Underwater Archaeological 3D Surveys Validation within the Removed Sets Framework

Julien Hué[3], Mariette Sérayet[1], Pierre Drap[1], Odile Papini[1], and Eric Würbel[2]

[1] LSIS-CNRS 6168, Université de la Méditerranée, ESIL - Case 925, Av de Luminy, 13288 Marseille Cedex 9 France
{serayet,drap,papini}@esil.univmed.fr
[2] LSIS-CNRS 6168, Université du Sud Toulon -Var, BP 132, 83957 La Garde Cedex France
wurbel@univ-tln.fr
[3] CRIL-CNRS 8188, Université d'Artois, Rue Jean Souvraz SP 18, 62307 Lens Cedex France
hue@cril.univ-artois.fr

This paper presents the results of the VENUS european project aimed at providing scientific methodologies and technological tools for the virtual exploration of deep water archaeological sites. We focused on underwater archaeological 3D surveys validation problem. This paper shows how the validation problem has been tackled within the Removed Sets framework, according to Removed Sets Fusion (RSF) and to the Partially Preordered Removed Sets Inconsistency Handling (PPRSIH). Both approaches have been implemented thanks to ASP and the good behaviour of the Removed Sets operations is presented through an experimental study on two underwater archaeological sites.

1 Introduction

The VENUS European Project (Virtual ExploratioN of Underwater Sites, IST-034924)[1] aimed at providing scientific methodologies and technological tools for the virtual exploration of deep underwater archaeology sites. In this context, digital photogrammetry is used for data acquisition. The knowledge about the studied objects is provided by both archaeology and photogrammetry. One task of the project was to investigate how artificial intelligence tools could be used to perform reasoning with underwater archaeological 3D surveys. More specifically, this task focused on the validation problem of underwater artefacts 3D surveys. Within this project two different conceptual descriptions of the surveyed artefacts have been proposed leading to two different solutions both developed within the Removed Sets framework. This syntactic approach is more suitable than a semantic one, in order to pinpoint the errors that cause inconsistency. The present paper provides a synthesis of these two solutions. The first solution stems from the Entity Conceptual Model for modeling generic knowledge and uses instanciated predicate logic as representation formalism and Removed Sets Fusion (RSF) with *Sum* strategy for reasoning [9]. The second one is based on an application ontology for modeling generic knowledge and the belief base is represented in instanciated predicate logic equipped with a partial preorder

[1] http://www.venus-project.eu

W. Liu (Ed.): ECSQARU 2011, LNAI 6717, pp. 663–674, 2011.

and Partially Preordered Removed Sets Inconsistency Handling (PPRSIH) for reasoning [17]. The paper is organized as follows. After describing in Section 2 the validation problem in the context of the VENUS project, Section 3 gives a brief synthetic presentation of the Removed Sets framework. Section 4 shows how the validation problem is expressed as a RSF problem while Section 5 shows that how the validation problem can be reduced to a PPRSIH problem. Finally, Section 6 discusses the results of the experimental study before concluding.

2 The Validation Problem in VENUS

In the context of the VENUS project, digital photogrammetry is used for data acquisition. Usual commercial photogrammetric tools only focus on geometric features and do not deal with the knowledge concerning the surveyed objects. The general goal is the integration of knowledge about surveyed objects into the photogrammetric tool ARPENTEUR [5] in order to provide more "intelligent" 3D surveys. In this project, we investigated how Artificial Intelligence tools can be used for representing and reasoning with 3D surveys information.

Within the context of underwater archaeological surveys, we deal with information of different nature. Archaeologists provide expert knowledge about artefacts, in most of the cases amphorae. Archaeological knowledge takes the form of a characterization of amphorae thanks to a typology hierarchically structured. For each type corresponds a set of features or attributes which we assign an interval representing the expected values for an amphora of this type.

The data acquisition process provides measures coming from the photogrammetric restitution of surveyed amphorae pictures on the underwater site (see ① in figure 1). These observations usually are uncertain, inaccurate or imprecise since the pictures are taken "in situ", their quality could not be optimal, because of the hostile environment: weather conditions, visibility, water muddying, site not cleaned, ... Moreover, errors could occur during the restitution step. For all these reasons, the archaeological knowledge (see ② in figure 1) and the data coming from the photogrammetric acquisition process could conflict. This special case of inconsistency handling is a validation problem because the measured values of attributes of a surveyed amphora "in situ", an instance, may not fit with the characterization of the amphorae type it is assumed to belong to. The VENUS project does not use image recognition. The generic knowledge is inserted in the system by the experts. There is no automatic image recognition since the experts recognise the objects in the image during the measuring step thanks to their a priori knowledge.

Example 1. We illustrate the validation problem with the Pianosa island site [12]. There are 8 types of amphorae: *Dressel20*, *Beltran2B*, *Gauloise_3*, ... and each type of amphorae is characterized by 9 attributes, *totalHeight*, *totalWidth*, *totalLength*, *bellyDiameter*, *internalDiameter*, ... [14]. However, the only measurable attributes are *totalHeight*, *totalLength*[2]. Default values for these attributes

[2] For amphorae the attributes *totalWidth* and *totalLength* have the same value since there are revolution objects.

take the form of a range of values $[v - v.t\%, v + v.t\%]$ centered around a typical value v (expressed in m.) where t is a tolerance threshold. For example, the default values for the attributes *totalHeight* and *totalLength* for the Dressel20 type are [0.5328, 0.7992] and [0.368, 0.552], while for a Beltran2B type they are [0.9008, 1.3512] and [0.3224, 0.4836]. Suppose, during the photogrammetric restitution process, the expert focuses on a given amphora, he recognizes as a *Beltran2B*. When the survey provides the values 1.13 as *totalHeight* and 0.27 as *totalLength*, the question is do these values fit with the characterization of the *Beltran2B*? When the values do not fit, the most probable reason is that the measures are incorrect due to bad conditions of acquisition.

In order to provide a qualitative representation of this validation problem, a conceptual description of archaeological knowledge is required (see ② in figure 1). Several conceptual descriptions have been used within the VENUS project. At the beginning of the project, we used a object oriented conceptual description, restricted form of the Entity Model approach [16]. The restricted Entity Model is denoted by $\mathcal{E} = \{\mathcal{C},\ \mathcal{V}^d,\ \mathcal{C}^I\}$ where $\mathcal{C}$ is a concept (or a class), $\mathcal{V}^d$ is the set of default values for the attributes, $\mathcal{C}^I$ is a set of constraints on attributes. The concepts are the types of amphorae surveyed on the archaeological site. For each concept, that is each type of amphorae, we represent the measurable attributes. The default values for these attributes take the form of a range of values and $\mathcal{V}^d$ is a set of intervals, each interval corresponding to the possible values of attributes for a given type of amphorae. The set of constraints on the attributes $\mathcal{C}^I$ consists in integrity contraints, domain constraints and conditional constraints which express the compatibility of the measured values of attributes with the default values of attributes for a given type. The belief profile consists of the generic knowledge according to the restricted Entity Model provided by the `typology.xml` file and of the instances of amphorae provided by the `amphora.xml` file.

During the project, we constructed an application ontology [13] from a domain ontology which describes the vocabulary on the amphorae (the studied artefacts) and from a task ontology describing the data acquisition process. This ontology consists of a set of concepts, relations, attributes and constraints like domain constraints. The belief base contains the application and ontology, constraints and observations. The ontology represents the generic knowlegde which is preferred to observations. Due to the lack of space, we only consider a small part of the ontology (Figure 2).

3 The Removed Sets Framework

The Removed Sets framework provides a syntactic belief change approach for revision and fusion. When dealing with belief change operations since we deal with uncertain, incomplete, dynamic information, inconsistency can result. In order to provide a consistent result of the change operation, the Removed Sets approach focuses on the minimal set of formulae to remove, called *Removed sets*, in order to restore consistency. The Removed Sets operations have been proved to be equivalent to the ones based on maximal consistent subsets [15,4,1]. However, in the context of applications where few inconsistencies may occur, the Removed Sets approach seems to be more efficient when implementing large belief bases.

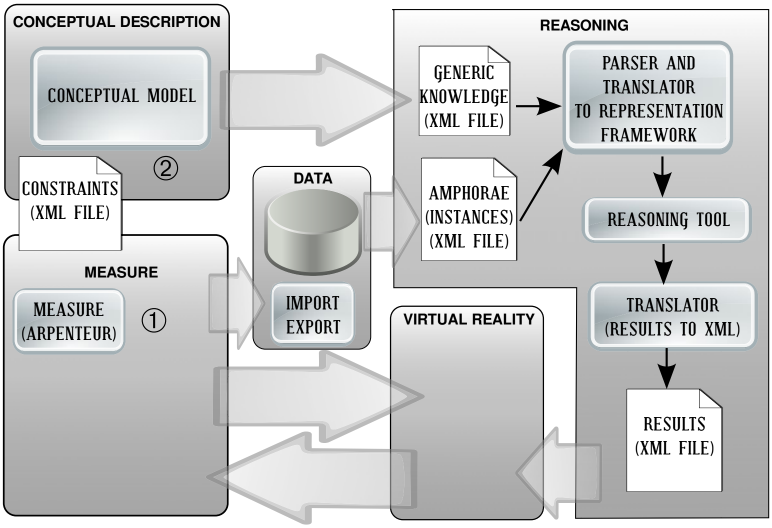

Fig. 1. General scheme

Initially, the Removed Sets approach has been proposed for revising propositional formulae in CNF (RSR [11,18]). It has then been generalized to arbitary propositional formulae for revision and fusion (RSF [9]). The Removed Sets approach has been extended to totally preordered belief bases (PRSR [2]), (PRSF [8]) and more recently to partially preordered belief bases for revision (PPRSR [17]). A central notion is the one of potential Removed Set[3] which are sets of formulas whose removal restores consistency into the union of belief bases.

Definition 1. *Let* $E = \{K_1, \ldots, K_n\}$ *be a belief profile such that* $K_1, \ldots, K_n$ *are propotional belief bases and* $K_1 \sqcup \ldots \sqcup K_n$ *is inconsistent (*$\sqcup$ *denotes set union with accounting for repetitions).* $X \subseteq K_1 \sqcup \ldots \sqcup K_n$ *is a* potential Removed Set *of* E *if and only if* $(K_1 \sqcup \ldots \sqcup K_n) \backslash X$ *is consistent.*

The collection of potential Removed Sets of E is denoted by $\mathcal{PR}(E)$. Since the number of potential Removed Sets of E is exponential w.r.t. the number of formulae, we only consider the minimal potential Removed Sets w.r.t. set inclusion. Moreover belief change operations or belief change strategies are formalized in terms of total preorders or partial preorders on potential Removed Sets minimal w.r.t. set inclusion.

3.1 Removed Sets Fusion

For Removed Sets Fusion, the fusion strategies ($Card$, Sum, Max, $GMax$) are formalized thanks to a total preorder over $\mathcal{PR}(E)$. Let X and Y be two potential

[3] We give the definitions in the general setting of fusion where revision is a special case.

Removed Sets, for each strategy P a total preorder $\leq_P$ over the potential Removed Sets is defined. $X \leq_P Y$ means that X is preferred to Y according to the strategy P. We define $<_P$ as the strict total preorder associated to $\leq_P$ (i.e. $X <_P Y$ if and only if $X \leq_P Y$ and $Y \not\leq_P X$).

Definition 2. *Let $E = \{K_1, \ldots, K_n\}$ be a belief profile such that $K_1 \sqcup \ldots \sqcup K_n$ is inconsistent. $X \subseteq K_1 \sqcup \ldots \sqcup K_n$ is a* Removed Set *of E according to the strategy P if and only if i) X is a potential Removed Set of E; ii) $\nexists X' \in \mathcal{PR}(E)$ such that $X' \subset X$; iii) $\nexists X' \in \mathcal{PR}(E)$ such that $X' <_P X$.*

The collection of Removed Sets of E according to the strategy P is denoted by $\mathcal{R}_P(E)$. The Removed Sets Fusion operation is defined by:

Definition 3. *Let $E = \{K_1, \ldots, K_n\}$ such that $K_1 \sqcup \ldots \sqcup K_n$ is inconsistent. The merging operation is defined by: $\Delta_P^{RSF}(E) = \bigcup_{X \in \mathcal{R}_P(E)} \{(K_1 \sqcup \ldots \sqcup K_n) \backslash X\}$.*

3.2 Partially Preordered Removed Sets Inconsistency Handling

Let K be a finite set of arbitrary formulae and $\preceq_K$ be a partial preorder on K. Restoring the consistency of a partially preordered belief bases involves the definition of a partial preorder on subsets of formulae, called comparators [3,19]. Several ways have been proposed for defining a preference relation on subsets of formulae of K, from a partial preorder $\preceq_K$. In the VENUS project, we focus on the lexicographic preference [19] which extends the lexicographic preorder initially defined for totally preordered belief bases to partially preordered belief bases. The belief base K is partitionned such that $K = E_1 \sqcup \ldots \sqcup E_n$ $(n \geq 1)$ where each subset E_i represents an equivalence class of K with respect to $=_K$ which is an equivalence relation. A preference relation between the equivalence classes E_i's, denoted by $\prec_s$ is defined by $E_i \prec_s E_j$ iff $\exists \varphi \in E_i$, $\exists \varphi' \in E_j$ such that $\varphi \prec_K \varphi'$. This partition can be viewed as a generalization of the idea of stratification defined for totally preordered belief bases. We rephrase the lexicographic preference defined in [19] as follows:

Definition 4. *Let $\preceq_K$ be a partial preorder on K, $Y \subseteq K$ and $X \subseteq K$. Y is said to be lexicographically preferred to X, denoted by $Y \trianglelefteq_\vartriangle X$, iff $\forall i$, $1 \leq i \leq n$: if $|E_i \cap Y| > |E_i \cap X|$ then $\exists j$, $1 \leq j \leq n$ such that $|E_j \cap X| > |E_j \cap Y|$ and $E_j \prec_s E_i$.*

Let $\mathcal{PR}(K)$ be the set of potential removed sets. Among them, we want to prefer the potential removed sets which allow us to remove the formulae that are not preferred according to $\preceq_K$. Therefore we generalize the notion of Removed Sets to subsets of partially preordered formulae. We denote by $\mathcal{R}_\vartriangle(K)$ the set of removed sets of K.

Definition 5. *Let K be an inconsistent belief base equipped with a partial preorder $\preceq_K$. $R \subseteq K$ is a removed set of K iff i) R is a potential removed set; ii) $\nexists R' \in \mathcal{R}_\vartriangle(K)$ such that $R' \subset R$; iii) $\nexists R' \in \mathcal{R}_\vartriangle(K)$ such that $R' \vartriangleleft_\vartriangle R$.*

Definition 6. *Let K be an inconsistent belief base equipped with a partial perorder $\preceq_K$. Restoring the consistency leads to a consistent belief base K' such that $K' = \bigcup_{X \in \mathcal{R}_\vartriangle(K)} \{K \backslash X\}$.*

3.3 ASP Implementation

In order to implement belief change operations within the Removed sets framework, we translate the belief change problem into a logic program with answer set semantics. This method proceeds in two stages. The first stage consists in the translation of E into a logic program Π_E and we have shown that the answer sets of Π_E correspond to the potential removed sets of E [9].

Let E be a belief profile[4]. Each propositional variable a occuring in E is represented by an ASP atom $a \in \mathcal{A}$ in Π_E. The set of all positive, (resp. negative) literals of Π_E is denoted by V^+, (resp. V^-). The set of rule atoms representing formulae is defined by $R^+ = \{r_f \mid f \in E\}$ and $F_O(r_f)$ represents the formula of E corresponding to r_f in Π_E, namely $\forall r_f \in R^+, F_O(r_f) = f$. This translation requires the introduction of intermediary atoms representing subformulae. We denote by ρ_f^j the intermediary atom representing f^j which is a subformula of $f \in E$. The first part of the construction has two steps:

1. We introduce rules in order to build a one-to-one correspondence between answer sets of Π_E and interpretations of V^+. For each atom, $a \in V^+$ two rules are introduced: $a \leftarrow not\ a'$ and $a' \leftarrow not\ a$ where $a' \in V^-$ is the negative atom corresponding to a.
2. We introduce rules in order to exclude the answer sets S corresponding to interpretations which are not models of $(E \backslash F)$ with $F = \{f \mid r_f \in S\}$. According to the syntax of f, the following rules are introduced: (i) If $f =_{def} a$ then $r_f \leftarrow not\ a$ is introduced; (ii) If $f =_{def} \neg f^1$ then $r_f \leftarrow not\ \rho_{f^1}$ is introduced; (iii) If $f =_{def} f^1 \vee \ldots \vee f^m$ then $r_f \leftarrow \rho_{f^1}, \ldots, \rho_{f^m}$ is introduced; (iv) If $f =_{def} f^1 \wedge \ldots \wedge f^m$ then it is necessary to introduce several rules: $\forall 1 \leq j \leq m$, $r_f \leftarrow \rho_{f^j}$.

This stage is common to any belief change operation while the next one depends on the chosen belief change operation.

In case of fusion the second stage provides, according to selected strategy P, another set of rules that leads to the program Π_E^P and we have shown [9] that the answer sets of Π_E^P correspond to the removed sets of E for a strategy P. In the validation problem since we have to minimize the number of formulae to remove, therefore the number of formulae occuring in a removed set, we select the Sum strategy. This strategy is expressed by the $minimize\{\}$ statement and the new logic progam $\Pi_E^{Sum} = \Pi_E \cup minimize\{r_f \mid r_f \in R^+\}$ is such that the answer sets of Π_E^{Sum} which are provided by the CLASP solver [7] correspond to the removed sets of $\Delta_{Sum}^{RSF}(E)$ [9].

In case of partially Preordered Removed Sets Inconsistency Handling the CLASP solver [7] gives the answer sets of Π_E. We then construct a partial preorder between them using the lexicographic comparator $\trianglelefteq_\Delta$. We have shown in [17] that the preferred answer sets according to $\trianglelefteq_\Delta$ correspond to the removed sets of E. We used a java program to partially preorder the answer sets to obtain the preferred answer sets. Since the lexicographic comparator satisfies the monotony property [19], it is sufficient to compare the answer sets which are minimal according to the inclusion. Moreover, the determination of the minimal

[4] In case of inconsistency handling the profile E is reduced to a belief base K.

answer sets according to this partial preorder does not increase the computationnal cost, since this cost is insignificant compared to the cost of answer sets computation by CLASP.

4 The Validation Problem within RSF

In order to represent the validation problem within the RSF framework and to implement it with ASP, we represented this problem with instanciated predicate logic. The belief profile consists of two belief bases. The first one stems from the restricted Entity Model conceptual description and represents the generic knowledge. We introduce the predicates $type(x, y)$ and $cmp(z, y, x)$ where x is an amphora, y is a type of amphorae and z is an attribute. $type(x, y)$ expresses that an amphora x belongs to a type y and $cmp(z, y, x)$ expresses that an attribute z of an amphora x of type y has a value compatible with the possible values for the type y, as specified in 2. The domain constraints specify that an amphora must have one and only one type. For n types of amphorae, for each amphora there is one disjunction $type(x, y_1) \vee \ldots \vee type(x, y_n)$ and $n(n-1)/2$ mutual exclusion formulae $\neg type(x, y_i) \vee \neg type(x, y_j)$. The conditional constraints specify the compatibility of the attributes values with respect to the type. For each amphora x, for each attribute z and for each type y, there is a formula $type(x, y) \rightarrow cmp(z, y, x)$. Let m be the number of attributes, the incompatibility of type specifies that for each amphora and each type there is a formula $\neg cmp(z_1, y, x) \wedge \ldots \wedge \neg cmp(z_m, y, x) \rightarrow \neg type(x, y)$. The second belief base represents the instances of amphorae: the type the observed amphora belongs to (namely $type(x, y)$) and the compatible attributes with the type (namely $cmp(z, y, x)$). We illustrate the RSF approach with the example 1.

Example 2. We limit ourselves to only two types of amphorae $Beltran2B$ and $Dressel20$, respectively denoted by $B2B$ and $D20$ thereafter, and to the survey of one observed amphora (denoted by 4 hereafter). Two attributes are used: totalHeight (denoted by tH) and totalLength (denoted by tL). The first belief base is automatically generated from the typology.xml file and $K_1 = \{\neg type(4, B2B) \vee \neg type(4, D20),\ type(4, B2B) \vee type(4, D20),\ type(4, D20) \rightarrow cmp(tH, D20, 4),\ type(4,\ D20) \rightarrow cmp(tL, D20, 4),\ type(4, B2B) \rightarrow cmp(tH,\ B2B,\ 4),\ type(4, B2B) \rightarrow cmp(tL,\ B2B, 4),\ \neg cmp(tH, B2B, 4) \wedge \neg cmp(tL,\ B2B, 4) \rightarrow \neg type(4, B2B),\ \neg cmp(tH, D20, 4) \wedge \neg cmp(tL, D20, 4) \rightarrow \neg type(4, D20)\ \}$. The second belief base corresponding to the observed amphora is automatically generated from typology.xml and amphora.xml files and $K_2 = \{type(4, B2B),\ cmp(tH, B2B, 4)\}$. The operation $\Delta^{RSF}_{Sum,\top}(E)$ where $E = \{K_1, K_2\}$ is translated into Π^{Sum}_E as follows:

$cmp(tH, B2B, 4).$	$1\ \{type(4, d20), type(4, B2B)\}\ 1.$
$r(x_0) \leftarrow not\ type(4, B2B).$	$n\_type(4, B2B) \leftarrow r(x_0).$
$r(x_5) \leftarrow type(4, d20), not\ cmp(tH, d20, 4), not\ cmp(tL, d20, 4).$	$\leftarrow n\_type(4, d20), type(4, d20).$
$r(x_1) \leftarrow type(4, d20), not\ cmp(tH, d20, 4).$	$n\_cmp(tH, d20, 4) \leftarrow r(x_1).$

$$r(x_2) \leftarrow type(4, d20), not\ cmp(tL, d20, 4). \qquad n\_cmp(tL, d2, 4) \leftarrow r(x_2).$$
$$r(x_6) \leftarrow type(4, B2B), not\ cmp(tH, B2B, 4), \ \leftarrow n\_type(4, B2B), type(4, B2B).$$
$$not\ cmp(tL, B2B, 4).$$
$$r(x_3) \leftarrow type(4, B2B), not\ cmp(tH, B2B, 4). \ n\_cmp(tH, B2B, 4) \leftarrow r(x_3).$$
$$r(x_4) \leftarrow type(4, B2B), not\ cmp(tL, B2B, 4). \ n\_cmp(tL, B2B, 4) \leftarrow r(x_4).$$
$$minimize\ \{r(x_0), r(x_1), r(x_2), r(x_3), r(x_4)r(x_5), r(x_6)\}.$$

Note that the ASP translation uses some shortcuts compared to the translation scheme depicted in section 3.3. Thanks to the cardinality literals by recent ASP solvers, the unique type constraint is reduced to a single rule $1\ \{type(4, d20), type(4, B2B)\}\ 1$. Also, the generation of the rule corresponding to $type(4, B2B)$ and the mutual exclusion between this atom and its classical negation are compacted into a single rule.

The only answer set of the above program is $\{cmp(tH, B2B, 4), type(4, B2B), r(x_4), n\_cmp(tL, B2B, 4)\}$ which corresponds to the removed set $\{type(4, B2B) \rightarrow cmp(tL, B2B, 4)\}$ that pinpoints a bad measure for the total length attribute under the hypothesis of an amphora of type $Beltran2B$.

5 The Validation Problem within PPRSIH

The conceptual description in this approach is represented in terms of an application ontology and an extract is illustrated in Figure 2.

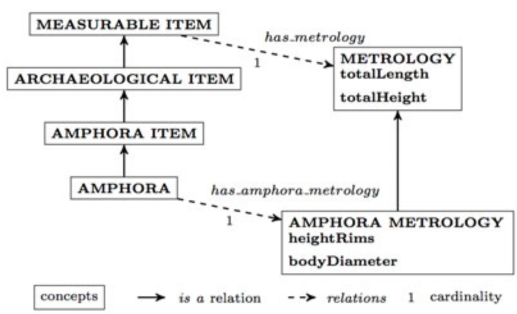

Fig. 2. Extract of the application ontology

The belief base consists of the application ontology, the constraints and the instances of amphorae represented in predicate logic. The introduced predicates are shown in an instanciated version in Table 1. The formulae corresponding to the extract of the ontology are given below where $amph$, $amph\_it$, $arch\_it$, $meas\_it$, $metro$, has_metro, tL, tH, $type$ denote $amphora$, $amphora\_item$, $archaeological\_item$, $measurable\_item$, $metrology$, $has\_metrology$, $totalLenght$, $total\ Height$, $typology$ respectively: $\forall x\ arch\_it(x) \rightarrow meas\_it(x)$, $\forall x\ amph\_it(x) \rightarrow arch\_it(x)$, $\forall x\ amph(x) \rightarrow amph\_it(x)$, $\forall x\ meas\_it(x) \rightarrow \exists z\ has\_metro(x, z)$, $\forall x \forall z\ has\_metro(x, z) \rightarrow metro(z)$, $\forall z\ \ metro(z) \rightarrow \exists l\ tL(z, l) \wedge \exists h\ tH(z, h)$, $\forall x\ amph(x) \rightarrow amph\_it(x) \wedge (type(x, y_1) \vee \cdots \vee type(x, y_n))$. The set of constraints consists in integrity constraints which specify that the value of attributes do not exceed a given value, domain constraints are specified by cardinality constraints within the application ontology and conditional constraints express the compatibility of the attribute values with respect to the type. The domain constraints are expressed like in Section 4 by one disjunction $\forall x\, type(x, y_1) \vee \cdots \vee type(x, y_n)$ and $n(n-1)/2$ mutual exclusion formulae $\neg type(x, y_i) \vee \neg type(x, y_j)$. The integrity constraints are expressed by the formulae: $\forall x\; meas\_it(x) \rightarrow \exists z \exists h(tH(z, h) \wedge cmpMItH(h, x))$, $\forall x\; meas\_it(x) \rightarrow \exists z \exists l(tL(z, l) \wedge cmpMItL(l, x))$, $\forall x\, arch\_it(x) \rightarrow \exists z \exists h(tH(z, h) \wedge cmpARItH(h, x))$, $\forall x\; arch\_it(x) \rightarrow \exists z\; \exists l\; (tL(z, l) \wedge cmpARItL(l, x))$, $\forall x\, amph\_it(x) \rightarrow \exists z \exists h(tH(z, h) \wedge cmpAItH(h, x))$, $\forall x\, amph\_it(x) \rightarrow \exists z \exists l(tL(z, l) \wedge cmpAItL(l, x))$. The conditional constraints are expressed by the formulae: $\forall x\, type(x, y_i) \rightarrow \exists z \exists h(tH(z, h) \wedge cmptH(h, y_i))$ $\forall x, type(x, y_i) \rightarrow \exists z \exists l(tL(z, l) \wedge cmptL(l, y_i))$. The formulae corresponding to the instances of amphorae are $amph(x)$, $type(x,y)$, $metro(z)$, $meas\_it(x)$, $arch\_it(x)$, $amph\_it(x)$, $has\_metro(x, z)$, $tL(z, l) \wedge cmpMItL(l, x) \wedge cmpARItL(l, x) \wedge cmpAItL(l, x) \wedge \neg cmptL(l, y_i)$ and $tH(z, h) \wedge cmpMItH(h, x) \wedge cmpARItH(h, x) \wedge cmpAItH(h, x) \wedge \neg cmptH(h, y_i)$. The belief base is equipped with a partial preorder which reflects the hierarchy of concepts in the ontology. Moreover constraints are preferred to the ontology which is preferred to the instances. We illustrate the PPRSIH approach thanks to example 1.

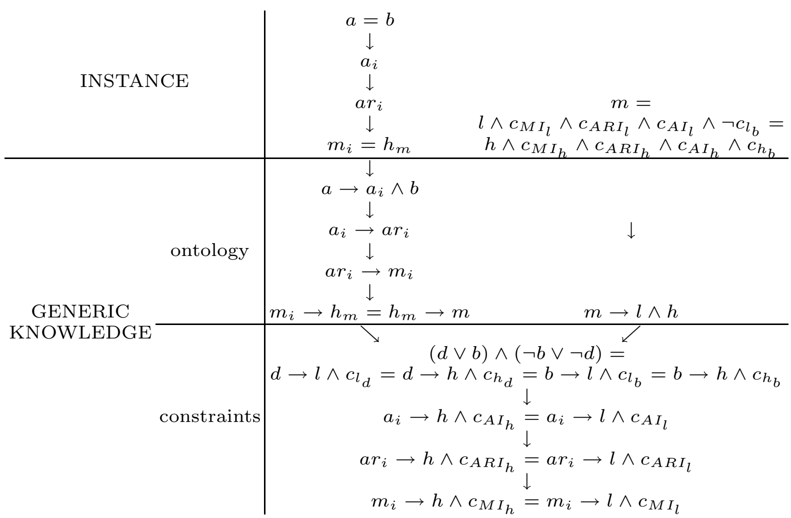

Fig. 3. Partial preorder on formulae of the belief base

Example 3. We limit ourselves to the amphorae types $Beltran2B$ and $Dressel20$ and to the survey of the observed amphora denoted by 4. Table 1 presents the instanciated predicates and Figure 3 illustrates the partially preordered belief base.

The validation problem is translated into a logic progam Π_E in the same spirit than the one presented in section 3.3. CLASP provides 1834 answer sets. However, if only focusing on the minimal answer sets with respect to inclusion we have to partially preorder 320 answer sets. According to the lexicographic

Table 1. Instanciated predicates and their corresponding proposition p

predicate	p	predicate	p	predicate	p
$meas\_it(4)$	m_i	$arch\_it(4)$	ar_i	$amph\_it(4)$	a_i
$amph(4)$	a	$metro(m)$	m	$type(4, Dressel20)$	d
$type(4, Beltran2B)$	b	$has\_metro(4, m)$	h_m	$tL(m, l)$	l
$tH(m, h)$	h	$cmpMItL(l, 4)$	c_{MI_l}	$cmpARItL(l, 4)$	c_{ARI_l}
$cmpAItL(l, 4)$	c_{AI_l}	$cmpMItH(h, 4)$	c_{MI_h}	$cmpARItH(h, 4)$	c_{ARI_h}
$cmpAItH(h, 4)$	c_{AI_h}	$cmptL(l, Dressel20)$	c_{l_d}	$cmptH(h, Dressel20)$	c_{h_d}
$cmptL(l, Beltran2B)$	c_{l_b}	$cmptH(h, Beltran2B)$	c_{h_b}	$amph(4)$	a
$type(4, Beltran2B)$	b	$metro(am)$	a_m		

comparator $\trianglelefteq_\Delta$, we obtain two uncomparable preferred answer sets S_1 and S_2 such that $F_O(S_1 \cap R^+) = \{a, b\}$ and $F_O(S_2 \cap R^+) = \{l \wedge c_{ARI_l} \wedge c_{AI_l} \wedge c_{MI_l} \wedge \neg c_{l_b}\}$. Therefore, there are two removed sets $R_1 = \{a, b\}$ and $R_2 = \{l \wedge c_{ARI_l} \wedge c_{ARI_l} \wedge c_{MI_l} \wedge \neg c_{l_b}\}$. The removed set R_1 pinpoints the typology while R_2 pinpoints that the value of TotalLength attribute may be wrong. This approach provides 2 removed sets while the RSF one only provides one removed set. The reason is that in PPRSIH approach the typology is only suspected if the value of one of the attributes is incompatible while in RSF approach the typology is suspected if the values of more than one attributes are incompatible.

6 Concluding Discussion

We now present the results of the experimental study, first on the full Pianosa survey which contains 40 amphorae then on the Port-Miou survey which contains 500 amphorae. We used 4 different tolerance thresholds t around the typical values of each type: 20%, 10%, 5% and 1% and N denotes the number of inconsistent amphorae. The CPU times $T1$, $T2$, $T3$ and T correspond to the translation from the XML files to the logic program, the ASP implementation of RSF, the translation from ASP to an XML file and the total time $T1 + T2 + T3$ respectively. The tests were conducted on a Centrino Duo cadenced at 1.73GHz and equipped with 2GB of RAM. The results are summarized in Table 2.

Table 2. CPU times (s) for RSF and PPRSIH on two surveys

(a) Pianosa survey (40 amphorae)

		RSF				PPRSIH			
t	**N**	**T1**	**T2**	**T3**	**T**	**T1**	**T2**	**T3**	**T**
20	5	0.05	0.62	0.95	1.62	0.24	1.12	0	1.36
10	26	0.05	0.60	0.64	1.29	0.27	5.13	0	5.40
5	30	0.05	0.61	0.45	1.11	0.29	5.87	0	6.16
1	36	0.05	0.60	0.33	0.98	0.31	6.91	0	7.22

(b) Port Miou survey (500 amphorae)

	RSF				PPRSIH			
N	**T1**	**T2**	**T3**	**T**	**T1**	**T2**	**T3**	**T**
44	0.43	5.26	0.14	5.83	0.68	9.38	0	10.06
65	0.43	5.06	0.04	5.53	0.75	13.69	0	14.44
72	0.43	4.99	0	5.42	0.81	15.03	0	15.84
81	0.43	5.06	0	5.49	0.88	16.80	0	17.68

Concerning the knowledge representation aspect the RSF approach stems from the Entity Model conceptual description and uses instanciated predicate logic. It creates a flat knowledge base, with numerous formulae, where all the objects are at the same level. In the full Pianosa survey involving 40 amphorae, the traduction of the problem requires 8462 formulae and 4160 atoms and in the full Port Miou involving 500 amphorae, the traduction of the problem requires 105775 formulae and 52000 atoms. Moreover, it only considers the intrinsic constraints between objects. However, the lack of expressivity and the high number of formulae are compensated by the good computational behaviour of the reasoning tasks expressed in this language. The PPRSIH approach stems from the application ontology and uses instanciated predicate logic equipped with a partial preorder. It creates a more structured belief base, involving less formulae than the first approach. In the full Pianosa survey involving 40 amphorae, the traduction of the problem requires 1080 formulae and 840 atoms and in the full Port-Miou survey involving 500 amphorae and the traduction requires 6021 formulae and 4683 atoms. It allows for representing the intrinsic constraints as well as the taxonomic relations between objects, and relations between objects. The partial preorder defined on the finite set of formulae expresses more structure than the first solution. This approach takes advantage of the good computational behaviour of instanciated predicate logic while expressing, in the same time, a more structured belief base.

Concerning the reasoning aspect, both implementations rely on CLASP which is one of the most efficient current ASP solver. The results obtained on Pianosa as well as on Port Miou survey given in Table 2 clearly show that both approaches deal with the full survey with a very good time. However, the first solution gives the best running times. Moreover, reducing the tolerance intervals increases the number of inconsistencies as illustrated in table 2 and the first solution seems to be not sensitive to this increasing while the running time of the second solution grows with the number of inconsistencies. The consuming task comes from the reading of the answer sets before partially ordering them in order to only select the preferred ones. In order to improve this approach we have to investigate how to directly encode the partial preorder on answer sets within the logic program. Another direction to follow in order to reach a trade-off between representation and reasoning could be to represent the validation problem in Description Logic, since the generic knowledge in expressed in terms of ontology. However, we have to study which low complexity Description Logic could be suitable. Moreover, we have to study to which extent the approach combining Description logic and ASP [6] could be used for implementation as well as the extended ASP solver to first order logic[10].

Acknowledgements

Work partially supported by the European Community under project VENUS (Contract IST-034924) of the "Information Society Technologies (IST) program of the 6th FP for RTD". The authors are solely responsible for the content of this paper. It does not represent the opinion of the European Community, and the European Community is not responsible for any use that might be made of data appearing therein.

References

1. Baral, C., Kraus, S., Minker, J., Subrahmanian, V.S.: Combining knowledge bases consisting of first order theories. In: Proc. of ISMIS, pp. 92–101 (1991)
2. Benferhat, S., Ben-Naim, J., Papini, O., Würbel, E.: An answer set programming encoding of prioritized removed sets revision: application to gis. Applied Intelligence 32(1), 60–87 (2010)
3. Benferhat, S., Lagrue, S., Papini, O.: Revision of partially ordered information: Axiomatization, semantics and iteration. In: Proc. of IJCAI 2005, Edinburgh, pp. 376–381 (2005)
4. Brewka, G.: Preferred sutheories: an extended logical framework for default reasoning. In: Proc. of IJCAI 1989, pp. 1043–1048 (1989)
5. Drap, P., Grussenmeyer, P.: A digital photogrammetric workstation on the web. Journal of Photogrammetry and Remote Sensing 55(1), 48–58 (2000)
6. Eiter, T., Ianni, G., Lukasiewicz, T., Schindlauer, R., Tompits, H.: Combining answer set programming with description logics for the semantic web. Artificial Intelligence 172(12-13), 1495–1539 (2008)
7. Gebser, M., Kaufmann, B., Neumann, A., Schaub, T.: *clasp*: A conflict-driven answer set solver. In: Baral, C., Brewka, G., Schlipf, J. (eds.) LPNMR 2007. LNCS (LNAI), vol. 4483, pp. 260–265. Springer, Heidelberg (2007)
8. Hue, J., Papini, O., Würbel, E.: Implementing prioritized merging with ASP. In: Hüllermeier, E., Kruse, R., Hoffmann, F. (eds.) IPMU 2010. CCIS, vol. 80, pp. 138–147. Springer, Heidelberg (2010)
9. Hué, J., Würbel, E., Papini, O.: Removed sets fusion: Performing off the shelf. In: Proc. of ECAI 2008, pp. 94–98 (2008)
10. Lefèvre, C., Nicolas, P.: A first order forward chaining approach for answer set computing. In: Erdem, E., Lin, F., Schaub, T. (eds.) LPNMR 2009. LNCS, vol. 5753, pp. 196–208. Springer, Heidelberg (2009)
11. Papini, O.: A complete revision function in propositionnal calculus. In: Neumann, B. (ed.) Proc. of ECAI 1992, pp. 339–343. John Wiley and Sons. Ltd., Chichester (1992)
12. Papini, O.: D3.1 archaeological activities and knowledge analysis. Technical report, Delivrable, VENUS project. january (2007), `http://www.venus-project.eu`
13. Papini, O., Curé, O., Drap, P., Fertil, B., Hué, J., Roussel, D., Sérayet, M., Seinturier, J., Würbel, E.: D3.6 reasoning with archaeological ontologies. technical report and prototype of software for the reversible fusion operations. Technical report, Delivrable, VENUS project (July 2009), `http://www.venusproject.eu`
14. Papini, O., Würbel, E., Jeansoulin, R., Curé, O., Drap, P., Sérayet, M., Hué, J., Seinturier, J., Long, L.: D3.4 representation of archaeological ontologies 1. Technical report, Delivrable, VENUS project (July 2008), `http://www.venus-project.eu`
15. Rescher, N., Manor, R.: On inference from inconsistent premises. Theory and Decision 1, 179–219 (1970)
16. Seinturier, J.: Fusion de connaissances: Applications aux relevés photogrammétriques de fouilles archéologiques sous-marines. PhD thesis, Université du Sud Toulon Var (2007)
17. Sérayet, M., Drap, P., Papini, O.: Extending removed sets revision to partially preordered belief bases. International Journal of Approximate Reasoning 52(1), 110–126 (2011)
18. Würbel, E., Papini, O., Jeansoulin, R.: Revision: an application in the framework of gis. In: Proc. of KR 2000, Breckenridge, Colorado, USA, pp. 505–516 (2000)
19. Yahi, S., Benferhat, S., Lagrue, S., Sérayet, M., Papini, O.: A lexicographic inference for partially preordered belief bases. In: Proc. of KR 2008, pp. 507–516 (2008)