A Theory Belief Model for Cognitive Agents [DRAFT no cites]

Todd Shimoda
Colorado State University

Introduction

The Theory Belief Model (TBM) provides a framework for using, testing, and developing cognitive theory within the arena of human-computer interaction (HCI). In essence, it can be stated as "Cognitive agents (biological or artificial) perform cognitive acts using sets of beliefs that can be viewed as theory." Theories can explain and predict cognitive acts, allow selection of one cognitive act over another, and can be disconfirmed or revised. The main purpose for the TBM is to test and build theories using interactions with humans and software agents. The framework is intended to inform both cognitive theory and intelligent software design.

Cognitive agents and acts

Cognitive agents are entities that perform cognitive acts. A cognitive act consists of three general actions: perceiving information in the environment, reasoning about those perceptions using existing knowledge, and acting to make a reasoned change to the external or internal environment.

Cognitive acts under this definition occur as a loop of all three steps. In other words, if an action occurs after perception with no intermediate reasoning, it would not be considered a cognitive act. Rather, it would be a reflexive act; many agent actions can be classified as such. On the other hand, reasoning can be superficial and obvious, and take little time to perform, but is still considered to be reasoning and qualifies as part of a cognitive act.

Cognitive agents are either biological (people or non-human animals) or artificial (robots or software agents). The discussion of cognitive agents presented here is concerned with people and software agents. Software agents are computer programs using artificial intelligence to perform cognitive acts that emulate human cognitive behavior. Obviously, people and software agents have considerable differences:

Cognitive action	Human attributes	Software agent attributes
Perceiving	• Sensory organs and nerve pathways • Neural activation • Conscious awareness of perception	• Data input from hardware and software • Variable binding • No conscious awareness of data input
Reasoning	• Rational and irrational • Conscious and subconscious processes • Knowledge in neural structures and connections	• Rational and logical • Transparent algorithmic procedures • Knowledge in digital data structures
Acting	• Physical change in the natural environmental • Communication with other agents through natural language process • Knowledge construction or change through new or modified neural connections	• Virtual change in the software environment • Communication with other agents through algorithmic productions • Knowledge construction or change through updated data structures

The list of attributes is not complete, but gives an overview of several main attributes of cognitive agents. These attributes are complex, volumes have been written about each. Also, the descriptions of the attributes are sketchy and abstract; concrete examples of how these attributes manifest in specific applications are presented later. Finally, social factors and emotional attributes are not specifically mentioned but are *not* excluded in cognitive acts for people or software agents, indeed, they may be very important in many cognitive acts.

Some of the changes to an agent's environment due to a cognitive act may be internal, specifically, changing or constructing knowledge. There would be no immediate physical evidence of this action to other agents in the external environment (or to agents observing the external environment.) The change, of course, may affect an agent's future cognitive acts (a delayed cognitive act).

Theories in the Theory Belief Model

The theories used in the TBM are concerned with any phenomena related to one or more of the cognitive actions or attributes of cognitive agents. For example, the Theory of Reasoned Action qualifies, the Theory of General Relativity does not. Cognitive agents may be doing something with the Theory of Relativity, such as learning about it, but that theory does not explain how they are learning about it.

Theories allow cognitive agents to act in a coherent manner, at least when compared with acting randomly or haphazardly. Theories are generally thought of as formal or scientific theories; however, informal theories are included in the TBM. Formal theories are well-formed and articulated sets of beliefs largely developed through scientific method or other disciplined procedures. Informal theories are also sets of beliefs, which can be well-formed and articulated, although most tend to be less well-formed and less-precisely articulated.

Informal theories generally develop through an agent's casual interaction with the environment. For example, people develop sets of beliefs about attitude based on various situations during which their attitude played a role (schemata and scripts), as well as observing other people in those kinds of situations. These beliefs allow them to explain and predict how attitude might affect people acting under various conditions. Scientists, on the other hand, have formalized these beliefs into theories about self-efficacy, for one, primarily through more rigorous scientific procedures involving hypothesis testing, statistical analysis, and model building.

Some argue there is little difference in principal between formal and informal theories and how they are formed, others argue there is considerable difference. Regardless, the TBM can incorporate formal and informal theories. The theories can be tested through creation of an environment where software agents interact with people in controlled contexts.

Implementation of TBM begins with determining a goal and a context in which to test or develop theories. The types of cognitive acts that are likely occur are defined. For example, two competing theories of learning and instruction can be implemented and tested in educational technology. The technology would incorporate software agents as instructional agents to give

students advice based on one or other of the theories. The effects of the differing instructional acts on learning acts could be compared.

Another example is in health communication. At least four theories have been used to predict the most effective kinds of health messages for, say, helping people quit smoking. A software agent system could be designed that has four competing agents, each acting under the beliefs of one of the theories. A test could be designed to provide evidence which of the agents performs more effectively. Or, an entirely new theory may emerge based on a software agent designed to incorporate all four theories in a hybrid theory.

The challenge in terms of theory-testing includes several issues:

(a) Has the theory been operationalized successfully?
(b) Has the theory been implemented successfully in the software agent knowledge and reasoning algorithms?
(c) Do the theories lend themselves equally readily to use with agents? In other words, might some theories describe people, but not describe an easily-derived model for building an agent?

The following examples illustrate how the TBM can be implemented and tested.

Educational Technology Example

In this example, the goal is to design educational technology using software agents that evaluates two competing theories of instruction. The context of the example is middle school students learning through doing scientific inquiry projects. Thus, the two types of cognitive agents involved are students and software agents that provide instructional advice on doing and learning scientific inquiry.

The main types of cognitive acts in the example are instructional acts, task completion acts, and learning acts. The software agents perform the instructional acts by giving advice. Students perform the task completion acts by doing their projects. Both software agents and students perform learning acts. Software agents learn about the students they are advising, students learn about science (and about the agents).

The two competing instructional and learning theories in this example are information processing and constructivism. Information processing theory focuses on fundamental mental operations, mainly how we perceive and remember events and information. It is primarily concerned with the three basic components of memory: the sensory register, short-term or working memory, and long-term memory. Implications for learning and instruction include providing organized instruction, recognizing the limits of attention, recognizing the limitations of short-term memory, and matching encoding strategies with the material to be learned.

Constructivism, a less-well formed theory than information processing, focuses on the idea that an individual actively constructs meaning in the world where there is no inherent meaning. Meaning is a highly individual process of actively building on or modifying existing knowledge.

As a set of instructional practices, constructivism favors processes over end products; guided discovery over expository learning; and authentic, embedded learning situations over abstracted, artificial ones.

The computer environment (SCI-WISE) in which the students do their work and where the advisors reside was designed with two types of agents: one that presents advice based on information processing theory (IP) and one using constructivist principles (C). In brief, the IP agents gave specific advice at the time it was considered appropriate. The C agents allowed the students to select what kind of advice they desired, when they received advice, and allowed students to modify the advice the agent gave.

Because the C agent system is a more self-guided discovery system it was hypothesized that student goal orientation toward learning and doing science would affect use and evaluation of the system. Goal orientation is a continuum with task completion on one pole to knowledge building on the other.

A study of sixth grade students was conducted. Students took a pretest questionnaire that measured their goal orientations toward science projects as well as their inquiry skills. The students worked in pairs (simulating the collaborative aspects of science) on an open-ended inquiry project that requires complex reasoning about human memory. The students used either the IP or the C version of SCI-WISE. After finishing the project, the students took a posttest questionnaire similar to the pretest, and evaluated the version of the system they used.

The main results showed that knowledge-oriented students rated either system useful, but at least in the beginning, getting advice in an IP fashion tended to help some students better learn the subject matter. Knowledge-oriented students tended to make more use of the modifiable features of the C system. Task oriented students clearly preferred the IP system over the C system, likely because it gave them immediate advice to help them complete their tasks.

The main implication for whether to use the IP theory or the C theory is that there may be situations where one or the other is preferable. Individual factors, such as goal orientation, as well as the current context are important in determining which theory would be preferable.

Health Communication Example

The goal of the example is to develop a system incorporating software agents that deliver highly-individualized health messages. The context is helping people of any age to quit smoking. The two types of agents are smokers and software agents who act to emulate smoking cessation counselors.

The main cognitive acts in this example are communication acts, health behavior acts, and learning acts. People’s health behavior acts are actions that positively or detrimentally affect their health. Learning acts can be performed by either smokers or software agents. Smokers learn about their habit and ways to quit; agents learn about smokers and the most effective messages.

Communication acts are any events that transmit messages containing information, including questions or requests for clarification. Communication acts are more encompassing than the more common term “speech acts,” primarily because the ability to speak should be attributed to humans and not to software agents (at this time). Also, communication in this system is not limited to speech but could incorporate other media such as visual representations or graphics, or even facial expressions if the software agents are given an animated onscreen presence.

Human behavior change is a difficult process to describe and assess, largely because of the many individual and social factors involved. The theories or models that have been developed largely from psychology attempt to capture general patterns of empirical research results. Each of four theories will be briefly introduced and its uses and limitations discussed. This section is not an exhaustive review of the theories, but illustrates the many factors that must be considered and how difficult it is to capture the mechanisms of human behavior change to use in computer-mediated systems.

Health Belief Model

The Health Belief Model (HBM) was developed in the 1950s to explain and predict health behaviors using the attitudes and beliefs of individuals. Key variables of the HBM include (a) perceived threat of a health condition, (b) perceived susceptibility contracting the condition, (c) perceived severity of the condition, (d) perceived benefits of reducing the threat of illness, (e) perceived barriers of particular health actions, and (f) demographic, socio-psychological, and environmental variables that affect an individual's perceptions and self-efficacy (belief in being able to successfully execute the behavior required to produce the desired outcomes).

HBM research has been used to explore a variety of health behaviors including smoking cessation in diverse populations. Perceived barriers have been identified as the most influential variable for predicting and explaining health-related behaviors. Other significant HBM variables are perceived benefits and perceived susceptibility, with perceived severity identified as the least significant variable. Using this model for smoking cessation would indicate focusing messages on overcoming perceived barriers toward smoking cessation, as well as the benefits of quitting. Focusing on the severity of cancers that are attributed to smoking would have a lesser effect.

Limitations of the HBM include: (a) the usefulness of the model as a whole has not been tested as researchers have tested only selected variables, (b) other factors such as environmental or economic are not included in the model, and (c) the model does not incorporate the influence of social norms and peer influences on people’s decisions regarding their health behaviors.

Theory of Reasoned Action

The Theory of Reasoned Action (TRA), developed in the 1960s, is based on the premise that humans are rational and that the behaviors being explored are under volitional control. The theory links individual beliefs, attitudes, intentions, and behavior. The theory’s variables are (a) behavior, defined in terms of action, target, context, and time, (b) intention to perform the behavior, (c) attitude toward performing the behavior, (d) beliefs about the potential outcomes of a defined behavior, and (e) norms and normative beliefs regarding a person’s beliefs of how the behavior is viewed by others.

The TRA proposes a linear process in which changes in an individual's behavioral and normative beliefs will affect the individual's actual behavior. Behavioral and normative beliefs influence individual attitudes and subjective norms, which shape a person's intention to perform a behavior. Intention, in this theory, is the strongest indicator that the desired behavior will be achieved. To develop appropriate interventions for a specific population and behavior, the variable and its corresponding beliefs and norms that exerts the greatest influence on the population should first be determined. For example, in the case of smoking cessation, intention toward quitting smoking may be the most important indicator behavioral change, and are dependent on an individual's beliefs about the behavior and norms associated with it.

Some limitations of the TRA include the inability of the theory, due to its individualistic approach, to consider the role of a person's environmental, and the linearity of the theory components. Specifically, individuals may first change their behavior and then their beliefs and attitudes about it.

Elaboration Likelihood Model

The Elaboration Likelihood Model (ELM) developed in the 1980s is a framework for organizing and understanding the basic processes of attitude change. ELM incorporates many of the major classic approaches to attitude and persuasion. According to ELM, people take either the central route (attributed to highly motivated individuals) or the peripheral route (less motivated) to attitude change. Attitude change via the central route, based on a thoughtful consideration of the issues presented, is relatively permanent, resistant to counter-argument, and predictive of behavior. Factors of the central route include (a) motivation to pay attention to the message, (b) ability to understand the message, (c) prior attitude to accept messages, and (d) the argument strength.

The peripheral route results in an attitude change that is relatively temporary, susceptible to counter-argument, and less predictive of behavior. Factors of this route include (a) reciprocation or prior obligation, (b) consistency in feelings, (c) social proof that others think or feel a certain way, such as liking the message deliverer, (d) authority of the message deliverer, and (e) scarcity of opportunity.

High motivation and high ability are necessary for a high probability of following the central route. Under other circumstances, where either one is absent, the peripheral route will likely be followed. People taking peripheral route might change to the central route, but this is not usual considered when using ELM, ignoring possible dynamic interaction between steps. The model also assumes people can be classified into categories such as having the ability to process a message or not having the ability. An individual likely understands some portion of messages or issues, and partially understands other portions of the messages. For smoking cessation, ELM indicates that messages should be understandable and focused on increasing motivation and ability.

Stages of Change Theory

Psychologists developed the Stages of Change Theory in the 1980s to compare smokers in therapy and self-changers along a behavior change continuum. The rationale behind the theory was to tailor therapy to individual's needs at their stage in the change process. The five

components of the Stages of Change Theory—precontemplation, contemplation, preparation for action, action, and maintenance—were identified and presented as a linear process of change.

A description of each stage are: (a) precontemplation—an individual has the problem, recognized or not, and has no intention of changing, (b) contemplation—the individual recognizes the problem and is seriously considering changing, (c) preparation for action—the individual recognizes the problem and intends to change the behavior within the next month, (d) action—the individual has enacted consistent behavior change for less than six months, and (e) maintenance—the individual maintains new behavior for six months or more. Stages of Change indicates messages should be tailored to an individual's current stage to move them toward the next.

As a psychological theory, the Stages of Change focuses on the individual without assessing the role that environmental issues may have on a person's ability to enact behavior change. Because the theory presents a descriptive rather than a causative explanation of behavior, the relationship between stages is not always clear. Indeed, the stages are no longer considered linear; rather, they are components of a cyclical process that varies for each individual. Finally, each of the stages may not be suitable for characterizing every population.

Theory/model	Description	Smoking cessation use	Limitations
Health Belief Model	explains and predicts health behaviors using the attitudes and beliefs toward disease, especially perceived barriers, perceived benefits, and perceived susceptibility	focus messages on overcoming perceived barriers toward smoking cessation, as well as the benefits of quitting	not tested as a whole, environmental or economic factors are not included, does not incorporate the influence of social norms and peer influences
Theory of Reasoned Action	links individual beliefs, attitudes, and intentions, and assumes that behaviors are under volitional control, and intention of quitting smoking is the most important indicator of behavioral change	focus messages on an individual's beliefs and attitudes to increase intention of quitting smoking	does not consider environmental issues, and assumes linearity of the theory components when they may be cyclical
Elaboration Likelihood Model	attitude change via the central route (individuals are highly motivated) is relatively permanent, resistant to counter-argument, and predictive of behavior; the peripheral route (individuals are less motivated) is less so	create messages that are understandable and focused on increasing motivation and ability	high motivation and high ability are necessary for a high probability of following the central route, dynamic interaction between steps, assumes people can be classified into categories
Stages of Change Theory	five stages are precontemplation, contemplation, preparation for action, action, and maintenance; no longer considered linear; rather, stages are components of a cyclical process that varies for each individual	tailor messages to an individual's cyclical stage of change process	doesn't account for environmental factors, presents a descriptive rather than causative explanation of behavior, each stage may not be suitable for characterizing every population

In the smoking cessation example, as in the educational technology example, a software agent for each of the theoretical stances would be designed. A specific example of this process follows for the Stages of Change Theory.

The system asks users to provide demographic information (e.g., age, gender, education level, ethnicity). They will be asked what tobacco products they use and whether they are interested in quitting smoking. They will complete questions to assess their readiness to quit, number of prior quit attempts, and previous use of primary cessation counseling and its success. The scripts will begin to classify smokers according to stage of change. For example, *precontemplators* will be defined as smokers who do not express an interest in quitting within the next 30 days. *Contemplators* are those smokers who are thinking about quitting in the next 30 days. *Preparers* are those definitely planning to quit in the next 30 days, while those in the *action* stage are recent quitters.

In general, messages attempt to move smokers to the next stage of quitting. For example, precontemplators will receive tailored-messages that give reasons for quitting, e.g., health, cost savings, environmental tobacco exposure within the home and work place, being a role model for the family, smell, appearance factors, food tasting better, among others. The messages can be evaluated by the smoker, or the smoker can request more information. The software agents learn more about the individual smoker, as well as develop more data about the messages given in the different stages.

Several research questions could be investigated using the software agents. One would be to strengthen the theory by discovering how overcome its limitations. Another would be to discover the individual and situational differences that might make one theory more useful than another. Another would be to develop and test new or hybrid theories.

Research investigations could compare software agents using different aspects or versions of a theory, or head-to-head comparisons of different theories (in which agents compete), or cooperative tests (in which agents collaborate). Primary effects on smokers that could be measured are success in quitting, intention to quit, reported relevance of the messages received. Secondary measures include process satisfaction and message recall.

Conclusion

The Theory Belief Model posits that cognitive agents (people or software agents) perform cognitive acts (perceiving-reasoning-taking action) based on theories (sets of beliefs). When designing computer systems with software agents using artificial intelligence techniques, these theories can be incorporated to allow rational cognitive acts to occur within the framework presented by the theories. In other words, the software agents act according to the beliefs indicated by the theories.

When people use the software agent systems, the theories can be evaluated through research investigations. Competing theories can be compared, such as information processing theory versus constructivist theory in designing educational technology. In such an investigation, it was shown that individual and situational differences may make one theory more relevant than another. Another area ripe for such theory comparison is health communication. A review of four health behavior theories indicates many research questions that could be investigated to improve health message delivery.