Carnegie Mellon

Theoretical and empirical guidance for a chunk valuation mechanism in ACT-R

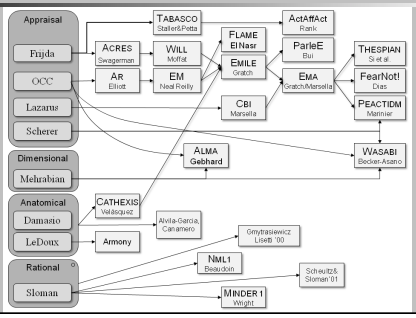
Ion Juvina, Alessandro Oltramari, & Christian Lebiere
Department of Psychology
Carnegie Mellon University

Why babies on ACT-R website

Outline

- Review of existing approaches
- Theoretical and empirical guidance
- Proposed design
- Discussion

Review of existing approaches



Modeling objectives

- Largely pragmatic
  - Improve functionality
    - Autonomy (Ritter, Marinier III)
  - Create humanlike agents / interfaces
- Theoretic
  - A way to define and understand emotions (Ritter)

Lack of guidance

- "although there is little theory or empirical evidence to guide us, we define…" (Marinier III et al., 2008)
- "computational models must fill in those details, but with little or no direction from appraisal theory, the details are often arbitrary" (Marinier & Laird, 2008)
- "Appraisal theory provides, at best, a high-level specification for a computational model of emotion, forcing modelers to adopt representational and process assumptions to create a working system" (Gratch et al., 2009)

Algorithmic level analysis

- Proliferation of algorithms and representations
  - Little convergence between approaches
- Functional level (Marr's computational)
  - Poorly guided assumptions
    - Philosophical legacy
    - Controversial or biased science
      - E.g., amygdala

Emotion has a mind of its own

- Emotion and cognition are separate systems
  - Emotion precedes cognition (Zajonc, 1980)
  - Basic emotions are automatic, universal, fast, effortless, irrational, unmediated, primitive (innate),
    - Thus, they must be encapsulated

Emotion specificity

- Specific emotions (fear, anger, joy, etc) are architectural primitives
  - Each basic emotion has a unique neural circuit (e.g., fear - amygdala)
  - Basic emotions are universally recognized from facial expressions
  - Specific emotions have distinct ANS signatures

Not much learning

- Need for learning is recognized
- Learning is "work in progress"
  - First priority: a "working system"

Guidance

- Theory and empirical data <-> modeling
- Lots of research on emotions
  - False positives, publication bias, etc.
  - Hard to sort out
  - Focus on comprehensive reviews and meta-analyses

Functional role of emotion

- Emotion guides information processing
  - Fast and efficient assessment of value (Pham, 2004)
    - what feels good must be desirable
  - Supports ecological rationality (Pham, 2007)
- Brain regions associated with emotions are also associated with regions involved in goal-directed and adaptive behavior (amygdala, OFC, ACC) (Murray, 2007)

Emotion vs cognition

- Emotion does NOT have a mind of its own
  - There is no evidence of any clear dissociation between emotion and cognition
    - animal models, human behavioral, physiological and brain imaging data
      - same brain regions are involved in both emotional and cognitive executive control (LPFC, OFC, MPFC, ACC, Amygdala) (Ochsner & Gross, 2005; Pessoa, 2008).
  - No clear evidence of specialization and encapsulation of affective processes
    - Automaticity results from practice

Emotion specificity

- Emotions are not psychological primitives (Ortony & Turner, 1990)
  - They are not biologically given (Barrett, 2006)
- Specific emotions result from learning and categorization (Barrett, 2006)
  - Context and learning history (Ekman, 1992)
  - Situated conceptualizations (Barsalou, 2005)
- Evidence in favor of core affect (Cacioppo et al., 2000; Russell, 2003; Paton et al. 2006)
  - Approach-avoidance (valence) and Arousal

Emotion specificity cont'd

- No objective means to measure the experience of emotions
  - No way to scientifically confirm when a person is happy or angry or sad (Barrett, 2006).
  - No distinct pattern for specific emotions
  - Behavioral responses correspond to situational demands rather than to specific emotions (Bouton, 2005)

Learning is essential

- Conditioning
  - Aversive signal + neutral stimulus <--> arousal (change in skin conductance)
- Instructed and observational learning
  - Symbolically acquired fear results in the same physiological fear responses as conditioned fear (Amygdala + SCR)
- Linguistic interpretation and episodic memory are critical for emotional learning
  - Associative learning
- Emotion regulation:
  - Through conscious strategies and practice individuals can change their interpretation of specific stimuli (reappraisal)

Suggestions

- Emotional module?
  - No (unless needed for practical reasons)
- Specific emotions?
  - Not in the architecture
  - Yes in specific models (but after perception, associative learning, categorization, language)
- Core affect?
  - Yes !!!

What do we need?

- Arousal and valence
- Arousal = chunk activation
- Valence = chunk valuation
  - We use it in specific models:
    - Lebiere & Gonzalez ("utility"); Juvina & Taatgen
    - Constrained retrieval requests contain implicit valuation judgments
  - We need an architectural mechanism for learning and extinction of chunk valuations
    - RL

How would it work?

- Retrieval = f(activation,valuation)

Example 1

- An obstacle
  - is goal-relevant (high activation)
  - Is not goal-congruent (negative valuation)

Example 2

- Base-level inhibition:
  - If we don't have it: perseverative retrieval
  - If we have it: no repetition priming
- Retrieval = f(activation,valuation)
  - High valuation -> repetition priming
  - Low valuation -> no perseverative retrieval

Summary and conclusion

- The current hype on emotion modeling
  - Partially misguided
  - Partially relevant
    - Universal mechanism of valuation
      - Goal-directedness, autonomy, adaptivity
- Further work needed to develop a learning mechanism for chunk valuation

Discussion

- Questions?

End

- Thank you for your attention!

Extra slides

25

Definition Emotions

- Valenced responses to external / internal stimuli
- Identifiable objects / triggers
- Stimuli
  - Intrinsic affective properties: aversive shock
  - Stimuli with acquired emotional value

26

Primacy of emotion?

- Amygdala activation to fearful faces
  - Even subliminally presented
- However, an emotional context conveyed verbally can alter the amygdala response to a facial expression
  - E.g., “she just won/lost $500”

27

Amygdala

- The most thoroughly investigated (since Kluver & Bucy, 1937)
- Traditionally thought to be involved in aversive conditioning
  - Individuals with bilateral damage to the amygdala can experience fear (LeDoux, 1996)
- Also involved in
  - appetitive conditioning and reward learning (Murray, 2007)
  - updating the current value of stimuli (reinforcer-devaluation effects)
    - Monkeys with inactive amygdalas keep eating after they had enough food (Wellman et al., 2005)
  - computing the predictive value of a stimulus (i.e., likelihood that it predicts reward or threat) (Kim et al., 2004)
  - Boosting processing in uncertain and ambiguous contexts (LeDoux, 1996; Barrett, 2006)

28

Valence

- (Paton et al. 2006) single neuron recording in monkeys:
  - Separate (amygdala) neurons encode positive vs. negative valence

29

Emotions as (embodied) concepts

- Emotional responses evoked by objects are stored with memory representations of these objects as somatic markers of these objects’ value (Damasio, 1994; Bechara, 2004)

30

Categorization

- The affective system assigns value primarily through processes of categorization: objects or events are mapped onto existing categories or schemas (Pham, 2007)
- Emotion is generated by
  - Associative learning (Barrett, 2006)
    - Core affect + categorization
  - Comprehension
  - Interpretation (Blanchette & Richards, 2011)

31

What is core affect

- Perceptual–cognitive phenomenon
  - Like color perception
  - Unconditioned stimuli -> affective reactions (arousal & approach/avoidance)
  - Conditioned stimuli -> affective reactions
  - Categorization: conditioned stimuli + context -> specific emotions

32