6. Cognitive and Affective Computing

Toyoaki Nishida
Kyoto University
Studies on conversational interactions

(1) Narratives and content flow
(2) Social discourse
(3) Verbal communication
(4) Nonverbal communication
(5) Cognitive process

[Nishida et al 2014]
The question is not whether intelligent machines can have any emotions, but whether machines can be intelligent without any emotions” (Minsky: Society of Mind).

BDI does not provide any theoretical ground to answer question such as: What should be an agent’s desires (and why)? Why should one desire be preferred over another? [Lim 2010]

Deeper source of behaviors ➞ Affective computing

[Georgeff 1990]
The vehicle on the left may appear to fear the light, while that on the right may appear as if it liked the light.

[Braitenberg 1984]
ALIVE

[Maes 1995; Nishida et al 2014]
Silas the dog

Interests / Goals / Motivations

Releasing mechanism
Sensory system

Behavior

Inhibition

Motor commands

World

[Blumberg 1997]
Emergent emotions and emotional behavior
  - Attributed to systems based on their observable emotional behaviors

Fast primary emotions
  - Innate, quick and dirty reactions
  - Include at least fear, surprise, and anger

Cognitively generated emotions
  - Explicit reasoning is typically involved in their generation.
  - “Don’t worry; be happy”
  - Reason about emotions (e.g., using the Ortony Clore Collins model)

Emotional experience
  - The ability to be cognitively aware of its emotional state. It consists of cognitive awareness, physiological awareness, and subjective feelings.

Body-mind interactions
  - Emotions influence memory and memory retrieval, cognition and decision making.
  - Cognitive thoughts, which include concerns, goals, and motivations, can generate emotions.

[Picard 1997]
## The OCC (Ortony, Clore, Collins) Model

<table>
<thead>
<tr>
<th>Consequences of events</th>
<th>Consequences for others</th>
<th>Desirable for other</th>
<th>“pleased”</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Undesirable for other</td>
<td>“happy-for”</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Prospects relevant</td>
<td>“hope”</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Confirmed</td>
<td>“satisfaction”</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Disconfirmed</td>
<td>“relief”</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Prospects irrelevant</td>
<td>“joy”</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Attributed to self</td>
<td>“gratification”</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Attributed to others</td>
<td>“gratitude”</td>
<td></td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>Consequences for self</th>
<th>Prospects relevant</th>
<th>“displeased”</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Confirmed</td>
<td>“fears-confirmed”</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Disconfirmed</td>
<td>“disappointment”</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Prospects irrelevant</td>
<td>“distress”</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Attributed to Self</td>
<td>“remorse”</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Attributed to others</td>
<td>“anger”</td>
<td></td>
</tr>
</tbody>
</table>

### Actions of agents

<table>
<thead>
<tr>
<th>Consequences of events</th>
<th>Focusing on self agent</th>
<th>“approving”</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Focusing on other agent</td>
<td>“admiration”</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>“disapproving”</td>
<td></td>
</tr>
</tbody>
</table>

| Focusing on self agent | “disapproving” |         |
|                        | “shame”         |         |
|                        | “reproach”      |         |

### Aspects of Objects

<table>
<thead>
<tr>
<th>Consequences of events</th>
<th>Attraction</th>
<th>“liking”</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>“love”</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Attraction</th>
<th>“disliking”</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>“hate”</td>
</tr>
</tbody>
</table>
The Affective Reasoner by Clark Eliott demonstrates how modeling personalities of agents and their social relationships can interact with the generation of emotions.

Model of personality
- How events, acts, and objects are interpreted with respect to an individual agent’s goals, standards, and preferences.
- How an agent will act or feel in response to an emotional state.

Model of social relationships and their influences on emotions
- Friendship. An agent will tend to have similarity valenced emotions in response to the emotions of another agent.
- Animosity. An agent will tend to have oppositely valenced emotions in response to the emotions of another agent.
- Empathy. An agent will temporarily substitute the presumed goals, standards, and preferences of another agent for its own. It will then synthesize emotions based on these presumed goals, standards and preferences, in an effort to feel what it thinks the other agent would feel.
User: “Ruth is unhappy”
Agent: “I am sorry to hear that. You must feel sorry for her.”

(days later)

Agent: “How is Ruth?”
User: “Ruth is jealous of her rival.”
Agent: “How much is Ruth feeling this resentment?”
User: “Very much.”
Agent: “Perhaps you are worried about Ruth?”
...
# Emotion types used in the Affective Reasoner

<table>
<thead>
<tr>
<th>Group</th>
<th>Specification</th>
<th>Name and emotion type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Well-being</td>
<td>appraisal of a situation as an event</td>
<td>joy: pleased about an event</td>
</tr>
<tr>
<td></td>
<td></td>
<td>distress: displeased about an event</td>
</tr>
<tr>
<td>Fortunes-of-others</td>
<td>presumed value of situation as an event affecting another</td>
<td>happy-for: pleased about an event desirable for another happy-for: pleased about an event</td>
</tr>
<tr>
<td></td>
<td></td>
<td>gloating: pleased about an event undesirable for another</td>
</tr>
<tr>
<td></td>
<td></td>
<td>resentment: displeased about an event desirable for another</td>
</tr>
<tr>
<td></td>
<td></td>
<td>jealousy: resentment over a desirable mutually exclusive goal</td>
</tr>
<tr>
<td></td>
<td></td>
<td>envy: resentment over a desired non-exclusive goal</td>
</tr>
<tr>
<td></td>
<td></td>
<td>sorry-for: displeased about an event undesirable for another</td>
</tr>
<tr>
<td>Prospect-based</td>
<td>appraisal of a situation as a prospective event</td>
<td>hope: pleased about a prospective desirable event</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fear: displeased about a prospective undesirable event</td>
</tr>
<tr>
<td>Confirmation</td>
<td>appraisal of a situation as confirming or disconfirming an expectation</td>
<td>satisfaction: pleased about a confirmed desirable event</td>
</tr>
<tr>
<td></td>
<td></td>
<td>relief: pleased about a disconfirmed undesirable event</td>
</tr>
<tr>
<td></td>
<td></td>
<td>fears-confirmed: displeased about a confirmed undesirable event</td>
</tr>
<tr>
<td></td>
<td></td>
<td>disappointment: displeased about a disconfirmed desirable event</td>
</tr>
<tr>
<td>Attribution</td>
<td>appraisal of a situation as containing an attractive or unattractive object</td>
<td>pride: approving of one's own act</td>
</tr>
<tr>
<td></td>
<td></td>
<td>admiration: approving of another's act</td>
</tr>
<tr>
<td></td>
<td></td>
<td>shame: disapproving of one's own act</td>
</tr>
<tr>
<td></td>
<td></td>
<td>reproach: disapproving of another's act</td>
</tr>
<tr>
<td>Attraction</td>
<td>appraisal of a situation as containing</td>
<td>liking: finding an object appealing</td>
</tr>
<tr>
<td></td>
<td></td>
<td>disliking: finding an object unappealing</td>
</tr>
<tr>
<td>Well-being / attribution</td>
<td>compound emotions</td>
<td>gratitude: admiration + joy</td>
</tr>
<tr>
<td></td>
<td></td>
<td>anger: reproach + distress</td>
</tr>
<tr>
<td></td>
<td></td>
<td>gratification: pride + joy</td>
</tr>
<tr>
<td></td>
<td></td>
<td>remorse: shame+distress</td>
</tr>
<tr>
<td>Attraction / attribution</td>
<td>compound emotion extensions</td>
<td>love: admiration + liking</td>
</tr>
<tr>
<td></td>
<td></td>
<td>hate: reproach + disliking</td>
</tr>
</tbody>
</table>

[Picard 1997]
If you come to know that animal or object or situation X causes fear, you will have two ways of behaving toward X. The first way is innate; you do not control it. Moreover, it is not specific to X; a large number of creatures, objects, and circumstances can cause the response. The second way is based on your own experience and is specific to X. Knowing about X allows you to think ahead and predict the probability of its being present in a given environment, so that you can avoid X, preemptively, rather than just have to react to its presence in an emergency. ...

Primary emotions depend on limbic system circuitry, the amygdala and anterior cingulate being the primary players. After an appropriate stimulus activates the amygdala, a number of responses ensue: internal responses, muscular responses, visceral responses, and responses to neurotransmitter nuclei and hypothalamus.

Secondary emotions utilize the machinery of primary emotions. The stimulus may still be processed directly via the amygdala but is now also analyzed in the thought process, and may activate frontal cortices (VM). VM acts via the amygdala. (p. 133-137)

H: Hypothalamus,
VMF: ventromedial prefrontal cortex

[Damasio 1994]
Affective Computing

Cognitive processing

High level

Inference and decision making

Low level

Pattern recognition and synthesis

Representations / signals

Emotional states

Representations / signals

[Picard 1997]
<table>
<thead>
<tr>
<th>Applications of affective computing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Affective mirror</td>
</tr>
<tr>
<td>Beyond emoticons</td>
</tr>
<tr>
<td>Text to speech</td>
</tr>
<tr>
<td>Helping autistic people</td>
</tr>
<tr>
<td>Consumer feedback</td>
</tr>
<tr>
<td>Points for courage</td>
</tr>
<tr>
<td>Emotions in learning</td>
</tr>
<tr>
<td>“No pain, no gain”</td>
</tr>
<tr>
<td>Classroom barometer</td>
</tr>
<tr>
<td>Emotions on the virtual stage</td>
</tr>
<tr>
<td>Music: listening to what you like</td>
</tr>
<tr>
<td>“Fast forward to the interesting part”</td>
</tr>
<tr>
<td>Agents that learn your preferences</td>
</tr>
<tr>
<td>Learning when to interrupt</td>
</tr>
<tr>
<td>Small talk</td>
</tr>
<tr>
<td>Animated agent faces</td>
</tr>
<tr>
<td>The audience performance</td>
</tr>
<tr>
<td>Film/video</td>
</tr>
<tr>
<td>Sensitive toys</td>
</tr>
</tbody>
</table>

[Picard 1997]
Dynamical system model for primary and secondary emotion

Valence

$PAD(x_t, y_t, z_t) = \left( \frac{1}{2}(x_t + y_t), |x_t| + z_t, d(t) \right)$

$x_t$ : valence of emotions
$y_t$ : valence of mood
$z_t$ : boredom

Awareness likelihood

[Becker 2007]
FAtiMA

Reactive layer

Goals

Intentions

Focus and Coping

Deliberative layer

Intention generation

Intention selection

Focus and Coping

Emotion focused coping

Autobiographic memory

Emotional state

Appraisal

Effectors

Sensors

Perception

Action

Knowledge base

Properties

Relations

Appraisal

Action tendencies

[Lim 2012]
FAtiMA-PSI

Reactive layer

Goals

Intention generation

Intentions

Intentions selection

Focus and Coping

Deliberative layer

Energy

Integrity

Affiliation

Certainty

Competence

Personality

thresholds

Autobiographic
memory

Emotional
state

Thresholds

Properties

Relations

Knowledge base

Perception

Appraisal

Action

PSI

Effectors

Modulating parameters

[Lim 2012]
FearNot!

[Aylett et al 2005]
Understanding of another person’s wrong belief requires explicit representation of the wrongness of this person’s belief in relation to one’s own knowledge. ... [Wimmer 1983]

[Theory of mind: an ability to] impute[] mental states to themselves and others. A system of inferences of this kind is properly viewed of as a theory because such states are not directly observable, and the system can be used to make predictions about the behavior of others. As to the mental states ... for example, purpose, or intention, as well as knowledge, belief, thinking, doubt, guessing, pretending, liking, and so forth. ... [Premack 1978]
Theory of mind – false belief test

Maxi is a boy who likes chocolate.

[Wimmer 1983]
Theory of mind – false belief test

Mother bought a chocolate.

Where should I put the chocolate?

In the blue cupboard.

Mother bought a chocolate.

[Wimmer 1983]
Theory of mind – false belief test

[Wimmer 1983]
Theory of mind – false belief test

[Image of a boy and girl looking at three boxes, one of which contains a chocolate bar]
Maxi leaves for the playground.

[Wimmer 1983]
Theory of mind – false belief test

[Wimmer 1983]
Theory of mind – false belief test

[Wimmer 1983]
Theory of mind – false belief test

Mother takes chocolate out of the blue cupboard.

[Wimmer 1983]
Mother grates a bit into the dough.
Theory of mind – false belief test

Mother does not put it back to the blue but into the green cupboard.

[Wimmer 1983]
Theory of mind – false belief test
Theory of mind – false belief test

[Wimmer 1983]
There comes Maxi back from the playground, hungry, and he wants some chocolate.

[BELIEF-question] Where will Maxi look for the chocolate?
There comes Grandpa.
Maxi says: “Please help me get the chocolate”
Grandpa asks: “Which cupboard?”

[UTTERANCE-question] Where will Maxi say the chocolate is?
[REALITY-question] [MEMORY-question]

[Wimmer 1983]
His big brother comes into the kitchen. He asks Maxi where the chocolate is. Maxi wants to tell him something wrong.

[UTTERANCE-question] Where will Maxi say the chocolate is? [REALITY-question] [MEMORY-question]
The computational architecture of theory of mind

**Intention Detector (ID)**
Interprets motion stimuli (stimuli with self-propulsion and direction) in terms of the mental states of goal and desire.

**Eye Direction Detector (EDD)**
Detects the presence of eye-like stimuli, detects the direction of eyes, and interprets gaze as seeing (attribution of perceptual states).

**Shared Attention Mechanism (SAM)**
Allows to build triadic representations: relations between an agent, the self, and a third object.

**Theory of Mind Mechanism (ToMM)**
Represents the full range of mental states and allows one to make sense of an agent's current behavior and predict an agent's future action.

**Visual, auditory and tactile cues**

**[Agent likes chocolate] ...**

**Visual cues**

**[Agent sees mother's-first-act]**
**[Agent not-see mother's-second-act]**

**Joint attention:**
Eye direction in terms of intention

**Referential opacity:**
Separation of mental from physical

**[I see [Agent sees chocolate-in-blue-cupboard]]**
**[I see [Agent not-see chocolate-in-green-cupboard]]**

**[Agent believes chocolate-in-blue-cupboard]**
**[Agent falsely-believes chocolate-in-blue-cupboard]**

[Baron-Cohen 1995]
Theory of mind

Perceptual processes

Expression raiser
“This is a banana.”

Central cognitive systems

decoupler

Manipulator
“I pretend
“This banana is a telephone.”

This is a telephone

This is a telephone

References

References