Developing motoric-perceptual categories for support relations

Nick Wilkinson

Supervisor; Prof. Giorgio Metta

Instituto Italiano di Tecnologia

nicholas.wilkinson@iit.it

The Brief

Task 5.6 Motoric and contextual information in object recognition

This line of research will address the problem of analyzing, modelling and building systems that can recognize objects by including contextual and motoric information.

This requires a new way of looking at the problem of learning to extract relevant information at each stage of processing (e.g. how to determine the visual features? What is the relationship between vision and touch in object recognition? How motor information structures the recognition process?).

Of particular interest is the determination of the link between object-related information and the pragmatic use of this information for the control of manipulation.

My angle...
Content before context – a blind alley?

- Traditionally, contextually embedded cognition has been regarded in AI as 'too hard', something to be approached once the 'basics' of object recognition are sorted.

- The result has been toy problems, object recognition in a vacuum, and the deliberate removal of contextual effects from experimental designs.

- Li et al (2004) show that both sensory and behavioural context causes significant changes in neural mode of function as early as V1 in monkeys.

- A widespread view is that contextual understanding is mediated by associating bits of content into 'context frames' (Bar [2004]) or sensorimotor 'simulations' (Barsalou [2009], Grush [2004]).

- The counter argument from Gestalt and Gibsonian psychology and phenomenology that contextual perception is not reducible to accumulations of features also has a long history.

- My view is more in line with the latter. I think we need to *start with context* or we are in danger of missing the point.

- Torralba & Oliva at MIT have published several papers describing computational models of large scale scene analysis which outperform content first saliency based models in predicting human eye movements and pure performance (e.g. Torralba, Murphy & Freeman [2010]).

Objects before relations? Maybe...

- A popular assumption is that cognitive systems learn to recognise and classify objects/features, *and then* abstract relationships between objects.

- So objects are concrete, while relations are abstract (because the same relation can hold between many different objects).

- This may be the case, but then again it may not. It is not true *a priori.* Certainly it has proved enormously difficult to approach the learning of relationships in this manner with computers and robots.

- Recognition of relations may be less, more, or equally fundamental and concrete as recognition of objects (because the same object can link many different relations).

- Perhaps cognitive systems learn to recognise and classify relations, *and then* abstract objects as nodes between relationships. Perhaps they do both.

- I am exploring the idea that objects are afforded the opportunity of individuation and recognition as nodes in the networks of relationships in which they play a part, an idea related to Gibson's notion of 'affordances' (Gibson [1977]), though not identical.

– This approach implies that objects are always situated within a relational context, or they are never individuated from the background gist, establishing the following relevance criteria for object recognition;

Objects in a relational vacuum are irrelevant.
General relevance increases as an object connects to more relations.
Immediate relevance increases as an object connects to more immediate relations.

Support – an intimate relationship with a forgotten object

- A land animal's relationship with the ground is general, immediate and vital. As such it may give relevance to a wide variety of objects.

- How can we artificially establish a similar, motivated relationship in a dead mechanism like a robot?

- di Paolo (2003) suggests a short-cut to pseudo-motivation by which we may be able to skirt around the ultimate problem of the robot's lack of metabolism.

- This is based on the development, homeostatic entrenchment and perturbation of habitual behavioural preferences. A 'way of life', to compensate for the lack of a life...

- Radical perturbations push sensorimotor activity beyond homeostatic boundaries, triggering increased plasticity in the control network, which is then subject to change until a within bounds behavioural trajectory and thus sensorimotor alignment with the stimulus stream is regained.

- The key to generating pseudo-motivation is then that these entrenched sensorimotor habits are *continuously challenged*, requiring and initiating ongoing adaptive change in response.

- This establishes a bi-causal relationship between goals and the agents ongoing internal structure, rooted in the agents historical context, rather than the more usual case where 'goals' are epiphenomena of human programming and arbitrary value systems.

- We can apply this method to the development of a 'motivated' relationship with the ground in the iCub using gravity, buoyancy, and principles of dissipative systems.

Prenatal Robotics??

- Birth is not the beginning of a life. For the neonate animal, it is more a radical change of conditions.

- There is extensive evidence that behaviours such as locomotor like movement emerge prenatally (e.g. Robinson & Kleven [2008]).

- Land animals begin their motor development suspended in dense fluid, close to neutrally buoyant.

- Through neutral buoyancy suspension, the parent system provides a 3 dimensional supporting surface to the child system, effectively providing equilibrium conditions for movement in three dimensions.

- Within bounds sensorimotor activity naturally comes to depend on this environment, where the supporting medium may be regarded as *part of the controller* for the activity.

- The natal change of medium, resulting in transference to a 2D supporting surface presents a *continuous challenge* to the stability of these behaviours, triggering plasticity which will engender motor control adapation until the lost third dimension of support is compensated by motor activity.

- The change in buoyancy relations changes the dissipative form of the system; gravity now dissipates potential energy, such that the systems DOF's tend towards 2D equilibrium configurations parallel to the supporting surface. The supply rate rate comes to depend fundamentally on the energetic contribution of the ground. The storage function is the tension maintained in the musculo-skeletal system.

- Thus an intimate physical relationship develops between the robot and the ground, which provides the only means of obtaining the continuous energy supply necessary to compensate gravity and sustain or recover the dynamical stability of behaviours which are now operating far from equilibrium.

Neural and postural context

- The musculo-skeletal posture taken by an animal helps to define the possible sensory and motor dynamics available to it.

- I think that a similar situation may hold for what one might call 'neural posture'; the dynamical establishment of sensitive surfaces in the brain through mass interaction of populations of excitatory and inhibitory neurons, which play a similar role to agonist-antagonist muscular activity.

- Cell membranes held close to threshold are *sensitive* to input and information rich in output while over excited/inhibited cells are *insensitive* to input and information poor in output, as are those whose signal is drowned in neural noise.

- These sensitive surfaces move actively and dynamically through the neural mass, helping to define the sensory and motor dynamics available to the system.

- Brain events interact by perturbing the continuous field defined by coupled sensitive units.

- Neural sensitive surfaces can be extended to parts of the brain much as our external sensitive surfaces (retina, hand, eardrum etc.) can be extended to parts of the world.

- The engineering task is then to design a neuro-muscular-skeletal system which evolves towards a postural trajectory appropriate to it's ongoing situation, such that useful sensorimotor dynamics are made available, while irrelevant ones are filtered by insensitivity, a state referred to by Merleau-Ponty (1962) as 'maximal grip'.

- This process generates context by selective sensitivity to input and selective availability of output.

- Sensorimotor 'simulation' is content generative rather than directly context generative.

Implementation

- 'Equilibrium point' models of motor control (Feldman [2009]) are highly compatible with this approach.

- The neural mass should consist of a highly connected, loosely hierarchical, strongly non-uniform network of dynamical units connecting exo-sensory (inc. proprioceptive) and motor surfaces. The literature on liquid state machines provides a strong theoretical and practical basis for development.

- Global excitatory and inhibitory populations establish sensitive surfaces in the neural mass. Evolution of the trajectories of these manifolds is a geometric and topological issue.

- 'Learning media' undergo long-term changes in structure in response to patterns of interaction at sensitive surfaces which are accompanied by reinforcement signals. Useful mechanisms include Hebbian learning (e.g. STDP), frequency analysis and self-organising maps.

- Ongoing changes of state in trained learning media can perturb neural posture at their level in the hierarchy directly and at other levels indirectly. So content affects context.

- Plasticity and reinforcement are regulated by developmentally established distributed homeostatic boundaries to the activity of the neural mass.

- Beyond or near bounds activity in some brain area attracts sensitive surfaces and mediates reinforcement signals to the learning media.

- These functional modules are not necessarily separate physical entities. A given unit or population may contribute to more than one module.

Experiments and tests

- A. Initial development and testing of a simple version of the control architecture on head and eye control in the iCub – learn useful seeing/searching postures and techniques whilst dealing with the 'heavy head' problem.

- B. Develop homeostatic boundaries in the neural mass and appropriate structure in the learning media while performing whole body behaviours such as crawling in dense fluid suspension (in simulation). Radically and continuously perturb these entrenched structures by removal of the dense fluid. Analyse the iCub's ability to dynamically adapt to the perturbation by the development of new structures and boundaries. Use the information gained to improve the control architecture.

- C. Learn about different types of ground in terms of the support relations which they afford and the states they induce in multi-modal learning media. Interesting possibilities include supportive vs. subsiding grounds, visual cliffs, slopes and uneven terrains. Develop adaptive structures to deal with different ground types with 'maximal grip'.

- D. Extend the range of entrenched behavioural preferences, the range of ground types (e.g. include a couch), and the level of human-robot interaction to teach the iCub to use objects in the world to help it stand up or maintain other unstable far from equlibrium positions and behaviours.

- E. Use the sensorimotor categories so far generated to try to extend the enactive 'concept' of support to relations between objects in the world which are not the agent, working from the ground up. A nice example would be playing the 'I knock it off, you pick it up again' game.

- F. Attempt to integrate disparate, implicit sensorimotor categories into abstract explicit categories through the use of linguistic and other gestural symbols in interaction with a human.

Shared interests and opportunities for collaboration

- This project shares a lot of ground with other parts of the RobotDoc program.

- A serious attempt to develop and exploit contextually embedded robotic cognition and integrate multi-modal motor and sensory function will have to involve;

a. understanding and exploitation of physical dynamics and bio-mechanical constraints,
b. endogenously generated motivational and affective mechanisms rooted in historical context and continuous developmental learning,
c. neural-computational frameworks for effective implementation,
d. the explication and abstraction of emergent, implicit sensorimotor categories through human-robot interaction and the use of world-based symbol systems.
e. the testing of robot behaviour against that of real infants and the use of robotic models to generate predictions to guide experimental research.

- I think that's just about everybody!

- So if anyone would be interested in collaborations, or can tell me why this or that won't work before I waste my time trying, please come and talk to me about it. Let's help each other get to grips with the various fiendishly difficult problems which we have somehow gotten ourselves involved in...

Thank you all for your time and attention.

{:O)

References

Barsalou, L.W. (2009). Simulation, situated conceptualization, and prediction. Philosophical Transactions of the Royal Society of London: Biological Sciences, 364, 1281-1289.

Di Paolo, E. A., (2003). Organismically-inspired robotics: Homeostatic adaptation and natural teleology beyond the closed sensorimotor loop, in: K. Murase & T. Asakura (Eds) Dynamical Systems Approach to Embodiment and Sociality, Advanced Knowledge International, Adelaide, Australia, pp 19 – 42.

Feldman AG (2009) New insights into action-perception coupling. Exp Brain Res 194(1):39–58

James J. Gibson (1977), The Theory of Affordances. In Perceiving, Acting, and Knowing, Eds. Robert Shaw and John Bransford

Grush, Rick (2004). The emulation theory of representation: motor control, imagery, and perception. Behavioral and Brain Sciences 27:377-442.

Li, W. et al. (2004) Perceptual learning and top-down influences in primary visual cortex. Nat. Neurosci. 7, 651–657

Merleau-Ponty, M. (1962) Phenomenology of Perception, trans. C. Smith, London: Routledge & Kegan Paul.

Robinson, S.R., Kleven, G.A., & Brumley, M.R. (2008). Prenatal development of interlimb motor learning in the rat fetus. Infancy, 13, 204-228.

A. Torralba, K. Murphy, W. T. Freeman (2010) Using the forest to see the trees: object recognition in context. Communications of the ACM, Research Highlights, 53(3): 107-114, 2010.