Health Behavior Coaching: A Motivating Domain for Human-Aware Artificial Intelligence Research

Shiwali Mohan, Anusha Venkatakrishnan, Danny Bobrow, Peter Pirolli
Interaction Analytics Lab
Palo Alto Research Center
3333 Coyote Hill Road, Palo Alto, California 94306

Abstract

Human-agent interaction has been studied for a while in AI research. However, very little attention has been paid to developing methods that are *human-aware* - that can model, reason about, and make decisions based on changes in human physiological, cognitive, and affective states. We argue that health behavior coaching is a domain that requires an agent to be *human-aware*, in order to be effective and therefore, is a fruitful domain to pursue. We have delineated a set of desiderata for agents that can coach people to develop healthy behaviors. This set results from our ongoing work on developing AFYA (Mohan et al. 2017) - an interactive coach that resides in a smartphone app and coaches people exercise more. This paper describes each desiderata, the challenge it poses to AI research, and provides examples from our work demonstrating how it can be met.

Introduction

As AI algorithms and technologies become reliable and robust, they are being employed in a variety of domains to solve complex problems in transportation, health, education, banking etc. Workflows in these domains have started to incorporate intelligent mechanisms to solve parts of the problems. Not surprisingly, the questions related to relationships between intelligent agents or systems and humans have taken center stage in AI research. Several communities in AI - human-robot interaction, conversational agents, personal assistants - have addressed some of these questions. The human-robot interaction community (Leite et al. 2013) has studied embodiment of physical agents in human spaces, impact of scripted verbal and non-verbal behaviors, understanding as well as producing gestures and emotions, and performing tasks in human environments. Research on design and analysis of conversational agents (Rossen and Lok 2012) has focused methods for robust verbal conversations for a variety of domains. The conversational content is usually fit to a specific domain by experts and the agent is designed to deliver the content in an interactive fashion. Similarly, the design of personal assistants (Myers et al. 2007) has focused on representations, planning, and learning of tasks in various domains. While these research directions have been fruitful in the development of interactive agents, they crucially miss a challenge that is critical for robust human-aware AI - modeling and reasoning about human decision making and behavior.

We argue that health-behavior coaching - helping people to develop helpful health-related behaviors and to curtail harmful ones - is a challenging as well as fruitful domain to conduct human-aware AI research. This domain puts modeling and reasoning about a human at the core of AI design. The domain requires that a health coach understand the cognitive, emotional, physical, situational, and other aspects of a trainee's health behaviors. An effective coach is knowledgeable about a wide range of behavioral interventions and how they influence a trainee. The interventions may provide informational support, encourage practice of helpful behaviors in different contexts, help with remembering behaviors when the right context arises etc. To be impactful, these interventions are personalized to each individual trainee and gradually adapted for their specific circumstances. The interventions are delivered through mixed-initiative communication and other forms of interaction. Solutions for this domain necessitate developing computational approaches for the multi-dimensional representation of a trainee's dynamically changing state. Additionally, the solutions must address how the representation integrates with long-term planning for trainee-coach joint tasks, communicating with the trainee, and incremental learning models that explain their behavior. Developing a health coach brings together research in knowledge representation and reasoning, computational cognitive models, model-based reasoning, and online learning.

In this paper, we outline the representational and behavioral desiderata for a health coaching agent and discuss challenges they present for AI research. We describe the progress we have made in designing AFYA (Mohan et al. 2017) - an interactive coach to help sedentary individuals develop regular aerobic exercise behaviors and evaluate it along the desiderata proposed.

Health Behavior Coaching

Health behaviors account for an estimated of 60% of the risks associated with chronic illnesses such as diabetes and cardiovascular disease. As the at-risk population grows around the world, the challenge of developing and disseminating effective methods for improving health behaviors is becoming critically important. A wide range of strategies (or *interventions*) have been investigated for promoting better health behaviors (Kahn et al.) including informational approaches to change knowledge and attitudes toward health, behavioral counseling approaches to teach and maintain health-related skills, creating social environments to support health behavior change, and environmental and policy changes to provide resources that enable good health behaviors. Of these, individual counseling received in personal meetings (Karen B. Eden et al. 2002) or over telephone (Eakin et al. 2007) have been shown to be very effective in promoting behavior change. Although useful, such counseling is very resource-intensive, in terms of both training necessary personnel and delivering to a large population.

To make such counseling pervasive and cost-effective, we are motivated to develop an interactive, intelligent agent that can reside on a mobile device and provide counseling in a manner similar to a human coach. Health-behavior change literature describes several interventions based on constructs such as goal setting (Kay Shilts et al. 2004), boosting self-efficacy (Stacey et al. 2015), reward shaping and incentive design (West et al. 2010), implementation intentions (Gollwitzer and Sheeran 2006; Bélanger-Gravel, Godin, and Amireault 2013), reminding (Fry and Neff 2009), that produce large effect sizes in positively influencing health behaviors. These interventions are usually delivered by expert humans through a prolonged interaction with their trainees.

While this prior research on health behavior change has been crucial in identifying practical methods that can produce health behavior change, the theories developed therein are inadequate to support the development of computational delivery methods (Riley et al. 2011). Further, the proposed theories model human health behavior as a function of constructs such as *motivation*, *attitude* as opposed to a product of a dynamic cognitive system that is influenced by physiological, affective, environmental, social, and experiential states (Spruijt-Metz et al. 2015). Consequently, there is a need for computational methods that can not only model and predict the changes in the human physiological and cognitive system, but also for methods that can coach this human system toward a beneficial goal.

Desiderata

To design an intelligent coach for affecting health behavior change, we must leverage prior design and analyses of intelligent agents by prior AI research. In the sections below, we describe the representational and behavioral desiderata of a coach. These are obtained by aligning aspects of this domain to how typical intelligent agents are designed. These desiderata motivate a human-aware design in which the agent's sensing, manipulation, and decision making are targeted toward affecting a human system. To ground this discussion in a concrete domain, we focus on healthy behaviors for sedentary, overweight individuals. These include exercising regularly (30 minutes for 5 days a week) incorporating a mix of aerobic and endurance exercises as well as eating well. Our ongoing work on developing AFYA aims at helping sedentary, overweight individuals develop these behaviors. In particular, we

Representational Desiderata

To develop a health coach, a critical challenge that must be addressed is that of representation. A typical AI agent encodes its sensory perceptions into a state representation of its environment. It also represents the actions it can take to manipulate its sensors, effectors, and the environment in pursuit of a goal. The decision making involves identifying an action or a sequence of action that will lead to perceptible progress toward the goal.

Trainee state In this domain, the coach's *environment* is the human system that it is trying to affect. The agent must represent the environment capturing information that is necessary to not only formulate a beneficial goal state but also to decide over its action space (described later). A crucial challenge is to identify what this information is as prior research in AI does not characterize representation of a human system. The interventions investigated in the health behavior literature provides some guidance toward how a useful state representation can be formulated. For a wide variety of interventions, representing two aspects of behavior are critical: the physical capability (e.g., *aerobic capacity*, *number of steps a person can take* etc.) and the cognitive behavior determinants (e.g., *beliefs about success at a task*, *beliefs about value of a task*). Additional descriptions about a trainee's situation (*access to fresh vegetables*), concepts known to them (*proteins of vegetable origin*), and their preferences (*dislikes swimming*) may also be useful in reasoning about interventions. In addition to identifying and describing multiple aspects about a trainee, another challenge is identifying the right level of abstraction. For example, as a trainee starts exercising regularly their physical capability changes due to changes in muscle physiology. Do we need to model individual muscles to represent a trainee's capability? Finally, the agent must describe and identify desirable states of the human system.

Our work (Mohan et al. 2017) studies how AFYA can prescribe aerobic exercises such that it is sensitive to *personal* differences and *temporal* changes in physiology. As AFYA's goal is to promote aerobic activities, it only represents its trainee's physical, aerobic capability. AFYA represents a trainee's aerobic capability as the volume (intensity $\times$ duration $\times$ frequency) the trainee can successfully undertake in a week. This is inspired by how experts (physical therapists, exercise coaches) reason about prescribing exercises to their trainees. This is a high-level representation when compared to representing muscle physiology. We show that this abstraction is sufficient to produce recommendations that are judged favorably by experts. This representation also allows for formulating a desirable state that the coach must nudge its trainees toward. We have adopted the American Heart Association (AHA) recommendation that adults must exercise at moderate intensity for at least 150 minutes in a week. In the future, we will expand AFYA's representation to include the trainee's cognitive and affective states.

Interventional actions Behavior change interventions describe several interactive actions that have various purposes. For example, motivational interviewing interventions employ conversational acts that ask a person to elaborate on their problems, summarize their problems etc. Similarly, goal setting interventions employ interactive actions aimed at describing a behavioral task to the trainee (*Walk briskly 15 minutes this evening*) or providing instructions about how to perform it (*walk at a speed that tires you, but you still can talk*). Actions such as these influence the state of the trainee. Some actions such as instructions may change their knowledge state while others may trigger retrospection. Some may result in the trainee changing their environment (*place a reminder about eating healthy on your fridge door*) others are aimed at gathering more information about the trainee. A critical challenge that must be addressed is to characterize the variety of actions described in the literature by which aspect of the trainee's state they modulate.

AFYA implements a goal-setting intervention (Kay Shilts et al. 2004). Here, the coach sets relevant and appropriate goals for their trainee. These goals are presented to the trainee and tracked through a smartphone application. AFYA provides for coach-trainee interaction on the phone to gather more information about the trainee's performance on these goals as well as to estimate their aerobic capability. This information is useful to generate a personalized schedule of exercises.

Trainee model

The coach must incorporate a model for the trainee it is working with. The trainee model plays two roles in the design of the coach. First, it drives expectations of the coach. Given a specific trainee state, the model predicts what state will result if the trainee follows the intervention. If this state is not observed, the coach can then begin to reason about its trainee state estimates and guide diagnostic behavior. The second role is during intervention planning. The model must be designed to compute the trainee states and their likelihood that would result from providing a specific intervention in a given trainee state. The model must also compute the likelihood of occurrence of the target behavior. These quantities are useful in evaluating the value of different intervention actions. The resultant state likelihood plays an important role in classical planning by defining a state transition function to help find a path (sequence of intervention actions to be taken) to achieve a desired state. The likelihood of behavior can be used to estimate the utility of taking an intervention action the coach can take for a trainee state. Prior research in AI has ignored development of such models for state changes in the human system.

Prior work in the cognitive modeling community, especially the interactive tutoring systems (Graesser et al. 2001) has proposed some relevant models. This research has looked at measuring and representing what is known by a human user (in domains such as algebra, programming) and adapting a lesson to maximize learning. This research exclusively focuses on building up cognitive skills and conceptual knowledge. This research makes a critical assumption that these tasks are performed in focused sessions and all cognitive resources employed exclusively. Consequently, the research focuses on building up cognitive skills and conceptual knowledge relevant to the domain. Health-related behaviors are significantly different in that they occur in ecological settings in which other behaviors and tasks may compete for cognitive resources. Beyond developing skills and knowledge required for health behaviors, they need to be made more accessible, salient, and valued as compared to other tasks a person might be involved in. This challenge requires us to adopt a more holistic view of behavior, going above and beyond *knowledge-level* models explored in interactive tutoring systems research.

In this context, AFYA employs a parameterized model of the trainee's aerobic capability that drives its expectation of the trainee's performance. It models a trainee's growth in aerobic capability every week as a staircase function of uniform step height. The height of the model at a week captures the aerobic capability in that week. The step height represents the expected growth in a trainee's physical capability. The step height is a model parameter that is updated based on how the trainee performs the exercise tasks prescribed in a week. For example, if they fail to exercise at the recommended level, the step height is reduced. Updating the parameters influences what aerobic capability the model predicts. As it is apparent, this is a partial model that only captures physiological changes in the human system. We are currently extending this model to predict changes in the cognitive and affective states.

Behavioral Desiderata

Now we identify behaviors that must be encoded in the intelligent coach such that it is able to coach people to develop healthier habits.

Communicative sensing Communication in coaching serves two functions - measuring and estimating current state of a person and delivering interventions through instructions. A coach may seek information about the perceived task difficulty, perceived success on task performance, or how much resources were expended in the task in addition to observing task performance directly (visual inspection) or indirectly (aiding memory recall). This information is then used to estimate the current state, update the causal model maintained by the coach, adapt the tasks to the right level of difficulty, explore variations of tasks etc.

Research along this direction must explore interactive strategies that work over a joint space of physiological sensing (with tools such as *FitBit*) and human-agent communication. While there are several sensors available to measure some aspect of the task (such as number of steps taken or heart rate), they might not be sufficient to capture all the information required by the coach to develop or adapt an interventional strategy. For example, heart rate sensing may provide information about how physically challenging a task was. However, future tasks prescribed by the coach need to be moderated by how difficult the trainee perceives the task to be and if they think they can achieve success at the given task. Coach-trainee communication may alleviate some of such sensing challenges.

To develop solutions for this challenge, sensing and communication must be employed together to generate estimates of a trainee's state. While physiological sensors can contribute to estimating the physical state of the trainee, communication must be leveraged to estimate the cognitive and situational state of the trainee. Uncertainties and ambiguities in estimating these states will be resolved by further communication. Over the lifetime of the coach, communication should become more directed as the coach begins to model the trainee accurately.

Presently, AFYA relies on trainee-coach mobile interaction to prescribe walking tasks. We developed interaction exchanges that leverage standard clinical practice of assessing a trainee's physical capability and measuring rate of perceived exertion while performing exercises. These interactions are then used to infer the trainee's state and to adapt the schedule of exercises to benefit the trainee the most. We show that without a sensor, communication with the trainee can be used for state estimation and adaptation.

Causal diagnosis There are several reasons for why an expected behavior does not occur. The trainee may not have sufficient skills or capabilities to execute the behavior, they may not believe that they can successfully perform the behavior (low *self-efficacy*), they may not remember the behavior at an opportune time, their environment may not afford the behavior etc. Each of these causes motivates a different interventional strategy. On observing a behavioral failure, the coach must generate hypotheses about the causes of failure to narrow down on the most likely cause. This may also involve further communication with the trainee.

The challenge here is to develop a formal framework for causal diagnosis of non-performance of a health behavior. An expectation-driven framework can be leveraged here. The coach explicitly represents what it expects the trainee will do given their current state and their trainee model. For example, the coach may believe that a person will be able to walk briskly for 15 minutes with 80% likelihood in the next week. This estimate may be based on two determinants of behavior: a) self-efficacy - the trainee believes that they can walk briskly for 30 minutes and their aerobic capability - given past exercise schedule and b) the trainee is physically capable of this task. In a case where the trainee is unable to fulfill the coach's expectation, the coach must reason about each of those determinants. If the person has been walking for the past month, it is likely that their self-efficacy is high. Therefore, the actual cause may be that they are incapable of sustaining this activity for 30 minutes. The coach, then, can gather more information about performance through human-coach interaction such as rate of perceived exertion to verify the cause.

As AFYA is designed only for prescribing aerobic exercises, it currently does not perform a causal diagnosis. In case that the trainee fails to pursue the exercise schedule prescribed by AFYA, it assumes that this occurs because the trainee isn't physically capable. This simplistic diagnosis is then followed by a revision in the trainee model and consequently in the prescription of the next exercise goal.

Intervention delivery Once a likely cause of failure is determined, the coach must deliver interventions that can resolve or alleviate the cause. An intervention is a sequence of interactive actions that influences the trainee's state. It may directly impact the cause or produce changes in trainee's behaviors that impact the cause. Once the representation desiderata is achieved, various AI approaches can be adapted to address this challenge. Planning algorithms, rule-driven reasoning, case-based approaches may find use in developing solutions for this problem. The environment the agent operates in is extremely noisy - estimating human state can be error prone due to noisy sensors and incomplete representations. Therefore, it crucial to develop approaches that are: a) *robust* - they produce reasonable behavior even when they have minimal information about the trainee and b) *adaptive* - their outputs can be influenced as the coach accumulates more information about the trainee.

To develop AFYA, we have adopted an incremental approach. For an intervention (e.g., goal setting), we consulted with experts to determine how they reason about prescribing aerobic exercises. Experts reason about the exercise's intensity (i), session length (d), and frequency in a week (f). They continually monitor how successful the trainee is in achieving the prescribing volume. They continually revise the volume as they gather more information about the trainee. We framed the goal setting problem as a planning problem in which the coach generates a schedule of exercises of increasing difficulty (computed based on i, d, f). The challenge in formulating goal setting as a planning problem is that the state transition function is undefined. Here, heuristics-based trainee model was developed to represent changes in the trainee's aerobic capability. This model predicted the state that will occur if the trainee successfully performs the assigned task. This prediction can be used to prescribe the right volume of exercise.

Experiential learning Any behavior change occurs over several months. If the interventions are successful, the state of the human system will undergo a change. Consequently, interventions must be adapted so that only relevant interventions are delivered to the trainee. A coach must continually learn about its trainee and adapt its behavior in response. Several AI learning algorithms can be explored to address this desiderata. However, the domain has some important characteristics that must be considered while adapting a learning algorithm. First, different trainees will have different needs and therefore, the coach must adapt its interventions to each specific trainee. Second, it is likely that the data available to learn from is very small - the coach may only get very few observations of its trainee. Additionally, the behavior of the coach must be reasonable to be perceived as trustworthy even in the very beginning. If it isn't, the trainee is unlikely to stick with the coach.

A model-based approach presents a good solution to this problem. A parameterized model is hand-designed based on observations of several human coach-trainee scenarios and in consultations with experts. The structure of the model encodes the determinants that influence human behavior and the parameters provide knobs that can be turned based on observations to fit the model to each individual trainee. A model-based intervention approach affords two kinds of learning: one where the model's parameters are updated in light of evidence observed about the trainee and second, through adapting the reasoning strategy such as by changing the planning heuristics. The challenge here is to explore the efficacy of these methods, individually and in combination, towards adapting interventions to each trainee.

AFYA is an example of personalization by updating the model. As described earlier, it models a person's growth in aerobic capability every week as a staircase function of uniform step height. The height of the model at a week captures the aerobic capability in that week. The step height represents the expected growth in a trainee's physical capability. The step height is a model parameter that is updated based on how the trainee performs the exercise tasks prescribed in a week. For example, if they fail to exercise at the recommended level, the step height is reduced. Updating the parameters influences what aerobic capability the model predicts. In this example, decreasing the height lowers the predicted aerobic capability. Consequently, during planning easier exercises are prescribed. We show that this type of adaptation is sufficient to produce plans that are judged favorably by experts.

Conclusion

AI research has begun to explore the role of humans in human-agent collaborative settings. While fruitful, prior work minimally studies the problem of reasoning about humans. In this paper, we argued that designing a coach for health behavior change is a motivating domain for human-aware AI research. This domain puts reasoning about a human system at the core of agent design. We delineated the representational and behavioral desiderata of an interactive intelligent coach and identified challenges for AI research. Additionally, we explained how our ongoing research in designing AFYA addresses these desiderata. Future research in this domain not only will encourage research on AI systems that are more *human-aware*, but also develop solutions for a problem that affects a large population.

References

Bélanger-Gravel, A.; Godin, G.; and Amireault, S. 2013. A meta-analytic review of the effect of implementation intentions on physical activity. *Health Psychology Review* 7(1):23–54.

Eakin, E. G.; Lawler, S. P.; Vandelanotte, C.; and Owen, N. 2007. Telephone Interventions for Physical Activity and Dietary Behavior Change A Systematic Review. *American Journal of Preventive Medicine* 32(5):419–434.

Fry, J. P., and Neff, R. A. 2009. Periodic Prompts and Reminders in Health Promotion and Health Behavior Interventions: Systematic Review. *Journal of Medical Internet Research* 11(2).

Gollwitzer, P. M., and Sheeran, P. 2006. Implementation Intentions and Goal Achievement: A Metaanalysis of Effects and Processes. *Advances in Experimental Social Psychology* 38:69–119.

Graesser, A. C.; VanLehn, K.; Rose, C. P.; Jordan, P. W.; and Harter, D. 2001. Intelligent Tutoring Systems with Conversational Dialogue. *AI Magazine* 22(4):39.

Kahn, E. B.; Ramsey, L. T.; Brownson, R. C.; Heath, G. W.; Howze, E. H.; Powell, K. E.; Stone, E. J.; Rajab, M. W.; and Corso, P. The Effectiveness of Interventions to Increase Physical Activity A Systematic Review and the Task Force on Community Preventive Services.

Karen B. Eden; C. Tracy Orleans; Cynthias Mulrow; Nola Pender; and Steven Teutsch. 2002. Does Counseling by Clinicians Improve Physical Activity? A Summary of the Evidence. *Annals of Internal Medicine* 137(3):208–215.

Kay Shilts, M.; Horowitz, M.; Townsend, M. S.; and Townsend, M. 2004. Goal Setting as a Strategy for Dietary and Physical Activity Behavior Change: A Review of the Literature. *American Journal of Health Promotion* 19(2).

Leite, I.; Martinho, C.; Paiva, A.; and Rv, S. U. 2013. Social Robots for Long-Term Interaction: A Survey. *International Journal of Social Robotics*.

Mohan, S.; Venkatakrishnan, A.; Silva, M.; and Pirolli, P. 2017. On Designing a Social Coach to Promote Regular Aerobic Exercise. In *Proceedings of the 29th Conference on Innovative Applications of Artificial Intelligence (to appear)*. AAAI.

Myers, K.; Berry, P.; Blythe, J.; and Conley, K. 2007. An intelligent personal assistant for task and time management. *AI Magazine* 28(2):47–62.

Riley, W. T.; Rivera, D. E.; Atienza, A. A.; Nilsen, W.; Allison, S. M.; and Mermelstein, R. 2011. Health behavior models in the age of mobile interventions: are our theories up to the task? *Translational Behavioral Medicine* 1(1):53–71.

Rossen, B., and Lok, B. 2012. Author's personal copy A crowdsourcing method to develop virtual human conversational agents. *Int. J. Human-Computer Studies* 70:301–319.

Spruijt-Metz, D.; Hekler, E.; Saranummi, N.; Intille, S.; Korhonen, I.; Nilsen, W.; Rivera, D. E.; Spring, B.; Michie, S.; Asch, D. A.; Sanna, A.; Salcedo, V. T.; Kukakfa, R.; and Pavel, M. 2015. Building new computational models to support health behavior change and maintenance: new opportunities in behavioral research. *Translational Behavioral Medicine* 5(3):335–346.

Stacey, F. G.; James, E. L.; Chapman, K.; Courneya, K. S.; and Lubans, D. R. 2015. A systematic review and meta-analysis of social cognitive theory-based physical activity and/or nutrition behavior change interventions for cancer survivors. *Journal of Cancer Survivorship* 9(2):305–338.

West, R.; Walia, A.; Hyder, N.; Shahab, L.; and Michie, S. 2010. Behavior change techniques used by the English Stop Smoking Services and their associations with short-term quit outcomes. *Nicotine & Tobacco Research* 12(7):742–747.