

Supporting Physicians in Taking Decisions in Clinical Guidelines: the GLARE “What if” Facility

Paolo Terenziani¹, Stefania Montani¹, Alessio Bottrighi¹, Mauro Torchio², Gianpaolo Molino²
¹DI, Univ. Piemonte Orientale “A. Avogadro”, Spalto Marengo 33, Alessandria, Italy
²Lab. Informatica Clinica, Az. Ospedaliera S. G. Battista, C.so Bramante 88, Torino, Italy

GLARE (GuideLine Acquisition, Representation and Execution) is a domain-independent system for the acquisition, representation and execution of clinical guidelines. GLARE is unique in its approach to supporting the decision-making process of users/physicians faced with various alternatives in the guidelines. In many cases, the best alternative cannot be determined on the basis of “local information” alone (i.e., by considering just the selection criteria associated with the decision at hand), but must also take into account information stemming from relevant alternative pathways. Exploitation of “global information” available in the various pathways is made possible by GLARE through the “what if” facility, a form of hypothetical reasoning which allows users to gather relevant decision parameters (e.g., costs, resources, times) from selected parts of the guideline in a semi-automatic fashion. In particular, the extremely complex task of coping with temporal information involves the extension and adaptation of various techniques developed by the Artificial Intelligence (AI) community.

1 INTRODUCTION

Optimizing clinical guidelines management and exploitation is a key area of research in Artificial Intelligence (AI) in medicine and medical decision-making [1-3]. Guidelines reflect the current understanding of best clinical practice and can assist physicians in the treatment of disease. They can also serve as a system for critical review and evaluation or for educational aims. Numerous approaches to the creation of computer-assisted tools for guidelines management have been proposed in the literature (e.g.: Asbru [4-5], EON [6], GEM [7], GLIF [8-9], GUIDE [10], ONCOCIN [11], PROforma [12], T-HELPER [13], and also [14-15]). In this paper, we describe GLARE [16,17], a user-friendly and domain-independent system supporting guideline acquisition, representation and execution.

2 MAIN FEATURES OF GLARE

Representation formalism. In contrast to several representation formalisms described in the literature which provide a wide set of primitives (see, e.g., EON [6]), in GLARE we focused on the trade-off between expressiveness and complexity. In order to guarantee

usability of the program to user-physicians not expert in Computer Science we defined a *limited set* of clear representation primitives, covering most of the relevant aspects of a guideline [16; see also 12]. We distinguish between *atomic* and *composite actions (plans)*, where *atomic actions* represent simple steps in a guideline, and *plans* represent actions which can be further broken-down and defined in terms of their components via the *has-part* relation. The *has-part* relation permits top-down refinement: a guideline itself can be seen as a composite action which can be broken down hierarchically into sub-parts. *Control relations* establish which actions can be executed next, and in what order. We distinguish between four different control relations: *sequence*, *controlled*, *alternative* and *repetition*. In the sequence and alternative relations, one can choose to indicate the minimum and/or maximum delay between actions. In a set of actions which are in a controlled relation, one can specify the minimum and/or maximum distance between any pair of endpoints of such actions. Finally, two different ways of specifying repetitions are defined (and can be combined): one can state that the action has to be performed until a given *exit condition* becomes true, or can specify a duration (*frame-time*) for the repetitions. In both cases, the *frequency* of the repetitions in time can be specified; then, several other parameters must/can be provided. For example, in the the frequency “3 times every 2 days” it is necessary to provide the *granularity* for the repetition (days in the example), the *grouping* to be considered (2 in the example), and, the *number of executions* of the action in the given periodicity (3 in the example). Four different types of *atomic actions* have been defined: *work actions* (actions to be performed at a certain step of the guideline), *query actions* (requests for information), *decisions* (selections among alternatives) and *conclusions* (explicit output of a decision process). Actions can be described in terms of their attributes. In particular, *work actions* are characterized by name, description, cost, time, resources, goals. *Decision actions* can be classified as *diagnostic decisions* or *therapeutic decisions*. *Diagnostic decisions* are represented as an open set of triples <diagnosis, parameter, score>, plus a threshold to be compared with the different diagnoses’ scores in order to select the most suitable alternative. On the other hand, *therapeutic decision* are characterized by a set of parameters (effectiveness, cost, side-effects,

compliance, duration), shown to the user at execution time and allowing for an informed therapy selection.

System architecture. As in most approaches in the literature, GLARE distinguishes between the *acquisition* phase (when a guideline is introduced into the system –e.g., by a committee of expert physicians) and the *execution* phase (when a guideline is applied to a specific patient). The system architecture is therefore composed of two main modules, the *acquisition tool* and the *execution tool*. The tools interact strictly with a set of databases, including the *clinical* database, which provides physicians with standard terminology, and the *patient* database, which contains the patients' data (see [16] for details).

Acquisition tool. The graphical interface of the acquisition tool is used to acquire atomic actions, *has-part* relations and control relations between the components of plans. The guideline is depicted as a graph, with each action represented as a node (different forms and colours are used to distinguish among different types of actions), while control relations are represented as arcs. By clicking on the nodes in the graph, the user can trigger other windows in order to acquire the internal descriptions (attributes) of the nodes. The interface also shows the hierarchical structure of the guideline in tree-form, where plans can be seen as parents of their components. The acquisition tool provides different forms of consistency checking. *Name* and *range checking* is automatically triggered whenever the expert physician introduces a new term or value within the description of an action, by forcing her/him to use only terms/values that have already been defined within the *clinical* database. Whenever the expert physician introduces a node or arc, different controls are automatically activated to check whether the new element is consistent with several *logical design criteria*. For example, alternative arcs may only exit from a decision action.

Execution tool. The execution tool is typically used "on-line": a user physician applies a guideline with reference to a specific patient; the patient's data is automatically retrieved from the *patient* database, and exploited by the system. This method is used for integrating guidelines in clinical practice. GLARE is also available for "off-line" execution (for education, critical review and evaluation purposes). During on-line execution, delays between actions in the guideline must be respected at execution time, while during off-line execution the engine must jump directly from one action to the next, without waiting for the given delay. In order to support both modalities, we adopt a technique relying on an "agenda" [16-17], a data structure which contains the next actions to be performed. The execution tool also incorporates a decision support facility, described in section 3.

Testing. We have already tested our prototype acquisition and representation system. Several groups of expert physicians, following a few-hour training session, used GLARE in order to acquire algorithms concerning different clinical domains, including bladder cancer, reflux esophagitis, and heart failure. In all the tests, our representation formalism and acquisition tool proved expressive enough to cover the clinical algorithms, and the acquisition of a clinical guideline was reasonably fast (e.g., the acquisition of the guideline on heart failure required 3 days).

3 THE "WHAT-IF" FACILITY

GLARE's execution tool incorporates a decision support facility to assist physicians in choosing among different therapeutic or diagnostic alternatives.

The default execution of decision actions works as follows. When a *diagnostic decision* is pushed onto the agenda, the execution module automatically retrieves the parameter values from the *patient* database, evaluates the scores for every alternative diagnosis, and then compares them with the corresponding threshold. All alternative diagnoses are then shown to the user-physician, together with their scores and the threshold, and the tool lets the user choose among them (a warning is given if the user chooses a diagnosis which does not exceed the threshold). The execution of a *therapeutic decision* consists in presenting the effectiveness, cost, side-effects, compliance, and duration of each alternative to the physician, thus allowing her/him to select one of them. GLARE's "what-if" facility is the implementation of a form of *hypothetical reasoning* that user-physicians can use to gain additional clues as to what would be the best course of action. In particular, users are helped in gathering the various types of information necessary for discriminating among the alternatives at any stage of the guideline. In many cases, therapeutic and/or diagnostic conclusions should not be taken on the basis of "local information" alone (i.e. by considering just the decision criteria associated with the specific decision at hand) but should also take into account information stemming from relevant alternative pathways. The unique feature of GLARE's "what if" facility is the capability of retrieving such "global information". This facility can be used both in the on-line and in the off-line execution mode. Essentially, it provides an idea of what could happen in the rest of the guideline were the user to select a given alternative for the patient at hand, and supports for comparisons of the alternatives. The graphical interface of the execution tool allows the user to select the portion of the guideline s/he believes to be relevant to the decision-making process. Figure 1 shows part of the *gallstone treatment* guideline. If symptomatic gallstones are diagnosed, a

surgical approach has to be chosen. Laparoscopic surgery and laparotomic surgery are both suitable methods and therefore indications about the implications of each alternative (in terms of e.g. time and money spent, and resources exploited), could be useful in deciding. The “what if” facility is even more helpful if symptomless gallstones are diagnosed. In this case, three choices are available: expectant management, litholitic therapy, and surgery, which, again, can be either laparoscopic or laparotomic. Thus, in this case, a series of two consecutive therapeutic decisions could be presented to the physician. The exploitation of the “global information” for therapy selection proves extremely relevant, since in this example, as in several other situations, no alternative is really “better” than the others from a strictly clinical viewpoint. The possibility of obtaining a complete scenario of the decision consequences is clearly an added value.

Generally speaking, when the “what-if” facility is activated, the physician is asked to indicate on the graph the starting node of the paths to be compared and (preferably) the ending nodes (otherwise, all the possible paths exiting the starting node will be taken into consideration). Relevant decision parameters (costs, resources, times) will be gathered from the selected portions of the guideline in a semi-automatic way. In particular, whenever a decision action is reached within each path, the user is allowed to choose a subset of alternatives by selecting the corresponding buttons in a pop-up window. If a selection is made, the other alternatives will be ignored by the reasoning process. If a plan is found, it is expanded into its components, and the “what-if” facility is recursively applied to each of them by analyzing all the decision actions that appear at the various decomposition levels. At the end of this process, the tool displays the values of the chosen parameters for each of the selected paths. While resources in a path are simply collected and costs summed up, coping with temporal information is a very complex task which involves the extension and adaptation of different temporal reasoning techniques developed within the AI community (see section 4).

4. TEMPORAL REASONING ISSUES

The “what if” facility can be used to compare the duration of different paths (e.g., to discover the temporally shortest path). The treatment of the “what if” facility involves quite subtle technical work when taking into account the temporal parameter. In fact, three different types of temporal constraints are involved within clinical guidelines dealt with by the GLARE system:

1. “standard” qualitative and quantitative constraints between actions belonging to the same plan.

These constraints are expressed within the *sequence*, *alternative* and *controlled* relations between actions, and by the durations of actions;

2. implicit constraints implied by the *part-of* relations between actions;
3. constraints on the *repetition* of actions, and on their periodicity.

Of course, all these constraints must be *represented*, and the interplay between them must be taken into account by the *temporal reasoning* process, in order to check consistency and to determine the temporally shortest paths in a *correct* and *complete* way. Besides completeness, the *tractability* of the temporal reasoning process is another main goal we set out to achieve. We based our work on “classical” AI approaches to temporal reasoning based on *bounds on differences* and on the STP framework [18-20], and extended them to cope with issues (1), (2), and (3) above. The STP framework takes into account conjunctions (sets) of bounds on differences, and can be used to model *precise* or *imprecise* temporal locations (dates), durations, delays between points, and different forms of qualitative temporal constraints between time points and/or time intervals [20-21]. This framework has very elegant computational properties: correct and complete propagation of the constraints (e.g., for consistency checking) can be performed in cubic time by a classical all-to-all-shortest-paths algorithm (such as Floyd-Warshall’s), and can furnish the *minimal network* of the constraints as output (i.e., the minimal constraints between each pair of entities) [18].

Temporal constraints between actions in a GLARE plan (i.e. constraints of type (1) above) can be mapped onto an STP framework, in which each action is represented by its starting and its ending point, and the delays associated with sequence and alternative relations can be mapped as distances between points. From the temporal point of view, we also represent a plan via its starting and its ending point. In this way, the *temporal constraints between a plan and the other actions* (atomic actions and/or plans) at the same hierarchical level can be simply represented as bound on differences. Moreover, *temporal constraints between each plan and its subactions* (i.e., constraints of type (2) above) are also represented as constraints between the starting/ending points of the plan and the endpoints of its (direct) subactions. In particular, the starting point of a plan is temporally equal to the starting point of its starting action, while its ending point is after or simultaneous with the ending point of the last action in each alternative path of sub-actions composing it. Thus, the constraints involving actions and plans (which are *not repeated*) can be homogeneously represented within a unique STP framework, containing a starting and an ending point for each plan and atomic action in the guideline.

Unfortunately, such an homogeneous approach must be significantly extended if one wishes to deal with *repeated actions* (i.e. with constraints of type (3)). For instance, if P is a repeated plan composed of a sequence of two subactions A_1 and A_2 , what we actually need to represent is the fact that P 's starting point coincides with the starting point of the *first repetition* of A_1 , and P 's ending point with the ending point of the *last repetition* of A_2 . Such constraints cannot be trivially modeled into the homogeneous STP framework discussed above. We thus chose to represent the constraints regarding repeated actions into separate STP frameworks, one for each repeated plan. Thus, in GLARE, the overall set of constraints in a guideline is represented by a *tree of STP frameworks* (STP-tree henceforth). The root of the tree is the STP which homogeneously represents the constraints between all the actions (composite and atomic) in the guideline, except repeated actions (which are plans, by our definition). Each node in the STP-tree is an STP, and has as many children as the number of repeated actions it contains. Each arc in the tree connects a pair of endpoints in an STP (the starting and ending point of a repeated action) to the STP containing the constraints between its subactions, and is labeled with a list of properties describing the temporal constraints on the repetitions (granularity, grouping etc.). Temporal consistency checking proceeds in a top-down fashion, starting from the root node of the STP-tree. In fact, the root is a "standard" STP, so that Floyd-Warshall's all-to-all shortest path standard algorithm can be applied to check its consistency, and to produce the minimal network of

the constraints it contains (*local minimal network* henceforth). Thereafter, we proceed towards the leaves of the tree. For each node in the tree other than the root, we progress in three steps: (1) first, we check the consistency of the constraints used to specify the repetition taken in isolation; (2) second, we map the "extra" temporal constraints regarding the repetition onto bounds on difference constraints; (3) third, we apply the Floyd-Warshall algorithm to the constraints in the STP plus the "extra" bounds on difference constraints determined in the previous step. This also provides a local minimal network.

While the third step is trivial, the first two steps are performed by ad-hoc specialised algorithms [22].

Property 1. The top-down visit of the STP-tree is complete as regards consistency checking.

Besides checking the consistency of the temporal constraints, our approach also produces a tree of local minimal networks. This is exploited by another specialised algorithm in order to efficiently evaluate the minimum and maximum duration of a user-selected path starting from an action A_1 and ending with an action A_n , as required by the "what if" facility operating on the time parameter. The distance between A_1 and A_n is evaluated on the basis of the path on the STP-tree corresponding to the path selected by the user in the guideline, exploiting the local minimal networks in order to look for (1) the duration of a whole repetition and (2) the distance between the endpoint of the reference action (which is initially A_1) and the end of one complete repetition (see [22] for more details).

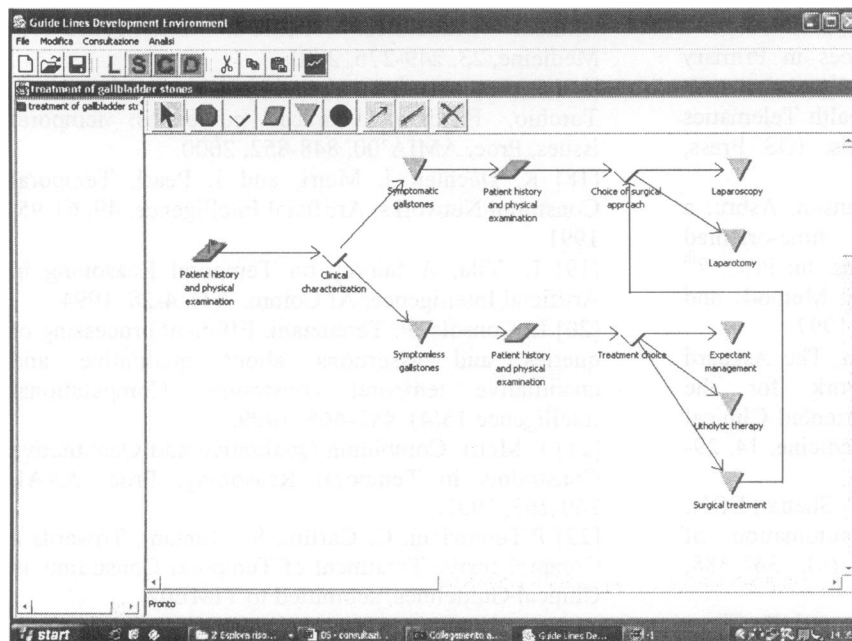


Figure 1: Part of the gallstone treatment guideline. The "what if" facility can be used to compare different paths.

5 COMPARISONS AND FUTURE WORK

In the latest years, many (semi-)automatic approaches dealing with clinical guidelines have been proposed: among them, we consider PROforma [12] and Asbru [4-5] to be the most similar to GLARE (see [16] for comparisons). Nevertheless, to the best of our knowledge, the literature describes no analogous approach to our “what if” facility which supports medical decision-making on the basis of “global information” in the guideline. Moreover, from the temporal point of view, GLARE also offers a very expressive representation formalism (dealing with inexact temporal constraints and repeated events) as well as complete and tractable algorithms for temporal consistency checking and for implementing the “what if” facility in the temporal context.

In the short term, we plan to systematically test the temporal reasoning and “what-if” facilities described in sections 3 and 4 on practical clinical cases. A long-term goal will be further extension of the temporal formalism, while striving to maintain an appropriate balance between expressiveness and computational complexity.

REFERENCES

- [1] J.M. Grimshaw, I.T. Russel, Effect of clinical guidelines on medical practice: a systematic review of rigorous evaluation. *Lancet*, 342-342, 1317-1322, 1993.
- [2] G. Molino, From Clinical Guidelines to Decision Support, in: *Proc. Joint European Conference on Artificial Intelligence in Medicine and Medical Decision Making*, LNAI 1620, Aalborg, 3-12, 1999.
- [3] I. Purves, Computerised Guidelines in Primary Health Care: Reflections and Implications. in: C. Gordon and J.P. Christensen, eds., *Health Telematics for Clinical Guidelines and Protocols*, IOS Press, Amsterdam, 57-74, 1995.
- [4] S. Miksch, Y. Shahar, and P. Johnson, Asbru: a task-specific, intention-based, and time-oriented language for representing skeletal plans, in: *Proc. 7th Workshop on Knowledge Engineering Methods and Languages*, Milton Keynes, UK, 9-20, 1997.
- [5] Y. Shahar, S. Mirksch, P. 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, 1998.
- [6] M.A. Musen, S.W. Tu, A.K. Das, Y. Shahar, EON: a component-based approach to automation of protocol-directed therapy, *JAMIA*, 3(6), 367-388, 1996.
- [7] R.N. Shiffman, B.T. Karras, A. Agrawal, R. Chen, L. Menco, and S. Nath, GEM: a proposal for a more comprehensive guideline document model using XML, *JAMIA*, 7(5), 488-498, 2000.
- [8] L. Ohno-Machado, J.H. Gennari, S. Murphy, N.L. Jain, S.W. Tu, D.E. Oliver, et al., The Guideline Interchange Format: A Model for Representing Guidelines, *JAMIA*, 5(4), 357-372, 1998.
- [9] M. Peleg, A.A. Boxawala, et al., GLIF3: The evolution of a Guideline Representation Format, in: *Proc. AMIA'00*, 645-649, 2000.
- [10] S. Quaglini, M. Stefanelli, A. Cavallini, G. Miceli, C. Fassino, and C. Mossa, Guideline-based careflow systems, *Artificial Intelligence in Medicine*, 20(1), 5-22, 2000.
- [11] S.W. Tu, M.G. Kahn, M.G. Musen, J.K. Ferguson, E.H. Shortliffe, and L.M. Fagan, Episodic Skeletal-plan Refinement on Temporal Data, *Comm. ACM*, 32, 1439-1455, 1989.
- [12] J. Fox, N. Johns, A. Rahmzadeh, R. Thomson, Disseminating medical knowledge: the PROforma approach, *Artificial Intelligence in Medicine*, 14, 157-181, 1998.
- [13] M.A. Musen, C.W. Carlson, L.M. Fagan, S.C. Deresinski, E.H. Shortliffe, Automated Support for Community-Based Clinical Research, in: *Proc. 16th Annual Symposium on Computer Applications in Medical Care*, 719-723, 1992.
- [14] C. Gordon and J.P. Christensen, *Health Telematics for Clinical Guidelines and Protocols*, IOS Press, 1995.
- [15] Special Issue on Workflow Management and Clinical Guidelines, D.B. Fridsma (Guest ed.), *JAMIA*, 22(1), 1-80, 2001.
- [16] P. Terenziani, G. Molino, and M. Torchio, A Modular Approach for Representing and Executing Clinical Guidelines, *Artificial Intelligence in Medicine*, 23, 249-276, 2001.
- [17] P. Terenziani, F. Mastromonaco, G. Molino, M. Torchio, Executing clinical guidelines: temporal issues, *Proc. AMIA'00*, 848-852, 2000.
- [18] R. Dechter, I. Meiri, and J. Pearl, Temporal Constraint Networks, *Artificial Intelligence*, 49, 61-95, 1991.
- [19] L. Vila, A Survey on Temporal Reasoning in Artificial Intelligence, *AI Comm.* 7(1), 4-28, 1994.
- [20] L. Console, P. Terenziani, Efficient processing of queries and assertions about qualitative and quantitative temporal constraints, *Computational Intelligence* 15(4), 442-465, 1999.
- [21] I. Meiri, Combining Qualitative and Quantitative Constraints in Temporal Reasoning, *Proc. AAAI*, 260-267, 1991.
- [22] P. Terenziani, C. Carlini, S. Montani, Towards a Comprehensive Treatment of Temporal Constraints in Clinical Guidelines, submitted to TIME02.