Decentralized case-based reasoning and Semantic Web technologies applied to decision support in oncology

Mathieu d'Aquin, Jean Lieber, Amedeo Napoli

The Knowledge Engineering Review, Vol. 00:0, 1–27. © 2013, Cambridge University Press
DOI: 10.1017/S0000000000000000 Printed in the United Kingdom

Decentralized case-based reasoning and Semantic Web technologies applied to decision support in oncology

Mathieu d'Aquin[1], Jean Lieber[2], and Amedeo Napoli[2]

[1] *Knowledge Media Institute, The Open University, Walton Hall, Milton Keynes, MK7 6AA, UK*
[2] *LORIA (CNRS – Inria Nancy Grand-Est – Université de Lorraine), Campus scientifique - Équipe Orpailleur, BP 239, 54506 Vandœuvre-lès-Nancy France*
E-mail: m.daquin@open.ac.uk, Jean.Lieber@loria.fr, Amedeo.Napoli@loria.fr

Abstract

This article presents the KASIMIR system dedicated to decision knowledge management in oncology and which is built on top of Semantic Web technologies, taking benefit from standard knowledge representation formalisms and open reasoning tools. The representation of medical decision protocols, in particular for breast cancer treatment, is based on concepts and instances implemented within the description logic OWL DL. The knowledge units related to a protocol can then be applied for solving specific medical problems, using instance or concept classification. However, the straight application of a protocol is not always satisfactory, e.g., because of contraindications, necessitating an adaptation of the protocol. This is why the principles and methods of case-based reasoning in the framework of description logics have been used. In addition, the domain of oncology is complex and involves several specialties, e.g. surgery and chemotherapy. This complexity can be better undertaken with a viewpoint-based representation of protocols and viewpoint-based reasoning, for either application or adaptation of the protocols. Accordingly, a distributed description logic has been used for representing a viewpoint-based protocol. The application and the adaptation of the viewpoint-based protocol to medical cases is carried out using global instance classification and decentralized case-based reasoning.

1 Introduction

This article presents the KASIMIR system dedicated to decision knowledge management in oncology. The system is built on top of advanced Semantic Web technologies [Hitzler et al., 2009], where knowledge representation formalisms are based on description logics –OWL DL [Baader et al., 2003, Antoniou and van Harmelen, 2004] and C-OWL [Bouquet et al., 2004]– with associated reasoning formalisms such as subsumption and instantiation, and in addition case-based reasoning [Riesbeck and Schank, 1989, Aamodt and Plaza, 1994, Kolodner, 1993]. The KASIMIR system can be viewed as an intelligent assistant for physicians in their everyday practice of decision making [d'Aquin et al., 2008]. In Lorraine, i.e. the region of North-East France where the KASIMIR system is currently developed, the decisions on cancer treatment, surveillance, etc., are based on decision protocols. For example, the protocol for breast cancer treatment associates a specific treatment recommendation with the clinical status of a patient. The main example used throughout the paper is related to breast cancer treatment. The associated protocol is called hereafter either decision protocol or simply protocol.

A protocol contains the standard knowledge for decision support in oncology and is structured as a kind of decision tree. Roughly speaking, the protocol can be considered as a set of associations relating a class of patients suffering from cancer and a set of therapeutic decisions that can be applied to these patients. A patient suffering from cancer is called a "cancer patient", a "tumor patient" or simply a "patient" when there is no ambiguity. For most medical cases –about 60 to 70%– the protocol is simply and straightforwardly applied. The remaining non typical medical cases, called *out-of-protocol cases*, are examined by a committee of physicians called the *breast therapeutic decision committee*, for adapting the protocol and making a therapeutic decision. More precisely, given an out-of-protocol patient description, say P, the breast therapeutic decision committee generally selects a class of patients whose description is *close* to the patient description P and adapts the associated therapeutic decision to propose a treatment for the patient. This adaptation of the protocol can be modeled within a case-based reasoning (CBR) process, where the problem to be solved corresponds to a patient description and the solution to a therapeutic decision to be applied to the patient.

Oncology has to be regarded as a complex domain where several specialties, e.g. chemotherapy, surgery, and radiotherapy, are involved. For each specialty, different characteristics of the patient are analyzed and taken into account for setting on a specific treatment within the whole treatment process. Accordingly, during a meeting of the breast therapeutic decision committee, each expert provides a view specific to his/her specialty on the treatment, as a part of a collective recommendation. The oncology specialties determine interrelated viewpoints for a patient treatment, i.e. units of information on a given patient in a local viewpoint can be shared and combined with other units of information in another local viewpoint to build a global treatment. Moreover, a decision taken in a local viewpoint, i.e. for a particular oncology specialty, may have an influence on the decision to be taken in another local viewpoint. Hence, knowledge representation and reasoning within the KASIMIR system have to consider –and to take advantage of– the multiple viewpoints involved in the decision.

Until now, two versions of the KASIMIR system have been developed. The first version is based on an object-based representation formalism [d'Aquin et al., 2004], and the second version is based on a semantic portal using description logics (DLs) and knowledge formalisms dedicated to the Semantic Web [d'Aquin et al., 2005a]. Being a standard for knowledge representation and exchange for Semantic Web applications, the Web ontology language OWL[1], particularly the sub-language OWL DL, provide an adequate formalism for a reusable representation of knowledge contained in a decision protocol. Reasoning mechanisms associated with OWL DL, such as subsumption, classification, and instantiation, can be efficiently used for decision support in oncology. Finally, it should be noticed that there are still new developments in the KASIMIR system [Meilender et al., 2012].

Applying Semantic Web technologies in such a challenging domain leads to interesting issues, emphasizing some of the distinctive strengths and limitations of available formalisms such as OWL. Actually, an originality of the present research work is that the KASIMIR system implements various non-typical elements within a DL framework, namely:

- The representation of a decision protocol for associating a treatment recommendation with a patient description is based on a domain ontology implemented within the OWL DL description logic, and used for knowledge representation and reasoning.
- The representation of adaptation knowledge for protocol adaptation and CBR within a DL framework uses *similarity paths* for comparing two problems and *adaptation paths* for adapting the solution of a *source* problem to a *target* problem considered as being close to the source problem.
- The representation of viewpoints within the KASIMIR system –one viewpoint corresponding to one oncology specialty– is based on a distributed DL formalism, namely C-OWL. The C-OWL formalism provides means for local representation and reasoning within a viewpoint, and global reasoning across several viewpoints, leading to *decentralized CBR*.

Accordingly, the main contributions of the paper are the following:

- We define and use CBR in a distributed environment based on distributed description logics and implemented in C-OWL. This leads to the original notion of "decentralized CBR", which combines case-based reasoning and distributed description logics.

[1] http://www.w3.org/TR/owl-features/

- We make precise and give details on a real-world application in the medical domain, thanks to the KASIMIR system. This application illustrates and shows the capabilities of decentralized CBR.
  To the best of the authors' knowledge, decentralized CBR is an original approach, and one of the few implementing CBR within a distributed description logic framework in a working system.

These non typical items show that the KASIMIR system has a particular situation in the family of systems for decision support in medicine based on Semantic Web technology. Some systems are close to the KASIMIR system without necessarily having the same focus on knowledge representation and reasoning. In [Patkar et al., 2006] the focus is mainly on the implementation of oncology guidelines in practice with an evaluation of "computerized decision support" systems (a related distributed architecture is presented in [Besana et al., 2008]). An analysis of the decision systems is given and the possible improvements of physician decisions when they use a computerized decision support system. An evaluation carried out on a version of the KASIMIR system led to the same conclusions (see section 5). In the same way, the representation, evolution, and management of clinical practice guidelines are discussed in [Bouaud and Séroussi, 2008], following a series of work on the subject. The focus here relates to the management of the different temporal versions of the clinical guideline rather than decision support and problem solving. A Semantic Web based approach for synthesizing health knowledge through ontologies is presented in [Hussain and Abidi, 2009], which focuses on the integration of different knowledge sources. The most influential work for us probably remains [Bouquet et al., 2004], mainly for introducing contextualized ontologies and C-OWL, as this is explained throughout the paper. An important research work is also carried out in biomedicine and Semantic Web by the team which is currently developing the PROTÉGÉ editor and Bioportal[2], an open repository of biomedical ontologies [Noy et al., 2009, Ghazvinian et al., 2009]. Finally, let us mention a recent "Workshop on Semantic Interoperability in Medical Informatics (SIMI 2012)" including contributions about knowledge acquisition, ontology mapping of disperse biomedical resources, annotation and semantic search in the biomedical domain, and automated decision support services (`http://grid.ece.ntua.gr/sites/simi2012/SIMI2012_proc.pdf`).

The summary of the paper is as follows. First, the framework for protocol representation based on DLs is introduced. It is shown how a problem matching the protocol, i.e. the description of a patient satisfying the protocol, is solved (section 2). Then, CBR for protocol adaptation is motivated and described. The reformulation model underlies the CBR problem-solving process, involving both similarity and adaptation paths (section 3). The representation of viewpoints in oncology and distributed reasoning within viewpoints are then detailed (section 4). A discussion follows, focusing on the use of a DL framework for tasks such as: (i) managing real-world knowledge in oncology, (ii) designing the associated KASIMIR system, (iii) implementing viewpoints and a CBR process within the KASIMIR system (section 5). A paragraph explaining the implementation and the evaluation of the KASIMIR system concludes the paper (section 5).

2 The DL Framework for Decision Support guided by a Protocol

A *clinical decision support* system is defined in [van Bemmel and Musen, 1997] as a program for which the input is the description of a medical situation and the output corresponds to information units helping the practitioner in making decisions concerning this situation. In general, a clinical decision support system is a *knowledge-based system*, relying on a formalized medical knowledge base and on inference mechanisms. The KASIMIR system is compliant with the above definition: the knowledge base corresponds to the representation of a medical protocol in OWL DL, while standard DL reasoning mechanisms control decision support for protocol application.

2.1 A Brief Introduction to Description Logics

In this section, we recall the basics of description logics which are useful and necessary for understanding the DL expressions given in the paper. This short presentation is based on the DL Handbook [Baader et al., 2003] and on another more recent introduction to DLs [Krötzsch et al., 2012].

[2] `http://bioportal.bioontology.org`

Description logics are a family of knowledge representation languages that are widely used for designing knowledge-based systems. In DLs, there are three kinds of entities: concepts, roles and individual names. Concepts denote sets of individuals (classes in OWL), roles denote binary relations between the individuals (properties in OWL), and individual names denote simple individuals in the domain. Concepts correspond to unary predicates in first-order logic, roles to binary predicates and individuals to constants. Two levels of knowledge are taken into account: conceptual knowledge ("Tbox")including introductions of concepts and roles, and assertional knowledge ("Abox") including expressions involving individuals. DLs are equipped with a formal logic-based semantics, and DL systems provide their users with various inference capabilities mainly based on satisfiability, subsumption, and classification of concepts and of individuals.

Elementary descriptions are atomic concepts and atomic roles while complex descriptions can be built using concept contructors. More precisely, let **C** be a set of concept names and **R** be a set of role names disjoint from **C**. For the so-called $\mathcal{ALC}$ description logic, the set of concepts over **C** and **R** is inductively defined as follows:

- Every concept name is an $\mathcal{ALC}$ concept description.
- $\top$ (top) and $\bot$ (bottom) are $\mathcal{ALC}$ concept descriptions.
- If $\texttt{C}$ and $\texttt{D}$ are $\mathcal{ALC}$ concept descriptions and $\texttt{r}$ is a role name, then the following are also $\mathcal{ALC}$ concept descriptions:
  - $\texttt{C} \sqcap \texttt{D}$ (conjunction)
  - $\texttt{C} \sqcup \texttt{D}$ (disjunction)
  - $\neg\texttt{C}$ (complete negation)
  - $\exists\texttt{r.C}$ (existential restriction)
  - $\forall\texttt{r.C}$ (universal restriction)

Other concept constructions can be defined such as: cardinality restrictions with $(\geq\texttt{n r})$ ("at least" restriction) and $(\leq\texttt{n r})$ ("at most" restriction), and set of individuals (nominals). Role constructions can be used for extending $\mathcal{ALC}$, such as conjunctions of roles, inverse roles, role hierarchy, etc. Each $\mathcal{ALC}$ extension has a name: e.g. $\mathcal{ALCN}$ for cardinality restrictions, $\mathcal{ALCO}$ for nominals, $\mathcal{ALCR}$ for conjunctions of roles, $\mathcal{ALCI}$ for inverse roles, and finally $\mathcal{ALCH}$ for role hierarchy. Below, the $\mathcal{S}$ notation is used where $\mathcal{S}$ stands for $\mathcal{ALC}$ and thus $\mathcal{SHOIN}$ stands for $\mathcal{ALCHOIN}$. Moreover, in $\mathcal{SHOIN}(\mathcal{D})$, $\mathcal{D}$ indicates a concrete domain and this is explained in the next subsection.

An *interpretation* $\mathcal{I} = (\Delta^\mathcal{I}, .^\mathcal{I})$ is based on an interpretation domain $\Delta^\mathcal{I}$ and on an interpretation function $.^\mathcal{I}$ mapping a concept onto a subset of $\Delta^\mathcal{I}$, and a role onto a subset of the product $\Delta^\mathcal{I} \times \Delta^\mathcal{I}$, with respect to the following equations (where $card(X)$ returns the cardinality of set X).

$$
\begin{array}{l}
\top^\mathcal{I} = \Delta^\mathcal{I} \\
\bot^\mathcal{I} = \emptyset \\
(\texttt{C} \sqcap \texttt{D})^\mathcal{I} = \texttt{C}^\mathcal{I} \cap \texttt{D}^\mathcal{I} \\
(\texttt{C} \sqcup \texttt{D})^\mathcal{I} = \texttt{C}^\mathcal{I} \cup \texttt{D}^\mathcal{I} \\
(\neg\texttt{C})^\mathcal{I} = \Delta^\mathcal{I} \setminus \texttt{C}^\mathcal{I} \\
(\forall\texttt{r.C})^\mathcal{I} = \{x \in \Delta^\mathcal{I} \mid \forall y : (x, y) \in \texttt{r}^\mathcal{I} \rightarrow y \in \texttt{C}^\mathcal{I}\} \\
(\exists\texttt{r.C})^\mathcal{I} = \{x \in \Delta^\mathcal{I} \mid \exists y : (x, y) \in \texttt{r}^\mathcal{I} \wedge y \in \texttt{C}^\mathcal{I}\} \\
(\geq\texttt{m r})^\mathcal{I} = \{x \in \Delta^\mathcal{I} \mid card(\{y \in \Delta^\mathcal{I}/(x, y) \in \texttt{r}^\mathcal{I}\}) \geq m\} \\
(\leq\texttt{n r})^\mathcal{I} = \{x \in \Delta^\mathcal{I} \mid card(\{y \in \Delta^\mathcal{I}/(x, y) \in \texttt{r}^\mathcal{I}\}) \leq n\}
\end{array}
$$

Then, a concept $\texttt{C}$ is *satisfiable* iff there exists an interpretation $\mathcal{I} = (\Delta^\mathcal{I}, .^\mathcal{I})$ such as $\texttt{C}^\mathcal{I} \neq \emptyset$, otherwise $\texttt{C}$ is not satisfiable. A concept $\texttt{C}$ is *subsumed* by a concept $\texttt{D}$ iff for all interpretation $\mathcal{I}$, $\texttt{C}^\mathcal{I} \subseteq \texttt{D}^\mathcal{I}$. Satisfiability and subsumption are the base of DL-based reasoning modes. Examples are proposed below (see for example § 2.4).

Hereafter, the usual syntax and semantics of DLs, introduced in [Baader et al., 2003] as well as the naming conventions and recommendations are respected: i.e. a concept name such as `MultifocalTumor`

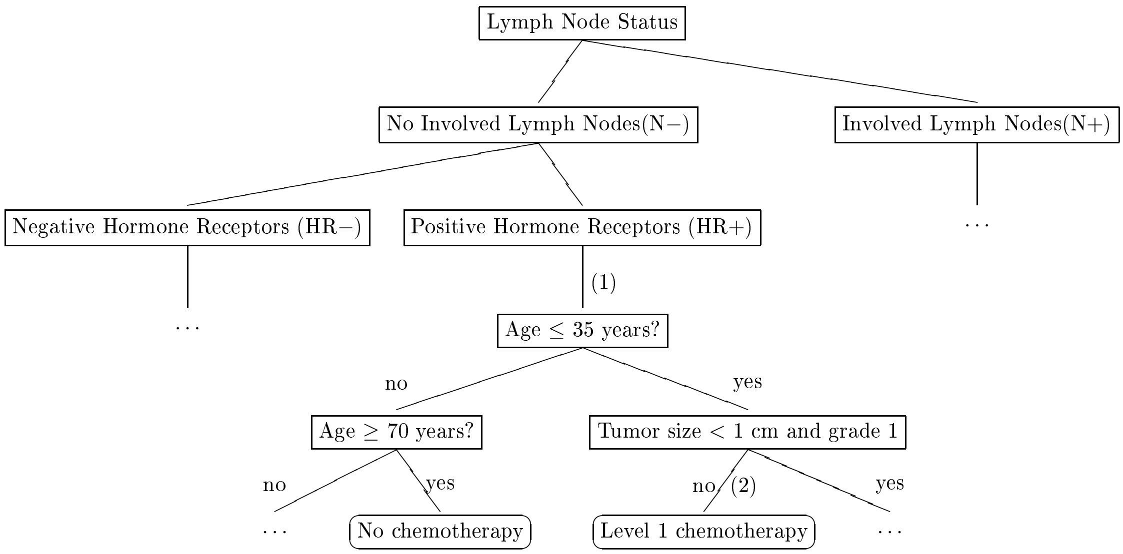

Figure 1 Excerpts from the breast cancer treatment protocol used at the CAV Hospital in Nancy, France (`www.oncolor.org`).

has a capitalized first letter, an instance name such as `PAT` has every letters capitalized, and a role name such as `hasRecommendation` has capitalized first letters except for the first one.

2.2 Description Logics and OWL DL

The knowledge associated with a protocol is represented within OWL DL 1.0, which can be seen as an expressive DL such as $\mathcal{SHOIN}(\mathcal{D})$, integrated within standard technologies of the Web such as XML and RDF(S) [Horrocks et al., 2003]. The notion of an *ontology* is considered here as an equivalent for a DL *knowledge base*.

In addition, the system described here makes use of concrete domain within the description logic formalism [Lutz et al., 2005, Lutz and Milicic, 2007]. Following the notations in [Lutz, 2003], a concrete domain is a pair $(\Delta^{\mathcal{D}}, \Phi^{\mathcal{D}})$ where $\Delta^{\mathcal{D}}$ is a set and $\Phi^{\mathcal{D}}$ is a set of predicate names on $\Delta^{\mathcal{D}}$. Here, the concrete domain $\mathcal{D} = (\mathbb{R}, \Phi^{\mathbb{R}})$ is considered, where $\mathbb{R}$ denotes real numbers, and $\Phi^{\mathbb{R}}$ is the set of predicate names for comparing real numbers: $\{<_a, \leq_a, =_a, \neq_a, \geq_a, >_a\}$, with $a \in \mathbb{R}$. The predicate P_a where $P \in \{<, \leq, =, \neq, \geq, >\}$ is interpreted as $P_a^{\mathcal{D}} = \{y \in \Delta^{\mathcal{D}} | y\ P\ a\}$. For example, ${\geq_3}^{\mathcal{D}} = [3; +\infty[$. The DL constructors are extended for integrating concrete domains. It becomes then possible to define a concept $\exists \mathtt{g}.\varphi$, where $\mathtt{g}$ is a *concrete role* and $\varphi \in \Phi^{\mathcal{D}}$. The concrete role $\mathtt{g}$ is interpreted as a partial function $\mathtt{g}^{\mathcal{I}} : \Delta^{\mathcal{I}} \mapsto \Delta^{\mathcal{D}}$ where $\Delta^{\mathcal{I}} \cap \Delta^{\mathcal{D}} = \emptyset$ and $(\exists \mathtt{g}.\varphi)^{\mathcal{I}} = \{\mathtt{x} \in \Delta^{\mathcal{I}} | \mathtt{g}^{\mathcal{I}}(\mathtt{x}) \in \varphi^{\mathcal{D}}\}$.

Accordingly, the concept of a “patient older than 70 having a tumor whose grade is different from 1” can be represented as:

$$\mathtt{TumorPatient} \sqcap \exists \mathtt{hasAge}. >_{70} \sqcap \exists \mathtt{hasTumor}.(\exists \mathtt{hasGrade}. \neq_1)$$

where `TumorPatient` is a concept representing patients suffering from a tumor.

2.3 Representation of a Protocol using DLs

A protocol can be likened to a decision tree, where internal nodes are related to the patient characteristics and leaf nodes are recommendations of treatments (see Figure 1). An internal node corresponds to a binary or Yes/No question related to one or several patient characteristics or features. Examples of features are the age, the tumor size (size), the tumor grade (grade 1, 2 or 3), the status of the lymph nodes (N+ for involved and N- for non involved lymph nodes), the status of the hormone receptors (HR+ for positive and

HR- for negative hormone receptors). The protocol in Figure 1 shows that, for a patient without involved lymph nodes (N-), with positive hormone receptors (HR+), younger than 35 years, and with a tumor size greater than 1 cm or having a grade $g \neq 1$, a chemotherapy of level 1 is recommended.

A path in the tree corresponds to a conjunction of conditions satisfied by a patient. The set of patients for which all these conditions apply constitutes a class of patients. Such a path is represented by a concept in DL and is defined by characteristics checked at the nodes of the path. For example, the concept C1 represents the set of patients having no involved lymph nodes and positive hormone receptors:

$$\begin{aligned}\texttt{C1} \equiv\ & \texttt{TumorPatient} \\ & \sqcap\ \exists\texttt{hasLymphNodes}.\{\texttt{n-}\} \\ & \sqcap\ \exists\texttt{hasHormoneReceptors}.\{\texttt{hr+}\}\end{aligned} \qquad \textbf{[ax1]}$$

The C1 concept reifies the path from the root of the tree to the node marked (1) in Figure 1. In particular, this concept includes the two nominals n- and hr+. Reasoning in presence of nominals is discussed in § 4.1 and detailed for example in [Lutz et al., 2005, Sirin et al., 2006, Kazakov et al., 2012]. An instance PAT of the TumorPatient concept satisfying every condition on the path from the root of the tree to the node (1) is recognized as an instance of C1. The path leading to the edge marked by (2) in Figure 1 is represented by the concept C2:

$$\begin{aligned}\texttt{C2} \equiv\ & \texttt{C1} \\ & \sqcap\ \exists\texttt{hasAge}.<_{35} \\ & \sqcap\ \neg(\exists\texttt{hasTumor}.(\exists\texttt{hasSize}.<_{1} \sqcap \exists\texttt{hasGrade}.=_{1}))\end{aligned} \qquad \textbf{[ax2]}$$

Following the physician way of doing, a cancer patient is considered as a person and as a medical case. A medical case is, in turn, considered as a problem case to be solved, or for which a solution has to be found, i.e. a cancer patient has a tumor for which a treatment has to be recommended. This leads to the notion of *medical case-treatment* axiom (see **[mct-ax]**) in the KASIMIR system (as made precise below, **[mct-ax]** is a kind of *problem-solution* axiom, see **[ps-ax]**). In the DL framework, the concepts to be considered are MedicalCase, TumorPatient, Tumor, and Treatment.

$$\begin{aligned}& \texttt{TumorPatient} \equiv \texttt{Person} \sqcap \texttt{MedicalCase} \sqcap \exists\texttt{hasTumor}.\texttt{Tumor} \\ & \texttt{MedicalCase} \sqsubseteq \exists\texttt{hasRecommendation}.\texttt{Treatment}\end{aligned} \qquad \textbf{[mct-ax]}$$

A recommendation corresponds to a treatment to be applied: a specific treatment is represented by a subconcept of Treatment. For example, the **[ax3]** axiom represents the link existing between the class of patients C2 and a recommendation for a treatment based on a chemotherapy of level 1.

$$\texttt{C2} \sqsubseteq \exists\texttt{hasRecommendation}.\texttt{Level1Chemotherapy} \qquad \textbf{[ax3]}$$

On a general level, the application of the protocol to a particular patient can be seen as a problem-solving task, where the problem corresponds to the description of the state of a cancer patient, i.e. a medical case, and the solution to the description of a treatment that can be recommended according to the characteristics of the patient. The following elements are involved in the problem-solving model:

- *Problems* are represented by the ProblemCase concept.
- *Solutions* are represented by the SolutionCase concept.
- *Axioms* relate specific problems to specific solutions.

In terms of DL representation, the concepts to be considered are:

$$\begin{aligned}& \texttt{ProblemCase} \sqsubseteq \exists\texttt{hasSolution}.\texttt{SolutionCase} \qquad \textbf{[ps-ax]} \\ & \texttt{MedicalCase} \sqsubseteq \texttt{ProblemCase} \\ & \texttt{hasRecommendation} \sqsubseteq \texttt{hasSolution}\end{aligned}$$

The **[ps-ax]** axiom is called a *problem-solution axiom* and can be used to guide DL reasoning for solving domain-specific problems, e.g. decision support in oncology. In addition, the `MedicalCase` concept is a subconcept of `ProblemCase` and the `hasRecommendation` role is a subrole of the `hasSolution` role. The introduction of problem-solution axioms is a generic technique for problem-solving, and can be reused in any other specific domain, e.g. diagnosis, recognition, classification. In addition, problem-solution axioms provide a basis for integrating DLs and CBR as explained in Section 3.

2.4 DL Reasoning for Decision Support

Given two concepts C and D in an ontology O, the *subsumption test* checks whether C is subsumed by D (or C is more specific than D): formally, C is subsumed by D, denoted by $\mathtt{C} \sqsubseteq \mathtt{D}$ if, for every model $\mathcal{I}$ of O, $\mathtt{C}^{\mathcal{I}} \subseteq \mathtt{D}^{\mathcal{I}}$. Based on the subsumption test, the *classification* process consists in searching, for a given concept C, the concepts subsuming C and the concepts subsumed by C. The classification process allows to organize the concepts of an ontology into a concept hierarchy. In the KASIMIR system, the concept hierarchy is used as a structure for representing a protocol.

The *instance checking* test checks whether an individual a is an instance of a concept C in an ontology O, i.e. if for every model $\mathcal{I}$ of O, $\mathtt{a}^{\mathcal{I}} \in \mathtt{C}^{\mathcal{I}}$. This test is a basis for the *instantiation* reasoning service that, given an individual a, consists in searching for the concepts C for which a is an instance of C. Decision support for protocol application is based on instantiation as follows. Relying on the problem-solution axiom **[ps-ax]**, an individual recognized as an instance of the `ProblemCase` concept is linked to an instance of the `SolutionCase` concept through the `hasSolution` role.

Accordingly, an instance representing a cancer patient is first recognized as an instance of the concept `MedicalCase` and is then associated with an instance of the `Treatment` concept through the role `hasRecommendation`. For example, let us consider that the instance PAT is a cancer patient verifying the following assertions:

$\mathtt{TumorPatient}(\mathtt{PAT})$	
$\mathtt{Tumor}(\mathtt{T})$	
$\mathtt{hasTumor}(\mathtt{PAT}, \mathtt{T})$	
$\mathtt{hasLymphNodes}(\mathtt{PAT}, \mathtt{n-})$	*no involved lymph node*
$\mathtt{hasHormoneReceptors}(\mathtt{PAT}, \mathtt{hr+})$	*positive hormone receptors*
$\mathtt{hasAge}(\mathtt{PAT}, 27)$	*age* ≤ 35
$\mathtt{hasGrade}(\mathtt{T}, 1)$	*tumor of grade 1*
$\mathtt{hasSize}(\mathtt{T}, 1.2)$	*tumor size* ≥ 1

The instantiation reasoning mechanism recognizes PAT as being an instance of the concept C2 (see axiom **[ax2]**). The axiom **[ax3]** states that a chemotherapy of level 1 is recommended for patients whose tumor satisfies the characteristics of the C2 concept. Formally, it can be inferred from the instantiation of C2 by PAT that there exists $\mathtt{x} \in \mathtt{Level1Chemotherapy}^{\mathcal{I}}$ such that $(\mathtt{PAT}, \mathtt{x}) \in \mathtt{hasRecommendation}^{\mathcal{I}}$, i.e. the instance PAT is associated with the object x in $\mathtt{Level1Chemotherapy}^{\mathcal{I}}$ through the `hasRecommendation` role.

The *concept satisfiability* test checks whether a concept C can be instantiated, i.e. if there exists a model $\mathcal{I}$ of the ontology for which $\mathtt{C}^{\mathcal{I}} \neq \emptyset$. In addition, an ontology O is said to be *unsatisfiable* or *inconsistent* if O does not have any model, e.g. O includes an instance of an unsatisfiable concept. In the KASIMIR system, the concept satisfiability test is used during the validation of the protocol representation, e.g. for checking whether a decision rule can be triggered. Moreover, it is assumed that the description of a cancer patient is correctly defined with respect to the protocol, i.e. the protocol and the description of the patient constitute a consistent knowledge base. The DL reasoner is then used during the design of the protocol and the elaboration of the patient description for detecting inconsistencies to be (manually) repaired.

3 Protocol Adaptation: from DL Reasoning to Case-Based Reasoning

The previous section explains how traditional DL inferences are carried out in the KASIMIR system for decision support associated with the application of medical protocols. These protocols take into account

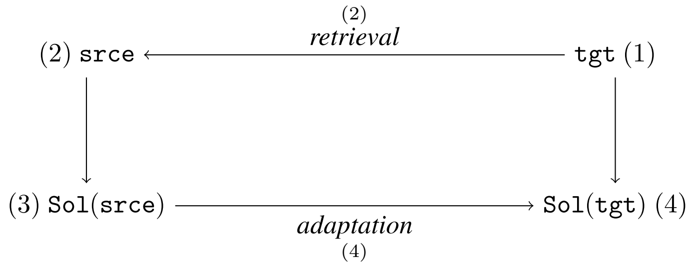

Figure 2 The case-based inference. (1) A target problem is provided to the CBR system. (2) The system retrieves a source problem similar to the target problem. (3) The source problem is a problem associated with a solution. (4) The system adapts this solution in order to solve the target problem.

the most common medical situations, i.e. about 60–70% of the medical cases. For the non standard medical cases, or *out-of-protocol patients*, the protocol cannot be directly applied. A decision has to be taken by a committee of physicians (having multiple specialties). During a meeting of this committee, a recommendation is established by adapting the solutions provided by the protocol for similar cases. For example, the breast cancer treatment protocol is designed for a female patient and cannot be directly applied to a male patient suffering from a breast cancer. A treatment for such a male patient can be established by reusing the recommendation provided by the protocol for a "similar female patient", i.e. considering the male patient as a female patient and adapting this recommendation. Actually, a precise and formal definition of an out-of-protocol patient does not exist. When the protocol cannot be applied in a direct and satisfactory way, the fact that a patient is declared as being out-of-protocol is decided by physicians. In the context of the KASIMIR system, the protocol cannot be directly applied to an out-of-protocol patient because some characteristics of the patient are missing or are unknown. Then, the protocol has to be adapted.

The retrieval of a similar patient and the adaptation of the protocol can be taken into account by case-based reasoning. This is the purpose of the next sections to introduce the principles of case-based reasoning and to show how CBR is integrated within a DL formalism for protocol adaptation and decision support.

3.1 Principles of Case-Based Reasoning

Case-based reasoning (CBR) is used in many applications to solve problems in so-called "weak theory domains". CBR solves new problems by retrieving and adapting solutions of previously solved cases stored in a case base [Riesbeck and Schank, 1989, Kolodner, 1993, Aamodt and Plaza, 1994].

A typical CBR cycle is composed of four main steps: given a new problem, called a *target problem*, the *retrieval* step selects a *source case* from a case base; the *adaptation* step builds a solution to the target problem, relying on the differences between the source and target problems, and modifying accordingly the solution of the retrieved source case; the *revision* step improves the current target solution depending on actual and expected system results; finally the *retention* (*memorization*) step stores the newly solved case for further reuse. As shown on Figure 2, among these steps, retrieval and adaptation are closely related and have a central place.

In CBR, the basic knowledge unit is the case which represents a problem solving episode. More precisely, a case $(\mathtt{pb}, \mathtt{Sol}(\mathtt{pb}))$ is a pair composed of a *problem* $\mathtt{pb}$ and a *solution* $\mathtt{Sol}(\mathtt{pb})$ of $\mathtt{pb}$. A case base is a structured set of *source cases* denoted by $(\mathtt{srce}, \mathtt{Sol}(\mathtt{srce}))$. The CBR process aims at solving a *target problem*, denoted by $\mathtt{tgt}$, with respect to a case base. During the *retrieval* step, a source problem $\mathtt{srce}$, considered to be similar to the target problem $\mathtt{tgt}$, is searched for in the case base. During the *adaptation* step, the solution $\mathtt{Sol}(\mathtt{srce})$ of $\mathtt{srce}$ is adapted for building a solution $\mathtt{Sol}(\mathtt{tgt})$ of $\mathtt{tgt}$.

Then, the solution $\mathtt{Sol(tgt)}$ may be tested and repaired. A learning step may complete the process where the new case $(\mathtt{tgt}, \mathtt{Sol(tgt)})$ can be memorized for future reuse.

In knowledge intensive CBR (KI-CBR) [Aamodt, 1990, 1991, Gómez-Albarrán et al., 1999, Aamodt, 2004], the CBR process relies on a formalized model of domain knowledge, such as a domain ontology, for case organization and retrieval [Lieber and Napoli, 1996]. In the next section, it is shown how a KI-CBR system takes advantage of domain knowledge for adapting a decision protocol, on the basis of the *reformulation model*.

3.2 Principles of CBR in the KASIMIR *system*

The model of reformulations introduced in [Melis et al., 1998] is a general framework for modeling and representing adaptation knowledge into simple and separated components. A reformulation is a pair $(\mathtt{r}, \mathcal{A}_\mathtt{r})$ where $\mathtt{r}$ is a relation between problems and $\mathcal{A}_\mathtt{r}$ is an *adaptation function*: when $\mathtt{r}$ relates $\mathtt{srce}$ to $\mathtt{tgt}$, then any solution $\mathtt{Sol(srce)}$ of $\mathtt{srce}$ can be adapted into a solution $\mathtt{Sol(tgt)}$ of $\mathtt{tgt}$, thanks to the adaptation function $\mathcal{A}_\mathtt{r}$ relating $\mathtt{Sol(srce)}$ to $\mathtt{Sol(tgt)}$ w.r.t. $\mathtt{srce}$ and $\mathtt{tgt}$. Actually, $\mathcal{A}_\mathtt{r}$ is a function depending on three inputs, namely $\mathtt{srce}$, $\mathtt{Sol(srce)}$, and $\mathtt{tgt}$ (this is made precise below).

The operations corresponding to problem relations $\mathtt{r}$ and adaptation functions $\mathcal{A}_\mathtt{r}$ have to be designed for a particular application, and, as well, a relevant strategy should be provided for building reformulations. Most of the time, and this is also true for the KASIMIR system, reformulations are based on operations such as specialization, generalization, and substitution.

In the reformulation model, retrieval consists in finding a *similarity path* relating $\mathtt{srce}$ to $\mathtt{tgt}$. Accordingly, a sequence of adaptation functions following the similarity path is reified into a corresponding *adaptation path* (see Figure 3).

$$
\begin{array}{ccccccccc}
\mathtt{srce} & \xrightarrow{\mathtt{r}_1} & \mathtt{pb}_1 & \xrightarrow{\mathtt{r}_2} & \mathtt{pb}_2 & \text{———} \cdots \text{———} & \mathtt{pb}_{q-1} & \xrightarrow{\mathtt{r}_q} & \mathtt{tgt} \\
\downarrow & & \downarrow & & \downarrow & & \downarrow & & \downarrow \\
\mathtt{Sol(srce)} & \xrightarrow{\mathcal{A}_{\mathtt{r}_1}} & \mathtt{Sol(pb_1)} & \xrightarrow{\mathcal{A}_{\mathtt{r}_2}} & \mathtt{Sol(pb_2)} & \text{———} \cdots \text{———} & \mathtt{Sol(pb_{q-1})} & \xrightarrow{\mathcal{A}_{\mathtt{r}_q}} & \mathtt{Sol(tgt)}
\end{array}
$$

Figure 3 On the first line, a similarity path from $\mathtt{srce}$ to $\mathtt{tgt}$. On the second line, the corresponding adaptation path from $\mathtt{Sol(srce)}$ to $\mathtt{Sol(tgt)}$.

Similarity paths and adaptation paths

A similarity path SP from a problem $\mathtt{srce}$ to a problem $\mathtt{tgt}$ is a set SP of triples $(\mathtt{pb_{i-1}}, \mathtt{r_i}, \mathtt{pb_i})$, $\mathtt{i} \in \{1, 2, \ldots, \mathtt{q}\}$, where $\mathtt{pb_0} = \mathtt{srce}$ and $\mathtt{pb_q} = \mathtt{tgt}$, $\mathtt{r_i}$ relates the $\mathtt{pb_{i-1}}$ and $\mathtt{pb_i}$ problems and is such that there exists a reformulation $(\mathtt{r_i}, \mathcal{A}_{\mathtt{r_i}})$ in the available adaptation knowledge (constituted by a finite set of reformulations). Given $\mathtt{Sol(srce)}$, a solution of $\mathtt{srce}$, and a similarity path from $\mathtt{srce}$ to $\mathtt{tgt}$, the solution of $\mathtt{tgt}$ can be built thanks to a sequential application of the adaptation functions $\mathcal{A}_{\mathtt{r_1}}$, $\mathcal{A}_{\mathtt{r_2}}$, ..., and $\mathcal{A}_{\mathtt{r_q}}$:

$$
\begin{array}{l}
\mathtt{Sol(pb_1)} = \mathcal{A}_{\mathtt{r_1}}(\mathtt{srce}, \mathtt{Sol(srce)}, \mathtt{pb_1}) = \mathcal{A}_{\mathtt{r_1}}(\mathtt{pb_0}, \mathtt{Sol(pb_0)}, \mathtt{pb_1}) \\
\mathtt{Sol(pb_2)} = \mathcal{A}_{\mathtt{r_2}}(\mathtt{pb_1}, \mathtt{Sol(pb_1)}, \mathtt{pb_2}) \\
\ldots \\
\mathtt{Sol(tgt)} = \mathtt{Sol(pb_q)} = \mathcal{A}_{\mathtt{r_q}}(\mathtt{pb_{q-1}}, \mathtt{Sol(pb_{q-1})}, \mathtt{pb_q})
\end{array}
$$

A similarity path is built according to a collection of available reformulations. Then several similarity paths relating two problems $\mathtt{srce}$ and $\mathtt{tgt}$ may exist. The choice between two similarity paths can be made either by the system according to an evaluation function or to an "external" domain expert (as this is the case here). In the current implementation (see section 5 and [d'Aquin et al., 2006c] for details), the set of available applicable operations, e.g. specialization, generalization, substitution, is recursively used to enumerate the possible (non cyclic) similarity paths. All possible solutions based on the different existing similarity paths are provided to the expert in charge of the decision.

This search for similarity paths respects the *adaptation-guided retrieval principle* [Smyth, 1996], where only source cases for which a solution is adaptable are retrieved, meaning and implying that adaptation knowledge is available. According to this principle, a similarity path provides a "symbolic reification" of similarity between problems, allowing a case-based reasoner to build an understandable explanation of the results.

An example of application of reformulations to breast cancer treatment

Figure 4 presents a simplified example of the reformulation model for adapting the breast cancer treatment protocol. In this example, the `tgt` problem corresponds to the description of a patient for which the protocol cannot be directly applied. There are two reasons for that: (i) the patient is a man, and the protocol for breast cancer treatment takes into account only female patients, (ii) the tumor localization of this patient cannot be determined with precision in the breast, linked to the fact that the patient is a man. The other characteristics used for making a decision are related to the age of the patient, the size of the tumor, and the fact that the tumor has only one focal point.

Given the target problem `tgt` and the source case $(\mathtt{srce}, \mathtt{Sol}(\mathtt{srce}))$, two reformulations, namely $(\mathtt{r}_1, \mathcal{A}_{\mathtt{r}_1})$ and $(\mathtt{r}_2, \mathcal{A}_{\mathtt{r}_2})$, are applied for solving `tgt`. Considering the similarity path, i.e. reading the figure from right to left, from `tgt` to `srce`, relation $\mathtt{r}_2$ denotes a substitution of `Male Patient` with `Female Patient`, while relation $\mathtt{r}_1$ denotes a substitution of `Unknown Localization` with `Internal Localization`. Considering the adaptation path, the $(\mathtt{r}_1, \mathcal{A}_{\mathtt{r}_1})$ reformulation indicates that, when the localization of the tumor is unknown, the patient has to be considered as if she/he had an internal tumor, because it is the worst case, and thus, the one for which the *largest* treatment should be envisioned. The adaptation function $\mathcal{A}_{\mathtt{r}_1}$ simply corresponds to a copy of the solution $\mathtt{Sol}(\mathtt{srce})$. The $(\mathtt{r}_2, \mathcal{A}_{\mathtt{r}_2})$ reformulation expresses that, since an ovary ablation is not applicable for a man, it has to be replaced with an hormone therapy having similar effects, i.e. cures of an anti-oestrogen drug called tamoxifen.

For implementing a CBR system following this framework in a particular application domain, adaptation knowledge is needed, taking the form of a set of reformulations $(\mathtt{r}, \mathcal{A}_{\mathtt{r}})$. This knowledge acquisition issue is addressed for the KASIMIR system in two complementary ways: (1) the acquisition of adaptation knowledge from experts is examined in [d'Aquin et al., 2006b], i.e. a kind of "manual" knowledge acquisition process, whereas (2) a knowledge discovery approach is proposed in [d'Aquin et al., 2007], i.e. a kind of semi-automatic learning process based on frequent itemset search. The work presented here focuses on the reasoning mechanism making use of such adaptation knowledge and therefore assumes that it is available.

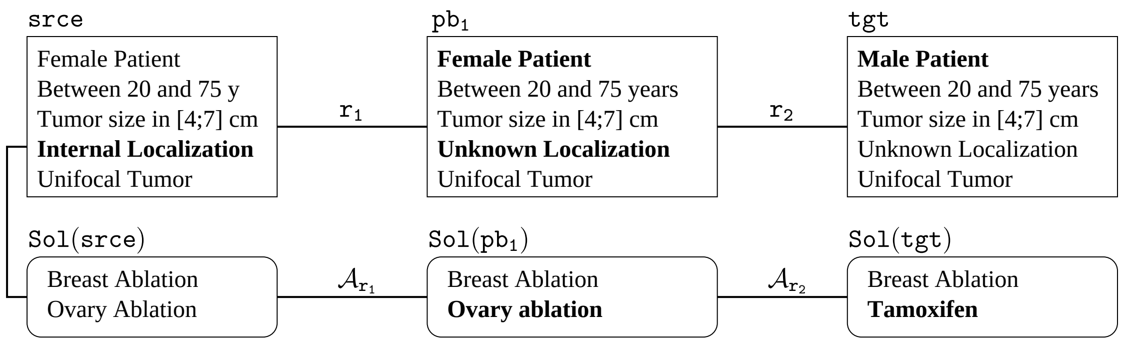

Figure 4 An example of CBR problem-solving using reformulations, with similarity path from `srce` to `tgt` on first line and corresponding adaptation path from $\mathtt{Sol}(\mathtt{srce})$ to $\mathtt{Sol}(\mathtt{tgt})$ on the second line.

3.3 DL Reasoning for supporting CBR Inferences

The inference mechanisms available within DL formalisms, namely subsumption and instantiation, can be used for implementing a CBR system [Salotti and Ventos, 1998, Gómez-Albarrán et al., 1999]. The subsumption relation organizes the concepts of an ontology into a hierarchy, that can be used in CBR as an indexing structure for the case base, for facilitating the retrieval of relevant source cases [Lieber and Napoli, 1996, Napoli et al., 1996]. In this hierarchy, an index `idx(srce)` is an abstraction of `srce`, encoding the relevant elements in the `srce` description leading to the `Sol(srce)` solution. The index `idx(srce)` is represented by the concept `Idx_Srce` and is linked to a solution represented by the `Idx_Sol_Srce` concept. A problem-solution axiom relates the index concept `Idx_Srce` to an index solution concept `Idx_Sol_Srce` as follows:

$$\mathtt{Idx\_Srce} \sqsubseteq \exists \mathtt{hasRecommendation.Idx\_Sol\_Srce}$$

Moreover, the source and target problems, respectively denoted as `srce` and `tgt` in the CBR notation, are represented by instances in DLs, respectively denoted by `SRCE` and `TGT`, while their solutions are represented by the instances `SOL_SRCE` and `SOL_TGT`. The individual `SRCE` is an instance of the index concept `Idx_Srce`, while the solution `SOL_SRCE` can be reused for any instance of the index concept `Idx_Srce`, i.e. as soon as a problem `TGT` is recognized as an instance of `Idx_Srce`, the solution `SOL_SRCE` of `SRCE` can be reused to solve `TGT`.

The retrieval operation of CBR relies on instantiation for finding the index of a source problem having an adaptable solution. The similarity path built during retrieval relates the problem `TGT` to a problem $\mathtt{PB}_0$ (i.e. the source problem), such that $\mathtt{PB}_0$ is an instance of `Idx_Srce`:

$$\mathtt{SRCE} \hookrightarrow \mathtt{Idx\_Srce} \hookleftarrow \mathtt{PB}_0\ \mathtt{r}_1\ \mathtt{PB}_1\ \mathtt{r}_2 \ldots \mathtt{PB}_\mathtt{q}\ \mathtt{r}_\mathtt{q}\ \mathtt{TGT}$$

The $\hookrightarrow$ arrow stands for “is an instance of”. The source problem represented by an instance of the index concept `Idx_Srce` is attached to a source solution through a problem-solution axiom similar to the axioms **[mct-ax]** and **[ps-ax]** introduced above. Thus, as soon as $\mathtt{PB}_0$ is recognized as an instance of `Idx_Srce`, the solution $\mathtt{SOL\_PB}_0$ is put in correspondence with the solution `SOL_SRCE` of `SRCE`. Therefore, the associated adaptation path is:

$$\mathtt{SOL\_SRCE} = \mathtt{SOL\_PB}_0\ \mathcal{A}_{\mathtt{r}_1}\ \mathtt{SOL\_PB}_1\ \mathcal{A}_{\mathtt{r}_2} \ldots \mathtt{SOL\_PB}_\mathtt{q}\ \mathcal{A}_{\mathtt{r}_\mathtt{q}}\ \mathtt{SOL\_TGT}$$

As an application, let us consider the example introduced in Figure 4. The target problem (on the right) can be represented as:

`TumorPatient(TGT)`	
`Male(S)`	
`hasGender(TGT, S)`	*male gender*
`hasAge(TGT, 55)`	*age between 20 and 75*
`hasTumor(TGT, T)`	
`¬MultifocalTumor(T)`	*unifocal tumor*
`UnknownLoc(L)`	
`hasLocalization(T, L)`	*unknown tumor localization*
`hasSize(T, 5.2)`	*tumor size between 4 and 7*

The index concept of the source problem in Figure 4 is represented by the concept `Idx_Srce` (**[ax4]**) with the problem-solution axiom introducing the concept `Idx_Sol_Srce`, i.e. the recommended treatment (**[ax45]**):

$$\begin{aligned}\mathtt{Idx\_Srce} \equiv\ & \mathtt{TumorPatient}\\ & \sqcap \exists \mathtt{hasGender.Female}\\ & \sqcap \exists \mathtt{hasAge.} \geq_{20}\end{aligned}$$

$$\begin{aligned}&\sqcap \exists \mathtt{hasAge}.\leq_{75}\\&\sqcap \exists \mathtt{hasTumor}.(\neg\mathtt{MultifocalTumor}\\&\qquad \sqcap \exists \mathtt{hasLocalization}.\mathtt{Internal}\\&\qquad \sqcap \exists \mathtt{hasSize}.\geq_{4}\\&\qquad \sqcap \exists \mathtt{hasSize}.\leq_{7})\end{aligned} \quad \textbf{[ax4]}$$

$$\mathtt{Idx\_Srce} \sqsubseteq \exists \mathtt{hasRecommendation}.\mathtt{BreastAndOvaryAblation} \quad \textbf{[ax5]}$$

The BreastAndOvaryAblation concept corresponds to Idx_Sol_Srce and represents treatments composed of a breast ablation and an ovary ablation. The similarity path built by the CBR retrieval operation corresponds to the similarity path shown in Figure 4:

$$\mathtt{Idx\_Srce} \hookleftarrow \mathtt{PB}_0\ \mathtt{r}_1\ \mathtt{PB}_1\ \mathtt{r}_2\ \mathtt{TGT}$$

TGT is an instance of TumorPatient with the properties corresponding to the characteristics of the patient to be treated. $\mathtt{PB}_0$ and $\mathtt{PB}_1$ are also instances of TumorPatient. According to $\mathtt{r}_2$, $\mathtt{PB}_1$ is introduced with the same set of assertions as TGT, except for the gender:

$$\begin{aligned}&\mathtt{Female(S1)}\\&\mathtt{hasGender}(\mathtt{PB}_1, \mathtt{S1}) \quad \textit{female gender}\end{aligned}$$

According to $\mathtt{r}_1$, $\mathtt{PB}_0$ (the source problem) is introduced with the same set of assertions as $\mathtt{PB}_1$, except for the localization of the tumor:

$$\begin{aligned}&\neg\mathtt{MultifocalTumor}(\mathtt{T}_0)\\&\mathtt{hasTumor}(\mathtt{PB}_0, \mathtt{T}_0)\\&\mathtt{Internal}(\mathtt{L}_0)\\&\mathtt{hasLocalization}(\mathtt{T}_0, \mathtt{L}_0)\end{aligned}$$

Then, $\mathtt{PB}_0$ is recognized as an instance of the index concept Idx_Srce (see axiom **[ax4]**) and is linked to the solution $\mathtt{SOL\_PB}_0$ of the Idx_Sol_Srce treatment concept, i.e. a surgery consisting of a breast ablation and an ovary ablation (see axiom **[ax5]**). The adaptation of the solution is based on the adaptation path:

$$\mathtt{BreastAndOvaryAblation} \hookleftarrow \mathtt{SOL\_PB}_0\ \mathcal{A}_{\mathtt{r}_1}\ \mathtt{SOL\_PB}_1\ \mathcal{A}_{\mathtt{r}_2}\ \mathtt{SOL\_TGT}$$

The instance $\mathtt{SOL\_PB}_1$ is a copy of $\mathtt{SOL\_PB}_0$ (application of $\mathcal{A}_{\mathtt{r}_1}$), and SOL_TGT is obtained by replacing the ovary ablation in $\mathtt{SOL\_PB}_1$ by cures of tamoxifen (application of $\mathcal{A}_{\mathtt{r}_2}$):

$$\mathtt{SOL\_TGT} \hookrightarrow \mathtt{BreastAblationAndTamoxifen}$$

4 Using Distributed Description Logics for Reasoning with Multiple Viewpoints

Oncology, like many other medical domains, involves several medical specialties. Thus, decision knowledge embedded in the protocols combines aspects from different specialties within several viewpoints. Each of these viewpoints constitutes a suitable way within a specialty for representing, organizing, and using decision knowledge. Indeed, a cancer specialist will use different elements of knowledge, and different representations of the patient according to his/her interest, e.g. for establishing a surgical treatment or a chemotherapy.

In this section, we introduce an original way of representing viewpoints in a distributed environment, namely distributed description logics. In addition, we show how CBR can be used for problem-solving in a particular application domain.

4.1 The Modularization of Oncology Protocols

Multiple viewpoints are visible in the structure of protocols. For example, the protocol for breast cancer treatment is based on several "sub-protocols", each of them focusing on a particular phase of the medical treatment, i.e. preoperative chemotherapy, surgery, radiotherapy, and complementary treatment, involving one or more specialties of oncology, among chemotherapy, surgery, radiotherapy, and hormonotherapy. A sub-protocol relies on its own vocabulary and provides a particular recommendation, in a given phase of the treatment, using only specific characteristics for this task. In addition, the breast therapeutic decision committee whose role is to adapt the content of the protocol for out-of-protocol cases, includes experts from different oncology specialties. An expert brings his/her own contribution to the solution, according to the specialty he/she depends on, and collaborates with the experts from other specialties, in the construction of a global and satisfactory solution.

Distributed DLs (DDLs) are extensions of DLs in which several local ontologies are considered to be related through semantic mappings [Borgida and Serafini, 2002, 2003]. Hereafter, it is shown how knowledge representation and reasoning based on the DDL framework is used within the KASIMIR system for taking into account the multiple viewpoints of the oncology specialties. Actually, the representation of the protocol presented in section 2.3 is *modularized*, meaning that it is reified as a multi-viewpoint representation relying on the principles of DDLs. Nevertheless, the characteristics of the representation introduced in section 2.3 remain valid and are used locally in each viewpoint.

Note on the DDL implementation of viewpoints

In the following, the emphasis is put on the design principles for multi-viewpoint representation of medical protocols, actually on what is expected to be achieved with DDLs. Consequently, some of the examples introduced in this section, as well as their representation in DDL, relies on characteristics going beyond the possibilities offered by the current implementation of the DDL formalism (e.g., the DRAGO system [Serafini and Tamilin, 2005, Tamilin, 2007]). In addition to showing what is expected to be achieved with DDLs and the DRAGO system, these examples are intended to provide real-world use-cases to be used for guiding further developments in DDLs.

In particular, the DDL formalism did not allow the use of nominals in bridge rules and this is still the case in the current version of the DRAGO system. Luciano Serafini, one of the initiators of the DDL formalism, has proposed a formalism and a tableau algorithm for dealing with mappings and axioms involving nominals in DDLs [Serafini, 2007, Homola and Serafini, 2010, Serafini and Homola, 2012]. Thus, the use of nominals in axioms hereafter is justified. However, in practice, the multi-viewpoint representation of the breast cancer treatment protocol implemented within the KASIMIR system is a simplification of what is presented in the following. In particular, nominals are replaced with atomic concepts having only one instance, namely the nominal that has to be considered. When the system has to deal with multiple nominals in expressions such as sets of nominals, an atomic concept is introduced for every nominal, with the constraint that these atomic concepts are mutually disjoint. Replacing nominals with atomic concepts is a simplification leading to an inelegant way of implementing the system: it is not easy to get a generic construction and many atomic concepts have to be introduced. Fortunately, the proposition in [Serafini, 2007, Homola and Serafini, 2010] allows the implementers to adapt the coding of the system and to propose more generic and efficient constructions for DDLs.

4.2 Distributed Description Logics and C-OWL

DDLs have been introduced in [Borgida and Serafini, 2002] and are the basis of the C-OWL extension of OWL for representing *contextualized ontologies* [Bouquet et al., 2004]. Contextualized ontologies are local representations of a domain, called *contexts*, that are semantically related through mappings. In the present framework, DDLs are used for formalizing and implementing several complementary representations of the oncology domain called *viewpoints*.

4.2.1 Syntax and Semantics

In a DDL, the knowledge about a domain is distributed over a set of contexts. A context O_i is an ontology, with a proper language and a proper interpretation. A mapping between contexts O_i and O_j is expressed through a set of bridge rules from O_i to O_j allowing to declare a correspondence between the interpretation domains of the two contexts. In this way, knowledge units in O_i can be interpreted and reused in (terms of) O_j.

Formally, a *context space* includes a set of contexts $\{O_i\}_{i \in I}$, where I is a set of indexes for contexts. The indexes in I are used to prefix the expressions, in associating an expression with the context in which it is defined. For example, $\mathtt{i:C}$, $\mathtt{i:}\exists\mathtt{R.C}$, $\mathtt{i:a}$, $\mathtt{i:C} \sqsubseteq \mathtt{D}$, and $\mathtt{i:C(a)}$, are expressions of the local language in O_i.

The semantics of a context space is given by a distributed interpretation $\mathfrak{I}$, including an interpretation $\mathcal{I}_i$ for each $i \in I$. $\mathcal{I}_i$ is defined by a local interpretation domain $\Delta^{\mathcal{I}_i}$ and a local interpretation function $\cdot^{\mathcal{I}_i}$. A context is interpreted with the corresponding local interpretation, i.e. an axiom or an assertion of O_i is satisfied by $\mathfrak{I}$ if it is satisfied by $\mathcal{I}_i$.

A mapping $\mathcal{M}_{ij}$ is a set of *bridge rules* from O_i to O_j. There are different types of bridge rules between concepts, instances or roles, of two contexts. An *into rule* is a bridge rule of the form $\mathtt{i:C} \xrightarrow{\sqsubseteq} \mathtt{j:D}$, where $\mathtt{i:C}$ and $\mathtt{j:D}$ are concepts respectively from O_i and O_j. This rule means that the concept $\mathtt{i:C}$ of O_i is considered, from the viewpoint of O_j, as being more specific than the concept $\mathtt{j:D}$. The *onto rule* $\mathtt{i:C} \xrightarrow{\sqsupseteq} \mathtt{j:D}$ means that the concept $\mathtt{i:C}$ is considered in O_j to be more general than $\mathtt{j:D}$. Bridge rules are directional: a bridge rule from O_i to O_j is considered from the viewpoint of O_j, and thus, $\mathtt{i:C} \xrightarrow{\sqsubseteq} \mathtt{j:D}$ is not equivalent to $\mathtt{j:D} \xrightarrow{\sqsupseteq} \mathtt{i:C}$.

Formally, the distributed interpretation $\mathfrak{I}$ of a context space is associated with a set of *domain relations*. A domain relation $r_{ij} \subseteq \Delta^{\mathcal{I}_i} \times \Delta^{\mathcal{I}_j}$ states, for an arbitrary object of $\Delta^{\mathcal{I}_i}$, the object of $\Delta^{\mathcal{I}_j}$ it corresponds to. The notation $r_{ij}(\mathtt{C}^{\mathcal{I}_i})$ denotes the interpretation of the concept $\mathtt{i:C}$ of O_i considered in the interpretation domain of O_j. The semantics of a bridge rule is given with respect to domain relations: $\mathfrak{I}$ satisfies $\mathtt{i:C} \xrightarrow{\sqsubseteq} \mathtt{j:D}$ if $r_{ij}(\mathtt{C}^{\mathcal{I}_i}) \subseteq \mathtt{D}^{\mathcal{I}_j}$ and $\mathfrak{I}$ satisfies $\mathtt{i:C} \xrightarrow{\sqsupseteq} \mathtt{j:D}$ if $r_{ij}(\mathtt{C}^{\mathcal{I}_i}) \supseteq \mathtt{D}^{\mathcal{I}_j}$.

Another form of bridge rules can be used to specify a correspondence between instances. $\mathtt{i:a} \xrightarrow{\equiv} \mathtt{j:b}$ means that the instance $\mathtt{i:a}$ in O_i corresponds to the instance $\mathtt{j:b}$ in O_j. Formally, $\mathfrak{I}$ satisfies $\mathtt{i:a} \xrightarrow{\equiv} \mathtt{j:b}$ if $r_{ij}(\mathtt{a}^{\mathcal{I}_i}) = \mathtt{b}^{\mathcal{I}_j}$.

4.2.2 Global and Local Reasoning

Local reasoning services in DDL are the standard DL reasoning services, performed in a particular context, without taking into account the bridge rules. A *global reasoning service* takes advantage of bridge rules for inferring statements in a context in using knowledge from other contexts. In [Serafini and Tamilin, 2004] and [Tamilin, 2007], an extension of the standard tableau algorithm for the computation of the global subsumption test in DDLs is presented. *Global subsumption* relies on the principle of "subsumption propagation" that, in its simplest form, can be expressed as:

if the mapping $\mathcal{M}_{ij}$ contains "$\mathtt{i:E} \xrightarrow{\sqsupseteq} \mathtt{j:C}$" and "$\mathtt{i:F} \xrightarrow{\sqsubseteq} \mathtt{j:D}$"
then "$\mathfrak{I}$ satisfies $\mathtt{i:E} \sqsubseteq \mathtt{F}$" implies that "$\mathfrak{I}$ satisfies $\mathtt{j:C} \sqsubseteq \mathtt{D}$"

Intuitively, this means that subsumption in a particular context can be inferred from subsumption in another context thanks to bridge rules. Similarly, *global instance checking* is based on an instantiation propagation rule:

if $\mathcal{M}_{ij}$ includes "$\mathtt{i:C} \xrightarrow{\sqsubseteq} \mathtt{j:D}$" and "$\mathtt{i:a} \xrightarrow{\equiv} \mathtt{j:b}$"
then "$\mathfrak{I}$ satisfies $\mathtt{i:C(a)}$" implies that "$\mathfrak{I}$ satisfies $\mathtt{j:D(b)}$"

Instantiation is extended for global instance checking. Based on bridge rules, information known about an instance in a particular context can be completed using inferences made in other contexts. It can be

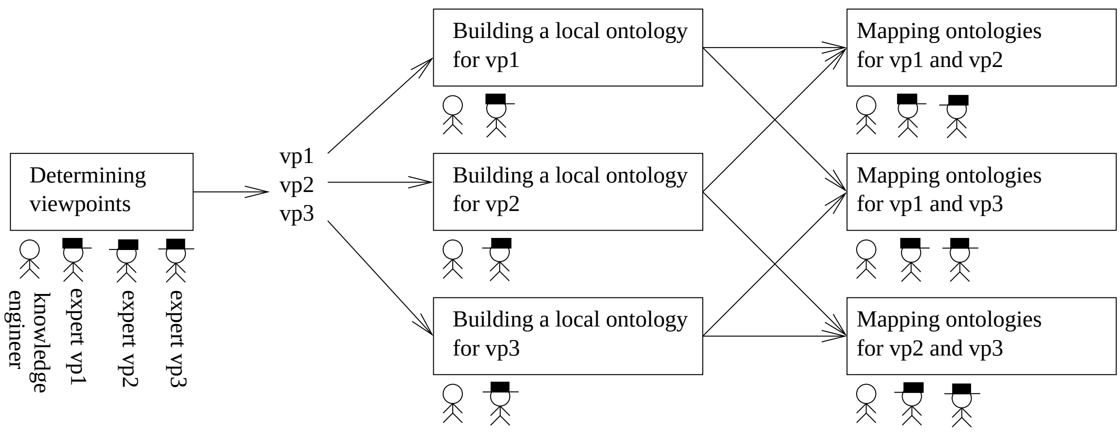

Figure 5 Three main steps can be considered for representing knowledge according to multiple viewpoints: determining viewpoints, building a local ontology, and mapping ontologies.

noticed that bridge rules relating instances, as well as global instance checking, were introduced in the DRAGO system [Tamilin, 2007], being in part motivated by the needs of the KASIMIR system.

4.3 Representation of Oncology Viewpoints using DDLs

Modeling and formalizing a multiple viewpoint representation within a DDL is a *decentralized* task, i.e. there is no need to set up a consensus, but there is a need to distinguish viewpoints. Three main steps can be considered (see Figure 5): (1) *determine the relevant viewpoints* in the domain, (2) *build a local ontology* for every viewpoint, and (3) *establish mappings* between local ontologies, reifying correspondences between viewpoints.

Determining viewpoints

A large and complex domain is often organized according to various sub-domains, e.g. services, tasks, working groups, communities. Such an organization provides a prior division of the domain into viewpoints, that can be reused by determining the viewpoints to be represented in an application. In the same way, the protocol for breast cancer treatment is considered to be composed of four main treatment phases, involving the following oncology specialties: presurgical chemotherapy (pc), surgery (s), radiotherapy (r), and complementary treatment (ct: hormonotherapy and/or chemotherapy). Although they exist for themselves, these four treatment phases are interrelated, i.e. in some situations, the decision taken for a particular phase depends on decision units lying in another one. This division of oncology knowledge is daily used by physicians and is well-suited for their work. Accordingly, in the KASIMIR system, this division supports four viewpoints, namely O_{pc}, O_s, O_r, and O_{ct}, respectively corresponding to the specialties pc, s, r, and ct. A fifth viewpoint could be considered as well: viewing the patient as a person, with, e.g. administrative data, habits, geographical data. However, this supplementary viewpoint was not introduced for keeping the example as simple as possible.

Building a local ontology

Building a local ontology consists in formalizing a context in DDL corresponding to a viewpoint. A local ontology includes the knowledge units considered to be useful within the current viewpoint, *independently* from the other viewpoints, i.e. a context in DDL formalizing a local ontology only implements knowledge units relevant to the viewpoint.

Decisions in the different specialties depend on different characteristics of the patient. Therefore, different concepts, representing a patient with different roles, are used for taking into account the protocol according to a given viewpoint. For example, in surgery, a patient with a directly operable and non multifocal tumor has to be treated with a partial ablation of the breast. This is represented by the following axioms in O_s:

$$\texttt{s:C} \equiv \texttt{TumorPatient} \sqcap \exists\texttt{hasTumor}.(\texttt{OperableTumor} \sqcap \exists\texttt{hasFocus.SimpleFocus}) \quad \textbf{[axS1]}$$

$$\texttt{s:C} \sqsubseteq \exists\texttt{hasRecommendation.PartialAblation} \quad \textbf{[axS2]}$$

In radiotherapy, the considered characteristics are related to the recommended surgery, to the localization of the tumor in the breast, the size, and the presence of malignant cells in the lymph nodes:

$$\texttt{r:C} \equiv \texttt{TumorPatient} \sqcap \exists\texttt{hasRecommendedSurgery.PreservingAblation} \sqcap \exists\texttt{hasLymphNodes}.\{\texttt{n-}\} \sqcap \exists\texttt{hasTumor}.(\exists\texttt{hasLocalization.InferoInternal}) \quad \textbf{[axR1]}$$

$$\texttt{r:RadiotherapyBreastAndEIMC} \equiv \texttt{Irradiation} \sqcap \exists\texttt{hasZone.Breast} \sqcap \exists\texttt{hasZone.EIMC} \quad \textbf{[axR2]}$$

$$\texttt{r:C} \sqsubseteq \exists\texttt{hasRecommendation.RadiotherapyBreastAndEIMC} \quad \textbf{[axR3]}$$

These axioms state that a patient with a preserving surgery recommendation, non involved lymph nodes, and a tumor located in the infero-internal part of the breast, has to be treated by a radiotherapy of the breast and a radiotherapy of the extended internal mammary chain (EIMC).

In the context O_{ct} the choice of the complementary treatment is guided by characteristics such as the status of the hormone receptors, the number of involved lymph nodes, the age of the patient, the size of the tumor, and the grade (1, 2, or 3). For example, the protocol states that a patient whose age is less than 35, with no involved lymph nodes, positive hormone receptors, a tumor of grade equal to or higher than 2, or of size equal to or higher than 1 cm, has to be treated by a chemotherapy of level 1:

$$\texttt{ct:C} \equiv \texttt{TumorPatient} \sqcap \exists\texttt{hasAge}.\leq_{35} \sqcap \exists\texttt{numberInvolvedLymphNodes}.\{\texttt{0}\} \sqcap \exists\texttt{hasHormoneReceptors}.\{\texttt{hr+}\} \sqcap \neg\exists\texttt{hasTumor}.(\exists\texttt{hasGrade}.=_1 \sqcap \exists\texttt{hasSize}.<_1) \quad \textbf{[axCT1]}$$

$$\texttt{ct:C} \sqsubseteq \exists\texttt{hasRecommendation.Level1Chemotherapy} \quad \textbf{[axCT2]}$$

Mapping ontologies

The relations existing between viewpoints are materialized through mappings between local ontologies. An important question concerns the knowledge units that can be *shared* and *reused* from a given viewpoint in another viewpoint. Thus, it is important to make precise the correspondences between viewpoints, indicating how a given context may take advantage of knowledge units in another context.

In the current example, a local ontology represents the vision of one or two specialties of oncology, playing a role in a particular phase of the treatment. Surgery is the central phase of the treatment and takes place after presurgical chemotherapy: then, decisions in surgery follow decisions in presurgical chemotherapy accordingly. In the same way, radiotherapy and complementary treatment aim at complementing surgery (for eliminating remaining involved cells with radiotherapy, chemotherapy, or hormonotherapy). The decision in these two viewpoints depends on the knowledge units in the surgery viewpoint. Figure 6 summarizes the existing mappings between the viewpoints introduced above. An arrow from a context O_i to a context O_j indicates the existence of a non empty mapping $\mathcal{M}_{ij}$.

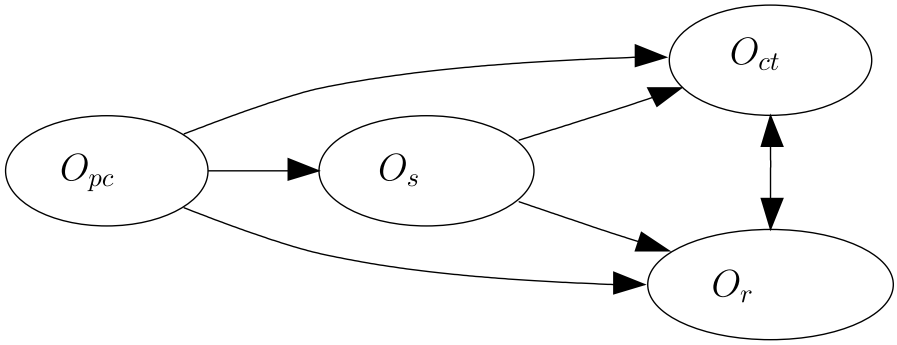

Figure 6 The mappings existing between viewpoints (contexts) in the multiple viewpoint representation of the protocol for breast cancer treatment.

In this framework, bridge rules are used either for sharing knowledge on patients between the viewpoints or for representing the influence of a decision in a particular viewpoint on the decision in another viewpoint. For example, decisions in radiotherapy and in complementary treatment both rely on the status of the lymph nodes –either involved or not–, even if this information unit is represented differently in O_r and O_{ct}. The sharing of this characteristic of the patient between O_r and O_{ct} can be represented by the bridge rules **[br1]** and **[br2]**.

$$\texttt{r:}\exists\texttt{hasLymphNodes.}\{\texttt{n-}\} \xrightarrow{\sqsubseteq} \texttt{ct:}\exists\texttt{numberInvolvedLymphNodes.}\{0\} \quad \textbf{[br1]}$$

$$\texttt{r:}\exists\texttt{hasLymphNodes.}\{\texttt{n+}\} \xrightarrow{\sqsubseteq} \texttt{ct:}\exists\texttt{numberInvolvedLymphNodes.}\geq_1 \quad \textbf{[br2]}$$

In sequence, a decision in radiotherapy depends on the recommended treatment in surgery, either preserving, i.e. a partial ablation preserving the essential part of the breast, or not preserving. The influence of the treatment in surgery on the treatment in radiotherapy is reified in the following bridge rule, allowing to complete a decision for the patient in O_r with the recommendation decided in O_s (bridge rule **[br3]**):

$$\texttt{s:}\exists\texttt{hasRecommendation.PartialAblation} \xrightarrow{\sqsubseteq} \texttt{r:}\exists\texttt{hasRecommendedSurgery.PreservingAblation} \quad \textbf{[br3]}$$

4.4 DDL Reasoning for Decision Support in Oncology

In the KASIMIR system, decision support based on multiple viewpoints is performed by global reasoning services provided by DDLs, mainly global instantiation (see section 4.2.2). This reasoning service is extended to a decentralized CBR mechanism for case adaptation.

4.4.1 Global Instantiation for Protocol Application

Solving a problem in a multi-viewpoint framework with CBR is a *decentralized* process: inferences occur locally, in each viewpoint represented by a context in DDL, taking advantage of bridge rules for reusing the knowledge and inferences from the other viewpoints. This problem-solving approach is similar to a reasoning process in *decentralized artificial intelligence*, defined in [Demazeau and Müller, 1989] as concerning the activity of autonomous intelligent agents that coexist and collaborate with other agents, each agent having proper goals and proper knowledge.

O_s	O_r	O_{ct}
TumorPatient(P) DirectlyOperableTumor(T) hasTumor(P, T) SimpleFocus(F) hasFocus(T, F)	TumorPatient(P) hasLymphNodes(P, n-) Tumor(T) hasTumor(P, T) InferoInternal(L) hasLocalization(T, L)	TumorPatient(P) Tumor(T) hasTumor(P, T) hasGrade(T, 2) hasSize(T, 1.8) hasHormoneReceptors(P, hr+) hasAge(P, 33)

Figure 7 The assertions describing the patient P to be treated in the three considered contexts.

Hereafter, a specific example shows how global instantiation is used for applying the protocol in a decentralized reasoning approach. In the example, the considered patient is 33 years old, is suffering from breast cancer, with a tumor having the following characteristics: non multifocal tumor of grade 2, with a size of 1.8 cm, directly operable, localized in the infero-internal part of the breast, with positive hormone receptors, and no involved lymph node. The physician having to establish a treatment for this patient may consider the above characteristics according to one or several viewpoints, among the four available viewpoints, i.e. O_{pc}, O_s, O_r, and O_{ct}. However, according to domain knowledge, the viewpoint "presurgical chemotherapy" cannot be considered in the present case. Thus, the physician is interested in the recommendations of the protocol for the three other viewpoints. The patient is represented by three local instances of the TumorPatient concept, namely: s:P (surgery), r:P (radiotherapy), and ct:P (complementary treatment) (see Figure 7).

Equivalence bridge rules make precise the fact that the different instances introduced in the viewpoints denote a same patient:

$$\texttt{s:P} \xrightarrow{\equiv} \texttt{r:P}$$
$$\texttt{r:P} \xrightarrow{\equiv} \texttt{ct:P}$$
$$\texttt{ct:P} \xrightarrow{\equiv} \texttt{s:P}$$

Based on the assertions introduced in O_s and on the axioms associated with the protocol in O_s, namely **[axS1]** and **[axS2]**, it can be inferred that a partial ablation is recommended for the patient in surgery (knowing that a directly operable tumor is an operable tumor). Accordingly, the bridge rule **[br3]** is used to complete the information known on the patient in radiotherapy, r:P is recognized as an instance of r:∃hasRecommendedSurgery. PreservingAblation.

Relying on the axioms **[axR1]**, **[axR2]**, and **[axR3]**, in O_r, a radiotherapy of the breast and of the extended internal mammary chain is recommended (RadiotherapyBreastAndEIMC). Now, the bridge rule **[br1]** allows to reuse information in O_r concerning involved lymph nodes in O_{ct}. Then, ct:P is considered as an instance of ct:∃numberInvolvedLymphNodes.{0}. The axioms **[axCT1]** and **[axCT2]** in O_{ct} allow to infer that a chemotherapy of level 1 is recommended for the patient ct:P. Finally, the global solution recommended by the protocol is composed of three instances of Treatment, i.e. three complementary viewpoints on the solution:

```
s:PartialAblation
r:RadiotherapyBreastAndEIMC
ct:Level1Chemotherapy
```

This decentralized problem-solving approach provides a *decentralized solution* composed of elements of solution for every viewpoint that has been considered during the problem-solving process. These three treatments do not represent alternatives, but different recommendations made by the system for the different phases of treatment, built from local knowledge in each viewpoint and taking into account the influence of viewpoints on one another.

4.4.2 Decentralized Case-Based Reasoning for Protocol Application

Decentralized CBR (DzCBR) is the CBR mechanism defined in a framework based on multiple viewpoint representation, such as a multi-viewpoint protocol. DzCBR relies on DDL bridge rules. The target problem is divided into a set of local target problems $\{\texttt{i:TGT}\}_i$, one local target problem $\texttt{i:TGT}$ for a particular viewpoint (context O_i). Equivalence bridge rules such as $\texttt{i:TGT} \xrightarrow{\equiv} \texttt{j:TGT}$ are declared between local target problems $\texttt{i:TGT}$ and $\texttt{j:TGT}$ ($i \neq j$), for making precise the fact that a local target problem is a view of the global target problem in a given viewpoint. The DzCBR mechanism relies on two main operations, materializing respectively the *localized* and the *collaborative* components of the reasoning process:

(i) *Local CBR* works on local knowledge for building a local solution $\texttt{i:SOL\_TGT}$ of the local target problem $\texttt{i:TGT}$. Local CBR relies on the CBR problem-solving approach introduced in section 3. Since local CBR is restricted to one particular viewpoint, it can be carried out in each context independently from the others.

(ii) *Case completion* is the *collaborative* step of DzCBR. It is based on bridge rules and global instantiation for completing the local target case, either the problem $\texttt{i:TGT}$ or the solution $\texttt{i:SOL\_TGT}$, thanks to knowledge sharing with the other viewpoints.

Local CBR and case completion are performed in each context O_i, until no more inferences can be drawn. The solution finally delivered corresponds to a set of local solutions, $\{\texttt{i:SOL\_TGT}\}_i$, provided by the local CBR processes and by the collaboration between viewpoints through bridge rules (as this is described in the preceding section on global instantiation for protocol application). Therefore, DzCBR is applied in the KASIMIR system for protocol adaptation relying on a multiple viewpoint representation of the protocol [d'Aquin et al., 2005b].

5 Discussion

Representation and reasoning with medical knowledge

There is a long history of research in medical informatics concerning formalization, capitalization, and automated use of medical knowledge, from the first *expert systems* [Clancey and Shortliffe, 1984, Marcus, 1988] to Semantic Web portals such as the KASIMIR semantic portal. In relation with the topic of this paper, several languages have been developed for the representation of *medical guidelines and protocols*, allowing a partial automated use of the protocol content. As shown in the survey [Peleg et al., 2003], these languages are generally focused on *task-centric* protocols, in the sense that the medical knowledge is represented in terms of sequences of actions. By contrast, the protocols represented within the KASIMIR system are *patient-centered* and based on knowledge representation languages such as DLs and DDLs. The representation of sequences of decisions and actions was not taken into account in the present paper, but could be considered in the future, for e.g. reifying the sequence in which treatments have to be applied. Actually, there is an on-going research work on the analysis and representation of patient pathways in the team of the authors [Jay et al., 2006, 2008, Egho et al., 2011]. The sequence of treatments are not so easy to integrate in a DL-based representation system, and requires the use of adapted temporal extensions, such as the one introduced in [Schulz and Hahn, 2004].

Using description logics for a real-world application in the domain of oncology

In this paper, a number of features of the KASIMIR system are detailed. Although these features are implemented for specific and actual purpose needs, namely decision support in oncology, they are of general interest and may be reused in other domains where similar problems are raised. Firstly, a protocol can be seen as a decision tree and can be represented as such within a description logic framework. In addition, the problem-solving process relies on the association of a medical case description (problem case) with a treatment description (solution case) through DL axioms. The DL used for implementing the KASIMIR system is the representation language OWL DL, a semantically well-founded knowledge representation language. OWL DL allows interoperability with Web applications, such as the set of tools linked to a semantic portal for, e.g. editing, retrieving, and visualizing knowledge units.

Meanwhile, the KASIMIR system is aimed at working in a complex real-world domain, involving elements that are difficult to represent and to manipulate. A first remark concerns the non-easiness of manipulating concrete domains (datatypes) as integrated elements within DLs (see e.g. [Pan and Horrocks, 2005]). In the KASIMIR system, real numbers, intervals of numbers, and procedures for managing these datatypes, are needed. The actual implementation of the system shows that this is neither easy nor convenient to work with these datatypes, without being obliged to build ad hoc constructions for satisfying reasoning and computational services, e.g. interval comparison (inclusion, equality), computation of attribute values such as the age given the date of birth, the body mass index depending on the height and the mass, etc.

Embedding case-based reasoning within a description logic framework

Another element of importance is the introduction of a CBR process within the KASIMIR system. Here, CBR may be regarded as an extension of DL reasoning based on satisfiability, subsumption, and instantiation. The CBR process allows a flexible problem-solving process, not based on exact matching but on partial matching. For that purpose, similarity paths and adaptation paths have been defined for reusing already solved problems. The idea of *similarity path* may be imported in other contexts, where the similarity between two problems or two elements to be compared can be defined and reified in terms of a set of transformations in the problem space. The transformations have a counterpart in the solution space through an adaptation function that corresponds to each transformation function. Then, the sequence of applied adaptation functions yields to an *adaptation path*. The framework of similarity and adaptation paths is of general interest, and provides an efficient operational problem-solving approach, that can be embedded without major changes in the framework of a DL. Moreover, the reasoning procedures of the DL system, namely concept classification, concept satisfiability, and instance recognition, support the CBR process. In this way, classification-based reasoning and CBR are fully integrated, and provide a realistic reasoning framework for a real-world working system. For example, in the KASIMIR system, they allow the full exploitation of the protocols and the guidance for the adaptation of source cases for solving medical out-of-protocol cases.

On the design of a distributed or viewpoint-based knowledge base

One of the first elements to be considered for evaluating the strength of the viewpoint representation is the easiness and convenience of expressiveness provided by a distributed representation language, such as DDLs and C-OWL. First of all, it can be noticed that, formally, the DDL formalism is not more expressive than a DL formalism, as shown in [Borgida and Serafini, 2002, 2003]: any knowledge base represented within a DDL can be transformed without loss of information into a knowledge base represented within a standard DL. A set of contexts in a context space implemented in C-OWL can be organized in a standard ontology implemented in OWL. However, the principle of *non propagation of the inconsistency*, introduced in C-OWL [Bouquet et al., 2004, Serafini and Tamilin, 2005], is not taken into account in the transformation described in [Borgida and Serafini, 2002]. This principle has not yet been directly used in the KASIMIR system.

A main interest of the representation of viewpoints in DDL can be understood in terms of knowledge engineering and knowledge management. The decentralized approach based on viewpoints or local ontologies represented by contexts, consists in building a set of local and modular ontologies. These ontologies are presenting an interest for a given group of persons and correspond to a part of the domain or to a given problem-solving task. Then, the ontologies are mapped to each other by the mean of bridge rules. A prior global consensus on domain knowledge is no more necessary. A viewpoint materializes an oriented and homogeneous representation of the knowledge considered to be useful according to a given interest. Moreover, the maintenance of a set of contexts is simpler than the maintenance of a unique ontology aggregating the whole domain knowledge. The knowledge units represented in a particular viewpoint may be locally updated, without making any reference to other representation units in other contexts (in certain cases, bridge rules have to be updated as well). Furthermore, the actors of the domain focus on the relevant and/or well-suited context, with a simplified and direct access to the adequate knowledge units. The collaboration and knowledge exchange between viewpoints are then carried out and controlled automatically, through mappings. Accordingly, in the KASIMIR system, the viewpoint-based representation of the protocol allows a simpler and focused knowledge acquisition and representation, knowing that a viewpoint is composed of only a few hundreds of concepts, built and evolving independently from other viewpoint's concepts.

The implementation of the KASIMIR system

The KASIMIR system is embedded within a *semantic portal*, committed to knowledge management and decision support in oncology [d'Aquin et al., 2005a]. The architecture of the KASIMIR semantic portal relies on a knowledge server, implemented as a set of Web services and embedding a DL reasoner. The PELLET OWL reasoner [Sirin et al., 2007] and the JENA API [McBride, 2002] are currently used to manage the representation of the protocols in OWL and DL reasoning. Practically, several protocols are modeled and formalized (see next paragraph). For example, the representation of the breast cancer treatment protocol, according to the principles introduced in section 2, includes about one thousand concepts and one hundred roles. The implementation was realized using the PROTÉGÉ ontology editor, extended with features for visualization and maintenance specific to the KASIMIR system. Furthermore, the DDL reasoner DRAGO, which is currently developed within PELLET at the Trento University [Serafini and Tamilin, 2005], was tested on a multi-viewpoint representation of the breast cancer treatment protocol (actually a simplification of the representation presented in section 4.3).

Finally, as introduced in section 3, a CBR prototype system was designed and implemented [d'Aquin et al., 2006c]. It is based on two Web services for case retrieval and adaptation, and relies on an OWL representation of the knowledge units used by the CBR process, e.g. cases, reformulations.

An evaluation of the KASIMIR system

The use of the KASIMIR system by physicians has been evaluated for the object-based representation version of the system (mentioned in the introduction and implemented within an object-based representation or OBR formalism). The OBR formalism includes the concept constructors $\top$, $\mathtt{C} \sqcap \mathtt{D}$, $\exists\mathtt{r.C}$, and $\exists\mathtt{g}.\varphi$ (see § 2.2, in this way, it can be compared to the description logic $\mathcal{EL}(\mathtt{D})$). The OBR formalism was sufficient for representing the protocols and for standard reasoning (as explained in § 2.4). However, the use of the OBR formalism requires the introduction of much more atomic concepts than this is necessary for the OWL DL version. From this point of view, the use of the OWL DL formalism can be considered as a substantial improvement. Nevertheless, the results of the evaluation of the OBR version of KASIMIR can be understood in the same way in the context of the OWL DL version of the KASIMIR system (`http://katexowl.loria.fr`).

An evaluation of the OBR version of KASIMIR is presented in [Rios et al., 2003] and was carried out by physicians on breast and prostate cancer protocols. The protocols were presented under three formats: (i) the original documents describing the protocols called the "paper protocols", (ii) an HTML version of these documents as available on the `www.oncolor.org` website called the "HTML protocols", (iii) the protocols as implemented in the OBR version of the KASIMIR system called the "KASIMIR protocols". The use of these three formats of the protocol was measured by the so-called *compliance*. For a given physician and a given format, the compliance on a test set of medical cases is the proportion of answers of the physician that are in agreement with the protocol. According to the principles of evidence-based medicine [Evidence-based medicine working-group, 1992], the compliance has to be maximized whenever the medical cases are not out-of-protocol (and this was the case for the test sets).

The results of this evaluation show that the mean compliance with the KASIMIR format was better for breast and prostate cancer protocols than the two other formats (paper and HTML), though this "superiority" was statistically significant only for the breast cancer protocol, certainly because not enough data were collected for the prostate cancer protocol. More precisely, the compliance values for the breast cancer protocol were 77.8, 75.8, and 87.1%, respectively for the paper, HTML, and KASIMIR formats. This shows that the development of the KASIMIR system has to be continued.

The evaluation of the DL version of the KASIMIR system has to be considered at a qualitative level. Actually, this DL version is viewed as a "proof of concept" of the following statements:

- Medical decision protocols can be represented with description logics such as $\mathcal{SHOIN}(\mathcal{D})$ and distributed description logics, with a concrete domain for numerical intervals.
- A decision problem, say `pb`, consists of the description of a medical case, e.g. the state of a patient, while the associated solution `Sol(pb)` consists of a medical decision, e.g. a treatment. The inference relating `pb` to `Sol(pb)` relies on instance classification.
- Inferences are efficiently carried out with an inference engine such as PELLET.

Several decision protocols were implemented and their use, e.g. by querying, is made possible through PELLET. These computerized medical decision protocols are related to the treatment of:

- prostate cancer,
- melanoma cancer,
- rectum cancer,
- larynx cancer,
- non metastatic breast cancer.

In addition, other existing protocols are related to:

- Breast cancer surveillance: where the problem consists of the description of a patient who has been cured of her/his breast cancer, and the solution consists of the forthcoming exams to prevent the patient from having a relapse.
- Neutropenia with fever: where the problem consists of the description of a patient having fever and who is currently under chemotherapy treatment –which involves less defense of the organism–, and the solution consists of actions to be carried out for taking in charge the patient.
- Eligibility of patients for some clinical trials where the problem consists of a description of a patient ill with cancer, and a solution consists of clinical trials she/he is eligible for.

All these "computerized protocols" belong to the Oncolor network (`www.oncolor.org`) and are not directly available (though they can be consulted on proper site). Moreover, their validity is restricted in time as they compile the state of the art about a medical decision protocol at a given time and they have to be updated when new results about this decision problem are published.

6 Conclusion and future work

The KASIMIR system is aimed at knowledge management and decision support in oncology. In this paper, topics such as representation, reasoning, and viewpoints, have been detailed. The KASIMIR system implements medical decision protocols, in particular for breast cancer treatment, using a DL formalism, namely the OWL DL language. In addition, a way of associating the characteristics of patients with recommended treatments is shown. The direct application of standard patient-treatment associations in the protocol is not always satisfactory, leading to the need of adapting the protocol. A CBR process embedded within a DL framework is used for solving the adaptation problem.

Moreover, a viewpoint-based representation of the protocols has been set up, using a DDL, for taking into account the complexity and diversity of oncology. The representation is based on four main viewpoints: surgery, chemotherapy, radiotherapy, and complementary treatment. A decentralized reasoning mechanism, relying on classification and decentralized CBR, takes advantage of viewpoints and of bridge rules between viewpoints for solving adaptation problems.

The CBR mechanism presented in section 3 is considered as a *knowledge intensive* process, relying on a DL-based representation of domain knowledge and on *adaptation knowledge*. Accordingly, the reformulation model is implemented within a DL framework for representing and making operational adaptation knowledge. In addition, efforts are currently carried out for designing methods, techniques, and tools for the acquisition of adaptation knowledge in the KASIMIR system. In particular, a part of this research work is related to the reuse of the results of adaptation knowledge acquisition sessions gathering oncology experts, researchers in computer science and in psycho-ergonomy [Lieber et al., 2003]. Another part involves knowledge discovery in databases techniques: the so-called CABAMAKA system is a tool for semi-automatic adaptation knowledge acquisition. The CABAMAKA system relies on a data mining technique for extracting adaptation knowledge from a case base, namely the search for frequent closed itemsets [d'Aquin et al., 2007].

Finally, in the same way, the integration of semi-automatic adaptation knowledge extraction and acquisition methods driven by experts is another topic of interest. Besides knowledge discovery, another ongoing work in the KASIMIR system concerns the definition and the implementation of fuzzy predicates over numerical concrete domains, with an associate inference engine (for extending work introduced in [d'Aquin et al., 2006a]).

References

Agnar Aamodt. Knowledge-Intensive Case-Based Reasoning and Sustained Learning. In *Proc. of the 9th European Conference on Artificial Intelligence (ECAI'90)*, pages 1–6, 1990.

Agnar Aamodt. *A Knowledge-Intensive, Integrated Approach to Problem Solving and Sustained Learning*. Doctoral dissertation, University of Trondheim, Norway, 1991.

Agnar Aamodt. Knowledge-Intensive Case-Based Reasoning in CREEK. In P. Funk and P. A. Gonzàlez-Calero, editors, *Proc. of the European Conference on Case-Based Reasoning, ECCBR'04*, volume 3155 of *Lecture Notes in Artificial Intelligence*, pages 1–15. Springer, 2004.

Agnar Aamodt and Enric Plaza. Case-based reasoning: Foundational issues, methodological variations, and system approaches. *AICom - Artificial Intelligence Communications*, 7(1):39–59, 1994.

Grigoris Antoniou and Frank van Harmelen. *A Semantic Web Primer*. The MIT Press, 2004.

Franz Baader, Diego Calvanese, Deborah McGuinness, Daniele Nardi, and Peter Patel-Schneider, editors. *The Description Logic Handbook*. Cambridge University Press, cambridge, UK, 2003.

Paolo Besana, Vivek Patkar, David Glasspool, and Dave Robertson. Distributed workflows: The open-knowledge experience. In Robert Meersman, Zahir Tari, and Pilar Herrero, editors, *OTM Workshops*, volume 5333 of *Lecture Notes in Computer Science*, pages 965–975. Springer, 2008. ISBN 978-3-540-88874-1.

Alex Borgida and Luciano Serafini. Distributed Description Logics: Assimilating Information from Peer Sources. *Journal of Data Semantics*, 1:153–184, 2003.

Alexander Borgida and Luciano Serafini. Distributed description logics: Directed domain correspondences in federated information sources. In *CoopIS/DOA/ODBASE*, pages 36–53–, 2002.

Jacques Bouaud and Brigitte Séroussi. Characterizing the dimensions of clinical practice guideline evolution. In *Proceedings of MIE2008*, Studies in Health Technology and Informatics 136, pages 139–144. IOS Press, 2008.

Paolo Bouquet, Fausto Giunchiglia, Frank van Harmelen, Luciano Serafini, and Heiner Stuckenschmidt. Contextualizing Ontologies. *Journal of Web Semantics*, 1(4):325–343, 2004.

William J. Clancey and Edward H. Shortliffe, editors. *Readings in Medical Artificial Intelligence: The First Decade*. Addison-Wesley, Reading, Massachusetts, 1984.

Mathieu d'Aquin, Sébastien Brachais, Jean Lieber, and Amedeo Napoli. Decision Support and Knowledge Management in Oncology using Hierarchical Classification. In *Proc. of the Symposium on Computerized Guidelines and Protocols, CGP-2004*, volume 101 of *Studies in Health Technology and Informatics*, pages 16–30. IOS Press, 2004.

Mathieu d'Aquin, Sébastien Brachais, Christophe Bouthier, Jean Lieber, and Amedeo Napoli. Knowledge Editing and Maintenance Tools for a Semantic Portal in Oncology. *International Journal of Human-Computer Studies (IJHCS)*, 62(5):619–638, 2005a.

Mathieu d'Aquin, Jean Lieber, and Amedeo Napoli. Decentralized Case-Based Reasoning for the Semantic Web. In *Procedings of International Semantic Web Conference, ISWC 2005*, volume 3729 of *Lecture Notes in Computer Science*, pages 142–155. Springer, 2005b.

Mathieu d'Aquin, Jean Lieber, and Amedeo Napoli. Towards a Semantic Portal for Oncology using a Description Logic with Fuzzy Concrete Domains. In E. Sanchez, editor, *Fuzzy Logic and the Semantic Web*, pages 379–393. Elsevier, 2006a.

Mathieu d'Aquin, Jean Lieber, and Amedeo Napoli. Adaptation Knowledge Acquisition: a Case Study for Case-Based Decision Support in Oncology. *Computational Intelligence (an International Journal)*, 22(3/4):161–176, 2006b.

Mathieu d'Aquin, Jean Lieber, and Amedeo Napoli. Case-based reasoning within semantic web technologies. In Jérôme Euzenat and J. Domingue, editors, *Proceedings of the 12th International Conference on Artificial Intelligence: Methodology, Systems, Applications (AIMSA 2006), Varna, Bulgaria, 13-15th September, 2006*, LNAI 4183, pages 190–200. Springer, Berlin, 2006c.

Mathieu d'Aquin, Fadi Badra, Sandrine Lafrogne, Jean Lieber, Amedeo Napoli, and Laszlo Szathmary. Case base mining for adaptation knowledge acquisition. In M.M. Veloso, editor, *IJCAI 2007, Hyderabad, India*, pages 750–755. Morgan Kaufman, 2007.

Mathieu d'Aquin, Jean Lieber, and Amedeo Napoli. Decision support systems in oncology: Are we there yet? *Oncology*, 9(11), 2008. www.hcplive.com/mdnglive/ONCNG-Oncology/Nov2008.

Yves Demazeau and Jean-Pierre Müller. Decentralized Artificial Intelligence. In Y. Demazeau and J.-P. Müller, editors, *Decentralized A.I. – Proc. of the First European Workshop on Modelling Autonomous Agents in a Multi-Agent World*, pages 3–13. North-Holland, 1989.

Elias Egho, Nicolas Jay, Chedy Raïssi, and Amedeo Napoli. A FCA-based analysis of sequential care trajectories. In Amedeo Napoli and Vilem Vychodil, editors, *The Eighth International Conference on Concept Lattices and their Applications - CLA 2011*, pages 363–376, Nancy, France, 2011. INRIA Nancy Grand Est - LORIA.

Evidence-based medicine working-group. Evidence-based medicine. A new approach to teaching the practice of medicine. *Journal of the American Medical Association*, 17:268, 1992.

Amir Ghazvinian, Natalya Fridman Noy, Clement Jonquet, Nigam H. Shah, and Mark A. Musen. What four million mappings can tell you about two hundred ontologies. In *International Semantic Web Conference*, pages 229–242, 2009.

M. Gómez-Albarrán, P.A. Gonzàles-Calero, B. Díaz-Agudo, and C. Fernàndez-Conde. Modelling the CBR Life Cycle Using Description Logics. In K.-D. Althoff, R. Bergamnn, and L.K. Branting, editors, *Proc. of the International Conference on Case-Based Reasoning, ICCBR'99*, volume 1650 of *Lecture Notes in Artificial Intelligence*, pages 147–161. Springer, 1999.

Pascal Hitzler, Markus Krötsch, and Sebastian Rudolph. *Foundations of Semantic Web Technologies*. CRC Press, Boca Raton (FL), 2009.

Martin Homola and Luciano Serafini. Augmenting Subsumption Propagation in Distributed Description Logics. *Applied Artificial Intelligence*, 24:39–76, 2010.

Ian Horrocks, Peter F. Patel-Schneider, and Frank van Harmelen. From SHIQ and RDF to OWL: the making of a Web Ontology Language. *Journal of Web Semantics*, 1(1):7–26, decembre 2003.

Sajjad Hussain and Syed Sibte Raza Abidi. Integrating healthcare knowledge artifacts for clinical decision support: Towards semantic web based healthcare knowledge morphing. In Carlo Combi, Yuval Shahar, and Ameen Abu-Hanna, editors, *AIME*, volume 5651 of *Lecture Notes in Computer Science*, pages 171–175, 2009. ISBN 978-3-642-02975-2.

Nicolas Jay, Amedeo Napoli, and François Kohler. Cancer patient flows discovery in drg databases. In A. Hasman, R. Haux, J. Van Der Lei, E. De Clercq, and F.H. Roger, editors, *Ubiquity: Technologies for Better Health in Aging Societies. Proceedings of MIE 2006*, Studies in Health Technology and Informatics, pages 725–730. IOS Press, 2006.

Nicolas Jay, François Kohler, and Amedeo Napoli. Analysis of social communities with iceberg and stability-based concept lattices. In Raoul Medina and Sergei Obiedkov, editors, *Formal Concept Analysis – Proceedings of the 6th International Conference on FCA (ICFCA 2008), Montréal*, Lecture Notes in Artificial Intelligence 4933, pages 258–272. Springer, Berlin, 2008.

Yevgeny Kazakov, Markus Kroetzsch, and Frantisek Simancik. Practical reasoning with nominals in the el family of description logics. In Gerhard Brewka, Thomas Eiter, and Sheila A. McIlraith, editors, *KR*. AAAI Press, 2012.

Janet Kolodner. *Case-Based Reasoning*. Morgan Kaufmann, 1993.

Markus Krötzsch, Frantisek Simancik, and Ian Horrocks. A description logic primer. Technical report, Department of Computer Science, University Of Oxford, UK, 2012. `http://arxiv.org/abs/1201.4089`.

Jean Lieber and Amedeo Napoli. Using classification in case-based planning. In Wolfgang Wahlster, editor, *Proc. of the 12th European Conference on Artificial Intelligence (ECAI-96)*, pages 132–136. John Wiley & Sons, Ltd., 1996.

Jean Lieber, Mathieu d'Aquin, Pierre Bey, Amedeo Napoli, Maria Rios, and Catherine Sauvagnac. Acquisition of Adaptation Knowledge for Breast Cancer Treatment Decision Support. In *Procedings of the 9th Conference on Artificial Intelligence in Medicine in Europe, AIME 2003*, 2003.

Carsten Lutz. Description Logics with Concrete Domains – A Survey. In *Advances in Modal Logics Volume 4*. King's College Publications, 2003.

Carsten Lutz and Maja Milicic. A tableau algorithm for description logics with concrete domains and general tboxes. *Journal of Automated Reasoning*, 38(1-3):227–259, 2007.

Carsten Lutz, Carlos Areces, Ian Horrocks, and Ulrike Sattler. Keys, nominals, and concrete domains. *Journal of Artificial Intelligence Research (JAIR)*, 23:667–726, 2005.

Sandra Marcus, editor. *Automatic Knowledge Acquisition for Expert Systems*. Kluwer Academic Publishers, Boston, 1988.

Brian McBride. Jena: a semantic Web toolkit. *Internet Computing, IEEE*, 6(6):55–59, 2002.

Thomas Meilender, Jean Lieber, Fabien Palomares, and Nicolas Jay. From web 1.0 to social semantic web: Lessons learnt from a migration to a medical semantic wiki. In *ESWC*, pages 618–632, 2012.

Erica Melis, Jean Lieber, and Amedeo Napoli. Reformulation in case-based reasoning. In Barry Smyth and Peter Cunningham, editors, *Proceddings of the Fourth European Workshop on Case-Based Reasoning, EWCBR-98*, Lecture Notes in Artificial Intelligence 1488, pages 172–183. Springer, 1998.

Amedeo Napoli, Jean Lieber, and Régis Curien. Classification-Based Problem-Solving in Case-Based Reasoning. In B. Faltings and I. Smith, editors, *Third European Workshop on Case-Based Reasoning (EWCBR'96), Lausanne*, Lecture Notes in Artificial Intelligence 1168, pages 295–308. Springer-Verlag, Berlin, 1996.

Natalya Fridman Noy, Nigam H. Shah, Patricia L. Whetzel, Benjamin Dai, Michael Dorf, Nicholas Griffith, Clement Jonquet, Daniel L. Rubin, Margaret-Anne D. Storey, Christopher G. Chute, and Mark A. Musen. Bioportal: ontologies and integrated data resources at the click of a mouse. *Nucleic Acids Research*, 37(Web-Server-Issue):170–173, 2009.

Jeff Pan and Ian Horrocks. OWL-Eu: Adding Customised Datatypes into OWL. In Asun Gómez-Pérez and Jérôme Euzenat, editors, *Proceedings of The Second European Semantic Web Conference, ESWC 2005*, number 3532 in Lecture Notes in Computer Science, pages 153–166. Springer, 2005.

V. Patkar, V. Hurt, R. Steele, S. Love, A. Purushotham, M. Williams, R. Thomson, and J. Fox. Evidence-based guidelines and decision support services: a discussion and evaluation in triple assessment of suspected breast cancer. *British Journal of Cancer*, 95:1490–1496, 2006.

M. Peleg, S. Tu, J. Bury, P. Ciccarese, J. Fox, R.A. Greenes, R. Hall, P.D. Johnson, N. Jones, A. Kumar, S. Miksch, S. Quaglini, A. Seyfang, E. H. Shortliffe, and M. Stefanelli. Computer-interpretable Guideline Models : A Case-study Approach. *Journal of the American Medical Informatics Association*, 10:52–68, 2003.

Chris K. Riesbeck and Roger C. Schank. *Inside Case-Based Reasoning*. Lawrence Erbaum Associates, Inc., Hillsdale, New Jersey, 1989.

M. Rios, E. Desandes, B. Bresson, I. Klein, A. Lesur, F. Boisson, V. Demange, and P. Bey. Référentiels de bonnes pratiques cancérologiques : étude comparative de trois supports d'aide à la proposition thérapeutique pour les cancers du sein et de la prostate en Lorraine. *Bulletin du cancer*, 90(4):363–370, 2003. Clinical guidelines in oncology: a comparative study of three decision support systems for breast and prostate cancer in Lorraine French region, in French.

Sylvie Salotti and Véronique Ventos. Study and Formalization of a Case-Based Reasoning System Using a Description Logic. In Barry Smyth and Padraig Cunningham, editors, *Advances in Case-Based Reasoning, Proc of the 4th European Workshop, EWCBR-98*, volume 1488 of *Lecture Notes in Artificial Intelligence*, pages 286–297. Springer, 1998.

Stefan Schulz and Udo Hahn. A Description Logic Approach to Clinical Guidelines and Protocols. In *Procedings of the Symposium on Computerized Guidelines and Protocols (CGP-2004)*, 2004.

Luciano Serafini. Distributed reasoning for semantically related domains. Technical report, ITC-IRST, Trento, 2007.

Luciano Serafini and Martin Homola. Contextualized knowledge repositories for the semantic web. *Journal of Web Semantics*, 12:64–87, 2012.

Luciano Serafini and Andrei Tamilin. Local Tableaux for Reasoning in Distributed Description Logics. In Volker Haarslev and Ralf Moeller, editors, *Proceedings of the International Workshop on Description Logics, DL'04, Whistler, BC, Canada*, pages 100–109. CEUR-WS, Volume 104, 2004.

Luciano Serafini and Andrei Tamilin. DRAGO: Distributed reasoning architecture for the semantic web. In A. Gomez-Perez and J. Euzenat, editors, *Proc. of the Second European Semantic Web Conference (ESWC'05)*, volume 3532 of *Lecture Notes in Computer Science*, pages 361–376. Springer, May 2005.

Evren Sirin, Bernardo Cuenca Grau, and Bijan Parsia. From wine to water: Optimizing description logic reasoning for nominals. In Patrick Doherty, John Mylopoulos, and Christopher A. Welty, editors, *KR*, pages 90–99. AAAI Press, 2006.

Evren Sirin, Bijan Parsia, Bernardo Cuenca Grau, Aditya Kalyanpur, and Yarden Katz. Pellet: A Practical OWL-DL Reasoner. *Journal of Web Semantics*, 5(2):51–53, 2007.

Barry Smyth. *Case-Based Design*. PhD. thesis, Trinity College, University of Dublin, April 1996.

Andrei Tamilin. *Distributed ontological reasoning: theory, algorithms, and applications*. Phd dissertaion, ITC - IRST, Trento, 2007.

Jan H. van Bemmel and Mark A. Musen, editors. *Handbook of Medical Informatics*. Springer-Verlag, Heidelberg, 1997.