Wine Quality Assessment under the Eindhoven Classification Method

Ana Pereira, Ana Crespo, Ana Fernandes
Departamento de Química, Escola de Ciências e Tecnologia, Universidade de Évora, Évora, Portugal
{anafcmpereira; anacfcrespo; anavilafernandes}@gmail.com

Inês Aranha
Herdade do Esporão
Reguengos de Monsaraz, Portugal
ines.aranha@esporao.com

Margarida Figueiredo
Departamento de Química,
Escola de Ciências e Tecnologia
Centro de Investigação em Educação e Psicologia
Universidade de Évora, Évora, Portugal
mtf@uevora.pt

Jorge Ribeiro
Escola Superior de Tecnologia e Gestão,
ARC4DigiT – Applied Research Center for Digital Transformation
Instituto Politécnico de Viana do Castelo, Portugal
jribeiro@estg.ipvc.pt

Humberto Chaves
Escola Superior Agrária de Beja
Instituto Politécnico de Beja
Beja, Portugal
hc@ipbeja.pt

José Neves
Centro ALGORITMI
Universidade do Minho
Braga, Portugal
jneves@uminho.pt

Henrique Vicente
Departamento de Química, Escola de Ciências e Tecnologia,
Centro de Química de Évora, Universidade de Évora, Évora, Portugal
Centro ALGORITMI, Universidade do Minho, Braga, Portugal
hvicente@uevora.pt

KEYWORDS
Eindhoven Classification Method, Logic Programming, Knowledge Representation and Reasoning, Entropy, Wine Quality.

ABSTRACT

The identification, classification and recording of events leading to deterioration of wine quality is essential for developing appropriate strategies to avoid them. This work introduces an adverse event reporting and learning system that can help preventing hazards and ensure the quality of the wines. The *Eindhoven Classification Method* (*ECM*) has been extended and adapted to the incidents of the wine industry. *Logic Programming* (*LP*) was used for *Knowledge Representation and Reasoning* (*KRR*) in order to model the universe of discourse, even in the presence of incomplete data, information or knowledge. On the other hand, the evolutionary process of the body of knowledge is to be understood as a process of energy devaluation, enabling the automatic extraction of knowledge and the generation of reports to identify the most relevant causes of errors that can lead to a poor wine quality. In addition, the answers to the problem are object of formal evidence through theorem proving.

INTRODUCTION

The wine industry may use the philosophy of lean thinking to minimize and eliminate waste and errors in order to create value (Chong-Fong 2015, George 2003). In order to accomplish these goals, its critical functions should be monitored, with a focus on the quality of the final product. Indeed, the wine sector is very complex and diverse, requiring a variety of operations, people, processes, equipment and structures in which a variety of adverse events may occur. Undeniably, an unwanted event may be described as a failure to perform a specific action or to use a wrong plan to attain a particular goal. The most efficient strategy to prevent adverse events is recognizing their causes. Such causes may be related to practical problems, human relationships, company policies, action plans, products, strategies or leadership.

People continue learning from their own mistakes and not from their successes. However, they do not like to share their errors or what they have learned with them. As a result, similar blunders may occur repeatedly and wine quality may be affected by avoidable faults. Some studies argue that reporting can be an achievable solution to this problem (Mushtaq et al 2018, van der Schaaf 1995, Vicente et al 2015, World Alliance for Patient Safety 2005), where the basic idea is based on an experience-based learning process. It must be stressed, however, that registering errors is not enough to guarantee the wine`s quality. In fact, collecting data is not enough to improve the practice.

To make the difference, it is important to conduct the technical review of the data in order to identify trends and patterns (Mushtaq et al 2018, Vicente et al. 2015), where the combination of reporting systems and machine learning methods for problem solving may be an answer to the

Communications of the ECMS, Volume 33, Issue 1,
Proceedings, ©ECMS Mauro Iacono, Francesco Palmieri,
Marco Gribaudo, Massimo Ficco (Editors)
ISBN: 978-3-937436-65-4/978-3-937436-66-1(CD) ISSN 2522-2414

problem. Under the present approach to solve the problem it is assumed that humans are fallible and that errors are to be predictable to occur in any organization. It focusses on the conditions under which individuals work and attempt to build defenses to avert errors or to mitigate their effects (Reason 2000). If compared with similar systems its advantages rely mainly on the fact that the approach followed here to knowledge representation is set in a continuous mode (i.e., it is given in the form of energy transfer operations as it will be shown below), therefore allowing for the handling of qualitative and quantitative data or knowledge, being it either incomplete, self-contradictory, or even error sensitive. It is a learning system that enables data analysis, ensures continuous improvement of wine quality and ultimately contributes to consumer satisfaction, being also object of formal proof (Neves 1984, Neves et al. 2007), something that similar methods or methodologies for problem solving used in the wine sector do not contemplate.

COMPUTATIONAL MODEL

To avoid the incidence of adverse events, the understanding of its main causes is essential. Thereby, when developing a framework that can be used in the wine industry the focus should be on methods that use analytical technics in the description of the adverse occurrences, to look at their main causes and the assessment of the attainment of the preventive actions implemented. Considering the concerns about the problem just referred to above, the *ECM* was selected; it uses *Root Cause Analysis* (*RCA*) that allows for the classification of the main causes according to pre-defined codes (van der Schaaf 1995).

The *ECM* enclose two types of errors, namely the *active* and the *latent* ones. The actives refer to human error and are considered at three levels of behavior, (i.e., skills, rules and knowledge), which are in accordance with the *Rasmussen SRK Model* (Rasmussen 1976). The latent ones, in turn, contemplate the technical and organizational errors (van der Schaaf 1995). The former ones arise from problems associated with physical components such as equipment or devices. The subsequent are due to mistakes related to knowledge transfer, procedures or protocols.

Recognizing the causes of a particular item is the first step in developing an *ECM-based* system. To achieve this goal, *Causal Trees* (*CTs*) were considered and *RCA* techniques applied (Figure 1). The *CTs* provide a global picture of the problem through a hierarchical structure and enable the implementation of useful and long-term solutions. For example, the unwanted *event A* (Figure 1) is due to three possible causes. It is known that cause 2's contribution to the adverse event is *high* (known value), while the contribution of causes 1 and 3 is *unknown*, which sets two different types of *null* or *unknown* values. With respect to cause 1, it is not possible to enforce the value to be considered, but it is known that it can only take two values (*low*/*medium*), i.e., an *unknown* value in a finite set of values. With regard to cause 3, it is not possible to be clear about its contribution to the adverse event, all values are plausible, i.e., an *unknown* value (not necessarily from a finite set of values).

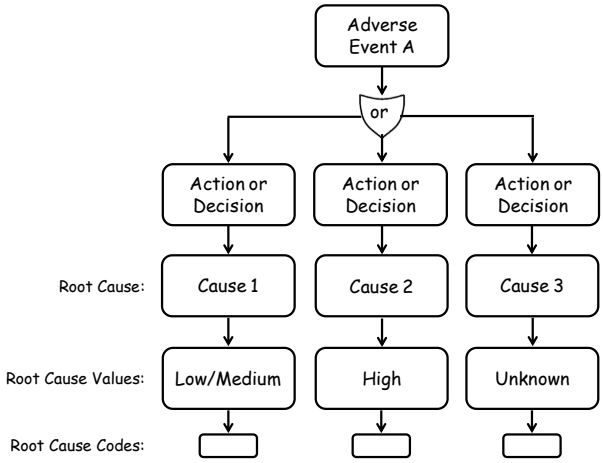

Figure 1: General structure of the Causal Tree for the adverse event A

Knowledge Representation and Reasoning

In this work *Knowledge Representation and Reasoning* (*KRR*) practices will be understood as a process of energy devaluation (Wenterodt and Herwig 2014). Indeed, the predicates' extensions that elicit the universe of discourse will be given as productions of the type (Pereira and Anh 2009), viz.

$$
\begin{aligned}
&\{\\
&\quad \neg\, p \leftarrow not\ p, not\ exception_p\\
&\quad p \leftarrow p_1, \cdots, p_n, not\ q_1, \cdots, not\ q_m\\
&\quad ?\,(p_1, \cdots, p_n, not\ q_1, \cdots, not\ q_m)\ \ (n, m \geq 0)\\
&\quad exception_{p_1},\ \cdots\ , exception_{p_j}\ \ (0 \leq j \leq k),\\
&\qquad\qquad\qquad\qquad being\ k\ an\ integer\ number\\
&\}
\end{aligned}
$$

where *n* and *m* stand for the cardinality of the predicates' set and predicate's arguments, respectively. "?" denotes *falsity*. The other symbols stand for themselves. In order to make the process comprehensible, it will be presented in a graphical form. Taking as an example a group of 3 (three) causes fixed as an *Adverse_Wine_Assessment Questionnaire-Three-Item (AWA – 3)*, viz.

- *Cause 1* – A *vertical tasting* involves wines from the same year but from different vineyards or winemakers;
- *Cause 2 – Bottle Stink* does not necessarily mean a spoiled bottle of wine, and;
- *Cause 3 – Reducing* the grape crop usually results in wines with lowers levels of alcohol.

designed to assess the workers' literacy level in the wine sector, varying on the interval 0…1, on the assumption that low *stakes* will trigger positive outcomes and benefits their corporations. In order to accomplish this goal, it will be used the scale, viz.

Very low (1), Low (2), Medium (3), High (4)

which leads to (Figure 2), viz.

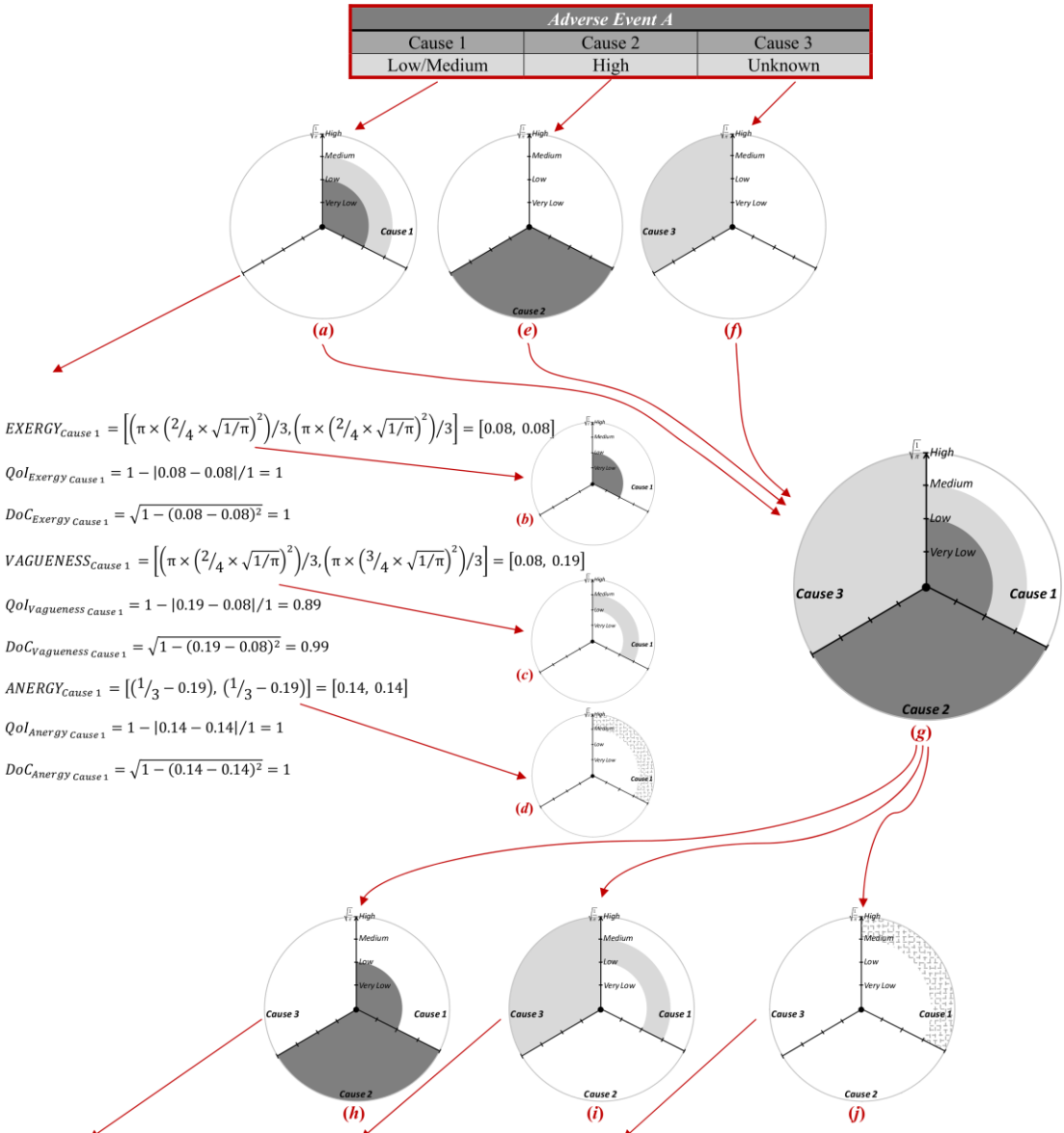

Figure 2: Going from a *Qualitative* setting to a *Quantitative* one

Indeed, aiming to the quantification of the qualitative information presented in the *CT* (Figure 1) and in a way to make the process intelligible, it was given in a graphical form (Figure 2). Once the contribution of cause 1 for *adverse event A* was *low/medium*, the correspondent numeric value is given by the colored area ranging between (Figure 2(a)), viz.

$$\left[\left(\pi \times \left(2/4 \times \sqrt{1/\pi}\right)^2/3,\ \pi \times \left(3/4 \times \sqrt{1/\pi}\right)^2/3\right)\right]$$

In terms of the energy's transfer operations, *exergy* for cause 1 $(exergy_{cause_1})$ corresponds to the dark colored area (Figure 2(b)), while *vagueness* $(vagueness_{cause_1})$ is given by the gray colored area (Figure 2(c)). Finally, *anergy* $(anergy_{cause_1})$ corresponds to the dashed area (Figure 2(d)). The contribution of cause 2 to the *adverse event A* is *high*, and the correspondent area is (Figure 2(e)), viz.

$$\pi \times \left({}^{4}/_{4} \times \sqrt{1/\pi}\right)^{2}/3$$

In this case $exergy_{cause_2}$ is given by the dark colored area, i.e., [0.33, 0.33] while $vagueness_{cause_2}$ and $anergy_{cause_2}$ are 0 (zero). Finally, the contribution of cause 3 is *unknown*, all the possibilities should be considered and the corresponding area is in the range (Figure 2(f)), viz.

$$\left[\left(\pi \times \left({}^{0}/_{4} \times \sqrt{1/\pi}\right)^{2}/3,\ \pi \times \left({}^{4}/_{4} \times \sqrt{1/\pi}\right)^{2}/3\right)\right]$$

Once the energy values that have been transferred and consumed are unknown, the $exergy_{cause_3}$ and $anergy_{cause_3}$ are 0 (zero), while $vagueness_{cause_3}$ is given by gray colored area, ranging between [0, 0.33]. The global view of the *adverse event A* is given in Figure 2(g) and the global values of *exergy*, *vagueness* and *anergy* are given by the areas shown in the Figure 2(h), (i) and (j), respectively. The adverse *event A* may now be set as the predicate *adverse_event_a, and given* in the form, viz.

$$adverse\_event\_a : \boldsymbol{EX}ergy,\ \boldsymbol{VA}gueness,\ \boldsymbol{AN}ergy,\ \boldsymbol{Q}uality\text{-}\boldsymbol{of}\text{-}\boldsymbol{I}nformation, \boldsymbol{D}egree\text{-}\boldsymbol{of}\text{-}\boldsymbol{C}onfidence \rightarrow \{True, False\}$$

where the variables ***EX**ergy*, ***VA**gueness* and ***AN**ergy* denote the *entropic states* or *sustainability* factors of the terms or clauses that make the logic program, whose extension is given below, viz.

$$\begin{aligned}
&\{\\
&\quad \neg\, adverse_event_a\,(EX,\ VA,\ AN,\ QoI,\ DoC), \leftarrow not\ adverse_event_a(EX,\ VA,\ AN,\ QoI,\ DoC),\\
&\qquad\qquad not\ exception_{adverse_event_a}(EX,\ VA,\ AN,\ QoI,\ DoC)\\
&\quad adverse_event_a\,(0.42,\ 0.44,\ 0.14,\ 0.85,\ 0.97).\\
&\}
\end{aligned}$$

The arguments *quality-of-information* (*QoI*) and *degree-of-confidence* (*DoC*) stand for themselves, however its evaluation may be found in Vicente et al. (2018).

CASE STUDY

To adjust the *ECM* to the *Wine Industry* occurrences, a new version of the model was conceived with extensions and adaptations for the sector. Furthermore, the *CTs* for the classification of the adverse events' root causes were drew. Such extensions and adaptations made possible to fit each category of the wine industry streamlining the classification process. The flow diagram of the classification process is portrayed in Figure 3, as well as the codes to categorize each adverse event (Vicente et al. 2015). Taking into consideration the adverse events classified as *"Human behavior – Knowledge-based errors"* (code *HKK*), they can arise from difficulties in execution, interpretation or reporting procedures. *Chemical analysis badly performed*, *chemical analysis unfinished* or *chemical analysis not validated* are examples of adverse events that falls into this class.

The *CT* regarding the adverse event *wine fault* is shown in Figure 4. Considering that the adverse event under consideration may occur due to various causes that should be taken simultaneously or separately, *AND/OR-nodes* are used to include such features in the *CT*. In addition, the *unknown* and *forbidden* operators were used to describe events for which the event's causes are *unknown/forbidden/not allowed* (e.g., due to internal policies). Thereby, based on the information presented in Figure 4, it is possible to identify all feasible situations, viz.

(i) Who made the registration of the occurrence report *wine fault* due to the *presence of **D**ekkera/**B**rettanomyces yeasts* (*D/B*), i.e., a *known value*;

(ii) The professional who recorded the adverse event only recorded *wine fault* due to *organoleptic changes*. It is not possible to be constructive about the origin of the adverse event, but it is known that it can only be the occurrence of ***D**ekkera/**B**rettanomyces yeasts* (*D/B*), existence of ***T**ri**C**hloro**A**nisole* (*TCA*) or ***OX**idation* of the wine (*OX*) for which the respective values are *very low/low*, *medium/high* and *low/medium*. This case corresponds to an *unknown value from a finite set of values*; and

(iii) It was only registered *wine fault*. All hypotheses are admissible, corresponding to an ***U**nknown* or a ***F**orbidden **V**alue* (*UFV*).

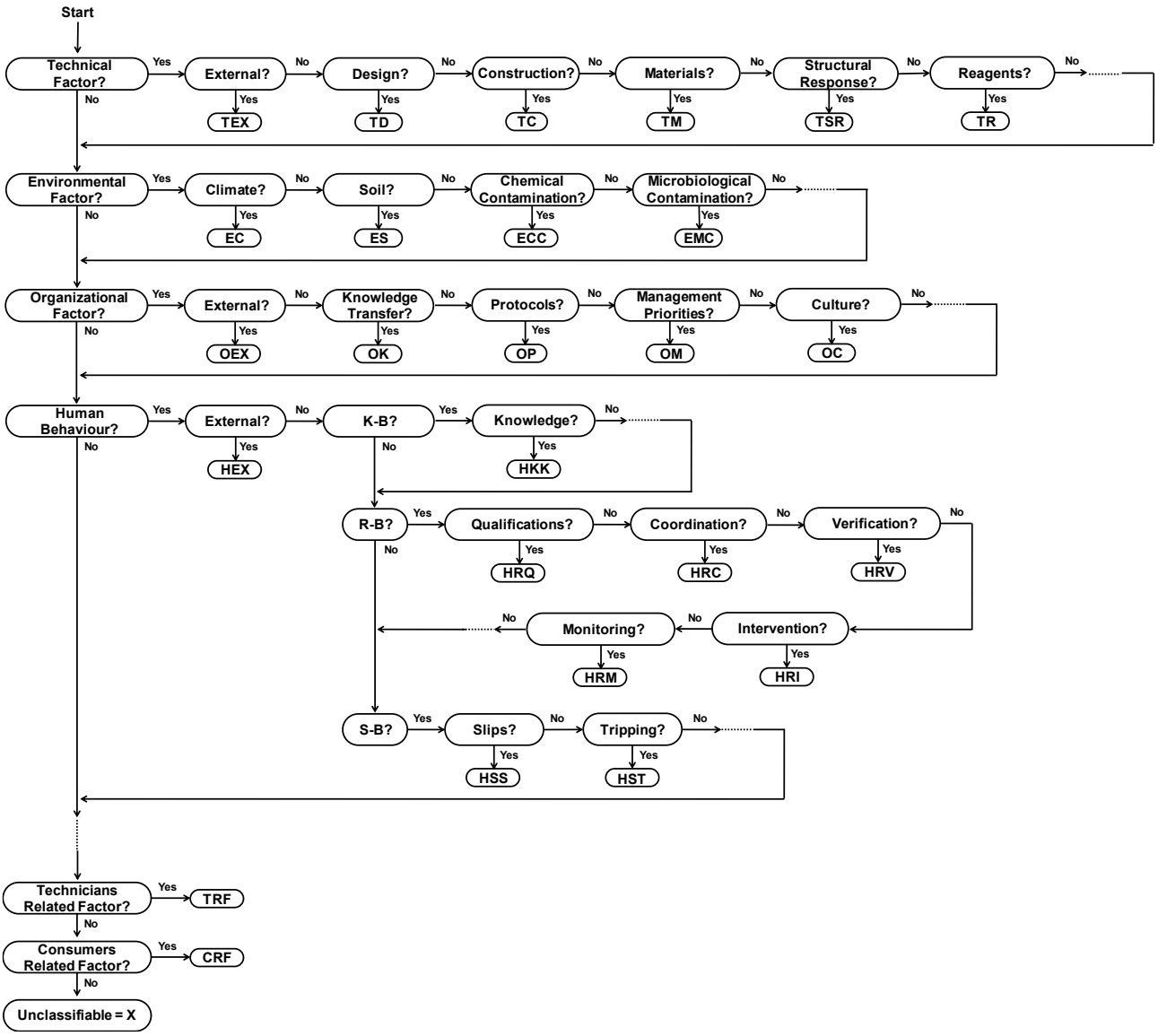

Figure 3: Flow chart of the *Eindhoven Classification Model* for the *Wine Industry*

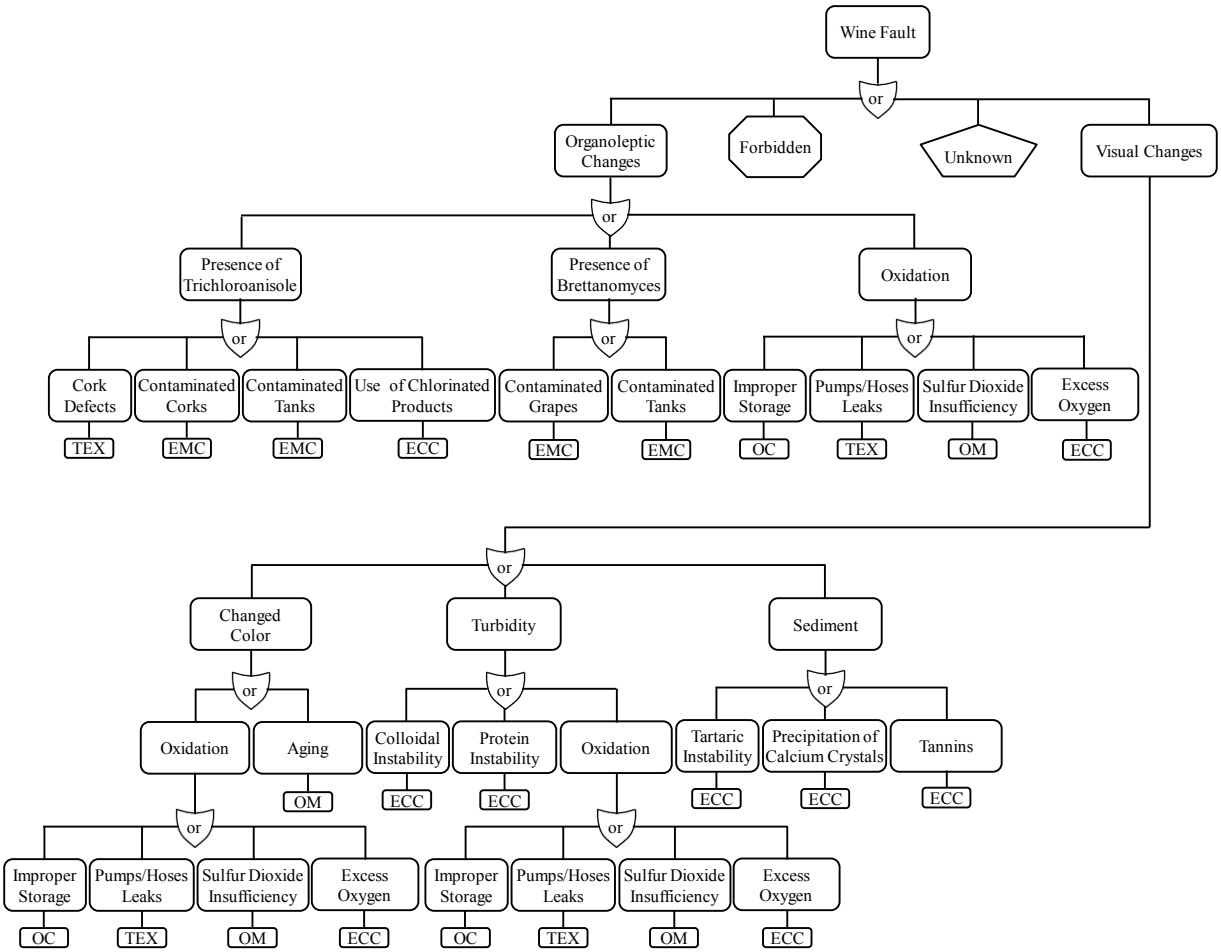

Figure 4: The adverse event *wine fault* in terms of an *Extended Causal Tree*

The logic program epitomized below, built in terms of the extents of predicates *action_or_decision_a*, *action_or_decision_b*, *action_or_decision_c* stands for a formal description of the situations *(i)*, *(ii)* and *(iii)* referred to above.

$\{$

$\{$

$$\neg\, action\_or\_decision\_a\ (EX,\ VA,\ AN,\ QoI,\ DoC), \leftarrow not\ action\_or\_decision\_a\ (EX,\ VA,\ AN,\ QoI,\ DoC),\ not\ exception_{action\_or\_decision\_a}\ (EX,\ VA,\ AN,\ QoI,\ DoC)$$

$$action\_or\_decision\_a\ (1,\ 0,\ 0,\ 1,\ 1).$$

$\}$

$\{$

$$\neg\, action\_or\_decision\_b\ (EX,\ VA,\ AN,\ QoI,\ DoC), \leftarrow not\ action\_or\_decision\_b\ (EX,\ VA,\ AN,\ QoI,\ DoC),\ not\ exception_{action\_or\_decision\_b}\ (EX,\ VA,\ AN,\ QoI,\ DoC)$$

$$action\_or\_decision\_b\ (0.29,\ 0.31,\ 0.40,\ 0.78,\ 0.89).$$

$\}$

$\{$

$$\neg\, action\_or\_decision\_c\ (EX,\ VA,\ AN,\ QoI,\ DoC), \leftarrow not\ action\_or\_decision\_c\ (EX,\ VA,\ AN,\ QoI,\ DoC),\ not\ exception_{action\_or\_decision\_c}\ (EX,\ VA,\ AN,\ QoI,\ DoC)$$

$$action\_or\_decision\_c\ (0,\ 0.33,\ 0,\ 0.67,\ 0.89).$$

$\}$

$\}$

that stands for an adverse event reporting and learning computational system. The *Adverse Event Reporting Forms for Wine Industry* (*AERF-WI*) is an interface web for adverse event registration. The registration can be done by professionals and/or by the consumers, through pre-defined forms conceived to each user profile.

CONCLUSIONS

This study presents an intelligent system enabling to deal with the problem of *KRR* under a qualitative and quantitative approach to incomplete, unknown, or even self-contradictory data, information or knowledge. It was shown how the fields of *Computer Science* and *Mathematical Logic* may be used to promote excellence in very dynamics and uncertain environments like the *Wine Industry*. This system offers some advantages, like simpler, faster, and more reliable analysis of adverse events. This information may be helpful to identify trends and areas for improvement. Furthermore, its formal description provides a path for knowledge acquisition, identification of the adverse events' main causes, and may inspire changes in wine industry procedures. Another advantage relies on its modularity, i.e., it offers the possibility of adding new categories and/or sub-categories at any time, without changing the structure or the working mode of the system.

Future work includes the development of the *Adverse Events Manager Reports for the Wine Industry* (*AEMR-WI*) module. Such component aims at the analysis of the adverse events recorded by *AERF-WI*. The *AEMR-WI* will provide automatic reports of the adverse events, supplemented with charts and statistical information about the events recorded. Finally, the *Adverse Events Knowledge Manager for Wine Industry* (*AEKM-WI*) module uses the data from the system database in order to identify trends, using data mining tools, a path to *Data Science*.

ACKNOWLEDGMENTS

This work has been supported by FCT – Fundação para a Ciência e Tecnologia within the Project Scope: UID/CEC/00319/2019

REFERENCES

Chong-Fong, Y. 2015. *Level of Service Quality of SME Laboratory Services on Customer Satisfaction*. MBA Thesis, University Sains Malaysia, Penang.

George, M. 2003. *Lean Six Sigma for Service: How to Use Lean Speed and Six Sigma Quality to Improve Services and Transactions*. McGraw-Hill, New York.

Mushtaq, F., C. O'Driscoll, F.C.T. Smith, D. Wilkins, N. Kapur and R. Lawton. 2018. "Contributory factors in surgical incidents as delineated by a confidential reporting system." *Annals of the Royal College of Surgeons of England* 100, 401-405.

Neves, J. 1984. "A logic interpreter to handle time and negation in logic databases". In *Proceedings of the 1984 annual conference of the ACM on The Fifth Generation Challenge*, R.L. Muller and J.J. Pottmyer (Eds.). ACM, New York, 50-54.

Neves, J., H. Vicente, M. Esteves, F. Ferraz, A. Abelha, J. Machado, J. Machado, J. Neves, J. Ribeiro and L. Sampaio. 2018. "A Deep-Big Data Approach to Health Care in the AI Age." *Mobile Networks and Applications* 23, 1123-1128.

Neves, J., J. Machado, C. Analide, A. Abelha and L. Brito. 2007. "The halt condition in genetic programming". In *Progress in Artificial Intelligence*, J. Neves, M.F. Santos and J. Machado (Eds.). Lecture Notes in Artificial Intelligence, Vol. 4874, Springer, Berlin, 160-169.

Pereira, L. and H. Anh. 2009. "Evolution prospection". In *New Advances in Intelligent Decision Technologies – Results of the First KES International Symposium IDT 2009*, K. Nakamatsu, G. Phillips-Wren, L. Jain and R. Howlett (Eds.). Studies in Computational Intelligence, Vol. 199, Springer, Berlin, 51-64.

Rasmussen, J. 1976. "Outlines of a hybrid model of the process plant operator". In *Monitoring Behavior and Supervisory Control*, T.B. Sheridan and G. Johannsen, G. (Eds.). Vol. 1, NATO Conference Series, Plenum Press, London, 371-383.

Reason, J.T. 2000. "Human error: models and management." *British Medical Journal* 320, 768-770.

van der Schaaf, T.W. 1995. "Near miss reporting in the chemical process industry: an overview." *Microelectronics Reliability* 35, 1233-1243.

Vicente, H., F. Borralho, C. Couto, G. Gomes, V. Alves and J. Neves. 2015. "An Adverse Event Reporting and Learning System for Water Sector Based on an Extension of the Eindhoven Classification Model." *Water Resources Management* 29, 4927-4943.

Vicente, H., M. Figueiredo, A. Dias, J. Marques, I. Araújo, N. Maia, J. Ribeiro and J. Neves. 2018. "Improving the Perception of Chemistry in Higher Education Programs through Many-Valued Empirical Machines." *Frontiers in Artificial Intelligence and Applications* 309, 563-571.

Wenterodt, T. and H. Herwig. 2014. "The Entropic Potential Concept: A New Way to Look at Energy Transfer Operations." *Entropy* 16, 2071-2084.

World Alliance for Patient Safety. 2005. "WHO draft guidelines for adverse event reporting and learning systems: from information to action". WHO Document Production Services, Geneve.