Fuzzy inference stage* using fuzzy rules, which incorporate the description of expert’s knowledge concerning the student’s answers to questions (position-justification pair values), infers the argument characteristics: completeness and quality. This stage also expresses argument characteristics with fuzzy sets, which are used in the next stage for defining similarity values. In the *Fuzzy-CBR inference stage* student’s behavior represented by the argument characteristics constitutes the corresponding case, where the argument characteristics are the problem description and the cognitive profile and profile descriptor are the solution. In this stage, reasoning steps are based on the hypothesis that similar problems have similar solutions. A fuzzy k-nn algorithm for flexible case retrieval from the case base, handles case adaptation by exploiting the defined as fuzzy sets similarity values between the arguments characteristics and infers the student’s cognitive profile and his profile descriptor by adapting the query case to the most similar of the retrieved cases.

![Figure 2: Structure modules of the F-CBR-DHTC system](image)

Our system’s intelligence stems from the application of Fuzzy-CBR techniques in the student modelling process. Fuzzy-CBR reinforces the student model to dynamically assess the student’s cognitive profile and the profile descriptor, which describe the student’s learning difficulties in details. Our student model is open to the student and to the instructor. It facilitates the student to keep an eye on his personal performance evaluation and the instructor to identify students with similar difficulties in order to provide them with the appropriate teaching strategies.

3.2 The Fuzzy Inference Stage

This stage aims to transform the symbolic input data (student’s responses) into linguistic terms. Using fuzzy rules, which take into account the quality and the completeness of an
argument infers the student’s argument characteristics. It also aims to facilitate the
definition of similarity values using fuzzy sets.

**Fuzzifier Stage.** Fuzzy inference stage represents human expert’s subjective linguistic
description of student’s responses. We define two sets of types of responses: \( A = \{ A_1, A_2, \ldots, A_n \} \) and \( B = \{ B_1, B_2, \ldots, B_n \} \), where \( A_1, A_2, \ldots, A_n \) are linguistic variables and represent observable responses concerning position, whereas \( B_1, B_2, \ldots, B_n \) represent observable responses concerning justification. The term set of \( A \) is \( T(A_n) = \{ D_{nk} \} \) with \( k \) linguistic values and the term set of \( B \) is \( T(B_n) = \{ E_{nl} \} \) with \( l \) linguistic values. The sets \( T_A = \{ T(A_1), T(A_2), \ldots, T(A_n) \} \) and \( T_B = \{ T(B_1), T(B_2), \ldots, T(B_n) \} \) are the fuzzy sets of all term sets that represent the observable behaviour. The symbolic input is fuzzified by means of the linguistic variable values \( D_{nk} \) and \( E_{nl} \). The numbers \( k, l, D_{nk} \) and \( E_{nl} \) are defined by the developer with the help of the human expert. Student’s argument completeness \( C_n \) result as combinations of the independent input data: position and justification and are represented as linguistic variables with \( m \) linguistic values and corresponding term set: \( T(C_n) = \{ C_{nm} \} \).

Let consider an example, of a historical text with \( n=5 \) arguments, \( k=3 \) linguistic values concerning position, \( l=3 \) linguistic values concerning justification and \( m=5 \) linguistic values concerning argument completeness. Then the sets are \( A_1 = \text{position for event1}, B_1 = \text{justification for event1} \) and \( C_1 = \text{argument for event1} \) and the term sets, according to Table 2, are \( T(A_1) = T(\text{position for event1}) = \{ D_{11}, D_{12}, D_{13} \} = \{ \text{right, mediocre, wrong} \} \), \( T(B_1) = T(\text{justification for event1}) = \{ E_{11}, E_{12}, E_{13} \} = \{ \text{right, mediocre, wrong} \} \), \( T(C_1) = T(\text{argument for event1}) = \{ C_{11}, C_{12}, C_{13}, C_{14}, C_{15} \} = \{ \text{incomplete, almost incomplete, intermediate, nearly complete, complete} \} \).

**Rule evaluation.** Each fuzzy rule is of the form [26]:

\[
\text{IF } A_n \text{ is } D_{nk} \text{ AND } B_n \text{ is } E_{nl} \text{ THEN } C_n \text{ is } C_{nm}
\]

Where \( A_n, B_n \) are linguistic input variables, \( D_{nk}, E_{nl} \) are linguistic input values, \( C_n \) are linguistic output variables and \( C_{nm} \) are linguistic output values of argument completeness taken from the Table 2.

For every argument a fuzzy rule, which takes into account the argument quality (table 3), is expressed in the form of Table 4. Table 4 depicts the consequents of the fuzzy IF-THEN rules for the linguistic variable \( C_1 \): argument completeness \( C_1 \). The fuzzy rule: “IF \( A_1 \) is \( D_{11} \) AND \( B_1 \) is \( E_{13} \) THEN \( C_1 \) is \( C_{13} \)” shows that the student’s argument completeness \( C_1 \) takes the value: complete.

**Table 4:** Consequents of the fuzzy IF-THEN rules expressing argument completeness for the inferior quality argument for event1.

<table>
<thead>
<tr>
<th></th>
<th>( E_{11} )</th>
<th>( E_{12} )</th>
<th>( E_{13} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( D_{11} )</td>
<td>right</td>
<td>incomplete</td>
<td>nearly incomplete</td>
</tr>
<tr>
<td>( D_{12} )</td>
<td>mediocre</td>
<td>incomplete</td>
<td>nearly incomplete</td>
</tr>
<tr>
<td>( D_{13} )</td>
<td>wrong</td>
<td>incomplete</td>
<td>nearly incomplete</td>
</tr>
</tbody>
</table>

In the computation of the fuzzy inference, aggregation component uses the \( \text{min} \) operator (logical AND) and numeric values of membership functions \( i_{kn} \) and \( i_{ln} \), whereas
composition component results in numerical values that consist the k x l fuzzy matrix R, taking into account argument completeness [16]. Each fuzzy rule of the form (1) results in one fuzzy set to each output variable. The human expert judges that an argument quality is more important than another and results in numerical values that consist the l x k fuzzy matrix W. For example, let assume all possible values, of “how complete the argument for event1 is” =\{10, 7, 4, 2, 1\} and consider the fuzzy sets A’, B’, C’, D’, E’; A’=\{1, 0.7, 0.3, 0.1, 0\} is the fuzzy set of the output variable value: complete. Figure 2, demonstrates values of the membership functions that express to what degree the values in the universe of discourse belong to the fuzzy set “how complete an argument for event1 is”.

Figure 2: Discrete fuzzy sets for the inferior quality argument for event1 with respect to “how complete the argument for event1 is”

By exploiting R and W, which represent the human expert’s estimations about argument completeness and quality student’s argument characteristics are translated into fuzzy assessments. The output of the fuzzy inference stage is the n-dimensional information vector \( S_n = \{S_1, S_2, \ldots, S_n\} \) in \([0,1]\) of student’s argument characteristics, where \( S_1, S_2, \ldots, S_n \) in \([0,1]\).

3.3 The Diagnostic Process in the Fuzzy-CBR Inference Stage

After the fuzzy inference stage infers the student’s argument characteristics \( S_n = \{S_1, S_2, \ldots, S_n\} \) the diagnostic process begins in the fuzzy-CBR inference stage. The aims of this stage are to define the case structure using the argument characteristics, define the similarity measures based on the fuzzy sets, which have been defined in the previous stage, and proceed to the case retrieval and case adaptation process [3][15].

Modelling the Case Structure. A case is viewed as a set of attributes, which is divided into two non-empty sets: the case name, the problem description attributes subset \( S \) (argument characteristics) and the solution description attributes subset \( T \) (cognitive profile and profile descriptor) [8]. To perform CBR we relate problems to solutions. We assume a finite set \( M \) of stored cases called a case base and a current problem description denoted by \( q \). Expressed in terms of the fuzzy-CBR framework, the case is declared in the following form: \(<\text{casename}, S_1, S_2, \ldots, S_n, T_1, T_2>\). An example of a case concerning a learner and a historical text with \( n=5 \) question-pairs is the following: \(<\text{Petropoulos}, \text{nearly incomplete}, \text{nearly incomplete}, \text{nearly incomplete}, \text{incomplete}, \text{incomplete}, \text{very low}, \text{two nearly incomplete and superior arguments, one incomplete and inferior}\)
argument, one incomplete and superior argument and one incomplete argument of medium quality>

**Definition of Attributes and Similarity Measures.** The challenging problem is to determine the degree to which stored cases are similar. The local similarity measures between argument characteristics of two cases are calculated according to the previously defined fuzzy sets. For example, the values of the symmetric Table 5 express the similarity values between the arguments using the fuzzy sets of Figure 2. The global similarity between two cases is computed using a fuzzy k-nn algorithm.

**Table 5.** Example of the symmetric matrix of local similarity values between arguments for the attribute \( S_1: \text{argument for the event} \).

<table>
<thead>
<tr>
<th></th>
<th>complete</th>
<th>almost complete</th>
<th>intermediate</th>
<th>nearly incomplete</th>
<th>incomplete</th>
</tr>
</thead>
<tbody>
<tr>
<td>complete</td>
<td>1</td>
<td>0.7</td>
<td>0.3</td>
<td>0.1</td>
<td>0</td>
</tr>
<tr>
<td>almost complete</td>
<td>0.7</td>
<td>1</td>
<td>0.7</td>
<td>0.3</td>
<td>0.1</td>
</tr>
<tr>
<td>intermediate</td>
<td>0.3</td>
<td>0.7</td>
<td>1</td>
<td>0.7</td>
<td>0.3</td>
</tr>
<tr>
<td>nearly incomplete</td>
<td>0.1</td>
<td>0.3</td>
<td>0.7</td>
<td>1</td>
<td>0.7</td>
</tr>
<tr>
<td>incomplete</td>
<td>0</td>
<td>0.1</td>
<td>0.3</td>
<td>0.7</td>
<td>1</td>
</tr>
</tbody>
</table>

**Retrieval and Adaptation Process- the Fuzzy k-nn Algorithm.** We have built an expert-emulating similarity-measuring fuzzy k-nn algorithm, which is used to find \( k \) nearest neighbour cases in the case base [12][9]. The fuzzy k-nn algorithm helps the student model to handle case adaptation by facilitating the retrieval process from the case base. The case retrieval fuzzy k-nn pseudocode with the local procedure of the similarity-measuring function is as follows.

---The case retrieval pseudocode

```plaintext
read(q)
{
    \( \hat{a} = 0.0 \)
    length_of_list = 0

    ---Similarity-measuring function (local procedure)
    for i = 1 to r
        { 
            outcome=1.0
            for (m = 1 to n)
            { 
                outcome = \( i_m (Y_i (m), q(m)) \)
                if (outcome < \( \hat{a} \))
                    return (0.0)
            }
            return(outcome)
        }
    sim(Y, q)=outcome

    ---end of similarity-measuring function
    if (sim(Y, q) > \( \hat{a} \))
        { 
            \( \hat{a} = \text{sim}(Y, q) \)
            insert_to_list(Y)
            length_of_list = length_of_list + 1
        }
```

---
if length_of_list > k
{
    remove_from_list_the_least_similar_member()
    length_of_list = length_of_list - 1
}
}
return(cases weighted according to their degree of similarity)
}

The similarity-measuring function is the local procedure, which calculates the global similarity between two cases. Let \( Y = \{ Y_1, Y_2, \ldots, Y_r \} \) be the set of \( r \) cases of the case base representing \( r \) potential similar cases, called *similar*, and \( q \) the query case. The function \( \text{sim}(Y, q) \) is given two cases and it returns a real number in \([0,1]\) proportional to the degree of similarity of the two cases. As the case base is traversed, and each prospective *similar* \( Y_i \) is subjected to a series of similarity tests progressively. The degree of membership \( \text{im}(Y_i(m), q(m)) \) is the argument characteristic-specific fuzzy comparison operation, for the \( m \)-th of the \( n \) argument characteristic. The argument-by-argument fuzzy comparison operation between \( q \) and a case \( Y_i \) results in the aggregation of the global similarity value.

Cases are either summarily ruled out of contention or finally ruled into the \( k \)-nn set. The similarity measuring operation is halted if the similarity falls below the \( \alpha \)-level. \( \alpha \)-level of a fuzzy set describes the current subset of cases that have membership above a prescribed threshold \( \alpha \) [10]. As better and better *similar* are collected during the traversal of the case base a priority queue in a list is continuously updated including the \( k \) most similar cases and the \( \alpha \)-level rises accordingly. As the \( \alpha \)-level rises, the case-to-case comparison can be aborted after fewer and fewer argument-to-argument comparisons.

The solution description attribute subset \( T = \{ T_1, T_2, \ldots, T_n \} \) of problem description vector \( S \) of \( q \) contains the attribute-values and \( T \) is the diagnosis - the requested value of cognitive profile and the profile descriptor [1][24]. The retrieval process finds a case with an information vector \( S' \) most similar to \( S \) which forms the solution \( T \).

### 3.4 Initialization of the Case Base

To initialize the case base we used a sample of 60 most representative cases of the problem domain. We assumed the construction of a homogeneous case base where all cases share the same record structure [25]. We conducted a research with 40 high school students and appropriate historical text and questions in order to have a sample of the distribution of cases to cognitive profiles. The cognitive profiles were judged by hand and the results appear in Figure 3. Taking into account the experimental results we identified that the frequency of occurrence of cases with *Very Low*, *Low* or *Nearly Low* cognitive profiles is greater than other profiles. Consequently, as most of the students are expected to have *Very Low*, *Low* or *Nearly Low* cognitive profile, the majority, almost 70%, of initial cases in the case base must belong to the corresponding subgroups. The knowledge contained in similarity measures was enriched with appropriate knowledge of the 40 episodic cases and was complemented with 20 prototypical cases that the expert judged necessary for the problem domain [1]. Finally, the knowledge base contains the case base of 60 cases allowing it to be managed that is all the essential active knowledge of the domain.
Figure 3: Frequency of cognitive profiles in a sample of 40 cases. The horizontal axis shows the cognitive profiles, {Very Low, Low, Nearly Low, Below Intermediate, Above Intermediate, Nearly High, High, Very High}, which correspond to {VL, L, NL, BI, AI, NH, H, VH}.

4. The Open F-CBR-DHTC Student Model

The environment of F-CBR-DHTC offers two modes of interaction with the user: The open to the student student model and the open to the instructor student model.

4.1 The Open to the Student Student Model

The student model contains representations of the student’s performance on a set of questions concerning different concepts in a historical text. The environment provides students with an easy-to-use interface through which they are given the historical text to read and questions with alternative answers. Students are encouraged to respond by selecting the right answers according to their opinion. Based on student’s answers: position-justification pairs the system infers the corresponding arguments by using the fuzzy rules in the Fuzzy-inference module. We used a historical text concerning the outbreak of French Revolution with 5 question-pairs each corresponding to an argument. The five arguments for each student constitute the case. Based on student’s arguments the model formulates the student’ cognitive profile and the profile descriptor by retrieving and adapting the most similar case from the case base using the fuzzy-k-nn algorithm in the Fuzzy-CBR inference module. At the end of this process the student can have access to his personal evaluation card, which reports his comprehension status concerning the historical text through the recognition or not of the cognitive categories, which reflect the levels of difficulty he faces in solving problems. Figure 5 demonstrates an example of a personal evaluation card.

<table>
<thead>
<tr>
<th>Text: French Revolution</th>
<th>Personal evaluation card of John Petropoulos</th>
</tr>
</thead>
<tbody>
<tr>
<td>Argument</td>
<td>Argument completeness</td>
</tr>
<tr>
<td>1</td>
<td>complete</td>
</tr>
<tr>
<td>2</td>
<td>incomplete</td>
</tr>
<tr>
<td>3</td>
<td>almost complete</td>
</tr>
<tr>
<td>4</td>
<td>intermediate</td>
</tr>
<tr>
<td>5</td>
<td>nearly incomplete</td>
</tr>
</tbody>
</table>

Figure 5: Contents of a personal evaluation card
The card reflects back to the student evaluation results, which are useful in helping him better understand his beliefs and change his reasoning. The corresponding card is formulated by retrieving from the case base and adapting the query case with the case, which has the most similar argument characteristics and consequently approximates the student’s learning difficulties. The two cases may have different inputs (answers to questions) but similar cognitive profile and profile descriptor.

4.2 The Open to the Instructor Student Model

The model provides the human instructor an administration mode interface, which allows him easily and quickly suspect, which means see at a glance areas in which the student is strongest or weakest, a student’s personal evaluation card, his cognitive profile and the attached profile descriptor. The profile descriptor provides the instructor with information concerning student’s difficulties in the recognition of the cognitive categories, the quality and the degree of recognition of the cognitive categories. So, the system is open to the instructor since it facilitates him to identify groups of students with similar or particular learning difficulties in order to adapt and schedule the appropriate instructional strategies for particular groups. Administration mode interface also permits easily modification of the case-based application interactively. The instructor can flexibly import a new historical text with different number of paragraphs and thus with different number of instances of cognitive categories, define a new structure of the case base, attributes, relationships between attributes and similarity values and use it for student diagnosis.

5 Preliminary Experiment and Evaluation by Experts

We implemented F-CBR-DHTC model using the application development tool CBR-Works, which was chosen for its flexibility in solving problems [20]. During formative evaluation, aiming to evaluate the diagnosis accuracy, we experimented F-CBR-DHTC using a historical text concerning the outbreak of French Revolution and 10 questions. The system learned the domain of diagnosis of student’s HTC from 60 cases: 40 from real students (episodic cases) and 20 prototypical cases. During the preliminary trials, presenting it with 20 new cases we tested the validity of the performance of the diagnostic system [3]. We divided the test cases into 4 groups of 5 cases each, to assess the improvement in F-CBR-DHTC model performance as it gained experience during the learning process. After each case was entered, the system attempted to assign the correct cognitive profile and profile descriptor to each case. If the model assigned to a case a profile descriptor, which is not correct we added that case into the case base. F-CBR-DHTC, as an incremental knowledge acquisition system, demonstrated satisfactory performance in the four test series.

After the trials we tested the system for its diagnosis accuracy by 4 human experts in AI (2 historians and 2 cognitive scientists). In a CBR diagnostic system the problem of noise, which means inherent ambiguities, can manifest itself as two or more cases with identical diagnosis but different inputs [2]. In our diagnostic system of students’ cognitive profiles and profile descriptors the diagnosis accuracy has to do with handling the noise in occurrences like O: two or more students with identical cognitive profile and profile descriptor have different answers to questions. To account for noise within the CBR system’s memory, we count the number of occurrences like O and change a certain
amount of attributes in the involved cases. To test our systems’ ability to handle noisy data during the adaptation process and consequently its utility and acceptability by human experts we conducted a research. The goal of the evaluation research was to examine the extent to which the system and the experts’ decisions agree. The following evaluation scenario was applied. We presented the system with 25 new (potentially real) cases with slightly different to real cases attributes. After the diagnosis, occurrences like O were found in 20 out of the 25 cases. The system was tested for its ability to handle noise by the human experts. The 4 participants engaged independently in the evaluation and were asked to diagnose the cognitive profiles and profile descriptors and judge explaining their view.

![Figure 6](image_url)

**Figure 6**: Number of cases per cognitive profile with agreement in the assessment of the profile descriptor using different methods of assessment: 4 human experts and the F-CBR-HTC diagnostic system for 20 cases. The horizontal axis shows the cognitive profiles {Very Low, Low, Nearly Low, Below Intermediate, Above Intermediate, Nearly High, High, Very High}, which correspond to {VL, L, NL, BI, AI, NH, H, VH}.

There was agreement in the estimation of the cognitive profiles and small disagreements in assessing the profile descriptors. From the results demonstrated in Figure 6 we can observe that the estimation made by the F-CBR-DHTC diagnostic system and the human expert coincide (on average) in 18 out of the 20 cases. Even though the sample is rather small to reach a safe conclusion and given the small amount of cases in the case base, the results indicate that F-CBR-HTC can indeed perform diagnosis in a way that gives results similar to the way human experts evaluated students.

### 6 Conclusions and Future Plans

In this work, we have utilized Fuzzy Case-Based Reasoning methodology for the construction of an open student model of students’ cognitive profiles for HTC. Fuzzy-CBR solves the problem of automating the process of student diagnosis in DHTC for inferring the student’s cognitive profiles. The combination of the advantages of fuzzy logic for managing the uncertainty involved in linguistic expressions an expert uses and
for case retrieval with the advantages of CBR to imitate human expert’s reasoning resulted in the successful construction of a computational model for diagnosis of HTC, which improves the existing diagnostic model DHTC. The system is open to the students and reflects them back their personal performance evaluation. The system is open to the instructors and facilitates them to observe students with similar difficulties in order to recommend what should be studied further. So, this model is expected to interest mainly history instructors, who wish to assess their students’ cognitive profiles and learning difficulties in order to design and experiment a different personalised teaching strategy. The model enables them to reuse experts’ experience from similar situation that occurred in the past and have been documented in the form of cases.

The results of the evaluation scenario proved the diagnosis accuracy of the model and are encouraging for the system’s educational impact on students, even preformed on a limited number of students in a small amount of cases. Future work focuses on the evaluation of the model in terms of scaling validity. We plan to provide the student model with more interactivity, make the student have control over the learning process by analysing his interaction with the system and make the student model open to peers.

References