Notes for Meeting 25
Cognition and Personality

Personality and Behavior

In many cases, different people behave differently in the same setting.

We often attribute these behavioral differences to distinctions in the people’s personalities.

Psychology has long considered the scientific study of personality to be an important endeavor.

However, there has been little AI work on this topic, except in the area of synthetic characters.

What is Personality?

We can define personality as a set of stable behaviors in certain classes of situations.

Not surprisingly, some psychological theories explain personality in behaviorist terms, treating it as a set of stimulus-response pairs.

Other theories instead posit a set of internal personality "traits" that influence behavior.

The first framework views personality as learned and changing, while the second views it as innate and stable.

Few accounts of personality make contact with results from cognitive psychology, especially those using computational models.

An Example Trait Theory

One widely known account of human personality is the Five Factor theory (Digman, 1990), which proposes five high-level traits:

- Openness (inventive/curious vs. consistent/cautious). Tendency to appreciate new and varied experiences.

- Conscientiousness (efficient/organized vs. easy going/careless). Tendency to exhibit self discipline and planned behavior.

- Extraversion (outgoing/energetic vs. solitary/reserved). Tendency to be stimulated in others' presence.

- Agreeableness (friendly/compassionate vs. cold/unkind). Tendency to be compassionate and cooperative toward others.

- Neuroticism (sensitive/nervous vs. secure/confident). Tendency to experience unpleasant emotions like anger and anxiety.

These traits appear to describe personality differences across many culture, and they appear to be stable in adulthood.

An older theory (Catell, 1947) posited 16 distinct personality traits.

* Adapted from Wikipedia entry http://en.wikipedia.org/wiki/Big_Five_personality_traits

Emotions and Personality

We have seen that emotions are transient while personalities are reasonably stable.

Nevertheless, some accounts assume close relations between personality and emotion.

In this framework, personality depends on factors like:

- How easily one exhibits certain emotions
  - E.g., some people are easily angered, others are often happy
- How one responds to a given emotion
  - E.g., some people raise their voice when angry, others are quiet

According to this view, personality traits are simply descriptions of these relations.

Dimensional Theories and Synthetic Characters

Personality traits need not be all or none; they are more naturally viewed as points along a continuum.

In this view, one can describe an agent’s personality as a point in multi-dimensional space.

For example, the five factor theory implies a five-dimensional space.

Some research on synthetic characters adopts this approach to specifying agent personality.

Such systems typically relate stable personality variables to transient emotional variables in another space.

A Cognitive Systems Account of Personality

One drawback of trait/dimensional theories is that they provide no explanation of the traits' origins.

Rizzo et al. (1997) offer a deeper account based on Ford (1992) that casts personalities as:

- Abstract goals and associated priorities

- Preferences over operators or plans for achieving these goals

They have implemented this theory within the Prodigy architecture, which is organized around means-ends problem solving.

The authors demonstrated its potential by showing how different personalities respond in helping scenarios.

More on Goals and Personality

Rizzo et al.’s account explains the origins of traits in the agent’s abstract, high-level goals.

A more advanced version of this theory would appear to require:

- Long-term, generalized conditional goals that the agent can use to generate specific, concrete goals.
  - The former play the role of stable traits, while the second drive behavior in specific situations.
- An expanded goal language that refers to other agents’ beliefs, goals, intentions, and even emotions.
  - This suggests that personality "traits" are also metacognitive, but operate an an even higher-level than emotions.

To the extent that the agent can gradually learn/revise such structures, this scheme unifies trait and learning approaches to personality.

This framework retains the notion of personality dimensions, but the number of dimensions may change over time.

Assignments for Meeting 26
Metareasoning and Metacognition

Read the article:

* Cox, M. T. (2005). Metacognition in computation: A selected history. Proceedings of the AAAI Spring Symposium on Metacognition in Computation (pp. 1-17). Stanford, CA: AAAI Press. [required]

* Cox, M. T. (2007). Perpetual self-aware cognitive agents. AI Magazine, 28, 32-45. [optional]