Reasoning about Perception, Space and Motion: a Cognitive Robotics Perspective

David A Randell
Intelligent Systems & Networks
Dept E&EE
Imperial College London
London, UK

www.ee.ic.ac.uk/research/neural/iis/research.htm
1: Overview

• Brief overview of Cognitive Robotics
  – emphasis: modelling and reasoning about space and visual perception, less so about motion

• Current Cog Rob research at Imperial
  – viewpoint based QSR logics, active perception
    • Other PHD research work: tracking objects, sensor fusion, modelling shape,…

• Applications & Implementation
  – Challenges implementing logical formalisms on real-world robots

• Retrospective thoughts
  – What does this tell us, if anything, about applied AI research?
2: Background

• What is Cognitive Robotics?
  – endowing robotic/software agents with high-level cognitive functions
    • theory/implementation of robots/agents that reason, act and perceive in a changing, incompletely known, unpredictable environment
    • Reasoning about goals, actions, perception, what to expect and look for, cognitive states of robot itself and others,…
    • Integrating reasoning, action, perception within uniform theoretical framework
3: Background - contd

• Simple approach
  – high-level control as automated planning?
    • Initial state and goal; find set of actions satisfying
      goal, then execute
  – leaves several unanswered questions, e.g.
    • sensing during task execution?
    • Lack of reactivity, e.g.
      – need high-priority interrupts, respond to failures of
        execution modules, unexpected situations, ...
    • Incompatibility with conventional robotics
      – dealing with micro actions, noise and uncertainty
  – development of different architectures ...
4: Background - contd

• Classic architectures and logics
  – Mainstream robotics: Sensing, path-planning, low-level control
  – Cog Rob: different emphasis...
    • Common architectures: **Deliberative**: Simple task-level planner and executor, functional decomposition of problem vs. **Reactive**: behaviour-based decomposition of problem. Hybrid architectures
    • Use of high-level programming
      – E.g. task-level controller guided by model of domain, hierarchical planning, layered abduction model,...
  – Application of specialised KR&R languages/logics
    • Situation/Event Calculus for change and dynamic updates
5: Background - contd

- Some key players:
  - Levesque, Reiter, ... (Uni Toronto)
    - high-level robotic control, Golog
  - Kupiers, Remolina (Uni Texas)
    - Spatial Semantic Hierarchy using an explicit ontological hierarchy: sensorimotor, control, causal, topological and metric levels e.g. topological level assuming places and regions, metrical level: distance, direction, ...
  - Shapiro et al (Buffalo, NY)
    - e.g. “Embodied Cassie” – NASA sponsored, robot assistant, hardware and software simulated agents
6: Background - contd

• **E&EE robotics group**
  - **Cog Rob**: Murray Shanahan, Mark Witkowski, myself,...
    - three consecutive Cog Rob EPSRC projects: two completed (using mobile robots), one current (humanoid robot),...
    - General approach: interleaving Event Calculus, Abduction, QSR, ...
  - **Biologically inspired robotics**: Yiannis Demiris,..
    - Royal Society & EPSRC funded projects
    - Approach: Robot imitation, perception, learning,...

• **DoC (Bioinformatics lab)**
  - Stephen Muggleton et al: Closed Loop Machine Learning; ILP and ASE-Progol; EPSRC project
7: Background - contd

- Some ISN robots ...

Left to right: Khepera, LinuxBot, walking robot, PeopleBot, Nomad and ‘Ludwig’ complete with toys
8: Background - contd

• Ludwig
  • An upper-torso humanoid robot; ~80% adult human size, each arm with three degrees-of-freedom, stereo-vision sensor

• Video clips of Ludwig in action
  • http://casbah.ee.ic.ac.uk/~mpsha/ludwig/Video.html
9: Motivation

• Early work (proof of concept)
  – Khepera robots
    • Limited robot sensors, heavily engineered robot environment

• Later work to overcome sensor restrictions and lack of interaction with environment
  – LinuxBots
    • more suited to true office environment, stereo-vision
  – Ludwig (humanoid)
    • Visual perception, interaction with objects,...
10: Motivation - contd

- Abduction [Shanahan et al IJCAI’89, ECAI’96, AAAI’96, AAAI’97, ETAI’98, ..., ATAL, 00]
  - Background theory $\Sigma$, observation $\Gamma$, infer hypotheses $\Delta$, so that

$$\Sigma \land \Delta \models \Gamma$$

where: $\Delta$ is consistent and comprised of abducibles

- used for sensor data assimilation
  - E.g. abducting existence of objects, from their images
  - Then use deduction for predictions...
• Identifying useful spatial ontologies and structures
  • Esp. information related to depth (from stereo camera)
• Exploit psychophysical depth cues
  • \textit{spatial occlusion and motion parallax}
    – Related to: object permanence, reversible actions for object recovery,…
    – topological invariants
    – Viewpoint-centred description of objects in space
• QSR ontologies, structures and techniques
  – region-based ontologies, JEPD relations, \textit{conceptual neighbourhood diagrams} (CNDs), …
• Viewpoint centred logics of space
12: Occlusion Calculi

- ROC20/6 [Randell et al, IJCAI’01; Randell and Witkowski, KR’02]
  - first-order theory; sorts: points, regions, bodies
  - 20 JEPD relations, embeds RCC8/5 calculi
  - primitives: TotallyOccludes(x,y,v), image(x,v), ...
  - Handles non-convex bodies/regions, c.f. LOS14

\[ \text{e.g. } \text{TotallyOccludes}(a,b,v) \]
13: Occlusion Calculi - contd

• Theory extended
  – additional primitives: \( N(v,x,y), \text{Left}(x,y,v) \)
  – also: integrating multiple viewpoints, e.g. for collaborative task execution, interpreting visual scenes,...

• Composition (look-up) tables computed
  – Verified using bespoke shell program interface to SPASS [Randell & Witkowski, KR’02]
    • Encodes theorems of the form:
      \[ R_1(x,y) \& R_2(x,y) \rightarrow R_3(x,z) \lor \ldots \lor R_n(x,z) \]
      – Where \( R_1, \ldots, R_n \) are a set of JEPD relations
    • Used for consistency check of sets of ground relations
### 14: Occlusion Calculi - contd

<table>
<thead>
<tr>
<th></th>
<th>NO</th>
<th>PO</th>
<th>POI</th>
<th>TO</th>
<th>MO</th>
<th>TOI</th>
</tr>
</thead>
<tbody>
<tr>
<td>NO</td>
<td>NO PO</td>
<td>PO</td>
<td>POI</td>
<td>TO</td>
<td>MO</td>
<td>TOI</td>
</tr>
<tr>
<td>PO</td>
<td>NO PO</td>
<td>PO</td>
<td>POI</td>
<td>TO</td>
<td>MO</td>
<td>TOI</td>
</tr>
<tr>
<td>POI</td>
<td>NO PO</td>
<td>PO</td>
<td>POI</td>
<td>TO</td>
<td>MO</td>
<td>TOI</td>
</tr>
<tr>
<td>TO</td>
<td>NO PO</td>
<td>PO</td>
<td>POI</td>
<td>TO</td>
<td>MO</td>
<td>TOI</td>
</tr>
<tr>
<td>MO</td>
<td>NO PO</td>
<td>PO</td>
<td>POI</td>
<td>TO</td>
<td>MO</td>
<td>TOI</td>
</tr>
<tr>
<td>TOI</td>
<td>NO PO</td>
<td>PO</td>
<td>POI</td>
<td>TO</td>
<td>MO</td>
<td>TOI</td>
</tr>
</tbody>
</table>

TO: TotallyOccludes
PO: PartiallyOccludes
NO: NonOccludes

E.g. the set:

\{	ext{TO}(a,b,v1), 
\text{PO}(b,c,v1), 
\text{NO}(a,c,v1)\}

is inconsistent.

Composition table for ROC6
15: Occlusion Calculi - contd

**Conceptual Neighbourhood Diagram (CND)**

- encodes direct continuous transitions
- interpreted as state-state envisionment, reworked in Event Calculus
- CND Structure as basis for prediction, expectation, abduced explanations

```
Conceptual Neighbourhood Diagram (CND)

CND for ROC6/20
```

- **NonOccludes**
  - **PartiallyOccludes**
    - **PartiallyOccludes\(^1\)**
  - **TotallyOccludes**
    - **MutuallyOccludes**
    - **TotallyOccludes\(^1\)**
16: Occlusion Calculi - contd

From this ...

ROC6/20 CND relates to occlusion event transitions

to this ...

NonOccludes \( \cup \) PartiallyOccludes \( \cup \) PartiallyOccludes\(^{-1}\) 

TotallyOccludes \( \cup \) MutuallyOccludes \( \cup \) TotallyOccludes\(^{-1}\) 

NonOccludesDC\( (a,b,v1) \) \& Left\( (a,b,v1) \) 

NonOccludesEC\( (a,b,v2) \) \& Left\( (a,b,v2) \) 

PartiallyOccludesPO\(^{-1}\)\( (a,b,v3) \) \& Left\( (a,b,v3) \) 

PartiallyOccludesTPP\(^{-1}\)\( (a,b,v4) \) \& Left\( (a,b,v4) \)
17: Applications

- But how does this work out in practice?
  - Early work largely directed toward theory development, ‘paper-based’ proof of concept,…

- Re-assessment of early assumptions made
  - e.g. Region-based ontologies fine in theory, but another level required to reliably extract them from sequences of digital images

- Re-examine ontologies, primitives and logical formalisms, and inference strategies used …
18: Applications - contd

- Hypothesising Object Relations from Image Transitions [Santos and Shanahan, ECAI’02]
  - Platform: LinuxBot: proprioceptors, stereo vision
  - ‘Jackson Pollock’ textured paper cylinders

LinuxBot among cylinders and typical extracted stereo image
19: Applications - contd

• Reasoning about ‘blobs’
  – Displacement, size and depth information from stereo-disparities
  – Uses restricted set of RCC-like relations: DC(x,y), EC(x,y), CO(x,y)
  – Axiomatised relations encoding transitions: approaching(x,y), coalescing(x,y),...

• Augmented by depth-profiles extracted from horizontal cut in images
  – uses Depth Profile Calculus (DPC)

• JEPD sets of relations, conceptual neighbourhoods, ...
20: Applications - contd

- Abducing motion of robot from sequences of images

Interpreting sequences of images: approaching and coalescing ‘blobs’
21: Applications - contd

- Active Vision [Shanahan, KR’02; Shanahan and Randell, KR’04]
  - using high level knowledge and reasoning to inform low-level perception
  - Platform: stereo-camera used for 2D image capture

Simple edge extraction, how can we recover or account for missing or incomplete edges?
• Use abduced top-level object descriptions to guide low-level processes
  - E.g. abducting edges from raw visual sensor data (Abduction/Hypothesis formation), then predicting existence of other edges from domain model (deduction)

• Prediction sets up the expectation
  - Use feedback
  - Adjust gain of camera to confirm/disconfirm existence of lines
23: Applications - contd

- Hypotheses:
  - (H1) A and D true corner
  - (H2) Edge AD surface feature

- Different expectations:
  - (E1) Vertical edge from A
  - (E2) Edge DC extends beyond D

- Selectively adjust sensitivity of edge detection

Hypothesised (light) and given (dark) line segments of an abduced cuboid
24: Applications - contd

- Augmented abduction using feedback and expectation
  - Abduced hypotheses weighted according to their explanation value
  - Numerical computation overlaid on abduction assigns explanatory value to competing hypotheses
  - Fulfilment of expectations adds to the *explanatory* value of a hypotheses

- Cast into hypothetico-deductive model
  - Process sensor data, abduce explanations, deduce expectations, confirm/disconfirm these, feedback results into loop,..
• Region-based spatial logics for practical image interpretation?
  
  – Assumptions examined:
    • Identity and extent of bodies assumed *a priori*
      – Opaque bodies, many relations not directly observable
    • Early QSR theories based on continuous models
      – But digital images map better to a discrete space
      – Ignores physical basis of imaging systems e.g. resolution and image scale
    • Reliable detection of region boundary information?
      – poor segmentation, incomplete boundary information,.. 
  
  – Other assumptions
    • fast sampling rates, sensor noise, variable lighting conditions, ...
26: Applications - contd

- Use discrete variants QSR theories
  - E.g. [Galton, COSIT ’99] discrete mereotopology
    - Ontology: Pixel elements and regions
    - Defines RCC8D set of spatial relations

- Augment theory
  - Monadic colour predicates defined on regions
    - exploit hue colour-space information

- specification language
  - logical properties of region, region-relation detectors

- Region/region-relation detectors
  - Visible*(x,v), DC*(x,y), EC*(x,y), NTPP*(x,y),...
27: Applications - contd

- Map LOS14 relations on bodies to RCC8D relations on their images and to the implemented detectors e.g.
  \[ \text{Clears}(x,y,v) \rightarrow DC^*(i(x,v),i(y,v)) \]
- Rework LOS14’s relation set and associated CND using the *Event Calculus*
  - Abduce object relations from 2D spatial relations extracted from images
- use CND to interpolate spatial relations not directly detected via the sensors
28: Applications - contd

- **Example path:** \( C \rightarrow JC \rightarrow PH \rightarrow JF \rightarrow F \)

- \( C \) maps to \( DC^* \); \( F \) maps to \( NTPP^* \) but \( JC, PH \) and \( JF \) are ambiguous wrt \( EC^* \). How do we recover the sequence of relations?

From \( DC^* \) at \( t_0 \) abduce \( C \) and \( JC \) at \( t_1 \) from \( EC^* \). At \( t_4 \) we abduce \( F \) from \( NTPP^* \). Only \( EH \) and \( JF \) neighbour \( F \), but \( EH \) ruled out as both regions continuously detected; so interpolate \( JF \) at \( t_3 \). Given \( JF \) at \( t_3 \), interpolate \( PH \) at \( t_2 \). This recovers the sequence: \( C \rightarrow JC \rightarrow PH \rightarrow JF \rightarrow F \).
29: Applications - contd

- Tracking regions using spatial relational information
  - Which regions in a sequence of images are images of the same object? Requires theory of region-identity

- One approach
  - use extra-logical properties for object recognition and tracking, e.g. colour, texture
  - Can we use relational information?

- Use Region-based QSR theories
  - e.g. RCC8, spatial variants of Allen’s interval logic
  - Exploit structure of CNDs for measuring similarity between relations
30: Applications - contd

Similarity matrix for CND$_{RCC8}$

\[ d_{\text{min}} \Phi(x,y) : \text{maps pair of individual ground relations } x,y \text{ to min CND node-node path distance for theory } \Phi, \text{ e.g.} \]
\[ d_{\text{min}} \Phi(\text{EC}(a,b),\text{DC}(a',b'))=1 \]
\[ \text{where } a \rightarrow a', b \rightarrow b' \]

•CNDs used for querying spatial databases [Papadias and Delis ’97] but here we use this for establishing region-identity over time
31: Applications - contd

Permutation of \((a',b',c')\) and assignments for model2

Uncorelated models: model1 (left), model2 (right)

**CND\textsubscript{RCC8}** similarity measures for model1 and model2

\[
\text{Dmin} \Phi(X,Y): \text{sums } d\text{min} \Phi(x,y) \text{ values for compared sets } X,Y \text{ of ground spatial relations for theory } \Phi, \text{ e.g.} \\
\text{Dmin} \Phi(X,Y)=14 \\
\text{where: } X=\{\text{EC}(a,b),... , \text{DC}(c,a)\}, \\
Y=\{\text{DC}(a',b'),... , \text{NTPPi}(c',a')\} \\
\text{and where: } a \rightarrow a', b \rightarrow b', c \rightarrow c'
\]
32: Applications - contd

<table>
<thead>
<tr>
<th>(a,b,c)</th>
<th>EC (a,b)</th>
<th>EC (b,c)</th>
<th>DC (a,c)</th>
<th>EC (b,a)</th>
<th>EC (c,b)</th>
<th>DC (c,a)</th>
<th>DminRCC8</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a',b',c')</td>
<td>DC (a',b')</td>
<td>TPP (b',c')</td>
<td>NTPP (a',c')</td>
<td>DC (b',a')</td>
<td>TPPi (c',b')</td>
<td>NTPPi (c',a')</td>
<td>14</td>
</tr>
<tr>
<td>(a',c',b')</td>
<td>DC (a',c')</td>
<td>TPP (c',b')</td>
<td>NTPP (a',b')</td>
<td>DC (c',a')</td>
<td>TPPi (b',c')</td>
<td>NTPPi (b',a')</td>
<td>10</td>
</tr>
<tr>
<td>(b',a',c')</td>
<td>DC (b',a')</td>
<td>TPP (a',c')</td>
<td>NTPP (b',c')</td>
<td>DC (a',b')</td>
<td>TPPi (c',a')</td>
<td>NTPPi (c',b')</td>
<td>14</td>
</tr>
<tr>
<td>(b',c',a')</td>
<td>DC (b',c')</td>
<td>TPP (c',a')</td>
<td>NTPP (c',b')</td>
<td>DC (c',b')</td>
<td>TPPi (a',c')</td>
<td>NTPPi (b',c')</td>
<td>10</td>
</tr>
<tr>
<td>(c',a',b')</td>
<td>DC (c',a')</td>
<td>TPP (a',b')</td>
<td>NTPP (c',b')</td>
<td>DC (a',c')</td>
<td>TPPi (b',a')</td>
<td>NTPPi (b',c')</td>
<td>14</td>
</tr>
<tr>
<td>(c',b',a')</td>
<td>DC (c',b')</td>
<td>TPP (b',a')</td>
<td>NTPP (c',a')</td>
<td>DC (b',c')</td>
<td>TPPi (a',b')</td>
<td>NTPPi (a',c')</td>
<td>14</td>
</tr>
</tbody>
</table>

Min combined DminRCC8 + DminLeft value = 14

<table>
<thead>
<tr>
<th>(a,b,c)</th>
<th>L(a,b)</th>
<th>L(b,c)</th>
<th>L(a,c)</th>
<th>R(b,a)</th>
<th>R(c,b)</th>
<th>R(c,a)</th>
<th>DminLeft</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a',b',c')</td>
<td>R(a',b')</td>
<td>O(b',c')</td>
<td>O(a',c')</td>
<td>L(b',a)</td>
<td>O(c',b')</td>
<td>O(c',a')</td>
<td>8</td>
</tr>
<tr>
<td>(a',c',b')</td>
<td>R(a',c')</td>
<td>O(c',b')</td>
<td>O(a',b')</td>
<td>L(c',a)</td>
<td>O(c',b')</td>
<td>O(b',a')</td>
<td>8</td>
</tr>
<tr>
<td>(b',a',c')</td>
<td>R(b',a')</td>
<td>O(a',c')</td>
<td>O(b',c')</td>
<td>L(a',b)</td>
<td>O(c',a')</td>
<td>O(c',b')</td>
<td>4</td>
</tr>
<tr>
<td>(b',c',a')</td>
<td>R(b',c')</td>
<td>O(c',a')</td>
<td>O(b',a')</td>
<td>L(c',b)</td>
<td>O(a',c')</td>
<td>O(a',b')</td>
<td>8</td>
</tr>
<tr>
<td>(c',a',b')</td>
<td>R(c',a')</td>
<td>O(a',b')</td>
<td>O(c',b')</td>
<td>L(a',c)</td>
<td>O(b',a')</td>
<td>O(c',b')</td>
<td>4</td>
</tr>
<tr>
<td>(c',b',a')</td>
<td>R(c',b')</td>
<td>O(b',a')</td>
<td>O(c',a')</td>
<td>L(b',c)</td>
<td>O(a',b')</td>
<td>O(a',c')</td>
<td>4</td>
</tr>
</tbody>
</table>

CND<sub>Left</sub> similarity measures

CND<sub>RCC8</sub> similarity measures
33: Applications - contd

![Graphs showing ambiguity levels for different categories: RCC8, LOS14, L/R, A/B, and Total. The y-axis represents log ambiguity in number of models, and the x-axis represents steps.](image)
34: Retrospective thoughts

• What does this tell us about modelling and reasoning about perception and space?
  – Limit to which theory development can be done in isolation from intended application, c.f. Hayes’ Naïve Physics programme

• Cog Rob applications highlights physical interpretation of denotation and reference
  – symbol-grounding/anchoring
    • tracking of objects, handling similarity and change

• Advantages of axiomatic approach
  – Primitives explicitly factored out, formal theory as specification language for implemented sensors
35: Retrospective thoughts - contd

• Ontological issues
  – region-based ontology useful abstraction
    • but requires much pre-processing of visual sensor data
  – Topological, geometrical and metrical information equally important

• Representational and computational issues
  – Soundness, completeness, complexity,…
  – Other pragmatic constraints
    • E.g. resource-bounded systems
36: Summary

- Overview of Cog Rob research
- Discussed limitations, challenges and revision of ideas leading to a set of open questions:
  - Re-examining primitives, ontologies, need for specialised inference, matching theory to practice
- Cognitive Robotics useful domain for tackling big AI questions, even for spatial logics!