Learning Components of Computational Models from Texts*

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

1 **Rensselaer Polytechnic Institute**
Troy, NY, USA
{mcsham2, nirens}@rpi.edu

2 **University of Maryland, Baltimore**
Baltimore, MD, USA
bjarrell@umaryland.edu, Gfantry@medicine.umaryland.edu

Abstract

The mental models of experts can be encoded in computational cognitive models that can support the functioning of intelligent agents. This paper compares human mental models to computational cognitive models, and explores the extent to which the latter can be acquired automatically from published sources via automatic learning by reading. It suggests that although *model components* can be automatically learned, published sources lack sufficient information for the compilation of fully specified models that can support sophisticated agent capabilities, such as physiological simulation and reasoning. Such models require hypotheses and educated guessing about unattested phenomena, which can be provided only by humans and are best recorded using knowledge engineering strategies. This work merges past work on cognitive modeling, agent simulation, learning by reading, and narrative structure, and draws examples from the domain of clinical medicine.

1998 ACM Subject Classification I.6 Simulation and Modeling

Keywords and phrases cognitive modeling, simulation, clinical medicine, learning by reading

Digital Object Identifier 10.4230/OASIcs.CMN.2015.108

1 Introduction

New scientific findings are being published much faster than domain experts can read or developers of intelligent systems can integrate. One way to address this information onslaught is through automation: by configuring intelligent agents that engage in lifelong learning by reading. Ideally, such agents will initially be endowed with a cognitive model corresponding to the models held by domain experts; then, as the agents read new texts, they will compare the information reported in those texts to the current state of their cognitive model, incorporating time-stamped, source-stamped updates into the model. Agents thus modified will not only themselves show increasingly sophisticated behavior, they will be able to pass on this learning to both people and intelligent systems via updating applications. Although a human-quality realization of this vision is not achievable overnight, learning by reading *is* realistic and can be pursued in a way that offers benefits in the near-, mid- and long-terms.

* This research was supported in part by Grant N00014-09-1-1029 from the U.S. Office of Naval Research. Any opinions or findings expressed in this material are those of the authors and do not necessarily reflect the views of the Office of Naval Research. Thanks to Mark Finlayson for useful conversations leading to our planned joint work.

6th Workshop on Computational Models of Narrative (CMN'15).
Editors: Mark A. Finlayson, Ben Miller, Antonio Lieto, and Remi Ronfard; pp. 108–123

OpenAccess Series in Informatics
OASICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

In this paper, we explore the nature of computational cognitive models that are sufficient to support the physiological and cognitive simulation of human-like intelligent agents, as developed for a prototype virtual patient application. We describe how these models, like the human mental models that underlie them, are comprised of a data-attested sketch filled in by clinical reasoning and educated guessing. We show how automatic learning by reading has the potential to automate the acquisition and updating of the data-attested portions, but argue that the backbones of the models – which derive of largely unwritten human expertise – are still best crafted manually.

The clinical models of diseases to be discussed here have features both of scripts (in the Schankian sense [28]) and of narratives, which informs how we approach the task of learning by reading.

Like *scripts*, the models record typical sequences of events and the objects that participate in them. They also allow for extensive individualization of the dynamically simulated cases based on two factors: (1) the physiological, psychological, emotional and circumstantial features of each virtual patient instance, and (2) the "moves" of the virtual patient and the clinician with respect to diagnosis, treatment and patient lifestyle, which can be undertaken at any point in the patient's simulated life. While selecting individualizing features for each virtual patient leads to some aspects of determinism in the simulation, much of the simulation is open-ended because the moves of the live clinician interacting with the virtual patient are not known beforehand and can fundamentally change patient outcome.

Like *narratives*, clinical disease models involve a non-trivial – in fact, sometimes life-and-death – plot. Ideally, the patient and clinician cooperate to cure the patient, but conflict can also occur: e.g., the virtual patient can choose to lie to the doctor to cover up non-compliance with a treatment protocol, or it can refuse medical intervention due to its personality traits or phobias [14]. Although, from a developer's point of view, such behavior is expected (the virtual patient will have been endowed with personality traits giving rise to this behavior), from the point of view of a system user, such outcomes are expected to be viewed as unexpected plot elements.

At the junction of script and narrative are two additional features of our clinicial disease models. First, the models include attested but atypical – i.e., story-worthy – events. In fact, one of the motivating factors in developing this virtual-patient-oriented clinician training system was to expose medical trainees to the broadest possible set of disease manifestations during a temporally compact training experience. The second script-narrative bridge derives from the constant influx of newly reported medical knowledge that must be incorporated into the models. Such new findings, which are often reported in case studies, are similar to the unexpected plot twists of narratives which, once encountered, must be recorded as modifications to scripts.

Our goal of learning by reading involves the automatic detection of such new information, particularly from case studies, and its seamless incorporation into the core disease models. An enabling factor is the canonical plot-like structure of case studies, which provide summarized background knowledge supplemented by the plot twist of an unexpected patient experience.

The work reported here dovetails with several programs of research and development. Our focus on the medical domain reverberates with Sileno et al.'s [29] focus on the legal domain, and they, like us, seek to ultimately support automatic knowledge acquisition from narrative; however, whereas our work involves a formal knowledge base, language processing, and agent simulation, Sileno et al.'s contribution is at a more theoretical level. O'Neill and Riedl [27] and Finlayson [4] both present methods of generating narrative structures using a manually annotated corpus as input. Whereas O'Neill and Riedl do not commit to any particular

knowledge representation formalism, Finlayson does, and uses it in the implementation of his Analogical Story Merging algorithm. Lieto and Damiano [6] discuss methods of detecting minimally different roles of participants in a narrative, such as hero vs. antihero. This aligns in spirit with our goal of detecting minimal differences between our disease models and the minimally different information presented in medical case studies. In terms of the ontologically-grounded modeling of complex events, the work of Schank and Abelson [28] was an early influence for the Theory of Ontological Semantics [21] that underpins the work reported here.

The paper is organized as follows. Section 2 sets the stage with an overview of the prototype medical teaching application – Maryland Virtual Patient (MVP) – that gave rise to our methodology of cognitive modeling. Section 3 draws a four-way comparison between human mental models, manually compiled cognitive models, the model components that can be semi-automatically elicited from human experts, and the model components that can be extracted from texts. Based on this comparison, we suggest a practical balance of effort between manual, semi-automatic and automatic knowledge acquisition strategies in support of agent configuration. Section 4 provides an overview of computational cognitive modeling in the OntoAgent environment, including excerpts from a disease model that successfully supported agent simulation in the MVP application. Section 5 describes how model components can be learned from texts, particularly by exploiting the predictable structure of genres such as case studies and disease overviews. Section 6 concludes the paper with the broader implications of this program of R&D.

2 The Maryland Virtual Patient (MVP) Application

Our modeling strategy developed during work on the prototype Maryland Virtual Patient (MVP) clinician training application [8] [9] [10] [13] [14] [22] [25] [26]. MVP is an agent-oriented system for automating certain facets of medical education and certification. It includes a network of human and software agents, at whose core is a virtual patient – a knowledge-based model of a person suffering from one or more diseases. The virtual patient is a "double agent" in that it displays both physiological and cognitive function. Physiologically, it undergoes both normal and pathological processes in response to internal and external stimuli, and shows realistic responses both to expected and to unexpected interventions; so if a trainee launches an inappropriate (unexpected) treatment, the patient's state will not improve and may even deteriorate, in which case the trainee must attempt to recover from his mistake.[1] Cognitively, the virtual patient experiences symptoms, has lifestyle preferences, can communicate with the human user in natural language, has memories of language interactions and simulated experiences, and can make decisions based on its knowledge of the world, its physical, mental and emotional states, and its current goals and plans. An optional tutoring agent provides advice and feedback to the trainee during the simulation.

Development of MVP follows the demand-side approach, meaning that it seeks to address a problem (detailed in [30]) that needs a solution rather than a problem that can be easily solved using standard methods (the supply-side approach). The specific problem MVP addresses is that medical educators, current training literature and pedagogical practice cannot provide medical students with adequately broad and varied training in cognitive analysis and problem solving. MVP seeks to permit trainees to diagnose and treat a large

[1] Compare this dynamic behavior with the static options in educationally-oriented branching scenarios that have also been called "virtual patients".

number of patient cases in a short amount of time, with the expectation that training results would mirror those of the SHERLOCK II electronic troubleshooting system for F16 aircraft of the US Air Force: participants using SHERLOCK II are reported to have learned more in 20 hours of tutoring than in 4 years of field experience [2].

Although many different paradigms of research and development involve entities called "virtual patients" (defined as mannekins, live actors, or branching scenarios), only MVP involves a knowledge environment that can support the approach to automatic lifelong learning described here. Key to this knowledge environment is reuse of the same knowledge representation language and static knowledge resources to support the wide range of agent functionalities described above [15]. Our prototype system has demonstrated that this AI-oriented, knowledge-based approach goes beyond theoretical status: we have worked out the details of knowledge representation and processing in implementations using realistic subject matter.

3 The Nature of Models

In this section we consider, in turn, human mental models, manually crafted computational cognitive models that seek to encode them, and the extent to which semi-automatic and automatic knowledge acquisition methods can realistically contribute to the computational modeling enterprise.[2]

Human mental models. Human mental models develop from a combination of experience, reading facts and stories, being told facts and stories, hypothesizing, reasoning, and even misrembering and forgetting. Although this wealth of contributors seems obvious, it is brought into relief when, as a non-specialist, one attempts to build a comprehensive computational model using only one of these sources as input: published texts. When working on modeling diseases and clinicial practices for MVP, the insufficiency of a "text-only" approach was immediately evident. Some gaps in knowledge represent facts that are actually not known because they are never measured: e.g., the physiological manifestations of the pre-clinical (non-symptomatic) stage of a disease. Other gaps reflect information that is not published in the literature for a *given* disease because it represents a broader generalization: e.g., a large tumor begins as a small tumor. Still other gaps reflect details that are not needed clinically (and are probably not known) but must be asserted if a realistic end-to-end simulation is to be implemented: e.g., does medication M, which ultimately cures disease D, improve property values at a steady rate or according to some non-linear function? The point is that humans somehow fill in these gaps sufficiently – albeit with a certain degree of uncertainty – to permit them to practice medicine effectively; and if they can do it, so must intelligent agents tasked with carrying out tasks requiring human-level reasoning.

Manually compiled, computational cognitive models. To develop computational cognitive models that were sufficient to support realistic patient simulations in MVP, a knowledge engineer led physican-informants through the process of distilling their extensive and tightly coupled physiological and clinical knowledge into the most relevant subset and expressing it in the most concrete terms. Not infrequently, specialists were also called upon to hypothesize about the unknowable, such as the preclinical stage of a disease and the values of physiological properties between the times when tests are run to measure them. Such hypotheses are

[2] See [5] for relevant discussion of manual vs. semi-automatic ontology development.

by nature somewhat vague, and could differ from expert to expert. However, rather than permit this imprecision to grind agent building to a halt, we proceed in the same way as live clinicians – and, presumably, any domain experts – do: by configuring *a* model that is reasonable and useful, with no claims that it is the only model possible or that it precisely replicates human functioning (cf. [1] for a discussion of modeling in the philosophy of science).

Decisions regarding what to include in our models derived from five desiderata: (1) that the models support realistic, interactive simulations; (2) that they not be unnecessarily detailed – i.e., if a detail would not be manifest in simulation (e.g., the firing of individual nerves), it was not included; (3) that they be easily updated to reflect new research findings; (4) that they be inspectable and explanatory, to support the pedagogical goals of the environment; and (5) that they be incorporated into an ontologically-grounded knowledge enviroment that supports all functionalities of all agents.

Taking these desiderata into account, and working within the OntoAgent cognitive architecture [15], we model diseases using an inventory of salient parameters whose values change over time in response to both internal stimuli (i.e., what the body does) and external stimuli (i.e., what the patient, doctor or outside world does). The selection of parameters to be included in a disease model is guided by practical considerations. Parameters are included because (a) they can be measured by tests, (b) they can be affected by medications or treatments, and/or (c) they are central to a physician's mental model of the disease. In addition to using parameters that directly reflect medically attestable properties, we also include abstract parameters that foster the formulation of a compact, comprehensible model (see Section 4 for examples).[3] Such features are particularly important at this stage of the discussion because they reflect the creative, unattested, aspect of computational modeling that naturally lies beyond automatic knowledge extraction methods since the information cannot be found explicitly in texts.

However, even if human reasoning is needed to build the more creative, hypothesis-driven aspects of computational models, the more concrete aspects can be aquired in semi-automatic and automatic ways, and it is to those that we now turn.

Semi-automatically acquirable model components. Since the collaboration between knowledge engineers and specialists is labor-intensive, the question arises, *To what extent can automation foster the process?* One way in which we experimented with reducing labor was by configuring a prototype knowledge elicitation system, called OntoElicit, to guide specialists through the process of independently recording "the basics" as preparation for work with a knowledge engineer [24]. The output of this work would then serve as input to the collaborative effort.

OntoElicit asks a domain expert to divide the given disease into conceptual stages correlating with important events. (The most obvious example of disease staging involves cancer, with its well-known stages 1 through 4; however, not all diseases are described in the literature as having a fixed inventory of stages.) Next, the system leads the expert through the process of providing – in a semi-formal way, guided by templates – details about disease progression, diagnosis and treatment. For example, when describing *physiology* and *symptoms*, the expert provides the inventory of properties that change over time, their start value before the disease begins and their expected values at end of each conceptual stage. Most values are recorded as a range of values covering different individual patients in the

[3] These features can be likened to the inclusion of intermediate categories in ontologies: although one does not typical talk about WHEELED-AIR-VECHILEs, this can be an appropriate node in an ontology.

population along with a default representing the most typical value. When describing *test results*, the expert indicates (a) which physiological properties are measured by each test, (b) any results that cannot be directly measured from the physiological model – e.g., visual findings by the administrator of the test, and (c) a "specialist's interpretation" of what the test results returned at that stage would indicate –e.g., "Suggestive of disease X." For *interventions* (medications, lifestyle changes, surgery, etc.), the expert indicates (a) which properties and/or symptoms are affected by the intervention, (b) the possible outcomes of the intervention, (c) possible side effects, and (d) if known, the percentage of the population expected to have each outcome and side effect. And for *diagnosis* and *treatment*, the expert provides fillers for ontological properties such as SUFFICIENT-GROUNDS-TO-SUSPECT (the given disease), SUFFICIENT-GROUNDS-TO-DIAGNOSE and SUFFICIENT-GROUNDS-TO-TREAT.

As mentioned earlier, the information acquired through OntoElicit is better described as *model components* than full models, since (a) some of the conceptual glue needed to hold the model together – most notably, causal chains – is absent and (b) the information is not written in the ontological metalanguage. However, the elicited information does include many aspects of a human mental model that would not be found in published sources, such as hypotheses about stage-by-stage disease progression despite the likely absence of actual attested property values for all stages. For this reason, the results of OntoElicit lie somewhere between a formal computational model and what we can expect to find in published sources.

Model components acquirable by agent reading. Published reports in the field of medicine typically contain only what is attested, making them insufficient as the sole source of knowledge for a comprehensive computational model. We might think of a complete computational model as a picture covered by a clear stencil whose holes represent model components that can be learned from the literature. As described in Section 5, the automatic learning of model components can be used either to update existing models or as the building blocks for more comprehensive, manually acquired models.

4 Modeling in OntoAgent

In the OntoAgent knowledge environment, disease models are recorded as complex events in the ontology. The ontology is a formal model of the world that is organized as a multiple-inheritance hierarchical collection of frames headed by concepts (OBJECTS and EVENTS) that are named using language-independent labels [7] [15] [21]. It currently contains approximately 9,000 concepts. The OBJECTS and EVENTS are described using PROPERTIES, both ATTRIBUTES and RELATIONS. The PROPERTIES themselves are primitives, i.e., their meaning is understood to be grounded in the real world without the need for further ontological decomposition. A short excerpt from the frame for the ontological concept SURGERY (which actually contains over a dozen more properties) is shown in Listing 1.

One of the properties not shown in this excerpt is the one that is key to modeling complex events: HAS-EVENT-AS-PART. The filler of this slot is an event script of the type introduced by Schank and Abelson [28]. Scripts represent typical sequences of events and their causal and temporal relationships. In other words, they encode how individual events hold well-defined places in routine, typical sequences of events that happen in the world, with a well-specified set of objects filling different roles throughout that sequence. Scripts require expressive means not provided in the simple slot-facet-filler formalism shown in Listing 1, and are recorded in a sister knowledge base. Scripts both drive agent simulation and support agent reasoning. For example, the script that describes a disease (its causes, variable paths of

Listing 1 Excerpt from the concept SURGERY in the OntoAgent ontology.

```
SURGERY
    IS-A (value MEDICAL-PROCEDURE)
    AGENT (default SURGEON) (sem PHYSICIAN) (relaxable-to HUMAN)
    THEME (default MEDICAL-PATIENT) (sem ANIMAL)
    LOCATION (default OPERATING-ROOM) (sem MEDICAL-BUILDING)
             (relaxable-to PLACE)
    INSTRUMENT (default SURGICAL-INSTRUMENT)
    DURATION (sem .5 - 8 (MEASURED-IN HOUR))
```

progression across patients, potential responses to interventions, etc.) permits (a) simulation of the disease in virtual patients, (b) reasoning about disease processes by the virtual medical tutor and (c) natural language dialog about the disease, since semantically-oriented natural language processing requires real-world knowledge support [21]. In short, a theoretically and practically motivated aspect of knowledge acquisition in OntoAgent is that knowledge, once recorded, should enable the maximum number of functionalities in the maximum number of agents [15].

For reasons of space, this discussion will focus primarily on the modeling of disease processes themselves, without as much detail about the modeling of interventions, clinical decision-making, agent decision-making, simulated agentive action, or any of the other necessary functionalities of agents, which are all handled in a corresponding way, as reported in the references cited earlier. It is important to understand the nature of the disease models in order to appreciate why they serve as a useful knowledge substrate for automatic knowledge acquisition from text. For this reason, we present select excerpts from our model for gastroesophageal reflux disease (GERD) by way of illustration.

4.1 An Excerpt from the Model for GERD

Gastroesophageal reflux disease, or GERD, can be defined as any symptomatic clinical condition that results from the reflux of stomach or duodenal contents into the esophagus. In laymen's terms, acidic stomach contents backwash from the stomach into the esophagus because the sphincter between the two – called the lower esophageal sphincter (LES) – is not functioning properly. The two sphincter abnormalities that give rise to GERD are abnormally low basal pressure of the LES (< 10 mmHg), or an abnormally large number or duration of so-called transient relaxations of the LES. Both of these lead to an increase in acid exposure to the lining of the esophagus. Clinically speaking, it does not matter which LES abnormality gives rise to excessive acid exposure, what matters is the amount of time per day this occurs. We record this feature as the variable "total time in acid reflux", or TTAR.

Although TTAR earns its place in the model as the variable that holds the results of the test called pH monitoring, it does not conveniently capture – for physicians or knowledge engineers – relative GERD severity. For that we introduced the abstract variable GERD-LEVEL. The values for GERD-LEVEL conveniently correlate with LES pressure as follows. If GERD is caused by a hypotensive LES, then GERD-LEVEL equals LES pressure. If GERD is caused by excessive transient relaxations, then the GERD-LEVEL reflects the same amount of acid exposure as would have been caused by the given LES pressure. So a GERD-LEVEL of 5 can indicate an LES pressure of 5 mmHg or a number/duration of transient relaxations per day that would expose the esophagus to that same amount of acid. Key aspects of the model then orient around GERD-LEVEL (rather than LES pressure, transient relaxations, or TTAR):

Table 1 Sample GERD levels and their associated total time in acid reflux (TTAR) per day. It also shows the baseline duration of each conceptual stage of the disease due to that TTAR, with more acid exposure leading to faster disease progression.

GERD level	ttar in hrs. per day	Stage duration in days
10	*less than 1.2*	*a non-disease state*
8	1.92	160
5	3.12	110
3	4.08	60

e.g., GERD-LEVEL is used to determine the pace of disease progression, with lower numbers reflecting more acid exposure and faster disease progression.

The stages of GERD are listed below. Each stage can be the end stage for some patients; that is, some lucky patients, even if left untreated, will never experience more than an inflamed esophagus, whereas others will end up with esophageal cancer. There is a bifurcation in disease path for patients experiencing late-stage disease, for reasons that are unknown.

- Preclinical: non-symptomatic inflammation of the esophagus.
- Inflammation: more severe inflammation of the esophagus, the beginning of symptoms.
- Erosion: one or more erosions occur in the esophageal lining.
- Ulcer: one or more erosions have progressed to the depth of an ulcer.
- *Post-ulcer path 1.* Barrett's metaplasia: a premalignant condition; progresses to cancer (an additional stage) in some patients.
- *Post-ulcer path 2.* Peptic stricture: an abnormal narrowing of the esophagus due to changes in tissue caused by chronic overexposure to gastric acid; does not lead to cancer.

The ontological scripts that support each stage of simulation include the basic physiological property changes, responses to interventions (if administered), and the effects of lifestyle choices. Sparing the reader the LISP code in which scripts are written, here is an example, in plain English, of how GERD progresses in an untreated patient who is predisposed to having erosion as the end stage of disease. During PRECLINICAL-GERD, the value of the property PRECLINICAL-IRRITATION-PERCENTAGE (an abstract property whose domain is MUCOSA-OF-ESOPHAGUS) increases from 0 to 100. When the value of PRECLINICAL-IRRITATION-PERCENTAGE reaches 100, the script for the PRECLINICAL-GERD is is unasserted, with the simultaneous assertion of the INFLAMMATION-STAGE script. During the INFLAMMATION-STAGE, the mucosal layer of the esophageal lining (recorded as the property MUCOSAL-DEPTH applied to the object ESOPHAGEAL-MUCOSA) 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 the EROSION-STAGE. At the start of the EROSION-STAGE, between 1 and 3 EROSION objects are 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 EROSION-STAGE script remains in a holding pattern since the patient we are describing does not have a predisposition to ulcer.

Over the course of each stage, property values are interpolated using a linear function, though other functions could be used if they were found to produce more lifelike simulations. So, halfway through PRECLINICAL-GERD, the patient's PRECLINICAL-IRRITATION-PERCENTAGE will be 50, and three quarters 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 1): e.g., a patient with a GERD-LEVEL of 8 will have a total time in acid reflux of 1.92 hours a day and each stage will last 160 days.

Some lifestyle habits, such as consuming caffeine, mints and fatty foods, increase GERD-LEVEL manifestation in some patients. In the model, if a patient is susceptible to GERD-influencing lifestyle habits and is engaging in those habits in simulation, then the effective GERD-LEVEL reduces by one. This results in an increase in acid exposure and a speeding up of each stage of the disease. If the patient is not actively engaging in the habit – e.g., after following the advice of a doctor to stop drinking caffeine – the GERD-LEVEL returns to its basic level. This is just one example of the utility of introducing the abstract property GERD-LEVEL into the model.

Let us now turn to two aspects of patient differentiation that highlight some more complex aspects of modeling: modeling why patients have different end stages of the disease, and modeling partial responses to medications. It is worth mentioning that we did not undertake either of these aspects of modeling in our initial model of GERD (published in [9]). The fact that we could seamlessly incorporate these enhancements, without perturbation to the base model, is evidence of the inherent extensibility of the models developed using this modeling strategy.

Modeling different end stages of disease across patients. It is unknown why patients have different end stages of GERD if the disease is left untreated. However, physicians can and do hypothesize about the reasons for cross-patient differentiation, which could include genetic, environmental, physiological and even emotional factors.[4] To capture some practically and pedagogically useful hypotheses, we introduced three abstract parameters into the model:

- MUCOSAL-RESISTANCE reflects the hypothesis that patients differ with respect to the degree to which the mucosal lining of the esophagus protects the esophageal tissue from acid exposure and fosters the healing of damaged tissue. A higher value on the abstract (0-1) scale of MUCOSAL-RESISTANCE is better for the patient.
- MODIFIED-TTAR combines MUCOSAL-RESISTANCE with the baseline TTAR to capture the hypothesis that a strong mucosal lining can functionally decrease the effect of acid exposure. For example, patients with an average MUCOSAL-RESISTANCE will have the stage durations shown in Table 1 above. Patients with an above-average MUCOSAL-RESISTANCE will have a lower MODIFIED-TTAR: e.g., if a patient's TTAR is 3.12 hours, but the patient has a mucosal resistance of 1.2, we model that as an MODIFIED-TTAR of 2.5 hours (3.12 multiplied by .8), and the disease progresses correspondingly slower. By contrast, if the patient's TTAR is 3.12 hours but it has a MUCOSAL-RESISTANCE of .8, then the MODIFIED-TTAR is 3.75 hours (3.12 multiplied by 1.2), and disease progression is correspondingly faster.
- DISEASE-ADVANCING-MODIFIED-TTAR is the total time in acid reflux required for the disease to manifest at the given stage. This variable permits us to indicate the end stage of a patient's disease in a more explanatory way that by simply asserting it. That is, for each patient, we assert how much acid exposure is necessary to make the disease progress into each stage, as shown in Table 2. If the acid exposure is not sufficient to support disease progression into a given stage (as shown by the italicized cells), the patient's

[4] For a medical description of the emotional effects on GERD, see [20]. For our incorporation of these factors into the clinical model, see [17].

Table 2 The first column indicates the patient's actual total time in acid reflux per day. The cells in the remaining columns indicate the total time in acid reflux needed for GERD to advance in that stage. Cells in italics show that the disease will not advance to this stage unless the patient's MODIFIED-TTAR changes – which could occur, e.g., if the patient took certain types of medications, changed its lifestyle habits or had certain kinds of surgery.

Patient	Modified-ttar	Preclin.	Inflamm.	Erosion	Ulcer	Pep.Strict.
John	1.92	1.92	1.92	*2.3*	*2.5*	*3.12*
Fred	2.8	1.92	1.92	2	2.7	*3.12*
Harry	4.08	1.92	1.92	3	3.5	4.0

Table 3 Effects of medications on MODIFIED-TTAR. The resulting MODIFIED-TTAR is written in brackets.

Patient	Modified-ttar	H2 blocker reduction	PPI once daily	PPI twice daily
John	1.92	.5 [1.42]	1.25 [.67]	1.5 [.42]
Fred	2.8	.3 [2.5]	1[1.8]	2.25 [.55]
Harry	4.08	.1 [3.98]	.8 [3.28]	2.2 [1.88]

disease will hit its end stage. For example, John is a patient whose disease will not progress past the Inflammation stage, even if left untreated, because his MODIFIED-TTAR is not high enough to support the erosion stage of GERD. Fred's disease will advance into the ulcer stage, and Harry's disease will advance to peptic stricture.

Modeling Complete and Partial Responses to Medication. In order to capture complete and partial responses to medications, medication effects are modeled as decreases in MODIFIED-TTAR, as shown in Table 3.

The table indicates the decrease in acid exposure caused by each medication for each patient, along with the resulting MODIFIED-TTAR. So, for each day that **John** takes an H2 blocker, his MODIFIED-TTAR will be 1.42, which is not a disease state. If he already has the disease, healing will occur. The other, stronger, medication regimens will also be effective for him. For **Fred**, the H2 blocker is not sufficient to promote complete healing (it brings the MODIFIED-TTAR down to 2.5), but it would be sufficient to not permit his disease to progress to the ulcer stage; or, if Fred were already in the ulcer stage, the ulcers would heal to the more benign level of erosions. If Fred took a PPI once or twice daily, his MODIFIED-TTAR would be < 1.92, meaning that his esophagus would heal completely. For **Harry**, the H2 blocker would not help at all – he would still progress right through the stricture stage. Taking a PPI once a day would heal ulcers and block late stages of disease. Taking a PPI twice a day would heal the disease completely, unless Harry had already experienced a stricture: there is no non-operative cure for a peptic stricture, a detail we will not pursue at length here but that is covered in the model (the STRICTURE object generated by the simulation remains a part of the patient's anatomy).

In sum, the physiologically-grounded parameter MUCOSAL-RESISTANCE permits each patient's end stage of disease progression to be calculated rather than asserted; and the parameters MODIFIED-TTAR and DISEASE-ADVANCING-MODIFIED-TTAR permit us to model full and partial efficacy of medications. As additional objective evidence becomes available through experimentation, the actual numerical values of these features can be modified accordingly.

Given models like this, the system need not exhaustively list all permutations of paths a

trainee could take when diagnosing and treating a virtual patient, or all responses of the virtual patient to interventions. Instead, the system relies on these ontologically-grounded descriptions of basic physiology, disease processes, and effects of treatments and their interactions, so that the state of an MVP at any given time is dynamically computed by the system's reasoning module. Similarly, any of the tests available in the system can be run at any time, as they measure physiological properties of the patient as it lives its simulated life.

Let us conclude this section by returning to the question of how closely simulation-supporting computational models like these align with what is available in the published literature. The most striking difference is that much of our computational model is neither directly attested nor attestable, there being no widescale monitoring of people's physiology on a daily basis over the course of years. So, even those properties that are in principle measurable (such as TTAR and SUBMUCOSAL-DEPTH) are only a starting point for a picture that must be largely filled in by educated guesses. This is in addition to properties that are not currently measurable (such as PRECLINICAL-IRRITATION-PERCENTAGE) and properties that are introduced in order to capture specialists' generalizations about phenomena (e.g., GERD-LEVEL). The fact that clinicians' mental models are largely comprised of evidence-supported educated guesses does not impede effective clinical practice, but it does represent a divergence from the small subset of actually attested information in the literature. So, the question becomes, to what extent can we learn aspects of such models from texts?

5 Learning Model Components from Texts

The answer is that we can learn from texts *model components*, defined as ontologically-grounded property-value pairs that direcly contribute to full computational models. Learnable features have the following properties:

- They are straightforward and concrete, such as LES-PRESSURE (measurable by a test) or SENSITIVITY-TO-CAFFEINE (knowable based on patient reports); they are not abstract modeling properties (MODIFIED-TTAR, MUCOSAL-RESISTANCE), which will have no precise equivalents in published texts.
- They are known to be changeable over time, based on our ontological knowledge of the domain. For example, since we know that new medications and tests are constantly being invented, we know that the properties TREATED-BY-MEDICATION and ESTABLISHED-BY-TEST must have an open-ended inventory of values. By contrast, we do not expect the need to change the fact that heartburn can be a symptom of GERD, or that HEARTBURN-SEVERITY is modeled as having values on the abstract scale (0-1).
- (For knowledge involving causal chains only) If a sequence of events is modeled temporally rather than causally (using what we call "clinical knowledge bridges"), these can be automatically replaced by attested causal chains. However, if the model already records casual chains, their modification is likely to be too complex to be learned automatically without inadvertently perturbing the model.

Table 4 shows some examples of properties (associated with their respective concepts) whose values we believe can be learned from the literature.

The fillers for each property are formal, ontologically-grounded knowledge structures, which are produced during the automatic analysis of text by the OntoSem language processor. For example, all of the following text strings, and many more, will result in text meaning representations that permit the system to insert PROTON-PUMP-INHIBITOR as the value for the property HAS-TREATMENT of the concept GASTROESOPHAGEAL-REFLUX-DISEASE:

Table 4 Examples of properties, associated with their respective concepts, whose values can be learned from the literature.

Concept	Properties
DISEASE	HAS-EVENT-AS-PART, AFFECTS-BODY-PART, CAUSED-BY, HAS-SYMPTOMS, HAS-DIAGNOSTIC-TEST, HAS-TREATMENT
DIAGNOSTIC-TEST	MEASURES-PROPERTY, NORMAL-RESULT, ABNORMAL-RESULT, SIDE-EFFECTS, PAIN-INDUCED
MEDICAL-TREATMENT	HAS-EVENT-AS-PART, EFFICACY, HAS-RISKS, PAIN-INDUCED

- a proton pump inhibitor treats <can treat, can be used to treat, can be prescribed to treat, is often prescribed to treat> GERD
- GERD is <can be> treated by <cured by> (taking) a proton pump inhibitor
- doctors <your doctor may> recommend <prescribe> (taking) a proton pump inhibitor
- patients may <can, may be advised to> take a proton pump inhibitor

Establishing the functional equivalence of these strings is not done by listing; instead, it is done by combining our general approach to natural language understanding with algorithms for paraphrase detection ([11, 12]) and ontologically-grounded reasoning.

Let us consider just three examples of how natural language analysis supports the knowledge extraction process we are describing. Assume we are seeking to automatically learn or verify the veracity of the previously discussed fact "GASTROESOPHAGEAL-REFLUX-DISEASE (HAS-TREATMENT PROTON-PUMP-INHIBITOR)". As we said, all of the inputs above provide this information, albeit some more directly than others. The input *GERD is treated by a proton pump inhibitor* perfectly matches the lexical sense for the verb *treat* that is defined by the structure "DISEASE is treated by MEDICATION", and the analyzer generates exactly the text meaning representation we are seeking: GASTROESOPHAGEAL-REFLUX-DISEASE (HAS-TREATMENT PROTON-PUMP-INHIBITOR). In other cases, the basic text meaning representation includes additional "benign" information, which does not affect the truth value of the main proposition: e.g., the potential modality scoping over the proposition *GERD can be treated by a proton pump inhibitor* does not affect the truth value of the main proposition, which is the same as before and matches the expectation we seek to fill. In still other cases, the meaning we are looking for must be inferred from what is actually written. For example, the input *Your doctor may recommend a proton pump inhibitor* does not explicitly say that a proton pump inhibitor treats GERD, but it implies this based on the general ontological knowledge that a PRECONDITION for a physician advising a patient to take a medication is (DISEASE (HAS-TREATMENT MEDICATION)). Because the system has access to this ontological knowledge, it can make the needed inference and fill in our slot as before. It should be noted that these types of reasoning rules are not spontaneously generated – they must be recorded, like any other knowledge. However, once recorded, they can be used for any applicable reasoning need of the agent.

When investigating what information could be extracted from medical texts, we focused on two genres that offer different opportunities for knowledge extraction: case studies and disease overviews. Like narratives, both of these have largely predictable content and structure, which should support the automatic identification of disease model component information.

Case studies do not present all disease mechanics. Instead, they typically begin with a broad overview of the disease to serve as a reminder to readers who are expected to be familiar with "the script". Then they focus on a single new or unexpected aspect of the disease as manifest in one or a small number of patients (cf. the story-worthy aspects of

Table 5 Application for updating clinicians from case studies.

Case study: "Meditation as medication for GERD"
Author: Dr. J. Physician
Date: Jan. 11, 2018
Therapies for GERD
Mild: lifestyle modifications, H2 blocker, PPI QD, **MEDITATION-new**
Severe: PPI BID

narratives). For example, [3] is a case study that reports that a mother and daughter both suffer from the same rare disease, achalasia, and suggests that this case supports previous hypotheses of a genetic influence on disease occurrence. The new findings are typically repeated in the Abstract, Case Report, and Discussion sections, offering useful redundancy to improve system confidence.

The system can automatically compare the information in a case study with the ontologically grounded computational model as follows. First it can semantically analyze the case study, focusing on the TMR chunks representing the types of learnable property values listed above. (This focusing means that the system need not achieve a perfect analysis of every aspect of the text: it knows what it is looking for.) Then, it can compare the learned property values with the the values in the model. Continuing with our example of mother-daughter achalasia, our current model of achalasia has no filler for the value of CAUSED-BY since, when we developed the model, the cause was not definitively known (it still is not; the genetic influence remains to be validated). Automatically filling an empty slot with a new filler can be carried out directly, with no extensive reasoning necessary. However, the *nature* of that slot filler must be understood: it represents an instance, not a generic ontological fact. The system has two sources of evidence that this information is an instance: (1) the individuals spoken about are instances, so the features applied to them are also instances (compare this with assertions about about generic *people* or generic *you*); (2) the genre of case study sets up the expectation that reported information will be at the level of instance.

We believe it would be useful to configure an application that would alert clinicians to new findings in a "snaphot" formalism like that shown in Table 5. This presentation style encapsulates the expectations that: (a) clinicians know, without explanation, that one of the ontological properties of diseases is that they might have effective therapies; (b) when providing new information, it is useful to provide old information as the backdrop, with a clear indication of whether the new information adds to or overwrites the old information; (c) clinicians understand that information provided in case studies represents instances and not cross-the-boards generalizations; (d) modern-day users understand that entities can be clicked on for more information (e.g., which lifestyle modifications are being referred to), (e) terseness is appreciated by busy people operating within their realm of specialization.

Let us turn now to the other genre from which model information can be extracted: disease overviews. They typically present a stable inventory of properties of interest, often even introduced by subheadings, such as causes of the disease, risk factors, physiological manifestations, symptoms, applicable tests and procedures, and so on. Not surprisingly, these categories align well with the knowledge elements we seek to extract from texts, shown in Table 4. The natural language processing of disease overviews would proceed as described above. However, we envision applications for this processing to be somewhat different. For example, an application could respond to a clinician's request for a thumbnail sketch of a disease by reading overviews, populating the inventory of key property values, and presenting them in a semi-formal manner, such a list of concept-property-value triples.

6 Discussion

This paper has presented a combination of work completed and work in the planning stages. The knowledge substrate and language processing capabilities are quite advanced, whereas the approach to mining new information from text is algorithmic.[5]

We present this work now as a contribution to a discussion that is key to computational narrative and agent building overall: to what extent can agents *in principle* learn models from text? And, if not full models, what *can* they learn through lifelong learning by reading?

In this paper we have suggested that although full models cannot be learned (they are largely unattested and rely centrally on educated guessing) certain model components can be automatically learned even in the near term, using currently available language processing technologies and achievable types of machine reasoning. This is a revolutionary idea, considering that we are talking about learning *ontologically-grounded knowledge structures* rather than extracting uninterpreted natural language strings from text.

If, by contrast, we want intelligent agents to learn full models from texts, then domain experts will need to write down fully specified mental models – an interesting prospect, particularly as it requires experts to boldly hypothesize about the unknown in the same way as they did to engineer the disease models for MVP. In short, modeling – be it recorded using an ontological metalanguage or a natural language like English – involves theorizing in an uncertain data space, something that is done as a matter of course in daily clinical practice but is not typically converted into published form. However, the potential rewards of fully specified (albeit with an understood tolerance for imprecision) models are tantelizing. Consider just a short excerpt from a committee report that lays out desiderata for virtual patient systems:

> "The clinician interacts with models and abstractions of the patient that place the raw data in context... These virtual patient models are the computational counterparts of the clinician's conceptual model of a patient... [The data] depict and simulate a theory about interactions going on in the patient and enable patient-specific parameterization... They build on submodels of biological and physiological systems..." [30].

Capabilities such as these directly motivate the need for inspectable, model-based artificial intelligence, not only in virtual patient applications but far beyond. It is our hope that the research reported here contributes to this vision, offering evidence of how component problems can be solved over time if we soberly analyze the necessary collaboration between human knowledge engineering and the potential for automatic agent learning.

References

1 Daniela Bailer-Jones. *Scientific Models in Philosophy of Science*. Pittsburgh, Pa: University of Pittsburgh Press, 2009.

2 Martha Evens and Joel Michael. *One-on-One Tutoring by Humans and Computers*. New Jersey and London: Lawrence Erlbaum and Associates, Publishers, 2006.

[5] We hope to further develop and implement the algorithms as a collaboration with Mark Finlayson, bringing to bear his Story Merging Algorithm [4], which will assist in comparing candidate model enhancements with our base models.

3 Yulia Viktorovna Evsyutina, Alexander Sergeevich Trukhmanov, and Vladimir Trophimovich Ivashkin. *Family case of achalasia cardia: Case report and review of literature*. World J Gastroenterol. Jan 28, 2014; 20(4): 1114–1118. Published online Jan 28, 2014.

4 Mark Finlayson. *Learning Narrative Structure from Annotated Folktales.* Doctoral Dissertation. Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, 2012.

5 Antonio Lieto. *Manually vs Semiautomatic Domain Specific Ontology Building.* Thesis in Information and Commercial Electronics, University of Salerno, 2008.

6 Antonio Lieto and Rossana Damiano. *A hybrid representational proposal for narrative concepts: A case study on characters' roles*. In: Mark A. Finlayson, Jan Christoph Meister, and Emile Bruneau (Eds.), Proceedings of the 5th International Workshop on Computational Models of Narrative, 2014.

7 Marjorie McShane, Sergei Nirenburg and Stephen Beale. *An NLP lexicon as a largely language independent resource*. Machine Translation 19(2): 139–173, 2005.

8 Marjorie McShane, George Fantry, Stephen Beale, Sergei Nirenburg, and Bruce Jarrell. *Disease interaction in cognitive simulations for medical training*. In: Proceedings of MODSIM World Conference and Expo, Virginia Beach, Sept. 11–13, 2007.

9 Marjorie McShane, Sergei Nirenburg, Stephen Beale, Bruce Jarrell, and George Fantry. *Knowledge-based modeling and simulation of diseases with highly differentiated clinical manifestations*. In: R. Bellazzi, A. Abu-Hanna, J. Hunter (Eds.), Proceedings of the 11th Conference on Artificial Intelligence in Medicine (AIME 07), Amsterdam, The Netherlands, 2007. Berlin, Heidelberg: Springer-Verlag. pp. 34–43. Amsterdam, The Netherlands, July 7–11, 2007.

10 Marjorie McShane, Bruce Jarrell, George Fantry, Sergei Nirenburg, Stephen Beale and Ben Johnson. *Revealing the conceptual substrate of biomedical cognitive models to the wider community*. In: J.D. Westwood, R.S. Haluck, H.M. Hoffman, G.T. Mogel, R. Phillips, et al. (Eds.), Medicine Meets Virtual Reality 16. Amsterdam, Netherlands: IOS Press, pp. 281–286, 2008.

11 Marjorie McShane, Sergei Nirenburg and Stephen Beale. *Resolving paraphrases to support modeling language perception in an intelligent agent*. In: J. Bos and R. Delmonte (Eds.), Semantics in Text Processing: STEP 2008 Conference Proceedings, Venice, Italy. London: College Publications, 2008.

12 Marjorie McShane, Sergei Nirenburg and Stephen Beale. *Two kinds of paraphrase in modeling embodied cognitive agents.* In: A.V. Samsonovich (Ed.), Biologically Inspired Cognitive Architectures. Papers from the AAAI Fall Symposium, Washington, DC, Nov. 7–9, 2008. AAAI Technical Report FS-08-04. Menlo Park, CA: AAAI Press, pp. 62–167, 2008.

13 Marjorie McShane and Sergei Nirenburg. *Dialog modeling within intelligent agent modeling*. Proceedings of the IJCAI-09 Workshop on Knowledge and Reasoning in Practical Dialog Systems, pp. 52–59. Pasadena, CA, USA, 2009.

14 Marjorie McShane, Stephen Beale, Sergei Nirenburg, Bruce Jarrell and George Fantry. *Inconsistency as diagnostic tool in a society of intelligent agents*. Artificial Intelligence in Medicine (AIIM), 55(3):137–48, 2012.

15 Marjorie McShane and Sergei Nirenburg. *A knowledge representation language for natural language processing, simulation and reasoning*. International Journal of Semantic Computing 6(1), 2012.

16 Marjorie McShane and Sergei Nirenburg. *Use of ontology, lexicon and fact repository for reference resolution in Ontological Semantics*. In: A. Oltramari, P. Vossen, L. Qin, and E. Hovy (Eds.), New Trends of Research in Ontologies and Lexical Resources, pp. 157–185, Springer, 2013.

17 Marjorie McShane, Sergei Nirenburg, Stephen Beale, Bruce Jarrell, George Fantry, and David Mallott. *Mind-, body- and emotion-reading*. Proceedings of IACAP 2013 (International Association for Computing and Philosophy), University of Maryland College Park, July 15–17, 2013.

18 Marjorie McShane, Sergei Nirenburg and Bruce Jarrell. *Modeling decision-making biases*. Biologically-Inspired Cognitive Architectures (BICA) Journal. Volume 3:39–50, 2013.

19 Nighat F. Mehdi, Miles M. Weinberger and Mutasim N. Abu-Hasan. *Achalasia: unusual cause of chronic cough in children*. Cough 2008, 4:6.

20 I. Mizyed, S.S. Fass and R. Fass. *Review article: gastro-oesophageal reflux disease and psychological comorbidity*. Alimentary Pharmacology and Therapeutics, 29: 351–358. Blackwell Publishing, Ltd., 2009.

21 Sergei Nirenburg and Victor Raskin. *Ontological Semantics*. The MIT Press, 2004.

22 Sergei Nirenburg, Marjorie McShane and Stephen Beale. *A simulated physiological/cognitive "double agent"*. Proceedings of the Workshop on Naturally Inspired Cognitive Architectures, AAAI 2008 Fall Symposium, Washington, D.C., Nov. 7–9, 2008.

23 Sergei Nirenburg, Marjorie McShane and Stephen Beale. *Aspects of metacognitive self-awareness in Maryland Virtual Patient*. Proceedings of the AAAI Fall Symposium on Cognitive and Metacognitive Educational Systems, Nov. 11–13, Arlington, VA, 2010.

24 Sergei Nirenburg, Marjorie McShane and Stephen Beale. *Hybrid methods of knowledge elicitation within a unified representational knowledge scheme*. In: J. Filipe and J.L.G. Dietz (Eds.), KEOD 2010: Proceedings of the International Conference on Knowledge Engineering and Ontology Development, October 25-28,Valencia, Spain, 2010. SciTePress. pp. 177–182, 2010.

25 Sergei Nirenburg, Marjorie McShane, Stephen Beale, Jesse English and Roberta Catizone. *Four kinds of learning in one agent-oriented environment*. In: A.V. Samsonovich et al. (Eds.), Proceedings of the First International Conference on Biologically Inspired Cognitive Architectures, Arlington, VA, Nov. 13-14, 2010. Amsterdam, Netherlands: IOS Press, pp. 92–97, 2010.

26 Sergei Nirenburg, Marjorie McShane, Stephen Beale, Bruce Jarrell and George Fantry. *Intelligent agents in support of clinical medicine*. Proceedings of MMVR18, February 9–12, 2011, Newport Beach, California, 2011.

27 Brian O'Neill and Mark Riedl. *Applying qualitative research methods to narrative knowledge engineering*. In: Mark A. Finlayson, Jan Christoph Meister, and Emile G. Bruneau (Eds.), Proceedings of the 5th Workshop on Computational Models of Narrative, pp. 139–153, 2014.

28 Roger Schank and Robert Abelson. 1977. *Scripts, Plans, Goals, and Understanding: An Inquiry into Human Knowledge Structures*. New York: L. Erlbaum Associates.

29 Giovanni Sileno, Alexander Boer, and Tom van Engers. *Legal Knowledge Conveyed by Narratives: Towards a Representational Model*. In: Mark A. Finlayson, Jan Christoph Meister, and Emile G. Bruneau (Eds.), Proceedings of the 5th Workshop on Computational Models of Narrative, pp. 182–191, 2014.

30 W. W. Stead and H. S. Lin (Eds). *Computational Technology for Effective Health Care: Immediate steps and strategic directions*. National Research Council. Washington, DC: The National Academies Press, 2009.

31 C. B. Thielst. *The new frontier of electronic, personal and virtual health records*. Journal of Healthcare Management, vol. 52, pp. 75–78, 2007.