WP10: PlayMate
WP1: Architectures & Representations

Jeremy Wyatt

@ Birmingham: Nick Hawes, Aaron Sloman, Michael Zillich, Marek Kopicki, Somboon Hongeng, Mohan Sridharan (from July 1st)

Summary

- WP10 PlayMate Scenario + Joint Scenario

- WP1 Architectures & Representations:
  - How to represent actions
  - Experimental profiling of architectures
  - Deliberative control of information processing
  - Representation of object shape to support learning of affordances
  - PlayMate system as an implementation of Global Workspace theory

Collaborative Manipulation

- Target Month 36:
  - Grouping/Arranging items by colour or size (coloured packets, blocks)
  - Human interventions (helpful and non-helpful)
  - Learning and recognition of action sequences
  - Early integration of manipulator with Explorer (accept offered objects, place held objects on a table)

- Target Month 48:
  - Laying the table together
  - Take a cup, bowl, jug, spoon
  - Human shows the robot how the pieces should go together. Place the spoon to the right of the bowl. Bowl on the table, cup on the table behind bowl. Pour contents of jug into cup.

PlayMate Scenario

- Key scientific challenges we could tackle/ are tackling:
  - Action Representation: Representing and recognising complex action sequences (several grasps and deposits) (WP1,WP7)
  - Object modelling: How should we represent how object shape determines action outcomes (WP7)
  - Interruption: How can we deal with unexpected events e.g. human intervention, execution failure (execution monitoring, continual planning) (WP1,WP4)
  - How can we plan information gathering (where to look, what information to extract) (WP1,WP4)
  - Recognition of failure or inability to perform an action (can't pick up spoon) (WP 1)
  - How can we represent that objects and parts have associated conventional actions (jug will be used to pour, handle can be grasped) (WP1,WP7)

PlayMate Integrated System: Where we are now

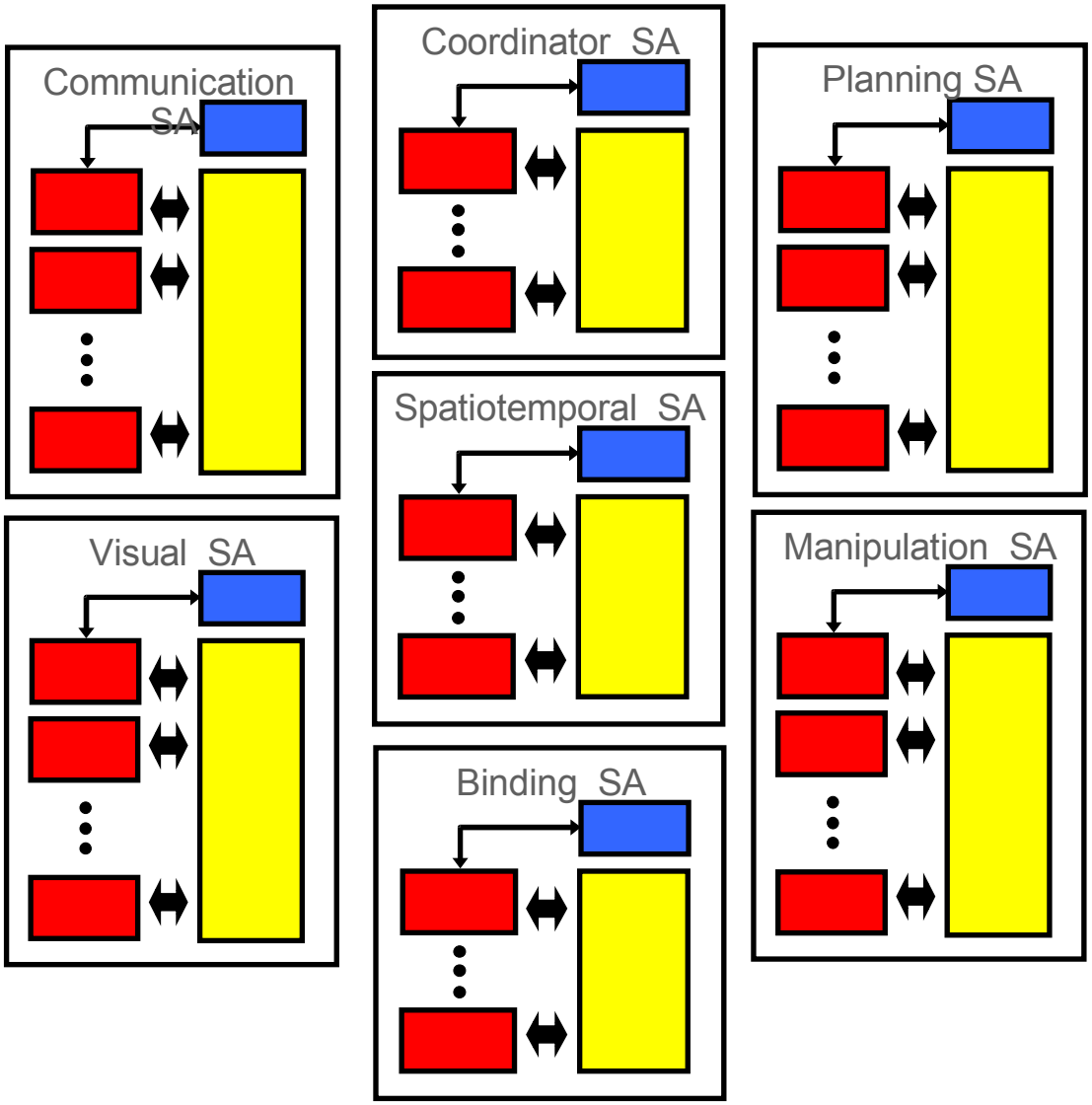
- More reliable grasping of simple objects
- Converted to tracker-based framework
- Action analysis integrated into spatio-temporal working memory
- New version of motion planning software + visual servoing for release to consortium: June 27
- First public release of CAST/BALT: June 27
- We are about 4 weeks from PlayMate system for 3rd year review demonstration

Integrating PlayMate and Explorer

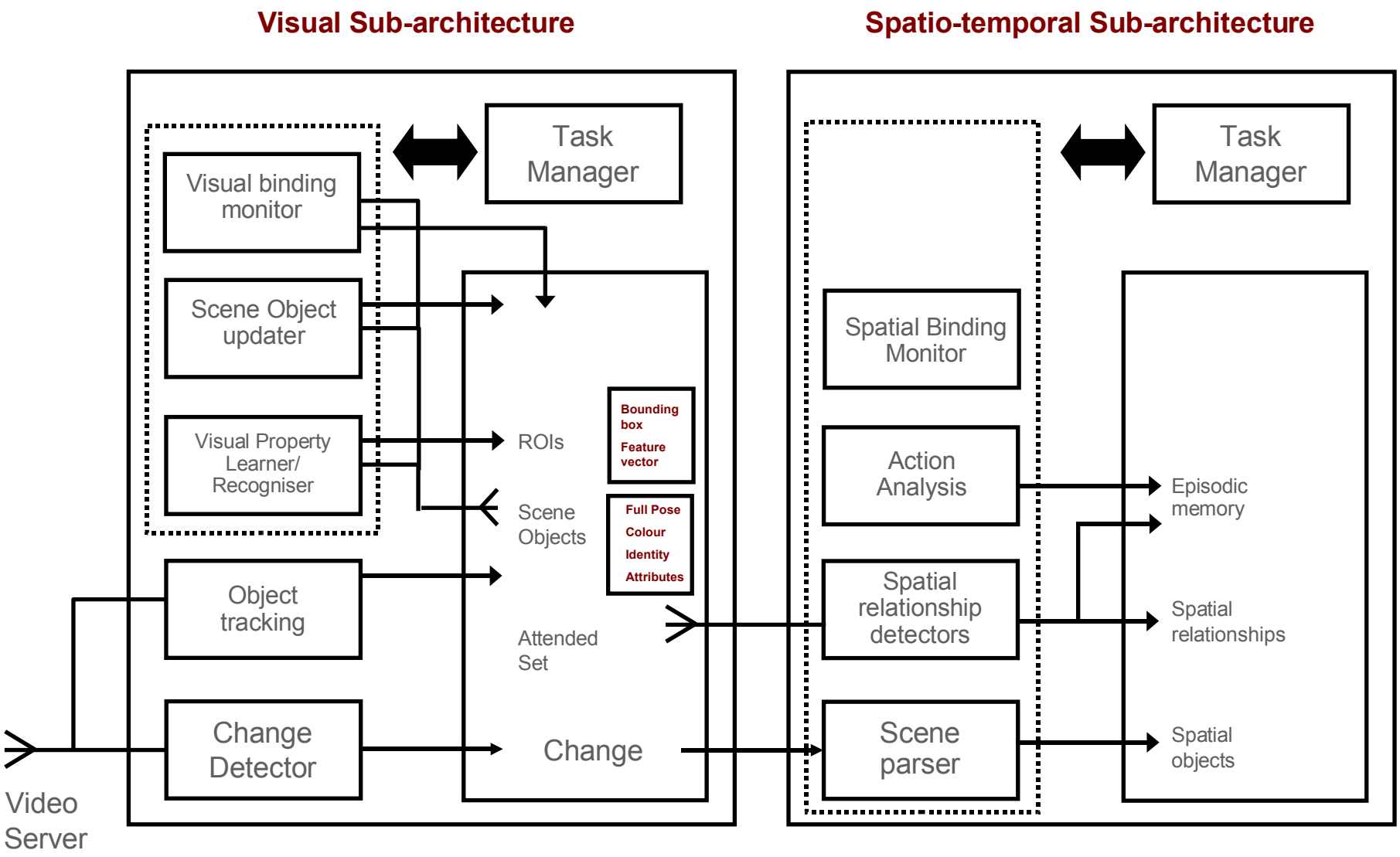
- We want initial integration for month 36, full integration by month 48
- For month 36:
  - Scenarios in which human hands object to robot, robot transports and places object on tables of known height
- For month 48:
  - Robot is able to retrieve a limited range of objects from a table of known height
- Scientific Qs: allows us to explore issues around integrate spatial and action representations of very different types

Current architectural instantiation

- Main Q: How to represent and manage information about action
- Representation of continuous change in visual WM
- Episodic representation of action stored in the spatio-temporal memory
- Planning makes reference to episodic memory through binding SA
- VM is used to raise alarms

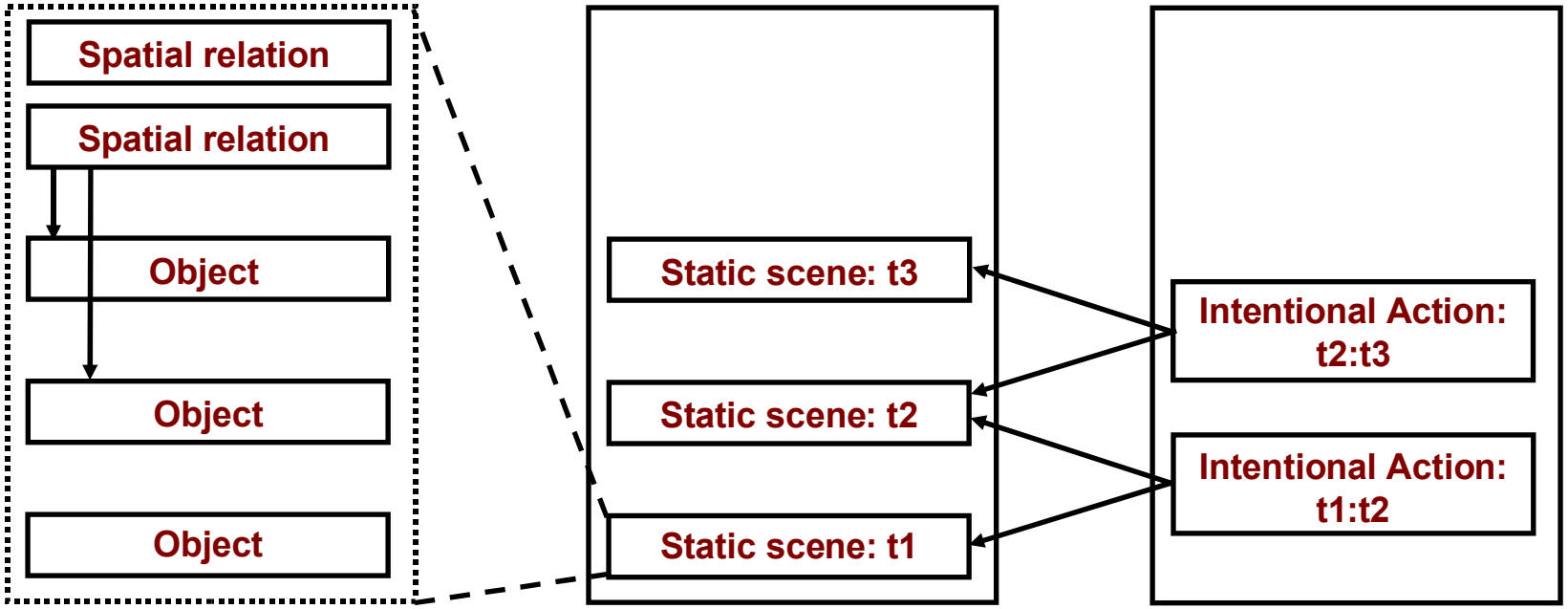


Architectures: how to represent action?



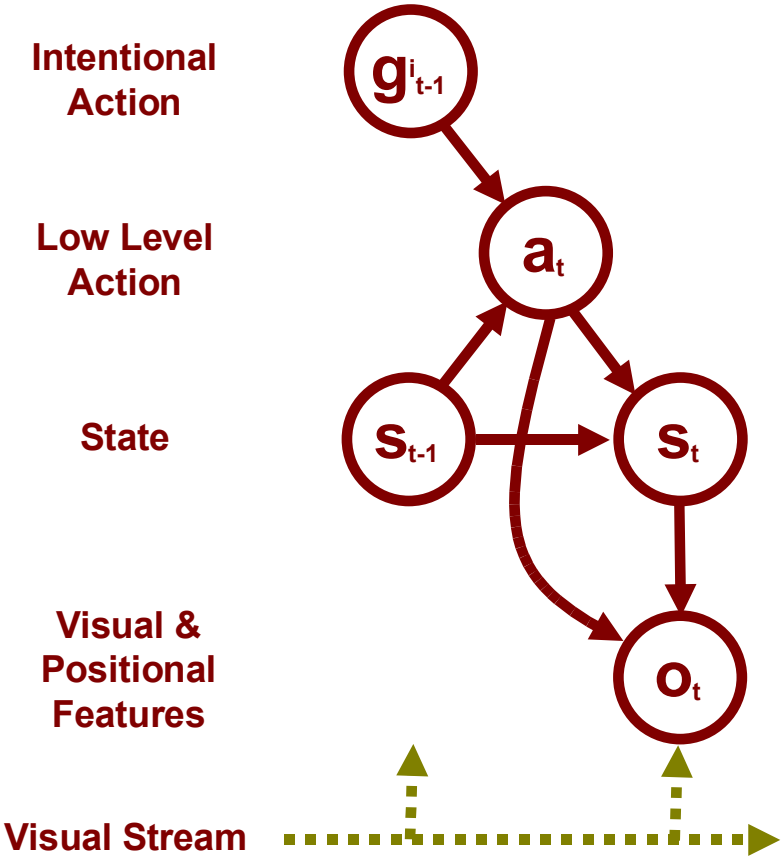
Spatial temporal (episodic) memory

- Episodic representation of action is stored in the spatio-temporal memory
- Each static snapshot is a set of objects and spatial relations
- Action analyser creates high level intentional action labels for the activity between static scenes
- These link the static scenes



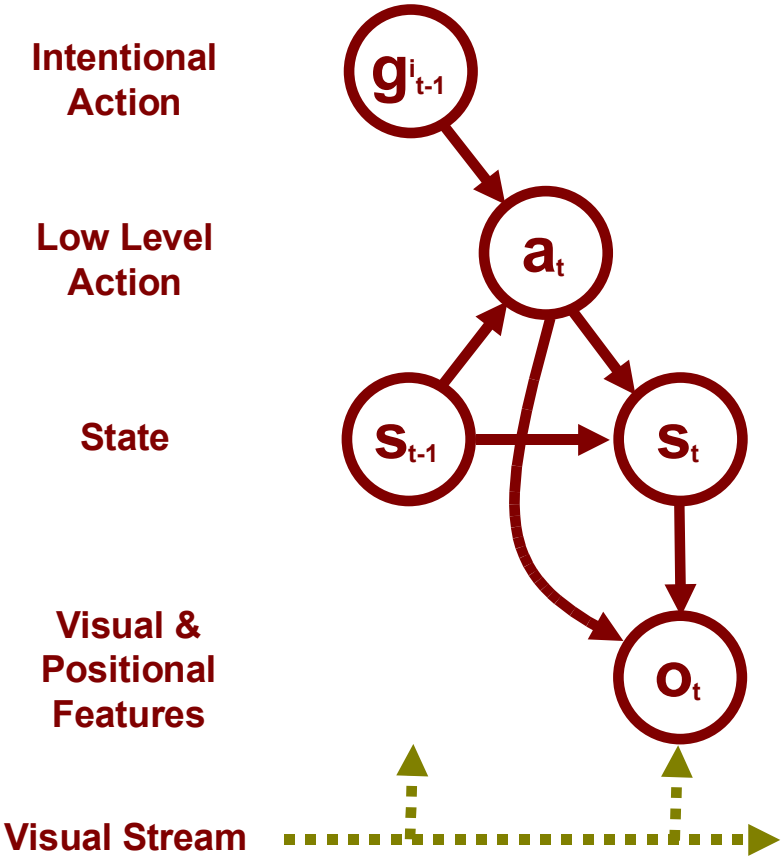
A hierarchical graphical model of action

- High level actions:
  - push, pull, reach, retract
- Low level manoeuvring actions
  - immediate hand velocity relative to objects
- Visual features
  - quantitative spatial relationships between objects, agent and objects
- Visual Stream
  - image stream



Problems

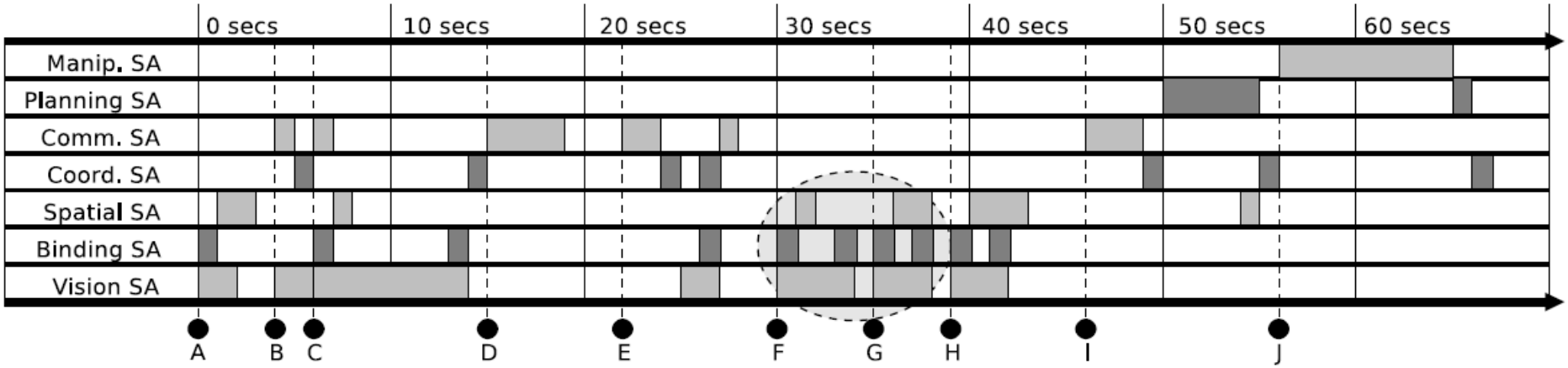
- Learning what intentional actions are composed of (in terms of sequences of low level actions and states)
- Learning what low level actions and states are composed of in terms of quantitative visual and positional features
- Recognising intentional actions from video streams

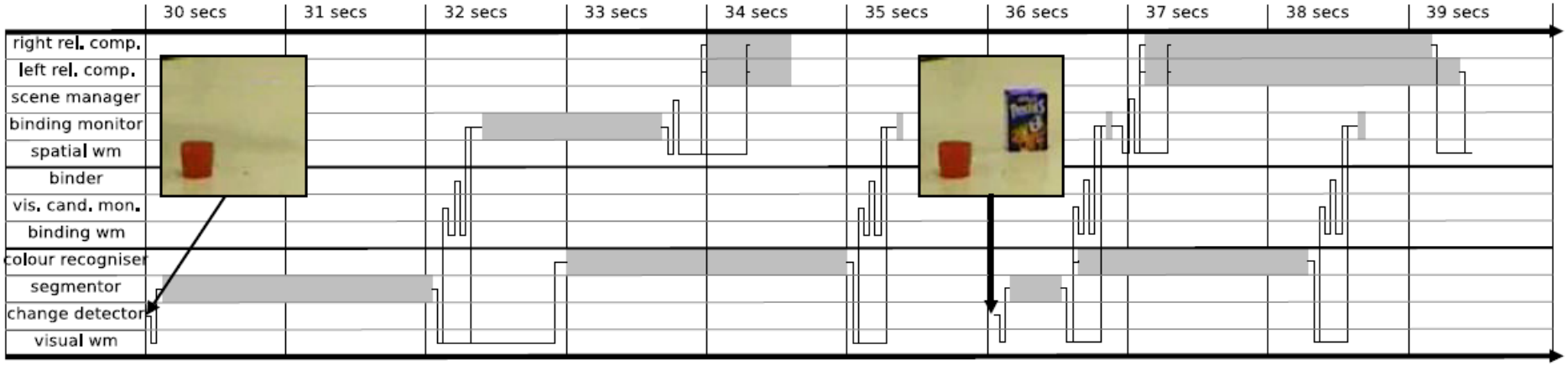


Recognition using learned action models

Current approach (AAAI 07)

A Red object placed on table.
B Tutor (T): "This is a red thing."
C Red object replaced with blue object.
D Robot (R): "Is that red?"
E T: "No, this is a blue thing."
F Blue object replaced with red object.
G Blue object placed to right of red object.
H Blue object placed to left of red object.
I T: "Put the blue things to the left of the red thing."
J R moves right hand blue object to left of red object.





Work in progress

- We want to characterise system behaviour as we move through the space of architectures for systems that satisfy the same design niche

- Easy architectural changes in CAS/CAST:
  - Variations in control: e.g. parallel versus sequential control in task managers.
    - Fully parallel; Sub-architecture parallel/component sequential;fully sequential.

  - Variations in connectivity (assume n components)
    - N components, 1 sub-architecture
    - N components, M sub-architectures (N>M, M>1)
    - N components, N sub-architectures

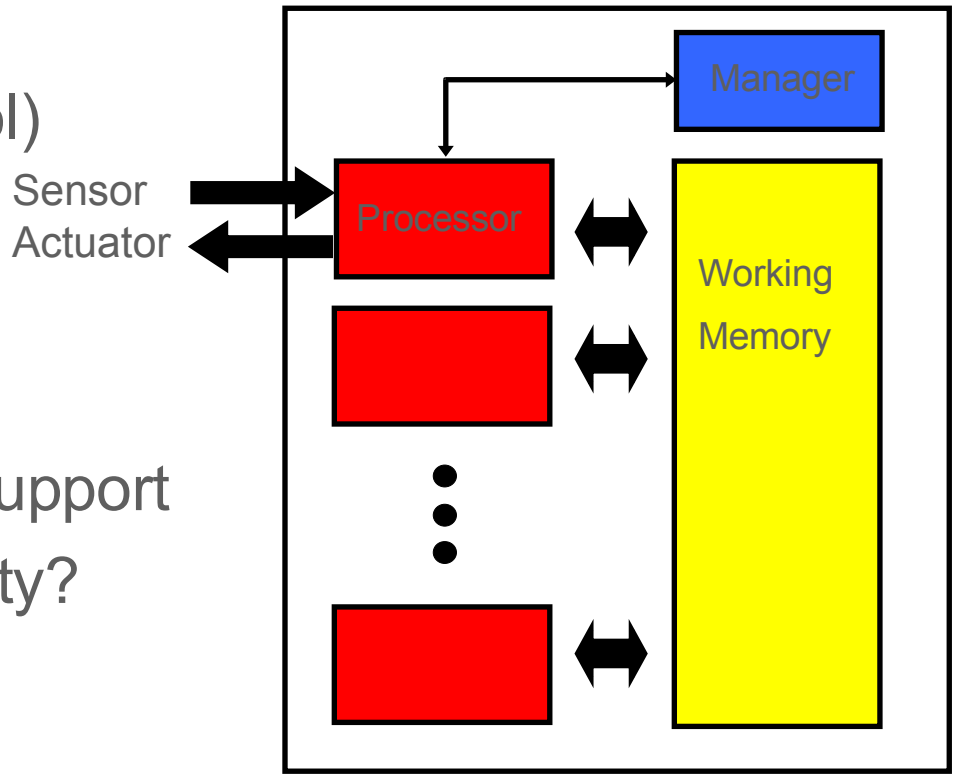
Measures of architectural behaviour

- Proportion of system level task steps completed
- Time to complete system level task
- Number of component level tasks that complete within a given time period
- Average time taken to read/write from working memory
- Average time taken for a component to complete
- Ratio of utilised to non-utilised change notifications

A Red object placed on table.
B Tutor (T): "This is a red thing."
C Red object replaced with blue object.
D Robot (R): "Is that red?"
E T: "No, this is a blue thing."
F Blue object replaced with red object.
G Blue object placed to right of red object.
H Blue object placed to left of red object.
I T: "Put the blue things to the left of the red thing."
J R moves right hand blue object to left of red object.

Styles of task management

- Permissive
- Reactive (Finite State Control)
- Deliberative (Planned)
- What kind of planner could support planning of component activity?

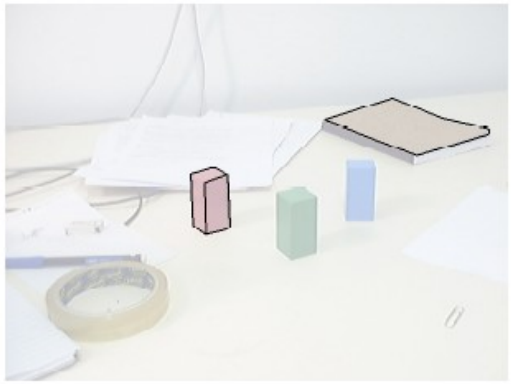


Planned approach

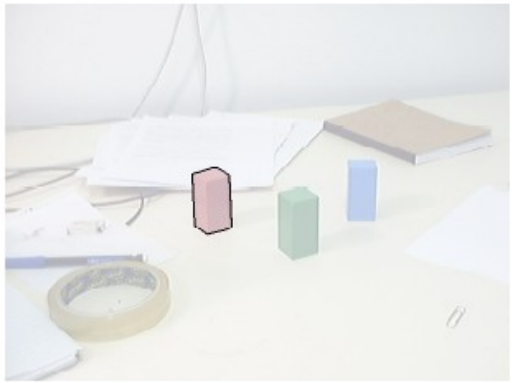
- Focus on visual processing components
- High level planner sets goals for Visual SA (e.g. find a particular object in the scene, find grasp points on an object)
- Learn models for visual actions that transition between distributions over states (e.g. probability of classifications for ROIs)
- Pose problem of planning with these uncertain actions as decision theoretic planning problem through an information state MDP
- Re-plan as information arrives

Representations: modelling object shape for affordances

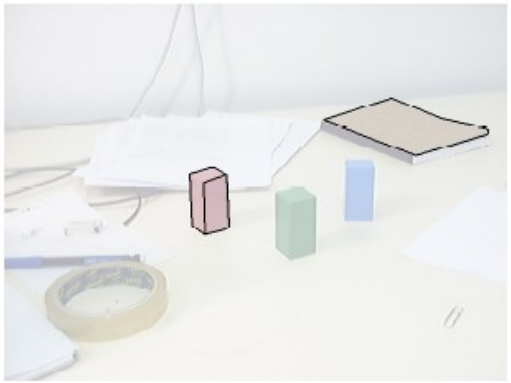
- Goal: predict what objects do under action
- How to represent shape sufficient for this prediction task?
- Approach:
  - Use edge information to recover some surface shape
  - Build on our current work on parameter free perceptual grouping
  - Current results:
    - Convex contour completion
    - Parameter free
    - Incremental & anytime
    - Modulated by saliency



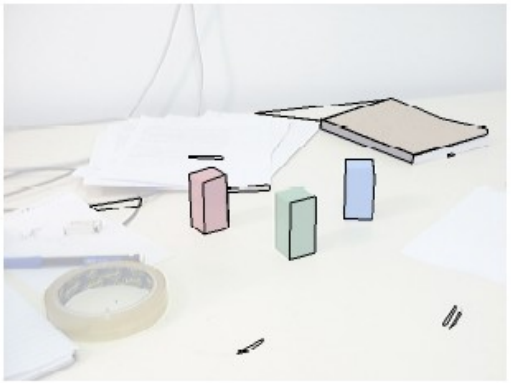
Perceptual Grouping



(a) 15 ms

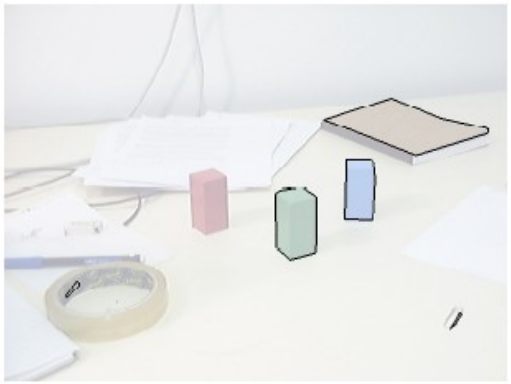


(b) 26 ms

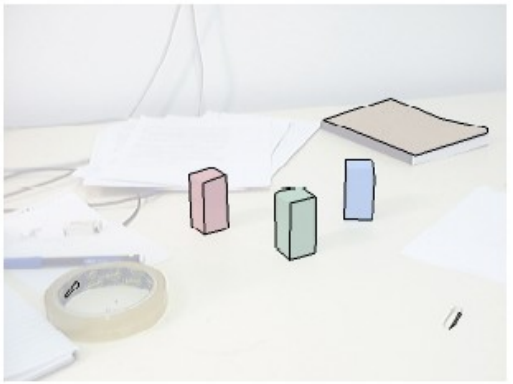


(c) 70 ms

(d) 19 ms



(e) 37 ms



(f) 45 ms

WP1: Architectures & Representations

- Recently published results:
  - Joint work with DFKI,ALU on mediating between representations now published (joint with WP4) (IJCAI07)
  - Main architectural ideas published in AAAI-07, with DFKI,ALU,UOL
  - Paper on CAST/BALT published in IEEE RO-MAN 07
  - AAAI 07 Workshop Paper on Evaluation methods for architectures
  - Hierarchical representation of action published in Humanoids Dec 06 + invited book chapter (to appear 2007) from Dagstuhl seminar.
  - Parameter free perceptual grouping (OAGM 07: best paper award)