Available online at www.sciencedirect.com
ScienceDirect
Journal of Biomedical Informatics 41 (2008) 432–441
Journal of Biomedical Informatics
www.elsevier.com/locate/yjbin

Methodological Review

Evolution and challenges in the design of computational systems for triage assistance

María M. Abad-Grau [a,*], Jorge Ierache [b], Claudio Cervino [c], Paola Sebastiani [d]

[a] Department of Computer Languages and Systems, University of Granada, c/Periodista Daniel Saucedo Aranda, 18071 Granada, Spain
[b] Institute of Intelligent Systems, FICCTE, University of Morón, Morón, Argentina
[c] School of Medicine, University of Morón, Morón, Argentina
[d] Department of Biostatistics, Boston University, School of Public Health, Boston, MA, USA

Received 1 May 2007
Available online 5 February 2008

Abstract

Compared with expert systems for specific disease diagnosis, knowledge-based systems to assist decision making in triage usually try to cover a much wider domain but can use a smaller set of variables due to time restrictions, many of them subjective so that accurate models are difficult to build. In this paper, we first study criteria that most affect the performance of systems for triage assistance. Such criteria include whether principled approaches from machine learning can be used to increase accuracy and robustness and to represent uncertainty, whether data and model integration can be performed or whether temporal evolution can be modeled to implement retriage or represent medication responses. Following the most important criteria, we explore current systems and identify some missing features that, if added, may yield to more accurate triage systems.

Keywords: Triage; Decision support system; Expert system; Machine learning; Decision tree; Bayesian network

1. Introduction

The triage emergency service is becoming a crucial part of the Emergency Department (ED) in every single hospital as a way to better distribute hospital resources. Triage—a French word meaning "sorting"—is the phase of personal interviews with a health professional at the ED to categorize patients by emergency level so that those in most need of treatment will be assisted first. Some examples of triage protocols that are widely applied are the Australian National Triage Scale (ATS) created in 1993 [1]; the Manchester Triage System (MTS) (since 1997); the Canadian Emergency Department Triage & Acuity Scale (CTAS) influenced by ATS and MTS [2]; the Emergency Severity Index (ESI) used in the United States and based on ATS, MTS and CTAS; the Andorran Model of Triage (MAT) or the Spanish Triage System (SET) based on MAT [3]. All of them use a 5-level scale (see Table 1) for patient categorization and are based mainly on symptoms. At least for CTAS and MAT—used in some European countries—concordance analysis and validity and usability studies have been conducted. Moreover, MAT groups symptoms in symptomatic categories and clinical algorithms and an electronic version (e-PAT) are available for it.

Although medical experts have developed several Clinical Decision Support Systems (CDSSs) for triage assistance by using their experience and knowledge, there is not yet an appropriate and positive evaluation about a significant impact of CDSSs directly built by experts [4]. Either they seem to build too simple models or the model has so many variables to take into account. For instance, MAT (and its electronic version e-PAT) and SET triage protocols use a standard reason-to-visit categorization system (PAT V 3.0) with 576 different reasons to go to the ED and not even

* Corresponding author. Fax: +34 958 243179.
E-mail addresses: mabad@ugr.es (M.M. Abad-Grau), jierache@unimoron.edu.ar (J. Ierache), ccervino@unimoron.edu.ar (C. Cervino), sebas@bu.edu (P. Sebastiani).

1532-0464/$ - see front matter
doi:10.1016/j.jbi.2008.01.007

Table 1
The five levels in hospital triage

(1) Resuscitative	(2) Emergent	(3) Urgent	(4) Less urgent	(5) Non-urgent

a team of experienced doctors has been able to build a valid model upon them [3]. It has to be noted that triage level I is considered of such an emergency that CDSSs for triage assistance cannot be used for time restrictions.

An alternative approach to develop CDSSs for triage assistance consists of utilizing algorithms able to directly learn the model from data. Human experts can use their knowledge to define sufficiently large triage protocols or to indicate the variables likely to affect triage so that information in the data to be used by learning algorithms is sufficient to build valid models. Well established machine learning algorithms are known to have a high expected accuracy and a robust performance in the presence of missing, redundant or inconsistent information as it will be seen in Section 2. Besides the issue of using learning-from data algorithms instead of domain experts to feed the model, we will study in the same section another important criterion that affects the performance of triage systems: different approaches for model building. We will focus on decision trees and Bayesian networks, perhaps the most widely used models that are user-understandable so that models learned from data can be modified by introducing expert knowledge [5]. In Section 3, we will classify current triage systems based on the type of model used and the learning algorithm. The classification will allow us to identify the features they lack that may contribute to a better performance of future systems. Directions for development of more reliable systems to aid triage will be discussed in Section 4.

2. Modeling CDSSs for triage assistance

The most important quality criterion for any CDSS is the accuracy of their inferences. Together with the implementation of triage protocols, some computer-aided systems are already used in the triage emergency service. Most of them, however, are mainly used for helping the triage personnel to ask the relevant questions, but not to make the final decisions [6].

Besides accuracy, other important features a CDSS should include are robustness when some information is missing, redundant or inconsistent; human readability so that medical experts can understand and even modify the underlying model and adaptability when new information is added to the knowledge base. Some particularly noticeable features of CDSSs for triage assistance are:

- the speed to return a decision, i.e. a triage level, and its effect on patient outcome and overcrowding reduction [7,8];
- the specificity to assign a low triage level to true non-urgent cases which translates into cost reduction [7];
- the ability to integrate data from different systems to build more accurate systems [9];
- the ability to model dynamic changes so that retriage and responses to medication can be represented [10].

In this section, we will study two different dimensions in modeling a CDSS that strongly affect all these features: the type of model used (the model approach) [11] and how the model is built using either expert knowledge or data (model building). We will also explore how these two modeling steps are taken into account in the design of CDSSs for triage assistance.

2.1. Dimension 1: choice of the model approach

The simplest CDSS can be modeled by a function called a *classifier*, a function C that for each configuration of the input variables $\mathbf{i} = i_1, i_2, \ldots i_n$ returns a *class* value, i.e., a value of a discrete variable called the *class variable*. When the configuration of input variables describes the symptoms of a patient, the classifier will return a triage level.

In applications such as medical triage it is important that the underlying model in the classifier be interpretable by domain experts. This feature is not required in general classifiers able to predict class values using of variables [12]. This is the case of *neural networks* classifiers, a nature-based approach for building often labeled *black-box* models that are not easily understood by humans. *Support vector machines (SVM)* [13] provide another example of classifiers that are not easily interpretable even by a domain expert. These models define hyperplanes dividing the decision surface that are based on non-linear transformations of the input variables and make the resulting model difficult to interpret. These models have been barely used for triage assistance—see as an example one using a multi-variate logistic classifier [14]. From now on we will focus only on interpretable or white-box models. A model providing this feature can be easily updated by medical experts or be built by a combination of human expertise and a learning machine.

Because of their simplicity, decision trees or the more generic rule-based models are commonly used to describe information that is directly provided by an expert. In fact, most of the CDSSs for triage assistance represent knowledge by means of decision rules [4,6,7,15–19]. A decision tree is a model based on a set of if-then rules linking variables to the class. The if-then rules are created with a recursive procedure that groups data into sets to maximize the overall information. Fig. 1 shows an example of a decision tree to assist in the decision making process about emergency triage of cardiovascular and respiratory diseases [19]. Each leaf node represents a decision, with only two different outcomes: ‘emergent’ (triage level II) or ‘non emergent’ (triage level III or higher). Although expected values, usually probabilities, are not provided in the figure, they must be also computed by the algorithm for each leaf node.

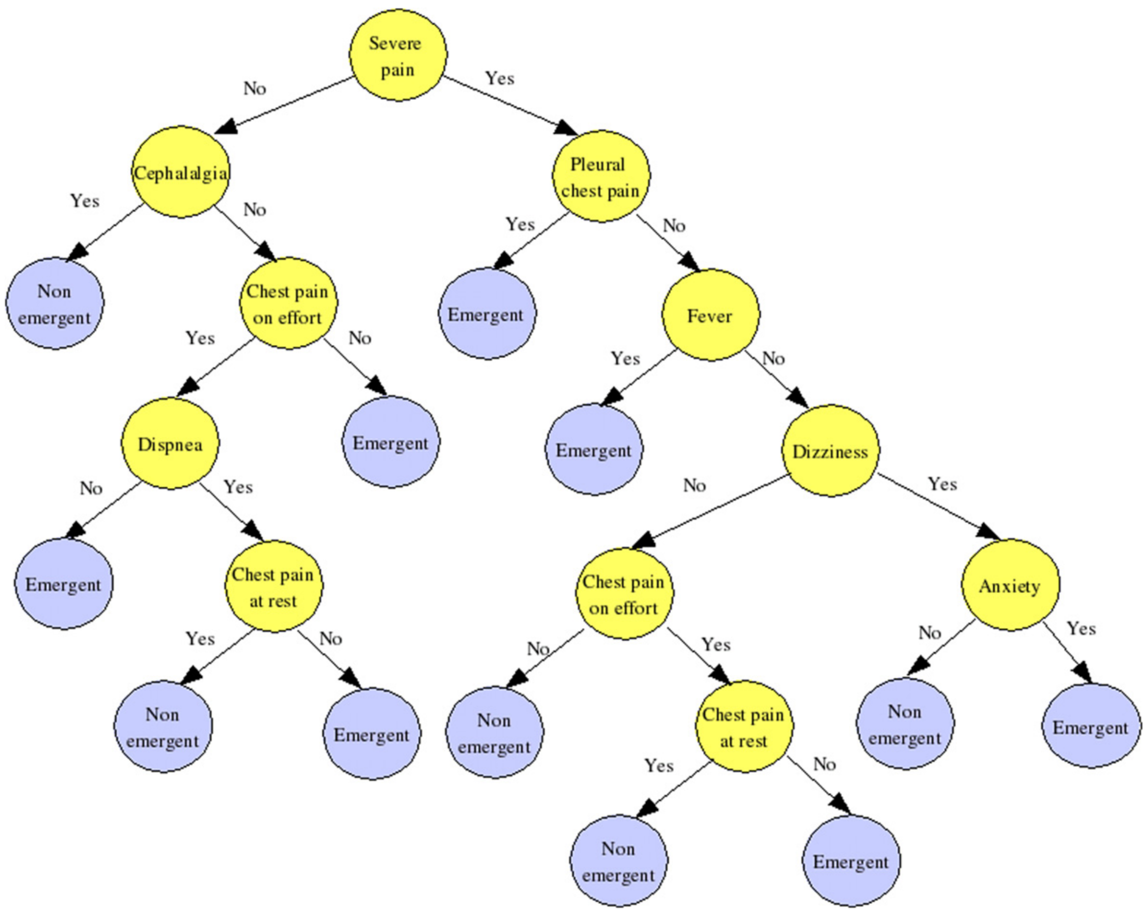

Fig. 1. A decision tree for triage level II learned with the C4.5 algorithm. Nodes in yellow represent input variables while nodes in blue represent the class variable. Source: own processing.

A Bayesian network is a graphical model described by a directed acyclic graph in which nodes represent variables in a knowledge domain and arrows from parents to children nodes represent probabilistic dependences between them. The directed acyclic graph describes properties of conditional independence that determine a factorization of the joint probability of the nodes variables [20]. Although less intuitive than a rule-based system and thus less used in CDSSs built by medical experts, Bayesian networks offer many features and have started to be widely used to model CDSSs in the last few years [21–25].

When a Bayesian network is built to optimize inferences for only one variable—the class—it works as a classifier. Bayesian classifiers are being used for medical diagnosis of Acute Coronary Syndrome [26], risk of death in sickle cell disease [27] and triage assistance [19,28,29]. One of the most popular classifiers based on a Bayesian network is the *naive Bayes classifier*. The Bayesian network used by this classifier is shown in Fig. 2a and represents the assumption that all the attributes are conditionally independent given the class. This assumption of conditional independence is represented by the attributes that are all children nodes of the same parent node that is the variable to be predicted. A slightly more complex model, called *Tree Augmented Naive Bayesian network (TAN)* [30], allows a

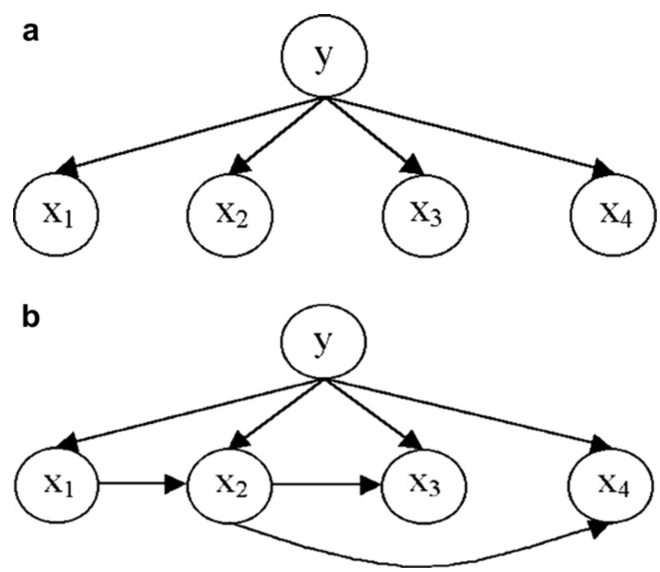

Fig. 2. (a) The structure of the naive Bayes classifier with four input attributes. (b) The structure of a TAN classifier. Source: own processing.

second parent for each input attribute thus including attributes that are marginally independent of the class variable but become conditionally dependent on some other attri-

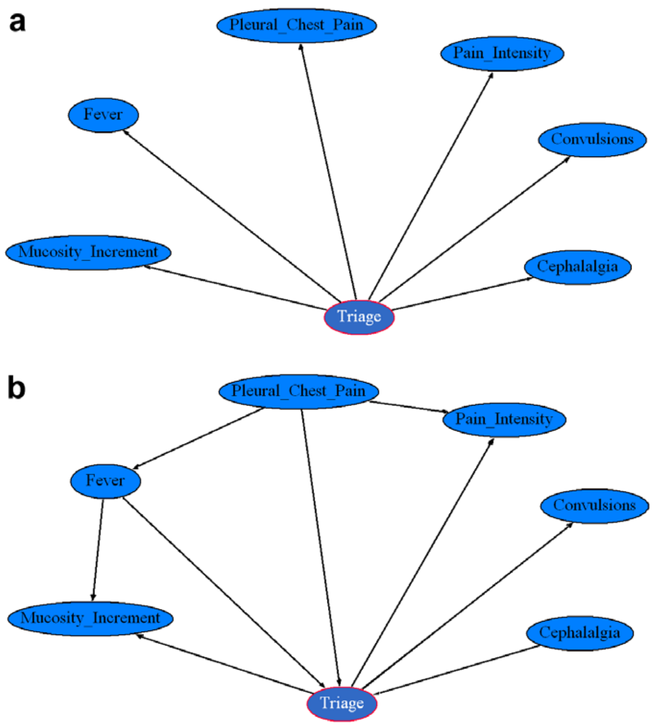

Fig. 3. The structure of the Bayesian network classifier for emergency triage (triage level II) that was learned using (a) the naive Bayes algorithm and (b) the K2 algorithm. Source: own processing.

butes. Fig. 2b shows an example. Fig. 3a shows a naive Bayes classifier for triage assistance while Fig. 3b shows a more complex model, both of them defined over the same set of variables.

Besides interpretability, the model used to build a CDSS influences many other performance features. For both decision trees and Bayesian networks, there exist algorithms able to learn accurate models from data which are also robust in presence of redundant, inconsistent or missing variables. While Bayesian networks model stochastic and therefore uncertain relations between variables, traditional decision trees do not represent uncertain relationships. A substantial effort has been made in the machine learning field to enhance decision tree algorithms so that uncertainty can be treated. More advanced approaches such as *rough set-based decision trees* [31], *belief decision trees* [32] or *credal decision trees* [33] have appeared to better handle uncertainty.

Both decision trees and Bayesian networks can represent temporal relationships between variables. Thus, both of them can be used to model symptom evolution, medication responses or retriage, see as an example the recommendations given by the CTAS for retriage frequencies [2] in Table 2. While symptom evolution and/or medication responses are implicitly used by emergency physicians, they are usually absent in triage protocols or CDSSs for triage and this fact has been reported as one possible major cause of their low accuracies [10]. Time-series-based decision trees have been successfully applied in other

Table 2
CTAS retriage frequencies depending on the triage level

Level I	Level II	Level III	Level IV	Level V
Continuous care	Every 15 min	Every 30 min	Every 60 min	Every 120 min

health-care computer systems, such as in diagnosis of liver cirrhosis in chronic hepatitis patients [34]. Dynamic Bayesian networks have been successfully used in [24] to model influenza surveillance.

The need for data integration and model updating by using information and communication technologies in triage and other emergency services has been widely acknowledged as a way to share data collected from different hospitals and increase accuracy in the models learned from them [35,36]. Incorporation of the expert opinion is straightforward to do in Bayesian networks [20]. Data integration can enhance the definition of more complete domains, including also diseases, so that the system could be used for both triage and diagnosis and even for training. The use of standardized ontologies for information codification is a main issue for data integration to be correctly implemented [35]. This topic will be further examined in the next section. Bayesian networks are naturally more suitable for modeling together triage and diagnosis, as inference can be performed on every single node and they can represent the triage level, a disease, a symptom, a drug, etc. On the contrary, a decision tree implements only one classifier, and integrability with a CDSS for diagnosis would mean to construct a new decision tree for every disease whose diagnosis is to be inferred.

Efficiency in determining a triage level may be improved when not all the input variables need to be assessed. Thus, while patients at resuscitation triage level must be directly recognized by the person performing triage without having to use any computational assistance, the less urgent a patient is the more time can be spent in triage. The *Take The Best (TTB)* heuristic uses a reduced set of variables that is large enough to discriminate between different triage levels so that the reduction in the information to be collected yields to a significant reduction in time. The implementation of this heuristic is straightforward in a decision tree [17]. However, depending on the tree structure, the number of variables needed to discriminate can significantly change. Algorithms to build decision trees usually place variables that better discriminate among the class values in the first levels of the tree while the less informative variables are used further down in the tree branches. Trees with such a structure may force the health professional to ask for more data than necessary if a more parsimonious structure were used. In contrast, Bayesian networks classifiers do not impose any order to introduce the evidence on the user and the TTB heuristic can be optimally implemented in Bayesian networks due to the flexibilty of inference algorithms [29]. The health professionals can start selecting the most acute symptoms they

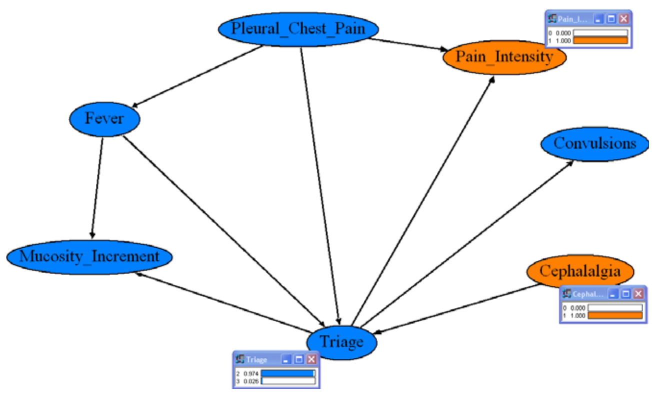

Fig. 4. An example in which triage level II is inferred (posterior probability is 0.975) with the only introduction of high pain intensity (value 1) and presence of cephalalgia (value 1). Source: own processing.

observe and whenever these values are sufficient to discriminate a level (by defining utility functions or loss functions for each level and some rejection strategy), the system will return a triage level. Fig. 4 shows an example where TTB can be applied for the Bayesian classifier in Fig. 3b. By selecting high pain intensity (value 1) and presence of cephalalgia (value 1) there is strong evidence for triage level II (posterior probability is 0.975) and the system could be setup to return the triage level without collecting any other information.

2.2. Dimension 2: model building

An important feature of a CDSS is the type of information used to build the model. This feature affects the accuracy and robustness of a CDSS but also specific quality criteria of CDSSs for triage assistance, such as the speed of patient processing—i.e. the time needed to return a triage level—data integration, specificity and dynamic events modeling. There are mainly two different sources of information to build the model: to directly provide knowledge by an expert or to use information from data.

2.2.1. Domain expert feeding

In this case, the model encodes expert knowledge. A limitation of models built by human experts is that their complexity is not chosen using approaches to guarantee a high performance. These models may be too simplistic to capture true relationships [4] or too complex to be valid given an insufficient amount of data [3]. Even when based on triage protocols, they may be too simple when they use a small and insufficient number of physiological attributes [17]. For instance, ATS uses only 8 physiological attributes and this limitation seems to be the cause of a low accuracy in CDSSs built upon it [17,37]. Contrary to the simplicity of ATS, MAT and SET triage protocols use a standard reason-to-visit categorization system (PAT V 3.0) which establishes 576 different reasons to go to the ED. These reasons are grouped in 32 symptomatic categories and 14 subcategories. This categorization system seems to be too complex and building CDSSs upon it has proven to be a very difficult task [3,19]. Models built by experts may not be sufficiently general, when experts base their knowledge on only few cases and introduce complex rules that barely generalize to new cases. In this case it is said that the rules 'overfit' to the cases used to build the model. Information can be incomplete or redundant and this has to be taken into account whenever the model is to be constructed.

2.2.2. Data driven models

Machine learning, a wide sub-field of Artificial Intelligence, is the study of algorithms and techniques that allow computers to 'learn' from data and extract patterns from data sets. When some information is used by an algorithm to identify similarity patterns, the learning process is referred to as *supervised learning*.

An algorithm to build a classifier C uses a *training sample* for model building. The training sample consists of n instances each one describing a class value and the values of p input variables that may affect the class. The training sample is used by the algorithm to derive classification rules that can be applied to new cases. Fig. 5 shows a general model of automatic supervised classifiers.

Data mining applies machine learning techniques to extract useful information from large data sets or databases. Machine learning and data mining have a wide spectrum of applications including bioinformatics, natural language processing, speech and handwritten recognition and object recognition in computer vision. In medicine, models built using machine learning techniques are used for medical diagnosis, such as detection of acute Coronary Syndrome [26,28] and prognosis, as the Bayesian network to predict stroke risk in sickle cell anemia patients [23].

CDSSs built from data can reach higher levels of accuracy and complexity than those expert-based CDSSs. For example, the accuracy of the triage system based on

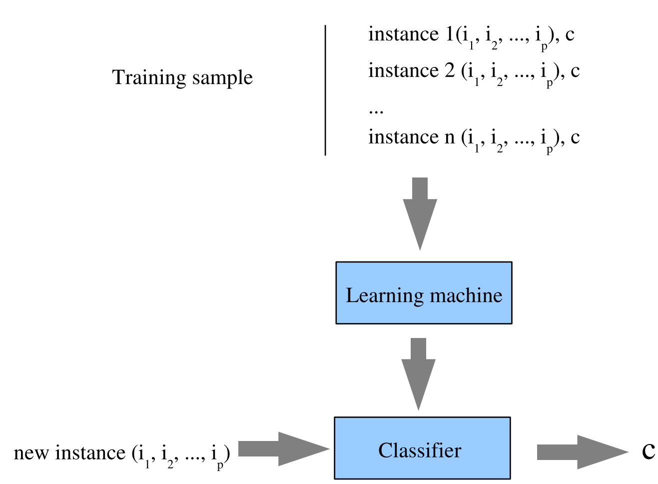

Fig. 5. General model of automatic supervised classifiers. The learning machine or algorithm learns the classifier from a training set, i.e., without human intermediation. The classifier returns a class value c for each new configuration $\mathbf{i} = i_1, i_2, \ldots i_n$ of the input variables. Source: own processing.

multi-variate logistic regression described in [14] was higher than the accuracy of triage made by an emergency specialist. Particularly, the sensitivity—the percentage of emergency patients correctly identified by the CDSS—was higher compared with the expert decision (92.4% versus 85.7%) and the specificity—the percentage of non-emergency patients correctly identified—was also higher in the CDSS (90% versus 40%).

Machine learning offers principled approaches for developing algorithms to learn complex models from high dimensional samples, possibly with redundant, inconsistent or missing data. Examples are algorithms based on the Statistical Learning Theory [13], the Computational Learning Theory or the Model of Probably Approximately Correct Learning by Valiant [38], the Minimum Description Length principle [39] and the Bayesian Inference [40]. One issue addressed by all of them is the bias/variance trade off, i.e., the smaller the bias of an algorithm the greater the chance of building a classifier with high accuracy in the training set. This high accuracy however can be the result of overfitting the data so that the performance of the classifier can be much worse in new data. Therefore, the use of sound algorithms based on one of the aforementioned principles should protect against overfitting. With the Conservation Law for Generalization Performance [11] it was better understood that the nature of the problem itself imposes a restriction in the performance of an applied method. How to choose a method whose hypothesis space is large enough to contain a solution to the problem and yet small enough to ensure reliable generalization from reasonable-size training sets is one of the most important issues in machine learning [41].

Simple classifiers based on decision trees or Bayesian networks can be fitted using relatively small data sets. As an example, the decision tree shown in Fig. 1 was built from a small simulated data set of only 124 cases and 40 input variables chosen by medical experts [19]. The algorithm used to learn the tree was C4.5 [42], still a benchmark in generalization capacity, which uses Minimum Description Length to trade off between samples size and model complexity. As the number of instances was small, the decision tree was 'pruned' keeping only nine variables to avoid overfitting.

From the same simulated data set, two Bayesian networks were also built, one is a Naive Bayes classifier and the other a more complex network learned by the *K2* algorithm [43], which was introduced to build Bayesian networks (see Fig. 3). From the 40 variables affecting triage in the data set, the final classifier learned by the K2 algorithm only chose six of them. Moreover, the number of bivariate dependencies among them was reduced from 21 to only 3.

3. Bidimensional review of CDSSs for triage assistance

In this section, we will use the two dimensions affecting the quality of a CDSS mentioned in the last section as classification criteria of CDSSs for triage assistance, following an historical order.

3.1. Domain expert/decision rules

This group consists of those CDSSs built by medical experts using a rule-based or decision tree model. Reported accuracies of systems in this group are usually low. Most of the evaluated and currently in use CDSSs are rule-based models built by a medical expert, an expert team or some triage protocol and their average accuracy is usually under 60% [7,15,6]. The CDSSs with reported larger accuracy do not outperform results achieved by triage nurses. For example, a computerized rule-based decision support algorithm to assist nursing triage of potential acute bronchopulmonary events in lung transplant recipients achieved accuracies greater than 90% but the same accuracy was observed in the triage conducted by nurses [44]. In fact, the high accuracies seem to be common only in those triage systems that somehow reduce the problem complexity, usually by strongly limiting the nature of the patients health complaints.

Some examples of such a system are the 'Automated Triage Management' (ATM) [7] developed by emergency doctors of the School of Medicine at UCLA; the 'Symptoms, Advice, Measure' (SAM) [15] developed by a family physician, and the 'Ped's Advice' (PA) [6], developed by medical doctors and nurses at the Academic Children Hospital in Uppsala (Sweden) to assist triage when the patients are children. In both SAM and PA the user, typically a nurse, introduces a word about the most noticeable symptom, i.e., cough, into the system that provides a yes/no questionnaire to be completed step-by-step by the user. At the end, the system recommends a triage level. Both systems are currently in use and have been criticized by their users for several limitations: (1) they are not used as decision support systems but as a memory helper, (2) they are not fully adapted to current clinical practice as they focus mainly on acute conditions and (3) they do not consider ethical and psychological knowledge, something that is worth in practical daily situations [6].

An enhanced subset of expert-based CDSSs for triage consists of those systems that suggest a diagnosis as well as a triage level. They usually provide an ordered list of possible diagnoses with their estimated probabilities. Two examples of these systems currently at exploitation phase are 'Quick Medical Reference' (QMR) and 'Iliad'—that use Bayesian algorithms to estimate the probabilities of the different diagnoses. Their validity in EDs has been studied by Graber and VanScoy [10]. Although Iliad performs better than QMR, mean accuracies in diagnostic prediction were too low to be reliable. As an example, Iliad provided the true diagnosis in its list in only 54% of the instances while QMR did in 51% of the cases. Moreover, only in 36% of the cases Iliad referred the true diagnosis among the five most probable and QMR only in 32% of the examples. One of the possible reasons for this low accuracy is the limited information: they do not take into account medications that patients are taking, the duration of symptoms, or the sequences in which symptoms appeared. Another drawback also in common with other systems cited above, such as SAM and PA, is that the list of symptoms is not comprehensive.

'eTRIAGE' is another example of CDSS for triage that was implemented in all EDs in the Edmonton Capital Health region in 2003 [45]. eTRIAGE uses rules from the CTAS protocol as the knowledge base. For research purpose, it also allows modification of the triage level based on clinical judgment of the user, typically a registered nurse. Comparison with a review panel showed 64.9% agreement [46].

'iTRIAGE' is a system still in validation phase that is based on the ATS [17,37]. This system can be used from a personal digital assistant (PDA) and synchronized with a server to update new incomes and decisions. The rules use fuzzy scores to convert linguistic terms such as 'moderate pain' to a numerical scale and are defined for each one of the eight different physiological attributes in ATS and a multi-criteria heuristic algorithm returns a decision. The system produces robust decisions for urgent scenarios and helps to reduce ambiguity in non-urgent ones. iTRIAGE was compared with the paper-based system by selecting 29 nursing students in their final year to use iTRIAGE and 21 to use the paper-based system. Although the 67% accuracy of iTRIAGE is an improvement compared to the other systems, it is still low and can be improved by increasing the information about patient condition that are currently limited to airway, breathing, circulation, conscious state, pain, neurovascular status, mental health emergencies and ophthalmic emergencies. The low accuracy even when using a CDSS implementing some heuristic strategies and rules allowing uncertainty seems to be due to an insufficient amount of information. The TTB heuristic is straightforward implemented in iTRIAGE [17].

An *expert system* is an advanced DSS with a knowledge base where both data and data model—a decision tree or a Bayesian network—are located separately from the inference engine or algorithm used to assist decision making. Modern expert systems are capable of expanding the knowledge base when new information is introduced. eTRIAGE and iTRIAGE are enhanced CDSSs, as they both separate code from knowledge. However, they cannot be considered modern expert systems as they both lack a *knowledge acquisition component* for the model to be updated by the domain expert [47,48,5], so they are static. OSGi, a software platform based on Java technology, was used to build a domain expert rule-based dynamic system for emergency triage [18]; quality measures were not reported though.

3.2. Domain expert/Bayesian network

Some expert systems use Bayesian networks to represent the knowledge base and probability theory—Bayes theorem in particular—for reasoning. To build Bayesian classifiers—the simplest Bayesian network for classification—medical experts have to elicit the probability distribution of the attributes (the symptoms) given the class (triage level or disease).

This approach was used to build a medical expert system focusing only on patients complaining of a non-traumatic abdominal pain [29]. The system used a Bayesian network as the knowledge base that was built by physicians who were previously trained in knowledge engineering. Although triage does not need to build a diagnosis list for a patient, this network included possible medical pathologies used to infer the triage value conditional on the possible causes of the chief complaint. Compared with results obtained when the triage was made by an emergency specialist, it reached a higher sensitivity—percentage of truly emergency patient correctly classified—(90% versus 64%) but a lower specificity—percentage of classified as non emergency patients that truly were non emergent—(25% versus 48%).

'PROSTANET', a system for diagnosing prostate cancer, is another example [49] of expert system that uses a Bayesian network. The system was built on top of a software tool for Bayesian network modeling and inference that lets the users change the structure and/or the parameters every time they want [50]. The user should be a domain expert with training in knowledge engineering as well. PROSTANET has been updated several times by using the expert knowledge.

3.3. Machine learning/decision rules

These systems are based on decision-rule models that are built from data. Fig. 1 shows an example of a system under this category developed for research purpose from a small simulated data set of only 124 cases and 40 input variables. The system reported an accuracy of 79.7% [19] using 5-fold cross-validation.

A rule-based system still at evaluation stage, 'Mobile Emergency Triage' (MET) [51,16] and some specific applications such as 'MET-Abdominal Pain' (MET-AP) [52] try to improve conventional triage by creating a CDSS whose rules go beyond the knowledge of a triage nurse. This is accomplished by incorporating knowledge from emergency physicians as well as knowledge that results from the combination of key historical information with physical findings [4]. Thus, the system is fed with both information from data sets and expert knowledge. Accuracies reported for MET were a little bit lower than those obtained by physicians (65.4% versus 70.02%) and specificity for critical patients was comparable with the one achieved by physicians [51]. The accuracy of the more specific triage system for children with abdominal pain [52] was 66% using 5-fold cross-validation, still too low to be considered accurate. The low accuracy can be attributed to an insufficient knowledge base: the system is built with only eight variables. Another feature of MET is the use of *ontologies* to

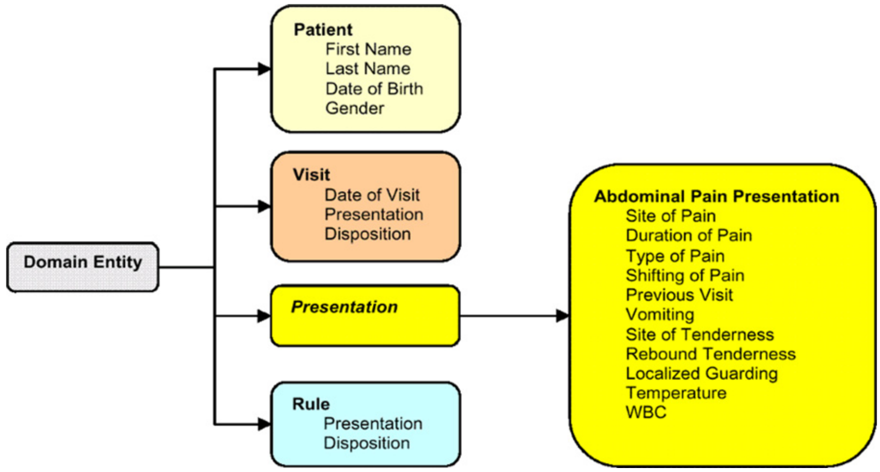

Fig. 6. The ontology domain used by the MET system for triage of abdominal pain presentation. Source: [16].

handle data integration from different and heterogeneous sources. An ontology is "the specification of a conceptualization" [53] so that concepts and their relationships to other concepts are specified precisely to support machine interpretation. In ontologies used for triage, the triage level could be inferred using concepts in different data sources, such as one with symptoms and another with information about drugs. In more advanced systems including diagnosis, molecular information will be used instead of overall symptoms [9] and the clinical patient record will require integration of clinical and genomic data [54] so that ontologies will become a fundamental tool. Fig. 6 shows the simple domain ontology used in MET for triage of abdominal pain presentation. Another important feature of MET is the use of rough-set-based decision trees [31] instead of conventional decision trees in order to represent uncertainty in the underlying rules. Therefore, the system can be more robust in presence of redundant, inconsistent or missing attributes [31].

Several algorithms for incremental updating of a decision tree have been developed [55]. However, most of the current clinical systems under this category do not take account of them.

3.4. Machine learning/Bayesian networks

Two systems under this category were developed [19], one using the Naive Bayes classifier and the other using the K2 algorithm implemented in the program Bayesware Discoverer [56]. From the generic Bayesian network learned with the K2 algorithm, we can select a subset of the variables that are sufficient for inferences on the triage variable. This set of variables consists of the *Markov blanket* of the variable to be predicted and makes it independent of all the other variables in the network [57]. Fig. 3b shows this Bayesian network. Accuracies, computed by using 5-fold cross-validation [58], were 87.9% and 86.9%, respectively, both significantly higher than the accuracies of a decision tree learned with the C4.5 algorithm (79.7%). This result is promising about the use of Bayesian networks for triage assistance and suggests that the high amount of uncertainty in the process of triage may be better handled by models that use probability theory to account for uncertainty. An example of a CDSS under this bidimensional category not used for triage at the ED is a Bayesian classifier recently developed to predict the risk of death in sickle cell anemia with accuracy above 90% [27].

Updating the parameters of a Bayesian network from a data set is straightforward. However, updating the network structure is still an open problem [59–61]. To the best of our knowledge, current clinical systems under this category are not provided with a knowledge acquisition component for either parameter or model updating.

4. Conclusion and future trends

Most of the current systems to assist triage in EDs built by domain experts have shown little accuracy, if ever validated. The low accuracy was not improved by imposing limitations in the domain either by reducing the type of different complains causing the patient to go to the ED or by limiting the type of diagnostic on the basis of those complains. As one of the problems seems to be the difficulty for medical experts to build valid models whenever too many variables affect the triage process, the use of principled approaches from the Machine Learning field allowing the direct construction of accurate models from data sets appears to be a more promising alternative. Bayesian networks constitute a worthy competitor of the more traditional rule-based systems as they can be interpretable, they can be modified by experts and they do not impose any order in the current evidence reported by a patient to infer the triage level. There are many available procedures that can be used to enhance CDSS for triage. For instance, dynamic models can lead to implement retriage or medication response. Moreover, ontologies for data integration have started to be used in triage systems and other technologies are ready to be applied. Thus, heterogeneous data sources such as medication and drug taken, familial risk or clinical records could feed a triage system and could be used to link a triage system with other systems such as CDSSs for disease diagnosis. Telecommunication technologies, already in use by some triage systems [16,17,37], will improve the collection of information and provide better information for machine learning modeling.

There is still a big gap between the systems currently in use and the advanced approaches described in this paper. Predicting the triage level will be still a very challenging task in the short-term. Perhaps the task is one of the most difficult ones in Artificial Intelligence, requiring a very large amount of information and complex knowledge but little time to collect the information and infer knowledge.

We can conjecture, however, that these technologies together with graphical informatics for simulation will also be used to handle genomic data and integrate triage as part of complex models relating diseases with molecular defects and with an emergency level. From molecular to organic levels, from conception to elderly, computer simulations and biologic and medical researchers will feed-forward each other in order to understand all the mechanisms occurring in the human body. Thus, not only triage but also diagnostic models will be part of far complex multivariate and multi-modal systems using symptoms, diseases, medication, vital signs, tests results, medical records, genomic data [62], environmental information, etc.

Acknowledgments

The authors have been partially supported by the Spanish Research Program under projects TIN2004-06204-C03-02, TIN2005-02516 and TIN2007-67418-C03-03, by the European Regional Development Fund (ERDF) and by the NIH grant NHGRI R01 HG003354-01A2. The authors are grateful to the reviewers, for their comments and suggestions.

References

[1] Victorian Department of Human Services. Consistency of Triage in Victoria's Emergency Departments: Summary Report. Melbourne, Australia: Victoria's Hospital; 2001 [TD-1106].

[2] Murray M, Bullard M, Grafstein E. Revisions to the Canadian emergency department triage and acuity scale implementation. Can J Emerg Med 2004;6(6):421–7.

[3] Gómez-Jiménez J, Segarra X, Prat J, Fernando JB, Albert E, Borras M. Concordancia, validez y utilidad del programa informtico de ayuda al triaje (PAT) del Modelo Andorrano de Triaje (MAT) (In Spanish). Emergencias 2003;1(5):339–44.

[4] Farion KJ, Michalowski W, Rubin S, Wilk S, Correll R, Gabory I. Prospective evaluation of the MET-AP system providing triage plans for acute pediatric abdominal pain. Int J Med Inf 2008;77(3):208–18.

[5] Cowell RC, Dawid AP, Lauritzen SL, Spiegelhalter DJ, editors. Probabilistic networks and expert systems. New York: Springer; 1999.

[6] Holmstrom I. Decision aid software programs in telenursing: not used as intended? Experiences of Swedish telenurses. Nurs Health Sci 2007;9:23–8.

[7] Guterman JJ, Mankovich NJ, Hiller J. Assessing the effectiveness of a computer-based decision support system for emergency department triage. In: Engineering in Medicine and Biology Society, 2003. Proceedings of the 15th Annual International Conference of the IEEE, 2003. p. 592–93.

[8] Hoot N, Zhou C, Jones I, Aronsky D. Predicting emergency department overcrowding: a prospective validation study. In: Supplement of the Academic Emergency Medicine, vol. 14; 2007, p. 53.

[9] Gerling IC, Solomon SS, Bryier-Ash M. Genomes, transcriptomes, proteomes. Arch Intern Med 2003;163(2):190–8.

[10] Graber MA, VanScoy D. How well does decision support software perform in the emergency department? Emerg Med J 2003;20:426–8.

[11] Schaffer C. A conservation law for generalization performance. In: Proceedings of the eleventh international conference on machine learning. Morgan Kaufman; 1994. p. 259–65.

[12] Sajda P. Machine learning for detection and diagnosis of disease. Annu Rev Biomed Eng 2006;8(8):1–29.

[13] Vapnik V. Statistical learning theory. New York: J. Wiley; 1998.

[14] Gellerstedt M, Bang A, Herlitz J. Could a computer-based system including a prevalence function support emergency medical systems and improve the allocation of life support level. Eur J Emerg Med 2006;13(5):290–4.

[15] Marklund B. Symptoms, råd, åtgärd. (Symptoms, advice, measure.), Software program. Vänersborg: Vårdutvecling AB, 2000.

[16] Michalowski W, Wilk S, Farion K, et al JP. Development of a decision algorithm to support emergency triage of scrotal pain and its implementation in the MET system. In: Information systems and operational research, vol. 43; 2005. p. 287–301.

[17] San Pedro J, Burstein F, Wassertheil J, Arora N, Churilov L, Zaslavsky A. On development and evaluation of prototype mobile decision support for hospital triage. In: Proceedings of the 38th Annual Hawaii International Conference on System Sciences (HICSS'05)—Track 6, 2005. p. 157c.

[18] Sheng YH, Chang P. A DSS with dynamically pluggable rules take emergency triage as example. In: AMIA symposium proceedings, 2006. p. 1091.

[19] Abad-Grau MM, Ierache J, Cervino C. Aplicación de Redes Bayesianas en el Modelado de un Sistema Experto de Triaje en Servicios de Urgencias Médicas. In: IX Workshop de Investigadores en Ciencias de la Computación, Trelew, Argentina, vol. 1; 2007. p. 43–47 [in Spanish].

[20] Jensen FV, Huirne RBM, et al., editors. Bayesian networks and influence diagrams Risk Management Strategies in Agriculture. Wageningen, Holland: Mansholt Studies; 1997.

[21] Wong WK, Moore A, Cooper G, Wagner M. Bayesian network anomaly pattern detection for disease outbreaks. In: Fawcett T, Mishra N, editors. Proceedings of the twentieth international conference on machine learning. Menlo Park, California: AAAI Press; 2003. p. 808–15.

[22] Burnside ES, Rubin DL, Shachter RD. Improving a Bayesian network's ability to predict the probability of malignancy of microcalcifications on mammography. In: International congress series: CARS 2004—computer assisted radiology and surgery. Proceedings of the 18th international congress and exhibition, vol. 1268; 2004. p. 1021–6.

[23] Sebastiani P, Ramoni MF, Nolan V, Baldwin CT, Steinberg MH. Genetic dissection and prognostic modeling of overt stroke in sickle cell anemia. Nat Genet 2005;37:435–40.

[24] Sebastiani P, Mandl KD, Szolovits P, Kohane I, Ramoni MA. Bayesian dynamic model for influenza surveillance. Stat Med 2005;25:1803–16. [157].

[25] Cruz-Ramírez N, Acosta-Mesa HG, Barrientos-Martínez RE, Nava-Fernández LA. Diagnosis of chronic idiopathic inflammatory bowel disease using Bayesian networks. In: Trinidad JFM, Carrasco-Ochoa JA, Kittler J, editors. CIARP. Lecture Notes in Computer Science, vol. 4225. Springer; 2006. p. 706–15.

[26] A AN, Fesmire FM, Sonnemaker D. Mining Bayesian networks to forecast adverse outcomes related to acute coronary syndrome. In: Proceedings of the 17th international FLAIRS conference, 2004. p. 1–6.

[27] Sebastiani P, Nolan V, Baldwin C, Abad-Grau MM, Wang L, Adewoye A, et al. A network model to predict the risk of death in sickle cell disease. Blood 2007;6:1–17.
[28] Fesmire FM, Novobilski A. First step in the Erlanger artificial intelligence initiative: development of a Bayesian network utilizing initial triage history to risk stratify chest pain patients for thirty-day adverse outcome. Ann Emerg Med 2003;42:857–88.
[29] Sadegui S, Barzi A, Sadeghi N, King B. A Bayesian model for triage decision support. Int J Med Inform 2006;75(5):403–11.
[30] Friedman N, Geiger D, Goldszmidt M. Bayesian network classifiers. Mach Learn 1997;29:131–63.
[31] Jinmao W, Dao H, Shugin W, Zhuyand M. Rough set based decision tree. In: Proceedings of the 4th world congress on intelligent control and automation, 2002. p. 426–31.
[32] Elouedi Z, Mellouli K, Smets P. Belief decision trees: theoretical foundations. Int J Approximated Reasoning 2001;28:91–124.
[33] Abellán J, Moral S. Building classification trees using the total uncertainty criterion. Int J Intell Syst 2003;18(12):1215–25.
[34] Yamada Y, Suzuki E, Yokoi H. Decision-tree induction from time-series data based on a standard-example split test. In: Proceedings of the twentieth international conference on machine learning (ICML 2003), Washington, DC; 2003. p. 840–7.
[35] Bartell EN, Coonan K, Finnel J, Pollock D, Cochrane D. Disparate systems, disparate data: integration, interfaces and standards in emergency medicine information technology. Acad Emerg Med 2004;11:1142–8.
[36] Wann TT, Lin BY, Ma A. Integration mechanisms and hospital efficiency in integrated health care and delivery systems. J Med Syst 2002;26:127–43.
[37] Padmanabhan N, Burstein F, Churilov L, Wassertheil J, Hornblower N, Parker N. A mobile emergency triage decision support system evaluation. In: Proceedings of the 39th hawaii international conference on system sciences, vol. 1; 2006. p. 3–4.
[38] Shavlik JW, Dietterich TG, editors. Readings in machine learning. San Mateo, CA: Morgan Kaufman; 1990.
[39] Rissanen J. Modeling by shortest data description. Automatica 1978;14:222–8.
[40] Schwartz G. Estimating the dimension of a model. Ann Stat 1978;6:461–4.
[41] Baxter J. A model of inductive bias learning. J Artif Intell Res 2000;12:149–98.
[42] Quinlan JR. Improved use of continuous attributes in C4.5. J Artif Intell Res 1996;4:77–90.
[43] Cooper GF, Herskovits E. A Bayesian method for construction Bayesian belief networks from databases. In: Proceedings of the eighth conference on uncertainty in artificial intelligence, 1991. p. 86–94.
[44] Finkeistein SM, Scudiero A, Lindgren B, Snyder M, Hertz MI. Decision support for the triage of lung transplant recipients on the basis of home-monitoring spirometry and symptom reporting. Heart Lung 2005;34(30):201–8.
[45] Dong SL, Bullard MJ, Meurer DP, Blitz S, Akhmetshin E, Ohinmaa A, et al. Predictive validity of a computerized emergency triage tool. Acad Emerg Med 2007;14(1):16–21.
[46] Jelinek GA, Little M. Inter-rater reliability of the National Triage Scale over 11,500 simulated occasions of triage. Emerg Med 1996;8:226–30.
[47] Neapolitan RE, editor. Probabilistic reasoning in expert systems. New York: John Wiley and Sons; 1989.
[48] Castillo E, Gutiérrez JM, Hadi AS, editors. Expert systems and probabilistic networks models. Springer Verlag; 1997.
[49] Lacave C, Díez FJ. Knowledge acquisition in PROSTANET—a Bayesian network for diagnosing prostate cancer. In: Knowledge-based intelligent information and engineering systems (LNCS), vol. 2774; 2003. p. 1345–50.
[50] Consorium TE. Elvira: an environment probabilistic graphical model. In: Proceedings of the first international workshop on probabilistic graphical models, Cuenca, Spain; 2002. p. 220–30.
[51] Michalowski W, Słowiński R, Wilk S. MET system: a new approach to m-health in emergency triage. J Inf Technol Healthcare 2004;2(4):237–49.
[52] Wilk S, Sjowinski R, Michalowski W, Greco S. Supporting triage of children with abdominal pain in the emergency room. Eur J Oper Res 2005;160:696–709.
[53] Gruber TR. A translation approach to portable ontology specifications. Knowledge Acquisition 1993;5(2):199–200.
[54] Louie B, Mork P, Martín-Sánchez F, Halevy A, Tarczy-Hornoch P. Data integration and genomic medicine. J Biomed Inform 2007;40:5–16.
[55] Utgoff P. Incremental induction of decision trees. Mach Learn 1989;4:161–89.
[56] Sebastiani P, Ramoni M, Crea A. Profiling your customer using Bayesian networks. SIGKDD Explorations 2000;1(2):91–6.
[57] Sebastiani P, Abad-Grau MM, Ramoni MF. Bayesian networks for genomic analysis. In: Dougherty ER, Schmulevich LC, Chen J, Wang ZJ, editors. Genomic signal processing and statistics. Eurasip; 2005. p. 281–320.
[58] Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: Proceedings of the 15th international joint conference on artificial intelligence, 1995. p. 114–9.
[59] Friedman N, Goldszmidt M. Sequential update of Bayesian network structure. In: Proceedings of thirteenth conference on uncertainty in artificial intelligence (UAI 97), 1997. p. 1–10.
[60] Huang H, Song H, Tian F, Lu Y, Wang Q. A comparatively research in incremental learning of Bayesian networks. In: Proceedings of the fifth world congress on intelligent control and automation, vol. 1. Hangzhou, China; 2004. p. 4260–5.
[61] Pan R, Peng Y, Ding Z. Belief update in Bayesian networks using uncertain evidence. In: 18th IEEE international conference on tools with artificial intelligence (ICTAI'06), vol. 1; 2006. p. 441–4.
[62] Kanehisa M, Bork P. Bioinformatics in the post-sequence era. Nat Genet 2003;33:305–9.