IST-FP6-508794

PROTOCURE II

Integrating formal methods in the development process of medical guidelines and protocols

Specific Targeted Research Project
Information Society Technologies

Deliverable D2.5: Library of design patterns for guidelines

Due date of deliverable: 31 December 2004
Actual submission date: 21 December 2004

Start date of project: 1 January 2004 **Duration:** 30 months

Organisation name of lead contractor: Universitat Jaume I, Castellon, Spain

Revision: final

Project co-funded by the European Commissions within the Sixth Framework Programme(2002-2006)		
Dissemination Level		
PU	Public	√
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Deliverable D2.5:
Library of Design Patterns for Guidelines

Radu Serban, Annette ten Teije, Frank van Harmelen
Vrije Universiteit Amsterdam

Mar Marcos, Cristina Polo, Juan Carlos Galan
Universitat Jaume I

Peter Lucas, Arjen Hommersom
Radboud Universiteit Nijmegen

Kitty Rosenbrand, Jolanda Wittenberg, Joyce van Croonenborg
Dutch Institute for Healthcare Improvement

Contents

List of Figures

Part I
Introduction

1 Scope of our work

Guidelines represent a way of disseminating a particular code of conduct in a specific field.

Clinical guidelines describe the best practice in the medical domain in a narrative form, including recommendations and regulations that cover either typical or exceptional situations and background information showing alternative methods and their effects.

Definition 1 (Clinical Evidence-based Guideline) *An* evidence-based clinical guideline *is a document with recommendations, advice and management instructions aimed at supporting the decision-making process of health-care professionals and patients, based on the results of scientifi c research with associated discussions and opinion formation, with the objective of making suitable and effective medical management explicit.*

It was argued ([Pro]) that background medical knowledge (concerning the symptoms of a disease, available drugs, methods for treatment and their possible effects, medical scenarios relevant when dealing with a particular disease) is captured using a similar format, regardless of the disease the treatment is prescribed for. Furthermore, the guidelines have a similar structure, containing scientific justification (emphasizing background knowledge relevant for the guideline), summary statement (conclusion), other considerations and recommendations.

This triggered a search for *re-usable components in evidence-based clinical guidelines*, which can be used as building blocks for producing a sketch of a clinical guideline, or for a more generic guideline. These building blocks are even more important if one considers their role in reducing the effort necessary for formalizing a medical guideline, in order to bring it into an executable form.

Finding re-usable components in clinical guidelines expressed in plain text is a difficult problem, partly due to the ambiguity of the language, and partly due to a lack of modularity and clear relations between the contexts that are described as part of the scenarios covered by the textual description.

We describe how an instance of this problem can be solved for a particular category of guidelines, namely oncology guidelines. More precisely, we study different clinical guidelines for treatment of breast cancer and lung cancer, respectively. The components common to these guidelines seem to be general enough to apply to a larger set of clinical guidelines. These guidelines have a modular structure, with one section dedicated to each medical procedure and medical scenario that is likely to employ difficult, varied or critical actions, where knowledge gaps are likely to make difficult for the right alternative to be chosen. The vocabulary of the guidelines is rather concise, including a compact set of medical terms from a medical ontology shared by healthcare professionals.

The problem is therefore simpler: given a text that contains terms from two vocabularies - a medical one, and a guideline-specific one -, and conforms to a limited number of syntactical constructions, the problem is to identify the most frequently used linguistic components that include terms from both vocabularies and are likely to capture some of the essential knowledge the guideline is designed to convey.

The identification of frequently used semantic patterns in clinical guidelines is the first step of one of the recently proposed methods for guideline formalization. Related work ([Svatek2004]) list a few requirements for these semantic patterns and is overviewed in section 4.6.

1.1 Living Guidelines

Authoring and maintenance of medical guidelines, which represent a way of sharing the state-of-the-art in medical knowledge, has been recognized as an important factor for improving the quality of healthcare, but at the same time also as a resource-consuming process. Guidelines for healthcare processes are frequently changing due to new research and technology improvements. The changes concern mostly parts of the guideline, therefore not requiring rewriting of the entire guideline.

The Protocure-II project focuses on **living clinical guidelines**, that is, on clinical guidelines which are regularly reviewed and updated to incorporate the latest scientific results in medicine.

To enable the update of a guideline after a revision, with a minimum of effort, we believe that a modularization of the guideline should be based upon a software engineering approach: the key questions, requirements, literature references and the relevant medical background knowledge are to be captured in a medical guideline project (MGP), together with a modular text representation of the guideline. The requirements and goals of the guideline should be linked with their textual representation, therefore a change of these constraints can immediately be operated at the textual level.

Protocure-I ([Pro]) project concluded that the splitting of the formalization process into finer-grained steps could bring benefits to the quality of the formal model of a guideline. It was also believed that experience in formalization can be captured in so-called guideline patterns, that improve the effectiveness of the formalization when the same steps are repeated for other guidelines.

For the modularization of the text fragments that describe similar knowledge and in order to make possible mappings between this knowledge and a more formal representation of it, we use a component called **guideline pattern** (GP). The project Protocure II ([Pro]) investigates the role of GPs in improving formalization and update of medical guidelines.

In this report we propose a set of pattern templates to support a multiple-step formalization process of a medical guideline and show how guideline authoring, formalization and maintenance of clinical guidelines can benefit from the use of medical guideline patterns.

1.2 Motivation

The translation of the text of a guideline into a formal and/or executable form is called *guideline formalization*. The guideline formalization helps in achieving several goals:

1. validation and critiquing of the guideline recommendations: by simulating the execution of the formal model of a guideline, guideline developers can discover inconsistencies, redundancies, ambiguities caused by statements in the text and can suggest improvements.

2. verification of guidelines: guideline developers can verify whether desirable properties which the guideline is set to achieve (such as, reducing the number of steps employed by a treatment) hold

3. decision-support for medical professionals: by simulating the execution of the formal model, the medical specialists can obtain suggestions about the feasible possibilities available at a particular point in the treatment.

4. update or authoring of guidelines: the formal model of a guideline can help in deciding which processes are essential and which processes need to be further refined/documented in the text of the guideline, to increase the comprehensibility of the text.

So far, the formalization process has been performed from scratch each time when a change of the guideline occured. This process would require less effort if not all translations are performed, but some shortcuts for translation are stored and used, in the form of guideline patterns. The guideline patterns define a mapping between a textual representation of a concept or situation that occurs frequently in a medical scenario and an equivalent, but more formal, representation of the same entity at another level of abstraction.

The main goal of this report is to investigate to what extent a set of *medical guideline patterns* (MGPs) can be built, that:

- makes guideline formalization more effective, through the reuse of the guideline structures, such as recommendations, in several contexts, independent of the disease addressed by the guideline;

- aids authoring of new guidelines.

Furthermore, we study what knowledge is necessary for finding guideline patterns and whether finding MGPs is a process that can partly be automated. For this goal, we define an algorithm for searching pattern templates and their instances in the guideline text. The algorithm makes explicit the fact that there are steps when a human decision is needed in order to validate a pattern template suggested as frequent template by the program, or to confirm that a given text fragment (representing a pattern instance) actually represents the instantiation of a pattern template.

1.3 Structure of the deliverable

This report is organized as follows.

In the rest of this section we give an introduction to medical guideline patterns, guideline formalization, and summarize the objectives of this deliverable. In section 4 of this chapter we describe the different categories of patterns analyzed. Each of these categories are dealt with in a separate section in the document.

Sections II, III, IV, V and VI present several complementary approaches to searching patterns in order to support guideline formalization.

In section VII we discuss how the different classes of patterns can interact and can be used in guideline formalization. Then we summarize the lessons learned from the task of this deliverable and make a few suggestions on the directions that can be investigated in other tasks of the Protocure-II project.

The appendix contains some results obtained when building a program for automated search of linguistic patterns and running it on a selected guideline. Section A provides some statistics concerning the medical vocabularies shared by three oncology guidelines. In fact, it is the rather high overlapping of medical terms found in these vocabularies specific to different guidelines, which motivated and supported our search for linguistic patterns. Section B.1 provides the mini-ontology of the medical domain used by the pattern search program. Section B.2 provides a sample of automated analysis output on a piece of the guideline selected by us for formalization. Sections C, D and E contain the lists of control, background and linguistic patterns identified by us.

2 Background

2.1 Guideline document

A guideline can contain several types of information:

- Data abstractions: concept definitions (for instance, drug terminology, medical procedures)
- Data values: range of accepted values for a clinical parameter, and values considered abnormal
- Medical procedures:
  - A general division in steps and subplans that have to be executed by medical staff, including concurrent, cyclical or iterative actions
  - Decisions and choices that have to be taken by the medical staff executing the protocol
- Policies independent of the medical implementation of the guideline

The medical process captured by a clinical guideline, seen at a very high level of abstraction, consists of the following processes:

```
Do (sequentially)
(optional) Pre-treatment checks
Do (in parallel)
    Diagnosis | Monitoring parameters | Treatment | Handling of critical situations
(optional) Post-treatment patient supervision
```

The guideline places constraints on at least one of these processes, indicating: what subprocesses they consist of, what medication is used, in which order and under which timing constraints they are to be executed, which combination of processes should be avoided, etc.

2.2 Guideline Formalization Process

In our proposal, the formalization process consists of a step-by-step guideline refinement method that starts from the text of the guideline and a set of guideline development-specific vocabularies, and at each refinement step first identifies pattern instances at the current representation level, and then uses macros to transform all instances of these pattern templates in their corresponding (more formal) representations. Using this *pattern-centered guideline formalization process*, one can generate a formal representation sketch of the guideline more effectively than not using any transformation shortcuts in formalization (as it was the case until now). The sketch of the guideline resulted after applying the formalization patterns often represent an incomplete image of the knowledge designed to be communicated by the guideline, on which subsequent refinement steps are necessary. The refinement of the formalization has to take into account details and less frequent constructs that could not be captured by the formalization patterns and are more specific to the medical scenario modelled by the guideline.

2.3 Living Guidelines: a knowledge management problem

A (narrative) guideline presents the medical care process from several perspectives and in a variety of forms: narrative textual description, tabular form, process flow diagrams, hierarchical tasks diagrams. Irrespective of the format used to convey a set of clinical statements, these linguistic wrappers of knowledge have one thing in common: disseminate knowledge that answers the following questions:

1. What information, processes and resources are needed for a medical organisation to follow the standard medical practice when performing tasks such as management of a disease? This requires definition of key concepts, goals to be achieved, medical actions available to achieve a task, etc.

2. What resources are used,changed, created or destroyed when a medical process is executed? The dependencies between medical processes are important when deciding which resources are essential and where the changes have to take place.

3. Who carries out the medical processes? It may require a description of the roles involved, including issues of authority (who is responsible for a process), per-

mission (who is allowed to carry out a process) and obligation (who is liable for the consequences of a process).

4. Under which circumstances should a process be carried out? This is done by specifying some control (timing, ordering, synchronization, hierarchical) constraints on clinical processes involved: which processes must come before others, which have to be carried out within a certain time, etc.

5. Why a process has to be carried out? It represents an explanation of the rationale for performing a process, in a narrative textual form.

Guidelines are a form of knowledge sharing. Living guidelines, which involve maintenance of up-to-date knowledge over a longer period, depend on a much more ellaborate knowledge management for knowledge sharing.

This problem of maintaining knowledge for dissemination of medical practice is a cyclic process of identifying the key questions the guideline is set to answer, analyzing the knowledge issues that are potentially relevant in order to answer these questions, select the knowledge components that need to be improved (better covered) and the actions required to improve them, then evaluate whether the actions are feasible and whether the new knowledge has eliminated all problems. This process is summarized in figure 1.

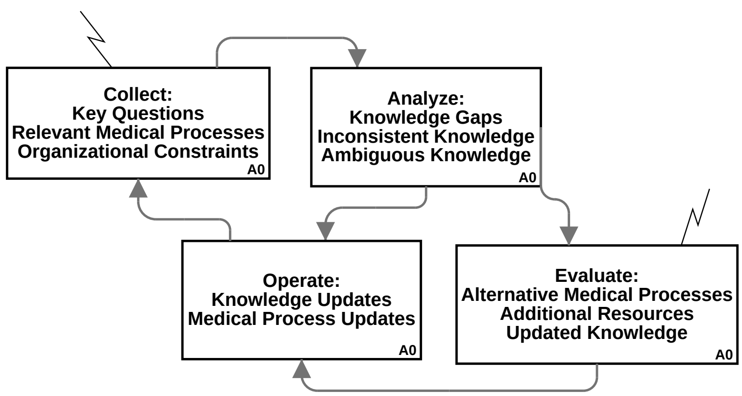

Figure 1: *The Knowledge Management Process of Living Guidelines*

If a knowledge problem can be solved better than at present, then updates are operated in the knowledge captured by the guideline and in the prescribed medical processes. This triggers a new loop in the knowledge management cycle.

To determine whether a knowledge item still remains relevant for a healthcare process (and therefore should be considered to be included in the guideline text), that piece of knowledge has to play one of the following roles:

1. avoids/reduces a knowledge bottleneck;

2. eliminates a knowledge gap;
3. provides better quality knowledge;
4. communicates effectively and unambiguously the existing knowledge.

Knowledge Bottlenecks occur where there is a problem with the availability of knowledge. This is often because the knowledge is very specialised, scarce, or not explicit.

Knowledge Gaps occur where the knowledge needed for making decisions is not explicit or not present within their organisation: practitioners need knowledge that they do not have. In some cases, this may occur because new knowledge has arisen which requires a refinement of existing knowledge, or, more often, because no-one has investigated the clinical target process before.

Knowledge Quality issues usually arise when the knowledge is dynamic. Providing the best quality knowledge requires frequent updates to the knowledge, which are expensive and may consume more time than the knowledge is worth. Knowledge quality issues may also arise in an organisation which is not using the best available knowledge because of individual resistance (practitioners prefer 'traditional methods'), high costs for refining/updating the knowledge, fear of negative consequences of applying that knowledge, or information overload (too much information for a practitioner to process in the time available).

Knowledge Communication is concerned with how to make knowledge understandable for the guideline target group and how to communicate effectively conclusions and recommendations.

A high-quality clinical guideline has the goals mentioned above, and one of our objectives is to develop mechanisms that allow guideline developers to achieve these goals.

2.4 Guideline pattern

A *pattern* is a recurring solution to a standard problem, which captures experience in a way that makes possible for others to reuse this experience. A guideline pattern in our case represents a relation between a medical context (defined by the application of a particular treatment, a special condition of a patient, a decision for a particular health-care path) and at least one other (typically, more formal) representation, at another level of abstraction or in another stage of the guideline development process.

Basically, the guideline patterns facilitate the translations between statements in the medical guideline (the textual representation, denoted TR), their intermediate representation (denoted IR) and/or their formal representation (denoted FR), which would reduce the formalization effort for transforming guidelines into a machine-processable form and also support understanding and explanation of properties for formally represented guidelines.

3 Objectives

3.1 Identification of re-usable components in guidelines

This deliverable is the result of completing task T2.5 of Protocure-II project. The goal of this task is to provide guideline developers with a number of patterns that appear frequently in medical guidelines, which can be reused (e.g. by filling in/adapting parts) in the development of a new guideline.

These re-usable guideline constructs called guideline patterns, are meant to support:

- the formalization and validation of medical guidelines in a systematic, structured way
- the verification of properties, critiquing and execution of guidelines.

The patterns that were initially investigated were primarily action patterns, and we assumed they could be either very abstract (e.g., describing relations between diagnostic and treatment or between treatment methods and available medication) or very specific to a disease (or category of diseases). Furthermore, as the target of the formalization process was a semi-formal, executable, ASBRU representation ([**?**, Seyfang2002]), it was assumed that the patterns will be derived based on our modelling experiences in Protocure I and II projects, and will mainly consist of ASBRU components.

However, since an important feature of Protocure is the ”hand-in-hand” development of guidelines and their Asbru counterparts, there has been a shift in perspective, rather an enlargement, concerning the activities to deal with task T2.5. Mainly, we have decided not to restrict pattern search to our past and current modelling experiences, but also to directly analyse guideline documents searching for recurrent elements. The intuition is that there might be similarities in the original documents, independently of the fact that the same modelling artifacts are used to describe them is Asbru. Furthermore, since textual representation of the guideline is the result of an elaborated process of knowledge refinement, we considered as patterns the knowledge refinement that results in a particular type of information (for instance, what is the effectiveness of a particular drug in treating a disease, what are the alternatives for dealing with a disease, etc) being included in the guideline text.

The patterns we have identified so far using these approaches (namely knowledge driven, text driven and Asbru model driven), are diverse in nature. They range from basic control/action patterns to domain-knowledge ones, to document formatting/structure patterns.

It seems that the patterns identified from a textual analysis of original guidelines are less complex than the ones obtained from exisiting Asbru models. The reason is that building an Asbru model requires a top-down design effort to make all the individual actions/plans fit in a hierarchical plan. Therefore, Asbru models include many complex structures which, on the other hand, are not often explicit in the original guideline.

3.2 Use of the library in guideline formalization, validation and authoring

The patterns that will make up the library of guideline components are used for several purposes:

- For simplifying the validation of the most frequently used constructs, by comparing the Textual and the Intermediate Representation of the same pattern instance;
- For making the formalization more straightforward and effective, by reusing knowledge about the complexity of certain transformations and storing them in a mapping pattern template;
- For allowing translation of formal fragments into a language understood by medical experts, thus facilitating verification of properties of the healthcare process prescribed by the guideline;
- For simplifying verification of properties of patterns: by being modular and having typed arguments (belonging to a small number of semantic categories), the properties of a pattern template can be verified and some of them can be assumed to hold for all instances of that pattern.

4 Classes of Patterns

There are several approaches to searching guideline patterns. The approach adopted by us is that of a step-refinement approach, in which the knowledge captured in a guideline is transformed according to one viewpoint in each step, in order to structure it to a sufficient extent to make it formalizable.

In the context of living-guidelines, the transformations should ideally represent knowledge invariants, but this situation has not been achieved in practice by the existing guideline development methods, due to the lack of a balance between expressivity and complexity between the language used to design guidelines (most often, natural language) and the language used to validate and verify properties of the guidelines (semi-formal and executable languages, such as ASBRU).

We considered the following perspectives in which the information captured in the guideline text can be refined:

The knowledge approach : the guideline text is viewed as interaction between several semantic structures capturing knowledge from the background domain; this approach produces the so-called background patterns;

The functional approach : the guideline text is viewed as a composition of several instances of programming-like control structures; these structures produce control patterns;

The linguistic approach : the guideline text is seen as an interaction of syntactical patterns that either capture background knowledge, or correspond to control structures; they have a very fine level of granularity and make up the linguistic patterns;

The knowledge-driven refinement approach : the guideline text and its executable representation is seen as a sequence of transformations/mappings between narrative structures, medical background knowledge, control structures, and executable modules.

In the Protocure II project, the guideline formalization process is done in multiple steps. First, the textual guideline is transformed into a set of narrative structures, then decomposed according to several dimensions (domain-knowledge, functional aspects, data structures and synchronization constraints), then re-composed into a semi-formal guideline representation using the ASBRU language and possibly furher refined into a formal representation in KIV, the formal language that is used to verify guideline properties.

These dimensions and formalization steps are captured by the Intermediate Representation language 8HB.

Definition 2 (Generic definition of pattern) *Given two representations of the same guideline, as a set of facts expressed in two different languages, we distinguish two types of patterns:*

- *a* pattern of representation *is defi ned as a frequently-used element or combination of elements in one of the representations;*
- *a* transformation pattern *is defi ned as a mapping between a set of elements in the fi rst representation and a set of elements in the second representation.*

In guideline formalization we are interested to obtain patterns of transformation, but in order to achieve this goal we have to identify the right representations that can allow us to search for patterns and to identify the representation patterns first.

Identification of patterns can take place at several levels. Based on the abstraction level at which they occur, the following classes of patterns (shown in figure 2) can be identified:

structural patterns : syntactical patterns in the original text of the guideline, i.e., source fragments that have similar formulation and are used several times in the guideline.

linguistic patterns : syntactical patterns at the intermediate representation level, i.e., source fragments that are translated into the same IR representation.

background patterns : patterns at the intentional level; intermediate representation or source fragments that correspond to similar control structures, even if their source and IR representation differ; these patterns require medical expertise from a medical expert or the guideline author.

control patterns : patterns that use the same control structures and describe actions with the same complexity;

implementation patterns : patterns at the implementation (formal representation) level, e.g., plans that are made up of the same control structures, that depend on the same activation conditions, that take similar arguments and depend on the same parameters, that invoke the same plan activations, etc.

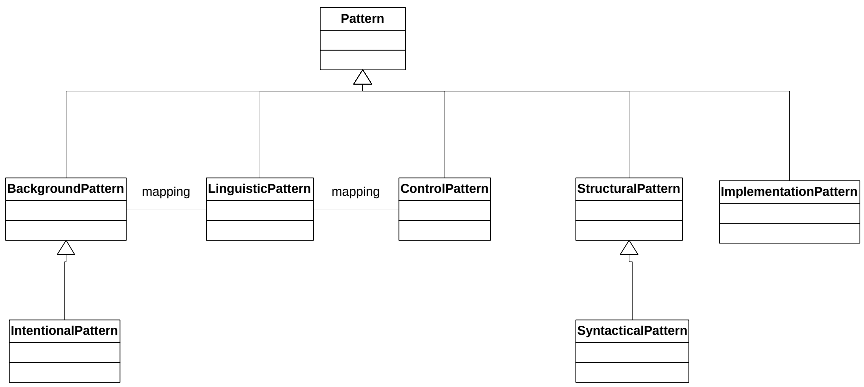

Figure 2: *Categories of patterns*

These classes are extracted from representations of guideline text situated at different levels of abstraction. As the figure suggests, a clear mapping can be established between linguistic pattern instances and control and background pattern instances. This means that: 1. a linguistic pattern template can be identified for some (but not all) control and background pattern templates; 2. a text fragment in the medical guideline can be at the same time an instance of a control pattern instance, a background pattern instance, and a linguistic pattern instance.

4.1 Structural patterns

Structural patterns represent syntactical structures extracted from the guideline text or synthesized by the guideline authors. They can be viewed as classification of the fragments in the guideline text along very general narrative structures. This category of patterns is discussed in section 5.

4.2 Linguistic patterns

Linguistic patterns are frequently-used components that are abstracted from the Intermediate Representation (IR) of the guideline using syntactic rules and knowledge specific to the domain of the guideline. This category of patterns is discussed in section 6.

4.3 Background patterns

These are components which describe relations between the entities in the medical domain that play an essential role in dealing with the disease handled by the guideline. They are abstracted from the guideline text or pre-compiled from existing background knowledge by the guideline (GL) authors.

This category of patterns is discussed in section 7.

4.4 Control Patterns

These are components that describe programming-like structures which can be mapped into the target (formal or executable) model. They are extracted by knowledge engineers from the Intermediate Representation (IR) of the guideline (GL) text and depend to some extent on the target Formal Representation used. However, the control patterns identified by us are rather ASBRU-independent.

This category of patterns is discussed in section 8.

4.5 Implementation Patterns

These are components extracted from the formal/executable model of the medical process described in the guideline. They are based on existing formalizations of guidelines, and not on the actual text of the guideline. Knowledge Engineers establish certain action timing, ordering and synchronization constraints between the processes involved, that can be reused in other guidelines as well. This category of patterns is discussed in section 9.

4.6 Related Work on Semantic Patterns

The area of extracting information from annotated and not annotated text for supporting information retrieval and querying and natural language understanding is one of the research areas related to our text-driven approach to patterns ([Riloff1996, Riloff and Shoen1995, Moreno and Perezmay 2001, Huffman1995]). Text and data mining approaches have been used to extract vocabularies or more complex syntactic constructs from untagged text ([Frank and Nevill-Manning1999, Riloff1996, Riloff and Shoen1995]), or guided by the use of an explicit dictionary, ontology, or positive examples ([Huffman1995]).

In another related area of research, the topic of software design patterns [Budinsky *et al.*1996, Beck *et al.*1996, Vlissides1997] has received a lot of attention in recent years in the Software Engineering community: [Fernandez and Yuan2000, Zimmer2001, Fowler1997, Florijn *et al.*1997, Coplien1998, Atkinson, von Dincklage, Ben-Yehuda, Hirschfeld and Lämmel2004, Bunnig *et al.*, Forbrig and Lammel, Appleton1997, Eden *et al.*1997, Riedewald]. Software design pattern catalogs exist, in which each pattern collects relatively independent solutions to a common software design problem.

Similar to these software design patterns, design patterns for the medical domain can be devised, which capture solutions to problems when designing and formalizing medical guidelines. This approach, of reusing formalized knowledge that appears frequently in different scenarios, has been proposed in the Protocure-I project and is pursued within Protocure-II project ([**?**]). Based on the essential concepts captured by these patterns, one can build a pattern language for the medical domain.

To become usable, these patterns have to be grouped according to the particular domains and tasks they cover, ranging from coverage of guideline development phases, coverage of healthcare process (and its quality), coverage of organizational aspects, modularization, complexity, decision-support, understandability of the text.

Part II
Structural Patterns

5 Structural Patterns

5.1 Characterization

Structural patterns are syntactic constructs that occur frequently in the text of the guideline. They include document structure patterns, for instance, the ”pattern” that each chapter contains the following sections: **Scientific Background**, **Conclusions**, **Recommendations** and **Other Considerations**; or the fact that an introductory sentence says what is the aim of the actions described in the current chapter; or the existence of an appendix section that lists references to literature and describes the terminology used in the guideline.

So far, the structural patterns did not play a significant enough role in authoring and formalization of clinical guidelines. Different organizations use different formats for their guidelines. However, it is very likely that in the future the structural patterns can play a significant role in maintenance of clinical guidelines, by facilitating replacement of a small section having a well defined subject, with an ”improved” section about the same subject.

5.2 Extraction

Structural patterns are obtained directly from the text of the guideline. It is possible that the guideline document is structured according to the hierarchical structure of the medical processes involved, or to the ordering of these processes.

When the guideline authors propose a particular structure of the guideline document, this facilitates a more effective maintenance of knowledge by the knowledge engineers that have to update the (living) guideline. When the structure of the document has a well-defined mapping with the medical knowledge, easier updates of medical knowledge in the document are possible, while leaving other sections unchanged.

Below we give an example of a guideline document structure:

5.3 Organization

In figure 4 the main narrative structures present in a guideline are shown. Each of these elements corresponds to a different formalization path. For instance, quantitative background information typically cannot be included in the ASBRU formalization other than as a range of values for a laboratory parameter.

Some examples are given for the statements that can occur in some of the guideline sections. These statements correspond to linguistic, background and control patterns described in the other parts of this document.

```
Introduction
   Title; Author;
   Background + History of disease;
   Aim of guideline; Guideline users;
   Working group + structure; Working method;
   Scientific proof; Legal binding; Update policy
Procedures
   Diagnosis; Treatment; Patient supervision and Handling of critical
situations
        Background
           Concept definitions (drug terminology, medical procedures)
           Range of accepted/unacceptable values for one parameter
           Medical actions: sequential, concurrent, cyclical, iterative
           Decisions, preferences, good practice
        Conclusions
           [Value-range] for [parameter] is considered normal.
        Recommendations
           In [context] it is recommended to do [action].
           [Action1] should not be combined with [action2].
           In case of [abnormal values] for [parameter], do [action].
        Considerations
           Effectiveness of [medication] for treatment [action] is
           [success-percentage].
Appendix
   References;
   Terminology used;
   Parameter values and ranges;
   Effectiveness medication/medical intervention;
   Summary procedures/conclusions
```

Figure 3: *Example of guideline document structure*

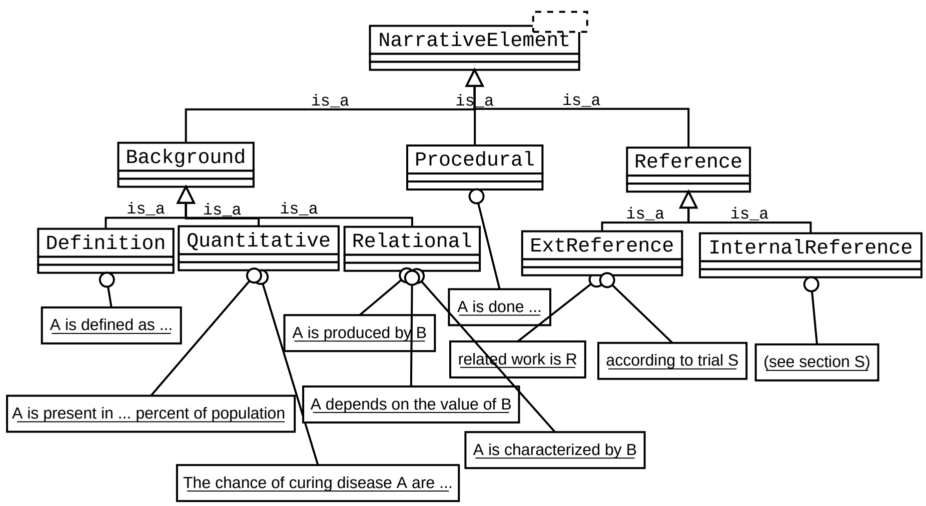

Figure 4: *Relations between the narrative elements present in the textual representation of a guideline and examples of instances*

5.4 Usage

In the first step of guideline formalization, the guideline text is divided into narrative structures, that are subsequently decomposed into control and background structures.

The structural patterns place some constraints on these narrative structures. For instance, that all text fragments that describe relevant medical scenarios have obey a fixed hierarchical structure.

Based on such constraints it can be decided what further transformations are needed to obtain an executable/formal model of the guideline text.

5.5 Conclusions and lessons learned

Narrative structures are rather close to the MHB constructs that correspond to the background and data dimensions.

They are used in the guideline formalization process as an intermediate step between text representation and control and background representations.

The mapping between structural patterns and ASBRU can be done only at a very high level of abstraction, with rules such as:

1. each chapter of a guideline will produce one plan group;
2. reference to common actions makes two plan groups connected;
3. if a text fragment TF_1 is part of a text fragment TF_2, then the resulting set of plans SP_1 used to represent TF_1 will be a part of the set of plans SP_2 used to represent SP_2.

Part III
Linguistic Patterns

6 Linguistic Patterns

6.1 Characterization

Linguistic patterns represent frequently-used language constructs that capture background, control or structure information. Background and control patterns frequently contain a standard linguistic representation.

The linguistic patterns identified by us correspond to a refinement of at least one category of patterns:

Control Patterns contain programming-like control structures. Due to the fact that a clear mapping can be defined between the general control structures captured by patterns and the more specific control patterns employed by an executable language (such as, ASBRU), an executable specification can be associated with most of the control pattern templates identified by us. However, it is rather difficult to identify them in the guideline text without a semantic annotation and without performing at least one abstraction step from the original text towards the intermediate representation.

Background Patterns describe relations between medical contexts, definition of parameters, causal linking between statements, etc, which summarize the background knowledge employed by the guideline to make recommendations and draw conclusions. They are more difficult to capture in an executable form than the control patterns, partly due to their non-operational nature.

Combined (Instantiated) Patterns contain more complex instantiations and combinations of control and background patterns.

In this report, linguistic, background and control patterns are treated in different sections, nonetheless it should be clear that few ”pure” linguistic patterns exist, that do not have a corresponding element in the list of control patterns and/or in the list of background patterns. All linguistic patterns capture fragments from the background component (associated with the background patterns), the control component (associated with the control patterns), or from both categories. The labels of patterns presented in section C indicate which categories of patterns a linguistic pattern belongs. The IR component of a control pattern is a linguistic representation of it.

The remainder of this section is organized as follows:

In section 6.1.1 we elaborate the motivation for searching linguistic patterns using an ontology with knowledge from the medical domain. Section 6.1.2 summarizes the research questions we are set to answer. Briefly, these represent a refinement of the objectives of this report, namely identifying re-usable linguistic components in the text of the guideline that are used frequently in the process of guideline formalization.

In section 6.2.1 we explain how the pattern templates found using background knowledge are chosen and report our results obtained so far. Section 6.6 makes concluding remarks and suggests future research directions for our work.

Part A of the appendix provides more details on what medical background knowledge has to be present in an ontology, to support finding of linguistic patterns. Section E contains a list of the most frequently encountered pattern templates and some instantiations of these templates in the text of the guideline.

6.1.1 Motivation for a linguistic pattern-driven approach to medical guideline formalization

The search for patterns as re-usable components of knowledge and control are motivated by the observation that particular constructs in the text of the guideline have a modular structure that is suitable for automatic translation into a more formal representation.

For instance, intentional fragments that occur frequently in the guideline as recommendations have a rather modular structure that can be extended or reduced to accomodate more complex or less complex recommendations. Examples:

```
Recommendation 1. In the event of isolated local recurrent breast
cancer following GRM and/or isolated regional recurrence following GRM
or MST, in previously not irradiated areas, the treatment of choice
consists of high-dose radiotherapy, if possible preceded by surgical
removal of the tumour.

Recommendation 2. In the event of isolated local recurrence following
MST, salvage mastectomy is recommended.

Recommendation 3. In the event of local recurrent breast cancer in a
previously irradiated area, the treatment of choice is low-dose
re-radiation with hyperthermia.

Recommendation 4. There are indications that in the event of local
recurrence in an irradiated area cytoreductive surgery preceded by
hyperthermia with re-radiation achieves better local control.
```

If constructs such as these:
$In\ the\ event\ of\ [MedContext],\ the\ treatment\ of\ choice\ is\ [Treatment].$ or
$In\ the\ event\ of\ [MedContext],\ [Treatment]\ is\ recommended.$
turn out to be often encountered in practice, their automatic translation into a more formal representation would greatly improve the effectiveness of translation, and subsequently the validation of the model. This so-called linguistic pattern templates would aid us in identification of inconsistencies, incomplete information, and other anomalies and would enable a more effective update of these modular parts.

Let us see how pattern templates can be extracted from guideline text based on an example from the 2002 **CBO guideline for treatment of mammacarcinoma** ([CBO2002]):

```
Recommendation (first recommendation in section 3.4):
Patients with locoregionally advanced breast cancer should receive
multidisciplinary treatment with curative intent.
```

Let us assume that we have the following ontology fragment of the medical domain:

```
inst_of("locoregionally advanced breast cancer",disease).
inst_of("multidisciplinary treatment",treatment).
inst_of("curative intent",med_goal).
inst_of(X,target_group) :-
     starts_with(X,"patients"),
     contains(X,Y),
     inst_of(Y,disease).
inst_of(X,recommendation_op) :-
     starts_with(X,"should").
inst_of(X,med_context) :-
     contains(X,Y),
     inst_of(Y,target_group).
inst_of(X,complex_treatment) :-
     contains(X,Y),
     inst_of(Y,treatment),
     contains(X,Z),
     inst_of(Z,med_goal).
```

In a real ontology of the medical domain there are many more such rules which describe how instances of medical concepts can be recognized based on their lexical and semantic content.

If we replace instances present in the recommendation text with their categories in the ontology, we obtain:

```
Recommendation:
Patients with [disease] should receive [treatment] with [medical_goal].
```

Then, if we apply the categorization rules in the ontology, the recommendation is rewritten as:

```
Recommendation:
[Target_group] [recommendation_op] receive [treatment] with [medical_goal].
```

and finally, if we ignore the linking words:

```
Recommendation:
[med_context] [recommendation_op] [complex_treatment].
```

Figure 5 depicts a compositional sketch of a generic recommendation, from a conceptual perspective. The recommendation contains:

1. an instance of a MedContext (”Patients with locoregionally advanced breast cancer” is a medical context), which is further refined as ”at least one of” (special association relation - recommendantion operator, in our terminology) an instance of TargetGroup (”Patients with ...”), possibly combined with a medical condition MedCondition (”locoregionally advanced breast cancer”);

2. an instance of RecOp (”should” is a modifier for recommendations);

3. an instance of MedAction (”receive multidisciplinary treatment with curative intent”), which further can be instantiated as a composition of a (optional) Treatment (”multidisciplinary treatment”), followed by (optionally) Medication (in our example, this element is not present), and MedGoal (”with curative intent”).

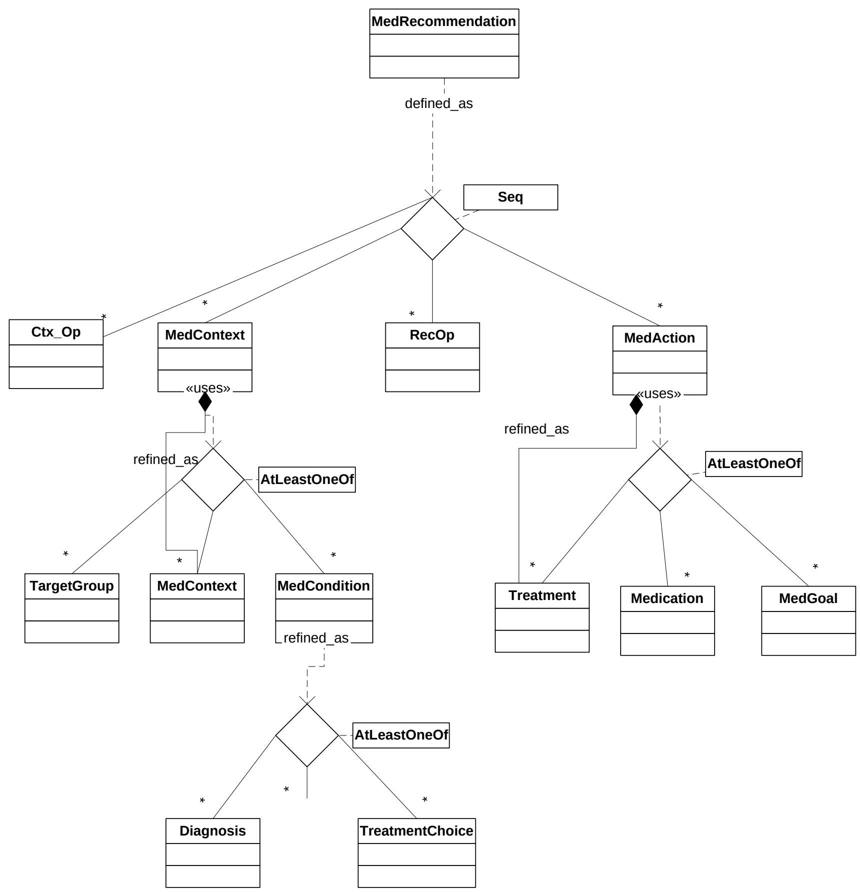

Figure 5: *A conceptual model of a recommendation*

The advantage of having such a conceptual sketch of the linguistic construct called ”medical recommendation” is that the template of any recommendation will include one of the following ordered lists of medical categories, obtained by traversing the tree:

1. (MedContext,RecOp,MedAction)

2. (TargetGroup,Diagnosis,RecOp,MedAction,MedGoal)

3. (TargetGroup,Diagnosis,RecOp,MedAction,Medication)

4. and so forth, all possible traversing paths through the graph starting with element ”MedRecommendation”.

The goal of finding linguistic pattern templates in the text amounts to finding the longest n-grams with elements belonging either to the medical categories (such as, $TargetGroup$, $Diagnosis$, $MedGoal$) or to syntactical categories (such as, Ctx_Op, $RecOp$). By having a list of syntactic terms that are most likely to occur in any guideline, a grammar for defining linguistic pattern templates can be inferred based on detected n-grams.

In section A we explain how such templates can be generated and instantiated automatically with the use of an ontology of the medical domain. It will become clear that producing meaningful linguistic pattern templates and translating them in a formal representation cannot be fully automated.

In section 6.2.1 we describe the method used by us to identify compositional linguistic patterns: having an ontology of the medical domain, a program can generate seqences containing combinations of medical terms that are connected using standard relations (such as, causal relationships, ordering and decomposition of actions, correlations action-intention, etc). Then instances of these pattern templates are sought in the guideline text. From the medically-relevant pattern templates generated we choose as basic pattern templates those which have sufficient support in more than one guideline text. After the set of basic pattern templates has been established, we analyze which combination of basic patterns can be instantiated based on common operators or categories they share. Those combination of basic pattern templates that have sufficient support in the guideline text are considered derived pattern templates.

6.1.2 Research questions

To pursue our goal, we need to answer the following questions:

1. Are there linguistic constructs in the text of a medical guideline that: a. correspond to a clear conceptual model; and b. are suitable to be reused because they occur frequently in specific locations of the guideline text and/or in more than one guideline? We will answer this question in section 6.1.1 by looking at oncology guidelines.

2. Do these frequent linguistic constructs correspond or can be associated with meaningful background knowledge from the medical domain? Can medical background knowledge be of help in automatically generating candidates for linguistic constructs? Or, given an existing pattern template, can new similar or extended pattern templates be derived using medical background knowledge? Do the templates generated in this way correspond to frequent pattern templates? This question is dealt with in section 6.2.1.

3. Can some of the frequent linguistic pattern templates be associated with a programming control structure (one of: sequencing, decomposition, logical combination, cyclical repetition, iteration) that has a well-defined formal equivalent? Do linguistic patterns have a control structure with a well-defined formal translation, which can be used to improve guideline formalization?

   Are there sufficient instances of such pattern templates (called linguistic-control pattern templates) in the text of a medical guideline that can be translated into programming structures in an algorithmic manner to consider this formalization support useful? The latter question is concerned with establishing the benefits of using linguistic patterns relative to the costs of identifying and maintaining them.

4. If we are to categorize the linguistic patterns in two more specific (not necessarily disjunctive) categories, namely linguistic-background patterns and linguistic-control patterns, which category of patterns is more likely to support us in the two main goals we are set to achieve by using patterns: a. building new guidelines; and b. using patterns in guideline formalization?

6.1.3 Examples of candidates for linguistic pattern instances

Linguistic background pattern instances from the **SIGN guidelines for breast cancer**:

1. There are three established surgical procedures for invasive breast cancers: (1) wide local excision (2) quadrantectomy or segmentectomy (3) mastectomy.

2. Mastectomy is indicated for operable breast cancer which is either large or at multiple sites, when radiotherapy is to be avoided, or by patient preference.

Linguistic background pattern instances from the **COIN oncology guidelines for breast cancer**:

1. Patients who appear on initial examination and investigation to have potentially curable disease should have a CT scan of thorax and upper abdomen with contrast enhancement if necessary. (Grade B)

2. Diagnostic imaging of bones, liver and brain should not be part of routine staging and should only be carried out if there is clinical or radiological suspicion of metastatic disease. (Grade B)

Linguistic background pattern instances from **NGC Guidelines for breast cancer**:

1. Consideration of axillary nodal sampling or partial axillary dissection should be given in these instances.

2. Axillary dissection is routinely performed for Stage II breast cancers for staging the disease and regional control of tumor.

6.2 Extraction

Linguistic patterns are abstracted from the IR representation of the guideline text using background knowledge, such as represented in figure 6 and 7.

There are two types of relations between medical concepts: ISA relations, which do not have labels, represent more specific concepts; and labeled relations, which describe a semantic connection between two medical concepts.

This type of background knowledge is needed in guideline analysis and formalization, as the authors of the guideline use these relations to produce the text of the guideline. Guideline formalization can benefit from reverse engineering the domain referred to by the guideline (when a model of this domain is not available), to improve the quality of the formal representation of the guideline.

In figure 7, an example of an unlabeled relation is explicited, together with two examples of instantiations that exist in the text, which connect medical terms with medical concepts.

6.2.0.1 Approach .

The paper makes (in some cases reiterates) the following statements:

1. Certain statements of the textual representation of a guideline fulfill the same conceptual role with respect to the categories of background knowledge used, and have a standard translation into a more formal guideline representation. For instance, a definition of the available medical actions and a list of constraints concerning their ordering is frequently given, regardless of the domain addressed by the clinical guideline. Similarly, the range of normal and abnormal values for specific laboratory parameters is defined, or medical factors influencing a medical decision for one treatment or another are mentioned either in the main body of the guideline, or in the appendix.

2. Identifying frequent textual constructs that can be translated in an algorithmic manner into their formal equivalent would greatly improve the effectiveness of the guideline formalization process. This is the main incentive for searching for linguistic patterns in medical guidelines.

3. The underlying model of many common textual constructs is supported by conceptual relations between the medical terms involved in the statement which the textual representation corresponds to. Therefore, using medical background knowledge in the form of 1. conceptual relationships between medical concepts present in the text (such as, *medical_action has_goal medical_goal*) and 2. simple is_a and instance_of relationships, one can generate potentially useful candidates for an easier translation of textual constructs into their formal equivalent.

4. The presence of programming structures in the guideline text is an indication of a possibly relevant piece of background knowledge that needs to be considered for the formal representation of the guideline.

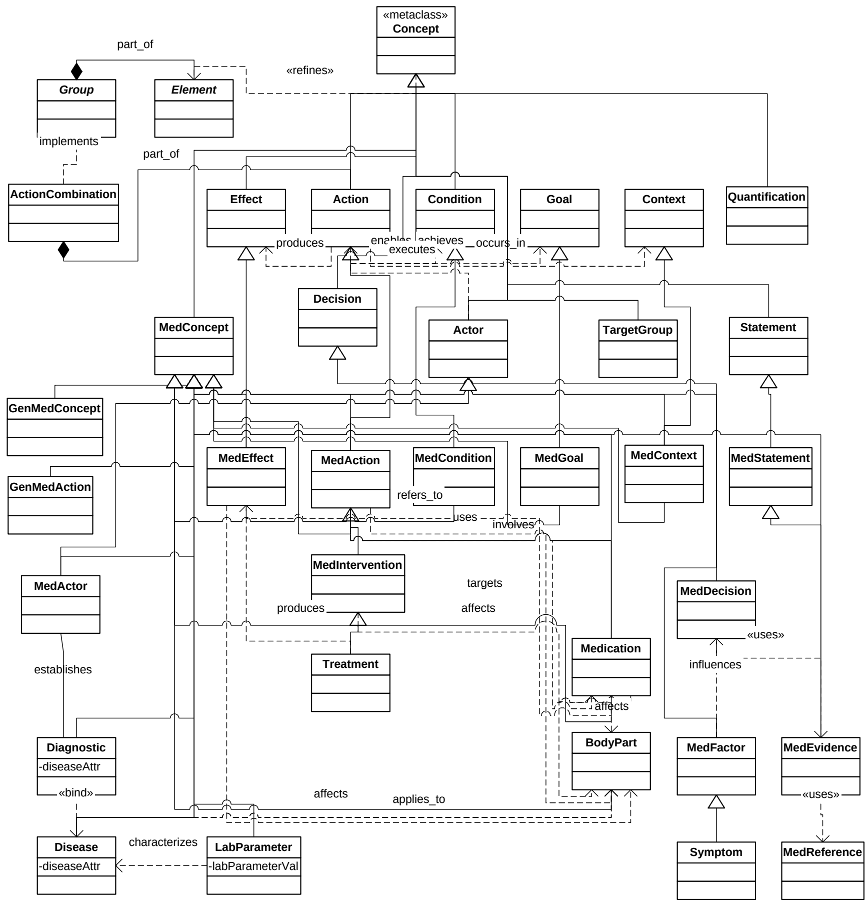

Figure 6: *Relations between the medical categories present in the medical background knowledge*

We propose a method of searching for linguistic (conceptual) patterns in the text of a guideline text, with the explicit aim of improving the effectiveness of the guideline formalization process and the quality of the resulting formal representation.

The method starts from language constructs that are frequently used in the medical domain (or a subdomain thereof), and automatically searches instances of these sentence templates in the text of the guideline. The most frequently used constructs are added to a list of pattern candidates which are then validated against other guidelines as well. For some of these constructs a formal representation is devised once, and each time an instance of the pattern is found, it can automatically be translated into its formal representation using the formal template prescribed by the pattern.

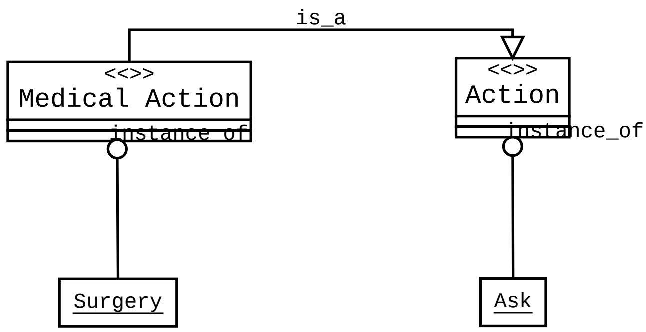

Figure 7: *Mappings between medical terms and medical categories in the background knowledge*

The pattern search process is heavily dependent on a medical ontology available. To search frequently used concepts and medically-relevant relationships between them we need an ontology containing the relevant medical terms that can be present in the guideline text and the medical categories they belong to. For instance, in order to classify pattern template $medical\_action\ sequence\ medical\_action$ as a frequent linguistic template, we need to be able to associate $mastectomy\ inst\_of\ medical\_action$, $radiotherapy\ inst\_of\ medical\_action$, $follows\ inst\_of\ sequence$ and to have sufficient coverage for $mastectomy\ follows\ radiotheraphy$ in the text of the guideline.

Using an ontology, possibly built or extracted automatically with the help of text mining and keyphrase extraction tools, has the advantage that it facilitates understanding of the aspects covered by the guideline, helps to identify inconsistencies and supports maintenance and reuse of the background knowledge covered by the guideline, thus enabling *living guidelines*.

A practical experiment in guideline formalization ([ten Teije; Anna Puig2004]) using the Stepper tool ([Ste]) sketched the basic requirements for the pattern search process, which has to consist of the following steps:

Algorithm 3 (Pattern Search Algorithm) *The steps applied by us for searching pattern instances in guideline texts:*

1. learn an ontology from a large set of guideline texts, by extracting keyphrases from these texts and assigning mappings of the medical terms into medical categories (MESH or UMLS categories).
2. use the ontology to semanticallly tag the guideline text:
   1. recognize medical terms of the ontology in the text of the guideline text analyzed;
   2. look up the medical categories of the medical terms recognized in the text, in the ontology of the medical domain;
   3. annotate medical terms in the guideline text with their corresponding medical categories (from the ontology);
3. search for meaningful low-level pattern templates, such as:

- laboratory parameters that are assigned values
- temporal ordering of actions, action decomposition
- associations between actions and goals to be achieved
- enumerations of medical factors that influence a clinical decision

4. adjust the level of abstraction and complexity of the pattern templates sought, by:
   - adding new relations that occur frequently in clinically-relevant pattern instances (validated by the medical expert)
   - adjusting the level of abstraction and complexity of the patterns sought
   - adding more complex pattern template candidates, containing: abstractions of the concepts used in existing pattern instances; or additional constraints for defining a new template based on two instances of some simpler pattern templates (for instance, the fact that two or more pattern instances detected in the same sentence which contain a common word actually represent parts of a larger linguistic pattern)
5. finding combinations of terms shared by several pattern instances, words that belong to different medical categories, etc, and trying: i. to invalidate some of the pattern candidates found; ii. to build more complex pattern, whose instances are then sought in other texts.
6. finally, collecting all pattern templates that have instances in more guideline texts, in a pattern library, and selecting the meaningful ones as building blocks for future guidelines.

In this report we investigate two special classes of pattern templates:
1. templates that describe medical background knowledge; and
2. templates that describe control knowledge.

Both classes of templates, particularly the templates for background knowledge, can be generated (or, alternatively, the instances can be validated) using an ontology for the medical domain.

6.2.1 Searching guideline patterns

Sentences in the guideline text often contain infrequent medical terms or other information important for the formalization process, such as, ordering of actions which were previously mentioned in the guideline text.

To minimize the information loss during formalization and to allow such sentences to be included among the ”relevant” terms, we defined a mini-ontology, containing categorized terms from the medical domain, and relational operators, corresponding to relations between actions, recommendations, factors influencing a decision, etc.

For all terms extracted, using TextToOnto, from the three input guidelines (CBO, RCR, SIGN), we generated a set of facts $is\_relevant\_term(<Word>)$:

```
is_relevant_term(ablation).
is_relevant_term(access).
is_relevant_term(adjuvant_chemotherapy).
...
```

These facts are used to filter out uninteresting sentences of the test guidelines, i.e. sentences that do not cover any relevant (medical or non-medical) terms will not be analyzed for patterns.

Below we give a few examples of facts from the ontology listed in appendix B.1. It contains terms categorized according to our own extension of the NGC semantic categories ([NGC]). The categories and the semantic relations between them is depicted in figure 8 and in figure 9.

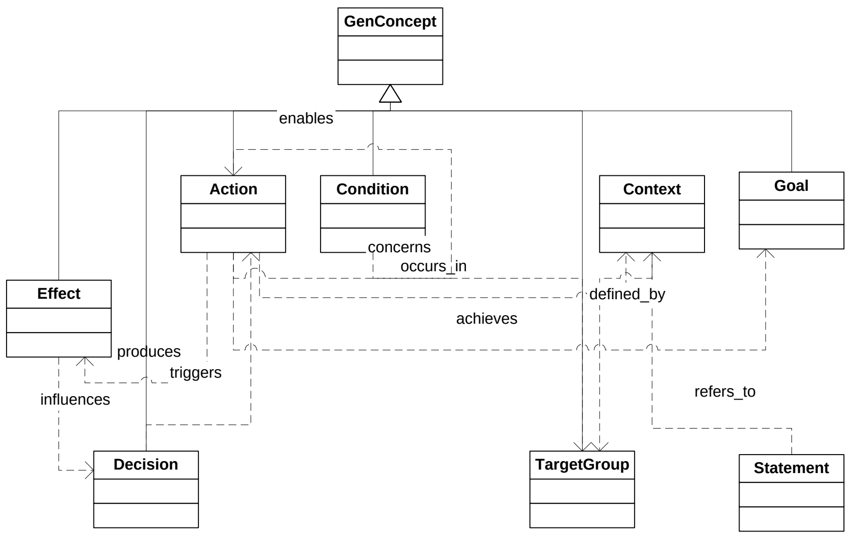

Figure 8: *Conceptual model of the generic domain, including the generic categories and their relations*

1. **medical specific categories**:
   - NGC categories: lab_parameter, symptom, disease, medication, body_part, med_effect, med_action;
   - custom medical categories: med_intervention (ISA med_action), med_condition, med_context, med_factor, med_concept, med_reference
2. **general categories**:
   - basic categories: action, condition, quantification
   - composed categories: recommendation, evidence, conclusion, pico_requirement
3. **operator categories**:
   - intra-phrase relational operators: generic_op, rel_op (assoc_rel_op, temp_rel_op, causal_rel_op)

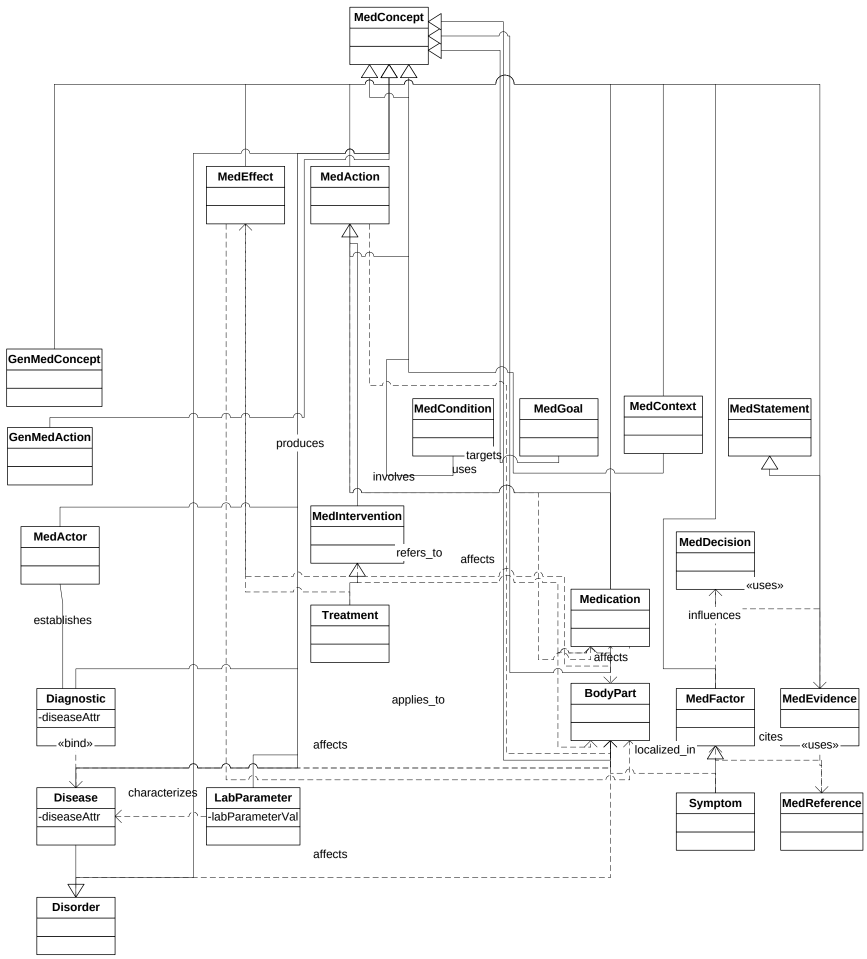

Figure 9: *Conceptual model of the medical domain, including the medical categories and their relations*

- action operators: decomp_op, act_op (seq_op, parallel_op)
- sentence operators: sentence_op

The relations between semantic categories which can occur in patterns are called operators, and represented in figure 10.

Having this $inst\_of$ relationships between terms and medical categories, we are able to process a sentence such as:

`Conclusions`

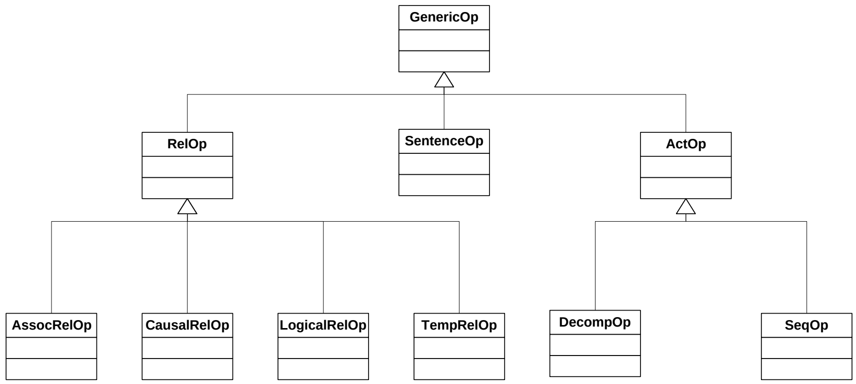

Figure 10: *The taxonomy of operators*

```
inst_of(med_intervention,axillary_surgery).
inst_of(med_intervention,radiotherapy).
inst_of(med_intervention,mastectomy).

inst_of(med_action,boost_dose).

inst_of(act_op,followed_by).

inst_of(body_part,breast).

inst_of(disease,breast_cancer).

inst_of(lab_parameter,margin_of_excision).

inst_of(decomp_op,consists_of).

inst_of(rel_op,which_is).

inst_of(sentence_op,eos).

% background knowledge
is_a(med_intervention,concept).
is_a(act_op,concept).
is_a(decomp_op,concept).

is_a(med_intervention,med_action).
is_a(med_action,action).
```

Figure 11: *Background knowledge used to formalize the example conclusion*

```
Grade 1:
For resectable locally advanced breast cancer, radiotherapy following
chemotherapy and surgery can reduce the risk of locoregional
recurrence by a factor of three, and can thereby improve the
disease-related survival.

A1 EBCTCG:58
```

```
A2 Overgaard:55,57 Ragaz:56
```

and extract the following pattern instances: $radiotherapy\ following\ chemotherapy$ and $radiotherapy\ following\ surgery$ which both instantiate pattern template $med\_action\ seq\_act\_op\ med\_action$.

This is possible because the medical background knowledge contains the following facts:

```
inst_of(radiotherapy,med_intervention).
inst_of(chemotherapy,med_intervention).
inst_of(surgery,med_intervention).
inst_of(following,seq_act_op).
inst_of(and,and_logical_rel_op).
is_a(med_intervention,med_action).
is_a(med_action,action).
is_a(seq_act_op,operator).
is_a(and_logical_rel_op,logical_rel_op).
is_a(logical_rel_op,rel_op).
is_a(rel_op,operator).
```

Note that seq_act_op stands for 'sequential action operator', and med_action, $med\_intervention$ stand for 'medical action' and 'medical intervention', respectively.

On the other hand, the background knowledge for the formalization process indicates that actions can be mapped to ASBRU plans and sequences of actions can be mapped into ASBRU plans with a sequential plan body. The following is an excerpt of Prolog code that performs this translation:

```
matches_template(InstanceLabel,Template,Instance) :-
   label(InstanceLabel),
   (text_fragment(InstanceLabel,Instance),
    inst_of(Instance,Template);
    Template=[MedCategory|MedCategories],
    Instance=[Term|Terms],
    matches_template(InstanceLabel,Term,MedCategory),
    matches_template(InstanceLabel,MedCategories,Terms)).

mapping(Action,Plan) :-
   inst_of(Action,action),
   inst_of(Plan,asbru_plan),
   has(Plan,[(plan-name,Action),(plan-body,to_be_defined)].

mapping(SeqActions,Plan) :-
   SeqActions=seq_act_op(Action1,Action2),
   matches_template(SeqActions,[action,seq_act_op,action]),
   mapping(Action1,Plan1),
   mapping(Action2,Plan2),
   inst_of(Plan,asbru_plan),
   has(Plan,[(plan-name,Action1-seq-Action2),
             (plan-body,sequential([plan-activation(Plan1),
                                    plan_activation(Plan2)]))]).

mapping(AndActions,Plan) :-
   AndActions=and_act_op(Action1,Action2),
   matches_template(AndActions,[action,and_logical_rel_op,action]),
   mapping(Action1,Plan1),
   mapping(Action2,Plan2),
   inst_of(Plan,asbru_plan),
   has(Plan,[(plan-name,Action1-seq-Action2),
             (plan-body,sequential([plan-activation(Plan1),
                                    plan_activation(Plan2)]))]).
```

```
translation(IR,FR) :-
   annotate(IR,AnnotatedIR),
   forall(
    contains(AnnotatedIR,text_fragment(InstLabel,TextInstance)),
    (
      matches_template(InstLabel,PattTemplate,TextInstance),
      mapping(TextInstance,FormalForm),
      replace(AnnotatedIR,TextInstance,FormalForm,NewAnnotatedIR),
      (
       NewAnnotatedIR=AnnotatedIR,FR=NewAnnotatedIR,!;
       translation(NewAnnotatedIR,FR)
      ))).
```

Using the pattern language we can express the mapping constraints and the possible result of translation:

```
text_fragment(Chapter3ConclusionsC1,"Grade 1:For resectable...").
contains(Chapter3ConclusionsC1,
        [text_fragment(PattInst1,"radiotherapy following chemotherapy"),
         text_fragment(PattInst2,"radiotherapy following surgery")]).
matches_template(PattInst1,[action,seq_act_op,action],
                           [radiotherapy,following,chemotherapy]).
matches_template(PattInst2,[action,seq_act_op,action],
                           [radiotherapy,following,surgery]).
translation(Chapter3ConclusionsC1,
        "Grade 1: For resectable locally advanced breast cancer,
        <reference:plan "radiotherapy-seq-chemotherapy and surgery">
        can reduce the risk of locoregional recurrence by a factor of
        three, and can thereby improve the disease-related survival.

         A1 EBCTCG:58
         A2 Overgaard:55,57 Ragaz:56
        ").
```

6.2.2 Implementation Details

The text of the guideline is pre-processed using a Java application that splits the text into sentences which contain word-level chunks. A chunk (concept proposed by Seyfang ([et al2004])) represents the maximal text fragment at which a pattern has been identified. Initially, the chunk is set at the word level, but following the search of patterns, it may grow towards combination of several words, sentences, paragraphs and sections (also depending of the level of granularity at which patterns are sought).

The document model used by the implementation is shown in figure 12.

We illustrate how a sentence is transformed into a list of chunks by our Java text pre-processor: the sentence ”Treatment1 consists of axillary surgery followed by radiotherapy.” is transformed into the following chunks:

```
[chunk(treatment1,[]),chunk(consists,[]),chunk(of,[]),chunk(axillary,[]),
chunk(surgery,[]),chunk(followed,[]),chunk(by,[]),
chunk(radiotherapy,[]),chunk(eos,[])].
```

A special chunk $chunk(eos, [])$ marks the end of a sentence. Each chunk consists of a tuple that includes the word assigned to that chunk and a list containing annotations for that chunk. Initially, all chunk annotations only contain the position of each word in the chunk in the guideline text), but after each processing of a text fragment (in

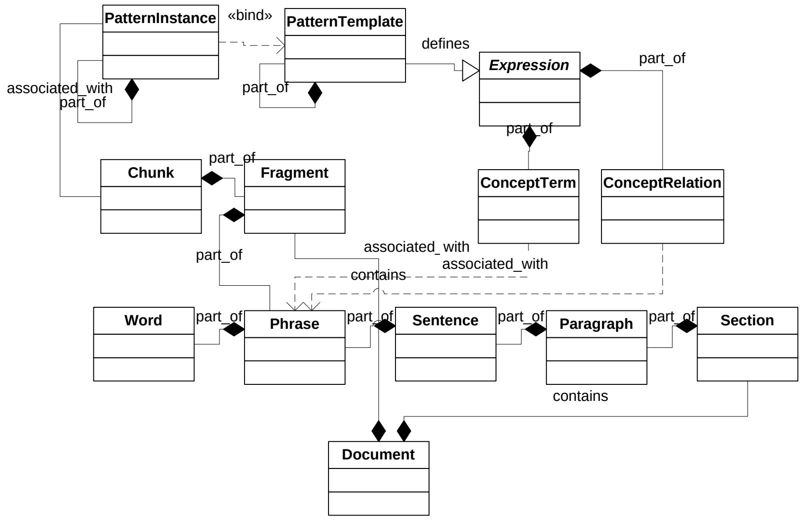

Figure 12: *The structural document model used in pattern search*

search for patterns) this list can be expanded as follows: when the word is found to be an instance of a medical term, its semantic category is added to its annotation; when a pattern is found, of which the chunk is a part, it is added as an annotation of that chunk, etc.

We define patterns at several levels of granularity:

1. at group-of-words level: amounts to semantically tagging of words in the guideline text
2. at sentence level: define concepts in different semantic category that correspond to (or can be refined into) formal constructs; for instance:
   $action_1\ consists\_of\_action_2, action\_3$
   contains a standard decomposition, with a well-defined formal translation (see appendix, section C).
3. at paragraph level: define sequences of sentences linked through a relevant word (containing the same concept, or referring to the same medical context).
4. at section level: define relations between paragraphs that contain concepts related through a particular relation.
5. at chapter level: define relations between sections that contain references to medical actions that are known (from background knowledge) to be related through a hierarchical decomposition or temporal ordering relation.

The analysis starts with chunks at word-level. At word-level, the annotation of each chunk contains its semantic category, position in the text, and patterns it is part of. When keyphrases from the medical domain are recognized, several chunks are merged into one, together with their annotations.

In the next step the analysis focuses on sentence-level chunks, which contain word-level chunks and are annotated with pattern instances found in that sentence. Within the limits of a sentence, several overlapping pattern instances can be found. Before proceeding with analysis at the next level, we have to decide which overlapping pattern instances in a sentence actully belong to a more complex pattern template.

Further, the guideline text is investigated for relations between pattern instances at paragraph-level, section-level, chapter-level patterns. However, it seems that patterns at levels higher than sentence level are not very likely to be found by this very simple pattern search.

6.2.3 Experimental results

We have performed several analyses on chapter 3 of the CBO guideline, with the goal of producing an executable (semi-formal) representation of the guideline text. The chapter has been selected as relevant for formalization because formalization of this chapter was previously available, and useful for validation of result.

The parameters of analysis were:

- the size of the ontology (which contains only is_a and $inst\_of$ relations);
- the set of pre-defined pattern templates to be searched in the text.

We have run a sentence analysis with a small pre-defined sequence of pattern templates to be sought within the guideline text. Then we updated the list of patterns sought. Subsequently, we improved the program by testing all possible combinations of 2, 3 and 4 medical terms and making a list of the most encountered combinations.

Then, for all patterns identified, we generated:

1. a MHB model containing the chunks identified as pattern instances
2. an executable (ASBRU) set of plans corresponding to the pattern instances

The following pattern templates have been used (where CategList is the list of semantic categories that were sought):

The result of analysis of chapter 3 in the CBO guideline, which contains 179 ”sentences”, at the sentence level, produces output of the following form: A condensed summary of the automated sentence analysis is available in section B.2 of the appendix.

By processing the pattern instances resulted from this analysis of the guideline text at sentence level, more complex patterns can be found in the following ways:

1. when two pattern instances share one or more words, it is possible that they cover either the same knowledge template, or a complimentary one; therefore their combination is likely to form a more complex pattern;

```
p_act_rel: [action,temp_rel_op,action]; [action,act_op,action]
   % relations between actions
   % [action,act_op,action] covers 'do action1 followed_by|preceded_by action2'
p_act_decomp: [action,decomp_op,action,action,action];
   % decomposition of actions
   % [action,decomp_op,action] covers 'action1 consisting_of action2'
p_spl_int_act: [intention,assoc_rel_op,action,assoc_rel_op];
p_complex_int_act:
   [intention,assoc_rel_op,action,assoc_rel_op,disease,assoc_rel_op,
              action,assoc_rel_op,disease_attr];
p_cond_act: [med_condition,med_action]; % if-then
p_rec: [med_context,recommendation,med_action]; % recommendation
p_lab_val: [lab_parameter,assoc_rel_op,lab_parameter_val];
p_qresult: ([action,quantification,result];
           [action,quantification,assoc_rel_op,result]);
           % quantification of result
p_disease_attr: [disease_attr,assoc_rel_op,disease];
p_labpar: [lab_parameter,assoc_rel_op,diagnostic];
p_body:(
        [lab_parameter,assoc_rel_op,body_part,assoc_rel_op,lab_parameter_val];
        [med_action,assoc_rel_op,body_part]);
        % [lab_parameter,assoc_rel_op,body_part] covers
        %'tumour size is|can_be large'
        % [med_action,assoc_rel_op,body_part] covers
        % 'surgery is_applied_to axilla'
p_causal: [causal_rel_op,med_condition,action];
        % 'in the event of ... do ...'

p_act_composition4: [treatment,decomp_op,med_intervention,act_op,med_intervention];
 'treatment consists of neoadjuvant_chemotherapy followed_by surgery '
p_excl_treatment: [med_intervention,excl_rel_op,assoc_rel_op,med_action];
 'treated solely with radiotherapy'
p_factors4treatment: [evidence,med_factor,assoc_rel_op,med_action];
 'there is no age limit for neoadjuvant chemotherapy'
p_rec2: [recommendation,med_action,assoc_rel_op,target_group];
 'some authors recommend just chemotherapy and radiotherapy for this group of patients'
p_rec3: [recommendation,assoc_rel_op,med_effect];
 'attention should be paid to side effects '
p_causal_act: [causal_rel_op,med_action,rel_op,med_action];
'therefore recommend just chemotherapy and radiotherapy'
p_neg_evidence: [evidence,recommendation,med_intervention];
 'insufficient evidence to recommend ...(mastectomy)'
p_symptom_act: [symptom,temp_rel_op,med_action];
'local recurrence following MST'
p_impossibility1: [neg_rel_op,action,rel_op,med_intervention];
'no chemotheraphy after radiotherapy'
p_evid_act_result: [evidence,assoc_rel_op,result_op,med_action,disease];
'no evidence that effective chemoterapy on disease'
```

2. when each of two pattern instances contain at least one word that belongs to the same (medical) category, it is possible that they both represent an extension of a simpler pattern template (one that includes the shared category)

For instance, the sentence:

”Treatment1 consists of axillary surgery followed by radiotherapy.”

contains the word $axillary\_surgery$ as part of two patterns:

```
===
Relevant sentence S(12) has 6 medical terms:
[therapy, surgery, treatment, improve, locoregional, control]

Sentence(12): [neoadjuvant_systemic_therapy, opand, surgery, were,
also, added_to_the, treatment_options, to_improve, the, disappointing,
locoregional_control, achieved, eos]

1: [med_intervention, logical_rel_op, med_intervention, decomp_op,
treatment] (p_act_composition1)......
   most specific:[med_intervention,logical_rel_op, med_intervention,
decomp_op, prescription]......
   least specific:[concept, generic_op, med_concept, generic_op, concept]
    [neoadjuvant_systemic_therapy, opand, surgery, also, treatment_options]

2: [treatment, intention, lab_parameter] (p_intention_of_treatment)......
   most specific:[prescription, intention, lab_parameter]......
   least specific:[concept, concept, concept]
    [treatment_options, to_improve, locoregional_control]

===
Relevant sentence S(21) has 12 medical terms:
[advanced, breast, cancer, unresectable, criteria, skin, carcinoma,
chest, wall, lymph, nodes, supraclavicular]

Sentence(21): [definition, locoregionally_advanced_breast_cancer,
opis_used_to_describe, breast_cancer, which_opis, unresectable,
on_the_basis_of, the, classic, unresectability_criteria, dblpnt,
oedema, of, the, skin, lpar_peau, d_quote_orange_rpar, ulceration,
satellite, skin, nodules, inflammatory, carcinoma, infiltration, of,
the, chest, wall, lpar_t4_rpar, lymph, nodes, fixed, to, one, another,
and_bkslash_or, to, deeper, structures, lpar_n2_rpar, opor, palpable,
internal, mammary, parasternal, infraclavicular, and_bkslash_or,
supraclavicular, lymph, nodes, lpar_n3_rpar, eos]

1: [disease, assoc_rel_op, disease] (p_disease_def)......most
specific:[disease, assoc_rel_op, disease]......least
specific:[concept, generic_op, med_concept]
   [locoregionally_advanced_breast_cancer, opis_used_to_describe,
breast_cancer]

2: [disease, assoc_rel_op, disease_attr] (p_def1)......most
specific:[disease, assoc_rel_op, disease_attr]......least
specific:[concept, generic_op, med_concept]
    [locoregionally_advanced_breast_cancer, opis_used_to_describe,
unresectable]

3: [disease, assoc_rel_op, disease, assoc_rel_op, disease_attr]
(p_def2)......most specific:[disease, assoc_rel_op, disease,
assoc_rel_op, disease_attr]......least specific:[concept, generic_op,
med_concept, generic_op, med_concept]
    [locoregionally_advanced_breast_cancer, opis_used_to_describe,
breast_cancer, which_opis, unresectable]

4: [disease_attr, assoc_rel_op, med_factor] (p_factor1)......most
specific:[disease_attr, assoc_rel_op, med_factor]......least
specific:[med_concept, generic_op, med_concept]
    [unresectable, on_the_basis_of, unresectability_criteria]



pattern1: template = [action, act_op, action],
          instance = [axillary_surgery, followed_by, radiotherapy]
```

```
pattern2: template = [action, decomp_op, action],
          instance = [treatment1, consists_of, axillary_surgery]
```

therefore a new pattern candidate can be derived by combining the two pattern templates:

$$< action - 1 >< decomposition - operator >< action - 2 >< action - operator > < action - 3 >$$

and an instance of that is *"Treatment1 consists of axillary surgery followed by radiotherapy"*, i.e. a complex pattern instance that covers the entire sentence.

New pattern candidates can be found by interogating the list of all patterns. Not all of them are meaningful combinations of data. For instance, some of them do not have a linguistic knowledge pattern underlying them, and are therefore discarded. In our case, the combination:
$< action - 1 >< action - operator >< action - 2 >< decomposition - operator >< action - 3 >$
can be discarded, as "Axillary surgery followed by radiotherapy consists of treatment1" does not represent a linguistically valid instance.

By associating a formal representation with each pattern that was validated by a medical expert, the text can be annotated with a formal representation in considerably less time than done manually. The formal representation still has to be adjusted by a knowledge engineer that has to deal with special cases and exceptions in translation.

It becomes clear from the first analysis that temporal ordering relation represented by pattern template $p1(A : action, \{following\} : operator, B : action)$ is probably the most frequent linguistic construct that can be found using the pre-defined set of templates. For example, in chapter 3 of the CBO guideline, 5 out of 11 instances of patern template $< action - 1 >< action - operator >< action - 2 >$ contain the word $following$ as instantiation of act_op action operator. This means that p1 is a frequent linguistic pattern, as it has a high support among the pattern instances found in the text (as a matter of fact, p1 is not only found in chapter 3, but also in other chapters as well).

The same characterization can be made about the linguistic pattern
$p2(A : action, \{after\} : operator, B : action)$.

In figure 13 we have depicted the piece of background knowledge uncovered by the discovery of pattern instances

```
p1(axillary_surgery, following, excision))
p1(biopsy, following, excision))
p1(breast_reconstruction, following, mastectomy))
```

which instantiate pattern template $p1(action,\ act\_op,\ action)$. This is the essence of our search for patterns: by grouping together such pattern instances that share common words and their corresponding categories, we are able to decide which linguistic constructs are most frequent.

Section B.2 contains a complete list of the pattern instances identified within the text of chapter 3 of the CBO guideline. Out of 179 sentences, 100 contain more than 2

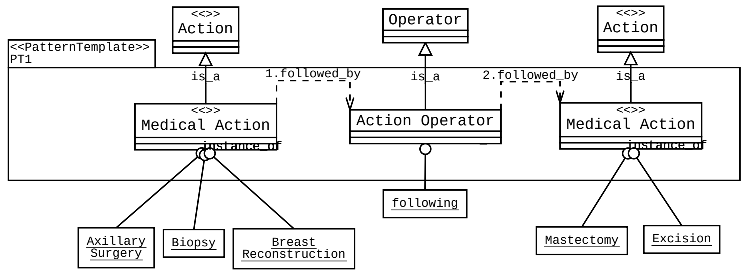

Figure 13: *Linguistic pattern fragment identified in guideline text*

medical terms, and therefore deemed relevant for analysis. Among these 100 sentences, 226 pattern instances were identified, a few of which are listed below:

```
1: p_act_decomp([action, decomp_op, action, action, action])
    [multidisciplinary_treatment, comprising, surgery, radiotherapy, chemotherapy]
12: p_act_rel([action, temp_rel_op, action])
    [radiotherapy, after, surgery]
13: p_act_rel([action, temp_rel_op, action])
    [radiotherapy, following, chemotherapy]
14: p_act_rel([action, temp_rel_op, action])
    [radiotherapy, following, neoadjuvant_chemotherapy]
15: p_act_rel([action, temp_rel_op, action])
    [radiotherapy, following, surgery]
16: p_act_rel([action, temp_rel_op, action])
    [support, after, treatment]
17: p_act_rel([action, temp_rel_op, action])
    [surgery, following, neoadjuvant_chemotherapy]
18: p_act_rel([action, temp_rel_op, action])
    [surgery, following, radiotherapy]
19: p_body([med_action, assoc_rel_op, body_part])
    [biopsy, of, axilla]
21: p_body([med_action, assoc_rel_op, body_part])
    [investigation, of, chest]
32: p_labpar([lab_parameter, assoc_rel_op, diagnostic])
    [tumour, with, clinical_invasion]
33: p_qresult([action, quantification, result])
    [multidisciplinary_treatment, better, survival]
53: p_temp_order([action, temp_rel_op, action])
    [surgery, following, neoadjuvant_chemotherapy]
54: p_temp_order([action, temp_rel_op, action])
    [surgery, following, radiotherapy]
```

An example of a relevant sentence that contained several pattern instances:

```
Sentence(58): [nr28, two, trials, investigated, the, addition_of,
hormonal, treatment, to, a, combination_of, chemotherapy,
radiotherapy, opand, surgery, eos]

1: p_act_decomp([action, decomp_op, action, action, action])
    [treatment, combination_of, chemotherapy, radiotherapy, surgery]
```

It becomes clear, after running the automated analysis, that not all pattern instances found represent meaningful combinations. Some of them overlap, others share only one word.

```
1. Pattern [cond_op, disorder] has 2 instances:

 ---> 1.1:   Patt p_ctx4([cond_op(in_the_event_of),
disorder(problems)]) instance of [cond_op, disorder] occurs in [in,
the, event, of, interpretation, problems, additional, diagnostic,
procedures, with, the, aid, of, blank, ct_examination, opis, usually,
sufficient, but, mri, opis, more, sensitive, eos]

 ---> 1.2:   Patt p_ctx4([cond_op(in_the_event_of),
disorder(problems)]) instance of [cond_op, disorder] occurs in [nr39,
in, the, event, of, cranial, nerve, problems, as, a, result, of,
metastases, of, the, base, of, the, skull, urgent, irradiation, opis,
indicated, in, order, to, prevent, any, further, irreversibility, eos]
```

The list of pattern templates that have the highest number of instances is also produced as a result of the analysis:

```
template [target_group, assoc_rel_op, disease] has 37 instances
template [treatment, assoc_rel_op, diagnostic] has 29 instances
template [treatment, assoc_rel_op, disease] has 24 instances
template [treatment, assoc_rel_op, med_condition] has 29 instances
template [treatment, assoc_rel_op, med_context] has 16 instances
template [treatment, assoc_rel_op, target_group] has 11 instances
template [treatment, generic_op, diagnostic] has 32 instances
template [med_action, prescription] has 15 instances
template [med_intervention, prescription] has 16 instances
template [med_action, assoc_rel_op, disease] has 53 instances
template [med_action, rel_op, body_part] has 11 instances
template [med_intervention, logical_rel_op, med_intervention] has 13 instances
template [med_intervention, rel_op, med_intervention] has 47 instances
template [treatment, assoc_rel_op, med_action] has 23 instances
template [med_context, treatment] has 21 instances
template [target_group, recommendation, med_action] has 4 instances
template [treatment, generic_op, prescription] has 13 instances
template [med_context, med_condition] has 30 instances
template [target_group, assoc_rel_op, lab_parameter] has 5 instances
template [med_intervention, diagnostic] has 12 instances
template [med_action, assoc_rel_op, body_part] has 14 instances
template [med_effect, assoc_rel_op, med_condition] has 11 instances
template [decomp_op, med_intervention, assoc_rel_op, med_intervention] has 2 instances
template [lab_parameter, assoc_rel_op, diagnostic] has 5 instances
template [result_op, assoc_rel_op, disease] has 4 instances
template [med_action, assoc_rel_op, disease, rel_op, med_intervention] has 2 instances
template [target_group, assoc_rel_op, symptom] has 3 instances
```

Figure 14 depicts this result in a more compact way. The list of most frequently encountered linguistic pattern templates can be found in section E of the appendix.

In order to expand the set of pattern templates sought or to search for new pattern templates, we need to establish which words are relevant and occur most frequently in the pattern instances found.

Therefore we built a summary of the all words occuring in the pattern instances found, and a list of words occuring in more than one pattern instance. We found a total of 105 medical terms, non-specific to the domain of breast cancer:

```
adding, addition_of, advanced, after, age_limit, alone, also, any,
assessment, attention_should_be_paid, before, benefit, best, better,
by_means_of, cases_of_patients, combination_of, combining,
complications, consists_of, current_treatment,
disease_related_survival, examination, factor_of, followed_by,
```

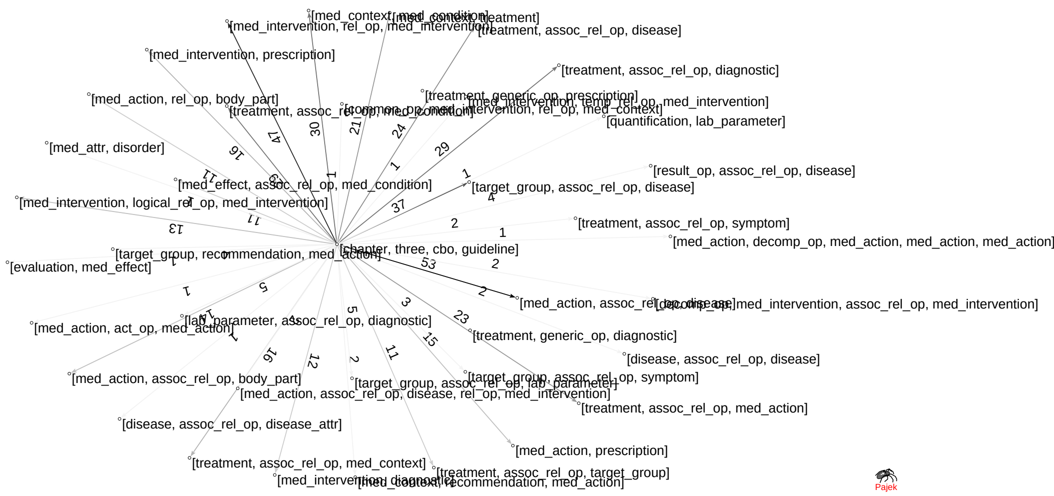

Figure 14: *Linguistic pattern templates identified in chapter 3 of the CBO guideline*

```
following, general_practitioner, group_of_patients, high_risk, higher,
if, in, in_the_case_of, in_the_event_of, in_this, increases_the,
indications, initially, inoperable_patients, lack_of,
local_recurrence, local_treatment, locoregional, locoregional_control,
may_opnot_be_possible, months, multidisciplinary_treatment, negative,
of, on_the_basis_of, oncology_patients, only, opand, operable, opfor,
opis, opis_used_to_describe, opor, pain, patient, patients, problems,
recommend, recurrence, regimens, resection, residual_tumour, response,
response_rate, risk, risk_of, screening, severe, should, should_be,
side_effects, since, skin_reactions, solely, stage_of_disease, strong,
systemic_treatment, that, therapy, there_opis_no,
this_group_of_patients, to, to_improve, to_reduce, treated, treatment,
treatment_modalities, treatment_of_choice, treatment_options, trial,
tumour, tumour_characteristics, tumour_load, unresectability_criteria,
unresectable, usually, versus, when, which_opis, with, without
```

It turns out that the words that occur in patterns are actually instances of only 18 medical categories. The mappings of these words to categories is given in figure 15. For each medical term MT present in the pattern instances identified in the text, which is known to belong to medical category MC, an arrow $MC \rightarrow MT$ is present in the graph. The more often a medical term occurs in the identified patterns, the more emphasized the arrow that targets that node in the graph becomes.

If we generate a graph from the sequences discovered as instances of pattern templates, such as the one shown in figure 16, the nature of linguistic pattern instances

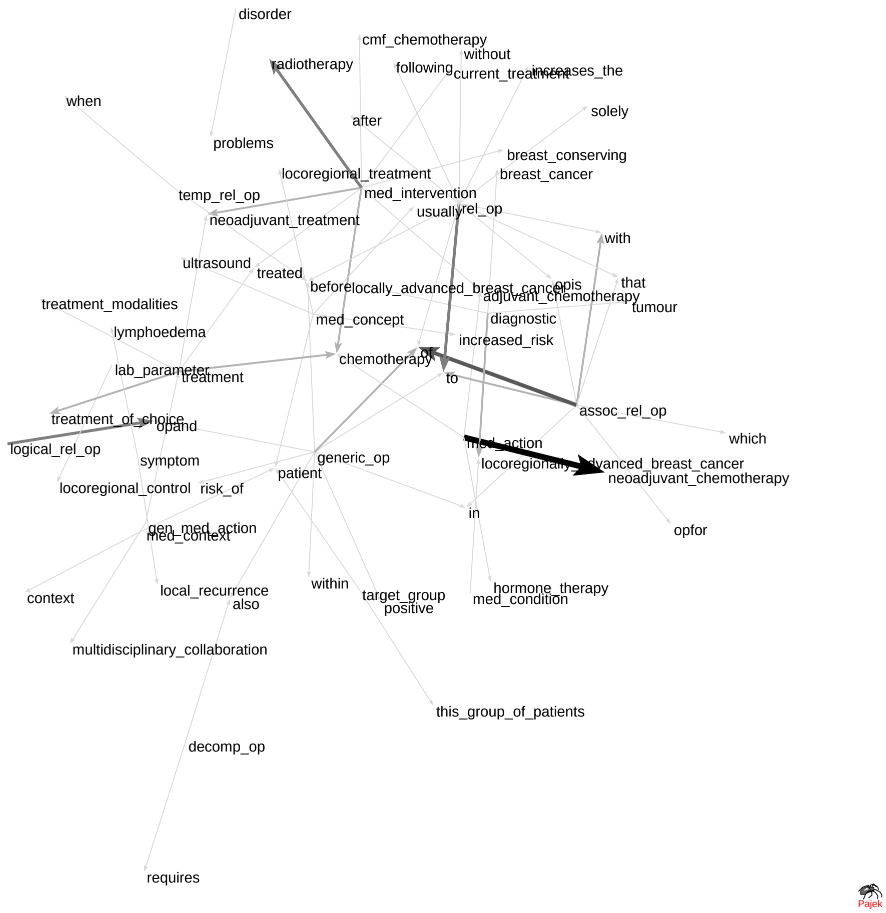

Figure 15: *Mappings between medical categories and words present in pattern instances identified in chapter 3 of the CBO guideline*

discovered at sentence level becomes clearer: on the first position most often a medical intervention is encountered, followed by action sequencing, decomposition, or logical operators, then on the last position either a medical goal, a medical action, reference to a disorder or a more generic medical concept. This is due to the fact that the search for pattern instances took a pre-defined list of pattern templates, mostly containing medical actions.

From the list of most frequently used words that occur in patterns, but do not belong to the medical domain, we have found: of, with, after, following. They are relational operators that denote a generic association between objects, or ordering of actions,

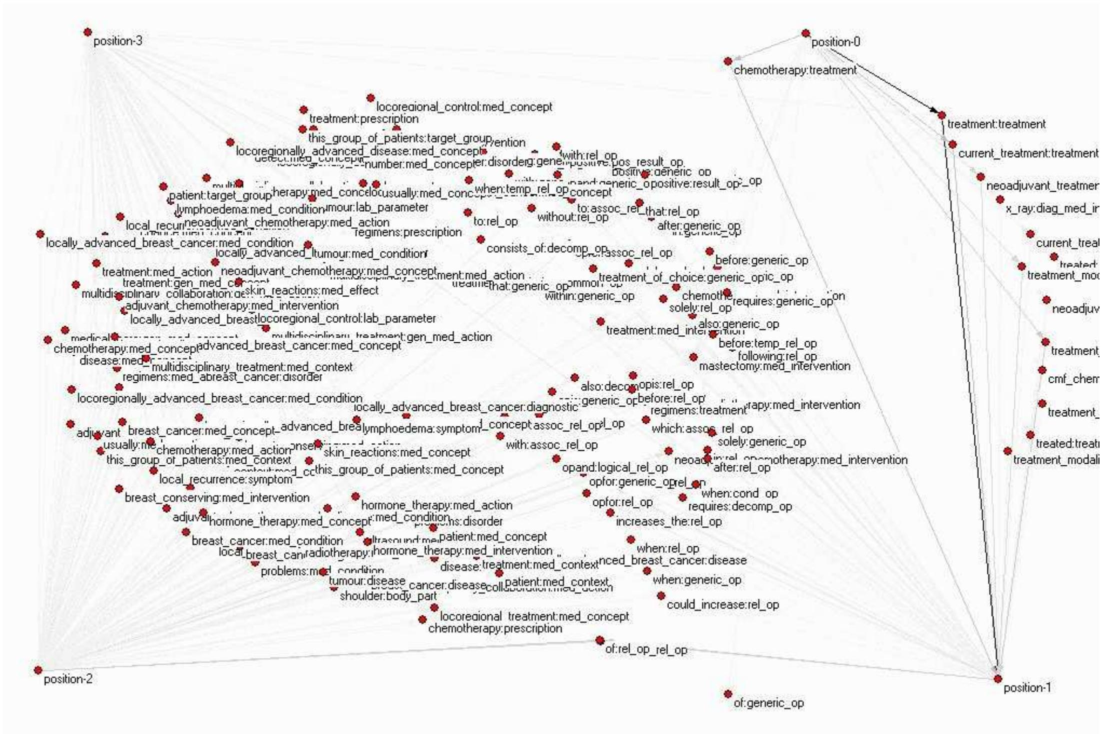

Figure 16: *Sequences of medical terms and medical categories present in the pattern instances identified in chapter 3 of the CBO guideline*

respectively. In an ontology they would correspond to terms instantiating relations between medical concepts.

Particularly for the words "after" and "following", we noticed that the same phrase was discovered as a pattern instance of two distinct pattern templates:
$p\_act\_rel = [action, act\_op, action]$, which describes relations between actions, and
$p\_temp\_rel\_op = [action, temp\_rel\_op, action]$ - which describes temporal constraints between actions.

This anomaly is explained by a modelling mistake, namely the fact that we assigned the terms "after" and "following" to two classes: act_op and $temp\_rel\_op$. We can decide that this is not a serious problem in searching for patterns, and leave them as such, or, alternatively, to create a new pattern category called $p\_temp\_act\_rel\_op$, which looks for occurences of words such as "after" and "following" only in the context of actions. The decision to change the ontology (if we decide to transfer the terms "after" and "following" to a new subcategory of $temp\_act\_rel\_op$ and remove them from their previous categories) changes the pattern instances that are recognized in the text.

This process can be avoided if a well-defined ontology of the medical domain exists, in which the categories of the terms corresponding to the most used relations are pre-defined and fixed. This will be the case in a living guideline, therefore the search for linguistic patterns will always produce the same list of simple pattern instances.

6.2.4 Validating the patterns againts other guidelines

After building a very simple ontology for the medical domain, with terms specific to breast cancer protocols, and developing a list of useful linguistic pattern templates, we have searched instances of these templates in the test set - the Royal College of Radiologists' Guideline for the Treatment of Breast Cancer.

Out of 2269 sentences in the text of RCR Guideline, only 1200 sentences contained one relevant term, i.e. present in the the list of terms to which categories could previously be assigned, using the simple ontology. Much fewer sentences, namely 314, contained 3 or more relevant terms and therefore the chance of containing a pattern.

However, only a small number of pattern templates were instantiated in the text of the test guideline.

The output of the program:

```
There were 314 relevant sentences processed
====
List of 521 patterns:
====
The most used pattern templates are:

template [med_action, assoc_rel_op, disease] has 53 instances
template [med_intervention, rel_op, med_intervention] has 47 instances
template [target_group, assoc_rel_op, disease] has 37 instances
template [treatment, generic_op, diagnostic] has 32 instances
template [med_context, med_condition] has 30 instances
template [treatment, assoc_rel_op, diagnostic] has 29 instances
template [treatment, assoc_rel_op, med_condition] has 29 instances
template [treatment, assoc_rel_op, disease] has 24 instances
template [treatment, assoc_rel_op, med_action] has 23 instances
template [med_context, treatment] has 21 instances
template [treatment, assoc_rel_op, med_context] has 16 instances
template [med_intervention, prescription] has 16 instances
template [med_action, prescription] has 15 instances
template [med_action, assoc_rel_op, body_part] has 14 instances
template [med_intervention, logical_rel_op, med_intervention] has 13 instances
template [treatment, generic_op, prescription] has 13 instances
template [med_intervention, diagnostic] has 12 instances
template [treatment, assoc_rel_op, target_group] has 11 instances
template [med_effect, assoc_rel_op, med_condition] has 11 instances
template [med_action, rel_op, body_part] has 11 instances
template [target_group, assoc_rel_op, lab_parameter] has 5 instances
template [lab_parameter, assoc_rel_op, diagnostic] has 5 instances
template [result_op, assoc_rel_op, disease] has 4 instances
template [target_group, recommendation, med_action] has 4 instances
template [target_group, assoc_rel_op, symptom] has 3 instances
template [med_action, assoc_rel_op, disease, rel_op, med_intervention] has 2 instances
template [treatment, assoc_rel_op, symptom] has 2 instances
template [disease, assoc_rel_op, disease] has 2 instances
template [med_context, recommendation, med_action] has 2 instances
template [decomp_op, med_intervention, assoc_rel_op, med_intervention] has 2 instances
template [disease, assoc_rel_op, disease_attr] has 1 instances
template [med_action, act_op, med_action] has 1 instances
```

```
template [med_action, decomp_op, med_action, med_action, med_action] has 1 instances
template [quantification, lab_parameter] has 1 instances
template [common_op, med_intervention, rel_op, med_context] has 1 instances
template [evaluation, med_effect] has 1 instances
template [med_attr, disorder] has 1 instances
template [med_intervention, temp_rel_op, med_intervention] has 1 instances
```

6.2.5 Related Work

Our approach to searching linguistic patterns in the guideline text is based on simple processing of text: 1. a simple sentence parsing; 2. semantic tagging; 3. deriving a set of collocations (recurrent combinations of words with a particular semantic equivalent) by applying a set of rules derived from the background knowledge of the domain; 4. merging phrases that represent concordances (co-occurences of the same phrase) among the previously identified collocations.

The task of extracting structure and semantics from text has received a great deal of attention in recent years. Assigning domain-specific categories to text is done using background knowledge in the form of conceptual graphs ([Witten, Zhong *et al.*2002]) or simpler mappings between concepts in an ontology and terms in the target domain of the textual description. Additionally, statistical and probabilistic models ([Frank and Nevill-Manning1999]) are used to increase the performance when ambiguous textual constructions are used.

6.3 Organization

6.3.1 Relations between linguistic patterns

Consider as input a text T_{GL} of a clinical guideline. Consider a transformation $SemAnnot(Text, Cat)$ of the text that produces a collection $AnnT_{GL}$ of primitive items from the set of categories Cat.

A schema is a collection of primitive items in Cat connected by relations between items or sets of items. The set of all schemas produced by Cat is denoted S_{Cat}.

For a given schema $S \in S_{Cat}$, we denote $P \subseteq S$ the fact that P is a component (subset) of S, i.e. S contains P. S is said to contain P if S has at least as many elements as P, and for each element of P a less specific element of S can be found, with the same ordering constraint in S as the source element in P (i.e., for each two elements $p_1, p_2 \in P$ with p_1 occuring before p_2 (denoted $p_1 \prec p_2$), two elements can be found in S, $s_1, s_2 \in S$ such that:
$is - more - specific([p_1], [s_1]) \wedge is - more - specific([p_1], [s_1]) \wedge s_1 \prec s_2$).

A schema $S \in S_{Cat}$ is called maximal if it is not a subschema of any other schema $S_1 \in S_{Cat}$.

A schema $S \in S_{Cat}$ is called minimal if it has no subschema in S_{Cat}.

The set of all facts capturing relevant medical knowledge present in the guideline is denoted MK. MK includes the categories in Cat. The set of all facts capturing control knowledge is denoted CK. Background knowledge is defined as the union of domain-specific knowledge and control knowledge $BK = MK \cup CK$.

The vocabulary used to define the content and the relations between linguistic patterns are simpler than in the case of control patterns:

covers : $PattTemplate\, covers\, F$, where $F \in BK$, is a mapping between S_{Cat} and BK

defined-as : $PattTemplate\, defined{-}as\, [C_1, \ldots, C_n]$, where $C_1, \ldots, C_n \in Cat, PattTemplate \in S_{Cat}$ and $PattTemplate\, covers\, F \in BK$

is-more-specific : the relation $is - more - specific$ is defined between two pattern templates $P_1 = [C_{11}, C_{12}, \ldots, C_{1n}]$ and $P_2 = [C_{21}, C_{22}, ...C_{2n}]$:
$is - more - specific(P_1, P_2)\ iff\ forall\ i = \overline{1, n}\ holds:\ is_a(C_{1i}, C_{2i})$

contains : pattern template $P_1\ contains\ P_2$, where $P_1 = [C_{11}, C_{12}, \ldots, C_{1n}]$, $P_2 = [C_{21}, C_{22}, ...C_{2n}]$ if $\{C_{21}, C_{22}, ...C_{2n}\} \subset \{C_{11}, C_{12}, \ldots, C_{1n}\}$

6.3.2 Action-centered patterns

The guidelines analyzed have a high concentration of references to medical actions. It seems therefore natural that the linguistic pattern instances that contain names of actions are among the most frequently identified pattern instances.

The pattern templates instantiated by these linguistic pattern instances can, to a large extent, be associated one-to-one with the control patterns listed in appendix section C.

6.3.2.1 Temporal ordering and decomposition of actions.

```
[med_intervention, rel_op, med_intervention]
--> [radiotherapy, following, neoadjuvant_chemotherapy]
```

6.3.2.2 Effects of actions.

```
[action, recommendation, med_effect]
--> [multidisciplinary_treatment, attention_should_be_paid, skin_reactions]
```

6.3.2.3 Associations action-goal.

```
[med_intervention, med_effect_op, lab_parameter]:
--> [surgery, to_reduce, tumour_load]
```

6.3.2.4 Recommendations for specific actions for target groups.

```
[treatment, assoc_rel_op, target_group]
--> [chemotherapy, opfor, this_group_of_patients]
```

6.3.2.5 Combinations of actions.

```
[decomp_op, med_intervention, logical_rel_op, med_intervention]
--> [addition_of, cmf_chemotherapy, opand, radiotherapy]

[med_action, decomp_op, med_action, med_action]
--> [current_treatment, consists_of, surgery, locoregional_radiotherapy]
```

6.3.2.6 Preferences for a specific medical intervention.

```
[treatment, assoc_rel_op, med_action]
--> [treatment_of_choice, opis, neoadjuvant_chemotherapy]
```

6.3.3 Background knowledge-centered patterns

A lot more of the pattern instances identified capture fragments of background knowledge that are not so useful in the formalization, as they do not have an operational component, but nonetheless they are relevant if maintenance of background knowledge is considered as the great challenge the implementation of living guidelines is going to face.

6.3.3.1 Associations disease-treatment.

```
[target_group, assoc_rel_op, disease]
--> [inoperable_patients, with, tumour]
```

6.3.3.2 Associations disease-target group.

```
target_group, assoc_rel_op, disease]
--> [oncology_patients, with, locoregionally_advanced_breast_cancer]
target_group, assoc_rel_op, disease]
--> [inoperable_patients, with, tumour]
```

6.4 Usage

6.4.1 Connecting patterns to formal representation

Once the patterns of a particular granularity level (word-level, sentence-level, paragraph-level) have been identified, the next important step is to generate the formal ASBRU constructs that correspond to each pattern.

For instance, an instance $i1 = (action1, following, action2)$ of pattern $p1(action, act-op, action)$ above is transformed into the ASBRU plan:

```
plan i1
comments
  "auto-generated from pattern instance i1"
plan-body
  subplans type=sequentially wait-for=all retry-aborted=no
     plan-activation action1
     plan-activation action2

plan action1
plan-body to-be-defined

plan action2
plan-body to-be-defined
```

in accordance to the prescription of the library of patterns which we have built so far ([Serban2004]).

A very general sketch of the ASBRU set of plans that describe the main activities can be generated by considering only the sentence-level patterns that include actions and action operators that are instance of temporal ordering relations.

The details of the ASBRU set of plans can further be filled in using instances of patterns that include laboratory parameter names: these pattern instances will be used to generate $parameter - def$ entries in the $domain - defs/domain$ section of the ASBRU set of plans. Not all these parameters can be filled in automatically.

6.5 Evaluation

The overall impression is that the pattern occurrences/instances detected through the linguistic approach are rather low-level, overlapping (indeed it has been mentioned that this is mostly due to the redundancy in the relations described in the ontology used as a starting point), and mainly focussed on temporal relations between actions, such as ”before” and ”after”. Although this fact can be explained by an interest in identifying the knowledge elements necessary for obtaining an operational version of the guideline, it is important to note that other fundamental types of knowledge, such as scientific conclusions, seem to have been disregarded. This observation is surprising since in this particular case, which corresponds to knowledge where the phrasing is rather structured (to the point that some form of standardisation has been discussed), one would expect that the pay-off of pattern search/identification is higher than in the case of action sequencing.

In this scenario it becomes relevant to evaluate if it is possible to build a coherent (fragment of an) MHB model from the pattern instances detected, which somehow amounts to assessing if the patterns employed in the detection are adequate to be used as a sort of knowledge acquisition pre-processing. The basic idea would be analysing if we can build an acceptable MHB model by using the pattern instances identified in relevant paragraphs to devise a MHB structure capturing the knowledge described in those parts. The “golden standard” we will use to determine whether the model is acceptable or not is the MHB model manually built by knowledge engineers (KEs).

In the light of the above, the evaluation we have carried out consisted of **(1)** a rough comparison (quantitative) of the amount of knowledge (automatically) identified by using patterns with respect to the knowledge modelled by (manual) knowledge acquisition; **(2)** an analysis (qualitative) of the utility of identified pattern instances, performed on specific fragments of the guideline; and **(3)** in connection with the two previous steps, an identification of the essential knowledge elements that current patterns overlook. The same steps in more detail:

(1) Some comparison of the amount of sentences in which the Java application has detected patterns with respect to the sentences modelled by the KE can give us an idea of the coverage of the detection process. To facilitate matters, the comparison will be restricted to procedural knowledge.

(2) Our analysis of the utility of the pattern instances detected by the application has consisted in studying these instances to determine whether a significant piece

	(autom.) processed sentences	(manual.) modelled sentences	modelled sentences processed	modelled sentences with patterns	modelled sentences with significant patterns
chapter 2	130	41	30 (73%)	4 (9.7%)	0
chapter 3	134	20	16 (80%)	7 (35%)	2 (10%)
chapter 4	91	25	18 (72%)	2 (8%)	0

Table 1: *Results summarizing the role of linguistic patterns in modelling for ASBRU*

of MHB model can be obtained from them. By means of this exercise we can have an idea of the potentials of the detection process for use as a knowledge acquisition tool.

(3) Finally, the identification of knowledge elements that are essential for obtaining a model of quality and that for some reason (e.g. there are no patterns dealing with them) have been disregarded by the Java application, will allow us to better direct our future research efforts.

6.5.1 Coverage of procedural knowledge by pattern detection

We have evaluated the coverage of the detection process with respect to the procedural parts modelled by the KE by calculating the percentage of sentences in those parts for which some pattern has been detected. Table 1 shows the numbers we have obtained in the different chapters.

The first column of table 1 presents the number of sentences processed by the application, which amounts to the number of sentences considered as relevant to the guideline topic (using a keyword list as criteria). The second and third columns give respectively the number of sentences actually modelled by the KE (i.e. the sentences considered relevant from the KE's viewpoint) and, among them, the amount of sentences processed by the application (both the number and the percentage with respect to the modelled sentences). Finally, the last column shows the amount of sentences among the latter (modelled by the KE and also processed by the application) where some pattern has been found.

Some conclusions follow. Firstly, it is important to note that although the amount of sentences that are considered relevant by the application exceeds by far the modelled knowledge, the latter is covered to a significant extent (between 70% and 80%). Nevertheless, the percentages drop considerably when we look at those modelled and processed sentences where patterns have been detected. The numbers get dramatically bad when we (manually) analyse the patterns to determine if they contain significant knowledge, i.e. if there is a chance to produce a coherent piece of MHB. An example of non-significant pattern is shown below:

```
=== Relevant sentence S(48) has 6 medical terms: [adjuvant, che-
motherapy, risk, recurrence, rate, survival]

Sentence(48): [adjuvant_chemotherapy, reduces, the, risk_of,
recurrence, opand, improves, the, rate, of, survival, eos]

1: [med_intervention, med_effect_op, lab_parameter] (p_result)......
    most specific:[med_intervention, med_effect_op, symptom]......
    least specific:[concept, med_effect_op, concept]
    [adjuvant_chemotherapy, risk_of, recurrence]
```

In this example it is obvious that the pattern instance not only is missing the kind of relation between adjuvant chemotherapy and the risk of recurrence (which is reduced) but also its relation with the rate of survival (it is improved). Therefore it would be difficult constructing an MHB fragment that properly represents (not even partially) the original sentence from that pattern.

All the above implies that the patterns used as starting point not only are not suitable for detecting the main bulk of modelled knowledge but also fail in capturing the essential parts thereof. The next section tries to evaluate the quality of some of the pattern instances detected.

6.5.2 Potential of pattern detection in knowledge acquisition

As it has been skeched before, in order to evaluate the utility of detected pattern instances, we have studied some of them to determine if a coherent MHB fragment could be obtained.

Since the results of the previous step reveal a bad coverage of procedural knowledge with patterns, in this step we have widened the scope of our analysis by considering other types of knowledge (e.g. data) and even non-modelled knowledge. In addition to this, we have included sentences with non-relevant pattern instances. Some examples of follow (the first two exapmples are modelled sentences with significant patterns and the other two refer to a data definition and a sentence with non-significant patterns, respectively).

6.5.2.1 Example 1 This example corresponds to a sentence within a conclusions' section, which is *Grade 3 — Adjuvant hormone therapy for locally advanced breast cancer results in improved survival in the long-term — A2 Bartelink25*.

The pattern detection process generates the following instances:

```
Sentence(93): [c, kuerer44, grade, nr3, adjuvant_hormone_therapy,
opfor, locally_advanced_breast_cancer, results_in, improved, survival,
in, the, long_term, eos]

1: [med_effect_op, med_effect] (p_result1)......
    most specific: [med_effect_op, med_effect]......
    least specific:[med_effect_op, result]
    [results_in, survival]
2: [med_action, assoc_rel_op, disease] (p_assoc_act_disease)......
    most specific:[med_action, assoc_rel_op, disease]......
```

```
    least specific:[concept, generic_op, med_concept]
    [adjuvant_hormone_therapy, opfor, locally_advanced_breast_cancer]
```

In summary, the first two patterns describe an effect (keyword `results_in`) which refers to survival, and the third one associates a therapy to a specific disease. Since the patterns come from the same sentence, we can assume the former refer to the latter, which would lead to the tentative MHB fragment shown below:

```
<chunk chunk-id=''CHUNK-001''>
        <control>
                <clinical-activity description=''undefined''
                        name=''adjuvant-hormone-therapy''>
                </clinical-activity>
        </control>
</chunk>
<chunk chunk-id=''CHUNK-002''>
        <background>
                <effect value-name=''survival''
                        expected-change=''results-in''
                        cause=''adjuvant-hormone-therapy''>
                </effect>
        </background>
</chunk>
```

The above fragment resembles to what the KE modelled, except for the framework of the background knowledge, which in the case of the manual model is an evidence chunk, and for the kind of change in the cause-effect relation, which should be `improves` rather than `results_in`. We can conclude that patterns are needed to detect evidence knowledge, and that the list of keywords used to detect patterns with medical effects should be enriched, or the corresponding patterns further qualified, so that they can be properly identified.

6.5.2.2 Example 2 This example comes from a recommendations' sentence: *Patients with locoregionally advanced breast cancer should receive multidisciplinary treatment with curative intent*.

```
Sentence(140): [recommendations, patients, with, locoregiona-
lly_advanced_breast_cancer, should, receive, multidisciplina-
ry_treatment, with, curative_intent, eos]

1: [target_group, assoc_rel_op, disease] (p_ctx1)......most
specific:[target_group, assoc_rel_op, disease]......least
specific:[concept, generic_op, concept]
    [patients, with, locoregionally_advanced_breast_cancer]
2: [med_context, recommendation, med_action] (p_rec1)......
most specific:[med_context, recommendation, gen_med_action]
......least specific:[concept, concept, med_concept]
    [patients, should, multidisciplinary_treatment]
3: [target_group, recommendation, med_action] (p_rec_target-
group)......most specific:[target_group, recommendation,
gen_med_action]......least specific:[concept, concept,
med_concept]
    [patients, should, multidisciplinary_treatment]
```

Below we present the tentative MHB fragment produced from the previous pattern instances, which we assume to be related (on the basis of the common word `patients`):

```
<chunk chunk-id=''CHUNK-001''>
        <control>
                <clinical-activity description="undefined"
                        name="multidisciplinary_treatment">
                </clinical-activity>
        </control>
</chunk>
```

In this case, the model by the KE additionally includes a background element that describes the intention of the multidisciplinary treatment and thus is richer. In our view, a refinement of the recommendation patterns is something to consider in order to improve the detection process.

6.5.2.3 Example 3 This sentence is a good example of the kind of definitions that can be found in a guideline: *Locoregionally advanced breast cancer is used to describe breast cancer which is unresectable on the basis of the classic unresectability criteria: oedema of the skin (peau d'orange), ulceration, satellite skin nodules, inflammatory carcinoma, infi ltration of the chest wall (T4), lymph nodes fi xed to one another and/or to deeper structures (N2), or palpable internal mammary, parasternal, infraclavicular and/or supraclavicular lymph nodes (N3).*

The pattern detection process produces the following output:

```
Sentence(21): [definition, locoregionally_advanced_breast_
cancer, opis_used_to_describe, breast_cancer, which_opis,
unresectable, on_the_basis_of, the, classic, unresectability_
criteria, dblpnt, oedema, of, the, skin, lpar_peau, d_quote_
orange_rpar, ulceration, satellite, skin, nodules, inflamma-
tory, carcinoma, infiltration, of, the, chest, wall, lpar_t4_
rpar, lymph, nodes, fixed, to, one, another, and_bkslash_or,
to, deeper, structures, lpar_n2_rpar, opor, palpable, inter-
nal, mammary, parasternal, infraclavicular, and_bkslash_or,
supraclavicular, lymph, nodes, lpar_n3_rpar, eos]

1: [disease, assoc_rel_op, disease](p_disease_def)......
    most specific:[disease, assoc_rel_op, disease]......
    least specific:[concept, generic_op, med_concept]
    [locoregionally_advanced_breast_cancer, opis_used_to_
describe, breast_cancer]
2: [disease, assoc_rel_op, disease_attr] (p_def1)......
    most specific:[disease, assoc_rel_op, disease_attr]......
    least specific:[concept, generic_op, med_concept]
    [locoregionally_advanced_breast_cancer, opis_used_to_
describe, unresectable]
3: [disease, assoc_rel_op, disease, assoc_rel_op, disease_attr]
(p_def2)......
    most specific:[disease, assoc_rel_op, disease, assoc_rel_op,
disease_attr]......
    least specific: [concept, generic_op, med_concept, generic_op,
med_concept]
    [locoregionally_advanced_breast_cancer, opis_used_to_describe,
breast_cancer, which_opis, unresectable]
4: [disease_attr, assoc_rel_op, med_factor] (p_factor1)......
    most specific:[disease_attr, assoc_rel_op, med_factor]......
    least specific:[med_concept, generic_op, med_concept]
    [unresectable, on_the_basis_of, unresectability_criteria]
```

From the above pattern instances we can derive the MHB fragment below, after discarding the first two instances (the first one because it is subsumed by another instance, and the second because it can be seen as an anomaly in the detection process):

```
<chunk chunk-id="#CHUNK-001">
        <data>
              <abstraction abstraction-rule="unresectableBC"
                  result="locoregionally-advancedBC">
              </abstraction>
        </data>
</chunk>
<chunk chunk-id="#CHUNK-002">
        <data>
              <abstraction abstraction-rule="unresectability-criteria"
                  result="unresectableBC">
              </abstraction>
        </data>
</chunk>
```

The resulting MHB code is fully coherent with the sentence, the problem is that the specification of the unresectability criteria is missing, because no pattern instance was found in the sentence from the collon. This is in contrast with what the KE modelled, which not only includes a disjunctive expression as the abstraction rule of the second chunk, but also incorporates an additional concept within the rule of the first chunk. This concept is derived from the knowledge contained in a different sentence, next to the analysed one: *In addition, large primary tumours (>5 cm; T3) are also included in this category (T3, T4, all N classes, M0; all T classes, N2 or 3, M0)*. Since the latter issue is hard to solve if we stick to the (simpler) sentence-level analysis, the main suggestion would be enriching the definition/description pattern so that they are able to detect lists of items.

6.5.3 Focus of future research on pattern detection

In the course of the previous steps we have concluded that the patterns used by the application are often not suitable. Here we try to recapitulate on our experiences and give indications that can help in improving the design of the linguistic patterns and thus of the process itself. According to this, we have distinguished two kinds of indications/comments, namely suggestions about patterns and comments on the difficulties of the detection process.

6.5.3.1 Suggestions about patterns Based on the examples we have analysed (and also on the examination of the output of the pattern detection process) we have identified a series of additional patterns which we think could be of help:

1. A pattern to detect **definitions by enumeration** should be considered. An instance example: ... *unresectability criteria: oedema of the skin (peau d'orange), ulceration, satellite skin nodules, inflammatory carcinoma, ...* .
2. A series of patterns to detect **conclusions** should be defined taken into account the homogeneous structure thereof.

3. **Intentions** should be captured by specific patterns defined from keywords such as `intention`, etc.

4. **Cause-effect relations** should be detected on the basis of the keywords typically used, such as `improves`, `reduces`[1], `increases`, `decreases`, etc, and combinations thereof possibly including "weasel" words.

5. Patterns dealing with **comparisons** should be considered, because the comparison between treatments, trials, results, etc is frequent.

6. Additional patterns should be included to detect **non-conventional conditionals**, such as in the following example: *Hormonal adjuvant therapy, if the hormone receptors are positive*.

7. Patterns for **recommendations** should ensure that keywords like `recommended to`, `advised to take`, `be considered`, `be recommended`, `qualify for`, etc are taken into account.

8. **Additive treatment combinations** should be detected in cases such as *...adding surgery to neoadjuvant chemotherapy* or *...addition of surgery to chemotherapy*[2]

6.5.3.2 Comments to the detection process As result of the evaluation, some problems in the pattern detection application have been uncovered, which are listed below:

1. Conclusion boxes are split in the wrong way (as they are extracted from tables), resulting in sentences that group a conclusion together with the literature information of the previous one. For example, the text:
   Grade 3 / Anthracycline-containing chemotherapy results in higher response rates than other forms of chemotherapy / C Kuerer44

   Grade 3 / Adjuvant hormone therapy for locally advanced breast cancer results in improved survival in the long-term / A2 Bartelink25
   . . . is split like follows:

   ```
   Sentence(93): [c, kuerer44, grade, nr3, adjuvant_hormone_therapy,
   opfor, locally_advanced_breast_cancer, results_in, improved, survival,
   in, the, long_term, eos]
   ```

2. Some sentences are ignored in the detection process, as they are not split correctly due to the fact that ":" is regarded as a sentence splitter. For instance, the highlighted parts were not included as parts of the first sentence, but treated apartly:

[1] Although used "result" patterns seem to be dealing with these cases, we have noticed problems.

[2] Although "action composition" patterns seem to deal with these cases, we have noticed they have problems.

Patients with locoregionally advanced breast cancer should receive multidisciplinary treatment with curative intent. ***The treatment should comprise the following components:***
**** Neoadjuvant chemotherapy, preferably six courses containing anthracycline.***
**** Surgery to reduce the tumour load. It is unclear whether limited or radical surgery is better.*** *In case of a clinically negative axilla, omission of axillary node dissection appears justifi able.**
*** ***Locoregional radiotherapy.***
** Hormonal adjuvant treatment, if the hormone receptors are positive.*

6.6 Conclusions and lessons learned

Searching for linguistic patterns is a text-driven method to obtain patterns. It takes as input the text of an existing guideline, and attempts to reverse engineer the linguistic pattern templates that were used to produce the text. A linguistic pattern template is an ordered list of semantic categories specific to the domain referenced by the guideline, which are linked by semantic relations relevant for the domain, which in turn can be expressed in a form that obeys some syntactic rules.

We propose a method of searching for linguistic patterns in the text of a guideline text, which 1. reduces the time spent in the guideline formalization process and improves the quality of the resulting formal representation; 2. facilitates understanding of the aspects covered by the guideline and enables their validation with respect to the requirements for the guideline; 3. allows a more effective maintenance of the background knowledge covered by the guideline, thus enabling living guidelines.

The method was inspired from a practical experiment with a tool for guideline formalization (Stepper [Ste]), and basically consists of:

1. learning of an ontology from a large body of guideline text;
2. assigning semantic categories to classes learnt automatically, and adjusting/filtering the ontology;
3. semantically tagging the text, using the semantic categories defined above;
4. searching for patterns obtained from meaningful formal constructs (i.e., laboratory parameters that are assigned values, temporal ordering of actions, etc) and in parallel:

   i. adjusting the ontology (by adding new relations that occur frequently in patterns deemed to be relevant by the medical expert);

   ii. adjusting the level of abstraction of the patterns sought or adding more complex patterns to be sought, which contain: abstractions of the concepts used in existing patterns, patterns which specify constraints between the detected patterns (such as, that two or more patterns are detected in the same sentence and they contain a common word).

5. finding common words and common prefixes shared by several patterns, words shared that belong to different semantic categories, etc, and trying: i. to invalidate some of the pattern candidates found; ii. to build more complex pattern, whose instances are then sought in other texts.

6. finally, collecting all pattern templates that have instances in more guideline texts, in a pattern library, and selecting the meaningful ones as building blocks for future guidelines.

Steps 1, 2 and 5 seem to be the most difficult. If an ontology for the medical domain would be used (even one of the existing medical thesauri - MeSH ([MeSa]), UMLS ([UML]), or NCIOncology ([NCI])), which actually will be the case in a living guideline, it would be much easier to obtain linguistic patterns instances, as the most basic building blocks for more complex patterns.

The conclusions of the evaluation of the linguistic patterns show that they are too finely-grained to be effectively used for direct production of ASBRU plans. However, as these patterns are closer to the Intermediate Representation language MHB, they can be used to transform the guideline text in an Intermediate Representation first. Subsequently, other classes of patterns (control or background patterns) can be used to further transform this Intermediate Representation of the guideline into ASBRU. It can thus be concluded that the linguistic patterns can play an important role in the beginning of the guideline formalization.

Among other advantages of the linguistic patterns, we list the following:

1. they represent the most finely-grained building blocks for larger patterns, underlying other text-oriented patterns, namely the control patterns and the background patterns;

2. they can be validated in a straightforward way, by means of marking them up in the text of the guideline and inspecting the context in which they appear to decide whether they are in fact relevant for formalization;

3. they are compositional, and some simple rules for extending them exist, which are based on co-references to the same medical terms, the same medical categories, or the same medical context.

However, it should be noted that the results of searching for linguistic patterns depends heavily on the existence of an ontology of the medical obtain and how easy is for semantic relations in that ontology to be mapped onto syntactical constructs. Medical thesauri exists, but only small parts of the semantic relations they contain are used in practice by the guideline authors to produce the text of the guideline.

Part IV
Background Patterns

7 Background Patterns

7.1 Characterization

Evidence-based guidelines are built based on evidence collected from the medical literature and structured in so-called evidence tables. Having said that, the knowledge involved in the actual composition of a guideline has a wider scope than simply the evidence tables. The most obvious example is the situation where the guideline developers need to answer a question, but no literature is available to gain solid evidence for any answer to this question. In such a case, the working group will try to reach a consensus.

Evidence and consensus-based conclusions are by itself not enough to formulate recommendations. Sometimes this reasoning from evidence to recommendations is made explicit, but in most cases it is not.

The goal of this section is to provide insight in the structure of evidence and the structure of reasoning. These patterns are never made explicit in the text itself, and in that sense they are background patterns. The underlying structure of the recommendations are handled in chapter 8.

7.2 Extraction

Background patterns are obtained by GL authors and by knowledge engineers, by directly using background knowledge of the medical domain covered by the guideline, or by extracting this knowledge from the guideline text.

For each medical process we have to identify its dependencies, in terms of other processes, resources and information required for its execution.

1. a description of the task or process that the knowledge is used for, e.g. assessing life expectancy of cancer patients;
2. the high-level goal this process contributes to; e.g. diagnosis or treatment;
3. which agents/roles should execute the process and under which circumstances, e.g. a nurse or a team of care practicians;
4. frequency of use of the knowledge;
5. the duration of the task and the validity of the observations.

7.2.1 The linguistic and control component of background patterns

A background pattern is one that either has a well-defined underlying linguistic pattern, or one that describes parameters of a standard knowledge relation. For instance, the definitions of parameters always have a range of values and values within a given interval are deemed as ”normal”, while others are ”abnormal”.

Examples of background patterns having a linguistic pattern:

```
1. [Histological diagnosis]:ACTION can be executed by means of [core
biopsy]:ACTION and [FNAC]:ACTION.

2. Since []:SITUATION, it is advisable to do []:ACTION.

3. It is advisable to achieve []:GOAL by means of []:ACTION.

4. For []:CONTEXT the risk of []:DIAGNOSTIC is higher than for
[]:CONTEXT.

5. Alternatives for []:TREATMENT of []:DISEASE depend on
[]:MED_FACTOR.

6. Prognosis of [patients with []:DISEASE]:TARGET_GROUP is poorer than
prognosis of [patient with []:DISEASE]:TARGET_GROUP.

7. []:ACTION consisting of [[]:ACTION, ..., []:ACTION]:SET(ACTION)
is recommended, in order to achieve []:GOAL.
```

7.2.2 The breast-cancer example

Treatment of breast cancer centres around the question whether the condition is curable or incurable. Whereas in the former situation treatment will be intensive, trading unpleasant consequences for the patients against gain in life expectancy, treatment in the latter situation will be palliative. Note that curable breast cancer still means that some 10-40patients may die due to the condition within 5 years after diagnosis. Hence, distinguishing between patients with a good and poor prognosis is far from easy, which explains the role and importance of clinical staging. Given our current knowledge about the prognosis in relationship with particular stages of the disease, the issue is very much which life expectance threshold is considered acceptable (for example, is 90% survival after 5 years acceptable, or is 80% survival after 5 years already acceptable)? Only for early breast cancer, the decision is easy, as there is much to gain from treatment with curative intent. Yet, with nonlocalised breast cancer, the decision-making process is much more difficult. Hence, underlying the design of the breast cancer guideline are consideration about curability, as unjustified treatment of a patient with incurable breast cancer with curative intent should be avoided as much as treatment of curable breast cancer with palliative intent. This is also a potential source of medical mistakes in the guideline, and thus one question is how such mistakes can be identified.

One problem is that this is partly a matter of opinion: whereas one clinical expert would consider a survival rate with intensive treatment of 40% acceptable, other experts would find this rate too low.

This implies that underlying a guideline are considerations why particular interventions, diagnostic or thepapeutic, are undertaken. These considerations may involve background knowledge, for example about the biological behaviour of tumour cells, which in many cases will not be incorporated into the guideline, whereas sometimes it is. In terms of patterns, the structures used in medical decision-making theory, where problems are described in terms of sequences of actions, taking into account known or unknown evidence and a weighing of the outcomes of actions given the patient's condition, expressed by utilities, seems appropriate.

One of the problems of locally advanced breast cancer the topic of Chapter 3 of the breast-cancer guideline is that it is only recent that clinical opinion has changed from taking this stage of breast cancer as being incurable (the cancer was too much advanced) to curable. This is mainly due to treatment improvements, such as neoadjuvant chemotherapy, which aims to shrink the tumour before the actual treatment starts. One of the clinical laws underlying management of locally advanced breast cancer is that it is treated with curative intent, whereas there are signs and symptoms, for example distant metastases, rendering the curative intent unrealistic. This implies that the patient received a much more intensive (and therefore unpleasant) treatment than is clinically warranted. This is clearly something that must be avoided. This explains the care taken in the guideline to identify patients that should be excluded from curative treatment, such as patients with liver metastases.

7.3 Organization

7.3.1 Patterns in evidence

The guidelines developed by the CBO are evidence-based, which means that in practice, when a health-care professional or patient need to make a decision, he or she should be aware of the evidence and its strength that support this decision. The strength of the evidence is judged based on the quality of the studies that have been performed. For example, strong evidence comes from randomised controlled trials (RCTs), while opinions based on clinical experience are considered as being much weaker. The first step in collecting relevant evidence from literature consists of drafting an appropriate question, which is done in a structured fashion. A question always consists of four subcomponents, which is expressed by the acronym PICO; PICO stands for:

P: Patient group , i.e. the group of patients having particular characteristics (have a particular disease, e.g. patients with breast cancer, or have undergone a particular treatment);

I: Intervention , i.e. an action such as treatment;

C: Comparison with other interventions , e.g. one treatment against another one, or one diagnostic test against another one;

O: Outcome : the result of the intervention, for example, diagnosis, cure, survival or death.

The utilisation of the PICO method is done for a number of reasons: it forces one to focus on the issues that really matter with regard to the scope of the guideline, producing questions which can be answered in principle if enough literature is available. As a result it creates an upperbound on the number of papers that have to be examined.

The results of PICOs are put into evidence tables and contain additional information to establish the weight of this evidence. Once this evidence has been selected, it is often written down in guideline as a conclusion. Not surprisingly, the structure of PICOs can be found in the guideline itself as well. On a more linguistic side they can take many forms, but in essence many conclusions that we find in a guideline can be seen as a qualitative comparison between interventions given some outcome that is apparently intended to be reached for a certain patient-group.

7.3.2 Process patterns

A guideline provides information about the proper management of patients with a particular disease, based on scientific evidence from the medical literature. This implies that in essence there are two processes that come together in a guideline: (1) the disease process in a group of patients, for example patients with a particular type of breast cancer; (2) medical management, described in terms of the actions, their effects, and the order of actions to be undertaken. Whereas the former process is based on a mixture of scientific evidence and clinical experience, the content of the latter process is, or should be, based on scientific evidence. In our example a description of a patient with locally advanced breast cancer is an instance of the disease process, whereas a description of the curative recommendation would be an instance of the medical management process. These patterns are related to medical decision-making theory, as mentioned above.

7.3.3 Invariants

Any intervention, diagnostic or therapeutic, is mentioned in the guideline because it is expected that it will have some beneficial effect or produce useful information about the patient's condition. While no general quality indicators are being written down specifically, the guideline gives us hints on which parameters should be invariant throughout the use of this guideline. For example, we can easily deduce that it is important that histological diagnosis can be performed without surgery at any time in the treatment. Another example is the invariant that locoregional treatment can only be administered if the tumour has been reduced to a certain level by neoadjuvant chemotherapy. Hence, in general it appears that underlying the guideline are semantics patterns of the form:
"if an intervention is undertaken, it is the best intervention in the given situation"

7.3.4 Causal patterns

Much of the background knowledge of a guideline is essentially physiological in nature, having to do with the underlying mechanisms of cancer and the mechanisms activated by treatment.

Much of this knowledge has the form of conditional cause-effect relationships, which, for example, have been studied in some detail in diagnostic reasoning. This gives rise to causal reasoning patterns that help understand both what happens with the patient. For example:

"lymphatic obstruction after irradiation CAUSES lymphoedema"

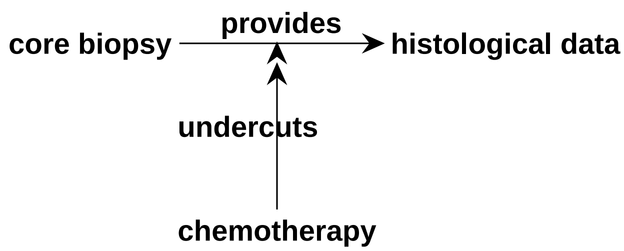

Figure 17: *Short semantic sketch of a guideline fragment*

The subphrase ‘after’ indicates that time involved in this type of causal reasoning, although there are few instances in the breast-cancer guideline where explicit references to time are given.

7.3.5 Correlation of signs and symptoms

Consider the case of imflammatory breast cancer:

> *"Inflammatory breast cancer, characterized by diffuse redness, peau d’orange and possibly swelling ..."*

this might be translated into elementary propositions:

- inflammatory-BC
- diffuse-redness
- peau-dorange
- possibly-swelling

However, this would imply that interpretation of these subphrases is not taken into account. Now interpretation of these subphrases is not at all trivial: what is the meaning of diffuse redness (does localised redness exclude inflammatory breast cancer?), why is possibly swelling mentioned, is it becuase the other signs are mandatory?

7.3.6 Reasons for ordering of interventions

When reading the text one can recognize certain constraints on the treatment. A core biopsy provides histological data if the patient did not receive chemotherapy and because of this the core biopsy is performed before chemotherapy. In principle, histological data and information about the receptor status is obtained after (during) surgery, by examing the resected tumour mass. If the patient received chemotherapy before surgery, then this is no longer possible. So there is some kind of structure that specifies how the concepts ‘core biopsy’, ‘resected tumour’, ‘histological data’, ‘receptor status’ and ’neo(adjuvant) chemotherapy’ are influenced by each other (see figure 17). Hence, we have certain relationships between the concepts in our model.

7.3.7 Sloppy meaning

Even though understanding the meaning of the words of a guideline requires access to an ontology, there are situations where for historic reasons misnomers are still used. For example, inflammatory breast cancer is not al all inflammatory, it is invasion of the lymphatic channels. Yet the term is still being used. Thus sloppiness of meaning is typical for medicine. Whereas it is something one will expect in every-day use of a language, one would not expect this in an area that has a scientific basis. This renders the use of a standard ontology difficult.

7.4 Usage

Knowledge from the target (medical) domain is used to build the structure of the guideline document.

The relations between target groups covered by the guideline and the relations between domain-specific processes define the structure of the document.

Whenever there is evidence for a medical statement that is relevant to treating the disease handled by a clinical guideline, this evidence is captured as a recommendation in the guideline.

A mapping between particular facts in the target domain and narrative structure can be devised, but this mapping depends to a large extent on the target domain of the guideline.

7.5 Conclusions and lessons learned

Having a clear terminology defined for the domain for which a guideline make prescriptions is of a great advantage when maintaining the knowledge embedded in a guideline.

Ambiguities caused by using a slightly different synonym can be eliminated, and relations in which different terms are with each other can be easier established.

This is likely to improve the quality of the guideline document and to ease the knowledge management process of living guidelines.

Part V
Control patterns

8 Control Patterns

8.1 Characterization

These are programming constructs that are less dependent of a certain linguistic form, but are often present in the text of the guideline. They include action decomposition, ordering, and repetition, or describe constraints such as the fact that an action may be associated with a time frame, intention and medical effect.

They are the closest constructs to an executable model, and therefore the formalization would benefit most if many of these constructs can be identified in the guideline text.

8.2 Extraction

Control patterns are independent from the text of the guideline. They are generated by the knowledge engineers that produce the formalization. The type and the level of detail of the control constructs used in the control patterns are influenced by the target executable language chosen by the team performing the formalization of the guideline. In our case, the control patterns are very similar to the ASBRU constructs.

The list of frequently encountered control patterns listed in section C has been obtained as result of one or two abstraction steps from the actual pattern instances that can be isolated from the guideline text:

1. replacing medical concepts with conditions, actions, goals, etc. and
2. identifying their intentions, and ignoring the components that do not match the intermediate representation.

These transformation steps require a considerable amount of background knowledge, not only about the medical domain, but also about the action execution model envisioned by the guideline authors.

8.2.1 Vocabulary for defining patterns

To describe the correspondence between terms present in text fragments and medical concepts, and to define the internal structure of a (control) pattern, we use a small vocabulary, consisting of the following elements:

1. standard taxonomy relations: is_a, instance_of, part_of; ex: $MedAction\ is\_a\ Action$; $CyclicalAction\ is\_a\ Action$; $'repeat'\ instance\_of\ CyclicalAction$;$'mastectomy'\ instance\_of\ MedAction$;

2. defined_as, refined_as; ex:
   $PattTemplate(medical - action - sequence)\ defined\_as$
   $MedAction\ followed\_by\ SequenceOperator\ followed\_by\ MedAction$;
   $SequenceOperator\ refined\_as$
   $select(one\_of)\ from\ group(TemporalOperator, ActionOrderOperator)$
3. applies_to
4. supported_by
5. uses, requires
6. binds, implements
7. group_of, combination_of
8. select (first n, last n, one_of, all) from list
9. group $elem_1, \ldots, elem_n$ as label
10. eval(expression)
11. execute
12. seq,parallel,followed_by
13. and,or,xor,not
14. excludes
15. forbidden, permitted, recommended, possible, impossible

8.2.1.1 Programming structures present in control and implementation patterns. The following ASBRU-inspired basic control constructs allow us to represent the control patterns. By using the vocabulary defined in section 8.2.1, we are able to represent in a more compact form and at a higher level of abstraction than ASBRU, patterns that include programming structures and decomposition/sequences of actions.

1. Sequence: seq(A1:ACTION,A2:ACTION)
2. Parallel split: parallel(A1:ACTION,A2:ACTION)
3. Assign: assign(AT:ATTR,V:VALUE)
4. Effect: effect(A:ACTION,P:PROPOSITION)
5. Exclusive choice decision: decision(P1:PROPOSITION?,A1:ACTION,A2:ACTION) Based on the truth of P1 one of the possible branches is chosen.
6. Selection: select(F:FUNCTION,S:SET,V:VALUE)

7. Loop: repeat_until(P:PROPOSITION,A:ACTION) Action A has to be repeated until proposition P becomes true. If P is already true, A does not execute. P may be specified loosely, such as: any—undefined—context(C)

8. Wait: wait(T:TIME)

9. Interrupt action: stop(P:PROPOSITION,A:ACTION) Stops execution of A if proposition P is true.

10. Cancel action: cancel(A:ACTION) Action A is unconditionally stopped stop(true,A) and additionally its effect becomes undefined:
$cancel(A) \stackrel{def}{=} seq(stop(true, A), effect(A, undefined))$

11. Trigger: trigger(P:PROPOSITION,A:ACTION) Action A is triggered when proposition P becomes true

12. User action: user(A:ACTION) Action A will be performed by the user.

13. Data definition: var(V:ATTR,C:CONTEXT) A variable V is defined within the context C.

14. Value definition: val(V:ATTR,V:VALUE,C:CONTEXT)

15. Set of elements: set(E:ELEMENT,[E_i:ELEMENT])

    The following constructs place an action in a context and indicate the deontic mode for that context:

16. Context: context(C:CONTEXT,A:ACTION) Context can be one of:

    - any: matches any context
    - none: there is no context that matches this one
    - before(A:ACTION)

17. Mode: mode(NEC—OBL—OPT—NOT—XOR A1:ACTION,C:CONTEXT,A:ACTION) This indicates the strength that action A is executed in context C:

    - REC: A is recommended (optional, but preferred)
    - NEC: A is necessary (the action cannot complete without it)
    - OBL: A is obligatory (the action can be bypassed, but with penalties)
    - OPT: A is optional (the action can be bypassed without penalties)
    - NOT: A should be avoided
    - XOR A1:ACTION: execution of A should be xor-ed with execution of A1 (either A or A1 will be executed)

    Relations between actions and actions/goals:

18. achieves(A:ACTION,GOAL:PROPOSITION)
19. has_intention(PLAN:ACTION,GOAL:PROPOSITION)
20. begins_with(PLAN:ACTION,ACT:ACTION)
21. defined_as(PLAN:ACTION,ACT:ACTION)

8.3 Organization

In figure 18 we have depicted the main programming-like structures that make up the elements that can be transformed into a formal/executable representation.

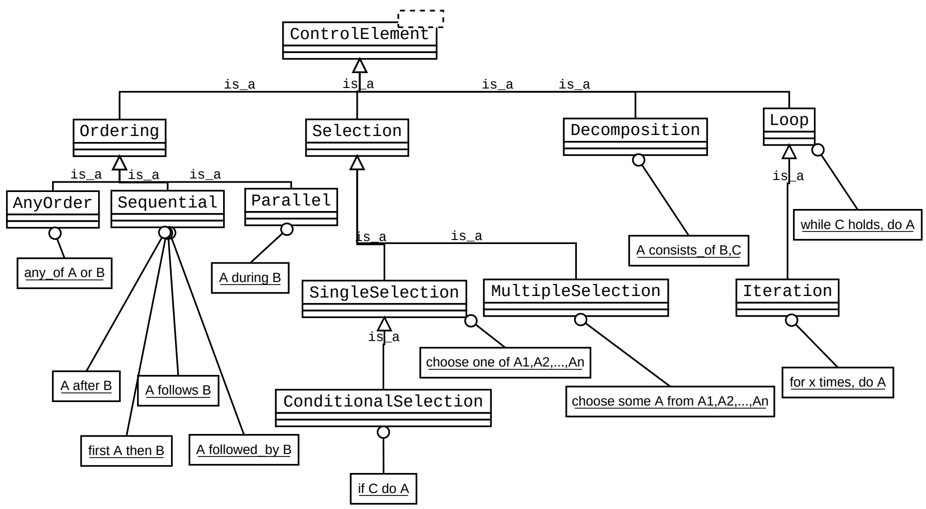

Figure 18: *Relations between the basic control structures (patterns) used in guideline formalization*

In our approach, a guideline pattern is represented as a tuple (L, C, TR, IR, FR, P, O, Loc, Arg, Voc) where:

- L is a Label/Description of the guideline pattern
- C is a set of orthogonal categories to which the pattern belongs (according to different criteria)
- TR is a Text Fragment/Instances (in the language used to disseminate the guideline) representing an instance of the guideline pattern ready to use in the guideline; in some cases, an instance of the pattern is a triple (TF, IR, FR)
- IR is an Intermediate Representation of the guideline pattern
- FR is a Formal Representation of the guideline pattern

- P is a proposition (or set of propositions) about the Properties of the pattern (location, compositionality) and Relations to other patterns of the library
- O is a set of observations about the pattern that is not covered by the other attributes of the pattern (i.e., exceptions, additional parameters, definition of possible values that occur in the pattern)
- Loc is the typical location in the guideline text where the pattern can be encountered
- Arg is a list of arguments passed as parameters to the pattern template
- Voc is a vocabulary of terms in guideline text that can instantiate the elements of IR or Arg

Except for L and at least one other arbitrary element of the tuple, the tuple may contain empty objects. In this report most tuples will contain at least the IR component, which represents a compromise solution for a representation that is general enough and structured enough to support formalization.

8.3.1 Action ordering

This class of patterns describes different specifications of ordering between sets of actions. The basic action pattern templates indicate ordering constraints between two actions: sequential, parallel, in any order.

8.3.2 Cyclical actions

Specification of repetitive actions is quite frequent in medical guidelines to be considered as a particular class of control patterns. The action loops are represented by iterations of actions with different parameters, or termination conditions for a cyclical action.

Examples of cyclical actions: monitoring of laboratory parameters is often specified as a cyclical action with unspecified number of repetitions, but with a termination condition attached; administration of drugs is specified as an iteration, associated with a time constraint.

8.3.3 Time constraints for action execution

To indicate the expected order of execution envisioned by the guideline authors, synchronization between actions is necessary, taking into account timeouts and validity of data over time.

The constraint that action A_1 must be terminated within d days before action A_2 can be started, or the fact that the measured value of a clinical parameter (patient's temperature) is valid for a specific duration, are examples of temporal constraints for action execution.

8.3.4 Hierarchical decomposition

This class of patterns contains refinement of an element (a medical action or a medical context) into a set of components.

8.3.5 Selection

Selection of elements from a particular set of elements of the same type is also frequent in clinical guidelines. In fact, the guidelines represent sets of choices for particular time points, in which the practitioner can choose between the most usual/verified solutions to a particular problem.

Most often, the selection includes selection of an action from a set of possible actions. The selection can be: choosing of a particular element, identified by name or satisfying a condition, choosing one random element, choosing a subset of the elements, etc.

8.3.6 Associations condition-action

To indicate how the healthcare process can proceed, the guidelines sometimes indicate the actions that should be executed in a particular medical context.

This construction is frequently found in conclusions and recommendations of the guidelines.

8.3.7 Associations action-goal

The connection between action and the clinical goal it is intended to achieve provides additional information about how the medical contexts covered by the guideline relate to each other.

Ideally, a guideline should contain a clear correlation between the high-level clinical goals the guideline is set to achieve (for instance, have the tumour removed as result of a treatment) and the actions prescribed by the guideline.

8.3.8 Relations between patterns

We need a language for classifying the components of the library of guideline patterns. There are two requirements for this language:

1. it must support understanding by humans (knowledge engineer, guideline author)
2. it must support inferring of pattern properties based on the properties of (simpler) inherited patterns

The relationships between our control patterns can be organized using the principles proposed by Zimmer ([Zimmer2001]) to describe relationships between design patterns. In his approach, at least three relations are important: **X (optionally—mandatorily) uses Y in its solution**; **X is similar to Y**; and **X can be combined with Y**.

We refine the relation $X\ is\ similar\ to\ Y$ with more precise attributes. Given the fact that the control pattern templates accept typed arguments that belong to either one of the medical categories in the ontology of the target domain, or to the taxonomy of linguistic operators (link words that connect medical terms and denote control operations), this relation can be refined into: $X\ is\ covered\ by\ Y$, with the understanding that the vocabulary of X and that of Y have common elements, and that the medical categories and operators of Y are covered by the ones in Y, on a one-to-one mapping.

We use the following relations between (control) pattern templates:

1. **X shares component C with Y**
2. **X extends Y with component C**
3. **X is covered by Y**
4. **X is equivalent to Y in which A is replaced by B**
5. **X is part of Y**

The language used by us for organizing the library and describing relations between pattern templates, pattern instances and text fragments contains the following constructs:

```
ELEMENT has\_type TYPE (an ELEMENT can be of a particular TYPE)
ELEMENT is any class or class instance  from the list:
    ATTR,(ATTR,VAL), FUNCTION, PROPOSITION, FUNCTION(VAL*), FUNCTION*(VAL)
P_INSTANCE1 part_of/used_by P_INSTANCE2
TEXT_FRAGMENT | P_INSTANCE instance_of P_TEMPLATE
P_TEMPLATE1 extends P_TEMPLATE2 with FEATURE*
P_TEMPLATE1 uses_component P_TEMPLATE2
P_TEMPLATE accepts_parameter (ATTR,VAL)
P_TEMPLATE has_constraint/matches_constraint PROPOSITION
P_INSTANCE has_attribute (ATTR,VAL)
P_INSTANCE implements/result_of P_TEMPLATE using (ATTR,VAL)*
P_TEMPLATE defined_as ACTION in CONTEXT |
                      ACTION for GOAL |
                      FUNCTION(P_TEMPLATE1*)
   FUNCTION in {first, next, before, obligatory, optional,
                        followed-by, in-parallel-with}
P_TEMPLATE refined_as FUNCTION1(TEXT*, P_TEMPLATE1) |
                      FUNCTION2*(P_TEMPLATE1, P_TEMPLATE2)
FUNCTION1 in {instantiated-with}; FUNCTION2 ? {applied-to,group-of}
```

Organization of control patterns The library of guideline formalization patterns contains the following attributes for each control pattern:

- NAME: shortcut identifying the pattern and (optionally) describing the category of pattern (i.e., abstract, instance)
- SEMANTIC CATEGORY: list of categorical names to which this pattern can be assigned (based on the axes of a semantic classification system)

- COMPLEXITY: it can be a basic pattern, an aggregation of several patterns, a parameterized component, or a static (fully instantiated) component
- SCOPE: constraint on the context in which the pattern occurs
- GOALS: the intentions that the execution of pattern tries to fulfill
- DESCRIPTION: a more detailed description of the pattern and its relations with other patterns
- REFERENCES: elements imported or re-used from external elements
- BODY: the content of the pattern
- ARGUMENTS: the names and the types/ranges of the variables that are received as input for the pattern
- EXAMPLES: instances of the pattern

One control pattern can belong to more than one categories:

- From the complexity perspective, patterns are split into:
  - Basic patterns: atomic blocks
  - Derived patterns: extending basic patterns with new attributes
  - Complex patterns: combining two or more basic and/or derived patterns
- From the abstraction level perspective:
  - Generic patterns: patterns whose control structure can accommodate patterns, as well as elements from the IR. Example: ”Do ¡action¿” is generic, as ¡action¿ can be replaced with the result of any action pattern, not only with instances of action class.
  - Instance patterns
- From the independence of meaning perspective:
  - Standalone patterns
  - Fragment patterns: cannot appear as a standalone sentences in text, but only as part of another pattern.
- From the conceptual modeling perspective:
  - Action patterns
  - Intention patterns
  - Time patterns

Pattern dependency language Knowing the dependencies between patterns helps developers understand their purpose, and their complexity, and to establish which patterns should be used in a specific situation.

From the perspective of implementation-time execution dependencies, there are only a few dependencies between patterns:

P1:PATTERN goal_dependent_on P2:PATTERN: Pattern P1 has a goal that depends on achieving the goal of pattern instance P2.

P1:PATTERN has_goal G:GOAL common_goal_with P2:PATTERN: Patterns P1 and P2 have the same goal G.

P1:PATTERN resource_dependent_on P2:PATTERN: Pattern P1 is data dependent on pattern P2.

P1:PATTERN task_dependent_on P2:PATTERN: Pattern P1 is execution-subordinated to pattern P2 (parts of P1 can only be activated after P2 has been activated)

P1:PATTERN result_dependent_on P2:PATTERN: Pattern P1 can only complete if and only after P2's execution has finished

8.4 Usage

Example: In the example depicted in figure 19, a composed pattern $Pattern1$ corresponding to the text fragment:

> *"After investigation(+treatment such as mastectomy, conservation treatment), it is recommended to repeat mammography at least once every two years [to detect recurrence of breast tumours]"*

is decomposed into several components, each of them being an instance of a pattern template.

Components C_i represent the building blocks on which more complex control patterns, such as the one instantiated by the text fragment above, can be built. The $implements$ relations between pattern templates $Pattern_j$ convey the following relation between control patterns: pattern template X is equivalent to pattern template Y, but with attributes $A_1, A_2, \ldots, A_n$ bound to the values $V_1, V_2, \ldots, V_n$. For instance, $Pattern_2$ is equivalent to $Pattern_5$, with attribute A1 mapped to value $action_1$ and attribute A2 mapped to value $action_2$.

8.4.1 Examples: translating guideline text using control patterns

A few of the control patterns identified by us can be represented using the programming structure introduced in section 8.2.1.

Example:
Suppose our library of control patterns contains only the following pattern templates:

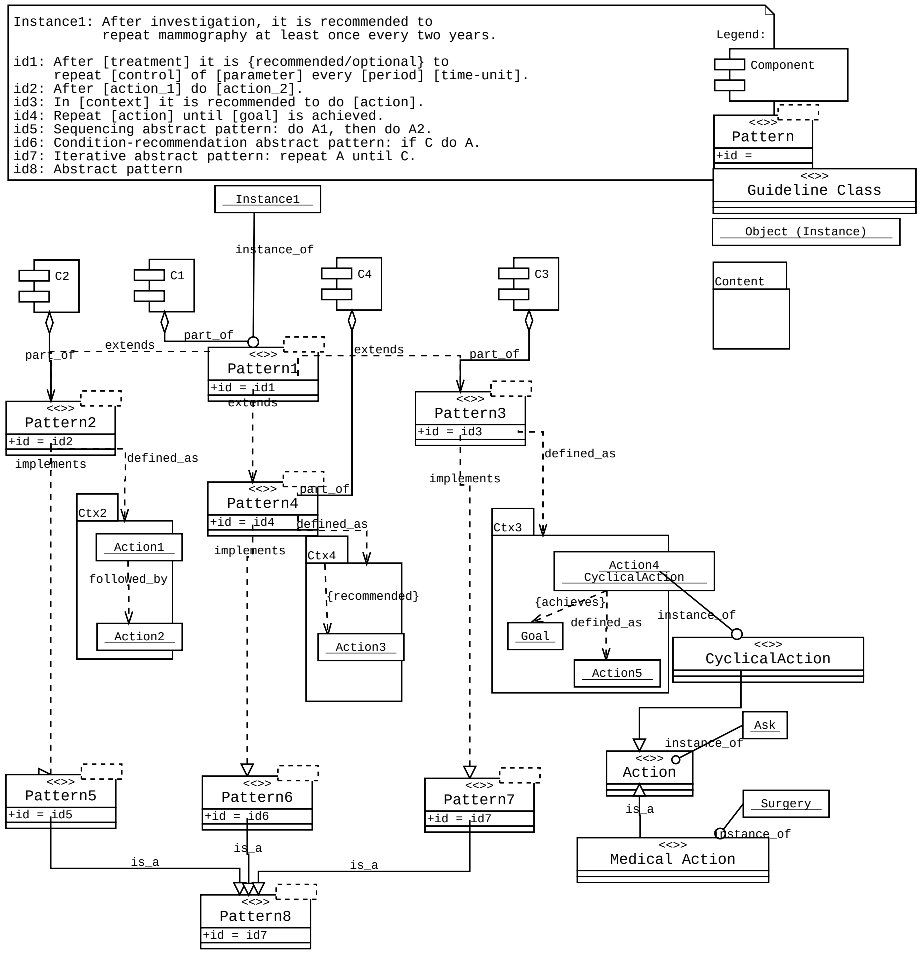

Figure 19: *Relations between the patterns used to represent a complex pattern*

```
mapping(IR: "[Action-A] consists_of [action-1],?,[action-n]",
        FR: definition(action-A,seq(action-1,?,action-n)) ).

mapping(IR: "[Condition-1] defines exclusion status of alternative
             [action-B] and preferred status for [action-1]",
        FR: definition(condition-1,and(disabled-status(action-B)?,preferred-status(action-1)?))).
mapping(IR: "If [condition] then do [action-A] else do [action-B]",
        FR: decision(condition-1?,action-A,action-B) ).
```

The following is the illustration of how a piece of text that instantiates this pattern template is transformed into its formal representation.

Text:

> Modified radical mastectomy (MRM) involves a mastectomy and dissection of the axillary nodes. If BCT is considered to be contra-indicated, and if the patient expresses a preference for mastectomy, MRM is the indicated treatment.

```
IR:
[MRM:med_action] consists-of:decomp-op [mastectomy:action]
                  and:logical-rel-op [dissection:med_action].
[C:condition] defined-as
   [BCT is considered to be contra-indicated:condition]
and:logical-rel-op
   [patient expresses a preference for mastectomy:condition]
If [C:condition] then do [MRM:med_action] else do [BCT:med_action].

FR:
definition(MRM,seq(mastectomy,dissection of the axillary nodes)).
definition(C,and(disabled(BCT),preferred(mastectomy)).
decision(C?,MRM,BCT).
```

In this case, the guideline text has been rewritten to a form that can be mapped into ASBRU.

8.5 Conclusions and lessons learned

Control patterns can be obtained in two ways:

1. by performing a two-step abstraction starting from the text (text-driven method): replacing domain-specific relevant terms with their corresponding domain-specific semantic categories; then replacing domain-independent syntactical terms in the text, which can be mapped to control structures, with their corresponding syntactic-control category.
2. by starting from a programming-like control structure (control-driven method) and performing a one step explicitation of all possible syntactic representations of the "control" words, followed by a validation of the resulting alternatives against the syntactical constructs allowed by a specific domain.

Most of the patterns identified by us, which are listed in the appendix, section C, have been obtained starting from analysis of textual representation of existing guidelines. Nonetheless, due to the abstraction step in which medical terms are replaced by the medical categories they belong to, the control patterns obtained are very generic, and their application not limited only to oncology guidelines. They have a representation in the Intermediate Representation language MHB. Due to the operational nature of most of the control patterns, they also have an ASBRU representation.

Extracting instances of control patterns from the guideline text must be performed within several smaller steps. This is the reason why we introduced linguistic patterns, which are abstractions of domain-specific terms in a domain-specific language of semantic categories. Linguistic patterns are intended to serve as fragments that can be used by control and background patterns.

Even though a direct evaluation of the usefullness of applying the control patterns in guideline formalization has not been performed, many lessons can be learned from validation of linguistic patterns (see section 6.5). Many of the linguistic patterns identified have a strong control component with the control patterns, only are situated at a lower level of granularity.

Control patterns are essential in defining transformation patterns from narrative structures to the formal representation, especially when the "semantics" of the formal model can be described in terms of these control strucures used to define the patterns.

Part VI
Implementation Patterns

9 Implementation Patterns

9.1 Characterization

These are patterns that have been identified by analysing the (intended) behaviour (in terms of the hierarchical decomposition and temporal ordering of actions, the cardinality constraints on them or the granularity of actions) of the semi-formal (Asbru) representation of the protocol.

In addition to the control aspects that are captured by control patterns, the implementation patterns can be characterized using the following Asbru-inspired attributes:

1. whether the plan body consists of one or more sub-plans, or whether it refers to actions to be performed by human intervention: plan-body?single-step, [subplans], user-performed
2. which kind of order of execution is specified for subplans: type?[sequentially], any-order, unordered
3. whether the plan activation needs human confirmation: activate-mode?[automatic], manual
4. whether the successful completion of one or more sub-plans is required for the completion of the plan: wait-for?[all], one, none
5. to some extent, the phase of the medical intervention to which the pattern belong, e.g. diagnosis, treatment, monitoring, post-treatment: context?screening, diagnosis, treatment, monitoring, post-treatment, [undefined]

9.2 Extraction

The implementation patterns have been extracted by knowledge engineers from pre-existing Asbru representations of guidelines, hence the name *Asbru model driven* approach. The patterns we initially aimed at were primarily action patterns, and we assumed they could be either very general (e.g. a description of a particular type of interaction between diagnosis and treatment phases) or more specific to a disease or category of diseases.

Some of the identified patterns are similar to control patterns, and can be viewed as the result of the composition of several instances of control patterns with a clear background knowledge underlying. However, other patterns are rather complex. This is in close connection with the approach of using existing Asbru models, which incorporate complex structures resulting from a top-down design effort to make all the individual guideline actions fit in a hierarchical plan. These structures, needless to say, are often not explicit in the original guideline texts.

9.3 Organization

This section contains a collection of action patterns identified from the Asbru models of the reference protocols (for the diagnosis and management of new-born jaundice and diabetes) used during the assessment phase of Protocure. Consequently, the patterns are concerned with plans (control/action patterns), and have often a high degree of complexity. As mentioned above, this feature is closely related to the top-down design necessary for building a hierarchical Asbru model.

For each pattern, the following information is given:

1. a descriptive name, and a short explanation
2. a reference pointing to the concrete part of the original protocol with which it bears relation
3. a reference to the part (or parts) of the Asbru model where it comes from, and to other parts where it appears(in the same or in another model), if any
4. label indicating its degree of complexity in terms of Asbru encoding, ranging from macro/low-level complexity to or high-level complexity
5. a reference to the pattern (or patterns) with which it interacts, if this is the case[3]
6. an excerpt of the Asbru model corresponding to the part where the pattern comes from, which can be considered as an instance thereof

9.3.1 Action along with a continuous monitoring of critical conditions requiring the interruption of the process

It consists in doing something and checking for one or several critical conditions in the meantime. If a particular critical condition occurs, the overall process should be interrupted.
Appears in model: Jaundice, Hyperbilirubinemia plan.
Related to protocol part: No explicit relation.
Also appears in model: –
Complexity: high-level.
Interactions/restrictions: Interacts with pattern 9.3.2.
Asbru excerpt:
plan Hyperbilirubinemia

. . .

conditions

abort-condition: (possibility-of-hemolytic-disease = yes)

plan-body type = unordered, wait-for-optional-subplans = yes

wait-for Diagnostics-and-Treatment-hyperbilirubinemia

Check-for-rapid-TSB-increase

Check-for-jaundice-after-2-weeks

Check-for-jaundice-after-3-weeks

Diagnostics-and-Treatment-hyperbilirubinemia

9.3.2 Monitoring of critical conditions

An action triggered whenever a critical condition deserving attention occurs, with the purpose of performing some management actions. In this pattern, slight variations are possible, e.g. depending on the relations between the monitoring subplan and the plan.

[3] Usually, a pattern of high complexity will interact with some other pattern.

Appears in model: Jaundice, Check-for-jaundice-after-2-weeks.
Related to protocol part: Somehow derived from box 22 of algorithm.
Also appears in model: In the same model: Check-for-jaundice-after-3-weeks and Check-for-rapid-TSB-increase plans.
Complexity: medium-level.
Interactions/restrictions: Interacts with pattern 9.3.1.
Asbru excerpt:
plan Check-for-jaundice-after-2-weeks

. . .

conditions

filter-precondition: (jaundice-clinically-significant = yes) *in* ([-, 2w] [2w, -] [-, -] birth-date)

plan-body type = sequentially

wait-for all

do type = any-order

wait-for all

ask physical-exam-OK
ask colour-stools
ask colour-urine

if (or (physical-exam-OK = no) (colour-stools = ”light”) (colour-urine = ”dark”)**)**

then

ask direct-serum-bilirubin
possibility-of-cholestatic-disease ← yes
Exit-possibility-of-cholestatic-jaundice

else

possibility-of-cholestatic-disease ← no
Exit-provide-routine-care

9.3.3 Sequence of actions, none of them must succeed necessarily

Do several things in the indicated order, being the completion of the plan dependent on conditions different from the completion of subplans. As a matter of fact, it could be the case that none of the subplans completes, because of the different complete and abort conditions of both the plan and the subplans. Using this alternative requires a careful analysis of the interactions between the subplans and the plan, in order to adjust the resulting behaviour.
Appears in model: Jaundice, Diagnostics-and-Treatment-hyperbilirubinemia plan.
Related to protocol part: No explicit relation.
Also appears in model: –
Complexity: low-level.
Asbru excerpt:

plan Diagnostics-and-Treatment-hyperbilirubinemia

. . .

conditions

complete-condition: (or (jaundice-clinically-signiﬁ cant = no) *explanation* "Follow this infant into routine clinical supervision."
completed(Treatment-hyperbilirubinemia) *in* NOW**)**

abort-condition: (or (term-child = no) *explanation* "Exiting the protocol to individualized clinical evaluation, including assessment of jaundice in light of prematurity."
(age = day1) *explanation* "Exiting the protocol to individualized clinical evaluation, including assessment of jaundice and non-isoimmune hemolytic disease."
(pathologic-reason = yes)**)**

plan-body type = sequentially

wait-for none

ask term-child

ask age-child

Diagnostics-hyperbilirubinemia

Treatment-hyperbilirubinemia

9.3.4 Sequence of actions, all of them must succeed

Do several things, all of them, in the indicated order. It is a very common structure, which appears e.g. when anamnesis questions must be followed by a decision.
Appears in model: Jaundice, Diagnostics-hyperbilirubinemia plan.
Related to protocol part: Somehow derived from boxes 2 to 18.
Also appears in model: Many plans.
Complexity: low-level.
Asbru excerpt:
plan Diagnostics-hyperbilirubinemia

. . .

plan-body type = sequentially

wait-for all

pathologic-reason ← no

Anamnesis-abnormal-signs

Blood-tests

Anamnesis-hemolytic-disease

Jaundice-determination

9.3.5 Unspecified-order set of actions, all of them must succeed

Do several things, all of them, one after the other but with unspeciﬁ ed order. It is also a very common control structure that has been mostly used to model anamnesis questions, in which the ordering can be assumed to be irrelevant.

Appears in model: Jaundice, Anamnesis-abnormal-signs plan.
Related to protocol part: Somehow in box 2.
Also appears in model: Many plans.
Complexity: low-level.
Asbru excerpt:
plan Anamnesis-abnormal-signs

. . .

do type = any-order

wait-for all

ask lethargy
ask apnea
ask tachypnea
ask instable-temperature
ask behavior-changes
ask hepatosplenomegaly
ask vomitting
ask feeding-diffi culty
ask excessive-weight-loss

. . .

9.3.6 Sequence of actions, only applicable ones, one action must succeed

Do all applicable things, in the indicated order. At least one of them should complete so that the overall plan is considered successful. On the other hand, all the applicable actions are given a chance to complete, i.e. the plan does not not reach the completed state if there are still applicable subplans running.
Appears in model: Jaundice, Blood-tests plan.
Related to protocol part: Somehow in boxes 6 to 12.
Also appears in model: –
Complexity: medium-level.
Interactions/restrictions: The fi lter preconditions of at least one of the subplans must be satisfi able in every case.
Asbru excerpt:
plan Blood-tests

. . .

plan-body type = sequentially, wait-for-optional-subplans = yes

wait-for (or Check-blood-test-mother Perform-blood-test-child**)**

Check-blood-test-mother

Perform-blood-test-child

9.3.7 Choice of an action among a list by the user

Perform one action from a predefi ned list of alternatives, depending on the user choice. Again, this pattern constitutes a very frequent control structure.

Appears in model: Jaundice, Jaundice-determination plan.
Related to protocol part: Methods mentioned in point 6 of Evaluation.
Also appears in model: In the same model: Choose-feeding-alternative and Phototherapy-normal-recommendation plans. In diabetes one: Glucosedetermination and Insulin-with-or-without-antidiabetics-treatment plans.
Complexity: medium-level.
Interactions/restrictions: The activation mode of the sub-plans must be manual, to implement the user selection.
Asbru excerpt:
plan Jaundice-determination

...

do type = any-order

wait-for one

Blanching-skin-with-digital-pressure-test

Icterometer-test

Transcutaneous-jaundice-meter-test

Determine-extent-cephalocaudad-progression

...

9.3.8 Action along with periodic tests and/or controls

Do several things in parallel, basically the relevant management actions along with some periodic tests and/or controls.
Appears in model: Jaundice, Treatment-hyperbilirubinemia plan.
Related to protocol part: No explicit relation in the case of jaundice protocol. In the diabetes case there are separate parts for the quartely and annual controls.
Also appears in model: In diabetes model, within Treatments-and-Controls plan.
Complexity: high-level.
Interactions/restrictions: The termination of the parallel block must be ensured.
Asbru excerpt:

plan Treatment-hyperbilirubinemia

. . .

plan-body type = parallel

wait-for one

do type = any-order

wait-for (or Regular-treatments Exchange-transfusion**)**

Regular-treatments **on-abort** Exchange-transfusion
Exchange-transfusion

cyclical-plan

do type = sequentially
wait-for all
ask TSB-value
ask age-child

retry-delay: min = 12h, max = 24h

9.3.9 Alternative actions, one action must succeed

Try several actions from a list, one at a time and depending on their applicability, until one succeeds. Normally one of the actions should suffice to perform the task.
Appears in model: Jaundice, first nested block within Treatment-hyperbilirubinemia plan.
Related to protocol part: No explicit relation.
Also appears in model: –
Complexity: medium-level.
Interactions/restrictions: The filter preconditions of at least one of the subplans must be satisfiable in every case.
Asbru excerpt: See excerpt of pattern 9.3.8.

9.3.10 Sequence of two trials

Try an action and, if it does not succeed, try a second one.
Appears in model: Jaundice, first part of first nested block within Treatmenthyperbilirubinemia plan.
Related to protocol part: No explicit relation in the case of jaundice protocol. In the diabetes case there is explicit sentence in the insulin treatment part: "Start with insulin in case of insufficient regulation with oral drugs".
Also appears in model: In diabetes model, treatment part within Treatments-and-Controls plan.
Complexity: low-level.
Asbru excerpt: See excerpt of pattern 9.3.8.

9.3.11 Series of trials, possibly more than once, a particular action must succeed

Try several actions, one at a time and depending on their applicability, until a particular action succeeds. It is possible trying an action more than once.
Appears in model: Jaundice, main part of Regular-treatments plan.

Related to protocol part: No explicit relation.
Also appears in model: –
Complexity: high-level.
Interactions/restrictions: Since some actions can be tried more than once, the termination of the main block must be expressed in terms of the termination of the action that must succeed.
Asbru excerpt:
plan Regular-treatments

. . .

do type = any-order, retry-aborted-subplans = yes

wait-for Observation

Phototherapy-intensive
Phototherapy-normal-prescription
Phototherapy-normal-recommendation
Observation

9.3.12 Sequence of actions, only applicable ones, a particular action must succeed

Do all applicable things, in the indicated order. At least a particular action should succeed.
Appears in model: Diabetes, nested block within Anamnesis plan.
Related to protocol part: Part within the diagnostic to evaluate if blood glucose determination is necessary.
Also appears in model: –
Complexity: low-level.
Asbru excerpt:
plan Anamnesis

. . .

do type = sequentially, wait-for-optional-subplans = yes

wait-for Anamnesis-typical-signs

Anamnesis-typical-signs
Anamnesis-olderthan-45

if (or (typical-signs = true) (risk-factors = true)**)**

then

glucose-determination-needed ← true

else

glucose-determination-needed ← false

9.3.13 Action along with a continuous monitoring of critical conditions requiring treatment

Do something during the whole life of the patient, and react to critical situations whenever they occur, possibly more than once.
Appears in model: Diabetes, nested block within Policy plan.

Related to protocol part: Somehow, since there are separate sections for the management of concurrent diseases, hypoglycemic comma and consultations and referrals.
Also appears in model: –
Complexity: high-level.
Interactions/restrictions: Interacts with pattern 9.3.20.
Asbru excerpt:
plan Policy

...

do type = unordered, retry-aborted-subplans = yes

wait-for Treatments-and-Controls

Treatments-and-Controls
Policy-for-concurrent-diseases
Policy-for-hypoglycemic-coma
Policy-for-consultation
Policy-for-chiropodist-referral
Policy-for-nurse-referral

9.3.14 Sequence of trials, at least the first action must succeed

Try several actions, in the indicated order and possibly with time delays in between some of them. At least the first action must succeed.
Appears in model: Diabetes, Non-insulin-DMT2-treatments plan.
Related to protocol part: Sequence of steps in the initial treatment part.
Also appears in model: –
Complexity: high-level.
Interactions/restrictions: Since only the first action must necessarily complete, the success/failure must be determined in terms of some parameter value after the final trial.
Asbru excerpt:
plan Non-insulin-DMT2-treatments

...

conditions

abort-condition: (and (glucose-monitoring = bad) completed(SU-derivative-plus-metformin-treatment) *in* NOW**)**

plan-body type = sequentially, wait-for-optional-subplans = yes

wait-for Diet-specialist-referral

Diet-specialist-referral

Fasting-glucose-test *in* ([11w, 13w] [-, -] [-, -] plan-state-transition(leave, first, activated(Diet-specialist-referral)))

SU-derivative-or-metformin-treatment

SU-derivative-plus-metformin-treatment

9.3.15 Search for treatment

Set an initial treatment (drugs and doses) and adjust it until the goal is achieved.
Appears in model: Diabetes, SU-derivative-or-metformin-treatment plan.
Related to protocol part: Step 2 in the initial treatment part.
Also appears in model: In the same model: Microalbuminuria-treatment plan. Very similar: Only-insulin-treatment plan.
Complexity: medium-level.
Interactions/restrictions: Interacts with pattern 9.3.16.
Asbru excerpt:
plan SU-derivative-or-metformin-treatment

...

plan-body type = sequentially

wait-for all

DMT2-treatment ← antidiabetics

if (quetelet-index $\leq$ 27)

then

ask drug-name

put-last(drug-name, antidiabetics**)**

else

put-last(metformin, antidiabetics**)**

antidiabetic-doses ← Initialise-drug-doses(antidiabetics, iterator-antidiabetics, antidiabetic-doses, iterator-antidiabetic-doses)

antidiabetic-doses
← Find-antidiabetic-doses(antidiabetics, iterator-antidiabetics, antidiabetic-doses, iterator-antidiabetic-doses)

9.3.16 Periodic adjustment of treatment

In relation to a treatment, increase the drug doses until the goal has been achieved or the maximal doses have been attained.
Appears in model: Diabetes, Find-antidiabetic-doses plan.
Related to protocol part: Details of step 2 in the initial treatment part.
Also appears in model: In the same model, although with a different periodicity: Find-ACE-inhibitor-doses plan. Very similar: Find-evening-insulin-dose, Find-morning-insulin-dose, and Find-insulin-doses plans.
Complexity: medium-level.
Interactions/restrictions: Used in pattern 9.3.15, and also in pattern 9.3.18. The suspend and reactivate conditions are relevant only in the case of the latter. Finally, it interacts with pattern 9.3.17.
Asbru excerpt:

plan Find-antidiabetic-doses

. . .

conditions

suspend-condition: **(and** `antidiabetic-problems` activated(Check-for-antidiabetic-problems) *in* NOW**)**

reactivate-condition: completed(Check-for-antidiabetic-problems) *in* NOW

. . .

plan-body

cyclical-plan

drug-doses, antidiabetic-maximal-doses ← Increase-drug-doses(drugs, iterator-drugs, drug-doses, iterator-drug-doses)

retry-delay: min = 2w, max = 4w

returns drug-doses

9.3.17 Increase of doses of a drug/list of drugs

Appears in model: Diabetes, Increase-drug-doses plan.
Related to protocol part: Standard drug increase procedure in the initial treatment part.
Also appears in model: –
Complexity: low-level.
Interactions/restrictions: Interacts with pattern 9.3.16.
Asbru excerpt:
plan Increase-drug-doses

. . .

plan-body type = sequentially

wait-for all

ask maximal-drug-doses

reset-iterator(iterator-drugs**)**

reset-iterator(iterator-drug-doses**)**

do-repeatedly

ask drug-dose-increase
put-last((+ drug-dose-increase **get(**iterator-drug-doses**)**), aux-doses**)**
go-to-next(iterator-drugs**)**
go-to-next(iterator-drug-doses**)**

termination-condition: is-at-end(iterator-drugs**)**

returns aux-doses, maximal-drug-doses

9.3.18 Search for treatment, with management of treatment problems

Set initial treatment (drugs and doses) and adjust it until the goal is achieved, checking for possible treatment problems meanwhile. In case a problem is detected, a management procedure is triggered.
Appears in model: Diabetes, SU-derivative-plus-metformin-treatment plan.
Related to protocol part: Details of step 3 in the initial treatment part.
Also appears in model: –
Complexity: high-level.
Interactions/restrictions: Interacts with patterns 9.3.16 and 9.3.19.
Asbru excerpt:
plan SU-derivative-plus-metformin-treatment

. . .

plan-body type = sequentially

wait-for all

ask drug-name

put-last(drug-name, antidiabetics**)**

put-last(metformin, antidiabetics**)**

antidiabetic-doses ← Initialise-drug-doses(antidiabetics, iterator-antidiabetics, antidiabetic-doses, iterator-antidiabetic-doses)

do type = unordered

wait-for Find-antidiabetic-doses

antidiabetic-doses ← Find-antidiabetic-doses(antidiabetics, iterator-antidiabetics, antidiabetic-doses, iterator-antidiabetic-doses)

antidiabetics, antidiabetic-doses ← Check-for-antidiabetic-problems(antidiabetics, iterator-antidiabetics, antidiabetic-doses, iterator-antidiabetic-doses)

9.3.19 Management of treatment problems

Simple management of treatment problems, consisting in the substitution of the problematic drug.
Appears in model: Diabetes, Check-for-antidiabetic-problems plan.
Related to protocol part: Again, details of step 3 in the initial treatment part.
Also appears in model: –
Complexity: high-level.
Interactions/restrictions: Interacts with pattern 9.3.18.
Asbru excerpt:

plan Check-for-antidiabetic-problems

...

conditions

filter-precondition: `antidiabetic-problems:` **(or** (contraindications = true) (side-effects = true)**)**

plan-body type = sequentially

wait-for all

ask problematic-antidiabetic

reset-iterator(iterator-drugs**)**

reset-iterator(iterator-drug-doses**)**

do-repeatedly

go-to-next(iterator-drugs**)**
go-to-next(iterator-drug-doses**)**

termination-condition: **(or** (**get(**iterator-drugs**)** = problematic-antidiabetic) **is-at-end(**iterator-drugs**))**

remove-at-iterator(iterator-drugs**)**

remove-at-iterator(iterator-drug-doses**)**

put-last(acarbose, drugs**)**

ask drug-dose

put-last(drug-dose, drug-doses**)**

returns drugs, drug-doses

9.3.20 Management of concurrent situations

Actions triggered whenever a concurrent situation deserving attention occurs. Once the situation is over, the action completes and the process continues.
Appears in model: Diabetes, Policy-for-concurrent-diseases.
Related to protocol part: Section for the management of concurrent diseases.
Also appears in model: In the same model: Policy-for-hypoglycemic-comma, Policy-for-consultation and Policy-for-chiropodist-referral plans.
Complexity: medium-level.
Interactions/restrictions: Interacts with pattern 9.3.13.
Asbru excerpt:

plan Policy-for-concurrent-diseases

...

conditions

filter-precondition: `concurrent-diseases:` **(or** (fever = true) (vomiting = true) (diarrhea = true)**)**

abort-condition: (concurrent-diseases-over = true)

plan-body type = sequentially

wait-for all

Extra-fluid-intake-prescription

Glucose-lowering-therapy-advice

ask concurrent-diseases-over

9.4 Usage

The intended use of the Asbru (implementation) patterns described in this section is serving as a basis for the construction of new Asbru models. Our claim is that these patterns can be of great help in the advanced stages of Asbru modelling, during which knowledge engineers have to devise more or less complex plans that represent in a proper way the procedural knowledge contained in the guideline, and structure them in higher-level plans. This is in connection with the fact that implementation patterns come from existing Asbru models, as opposed to e.g. linguistic or control patterns, which have been derived from guideline texts (and hence are far away from Asbru). With the purpose of determining upto which point our implementation patterns are reusable, we have carried out a small evaluation exercise which is presented next.

9.5 Evaluation

The evaluation of the implementation patterns has consisted in modelling in Asbru a subset of the chapters of CBO's breast cancer guideline, using the list of patterns presented in this section as support, and subsequently studying the degree of reusability (whether they are used or not, and the frequency of reuse) of the different patterns. The part of the guideline we have concentrated on comprises chapters 2, 3 and 4, dedicated respectively to the adjuvant therapy for operable breast cancer, the treatment of locally advanced breast cancer, and the follow-up. Lastly, it is important to know that we have used as starting point for the Asbru modelling the MHB mark-up of the corresponding chapters, rather that the guideline text alone. This means we have exploited the MHB descriptions of the procedural knowledge of the chapters, including the hierarchical draft structure that the (MHB) cross-reference tool produces.

The number of plans that we have studied amounts to 52, which includes the 24 plans of chapter 2 (4 of them user-performed), 12 plans of chapter 3 (4 user-performed), and 16 plans of chapter 4 (6 user-performed). Only two different patterns have been reused in those plans, namely pattern 9.3.6 (sequence of actions, all the applicable ones, with one or more mandatory actions) and, more often often, pattern 9.3.7 (choice of an action among a list). Although the reuse percentage is significant –actual reuse in 12 out of the 38 non user-performed plans, i.e. the 31% of cases– neither the variety nor the complexity of the applied patterns is high. One reason to explain the small number of used patterns might be their specificity with respect to jaundice and/or diabetes protocols. For example, (diabetes-derived) patterns for searching for the right doses of medication seem to have limited possibilities of reuse in the framework of breast cancer management. Concerning the low complexity of used patterns, it could be explained by the fact that the modelling has been restricted to individual chapters dealing with specific parts of the treatment, which makes the reusability of e.g. high-level coordination patterns less likely.

Regardless of the above, it is important to recall that the reuse of implementation patterns facilitates the modelling task to a great extent. As a matter of fact, it has allowed us to directly apply a few existing Asbru solutions to standard modelling problems. Related to this, we have found new problems/situations which had not been modelled in the past and hence required much more modelling effort. An example is the group of activities related to the follow-up of patients within chapter 4. Once we will have devised (and validated) suitable Asbru solutions for them, we will incorporate them as new Asbru (implementation) patterns that will facilitate the future modelling of similar aspects.

9.6 Conclusions

In this section we have described a number of action patterns identified from an analysis of the Asbru models of jaundice and diabetes protocols, and evaluated their utility by means of a modelling exercise based on chapters 2 to 4 of CBO's breast cancer guideline. In the following paragraphs we present some conclusions and lessons derived from both the process of pattern identification and the study of the (re)usability of patterns.

9.6.1 Conclusions on pattern identification

Although the analysis has been performed long after the modelling activity, it is important to note that it was heavily influenced by the past modelling experiences. Regardless of this observation, a number of lessons can be drawn from our experiences in scanning Asbru models searching for reusable guideline patterns. Firstly, the two Asbru models contain a high number of action patterns with potential for reuse. Indeed, already in this stage of pattern identification we observed that some of these patterns had been actually reused several times in the two models.

Concerning the nature/kind of patterns, as it has been mentioned before, they mainly fall in the category of action patterns. However, we can distinguish different types of patterns within this general category, at least:

- Patterns for more or less simple action sequencing. Examples here include sequences with none or all as waiting strategy (i.e. patterns 9.3.3 and 9.3.4).
- Patterns for more complex action sequencing, such as a series of possibly repeated trials (pattern 9.3.11) or an ordered sequence of trials (pattern 9.3.14).
- Patterns implementing the search for the right dosage of medication, mainly the patterns 9.3.15 to 9.3.19.
- Patterns for high-level coordination of actions, e.g. policy with monitoring of conditions and/or periodic checks. As examples we can cite patterns 9.3.1, 9.3.8 and 9.3.13, and their lower-level counterparts.

Notice that the types of patterns in the above list are listed in an increasing order of complexity. Another important observation we already mentioned is that the patterns obtained from exisiting Asbru models can be rather complex, because of the design efforts of the modeller.

In summary, existing guideline and protocol models contain a wealth of structures with potentials for reuse. At present we have restricted our analysis to control structures, resulting in a collection of action patterns of several types. In addition to action patterns, we believe that the same strategy can be applied to other parts of the Asbru model, with the purpose of obtaining patterns covering features different from plan control.

9.6.2 Conclusions on pattern reusability

Although the utility of some of the patterns was already clear during pattern identification stage, as we have mentioned before, we have carried out an small-scale evaluation exercise based on some chapters of the evidence-based breast cancer guideline developed by CBO. The results of this exercise suggest that reuse in modelling can be achieved to a significant extent, even considering a relatively small number of implementation patterns.

Finally, we have also shown that reuse is possible accross different types of guideline and protocol documents. As a matter of fact, the patterns extracted from the models of jaundice and diabetes (non evidence-based) protocols have been reused to model parts of the evidence-based breast cancer guideline. This is an important result in itself, since the recommendations of evidence-based guidelines seemed to be completely different in nature.

Part VII
Concluding remarks

10 Main Conclusions

The library of guideline patterns aims at:

1. facilitating the understanding of the formalization process, by splitting it into smaller, more manageable steps; according to this *knowledge-driven refinement process*, the guideline formalization process can be described in terms of the classes of patterns described so far, as follows:

   Algorithm 4 (Guideline formalization steps using a library of guideline patterns) *(a) Analyze guideline document and extract (manually or automatically) the document structure based on the narrative elements and relations between them.*

   (b) Analyze guideline document at the sentence level and extract sentences that contain frequent item sets from:
   1. the ontology of the medical domain;
   2. the ontology of programming control structures.

   (c) Identify, among the text fragments that contain frequent item sets from both medical and control domain, sequences of elements that are frequently combined.

   (d) Establish a mapping between narrative frequent items and control+medical frequent item sets identified above.

   (e) Establish a mapping between control item sets and their formal (ASBRU equivalent) translation.

   (f) Transform the narrative instances in the document into an ASBRU equivalent, using, whenever possible, the mappings above.

2. facilitating a more effective verifi cation and maintenance of a ”living guideline”, by providing a translation of an IR construct into its formal equivalent. So far, we have validated some of the guideline patterns through the use of the Stepper and GMT tools, but the integration of the library with any of these tools has not been accomplished yet.

The report addresses several topics and draws conclusions about:

1. the source for patterns: the text-driven approach searches for patterns in the text of guidelines from the same medical domain, primarily oncology guidelines; the implementation-driven approach starts from existing formalization of guidelines.
2. the language for describing patterns: for many of the patterns we can use the intermediate language MHB. However, due to its nature of being closer to a textual representation, it is not suffi cient, or too abstract, to capture fi ner-grained control structures and background relations between medical concepts.
3. the organisation of the library: we have studied several classes of patterns and elaborated relations between them. It is clear that different types of patterns cover different parts of the guideline formalization process.

   From the perspective of guideline execution, several types of patterns can be distinguished:

(a) patterns defining important concepts and actions
(b) patterns describing decomposition of actions
(c) patterns describing restrictions between actions
(d) patterns describing intended effects of actions
(e) patterns describing exceptional situations
(f) patterns describing user interaction and user responsibilities

On the other hand, it seems natural to categorize patterns based on the category of medical actions they belong to. This would make a medical expert more able to retrieve and select the right candidate pattern for a situation he/she wants to describe when authoring the guideline. For this goal, we can re-use the semantic relations between medical terms used by existing medical thesauri, such as MeSH ([MeSa]), UMLS ([UML]), or NCIOncology ([NCI]).

4. retrieval of relevant parts of the guideline based on keyword search can represent a possible use of the library. This goal has been explored in existing medical projects, such as DeGeL2 project ([DeG]), in which the guidelines are classified, at a very abstract level, according to the category of activities that are most dominant in the guideline: **Counseling; Prevention; Evaluation; Screening; Diagnosis; Treatment; Management**.

   The medical concepts present in the guidelines are categorized based on the types of medical concepts they express: **Disorders; Laboratory Parameters; Symptoms and Signs; Treatment Methods; Body Parts; Diagnostic Findings**.

   For each clinical guideline these medical concept categories can be associated with medical terms specific to the domain addressed by the guideline, using background medical knowledge. This knowledge is provided by the group of medical experts that write the guideline, but it is likely that the retrieval of these concepts can take place automatically from medical thesauri, such as MeSH ([MeSa]), UMLS ([UML]), or NCIOncology ([NCI]).

   These mappings between key medical concepts and a range of possible values that they can take within the medical domain are used as criteria for search and retrieval. For instance, in a breast cancer guideline, one could search whether the guideline contains references to:

   - Treatment_method in mastectomy, radiotheraphy
   - Disorder contains "tumour"
   - Diagnostic not in "invasive breast cancer", "operable breast cancer"

5. the use of the library: for some examples we illustrate how one text fragment that instantiates a pattern can be translated into its semi-formal ASBRU equivalent. However, not all patterns have a textual representation, and not all have an ASBRU representation either. Nonetheless, they are useful in various phases of the guideline formalization process.

6. automatic identification of patterns: we have built an application that uses an ontology to generate abstract linguistic patterns, and to search for instances of these patterns within the guideline text. With the help of a medical expert, these can further be refined into more specific patterns or composed with other patterns to built more complex patterns.

10.1 Lessons learned

1. As shown in the appendix, there is no clear border between the linguistic patterns, background patterns, control patterns and implementation patterns. It happens frequently that the same text fragment represents an instantiation of several types of pattern templates. A possible guideline formalization path includes the following use of the classes of patterns identified:

   (a) first, the application of structural patterns, to establish a mapping between components in the guideline text and groups of plans in ASBRU;

   (b) then, for each of the text structures produced, linguistic patterns are applied, to refine the content of the ASBRU sets of plans generated in the first phase;

   (c) not for all linguistic patterns instances found an ASBRU component can be generated; however, MHB components can also be generated by application of linguistic patterns, which can be used as input by control and background patterns;

   (d) by applying implementation patterns, some of the resulting MHB and ASBRU components can be combined into more detailed ASBRU components.

2. The mappings between background patterns and linguistic patterns are difficult to build automatically. The number of meaningless combinations generated is too large and the number of meaningful (medically relevant) pattern templates too small to consider automatic generation of pattern templates for linguistic background patterns successful.

3. Not all control patterns have a formal ASBRU equivalent. From background pattern templates, much fewer can generate an ASBRU component. Therefore, the use of patterns seems to revolve around the quality of the Intermediate Representation that is used before generating ASBRU.

4. **Interaction with the MHB Intermediate Representation Language**

   Our task was closely connected with task 2.2 - "Specification of formats of intermediate, ASBRU and KIV representations"-, which produced the Intermediate Representation (IR) language *MHB* ([Seyfang *et al.*2004]).

   A large part of the patterns identified (from all categories) can be described using the Intermediate Representation language MHB ([Seyfang *et al.*2004]). In fact, especially the search for control patterns and the development of the IR language have influenced each other: on one hand, the IR has evolved indirectly based on the type of frequently-used knowledge and control constructs. On the other hand, what was considered to represent a control pattern has been represented in such a way to allow for a possible representation in IR for that pattern.

   However, there are a few implementation and control patterns that are not based on textual representation, but rather on concepts that cannot be represented in MHB with sufficient precision to allow for their formalization. One of the reasons is the complexity of the patterns, which include correlations between the intentional aspect (what are the intentions of a particular medical action), the knowledge aspect (why is this medical action present in the current context), and the control aspect (how this medical action is supposed to take place), which challenge the limited expressivity of the ASBRU language for modelling non-procedural knowledge.

5. The organization of the library is dependent of the goal for which the patterns are used: a. authoring; b. formalization.

6. The low-level granularity of linguistic pattern templates and the higher support for instances of background pattern templates in comparison to control patterns, makes them rather difficult to use in the last steps of the formalization, but more appropriate in the starting step.
7. We considered control patterns as representations of control structures - iterations, synchronization of actions, selection, etc - at the level of an Intermediate Representation that is enriched with domain-specific categories. This distinguishes them from implementation patterns, which are representations of an Intermediate Representation enriched with a higher-level intentional component and "aware" of the target domain.

   However, the control patterns and the implementation patterns seem to be the most useful for guideline formalization, as they both can be mapped to abstract programming-like control structures and therefore a mapping between them can be established. Control patterns are closer to an abstract, but text-centered, Intermediate Representation, while implementation patterns are closer to an abstract Semi-formal Representation of the guideline.
8. Extracting instances of control patterns from the guideline text is easier than extracting implementation patterns, but must be performed within several smaller steps.

   We introduced linguistic patterns as simplifications of control patterns, as they represent abstractions of domain-specific terms in a domain-specific language of semantic categories. Linguistic pattern instances are the building blocks for control and background patterns.

10.2 Future directions for research

Besides integration of the patterns in the existing tools for guideline formalization (GMT), the role of pattern templates in testing and validating guideline components will be further investigated as other tasks of the project which can use the patterns are started.

Based on the instances of action pattern templates (linguistic, control, implementation) that have a higher complexity, combinations of actions can be generated, to aid in testing which sets of plans are activated/executed most often, whether all of them are reachable, etc.

Visualization of the knowledge captured by the guideline can also be improved using the support provided by the linguistic pattern search program.

Validation of patterns. The effort spent in the first step of the formalization process - semantic mark-up of the text, can be reduced considerably if MHB chunks or ASBRU plan schemata are generated from the instances of patterns that are recognized in the text.

In guideline authoring, the library of patterns can be used to generate as a typed, parameterized formatting template.

At a minimum, the vocabulary collected for linguistic patterns can be used as a means to validate the linguistic patterns.

By validation, we understand at least two types of operations:

validation of representation patterns by medical experts and knowledge engineers : checking whether the identified representation patterns actually capture clinically relevant information and meaningful formalization steps; this includes validation of linguistic patterns by visually inspecting the marked up guideline text and deciding whether it contains the pattern found automatically;

cross-validation of linguistic patterns with respect to control and implementation patterns : checking whether there are overlappings between the background, control and implementation pattern instances on one hand and linguistic pattern instances on the other hand.

The first type of validation has been performed with the help of the GATE tool ([GAT]). We generated several lists with medical terms, each belonging to different semantic categories and marked them up with different colours in the text of the guideline. We tried to answer the question: Are there pieces of knowledge that actually represent background, control and implementation knowledge, or these are rather disjunctive categories? Can we use all of these categories of patterns in our guideline formalization effort?

References

[A.2002] Shahar Y; Young O.; Shalom E.; Galperin M.; Mayaffit A.; Moskovitch R.; Hessing A. A framework for a distributed, hybrid, multiple-ontology clinical-guideline library and automated guideline-support tools. Technical report, Ben-Gurion University, 2002. URL: http://www.ise.bgu.ac.il/artic/artfiles/318/DEGEL.JBI.resubmit.2.pdf.

[Appleton1997] B. Appleton. Patterns and software: Essential concepts and terminology, 1997.

[Atkinson] Irving Atkinson. The graphical depiction of design patterns.

[Beck *et al.*1996] Kent Beck, James O. Coplien, Ron Crocker, Lutz Dominick, Gerard Meszaros, Frances Paulisch, and John Vlissides. Industrial experience with design patterns. In *Proceedings of the 18th International Conference on Software Engineering*, pages 103–114. IEEE Computer Society Press, 1996.

[Ben-Yehuda] Shai Ben-Yehuda. Pattern language for framework construction.

[Budinsky *et al.*1996] Frank J. Budinsky, Marilyn A. Finnie, John M. Vlissides, and Patsy S. Yu. Automatic code generation from design patterns. *IBM Systems Journal*, 35(2):151–171, 1996.

[Bunnig *et al.*] Stefan Bunnig, Peter Forbrig, Ralf Lammel, and Normen Seemann. Design pattern oriented programming.

[CBO2002] CBO. *Guideline for the Treatment of Breast Carcinoma*. CBO, 2002. PMID: 12474555.

[Coplien1998] James O. Coplien. Software design patterns: Common questions and answers, 1998.

[DeG] Degel project (digital electronic guideline library).

[Eden *et al.*1997] Amnon H. Eden, Amiram Yehudai, and Joseph Gil. Precise specification and automatic application of design patterns. In *Automated Software Engineering*, pages 143–152, 1997.

[et al2004] Andreas Seyfang et al. Specification of formats of intermediate, asbru and kiv representations. Technical Report Deliverable D2.2a, Protocure II Project, Viena University of Technology, Austria; Dutch Institute for Healthcare Improvement, The Netherlands; University of Augsburg, Germany; University of Nijmegen, The Netherlands, July 1 2004. URL: http://www.keg.uji.es/protocure/D22a-specification-of-formats.pdf.

[Fernandez and Yuan2000] E.B. Fernandez and X. Yuan. Semantic analysis patterns. In *Procs. 19th Int. Conf. on Conceptual Modeling, ER2000, Salt Lake City, USA*, pages 183–195. Florida Atlantic University, October 2000.

[Florijn *et al.*1997] Gert Florijn, Marco Meijers, and Pieter van Winsen. Tool support for object-oriented patterns. *Lecture Notes in Computer Science*, 1241:472–??, 1997.

[Forbrig and Lammel] P. Forbrig and R. Lammel. Programming with patterns.

[Fowler1997] Martin Fowler. *Analysis Patterns: Reusable Object Models*. Addison-Wesley, 1997.

[Frank and Nevill-Manning1999] Paynter G.W. Witten I.H. Gutwin C. Frank, E. and C. Nevill-Manning. Domain-specific keyphrase extraction. In *Proceedings of the International Joint Conference on Artificial Intelligence*, pages pp.668–673. Morgan Kaufmann Publishers, San Francisco, CA, 1999.

[GAT] Gate tool (generic architecture for text engineering.

[Hirschfeld and Lämmel2004] R. Hirschfeld and R. Lämmel. Reflective Designs. Draft, 32 pages; Submitted to IEE Proceedings Software, 4 July 2004.

[Huffman1995] Scott B. Huffman. Learning information extraction patterns from examples. In *Learning for Natural Language Processing*, pages 246–260, 1995.

[MeSa] Mesh (medical subject headings).

[MeSb] Mesh tree structure.

[Moreno and Perezmay 2001] Antonion Moreno and Chantal Perez. From text to ontology: Extraction and representation of conceptual information. In *Proceedings of the Conference TIA*, may 2001. URL: $http://www.cs.utah.edu/~riloff/psfiles/aaai96.pdf$.

[NCI] National cancer institute (nci) ontology.

[NGC] National guideline clearinghouse (ngc) homepage.

[Pro] Protocure ii project.

[Riedewald] Ralf Lammel; Gunter Riedewald. A calculus for modular and extensible language definition in ldl.

[Riloff and Shoen1995] Ellen Riloff and Jay Shoen. Automatically acquiring conceptual patterns without an automated corpus. In David Yarovsky and Kenneth Church, editors, *Proceedings of the Third Workshop on Very Large Corpora*, pages 148–161, Somerset, New Jersey, 1995. Association for Computational Linguistics.

[Riloff1996] E. Riloff. Automatically generating extraction patterns from untagged text. In *Proceedings of the Thirteenth National Conference on Artificial Intelligence (AAAI-96)*, pages 1044–1049, 1996.

[S.B.; and Goldstein2003] Shahar Y; Shalom E.; Mayaffit A.; Young O.; Galperin M.; Martins S.B.; and M.K. Goldstein. A distributed, collaborative, structuring model for a clinical-guideline digital-library. In *Proceedings of the 2003 AMIA Annual Fall Symposium, Washington, DC*, 2003.

[Serban2004] Radu Serban. *Task 2.5: Building a library of guideline patterns (Progress report, Protocure II Project)*. May 2004.

[Seyfang *et al.*2004] Andreas Seyfang, Sylvia Miksch, and Peter Votruba. Deliverable d2.2a specification of formats of intermediate, asbru and kiv representations. Technical Report Deliverable D2.2a, Protocure-II, June 2004. URL: http://www.keg.uji.es/protocure/DeliverableV3.pdf.

[Seyfang2002] R.; Miksch S. Seyfang, A.; Kosara. Asbru's reference manual, asbru version 7.3. Technical report, Vienna University of Technology, Institute of Software Technology and Interactive Systems, 1 2002. URL: http://www.ifs.tuwien.ac.at/asgaard/asbru/asbru$_{73}/$.

[Ste] Stepper tool: A mark-up tool for narrative guidelines.

[Svatek2004] Vojtech Svatek. Design patterns for semantic web ontologies: Motivation and discussion. In Witold Abramowicz, editor, *Proceedings of Business Information Systems Conference (BIS2004), Poznan, Poland*, 2004.

[ten Teije; Anna Puig2004] Radu Serban; Annette ten Teije; Anna Puig. Experiments on guideline formalization with tool support. Technical report, Vrije Universiteit, 2004. URL: http://www.cs.vu.nl/ serbanr/ExperGuidFormal.pdf.

[Tex2001] *TextToOnto Software Tool*. 2001.

[UML] Unified medical language system.

[Vlissides1997] John Vlissides. Patterns: The top ten misconceptions. *Object Magazine*, 1997.

[von Dincklage] Daniel von Dincklage. Making patterns explicit with metaprogramming.

[Witten] Ian H. Witten. Adaptive text mining: Inferring structure from sequences.

[Zhong *et al.*2002] Jiwei Zhong, Haiping Zhu, Jianming Li, and Yong Yu. Conceptual graph matching for semantic search. In *ICCS*, pages 92–196, 2002.

[Zimmer2001] Walter Zimmer. Relationships between design patterns. pages 345–364. 2001.

A Building an ontology for medical background knowledge

Comprehensive ontologies for the medical domain already exist ([**?**]), but quite frequently the same terms are viewed as belonging to different categories.

We built a smaller ontology containing terms relevant to the oncology domain, by using as *input (training set) three evidence-based cancer guidelines*:

1. the Scottish Guideline for Treatment of Breast Cancer, denoted SIGN;
2. the Royal College of Radiologists Guideline for Treatment of Breast Cancer, denoted RCR;
3. the Dutch Institute for Quality of Healthcare Guideline for Treatment of Breast Cancer, denoted CBO.

The categories and the patterns identified in the training set were then validated on a test set which consists of two evidence-based guidelines:

1. the RCR (Royal College of Radiologists) Guideline for Treatment of Lung Cancer;
2. the ACCP (American College of Chest Physicians) Guideline for Diagnosis and Treatment of Lung Cancer.

We verified how many of the terms found in one type of oncology guidelines are generic medical terms for which background knowledge is available (by checking their presence in the MESH thesaurus) and verified how many of the terms can actually be found in another type of oncology guidelines. Furthermore, we attempted to validate pattern templates found in the training set on the test set.

The first step was to build a simple (containing mostly is_a and inst_of relationships) ontology containing terms from these texts, which can be used to identify patterns in the guideline text. Each term represents an instance of at least one semantic category (class), and categories are connected through is_a relationships.

For identifying patterns, a semantic tagging of the words in the text must be performed. As a pre-requisite for performing semantic tagging, we must have a classification of several relevant medical terms according to several categories.

We chose a categorization which extends the NGC semantic axes, as it is used in the DeGeL ([DeG, A.2002, S.B.; and Goldstein2003]) and has common axes with the MeSH categorization ([MeSb]). For categorization we used the MESH2004 ([MeSa]) descriptor set, containing more than 150.000 terms from the medical domain with their corresponding semantic category.

This medical thesaurus was chosen as it provides a semantic category attribute for each medical term it contains. Provided that enough words in the guidelines are covered by MESH, automatic tagging of guideline text would be possible, which represents the first step for automatic searching of guideline patterns.

A set of medical terms has been extracted from MESH using an XSL stylesheet, and pre-processing of the output, which included: removing of duplicates, renaming of strange characters to Prolog-allowed constant names and sorting of terms. As a result, from approx. 155,000 terms contained in MESH2004Descr, we have extracted 110,000 unique terms in MESHSET, which is further used as containing all terms in MESH.

Note that MESH2004Descr contains the list of medical terms accepted as of 2004 by the US National Library of Medicine, after a pre-processing that removed irrelevant terms (very long names, words containing non-alphanumeric characters, etc).

For about 150 relevant medical terms (extracted as described in section A) we have extracted their semantic categories from the MESH.

Building the medical ontology - Statistics The step of extracting a medical ontology from a corpus of medical texts is a typical knowledge extraction task, and has been addressed by recent work in this fi eld ([Frank and Nevill-Manning1999, Moreno and Perezmay 2001, Riloff1996, Huffman1995]). Our work is related to that of Moreno & Perez ([Moreno and Perezmay 2001]) to the extent that we aim at extracting conceptual information from existing medical texts.

We used the TextToOnto tool ([Tex2001]) to extract the vocabularies containing the most frequently used and relevant terms in the each of the texts. The sizes of these vocabularies (having a relevance threshold above 0.5) are:

- SIGN: 200 terms
- CBO: 323 terms
- RCR: 225 terms

Note that from a total of 487 unique relevant terms used by the three input guidelines, only 74 (about 15%) terms were shared by all three ontologies:

```
CBO n RCR n SIGN have 74 common elements: [advanced_breast_cancer,
assess, axilla, benefit, biopsy, bone, breast, cancer, carcinoma,
care, chemotherapy, chest, clin, combin, control, core_biopsy, data,
diagnosi, diseas, dose, effect, evid, examin, excision, grade, group,
high_risk, hormon, hormone_therapy, hospit, inform, level, life,
lymph, mammographi, manag, med, menopaus, node, oncologi,
operable_breast, operable_breast_cancer, overview, pain, patient,
patients, practic, primari, qual, radiat, radiation_therapy,
radiotherapi, rate, receptor, recurr, recurrence, report, research,
respons, risk, size, stage, statu, studi, support, surgeon, surgeri,
surviv, therapy, treatment, trial, tumour, tumour_size, year]
```

In fi gure 20 we show additional medical terms that are shared by only two-out-of-the-three vocabularies.

The number of relevant medical terms that are specifi c for the oncology domain is relatively low, below 70 concepts, which amounts to 10–15% of relevant terms identifi ed. This suggests that the categorization (mapping between medical terms and medical categories) used to search for patterns can be applied to a wider range of medical guidelines, not only to oncology guidelines.

It would be interesting to investigate what coverage of these relevant terms is offered by the NCIOncology ([]) ontology and by the WordNet ([]) taxonomy. This would provide a second estimate on the available background knowledge for linking concepts in medical guidelines automatically, in search for patterns.

Let us now focus on relations of the input ontologies (CBO,RCR,SIGN) with the set of terms present in an existing ontology for medical domain, namely MESH.

The intersection of terms from the three vocabularies extracted from the input guidelines (CBO, RCR, SIGN) with the terms in MESHSET (110.000+ terms) yields 40 common elements, which is less than 10% of all (487) relevant terms found in the three input guidelines:

```
MeshShared = 'SharedTerms in CBRnSIGNnRCR' and 'MESHSET' share 40
elements:
[axilla, benefit, biopsy, bone, breast, cancer, carcinoma, care,
chemotherapy, chest, control, data, dose, effect, excision, group,
```

```
Vocabularies 'CBO'(323 elements) and 'RCR'(225 elements) share 100 elements; most
specific ones:
[adjuvant_chemotherapy, age, early_breast_cancer, breast_cancer,
menopausal_status, metastatic_breast, metastatic_breast_cancer,
recurrence_rate, result, review, survival_benefit, systemic_therapy,
systemic_treatment]

Vocabularies 'CBO'(323 elements) and 'SIGN'(200 elements) share 105 elements;
most specific ones:
[breast_cancer_patients, breast_conservation, breast_reconstruction,
cancer_patients, clinical_trial, conservation, increased_risk,
metastasis, morbidity, outcome, symptom,systemic_therapy ]

Vocabularies 'RCR' and 'SIGN' share 103 elements; most specific:
[ablation, breast_care, breast_disease, care_team, death, decis,
excision, health, health_care_team, local_excision, local_recurrence,
metastatic_disease, response_rate, risk_disease, sole_therapy,
specialist]
```

Figure 20: *Medical terms in shared guideline vocabularies*

```
high_risk, level, life, lymph, node, overview, pain, patient,
patients, radiation_therapy, rate, receptor, recurrence, report,
research, risk, size, stage, support, surgeon, therapy,
treatment, trial, year]

Set 'MESHSET' has 110666 elements.

'CBO' n 'MESHSET' have 87 common elements: MeshShared u
[adjuvant_chemotherapy, age, anxiety, appendix, arm, breast_cancer,
breast_reconstruction, clinical_trial, dissection, ductal, liver,
lymph, lymph_node, method, morbid, node, number, order]

'RCR' n 'MESHSET' have 79 common elements: MeshShared u
[breast_cancer, breast_disease, care_team, chest_wall,
combination_chemotherapy, death, disease, factor, health,
health_care_team, interval, local, local_excision,
lymph_node, management, mastectomy, measurement,
nurse, oncology, organization, polychemotherapy]

'SIGN' n 'MESHSET' have 87 common elements: MeshShared n [age_group,
audit, breast_disease, breast_reconstruction, care_team, clinic,
clinical_trial, complication, conservation, death, disease, health,
health_care_team, health_service, hospital, involvement, local,
local_excision, morbidity, network, nurse, nursing, palliative_care,
radiotherapy, reconstruction, role, screen, screening, section,
service, skin, specialist, specimen, staff, surgery, symptom, team, test]
```

The union of three guidelines (CBO, SIGN, RCR) produced a list of 487 ”relevant” terms (relevance being measured as a minimal occurrence threshold in TextToOnto), of which 142 can be found back in MESH and therefore assigned a semantic category.

```
'Union of CBRxRCRxSIGN' n 'MESHSET' have 142 common elements:
[ablation, access, adjuvant_chemotherapy, age, age_group, anxiety,
appendix, arm, assessment, audit, axilla, benefit, biopsy, bone,
breast, breast_cancer, breast_disease, breast_reconstruction, cancer,
carcinoma, care, care_team, chapter, chemotherapy, chest, chest_wall,
clinic, clinical_trial, coin, college, combination_chemotherapy,
complication, conservation, control, course, data, death, disease,
```

```
dissection, dose, ductal, effect, event, excision, factor, form,
function, group, health, health_care_team, health_service, high_risk,
hospital, impact, interval, involvement, level, life, line, liver,
local, local_excision, lymph, lymph_node, management, mastectomy,
mean, measurement, method, morbid, morbidity, network, node, number,
nurse, nursing, oncology, order, organization, overview, pain,
palliative_care, patient, patients, period, phase, polychemotherapy,
post, problem, procedure, progress, quality, radiation_therapy,
radiotherapy, rate, receptor, reconstruction, recurrence, report,
research, review, risk, role, screen, screening, second, section,
sentinel, service, set, shoulder, sign, size, skin, specialist,
specimen, staff, stage, standard, status, subject, support, surgeon,
surgery, survival, survival_rate, symptom, team, test, therapy, time,
time_interval, toxic, treatment, trial, tumor, type, ultrasound, unit,
uses, wall, year]
```

The intersection of the relevant terms in the three guidelines produced a list of 74 ”relevant” terms, of which only 40 are medical terms with an established semantic category.

Any sentence containing words from the union of relevant terms from the three input guidelines (CBO, RCR, SIGN) is viewed as potentially relevant for pattern search.

For the 142 terms present in MESHSET a category could automatically be assigned. For about 40 others we could semi-automatically generate a category using the learning feature of the Taxonomy Builder component in TextToOnto tool. But for most of the other more than $487 - 142 - 40 = 305$ terms we had to assign manually a category.

The result of the Taxonomy Builder is a list $subConceptOf(concept_1, concept_2, likelihood)$, which will be incorporated into our ontology as a set of $inst\_of(concept_1, category_2)$ facts.

Examples:

```
subConceptOf(antidepressant,medication,1.0)
subConceptOf(needle biopsy,procedure,1.0)
subConceptOf(pain,symptom,1.0)
subConceptOf(tumour size,factor,1.0)
subConceptOf(practitioner,care professionals,1.0)
subConceptOf(laser therapy,technique,1.0)
```

This $subConceptOf$ relation is translated into an $inst\_of$ relation, using the following convention: $A\ is\ inst\_of\ B$ if concept B belongs to the list of semantic categories; otherwise $A\ is\ inst\_of\ B$ will be rewritten as $A\ is\ inst\_of\ C$ where C is the semantic category of B.

At the end of this process, we have obtained a mini-ontology comprised of is_a relations between the main semantic categories (built by hand) and $inst\_of$ relations between text fragments and semantic categories (partly extracted automatically, partly filled in manually).

B Implementation details

B.1 Ontology of the medical domain

The ontology is available as annex of the deliverable, in file categ_terms.pl of the archive deliverable25.zip, which is available at URL: www.protocure.org/deliverables. Here only the medical categories and the relations between them are given.

```
:- dynamic inst_of/2.
:- dynamic is_a/2.
:- dynamic is_relevant_word/3.
:- dynamic split_into_parts/2.
:- dynamic is_category/1.
:- dynamic is_med_category/1.
:- dynamic is_gen_category/1.
:- dynamic is_operator/1.
:- dynamic der_is_a/2.

% dictionary terms
inst_of(med_intervention,axillary_surgery).
inst_of(med_intervention,radiotherapy).

inst_of(diag_med_intervention,bronchoscopy). % lung cancer
inst_of(diag_med_intervention,mammographic_check).

inst_of(med_factor,age_limit).
inst_of(med_factor,tumour_size).

inst_of(action,produces).
inst_of(action,receive).

inst_of(cond_op,when).
inst_of(cond_op,in_case_of).
inst_of(cond_op,in_the_event_of).

inst_of(result_op,risk_of).
inst_of(result_op,signs_of).
inst_of(result_op,presence_of).
inst_of(result_op,by_means_of).
inst_of(result_op,based_on).

% relations between medical categories of our model
is_a(prescription,treatment).
is_a(treatment,med_intervention).
is_a(med_intervention,med_action).
is_a(gen_med_action,med_action).
is_a(med_action,action).
is_a(med_actor,target_group).
is_a(symptom,med_condition).
is_a(med_factor,med_condition).
is_a(med_condition,condition).
is_a(neg_evidence,evidence).
is_a(med_preference,med_concept).
is_a(med_practice,med_concept).
is_a(diagnostic,med_concept).
is_a(disease,disorder).
is_a(disorder,med_concept).

is_a(rel_op,generic_op).
is_a(act_op,generic_op).
is_a(result_op,generic_op).
is_a(effect_op,generic_op).
is_a(cond_op,generic_op).
```

```
is_a(target_group,med_context).
is_a(temp_rel_op,rel_op).
is_a(causal_rel_op,rel_op).
is_a(assoc_rel_op,rel_op).
is_a(excl_rel_op,rel_op).
is_a(comp_rel_op,rel_op).
is_a(neg_rel_op,rel_op).
is_a(def_rel_op,rel_op).
is_a(neg_result_op,result_op).
is_a(pos_result_op,result_op).

is_med_category(X) :-
    X=treatment;
    X=med_intervention;
    X=gen_med_action;
    X=med_action;
    X=med_context;
    X=med_reference;
    X=gen_med_concept;
    X=lab_parameter;
    X=lab_parameter_val;
    X=med_condition;
    X=diagnostic;
    X=symptom;
X=disorder;
    X=disease;
    X=disease_attr;
    X=disease_attr_val;
X=prescription;
X=medication;
X=med_factor;
X=med_preference;
X=med_practice;
X=med_actor;
    X=med_effect;
    X=body_part.

is_gen_category(X) :-
   X=action;
X=target_group;
    X=recommendation;
    X=condition;
    X=quantification;
    X=evidence;
X=time_spec;
    X=neg_evidence;
X=definition;
    X=intention.

is_category(X) :-
   (
    is_med_category(X);
is_gen_category(X);
is_operator(X);
X=generic_op;
X=med_concept
).

is_operator(X) :-
    X=rel_op;
X=act_op;
X=decomp_op;
X=assoc_rel_op;
X=order_rel_op;
X=logical_rel_op;
```

```
X=causal_rel_op;
X=cond_op;
X=excl_rel_op;
X=neg_rel_op;
X=comp_rel_op;
X=common_op;
X=result_op;
X=neg_result_op;
X=pos_result_op;
X=def_rel_op;
X=effect_op;
X=temp_rel_op.
```

B.2 Analysis of CBO Breast Cancer Guideline

This section contains a summary of the output generated by the program for pattern search for the analysis of chapter 3 of the CBO guideline ([**?**]). The entire output produced by the program is available in the annex of this deliverable, as file output_pattsearch_analysis_cbo_ch3.txt in archive www.protocure.org/deliverables/deliverable25.zip.

Below only a sample containing the most relevant pattern instances identified is provided. For each patterns instance the following attributes are given:

1. list of medical categories present in the pattern template instantiated;
2. the generic name of the pattern template;
3. the most specific list of categories, i.e., a list of the most ”specialized” pattern template that still is instantiated by the same pattern instance; the least specific list of categories, i.e., corresponding to a more ”abstract” pattern template, is present in the output of the program, but was suppressed in the list, as almost all pattern templates correspond to the generic pattern template $[med\_concept, med\_concept]$ or $[med\_concept, generic\_op, med\_concept]$;
4. the text present in the pattern instance, presented as a list of the same length as that of the medical categories of the pattern template.

```
There were 179 relevant sentences processed
====
List of 226 patterns:
4: [disease, assoc_rel_op, disease, assoc_rel_op, disease_attr] (p_def2)
    most specificific:[disease, assoc_rel_op, disease, assoc_rel_op, disease_attr]
    [locoregionally_advanced_breast_cancer, opis_used_to_describe,
     breast_cancer, which_opis, unresectable]
5: [disease_attr, assoc_rel_op, med_factor] (p_factor1)
    most specificific:[disease_attr, assoc_rel_op, med_factor]
    [unresectable, on_the_basis_of, unresectability_criteria]
7: [evaluation, med_effect] (p_result3)
    most specificific:[evaluation, med_effect]
    [higher, response_rate]
11: [med_action, assoc_rel_op, body_part] (p_body)
    most specificific:[med_action, assoc_rel_op, body_part]
    [biopsy, of, axilla]
12: [med_action, assoc_rel_op, disease] (p_assoc_act_disease)
    most specificific:[prescription, assoc_rel_op, disease]
    [current_treatment, opfor, locoregionally_advanced_breast_cancer]
17: [med_context, med_condition] (p_ctx3)
```

```
    most specificific:[med_context, disease]
    [in_the_case_of, inflammatory_breast_cancer]
19: [med_action, assoc_rel_op, disease] (p_assoc_act_disease)
    most specificific:[gen_med_action, assoc_rel_op, disease]
    [multidisciplinary_treatment, of, locoregionally_advanced_breast_cancer]
26: [med_action, assoc_rel_op, disease] (p_assoc_act_disease)
    most specificific:[gen_med_concept, assoc_rel_op, disease]
    [treatment, of, locoregionally_advanced_breast_cancer]
28: [lab_parameter, assoc_rel_op, diagnostic] (p_labpar)
    most specificific:[lab_parameter, assoc_rel_op, diagnostic]
    [tumour, with, clinical_invasion]
33: [med_intervention, logical_rel_op, med_intervention] (p_discovered12)
    most specificific:[med_intervention, logical_rel_op, prescription]
    [radiotherapy, opand, chemotherapy]
35: [med_intervention, logical_rel_op, med_intervention, decomp_op, treatment]
(p_act_composition1)
    most specificific:[med_intervention, logical_rel_op,
med_intervention, decomp_op, prescription]
    [neoadjuvant_systemic_therapy, opand, surgery, also, treatment_options]
38: [med_intervention, med_effect_op, lab_parameter] (p_result)
    most specificific:[med_intervention, med_effect_op, lab_parameter]
    [surgery, risk_of, locoregional_recurrence]
40: [med_intervention, med_effect_op, lab_parameter] (p_result)
    most specificific:[med_intervention, med_effect_op, lab_parameter]
    [surgery, to_reduce, tumour_load]
41: [med_intervention, prescription] (p_discovered15)
    most specificific:[prescription, prescription]
    [chemotherapy, regimens]
46: [decomp_op, med_intervention, logical_rel_op, med_intervention] (p_act_composition2)
    most specificific:[decomp_op, med_intervention, logical_rel_op, med_intervention]
    [addition_of, cmf_chemotherapy, opand, radiotherapy]
47: [decomp_op, med_intervention, assoc_rel_op, med_intervention] (p_act_composition3)
    most specificific:[decomp_op, med_intervention, assoc_rel_op, med_intervention]
    [addition_of, cmf_chemotherapy, to, mastectomy]
56: [med_action, decomp_op, med_action, med_action, med_action] (p_act_decomp)
    most specificific:[prescription, decomp_op, med_action, med_action, med_action]
    [current_treatment, consists_of, surgery, surgery, locoregional_radiotherapy]
57: [med_action, assoc_rel_op, disease] (p_assoc_act_disease)
    most specificific:[prescription, assoc_rel_op, disease]
    [current_treatment, of, locoregionally_advanced_breast_cancer]
77: [med_intervention, rel_op, med_intervention] (p_act_rel)
    most specificific:[med_intervention, temp_rel_op, med_intervention]
    [radiotherapy, following, neoadjuvant_chemotherapy]
78: [lab_parameter, assoc_rel_op, diagnostic] (p_labpar)
    most specificific:[lab_parameter, assoc_rel_op, diagnostic]
    [tumour_load, of, residual_tumour]
81: [med_intervention, excl_rel_op, assoc_rel_op, med_action] (p_excl_treatment)
    most specificific:[prescription, excl_rel_op, assoc_rel_op, med_action]
    [treated, solely, with, neoadjuvant_chemotherapy]
94: [causal_rel_op, med_action, rel_op, med_action] (p_causal_act)
    most specificific:[causal_rel_op, common_op, assoc_rel_op, med_action]
    [since, treatment_of_choice, opis, neoadjuvant_chemotherapy]
98: [med_action, assoc_rel_op, disease] (p_assoc_act_disease)
    most specificific:[common_op, assoc_rel_op, lab_parameter]
    [treatment_of_choice, to, tumour]
100: [med_action, assoc_rel_op, med_action, temp_rel_op, med_action] (p_ctx_assoc)
    most specificific:[prescription, assoc_rel_op, med_action, temp_rel_op, prescription]
    [neoadjuvant_treatment, with, adjuvant_chemotherapy, before, chemotherapy]
104: [quantification, lab_parameter] (p_quantif_disease)
    most specificific:[evaluation, lab_parameter]
    [better, locoregional_control]
106: [med_action, assoc_rel_op, disease, rel_op, med_intervention] (p_derived7)
    most specificific:[lab_parameter, assoc_rel_op, diagnostic,
temp_rel_op, med_intervention]
    [resection, of, residual_tumour, following, neoadjuvant_chemotherapy]
```

```
107: [quantification, lab_parameter] (p_quantif_disease)
    most specificific:[quantification, lab_parameter]
    [high_risk, locoregional_recurrence]
110: [recommendation, assoc_rel_op, med_effect] (p_rec3)
    most specificific:[recommendation, assoc_rel_op, med_effect]
    [attention_should_be_paid, to, side_effects]
113: [action, recommendation, med_effect] (p_rec4)
    most specificific:[gen_med_action, recommendation, med_effect]
    [multidisciplinary_treatment, attention_should_be_paid, skin_reactions]
114: [evaluation, med_effect] (p_result3)
    most specificific:[evaluation, med_effect]
    [severe, skin_reactions]
115: [med_action, assoc_rel_op, body_part] (p_body)
    most specificific:[gen_med_action, assoc_rel_op, body_part]
    [multidisciplinary_treatment, of, shoulder]
116: [med_intervention, temp_rel_op, med_intervention] (p_temp_order)
    most specificific:[med_intervention, temp_rel_op, prescription]
    [radiotherapy, following, chemotherapy]
123: [recommendation, med_action, assoc_rel_op, target_group] (p_rec2)
    most specificific:[recommendation, prescription, assoc_rel_op, target_group]
    [recommend, chemotherapy, opfor, this_group_of_patients]
128: [target_group, assoc_rel_op, disease] (p_ctx1)
    most specificific:[target_group, assoc_rel_op, lab_parameter]
    [inoperable_patients, with, tumour]
129: [target_group, assoc_rel_op, disease] (p_ctx1)
    most specificific:[target_group, assoc_rel_op, disease]
    [oncology_patients, with, locoregionally_advanced_breast_cancer]
141: [target_group, assoc_rel_op, lab_parameter] (p_tgroup)
    most specificific:[target_group, assoc_rel_op, lab_parameter]
    [patients, with, locoregional_recurrence]
142: [target_group, recommendation, med_action] (p_rec_targetgroup)
    most specificific:[target_group, recommendation, gen_med_action]
    [patients, should, multidisciplinary_treatment]
143: [treatment, assoc_rel_op, diagnostic] (p_discovered1)
    most specificific:[prescription, assoc_rel_op, disease]
    [current_treatment, of, locoregionally_advanced_breast_cancer]
167: [treatment, assoc_rel_op, med_action] (p_discovered3)
    most specificific:[prescription, assoc_rel_op, med_action]
    [treated, with, radiotherapy]
172: [treatment, assoc_rel_op, med_action] (p_discovered3)
    most specificific:[common_op, assoc_rel_op, med_action]
    [treatment_of_choice, opis, neoadjuvant_chemotherapy]
191: [treatment, assoc_rel_op, med_practice] (p_discovered6)
    most specificific:[common_op, assoc_rel_op, common_op]
    [treatment_of_choice, opis, usually]
193: [treatment, assoc_rel_op, target_group] (p_discovered8)
    most specificific:[prescription, assoc_rel_op, target_group]
    [chemotherapy, opfor, this_group_of_patients]
196: [treatment, excl_rel_op, med_action] (p_discovered9)
    most specificific:[prescription, excl_rel_op, med_action]
    [treated, solely, neoadjuvant_chemotherapy]
```

C List of Control Patterns

This section contains a compact representation of the most relevant attributes of the control patterns. The entire pattern template record saved in the patterns database, is available in the annex as document $db\_patterns.xml$ at URL: $www.protocure.org/deliverables/deliverable25.zip$.

The starting letters of the pattern template identifiers indicate whether the pattern has a strong control component (the pattern starting with letter C) and was built from an abstract control structure, or whether it was derived from a text fragment (and therefore contains a mix of several pattern instances) - the patterns starting with letter I.

The control patterns are stored in a MySQL database with the structure depicted in table 2, which can be exported into an XML document. Except from the textual and ASBRU representation (if any), for all pattern templates a list of medical categories present in the template is defined, and a vocabulary of terms which can be used to recognize an instance of the template (for instance, as operators) is given.

By means of XSL scripts, the content of the database can be used (as GMT macros, for instance) to assist in the guideline formalization process, or to mark the terms in the guideline text that are known to belong to patterns.

Table 2: Structure of table patterns

Field	Type	Comments
pid	int(11)	DB counter
pattid	varchar(100)	pattern identifi er
plabel	varchar(100)	name of pattern
pargs	varchar(200)	arguments taken by pattern
pattr	varchar(200)	mapping context attributes-arguments
pcateg	varchar(200)	list of categories pattern belongs to
pdescr	text	description of the pattern
ptext	text	textual instances of the pattern
pir	text	intermediate representation of the pattern
pformal	text	formal representation (ASBRU) of the pattern
ploc	varchar(100)	location in guideline text where it occurs
pprops	text	properties of the pattern
prela	text	relations to other patterns
pobs	text	observations about this pattern
pdatin	datetime	date of fi rst entry
pdatupd	datetime	date of last update
plevel	varchar(30)	level of abstraction where it occurs
p8hb	varchar(200)	equivalent in MHB
pdict	text	dictionary terms used in this pattern
pguideline	varchar(100)	
psesame	text	sesame annotation
psupport	varchar(50)	support for pattern in texts
pconfi dence	varchar(50)	confi dence that pattern is really useful
powner	varchar(10)	owner (category) of the pattern

The entries in the control pattern database are given below.

Table 3: Pattern 1

Identifier/Label	$C01$ **Simple action execution**
Arguments	[action],[context]
IR	$Execute\ [action]\ in\ [context].$
Category	
Samples	$(CBO - BC):\ Administration\ of\ a\ boost\ dose\ in\ the\ case\ of\ BCT\ for\ DCIS\ may\ be\ beneficial.$
ASBRU	$plan\ [action]$ $plan - body\ to - be - defined.$ $plan\ callee - [action]$ $plan - body\ do\ sequentially$ $plan - activation\ [action].$

Table 4: Pattern 2

Identifier/Label	$C02$ **Exclusive alternative branching actions**
Arguments	[action-1],[action-2]
IR	1. $Do\ either\ [action - 1]\ or\ [action - 2].$ 2. $[Action] := [action - 1]\ xor\ [action - 2].$ 3. $[action - 1]\ excludes\ [action - 2].$ 4. $[action - 1]\ is\ incompatible\ with\ [action - 2].$
Category	
Samples	
ASBRU	$plan\ branching - [action - 1] - [action - 2]$ $plan - body\ subplans\ order = any - order\ wait - for = one$ $choice := 0$ $plan - activation\ [action - 1] - branch$ $plan - activation\ [action - 2] - branch.$ $plan\ [action - 1] - branch$ $conditions$ $abort - condition : choice = 2$ $plan - body\ subplans\ sequentially$ $if\ choice = 0$ $choice := 1$ $endif$ $if\ choice = 1$ $plan - activation\ [action - 1]$ $endif.$ $plan\ [action - 2] - branch$ $conditions$ $abort - condition : choice = 1$ $plan - body$ $subplans\ sequentially$ $if\ choice = 0$ $choice := 2$ $endif$ $if\ choice = 2$ $plan - activation\ [action - 2]$ $endif.$ $plan\ [action - 1]$ $plan - body\ to - be - defined.$ $plan\ [action - 2]$ $plan - body\ to - be - defined.$

Table 5: Pattern 3

Identifier/Label	$C03$ **Trigger condition-action**
Arguments	[condition],[action]
IR	1. $When\ [condition]\ holds,\ execute\ [action].$ 2. $If\ [condition],\ do\ [action].$ 3. $[Condition]\ holds\ when\ executing\ [action]$
Category	
Samples	$There\ is\ no\ randomised\ evidence\ to\ indicate\ a\ beneficial\ effect$ $when\ adding\ surgery\ to\ neoadjuvant\ chemotherapy\ and\ radiotherapy.$
ASBRU	$plan\ [action]$ $conditions$ $filter - condition : [condition]$ $plan - body\ [to - be - defined].$

Table 6: Pattern 4

Identifier/Label	$C04$ **Conditional action execution**
Arguments	[condition],[action-1],[action-2]
IR	$If\ [condition]\ then\ do\ [action - 1]\ else\ do\ [action - 2].$
Category	
Samples	
ASBRU	$plan\ [action - 1]$ $conditions$ $filter - precondition :\ [condition]$ $plan - body\ to - be - defined$ $plan\ [action - 2]$ $conditions$ $filter - precondition :\ not([condition])$ $plan - body\ to - be - defined$

Table 7: Pattern 5

Identifier/Label	$C05$ **Association (context,execution mode,action)**
Arguments	[medical-context],[action]
IR	$In\ [medical - context],\ [action]\ is\ recommended/optional/forbidden.$
Category	
Samples	
ASBRU	

Table 8: Pattern 6

Identifier/Label	$C06$ **Modal action execution**

Arguments	[modal-qualifi er],[action]
IR	1. $Possible\ to\ do\ [action]$. 2. $Obligatory\ to\ do\ [action]$. 3. $Recommended\ to\ do\ [action]$. 4. $Forbidden\ to\ do\ [action]$. 5. $Impossible\ to\ do\ [action]$. 6. $It\ is\ recommended/optional/necessary/forbidden\ to\ perform\ [action]$.
Category	
Samples	
ASBRU	

Table 9: Pattern 7

Identifi er/Label	$C07$ **Association (context,condition,modal-action)**
Arguments	[modal-action-1],[medical-context],[condition]
IR	$Do\ [modal - action - 1]\ in\ [medical - context]\ if\ [condition]$.
Category	
Samples	
ASBRU	

Table 10: Pattern 8

Identifi er/Label	$C08$ **Repetitive action**
Arguments	[action]
IR	
Category	
Samples	$Repeat\ [action]$.
ASBRU	$plan\ repeat - [action]$ $plan - body\ \ cyclical - plan\ \ \ \ plan - activation\ [action]$ $plan\ [action]$ $plan - body\ to - be - defined$.

Table 11: Pattern 9

Identifi er/Label	$C09$ **Repetitive action with delay specifi cation**
Arguments	[action],[time-specifi cation]
IR	$Do\ [action]\ every\ [time - specification]$.
Category	
Samples	
ASBRU	$plan\ repeat - [action]$ $plan - body\ type = sequentially\ wait - for - all$ $cyclical - plan$ $plan - activation\ [action]$ $retry - delay\ [time - specification]$. $plan\ [action]$ $plan - body\ to - be - defined$.

Table 12: Pattern 10

Identifier/Label	$C10$ **Repetitive action with duration specification**
Arguments	[action],[duration-time-specification]
IR	$Repeatedly\ do\ [action]\ for\ [duration-time-specification].$
Category	
Samples	
ASBRU	$plan\ repeat-[action]$ $plan-body\ type = sequentially\ wait-for-all$ $cyclical-plan$ $cyclical-plan-body$ $plan-activation\ [action]$ $repeat\ specification$ $duration\ [duration-time-specification].$ $plan\ [action]$ $plan-body\ to-be-defined.$

Table 13: Pattern 11

Identifier/Label	$C11$ **Repetitive action with goal specification**
Arguments	[action],[goal-specification]
IR	$Do\ [action]\ until\ [condition]\ becomes\ true.$
Category	
Samples	
ASBRU	

Table 14: Pattern 12

Identifier/Label	$C12$ **Sequential action composition (Sequentially ordered actions)**
Arguments	[action-1],[action-2]
IR	1. $After\ [action-1]\ do\ [action-2].$ 2. $Before\ [action-2]\ do\ [action-1].$ 3. $Do\ [action-1],\ then\ [action-2].$ 4. $[Action-1]\ has\ to\ be\ finished\ before\ [action-2]\ can\ start.$
Category	
Samples	
ASBRU	$plan\ seq-[action-1]-[action-2]$ $plan-body\ wait-for\ all$ $subplans\ type = sequentially$ $plan-activation\ [action-1]\quad plan-activation\ [action-2]$ $plan\ [action-1]$ $plan-body\ to-be-defined$ $plan\ [action-2]$ $plan-body\ to-be-defined$

Table 15: Pattern 13

Identifi er/Label	$C13$ **(Unspecifi ed order) execution of a group of actions**
Arguments	[action-1],...,[action-n]
IR	$Perform\ unordered\ the\ following\ actions:\ [action-1],\ \ldots,\ [action-n].$
Category	
Samples	
ASBRU	$plan\ seq-[action-1]-[action-2]$ $plan-body\ wait-for\ all$ $subplans\ type=[unordered]$ $plan-activation\ [action-1]$ $plan-activation\ [action-2]$ $\ldots$ $plan-activation\ [action-n]$ $plan\ [action-1]$ $plan-body\ to-be-defined$ $plan\ [action-2]$ $plan-body\ to-be-defined$ $plan\ [action-n]$ $plan-body\ to-be-defined$

Table 16: Pattern 14

Identifi er/Label	$C14$ **Ordered action group execution**
Arguments	[action-1],...,[action-n],[exec-order-spec]
IR	$1.\ Do\ sequentially:\ [action-1],\ \ldots,\ [action-n].$ $2.\ Do\ in\ parallel:\ [action-1],\ \ldots,\ [action-n].$
Category	
Samples	
ASBRU	$plan\ seq-[action-1]-[action-2]$ $plan-body\ wait-for\ all$ $subplans\ type=[exec-order-spec]$ $plan-activation\ [action-1]$ $plan-activation\ [action-2]$ $\ldots$ $plan-activation\ [action-n]$ $plan\ [action-1]$ $plan-body\ to-be-defined$ $plan\ [action-2]$ $plan-body\ to-be-defined$ $plan\ [action-n]$ $plan-body\ to-be-defined$

Table 17: Pattern 15

Identifi er/Label	$C15$ **Sequentially ordered actions associated with context**
Arguments	[medical-context],[action-1],..,[action-n]

IR	$In\ [medical - context],\ perform\ sequentially:\ [action - 1],\ \ldots,\ [action - n].$
Category	
Samples	
ASBRU	$plan\ seq - [action - 1] - \ldots - [action - n]$ $conditions$ $filter - precondition:\ [med - context]$ $plan - body\ wait - for\ all$ $subplans\ type = [unordered]$ $plan - activation\ [action - 1]$ $plan - activation\ [action - 2]$ $\ldots$ $plan - activation\ [action - n]$ $plan\ [action - 1]$ $plan - body\ to - be - defined$ $plan\ [action - 2]$ $plan - body\ to - be - defined$ $plan\ [action - n]$ $plan - body\ to - be - defined$

Table 18: Pattern 16

Identifi er/Label	$C16$ **Random choice of element from group**
Arguments	[element],[group-of-element]
IR	$Select\ [element]\ from\ [group - of - element].$
Category	
Samples	
ASBRU	

Table 19: Pattern 17

Identifi er/Label	$C17$ **Conditional selection of element from group**
Arguments	[element], [group-of-element]
IR	$Select\ [element]\ from\ [group - of - element]\ satisfying\ [condition].$
Category	
Samples	
ASBRU	

Table 20: Pattern 18

Identifi er/Label	$C18$ **Selective action execution**
Arguments	[action-quantifi cation],[action-1],..,[action-n]
IR	$Perform\ all/some/first\ \ldots/at\ least\ first\ [action - quantification]$ $at\ most\ [action - quantification]\ of\ the\ following\ group - of\ actions:$ $[action - 1], \ldots, [action - n].$

Category	
Samples	
ASBRU	$plan\ seq - [action - 1] - \ldots - [action - n]$ $plan - body\ wait - for\ all$ $subplans\ type = seq\ wait - for\ selected - plans - [action - quantification]$ $plan - activation\ [action - 1]$ $plan - activation\ [action - 2]$ $\ldots$ $plan - activation\ [action - n]$ $plan\ [action - 1]$ $plan - body\ to - be - defined$ $plan\ [action - 2]$ $plan - body\ to - be - defined$ $plan\ [action - n]$ $plan - body\ to - be - defined$

Table 21: Pattern 19

Identifi er/Label	$C19$ **Trigger action-condition**
Arguments	[action],[condition]
IR	1. $After\ [action],\ [condition]\ holds.$ 2. $As\ a\ result\ of\ [action],\ [condition]\ holds.$
Category	
Samples	
ASBRU	$plan\ [action] effect\ \ [condition]$ $plan - body\ to - be - defined.$

Table 22: Pattern 20

Identifi er/Label	$C20$ **Association action-goal (intention)**
Arguments	[action],[proposition]/[goal]
IR	1. $To\ achieve\ [goal],\ perform\ [action].$ 2. $[Action]\ achieves\ [goal].$ 3. $By\ performing\ [action],\ [goal]\ can\ be\ achieved.$ 4. $The\ goal\ of\ procedure\ [action]\ is\ to\ ensure\ that\ [proposition]\ holds.$
Category	
Samples	
ASBRU	$plan\ [action]$ $intention$ $achieve - overall - state:\ \ [proposition]$ $plan - body\ to - be - defined$

Table 23: Pattern 21

Identifi er/Label	$C21$ **Negative association action-goal**

Arguments	[goal],[action]
IR	1. $[Goal]\ cannot\ be\ achieved\ by\ performing\ [action].$ 2. $[Action]\ does\ not\ lead\ to\ [goal].$
Category	
Samples	
ASBRU	

Table 24: Pattern 22

Identifier/Label	$C22$ **Composition of actions**
Arguments	[action], [action-1], [action-2]
IR	$[Action]\ is\ action-composition([action-1],[action-2]).$
Category	
Samples	
ASBRU	

Table 25: Pattern 23

Identifier/Label	$C23$ **Hierarchical decomposition of actions**
Arguments	[action], [action-1], [action-2]
IR	$[Action]\ consists\ of\ action-decomposition([action-1],\ldots[action-2]).$
Category	
Samples	
ASBRU	

Table 26: Pattern 24

Identifier/Label	$C24$ **Constraints on multiple modal actions**
Arguments	[context], [action-1], [action-2]
IR	1. $[Action-1]\ should\ be\ done\ in\ parallel\ with/before/after/never$ $in\ combination\ with/preferred\ to\ [action-2].$ 2. $In\ [context],\ [action-1]\ must/can\ be/must\ not\ be$ $done\ before/after/during\ [action-2].$
Category	
Samples	
ASBRU	

Table 27: Pattern 25

Identifier/Label	$C25$ **Modal decomposition of actions**
Arguments	[action], [group-of-actions]

IR	$[Action]\ should\ preferably/mandatory/optionally\ consist\ of\ action - composition\ of\ [group - of - actions].$
Category	
Samples	
ASBRU	

Table 28: Pattern 26

Identifier/Label	$C26$ **Subordination of actions**
Arguments	[action-1], [action-2]
IR	1. $[Action - 1]\ is\ part\ of\ [action - 2].$ 2. $[Action - 1]\ is\ executed\ as\ part\ of\ [action - 2].$
Category	
Samples	
ASBRU	

Table 29: Pattern 27

Identifier/Label	$C27$ **Explicit exclusion of action**
Arguments	[action]
IR	1. $Do\ not\ perform\ [action].$ 2. $Avoid\ do\ [action].$ 3. $Never\ do\ [action].$
Category	
Samples	
ASBRU	

Table 30: Pattern 28

Identifier/Label	$C28$ **Defining optional actions**
Arguments	[context], [action]
IR	1. $In\ [context],\ [action]\ is\ optional.$ 2. $In\ [context],\ [action]\ may\ be\ omitted/skipped.$
Category	
Samples	$Ex\ (NGC):\ In\ rare\ instances\ of\ small\ low\ grade\ cancers\ (i.e.,\ tubular\ carcinoma\ < 1\ cm),\ particularly\ in\ elderly\ or\ debilitated\ patients\ with\ a\ benign\ axillary\ exam,\ axillary\ dissection\ may\ be\ omitted.$
ASBRU	

Table 31: Pattern 29

Identifier/Label	$C29$ **Conditional recommendation for action**

Arguments	[condition], [modal-action]
IR	1. $If\ [condition],\ [modal-action].$ 2. $In\ [context]\ [modal-action].$ 3. $[Modal-action]\ when\ [condition].$
Category	
Samples	$(SIGN)\ Mastectomy\ is\ indicated\ for\ operable\ breast\ cancer\ which\ is\ either\ large$ $or\ at\ multiple\ sites,\ when\ radiotherapy\ is\ to\ be\ avoided,\ or\ by\ patient\ preference.$
ASBRU	$plan\ [modal-action]$ $conditions$ $filter-precondition:\ [condition]$ $plan-body\ to-be-defined.$

Table 32: Pattern 30

Identifi er/Label	$C30$ **Conditional exclusion of action in context**
Arguments	[medical-context], [medical-action], [condition]
IR	1. $In\ [medical-context]\ [medical-action]\ should\ not\ take\ place\ if\ [condition].$ 2. $Avoid\ [medical-action]\ if\ [condition]\ in\ [medical-context].$
Category	
Samples	
ASBRU	

Table 33: Pattern 31

Identifi er/Label	$C31$ **Recommendation of preferred and excluded actions**
Arguments	[condition], [medical-context], [preferred-medical-action], [alternative-medical-action]
IR	$If\ [condition]\ is\ true\ in\ [medical-context],\ then\ it\ is\ recommended\ to\ do$ $[preferred-medical-action]\ and\ never\ to\ do\ [alternative-medical-action].$
Category	
Samples	
ASBRU	

Table 34: Pattern 32

Identifi er/Label	$C32$ **Variable initialization/assignment pattern**
Arguments	[variable], [variable-value]
IR	1. $[Variable]\ :=\ [value]$2. $Initial\ value\ of\ [variable]\ is\ [variable-value].$
Category	
Samples	
ASBRU	$plan-body\ \ [variable]\ :=\ [variable-value]$

Table 35: Pattern 33

Identifi er/Label	$C33$ **Conditional variable assignment**
Arguments	[variable], [value-1], [value-2]
IR	$if\ [condition]\ then$ $[variable]\ :=\ [value-1]$ $else$ $[variable]\ :=\ [value-2]$
Category	
Samples	
ASBRU	$plan\ [plan-name]$ $plan-body\quad if\ [condition]\quad then\ [variable]\ :=\ [value1]\quad else\ [variable]\ :=\ [value2]$

Table 36: Pattern 34

Identifi er/Label	$C34$ **Conditional execution based on interval values**
Arguments	[discrete-variable],[min-range-i], [max-range-i], [action-i]
IR	$in\ case\ [discrete-variable]\ has\ value\ in\ range-1=[min-range-1, max-range-1]:$ $execute[action-1]$ $range-2=[min-range-2, max-range-2]:$ $execute[action-2]$ $\ldots$ $range-n=\ [min-range-n, max-range-n]:$ $execute[action-n]$ $unknown:$ $execute[default-action]$
Category	
Samples	
ASBRU	$plan\ [plan-name]$ $plan-body\ type=unordered$ $wait-for\ one$ $plan-activation\ plan-1$ $plan-activation\ plan-2$ $\ldots$ $plan-activation\ [procedure-n]$ $plan\ [plan-1]$ $conditions$ $filter-condition:[discrete-var]\ in\ [min-range-1, max-range-1]$ $plan-body\ to-be-defined$ $plan\ [plan-2]$ $conditions$ $filter-condition:[discrete-var]\ in\ [min-range-2, max-range-2]$ $plan-body\ to-be-defined$ $plan\ [plan-n]$ $conditions$ $filter-condition:[discrete-var]\ in\ [min-range-n, max-range-n]$ $plan-body\ to-be-defined$

Table 37: Pattern 35

Identifi er/Label	$C35$ **Evaluation of expression**

Arguments	[variable], [expression], [context]
IR	1. $[Variable] := value([expression])\ in\ [context]$ 2. $[Variable] := truth([proposition-expression])\ in\ [context]$
Category	
Samples	
ASBRU	

Table 38: Pattern 36

Identifi er/Label	$C36$ **Direct querying parameter value**
Arguments	[lab-parameter]
IR	$Ask\ value\ of\ [lab-parameter].$
Category	
Samples	
ASBRU	$plan-body\ \ ask\ [lab-parameter]$

Table 39: Pattern 37

Identifi er/Label	$C37$ **Time specifi cation pattern**
Arguments	[time-spec], [period], [time-unit]
IR	1. $[Time-spec] := every\ [period]\ [time-unit]$ 2. $[Time-spec] := for\ minimum/maximum/exactly\ [period]\ [time-unit]$
Category	
Samples	
ASBRU	

Table 40: Pattern 38

Identifi er/Label	$C38$ **Condition specifi cation pattern**
Arguments	[condition-def], [condition-1], [condition-2], [context]
IR	1. $[Condition-def] := boolean-operator([condition-1], [condition-2])$ 2. $[Condition-def] := [condition-1]\ in\ [Context]$
Category	
Samples	
ASBRU	

Table 41: Pattern 39

Identifi er/Label	$C39$ **Refi ned context**
Arguments	[context-def], [condition], [context]
IR	$[Context-def] := [Condition]\ in\ [Context]$

Category	
Samples	
ASBRU	

Table 42: Pattern 40

Identifier/Label	$C40$ **Synchronization of actions**
Arguments	[action-1], [action-2]
IR	1. $[Action-2]\ can\ start/complete/abort/suspend/resume\ after\ value\ of\ [parameter]$ $has\ been\ obtained,\ by\ performing\ [action-1].$ 2. $From\ [action-1],\ [action-2],\ ...,\ [action-n],\ [action-1]\ has\ to\ finish\ first.$
Category	
Samples	
ASBRU	

Table 43: Pattern 41

Identifier/Label	$C41$ **Conditional recommendation for repetitive action**
Arguments	[treatment-action], [control-action], [lab-parameter]
IR	$After\ [treatment-action]\ it\ is\ recommended/optional\ to\ repeat\ [control-action]$ $of\ [lab-parameter]\ every\ [period]\ [time-unit].$
Category	
Samples	
ASBRU	

Table 44: Pattern 42

Identifier/Label	$C42$ **Clustering of actions**
Arguments	[medical-action-1], [medical-action-2]
IR	1. $When\ [medical-action-1]\ is\ carried\ out\ [medical-action-2]\ is\ usually\ performed.$ 2. $[Medical-action-1]\ is\ accompanied\ by\ [medical-action-2].$
Category	
Samples	$(SIGN)\ When\ mastectomy\ is\ carried\ out\ axillary\ clearance\ is\ usually\ performed.$
ASBRU	

Table 45: Pattern 88

Identifier/Label	$F01$ **Action application**
Arguments	[med-action],[body-part],[medication]
IR	1. $[Med-action]\ applied-to\ [body-part].$ 2. $[Med-action]\ using\ [medication].$

Category	
Samples	
ASBRU	$plan\ [med-action]$ $plan-body\ to-be-defined.$

Table 46: Pattern 89

Identifi er/Label	$F02$ **Quantifi cation of effects**
Arguments	[quantifi cation],[med-action],[result]
IR	1. $[Med-action]\ produces\ [quantification]\ [result].$
Category	
Samples	
ASBRU	

Table 47: Pattern 57

Identifi er/Label	$I01$ **Checking of a disjunctive condition**
Arguments	[patient],[disease],[med-factor](1,n)
IR	$Any\ one\ of\ the\ following\ risk\ factors\ indicate/suggest\ that\ [patient]\ has\ [disease]:$ $[factor-1];\ [factor-2];\ \ldots[factor-n].$
Category	
Samples	$Are\ there\ any\ of\ the\ following\ risk-factors\ present\ to\ suggest\ that$ $non-immune\ hemolytic\ disease\ is\ possible\ in\ patient:\ \ldots?$
ASBRU	$plan\ treatment-for-[disease]$ $conditions$ $filter-condition:\ [disease]-risk-factor=true$ $plan-body\ to-be-defined$ $plan\ check-disjunctive-possibility-[disease]$ $conditions$ $complete-condition:\ [disease]-risk-factor=true$ $intentions$ $achieve-overall-state:$ $is-known-parameter([disease]-risk-factor)$ $plan-body\ type=any-order$ $wait-for\ one$ $plan-activation\ check-[disease]-risk-factor([factor-1])\ \ \ldots$ $plan-activation\ check-[disease]-risk-factor([factor-2])$ $plan\ check-[disease]-risk-factor$ $arguments\ factor$ $plan-body\ type=sequentially\ wait-for\ all$ $ask\ [factor]$ $if\ present([factor])$ $then\ [disease]-risk-factor\ :=\ true$

Table 48: Pattern 58

Identifier/Label	$I02$ **Sequence of parameter requests**
Arguments	
IR	$Ask\ the\ values\ of\ the\ following\ parameters:\ [lab-parameter-1],\ \ldots,\ [lab-parameter-N].$
Category	
Samples	$Ask\ the\ values\ of\ all\ the\ following\ parameters:\ thirst,\ polyuria,\ weight-loss.$
ASBRU	$plan-body\ type = any-order\ \ wait-for\ all$ $ask\ [lab-parameter1]\ \ ask\ [lab-parameter2]\ \mid\ ask\ [lab-parameterN]$

Table 49: Pattern 59

Identifier/Label	$I03$ **Explicit action exclusion in specific context**
Arguments	[medical-action],[target-group],[treatment],[lab-parameter],[lab-parameter-value]
IR	$In\ context\ of\ [medical-action]\ applied\ to\ [target-group],$ $[treatment]\ should\ not\ take\ place\ if\ [lab-parameter]\ has\ [value].$
Category	
Samples	$(CBO\ hypertension):\ Het\ opsporen\ van\ verhoogde\ bloeddruk\ heeft\ geen\ zin$ $bij\ personen\ die\ bij\ een\ eventuele\ vastgestelde\ verhoogde\ bloeddruk$ $niet\ in\ het\ indicatiegebied\ "behandelen\ (overwegen)"\ vallen.$
ASBRU	$plan\ [medical-action]$ $plan-body\ type = parallel$ $wait-for\ one$ $plan-activation\ [treatment]$ $plan-activation\ [treatment]-alternative$ $plan\ [treatment]$ $conditions$ $abort-condition:\ [parameter] = [value]$ $plan-body\ to-be-defined$ $plan\ [treatment]-alternative$ $plan-body\ to-be-defined$

Table 50: Pattern 60

Identifier/Label	$I04$ **Post-treatment periodic control within limited time frame**
Arguments	
IR	$After\ [treatment]\ it\ is\ recommended/optional$ $to\ repeat\ [control]\ of\ [parameter]\ every\ [period]\ [time-unit].$
Category	
Samples	$Example\ (CBO):\ Na\ het\ bereiken\ van\ de\ streefwaarden\ is\ controle$ $eens\ per\ zes\ tot\ twaalf\ maanden\ aangewezen.$

ASBRU	$plan\ [medical - procedure]$ $plan - body\ type = sequentially$ $wait - for\ all$ $[treatment]$ $[optional - control]$ $plan\ [optional - control - post - treatment]$ $plan - body\ type = sequentially$ $wait - for\ one$ $cyclical - plan$ $[post - treatment - control]$ $retry - delay:\ [period]\ [time - unit]$ $plan\ [treatment]$ $plan - body\ to - be - defined$ $plan\ [post - treatment - control]$ $plan - body\ to - be - defined$

Table 51: Pattern 61

Identifi er/Label	$I05$ **Action enumeration with success specifi cation**
Arguments	
IR	$To\ complete\ procedure\ [proc],\ one\ of\ the\ following\ steps:$ $[alternative - plan - name - 1],\ or$ $\ldots,\ or$ $[alternative - plan - name - n]$ $must\ be\ executed.$
Category	
Samples	
ASBRU	$plan\ [proc]$ $plan - body\ type = any - order$ $wait - for\ one$ $plan - activation\ [alternative - plan - name - 1]$ $\ldots$ $plan - activation\ [alternative - plan - name - n]$

Table 52: Pattern 62

Identifi er/Label	$I06$ **User activated plan**
Arguments	
IR	$[Action]\ is\ initiated\ by\ doctor.$
Category	
Samples	
ASBRU	$plan\ [plan - name]\ \ [comment]$ $conditions$ $activate - mode: manual$ $plan - body\ to - be - defined$

Table 53: Pattern 63

Identifier/Label	$I07$ **Cancelable plans**
Arguments	[proposition],[lab-parameter],[lab-par-value],[medical-action]
IR	1. $If\ [proposition]\ then\ abort\ [medical-action].$ 2. $If\ [lab-parameter]\ has\ value\ [lab-par-value]\ then\ stop\ [medical-action].$
Category	
Samples	
ASBRU	$plan\ [medical-procedure]\ \ [comment]$ $conditions$ $abort-condition:\ [proposition]$ $plan-body\ to-be-defined$

Table 54: Pattern 64

Identifier/Label	$I08$ **Termination of plans**
Arguments	
IR	$If\ [proposition]\ then\ [medical-procedure]\ finished\ successfully.$
Category	
Samples	
ASBRU	$plan\ [medical-procedure]\ \ [comment]$ $conditions$ $complete-condition:\ [proposition]$ $plan-body\ to-be-defined$

Table 55: Pattern 65

Identifier/Label	$I09$ **User performed plans**
Arguments	
IR	$Doctor\ has\ to\ perform\ [medical-procedure].$
Category	
Samples	
ASBRU	$plan\ [medical-procedure]\ \ [comment]$ $conditions$ $activate-mode: manual$ $plan-body\ user-performed$

Table 56: Pattern 81

Identifier/Label	$I10$ **Pattern conformance to regulation**
Arguments	
IR	$[Action-1]\ should\ follow\ the\ recommendations\ in\ \ [reference\ to\ document\ fragment].$
Category	
Samples	
ASBRU	$plan\ [action-1] comments\ \ refer-to\ [reference\ to\ document\ fragment]$ $plan-body\ to-be-defined$

Table 57: Pattern 67

Identifi er/Label	$I11$ **Pattern excluding membership**
Arguments	
IR	$[Action - 1]$ $should\ not\ be\ part\ of$ $[action - 2]$.
Category	
Samples	$(SIGN):$ $Diagnostic\ imaging\ of\ bones,\ liver\ and\ brain\ should\ not\ be\ part\ of$ $routine\ staging\ and\ should\ only\ be\ carried\ out\ if\ there\ is\ clinical\ or\ radiological$ $suspicion\ of\ metastatic\ disease.$ $(Grade\ B)$
ASBRU	$plan\ [action - 1]$ $conditions$ $filter - condition:\ not(parent.name = [action - 2])$ $plan - body\ to - be - defined$ $plan\ [action - 2]$ $plan - body\ to - be - defined$

Table 58: Pattern 68

Identifi er/Label	$I12$ **Plan execution based on Boolean value condition**
Arguments	
IR	if $[symptom]\ is\ present$ $then$ $execute\ [procedure - 1]$ $else$ $execute\ [procedure - 2]$
Category	
Samples	
ASBRU	$plan\ [handling - [symptom]]$ $plan - body\ type = sequentially$ $wait - for\ all$ $ask\ [symptom]$ $plan - activation\ alternative - procedures$ $plan\ alternative - procedures$ $plan - body\ type = unordered$ $wait - for\ one$ $plan - activation\ [procedure - 1]$ $plan - activation\ [procedure - 2]$ $plan\ [procedure - 1]$ $conditions$ $filter - condition:\ present([symptom])$ $plan - body\ to - be - defined$ $plan\ [procedure - 2]$ $conditions$ $filter - condition:\ not\ present([symptom])$ $plan - body\ to - be - defined$

Table 59: Pattern 69

Identifi er/Label	$I13$ **Assigning action for treating a specifi c target group**
Arguments	[patient-category],[medical-action]
IR	1. $Consider\ [patient\ category]\ for\ applying\ [medical-action].$ 2. $[Method]\ can\ be\ applied\ for\ patients\ in\ [patient\ category].$
Category	
Samples	$(SIGN):\ All\ patients\ with\ a\ clinical\ diagnosis\ of\ lung\ cancer\ should\ be\ considered$ $for\ bronchoscopy\ or\ another\ test\ aimed\ at\ pathological\ confirmation\ unless$ $a\ respiratory\ physician\ or\ thoracic\ surgeon\ has\ indicated$ $it\ would\ not\ be\ appropriate\ (Grade\ C)$
ASBRU	$plan\ [medical-action]$ $conditions$ $filter-condition:\ patient\ in\ [patient\ category]$ $plan-body\ to-be-defined$

Table 60: Pattern 70

Identifi er/Label	$I14$ **Checking of a conjunctive condition**
Arguments	[risk-factors],[patient],[disease]
IR	$If\ the\ following\ [risk\ factors]\ are\ present\ in\ [patient],$ $this\ suggests\ that\ [patient]\ has\ [disease]:$ $[factor-1];\ [factor-2];\ \ldots[factor-n]?$
Category	
Samples	
ASBRU	$plan\ treatment-for-[disease]$ $conditions$ $filter-condition:\ [disease]-risk-factor=true$ $plan-body\ to-be-defined$ $plan\ check-disjunctive-possibility-[disease]$ $conditions$ $complete-condition:\ [disease]-risk-factor=true$ $abort-condition:\ [disease]-risk-factor=false$ $intentions$ $achieve-overall-state:$ $is-known-parameter([disease]-risk-factor)$ $plan-body\ type=any-order$ $wait-for\ all$ $plan-activation\ check-[disease]-risk-factor([factor-1])\ \ \ldots$ $plan-activation\ check-[disease]-risk-factor([factor-2])$ $plan\ check-[disease]-risk-factor$ $arguments\ factor$ $plan-body\ type=sequentially\ wait-for\ all$ $ask\ [factor]$ $if\ present([factor])$ $then\ [disease]-risk-factor\ :=\ true$ $else\ [disease]-risk-factor\ :=\ false$

Table 61: Pattern 71

Identifi er/Label	$I15$ **Parallel action and monitoring**
Arguments	[lab-parameter],[medical-procedure]

IR	$Check\ the\ value\ of\ [lab-parameter]\ during\ application\ of\ [medical-procedure].$
Category	
Samples	
ASBRU	$plan\ parallel-action-monitoring$ $plan-body\ type=parallel$ $wait-for\ all$ $plan-activation\ [monitor-parameter]$ $plan-activation\ execute-[medical-procedure]$ $plan\ [monitor-parameter]$ $conditions$ $filter-condition:\ started-medical-procedure$ $complete-condition:\ completed-medical-procedure=true$ $plan-body\ type=sequentially$ $ask\ [lab-parameter]$ $plan\ execute-[medical-procedure]$ $plan-body\ type=sequentially$ $wait-for\ all$ $started-medical-procedure:=true$ $plan-activation\ medical-procedure$ $completed-medical-procedure:=true$ $plan\ medical-procedure$ $plan-body\ to-be-defined$

Table 62: Pattern 72

Identifi er/Label	$I16$ **Repeated parameter monitoring in parallel with medical action**
Arguments	
IR	$During\ the\ administration\ of\ [medication]$ $the\ doctor\ has\ to\ collect\ the\ values\ of\ [parameter].$
Category	
Samples	
ASBRU	$plan\ repeated-action-monitoring$ $plan-body\ type=parallel$ $wait-for\ all$ $plan-activation\ [collect-[parameter]-values]$ $plan-activation\ [medical-procedure]$ $plan\ [collect-[parameter]-values]$ $conditions$ $filter-condition:\ started([medical-procedure])$ $complete-condition:\ finished([medical-procedure])$ $plan-body\ type=sequentially$ $cyclical-plan$ $ask\ [parameter]$ $end-condition:\ finished([medical-procedure])$ $plan\ medical-procedure$ $plan-body\ [user-performed]$

Table 63: Pattern 73

Identifier/Label	$I17$ **Repeated parameter monitoring with time limit until condition reached**
Arguments	[treatment],[monitoring],[lab-parameter],[interval],[time-unit]
IR	$When\ applying\ [treatment],\ \ [monitoring]\ of\ [lab-parameter]\ should\ be\ repeated$ $every\ [interval]\ [time-unit]\ for\ [time]\ [time-unit]\ \ until\ [target-values]$ $are\ reached\ for\ [lab-parameter].$
Category	
Samples	$(CBO):\ Bij\ medicamenteuze\ behandeling\ is\ twee-\ tot\ vierwekelijkse\ controle$ $aangewezen\ tot\ de\ streefwaarden\ zijn\ bereikt.$
ASBRU	$plan\ repeated-[parameter]-monitoring-with-termination-condition$ $plan-body\ type=parallel$ $wait-for\ [treatment]$ $plan-activation\ monitoring-action$ $plan-activation\ [treatment]-action$ $plan\ monitoring-action$ $conditions$ $complete-condition:\ [parameter]\ in\ [target\ values]$ $plan-body$ $type=sequentially$ $wait-for\ all$ $cyclical-plan$ $plan-activation\ monitoring$ $end-time:\ [time]\ [time-unit]$ $end-condition:\ finished-[treatment]=true$ $retry-delay:\ [interval]\ [time-unit]$ $plan\ monitoring$ $conditions$ $complete-condition:\ [parameter]\ in\ [target\ values]$ $plan-body\ to-be-defined$ $plan\ [treatment]-action$ $plan-body\ type=sequentially$ $wait-for\ all$ $finished-[treatment]\ :=\ false$ $plan-activation\ [treatment]$ $finished-[treatment]\ :=\ true$ $plan\ treatment$ $plan-body\ to-be-defined$

Table 64: Pattern 74

Identifier/Label	$I18$ **Additional action for specific parameter values**
Arguments	[lab-parameter],[lab-parameter-value-1],[lab-parameter-value-2]
IR	$For\ [patients]\ having\ [lab-parameter]\ <=\ [lab-parameter-value-1]\ or$ $[lab-parameter]\ >=\ [lab-parameter-value-2]\ additional\ [procedure]\ must\ be\ executed$ $(at\ the\ end\ of\ the\ [standard-procedure]).$
Category	
Samples	$Example:\ For\ kids\ younger\ than\ 10\ years\ and\ old\ people\ above\ 70\ years\ old,$ $the\ following\ precautions\ must\ be\ taken:\ ....$

ASBRU	$plan\ handling - aditional - [procedure]$ $plan - body\ type = sequentially$ $wait - for\ all$ $ask\ [lab - parameter] - value$ $plan - activation\ special - handling - [age]$ $[procedure]([lab - parameter - value - 1], [lab - parameter - value - 2])$ $plan\ special - handling - [age]$ $plan - body\ type = sequentially$ $wait - for\ all$ $plan - activation\ [standard - procedure]$ $finished - [standard - procedure]\ := \ true$ $plan - activation\ [procedure]([lab - parameter - value - 1], [lab - parameter - value - 2])$ $plan\ [standard - procedure]$ $plan - body\ to - be - defined$ $plan\ [procedure]\ arguments\ v - 1, v - 2$ $conditions$ $filter - condition:$ $[age]\ in\ [min - [age], v - 1]\ or\ \ \ [age]\ in\ [v - 2, max - [age]]$ $plan - body\ to - be - defined$

Table 65: Pattern 75

Identifier/Label	$I19$ **Cancelable treatment**
Arguments	[lab-parameter],[lab-param-threshold],[medical-procedure]
IR	$[Medical\ procedure]\ should\ be\ aborted\ if$ $[lab - parameter]\ has\ value\ greater\ than \mid smaller\ than\ [lab - param - threshold].$
Category	
Samples	
ASBRU	$plan\ [medical - procedure]$ $conditions$ $abort - condition:\ [lab - parameter] greater - than [lab - param - threshold]$ $plan - body\ to - be - defined$

Table 66: Pattern 76

Identifier/Label	$I20$ **Definition of normal values for a parameter**
Arguments	
IR	$[Parameter]\ is\ characterized\ as\ [normal]\ for\ value\ range\ [min - range - 1,\ max - range - 1]$ $or\ in\ [special\ context]\ if\ values\ range\ in\ interval\ [min - range - 2, max - range - 2].$
Category	
Samples	$(CBO):\ High - blood - pressure:\ De\ bloeddruk\ is\ verhoogd$ $wanneer\ deze\ hoger\ is\ dan\ of\ gelijk\ is\ aan\ 140\ mmHg\ systolisch\ en/of\ 90\ mmHg\ diastolisch.$ $Voor\ personen\ van\ 60\ jaar\ en\ ouder\ die\ geen\ diabetes,\ familiaire\ hypercholesterolemie$ $of\ manifeste\ hart -\ en\ vaatziekte\ hebben,\ geldt\ en\ grens$ $van\ 160\ mmHg\ systolisch.$

ASBRU	*domain − def parameter − group* *parameter − def* *name [parameter]* *type continuous* *parameter − def* *name [normal] − [parameter]* *type [parameter] − scale* *qualitative − parameter − def use − as − context* *limits* *context [special − context]* *limit − entry value =′ normal′ [min − range − 1, max − range − 1]* *limits* *context not([special − context])* *limit − entry value =′ normal′ [min − range − 2, max − range − 2]* *comment* *if not([special − context]) and [parameter] in [min − range − 1, max − range − 1]* *then abstr.[parameter] := ′normal′ else* *if [special − context] and [parameter] in [min − range − 2, max − range − 2]* *then abstr.[parameter] := ′normal′ else* *abstr.[parameter] := ′abnormal′*

Table 67: Pattern 77

Identifi er/Label	*I21* **Intention-action plan**
Arguments	
IR	1. *[Medical procedure] has to be performed in order to achieve [goal − state].* 2. *For achieving [purpose − state] it is optional/recommended/obligatory to perform [procedure].*
Category	
Samples	*(CBO)* : 1. *Bij behandeling wordt gestreefd naar een diastolische bloeddruk < 90 mmHg en (afhankelijk van de leeftijd) een systolische bloeddruk < 140 mmHg of < 160 mmHg* 2. *Om de grootte/uitbreiding van de tumor zo goed mogelijk te kunnen voorspellen worden vergrotingsopnamen geadviseerd.*
ASBRU	*plan [medical − procedure]* *intentions* *achieve final state [goal]* *plan − body to − be − defined*

Table 68: Pattern 78

Identifi er/Label	*I22* **Mandatory action**
Arguments	
IR	*The [medical − procedure] consists of the following [mandatory] actions :* *[step − 1], [step − 2], ..., [step − n].*
Category	
Samples	*Maternal prenatal testing should include* *ABO and Rh(D) typing and a serum screen for unusual isoimmune antibodies.*
ASBRU	*plan [medical − procedure]* *plan − body type = sequentially* *wait − for all* *plan − activation step − 1* *plan − activation step − 2* *...* *plan − activation step − n*

Table 69: Pattern 79

Identifier/Label	$I23$ **Optional action enumeration**
Arguments	
IR	1. $[Medical - procedure]:\ [step-1];\ step-2];\ ...;\ [step-n].$ 2. $The\ [medical - procedure]\ consists\ of\ the\ following\ [optional]\ actions:$ $[step-1],\ [step-2],\ ...,\ [step-n].$
Category	
Samples	$(CBO\ diabetes):\ Behandeling\ met\ insuline\ (facultatief)$ $Start\ met\ insuline\ bij\ onvoldoende\ regulatie\ met\ orale\ middelen.$ $Leer\ patient\ vooraf\ zelfcontroles.$ $Geef\ instructies\ over\ beleid\ bij\ hypoglycemie.$ $Verwijs\ zo\ nodig\ naar\ dietist.$
ASBRU	$plan\ [medical - procedure]$ $plan - body\ type = sequentially$ $wait - for - optional - subplans\ no$ $plan - activation\ step - 1$ $plan - activation\ step - 2$ $...$ $plan - activation\ step - n$

Table 70: Pattern 80

Identifier/Label	$I24$ **Pattern constraint on action executions**
Arguments	
IR	$In\ [context],\ [action-1]\ must/can\ be/must\ not\ be$ $executed\ before/after/during\ [method].$
Category	
Samples	
ASBRU	$plan\ [procedure]$ $plan - body\ type = parallel \quad wait - for\ all$ $[action-1]$ $[action-2]$ $plan\ [action-1]$ $conditions$ $filter - condition:\ [action-2]\ execution\ state\ constraint$ $plan - body\ to - be - defined$ $plan\ [action-2]$ $plan - body\ to - be - defined$

Table 71: Pattern 82

Identifier/Label	$I26$ **Constraint on ordering of actions**
Arguments	
IR	$From\ [action-1],\ [action-2],\ ...,\ [action-n],$ $[action-1]\ has\ to\ finish\ first.$
Category	

Samples	
ASBRU	$plan\ [ordered-action-set]$ $plan-body\ type = parallel$ $wait-for\ all$ $[action-1]$ $[other-actions]$ $plan\ [other-actions]$ $conditions$ $filter-condition:\ done-action-1]$ $plan-body\ type = parallel$ $wait-for\ one$ $[action-2]$ $\ldots$ $[action-n]$ $plan\ [action-1]$ $plan-body\ to-be-defined$

Table 72: Pattern 83

Identifi er/Label	$I27$ **Explicit enumeration of parameters**
Arguments	
IR	$In\ [context],\ the\ following\ [parameters]\ must\ be\ evaluated/monitored.$
Category	
Samples	$(CBO\ breast\ cancer):\ Bij\ histologisch\ onderzoek\ moet\ worden\ gelet\ op\ de$ $uitbreiding\ van\ de\ afwijking\ en\ de\ relatie\ tot\ de\ resectieranden,$ $bij\ voorkeur\ de\ minimale\ afstand\ tot\ de\ resectieranden,$ $voorts\ naar\ de\ differentiatiegraad,\ het\ type\ DCIS$ $en\ de\ aan-\ of\ afwezigheid\ van\ micro-invasie.$
ASBRU	$plan\ [ordered-action-set]$ $plan-body\ type = parallel$ $wait-for\ all$ $[action-1]$ $[other-actions]$ $plan\ [other-actions]$ $conditions$ $filter-condition:\ done-action-1]$ $plan-body\ type = parallel$ $wait-for\ one$ $[action-2]$ $\ldots$ $[action-n]$ $plan\ [action-1]$ $plan-body\ to-be-defined$

Table 73: Pattern 84

Identifi er/Label	$I28$ **Pattern success rate for action**

Arguments	
IR	$In\ [context],\ [procedure]\ has\ [100\%]\ chance\ of\ [success/failure].$
Category	
Samples	$(CBO\ breast\ cancer):\ Verwijdering\ van\ de\ gehele\ borstklier\ (ablation\ mammae)$ $geeft\ een\ vrijwel\ 100\%\ genezingskans.$
ASBRU	

Table 74: Pattern 85

Identifi er/Label	$I29$ **Pattern preferences of actions**
Arguments	
IR	$In\ [context],\ the\ preferred\ method\ of\ performing\ [procedure]\ is\ by\ [method].$
Category	
Samples	$(CBO\ breast\ cancer):\ Het\ merendeel,\ naar\ schatting\ 80-85\%\ van\ het\ DCIS,$ $is\ niet\ palpabel\ en\ wordt\ op\ basis\ van\ de\ mammografische\ bevindingen\ ontdekt,$ $meestal\ aan\ de\ hand\ van\ geclusterde\ microcalcificaties.$
ASBRU	$plan\ [procedure]$ $plan-body\ type=parallel$ $wait-for\ one$ $[action-1]$ $[other-actions]$ $plan\ [other-actions]$ $preferences$ $[low]$ $plan-body\ to-be-defined$ $plan\ [method]$ $preferences$ $[high]$ $plan-body\ to-be-defined$

D List of Background Patterns

Although chapter 7 gives a number of background patterns, such as ”process patterns”, ”causal patterns’, ”correlation of signs and symptoms”, in this section we give a set of patterns that are at a finer level of granularity and closer to control patterns. They are in the same format as the control patterns, so that it is easier to compare them.

The major difference between background patterns and control patterns is their different emphasis on the knowledge aspect and control aspect, respectively. Essentially, the background patterns describe fragments of medical knowledge that typically is not directly formalizable, while the control patterns emphasize the control aspect (means, ends and constraints to perform an action) and most of them can be formalized.

Table 75: Pattern 86

Identifi er/Label	*B*01 **Background knowledge**
Arguments	[action], [outcome], [effect], [medication], [disease], [symptom], [condition], [action], [proposition], [goal], [medical-context]
IR	
Category	
Samples	
ASBRU	

Table 76: Pattern 56

Identifi er/Label	*B*02 **Background patterns - causal linking**
Arguments	[action], [outcome], [effect], [medication], [disease], [symptom], [condition], [action], [proposition], [goal], [medical-context]
IR	1. *Because [condition], [action] is to be performed.* 2. *[Proposition − 1] because [proposition − 2].* 3. *When [proposition − 1] it means that [proposition − 2].* 4. *Because of [proposition − about − parameter], [goal] cannot be achieved by performing [action].* 5. *Because [goal] is difficult to achieve by [procedure], [action] needs to be performed.* 6. *Because of [medical − context], [action] must be done.*
Category	
Samples	
ASBRU	

Table 77: Pattern 87

Identifi er/Label	*B*03 **Background pattern - evidence and negative evidence**
Arguments	[neg-evidence],[med-action],[quantifi cation],[disease]
IR	1. *There is [neg − evidence] that [med − action] results in [quantification] risk of [disease].*
Category	
Samples	*(CBO − BC) : [Grade 3]There are no signs that either primary or secondary breast reconstruction results in a higher risk of recurrent breast cancer.* *(CBO − BC) : There are no indications to suggest that a skin − sparing mastectomy followed by immediate reconstruction leads to a higher risk of local or systemic recurrence of breast cancer.C Kroll 90, 92 > C Kroll 90, 92*
ASBRU	

Table 78: Pattern 44

Identifier/Label	$B04$ **Background pattern - reference to table/diagram**
Arguments	[action], [reference]
IR	$1.\ For\ instructions\ on\ how\ to\ do\ [action],\ consult\ [reference\ to\ plan/table/diagram].$ $2.\ The\ values\ of\ [parameters]\ listed\ in\ [conclusion/table/appendix]\ are\ necessary$ $for\ [action]\ to\ be\ executed.$ $3.\ The\ body\ of\ [specification-element]\ can\ be\ found\ in$ $[reference-to-specification-element].$
Category	
Samples	$3.\ (NGC):\ For\ more\ information\ about\ sentinel\ lymph\ node\ biopsy,$ $please\ refer\ to\ Institute\ for\ Clinical\ Systems\ Improvement\ Technology\ Assessment\ nr45$ $"Lymphatic\ Mapping\ with\ Sentinel\ Lymph\ Node\ Biopsy\ for\ Breast\ Cancer"$ $(available\ from\ Institute\ ...).$
ASBRU	

Table 79: Pattern 46

Identifier/Label	$B05$ **Background - declarations and definitions of parameters**
Arguments	[medical-context], [lab-parameters]
IR	$1.\ In\ [medical-context],\ the\ following\ [lab-parameters]\ must\ be\ evaluated/monitored.$ $2.\ [Lab-parameters]\ are\ relevant\ for\ [medical-context].$
Category	
Samples	
ASBRU	

Table 80: Pattern 55

Identifier/Label	$N01$ **Optional exclusion**
Arguments	[action], [medical-context]
IR	$[Action]\ is\ not\ necessary\ for\ [medical-context].$
Category	
Samples	$Ex\ (NGC\ guideline):\ Axillary\ dissection\ is\ not\ usually\ necessary$ $for\ intraductal\ carcinoma\ in\ situ\ (DCIS).$
ASBRU	

Table 81: Pattern 54

Identifier/Label	$N02$ **Defining default action**
Arguments	[medical-action], [medical-action-choice]
IR	$[Medical-action]\ is\ the\ default\ for\ [medical-action-choice].$
Category	
Samples	$Ex\ (NGC\ guideline):\ Traditionally,\ axillary\ dissection\ has\ been\ the\ standard\ of\ practice.$
ASBRU	

Table 82: Pattern 53

Identifier/Label	$N03$ **Quantification of action effects**
Arguments	[situation-quantifier, [action], [result-value]
IR	$In\ [situation - quantifier]\ cases\ when\ [action]\ is\ applied,\ the\ result\ is\ [result - value].$
Category	
Samples	
ASBRU	

Table 83: Pattern 51

Identifier/Label	$N04$ **Exclusion of action from a group of actions**
Arguments	[action-1],[action-2]
IR	$[Action - 1]\ should\ not\ be\ part\ of\ [action - 2].$
Category	
Samples	
ASBRU	

Table 84: Pattern 47

Identifier/Label	$N05$ **Restriction on execution states of groups of actions/plans**
Arguments	[group-of-actions-1], [group-of-actions-2]
IR	1. $[Group - of - actions - 1]\ has\ to\ finish\ before/after\ [group - of - actions - 2].$ 2. $[Group - of - actions - 1]\ has\ to\ be\ inactive\ during/after/before\ [group - of - actions - 2].$
Category	
Samples	
ASBRU	

Table 85: Pattern 48

Identifier/Label	$N06$ **Rate of success for action**
Arguments	[action], [result-quantifier]
IR	$[Action]\ has\ [result - quantifier]\ chance\ of\ success/failure.$
Category	
Samples	
ASBRU	

Table 86: Pattern 49

Identifier/Label	N07 **Explicit preference for action**
Arguments	[context],[action],[preferred-subaction]
IR	1. In [context], the preferred method of performing [action] is by [preferred − subaction]. 2. [Medical − procedure] should preferably include : [preferred − action] and not [alternative − action]. 3. In [medical − procedure] action [preferred − action] is preferred to [alternative − action].
Category	
Samples	Ex(CBO) : Bij gediagnosticeerde verhoogde bloeddruk wordt eerst gekozen voor leefstijladviezen en niet − medicamenteuze behandeling.
ASBRU	1. plan [medical − procedure] preferences preferred − [medical − procedure] plan − body type = sequentially wait − for one preferred − [medical − procedure] alternative − [medical − procedure] plan preferred − [medical − procedure] plan − body type = sequentially wait − for all step − 1 step − 2 ... step − n plan alternative − [medical − procedure] plan − body to − be − defined 2, 3. plan [medical − procedure] plan − body type = parallel wait − for one preferred − [medical − procedure] alternative − [action] plan preferred − [medical − procedure] plan − body type = sequentially wait − for all preferred − action − start := true plan − activation [preferred − action] preferred − action − start := false plan [preferred − action] plan − body to − be − defined plan [alternative − action] conditions filter − condition : preferred − action − start = false abort − condition : preferred − action − start = true plan − body to − be − defined

Table 87: Pattern 50

Identifier/Label	N08 **Assigning preferred-action to target group**
Arguments	[patient-category],[action]
IR	1. Consider [patient category] for applying [action]. 2. [Action] can be applied for patients in [patient category].
Category	

Samples	
ASBRU	

Table 88: Pattern 52

Identifier/Label	$N10$ **Support for hypothesis**
Arguments	[proposition], [proposition-1], [hypothesis-condition]
IR	$[Proposition] := [Proposition - 1]\ supports\ [Hypothesis - condition]$
Category	
Samples	
ASBRU	

E List of Linguistic Patterns

The most frequently encountered linguistic pattern templates can be mapped to background knowledge. In a few cases :

```
1. [common_op, med_intervention, rel_op, med_context]
2. [decomp_op, med_intervention, assoc_rel_op, med_intervention]
3. [disease, assoc_rel_op, disease]
4. [disease, assoc_rel_op, disease_attr]
5. [evaluation, med_effect]
6. [lab_parameter, assoc_rel_op, diagnostic]
7. [med_action, act_op, med_action]
8. [med_action, assoc_rel_op, body_part]
9. [med_action, assoc_rel_op, disease, rel_op, med_intervention]
10. [med_action, assoc_rel_op, disease]
11. [med_action, decomp_op, med_action, med_action, med_action]
12. [med_action, prescription]
13. [med_action, rel_op, body_part]
14. [med_attr, disorder]
15. [med_context, med_condition]
16. [med_context, recommendation, med_action]
17. [med_context, treatment]
18. [med_effect, assoc_rel_op, med_condition]
19. [med_intervention, diagnostic]
20. [med_intervention, logical_rel_op, med_intervention]
21. [med_intervention, prescription]
22. [med_intervention, rel_op, med_intervention]
23. [med_intervention, temp_rel_op, med_intervention]
24. [quantification, lab_parameter]
25. [result_op, assoc_rel_op, disease]
26. [target_group, assoc_rel_op, disease]
27. [target_group, assoc_rel_op, lab_parameter]
28. [target_group, assoc_rel_op, symptom]
29. [target_group, recommendation, med_action]
30. [treatment, assoc_rel_op, diagnostic]
31. [treatment, assoc_rel_op, disease]
32. [treatment, assoc_rel_op, med_action]
33. [treatment, assoc_rel_op, med_condition]
34. [treatment, assoc_rel_op, med_context]
35. [treatment, assoc_rel_op, symptom]
36. [treatment, assoc_rel_op, target_group]
37. [treatment, generic_op, diagnostic]
38. [treatment, generic_op, prescription]
```