Information for decision-making is ubiquitous: Revisiting the reverse engineering mode in breadmaking technology

Rallou Thomopoulos, Ahmed Chadli, Madalina Croitoru, Joel Abecassis, Gérard Brochoire, Hubert Chiron

To cite this version:

HAL Id: lirmm-01382351
https://hal-lirmm.ccsd.cnrs.fr/lirmm-01382351
Submitted on 16 Oct 2016

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.
Information for decision-making is ubiquitous: revisiting the reverse engineering mode in breading technology

Rallou Thomopoulos¹,², Ahmed Chadli¹, Madalina Croitoru², Joël Abécassis³, Gérard Brochoire³ and Hubert Chiron⁴

¹ IATE joint research unit, CIRAD-INRA-Supagro-UM, Montpellier, France
² INRIA project team GraphIK, LIRMM (CNRS-UM), Montpellier, France
³ INBP, Rouen, France
⁴ BIA research unit, INRA, Nantes, France
Contact author email: rallou.thomopoulos@supagro.inra.fr

Abstract—This paper deals with the process of decision making in the reverse engineering mode and highlights the need for polyvalent information. Three aspects are considered. 1) Reverse engineering implies a preliminary assumption: having defined a desired outcome of the decision process. Defining goals on the possible outcomes is a complex, multi-actor process based on ubiquitous information. Once identified at best, several alternative scenarios may lead to the desired outcome. The first issue consists in evaluating these alternative scenarios. 2) While taking into consideration the positive consequences that the different alternatives will generate, the decision process has to allow for possible negative impacts, which are not explicitly expressed in the defined goals. We thus consider the reverse engineering process has to be bipolar and take rejections into account. 3) Finally, the simultaneous achievement (respectively, avoidance) of several goals (respectively, rejections) is not always possible and depends, in particular, on whether the actions leading to each of these goals (respectively avoiding these rejections) are compatible or not. We thus seek the “best” compatible set of actions and propose to define it as optimizing the bipolar preferences expressed on the outcomes. The approach is both graphical and logical and is focused on a case study in breading technology.

Keywords: Conceptual Modelling and Ontologies, Collaborative Analysis and Decision, Decision Information Systems, Food Science and Technology Application, Multi-Source Information.

I. INTRODUCTION

This paper deals with the generic question of decision making. It considers three aspects of decision making:

1) decision scenarios can be seen as possible actions that will lead to different outcomes. Making a decision implies to have previously defined goals or preferences on these outcomes;
2) once the goals defined, several ways (i.e. actions) may be possible to achieve them. The choice between these ways has to take into consideration the consequences that these actions will generate, positive or negative, including aspects outside of the defined goals;
3) finally, the simultaneous achievement of several goals is not always possible. It depends whether the actions leading to each of these goals are compatible or not.

How to maximize the set of reached goals is thus a relevant question.

Why is information ubiquitous? The paper deals with a case study in breading technology. Choices concerning the best kind of bread to produce on a large scale rely on various sources and cover different points of view. They include sensorial, nutritional, cost and practicality aspects, as well as hygienic or environmental impact concerns. Information sources go from websites, project meetings, expert interviews, scientific articles, manufacturing practices to consumers’ habits, opinions, preferences and choices, available through online forums, sales statistics, new marketing trends, etc. We will explain why we chose a graphical and logical approach, what it allows to do that other approaches do not.

In agrifood chains, the products traditionally go through the intermediate stages of processing, storage, transport, packaging and reach the consumer (the demand) from the producer (the supply). More recently, due to an increase in quality constraints, several parties are involved in production process, such as consumers, industrials, health and sanitary authorities, etc. expressing their requirements on the final product as different point of views which could be conflicting. The notion of reverse engineering control, in which the demand (and not the supply) sets the specifications of desired products and it is up to the supply to adapt and find its ways to respond, can be considered in this case.

In this article, we discuss two aspects of this problem. First, we accept the idea that specifications cannot be established and several complementary points of view - possibly contradictory - can be expressed (nutritional, environmental, taste, etc.). We then need to assess their compatibility (or incompatibility) and identify solutions satisfying a maximum set of viewpoints. To this end we propose a logical framework based on argumentation and introduce a method of decision making based on backward chaining for the bread industry.

Since a joint argumentation - decision support approach is highly relevant to the food sector [1], the contribution of the paper is twofold. First we present a real use case of an argumentation process in the agrifood domain. Second we
introduce the notion of viewpoint / goal in this setting based on the notion of backwards chaining reasoning and show how to use those techniques in a concrete application.

The main alternative method to deal with the problem is the multicriteria decision approach. However multicriteria decision aims at evaluating several alternative options, whereas argumentation-based decision focuses on whether several options make sense together, which is a different perspective, addressed in this paper. Moreover, multicriteria decision is not connected to the backward chaining procedure as the argumentative approach is, by construction of the arguments, as will be explained in Section V-B.

In Section II, we introduce the real scenario considered in the application and the intuitive principle of the approach. In Section III, we motivate our technical and modeling choices. In Section IV, the developed approach is introduced. It relies on an instantiation of a logic based argumentation framework based on a specific fragment of first order logic. In Section V, we explain the technical results that ensure the soundness and completeness of our agronomy application method. In Section VI, some evaluation results are presented. Finally, Section VII concludes the paper.

II. SCENARIO AND INTUITIVE PRINCIPLE

A. Breadmaking scenario

The case of study considered in this paper relates to the debate around the change of ash content in flour used for common French bread. Various actors of the agronomy sector are concerned, in particular the Ministry for Health through its recommendations within the framework of the PNNS (“National Program for Nutrition and Health”), the millers, the bakers, the nutritionists and the consumers.

The PNNS recommends to privilege the whole-grain cereal products and in particular to pass to a common bread of T80 type, i.e made with flour containing an ash content (mineral matter rate) of 0.8%, instead of the type T65 (0.65% of mineral matter) currently used. Increasing the ash content comes down to using a more complete flour, since mineral matter is concentrated in the peripheral layers of the wheat grain, as well as a good amount of components of nutritional interest (vitamins, fibers). However, the peripheral layers of the grain are also exposed to the phytosanitary products, which does not make them advisable from a health point of view, unless one uses organic flour.

Other arguments (and of various nature) are in favour or discredit whole-grain bread. From an organoleptic point of view for example, the bread loses out in its “being crusty”. From a nutritional point of view, the argument according to which the fibers are beneficial for health is discussed, some fibers could irritate the digestive system. From an economic point of view, the bakers fear selling less bread, because whole-grain bread increases satiety – which is beneficial from a nutritional point of view, for the regulation of the appetite and the fight against food imbalances and pathologies. However whole-grain bread requires also less flour and more water for its production, thus reducing the cost. The millers also fear a decrease in the quality of the technical methods used in the flour production.

Beyond the polemic on the choice between two alternatives (T65 or T80), one can take the debate further by distinguishing the various points of view concerned, identifying the desirable target characteristics, estimating the means of reaching that point. The contribution of this paper is showing how using argumentation can help towards such practical goals.

B. Intuitive principle

Information about the technical steps of an agrifood chain and their impact on end-product properties may be summarized as shown in Figure 1. The meaning of the diagram is the following: if a product goes through Action 1, followed by Action 1.1, then it will have the properties Property 1 and Property 2. Of course the process may include more than two consecutive actions and lead to a various number of properties. The diagram is limited for simplicity reasons.

Reverse engineering firstly consists in defining the expected properties of the end-product, for instance in Figure 2 Property 3 is expected.

For a given expected property, several paths – i.e. several sequences of actions – may be possible to obtain it. For instance, in Figure 3, in order to obtain Property 3, two sequences are possible: 1) Action 1 followed by Action 1.2 or 2) Action 2 followed by Action 2.1.
The choice between these potential sequences has to be analyzed, since it is likely that they do not have the same consequences. Let us suppose that the sequence Action 2 - Action 2.1 is chosen, as illustrated in Figure 4.

We then have to consider that the choice of this sequence will also have as a result obtaining Property 4 and Property 5, as shown in Figure 5. However, these properties may be wanted or not, which has to be taken into account.

The case where several properties are expected is illustrated in Figure 6. Property 1 and Property 5 are wanted in this example.

In this case, several sequences of actions will have to be conducted in parallel, as shown in Figure 7: 1) Action 1 followed by Action 1.1 and 2) Action 2 followed by Action 2.2. The question is: are those actions compatible, i.e. is it possible to achieve them simultaneously, or are they exclusive. Both cases are to be considered.

In the following sections, we will show how we can compute the compatibility of those actions, using an argumentation framework as a means of computation.

III. MOTIVATION

In this paper we will elicit the points of view and the desirable target characteristics by the means of interviews with agronomy experts. Once the target characteristics identified, finding the means of reaching them will be done automatically by a combination of reverse engineering and argumentation. The reverse engineering will be used in order to find the complete set of actions to take towards a given characteristic, for all characteristics. In certain cases the actions to take will be inconsistent. Argumentation will then be employed in order to identify actions that can be accepted together.
A. Reverse Engineering

While reverse engineering has been widely employed in other Computer Science domains such as multi agent systems or requirements engineering (e.g. [2]), it is quite a novel methodology when applied in agronomy. In agrifood chains, the products traditionally go through the intermediate stages of processing, storage, transport, packaging and reach the consumer (the demand) from the producer (the supply). It is only recently, due to an increase in quality constraints, that the notion of reverse engineering control has emerged [3]. In this case the demand (and not the supply) sets the specifications of desired products and it is up to the supply to adapt and find its ways to respond. In what follows, starting from the desired target criteria for the final product, the methods allowing one to identify ways to achieve these criteria (by intervention on the various stages of the supply chain) are named “reverse engineering”.

Reverse engineering is known to be challenging from a methodological viewpoint. This is due to two main aspects. First, the difficulty of defining the specifications for the expected finished product. The desired quality criteria are multiple, questionable, and not necessarily compatible. The second difficulty lies in the fact that the impact of different steps of food processing and their order is not completely known. Some steps are more studied than others, several successive steps can have opposite effects (or unknown effects), the target criteria may be outside of the characteristics of products. Second, reconciling different viewpoints involved in the food sector still raises unaddressed questions. The problem does not simply consist in addressing a multi-criteria optimisation problem [4]: the domain experts would need to be able to justify why a certain decision (or set of possible decisions) is taken.

B. Argumentation

Argumentation is a reasoning model based on the construction and the evaluation of interacting arguments. It has been applied to nonmonotonic reasoning, decision making, or for modeling different types of dialogues including negotiation. Most of the models developed for these applications are grounded on the abstract argumentation framework proposed by Dung in [5]. This framework consists of a set of arguments and a binary relation on that set, expressing conflicts among arguments. An argument gives a reason for believing a claim, for doing an action.

Argumentation theory in general [5], [6], [7] is actively pursued in the literature. Some approaches combine argumentation and multi criteria decision making [8].

Value based Argumentation Frameworks [9] have been proposed and applied to real-world case studies [10], where the strength of an argument corresponds to the values it promotes. What we call viewpoint later on in this paper would then correspond to the notion of audience in such setting. Although intuitive, this approach is not adapted in the case of the considered application. Here a value can be “split” into several audiences: there could be contradictory goals even from the same viewpoint. The notion of viewpoint and goals introduced in this setting also reminds us of those proposed by [11].

1) Logic-based Argumentation: In this paper we present a methodology combining reverse engineering and logical based argumentation for selecting the actions to take towards the agronomy application at hand. The logical instantiation language is a subset of first order logic denoted in this paper $\mathcal{S}RC$ equivalent to Datalog+ [12], Conceptual Graphs or Description Logics (more precisely the $\mathcal{EL}$ fragment [13] and DL-Lite families [14]). All above mentioned languages are logically equivalent in terms of representation or reasoning power. The reason why this application is using $\mathcal{S}RC$ is the graph based representation specific to $\mathcal{S}RC$ (and not to the other languages). This graph based representation (implemented in the Cogui tool [15], [16]) makes the language suitable for interacting with non computing experts [16].

Here we use the instantiation of [17] for defining what an argument and an attack are. While other approaches such as [18], [19], [20] etc. address first order logic based argumentation, the work of [17] uses the same $\mathcal{S}RC$ syntax and graph reasoning foundations. In Figure 8 the visual interface of Cogui is depicted: knowledge is represented as graph which is enriched dynamically by rule application. More on the visual appeal of Cogui for knowledge representation and reasoning can be found in [16].

IV. Approach

As mentioned above, in this paper we use an instantiation of logic based argumentation based on a specific fragment of first order logic. This subset is equivalent to Datalog+ [12], Conceptual Graphs or Description Logics (the $\mathcal{EL}$ fragment [13] and the DL-Lite families [14]). The reason for which our application required this specific logic fragment is related to the information capitalisation needs of the food sector. The long term aim is to enrich ontologies and data sources based on these ontologies and join the Open Data movement. This entails that the language used by the food applications needs to be compatible with the Semantic Web equivalent languages.

The choice of the $\mathcal{S}RC$ syntax and graph reasoning mechanism is justified by the visual appeal of this language for non computing experts.

In a nutshell our methodology is as follows. The set of goals, viewpoints as well as the knowledge associated with the goals / viewpoints is elicited either by the means of interviews with the domain experts or manually from different scientific papers. This step of the application is the most time consuming but the most important. If the knowledge elicited is not complete, sound or precise the outcome of the system is compromised. Then, based on the knowledge elicited from the knowledge experts and the goals of the experts, we enrich the knowledge bases using reverse engineering (implemented using backwards chaining algorithms). Putting together the enriched knowledge bases obtained by backwards chaining from the different goals will lead to inconsistencies. The argumentation process is used at this step and the extensions yield by the applications computed. Based on the extensions and the associated viewpoints we can use voting functions to determine the application choice of viewpoints.

A. Use Case Real Data

Expressing the target characteristics – or goals – according to various points of view consists of identifying the facets
involved in the construction of product quality: points of view, topics of concern such as nutrition, environment, technology, etc. In addition, such viewpoints have to be addressed according to their various components (fibers, minerals, vitamins, etc). Desirable directions need to be laid down, and in a first step we consider them independent from each other.

The considered sources of information include, from most formal to less formal: (1) peer reviewed scientific papers; (2) technical reports or information posted on websites; (3) conferences and scientific meetings around research projects; (4) expert knowledge obtained through interviews. The scientific articles we have analysed – with the supervision of experts in agrifood – include: [21], [22], [23], [24], [25]. [21] compares the different types of flour from a nutritional point of view. [22] explores the link between fiber and satiety. [23], [24] deal with consumer behaviour and willingness to pay. They focus on French baguette when information concerning the level of fibers is provided, and they base their results on statistical studies of consumer panels. [25] provides a summary of the nutritional aspects of consumption of bread and the link with technological aspects.

We also reviewed technical reports available on official websites on health policy: the public PNNS (National Program for Nutrition and Health, www.mangerbouger.fr/pnns) [26], the European project Healthgrain (looking at improving nutrition and health through grains) [27], [28], as well as projects and symposia on sanitary measures regarding the nutritional, technological and organoleptic properties of breads [29], [30], [31], [32]. Finally, several interviews were conducted to collect domain expert knowledge, in particular for technology specialists in our laboratory.

A summary of the results obtained in the baking industry is synthesised in Figure 9 regarding nutritional and organoleptic aspects. Figure 9(a) shows the main identified goals to reach for a nutritionally optimised bread (for instance, containing a high level of soluble fibers, vitamins and minerals, low salt, etc.), whereas Figure 9(b) sums up the main goals to achieve for an enjoyable bread regarding sensorial concerns (for example, crusty, etc.).

V. TECHNICAL SOUNDNESS

In this section we explain the technical results that ensure the soundness and completeness of our agronomy application method. The section is composed of three parts. A first subsection explains the logical subset of first order logic language employed in the paper. The second subsection shows how to construct arguments and attacks in order to obtain extensions when a knowledge base expressed under this language is inconsistent. Last, the third section shows how we used reverse engineering to complete the knowledge base with all possible actions and how argumentation can be used in order to select consistent subsets of knowledge which support given actions.

A. The Logical Language

In the following, we give the general setting knowledge representation language used throughout the paper.

A knowledge base is a 3-tuple $\mathcal{K} = (\mathcal{F}, \mathcal{R}, \mathcal{N})$ composed of three finite sets of formulae: a set $\mathcal{F}$ of facts, a set $\mathcal{R}$ of
rules and a set $N$ of constraints. Let us formally define what we accept as $F$, $R$ and $N$.

**Facts Syntax.** Let $C$ be a set of constants and $P = P_1 \cup P_2 \ldots \cup P_n$ a set of predicates of the corresponding arity $i = 1, \ldots, n$. Let $V$ be a countably infinite set of variables. We define the set of terms by $T = V \cup C$. As usual, given $i \in \{1 \ldots n\}$, $p \in P_i$ and $t_1, \ldots, t_i \in T$ we call $p(t_1, \ldots, t_i)$ an atom. A fact is the existential closure of an atom or an existential closure of a conjunction of atoms. (Note that there is no negation or disjunction in the facts and that we consider a generalised notion of facts that can contain several atoms.)

- $\text{Bread}$, $\text{Cereal}$, $\text{LowSalt}$, $\text{ContaminantFree}$ are examples of unary predicates (arity 1) and $\text{IsIngredientOf}$ is a binary predicate (arity 2).
- $\text{Wheat}$, oats, rye, barley are constant examples.
- $\text{Cereal}$ (wheat) is an atom.
- $\exists x (\text{Bread}(x) \land \text{IsIngredientOf}(\text{wheat}, x))$ is a fact.

Due to lack of space we do not show the full semantic definitions of facts (or rules and constraints in the following section). For a complete semantic depiction of this language please check [15], [16], [17]. It is well known that $F' \models F$ (read the fact $F'$ entails the fact $F$) if and only if there is a homomorphism from $F$ to $F'$ [15].

**Rules.** A rule $R$ is a formula of the form

$$\forall x_1, \ldots, \forall x_n \forall y_1, \ldots, \forall y_m (H(x_1, \ldots, x_n, y_1, \ldots, y_m) \rightarrow \exists z_1, \ldots, \exists z_k C(y_1, \ldots, y_m, z_1, \ldots, z_k))$$

where $H$, the hypothesis, and $C$, the conclusion, are atoms or conjunctions of atoms, $n, m, k \in \{0, 1, \ldots\}$, $x_1, \ldots, x_n$ are the variables appearing in $H$, $y_1, \ldots, y_m$ are the variables appearing in both $H$ and $C$ and $z_1, \ldots, z_k$ the new variables introduced in the conclusion. An example of a rule is the following:

$$\forall x (\text{Bread}(x) \land \text{PesticideFree}(x) \land \text{MycotoxinFree}(x) \rightarrow \text{ContaminantFree}(x)).$$

In the following we will consider rules without new existential variables in the conclusion.

Reasoning consists of applying rules on the set $F$ and thus inferring new knowledge. A rule $R = (H, C)$ is applicable to set $F$ if and only if there exists $F' \subseteq F$ such that there is a homomorphism $\sigma$ from the hypothesis of $R$ to the conjunction of elements of $F'$. A rule $R = (H, C)$ is inversely applicable to a fact $F$ if there is a homomorphism $\pi$ from $C$ to $F$. In this case, the inverse application of $R$ to $F$ according to $\pi$ produces a new fact $F''$ such that $R(F'') = F$. We then say that the new fact is an immediate inverse derivation of $F$ by $R$, abusively denoted $R^{-1}(F)$.

Note that this technique is commonly used, for example, for backward chaining query answering [33], [34] where a query is rewritten according to the rules. The same mechanism is also discussed by abductive reasoning algorithms [35] where minimal sets of facts (in the set inclusion sense) are added to the knowledge base in order to be able to deduct a query.

Let $F = \text{Bread(bleuette)} \land \text{PesticideFree(bleuette)} \land \text{MycotoxinFree(bleuette)}$ and $R$ the rule $\forall x (\text{Bread}(x) \land \text{PesticideFree}(x) \land \text{MycotoxinFree}(x) \rightarrow \text{ContaminantFree}(x))$.

$R$ is applicable to $F$ and produces by derivation the following fact: $\text{Bread(bleuette)} \land \text{PesticideFree(bleuette)} \land \text{MycotoxinFree(bleuette)} \land \text{ContaminantFree(bleuette)}$.

Let $F = \text{Bread(bleuette)} \land \text{ContaminantFree(bleuette)}$ and $R$ the rule $\forall x (\text{Bread}(x) \land \text{PesticideFree}(x) \land \text{MycotoxinFree}(x) \rightarrow \text{ContaminantFree}(x))$.

$R$ inversely applicable to $F$ and produces by inverse derivation the fact: $F' = \text{Bread(bleuette)} \land \text{PesticideFree(bleuette)} \land \text{MycotoxinFree(bleuette)}$.

Let $F$ be a subset of $F$ and let $R$ be a set of rules. A set $F_0$ is called an $R$-derivation of $F$ if there is a sequence of sets (called a derivation sequence) $(F_0, F_1, \ldots, F_k)$ such that $F_0 \subseteq F$, $F_0$ is $R$-consistent, for every $i \in \{1, \ldots, n - 1\}$, it holds that $F_i$ is an immediate derivation of $F_{i-1}$.

Given a set $\{F_0, \ldots, F_k\} \subseteq F$ and a set of rules $R$, the closure of $\{F_0, \ldots, F_k\}$ w.r.t. $R$, denoted $\mathcal{C}_R(\{F_0, \ldots, F_k\})$, is defined as the smallest set (with respect to $\subseteq$) which...
contains \( \{F_0, \ldots, F_k\} \), and is closed for \( R \)-derivation (that is, for every \( R \)-derivation \( F_n \) of \( \{F_0, \ldots, F_k\} \), we have \( F_n \subseteq \bigcap R(\{F_0, \ldots, F_k\}) \)). Finally, we say that a set \( F \) and a set of rules \( R \) entail a fact \( G \) (and we write \( F, R \models G \)) iff the closure of the facts by all the rules entails \( F \) (i.e. if \( \bigcap R(\{F_0, \ldots, F_k\}) \models G \)).

**Constraints.** A constraint is a formula \( \forall x_1 \cdots \forall x_n (H(x_1, \ldots, x_n) \rightarrow \bot) \), where \( H \) is an atom or a conjunction of atoms and \( n \in \{0, 1, 2, \ldots\} \). Equivalently, a constraint can be written as \( \neg (\exists x_1 \cdots \exists x_n H(x_1, \ldots, x_n)) \). As an example of a constraint, \( N \) denotes the existential closure of the hypothesis of \( N \). A set is consistent if and only if it is not inconsistent. A set \( \{F_1, \ldots, F_k\} \subseteq F \) is \( R \)-inconsistent if and only if there exists a constraint \( N \in N \) such that \( \{F_1, \ldots, F_k\} \models H_N \), where \( H_N \) denotes the existential closure of the hypothesis of \( N \).

Let \( K = (F, R, N) \) where:

- \( F \) contains the following facts:
  - \( F_1 = \text{Bread(bleuette)} \land \text{PesticideFree(bleuette)} \)
  - \( F_2 = \exists e \text{ExtractionRate}(e, \text{bleuette}) \)
  - \( F_3 = \exists f (\text{FiberContent}(f, \text{bleuette}) \land \text{High}(f)) \)
- \( R \) consists of the following rules:
  - \( R_1 = \forall x,y (\text{Bread}(x) \land \text{ExtractionRate}(y,x) \land \text{PesticideFree}(x) \rightarrow \text{Decrease}(y)) \)
  - \( R_2 = \forall x,y,z (\text{Bread}(x) \land \text{ExtractionRate}(y,x) \land \text{FiberContent}(z,x) \land \text{High}(z) \rightarrow \text{Growth}(y)) \)
  - \( R_3 = \forall x (\text{Bread}(x) \land \text{ContaminantFree}(x) \rightarrow \text{PesticideFree}(x) \land \text{MycotoxinFree}(x)) \)
- \( N \) contains the following negative constraint:
  - \( N = \neg (\exists x (\text{Growth}(x) \land \text{Decrease}(x))) \)

\( K \) is consistent since \( (F, R) \models N \). Indeed, \( F_1 \) and \( R_3 \) allow to deduce \( \text{PesticideFree}(\text{bleuette}) \). Combined to \( F_2 \) and \( R_1 \) we obtain \( \text{Decrease}(e) \). \( F_3 \) and \( R_2 \) deduce \( \text{Growth}(e) \), violating the negative constraint \( N \)

Given a knowledge base, one can ask a conjunctive query in order to know whether something holds or not. Without loss of generality we consider boolean conjunctive queries (which are facts). As an example of a query, take \( \exists x_1 \text{cat}(x_1) \). The answer to query \( \alpha \) is positive if and only if \( F, R \models \alpha \).

Answering \( Q \), traditionally, has two different algorithmic approaches: either forward chaining or backwards chaining. The two approaches come to either (1) finding an answer of \( Q \) in the \( R \)-derivations of the facts in the knowledge base or (2) computing the inverse \( R \)-derivations of the query and finding if there is a match in the facts. We will focus on the latter approach in the following.

**B. Arguments and Attacks**

This section shows that it is possible to define an instantiation of Dung’s abstract argumentation theory [5] that can be used to reason with an inconsistent ontological KB.

We first define the notion of an argument. For a set of formulae \( G = \{G_1, \ldots, G_n\} \), notation \( \bigwedge G \) is used as an abbreviation for \( G_1 \land \ldots \land G_n \).

**Definition 1**: Given a knowledge base \( K = (F, R, N) \), an argument \( a \) is a tuple \( a = (F_0, F_1, \ldots, F_n) \) where:

- \( (F_0, \ldots, F_{n-1}) \) is a derivation sequence with respect to \( K \)
- \( F_n \) is an atom, a conjunction of atoms, the existential closure of an atom or the existential closure of a conjunction of atoms such that \( F_{n-1} \models F_n \).

This definition, following the definition of [17] is a straightforward way to define an argument, since an argument corresponds to a derivation.

To simplify the notation, from now on, we suppose that we are given a fixed knowledge base \( K = (F, R, N) \) and do not explicitly mention \( F, R \) nor \( N \) if not necessary. Let \( a = (F_0, \ldots, F_n) \) be an argument. Then, we denote \( \text{Supp}(a) = F_0 \) and \( \text{Conc}(a) = F_n \).

Arguments may attack each other, which is captured by a binary attack relation \( \text{Att} \subseteq \text{Arg}(F) \times \text{Arg}(F) \).

**Definition 2**: Let \( K = (F, R, N) \) be a knowledge base and let \( a \) and \( b \) be two arguments. The argument \( a \) attacks argument \( b \), denoted \( (a, b) \in \text{Att} \), if and only if there exists \( \varphi \in \text{Supp}(b) \) such that the set \( \{\text{Conc}(\varphi), \varphi\} \) is \( R \)-inconsistent.

This attack relation is not symmetric. To see why, consider the following example. Let \( F = \{p(m), q(m), r(m)\} \), \( R = \emptyset \), \( N = \{\forall x_1 (p(x_1) \land q(x_1) \land r(x_1) \rightarrow \bot)\} \). Let \( a = (\{p(m), q(m)\}, p(m) \land q(m)) \) and \( b = (\{r(m)\}, r(m)) \). We have \( (a, b) \in \text{Att} \) and \( (b, a) \notin \text{Att} \). This will ensure that the naive extension is different, at least in theory, from the preferred, stable, etc. semantics. However, in our application they all entail the same information as shown later on.

**Definition 3**: Given a knowledge base \( K = (F, R, N) \), the corresponding argumentation framework \( \text{AF}_K \) is a pair \( (A = \text{Arg}(F), \text{Att}) \) where \( \text{Arg}(F) \) is the set of all arguments that can be constructed from \( F \) and \( \text{Att} \) is the corresponding attack relation as specified in Definition 2.

Let \( E \subseteq A \) and \( a \in A \). We say that \( E \) is conflict free iff there exists no arguments \( a, b \in E \) such that \( (a, b) \in \text{Att} \). \( E \) defends \( a \) iff for every argument \( b \in A \), if we have \( (b, a) \in \text{Att} \) then there exists \( c \in E \) such that \( (c, b) \notin \text{Att} \).

\( E \) is admissible iff it is conflict free and defends all its arguments. \( E \) is a complete extension iff \( E \) is an admissible set which contains all the arguments it defends. \( E \) is a preferred extension iff it is maximal (with respect to set inclusion) admissible set. \( E \) is a stable extension iff it is conflict-free and for all \( a \in A \setminus E \), there exists an argument \( b \in E \) such that \( (b, a) \in \text{Att} \).

\( E \) is a grounded extension iff \( E \) is a minimal (for set inclusion) complete extension.

For an argumentation framework \( AS = (A, Att) \) we denote by \( \text{Ext}_x(AS) \) (or by \( \text{Ext}_x(A, Att) \)) the set of its extensions with respect to semantics \( x \). We use the abbreviations \( c, p, s, q \) for respectively complete, preferred, stable and grounded semantics.
An argument is sceptically accepted if it is in all extensions, credulously accepted if it is in at least one extension and rejected if it is not in any extension.

Based on this definition of arguments and attacks in [17] was also shown that the rationality postulates of [36] are respected. This instantiation respects the direct, indirect consistency as well as the closure.

C. Formalising the use case

In this subsection we formalise the notions presented in section IV.

Let \( K = (F, R, N) \) be a consistent knowledge base. This is the knowledge base that all actors share and agree upon. In this paper we assume that the rules and negative constraints are common to everybody.

The goals of the different actors can be seen as a set of existentially closed conjuncts. We denote them by \( G_1, G_2, ..., G_n \).

Let \( G_i \) be a goal and \( K \) the knowledge base. \( K \) is consistent and \( K \) does not entail \( G_i \). We compute the inverse \( R \)-derivations of \( G_i \) (where \( R \) is the set of rules of the knowledge base). We add all of the \( R^{-1}(G_i) \) to the facts. We thus obtain a new knowledge base \( K_i \), which differs from \( K \) solely by its facts set (which now also includes \( R^{-1}(G_i) \)).

We also impose that \( K_i \) is consistent.

Given \( G = \{G_1, G_2, ..., G_n\} \), the goals correspond to a set of viewpoints \( V \) (there exists a function \( \kappa : G \rightarrow 2^V \)). This function can assign a goal to one or more viewpoints and each viewpoint can be associated with one or more goals. Given a goal \( G_i \), the (set of) viewpoint(s) associated with this goal is denoted by \( \kappa(G_i) \). Similarly, given a viewpoint \( v_i \), the set of goals associated with it is denoted by \( \kappa^{-1}(v_i) \).

Example 1: Let the set of viewpoints \( V = \{\text{nutrition, sanitary, organoleptic}\} \) and \( G \) consisting of the following goals: 
\[ G_1 = \exists x \ (\text{Bread}(x) \land \text{LowSalt}(x)), \ G_2 = \exists x \ (\text{Bread}(x) \land \text{ContaminantFree}(x)), \ G_3 = \exists x \ (\text{Bread}(x) \land \text{Crusty}(x)) \]
\[ G_4 = \exists x \ (\text{Bread}(x) \land \text{TraceElementRich}(x)) \]

We have \( \kappa(G_1) = \kappa(G_3) = \text{nutrition}, \ \kappa(G_2) = \text{sanitary} \) and \( \kappa(G_4) = \text{organoleptic} \). Conversely \( \kappa^{-1}(\text{nutrition}) = \{G_1, G_3\}, \ \kappa^{-1}(\text{sanitary}) = \{G_2\} \) and \( \kappa^{-1}(\text{organoleptic}) = \{G_4\} \).

The rules will correspond to the set of sufficient conditions needed for the goal \( G_i \). In the context of our practical application this is illustrated in Figure 10 (with respect to nutrition goals).

Example 2: To reach the goal \( G_1 = \exists x \ (\text{Bread}(x) \land \text{LowSalt}(x)) \), the knowledge base \( K \) contains the following rule:
\[ \forall x,y \ (\text{Bread}(x) \land \text{SaltAdjunction}(y,x) \land \text{Decrease}(y) \Rightarrow \text{LowSalt}(x)) \]

Let us now consider the set of goals \( G = \{G_1, G_2, ..., G_n\} \) and the initial knowledge base \( K = (F, R, N) \). As described above we compute the \( n \) knowledge bases, corresponding to each goal: \( K_i = (F \cup R^{-1}(G_i), R, N) \) for each \( i = 1, ..., n \). We consider the union of all these knowledge bases:
\[ K_{agg} = (F \bigcup_{i=1}^{n} R^{-1}(G_i), R, N) \]

Example 3: Let \( K = (F, R, N) \) where :
- \( F = \{F_1\} = \{\text{CurrentExtractionRate(T65)}\} \)
- \( R \) contains the following rules:
  - \( R_1 = \forall x,y \ (\text{Bread}(x) \land \text{ExtractionRate}(y,x) \land \text{Decrease}(y) \Rightarrow \text{Digestible}(y)) \)
  - \( R_2 = \forall x,y \ (\text{Bread}(x) \land \text{SaltAdjunction}(y,x) \land \text{Decrease}(z) \Rightarrow \text{LowSalt}(x)) \)
  - \( R_3 = \forall x,y \ (\text{Bread}(x) \land \text{ExtractionRate}(y,x) \land \text{Growth}(y) \Rightarrow \text{TraceElementRich}(x)) \)
  - \( R_4 = \forall x,y \ (\text{Bread}(x) \land \text{ExtractionRate}(y,x) \land \text{Decrease}(y) \Rightarrow \text{PesticideFree}(x)) \)

Let the goal set \( G \) as follows:
- \( G_1 = \exists p \ (\text{Bread}(p) \land \text{Digestible}(p)), \ \kappa(G_1) = \text{nutrition} \)
- \( G_2 = \exists p \ (\text{Bread}(p) \land \text{LowSalt}(p)), \ \kappa(G_2) = \text{nutrition} \)
- \( G_3 = \exists p \ (\text{Bread}(p) \land \text{TraceElementRich}(p)), \ \kappa(G_3) = \text{nutrition} \)
- \( G_4 = \exists p \ (\text{Bread}(p) \land \text{PesticideFree}(p)), \ \kappa(G_4) = \text{sanitary} \)

Then:
- \( K_1 = (F, R, N) \) where \( F_1 = F \cup R^{-1}(G_1) \) contains the following facts:
  - \( F_1 = \text{CurrentExtractionRate(T65)} \)
  - \( F_2 = \text{Bread}(p) \land \text{ExtractionRate}(\tau,p) \land \text{Decrease}(\tau) \)
- \( K_2 = (F, R, N) \) where \( F_2 = F \cup R^{-1}(G_2) \) contains the following facts:
  - \( F_1 = \text{CurrentExtractionRate(T65)} \)
  - \( F_3 = \text{Bread}(p) \land \text{SaltAdjunction}(s,p) \land \text{Decrease}(s) \)
- \( K_3 = (F, R, N) \) where \( F_3 = F \cup R^{-1}(G_3) \) contains the following facts:
  - \( F_1 = \text{CurrentExtractionRate(T65)} \)
  - \( F_4 = \text{Bread}(p) \land \text{ExtractionRate}(\tau,p) \land \text{Growth}(\tau) \)
- \( K_4 = (F, R, N) \) where \( F_4 = F \cup R^{-1}(G_4) \) contains the following facts:
  - \( F_1 = \text{CurrentExtractionRate(T65)} \)
  - \( F_2 = \text{Bread}(p) \land \text{ExtractionRate}(\tau,p) \land \text{Decrease}(\tau) \)

Finally \( K_{agg} = (F \bigcup_{i=1}^{n} R^{-1}(G_i), R, N) \) where
\[ F \bigcup_{i=1}^{n} R^{-1}(G_i) = \{F_1, F_2, F_3, F_4\} \]

As observed in the previous example, it may happen that \( K_{agg} \) is inconsistent (and it does so even for goals belonging to the same viewpoint). We then use argumentation, which, by the means of extensions will isolate subsets of facts we can accept.
together (called extensions). Furthermore, the extensions will allow us to see which are the viewpoints associated to each maximal consistent subset of knowledge (by the means of the function $\kappa$). A choice procedure then has to be used (see example below).

The argument framework we can construct from the above knowledge base is $(\mathcal{A}, \mathcal{A})$ where $\mathcal{A}$ contains the following:

- $a = (\{F_2\}, F_2, R_1(F_2))$ where $R_1(F_2) = \text{Bread}(p) \land \text{ExtractionRate}(\tau, p) \land \text{Decrease}(\tau) \land \text{Digestible}(p)$.

- $b = (\{F_4\}, F_4, R_2(F_4))$ where $R_2(F_4) = \text{Bread}(p) \land \text{ExtractionRate}(\tau, p) \land \text{Growth}(\tau) \land \text{TraceElementRich}(p)$.

- $c = (\{F_2\}, F_2, R_4(F_2))$ where $R_4(F_2) = \text{Bread}(p) \land \text{ExtractionRate}(\tau, p) \land \text{Decrease}(\tau) \land \text{PesticideFree}(p)$.

- $d = (\{F_3\}, F_3, R_2(F_3))$ where $R_2(F_3) = \text{Bread}(p) \land \text{SaltAdjunction}(s, p) \land \text{Decrease}(s) \land \text{LowSalt}(p)$ and $\mathcal{A} = \{(a, b), (b, a), (b, c), (c, b)\}$.

In this argumentation system defined we now obtain:

- $\text{Ext}_{\text{stable}}(\mathcal{A}, \mathcal{A}) = \text{Ext}_{\text{semi-stable}}(\mathcal{A}, \mathcal{A}) = \text{Ext}_{\text{preferred}}(\mathcal{A}, \mathcal{A}) = \{\{a, c, d\}, \{b, d\}\}$.

Starting from the extensions $\text{Ext}_{\mathcal{A}}(\mathcal{A}, \mathcal{A})$, the proposed decision support system functions as follows: for every extension $\varepsilon \in \text{Ext}_{\mathcal{A}}(\mathcal{A}, \mathcal{A})$:

- Consider the facts occurring in the arguments of $\varepsilon$;
- Identify the knowledge bases $\mathcal{K}_i$ where these facts occur;
- Obtain the goals $G_i$ which are satisfied by the extension;
- Using the $\kappa$ function to obtain the viewpoints corresponding to these goals;
- Show domain experts the set of goals, and compatible viewpoints corresponding to the given extension.

This method allows us to obtain a set of options equal to the cardinality of $\text{Ext}_{\mathcal{A}}(\mathcal{A}, \mathcal{A})$. For taking a final decision several possibilities can be considered and presented to the experts:

- Maximise the number of goals satisfied;
- Maximise the number of viewpoints satisfied;
- Use preference relations of experts on goals and / or viewpoints.

In the previous example (please recall that the goals $G_1$ and $G_2$ are associated with the nutritional viewpoint while $G_4$ is associated with the sanitary viewpoint) we have:

- The first extension $\{a, c, d\}$ is based on the facts $F_2$ and $F_3$ obtained from $\mathcal{K}_1$, $\mathcal{K}_2$ and $\mathcal{K}_4$ that satisfy the goals $G_1$, $G_2$ and $G_4$.
- The second extension $\{b, d\}$ is based on $F_3$ and $F_4$ obtained from $\mathcal{K}_2$ and $\mathcal{K}_3$ satisfying $G_2$ and $G_3$ both associated with the nutritional viewpoint.

One first possibility (corresponding to the extension $\{a, c, d\}$) consists of accomplishing $F_2$ and $F_3$ and allows to satisfy the biggest number of goals and viewpoints.
The second possibility (corresponding to the extension \{b, d\}) consists of accomplishing \( F_2 \) and \( F_4 \). It would satisfy two goals and one viewpoint. It could be considered though if the goal \( G_3 \) (not satisfied by the first option) is preferred to the others.

VI. Evaluation

The evaluation of the implemented system was done via a series of interviews with domain experts. The above knowledge and reasoning procedures were implemented using the Cogui knowledge representation tool [16], with an extension of 2000 lines of supplemental code. Three experts have validated our approach: two researchers in food science and cereal technologies of the French national institute of agronomic research, specialists respectively of the grain-to-flour transformation process and of the breadmaking process, and one industrial expert - the Director of the French National Institute of Bread and Pastry.

The first meeting dealt with the delimitation of the project objectives and addressed fundamental questions such as: Is it possible to uniquely define a “good” bread? Which scenarios of “good bread” should be considered? How could they be defined from a nutritional, sanitary, sensorial and economic point of view? Which are the main known ways to achieve them?

Then a series of individual interviews constituted the elicitation phase. Each expert gave more arguments which were complementing one each other. In parallel, the writing of specifications for the demonstrator and the definition of the knowledge base structure were conducted.

In the following plenary meeting the real potential of the approach was shown. The experts were formulating goals and viewpoints they were interested in and the Cogui system together with the argumentation extension was yielding the associated possible propositions. Figure 11 shows a screenshot of the demonstrator answers for a two-goal query: a nutritional goal (high fiber content) and an organoleptic goal (crusty bread). Two sets of compatible actions are proposed, some choices (such as increasing or decreasing the extraction rate) being incompatible for both goals, and thus separated in the two alternative sets.

Four scenarios were more specifically evaluated. These scenarios concern four kinds of consumers: obeses (fiber preference), people with iron deficiency (micronutrient preference), people with cardiovascular disease (decreased salt preference) and vegetarians (limited phytic acid), which produces different sets of goals. For each scenario, the system proposes several outputted recommendations. The audience for decreasing salt tips the balance in favour of a recommendation for the T80 bread, while the audience for decreasing phytic acid pushes to specify recommendations towards a natural sourdough bread or a conservative T65 bread. Other audiences are in favor of a status quo. The results were considered as explainable by experts, but not obvious, since many considerations had to be taken into account.

Two interests of the approach were more particularly highlighted. They concern cognitive considerations. Firstly, experts were conscious that the elicitation procedure was done according to their thought processes, that is, in a forward way which is more natural and intuitive. The system was thus able to restitute the knowledge in a different manner than the experts usually do. Secondly, from a problem that could initially seem simple, the experts realized that it covered a huge complexity that a human mind could hardly address alone. The tool is currently available to them under restricted access.

The knowledge modeling task can be a very time-consuming step. As presented in Section IV-A, several sources of information were used, from peer reviewed scientific papers and technical reports, to conference meetings and expert interviews. On the one hand, expert interviews appeared to be the least expensive ones in terms of time. A one-day period allows both elicitating knowledge through an interview and formalizing it in the software system – which constitutes the longest part of the work. However, this relatively short time hides a strong prerequisite: having already a clear view of the case study, a synopsis of the questions to ask the expert and an implemented knowledge model. On the other hand, websites, technical reports and scientific articles are more costly to analyze. For instance, the critical reading a scientific paper of the domain may require a one-day period on its own, for a discerning reader. However they allow one to to grasp the ins and outs of the question.

During the evaluation step, the experts raised the question of the importance attached to the different pieces of knowledge modeled in the system. Moreover, in some cases experts may hesitate on the relevance of some facts or rules. A possibility would thus be to adopt a preference-based argumentation system, as proposed in several works such as [37], [38], [39], [40], [41], able to take into account different levels of importance among arguments.

VII. Conclusion

The breadmaking case exposed in this paper was studied using very eclectic and ubiquitous sources of information. It addresses a reverse engineering issue which is very generic to various kinds of processing systems and is not restricted to the agrifood context. It uses argumentation in an unusual way, as a computational method to determine compatible objectives and actions in the studied sector.

Computing best compatible set of actions was targeted as a result. We proposed to define it as optimizing the bipolar preferences expressed, which was a major outcome of this research.

This case study represents an original application and an introspective approach in the agronomy field by providing an argumentation based decision-support system for the various food or non-food sectors. It requires nevertheless the very expensive task of knowledge modeling. Such task, in its current state cannot be automated. It strongly depends on the quality of expert opinion and elicitation (exhaustiveness, certainty, etc). The current trend for decision-making tools includes more and more methods of argumentation as means of including experts in the task of modeling and the decision-making processes.

Another element to take into account, not discussed in this paper, is the difficulty of technologically (from an agronomy viewpoint) putting in place the facts of each option. Modeling this aspect in the formalism is still to be studied.
REFERENCES


