Perception for Grasping and Manipulation


Early 1980s at UMass Amherst: *Laboratory for Perceptual Robotics*: Salisbury hand, visual input, tactile sensors.

Michael Arbib
arbib@usc.edu
Control Architecture for the Belgrade/USC Hand

George A. Bekey¹
Rajko Tomovic²
Ilija Zeljkovic¹

Figure 7.3: Schematic diagram of finger-side view.

Figure 7.4: Finger motion during flexing (©IEEE, Rao et al 1988).
Part I

Perceptual and Motor Schemas

Affordances and Effectivities
Coordinated control program for reaching and grasping

Visual Location

- Visual input
- Target location
- Visual and kinesthetic input

Size Recognition

- Visual input
- Size recognition criteria

Orientation Recognition

- Visual input
- Orientation recognition criteria

Fast Phase Movement

Hand Preshape

- Visual input
- Orientation

Hand Rotation

- Visual input
- Visual, kinesthetic, tactile input

Actual Grasp

- Slow Phase Movement

Grasping

- Activation of reaching
- Activation of visual search

Perceptual schemas

Motor schemas

Jeannerod and Biguer 1979

Arbib 1981
Grasping & Opposition Spaces

Iberall, Bingham and Arbib 1986: An object may have different representations for different tasks.

Opposition axes may be set for:
- different parts of the object
- setting goals for preshaping and moving the hand
- .... and for manipulation

A Compound Grasp
Another Vocabulary

Deploying the terminology of J.J. Gibson and his followers:

★ Perceptual schemas identify affordances relevant to the current task
★ Motor schemas deploy effectivities to exploit affordances

A coordinated control program integrates and schedules these but not only these schemas
$\mathcal{RS}$ (Robot Schemas) is a model of distributed computation embodying our nested network approach to robot programming. This paper introduces the $\mathcal{RS}$ model and shows how it can be used to represent robot programs in an efficient and concise manner.
Integrated Learning of Grasps and Affordances: The ILGA Model
James Bonaiuto, Michael A. Arbib
Part II

Basic Parieto-Frontal Interactions for Visually Directed Hand Movements
Introducing AIP and F5 (Grasping) in Monkey

F5 - grasp commands in premotor cortex
Giacomo Rizzolatti

AIP - grasp affordances in parietal cortex
Hideo Sakata
The FARS (Fagg-Arabib-Rizzolatti-Sakata) Model
A Focus on “How” and “What”
Mirror Neurons: Learning from the Macaque

This example: a precision pinch

A mirror neuron is active for execution of a limited set of actions & observation of a congruent set of actions
Recognition of oro-facial actions by humans

Actions belonging to the observer’s motor repertoire are mapped on the observer's motor system.

My Addendum: All actions can be recognized by other routes, whether or not in cooperation with mirror systems.
MNS Model of Learning in the Mirror Neuron System

The first demonstration of mirror neurons as formed through learning

Original Model: Oztop & Arbib, Biological Cybernetics, 2002
MNS2 extends MNS: Again, much more than just Mirror Neurons

Examples: (1) Bringing in audition
(2) Partially Hidden Grasps: Data from Umiltà et al., 2001

Working memory and dynamic remapping of hand working memory allows hidden grasps to be recognized

No response if object is not visible and not in working memory
The usual story: Recognizing the actions of others supports
1) cooperation and competition between individuals
2) Imitation
   ✴ Caveat: Monkeys have little if any capacity for imitation as
   compared to apes. So imitation involves further mechanisms
   than mirror neurons *simpliciter*

A complementary (and perhaps evolutionarily prior) function
implicates mirror neurons in observing one’s own actions
Alstermark’s Cat – Flexible Action Patterns and their Rapid Reorganization

From Alstermark et al. (1981)
Three Key Ideas

Motor schema activation determined by

- **Desirability** – Dynamically updated via reinforcement learning
- **Executability** – Determined by affordances and probability of action’s success

A New Role for Mirror Neurons: What Did I Just Do?

- An observation/execution matching (mirror) system may contribute to rapid reorganization of motor programs in the face of disruption when a known schema can be recognized as “filling the gap” for disrupted schemas by updating executability and observability estimates
The Augmented Competitive Queuing (ACQ) system

Bonaiuto & Arbib, *Biological Cybernetics*, 2010
Part III

Evolution, All too Briefly

Macaque

Chimpanzee

Human
Oxford University Press, 2012

Several steps beyond
Arbib (2002, 2005)

For 12 critiques and a Response

Language and Cognition
2013 5(2/3)
Edited by David Kemmerer

This work was supported in part by
National Science Foundation
Grants No. 0924674 & BCS-1343544

how the Brain got Language
The Mirror System Hypothesis
A Key Hypothesis: Bringing in Complex Imitation

i) Monkeys (and LCA-m): Little capacity for imitation

ii) Apes (and LCA-c): A capacity for simple imitation based more on attention to subgoals than to how movements are shaped to achieve them

ii) Hominid Evolution yields a complex imitation system:
The abilities to

- recognize another's performance as a set of familiar movements
- use this recognition to repeat the performance, and

More generally: the ability to

- recognize that another’s performance combines variants
- approximate the performance on this basis, with increasing practice yielding increasing skill.

Note utility for language learning and use once this can be applied to words and word streams – but it evolved (we claim) to support praxis
The Mirror System Hypothesis (2012):
A) Evolving the Language-Ready Brain and Protolanguage

**Pre-Hominid:**
Grasping
A mirror system for grasping: LCA-m
A simple imitation system for grasping: LCA-c

**Hominid Evolution**
A complex imitation system
“Naïve” Pantomime: Adapting the action repertoire to open up communication

- **Protosign:** a manual-based communication system, breaking through the fixed repertoire of primate vocalizations to yield an open repertoire for semantic expression
- **Protospeech and multi-modal protolanguage:** resting on the invasion of the vocal apparatus by collaterals from the protosign system
How the Brain Got Language, Chapter 10: Once early Homo sapiens emerges, cultural evolution dominated biological change:

From protolanguage to language; Emergence of grammar: Complex imitation supports co-emergence of

- Phonology
- Lexicon
- Constructions
Dyadic Brain Modeling
A Hypothetical Example: Beckoning

Child has distal goal: Social bonding – getting mother to hug him:

1) Child reaches out, grabs, and tugs on Mother, leading Mother to move towards Child as a response.

6) Child beckons Mother to move towards her.

Computational Comparative Neuroprimatology: Each Brain Script Extends the ACQ Model in the same way
Not only do hands evolve, but so too do the pathways that open up new possibilities for perception and control.

Process Versus Product in Social Learning:
Comparative Diffusion Tensor Imaging of Neural Systems for Action Execution–Observation Matching in Macaques, Chimpanzees, and Humans.

Part IV

Distalization of the End Effector
Distalizing the End-Effector

**Grasping:** the hand as end-effector

**Manipulation:** First hand then object as end-effector
Another two-stage action

Contrast doing this with vision and in the dark

- Feedback &/vs feedforward
- Bringing in haptics and proprioception
Tool Use

Phase 1: Grasping the Tool: The hand is the end effector: attention to the relation of hand and tool guides the action

Subsequent Phases: Using the Tool: The tool becomes the end effector: attention to the relation of the “business end” of the tool and object guides the action

Another issue: Bimanual coordination
Case Study 1: Macaque postcentral neurons coding a body schema modified by tool use

Umiltà et al. (2008) recorded F5 and F1 neuron activity in monkeys trained to grasp objects using “normal pliers” and “reverse pliers.”

For the F5 neurons studied: activity correlated with the movements of the end effector, the jaws of the pliers.

For M1/F1 neurons: some discharged in relation to hand movements, others in relation to end effector motion.
Tool Use and the extension of perception

*The end effector migrates distally from the hand.*

- Yet it is still the hand that has to be controlled
- Visual attention is directed to the tip of the tool rather than to the hand itself
- Haptic feedback is still provided via the hand but its meaning (the deployed perceptual schema) depends strongly on the phase of the current task
Tool Use and the extension of perception

When Justus plays the cello:

- Visual attention?
- Haptic feedback
- Other senses ...