LASSIE

Modeling Treatment Processes Using Information Extraction

Conducted for the purpose of receiving the academic title
'Doktorin der Sozial- und Wirtschaftswissenschaften'

Supervisor
Silvia Miksch
Institute of Software Technology & Interactive Systems [E188]
Submitted at the Vienna University of Technology
Faculty of Informatics

by

Katharina Kaiser

Vienna, October 28, 2005

Contents

Abstract

Modeling clinical guidelines and protocols in a computer-interpretable format is a challenging, but burdensome and time-consuming task. Existing methods and tools to support this task demand detailed medical knowledge, knowledge about the formal representations, and a manual modeling. Furthermore, formalized guideline documents mostly fall far short in terms of readability and understandability for the human domain modeler.

In this thesis we propose a methodology to support the human modeler by both automating parts of the modeling process and making the modeling process traceable and comprehensible.

Our methodology called LASSIE[1], represents a novel step-wise procedure that uses Information Extraction to semi-automatically model treatment processes. We have developed several heuristics without the need to apply Natural Language Understanding. Finally, we integrated our heuristics in a form of a framework and applied them to several guidelines from the medical subject of otolaryngology. The framework has been applied to formalize the guidelines in the formal *Asbru* plan representation.

Findings of our evaluation indicate that using semi-automatic, step-wise Information Extraction methods are a valuable instrument to formalize clinical guidelines and protocols.

[1] modeLing treAtment proceSSes using Information Extraction

Zusammenfassung

Klinische Leitlinien und Protokolle in einer computer-interpretierbaren Form zu modellieren ist eine anspruchsvolle, aber mühsame und zeitaufwändige Tätigkeit. Um Methoden und Werkzeuge, die zur Unterstützung dieser Aufgabe entwickelt wurden, anwenden zu können bedarf es medizinischen Wissens sowie Kenntnis der formalen Darstellung. Außerdem muss die gesamte Modellierung von einem menschlichen Modellierer durchgeführt werden. Ein weiteres Problem besteht darin, dass formalisierte Leitlinien für den menschlichen Modellierer meist nicht mehr lesbar und verständlich sind.

In dieser Arbeit stellen wir eine Methode vor, die einerseits den Modellierer durch das Automatisieren von Teilen des Modellierungsprozesses unterstützt und andererseits den gesamten Prozess nachvollziehbar und verständlich macht.

Unsere Methode, LASSIE, stellt ein neuartiges schrittweises Verfahren dar, das mittels Informationsextraktion Therapieabläufe semi-automatisch modelliert. Wir haben dazu einige Heuristiken entwickelt, die ohne die Anwendung von Natural Language Understanding auskommen. Diese Heuristiken haben wir in einem System integriert und dieses für Leitlinien aus der Otolaryngologie (Hals-Nasen-Ohren-Heilkunde) eingesetzt. Wir haben das System zur Formalisierung von medizinischen Leitlinien in der *Asbru* Repräsentation benutzt.

Ergebnisse aus unserer Evaluierung haben deutlich gemacht, dass die Anwendung von semi-automatischen, schrittweisen Informationsextraktionsmethoden ein nützliches Instrument darstellen, um medizinische Leitlinien und Protokolle zu formalisieren.

Chapter 1

Introduction

Errors in healthcare are a leading cause of death and injury. Kohn et al. [47] mention that, for example, preventable adverse events are a leading cause of death in the United States. In their studies they state that at least 44,000 and perhaps as many as 98,000 Americans die in hospitals each year as a result of medical errors. Significantly, there is no data about Central Europe, which is doubtlessly a consequence of our exposure to errors.

This leads us to efforts for increasing the quality of care, concurrently reducing the costs of care, and improving workflow for health care providers.

Clinical guidelines and protocols (CGPs) are ”systematically developed statements to assist practitioners and patient decisions about appropriate healthcare for specific circumstances” [29]. CGPs typically address a specific health condition and provide recommendations to the physician about issues such as who to investigate for the problem, how to investigate it, how to diagnose it, and how to treat it. In Appendix A two example guidelines from the clinical specialty of otolaryngology are shown. Research has shown that, if properly developed, communicated, and implemented, guidelines can improve patient care [38]. CGPs provide not only decision support for the medical personnel (physicians, nursing staff, etc.), patients, and relatives, but also promulgate the most effective and efficient treatment. Therefore, CGPs are an important issue in quality assurance.

Because these guidelines are largely in narrative form, they are sometimes ambiguous and generally lack the structure and internal consistency that would allow execution by a computer. But the method of delivery affects the ability of a guideline to have an effect on clinical decision-making. Guideline implementation strategies that provide patient-specific advice automatically at point-of-care are likely to be effective. Therefore, many researchers have proposed frameworks for modeling CGPs in a computer-interpretable and -executable format [35, 33, 82, 30, 96, 71, 95]. These frameworks are tailored for specific classes of guidelines, specific users, and specific organizations.

Tu and Musen [97] define five tasks that computer-supported guideline-based care must provide: (1) making decisions, (2) sequencing actions and decisions, (3) defining goals (e.g., achieve/maintain/avoid special patient states), (4) interpreting data (i.e., deriving abstract concepts from concrete data), and (5) refining actions. Therefore, CGPs are not only an important issue in quality assurance (during the

therapy), but also in quality control (during and after the therapy). But to execute CGPs in a computer-supported way, the information in the guideline, which is in plain textual form, in tables, or represented in flow charts, has to be formalized. That means, a formal representation is required in order to make the information computable. Thus, several so-called *guideline representation languages* were developed for the assistance of varied guidelines and protocols and to make different kinds of applications possible.

1.1 Problem Domain

In the course of time most of these languages have produced a very extensive syntax. Due to this complexity the modeling process of a CGP into such a language is a very challenging venture. Nowadays, researchers know the requirements to such a language and the resulting complexity. Often, they develop both a particular guideline representation language and methodologies to model guidelines together. But prior languages often lack the tools and methods that support the guideline modeling.

Thus, research has to be directed in such a way where tools and methods are developed for supporting the modeling process especially for existing languages.

Many of these tools that are already developed or in the process of development demand

- Medical background knowledge
- Knowledge about the formal methods (i.e., the guideline representation language)
- Manual modeling task

Furthermore, they mostly lack readability/understandability once the guideline document is formalized in the guideline representation language. If the modeling process is carried out in multiple steps, often the traceability of each step is not possible. Some languages provide no or only rudimental support to select the appropriate syntax for modeling a particular part of a guideline.

Therefore, guideline modelers demand new approaches that can facilitate the modeling process and support them by providing:

- Intelligent methods for a simplified plan modeling process
- Simple formulated modeling templates without the need to know the exact syntax of the representation language
- A traceable modeling process
- (Semi-)automatic modeling.

1.2 Our Approach

Inspired by multi-step modeling methodologies, such as Stepper (see Section 2.2.2), SAGE (see Section 2.1.7), or GEM (see Section 2.1.3), we defined a step-wise approach, which should be applicable for various representation languages (see Figure 1.1) [41]. It facilitates the formalization process by using various intermediate representations that are obtained by stepwise procedures. The intermediate representations are specific templates that can present the desired information. The transformation from one representation to another is applied using Information Extraction (IE) methods. Heuristic methods detect relevant information, which is filled then into the templates' slots for subsequent processing.

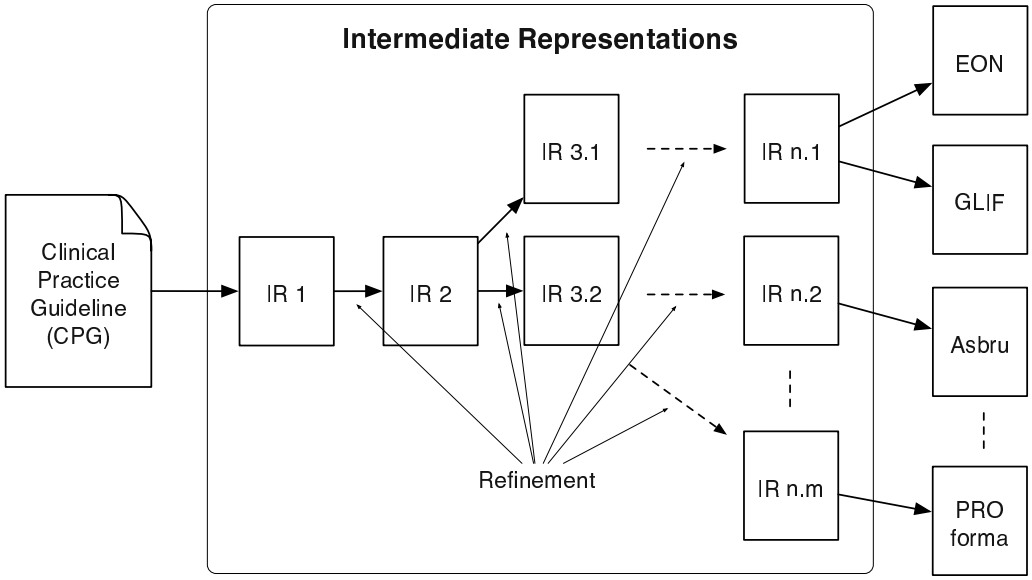


Figure 1.1: Idea of the modeling approach. Starting from the guideline document at the left side the stepwise approach using intermediate representations obtains a formalized guideline in a guideline representation language on the right side.

CGPs present effective treatment processes. One challenge when authoring CGPs is the detection of individual processes and their relations and dependencies. We want to demonstrate that it is possible to formalize processes using IE and develop a framework for modeling guidelines in Asbru (see Section 2.1.1) as proof of concept.

Chapter 2 with related work in guideline representation languages and modeling tools, Chapter 3 with relevant Information Extraction systems, and Chapter 4 describing the common structure of Information Extraction systems, build the basis for the design of the LASSIE methodology. Chapter 5 describes our considerations and definitions of rules for modeling processes in the *Asbru* representation.

Before going into detail concerning the methodology we describe the prerequisites and preprocessing done in Chapter 6. The LASSIE methodology is covered in Chapter 7. Within this chapter we describe how the data obtained by IE can be transformed in Asbru [82].

Chapter 8 describes the evaluation and their results. In Chapter 9 we summarize our work and outline future plans. Finally, Chapter 10 describes our conclusions.

Parts of this work were published at the following conferences: 10th Doctoral Consortium on Advanced Information Systems Engineering (CAiSE'03), June 16-17, 2003, Velden, Austria; Symposium on Computerized Guidelines and Protocols (CGP 2004), April 13-14, 2004, Prague, Czech Republic; European Conference on Artificial Intelligence (ECAI 2004), August 22-27, 2004, Valencia, Spain; 10th Conference on Artificial Intelligence in Medicine (AIME 2005), July 23-27, 2005, Aberdeen, UK; 1st Doctoral Consortium at the Conference on Artificial Intelligence in Medicine (AIME 2005), July 23, 2005, Aberdeen, UK.

Figure 1.2: Lassie, the smart dog.

Chapter 2

Related Work in Formal Guideline Representations

Various guideline representation languages have been and are developed for different purposes, with different goals and intentions, different intended users, and different applications. Shiffman et al. [84] describe the requirements to a guideline knowledge model as following:

- **Comprehensive**, i.e., capable of expressing all the knowledge contained in the guideline
- **Expressively adequate** to convey the complexities and nuances of clinical medicine while remaining informationally equivalent to the original guideline
- **Flexible**, i.e., a useful model must be able to deal with the variety and complexity of guidelines. The representation should permit modeling at high and low levels of granularity, so that guidelines can be interpreted at different levels of abstraction.
- **Comprehensible**, i.e., it should match the stakeholders' normal problem-solving language and allow domain experts to describe their knowledge with little effort
- **Shareable** across institutions
- **Reusability** across all phases of the guideline life cycle

Figure 2.1 presents the history of various guideline modeling methods. We describe the most significant to demonstrate the requirements and considerations for tools and methods that were developed to support the modeling process. We also describe special methodologies and tools used to formalize guidelines.

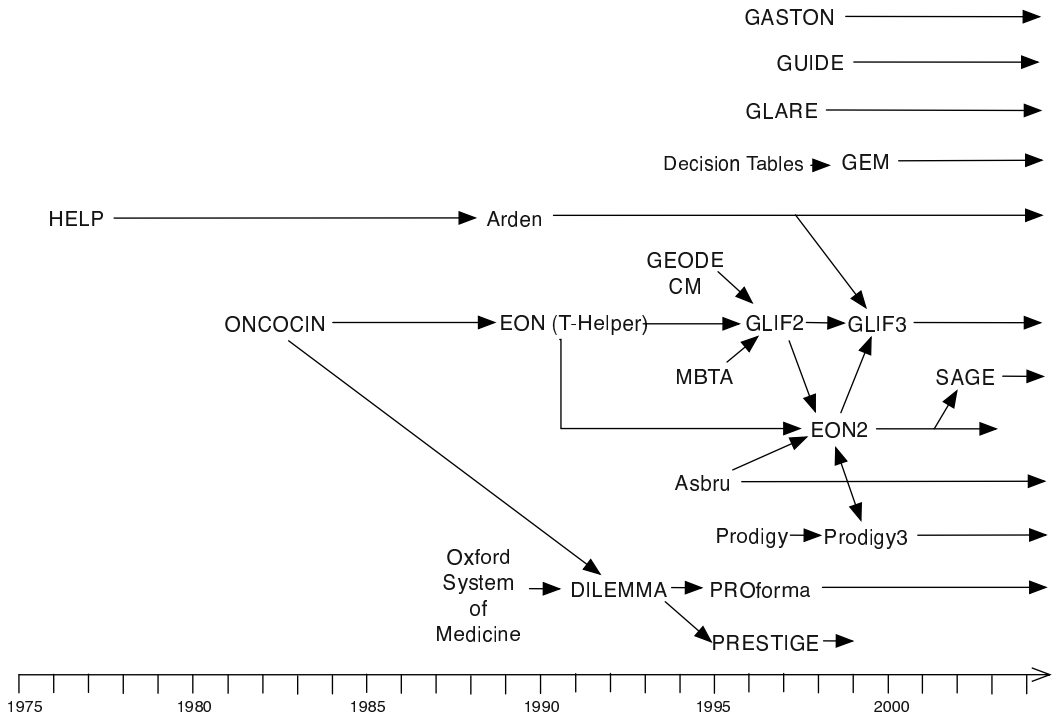


Figure 2.1: History of guideline modeling methods. The guideline modeling methods are positioned on a time axis according to the time at which they started being developed. An arrow between two methods originates from a method that influenced the method depicted next to the arrowhead. (Adapted from [28].)

2.1 Guideline Representation Languages

2.1.1 Asbru

Asbru is a task-specific and intention-based plan representation language that is used in the *Asgaard* project[1] to embody clinical guidelines and protocols as time-oriented skeletal plans [82]. Skeletal plans provide a powerful way to reuse existing domain-specific procedural knowledge, while leaving room for execution-time flexibility to achieve particular goals.

Asbru is designed to represent protocols rather than guidelines. Protocols are local tools that set out specifically what should happen, when and by whom in the care process. They can be seen as the local definition of a particular care process derived from a more generic guideline. Although, there is a difference between protocols and guidelines, it was shown that Asbru can represent not only protocols but also guidelines [81].

Asbru was designed specifically for the set of plan management tasks [63]. It enables the designer to represent both the prescribed actions of a skeletal plan and the knowledge roles required by the various problem-solving methods performing the intertwined supporting subtasks. The major features of Asbru are that

[1] In Norse mythology, *Asgaard* was the home of the gods. It was located in the heavens and was accessible only over the rainbow bridge, called *Asbru* (or *Bifrost*).

- Prescribed actions and states can be continuous
- Intentions, conditions, and world states are temporal patterns
- Uncertainty in both temporal scopes and parameters can flexibly be expressed by bounding intervals
- Plans might be executed in sequence, all plans or some plans in parallel, all plans or some plans in a particular order, or periodically
- Particular conditions and operators are defined to monitor the plans' execution
- Explicit intentions and preferences can be stated for each plan separately

Basically, an Asbru plan can be seen as a template. This template gets instantiated whenever the plan gets executed. Furthermore, more than one instance might be created for a single plan. This pattern can be seen as analogous to the Class–Object relationship in object-oriented Programming.

An Asbru plan consists of a *plan name*, *arguments* (e.g., time annotation), *knowledge roles*, and a *plan body*. Knowledge roles are preferences (i.e., constraints concerning the plan execution), intentions (i.e., goals that have to be achieved, maintained, or prevented), conditions (i.e., constraints concerning the transition between plan states and the switching between plans), effects (i.e., relation between plan arguments and measurable quantities), and plan layout (i.e., the order of executing subplans is defined).

The basic construct is the temporal pattern. One important part of such a temporal pattern is the time annotation, which is represented in Figure 2.2.

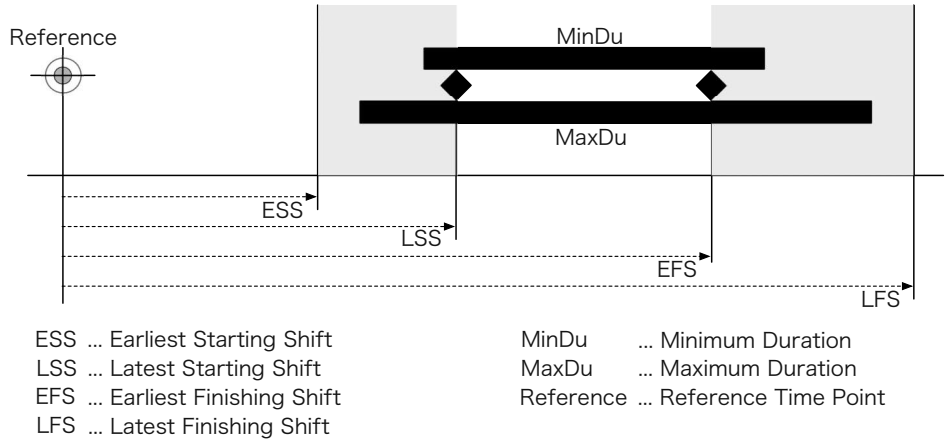


Figure 2.2: Time interval in Asbru. The gray areas indicate the periods when the action has to start and accordingly finish.

Thereby, uncertainties exist in respect of the beginning, the end, and the duration of the interval. Furthermore, no time points for the beginning and the end are defined, but shifts from an arbitrary definable reference point. This point can individually be assigned for any interval.

Furthermore, we can differentiate among six plan states in Asbru: *considered, possible, activated, suspended, completed,* and *aborted*. An instantiated plan can have only one plan state at any one time (see Figure 2.3).

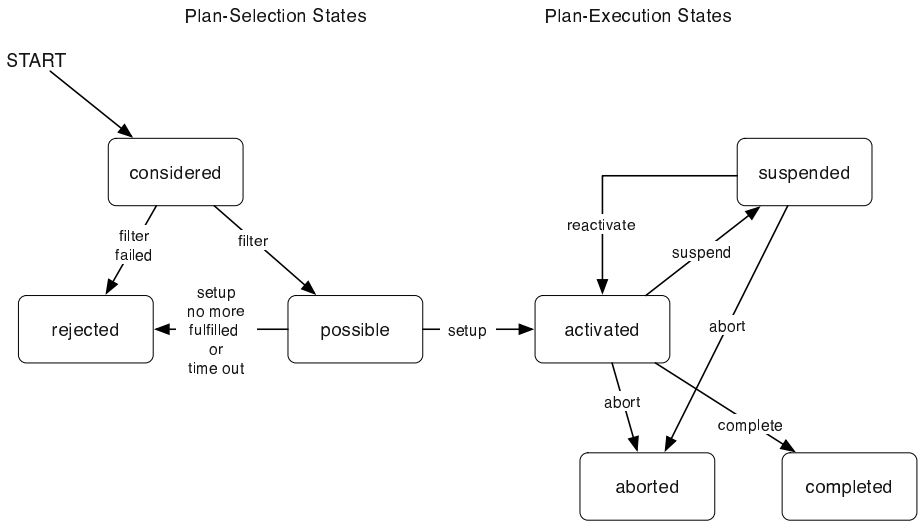


Figure 2.3: Plan states and conditions in Asbru.

For representing different relations among plans, Asbru provides several *plan types*, which describe the behavior concerning the execution of the plans and their synchronization [79]:

- **Sequential plans:** Subsequent plans can be activated if the preceding plan is finished. That means its state must be *completed* or *aborted*.
- **Parallel plans:** Plans are activated at the same starting time. They do not need to be finished at the same time.
- **Any-order plans:** Only one plan can be in state *activated* at one time. Other plans must be in preselection phase, *suspended, completed*, or *aborted* state.
- **Unordered plans:** All plans are executed without any synchronization.
- **Subplans:** Plans can be nested. That means a plan can invoke subplans, which are again subject to a particular ordering by assigning one of the above plan types. A subplan can be activated during its parent's *activated* or *suspended* state and can last during its parent's *completed* or *aborted* state.
- **Cyclical plans:** A plan is repeated several times. The repeat specification can be stated by the period between two iterations or the frequency (i.e., the number or recurrences within a certain period).

Since a plan is represented in XML, it is basically human readable. But understanding such a plan needs a lot of training, semantic and syntactic knowledge about the representation language is cumbersome, and surely not suitable for physicians.

2.1.2 EON

Between 1996 and 2003 the EON model was developed at Stanford University. The EON guideline modeling and execution system forms part of the EON architecture, a component-based suite of models and software components for the creation of guideline-based applications [68].

EON includes an extensible suite of models to represent parts of a clinical practice guideline, domain ontologies, a view of patient data (virtual medical record), and other entities (e.g., those that define roles in an organization). The guideline model (called the Dharma model) defines guideline knowledge structures such as eligibility criteria, abstraction definitions, guideline algorithm, decision models, and recommended actions. The EON guideline execution system obtains patient data through a specified temporal database manager or from user input, and generates recommendations according to the contents of the specific guideline.

Encoding of EON guidelines is done in the Protégé-2000 (see Section 2.2.5) knowledge engineering environment [67]. The encoding process is facilitated by specialized views of the EON guideline model designed to satisfy specific requirements of different classes of guidelines. These requirements are conceptualised in terms of a set of guideline tasks (e.g., decision-making, specification of work to be performed, interpretation of data, setting goals). A guideline developer using EON creates specialized views of the guideline model by selecting modeling solutions to these tasks.

2.1.3 GEM

GEM (the Guideline Elements Model) [84] is an XML-based guideline document model that can store and organize the heterogeneous information contained in practice guideline documents.

GEM has been under continued development at Yale University since 2000. It is an international ASTM standard for the representation of practice guidelines in XML format since 2002 [10].

The GEM project consists of the Guideline Elements Model itself, the editing tool Gem Cutter (see Section 2.2.3), and the quality evaluation method GEM-Q.

GEM is intended to facilitate the translation of natural language guideline documents into a standard computer interpretable format. It encodes considerable information about guideline recommendations in addition to the recommendations themselves, including the reason for each recommendation, the quality of evidence that supports it, and the recommendation strength assigned by the developers. For encoding guideline knowledge no programming knowledge is required, but a markup process is applied.

The authoring process for GEM guidelines takes place in three steps.

1. The GEM document, which has an XML-based syntax, is created based on the original guideline using the GEM Cutter. The elements of the GEM document are then stored in a relational design database.

2. *Knowledge Customization*: meta-information is added, the guideline can be

locally adapted, and abstract concepts of the guideline can be implemented. This step is guided by the *knowledge customization wizard*.

3. *Knowledge Integration* into the clinical workflow depending on local circumstances.

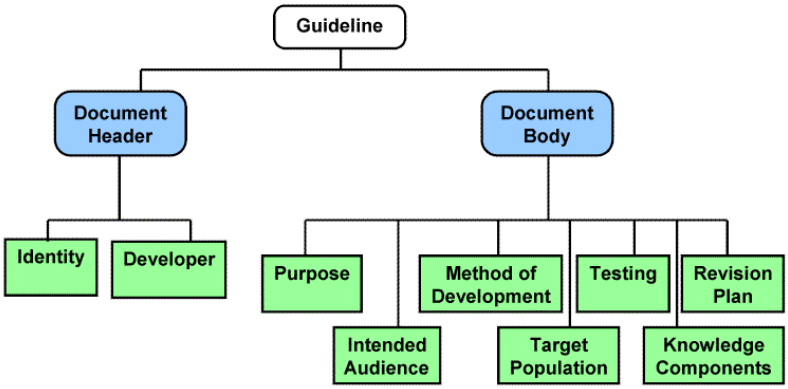


Figure 2.4: GEM Hierarchy [3].

GEM is constructed as a hierarchy with more than 100 discrete elements and more than nine major branches (see Figures 2.4 and 2.5). The majority of the elements describes properties of the guideline as a whole (e.g., title, developer, purpose, target, target population). The content of the guideline can be described in detail by three groups: Recommendations, definitions, and algorithm. *Recommendations* can be conditional or mandatory. For the recommended action the benefit, risk, and cost are stored and reason, evidence quality, strength of recommendation, costs, and so on are annotated. *Definitions* consist of term and term meaning, which are both free text. *Algorithm* consists of an action step, a conditional step, a branch step, and a synchronization step.

2.1.4 GLARE

GLARE stands for GuideLine Acquisition, Representation, and Execution and is a domain-independent system. GLARE technology consists of a representation formalism, a knowledge authoring tool, and a guideline execution tool [94]. GLARE provides consistency checking facilities during the guideline acquisition phase, supports the implementation of temporal reasoning [93], and provides a hypothetical reasoning facility.

The system is based on a modular architecture, which includes an acquisition tool and an execution tool.

The GLARE representation language is designed to achieve a balance between expressiveness and complexity. The formalism consists of a limited, but very focused and clearly understandable set of primitives. It is made up of different types of actions: plans (i.e., composite actions, hierarchically decomposable in their subactions) and atomic actions. Atomic actions can be query actions, decisions, work

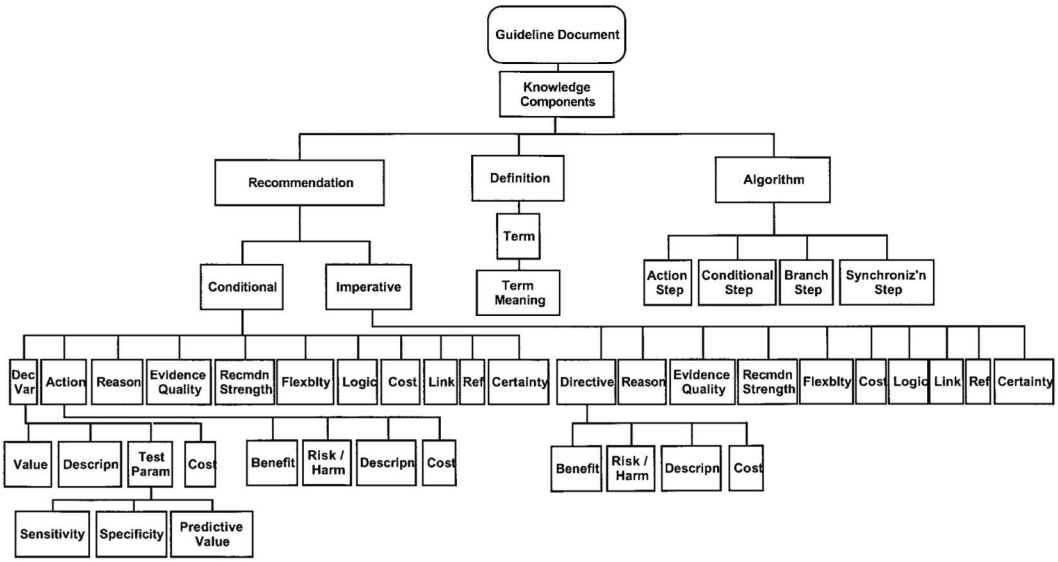


Figure 2.5: GEM: Model of the knowledge components hierarchy [84].

actions, and conclusions. *Work actions* refer to activities of care personnel, similar to user-performed plans in Asbru or actions in PROForma. *Query actions* request information as does the *ask*-parameter in Asbru or enquiries in PROforma. *Conclusions* are the different outcomes of a decision process and resemble choices in GLIF or PROforma. *Decisions* are similarly modeled as in GLIF and PROforma. All actions are linked by *control relations*, which define the order in which they are performed: sequential, concurrent, alternative, or repetition [94].

GLARE provides expert physicians with an ”intelligent” guideline acquisition interface. This provides different types of checks to help in developing a consistent guideline: syntactic and semantic tests verify the ”well-formedness” of a guideline. Further, extended Artificial Intelligence (AI) temporal reasoning techniques are used to check the consistency of temporal constraints imposed between actions [94].

2.1.5 GLIF

GLIF (the Guideline Interchange Format) [71] was developed by the InterMed Collaboratory at Stanford Medical Informatics, Harvard University, McGill University, and Columbia University, and supports sharing of computer-interpretable clinical guidelines across different medical institutions and system platforms. GLIF has a formal representation. It defines an ontology for representing guidelines, as well as a medical ontology for representing medical data and concepts. Tools are under development to support guideline authoring and execution.

GLIF3 is a new version of GLIF designed to support computer-based execution. GLIF3 includes a formal expression language for specifying decision criteria and patient states named GELLO [88]. GELLO, an object-oriented expression language, has a layered patient data model to refer to patient data items defined

by a controlled terminology that includes standard medical vocabularies (such as UMLS), as well as standard data models for medical data implemented.

GLIF3 enables guideline encoding at three levels of abstraction. At a conceptual level the guideline is presented as a flowchart without full details of decisions. The second level, a computable specification that can be verified for logical consistency and completeness, is reached by adding decisions, patient data, and iteration information. The third level, an implementable specification that can be incorporated into particular institutional information systems, is reached by mapping actions to institutional procedures at a certain site and patient data references to a certain electronic patient data record.

GLIF3 has been designed to support computer-based guideline execution: it has a *computable* level of specification, which formally defines logical criteria, definitions of patient data items, clinical actions and flow of the guidelines. The *computable* level of the specification may be regarded as coming between the *abstract flowchart* level and the *implementation level*. The *abstract flowchart* level helps authors and users view and understand a guideline. The *implementation level* includes non-shareable, institution-specific details, which enable guidelines to be incorporated into operational clinical information systems. Shareable components of a guideline are therefore explicitly separated from institution-specific or vendor platform-specific (non-shareable) components.

For representing the guideline's content its algorithm is composed of guideline steps. These can be either *action*, *decision*, *branch*, *synchronization*, or *patient state*.

Action. This step specifies a set of tasks to be performed and may contain several attributes (e.g., iteration information, duration, etc.). Actions can be nested. There are two types of actions: (1) *medically-oriented actions*, which specify a medical task as defined in the Reference Information Model (RIM) layer of GLIF's data model, and (2) *programming-oriented actions*.

Decisions. These steps direct the control flow between alternative steps. There are two types of decision steps: (1) *case steps*, which model deterministic decisions, and (2) *choice steps*, which model non-deterministic decisions.

Branch. This step models concurrency of multiple guideline steps which may be performed in parallel or in any order.

Synchronization. These steps are used in conjunction with branch steps. They mark the place where the different branches of execution meet again and specify the conditions to proceed (i.e., whether all, some, or one of the preceding steps must have been completed before continuing).

Patient State. This step labels its position in the guideline for two purposes. Firstly, it shows the progress of the patient state. And secondly, it serves as an

entry point of the guideline. This means that guidelines can be started at any place which contains a patient state.

2.1.6 PROforma

PROforma [91] has been developed by the Advanced Computation Laboratory, Cancer Research UK since 1992. It is the basis of both a method and a technology for developing and publishing executable clinical guidelines. Applications built using PROforma software are designed to support the management of medical procedures and clinical decision making at the point of care. One aim of the PROforma project is to explore the expressiveness of a deliberately minimal set of modeling constructs.

In PROforma, a guideline application is modeled as a set of tasks and data items. The PROforma task model (Figure 2.6) divides from the keystone (generic task) into four types: plans, decisions, actions, and enquiries.

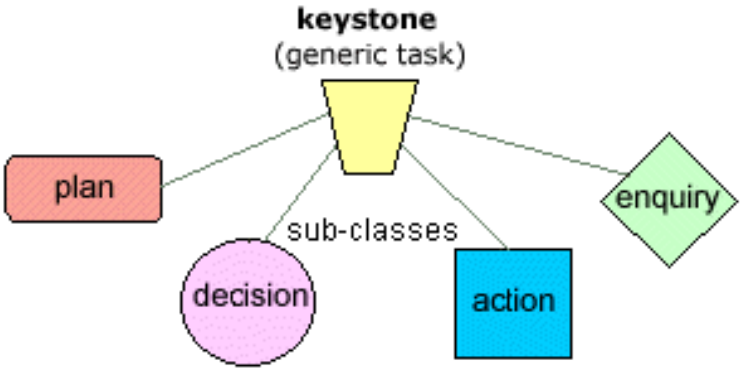


Figure 2.6: The PROforma task model [1].

Plans. They are the basic building blocks of a guideline and may contain any number of tasks of any type, including other plans.

Decisions. They are taken at points where options are presented, e.g., whether to treat a patient or carry out further investigations.

Actions. They represent some procedure that needs to be executed in the external environment (such as the administration of an injection).

Enquiries. They are typically requests for further information or data, required before the guideline can proceed.

All tasks share attributes describing goals, control flow, pre-, and post-conditions. The simple task ontology should make it easier to demonstrate soundness and to teach the language to encoders.

PROforma processes may be represented diagrammatically as directed graphs in which nodes represent tasks and arcs represent scheduling constraints.

PROforma software consists of a graphical editor to support the authoring process, and an engine to execute the guideline specification. The engine can also be used as a tester during the application development phase.

The modeling process consists of two steps:

1. A PROforma editor supports the construction of a guideline in terms of the four task types. Using the icons shown in the diagram, networks of tasks can be composed that represent plans or procedures carried out over time.

2. The resulting populated graphical structure is automatically converted into a database ready for execution.

2.1.7 SAGE

The Standards-based Shareable Active Guideline Environment (SAGE) was introduced in 2002. It is developed by IDX Systems Corporation, Apelon Inc., Intermountain Health Care, Mayo Clinic, Stanford Medical Informatics, and University of Nebraska Medical Center.

Past efforts have gone into developing shared models for representing medical decisions and clinical guidelines. Reuse of a guideline knowledge base also requires that an infrastructure that includes medical record query interface, terminology mediation, and act interface is in place. Furthermore, it often lacks integrating guideline-based decision support with the workflow of care process. The success of clinical decision-support systems (DSSs) depends heavily on how the system is integrated into the care process. In recent years, much interchange and cross-fertilization have taken place in the guideline modeling community. These considerations led to the decision to start development of the SAGE model [19] – The Standards-based Shareable Active Guideline Environment – by giving the opportunity to synthesize prior work and, wherever possible, to establish mappings between the SAGE model and other models. It was introduced in 2002 and is developed by IDX Systems Corporation, Apelon Inc., Intermountain Health Care, Mayo Clinic, Stanford Medical Informatics, and University of Nebraska Medical Center.

In summary, the SAGE project seeks to create a guideline model that

- Uses standardized components that allow interoperability of guideline execution elements with the standard services provided within vendor clinical information systems

- Includes organizational knowledge to capture workflow information and resources needed to provide decision-support in enterprise setting

- Synthesizes prior guideline modeling work for encoding guideline knowledge needed to provide situation-specific decision support and to maintain linked explanatory resource information for the end-user.

The SAGE model organizes guideline recommendations as *recommendation sets* consisting of either Activity Graphs that represent guideline-directed processes or Decision Maps that represent recommendations involving decisions at a time point. Within a particular context, a recommendation may either describe the preferred choice in a management decision or it may recommend a series or actions to be carried out.

SAGE uses a suite of data models and services as interfaces to clinical information systems to achieve interoperability of guideline decision support systems (DSS). It uses standard terminologies and a deployment-driven guideline modeling methodology.

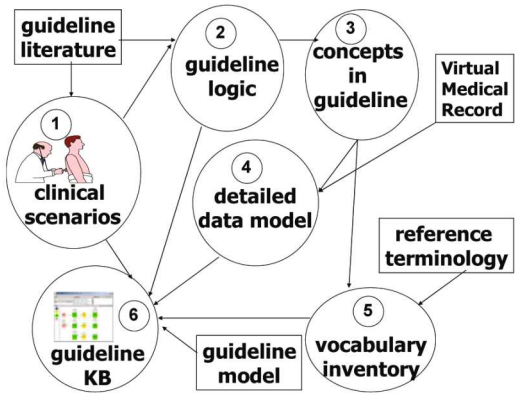


Figure 2.7: Steps in modeling clinical practice guidelines for integration into workflow. The arrows represent information flow [98].

The SAGE guideline knowledge base development methodology consists of six main steps (cp. Figure 2.7).

1. Clinicians must create clinical scenarios that are detailed enough to support integration of executable guideline content into real clinical workflow. For each scenario, user-interface screens are created to simulate the interactions between care providers and the clinical information system.

2. Clinicians analyze the information content of the desired guideline recommendations and extract the knowledge and logic needed to generate these recommendations from guideline texts, medical literature, and clinical expertise. The extraction process requires clinicians to select, interpret, augment, and operationalize guideline statements to disambiguate concepts.

3. Clinical concepts used in the extracted guideline logic are identified.

4. Concepts identified as part of the required guideline logic are instantiated as detailed data models that correspond to constraints on classes of ”virtual medical record” (vMR). The vMR supports a structured data model for representing information related to individual patients, domains for values of attributes in the data model, and queries through which guideline DSS can test the various patient states.

5. Guideline concepts in terms of standard terminologies are specified. To implement a computerized guideline in a particular institution, terms used in a guideline knowledge base to describe patient states must be mapped to terms in that institution's electronic patient record. Standard terminologies, such as SNOMED CT[2] (the Systematized Nomenclature of Medicine Clinical Terms) and LOINC[3] (Logical Observation Ientifiers, Names and Codes), provide the necessary shared semantics for such mappings.

6. Clinical scenarios and guideline logic are translated into a computer-interpretable model of guidelines. The SAGE methodology calls for explicit modeling of guideline usage as part of the executable guideline specification. As such, it assumes that a guideline does not dictate the workflow in a clinic, but the guideline knowledge base specifies how a DSS reacts to events in the care process.

When encoding a guideline for SAGE, clinical experts must interpret the guideline statements and create one or more plans that will support the guideline goals in the specific work environment of their healthcare organization. To achieve this, the recommendation set employs four "nodes": Context, Decision, Action, and Route.

1. **Context Node.** Specifies and declares the assumptions made about the healthcare enterprise work model that are otherwise implicit in every instance of a guideline implementation; their defining attributes specify the trigger events, clinical setting, and patient state.

2. **Action Node.** Models one or more information system activities employed in support of a recommendation set.

3. **Decision Node.** Describes the acquisition of some data (directly from the patient EMR or interactively by asking the clinician) and the employment of a decision model to evaluate branching logic.

4. **Routing Node.** Synchronizes multiple activity paths.

[2]http://www.snomed.org
[3]http://www.regenstrief.org/loinc/

2.2 Tools Supporting the Formalization Process

Several tools have been developed to support the formalization process. They range from markup-based tools, such as DELT/A, Stepper, and GEM-Cutter, to graphical tools using symbols to model diagrams, such as Protégé or the Plan Body Wizard of the DeGel framework, or graphical metaphors, such as AsbruView.

2.2.1 Document Exploration and Linking Tool / Addons (DELT/A)

The Institute of Software Technology and Interactive Systems at the Vienna University of Technology is developing a tool to provide a relatively easy way to translate free text into various (semi-)formal, XML-based representations. It achieves this by displaying both the original text and the translation, and showing the user which parts of the formal code correspond to which elements of the original text. This not only makes it easier to author plans, but also to understand the resulting constructs in terms of the original guideline.

DELT/A[4] [51, 103, 104] (formerly known as Guideline Markup Tool – GMT) provides two main features: (1) linking between a textual guideline and its formal representation, and (2) applying design patterns in the form of macros.

DELT/A allows the definition of links between the original guideline and the target representation, which gives the user the possibility to find out where a certain value in the XML-language notation comes from. Therefore, if someone wants to know the origin of a specific value in the XML file DELT/A can be used to jump to the correlating point in the text file where the value is defined and the other way round.

The second feature of DELT/A is the usage of macros. A macro combines several XML elements, which are usually used together. Thus, using macros allows creating and extending specific XML files more easily through the usage of common design patterns.

DELT/A supports the following tasks:

Authoring and augmenting guidelines. We want to be able to take a new guideline in plain text and create an (XML-based) representation of it, and to add links to the corresponding parts of a guideline to an already existing XML file.

Understanding the (semi-)formal representation of guidelines. For a guideline in a (semi-)formal representation, we want to be able to see where values in the different parts of the representation's code come from, and how parts of the original text were translated into it. This is important not just for knowledge engineers, but also for physicians wanting to get an understanding of the language.

Structuring the syntax of the (semi-)formal representation. DELT/A provides a structured list of elements of the target language – the macros – that need to be

[4] http://ieg.ifs.tuwien.ac.at/projects/delta/

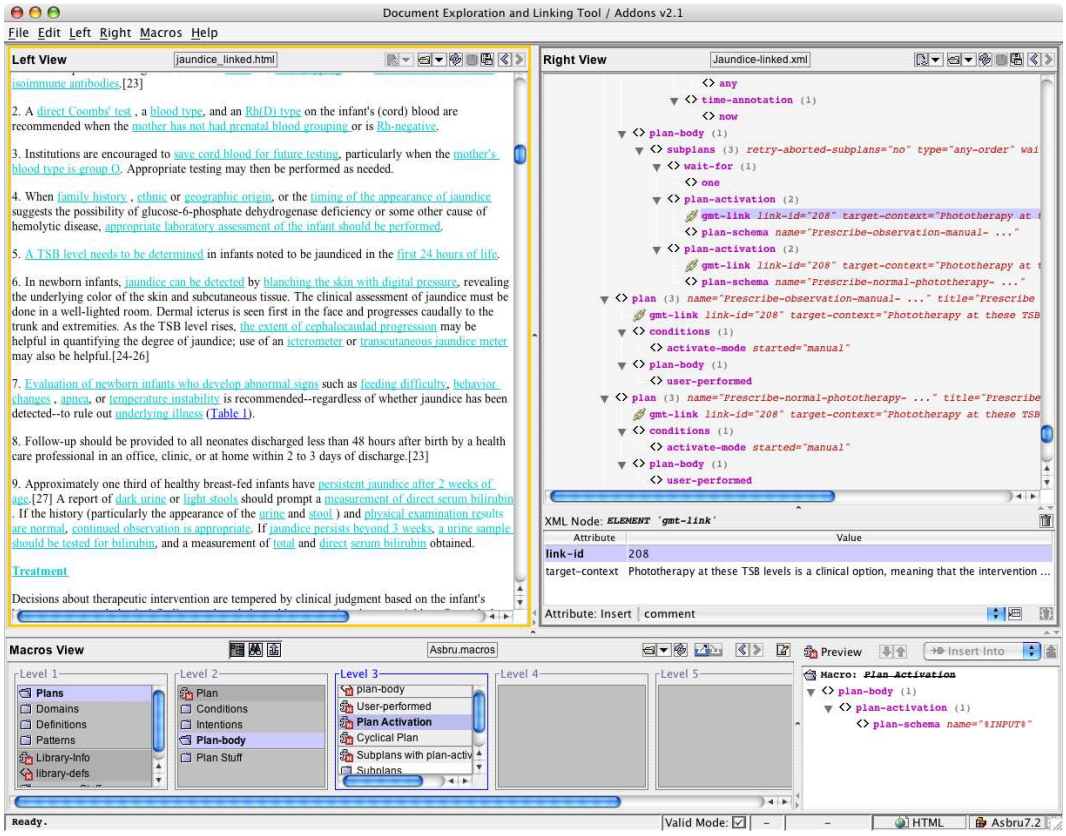


Figure 2.8: Document Exploration and Linking Tool / Addons. The left pane shows a guideline document in textual format. The right pane shows a formal representation of the document. The bottom pane shows the Macros that can be used to support the formalization task by templates of several models.

done in a way that best supports the authoring of plans. This list will also provide a good starting point for teaching material and possible subsets of the language for special purposes.

By means of these features the original text parts need not be stored as part of the target representation elements. The links clearly show the source of each element in the target representation. Additionally, there is no need to produce a guideline in natural language from the target representation since the original text remains unaltered.

DELT/A's user interface (see Figure 2.8) consists of various panes. The top left and right panes provide equivalent views to either edit XML files or HTML files. The Macros pane provides either a structure view, search view, or insertable macros view, as well as a preview of the current macro.

2.2.2 Stepper

Stepper[5] [92, 74] is a mark-up tool for narrative guidelines, developed by the Eu-

[5] http://euromise.vse.cz/stepper-en/aplikace_stepper/index.php

roMISE centrum – Kardio and the University of Economics, Prague, Czech Republic. Its development started in 2001; a first beta version was introduced in 2003. The Stepper project has two main goals:

1. To develop a stepwise method for formalization (in this context, XML transformation) of text documents of clinical guidelines

2. To develop the Stepper tool, an XML editor enhanced with features to support the above method

Stepper has been designed as a document-centric tool, which takes a guideline text as its starting point and splits the formalization process into multiple user-definable steps, each of which corresponds to an interactive XML transformation. The result of each step is an increasingly formalized version of the source document. An embedded XSLT processor carries out non-interactive transformation. Both the mark-up and the iterative transformation process are carried out by rules expressed in a new transformation language based on XML, the so-called XKBT (XML Knowledge Block Transformation). This was because the well-known standard for transformation XSLT did not solve all problems when explicitly expressing transformations of knowledge in each step. Hence, a tailored transformation language was developed.

The transformation process with Stepper consists of six steps:

1. *Input text format.* The format of the original guideline text is XHTML, the XML version of HTML.

2. *Coarse-grained semantic mark-up.* Basic blocks of the text are marked (e.g., headings, sentences) and parts without operation semantics are removed.

3. *Fine-grained semantic mark-up.* Complex sentences are rearranged into simpler ones and background knowledge is added. In addition, a data dictionary is created, which describes the clinical parameters involved.

4. *Universal knowledge base.* The original document is transformed into a universal knowledge base. This involves changing the structure of the document to achieve modularity, which is assumed to involve medical experts in part.

5. *Export-specific knowledge base.* The representation is adapted to ease the export to the target representation. Therefore, an export-specific knowledge base is produced from the universal one.

6. *Target computational representation.* The ultimate format is produced by the knowledge engineer. This step is assumed to be performed fully automatically using XSL style sheets.

By using the Stepper method and tool it is possible to transform CGPs into fragments of operational code (e.g., Java) or into parts of a guideline representation language (e.g., Asbru).

Stepper's main advantage is the documentation of all activities, which allows to easily review the transformation process. Stepper also provides an interface showing the interconnection between the source text and the model.

2.2.3 GEM-Cutter

The *GEM Cutter* [73] is a tool with the aim to facilitate the transformation of CGPs into the GEM format. It was developed by Yale Center for Medical Informatics at Yale University School of Medicine. GEM (the Guideline Elements Model – see Section 2.1.3) is an XML-based guideline document model.

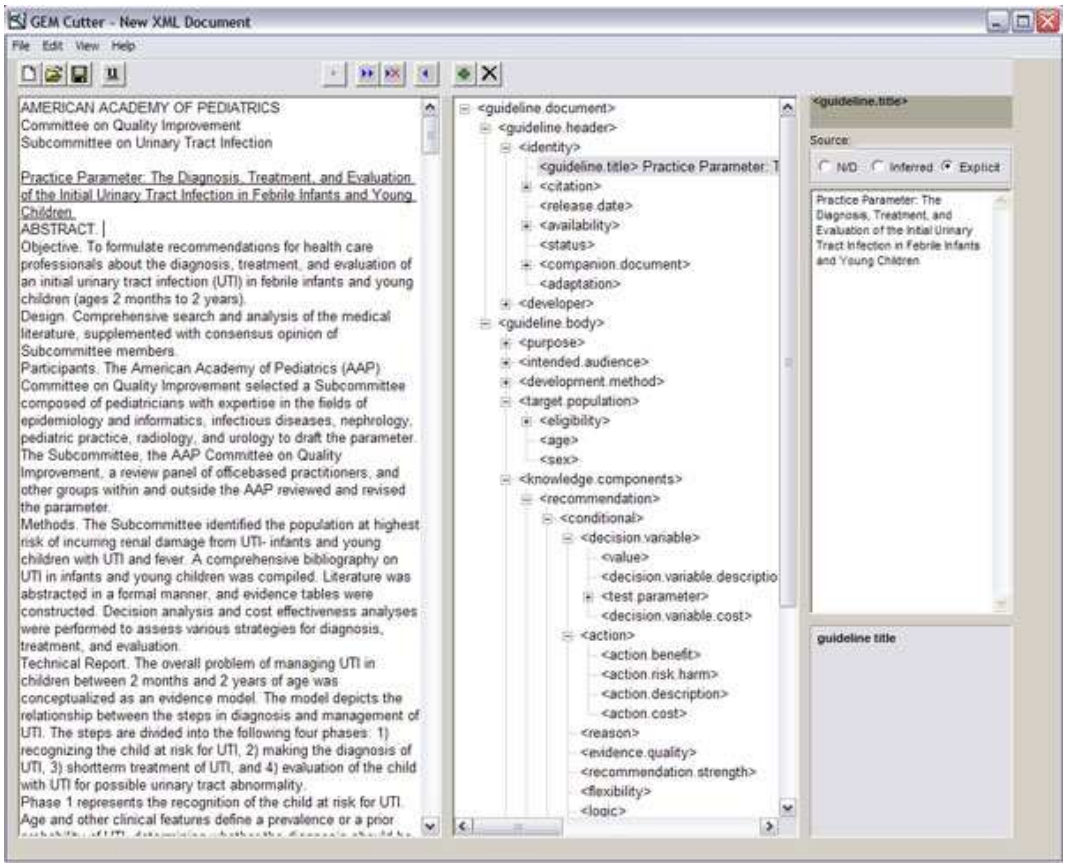


Figure 2.9: GEM Cutter [2]. The left pane shows the guideline text, the middle pane shows the GEM tree segment and the right pane shows the Element Segment.

GEM Cutter's Main Screen consists of three vertical segments (cp. Figure 2.9). In the left pane – the *Guideline Text Segment* – you see the original text of the guideline, in the middle pane – the *GEM Tree Segment* – a tree view of the developing GEM file is displayed, and in the right pane – the *Element Segment* – there is additional, important information about the GEM file.

From the Guideline Text Segment text is copied into elements in the GEM Tree Segment. Text that has been copied is underlined. The GEM Tree Segment displays the contents of the GEM document in tree structure format. Each item on the tree represents a GEM Element, which can contain text and can be edited. The Element Segment consists of various parts: the Element Name Bar, the Element Source Bar, the Element Text Box, and the Element Definitions Box. These contain additional information of the element selected in the GEM document.

2.2.4 Degel – Digital Electronic Guideline Library

Degel[6] [83] is a generic framework with tools to support guideline classification, semantic markup, context-sensitive search, browsing, run-time application, and retrospective quality assessment. It is applicable for any XML-based guideline representation, currently supporting Asbru [82] and GLIF [71]. It supports the gradual migration of free text guidelines to formal representations.

Semantic markup is performed using the Uruz web-based guideline markup-tool, which resembles the DELT/A but does not maintain links between different representations of the guideline. Uruz can also be used to create a guideline document de-novo (i.e., without using any source) by directly writing into the knowledge roles of a selected target ontology. The editor can modify the contents or add new content. This enables implicit knowledge to become more explicit, further facilitating the task of the knowledge engineer who fully formalizes the guideline.

Several features are especially tailored to Asbru, such as the plan-body wizard (PBW), which is used for defining the guideline's control structure. The PBW enables a user to decompose the actions embodied in the guideline into atomic actions and other sub-guidelines, and to define the control structure relating to them (e.g., sequential, parallel, repeated application). The PBW, used by medical experts, significantly facilitates the final formal specification by the knowledge engineer.

To be truly sharable, guidelines need to be represented in a standardized fashion. Thus, Uruz enables the user to embed in the guideline document terms originating from standard vocabularies, such as ICD-9-CM (International Classification of Diseases, Ninth Revision, Clinical Modification) for diagnosis codes, CPT-4 (Current Procedural Terminology) for procedure codes, and LOINC-3 for observations and laboratory tests. In each case, the user selects a term when needed, through a uniform, hierarchical search interface to our Web-based vocabulary server.

2.2.5 Protégé

Protégé[7] is an open source ontology development and knowledge acquisition environment developed by Stanford Medical Informatics [32].

Protégé is a Java tool, which provides an extensible architecture for the creation of customized knowledge-based tools. It assists users in the construction of large electronic knowledge bases. It has an intuitive user interface that enables developers to create and edit domain ontologies and supports customized user-interface extensions, incorporates the Open Knowledge Base Connectivity (OKBC) knowledge model, and interacts with standard storage formats such as relational databases, XML, and RDF.

Protégé is also used to author guidelines in various models (e.g., EON, GLIF, Prodigy, Proforma). Thereby, a part of the modeling can be accomplished using predefined graphical symbols. These symbols are arranged in a diagram and linked by graphs. See Figure 2.10 for an example. The underlying data are entered by web masks.

[6] http://medinfo.ise.bgu.ac.il/medlab/ResearchProjects/RP_DeGeLhtm.htm
[7] http://protege.stanford.edu

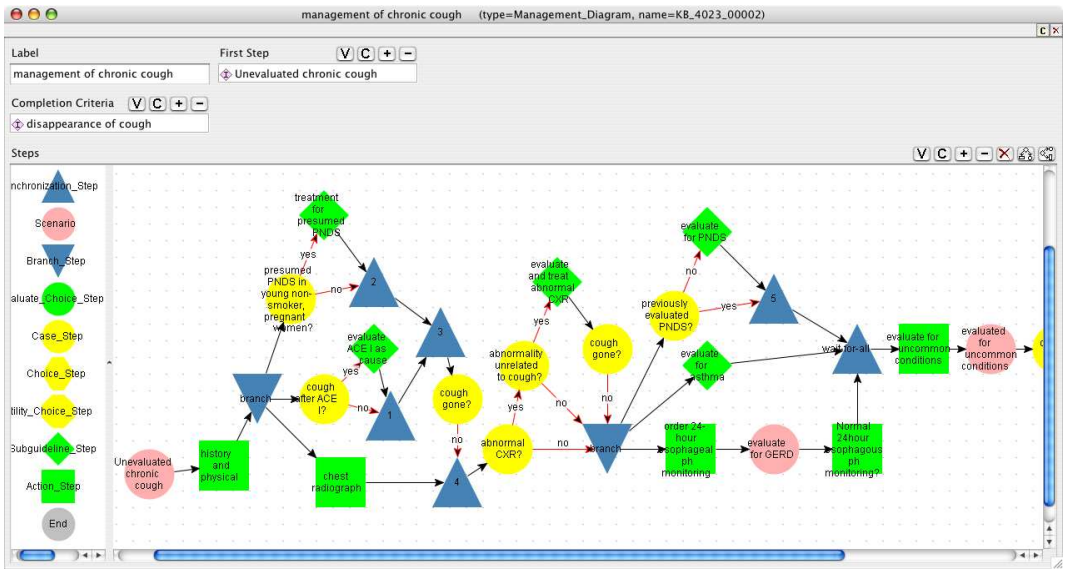


Figure 2.10: View of Protégé being used to author a guideline for managing chronic cough. The guideline model being used in this application is Dharma, part of the EON framework.

2.2.6 AsbruView

Information Visualization is the use of computer-supported interactive visual representations of abstract data to facilitate cognition: The Asgaard/Asbru project focuses on data and plan visualization during the design and execution of guideline and protocol applications. They developed a graphical user interface to Asbru, which supports the development of guidelines and protocols, called AsbruView[8] [64, 52, 48, 49]. Different methods are in development, which are dealing with the visualization of data and plans during the execution phase and are utilizing Focus+Context techniques (e.g., Semantic Depth of Fields (SDOF) [50]).

Asbru is a complex language, which cannot be fully understood by physicians, who have no or hardly any training in formal methods. AsbruView is a tool to make Asbru accessible to physicians, and to give any user an overview of a plan hierarchy. AsbruView is based on visual metaphors to make the underlying concepts easier to grasp. This was done because not only is the notation foreign to physicians, but also the underlying concepts.

AsbruView provides four views: Topological View, Temporal View, SOPO-View, and XML View.

The metaphors and graphical representation of AsbruView have proved to be useful in communicating Asbru's concepts to physicians. Users get a better overview of the therapy steps than from tables, while at the same time being able to see the precise temporal constraints of plans (which is not the case with flowcharts).

[8] http://www.asgaard.tuwien.ac.at/tools/asbruview.html

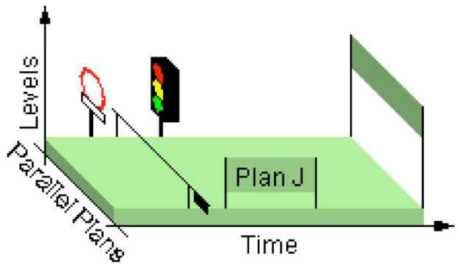


Figure 2.11: AsbruView: Anatomy of a Plan in the Topological View.

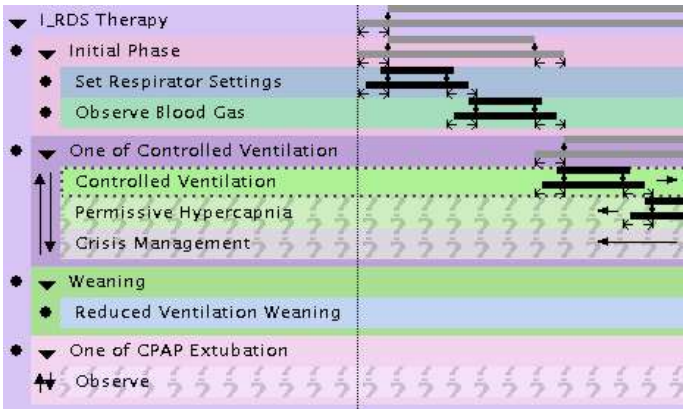


Figure 2.12: AsbruView: Example screenshot of the Temporal View.

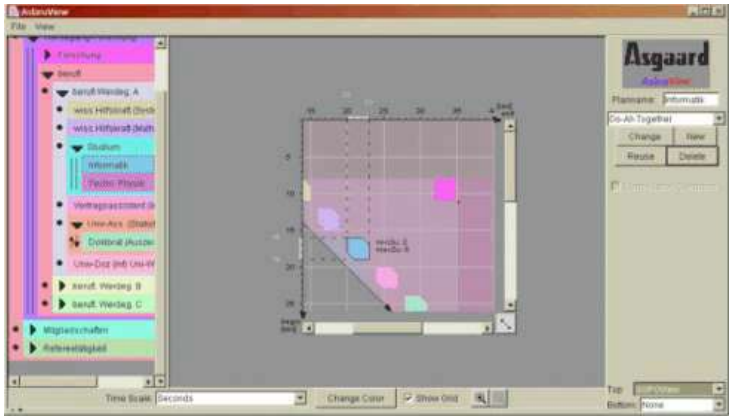


Figure 2.13: AsbruView: Part of an example screenshot of SOPOView.

Chapter 3

Related Work in Information Extraction

Information Extraction (IE) is an emerging NLP technology whose function is to process unstructured, natural language text, to locate specific pieces of information, or facts in the text, and to use these facts to fill a database.

Peshkin and Pfeffer [72] define IE as the task of filling template information from previously unseen text which belongs to a pre-defined domain.

Similar to IE systems are *wrappers*. A wrapper is a program carrying out retrieving information from different repositories, merging, and unifying them. But often wrappers are only applied to the latter two activities. The aim of a wrapper is to locate relevant information in semi-structured data and to put it into a self-described representation for further processing [56]. It seems, as if IE systems and wrappers do just the same, but the application areas are different. Besides, many information resources do not exhibit the rich grammatical structure NLP techniques applied in IE systems are designed to exploit.

The most widespread application area of wrappers is the World Wide Web with its unlimited amount of web sites that are mostly semi-structured. The differences between the structure of each document in the web and the fact that sites are changed periodically makes it obvious that building such programs by hand is not a feasible task. This leads to two main problems in this field: ”wrapper generation” and ”wrapper maintenance” [21].

In this chapter we discuss two different approaches of developing IE and wrapper systems: (1) the Knowledge Engineering approach, and (2) the learning approach.

3.1 The Knowledge Engineering Approach

At the beginning of developing IE systems the systems were customized manually to a given task. Due to the very high costs for computational linguists approaches were demanded to automate this task. Anyhow, the best performing systems for various IE tasks have been hand crafted.

FASTUS

FASTUS [34], developed in the early 1990's, is a (slightly permuted) acronym for Finite State Automaton Text Understanding System, although it does not deal with text understanding but IE. It is a cascaded, nondeterministic finite state automaton.

FASTUS, as used for the fourth Message Understanding Conference (MUC-4), is based on a 4-step processing: (1) triggering, (2) recognizing phrases, (3) recognizing patterns, and (4) merging of incidents. In the first pass, trigger words are searched for in every sentence. Additionally, person's names identified in previous sentences are also treated as trigger words for the remainder of the text. In the second step, noun groups, verb groups, and several critical word classes (e.g., prepositions, conjunctions, relative pronouns) are recognized by a nondeterministic finite state automaton. In this phase no parts-of-speech (POS) tagger is necessary as it was shown not to improve the results, but decreases the system's performance. In the third step, patterns are recognized. Pattern generation was done completely hand-coded. For MUC-4 95 patterns were implemented to extract incidents detected in the documents. In the last step incidents are merged. The application area of MUC-4 has been news articles about terrorist activities. Merging is accomplished for incidents of the same sentence and for incidents remaining at the end of the processing which are merged with those incidents found in previous sentences. In case of incompatible incident types, dates, or locations, merging is blocked.

The basic system is relatively small, although the dictionary used is very large. The manually developed rules were very effective and performed very well.

GE NLTOOLSET

The GE NLTOOLSET [53] is an IE system using a knowledge-based, domain-independent core of text processing tool. The processing of the toolset is divided into three stages: (1) pre-processing, (2) linguistic analysis, and (3) post-processing.

In the pre-processing phase the text is segmented and irrelevant parts are filtered out, phrases that are template activators are identified, and portions of text are marked that could describe discrete events. In the analysis phase parsing and semantic interpretation is performed. The post-processing module selects the templates and maps semantic categories and roles into those templates.

The system's knowledge base consists of a core sense-based lexicon and a feature and function grammar. The core lexicon contains over 10,000 entries, of which a small set is restricted to the application domain. The core grammar consists of 170 rules, again with a small set of MUC-specific rules.

PLUM

PLUM (Probabilistic Language Understanding Model) [11] as used in MUC-4 applied manually generated rules. The system architecture contains a preprocessing module, a morphological analysis module, a parsing module, a semantic interpreter, a discourse processing module, and a template generation module.

Within the *preprocessing module* message boundaries were determined, the header is identified, and paragraph and sentence boundaries are determined. The

morphologic analysis module assigns parts-of-speech (POS) information, whereby the POS tagging is augmented by (automatically trained) probabilistic models for recognizing words of Spanish and English origin. The *parsing module* generates one or more non-overlapping parse fragments spanning the input sentence. These fragments are then processed by the *semantic interpreter*. This module uses semantic components, such as lexical semantics and semantic rules. Lexical semantics are constructed by an automatic case frame induction procedure. They indicate the word's semantic type and predicates pertaining to it. Semantic rules are general syntactic patterns. Their basic elements are 'semantic forms', which can be either entities of the domain, events, or states of affairs. Entities correspond to people, places, things, and time intervals of the domain and arise from noun phrases. Events (who did what to whom) and states of affairs (properties of entities) may be described in clauses. The *discourse module* constructs event objects corresponding to relevant events in the message. Thereby, it must infer indirect relations not explicitly found by the interpreter and resolve any references in the text. The *template generator* then uses the structures created by the discourse module to generate the final templates.

PROTEUS

PROTEUS [105] is a core extraction engine consisting of seven modules (see 3.1): (1) lexical analysis, (2) name recognition, (3) partial syntactical analysis, (4) scenario pattern analysis, (5) reference resolution, (6) discourse analysis, and (7) output generation.

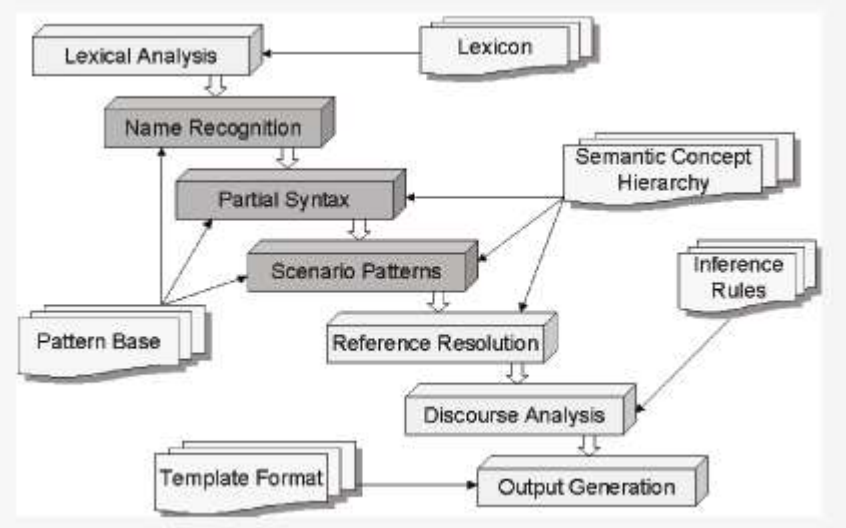


Figure 3.1: Proteus system architecture [105].

The *lexical analysis* module splits the document into sentences and tokens. Each token is assigned a *reading* using dictionaries. Optionally, a parts-of-speech tagger can be invoked to eliminate unlikely readings from tokens. The *name recognition*, *partial syntax*, and *scenario patterns* modules use deterministic, bottom-up, partial parsing, or pattern matching. Patterns are regular expressions. Patterns for

name recognition identify proper names. The *partial syntax* module finds noun phrases and verb phrases. The *scenario patterns* module finds higher-level syntactic constructions. The *reference resolution* module links anaphoric pronouns or their antecedents and merges other co-referring expressions. The *discourse analysis* module builds more complex event structures using higher-level inference rules. Thereby, several clauses contain information about a single complex fact. The *template generation* module performs a transformation of the gathered information into the final template structure.

The pattern acquisition consists of several steps. First, the user enters a sentence containing an event and selects an event template from a list of events. Then, the system applies current patterns to the example to obtain an initial analysis. Thereby, it identifies noun/verb groups and their semantic types and applies a minimal generalization. The system presents the result to the user, who can modify each pattern element (e.g., choose the appropriate level of generalization, make the element optional, remove it). The user then has to specify how pattern elements are used to fill slots in the event template. Now the system builds a new pattern to match the example and compiles the associated action, which will fire when the pattern matches and will fill the slots of the event template. The new pattern is added to the pattern base.

3.2 The Automatic Learning Approach

Manually generating extraction rules is a cumbersome and time-consuming task. Thus, research has been directed towards automating this task. The automatic approach takes a set of documents and outputs a set of extraction patterns by using Machine Learning techniques. Automatic learning systems can be categorized in three groups: (1) *supervised learning systems* [75, 87, 46, 37, 86], where a large set of training data is required to learn the rules using Machine Learning techniques, (2) *semi-supervised learning systems* [78, 4, 106], and (3) *unsupervised learning systems* [77, 22, 89], where rules are learned by a small set of seed rules and an annotated corpus using bootstrapping methods.

To cope with the problems of ”wrapper generation” and ”wrapper maintenance” rule-based methods have been especially popular in recent years. Some techniques for generating rules in the realm of text extraction are called ”wrapper induction” methods. These techniques have proved to be rather successful for IE tasks in their intended domains, which are collections of documents such as web pages generated from a template script [69, 55, 58, 14]. However, wrapper induction methods do only extend well to documents specific to the induced rules.

In semi-automatic wrapper generation Machine Learning approaches are applied. Tools may support the design of the wrapper. Some approaches offer a declarative interface where the user shows the system what information to extract. Representative systems for semi-automatic wrapper generation are WIEN [56, 54], SoftMealy [36], STALKER [70], Lixto [14], and XWrap [58].

Automatic wrapper generation tools use unsupervised learning techniques. Therefore, no training sets are necessary, just a post-generation tuning. Important

Table 3.1: Sample of AutoSlog heuristics.

Linguistic Pattern
<subject>passive-verb
<subject>active-verb
<subject>verb infinitive
<subject>auxiliary noun
passive-verb <direct-object>
active-verb <direct-object>
infinitive <direct-object>
...

systems are Shopbot, RoadRunner, and IEPAD.

3.2.1 Supervised Learning

Supervised learning uses annotated training data to induce extraction rules. Thereby, almost no knowledge about the domain is necessary, but a large set of training data has to be annotated according to the underlying structure of information to be extracted.

AutoSlog

AutoSlog [75] was the first system to learn text extraction rules from training examples. It extracts a domain-specific dictionary of concept nodes for extracting information from text. A concept node is a rule which includes a ”trigger” word or words and a semantic constraint. If the system finds the trigger in the text and the concept node’s conditions are satisfied, the concept node is activated and the concept node definition is extracted from the context.

The system identifies a sentence annotated with a slot filler and semantic tag. Then, it looks up its list of heuristics (for samples see Table 3.1) and sees if any of the heuristics match the part of the sentence containing the slot filler.

Each heuristic handles only a single-slot extraction. AutoSlog also uses a semantic tagger and semantic constraints in the extraction rule; it does not merge similar concept nodes and handles only free text.

PALKA

The PALKA system [46] uses an induction method similar to Mitchell’s candidate elimination algorithm [65]. It produces the extraction rule as a pair of a meaning frame and a phrasal pattern, called *Frame-Phrasal pattern structure (FP-structure)*.

If existing rules cannot be applied, PALKA creates a new rule and tries to generalize it with existing ones to include a new positive instance. It specializes existing rules to avoid a negative instance if it generates the wrong interpretation when applying existing rules.

CRYSTAL

The CRYSTAL system [87] takes texts, which are processed by a syntactic parser. It needs training documents annotated by a domain expert, as well as a semantic hierarchy. CRYSTAL starts learning by creating the extraction rules for each instance of the target event in the training data. Using inductive learning it finds the most similar pair or rules and merges them together by finding the most restrictive constraints that cover both rules.

LIEP

The LIEP system [37] uses heuristics in a manner similar to AutoSlog, but learns multi-slot rules and cannot handle single slot extraction. It allows a user to interactively identify events in text, based on the assumption that a large annotated training corpus is hard to obtain. For each potential training sentence, entities of interest (e.g., people, companies, etc.) are identified.

LIEP tries to choose extraction patterns that will maximize the number of extractions of positive examples and minimize spurious extractions. If a new example cannot be matched by a known pattern, LIEP attempts to generalize a known pattern to cover the example. If the generalization is not possible or the resulting pattern decreases the quality, a new pattern is constructed.

WHISK

The WHISK system [86] uses Machine Learning algorithms known as covering algorithms [62].

WHISK uses a covering algorithm to induce regular expressions in a top-down induction. It begins with the most general rule and continues by progressively specializing the available rules. The process stops when a set of rules is generated covering all positive training examples. Afterwards, post-pruning is achieved to remove rules that generate overfitting.

RAPIER

The RAPIER (Robust Automated Production of Information Extraction Rules) system [18] uses pairs of sample documents and filled templates to induce pattern-match rules that directly extract fillers for the slots of the template. It employes a bottom-up learning algorithm in order to limit search without imposing artificial limits on the constants to be considered, and in order to prefer high precision by preferring more specific rules. Pairs of rules are randomly chosen and a beam search is achieved to find the best generalization of the two rules, taking a least general generalization (LGG), then adding constraints until the proposed rule operates

correctly on the training data. RAPIER can only handle single-slot extraction on semi-structured text.

GATE – ANNIE

GATE [25] is a framework and graphical development environment which enables users to develop and deploy language engineering components and resources. A set of reusable processing resources is packed together to ANNIE, A Nearly-New IE system. These processing resources can be used individually or coupled together with new modules.

ANNIE consists of a tokeniser, a sentence splitter, a parts-of-speech tagger, a gazetteer, a finite state transducer, an orthomatcher, and a coreferencer. The *tokeniser* splits text into tokens (i.e., numbers, punctuation, symbols, and words of different types). The *sentence splitter* segments the text into sentences, which are the input of the *tagger*. It produces a parts-of-speech tag as an annotation on each word or symbol. The *gazetteer* consists of lists of cities, organizations, days of the week, and so on. The *semantic tagger* consists of handcrafted rules, which describe patterns to match and annotations to be created as a result. The *orthomatcher* performs co-reference, or entity tracking, by recognizing relations between entities. The *coreferencer* finds identity relations between entities in the text.

GATE provides easy-to-use and extendable facilities for text annotation to annotate required training data for NLP algorithms. The annotation can be done manually by the user or semi-automatically by running some processing resources over the corpus and then correcting and adding new annotations manually. Depending on the information to be annotated, some ANNIE modules can be used or adapted to bootstrap the corpus annotation task.

GATE addresses the complete range of issues in NLP application development in a flexible and extensive way. It promotes robustness, re-usability, and scalability.

WIEN

WIEN (Wrapper Induction ENvironment) [56, 54] is designed for automatic learning of Web pages and is strongly influenced by ShopBot (see Section 3.2.2 for details). It works on structured text containing tabular information.

WIEN looks for uniform delimiters that identify the beginning and end of each slot and for delimiters separating the tabular information from the surrounding text.

In order to automate wrapper construction as much as possible and thereby avoiding manually labeling training data, a set of techniques for automatic labeling has been developed. The labeling algorithm takes a set of heuristics for recognizing instances of attributes to be extracted as input, but it is not concerned how they are obtained.

Wrapper induction uses a bottom-up induction algorithm, which takes a set of labeled pages as input. WIEN uses only delimiters immediately preceding and following the data to be extracted. It cannot wrap sources in which some items are missing or sources where items may appear in varying order.

SoftMealy

SoftMealy [36] is based on non-deterministic finite state automata (NDFA). It uses a bottom-up inductive learning algorithm to produce extraction rules using labeled training pages. It can deal with missing values and is able to handle permutations of values if the training set includes all possible permutations.

STALKER

STALKER [70] is a hierarchical wrapper induction algorithm. It transforms complex documents into a series of simpler extraction tasks handling both missing data and permutations of values.

Web documents are described by *Embedded Catalog (EC)* trees. With the EC and a set of extraction rules STALKER is able to extract the data. To induce the extraction rules a user marks up relevant data on sample pages and STALKER applies a sequential covering algorithm for rule induction. Thereby, STALKER starts generating a rule that covers as many positive examples as possible. Then it tries to create another rule for the remaining examples, and so on. This procedure is accomplished until all examples are covered. After having a list of rules it refines the rules and creates a *disjunctive rule* to cover all examples.

STALKER handles only single-slot extraction, but requires fewer sets of training examples than other algorithms.

Lixto

Lixto [14] (see Figure 3.2) is a declarative wrapper program for supervised wrapper generation and automated web information extraction. The system assists the user to semi-automatically create wrappers by providing a visual and interactive user interface. It provides a flexible hierarchical extraction pattern definition and deals with various kinds of conditions, such as contextual, internal, and range conditions, predefined concepts, references to other user-defined patterns, or comparison conditions.

The generated wrappers are relatively stable and easy to update. The definition of extraction patterns works on both the tree structure of the document and flat strings (i.e., regular expressions). The rules are presented in an internal rule language named ELOG, whereas the wrapper designer does not deal with ELOG. The wrapper designer can label instances directly on the web page and no working inside the HTML source or the tree representation is necessary.

XWrap

XWrap [58] is a semi-automatic wrapper generation framework. It consists of four components (Figure 3.3) for data wrapping: Syntactical Structure Normalization, Information Extraction, Code Generation, and Testing and Packing.

The Information Extraction component derives extraction rules using declarative specification to describe how to extract information content of interest from its HTML formatting. It performs this task in three steps: (1) identifying interesting

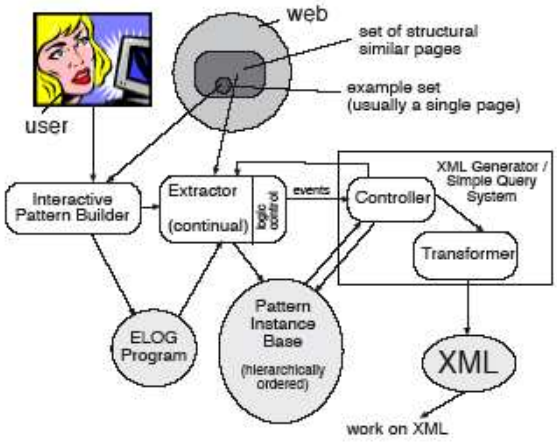


Figure 3.2: Overview of the Lixto system [14].

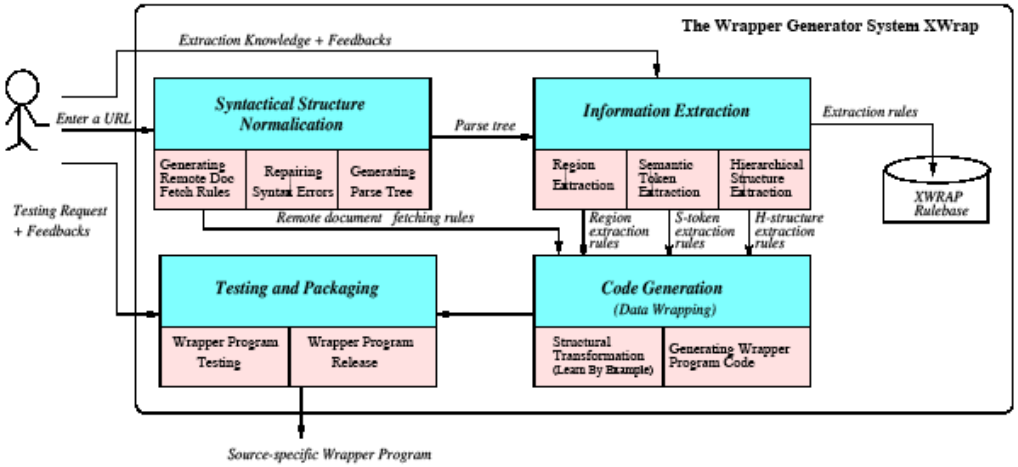


Figure 3.3: The XWrap system architecture [58].

regions in retrieved documents, (2) identifying important semantic tokens and their logical paths and node positions in the parse tree, and (3) identifying useful hierarchical structures of the retrieved document. Each step results in a set of extraction rules specified in declarative languages which are used by the Code Generation component to generate the wrapper program code.

3.2.2 Semi- and Unsupervised Learning

Semi-supervised Learning systems are based on bootstrapping methods, expanding an initial small set of extraction patterns. Bootstrapping is a general framework for improving a learner using unlabeled data [39]. Typically, bootstrapping is an iterative process where labels for the unlabeled data are estimated at each round in the process, and the labels are then incorporated as training data into the learner.

Unsupervised Learning systems reduce the burden of the user to require only

a statement of the required information. No extraction patterns are given in advance by the user. The main challenge is to realize the user's need into a set of the extraction patterns. No annotated set of documents is necessary.

Mutual Bootstrapping

Riloff and Jones [78] propose a co-training algorithm [15] using mutual bootstrapping for lexical discovery. Lexicons and extraction patterns are thereby used as separate features. Mutual bootstrapping is used due to the assumption that a good pattern can find a good lexicon and a good lexicon can find a good pattern.

Given a handful of lexical entries as initial data, patterns are discovered that extract the initial lexicon. The extracted patterns are ranked and the most reliable are used to extract more lexical items.

A strong limitation of mutual bootstrapping is that a minor error can cause a large amount of errors during the following iteration. A touching up process was introduced by *meta-bootstrapping*. Thereby, each iteration takes only the five best noun phrases for adding to the extracted lexicons.

EXDISCO

EXDISCO [106] also applies a mutual bootstrapping strategy. It is based on the assumption that the presence of relevant documents indicates good patterns and good patterns can find relevant documents.

Given an unannotated corpus and a handful of seed patterns, the document set is divided into a *relevant document set* containing at least one instance of patterns and a *non-relevant document set* not containing any seed patterns. Now candidate patterns are generated from the clauses in the documents and ranked in correlation with the relevant documents. The highest pattern is added to the pattern set and each document is re-ranked using the newly obtained pattern set. Again, the entire document set is split into *relevant* and *non-relevant* and the system keeps iterating.

Snowball

The Snowball [4] system is based on Dual Iterative Pattern Expansion (DIPRE) algorithm [17]. DIPRE is similar to co-training and works well on data with two distinct features, each of which can independently distinguish the class of instances from the other.

Figure 3.4 shows the main components of the Snowball system. Given a handful of initial relation instances and a general regular expression that the entities must match, Snowball generates patterns from text documents. A key improvement from DIPRE is that Snowball's patterns include named-entity tags (e.g., <LOCATION>-based <ORGANIZATION> instead of <STRING1>-based <STRING2>). Snowball patterns are generated by clustering similar tuples using a simple single-pass clustering algorithm.

After generating patterns, the system discovers new tuples that match the patterns to a certain degree. Each candidate tuple will then have a number of patterns

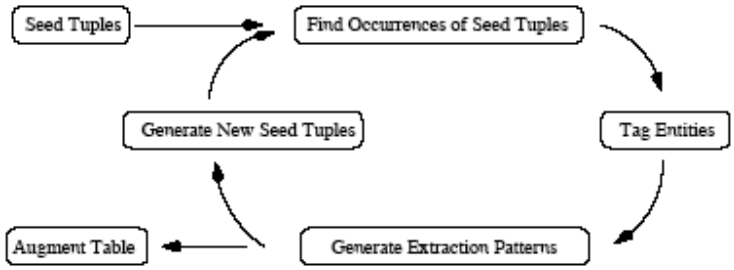


Figure 3.4: Main components of Snowball [4].

that helped generate it associated with a degree of match. This information helps Snowball to decide what candidate tuples to add to the final template.

AutoSlog-TS

AutoSlog-TS [76] is an extension of AutoSlog (see Section 3.2.1 for details). AutoSlog-TS (cp. Figure 3.5) needs only a corpus pre-classified with respect to each document's relevance to the task of interest. It generates extraction patterns for every noun phrase in the training corpus by means of heuristics. A major extension to AutoSlog is that it allows multiple rules to fire if more than one matches the context and thus multiple extraction patterns may be generated. Statistics will reveal which pattern is needed to be reliable for the domain. In a second stage these

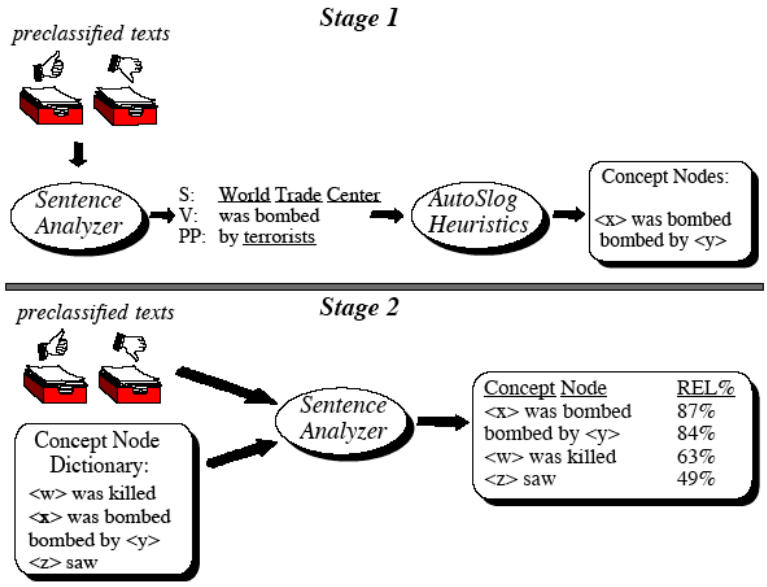


Figure 3.5: AutoSlog-TS flowchart [76].

patterns are evaluated. Thereby, it processes the corpus a second time and generates relevance statistics for each pattern by which the extraction patterns are ranked.

QDIE

The Query-Driven Information Extraction (QDIE) framework [89] tries to minimize human intervention by using a set of keywords as input. It parses the documents by a dependency parser and a Named Entity tagger and retrieves relevant documents specified by the user's query. Dependency trees of the sentences are used for pattern extraction. Each dependency subtree of a sentence that conforms to the pattern model becomes a *pattern candidate*. QDIE calculates the relevance score for each pattern candidate using tf/idf scoring in Information Retrieval (IR) literature. A pattern is more relevant the more it appears in the relevant document set and less across the entire document set.

QDIE uses the Subtree model, a generalization of the Predicate-Argument model [106] and the Chain model [90], so that any subtree of a dependency tree in a source sentence can be regarded as an extraction pattern candidate.

ShopBot

ShopBot [26] is a domain-independent comparison-shopping agent. ShopBot autonomously learns how to shop at vendors given the home pages of several on-line stores. It relies on a combination of heuristic search, pattern matching, and inductive learning techniques and does not need to apply sophisticated natural language processing.

ShopBot operates in two phases: the learning phase, which is performed offline, and the online comparison-shopping phase. During the learning phase the learner module (see Figure 3.6) automatically generates symbolic vendor descriptions of each site. Together with the domain description this is all the knowledge required by the comparison-shopping phase for finding products at this vendor.

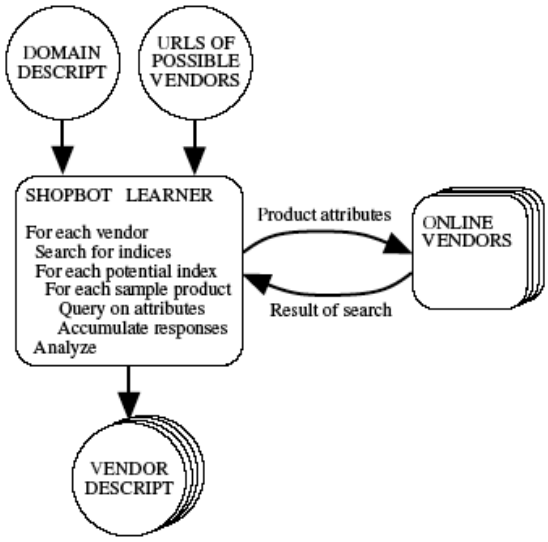


Figure 3.6: ShopBot learner module [26].

To learn the vendor description three components, which strongly interdepend,

have to be considered: (1) identifying an appropriate search form, (2) determining how to fill in the form, and (3) distinguish the format of product descriptions from the resulting page.

First, the module searches for a set of candidate forms and computes an estimate for each form of how successful the comparison-shopping phase would be if the particular form were chosen by the learner. The estimation is done by filling in the form and making several test queries using the form to search for several popular products. The test queries' results provide training examples from which the learner induces the format of product descriptions in the result pages from the used form as well as to compute the estimate measure for the form. After obtaining estimates for all forms, the learner picks up the form with the best estimate and records the vendor description, how to fill the form, and the result pages of the form.

RoadRunner

RoadRunner [24, 23] is based on an unsupervised learning algorithm. Its goal is to automatically extract data from Web sources by exploiting similarities in page structure across multiple pages. RoadRunner works by inducing the grammar of Web pages by comparing several pages containing long lists of data. Its grammar is expressed at the HTML tag level. RoadRunner works well on data-intensive sites.

Given a set of sample HTML pages belonging to the same class, the nested type of the source dataset is found. The system compares the HTML codes of the pages and infers a common structure and a wrapper. These are used to extract the source dataset.

IEPAD

IEPAD (Information Extraction based on PAttern Discovery) [20] processes semi-structured texts by means of unsupervised inductive learning. It is more expressive than WIEN (cp. Section 3.2.1) and discovers extraction patterns from web pages without user-labeled examples. IEPAD applies several pattern discovery techniques such as PAT trees, multiple string alignments, and pattern matching algorithms. Its extraction rules are pattern-based instead of delimiter-based and it can handle exceptions such as missing attributes, multiple attribute values, and variant attribute permutations.

Chapter 4

Architecture of an Information Extraction Framework

A typical Information Extraction system has phases for input tokenization, lexical and morphological processing, some basic syntactic analysis, and identifying the information being sought in the particular application (cf. Figure 4.1) [8]. Depending on what one is interested in some phases may not be necessary. In addition to the modules in the left-hand column, IE systems may include modules from the right-hand column, depending on the particular requirements of the application.

In this chapter we describe the most relevant modules in detail.

4.1 Tokenization and Text Sectioning

The Tokenization module is responsible for splitting the input text into sentences and tokens. Tokenization is a trivial problem for European languages with clear word borders. However, in processing some languages like Chinese or Japanese, it is not evident from the orthography where the word boundaries are. Therefore, systems for these languages must necessarily be expanded by a **Word Segmentation module**.

4.2 Lexical and Morphological Processing

Morphology is a sub discipline of linguistics and is interested in the structure of word forms. Many IE systems for languages with simple inflectional morphology, like English, do not have a morphological analysis component at all. In English, it is easy to simply list all inflectional variants of a word explicitly in the lexicon. For languages like French, with more complex inflectional morphology, a morphological analysis component makes more sense, but for a language like German, where compound nouns are agglutinated into a single word, morphological analysis is essential.

In the lexical analysis the tokens determined by the Tokenization module are looked up in the dictionary to determine their possible parts-of-speech and other lexical features that are required for subsequent processing. The most important job

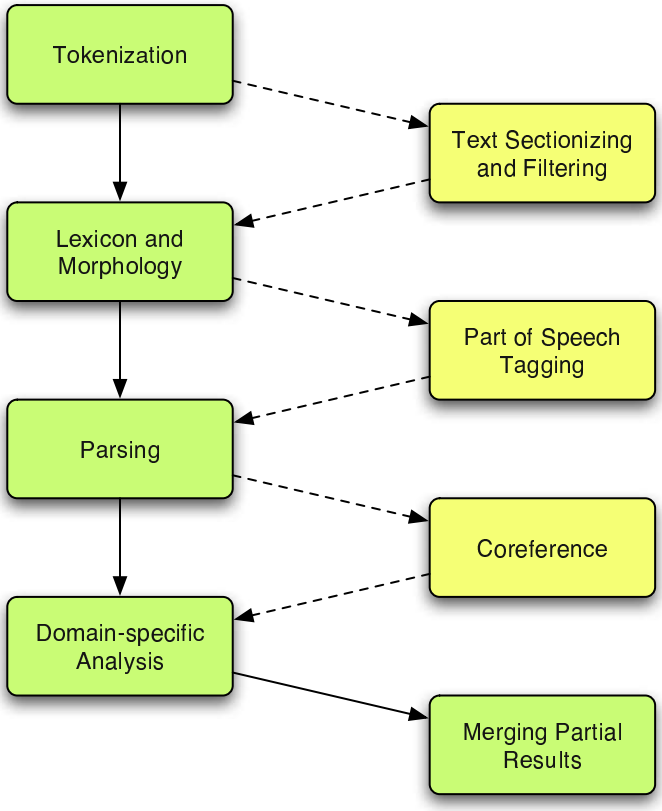


Figure 4.1: Modules of an Information Extraction System. The green items symbolize modules which are frequently used in IE systems. The yellow items present rarely used modules.

of this module is the handling of proper names. Recognizing names can thereby be done with either handcrafted rules under the Knowledge Engineering approach or automatically trained rules derived from an annotated corpus.

In addition to name recognition, this module must assign lexical features to words required for subsequent processing. This can be accomplished by either a lookup in a lexicon, or by automatic taggers, like a parts-of-speech tagger. A **parts of speech tagger** annotates each word of a sentence with its parts of speech tag, such as noun, verb, adjective, and so on. It can avoid incorrect analysis based on disambiguation caused by rare word senses in comprehensive lexicons.

4.3 Parsing

Syntactic Analysis has the aim to identify the syntactic structure of the analyzed document. Most IE systems only accomplish a shallow, fragmented analysis of the texts. But there have even been IE systems which totally skip the phase of syntactic analysis.

IE systems are only interested in specific types of information in a text and ignore those portions of text which are not relevant for their task. Therefore, parsing these portions and finding irrelevant grammatical relationships will be unnecessary.

4.4 Coreference

The reason for this module is simply that application relevant entities will be referred to in many different ways throughout a given text and thus success on the IE task was, at least to some extent, conditional on success at determining when one noun phrase referred to the very same entity as another noun phrase. Thus, a Coreference module should handle the following coreference problems:

Name-alias coreference. Names and their common variants must be recognized as coreferring, e.g., 'Silvia Miksch' and 'Prof. Miksch'.

Pronoun-antecedent coreference. Pronouns like 'she', 'he', 'they', and so on must be associated with their antecedents, resolving them to a domain relevant named entity if possible.

Definite description coreference. Handling this problem requires arbitrary world knowledge to make the connection among descriptors. When building an IE system, it is reasonable to include ontological information for domain-relevant entities that enable such resolution in restricted cases, but doing it in general is unrealistic. Examples are 'Apple Computer' and 'the company' or 'Apple' and 'the Cupertino computer manufacturer'.

As with other IE modules, it is possible to build Coreference modules by both knowledge engineering and automatic learning approach. The knowledge engineering approach is, for instance, applied by FASTUS [9]. For accomplishing the coreference task with the learning approach Decision Trees can be applied [59].

4.5 Domain-specific Analysis

The domain analysis is the core of most IE systems. The preceding modules prepare the text for the domain analysis by adding semantic and syntactic features to it.

This module fills the templates, which are in general constructed as attribute-value pairs. Templates consist of a collection of slots (i.e., attributes), each of which may be filled by one or more values. These values can consist of the original text, one or more of a finite set of predefined alternatives, or pointers to other template objects. Typically, slot fillers representing dates, times, job titles, and so on are standardized by normalization rules.

For extracting facts and events, the system needs domain specific extraction patterns (i.e., extraction rules or case frames). These patterns can be generated manually (by means of Knowledge Engineering) or automatically (by means of automatic learning). The portion of text that matches a defined linguistic pattern is memorized and the information is extracted by the guidance of the extraction rule from this portion of text to fill the template.

For designing the domain-relevant pattern rules there exist two approaches, which can be characterized as (1) the *atomic approach* and (2) the *molecular approach*.

The *molecular approach* is the most common one. It involves matching all or most of the arguments in an event in a single pattern. The development cycle of this approach starts with a small number of highly reliable rules that capture the common core cases of the domain, but ignoring a broad class of less frequently occurring relevant patterns. Further development is characterized by capturing ever-larger numbers of increasingly rare and problematic cases with increasingly general and possibly overgenerating rules. Thus, the system starts with high precision and low recall scores and evolves in increasing recall and progressively lower precision.

The *atomic approach* builds a domain module that recognizes the arguments to an event and combines them into template structures strictly on the basis of intelligent guesses rather than syntactic relationships. The development cycle can be characterized by assuming domain-relevant events for any recognized entities, leading to high recall, but much overgeneration, and thus low precision. Further development would result improving filters and heuristics for combining the atomic elements, improving precision

The atomic approach make sense at tasks characterized by the following features: (1) entities in the domain have easily determined types and (2) the templates are structured so that there is only one or a very small number of possible slots that an entity of a given type can fill and only entities of a given type can fill those slots.

4.6 Merging Partial Results

Usually, the sought-after information is spread among different sentences. In these cases information should be combined before creating the final templates. For this purpose some IE systems use a Merging module. This module uses an algorithm to decide which templates can be merged.

Again, there are two merging strategies: (1) the Knowledge Engineering approach, where one performs data analyses to identify good merging principles, defines rules, and tests the results; and (2) using some training mechanism to acquire the merging strategies automatically by the system.

Chapter 5

Analysis of *Asbru* in Respect of Plan Modeling

This chapter addresses the definition of rules that can deal with various temporal information of processes, such as metric or qualitative information and point or interval relations among processes, uncertain and incomplete temporal information for representing these processes in *Asbru* (see Section 2.1.1 for details).

Clinical guidelines contain treatment instructions in natural language. Each episode in the context of a therapy is often temporally related to other episodes. This relationship is called *interval relation*, as every episode can be presented as an interval.

Natural language often contains information about interval relations in an incomplete and uncertain form. **Incompleteness** is thereby defined as missing information about the relation between two intervals (e.g., there is only information available that interval X starts before interval Y is started and there is no information about the ending of the intervals). **Uncertainty** denotes imprecise information of time points, such as beginning and end, or durations.

The emphasis of Asbru is the modeling of CGPs that are *process-oriented*, that is, temporal aspects of actions and events are explicitly and often mentioned. In order to offering support for modeling such processes we have to systematically analyze Asbru elements and process components. Based on these preparations we can define rules for an automized modeling process.

5.1 Information About Time

To represent temporal relationships between events it demands for a temporal logic. We can differentiate among instant-based time theories [60, 85, 12], which cannot represent events taking time, interval-based theories [6, 7, 31], and theories treating both instances and intervals as independent primitives [16, 101].

Defining flows in a formal representation is a very complex and difficult task. Dependencies among specific activities and processes rule which have to be considered. In doing so the complete interval of an activity or individual time points that define the limiting points of the interval (i.e., the start or the end of an activity

or process) have to be considered. Therefore, we adopted the latter theory based on both instances and intervals for this work.

In planning activities and actions are modeled and their flows are structured. These activities and actions are represented by plans that can specify both discrete actions and composite actions and in this chapter we will primarily focus on their temporal aspects. That is, a plan is represented by an interval that is expressed by temporal constraints. Thereby, different kinds of constraints have to be considered that are used in the following expressions [100]: (1) metric information, (2) qualitative relations, and (3) combinations of both.

Metric Information. Thereby, temporal information is available in dates or another precise numeric form. Durations can be numerically represented and easily computed. Assertions are time-stamped to absolute numeric values.

Qualitative Relations. When specifying time annotations we can describe time points and intervals. Exact time points can be differentiated in a simple way by relations *before* ($\prec$), *equal* ($=$), and *after* ($\succ$). Relation vectors are composed of simple point relations, for instance ($p_1 \prec = p_2$). The constraints closure can be computed in polynomial-time algorithms [102].

Based on the 13 relations that correspond to the simple definite mutually exclusive relations that may exist between two intervals, Allen developed an *algebra of intervals* [5]. The relations of these algebra are *before/after* ($</>$), *meets/met-by* (m/mi), *equal* ($=$), *starts/started-by* (s/si), *during/contains* (d/di), *finishes/finished-by* (f/fi), and *overlaps/overlaped-by* (o/oi).

But what if the exact time of each event is not necessarily known? Especially in planning and scheduling there are often uncertainties or even incompleteness of the knowledge in respect of the beginning, ending, and duration of an action or event. In such a case it is not always possible to give exact statements regarding the relation to other actions or events.

Thus, Allen takes vectors of simple relations, which are interpreted as the disjunction of relations. For instance, the vector ($I_1 \ before\, meets\, overlaps \ I_2$) means that the interval I_1 either occurs before, meets, or overlaps I_2. As the constraint consistency of statements in the interval algebra is *NP-hard* [102], a suggestion to work in practical systems is amongst other things moving to a less expressive formalism, the point algebra.

Another considerable point in the interval algebra is incomplete knowledge about the relationship between intervals. The less knowledge is available about the relationship the longer is the symbolic representation of that knowledge. Therefore, alternative sets of relationships were developed. For instance Freksa's [31] approach is based on the concept of neighborhood. He generalizes relations with *semi-intervals*. Thereby, he combines several Allen-relations to new relations altogether. This offers the possibility to compare uncertain intervals and represent them by only one relation.

Combinations of Qualitative and Metric Information. Both types can be integrated in a single system in order to cope with the availability of knowledge at

different degrees of precision. Thereby, one approach is to address qualitative and quantitative interval based components separately and circulate the results of the reasoning tasks between them to approximate to a global solution [45]. Another approach is defining *general temporal networks* integrating both types of constraints. The problem of reasoning on general temporal networks is solved by applying classical constraint satisfaction techniques [61].

We will deal with interval-based theories as both within the clinical guidelines the particular actions are formulated as such and Asbru represents actions as intervals.

5.2 Time Annotations and Their Semantics

For the formal definition of complex flows we need a powerful representation. Asbru provides a powerful means to represent intervals: the *time annotation* (see Figure 5.1).

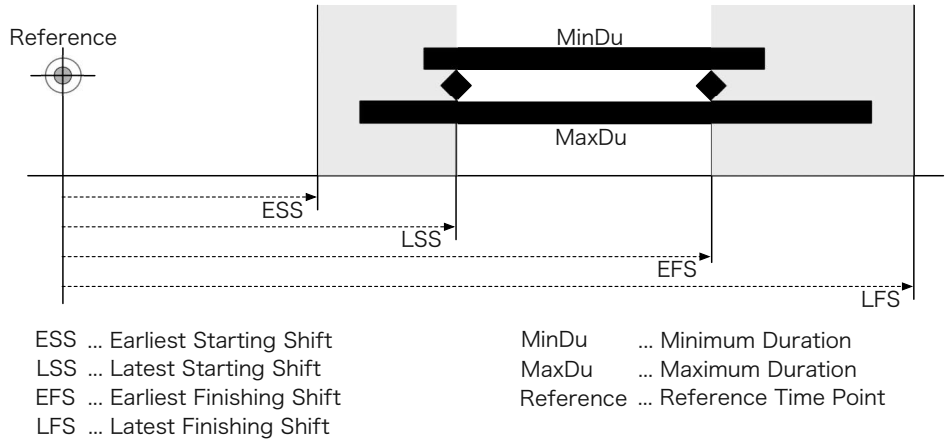


Figure 5.1: Time interval in Asbru. The grey areas indicate the periods when the action has to start and accordingly finish.

This figure shows that uncertainties exist in respect of the beginning, the end, and the duration of the interval. Furthermore, no time points for the beginning and the end are defined, but shifts from an arbitrary definable reference point. This point can individually be assigned for any interval. To cope with incomplete information each piece of information is optional. Starting from this definition the following implicit constraints can be extracted [27]:

$$ESS \leq LSS \tag{5.1}$$

$$EFS \leq LFS \tag{5.2}$$

$$ESS \leq EFS \tag{5.3}$$

$$LSS \leq LFS \tag{5.4}$$

$$MinDur \leq MaxDur \tag{5.5}$$

$$EFS - LSS \leq MinDur \leq Min[(EFS - ESS), (LFS - LSS)] \tag{5.6}$$

$$Max[(EFS - ESS), (LFS - LSS)] \leq MaxDur \leq LFS - ESS \tag{5.7}$$

5.3 Representing Plans in *Asbru*

Before plans and flows are displayed in a formal representation, their individual temporal aspects are described in different ways. We want both to enable different and consistent forms of representation and to identify coherences among several plans. This is not only helpfully for further automated processing, but also to enable the human plan generator a better comprehension of coherences.

To represent relations between intervals Allen introduced a temporal logic based on intervals and their qualitative relationships in time [5]. His approach is simple, transparent, and easy to implement. It is based on intervals corresponding to events, qualitative relations between these intervals, and an algebra for reasoning about these interval relations. He defines 13 mutually exclusive interval relations (see Figure 5.2).

Relation	Label	Inverse	Illustration
X *before* Y	X < Y		XXX
Y *after* X		Y > X	YYY
X *meets* Y	X m Y		XXXX
Y is *met-by* X		Y mi X	YYYY
X *equals* Y	X = Y		XXXXXXXX
		Y = X	YYYYYYYY
X *during* Y	X d Y		XXXX
Y *contains* X		Y di X	YYYYYYYY
X *starts* Y	X s Y		XXXX
Y is *started-by* X		Y si X	YYYYYYYY
X *finishes* Y	X f Y		XXXX
Y is *finished-by* X		Y fi X	YYYYYYYY
X *overlaps* Y	X o Y		XXXXX
Y is *overlapped-by* X		Y oi X	YYYYY

Figure 5.2: Allen's 13 mutually exclusive interval relations.

In [43] we showed how these relations can be modeled in Asbru. Therefore, Asbru provides three possibilities to model relationships between events: (1) plan states and state transitions, (2) plan types, and (3) time points.

5.3.1 Plan States and Plan State Transitions

As mentioned in Section 2.1.1 Asbru provides several plan types (sequentially, parallel, any-order, unordered, subplans), which describe the behavior concerning the

execution of the plans and their synchronization. Moreover, each instance of a plan can take six plan states [79] (cp. also Figure 2.3 on page 8):

1. Considered
2. Possible
3. Activated
4. Suspended
5. Completed
6. Aborted

The first two states (i.e., *considered* and *possible*) are combined by the superior state *selected*. The states *activated* and *suspended* can be addressed by the superior state *executed*.

An instantiated plan can only have one plan state at any one time. Based on these specifications we will now describe the modeling possibilities. To simplify matters only seven of Allen's thirteen relations are stated. The missing six relations can easily be created by exchanging the attributes.

***Before* relation.** The subsequent interval B can not start until the preceding interval A is in *completed* or *aborted* state.

***Meets* relation.** The subsequent interval B must start when the preceding interval A switches from *executed* to *completed* or *aborted* state.

***Equal* relation.** Interval B starts when interval A switches to *executed* state and finishes when interval A switches to *completed* or *aborted* state.

***Contains* relation.** Interval B can not start until interval A is in *executed* state and it must finish while interval A is in *executed* state.

***Starts* relation.** When interval A switches to the *executed* state interval B can start. Interval B can not finish until interval A is in *completed* or *aborted* state.

***Finishes* relation.** Interval B must start before interval A is in *executed* state. Both intervals finish at the same time, that is, when interval A switches to the *completed* or *aborted* state interval B also switches to the *completed* or *aborted* state.

***Overlaps* relation.** Interval B starts when interval A is in *executed* state and finishes when interval A is in *completed* or *aborted* state.

5.3.2 Plan Types

Based on the various plan types in *Asbru* (cp. Section 2.1.1) we can model interval relations in the following way:

Before/after relation. In these relations the preceding plan must be in *completed* or *aborted* state when the subsequent plan switches into *executed* state, which happens in **sequential plans** and **any-order plans**.
The relation can also be performed by **unordered plans**, if an additional restriction like 'preceding plan in state *completed* or *aborted*' is defined.

Meets/met-by relation. Basically, the same plan types can be used like for 'before/after'-relations. The difference between the 'meets/met-by'-relations and the 'before/after'-relations is that the intervals of the first relations adjoin and between the intervals of the latter relations are again a period of time. Therefore, it is necessary to set the starting time of the subsequent interval to the finishing time of the preceding interval to define 'meets/met-by'-relations.

Equal relation. In this relation the limiting points of both intervals must correspond. This concurrence is found in **parallel plans** where the starting times are synchronized. Additional constraints have to be set regarding the finishing time.
In the same way this kind of relation can be modeled by **subplans**. In doing so the starting and finishing times of the subplan have to correspond to the particular times of the parent plan.
The same constraints have to be applied for a realization with **unordered plans**.

Contains/during relation. Similar to the 'equal relation' the 'contains'-relation can be modeled by **subplans** and **unordered plans**. The difference of this relation is that the starting time of the subplan is greater than the parent plan's starting time and the finishing time of the subplan is less than the parent plan's finishing time. In other words the subplan can only be activated if the parent plan is in *executed* state and the parent plan can not finish until the subplan is in *completed* or *aborted* state.
A 'during'-relation has to be converted into a 'contains'-relation in order to model it by **subplans**.

Starts/started-by relation. For this kind of relations the same plan types are applied like for the 'equal'-relation, but with different constraints regarding the finishing time. The second plan can only finish if first plan is already in *completed* or *aborted* state.

Finishes/finished-by relation. The 'finished-by'-relation can be applied by means of a **subplan**. Thereby, the plan with the shorter duration is the subplan. An additional constraint must be defined to finish both plans at the same time.
Both relations can be constituted by **unordered plans**.

Overlaps/overlaped-by relation. The 'overlaps'-relation can be implemented by a **subplan** type, whereby the subplan can be finished when the parent plan is already in *completed* or *aborted* state.
Both relations can be modelled by the **unordered plan** type with additional constraints for the beginning and finishing points.

Considering a simplified modeling of interval relations in Asbru, we can part Allen's thirteen mutual exclusive interval relations [5] in two categories:

1. Interval relations, where the intervals appear in sequential order; those are *before, after, meets*, and *met-by*.

2. Interval relations, where the intervals are overlapping. That means, at least one subinterval of an interval is equal to a subinterval of the related interval. These interval relations are *equal, starts, started-by, finishes, finished-by, during, contains, overlaps*, and *overlapped-by*.

From the latter category we can again deduce different classes of relations regarding the position of the various starting and finishing points. Figure 5.3 gives an outline of these classes.

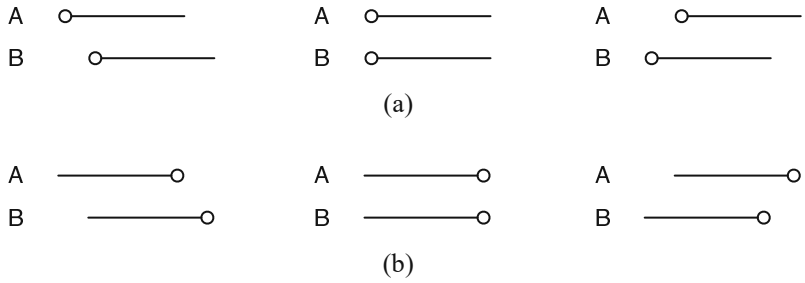


Figure 5.3: Different formations regarding (a) the starting points and (b) the finishing points of two intervals; so called semi-intervals.

Now we want to gather special insights by considering only the limiting points of the intervals for a simplified model based on these plan types.

Starting Points. Considering the *equal, starts/started-by, finishes/finished-by*, *during/contains,* and *overlaps/overlapped-by* relations we can differentiate among the three configurations regarding the starting points (cp. Figure 5.3(a)). Possible Asbru plan types that we can use to model interval relations categorized by the relation of their starting points are cited in Figure 5.4.

Finishing Points. Equivalent to the above mentioned considerations we can define the plan types for interval relations categorized by the relation of their finishing points (cp. Figure 5.3(b)). The particular plan types are represented in Figure 5.5.

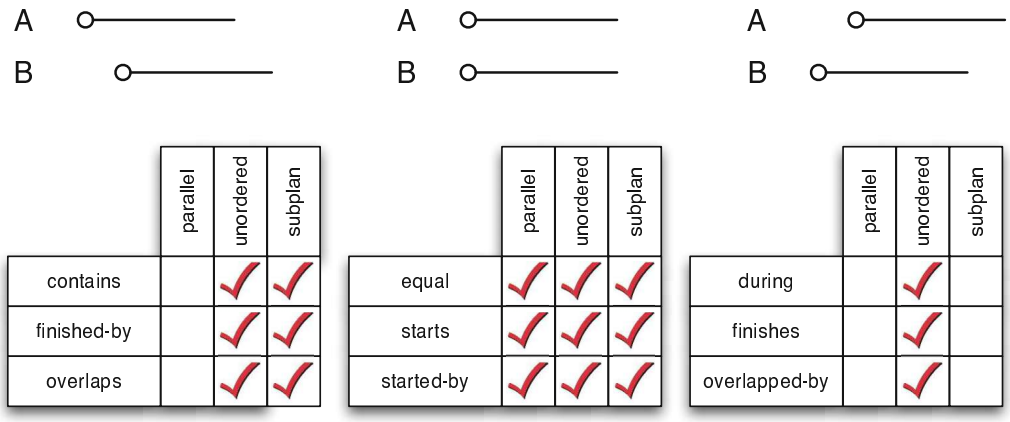


Figure 5.4: Plan types used to model several semi-interval relations. This array is resulting from considerations regarding the starting points.

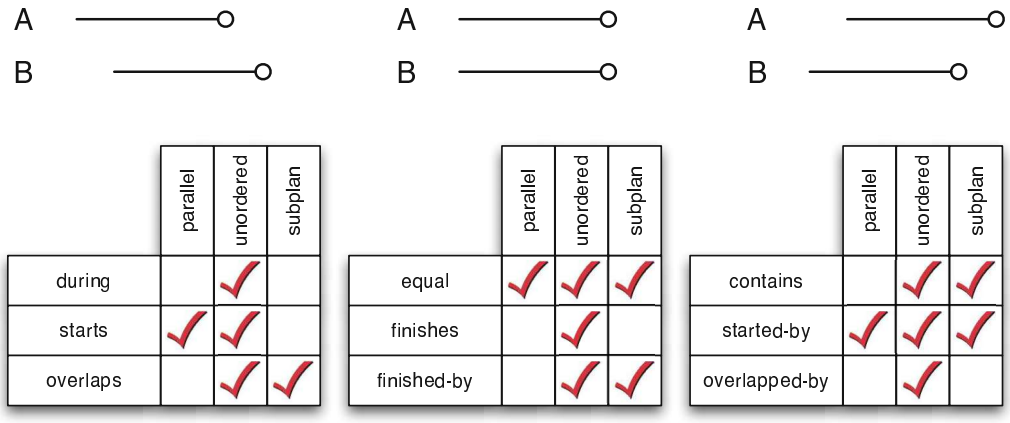


Figure 5.5: Plan types used to model several semi-interval relations. This array is resulting from considerations regarding the finishing points.

The types of synchronization that define the different plan types in Asbru are oriented by the start of the particular plans – and not by the finishing of the plans (cp. [79]). This is apparent in Figures 5.4 and 5.5.

Based on these specifications we can conclude that the set of possible Asbru plan types only depend on the relation of the starting points of the intervals.

5.3.3 Time Points

Apart from the modeling of relations based on plan states and plan types, they can be modeled by referencing timepoints between the intervals.

First, we will declare implicit constraints, which must hold during the execution of the plans. We will only state the first seven relations as the inverse relations can be defined by exchanging the attributes. These constraints are based on the simplified form of intervals by time points (SP ... starting point; FP ... finishing point) and the duration.

1. **A before B**
   $\mathtt{SP}_B \succ \mathtt{FP}_A$
2. **A meets B**
   $\mathtt{SP}_B = \mathtt{FP}_A$
3. **A equal B**
   $\mathtt{SP}_B = \mathtt{SP}_A$
   $\mathtt{FP}_B = \mathtt{FP}_A$
   $\mathtt{duration}_B = \mathtt{duration}_A$
4. **A during B**
   $\mathtt{SP}_B \prec \mathtt{SP}_A$
   $\mathtt{FP}_B \succ \mathtt{FP}_A$
   $\mathtt{duration}_B > \mathtt{duration}_A$
5. **A starts B**
   $\mathtt{SP}_B = \mathtt{SP}_A$
   $\mathtt{FP}_B \succ \mathtt{FP}_A$
   $\mathtt{duration}_B > \mathtt{duration}_A$
6. **A finishes B**
   $\mathtt{SP}_B \prec \mathtt{SP}_A$
   $\mathtt{FP}_B = \mathtt{FP}_A$
   $\mathtt{duration}_B > \mathtt{duration}_A$
7. **A overlaps B**
   $\mathtt{SP}_B \succ \mathtt{SP}_A$
   $\mathtt{FP}_B \succ \mathtt{FP}_A$

As the intervals might have uncertainties regarding the start and end time points as well as the duration, it is very difficult to define relations by time points.

Considering the previous constraints, we can make the following declarations about intervals for the more complex form that is used in *Asbru* [79]:

A before B. The reference point of B (ref_B) is set to the time point, when plan state of interval A passes from *executed* to *completed* or *aborted*.

$$ref_B = FP_A \tag{5.8}$$

B's earliest starting shift has to be greater 0.

$$ESS_B > 0 \tag{5.9}$$

A meets B. This relation is similar to the 'before/after relation', but the starting shift has to be set to 0:

$$ref_B = FP_A \tag{5.10}$$

$$ESS_B = 0, LSS_B = 0 \tag{5.11}$$

A equal B. The reference point of interval B is set to the time point, when interval A passes to *executed* state.

$$ref_B = SP_A \tag{5.12}$$

The starting shift is set to 0.

$$ESS_B = 0, LSS_B = 0 \tag{5.13}$$

The time points for the finishing shift cannot be set exactly.

$$EFS_B \geq MinDur_A, LFS_B \leq MaxDur_A \tag{5.14}$$

A during B. For this relation the starting shift can not be defined and for the finishing shift no exact definitions can be made, too. On this account the inverse relation **A contains B** will be defined and can be used for modeling the *during* relation:
B's reference point is set to the time point when plan A passes to the *executed* state.

$$ref_B = SP_A \tag{5.15}$$

B's starting shift must be after A's starting shift.

$$ESS_B > 0 \tag{5.16}$$

B's finishing shift must be less than A's finishing shift

$$LFS_B < LFS_A \tag{5.17}$$

and the duration of B must be less than the duration of A.

$$MaxDur_B < MaxDur_A \tag{5.18}$$

A starts B. B's reference point is set to the time point when plan A passes to the *executed* state.

$$ref_B = SP_A \tag{5.19}$$

The starting shift is set to 0.

$$ESS_B = 0, LSS_B = 0 \tag{5.20}$$

The finishing shift cannot be exactly defined.

$$EFS_B > MinDur_A \tag{5.21}$$

$$MinDur_B > MinDur_A \tag{5.22}$$

A finishes B. Due to the possiblitiy of different reference points of the intervals the finishing shift cannot be set exactly.

$$MinDur_B > MinDur_A \tag{5.23}$$

A overlaps B. B's reference point is set to the time point when A is *activated*.

$$ref_B = SP_A \tag{5.24}$$

Both the starting and the finishing shift cannot be set exactly.

$$0 < ESS_B < MinDur_A \tag{5.25}$$

$$EFS_B > MinDur_A \tag{5.26}$$

5.3.4 Uncertain and Incomplete Intervals

If the temporal information about intervals is incomplete there are several possible relations. Which relation vectors are resulting can be derived from the concept of neighborhood [31]. **Neighborhood** is defined, if one time point of one interval is changing to the next time point that defines another relation. In Figure 5.6 all vectors are shown that result from non-existing knowledge about one time point. Comparing the vectors with Figure 5.7 we can see that all these relations of one vector are neighboring.

As well the possible Asbru plan types are stated. They are resulting from the intersection of possible plan types which can be used to model each individual relation.

If only incomplete knowledge about this one time point is available the resulting vector is changing and thus, also the possible plan types in Asbru change. Figure 5.8 shows an example where the first four vectors describe relations with given Latest Starting Shift (LSS) of interval Y that is moving from left to right. The fifth vector in the middle describes relations where for example, no definitions about the Starting Shift (ESS and LSS) of interval Y are given. The latter four vectors describe relations with given Earliest Starting Shift (ESS) of interval Y that is moving from left to right.

In case there are more uncertain or incomplete time annotations the vectors are growing more and more complex. Thereby, the only possibility to model the relations between intervals can be to connect each interval by means of its metric information and represent them using 'unordered' plan types.

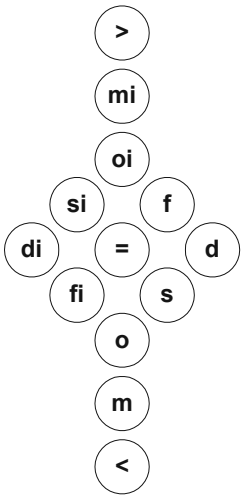


Figure 5.6: Neighboring qualitative relations.

Relation Vectors		Asbru Plan-Types				
		sequential plans	parallel plans	any-order plans	unordered plans	subplans
X Y	X **<** **m** **o** **s** **d** Y				✓	
X Y	X **fi** **=** **f** Y				✓	
X Y	X **di** **si** **oi** Y				✓	
X Y	X **mi** Y	✓		✓	✓	
X Y	X **>** Y	✓		✓	✓	
X Y	X **>** **mi** **oi** **f** **d** Y				✓	
X Y	X **si** **=** **s** Y		✓		✓	✓
X Y	X **di** **fi** **o** Y				✓	✓
X Y	X **m** Y	✓		✓	✓	
X Y	X **<** Y	✓		✓	✓	

Figure 5.7: Interval relation vectors and possible Asbru plan types. Fixed time points are displayed by black circles and time points illustrating possible states in an uncertain area are denoted by plotted circles.

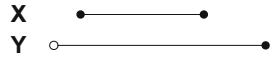
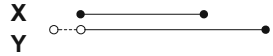
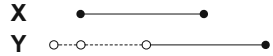
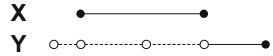
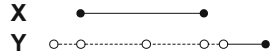
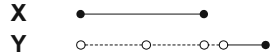
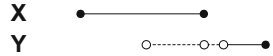
Relation Vectors		Asbru Plans				
		sequential plans	parallel plans	any-order plans	unordered plans	subplans
	X **d** Y				✓	✓
	X **s d** Y				✓	✓
	X **o s d** Y				✓	
	X **m o s d** Y				✓	
	X **< m o s d** Y				✓	
	X **< m o s** Y				✓	
	X **< m o** Y				✓	
	X **< m** Y	✓		✓	✓	
	X **<** Y	✓		✓	✓	

Figure 5.8: Changing vectors with uncertain knowledge about intervals and particular possible Asbru plans. Fixed time points are displayed by black circles and possible time points of the uncertain starting area are denoted by plotted circles.

5.3.5 Semi-Intervals

To better cope with the information incompleteness relations between intervals can also be modeled by Freksa's semi-intervals [31]. Semi-intervals correspond to temporal beginnings or endings of events. The advantages of semi-intervals are that coarse knowledge can be processed directly, incomplete knowledge can be fully exploited, and knowledge about relations between events can be represented much more compactly. Freksa defines eleven semi-intervals depicted in Figure 5.9.

Based on these specifications we will now describe the modeling possibilities:

X is *older* than Y; Y is *younger* than X

(1) Y starts if X is in state *executed, completed*, or *aborted*

(2) Y is **subplan** of X (only if X overlaps Y)

(3) Y's reference point for its starting shifts are set to X's starting point (SP) and Y's earliest starting shift (ESS) is set to '1'

(4) X and Y are sibling plans of type **unordered** and additional constraints of (1) or (3) are necessary

X is *head to head* with Y

(1) Y starts if X switches from state *selected* to state *activated*

Relation	Label	Inverse	Illustration
X is *older* than Y	ol		XXX????
Y is *younger* than X		yo	YY
X is *head to head* with Y	hh		XXX??
		hh	YYYY
X *survives* Y	sv		????XXX
Y is *survived by* X		sb	YY
X is *tail to tail* with Y	tt		??XXX
		tt	YYYY
X *precedes* Y	pr		XXX?
Y *succeeds* X		sd	YYY
X is a *contemporary* of Y	ct		?XXX???
		ct	???YYY?
X is *born before death of* Y	bd		XXX?????
Y *died after birth of* X		db	?????YYY

Figure 5.9: Eleven semi-interval relationships. Question marks (?) in the pictorial illustration stand for either the symbol denoting the event depicted in the same line (X or Y) or for a blank. The number of question marks reflects the number of qualitatively alternative implementations of the given relation. [31]

(2) X and Y are sibling plans of type **parallel** or **unordered**

(3) Y's reference point for its starting shifts is set to X's starting point (SP) and Y's earliest and latest starting shifts (ESS, LSS) are set to '0'

X *survives* Y; Y is *survived by* X

(1) X finishes when Y is already in state *completed* or *aborted*

(2) X's reference point for its finishing shifts is set to Y's finishing point (FP) and X's earliest finishing shift (EFS) is set to '1'

(3) As the plan types only influence the synchronization of the start of the plans, we cannot declare a plan type for this semi-interval relationship.

X is *tail to tail* with Y

(1a) X finishes when Y switches to state *completed* or *aborted*

(1b) Y finishes when X switches to state *completed* or *aborted*

(2a) X's reference point for its finishing shift is set to Y's finishing point (FP) and X's earliest and latest finishing shifts (EFS, LFS) are set to '0'

(2b) Y's reference point for its starting shifts is set to X's finishing point (SP) and Y's earliest and latest finishing shifts (EFS, LFS) are set to '0'

(3) X and Y are sibling plans from type **unordered** and additional constraints of (1a/1b) or (2a/2b) are necessary

X *precedes* Y; Y *succeeds* X

(1) Y starts if X is in state *completed* or *aborted*

(2) X and Y are sibling plans of type **sequential** with X stated before Y

(3) Y's reference point for its starting shifts is set to X's finishing point (SP) and Y's earliest starting shift (ESS) is set to '0'

X is *contemporary* of Y This relation is hard to obtain. We cannot definitely declare how it can be modeled.

(1) X and Y are sibling plans of type **unordered**.
It is also possible for plan type **parallel**, if the plans are *head to head*.
X and Y are parent plan and **subplan**

X is *born before death* of Y; Y *died after birth* of X

(1a) X starts when Y is in state *selected*, or *executed*

(1b) Y finishes when X is in state *executed, completed*, or *aborted*

(2) As the plan types only influence the synchronization of the start of the plans, we cannot declare a plan type for this semi-interval relationship.

(3) Based on Asbru's time annotation the following implicit constraints concerning the finishing shift of an interval are defined: (1) $EFS \leq LFS$; (2) $ESS \leq EFS$; and (3) $LSS \leq LFS$.
Modeling a 'died after birth of' relation requires additional constraints: the reference point of Y's finishing shift is set to the starting point of X and Y's earliest finishing shift (ESS) is set to '1'.

Freksa's semi-intervals better map incomplete and uncertain information about the relations than Allen's interval relations. Therefore, we will rather apply this concept for modeling relations.

5.4 Other Relation Concepts: *Selection* and *Iteration*

Treatment instructions often contain special forms of dependencies: the *selection* and the *iteration*.

5.4.1 Select Relations

Select-one-of actions often appear in guidelines, but unfortunately Asbru does not provide a separate modeling concept. Thus, the modeling is elaborate.

Select-one-of actions must have the same parent plan. This is similar to Moser's XOR structure [66]. She defines it by utilizing `any-order plans` so that only one plan can be executed at any one time (but not excluding one from execution). The decision of which plan starts executing is done in the pre-selection phase. In order to have only one action executed we have to implement additional constraints. Thus, we define the complete-condition for the parent plan when finishing the first child plan.

5.4.2 Iterative Events

Basically, we can differentiate between two kinds of iterations:

1. Iterations defined by the frequency (i.e., the number of recurrences within a certain period)
   Example: *twice a day*

2. Iterations defined by the period between two recurrences
   Example: *every 4-6 hours*

For iterative events Asbru provides the concept of 'cyclical plan'. With this plan we are able to model both kinds of iterations. A plan containing recurrent instructions defines a *cyclical plan* containing the following definitions:

cyclical-plan-body is a single plan step. In most cases it will be a `plan-activation`.

any-repeat-specification is either a `repeat-specification` or a `cyclical-time-annotation` specifying the times at which the step defined by `cyclical-plan-body` is repeated.

The **cyclical-time-annotation** is thereby used for iterations with given frequency. Although we have to mention that the frequency has to be stated as the period between two repetitions (e.g., the frequency for 'twice a day' is '12 hours').

The `cyclical-time-annotation` guarantees that even after one week the iterations are synchronized compared to the `repeat-specification`. See Listing 5.1 for an example of a cyclical plan with a cyclical-time-annotation.

The **repeat-specification** is used for iterations with an individual interval between two occurrences. It makes a difference whether you take a drug every four hours or every six hours, as you can take more drugs if you take it every four hours. This model can be used assuming that the duration of the iteration is not important. Only the duration from the end of one repetition to the beginning of the succeeding repetition is measured. Asbru does not provide a model measuring the duration from the beginning of one repetition to the beginning of the succeeding repetition. See Listing 5.2 for an example of a cyclical plan with a repeat-specification.

Listing 5.1: Example of an iterative plan with given frequency of repetitions. The cyclical plan contains the nodes `cyclical-plan-body` and `cyclical-time-annotation`. Within the latter one we can define the frequency of the iterations by means of the period between two iterations.

```
<plan name="PLAN_WITH_ITERATIVE_INSTRUCTIONS_twice_a_day" >
   <plan-body>
      <cyclical-plan>
         <cyclical-plan-body>
            <plan-activation>
               <plan-schema name="CYCLICAL_PLAN" />
            </plan-activation>
         </cyclical-plan-body>
         <cyclical-time-annotation>
            <time-range>
               <starting-shift>
                  <minimum>
                     <numerical-constant value="4" unit="h" />
                  </minimum>
                  <maximum>
                     <numerical-constant value="10" unit="h" />
                  </maximum>
```

```
18                </starting-shift>
19             </time-range>
20             <set-of-cyclical-time-points>
21                <time-point>
22                   <numerical-constant value="0" />
23                </time-point>
24                <offset>
25                   <numerical-constant value="0" />
26                </offset>
27                <frequency>
28                   <numerical-constant value="12" unit="h" />
29                </frequency>
30             </set-of-cyclical-time-points>
31          </cyclical-time-annotation>
32       </cyclical-plan>
33    </plan-body>
34 </plan>
```

Listing 5.2: Example of an iterative plan with given interval between repetitions. The cyclical plan contains the nodes `cyclical-plan-body`,and `repeat-specification`. Within the latter one we can define the period between two iterations.

```
1  <plan name="PLAN_WITH_ITERATIVE_INSTRUCTIONS_every_4-6_hours" >
2     <plan-body>
3        <cyclical-plan>
4           <cyclical-plan-body>
5              <plan-activation>
6                 <plan-schema name="CYCLICAL_PLAN" />
7              </plan-activation>
8           </cyclical-plan-body>
9           <repeat-specification>
10             <retry-delay>
11                <minimum>
12                   <numerical-constant value="4" unit="h" />
13                </minimum>
14                <maximum>
15                   <numerical-constant value="6" unit="h" />
16                </maximum>
17             </retry-delay>
18          </repeat-specification>
19       </cyclical-plan>
20    </plan-body>
21 </plan>
```

Chapter 6

Preprocessing

Before starting with the developing process we had to do some preprocessing. This preprocessing refers to providing required resources, such as the guidelines used for developing rules and lexical resources used by these rules.

6.1 Guideline Selection Process

Asbru is designed to represent protocols rather than guidelines. Unfortunately, protocols are not publicly available from a health carer and thus difficult to obtain. So, we had to use guidelines for our tasks. We first started to establish requirements for choosing guidelines. Using these prerequisites we chose a set of guidelines and divided it into two parts. One part was then used to develop out extraction and transformation rules, the other part was then used to test the rules generated.

6.1.1 Selecting Guidelines

Guideline representation languages are developed for different purposes, with different goals and intentions, different intended users, different applications, and even different specialties. Not every guideline can be optimally represented with Asbru. To develop a system which formalizes treatment flows in Asbru, it is therefore necessary to use guidelines that are more protocols than guidelines. Protocols set out specifically what should happen, when and by whom in the care process. They can be seen as the local definition of a particular care process derived from a more discretionary guideline.

These considerations are included in the criteria that have to be taken into account:

Guideline category. The guideline category is the core of a guideline. There exist various guideline categories, such as diagnosis, education, management, prevention, rehabilitation, or treatment. As the system has to formalize treatment instructions from guidelines, we chose the categories *management* and *treatment*. The scope of guidelines of the management category integrates diagnosis, treatment, and monitoring or follow-up. Guidelines of the treat-

ment category recommend procedures or practices that are intended to relieve physical or mental illness or injury.

Quality. Treatment instructions should be chosen by best medical evidence to provide most effective and efficient therapy for the patient. This requirement is met by medical evidence.

Clinical specialty. To obtain a consistent class of guidelines they should describe the treatment of similar diseases to support the comparison of treatment instructions. Thus, the guidelines should be from only one clinical specialty. Due to the lack of specific knowledge in any clinical domain, the specialty can be chosen as a personal interest.

Temporal dimension. Guidelines should include treatment instructions featuring temporal aspects of flows.

Structure. Structural criteria enable the detection of text modules like tables, lists, and paragraphs. We can assume that the more modules exist the better particular instructions can be detected and arranged.

Number of guidelines. To gain significant results the number of guidelines should be as large as possible. As the number of guidelines is split into two sets at a ratio of 1 : 2 we demand for approximately 20 guidelines.

Using these criteria we searched for appropriate guidelines at the repository of the National Guideline Clearinghouse (NGC)[1]. We obtained several guidelines, especially from the field of otolaryngology. From this specialty we received 18 guidelines that are used for developing and testing the system. Appendix A shows two examples of these guidelines.

6.1.2 Choosing the Test Set

Table 6.1 shows the obtained guidelines with their according developing organizations and the release date of the guideline. These 18 guidelines have been prepared from ten different organizations. To make a choice for an adequate test set of six guidelines, we cannot consult this criteria alone. The organizations have developed between one and six guidelines. After a detailed analysis of each particular guideline it is obvious that the structure of guidelines of a certain developer organization need not be consistent. Thus, we had to make a choice by means of (1) the document structure of the guideline and (2) the disease to be treated.

[1] http://www.guidelines.gov
[2] Centers for Disease Control and Prevention
[3] Institute for Clinical Systems Improvement
[4] University of Michigan Health Systems
[5] Scottish Intercollegiate Guidelines Network
[6] Cincinnati Children's Hospital Medical Center
[7] American Acadademy of Family Physicians
[8] American Academy of Otolaryngology–Head and Neck Surgery
[9] American Academy of Pediatrics

Table 6.1: Obtained guidelines from the specialty of otolaryngology.

#	Title	Organisation	Year	Pages
1	Acute otitis media: management and surveillance in an era of pneumococcal resistance	CDCP[2]	1999	2
2	Acute pharyngitis	ICSI[3]	2003	6
3	Acute rhinosinusitis in adults	ICSI	1999	2
4	Acute sinusitis in adults	ICSI	2004	9
5	Allergic rhinitis	UMHS[4]	2002	5
6	Diagnosis and management of childhood otitis media in primary care.	SIGN[5]	2003	2
7	Diagnosis and treatment of obstructive sleep apnea	ICSI	2004	14
8	Diagnosis and treatment of otitis media in children	ICSI	2004	8
9	Evidence based clinical practice guideline for children with acute bacterial sinusitis in children 1 to 18 years of age	CCHMC[6]	2001	9
10	Evidence based clinical practice guideline for medical management of otitis media in children 2 months to 6 years of age	CCHMC	1998	6
11	Management of sore throat and indications for tonsillectomy.	SIGN	1999	7
12	Otitis media	UMHS	2002	2
13	Otitis media with effusion	AAFP[7]; AAO[8]; AAP[9]	2004	4
14	Pneumococcal vaccination for cochlear implant candidates and recipients: updated recommendations of the Advisory Committee on Immunization Practices	CDCP	2003	1
15	Reduction of the influenza burden in children	AAP	2002	2
16	Rhinitis	ICSI	2003	12
17	Sore throat and tonsillitis	FMSD[10]	2004	4
18	Symptomatic treatment of radiation-induced xerostomia in head and neck cancer patients	PGI[11]	2004	1

Table 6.2: Set of guidelines used for development.

#	Title
2	Acute pharyngitis
4	Acute sinusitis in adults
7	Diagnosis and treatment of obstructive sleep apnea
10	Evidence based clinical practice guideline for medical management of otitis media in children 2 months to 6 years of age
12	Otitis media
16	Rhinitis

Table 6.2 shows the set of guidelines used for developing extraction and transformation rules. The remaining guidelines are then used to test the system.

6.2 Lexical Resources

Medical treatment instructions mostly refer to the administration of drugs, surgical procedures, or other care and prevention measures. Information about drug administration often happens not by the trade name of a drug, but by the name of the drug agent (i.e., *amoxicillin/clavulanate potassium* instead of *Augmentin*). Thus, we need a lexicon containing drug agents and their various pharmacological actions.

The Medical Subject Headings (MeSH) is a controlled vocabulary thesaurus of the National Library of Medicine (NLM)[12]. It consists of sets of terms naming descriptors in a hierarchical structure. 22,997 descriptors are arranged within eleven levels of hierarchy. Additionally, there are more than 151,000 headings called Supplementary Concept Records within a separate thesaurus. For finding the most appropriate MeSH heading there are also thousands of cross-references. These include 24,050 *see* references and 112,012 other entry points [99].

From the regular MeSH thesaurus we used the following:

- Agent terms and their according pharmacological actions
- Surgical procedures
- Diseases

The largest part of the used data are agent terms and pharmacological actions. These were sorted by the pharmacological action (e.g., *analgesics*, *antibiotics*, *anti-infective agents*), whereby an agent may have various pharmacological actions. For example, *aspirin* has four pharmacological actions (1) anti-inflammatory agents,

[10] Finnish Medical Society Duodecim
[11] Practice Guidelines Initiative
[12] http://www.nlm.nih.gov/

non-steroidal; (2) fibrinolytic agents; (3) platelet aggregation inhibitors; and (4) cyclooxygenase inhibitors.

We received 6.029 substances classified in 427 pharmacological actions. Furthermore, we adapted them according for missing terms, different wordings, acronyms, and varying categorization.

Chapter 7

The LASSIE Methodology

Most guideline representation languages are very powerful and thus very complex. They can contain many different types of information and data. We therefore decided to apply a multi-step transformation process [42] called LASSIE (see Figure 7.1). LASSIE is an acronym for 'modeLing treAtment proceSSes using Information Extraction'. The multi-step methodology is necessary, as a one-step or even a two-step modeling process was shown to be not sufficient to the modeler [13, 80].

LASSIE is intended to be a semi-automatic approach. The intermediate representations support the step-wise procedure. To enable the user to observe the particular steps and to encroach and make adaptations after each process step we propose using the Document Exploration and Linking Tool / Addons (DELT/A) (see Section 2.2.1 for details).

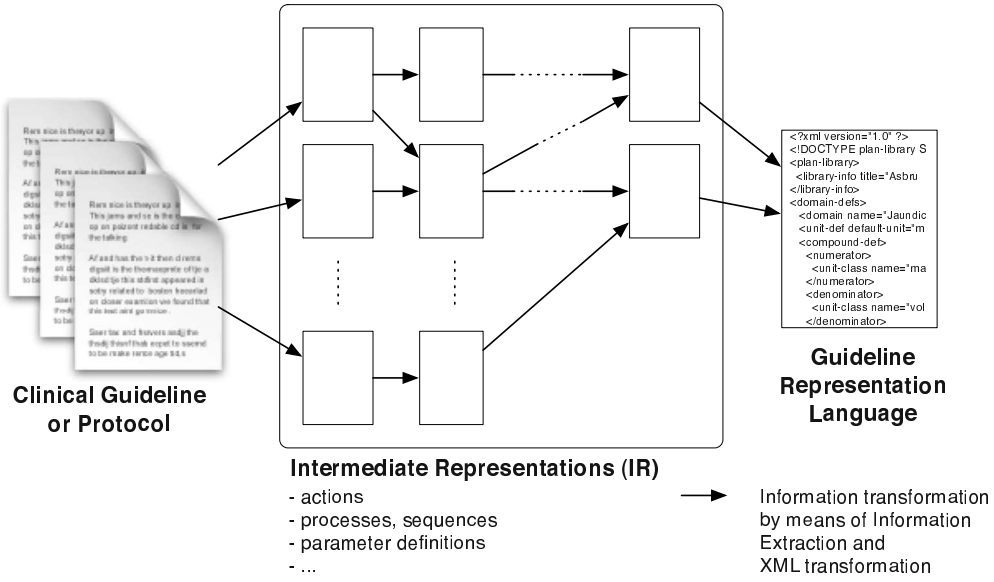


Figure 7.1: LASSIE's guideline transformation process. A multi-step process using intermediate representations to transform clinical guidelines and protocols (CGPs) into a formal representation language.

The following section describes the concept of the intermediate representations. In Section 7.3 we present the steps that extract process information from clinical

guidelines of otolaryngology using heuristic algorithms. These algorithms are based on extraction patterns whose obtaining is described in Section 7.2. The output of this method is a unified format, which can be transformed into the final representation, but which is independent of the final guideline representation language.

In order to create the guideline in the Asbru format Section 7.4 describes how the intermediate representation can be transformed into Asbru.

7.1 Intermediate Representations

LASSIE facilitates the formalization process by using various *intermediate representations* that are obtained by stepwise procedures. Intermediate representations cover the representation of a particular piece of information (i.e., actions, processes, sequences, etc.) [40, 42]. They are XML-based documents and are constructed by refining the preceding representation. Their benefits are:

- Concise formalization process
- Different formats for various kinds of information
- Separate views and procedures for various kinds of information
- Application of specific heuristics for each particular kind of information
- Simpler and more concise evaluation and tracing of each process step

To process as large a class as possible of documents and information we need specific heuristics. These are applied to a specific form of information, for instance:

Different kinds of information. Each kind of information (e.g., processes, parameters) needs specific methods for processing. By presenting only one kind of information the application of the associated method is simpler and easier to trace.

Different representations of information. We have to take into account various ways in which the information might be represented (i.e., structured, semi-structured, or free text).

Different kinds of guidelines. CGPs exist for various diseases, various user-groups, various purposes, various organizations, and so on, and have been developed by various guideline developers' organizations. Therefore, we can speak about different classes of CGPs that may contain similar guidelines.

To transform information by applying Information Extraction (IE) methods, we generated specific templates that can present the desired information. Heuristic methods detect relevant information, which is filled into the templates' slots for subsequent processing.

To put the user in a position to trace each processing step, to encroach upon the process, and to alter or augment the output we have to consider the following items:

- Templates have to be simple and concise – even if the consequence are redundant operations
- For the user it is important to discover, which parts of a guideline are processed and where the corresponding instances in the intermediate representations are. Furthermore, it has to be identifiable which instances of an intermediate representation have been processed to which instances of a succeeding representation.

Hereby, we want to emphasize that is not not our intention to develop new representations for clinical guidelines. The intermediate representations are only means in our approach to semi-automatically generate a formal representation in any guideline representation language.

7.2 Developing Extraction Rules

7.2.1 Pattern Discovery: Choosing the Approach

To apply IE we have to develop rules that – disposed on the text – output relevant information. To develop such rules two approaches are available: (1) the Knowledge Engineering approach and (2) the Automatic Learning approach. Appelt [8] defined a number of considerations that influence the decision to utilize a particular approach:

1. **The availability of training data.** If the required training data is available or cheaply and easily obtainable, the learning approach should be chosen. For complex domain-level tasks, where the annotation task is much slower, more difficult, more expensive, or requires extensive domain expertise from annotators, the KE approach may be favored.
2. **The availability of linguistic resources.** If linguistic resources like lexicons and dictionaries are available developing rules by a knowledge engineer may be possible. Otherwise, it may be necessary to rely on training from the annotated corpus.
3. **The availability of knowledge engineers.** If there is no skilled knowledge engineer, the learning approach should be chosen.
4. **The stability of the final specifications.** If specifications change, it is often easier to make minor changes to a set of rules than to reannotate and retrain the system. However, other changes in specifications may be easier to accomplish with a trainable system.
5. **The level of performance required.** Human skills count for a lot. The best performing systems for various IE tasks have been hand crafted. The performance of automatically trained systems depend on the quantity of available training data. Evaluations at MUCs show that with enough data the automatic training approach can achieve equivalent results to the Knowledge Engineering approach.

As a result of the small set of training documents it is not possible to apply an Automatic Learning approach. The tasks are very complex and difficult and thus a knowledge engineer is better qualified to accomplish them. Linguistic resources, such as a lexicon are also available.

Thus, we will apply the Knowledge Engineering approach. Therefore, we develop extraction rules by analyzing a set of training documents and identify patterns. These patterns are both delimiter-based and linguistic-based. This makes LASSIE a hybrid between NLP-based IE systems and wrapper systems.

7.2.2 Finding the Happy Medium

Usually, we learn rules using some set of training examples. We assume it is possible to reach a state where we will also be able to extract the correct information for other examples, thus generalizing to situations not presented during training. We aim for a compromise between bias, where our model does not follow the right trend in the data, and variance, where our model fits the data points too closely. These two extremes are known as *underfitting* and *overfitting*.

An important concept in this context is the number of parameters in a model. As this number increases, the model can bend in more complicated ways. If the number of parameters in our model is larger than in reality, then we risk overfitting, and if our model contains fewer parameters, we could underfit. However, especially in cases where training examples are rare, we may adjust to very specific random features of the training data that have no causal relation to the target function.

In both statistics and machine learning, in order to avoid overfitting, additional techniques (e.g., cross-validation, early stopping) are used that can indicate when further training does not result in better generalization. By manually developing rules it relies on the knowledge engineer to generalize rules in the right way.

Another important aspect when developing rules is the weight of *recall* and *precision*. These two scores are statistical measures and form an interdependent pair. The recall score measures the ratio of correct information extracted from the texts (COR) against all the available information present in the text (POS). The precision score measures the ratio of correct information that was extracted (COR) against all the information that was extracted (ACT) [57].

$$recall = \frac{COR}{POS}$$

$$precision = \frac{COR}{ACT}$$

In other words: Recall measures how well a system finds what you want, and precision measures how well it weeds out what you do not want. Ideal systems are those with both high recall and high precision. But often this is not possible: there has to be a trade-off. Only one measure can be optimized at the cost of the other. Therefore, we have to consider for each task how to weight the scores.

7.2.3 Obtaining Patterns

Modeling a guideline is not an easy task with only one valid outcome. If human modelers accomplish the tasks they may obtain varying instances of the templates from one guideline. Even if a modeler repeats the modeling of one guideline different instances may result. Thus, we can only develop an approximation of rules and we will consider this in our evaluation.

Guidelines are a mixture of grammatical and telegraphic text and have additional formatting information. For detecting patterns we used OpenNLP[1], a package of java-based NLP tools which perform sentence detection, tokenization, parts-of-speech tagging, chunking and parsing, and named-entity detection using maximum entropy models. We used its parts-of-speech tagging to obtain grammatically induced rules. Unfortunately, the structure of the text in the guidelines does not afford inferences on grammatical relationships. But we can establish patterns by means of the appearance of particular words and phrases similar to Snowball [4] (compare Section 3.2.2).

Extraction patterns for these guidelines are based on syntactical and semantical constraints as well as delimiters. Patterns are defined on three levels, whereas patterns at a certain level serve as concept classes in the preceding levels: (1) phrase level patterns, (2) sentence level patterns, and (3) discourse level patterns.

Phrase Level Patterns

They are used for identifying basic entities, which build the attributes of actions. They are defined by regular expressions. Table 7.1 shows examples of phrase level patterns.

Sentence Level Patterns

They use phrase level patterns, medical terms obtained from the lexicon, and trigger words for the medical terms to identify medical actions and their attributes. The trigger words are mainly verbs and indicate the application of a therapy (e.g., the administration of a drug agent or the implementation of a surgical procedure) or the avoidance of a therapy. Sentence level patterns are delimiter-based and use syntactic constraints. We can categorize the patterns in two groups: (1) patterns for free text and (2) patterns for telegraphic text.

Free text patterns. These patterns can be applied to free text, which has a grammatical structure and is usually identified in paragraphs (i.e., between `<p>` and `</p>` tags), but also in list elements (i.e., between `<li>` and `</li>` tags).

These patterns indicate that therapy instruments (i.e., agent terms and surgical procedures) combined with trigger terms (e.g., activate, indicate, perform, prescribe) appearing in the same clause identify relevant sentences. The particular clauses must not be `<condition>` clauses. Phrase level patterns, such

[1] http://opennlp.sourceforge.net

Table 7.1: Examples of phrase level patterns.

<number>	([\d]+(([\.]([\d]+))|((\s*[\d]+)?/[\d]+))?)
<numberOrRange>	<number>(((␣to␣)|(\s*-\s*))<number>)?
<time-unit>	m(illi)?)?sec(ond)?(s)?|min(ute)?(s)?|hour(s)?| day(s)?|week(s)? ...
<dose-unit>	(m|c|d)?(l|g)(/kg(/<time-unit>)?)?|drop(s)?| tab(s)? ...
<dosage>	<numberOrRange>[\s]*<dose-unit>
<time>	<numberOrRange>[\s]*<time-unit>
<iteration>	TID|BID|QD|(Q|every)␣<time>| <numberOrRange>␣(times|doses)␣(per|a)␣<time-unit>
<person>	those|patient(s)?|person(s)?|child(ren)? ...
<condition>	(in␣(case(s)?|areas)|if|unless|who(m)?)␣[^,:]+ | in␣.*allergic [^,\.:]+ | (for|in)␣(a␣)?(<person>) [^,\.:]+

as <dosage>, <duration>, <condition>, and so on can be arbitrarily combined with <therapy instrument> <trigger> pairs.

Information concerning a treatment recommendation can be distributed in several sentences. These sentences including additional information (e.g., *'The standard dose is 40 to 45 mg/kg/day.'*) neither contain a therapy instrument nor a trigger term. Listing 7.1 shows some sentence level patterns for free text.

Listing 7.1: Examples of sentence level patterns for both free and telegraphic text. The ordering of the items is irrelevant.

```
<therapy instrument><trigger>
<therapy instrument><trigger><duration>
<therapy instrument><trigger><iteration>
<therapy instrument><duration>
<therapy instrument><iteration>
<therapy instrument><dosage>
<condition><therapy instrument><trigger>
<dosage>
<iteration>
<trigger><duration>
```

Telegraphic text patterns. These patterns are applicable in list elements. In list elements often ungrammatical text is formulated and therefore, there is no need for trigger terms. Often, only a therapy instrument indicates the relevancy of an element (cp. Figure 7.2).

Other patterns exist for list elements indicating that these elements are rele-

- Amoxicillin/clavulanate potassium
- Cefuroxime axetil
- Ceftriaxone sodium

Figure 7.2: Example of list elements in a guideline for diagnosis and treatment of otitis media.

vant if within their context or in the paragraph preceding the list special terms appear (see Listing 7.2). These terms (i.e., remedy, remedies, measure, measures, medication, medications) are important, because they specify actions that may not contain therapy instruments in the form of agent terms or surgical procedures (e.g., *'Maintain adequate hydration (drink 6 to 10 glasses of liquid a day to thin mucus)'*).

Listing 7.2: Examples of sentence level patterns for telegraphic text. The ordering of the items is irrelevant.

```
<therapy instrument>
<therapy instrument><dosage><duration>
<therapy instrument><dosage><iteration>
<therapy instrument><dosage><duration><iteration>
<therapy instrument><duration><iteration>
<condition><therapy instrument><duration><iteration>
<condition><therapy instrument><dosage><iteration>
<condition><therapy instrument><dosage>
<measure term in context>
```

Discourse Level Patterns

These patterns are used to

1. Categorize sentences
2. Merge them to actions
3. Find relationships between actions to structure them.

They are based on *sentence level patterns*, but are augmented to consider the structure and the layout of the documents.

7.3 Gaining Process Information

7.3.1 Some Remarks to the Content of Guidelines

Medical knowledge can be organized in three categories:

1. Common knowledge that the medical personnel (i.e., physicians, nurses, etc.) must know in order to accomplish their daily work

2. Knowledge that is not commonly known, but that is required for the most effective and efficient treatment

3. Knowledge that is not commonly known and that is not required for the personnel's daily routine (e.g., knowledge about diseases that are no longer occurring)

Guidelines are developed for various user groups. Guidelines for intended users of physicians, nurses, and other care personnel contain information belonging to the second category. Knowledge of the first category (e.g., what is 'fever' and at which blood temperature do we define 'high fever'?) is not explicitly described in most of these guidelines.

If guidelines have to be executed in a computer-supported way it is necessary to model this implicit information. Due to the lack of this knowledge in an explicit form, it is not possible to model this information automatically. In our system we are restricted to the processing of explicitly described information in the guidelines. If additional knowledge is required it has to be added by a domain expert.

7.3.2 How to Extract Processes

To extract processes from clinical guidelines we proceed in several steps. These serve to filter segments of text containing treatment instructions from the documents and to generate processes.

We propose a two-step approach to gain a representation that is independent of the subsequent guideline representation language.

The first step is to extract the relevant sentences containing treatment instructions by marking-up the original guideline document. This is explained in Section 7.3.3. A marked-up guideline and the corresponding representation containing the extracted sentences is shown in Sections B.1 and B.2.

The subsequent step is to combine several sentences to one action and to structure the actions and detect relations among them. This step is described in Section 7.3.4. A resulting representation of this task's output is shown in Section B.3.

These two steps should provide a basis for the subsequent transformation of the process information into any guideline representation language.

7.3.3 Extracting Relevant Sentences

This task is a first step towards our final guideline representation. We will achieve it by two modules: (1) the segmentation & filtering module and (2) the template generation module (see Figure 7.3 for an overview).

This first intermediate step is especially important as not the entire content of a guideline contains processes, which are to be modeled. Only about 20 % of sentences of a guideline are of interest for modeling processes. On this account it is important to select the relevant sentences for modeling.

Thus, this task performs an automatic mark-up of sentences that are utilized to process the subsequent steps.

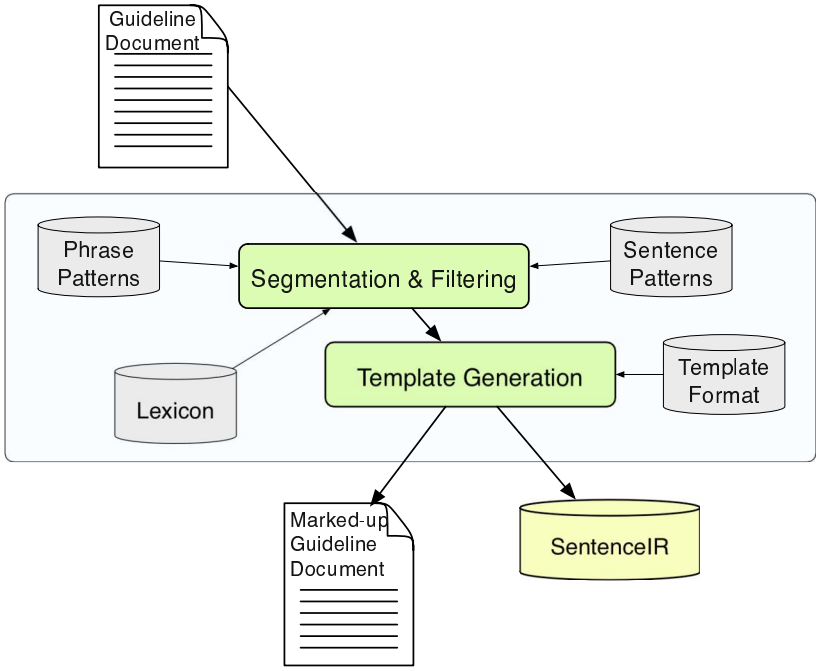


Figure 7.3: Detecting relevant sentences. We split this task into two modules: (1) segmentation & filtering module and (2) template generation module.

Segmentation & Filtering

Detecting relevant sentences is a challenging task, which we undertake in two steps: (1) detecting irrelevant sentences to exclude them from further processing and (2) detecting relevant sentences.

As in many IE systems only parts of the document are of interest. We divide the document text into relevant and irrelevant sentences. Relevant sentences describe actions of a treatment processes, irrelevant sentences are associated with diagnosis, symptoms, or etiology.

The first filtering occurs at the section level. Sections in the document with captions indicating diagnosis or symptom declarations have to be omitted in further processing. We can identify these captions by keywords such as 'history', 'diagnosis', 'criteria', 'symptom', 'clinical assessment', 'risk factor', 'complicating factor', 'etiology', and so on.

Detecting relevant sentences is not a trivial task. First, we parse the entire document and split it into sentences. Then we process every sentence with regard to its context within the document and its group affiliation. Thereby, the context is obtained by captions (e.g. *'Acute Pharyngitis Algorithm Annotations | Treatment | Recommendations:'*) and a group contains sentences from the same paragraph or the same list, if there are no sublists.

Each sentence is now checked for relevance. This task is accomplished by applying *sentence level patterns*.

Template Generation

Now that we have collected the relevant sentences from the guideline, we can proceed with generating the intermediate representation *SentenceIR*. We generate two files: one file listing all relevant sentences (Listing 7.3 shows an example instance of the representation of a sentence) and the marked-up guideline document (Listing 7.4 shows the source of a marked-up guideline document). Both are linked by applying the same id to the same sentences.

Thus, the presentation of the template file and the guideline document are as simple as possible in order to support the user by detecting all relevant sentences.

Listing 7.3: Sentence instances of a *SentenceIR* template.

```
<sentence>
   <delta-link link-id="8"/>
   <description>In children with risk factors for Streptococcus
       pneumoniae, it is recommended that Amoxicillin, high
      dose (80 to 90 mg/kg/day) or Augmentin (with high dose
      amoxicillin component) be utilized as first-line therapy
       (Nash and Wald, 2001 [S]; Wald, Chiponis, and
      Ledesma-Medina, 1986 [B]; Nelson, Mason, and Kaplan,
      1994 [C]; Dowell et al., 1999 [E]; Dowell, 1-1998 [E];
      Friedland and McCracken, 1994 [E]; Local Expert
      Consensus [E]).
   </description>
</sentence>
<sentence>
   <delta-link link-id="9"/>
   <description>Note: Failure with amoxicillin is likely to be
      due to resistant Streptococcus pneumoniae, Haemophilus
      influenzae, or Moraxella catarrhalis.
   </description>
</sentence>
<sentence>
   <delta-link link-id="10"/>
   <description>High dose amoxicillin will overcome
      Streptococcus pneumoniae resistance (changes in
      penicillin-binding proteins) (Dowell et al., 1999 [E];
      Whitney et al., 2000 [D]).
   </description>
</sentence>
```

Listing 7.4: Excerpt of a source listing of a marked-up guideline document. Relevant sentences are enclosed by HTML-like ”a” tags.

```
<li>
   <a id="delta:8">In children with risk factors for
      Streptococcus pneumoniae, it is recommended that
      Amoxicillin, high dose (80 to 90 mg/kg/day) or Augmentin
       (with high dose amoxicillin component) be utilized as
      first-line therapy (Nash and Wald, 2001 [S]; Wald,
      Chiponis, and Ledesma-Medina, 1986 [B]; Nelson, Mason,
      and Kaplan, 1994 [C]; Dowell et al., 1999 [E]; Dowell, 1
```

```
          -1998 [E]; Friedland and McCracken, 1994 [E]; Local
          Expert Consensus [E]).
3     </a>
4     <ul type="disc">
5        <li>
6           <a id="delta:9">Note: Failure with amoxicillin is
               likely to be due to resistant Streptococcus
               pneumoniae, Haemophilus influenzae, or Moraxella
               catarrhalis.
7           </a>
8           <a id="delta:10">High dose amoxicillin will overcome
               Streptococcus pneumoniae resistance (changes in
               penicillin-binding proteins) (Dowell et al., 1999
               [E]; Whitney et al., 2000 [D]).
9           </a>
10          The clavulanic acid component of Augmentin is active
11          against resistant Haemophilus influenzae and Moraxella
12          catarrhalis (B-lactamase enzyme) (Wald, Chiponis, and
13          Ledesma-Medina, 1986 [B]; Dagan et al., 2000 [A]).
14       </li>
15    </ul>
16 </li>
```

An entire marked-up guideline and the corresponding *SentencIR* representation is shown in Sections B.1 and B.2.

7.3.4 Extracting Required Information and Finding Processes

The information contained in *SentenceIR* and the marked-up guideline document are the input for the next task (see Figure 7.4 for an overview). Its intention is to structure relevant sentences and find relationships between sentences. Again, the output of this task should be represented in a format that is independent of any desired guideline representation format.

Structure Extraction

In the preceding task we only obtained the sentences necessary for building a formal model of the treatment recommendations. But we have to know their context and how they are arranged in the document to receive a structure. This structure is very important as we will be geared to it during the other subtasks, especially the merging and grouping and the process extraction.

Thus, we obtain the structure by means of hierarchical groups. Every action is assigned to one group. The context of a sentence defines the affiliation to a group. A sentence's context can be defined by the sentence's position in the hierarchal structure. We use the superior headings that establish several context items.

Slot Extraction

This module is used to extract the following information:

- Therapy instruments (i.e., agent terms and surgical procedures)

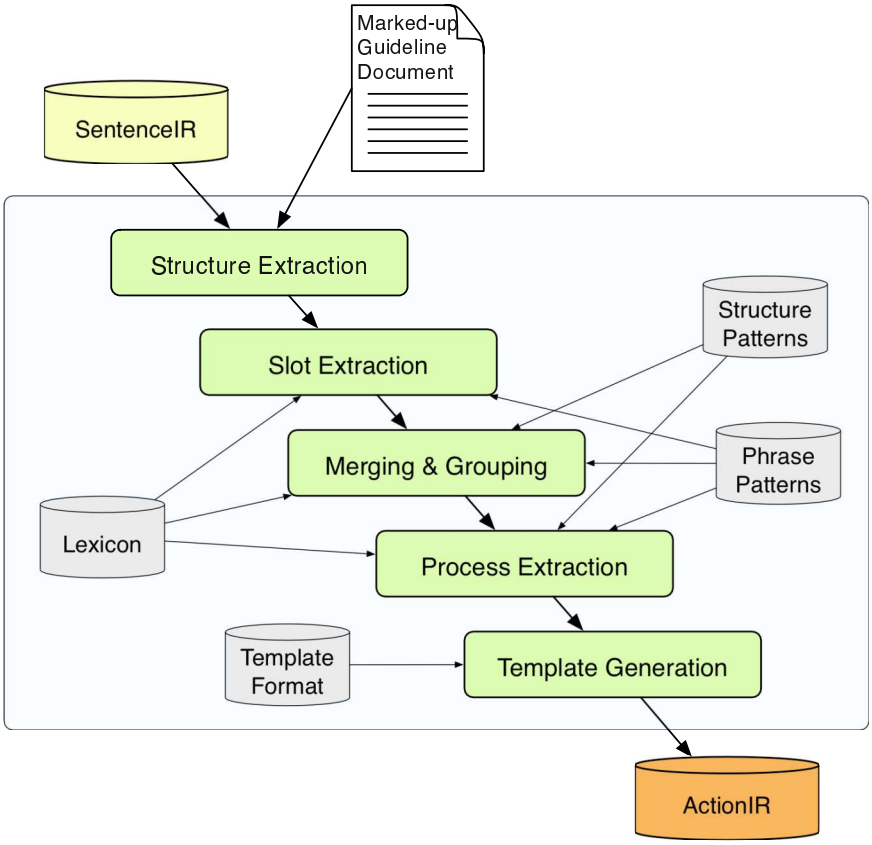


Figure 7.4: Finding processes and extracting required information. We split this task into five modules (i.e., the structure extraction module, the slot extraction module, the merging & grouping module, the process extraction module, and the template generation module).

- Dosage information in case of a drug administration
- Duration of the therapy action
- Iteration information of the action
- Conditions which have to be fulfilled to perform an action

It uses both the lexicon and the *phrase level patterns*.

Merging & Grouping

In this module we categorize sentences in actions or negative actions and annotations. Annotations always belong to at least one action (or negative action). They cannot exist alone. This module extensively applies *discourse level patterns*.

First, we check whether a sentence describes an action or a negative action. Negative actions are instructions that an action should not be performed, often under specific conditions (e.g., *'Do not use aspirin with children and teenagers because it may increase the risk of Reyes syndrome.'*). Most guideline representation languages will handle such actions by inverting the condition. Languages may exist which will handle these in other ways. Therefore, we will provide a representation for such actions that can be used in a general way.

Explicit annotations containing keywords, such as 'Note:' or 'Notice', are easy to discover. To detect implicitly described annotations we have developed a special heuristic. This is done by checking whether the therapy instrument in the form of an agent term or a surgical procedure in a sentence also appears in processes belonging to the same group appearing above this sentence in the text. If this happens, the sentence is added as an annotation to all these processes. The assigning of an annotation to actions falls back on approaches of coreferencing. We use name-alias coreferencing and definite description coreferencing based on therapy instruments and their hypernyms. We do not apply a pronoun-antecedent coreference.

Explicitly mentioned annotation sentences are added to each process of the same group appearing before the annotation sentence. If there are no sentences in front of the annotation, the annotation sentences are added to each process of the parent group.

If the sentence is not an annotation, it is added as a new process and augmented by additional information extracted in the preceding module.

Process Extraction

By analyzing treatment processes of otolaryngology contained in the guidelines we detected the following processes:

- Sequential processes
- Processes without temporal dependencies
- Processes which exclude each other (i.e., one process has to be selected from several)

- Processes containing subprocesses
- Recurring processes

To group actions and to detect relationships between actions we use *discourse level patterns*. We will describe those used by this module below.

The default relationship among processes is that there is no synchronization in their execution.

To group actions to a *selection* they must fulfill the following requirements: (1) the actions have to belong to the same group, and (2) agents or surgical procedures must have the same superordinate. For instance, processes describing the administration of *Erythromycin*, *Cephalexin*, and *Clindamycin* within one group are combined in a *selection*, as all these agents are antibiotics. If actions are grouped in a selection, one of these actions has to be selected to be executed. This algorithm works well for CGPs of otolaryngology. There may be other specialties where this algorithm has to be altered for other agent categories.

Furthermore, we try to detect relations between actions that are explicitly mentioned within the text as well as relations that are implicitly given by the document structure. The former is very difficult to detect, as we often cannot detect the reference of the relation within the CGP (e.g., *'After 10 to 14 days of failure of first line antibiotic ...'*). Nevertheless, we found heuristics that arrange actions or action groups if the reference is unambiguously extractable out of the text. These heuristics can be grouped in two categories: (1) detecting sentences describing relations between actions, and (2) detecting actions that are described in the preceding heuristic.

A relation is mainly identifiable by a relation term (e.g., 'before', 'after', 'during', 'while'). If such a term appears, we are searching for therapy instruments, as these describe most of our actions. After we have detected these terms, we search for actions containing the particular instruments. If we have found both the source action and the destination action we can create a new relation.

Implicitly given relations by the document structure are detected by patterns of the document structure (e.g., *'Further Treatment'* appears `after` *'Treatment'* or *'Treatment'* appears `before` *'Follow-Up'*). Thereby, we developed structure patterns which are part of discourse level patterns. They are used to determine the relations between several groups.

Template Generation

The template of this intermediate representation has to contain actions as well as their relations. It has to be simple and concise and it has to illustrate from which original data the current information was built. We split the new *ActionIR* template in three parts: (1) an area for actions, (2) an area for relations, and (3) an area for the structure illustrating the hierarchy and nesting of groups.

An **action** contains the sentence describing the action and the annotation sentences that were assigned. Furthermore, it contains the treatment instruments and their MeSH ids. If information about the dosage, duration, or iteration of a drug administration was extracted it is stated, too. Also, any conditions are specified. If

the action is part of a selection, it is stated by the selection id. DELT/A links are inherited from the *SentenceIR* representation in order to provide the traceability of the process from both the original guideline document and the *SentenceIR* document. Listing 7.5 shows an example instance.

Listing 7.5: Action instance of an *ActionIR* template.

```
<action id="8" parent="5" group="18" selection="0">
   <delta-link link-id="8"/>
   <description>In the child with no risk factors for
      penicillin-resistant Streptococcus pneumoniae standard
      dose amoxicillin or Augmentin (with standard dose
      Amoxicillin component) may be considered as initial
      therapy.
   </description>
   <agents>
      <agent MeSH="D000658" name="amoxicillin"/>
      <agent MeSH="D019980" name="Augmentin"/>
   </agents>
   <condition>
      <item>In the child with no risk factors for
         penicillin-resistant Streptococcus pneumoniae
      </item>
   </condition>
   <annotations>
      <annotation>Note: Forty-six percent of isolates at
         Children's Hospital Medical Center of Cincinnati,
         Ohio have intermediate or high Penicillin-resistant
         Streptococcus pneumoniae and local data supports that
          15% of children locally may fail initial therapy
         with standard dose amoxicillin.
         <delta-link link-id="9"/>
      </annotation>
   </annotations>
   <context>
      <item>Antibiotic Treatment</item>
   </context>
</action>
```

Relations are stated by their type (e.g., succeeding, preceding, overlapping) and the concerned actions by their DELT/A ids. Listing 7.6 shows an example.

Listing 7.6: Relation instance of an *ActionIR* template.

```
<relation type="SUCCEEDING">
   <source id="13">
      <delta-link link-id="13"/>
   </source>
   <destination id="17">
      <delta-link link-id="17"/>
   </destination>
</relation>
```

Apart from actions and their relations the **structure** of the document is given illustrating the nesting of the groups and selections.

The entire *ActionIR* representation of a guideline is shown in Section B.3.

7.4 Application: Modeling Plans in *Asbru*

Now we want to verify whether the information obtained by the process extraction is in a format that can be utilized to transform it into a guideline representation. We have chosen the guideline representation language **Asbru** (see Section 2.1.1 for details), which was developed at Vienna University of Technology, Stanford Medical Informatics, University of Newcastle, and Ben Gurion University, to embody CPGs as time-oriented skeletal plans [82].

We want to process the intermediate representation *ActionIR* to generate an *Asbru* guideline. In Chapter 5 we have analyzed the requirements for modeling plans in the *Asbru* language. This is the basis for our next steps.

7.4.1 A Step Towards *Asbru*

Asbru is a task-specific and intention-based guideline representation language. Its syntax is very extensive and its grammar is very complex. It allows for the modeling of time-oriented and skeletal plans.

To obtain the necessary information required for Asbru, we have to paste additional information. This – still intermediate – representation will be a preliminary stage to Asbru. Figure 7.5 gives an outline.

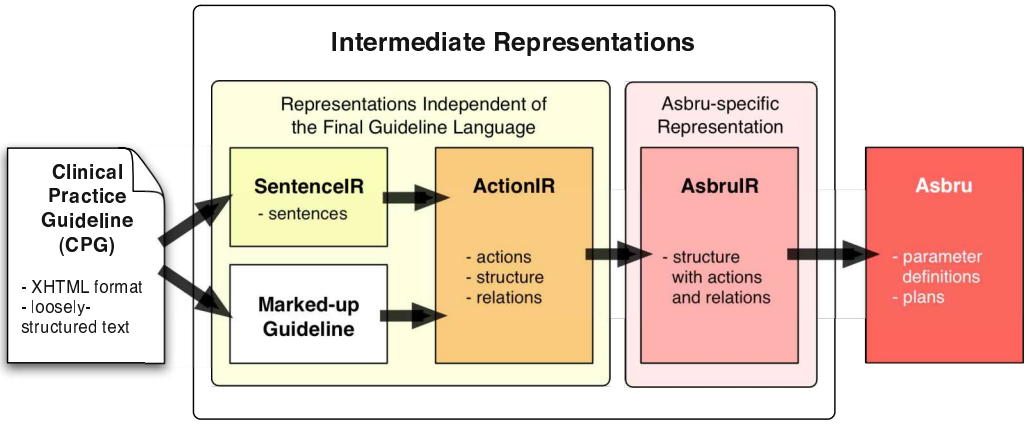


Figure 7.5: Steps to (semi-)automatically obtain an Asbru guideline.

The *TimeWrap* method [44] has been developed to support the transformation of structured treatment information into Asbru. Starting from a formal base representation, which is similar to *ActionIR*, but less expressive, we extract Asbru-specific information and integrate it in an other intermediate representation. This new format contains several actions and their temporal specifications to generate both atomic and cyclical plans in Asbru.

For our current purposes we need a more powerful representation to state conditions as well as the relationships between several actions. The intermediate representation **AsbruIR** contains the following refined Asbru-specific data:

- Temporal information of durations and iterations are refined and calculated if necessary
- Conditions are refined and classified in categories (e.g., 'diagnosis', 'patient', 'allergy') if possible
- Interval relations are modeled in a way defined in the preceding section
- Actions and structure are merged
- Auxiliary actions are inserted

Figure 7.6 gives an overview of the steps required to obtain the new representation.

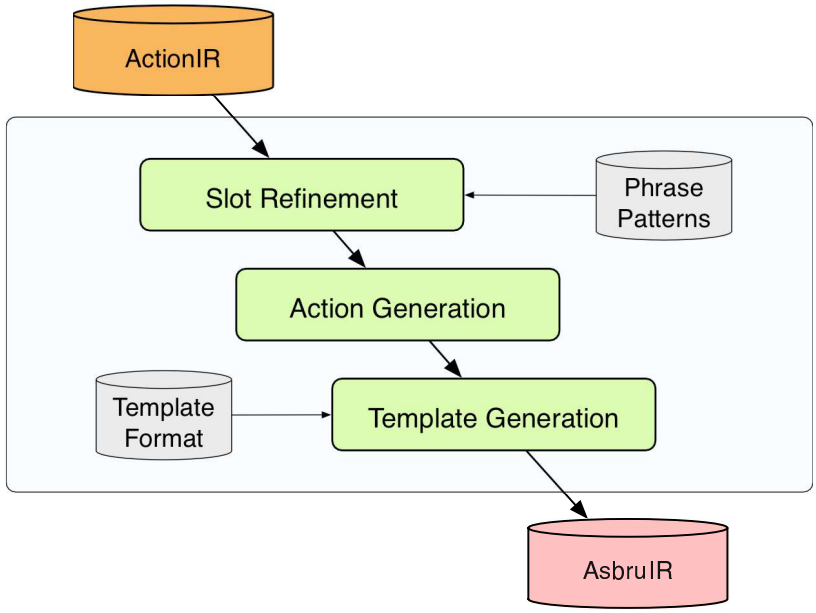


Figure 7.6: Refining process information and generating auxiliary actions. We split this task into three modules (i.e., slot refinement module, action generation module, and template generation module).

Slot Refinement. A lot of the information we have extracted so far is not in an Asbru-interpretable format. For instance, we extracted phrases, such as *every 4 to 6 weeks*, to describe an iteration. To use it in Asbru the information has to be itemized. Therefore, we use phrase patterns with an Asbru-specific emphasis. Until now, we only 'know' that the phrase contains information about an iteration. So we itemize values and units. In case of an iteration we also define its type (see Section 5.4.2 for details). Similarly we proceed with 'duration' instances.

To better cope with various conditions we classify them into patient-, disease-, and allergy-specific conditions. Furthermore, in case of a negative action we assign each condition.

In case of interval relations we generate new conditions. For instance, we model sequential relations by means of a simple condition: the succeeding action gets a 'setup-precondition' with the constraint that the preceding action has to be in state 'completed'.

Action Generation. This module addresses the merging of the action section and the structure section as well as the generation of auxiliary actions.

The *structure* section of the *ActionIR* representation contains the hierarchy of the guideline document and is now filled with the particular actions.

In case we detected *select-one-of* actions we have to generate a special action which controls the choice of the particular actions. Although this kind of plan often appears in guidelines Asbru does not provide a separate modeling concept. Thus, the modeling is elaborate and we will support it in this intermediate representation.

Select-one-of actions must have a parent action able to invoke each single action. To sustain the constraint to execute only one action the parent action has to complete after the finishing of the selected action. For this intermediate representation it will be sufficient to define an additional condition in the newly generated parent action. Listing 7.7 shows an example.

Listing 7.7: Example of a parent action constituting a *select-one-of* relation for actions with plan id '6' and '7'.

```
<action id="SELECT_0" parent="5" group="16">
   <condition>
      <item type="complete-condition" plan="6"
         state="completed"/>
      <item type="complete-condition" plan="7"
         state="completed"/>
   </condition>
</action>
```

Further auxiliary actions have to be generated in order to represent the entire guideline hierarchy. We can achieve this by means of the action's context items. If an action contains context items, but no parent action, we generate such parent actions using the context items.

Template Generation. For this representation we focus on automatic processing to generate an Asbru protocol. For the human user it is difficult to get an overall outline of the guideline as actions are displayed in their hierarchical structure. Due to the refined itemization the information is not so easily readable (in the sense of understandable). Therefore, we additionally state it as string information.

Nevertheless, an evaluation of this representation is less laborious as the information is only refined and only some simple rules are applied for this task.

Listing 7.8 shows an example of an *AsbruIR* instance.

Listing 7.8: Instance of an *AsbruIR* template.

```
<action id="25" parent="SELECT_1" group="18" selection="1">
   <delta-link link-id="25" />
```

```
3     <description>For those allergic to amoxicillin:
          Trimethoprim-sulfamethoxazole (TMP/SMX): one double
          strength tab BID 10 days
4     </description>
5     <agents>
6        <agent MeSH="D015662" name="Trimethoprim-sulfamethoxazole
             ">
7           <iteration term="BID" specification="CYCLICAL">
8              <frequency value="12" unit="h" />
9              <minimum value="2"  unit="h" />
10             <maximum value="10" unit="h" />
11          </iteration>
12          <duration term="10 days">
13             <minimum value="10" unit="d" />
14             <maximum value="10" unit="d" />
15          </duration>
16       </agent>
17    </agents>
18    <condition>
19       <item allergy="amoxicillin">For those allergic to
             amoxicillin</item>
20    </condition>
21    <annotations>
22       <annotation>Trimethoprim-sulfamethoxazole (TMP/SMX) is a
             potential first-line antibiotic.
23          <delta-link link-id="28" />
24       </annotation>
25       <annotation>Studies have shown effectiveness with 3 to 14
              days.
26          <delta-link link-id="32" />
27       </annotation>
28    </annotations>
29 </action>
```

An entire *AsbruIR* representation of a guideline is shown in Section B.4.

7.4.2 The Last Step: Obtaining *Asbru*

The final step in our approach is the generation of an Asbru guideline. We will accomplish this task using XSLT[2].

We will therefore build several XSLT templates to transform the required Asbru segments. If these segments are combined they might represent an Asbru guideline. Starting from the intermediate representation *AsbruIR* the templates have to map the following concepts: plans and parameter definitions.

Plans. They are the basic building blocks of Asbru guidelines. Our XSLT templates have to constitute several kinds of plans which arise from *AsbruIR*.

- *Atomic plans.* This plan cannot be refined. Example: user-performed plans.

[2]XSLT is an acronym for XSL transformation. XSL is an acronym for Extensible Stylesheet Language. It is a XML-based language used for the transformation of XML documents.

- *Cyclical plans.* This plan iteratively calls another plan. The modeling depends on whether the iteration is frequency-based or period-based.

- *Plans containing subplans.* Plans can be nested. Due to the hierarchical structure of guidelines this is a very frequent concept. Asbru provides the possibility to model various kinds of synchronization among the subplans.

Plans contain several optional child elements. For our modeling we only use `conditions` and `plan-body`.

- *Conditions.* They control the execution of a plan by enabling plan instances to receive a certain plan state. Figure 2.3 on page 8 describes the conditions that must apply to switch from one state to an other. Using the `conditions` concept it is also possible to accomplish a state transition manually.

  We will use the `setup-precondition` to activate a plan and the `complete-condition` to finish a plan.

  Conditions can only map temporal constraints. The concept of `conditions` as appearing in *ActionIR* is not equal to Asbru's `conditions` concept and thus has to be modeled by other concepts which appear in the `plan-body` element.

- *Plan-body.* It contains the information about the events when executing the plan. There are several possibilities, which are all described in [79]. For our modeling we only use

  – Subplans. A set of plan steps performed in parallel or sequentially.

  – Cyclical-plan. A plan repeated several times.

  – Single-step. A single step of plan execution; for instance, a plan activation.

  – User-performed. This plan is executed through some action by the user, which is not further modeled in the system.

Parameter definitions. In Asbru variables are referred to as parameters. Parameters may have quantitative as well as qualitative character. They are defined in an own section (i.e., the `domain-defs`). In many cases the definition of parameters for guidelines has to be done manually as they often reflect implicit or tacit knowledge.

In many condition statements we use parameters. Therefore, we will define the necessary parameters. Listing 7.9 shows an example of parameter definitions for a guideline.

Listing 7.9: Parameter definitions for an Asbru guideline.

```
1 <domain-defs>
2    <domain name="Acute_pharyngitis">
3       <parameter-group title="Diagnosis parameters">
4          <parameter-def name="for_treatment_of_culture_
              positive_cases_of_GABS_pharyngitis" required="no"
              type="any">
```

```
5              <raw-data-def mode="manual" use-as-context="no"
                   user-text="Diagnosis is GABS pharyngitis?"/>
6           </parameter-def>
7           <parameter-def name="for_streptococcal_pharyngitis"
               required="no" type="any">
8              <raw-data-def mode="manual" use-as-context="no"
                   user-text="Diagnosis is streptococcal
                   pharyngitis?"/>
9           </parameter-def>
10          <parameter-def name="for_strep_pharyngitis"
               required="no" type="any">
11             <raw-data-def mode="manual" use-as-context="no"
                   user-text="Diagnosis is strep pharyngitis?"/>
12          </parameter-def>
13          <parameter-def name="for_treatment_of_
               GABS_pharyngitis" required="no" type="any">
14             <raw-data-def mode="manual" use-as-context="no"
                   user-text="Diagnosis is GABS pharyngitis?"/>
15          </parameter-def>
16       </parameter-group>
17       <parameter-group title="Allergy parameters">
18          <parameter-def name="In_PCN-allergic_patients"
               required="no" type="any">
19             <raw-data-def mode="manual" use-as-context="no"
                   user-text="Patient is allergic to PCN?"/>
20          </parameter-def>
21          <parameter-def name="In_PCN-_and_erythromycin-
               allergic_patients" required="no" type="any">
22             <raw-data-def mode="manual" use-as-context="no"
                   user-text="Patient is allergic to PCN and
                   erythromycin?"/>
23          </parameter-def>
24       </parameter-group>
25    </domain>
26 </domain-defs>
```

Using different templates we are able to generate Asbru plans with various levels of detail. For instance, we can create plans at a very low level where an action in an *AsbruIR* representation corresponds to exactly one Asbru plan – no matter if the action contains iterative instructions or a bundle of treatment instructions. The highest level in Asbru modeling would be the itemization of iterative instructions and further subplans.

AsbruIR is a mixture between single- and multi-slot templates. It consists of actions, whereas one action might consist of several sentences (i.e. action sentence and several annotation sentences), which form one action. But this action may consist of several sub-actions or several alternative sub-actions. This is shown in the `agents` section of the template: each agent is separately presented and can therefore form an own plan in Asbru.

Listing 7.10 shows an *AsbruIR* representation of a plan. Listings 7.11 to 7.13 show the corresponding Asbru plans.

Listing 7.10: Instance of a plan represented in *AsbrulR*.

```
<action id="24" parent="SELECT_1" group="18" selection="1">
   <delta-link link-id="24"/>
   <description>Amoxicillin: 500 mg tab three times per day (
      TID) 10 days or 875 mg tab two times per day (BID) 10
      days
   </description>
   <agents>
      <agent MeSH="D000658" name="Amoxicillin">
         <dosage rate="500 mg"/>
         <iteration term="three times per day"
             specification="CYCLICAL">
            <frequency value="8" unit="h" />
            <minimum value="1" unit="h" />
            <maximum value="7" unit="h" />
         </iteration>
         <duration term="10 days">
            <minimum value="10" unit="d"/>
            <maximum value="10" unit="d"/>
         </duration>
      </agent>
      <agent MeSH="D000658" name="Amoxicillin">
         <dosage rate="875 mg"/>
         <iteration term="two times per day"
             specification="CYCLICAL">
            <frequency value="12" unit="h" />
            <minimum value="2" unit="h" />
            <maximum value="10" unit="h" />
         </iteration>
         <duration term="10 days">
            <minimum value="10" unit="d"/>
            <maximum value="10" unit="d"/>
         </duration>
      </agent>
   </agents>
   <annotations>
      <annotation>Amoxicillin is a potential first line agent.
         <delta-link link-id="26"/>
      </annotation>
      <annotation>Studies have shown effectiveness with 3 to 14
          days.
         <delta-link link-id="32"/>
      </annotation>
   </annotations>
</action>
```

Listing 7.11 shows the main plan. It consists of a `conditions` section and a `plan-body` section. The plan recommends the administration of Amoxicillin and provides two different methods. They differ in the dosage and the number of doses within a certain period. Thus, we can itemize the plan in selective subplans (see Listing 7.12 for one of these subplans).

To permit only the execution of one subplan, we have to model this plan in a way that it is finished once one subplan has been completed. This is accomplished by

the `complete-condition`. If any of the plans mentioned in the complete-condition reaches the state *completed* the actual instance of the plan `PLAN_24` completes.

In the `plan-body` section we model the subplans of type *any-order*. This means that only one subplan can be executed at one time while the other subplans are waiting. Due to the `complete-condition` the waiting plans are never executed. The various subplans are listed using the `plan-activation` model. Within this model a time-annotation can be stated, for instance to declare the plan's duration.

Listing 7.11: Representation of the superior Asbru plan of the *AsbruIR* instance stated above.

```
<plan name="PLAN_24" title="Amoxicillin: 500 mg tab three times
    per day (TID) 10 days or 875 mg tab two times per day (BID
   ) 10 days">
   <delta-link link-id="24"/>
   <delta-link link-id="26"/>
   <delta-link link-id="32"/>
   <explanation text="Amoxicillin is a potential first line
      agent. Studies have shown effectiveness with 3 to 14
      days."/>
   <conditions>
      <complete-condition>
         <constraint-combination type="or">
            <plan-state-constraint state="completed">
               <plan-pointer>
                  <static-plan-pointer plan-name="
                     ATOMIC_PLAN_24_Amoxicillin_500_mg_
                     three_times_per_day"/>
               </plan-pointer>
               <time-annotation>
                  <now/>
               </time-annotation>
            </plan-state-constraint>
            <plan-state-constraint state="completed">
               <plan-pointer>
                  <static-plan-pointer plan-name="
                     ATOMIC_PLAN_24_Amoxicillin_875_mg_
                     two_times_per_day"/>
               </plan-pointer>
               <time-annotation>
                  <now/>
               </time-annotation>
            </plan-state-constraint>
         </constraint-combination>
      </complete-condition>
   </conditions>
   <plan-body>
      <subplans type="any-order">
         <wait-for>
            <one/>
         </wait-for>
         <plan-activation>
            <plan-schema name="
               ATOMIC_PLAN_24_Amoxicillin_500_mg_
```

```
                three_times_per_day">
35              <delta-link link-id="24"/>
36              <time-annotation>
37                 <time-range>
38                    <duration>
39                       <minimum>
40                          <numerical-constant value="10"
                               unit="d"/>
41                       </minimum>
42                       <maximum>
43                          <numerical-constant value="10"
                               unit="d"/>
44                       </maximum>
45                    </duration>
46                 </time-range>
47                 <self/>
48              </time-annotation>
49           </plan-schema>
50        </plan-activation>
51        <plan-activation>
52           <plan-schema name="
                ATOMIC_PLAN_24_Amoxicillin_875_mg_
                two_times_per_day">
53              <delta-link link-id="24"/>
54              <time-annotation>
55                 <time-range>
56                    <duration>
57                       <minimum>
58                          <numerical-constant value="10"
                               unit="d"/>
59                       </minimum>
60                       <maximum>
61                          <numerical-constant value="10"
                               unit="d"/>
62                       </maximum>
63                    </duration>
64                 </time-range>
65                 <self/>
66              </time-annotation>
67           </plan-schema>
68        </plan-activation>
69     </subplans>
70  </plan-body>
71 </plan>
```

The subplans for the administration of the drug in a particular dosage also have to be modeled. Listing 7.12 shows one of these subplans. It also contains a `conditions` and a `plan-body` section. In the `conditions` section we model a `setup-precondition` without any content to manually perform the state transition from *selected* to *activated*.

The `plan-body` contains a cyclical plan to model the iteration of 'three times per day'. It contains the activation of the cyclical plan and the cyclical time annotation for a frequency-based iteration. Within the latter we declare the frequency of the iterations and the time range that states the period within the cyclical plan that can

be executed.

Listing 7.12: Subplan of *PLAN_24* representing a cyclical plan.

```
<plan name="
    ATOMIC_PLAN_24_Amoxicillin_500_mg_three_times_per_day"
    title="Amoxicillin: 500 mg, three times per day">
    <delta-link link-id="24"/>
    <conditions>
       <setup-precondition confirmation-required="yes">
          <none/>
       </setup-precondition>
    </conditions>
    <plan-body>
       <cyclical-plan>
          <cyclical-plan-body>
             <plan-activation>
                <plan-schema name="
                   CYCLICAL_PLAN_Amoxicillin_500_mg">
                   <delta-link link-id="24"/>
                </plan-schema>
             </plan-activation>
          </cyclical-plan-body>
          <cyclical-time-annotation>
             <time-range>
                <starting-shift>
                   <minimum>
                      <numerical-constant value="1" unit="h"/>
                   </minimum>
                   <maximum>
                      <numerical-constant value="6" unit="h"/>
                   </maximum>
                </starting-shift>
             </time-range>
             <set-of-cyclical-time-points>
                <time-point>
                   <numerical-constant value="0"/>
                </time-point>
                <offset>
                   <numerical-constant value="0"/>
                </offset>
                <frequency>
                   <numerical-constant value="8" unit="h"/>
                </frequency>
             </set-of-cyclical-time-points>
          </cyclical-time-annotation>
       </cyclical-plan>
    </plan-body>
</plan>
```

The plan that is iteratively executed is modeled by a user-performed plan (see Listing 7.13). This plan is executed through some action by the user and is not further modeled in the system.

Listing 7.13: Asbru: Iteratively executed user-performed plan.

```
<plan name="CYCLICAL_PLAN_Amoxicillin_500_mg"
   title="Amoxicillin: 500 mg">
   <delta-link link-id="24"/>
   <plan-body>
      <user-performed/>
   </plan-body>
</plan>
```

Some plans should only be executed if a certain condition applies, for instance:

> *In case of penicillin allergy: oral cephalexin 750 mg x 2 or cefadroxil 1 g x 1 (Deeter et al., 1992; DARE-953519, 1999) [A].*

We can model this in Asbru using the `if-then-else` statement. Listing 7.14 shows the Asbru plan of the above sentence.

Listing 7.14: Asbru: Plan for an `if-then-else` statement. The plan contains a subplan including an `ask` statement, where the parameter value is queried, and an `if-then-else` statement, where the value is compared to a constant. In case of match the `then-branch` gets executed.

```
<plan name="PLAN_7" title="In case of penicillin allergy: oral
   cephalexin 750 mg x 2 or cefadroxil 1 g x 1 (Deeter et al.,
    1992; DARE-953519, 1999) [A].">
   <delta-link link-id="7"/>
   <delta-link link-id="8"/>
   <delta-link link-id="9"/>
   <explanation text="It is not necessary to start antibiotics
      immediately: a delay of 1 (-3) day(s) does not increase
      complications or delay the resolution of acute disease.
      An analgesic (Thomas, Del Mar, and Glasziou, 2000;
      DARE-20018156, 2002) [B] (paracetamol and ibuprofen are
      the safest) is more effective than antibiotics against
      symptoms."/>
   <plan-body>
      <subplans retry-aborted-subplans="no" type="any-order"
         wait-for-optional-subplans="no">
         <wait-for>
            <all/>
         </wait-for>
         <ask>
            <parameter-ref name="In_case_of_penicillin_allergy"
               />
            <time-out>
               <now/>
            </time-out>
         </ask>
         <if-then-else>
            <simple-condition>
               <comparison type="equal">
                  <left-hand-side>
                     <parameter-ref name="
                        In_case_of_penicillin_allergy"/>
```

```
                    </left-hand-side>
                    <right-hand-side>
                        <qualitative-constant value="yes"/>
                    </right-hand-side>
                </comparison>
            </simple-condition>
            <then-branch>
                <plan-activation>
                    <plan-schema name="ATOMIC_PLAN_7">
                        <delta-link link-id="7"/>
                    </plan-schema>
                </plan-activation>
            </then-branch>
        </if-then-else>
    </subplans>
</plan-body>
</plan>
```

The `plan-body` forms a subplan of type *any-order* containing an `ask` statement and an `if-then-else` statement. The former is necessary to get a value for the parameter defined in the parameter-definition section (see Listing 7.9). This parameter is then used in the condition part of the `if-then-else` statement, where the parameter's value is compared with a certain value. If the comparison is correct the `then-branch` is executed activating another plan. Unfortunately, it is not possible to model subplans in a `then-branch`. Therefore, we have to activate one plan that again activates possible subplans.

As we can see modeling Asbru is very tedious. But based on well structured and extensive data it is possible to create the necessary Asbru statements using sophisticated XSLT templates. Due to their size and complex syntax Asbru guidelines are not comprehensible for a human user, but they can be executed in a computer-supported way. An *Asbru* representation of a whole guideline is shown in Section B.5.

7.5 Benefits and Features

Currently existing methodologies for modeling computer-interpretable CPGs are not sufficient. This results from presently existing modeling processes, where all tasks rely on human user guidance and interaction. Furthermore, it is necessary to assure the modeling of various representation languages by means of one methodology.

In summary, LASSIE offers a number of benefits and features:

- Automating parts of the modeling process
- Disburdening the physicians in the modeling process due to the providing of a medical ontology
- Structuring the guideline information

- Decomposing the guideline into parts containing various kinds of information (e.g., treatment processes, diagnosis methods, definitions)
- Making the modeling process traceable and comprehensible
- Applicable for many guideline representation languages
- Supporting the guideline development process to better structuring guidelines, identifying ambiguities, inconsistencies, and incompleteness

Due to these features it is also possible to put the concept of '*living guidelines*' into action. '*Living guidelines*' are an approach to update guidelines on a more continuous basis than the usual practice of revision every two to five years. Scientific and pragmatic knowledge is growing faster every year and therefore a guideline is a static document, which cannot be modified easily. To become flexible, adaptable documents the aim is to develop guidelines, which present up-to-date and state-of-the art knowledge to practitioners. To make this possible, guidelines have to be modular in structure, so that only part of a guideline must be adjusted and not the whole document needs revision.

In order to make the modeling process traceable and each modeled construct better understandable one can use DELT/A. Figures 7.7 to 7.12 show how the particular representations are linked together. Thereby, any representation can be linked to any other representation in order to show how one sentence or action is specified in the different formats.

7.6 Shortcomings

Some problems and shortcomings of guideline modeling with LASSIE are not solved so far. Although the review of the modeling process using DELT/A is a great support, its representation and usage is unfamiliar for physicians. They have difficulties using DELT/A as most of them are not familiar with XML, the concept of macros, the insertion of XML nodes using macros, and any representation format that is not pure natural language text.

Although the developed intermediate representations have a simple structure their correction is a difficult task for many users. The insertion of more intermediate steps and thus further intermediate representations could relieve this shortcoming.

Otolaryngology guidelines are well structured and their sentences mostly describe only one action. Guidelines from other specialties may describe a sequence of actions within one sentence. In such cases a sentence-based action generation will be too imprecise. Such sentences have to be decomposed in order to identify actions and their relationships adequately.

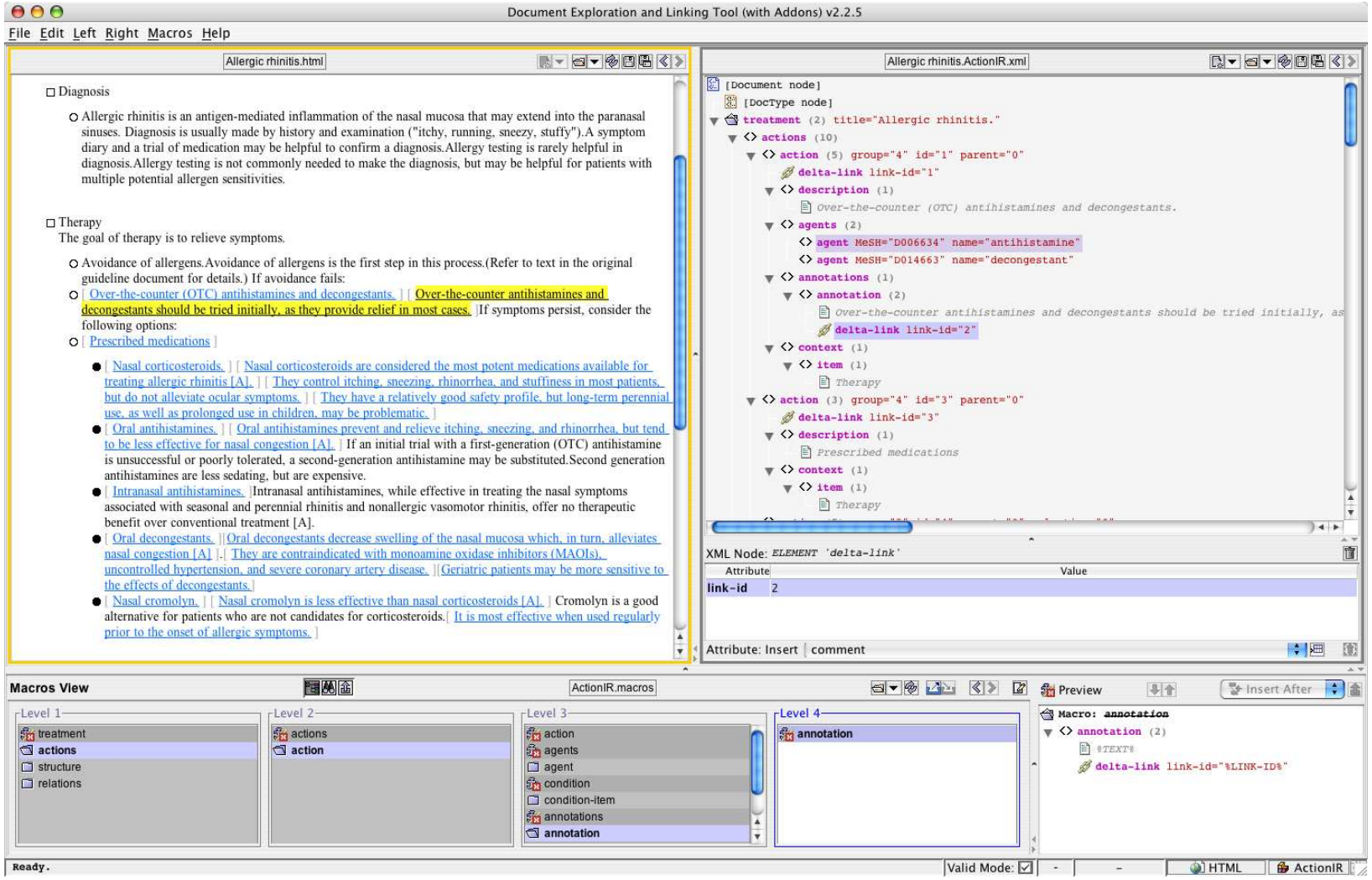


Figure 7.7: DELT/A showing the guideline document and its *ActionIR* representation.

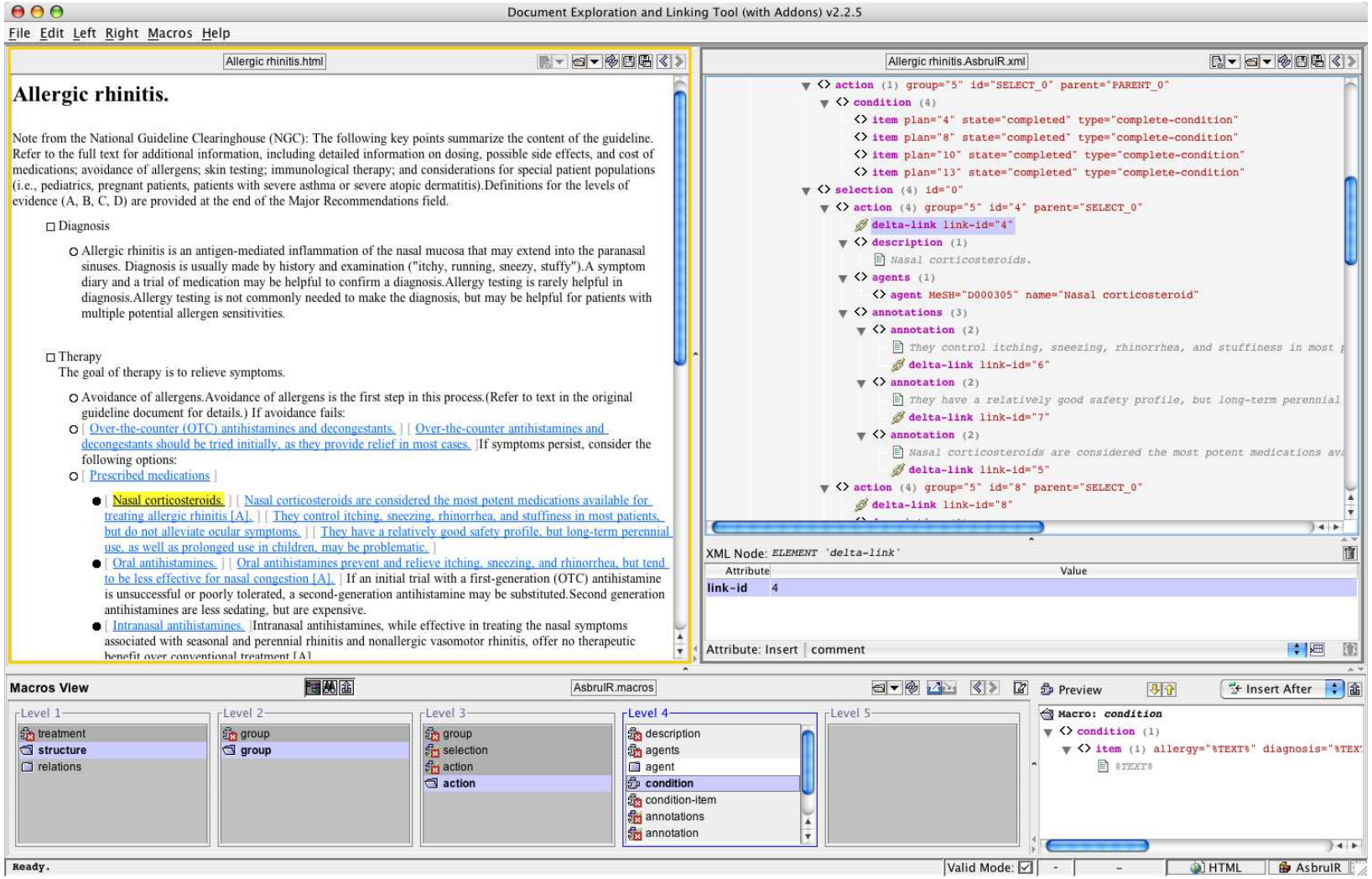


Figure 7.8: DELT/A showing the guideline document and its *AsbrulR* representation.

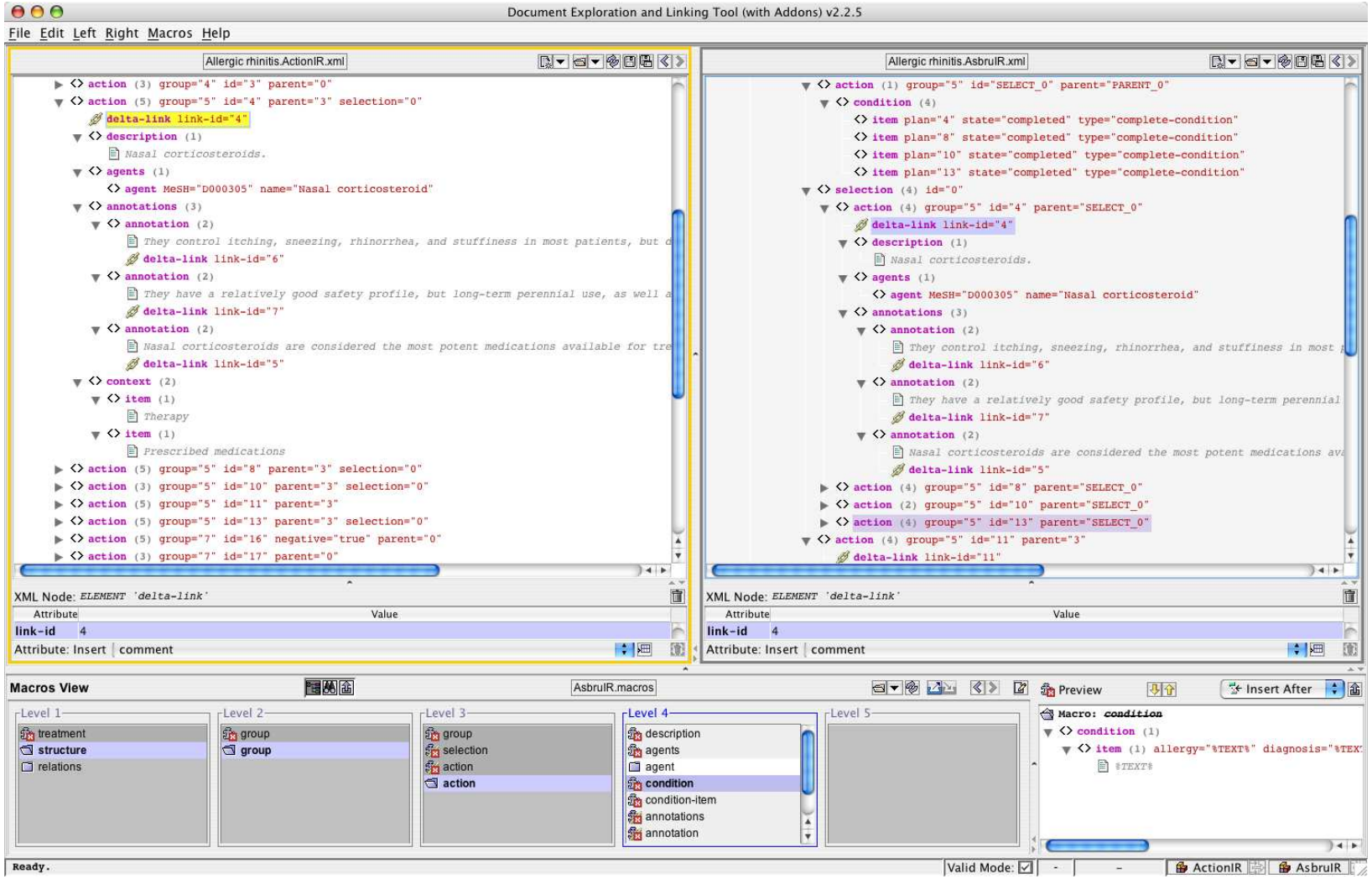


Figure 7.9: DELT/A showing the *ActionIR* representation and its corresponding *AsbruIR* representation.

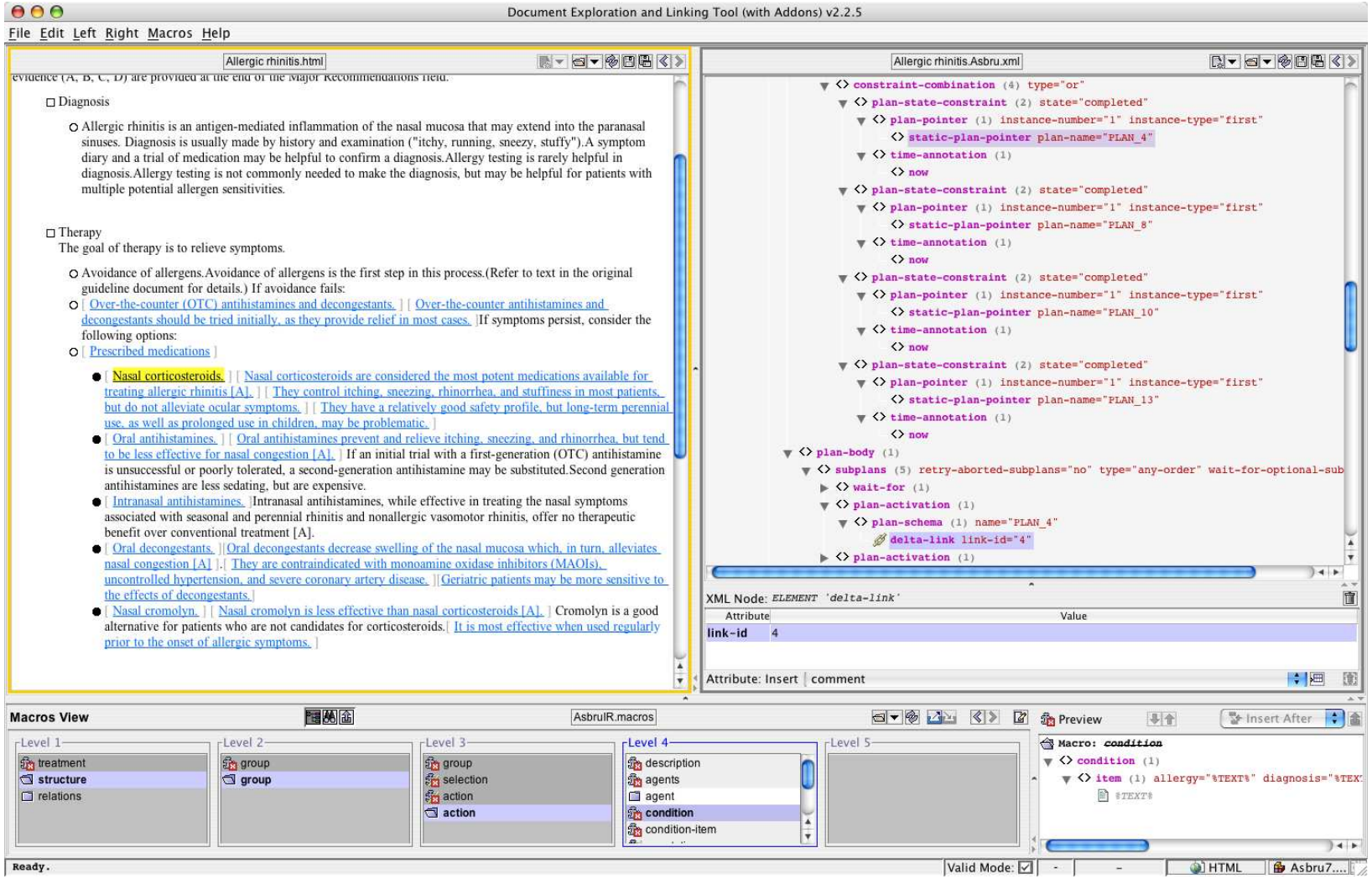


Figure 7.10: DELT/A showing the guideline document and its *Asbru* representation.

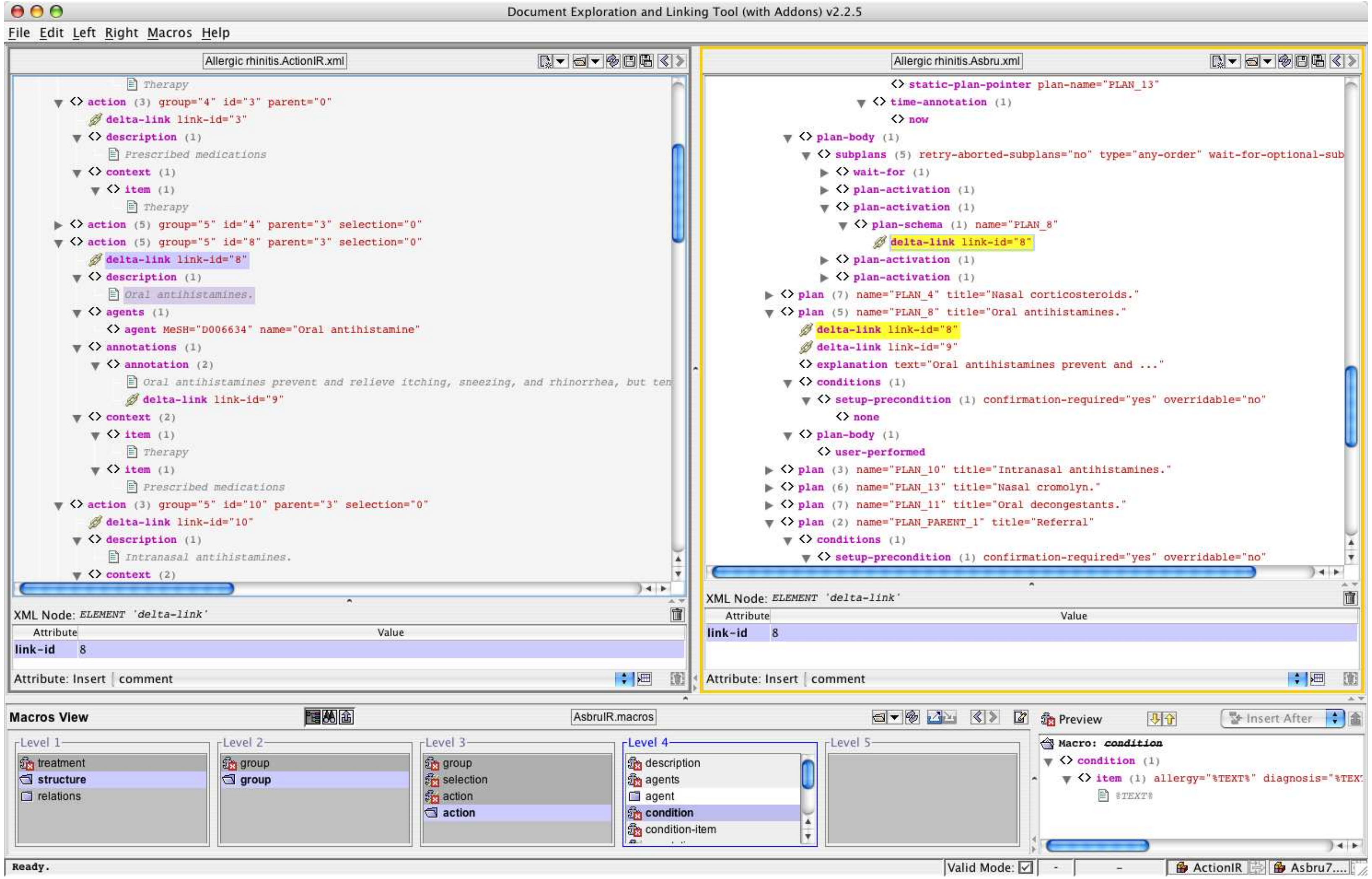


Figure 7.11: DELT/A showing the *ActionIR* representation and its corresponding *Asbru* representation.

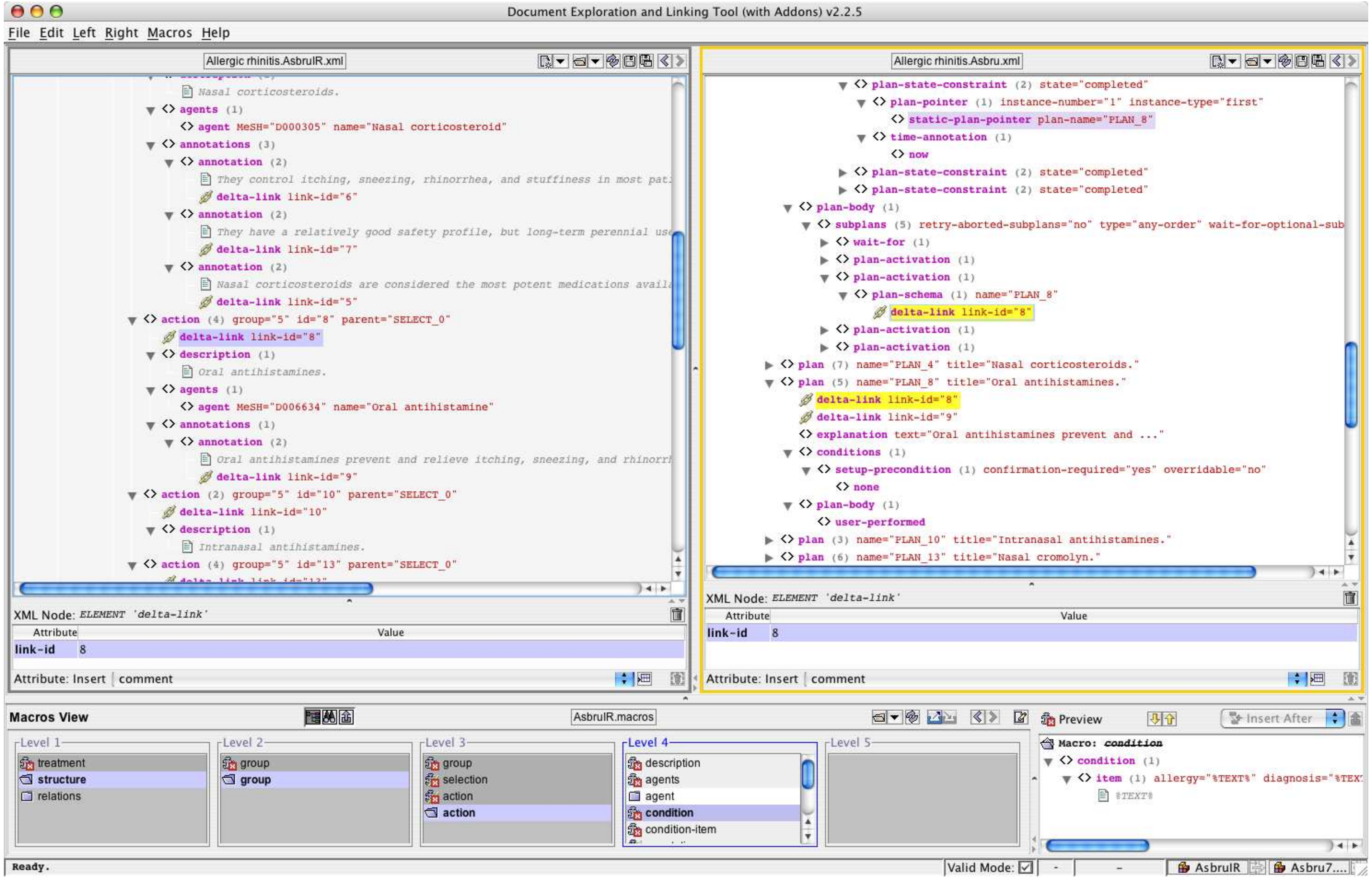


Figure 7.12: DELT/A showing the *AsbruIR* representation and its corresponding *Asbru* representation.

Chapter 8

Evaluation

In order to test the applicability of LASSIE we propose a two-step evaluation:

1. Evaluation of the IE components using statistical measures
2. Evaluation of the methodology by generating correct Asbru guidelines

As mentioned in Section 6.1 we chose six guidelines to develop extraction and transformation rules for LASSIE. Another twelve guidelines were used to accomplish the evaluation task (see Table 8.1).

By means of the evaluation we also want to constitute the benefit of the process steps for the users. We measure the benefit on the basis of the effort manually modeling or adjusting the processes using DELT/A. Therefore, we have developed special macros tailored to the intermediate representations. For the modeling we load both the HTML representation and the particular intermediate representation of the guideline. Thereby, the text of the guideline and the corresponding formal representation are linked together by DELT/A links. Figures 8.2 and 8.3 present a guideline and its formalizations in *SentenceIR* and *ActionIR*.

8.1 Evaluation of the IE Components

8.1.1 Setting

Rules in IE systems are developed using training examples. These rules have to reach a state where they are able to extract the correct information from other examples, too. In order to test these acquirements we developed Java applications that generate the intermediate representations.

The particular intermediate representations generated from the test set were evaluated by two persons using the DELT/A tool (see Section 2.2.1). The participants are computer scientists, who are familiar with guidelines, guideline formalization, and the DELT/A tool, but have no medical background. However, the chosen guidelines do not require specific medical knowledge to evaluate the IE tasks.

For the evaluation we provided the guidelines in XHTML format and the necessary language and macro files for DELT/A.

During the evaluation the participants had to mind the types of errors appearing. Thus, we have predefined error classes which are shown in Table 8.2.

Table 8.1: Set of guidelines used for the evaluation.

#	Title
1	Acute otitis media: management and surveillance in an era of pneumococcal resistance
3	Acute rhinosinusitis in adults
5	Allergic rhinitis
6	Diagnosis and management of childhood otitis media in primary care. A national clinical guideline.
8	Diagnosis and treatment of otitis media in children
9	Evidence based clinical practice guideline for children with acute bacterial sinusitis in children 1 to 18 years of age
11	Management of sore throat and indications for tonsillectomy. A national clinical guideline.
13	Otitis media with effusion
14	Pneumococcal vaccination for cochlear implant candidates and recipients: updated recommendations of the Advisory Committee on Immunization Practices
15	Reduction of the influenza burden in children
17	Sore throat and tonsillitis
18	Symptomatic treatment of radiation-induced xerostomia in head and neck cancer patients

Table 8.2: Error classes.

A1	Sentence was classified as **action** instead of annotation
A2	Sentence was classified as **annotation** instead of action
A3	Irrelevant sentence is classified as **action**
A4	Irrelevant sentence is classified as annotation
A5	Sentence is missing
A6	Annotation was assigned to a wrong action
A7	Spurious slot extraction due to **different** classification/categorization in the ontology
A8	Spurious slot extraction due to **missing** classification/categorization in the ontology
A9	Spurious slot extraction due to bad patterns
A10	Spurious slot extraction due to incorrect sentence (grammar, punctuation, etc.)

Error Classes A1 to A6. These errors appear in the *Process Extraction* task when classifying and merging sentences. Error A6 covers two cases: (1) missing assignment of an annotation to an action and (2) assigning an annotation to a wrong action.

Error Classes A7 to A10. These errors appear also in the *Process Extraction* task when extracting slot information.

Errors A7 and A8 mainly affect drug agents and other medical therapy recommendations, such as surgical procedures, and so on.

Different classification or categorization in the ontology is a problem which often appears. For many domains knowledge bases are available that represent their knowledge differently and classify it by different concepts. Especially, in such complex domains, such as the medicine, some concepts cannot be classified unambiguously and generally. One example is the concept *antibiotic*. This concept can be defined in different ways:

1. A simple translation from Greek 'anti' and 'bios' which means 'against life'. Thus, an antibiotic is an agent destroying an infection caused by bacteria, viruses, fungi, parasites, and so on.

2. Antibiotic is a synonym for anti-bacterial agent.

3. Antibiotics are a subgroup of anti-bacterial agents.

MeSH (which is the basis of the ontology) uses the second definition. Some guideline developing organizations use other definitions, and therefore it is possible that the outcome of the application is spurious due to this fact.

Error A8 appears if a term does not occur or is in a different form in the ontology. MeSH, for instance, does not differentiate among the various forms of *corticosteroids*. Both *intranasal corticosteroids* and *nasal corticosteroids* are only identified as *corticosteroids*. As such, spurious extractions emerge that affect subsequent modules.

Measurement

Over the course of several Message Understanding Conferences (MUCs) organizers and participants agreed how IE systems should be evaluated [57]. So, the extracted output is presented as hierarchical attribute-value structures. Human annotators provide a set of key templates for the training data and the test data that is compared to the system's output. Values that correctly match the key values are counted as *correct*, whereas values that do not match are *incorrect*. Attributes with non-empty values not aligning with a key attribute are considered *overgeneration*. Thus, it is possible to define recall and precision scores for the output of an IE system given the total possible correct responses (POS), number of correct values (COR), number of incorrect values (INC), and overgenerated values (OVG) as follows:

$$recall = \frac{COR}{POS}$$

$$precision = \frac{COR}{COR + INC + OVG}$$

Recall measures the ratio of correct information extracted from the texts against all the available information present in the text. Precision measures the ratio of correct information that was extracted against all the information that was extracted.

Recall and precision are the most frequently used metrics when referring to an IE system's performance.

Proceeding the Evaluation Task

The participants had to evaluate the intermediate representations *SentenceIR* and *ActionIR* generated by the system. The evaluation was accomplished in two steps:

1. Evaluation of the intermediate representation *SentenceIR* by comparing the files generated by LASSIE with the key target template generated by the participants. Hereby, we compile the following measures for each document:

   - *POS*: number of relevant sentences according to the key target template
   - *ACT*: number of relevant sentences generated by the system
   - *COR*: number of relevant sentences correctly detected by the system

   Using these values we are able to generate the recall and precision scores. The recall score measures the ratio of relevant sentences correctly detected by the system against all relevant sentences present in the text. The precision score measures the ratio of relevant sentences correctly detected by the system against all the sentences that were extracted.

2. Evaluation of the intermediate representation *ActionIR* using the key target templates of the *SentenceIR* representation. Again, the files generated by LASSIE are compared to the *ActionIR*'s key target templates generated by the participants. We compiled the following measures for each document:

   - *POS*: number of slot fillers according to the key target template
   - *ACT*: number of slot fillers generated by the system
   - *COR*: number of correct slot fillers generated by the system
   - *PAR*: number of partially correct slot fillers generated by the system

   These values were used to generate recall and precision scores. Furthermore, we also compiled the measures for the following slots: action/annotation, agent, dosage, duration, iteration, condition, and relation.

We have chosen to process a step-by-step evaluation as LASSIE performs a semi-automatic, step-wise approach, where the output will be traced, evaluated, and augmented after each process step. A one-step evaluation (i.e., the output of one task is the unaltered input of the succeeding task) is contradictory to LASSIE's concept.

8.1.2 The Mark-Up Task

The participants generated key target templates for all guidelines using the DELT/A tool. These were then compared to the templates generated by LASSIE. We compiled the number or relevant sentences according to the key target template (POS), the number or relevant sentences generated by the system (ACT), and the number of correctly detected relevant sentences generated by the system (COR). Out of these values we were able to compute the recall and precision scores (see Table 8.3 for the values and evaluation scores and Figure 8.1 for the dissemination of the outputs of the various guidelines in the recall-precision graph).

As it is hardly possible to develop a system which delivers only correct results, systems are often optimized with respect to one score. Thus, we have to consider two cases:

1. The system should detect almost all relevant sentences and probably will spuriously detect some irrelevant sentences, too.

2. The system should hardly detect irrelevant sentences as relevant and probably will ignore some relevant sentences.

Systems with an emphasis on the first case will gain a higher recall, but may concurrently derive a decreased precision. Systems with an emphasis on the second case will gain a higher precision at the expense of a decreased recall.

To optimize the benefit of this task it is more important to provide almost all relevant sentences rather than omitting irrelevant ones. For the user it is easier to remove spuriously detected sentences than reading the remaining ones, about 80 %, in the guideline to detect the lacking relevant sentences.

During the developing phase a pre-evaluation of the guidelines from the developing set produced a recall score of 78.18 % and a precision score of 95.56 %. In order to derive more relevant sentences we adjusted our extraction rules and achieved an increased recall by concurrently decreasing precision.

The resulting values of the test set for recall of **90.8 %** and precision of **94.9 %** are an enormous improvement and increase the benefit of the mark-up task.

Many users claim that IE systems performing a detection of relevant text parts are only of use if they detect all the relevant parts (i.e., a recall score of 100 %). Otherwise the user has to read the entire document to find the remaining relevant sentences. We are of opinion that even with the current performance a benefit exists. Relevant sentences are not equally spread on the whole document, but mostly constitute clusters which have to be verified then.

8.1.3 The Process Extraction Task

The key target templates of the preceding mark-up task were the input of the process extraction task. The system's output templates were again compared to this task's key target templates generated by the participants. We compiled the number of filled slots according to the key target template (POS), the number of slot fillers generated by the system (ACT), the number of correct slot fillers generated by the

Table 8.3: Evaluation results of the mark-up task.

#	Title	SEN	POS	ACT	COR	REC	PRE
1	Acute otitis media: management and surveillance in an era of pneumococcal resistance	17	5	4	4	0.8	1
3	Acute rhinosinusitis in adults	44	10	10	9	0.9	0.9
5	Allergic rhinitis	75	16	20	15	0.938	0.75
6	Diagnosis and management of childhood otitis media in primary care	57	12	9	9	0.75	1
8	Diagnosis and treatment of otitis media in children	269	49	49	49	1	1
9	Evidence based clinical practice guideline for children with acute bacterial sinusitis in children 1 to 18 years of age	171	32	32	31	0.969	0.969
11	Management of sore throat and indications for tonsillectomy	96	20	18	18	0.9	1
13	Otitis media with effusion	87	7	5	5	0.714	1
14	Pneumococcal vaccination for cochlear implant candidates and recipients	20	4	3	3	0.75	1
15	Reduction of the influenza burden in children	50	6	7	6	1	0.857
17	Sore throat and tonsillitis	120	19	17	16	0.842	0.941
18	Symptomatic treatment of radiation-induced xerostomia in head and neck cancer patients	5	5	3	3	0.6	1
		1,011	185	177	168	**0.908**	**0.949**

SEN number of sentences of the guideline document
POS number of relevant sentences according to the key target template
ACT number of relevant sentences detected by the system
COR number of correctly detected relevant sentences detected by the system
REC ratio of COR to POS
PRE ratio of COR to ACT

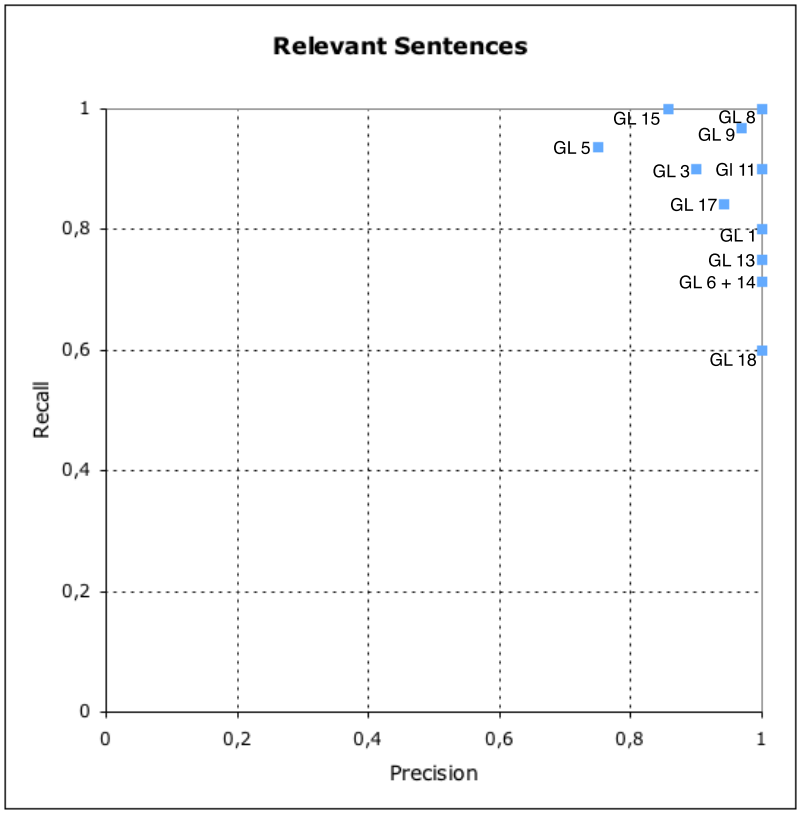


Figure 8.1: Recall & precision scores of the mark-up task for each guideline document.

system (COR), and the number of partially correct slot fillers generated by the system (PAR). Starting from these values we computed the recall (REC) and precision (PRE) scores (see Table 8.4). The overall scores of this task are **84 %** recall and **86.8 %** precision.

For adjusting the intermediate representation generated by LASSIE we use the DELT/A tool. Figure 8.3 presents a guideline and its formalization in *ActionIR*.

Apart from the evaluation of the overall task we also analyzed the following subtasks:

- Slot extraction for *agent*, *dosage*, *duration*, *iteration*, *condition*, and *relation* slots
- Sentence categorization and assignment

Slot Extraction

In this evaluation task we have analyzed the following slot extractions: *agent, dosage, duration, iteration, condition*, and *relation* slots. Therefore, we used the error classes A7 to A10 (see Table 8.5).

***Agent* Slots.** The extraction of several therapy instruments, such as drug agents or surgical procedures, is an important task to correctly execute the subsequent trans-

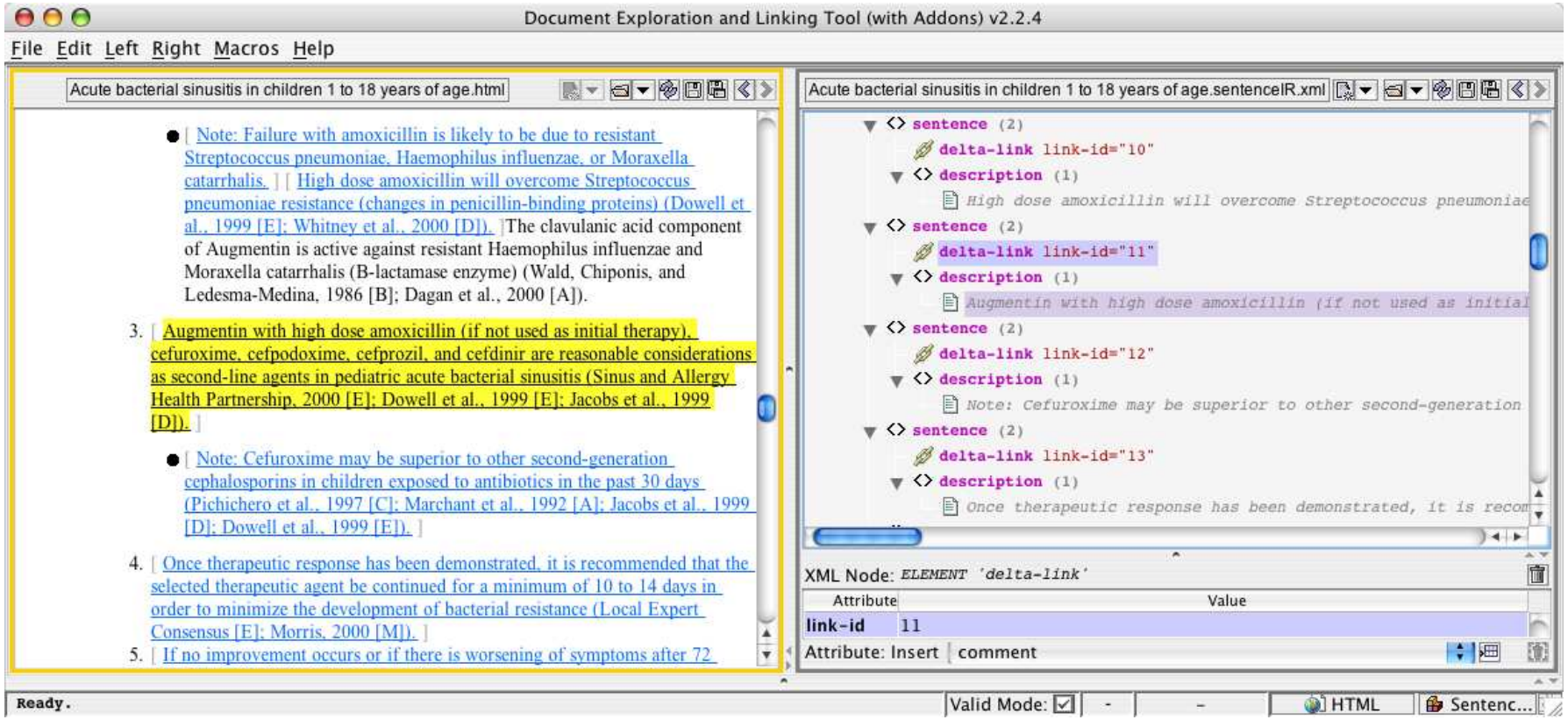


Figure 8.2: The DELT/A tool showing the marked-up guideline on the left side and the *SentenceIR* file containing a list of all extracted sentences on the right side.

Table 8.4: Evaluation results of the process extraction task.

#	Title	POS	ACT	COR	PAR	REC	PRE
1	Acute otitis media: management and surveillance in an era of pneumococcal resistance	13	15	11	0	0.846	0.733
3	Acute rhinosinusitis in adults	22	21	21	0	0.955	1
5	Allergic rhinitis	24	24	20	0	0.833	0.833
6	Diagnosis and management of childhood otitis media in primary care	27	24	19	0	0.704	0.792
8	Diagnosis and treatment of otitis media in children	107	105	97	2	0.916	0.933
9	Evidence based clinical practice guideline for children with acute bacterial sinusitis in children 1 to 18 years of age	56	60	50	1	0.902	0.842
11	Management of sore throat and indications for tonsillectomy	35	35	28	1	0.814	0.814
13	Otitis media with effusion	12	12	9	1	0.792	0.792
14	Pneumococcal vaccination for cochlear implant candidates and recipients	13	9	7	0	0.538	0.778
15	Reduction of the influenza burden in children	16	14	14	0	0.875	1
17	Sore throat and tonsillitis	38	33	27	0	0.711	0.818
18	Symptomatic treatment of radiation-induced xerostomia in head and neck cancer patients	8	7	6	0	0.75	0.857
		371	359	309	5	**0.840**	**0.868**

POS number of slot fillers according to the key target template
ACT number of slot fillers generated by the system
COR number of correct slot fillers generated by the system
PAR number of partially correct slot fillers generated by the system
REC ratio of COR plus (.5 x) PAR slot fillers to POS slot fillers
PRE ratio of COR plus (.5 x) PAR slot fillers to ACT slot fillers

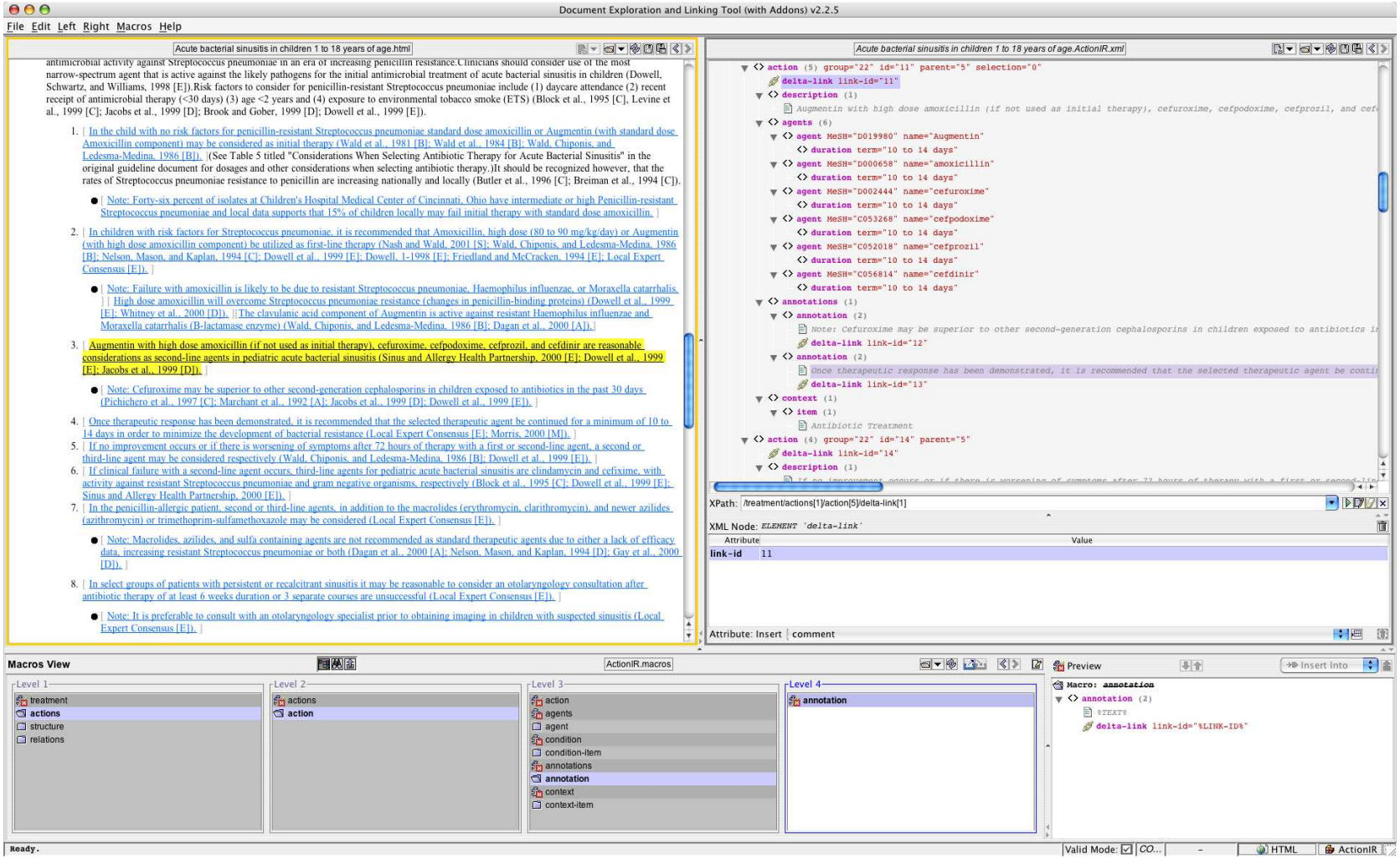


Figure 8.3: DELT/A showing the guideline document and its *ActionIR* representation.

Table 8.5: Error classes A7 to A10 for the Process Extraction task.

Guideline/ Error Class	**1**	**3**	**5**	**6**	**8**	**9**	**11**	**13**	**14**	**15**	**17**	**18**	
A7													**0**
A8											1		**1**
A9	1	1			1	1	1	1	6		6		**18**
A10													**0**

formation processes. Using our MeSH-based ontology we are able to identify and classify the instruments. Each instrument has a unique MeSH id.

Table 8.6: Evaluation results for the *agent* slot extraction.

#	POS	ACT	COR	REC	PRE
1	1	1	1	1	1
3	9	9	9	1	1
5	7	7	7	1	1
6	8	7	7	0.875	1
8	22	24	21	0.955	0.875
9	16	16	15	0.938	0.938
11	9	10	9	1	0.9
13	4	4	4	1	1
14	2	2	2	1	1
15	4	4	4	1	1
17	5	4	4	0.8	1
18	1	1	1	1	1
	88	89	84	**0.955**	**0.944**

#	id of the guideline
POS	number of slot fillers according to the key target template
ACT	number of slot fillers generated by the system
COR	number of correct slot fillers generated by the system
REC	ratio of COR slot fillers to POS slot fillers
PRE	ratio of COR slot fillers to ACTslot fillers

A spurious *agent* slot extraction can be assigned to error classes A7 and A8. One source of error, which appears only very infrequently, is the lack of an extensive coreference module (see Section 4.4). Figure 8.4 shows an example, where the agent is not stated due to missing coreferences.

As a result of the comprehension of the MeSH ontology we are able to extract almost all therapy instruments from the text. From the 88 possible therapy instruments we only missed four slots from guidelines 6, 8, 9, and 17. Thus, the recall score for this task is **95.5 %**. But we also over-generated some agent slots. In guidelines 8, 9, and 11 we extracted five slots that caused a precision score of **94.4 %**.

If sentences are categorized as annotations the instruments are not appearing in the template. In case of a wrong classification the instruments have to be added

1. First-Line Medications
   - amoxicillin (40 mg/kg/day) if low risk (> 2 years, no day care, and no antibiotics for the past three months).
   - 80 mg/kg/day if no low risk or for resistant AOM if the lower dose was used initially.

Figure 8.4: Part of a guideline treating otitis media in children. The second item does not name the drug agent and therefore the agent and the dosage are not extracted.

manually. This applies also for dosage, duration, and iteration information that are presented as child elements of the agent slot.

***Dosage* Slots.** The cause of this error is either a spurious annotation assignment or bad extraction patterns (error class A9). Table 8.7 shows the evaluation results for this subtask. In guidelines 8 and 17 we missed one dosage slot each time which caused a recall score of **84.6 %**. At the same time we over-generated one dosage slot each time in guidelines 1 and 9. The resulting precision score therefore is at **84.6 %**.

In case of a manual administration of a drug a missing dosage indication does not have a collateral effect on the execution. If the execution unit controls a machine administering a drug (e.g., artificial respiration) the dosage has to be specified. The insertion of the dosage term is handled with marginal effort using the appropriate DELT/A macro.

Table 8.7: Evaluation results for the *dosage* slot extraction.

#	POS	ACT	COR	REC	PRE
1	2	3	2	1	0.667
3	0	0	0	1	1
5	0	0	0	1	1
6	0	0	0	1	1
8	3	2	2	0.677	1
9	1	2	1	1	0.5
11	1	1	1	1	1
13	0	0	0	1	1
14	0	0	0	1	1
15	0	0	0	1	1
17	5	4	4	0.8	1
18	1	1	1	1	1
	13	13	11	**0.846**	**0.846**

***Duration* Slots.** The duration of a therapy is an important statement for both the patients and the medical personnel. In case of a missing or spurious *duration* information it can be applied with minimal effort using the DELT/A tool. Table 8.8 shows the evaluation results for this subtask.

Table 8.8: Evaluation results for the *duration* slot extraction.

#	POS	ACT	COR	PAR	REC	PRE
1	0	0	0	0	1	1
3	1	0	0	0	0	1
5	0	0	0	0	1	1
6	1	0	0	0	0	1
8	5	5	4	1	0.9	0.9
9	1	1	0	0	0	0
11	2	1	1	0	0.5	1
13	0	0	0	0	1	1
14	0	0	0	0	1	1
15	0	0	0	0	1	1
17	2	1	1	0	0.5	1
18	0	0	0	0	1	1
	12	8	6	1	**0.542**	**0.813**

#	id of the guideline
POS	number of slot fillers according to the key target template
ACT	number of slot fillers generated by the system
COR	number of correct slot fillers generated by the system
PAR	number of partially correct slot fillers generated by the system
REC	ratio of COR plus (.5 x) PAR slot fillers to POS slot fillers
PRE	ratio of COR plus (.5 x) PAR slot fillers to ACT slot fillers

Although the guidelines contained twelve duration slots, we only extracted six of them correctly and one partially correctly. The six missing or partially correct slots are spread to six guidelines (i.e., 3, 6, 8, 9, 11, and 17) and thus the recall score for this task is at **54.2 %**. As we only had one over-generated and one partially correct slot in guidelines 8 and 9 the precision is at **81.3 %**.

Anyhow, the recall score is very low and we have to improve our patterns for this slot in order to gain better results.

***Iteration* Slots.** The causes for errors in this task are similar to the ones in the *duration* slot extraction. A missing iteration term can be inserted with minimal effort. The insertion into a subsequent intermediate representation is already more extensive. Modeling cyclical plans in Asbru is cumbersome and demands detailed knowledge about Asbru's syntax (see Section 5.4.2 for details). Thus, correctly extracting iteration information is a crucial task within the modeling process. Table 8.9 shows the evaluation results for this subtask.

Unfortunately, we missed three out of five iterations which are all contained in guideline 17. Thus, the recall score is only **40 %**. The missing extractions are caused by missing extraction patterns for these iteration terms. But we also over-generated one slot of guideline 8 whereby the precision score was set to **66.7 %**.

Table 8.9: Evaluation results for the *iteration* slot extraction.

#	POS	ACT	COR	REC	PRE
1	0	0	0	1	1
3	0	0	0	1	1
5	0	0	0	1	1
6	0	0	0	1	1
8	0	1	0	1	0
9	0	0	0	1	1
11	1	1	1	1	1
13	0	0	0	1	1
14	0	0	0	1	1
15	0	0	0	1	1
17	3	0	0	0	1
18	1	1	1	1	1
	5	3	2	**0.4**	**0.667**

#	id of the guideline
POS	number of slot fillers according to the key target template
ACT	number of slot fillers generated by the system
COR	number of correct slot fillers generated by the system
REC	ratio of COR slot fillers to POS slot fillers
PRE	ratio of COR slot fillers to ACT slot fillers

***Condition* Slots.** The extraction of condition terms is an important factor in modeling guidelines. Table 8.10 presents the evaluation results of this subtask. We only extracted 23 correct and 4 partially correct conditions out of 34 conditions which causes a recall score of **73.5 %**. The missing conditions were spread to three guidelines (i.e., guidelines 1, 14, and 15). Unfortunately, we over-generated some condition slots. Guidelines 8, 14, and 17 contain this information that caused a precision score of **73.5 %**.

The causes of errors are mainly based on insufficient patterns. Spurious slots are easy to discover, as only *action* sentences have to be checked and the correction of condition slots is easy to maintain.

***Relation* Slots.** Within this subtask we evaluate both the *relation* concepts and the *selection* concept. The former are more difficult to obtain and thus the errors mainly appear when detecting interval relations. But an adjustment of spurious interval relations is easier to accomplish than an adding of missing selection relations. Table 8.11 presents the results of the relations' evaluation.

We correctly extracted 34 out of 39 relations. Guidelines 6, 14, and 17 contain information regarding relations that our patterns could not detect. Thus, the obtained recall score is **87.2 %**. At the same time the system over-generated five slots in guidelines 1, 6, and 9 which were all *relation* concepts. This caused a precision score of **87.2 %**.

Table 8.10: Evaluation results for the *condition* slots.

#	POS	ACT	COR	PAR	REC	PRE
1	2	1	1	0	0.5	1
3	2	2	2	0	1	1
5	0	0	0	0	1	1
6	1	1	1	0	1	1
8	6	9	5	1	0.917	0.611
9	8	8	7	1	0.938	0.938
11	2	2	1	1	0.75	0.75
13	1	1	0	1	0.5	0.5
14	4	2	0	0	0	0
15	4	2	2	0	0.5	1
17	4	6	4	0	1	0.667
18	0	0	0	0	1	1
	34	34	23	4	**0.735**	**0.735**

#	id of the guideline
POS	number of slot fillers according to the key target template
ACT	number of slot fillers generated by the system
COR	number of correct slot fillers generated by the system
PAR	number of partially correct slot fillers generated by the system
REC	ratio of COR plus (.5 x) PAR slot fillers to POS slot fillers
PRE	ratio of COR plus (.5 x) PAR slot fillers to ACT slot fillers

Table 8.11: Evaluation results for the *relation* slots.

#	POS	ACT	COR	REC	PRE
1	2	4	2	1	0.5
3	2	2	2	1	1
5	1	1	1	1	1
6	4	3	2	0.5	0.667
8	14	14	14	1	1
9	7	9	7	1	0.778
11	0	0	0	1	1
13	2	2	2	1	1
14	3	2	2	0.667	1
15	2	2	2	1	1
17	2	0	0	0	1
18	0	0	0	1	1
	39	39	34	**0.872**	**0.872**

Sentence Categorization and Assignment

We evaluate the correctness of both the categorization of sentences to actions and annotations and the assignment of annotations to actions. Table 8.12 shows the evaluation results for this subtask and Figure 8.5 shows the recall and precision graph.

Table 8.12: Evaluation results for the *sentence* category.

#	POS	ACT	COR	REC	PRE
1	6	6	5	0.833	0.833
3	8	8	8	1	1
5	16	16	12	0.75	0.75
6	13	13	9	0.692	0.692
8	49	49	46	0.939	0.939
9	23	24	20	0.870	0.833
11	20	20	15	0.75	0.75
13	5	5	3	0.6	0.6
14	4	3	3	0.75	1
15	6	6	6	1	1
17	17	18	14	0.824	0.778
18	5	4	3	0.6	0.75
	172	172	144	**0.837**	**0.837**

#	id of the guideline
POS	number of slot fillers according to the key target template
ACT	number of slot fillers generated by the system
COR	number of correct slot fillers generated by the system
REC	ratio of COR to POS slot fillers
PRE	ratio of COR to ACT slot fillers

To understand how these outcomes emerged we have analyzed the results by means of the error classes. Table 8.13 shows the errors categorized in classes A1 to A6 for each guideline.

Table 8.13: Error classes A1 to A6 for the Process Extraction task.

Guideline/ Error Class	**1**	**3**	**5**	**6**	**8**	**9**	**11**	**13**	**14**	**15**	**17**	**18**	
A1	1		3	2	1	2	1	2			2	1	**15**
A2			1	2	1		4				1		**9**
A3						1					1		**2**
A4													**0**
A5									1			1	**2**
A6				1	5		6						**12**

The 185 sentences we obtained from the *mark-up* task had to be categorized in actions, annotations, and irrelevant sentences (i.e., annotations that had no action to

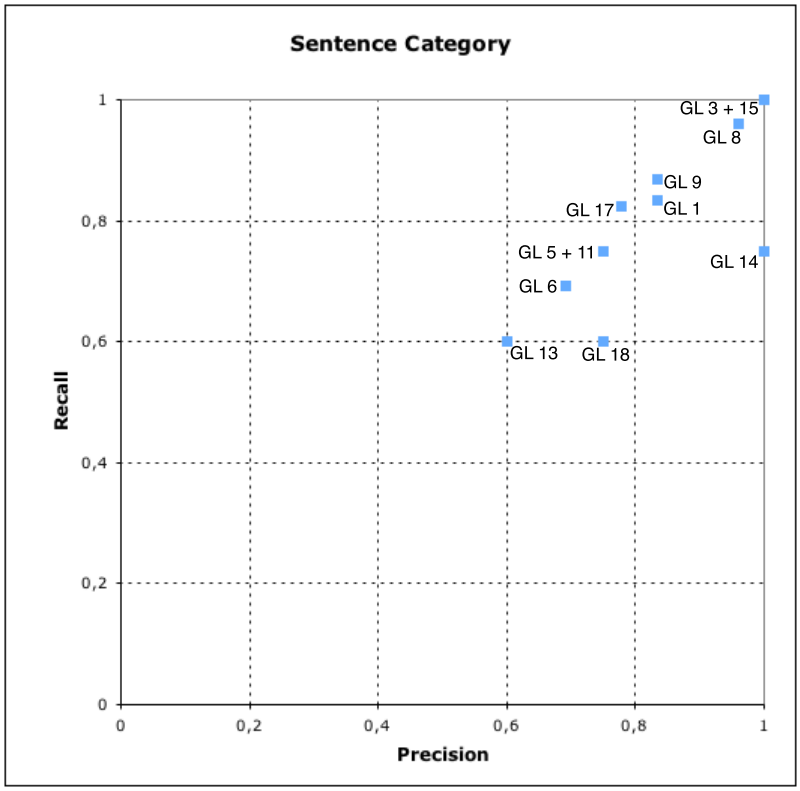


Figure 8.5: Recall & precision scores for the sentence categorization and assignment.

be assigned to). Additionally, the annotations had to be assigned to actions. Less the irrelevant sentences we had 172 slots whereas 145 were classified and assigned correctly. Thus, the resulting recall score is **84.3 %**. But we also had over-generated slots, for instance in guidelines 9 and 17, and lost sentences in guidelines 14 and 18. These caused a precision score of **84.3 %**.

It shows that spurious categorizations are the most frequently appearing errors. The wrong categorization to actions thereby appears more frequently than the wrong categorization to annotations. Although, the latter has far more consequences for the user adaption: the user has to create the action and has to manually add the necessary slot information for agents, and so on. Manually creating annotations is associated with less effort: the user has to assign the annotation to the particular actions which is rather easy using DELT/A.

Another frequently appearing error is the spurious assignment of an annotation sentence to an action which is expressed by error class A6. The effect of such an error on the effort of the adaptation is varying. An assigning to an action or the removal of an annotation asks only for a minimum of effort. But if the annotation contains temporal information (e.g., about the duration or the repetition of a therapy instruction) it must be added manually.

The causes for the spurious classifications are manifold. It can be due to insufficient rules for the annotation class, and it can be due to the missing pronoun-antecedent coreferencing (see Sections 4.4 and 7.3.4).

With correct classification the benefit for the user increases by concurrently de-

creasing the effort of modeling.

8.2 Evaluation of the Methodology

The second part of the evaluation deals with the applicability of the methodology by transforming the information presented in *ActionIR* into the formal guideline representation language *Asbru*.

The transformation process is thereby split into two parts: (1) obtaining an Asbru-specific representation *AsbruIR* and (2) transforming the latter into Asbru.

8.2.1 Obtaining an Asbru-specific Representation

Obtaining the *AsbruIR* representation is a rather trivial task. The structure and action parts are merged, temporal information such as duration and iteration, is itemized, and interval relations are modeled by temporal conditions. Thus, we will now evaluate the itemization of the temporal information and the modeling of temporal conditions.

Itemization of Temporal Information

We checked both the twelve duration slots and the five iteration slots. The itemization is based on rather simple rules and thus LASSIE transformed the information correctly.

Also, for the user the benefit of this task is given. The itemization is trivial, but the user has to assign the particular items to their corresponding elements which can be an annoying task. Thus, the automating of this task supports the user and enhances the benefit.

Modeling Temporal Conditions

Our test set of twelve guidelines contains 13 interval relations which were all transformed correctly. All of them describe a sequential ordering of the actions and thus can be modeled by a condition for the starting point of the subsequent action. Again, the transformation is very trivial, but the user profits from the automatization as she/he does not have to know how to model those relations.

8.2.2 Transformation into *Asbru*

We accomplished the evaluation in two steps:

1. Evaluation of whether the transformation of the particular components has been correctly accomplished. We will evaluate this only by random examination.

2. Evaluation of the correctness of the Asbru guideline in respect of

   - Whether the XML document is valid against the Asbru DTD

- Whether the hierarchical structure of the Asbru guideline is correct. This can be accomplished using AsbruView (see Section 2.2.6 for details).

Transformation of Various Concepts

For evaluating the transformation of the concepts we analyzed parameter definitions, conditions, duration, iteration, selection plans, and subplans for multiple therapy instructions within one sentence. They were accomplished almost errorless. The only errors we received were duplicate parameter names. This emerges when a condition term appears more than once in a guideline. Sometimes the text used to query the executing user has to be adjusted slightly.

The benefit of these transformations is enormous. Some modeling concepts are very small, but still without knowledge about Asbru the manual generation is cumbersome. In case of larger and more complex concepts its hard up to impossible to keep track of the modeling. One action in the *AsbruIR* representation often has to be modeled by multiple *Asbru* plans.

Correctness of *Asbru* Guidelines

First we checked the validity of the *Asbru* XML document. Thereby, we can verify whether the XML document complies the specifications defined in *Asbru*'s DTD.

When loading the XML document into AsbruView it is checked whether a root plan exists. A root plan is a plan which has no parent plan. A correct *Asbru* document may contain only one root plan with which the execution starts.

Using the XML representation of *Asbru* guidelines a hierarchy or structure of the plans is not or at least difficult to obtain. Using AsbruView the plan structure can be visualized (see Figure 8.6). The hierarchy can then be cross checked with the original guideline, the *Asbru* representation, or any of the guideline's representations.

Every tested guideline was both valid against the language specification and correct in terms of the structure of the plans and the number of plans.

8.3 Discussion

Analyzing the results of the evaluation we can see that the crucial factor for the methodology's performance is the process extraction task (i.e., the output reflected in the *ActionIR* format). Based on the correct representation in *ActionIR* an almost errorless transformation into *Asbru* is possible.

Thereby, the emphasis has to be the detection of relevant text segments and the merging of particular sentences to actions.

But the evaluation's results have also to be seen in terms of the following aspects:

Participants of the evaluation. Depending on the participant the results of the evaluation will vary. These variations might be different according to the medi-

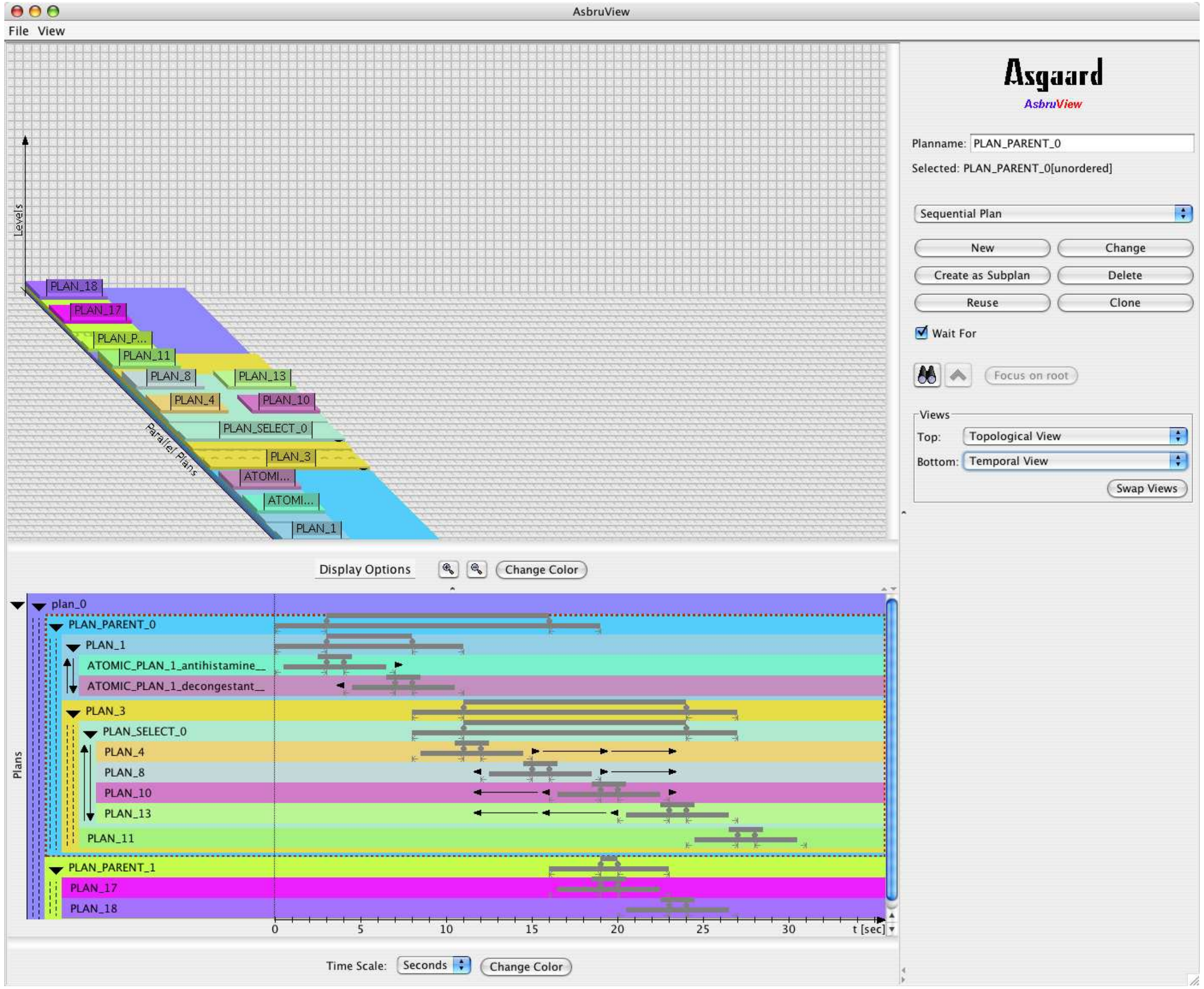


Figure 8.6: Representing the hierarchy of an *Asbru* guideline using AsbruView.

cal knowledge of the participants and the complexity of the medical topic of the guidelines. According to these parameters we only face minor deviations.

Underlying medical ontology. Treatment instructions in guidelines of otolaryngology mostly consist of the administration of drugs, surgical procedures, or remedies. Other specialties, such as oncology, consist of more complex treatment processes (e.g., *Infuse etoposide over 30 - 60 min to avoid hypotension; monitor patient's blood pressure during infusion. Prior to use, dilute the drug to a final concentration of 0.2-0.4 mg/ml to prevent precipitation. Monitor for crystallization during infusion.*). Thereby, actions have to be executed in the context of the therapy which do not deal with therapy in a traditional sense (e.g., *Monitor blood pressure during infusion.*). In order to detect these actions and to process them adequately we have to augment the ontology in terms of these new actions.

Clinical specialty and selected guidelines. For developing and testing our approach we used guidelines of the medical subject of otolaryngology. These guidelines are mostly well structured and clearly formulate the treatment instructions. The evaluation results also correlate with the portion of semi-structured data. Less structured documents with a major portion of only free text perform worse than well-structured documents.

Chapter 9

Summary and Future Work

First results of applying Information Extraction and a step-wise procedure for supporting the modeling of treatment processes of clinical practice guidelines seem promising. They support the requirements preliminarily defined in Chapter 1 and offer the needed functionality for a traceable modeling process for the users.

Findings of our evaluation discussed in Chapter 8 indicate that using semi-automatic, step-wise Information Extraction methods are a valuable instrument to formalize clinical guidelines and protocols.

We have developed several heuristics, which we integrated in a framework and applied them to several guidelines of the specialty of otolaryngology. Thereby, we firstly generate a simple representation of treatment instructions (i.e., actions), which are independent from the final guideline representation language. Based from this independent representation we can secondly transform the information in further steps into the guideline languages. To proof our methodology we applied the framework to formalize guidelines in the formal *Asbru* plan representation.

For automatically modeling guidelines in Asbru we have analyzed the Asbru representation in respect of modeling and synchronizing of processes. As Asbru can represent both incomplete and uncertain temporal information the definition of rules or modeling templates is a difficult task. However, we have analyzed the various possibilities to model temporal dependencies between actions. These contained representations for Allen's *algebra of intervals* [5] and its possibilities to deal with incomplete and uncertain information as well Freksa's concept of neighborhood [31] and his generalization of relations with *semi-intervals*. Additionally, we have defined modeling templates for other temporal concepts appearing in guidelines: the *selection* and the *iteration*.

However, some open issues remain: in the near future we want to continue the development of our IE methods in order to apply them to more complex guidelines, such as from the medical subject of oncology. These guidelines have less portions of semi-structured layout and consist mostly of free text. They are more complex and processes have to be extracted by more sophisticated heuristics. To apply the methodology to various guideline representation languages, we will analyze and categorize them according to their syntax, structure, and complexity. The development of the intermediate representations is geared to this categorization and provides representations that will comply as far as possible. In particular we will

implement a transformation process to the *PROforma* representation (see Section 2.1.6 for details).

Some problems and shortcomings of guideline modeling with LASSIE are not solved so far. A formal representation is unfamiliar for physicians. Especially, the representations in the XML format means a hurdle for many of them. To verify each modeling step we have to provide other representation methods which use XML only for internal representation. Although the developed intermediate representations have a simple structure their correction and adjustment is a difficult task for many users. The insertion of more intermediate steps and thus further intermediate representations could relieve this shortcoming. Anyhow, this has to be thoroughly reflected in order to have the modeling process not appear too complex by a high number of intermediate representations.

Otolaryngology guidelines are well structured and their sentences mostly describe only one action. Guidelines from other specialties may describe a sequence of actions within one sentence. In such cases a sentence-based action generation will be too imprecise. Such sentences have to be decomposed in order to better identify relationships between actions within one sentence.

Chapter 10

Conclusions

Modeling clinical guidelines and protocols is a complex task which has to be assisted by both physicians and knowledge engineers. Bearing those two user groups in mind a method is demanded supporting them in their particular fields of functions: the physicians have to be less overcharged by the formal specifications and the knowledge engineers have to be fostered by providing medical knowledge. Apart from this interesting conceptual formulation we have developed extensive ideas and visions. In the end we have implemented one of the most interesting and auspicious ideas: the modeling of treatment processes using the LASSIE methodology.

Based on the findings of our evaluations we can say that LASSIE might offer distinct benefits. In particular, LASSIE

1. Automates parts of the modeling process
2. Disburdens the physicians in the modeling process due to the providing of a medical ontology
3. Structures the guideline information
4. Decomposes the guideline into parts containing various kinds of information (e.g., treatment processes, diagnosis methods, definitions)
5. Makes the modeling process traceable and comprehensible
6. Is applicable for many guideline representation languages
7. Supports the guideline development process in order to better structure guidelines, identifies ambiguities, inconsistencies, and incompleteness

According to our prerequisites for supporting guideline modelers and facilitating their tasks (compare Section 1.1) we introduce the LASSIE methodology. Furthermore, the methodology may also influence the application of the concept of '*living guidelines*', an approach to have guidelines become flexible, adaptable documents which present up-to-date and state-of-the-art knowledge to practitioners.

Chapter 11

Acknowledgements

First of all I would like to thank my supervisor **Silvia Miksch**, who gave me the opportunity to do my Ph.D. and for her support, motivation, and encouragement throughout my total work.

Furthermore, I would like to thank my colleagues **Wolfgang Aigner, Klaus Hinum, Alime Öztürk**, Andreas Seyfang, Peter Votruba, and Burcu Yildiz at the Information Engineering group for their support. Many thanks go to **Andreas Seyfang** for his patience in explaining aspects of the *Asbru* language to me, to **Peter Votruba** for eliminating bugs in DELT/A and for accomplishing parts of the evaluation task, and to **Burcu Yildiz** for doing a great job at the evaluation.

The student **Cem Akkaya** gave valuable inputs when coding a first prototype of the Information Extraction package, according to our specifications.

Special thanks go to my family **Lena, Elsa** and **Michael** for their patience and support and for being sympathetic when I was down or had very little time to spend with them. They made it possible for me to reach these days.

This work has been carried out within the Asgaard Project, which is supported by *Fonds zur Förderung der wissenschaftlichen Forschung* (Austrian Science Fund), grant P15467-INF.

Appendix A

Clinical Practice Guidelines

Evidence based clinical practice guideline for children with acute bacterial sinusitis in children 1 to 18 years of age.

Cincinnati Children's Hospital Medical Center; year of publication: 2001

The levels of evidence (A-F, S, M, Q, L, O, and X) are defined at the end of the "Major Recommendations" field.

Clinical Assessment

1. No single symptom or sign is specific for the diagnosis of acute bacterial sinusitis. The diagnosis is best made clinically in the presence of a constellation of signs and symptoms of **at least 10 days duration without improvement** (Wald et al., 1981 [B]; Wald et al., 1984 [B]; Wald, Guerra, and Byers, 1991 [C]; Aitken and Taylor, 1998 [C]; Garbutt, Gellman, and Littenberg, 1999 [C]). The 10-day duration is suggested because it has been shown that in most children with uncomplicated upper respiratory infection improvement is seen on average, by 10 days (Wald et al., 1981 [B]; Wald et al., 1984 [B]; Wald, Guerra, and Byers, 1991 [C]). On the other hand acute severe bacterial sinusitis represents a more toxic form and may have a shorter duration of symptoms (Wald, 1994 [S]; Fireman, 1992 [S]; Giebink, 1994 [S]).

2. The quantity, quality, and color of nasal discharge are not helpful in differentiating acute bacterial sinusitis from other upper respiratory illnesses (e.g., common cold, allergic rhinitis). It is therefore recommended that the character of the nasal discharge **not** be used to make a diagnosis or as an indication for antibiotic treatment (McLean, 1970 [D]; Aitken and Taylor, 1998 [C]; Gungor and Corey, 1997 [S]; Wald et al., 1981 [B]; Wald, Guerra, and Byers, 1991 [C]; Wald, 1994 [S]).

    - **Note 1**: Physical exam is likely to reveal purulent nasal discharge and/or

posterior oropharyngeal drainage. These findings however, are non-specific and of little diagnostic usefulness (Ott et al., 1991 [S]; Fireman, 1992 [S]; Wald et al., 1981 [B]; McLean, 1970 [D]; Kogutt and Swischuk, 1973 [D]; Williams and Simel, 1993 [S]).

- **Note 2**: Transillumination may be useful above the age of 10 in the hands of an experienced clinician (Wald, 1994 [S]), but may be of limited value in the younger, uncooperative child (Brook et al., 2000 [E]; Ott et al., 1991 [S]; Wald, Chiponis, and Ledesma-Medina, 1986 [B]; McLean, 1970 [D]; Otten and Grote, 1989 [C]).

Acute bacterial sinusitis is characterized by:

- Persistence of upper respiratory symptoms for greater than 10 days without improvement (Wald et al., 1981 [B]; Wald et al., 1984 [B]; Wald, Chiponis, and Ledesma-Medina, 1986 [B]; Isaacson, 1996 [S]; Brook et al., 2000 [E]; Wald, 1994 [S]).
- Nasal congestion and nasal discharge of any quality (i.e., may be either thin and milky or thick and purulent) (McLean, 1970 [D]; Aitken and Taylor, 1998 [C]; Dowell, Schwartz, and Phillips, 1998 [E]; Gungor and Corey, 1997 [S]; Wald et al., 1981 [B]; Wald, 1994 [S]; Ott et al, 1991 [S]; Fireman, 1992 [S]; Kogutt and Swischuk, 1973 [D]; Williams and Simel, 1993 [S]).
- Persistent cough which is often more severe at night (Wald et al., 1984 [B]).

  Note 1: Less-common complaints may include low-grade fever (McLean, 1970 [D]; Aitken and Taylor, 1998 [C]), sore throat or ear discomfort (Brook et al., 2000 [E]; Zacharisen and Kelly, 1998 [S]; Lusk and Stankiewicz, 1997 [S]), fatigue, malodorous breath, intermittent periorbital edema or facial swelling, and facial or tooth pain (Zacharisen and Kelly, 1998 [S]; Wald, 1994 [S]; Isaacson, 1996 [S]).

Acute severe bacterial sinusitis may present with the same symptoms described above but is differentiated from acute bacterial sinusitis by the following:

- Toxic appearing child (Wald, 1994 [S]; Fireman, 1992 [S]; Giebink, 1994 [S]).
- Fever higher than 102.2 degrees Fahrenheit (Wald, 1994 [S]; Fireman, 1992 [S]; Giebink, 1994 [S]).
- Duration of symptom observation may be $<$10 days (Wald, 1994 [S]; Fireman, 1992 [S]; Giebink, 1994 [S]).

Other conditions that present with symptoms similar to acute bacterial sinusitis include recurrent viral upper respiratory infection, allergic rhinitis,

cough-variant asthma, enlarged adenoids, deviated nasal septum, choanal atresia, nasal foreign body, neoplasm (Zacharisen and Kelly, 1998 [S]) gastroesophageal reflux disease, and nasal polyp (Isaacson, 1996 [S]; Brook et al., 2000 [E]; Gungor and Corey, 1997 [S]) and should be considered in the differential diagnosis.

Radiologic Assessment

1. Routine radiologic studies **are not indicated** in the initial management of a patient suspected of uncomplicated acute bacterial sinusitis. It is recommended that patients with suspected acute bacterial sinusitis be treated based on the clinical impression without the addition of imaging studies (Schwartz, Pitkaranta, and Winther, 2001 [C]; Diament, 1992 [E]; American College of Radiology Appropriateness Criteria [ACR], 2000 [E]; Local Expert Consensus [E]). Paranasal sinus imaging abnormalities are non-specific, often present without sinusitis, and may last longer than clinical symptoms when sinusitis is present (Maresh, 1940 [D]; Shopfner and Rossi, 1973 [C]; Odita et al., 1986 [C]; Glasier, Ascher, and Williams, 1986 [C]; Glasier, Mallory, and Steele, 1989 [C]; Diament et al., 1987 [C]; Rak et al., 1991 [C]; Gwaltney et al., 1994 [D]).

    - **Note 1**: Abnormalities of the paranasal sinuses are found frequently on conventional radiographs and computed tomography scans in children **without** clinical evidence of sinusitis (see table below) (Maresh, 1940 [D]; Shopfner and Rossi, 1973 [C]; Odita et al., 1986 [C]; Glasier, Ascher, and Williams, 1986 [C]; Glasier, Mallory, and Steele, 1989 [C]; Diament et al., 1987 [C]; Rak et al., 1991 [C]). The presence of an upper respiratory infection alone (without sinusitis), can result in mucosal thickening and abnormal findings in the paranasal sinuses on plain radiographs and computed tomography scans (Shopfner and Rossi, 1973 [C]; Glasier, Ascher, and Williams, 1986 [C]; Glasier, Mallory, Steele, 1989 [C]; Gwaltney et al., 1994 [D]).
    - **Note 2**: Imaging findings may persist well after symptoms improve. Computed tomography abnormalities with the common cold may last up to 2 weeks after symptomatic improvement (Gwaltney et al., 1994 [D]). Magnetic resonance imaging changes in patients with symptoms of acute sinusitis may last more than 8 weeks (Leopold, 1994 [D]).
    - **Note 3**: "Limited" sinus computed tomography lacks sensitivity in identifying air-fluid levels (Gross et al., 1991 [C]), suboptimally visualizes the osteomeatal complex 30% of the time and misses 20-30% of the findings found on full computed tomography (Wippold et al., 1995 [C]).

Table. Abnormal Imaging in Children with Upper Respiratory Symptoms

Age Range	Imaging Mechanism	% Abnormal
*6 months-15 years	Plain films	15-57%
**Infants and children	Computed tomography scan	18-67%
***15-85 years	Magnetic resonance imaging	80%

* (Maresh, 1940 [D]; Odita et al., 1986 [C]; Diament et al., 1987 [C]; Shopfner and Rossi, 1973 [C];
**Glasier, Ascher, and Williams, 1986 [C]; Glasier, Mallory, and Steele, 1989 [C];
***Rak et al., 1991 [C])

2. Imaging may be considered appropriate under the following circumstances: (Local Expert Consensus [E]) (see table below for radiologic modalities)

   - In children with acute sinusitis and suspected subperiosteal or orbital abscess. An Otolaryngology (or Ophthalmology if orbital process is present) consultation prior to obtaining radiologic studies in this patient population may reduce the need for an early study and limit repeat radiation exposure (Local Expert Consensus [E]).
   - In children with acute sinusitis and suspected intracranial complications (Local Expert Consensus [E]).
   - In the older child, a clear or normal Water's view may be helpful in ruling out significant maxillary sinus disease taking into consideration clinical findings after unsuccessful therapy (Local Expert Consensus [E]; Lau et al., 1999 [S]; Wald, 1988 [E]).

Table. Radiologic Modalities for Suspected Complications of Acute Sinusitis

Indications	Modalities
Suspected subperiosteal or orbital abscess	Contrast enhanced computed tomography scan or orbits (thin section)
Suspected intracranial complications	Contrast enhanced computed tomography or magnetic resonance imaging of brain

(Local Expert Consensus [E]; ACR, 2000 [E])

Laboratory Assessment

1. Routine laboratory testing such as complete blood count or nasopharyngeal culture is **not** recommended in the initial evaluation in children with uncomplicated sinusitis (Clement et al., 1998 [E]).

   - **Note**: Organisms recovered from nasopharyngeal washings and throat culture do not reflect the organisms found in sinus aspirate (Wald et al., 1981 [B]).

2. Although sinus aspiration and bacterial culture are recognized as the "gold standard" for definitive diagnosis of bacterial sinusitis they are not recom-

mended for use in the initial evaluation and management of the child with uncomplicated acute bacterial sinusitis. Sinus aspiration and culture may need to be considered under the following situations: (Wald et al., 1981 [B])

- Severe illness or toxic looking child
- Immunocompromised child
- Presence of suppurative or intracranial complications

Antibiotic Treatment

The diagnosis and treatment of acute pediatric sinusitis is best considered in light of the duration and severity of symptoms and the increasing prevalence of resistant strains of a common sinus pathogen, *Streptococcus pneumoniae*. The treatment algorithm for this guideline was developed with a focus on antimicrobial activity against *Streptococcus pneumoniae* in an era of increasing penicillin resistance. Clinicians should consider use of the most narrow-spectrum agent that is active against the likely pathogens for the initial antimicrobial treatment of acute bacterial sinusitis in children (Dowell, Schwartz, and Williams, 1998 [E]). Risk factors to consider for penicillin-resistant Streptococcus pneumoniae include (1) daycare attendance (2) recent receipt of antimicrobial therapy (<30 days) (3) age <2 years and (4) exposure to environmental tobacco smoke (ETS) (Block et al., 1995 [C], Levine et al., 1999 [C]; Jacobs et al., 1999 [D]; Brook and Gober, 1999 [D]; Dowell et al., 1999 [E]).

1. In the child with no risk factors for penicillin-resistant Streptococcus pneumoniae standard dose amoxicillin or Augmentin (with standard dose Amoxicillin component) may be considered as initial therapy (Wald et al., 1981 [B]; Wald et al., 1984 [B]; Wald, Chiponis, and Ledesma-Medina, 1986 [B]). (See Table 5 titled "Considerations When Selecting Antibiotic Therapy for Acute Bacterial Sinusitis" in the original guideline document for dosages and other considerations when selecting antibiotic therapy.) It should be recognized however, that the rates of *Streptococcus pneumoniae* resistance to penicillin are increasing nationally and locally (Butler et al., 1996 [C]; Breiman et al., 1994 [C]).

   - **Note**: Forty-six percent of isolates at Children's Hospital Medical Center of Cincinnati, Ohio have intermediate or high Penicillin-resistant *Streptococcus pneumoniae* and local data supports that 15% of children locally may fail initial therapy with standard dose amoxicillin.

2. In children with risk factors for *Streptococcus pneumoniae*, it is recommended that Amoxicillin, high dose (80 to 90 mg/kg/day) or Augmentin (with high dose amoxicillin component) be utilized as first-line therapy (Nash and Wald, 2001 [S]; Wald, Chiponis, and Ledesma-Medina, 1986 [B]; Nelson, Mason, and Kaplan, 1994 [C]; Dowell et al., 1999 [E]; Dowell, 1-1998 [E]; Friedland and McCracken, 1994 [E]; Local Expert Consensus [E]).

- **Note**: Failure with amoxicillin is likely to be due to resistant *Streptococcus pneumoniae*, *Haemophilus influenzae*, or *Moraxella catarrhalis*. High dose amoxicillin will overcome *Streptococcus pneumoniae* resistance (changes in penicillin-binding proteins) (Dowell et al., 1999 [E]; Whitney et al., 2000 [D]). The clavulanic acid component of Augmentin is active against resistant *Haemophilus influenzae* and *Moraxella catarrhalis* (B-lactamase enzyme) (Wald, Chiponis, and Ledesma-Medina, 1986 [B]; Dagan et al., 2000 [A]).

3. Augmentin with high dose amoxicillin (if not used as initial therapy), cefuroxime, cefpodoxime, cefprozil, and cefdinir are reasonable considerations as second-line agents in pediatric acute bacterial sinusitis (Sinus and Allergy Health Partnership, 2000 [E]; Dowell et al., 1999 [E]; Jacobs et al., 1999 [D]).

   - **Note**: Cefuroxime may be superior to other second-generation cephalosporins in children exposed to antibiotics in the past 30 days (Pichichero et al., 1997 [C]; Marchant et al., 1992 [A]; Jacobs et al., 1999 [D]; Dowell et al., 1999 [E]).

4. Once therapeutic response has been demonstrated, it is recommended that the selected therapeutic agent be continued for a minimum of 10 to 14 days in order to minimize the development of bacterial resistance (Local Expert Consensus [E]; Morris, 2000 [M]).

5. If no improvement occurs or if there is worsening of symptoms after 72 hours of therapy with a first or second-line agent, a second or third-line agent may be considered respectively (Wald, Chiponis, and Ledesma-Medina, 1986 [B]; Dowell et al., 1999 [E]).

6. If clinical failure with a second-line agent occurs, third-line agents for pediatric acute bacterial sinusitis are clindamycin and cefixime, with activity against resistant *Streptococcus pneumoniae* and gram negative organisms, respectively (Block et al., 1995 [C]; Dowell et al., 1999 [E]; Sinus and Allergy Health Partnership, 2000 [E]).

7. In the penicillin-allergic patient, second or third-line agents, in addition to the macrolides (erythromycin, clarithromycin), and newer azilides (azithromycin) or trimethoprim-sulfamethoxazole may be considered (Local Expert Consensus [E]).

   - **Note**: Macrolides, azilides, and sulfa containing agents are not recommended as standard therapeutic agents due to either a lack of efficacy data, increasing resistant *Streptococcus pneumoniae* or both (Dagan et al., 2000 [A]; Nelson, Mason, and Kaplan, 1994 [D]; Gay et al., 2000 [D]).

8. In select groups of patients with persistent or recalcitrant sinusitis it may be reasonable to consider an otolaryngology consultation after antibiotic therapy

of at least 6 weeks duration or 3 separate courses are unsuccessful (Local Expert Consensus [E]).

- **Note**: It is preferable to consult with an otolaryngology specialist prior to obtaining imaging in children with suspected sinusitis (Local Expert Consensus [E]).

9. Toxic appearing children who demonstrate poor tolerance of oral intake may require initial parenteral therapy either as an outpatient or short inpatient stay. Reassessment of such patients after initial stabilization may avoid unnecessary imaging and referral early in the course of therapy.

Consultation for Complications of Acute Bacterial Sinusitis

Although children with the complications discussed below (see also Table 4 titled "Complications of Sinusitis" in original guideline document) are listed as exclusions to this guideline, several recommendations are included here to assist the practitioner in decisions regarding consultation to specialists for these key complications.

1. It is recommended that otolaryngology and/or ophthalmology consultation be sought for signs of impending suppurative complications of acute sinusitis (Local Expert Consensus [E]). Such complications are rare but very serious and often result from orbital or intracranial spread of infection (Rosenfeld and Rowley, 1994 [D]).

   - **Note 1**: Preseptal cellulitis, involving only tissue anterior to the orbital septum, manifests as lid edema/erythema, conjunctivitis, and fever. It may be treated with oral antibiotics and close follow-up except where toxicity or specific symptoms preclude adequate antimicrobial effectiveness by mouth.
   - **Note 2**: In cases of orbital spread, as listed above, an otolaryngology/ophthalmology consultation is recommended. Consultation should be considered prior to imaging, so as to limit repeat radiation exposure (Local Expert Consensus [E]).

2. It is recommended that otolaryngology consultation be considered in cases of a moderately to severely ill child with suspected acute frontal or sphenoid sinusitis due to the potential for intracranial spread. Infection arising in either site will generally occur in a relatively older age group (>6 years), and based on the developmental anatomy of these sinuses, the clinical presentation is likely to be more severe (Wolf, Anderhuber, and Kuhn, 1993 [F]).

   - **Note 1**: Acute frontal sinusitis manifests as an intense frontal headache with tenderness over the sinus itself. Spread of infection anteriorly produces periosteal edema and osteomyelitis and may manifest as doughiness of the forehead skin, known as Pott puffy tumor. Spread of infection to the cranial vault results in meningitis or intracranial abscess.

- **Note 2**: Acute isolated sphenoid sinusitis is rare, with an estimated incidence of <1% of all sinusitis cases. (Hnatuk, Macdonald, and Papsin, 1994 [S]; Fearon, Edmonds, and Bird, 1979 [S]; Wyllie, Kern, and Djalilian, 1973 [S]). Acute sphenoid sinusitis represents an elusive diagnosis, (Postma, Chole, and Nemzek, 1995 [S]; Sellars, Goldberg, and Seid, 1975 [S]; Myer et al., 1982 [S]) as signs and symptoms are more variable and non-specific than those of frontal sinus disease. Nasal symptoms may be absent. Headache is severe, deep-seated and worse at night, with the pain radiating to any craniofacial region (Sellars, Goldberg, and Seid, 1975 [S]; Myer et al., 1982 [S]). Suppurative complications may involve any of the vital juxtaposing structures, including the cavernous sinus, intracranial cavity, orbit, pituitary gland and abducens nerve.

Symptomatic Treatment

1. Given the lack of evidence for effectiveness of common agents used for symptomatic treatment of cough or congestion (i.e. reduction in frequency or severity), use of such treatments are not recommended in the routine management of patients with acute sinusitis (Hutton et al., 1991 [B]; Taylor et al., 1993 [B]; McCormick et al., 1996 [B]; Clemens et al., 1997 [B]; Chang et al., 1998 [B]; Bernard et al., 1999 [B]; Davies et al., 1999 [B]; Gadomski and Horton, 1992 [O]).]

   - **Note 1**: Studies measuring a decrease in frequency, severity, and time to resolution of cough or congestion in children with symptoms from upper respiratory infection found no significant difference between any of the therapeutic interventions and placebo. The therapies evaluated were inhaled steroids, inhaled and oral beta-2 agonists, anti-histamines/decongestants (brompheniramine, phenylephrine, phenylpropanolamine, dextromethorphan/guaifenesin, oxymetolazine or ”afrin”), and morphine derivatives (codeine) (Hutton et al., 1991 [B]; Taylor et al., 1993 [B]; McCormick et al., 1996 [B]; Clemens et al., 1997 [B]; Chang et al., 1998 [B]; Bernard et al., 1999 [B]; Davies et al., 1999 [B]; Gadomski and Horton, 1992 [O]).
   - **Note 2**: Given the recent association of one previously common ingredient (phenylpropanolamine) with stroke and that most anti-histamine, decongestants, and antitussives have not been U.S. Food and Drug Administration (FDA) approved in children, their use is discouraged (American Academy of Pediatrics [AAP], 1997 [S]; Kernan et al., 2000 [D]).
   - **Note 3**: Although hypertonic and normal saline and balanced physiological saline nasal washes are commonly used in postoperative patients and in children with chronic sinusitis (Pigret and Jankowski, 1994 [B]; Shoseyov et al, 1998 [B]; Nuutinen et al., 1986 [C]), there is no evidence for their effectiveness in pediatric acute sinusitis.

Parental Expectations

1. It is recommended that physicians explore parental expectations concerning the office visit for a child with an upper respiratory infection.Antibiotic prescription for upper respiratory infections may often be provided because physicians believe parents expect it (MacFarlane et al., 1997 [C]; Barden et al., 1998 [C]; Mangione-Smith et al., 1999 [C]).

    - **Note 1**: Factors that influence the decision to provide an antibiotic prescription include parental expectations, parental assumptions that infections equate to antibiotic need, parental lack of knowledge regarding bacteria and viruses, the role of antibiotics, and pressure from day care centers and employers (MacFarlane et al., 1997 [C]; Barden et al., 1998 [C]; Mangione-Smith et al., 1999 [C]).
    - **Note 2**: Receipt of an antibiotic prescription is not related to satisfaction with care. Satisfaction is related to responsiveness of the physician to the parents' concerns (Mangione-Smith et al., 1999 [C]; Barden et al., 1998 [C]; Hamm, Hicks, and Bemben, 1996 [C]).

2. It is recommended that the physician, staff, or office nurse educate parents regarding the role of viruses in infections and antibiotic resistance risks, thereby minimizing need for inappropriate use of antibiotics (Trepka et al., 2001 [C]; Roberts et al., 1983 [A]; Barden et al., 1998 [C]; Braun et al., 2000 [C]; Braun and Fowles, 2000 [C]; Palmer and Bauchner, 1997 [C]).

    - **Note**: In two double blind, randomized, controlled trials of antibiotic compared to placebo in acute pediatric sinusitis, more than two thirds of children in the placebo arms experienced clinical improvement (Wald, Chiponis, and Ledesma-Medina, 1986 [B]; Garbutt et al., 2001 [B]).

Definitions:

Evidence Based Grading Scale:
A: Randomized controlled trial: large sample
B: Randomized controlled trial: small sample
C: Prospective trial or large case series
D: Retrospective analysis
E: Expert opinion or consensus
F: Basic laboratory research
S: Review article
M: Meta-analysis
Q: Decision analysis
L: Legal requirement
O: Other evidence
X: No evidence

Diagnosis and treatment of otitis media in children.

Institute of Clinical Systems Improvement; year of publication: 2004

The recommendations for the diagnosis and treatment of otitis media and otitis media with effusion in children are presented in the form of an algorithm with 23 components, accompanied by detailed annotations. An algorithm for Diagnosis and Treatment of Otitis Media in Children and an algorithm for Otitis Media with Effusion are provided; clinical highlights and selected annotations (numbered to correspond with the algorithm) follow.

Class of evidence (A-D, M, R, X) ratings are defined at the end of the Major Recommendations field.

Clinical Highlights

1. Schedule an appointment for the child within 24 hours of the call to the clinic. (Annotation #4)

2. A clinical examination is necessary to diagnose acute otitis media. Diagnosis made over the phone is generally discouraged. (Annotation #4)

3. Educate parents on measures to prevent the occurrence of otitis media. (Annotation #6)

4. Prescribe first-line antibiotics (amoxicillin) when the diagnosis of otitis media is made. (Annotation #7)

5. Prescribe second-line antibiotics when the patient fails to respond to first-line drugs, has a history or lack of response to first-line drugs, is hypersensitive to first-line medications, has a resistant organism as defined by culture, or has a coexisting illness requiring a second-line medication. (Annotation #7)

6. Recheck in 3 to 4 weeks or at next well child visit (if within 4 to 6 weeks) for all children <5 years of age, and those 5 years of age or older if risk factors are identified, there is a history of previous ventilation tubes or ear surgery, or if there is a history of speech or development delay. (Annotation #10)

7. Refer the patient to an ear, nose, and throat (ENT) physician when the criteria are met. (Annotation #12)

Diagnosis and Treatment of Otitis Media in Children Algorithm Annotations

1. **Symptoms Suggestive of Otitis Media?**

   Children less than 3 years old more often present with non-specific symptoms (irritability, fever, night waking, poor feeding, coryza, conjunctivitis, and occasionally balance problems). Ninety percent (90%) of infants and toddlers with otitis media have associated rhinitis symptoms.

For more information on symptoms of acute otitis media (AOM) please refer to Annotation Appendix A of the original guideline document, "Recommended Patient Education Content."

Ear pulling without associated symptoms is usually not a symptom of otitis media.

Evidence supporting this recommendation is of classes: C, R

2. **Schedule Appointment Within 24 Hours**

   While symptoms of acute otitis media are often dramatic, the illness is rarely an emergency. Most children can be treated symptomatically through the night unless symptoms of a more serious illness are present. Comfort measures can be discussed with parent/caretaker. (Refer to Annotation Appendix A, "Recommended Patient Education Content" in the original guideline document.)

   Diagnosis of otitis media is made by exam. Diagnosis by phone should be avoided except in special circumstances (children with a history of multiple sets of ventilating tubes or children in high-risk categories such as cleft palate or Down's syndrome who present with bloody or purulent drainage and who are well known to the provider, and in whom follow-up is assured).

3. **Meets Diagnostic Criteria for Acute Otitis Media (AOM)?**

   Middle ear effusion (seen on examination and/or confirmed by pneumatic otoscopy) with:

   a. Local signs of inflammation (redness, bulging)
   b. Symptoms associated with AOM:
      - otalgia
      - otorrhea
      - irritability
      - restlessness
      - poor feeding
      - fever

   AOM is characterized by middle ear effusion with acute inflammation. (The tympanic membrane is usually full or bulging [decreased mobility by pneumatic otoscopy]. Color is usually red, yellow or cloudy.) Symptoms may include otalgia, otorrhea, irritability, restlessness, poor feeding or fever. Tympanometry is usually not necessary to establish the diagnosis of AOM.

4. **Discuss Prevention of Otitis Media**

   Parents/caretakers should be counseled about otitis media prevention. Elimination of controllable risk factors should be encouraged whenever possible.

   Otitis media prevention measures to discuss include:

- encouraging breast feeding
- feeding child upright if bottle fed
- avoiding exposure to passive smoke
- limiting exposure to numbers of children to the extent possible
- teaching adults and children careful hand washing technique
- limiting exposure to viral upper respiratory infections
- avoid pacifier use beyond 10 months of age
- ensure immunizations are up-to-date; including influenza and Prevnar

For more information on prevention of otitis media (OM), please refer to Annotation Appendix A, "Recommended Patient Education Content" in the original guideline document.

Evidence supporting this recommendation is of classes: B, C, D

5. **Initiate Appropriate Treatment**

- **Treatment Options for Acute Otitis Media**
  - Antibiotic regimen using criteria for first- versus second-line antibiotics.
  - Observation of mildly symptomatic children is encouraged in the absence of risk factors. Risk factors may include: severity of symptoms, age <2 years, and parental acceptance.

Options for treatment include:

A. Therapeutic (10 day) course of antibiotics. Consideration may be given to a shortened course of antibiotics (5 days) for children who are at low risk (i.e., age >2 years, no history of chronic or recurrent otitis media and intact tympanic membranes).
   i. First-Line Medications
      a. amoxicillin (40 mg/kg/day) if low risk (>2 years, no day care, and no antibiotics for the past three months).
      b. 80 mg/kg/day if not low risk or for resistant AOM if the lower dose was used initially.
   ii. Recommended second-line medications include: (Check the health plan formulary listing for currently available medications.)
      - amoxicillin/clavulanate potassium (Augmentin)
      - cefuroxime axetil (Ceftin)
      - ceftriaxone sodium (Rocephin): prescribe one dose for new onset otitis media and a three-day course for a truly resistant pattern of otitis media or if oral treatment cannot be given.
      - cefprozil (Cefzil)
      - loracarbef (Lorabid)

- cefdinir (Omnicef)
- cefixime (Suprax)
- cefpodoxime proxetil (Vantin)

iii. Indications for second-line medications include:
- failure to respond to first-line drugs (resistant or persistent acute otitis media)
- history of lack of response to first-line drug (failure of medication on at least two occasions in the current respiratory season)
- hypersensitivity to first-line medications
- presence of resistant organism determined by culture
- coexisting illness requiring a second-line medication

iv. Second-line medications that are currently used but are not as strongly supported in the literature are listed below. These medications are **not** recommended when the patient has failed a course of amoxicillin.
- trimethoprim sulfa (Bactrim, Septra)
- clarithromycin (Biaxin)
- erythromycin ethylsuccinate and sulfisoxazole acetyl (Pediazole)
- azithromycin (Zithromax)

Observation with or without provisional prescription if symptoms of AOM should worsen

This option is not recommended in the acutely ill child but may be considered in an asymptomatic or only mildly symptomatic child with mild findings on exam. Parents should be instructed to call back if symptoms persist, if the child is inconsolable, or if the child is becoming more ill.

For a child with a draining ear, whether from ventilation tubes or perforation, a nontoxic drop (such as ciprofloxin or ofloxacin) may be added to oral antibiotic treatment.

The use of nasal decongestants and corticosteroids is not supported in the literature.

Treatment of Resistant Acute Otitis Media

Resistant acute otitis media (AOM) is defined as persistence of moderately severe symptoms (pain and fever) after 3 to 5 days of antibiotic therapy with findings of continued pressure and inflammation (bulging) behind the tympanic membrane. A second antibiotic should be chosen; the alternative first-line medication may be an appropriate choice. (Referral to ENT specialist may be indicated if significant pain and fever continue for 4 to 5 days on the second medication or if complications of otitis media occur.)

Treatment of Persistent Acute Otitis Media

Persistent AOM is defined as continued findings of AOM present within 6 days of finishing a course of antibiotics. A second course of therapy with a different antibiotic is indicated for persistent AOM.

Evidence supporting this recommendation is of classes: **First-line medications: A, M, R Second-line medications: A, D Treatment of resistant acute otitis media: A, M, R Treatment of persistent acute otitis media: R**

(a) **History of Recurrent Acute Otitis Media?**

History should be reviewed or elicited at the time of diagnosis of AOM. If criteria of recurrent AOM are present, a prophylactic antibiotic regimen follows the therapeutic course of antibiotics. Children in high-risk categories may be considered for more aggressive or earlier intervention with prophylactic antibiotics. The decision for prophylaxis should be based on both the diagnostic criteria and the child's risk factors.

Diagnostic Criteria for Recurrent Acute Otitis Media

- A minimum of three or more episodes of AOM in a 6-month period or during a respiratory season or 4 or more in a year

Children at Increased Risk of Recurrent Acute Otitis Media

- Cleft palate, craniofacial abnormalities and Down's syndrome (very high risk category)
- First episode early (under 6 months)
- Family history of recurrent AOM in a sibling or parent
- Day care attendance
- Exposure to tobacco smoke
- Not breast-fed
- Ethnic origin: Native American or Innuit (Eskimo)

Evidence supporting this recommendation is of classes: B, C, D, R

(b) **Consider Prophylactic Regimen**

Prophylactic Treatment Options

- amoxicillin (20 mg/kg QD [once a day])

The usual duration of antibiotic prophylaxis is 2 to 6 months. Parents should be advised that prophylaxis has been shown to reduce the frequency of AOM by 40% to 50% but will not eliminate its occurrence.

Evidence supporting this recommendation is of class: A

(c) **Schedule Follow-Up in 3 to 4 Weeks**

Follow-up Considerations

- Recheck all children <5 years old
- Recheck children >5 years old if:
  - risk factors identified

* history of previous ventilating tubes or ear surgery
* history of speech or developmental delay

Timing of Rechecks

- Recheck in 3 to 4 weeks or at next well child visit if within the next 4 to 6 weeks.
- Reassess for symptoms of unresponsive otitis: pain, fever, or irritability continuing after 3 to 5 days of treatment. (Refer to Annotation #7, "Initiate Appropriate Treatment.")

Evidence supporting this recommendation is of class: D

(d) **Acute Otitis Media Resolved?**
Resolution is defined as a return to normal on exam with no evidence of effusion or inflammation and/or normal mobility. Tympanometry is not routinely needed to document resolution.

(e) **Criteria for Ear, Nose, and Throat Referral Met?**
A child needs to meet one of the following nine criteria for ear, nose, and throat (ENT) referral for consideration of ventilating tubes:

i. Patients in high-risk categories should be referred immediately to ENT; patients with craniofacial anomalies, Downs' syndrome, cleft palate, and patients with speech and language delay.
ii. Recurrent AOM which fails medical management (>3 episodes in 6 months or >4 episodes in one year) with failure of prophylaxis defined as recurrence x 2 on prophylaxis in a 2 to 6 month time period. (A prophylactic regimen is described in Annotation #9, "Consider Prophylactic Regimen.")
iii. Refractory acute otitis media with moderate to severe symptoms unresponsive to at least 2 antibiotics. (Refer to Annotation #7, "Initiate Appropriate Treatment.")
iv. Bilateral or unilateral otitis media with effusion (OME) persisting for at least 3 months with hearing threshold of 20 dB or worse.
v. Development of advanced middle ear disease involving tympanic membrane atrophy, retraction pockets, ossicular erosion or cholesteatoma.
vi. Medical treatment failure secondary to multiple drug allergy or intolerance.
vii. At least 2 recurrences of otitis media within 2 to 3 months following ventilating tube extrusion with failed medical management.
viii. Impending or actual complication of otitis media including:
   - Mastoiditis
   - Facial nerve paralysis
   - Lateral (sigmoid) sinus thrombosis

- Meningitis
- Brain abscess
- Labyrinthitis

ix. History of six or more months of effusions out of the previous twelve months.

Children at increased risk for otitis media include those under two years of age, those who have an episode of otitis media at less than 6 months of age, children in day care, and children who have a positive family history of otitis media.

Counseling Messages

When counseling parents/caregivers about otitis media prevention, encourage measures to diminish risk factors when possible. (Refer to Annotation #6, "Discuss Prevention Otitis Media.") Discussions with parents should take place regarding medical versus surgical treatment.

Evidence supporting this recommendation is of classes: R, X

(f) **Meets Diagnostic Criteria for Otitis Media with Effusion?**

Middle ear effusion (seen on examination and/or confirmed by pneumatic otoscopy) or abnormal tympanometry or acoustic reflectometry without signs or symptoms of AOM.

The diagnosis of otitis media with effusion (OME) is distinguished from AOM by the presence of an effusion with a lack of signs or symptoms of inflammation or pressure behind the eardrum. Tympanic membrane findings: opaque or yellow, position neutral or retracted, decreased mobility or air fluid level. Tympanometry or pneumatic otoscopy may be useful in establishing the diagnosis.

Evidence supporting this recommendation is of classes: C, R, X

Otitis Media with Effusion Algorithm Annotations

(a) **Consider Treatment Options**

Treatment options to be considered include:

i. Observe–rechecking in 4 to 6 weeks.
   Course of antibiotics should be given as a trial prior to referral for ventilating tubes. Ten-day course of antibiotics using first- and second-line criteria. (Refer to Annotation #7, "Initiate Appropriate Treatment.")

ii. Referral for ventilating tubes if patient meets ENT referral criteria.
   Course of antibiotics should be given as a trial prior to referral for ventilating tubes. Ten-day course of antibiotics using first- and second-line criteria. (Refer to Annotation #7, "Initiate Appropriate Treatment.")

Evidence supporting this recommendation is of class: R

(b) **Follow-up 4 to 6 Weeks**

More frequent rechecking than every 4 to 6 weeks of OME is unnecessary and inappropriate. Ninety (90%) to 95% of OME will

resolve in 3 to 4 months. Continued observation to assure complete resolution is appropriate since hearing loss accompanies OME.
Evidence supporting this recommendation is of classes: A, D

(c) **Otitis Media with Effusion Resolved?**
Mobility of the eardrum should be normal or results of tympanogram or pneumatic otoscopy should confirm resolution.
Evidence supporting this recommendation is of classes: C, X

(d) **Criteria for ENT Referral Met?**
Refer to Annotation #12, "Criteria for ENT Referral Met?"

- **Definitions:**

  Classes of Research Reports

  (a) Primary Reports of New Data Collection:
  - Class A:
    - Randomized, controlled trial
  - Class B:
    - Cohort study
  - Class C:
    - Non-randomized trial with concurrent or historical controls
    - Case-control study
    - Study of sensitivity and specificity of a diagnostic test
    - Population-based descriptive study
  - Class D:
    - Cross-sectional study
    - Case series
    - Case report

  (b) Reports that Synthesize or Reflect upon Collections of Primary Reports:
  - Class M:
    - Meta-analysis
    - Systematic review
    - Decision analysis
    - Cost-effectiveness analysis
  - Class R:
    - Consensus statement
    - Consensus report
    - Narrative review
  - Class X:
    - Medical opinion

Appendix B

Intermediate Representations

B.1 Marked-up Guideline Document

Diagnosis and treatment of otitis media in children.

The recommendations for the diagnosis and treatment of otitis media and otitis media with effusion in children are presented in the form of an algorithm with 23 components, accompanied by detailed annotations. An algorithm for Diagnosis and Treatment of Otitis Media in Children and an algorithm for Otitis Media with Effusion are provided; clinical highlights and selected annotations (numbered to correspond with the algorithm) follow.

Class of evidence (A-D, M, R, X) ratings are defined at the end of the Major Recommendations field.

Clinical Highlights

1. Schedule an appointment for the child within 24 hours of the call to the clinic. (Annotation #4)

2. A clinical examination is necessary to diagnose acute otitis media. Diagnosis made over the phone is generally discouraged. (Annotation #4)

3. Educate parents on measures to prevent the occurrence of otitis media. (Annotation #6)

4. Prescribe first-line antibiotics (amoxicillin) when the diagnosis of otitis media is made. (Annotation #7)

5. Prescribe second-line antibiotics when the patient fails to respond to first-line drugs, has a history or lack of response to first-line drugs, is hypersensitive to first-line medications, has a resistant organism as defined by culture, or has a coexisting illness requiring a second-line medication. (Annotation #7)

6. Recheck in 3 to 4 weeks or at next well child visit (if within 4 to 6 weeks) for all children <5 years of age, and those 5 years of age or older if risk factors are identified, there is a history of previous ventilation tubes or ear surgery, or if there is a history of speech or development delay. (Annotation #10)

7. Refer the patient to an ear, nose, and throat (ENT) physician when the criteria are met. (Annotation #12)

- Diagnosis and Treatment of Otitis Media in Children Algorithm Annotations

  1. Symptoms Suggestive of Otitis Media?

     Children less than 3 years old more often present with non-specific symptoms (irritability, fever, night waking, poor feeding, coryza, conjunctivitis, and occasionally balance problems). Ninety percent (90%) of infants and toddlers with otitis media have associated rhinitis symptoms.

     For more information on symptoms of acute otitis media (AOM) please refer to Annotation Appendix A of the original guideline document, "Recommended Patient Education Content."

     Ear pulling without associated symptoms is usually not a symptom of otitis media.

     Evidence supporting this recommendation is of classes: C, R

  2. [**Schedule Appointment Within 24 Hours**]
     While symptoms of acute otitis media are often dramatic, the illness is rarely an emergency. Most children can be treated symptomatically through the night unless symptoms of a more serious illness are present. Comfort measures can be discussed with parent/caretaker. (Refer to Annotation Appendix A, "Recommended Patient Education Content" in the original guideline document.)

     Diagnosis of otitis media is made by exam. Diagnosis by phone should be avoided except in special circumstances (children with a history of multiple sets of ventilating tubes or children in high-risk categories such as cleft palate or Down's syndrome who present with bloody or purulent drainage and who are well known to the provider, and in whom follow-up is assured).

  3. Meets Diagnostic Criteria for Acute Otitis Media (AOM)?

     Middle ear effusion (seen on examination and/or confirmed by pneumatic otoscopy) with:

     a. Local signs of inflammation (redness, bulging)
     b. Symptoms associated with AOM:
        - otalgia
        - otorrhea
        - irritability
        - restlessness
        - poor feeding
        - fever

     AOM is characterized by middle ear effusion with acute inflammation. (The tympanic membrane is usually full or bulging [decreased mobility by pneumatic otoscopy]. Color is usually red, yellow or cloudy.)

Symptoms may include otalgia, otorrhea, irritability, restlessness, poor feeding or fever. Tympanometry is usually not necessary to establish the diagnosis of AOM.

4. [**Discuss Prevention of Otitis Media**]

   Parents/caretakers should be counseled about otitis media prevention. Elimination of controllable risk factors should be encouraged whenever possible.

   Otitis media prevention measures to discuss include:

   - [**encouraging breast feeding**]
   - [**feeding child upright if bottle fed**]
   - [**avoiding exposure to passive smoke**]
   - [**limiting exposure to numbers of children to the extent possible**]
   - [**teaching adults and children careful hand washing technique**]
   - [**limiting exposure to viral upper respiratory infections**]
   - [**avoid pacifier use beyond 10 months of age**]
   - [**ensure immunizations are up-to-date; including influenza and Prevnar**]

   For more information on prevention of otitis media (OM), please refer to Annotation Appendix A, "Recommended Patient Education Content" in the original guideline document.

   Evidence supporting this recommendation is of classes: B, C, D

5. [**Initiate Appropriate Treatment**]

   - [**Treatment Options for Acute Otitis Media**]
     * Antibiotic regimen using criteria for first- versus second-line antibiotics.
     * Observation of mildly symptomatic children is encouraged in the absence of risk factors. Risk factors may include: severity of symptoms, age <2 years, and parental acceptance.

     Options for treatment include:

     A. [**Therapeutic (10 day) course of antibiotics**.] [**Consideration may be given to a shortened course of antibiotics (5 days) for children who are at low risk (i.e., age >2 years, no history of chronic or recurrent otitis media and intact tympanic membranes).**]

        i. [**First-Line Medications**]

           a. [**amoxicillin (40 mg/kg/day) if low risk (>2 years, no day care, and no antibiotics for the past three months).**]

           b. [**80 mg/kg/day if not low risk or for resistant AOM if the lower dose was used initially.**]

ii. [**Recommended second-line medications include: (Check the health plan formulary listing for currently available medications.)**]
   * [**amoxicillin/clavulanate potassium (Augmentin)**]
   * [**cefuroxime axetil (Ceftin)**]
   * [**ceftriaxone sodium (Rocephin): prescribe one dose for new onset otitis media and a three-day course for a truly resistant pattern of otitis media or if oral treatment cannot be given.**]
   * [**cefprozil (Cefzil)**]
   * [**loracarbef (Lorabid)**]
   * [**cefdinir (Omnicef)**]
   * [**cefixime (Suprax)**]
   * [**cefpodoxime proxetil (Vantin)**]

iii. Indications for second-line medications include:
   * failure to respond to first-line drugs (resistant or persistent acute otitis media)
   * history of lack of response to first-line drug (failure of medication on at least two occasions in the current respiratory season)
   * hypersensitivity to first-line medications
   * presence of resistant organism determined by culture
   * coexisting illness requiring a second-line medication

iv. [**Second-line medications that are currently used but are not as strongly supported in the literature are listed below.**]
   [**These medications are not recommended when the patient has failed a course of amoxicillin.**]
   * [**trimethoprim sulfa (Bactrim, Septra)**]
   * [**clarithromycin (Biaxin)**]
   * [**erythromycin ethylsuccinate and sulfisoxazole acetyl (Pediazole)**]
   * [**azithromycin (Zithromax)**]

   Observation with or without provisional prescription if symptoms of AOM should worsen
   This option is not recommended in the acutely ill child but may be considered in an asymptomatic or only mildly symptomatic child with mild findings on exam. Parents should be instructed to call back if symptoms persist, if the child is inconsolable, or if the child is becoming more ill.
   For a child with a draining ear, whether from ventilation tubes or perforation, a nontoxic drop (such as ciprofloxin or ofloxacin) may be added to oral antibiotic treatment.

[**The use of nasal decongestants and corticosteroids is not supported in the literature.**]

* Treatment of Resistant Acute Otitis Media

  Resistant acute otitis media (AOM) is defined as persistence of moderately severe symptoms (pain and fever) after 3 to 5 days of antibiotic therapy with findings of continued pressure and inflammation (bulging) behind the tympanic membrane. [**A second antibiotic should be chosen; the alternative first-line medication may be an appropriate choice.**] (Referral to ENT specialist may be indicated if significant pain and fever continue for 4 to 5 days on the second medication or if complications of otitis media occur.)

* Treatment of Persistent Acute Otitis Media

  Persistent AOM is defined as continued findings of AOM present within 6 days of finishing a course of antibiotics. [**A second course of therapy with a different antibiotic is indicated for persistent AOM.**]

  Evidence supporting this recommendation is of classes:
  First-line medications: A, M, R
  Second-line medications: A, D
  Treatment of resistant acute otitis media: A, M, R
  Treatment of persistent acute otitis media: R

B. History of Recurrent Acute Otitis Media?

History should be reviewed or elicited at the time of diagnosis of AOM. If criteria of recurrent AOM are present, a prophylactic antibiotic regimen follows the therapeutic course of antibiotics. Children in high-risk categories may be considered for more aggressive or earlier intervention with prophylactic antibiotics. The decision for prophylaxis should be based on both the diagnostic criteria and the child's risk factors.

* Diagnostic Criteria for Recurrent Acute Otitis Media
  - A minimum of three or more episodes of AOM in a 6-month period or during a respiratory season or 4 or more in a year
* Children at Increased Risk of Recurrent Acute Otitis Media
  - Cleft palate, craniofacial abnormalities and Down's syndrome (very high risk category)
  - First episode early (under 6 months)
  - Family history of recurrent AOM in a sibling or parent
  - Day care attendance
  - Exposure to tobacco smoke
  - Not breast-fed

· Ethnic origin: Native American or Innuit (Eskimo)

Evidence supporting this recommendation is of classes: B, C, D, R

C. [**Consider Prophylactic Regimen**]

∗ Prophylactic Treatment Options

· [**amoxicillin (20 mg/kg QD [once a day])**]

[**The usual duration of antibiotic prophylaxis is 2 to 6 months.**] Parents should be advised that prophylaxis has been shown to reduce the frequency of AOM by 40% to 50% but will not eliminate its occurrence.

Evidence supporting this recommendation is of class: A

D. [**Schedule Follow-Up in 3 to 4 Weeks**]

∗ Follow-up Considerations

· Recheck all children <5 years old

· Recheck children >5 years old if:
- risk factors identified
- history of previous ventilating tubes or ear surgery
- history of speech or developmental delay

∗ Timing of Rechecks

· [**Recheck in 3 to 4 weeks or at next well child visit if within the next 4 to 6 weeks.**]

· Reassess for symptoms of unresponsive otitis: pain, fever, or irritability continuing after 3 to 5 days of treatment. (Refer to Annotation #7, "Initiate Appropriate Treatment.")

Evidence supporting this recommendation is of class: D

E. Acute Otitis Media Resolved?

Resolution is defined as a return to normal on exam with no evidence of effusion or inflammation and/or normal mobility. Tympanometry is not routinely needed to document resolution.

F. Criteria for Ear, Nose, and Throat Referral Met?

∗ A child needs to meet one of the following nine criteria for ear, nose, and throat (ENT) referral for consideration of ventilating tubes:

i. Patients in high-risk categories should be referred immediately to ENT; patients with craniofacial anomalies, Downs' syndrome, cleft palate, and patients with speech and language delay.

ii. Recurrent AOM which fails medical management (>3 episodes in 6 months or >4 episodes in one year) with failure of prophylaxis defined as recurrence x 2 on prophylaxis in a 2 to 6 month time period. (A prophylactic regimen is described in Annotation #9, "Consider Prophylactic Regimen.")

iii. Refractory acute otitis media with moderate to severe symptoms unresponsive to at least 2 antibiotics. (Refer to Annotation #7, "Initiate Appropriate Treatment.")

iv. Bilateral or unilateral otitis media with effusion (OME) persisting for at least 3 months with hearing threshold of 20 dB or worse.

v. Development of advanced middle ear disease involving tympanic membrane atrophy, retraction pockets, ossicular erosion or cholesteatoma.

vi. Medical treatment failure secondary to multiple drug allergy or intolerance.

vii. At least 2 recurrences of otitis media within 2 to 3 months following ventilating tube extrusion with failed medical management.

viii. Impending or actual complication of otitis media including:
- Mastoiditis
- Facial nerve paralysis
- Lateral (sigmoid) sinus thrombosis
- Meningitis
- Brain abscess
- Labyrinthitis

ix. History of six or more months of effusions out of the previous twelve months.

Children at increased risk for otitis media include those under two years of age, those who have an episode of otitis media at less than 6 months of age, children in day care, and children who have a positive family history of otitis media.

∗ Counseling Messages

When counseling parents/caregivers about otitis media prevention, encourage measures to diminish risk factors when possible. (Refer to Annotation #6, "Discuss Prevention Otitis Media.") Discussions with parents should take place regarding medical versus surgical treatment.
Evidence supporting this recommendation is of classes: R, X

G. Meets Diagnostic Criteria for Otitis Media with Effusion?

Middle ear effusion (seen on examination and/or confirmed by pneumatic otoscopy) or abnormal tympanometry or acoustic reflectometry without signs or symptoms of AOM.
The diagnosis of otitis media with effusion (OME) is distinguished from AOM by the presence of an effusion with a lack of signs or symptoms of inflammation or pressure behind the eardrum. Tympanic membrane findings: opaque or yellow, po-

sition neutral or retracted, decreased mobility or air fluid level. Tympanometry or pneumatic otoscopy may be useful in establishing the diagnosis.
Evidence supporting this recommendation is of classes: C, R, X

- Otitis Media with Effusion Algorithm Annotations

  1. Consider Treatment Options

     Treatment options to be considered include:

     (a) [**Observe–rechecking in 4 to 6 weeks.**]

     [**Course of antibiotics should be given as a trial prior to referral for ventilating tubes.**] [**Ten-day course of antibiotics using first- and second-line criteria.**] (Refer to Annotation #7, "Initiate Appropriate Treatment.")

     (b) Referral for ventilating tubes if patient meets ENT referral criteria. Course of antibiotics should be given as a trial prior to referral for ventilating tubes. Ten-day course of antibiotics using first- and second-line criteria. (Refer to Annotation #7, "Initiate Appropriate Treatment.")

     Evidence supporting this recommendation is of class: R

  2. [**Follow-up 4 to 6 Weeks**]

     [**More frequent rechecking than every 4 to 6 weeks of OME is unnecessary and inappropriate.**] Ninety (90%) to 95% of OME will resolve in 3 to 4 months. Continued observation to assure complete resolution is appropriate since hearing loss accompanies OME.

     Evidence supporting this recommendation is of classes: A, D

  3. Otitis Media with Effusion Resolved?

     Mobility of the eardrum should be normal or results of tympanogram or pneumatic otoscopy should confirm resolution.

     Evidence supporting this recommendation is of classes: C, X

  4. Criteria for ENT Referral Met?

     Refer to Annotation #12, "Criteria for ENT Referral Met?"

- Definitions:

  Classes of Research Reports

  A. Primary Reports of New Data Collection:
     - Class A:
       * Randomized, controlled trial
     - Class B:
       * Cohort study

- Class C:
  - Non-randomized trial with concurrent or historical controls
  - Case-control study
  - Study of sensitivity and specificity of a diagnostic test
  - Population-based descriptive study
- Class D:
  - Cross-sectional study
  - Case series
  - Case report

B. Reports that Synthesize or Reflect upon Collections of Primary Reports:

- Class M:
  - Meta-analysis
  - Systematic review
  - Decision analysis
  - Cost-effectiveness analysis
- Class R:
  - Consensus statement
  - Consensus report
  - Narrative review
- Class X:
  - Medical opinion

B.2 *SentenceIR* Representation

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE treatment SYSTEM "../../delta/xmlLanguages/SentenceIR/
    SentenceIR.dtd">
<treatment title="Diagnosis and treatment of otitis media in
    children.">
     <sentences>
         <sentence>
             <delta-link link-id="1"/>
             <description>Schedule Appointment Within 24 Hours</
                 description>
         </sentence>
         <sentence>
             <delta-link link-id="2"/>
             <description>Discuss Prevention of Otitis Media</
                 description>
         </sentence>
         <sentence>
             <delta-link link-id="3"/>
             <description>encouraging breast feeding</description>
         </sentence>
         <sentence>
             <delta-link link-id="4"/>
             <description>feeding child upright if bottle fed</
                 description>
         </sentence>
         <sentence>
             <delta-link link-id="5"/>
             <description>avoiding exposure to passive smoke</
                 description>
         </sentence>
         <sentence>
             <delta-link link-id="6"/>
             <description>limiting exposure to numbers of children
                 to the extent possible</description>
         </sentence>
         <sentence>
             <delta-link link-id="7"/>
             <description>teaching adults and children careful hand
                  washing technique</description>
         </sentence>
         <sentence>
             <delta-link link-id="8"/>
             <description>limiting exposure to viral upper
                 respiratory infections</description>
         </sentence>
         <sentence>
             <delta-link link-id="9"/>
             <description>avoid pacifier use beyond 10 months of
                 age</description>
         </sentence>
         <sentence>
             <delta-link link-id="10"/>
             <description>ensure immunizations are up-to-date;
```

```
        including influenza and Prevnar</description>
</sentence>
<sentence>
    <delta-link link-id="11"/>
    <description>Initiate Appropriate Treatment</
        description>
</sentence>
<sentence>
    <delta-link link-id="12"/>
    <description>Treatment Options for Acute Otitis Media
        </description>
</sentence>
<sentence>
    <delta-link link-id="13"/>
    <description>Observation of mildly symptomatic
        children is encouraged in the absence of risk
        factors.</description>
</sentence>
<sentence>
    <delta-link link-id="14"/>
    <description>Therapeutic (10 day) course of
        antibiotics.</description>
</sentence>
<sentence>
    <delta-link link-id="15"/>
    <description>Consideration may be given to a shortened
         course of antibiotics (5 days) for children who
        are at low risk (i.e., age &gt; 2 years, no history
         of chronic or recurrent otitis media and intact
        tympanic membranes).</description>
</sentence>
<sentence>
    <delta-link link-id="16"/>
    <description>First-Line Medications</description>
</sentence>
<sentence>
    <delta-link link-id="17"/>
    <description>amoxicillin (40 mg/kg/day) if low risk (&
        gt; 2 years, no day care, and no antibiotics for
        the past three months).</description>
</sentence>
<sentence>
    <delta-link link-id="18"/>
    <description>80 mg/kg/day if not low risk or for
        resistant AOM if the lower dose was used initially
        .</description>
</sentence>
<sentence>
    <delta-link link-id="19"/>
    <description>Recommended second-line medications
        include: (Check the health plan formulary listing
        for currently available medications.)</description>
</sentence>
<sentence>
    <delta-link link-id="20"/>
```

```
        <description>amoxicillin/clavulanate potassium (
            Augmentin)</description>
    </sentence>
    <sentence>
        <delta-link link-id="21"/>
        <description>cefuroxime axetil (Ceftin)</description>
    </sentence>
    <sentence>
        <delta-link link-id="22"/>
        <description>ceftriaxone sodium (Rocephin): prescribe
            one dose for new onset otitis media and a three-day
             course for a truly resistant pattern of otitis
            media or if oral treatment cannot be given.</
            description>
    </sentence>
    <sentence>
        <delta-link link-id="23"/>
        <description>cefprozil (Cefzil)</description>
    </sentence>
    <sentence>
        <delta-link link-id="24"/>
        <description>loracarbef (Lorabid)</description>
    </sentence>
    <sentence>
        <delta-link link-id="25"/>
        <description>cefdinir (Omnicef)</description>
    </sentence>
    <sentence>
        <delta-link link-id="26"/>
        <description>cefixime (Suprax)</description>
    </sentence>
    <sentence>
        <delta-link link-id="27"/>
        <description>cefpodoxime proxetil (Vantin)</
            description>
    </sentence>
    <sentence>
        <delta-link link-id="28"/>
        <description>Second-line medications that are
            currently used but are not as strongly supported in
             the literature are listed below.</description>
    </sentence>
    <sentence>
        <delta-link link-id="29"/>
        <description>These medications are not recommended
            when the patient has failed a course of amoxicillin
            .</description>
    </sentence>
    <sentence>
        <delta-link link-id="30"/>
        <description>trimethoprim sulfa (Bactrim, Septra)</
            description>
    </sentence>
    <sentence>
        <delta-link link-id="31"/>
```

```
        <description>clarithromycin (Biaxin)</description>
    </sentence>
    <sentence>
        <delta-link link-id="32"/>
        <description>erythromycin ethylsuccinate and
            sulfisoxazole acetyl (Pediazole)</description>
    </sentence>
    <sentence>
        <delta-link link-id="33"/>
        <description>azithromycin (Zithromax)</description>
    </sentence>
    <sentence>
        <delta-link link-id="34"/>
        <description>The use of nasal decongestants and
            corticosteroids is not supported in the literature
            .</description>
    </sentence>
    <sentence>
        <delta-link link-id="35"/>
        <description>A second antibiotic should be chosen; the
             alternative first-line medication may be an
            appropriate choice.</description>
    </sentence>
    <sentence>
        <delta-link link-id="36"/>
        <description>A second course of therapy with a
            different antibiotic is indicated for persistent
            AOM.</description>
    </sentence>
    <sentence>
        <delta-link link-id="37"/>
        <description>Consider Prophylactic Regimen</
            description>
    </sentence>
    <sentence>
        <delta-link link-id="38"/>
        <description>amoxicillin (20 mg/kg QD [once a day])</
            description>
    </sentence>
    <sentence>
        <delta-link link-id="39"/>
        <description>The usual duration of antibiotic
            prophylaxis is 2 to 6 months.</description>
    </sentence>
    <sentence>
        <delta-link link-id="40"/>
        <description>Schedule Follow-Up in 3 to 4 Weeks</
            description>
    </sentence>
    <sentence>
        <delta-link link-id="41"/>
        <description>Recheck in 3 to 4 weeks or at next well
            child visit if within the next 4 to 6 weeks.</
            description>
    </sentence>
```

```
        <sentence>
            <delta-link link-id="42"/>
            <description>Observe--rechecking in 4 to 6 weeks.</
               description>
        </sentence>
        <sentence>
            <delta-link link-id="43"/>
            <description>Course of antibiotics should be given as
               a trial prior to referral for ventilating tubes.</
               description>
        </sentence>
        <sentence>
            <delta-link link-id="44"/>
            <description>Ten-day course of antibiotics using
               first- and second-line criteria.</description>
        </sentence>
        <sentence>
            <delta-link link-id="45"/>
            <description>Referral for ventilating tubes if patient
                meets ENT referral criteria.</description>
        </sentence>
        <sentence>
            <delta-link link-id="46"/>
            <description>Course of antibiotics should be given as
               a trial prior to referral for ventilating tubes.</
               description>
        </sentence>
        <sentence>
            <delta-link link-id="47"/>
            <description>Ten-day course of antibiotics using
               first- and second-line criteria.</description>
        </sentence>
        <sentence>
            <delta-link link-id="48"/>
            <description>Follow-up 4 to 6 Weeks</description>
        </sentence>
        <sentence>
            <delta-link link-id="49"/>
            <description>More frequent rechecking than every 4 to
               6 weeks of OME is unnecessary and inappropriate.</
               description>
        </sentence>
    </sentences>
</treatment>
```

B.3 *ActionIR* Representation

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE treatment SYSTEM "../../delta/xmlLanguages/ActionIR/
    ActionIR.dtd">
<treatment title="Diagnosis and treatment of otitis media in
    children.">
     <actions>
         <action group="3" id="1" parent="0">
             <delta-link link-id="1"/>
             <description>Schedule Appointment Within 24 Hours</
                 description>
             <context>
                 <item>Diagnosis and Treatment of Otitis Media in
                     Children Algorithm Annotations</item>
             </context>
         </action>
         <action group="3" id="2" parent="0">
             <delta-link link-id="2"/>
             <description>Discuss Prevention of Otitis Media</
                 description>
             <context>
                 <item>Diagnosis and Treatment of Otitis Media in
                     Children Algorithm Annotations</item>
             </context>
         </action>
         <action group="6" id="3" parent="2">
             <delta-link link-id="3"/>
             <description>encouraging breast feeding</description>
             <context>
                 <item>Diagnosis and Treatment of Otitis Media in
                     Children Algorithm Annotations</item>
                 <item>Discuss Prevention of Otitis Media</item>
             </context>
         </action>
         <action group="6" id="4" parent="2">
             <delta-link link-id="4"/>
             <description>feeding child upright if bottle fed</
                 description>
             <condition>
                 <item>if bottle fed</item>
             </condition>
             <context>
                 <item>Diagnosis and Treatment of Otitis Media in
                     Children Algorithm Annotations</item>
                 <item>Discuss Prevention of Otitis Media</item>
             </context>
         </action>
         <action group="6" id="5" negative="true" parent="2">
             <delta-link link-id="5"/>
             <description>avoiding exposure to passive smoke</
                 description>
             <context>
                 <item>Diagnosis and Treatment of Otitis Media in
                     Children Algorithm Annotations</item>
```

```
            <item>Discuss Prevention of Otitis Media</item>
        </context>
    </action>
    <action group="6" id="6" parent="2">
        <delta-link link-id="6"/>
        <description>limiting exposure to numbers of children
            to the extent possible</description>
        <context>
            <item>Diagnosis and Treatment of Otitis Media in
                Children Algorithm Annotations</item>
            <item>Discuss Prevention of Otitis Media</item>
        </context>
    </action>
    <action group="6" id="7" parent="2">
        <delta-link link-id="7"/>
        <description>teaching adults and children careful hand
            washing technique</description>
        <context>
            <item>Diagnosis and Treatment of Otitis Media in
                Children Algorithm Annotations</item>
            <item>Discuss Prevention of Otitis Media</item>
        </context>
    </action>
    <action group="6" id="8" parent="2">
        <delta-link link-id="8"/>
        <description>limiting exposure to viral upper
            respiratory infections</description>
        <context>
            <item>Diagnosis and Treatment of Otitis Media in
                Children Algorithm Annotations</item>
            <item>Discuss Prevention of Otitis Media</item>
        </context>
    </action>
    <action group="6" id="9" negative="true" parent="2">
        <delta-link link-id="9"/>
        <description>avoid pacifier use beyond 10 months of
            age</description>
        <context>
            <item>Diagnosis and Treatment of Otitis Media in
                Children Algorithm Annotations</item>
            <item>Discuss Prevention of Otitis Media</item>
        </context>
    </action>
    <action group="6" id="10" parent="2">
        <delta-link link-id="10"/>
        <description>ensure immunizations are up-to-date;
            including influenza and Prevnar</description>
        <context>
            <item>Diagnosis and Treatment of Otitis Media in
                Children Algorithm Annotations</item>
            <item>Discuss Prevention of Otitis Media</item>
        </context>
    </action>
    <action group="3" id="11" parent="0">
        <delta-link link-id="11"/>
```

```
        <description>Initiate Appropriate Treatment</
            description>
        <context>
            <item>Diagnosis and Treatment of Otitis Media in
                Children Algorithm Annotations</item>
        </context>
    </action>
    <action group="7" id="12" parent="11">
        <delta-link link-id="12"/>
        <description>Treatment Options for Acute Otitis Media
            </description>
        <condition>
            <item>for Acute Otitis Media</item>
        </condition>
        <context>
            <item>Diagnosis and Treatment of Otitis Media in
                Children Algorithm Annotations</item>
            <item>Initiate Appropriate Treatment</item>
        </context>
    </action>
    <action group="8" id="13" parent="12">
        <delta-link link-id="13"/>
        <description>Observation of mildly symptomatic
            children is encouraged in the absence of risk
            factors.</description>
        <context>
            <item>Diagnosis and Treatment of Otitis Media in
                Children Algorithm Annotations</item>
            <item>Initiate Appropriate Treatment</item>
            <item>Treatment Options for Acute Otitis Media</
                item>
        </context>
    </action>
    <action group="9" id="14" parent="12">
        <delta-link link-id="14"/>
        <description>Therapeutic (10 day) course of
            antibiotics.</description>
        <agents>
            <agent MeSH="D000900" name="antibiotic">
                <duration term="10 day"/>
            </agent>
            <agent MeSH="D000900" name="antibiotic">
                <duration term="5 days"/>
            </agent>
        </agents>
        <annotations>
            <annotation>Consideration may be given to a
                shortened course of antibiotics (5 days) for
                children who are at low risk (i.e., age &amp;gt
                ; 2 years, no history of chronic or recurrent
                otitis media and intact tympanic membranes).<
                delta-link link-id="15"/>
            </annotation>
            <annotation>The use of nasal decongestants and
                corticosteroids is not supported in the
```

```
                literature.<delta-link link-id="34"/>
            </annotation>
        </annotations>
        <context>
            <item>Diagnosis and Treatment of Otitis Media in
                Children Algorithm Annotations</item>
            <item>Initiate Appropriate Treatment</item>
            <item>Treatment Options for Acute Otitis Media</
                item>
        </context>
    </action>
    <action group="10" id="16" parent="14">
        <delta-link link-id="16"/>
        <description>First-Line Medications</description>
        <context>
            <item>Diagnosis and Treatment of Otitis Media in
                Children Algorithm Annotations</item>
            <item>Initiate Appropriate Treatment</item>
            <item>Treatment Options for Acute Otitis Media</
                item>
            <item>Therapeutic (10 day) course of antibiotics
                .</item>
        </context>
    </action>
    <action group="11" id="17" parent="16">
        <delta-link link-id="17"/>
        <description>amoxicillin (40 mg/kg/day) if low risk (&
            gt; 2 years, no day care, and no antibiotics for
            the past three months).
</description>
        <agents>
            <agent MeSH="D000658" name="amoxicillin">
                <dosage rate="40 mg/kg/day"/>
            </agent>
        </agents>
        <condition>
            <item>if low risk (&amp;gt; 2 years, no day care,
                and no antibiotics for the past three months)</
                item>
        </condition>
        <context>
            <item>Diagnosis and Treatment of Otitis Media in
                Children Algorithm Annotations</item>
            <item>Initiate Appropriate Treatment</item>
            <item>Treatment Options for Acute Otitis Media</
                item>
            <item>Therapeutic (10 day) course of antibiotics
                .</item>
            <item>First-Line Medications</item>
        </context>
    </action>
    <action group="11" id="18" parent="16">
        <delta-link link-id="18"/>
        <description>80 mg/kg/day if not low risk or for
            resistant AOM if the lower dose was used initially
```

```
            .</description>
        <agents>
            <agent MeSH="D000658" name="amoxicillin">
                <dosage rate="80 mg/kg/day"/>
            </agent>
        </agents>
        <condition>
            <item>if not low risk or for resistant AOM if the
                lower dose was used initially.</item>
        </condition>
        <context>
            <item>Diagnosis and Treatment of Otitis Media in
                Children Algorithm Annotations</item>
            <item>Initiate Appropriate Treatment</item>
            <item>Treatment Options for Acute Otitis Media</
                item>
            <item>Therapeutic (10 day) course of antibiotics
                .</item>
            <item>First-Line Medications</item>
        </context>
    </action>
    <action group="10" id="19" parent="14">
        <delta-link link-id="19"/>
        <description>Recommended second-line medications
            include: (Check the health plan formulary listing
            for currently available medications.)</description>
        <context>
            <item>Diagnosis and Treatment of Otitis Media in
                Children Algorithm Annotations</item>
            <item>Initiate Appropriate Treatment</item>
            <item>Treatment Options for Acute Otitis Media</
                item>
            <item>Therapeutic (10 day) course of antibiotics
                .</item>
        </context>
    </action>
    <action group="12" id="20" parent="19" selection="0">
        <delta-link link-id="20"/>
        <description>amoxicillin/clavulanate potassium (
            Augmentin)</description>
        <agents>
            <agent MeSH="D019980" name="amoxicillin/
                clavulanate"/>
        </agents>
        <context>
            <item>Diagnosis and Treatment of Otitis Media in
                Children Algorithm Annotations</item>
            <item>Initiate Appropriate Treatment</item>
            <item>Treatment Options for Acute Otitis Media</
                item>
            <item>Therapeutic (10 day) course of antibiotics
                .</item>
            <item>Recommended second-line medications include:
                (Check the health plan formulary listing for
                currently available medications.)</item>
```

```
        </context>
    </action>
    <action group="12" id="21" parent="19" selection="0">
        <delta-link link-id="21"/>
        <description>cefuroxime axetil (Ceftin)</description>
        <agents>
            <agent MeSH="D002444" name="cefuroxime"/>
        </agents>
        <context>
            <item>Diagnosis and Treatment of Otitis Media in
                Children Algorithm Annotations</item>
            <item>Initiate Appropriate Treatment</item>
            <item>Treatment Options for Acute Otitis Media</
                item>
            <item>Therapeutic (10 day) course of antibiotics
                .</item>
            <item>Recommended second-line medications include:
                 (Check the health plan formulary listing for
                currently available medications.)</item>
        </context>
    </action>
    <action group="12" id="22" parent="19" selection="0">
        <delta-link link-id="22"/>
        <description>ceftriaxone sodium (Rocephin): prescribe
            one dose for new onset otitis media and a three-day
             course for a truly resistant pattern of otitis
            media or if oral treatment cannot be given.</
            description>
        <agents>
            <agent MeSH="D002443" name="ceftriaxone"/>
        </agents>
        <context>
            <item>Diagnosis and Treatment of Otitis Media in
                Children Algorithm Annotations</item>
            <item>Initiate Appropriate Treatment</item>
            <item>Treatment Options for Acute Otitis Media</
                item>
            <item>Therapeutic (10 day) course of antibiotics
                .</item>
            <item>Recommended second-line medications include:
                 (Check the health plan formulary listing for
                currently available medications.)</item>
        </context>
    </action>
    <action group="12" id="23" parent="19" selection="0">
        <delta-link link-id="23"/>
        <description>cefprozil (Cefzil)</description>
        <agents>
            <agent MeSH="C052018" name="cefprozil"/>
        </agents>
        <context>
            <item>Diagnosis and Treatment of Otitis Media in
                Children Algorithm Annotations</item>
            <item>Initiate Appropriate Treatment</item>
            <item>Treatment Options for Acute Otitis Media</
```

```
                item>
            <item>Therapeutic (10 day) course of antibiotics
                .</item>
            <item>Recommended second-line medications include:
                 (Check the health plan formulary listing for
                currently available medications.)</item>
        </context>
    </action>
    <action group="12" id="24" parent="19" selection="0">
        <delta-link link-id="24"/>
        <description>loracarbef (Lorabid)</description>
        <agents>
            <agent MeSH="C054920" name="loracarbef"/>
        </agents>
        <context>
            <item>Diagnosis and Treatment of Otitis Media in
                Children Algorithm Annotations</item>
            <item>Initiate Appropriate Treatment</item>
            <item>Treatment Options for Acute Otitis Media</
                item>
            <item>Therapeutic (10 day) course of antibiotics
                .</item>
            <item>Recommended second-line medications include:
                 (Check the health plan formulary listing for
                currently available medications.)</item>
        </context>
    </action>
    <action group="12" id="25" parent="19" selection="0">
        <delta-link link-id="25"/>
        <description>cefdinir (Omnicef)</description>
        <agents>
            <agent MeSH="C056814" name="cefdinir"/>
        </agents>
        <context>
            <item>Diagnosis and Treatment of Otitis Media in
                Children Algorithm Annotations</item>
            <item>Initiate Appropriate Treatment</item>
            <item>Treatment Options for Acute Otitis Media</
                item>
            <item>Therapeutic (10 day) course of antibiotics
                .</item>
            <item>Recommended second-line medications include:
                 (Check the health plan formulary listing for
                currently available medications.)</item>
        </context>
    </action>
    <action group="12" id="26" parent="19" selection="0">
        <delta-link link-id="26"/>
        <description>cefixime (Suprax)</description>
        <agents>
            <agent MeSH="D020682" name="cefixime"/>
        </agents>
        <context>
            <item>Diagnosis and Treatment of Otitis Media in
                Children Algorithm Annotations</item>
```

```
            <item>Initiate Appropriate Treatment</item>
            <item>Treatment Options for Acute Otitis Media</
               item>
            <item>Therapeutic (10 day) course of antibiotics
               .</item>
            <item>Recommended second-line medications include:
                (Check the health plan formulary listing for
               currently available medications.)</item>
        </context>
    </action>
    <action group="12" id="27" parent="19" selection="0">
        <delta-link link-id="27"/>
        <description>cefpodoxime proxetil (Vantin)</
           description>
        <agents>
            <agent MeSH="C053268" name="cefpodoxime"/>
        </agents>
        <context>
            <item>Diagnosis and Treatment of Otitis Media in
               Children Algorithm Annotations</item>
            <item>Initiate Appropriate Treatment</item>
            <item>Treatment Options for Acute Otitis Media</
               item>
            <item>Therapeutic (10 day) course of antibiotics
               .</item>
            <item>Recommended second-line medications include:
                (Check the health plan formulary listing for
               currently available medications.)</item>
        </context>
    </action>
    <action group="10" id="28" parent="14">
        <delta-link link-id="28"/>
        <description>Second-line medications that are
           currently used but are not as strongly supported in
            the literature are listed below.</description>
        <annotations>
            <annotation>These medications are not recommended
               when the patient has failed a course of
               amoxicillin.<delta-link link-id="29"/>
            </annotation>
        </annotations>
        <context>
            <item>Diagnosis and Treatment of Otitis Media in
               Children Algorithm Annotations</item>
            <item>Initiate Appropriate Treatment</item>
            <item>Treatment Options for Acute Otitis Media</
               item>
            <item>Therapeutic (10 day) course of antibiotics
               .</item>
        </context>
    </action>
    <action group="14" id="30" parent="15" selection="1">
        <delta-link link-id="30"/>
        <description>trimethoprim sulfa (Bactrim, Septra)</
           description>
```

```
        <agents>
            <agent MeSH="D014295" name="trimethoprim"/>
        </agents>
        <context>
            <item>Diagnosis and Treatment of Otitis Media in
                Children Algorithm Annotations</item>
            <item>Initiate Appropriate Treatment</item>
            <item>Treatment Options for Acute Otitis Media</
                item>
            <item>Therapeutic (10 day) course of antibiotics
                .</item>
            <item>Second-line medications that are currently
                used but are not as strongly supported in the
                literature are listed below.</item>
        </context>
    </action>
    <action group="14" id="31" parent="15" selection="1">
        <delta-link link-id="31"/>
        <description>clarithromycin (Biaxin)</description>
        <agents>
            <agent MeSH="D017291" name="clarithromycin"/>
        </agents>
        <context>
            <item>Diagnosis and Treatment of Otitis Media in
                Children Algorithm Annotations</item>
            <item>Initiate Appropriate Treatment</item>
            <item>Treatment Options for Acute Otitis Media</
                item>
            <item>Therapeutic (10 day) course of antibiotics
                .</item>
            <item>Second-line medications that are currently
                used but are not as strongly supported in the
                literature are listed below.</item>
        </context>
    </action>
    <action group="14" id="32" parent="15" selection="1">
        <delta-link link-id="32"/>
        <description>erythromycin ethylsuccinate and
            sulfisoxazole acetyl (Pediazole)</description>
        <agents>
            <agent MeSH="D015643" name="erythromycin
                ethylsuccinate"/>
            <agent MeSH="D013444" name="sulfisoxazole"/>
        </agents>
        <context>
            <item>Diagnosis and Treatment of Otitis Media in
                Children Algorithm Annotations</item>
            <item>Initiate Appropriate Treatment</item>
            <item>Treatment Options for Acute Otitis Media</
                item>
            <item>Therapeutic (10 day) course of antibiotics
                .</item>
            <item>Second-line medications that are currently
                used but are not as strongly supported in the
                literature are listed below.</item>
```

```
        </context>
    </action>
    <action group="14" id="33" parent="15" selection="1">
        <delta-link link-id="33"/>
        <description>azithromycin (Zithromax)</description>
        <agents>
            <agent MeSH="D017963" name="azithromycin"/>
        </agents>
        <context>
            <item>Diagnosis and Treatment of Otitis Media in
                Children Algorithm Annotations</item>
            <item>Initiate Appropriate Treatment</item>
            <item>Treatment Options for Acute Otitis Media</
                item>
            <item>Therapeutic (10 day) course of antibiotics
                .</item>
            <item>Second-line medications that are currently
                used but are not as strongly supported in the
                literature are listed below.</item>
        </context>
    </action>
    <action group="15" id="35" parent="0">
        <delta-link link-id="35"/>
        <description>A second antibiotic should be chosen; the
            alternative first-line medication may be an
            appropriate choice.</description>
        <agents>
            <agent MeSH="D000900" name="antibiotic"/>
        </agents>
        <context>
            <item>Diagnosis and Treatment of Otitis Media in
                Children Algorithm Annotations</item>
            <item>Initiate Appropriate Treatment</item>
            <item>Therapeutic (10 day) course of antibiotics
                .</item>
            <item>Treatment of Resistant Acute Otitis Media</
                item>
        </context>
    </action>
    <action group="16" id="36" parent="0">
        <delta-link link-id="36"/>
        <description>A second course of therapy with a
            different antibiotic is indicated for persistent
            AOM.</description>
        <agents>
            <agent MeSH="D000900" name="antibiotic"/>
        </agents>
        <condition>
            <item>for persistent AOM</item>
        </condition>
        <context>
            <item>Diagnosis and Treatment of Otitis Media in
                Children Algorithm Annotations</item>
            <item>Initiate Appropriate Treatment</item>
            <item>Therapeutic (10 day) course of antibiotics
```

```
                .</item>
            <item>Treatment of Resistant Acute Otitis Media</
                item>
            <item>Treatment of Persistent Acute Otitis Media</
                item>
        </context>
    </action>
    <action group="9" id="37" parent="12">
        <delta-link link-id="37"/>
        <description>Consider Prophylactic Regimen</
            description>
        <context>
            <item>Diagnosis and Treatment of Otitis Media in
                Children Algorithm Annotations</item>
            <item>Initiate Appropriate Treatment</item>
            <item>Treatment Options for Acute Otitis Media</
                item>
        </context>
    </action>
    <action group="22" id="38" parent="37">
        <delta-link link-id="38"/>
        <description>amoxicillin (20 mg/kg QD [once a day])</
            description>
        <agents>
            <agent MeSH="D000658" name="amoxicillin">
                <dosage rate="20 mg/kg"/>
                <iteration term="QD"/>
                <duration term="2 to 6 months"/>
            </agent>
        </agents>
        <annotations>
            <annotation>The usual duration of antibiotic
                prophylaxis is 2 to 6 months.
        <delta-link link-id="39"/>
            </annotation>
        </annotations>
        <context>
            <item>Diagnosis and Treatment of Otitis Media in
                Children Algorithm Annotations</item>
            <item>Initiate Appropriate Treatment</item>
            <item>Treatment Options for Acute Otitis Media</
                item>
            <item>Consider Prophylactic Regimen</item>
            <item>Prophylactic Treatment Options</item>
        </context>
    </action>
    <action group="9" id="40" parent="12">
        <delta-link link-id="40"/>
        <description>Schedule Follow-Up in 3 to 4 Weeks</
            description>
        <context>
            <item>Diagnosis and Treatment of Otitis Media in
                Children Algorithm Annotations</item>
            <item>Initiate Appropriate Treatment</item>
            <item>Treatment Options for Acute Otitis Media</
```

```
            item>
        </context>
    </action>
    <action group="27" id="41" parent="40">
        <delta-link link-id="41"/>
        <description>Recheck in 3 to 4 weeks or at next well
            child visit if within the next 4 to 6 weeks.</
            description>
        <context>
            <item>Diagnosis and Treatment of Otitis Media in
                Children Algorithm Annotations</item>
            <item>Initiate Appropriate Treatment</item>
            <item>Treatment Options for Acute Otitis Media</
                item>
            <item>Schedule Follow-Up in 3 to 4 Weeks</item>
            <item>Timing of Rechecks</item>
        </context>
    </action>
    <action group="34" id="42" parent="0">
        <delta-link link-id="42"/>
        <description>Observe--rechecking in 4 to 6 weeks.</
            description>
        <context>
            <item>Otitis Media with Effusion Algorithm
                Annotations</item>
            <item>Consider Treatment Options</item>
        </context>
    </action>
    <action group="34" id="43" parent="42">
        <delta-link link-id="43"/>
        <description>Course of antibiotics should be given as
            a trial prior to referral for ventilating tubes.</
            description>
        <agents>
            <agent MeSH="D000900" name="antibiotic">
                <duration term="Ten-day"/>
            </agent>
            <agent MeSH="D000900" name="antibiotic">
                <duration term="Ten-day"/>
            </agent>
        </agents>
        <annotations>
            <annotation>Ten-day course of antibiotics using
                first- and second-line criteria.
        <delta-link link-id="44"/>
            </annotation>
        </annotations>
        <context>
            <item>Otitis Media with Effusion Algorithm
                Annotations</item>
            <item>Consider Treatment Options</item>
            <item>Observe--rechecking in 4 to 6 weeks.</item>
        </context>
    </action>
    <action group="34" id="45" parent="0">
```

```
        <delta-link link-id="45"/>
        <description>Referral for ventilating tubes if patient
            meets ENT referral criteria.</description>
        <condition>
            <item>if patient meets ENT referral criteria.</
                item>
        </condition>
        <annotations>
            <annotation>More frequent rechecking than every 4
                to 6 weeks of OME is unnecessary and
                inappropriate.
        <delta-link link-id="49"/>
            </annotation>
        </annotations>
        <context>
            <item>Otitis Media with Effusion Algorithm
                Annotations</item>
            <item>Consider Treatment Options</item>
        </context>
    </action>
    <action group="34" id="46" parent="45">
        <delta-link link-id="46"/>
        <description>Course of antibiotics should be given as
           a trial prior to referral for ventilating tubes.</
           description>
        <agents>
            <agent MeSH="D000900" name="antibiotics">
                <duration term="Ten-day"/>
            </agent>
        </agents>
        <annotations>
            <annotation>Ten-day course of antibiotics using
                first- and second-line criteria.<delta-link
                link-id="47"/>
            </annotation>
        </annotations>
        <context>
            <item>Otitis Media with Effusion Algorithm
                Annotations</item>
            <item>Consider Treatment Options</item>
            <item>Referral for ventilating tubes if patient
                meets ENT referral criteria.</item>
        </context>
    </action>
    <action group="33" id="48" parent="0">
        <delta-link link-id="48"/>
        <description>Follow-up 4 to 6 Weeks</description>
        <annotations>
            <annotation>More frequent rechecking than every 4
                to 6 weeks of OME is unnecessary and
                inappropriate.<delta-link link-id="49"/>
            </annotation>
        </annotations>
        <context>
            <item>Otitis Media with Effusion Algorithm
```

```
                Annotations</item>
            </context>
        </action>
    </actions>
    <structure>
        <group id="1"/>
        <group id="2">
            <group id="3">
                <group id="4">
                    <group id="5"/>
                </group>
                <group id="6"/>
                <group id="7">
                    <group id="8"/>
                    <group id="9">
                        <group id="10">
                            <group id="11"/>
                            <group id="12">
                                <selection id="0"/>
                            </group>
                            <group id="13"/>
                            <group id="14">
                                <selection id="1"/>
                            </group>
                        </group>
                        <group id="15"/>
                        <group id="16"/>
                        <group id="17">
                            <group id="18"/>
                        </group>
                        <group id="19">
                            <group id="20"/>
                        </group>
                        <group id="21">
                            <group id="22"/>
                        </group>
                        <group id="23">
                            <group id="24">
                                <group id="25"/>
                            </group>
                        </group>
                        <group id="26">
                            <group id="27"/>
                        </group>
                        <group id="28">
                            <group id="29">
                                <group id="30"/>
                            </group>
                        </group>
                        <group id="31"/>
                    </group>
                </group>
            </group>
        </group>
        <group id="32">
```

```
            <group id="33">
                <group id="34"/>
            </group>
        </group>
        <group id="35">
            <group id="36">
                <group id="37"/>
                <group id="38"/>
                <group id="39"/>
                <group id="40"/>
                <group id="41"/>
                <group id="42"/>
                <group id="43"/>
            </group>
        </group>
    </structure>
    <relations>
        <relation id="9" type="SUCCEEDING">
            <source id="14">
                <delta-link link-id="14"/>
            </source>
            <destination id="40">
                <delta-link link-id="40"/>
            </destination>
        </relation>
        <relation id="9" type="SUCCEEDING">
            <source id="37">
                <delta-link link-id="37"/>
            </source>
            <destination id="40">
                <delta-link link-id="40"/>
            </destination>
        </relation>
    </relations>
</treatment>
```

B.4 *AsbruIR* Representation

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE treatment SYSTEM "../../delta/xmlLanguages/AsbruIR/
   AsbruIR.dtd">
<treatment title="Diagnosis and treatment of otitis media in
   children.">
   <structure>
      <group id="2">
         <action id="PARENT_0" parent="0" group="2">
            <description>Diagnosis and Treatment of Otitis Media
               in Children Algorithm Annotations</description>
         </action>
         <group id="3">
            <action group="3" id="1" parent="PARENT_0">
               <delta-link link-id="1"/>
               <description>Schedule Appointment Within 24 Hours</
                  description>
            </action>
            <action group="3" id="2" parent="PARENT_0">
               <delta-link link-id="2"/>
               <description>Discuss Prevention of Otitis Media</
                  description>
            </action>
            <group id="6">
               <action group="6" id="3" parent="2">
                  <delta-link link-id="3"/>
                  <description>encouraging breast feeding</
                     description>
               </action>
               <action group="6" id="4" parent="2">
                  <delta-link link-id="4"/>
                  <description>feeding child upright if bottle fed
                     </description>
                  <condition>
                     <item>if bottle fed</item>
                  </condition>
               </action>
               <action group="6" id="6" parent="2">
                  <delta-link link-id="6"/>
                  <description>limiting exposure to numbers of
                     children to the extent possible</description>
               </action>
               <action group="6" id="7" parent="2">
                  <delta-link link-id="7"/>
                  <description>teaching adults and children
                     careful hand washing technique</description>
               </action>
               <action group="6" id="8" parent="2">
                  <delta-link link-id="8"/>
                  <description>limiting exposure to viral upper
                     respiratory infections</description>
               </action>
               <action group="6" id="10" parent="2">
                  <delta-link link-id="10"/>
```

```
            <description>ensure immunizations are up-to-date
                ; including influenza and Prevnar</
                description>
            <annotations>
               <annotation>avoiding exposure to passive
                   smoke
                  <delta-link link-id="5"/>
               </annotation>
               <annotation>avoid pacifier use beyond 10
                   months of age
                  <delta-link link-id="9"/>
               </annotation>
            </annotations>
         </action>
      </group>
      <action group="3" id="11" parent="PARENT_0">
         <delta-link link-id="11"/>
         <description>Initiate Appropriate Treatment</
             description>
      </action>
      <action id="PARENT_3" parent="PARENT_0" group="3">
         <description>Initiate Appropriate Treatment</
             description>
      </action>
      <group id="7">
         <action group="7" id="12" parent="11">
            <delta-link link-id="12"/>
            <description>Treatment Options for Acute Otitis
               Media</description>
            <condition>
               <item diagnosis="Acute Otitis Media">for
                  Acute Otitis Media</item>
            </condition>
         </action>
         <group id="8">
            <action group="8" id="13" parent="12">
               <delta-link link-id="13"/>
               <description>Observation of mildly
                  symptomatic children is encouraged in the
                  absence of risk factors.</description>
            </action>
         </group>
         <action id="PARENT_2" parent="PARENT_3" group="7">
            <description>Therapeutic (10 day) course of
               antibiotics.</description>
         </action>
         <group id="9">
            <action group="9" id="14" parent="12">
               <delta-link link-id="14"/>
               <description>Therapeutic (10 day) course of
                  antibiotics.</description>
               <agents>
                  <agent MeSH="D000900" name="antibiotic">
                     <duration term="10 day">
                        <minimum value="10" unit="d"/>
```

```
                        <maximum value="10" unit="d"/>
                    </duration>
                </agent>
                <agent MeSH="D000900" name="antibiotic">
                    <duration term="5 days">
                        <minimum value="5" unit="d"/>
                        <maximum value="5" unit="d"/>
                    </duration>
                </agent>
            </agents>
            <annotations>
                <annotation>Consideration may be given to
                    a shortened course of antibiotics (5
                    days) for children who are at low risk
                    (i.e., age &amp;gt; 2 years, no history
                     of chronic or recurrent otitis media
                    and intact tympanic membranes).
                    <delta-link link-id="15"/>
                </annotation>
                <annotation>The use of nasal decongestants
                     and corticosteroids is not supported
                    in the literature.
                    <delta-link link-id="34"/>
                </annotation>
            </annotations>
        </action>
        <group id="10">
            <action group="10" id="16" parent="14">
                <delta-link link-id="16"/>
                <description>First-Line Medications</
                    description>
            </action>
            <group id="11">
                <action group="11" id="17" parent="16">
                    <delta-link link-id="17"/>
                    <description>amoxicillin (40 mg/kg/day)
                         if low risk (&gt; 2 years, no day
                        care, and no antibiotics for the
                        past three months).</description>
                    <agents>
                        <agent MeSH="D000658" name="
                            amoxicillin">
                            <dosage rate="40 mg/kg/day"/>
                        </agent>
                    </agents>
                    <condition>
                        <item>if low risk (&amp;gt; 2 years,
                             no day care, and no antibiotics
                            for the past three months)</item>
                    </condition>
                </action>
                <action group="11" id="18" parent="16">
                    <delta-link link-id="18"/>
                    <description>80 mg/kg/day if not low
                        risk or for resistant AOM if the
```

```
                lower dose was used initially.</
                description>
            <agents>
                <agent MeSH="D000658" name="
                    amoxicillin">
                    <dosage rate="80 mg/kg/day"/>
                </agent>
            </agents>
            <condition>
                <item diagnosis="AOM">if not low
                    risk or for resistant AOM if the
                    lower dose was used initially.</
                    item>
            </condition>
        </action>
    </group>
    <action group="10" id="19" parent="14">
        <delta-link link-id="19"/>
        <description>Recommended second-line
            medications include: (Check the health
            plan formulary listing for currently
            available medications.)</description>
    </action>
    <group id="12">
        <action id="SELECT_0" parent="19" group="
            12">
            <condition>
                <item type="complete-condition" plan
                    ="20" state="completed"/>
                <item type="complete-condition" plan
                    ="21" state="completed"/>
                <item type="complete-condition" plan
                    ="22" state="completed"/>
                <item type="complete-condition" plan
                    ="23" state="completed"/>
                <item type="complete-condition" plan
                    ="24" state="completed"/>
                <item type="complete-condition" plan
                    ="25" state="completed"/>
                <item type="complete-condition" plan
                    ="26" state="completed"/>
                <item type="complete-condition" plan
                    ="27" state="completed"/>
            </condition>
        </action>
        <selection id="0">
            <action group="12" id="20" parent="
                SELECT_0" selection="0">
                <delta-link link-id="20"/>
                <description>amoxicillin/clavulanate
                    potassium (Augmentin)</
                    description>
                <agents>
                    <agent MeSH="D019980" name="
                        amoxicillin/clavulanate"/>
```

```
    </agents>
</action>
<action group="12" id="21" parent="
    SELECT_0" selection="0">
    <delta-link link-id="21"/>
    <description>cefuroxime axetil (
        Ceftin)</description>
    <agents>
          <agent MeSH="D002444" name="
              cefuroxime"/>
        </agents>
    </action>
<action group="12" id="22" parent="
    SELECT_0" selection="0">
    <delta-link link-id="22" />
    <description>ceftriaxone sodium (
        Rocephin): prescribe one dose for
         new onset otitis media and a
        three-day course for a truly
        resistant pattern of otitis media
         or if oral treatment cannot be
        given.</description>
    <agents>
        <agent MeSH="D002443" name="
            ceftriaxone" />
    </agents>
</action>
<action group="12" id="23" parent="
    SELECT_0" selection="0">
    <delta-link link-id="23" />
    <description>cefprozil (Cefzil)</
        description>
    <agents>
        <agent MeSH="C052018" name="
            cefprozil" />
    </agents>
</action>
<action group="12" id="24" parent="
    SELECT_0" selection="0">
    <delta-link link-id="24" />
    <description>loracarbef (Lorabid)</
        description>
    <agents>
        <agent MeSH="C054920" name="
            loracarbef" />
    </agents>
</action>
<action group="12" id="25" parent="
    SELECT_0" selection="0">
    <delta-link link-id="25" />
    <description>cefdinir (Omnicef)</
        description>
    <agents>
        <agent MeSH="C056814" name="
            cefdinir" />
```

```
            </agents>
         </action>
         <action group="12" id="26" parent="
            SELECT_0" selection="0">
            <delta-link link-id="26" />
            <description>cefixime (Suprax)</
               description>
            <agents>
               <agent MeSH="D020682" name="
                  cefixime" />
            </agents>
         </action>
         <action group="12" id="27" parent="
            SELECT_0" selection="0">
            <delta-link link-id="27" />
            <description>cefpodoxime proxetil (
               Vantin)</description>
            <agents>
               <agent MeSH="C053268" name="
                  cefpodoxime" />
            </agents>
         </action>
      </selection>
   </group>
   <action group="10" id="28" parent="14">
      <delta-link link-id="28" />
      <description>Second-line medications that
         are currently used but are not as
         strongly supported in the literature
         are listed below.</description>
      <annotations>
         <annotation>These medications are not
            recommended when the patient has
            failed a course of amoxicillin.
            <delta-link link-id="29" />
         </annotation>
      </annotations>
   </action>
   <group id="14">
      <action id="SELECT_1" parent="15" group="
         14">
         <condition>
            <item type="complete-condition" plan
               ="30" state="completed" />
            <item type="complete-condition" plan
               ="31" state="completed" />
            <item type="complete-condition" plan
               ="32" state="completed" />
            <item type="complete-condition" plan
               ="33" state="completed" />
         </condition>
      </action>
      <selection id="1">
         <action group="14" id="30" parent="
            SELECT_1" selection="1">
```

```
                    <delta-link link-id="30" />
                    <description>trimethoprim sulfa (
                        Bactrim, Septra)</description>
                    <agents>
                        <agent MeSH="D014295" name="
                            trimethoprim" />
                    </agents>
                </action>
                <action group="14" id="31" parent="
                    SELECT_1" selection="1">
                    <delta-link link-id="31" />
                    <description>clarithromycin (Biaxin)
                        </description>
                    <agents>
                        <agent MeSH="D017291" name="
                            clarithromycin" />
                    </agents>
                </action>
                <action group="14" id="32" parent="
                    SELECT_1" selection="1">
                    <delta-link link-id="32" />
                    <description>erythromycin
                        ethylsuccinate and sulfisoxazole
                        acetyl (Pediazole)</description>
                    <agents>
                        <agent MeSH="D015643" name="
                            erythromycin ethylsuccinate"
                            />
                        <agent MeSH="D013444" name="
                            sulfisoxazole" />
                    </agents>
                </action>
                <action group="14" id="33" parent="
                    SELECT_1" selection="1">
                    <delta-link link-id="33" />
                    <description>azithromycin (Zithromax
                        )</description>
                    <agents>
                        <agent MeSH="D017963" name="
                            azithromycin" />
                    </agents>
                </action>
            </selection>
        </group>
    </group>
    <action id="PARENT_1" parent="PARENT_2" group="9
        ">
        <description>Treatment of Resistant Acute
            Otitis Media</description>
    </action>
    <group id="15">
        <action group="15" id="35" parent="PARENT_1">
            <delta-link link-id="35" />
            <description>A second antibiotic should be
                chosen; the alternative first-line
```

```
                medication may be an appropriate choice
                .</description>
            <agents>
                <agent MeSH="D000900" name="antibiotic"
                    />
            </agents>
        </action>
    </group>
    <action id="PARENT_4" parent="PARENT_1" group="9
        ">
        <description>Treatment of Persistent Acute
            Otitis Media</description>
    </action>
    <group id="16">
        <action group="16" id="36" parent="PARENT_4">
            <delta-link link-id="36" />
            <description>A second course of therapy
                with a different antibiotic is
                indicated for persistent AOM.</
                description>
            <agents>
                <agent MeSH="D000900" name="antibiotic"
                    />
            </agents>
            <condition>
                <item diagnosis="persistent AOM">for
                    persistent AOM</item>
            </condition>
        </action>
    </group>
    <action group="9" id="37" parent="12">
        <delta-link link-id="37" />
        <description>Consider Prophylactic Regimen</
            description>
    </action>
    <group id="22">
        <action group="22" id="38" parent="37">
            <delta-link link-id="38"/>
            <description>amoxicillin (20 mg/kg QD [
                once a day])</description>
            <agents>
                <agent MeSH="D000658" name="amoxicillin
                    ">
                    <dosage rate="20 mg/kg"/>
                    <iteration term="QD"  specification=
                        "CYCLICAL">
                        <frequency value="1" unit="d"/>
                        <minimum value="4" unit="h"/>
                        <maximum value="20" unit="h"/>
                    </iteration>
                    <duration term="2 to 6 months">
                        <minimum value="2" unit="m" />
                        <maximum value="6" unit="m" />
                    </duration>
                </agent>
```

```
                        </agents>
                        <annotations>
                           <annotation>The usual duration of
                              antibiotic prophylaxis is 2 to 6
                              months.
                              <delta-link link-id="39"/>
                           </annotation>
                        </annotations>
                     </action>
                  </group>
                  <action group="9" id="40" parent="12">
                     <delta-link link-id="40" />
                     <description>Schedule Follow-Up in 3 to 4
                        Weeks</description>
                     <condition>
                        <item type="setup-precondition" plan="14"
                           state="completed" />
                        <item type="setup-precondition" plan="37"
                           state="completed" />
                     </condition>
                  </action>
                  <group id="27">
                     <action group="27" id="41" parent="40">
                        <delta-link link-id="41" />
                        <description>Recheck in 3 to 4 weeks or at
                           next well child visit if within the
                           next 4 to 6 weeks.</description>
                     </action>
                  </group>
               </group>
            </group>
         </group>
      </group>
      <group id="32">
         <action id="PARENT_6" parent="0" group="32">
            <description>Otitis Media with Effusion Algorithm
               Annotations</description>
         </action>
         <group id="33">
            <action id="PARENT_5" parent="PARENT_6" group="33">
               <description>Consider Treatment Options</
                  description>
            </action>
            <group id="34">
               <action group="34" id="42" parent="PARENT_5">
                     <delta-link link-id="42" />
                     <description>Observe--rechecking in 4 to 6
                        weeks.</description>
               </action>
               <action group="34" id="43" parent="42">
                  <delta-link link-id="43" />
                  <description>Course of antibiotics should be
                     given as a trial prior to referral for
                     ventilating tubes.</description>
                  <agents>
```

```
                <agent MeSH="D000900" name="antibiotic">
                  <duration term="Ten-day">
                    <minimum value="10" unit="d" />
                    <maximum value="10" unit="d" />
                  </duration>
                </agent>
              </agents>
            <annotations>
              <annotation>Ten-day course of antibiotics
                using first- and second-line criteria.
                <delta-link link-id="44" />
              </annotation>
            </annotations>
          </action>
          <action group="34" id="45" parent="PARENT_5">
            <delta-link link-id="45" />
            <description>Referral for ventilating tubes if
              patient meets ENT referral criteria.</
              description>
            <condition>
              <item patient="patient meets ENT referral
                criteria.">if patient meets ENT referral
                criteria.</item>
            </condition>
            <annotations>
              <annotation>More frequent rechecking than
                every 4 to 6 weeks of OME is unnecessary
                and inappropriate.
                <delta-link link-id="49" />
              </annotation>
            </annotations>
          </action>
          <action group="34" id="46" parent="45">
            <delta-link link-id="46" />
            <description>Course of antibiotics should be
              given as a trial prior to referral for
              ventilating tubes.</description>
            <agents>
              <agent MeSH="D000900" name="antibiotics">
                <duration term="Ten-day">
                  <minimum value="10" unit="d" />
                  <maximum value="10" unit="d" />
                </duration>
              </agent>
            </agents>
            <annotations>
              <annotation>Ten-day course of antibiotics
                using first- and second-line criteria.
                <delta-link link-id="47" />
              </annotation>
            </annotations>
          </action>
        </group>
        <action group="33" id="48" parent="PARENT_6">
          <delta-link link-id="48" />
```

```
                <description>Follow-up 4 to 6 Weeks</description>
                <annotations>
                   <annotation>More frequent rechecking than every
                      4 to 6 weeks of OME is unnecessary and
                      inappropriate.
                      <delta-link link-id="49" />
                   </annotation>
                </annotations>
             </action>
          </group>
       </group>
    </structure>
</treatment>
```

B.5 *Asbru* Representation

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE plan-library SYSTEM "../lib/Asbru_7.3h_DELTA.dtd">
<plan-library>
   <domain-defs>
      <domain name="
         Diagnosis_and_treatment_of_otitis_media_in_children">
         <parameter-group title="Patient parameters">
            <parameter-def name="
               if_patient_meets_ENT_referral_criteria" required="
               no" type="any">
               <raw-data-def mode="manual" use-as-context="no"
                  user-text="Patient meets ENT referral criteria?"
                  />
            </parameter-def>
         </parameter-group>
         <parameter-group title="Diagnosis parameters">
            <parameter-def name="for_Acute_Otitis_Media" required=
               "no" type="any">
               <raw-data-def mode="manual" use-as-context="no"
                  user-text="Diagnosis is Acute Otitis Media?"/>
            </parameter-def>
            <parameter-def name="for_persistent_AOM" required="no"
                type="any">
               <raw-data-def mode="manual" use-as-context="no"
                  user-text="Diagnosis is persistent AOM?"/>
            </parameter-def>
         </parameter-group>
         <parameter-group title="Various parameters">
            <parameter-def name="if_bottle_fed" required="no" type
               ="any">
               <raw-data-def mode="manual" use-as-context="no"
                  user-text="if bottle fed?"/>
            </parameter-def>
            <parameter-def name="if_low_risk_&amp;gt;
               _years_no_day_care_and_no_antibiotics_for_the_past_three_months
               " required="no" type="any">
               <raw-data-def mode="manual" use-as-context="no"
                  user-text="if low risk (&amp;gt; 2 years, no day
                   care, and no antibiotics for the past three
                  months)?"/>
            </parameter-def>
            <parameter-def name="
               if_not_low_risk_or_for_resistant_AOM_if_the_lower_dose_was_used
               " required="no" type="any">
               <raw-data-def mode="manual" use-as-context="no"
                  user-text="if not low risk or for resistant AOM
                  if the lower dose was used initially.?"/>
            </parameter-def>
         </parameter-group>
      </domain>
   </domain-defs>
   <plans>
      <plan-group>
```

```
<plan name="plan_0" title="Diagnosis and treatment of
    otitis media in children.">
   <plan-body>
      <subplans type="unordered">
         <wait-for>
            <all/>
         </wait-for>
         <plan-activation>
            <plan-schema name="PLAN_PARENT_0">
               <delta-link link-id="0"/>
            </plan-schema>
         </plan-activation>
         <plan-activation>
            <plan-schema name="PLAN_PARENT_6">
               <delta-link link-id="6"/>
            </plan-schema>
         </plan-activation>
      </subplans>
   </plan-body>
</plan>
<plan name="PLAN_PARENT_0" title="Diagnosis and Treatment
     of Otitis Media in Children Algorithm Annotations">
   <conditions>
      <setup-precondition confirmation-required="yes">
         <none/>
      </setup-precondition>
   </conditions>
   <plan-body>
      <subplans type="unordered">
         <wait-for>
            <all/>
         </wait-for>
         <plan-activation>
            <plan-schema name="PLAN_1">
               <delta-link link-id="1"/>
            </plan-schema>
         </plan-activation>
         <plan-activation>
            <plan-schema name="PLAN_2">
               <delta-link link-id="2"/>
            </plan-schema>
         </plan-activation>
         <plan-activation>
            <plan-schema name="PLAN_11">
               <delta-link link-id="11"/>
            </plan-schema>
         </plan-activation>
         <plan-activation>
            <plan-schema name="PLAN_PARENT_3">
               <delta-link link-id="3"/>
            </plan-schema>
         </plan-activation>
      </subplans>
   </plan-body>
</plan>
```

```
<plan name="PLAN_1" title="Schedule Appointment Within 24
    Hours">
   <delta-link link-id="1"/>
   <conditions>
      <setup-precondition confirmation-required="yes">
         <none/>
      </setup-precondition>
   </conditions>
   <plan-body>
      <user-performed/>
   </plan-body>
</plan>
<plan name="PLAN_2" title="Discuss Prevention of Otitis
   Media">
   <delta-link link-id="2"/>
   <conditions>
      <setup-precondition confirmation-required="yes">
         <none/>
      </setup-precondition>
   </conditions>
   <plan-body>
      <subplans type="unordered">
         <wait-for>
            <all/>
         </wait-for>
         <plan-activation>
            <plan-schema name="PLAN_3">
               <delta-link link-id="3"/>
            </plan-schema>
         </plan-activation>
         <plan-activation>
            <plan-schema name="PLAN_4">
               <delta-link link-id="4"/>
            </plan-schema>
         </plan-activation>
         <plan-activation>
            <plan-schema name="PLAN_6">
               <delta-link link-id="6"/>
            </plan-schema>
         </plan-activation>
         <plan-activation>
            <plan-schema name="PLAN_7">
               <delta-link link-id="7"/>
            </plan-schema>
         </plan-activation>
         <plan-activation>
            <plan-schema name="PLAN_8">
               <delta-link link-id="8"/>
            </plan-schema>
         </plan-activation>
         <plan-activation>
            <plan-schema name="PLAN_10">
               <delta-link link-id="10"/>
            </plan-schema>
         </plan-activation>
```

```
            </subplans>
         </plan-body>
      </plan>
      <plan name="PLAN_3" title="encouraging breast feeding">
         <delta-link link-id="3"/>
         <conditions>
            <setup-precondition confirmation-required="yes">
               <none/>
            </setup-precondition>
         </conditions>
         <plan-body>
            <user-performed/>
         </plan-body>
      </plan>
      <plan name="PLAN_4" title="feeding child upright if
          bottle fed">
         <delta-link link-id="4"/>
         <conditions>
            <setup-precondition confirmation-required="yes">
               <none/>
            </setup-precondition>
         </conditions>
         <plan-body>
            <subplans retry-aborted-subplans="no" type="
                any-order" wait-for-optional-subplans="no">
               <wait-for>
                  <all/>
               </wait-for>
               <ask>
                  <parameter-ref name="if_bottle_fed"/>
                  <time-out>
                     <now/>
                  </time-out>
               </ask>
               <if-then-else>
                  <simple-condition>
                     <simple-condition-combination type="and">
                        <comparison type="equal">
                           <left-hand-side>
                              <parameter-ref name="
                                 if_bottle_fed"/>
                           </left-hand-side>
                           <right-hand-side>
                              <qualitative-constant value="yes"
                                 />
                           </right-hand-side>
                        </comparison>
                     </simple-condition-combination>
                  </simple-condition>
                  <then-branch>
                     <plan-activation>
                        <plan-schema name="ATOMIC_PLAN_4">
                           <delta-link link-id="4"/>
                        </plan-schema>
                     </plan-activation>
```

```
            </then-branch>
         </if-then-else>
      </subplans>
   </plan-body>
</plan>
<plan name="ATOMIC_PLAN_4" title="feeding child upright
   if bottle fed">
   <plan-body>
      <user-performed/>
   </plan-body>
</plan>
<plan name="PLAN_6" title="limiting exposure to numbers
   of children to the extent possible">
   <delta-link link-id="6"/>
   <conditions>
      <setup-precondition confirmation-required="yes">
         <none/>
      </setup-precondition>
   </conditions>
   <plan-body>
      <user-performed/>
   </plan-body>
</plan>
<plan name="PLAN_7" title="teaching adults and children
   careful hand washing technique">
   <delta-link link-id="7"/>
   <conditions>
      <setup-precondition confirmation-required="yes">
         <none/>
      </setup-precondition>
   </conditions>
   <plan-body>
      <user-performed/>
   </plan-body>
</plan>
<plan name="PLAN_8" title="limiting exposure to viral
   upper respiratory infections">
   <delta-link link-id="8"/>
   <conditions>
      <setup-precondition confirmation-required="yes">
         <none/>
      </setup-precondition>
   </conditions>
   <plan-body>
      <user-performed/>
   </plan-body>
</plan>
<plan name="PLAN_10" title="ensure immunizations are
   up-to-date; including influenza and Prevnar">
   <delta-link link-id="10"/>
   <delta-link link-id="5"/>
   <delta-link link-id="9"/>
   <explanation text="avoiding exposure to passive
      smokeavoid pacifier use beyond 10 months of age"/>
   <conditions>
```

```
            <setup-precondition confirmation-required="yes">
               <none/>
            </setup-precondition>
         </conditions>
         <plan-body>
            <user-performed/>
         </plan-body>
      </plan>
      <plan name="PLAN_11" title="Initiate Appropriate
         Treatment">
         <delta-link link-id="11"/>
         <conditions>
            <setup-precondition confirmation-required="yes">
               <none/>
            </setup-precondition>
         </conditions>
         <plan-body>
            <plan-activation>
               <plan-schema name="PLAN_12">
                  <delta-link link-id="12"/>
               </plan-schema>
            </plan-activation>
         </plan-body>
      </plan>
      <plan name="PLAN_PARENT_3" title="Initiate Appropriate
         Treatment">
         <conditions>
            <setup-precondition confirmation-required="yes">
               <none/>
            </setup-precondition>
         </conditions>
         <plan-body>
            <plan-activation>
               <plan-schema name="PLAN_PARENT_2">
                  <delta-link link-id="2"/>
               </plan-schema>
            </plan-activation>
         </plan-body>
      </plan>
      <plan name="PLAN_12" title="Treatment Options for Acute
         Otitis Media">
         <delta-link link-id="12"/>
         <conditions>
            <setup-precondition confirmation-required="yes">
               <none/>
            </setup-precondition>
         </conditions>
         <plan-body>
            <subplans retry-aborted-subplans="no" type="
               any-order" wait-for-optional-subplans="no">
               <wait-for>
                  <all/>
               </wait-for>
               <ask>
                  <parameter-ref name="for_Acute_Otitis_Media"
```

```
                    />
                <time-out>
                    <now/>
                </time-out>
            </ask>
            <if-then-else>
                <simple-condition>
                    <simple-condition-combination type="and">
                        <comparison type="equal">
                            <left-hand-side>
                                <parameter-ref name="
                                    for_Acute_Otitis_Media"/>
                            </left-hand-side>
                            <right-hand-side>
                                <qualitative-constant value="yes"
                                    />
                            </right-hand-side>
                        </comparison>
                    </simple-condition-combination>
                </simple-condition>
                <then-branch>
                    <plan-activation>
                        <plan-schema name="PARENT_PLAN_IN_12">
                            <delta-link link-id="12"/>
                        </plan-schema>
                    </plan-activation>
                </then-branch>
            </if-then-else>
        </subplans>
    </plan-body>
</plan>
<plan name="PARENT_PLAN_IN_12">
    <plan-body>
        <subplans type="unordered">
            <wait-for>
                <all/>
            </wait-for>
            <plan-activation>
                <plan-schema name="PLAN_13">
                    <delta-link link-id="13"/>
                </plan-schema>
            </plan-activation>
            <plan-activation>
                <plan-schema name="PLAN_14">
                    <delta-link link-id="14"/>
                </plan-schema>
            </plan-activation>
            <plan-activation>
                <plan-schema name="PLAN_37">
                    <delta-link link-id="37"/>
                </plan-schema>
            </plan-activation>
            <plan-activation>
                <plan-schema name="PLAN_40">
                    <delta-link link-id="40"/>
```

```
            </plan-schema>
         </plan-activation>
      </subplans>
   </plan-body>
</plan>
<plan name="PLAN_13" title="Observation of mildly
   symptomatic children is encouraged in the absence of
   risk factors.">
   <delta-link link-id="13"/>
   <conditions>
      <setup-precondition confirmation-required="yes">
         <none/>
      </setup-precondition>
   </conditions>
   <plan-body>
      <user-performed/>
   </plan-body>
</plan>
<plan name="PLAN_PARENT_2" title="Therapeutic (10 day)
   course of antibiotics.">
   <conditions>
      <setup-precondition confirmation-required="yes">
         <none/>
      </setup-precondition>
   </conditions>
   <plan-body>
      <plan-activation>
         <plan-schema name="PLAN_PARENT_1">
            <delta-link link-id="1"/>
         </plan-schema>
      </plan-activation>
   </plan-body>
</plan>
<plan name="PLAN_14" title="Therapeutic (10 day) course
   of antibiotics.">
   <delta-link link-id="14"/>
   <delta-link link-id="15"/>
   <delta-link link-id="34"/>
   <explanation text="Consideration may be given to a
      shortened course of antibiotics (5 days) for
      children who are at low risk (i.e., age &amp;gt; 2
      years, no history of chronic or recurrent otitis
      media and intact tympanic membranes). The use of
      nasal decongestants and corticosteroids is not
      supported in the literature."/>
   <conditions/>
   <plan-body>
      <subplans type="unordered">
         <wait-for>
            <all/>
         </wait-for>
         <plan-activation>
            <plan-schema name="PLAN_16">
               <delta-link link-id="16"/>
            </plan-schema>
```

```
            </plan-activation>
            <plan-activation>
               <plan-schema name="PLAN_19">
                  <delta-link link-id="19"/>
               </plan-schema>
            </plan-activation>
            <plan-activation>
               <plan-schema name="PLAN_28">
                  <delta-link link-id="28"/>
               </plan-schema>
            </plan-activation>
         </subplans>
      </plan-body>
   </plan>
   <plan name="PLAN_16" title="First-Line Medications">
      <delta-link link-id="16"/>
      <conditions>
         <setup-precondition confirmation-required="yes">
            <none/>
         </setup-precondition>
      </conditions>
      <plan-body>
         <subplans type="unordered">
            <wait-for>
               <all/>
            </wait-for>
            <plan-activation>
               <plan-schema name="PLAN_17">
                  <delta-link link-id="17"/>
               </plan-schema>
            </plan-activation>
            <plan-activation>
               <plan-schema name="PLAN_18">
                  <delta-link link-id="18"/>
               </plan-schema>
            </plan-activation>
         </subplans>
      </plan-body>
   </plan>
   <plan name="PLAN_17" title="amoxicillin (40 mg/kg/day) if
       low risk (&gt; 2 years, no day care, and no
      antibiotics for the past three months).">
      <delta-link link-id="17"/>
      <conditions>
         <setup-precondition confirmation-required="yes">
            <none/>
         </setup-precondition>
      </conditions>
      <plan-body>
         <subplans retry-aborted-subplans="no" type="
            any-order" wait-for-optional-subplans="no">
            <wait-for>
               <all/>
            </wait-for>
            <ask>
```

```
                    <parameter-ref name="if_low_risk_&amp;gt;
                        _years_no_day_care_and_no_antibiotics_for_the_past_thr
                        "/>
                    <time-out>
                        <now/>
                    </time-out>
                </ask>
                <if-then-else>
                    <simple-condition>
                        <simple-condition-combination type="and">
                            <comparison type="equal">
                                <left-hand-side>
                                    <parameter-ref name="if_low_risk_
                                        &amp;gt;
                                        _years_no_day_care_and_no_antibiotics_for_
                                        "/>
                                </left-hand-side>
                                <right-hand-side>
                                    <qualitative-constant value="yes"
                                        />
                                </right-hand-side>
                            </comparison>
                        </simple-condition-combination>
                    </simple-condition>
                    <then-branch>
                        <plan-activation>
                            <plan-schema name="
                                ATOMIC_PLAN_17_amoxicillin_40_mgkgday
                                ">
                                <delta-link link-id="17"/>
                            </plan-schema>
                        </plan-activation>
                    </then-branch>
                </if-then-else>
            </subplans>
        </plan-body>
    </plan>
    <plan name="ATOMIC_PLAN_17_amoxicillin_40_mgkgday" title=
        "amoxicillin: 40 mg/kg/day, ">
        <delta-link link-id="17"/>
        <conditions>
            <setup-precondition confirmation-required="yes">
                <none/>
            </setup-precondition>
    </conditions>
        <plan-body>
            <user-performed/>
        </plan-body>
    </plan>
    <plan name="PLAN_18" title="80 mg/kg/day if not low risk
        or for resistant AOM if the lower dose was used
        initially.">
        <delta-link link-id="18"/>
        <conditions>
            <setup-precondition confirmation-required="yes">
```

```
            <none/>
         </setup-precondition>
      </conditions>
      <plan-body>
         <subplans retry-aborted-subplans="no" type="
            any-order" wait-for-optional-subplans="no">
            <wait-for>
               <all/>
            </wait-for>
            <ask>
               <parameter-ref name="
                  if_not_low_risk_or_for_resistant_AOM_if_the_lower_dose
                  "/>
               <time-out>
                  <now/>
               </time-out>
            </ask>
            <if-then-else>
               <simple-condition>
                  <simple-condition-combination type="and">
                     <comparison type="equal">
                        <left-hand-side>
                           <parameter-ref name="
                              if_not_low_risk_or_for_resistant_AOM_if_th
                              "/>
                        </left-hand-side>
                        <right-hand-side>
                           <qualitative-constant value="yes"
                              />
                        </right-hand-side>
                     </comparison>
                  </simple-condition-combination>
               </simple-condition>
               <then-branch>
                  <plan-activation>
                     <plan-schema name="
                        ATOMIC_PLAN_18_amoxicillin_80_mgkgday
                        ">
                        <delta-link link-id="18"/>
                     </plan-schema>
                  </plan-activation>
               </then-branch>
            </if-then-else>
         </subplans>
      </plan-body>
   </plan>
   <plan name="ATOMIC_PLAN_18_amoxicillin_80_mgkgday" title=
      "amoxicillin: 80 mg/kg/day, ">
      <delta-link link-id="18"/>
      <conditions>
         <setup-precondition confirmation-required="yes">
            <none/>
         </setup-precondition>
   </conditions>
      <plan-body>
```

```
            <user-performed/>
         </plan-body>
      </plan>
      <plan name="PLAN_19" title="Recommended second-line
         medications include: (Check the health plan formulary
         listing for currently available medications.)">
         <delta-link link-id="19"/>
         <conditions>
            <setup-precondition confirmation-required="yes">
               <none/>
            </setup-precondition>
         </conditions>
         <plan-body>
            <plan-activation>
               <plan-schema name="PLAN_SELECT_0"/>
            </plan-activation>
         </plan-body>
      </plan>
      <plan name="PLAN_SELECT_0">
         <conditions>
            <complete-condition>
               <constraint-combination type="or">
                  <plan-state-constraint state="completed">
                     <plan-pointer>
                        <static-plan-pointer plan-name="PLAN_20
                           "/>
                     </plan-pointer>
                     <time-annotation>
                        <now/>
                     </time-annotation>
                  </plan-state-constraint>
                  <plan-state-constraint state="completed">
                     <plan-pointer>
                        <static-plan-pointer plan-name="PLAN_21
                           "/>
                     </plan-pointer>
                     <time-annotation>
                        <now/>
                     </time-annotation>
                  </plan-state-constraint>
                  <plan-state-constraint state="completed">
                     <plan-pointer>
                        <static-plan-pointer plan-name="PLAN_22
                           "/>
                     </plan-pointer>
                     <time-annotation>
                        <now/>
                     </time-annotation>
                  </plan-state-constraint>
                  <plan-state-constraint state="completed">
                     <plan-pointer>
                        <static-plan-pointer plan-name="PLAN_23
                           "/>
                     </plan-pointer>
                     <time-annotation>
```

```
                        <now/>
                    </time-annotation>
                </plan-state-constraint>
                <plan-state-constraint state="completed">
                    <plan-pointer>
                        <static-plan-pointer plan-name="PLAN_24
                            "/>
                    </plan-pointer>
                    <time-annotation>
                        <now/>
                    </time-annotation>
                </plan-state-constraint>
                <plan-state-constraint state="completed">
                    <plan-pointer>
                        <static-plan-pointer plan-name="PLAN_25
                            "/>
                    </plan-pointer>
                    <time-annotation>
                        <now/>
                    </time-annotation>
                </plan-state-constraint>
                <plan-state-constraint state="completed">
                    <plan-pointer>
                        <static-plan-pointer plan-name="PLAN_26
                            "/>
                    </plan-pointer>
                    <time-annotation>
                        <now/>
                    </time-annotation>
                </plan-state-constraint>
                <plan-state-constraint state="completed">
                    <plan-pointer>
                        <static-plan-pointer plan-name="PLAN_27
                            "/>
                    </plan-pointer>
                    <time-annotation>
                        <now/>
                    </time-annotation>
                </plan-state-constraint>
            </constraint-combination>
        </complete-condition>
    </conditions>
    <plan-body>
        <subplans type="any-order">
            <wait-for>
                <all/>
            </wait-for>
            <plan-activation>
                <plan-schema name="PLAN_20">
                    <delta-link link-id="20"/>
                </plan-schema>
            </plan-activation>
            <plan-activation>
                <plan-schema name="PLAN_21">
                    <delta-link link-id="21"/>
```

```
                </plan-schema>
            </plan-activation>
            <plan-activation>
                <plan-schema name="PLAN_22">
                    <delta-link link-id="22"/>
                </plan-schema>
            </plan-activation>
            <plan-activation>
                <plan-schema name="PLAN_23">
                    <delta-link link-id="23"/>
                </plan-schema>
            </plan-activation>
            <plan-activation>
                <plan-schema name="PLAN_24">
                    <delta-link link-id="24"/>
                </plan-schema>
            </plan-activation>
            <plan-activation>
                <plan-schema name="PLAN_25">
                    <delta-link link-id="25"/>
                </plan-schema>
            </plan-activation>
            <plan-activation>
                <plan-schema name="PLAN_26">
                    <delta-link link-id="26"/>
                </plan-schema>
            </plan-activation>
            <plan-activation>
                <plan-schema name="PLAN_27">
                    <delta-link link-id="27"/>
                </plan-schema>
            </plan-activation>
        </subplans>
    </plan-body>
</plan>
<plan name="PLAN_20" title="amoxicillin/clavulanate
    potassium (Augmentin)">
    <delta-link link-id="20"/>
    <conditions>
        <setup-precondition confirmation-required="yes">
            <none/>
        </setup-precondition>
    </conditions>
    <plan-body>
        <user-performed/>
    </plan-body>
</plan>
<plan name="PLAN_21" title="cefuroxime axetil (Ceftin)">
    <delta-link link-id="21"/>
    <conditions>
        <setup-precondition confirmation-required="yes">
            <none/>
        </setup-precondition>
    </conditions>
    <plan-body>
```

```
            <user-performed/>
         </plan-body>
      </plan>
      <plan name="PLAN_22" title="ceftriaxone sodium (Rocephin)
          : prescribe one dose for new onset otitis media and a
          three-day course for a truly resistant pattern of
          otitis media or if oral treatment cannot be given.">
         <delta-link link-id="22"/>
         <conditions>
            <setup-precondition confirmation-required="yes">
               <none/>
            </setup-precondition>
         </conditions>
         <plan-body>
            <user-performed/>
         </plan-body>
      </plan>
      <plan name="PLAN_23" title="cefprozil (Cefzil)">
         <delta-link link-id="23"/>
         <conditions>
            <setup-precondition confirmation-required="yes">
               <none/>
            </setup-precondition>
         </conditions>
         <plan-body>
            <user-performed/>
         </plan-body>
      </plan>
      <plan name="PLAN_24" title="loracarbef (Lorabid)">
         <delta-link link-id="24"/>
         <conditions>
            <setup-precondition confirmation-required="yes">
               <none/>
            </setup-precondition>
         </conditions>
         <plan-body>
            <user-performed/>
         </plan-body>
      </plan>
      <plan name="PLAN_25" title="cefdinir (Omnicef)">
         <delta-link link-id="25"/>
         <conditions>
            <setup-precondition confirmation-required="yes">
               <none/>
            </setup-precondition>
         </conditions>
         <plan-body>
            <user-performed/>
         </plan-body>
      </plan>
      <plan name="PLAN_26" title="cefixime (Suprax)">
         <delta-link link-id="26"/>
         <conditions>
            <setup-precondition confirmation-required="yes">
               <none/>
```

```
            </setup-precondition>
         </conditions>
         <plan-body>
            <user-performed/>
         </plan-body>
      </plan>
      <plan name="PLAN_27" title="cefpodoxime proxetil (Vantin)
          ">
         <delta-link link-id="27"/>
         <conditions>
            <setup-precondition confirmation-required="yes">
               <none/>
            </setup-precondition>
         </conditions>
         <plan-body>
            <user-performed/>
         </plan-body>
      </plan>
      <plan name="PLAN_28" title="Second-line medications that
          are currently used but are not as strongly supported
          in the literature are listed below.">
         <delta-link link-id="28"/>
         <delta-link link-id="29"/>
         <explanation text="These medications are not
             recommended when the patient has failed a course of
              amoxicillin."/>
         <conditions>
            <setup-precondition confirmation-required="yes">
               <none/>
            </setup-precondition>
         </conditions>
         <plan-body>
            <user-performed/>
         </plan-body>
      </plan>
      <plan name="PLAN_SELECT_1">
         <conditions>
            <complete-condition>
               <constraint-combination type="or">
                  <plan-state-constraint state="completed">
                     <plan-pointer>
                        <static-plan-pointer plan-name="PLAN_30
                            "/>
                     </plan-pointer>
                     <time-annotation>
                        <now/>
                     </time-annotation>
                  </plan-state-constraint>
                  <plan-state-constraint state="completed">
                     <plan-pointer>
                        <static-plan-pointer plan-name="PLAN_31
                            "/>
                     </plan-pointer>
                     <time-annotation>
                        <now/>
```

```
                    </time-annotation>
                </plan-state-constraint>
                <plan-state-constraint state="completed">
                    <plan-pointer>
                        <static-plan-pointer plan-name="PLAN_32
                            "/>
                    </plan-pointer>
                    <time-annotation>
                        <now/>
                    </time-annotation>
                </plan-state-constraint>
                <plan-state-constraint state="completed">
                    <plan-pointer>
                        <static-plan-pointer plan-name="PLAN_33
                            "/>
                    </plan-pointer>
                    <time-annotation>
                        <now/>
                    </time-annotation>
                </plan-state-constraint>
            </constraint-combination>
        </complete-condition>
    </conditions>
    <plan-body>
        <subplans type="any-order">
            <wait-for>
                <all/>
            </wait-for>
            <plan-activation>
                <plan-schema name="PLAN_30">
                    <delta-link link-id="30"/>
                </plan-schema>
            </plan-activation>
            <plan-activation>
                <plan-schema name="PLAN_31">
                    <delta-link link-id="31"/>
                </plan-schema>
            </plan-activation>
            <plan-activation>
                <plan-schema name="PLAN_32">
                    <delta-link link-id="32"/>
                </plan-schema>
            </plan-activation>
            <plan-activation>
                <plan-schema name="PLAN_33">
                    <delta-link link-id="33"/>
                </plan-schema>
            </plan-activation>
        </subplans>
    </plan-body>
</plan>
<plan name="PLAN_30" title="trimethoprim sulfa (Bactrim,
    Septra)">
    <delta-link link-id="30"/>
    <conditions>
```

```
            <setup-precondition confirmation-required="yes">
               <none/>
            </setup-precondition>
         </conditions>
         <plan-body>
            <user-performed/>
         </plan-body>
      </plan>
      <plan name="PLAN_31" title="clarithromycin (Biaxin)">
         <delta-link link-id="31"/>
         <conditions>
            <setup-precondition confirmation-required="yes">
               <none/>
            </setup-precondition>
         </conditions>
         <plan-body>
            <user-performed/>
         </plan-body>
      </plan>
      <plan name="PLAN_32" title="erythromycin ethylsuccinate
          and sulfisoxazole acetyl (Pediazole)">
         <delta-link link-id="32"/>
         <conditions/>
         <plan-body>
            <subplans type="any-order">
               <wait-for>
                  <one/>
               </wait-for>
               <plan-activation>
                  <plan-schema name="
                     ATOMIC_PLAN_32_erythromycin_ethylsuccinate
                     ">
                     <delta-link link-id="32"/>
                  </plan-schema>
               </plan-activation>
               <plan-activation>
                  <plan-schema name="
                     ATOMIC_PLAN_32_sulfisoxazole">
                     <delta-link link-id="32"/>
                  </plan-schema>
               </plan-activation>
            </subplans>
         </plan-body>
      </plan>
      <plan name="ATOMIC_PLAN_32_erythromycin_ethylsuccinate"
          title="erythromycin ethylsuccinate">
         <delta-link link-id="32"/>
         <conditions>
            <setup-precondition confirmation-required="yes">
               <none/>
            </setup-precondition>
         </conditions>
         <plan-body>
            <user-performed/>
         </plan-body>
```

```
</plan>
<plan name="ATOMIC_PLAN_32_sulfisoxazole" title="
    sulfisoxazole">
   <delta-link link-id="32"/>
   <conditions>
      <setup-precondition confirmation-required="yes">
         <none/>
      </setup-precondition>
   </conditions>
   <plan-body>
      <user-performed/>
   </plan-body>
</plan>
<plan name="PLAN_33" title="azithromycin (Zithromax)">
   <delta-link link-id="33"/>
   <conditions>
      <setup-precondition confirmation-required="yes">
         <none/>
      </setup-precondition>
   </conditions>
   <plan-body>
      <user-performed/>
   </plan-body>
</plan>
<plan name="PLAN_PARENT_1" title="Treatment of Resistant
    Acute Otitis Media">
   <conditions>
      <setup-precondition confirmation-required="yes">
         <none/>
      </setup-precondition>
   </conditions>
   <plan-body>
      <subplans type="unordered">
         <wait-for>
            <all/>
         </wait-for>
         <plan-activation>
            <plan-schema name="PLAN_35">
               <delta-link link-id="35"/>
            </plan-schema>
         </plan-activation>
         <plan-activation>
            <plan-schema name="PLAN_PARENT_4">
               <delta-link link-id="4"/>
            </plan-schema>
         </plan-activation>
      </subplans>
   </plan-body>
</plan>
<plan name="PLAN_35" title="A second antibiotic should be
     chosen; the alternative first-line medication may be
    an appropriate choice.">
   <delta-link link-id="35"/>
   <conditions>
      <setup-precondition confirmation-required="yes">
```

```
            <none/>
         </setup-precondition>
      </conditions>
      <plan-body>
         <user-performed/>
      </plan-body>
   </plan>
   <plan name="PLAN_PARENT_4" title="Treatment of Persistent
       Acute Otitis Media">
      <conditions>
         <setup-precondition confirmation-required="yes">
            <none/>
         </setup-precondition>
      </conditions>
      <plan-body>
         <plan-activation>
            <plan-schema name="PLAN_36">
               <delta-link link-id="36"/>
            </plan-schema>
         </plan-activation>
      </plan-body>
   </plan>
   <plan name="PLAN_36" title="A second course of therapy
      with a different antibiotic is indicated for
      persistent AOM.">
      <delta-link link-id="36"/>
      <conditions>
         <setup-precondition confirmation-required="yes">
            <none/>
         </setup-precondition>
      </conditions>
      <plan-body>
         <subplans retry-aborted-subplans="no" type="
             any-order" wait-for-optional-subplans="no">
            <wait-for>
               <all/>
            </wait-for>
            <ask>
               <parameter-ref name="for_persistent_AOM"/>
               <time-out>
                  <now/>
               </time-out>
            </ask>
            <if-then-else>
               <simple-condition>
                  <simple-condition-combination type="and">
                     <comparison type="equal">
                        <left-hand-side>
                           <parameter-ref name="
                              for_persistent_AOM"/>
                        </left-hand-side>
                        <right-hand-side>
                           <qualitative-constant value="yes"
                              />
                        </right-hand-side>
```

```
                        </comparison>
                    </simple-condition-combination>
                </simple-condition>
                <then-branch>
                    <plan-activation>
                        <plan-schema name="
                            ATOMIC_PLAN_36_antibiotic">
                            <delta-link link-id="36"/>
                        </plan-schema>
                    </plan-activation>
                </then-branch>
            </if-then-else>
        </subplans>
    </plan-body>
</plan>
<plan name="ATOMIC_PLAN_36_antibiotic" title="antibiotic"
    >
    <delta-link link-id="36"/>
    <conditions>
        <setup-precondition confirmation-required="yes">
            <none/>
        </setup-precondition>
    </conditions>
    <plan-body>
        <user-performed/>
    </plan-body>
</plan>
<plan name="PLAN_37" title="Consider Prophylactic Regimen
    ">
    <delta-link link-id="37"/>
    <conditions>
        <setup-precondition confirmation-required="yes">
            <none/>
        </setup-precondition>
    </conditions>
    <plan-body>
        <plan-activation>
            <plan-schema name="PLAN_38">
                <delta-link link-id="38"/>
            </plan-schema>
        </plan-activation>
    </plan-body>
</plan>
<plan name="PLAN_38" title="amoxicillin (20 mg/kg QD [
    once a day])">
    <delta-link link-id="38"/>
    <delta-link link-id="39"/>
    <explanation text="The usual duration of antibiotic
        prophylaxis is 2 to 6 months."/>
    <conditions/>
    <plan-body>
        <plan-activation>
            <plan-schema name="
                ATOMIC_PLAN_38_amoxicillin_20_mgkg_QD">
                <delta-link link-id="38"/>
```

```
                        <time-annotation>
                            <time-range>
                                <duration>
                                    <minimum>
                                        <numerical-constant value="2"
                                            unit="m"/>
                                    </minimum>
                                    <maximum>
                                        <numerical-constant value="6"
                                            unit="m"/>
                                    </maximum>
                                </duration>
                            </time-range>
                            <self/>
                        </time-annotation>
                    </plan-schema>
                </plan-activation>
            </plan-body>
        </plan>
        <plan name="ATOMIC_PLAN_38_amoxicillin_20_mgkg_QD" title=
            "amoxicillin: 20 mg/kg, QD">
            <delta-link link-id="38"/>
            <conditions>
                <setup-precondition confirmation-required="yes">
                    <none/>
                </setup-precondition>
            </conditions>
            <plan-body>
                <cyclical-plan>
                    <start-time>
                        <cyclical-time-annotation>
                            <time-range/>
                            <set-of-cyclical-time-points>
                                <time-point>
                                    <now/>
                                </time-point>
                                <offset>
                                    <now/>
                                </offset>
                                <frequency>
                                    <now/>
                                </frequency>
                            </set-of-cyclical-time-points>
                        </cyclical-time-annotation>
                    </start-time>
                    <cyclical-plan-body>
                        <plan-activation>
                            <plan-schema name="
                                CYCLICAL_PLAN_amoxicillin_20_mgkg">
                                <delta-link link-id=""/>
                            </plan-schema>
                        </plan-activation>
                    </cyclical-plan-body>
                    <cyclical-time-annotation>
                        <time-range>
```

```
                        <starting-shift>
                           <minimum>
                              <numerical-constant value="4" unit="
                                 h"/>
                           </minimum>
                           <maximum>
                              <numerical-constant value="20" unit=
                                 "h"/>
                           </maximum>
                        </starting-shift>
                     </time-range>
                     <set-of-cyclical-time-points>
                        <time-point>
                           <numerical-constant value="0"/>
                        </time-point>
                        <offset>
                           <numerical-constant value="0"/>
                        </offset>
                        <frequency>
                           <numerical-constant value="1" unit="d"
                              />
                        </frequency>
                     </set-of-cyclical-time-points>
                  </cyclical-time-annotation>
               </cyclical-plan>
            </plan-body>
         </plan>
         <plan name="CYCLICAL_PLAN_amoxicillin_20_mgkg" title="
            amoxicillin: 20 mg/kg">
            <delta-link link-id="38"/>
            <plan-body>
               <user-performed/>
            </plan-body>
         </plan>
         <plan name="PLAN_40" title="Schedule Follow-Up in 3 to 4
            Weeks">
            <delta-link link-id="40"/>
            <conditions>
               <setup-precondition confirmation-required="yes">
                  <none/>
               </setup-precondition>
            </conditions>
            <plan-body>
               <plan-activation>
                  <plan-schema name="PLAN_41">
                     <delta-link link-id="41"/>
                  </plan-schema>
               </plan-activation>
            </plan-body>
         </plan>
         <plan name="PLAN_41" title="Recheck in 3 to 4 weeks or at
             next well child visit if within the next 4 to 6 weeks
            .">
            <delta-link link-id="41"/>
            <conditions>
```

```
            <setup-precondition confirmation-required="yes">
               <none/>
            </setup-precondition>
         </conditions>
         <plan-body>
            <user-performed/>
         </plan-body>
      </plan>
      <plan name="PLAN_PARENT_6" title="Otitis Media with
          Effusion Algorithm Annotations">
         <conditions>
            <setup-precondition confirmation-required="yes">
               <none/>
            </setup-precondition>
         </conditions>
         <plan-body>
            <subplans type="unordered">
               <wait-for>
                  <all/>
               </wait-for>
               <plan-activation>
                  <plan-schema name="PLAN_PARENT_5">
                     <delta-link link-id="5"/>
                  </plan-schema>
               </plan-activation>
               <plan-activation>
                  <plan-schema name="PLAN_48">
                     <delta-link link-id="48"/>
                  </plan-schema>
               </plan-activation>
            </subplans>
         </plan-body>
      </plan>
      <plan name="PLAN_PARENT_5" title="Consider Treatment
          Options">
         <conditions>
            <setup-precondition confirmation-required="yes">
               <none/>
            </setup-precondition>
         </conditions>
         <plan-body>
            <subplans type="unordered">
               <wait-for>
                  <all/>
               </wait-for>
               <plan-activation>
                  <plan-schema name="PLAN_42">
                     <delta-link link-id="42"/>
                  </plan-schema>
               </plan-activation>
               <plan-activation>
                  <plan-schema name="PLAN_45">
                     <delta-link link-id="45"/>
                  </plan-schema>
               </plan-activation>
```

```
      </subplans>
   </plan-body>
</plan>
<plan name="PLAN_42" title="Observe--rechecking in 4 to 6
    weeks.">
   <delta-link link-id="42"/>
   <conditions>
      <setup-precondition confirmation-required="yes">
         <none/>
      </setup-precondition>
   </conditions>
   <plan-body>
      <plan-activation>
         <plan-schema name="PLAN_43">
            <delta-link link-id="43"/>
         </plan-schema>
      </plan-activation>
   </plan-body>
</plan>
<plan name="PLAN_43" title="Course of antibiotics should
   be given as a trial prior to referral for ventilating
   tubes.">
   <delta-link link-id="43"/>
   <delta-link link-id="44"/>
   <explanation text="Ten-day course of antibiotics using
       first- and second-line criteria."/>
   <conditions>
      <setup-precondition confirmation-required="yes">
         <none/>
      </setup-precondition>
   </conditions>
   <plan-body>
      <plan-activation>
         <plan-schema name="ATOMIC_PLAN_43_antibiotic">
            <delta-link link-id="43"/>
            <time-annotation>
               <time-range>
                  <duration>
                     <minimum>
                        <numerical-constant value="10"
                           unit="d"/>
                     </minimum>
                     <maximum>
                        <numerical-constant value="10"
                           unit="d"/>
                     </maximum>
                  </duration>
               </time-range>
               <self/>
            </time-annotation>
         </plan-schema>
      </plan-activation>
   </plan-body>
</plan>
<plan name="ATOMIC_PLAN_43_antibiotic" title="antibiotic"
```

```
    >
    <delta-link link-id="43"/>
    <conditions>
       <setup-precondition confirmation-required="yes">
          <none/>
       </setup-precondition>
    </conditions>
    <plan-body>
       <user-performed/>
    </plan-body>
 </plan>
 <plan name="PLAN_45" title="Referral for ventilating
    tubes if patient meets ENT referral criteria.">
    <delta-link link-id="45"/>
    <delta-link link-id="49"/>
    <explanation text="More frequent rechecking than every
        4 to 6 weeks of OME is unnecessary and
       inappropriate."/>
    <conditions>
       <setup-precondition confirmation-required="yes">
          <none/>
       </setup-precondition>
    </conditions>
    <plan-body>
       <subplans retry-aborted-subplans="no" type="
          any-order" wait-for-optional-subplans="no">
          <wait-for>
             <all/>
          </wait-for>
          <ask>
             <parameter-ref name="
                if_patient_meets_ENT_referral_criteria"/>
             <time-out>
                <now/>
             </time-out>
          </ask>
          <if-then-else>
             <simple-condition>
                <simple-condition-combination type="and">
                   <comparison type="equal">
                      <left-hand-side>
                         <parameter-ref name="
                            if_patient_meets_ENT_referral_criteria
                            "/>
                      </left-hand-side>
                      <right-hand-side>
                         <qualitative-constant value="yes"
                            />
                      </right-hand-side>
                   </comparison>
                </simple-condition-combination>
             </simple-condition>
             <then-branch>
                <plan-activation>
                   <plan-schema name="PLAN_46">
```

```
                        <delta-link link-id="46"/>
                    </plan-schema>
                </plan-activation>
            </then-branch>
        </if-then-else>
    </subplans>
</plan-body>
</plan>
<plan name="PLAN_46" title="Course of antibiotics should
    be given as a trial prior to referral for ventilating
    tubes.">
    <delta-link link-id="46"/>
    <delta-link link-id="47"/>
    <explanation text="Ten-day course of antibiotics using
        first- and second-line criteria."/>
    <conditions>
        <setup-precondition confirmation-required="yes">
            <none/>
        </setup-precondition>
    </conditions>
    <plan-body>
        <plan-activation>
            <plan-schema name="ATOMIC_PLAN_46_antibiotics">
                <delta-link link-id="46"/>
                <time-annotation>
                    <time-range>
                        <duration>
                            <minimum>
                                <numerical-constant value="10"
                                    unit="d"/>
                            </minimum>
                            <maximum>
                                <numerical-constant value="10"
                                    unit="d"/>
                            </maximum>
                        </duration>
                    </time-range>
                    <self/>
                </time-annotation>
            </plan-schema>
        </plan-activation>
    </plan-body>
</plan>
<plan name="ATOMIC_PLAN_46_antibiotics" title="
    antibiotics">
    <delta-link link-id="46"/>
    <conditions>
        <setup-precondition confirmation-required="yes">
            <none/>
        </setup-precondition>
    </conditions>
    <plan-body>
        <user-performed/>
    </plan-body>
</plan>
```

```
            <plan name="PLAN_48" title="Follow-up 4 to 6 Weeks">
               <delta-link link-id="48"/>
               <delta-link link-id="49"/>
               <explanation text="More frequent rechecking than every
                   4 to 6 weeks of OME is unnecessary and
                  inappropriate."/>
               <conditions>
                  <setup-precondition confirmation-required="yes">
                     <none/>
                  </setup-precondition>
               </conditions>
               <plan-body>
                  <user-performed/>
               </plan-body>
            </plan>
         </plan-group>
      </plans>
   </plan-library>
```

List of Figures

List of Tables

Listings

Bibliography

[1] Proforma tasks: four types (sub-classes) of task are defined in a class hierarchy. http://www.acl.icnet.uk/lab/proformaspec.html, 2003. [retrieved: Oct. 20, 2005].

[2] Gem cutter: Screenshot. http://gem.med.yale.edu/GEM_Cutter/gem_cutter.htm, 2005. [retrieved: Oct. 20, 2005].

[3] Gem hierarchy. http://gem.med.yale.edu/Hierarchy/hierarchy.htm, 2005. [retrieved: Oct. 20, 2005].

[4] Eugene Agichtein and Luis Gravano. Snowball: Extracting relations from large plain-text collections. In *Proceedings of the 5th ACM International Conference on Digital Libraries*, 2000.

[5] James F. Allen. Maintaining knowledge about temporal intervals. *Communications of the ACM*, 26(11), November 1983.

[6] James F. Allen. Towards a general theory of action and time. *Artificial Intelligence*, 23(2):123–154, July 1984.

[7] James F. Allen and Pat J. Hayes. Moments and points in an interval-based temporal logic. *Computational Intelligence*, 5:225–238, January 1990.

[8] Douglas E. Appelt. Introduction to information extraction. *AI Communications*, 12:161–172, 1999.

[9] Douglas E. Appelt, Jerry Hobbs, John Bear, David Israel, and Mabry Tyson. FASTUS: A finite-state processor for Information Extraction from real world text. In *Proceedings of the 13th International Joint Conference on Artificial Intelligence*, pages 1172–1178, 1993.

[10] ASTM International. *E2210-02 Standard Specification for Guideline Elements Model (GEM)-Document Model for Clinical Practice Guidelines*, 2003.

[11] Damaris Ayuso, Sean Boisen, Heidi Fox, Herb Gish, Robert Ingria, and Ralph Weischedel. BBN: Description of the PLUM system as used for MUC-4. In *Proceedings of the Fourth Message Understanding Conference (MUC-4)*, pages 169–176, 1992.

[12] Fahiem Bacchus, Josh Tennenberg, and Johannes Koomen. A non-reified temporal logic. In James F. Allen, Richard Fikes, and Erik Sandewall, editors, *KR'89: Principles of Knowledge Representation and Reasoning*, pages 2–10, San Mateo, California, 1989. Morgan Kaufmann.

[13] M. Balser, O. Coltell, J. van Croonenborg, C. Duelli, F. van Harmelen, A. Jovell, P. Lucas, M. Marcos, S. Miksch, W. Reif, K. Rosenbrand, A. Seyfang, and A. ten Teije. Protocure: Integrating formal methods in the development process of medical guidelines and protocols. In Katharina Kaiser, Silvia Miksch, and Samson W. Tu, editors, *Computer-based Support for Clinical Guidelines and Protocols. Proceedings of the Symposium on Computerized Guidelines and Protocols (CGP 2004)*, volume 101: Studies in Health Technology and Informatics, pages 103–107, Prague, Czech Republic, 2004. IOS Press.

[14] Robert Baumgartner, Sergio Flesca, and Georg Gottlob. Visual Web Information Extraction with Lixto. In *Proceedings of the Conference on Very Large Databases (VLDB)*, 2001.

[15] Avrim Blum and Tom Mitchell. Combining labeled and unlabeled data with co-training. In *Proceedings of the Eleventh Annual Conference on Computational Learning Theory (COLT)*, pages 92–100. ACM Press, 1998.

[16] A. Bochman. Concerted instance-interval temporal semantics: Temporal ontologies. *Notre Dame Journal of Formal Logic*, 31(3):403–414, 1990.

[17] Sergey Brin. Extracting patterns and relations from the world wide web. In *WebDB Workshop at 6th International Conference on Extended Database Technology, EDBT'98*, 1998.

[18] Mary Elaine Califf and Raymond J. Mooney. Relational learning of pattern-match rules for Information Extraction. In *Proceedings of the 16th National Conference on Artificial Intelligence (AAAI-99)*, pages 328–334, Orlando, FL, July 1999.

[19] James R. Campbell, Samson W. Tu, James G. Mansfield, Julie I. Boyer, James McClay, Craig Parker, Prabhu Ram, Sidna M. Scheitel, and Kevin McDonald. The SAGE guideline model: A knowledge representation framework for encoding interoperable clinical practice guidelines. In *Proceedings of the AMIA Annual Symposium*, November 2003.

[20] Chia-Hui Chang, Chun-Nan Hsu, and Shao-Cheng Lui. Automatic information extraction from semi-structured web pages by pattern discovery. *Decision Support Systems, Special Issue on Web Retrieval and Mining*, 35(1):129–147, April 2003.

[21] Boris Chidlovskii. Automatic repairing of web wrappers. In *Proceeding of the Third International Workshop on Web Information and Data Management*, pages 24–30, Atlanta, Georgia, USA, 2001.

[22] Michael Collins and Yoram Singer. Unsupervised models for named entity classification. In *Proceedings of the Joint SIGDAT Conference on Empirical Methods in Natural Language Processing*, pages 100–110, 1999.

[23] Valter Crescenzi, Giansalvatore Mecca, and Paolo Merialdo. Automatic web information extraction in the RoadRunner system. In *International Workshop on Data Semantics in Web Information Systems (DASWIS-2001) in conjunction with 20th International Conference on Conceptual Modeling (ER 2001)*, 2001.

[24] Valter Crescenzi, Giansalvatore Mecca, and Paolo Merialdo. Roadrunner: Towards automatic data extraction from large web sites. In *Proceedings of the Conference on Very Large Databases (VLDB'01)*, pages 109–118, 2001.

[25] Hamish Cunningham, Diana Maynard, Kalina Bontcheva, and Valentin Tablan. GATE: A framework and graphical development environment for robust NLP tools and applications. In *Proceedings of the 40th Anniversary Meeting of the Association for Computational Linguistics (ACL'02)*, Philadelphia, July 2002.

[26] Robert B. Doorenbos, Oren Etzioni, and Daniel S. Weld. A scalable comparison-shopping agent for the world-wide web. In W. Lewis Johnson and Barbara Hayes-Roth, editors, *Proceedings of the First International Conference on Autonomous Agents (Agents'97)*, pages 39–48, Marina del Rey, CA, USA, 1997. ACM Press.

[27] Georg Duftschmid. *Knowledge-Based Verification of Clinical Guidelines by Detection of Anomalies*. PhD thesis, Vienna University of Technology, Vienna, Austria, Jan. 1999.

[28] Peter L. Elkin, Mor Peleg, Ronilda Lacson, Elmer Bernstam, Samson Tu, Aziz Boxwala, Robert Greenes, and Edward H. Shortliffe. Toward Standardization of Electronic Guideline Representation. *MD Computing*, 17(6):39–44, 2000.

[29] M.J. Field and K.H. Lohr, editors. *Clinical Practice Guidelines: Directions for a New Program.* National Academy Press, Institute of Medicine, Washington DC, 1990.

[30] John Fox, Nicky Johns, and Ali Rahmanzadeh. Disseminating medical knowledge: The PROforma approach. *Artificial Intelligence in Medicine*, 14(1-2):157–181, 1998.

[31] Christian Freksa. Temporal reasoning based on semi-intervals. *Artificial Intelligence*, 54(1):199–207, March 1992.

[32] John H. Gennari, Mark A. Musen, Ray W. Fergerson, William E. Grosso, Monica Crubézy, Henrik Eriksson, Natalya F. Noy, and Samson W. Tu. The Evolution of Protégé: An Environment for Knowledge-based Systems Development. *International Journal of Human Computer Studies*, 58(1):89–123, 2003.

[33] Colin Gordon, Ian Herbert, and Peter Johnson. Knowledge representation and Clinical Practice Guidelines: the DILEMMA and PRESTIGE Projects. In *Medical Informatics Europe (MIE'96)*, Copenhagen, 1996.

[34] Jerry R. Hobbs, Douglas Appelt, Mabry Tyson, John Bear, and David Islael. SRI International: Description of the FASTUS system used for MUC-4. In *Proceedings fo the 4th Message Understanding Conference (MUC-4)*, pages 268–275, 1992.

[35] George Hripcsak, Paul B. Clayton, T. Allan Pryor, Peter Haug, Ove B. Wigertz, and Johan van der Lei. The Arden Syntax for medical logic modules. In R. A. Miller, editor, *Proc. of the 14th Annual Symposium on Computer Applications in Medical Care*, pages 200–204, Washington, D.C. New York, 1990. IEEE Computer Society Press.

[36] Chun-Nan Hsu and Ming-Tzung Dung. Wrapping semistructured web pages with finite-state transducers. In *Proceedings of the Conference on Automatic Learning and Discovery*, 1998.

[37] Scott B. Huffman. Learning information extraction patterns from examples. In *Lecture Notes in Computer Science. Connectionist, Statistical, and Symbolic Approaches to Learning for Natural Language Processing*, volume 1040, pages 246–260, London, UK, 1996. Springer Verlag.

[38] Mary E. Johnston, Karl B. Langton, R. Brian Haynes, and Alix Mathieu. Effects of computer-based clinical decision support systems on clinician performance and patient outcome: a critical appraisal of research. *Ann Intern Med*, 120(2):135–142, 1994.

[39] Rosie Jones, Andrew McCallum, Kamal Nigam, and Ellen Riloff. Bootstrapping for text learning tasks. In *Workshop on Text Mining: Foundations, Techniques and Applications at the International Joint Conference on Artificial Intelligence (IJCAI-99)*, pages 52–63, 1999.

[40] Katharina Kaiser. Knowledge-based methods for facilitating the creation of Asbru protocols. In Joerg Evermann, Eva Söderström, and Julia Kotlarsky, editors, *10th Doctoral Consortium on Advanced Information Systems Engineering (at CAiSE'03)*, pages 101–111, Velden, Austria, June 2003. University of Skövde.

[41] Katharina Kaiser. Semi-automatic transformation of structured guideline components into formal process representations. In *1st Doctoral Consortium at the Conference on Artificial Intelligence in Medicine (AIME)*, Aberdeen, UK, 2005.

[42] Katharina Kaiser, Cem Akkaya, and Silvia Miksch. Gaining process information of clinical practice guidelines using information extraction. In *Proceedings of the 10th Conference on Artificial Intelligence in Medicine (AIME), Lecture Notes in Computer Science (LNCS 3581)*, pages 181–190, Aberdeen, UK, 2005. Springer Verlag.

[43] Katharina Kaiser and Silvia Miksch. Temporal plan interdependencies. In R. López de Mántaras and Lorenza Saitta, editors, *Proc. of the 16th European Conference on Artificial Intelligence (ECAI 2004)*, pages 1039–1040, Valencia, Spain, 2004. IOS Press.

[44] Katharina Kaiser and Silvia Miksch. TimeWrap – a method for automatic transformation of structured guideline components into formal process-representations. In Katharina Kaiser, Silvia Miksch, and Samson Tu, editors, *Computer-based Support for Clinical Guidelines and Protocols. Proc. of the Symposium on Computerized Guidelines and Protocols (CGP 2004)*, volume 101: Studies in Health Technology and Informatics, pages 61–74, Prague, Czech Republic, 2004. IOS Press.

[45] Henry A. Kautz and Peter B. Ladkin. Integrating metric and qualitative temporal reasoning. In Thomas Dean and Kathleen McKeown, editors, *Proceedings of the Ninth National Conference on Artificial Intelligence (AAAI'91)*, pages 241–246, Menlo Park, California, 1991. AAAI, AAAI Press.

[46] Jun-Tae Kim and Dan I. Moldovan. Acquisition of linguistic patterns for knowledge-based information extraction. *IEEE Transactions on Knowledge and Data Engineering*, 7(5):713–724, October 1995.

[47] Lina T. Kohn, Janet M. Corrigan, and Molla S. Donaldson, editors. *To Err Is Human: Building a Safer Health System*. National Academy Press, Washington, D.C., 2000.

[48] Robert Kosara and Silvia Miksch. Visualization techniques for time-oriented, skeletal plans in medical therapy planning. In Werner Horn, Yuval Shahar, G. Lindberg, S. Andreassen, and J. Wyatt, editors, *Proceedings of the Joint European Conference on Artificial Intelligence in Medicine and Medical Decision Making (AIMDM'99)*, pages 291–300, Aalborg, Denmark, 1999. Springer Verlag.

[49] Robert Kosara and Silvia Miksch. Metaphors of Movement: A Visualization and User Interface for Time-Oriented, Skeletal Plans. *Artificial Intelligence in Medicine, Special Issue: Information Visualization in Medicine*, 22(2):111–131, May 2001.

[50] Robert Kosara, Silvia Miksch, and Helwig Hauser. Semantic depth of field. In *IEEE Symposium on Information Visualization (InfoVis 2001)*, pages 97–104. IEEE Computer Society Press, 2001.

[51] Robert Kosara, Silvia Miksch, Andreas Seyfang, and Peter Votruba. Tools for acquiring clinical guidelines in Asbru. In *Proc. of the Sixth World Conference on Integrate Design and Process Technology (IDPT'02)*, 2002.

[52] Robert Kosara, Silvia Miksch, Yuval Shahar, and Peter Johnson. AsbruView: Capturing complex, time-oriented plans — beyond flow-charts. In *Thinking with Diagrams 98*, 1998.

[53] George Krupka, Paul Jacobs, Lisa Rau, Lois Childs, and Ira Sider. GE NLTOOLSET: Description of the system as used for MUC-4. In *Proceedings of the 4th Message Understanding Conference (MUC-4)*, pages 177–185, 1992.

[54] Nick Kushmerick. *Wrapper Induction for Information Extraction*. PhD thesis, University of Washington, 1997.

[55] Nick Kushmerick. Wrapper induction: Efficiency and expressiveness. *Artificial Intelligence*, 118(1-2):15–68, 2000.

[56] Nick Kushmerick, Daniel S. Weld, and Robert B. Doorenbos. Wrapper Induction for Information Extraction. In *Proceedings of International Joint Conference on Artificial Intelligence (IJCAI-97)*, Nagoya, 1997.

[57] Wendy Lehnert, Claire Cardie, David Fisher, Joseph McCarthy, Ellen Riloff, and Steven Soderland. Evaluating an Information Extraction system. *Journal of Integrated Computer-Aided Engineering*, 1(6), 1994.

[58] Ling Liu, Calton Pu, and Wei Han. XWRAP: An XML-enabled Wrapper Construction System for Web Information Sources. In *Intern. Conference on Data Engineering (ICDE)*, pages 611–621, 2000.

[59] Joseph McCarthy and Wendy Lehnert. Using decision trees for coreference resolution. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence*, pages 1050–1055, 1995.

[60] Drew McDermott. A Temporal Logic for Reasoning about Processes and Plans. *Cognitive Science*, 6:101–155, 1982.

[61] Itay Meiri. Combining qualitative and quantitative constraints in temporal reasoning. In Thomas Dean and Kathleen McKeown, editors, *Proceedings of the Ninth National Conference on Artificial Intelligence (AAAI'91)*, pages 260–267, Menlo Park, California, 1991. AAAI, AAAI Press.

[62] Ryszard S. Michalski. A theory and methodology of inductive learning. *Artificial Intelligence*, 20(2):111–161, 1983.

[63] Silvia Miksch. Plan management in the medical domain. *AI Communications*, 12(4):209–235, 1999.

[64] Silvia Miksch, Robert Kosara, Yuval Shahar, and Peter Johnson. AsbruView: Visualization of time-oriented, skeletal plans. In *Proceedings of the 4th International Conference on Artificial Intelligence Planning Systems 1998 (AIPS-98)*. Carnegie Mellon University, AAAI Press, 1998.

[65] Tom M. Mitchell. Version spaces: A candidate elimination approach to rule learning. In *Proceedings of the 5th International Joint Conference on Artificial Intelligence (IJCAI-77)*, pages 305–310, Cambridge, MA, August 1977.

[66] Monika Moser and Silvia Miksch. Improving clinical guideline implementation through prototypical design patterns. In *Proceedings of the 10th Conference on Artificial Intelligence in Medicine (AIME), Lecture Notes in Computer Science (LNCS 3581)*, pages 126–130, Aberdeen, UK, 2005. Springer Verlag.

[67] Mark A. Musen. Domain ontologies in software engineering: Use of Protege with the EON architecture. *Methods of Information in Medicine*, 37:540–550, 1998.

[68] Mark A. Musen, Samson W. Tu, Amar K. Das, and Yuval Shahar. EON: A component-based approach to automation of protocol-directed therapy. *Journal of the American Medical Information Association (JAMIA)*, 3(6):367–388, 1996.

[69] Ion Muslea, Steve Minton, and Craig Knoblock. A hierarchical approach to wrapper induction. In Oren Etzioni, Jörg P. Müller, and Jeffrey M. Bradshaw, editors, *Proceedings of the Third International Conference on Autonomous Agents (Agents'99)*, pages 190–197, Seattle, WA, USA, 1999. ACM Press.

[70] Ion Muslea, Steven Minton, and Craig A. Knoblock. Hierarchical wrapper induction for semistructured information sources. *Autonomous Agents and Multi-Agent Systems*, 4(1/2):93–114, 2001.

[71] Mor Peleg, Aziz A. Boxwala, Omolola Ogunyemi, Qing Zeng, Samson Tu, Ronilda Lacson, Elmer Bernstam, Nachman Ash, Peter Mork, Lucila Ohno-Machado, Edward H. Shortliffe, and Robert A. Greenes. GLIF3: The Evolution of a Guideline Representation Format. In *Proceedings of the AMIA Annual Symposium*, pages 645–649, Los Angeles, CA, 2000.

[72] Leonid Peshkin and Avi Pfeffer. Bayesian information extraction network. In *Proc. of the 18th International Joint Conference on Artificial Intelligence (IJCAI)*, 2003.

[73] Kristi-Anne Polvani, Abha Agrawal, Bryant Karras, Aniruddha Deshpande, and Richard Shiffman. *GEM Cutter Manual*. Yale Center for Medical Informatics, 2000.

[74] Marek Růžička and Vojtech Svátek. Mark-up Based Analysis of Narrative Guidelines with the Stepper Tool. In Katharina Kaiser, Silvia Miksch, and Samson W. Tu, editors, *Computer-based Support for Clinical Guidelines and Protocols. Proceedings of the Symposium on Computerized Guidelines and Protocols (CGP 2004)*, volume 101: Studies in Health Technology and Informatics, pages 132–136, Prague, Czech Republic, 2004. IOS Press.

[75] Ellen Riloff. Automatically constructing a dictionary for information extraction tasks. In *Proc. of the 11th National conference on Artificial Intelligence*, pages 811–816, 1993.

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

[77] Ellen Riloff. An empirical study of automated dictionary construction for information extraction in three domains. *Artificial Intelligence*, 85(1-2):101–134, 1996.

[78] Ellen Riloff and Rosie Jones. Learning dictionaries for information extraction by multi-level bootstrapping. In *Proceedings of the 16th National Conference on Artificial Intelligence*, pages 474–479. AAAI Press/MIT Press, 1999.

[79] Andreas Seyfang, Robert Kosara, and Silvia Miksch. Asbru 7.3 Reference Manual. Technical Report Asgaard-TR-2002-1, Vienna University of Technology, Institute of Software Technology & Interactive Systems, Vienna, Austria, Europe, 2002.

[80] Andreas Seyfang, Silvia Miksch, Peter Votruba, Kitty Rosenbrand, Jolanda Wittenberg, Joyce von Croonenborg, Wolfgang Reif, Martin Balser, Jonathan Schmitt, Theo van der Weide, Peter Lucas, and Arjen Hommersom. MHB – a many-headed bridge between guideline formats. Technical Report Asgaard-TR-2004-5, Vienna University of Technology, Austria; Dutch Institute fo rHealthcare Improvements, the Netherlands; University of Augsburg, Germany; University of Nijmegen, the Netherlands, 2004.

[81] Andreas Seyfang, Silvia Miksch, Peter Votruba, Kitty Rosenbrand, Jolanda Wittenberg, Joyce von Croonenborg, Wolfgang Reif, Michael Balser, Jonathan Schmitt, Theo van der Weide, Peter Lucas, and Arjen Homersom. Specification of formats of intermediate, Asbru and KIV representations. Deliverable D2.2a, The Protocure Project, November, 24 2004.

[82] Yuval Shahar, Silvia Miksch, and Peter Johnson. The Asgaard project: A task-specific framework for the application and critiquing of time-oriented clinical guidelines. *Artificial Intelligence in Medicine*, 14:29–51, September 1998.

[83] Yuval Shahar, Ohad Young, Erez Shalom, Alon Mayaffit, Robert Moskovitch, Alon Hessing, and Maya Galperin. DEGEL: A hybrid, multiple-ontology framework for specification and retrieval of clinical guidelines. In Michel Dojat, Elpida Keravnou, and Pedro

Barahona, editors, *Artificial Intelligence in Medicine: 9th Conference on Artificial Intelligence, in Medicine in Europe, AIME 2003, LNAI 2780*, pages 122–131, Protaras, Cyprus, 2003. Springer-Verlag.

[84] Richard N. Shiffman, Bryant T. Karras, Abha Agrawal, Roland Chen, Luis Marenco, and Sujai Nath. GEM: a proposal for a more comprehensive guideline document model using XML. *Journal of the American Medical Informatics Association (JAMIA)*, 7(5):488–498, 2000.

[85] Yoav Shoham. Temporal logics in AI: Semantical and ontological considerations. *Artificial Intelligence*, 33(1):89–104, September 1987.

[86] Stephen Soderland. Learning Information Extraction Rules for Semi-Structured and Free Text. *Machine Learning*, 34(1-3):233–272, 1999.

[87] Stephen Soderland, David Fisher, Jonathan Aseltine, and Wendy Lehnert. CRYSTAL: inducing a conceptual dictionary. In *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence (IJCAI'95)*, pages 1314–1319, 1995.

[88] Margarita Sordo, Omolola Ogunyemi, Aziz A. Boxwala, Robert A. Greenes, and Samson Tu. GELLO: An object-oriented query and expression language for clinical decision support. Summary report, Openclinical.org, March, 17 2004.

[89] Kiyoshi Sudo. *Unsupervised Discovery of Extraction Patterns for Information Extraction*. PhD thesis, New York University, New York, September 2004.

[90] Kiyoshi Sudo, Satoshi Sekine, and Ralph Grishman. Automatic pattern acquisition for Japanese Information Extraction. In *Proceedings of Human Language Technology Conference (HLT2001)*, San Diego, CA, 2001.

[91] David R. Sutton and John Fox. The syntax and semantics of the pro*forma* guideline modeling language. *Journal of the American Medical Informatics Association (JAMIA)*, 10(5):433–443, Sep / Oct 2003.

[92] Voitech Svátek and Marek Růžička. Step-by-step mark-up of medical guideline documents. *International Journal of Medical Informatics*, 70(2-3):319–335, July 2003.

[93] Paolo Terenziani, Carlo Carlini, and Stefania Montani. Towards a comprehensive treatment of temporal constraints in clinical guidelines. In *Proc. of the Ninth International Symposium on Temporal Representation and Reasoning (TIME-02)*, pages 20–27, Manchester, 2002. IEEE Computer Society Press.

[94] Paolo Terenziani, Gianpaolo Molino, and Mauro Torchio. A modular approach for representing and executing clinical guidelines. *Artificial Intelligence in Medicine*, 23(3):249–276, November 2001.

[95] Paolo Terenziani, Stefania Montani, Alessio Bottrighi, Mauro Torchio, Gianpaolo Molino, and Gianluca Correndo. The GLARE approach to clinical guidelines: Main features. In Katharina Kaiser, Silvia Miksch, and Samson Tu, editors, *Computer-based*

Support for Clinical Guidelines and Protocols. Proceedings of the Symposium on Computerized Guidelines and Protocols (CGP 2004), volume 101: Studies in Health Technology and Informatics, pages 162–166, Prague, Czech Republic, 2004. IOS Press.

[96] Samson W. Tu and Mark A. Musen. A Flexible Approach to Guideline Modeling. In *Proceedings of the AMIA Annual Symposium*, pages 420–424, Washington DC, 1999.

[97] Samson W. Tu and Mark A. Musen. Representation formalisms and computational methods for modeling guideline-based patient care. In Barbara Heller, Markus Löffler, Mark Musen, and Mario Stefanelli, editors, *Proceedings of the First European Workshop on Computer-based Support for Clinical Guidelines and Protocols 2000*, volume 83: Studies in Health Technology and Informatics, pages 115–132, Leipzig, Germany, 2001. IOS Press.

[98] Samson W. Tu, Mark A. Musen, Ravi Shankar, James Campbell, Karen Hrabak, James McClay, Stanley M. Huff, Robert McClure, Craig Parker, Roberto Rocha, Robert Abarbanel, Nick Beard, Julie Glasgow, Guy Mansfield, Prabhu Ram, Qin Ye, Eric Mays, Tony Weida, Christopher G. Chute, Kevin McDonald, David Mohr, Mark A. Nyman, Sidna Scheital, Harold Solbrig, David A. Zill, and Mary K. Goldstein. Modeling guidelines for integration into clinical workflow. In Marius Fieschi, Enrico Coiera, and Yu-Chuan Jack Li, editors, *Proc. from the Medinfo 2004 World Congress on Medical Informatics*, pages 174–178. IOS Press, 2004.

[99] United States National Library of Medicine. Fact sheet: Medical Subject Headings (MeSH). http://www.nlm.nih.gov/pubs/factsheets/mesh.html, 27 May 2005. [retrieved: Sept. 19, 2005].

[100] Lluis Vila. A survey on temporal reasoning in artificial intelligence. *AI Communications*, 7(1):4–28, 1994.

[101] Lluis Vila and Eddie Schwalb. A theory of time and temporal incidence based on instants and periods. In *3rd Workshop on Temporal Representation and Reasoning (TIME'96)*, pages 21–28, Key West, Florida, May 19 - 20 1996.

[102] Mark Vilain and Henry Kautz. Constraint propagation algorithms for temporal reasoning. In *Proceedings of the 5th National Conference on Artificial Intelligence (AAAI'86)*, pages 377–382, Philadelphia, PA, August 1986. AAAI.

[103] Peter Votruba, Silvia Miksch, and Robert Kosara. Linking Clinical Guidelines with Formal Representations. In *Proceedings of the 9th Conference on Artificial Intelligence in Medicine in Europe (AIME 2003)*, pages 152–157, Berlin, 2003. Springer.

[104] Peter Votruba, Silvia Miksch, Andreas Seyfang, and Robert Kosara. Tracing the formalization steps of textual guidelines. In Katharina Kaiser, Silvia Miksch, and Samson Tu, editors, *Computer-based Support for Clinical Guidelines and Protocols. Proc. of the Symposium on Computerized Guidelines and Protocols (CGP 2004)*, volume 101: Studies in Health Technology and Informatics, pages 172–176, Prague, Czech Republic, 2004. IOS Press.

[105] Roman Yangarber and Ralph Grishman. NYU: Description of the Proteus/PET system as used for MUC-7 ST. In *Proceedings of the 7th Message Understanding Conference: MUC-7*, Washington, DC, 1998.

[106] Roman Yangarber, Ralph Grishman, Pasi Tapanainen, and Silja Huttunen. Automatic acquisition of domain knowledge for information extraction. In *Proceedings of the 18th International Conference on Computational Linguistics (COLING 2000)*, Saarbrücken, Germany, August 2000.

Curriculum Vitae

Address	Katharina Kaiser [redacted] kaiser@ifs.tuwien.ac.at
Date of Birth	[redacted]

Education

November 2002 to November 2005
Ph.D. studies in Information Science
at the Vienna University of Technology

October 1993 to October 2002
M.S. studies in Information Systems
at the Vienna University of Technology

Job Experience	since November 2002 Research assistant at the Institute of Software Technology & Interactive Systems March 2004 to July 2004 and March 2005 to July 2005 Student assistant at the Institute of Software Technology & Interactive Systems, Information & Software Engineering Group March 2001 to June 2002 Student assistant at the Institute of Information Systems, Database and Artificial Intelligence Group February 1996 Technical support at the Center of Technology Support Europe at Hewlett Packard Brussels, Belgium March 1995 to December 1997 and May 1999 to December 1999 Technical support at the IT department at Hewlett Packard Vienna, Austria
Professional Memberships	ÖGAI (Austrian Society for Artificial Intelligence)
Publications	Please see: http://www.asgaard.tuwien.ac.at/~kaiser/publications.html