

Identifying Control Parameters In Cheese Fabrication Process Using Precedence Constraints

Mélanie Munch, Pierre-Henri Wuillemin, Juliette Dibie-Barthelemy, Cristina Manfredotti, Thomas Allard, S Buchin, Elisabeth Guichard

▶ **To cite this version:**

Mélanie Munch, Pierre-Henri Wuillemin, Juliette Dibie-Barthelemy, Cristina Manfredotti, Thomas Allard, et al.. Identifying Control Parameters In Cheese Fabrication Process Using Precedence Constraints. 21st International Conference on Discovery Science (DS 2018), Oct 2018, Limassol, Cyprus. pp.421-434, 10.1007/978-3-030-01771-2_27 . hal-01923649

HAL Id: hal-01923649
https://hal.science/hal-01923649

Submitted on 15 Nov 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Identifying Control Parameters In Cheese Fabrication Process Using Precedence Constraints

Melanie MUNCH[1], Pierre-Henri WUILLEMIN[2], Juliette DIBIE[1], Cristina MANFREDOTTI[1], Thomas ALLARD[3], Solange BUCHIN[4], and Elisabeth GUICHARD[3]

[1] UMR MIA-Paris, AgroParisTech, INRA, Université Paris-Saclay, 75005 Paris, France
[2] Sorbonne University, UPMC, Univ Paris 06, CNRS UMR 7606, LIP6, 75005 Paris, France
[3] CSGA, AgroSupDijon, CNRS, INRA, Université Bourgogne Franche-Comté, 21000, Dijon, France
[4] UR342 INRA, 39800 Poligny, France

Abstract. Modeling cheese fabrication process helps experts to check their assumption on the domain such as finding which parameters (denoted as control parameters) can explain the final products and its properties. This modeling is however complex as it involves various parameters and a reasoning over different steps. Our previous work presents a method to learn a probabilistic relational model in order to check a user's (an expert on the considered domain) assumption on a transformation process domain, using a knowledge base of this domain and his expert knowledge. However this method did not include temporal information, and thus the learned model is not enough to reason on the cheese fabrication process. In this article we present an extension of our previous work that allows a user to integrate causal and temporal information represented by precedence constraints in order to model a cheese fabrication process. This allows the user to check his assumption to identify the transformation process control parameters.

Keywords: Ontology, Probabilistic Relational Model, Temporality

1 Introduction

Cheese processing is a complex domain involving many different variables. Their combination leads to final products that can differ in quality which can be assessed by different criteria (i.e. sensory, nutritional). Parameters that are enough to explain all these criteria are denoted as **control parameters**. In order to help experts assess and check their assumptions (e.g. identifying control parameters), tools and methods are needed to analyze data. In a previous work [7], we have defined a method helping an expert to check an assumption about possible causal relations between variables by combining a knowledge base and a probabilistic

relational model. However cheese processing being composed of a succession of different steps also includes temporal information, that was not considered in our method. In this article, we therefore propose a generalized version including temporality.

This work has been applied on a real application about cheese processing using data from the TrueFood project. The goal of the TrueFood project is to investigate to what extent the impact of some combinations of thermophile lactic bacteria (i.e. Streptococcus thermophiluss, Lactobacillus helveticus LH with 2 distinct levels and Lactobacillus delbrueckii LD with 2 distinct levels) on the characteristics of hard cooked cheese is affected by the use of milks with various compositions and by the use of different technological conditions (such as distinct temperature for the heating of the milk in the vat). Our study focuses on 24 hard cooked cheese of 10kg each manufactured during three weeks in January 2008, and made using 100 liters vats. Three kinds of milk, differing in their protein content and their production conditions, were used for the cheese making. During the cheese making, three different temperatures (53°C, 55°C and 57°C) were applied for the milk heating. During this study various parameters were monitored, such as different measures of proteolysis. In particular, the potentially bioactive peptides content of the cheeses were measured at several steps of the cheese ripening. Their sensory properties were also assessed at the end of the ripening step: texture and flavor were evaluated by 11 panelists on a 10 points scale.

The influence of milk heating and of combination of lactic bacteria during cheese manufacture on the formation of peptides has already been observed in the literature [12]. Moreover the impact of the type of milk used for the cheese manufacture (especially the influence of the cows feeding system) on the organoleptic properties of hard cheeses has been shown in [9].

In our study, the experts make the assumption that the three factors of variation of the experiments (i.e. type of milk used for the cheese making, combination of thermophile lactic bacteria added to it, and the milk temperature) are the control parameters for the potentially bioactive peptide content of the cheese and its sensory properties. Our aim is to check this assumption using our method extended to take into account both causal and temporal informations.

This paper is structured as follows. Sect. 2 presents the background on probabilistic relational models and related works. Sect. 3 presents the state of the art on dealing with temporal information in probabilistic models. Sect. 4 presents our improved method to help the experts check their assumption using causal and temporal constraints. Sect. 5 presents our study on the data of the TrueFood project. Sec. 6 concludes this article.

2 Background

2.1 BNs And PRMs

Probabilistic relational models (PRMs) extend Bayesian networks (BNs) with the notion of class of relational databases. A BN is the representation of a joint

probability over a set of random variables that uses a Directed Acyclic Graph (DAG) to encode probabilistic relations between variables. However, in the case of numerous random variables with repetitive patterns, it cannot efficiently represent every probabilistic relations.

PRMs extend the BN representation with a relational structure between potentially repeated fragments of BN called classes [14]. They define the high-level, qualitative description of the structure of the domain and the quantitative information given by the probability distribution over the different **attributes** [3], where the attributes represent the different possible values for the variables. In the following, we consider attributes as the objects we want to reason with: we want to assess whether a specific attribute's value can explain another attribute's value. A **class** is defined as a DAG over a set of attributes. These can be inner attributes or attributes from other classes referenced by so-called **reference slots**. The high level structure of a PRM (i.e. its **relational schema**) describes a set of classes C, associated with attributes $A(C)$ and reference slots $R(C)$. A slot chain is defined as a sequence of reference slots that allows one to put in relation attributes of objects that are indirectly related. The probabilistic models are defined on the low level structure (i.e. at the class level) over the set of inner attributes, conditionally to the set of outer attributes and represent generic probabilistic relations inside the classes. This is the **relational model** of the PRM. Classes can be instantiated for each specific situation. A **system** of a PRM provides a probability distribution over a set of instances of a relational schema [15] and, once instantiated, is equivalent to a BN.

2.2 Essential Graph

An instantiated system of a PRM is equivalent to a BN. As a consequence, alongside the construction of the PRM, we also learn an Essential Graph (EG). An EG is a semi-directed graph associated to a BN and composed of edges and oriented arcs. They both share the same skeleton, but the orientation of the EG's edges can vary. If the orientation of an edge is the same for all the Markov equivalent graphs of the BN, this edge is also oriented in the EG; if not, the edge remains unoriented. All directed edges in the EG are called essential arcs [5]. An example of a EG and its two possible interpretations is given by Fig. 1.

The EG expresses whether the orientation of an arc between two nodes can be reversed without modifying the probabilistic relations encoded in the graph. It is useful when presenting results to the user as it can help him visualizing the causal relations learned: when a model has been learned with causal constraints, if an edge is oriented in the EG, it could mean that there is a causal dependence.

2.3 Causal Relation Discovery Driven By An Ontology

In a previous work, we have proposed a method to learn a PRM to discover causal relations in a knowledge base $\mathcal{KB}$ relying on a user's assumption, the user being an expert of the studied domain. We consider, in the following, a

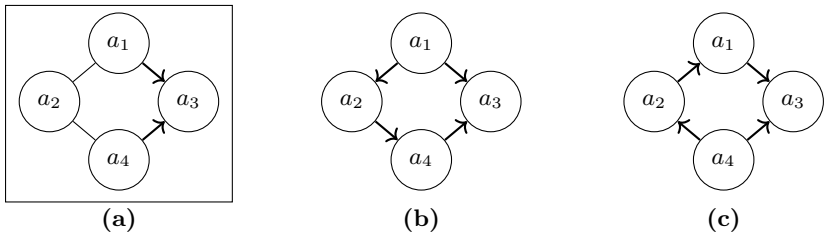


Fig. 1. Example of an essential graph **(a)** and two BNs **(a)** and **(b)** representing possible interpretations.

knowledge base $\mathcal{KB} = (\mathcal{O}, \mathcal{F})$ where the ontology $\mathcal{O}$ is represented in OWL[5] and the data $\mathcal{F}$ in RDF[6]. The user's assumption about possible causal relations between data is of the form *"$E_1, ..., E_n$ have a causal influence on $C_1, ..., C_p$"*, with E_i attributes the user has determined as **explaining** and C_j attributes the user has determined as **consequence**. From the assumption and the knowledge base $\mathcal{KB}$, a database B is created and, afterwards used for the learning. It is composed of the explaining and consequence attributes as well as other inferred attributes as presented in [7].

Given the distinction between explaining and consequence attributes, we introduce some constraints in the learning. In particular, explaining attributes may have an influence over consequence attributes but the inverse is not possible. As a result, if during the learning a relation is found between explaining and consequence attributes, then it has to be oriented from explaining to consequence. These causal constraints guide the probabilistic model construction: indeed, learning using constraints that reflect causality results in a model including causal information and allows the validation of the user's assumption.

Our method gives the user the possibility to check his assumption about possible causal relations between data of a knowledge base. The integration of explaining and consequence attributes helps him express his own knowledge of the domain, and guide the learning towards a coherent causal model. This model however does not take into account possible temporal relations between data and the fact that explaining attributes at one time step can become consequence attributes at the next time step. We denote by **event** a group of attributes that happen at the same time. When dealing with temporal information, it is possible that the consequence attributes of an event e_t at time t become the explaining ones of another attribute of another event e_{t+1} at time $t + 1$, which would be hardly represented by our previous explaining and consequence attributes. Moreover, we can suppose that all the attributes from an event can have an influence over all the attributes of the following events.

We propose, in this paper, an extension of our method dealing with both causality and temporality constraints.

[5] https://www.w3.org/OWL/
[6] https://www.w3.org/RDF/

3 State Of The Art

Both causality and temporality impose a direction to the relation between attributes. For this reason, we have to consider how to take into account constraints while learning PRMs. We first study works on precedence constraints and then how the temporality have been addressed in previous works.

3.1 Learning Under Constraints

Related works have established that using constraints while learning BNs brings more efficient and accurate results. Parameters learning can be improved by allowing users to specify their knowledge through constraints estimations and priors [10]. In [2] an exact structure learning algorithm that uses data and expert's knowledge constraint is presented by defining two types of constraints. In particular one of those identifies where arcs may or may not be included. In [4] it is argued that combining analogical generalization and structure mapping with statistical machine learning methods allows state-of-the-art performances on standards tasks.

In the K2 algorithm [1], a complete ordering of the attributes is required before learning a BN. In this way, the authors introduce **precedence constraints** between the attributes. If, in this order, an attribute A is before an attribute B, then a precedence constraints is applied between those two from A to B, meaning that during the learning we do not consider the possibility of a relation from B to A. If a relation is found between A and B, then the direction of this relation has to be from A to B. However K2 requires a complete knowledge over all the different attributes precedences (since all attributes have to be sorted), which is not always the case as we generally don't know everything about the domain. In this paper we present a method to learn with only partial constraints.

3.2 Integrating Temporality

Temporality has been expressed in Markov models such as Markov chains (MCs, [11]) or Hidden Markov models (HMMs, [11]). MC is a stochastic model describing a sequence of possible events in which the probability of each depends only on the state of the previous one. HMM is a MC with unobserved states, meaning that some attributes' values can vary with an unknown attribute. Temporal information can be gained from both of them: following the flow of time, we can deduce that if an event happens at time t, then it can have an influence over all events that happen at time $t + i$. Moreover we can also deduce subjective independence information between events: if an event only depends on the previous one, then there is no relation between two events which are not consecutive.

However MC and HMM are limited in our case as they cannot handle our need to represent numerous attributes and their relations in time. Both can be extended by dynamic BNs (DBNs, [11]). Widely used to model sequential data, in particular time-series, DBNs introduce the notion of relation between variables over adjacent time steps. For instance, in [6] DBNs are used to model a long-term

simulation of clinical complications in type 1 diabetes. They define two models, one Data-Driven only, and another designed with expert inputs. However, DBNs impose to look at the same attributes and their evolution through time, which is not our aim.

Moreover, we want to consider that every event can have an influence on the attributes of the events that happen after it. This would allow us to better study the possible influence of all attributes on the following ones, which is useful in our problem where we want to assess the relations between attributes of different events (and not only between attributes of consecutive events). This leads us to define a new kind of model, we call it **stack model** as presented below.

4 Stack Model

4.1 Determining Precedence Constraints

Using [7] where we defined explaining and consequence attributes, we propose to decompose the precedence constraints into two sub-constraints: the causal constraints and the temporal constraints. **Causal constraints** are information on the relations between attributes of the type *"The attribute A is a possible cause for the attribute B"*. **Temporal constraint** are information on the relations between attributes of the type *"The attribute A happens before the attribute B"*. These causal and temporal constraints both imply two things: (1) the value of B *can be* explained by A (but it doesn't have to); (2) B can never explain the value of A.

Causal and temporal constraints are differentiated by their nature: temporality is immediate and objective (i.e. the past can influence the future and not the contrary), while causality usually needs a supply of expert knowledge.

Temporal constraints. When possible, the temporal information is provided in the knowledge base through the time ontology[7] that helps anchoring its events in time. In some cases it is also possible to introduce temporal information from other ways (e.g. directly from experts). In all cases we suppose that attributes can be attached to a specific event in time, and as a consequence they also contain temporal information.

Causal constraints. Causal information can be brought by experts or by the ontology itself. In certain cases it is also possible to use statistical independence tests such as χ^2 test used for the construction of causal BNs in order to guess some possible causal relations [13].

4.2 Description

The main idea of our stack model is that it is built in order to graphically represent the two kinds of precedence constraints we defined in Sec. 4.1. If an attribute is put higher in the stack then it has precedence constraint on all

[7] https://www.w3.org/TR/owl-time/

attributes below it; if two attributes are on the same level then they do not have precedence constraints.

It is also possible to encounter parallel events. In this case, we suppose we have enough information from the knowledge base to differentiate the events, in order to know which attribute correspond to which event. In this case, we define paths for each parallel events. Events on the same path all have parenthood links: temporal constraints can be established between them. On the contrary, events that do not share the same parent events are on two separated paths, and we suppose they cannot influence each other. As a consequence, there cannot be precedence constraints between them, neither causal nor temporal.

Starting from a user's assumption, the model construction is based on the two operations described below[8].

1. Defining temporal constraints. Groups of attributes that happen at the same time are put at the same level. If they are from a same event, they are put in a same stack; on the contrary if they are from parallel event we create different paths, each with a stack, for each parallel event.

2. Defining causal constraints. Inside a stack some attributes might have a causal influence over others. In order to express those causal constraints, we sort the attributes such as higher attributes can explain lower attributes and that attributes at the same level share no causal influence between each other.

An example of this construction is given in Fig. 2. We consider here four events: one at time t_1, two parallel at time t_2 and one at time t_3 (a). When constructing the model we first only consider temporal constraints (b): two paths are created with on one side a stack with the group of attributes A and on another side two stacks with respectively the group of attributes B and C, the first being above the second. Finally, a fourth stack is created below all the others, including the group of attributes D. Temporal constraints are defined between the different stacks: since the group B is not on the same path as A, no temporal constraint is drawn between them. In the end, if needed, causal constraints are defined (c). In our example, we suppose that our expert distinguishes between explaining and consequence attributes in the group A, respectively subgroups A_1 and A_2.

In order to lighten the figure, arrows between groups of attributes inside different stacks are not represented: however, if two stacks are linked, then it means that each attribute on the higher stack have a temporal constraint over the lower.

4.3 From Stack Models To PRMs

The final model is used to construct a PRM's relational schema, which defines the classes and their attributes of the PRM. Each subgroup of attributes becomes a class, which are linked together with reference slots following the different

[8] For convenience and in order to ease the readability of the presentation we use in this article a *top-down* construction (from temporality to causality). However nothing prevents us to use the opposite *bottom-up* construction (from causality to temporality).

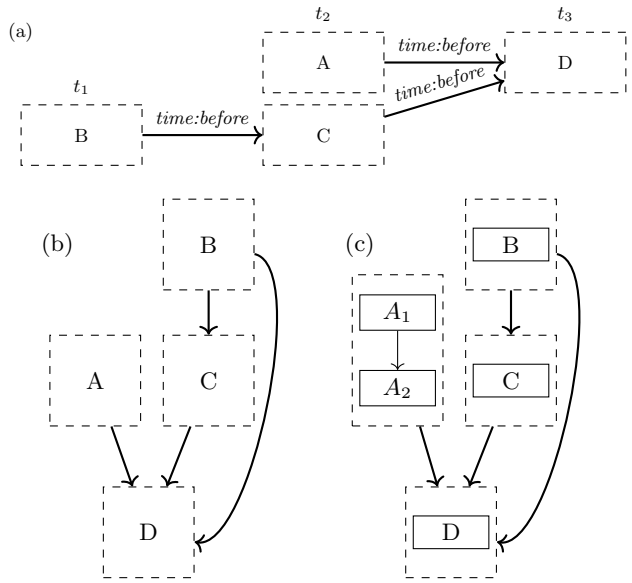


Fig. 2. (a) Example for a system with parallel events. (b) Definition of the temporal constraints. (c) Definition of the causal constraints.

precedence constraints. For instance in the model in Fig. 2 (c), it would lead to five classes and six reference slots.

Once the relational schema is defined, the PRM can be learned using the database B extracted from the knowledge base [8]. This PRM can then be instantiated in order to obtain a BN representing our learned model. It will include causal information as it was learned under precedence constraints; however, it is not a complete causal BN considering that the learning of dependencies between attributes inside the same group is dealt like a classical BN. In order to deal with causal information, the EG is used: if an arc is oriented in it (meaning that its orientation cannot be changed without changing the likelihood of the BN) then it can mean that there is a causal relation.

5 Experiment

5.1 Data Description

Considering the TrueFood project, the experts would like to model the different relations between the attributes in order to explain the products at the end and infer its characteristics. More particularly they want to check the formulated assumption: *"The temperature, ferments and type of milk have a causal influence on the potentially bioactive peptide content of the cheese and its sensory properties"*. Following the approach presented in [7] temperature, ferments and

type of milk are the only explaining attributes of the problem, while the other are consequences. Since those three are fixed at the beginning, they correspond to the **control parameters**.

The dataset is composed of data from three different steps that are part of a cheese fabrication and tasting process: Step in the vat, Ripening and Mastication.

- **Step in the vat:** is described by three processing control parameters (Temperature, Starters and Type of milk), and two measured (hardening and clotting times).
- **Ripening:** is described by the measured value of five different concentrations in cheese: butyric acid, propionic acid, acetic acid, free amino acids and free amino groups.
- **Mastication:** In this step, a panel of 11 judges has evaluated each cheese sample on 45 different criteria (e.g. spice aroma, sugar or fat perception). Those sensory notes can be divided into two categories, texture of the cheese (10 attributes) and flavor (35 attributes). The scores ranged from 0 to 10.

The times measured during the step in the vat are a pre-requisite to **study bioactive peptide contents**, even if they do not represent their quantities. On another hand the attributes measured during the ripening and the mastication steps are useful to **evaluate the cheese sensory properties**.

5.2 Model Construction

A first descriptive analysis over the notes attributes during the mastication step shows that some have a variance $\sigma < 0.25$. Given the standard variation calculated by $\sqrt{\sigma}$, it means that for these attributes the variation over the whole samples is less than ± 0.5 points. We consider it to be too low to observe meaningful variations among the different samples, and remove them from the studied set, leaving 39 attributes (9 texture attributes and 30 flavor attributes).

In order to apply our method to this dataset, we first separate the attributes per steps. We order them in stacks, following the temporal order: first *Step in the vat*, then *Ripening* and then *Mastication*. Temporal constraints are then drawn. Once this has been done, the only causal constraints that need to be introduced are taken into account in order to separate the control attributes from the rest. Temperature, Ferments and Type of Milk are stacked above the Hardening and clotting times inside the same step.

The obtained model is presented in Fig. 3, where the different steps are underlined by the dashed squares and $\times i$ denotes the number of attributes of the given type.

5.3 Analysis

While analyzing the PRM we focus on two types of relations: the intra-step and the inter-step relations. While the analysis of the intra-step relations in general has already been tackled in our previous work (leading to causal information

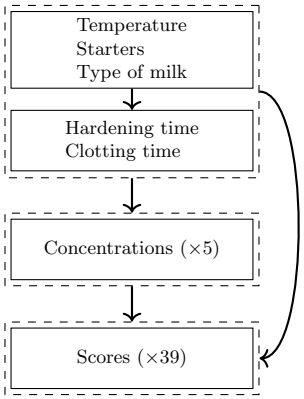


Fig. 3. Model constructed from the expert assumption.

analysis), the inter-step gives a whole new reading of the model. It indeed helps us generate new information about the temporal aspect, in particular discovering if some steps can explain all the other attributes, or, on the contrary, if a step has no influence on the process. In our case, we would like to see in what extent the control parameters are able to explain (in)directly the other attributes. As a consequence we extend our study on inter-step relations, also including the inter-subgroup relations between the control parameters and the two attributes Solidifying time and Clotting time.

In order to illustrate our results, we consider three attributes A, B and C with A a control parameter in a step, and B and C two attributes of the step after (Fig. 4). When checking whether A can explain C, two cases are possible:

1. A **has a complete or partial control over** C**.** In the first case (Fig. 4 (a)) there exists a inter-step relation directly from A to C: it means that knowing the state of A will give the maximum possible information on C. In the second case (Fig. 4 (b)) the inter-step relation between A and C is intercepted by other attributes, B in our example. Since there is no direct relation between A and C, then knowing A will only give partial information on C. Moreover, knowing B makes the knowledge of A obsolete, as B alone is enough to have a complete information on C.
2. A **has no control over** C**.** There are also two possible cases. The first (Fig. 4 (c)) is straight-forward: since there are no inter-step relations between A and B nor C, then A has no control over C. In the second (Fig. 4 (d)) a *v-structure* $A \rightarrow B \leftarrow C$ makes A and C independent: A and C are *d-separated* by B, meaning that controlling A cannot influence C, however fixing B gives partial information on both A and C.

It is important to note that we assume the direction of the relation between A and B, C because we previously defined A as a control parameter. However, since B and C are not control parameters, there is no precedence constraint that indicates whether B has an influence on C or otherwise. In order to do this

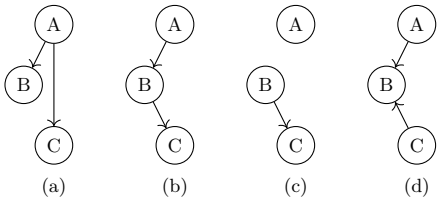


Fig. 4. Possible cases encountered during the results analysis of the assumption *"A has an influence over C"*.

analysis we, therefore, use the EG, that indicates whether the direction of the relation is sure (i.e. changing this relation direction would modify the likelihood of the learned BN). If a relation is oriented in the EG, we can suppose that this orientation is due to the precedence constraints used during the learning that brought causal information and, therefore, this relation orientation might be causal. This is not however automatic, and relations orientation in the EG can sometimes reflect other problem such as learning artifact or missing attributes. If, finally, a relation is not oriented in the EG then we cannot assume causal information from it. For the following we assume that, when talking about relations orientation, we are using the EG informations.

5.4 Results

The vast majority of the observed inter-steps relations found **confirms the experts assumption**: *"The temperature, ferments and type of milk have a causal influence on the potentially bioactive peptide content of the cheese and its sensory properties"*. Some of them are directly explained, while others are linked to attributes of the same group that are explained by the control parameters. Only three sensory notes are not linked at all to any parameter. Those results and the number of found relations are summarized by Fig. 5.

While the study of the times during processing and concentrations is pretty straight-forward, all being completely or partially explained by the control parameters, an interesting trend in the sensory notes attributes can be observed while looking at the EG. Indeed we can notice that a large group of 21 flavor attributes (over the 30) is d-separated from the control parameters by another sensory attribute, meaning that this part is in fact equally independent from the control parameters despite being part of the network. More generally we observe in the EG a difference between flavor and texture attributes. Fig. 6 shows an excerpt of the learned EG where texture attributes are denoted by T_i ($i \in [1, 9]$) and flavor attributes are separated into two groups denoted by F_j ($i \in [1, 30]$). This choice has been made in order to ease the reading: however one must keep in mind that two relations between F_j and different T_i do not involve the same attribute of F_j.

Flavor attributes in F_1 are d-separated from the control parameters, and the other in F_2 are partially explained by them. Moreover when looking at F_1 we

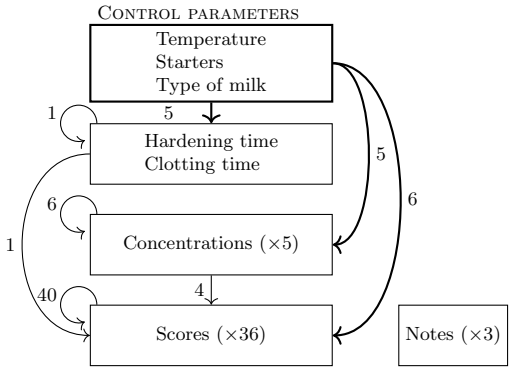


Fig. 5. Summary of the number of observed inter and intra step relations.

observe a large number of intra-step relations between them. This leads us to assume that (1) they are highly correlated with each other and (2) their relations are not causal due to the high number of attributes learned together without any precedence constraints. On the contrary texture attributes are mostly directly explained by the control parameters (especially the type of milk), and some are partially explained by them. Most of them do not have intra-step relations with each other as seen in Fig. 6. The two texture attributes not represented in the network are related to time and concentration attributes and not directly linked to the control parameters. These observations are validated by the experts: considering the milk differences in terms of production conditions and composition, milk on the cheese texture was expected. In addition, flavor attributes are indeed more likely to be correlated with each other.

Since nearly all flavor attributes are linked together, it could be interesting to profile the cheeses with their different flavor values. This way, instead of reasoning with all the numerous flavor attributes, we could directly check the influence of the control parameter on the cheese type.

6 Conclusion

In this article we present a new method able to help experts to study a domain represented by a knowledge base. Our aim is to provide to the experts a method to check the assumptions they can formulate on that domain by learning a PRM that presents the different probabilistic relations between its attributes. In order to guide this learning we also allow the experts to integrate causal and temporal informations by defining precedence constraints. Considering the TrueFood project we use our method to check the following assumption: *"The temperature, ferments and type of milk have a causal influence on the potentially bioactive peptide content of the cheese and its sensory properties"*.

To integrate precedence constraints, we extend our previous work, that already included causal information, in order to include temporal information. To

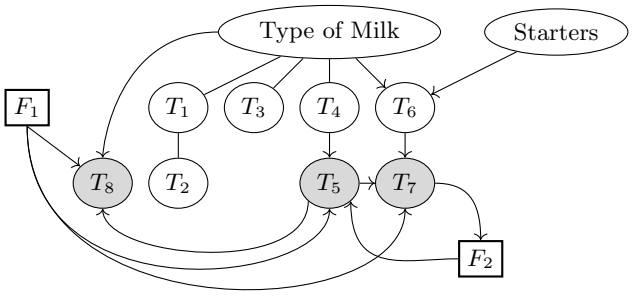


Fig. 6. Excerpt of the EG learned with T_i texture attributes and F_i groups of flavor attributes. Grey attributes d-separates F_1 from the control parameters.

do so, we define a new model, denoted by the **stack model**, where attributes are organized so that higher ones can have a precedence constraint over the ones below (i.e. they can be their cause). Using this model, we learn a PRM which, once instantiated, gives us a BN we can use to check the assumption.

The learned BN gives us two ways of analysis. First, using its EG (a graph that shows arcs whose direction is sure considering the BN structure), we can check the assumption. Considering that our control parameters are fixed at the beginning of the process, if a relation is found between them and an attribute, then we can conclude that the parameters may control this attribute. Second, once the model has been validated by the experts, it can be used to predict results. For instance if we want to control the cheese texture scoresin order to keep them in a certain range, we identify the control parameters we have to act on.

In future work we want to study in more detail the validation and introspection of the learned model in order to improve our help to the user.

References

1. Gregory F. Cooper and Edward Herskovits. A bayesian method for the induction of probabilistic networks from data. *Machine Learning*, 9(4):309–347, Oct 1992.
2. Cassio P. de Campos, Zhi Zeng, and Qiang Ji. Structure learning of bayesian networks using constraints. In *Proceedings of the 26th Annual International Conference on Machine Learning*, ICML '09, pages 113–120, New York, NY, USA, 2009. ACM.
3. Nir Friedman, Lise Getoor, Daphne Koller, and Avi Pfeffer. Learning probabilistic relational models. In *Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence, IJCAI 99, Stockholm, Sweden, July 31 - August 6, 1999. 2 Volumes, 1450 pages*, pages 1300–1309, 1999.
4. Chen Liang and Kenneth D. Forbus. Learning plausible inferences from semantic web knowledge by combining analogical generalization with structured logistic regression. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*, AAAI'15, pages 551–557. AAAI Press, 2015.

5. David Madigan, Steen A Andersson, Michael D Perlman, and Chris T Volinsky. Bayesian model averaging and model selection for markov equivalence classes of acyclic digraphs. *Communications in Statistics–Theory and Methods*, 25(11):2493–2519, 1996.
6. Simone Marini, Emanuele Trifoglio, Nicola Barbarini, Francesco Sambo, Barbara Di Camillo, Alberto Malovini, Marco Manfrini, Claudio Cobelli, and Riccardo Bellazzi. A dynamic bayesian network model for long-term simulation of clinical complications in type 1 diabetes. *Journal of Biomedical Informatics*, 57:369 – 376, 2015.
7. Melanie Munch, Pierre-Henri Wuillemin, Cristina E. Manfredotti, and Juliette Dibie. Towards interactive causal relation discovery driven by an ontology. Technical report, https://hal.archives-ouvertes.fr/hal-01823862v1, 2018.
8. Melanie Munch, Pierre-Henri Wuillemin, Cristina E. Manfredotti, Juliette Dibie, and Stéphane Dervaux. Learning probabilistic relational models using an ontology of transformation processes. In *On the Move to Meaningful Internet Systems. OTM 2017 Conferences - Confederated International Conferences: CoopIS, C&TC, and ODBASE 2017, Rhodes, Greece, October 23-27, 2017, Proceedings, Part II*, pages 198–215, 2017.
9. Tom F. O'Callaghan, David T. Mannion, Deirdre Hennessy, Stephen McAuliffe, Maurice G. O'Sullivan, Natasha Leeuwendaal, Tom P. Beresford, Pat Dillon, Kieran N. Kilcawley, Jeremiah J. Sheehan, R. Paul Ross, and Catherine Stanton. Effect of pasture versus indoor feeding systems on quality characteristics, nutritional composition, and sensory and volatile properties of full-fat cheddar cheese. *Journal of Dairy Science*, 100(8):6053 – 6073, 2017.
10. Cassio P. de Campos and Qiang Ji. Improving bayesian network parameter learning using constraints, 01 2009.
11. Kevin Patrick Murphy. Dynamic bayesian networks: Representation, inference and learning, 01 2002.
12. Lourdes Santiago-López, Jose E. Aguilar-Toalá, Adrián Hernández-Mendoza, Belinda Vallejo-Cordoba, Andrea M. Liceaga, and Aarón F. González-Córdova. Invited review: Bioactive compounds produced during cheese ripening and health effects associated with aged cheese consumption. *Journal of Dairy Science*, 101(5):3742 – 3757, 2018.
13. P. Spirtes, C. Glymour, and R. Scheines. *Causation, Prediction, and Search*. MIT press, 2nd edition, 2000.
14. Lionel Torti, Pierre-Henri Wuillemin, and Christophe Gonzales. Reinforcing the Object-Oriented Aspect of Probabilistic Relational Models. In *PGM 2010 - The Fifth European Workshop on Probabilistic Graphical Models*, pages 273–280, Helsinki, Finland, September 2010.
15. Pierre-Henri Wuillemin and Lionel Torti. Structured probabilistic inference. *Int. J. Approx. Reasoning*, 53(7):946–968, 2012.