Intelligent Control
and Cognitive Systems

brings you...

Emotions, Drives and Complex Control

Joanna J. Bryson
University of Bath, United Kingdom

Emotions in Cognitive Systems

- Natural Cognition
  - Organise behaviour / provide another different kind of control state.
    - This includes social behaviour.
- Artificial Cognition
  - Organisation, communication, variation.

What are emotions for?

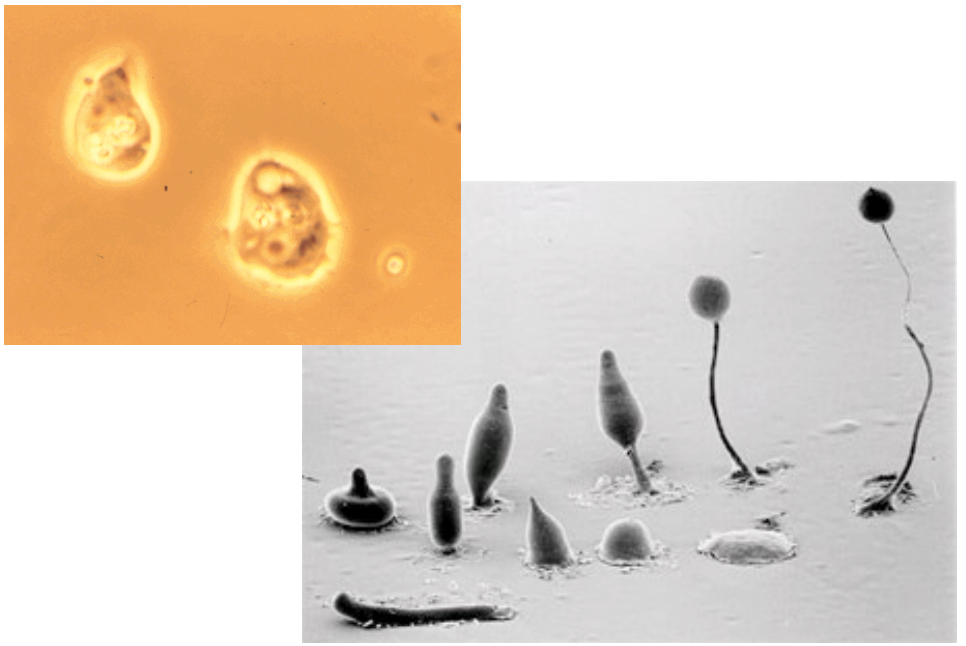
(Brutally functionalist answer.)

- Emotions are the original form of intelligence, and still the core organising structure of mammal intelligence.
- Intelligence is an evolved system that lets us change behaviour quickly.
  - Goal: do the right thing at the right time.

DODDER SMELLS ITS HOST; LOCATES BY VOLATILES FROM HOST

Very Simple Intelligence

Plants can wind & unwind (reversing decisions) in pursuit of support, light, prey.
(Anthony Trewavas, Edinburgh)



Single cell organisms also pursue multiple goals & hunt prey.

The Most Basic Emotions

- Emotions (or at least neurotransmitters) coordinate behaviour even in primitive animals that don't have neurons.
- The most basic emotions are excitement and depression.
  - Action Options: act urgently, withdraw, or act normally.

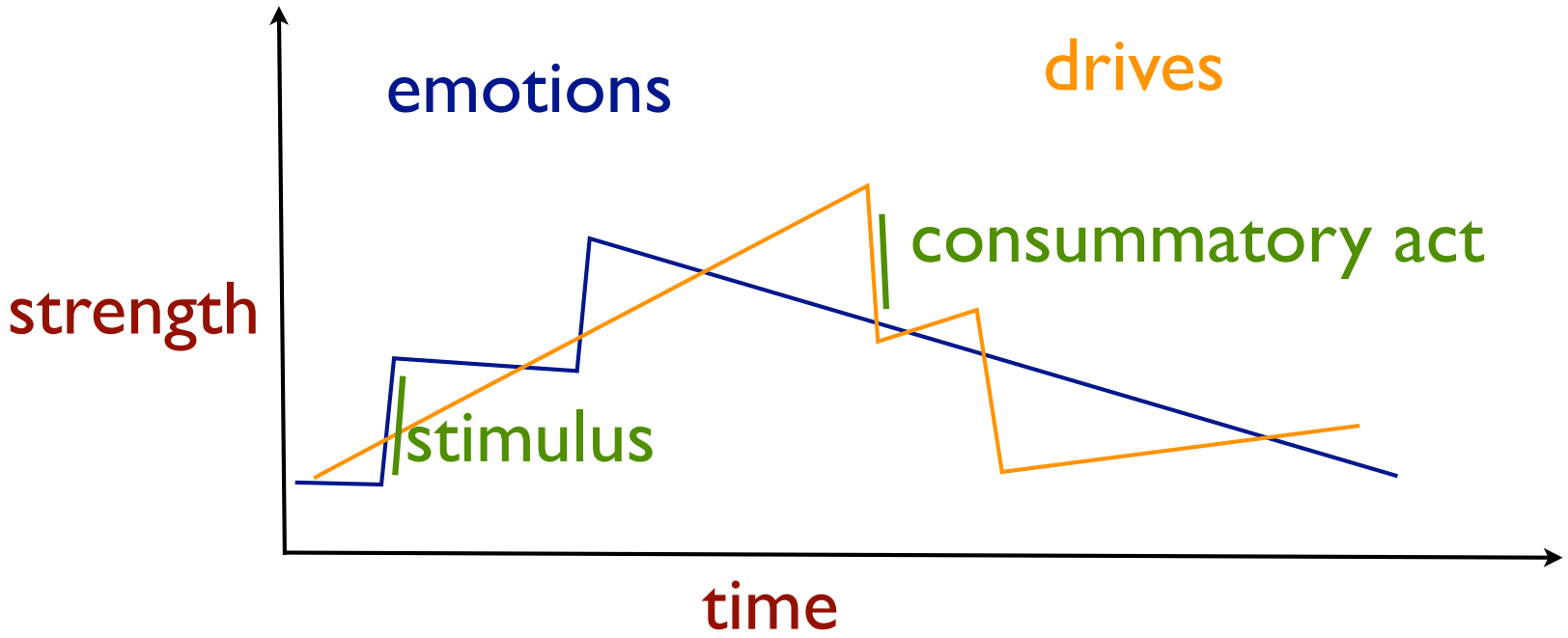
Carlson, *The Physiology of Behaviour,* Pearson, 2006.

Natural intelligence responds to a tiger

Intelligent Control

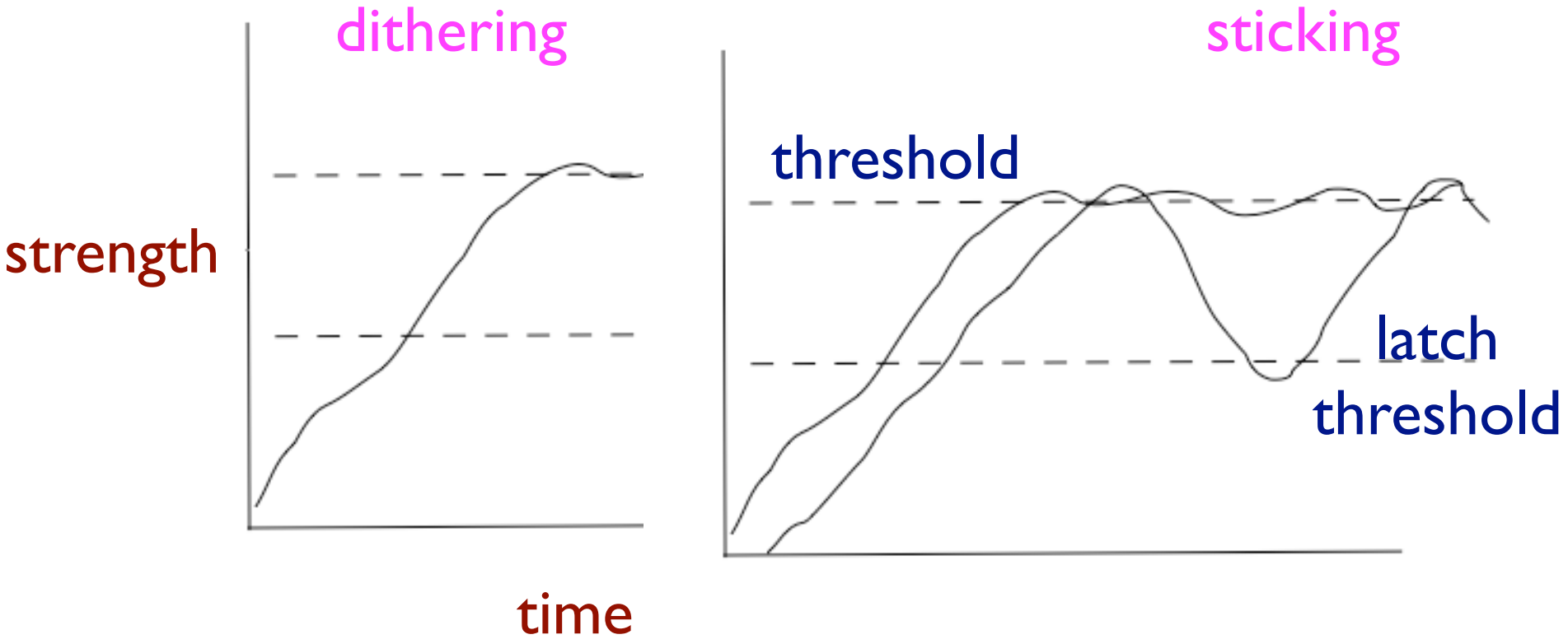
- Fast, complicated processes (e.g. perception) done by electrical state of neurons.
- Long-term learning done by growth/change of neurons.
- Intermediate action context – priorities – stored by chemical wash. Emotions & Drives

Emotions *vs.* Drives



- Emotions build acutely due to perceived events, decay with time or interference.
- Drives build with time (sometimes acutely due to perception like emotions), decay acutely with consummatory actions.

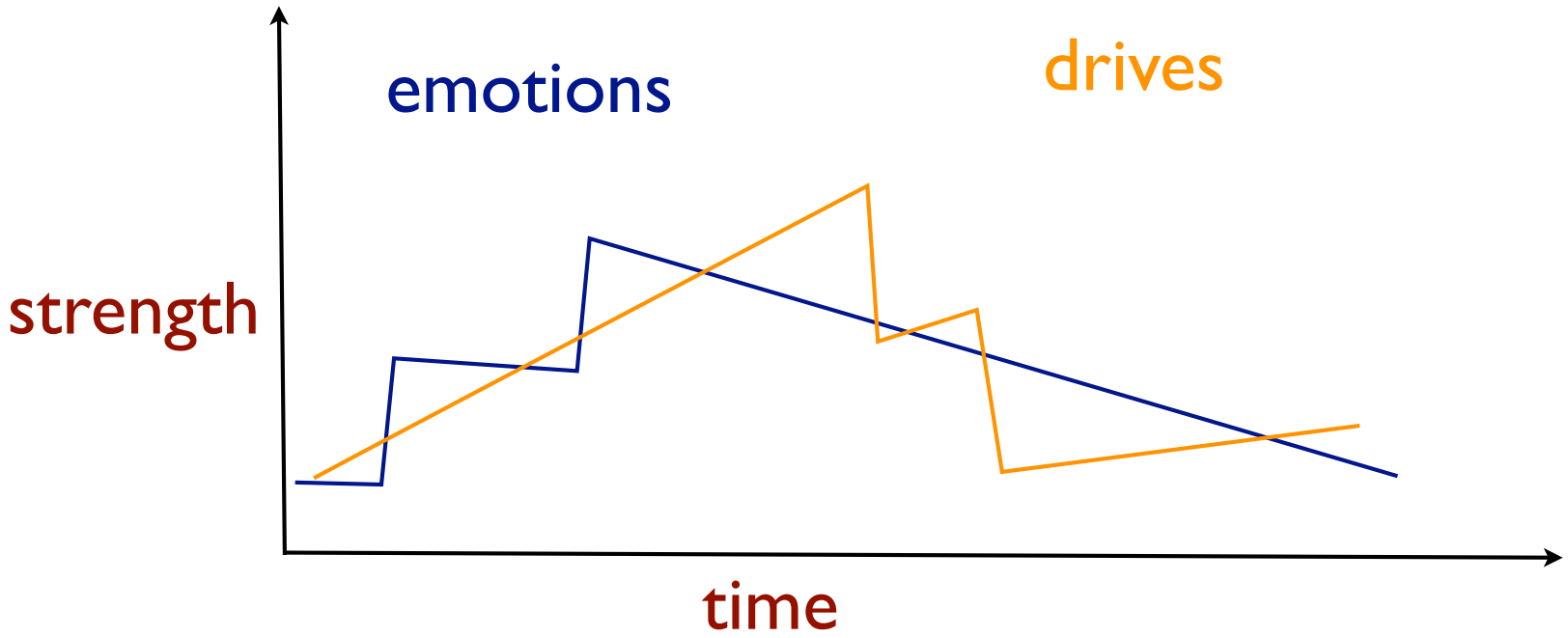
Basic Control: Latching



Note zone where action is not determined only by strength, but also by memory.

This is done in thermostats using a simple spring & magnet.

Emotions & Drives



- In NI, both emotions and drives use chemical "memory" (state) + attention as latches.
- Stick to one behaviour a while, reduce dithering.

Durative State in Synthetic Control

- Emotions and Drives are a sort of temporary memory system to help you arbitrate between goals.
- Simulate the chemical levels numerically.
- Other systems, e.g. going to just the most urgent goal, are inefficient, lead to dithering.

Experiments in Latching

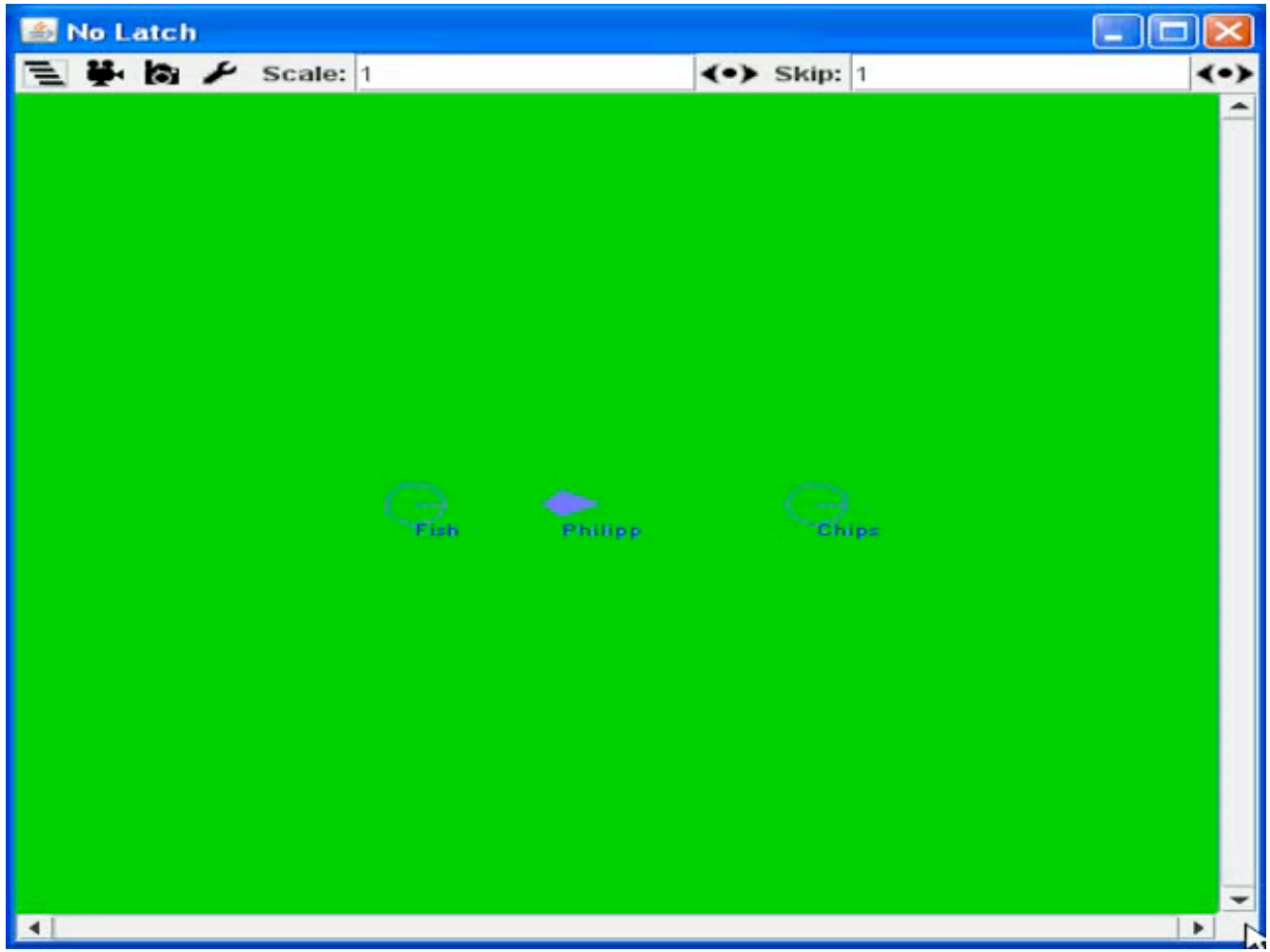
1. No latch
2. Strict latch
   - Trigger behaviour if internal state is below δ
   - Maintain behaviour until internal state is above $\varphi \geq \delta$
3. Strict latching with interruptions; can be very inefficient
   - Agents may persevere for minimum gain
   - Inefficiency first identified by Hagen Lehmann
4. Flexible latch:
   - Introduce a third threshold, ψ such that $\delta \leq \psi \leq \varphi$
   - Behaviour is triggered as before but if agent is interrupted:
     - if internal state is below ψ: continue,
     - otherwise: reset latch

Philipp Rohlfshagen and Joanna J. Bryson, "Improved Bio-Inspired Maintenance of Homeostatic Goals via Flexible Latching", *Cognitive Computation* **2**(3):230-241 2010.

Experimental Questions

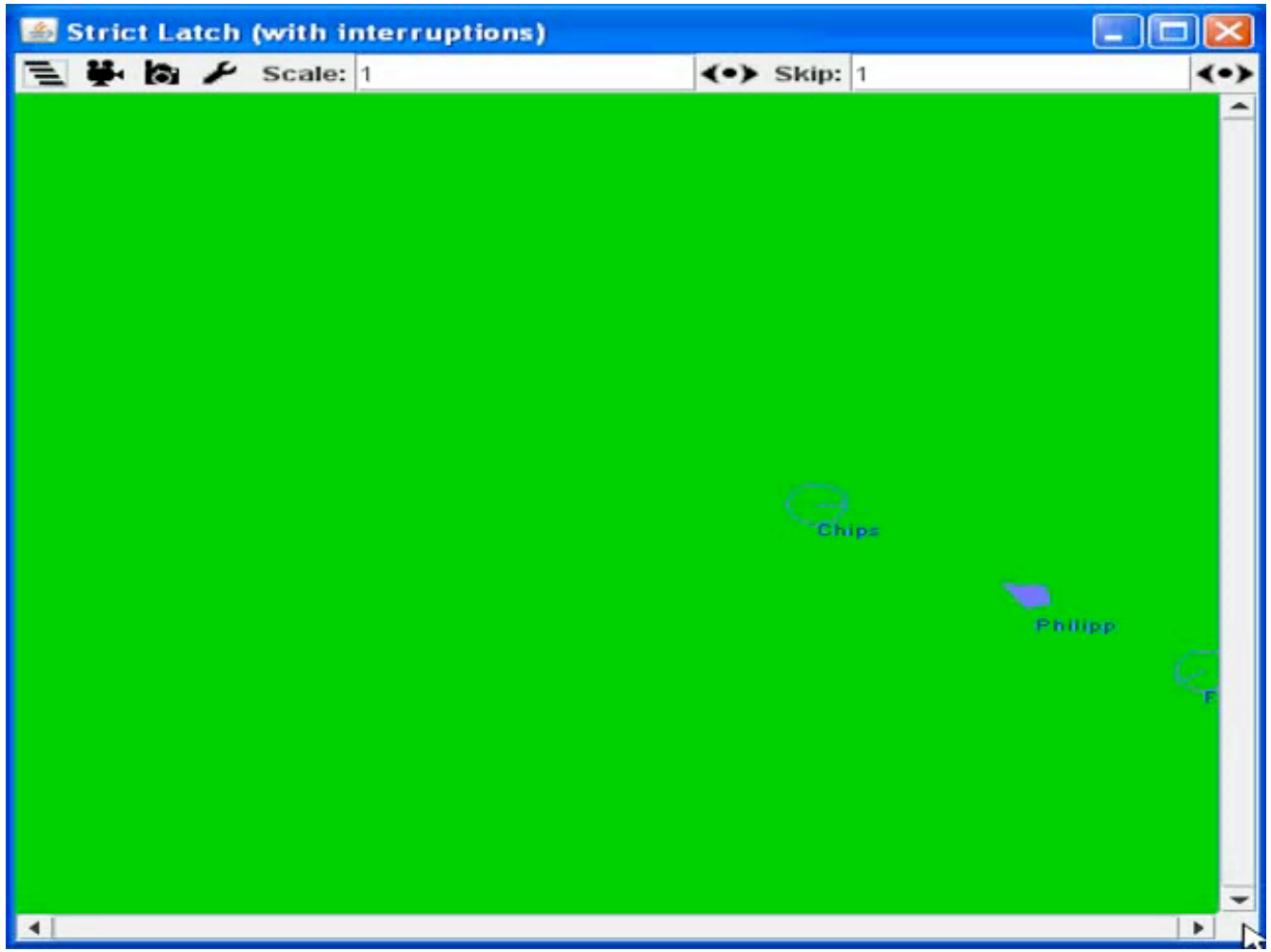
- Does flexible latching increase efficiency?
- If so, what is the optimal value for ψ?
- Experiment with three goals: two required resources + high-level goal of dancing.

No (Emotion-Like) Drives

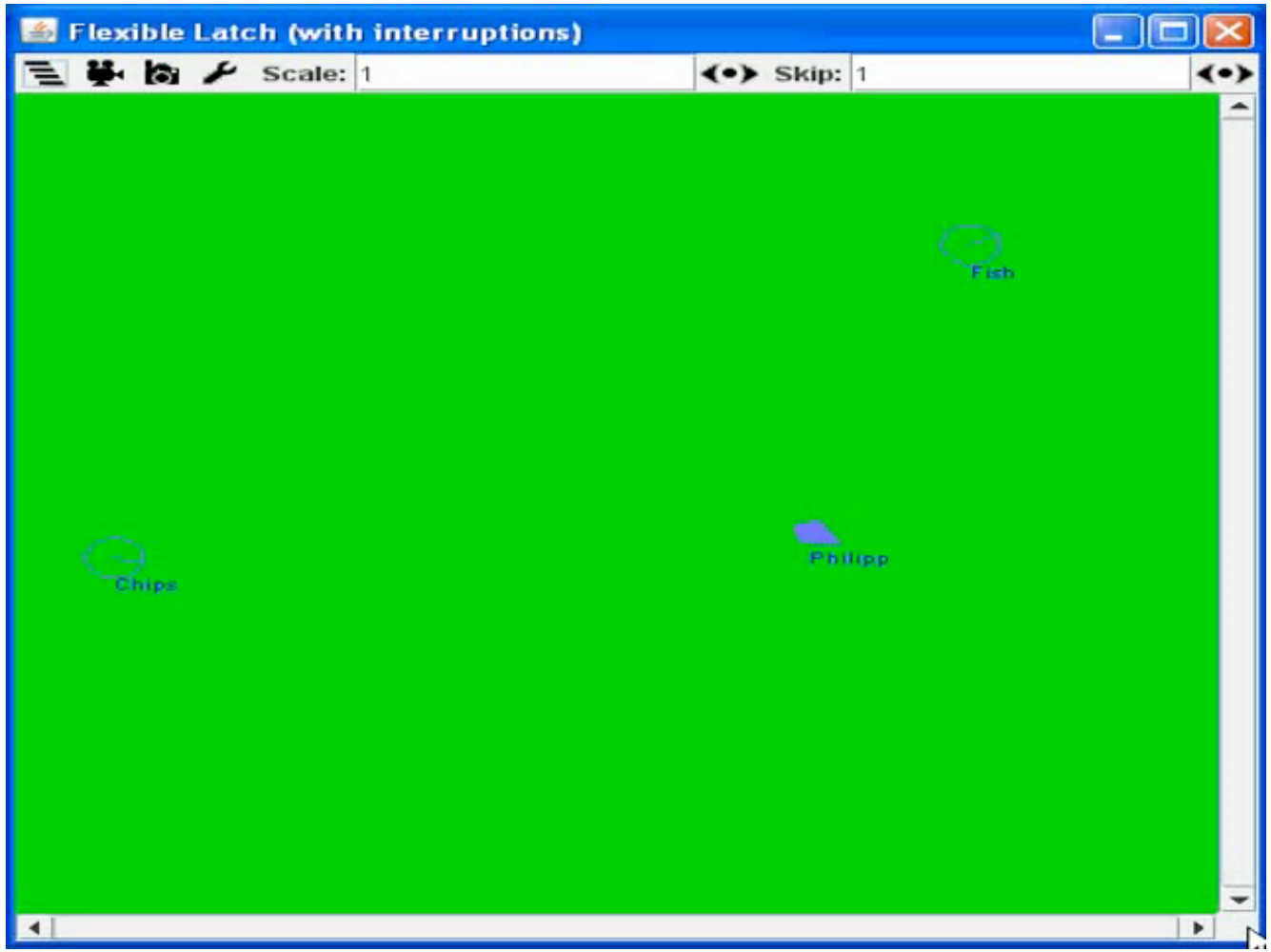


Strict Latches as Drives (no interruptions)

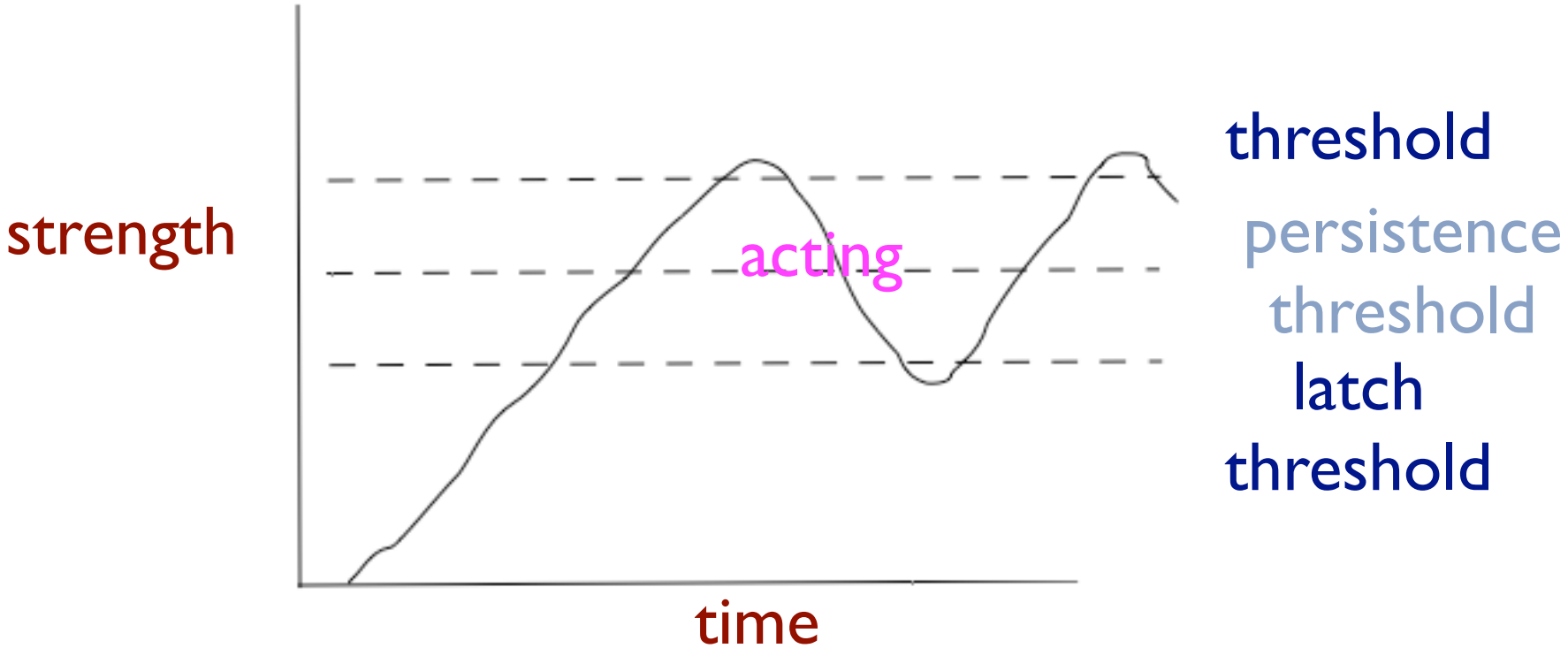
Strict Latch (interruptions)



Improved Representation: Flexible Latch



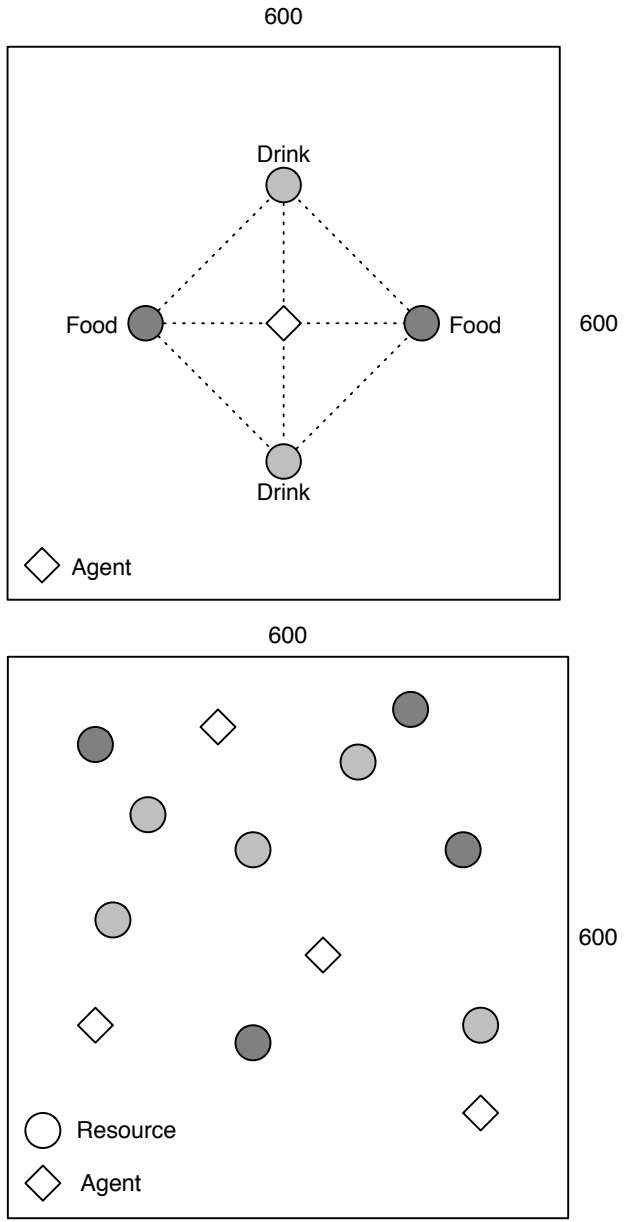
Flexible Latch: Where do you reset?



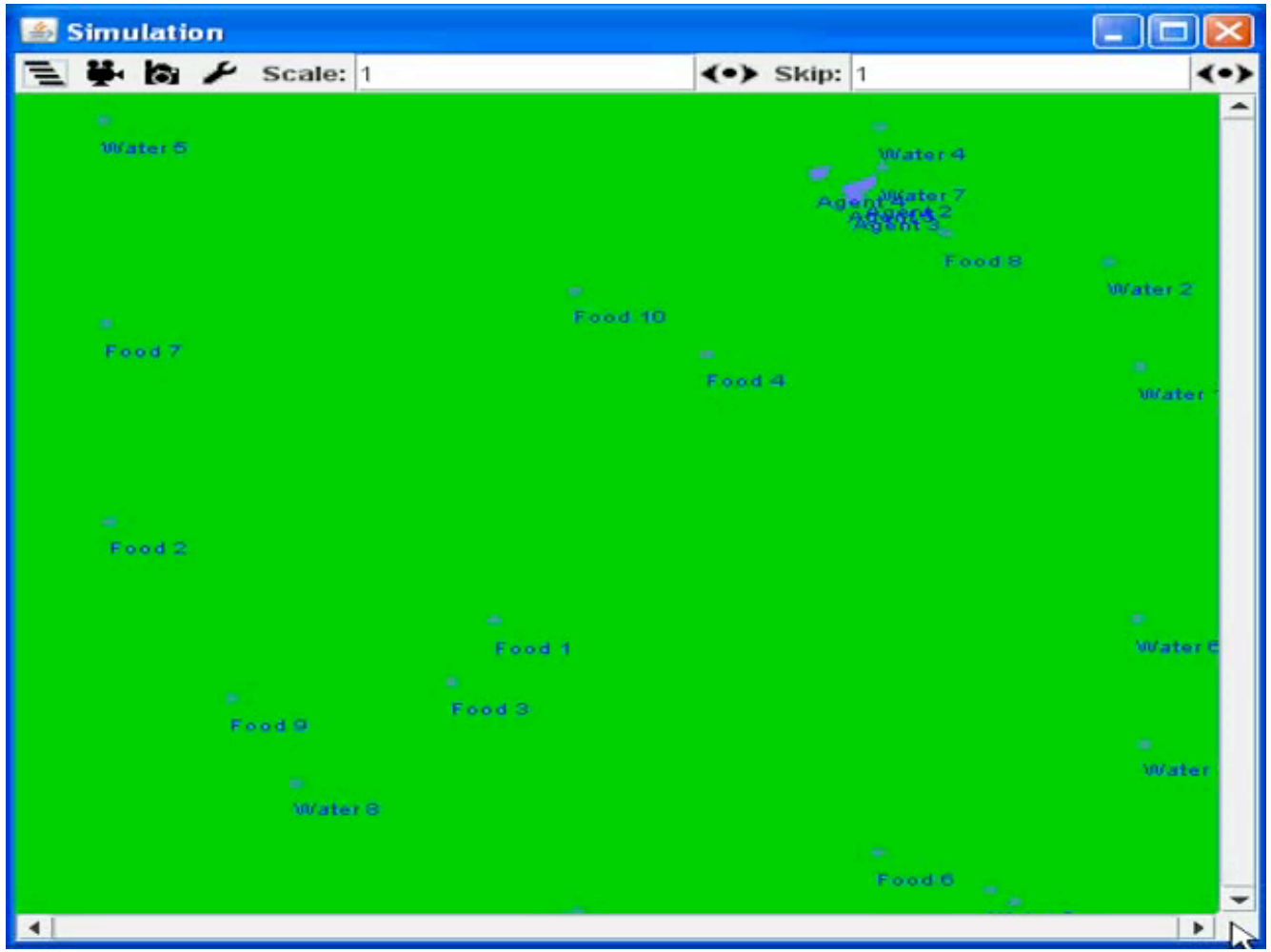
Want to know right threshold for reconsidering current direction if interrupted while acting.

Experiments

- Test and compare all variants
  - Check frequency of execution of low-priority goals
  - Also frequency ratio of primary and secondary actions
- Two simulation settings
  - Controlled environment
  - Random (more realistic environment)

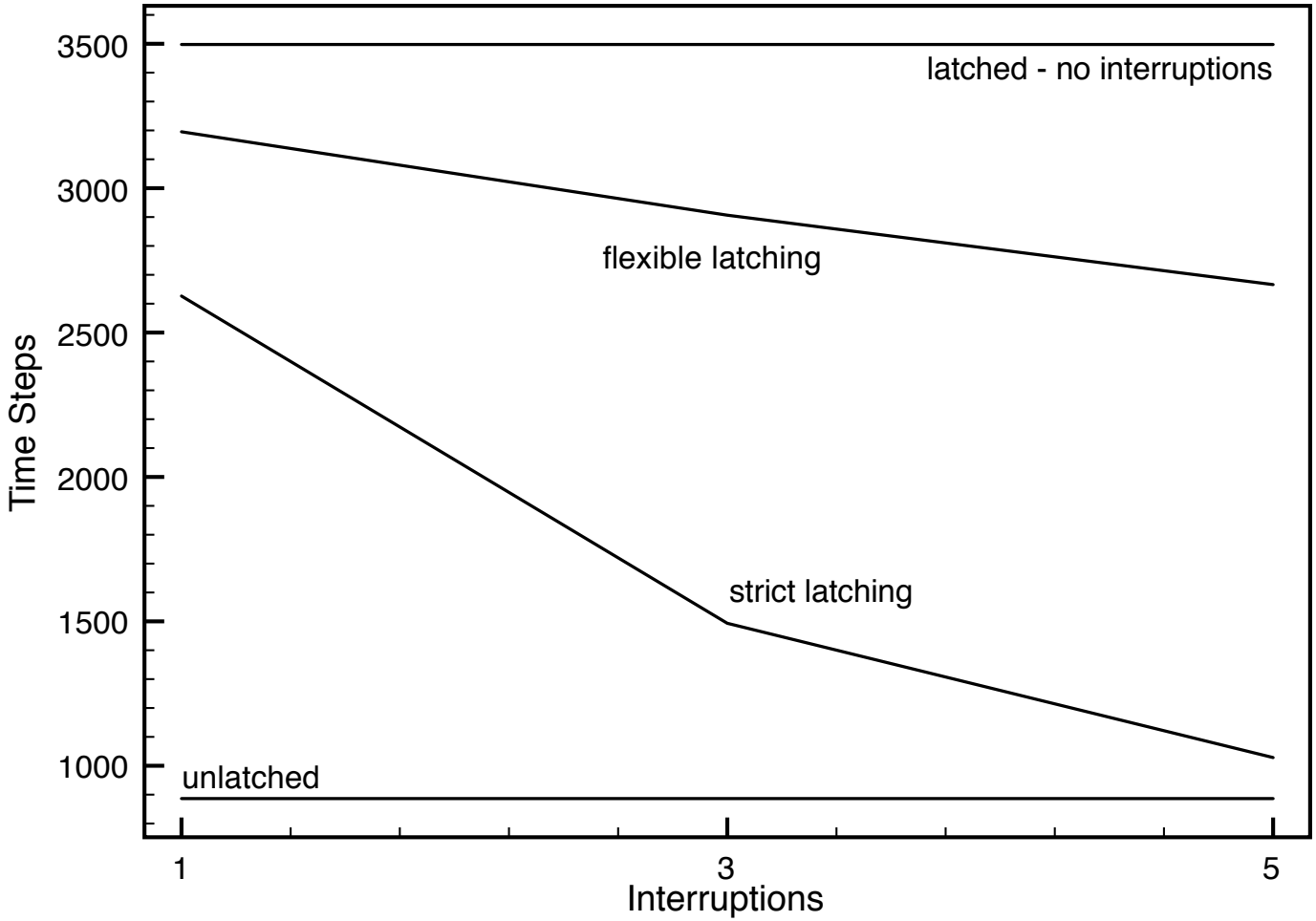


“Natural”


Simulation
Scale: 1
Skip: 1
Water 5
Water 4
Water 7
Food 8
Water 2
Food 10
Food 7
Food 4
Food 2
Food 1
Food 3
Food 9
Water 8
Food 6

Results

time steps available for "low-priority" drives



# of interruptions

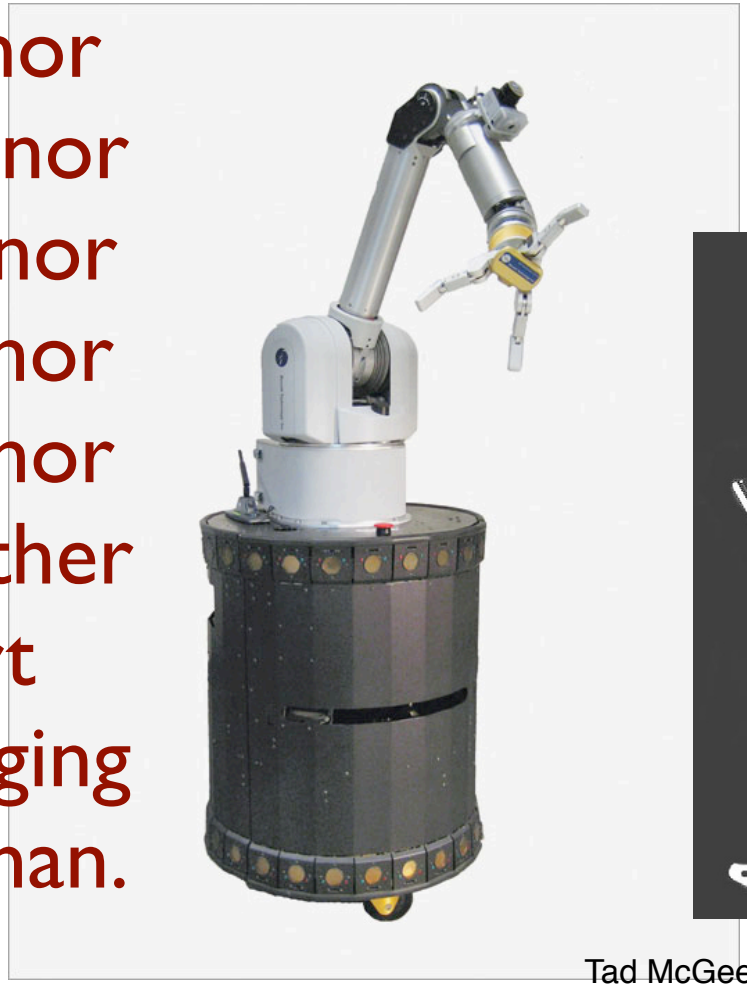
Conclusion

- Sometimes time should be allocated arbitrarily in order to prevent dithering.
- But arbitrary decisions should be easy to revisit.
  - The optimal value for the intervening? persistence threshold = latching threshold ⇒ revisit for all interrupts.

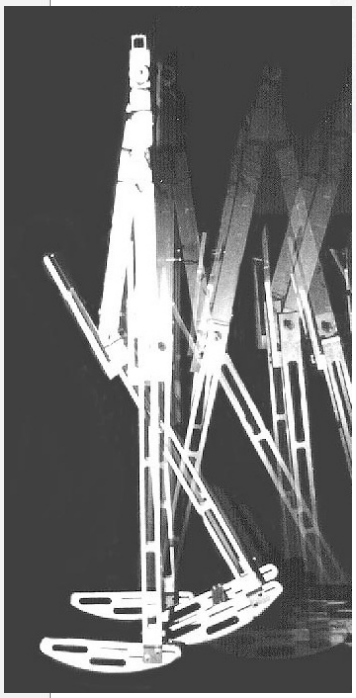
Do the blue diamonds really have emotions?

What's an emotion?

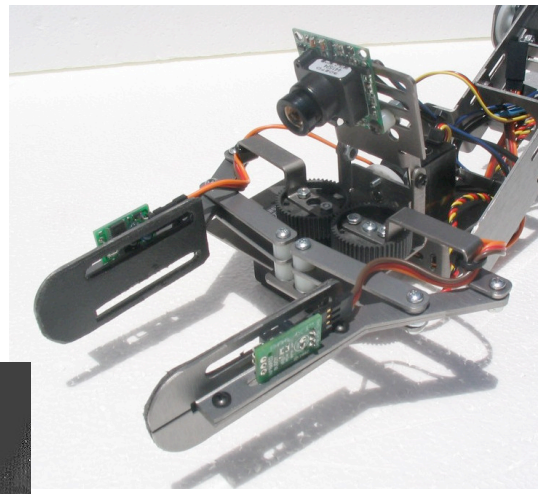
it is nor hand, nor foot, nor arm, nor face, nor any other part belonging to a man.



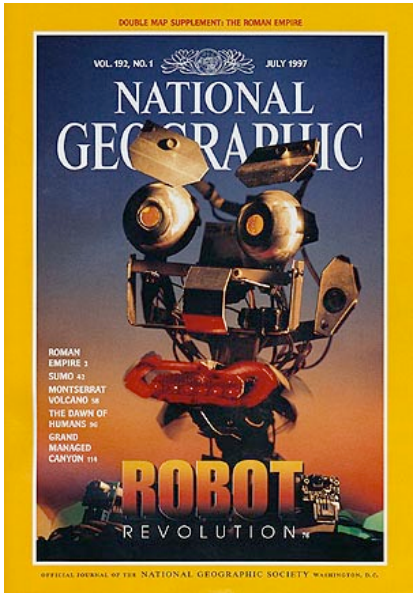
Glenn Matsumura, Wired 2007



Tad McGeer's passive dynamic walker



SG5-UT Robotic Arm

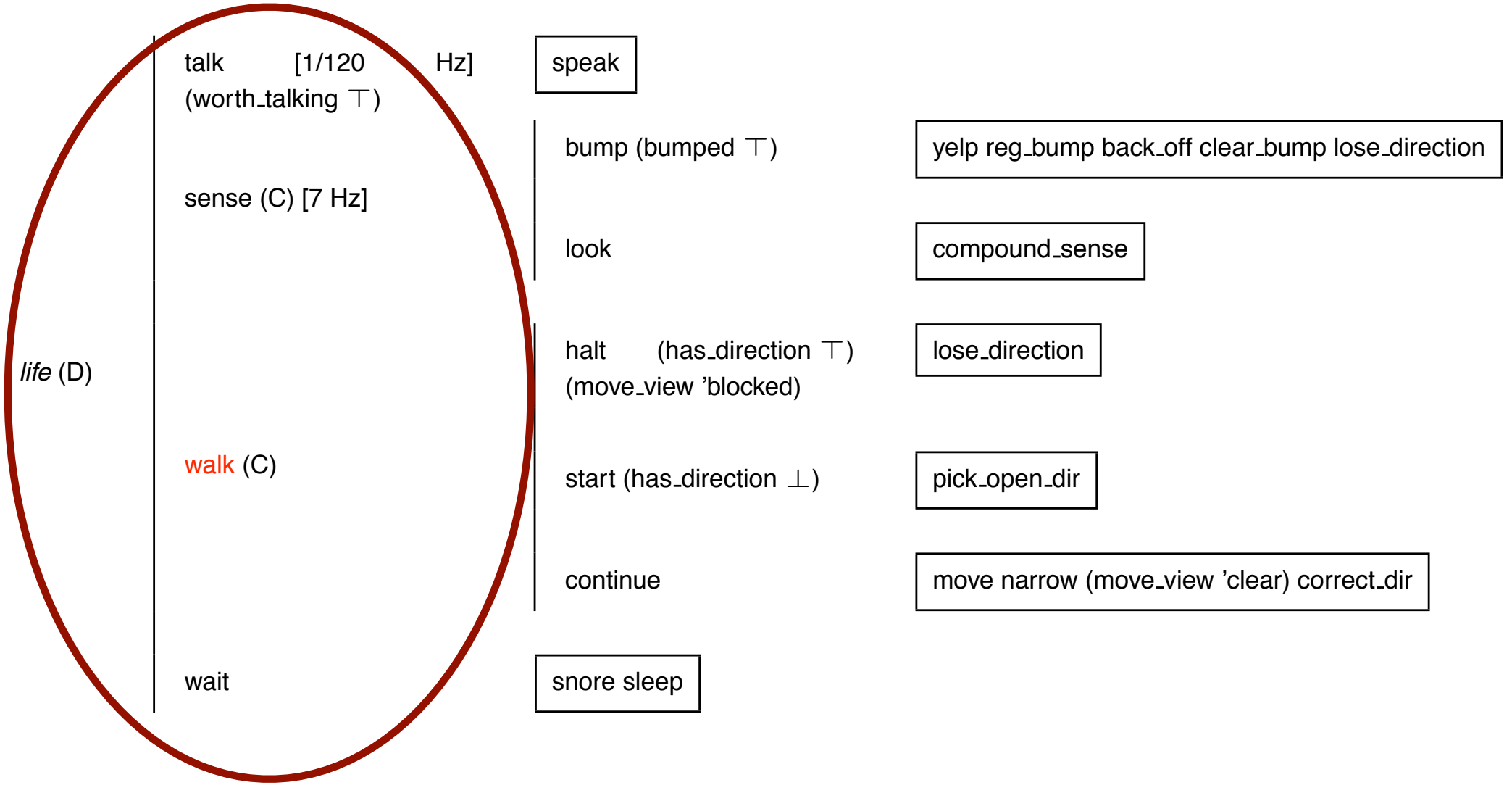


Chuck Rosenberg's IT, 1997

Implementing in BOD

- Create a super-class for behaviour modules containing drive or emotion state and latching logic.
- Add ‘sense’ to detect whether latched.
- Add means and extent to increment & decrement.

drive collection



```
((SDC life (goal (s-one_step (s-succeed 0)))
  (drives
   ((dead (trigger((s-is_dead 0))) a_stay_dead))
   ((drink (trigger((s-wants_drink))) a-drink) (eat (trigger((s-wants_food))) c-eat))
   ((groom (trigger((s-wants_to_groom))) c-groom))
   ((explore (trigger((s-succeed))) a-explore))))

(C a-groom (goal ((s-succeed 0)))
 (elements
  ((has-no-target (trigger((s-has_groom_target 0))) a-pick_groom_target))
  ((not-near-target (trigger((s-is_near_groom_target 0))) a-move_to_groom_target))
  ((default-groom (trigger((s-succeed))) a-groom_with_target))))

(C a-eat (goal ((s-succeed 0)))
 (elements
  ((has-no-food (trigger((s-has_food 0))) a-pick_food))
  ((not-near-target (trigger((s-is_near_food_target 0))) a-move_to_food))
  ((default-feeding (trigger((s-succeed))) a-eat))))

(C a-drink (goal ((s-succeed 0)))
 (elements
  ((has-no-drink (trigger((s-has_drink 0))) a-pick_drink))
  ((not-near-target (trigger((s-is_near_drink_target 0))) a-move_to_drink))
  ((default-feeding (trigger((s-is_near_drink_target))) a-drink)))))
```

- Drives can (& if they are, should) have same priority.
- mux due to latch state.
- (Rohlfshagen labelled senses & actions.)

Emotions and Coherence

- Fully reactive AI can change state much too rapidly to be comprehensible.
- Humans read goals largely by emotional facial expressions.
- Artificial emotions can be used to make reactive systems more comprehensible, easier to use.

(Sengers 1998,1999)

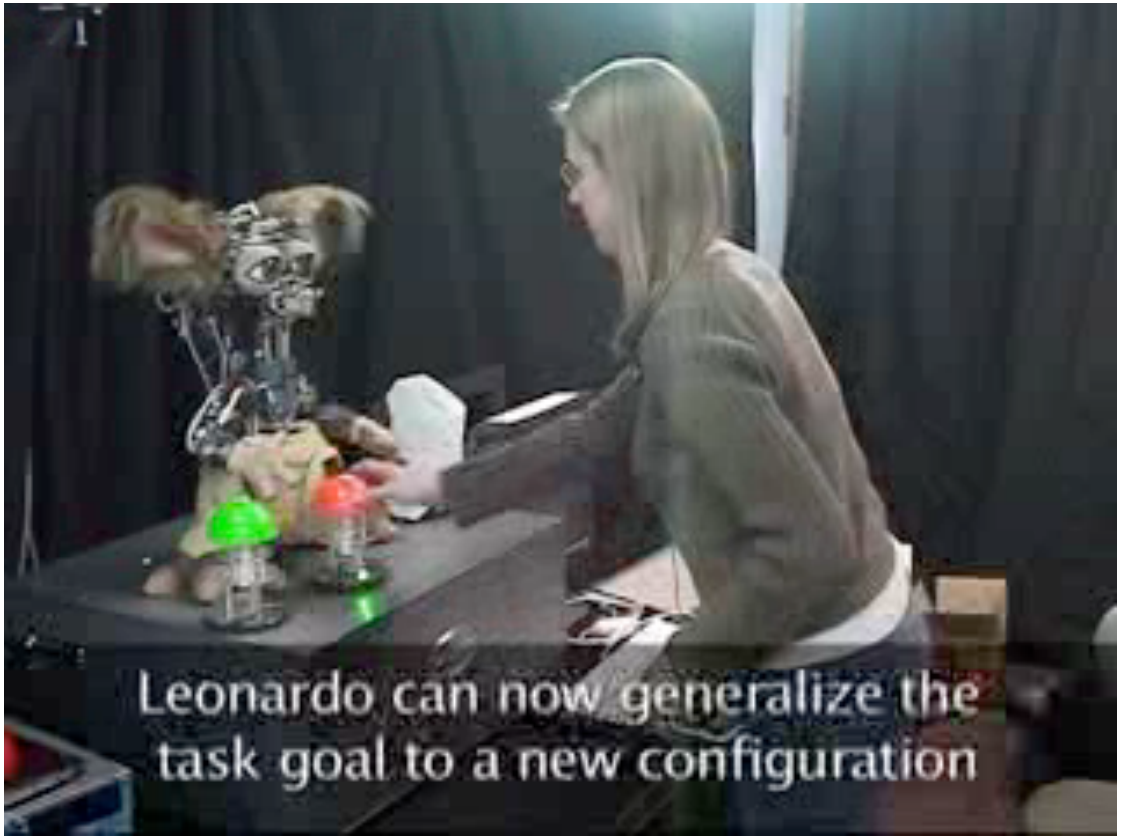
Emotions in Cognitive Systems

- Natural Cognition
  - Organise behaviour / provide another different kind of control state.
    - This includes social behaviour.
- Artificial Cognition
  - Organisation, communication, variation.

Emotions as Communication

- Most AI emotion systems are for HCI:
  - Increasing engagement,
  - Increasing believability,
  - Facilitating comprehensibility.

What are emotions doing for this robot?



Andrea Thomaz at MIT (she's now at Georgia Tech)

Emotions for Human-Robot Interaction

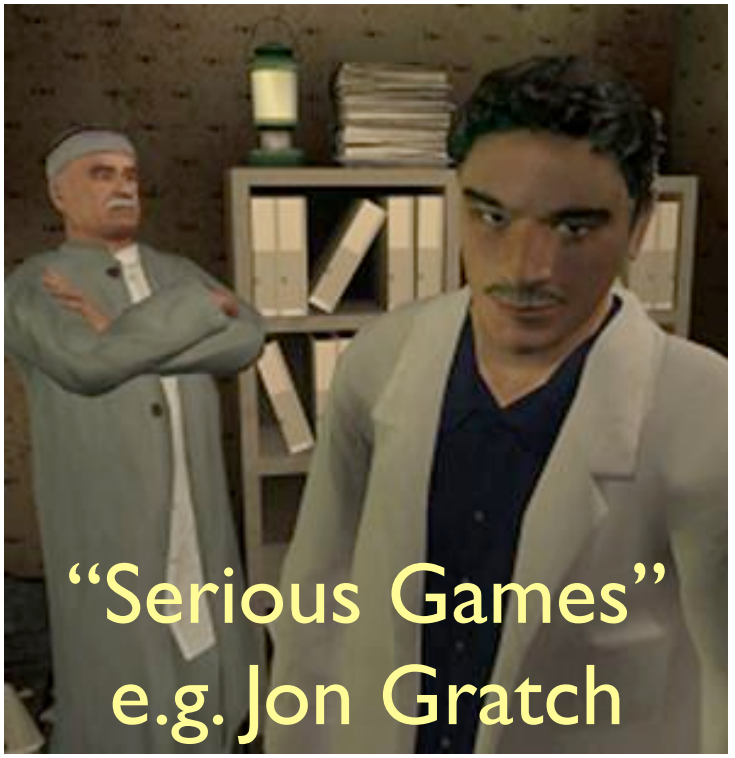
- Humans have very complex social lives, with associated skills and emotions.
- We read other's emotions to interact correctly.
- Hypothesis: we need the same interface for robots if we are to work or live with them.
- These emotions could be "fake", not aligned with real goals (just like for humans).

Robots for Human Emotions

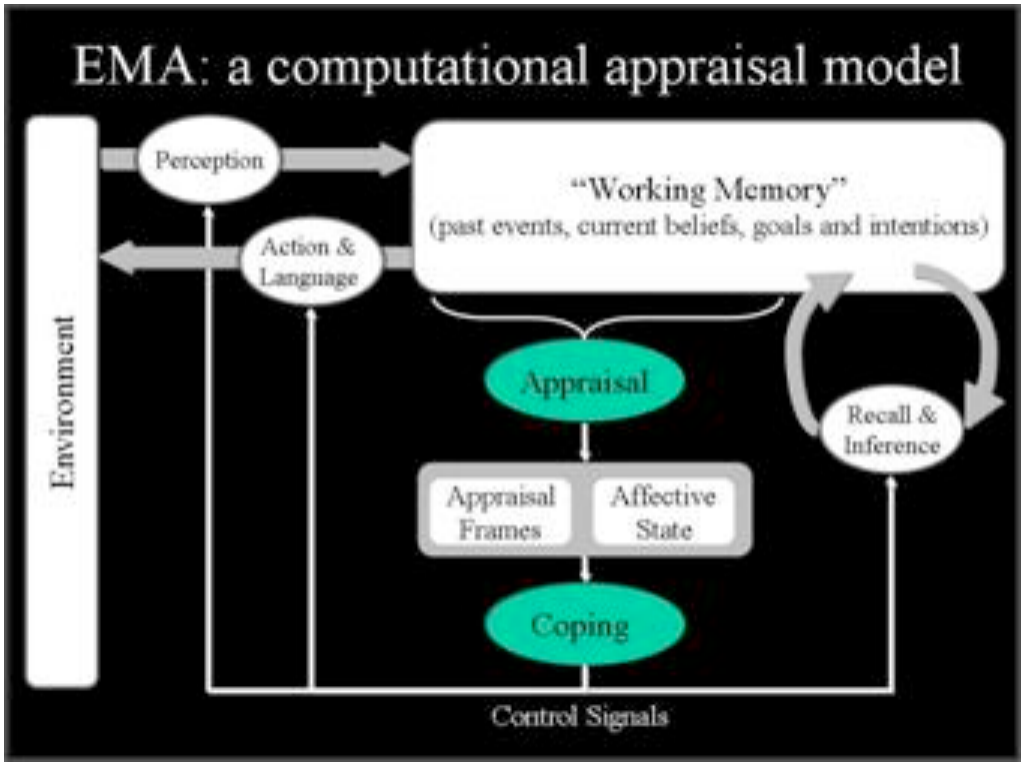
- Some robots are specifically designed to address attachment issues.
- Claim: maintaining engagement is not just for sales, but also necessary for therapy.

Paro "Robot Seal Healing Pet"

Emotions in Games

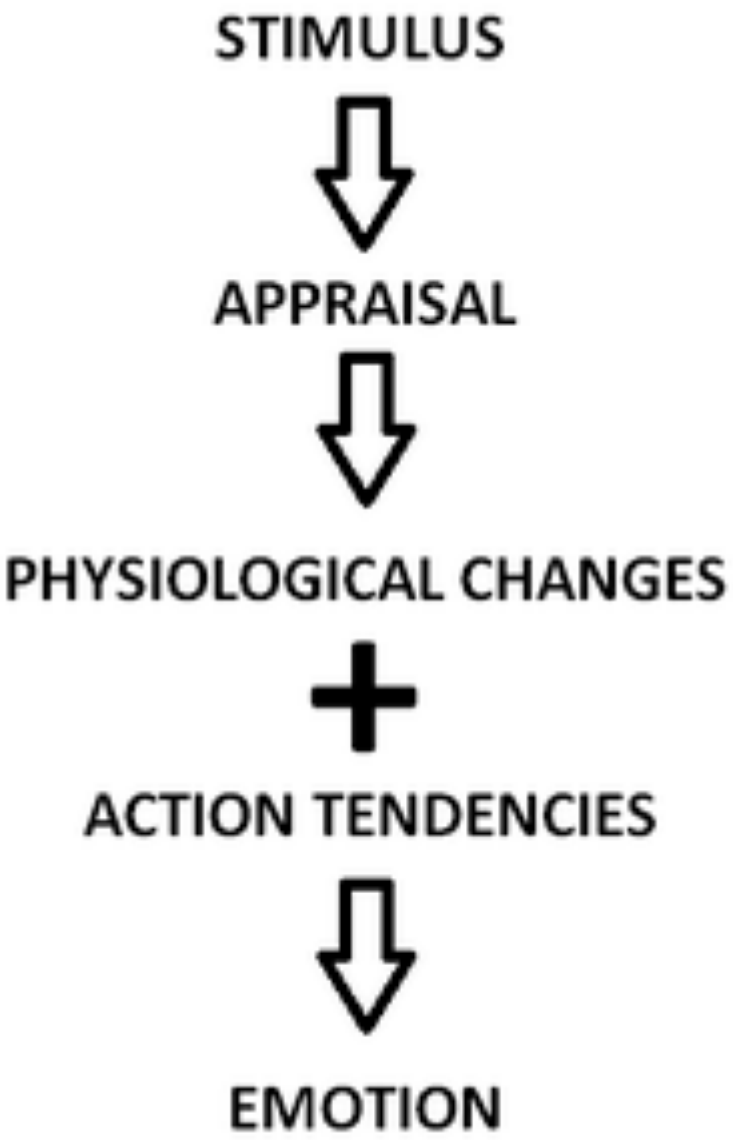


Fable / Lionhead

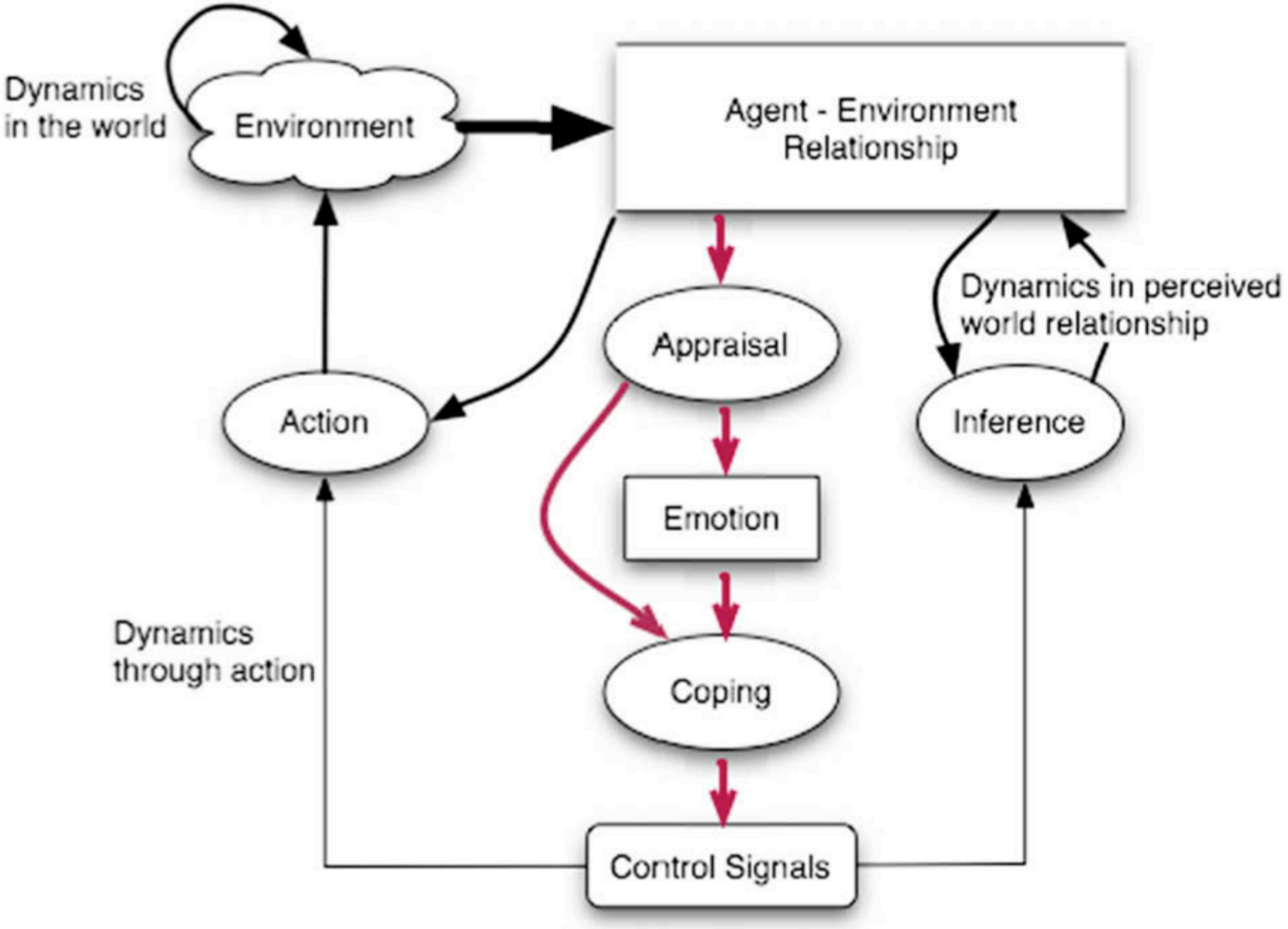


Appraisal

- Problems for AI: what stimuli should trigger which emotions?

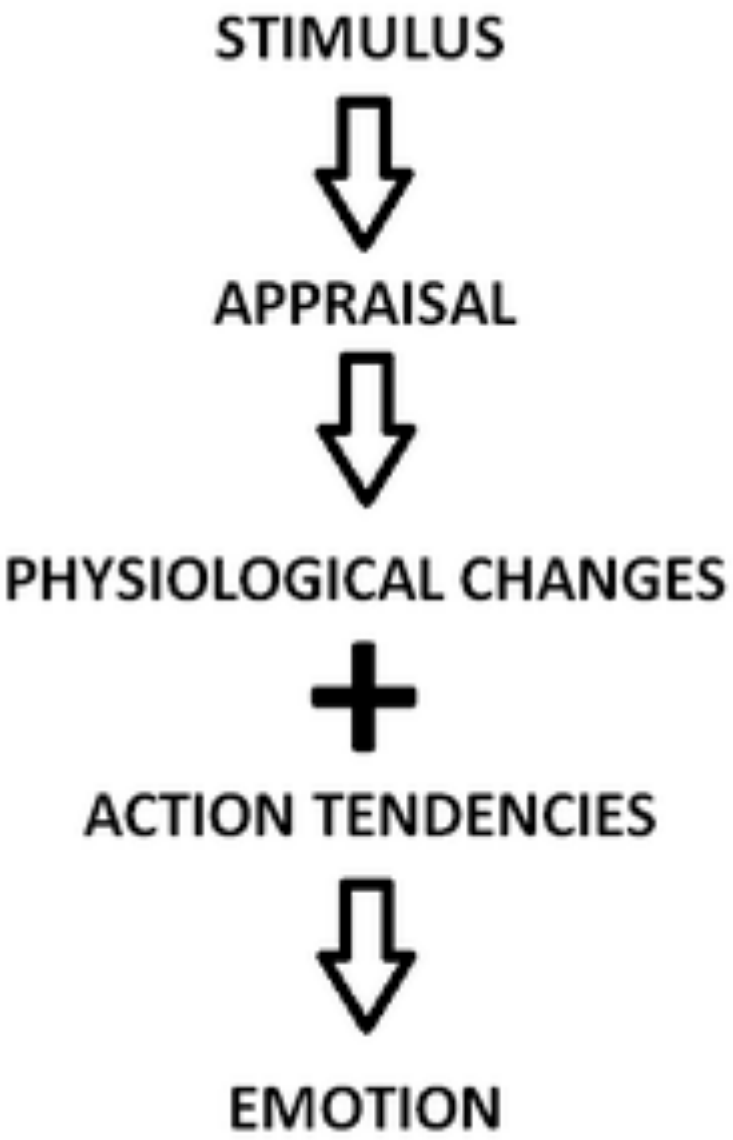


Psychology theory, based partly on human pathologies.


Dynamics
in the world
Environment
Agent - Environment
Relationship
Appraisal
Dynamics in perceived
world relationship
Action
Inference
Emotion
Dynamics
through action
Coping
Control Signals

Appraisal

- Problems for AI: what stimuli should trigger which emotions?
- Figuring out which emotions there are is necessary for this.



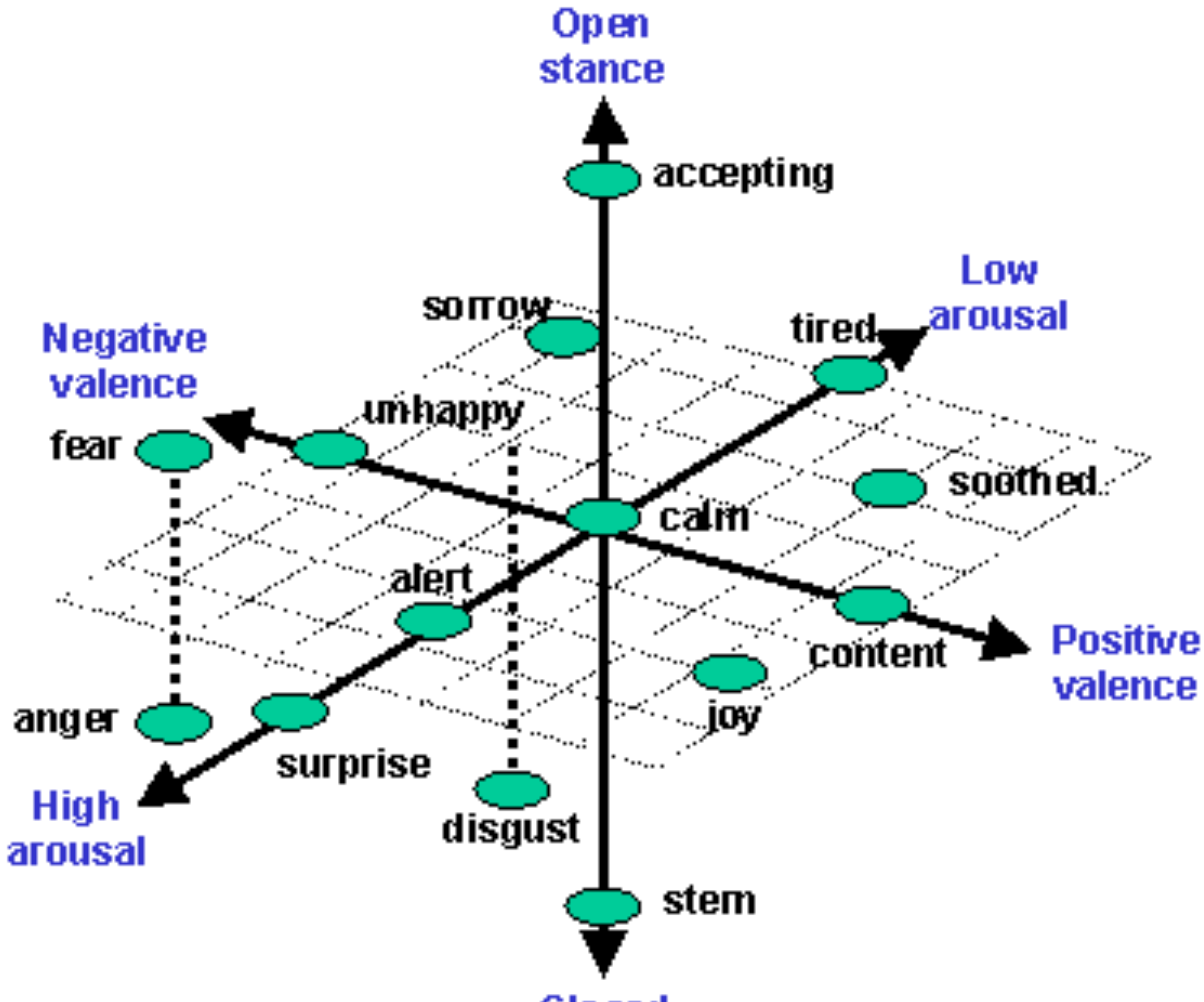
Psychology theory, based partly on human pathologies.

Standard Approaches

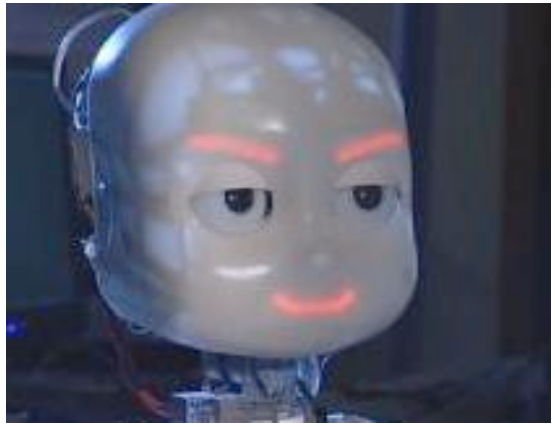
- Darwin hypothesised that emotional expressions were universal.
- Eckman (1978) described such using Facial Action Coding System (FACS)

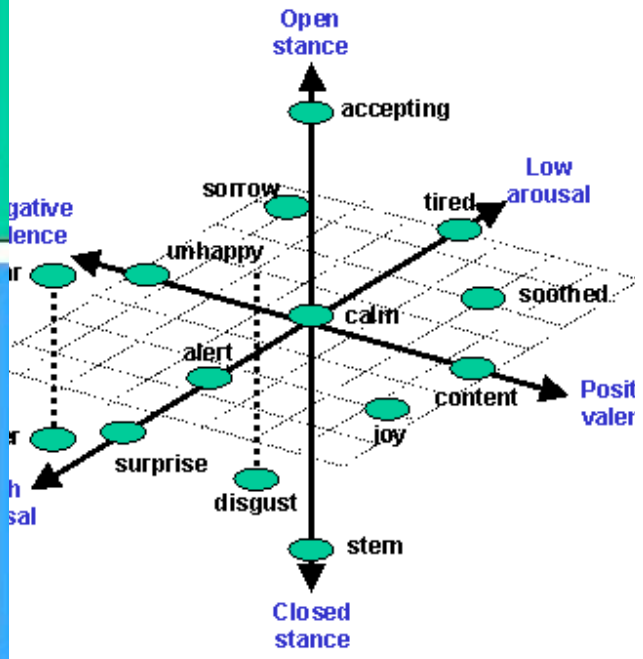
Done by actors

Representation – Discrete or Dimensional?



c.f. Hamann 2012, TiCS





Brezeal (2002)

Open Questions

- How many different emotion systems / axes are there underlying the space?
- How much of emotional experience is a consequence of cultural construction?
- Remember: we are category-learning machines. What would it be like if we didn't have labels for emotions? What is it like to experience emotions without labels?

Gesture and temporal coherence

- There's a lot more to communicating emotion than facial expression.
- In games, much research on posture.
- Also an issue in emotionally-neutral contexts.

Emotions in Cognitive Systems

- Natural Cognition
  - Organise behaviour / provide another different kind of control state.
    - This includes social behaviour.
- Artificial Cognition
  - Organisation, communication, variation.

Emotions as Memory

- Recent events:
  - episodic memory,
  - emotions.

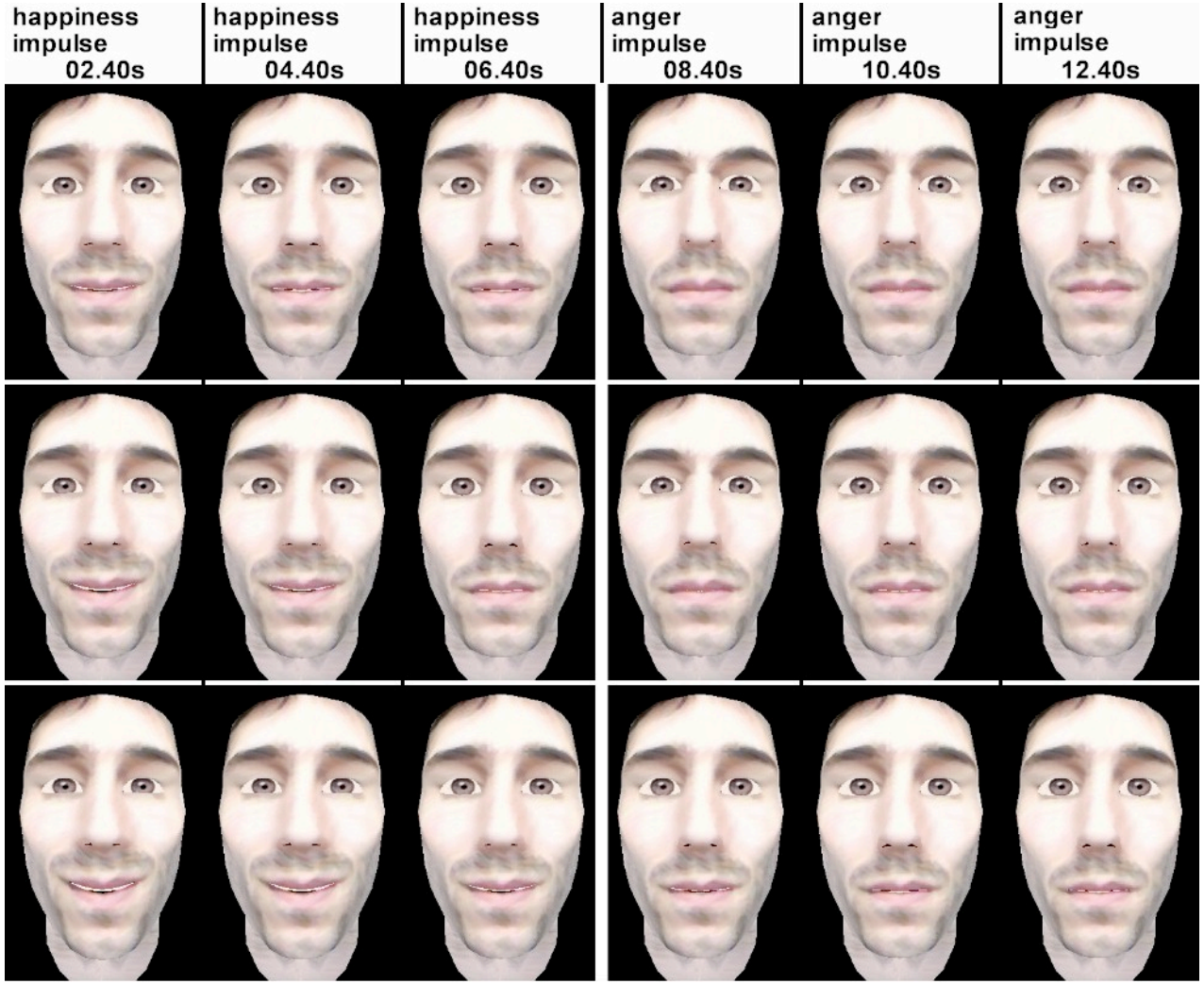
} These fade, get replaced.

- "Knowledge":
  - facts,
  - expectations.

} These only build (more or less).

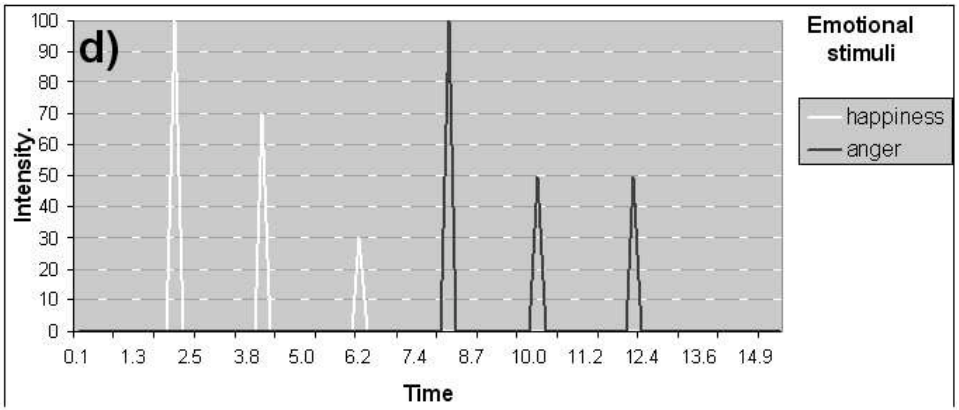
Emmanuel Tanguy, Philip Willis and Joanna J. Bryson, "Emotions as Durative Dynamic State for Action Selection", in *The Twentieth International Joint Conference on Artificial Intelligence (IJCAI)*, Hyderabad, India, pp. 1537–1542, Morgan Kaufmann 2007.

Example:
Emotions as Memory

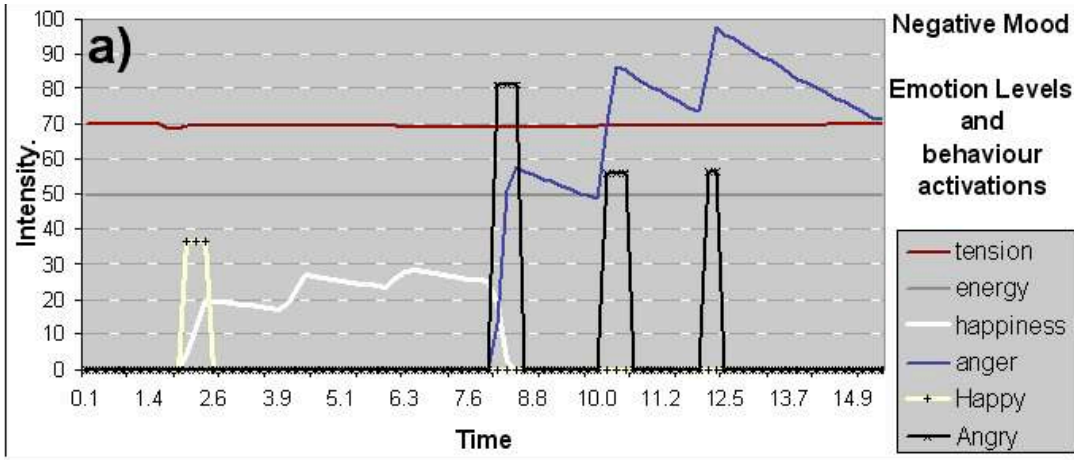


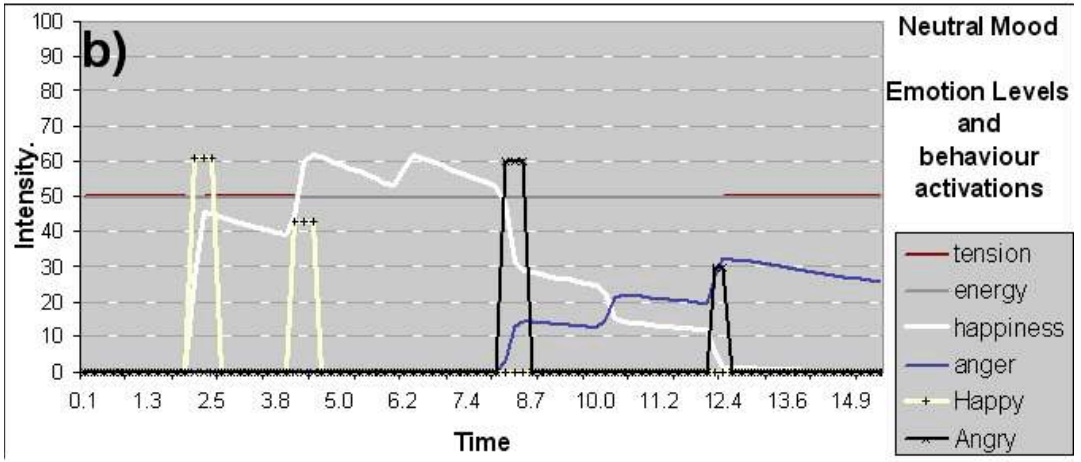
Tanguy (2006)

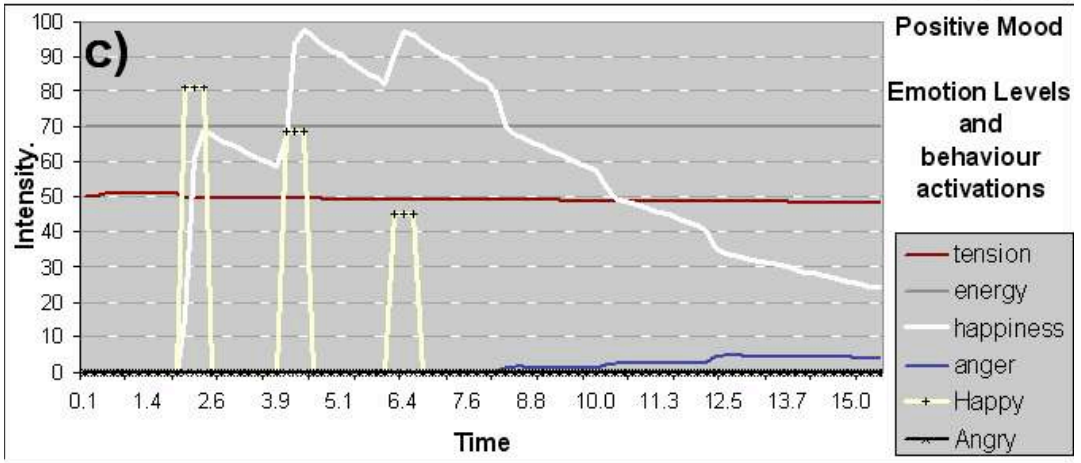
(Tanguy, Bryson & Willis 2007;
Bryson & Tanguy 2010)



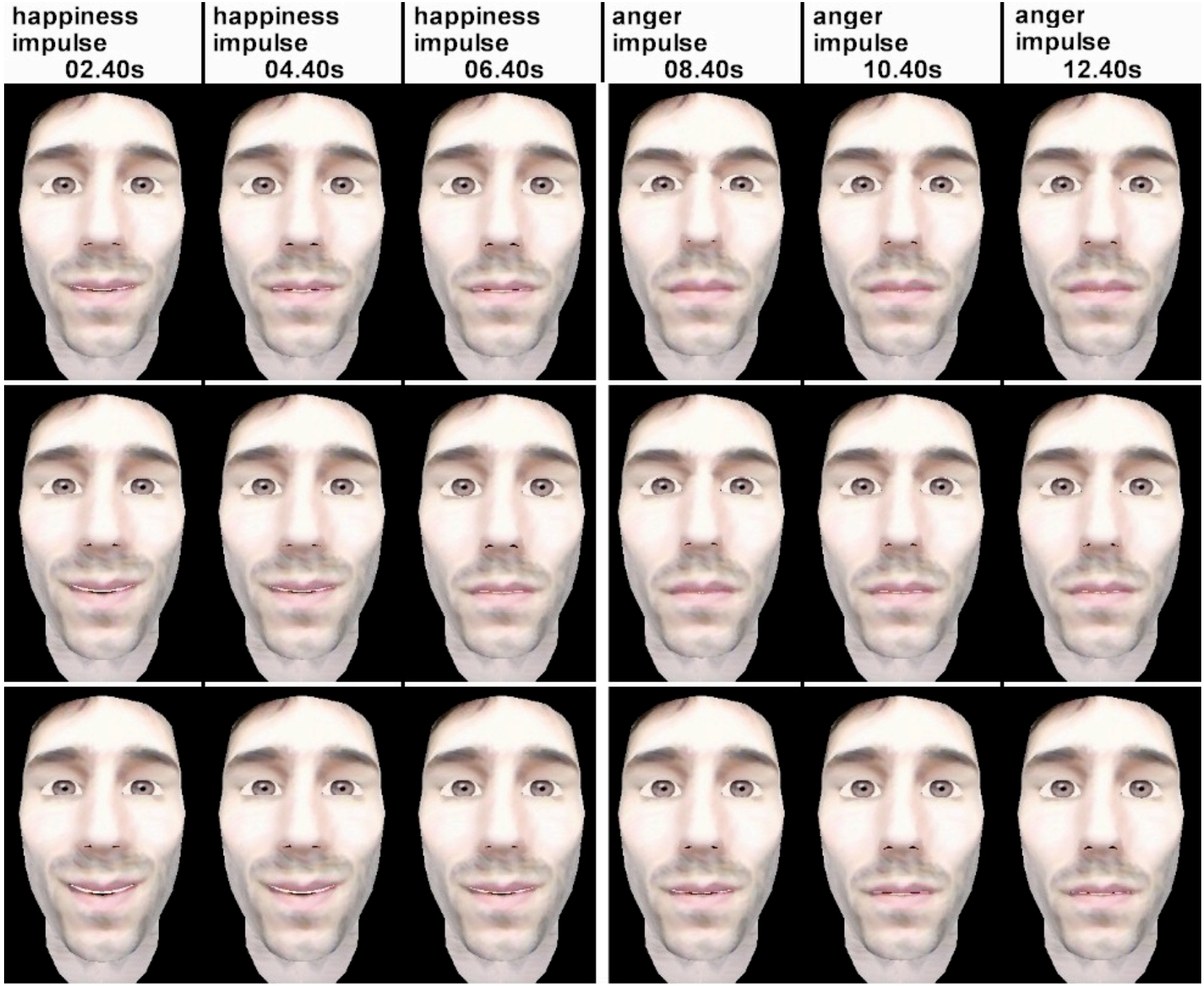
I've got good news
and bad news...







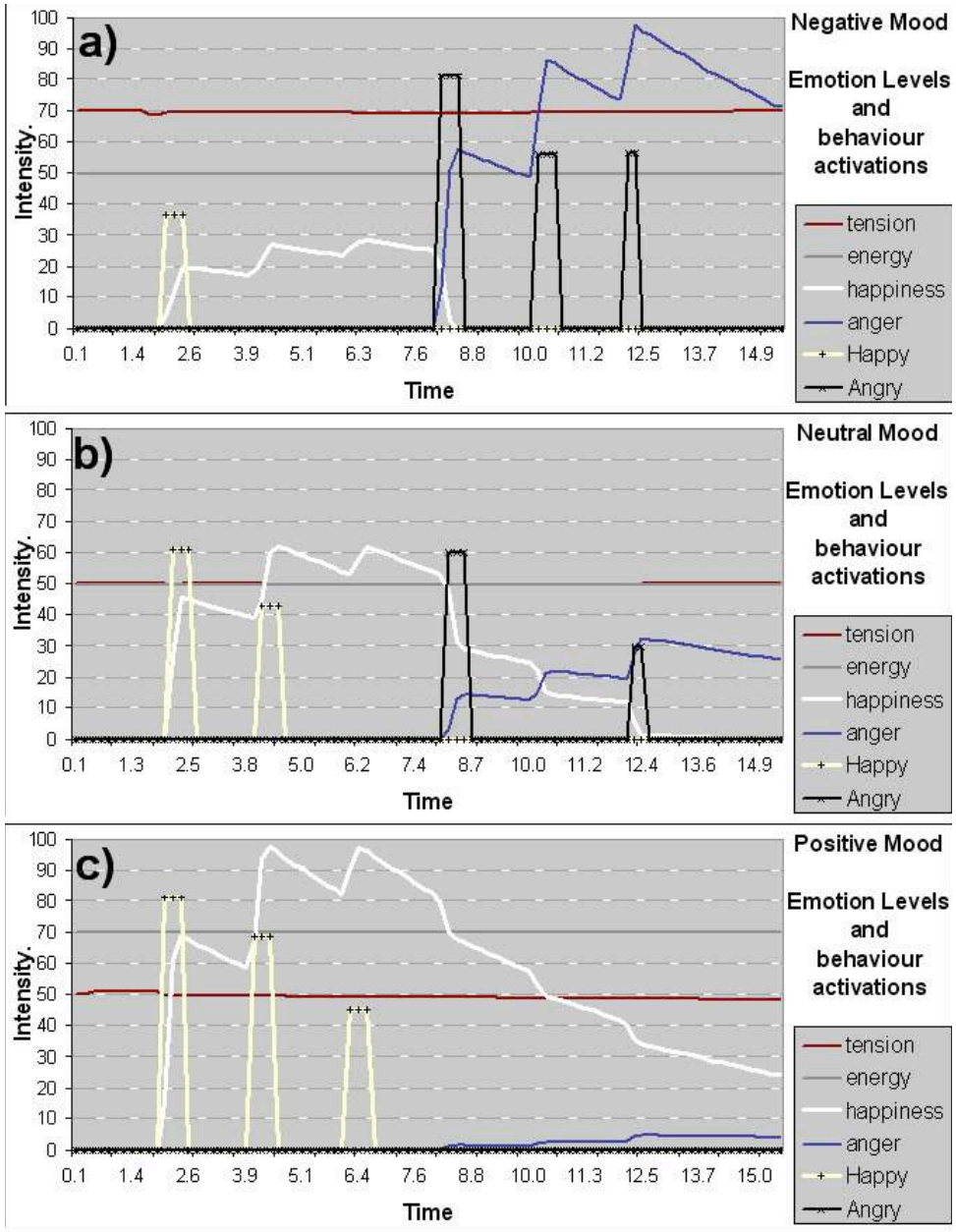
Code & video available online.


happiness impulse 02.40s
happiness impulse 04.40s
happiness impulse 06.40s
anger impulse 08.40s
anger impulse 10.40s
anger impulse 12.40s

AI Emotions

- Mood — longer term.
- Emotions — shorter term.
- Behaviour (e.g. expressions) is altered by these.

- Simplifies coding,
- increases variability.



Summary

- Emotions and drives in nature are key to coherent behaviour – another form of focus.
- Emotions in AI are mostly used for believability and engagement.
- Can also be used to add interest / variability by creating situation-dependent context for action selection.

Thanks!

Philipp Rohlfshagen

Their work supported by EPRSC grant GR/S79299/01

Extra Vocabulary

- 2011
  - phototaxis
  - (non)holonomic motion
  - Braitenberg's *Vehicles* (1984)
- 2012 mux :: mutually exclusive :: xor