



# WP7WG3: What is in an affective architecture for situated agents?

Paolo Petta (OFAI)

Ruth Aylett (HW)

## WP7WG3: Goal

### Deliverable D7c:

- “To explore the roles of emotions in connecting lower-level cognition and action to aspects of higher levels.”
- “...a critical analysis of main integration challenges and the potential contributions of emotions.”

# WP7WG3: Threads of activities

## Thread A:

### Conceptual clarifications

- Lower-level cognition and action:  
in particular, in virtual environments
- Higher-level cognition and action:  
focus on what is accepted from and  
propagated downwards to the lower level

# WP7WG3: Threads of activities

## Thread B:

### Operationalisation of theory

- Which scenarios require and warrant the providing of which emotional functionalities
- Uncover tacit assumptions of theory: theoretical constructs and environmental properties
- Deviations from natural theory due to engineering constraints

# WP7WG3: Threads of activities

## Thread C: Evaluation of systems & designs, scenario definitions

- Critical assessment of current achievements and issues
- Scenario descriptions relating environmental aspects to features of higher-level and lower-level cognition and action.
- Contributions to the “assumptions” and “guidelines” reports of the exemplar

WP7WG3: What is in an affective architecture for situated agents?

# WP7WG3: Threads of activities

## Thread D: Integration challenges

- Formulation of criteria of practical relevance for the selection of theories, designs, implementation methods, and scenarios

## Thread E: Guidelines for future research

- Directions for principled approaches

# Situatedness

- Predicament of being in the world
- Ongoing activity (cf. the beheaded chicken)
- Coupling to the external world
  - E.g.: anticipation, waiting (expectancies)
- Freedom in subjective conceptualising
- Social lifeworlds
- Necessity to **enact**, **defend** and **revise** subjective conceptualisations of current and future states of the social lifeworld

# Purposes of Emotions

- (Prospective) maintenance of the subjective lifeworld in the light of available evidence
  - Concern satisfaction (including social concerns)
  - Future-oriented subjective significance of events
  - Maintenance of processing context: situation-driven action tendencies
  - Sense of reality (grounding)
  - Degree of openness to world
  - ...
- Embedded in regulatory contexts (dual control)

# Complete agents

- Ecologically balanced
- Avoid gratuitous reifications
- Multiple continuous processes
  - Allows synchronised recruitment of resources

## Lower vs. higher level

- Scientifically, only a helpful simplification
- Resolution of operations (time, space)
- Temporal regularities of representations
- Degrees of dependence from outside world
  - Abstractions, generalisations
  - E.g. higher-level reification of duration

## Lower level

- Direct coupling to an agent's environment
- Time scale corresponds directly to temporal resolution of world interface
- No (symbolic) representations (needed)

## Higher level

- Reasoning (over abstract spaces)
- Planning, counterfactual reasoning (abstracted from current world state)
- Relevant properties of representations
  - Rate of change
  - Referent; internal or environmental conditions

# Interrelationship between lower and higher levels

- What meaningful information can be transferred?
- Management of ongoing activities of the agent at different levels
- Interaction coordination
  - Subsumes decision making and action selection

# What High Level systems give

- Sequencing
  - Especially where actions are inter-related
  - Painting yourself into corners
- Complex goal achievement
  - Where inter-related goals must be managed
- Generic models
- Efficient use of resources

WP7WG3: What is in an affective architecture for situated agents?

# What Low Level systems give

- Responsiveness
  - To the environment
  - To other agents
- Ability to use raw sensor data
  - Sensors do not deliver symbols
- Adaptiveness
  - No explicit expectations

WP7WG3: What is in an affective architecture for situated agents?

# Three-level hierarchies

- De facto standard in robotics
- Strategic (task)
  - Long time scales, planning
- Tactical (action)
  - Reactive systems, abstract execution (kinematics rather than dynamics)
- Executive (command)
  - Real-time, dynamics, direct control of hardware

# Relationships between levels

- Top-down control
  - One-to-many mapping down an action hierarchy
- Bottom-up control
  - HL called only when LL gets stuck
  - May still have a one-to-many mapping
- Supervisory control
  - HL sets context/creates constraints for LL

# Bi-directional interaction

- From HL to LL
  - Originally thought of as execution
- From LL-HL
  - Originally thought of as:
    - Sensor processing
    - Error handling when actions fail from HL perspective

# Top-down control

- Action expansion
  - Planner output of ‘primitive actions’
  - Passed to execution system
    - Maps onto control-system commands
    - In robotics, these are typically movements
- Sensing only at HL
  - Via modelling process
  - May include some LL self-correction
    - Local dynamics only
    - Safety issues
- Model hierarchies

# What is it good for?

- Largely abandoned in robotics
  - Too slow, too rigid, too fragile
- Still used in ECAs
  - Which are not in the physical world
  - Output of the HL system is markup
  - Sensing typically from language processing of user inputs

# Role of emotion

- Emotion from HL appraisal
  - Only affects HL
  - Propagated downwards to expressive behaviour
  - Which is in any case usually ‘chunked’ - animation/motion capture
    - Different from robotics
- Responsiveness issues
  - Difficult to deal with body-language interaction
    - Problems with turn-taking
  - Or dynamic environments

# Bottom-up control

- Planning-as-resource
  - When painted-into-corner
- Which you have to notice
  - Emotion may be part of the mechanism for doing this
- Many robots run without any HL system at all
  - Primitiveness of sensing does not support goal or action complexity
  - Once language use is included this becomes untenable

# Social robots

- LL can handle movement interaction
- Could invoke language system as needed
  - Propagate emotional state upwards
- But language content cannot be determined at low-level
  - Either for self or dialogue partner
- Hard to avoid traded control
  - Then how to link the two emotional levels?
- Persistence produces a goal management problem

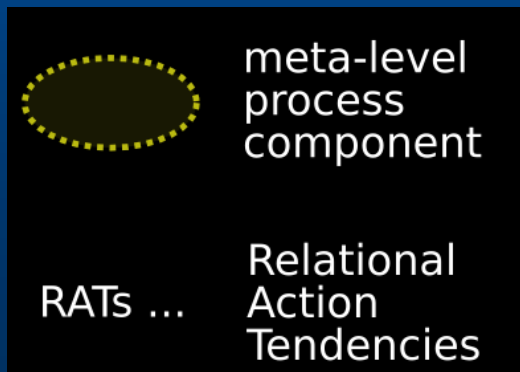
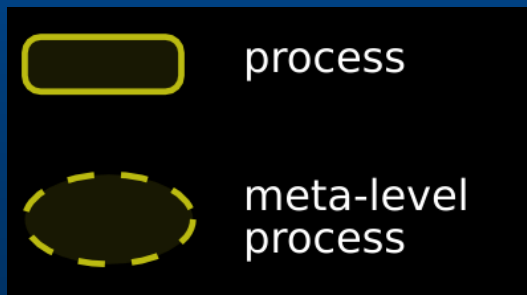
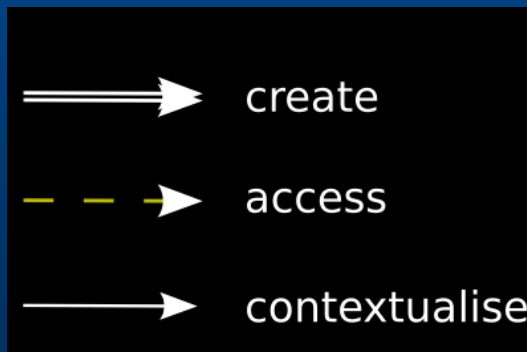
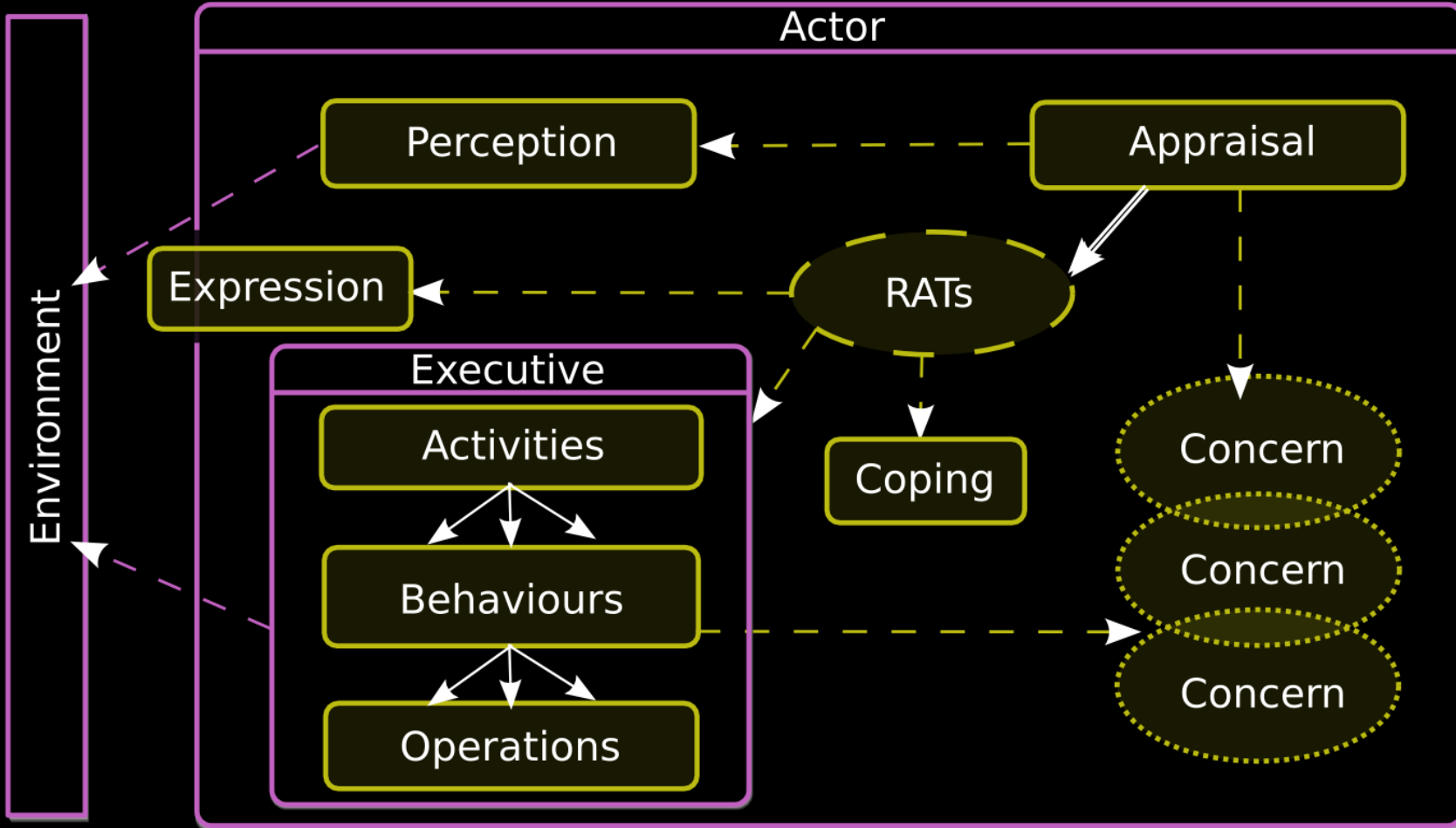
# Supervisory control

- Assume a repertoire of competences
  - Sets of behaviours whose goal-achievement possibilities modelled at HL
- HL maps primitive action onto set of competences
  - Enables these and disables others
- LL runs competences under sensor control
  - Example: bird nesting v bird seeking food
- Separates second-to-second variability from mode shifts

# A constraint model

- More general view of HL-LL relationship
  - Mutual constraints
- LL emotional systems constrain HL goal management
  - Project drives and motives upwards
  - LL emotion as a component of continuing appraisal process
  - HL may elaborate LL emotions
  - Or overrule them:  
Involuntary and controlled expressive behaviour

WP7WG3: What is in an affective architecture for situated agents?



# Terminology

- Executive
- Action & Interaction (Expectations)
- Activity (maintained high-level context)
- Task (specification of what to do)
- Goal (state to be achieved, avoided, maintained)
- Behaviour (unsupervised compound interaction)
- Operation (unit of interaction)

# Terminology

- Concern  
(Autotelic principle, no direct activity control)
- Motivation (Active part of concern)
- Emotion eliciting stimulus (against current context)
- Relational Action tendency  
(situation-driven active conceptualisation)
- ...

# Scenario template (for interaction coordination)

- Purpose of whole system
- Properties of environment
- User Interface
- Agents characteristics
  - Kinds, numbers, concerns, tasks
- Agent-world interaction
- Agent-agent interactions

## Evaluation criteria

- Scope of the emotion model
  - Coherence and completeness
- Degree of principled integration of emotion model in architecture
  - E.g., amount of “patching”

# Issues/challenges

- Nature of eliciting stimuli?
  - Sensors + early processing
  - Early/automatic stimulus analysis: valence, novelty,...
- Dynamic properties, calibration
  - Speeds, synchronisation points, intensities,...
- Integration of multiple timescales
- Flexible couplings of levels
- Tension between design-time configuration and control vs. run-time adaptivity and evolution

# Requirements analysis

- Systems differ in conceptual assumptions
  - Not always complementary/compatible
  - Requirements for application scenarios
    - What warrants the use of different levels in control mechanisms?
    - What warrants emotional processes as part of the architectures?
- Importance of overall system properties, “purpose”