Artificial Intelligence in Medicine 55 (2012) 137–148

Contents lists available at SciVerse ScienceDirect

Artificial Intelligence in Medicine

journal homepage: www.elsevier.com/locate/aiim

Inconsistency as a diagnostic tool in a society of intelligent agents

Marjorie McShane[a,*], Stephen Beale[a], Sergei Nirenburg[a], Bruce Jarrell[b], George Fantry[b]

[a] Department of Computer Science and Electrical Engineering, University of Maryland Baltimore County, Baltimore, MD 21250, USA
[b] University of Maryland School of Medicine, Baltimore, MD 21201, USA

ARTICLE INFO

Article history:
Received 27 October 2011
Received in revised form 26 March 2012
Accepted 25 April 2012

Keywords:
Cognitive modeling
Intelligent agents
Virtual tutor
Clinical decision making support

ABSTRACT

Objective: To use the detection of clinically relevant inconsistencies to support the reasoning capabilities of intelligent agents acting as physicians and tutors in the realm of clinical medicine.
Methods: We are developing a cognitive architecture, OntoAgent, that supports the creation and deployment of intelligent agents capable of simulating human-like abilities. The agents, which have a simulated mind and, if applicable, a simulated body, are intended to operate as members of multi-agent teams featuring both artificial and human agents. The agent architecture and its underlying knowledge resources and processors are being developed in a sufficiently generic way to support a variety of applications.
Results: We show how several types of inconsistency can be detected and leveraged by intelligent agents in the setting of clinical medicine. The types of inconsistencies discussed include: test results not supporting the doctor's hypothesis; the results of a treatment trial not supporting a clinical diagnosis; and information reported by the patient not being consistent with observations. We show the opportunities afforded by detecting each inconsistency, such as rethinking a hypothesis, reevaluating evidence, and motivating or teaching a patient.
Conclusions: Inconsistency is not always the absence of the goal of consistency; rather, it can be a valuable trigger for further exploration in the realm of clinical medicine. The OntoAgent cognitive architecture, along with its extensive suite of knowledge resources an processors, is sufficient to support sophisticated agent functioning such as detecting clinically relevant inconsistencies and using them to benefit patient-centered medical training and practice.

1. Introduction

The term *inconsistency* tends to imply a negative evaluation of a state of affairs, the lack of a state of consistency that is the implied goal. However, in societies of people, and in societies of intelligent agents modeled as people, inconsistency is not always a detriment, it can actually serve as a diagnostic tool. For example, in the domain of clinical medicine, inconsistencies between test results and the doctor's hypothesis about what is wrong can suggest that the hypothesis was incorrect or that the test results were flawed, and inconsistencies between a doctor's observation and a patient's report can suggest an intentional or unintentional misrepresentation by the patient. Similarly, in the domain of *teaching* clinical medicine, inconsistencies between an expert's preferred approach to solving a problem and a novice's approach can suggest room for improvement for the novice. Accordingly, if intelligent agents are modeled to function as clinicians or as tutors for clinicians, they should be prepared to exploit *diagnostic inconsistencies* in the same way as people do. This paper describes our recent work in configuring intelligent agents who serve in both of these roles and count among their stockpile of cognitive capabilities the detection of diagnostic inconsistencies.

1.1. OntoAgent agents

The intelligent agent environment to be used for illustration is OntoAgent, which supports the modeling of human-like behavior in artificial intelligent agents that collaborate with people (http://www.trulysmartagents.org/index.php). OntoAgent is a multi-agent environment being developed to support a suite of applications, including training and advising in the domain of clinical medicine. The development approach, which seeks to model agents that function like people, necessitates that our program of R&D be comprehensive, actively covering areas including the following:

- Physiological simulation using hybrid domain knowledge: physiological pathways when they are known and relevant for the goals of the simulation, and clinical "bridges" when pathways are

* Corresponding author at: CSEE Dept., ITE 325, University of Maryland Baltimore County, Baltimore, MD 21250, USA.
E-mail address: marge@umbc.edu (M. McShane).

0933-3657/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.artmed.2012.04.005

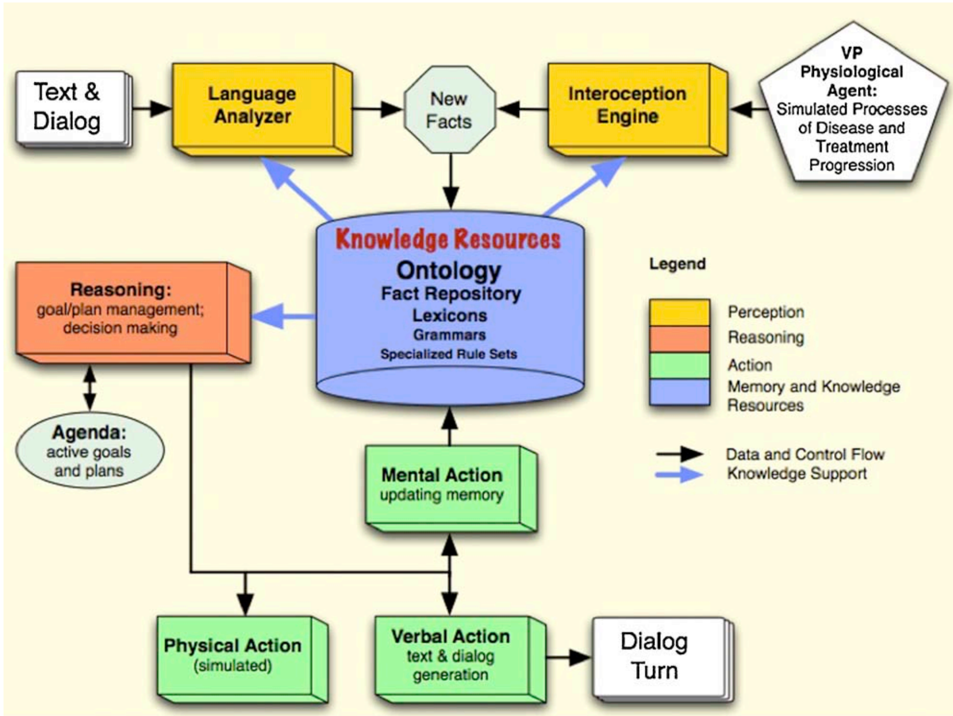

Fig. 1. The architecture of "double agents" in OntoAgent. The physiological agent is represented by the upper right pentagon. The rest of the figure – including the interoception engine, which perceives physiological stimuli generated by the physiological agent – is the realm of the cognitive agent.

either unknown or unnecessary for the needs of realistic simulation [1,2].
- Modeling interoception, which is the cognitive perception of physical signs and symptoms by intelligent agents [3].
- Decision theory, including goal and plan management and hybrid (rule-based, analogy-based and probabilistic) reasoning [3].
- Deep language processing that includes semantic and pragmatic analysis and language generation [4,5].
- Dialog modeling [6].
- Memory modeling and management [7–9].
- Agent individualization according to character traits (courage, suggestibility, boredom threshold etc.), personal preferences (likes coffee, is afraid of surgery, etc.), differing knowledge of the world, etc. [10].
- Agent learning of facts, concepts and language by being told, by reading and by experience [11].
- Use of a shared metalanguage of description and knowledge bases for agent reasoning, learning and language processing [12,13].
- Semi-automatic knowledge elicitation for knowledge-based systems [14].
- Integrating complex multipurposes knowledge systems.
- Ergonomic issues for developers and subject matter experts, such as the development and maintenance of a uniform knowledge representation substrate for core knowledge bases and processors and the development of a variety of efficiency-enhancing toolkits ([4] and http://www.trulysmartagents.org/dekade.php).
- Validation of our work by demonstrating feasible, proof-of-concept applications [1,2,10,11,15].

Here, we provide brief background about select aspects of OntoAgent that should suffice to support understanding of the new work being reported.

Intelligent agents in OntoAgent are "double agents", in that they have a cognitive side and, optionally, a physiological side. The architecture of agents is shown in Fig. 1.

To model the physiological agent – which "lives" a simulated life over time in applications – we encode knowledge about biophysical functions that have clinical relevance in the maintenance of health, the production of disease, and the bidirectional transitions between these two states. When biomechanisms are known and are deemed clinically important, they are modeled using causal chains. Gaps in our knowledge of biomechanisms are bridged with non-biomechanistic knowledge from the literature, practical clinical knowledge and, in some cases, probabilistic methods. The depth and granularity of the models are determined by the demands of automatic function and realism in our current and anticipated applications. As a rule of thumb, a feature value or process is included in the model if it can either be measured through tests, be affected by medication/interventions, or cause a change in some other clinically relevant feature value or process. Of central importance is the fact that our models can be easily modified, reflecting new medical findings or the assessment that a clinical "bridge" (i.e., a non-mechanistic observation of how a disease process manifests based on population-level clinical observation) needs to be replaced by a biomechanism in order to support the functionality required by an application.

The cognitive agent engages in the well-known triad of functionalities: perception, reasoning and action. It is ideologically close to, though methodologically not identical with, the belief-desire-intention (BDI) model of agency [16]. Unlike the classical BDI implementations [17], our approach centrally involves language comprehension and production as well as physiological simulation.

As shown in Fig. 1, agents undergo two types of *perception*: interoception and language understanding. The results of both modes of perception are interpreted and stored in the agent's knowledge resources: ontology, fact repository (memory of assertions), and language resources, such as lexicon and grammar (the latter can be learned, e.g., when the agent encounters new words through interaction with another agent). Each agent has its own knowledge resources, which can differ from those of other agents to represent different levels of education, different personal experience,

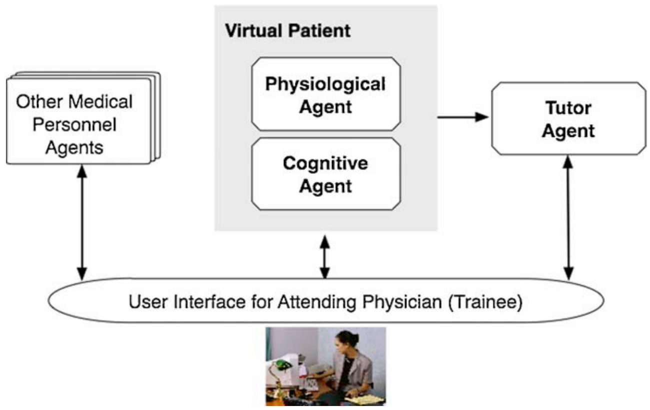

Fig. 2. The network of agents in the Maryland Virtual Patient (MVP) application.

etc. Agents also learn about aspects of other agents' knowledge and store these in their fact repository "profiles" of those agents.

Interoception is the agent's perception of its bodily signals, which are generated by the physiological agent. The interoception submodule operates a set of demons that are programmed (a) to notice the changes in values of specific physiological parameters and (b) if these values move outside a certain range, to instantiate corresponding symptoms in the VP's memory. Symptoms are represented as values of properties in the VP's profile of self, which is an instance of the ontological concept HUMAN stored in its fact repository.

The second type of agent perception is *language understanding*, which will be described in some detail in Section 2.

Agent *reasoning* is carried out at dozens of levels, from the many processes involved in deep natural language understanding to the processes involved in memory management to the manipulation of plans and goals. Fig. 1 highlights the goal- and plan-oriented aspects of agent reasoning, which is operationalized using knowledge structures like the ones to be described in Section 4.

Once we have a parameterizable double agent, what can we do with it? There are many options. We can take one instance of a body combined with one instance of a mind and create a virtual patient that can serve as a training case for physicians; and we can repeat that process thousands of times to create differentiated cases. Or we can take a mind (no body needed) with sophisticated medical knowledge and create a clinician's advisor. Or we can take a mind similar to the latter, supplement it with pedagogical knowledge, and create an interactive tutor. In short, not only does this "theme and variations" style of modeling permit wide variety among agents, it maximally reuses knowledge structures and modeling strategies across agents.

1.2. Societies of agents

Intelligent agents in OntoAgent function in societies of humans and other intelligent agents. A sample configuration is shown in Fig. 2, which is a high-level view of the prototype Maryland Virtual Patient (MVP) application.

MVP is a cognitive simulation and training system whose goal is to provide medical practitioners with the opportunity to develop clinical decision-making skills by managing many highly differentiated virtual patients (VPs) [1,2,10]. The human agent, who is typically a physician-in-training seeking to improve his cognitive decision making skills, plays the role of the attending physician. The two core artificial agents are the VP and the tutor, which share an inventory of cognitive skills including, non-exhaustively, the ability to reason in a context-sensitive way, to communicate in natural language, to learn, to manage memory, and to maintain a dynamic model of other agents. The VP and the tutor also have certain specialized capabilities. The tutor has specialized ontological knowledge and decision functions devoted to tutoring but has no need for a simulated body or personality. The VP, for its part, does have a simulated body and a personality. Physiologically, it undergoes both normal and pathological processes and responds realistically both to expected and to unexpected (e.g., caused by errors in a user's medical logic) internal and external stimuli. Its personality affects its decision making in the realms of lifestyle and health care.

Users of MVP can interview a VP using natural language; 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 such as medication and surgery; follow-up after those treatments to judge their efficacy; follow a patient's condition over an extended period of time; receive mentoring from the automatic tutor, if desired; and repeat the management of a given VP using different management strategies to compare their outcomes. The user can launch any intervention available in the system at any time during the simulation, be it clinically justified or not. In the latter case, if the user inadvertently worsens the VP's condition or initiates a new disease process, he must recover from the error in the continuing simulation by treating the new condition he has inadvertently caused. A prototype MVP system has been implemented, covering several diseases of the esophagus, and the system continues to be developed.

2. Technical background about the OntoAgent environment

This section provides background about the OntoAgent environment that will be useful for readers who want to fundamentally understand how the reported work is being implemented and why it is feasible. Readers preferring to focus more narrowly on the utility of diagnostic inconsistency may wish to skip this section and return later, if needed, should a deeper understanding of the illustrative knowledge structures be desired.

All OntoAgent applications use a primarily knowledge-based approach to agent modeling. All physiological, general cognitive and language processing capabilities of all intelligent agents rely on the same ontological substrate, the same organization of the fact repository (agent memory of assertions) and the same approach to knowledge representation. Since knowledge structures will be presented as an explanation of *how* we prepare our agents to function effectively in the face of inconsistency, a short overview of each knowledge base is warranted.

The OntoAgent ontology is a formal model of the world that provides a metalanguage for describing meaning derived from any source, be it language, intelligent agent perception, intelligent agent reasoning or simulation [4,12,13]. The metalanguage of description is unambiguous, permitting automatic reasoning about language and the world to be carried out without the interference of lexical and morphosyntactic ambiguities. The ontology is organized as a multiple-inheritance hierarchical collection of frames headed by concepts that are named using language-independent labels. It currently contains approximately 9000 concepts, most of which belong to the general domain. Concepts divide up into events, objects and properties (relations and attributes). The objects and events are described using the properties, while 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.

The expressive power of the ontology is enhanced by multivalued fillers for properties, implemented using facets. Facets permit the ontology to include information such as "the most typical

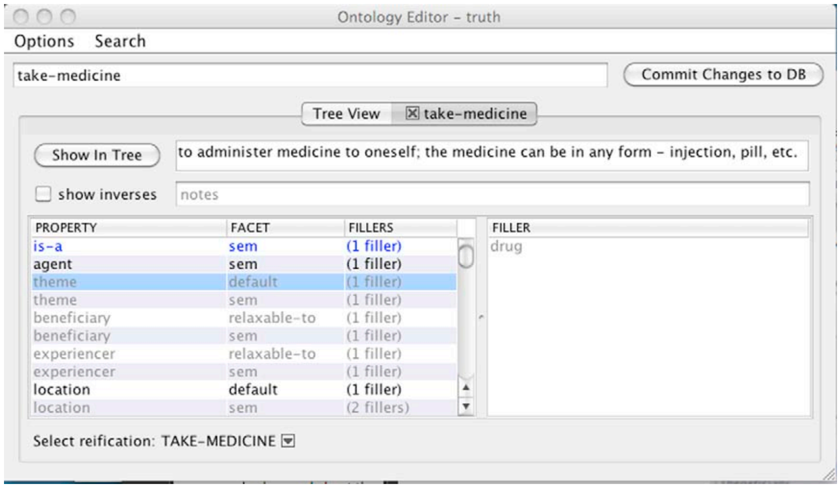

Fig. 3. Excerpt from the ontological frame for the concept TAKE-MEDICINE. The highlighted panel shows that the default constraint on the THEME is DRUG.

colors of a car are white, black, silver and gray; other normal, but less common, colors are red, blue, brown and yellow; rare colors are gold and purple." The inventory of facets includes: *default*, which represents the most restricted, highly typical subset of fillers; *sem*, which represents typical selectional restrictions; *relaxable-to*, which represents what is possible but is not typical; and *value*, which represents not a constraint but an actual, non-overridable value. Objects and events are defined for an average of 16 properties each, many of whose fillers are inherited rather than locally specified. Unlike most ontologies, the OntoAgent ontology includes complex events, otherwise known as scripts [18], that support both simulation and reasoning about language and the world.

To summarize, the *meaning* of an object or event in the OntoAgent ontology can be understood as the meaning of its set of property-facet-value triples.

Excerpts from the frame view and tree view of the ontological concept TAKE-MEDICINE, as presented in the DEKADE development environment (http://www.trulysmartagents.org/dekade.php), are shown in Figs. 3 and 4.

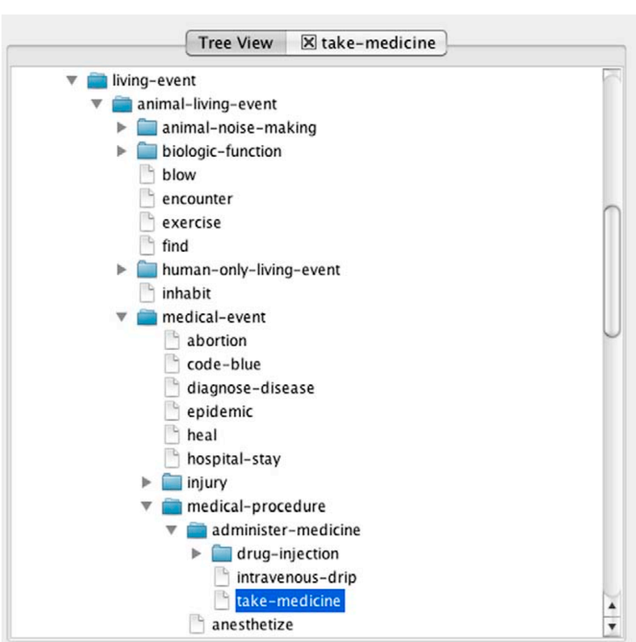

Fig. 4. The concept TAKE-MEDICINE shown in the ontological hierarchy.

Since the ontology is language independent, its link to any natural language must be mediated by a lexicon. Each lexical sense specifies which concept, concepts, property or properties of concepts defined in the ontology must be instantiated in the text-meaning representation to account for the meaning of a given lexical unit of input. For example, the English lexicon indicates that the one sense of *dog* maps to the concept DOG (a type of CANINE); another sense maps to HUMAN, further specified to indicate a negative evaluative modality (e.g., a woman can call her cheating ex-boyfriend a dog); and yet another sense maps to the event PURSUE.

Lexical senses for argument-taking words and modifiers are presented along with their typical syntactic configurations. For example, the lexicon entry for the 7th verbal sense of *have* is presented in Fig. 5, which shows the graphical user interface in which lexicon entries are viewed and edited in the OntoAgent environment. The syn–struc zone indicates the expected syntactic configuration, whereas the sem–struc zone indicates the meaning of the head word, its semantic relationship to the other elements of the structure, and any necessary constraints on other elements of the structure. Variables ($var1, $var2, etc.) are used to link the corresponding elements in the syn–struc and sem–struc zones.

In our example (Fig. 5), the syn–struc says that this is a transitive sense – i.e., the verb requires both a subject (referred to as $var1) and a direct object ($var2). The sem–struc indicates that the ANIMAL being realized by $var1 is the EXPERIENCER of the EVENT realized by $var2. The system knows that $var1 must be an ANIMAL and $var2 must be an EVENT because the ontological description of the property EXPERIENCER includes these constraints.[1] The sem–struc also includes the constraint that $var2 must be not just an EVENT, but specifically a MEDICAL-EVENT. If an input meets both the syntactic and the semantic expectations of this sense, then this sense offers a candidate interpretation of the input. To summarize, the OntoSem lexicon supports the combined syntactic and semantic analysis of texts, and the metalanguage of description in the sem–strucs of lexicon entries is identical to that used in the ontology. (For more on the lexicon and ontology, see [5,19].)

Text processing is carried out by the OntoSem text analysis system, whose top-level architecture is shown in Fig. 6. To give a taste of how intelligent agents in OntoAgent carry out text understanding, let us walk through the processing of a sentence that will be

[1] The semantic constraints recorded in the lexicon and ontology are used by the analyzer in the same way, with the resources being simultaneously leveraged during text processing.

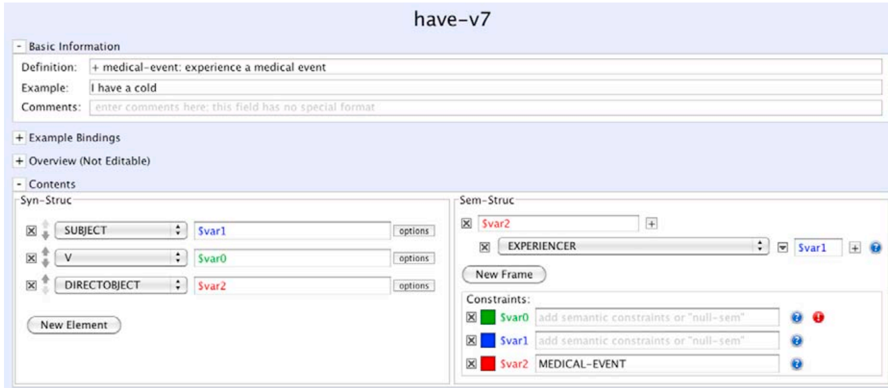

Fig. 5. The 7th verbal sense of *have*, used in contexts in which the direct object semantically represents a MEDICAL-EVENT.

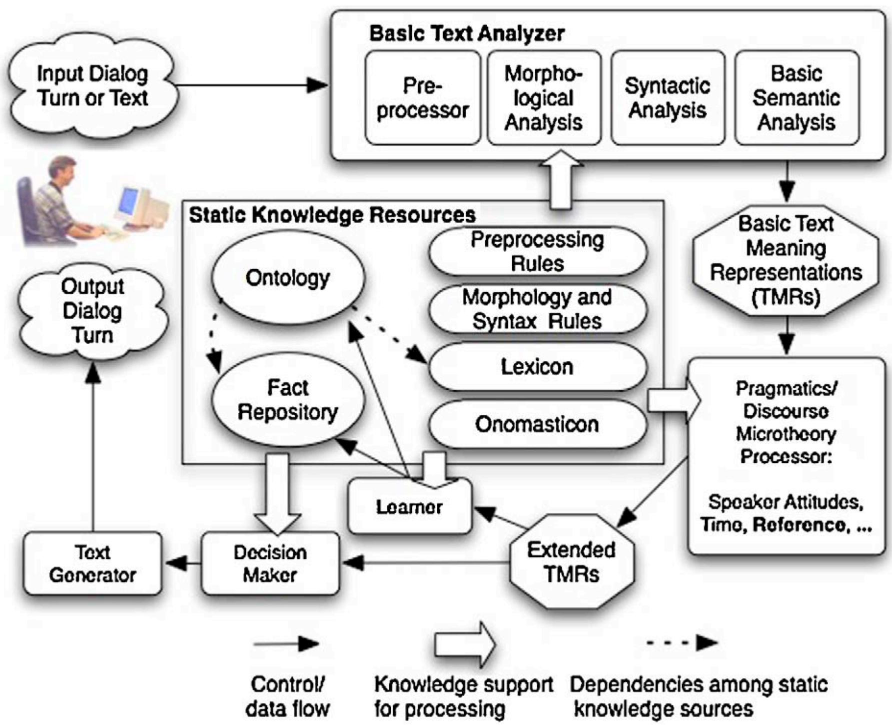

Fig. 6. The architecture of the OntoSem text analysis system.

returned to later as part of a more extended example of inconsistency detection: *Are you having any side effects*? This description is an encapsulation of several chapters of [5]. It necessarily assumes background knowledge of NLP and is intended to provide some idea of how OntoAgent agents understand text, with the understanding that many issues must remain unopened in this short space.

Step 1: Preprocessing and morphological analysis. The preprocessor tokenizes the text, then, for each word, the morphological analyzer determines the part of speech (PoS), root and morphological features (not listed here), as shown in Table 1. Phrasal entries, like “side effect”, are recognized as a complex root word since they are recorded as such in the OntoSem lexicon.

Step 2: Generating actual dependencies from the input sentence. We use the Stanford parser [20] to generate a syntactic dependency analysis. We call the output dependencies *actual* dependencies – i.e., dependencies from the actual input sentence – to distinguish them from *expected* dependencies recorded in the OntoSem lexicon for argument-taking words. The *actual* dependencies for our sentence are as follows:

aux(having-2, Are-0)
nsubj(having-2, you-1)
det(side_effects-4, any-3)
dobj(having-2, side_effects-4)
punct(having-2,?-5)

Step 3: Compiling *expected* dependencies from the OntoSem lexicon. Next, the OntoSem analyzer compiles a list of lexically-encoded *expected* dependencies for each lexical sense of each argument-taking word in the input. For

Table 1
Select results of preprocessing and morphological analysis.

Word #	String	PoS	Root
0	Are	aux	be
1	you	n	you
2	having	v	have
3	any	det	any
4	side_effects	n	side_effect
5	?	punct	?

example, for each transitive sense of a verb (like *have-v*7 above), it will extract from the lexicon the following *expected* dependency structure:

subject	$var1
v	$var0
directobject	$var2

Step 4: Converting *expected* dependencies into Stanford-compatible format. The analyzer then converts each *expected* dependency structure from the OntoSem lexicon into a Stanford-compatible format, like the following for our transitive, *have-v*7 example.

nsubj($var0, $var1)
dobj($var0, $var2)

This format permits more direct comparison between the *actual* dependency structure for the input sentence (generated by the Stanford parser) and the *expected* dependency structure for each argument-taking candidate word sense listed in the OntoSem lexicon.

Step 5: Linking Stanford's *actual* dependencies to OntoSem's *expected* dependencies. At this point the analyzer aligns the *actual* dependencies from the Stanford parse with the *expected* dependencies from the OntoSem lexicon and scores each alignment. We continue with the example of *have* from our input sentence. All senses of *have* that have a transitive syn–struc will be treated in the same way and receive the same Stanford-OntoSem linking score.

Expected dependencies from OntoSem transitive lexical senses of *have*, translated into Stanford-compatible format:

nsubj($var0, $var1)
dobj($var0, $var2)

Actual dependencies for *have* from the Stanford parse:

nsubj(have-2, you-1)
dobj(have-2, side_effects-4)

In this case, the alignment is straightforward, with $var1 binding to you-1, and $var2 binding to side_effects-4.[2] To summarize, the goal of this step is to generate an *exclusively syntactic* score (since the Stanford parser can only weight in about syntax) for each OntoSem sense of each argument-taking word in the input sentence. This syntactic score will later be combined with a semantic score to yield the best overall analysis of the text.

Step 6: Semantic analysis. All lexical senses that achieve a given scoring threshold based on their syntactic alignment with the Stanford dependency parse are then passed on to the semantic analyzer to be scored based on the semantic constraints recorded in the OntoSem lexicon and ontology.

Step 7: Finding the optimal solution. The text analyzer includes a constraint optimization engine called Hunter-Gatherer [21], which combines the syntactic and semantic scores of all candidate word senses and selects the optimal combination.

Step 8: Generating the text meaning representation. The output of text analysis is a text meaning representation (TMR) that is written in the metalanguage of Ontological Semantics [4]. The content of a given TMR is generated by combining the sem–struc representations of the lexical senses selected to convey the meaning of each word in the sentence. Each TMR frame is headed by an instance of an ontological concept.

Below is the TMR for our sample sentence *Are you having any side effects*?[3]

REQUEST-INFO		
theme	EXPERIENCER-1	
textstring	?	
from-sense	?-punct1	
EXPERIENCER-1		
DOMAIN	SIDE-EFFECT-1	
RANGE	HUMAN-1	
TIME	FIND-ANCHOR-TIME	; indicates present tense
textstring	have	
from-sense	have-v7	
SIDE-EFFECT-1		
DOMAIN-OF	EXPERIENCER-1	
reference-action	block-coreference	
HUMAN-1		
EXPERIENCER-OF	SIDE-EFFECT-1	
textstring	you	
from-sense	you-n1	

TMR frames contain both contentful, semantic elements and metadata. The contentful elements are instances of ontological concepts, written in small caps and followed by disambiguating numerical indices. The metadata includes the *textstring* slot, filled by the string that spawned the TMR frame, and the *from-sense* slot, filled by the lexical sense used for the interpretation.We have already discussed the disambiguation decision for *have* in some detail. Let us briefly mention the disambiguation of other elements of this input for the sake of completness.

Question mark. Question marks are treated like argument-taking words, being supplied with lexical entries that contain syntactic and semantic expectations for the question mark's "dependents". The question-mark sense used for this input is described as participating in a structure with a clause-initial auxiliary – such that the syntactic dependencies of the auxiliary are taken care of by this lexical sense as well. The expected dependencies listed in the OntoSem syn–struc of the entry – converted into Stanford-compatible format – are as follows.

Expected dependencies recorded in the OntoSem lexicon for "?" in sentences with a leading aux:

aux($var1,$var2)
punct($var1, $var0)

The *actual* dependencies generated by the Stanford parser for our input sentence:

aux(have-2, Are-0)
punct(have-2,?-5)

There is a perfect alignment between the expected and actual dependencies, with $var1 bound to *have*-2 and $var2 bound to *Are-0.*The sem–struc of this sense of question-mark indicates that the question mark instantiates a REQUEST-INFO event whose THEME is the meaning of the main verb – in our example, *have*.

Side effect. There is only one lexical sense for the word *side_effect*, which maps to the concept SIDE-EFFECT, a descendant of MEDICAL-EVENT.

Any. The meaning of *any* in our context is not rendered as an ontological concept; instead, this word triggers a

[2] An example of an imperfect Stanford-to-OntoSem linking would be if the input sentence contained [subject, verb, complement], as in *She said that he's running late*, whereas the lexical sense being considered expected [subject, verb, direct object]. The given lexical sense of *say* would not be a good fit for the input sentence and would get a low syntactic score.

[3] For ease of understanding, we represent *side effects* without the set notation typically used for plurals. Using set notation, the TMR frame would be headed by an instance of SET with MEMBER-TYPE: SIDE-EFFECT and CARDINALITY: (>1).

procedural semantic routine that blocks the search for a coreferential antecedent for *side effects*. Procedural semantic routines are encoded in the lexicon in a field not shown in Fig. 5 above. They are run during the stage of microtheory processing in Fig. 6. For more on procedural semantic routines in OntoSem, see [5,22]; and for more on reference resolution, see [5,9].

Step 9: Storing information in memory (the fact repository). The important (as determined by the application) aspects of text meaning representations are remembered by agents in their fact repository, which is the memory of assertions. Among the main differences between raw text meaning representations and fact repository entries is that the fact repository reflects the results of reference resolution. For example, if a given instance of SIDE-EFFECT is referred to many times, there will be one fact repository "anchor" – say, SIDE-EFFECT-FR8 – that will be the locus of all of the information about that event derived from all of the individual, coreferential text meaning representation frames.

Language processing wrap-up. This very fast walk through language processing in OntoAgent is intended to convey that agents carry out deep language understanding using a large suite of static knowledge resources and processors. The result of language processing is storing ontologically interpreted knowledge in the agent's fact repository (memory). The agent then carries out all reasoning and decision-making on the basis of storied memories.

3. Comparisons with others

Before moving from the overview of OntoAgent (Section 2) to the new work to be reported here (Section 4), let us briefly draw some comparisons with other approaches to related problems of artificial intelligence.

The work that most closely parallels that of OntoAgent is being pursued by James Allen and collaborators who, like us, pursue classical AI goals – modeling intelligent agents with human-like capabilities – using extensive knowledge bases, deep language processing, agent reasoning, and agent learning. Recent applications include Chester [23], which gives advice about taking medication, and PLOW [24], which learns task models from a collaborative learning session. These systems exploit the TRIPS [25] environment for dialog processing and agent reasoning.

Comparisons with other groups and approaches are far more distant. Consider, for example, dialog processing, which is a central component of OntoAgent. Most dialog systems cover narrow domains of discourse and encode only those lexical elements that pertain to a domain, thus artificially simplifying the very difficult problem of lexical ambiguity resolution. Examples of such systems include: DUDE, A Dialog and Understanding Development Environment [26], which supports the development of dialog systems covering routine questions in business domains, like ordering tickets; and ALFRED, active logic for reason enhanced dialog [27], which has been shown to be able respond to user needs in controlling pool temperature settings, moving a toy train, playing different movies, etc. In OntoAgent, by contrast, the domain of interest is much broader and the problem of lexical ambiguity is faced head-on, with no hand-picking of lexical senses for a specific domain.

Another point of contrast with others regards OntoAgent's definition of virtual patient. The most common type of ìvirtual patientî is not a simulation at all but, rather, a branching narrative scenario organized as a decision tree that depicts a specific medical case from beginning to end (e.g., MedCases, Inc.). In these, user options are restricted and responses are highly pre-scripted and delivered through multiple-choice questions. Most importantly, in such systems, (a) creating scenarios is very long and laborious, (b) scenarios are difficult, if at all possible, to modify in response to changes in medical knowledge and practice without completely rewriting the scenario (not so of our ontological models, which can be readily modified), and (c) patient outcomes are fully predetermined by the scenario. The center of gravity in "branching scenario" R&D is presentation issues [28,29]. Good visualization solutions – sometimes using off-the-shelf speech recognition and synthesis software – become the main avenue of enhancing the verisimilitude of the interactive experience. Videos of human actors, advanced graphics, including avatars, the incorporation of off-the-shelf speech recognition and synthesis software and visualization of results of medical tests (such as X-rays and MRIs) have been prominent among the means of strengthening the verisimilitude of the human-computer interactive experience. Indeed, the latest VP environments, e.g., [30] show significant progress even when compared to quite recent ones, e.g., [31].

The recent literature includes several contributions analyzing the needs of medical AI and comparing those needs with available technologies. As concerns virtual patients, Stead and Lin's [32] desiderata closely match the features of OntoAgent agents, and Stead and Lin's belief in the need for "radical change" – i.e., the development of truly sophisticated functionalities developed over a long time – aligns with the OntoAgent philosophy as well.

The emphasis in OntoAgent is on acquiring knowledge that is sufficiently deep to support the complex reasoning, simulation and language processing required by the application. This is in contrast to many recent and current approaches – notably, in natural language processing – that stress broad coverage of data at the expense of the depth of acquired knowledge. Specifically, MVP utilizes knowledge-rich approaches to NLP and agent modeling that were more widely pursued 20 or 30 years ago than they are today. Interest in plan- and goal-based R&D, as well as deep-semantic NLP, dwindled when investigators concluded that they were too labor-intensive to support practical applications. However, these conclusions must be reassessed, particularly when juxtaposing past efforts with the knowledge acquisition effort in MVP, which we managed to keep within realistic limits of labor intensity while maintaining sufficient depth and breadth of coverage for support of the complex simulation, reasoning and natural language communication application.

Much of the research on plan- and goal-based reasoning in AI was devoted to creating systems that *built* plans on the fly. In MVP, by contrast, we imposed the constraint that the system would not be required to develop plans, it would only be required to use preconstructed plans. This simplifying constraint is well-suited to MVP since system users will not be asked to solve never before seen types of cases, and the system itself – in the guise of the virtual mentor – will not be asked to invent novel approaches to patient care or fundamentally creative responses to questions.

4. Inconsistency as a trigger for exploration

As people, we use inconsistencies as a trigger for reanalyzing situations, questioning our understanding of them and/or seeking further information. Inconsistencies arise when something observed/reported conflicts with our expectations, or when multiple bits of evidence/reports conflict with each other. So, if a pancake recipe (for a family, not an army) calls for a cup of salt, you can confidently assume a typo based on knowledge of how far a little salt goes; and if one contractor offers to do a job for $1000 whereas another comes in at $12,000, you would be wise not to blindly choose the lower bidder. In the first example, the inconsistency stops you from making an inedible breakfast; in the second example, the inconsistency leads you to investigate whether different approaches to solving the problem are being

Table 2
Some diagnostic inconsistencies relevant for a clinician.

Type	Phenomena	Opportunities afforded
1	Test results do not support the doctor's hypothesis about the patient's condition.	Do more testing to verify results. Rethink hypothesis.
2	The result of a treatment trial does not support a clinical diagnosis.	Rethink hypothesis. Motivate patient to adhere to treatment regimen.
3	Information reported by the patient is inconsistent with observations.	Determine if patient is not telling the truth. Teach/correct/encourage patient. Reevaluate observations.

proposed (short-term vs. long-term), and/or whether you are dealing with a shady contractor.

Inconsistency can offer equally rich opportunities in the domain of clinical medicine, as shown by the examples in Table 2 and discussed in turn subsequently.

Type 1: Test results do not support the physician's hypothesis about the patient's condition. A common sequence of events in clinical medicine is for a clinician to hypothesize a diagnosis based on a patient interview and then follow up with medical testing. If the testing does not corroborate the hypothesis – thus representing an inconsistency in the doctor's mental model of the patient's condition – this could represent several situations: the hypothesis was wrong; the hypothesis was correct but the condition is not advanced enough to be corroborated by the test; the hypothesis was correct but the test results were flawed. The way the doctor responds to this inconsistency depends upon the interaction of many factors, including at least *the severity of the patient's condition* (Is "wait and see" an option?), the *strength of the hypothesis* (Could another hypothesis readily account for the available data?) and the understood *reliability of the test* (How likely is a testing error?).[4] The opportunities afforded by this type of inconsistency include rethinking the hypothesis and/or verifying the test results, both of which are in service of better management of the patient.

Type 2: The result of a treatment trial does not support the physician's hypothesis about the patient's condition. In clinical medicine, two types of diagnoses can be distinguished: *definitive diagnoses* are confirmed by medical testing, whereas *clinical diagnoses* are suggested based on the success of a therapeutic intervention as a diagnostic test. As an example of the latter, if a doctor believes a patient has gastroesophageal reflux disease (GERD) he can prescribe medication to reduce stomach acid; if the medication improves the symptoms, a clinical diagnosis of GERD can be posited. If a diagnostic treatment trial proves ineffective for a patient, this inconsistency could have several explanations: the patient could have shown poor adherence and admitted to it; the patient could have shown poor adherence but hidden it; the hypothesis was correct but the patient was a non-responder; or the hypothesis was incorrect to begin with. The case of admitted lack of adherence is relatively straightforward: barring significant changes in patient health or the patient's evaluation that he will again not be able to comply (as might occur if the treatment involves a difficult-to-change lifestyle habit), the treatment protocol is typically repeated. Managing the other three cases, however, involves a decision whose input parameters include, non-exhaustively: an evaluation of the strength of the hypothesis; the availability and likelihood of alternative hypotheses; the efficacy rate of the treatment; and whether or not the patient has a reason to misrepresent his compliance (e.g., a teenage girl afraid of gaining weight might refuse to take a drug with the side effect of weight gain, and might be afraid to admit that to the doctor). The opportunities afforded by this type of inconsistency include the clinician's rethinking of the hypothesis, which might have been incorrect, and the physician's encouraging the patient to comply with the treatment regime, if the patient ultimately admits to non-compliance. The matter of detecting and managing instances of patients not telling the truth is what we turn to next.

Type 3. Information reported by the patient is inconsistent with observations. If there was ever evidence for clinical medicine involving a combination of art and science, this situation is the star example because, intentionally or not, patients do not always tell the truth. Part of the art of clinical medicine is determining whether or not the patient's report is likely to be true and, if not, why. The "why" helps the physician to remedy the situation in a way that is both compassionate and effective. Consider some of the many reasons why a patient might not tell the truth:

- The patient fails to understand some information but is embarrassed to admit it. This can be due to many reasons, including insufficient medical literacy [33], difficulty processing verbal input, or a language barrier. The results include inadvertent misinterpretations of medication dosing, suboptimal post-operative home care, etc.
- The patient fails to understand the importance of something and therefore considers its misrepresentation to be inconsequential. For example, some medications must be taken in a specific temporal relationship to the ingestion of food. If a patient does not understand that the medication loses efficacy if taken otherwise, he might report that he is taking it on schedule even though he is not.
- The patient has beliefs that he values more highly than telling the whole truth. For example, patients of some socio-cultural background underrepresent their symptoms due the belief that it is not honorable to admit to symptoms [34].
- The patient has some priority that does not align with the common doctor–patient goal of achieving effective treatment. For example, if a teenage organ-transplant patient is prescribed a high dose of the steroid Prednisone to prevent rejection of the new organ, she might balk at the acne it causes as a side effect and decide that clear skin is more important to her than the potential risks of not taking the medication. Of course, her priorities would be seriously misguided since not taking the medication could lead to loss of the organ and possibly death, which is why the physician must be on the alert to detect lack of compliance and encourage a subsequent change of behavior.
- The patient does not want to admit to a lack of willpower, as is necessary for carrying out lifestyle modifications such as weight loss or conquering addiction.
- The patient is embarrassed by a symptom, such as loss of sex drive or flatulence.
- The patient is afraid of legal or other repercussions, as from illicit drug use.

This list could go on. The point is that there are a lot of reasons why a patient might not tell the doctor the truth, and the doctor must not only be able to detect such instances (*Why does this 16-year old girl on a high dose of Prednisone not have acne*?) but also respond to them in a way that supports his collaboration with the patient. This latter is a cornerstone of patient-centered medicine, which interprets the patient as an active partner in his own health care [35].

[4] The reliability of medical tests is statistically measured using *sensitivity* (the likelihood that the test will be positive given a patient with the condition) and *specificity* (the probability that the test will be negative given a patient without the condition). Errors can occur for many reasons, including flawed administration of the test, test results being interpreted incorrectly by specialists, and a failure to recognize interfering factors, such as the effects of current medications.

If the doctor does suspect that the patient is not telling the truth, the next step is to determine whether or not the suspicion is justified. The goal is to foster patient cooperation and collaboration, not to scold, embarrass or alienate the patient. In most cases, this will involve showing compassion and understanding. Among the strategies available to the clinician are: *verify* that the patient understands the information by engaging in a conversation about it; *explain* the importance of precisely reporting key information; *explain* why he is asking the question; *reassure* the patient that there is nothing to be embarrassed about; *sympathize* about the difficulties faced by the patient; *explain* the consequences of misreporting; *clarify* medical terminology; *learn about the patient* as a person, asking him about his beliefs, priorities, values, etc.; and, in cases in which a patient is endangering his own well-being, *scare* the patient with the possible ramifications of his actions. In all cases, the plan selected by the clinician should be in service of the joint goal of making the patient as healthy as possible. In many cases the inconsistency between the patient report and observed reality can be used by the clinician to explain his motivation for further pursuing specific factual details.

Of course, an inconsistency between an observed outcome and a patient report does not necessarily derive from a patient's misrepresentation of the state of affairs: the patient might be an actual outlier or might have made a genuine mistake. As such, a clinician's knowledge about possible causes for misreporting must be combined with other available data, such the patient's character traits and specifics of his condition.

Learning to skillfully manage these types of scenarios is central to mastering the art of clinical medicine. The MVP application described above seeks to foster the acquisition of this set of skills – in addition to many others. The optimal pedagogical support, we believe, is a tutoring agent endowed with the skills of an expert clinician who will follow the moves of the user (a clinician in training) as he or she practices interacting with VPs and will intervene with advice when necessary.

For this discussion, we select an example from a medical domain that is not yet incorporated into MVP: organ transplant surgery (MVP currently covers only diseases of the esophagus). We pursue this hypothetical example because it provides a clear illustration of the problem space under discussion – i.e., exploiting incongruity to clinical advantage. It must be emphasized, however, that all aspects of language understanding, memory augmentation and reasoning are the same no matter which medical domain is being addressed by MVP.

In our example scenario, the virtual patient is a 16-year old girl who had a kidney transplant a month ago. She was prescribed a high dose of Prednisone to prevent organ rejection and does not have acne. She is back for her first follow-up visit. An excerpt from the physician-patient dialog might go as follows, with the virtual tutor chiming in at the point at which it realizes that the physician has accepted the VP's affirmation of medication compliance without question.

Doctor	So, how have you been feeling?
VP	A little tired.
Doctor	Have you been taking your Prednisone?
VP	Yeah.
Doctor	Are you having any side effects?
VP	No.
Doctor	Has your wound healed?
Tutor	*Back up a step: the absence of side effects with this dose of Prednisone is unusual in patients her age.*
Doctor	Let's return to that last question for a second – have you been taking the full dose of Prednisone exactly as scheduled? The reason I ask is because, as we discussed before your surgery, people your age almost always get acne from this medication.
Tutor	*Nice bedside manner.*

Before discussing how this exchange is processed and interpreted by the tutoring agent, consider select aspects of the tutor's knowledge – stored in the fact repository (FR) – about the live doctor and the virtual patient prior to the exchange.

HUMAN-FR1 is the doctor in training, Frederick Jones. He has little medical experience, as shown by the value .3 on the abstract scale {0,1}. The tutor has developed a model of Frederick's ontology based on (a) expectations about students with this level of training and (b) evidence from Frederick's interactions. The latter information is associated with a higher confidence level than the former since it has been explicitly attested. Every time Frederick does something agentive while using the system – asks a question, orders a test, explains something – that action is stored as a filler of the AGENT-OF slot is his FR frame. The same is true of every other semantic role Frederick can play in an event – be he the BENEFICIARY-OF it, the THEME-OF it, and so on.

HUMAN-FR1

HAS-PERSONAL-NAME	FREDERICK
HAS-SURNAME	JONES
MEDICAL-EXPERIENCE	.3
HAS-ONTOLOGICAL-MODEL	[ontology.human-fr1]
HAS-FR	[fr.human-fr1]
AGENT-OF	REQUEST-INFO-FR1
AGENT-OF	ORDER-TEST-FR45
BENEFICIARY-OF	RESPOND-FR12
...	

The tutor creates a similar fact repository profile of the virtual patient, Sherry Palmeri, throughout the simulation, a short excerpt of which is as follows:

HUMAN-FR2

HAS-PERSONAL-NAME	SHERRY
HAS-SURNAME	PALMERI
AGE	16
GENDER	FEMALE
AGENT-OF	RESPOND-FR1
EXPERIENCER-OF	KIDNEY-FAILURE-FR1
EXPERIENCER-OF	SURGERY-FR1
...	

The concept instances used as fillers in these personal profiles are also expanded into full frames in the fact repository. The frames below say that the post-operative care for Sherry's surgery includes taking medicine; that the medicine prescribed is Prednisone with the indicated dosage and duration; and that the purpose of taking the medication is to prevent rejection of Sherry's kidney.

SURGERY-FR1

EXPERIENCER	HUMAN-FR2
POST-OP-CARE	TAKE-MEDICATION-FR1

TAKE-MEDICATION-FR1

POST-OP-CARE-FOR	SURGERY-FR1
AGENT	HUMAN-FR2
THEME	PREDNISONE
DOSAGE	60 MG/DAY
DURATION	1 MONTH
PURPOSE	PREVENT-FR1[5]

PREVENT-FR1

THEME	ORGAN-REJECTION-FR1	; a simplification

ORGAN-REJECTION-FR1

EXPERIENCER	HUMAN-FR2	
THEME	KIDNEY-FR1	; Sherry's new kidney

In addition to fact repository knowledge, the virtual tutor also has extensive ontological knowledge, both in the medical domain and in the general domain. Let us begin with the general domain, focusing on the issue of not telling the truth, which is referred to ontologically by the concept LIE. It is reasonable to assume that the average adult human knows that people can lie and understands

[5] We posit a "prevent" event for simplicity's sake. The notion of "prevent" is actually represented ontologically using modality scoping over a proposition.

at least some reasons why someone might lie. As such, part of the general ontology of all of our agents is a script about lying. Rather than present a metalanguage version of the script for lying, we will render its relevant contents in English. The PURPOSE of LIE is DECEIVE. Detection of lying is recorded as a series of conditional statements associated with probabilities, such as the following:

1. If a reported event or state is impossible, then the likelihood of that report being a lie is >99%. For example, an obese patient reports having eaten fewer than 500 calories a day for a month but has not lost any weight.
2. If an event is reported to have taken place but its high-probability effects or side effects do not occur, then the likelihood of that report being a lie is a function of: (a) the probability of the given effects or side effects, (b) the value of DIFFICULTY-ATTRIBUTE for the event, (c) the value of EMBARRASSMENT-LEVEL for the event, its subevents or effects, and (d) the value of DISTRESS-LEVEL for the event, its subevents or effects. For example, dieting is *difficult* and failing to stick to a diet (a potential subevent of dieting) can be *embarrassing*, so if a person reports that he has been on a strict diet but has not lost any weight – i.e., a very high probability effect did not occur – then there is a high probability that the person is lying.
3. An event is considered unacceptable or frowned upon by the group to which the individual belongs and the event is a typical symptom or subevent of an event that did happen or is happening. For example, a particularly macho man known to have bone cancer might report little or no pain.

The tutor's ontology is also rich with medical knowledge. Factoids that are important for our example are as follows, with a gloss preceding each knowledge structure.

There is a great potential for dying if an organ is rejected.

```
ORGAN-REJECTION
  EFFECT                    DIE (POTENTIAL (>.9))
```

Taking a large dose of Prednisone for a period of more than 2 weeks has a very high chance of causing severe acne in teens (similar structures exist for different durations, ages, etc.).

```
TAKE-MEDICATION
  AGENT HUMAN (AGE (> < 13 19))
  THEME PREDNISONE (QUANT (> .6))
  SIDE-EFFECT
    IF (DURATION (>2 (MEASURED-IN WEEK)))
    THEN (ACNE ((EXPERIENCER TAKE-MEDICATION.AGENT)
    (SEVERITY (>.7)) (PROBABILITY .9)))))
```

The distress and embarrassment caused by acne can range from mild (.2) to severe (.9).

```
ACNE
  DISTRESS-LEVE             .2 > < .9
  EMBARRASSMENT-LEVEL       .2 > < .9
```

Lying can be caused by fear, embarrassment, non-compliance, etc.

```
LIE
  CAUSED-BY                 FEAR, EMBARRASSMENT, NON-COMPLIANCE, . . .
```

Non-compliance can be caused by side-effects of medication.

```
NON-COMPLIANCE
  THEME                     TAKE-MEDICATION
  CAUSED-BY                 TAKE-MEDICATION.SIDE-EFFECT
```

It is important to understand that ontological knowledge can be recorded at any grain size. For example, eating a sandwich can be described from as few as 3 steps (bite, chew, swallow) to many hundreds of steps (extend hands toward sandwich, grasp sandwich, pick up sandwich, raise sandwich toward mouth, . . . followed by an almost infinitely complex rendering of the physiology of swallowing and digesting the food). We aim for a grain-size that is just detailed enough to support our simulation needs.

The final aspect of knowledge needed by the tutor to make a successful intervention in our dialog is his inventory of goals and their associated plans. On the one hand, he has the "expert" version of the goals and plans that the doctor in training is attempting to master by using the MVP system. Relevant for our example are the goals (1) KNOW-PATIENT-SYMPTOMS whose associated plans include REQUEST-INFO (i.e., ask a question), PHYSICAL-EXAM, DETECT-LYING, PURSUE-LYING-HYPOTHESIS, etc. (we exclude details about how these events are related in the DIAGNOSE-PATIENT script) and (2) COLLABORATE-WITH-PATIENT, whose associated plans include SHOW-EMPATHY, EXPLAIN-QUESTIONS, LEARN-PATIENT-PRIORITIES, etc. The second type of tutor goals and plans apply directly to tutoring. Among these are the goals (1) AVOID-OVERSIGHT whose plans include WARN-ABOUT-OVERSIGHT, and (2) PROVIDE-POSITIVE-FEEDBACK whose plans include REINFORCE-GOOD-BEDSIDE-MANNER.

Returning to our dialog exchange, when the tutor follows the conversation of the doctor and patient, it is playing two roles: the role of expert physician and the role of tutor. As an expert physician, it evaluates every move of by the doctor and evaluates if it correlates with the tutor's own mental model of what should be done. If there is an inconsistency between the action of the doctor and the good practices recorded in the tutor's mental model, the inconsistency triggers the need for a tutoring move. (Actually, it can trigger one of many tutoring moves – among them being "let the doctor continue to make his mistake and see what happens" – but we constrain the discussion to the move of alerting the doctor to his actual or potential mistake.)

Earlier, we showed the text meaning representation generated for the sentence *Are you having any side effects?* It would take us too far afield to describe how OntoSem interprets the fragment response "*No*" (for that, see [5,36]), but suffice it to say that the result of interpreting this pair of dialog turns is for the tutor's FR to be supplemented by the boldface information about the VP below:

```
HUMAN-FR2
  HAS-PERSONAL-NAME         SHERRY
  HAS-SURNAME               PALMERI
  AGE                       16
  GENDER                    FEMALE
  AGENT-OF                  RESPOND-FR1
  AGENT-OF                  TAKE-MEDICINE-FR1
  EXPERIENCER-OF            KIDNEY-FAILURE-FR1
  EXPERIENCER-OF            SURGERY-FR1
  EXPERIENCER-OF            [NOT] SIDE-EFFECT-FR1[6]
  . . .
```

Every time a VP responds to a question, the tutor's DETECT-LYING function is activated, representing the cognitive reality that people are always on the lookout for lies, even if they are not paying attention to that at all times. As described earlier, one of the conditional statements in the detect-lying function is: *If an event is reported to have taken place but its high-probability effects or side effects do not occur, then the likelihood of that report being a lie is a function of: (a) the probability of the given effects or side effects, (b) the value of* DIFFICULTY-ATTRIBUTE *for the event, (c) the value of* EMBARRASSMENT-LEVEL *for the event, its subevents or effects, and (d) the value of* DISTRESS-LEVEL *for the event, its subevents or effects.* Applying this to our context, the function looks as follows:

[6] We use "[not] SIDE-EFFECT" as a pretty-printed representation of our canonical method for indicating negation: scoping EPISTEMIC modality with a value of 0 over the event.

Likelihood of ACNE for a teenager on a high dose of Prednisone:	.9
Value of DIFFICULTY-ATTRIBUTE FOR TAKE-MEDICATION:	.05
Value of EMBARRASSMENT-LEVEL FOR TAKE-MEDICATION:	0
Value of EMBARRASSMENT-LEVEL FOR ACNE (an EFFECT OF TAKE-MEDICATION):	.2 > < .9
Value of DISTRESS-LEVEL FOR TAKE-MEDICATION:	0
Value of DISTRESS-LEVEL FOR ACNE:	.2 > < .9

The combination of the very high likelihood of ACNE and the potentially very high levels of embarrassment and distress caused by ACNE lead to a high value for this function, the mathematical details of which we leave aside. As such, the tutor suspects – and therefore believes that the doctor should suspect as well – that the patient might be lying. The fact that the doctor does not pursue this hypothesis but, rather, moves on to a new question, indicates to the tutor that the doctor failed to detect a potential lie, which triggers the tutoring intervention. Obviously, the intervention cannot occur before the next question is asked because the tutor cannot read the doctor's mind.

The tutor will recognize the doctor's good bedside manner using a similar strategy. One of the standing goals of the tutor is to reinforce good bedside manner, which includes, among other things, EXPLAIN-QUESTIONS, which is exactly what the doctor does. Whereas a well-configured tutor will not intervene every time a user does something right – a practice that would be highly annoying – this intervention shows that the tutor *is* evaluating every move as being either *con*sistent or *in*consistent with the tutor's expert knowledge of the processes of diagnosing and treating a patient.

Implementation of these agent functionalities described above is currently in progress in our broad program of work on goal- and plan-based reasoning by intelligent agents.

5. Discussion

This paper has shown that just as inconsistency is a central tool for human clinicians and tutors, so can it be for intelligent agents modeled to fulfill these roles. In fact, inconsistency is just one of many phenomena that, while most typically looked upon as "bugs" – or at least headaches – by developers, actually can be recast as useful features for intelligent agents. By way of situating the reported work in a broader context, let us briefly consider a few more traditionally marginalized phenomena that are actually useful in societies of human-like intelligent agents.

In the fields of semantic analysis and reference resolution, most practical work has pursued the goal of finding exactly one precise meaning for every word and exactly one precise sponsor for every referring expression, respectively. However, referential ambiguity and underspecification are not always mistakes – they afford the benefit of freeing speakers from introducing unnecessary details and making irrelevant distinctions, and they free listeners from interpreting such details and distinctions. For example, sentence (1) below shows referential ambiguity for the expression *it* since this expression can be understood to refer to the coffee, the cup, or the coffee and cup together. Since all of these interpretations would lead to the similar knowledge base amendments and would support similar reasoning by the intelligent agent, this ambiguity can be considered benign. (For discussion of related issues in corpus annotation, see [37].)

1. Please pour some fresh coffee into your dad's favorite mug and bring it to him upstairs.

Example (2) illustrates semantic underspecification: it is probably not the case that every minute of the school day was boring – the point is that the child's overall impression of school was boredom. Thus, the focus is on the child's experience, not the particular activities that made him feel that way.

2. [A child walking home from school with his pals] Man, that was boring as always!

Examples such as these suggest that the narrow, traditionally understood goals of semantic analysis and reference resolution – as pursued in various narrow-task competitions – are, at best, incomplete. Instead, the option for fuzzy reference and underspecification must be available to agents, with further specification being pursued only if agent goals should require that knowledge. The important point here is that the impetus to reinterpret certain intermediate goals of text analysis derives from the broader goal of creating intelligent agents that do something useful. To put a finer point on it: semantic disambiguation and reference resolution in isolation are of no use by themselves, and positing "end goals" for events that are, at base, not an end in themselves is not justified. Similarly, inconsistency is, in the abstract, neither positive nor negative – it simply *is* in human societies, and should, in turn, be exploited and managed in societies of agents that include or model humans.

As we said in the introduction, the term *inconsistency* tends to imply the absence of a goal state of consistency. However, in societies of people and human-like intelligent agents, inter-agent inconsistencies are norm: every agent will have a different ontology, lexicon and fact repository, as well as different decision functions deriving from its personality traits and physical and mental states. As such, imposing the expectation of consistency on a society of intelligent agents would simply be wrong. Whereas not all inconsistencies are as directly *useful* as the ones described above, they can be managed equally well when planned for properly.

In the context of clinical medicine, two types of inconsistency among agents are of particular importance: differences between the factual knowledge bases (ontology and lexicon) of physicians and patients, and differences in the priorities and preferences of physicians and patients, as reflected in their decision functions. That is, if a physician wants to make fast and effective progress with a patient, he should (a) attempt to predict what the patient does and does not know, (b) talk in a language the patient understands, (c) do his best to make sense of the expectedly non-technical descriptions provided by the patient and (d) be prepared to teach the patient what he needs to know. In addition, if the physician wants the patient to comply with a treatment protocol, he should collaborate with the patient, taking into consideration his priorities and preferences, even if they do not align with his own. Within the OntoAgent environment, we are working on four core capabilities that permit agents to manage inter-agent inconsistencies: the dynamic modeling of the knowledge bases of other agents; the management of linguistic and meta-language *paraphrase* [7,8]; the ability to teach and learn new ontological and lexical knowledge [11]; and the ability to collaborate in decision-making [3]. The modeling of these capabilities shows many of the same expectation-oriented characteristics as the modeling of the capabilities that permit agents to exploit the diagnostically useful inconsistencies described in this paper.

A natural question is, Is the OntoAgent approach really feasible? We believe that it is for, among others, the following reasons. Our development efforts are targeted toward specific applications: there is no attempt to develop a fully generalized, plug-in ready cognitive architecture (like, e.g., TRIPS [25]), or to implement a domain-independent dialog system, or to equip system agents with all of the plans and goals of human beings, or to endow them with the full spectrum of possible character traits (as is done in theoretical approaches to affective modeling). Instead, theoretical and practical advancements are geared toward the near- and long-term future of specific systems, with infrastructure decisions being made with a long-term view but knowledge support targeted at near- and mid-term goals. In addition, we are not attempting to model every known aspect of human physiology, we are modeling only those that are needed to support a lifelike simulation at a grain-size that fulfills all current and expected goals for the system.

References

[1] McShane M, Fantry G, Beale S, Nirenburg S, Jarrell B. Disease interaction in cognitive simulations for medical training. In: Oxley L, Kulasiri D, editors. MODSIM 2007 International congress on modelling and simulation. Modelling and Simulation Society of Australia and New Zealand; 2007. p. 74–80.

[2] McShane M, Nirenburg S, Beale S, Jarrell B, Fantry G. Knowledge-based modeling and simulation of diseases with highly differentiated clinical manifestations. In: Bellazzi R, Abu-Hanna A, Hunter J, editors. Proceedings of the 11th conference on artificial intelligence in medicine (AIME 07). Amsterdam, The Netherlands, Berlin, Heidelberg: Springer-Verlag; 2007. p. 34–43.

[3] Nirenburg S, McShane M, Beale S. A simulated physiological/cognitive "double agent". In: Beal J, Bello P, Cassimatis N, Coen M, Winston P, editors. Papers from the AAAI fall symposium, naturally inspired cognitive architectures. AAAI technical report FS-08-06. Menlo Park, CA: AAAI Press; 2008.

[4] Nirenburg S, Raskin V. Ontological semantics. Cambridge, MA: The MIT Press; 2004.

[5] McShane M, Nirenburg S, Beale S. Meaning-centric language processing. Technical report #01-12. Institute for Language and Information Technologies. University of Maryland Baltimore County; 2012.

[6] McShane M, Nirenburg S. Dialog modeling within intelligent agent modeling. In: Jönsson A, Alexandersson J, Traum D, Zukerman I, editors. Proceedings of the IJCAI-09 workshop on knowledge and reasoning in practical dialog systems. 2009. p. 52–9.

[7] McShane M, Nirenburg S, Beale S. Resolving paraphrases to support modeling language perception in an intelligent agent. In: Bos J, Delmonte R, editors. Semantics in text processing: STEP 2008 conference proceedings, Venice, Italy. London: College Publications; 2008.

[8] McShane M, Nirenburg S, Beale S. Two kinds of paraphrase in modeling embodied cognitive agents. In: Samsonovich AV, editor. Biologically inspired cognitive architectures. Papers from the AAAI Fall Symposium. AAAI Technical Report FS-08-04. Menlo Park, CA: AAAI Press; 2008. p. 162–7.

[9] McShane M, Nirenburg S, Beale S. Reference-related memory management in intelligent agents emulating humans. In: Langley P, editor. Advances in cognitive systems. Papers from the AAAI fall symposium. AAAI technical report FS-11-01. Menlo Park, CA: AAAI Press; 2011. p. 232–9.

[10] McShane M, Jarrell B, Fantry G, Nirenburg S, Beale S, Johnson B. Revealing the conceptual substrate of biomedical cognitive models to the wider community. In: Westwood JD, Haluck RS, Hoffman HM, Mogel GT, Phillips R, Robb RA, et al., editors. Medicine meets virtual reality 16. Amsterdam, Netherlands: IOS Press; 2008. p. 281–6.

[11] Nirenburg S, McShane M, Beale S, English J, Catizone R. In: Samsonovich AV, Jóhannsdóttir KR, Chella A, Goertzel B, editors. Four kinds of learning in one agent-oriented environment. Amsterdam, Netherlands: IOS Press; 2010. p. 92–7.

[12] Nirenburg S, McShane M, Beale S. A unified ontological-semantic substrate for physiological simulation and cognitive modeling. In: Proceedings of the first international conference on biomedical ontology (ICBO-2009). 2009.

[13] McShane M, Nirenburg S. A knowledge representation language for natural language processing, simulation and reasoning. International Journal of Semantic Computing, in press.

[14] Nirenburg S, McShane M, Beale S. Hybrid methods of knowledge elicitation within a unified representational knowledge scheme. In: Filipe J, Dietz JLG, editors. KEOD 2010 – Proceedings of the international conference on knowledge engineering and ontology development. Valencia, Spain: SciTePress; 2010. p. 177–82.

[15] McShane M, Nirenburg S, Jarrell B, Beale S, Fantry G. Maryland virtual patient: a knowledge-based, language-enabled simulation and training system. Bio-Algorithms and Med-Systems 2009;5(9):57–63.

[16] Bratman M. Faces of intention. Cambridge, UK: Cambridge University Press; 1999.

[17] Wooldridge M. Computationally grounded theories of agency. In: Durfee E, editor. Proceedings of the fourth conference on multiagent systems (ICMAS'00). Washington, DC, USA: IEEE Computer Society; 2000. p. 13–22.

[18] Shank R, Abelson R. Scripts, plans, goals and understanding. Hillsdale, NJ: Erlbaum; 1977.

[19] McShane M, Nirenburg S, Beale S. An NLP lexicon as a largely language independent resource. Machine Translation 2005;19(2):139–73.

[20] de Marneffe M., MacCartney B, Manning CD. Generating typed dependency parses from phrase structure parses. In: Proceedings of The fifth international conference on language resources and evaluation (LREC 2006). Distributed on CD by ELRA, Paris, France. 2006. Available at http://www.lrec-conf.org/proceedings/lrec2006/.

[21] Beale S. Hunter-gatherer: applying constraint satisfaction, branch-and-bound and solution synthesis to natural language semantics. In: Technical report, MCCS-96-289. New Mexico State University: Computing Research Lab; 1996.

[22] McShane M, Beale S, Nirenburg S. Some meaning procedures of Ontological Semantics. In: Lino MT, Xavier MF, Ferreira F, Costa R, Silva R, (Eds.), Proceedings of the fourth international conference on language resources and evaluation (LREC-2004). Conference CD distributed by European Language Resources Association (ELRA), Paris, France. 2004.

[23] Allen J, Ferguson G, Blaylock N, Byron D, Chambers N, Dzikovska M, et al. Chester: towards a personal medication advisor. Journal of Biomedical Informatics 2006;39:500–13.

[24] Allen J, Chambers N, Ferguson G, Galescu L, Jung H, Swift M, et al. PLOW: a collaborative task learning agent. In: Proceedings of the 22nd national conference on artificial intelligence (AAAI'07), Vancouver, BC. Menlo Park, CA: AAAI Press; 2007. p. 1514–9.

[25] Ferguson G, Allen J. TRIPS: an integrated intelligent problem-solving assistant. In: Proceedings of the 15th national conference on artificial intelligence (AAAI'98), Madison, Wisconsin. Menlo Park, CA: AAAI Press; 1998. p. 567–73.

[26] Lemon O, Liu X. DUDE: a dialogue and understanding development environment, mapping business process models to information state update dialogue systems. In: Proceedings of the 11th conference of the European chapter of the Association for Computational Linguistics (EACL-2006). 2006. p. 99–102.

[27] Anderson ML, Josyula D, Perlis D. Talking to computers. In: Tecuci C, Aha DW, Boicu M, Cox MT, Ferguson G, Tate A, editors. Proceedings of the workshop on mixed initiative intelligent systems, IJCAI 2003. 2003. p. 1–8.

[28] Ellaway R. The realities of the virtual patient. Proceedings of the international conference on virtual patients. Bio-Algorithms and Med-Systems 2009;vol. 5(9):8.

[29] Zary N. Virtual patients for education, assessment and research: a web-based approach. PhD Dissertation, Karolinska Institute, Stockholm, Sweden; 2007.

[30] Courteille O, Bergin R, Stockeld D, Ponzer S, Fors U. The use of a virtual patient case in an OSCE-based exam – a pilot study. Medical Teacher 2008:pe66–76.

[31] Chesher D. Exploring the use of a web-based virtual patient to support learning. PhD Thesis. The University of Sydney, 2005.

[32] Stead WW, Lin HS, editors. Computational Technology for Effective Health Care: immediate Steps and Strategic Directions. Washington, DC: The National Academies Press; 2009.

[33] Berkman ND, DeWalt DA, Pignone MP, Sheridan SL, Lohr KN, Lux L, et al. Literacy and health outcomes. In: Evidence reports/technology assessments, no. 87. Agency for Healthcare Research and Quality, Rockville, MD, Report # 04-E007-2; January. 2004.

[34] D'Avanzo C. Mosby's pocket guide to cultural health assessment. Mosby Inc.; 2007.

[35] Stewart M, Brown JB, Weston WW, McWhinney IR, McWilliam CL, Freeman TR. Patient-centered medicine: transforming the clinical method. Milton Keynes: Radcliffe Medical Press; 2006.

[36] McShane M, Nirenburg S, Beale S. Semantics-based resolution of fragments and underspecified structures. Traitement Automatique des Langues 2005;46(1):163–84.

[37] Artstein R, Poesio M. Inter-coder agreement for computational linguistics. Computational Linguistics 2008;34(4):555–96 [survey article].