A Linguistic Modeling Approach to Characterize Items in Computarized Adaptive Test for Intelligent Tutor Systems Based on Competency

Miguel Badaracco and Luis Martínez

Abstract An Intelligent Tutor System based on Competency education (ITS-C) aims to personalize teaching processes according to student's competency profile and learning activities by means of artificial intelligence (AI) techniques. One of the most challenging process in ITS-C is the diagnosis process, so far it has been carried out by computerized adaptive tests (CAT) based on item response theory (IRT), in spite of the good performance, its construction requires a hard statistical calibration of a huge bank of items. Such processes are usually intractable in small institutions. To overcome previous difficulties, enhance the accuracy of diagnosis, and the adaptation to student's competence level this contribution proposes the use of teachers' knowledge to replace statistical calibration by modeling such expert's knowledge linguistically using the fuzzy linguistic approach.

Keywords Intelligent tutor system based on competency education · Fuzzy linguistic approach · Computing with words

1 Introduction

An ITS provides direct customized instruction or feedback to students in their learning processes by means of artificial intelligence (AI) techniques, being mainly applied to knowledge representation, managing an instruction strategy as an expert both in the teaching and pedagogical issues in order to diagnose properly the

M. Badaracco (✉)
Faculty of Economics and Business Administration,
National University of Formosa (UNaF), Formosa, Argentina
e-mail: itfor@arnet.com.ar

L. Martínez
Department of Computer Science, University of Jaén (UJA), Jaén, Spain
e-mail: luis.martinez@ujaen.es

Z. Wen and T. Li (eds.), *Foundations of Intelligent Systems*,
Advances in Intelligent Systems and Computing 277,
DOI: 10.1007/978-3-642-54924-3_80,

student learning status at any time. To fulfill its objective, an ITS is organized by an architecture composed by a domain model, student model, diagnosis of the student [1], instructional model [2], and the interface [1, 3]. The importance of the pedagogical model utilized in the ITS, to achieve its goals, has favored the extension of the ITS into ITS based on competency-based education (CBE) (ITS-C) whose architecture was introduced in [4].

Although the use of CBE improves different processes in the student learning process, still the diagnosis process of ITS-C to update the student model is a very challenging process. Many ITS and the proposed for ITS-C used CAT based on item response theory (IRT), whose construction requires a hard and complex statistical calibration of a huge bank of items, that is intractable in small institutions [4, 5].

To overcome previous drawbacks this contribution proposes the use of teacher's knowledge to deal with CAT removing the statistical calibration of the bank of items. Such knowledge usually involves uncertainty related to qualitative aspects that characterize the items of the bank. Therefore, our proposal considers the use of the fuzzy linguistic approach [6–8] to model teacher's knowledge. Hence, a diagnosis process, that extends the proposed CAT introduced in [9] for ITS-C, is then presented and so-called fuzzy linguistic-CAT (FL-CAT).

The contribution is organized as follows: Sect. 2 reviews some necessary concepts about ITS-C. Section 3 presents a linguistic characterization for items in the bank of items and then a new diagnosis process dealing with such information called FL-CAT. Finally, some concluding remarks are pointed out.

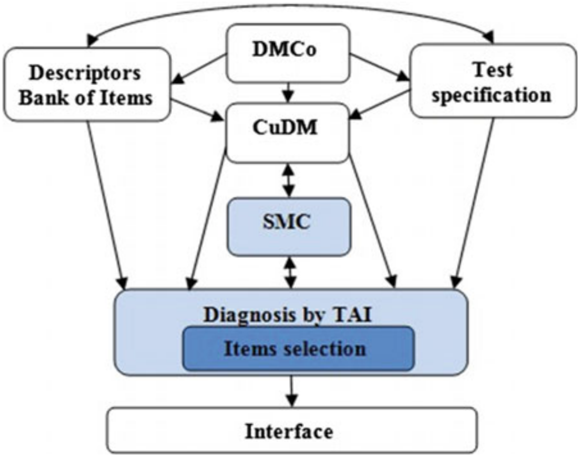
2 Intelligent Tutoring Systems Based on Competency-Based Education. Architecture and Diagnosis Process

An ITS-C [4] extends an ITS by linking it and the pedagogical model based on CBE using the architecture shown in Fig. 1. Here it is reviewed that the elements of interests of an ITS-C for our proposal as the domain model, the student model, and a detailed revision of the diagnosis of an ITS-C [9] together the updating process of the student model of competency (SMC).

2.1 Domain Model of ITS-C

The representation of the *domain model* in an ITS-C is based on the descriptors utilized in CBE [4] that reflect good professional practices to guide the development of the competency associated with an occupational role or profile [10–12]. Such a set of descriptors are:

Fig. 1 ITS-C architecture



- *Competency unit* (cu): It is a main function that describes and groups the different activities concerning the role or profile chosen.
- *Competency element* (ce): It is the disaggregation of a main function (cu) that aims to specify some critical activities.
- *Evidence of performance* (evd): It checks if a process is performed.
- *Evidence of product* (evp): It is a descriptor of tangible evidence in the results.
- *Evidence of knowledge* (evk): It is a descriptor of scientific-technologic knowledge.

Therefore the *domain model* contains the expert's competences profile about a knowledge domain, hence for an ITS-C it will consist of four components briefly detailed below, for further detailed description see [4]:

1. *A domain model of competency* (*DMCo*): It is represented by a semantic network whose nodes are competence units (cu), competence elements (ce), *descriptors* (evd, evp, evk), and their relations.
2. *Curriculum domain model* (*CuDM*): It is based on the CBE that takes a modular structure, in which each module (M_i) contains (ce) belonging to the *DMCo*.
3. *A set of descriptors*: The descriptors associated with the *ce* of the didactic modules are evd, evp, and evk that belong to a bank of items.
4. *Test specifications*: They are provided by the teachers and associated with the diagnosis process considering the scope of application and the student's necessities of learning.

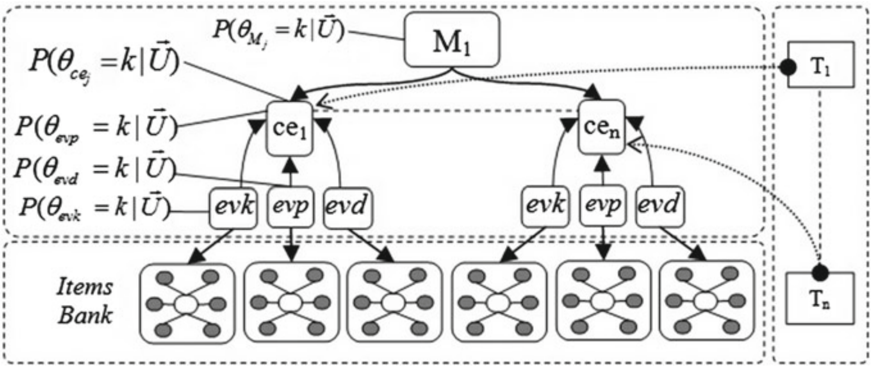


Fig. 2 Module of student model of competence

2.2 Student Model of ITS-C

In an ITS-C, the *SMC* stores student's information, whose data are updated through a diagnosis process. For the representation of the student's knowledge and learning process, the SMC uses an overlay model in the semantic network of the CuDM [9].

In such a semantic network the nodes evp, evd, and evk store a probability distribution $P(\theta_{\text{evp}} = k \mid \vec{u}_i)$, $P(\theta_{\text{evd}} = k \mid \vec{u}_i)$, and $P(\theta_{\text{evk}} = k \mid \vec{u}_i)$ regarding the student's level of competency k in the corresponding node, k can take values from 1 to the maximum number of levels of competency on which the student is evaluated. Being θ the student's level of technical-scientific knowledge about a descriptor for a response pattern $\vec{u}_i$ obtained from the responses provided by the student in the test T (See Fig. 2) during the diagnosis process.

2.3 Diagnosis for ITS-C Based on CAT

Due to the fact that our interest relies on the diagnosis process, a further detailed review of such process is done.

The diagnosis process estimates and updates the level of competency achieved by the student in the nodes of the SMC. In ITS-C, computerized adaptive test (CAT) based on the IRT [5] was adapted and extended [9].

In CAT systems the relationship between student outcomes in the test and her response to a certain item can be described by a monotone increasing function called the item characteristic curve (ICC). The ICC of an ITS-C coincides with the correct response option of the characteristic curve of option (CCO). Its main components are:

- *A response model associated to the items*: It describes the student's expected performance according to his/her estimated knowledge.
- *Bank of Items*: Each item I_i is associated to its descriptors (evd, evp, or evk) and each option of I_i corresponds to a CCO obtained by a calibration process based on the Ramsay algorithm [13]. Each CCO is represented by a probability distribution, $P(\vec{u}_i \mid \theta_0)$, where each component represents the probability that the student selects the response pattern $\vec{u}_i$, given her level of competence θ.

To develop a test, teachers must provide *test specifications* considering the scope of application and the student's necessities of learning, namely:

- *Initial level of Knowledge*: The initial knowledge estimation is crucial because it determines the length of the CAT for each student. It may be estimated by using different models based on previous information.
- *Criterion for selecting evidence node* (evp, evd, *or* evk): The algorithm selects the descriptor that has the level of knowledge associated with lower probability [5, 9]:

$$\min(\theta_{\text{ev}}) = \min(\text{MAP}(P(\theta_{\text{ev}}|\vec{u}_n))) \tag{1}$$

- *Criterion for selecting items*: A common method is the maximum information [5, 14] that selects the item, which maximizes the information in the provisional distribution of student's knowledge. The information function for the item, I_j, is calculated as follows:

$$PI_j(\theta_i) = \frac{\left(P'_j(\theta_i)\right)^2}{P_j(\theta_i)\left(1 - P_j(\theta_i)\right)} \tag{2}$$

  Being θ_i the knowledge level of the student i, $P_j(\theta_i)$ the value of the CCO for the student's level, and $P'_j(\theta_i)$ the derived function from the CCO at that point. Other selection criteria were proposed in [5, 14].
- *Stop criterion*: The test should stop when the student achieves a level of knowledge fixed a priori, though there are other criteria.

2.4 Updating the SMC

During the management of a test, the student's knowledge is estimated every time that he/she answers a question, by updating the student's knowledge distribution [14], as follows:

$$P(\theta_{\text{ev}}|\vec{u}_1,\ldots,\vec{u}_i) = \begin{cases} |P(\theta_{\text{ev}}|\vec{u}_1,\ldots,\vec{u}_{i-1})P_o(\vec{u}_i|\theta)| \\ \text{if } Q_i \text{ asseses evd}_j,\ \text{evk}_j \text{ or evp}_j. \\ P(\theta_{ev}|\vec{u}_1,\ldots,\vec{u}_{i-1}) \text{ in other case.} \end{cases} \tag{3}$$

Being $P(\theta_{\text{ev}}|\vec{u}_1,\ldots,\vec{u}_{i-1})$ the a priori student's knowledge estimation on evd, evp, or evk, and $P_o(\vec{u}_i|\theta)$ the CCO for the response pattern $\vec{u}_i$.

After the updating process, the system estimates the level corresponding to the distribution by using one out of two choices introduced in the CAT [5, 14]:

- Expectation a posteriori (EAP):

$$\theta_{\text{ev}} = \text{EAP}(P(\theta_{\text{ev}}|\vec{u}_n)) = \sum_{k=1}^{n} kP(\theta_{\text{ev}} = k|\vec{u}_n) \tag{4}$$

  being k the knowledge level.
- Maximum a posteriori (MAP):

$$\theta_{\text{ev}} = \text{MAP}(P(\theta_{\text{ev}}|\vec{u}_n)) = \max P(\theta_{\text{ev}} = k|\vec{u}_n) \tag{5}$$

A further detailed revision of operation and specifications of CAT can be found in [9]. It is clear that the diagnosis process depends on the statistical calibration of the items that obtains the CCO, which is costly and intractable in small institutions.

In order to overcome this drawback, in the next section, it is proposed an item characterization model to replace statistical calibration based on teacher's knowledge modeled linguistically by the fuzzy linguistic approach [6–8].

3 A Linguistic Characterization of Items for Diagnosis in ITS-C

Following, it introduced a new approach to characterize items in the bank of items from teachers' knowledge linguistically modeled that replaces the statistical calibration to obtain the items CCO. Moreover a new diagnosis process is then introduced to deal with such characterization in ITS-C (See Fig. 3). Therefore, first it briefly revised some concepts about fuzzy linguistic approach. Afterwards it introduced the new characterization of the items in the item bank by expert's knowledge.

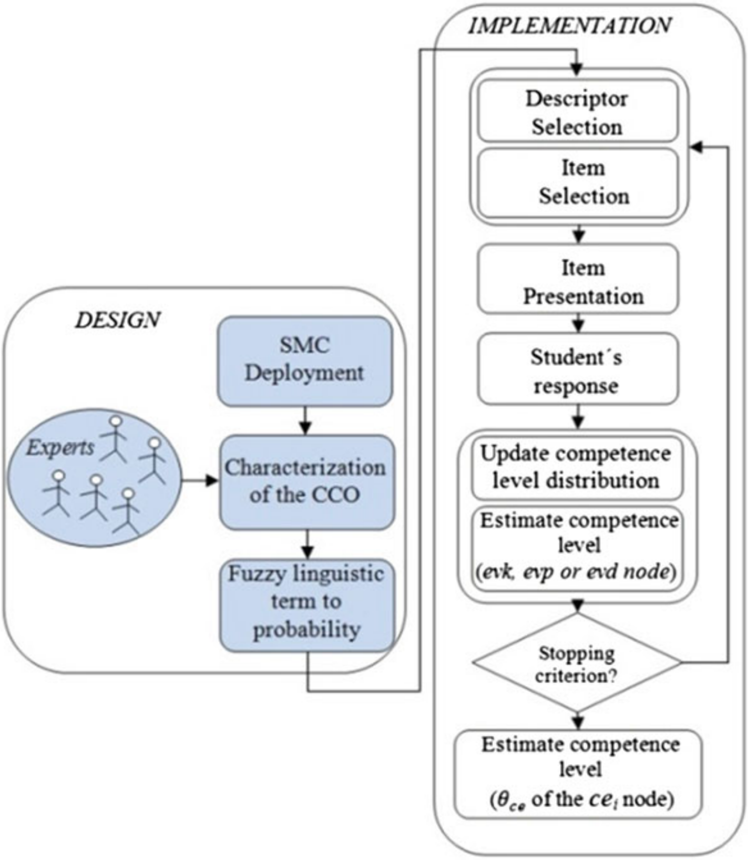


Fig. 3 New student model's update

3.1 The Fuzzy Linguistic Approach

Usually, we work in a quantitative setting, where the information is expressed by numerical values. However, many aspects of different activities in the real world cannot be assessed in a quantitative form, but rather in a qualitative one, i.e., with vague or imprecise knowledge. In that case, a better approach may be to use linguistic assessments instead of numerical values. The fuzzy linguistic approach represents qualitative aspects as linguistic values by means of linguistic variables [6–8].

To do so, a very important concept is the *granularity of uncertainty*, i.e., the level of discrimination among different degrees of uncertainty, typical values of cardinality used in the linguistic models are odd ones, such as 7 or 9, where the mid-term represents an assessment of "approximately 0.5", and the rest of the terms being placed symmetrically around it [15]. In the literature, several possibilities can be found [8, 16, 17]. One possibility of generating the linguistic term set consists of directly supplying the term set by considering all terms distributed on a scale on which a total order is defined. For example, a set of seven terms S, could be:

$$S = \{s_0 = \text{None}; s_1 = \text{Very Low}; s_2 = \text{Low}; s_3 = \text{Medium}; \\ s_4 = \text{High}; s_5 = \text{Very High}; s_6 = \text{Perfect}\}$$

Usually, in these cases, it is required that in the linguistic term set there exist:

1. A negation operator: $\text{Neg}(s_i) = s_j$ such that $j = g - i$ ($g + 1$ is the cardinality).
2. An order: $s_i \leq s_j <=> i \leq j$.

Therefore, there exists a minimization and a maximization operator.

The semantics of the terms are given by fuzzy numbers defined in the [0, 1] interval, which are described by membership functions. Since the linguistic assessments given by the users are just approximate ones, some authors consider that linear trapezoidal membership functions are good enough to capture the vagueness of those linguistic assessments, since it may be impossible and unnecessary to obtain more accurate values [16].

3.2 *Characterizing Item CCO from Expert's Knowledge Linguistically Modeled*

Here it is introduced to a new approach to characterize the item CCO by using the teachers' knowledge modeled linguistically. In this approach, first it defined its representation framework. Second it showd how to manage linguistic terms to build a probability distribution that will be used in the diagnosis operation by the FL-CAT [9].

3.2.1 Framework for Fuzzy Linguistic Representation of Items' CCO

The first step of the fuzzy linguistic representation approach defines a framework to establish the representation model, the domains, and scales in which experts provide their assessments:

- $E = \{e_1, e_2, \ldots, e_m\}$, panel of experts.
- $O_{ij}^{\text{ek}} = \left\{O_{ij}^{e1}, O_{ij}^{e2}, \ldots, O_{ij}^{\text{en}}\right\}$ option characteristic vector (OCV), where O_{ij}^{ek} is the teachers valuation that expresses the possibility that a student with level of competence k to select the option O_j to the item I_i.

The experts express their opinions by means of linguistic labels in a linguistic term set, S: $S = \{s_0$: Very Low; s_1: Low; s_2: Medium; s_3: High; s_4: Very High$\}$

3.2.2 Converting Fuzzy Linguistic Terms into a Probability Distribution

The fuzzy linguistic terms are transformed into a probability distribution according to [18]:

Definition 1 [18]. (*Mass assignment*) Let f be a subset of a finite universe Ω such that the range of membership function of f is $\{y_1, y_2, \ldots, y_n\}$ with $y_{\mathrm{i}} > y_{\mathrm{i+1}} > 0$. Then the mass assignment of f, denoted by m_f is a probability distribution on 2^{Ω} satisfying:

$$m_f(\emptyset) = 1 - y_i, m_f(F_i) = y_i - y_{i+1}, \quad \text{for } i = 1, .., n-1 \quad \text{and } m_f(F_n) = y_n,$$

where $F_i = \{x \in \Omega | \mu_f(x) \geq y_{\mathrm{i}}\}$ for $i = 1, \ldots, n$. $\{F_i\}_{i=1}^n$ are referred to as the focal elements (sets) of m_f.

The notion of mass assignment suggests a means of conditioning a variable relative to a fuzzy constraint. The following definition was proposed by Baldwin [19] and Lawry [18].

Definition 2 (*least prejudiced distribution*) For f a fuzzy subset of a finite universe Ω such that f is normalized then the least prejudiced distribution of f is a probability distribution on Ω given by

$$\forall x \in \Omega \quad \Pr(x|f) = \sum_{F_i : x \in F_i} \frac{m_f(F_i)}{|F_i|}, \tag{6}$$

where m_f is the mass assignment of f and $\{F_i\}$ is the corresponding set of focal elements.

The notion of least prejudiced distribution provides a mechanism by which we can, in a sense, convert a fuzzy set into probability distribution.

Lawry [18] shows an alternative interpretation of linguistic variables:

Definition 3 [6] A linguistic variable is a quadruple $\langle L, |T(L), \Omega, M\rangle$ in which L is the name of the variable, $T(L)$ is a countable term set of label or words (i.e., the linguistic values), Ω is a universe of discourse and M is a semantic rule.

The semantic rule M is defined as a function that associates a normalized fuzzy subset of Ω with each word in $T(L)$. The fuzzy set $M(w)$ can be viewed as encoding the meaning of w so that for $u \in U$ the membership value $\mu_{M(w)}(u)$ quantifies the suitability or applicability of the word w as a label for the value u. It is possible to regard the semantic function M as being determined by a group voting model across a population of voters as follows. Each voter is a asked to provide the subset of words from the finite set $T(L)$ which are appropriate as labels for the value u. The membership value $\mu_{M(w)}(u)$ is taken to be the proportion of voters who include w in their set of labels.

Definition 4 [19, 20] Let $x \in \Omega$. Then the linguistic description x relative to the linguistic variable L is the fuzzy subset of $T(L)$

$$\mathrm{des}_L(x) = \sum_{w \in T(L)} w \Big/ \mu_{M(w)}(x). \tag{7}$$

This notion can be extended to the case where the value given is a fuzzy subset of Ω in which case the appropriate linguistic description is defined as follows.

$$\forall w \in T(L) \;\; \mathbf{Pr}(w|\mathrm{des}_L(S)) = \frac{1}{|S|} \sum_{x \in S} \mathbf{Pr}(w|\mathrm{des}_L(x)). \tag{8}$$

The voting model interpretation of this definition would be that we select an element at random from S and present it to a randomly selected voter from the population and ask him or her to select a single word to label it. In this case $\mathbf{Pr}(w|\mathrm{des}_L(x))$ corresponds to the probability that word w is selected.

If $|S|=1$, i.e., if each label is taken as a subset and considering (7) and (8), then

$$\mathbf{Pr}(w|\mathrm{des}_L(x)) = \sum_{w \in T(L)} w \Big/ \mu_{M(w)}(x). \tag{9}$$

3.2.3 SMC's Diagnosis and Updating

During the management of FL-CAT, updating the student's knowledge distribution is computed by (3). The CCO is obtained by the transformation of expert teacher's valuation O_{ij}^{ek}, as follows. For simplicity we call k to O_{ij}^{ek}.

1. It is defined the linguistic description of k:

$$\mathrm{des}_L(k) = \sum_{w \in S(L)} w \Big/ \mu_{M(w)}(k),$$

where $\mu_{M(w)}(k)$ quantifies the suitability or applicability of the word w as a label for k.

2. Considering each level k as a subset and by (9), then:

$$\mathbf{Pr}(w|\mathrm{des}_L(k)) = \sum_{w \in S(L)} w \Big/ \mu_{M(w)}(k)$$

3. The aggregated probability of $w \in S(L)$ is obtained as follows:

$$\mathbf{Pr}_a(w|\mathrm{des}_L(k)) = A(\mathbf{Pr}(w|\mathrm{des}_L(k))), \tag{10}$$

where A is weighted average of the labels w that describe k.

Table 1 Experts teacher's valuation

k	O_{34}^{ek}	O_{34}^{ek}	O_{34}^{ek}	O_{34}^{ek}	O_{34}^{ek}
1	VL	VL	VL	VL	LO
2	VL	VL	VL	LO	LO
3	LO	LO	LO	VL	VL
4	VH	HI	HI	VH	VH
5	VH	VH	VH	VH	HI

3.3 Performance

Let us suppose that evaluating a student by FL-CAT in evidence node evk, whether the probability distribution of competence levels, is as follows:

$$\left.\begin{array}{l} P(\theta_{\mathrm{evk}} = 1|\vec{u}_i) = 0,10 \\ P(\theta_{\mathrm{evk}} = 2|\vec{u}_i) = 0,10 \\ P(\theta_{\mathrm{evk}} = 3|\vec{u}_i) = 0,30 \\ P(\theta_{\mathrm{evk}} = 4|\vec{u}_i) = 0,40 \\ P(\theta_{\mathrm{evk}} = 5|\vec{u}_i) = 0,10 \end{array}\right\} \theta_{\mathrm{evk}} = k = 4$$

Let us suppose the FL-CAT's adaptation mechanism selected the item I_3 and expert teacher's valuation (O_{ij}^{ek}) of I_3 for the option 4 (See Table 1):

The transformation of expert teacher's valuation O_{ij}^{ek}, is shown below:

$$O_{ij}^{e1} = \begin{cases} \mathrm{des}_L(1) = \Pr(w|\mathrm{des}_L(1)) = \mathrm{VL}/0.8 + \mathrm{LO}/0.2 \\ \Pr(\mathrm{VL}|\mathrm{des}_L(1)) = 0.8 \\ \Pr(\mathrm{LO}|\mathrm{des}_L(1)) = 0.2 \\ \Pr_a(w|\mathrm{des}_L(1)) = \dfrac{0 * 0.8 + 0.25 * 0.2}{0.8 + 0.2} = 0.05 \end{cases}$$

In the same way for the rest O_{ij}^{ek}'s k levels. Therefore, the new probability distribution of competence levels in evidence node evk as follows:

$$\left.\begin{array}{l} |P(\theta_{\mathrm{evk}} = 1|\vec{u}_i)P_3(\vec{u}_4|\theta_1)| = |0.10 * 0.05| = 0.01 \\ |P(\theta_{\mathrm{evk}} = 2|\vec{u}_i)P_3(\vec{u}_4|\theta_2)| = |0.10 * 0.10| = 0.02 \\ |P(\theta_{\mathrm{evk}} = 3|\vec{u}_i)P_3(\vec{u}_4|\theta_3)| = |0.30 * 0.15| = 0.09 \\ |P(\theta_{\mathrm{evk}} = 4|\vec{u}_i)P_3(\vec{u}_4|\theta_4)| = |0.40 * 0.90| = 0.70 \\ |P(\theta_{\mathrm{evk}} = 5|\vec{u}_i)P_3(\vec{u}_4|\theta_5)| = |0.10 * 0.95| = 0.18 \end{array}\right\} \theta_{\mathrm{evk}} = K = 4$$

Observe that de value of θ_{evk} converges to four now more precision (probability).

4 Conclusions

We have presented a new approach for modeled of the teachers' knowledge that replaces the CCO's statistical calibration, avoiding that costly process.

If in the design of an ITS-C the CCO's evaluations of the experts are converted into probabilities, as shown in Sect. 3, then the implementation of ITS-C the performance of the FL-CAT will have similar performance to the CAT.

Therefore, the proposal is a viable alternative for the ITS-C that can be implemented in small institutions, opening new possibilities for ITS-C application.

References

1. Wenger E (1987) Artificial intelligence and tutoring systems. Morgan Kaufmann Publishers, Inc, San Francisco
2. Bourdeau J, Grandbastien M (2010) Modeling tutoring knowledge. In: Nkambou R, Bourdeau J, Mizoguchi R (eds) Advances in intelligent tutoring systems. Springer, Berlin, Heidelberg, pp 123–143
3. Nkambou R (2010) Modeling the domain: an introduction to the expert module. In: Nkambou R, Bourdeau J, Mizoguchi R (eds) Advances in intelligent tutoring systems. Springer, Berlin, Heidelberg, pp 15–32
4. Badaracco M, Martínez L (2011) An intelligent tutoring system architecture for competency-based learning. In: König A et al (eds) Knowlege-based and intelligent information and engineering systems. Springer, Berlin, Heidelberg, pp 124–133
5. Guzmán E, Conejo R, Pérez de la Cruz J (2007) Adaptive testing for hierarchical student models. User Model User-Adap Inter 17(1):119–157
6. Zadeh LA (1975) The concept of a linguistic variable and its application to approximate reasoning—I. Inf Sci 8(3):199–249
7. Zadeh LA (1975) The concept of a linguistic variable and its application to approximate reasoning—II. Inf Sci 8(4):301–357
8. Zadeh LA (1975) The concept of a linguistic variable and its application to approximate reasoning—III. Inf Sci 9(5):43–80
9. Badaracco M, Martínez L (2012) A fuzzy linguistic algorithm for adaptive test in intelligent tutoring system based on competences. Expert Syst Appl 40(8):3073–3086
10. Catalano A, Avolio de Cols S, Sladogna M (eds) (2004) Curriculum design based on standards of occupational competence. Concepts and methodological guidelines. CINTERFOR OIT: Buenos Aires
11. Zalba E, Gutierrez N (2006) An approach to education based on competencies. Available: http://www.me.gov.ar/spu/guia_tematica/CPRES/cpres-comision.html. Accessed 20 Feb 2008
12. Europe TUNING E.S. (2000) Una introducción a la contribución de las universidades al proceso de Bolonia. Available: http://tuning.unideusto.org/tuningeu/images/stories/Template/General_Brochure_Spanish_version.pdf. Accessed 20 Sept 2009
13. Ramsay JO (1991) Kernel smoothing approaches to nonparametric item characteristic curve estimation. Psychometrika 56:611–630
14. Owen RJ (1969) A bayesian approach to tailored testing (research report no. 69). In: Linden WJvd (ed) Educational testing service. Princeton, New Jersey, pp 69–92

15. Bonissone PP, Decker KS (1986) Selecting uncertainty calculi and granularity: an experiment in trading-off precision and complexity. In: Kanal ELH, Lemmer JF (ed) Uncertainty in artificial intelligence. North-Holland Publishing Company, pp 217–247
16. Bonissone PP (1982) Approximate reasoning in decision analysis. In: Gupta MM, Sánchez E (eds) North-Holland Publishing Company
17. Bordogna G, Pasi G (1993) A fuzzy linguistic approach generalizing boolean information retrieval: a model and its evaluation. J Am Soc Inf Sci 44(2):70–82
18. Lawry J (2001) A methodology for computing with words. Int J Approximate Reasoning 28:51–89
19. Baldwin JF (1992) The management of fuzzy and probabilistic uncertainties for knowledge based systems. In: Shapiro SA (ed) Encyclopaedia of Al. Wiley, New York, pp 528–537
20. Baldwin JF, Lawry J, Martin TP (1998) Mass assignment based induction of decision trees on words. In: IPMU98