Institute for Systems and Robotics

Emotion-based Agents:
putting the puzzle together II

Rodrigo Ventura

http://www.isr.ist.utl.pt/~yoda

email: yoda@isr.ist.utl.pt

- Overview
  - Antonio Damasio: emotions are fundamental for appropriate decison-making
    - long-term
    - without anticipating all possible consequences
  - Somatic marker hypothesis
    - sensory images associated with body states
    - “gut feeling”
    - hunches towards-to / away-from alternatives
  - Cognition machinery evolved on top of emotion and regulatory mechanisms
    - two intertwined levels

- Overview
  - Anticipated emotions
    - consequentialist

      vs
  - Anticipatory emotions
    - brief changes in the body state

      e.g., SCR

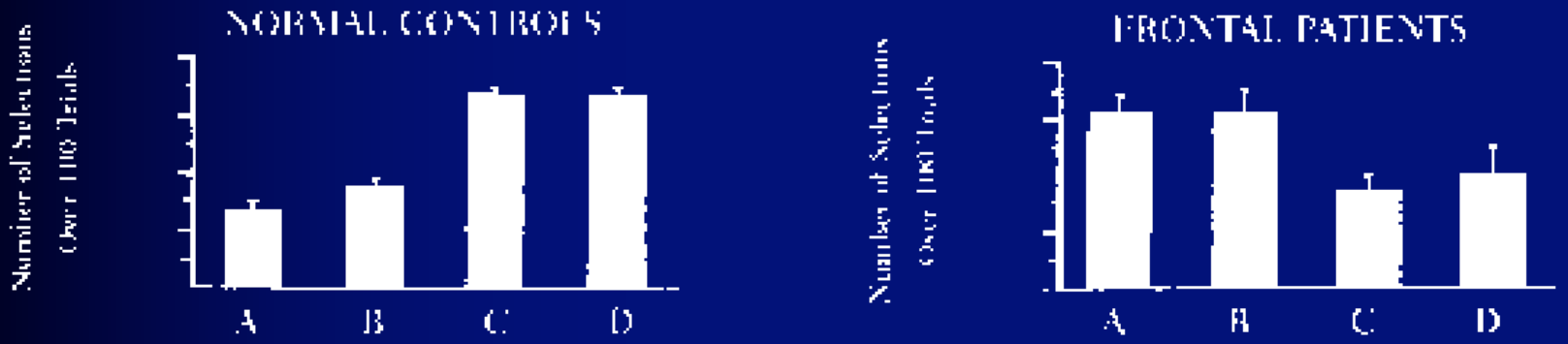
  - Two levels of decision-making:
    - cognitive – expected utility, outcome probability, rationality
    - emotional – insensitive to probabilities, anticipatory emotions, e.g. fears, panic, phobias, risk, gambling

- Examples

  – 4 decks of cards: A, B, C, and D

    A, B – gain $100, occasional high losses (≈$1250)

    C, D – gain $50, occasional small losses (<$100)



- "future myopia"
- all showed SCR after punishment/reward
- only normal patients showed anticipatory SCR

- Examples

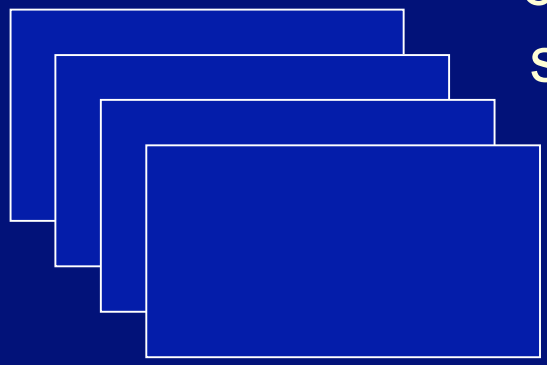
  – the good-guy/bad-guy experiment

    - patients with lesions in memory areas (LTM)

      remember the movie Memento?

    - emotional memory / no recognition

- Long-term research goals
  - Objective: to cope with complex and dynamic environments
  - Roles of the emotional system
    - relevance
    - meta-management
    - intuition

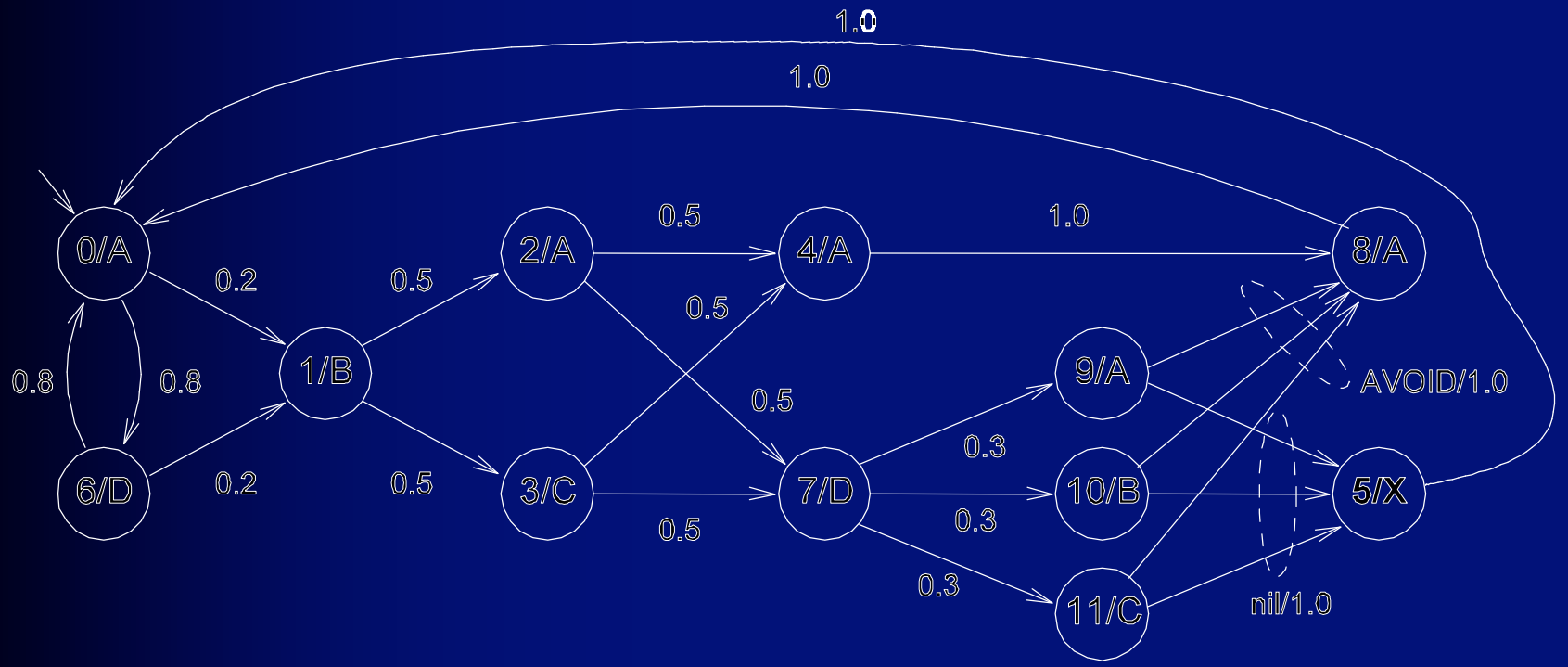


- Emotion-based agent model components
  - Two levels of representation:
    - cognitive: rich, complex, slow to process
    - perceptual: rough, simple, fast to process
  - Desirability vector (DV) – components represent various aspects, e.g. positiveness, negativeness, relevance, etc.
  - Homeostatic vector – representation of the body state, drive towards equilibrium, motivation

- Emotion-based agent model mechanisms
  - association between cognitive and perceptual images, and the DV – *somatic marking*
    - when/how to establish these associations
    - when/how to utilize them
  - to use a perceptual (simpler) representation to facilitate the search for matching cognitive representations – *indexing*
  - to confront the two representations when faced with a situation – *an extra dimension*

- Causal models
  - Goal: formulation and refinement of causal models, during interaction of the agent with the environment
  - Bootstrap: built-in association among certain stimuli and certain desirability vectors
  - Testbed: hidden MDP, designed such that there are world laws that allow the (deterministic) anticipation of desirable/undesirable states

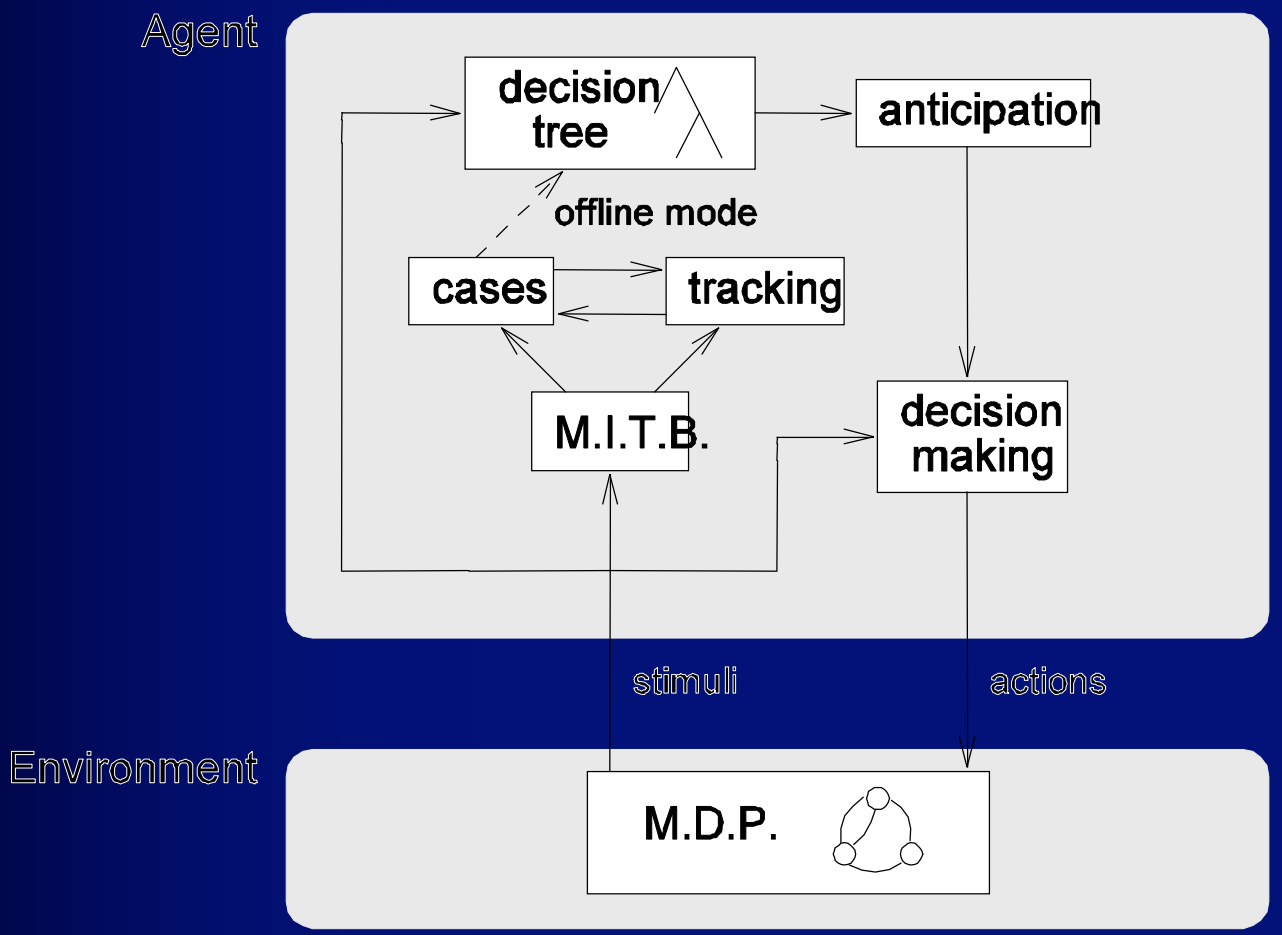
- Example



  - bootstrap: symbol ‘X’ means undesirable
  - world law: sequence [B, *, D, *] anticipates the X

- First approach
  - online mode:
    - selectively collect sub-sequences
    - use the causal model (if any) to anticipate and act
  - offline mode:
    - construct / refine the causal model
  - causal model implemented with a decision-tree (C4.5 to build it; *ad-hoc* algorithm to refine it)

- Implementation architecture



– it works, but hard to scale in complexity

- Second approach
  - Ideas:
    - to use cognitive representation to hold *causality*
      - ‹ cause, (in)action › → effect
    - to use perceptual representation to hold *evaluation*
      - in particular, an effect from a ic can only be evaluated using the perceptual memory
    - ic is an under-generalization
    - ip is an over-generalization
    - use online mode to generate hypothesis, and collect statistics
    - use offline mode to generalize ic's, and to get rid of useless information – ip's, indexes, etc.

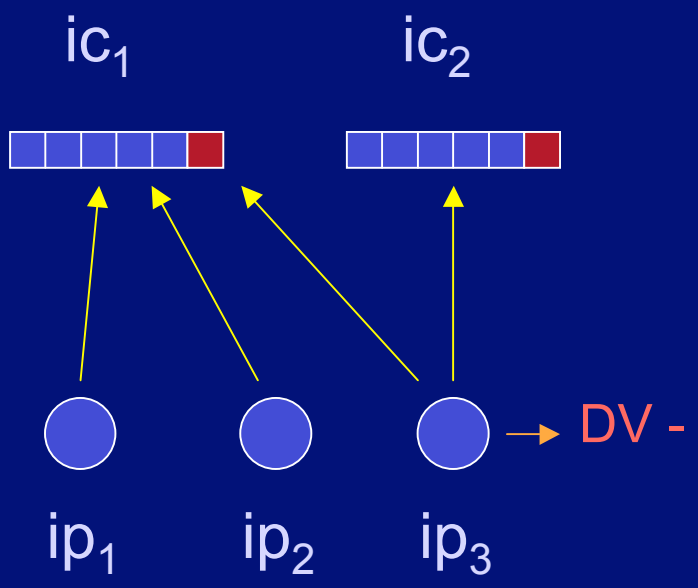
- Second approach

  – cognitive image (ic):
    - sub-sequences templates – cause and effect

  – perceptual image (ip):
    - set of features
    - indexes a set of ic's
    - association with DV's
    - built-in associations:

      'X' → DV-

      'Y' → DV+

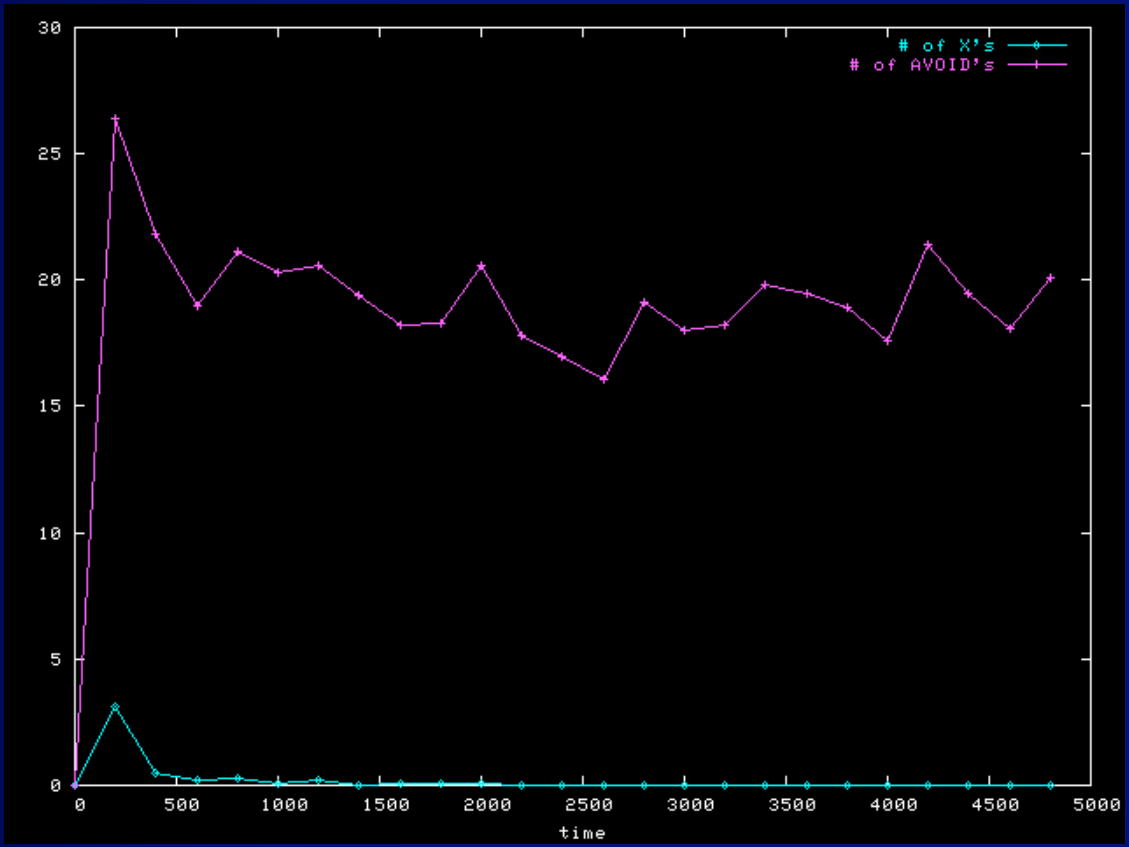


- Second approach

  - Online algorithm:
    1. evaluate stimulus (perceptual memory)
    2. anticipate future outcomes (cognitive memory) and evaluate them (perceptual)
    3. decide according to anticipations
    4. update past anticipations

  - Offline algorithm:
    - generalize ic's
    - eliminate useless ip→ic associations
    - eliminate useless ip's

- Second approach
  - Results (5000 step trial):
    - free-run: about 329 ‘X’s
    - agent: about 4 ‘X’s, about 348 actions



- Second approach

  - Results:
    - in simple environments, it works
    - ...still hard to scale to complex environments

  - What do we gain from using the double-representation framework?
    - efficiency – *indexing*
    - relevancy – *ip features*
    - confrontation of two representations with different degrees of specialization
    - intuition and meta-management – *to be explored*

- Current and future work
  - to implement a mechanism of anxiety (anticipatory emotions)
  - meta-management of cognitive processes driven by emotions: e.g. anxiety
  - to shift towards more pro-activity (to play with the environment)
  - chaining of (cause,effect) pairs – *planning*