Robot Emotions

Emotions of Living Creatures

- motivation system for complex organisms

- *determine the behavioral reaction to environmental (often social) and internal events of major significance for the needs and goals of a creature (Plutchik, 1991; Izard, 1977).*

Emotions of Living Creatures

- Positive emotions
  - elicited by events that satisfy some motive, enhance one's power of survival, or demonstrate the successful exercise of one's capabilities.
  - often signal that activity toward the goal can terminate, or that resources can be freed for other exploits.

- Negative emotions:
  - result from painful sensations or threatening situations.
  - motivate actions to set things right or to prevent unpleasant things from occurring.

Frijda, 1994

Theory of Basic Emotions

- There exists a set of basic or primary emotions that have been selected for through evolution.
  - anger, disgust, fear, joy, sorrow, and surprise (Ekman and Oster, 1982)

- Each basic emotion serves a particular function (biological or social) to prepare and motivate a creature to respond in adaptive ways.

Theory of Appraisal

- Emotion has evolved as a relevance-detection and response-preparation system.

- *Emotion is an appraisal system that assesses the perceived antecedent conditions with respect to the organism's well-being, its plans, and its goals*

  (Levenson, 1994; Izard, 1994; Frijda, 1994c; Lazarus, 1994).

- People *affectively* appraise events with respect to novelty, intrinsic pleasantness, goal/need significance, coping, and norm/self compatibility.
- These appraisals, along with other factors such as pain, hormone levels, drives, etc., evoke a particular emotion.

Scherer 1994

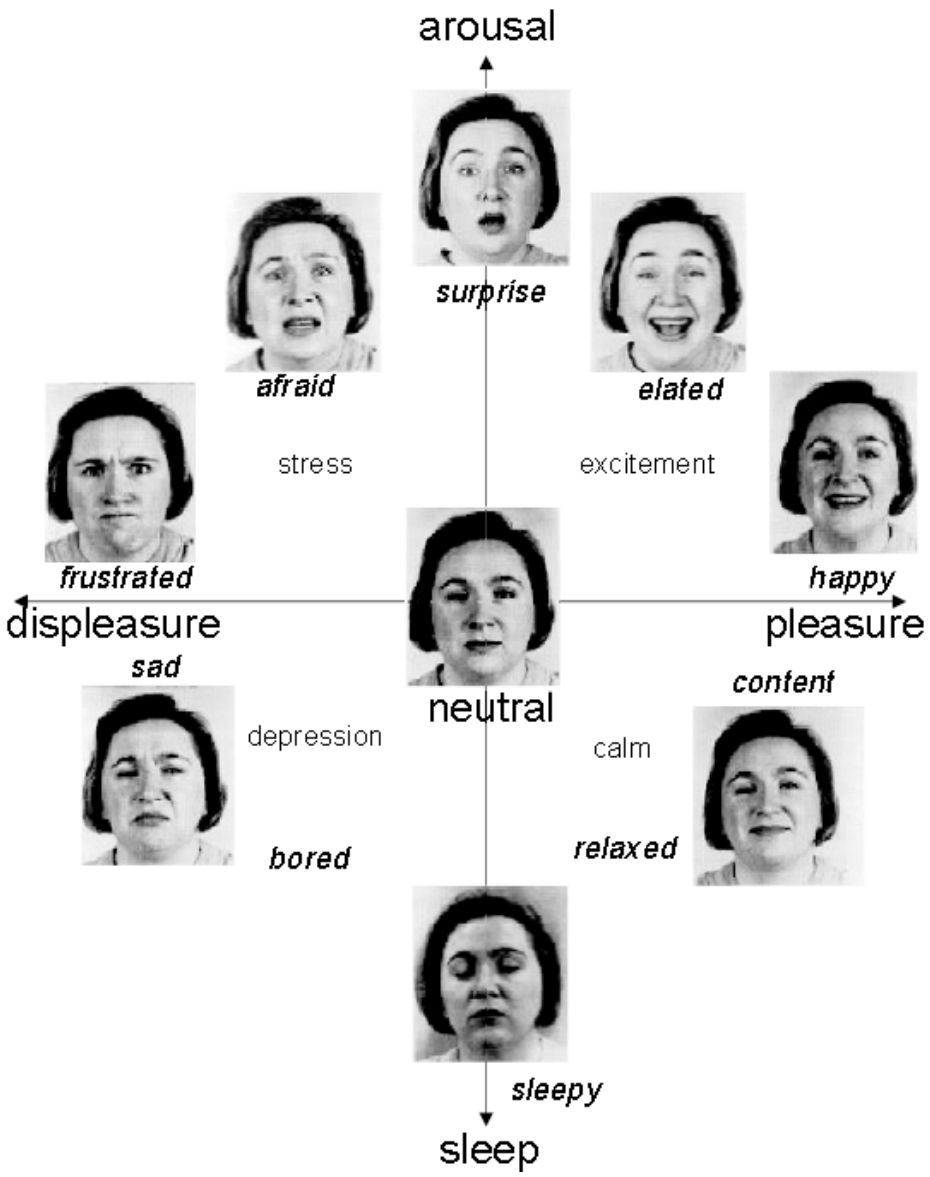
Definitions

- *Affect (n.)* – feeling or emotion
- *Drive (n.)* - an internal state of tension that motivates an organism to engage in activities that should reduce this tension
- *Valence (n.)* – the intrinsic attractiveness (positive valence) or averseness (negative valence) of an event, object, or situation

Theory of Componential Emotions

- Views emotion based on where they are located along different continuous dimensions (e.g., arousal and valence)

- Supporting argument:
  - facial expressions have a systematic, coherent, and meaningful structure that can be mapped to affective dimensions (Russell, 1997; Lazarus, 1991; Plutchik, 1984; Smith, 1989; Woodworth, 1938).


arousal
surprise
afraid
elated
stress
excitement
frustrated
happy
displeasure
pleasure
sad
content
neutral
depression
calm
bored
relaxed
sleepy
sleep

Effect of Affect

- physiological changes (e.g., modulating arousal level)
- adjustments in subjective experience
- elicitation of behavioral response (e.g., approach, attack, escape)
- displaying expression

- Together these factors represent a generalized solution for coping with the demands of the world

Communicating Emotion

- Emotion shown through voice, face, gesture, and posture

- Emotional signals serve to
  - Communicate our emotional state to others
  - Influence the behavior of others

Levenson 1994

- projection of how the others will react to these different possible courses of action largely determines the creature's behavioral choice

Scherer 1994

Behavioral Homeostasis

- Emotions establish a desired relation between the organism and the environment that pulls the creature toward certain stimuli and events and pushes it away from others.

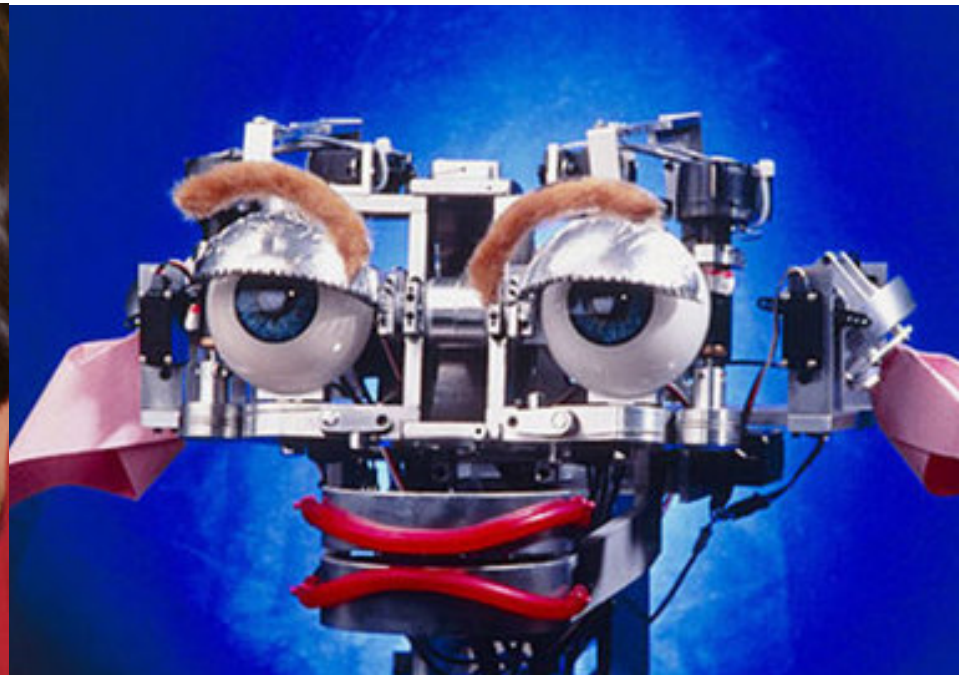
  Plutchik 1991

- Examples in social behavior:
  - proximity seeking, social avoidance, chasing off offenders

Emotion and Sociable Humanoid Robots

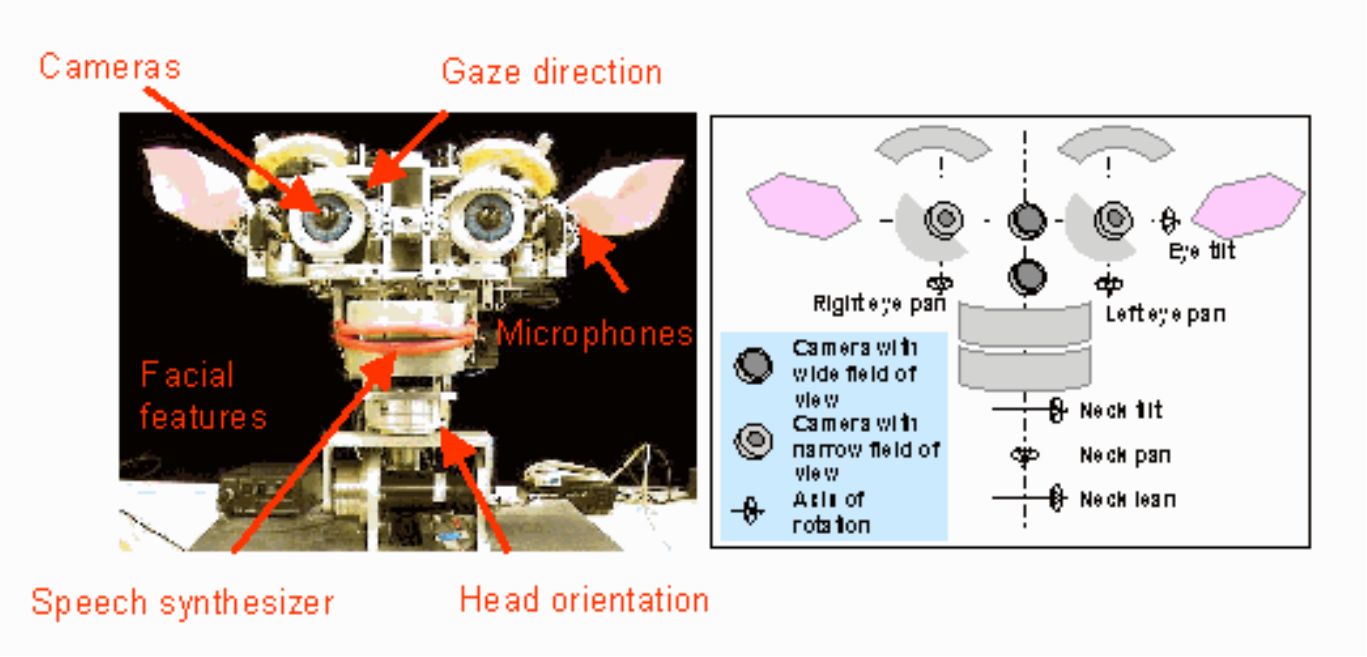
Cynthia Breazeal

Int. J. Human-Computer Studies, 2003.



Kismet

- 21 DOF
  - 3 each in head and eyes, 15 for facial features



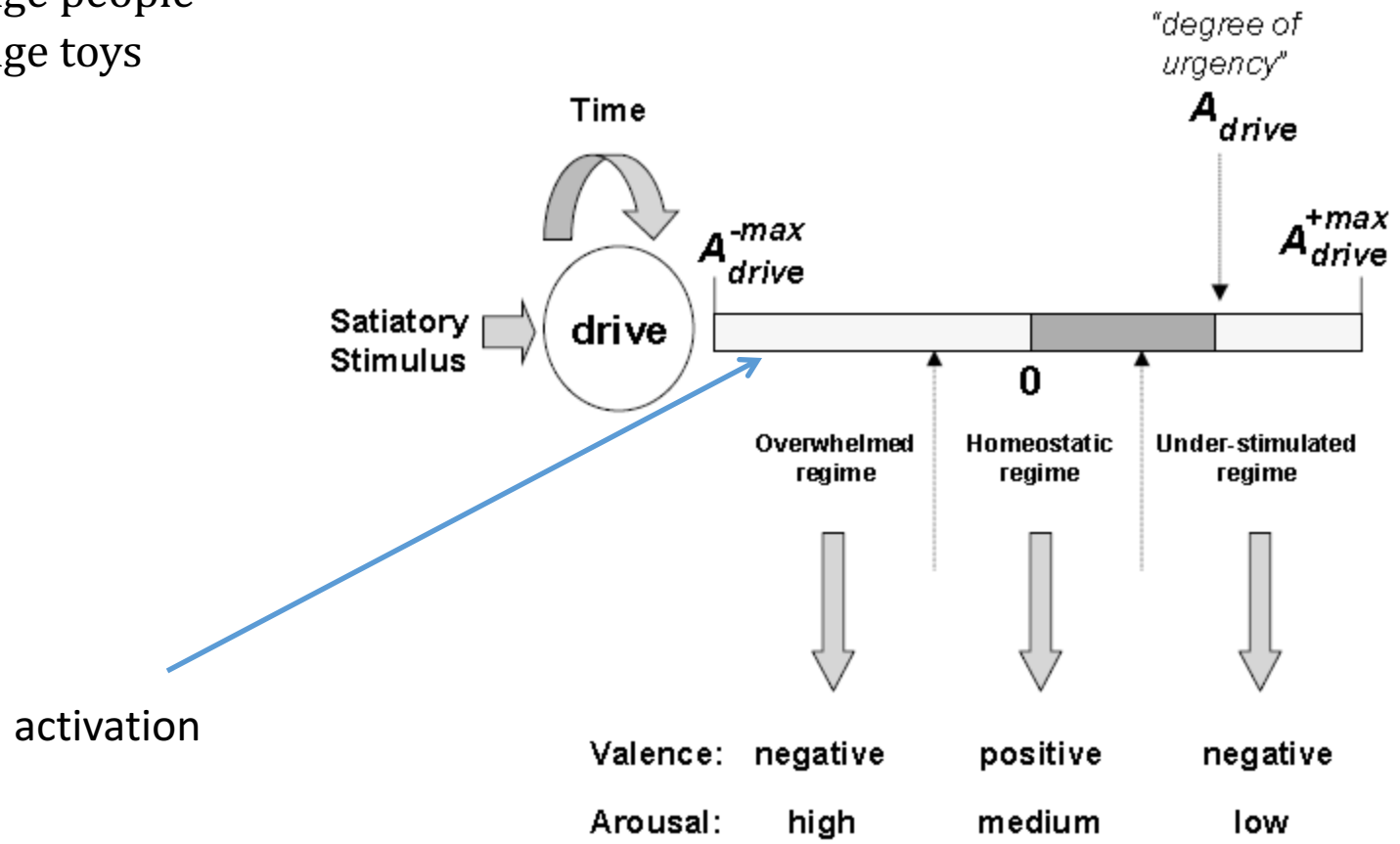
Kismet

- Goals:
  - Socially engage people
  - Eventually learn from people

- Behavior controlled by emotions and drives

Drives

- Kismet has three drives:
  - Engage people
  - Engage toys
  - Rest



Drives

- *Homeostatic regime* – encountering the satiatory stimulus and that stimulus is of appropriate intensity.
- *Understimulated regime* – absence of the satiatory stimulus (or if the intensity is too low)
- *Overwhelmed regime* - satiatory stimulus is too intense (e.g., moving too close or too fast)
- To remain in balance, it is not sufficient that the satiatory stimulus be present; it must also be of an appropriate intensity.

Drives

- Influence behavior selection by preferentially passing activation to some behaviors over others (i.e., those that serve to satiate the drive).

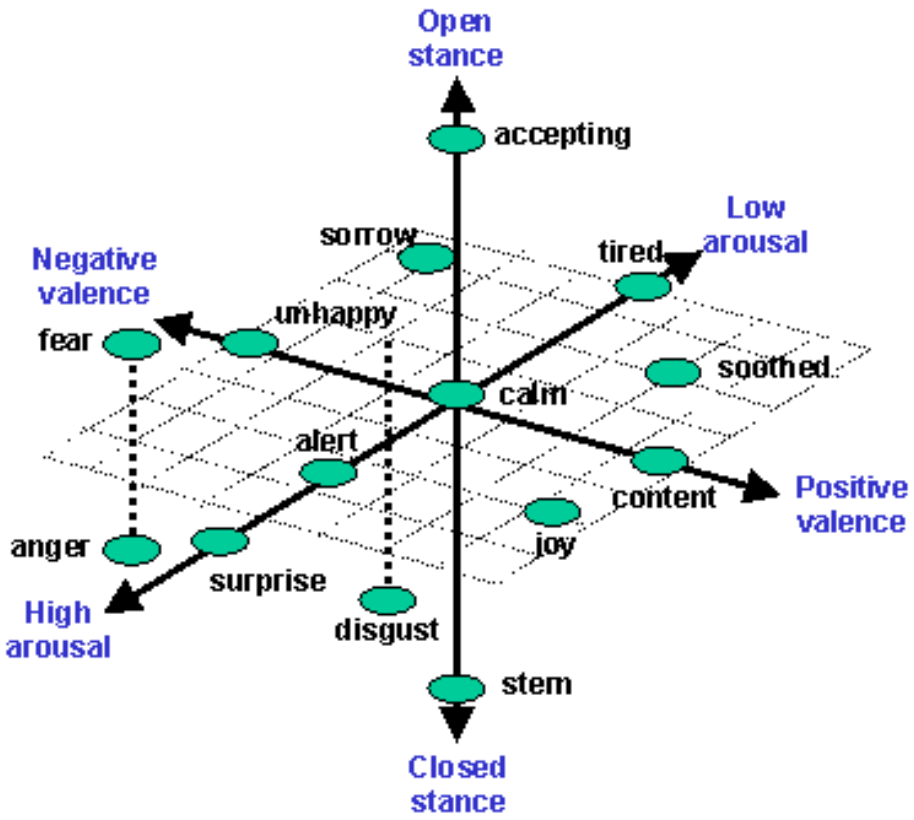
- Provide a functional context (i.e., the goal, namely which "need" the robot is actively trying to address) that organizes behavior and perception.

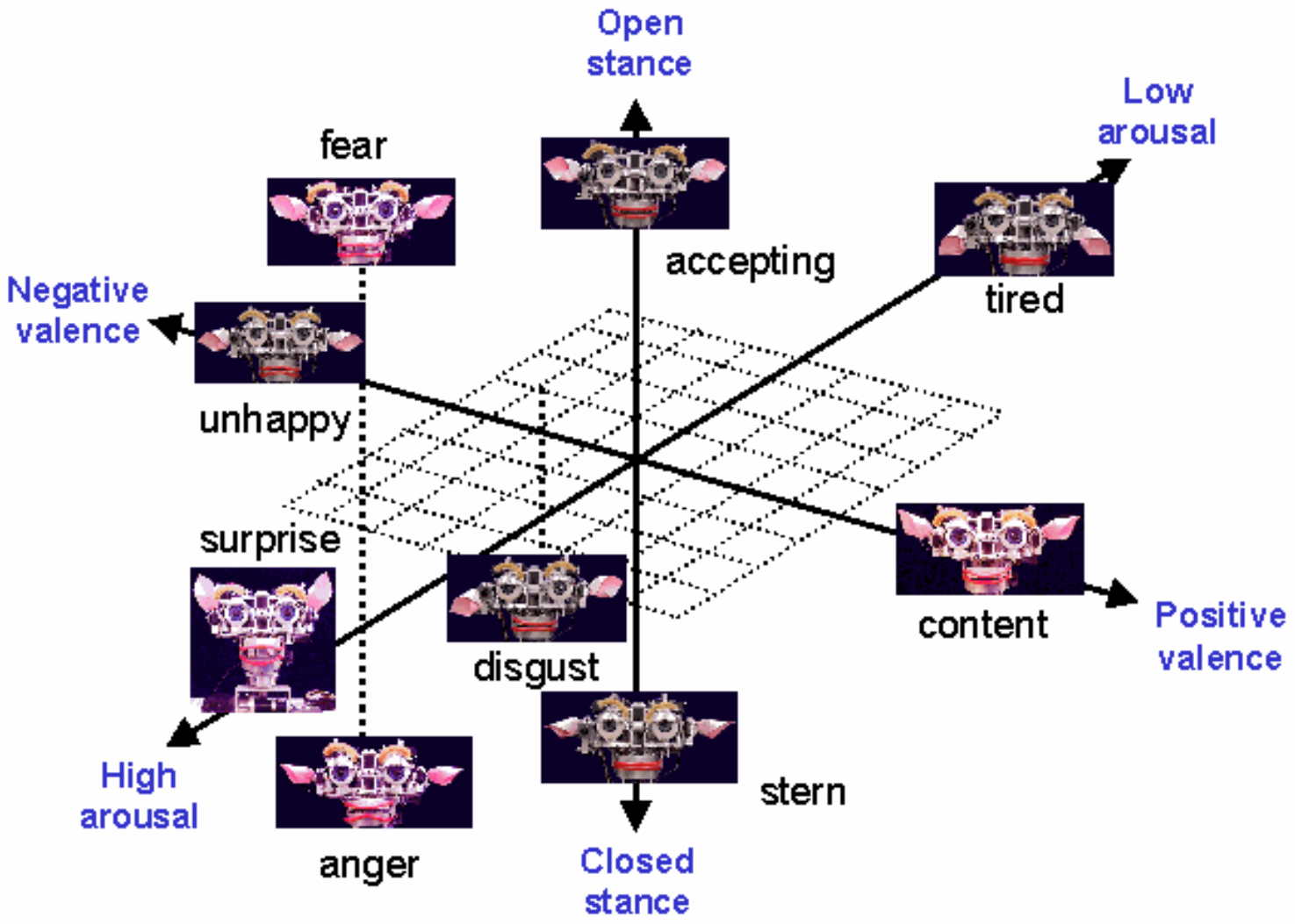
- Influence the robot's affective state by directly contributing to valence and arousal measures.

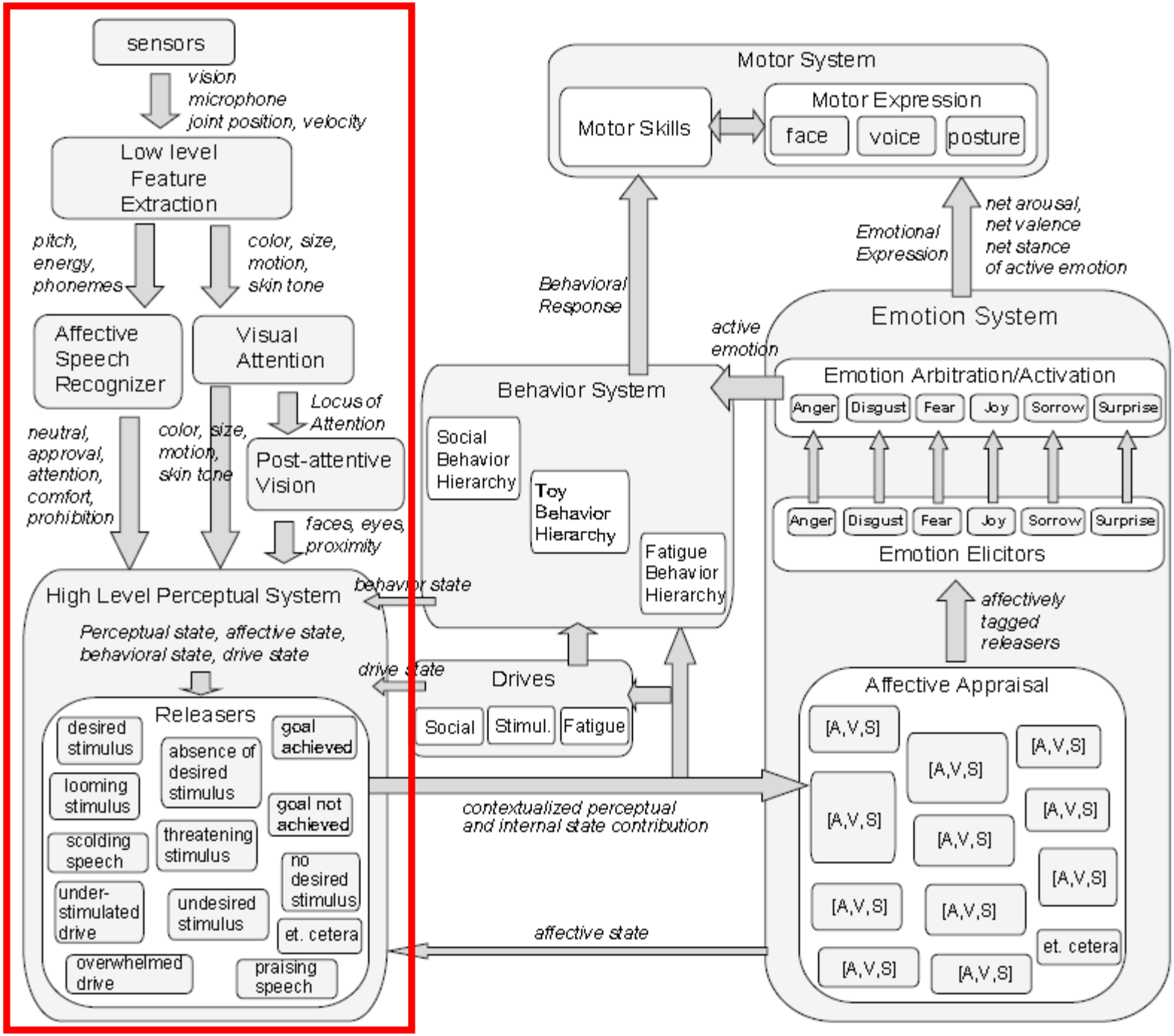
Emotions

- Triggered by various events that are evaluated as being of significance to the ‘‘well being’’ of the robot.
- Once triggered, each emotion serves to establish a desired relation between the robot and its environment.
- The emotion system contributes to satiating drives by bringing the robot into contact with things that benefit it and to avoid those things that are undesirable or potentially harmful.

Emotions


Open stance
accepting
Low arousal
sorrow
tired
Negative valence
unhappy
fear
soothed
calm
alert
content
Positive valence
joy
anger
surprise
High arousal
disgust
stem
Closed stance


Open stance
Low arousal
fear
accepting
Negative valence
tired
unhappy
surprise
content
Positive valence
disgust
High arousal
stern
anger
Closed stance


sensors
vision
microphone
joint position, velocity
Low level Feature Extraction
pitch, energy, phonemes
color, size, motion, skin tone
Affective Speech Recognizer
Visual Attention
Locus of Attention
Post-attentive Vision
faces, eyes, proximity
neutral, approval, attention, comfort, prohibition
color, size, motion, skin tone
High Level Perceptual System
Perceptual state, affective state, behavioral state, drive state
behavior state
drive state
Releasers
desired stimulus
absence of desired stimulus
goal achieved
looming stimulus
goal not achieved
scolding speech
threatening stimulus
no desired stimulus
under-stimulated drive
undesired stimulus
et. cetera
overwhelmed drive
praising speech
Motor System
Motor Skills
Motor Expression
face
voice
posture
Behavioral Response
Emotional Expression
net arousal, net valence net stance of active emotion
Behavior System
Social Behavior Hierarchy
Toy Behavior Hierarchy
Fatigue Behavior Hierarchy
active emotion
Drives
Social
Stimul.
Fatigue
Emotion System
Emotion Arbitration/Activation
Anger
Disgust
Fear
Joy
Sorrow
Surprise
Emotion Elicitors
affectively tagged releasers
Affective Appraisal
[A,V,S]
contextualized perceptual and internal state contribution
affective state

Releasers

- combine lower-level perceptual features with measures of the robot's internal state into behaviorally significant perceptual categories
- attributes
  - presence or absence of a stimulus (and for how long)
  - its nature (e.g., toy-related or person-related)
  - the quality of the stimulus (e.g., the intensity is too low, too high, or just right)
  - whether it is desired or not (e.g., it relates to the active goals or motivations)

Example

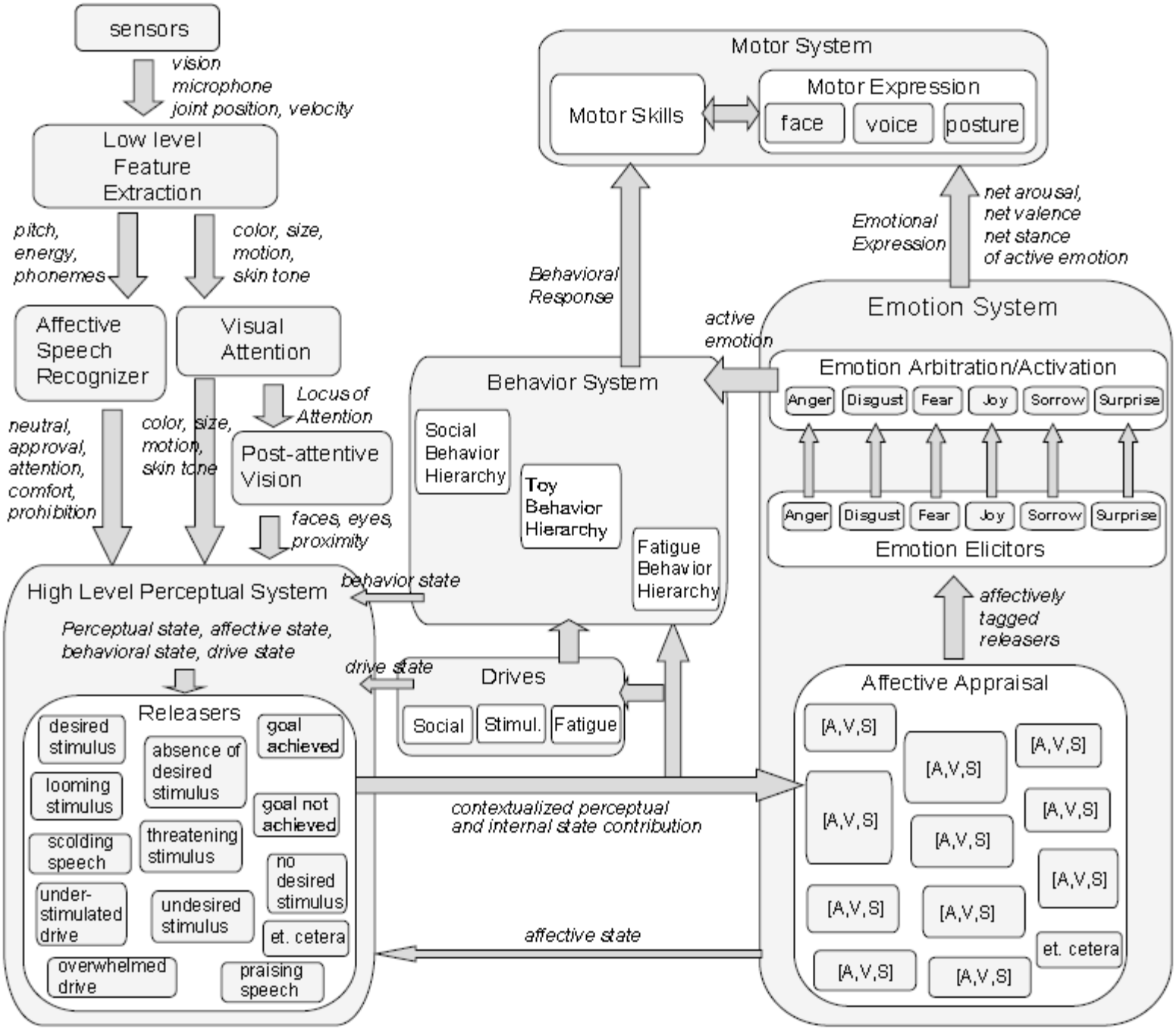
- Toy percept: color, size, motion and proximity

- If the stimulation drive is active and the toy is neither too fast nor too close to the robot, then the *desired-toy releaser* is active.
- If the social drive is active, then the *undesired-toy releaser* is active.
- If the toy has an aggressive motion (i.e., too close and moving too fast), then the *threatening-toy releaser* is active.

Releaser

- Each releaser calculates an activation level
- If activation > threshold, the perceptual system represents it in terms of 3 values [-1250,1250]:
  - Arousal (how arousing)
  - Valence (how desirable)
  - Stance (how approachable)

- Threatening toy: A = 1200, V = -1000, S = -1000


sensors
vision
microphone
joint position, velocity
Low level Feature Extraction
pitch, energy, phonemes
color, size, motion, skin tone
Affective Speech Recognizer
Visual Attention
Locus of Attention
Post-attentive Vision
neutral, approval, attention, comfort, prohibition
color, size, motion, skin tone
faces, eyes, proximity
High Level Perceptual System
Perceptual state, affective state, behavioral state, drive state
behavior state
drive state
Releasers
desired stimulus
absence of desired stimulus
goal achieved
looming stimulus
goal not achieved
scolding speech
threatening stimulus
no desired stimulus
under-stimulated drive
undesired stimulus
et. cetera
overwhelmed drive
praising speech
Motor System
Motor Skills
Motor Expression
face
voice
posture
Behavioral Response
Emotional Expression
net arousal, net valence net stance of active emotion
Emotion System
active emotion
Behavior System
Social Behavior Hierarchy
Toy Behavior Hierarchy
Fatigue Behavior Hierarchy
Drives
Social
Stimul.
Fatigue
Emotion Arbitration/Activation
Anger
Disgust
Fear
Joy
Sorrow
Surprise
Emotion Elicitors
affectively tagged releasers
Affective Appraisal
[A,V,S]
contextualized perceptual and internal state contribution
affective state

Emotion Elicitor

- Calculates activation level for emotions
- Determines how each releaser contributes to a given emotion

- Each emotion type has an associated [A,V,S] profile

Emotion Activation Level

- Activation level: $[0, A_{max}]$

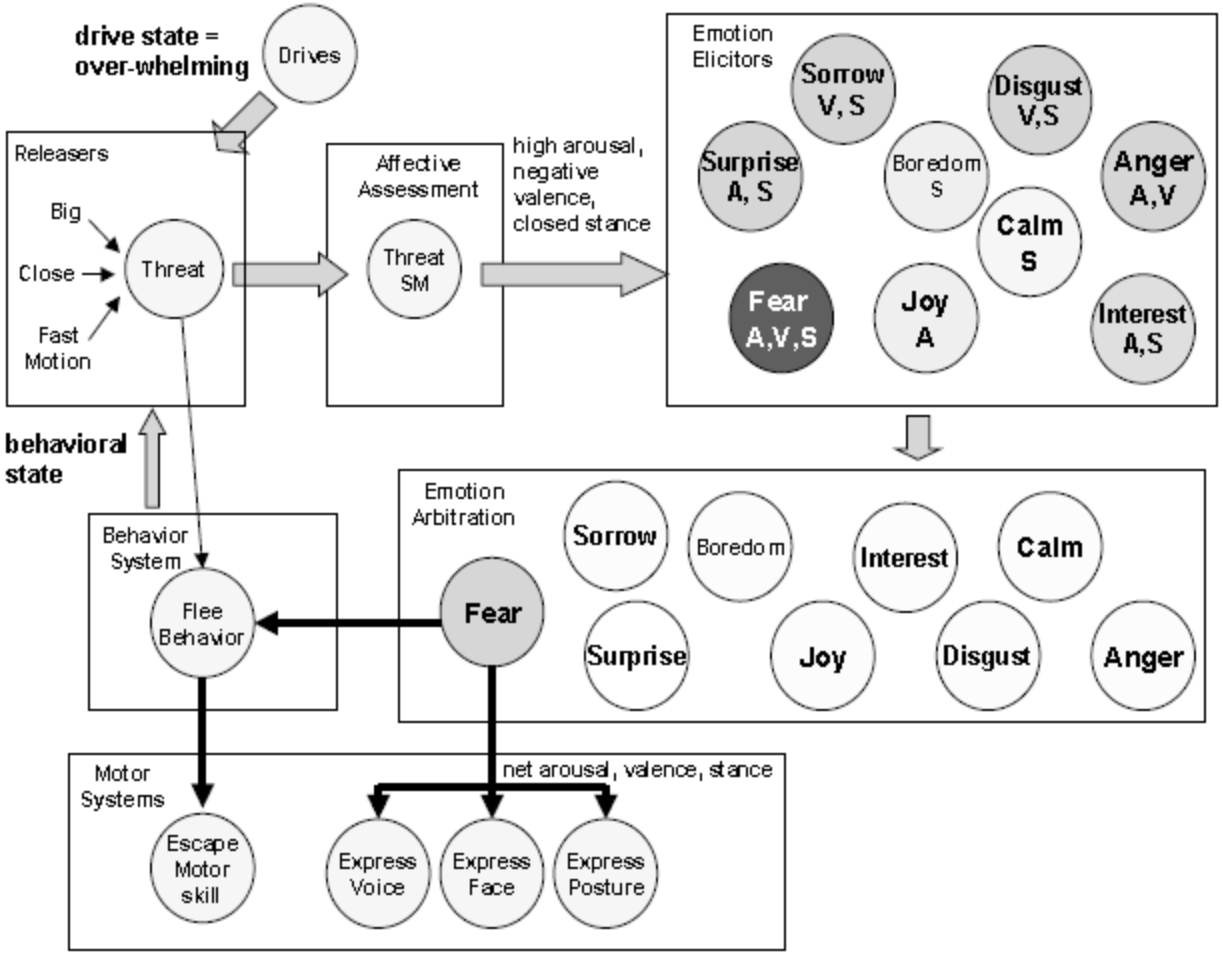
$$A = E+B+P-d$$

- E = elicitor activation level
- B = offset to make it easier to trigger some emotions
- P = persistence bias that maintains emotion for some time once its active
- d = decay term that returns emotion to its bias value once active

Emotion Activation

- The emotion with the highest activation is expressed
- Two activation thresholds:
  - Facial expression (strength relative to activation)
  - Behavior

Emotion Arbitration



Summary of the antecedents and behavioral responses that comprise Kismet's emotive responses

Antecedent conditions	Emotion	Behavior
Delay, difficulty in achieving goal of adaptive behavior	Anger, frustration	Complain
Presence of an undesired stimulus	Disgust	Withdraw
Presence of a threatening, overwhelming stimulus	Fear, distress	Escape
Prolonged presence of a desired stimulus	Calm	Engage
Success in achieving goal of active behavior, or praise	Joy	Display pleasure
Prolonged absence of a desired stimulus, or prohibition	Sorrow	Display sorrow
A sudden, dose stimulus	Surprise	Startle response
Appearance of a desired stimulus	Interest	Orient
Need of an absent and desired stimulus	Boredom	Seek

Image Evaluation Confusion Matrix

	Accepting	*Anger*	*Bored*	*Disgust*	*Fear*	*Joy*	*Interest*	*Sorrow*	*Stern*	*Surprised*	*% Correct*
Anger	5.9	76.5	0	0	5.9	11.7	0	0	0	0	76.5
Disgust	0	17.6	0	70.6	5.9	0	0	0	5.9	0	70.6
Fear	5.9	5.9	0	0	47.1	17.6	5.9	0	0	17.6	47.1
Joy	11.7	0	5.9	0	0	82.4	0	0	0	0	82.4
Sorrow	0	5.9	0	0	11.7	0	0	83.4	0	0	83.4
Stern	7.7	15.4	0	7.7	0	0	0	15.4	53.8	0	53.8
Surprise	0	0	0	0	0	17.6	0	0	0	82.4	82.4

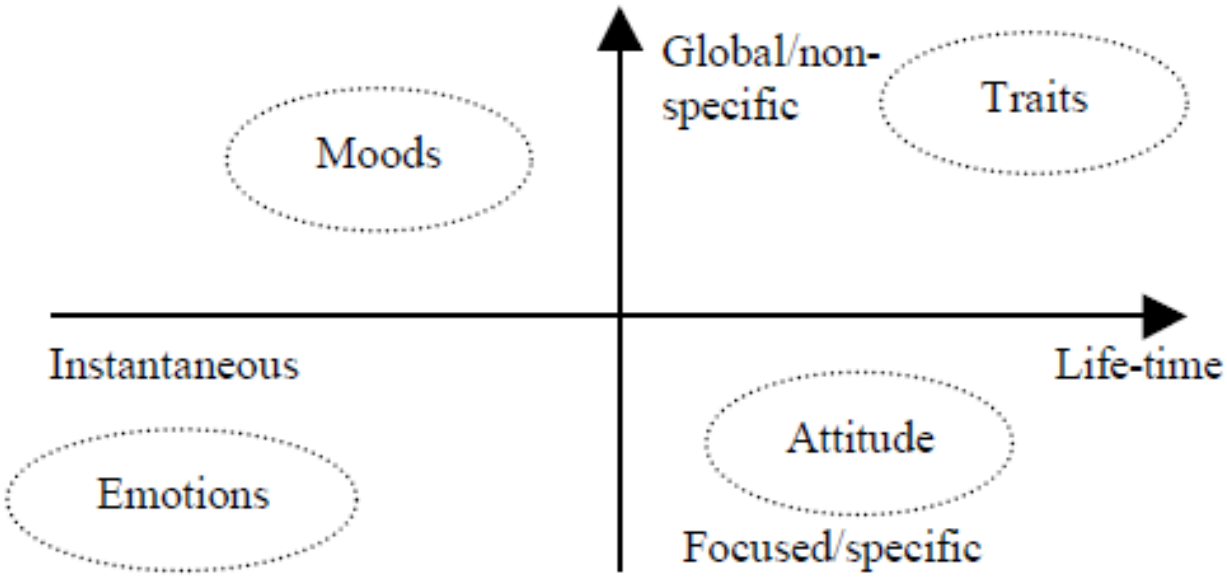
Video Evaluation Confusion Matrix

	Anger	*Disgust*	*Fear*	*Joy*	*Interest*	*Sorrow*	*Surprise*	*% Correct*
Anger	86	0	0	14	0	0	0	86
Disgust	0	86	0	0	0	14	0	86
Fear	0	0	86	0	0	0	14	86
Joy	0	0	0	57	28	0	15	57
Interest	0	0	0	0	71	0	29	71
Sorrow	14	0	0	0	0	86	0	86
Surprise	0	0	29	0	0	0	71	71

Summary

Other factors to consider

- Moods
  - typically of lower intensity than emotions and have fairly low variance over the course of a single day
- Attitude
  - (according to one view:) amalgamation of emotions experienced with a particular person or thing, reflecting one's relationship with that person over time
- Personality Traits
  - mainly inherited or imprinted by early experience, and can be assumed to be fairly constant


Global/non-specific
Traits
Moods
Instantaneous
Life-time
Attitude
Emotions
Focused/specific

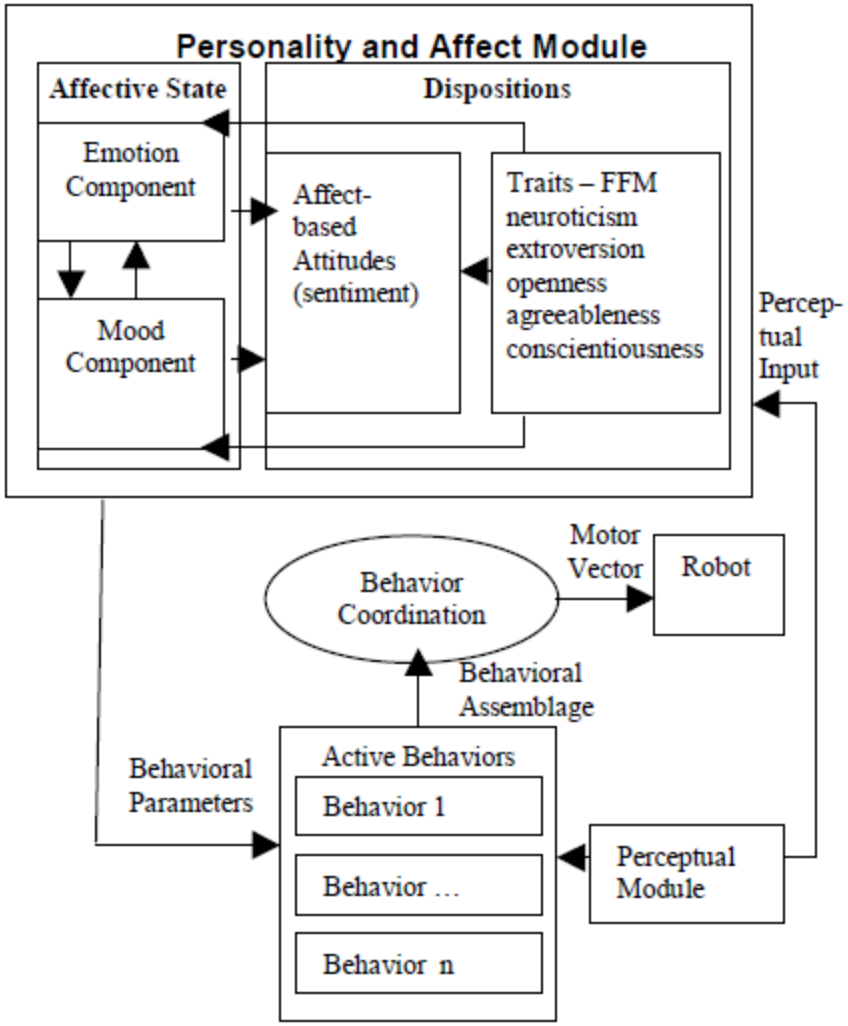
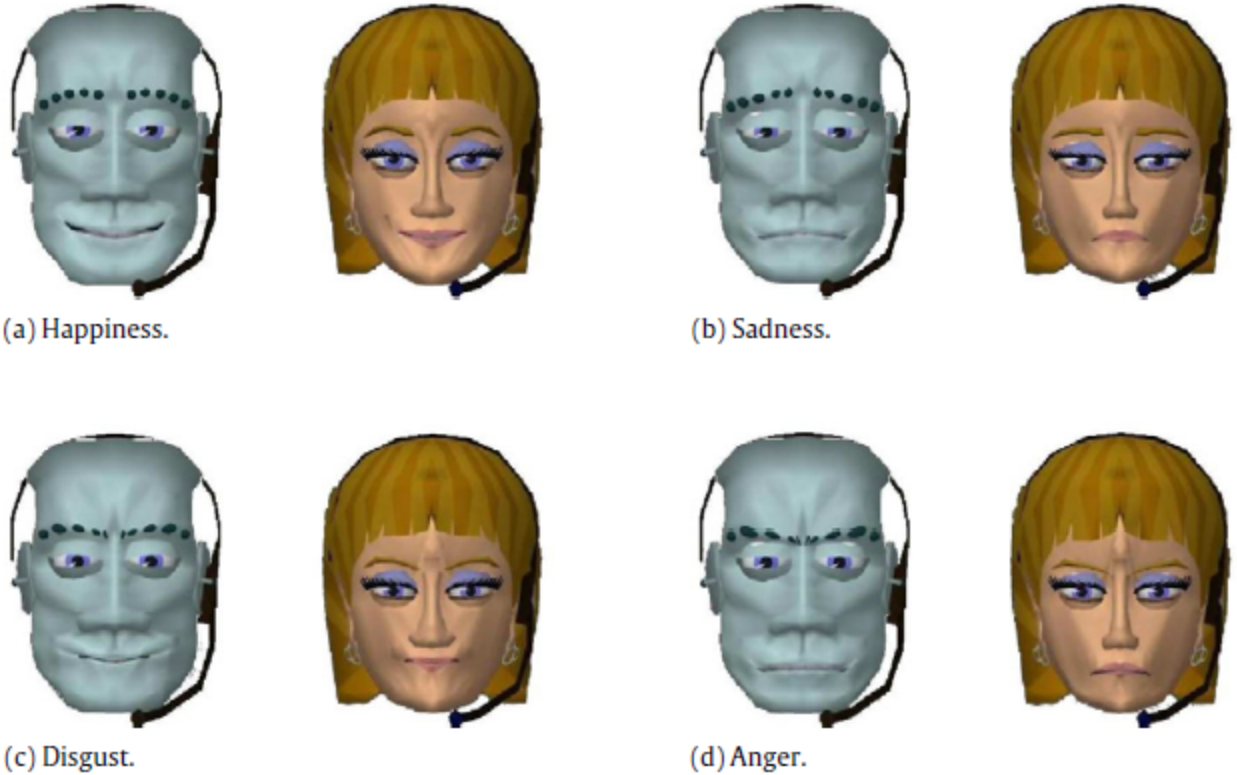


Figure 1: Integrated Model of Personality and Affect (TAME)

The Roboceptionist Project



(a) Happiness. (b) Sadness.

(c) Disgust. (d) Anger.

Fig. 3. Four emotional expressions, shown on both Tank and Valerie's faces.

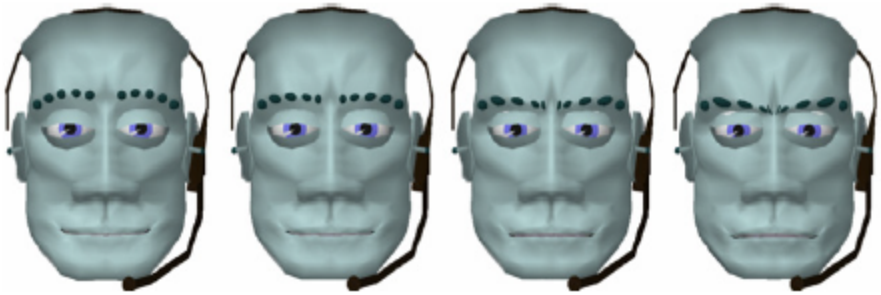
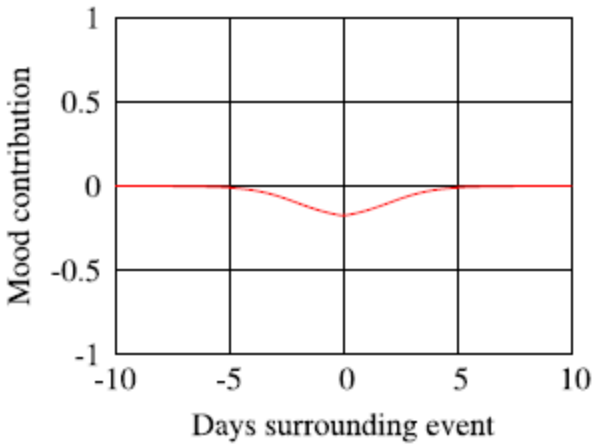


Fig. 4. Interpolation of emotional expressions between neutral (far left) and very angry (far right).

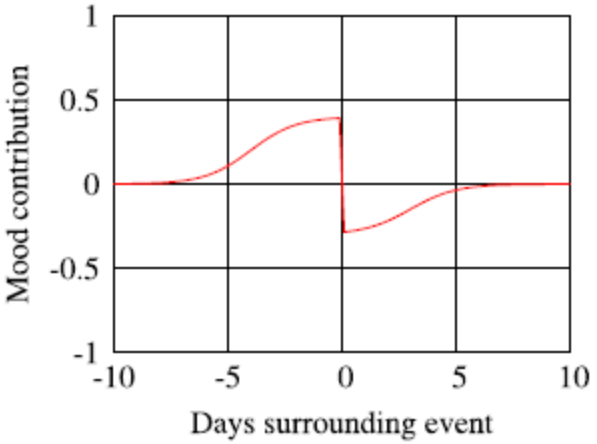
User	Good afternoon
Tank	Welcome to Newell Simon. [happy 0.1]
User	I am a new student here
Tank	Do you enjoy your job?
User	Yes
Tank	That's good. I like my new job here.
User	What's your job
Tank	I can do all sorts of tasks. I can give directions. I can help you find a room or a faculty or staff member in the School of Computer Science. I know the current weather for all the major cities of the world.
User	Could you please tell me where is Dr. [name]?
Tank	[name] is in [room number] Newell Simon Hall. Would you like directions?
User	Yes, please
Tank	You are already on the third floor of Newell Simon Hall. Go across the atrium... [further directions]
User	Do you know where I can find a map
Tank	Sorry. I don't know who or what that is
User	ok
Tank	[nods]
User	thanks
Tank	You're welcome. [happy 0.4]
User	Have a nice day
Tank	Glad to have helped. Have a pleasant day.
User	bye
Tank	Thanks for visiting Newell Simon. Come back and see me again
User	sure :)
Tank	I didn't understand that. It might help if you rephrase that. Or, ask me something simple. [disgusted 0.5]

User	idiot
Tank	How rude. I am not used to that type of abuse. [angry 0.5]
User	lol
Tank	I'm glad you are amused. [happy 0.3]
User	marion
Tank	I've told you people: DON'T CALL ME MARION! [angry 0.8]
User	lol
Tank	I'm glad you are amused. [happy 0.3]

Effect of life events on mood



(a) Impending dentist visit ($m = -0.2$).



(b) Blind date that went badly (anticipated $m = 0.4$, actual $m = -0.3$).

$$c = m - \frac{m}{1 + e^{-|d|+|10m|}}.$$

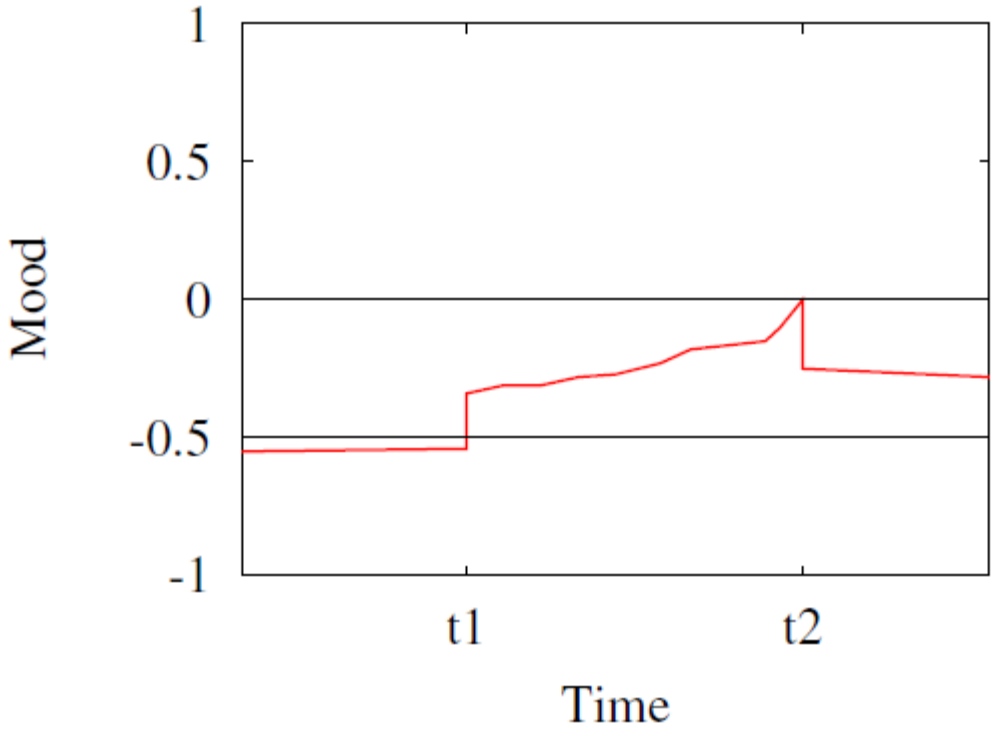


Fig. 6. An example of how attitude and interactions influence mood. Here, the baseline mood is -0.5, and a previous interaction has lowered the mood even further. A well-liked person ($A_m = 0.4$) approaches to begin interaction at time t1, which immediately improves the robot's mood. The person interacts from t1 to t2, largely trying to cheer up the robot, improving its mood considerably. When the person leaves, the mood drops once again, but remains at a higher level than before. After the interaction, the mood begins to decay toward the baseline.

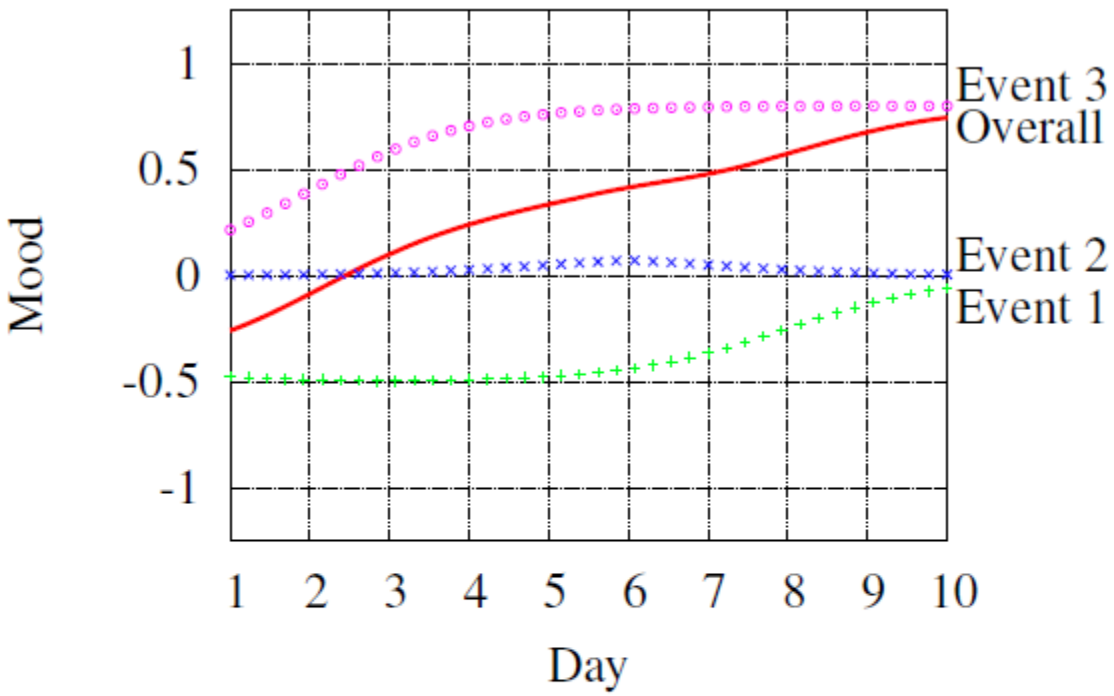


Fig. 4. Three hypothetical events and their contributions to the overall mood across ten days. Event 1, scheduled minor surgery, occurs on day 3 with contribution -0.5, event 2, an evening with friends, occurs on day 6 with contribution 0.1, and event 3, the start of a vacation, occurs on day 10 with contribution 0.8.

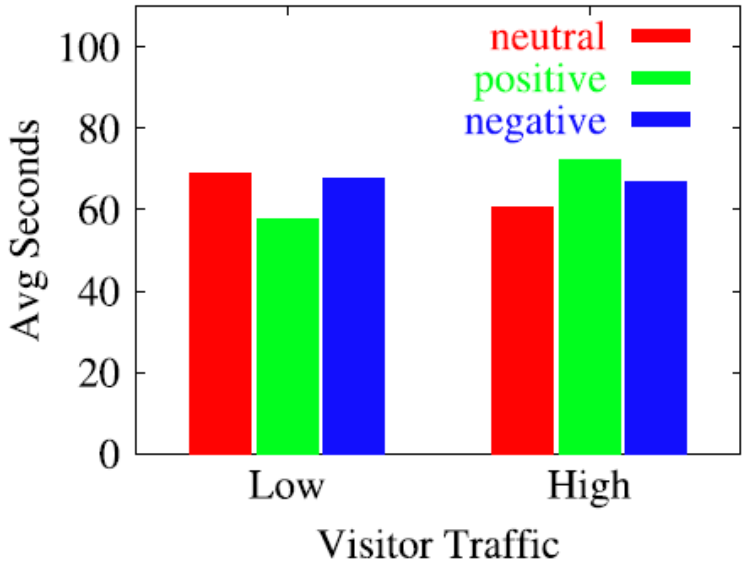
Evaluation results

- 100+ person user study showed that people can accurately identify the basic emotions based on facial expression

- 9-week interaction study, 8 hrs/day, 5 days/wk
  - Positive, negative and neutral mood

Evaluation results

- Over 1,600 interactions, 120 surveys
  - People could distinguish robot moods
  - People found the positive robot easier to understand and more natural
  - Length of interaction varied



Summary

- Overall approach much more mathematical

$$c = m - \frac{m}{1 + e^{-|d|+|10m|}}. \qquad s' = s\left(1 + \frac{1}{4}vm\right)$$

$$A_f = \frac{1}{2}\left(1 + \frac{1}{10}\min(\text{hours}, 10) - \frac{1}{30}\min(\text{days}, 30)\right)$$

$$m_d = \begin{cases} \dfrac{1 + m_b}{1 + e^{-B(m_o + M(1+m_b))}} - 1 & \text{if } m_o < 0 \\ m_b & \text{if } m_o = 0 \\ \dfrac{1 - m_b}{1 + e^{-B(m_o - M(1-m_b))}} + m_b & \text{if } m_o > 0 \end{cases}$$

$$A'_m = A_m + (\Delta m_o \text{ during interaction})(1 - A_f).$$

- Shows that people can accurately evaluate robot mood, but the effect of mood on the interaction is hard to pinpoint even in a very large study
- Applications?