1. Introduction

SLSs are intelligent systems that imitate the human mind. The main characteristics of these systems are the ability of inference, reasoning, perception, learning, and knowledge-based systems. To a limited degree, AI permits SLS to accept knowledge from human input, and then use that knowledge through simulated thought and reasoning processes to solve problems. Many types of SLSs are in existence today and are applied to different domains and tasks, e.g., geology, biological sciences, medical sciences, health care, commerce, and education.

Scientists have applied the KE techniques and methodologies to grow smart tutoring and learning systems [1]. Moreover, the confluence of AI, KE and web science is enabling the conception of a new era of web-based smart systems for all educational and learning tasks. Smart learning represents an assortment of smart services that are based on smart digital media and communication and information technologies for supporting learning, training and educational processes [2–5]. For that reason, SLSs are complex to build and complex to maintain.

Based on the recent research during the last five years, knowledge engineers and software developers have started to investigate the usage of KE to develop newer and robust SLS [6–8]. The main properties and characteristics of these systems are the ability of perception, inference, reasoning, learning, thinking, and knowledge-based. Our main objective is to extract and discover the critical aspects and main advantages of these paradigms for the SLS development.

The structure of the paper includes the following sections: the second section makes a review of research literature on the problem, the third section indicates the aim and the tasks of the paper, the fourth section introduces a summary of the knowledge management and representation techniques for developing the SLS. In the fifth section, we analyze three intelligent approaches applied by the knowledge engineers to develop SLS, namely data mining (DM), case-based reasoning (CBR), and ontological engineering (OE). Sections six and seventh discuss the benefits and challenges of such approaches. The last section draws conclusions and future work.

2. Literature review and problem statement

Many of KE and intelligent algorithms are used in the developing of smart learning systems, e.g., artificial neural networks, support vector machines, genetic algorithms, decision trees, and fuzzy logic. This section reviews some of the most important aspects of knowledge engineering
and AI techniques that are used for smart learning systems developments.

Knowledge could be an insufficiently clarified term. Research is still being done on how knowledge can be modeled so it can be manipulated and processed by a computer. Knowledge representation and reasoning methodologies receive increasing attention within the community of smart computing applications and AI in education [2, 10, 11, 28].

Among the many approaches applied by the knowledge engineers are semantic networks, frames, production rules and scripts [1, 9]. On the other side, memory organization packets, reminding and explanation patterns, cases, and ontological engineering are general knowledge representation techniques, accounting for the stereotypical and heterogeneous nature of episodic knowledge [1, 9].

From the knowledge engineering perspectives, the ontological engineering paradigm is frequently used by smart computing and information science communities [12]. Currently, there are many uses of ontologies in commercial, industrial, life sciences, and academical domains [13, 29]. The application of ontologies in smart educational and learning systems may be proposed from several points of view: as a chain between heterogeneous educational systems, as a common and approved vocabulary for multi-agent systems, ontologies for pedagogical resources sharing or for sharing data and ontologies used to mediate the search of the learning resources on the web-based environments. The brief specification of a system involves functional interconnected elements. These elements are connected using an intelligent interface and a shared vocabulary [13].

Recently, data mining (DM), case-based reasoning (CBR) and intelligent-agents systems (IAS) are the most common knowledge engineering techniques used by the developers of intelligent systems. DM is not a consistent field, it dwells upon already well-established machine learning and computational intelligence techniques. DM is focused on the discovery of hidden patterns and new rules from large databases [18, 30]. CBR concerns about the reasoning from experiences or “old cases” in an effort to solve new problems, critique solutions, and explain anomalous situations. From the psychological point of view [14], CBR refers to reasoning in which a human problem-solver relies on preceding cases that he or she has encountered. Psychologists have discovered the following facts: (a) people are good at using analogues to solve new problems, (b) people are not always able to remember well the right solutions (computers are better at remembering). From the computational perspective [14], CBR refers to a set of concepts and techniques that may be applied to fulfill the following operations: (a) record and index cases, (b) search for similar case in the case memory to recognize the ones that could be useful in solving new cases, (c) adjust earlier cases to better match new cases, and (d) synthesize and generate new cases. Nowadays, knowledge engineer’s and AI researchers have started to use the CBR concepts in enhancing human decision making through developing case-based reasoning systems [2, 15, 16]. On the other side, intelligent learning systems built based on the IAS approach consists in a set of intelligent agents, which have to communicate and collaborate through messages [17]. Software agents can comprehend and interpret the messages due to an accredited ontology or the interoperability of the private ontologies.

3. The aim and objectives of the study

The aim of the present study is to determine the benefits of the knowledge engineering paradigms to the knowledge engineer designers who are working in the development of smart learning systems.

To accomplish this aim, the following objectives have been set:
1. Evaluating and analyzing the knowledge representation and reasoning techniques, which are used to build the “knowledge base” and “inference engine” for any SLS for a specific task.
2. Exploiting the main features and advantages of three computational intelligence techniques applied by the knowledge engineers to develop SLS, namely data mining, case-based reasoning, and ontological engineering.

4. Knowledge Management and Representation Techniques for SLS

From the knowledge engineering perspective, the main components in developing an efficient and smart learning/educational system for any task involve the “knowledge base” and the “inference engine/reasoning mechanism”. From the KE perspective, Fig. 1 displays the knowledge representation, management and reasoning techniques, which are used to build the “knowledge base” and “inference engine and reasoning mechanism”. In this part of our paper, a summary of such techniques will be explained.

Fig. 1. Knowledge management and reasoning techniques

4.1 Knowledge Representation Techniques for Static and Hierarchical Knowledge

Semantic Networks as a knowledge structure. A Semantic Network (SN) shows relationships among various entities.
SNs are a variety of networks usually applied to structure more general kinds of information. These are groups of nodes and links in which the nodes refer to concepts and the links present the relationships between them. So, SN shows relationships among various entities. They got such a name since they were originally employed to represent the sense of natural language expressions. The application of some type of network for modeling of concepts and relationships is so widespread in AI systems and all types of knowledge-based systems.

**Frames as a knowledge structure.** The frame is grounded on Marvin Minsky’s theory about human thinking. “When one faces a new situation”, he says, “One chooses from memory a substantial structure called a frame”. Following to Minsky’s view, a frame is a data structure for promoting a typecast situation, like being in a certain kind of living room or visiting a child’s birthday party.

Production Rules as a knowledge structure. Production rule is the simplest and most popular method to present knowledge. “If-then” rules are considered as the most typical form of declarative knowledge representation applied in AI applications. Most people are comfortable reading rules, in contrast to knowledge represented in predicate logic. Every rule can be viewed as a separate piece of knowledge or element of information facts in a knowledge base. New knowledge could be easily added, and existing knowledge may be changed simply by creating or modifying individual rules. Therefore, the rule is simple, modular, appropriate size, procedural, and descriptive.

**Scripts as a knowledge structure.** Script is another kind of a knowledge representation technique, which is alike to a frame, but instead of depicting an object, the script depicts a sequence of events. Like the frame, the script represents a stereotyped situation. Dissimilar to the frame, it is typically presented in a specific context. To demonstrate a sequence of events, the script applies the application of a series of slots that include information about the actions, objects and people that are participating in the events. Scripts accounted for information about stereotypical events, e.g. visiting the dentist, catching a bus and visiting a restaurant. In stereotypical events (common situations), a person has a set of expectations of the props, goals, default setting and behaviors of the other people involved. Scripts are equivalent to Minsky’s frames [1, 9], which have been planned in the context of visual processing. Scripts are appropriate to autobiographical events and are inherently episodic in origin and use, i.e. scripts stem from experience and are used to construe new events. Scripts were suggested as a knowledge structure for a conceptual memory. As a psychological theory of memory, scripts proposed that people would recall an event in terms of its associated script.

### 4.2. Knowledge Representation Techniques for Stereotypical Knowledge

**Memory Organization Packets (MOP).** MOPs can be viewed as met scripts, e.g. instead of a physician script or a dentist script, there might be a professional-office-visit MOP that could be instantiated and identified for both the dentist episodes and the physician episodes. This MOP will involve a generic waiting room scene, thus providing the basis for confusion between dentist and physician episodes.

**Reminding and Explanation Patterns.** In [9], a theory of learning based on reminding is proposed. The main features of this theory can be abridged in the following points:

- **Conform-Driven Learning:** When the new situations (or experiences) conform to the past cases and events, Thus we can categorize a new episode in terms of previous cases.
- **Failure-Driven Learning:** WHEN the new situation does not adjust to the prior case, we have a failure. That is, we had an expectation according to a prior event that did not happen in the new situation. THUS we must classify this new experience as different from the precedent episode. We must store this new experience, and we must learn.
- **Discrepancy-Driven Learning:** WHEN we realize a discrepancy between our predictions and some event, THUS we have something to learn and consequently we need to review our knowledge structure.

The mechanism to evolve our knowledge requires explanation. In [9], presented an explicit knowledge structure “explanation patterns”, which is applied to generate, index, and test explanations in combination with an episodic memory is presented.

**Cases as a knowledge structure.** The “case” is a list of attributes/features that lead to a specific outcome, e.g. the information on a patient medical record. Fig. 2 shows the ideal structure of the “case” from the knowledge engineering intelligent perspective. The figure reveals the following: depending on the case structure, the case could be used for a variety of purposes. In Smart Learning Systems, the case can include: (a) a multi-media description of the problem, (b) a description of the correct actions to take including alternative steps, (c) multi-media interpretations of why these steps are correct, and (d) a list of methods to control whether learner/students correctly executed the steps.

Consequently, determining the appropriate case features is the principal knowledge engineering task in developing the case-based memory for any smart software. This task involves defining the domain terminology and collecting representative cases of problem solving by the knowledge engineers. Representation of cases may be in one of various forms (frames, predicate, scribes).

**Ontological Engineering (OE).** The concept ontology represents a common terminology in a specific task. OE refers to the group of activities that concentrate on the ontology development process, the ontology life cycle, and methodologies for building ontologies, as well as the tool suites and languages that support them.

**Fig. 2. The “case” structure**

### 4.3. Reasoning Methodologies for SLS

The field of reasoning is a critical and essential issue for the development of SLS software. The research area in this field covers a variety of topics, e.g. automated reasoning,
probabilistic reasoning, causal reasoning, case-based reasoning, commonsense reasoning, fuzzy reasoning, geometric reasoning, non-monotonic reasoning, model-based reasoning, qualitative reasoning, spatial reasoning and temporal reasoning. In fact, these methodologies receive increasing attention within the development of SLS community. This section deals with reasoning with rules, fuzzy-rules, and cases.

**Reasoning with Rules:** Rules are easily manipulated by reasoning systems. Forward chaining can be used to produce new facts (hence the term "production" rules), and backward chaining can deduce whether statements are true or not. Rule-based systems were one of the first large-scale commercial successes of artificial intelligence research. An expert system or knowledge-based system is the common term used to describe a rule-based processing system. It consists of three major elements, a knowledge base (the set of if-then rules and known facts), a working memory or database of derived facts and data, and an inference engine, which contains the reasoning logic used to process the rules and data.

Smart rule-based systems solve problems by taking an input specification and then "chaining" together the appropriate set of rules from the rule base to arrive at a solution. Given the same exact problem situation, the system will go through exactly the same amount of work to come up with the solution. In other words, rule-based SLSs don’t inherently learn. In addition, given a problem that is outside the system’s original scope, the system often can’t render any assistance. Finally, rule-based SLSs are very time-consuming to build and maintain because rule extraction from experts is labor-intensive and rules are inherently dependent on other rules, making the addition of new knowledge to the system a complex debugging task.

**Forward chaining reasoning (FCR):** FCR is a data-driven reasoning process where a set of rules is used to drive new facts from an initial set of data. It does not use the resolution algorithm used in predicate logic. The forward-chaining algorithm generates new data by the simple and straightforward application or firing of the rules. As an inference procedure, forward chaining is very fast. Forward chaining is also used in real-time monitoring and diagnostic systems where quick identification and response to problems are required.

**Backward chaining reasoning (BCR):** BCR is often called goal-directed inference, because a particular consequence or goal clause is evaluated first, and then we go backward through the rules. Unlike FCR, which uses rules to produce new information, backward chaining uses rules to answer questions about whether a goal clause is true or not. BCR is more focused than forward chaining, because it only processes rules that are relevant to the question. BCR is used for advisory systems, where users ask questions and get asked leading questions to find an answer.

**Reasoning with Fuzzy Rules:** Fuzzy logic deals with truth values which range continuously from 0 to 1. Thus, something could be half true 0.5 or very likely true 0.9 or probably not true 0.1. The use of fuzzy logic in reasoning systems affects not only the inference engine but also the knowledge representation itself. Reasoning with fuzzy rule systems is a forward-chaining procedure. The initial numeric data values are fuzzified, that is, turned into fuzzy values using the membership functions. Instead of a match and conflict resolution phase where we select a triggered rule to fire, in fuzzy systems, all rules are evaluated, because all fuzzy rules can be true to some degree (ranging from 0.0 to 1.0). The antecedent clause truth values are combined using fuzzy logic operators. Next, the fuzzy sets specified in the consequent clauses of all rules are combined, using the rule truth values as scaling factors. The result is a single fuzzy set, which is then defuzzified to return a crisp output value.

**Reasoning with Cases:** The idea of case-based reasoning (CBR) is becoming popular in developing Smart Learning Systems because it automates applications that are based on precedent or that contain incomplete causal models. In rule-based SLSs, an incomplete mode or an environment, which does not take into account all variables could result in either an answer built on incomplete data or simply no answer at all. The CBR approach attempts to get around this shortcoming by inputting and analyzing problem data.

5. Intelligent Approaches for Smart Learning Systems

This section analyzes three intelligent and robust approaches applied by the knowledge engineers to implement the smart learning systems, namely: data mining and knowledge discovery, case-based reasoning and intelligent agents.

5. 1. Data Mining and Knowledge Discovery (DM and KD) Approach

Fig. 3 shows the whole knowledge discovery processes. The preprocessing process is often defined as data cleaning. The cleaned data are located in the data warehouse. This is followed by the DM process and its results are provided to an output visualization generator producing action lists or monitor reports. Each process is maintained by various intelligent methodologies.

The overall KD process includes the evaluation and the adequate interpretation of the mined patterns to identify which patterns may be regarded as new knowledge. Central issues in KD occur from the very nature of databases and the objects (data) they handle. They are categorized in the following way: (a) vast amounts of data, (b) active nature of data, (c) imprecise data, (d) noisy data, (e) missing attribute values, and (f) redundant or insignificant data [19, 20].
In addition, all published studies reveal that [20] DM is supported by a host that acquires the character of data in various tasks, including:

a) Clustering: The main goal is to find natural groupings (clusters) in highly dimensional data. Clustering is an example of unsupervised learning and is a part of pattern recognition. In this respect, K-means algorithms are commonly used.

b) Regression Models: These derived from standard regression analysis and its applied part known as system identification. The underlying idea is to create a linear or nonlinear function. Machine learning techniques, Support Vector Machines, Decision Trees, Rule induction, Neural Networks are preferable techniques to perform this task.

c) Classification: This concerns learning that classifies data into the predefined categories. Regarding this task, a huge number of classifiers have been developed. Based on our analysis, we found that Support Vector Machines, Decision Trees, Neural Networks, Rule induction, and Genetic Algorithms are more appropriate techniques to perform classification tasks.

d) Summarization: This is an approach of describing data with a small number of attributes/features. This task is often applied in an automated report generation and interactive exploratory data analysis through the multivariate visualization approaches.

e) Link analysis: It is concerned about the conception of relationships among database fields. In a particular case, we may be interested in the determination of the link between the variables.

f) Sequence Analysis: this type of task is oriented to problems of modeling sequential data. Pertinent models support the temporal neural networks, time series analysis and time series models.

Recently, researchers investigate different DM methods to help administrators and instructors to enhance e-Learning systems [21–23]. Some of the major e-Learning problems or subjects to which data mining techniques have been applied are dealing with.

5.2. Case-Based Reasoning (CBR) Approach

The typical functional diagram of a CBR methodology is shown in Fig. 4. When a new case has appeared in the system, the problem is indexed, and therefore, the indexes are applied to extract previous similar case or cases from memory. These past cases lead to a set of prior solutions. Subsequently, the previous solutions are modified to adjust to the new situation. Then the proposed solution is attempted. If the solution succeeds, then it is saved as a working solution; if it fails, the working solution must be repaired and tested again. In support of CBR processes, the following rules knowledge structures are necessary to accomplish the resulting tasks: case indexing, case memory, similarity, modification, and repair. The effectiveness of CBR systems depends on the quality and quantity of cases in a case memory.

6. Benefits of Knowledge Engineering Approaches for Smart Learning Systems

Based on our analysis of the three approaches mentioned in the above sections, we can draw the major benefits of DM, CBR, and IA approaches for SLS (Table 1).

<table>
<thead>
<tr>
<th>KE Approach</th>
<th>Benefits for SLS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Mining</td>
<td>– Evaluation of candidate’s learning performance and learning recommendations based on the candidates’ learning behavior.</td>
</tr>
<tr>
<td></td>
<td>– Assessment of learning resources and web-based courses, give feedback to teachers and students of online courses.</td>
</tr>
<tr>
<td></td>
<td>– Detection of new, useful and interesting knowledge based on the candidate’s usage data.</td>
</tr>
<tr>
<td></td>
<td>– Grouping learners/users based on their skills and other characteristics.</td>
</tr>
<tr>
<td></td>
<td>– Identifying learners with little motivation and finding the suitable treatment.</td>
</tr>
<tr>
<td></td>
<td>– Identification of a typical candidate’s learning behavior.</td>
</tr>
<tr>
<td>Case-Based Reasoning</td>
<td>– With more cases available in the case-memory, a learner will have the opportunity to get advantage from the failures of others.</td>
</tr>
<tr>
<td></td>
<td>– Retrieval cases process will allow learners to better recognize what is important in a new situation.</td>
</tr>
<tr>
<td></td>
<td>– CBR system provides the learner with a model of the way decision-making needs to be done.</td>
</tr>
<tr>
<td></td>
<td>– CBR system can augment the learners’ memories of even educators.</td>
</tr>
<tr>
<td></td>
<td>– From the educational perspectives, both educators and learners tend to focus on too few possibilities when reasoning analogically or to concentrate on the wrong cases.</td>
</tr>
<tr>
<td>Intelligent agents</td>
<td>– Acting autonomously.</td>
</tr>
<tr>
<td></td>
<td>– Acting with other software agents.</td>
</tr>
<tr>
<td></td>
<td>– Mimics human interaction types, e.g. cooperation, coordination and negotiation.</td>
</tr>
</tbody>
</table>
7. Discussion and Challenges

From our study of the technical aspects of the KE paradigms, we can drive the main features and characteristics of smart learning systems in the following: (a) it is a knowledge-based system, (b) based on heuristic, declarative, interactive and symbolic processing, (c) convergent reasoning, i.e. producing a few results from the big data analytics, and (d) gives recommendations.

Case-based reasoning methodology organizes knowledge in "cases or examples" of previous problems and their solutions. Such knowledge structure overcomes the knowledge in a lesson-oriented manner and the automatic generation of tests and exercise. Moreover, the CBR methodology addresses the problems of rule-based systems, e.g. knowledge acquisition, performance experience, adaptive solutions and maintenance.

Information and data mining techniques are very encouraging approaches towards the data analytics of learner activities and behavior, which are collected by learning management systems. Smart data mining techniques could promote online learning for the learners. The big challenge in this respect is the choice of the appropriate mining tool to perform a specific task.

Nowadays, intelligent agents’ paradigms were proposed to reinforce the efficacy of smart learning systems through the following dimensions: (a) agents as a modeling and design paradigm for unconventional human-computer interaction and (b) agents for smart functional decomposition of complex systems. Agents’ technologies are often considered as incarnations of different types of AI, including reasoning, knowledge engineering, machine learning and information mining. Research interests in agent systems are extended to several topics such as modeling, design, and development of robust smart learning systems that are appealing for a range of smart applications.

8. Conclusions

The paper has presented a comprehensive analysis of a variety of knowledge engineering and machine learning techniques for developing smart learning systems. From our technical analysis of the knowledge engineering paradigms, we deduced the following conclusions:

1. The fusion of smart computing with the knowledge engineering paradigms, cognitive sciences, data science and web science solves the technical problems and difficulties in designing robust generation of smart learning systems. The web-based nature of such systems can enhance the online learning/education/training processes through the web. The key issues to the success of designing efficient smart learning systems are the selection of the suitable techniques of both knowledge representation and knowledge engineering technique. Case-based reasoning methodology addresses the problems of rule-based learning systems, e.g. knowledge acquisition, performance experience, adaptive solutions and maintenance.

2. Knowledge discovery process and data mining techniques are very encouraging approaches towards the data analytics of learner activities and behavior, which are collected by learning management systems. Intelligent agents have superb benefits to reinforce the efficacy of smart learning systems. The development of smart learning systems is a composite process that promotes a set of technological and research challenges that have to be treated in an interdisciplinary manner. Guaranteeing the success of smart learning systems to the cloud-computing environment is an interesting challenge.

Our future work would revolve around the design of hybrid methods for developing smart learning systems and their application in Web-based as well as mobile-based environments.

References


