A General Outpatient Triage System Based on Dynamic Uncertain Causality Graph

XUSONG BU¹, LIN LU², ZHAN ZHANG³, QIN ZHANG¹,⁴, AND YAN ZHU⁵

¹School of Computer Science and Engineering, Beihang University, Beijing 100191, China
²Department of Endocrinology, Peking Union Medical College Hospital, Beijing 100006, China
³Institute of Nuclear and New Energy Technology, Tsinghua University, Beijing 100084, China
⁴Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China
⁵School of Economics and Management, Tsinghua University, Beijing 100084, China

Corresponding authors: Xusong Bu (buxusong@buaa.edu.cn) and Qin Zhang (qinzhang@tsinghua.edu.cn)

ABSTRACT In order to reduce the misdiagnosis caused by outpatient triage error and help triage nurses improve their triage accuracy, a general outpatient triage system based on the hybrid Dynamic Uncertain Causality Graph (hybrid DUCG) is presented. The hybrid DUCG is a combination of the single-valued DUCG (S-DUCG) and the multivalued DUCG (M-DUCG). In this paper, the M-DUCG is utilized to construct the basic outpatient triage knowledge base. In the knowledge base, each department is modeled as one sub-M-DUCG. It describes the causal relations between the department and its associated clinical features. All the sub-M-DUCGs are compiled into one M-DUCG as the basic outpatient triage knowledge base. The S-DUCG is employed to adjust the basic knowledge base to match the variant triage requirements for various hospitals. To validate the triage system, we constructed a basic triage knowledge base and adjusted it to match the triage requirements of a county-level hospital. Verification experiments were performed based on the adjusted triage knowledge base. The result showed that the triage system has high triage accuracy, and it can effectively assist triage nurses in outpatient triage.

INDEX TERMS Probabilistic reasoning, knowledge representation, outpatient triage, causality.

I. INTRODUCTION

Triage systems can be classified into three types, the specific disease triage [1]–[3], the emergency triage [4], [5], and the outpatient triage [6], [7]. The purpose of the first two is to arrange the treatment priority for patients according to their fundamental physiological indicators and the urgency of their health. Different from the first two, the purpose of outpatient triage is to triage the patient to the proper department according to his clinical features. Research on triage systems has focused primarily on emergency triage and specific disease triage, and numerous computer-aided systems developed to help triage nurses to achieve that mission [8]–[14] but ignored the importance of the outpatient triage. Outpatient triage is the first step of clinical diagnosis, many misdiagnoses were caused by outpatient triage errors, most of which were due to a lack of experience and medical knowledge among triage nurses and to insufficient information provided by patients [15], [16]. Some common symptoms, such as fever, headache, and cough, can be caused by various diseases, and those diseases belong to different departments. If there are not enough clinical features of patients, outpatient triage errors could easily be made by an inexperienced nurse. For a patient, the direct consequence of an outpatient triage error is a waste of time, a bad medical experience, and property damage. In severe cases, it can cause misdiagnosis, which can lead the patient to miss the best treatment time, with irreversible consequences. Therefore, it is of great practical significance to study and develop an outpatient triage system to assist triage nurses in improving the outpatient triage accuracy and reducing misdiagnosis or missed diagnosis caused by outpatient triage errors.

The essence of outpatient triage is to classify patients into the proper department according to their clinical features. Various algorithms could be used for this classification, such as the K-NN [17], [18], Naive Bayes, Bayesian Network [19], [20], SVM [21]–[23], decision tree [24], [25], artificial neural network, and deep learning [26]–[29]. These classification algorithms have been fully verified and widely applied in the field of text classification and image...
recognition [30]–[34]. Unlike other fields, the algorithms used in clinical diagnosis not only require a high diagnosis accuracy but also good interpretability for its result. However, the algorithms cannot properly explain the calculation results, and this prevents them from performing well in clinical diagnosis. Therefore, an algorithm with good interpretability and extreme accuracy is needed. As a graphical causal probability model [35]–[40], the DUCG can intuitively express causal knowledge in various graphic symbols, present the result in the form of conditional probability and illustrate it graphically. With initial application in the fault diagnosis of complex large scale systems [41]–[45], the DUCG was later applied to clinical diagnoses, such as jaundice, vertigo, and nasal obstruction, and perfect results were achieved [46], [47]. Due to the good performance of DUCG in clinical diagnosis, this paper uses the DUCG to perform outpatient triage.

Section II is a brief description of the DUCG model, including M-DUCG and S-DUCG. Section III explains the construction method of the triage knowledge base. Section IV shows the calculation method and the calculation case of the hybrid DUCG. Section V describes a method of heuristic information collection and verifies the triage system. Section VI is a conclusion, and it outlines future work.

II. BRIEF INTRODUCTION TO THE DUCG

In the DUCG, different types of variables/events are represented by different shapes [41]. As shown in Figure 1, Bi (i) is the root cause event, such as a fault source in the fault diagnosis or a disease in the clinical diagnosis. The B-type variable has no parent variable but has at least one child variable. During the inference calculation, the B-type variable is used as a hypothesis. Xi (i) stands for a consequence variable or a procedure variable, for example, clinical features in a clinical diagnosis or abnormal signals in a fault diagnosis. It can be caused by any type of variables and can also be used as the cause of other variables. Di (i) represents the default or unspecified cause event of Xi. D works only when none of the known causes of Xi has occurred. BXi (i) represents an integrated cause variable. In a fault diagnosis, BXi is used as a collection of multiple root cause events [48], while in clinical diagnosis, BXi represents the root cause event affected by risk factors [45], [46]. During the inference calculation, if BXi exists, it replaces Bi as the hypothesis. Gi (i) is the logic gate variable representing the logic relations among its parent variables. The logical combinations of its parent variables are described by the logic gate specification table (LGSi). For example, the LGS1 of G1 is G1,0 = “Remnant”, G1,1 = X3,X6,0, and G1,2 = X3,X6,1. SGi (i) is also a logic gate, but it is only used in clinical diagnosis to express combinations of risk factors. In the DUCG, uncertain causality relationships are represented by different directed arcs. The uncertain causality between variables in S-DUCG is represented by linkage event Pn,i (i). Before “;” is the child variable, and n is the index of the child variable; after “;” is the parent variable and i is the index of the parent variable. Its causality intensities are coded in matrix pn,i, e.g., P3,1 is the linkage event from B1 to X3, with the parameter matrix as follows:

\[ p_{3,1} = \text{Pr}\{P_{3,1}\} = \begin{pmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{pmatrix} \]

Different from P3,1, P6,1 is a conditional linkage event( ). Its condition is denoted as Z6,1 = X8,1, which means that if X8,1 is true, P6,1 exists; otherwise, P6,1 does not exist. Weighted functional event Fn,i (i) represents the uncertain causality between variables in M-DUCG, e.g., F7,2: Fn,i will be explained in detail later. Zoom factor variable SAi,n (i) is specifically used between BXn and SG to represent the zoom effect of risk factors on BXn. Unlike Fn,i, the parameters of SAi,n can be greater than 1, and a dotted double directed arc ( ) is used to represent the conditional zoom factor variable. The black directed arc ( ) links the input variables of the G-type or SG-type variable.

S-DUCG is short for DUCG that is applied to single-valued variables [36]. In S-DUCG, all variables have two states, only the cause of the variable’s true state (1 state) needs to be defined, and the false state (0 state) regarded as the complement of the true state. Figure 2(a) is an example of S-DUCG, it shows that Xi and Xj have a certain probability to cause Xn. Figure 2(b) reveals the functional mechanism of Figure 2(a), it shows that the relationship between Xi and Xj is OR. Pn,i and Pn,j are two linkage events, and they show the causal intensity that Xi and Xj cause Xn to occur separately. Therefore, Xn can be expressed as Xn = Pn,i X i ∪ Pn,j X j, and this process called event expansion. Before probability
A basic triage knowledge base is used to describe the department division of the hospital and to model the causal relationships between departments and clinical features. In reality, the granularity of the department division for each hospital is different; the department division of large-scale general hospitals is detailed, and by contrast, the division of county-level hospitals or community hospitals is rough. To make the triage knowledge base have good scalability and match the different triage needs of different hospitals, we construct the triage knowledge base in the following way:

1) Construct a basic triage knowledge base. Modeling each department as a sub-DUCG according to the most detailed department division schemes, and all those sub-DUCGs are compiled into a complete DUCG as a basic triage knowledge base.

2) Configure the basic knowledge base to meet the triage needs of a target hospital. The purpose of the configuration is to match the target hospital departments with the departments of the basic triage knowledge base, and this process is achieved by adding a new S-DUCG without any modification of the basic triage knowledge base.

In the following section, the construction method and the configuration method of the basic knowledge base will be described.

A. USING M-DUCG TO BUILD A BASIC TRIAGE KNOWLEDGE BASE

A basic triage knowledge base is an integrated DUCG, which is composed of some sub-DUCGs, and each sub-DUCG stands for one independent department. The automatic integration of DUCG is made by our designed DUCG compiler, so we only need to concentrate on how to construct each department (sub-DUCG). For one department, we classify all of its clinical features into four types: department, risk factors, symptoms, and physical signs. Each sub-DUCG should clearly express the causality between the department and the other three types of clinical features. Because the clinical features include both single-valued variables and multivalued variables, M-DUCG is used to construct the department. The detailed department construction method is as follows:

A department is represented by a $B$-type variable. Risk factors are represented by $X$-type variables, and they are used as input to a double line logic gate ($SG$-type variable). The $SG$-type variable is used to express the combination of risk
The M-DUCG of otolaryngology, $B_5$, stands for otolaryngology. $X_{22}$ and $X_{23}$ are risk factors. Other variables are symptoms and physical signs related to otolaryngology. Labels show the medical significance of variables.

In Figure 4, $B_5$ represents otolaryngology, it has two states, state 0 being false (not belonging to this department) and state 1 being true (belonging to this department). $X_{22}$ (dust environment) and $X_{23}$ (noise environment) are risk factor variables, if they are true, they can increase the incidence of otolaryngology. $B_5$ represents the otolaryngology variable affected by risk factors. Its child variables are the symptoms and physical signs related to it. Red directed arcs stand for the intensity of causality between parents and children, and parameter matrices are given in Appendix A. Take the red direct arc from otolaryngology ($B_5$) to nasal obstruction ($X_1$) as an example, the parameter matrix between them is:

$$a_{1;5} = \begin{pmatrix} a_{1,0;5,0} & a_{1,0;5,1} \\ a_{1,1;5,0} & a_{1,1;5,1} \end{pmatrix} = \begin{pmatrix} \ast & \ast \\ \ast & 0.95 \end{pmatrix}$$

In this matrix, we only assign $a_{1,1;5,1} = 0.95$. It means that when otolaryngology related diseases occur, the probability that they can cause nasal obstruction to occur is 95%. For we do not concern the relationships between the other states, the parameters between them are given as “\(\ast\)”, which means it will not function during the inference calculation.

So far, the sub-M-DUCG representing for otolaryngology is completed, it clearly describes the causal relationships between the otolaryngology and its related clinical features. The other departments are also modeled as sub-M-DUCGs in the same way. After all the sub-M-DUCGs are finished, they are compiled into one complete M-DUCG by our DUCG compiler. During the compilation, some verification work is made to remove the redundant variables and lines in
different sub-M-DUCGs. The complete M-DUCG is shown in Figure 5. It includes 31 departments, and the detailed departments are listed in Table 1. 436 X-type variables are used to describe the clinical features (symptoms, physical signs, and risk factors) that are related to departments. 856 red lines present the causal relations between departments and clinical features. This M-DUCG is used as the basic triage knowledge base.

**Table 1. The departments contained in the basic triage knowledge base.**

| Variable  | Description                  |
|-----------|------------------------------|
| BX5       | Otorhinolaryngology          |
| BX6       | Cardiology                   |
| BX7       | Respiratory Medicine         |
| BX8       | Gastroenterology             |
| BX9       | Psychiatry                   |
| BX10      | Thoracic Surgery             |
| BX11      | Orthopedic                   |
| BX12      | Dermatology                  |
| BX13      | Neurology                    |
| BX14      | Neurosurgery                 |
| BX15      | Ophthalmology                |
| BX16      | Rheumatology & Allergy       |
| BX17      | Gynecology                   |
| BX18      | Breast Surgery               |
| BX19      | Stomatology                  |
| BX20      | Pain Department              |
| BX21      | General Surgery              |
| BX22      | Urologic                     |
| BX23      | Vascular Surgery             |
| BX24      | Andrology                    |
| BX25      | Nephrology                   |
| BX26      | Proctology                   |
| BX27      | Liver & Gall Surgery         |
| BX28      | Endocrinology                |
| BX29      | Hematology                   |
| BX30      | Nutritional Department       |
| BX31      | Infectious Disease           |
| BX32      | Venerology                   |
| BX33      | Head & Neck Surgery          |
| BX34      | Fever Clinic                 |
| BX35      | General Internal Medicine    |

**B. USING S-DUCG TO CONFIGURE THE BASIC TRIAGE KNOWLEDGE BASE**

In reality, the granularity of the department division for each hospital is different; the department division of large-scale general hospitals is detailed, while the division of county-level hospitals or community hospitals is rough. How to use our triage system to support the different triage needs in various hospitals? It needs our triage system to get the ability of universality. We equip our triage system with the universal ability by the “integration” and “disintegration” operations to our basic triage knowledge base.

1) INTEGRATION OPERATION

This operation will be employed when multiple departments of the triage knowledge base should be mapping into one department of the target hospital. In using the ear, nose, throat, pharynx, head, and neck diseases as an example. A county-level hospital usually classifies those diseases into one department, otolaryngology, while in our basic triage knowledge, they are divided into two departments, otolaryngology (Figure 4), and head & neck surgery (Appendix B). Therefore, department integration should be performed to make our triage knowledge base to match the department division of the target hospital before our basic triage system used in the target hospital. Department integration does not mean to construct one new department to replace the original two, instead, a new S-DUCG is constructed to integrate the two departments into one, therefore, the department integration does not require any modifications to the basic triage knowledge base. In fact, departmental integration is a logical OR relationship, that is, similar to the union operation of multiple sets, the department variable (BX-type variable) is single-valued, so the departmental integration can be transformed into an S-DUCG model. As shown in Figure 6, BX101 (Integrated Otolaryngology), BX5 (Otolaryngology), and BX34 (Head & Neck Surgery) form an S-DUCG, P101;5 and P101;34 represent the causal intensities between BX101 and its two parent variables. Since the relationship between the child variable and parent variable is inclusive, and all of them are single-valued, the parameter matrix of Pn;1 is always satisfied as follows:

\[
p_{n;1} = \Pr \{P_{n;1}\} = \left( \begin{array}{c}
-1 \\
-1 \\
1
\end{array} \right)
\]

That is, when the parent event occurs, the child event must occur. During the process of DUCG compilation, those three sub-DUCG were compiled into one hybrid DUCG, and it is shown as Appendix C. After department integration, BX101 will be used as a hypothesis (BX101 is called the hypothetical BX-type variable). BX5 and BX34 will not be employed as the hypothesis, just used as the intermediate results during the calculation (BX5 and BX34 are called the intermediate BX-type variables). To simplify the description, in the following section, we abbreviate the hypothetical BX-type variable as h-BX and an intermediate BX-type variable as i-BX. By this method, multiple departments can be merged into one.

2) DISINTEGRATION OPERATION

When a department in the triage knowledge base needs to be mapped to multiple departments in the target hospital, a disintegration operation is required. Using the fever clinic as an example, our triage knowledge base has a fever clinic (BX35) to triage patients with high fever, while the target community hospital does not. As a result, before our triage knowledge base is applied to the hospital, it is necessary to disintegrate the fever clinic to different departments of the target hospital associated with high fever according to the conditions. In the target hospital, departments with high fever symptoms are the infectious disease, rheumatology & allergy, and respiratory medicine, so the fever clinic should be disintegrated into these three departments according to the patient’s conditions. This situation can be modeled by the S-DUCG as showing...
FIGURE 5. The basic triage knowledge base based on M-DUCG. It is composed by 31 sub-M-DUCGs, each sub-M-DUCG stands for one department.
FIGURE 6. Otorhinolaryngology and Head & Neck Surgery are integrated into one department by using the S-DUCG model.

FIGURE 7. Fever Clinic are disintegrated into three departments by disintegration operation.

in Figure 7, and the parts of the parameters in Figure 7 are below it. In Figure 7, the fever clinical \((BX_{35})\) is disintegrated into three departments of the target hospital, infectious disease \((BX_{132})\), rheumatology & allergy \((BX_{116})\), and respiratory medicine \((BX_{107})\) by different conditions. These conditions are represented by conditional linkage events. For example, conditional \(P_{107;35}\) is the conditional linkage event between \(BX_{35}\) and \(BX_{107}\), and \(Z_{107;35} = X_{101,1} \cup X_{102,1} (X_{101,1} \text{ denotes the patient was from an infected area, } X_{102,1} \text{ denotes the patient was from the pastoral area or had a history of exposure to cattle and sheep})\) is its condition. If \(Z_{107;35}\) is true, then the relationship will be established, and the causal intensity is 0.8. Similarly, if \(X_{121,1}\) is true, then the relationship between \(BX_{35}\) and \(BX_{116}\) will be created. Otherwise, these relationships will not exist.

\[
p_{107;35} = \left(\begin{array}{c} - \\ - \\ 0.8 \end{array}\right), \quad Z_{107;35} = X_{101,1} \cup X_{102,1},
\]

\[
p_{116;35} = \left(\begin{array}{c} - \\ - \\ 0.6 \end{array}\right), \quad Z_{116;35} = X_{121,1}.
\]

By integration and disintegration operations, our basic knowledge base can be flexibly adjusted to meet the triage requirements of various hospitals. This method achieves the effect of integration and disintegration by adding some additional S-DUCGs, therefore, avoids modifications to the basic knowledge base, and equips our triage system with flexibility and scalability.

To verify the feasibility of this method, we used our knowledge base to achieve the triage requirements of a county-level hospital (target hospital). There are 25 departments in this hospital and all departments given in Table 2, but our knowledge base has 31 departments. To apply our knowledge base to this hospital, some departments in our knowledge base should integrate into the corresponding department of the target hospital. The relations of department integration show in Table 2. The Target Departments are the departments of the county-level hospital, and the Matching Relations show the matching relations between the departments in our basic triage knowledge base and that of the target hospital. The \(h-BX\) is the symbol of the department of the target hospital in the hybrid DUCG, it will be used as a hypothesis during the calculation. For example, the Respiratory Medicine of our knowledge base matches to the Respiratory Medicine of the target hospital exactly, so the Respiratory Medicine of our knowledge base was directly used as Respiratory Medicine of the target hospital, and \(BX_{7}\) is used as the symbol of the Respiratory Medicine in the hybrid DUCG. Another example, Dermatology & Venereology of the target hospital matched the Dermatology \((BX_{12})\), and Venereology \((BX_{33})\) in our knowledge base, so Dermatology \((BX_{12})\) and Venereology \((BX_{33})\) integrated as one department, Dermatology & Venereology, \(BX_{103}\) used as the symbol of Dermatology & Venereology. According to the matching relations in Table 2, we built one S-DUCG shown in Appendix D. After together compiling the M-DUCG in Figure 5 and S-DUCG in Appendix D, we get the final hybrid DUCG showed in Figure 8. This hybrid DUCG used as the outpatient triage knowledge base for the target hospital, and the inference calculation method for the hybrid DUCG model will be explained in the next section.

| Target Departments | Matching Relations | \(h-BX\) |
|--------------------|--------------------|--------|
| Respiratory Medicine | Respiratory Medicine \((BX_{7})\) | \(BX_{7}\) |
| Psychiatry | Gastroenterology \((BX_{4})\) | \(BX_{4}\) |
| Thoracic Surgery | Psychiatry \((BX_{33})\) | \(BX_{33}\) |
| Orthopedics | Thoracic Surgery \((BX_{45})\) | \(BX_{45}\) |
| Neurology | Neurology \((BX_{1})\) | \(BX_{1}\) |
| Neurosurgery | Neurosurgery \((BX_{6})\) | \(BX_{6}\) |
| Ophthalmology | Ophthalmology \((BX_{10})\) | \(BX_{10}\) |
| Rheumatology & Allergy | Rheumatology & Allergy \((BX_{15})\) | \(BX_{15}\) |
| Stomatolog | Stomatolog \((BX_{18})\) | \(BX_{18}\) |
| Pain Department | Pain Department \((BX_{5})\) | \(BX_{5}\) |
| Vascular Surgery | Vascular Surgery \((BX_{25})\) | \(BX_{25}\) |
| Nephrology | Nephrology \((BX_{24})\) | \(BX_{24}\) |
| Proctology | Proctology \((BX_{26})\) | \(BX_{26}\) |
| Liver & Gall Surgery | Liver & Gall Surgery \((BX_{47})\) | \(BX_{47}\) |
| Endocrinology | Endocrinology \((BX_{46})\) | \(BX_{46}\) |
| Fever Clinic | Fever Clinic \((BX_{35})\) | \(BX_{35}\) |
| Otorhinolaryngology | Otorhinolaryngology \((BX_{3})\) | \(BX_{3}\) |
| Head & Neck Surgery | Head & Neck Surgery \((BX_{26})\) | \(BX_{26}\) |
| Cardiovascular Medicine | Cardiology \((BX_{3})\), Hematology \((BX_{24})\) | \(BX_{3}, BX_{24}\) |
| Dermatology & Venereology | Dermatology \((BX_{13})\), Venereology \((BX_{15})\) | \(BX_{13}, BX_{15}\) |
| Gynecology | Gynecology \((BX_{14})\), Breast Surgery \((BX_{14})\) | \(BX_{14}\) |
| General Surgery | General Surgery \((BX_{27})\), Liver & Gall Surgery \((BX_{47})\) | \(BX_{27}, BX_{47}\) |
| Urology | Urologic \((BX_{3})\), Andrology \((BX_{24})\) | \(BX_{3}, BX_{24}\) |
| General Internal Medicine | Infectious Disease \((BX_{52})\), General Internal Medicine \((BX_{36})\) | \(BX_{52}, BX_{36}\) |
| Gastroenterology | Gastroenterology \((BX_{9})\), Nutritional Department \((BX_{55})\) | \(BX_{9}, BX_{55}\) |
Figure 8. A triage knowledge based on hybrid DUCG. It uses to describe the department division of a county level hospital. It is composed by the M-DUCG shown in Figure 5 and an S-DUCG shown in Appendix D.

IV. INFERENCE METHOD OF THE HYBRID DUCG

At present, the DUCG inference method is only appropriate for M-DUCG or S-DUCG single models. The reasoning steps include simplification, the acquisition of a hypothesis space, expression expansion, and probability calculation. Since the existing inference methods are no longer applicable to the hybrid DUCG model, a new inference algorithm for the hybrid DUCG described as follows:

1) Simplification. Simplify the hybrid DUCG by simplification rules 1-10 and 16 [36], [41] based on the received evidence $E$. Simplification can not only eliminate unrelated variables, causal relationships, and hypotheses
but also clearly express the causal relationship between possible hypotheses and evidence. Moreover, it can reduce the complexity of the probability calculation without losing accuracy.

2) Obtain the hypothesis set $S_H$. In the simplified hybrid DUCG, all the $h$-BXs are treated as possible hypothesis $H_{ij}$.

3) Expression expansion. Before probability calculations, evidence $E$ and $H_{ij}/E$ must be logically expanded along the causal chain to facilitate the subsequent numerical calculations. During the expansion process, logic operations such as AND, OR, XOR, NOT, absorption, exclusion, and complement are applied. By repeatedly using (1), (2), (3) and (4), the evidence is expanded to the form of sum-of-products composed of only \{B-, BX-, D-, A-, P-, r-\}-type variables. For each step in the logic expansion, if the relationship between the parent variable and child variable is of the $P$-type, the model between them is $S$-DUCG and equation (1) is used to expand them; if the relationship between them is of the $F$-type, the model between them is $M$-DUCG and equation (4) will be used.

4) Probability calculation. If there is more than one hypothesis in $S_H$, we need to rank them according to the hypothesis conditional probability. The formula for the conditional probability of $H_{ij}$ is defined as follows:

$$h^*_{ij} = \text{Pr} \{H_{ij}/E\} = \text{Pr} \{H_{ij}/E\} / \text{Pr} \{E\}$$   (5)

The formula for the rank probability of $H_{ij}$ is as follows:

$$h_{ij} = h^*_{ij} / \sum_{H_{ij} \in S_H} h^*_{ij}$$   (6)

The rank probabilities are the final result for causal reasoning.

A simple computational example is used to explain the reasoning algorithm. Figure 9 is an example used to illustrate the inference method of hybrid DUCG, and its related parameters are under it. This case shows when evidence $E = X_{13,1}$, which hypothesis is more likely to cause it to occur. The calculation process is as follows:

$$b_1 = (−0.03)^T, b_4 = (−0.05)^T,$$

$$b_7 = (−0.02)^T,$$

$$LGS_2 = LGS_5 = \begin{pmatrix} 0 & \text{Re} \\text{mnant} \\ 1 & X_{11,1} \end{pmatrix}, \quad sa_{3,2} = \begin{pmatrix} −1 & 1.2 \end{pmatrix},$$

$$LGS_5 = \begin{pmatrix} 0 & \text{Re} \\text{mnant} \\ 1 & 1.5 \end{pmatrix}, \quad sa_{9,8} = \begin{pmatrix} −1 \\ 1 \end{pmatrix}.$$  

$$p_{10.5} = p_{10.6} = \begin{pmatrix} −1 \\ 1 \end{pmatrix}, \quad a_{11,12} = (0 \ 1),$$

$$a_{13,3} = \begin{pmatrix} −1 & 0.4 \end{pmatrix}, \quad a_{13,6} = \begin{pmatrix} −1 & 0.5 \end{pmatrix},$$

$$a_{13,9} = \begin{pmatrix} −1 & 0.2 \end{pmatrix}, \quad a_{14,3} = \begin{pmatrix} −1 & 0.2 \end{pmatrix},$$

$$a_{14,9} = \begin{pmatrix} −1 & 0.8 \end{pmatrix}, \quad a_{15,3} = \begin{pmatrix} −1 & 0.9 \end{pmatrix},$$

$$a_{16,6} = \begin{pmatrix} −1 & 0.1 \end{pmatrix}, \quad a_{16,9} = \begin{pmatrix} −1 & 0.6 \end{pmatrix},$$

$$r_{ni,i} = 1, \quad a_{17,9} = \begin{pmatrix} −1 & 0.95 \end{pmatrix}.$$  

1) Simplification. In Figure 9, variables and relations not related to $X_{13,1}$ are eliminated, so $D_{12}$, $X_{11}$, $X_{14}$, $X_{15}$, $X_{16}$ and $X_{17}$ are eliminated. The simplified DUCG shows in Figure 10.

2) Obtain the hypothesis set. From the simplified DUCG, we obtain the hypothesis set $S_H = \{BX_{10.1}, BX_{9.1}\}$. $BX_3$ and $BX_6$ are $i$-BX type variable. They are not used as hypothesis.

3) According to the method of expression expansion, we obtain the expanded expression of $E$, $BX_{10.1}E$, and $BX_{9.1}E$ as follows:

$$E = X_{13,1}$$

$$= F_{13,1;3,1}BX_{3,1} + F_{13,1;6,1}BX_{6,1} + F_{13,1;9,1}BX_{9,1}$$

(7)

$$BX_{10.1}E$$

$$= F_{13,1;3,1}BX_{3,1} + F_{13,1;6,1}BX_{3,1}BX_{6,1}$$

$$+ F_{13,1;9,1}BX_{3,1}BX_{9,1}$$

$$+ F_{13,1;1,3}P_{10.1;3,1}P_{10.1;6,1}BX_{3,1}BX_{6,1}$$

$$+ F_{13,1;6,1}P_{10.1;3,1}P_{10.1;6,1}BX_{6,1}$$

$$+ F_{13,1;9,1}P_{10.1;3,1}P_{10.1;6,1}BX_{6,1}BX_{9,1}$$

$$+ F_{13,1;6,1}P_{10.1;3,1}P_{10.1;6,1}BX_{3,1}BX_{6,1}$$

$$+ F_{13,1;9,1}P_{10.1;3,1}P_{10.1;6,1}BX_{3,1}BX_{6,1}BX_{9,1}$$

(8)

$$BX_{9.1}E$$

$$= F_{13,1;3,1}BX_{3,1}BX_{9,1}$$

$$+ F_{13,1;6,1}BX_{6,1}BX_{9,1} + F_{13,1;9,1}BX_{9,1}$$

(9)

4) Probability calculation. By substituting the uppercase letters with their corresponding lowercase letters in the
Incompetent clinical information of patients may also lead to triage errors. To collect the clinical information of patients competently, a heuristic information collection method is designed for the triage system. The method is as follows:

1) Enter the patient’s self-reported symptoms into the system for an initial reasoning calculation and obtain the ranked result.

2) For the top 3 departments in the ranked result, select 6 unknown evidence ($X_i$) from each department ($BX_n$) with the highest occurrence probability as optional recommended items, and $\max (a_{i,k,n})$ is used as the occurrence probability of $X_i$. For each recommended item, according to the principle that the recommended item of departments with high ranked probability should be recommended first, the product of $\max (a_{i,k,n}) \times \Pr \{BX_n\}$ is used as the ranked value to rank the total 18 recommended items. Finally, the top 10 of the 18 recommended items and all the risk factors related to these 3 departments are selected as the recommended items for patients to answer, and the recommended evidence will no longer be recommended for the next time.

3) Repeat steps 1 and 2 until no new evidence is entered.

The final diagnosis result is usually existing in the top 3 hypotheses of each calculation, so the recommended items are chosen from the top 3 departments. 10 clinical features are chosen as the final recommended items are primarily based on the following two concerns. One is that too few recommended items are chosen as the final recommended items are primarily based on the following two concerns. One is that too few recommended items make the recommended evidence will no longer be recommended for the next time.

According to (5), we obtain the conditional probability of each hypothesis:

$$h_{10,1} = \Pr \{BX_{10,1}\}/\Pr (E) = 0.9092$$

$$h_{9,1} = \Pr \{BX_{9,1}\}/\Pr (E) = 0.1156$$

Their rank probabilities are as follows:

$$h_{10,1} = h_{10,1}^E (h_{10,1}^E + h_{9,1}^E) = 0.887$$

$$h_{9,1} = h_{9,1}^E (h_{10,1}^E + h_{9,1}^E) = 0.113$$

From the rank probabilities, we can see that $X_{13,1}$ is likely to be caused by $BX_{10,1}$.

V. HEURISTIC INFORMATION COLLECTION METHOD AND CASE STUDY

Incompetent clinical information of patients may also lead to triage errors. To collect the clinical information of patients competently, a heuristic information collection method is designed for the triage system. The method is as follows:

1) Enter the patient’s self-reported symptoms into the system for an initial reasoning calculation and obtain the ranked result.

2) For the top 3 departments in the ranked result, select 6 unknown evidence ($X_i$) from each department ($BX_n$) with the highest occurrence probability as optional recommended items, and $\max (a_{i,k,n})$ is used as the occurrence probability of $X_i$. For each recommended item, according to the principle that the recommended item of departments with high ranked probability should be recommended first, the product of $\max (a_{i,k,n}) \times \Pr \{BX_n\}$ is used as the ranked value to rank the total 18 recommended items. Finally, the top 10 of the 18 recommended items and all the risk factors related to these 3 departments are selected as the recommended items for patients to answer, and the recommended evidence will no longer be recommended for the next time.

3) Repeat steps 1 and 2 until no new evidence is entered.

The final diagnosis result is usually existing in the top 3 hypotheses of each calculation, so the recommended items are chosen from the top 3 departments. 10 clinical features are chosen as the final recommended items are primarily based on the following two concerns. One is that too few recommended items are likely to cause inaccurate recommendations. Another is that too many recommended items make the recommendation less concise. In practical applications, those parameters can be configured dynamically. An outpatient case was used to explain the heuristic information collection method and diagnosis process.

The patient’s self-report is as follows: male ($X_{73,1}$), joint pain ($X_{10,1}$) for two months and has had a fever ($X_{79,1}$) for two weeks. After entering this evidence $E = X_{40,1}X_{73,1}X_{79,1}$ into the triage system, we obtain the following top 3 results: rheumatology & allergy (44.13%), infectious disease (17.79%), and orthopedics (8.31%). Simultaneously, the system recommends that the patient should answer questions listed in Table 3 for further diagnosis. During the consultation, the patient said that he started raising sheep ($X_{102,1}$) 3 months ago, and some of the other breeders also showed the same symptoms as him. In addition, the patient had cold intolerance ($X_{115,1}$). With this evidence, $E = X_{102,1}X_{115,1}$ was input into our system for another calculation, and the result shows that the top 3 departments are infectious disease (55.42%), rheumatology & allergy (44.13%), and orthopedics (2.24%). Since no more evidence was provided by the patient, the triage finished. This result indicated that the patient should go to the infectious disease department to see a doctor. The patient was ultimately diagnosed with brucellosis, which

---

**TABLE 3. Question list recommended by the system for patients.**

| Question about history of past illness. | Question about symptom or physical sign. |
|----------------------------------------|-----------------------------------------|
| Tuberculosis history ($X_{90,1}$)?      | Dryness of mouth ($X_{66}$)?            |
| History of cattle and sheep exposure?   | Dryness of eyes ($X_{75}$)?             |
| from pastoral areas ($X_{102}$)?        | Lymphadenectasis ($X_{65}$)?            |
|                                           | Dyspnea ($X_{93}$)?                    |
|                                           | Lumbago ($X_{77}$)?                    |
|                                           | Skelalgia ($X_{60,1}$)?                |
|                                           | Cold intolerance ($X_{115,1}$)?         |
|                                           | Bellyache ($X_{21,1}$)?                |
|                                           | Diarrhea ($X_{31,1}$)?                 |
|                                           | Cough ($X_{61,1}$)?                    |
means that the outpatient triage was accurate. Figure 11 is a graphical explanation of the infectious disease department, it can explain all symptoms except the gender, for the onset of infectious diseases has nothing to do with gender.

To verify the performance of our triage system, we performed three experiments. The first was used to test the triage performance of our triage system. The second experiment was used as a comparison of our triage system and some machine learning methods mentioned above, and the last one was a comparison of our triage system and inexperienced triage nurses. All tests were run on a laptop with an Intel Core i7@2.70-GHz processor and 16-GB random access memory.

The cases used for experiments were gathered from the hospital information system (HIS) of three hospitals, 151 cases from Peking Union Medical College Hospital, 100 cases from Xuanwu Hospital Capital Medical University, 200 cases from Suining Central Hospital, total 451 cases, and all cases were selected randomly. During the experiments, only the self-reported symptoms of the patients recorded in cases were used as input.

Table 4 is the triage performance of our triage system. Test Cases are the number of cases used for testing each department. True is the number of test cases correctly triaged, and False is that of incorrectly triaged. Accuracy is the triage accuracy of each department. As shown in Table 4, the overall triage accuracy for our triage system was 96.8%, and the accuracy for each department is higher than 90% except for neurosurgery. 22 cases were used for testing neurosurgery, but 4 cases were incorrectly triaged into neurology. In reality, it is rather difficult to distinguish whether the patient should be triaged into neurology or neurosurgery only based on his self-reported symptoms. During the actual triage process, if he has the symptom of head trauma, he should be triaged into neurosurgery. Otherwise, he should be triaged into neurology. Therefore, the doctor doesn’t think these 4 cases are triage errors. If the 4 cases don’t be recognized as triage errors, then the accuracies for all departments were higher than 90%, and the overall accuracy would be 97.8%.

The experiment results for some machine learning algorithms are presented in Table 5. For each algorithm, 66% of the cases were employed as the training set, and the rest of the cases were used as the test set. Each experiment was repeated 10 times. As shown in Table 5, the average triage performances of the machine learning algorithms are mediocre, ranged from 43.9% to 75.8%. The SVM algorithm did the best triage performance with the highest accuracy, but it still 21% less accurate than our triage system. For the same algorithm, the triage accuracies for different departments varied greatly, some departments were lower than 50%, even 0%, while some reached to 90%. For our triage, the triage accuracies for different departments were stable.

The same cases were employed to test the triage performance of 7 triage nurses with less than two-year experience in outpatient triage. As shown in Table 6, the triage accuracy for all tested nurses ranged from 60.3% to 80.2%. The average accuracy was 68.9%, which means that more than 30% of patients were incorrectly triaged. If our triage system is employed to assist those inexperience nurses for outpatient triage, the accuracy of triage will be improved significantly.

Experimental results demonstrate that our triage system performs better than some machine learning algorithms and inexperienced triage nurses. Although machine learning algorithms can obtain the model through data learning, the quality of the model heavily depends on the data quantity and data quality of the training set. In essence, outpatient triage is a multi-classification problem with high-dimensional data. In addition to the excellent performance of the SVM algorithm in dealing with high-dimensional data, other algorithms have limited ability to cope with
TABLE 5. The triage result of some machine learning algorithms.

| Departments                      | KNN % | Naive Bayes % | Bayes Network % | SVM % | C4.5 % | Back Propagation % |
|----------------------------------|-------|---------------|-----------------|-------|--------|-------------------|
| Respiratory Medicine             | 87.9% | 66.7%         | 84.8%           | 87.9% | 78.8%  | 81.8%             |
| Psychiatry                       | 81.8% | 81.8%         | 0.0%            | 81.8% | 27.3%  | 81.8%             |
| Thoracic Surgery                 | 50.0% | 0.0%          | 0.0%            | 33.3% | 33.3%  | 50.0%             |
| Orthopedics                      | 68.8% | 81.3%         | 78.1%           | 78.1% | 71.9%  | 78.1%             |
| Neurology                        | 45.5% | 27.3%         | 0.0%            | 54.5% | 36.4%  | 45.5%             |
| Neurosurgery                     | 81.0% | 57.1%         | 33.3%           | 66.7% | 76.2%  | 76.2%             |
| Ophthalmology                    | 80.0% | 70.0%         | 0.0%            | 60.0% | 20.0%  | 80.0%             |
| Rheumatology & Allergy           | 85.7% | 50.0%         | 100.0%          | 89.3% | 96.4%  | 82.1%             |
| Stomatology                      | 70.0% | 80.0%         | 0.0%            | 80.0% | 50.0%  | 70.0%             |
| Pain Department                  | 28.6% | 42.9%         | 0.0%            | 42.9% | 28.6%  | 57.1%             |
| Vascular Surgery                 | 50.0% | 66.7%         | 0.0%            | 66.7% | 0.0%   | 50.0%             |
| Nephrology                       | 68.8% | 75.0%         | 0.0%            | 50.0% | 50.0%  | 43.8%             |
| Proctology                       | 80.0% | 66.7%         | 60.0%           | 80.0% | 80.0%  | 80.0%             |
| Endocrinology                    | 64.0% | 76.0%         | 40.0%           | 80.0% | 80.0%  | 68.0%             |
| Fever Clinic                     | 100.0%| 100.0%        | 0.0%            | 100.0%| 100.0% | 100.0%           |
| Otorhinolaryngology Head & Neck Surgery | 86.1% | 91.7% | 77.8% | 88.9% | 77.8% | 80.6% |
| Cardiovascular Medicine          | 65.2% | 69.6%         | 56.5%           | 52.2% | 60.9%  | 65.2%             |
| Dermatology & Venereology        | 82.6% | 60.9%         | 0.0%            | 82.6% | 43.5%  | 56.5%             |
| Gynecology                       | 44.4% | 50.0%         | 0.0%            | 88.9% | 66.7%  | 55.6%             |
| General Surgery                  | 74.3% | 80.0%         | 57.1%           | 68.6% | 74.3%  | 74.3%             |
| Urologic                         | 96.4% | 92.9%         | 35.7%           | 82.1% | 57.1%  | 78.6%             |
| General Internal Medicine        | 46.7% | 33.3%         | 0.0%            | 60.0% | 33.3%  | 40.0%             |
| Gastroenterology                 | 77.8% | 91.7%         | 55.6%           | 83.3% | 61.1%  | 75.0%             |
| Weight Average                   | 74.3% | 70.7%         | 43.9%           | 75.8% | 64.1%  | 71.0%             |

VI. CONCLUSIONS AND FUTURE WORK

In this study, we develop a general outpatient triage system based on the hybrid DUCG to aid triage nurses to achieve outpatient triage accurately. In this system, the M-DUCG is used to create a knowledge base as the basic outpatient triage base. The knowledge base contains 31 different departments, 436 clinical features related to those departments, and 856 lines are used to represent the causal relationships between departments and clinical features. The S-DUCG is
used to adjust the basic knowledge base to the different triage requirements of various hospitals. This makes the system have good scalability.

Validation experiments are made to compare the accuracy of our system with some machine learning algorithms and inexperienced triage nurses. Benefiting from the intuitive causal knowledge base and the chaining inference algorithm, the system achieves outpatient triage with high accuracy compared with some machine learning algorithms and the inexperienced outpatient triage nurses. In addition, the knowledge base is understandable to triage nurses, and the inference process and the results are intuitive representation by graphs. Those make triage conclusions and advice more explicable and convincing, further increasing the objectivity of outpatient triage. For the inexperienced triage nurses, they can obtain the triage knowledge by using this system constantly and thus improve their accuracy of triage.

In future work, we will further optimize the basic triage knowledge base and conduct actual triage tests in hospitals. Then, the system will be employed in hospitals to assist triage nurses with outpatient triage.

APPENDIX

A. Parameter matrices used to indicate the causal intensities between variables of otolaryngology.

\[ r_n;1 = 1, \quad b_5 = \begin{pmatrix} -0.0313 \end{pmatrix}, \quad a_{1;5} = \begin{pmatrix} -0.95 \end{pmatrix}, \]

\[ a_{3;5} = \begin{pmatrix} -0.95 \end{pmatrix}, \quad a_{4;5} = \begin{pmatrix} -0.8 \end{pmatrix}, \]

\[ a_{5;5} = \begin{pmatrix} -0.5 \end{pmatrix}, \]

\[ s_a_{5;5} = \begin{pmatrix} 1 & 1.3 & 1.6 \end{pmatrix}, \quad a_{6;5} = \begin{pmatrix} -0.9 \end{pmatrix}, \]

\[ a_{9;5} = \begin{pmatrix} -0.9 \end{pmatrix}, \quad a_{11;5} = \begin{pmatrix} -0.1 \end{pmatrix}, \]

\[ a_{12;5} = \begin{pmatrix} -0.95 \end{pmatrix}, \]

\[ a_{13;5} = \begin{pmatrix} -0.08 \end{pmatrix}, \quad a_{15;5} = \begin{pmatrix} -0.1 \end{pmatrix}, \]

\[ a_{16;5} = \begin{pmatrix} -0.9 \end{pmatrix}, \]

\[ a_{17;5} = \begin{pmatrix} -0.9 \end{pmatrix}, \quad a_{18;5} = \begin{pmatrix} -0.7 \end{pmatrix}, \]

\[ a_{19;5} = \begin{pmatrix} -0.8 \end{pmatrix}, \]

\[ a_{20;5} = \begin{pmatrix} -0.6 \end{pmatrix}, \quad a_{21;5} = \begin{pmatrix} -0.05 \end{pmatrix}, \]

\[ a_{22;D} = \begin{pmatrix} -1 \end{pmatrix}, \quad a_{23;D} = \begin{pmatrix} -1 \end{pmatrix}, \quad a_{24;D} = \begin{pmatrix} -1 \end{pmatrix}, \quad a_{30;5} = \begin{pmatrix} -0.06 \end{pmatrix}, \quad a_{38;5} = \begin{pmatrix} -0.2 \end{pmatrix}, \quad a_{50;5} = \begin{pmatrix} -0.02 \end{pmatrix}, \quad a_{51;5} = \begin{pmatrix} -0.1 \end{pmatrix}, \quad a_{67;5} = \begin{pmatrix} -0.1 \end{pmatrix}, \quad a_{79;5} = \begin{pmatrix} -0.2 \end{pmatrix}, \quad a_{90;5} = \begin{pmatrix} -0.6 \end{pmatrix}, \quad a_{97;5} = \begin{pmatrix} -0.8 \end{pmatrix}, \quad a_{98;5} = \begin{pmatrix} -0.7 \end{pmatrix}, \quad a_{99;5} = \begin{pmatrix} -0.9 \end{pmatrix}, \]

\[ LGS_5 = \begin{pmatrix} 0 & \text{Remnant} \\ 1 & X_{22,1} \\ 2 & X_{23,1} \end{pmatrix}, \quad s_a_{5;5} = \begin{pmatrix} 1 & 1.2 & 1.6 \end{pmatrix}. \]

B. The M-DUCG model of Head & neck surgery in the basic triage knowledge base.

C. The hybrid DUCG of integrated otolaryngology composed by otolaryngology (Figure 4), Head & neck surgery (Appendix B), and the S-DUCG in Figure 6.

D. The S-DUCG model used to represent the relationship of department integration shown in Table 2.
REFERENCES

[1] S. D. Pearson, L. Goldman, E. J. Orav, E. Guadagnoli, T. B. Garcia, P. A. Johnson, and T. H. Lee, “Triage decisions for emergency department patients with chest pain,” J. Gen. Internal Med., vol. 10, no. 10, pp. 557–564, Oct. 1995.

[2] W. Michalowski, S. Rubin, R. Slowinski, and S. Wilk, “Triage of the child with abdominal pain: A clinical algorithm for emergency patient management,” Paediatrics Child Health, vol. 6, no. 1, pp. 23–28, Jan. 2001.

[3] D. Mohan and A. E. Barnato, M. R. Rosengart, D. C. Angus, and R. Slowinski, “Triage of the child with abdominal pain: A clinical algorithm for emergency patient management,” Paediatrics Child Health, vol. 6, no. 1, pp. 23–28, Jan. 2001.

[4] A. I. Grouse, R. O. Bishop, and A. M. Bannon, “The Manchester triage system provides good reliability in an australian emergency department,” Emergency Med. J., vol. 26, no. 7, pp. 484–486, Jul. 2009.

[5] C.-J. Ng, K.-H. Hsu, J.-T. Kuan, T.-F. Chiu, W.-K. Chen, H.-J. Lin, “A fuzzy K-nearest-neighbor algorithm,” IEEE Trans. Syst., Man, Cybern., vol. SMC-6, no. 4, pp. 580–585, Aug. 1985.

[6] D. K. Nishijima, S. D. Gaona, T. Waechter, R. Maloney, T. Bair, A. Blitz, N. Friedman, D. Geiger, and M. Goldszmidt, “Bayesian network classifiers,” Mach. Learn., vol. 29, no. 2, pp. 131–163, Nov. 1997.

[7] J. Schmidhuber, “Deep learning in neural networks: An overview,” Proc. IEEE, vol. 100, no. 6, pp. 2203–2217, 2012.

[8] Y. Lecun, Y. Bengio, and G. Hinton, “Deep learning,” Nature, vol. 521, no. 7553, pp. 436–444, 2015.

[9] G. Wang, “The practice and effect of improving the quality of triage nursing in outpatient department,” J. Nursing Admin. China, vol. 8, no. 8, pp. 21–22, Aug. 2008.

[10] S. L. Liu, “Research on application of standard operating procedures in quality management of outpatient triage system,” J. Clin. Med. Pract. China, vol. 16, no. 18, pp. 101–103, May 2012.

[11] B. Mistry, S. Stewart De Ramirez, G. Kelen, P. S. K. Schmitz, K. S. Balhara, S. Levin, D. Martinez, K. Psoter, X. Antón, and J. S. Hinson, “Accuracy and reliability of emergency department triage using the emergency severity index: An international multicenter assessment,” Ann. Emergency Med., vol. 71, no. 5, pp. 581–587.e3, May 2018.

[12] D. R. Vinson, H. Garzon, J. S. Hinson, and D. Ballard, “Out-of-hospital triage of older adults with head injury: A retrospective study of the effect of adding ‘anti-coagulation or anti-platelet medication use’ as a criterion,” Ann. Emergency Med., vol. 70, no. 2, pp. 127–138, Feb. 2017.

[13] M. S. Ahmad and D. R. Wood, “Building a visual aid for a triage tool in a developing health service,” J. Clin. Epidemiol., vol. 94, pp. 27–34, Feb. 2018.

[14] J. M. Zachariasse, D. Nieboer, R. Oostenbrink, H. A. Moll, and E. W. Steyerberg, “Multiple performance measures are needed to evaluate triage systems in the emergency department,” J. Clin. Epidemiol., vol. 94, pp. 27–34, Feb. 2018.

[15] S. Levin, M. Toerper, E. Hamrock, J. S. Hinson, S. Barnes, H. Gardner, A. Dugas, B. Linton, T. Kirsch, and G. Kelen, “Machine-learning-based electronic triage more accurately differentiates patients with respect to clinical outcomes compared with the emergency severity index,” Ann. Emergency Med., vol. 71, no. 5, pp. 565–574.e2, May 2018.

[16] X. SiMa and Y. Wang, “General hospital out-patient clinic service points defects and Countermeasures,” J. Traditional Chin. Med. Univ. Hunan, vol. 30, no. 4, pp. 69–70 Apr. 2010.

[17] J. Liang et al., “Factors affecting accuracy rate of outpatient triage in the large hospital and the countermeasures,” Chin. Clin. Nursing, vol. 6, no. 6, pp. 525-526, Nov. 2014.

[18] S. A. Dudani, “The distance-weighted k-nearest-neighbor rule,” IEEE Trans. Syst., Man, Cybern., vol. SMC-6, no. 4, pp. 325–327, Apr. 1976.

[19] J. M. Keller, M. R. Gray, and J. A. Givens, “A fuzzy K-nearest-neighbor algorithm,” IEEE Trans. Syst., Man, Cybern., vol. SMC-15, no. 4, pp. 580–585, Aug. 1985.

[20] N. Majumder, S. Poria, A. Gelbukh, and E. Cambria, “Deep learning-based word with support vector machine,” Knowl.-Based Syst., vol. 61, pp. 85–117, Sep. 2016.

[21] X.-W. Chen and X. Lin, “Big data deep learning: Challenges and perspectives,” IEEE Access, vol. 2, pp. 514–525, 2014.

[22] Z. Sun, Z. Guo, C. Liu, W. Jiang, J. Liu, and S. Liu, “Fast extended one-versus-rest multi-label support vector machine using approximate extreme points,” IEEE Access, vol. 5, pp. 8526–8535, 2017.

[23] J. Zhang and Y. Wang, “A rough margin based support vector machine,” Inf. Sci., vol. 178, no. 9, pp. 2204–2214, May 2008.

[24] Z. Sun, Z. Guo, C. Liu, W. Jiang, J. Liu, and S. Liu, “Fast extended one-versus-rest multi-label support vector machine using approximate extreme points,” IEEE Access, vol. 5, pp. 8526–8535, 2017.

[25] J. Schmidhuber, “Deep learning in neural networks: An overview,” Neural Netw., vol. 61, pp. 85–117, Jan. 2015.

[26] Z.-H. Zhou, Y. Jiang, Y.-B. Yang, and S.-F. Chen, “Lung cancer cell identification based on artificial neural network ensembles,” IEEE Trans. Neural Netw., vol. 21, no. 1, pp. 85–117, Jan. 2010.

[27] Z. Sun, Z. Guo, C. Liu, X. Wang, J. Liu, and S. Liu, “Fast extended one-versus-rest multi-label support vector machine using approximate extreme points,” IEEE Access, vol. 5, pp. 8526–8535, 2017.

[28] J. Zhang and Y. Wang, “A rough margin based support vector machine,” Inf. Sci., vol. 178, no. 9, pp. 2204–2214, May 2008.

[29] M. S. Ahmad and D. R. Wood, “Building a visual aid for a triage tool in a developing health service,” J. Clin. Epidemiol., vol. 94, pp. 27–34, Feb. 2018.

[30] J. M. Keller, M. R. Gray, and J. A. Givens, “A fuzzy K-nearest-neighbor algorithm,” IEEE Trans. Syst., Man, Cybern., vol. SMC-6, no. 4, pp. 580–585, Aug. 1985.

[31] N. Majumder, S. Poria, A. Gelbukh, and E. Cambria, “Deep learning-based word with support vector machine,” Knowl.-Based Syst., vol. 61, pp. 85–117, Sep. 2016.
Q. Zhang, “Probabilistic reasoning based on dynamic causality trees/diagrams,” Rel. Eng. Syst. Saf., vol. 46, no. 3, pp. 209–220, Jan. 1994.

Q. Zhang, “Dynamic uncertain causality graph for knowledge representation and reasoning: Discrete DAG cases,” J. Comput. Sci. Technol., vol. 27, no. 1, pp. 1–23, Jan. 2012.

Q. Zhang, C. Dong, Y. Cui, and Z. Yang, “Dynamic uncertain causality graph for knowledge representation and probabilistic reasoning: Statistics base, matrix, and application,” IEEE Trans. Neural Netw. Learn. Syst., vol. 25, no. 4, pp. 645–663, Apr. 2014.

C. Dong and Q. Zhang, “Research on weighted logical inference for uncertain fault diagnosis,” (in Chinese) Acta Automatica Sinica, vol. 40, no. 12, pp. 2766–2781, 2014.

Q. Zhang, “Dynamic uncertain causality graph for knowledge representation and probabilistic reasoning: Directed cyclic graph and joint probability distribution,” IEEE Trans. Neural Netw. Learn. Syst., vol. 26, no. 7, pp. 1503–1517, Jul. 2015.

C. Dong and Q. Zhang, “The cubic dynamic uncertain causality graph: A methodology for temporal process modeling and diagnostic logic inference,” IEEE Trans. Neural Netw. Learn. Syst., early access, Jan. 3, 2020, doi: 10.1109/TNNLS.2019.2953177.

Q. Zhang and S. Geng, “Dynamic uncertain causality graph applied to dynamic fault diagnoses of large and complex systems,” IEEE Trans. Rel., vol. 64, no. 3, pp. 910–927, Sep. 2015.

C. Dong, Y. Zhao, and Q. Zhang, “Assessing the influence of an individual event in complex fault spreading network based on dynamic uncertain causality graph,” IEEE Trans. Neural Netw. Learn. Syst., vol. 27, no. 8, pp. 1615–1630, Aug. 2016.

Z. Zhou and Q. Zhang, “Model event/fault trees with dynamic uncertain causality graph for better probabilistic safety assessment,” IEEE Trans. Rel., vol. 66, no. 1, pp. 178–188, Mar. 2017.

Y. Qu, Q. Zhang, and Q. Zhu, “Application of dynamic uncertain causality graph to dynamic fault diagnosis in chemical processes,” (in Chinese), CAAI Trans. Intell. Syst., vol. 10, no. 3, pp. 354–361, 2015.

C. Dong, Z. Zhou, and Q. Zhang, “Cubic dynamic uncertain causality graph: A new methodology for modeling and reasoning about complex faults with negative feedbacks,” IEEE Trans. Rel., vol. 67, no. 3, pp. 920–932, Sep. 2018.

C. Dong, X. Wang, Q. Zhang, and N. Wang, “The methodology of dynamic uncertain causality graph for intelligent diagnosis of vertigo,” Comput. Methods Programs Biomed., vol. 113, no. 1, pp. 162–174, Jan. 2014.

S.-R. Hao, S.-C. Geng, L.-X. Fan, J.-J. Chen, Q. Zhang, and L.-J. Li, “Intelligent diagnosis of jaundice with dynamic uncertain causality graph model,” J. Zhejiang Univ.-Sci. B, vol. 18, no. 5, pp. 393–401, May 2017.

A. Geng and Q. Zhang, “Calculation method to diagnose integrated causes of faults in process systems by means of dynamic uncertain causality graph,” in Proc. Asia-Pacific Conf. Comput. Sci. Appl., Shanghai, China, Dec. 2014, pp. 306–311.

XUSONG BU was born in Shandong, China, in 1988. He received the B.S. degree in computer science and technology from Ludong University, Yantai, China, in 2012, and the M.S. degree in computer software and theory from Ningxia University, Yinchuan, China, in 2015. He is currently pursuing the Ph.D. degree with the School of Computer Science and Engineering, Beihang University. His current research interests include intelligent fault diagnosis for complex systems, uncertainty in artificial intelligence, computer-aided clinical decision support methods, and recommendation systems based on DUCG.

LIN LU was born in Shanxi, China, in 1974. She received the M.D. degree in clinical medicine from Peking Union Medical College, Beijing, China, in 1999. She has been a Resident, a Chief Resident, an Attending Physician, and an Associate Professor, successively, with the Department of Endocrinology, Peking Union Medical College Hospital, since 1999. She currently specializes in endocrinology as an Associate Professor.

ZHAN ZHANG was born in Chongqing, China, in 1985. He received the B.S. degree in electrical engineering from Tsinghua University, Beijing, China, in 2008, and the M.S. degree in information technology from the University of New South Wales, Sydney, Australia, in 2010. He is currently pursuing the Ph.D. degree with Tsinghua University. After that, he worked as a Network Engineer with URSYS Company, Sydney, for six years. He set up Beijing Tsingrui Intelligence Technology Company, Ltd., Beijing, China, two years ago, and is the CEO. He has long been involved in the development of DUCG theory and applications. He designed the DUCG cloud platform and conducts the development of the clinical knowledge base constructions and tests. His current research interests include reliability engineering, the fault diagnosis of industrial systems, and AI-add clinical diagnoses.

QIN ZHANG received the B.S., M.S., and Ph.D. degrees from Tsinghua University, Beijing, China, in 1982, 1984, and 1989, respectively, all in nuclear engineering. He was a Visiting Scholar with the University of Tennessee, Knoxville, TN, USA, and the University of California at Los Angeles, Los Angeles, CA, USA, from 1987 to 1989, where he was studying system reliability engineering and intelligent fault diagnoses. He was a Postdoctoral Fellow of the School of Economics and Management, Tsinghua University, from 1990 to 1991, and an Associate Professor with the Institute of Nuclear and New Energy Technology (INET), Tsinghua University, from 1991 to 1993, where he was conducting online safety-related expert system development supported by the IAEA. He was the Department Manager of China Technology Innovation Corporation and the General Manager of United Company, for Technology Innovation, Xiamen, China, from 1993 to 1997. He was the Head of the Chongqing Science and Technology Commission, from 1997 to 2003, and the Deputy Commissioner of the State Intellectual Property Office, China, from 2003 to 2009. From 2009 to 2016, he was an Executive Secretary of the China Association for Science and Technology (CAST), where he was elected as the Vice President. He is currently the Director of the specialized committee on uncertainties in AI with the China Association for Artificial Intelligence (CAAI), and the Deputy Director of the Specialized Committee of the China Institute of Electrical and Energy Engineers. He is also the Chief Scientist of Beijing Tsingrui Intelligence Technology Company, Ltd. He is a Professor and a Ph.D. Supervisor of the INET and the Department of Computer Science and Technology, Tsinghua University, and the School of Computer Science and Engineering, Beihang University, Beijing, therein working on the theory and platform development of uncertain causality representations and probabilistic reasoning, as well as applications in the diagnoses of faults of nuclear power plants, chemical plants, spacecraft, and conditions such as vertigo and jaundice. He is a Fellow of the CAAI and is a member of the International Nuclear Energy Academy.

YAN ZHU was born in Jilin, China. He received the Ph.D. degree in nuclear science and engineering from Tsinghua University, Beijing, China, in 1998. He is currently a Professor and Doctoral Supervisor with the Department of Management Science and Engineering, School of Economics and Management, Tsinghua University. He is the Dean of the Internet Industry Research Institute, Tsinghua University, where he is also the Director of the Advanced Information Technology Application Laboratory, School of Economics and Management, and the Executive Deputy Director of the Medical Management Research Center.