Knowledge-Based Modeling and Simulation of Diseases with Highly Differentiated Clinical Manifestations

Marjorie McShane[1], Sergei Nirenburg[1], Stephen Beale[1],
Bruce Jarrell[2], and George Fantry[2]

[1] University of Maryland Baltimore County,
1000 Hilltop Circle, Baltimore Maryland, 21250 USA
{marge,sergei,sbeale}@cs.umbc.edu
[2] University of Maryland School of Medicine,
655 West Baltimore Street Baltimore, Maryland 21201-1559, USA
BJarrell@som.umaryland.edu,
GFantry@medicine.umaryland.edu

Abstract. This paper presents the cognitive model of gastroesophageal reflux disease (GERD) developed for the Maryland Virtual Patient simulation and mentoring environment. GERD represents a class of diseases that have a large number of clinical manifestations. Our model at once manages that complexity while offering robust automatic function in response to open-ended user actions. This ontologically grounded model is largely based on script-oriented representations of causal chains reflecting the actual physiological processes in virtual patients. A detailed description of the GERD model is presented along with a high-level description of the environment for which it was developed.

Keywords: cognitive model, simulation, gastroesophageal reflux disease, virtual patient.

1 The Maryland Virtual Patient Environment

The Maryland Virtual Patient[1] (MVP) project is developing an agent-oriented environment for automating certain facets of medical education and certification. This environment is effectively a network of human and software agents, at whose core is a virtual patient – a knowledge-based model of a person with a disease. This model is implemented in a computer simulation. The virtual patient is a "double agent" that displays both physiological and cognitive function. Physiologically, it undergoes both normal and pathological processes in response to internal and external stimuli. Cognitively, it experiences symptoms, has lifestyle preferences, has memory (many of whose details fade with time), and communicates with the human user about its personal history and symptoms.

[1] Patent pending.

R. Bellazzi, A. Abu-Hanna, and J. Hunter (Eds.): AIME 2007, LNAI 4594, pp. 34–43, 2007.

Other software agents in the MVP environment include consulting physicians, lab technicians and a virtual mentor (tutor). What makes virtual patient modeling feasible – considering that comprehensively modeling human physiology would be a boundless endeavor – is our goal-oriented approach: we are not trying to recreate the human organism in all its details, we are modeling it to the extent necessary to support its realistic autonomous functioning in applications aimed at training and testing the diagnostic and treatment skills of medical personnel. Trainees can use the MVP simulation environment to interview a virtual patient; order lab tests; receive the results of lab tests from technician agents; receive interpretations of lab tests from consulting physician agents; posit hypotheses, clinical diagnoses and definitive diagnoses; prescribe treatments; follow-up after those treatments to judge their efficacy; follow a patient's condition over an extended period of time, with the trainee having control over the speed of simulation (i.e., the clock); and, if desired, receive mentoring from the automatic mentor.

The virtual patient simulation is grounded in an ontologically-defined model of human anatomy and physiology. Disease processes and treatments are modeled with a clear separation between direct and indirect effects. For example, if a trainee performs a Heller myotomy (a surgical procedure that cuts the lower esophageal sphincter (LES)) on *any* patient, whether or not his condition suggests the need for one, his basal LES pressure will decrease substantially, in most cases permitting an excessive amount of acid reflux per day and, over time, giving him gastroesophageal reflux disease (GERD).[2] To emphasize, there is no rule that says that a Heller myotomy turns a virtual patient into a GERD patient: the anatomical result of a Heller myotomy sets off a chain of events that turns most virtual patients into GERD patients. This level of anatomical and physiological automaticity permits the virtual patient to respond realistically even to completely unexpected and clinically incorrect interventions by the trainee. If the trainee launches such interventions, he must subsequently manage a more complex patient for the duration of the simulation.

Instances of virtual patients with particular diseases and particular physiological peculiarities are generated from core ontological knowledge about human physiology and anatomy by grafting a disease process onto a generic instance of human. Disease processes themselves are described as complex events in the underlying ontology.

This paper details the script-based modeling of diseases that are marked by distinctly different clinical manifestations – what we call *tracks*. We focus on the example of gastroesophageal reflux disease (GERD), which has six different tracks and a range of patient parameterization within each track. Disease modeling lies at the core of the MVP system, since the models must be both robust enough to support realistic function and constrained enough to be readily implemented. The models must also be generalizable enough to permit approaches and content modules to be reused in the modeling of new diseases. Our current disease models, which have been implemented and tested, fulfill all of these requirements.

---

[2] GERD would not arise if the Heller myotomy were incomplete or if the patient had achalasia, which could cause LES pressure to increase over time (see [1] for details of achalasia modeling).

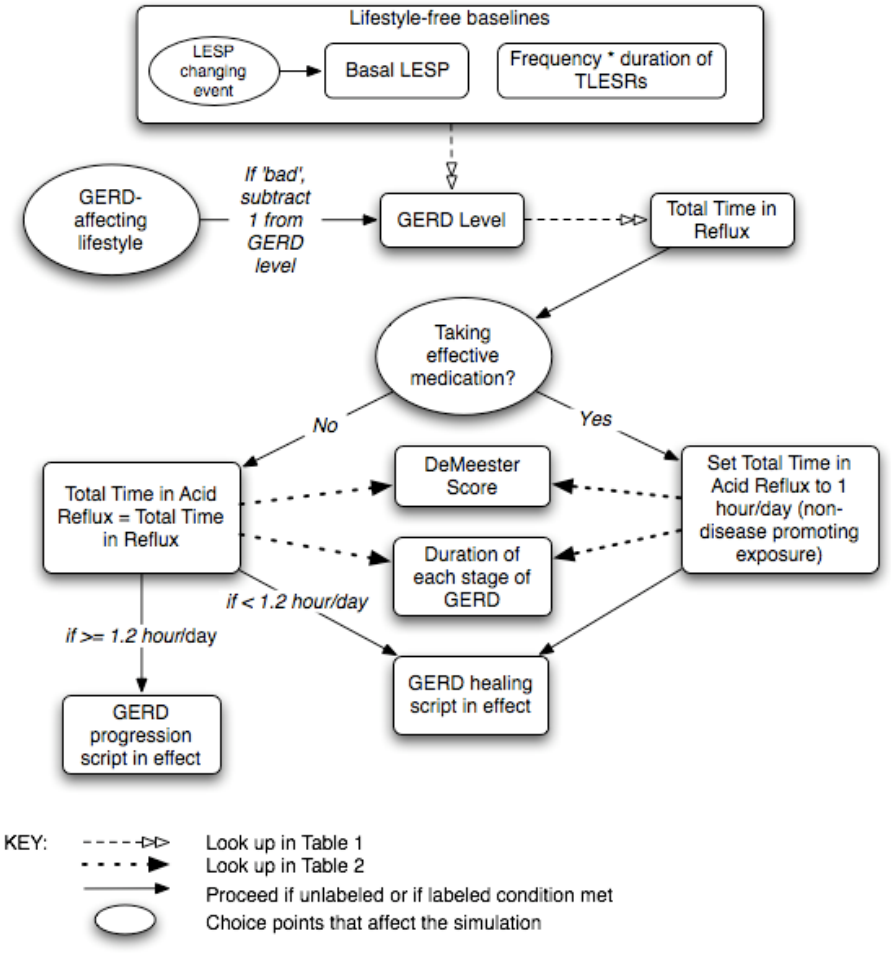

Fig. 1. A representation of how GERD-related property values are set and modified, and how the simulator determines whether GERD is progressing or healing

2 Multi-track Scripts: The Example of GERD

The MVP system currently covers six esophageal diseases: gastroesophageal reflux disease (GERD), laryngopharyngeal extraesophageal reflux disease (LERD), LERD-GERD (a combination of LERD and GERD), scleroderma esophagus, Zenker's diverticulum and achalasia. In this section we describe the knowledge-based model of one of these diseases, GERD, which can be defined as any symptomatic clinical condition that results from the reflux of stomach or duodenal contents into the esophagus. The two sources of GERD are abnormally low pressure of the lower esophageal sphincter (LES) (< 10 mmHg), or an abnormally large number or duration of transient relaxations of the LES (TLESRs), both of which result in increased acid

exposure of the esophageal lining. Both basal LES pressure (LESP) and TLESRs can be negatively affected by lifestyle habits such as consuming caffeine, chocolate or fatty foods.

A person can become a GERD patient in four ways: (a) by having an inherent predisposition to low LESP or excessive TLESRs; (b) by engaging in lifestyle habits that negatively impact LESP or TLESRs; (c) as a complication of another disease, like scleroderma esophagus, which decreases LESP; or (d) as a complication of an outside event or intervention, like a Heller myotomy, which results in a hypotensive LES.

The top levels of Figure 1 show the factors contributing to the inception of GERD. The severity of the GERD-producing factors is reflected by the attribute "GERD level", which was introduced to unify the model, abstracting away from which specific LES-related abnormality gave rise to the disease. The lower the GERD level, the higher the daily esophageal acid exposure and the more fast-progressing the disease. The reason for associating a low GERD level with severe GERD is mnemonic: the GERD levels are the same as the basal LESP for patients who have low-pressure GERD. For example, a patient with a LESP of 1 mmHg will have a GERD level of 1. If a patient has a GERD level of 1 due to TLESRs, that means his daily esophageal acid exposure from the transient relaxations is the same as it would have been if he had had a basal LESP of 1.

Using GERD level as the anchor for modeling provides a simple mechanism for incorporating a patient's lifestyle habits into the simulation: whenever he is engaging in bad lifestyle habits (assuming he has GERD-related sensitivities to those habits), his GERD level decreases by 1. For patients with a baseline GERD level of 10 – which is not a disease state – this means that engaging in bad habits is sufficient to initiate GERD and discontinuing them is sufficient to promote healing without the need for medication. For patients with a baseline GERD level of less than 10, lifestyle improvements can slow disease progression but not achieve the healing of previous esophageal damage. Once a patient's GERD level is established, total time in reflux can be determined from Table 1.

Table 1. Property-value correlations for GERD, Part I (an excerpt)

LESP or the equivalent number of TLESRs per day	GERD level	Total Time in Reflux (in hours and percentage of time per day)
10-15	10	1.2 hrs. {5%}
9	9	1.56 hrs. {6.5%}
8	8	1.92 hrs. {8 %}
…	…	…
1	1	5.28 hrs.{22%}
0	0	6.0 hrs. {25%}

However, it is not total time in reflux that actually causes GERD, it is total time in *acid* reflux.[3] The acidity of the reflux can be decreased (i.e., pH can be increased) by taking medication. If effective medication is taken, the patient's acid exposure is set

[3] New impedance testing permits the detection of GERD due to non-acid reflux, with this rare condition not currently being covered by the MVP system.

to 1 hour per day, which puts the patient in a healing state.[4] If no effective medication is being taken, then total time in acid reflux equals total time in reflux. Total time in acid reflux determines the rate of disease progression (i.e., duration of each conceptually delineated stage of the disease), as well as the patient's DeMeester score[5], as shown in Table 2. If the patient shows incomplete compliance, the simulator can switch him from a healing state to a disease state with great frequency, reflecting every instance of a taken or missed dose.

Table 2. Property-value correlations for GERD, Part II (an excerpt)

GERD level (for orientation)	Total Time in Acid Reflux	Stage Duration	DeMeester Score
10	<= 1.2 hrs. {5%}	na	10
9	1.56 hrs. {6.5%}	180 days	20
8	1.92 hrs. {8 %}	160 days	25
…	…	…	…
1	5.28 hrs. {22%}	40 days	80
0	6.0 hrs. {25%}	30 days	120

GERD can follow any of six clinical manifestations, or tracks:

1. non-erosive GERD, for which inflammation of the esophageal lining is the ending point of the disease
2. GERD with erosive esophagitis leading to erosion(s) but not ulcer(s)
3. GERD with erosive esophagitis leading to ulcer(s)
4. GERD with erosive esophagitis leading to peptic stricture
5. GERD with Barrett's metaplasia
6. GERD progressing past Barrett's metaplasia to adenocarcinoma

Which of these tracks a patient's disease will follow is determined by inherent predispositions. This means that no matter how long a track 1 patient experiences GERD, the disease will never progress past the level of inflammation, whereas if a track 4 patient remains untreated long enough, he will get a peptic stricture. Having a predisposition to later stages of GERD does not, however, necessitate experiencing those complications: ongoing effective treatment can reverse the disease course and ensure esophageal health indefinitely.

All GERD patients are assigned a set of GERD-related predispositions which determine how many and which stages of the disease they will experience. These predispositions can be asserted by patient authors (see below for patient authoring) or can be automatically set by the simulator for cases in which GERD arises spontaneously due to complications of a different disease or intervention.

Disease progression is modeled as changes over time of (a) physiological and symptom-related property values, and (b) the frequency, intensity, etc., of various

[4] It was decided that no pedagogical benefit would be gained by making the total time in acid reflux while on medication a variable: the important point is that the acid exposure is low enough to permit healing.

[5] DeMeester score is a complex reckoning of many factors measured during pH monitoring.

simulated events, like regurgitation. In addition, new objects (such as ulcers or tumors) can be created, modified and destroyed over time.

The simulation is driven by causal chains when they are known and are deemed useful to the goals of the simulation; otherwise, temporally oriented "bridges" reflecting clinical knowledge drive the simulation (e.g., if the disease has reached day 240 of untreated progression, the value of property *x* will be *y*).

2.1 Example: GERD with Erosive Esophagitis Leading to Erosion(s)

We illustrate the progression of GERD using the disease track "GERD with erosive esophagitis leading to erosion(s)", which is conceptually divided into three stages:

1. **preclinical GERD**, during which the value of the property *preclinical irritation percentage* (whose domain is *mucosa of distal esophagus*) increases from 0 to 100. When the *preclinical irritation percentage* reaches 100, the script for the preclinical stage is unasserted, with the simultaneous assertion of the script for
2. **the inflammation stage of GERD**, during which the mucosal layer of the esophageal lining is eroded, going from a depth of 1 mm. to 0 mm. over the duration of the stage. When mucosal depth reaches 0 mm., the script for the inflammation stage is unasserted, with the simultaneous assertion of the script for
3. **the erosion stage of GERD**, at the start of which an erosion object is created whose depth increases from .0001 mm. upon instantiation to .5 mm. by the end of the stage, resulting in a decrease in submucosal depth from 3 mm. to 2.5 mm. When submucosal depth has reached 2.5 mm. the script remains in a holding pattern since this patient does not have a predisposition to ulcer.[6]

Over the course of each stage, property values are interpolated using a linear function (though other functions could, in fact, be used). So halfway through the preclinical stage the patient's "irritation percentage" will be 50, and ¾ of the way through that stage it will be 75.

The length of each stage depends upon the patient's total time in acid reflux (cf. Table 2): e.g., for a patient with a total time in acid reflux of 1.92 hours a day, each stage will last 160 days. Such a patient will either have a baseline GERD level of 8 and not be engaging in bad habits, or have a baseline GERD level of 9 and be engaging in bad habits, with the bad habits "demoting" him to an effective GERD level of 8 (cf. Table 1 and Figure 1).

The experiencing of symptoms varies widely across patients but a fixed inventory of symptoms is associated with each disease, and expected ranges of values for each symptom can be asserted for each stage. For symptoms with abstract values ("On a scale of 1 to 10…"), we use the scale {0,1}, with decimals indicating intermediate values.

---

[6] Had the patient had a predisposition to ulcer, reaching a submucosal depth of < 2.5 mm. would have unasserted the erosion stage script and asserted the ulcer stage script, leading to the creation of an ulcer object and its increase in size over time.

Table 3 shows the symptom profile table for "GERD with erosive esophagitis leading to erosion(s)". The ranges shown indicate the possible values for each symptom by the end of the given stage, with default values being shown in parentheses. For example, the patient Beatrice Thompson might experience heartburn severity at the average (default) level of .4 by the end of the inflammation stage, but might experience it at a higher than average level, .8, by the end of the erosion stage. All values between these "stage end" points are interpolated using a linear function. No symptoms are experienced in the preclinical stage, reflecting the definition of a preclinical disease.

Table 3. Symptom profile table for one track of GERD

Stage	Preclinical	Inflammation	Erosion
Heartburn freq. (times per day)	0	3 – 5 (4)	6 – 8 (7)
Heartburn severity	0	.3 – .5 (.4)	.6 – .8 (.7)
Regurgitation (times per week)	0	3 – 5 (4)	6 – 8 (7)
Symptom correlation	0	0 – 1 (random .5 < > .8)	0 – 1 (random .5 < > .8)

2.2 Interactions and Interventions

What we have shown so far is the knowledge needed to simulate GERD, assuming no external interventions. However, external interventions are key to an educationally-oriented simulation, as are realistic patient responses to interventions, be they clinically appropriate (expected) or clinically inappropriate (unexpected).

The MVP system supports verbal interaction with the virtual patient as well as two types of interventions: diagnostic tests and treatments. Questioning the patient and carrying out diagnostic tests are similar in that the system must (a) interpret the question or the request for a test, (b) look up the appropriate physiological or symptom-related property values or stored events in the database populated during the simulation, (c) return a response as an English string. Currently, since we have not yet plugged in natural language processing (NLP) capabilities (see below), steps (a) and (c) are handled by menus and preconstructed strings, respectively. However, once NLP is incorporated, these steps will rely on the language understanding and generation capabilities of the virtual patient, lab technicians and outside specialists.

Typical tests carried out when GERD is suspected are:

- esophagogastroduodenoscopy (EGD), which returns information about the presence of inflammation, erosion, ulcer, etc.
- pH monitoring, which returns time in acid reflux, symptom correlation (the correlation between a patient's symptoms and the acidity of the distal esophagus), and DeMeester score (cf. footnote 5)
- barium swallow, which can be used to detect tumors.

All positive and pertinent negative results are returned. These tests contribute to a definitive diagnosis of GERD, although a clinical diagnosis can be made based on the efficacy of drugs in reducing symptoms. If tests that are atypical for GERD are launched on a GERD patient, the relevant property value(s) will be sought and, if no

abnormal values are found, the return value will be normal. In this way, the MVP system is always prepared for unexpected moves by users.

As concerns treatment, the typical treatment options for GERD are lifestyle modifications in combination with H2 blockers or PPIs (QD or BID).[7] Lifestyle modifications can be completely effective only for patients with a baseline GERD level of 10 and bad habits that they succeed in overcoming. All other manifestations of GERD require medication, and medications can be effective or ineffective for different patients, as recorded in their physiological profiles.

Treatments administered in the MVP environment can improve the patient's condition, be ineffective, or be detrimental. For example, whereas PPI BID typically works for a patient with a GERD level of 5, H2 blockers typically do not work and a Heller myotomy will certainly be detrimental. Since Heller myotomy is not an expected treatment for a GERD patient, there is nothing in his profile to indicate how he will respond to this surgery; instead, the ontologically defined default outcome is used, which is reduction of the basal LESP to 2 mmHg.

If an effective treatment is administered, the GERD progression script for the current stage (e.g., "erosion") is halted and the "heal GERD" script is launched, improving the physiological and symptom profiles over an appropriate amount of time depending upon the current stage of the disease. For example, healing that starts during the inflammation stage can take up to 4 weeks (depending on how far into the inflammation stage the patient has progressed), while healing that starts during the erosion stage takes up to 8 weeks.

It should be clear that many aspects of GERD are inherent parts of the disease model that cannot be modified by patient authors. For example, a patient with a GERD level of 2 will have a faster-progressing disease than a patient with a GERD level of 9, assuming no effective treatment. The correlation between the amount of acid exposure and the rate of disease progression was considered a fundamental aspect of the disease by the physicians involved in building the model, who saw no practical or pedagogical reason to permit patient authors to override the set correlations between GERD level and stage duration.[8] By contrast, there are practical and pedagogical reasons to permit different patients to display other kinds of differences provided for in the patient authoring process.

2.3 Patient Authoring

Patient authoring involves filling in a one-page on-line questionnaire that permits domain experts or teachers to create a population of patients that display clinically relevant variations on the disease theme. For GERD, the population of patients must include patients with predispositions to each of the six disease tracks as well as different GERD levels, sources of GERD (low LESP vs. TLESRs), food sensitivities,

[7] We have not yet included reflux surgery in the system's repertoire. See [1] for a description of achalasia, whose inventory of treatment options and their possible outcomes is more complex.

[8] For other diseases, such as achalasia, stage correlation can be set explicitly by patient authors. This reflects the fact that the causal chains for such diseases are not as well understood, so there is no physiological variable from which to derive the differences in stage durations among patients that are clinically observed.

symptom profiles, reactions to medications, and compliance to treatment. Once these choices are made, the simulation is fully prepared for interactive use.

Disease scripts can be considered the central axis of the MVP environment, since all interactive and educational functionalities depend upon the ability of the disease scripts to support a robust and realistic simulation. In fact, the emphasis in system development thus far has been on disease scripts, with automatic mentoring (described in [1]) being added just recently, and natural language support not yet integrated, although the latter is the forte of our group, which has spent the past twenty years working in the field of knowledge-based natural language processing.

3 Comparisons with Other Systems and Approaches

The benefits of simulation in medicine have been widely propounded.[9] To give just one example, Satish and Streufert [3] write: "We need to ensure that medical personnel have the factual content knowledge needed to respond to the task at hand, but we also need to make sure that they can respond to complex challenges by processing information optimally. Simulations, if used as part of an appropriate training system, provide an optimal opportunity to acquire both."

Although simulation using live actors has long been a part of medical training, computer-based simulation that provides sufficient authenticity is still in its infancy, particularly as regards simulation of decision-making tasks via cognitive modeling (as contrasted with mannequins trainers for the development of motor skills). As many quality overviews of simulation already exist in the literature (e.g., [4], [5]), we limit our comparisons to two systems that are directly comparable to the MVP along at least one parameter.

CIRCSIM-Tutor [5] is a system dedicated to training medical students about the baroreceptor reflex. Over the course of the system's history, it has shifted from being a dynamic mathematical model that could be interacted with by students but provided no tutoring (MacMan), to being a tutoring system that no longer relies on the mathematical model, instead using a set number of statically stored cases. In terms of overall pedagogical goals – expediting medical education through automatic mentoring using natural language – the CIRCIM and MVP systems are similar. However, whereas the CIRCSIM group moved away from using a live physiological model, the MVP group is committed to developing a robust physiological model that can permit many types of interaction by trainees. For example, whereas one user might prefer to enable the tutor immediately in order to more quickly learn to avoid mistakes, another might prefer to learn by trial and error. Both methods being equally supported by the interactive simulation.

Another striking difference between CIRCSIM and MVP is the approach to NLP. Whereas CIRCISM essentially operates in terms of strings, MVP will carry out the knowledge-rich text processing developed in the theory of Ontological Semantics [6] and implemented in the OntoSem system. Using OntoSem, text strings are automatically translated into interpreted ontological concepts that are combined in logically correct ways to form text-meaning representations. It is over these

[9] Previously we have reported on the potential of this work to further medical pedagogy [2].

text-meaning representations that the language-based reasoners operate. Notably, the same ontological substrate is used both for medical modeling/simulation and for language processing, leading to a highly integrated knowledge base for the MVP system.

A knowledge-based approach to modeling virtual patients to support recertification has been reported in [7,8,9]. The authors describe both the modeling of the patient simulation process and the task of creating knowledge to support such a system. They also consider the situation of disease co-occurrence, which the MVP system also includes (e.g., scleroderma esophagus can have GERD as a side effect, with both continuing their courses simultaneously). One of the many major distinctions between this approach and our project is that Sumner *et al.* create probabilistic models using Bayesian networks and modified Monte Carlo methods, while our approach stresses causal modeling and guided creation (authoring) of virtual patients.

Initial testing of the MVP system was carried out over six hours with 40 second and third year medical students from the University of Maryland School of Medicine. Among the salient observations were that students followed a pattern of evaluation and management that is parallel to the actual process used in clinical care, they managed patients as if their actions had real consequences, and they realized the importance of "observation over time" as a diagnostic maneuver. More extensive evaluation is planned for the near future.

References

1. McShane, M., Jarrell, B., Nirenburg, S., Fantry, G., Beale, S.: Training clinical decision making using cognitively modeled virtual patients (Forthcoming)
2. Jarrell, B., Nirenburg, S., McShane, M., Fantry, G., Beale, S., Mallott, D., Raczek, J.: An interactive, cognitive simulation of gastroesophageal reflux disease. In: Proceedings of medicine meets virtual reality 15 in vivo, in vitro, in silico: Designing the next in medicine, February 6-9, 2007, Long Beach, California (2007)
3. Satish, U., Streufert, S.: Value of a cognitive simulation in medicine: towards optimizing decision making performance of healthcare personnel. Qual. Saf. Health Care 11, 163–167 (2002)
4. Streufert, S., Satish, U., Barach, P.: Improving medical care: The use of simulation technology. Simulation & Gaming 32(2), 164–174 (2001)
5. Evens, M., Michael, J.: One-on-One Tutoring by Humans and Computers. Lawrence Erlbaum and Associates, Publishers, New Jersey and London (2006)
6. Nirenburg, S., Raskin, V.: Ontological Semantics. MIT Press, Cambridge, Mass (2004)
7. Sumner, W., Marek, V.W., Truszczynski, M.: Data transformations for patient simulations. In: Proceedings of the 19th annual symposium for computer applications in medical care, p. 970 (1995)
8. Sumner, W., Marek, V.W., Truszczynski, M.: A formal model of family medicine. Journal of american board of family practice 9, 41–52 (1996)
9. Marek, V.W., Sumner, W., Truszczynski, M.: Creating evolution scenarios for hybrid systems. In: Proceedings of IEEE-SMC Symposium on Control, Optimization and Supervision, at CESA96, Lille, pp. 512–516 (1996)