

Concern Processing in Autonomous Agents



Emotions, Personality and
other Woolly Concepts

What Are Emotions?

- *"An emotion is usually caused by a person consciously or unconsciously evaluating an event as relevant to a concern (a goal) that is important; the emotion is felt as positive when a concern is advanced and negative when a concern is impeded.*
- *The core of an emotion is readiness to act and the prompting of plans; an emotion gives priority for one or a few kinds of action to which it gives a sense of urgency – so it can interrupt, or compete with, alternative mental processes or actions. Different types of readiness create different outline relationships with others.*
- *An emotion is usually experienced as a distinctive type of mental state, sometimes accompanied or followed by bodily changes, expressions, actions."*

- Oatley and Jenkins, *Understanding Emotions* (page 96)

Why Do Animals Have Emotions?



- Emergency Responses
- Communication
- Survival (species vs. individual)

- Humans have evolved a vast richness in different types of emotion and their use

Why Give Computers Emotions?



- Entertainment / Simulation
- Human-Computer Interaction
- Solutions for incomplete knowledge
- Aid Our Understanding Human Emotions

- Each reason offers different challenges

The Synthetic and The Artificial



- Synthetic (observed) Emotions
 - Intentional Stance
 - Cognitive Structure of Emotions
 - Affective Reasoning Engine

- Artificial (analysed) Emotions
 - Cognitive Modelling
 - Capture the Emotion Process

Pros and Cons



Synthetic

- Rational Agents
- Homogeneous
- Approximate
- Specialised

Artificial

- Cognitive Agents
- Co-Evolution
- Human-Like
- Broad-but-Shallow

Synthetic Emotions: The OCC Approach



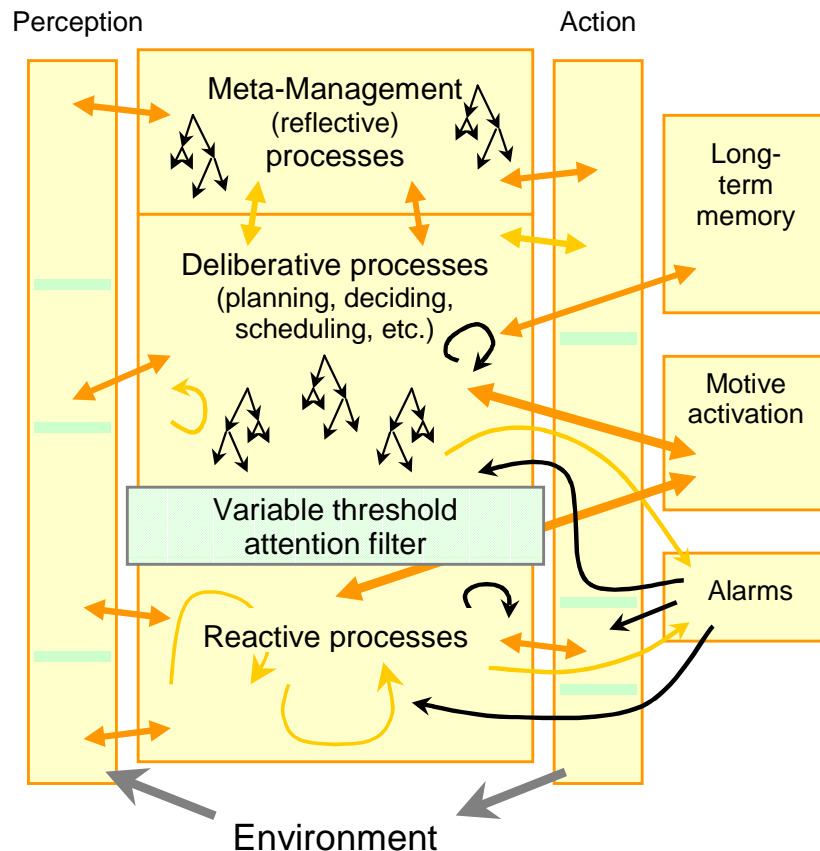
- Treat Emotion as a Black Box Process
 - Reason about Causal Triggers
 - Reason about Response
 - High-Level and Abstract
-
- Problems with apparent irrationality
(or rather automatic behaviour)

Artificial Emotions: Where To Start?



- Define An Agent Framework
 - Identify Emotion Classes
 - Map Cognitive Processes
 - Design/Build Emotional Agents
 - Refine the Design
-
- The Design-Based Approach

Motivated Agent Framework



- Three Layers
- Concurrent Processes
- Attention Filter(s)
- Global Alarms
- Supports Many Different Types of Control State

What Are Control States?

Control States of varying Scope and Duration

Long term

Relatively hard to change,
very slow learning, causes
and effects diffuse and
indirect.

Short term

Changeable, more specific
causes and effects,
semantic content.

Personality
Skills

Attitudes
Standards
Preferences

Moods
Emotions

Desires
Intentions
Plans

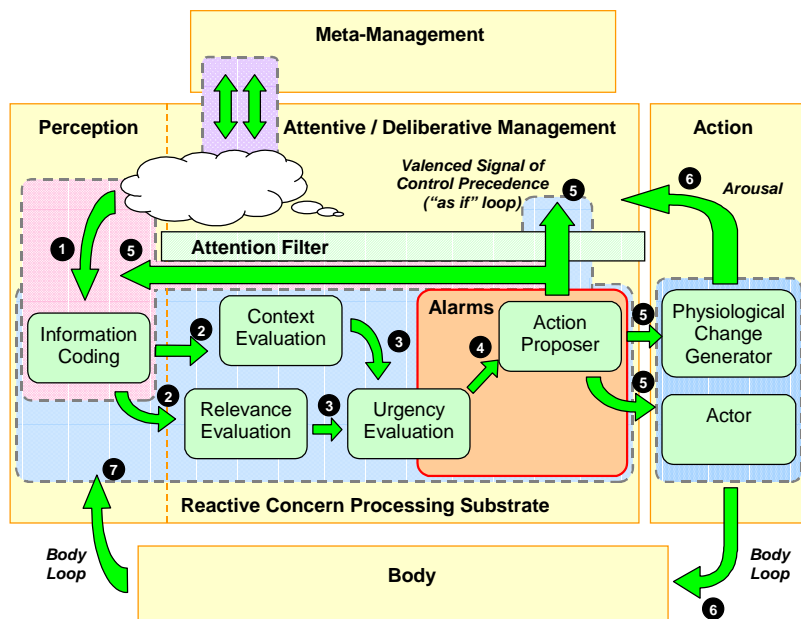
- “Information-bearing representations of an information-processing control system”
- Two Attributes
 - Dimensional
 - Structural
 - (Functional)

Three Classes of Emotional State

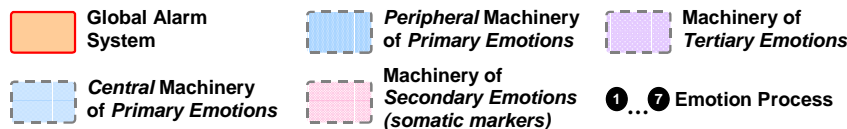


- **Primary emotional states:** such as being startled, terrified, or sexually stimulated, are typically triggered by patterns in the early sensory input and detected by a dedicated global alarm system. These emotional states are sometimes called primes or primary emotions.
- **Secondary emotional states:** such as being anxious, apprehensive, or relieved, depend on the existence of a deliberative layer in which plans (for future states) can be created and executed with relevant risks noticed, progress assessed, and success detected. Damasio [94] terms cognitively generated emotional states – secondary emotions.
- **Tertiary emotional states:** such as feeling humiliated, ashamed, or guilty, can be further characterised by a difficulty to focus attention on urgent or important tasks. These emotions cannot occur unless there is a meta-management layer to which the concept of “losing control” becomes relevant.

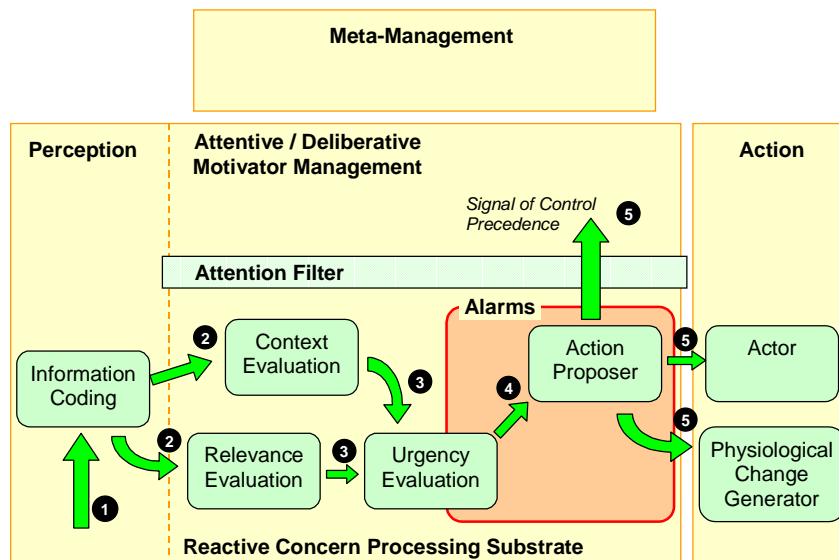
The Emotion Process



- Based on Frijda [86]
- Allows us to identify the active processes
- Emotions are an emergent phenomena
- No single Emotion System

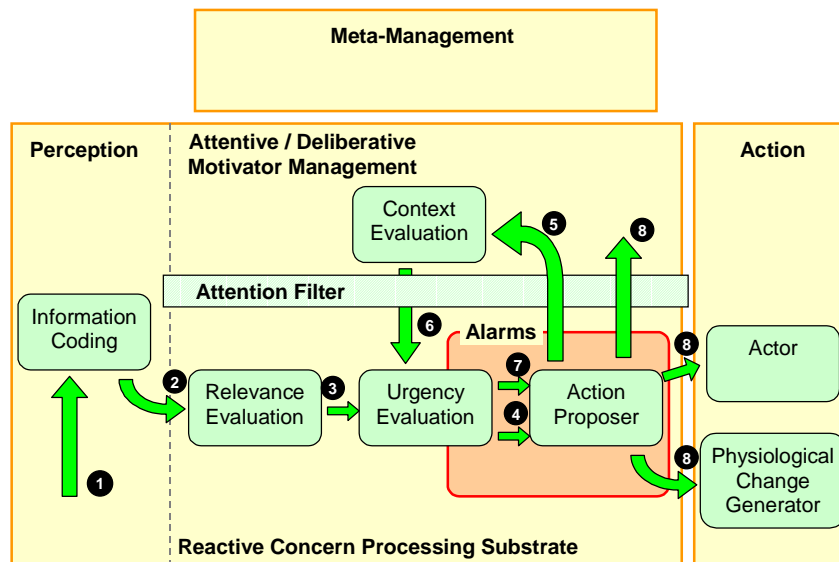


Information Flow in a Primary Emotion



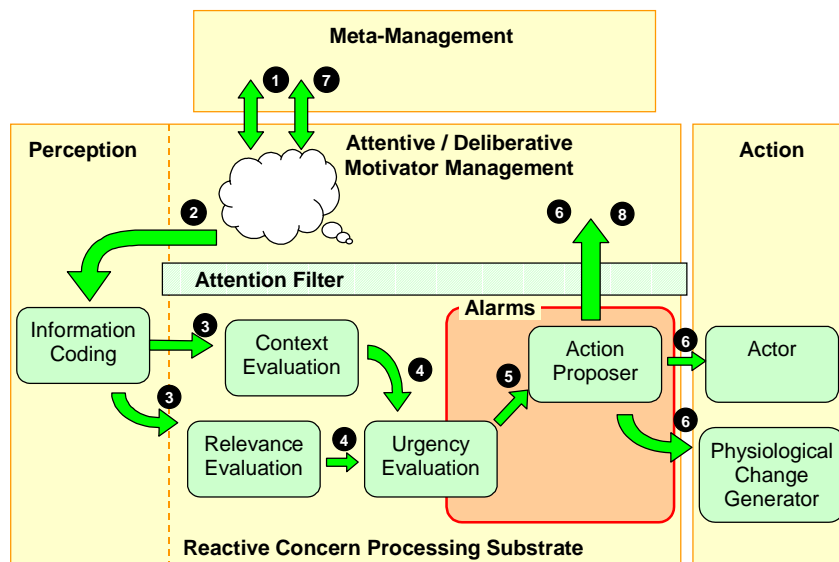
- External percepts encoded
- Relevance to concerns and context of situation evaluated
- Urgency of event evaluated
- Action readiness change is generated and presses for control precedence
- Attentive cognition is interrupted, involuntary action *might be* performed, and some form of physiological change *might be* instigated

Information Flow in a Secondary Emotion



- Requires deliberation
- Generated: (a) in response to deliberative thought process, or (b) with deliberative context evaluation
- Interruption of attention
- Uses mechanisms of primary emotions
- Many different pathways to trigger emotional response: somatic markers are only one possible path

Information Flow in a Tertiary Emotion



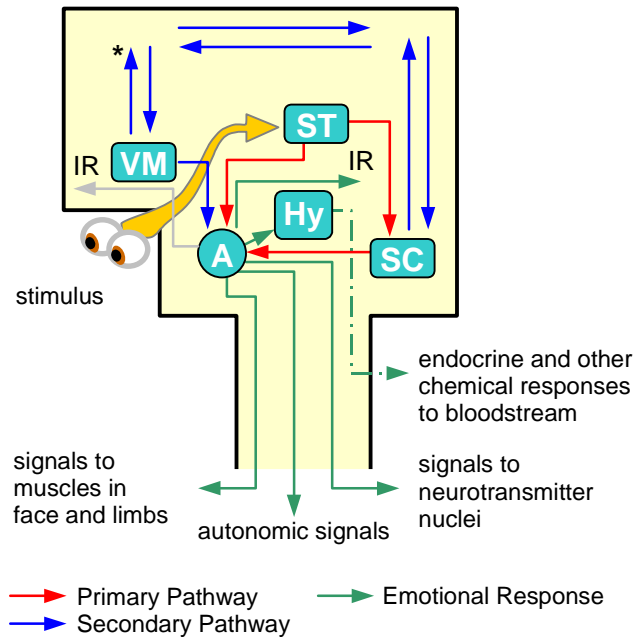
- Sub-class of secondary emotions
- Characterised by a loss of control of attentive processing
- Mechanisms to control attentive processing
- Arise from mismatch between cultural conditioning and the more universal mechanisms of primary emotions

Sources of Inspiration



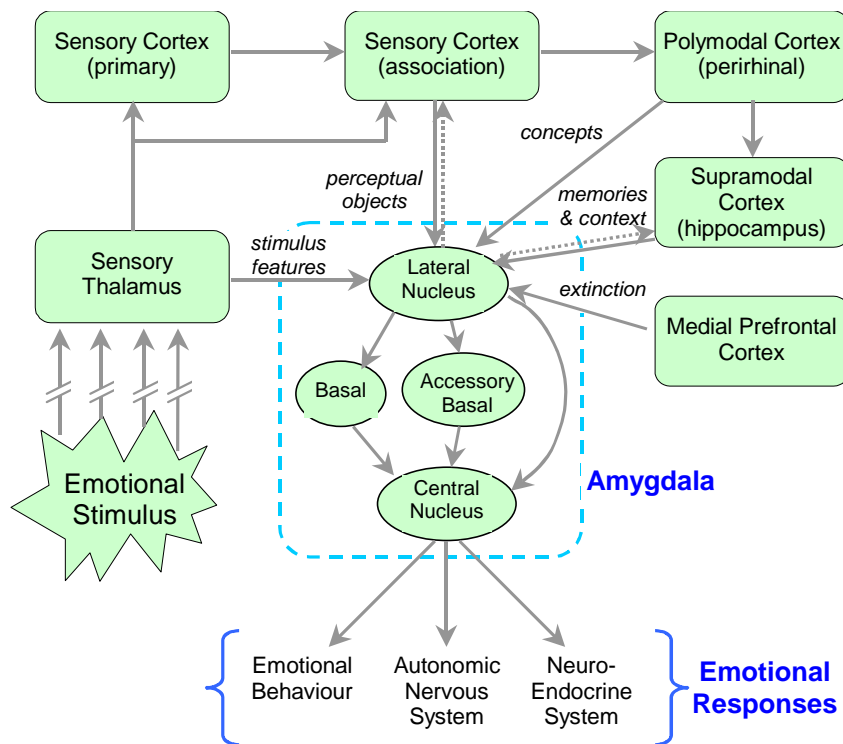
- Computer Science [Simon 67; Brooks 86]
- Cognitive Science [Sloman 99; Minsky 85; Cañamero 97; Moffat & Frijda 95; Velásquez 99; McCauley & Franklin 98]
- Psychology [Frijda 86]
- Neurology [Damasio 94; LeDoux 96 ; Rolls 99]

Somatic Markers



- Damasio [94]
- Primary and secondary emotions
- Learnt associations
- Falsely interpreted as emotions necessary for rational thought

Fear and the Amygdala



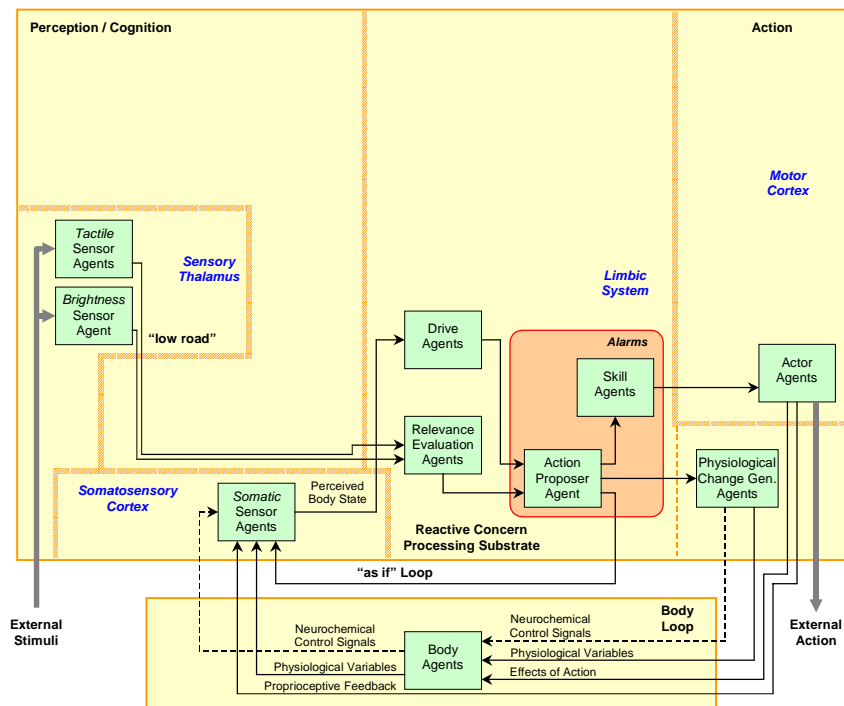
- LeDoux [96]
- Amygdala central to "fear" system, but not positive emotions.
- Mapping of function onto brain site
- Often a two-way communication channel

Co-Evolution of Emotion



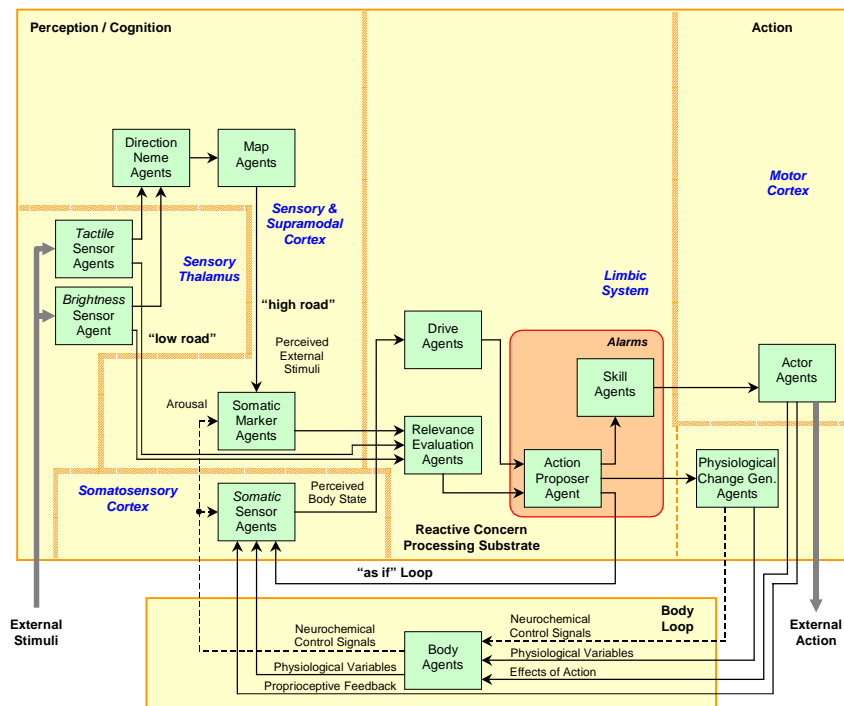
- Society-of-Mind Architecture [Cañamero 97; Minsky 85]
- Subsumption Style Philosophy [Brooks 86]
- Co-evolution of Competence Levels
 - Mapped to Brain Functions [LeDoux 86; Damasio 94]
- Leads to a Co-evolution of Emotion

Competence Level 0



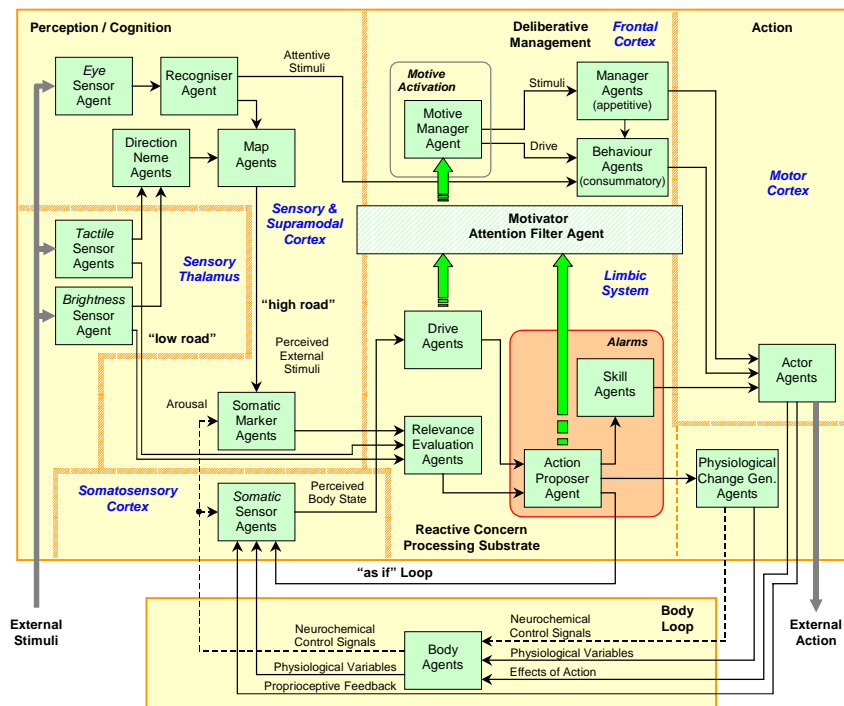
- Actively sense the environment and respond to its basic needs
- Homeostatic and non-homeostatic drives
- Inhibition of drives through body or "as if" loop

Competence Level 1



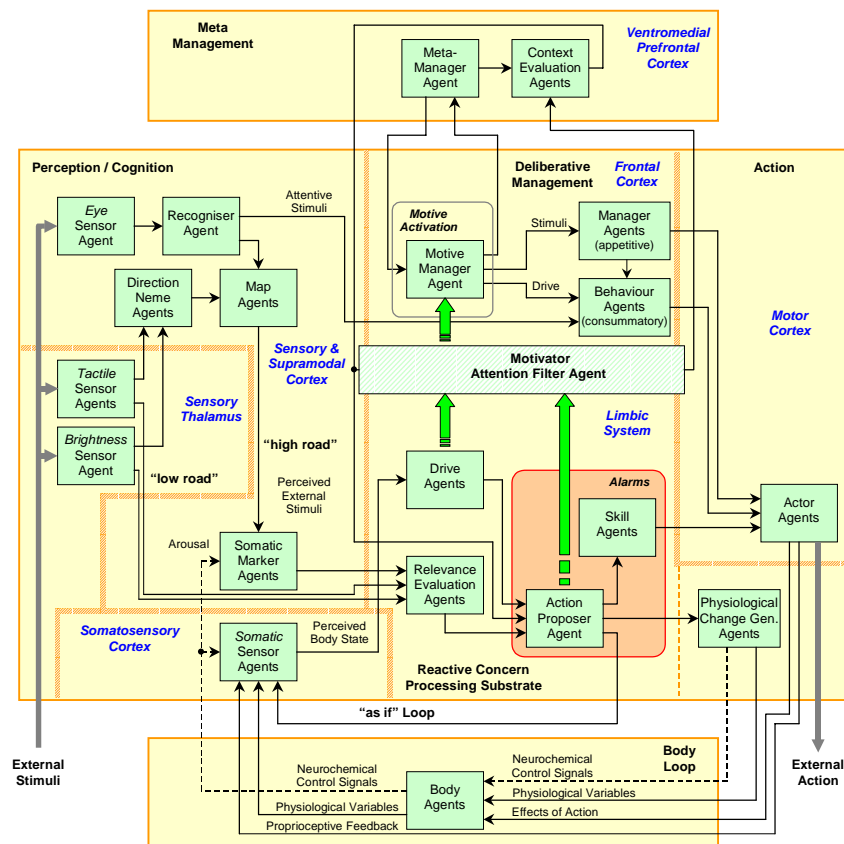
- Integration of senses to produce percepts
- Learnt associations between percepts and body state
- Rudimentary *primary* and *secondary* emotions
- No interruption of attentive processing

Competence Level 2



- Deliberative motive management layer
- Attention Filter
- Separation of deliberative and reactive percept processing
- Interruption of attentive processing

Competence Level 3



- Three layered model
- Active control of attentive processing
- Global alarm mechanism fully integrated into system
- Supports primary, secondary, and tertiary emotions

Emergence of Emotions

■ *Primary Emotions*

- “low road” - early *sensory* agents (tactile and brightness)
- “high road” - *direction neme, map, and somatic marker* agents

■ *Secondary Emotions*

- emergent emotional states that require deliberation
- *recogniser* agent and the “high road”
- *meta-manager* agent monitoring the *motive manager* agent

■ *Tertiary Emotions*

- emergent emotional states that require reflection
- *meta-manager* agent monitoring the *motive manager* agent

Where Next?



- Simulations
 - Design-Based Approach
 - Society-of-Mind Model
- Refinement
 - Relevance Evaluation Agents
 - Identify How Layers interact

A Quick Word On Personality



- Some determinates of agent personality?
 - Agent *concerns* (motivators, goals) (interpretive)
 - Motivational profile (interpretive)
 - Sensitivity of emotional states (interpretive)
 - Spontaneous vs. planned responses (manifest)
 - Persistence in attaining goals (manifest)
 - Choice of words and gestures (manifest)
- Properties of the architecture? Or are they best captured by a personality module?

Bibliography

- Allen, S. (2000) *Concern Processing in Autonomous Agents*. Submitted PhD Thesis, University of Birmingham.
- Brooks, R. (1986) A Robust Layered Control System for a Mobile Robot. *IEEE Journal of Robotics and Automation*, Vol. RA-2, No. 1, pages 12-23.
- Cañamero, D. (1997) Modeling Motivations and Emotions as a Basis for Intelligent Behavior. In *Proceedings of the First International Symposium on Autonomous Agents, AA'97*, Marina del Rey, CA, February 5-8, The ACM Press.
- Damasio, A. (1994) *Descartes' Error: Emotion, Reason and the Human Brain*. New York: G. P. Putman's Sons.
- Frijda, N. (1986) *The Emotions*. Cambridge: Cambridge University Press.
- LeDoux, J. (1996) *The Emotional Brain: The Mysterious Underpinnings of Emotional Life*. New York: Simon and Schuster.
- McCauley, L. & Franklin, S. (1998) An Architecture for Emotion. In *Working Notes of AAAI Fall 1998 Symposium*.
- Minsky, M. (1985) *The Society of Mind*. New York: Simon & Schuster.
- Moffat, D. & Frijda, N. (1995) Where there's a will there's an agent. In M. J. Wooldridge, and N. Jennings (Eds.), *Intelligent Agents, ECAI-94. Lecture Notes in Artificial Intelligence 890*. New York: Springer-Verlag, pages 245-260.
- Rolls, E. T. (1999) *The Brain and Emotion*. Oxford: Oxford University Press.
- Simon, H. (1967) Motivational and emotional controls of cognition. Reprinted in *Models of Thoughts*, Yale University Press, (1979), pages 29-38.
- Sloman, A. (1999) Architectural Requirements for Human-like Agents Both Natural and Artificial. (What sorts of machines can love?). To appear in K. Dautenhahn (Ed.) *Human Cognition And Social Agent Technology*, John Benjamins Publishing.
- Velásquez, J. (1999) When Robots Weep: Emotional Memories and Decision-Making. In *Proceedings of the Fifteenth National Conference on Artificial Intelligence*, Menlo Park, CA: AAAI Press, pages 70-75.