Bayesian Network Modelling for Improved Knowledge Management of the Expert Model in the Intelligent Tutoring System

Fatima-Zohra Hibbi, Otman Abdoun, El Khatir Haimoudi
Computer Science Department, Advanced Science and Technologies Laboratory
Polydisciplinary Faculty, Abdelmalek Essaadi University, Larache, Morocco

***Abstract*—The expert module is an essential part of the intelligent tutoring system. This module uses only declarative knowledge, excluding other types of domain knowledge: procedural and conditional. This elimination makes the expert module very delicate. To solve this issue, the authors propose to embed knowledge processing into the expert model. The contribution aims to empower the expert model via the fragmentation of the knowledge process into four categories: Analyzation, Application, Conceptualization, and Experimentation using the Bayesian Network method as an instrument for modelling expert systems in uncertain areas. According to the management of the expert system through a list of criteria, the expert module can suggest the correct type of knowledge and their following status.**

Keywords*—*Smart tutoring system; expert model; knowledge processing; Bayesian network

I. Introduction

The accompaniment of students is a very important function in a learning situation, and its guarantee promotes motivation, attendance and the smooth running of all learning activities. For this purpose, the researchers tried to integrate techniques of Artificial Intelligence to propose accompanying mechanisms to guarantee the users' assiduity. Several contributions have focused on improving the intelligent tutoring system to make it more efficient, such as the integration of the two tutors: implicit and explicit, or each tutor uses his strategies and toolbox. The choice of one of the tutors is based on the following criteria: learner profile, effective feedback and learning progress data [1]; the latter needs Learning Analytics to collect the traces that learners leave behind and the uses these traces to improve learning [2].

The tutoring system is based during their interventions on the interconnection of these different models and especially the expert model. The knowledge representation in the expert module explores a set of strategies such as schemes, conditions, etc. The orientation of the system is based on the use of the knowledge of the expert module. Some systems use many techniques to formalize the expert module.

The CREAM system is one of the creation devices that offer the user to write specific objectives independently of the subjects. Thus, it uses various terminologies for the acquired skills [17].

However, the classical expert model in the previous intelligent tutoring system handles just declarative knowledge, which means the tutor intervenes in the activities in which the learner needs only definition and examples. This implies a lack of the taxonomy of the knowledge domain. To enhance the learner evaluation the authors [3] proposed a pedagogical solution based on the combination of three bits of knowledge: Declarative, procedural and conditional according to their status.

To enlighten the purpose of this work it would be noteworthy that the work presented in this paper is related to our previous work named: "Knowledge Management in the Expert Model of the Smart Tutoring System". In other words, the objective of the previous work was the management of the knowledge and subdivided into three categories: declarative, procedural and conditional knowledge in the expert model. To confirm this assumption, they investigate the efficiency of the proposed solution presented in the paper above. This article presents the experiments and results of knowledge processing using a Bayesian network technique and shows how the proposed approach can be applied and enhance the learner evaluation. To measure the performance of knowledge processing; they will use metrics, that evaluate the efficiency of each node dynamically in real-time.

The article is organized as follow: the second section describes the problem of the expert model in the smart tutoring system and clarifies the proposed solution; after that, the authors illustrate the construction of the Bayesian network for our model; then, section four details the conditional probability table. Last but not least, the implementation phase is subdivided into three subsections presented as bellow: the Bayesian network program, the results and the evaluation part. Finally, this article will be ended with a conclusion and future works.

II. Related Works

A. *The Expert Model of Smart Tutoring System*

An intelligent tutoring system (ITS) is a system that provides feedback and helps learners in their learning process such as remediation in a specific field like mathematics [4], or facilitating computer theory [5]. ITS contains four components; the first one is the interface model which can help the learner in a task via a learning environment; the second one is the learner model that has the personal information of the learner, their preferences and characteristics; then, the instructional model which makes a decision about teaching method via a diagnostic process of the learner model; the last one is the expert model which has the representation of the declarative knowledge [3].

The expert module allows referencing an expert or a domain model. It provides a description of the knowledge or behaviours that represent the expertise in the domain [6]. It is a module where the main information is unrolled and will be taught. The best expert module is the most designed and appropriate [7]. This system must be performed in accordance with the knowledge that is available in the Learner module [8]. The expert module has two main functions [9], [10], [7] to elaborate questions, answers and explanations, and to act as a reference.

Several systems use different tools to model the expert module; among those techniques: the fuzzy cognitive map [11]. Another tool that uses thematic maps is the EON tool, which allows the creation of intelligent tutors oriented towards pedagogy [12] etc. The majority of these tools have a feature of authoring instructional objectives and the lack of the taxonomy of learning inside the expert modules. For that reason, in our previous work, the authors proposed to integrate the implicit and explicit intervention to ITS [2]. This tutor will select the appropriate strategy according to three criteria. This modification is concerned with the instructional model; this later has bidirectional communication with the expert model; this module is a computer representation of declarative knowledge, this knowledge allows the ITS to compare the learner's actions and choices with those of an expert in order to evaluate what he understands and what he does not understand [13], [14]. This module encounters a lot of issues, for that reason the authors propose the solution illustrated in Fig. 1.

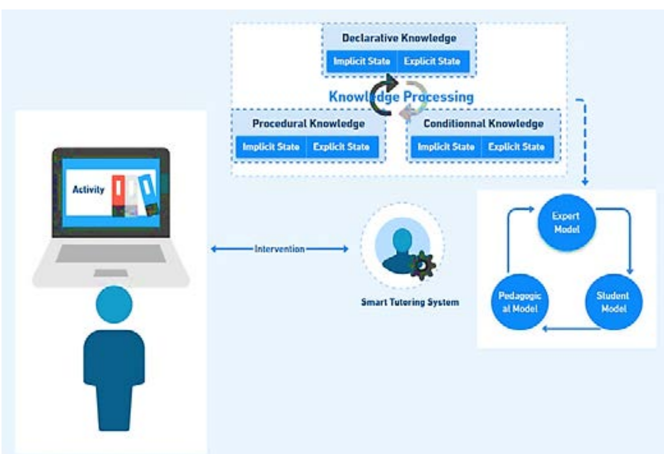

Fig. 1. Knowledge Processing in the Expert Model [3].

This solution combined the three categories of knowledge in the expert model. Declarative: definition of the concepts, procedural: problem-solving ability and conditional knowledge may have an implicit or explicit status, maybe both; and a type of assessment (memorization, administration, Expertise). The modelization of these large numbers of states that have a dispersed representation required the Bayesian network modelling. The choice of Bayesian Network is not arbitrary. This type of network is versatile: they can use the same model to evaluate, predict, diagnose, or optimize decisions, which helps to make the effort of building the Bayesian network profitable.

Therefore, the graphical representation of a Bayesian network is explicit, intuitive and understandable by a non-specialist, which facilitates both the validation of the model, its eventual evolutions and especially their uses.

B. The Development of Bayesian Network

Bayesian networks (BNs) are a tool for representing uncertainty using probabilities and robust mathematical foundations. BNs correspond to the probability distributions that can be generated by products of conditional probability distributions [3]. Several models are created based on knowledge representation and reasoning. Probabilistic graphical models, especially the Bayesian networks initiated by Pearl [15], have appeared as tools to describe uncertain knowledge and thoughts based on limited information.

Bayesian networks have been proven successful in modelling any kind of knowledge problem. Thus, Bayesian networks have been applied in numerous diverse fields, including medical diagnostics, information retrieval, and marketing. There are a lot of tools to create efficient Bayesian networks in an easy way like GeNIe [18] and SMILE [19]. Based on the literature, Bayesian Networks have been a positive effect in the student model of the ITS. ANDES is an example of an intelligent tutoring system which used Bayesian networks to conduct a long-term knowledge evaluation, and prediction of learners' behaviors while problem solving [20].

A Bayesian network BN = (G, N) is characterized by

G = (Y, E) directed acyclic graph with a set of vertices associated and random parameters, which is expressed by the following formula:

$$N = P\ (Y_i — Pa\ (Y_i)) \tag{1}$$

With $Y= (Y,. . ..,Yn)$

In a real-world application, the modelization of a large number of states that have a dispersed representation is indispensable. For that reason, to generate a classification with a representation that allows modelling with many variables a Bayesian network method is required [16]. The graphical representation of the BN describes the conditional independencies between variables which is easy to figure the joint probability. The Fig. 2 presents the Bayesian Network schema:

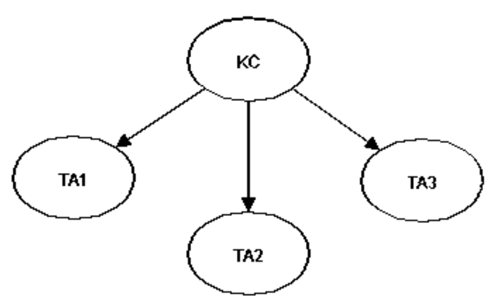

Fig. 2. Bayesian Network Schema.

To construct the Bayesian network model for our issue, two essential steps are required: specification of the structure model and specification of variable values.

1) The specification of the structure model: To describe the development of the Bayesian network, the authors start by defining the nodes of our model; the model involves an initial node "knowledge category", which is composed of several parent nodes presented as bellow: typeActivity1, typeActivity2 and typeActivity3. Each of these nodes contain child nodes. The links to these nodes are prerequisite relationships:

a) TypeActivity 1: The activity is based on a declarative knowledge; this type contains two nodes: Memorization: this activity uses a definition or examples of the concept and status; this is a type of evaluation, it reflects the learner's knowledge; the states of knowledge can be: Implicit: the learner knows what this is meaning, or explicit: the learner knows how to describe it. Fig. 3 presents the node "TypeActivity1" and their child.

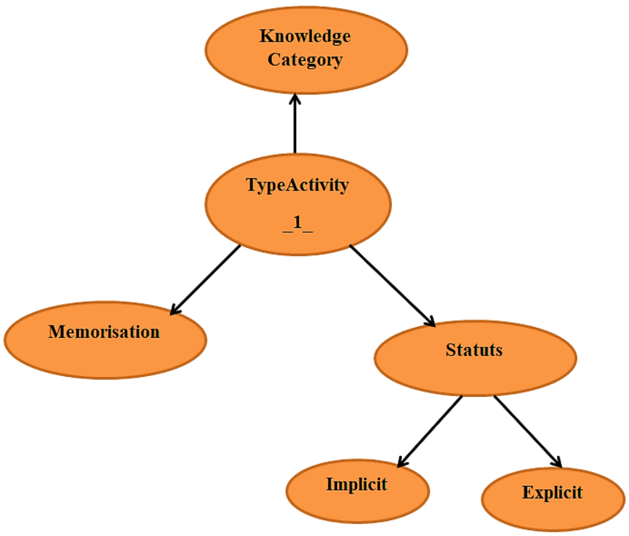

Fig. 3. The Node"TypeActivity1" and the Descendant.

b) TypeActivity 2: The activity is based on a procedural knowledge; it is composed of Administration: the activity applies case studies; and the learner tries to figure out how to do the exercise (Explicit status) and how can do it (implicit status). Fig. 4 describes the node "TypeActivity2" and their child: Administration and status (implicit and explicit).

c) TypeActivity 3: This type of activity is based on conditional knowledge; it comprises: Expertise exercises. The activity develops the ability to realize a project based on both activities described above: TypeActivity1 and TypeActivity2; and the states of knowledge. The Fig. 5 illustrates the Bayesian network of the "TypeActivity3" as a node and their child.

2) The specification of variable values: After presenting the network model, the authors will define the values of the variable. In the construction of Bayesian network, the authors observe that the knowledge category depends of the type of activity. Which make us deduce that is the diagnostic relationship. The initial node Knowledge Category (KC) comprises three parents: typeActivity1 (TA1), typeActivity2 (TA2) and typeActivity3 (TA3) which are corresponding to three weight: w1=0.2, w2=0.3 and w3=0.5. The conditional probability of KC is calculated using the formula bellow:

$$P\ (KC|TA1,\ TA2,\ TA3)=w1*h1+w2*h2+w3*h3 \tag{2}$$

With:

$$h1 = \begin{cases} 1\ if\ TA\ 1 = KC \\ 0\ Otherwhise \end{cases}$$

$$h2 = \begin{cases} 1\ if\ TA\ 2 = KC \\ 0\ Otherwhise \end{cases}$$

$$h3 = \begin{cases} 1\ if\ TA\ 3 = KC \\ 0\ Otherwhise \end{cases}$$

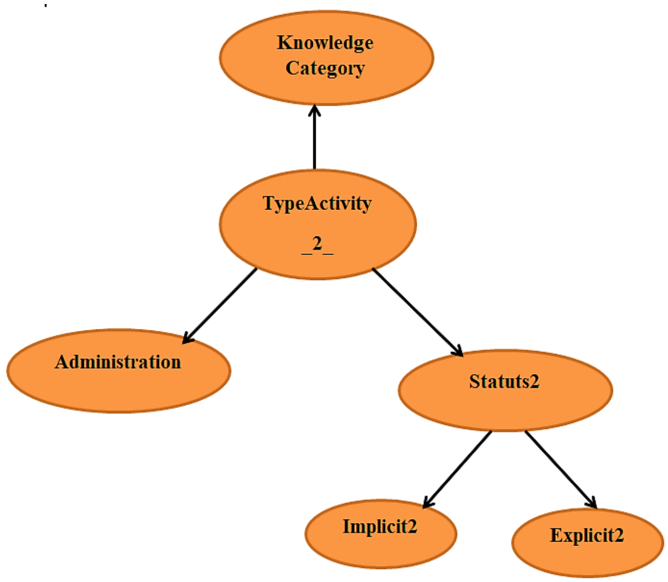

Fig. 4. The Node "TypeActivity2" and their Child.

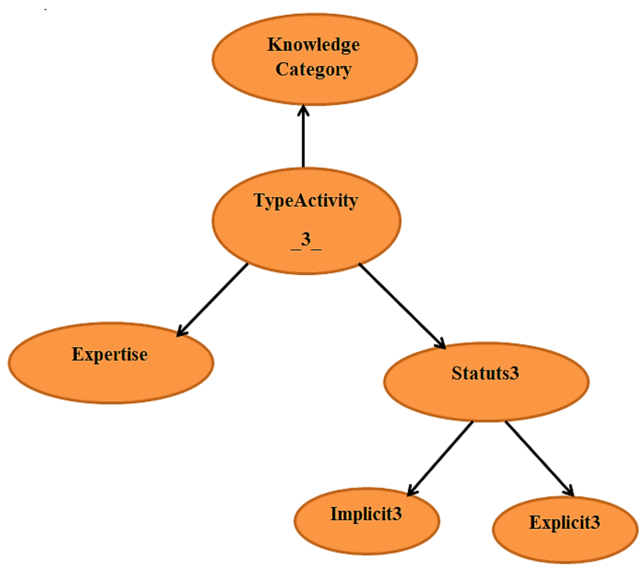

Fig. 5. The Node "TypeActivity3" and their Descendants.

The authors mention that {KC, AT1, AT2, AT3} is a complete set of mutually exclusive variables, whose variable is also random and binary. As a generalization on the formula presented in (3) and (4), it can be said that:

$$P(X=1|Y1, Y2, \dots Yn)= wi*hi, \{i=1,\dots,n\} \quad (3)$$

$$P(\text{not } X|Y1, Y2\dots Yn) =1-P(X=1|Y1, Y2, \dots Yn) \quad (4)$$

The Table I presents the conditional probability of the node “knowledge category”.

TABLE I. THE CONDITIONAL PROBABILITY TABLE OF THE NODE “KNOWLEDGE CATEGORY”

TA1	TA2	TA3	P(X=1)	P(X=0) =1-P(X=1)
0	0	0	0,0	1,0
0	0	1	0,5	0,5
0	1	0	0,3	0,7
0	1	1	0,8	0,2
1	0	0	0,2	0,8
1	0	1	0,7	0,3
1	1	0	0,5	0,5
1	1	1	1,0	0,0

III. IMPLEMENTATION PHASE

In this section, the authors will describe the environments, present the results and evaluate the efficiency of the proposed solution:

A. Environments

After the construction of our Bayesian network model, the authors validate the proposed solution using the ANACONDA navigator. Spyder 3.3.6 is a scientific Python development environment and they work with python 3.7 as a programming language. In the phase of coding, they apply the POMEGRANATE package for building probabilistic models. Based on the conditional probability table that they built in the previous section, they initialized the status and activity using Univariate distribution (), which made up of characters and their probabilities. The sum of the probability will be 1.0. The Fig. 6 presents the Parameters of the development environment:

B. The Results

Fig. 7 illustrates the probability of activity containing conditional knowledge; the result shows us that this activity should have an expertise activity, including the implicit and explicit status.

The probability of an activity containing declarative knowledge is that it should have a memorization activity, including the implicit and explicit status. The results are described in the Fig. 8.

The Fig. 9 presents the probability of activity containing procedural knowledge; the result shows us this activity should have an administration activity including the implicit and explicit status:

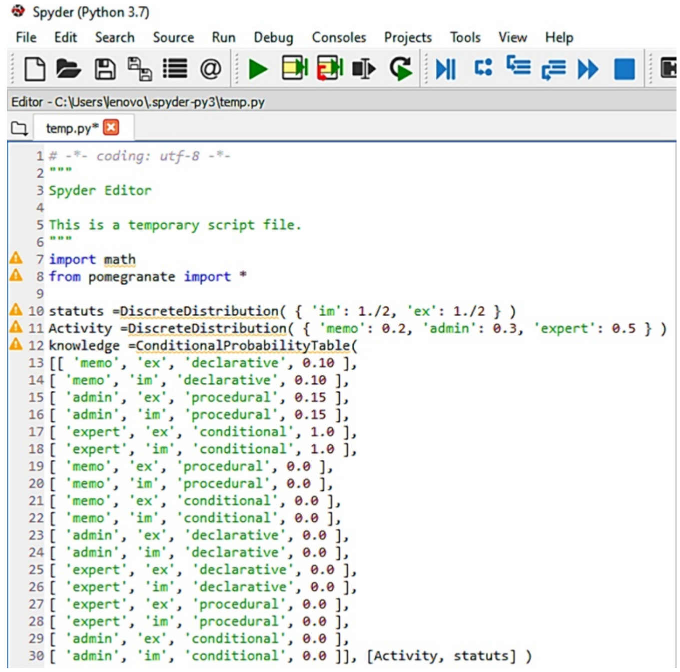

Fig. 6. Parameters of the Development Environment.

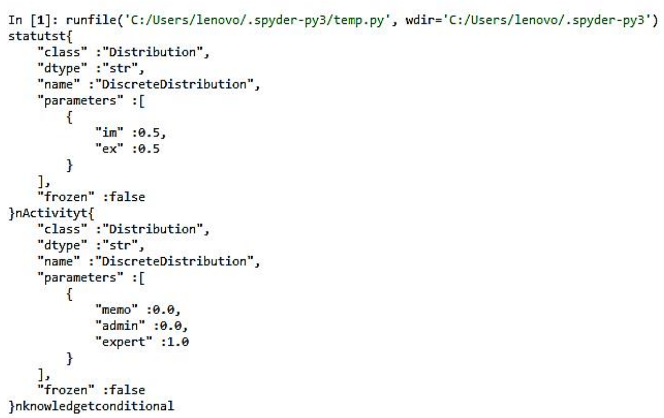

Fig. 7. The Probability to have a Conditional Knowledge.

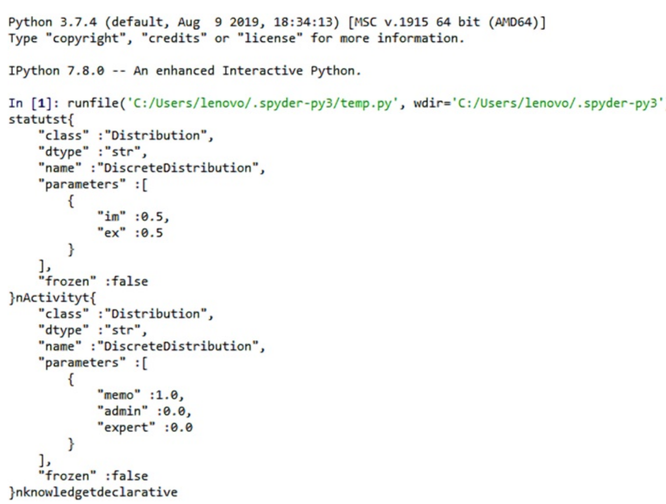

Fig. 8. The Probability to have a Declarative Knowledge.

```
In [2]: runfile('C:/Users/lenovo/.spyder-py3/temp.py', wdir='C:/Users/lenovo/.spyder-py3')
statutst{
    "class" :"Distribution",
    "dtype" :"str",
    "name" :"DiscreteDistribution",
    "parameters" :[
        {
            "im" :0.5,
            "ex" :0.5
        }
    ],
    "frozen" :false
}nActivityt{
    "class" :"Distribution",
    "dtype" :"str",
    "name" :"DiscreteDistribution",
    "parameters" :[
        {
            "memo" :0.0,
            "admin" :1.0,
            "expert" :0.0
        }
    ],
    "frozen" :false
}nknowledgetprocedural
```

Fig. 9. The Probability to have a Procedural Knowledge.

To summarize the results illustrated in the figures above, the authors can conclude the following points:

To create an activity that can attend the conditional knowledge; a pedagogical sequence must contain an expertise exercise, such as the ability to choose the appropriate method to solve a problem in a case study.

To achieve the declarative knowledge, the activity should have a memorization exercise like the recognition, categorization and differentiation between different objects (implicit status) and schematization, graphic, symbolic, linguistic (Explicit status).

A pedagogical sequence based on procedural knowledge should contain administration exercises in such a way that the learner tries to figure out how to do the exercise (Explicit status) and he can do it (Implicit status).

C. Evaluation and Discussion

To measure the performance of knowledge processing, the authors used a lot of metrics; the first one is the calculation of the target node and all the evidences using the global confusion matrix (GCM); the obtained results from the GCM helps to know the probability of the correct classification (PCC), after that, they calculated the marginal probability of the correct classification (MPCC), and finally they obtained the marginal improvement (MI), individual PCC (IPCC) and the cost rate.

To evaluate the performance of the Bayesian network, validation of each node is necessary. The Fig. 10 illustrates the node evaluation "Type Activity1" as a target and their two fragments of evidence: "Memorization" and "Status": implicit and explicit.

From the results, the authors find that by adding the evidence nodes in the evaluation of the target node, the percentage probability of correct classifications is rising. By measuring the probability of each node's correct classification, they see how each node contributes separately to the classification. In this evaluation, the "Memorization" node is the biggest contributor. In this evaluation, they find the influence of the node "Memorization" to the target node.

The results reflect the choice of the TypeActivity1 which means that the expert model will concentrate the knowledge on the declarative one; for example: definition and explanation with the uses of (58, 70%) as an explicit knowledge likes the graphs and schemas; and, (56, 90%) as an implicit knowledge which contains a recognition exercises.

According to the results and the validation of each node of the Bayesian network, the authors were able to handle globally the functioning of the network. According to the management of the expert system through a list of criteria that contains the type of knowledge as well as their status; the expert model can suggest knowledge in the right type as well as the appropriate status. If, for example, the expert system chooses the type of activity1, then suggestions can be demonstrated based on declarative knowledge (memorization) and appropriate status (implicit/explicit).

The Bayesian network modelling helps us to modelize the large number of states that have a dispersed representation. This modelization helps to make a prediction and evaluation of the three categories of knowledge and their child: type of assessment and status.

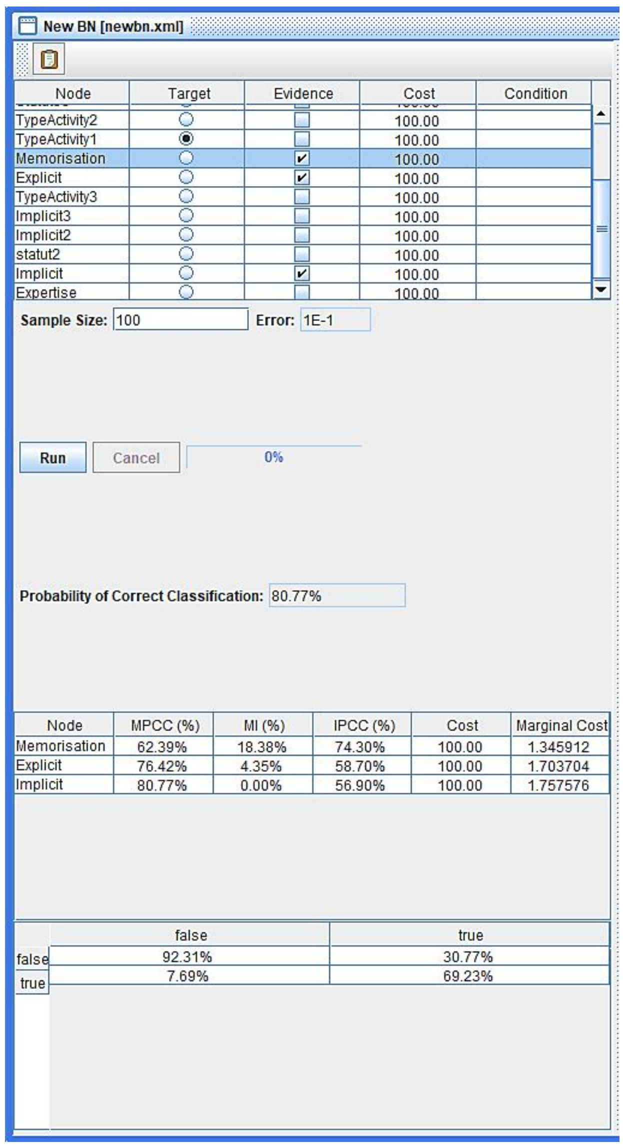

Fig. 10. The Node Evaluation "Type Activity1.

IV. Conclusion

The evaluation part is an essential step in the learning process. To enhance this step in the expert model of intelligent tutoring system, knowledge processing is required. In this paper, the authors have presented the problem of the expert model; after that, they illustrate the construction of the Bayesian network; then, they evaluate the efficiency of the proposed process and the results in the implementation phase. This paper aims to investigate the efficiency of the pedagogical solution presented in their previous research paper via the Bayesian network model. This study analyses the knowledge processing according to the type of activity; this later depends on the knowledge and the status. In future work, we aim to study the process of the selected knowledge, and we will integrate the intelligent tutoring system into the FPL E-learning platform and the different contributions in one solution.

References

[1] F.-Z. Hibbi, O. Abdoun, en H. El Khatir, “Extract tacit knowledge in the learner model of the smart tutoring system”, Int. J. Emerg. Technol. Learn., vol 15, no 04, bl 235, Feb 2020.

[2] F.-Z. Hibbi, O. Abdoun, en E. L. K. Haimoudi, “Exploration of Analytical Mechanisms in the Feedback model”, Procedia Comput. Sci., vol 148, bll 201–207, 2019.

[3] F.-Z. Hibbi, O. Abdoun, en E. K. Haimoudi, “Knowledge management in the expert model of the smart tutoring system”, in Proceedings of the 3rd International Conference on Networking, Information Systems and Security, Marrakech Morocco, 2020.

[4] M. Mohammed A en S. S. Al-Nakhal, Adaptive Intelligent Tutoring System for Learning Computer Theory. 2017.

[5] R. Nkambou en C. Frasson, “A New Approach to IT Scurriculum and Course Authoring: The Authoring Environment, Elsevier Science Ltd”, Pergamon Computers Educ, great Britian, vol 31, bll 105–130, 1998.

[6] M. A. Mark en J. E. Greer, “Evaluation Methodologies for Intelligent Tutoring Systems”, Journal of Artificial Intelligence in Education, 1999.

[7] M. Grandbastien, Teaching Expertise is at the Core of ITS Research, Interna-tional journal of Artificial Intelligence in Education. 1999.

[8] R. Mcnally en J. Shute, “Intelligent Tutoring Systems: Past, Present and Future”, in Communications and Technology, D. Jonassen, Red Macmillan, 1996.

[9] J. A. Self, “The Distinctive Characteristics of Intelligent tutoring Systems Re-search: ITSs Care, Precisely”, International Journal of Artificial Intelligence in Edu-cations, 1999.

[10] T. Murray, “Authoring knowledge-based tutors: Tools for content, instructional strategy, student model, and interface design”, J. Learn. Sci., vol 7, no 1, bll 5–64, Jan 1998.

[11] J. Pearl, Probabilistic reasoning in intelligent systems. London, England: Morgan Kaufmann, 2014.

[12] J. Pearl, Probabilistic Reasoning in Intelligent Systems. Elsevier Science, 2014.

[13] H. E. Kyburg, “Probabilistic reasoning in intelligent systems: Networks of plausible inference”, J. Philos., vol 88, no 8, bl 434, Aug 1991.

[14] Pipatsarun Phobun and Jiracha Vicheanpanya, ”Adaptive intelligent tutoring system for e-learning systems”, ELSEVIER. 2010.

[15] E. Castillo, J. M. Gutierrez, en A. S. Hadi, Expert Systems and Probabilistic Network Models”. Springer, 1997.

[16] R. Nkambou, R. Gauthier, en M. C. Frasson, “CREAM-Tools: an authoring environment for curriculum and course building in an ITS”, in Proceedings of the Third International Conference on Computer Aided Learning and Instructional Science and Engineering, New York: Springer Verlag, 1996.

[17] N. Jongsawat, Anunucha, en Wichian, “Dynamic data feed to Bayesian network model and SMILE web application”, in Bayesian Network, Sciyo, 2010.

[18] N. Jongsawat, P. Poompuang, en W. Premchaiswadi, “Dynamic data feed to Bayesian network model and SMILE web application”, in 2008 N Ninth ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing, Phuket, Thailand, 2008.

[19] M. J. Smile, “Structural modeling, inference, and learning engine and genie: A development environment for graphical decision-theoretic models”, in Proceedings of the Sixteenth National Conference on Artificial Intelligence (AAAI-99), 1999, bll 18–22.

[20] E. Millan en J. L. P ´ erez-De-La-Cruz, “A Bayesian diagnostic algorithm for student modeling and its evaluation”, User Modeling and UserAdapted Interaction, vol 12, no 2–3, bll 281–330, 2002.