Mind-, Body- and Emotion-Reading

Marjorie McShane[1], Sergei Nirenburg[1], Stephen Beale[1],
Bruce Jarrell[2], George Fantry[2] and David Mallott[2]
[1] University of Maryland Baltimore County; Baltimore, Maryland, USA
marge@umbc.edu, sergei@umbc.edu, sbeale@umbc.edu
[2] University of Maryland, Baltimore; Baltimore, Maryland, USA
bjarrell@umaryland.edu, gfantry@medicine.umaryland.edu,
dmallott@meded.umaryland.edu

Abstract

A core requirement of human-like intelligent agents is the ability to create and dynamically maintain mental models of other agents. We suggest that such modeling capabilities must extend beyond traditional notions of "mindreading" to include inferring unobservable features of other agents' bodies and emotions as well. We describe the modeling of agents endowed with a mind, body and emotions within the OntoAgent knowledge environment and discuss how the same ontologically-grounded, mentalist models that support the *simulation* of agents also permit those agents to carry out *mind-, body-* and *emotion-reading* of their human and artificial collaborators.

1 Introduction

Intelligent agents that can function like humans could offer great benefits to society, including in the domain of clinical medicine, where they could act as virtual patients, tutors and advisors (Stead and Lin 2009). However, to be truly useful in these roles, intelligent agents must reflect the diversity of human beings (Berkman et al. 2004; D'Avanzo 2007) and be able to function robustly in the complex doctor-patient relationship (Dorr Goold and Lipkin, Jr. 1999), as emphasized by the currently ascendant paradigm of patient-centered medicine (Stewart et al. 2006).

The long-term objective of the OntoAgent research program is to model societies of intelligent agents that are sufficiently sophisticated and multi-faceted to fulfill such roles. Ontologically-grounded, mentalist modeling strategies are employed to configure agents with simulated minds, bodies, and emotional profiles (Nirenburg et al. 2008).

- The **mind** is modeled with the functionalities traditionally posited for cognitive architectures, including perception, situation assessment, decision making, problem solving, communication, learning and action (Langley et al. 2009). OntoAgent emphasizes deep-semantic language understanding – provided by the OntoSem module of OntoAgent (Nirenburg and Raskin 2004) – as an input channel of perception. Results of language processing are stored in memory and serve as heuristic input to subsequent agent reasoning and action.

- The **body** is a non-robotic physiological simulation, modeled as a set of ontological parameters whose values change over time in response to external and internal stimuli (McShane et al. 2007). Physiological modeling involves a combination of causal chains and what we call "clinical knowledge bridges." Causal-chain modeling provides the highest degree of autonomous functioning of agents but is not always feasible, due to lacunae in medical understanding, and is not always necessary, since many known causal chains would describe functionalities that have no visible manifestation or utility in OntoAgent's current application domains. Knowledge bridges encapsulate population-level generalizations and reflect the mental models of expert clinicians. The interpretation of bodily signals by the mind is modeled by a simulated process of interoception.

- The **emotional** layer of OntoAgents is both affected by and affects the models of the mind and the body. Emotions are generated by functions that are similar in format to the decision functions that guide conscious agent behavior (Nirenburg et al. 2008). Specifically,
    - their input is a set of ontological property values derived within perception and reasoning modules;
    - their output is an assignment of values to specific features (such TYPE, INTENSITY, TIME, EXPERIENCER and CAUSED-BY) of instances of emotion; and
    - their composition is determined through an iterative process of introspection and testing (cf. a similar function generation strategy in PEACTIDIM; Gratch et al. 2009).

  The OntoAgent model of emotion – while inspired by human studies and theoretical models – does not commit to any single, previously propounded theory or attempt to resolve centuries' old debates spanning the fields of philosophy, psychology, cognitive science and evolutionary biology (Fellous and Arbib 2005). There are several reasons for this. First, people have limited ability to understand their own experiencing of emotion, to analyze with certainty their observations of affect in others, or to unambiguously verbalize either of these, which makes the modeling of emotion/affect an inherently imprecise enterprise (de Sousa 2013; Plutchik 2001). Second, practically all theories stop short of providing actual mappings from the theory to behavioral data, with the cavernous gap between theory and practice making the implementation of emotion/affect in agents still largely virgin territory (Gratch et al. 2009; Hudlicka 2011). Third, *our* core contribution to emotion/affect modeling involves incorporating a *functionally adequate* model into a richly multi-faceted agent. This pursuit of functional verisimilitude permits us to remain agnostic regarding some much-debated theoretical tenets, such as the role of cognitive appraisal in emotion generation: OntoAgents can experience emotions either with or without cognitive appraisal, as is deemed most appropriate for the given emotion and the contextual triggers for it.

OntoAgents are used in prototype applications that include both humans and other intelligent agents. The interplay of mind, body and emotion modeling in Ontogent is presented in this paper through an extended example situated in the Maryland Virtual Patient (MVP) application (McShane et al. 2007). MVP is a system for training clinicians to diagnose and treat diseases with the help of OntoAgents. MVP allows trainees to manage, over simulated time, a population of virtual patients (VPs). The trainee-VP interaction occurs through text-based natural language dialog. The trainee can interview, teach, negotiate with, diagnose, treat, and follow up with VPs over any number of simulated office visits. VPs, for their part, can learn, ask and answer questions, make decisions, and function in a realistic physiological manner in response to the specific medical management agreed upon by the VP and the trainee. A tutoring agent observes the trainee-VP interaction and provides contextual advice.

This paper reports two innovations to the well-established OntoAgent research paradigm. The first is the incorporation of emotional state modeling into OntoAgent. Although this enhancement would increase the verisimilitude of agents operating in any domain, it is particularly well motivated in the realm of clinical medicine due to the direct and indirect effects of emotional states on disease processes. The second innovation is the integration into a single theory and functional approach what we call "mind-, body- and emotion-reading", which follows the principle, but expands the purview, of mental-state ascription, or “mindreading” (i.e., hypothesizing about the beliefs, plans and goals of others; see, e.g., Stich and Nichols 2003; Bello 2012). We suggest that developing agent models that tightly integrate bodies, minds and emotions not only permits the modeled agents to behave more like people, it also provides the knowledge substrate to allow them to hypothesize about these states in others – i.e., carry out mind-, body- and emotion-reading. Applying the same knowledge substrate and modeling strategies across functionalities, agents and domains is central to the feasibility of the OntoAgent paradigm and, we believe, satisfactorily addresses the issue of scalability of this knowledge-based paradigm.

2 Example and Architecture

The exposition is structured as a walk through a simulation run in the MVP application, with Sherry serving as our VP and DOC serving as the trainee. Like all VP instances, Sherry is configured when a developer, teacher, or medical expert selects particular values for relevant parameters that comprise the ontological model of a patient (Jarrell et al. 2008).

The modeling strategies supporting the simulation are described using Figure 1, which is the situated architecture of a virtual patient (like Sherry) in the MVP application.* The modules inside the large dark rectangle are not directly observable to DOC, whereas the outputs of the modules in the four boxes to the left are. The modules surrounded by dotted lines await implementation, whereas the rest have either been implemented or are well underway. Mind-, body- and emotion-reading occurs when a live clinician, virtual clinician (albeit not in the MVP application as currently configured) or virtual tutor develops a mental model of unobservable aspects of the VP. Aspects of mind- and emotion-reading are carried out by the VP with respect to DOC as well; this is not explicitly reflected in the figure but, rather, is folded into **Language understanding and associated learning**, which is the main input channel from DOC to VP. (Architecture modules appear in boldface throughout the running text.)

Sherry’s physiological abnormality is a mildly hypotensive lower esophageal sphincter (the muscle at the junction of the esophagus and the stomach), which is insufficiently hypotensive to lead to a disease state given optimal life circumstances, but will lead to gastroesophageal reflux disease (GERD) given suboptimal emotional states or lifestyle habits (see McShane et al. 2007 for our simulation-supporting model of GERD).

At the start of the simulation, Sherry does not have GERD but she gets a new job that results in her experiencing a high level of stress. Since Sherry’s personality traits include a strong aversion to stress, she posits the goal of stress relief. Among her ontologically recorded plans to relieve stress are (a) indulging in comforting activities like drinking coffee, consuming alcohol, and overeating (b) and taking psychotropic drugs – namely, tricyclic antidepressants. Since Sherry sees no reason *not* to engage in these, she pursues them all and is rewarded with a reduction in perceived stress. What she does not know is that all of these promote the development of GERD in individuals with a physiological predisposition to GERD, the causal chains for which are encoded in the physiological model (**Physiology).** Specifically, taking tricyclic antidepressants [arrow 1], experiencing vital exhaustion [arrow 2] and eating/drinking in certain inappropriate ways [arrow 3] lead to hypotension and/or excessive transient relaxations of the lower esophageal sphincter, which result in the disease processes of GERD.

* We describe the architecture as “situated” because (a) certain types of actions have priority over others (e.g., medical actions are presented explicitly but professional ones are not) and (b) the collaborating agent is expected to be a clinician (if it

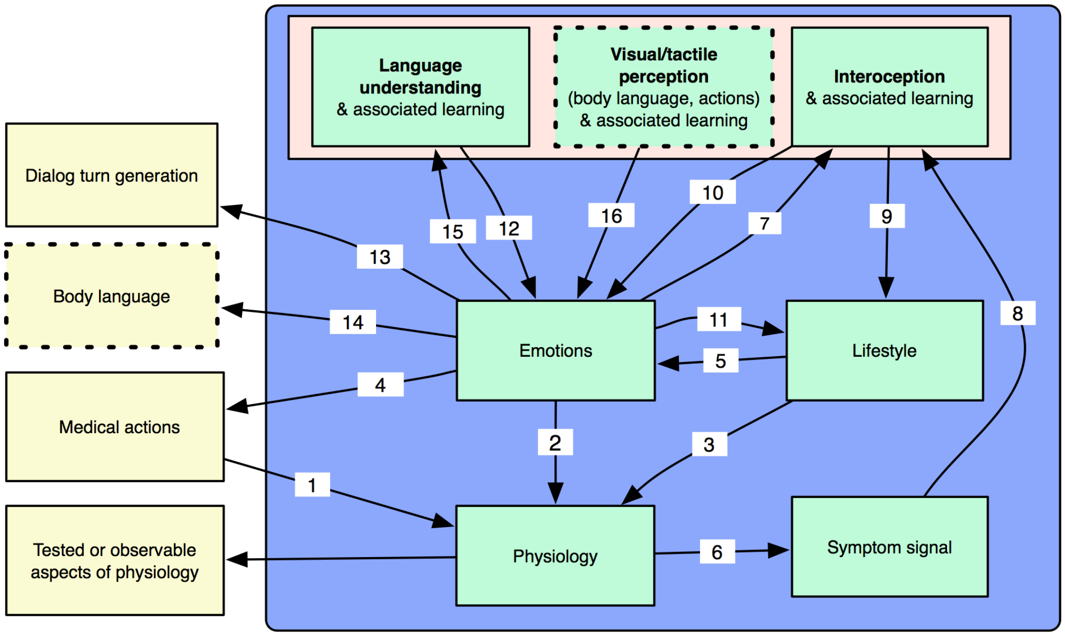

Figure 1. The model of an agent functioning as a virtual patient interacting with a clinician. The modules within the large dark rectangle are not transparent to the clinician: their content and interaction must be inferred via mind-, body- and emotion-reading. The outputs of the four modules to the left, by contrast, are directly observable to the clinician. The arrows represent data flow between agent modules, interpreted as channels for passing parameter values. Interactions between modules are marked by arrows and are described further in the text.

Note that lifestyle also indirectly affects **Physiology** [arrow 5], since certain choices (e.g., trying but failing to stop drinking coffee) induce stress, thus impacting **Emotions** (*I'm so mad at myself for not being able to kick my coffee habit!*), which in turn impacts **Physiology.**

As GERD progresses, at some point it causes sufficient inflammation to the lining of the esophagus to give rise to the symptom of heartburn. That is, the physiological simulation (**Physiology**) generates a **Symptom signal**, which represents the type, intensity and frequency of a symptom given two **Emotion**-related default conditions: (a) that Sherry is attentive – i.e., not distracted, in which case even a strong symptom signal might go unnoticed; and (b) that she is calm, since it has been experimentally established (Mizyed et al. 2009) that acute stress, prolonged stress or a history of trauma can heighten symptom perception. Whatever the values of the properties contributed by **Emotions,** they combine [arrow 7] with the **Symptom signal** [arrow 8] to generate the patient's actual experiencing of the symptom via **Interoception.**

Sherry's being bothered by her heartburn could lead [arrow 9] to **Lifestyle** decisions (*Let me try to reduce my caffeine habit, which seems to aggravate my heartburn*) that would positively affect [arrow 3] **Physiology**. However, that is not necessarily the case: Interoception can also negatively affect [arrow 10] **Emotions** (*I'm so tired of this heartburn!*) which can directly impact **Physiology** [arrow 2] or lead to **Lifestyle** decisions (*I'll have a little coffee to improve my mood*) that aggravate **Physiology**.

Even though Sherry realizes – thanks to her particular world knowledge and reasoning abilities – that certain lifestyle choices seem to aggravate her heartburn, she, in fact, cannot make herself change her lifestyle and decides to see the doctor instead. The mechanisms for this, and other, decision-making are described in Nirenburg et al. 2008.

When Sherry shows up at DOC's office, DOC (the medical trainee) begins to interview her by asking *What brings you here?* Sherry's **Language Understanding** module contextually interprets this input using the battery of knowledge resources and processors of the OntoSem text analysis system (Nirenburg and Raskin 2004). Her ontologically grounded, semantic interpretation of the input's meaning prepares her to answer the question [arrow 13, via arrow 12], and her stylistic interpretation permits one aspect of mindreading of the doctor: he is informal.[†] DOC continues to ask questions, which Sherry answers. With each question or explanation uttered by DOC, Sherry learns features of interest to her (for OntoAgent learning, see Nirenburg et al. 2010).

We do not explicitly posit DOC's mind-, body-, and emotion-reading of Sherry since DOC, in the example described here, is a live person, not an OntoAgent.[‡] We can, however, posit the corresponding actions of the virtual tutor, TUTOR, who constructs dynamic mental models of both Sherry and DOC throughout the simulation. TUTOR is endowed with the full complement of medical, world and interpersonal knowledge that underlies agent simulation, making it an expert in all relevant domains. This does not mean, however, that it is omniscient: it learns about Sherry using the same channels available to DOC, and it learns about DOC using the same channels available to Sherry (Nirenburg and McShane 2012). However, unlike Sherry, TUTOR can evaluate DOC's knowledge of clinical medicine by comparing DOC's actions with TUTOR's understanding of best practices. For example, DOC might not be aware of the emotional component in the diagnosis and treatment of GERD and, if he fails to cover related issues in the patient interview, TUTOR can point out the oversight and suggest associated investigation of these aspects of the expert model.

Returning to our simulation run, during the patient interview DOC asks Sherry if she consumes caffeine, chocolate or mints, all of which are known to promote GERD in some patients. She responds 'yes' to caffeine, reporting high coffee consumption. He asks if she eats big meals before bed ('yes'), and if she is taking any medication ('tricyclic antidepressants'). He says that she most likely has GERD, describes the disease to her (**Language understanding and associated learning)**, and says she'll get the best symptom relief by taking proton-pump inhibitors twice daily. When he says the latter, Sherry gets nervous [arrow 12], goes pale [arrow 14], and responds [arrow 13] that she doesn't want to take any more medication [arrow 4]. This last reaction and corresponding decision result from her remembered experiences of side-effects from medications, and her resulting aversion to medication use overall. Since Sherry's personality traits include boldness and the desire to understand the whole decision space before making decisions, she asks if there are any other options [arrow 13]. DOC explains that if she can make significant lifestyle changes – stop consuming coffee, alcohol and large meals before bed – that might be sufficient to reverse the disease course. Sherry agrees to give that a try [arrow 4].

Sherry leaves the virtual office and accelerated simulated time resumes. Sherry's personality traits (e.g., low self-control) and life circumstances (high stress, getting home from work late) remain the same, and the impact of the doctor's advice – interpreted as a plan to achieve the goal of symptom relief – is not sufficient to overturn her **Lifestyle** choices (coffee, alcohol, large evening meals). Her disease processes (**Physiology**) continue unabated, and her symptom perception increases [arrow 7] due to an increased level of anxiety (**Emotion**), generated by her failure to follow the doctor's advice.

As a result, a month later Sherry decides to see DOC again. He asks how her symptoms are and whether she has been able to modify her lifestyle. The latter question makes Sherry uncomfortable [arrow 12] because both options for answering are suboptimal. Telling the truth conflicts with the goal of saving face (recall that her personality traits make her susceptible to embarrassment), whereas lying conflicts with the goal of best collaborating with her doctor to get relief of her symptoms (for the

[†] More specifically, Sherry's decision function that permits her to assign a value to the property FORMALITY as attributed to DOC takes as input information including what he says (e.g., Does he make jokes?) and how he says it (e.g., Does he use colloquial idioms? informal words?).

[‡] Actually, we are implementing a demonstration system in which both DOC and the VP are virtual agents. This will be discussed in the conference talk.

modeling of lying in OntoAgent, as well as mindreading in service of lie detection, see McShane et al. 2012a.) The decision whether or not to tell the truth involves many input parameters, including the results of Sherry's mindreading of the doctor. From her first appointment, Sherry has a positive overall impression of DOC, generalized from limited heuristic evidence via the halo effect (for bias modeling in OntoAgent, see McShane et al. 2012b); this votes against lying. Another vote against lying is an **Observable aspect of physiology** that she believes DOC will notice: she has gained 8 pounds in the past month, which suggests bad eating habits. Deciding that lying is not in her best interests, Sherry admits that she has not changed his lifestyle and her symptoms have gotten worse.

DOC says he understands that lifestyle changes are very hard, especially when someone is under stress. DOC's sympathetic approach enhances Sherry's positive-halo impression of DOC (McShane et al. 2012b). DOC again suggests medication but Sherry's attitude toward medication is the same as before. Her character traits (high assertiveness, the desire to understand the full decision space before making a decision) lead her to ask for other options. DOC says there are no other options and says that if left untreated, GERD can, over time, lead to esophageal cancer. Hearing that (and, if available, seeing the doctor's suddenly serious facial expression [arrow 1]), Sherry's anxiety level skyrockets [arrow 12] and her ability to understand language input and remember it decreases [arrow 15]. As such, she doesn't pay attention to (i.e., fully semantically analyze or remember) the doctor's next utterance, which is, "Cancer is really quite rare, though, so the main reason to take medication is to help your symptoms." The cancer prospect alone (realized as high fear of death) is sufficient to change Sherry's decision about taking medication [arrow 4]. She tells the doctor she will take the medication, he prescribes it, and this appointment concludes.

In an actual system run, Sherry's simulated life would continue. Based on her character traits, she would or wouldn't take the medication as prescribed, which would affect her physiology and associated symptoms in the ways generated by the model. She would keep regularly assessing whether or not to make further appointments with DOC.

3 Discussion

This very abbreviated tour through a system run of MVP was not intended to adequately detail *how* OntoAgents are modeled – that will be presented in a full-length article in preparation, which will report on the ongoing implementation of all of the aspects of functioning described above. This paper, by contrast, was intended to give insight into the motivations for pursuing what might be considered by some to be an unusually broad program of research and development.

The domain of clinical medicine offers a rich space of opportunities for agent modeling because all aspects of being human (mind, body and emotions) are simultaneously relevant, as is social interaction with other human and artificial agents who are similarly endowed. The requirements of social cognition in such a domain extend to inferring unobservables relating to all aspects of multi-dimensional agents – i.e., mind-, body- and emotion-reading.

As is canonically true of all work in a knowledge-based paradigm, expanding coverage of knowledge about the world, language, decision-making, etc., becomes feasible only once the principles and methods have been well developed and demonstrated.

References

Bello, P. (2012) Cognitive foundations for a computational theory of mindreading. *Advances in Cognitive Systems* 1: 59-72.

Berkman, N. D., DeWalt, D. A., Pignone, M. P., Sheridan, S. L., Lohr, K. N., Lux, L., et al. (2004) *Literacy and Health Outcomes: Evidence Reports/Technology Assessments*, No. 87. Agency for Healthcare Research and Quality, Rockville, MD, Report # 04-E007-2, January 2004.

D'Avanzo, C. (2007) *Mosby's Pocket Guide to Cultural Health Assessment*. Mosby Inc.

de Sousa, R. (2013) "Emotion". *The Stanford Encyclopedia of Philosophy* (Spring 2013 Edition), Zalta, E. N. (ed.), URL = <http://plato.stanford.edu/archives/spr2013/entries/emotion/>

Dorr Goold, S., Lipkin, Jr., M. (1999) The doctor–patient relationship: Challenges, opportunities, and strategies. *J Gen Intern Med*. 14(S1): S26–S33.

Gratch, R. P., Marinier, III, R. P., Laird, J. E., Lewis, R. L. (2009) A computational unification of cognitive behavior and emotion. *Cognitive Systems Research* 10(1): 48-69.

Fellous, J., Arbib, M. A., eds. (2005) *Who Needs Emotions? The brain meets the robot.* Oxford University Press.

Hudlicka, E. (2011) Guidelines for designing computational models of emotions. *International Journal of Synthetic Emotions* 2(1).

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

Langley, P., Laird, J. E., Rogers, S. (2009) Cognitive architectures: Research issues and challenges. *Cognitive Systems Research*, 10: 141-160.

McShane, M., Fantry, G., Beale, S., Nirenburg, S., Jarrell, B. (2007) Disease interaction in cognitive simulations for medical training. *Proceedings of MODSIM World Conference, Medical Track*, Virginia Beach, Sept. 11-13 2007.

McShane, M., Beale, S., Nirenburg, S., Jarrell, B., Fantry, G. (2012a). Inconsistency as diagnostic tool in a society of intelligent agents. *Artificial Intelligence in Medicine* (AIIM), 55(3): 137-48.

McShane, M., Nirenburg, S., Jarrell, B. (2012b) Modeling decision-making biases. *Biologically-Inspired Cognitive Architectures (BICA) Journal*. Available online Nov. 3, 2012.

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

Nirenburg, S., McShane, M. (2012) Agents modeling agents: Incorporating ethics-related reasoning. *Proceedings of the Symposium on Moral Cognition and Theory of Mind*, Birmingham, UK, July.

Nirenburg, S., McShane, M., Beale, S. (2008) A simulated physiological/cognitive "Double Agent". Beal, J., Bello, P., Cassimatis, N., Coen, M., Winston, P. (Eds.), *Papers from the AAAI fall symposium, Naturally Inspired Cognitive Architectures*. AAAI technical report FS-08-06, Menlo Park, CA: AAAI Press.

Nirenburg, S., McShane, M., Beale, S., English, J., Catizone, R. (2010) Four kinds of learning in one agent-oriented environment. *Proceedings of the First International Conference on Biologically Inspired Cognitive Architectures* (BICA), Arlington, VA, Nov. 13-14.

Nirenburg, S., Raskin, V. (2004) *Ontological Semantics*. Cambridge, Mass.: The MIT Press.

Plutchik, R. (2001) The nature of emotions. *American Scientist,* 89(4): 344 - 350.

Stead, W. W., Lin, H. S. (Eds.) (2009) *Computational Technology for Effective Health Care: Immediate Steps and Strategic Directions*. Washington, D.C.: The National Academies Press.

Stewart, M., Brown, J. B., Weston, W. W., McWhinney, I. R., McWilliam, C. L., Freeman, T. R. (2006) *Patient-Centered Medicine: Transforming the clinical method*. Milton Keynes: Radcliffe Medical Press. 2006.

Stich, S., Nichols, S. (2003) Folk psychology. Stich, S., Warfield, T. (Eds.), *The Blackwell Guide to the Philosophy of Mind*. Oxford, UK: Blackwell, pp. 235-255.