

Regular Expression Based Online Aided Decision Making Knowledge Base for Quality and Security of Food Processing

Liu Wen, Liu Peng, Li Qiang, Duan Min, Wang Yan-Rong, Dai Yue

*Institute of Food and Agricultural Standardization, China National Institute of Standardization,
Beijing, 100088 China*

*Emails: liuwen@cnis.gov.cn liupeng@cnis.gov.cn liqiang@cnis.gov.cn duanmin@cnis.gov.cn
wangyr@cnis.gov.cn daiyue@cnis.gov.cn*

Abstract: *With regard to the inefficient application of a food processing information system due to shortage of the knowledge acquisition measure and self-updating function of knowledge, a method of constructing an online aided decision making knowledge base for quality and security of food processing, based on regular expression, is discussed in the paper. Firstly, the method establishes an online aided decision making knowledge base for quality and security of food processing based on regular expression; and then an automatic knowledge inference engine is applied to update the knowledge in the base, combined with industry experts' experience knowledge. Continuous deriving of food processing knowledge can be realized based on the inference engine. The research will greatly enhance the efficiency and applicability of obtaining knowledge from an online aided decision making system for quality and security of food processing.*

Keywords: *Food control, food safety, aided decision, knowledge base, regular expression.*

1. Introduction

An important aspect of food security is the quality and safety during the procedures of food processing. The food production procedure in modern food processing industry often involves multiple links and steps, each one having connection with food quality and safety. Therefore, it is of very important practical significance to

monitor the food production procedure online and to provide support for decision making. The parameters of food production process monitoring include various data, such as necessary test data, processing technology parameters, personnel operating records, records of the processing environment and other data. The parameters are numerous in kind and the data processing is huge in capacity. In order to response to the current status of quality and safety of the food production process timely, it is necessary to apply an online aided decision making system for quality and security of food processing in order to assist the operator to respond promptly and give advice in decision making.

At present, researches related to online supporting decision-making knowledge base are relatively less both at home and abroad, although there are some reports concerning other industries. Meng [1] has studied the warning knowledge model based on the knowledge base of a coal mining accident and then to better use it in the prevention of coal mining accidents. Feng [2] has combined the research methods and opinion digging of the knowledge base, focusing on developing the research of construction of knowledge base based on cyber car reviews corpus oriented to opinion digging. Zhong, Liu and Ding [3] have established a knowledge base on hybrid reasoning. The knowledge base is able to realize information inquiries on knowledge and semantic level, and it can also solve complex problems. Its rich semantic descriptions and mixing reasoning ability make up the shortcomings of the traditional knowledge base. Xu and Liu [4] analyzed the inadequacy of the classical first-order predicate logic as base of knowledge base's system logic. According to the characteristics and needs of the knowledge base system, they pointed out that logic of computability cannot only provide a reasonable logical basis for the knowledge base system; it can also make up the inadequacy of classical logic as the logical basis of knowledge base. Song et al. [5] have implemented a more in-depth study and discussion to the establishment of fault diagnosis expert system knowledge base for oil spectral analysis from both theoretical and practical aspects. Guo [6] has systematically introduced the development, achieved the mode and research development of open access in research and practice both home and abroad; and they have discussed the two main realization modes of open access – open access of journals and open repository. In addition to that, institutional repositories, subject repositories are the two main forms of open repository. The above-mentioned knowledge base of decision making is not able to be applied in food processing. The research in this paper is based on establishment of online supporting decision-making knowledge base basing on regular expression modes. After that to apply the knowledge inference engine to automatically update the knowledge in the system knowledge base in order to realize the construction of knowledge base of the quality and safety during the procedure of food production and processing, to efficiently improve the intelligence level of an online aided decision making system for quality and security of food processing. Liu and Sun [7] have considered that knowledge base is the core of the expert system of mine safety early warning. The construction of knowledge base is the key to establish the expert system of mine safety. This paper introduces the classification of knowledge classification and knowledge acquisition process of the

expert system of mine safety warning, and the knowledge and heuristic knowledge of mine safety field are represented by predicate logic and a generative rule, and then SQL Sever2000 Microsoft is used to establish the knowledge base, which lays a good foundation for the establishment and implementation of a mine safety early warning expert system. Ma [8] studied a mine water disaster early warning expert system and gave the mine water disaster early warning expert system knowledge generation and expression ways and means, and the design of inference rules of the mine water hazard warning. Finally, on the basis of the research of the developed mine water disaster early warning expert system, Ni, Zhang and Wang [9] put forward a kind of a combination forecasting model based on ontology, and solves the problem of semantic heterogeneity among the single forecasting model and index system on the basis of analyzing the deficiency of existing forecasting and warning system. Wang and Wang [10] have based an enterprise early warning support system composed of many intelligent systems which can run and cooperate with each other. Every qualitative or quantitative method is designed to be an early warning agent, each one having the ability to solve the corresponding solving method, knowledge processing and communication with other agents. Each agent is learning continuously to improve its ability, so as to improve the accuracy of the early warning system. Liu and Liu [11] have proposed the research idea of the “based on the “3S” in the decision support system of the Xinjiang snow melting flood early warning, based on the data warehouse, method base, model base and knowledge base. Fu Guixiang et al. have introduced the overall framework of the coal mine safety comprehensive intelligence warning system based on Internet of things. Reference [12] expounds the design of the early warning model, the gas early warning model, the fire warning model, the flood warning model and the roof early warning model. The system can carry out real-time monitoring and intelligent warning for the safety hazard in coal mine safety production process, and efficiently improve the safety production level of the coal enterprise. Li, Chen, and Huang [13], in a view point of the current problems in the research and application of network public opinion monitoring and early warning, based on the analysis of the necessity of establishing the knowledge base of public opinion, give in their paper the definition of a network public opinion, and analyze the composition of the knowledge base, and explain the structure design and the main properties of various kinds of knowledge bases. Wu [14] has analyzed the construction method of the knowledge base and the structure of the knowledge base system. The research on knowledge representation, the knowledge representation and application of a neural network were discussed, and the knowledge representation, storage and reasoning of the expert system based on a neural network were pointed out. These three groups were characterized by self organization, self-learning and self adaptation. A neural network in the construction industry knowledge expression, knowledge acquisition, using a neural network of architectural engineering quantity estimates for knowledge representation, knowledge acquisition research and application, gets good results.

2. Overview of the online aided decision-making knowledge base for quality and security of food processing based on regular expression

In the online aided decision-making system for quality and security of food processing, in [15-17] a knowledge set of industrial quality and security control experience and knowledge represented according to some kind of the form is stored, and is called an online aided decision-making knowledge base for quality and security of food processing based on regular expression. This knowledge base generally includes the next aspects of knowledge: 1) deterministic knowledge: refers to common-sense knowledge that is skilfully applied within industry; 2) experience-based knowledge: refers to the knowledge that is gained with the help of continuous accumulation by the industry experts in their practices of quality and safety control in long terms; 3) reasoning-based knowledge refers to knowledge that is produced after inference reasoning with the knowledge existing in the knowledge base. Deterministic knowledge and experience-based knowledge is knowledge gained from the mature experience of industry. They are summarized and comprehended constantly by the experts after years of practice in [18]. They are the main source of knowledge for logical reasoning in a knowledge base.

Large quantities of quality and safety control of knowledge and information are required in the design of online aided decision-making knowledge base for quality and security of food processing based on regular expression; the construction of the knowledge base is necessary to provide support for the system. The exact nature, availability and integrity of the quality and security control knowledge collected in the knowledge base is an important factor to evaluate whether the system is excellent or not. Wide range of the content, including live data, quality and safety control knowledge and other information are involved in the knowledge base. Therefore, the difficulty of constructing the knowledge base lies in how to carry out the acquisition, storage and presentation of the information. Since the establishment of the knowledge base is a very complex process, the base is usually composed of many sub knowledge bases thereof. Therefore, when establishing the knowledge base, a sub knowledge base must be created firstly, and then the knowledge base will be expanded continuously. In this paper, according to the actual need of online aided decision-making knowledge base for quality and security of food processing, different sizes of bases are used.

3. A summary of the knowledge base of food processing quality and safety

The assistant decision making system, which is stored in a certain form, is called knowledge base, which is called knowledge. The process knowledge database generally includes two aspects of knowledge:

1. Basic knowledge, that refers to the existing knowledge and knowledge in the long-term process of continuous accumulation of knowledge. Basic knowledge is the most primary source of knowledge in an expert system, which is summarized and realized by the experts in the field.

2. Knowledge is derived from the knowledge of the basic knowledge and knowledge base, and with the help of the industry experts recognizing the knowledge and knowledge of the library, repeat of the new knowledge is realized.

4. Food processing quality and safety online assistant decision knowledge representation

Representation of Knowledge is a kind of organization and description of the knowledge; it is the formalization, symbolization and modelling of the knowledge. There are many methods to represent the knowledge. As for regular expression methods, it means that each regular expression is corresponding to a piece of rule. Its basic form is " $P \rightarrow Q$ " or "IF P THEN Q ", wherein P refers to a prerequisite or premise, Q refers to concluding or a later member. The meaning of the basic form is if the premise P meets the requirement, it is capable to infer the later member Q or the performance of operation specified by the later member Q .

Advantages of regular expressions is that it is very close to human thinking, The content is easy to be understood, and the knowledge is very convenient to access to. At the same time, the rules are easy to modify, expand and they are independent. Large number of regular expression rules can constitute a knowledge base. The input questions will be compared with the premise of the rules in the knowledge base, one by one when seeking for a solution of the problem, and the reasoning will be made step by step. The system will use the regular expression notation to represent the knowledge, and construct the knowledge base of online aided decision-making system for quality and security of food processing.

The knowledge adopted in the knowledge base of online aided decision-making system for quality and security of food processing is in the form of $(R, F) \rightarrow C$, where:

F – represents the hazard phenomenon of quality and safety during the procedure of food processing, such as detecting excessive lead content of pasteurized dairy products, too high temperature of the outlet of a pasteurization machine;

R – represents likely reasons of hazard phenomena of quality and safety during the procedure of food processing, such as detected excessive lead content of pasteurized dairy products that may be due to the migration of heavy metals in the processing pipeline; too high temperature of the outlet of a pasteurization machine may be the result from too small pressure of the cooling system water and so on;

C – represents the corresponding control measures against a hazard phenomenon of quality and safety during the procedure of food processing; such control measures may appear if the pasteurized dairy products detect high levels of lead that exceed the replacement process piping and isolate product; pasteurizer outlet-temperature control measures are to increase the cooling system water pressure.

From the regular expression rules introduced in the previous section we can see that the former members or the later members are disjunction of the conjunction of a limited number of facts or conclusions. In the repository of rules, there may be

rules that have different premises but have the same conclusion, but the rules of regular expressions require consistency. There is no relationship between the former members of the rules that have the same later members.

5. Realization procedure of a knowledge inference machine

According to the characters of the knowledge base constructed on the present system, forward reasoning technique is applied in the system. This system uses forward inference (i.e., data-driven reasoning – the reasoning model is according to the mode of reasoning by conditions introduced by the conclusion of directions, a reasoning machine from the initial knowledge by using some inference rules, in order to prove the establishment of the objective fact or proposition. According to the rules of reasoning to produce new knowledge to add to the knowledge, the reasoning process is to first add to the knowledge base by industry experts, to determine the initial known knowledge. A knowledge inference engine based on the use of certain reasoning rules matches the traversal reasoning of the knowledge base in initial knowledge. Repeating the above process and updating the knowledge base of facts, another knowledge matching knowledge will update to the knowledge base of their conclusions until no matching new knowledge appears and there is no longer a new knowledge in the knowledge base so far. The diagram is shown as in Fig. 1. The reasoning process is in two steps.

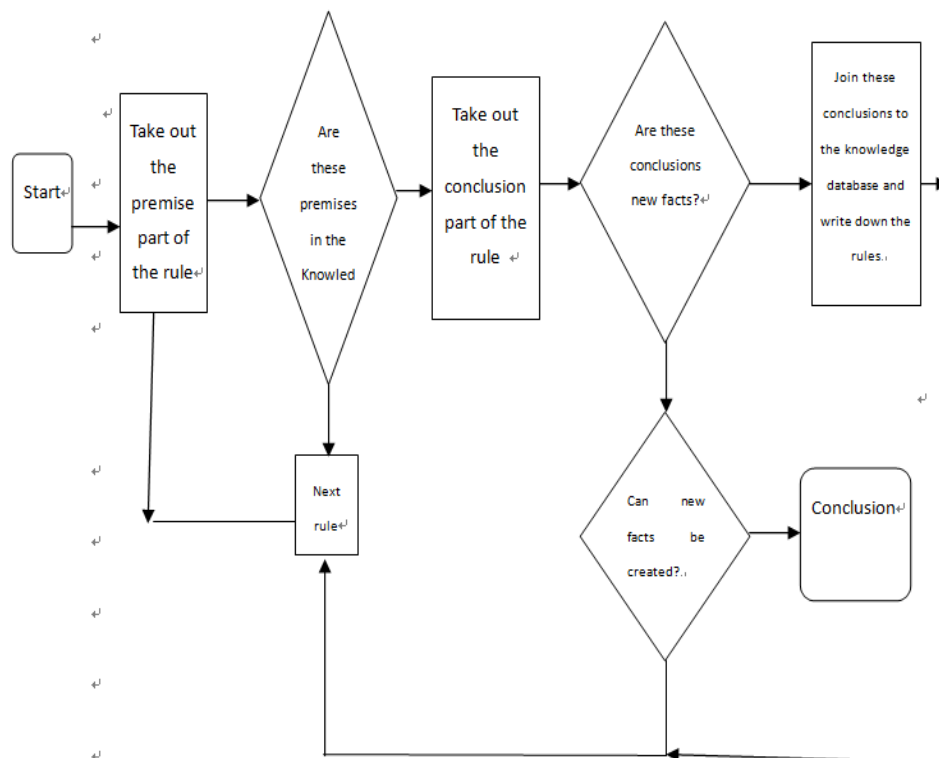


Fig. 1. Schematic diagram of the knowledge base reasoning machine

Step 1. The inference engine will match the facts provided by the users with the premise of the rules in the knowledge base.

Step 2. The conclusion of successfully matched rules will be added to the knowledge bases as new facts (in this case, the facts in the knowledge base that have increased).

With all the facts in the updated knowledge base, the above two steps will be repeated until a final conclusion is gained or no new facts can be added to the knowledge base.

If there are three pieces of knowledge in the knowledge base:

$(R1, R2, F1) \rightarrow C1,$

$(R3, R4, F2) \rightarrow C2,$

$(R3, R4, F3) \rightarrow C3,$

If $C1 = R3$.

Then new knowledge can be obtained.

$(R1, R2, C1, R4, F1, F2) \rightarrow C2,$

$(R3, R4, C2, R6, F2, F3) \rightarrow C3,$

$(R1, R2, C1, R4, R3, C2, R6, F1, F2, F3) \rightarrow C3.$

From the schematic diagram of the forward inference engine, at least the following features can be known:

(1) it knows how to use knowledge in the knowledge base based on the known facts;

(2) it can store the conclusions made through reasoning got in the knowledge base;

(3) it can determine when to end the reasoning process.

6. Verification by experiments

In the production process of pasteurized milk in a dairy plant in Beijing, abnormal disposal knowledge in the production process of the pasteurization machine will be taken as an example to verify the application effect of online aided decision-making knowledge base for quality and security of food processing based on regular expression.

In the production process of the Beijing dairy plant's pasteurized milk production process, the pasteurization machine has the following 8 pieces of primal knowledge. We can have a description according to the next 11 form $(R, F) \rightarrow C$:

1. (Error position of the allocated disk, ingredients GS error position) \rightarrow (check position of the dispensing tray)

2. (Error position of the allocated disk, pasteurization GS error position) \rightarrow (check position of the dispensing tray)

3. (Fault of the supply system, the pasteurization tank is not ready) \rightarrow (Check the water supply system)

4. (Inadequacy of the cooling system pressure, the pasteurization outlet temperature is too high) \rightarrow (increase the water pressure of the cooling system)

5. (Insufficient steam pressure and hot water pressure, the pasteurization temperature is too low) \rightarrow (increased steam pressure and hot water pressure)

6. (Steam pressure and water pressure is too large, the pasteurization temperature is too high) → (reducing the vapour pressure and hot water pressure))

7. (Insufficient soft water pressure, homogenizer cooling water pressure is low) → (increase the soft water pressure)

8. (Increase the water pressure of the cooling system, the pasteurization temperature is too low) → (increase the vapour pressure and water pressure)

The forward inference engine will reason the existing 8 pieces of knowledge based on the regular expression methods of the online supporting decision-making knowledge base of quality and security of food processing.

The later member of the first piece of knowledge C (check the position of the dispensing tray) does not belong to the former member of any knowledge; check the second piece of knowledge.

The later member of the second piece of knowledge C (check the position of the dispensing tray) does not belong to the former member of any knowledge, check the third piece of knowledge.

The later member of the third piece of knowledge C (check the water system) does not belong to the former member of any knowledge, check the forth piece of knowledge.

The later member of the forth piece of knowledge C (increase the water pressure of cooling system) belongs to the former member of the eighth piece of knowledge, new knowledge 9 is formed.

9. (Insufficient cooling system pressure (increase the water pressure of the cooling system, over temperature of pasteurization outlet, too low pasteurization temperature) → (increase the vapour pressure and water pressure)

The new knowledge 9 is traversal of the existing knowledge 8, no duplication, as the official new knowledge into the knowledge base 9.

Due to the emergence of new knowledge, start from the first traversal derivation.

The later member of the first piece of knowledge C (check the position of the dispensing tray) does not belong to the former member of any knowledge; check the second piece of knowledge.

The later member of the second piece of knowledge C (check the position of the dispensing tray) does not belong to the former member of any knowledge, check the third piece of knowledge.

The later member of the third piece of knowledge C (check the water system) does not belong to the former member of any knowledge, check the forth piece of knowledge.

The later member of the forth piece of knowledge C (increase the water pressure of the cooling system) belongs to the former member of the eighth piece of knowledge, new knowledge 9 is formed.

10. (Insufficient cooling system pressure (increase the water pressure of the cooling system, over temperature of the pasteurization outlet, too low pasteurization temperature) → (increase the vapour pressure and water pressure)

The new knowledge of 10 is traversal of the existing knowledge 9, and ninth repeat, not as a new knowledge to join the knowledge base.

Continue to traverse the fourth knowledge of C (increase the cooling system water pressure) is the first part of the ninth knowledge, forming a new knowledge 10.

11. (Insufficient cooling system pressure (increase the water pressure of the cooling system, over temperature of the pasteurization outlet, too low pasteurization temperature) → (increase the vapour pressure and water pressure).

The new knowledge of 11 is traversal of the existing knowledge 9, and ninth repeat, not as a new knowledge to join the knowledge base.

The later member of the fifth piece of knowledge C (increase the vapour pressure and water pressure) does not belong to the former member of any knowledge; check the six piece of knowledge.

The later member of the six piece of knowledge C (decrease the vapour pressure and water pressure) does not belong to the former member of any knowledge; check the Seventh piece of knowledge.

The later member of the Seventh piece of knowledge C (increase the soft water pressure) does not belong to the former member of any knowledge; check the eighth piece of knowledge.

The later member of the eighth piece of knowledge C (increase the vapour pressure and water pressure) does not belong to the former member of any knowledge; check the ninth piece of knowledge.

The later member of the eighth piece of knowledge C (increase the vapour pressure and water pressure) does not belong to the former member of any knowledge, no knowledge entry, end.

From the above given, to get new knowledge 9 (Insufficient cooling system pressure, increase the water pressure of the cooling system, over temperature of pasteurization outlet, too low pasteurization temperature) → (increase the vapour pressure and water pressure).

7. Conclusion

Based on the relevant theory of regular expression, the knowledge representation method of online assistant decision making for the quality safety of food processing is realized. Based on the representation method, the prototype of the knowledge object in the field of food processing quality security is expressed. Then, the intelligent reasoning method and the related personnel of the intelligent reasoning machine are studied, and the new knowledge is added to the knowledge base, so as to realize the quality and safety of food processing. The achievement of this research is derived from the National Technical Programs of Research and Demonstration of online safety and quality monitoring and control technique during food processing (2012BAD29B04).

References

1. Meng, X. Model and Application of Mining Accident Warning Knowledge Base Based on Ontology. PhD Dissertation, China University of Mining.
2. Feng, S. Establishment and Research of Car Ontology Knowledge Base Oriented to Opinion Digging. Master Dissertation, Shanxi University.
3. Zhong, X., Z. Liu, P. Ding. Construction and Application of Knowledge Basing on Mix Reasoning. – Chinese Journal of Computers, Vol. **35**, 2012, No 4, pp. 762-766.
4. Xu, W., S. Liu. Logical Basis of Knowledge Base System. – Chinese Journal of Computers, Vol. **32**, 2009, No 11, pp. 2124-2128.
5. Song, L., D. Tang, L. Chen, M. Mao. Establishment of Knowledge Base of Expert System for Aeroengine Oil Spectrum. – Journal of Aviation, Vol. **21**, 2000, No 5, pp. 454-457.
6. Guo, S. Research of Institutional Repositories Based on Open Access. – Master Dissertation, Northeast Normal University.
7. Liu, X., Q. Sun. Research on Knowledge Base of Expert System for Mine Safety Warning System. – Mining Safety and Environmental Protection, 2008, No 3, pp. 34-36+91.
8. Ma, J. The Mine Water Hazard Early Warning Expert System of. Master Dissertation, Xi'an University of Science and Technology.
9. Ni, Z., Y. Zhang, L. Wang. – Control Engineering Based on Ontology of Combination Forecasting and Early Warning System Model, 2011, No 5, pp. 815-819.
10. Huang, J., Q. Wang. Design. – Computer Engineering and Application Based on MAS, 2003, No 35, pp. 192-194+224.
11. Liu, Y., Z. Liu. Decision Support System for the Study of Xinjiang. – 2007, No 2, pp. 110-113.
12. Fu, G., H. Zhou, J. Guo. – Based on the Internet of Things of Coal Mine Safety Comprehensive Intelligent Warning System for Industrial Automation, 2014, No 4, pp. 99-101.
13. Li, C., Y. Chen, H. Huang. Research on the Knowledge Base of Network Public Opinion Monitoring and Early Warning. – Library and Information Work, 2011, No S2, pp. 262-266.
14. Wu, X. Research on Engineering Failure Knowledge Management and Early Warning. Hua Zhong University of Science and Technology.
15. Crosbie, M., G. Safford. Applying Genetic Programming to Intrusion Detection. – In: Proc. of AAAI Fall Symposium on Genetic Programming, 2003, pp. 1-8.
16. Shafer, G. A Mathematical Theory of Evidence. Princeton, Princeton University Press, 1976, pp. 133-185.
17. Pawlak, Z., A. Skomón. Rudiments of Rough Sets. – Information Sciences, Vol. **177**, 2007, No 1.
18. Jevsnik, M., V. Hlebee, P. Raspor. Consumer's Awareness of Food Safety from Shopping to Eating. – Food Control, Vol. **19**, 2008, No 8, pp. 737-745.
19. Goni, S. M., S. Oddone, J. A. Segura. Prediction of Foods Freezing and Thawing Times: Artificial Neural Networks and Genetic Algorithm Approach. – Journal of Food Engineering, Vol. **84**, 2008, No 1.