OVERVIEW OF THE USE OF ADAPTIVE ARTIFICIAL INTELLIGENCE TECHNIQUES FOR THE DESIGN OF LIFELIKE AGENTS

Toys and consumer electronics embodying interactive, lifelike agents are the rage. Sensory took the first big step toward natural interaction by giving consumer products the ability to speak and hear. Artificial intelligence (AI) techniques can provide the next step toward realistic, interesting behavior that will capture the attention and imagination of users.

Sensory's expertise in speech processing and AI, combined with its many years of design experience with toys and consumer electronics, enable Sensory to create a new breed of interactive, lifelike agent.

Adaptive artificial intelligence

Traditionally, AI agents are designed by specifying a set of agent behaviors i.e., how the agent is to respond to each situation it encounters. Such agents are not very interesting because their behavior is _static_: they follow fixed, preprogrammed rules.

Sensory's approach is to build agents that learn and whose behavior changes and evolves over time. We provide a set of adaptive AI (AAI) tools that facilitate the creation of adaptive agents with the following characteristics:

* Each agent has a unique personality that is shaped by its innate dispositions and by recent experience.

* The personality is characterized by _drives_ (e.g., the desire to eat, rest, play, entertain, etc.), _emotions_ (e.g., happiness, crankiness, anger), and _preferences_ (e.g., for one food over another, or one activity over another).

* At any moment in time, the agent has a _drive state_ that indicates the strength of the agent's drives. For example, the agent may have a strong hunger, but little need for rest.

* Likewise, the agent has an _emotional state_ that indicates the degree to which it is happy, angry, etc.

* The agent also has a _preference state_ that indicates whether it prefers bananas to apples, etc.

* Together, the agent's drive, emotional, and preference states are referred to as the _internal state_. The internal state affects the actions it performs. When an agent is cranky, it will prefer a response like "whatever" over "sure, let's do that!" When an agent is hungry, it will fixate on activities that provide it with food, or choose to sing songs or tell jokes that are related to food. Further, it will have distinctive preferences for one food item over another. Action choice is nondeterministic, meaning that the agent's behavior, while guided by its drives, emotions, and preferences, includes an element of randomness that makes it not entirely predictable.

* The internal state of an adaptive agent changes over time. The AAI tools specify how this change occurs, and how the internal state influences its choice of actions.

Some details

Internal state

AAI is based on an internal state that characterizes the mental life of the agent. The components of the internal state fall into three categories:

(1) The drive state is a vector with an element for each drive ranging from 0 to 1 indicating the intensity of the drive at a particular point in time.

(2) The emotional state is characterized by multiple dimensions. Sample dimensions, and examples of states lying on the dimension are:

    ANGRY - CRANKY - UPBEAT - JOYFUL
    TERRIFIED - FEARFUL - CONFIDENT - COURAGEOUS
    BORED - SURPRISED - INTERESTED

Each dimension is represented by a value ranging from 0 to 1 which indicates where on the continuum between the extremes the current state lies. The emotional state is then a point in a multidimensional hypercube.

(3) The preference state is characterized by multiple dimensions, each corresponding to some class of items it encounters in its universe, e.g., food, games, activities, friends, etc. One each dimension, the preference state is indicated by an n-dimensional vector, where "n" is the number of alternatives available to the agent. For example, on the food dimension, the alternatives might be bananas, oranges, and apples4 on the game dimension, the alternatives might be peek-a-boo, guess-my-number, and name-the-color. Associated with each alternative on each dimension is a numeric value in the range 0 to 1 which indicates the relative strength of preference for that alternative. For example, an agent can eat bananas, oranges, and apples could have a preference vector [0, .5, 1], meaning that the preference for apples is twice that of oranges, and that the agent does not like bananas.

Dispositions

The general personality of the agent is specified by _dispositions_, which are simply default values for the emotional states and preferences of the agent. One agent may intrinsically prefer bananas to apples, another may prefer apples to bananas. One agent may be fundamentally cranky, while another is fundamentally upbeat.

Via dispositions, each agent can have a unique personality.

Adaptive behavior

Given this framework, we can specify how behavior adapts with time and experience. We give an overview of various sorts of adaptation.

Drives:
- Drives such as hunger increase at a fixed rate over time, or with the activity level of the agent.
- Drives are reset by sensor inputs (e.g., receiving food, being shut off)

Preferences:
- Repeated selection of an item leads to habituation -- a lowering of the

preference for that item. For example, after eating a banana, the preference for the banana drops.
- The preference returns to disposition level gradually over time.
- Preferences can be modulated by emotional state (e.g., comfort foods are preferred when the agent is sad).
- Choice behavior: when faced with a choice among alternatives, the probability of choosing an alternative is proportional to its current preference level.

Emotions:
- Emotional states are determined by drive levels and by events in the world.
- Negative emotional states arise either from a strong drive not being fulfilled, or from some desired object/sensation being withdrawn.
- Positive emotional states arise either from the fulfillment of a strong drive, or from the appearance of a desired object/sensation.

Behavior
- Typically in toy products, an agent's behavioral loop consists of choosing a high-level activity (e.g., eating, playing ball, napping, etc.) either at random or according to a deterministic script. The AAI technology allows choice of activity to be guided by both current drives and emotional state. This choice is characterized by a probability distribution
  P(activity | primitive state).
- The same technique can be applied to choice of specific actions within an activity category (e.g., which song to sing).

- Choice of utterances is determined by emotional state: Any response of the agent can be conditioned on the sign of its affect. This choice is characterized by a probability distribution
  P(utterance | emotional state). For example, a response to a user in a positive affective state might be, "That's awesome" and in a negative affective state, "Yeah, whatever."

Reinforcement learning
- Reinforcement learning (RL) is a technique for allowing an agent to change its behavior so as to maximize its reward and minimize its punishment. To design an RL system, the agent must have some means of obtaining reinforcement (i.e., reward and/or punishment). Reinforcement can be internal (e.g., food, not being shut off) or external (e.g., user giving verbal feedback such as "bad dog", user turning agent upside down, user shouting at agent, etc.). Simple RL is concerned with choosing a single action to maximize reward; sequential RL is concerned with choosing a sequence of actions.
- Reward and punishment are also linked to changes in emotional state, as described above.
- Examples of reinforcement learning are:
  * An agent may "learn to speak" by starting with gibberish, and with reinforcement, its utterance would gradually be transformed into English. One type of reinforcement would be a user rewarding the agent for speaking English ("good dog" or "that's right"). Another type of reinforcement would simply be the amount of speech background sound in the environment.
  * An agent might learn to avoid bedtime by choosing actions that keep the user engaged.
  * The state-contingent choices of an agent, P(activity | state) and P(utterance | state) can be shaped by reward in the past.

The AAI Tools

Specifically, the AAI tools provide a set of procedures that perform updates. Given the following variables,

D : drive state
E : emotional state
P : preference state
T : passage of time
I : innate dispositions
S : sensory experience (sensors, speech input)
A : agent actions

we have the following update procedures:

D' = f1(D, T, A)
E' = f2(E, D, S, I)
P' = f3(P, T, I)
A' = f4(D, E, P, S)

where f1(.), f2(.), and f3(.), and f4(.) are functions (computer code) that achieve the adaptive AI behavior. Some aspects of this code are tailored to the particular application, e.g., the sensory signals available to the agent; other aspects are general proprietary algorithms Sensory has developed.