

Affect and Emotions in Intelligent Agents: Why and How?

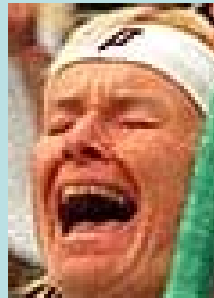


Andrew Ortony
Northwestern University

Affect and Emotions in Intelligent Agents: Why and How?

- **Affect:** value, good-bad, beneficial-harmful, etc.
- **Affect:** {emotions, moods, preferences, . . . }
- **Affect:** undifferentiated vs. differentiated
 - simple good-bad, positive-negative feelings
vs. complex feelings elaborated with meaning

Affect and Emotions in Intelligent Agents: Why and How?



Intelligent Agents (includes autonomous, humanlike robots, softbots, artificial characters, humans, etc.)

Affect and Emotions in Intelligent Agents:

Why?

- more lifelike, believable, and helpful agents
- computer-based models as theory test beds

Affect and Emotions in Intelligent Agents: Why and How?

Why we have affect and emotions, and how they work

- Effective Functioning
- Undifferentiated Affect and Effective Functioning
 - classic experiments: Attention, Altruistic Behavior
- Differentiated Affect (Emotions) and Effective Functioning

How to represent affect and emotions in agents

- Modeling the cognitive component
- Modeling behavioral manifestations

Affect and Emotions in Intelligent Agents: Why?

Why we have affect and emotions, and how they work

- ➡ – Effective Functioning
 - Undifferentiated Affect and Effective Functioning
 - Three classic experiments: Attention, Altruistic Behavior, Judgment
 - Differentiated Affect (Emotions) and Effective Functioning

How to represent affect and emotions in agents

- Modeling the cognitive component
- Modeling behavioral manifestations

Effective Functioning*

- Optimizing fit between an organism's functioning and environmental conditions
- Historically, affect thought to hurt EF
 - Under normal conditions, affect under control
 - Under abnormal conditions, affect out of control
 - emotional disorders (depression, GAD, phobias, bipolar disorder)

* Ortony, A., Norman, D. A. & Revelle, W. (2005). Affect and proto-affect in effective functioning. In J.M. Fellous & M.A. Arbib, *Who needs emotions: The brain meets the machine*. New York: Oxford University Press.

Effective Functioning

In biological systems, the interaction of

- **what** organisms “do” (**modes**)
- **how** they do it (**levels**)
- **where** they do it (**environments**)

Effective Functioning:

What organisms “do” – four modes

- feel
- want
- know, believe, think, etc.
- act
- perceive

Effective Functioning:

What organisms “do” – four modes

- feel – pure **affect** (value)
- want – **motivation** (action tendencies)
- know, believe, think, etc. – **cognition** (meaning)
- act – **behavior** (action)

Effective Functioning:

How organisms do it – three levels*

- fast, unlearned (e.g., reflexes) – **Reactive**
- learned automatic procedures – **Routine**
- controlled conscious activity – **Reflective**
- affect manifest differently at different levels

* For levels, see also Sloman, Minsky, etc.

Differentiated vs. Undifferentiated Affect

- progressively more differentiation from Reactive through to Reflective level
 - undifferentiated = diffuse
 - maximally differentiated = “full-blown” emotions
- both important for effective functioning – for responding to problems and prospects
 - attention
 - information processing
 - action readiness

Affect and Emotions in Intelligent Agents: Why?

Why we have affect and emotions, and how they work

- Effective Functioning

➡ – Undifferentiated Affect and Effective Functioning

- classic experiments: Attention, Altruistic Behavior

- Differentiated Affect (Emotions) and Effective Functioning

How to represent affect and emotions in agents

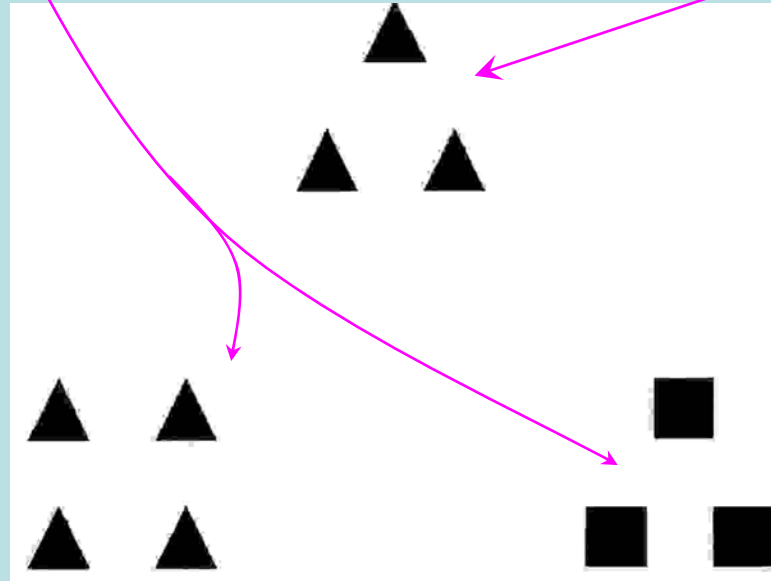
- Modeling the cognitive component
- Modeling behavioral manifestations

Undifferentiated Affect

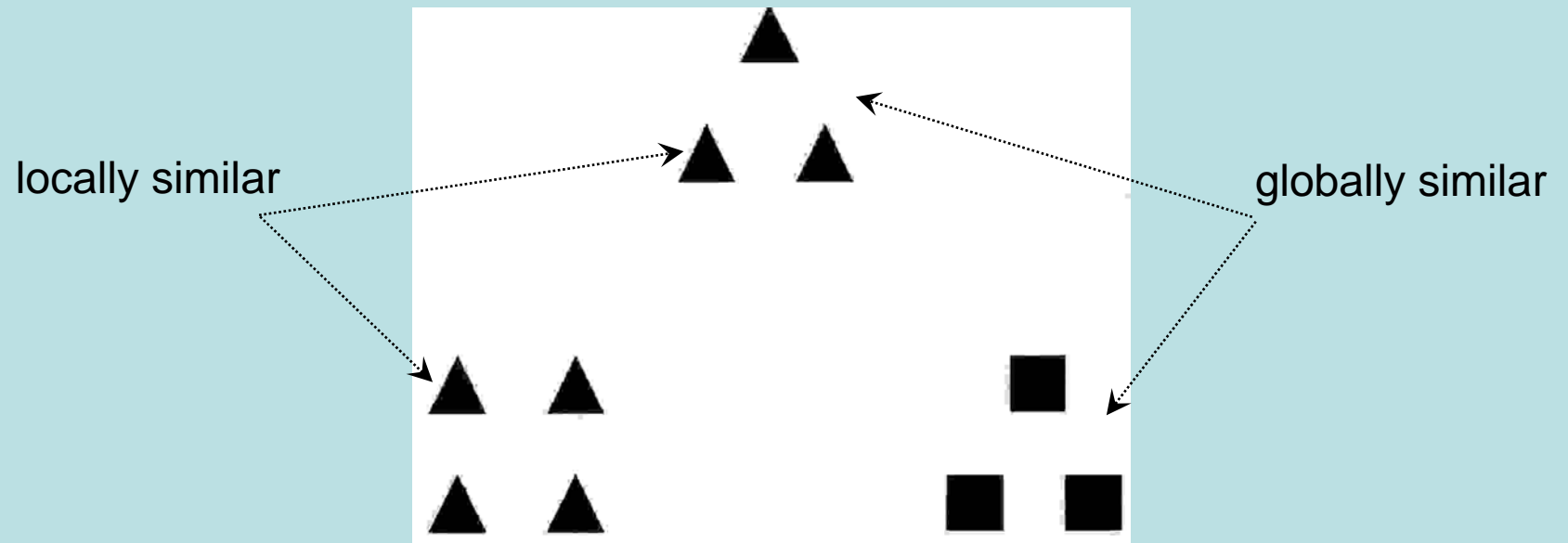
- negative
 - relatively more attention to local information (here and now)
 - useful for immediate threat detection
- positive
 - relatively more attention to global information (out there)
 - useful for curiosity and exploratory behaviors

Affect and **Attention**

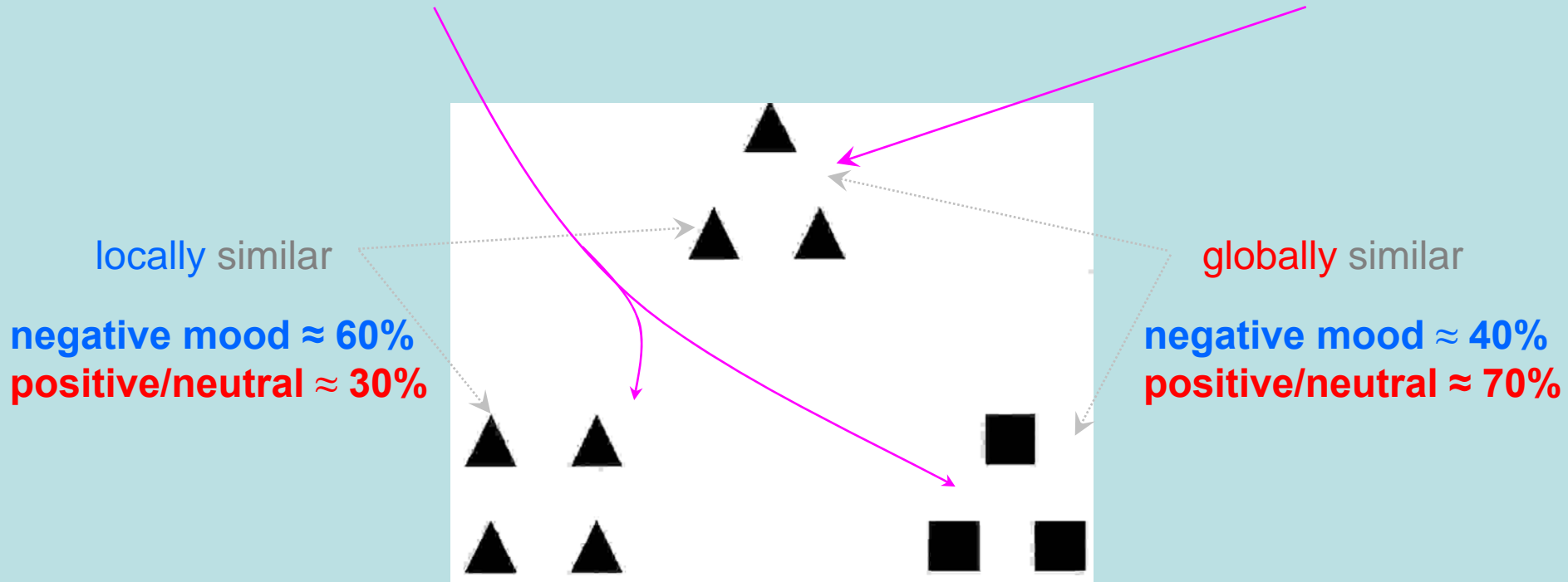
- mood induction (autobiographical event)
- which comparison figure is more similar to target ?



Gaspar, K., & Clore, G. L. (2002). Attending to the big picture: Mood and global versus local processing of visual information, *Psychological Science*, 13, 34-40.



which comparison figure is more similar to target ?



Affect and **Altruistic Behavior**

- 10¢ placed in the coin return slot of public phone
- confederate drops papers as caller leaves
- of those who found 10¢, 87.5% helped
- of those who didn't find 10¢, ?% helped

Isen, A. M., & Levin, P. F. (1972). Effect of feeling good on helping: Cookies and kindness. *JPSP*, 21, 384-388.

Affect and **Altruistic Behavior**

4%

Isen, A. M., & Levin, P. F. (1972). Effect of feeling good on helping: Cookies and kindness. *JPSP*, 21, 384-388.

Affect and Emotions in Intelligent Agents: Why and How?

Why we have affect and emotions, and how they work

- Effective Functioning
- Undifferentiated Affect and Effective Functioning
 - Three classic experiments: Attention, Altruistic Behavior, Judgment
- ➡ – Differentiated Affect (Emotions) and Effective Functioning

How to represent affect and emotions in agents

- Modeling the cognitive component
- Modeling behavioral manifestations

Deriving Differentiated from Undifferentiated Affect

Undifferentiated Affect: Just two simple kinds of feeling

- pure positive and pure negative affect
- only about the here and now

(1) a positive feeling about a **good** thing

(2) a negative feeling about a **bad** thing

From Undifferentiated Affect . . .

Undifferentiated

1. a positive feeling about a **good** thing (now)
2. a negative feeling about a **bad** thing (now)

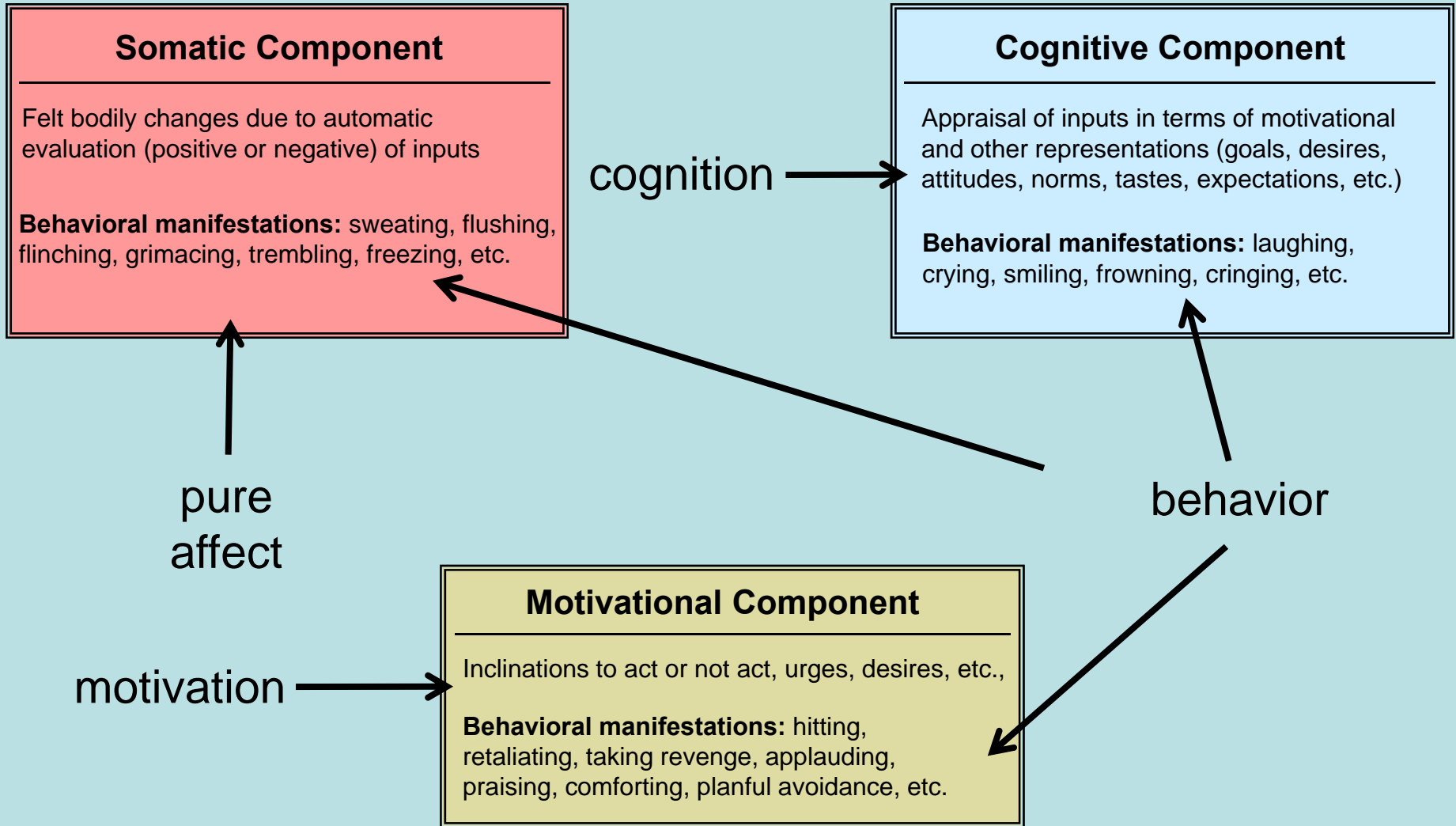
With some differentiation

3. a positive feeling about a **potential good** thing (possible future – primitive hope)
4. a negative feeling about a **potential bad** thing (possible future – primitive fear)

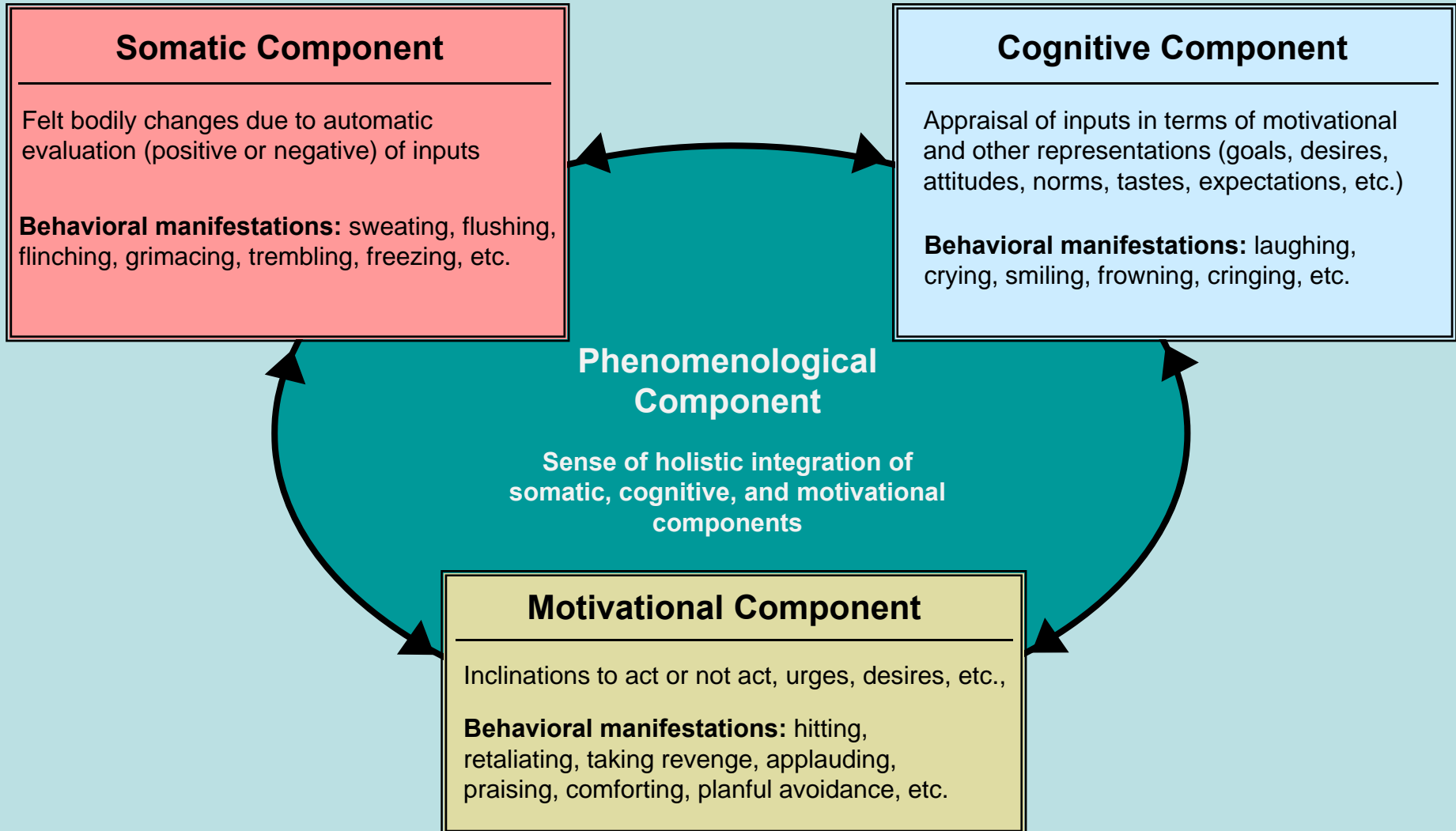
. . . to Maximally Differentiated Affect – “Full-blown” Emotions

Cognitively elaborated good/bad feelings

- adds **cognitive content** to undifferentiated affect
- potentially involves all modes: affect, cognition, motivation, and behavior



GENERAL STRUCTURE OF A TYPICAL FULL-BLOWN EMOTION



Affect and Emotions in Intelligent Agents: How?

Why we have affect and emotions, and how they work

- Effective Functioning
- Undifferentiated Affect and Effective Functioning
 - Three classic experiments: Attention, Altruistic Behavior, Judgment
- Differentiated Affect (Emotions) and Effective Functioning

⇒ How to represent affect and emotions in agents


- Modeling the cognitive component
- Modeling behavioral manifestations

Modeling the Cognitive Component

appraising the world in terms of concerns

- goals \Rightarrow goal-based emotions
- standards \Rightarrow standards-based emotions
- tastes \Rightarrow taste-based emotions

Cognitive content of goal-based motions

- 
- 1. a positive feeling about a **good** thing (pleased)
 - 2. a negative feeling about a **bad** thing (displeased)
 - 3. a positive feeling about a **potential good** thing (hope)
 - 4. a negative feeling about a **potential bad** thing (fear)
 - 3.1 a positive feeling because a **potential good** thing **happened** (satisfaction)
 - 3.2 a negative feeling because a **potential good** thing **didn't happen** (disappointment)
 - 4.1 a negative feeling because a **potential bad** thing **happened** (fears confirmed)
 - 4.2 a positive feeling because a **potential bad** thing **didn't happen** (relief)

Cognitive content of goal-based motions

-
- 1. a positive feeling about a **good** thing (pleased)
 - 2. a negative feeling about a **bad** thing (displeased)
 - 3. a positive feeling about a **potential good** thing (hope)
 - 4. a negative feeling about a **potential bad** thing (fear)
 - 3.1 a positive feeling because a **potential good** thing **happened** (satisfaction)
 - 4.1 a negative feeling because a **potential bad** thing **happened** (fears confirmed)
 - 4.2 a positive feeling because a **potential bad** thing **didn't happen** (relief)
 - 3.2 a negative feeling because a **potential good** thing **didn't happen** (disappointment)

Different kinds of concerns – e.g., goals vs. standards

Different classes of emotions – e.g., goal-based vs. . .

Different cognitive content – e.g., fear, relief, etc.

goals
(for events)

representations

hope
fear

relief
disappointment

satisfaction
fears-confirmed

**GOAL-BASED
EMOTIONS**

emotions

**EVENTS, AGENTS, OR
OBJECTS**

elicitors

appraised in terms of

goals

(for events)

norms/standards

(for agents' actions)

tastes/attitudes

(for objects)

representations

desirability

praiseworthiness

appealingness

criteria

hope
fear

relief
disappointment
satisfaction
fears-confirmed

**GOAL-BASED
EMOTIONS**

anger
gratitude

gratification
remorse

etc.

**COMPOUND
EMOTIONS**

pride
shame

admiration
reproach

etc.

**NORM-BASED
EMOTIONS**

love
hate

etc.

**TASTE-BASED
EMOTIONS**

emotions

DISAPPOINTMENT EMOTIONS

TYPE SPECIFICATION:

(displeased about) the disconfirmation of the prospect of a desirable event

(i.e., a negative feeling because a **potential good** thing **didn't happen**)

TOKENS:

dashed hopes, despair, disappointment, frustration, heartbroken, etc.

VARIABLES AFFECTING INTENSITY:

- (1) the intensity of the attendant hope emotion
- (2) the effort expended in trying to attain the event
- (3) the degree to which the event was realized

IF HOPE (p, e, t) > 0 AND BELIEVE (p, not (e), t2)

THEN set DISAPPOINTMENT-POTENTIAL (p, e, t2) =

f(|HOPE (p, e, t), EFFORT (p, e), REALIZATION (e, t2), lg (p, e, t2)|)

IF DISAPPOINTMENT-POTENTIAL (p, e, t2) > DISAPPOINTMENT-THRESHOLD (p, t2)

THEN set DISAPPOINTMENT-INTENSITY (p, e, t2) =

DISAPPOINTMENT-POTENTIAL (p, e, t2) – DISAPPOINTMENT-THRESHOLD (p, t2)

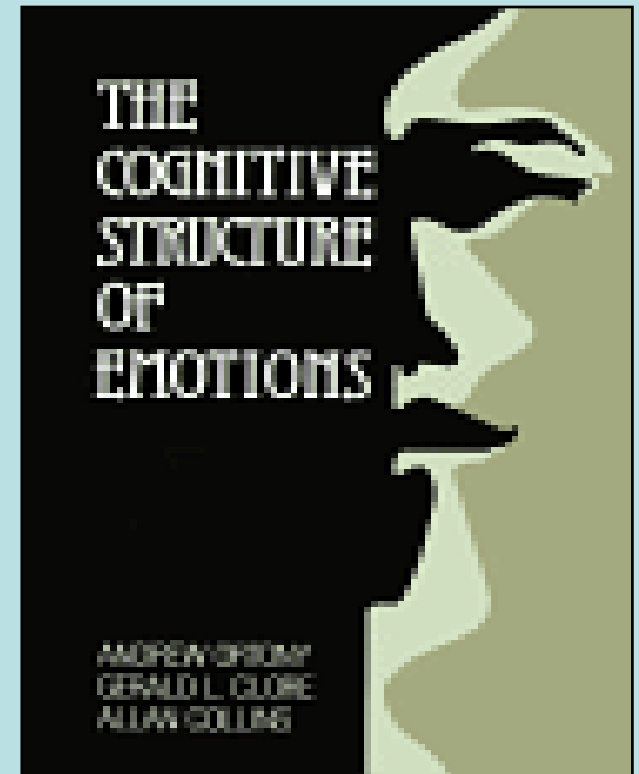
Modeling the Cognitive Component

One way . . .

. . . **o c c** model

r l o
t o l
o r i
n e n
y s

. . . more details in . . .





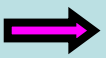
Mr. Bubb

Affect and Emotions in Intelligent Agents: How?

Why we have affect and emotions, and how they work

- Effective Functioning
- Undifferentiated Affect and Effective Functioning
 - Three classic experiments: Attention, Altruistic Behavior, Judgment
- Differentiated Affect (Emotions) and Effective Functioning

How to represent affect and emotions in agents

- Modeling the cognitive component
-  – Modeling behavioral manifestations

Modeling Behavioral Manifestations

- emotion-behavior linkage less constrained than cognition-emotion linkage



What misery!



What joy!

(Novotna, Wimbledon, 1998)

Modeling Behavioral Manifestations

- emotion-behavior linkage less constrained than cognition-emotion linkage
- affected by intensity and coping potential
- initial proposal for taxonomy of behavioral manifestations or “causal correlates”

Somatic Component

Felt bodily changes due to automatic evaluation (positive or negative) of inputs

Behavioral manifestations: sweating, flushing, flinching, grimacing, trembling, fleeing, freezing, and spontaneous aggression, etc.

Cognitive Component

Appraisal of inputs in terms of motivational and other representations (goals, desires, attitudes, norms, tastes, expectations, etc.)

Behavioral manifestations: laughing, crying, smiling, frowning, cringing, etc.

Motivational Component

Inclinations to act or not act, urges, desires, etc.,

Behavioral manifestations: hitting, retaliating, taking revenge, applauding, praising, comforting, planful avoidance, etc.

behavior

```
graph TD; behavior --> Somatic; behavior --> Cognitive; behavior --> Motivational;
```

Behavioral Manifestations

somatic

cognitive

motivational
(coping)

physiological
heart-racing

facial
grinning

physical
hyperactive
movement

vocal
shrieking

external

internal

emotion-oriented

problem-oriented
situation
perpetuation

facial
smiling

physical/postural
hand-slapping

vocal
cheering

self-regulating
calming down

other-modulating
telling others

evaluative
admiring

attentional
obsessing

reflective

reactive

Behavioral Manifestations of Joy Emotions

SOMATIC	physiological	heart racing/pounding
	facial	grinning
	physical	exaggerated, hyperactive, movements
	vocal	shrieking, screaming
EXTERNAL COGNITIVE INTERNAL	facial	smiling
	physical/postural	opening out, jumping for joy, “high-fiving”
	vocal	cheering, laughing
	evaluative	admiring/liking perceived facilitators
	attentional	obsessing about how it happened/what it means
MOTIVATIONAL	emotion: self-regulating	sharing the news, telling others
	emotion: other-regulating	soliciting praise/congratulations from others
	problem-oriented	maintaining/preserving cause, reaffirmation

Behavioral Manifestations of Anger Emotions

SOMATIC	physiological	shaking, flushing
	facial	baring of teeth
	physical	clenching of teeth or hands, body tense
	vocal	raising of voice
EXTERNAL COGNITIVE INTERNAL	facial	scowling, glaring
	physical/postural	throwing, banging, hitting
	vocal	shouting, swearing, sarcastic/sulky remarks, refusing to speak
	evaluative	disliking offender
	attentional	obsessing about event, ruminating about how and why
MOTIVATIONAL	emotion: self-regulating	calming down, getting a grip, talking it over
	emotion: other-regulating	causing distress to offender (direct aggression, denial of benefits)
	problem-oriented	preventing continuation or recurrence of problem

Modeling Behavioral Manifestations

- which particular behaviors are selected depends partly on motivational component – considerations of capacity, time, prudence, etc.
- initial rough proposal only – needs refining, validating, and eventual formalizing

Conclusions

- a way to think about the complex interactions of affect, motivation, cognition, and behavior
- evident role of affect and emotion in effective functioning in humans justifies incorporation in computational artifacts
- some suggestions about how to incorporate at least (non-physiological/somatic) aspects of relations between emotion, cognition, and motivation and behavior

thank you

