Emotions, Personality, and Cognition

CompSci 765
Meetings 24 and 25

Pat Langley
Department of Computer Science
University of Auckland
Why Study Emotions and Personality?

This week we examine two topics – emotions and personality – not usually included in AI courses.

There are two main reasons to study these important topics:

• **Scientific**: They may be necessary elements of a complete theory of intelligence;

• **Engineering**: They may be essential to building believable synthetic agents and effective tutors.

We maintain that both involve *higher-level* aspects of cognition, similar to explanation, planning, and dialogue.

Also, they are best studied in terms of their interaction with such abilities, which they serve to *modulate*. 
The Pervasiveness of Emotions

Emotions play a central role in most aspects of human life; they color and modulate our activities, both physical and mental. This raises an important and interesting scientific question:

- *How are emotions related to cognition?*

More broadly, what function do emotions serve in an integrated cognitive system?

Science fiction often depicts human-level AI systems as devoid of emotion, but does this really make sense?
Emotions and Rationality

The traditional view is that emotions are ‘irrational’ holdovers from our evolutionary precursors.

This perspective influenced much early AI work, which held emotions as being *detrimental* to intelligent systems.

- Clearly, we can build programs that – to some extent – reason, plan, and communicate without emotional components.
- But Simon (1967) has argued that affect and emotion play important roles in controlling cognitive attention.
- And Damasio (1994) describes brain-damaged humans who have little or no emotion who cannot make decisions.

This suggests that human-level cognitive systems may actually *require* emotions.
Some Distinctions

Both academic papers and everyday language often confuse key concepts that are quite distinct:

- **Affect.** The positive or negative aspect of some experience.
- **Mood.** A global variant of affect for the entire cognitive system.
- **Emotion.** A mental structure related to goals and beliefs about an event, agent, or object.
- **Feeling.** An affective or hormonal response that is associated with an emotion.

A complete account would relate each such factor to cognition. Here we will focus on emotions, which are the most complex and interesting from an AI perspective.
### Examples of Emotions

We view many emotions as important enough to name, such as:

<table>
<thead>
<tr>
<th>Happy</th>
<th>Sad</th>
<th>Angry</th>
<th>Afraid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Worried</td>
<td>Despairing</td>
<td>Love</td>
<td>Proud</td>
</tr>
<tr>
<td>Courageous</td>
<td>Disappointed</td>
<td>Relieved</td>
<td>Pleasantly surprised</td>
</tr>
<tr>
<td>Frustrated</td>
<td>Satisfied</td>
<td>Helpless</td>
<td>Annoyed</td>
</tr>
<tr>
<td>Irritated</td>
<td>Disgusted</td>
<td>Resentful</td>
<td>Envious</td>
</tr>
<tr>
<td>Jealous</td>
<td>Embarrassed</td>
<td>Guilty</td>
<td>Ashamed</td>
</tr>
<tr>
<td>Regretful</td>
<td>Offended</td>
<td>Self righteous</td>
<td>Sympathetic</td>
</tr>
<tr>
<td>Pitying</td>
<td>Amused</td>
<td>Wonder /Awe</td>
<td>Schadenfreude</td>
</tr>
</tbody>
</table>

Other mammals have emotions, but human variants are unique in their richness and complexity.

This suggests there is a strong cognitive component to emotion.
Representing Emotions

Before we can discuss emotional processes, we must consider how to represent them.

Marsella, Gratch, and Petta (2011) distinguish three main ways to encode emotional content:

• *Dimensional* models – points in N-dimensional space;
• *Anatomical* models – activations in neural circuits;
• *Appraisal* models – relations among cognitive structures.

These frameworks suggest some very different ways to explain emotional processes.
Dimensional Models

Most dimensional models characterize emotions as *points* in a three-dimensional space:

- *Pleasure* – measure of valence;
- *Arousal* – level of affective activation;
- *Dominance* – measure of control.

Synthetic characters often use such “PAD” models (Wachsmuth, 2008), but they ignore some important facts:

- Emotions are *about* some event, person, or object;
- We can have *mixed* emotions about the same target.

This suggests that they involve much richer cognitive structures.
Emotions as Cognitive Structures

Appraisal models view emotions as inferred relations among mental structures based on situations.

Ortony et al. (1988) describe 22 configurations that characterize emotions organized around events, objects, and other agents.

These patterns serve as ‘elicitation’ conditions on emotions that involve relations among:

- An agent’s goals, intentions, expectations, and beliefs;
- An agent’s inferences about others’ beliefs, goals, intentions.

This suggests that emotional structures are abstract and domain independent, much as the inference rules for dialogue.
Two Facets of Emotions

We can distinguish between two aspects of emotions that seldom appear in the literature:

- **Short-term** structures that encode emotional *instances*:
  - disappointed(John, failed(John, CompSci101))
  - resentful(John, passed(Sam, CompSci101))

- **Long-term** generic structures that *generate* these instances

This is analogous to the distinction between concrete beliefs and conceptual rules that define them.

Both appear necessary for a full cognitive account of emotion.
Long-Term Emotional Structures

We can state conditions for eliciting emotions as abstract rules:

- An agent is *disappointed* about an event if the agent wanted the event, expected the event, and believed it did not occur.
  
  \[
  \text{disappointed}(\text{Agent}, \text{Event}) : - \\
  \text{goal}(\text{Agent}, \text{Event}), \text{expect}(\text{Agent}, \text{Event}), \\
  \text{belief}(\text{Agent}, \text{not}(\text{Event})).
  \]

- An agent is *jealous* of another agent if he wants an object, believes he does not possess it, and believes the other does possess it.
  
  \[
  \text{jealous}(\text{Agent}, \text{Other}, \text{Object}) : - \\
  \text{goal}(\text{Agent}, \text{possess}(\text{Agent, Object})), \\
  \text{belief}(\text{Agent, not(possess}(\text{Agent, Object}))), \\
  \text{belief}(\text{Agent, possess}(\text{Other, Object})).
  \]

Complex emotions are specializations of basic ones that involve more conditions.
Two Emotion-Related Tasks

We can distinguish two cognitive tasks that use such long-term emotional structures:

• *Generation*, which produces emotions for a primary agent
  • disappointed(John, failed(John, CompSci101))

• *Understanding*, which infers the emotions of other agents
  • belief(John, disappointed(Jane, failed(John, CompSci101))))

This is analogous to the distinction between plan generation and plan understanding.

Again, both appear necessary for a full computational account of emotion’s relation to cognition.
Intensity of Emotional Feelings

Although emotions have a clear cognitive structure, they also have associated feelings.

- One aspect of feelings is their intensity, which begins high and gradually decays over time.
- Memories of past emotions can recreate these feelings, but usually not as strongly as the original.

Computational models that address intensity often calculate it with equations based on utilities and probabilities.

Such calculations are more central to dimensional models than to appraisal accounts, but both can include them.
Effects of Emotions

Most accounts of emotion model generation of emotions but not their effects.

More complete models clarify how influence either the agent’s:

• physical behavior
  • e.g., crying about loss or punching someone in anger
• cognitive processing
  • e.g., changing goal priorities or invoking planning

This suggests that emotions are not evolutionary relics; instead, they serve as high-level regulators of cognition.
We unusual aspect of human cognition is the ability to think about thinking, or *metacognition*, which:

- *Given*: Traces of base-level cognition (e.g., inferences of plans) and reasoning chains that support them;
- *Given*: Abstract structures that specify patterns over these base-level cognitive traces;
- *Find*: Instances of these patterns that alter base-level cognition.

Metacognition operates at a higher level than regular thinking, inspecting it and modulating base-level processes.
Emotions as Metacognition

Remember that emotional concepts specify abstract relations among goals, beliefs, and expectations.

Also recall there is evidence that emotions influence domain-level decision making and planning.

This leads naturally to the *emotional metacognition hypothesis*:

- *Emotions play a metacognitive role that operates over and influences base-level cognition.*

That is, they inspect traces of basic cognition and alter its course in response.

This view follows Simon (1967) in claiming that emotions play a *regulatory* role in cognition.
Emotions: Summary Remarks

Different theories of emotion make different assumptions, but the most promising claims that:

• Emotions are symbolic *cognitive structures*;
• Long-term emotional *rules* generate short-term *instances*;
• These rules are used to *generate* and *understand* emotions;
• Emotions play a *metacognitive* role in influencing behavior.

In this account, emotions are not irrational evolutionary relics; they are central to producing human-level intelligence.
Different people often behave quite differently in very similar circumstances.

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

This raises another important and interesting scientific question:

- How is personality related to cognition?

Psychology has long considered the study of personality to be a legitimate scientific endeavor.

But there has been little AI work on this topic, except in the area of synthetic characters.
What is Personality?

We associate personality with *stable behaviors* in certain types of situations, but we should also explain their basis.

- Some psychological theories explain personality in behaviorist terms, treating it as a set of *stimulus-response pairs*.
- Other theories instead posit a set of fixed 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, especially the computational ones, make contact with results from cognitive psychology.
Digman’s (1990) *five factor theory* of personality proposes five high-level traits:

- **Openness**. Tendency to appreciate new and varied experiences.
- **Conscientiousness**. Tendency to exhibit self discipline and planned behavior.
- **Extraversion**. Tendency to be stimulated in others’ presence.
- **Agreeableness**. Tendency to be compassionate and cooperative toward others.
- **Neuroticism**. Tendency to experience unpleasant emotions like anger and anxiety.

These traits appear to describe personality differences in many cultures. An older theory (Cattell, 1947) posited 16 distinct traits.
Emotions and Personality

Emotions (rather, emotion instances) are transient; personalities are reasonably stable.

But some accounts assume close relations between personality and emotion.

In these 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 raise their voices when angry, others are quiet

In this view, personality traits are simply descriptions of these long-term relational structures.
We view many aspects of personality as central enough to name:

<table>
<thead>
<tr>
<th>Friendly</th>
<th>Distant</th>
<th>Organized</th>
<th>Careless</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caring</td>
<td>Unconcerned</td>
<td>Thoughtful</td>
<td>Thoughtless</td>
</tr>
<tr>
<td>Selfless</td>
<td>Selfish</td>
<td>Giving</td>
<td>Greedy</td>
</tr>
<tr>
<td>Persistent</td>
<td>Relenting</td>
<td>Stubborn</td>
<td>Compromising</td>
</tr>
<tr>
<td>Judgmental</td>
<td>Forgiving</td>
<td>Relaxed</td>
<td>Tense</td>
</tr>
<tr>
<td>Loyal</td>
<td>Disloyal</td>
<td>Reliable</td>
<td>Unreliable</td>
</tr>
<tr>
<td>Energetic</td>
<td>Lazy</td>
<td>Confident</td>
<td>Timid</td>
</tr>
<tr>
<td>Brave</td>
<td>Cowardly</td>
<td>Open minded</td>
<td>Dogmatic</td>
</tr>
</tbody>
</table>

This very partial list suggests that trait theories are unlikely to cover observed variations.

And it also suggests a major cognitive component to personality.
One drawback of trait / dimensional theories is that they offer no explanation of traits’ origins.

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

• Abstract goals and associated priorities
  • E.g., material wealth, safety, help to others, entertainment
• Priorities on operators / plans for achieving them

They implemented this theory in a problem-solving architecture that uses goal-driven planning.

They also showed how agents with different personalities acted differently in the same scenarios.
Goals and Personality

A more complete account of personality would also include:

• Generalized, conditional goals that the agent uses to generate specific, concrete goals.  
  • The former encode stable traits, while the latter drive behavior in specific situations.

• An expanded goal language that refers to others’ beliefs, goals, intentions, and even emotions.
  • Personality “traits” are also abstract mental structures, but ones that operate at a higher-level than emotions.

If an agent gradually learns / revises such structures, this scheme unifies trait and learning approaches to personality.

The number of personality dimensions may change over time.
Personalities and Metacognition

Again, *metacognitive* processes inspect traces of base cognition and influence its operation.

Goal-based accounts of personality clarify how such abstract processes can influence an agent’s:

• physical behavior (e.g., tendency to flee or use force)
• cognitive processing (e.g., amount of planning before acting)

This suggests in turn the *personality metacognition hypothesis*:

• *Personality plays a metacognitive role that operates over and influences base-level cognition.*

But the effects of personality are even higher level than those of emotions, since they modulate the latter.
Different accounts of personality make different assumptions, but the most promising claims that:

- Personalities are symbolic *cognitive structures*;
- Abstract prioritized rules determine one’s top-level *goals*;
- *Variations* in this knowledge give personality *differences*;
- Personality exerts a *metacognitive* influence on behavior.

In this view, personality is not a point in an N-dimensional space, but rather a set of abstract rules for guiding cognition.
Moral Cognition and Consciousness

Two other critical aspects of human intelligence have received little attention from AI researchers:

• Moral judgement and decision making
  • These often involve generic relations among beliefs and goals of the primary and other agents

• Awareness of cognition / consciousness
  • Which operates over traces of base-level cognitive processing

These have the same abstract, domain-independent character as emotions and personality, but at even higher levels.

Both also appear to be *metacognitive* processes that examine and modulate lower-level behaviors.
Readings on Emotion and Personality


• http://en.wikipedia.org/wiki/Big_Five_personality_traits
End of Lecture