Contextual Knowledge Based System: A study and design in enology

AGABRA J. (1, 2), ALVAREZ I. (1, 2), BREZILLON P. (2)

(1) CEMAGREF
LISC
BP 121
92185 Antony Cedex, France
Tel: +33 1 4096 6235 Fax: +33 14096 6080
E-mail: {agabra, ia} @cemagref.fr

(2) LAFORIA, Box 169
University Paris VI
4, place Jussieu
75252 PARIS Cedex 05, France
Tel: +33 1 4427 7008 Fax: +33 1 4427 7000
E-mail: {ia, brezil}@laforia.ibp.fr

Abstract:
Wine making is a domain where formalization is very difficult, mainly because one has to deal with a huge number of heterogeneous pieces of knowledge that intervene at different steps of the wine-making process. We address the problem of stops in the alcoholic fermentation. This step of the wine-making process is strongly correlated with knowledge of the events that occur at previous steps. This knowledge of the previous steps intervenes in the wine-making process as contextual knowledge. In this paper, we present the results of our modeling of contextual knowledge and the design of a context-based system. In our application, we use the onion metaphor to model contextual knowledge. The problem to solve is the heart, and contextual knowledge is organized in layers around the heart in an order that relies on a qualitative distance.

Keywords:
Context, knowledge representation, contextual knowledge based system, wine making, alcoholic fermentation

1.INTRODUCTION

Expert knowledge holds an important place in the wine-making process. One reason is that wine making embraces a great deal of complex biological processes (vine-growing, grape ripening, alcoholic fermentation, etc.). Decision support systems can play a crucial role in helping the different actors intervening all along this process, from the vine-growing to the wine-making, mainly in the problem of stuck fermentations.

One of the most critical steps in wine making is the fermentation itself and some difficulties in its achievement may arise. Whatever the causes are, the consequence is the same: the fermentation process, in which sugar is transformed into alcohol, slows down (sometimes very much) or even stops completely, leaving too much sugar in the wine. These events are respectively called sluggish and stuck fermentations. They can result in wine spoilage because there is a severe control on wine quality (e.g., for quality reasons, the final sugar concentration in the wine must be lower than 2g/l, except for sweet wines).

The problem of stuck fermentations arises even in famous vineyards. The reason is that the different vine varieties are vinified separately. Conversely, in the more frequent case of blended fermentations, the good fermentativeness (fermentative qualities) of one vine variety can compensate for the faults of another. The consequences are:
- A great waste of money.
- A lack of room in vats. A fermentation that lasts one month instead of ten days implies that one or more vats cannot be used for another fermentation or for storage.

Proc. of the International and Interdisciplinary Conference on MOdeling and Using Context (CONTEXT-97), Federal University of Rio de Janeiro Ed., 1997, pp. 351-362.

- Critical time limits. For example, there are some wines that have to be put on sale at a particular period, and any delay can be financially prejudicial.
- Eventual noxious effects on the organoleptic characteristics of the resulting wine when a stuck fermentation is artificially restarted by addition of chemical activators.

Our work aims to help wine makers to predict a stop in the normal fermentation process as soon as possible. This is not a new preoccupation. Several approaches attempted to use either neural networks or knowledge-based systems techniques [Insa et al., 1995]. These attempts failed to give a satisfactory solution mainly because they had a too partial view of the problem. They did not take into account the fact that the vine-growing is not directly related to the alcoholic-fermentation, even if knowledge of the former step is used by the wine maker in the latter step. We call such knowledge contextual knowledge and we aim at introducing a Contextual Knowledge-Based System.

This paper describes the results obtained in the design and development of this Contextual Knowledge-Based System. This work comes within the COSINUS project, aiming at addressing the problems of the enologists. The goal of the COSINUS project is to develop a system that associates two subsystems, namely a neural network and a knowledge-based system. This association would make possible to complement the failures of one system by the successes of the other.

This paper is organized in four parts. We present in Sections 2 and 3 respectively the vine field and the human actors that intervene in this field. Then we describe what context is in our model (Section 4) and then discuss the consequences of introducing context in our application, as far as knowledge representation is concerned (Section 5.)

2.DOMAINPRESENTATION

Wine is the result of the transformation of grape sugar into alcohol, a process which is called alcoholic fermentation. The whole wine-making process is divided into six main steps:
- the vine-growing, which allows the production of grapes to be as good as possible,
- the harvest, when grape clusters are picked and crushed to release the juice called "must,"
- the prefermentative treatments, which consist of clarification (clearing of deposit), and adding of different products: to help the clarification (bentonite), to prevent oxidization (sulfur dioxide) or even to activate the coming fermentation (vitamines),
- the fermentation process, which is also called alcoholic fermentation, sometimes followed by a malolactic fermentation,
- the wine fining, whose goal is to remove the last particles in suspension in the wine,
- the bottling and commercialization.

We can easily imagine the complexity of the wine-making process since it involves natural (biological, chemical, biochemical) and industrial processes, which are tightly coupled. The consequences are long paths of causes and effects, with little available and totally reliable data, and few complete biochemical descriptions.

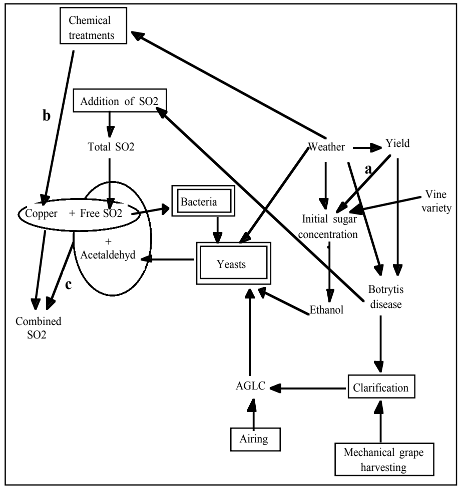

Figure 1: Part of the deep model

Figure 1 represents only a little part of the whole process and points out the existence of different kinds of links. For example:
- The link **a** is a statistical link, that has been established by experiment
- The link **b** is a qualitative link
- The link **c** is a chemical link, that represents the combination of free sulfur dioxide with acetaldehyde.

The rectangular boxes in Figure 1 stand for human actions during the different steps of the wine-making process. The double frame around "yeast" and "bacteria" means that very complex reactions take place inside these microorganisms. These reactions are not all known. We just can say that there are chemical reactions both in the cells' membrane and inside the cells.

The fermentation process itself first consists in the conversion of grape sugar into alcohol, either by natural yeast found on grape skins, or by commercial yeast strains added just before the fermentation begins. Many other chemical reactions lead to secondary products, which may interact with the yeast or the chemical elements added to the must. The equilibrium of these chemical reactions may be broken and lead to pathological situations. For example, the alcoholic fermentation produces acetaldehyde. In certain conditions, this acetaldehyde may combine with the sulfur dioxide previously added, making it inefficient to control the bacteria population. Another example is the ethanol, the first product of the fermentation, which may damage the yeast's membrane above a certain threshold.

Besides this complexity, due to both the chemical and biochemical mechanisms of the fermentation, many parameters (taken as initial conditions for the fermentation) are the result of other complex processes, involving many other parameters from different origins. For example, the initial sugar concentration depends on many variables, some of them involving human actions (vine shoot pruning, date of the harvest), others related to the natural conditions (rain, temperature, climate). As shown in Figure 1, there are other parameters, such as the vine variety or the kind of soil, that act as external constraints on parts of the whole process

3. ACTORS

Different categories of human actors work on wine: commercial actors such as wine merchants; some political actors such as the French Union of Enologists (UŒF) or the French National Institute of Appellation of Origin (INAO); some technical actors (vine-growers, enologists, microbiologists, etc.). Let us focus on those who intervene along the wine-making process. Figure 2 shows two different aspects: the part of the whole process people know, and the steps at which they are supposed to act.

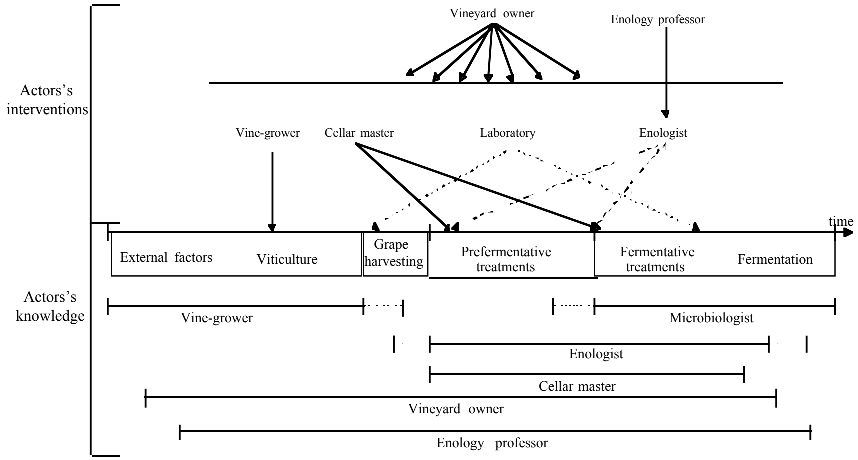

Figure 2: Diagram of experts' knowledge and interventions.
(The horizontal dotted lines in the actors' knowledge part of Figure 2 stand for limited knowledge.)

Chronologically, the actors are the following:
- The vine-grower is specialized in viticulture and has a global view of all that has been done on the grapes from the beginning. He tries to answer questions like: When to do the chemical treatments in order not to leave chemical residues on the grapes, according to weather and the vegetative cycle? What kind of tying or vine shoot pruning may lead to higher sugar concentration in grapes?
- The cellar master is in charge of nearly everything after the harvest. He may have to follow eventual instructions from the vineyard owner.
- The enologist is involved in the same steps as the cellar master. He may either supervise the prefermentative and fermentative operations, or come to the cellar to give advice when needed (about the products to add for increasing the fermentation rate, about the kind of prefermentative technology to use in order to have a good quality must, etc.)
- The microbiologist is interested in the biological mechanisms taking place in the yeast.
- The enology teacher teaches the theoretical knowledge of enologists.
- A laboratory may also have a specific role by making a chemical analysis of the grapes before harvest, or of the must during the alcoholic fermentation.

There are many other people involved as the numerous seasonal workers.

Figure 2 shows that there are several kinds of expert knowledge, which are sometimes completely separate. Therefore, each expert has only a partial view of the whole wine-making process. This means that when the expert of one sub-domain (e.g., the microbiologist) addresses the problem of stuck fermentation, he does not consider explicitly the weather at harvest time, in his analysis (for example).

4. CONTEXT IN OUR MODEL

4.1. Limits of classical approaches

Most of the research made in viticulture and enology has concentrated on one or more factors and tried to explain the relation between these factors and the fermentation process, in particular the problem of stuck fermentations. For example, Maigre et al. (1995) in Switzerland have been working since 1978 on the influence of green cover crops and nitrate fertilization on the alcoholic fermentation and the final quality of the wine produced. In the biochemical field, many researchers (from South-Africa, Australia, United-States...) have tried to show either the importance of the clarification technique, or the must's nitrogen concentration, or even the must's optanoic and the decanoic acids' concentration.

Four years ago, the French National Institute of Agronomical Research (INRA) of Montpellier began to study in Languedoc-Roussillon the variability of musts' chemical composition, especially the fermentation kinetics to see whether it was related to the problem of stuck fermentations or not [Sablayrolles, 1995]. This study has triggered different approaches: kinetics curves' interpretation [Grenier and Sablayrolles, 1990], expert system development [Grenier, Feuilloley, Sablayrolles, 1988], classification and learning techniques in neural networks [Insa G. et al., 1995]. The objective was to predict a stuck fermentation sufficiently in advance to make corrections possible. All these approaches failed to give interesting enough results, but they permitted at least to reveal the limits of using fermentation data only [Insa G. et al., 1995]**.**

Two other studies have tried, on the basis of the kinetics data, to predict the kinetic curves from additional information (initial sugar concentration, nitrogen concentration, etc.) [Marin R., 1995]. If the prediction was good at the beginning of the kinetics, it was inaccurate for the end of the curve, where the sluggish and stuck fermentations differ from good ones [Alvarez, 1995]. The author has tried to use symbolic rules extracted from the available cases to improve the results. However the extracted knowledge was too atomized to be used by a human expert.

4.2. What is contextual knowledge for the fermentation process?

Part of knowledge is explicitly used in the fermentation process (e.g., the temperature of the must during the fermentation). Another part of the knowledge intervenes in an implicit way (e.g., the weather at harvest time). We call this latter type of knowledge **contextual knowledge**. The use of contextual knowledge enables to account for all the possible causes for the stuck fermentation, and thus provide a correct explanation of the problem.

All the pieces of contextual knowledge are not at the same level, some being more directly related to the problem to solve than others. For instance, the weather during the last few months before harvest is "farther" than the temperature of the vat during the fermentation process. However, both of them participate in the fermentation process. Let us illustrate this aspect by the example of Figure1.

One possible explanation for a stuck fermentation is a conjunction of two causes: an insufficient resistance of the yeast's membrane, and the negative action of ethanol, naturally produced by the fermentation process, which leads to the death of cells. If we consider the insufficient resistance of cells'membrane, we know that it may come from a lack of AGLC (a kind of fatty acid) in the must. If we stop the explanation at this point, we have only used knowledge about the fermentation process. However, we may find another explanation related to the prefermentative techniques. For example, airing of the must may have an influence on this AGLC concentration. There is also another explanation: a too strong clarification often leads to a lack of AGLC in the must. Then we can then try to see in which particular case a too strong clarification may be done. One explanation is a mechanical grape harvesting. This often leads to a must full of « foreign bodies » (part of stalk, leaves, iron, etc.), which requires to be cleared. This explanation has been taken from the viticulture knowledge.

Proc. of the International and Interdisciplinary Conference on MOdeling and Using Context (CONTEXT-97), Federal University of Rio de Janeiro Ed., 1997, pp. 351-362.

This is also true for the effect of ethanol on the yeast. A high ethanol concentration comes from a high initial sugar concentration in the must. Such a high sugar concentration may have several origins: a high yield, or the weather during grape ripening for example. It may also depend on the vine variety itself, which is the most distant factor that we can find.

Thus, it is possible to classify the pieces of contextual knowledge. Tichiner (cited in [Jansen, 1995]) proposes, in another domain, a view very close to our. Tichiner said that the environment around an organism is the basis for the definition of context. Context must be considered outside the actual knowledge as a layer, or a set of layers, around it. Our model may thus be seen as an onion-like model, in which contextual deep knowledge is organized in layers according to an onion metaphor. The outer layer gathers parameters that are contextual to all the causes of stuck fermentations.

4.3. What is contextual knowledge for each actor?

We can represent on a diagram the origin of the knowledge we used to find possible causes for stuck fermentations. Figures 3a, 3b, and 3c give the viewpoints of the microbiologist, the enologist and the vine-grower on the domain knowledge represented in Figure 1.

Microbiologist: Enologist: Vine-grower:

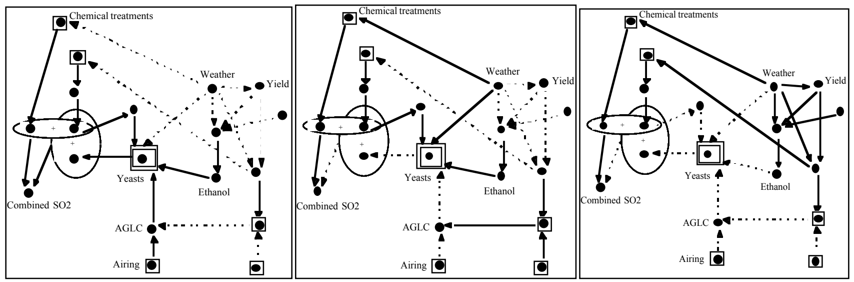

Figure 3a Figure 3b Figure 3c

The knowledge in Figure 1 has been obtained by mixing knowledge coming from all experts (e.g., vine-grower, enologist, and microbiologist in our example). Some pieces of knowledge are shared by different experts and some are not. However, they equally contribute to understand and solve the problem. That is the reason why we have to build a model that enables us to account for all knowledge pieces, the ones concerning directly the fermentation as well as those as faraway as the weather during vine-growing for example.

4.4. What is contextual knowledge for the problem solving?

From the expert knowledge gathered from interviews [Agabra J., 1996], we have implemented a model[1] of the domain (deep knowledge) and we are implementing a set of rules[2] (shallow knowledge). A simulation from the model provides an evaluation of the risk of stuck fermentation of a particular must. Before the simulation, we have a set of data that are a description of a must at the beginning of the fermentation phase or during it. We use all the available information about previous

[1] The implementation has been partially made with METAGEN [Sahraoui et al., 1995], which permits to express the knowledge in the different phases of a software development and produces automatically the code of the resulting system.
[2] In NEOPUS formalism [Pachet F., 1995]

phases (measured chemical parameters, actions already done, etc.). These data naturally include contextual knowledge as described in sections 4.2 and 4.3.

A simulation consists of different macro-steps that occur chronologically, from the vine-growing to the prefermentative treatments, ending with the fermentation itself (as seen in Figure 2). Each step inherits a part of its initial conditions from the former step. The derivative path to these initial conditions constitutes a static and fixed context to the present macro-step. For example, the initial nitrogen concentration is used during the step of the perfermentative treatments. This nitrogen concentration has been derived at the vine-growing step from many parameters, including the weather. The weather is thus contextual for the prefermentative step.

The other input data of a step are a subset of the initial available data. All the data available at the beginning of the simulation are not used at the same time. For example, the parameters concerning the prefermentative techniques (e.g., temperature and duration of the clarification) are available from the beginning of the simulation, but make up data input for the prefermentative step only, and will be thus used only when the simulation will have reached this particular macro-step.

5. DISCUSSION ABOUT CONTEXT

During the design of our system, we addressed three main issues:
- the definition of contextual knowledge for the problem to solve,
- the definition of contextualized knowledge, knowledge that is used for solving the problem, and
- the definition of a distance between pieces of contextual knowledge.

These questions arose at different steps of the system development and in different situations. We consider in this paper the discrimination between contextual and contextualized knowledge from two different viewpoints: the model building and the organization of domain knowledge, and the use of knowledge.

5.1. Knowledge organization and model building

Our objective is to make a prediction of stuck fermentations. Experts do not have all the domain knowledge that is needed to build a model of the stuck fermentations' problem (see sections 3 and 4.2). Deep knowledge is organized in a model used for a simulation purpose to predict a risk of stuck fermentation (see section 4.4).

There is a distinction to make between contextual and contextualized knowledge. Contextualized knowledge is the contextual knowledge that is used to derive a solution, i.e. there is a derivation path from this knowledge to the conclusion as illustrated in Figure 4a. Indeed, there are several layers of context that are stacked in our application, one for each main step of the wine making before the fermentation process (i.e., viticulture and prefermentative treatments).

Proc. of the International and Interdisciplinary Conference on MOdeling and Using Context (CONTEXT-97), Federal University of Rio de Janeiro Ed., 1997, pp. 351-362.

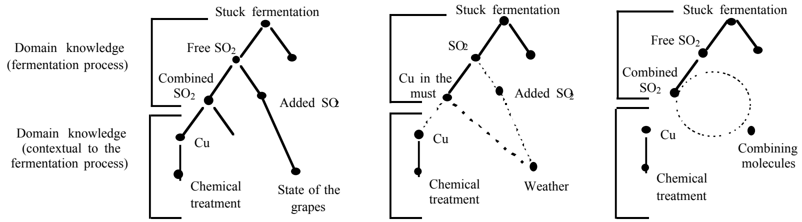

Figure 4a: A known derivation path to stuck fermentation Figure 4b: Missing links Figure 4c: Complex links

Unfortunately, it is not always possible (and almost never in the fermentation case) to derive a path between pieces of contextual knowledge and the goal, mainly because there is a lack of information. This lack of information can correspond to a lack of communication between experts (see Figure 4b) or deep knowledge too complex to be well organized (see Figure 4c).

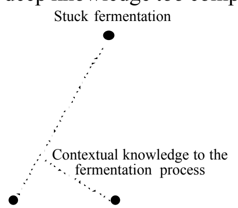

Figure 5: Heuristic knowledge

To overcome the problem, we can first use contextual knowledge as shallow knowledge. The incomplete path is replaced by a global link between known pieces of knowledge and the goal, as shown in Figure 5. For example: some enologists use the following rule : if the vat's temperature is higher than 30°C (for red wines), and if there is no frequent airing, this will be damageable to the yeast.

Second, we can weaken the conclusion to reach. Contextual knowledge thus becomes contextualized knowledge for the solving of a weaker problem.

For example, the complexity of the fermentation process implies that it is impossible to derive or infirm a stop due to the presence of copper in the must. Nevertheless, we know that copper is involved in chemical reactions that have negative effects on the success of the fermentation. Thus, if the desirable conclusion is the risk of stuck fermentation instead of the stop of the fermentation, we can derive a path between copper and the new goal, and the knowledge about copper becomes contextualized with regard to this new conclusion.

Relatively to the model building and knowledge organization issues, context for the problem solving (i.e., the prediction of stuck fermentation) has to be considered at a global level. For the reasons mentioned in previous sections, the domain knowledge is contextual to the problem to solve and cannot be used directly. It is necessary to find a way to *contextualize* this knowledge relatively to a weaker problem (risk of stuck fermentation).

5.2. Use of knowledge in the wine model

Once the knowledge is contextualized globally, it is possible to build a model (concepts at different levels of abstraction, and relations between concepts in each level) and use it for a simulation purpose (as seen in section 4.4). When using knowledge during a simulation, the notions of contextual and contextualized knowledge relatively to the reasoning step are different from the definition given above.

We have seen in section 4.4 that part of the input data is not relevant to the first steps of the simulation, such as the prefermentative treatments; these data and the knowledge related to them are external to the reasoning process until the prefermentative step. We have also seen that some pieces of knowledge are very important (e.g., part of chemical relations that act on the yeast), and can be derived from other (and distant) steps of the simulation. Each step of the simulation inherits a static and fixed context from the previous one: the derivative paths (in previous steps) of all initial conditions to the present step. During the simulation, there is a real process of contextualization (although very simple in our case) as proposed by [Edmondson and Meech, 1993]. For instance, the copper concentration is derived from steps very distant from the fermentation stage (see Figure 1). The knowledge involved in the determination of the copper concentration (chemical treatments, etc.) is clearly contextual to the step at which this result is used (fermentation process). In the model, we have represented this kind of knowledge as a couple (Figure 6). At the beginning of the fermentation step, there is a transfer of value from one element of the couple to the other.

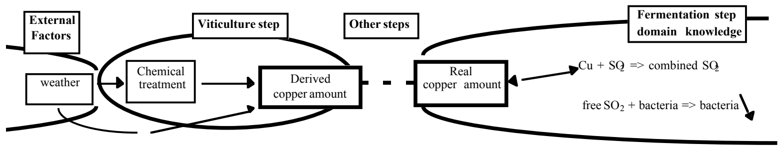

Figure 6: A couple of contextual -- contextualized pieces of knowledge

A measurement of the copper concentration may sometimes be made at the fermentation step (or at the previous step). Our purpose is not to deal with all the possibilities of validation or explanation that the comparison between the two data enables. The interest here is that contextualized knowledge in our model depends on the evolution of knowledge in the domain. The use of new sensors in cellars and systematic measurements may lead to abandon the process of contextualization of knowledge from the contextual knowledge concerned. Conversely, if more scientific knowledge is available on biochemical relations, some external knowledge (for instance, knowledge represented in the model but not linked with the goal, such as green cover crop) will become contextual.

To put it in a nutshell, there is a local definition of contextual and contextualized knowledge relatively to the step of the resolution process. The status (external, contextual, contextualized) of a piece of knowledge is not static during the whole simulation, only at each step (for a given state of art in the domain).

To conclude, we have to mention that, for a same model, a piece of knowledge can be contextual or not depending on the problem to solve. For instance, once the simulation has run and given a result (risk of stuck fermentation), the model will be used to give a measure of confidence and to test the sensibility of the result to initial conditions and some parameters. In the new problem of the confidence in the result, the context of each resolution step will be different from the one of the risk of stuck-fermentation resolution.

6. CONCLUSION

We have shown in this paper that contextual knowledge can be essential to KBS, to a point where it consists of the basis of the domain model. This is the case for the problem of stuck fermentations in the wine-making domain, where no expert in the domain can be expert on the biochemical problem. However, expert knowledge can be used as contextual information to explain a problem or to evaluate a risk. The use of contextual knowledge during the problem solving implies a discrimination between contextual and contextualized knowledge, and a process of "contextualization" to change the former into the latter.

Proc. of the International and Interdisciplinary Conference on MOdeling and Using Context (CONTEXT-97), Federal University of Rio de Janeiro Ed., 1997, pp. 351-362.

In a further work we will try to organize in the same onion model the superficial knowledge we have gathered from the experts, and to validate the system. We think that this Contextual Knowledge Based System allows to bypass limitations of more classical approaches, restricted to data and knowledge in context.

Acknowledgments:
The COSINUS project concerns researchers of the University Paris 6, Cemagref and INRA[1]. We thank particularly our colleagues that are in charge of the neural network part, F. d'Alche, P. Gallinari and H. Zaragoza, and J.-M. Sablayrolles, L. Blateyron from INRA and P. Grenier from Cemagref. We thank also B. Lesueur (Paris 6 University), who initiated us to METAGEN.

6.REFERENCES

Agabra J. (1996) "Interviews in Champagne and Bourgogne, and questionnaires", Internal Report, LISC, Cemagref. (In French)

Alvarez I. (1995) "Use of Intellisphere learning capability for the micro-project termination of kinetics", Internal Report, LISC, Cemagref. (In French)

Edmondson W. H. and Meech J. F. (1993) "A model of context for human-computer interaction", Proceeding of the IJCAI-93 Workshop on Using Knowledge in its Context, Technical Report 93/13, LAFORIA, University Paris 6, pp.31-38.

Grenier P., Feuilloley P., and Sablayrolles J.M., (1988) "Development of an expert system for the optimization of the wine alcoholic fermentation". International Conference on Agricultural Engineering, March 2-5, Paris, paper n° 88.399.

Grenier P. , **Sablayrolles J-M.** (1990) "Captation de vitesse d'évolution d'un procédé continu en IAA et son utilisation dans les systèmes experts de pilotage." International Symposium ACoFop II Conference, Edition ENSIAA, 313-324.

Insa G., Sablayrolles J.M., Douzal V. (1995) "Alcoholic fermentation under enological conditions. Use of a combination of data analysis and neural networks to predict sluggish and stuck fermentations", Bioprocess Engineering, 442, Springer Verlag.

Jansen B. (1995) "Context in context", http://mac145.syd.dit.csiro.au/Context/context.html (Working Draft V4).

Maigre D., Aerny J. and Murisier F. (1995) "Upkeeping of vine-growing soils and quality of Chasselas wines: influence of permanent green cover and nitrate fertilization" (In French), Revue Suisse Vitic. & Hortic., 27 (4): 237-251.

Marin R. (1995) "Post-doctoral study on prediction of stuck fermentation", Internal Report, Gemo, Cemagref. (In French)

[1] This study has been partially granted by INRA and Cemagref.

Pachet, F. (1995) "On the embeddability of production rules in object-oriented languages.", Journal of Object-Oriented Programming, Vol. 8, N. 4, pp. 19-24

Sablayrolles J.M., Salmon J. M. and Barre P. (1996) "Nutritional deficiency of musts. Efficacity of combined additions of oxygen and ammoniacal nitrogen", Revue française d'oenologie, vol. 159 : 25-29.

Sahraoui H., Revault N., Blain G. and Perrot J. F. (1995) "A Metamodeling technique: The MÉTAGEN system", Proceedings of TOOLS 16 : TOOLS Europe'95, pp. 127-139, Prentice Hall Versailles,France.