Washington University in St. Louis

Washington University Open Scholarship

All Computer Science and Engineering Research

Computer Science and Engineering

Report Number: WUCS-89-50

1989-03-09

Model-Based Interpretation of Time-Varying Medical Data

Michael G. Kahn, Lawrence M. Fagan Washington University in St. Louis, and Lewis B. Sheiner Washington University in St. Louis

Temporal concepts are critical is medical therapy-planning. If given early enough, specific therapeutic choices may abort or suppress evolving undesired changes in a patient's clinical status. Effective medical decision making demands recognition and interpretation of complex temporal changes that permeate the medical record. This paper presents a methodology for representing and using medical knowledge about temporal relationships to infer the presence of clinically relevant events, and describes a program, called TOPAZ, that uses this methodology to generate a narrative summary of such events. A unique feature of TOPAZ is the use of numeric and symbolic modeling techniques to perform... Read complete abstract on page 2.

Follow this and additional works at: https://openscholarship.wustl.edu/cse_research

Part of the Computer Engineering Commons, and the Computer Sciences Commons

Department of Computer Science & Engineering - Washington University in St. Louis
Campus Box 1045 - St. Louis, MO - 63130 - ph: (314) 935-6160.

Model-Based Interpretation of Time-Varying Medical Data

Michael G. Kahn, Lawrence M. Fagan Washington University in St. Louis, and Lewis B. Sheiner Washington University in St. Louis

Complete Abstract:

Temporal concepts are critical is medical therapy-planning. If given early enough, specific therapeutic choices may abort or suppress evolving undesired changes in a patient's clinical status. Effective medical decision making demands recognition and interpretation of complex temporal changes that permeate the medical record. This paper presents a methodology for representing and using medical knowledge about temporal relationships to infer the presence of clinically relevant events, and describes a program, called TOPAZ, that uses this methodology to generate a narrative summary of such events. A unique feature of TOPAZ is the use of numeric and symbolic modeling techniques to perform temporal reasoning tasks that would be difficult to encode and perform using only one modeling methodology.

MODEL-BASED INTERPRETATION OF
TIME-VARYING MEDICAL DATA

Michael G. Kahn, Lawrence M. Fagan
and Lewis B. Sheiner

WUCS-89-50

November 1989

Department of Computer Science
Washington University
Campus Box 1045
One Brookings Drive
Saint Louis, MO 63130-4899

Model-Based Interpretation of Time-Varying Medical Data

Michael G. Kahn
Department of Internal Medicine, Box 8121
Washington University School of Medicine
St. Louis, MO 63110

Lawrence M. Fagan
Medical Computer Science Group
Stanford University Medical Center
Stanford, CA 94305-5479

Lewis B. Sheiner
Department of Laboratory Medicine
University of California
San Francisco, CA 94143

March 9, 1989

Abstract

Temporal concepts are critical in medical therapy-planning. If given early enough, specific therapeutic choices may abort or suppress evolving undesired changes in a patient's clinical status. Effective medical decision making demands recognition and interpretation of complex temporal changes that permeate the medical record.

This paper presents a methodology for representing and using medical knowledge about temporal relationships to infer the presence of clinically relevant events, and describes a program, called TOPAZ, that uses this methodology to generate a narrative summary of such events. A unique feature of TOPAZ is the use of numeric and symbolic modeling techniques to perform temporal reasoning tasks that would be difficult to encode and perform using only one modeling methodology.

1   Introduction

Change is an essential feature of all medical decisions. For medical therapy-planning decisions, the options and choices for good, patient-specific treatment strongly hinge on the patient's *past clinical course, current clinical status,* and *predicted future course* (prognosis). A computer-based

1

therapy-planning system that creates a good patient-specific treatment plan must be capable of representing and reasoning about a patient's past, present, and potential future.

Clinical care generates a staggering volume of data. For the physician faced with finding salient prognostic and therapeutic features within a patient's medical record, interpreting and summarizing this large body of information is a complex and demanding task. Computer-based medical-record systems facilitate rapid access to more complete patient information [1,13,21,29,26]. However integrating, interpreting, abstracting, and summarizing the clinical data is still left to the physician.

We report the design of a computer-based decision support system that interprets time-ordered clinical data. A central tenet of this work is that intelligent interpretation of time-ordered data requires several different formalisms. At least three types of temporal knowledge must be encoded:

1. Knowledge about the static and temporal relationships among observations

2. Knowledge about the changing clinical context

3. Knowledge about the expected evolution of observations over time

We describe our methodology for representing and reasoning with diverse temporal knowledge and a program, called TOPAZ, that was implemented using this methodology. TOPAZ analyzes the temporal sequence of white blood cell (WBC) counts and drug dosages from a patient receiving cancer chemotherapy (Figure 1). The program produces a textual summary of the key temporal features of the patient's data (Figure 2). A physician or an expert planning system could use the clinical features detected by TOPAZ to develop or modify patient-specific therapy plans [11].

2   The TOPAZ Methodology

General medical knowledge describes the expected clinical course of a hypothetical, *average* patient. Because a patient has characteristics that distinguish him from the typical patient, each patient's clinical course will differ from the prototypical expected course. After following a patient

2

Weight	90.7		89.5	88.2	87.4	89.1		87.3	88.5	84
WBC X 1000	7.7	7.9	6.9	4.8	3.1	4.4	3.3	4.7	5.3	5.0
PCV	40.9	40.4	37.1	36.3	32.2	32.6	28.2	27.7	24.4	29.9
Hemoglobin	13.1	12.8	12.2	11.8	10.6	10.5	9.4	8.8	9.4	9.2
Platelets	574	329	368	282	305	120	112	193	150	106
BSA (m^2)		2.11								
Chemo Name		MOPP	MOPP	MOPP	MOPP	MOPP	MOPP	MOPP	MOPP	MOPP
Cycle #		1	1	2	2	3	3	4	4	5
Subcycle		A	B	A	B	A	B	A	B	A
Mustard		12	12	12	12	10	8	8	8	8
Vincristine		2.0	2.0	2.0	2.0	2.0	2.0	1.0	1.0	1.0
Procarbazine		200	200	200	100	75	75	75	75	75
Prednisone										
Day	19	26	4	30	6	28	4	9	16	8
Month	Oct	Oct	Nov	Nov	Dec	Dec	Jan	Feb	Feb	Mar
Year	83	83	83	83	83	83	84	84	84	84

· · ·

Figure 1: A portion of the medical record from a patient who has received MOPP chemotherapy. Columns contain data collected during a clinic visit. Rows contain clinical observations, laboratory values, and therapeutic drug dosages. WBC = white blood cell, PCV = packed cell volume, BSA = body surface area. A full record includes demographic data and additional pages of clinical measurements.

Patient Identifying Data:
This is a 25 y/o male who presented with clinical stage III-S-A, pathological stage III-S-A Hodgkins' lymphoma. The patient's medical record begins on 19 October 1983 and ends on 24 May 1984. He was assigned to MOPP chemotherapy and received 7 treatment cycles.

Analysis of Clinical Course:
Based on our analysis of the entire clinical course, the best fit of our model to the data suggests the following key patient-specific physiological parameter values:

1. A steady-state bone marrow mass that is approximately 0.6 times the population value. This is considered a moderate reduction in steady-state bone marrow mass.

2. A duration of drug effect that is approximately 2.2 times the population value. This is considered a moderate increase in the duration of drug effect.

3. A sensitivity to drug effect that is approximately 1.4 times the population value. This is considered a mild increase in the sensitivity to drug effect.

Based on differences between population-based predictions (the expected clinical course) and the patient-specific model predictions (the modeled clinical course), the patient had 2 remarkable clinical events. For each event, we provide a description of the event, a probable cause for the event, and an analysis of the ability of the patient-specific model to predict the actual observations during this event.

1. A series of 2 visits prior to receiving chemotherapy, starting on 19 October 1983 and ending on 26 October 1983, with predicted WBC counts that were systematically lower than expected for this period. The predicted WBC count with the largest deviation from expected occurred on 26 October 1983 (63% of the expected value).

    The patient's estimated steady-state marrow mass was lower than expected causing a lower than expected pre-treatment WBC count. The model cannot explain the underlying etiology of a lower than expected steady state marrow mass.

    There were no periods of poor model fit during this remarkable clinical event.

2. A series of 13 visits after starting chemotherapy, starting on 30 November 1983 and ending on 24 May 1984, with predicted WBC counts that were systematically lower than expected for this period. The predicted WBC count with the largest deviation from expected occurred on 28 December 1983 (64% of the expected value).

    After adjusting for the patient's steady-state bone marrow mass, the two drug-related parameters, duration of drug effect and drug sensitivity, jointly caused an increased drug effect which led to increased myelosuppression and lower than expected WBC counts. The increased drug effect had a peak myelosuppressive effect 13.1 times greater than expected on 19 April 1984. The accumulation of drug due to the prolonged duration of drug effect was responsible for a 9.4 fold increase in drug effect. The increased sensitivity to drug was responsible for a 1.4 fold increase in drug effect.

Figure 2: TOPAZ-generated summary. This summary was constructed from the patient record from which Figure 1 was abstracted.

for some time, an experienced clinician gradually forms an appreciation of the individual's idiosyncratic response to his illness and treatment. For decisions to be tailored to a specific individual, unique patient features must modify general medical knowledge. A skilled summary of the patient's medical record should recognize the same patient-specific implications that clinicians derive when they review a patient's chart. To meet this requirement, we need a method that tailors general medical knowledge based on an individual's clinical observations.

The TOPAZ methodology that we have developed emphasizes three specific tasks (Figure 3) [9]:

1. *Model-based data interpretation.* An explicit causal-based physiological model of the temporal relationships among observations and underlying clinical concepts can be used to detect significant temporal features in time-ordered data. This model encodes which temporal patterns are expected (predicted) and which model features (parameters), if modified, bring the model predictions closer to the actual observations.

2. *Interval-based data abstraction.* Abstraction is the process of combining a set of related features into a single related concept that encompasses the more detailed features. In abstracting time-ordered data, intervals can be used to for combining point observations and for combining multiple smaller intervals.

3. *Problem-based text generation.* An effective summary highlights the important abstractions and suppresses the irrelevant features. A summary also must include supporting and conflicting evidence for each abstraction that is concluded. Problem-based text generation produces textual descriptions of "interesting" abstractions and describes the support or conflict for the abstraction in the data.

The TOPAZ methodology uses an explicit temporal model to encode general medical knowledge. This model incorporates what is known about the underlying features, processes, and relationships that can cause specific measurements to change over time. In medicine, an understanding of causal

5

Observation

WBC:	12.9	9.2	13	3.4	2.1	1.7
Mustard:	11	11	11	8	8	0
Oncovin:	2.0	2.0	2.0	2.0	1.0	0
Procarb:	175	175	175	125	75	0
Pred:	80	80				

Model Fitting

Estimates patient-specific model parameter values and model predictions based on time-ordered clinical observations

Implication

Marrow:
Mass: 1.9 1.9 1.0 0.7 0.5 0.5
Drug
Effect: 0 0 1.4 1.6 2.4 2.7
WBC
(Pred) 1.9 1.9 1.7 0.7 0.6 0.5

Interval Abstraction

Combines intervals of significant deviation from expectation

Abstraction

DrugEffect
Elevated

Abnl Residuals

Problem-based Text Generation

Converts "interesting" internal structures into the domain language

Summarization

This patient had increased drug effect which led to increased myelosuppression and lower than expected WBC counts.

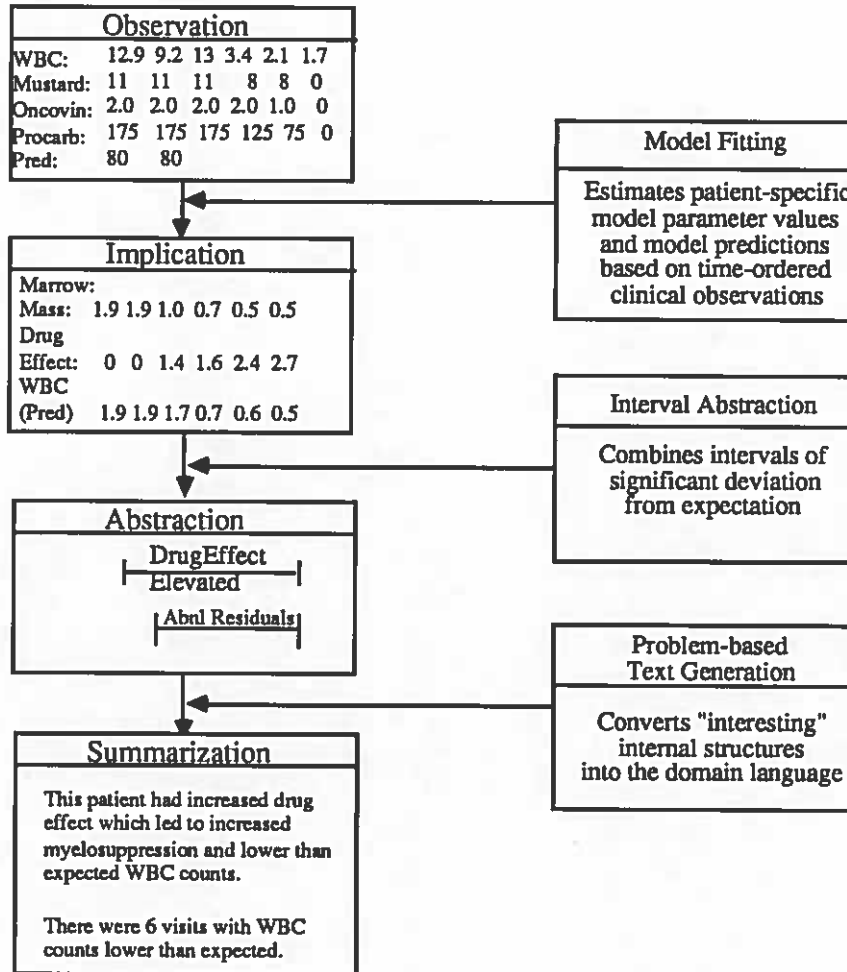There were 6 visits with WBC counts lower than expected.

Figure 3: Architecture of model-based summarization. Summarization of time-ordered data is decomposed to three steps: (1) model fitting, which estimates patient-specific model features from time-ordered observations using an explicit representation of the temporal relationships among inputs, outputs, and underlying features; (2) interval abstraction, which aggregates periods of time during which model predictions and observations show significant deviations from expectations; and (3) text generation, which selects "interesting" abstractions and presents them as text. Note that the observed WBC and the predicted WBC are not identical. This difference is recognized in the abstraction pass and is mentioned in the final summary text.

physiological processes is important for effective clinical problem solving. Therefore, the TOPAZ temporal model encodes physiological knowledge.

When the patient has no previous medical history, only general medical knowledge is available for making initial decisions. As data are obtained, patient-specific knowledge replaces general knowledge, and patient-specific decisions replace general decision-making guidelines. Like the clinician who revises his opinion as new patient features are observed, TOPAZ modifies an initial temporal model as patient-specific measurements are observed (Figure 3, step 1). The transition from the type of knowledge typically found in the medical literature (general medical knowledge) to the type of knowledge needed to treat a specific individual (patient-specific knowledge) occurs as new patient data generate a patient-specific physiological model.

We have defined the term *context* to be a period of time during which a meaningful state or situation exists. Because each domain determines what events and situations are meaningful, contexts are domain-specific. Contexts can occur once and then disappear, can appear intermittently or regularly, or can persist indefinitely once started. In TOPAZ, an interval-based symbolic context model detects periods of time when significant deviations from the expected clinical course were seen (Figure 3, step 2). When an unusual event is detected, the context model encodes the knowledge that attempts to explain the most plausible cause of the unusual events in terms of model parameters and relationships found in the physiological process model. The context model maps model predictions into clinical abstractions. It also serves as the bridge between number-based methods and symbol-based methods. Because this model deals with both numeric and symbolic elements, the knowledge encoded within this model includes both numeric and symbolic concepts.

Physicians describe patients to their colleagues using stereotypic presentation styles. There are two key parts to a typical patient presentation: (1) a description of *notable clinical events* that occurred during observation or treatment, and (2) a discussion of the *implications of notable events* for the patient's clinical status. The content of each part is both problem- and context-specific. The text-generation system must be flexible enough to accommodate records of patients with vastly

7

different clinical courses, but structured enough to impose the typical stylistic features expected by the clinical reader.

The interval abstractions created by the symbolic context model are used to construct a patient-specific summary that is focused on only those problems actually encountered during the clinical course (Figure 3, step 3). To construct text that embodies the expected form and content of a typical clinical summary, we use a third model, called a *presentation model* [14]. The presentation model is constructed from two generalized text schemas: (1) a schema of the general organizational structure of a clinical summary (the *organizational schema*) and (2) a schema of the specific structure of a clinical argument (the *rhetorical schema*). The organizational schema encodes the overall flow of topics that typically appear in clinical summaries. The rhetorical schema encodes the details of constructing a persuasive clinical argument that presents and supports clinical interpretations and inferences. The organizational schema provides the global structure in the clinical summary, whereas the rhetorical schema provides the logical structure of the summary contents.

3   The TOPAZ Program

In Section 2, a three-step methodology for analyzing and summarizing time-ordered medical data was presented. Using actual observations and a generic model of the underlying system structure, a patient-specific model is created. Differences between the predictions generated by the patient-specific model and the predictions generated by the population-based model define "unusual" clinical events. An interval-based model abstracts the patient's clinical course by aggregating intervals of unusual model predictions. The creation of an "unusual" event interval triggers an examination of the structural features of the patient-specific model for plausible explanations of the unusual findings. A presentation model represents the structural and logical features of a clinical presentation so that each abstraction discussed in the summary is presented in the proper clinical context (state). In this section, we present the implementation details of the three TOPAZ
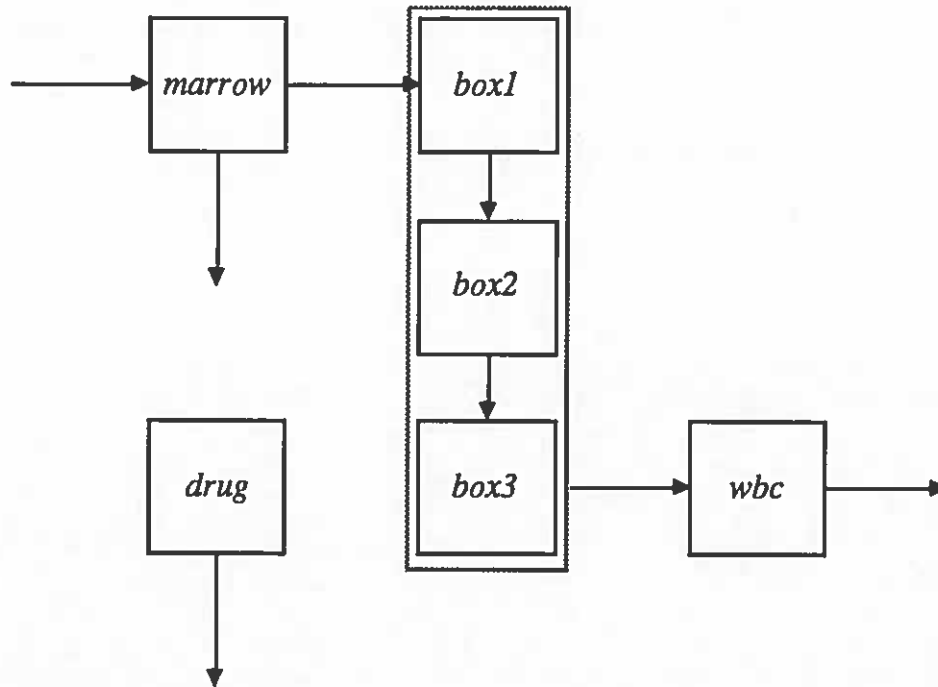
8

Figure 4: Graphical representation of the TOPAZ bone-marrow model. A compartment model is constructed from hypothetical compartments and transfers between compartments. Boxes represent compartments; lines represent transfers. This graph represents a non-linear system of differential equations. The meanings of the model parameters are given in Table 1.

models.

3.1   The Mechanistic Bone-Marrow Model

Using standard techniques of compartment models, a parameterized numeric model of the physiological processes in the production, maturation, and destruction of granulocytes was developed (Figure 4, Table 1). This model is called the *bone-marrow model*.

The graphical representation of the TOPAZ bone-marrow model in Figure 4 contains only the clinically meaningful features of granulocyte development. The bone-marrow compartment (*mar-*

Table 1: Parameters of the TOPAZ bone marrow model.

Parameter name	Parameter units	Initial value	Physiological interpretation
bm_{in}	days^{-1}	0.20	fractional bone-marrow input rate
bm_{ss}	normalized	2.0	bone-marrow steady-state level
$dtransit$	days	2.0	mean delay transit time
wbc_{out}	days^{-1}	0.45	fractional WBC decay rate
d_{out}	days^{-1}	0.075	fractional drug-effect decay rate
d_{senstv}	unitless	0.25	drug sensitivity slope

row), representing immature granulocytes in the marrow, is replenished by an exogenous source of replicating cells, which represents the omnipotent stem cells. Two losses occur from the bone-marrow compartment. One loss enters a chain of compartments, called the *delay chain*, which represents the maturation phase of granulocytic development. The three compartments that constitute the delay chain (*box1*, *box2*, and *box3*) have no losses except into the next delay compartment. This chain effectively delays and spreads out the effects of granulocytic production and MOPP toxicity as they propagate from the bone-marrow compartment to the peripheral-granulocyte (WBC) compartment (*wbc*). The second loss from the bone-marrow compartment represents the destruction of cells within the bone marrow from both ineffective hematopoesis and drug toxicity. Myelosuppression from chemotherapy is modeled as an additional increment to the second marrow-loss rate. This increased loss rate is a function of the amount of drug effect remaining in the patient (*drug*) and the sensitivity of this patient to the drug. The WBC compartment, representing the peripheral granulocytes, is the compartment from which observations (WBC counts) are drawn. A loss that is not sensitive to the presence of drug exits from the WBC compartment, which represents the normal turnover of peripheral granulocytes. Additional destruction of peripheral granulocytes by chemotherapeutic agents is not incorporated in this model.

The TOPAZ physiological model is implemented in FORTRAN using a fifth-order Runge–Kutta

numerical integrator [20] to solve the non-linear differential equations. Prior distributions for the model parameters (Table 1) represent general medical knowledge about typical patients. Patient-specific observations (WBC counts and dosage information) are combined with this prior knowledge using Bayes rule, assuming normally distributed parameters and observation errors, to estimate the mode of the posterior distribution of the parameters [24,23]. These estimated parameters are taken to be the patient-specific model parameters.

3.2   The Symbolic Context Model

The numeric techniques presented in the previous section provide an intuitive method for detecting unusual observations. Model predictions with the prior distributions of model parameters equal to the population values are the *expected* WBC counts for the *typical* patient. Measured WBC counts and patient-specific model predictions that are significantly different from the expected WBC counts are *unusual*. In this section, we present a second modeling technique, based on symbolic temporal pattern matching, designed to derive the abstract clinical concepts implied by unexpected patient observations, abnormal patient-specific model parameter estimates, and unusual model predictions.

Numeric process-based temporal models, such as the TOPAZ bone-marrow model encode time as a continuous quantity. Given a set of initial conditions, parameter values, new inputs, and constraints among variables, the bone-marrow model predicts model states at any time t. Missing from this perspective of time is the *meaning* of the predicted states. A very low predicted or observed WBC count has clinical importance that has no representation in a process-based model. A drug sensitivity that is estimated to be higher than is usual alters the TOPAZ bone-marrow model predictions, but nowhere does this model encode the clinical knowledge that a significant change in drug sensitivity has important therapeutic ramifications. Process models do not represent the *implications* of their predictions; context models encode the *domain-specific knowledge* required to interpret the process-model predictions.

The TOPAZ context model has two main tasks: (1) to abstract clinically meaningful intervals of time from the medical record, and (2) to explain these events using information supplied by the patient-specific bone-marrow model. The TOPAZ context model is encoded in a structure called ETNET. ETNET is derived from and is an extension of TNET [10], a temporal structure that supports context-sensitive temporal queries for time-ordered data in a cancer chemotherapy advising expert system called ONCOCIN [27]. ETNET extends the capabilities of TNET by including context-specific rules to conclude key features about an interval (such as maximum or minimum values within the interval) and rules that determine if other context-sensitive intervals can be created. In TOPAZ, ETNET is used to model changes in a patient's hematologic status as a set of abstraction intervals. Each interval represents the existence of a notable clinical event or physiological state (Figure 5).

TOPAZ examines pair-wise differences among patient observations, population-based predictions, and patient-based predictions to infer meaningful temporal abstractions. For example, if we assume our patient is typical, then a large difference between his patient-specific predicted WBC count and the population-based predicted WBC count is "surprising" (Table 2). TOPAZ creates an ETNODE to denote any interval of time during which large deviations from the expected clinical course were observed (Figure 5). The domain expert specifies how large a difference must be seen for the difference to be considered surprising or atypical. Once this threshold is exceeded, TOPAZ considers each episode to be notable, significant, and worthy of comment and explanation. Each ETNODE generated at this step contains rules that then search for possible explanatory causes of this significant deviation.

3.3 The Presentation Text Model

TOPAZ summary text is designed to mimic the style of a physician describing to a colleague a patient whose course is complicated. When focused on the patient's clinical measurements, physicians discuss more than just the numbers; they also interpret what these numbers mean
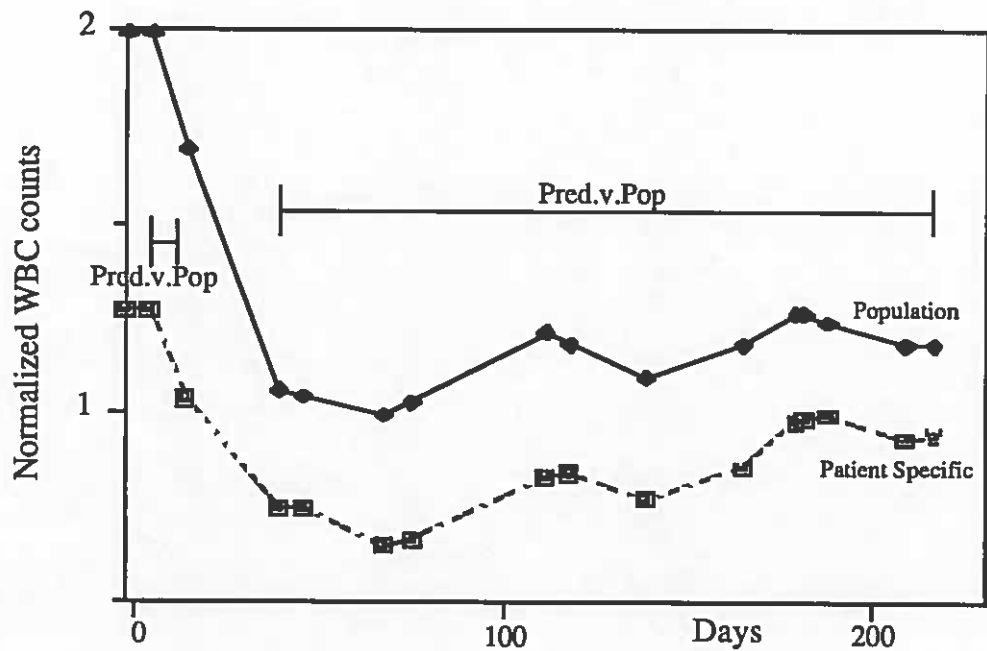
12

Figure 5: ETNODE intervals created by model prediction differences. Large differences between patient-specific predictions and population-based predictions causes the creation of an ETNODE with a Pred.v.Pop type-label. An ETNODE of this type represents an interval of time during which the patient's modeled clinical course differs from the expected (population-based) clinical course. Rules associated with Pred.v.Pop ETNODEs attempt to find an explanation for this discrepancy. Other rules attempt to critique the conclusion. In this figure, the dotted line plots the patient-specific model predictions, the solid line plots the population-based model predictions, and the intervals denote differences of sufficient magnitude to cause the creation of a Pred.v.Pop ETNODE.

Table 2: Pair-wise comparisons for summarization and explanation.

Summarization concept	Comparison used to derive concept	Notation
surprising observations	population-based model predictions vs patient-specific model predictions	Pred.v.Pop
explanatory states	population-based model parameters vs patient-specific parameter estimates	PopP.v.PredP
model critique	patient observations vs patient-specific model predictions	Obs.v.Pred

clinically. The purpose of the text-generation system is to convert abstractions represented as ETNET intervals into an output format acceptable to clinicians.

TOPAZ uses two presentation schemas to produce a clinical summary from ETNET abstractions: (1) an organizational schema that determines the overall logical structure of the summary text, and (2) a rhetorical schema that encodes the logical components of a clinical argument. The main purpose of the organizational schema is to provide overall coherence and structure to the summary text. The rhetorical schema is responsible for selecting features from the analysis that support a summary statement. Because there may be conflicting evidence present in the medical record, the rhetorical schema must be able to both support and critique each conclusion reached by TOPAZ.

The TOPAZ presentation model is implemented as a hierarchy of augmented transition networks (ATNs; Figure 6). An ATN is a collection of nodes connected by directed arcs. Nodes contain logical statements that determine which arc is to be traversed to reach the next node to be processed. A node or arc can contain actions that are executed only if that node or arc is traversed. ATNs have two features that are useful for prose generation: (1) nodes can be organized into hierarchical units, and (2) intermediate results can be shared among nodes and passed between network hierarchies. ATN hierarchies are useful for decomposing a task into logical subtasks,

where the initial ATN networks encodes the overall structure of the task and the SUBATNs encode specific details to perform a circumscribed component. For example, in text-generation, an ATN responsible for a specific paragraph may use various SUBATNs to construct individual sentences. These SUBATNs, in turn, may use addition SUBATNs to construct sentence fragments or phrases used in the sentences.

Miller has developed a text-generation program, called PROSENET, that uses the traversal of ATN nodes and arcs to control the production of text fragments (Figure 6) [15,16]. PROSENET uses ATN hierarchies to organize the text generation into logical segments, such as sentences, paragraphs, and main topics. This logical decomposition makes text generation with PROSENET easy to adapt to the organizational and rhetorical schemas in TOPAZ. Rennels reimplemented PROSENET using an object-oriented design [22,16]. The TOPAZ text-generation system is based on a simplified version of Rennels' system.

4 Limitations

TOPAZ uses both numeric and symbolic temporal models to interpret and abstract subtle clinical features that evolve over time. By combining a *processlike*, real-valued temporal model with an *eventlike*, interval temporal model, TOPAZ is able to represent and reason about clinical entities with markedly differing temporal characteristics. The TOPAZ methodology and implementation contain a number of design assumptions that can severely limit their use in certain tasks:

- The TOPAZ methodology assumes that an accurate structural model of the relationships among observations and underlying processes can be elucidated from experts. Many processes in medicine are not understood with sufficient precision to permit the expert to construct a detailed structural model. Qualitative modeling techniques probably are not the answer in this situation. In data-poor and error-prone settings, qualitative models do not provide sufficient inferencing capabilities to draw conclusions about unobserved system states. In the
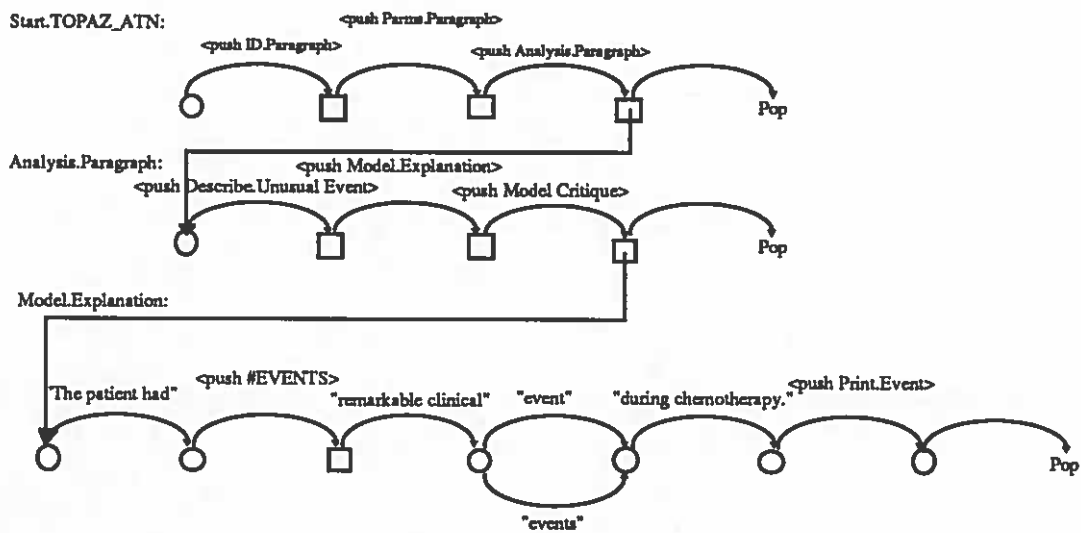
15

Figure 6: The PROSENET methodology for text generation. PROSENET generates text using hierarchies of augmented transition networks (ATNs). Nodes and arcs generate text fragments. Each node contains logical statements that determines which node to traverse next. Hierarchies of networks organize text generation into progressively larger units—for example, from sentences, to paragraphs, to topics. Squares denote ATN nodes that call sub-ATNs to generate specific text fragments.

absence of an accurate structural model, observations of interest could be encoded as rules in ETNODEs. This approach is the method used in VM and RX. It is equivalent to implicitly encoding a model as rules. In an area where good predictive models can be constructed, we believe encoding a structural model as rules is not the desired approach. In areas where no such structural model exists, ETNODE rules would be the only method available in TOPAZ to represent interesting temporal patterns.

- The TOPAZ temporal models represent future temporal planning sequences poorly. For example, there is no way to represent the phrase: "Patients with this cancer receive four cycles of treatment X, followed by two cycles of treatment Y, and, if they show improvement, continue to receive treatment X until all evidence of active tumor growth disappears." ETNET could represent the fact that a specific patient had received four cycles of treatment X followed by two cycles of treatment Y followed by n cycles of treatment X, but this would not capture the same temporal information as is contained in the quoted sentence. Musen et al. [19,18] describe a graphical language and an ATN-based representation for specifying procedural knowledge in cancer-chemotherapy protocols. Their techniques could encode the sequence information in the above example. TOPAZ only notes whether, given the therapy recorded in the medical record, the patient observations are unusual. Without an encoding of the *expected* course of chemotherapy, TOPAZ cannot detect when the dosage or timing of drugs was abnormal. Using the additional temporal sequencing knowledge contained in Musen's representation system, TOPAZ would be able to detect deviations from the expected therapeutic course.

- The TOPAZ temporal models represent temporal uncertainty poorly. The physiological model requires that the times of observations and therapies be known accurately. The TOPAZ bone-marrow model represents the presence of random error in patient observations, but it does not represent random error in the input (therapy) or temporal uncertainty. Statistical techniques

that incorporate these additional sources of uncertainty exist, but, in the data-poor TOPAZ environment, they are unlikely to be helpful. The TOPAZ interval model (ETNET) has no capability of representing temporal imprecision or uncertainty among events. Imprecision in the starting and ending times for events could be added by extending the definition of ETNET event-times to include a variance (or fuzz) term. We know of no simple extension that would support uncertain temporal relationships among events.

5 Comparison to Previous Research

Although temporal relationships are prominent in medicine, only a few computer-based decision support systems have focused on representing and reasoning with temporal concepts. VM was a rule-based expert system designed to interpret on-line physiological data using IF-THEN rules to derive clinical abstractions, such as hemodynamic stability, from patient data [6,7]. VM introduced the use of context-sensitive rules to convert primary observations into context-free symbolic abstractions. This conversion step was a key contribution of VM because it enabled the rest of VM's reasoning, which used only the clinical abstractions to derive therapy-management conclusions, to be context-free even though the underlying problem domain was clearly context-sensitive. TOPAZ extends the notion of context-sensitive reasoning by embedding rules within temporal intervals representing the clinical context in which these rules make sense. At any moment in time, multiple clinical contexts can be relevant, each with their own rules for reasoning about the current data.

RX combined symbolic and statistical methods to analyze sets of patient records for the presence of previously unknown causal relationships [2,3]. Although RX was concerned with a collection of patients, RX derived abstract interval events by examining data present in each patient visit, using a hierarchically arranged set of features. RX introduced the notion of using observations from one point in a patient's record to support a conclusion of a clinical abstraction during another one in time. For example, the presence of urinary protein at a visit 7 days ago could be used to

18

justify the clinical abstraction of nephrotic syndrome on the current visit, even if there was no value for urinary protein recorded for the current visit. TOPAZ has no equivalent feature because it uses an explicit physiological model to detect patterns among clinical features. Downs' extended the temporal abstraction capabilities of RX by incorporating the odds-ratio formulation of Bayes' formula to determine whether sufficient evidence was found in the record to support the conclusion of a clinical abstraction [4,5]. Each clinical abstraction had one or more predicates that represented evidence for or against the presence of the concept. If the rule succeeded, a posterior odds ratio updated the belief in the existence of the concept. With this extension, Downs' system could encode rules for both confirmatory and contradicting evidence. TOPAZ uses differences between the expected clinical course and the observed or patient-specific predicted clinical course to trigger abstractions. This technique is similar to Downs' odd-ratio threshold that had to be exceeded before an abstraction was concluded.

TOPAZ uses multiple temporal representations to encode differing views of time. Multiple symbolic temporal representations were the hallmark of Kenneth Kahn's TIME SPECIALIST [8] and Mittal's PATREC [17]. Unlike TOPAZ, neither the TIME SPECIALIST nor PATREC represented causal relationships between stored events. The combination of numeric causal models and symbolic temporal models was used in the DIGITALIS THERAPY ADVISOR [25] and in the HEART FAILURE PROJECT [12]. The DIGITALIS THERAPY ADVISOR combined a numeric model of digitalis kinetics with a symbolic model of digitalis toxicity. The HEART FAILURE PROJECT, which is an on-going research effort, combines a symbolic causal model of cardiovascular hemodynamics and linear differential equations. Unlike these programs, TOPAZ uses statistical techniques rather than heuristics to individualize the mechanistic model. In addition, neither program incorporates an interval-based view of time.

CASNET encoded physiological relationships in a symbolic causal network [28]. With this network, CASNET reasoned about state changes that are seen in glaucoma. Unlike TOPAZ, CASNET did not encode temporal relationships among state transitions. Thus, CASNET could

19

reason about changes in a patient's physiological status but could not discriminate between various states due to temporal differences.

6 Summary

The TOPAZ methodology uses three models to view the temporal aspects of a patient's record in order to summarize it. A structural model takes a continuous, process-oriented view; an interval-based model takes a discrete, event-oriented view; and a presentation model takes a state-based view. No single formalism would be sufficiently flexible to encode the temporal perspective embodied in the other models.

The explicit structural model is essential to the TOPAZ methodology. It is used to construct a patient-specific model from time-ordered observations, to define the concept of "unusual" observations, and to provide a plausible explanation for them.

Context limits the size of symbolic reasoning. In TOPAZ, an unusual finding triggers the search for a potential explanatory cause. Without the enabling abnormal context, the search for a cause would not (and should not) occur. Not all model features can contribute to a specific abnormal finding. Context-specific reasoning limits the search to only those model features that potentially could explain the lower-than-usual predicted drug level.

TOPAZ is a computer program that constructs textual summaries of context-sensitive, time-ordered data. The key themes of multiple temporal models, an explicit structural model, and context-specific reasoning are the centerpieces of the TOPAZ program.

Acknowledgments

20

pher Lane, Mark Musen, Janice Rohn, Samson Tu, and Cliff Wulfman, for making this project possible.

References

[1] BI Blum. *Clinical Information Systems*. Springer-Verlag, New York, NY, 1986.

[2] RL Blum. Discovery and representation of causal relationships from a large time-oriented clinical database: The RX project. In DA Lindberg and PL Reichertz, Editors, *Lecture Notes in Medical Informatics*, Volume 19. Springer-Verlag, New York, NY, 1982.

[3] RL Blum. Discovery, confirmation and incorporation of causal relationships from a large time-oriented clinical data base: The RX project. In WJ Clancey and EH Shortliffe, Editors, *Readings in Medical Artificial Intelligence: The First Decade*, Chapter 17, pages 399–425. Addison-Wesley, Reading, MA, 1984.

[4] SM Downs. A program for automated summarization of on-line medical records. Master's thesis, Medical Computer Science, Stanford University, Stanford, CA, 1986.

[5] SM Downs, MG Walker, and RL Blum. Automated summarization of on-line medical records. In R Salamon, B Blum, and M Jorgensen, Editors, *MEDINFO 86*, pages 800–804, Washington DC, 1986. North-Holland.

[6] LM Fagan. VM: Representing time-dependent relations in a medical setting. PhD thesis, Department of Computer Science, Stanford University, Stanford, CA, 1980.

[7] LM Fagan, JC Kunz, EA Feigenbaum, and JJ Osborn. Extensions to the rule-based formalism for a monitoring task. In BG Buchanan and EH Shortliffe, Editors, *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*, Chapter 22, pages 397–423. Addison-Wesley, Reading, MA, 1984.

[8] KM Kahn and GA Gorry. Mechanizing temporal knowledge. *Artificial Intelligence*, 9:87–108, 1977.

[9] MG Kahn. Model-based interpretation of time-ordered medical data. PhD thesis, Section on Medical Information Sciences, University of California–San Francisco, San Francisco, CA, 1988.

[10] MG Kahn, JC Ferguson, EH Shortliffe, and LM Fagan. Representation and use of temporal information in ONCOCIN. In MJ Ackerman, Editor, *Proceedings Symposium on Computer Applications in Medical Care*, pages 172–176, Baltimore, MD, 1985. IEEE Computer Society.

[11] CP Langlotz, LM Fagan, SW Tu, BI Sikic, and EH Shortliffe. A therapy planning architecture that combines decision theory and artificial intelligence techniques. *Computers and Biomedical Research*, 20:279–303, 1987.

[12] WJ Long. Reasoning about state from causation and time in the medical domain. In *Proceedings AAAI-83*, pages 251–254, Washington DC, 1983. Morgan Kaufmann.

[13] CJ McDonald, L Blevins, WM Tierney, and DK Martin. The Regenstrief medical records. *M.D. Computing*, 5:34–47, September/October 1988.

[14] KR McKeown. *Text generation: Using discourse strategies and focus constraints to generate natural language text.* Cambridge University Press, Cambridge, England, 1985.

[15] PL Miller. *A Critiquing Approach to Expert Computer Advice: ATTENDING.* Pitman, Boston, MA, 1984.

[16] PL Miller and GD Rennels. Prose generation for expert systems: An applied computational linguistics approach. *AI Magazine*, 9:37–44, Fall 1988.

[17] S Mittal, B Chandrasekaran, and J Sticklen. Patrec: A knowledge-directed database for a diagnostic expert system. *Computer*, 17:51–58, 1984.

[18] MA Musen. Generation of model-based knowledge-acquisition tools for clinical-trial advice systems. PhD thesis, Medical Computer Science Group, Stanford University, Stanford, CA, 1988.

[19] MA Musen, LM Fagan, and EH Shortliffe. Graphical specification of procedural knowledge for an expert system. In *Proceedings IEEE Computer Society Workshop on Visual Languages*, pages 167–178, Dallas, TX, 1986. The Computer Society Press. Reprinted in: *Expert Systems: The User Interface*; J Hendler, Editor; Ablex, Norwood, NJ.

[20] WH Press, BP Flannery, SA Teukolsky, and WT Vetterling. *Numerical Recipes: The Art of Scientific Computing.* Cambridge University Press, Cambridge, England, 1986.

[21] TA Pryor. The HELP medical record system. *M.D. Computing*, 5:22–33, September/October 1988.

[22] GD Rennels. A computational model of reasoning from the clinical literature. PhD thesis, Medical Computer Science Group, Stanford University, Stanford, CA, 1986.

[23] LB Sheiner and SL Beal. Bayesian individualization of pharmacokinetics: Simple implementation and comparison with non-Bayesian methods. *Journal of Pharmaceutical Sciences*, 71:1344–1348, 1982.

[24] LB Sheiner, H Halkin, C Peck, B Rosenberg, and KL Melmon. Improved computer-assisted digoxin therapy. A method using feedback of measured serum digoxin concentrations. *Annals of Internal Medicine*, 82:619–627, 1975.

[25] H Silverman. A digitalis therapy advisor. PhD thesis, Project MAC, Massachusetts Institute of Technology, Cambridge, MA, 1975.

[26] WW Stead and WE Hammond. Computer-based medical records; The centerpiece of TMR. *M.D. Computing*, 5:48–62, September/October 1988.

[27] SW Tu, MG Kahn, MA Musen, JC Ferguson, EH Shortliffe, and LM Fagan. Episodic monitoring of time-oriented data for heuristic skeletal-plan refinement. Technical Report KSL 87-70, Knowledge Systems Laboratory, Stanford University, Stanford, CA, 1988.

[28] SM Weiss, CA Kulikowski, S Amarel, and A Safir. A model-based method for computer-aided medical decision making. *Artificial Intelligence*, 11:145–172, 1978.

[29] QE Whiting-O'Keefe, A Whiting, and J Henke. The STOR clinical information system. *M.D. Computing*, 5:8–21, September/October 1988.