Sixth and Final Six-Monthly Meeting

Research Planning: The Creation of the Research Roadmap

Frankfurt Airport Conference Centre
Friday 11th February 2005

Background

Background – Research Roadmap Meetings

November 6th, 2002: Research Roadmap Workshop, Paris

February 13th, 2003: Research Roadmap Workshop, Amsterdam

[October 26th-31st, 2003: Dagstuhl workshop on cognitive vision]

[January 12th, 2004: Cognitive Vision Colloquium, Prague]

March 26th, 2004: Six-monthly meeting dedicated to the research roadmap

May 28th, 2004: Research Roadmap Workshop

September 24th, 2004: Six-monthly meeting dedicated to the research roadmap

Background – Research Roadmap Drafts

OFFICIAL VERSION

- Research Roadmap - V4.0 11-10-04 (pdf 416k)

DRAFT VERSION

- Research Roadmap - V3.3 04-10-04 (pdf 467k)
- Research Roadmap - V3.2 06-9-04 (pdf 415k)

LEGACY ROADMAPS

- Research Roadmap - V2.5 25-3-03 (pdf 145k)
- Research Roadmap - V2.4 02-3-03 (pdf 141k)
- Research Roadmap - V2.4 02-3-03 (doc 224k)
- Research Roadmap - V2.3 13-2-03 (pdf 61k)
- Research Roadmap - V2.3 13-2-03 (doc 127k)

The Emerging Discipline of Cognitive Vision

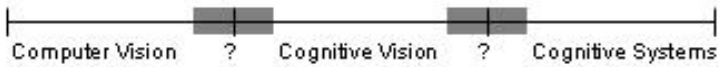
- Emerging discipline
- No one single established and widely-accepted paradigm
- We don't know what approach will win
- Need
  - An (accepted) definition of the discipline
  - Scientific foundations:
    a science of cognition (and a mathematical framework)
  - Development strategy
    - Research agenda
    - Transferable technology (algorithms, h/w, s/w)
    - Tools (development environments, languages, benchmarks)

Roadmap Objectives

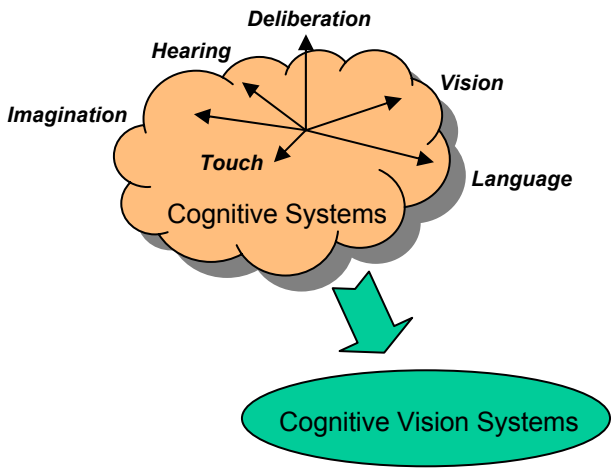
- Definition of cognitive vision
  - Paradigm-neutral
  - Requisite capabilities
  - Task-specific competences it enables
- Identify competing/alternative approaches
- Create inclusive research agenda
- Identify critical gaps in scientific understanding / technical know-how (& strategy for filling them)
- Highlight contentious and significant issues

Obstacles to progress

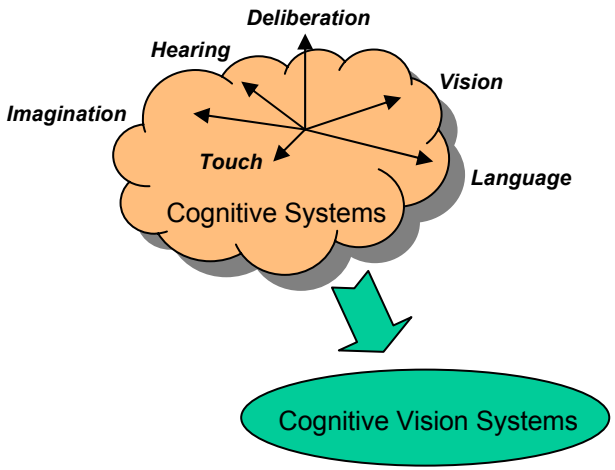
- Cognitive Vision as a part of the Computer Vision / Cognitive Systems spectrum



- Or Cognitive Vision as a projection of the space of Cognitive Systems

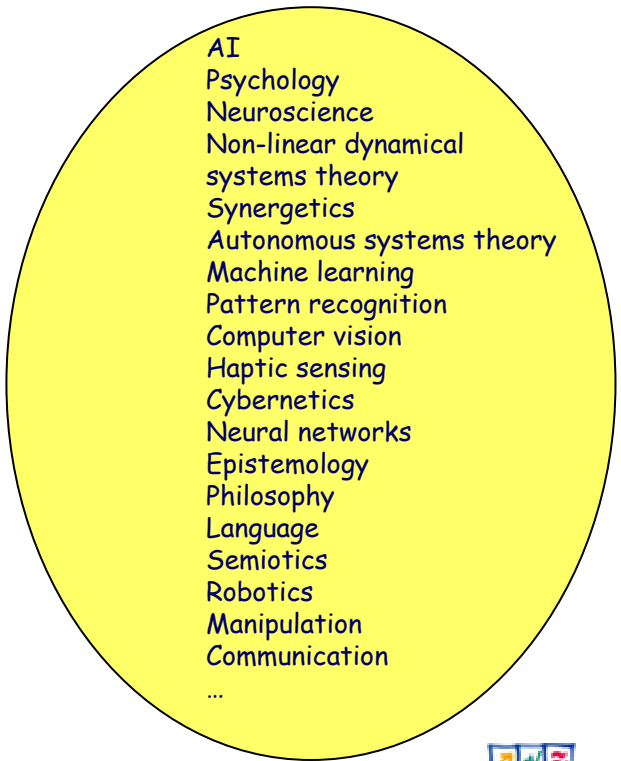


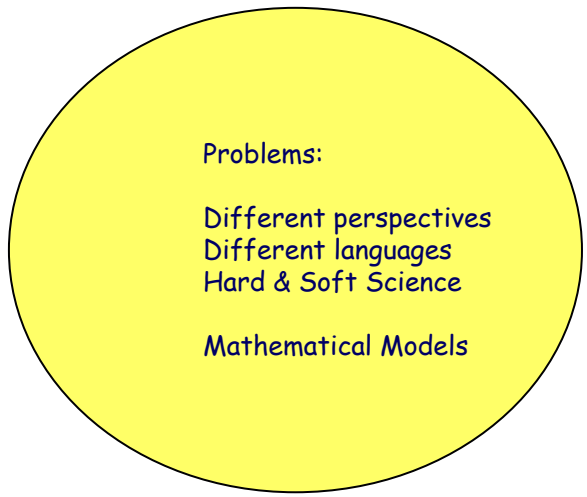
We chose to view cognitive vision as a
visually-enabled cognitive system



Obstacles to progress

- Multidisciplinary nature of cognitive vision
- Difficult to deal with, but essential



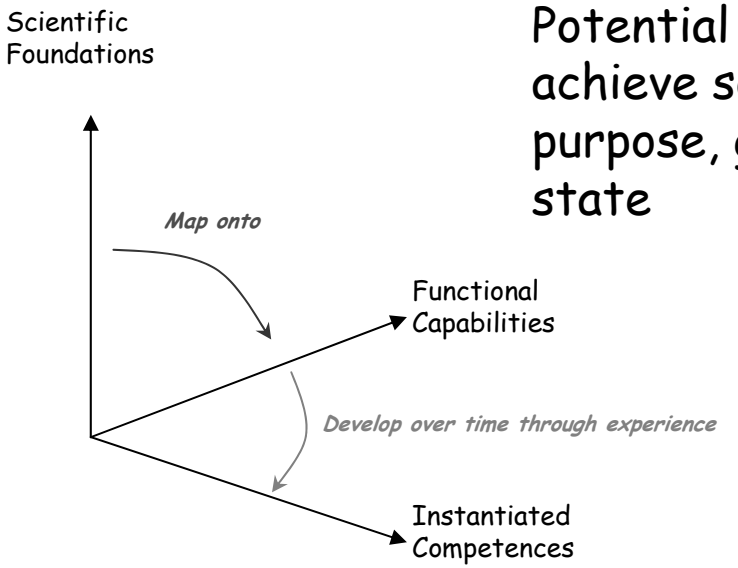


The Nature and Scope of Cognitive Vision

MULTI-FACETED

Constituent
theoretical models

The ECVision Model



Potential faculty to achieve some function, purpose, goal, or desired state

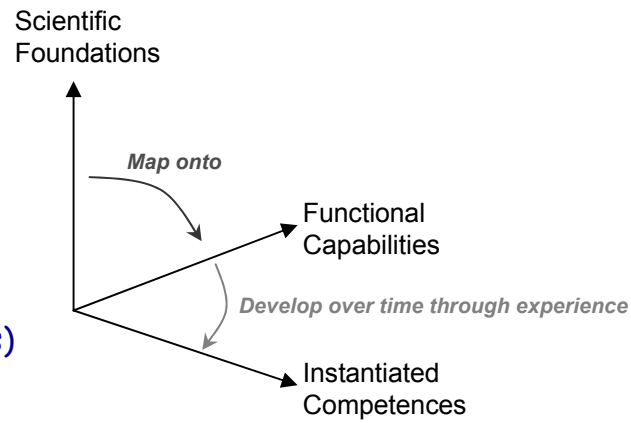
Actual power/faculty to achieve something (task-dependent)

Based on capabilities

Developed over time through experience, learning, practice

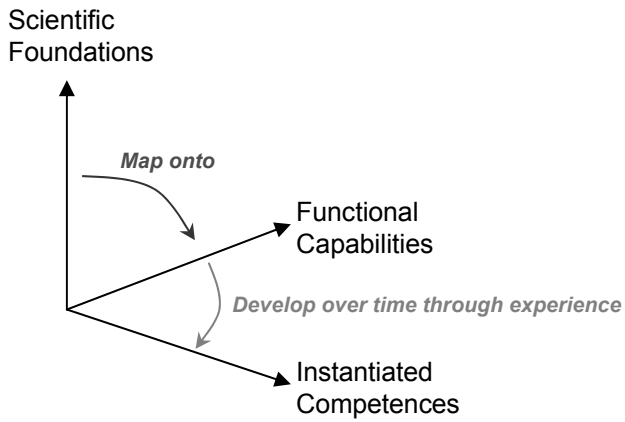
Scientific Foundations

- Sensing (visual, haptic, …)
  - Environmental perturbation of the system
- Architecture
  - Minimal configuration necessary so that development is possible (phylogeny/ontogeny trade-off)
- Representations
  - Stable states of a cognitive system (cognitivist vs. emergent representation; absolute vs. relative)
- Knowledge & Memory
  - Source of knowledge (designer, environment, emergent)
  - Persistence of states of a cognitive system
  - Episodic, modal, propositional, short-term, long-term
- Learning & Development
  - Acquisition of knowledge (model learning / parameter estimation)
  - Acquisition of capabilities (know-how, skill learning)
  - Modification of the system itself
- Recognition
  - Discrimination of sensed entities (perceptual events, inc. behaviours)
- Generalization
  - Class enlargement vs. categorization; perception of affordance
- Deliberation & Reasoning
  - Explicit weighing of alternative options in selecting a given action/behaviour (x 3)
- Planning
  - Ability to deal with future events (some of which may never happen!)
  - Anticipation, expectation, managing continencies
- Goal identification and achievement
  - Reactive vs learned self-initiated vs trained externally-initiated vs reflexive goals
  - Developmental driving forces
- Communication
  - As a consequence of the system's cognitive activities
  - Semiotic communication vs data-communication (direct external access to internal states)
- Action
  - Forcible interaction
  - Gesture & language



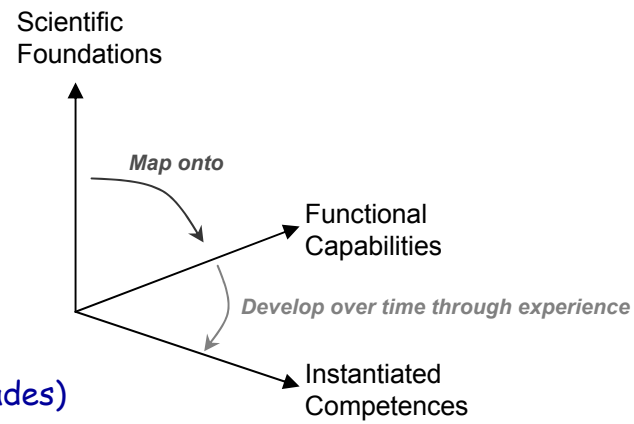
Functional Capabilites

- Detection and Localization
  - Discriminate between regions in the sensory field (vision: colour, texture, motion, distance, …)
  - Feature (*) extraction (hard-wired)
  - Detection of sensory entity (discrimination + features) including events
  - Localization (position & extent) cf FoR
  - Establish spatial relationships (occlusion / relative ordering)
  - Establish temporal relationships (simultaneity / relative ordering)
  - Novelty detection
- Tracking
  - Learned capabilities (tracking & entities)
  - Up to 4-D spatio-temporal paths
  - Hardwired & learned entities
  - Partial and complete occlusion (short-term prediction)
  - Long-term prediction (cf. expection and planning)
  - Spatial attention
    - Pre-attentive: detection and localization
    - Selective attention: classification, prediction, visualization
- Classification & Categorization
  - Entity grouping & meta-level functional classification
  - Classification of sensory entities
  - Classification of behaviours
  - Detect instances of a new class
  - Learn classes of entities and behaviours (from physical to social)
  - Form new classes from old ones (generalization)
  - Generalize to new contexts
  - Map from classes of objects to categories of affordances



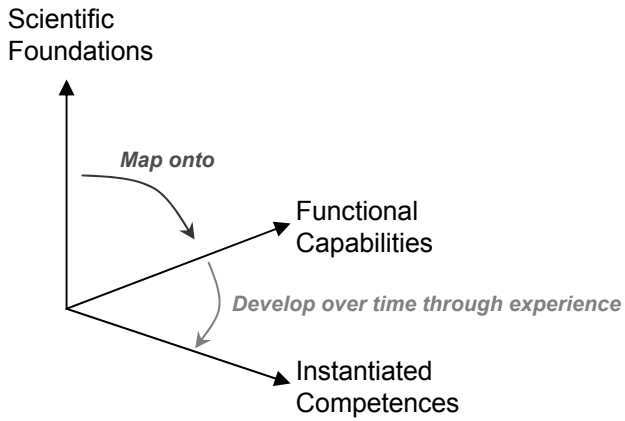
Functional Capabilites

- Prediction
  - Interpret intent of behaviour to predict future spatio-temporal configurations of the environment
  - Possibly form hypotheses and then deliberate (rhrough active exploration or passive assessment)
- Concept formation and visualization
  - Conceptualization (abstraction)
  - Visualization (instantiation)
  - Concept – Entity association
  - Conceptual entity decomposition (relational/hierarchical)
  - Extra-agent conceptualization
- Inter-agent communication & expression
  - Gesture / verbal
  - Understanding of the world
  - Role in driving the developmental process
- Visuo-motor coordination
  - Visual accommodation and focussing (continuous and after saccades)
  - Gaze control
  - Hand-eye / body-eye coordination (follow, reach for, grasp at, grip)
  - Learned coordination (hit, objects as body extensions)
  - Prediction of object, hand, body motion
  - Bodily self-awareness
  - Imitation of actions of others
- Embodied Exploration
  - Role in driving the developmental process
  - Discovery of new modes of interaction
  - 'Map formation' (eco- vs ego-)
  - Point to or gesture at entities (objects / people)
  - Learn affordances



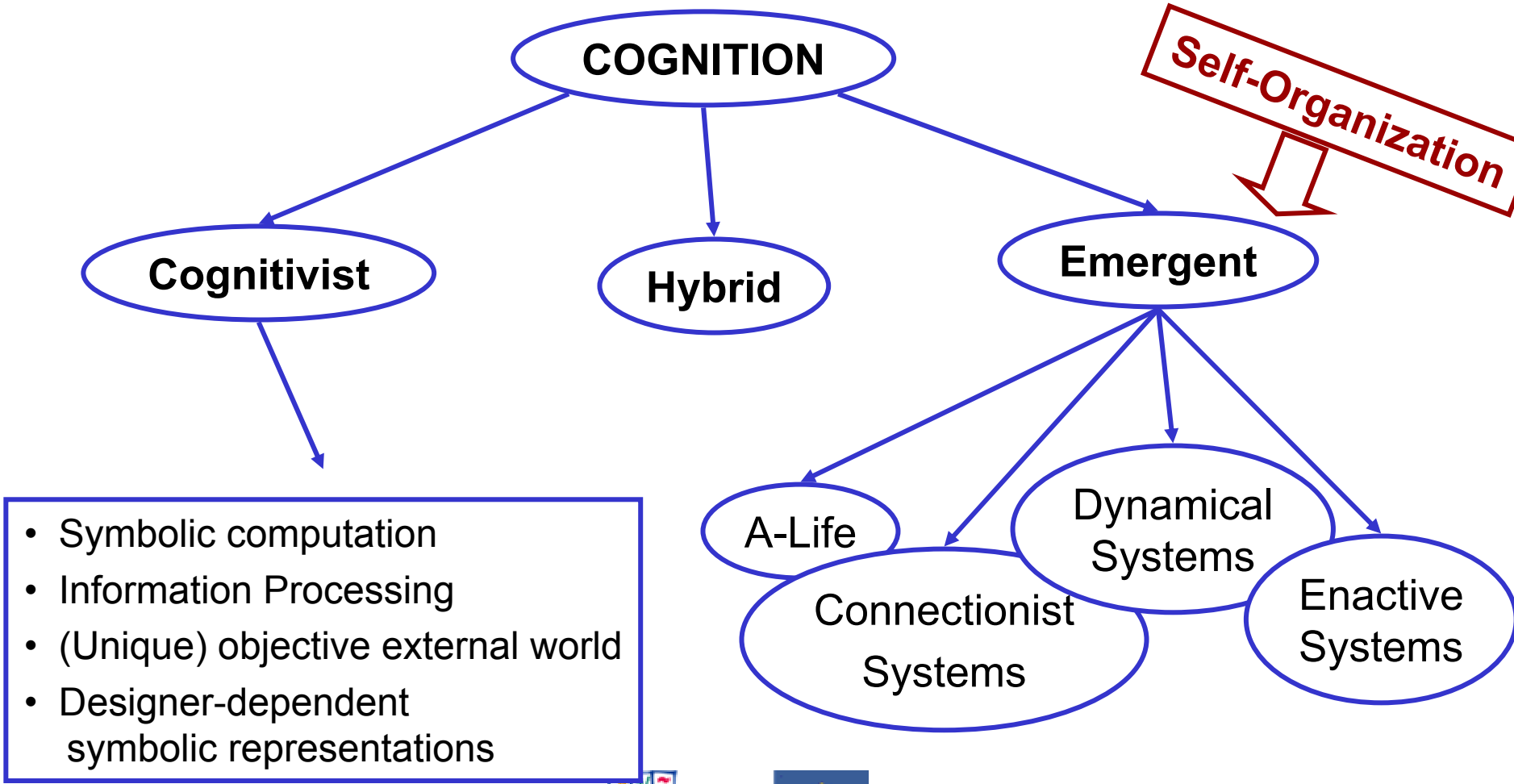
Instantiated Competences

- Cognition needs context
- Task-specific scenarios
- Provide a space for development
- Rich enough to require development
- Not too rich (overwhelm simple models!)
  - Cognitive Surveillance System
  - Cognitive Home Assistant
  - Advanced Driver Assistance System
  - Competences of Infants and Children



The Paradigms of Cognitive Systems

Paradigms of Cognitive Systems



- Symbolic computation
- Information Processing
- (Unique) objective external world
- Designer-dependent symbolic representations

Cognitivist Hybrid Connectionist Dynamical Enactive

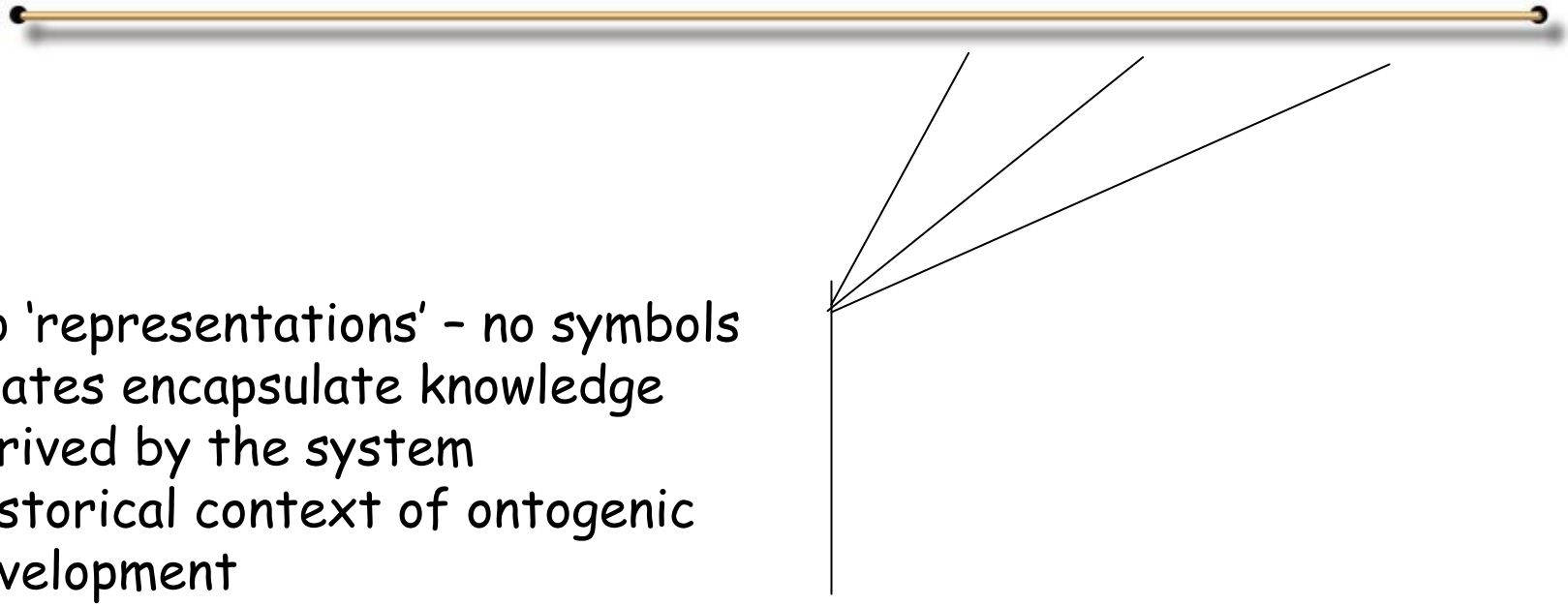
- Knowledge provided by human designers
- Possibly refined through learning
- Brittle
  Representations based axioms & assumptions of an external designer
- Fail when they stray far from the domain for which they were designed

Cognitivist Hybrid Connectionist Dynamical Enactive

- Generic representational schemes
- Populated entirely by the agent
- Incremental unsupervised learning
- Deliberative behaviour designed in

(link between perceptual-motor skill and knowledge representation controlled by designer)

Cognitivist Hybrid Connectionist Dynamical Enactive



- No 'representations' – no symbols
- States encapsulate knowledge derived by the system
- Historical context of ontogenic development

- Paradigms are not equally mature
- Dynamical systems
  - Arguments are compelling BUT ..
  - Not yet clear how to get higher-level cognition
- Cognitivist systems
  - More advanced
  - Not many achievements in generalization
  - More brittle (in principle)
- Enactive (& Dynamical)
  - SHOULD be much less brittle (mutual specification through co-development)
  - But very limited cognition at present
- Hybrid systems
  - Best of both worlds?
  - But unclear how one can really combine antagonistic philosophies

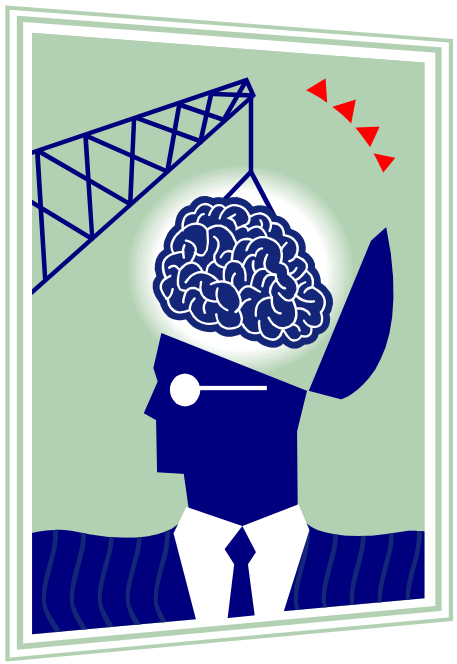
Core Scientific Concerns

- Phylogeny vs. Ontogeny
- The need for embodiment
  - An argument for embodied cognition
  - Types of embodiment
  - Impact of embodiment on perception
- Scientific challenges
  - Methods for continuous learning and development
  - Minimal architectures
  - Goal identification and achievement
  - Generalization
- Caveat

The Balance between Phylogeny and Ontogeny:

Hardwired Functionality
vs.
Learned Capabilities

Phylogeny and Ontogeny



Minimal set of sensory and motoric control structures

To boot-strap ontogenic development

No time to evolve a complete system

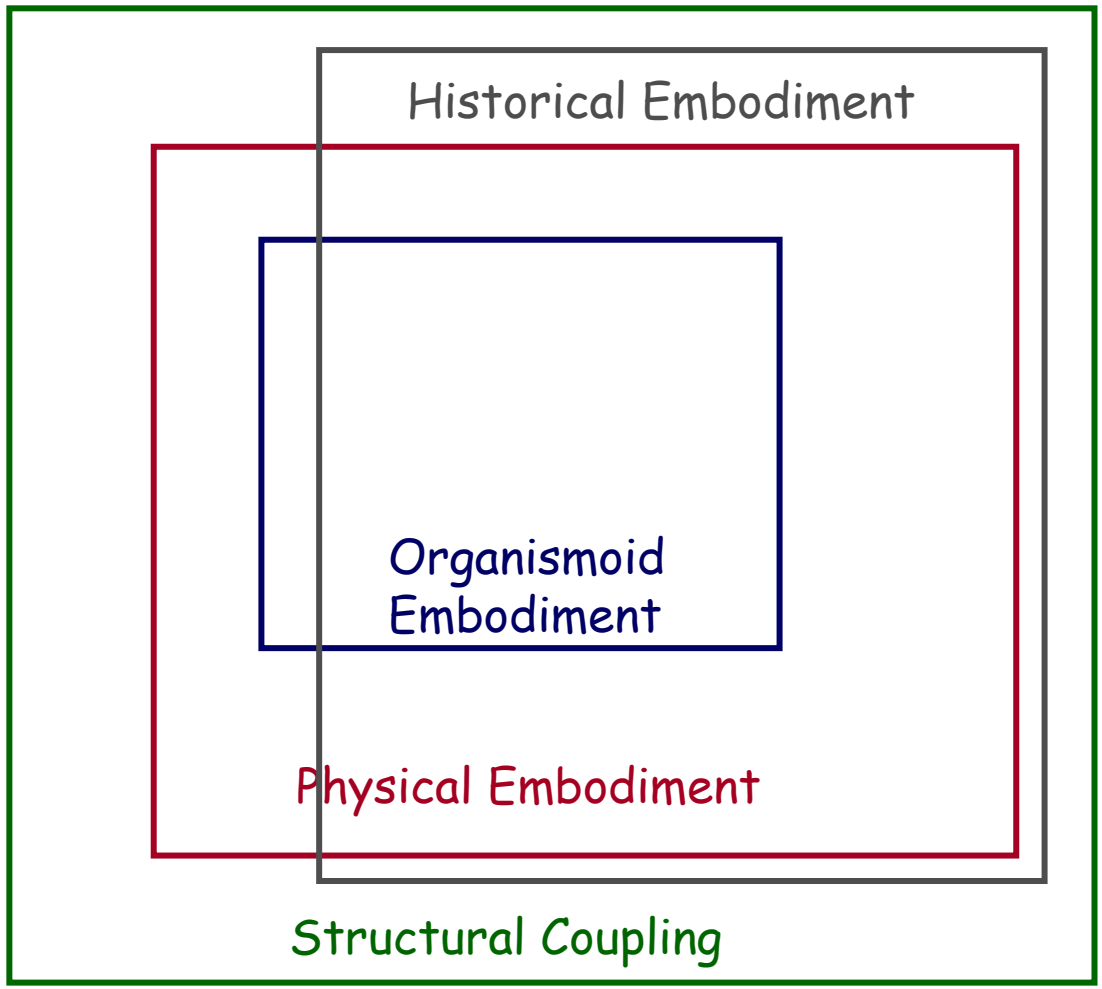
The Need for Embodiment

- Very divisive issue:
  different paradigms have different requirements

- Cognitivist:
  - there is no need for embodiment
    (cf. physical symbol system hypothesis)
  - Symbolic knowledge can be programmed in directly
  - Embodiment may be useful, but it's not necessary

- Emergent:
  - Must be embodied (by definition)
  - Embedded in their environment
  - Situated historical developmental context
  - Two complementary processes
    - Self-organization
    - Structural coupling
    - Co-development (ontogenic development in the context of environmental perturbations of the system)
  - Without physical embodied exploration, a cognitive system has no basis for development
  - Embodiment is a key component of system dynamics – morphology matters!

- Are there different forms of embodiment?
- What's an action?
- Does it have to be forcible?
- Is speech an act (or action)?
- Does action require mobility?
- Is a computer triggering a switch embodied?



From: T. Ziemke, 'What's this thing called embodiment?', 2003

- Structural coupling
  - System can be perturbed by the environment
  - System can perturb the environment
- Historical embodiment
  - History of structural coupling
- Physical embodiment
  - Forcible action (excluded software agents)
- Organismoid embodiment
  - Organism-like bodily form (e.g. humanoid robots)
- Organismic embodiment
  - Autopoietic living systems

Scientific Challenges

1. Methods for continuous learning and development
2. Minimal architectures
3. Goal identification and goal achievement
4. Generalization

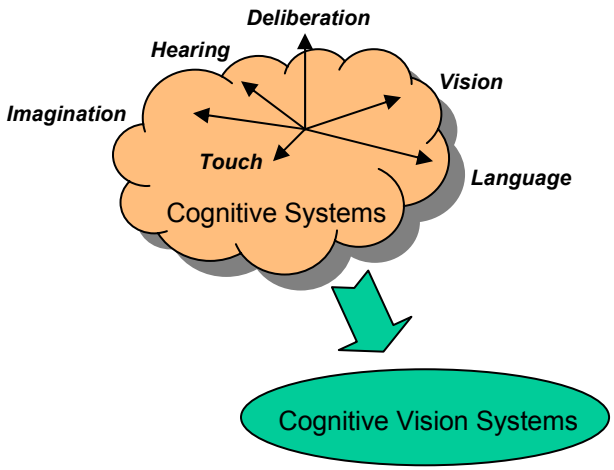
- Methods for continuous learning and development
  - Learning mechanisms: fast, incremental, continuous, large capacity, graceful degradation
  - Representations & feature sets: efficient and extendable
  - Learning domains: perceptual & conceptual, and mapping between them
    - Perceptual to conceptual mapping (to facilitate expression, deliberation, communication)
    - Conceptual to perceptual (to facilitate action)
  - Learn parameters (populate existing representations and mappings)
  - Learn system identification (develop new representations and mapping) … very hard!

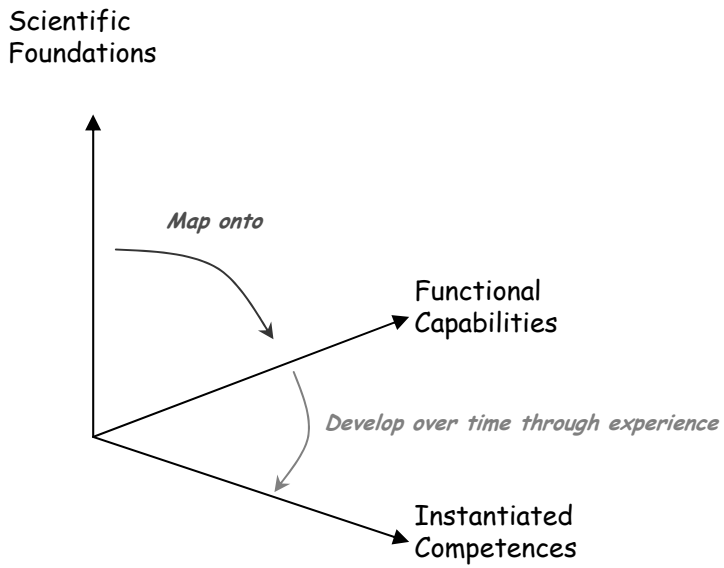
- Minimal Architectures
  - Balance between phylogeny and ontogeny
  - Difficulty in identifying constituents of each

- Goal identification and goal achievement
  - Cognitivist: easy as goals are specified explicitly as the outcome of cognitive behaviour
  - Emergent: hard as behaviour is a non-specific emergent consequence of a set of system dynamics
    - Stipulate goals as constraints or boundary conditions (phylogenic configuration or ontogenic development)
  - Four types of goals
    - Learned self-initiated goals (unsupervised)
    - Trained externally-initiated goals (supervised)
    - Reactive goals
    - Intentional or reflexive goals

- Generalization
  - Transferrability of competences or skills from one context to another
  - Requires faculty for generalization
    - Analogical reasoning
    - Metaphorical deliberation
    - Manipulation of visual memory
  - Paradigm-dependent
  - Does it require establishment of some form of inter-agent cognitive communication
    - Language
    - Speech
    - Gesture

Questions?


Deliberation
Hearing
Vision
Imagination
Touch
Language
Cognitive Systems
Cognitive Vision Systems


Scientific
Foundations
Map onto
Functional
Capabilities
Develop over time through experience
Instantiated
Competences