Consciousness and Creativity in Brain-Inspired Cognitive Architectures

Włodzisław Duch

Department of Informatics,
Nicolaus Copernicus University, Toruń, Poland

Google: W. Duch

CHIST-ERA, Rome, May 2010

A roadmap to human level intelligence

workshop organized by:

Włodzisław Duch (Google: W Duch)
Department of Informatics, Nicolaus Copernicus University, Torun, Poland
& School of Computer Engineering, Nanyang Technological Uni, Singapore

Nikola Kasabov (http://www.kedri.info)
KEDRI, Auckland, New Zealand

James Anderson, Paul Allopenna, Robert Hecht-Nielsen, Andrew Coward,
Alexei Samsonovich, Giorgio Ascoli, Kenneth De Jong, Ben Goertzel

WCCI'2006, Vancouver, , British Columbia, Canada, July 17, 2006

Steps Toward an AGI Roadmap

Włodek Duch (Google: W. Duch)

Artificial General Intelligence (AGI):
architectures that can solve many problems and transfer knowledge between the tasks.

Roadmaps:

- A Ten Year Roadmap to Machines with Common Sense (Push Singh, Marvin Minsky, 2002)
- Euron (EU Robotics) Research Roadmap (2004)
- Neuro-IT Roadmap (EU, A. Knoll, M de Kamps, 2006)

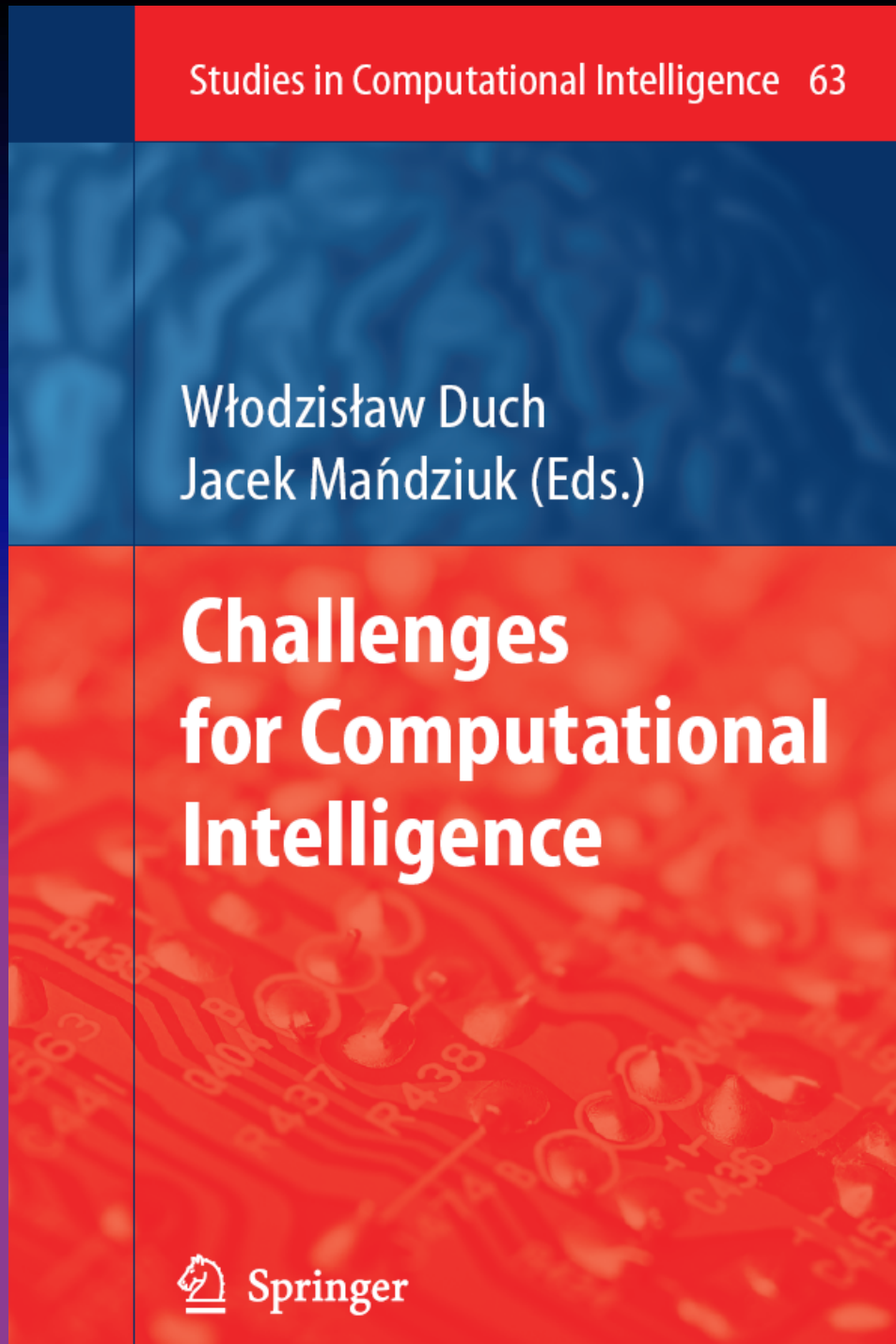
Challenges: Word games of increasing complexity:

- 20Q is the simplest, only object description.
- Yes/No game to understand situation.
- Logical entailment competitions.

Conference series, journal, movement.

In the year 1900 at the International Congress of Mathematicians in Paris David Hilbert delivered what is now considered the most important talk ever given in the history of mathematics, proposing 23 major problems worth working at in future. 100 years later the impact of this talk is still strong: some problems have been solved, new problems have been added, but the direction once set - identify the most important problems and focus on them - is still important.

It became quite obvious that this new field also requires a series of challenging problems that will give it a sense of direction.



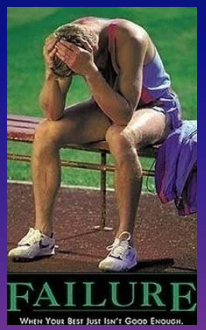
Failures of AI

Many ambitious general AI projects failed, for example:

A. Newell, H. Simon, General Problem Solver (1957).

Eduardo Caianiello (1961) – mnemonic equations explain everything.

5th generation computer project 1982-1994.



AI has failed in many areas:

- problem solving, reasoning
- flexible control of behavior
- perception, computer vision
- language ...

Why?

- Too naive?
- Not focused on applications?
- Not addressing real challenges?

Ambitious approaches…

CYC, started by Douglas Lenat in 1984, commercial since 1995.
Developed by CyCorp, with 2.5 millions of assertions linking over 150.000 concepts and using thousands of micro-theories (2004).

Cyc-NL is still a "potential application", knowledge representation in frames is quite complicated and thus difficult to use.

Hall baby brain – developmental approach, www.a-i.com, failed.

Open Mind Common Sense Project (MIT): a WWW collaboration with over 14,000 authors, who contributed 710,000 sentences; used to generate ConceptNet, very large semantic network.

Some interesting projects are being developed now around this network but no systematic knowledge has been collected.

Other such projects:

HowNet (Chinese Academy of Science),
FrameNet (Berkeley), various large-scale ontologies,
MindNet (Microsoft) project, to improve translation.

Mostly focused on understanding of all relations the in text/dialogue.

Challenges: language

- Turing test – original test is too difficult.
- Loebner Prize competition, for almost two decades played by chatterbots based on template or contextual pattern matching – cheating can get you quite far ...

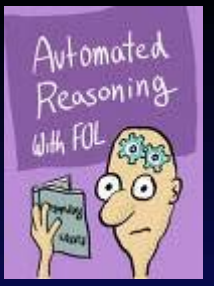
- A "personal Turing test" (Carpenter and Freeman), with programs trying to impersonate real personally known individuals.

- Question/answer systems; Text Retrieval Conf. (TREC) competitions.

- Word games, 20-questions game - knowledge of objects/properties, but not about complex relations between objects. Success in learning language depends on automatic creation, maintenance and the ability to use large-scale knowledge bases.

- Intelligent tutoring systems? How to define milestones?

Challenges: reasoning



- Super-expert system in a narrow domain (Feigenbaum), needs a lot of general intelligence to communicate, should reason in math, bioscience or law, experts will pose problems, probe understanding.

- Same direction, but without NLP: Automated Theorem Proving (ATM) System Competitions (CASC) in many sub-categories.

- General AI in math: general theorem provers, perhaps using meta-learning techniques with specialized modules + NLP.

- Automatic curation of genomic/pathways databases, creation of models of genetic and metabolic processes for bioorganisms.

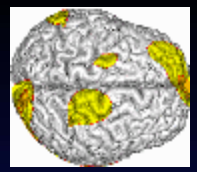
- Partners that advice humans in their work, evaluating their reasoning (theorem checking), adding creative ideas, interesting associations.

Real AI?

- General purpose systems that can be taught skills needed to perform human jobs, and to measure which fraction of these jobs can be done by AI systems (Nilsson, Turing's "child machine").
- Knowledge-based information processing jobs – progress measured by passing a series of examinations, ex. accounting.
- Manual labor requires senso-motoric coordination, harder to do?
- DARPA Desert & Urban Challenge competitions (2005/07), old technology, integration of vision, signal processing, control, reasoning.
- Humanoid robotics: understanding of perception, attention, learning casual models from observations, hierarchical learning with different temp. scales.
- "Cognitive Assistant that Learns and Organizes" (CALO), part of DARPA Personalized Assistant that Learns (PAL) call, SRI+21 inst.

  5-year project to create partners/personal assistants, rather than complete replacements for human workers (also CM RADAR).

Motivation & possibilities

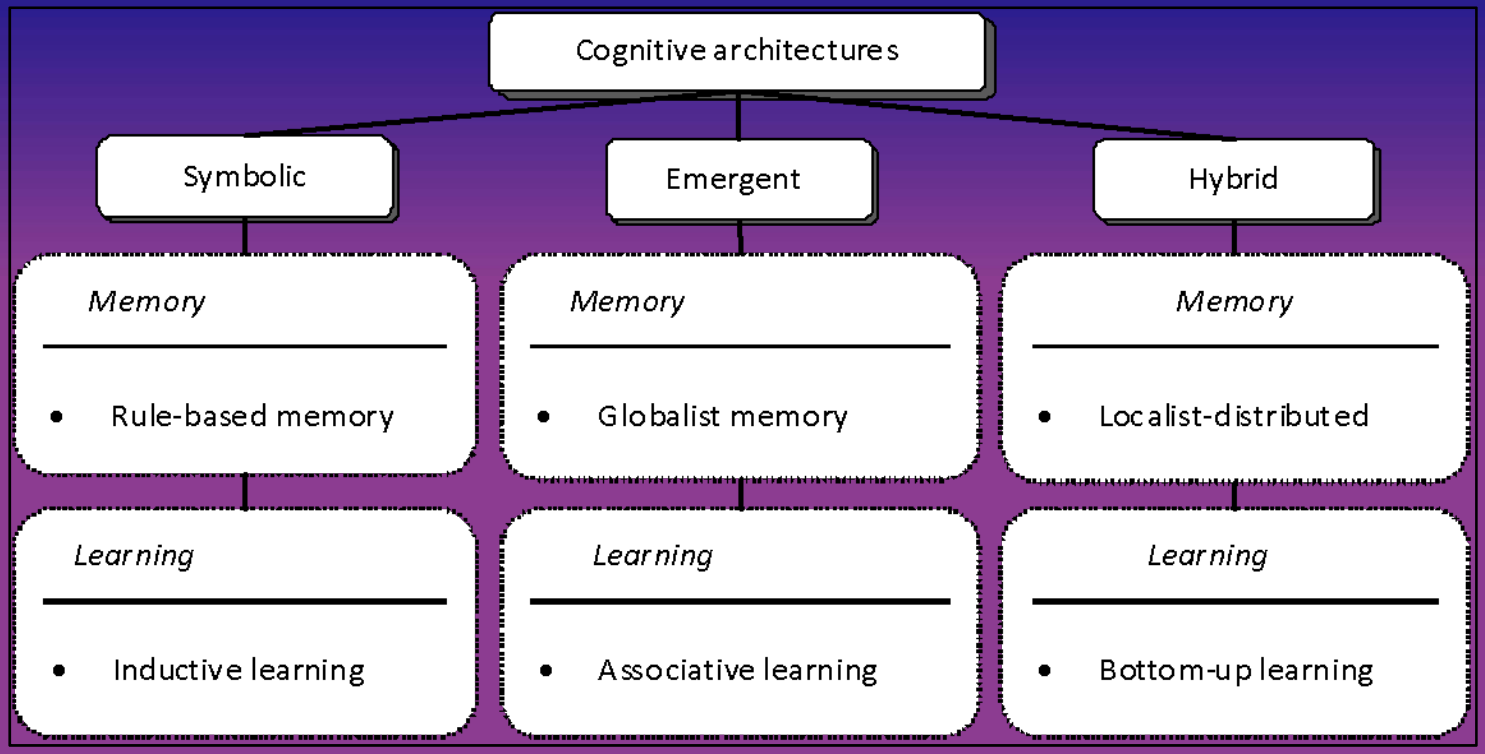


To reach human-level intelligence we need to go beyond pattern recognition, memory and control. How to reach this level?

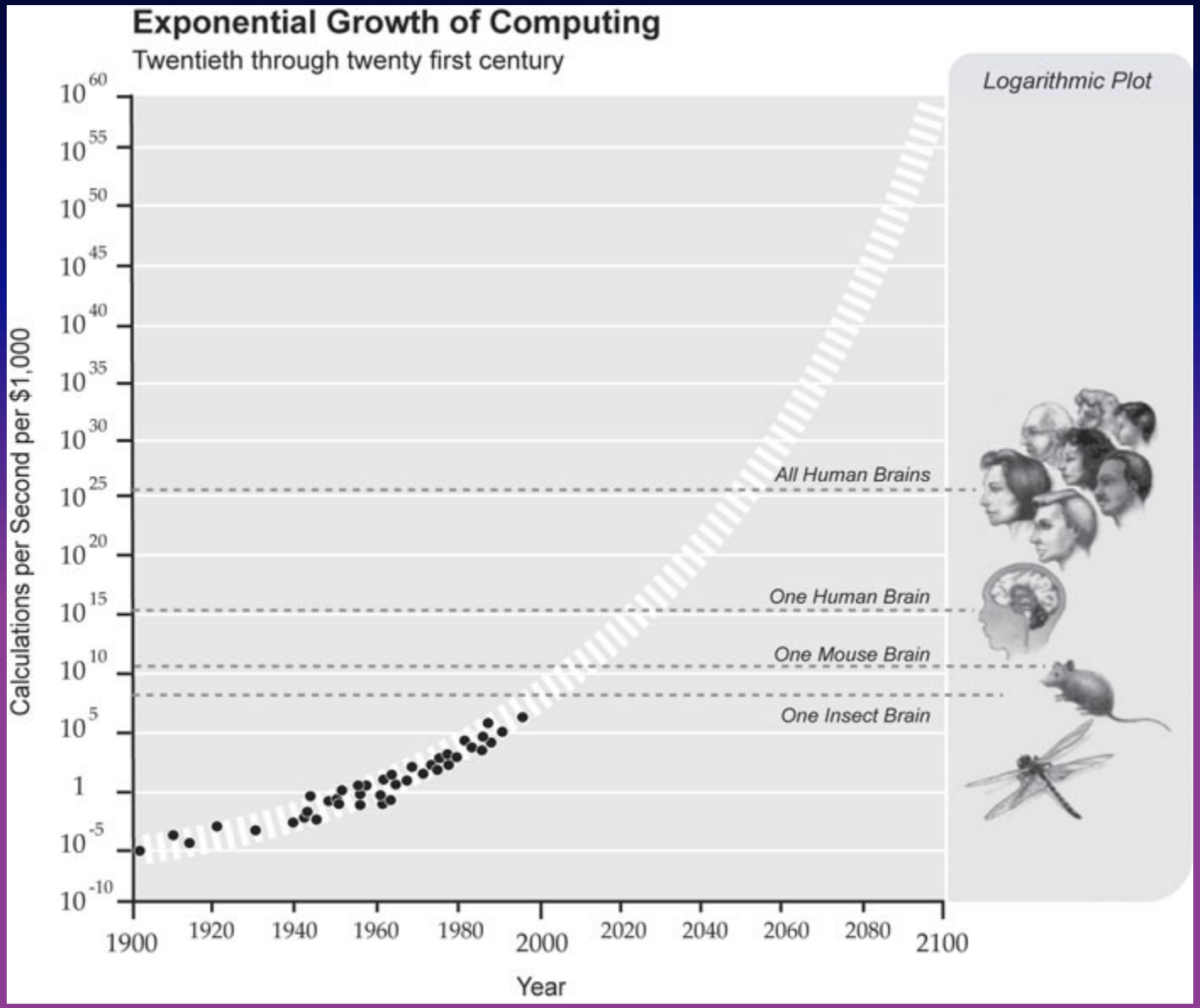
- **Top down style**, inventing principles: laminar & complementary computing (Grossberg), chaotic attractors (Freeman & Kozma), AMD (John Weng), confabulation (Hecht-Nielsen), symbolic cognitive architectures (SOAR, ACT-R, CLARION ...).
- **Bottom up style**: systematic approximations to brain activity: neuromorphic systems, CCN (Izhikevich), Blue Brain, Emergent ...
- **Mind models**: “shadows of neurodynamics”: models of mind processes based on psychological spaces, providing inner perspective as an approximation to neurodynamics.
- Designs for artificial brains based on cognitive/affective architectures Integration of perception, affect and cognition, large-scale semantic memory models, implementing control/attention mechanisms.

Cognitive architectures

- CA frequently created to model human performance in multimodal multiple task situations, rather than AGI.
- Newell, *Unified Theories of Cognition* (1990), 12 criteria for CS: behavioral: adaptive, dynamic, flexible; development, evolution, learning, knowledge integration, vast knowledge base, natural language, real-time performance, and brain realization.



Exponential growth of power



From
R. Kurzweil,

The Law of Accelerating Returns

By 2020 PC computers will match the raw speed of brain operations!

What about organization of info flow?

Special hardware?

What is needed: elements performing like a spiking biological neurons connected in the layered 2-D structures of mammalian cerebral cortex.

- Nanoscale Memristor Device as Synapse in Neuromorphic Systems (Univ. of Michingan); neural plasticity in hardware.
- ALAVLSI, Attend-to-learn and learn-to-attend with analog VLSI, EU IST Consortium 2002-2005, Plymouth, ETH, Uni Berne, Siemens.
- A general architecture for perceptual attention and learning based on neuromorphic VLSI technology. Coherent motion + speech categorization, ended in 2005, new projects to create low-poer analog spiking neuron chips along these lines are ongoing.
- RobotCub open-architecture for Cognition, Understanding, and Behavior, developing a humanoid robotic platform shaped as a 2.5 years old child, a lot of hardware for control.
- P-RAM neurons developed at KCL.
- ASSO-IC for Haikonen's associative architecture (VTT, planned).

Brain-like computing

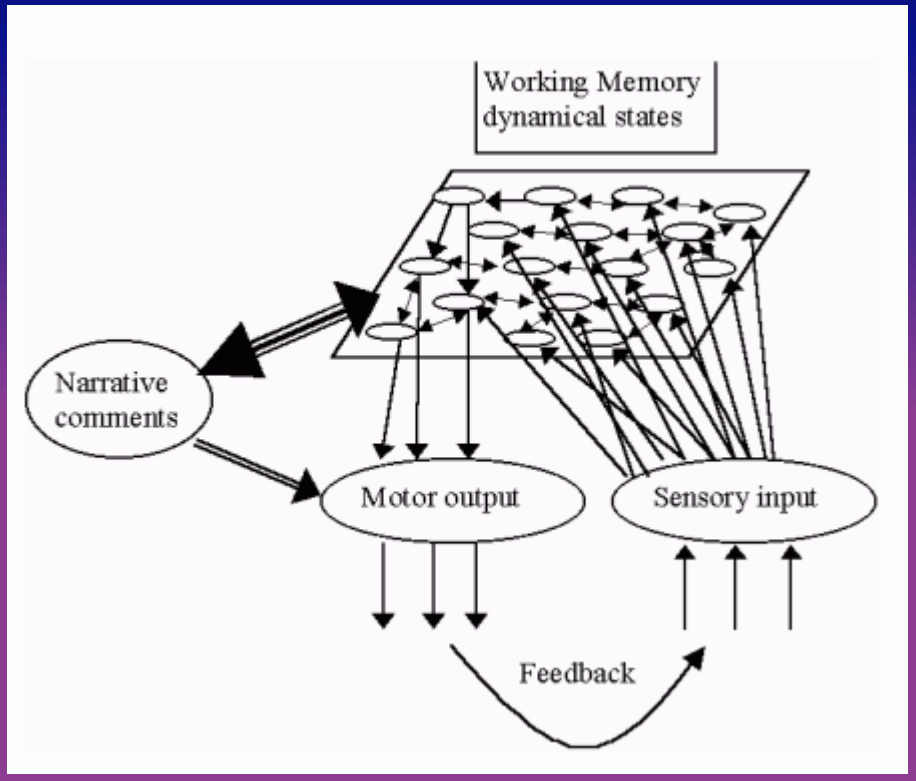
Brain states are physical, spatio-temporal states of neural tissue.

- I can see, hear and feel only my brain states! Ex: change blindness.
- Cognitive processes operate on highly processed sensory data.
- Redness, sweetness, itching, pain ... are all physical states of brain tissue.

In contrast to computer registers, brain states are dynamical, and thus contain in themselves many associations, relations.

Inner world is real! Mind is based on relations of brain's states.

Computers and robots do not have an equivalent of such WM.



Towards conscious robots

Few explicit attempts to build them so far.

Stan Franklin, "Conscious" Software Research Group, Institute of Intelligent Systems, University of Memphis, CMattie, LIDA projects:
an attempt to design and implement an intelligent agent under the framework of Bernard Baars' Global Workspace Theory.

My formulation: 1994, 1995 and: Brain-inspired conscious computing architecture. Journal of Mind and Behavior, Vol. 26(1-2), 1-22, 2005

Owen Holland, University of Essex: consciousness via increasingly intelligent behavior, robots with internal models, development of complex control systems, looking for “signs of consciousness”, 0.5 M£ grant.

Pentti Haikonen (Nokia, Helsinki) + consortium led by VTT (Finland)
The cognitive approach to conscious machines (Imprint Academic 2003).
Simulations + microchips, ASSO-IC FET Proactive STERP, submitted.

Robot development

Nomad, DB, Cog, Kismet, Hal – develop robot mind in the same way as babies' minds, by social interactions.

Cog: saccadic eye movements, sound localization, motor coordination, balance, auditory/visual signal coordination, eye, hand and head movement coordination, face recognition, eye contact, haptic (tactile) object recognition ...
Interesting model of autism!

Kismet: sociable humanoid with emotional responses, that seems to be alive and aware.

DB: learning from demonstration, dance, pole balancing, tennis swing, juggling ...
complex eye movements, visuo-motor tasks, such as catching a ball.

Machine consciousness

Owen Holland (Essex Univ), Tom Troscianko and Ian Gilchrist (Bristol Univ) received 0.5 M£ from the EPSR Council (UK) for a project 'Machine consciousness through internal modeling', 2004-2007.

http://www.machineconsciousness.org/

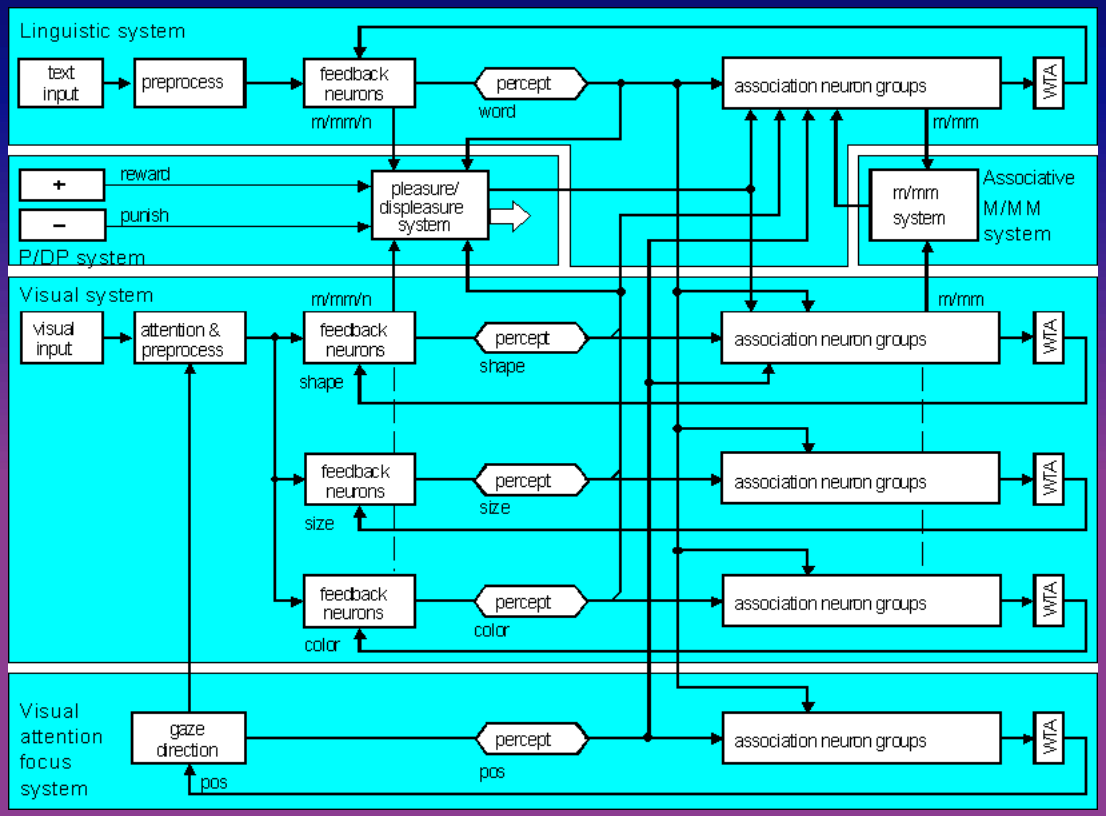
To survive robots will plan actions, build a model of the world and a model of itself - its body, sensors, manipulators, preferences, history …

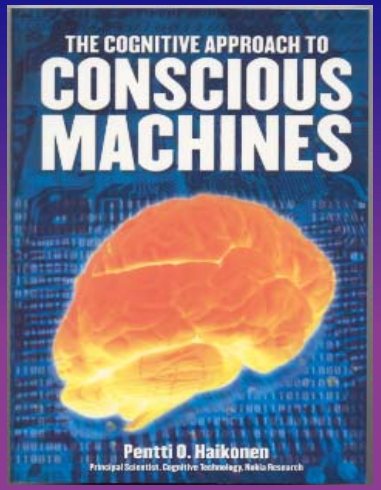
The main focus of interest will be the self-model; its characteristics and internal changes are expected to resemble those of the conscious self in humans, perhaps closely enough to enable some of the robots to be regarded as possessing a form of machine consciousness.
Increasingly complex biologically inspired autonomous mobile robots forced to survive in a series of progressively more difficult environments, and will then study the external and internal behavior of the robots, looking for signs and characteristics of consciousness.

Now Owen is more involved in mechanical design of robots …

Conscious machines: Haikonen

Haikonen has done some simulations based on a rather straightforward design, with neural models feeding the sensory information (with WTA associative memory) into the associative "working memory" circuits.





What do we really want?

Do we really want conscious robots/systems that feel?

- Conscious in what sense? C-like cognitive behavior, in the sense of being aware - yes, robots should have it.
- Phenomenal consciousness with inner life, self, unreliable processes? Is this desired in machines?
- How reliable may machines with phenomenal C be?
- First, can we build them?
  How to build a robot that feels, J.Kevin O'Regan at CogSys 2010 at ETH Zurich on 27/1/2010
- Sensorimotor account of action/perception shows that “hard problem” of consciousness is not such a problem after all.
  Applications of this approach: sensory substitution, as long as the structure is right the signals are correctly interpreted.

- We want machines to be: human like, creative, intuitive, but also following our orders without psychological suffering.

Some failed attempts

Many have proposed the construction of brain-like computers, frequently using special hardware.

- **Connection Machines** from Thinking Machines, Inc. (D. Hills, 1987) was commercially almost successful, but never become massively parallel and the company went bankrupt.

CAM Brain (ATR Kyoto) – failed attempt to evolve the large-scale cellular neural network; based on a bad idea that one can evolve functions without knowing them. Evolutionary algorithms require supervision (fitness function) but it is not clear how to create fitness functions for particular brain structures without knowing their functions first; if we know the function it is easier to program it than evolve.

Attention-Based Artificial Cognitive Control Understanding System (ABACCUS)

Large EU integrated project with 9 participants (never started):

- King's College London (John G. Taylor, coordinator), UK
- Centre for Brain & Cognitive Development, Berkbeck College, University of London, UK
- Cognition and Brain Sciences Unit, Medical Research Council, UK

- Robotics and Embedded Systems, Technical University of Munich, G
- Institute of Neurophysiology and Pathophysiology, Universitätsklinikum Hamburg-Eppendorf, G

- Institute of Computer Science, Foundation for Research and Technology – Hellas, Heraklion, Crete, GR
- National Center for Scientific Research "Demokritos", Athens, GR

- Dip. di Informatica, Sistemistica, Telematica, Universita di Genova, I

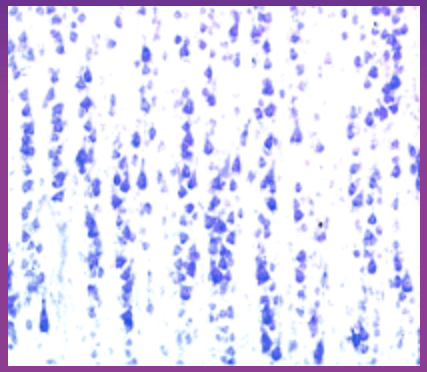
- Dep. of Informatics, Nicholaus Copernicus University, Torun, PL

Other ambitious attempts

Artificial Development (www.ad.com) is building CCortex™, a complete 20-billion neuron simulation of the Human Cortex and peripheral systems, on a cluster of 500 computers - the largest neural network created to date.

Artificial Development plans to deliver a wide range of commercial products based on artificial versions of the human brain that will enhance business relationships globally. Rather unlikely?

The Ersatz Brain Project – James Anderson (Brown University), based on modeling of intermediate level cerebral cortex structures - cortical columns of various sizes (mini ~10^2, plain ~10^4, and hypercolumns ~10^5).



NofN, Network of Networks approximation, 2D BSB network.
http://www.ersatzbrain.org/
not much is happening with this project (2010).

Darwin/Nomad robots

G. Edelman (Neurosciences Institute) & collaborators, created a series of Darwin automata, brain-based devices, “physical devices whose behavior is controlled by a simulated nervous system”.

(i) The device must engage in a behavioral task.

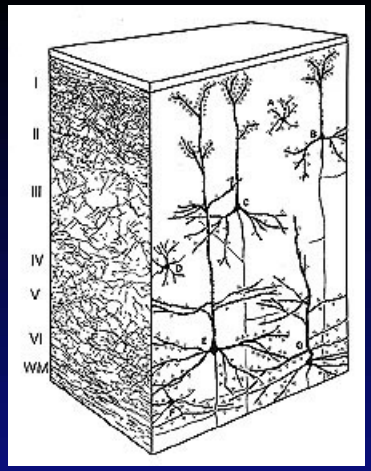
(ii) The device’s behavior must be controlled by a simulated nervous system having a design that reflects the brain’s architecture and dynamics.

(iii) The device’s behavior is modified by a reward or value system that signals the salience of environmental cues to its nervous system.

(iv) The device must be situated in the real world.

Darwin VII consists of: a mobile base equipped with a CCD camera and IR sensor for vision, microphones for hearing, conductivity sensors for taste, and effectors for movement of its base, of its head, and of a gripping manipulator having one degree-of-freedom; 53K mean firing +phase neurons, 1.7 M synapses, 28 brain areas.

Blue Brain



Models at different level of complexity:
http://bluebrainproject.epfl.ch

1. The Blue Synapse: A molecular level model of a single synapse.
2. The Blue Neuron: A molecular level model of a single neuron.
3. The Blue Column: A cellular level model of the Neocortical column with 10K neurons, later 50K, 100M connections (completed 2008).
4. The Blue Neocortex: A simplified Blue Column will be duplicated to produce Neocortical regions and eventually and entire Neocortex.
5. The Blue Brain Project will also build models of other Cortical and Subcortical models of the brain, and sensory + motor organs.

Blue Gene simulates~100M minimal compartment neurons or 10-50'000 multi-compartmental neurons, with 10^3-10^4 x more synapses.
Next generation BG will simulate >10^9 neurons with significant complexity.

Great expectations but not new papers since 2008 ?

Ccortex & Brain Corp

Artificial Development (www.ad.com) is building CCortex™, a complete 20G neuron 20T connection simulation of the Human Cortex and peripheral systems, on a cluster of 500 computers - the largest neural network created to date.

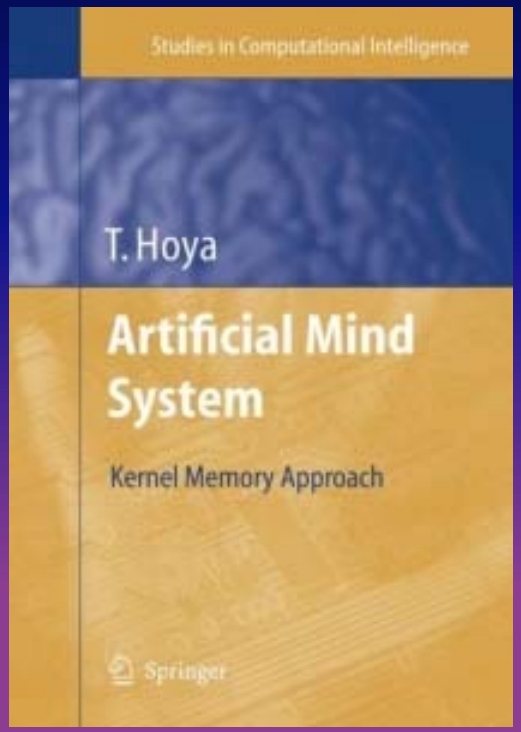
Artificial Development plans to deliver a wide range of commercial products based on artificial versions of the human brain that will enhance business relationships globally.

The project made bold statements in 2007, opened a lab in Kochi, Kerala, India, to "uncover relevant information on the functioning on the human brain, and help model and interpret the data." Died in 2008?

The company is run by Marcos Guillen, who made money as ISP in Spain but has no experience in neuroscience or simulations.

More serious: **Brain Corporation** http://braincorporation.com/ simulation of visual system and motor control using spiking neuron system. Eugene M. Izhikevich is CEO, already made whole 1 sec whole brain simulation.

Artificial Mind System (AMS)
Kernel Memory Approach



Series: **Studies in Computational Intelligence (SCI)**, Vol. 1 (270p)

Springer-Verlag: Heidelberg
Aug. 2005

available from:
http://www.springeronline.com/

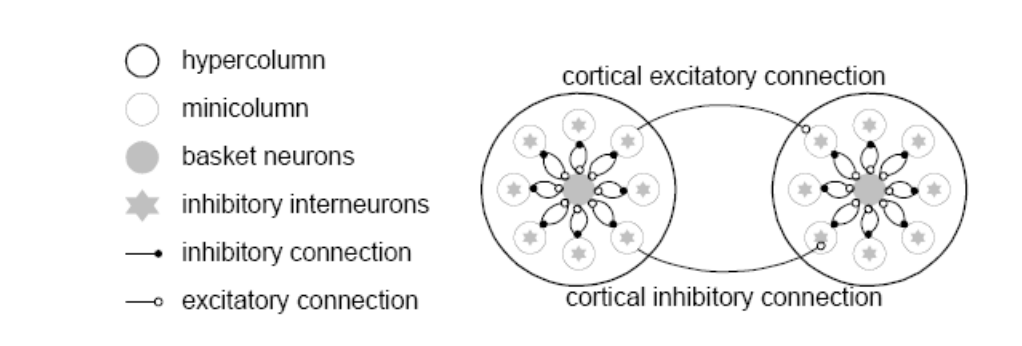
by **Tetsuya Hoya**
BSI-RIKEN, Japan
Lab. Advanced Brain Signal Processing

Hybrid approach, based on modularity of mind.

Bayesian Confidence Propagating NN.

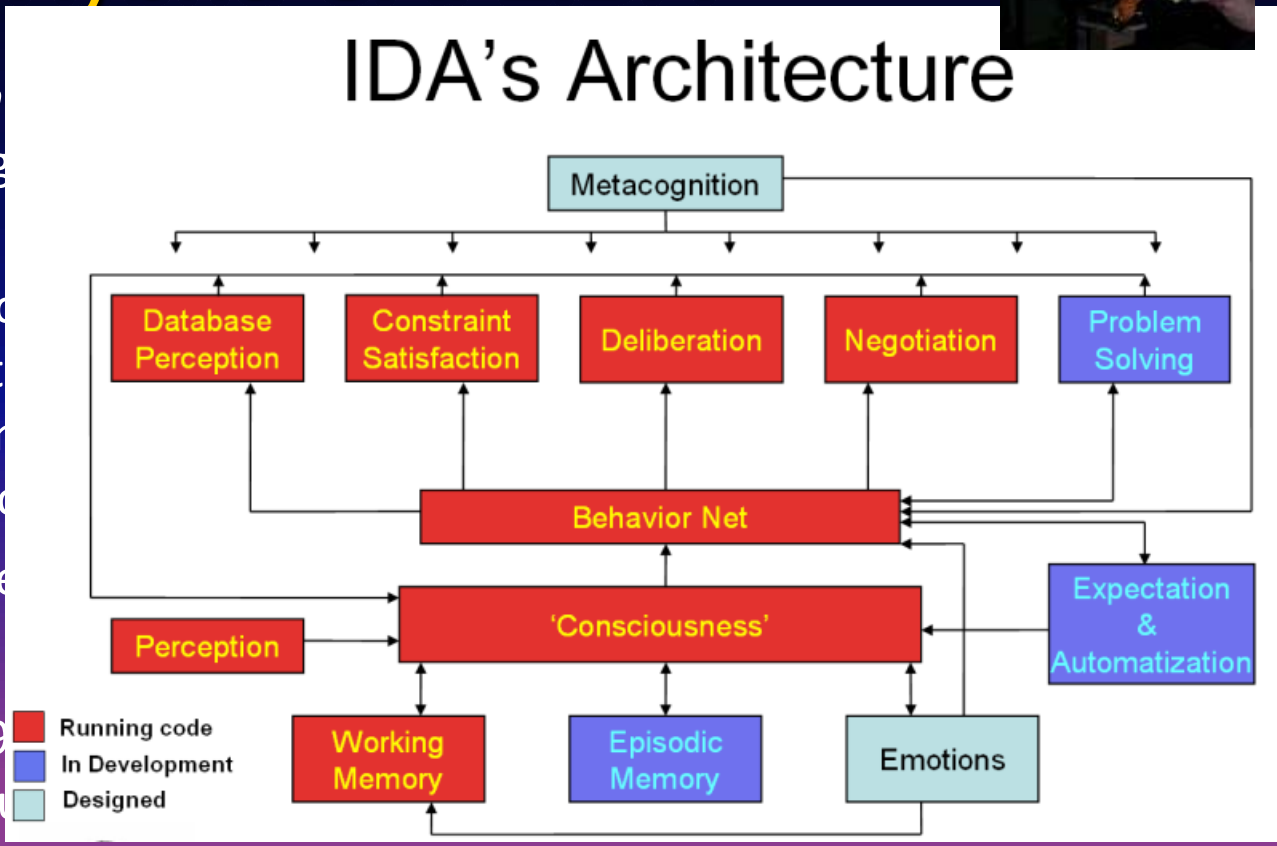
Johansson/Lansner ideas:

- Assumption: functional principles of cortex reside on a much higher level of abstraction than that of the single neuron i.e. closer to abstractions like ANN and connectionist models.
- Target: artificial brain, compact, low-power, multi-network NN.
- Mapping of cortical structure onto the BCPNN, an attractor network.
- Implementation of BCPNN based on hyper columnar modules.

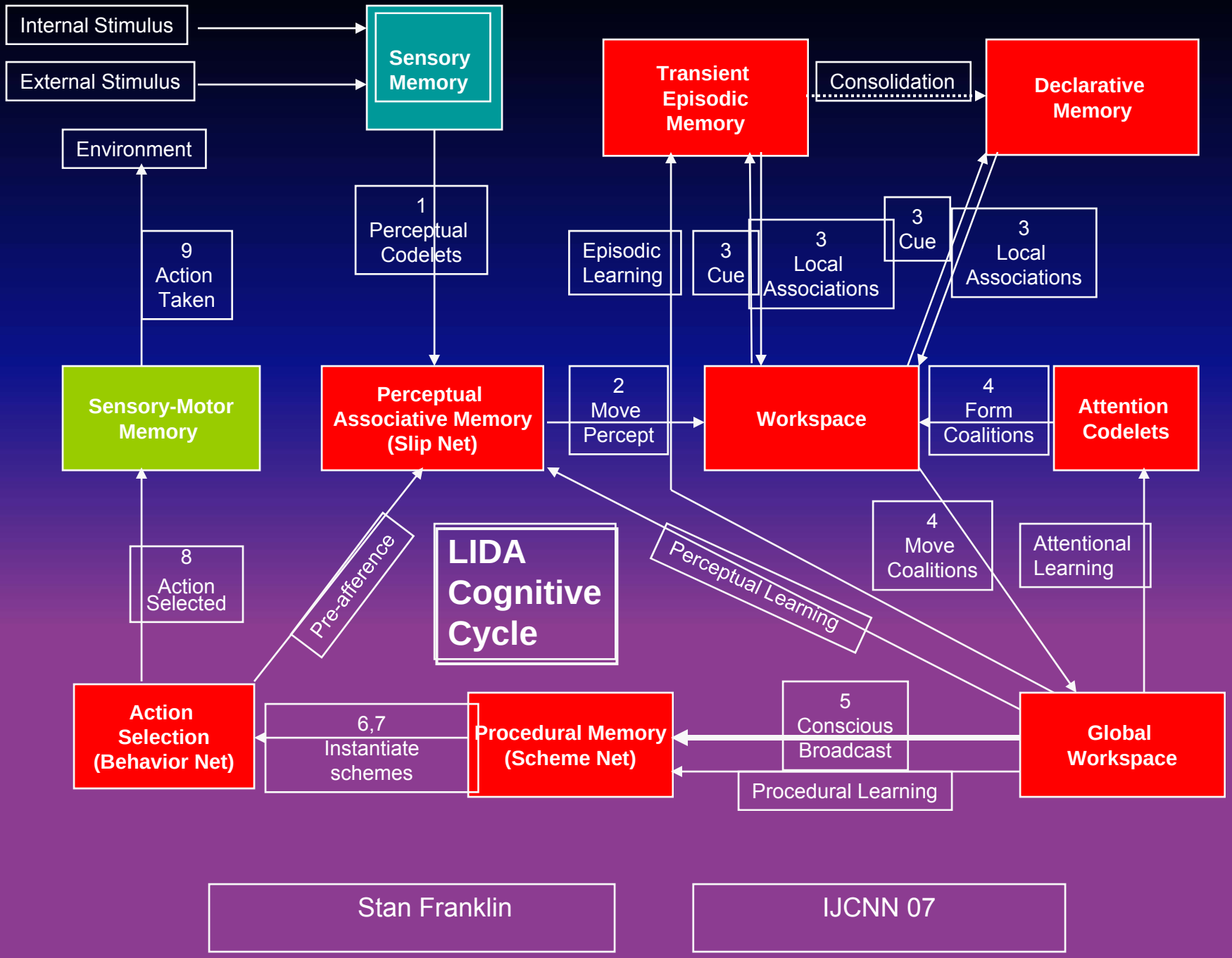


- Hypercolumn needs $5 \cdot 10^9$ ops, with about $2 \cdot 10^6$ hypercolumns in human cortex, giving about 10^{16} ops.
- No detailed structure proposed.

Hybrid CA: LIDA



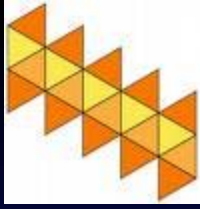
- **LIDA** (The *Learning In*
  framework for intellig
- LIDA: partly symbolic
  modules for perceptic
  episodic memory, act
  constraint satisfactior
  metacognition, and c
- Cooperation of codele
- Perceptual, episodic,
- **DUAL** (B. Kokinov 199
  multi-agent architectu
  micro-agents for memory and processing, agents form coalitions with emergent dynamics, at macrolevel psychological interpretations may be used to describe model properties.
- Micro-frames used for symbolic representation of facts, relevance in a particular context is represented by network connections/activations.


Internal Stimulus
External Stimulus
Sensory Memory
Environment
1 Perceptual Codelets
9 Action Taken
Transient Episodic Memory
Consolidation
Declarative Memory
Episodic Learning
3 Cue
3 Local Associations
3 Cue
3 Local Associations
Sensory-Motor Memory
Perceptual Associative Memory (Slip Net)
2 Move Percept
Workspace
4 Form Coalitions
Attention Codelets
8 Action Selected
Pre-afference
LIDA Cognitive Cycle
Perceptual Learning
4 Move Coalitions
Attentional Learning
Action Selection (Behavior Net)
6,7 Instantiate schemes
Procedural Memory (Scheme Net)
5 Conscious Broadcast
Procedural Learning
Global Workspace
Stan Franklin
IJCNN 07

Hybrid CA: DUAL

- **DUAL** (B. Kokinov, since 1994), inspired by Minsky's "Society of Mind".
- Good example of hybrid, multi-agent architecture.
- Dynamic emergent computations: interacting micro-agents for memory and processing, agents form coalitions with emergent dynamics.
- At macrolevel psychological interpretations may be used to describe model properties.
- Micro-frames used for symbolic representation of facts.
- Relevance in a particular context is represented by network connections/activations.
- Used in a model of reasoning and psychophysics, model of active vision, attention, eye movement simulation.
- Combined with AMBR, Associative Memory-Based Reasoning, analogy making and reasoning.
- Scaling? No larger scale attempts so far.

Hybrid CA: Polyscheme



- **Polyscheme** (N.L. Cassimatis, 2002) integrates multiple methods of representation, reasoning and inference schemes in problem solving. Polyscheme "specialist" models some aspects of the world.
- Scripts, frames, logical propositions, neural networks and constraint graphs represent knowledge, interacting & learning from other specialists; attention is guided by a reflective specialist, focus schemes implement inferences via script matching, backtracking search, reason maintenance, stochastic simulation and counterfactual reasoning.
- High-order reasoning is guided by policies for focusing attention. Operations handled by specialists include forward inference, subgoaling, grounding, with different representations but same focus, may integrate lower-level perceptual and motor processes.
- Both for abstract and common sense physical reasoning in robots.
- Used to model infant reasoning including object identity, events, causality, spatial relations. Meta-learning, different approaches to problem solving.
- No ambitious larger-scale applications yet.

Hybrid CA: 4CAPS

- **4CAPS** (M.A. Just 1992), designed for complex tasks, language comprehension, problem solving or spatial reasoning.
- Operating principle: "Thinking is the product of the concurrent activity of multiple centers that collaborate in a large scale cortical network".
- Used to model human behavioral data (response times and error rates) for analogical problem solving, human–computer interaction, problem solving, discourse comprehension and other complex tasks solved by normal and mentally impaired people.
- Has number of centers (corresponding to particular brain areas) that have different processing styles; ex. Wernicke's area is constructing and selectively accessing structured sequential & hierarchical reps. Each center can perform/be a part of **multiple cognitive functions**, but has a limited computational capacity constraining its activity. Functions are assigned to centers depending on the resource availability, therefore the topology of the whole large-scale network is not fixed.
- Activity of 4CAPS modules correlates with fMRI and other neuroimaging data.
- M. Anderson: neural reuse, important concept (BBS, forthcoming 2010).

Hybrid CA: ACT-R

- **ACT-R** (*Adaptive Components of Thought-Rational*) (Anderson, >20 y), aims at simulations of full range of human cognitive tasks.
- Perceptual-motor modules, memory modules, pattern matcher.
- Symbolic-connectionist structures for declarative memory (DM), chunks for facts; procedural memory (PM), production rules.
  Buffers - WM for inter-module communications and pattern matcher searching for production that matches the present state of buffers.
- Top-down learning approach, sub-symbolic parameters of most useful chunks or productions are tuned using Bayesian approach.
- Rough mapping of ACT-R architecture on the brain structures.
- Used in a large number of psychological studies, intelligent tutoring systems, no ambitious applications to problem solving and reasoning.

- SAIL: combining symbolic ACT-R with O'Reilley's Emergent architecture. Combines advantages of both approaches, but will it scale up?

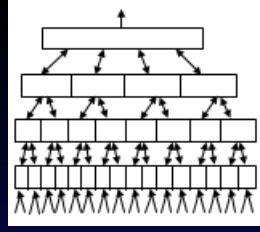
Hybrid CA: others 2

- **Shruti** (Shastri 1993), biologically-inspired model of human reflexive inference, represents in connectionist architecture relations, types, entities and causal rules using focal-clusters. These clusters encode universal/existential quantification, degree of belief, and the query status. The synchronous firing of nodes represents dynamic binding, allowing for representations of complex knowledge and inferences.
  Has great potential, but development is slow .
- **The Novamente AI Engine** (B. Goertzel, 1993), *psynet* model and "patternist philosophy of mind": self-organizing goal-oriented interactions between patterns are responsible for mental states.
- Emergent properties of network activations lead to hierarchical and relational (heterarchical) pattern organization.
- Probabilistic term logic (PTL), and the Bayesian Optimization Algorithm (BOA) algorithms are used for flexible inference.
- Actions, perceptions, internal states represented by tree-like structures.
- New architecture, scaling properties are not yet known.

Emergent CA: others

- **NOMAD** (Neurally Organized Mobile Adaptive Device) (Edelman >20y) based on "neural Darwinism" theory, emergent architectures for pattern recognition task in real time. Large (~10^5 neurons with ~10^7 synapses) simulated nervous system, development through behavioral tasks, value systems based on reward mechanisms in adaptation and learning, importance of self-generated movement in development of perception, the role of hippocampus in spatial navigation and episodic memory, invariant visual object recognition, binding of visual features by neural synchrony, concurrent, real-time control. Higher-level cognition?
- **Cortronics** (Hecht-Nielsen 2006), thalamocortical brain functions.
- Lexicons based on localist cortical patches with reciprocal connections create symbols, with some neurons in patches overlapping.
- Items of knowledge = linked symbols, with learning and information retrieval via confabulation, a competitive activation of symbols.
- Confabulation is involved in anticipation, imagination and creativity, on a shorter time scale than reasoning processes.

Emergent CA: directions



- The **NuPIC** (*Numenta Platform for Intelligent Computing*) (J. Hawking 2004), Hierarchical Temporal Memory (HTM) technology, each node implementing learning and memory functions. Specific connectivity between layers leads to invariant object representation. Stresses temporal aspects of perception, memory for sequences, anticipation.
- Autonomous mental development (J. Weng, ~10 y).
- M.P. Shanahan, internal simulation with a global workspace (2006) weightless neural network, control of simulated robot, very simple.
- COLAMN (M. Denham, 2006), and Grossberg "laminar computing".
- E. Korner & G. Matsumoto: CA controls constraints used to select a proper algorithm from existing repertoire to solve a specific problem.
- DARPA Biologically-Inspired Cognitive Architectures (BICA) program (2006), several consortia formed, for example:
- "TOSCA: Comprehensive brain-based model of human mind", Michigan, MIT Media Lab, AlgoTek, Dartmouth, Johns Hopkins, Harvard and Rutgers.

Where will it go?

- Many architectures, some developed over ~ 30 y, others new.
- Used in very few real-world applications.
- Grand challenges + smaller steps that lead to human and super-human levels of competence should be formulated to focus the research.
- Extend small demonstrations in which a cognitive system reasons in a trivial domain to results that may be of interest to experts, or acting as an assistant to human expert.

- What type of intelligence do we want?

  H. Gardner (1993), at least seven kinds of intelligence:
  logical-mathematical, linguistic, spatial, musical, bodily-kinesthetic, interpersonal and intrapersonal intelligence, perhaps extended by emotional intelligence and a few others.

- To some degree they are independent!
  Perhaps AGI does not have to be very general ... just sufficiently broad to achieve human-level competence in some areas and lower in others.

Trends

- Hybrid architectures dominate, but biological inspirations are very important, expect domination of BICA architectures.
- Focus is mainly on the role of thalamo-cortical and limbic systems, identified with cognitive and emotional aspects.
- Several key brain-inspired features should be preserved in all BICA:

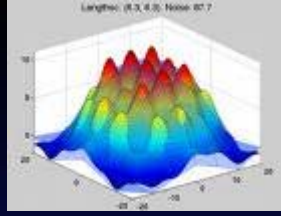
  hierarchical organization of information processing at all levels;

  specific spatial localization of functions, flexible use of resources, time scales; attention; role of different types of memory, imagination, intuition, creativity.

Missing so far:

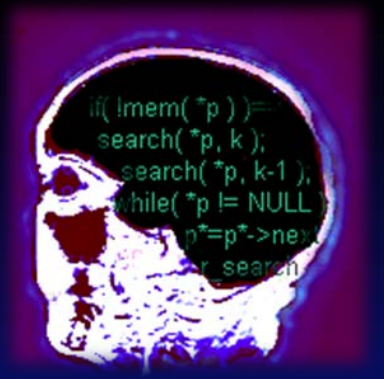
- Specific role of left and right hemisphere, brain stem etc.
- Many specific functions, ex. various aspects of self, fear vs. apprehension, processed by different amygdala structures.
- Regulatory role of the brain stem which may provide overall meta-control selecting different types of behavior is completely neglected.

BICA as approximation



- Significant progress has been made in drawing inspirations from neuroscience in analysis of perception, less in higher cognition.
- For example, neurocognitive approach to linguistics has been used only to analyze linguistic phenomena, but has no influence on NLP.
- "Brain pattern calculus" to approximate spreading neural activation in higher cognitive functions is urgently needed! How to do it?
  Neural template matching? Network-constrained quasi-stationary waves describing global brain states $\Psi(w,Cont)$?
- Practical algorithms to discover "pathways of the brain" has been introduced recently (Duch et al, in print) to approximate symbolic knowledge & associations stored in human brain.
- Efforts to build concept descriptions from electronic dictionaries, ontologies, encyclopedias, results of collaborative projects and active searches in unstructured sources are under way.
- Architecture that uses large semantic memory to control an avatar playing word games has been demonstrated.

Neurocognitive informatics



Use inspirations from the brain, derive practical algorithms!

My own attempts - see my webpage, Google: W. Duch

1. Mind as a shadow of neurodynamics: geometrical model of mind processes, psychological spaces providing inner perspective as an approximation to neurodynamics.

2. Intuition: learning from partial observations, solving problems without explicit reasoning (and combinatorial complexity) in an intuitive way.

3. Neurocognitive linguistics: how to find neural pathways in the brain.

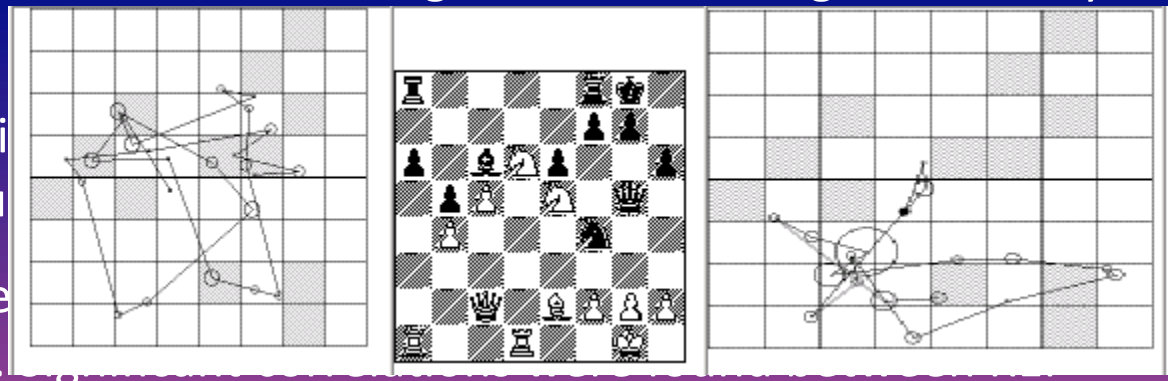
4. Creativity in limited domains + word games, good fields for testing.

Duch W, Intuition, Insight, Imagination and Creativity,
IEEE Computational Intelligence Magazine 2(3), August 2007, pp. 40-52

Intuition

Intuition is a concept difficult to grasp, but commonly believed to play important role in business and other decision making; „knowing without being able to explain how we know”.

Sinclair Ashkanasy (2005): intuition is a „non-sequential information-processing mode, with cognitive & affective elements, resulting in direct knowing without any use of conscious reasoning”.



3 tests measuring intuition: Rati
Type Inventory (MBTI) and Accu

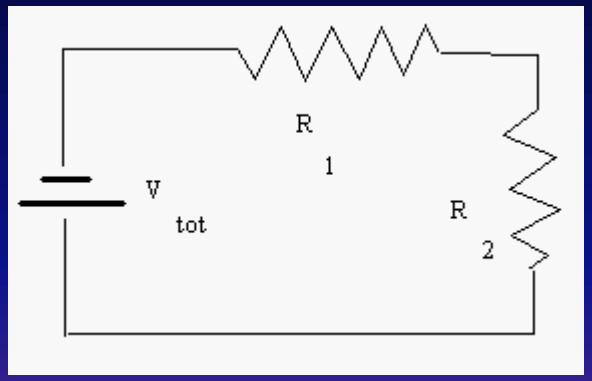
Different intuition measures are
theoretical concept of intuition.
intuition scale and some measures of creativity.

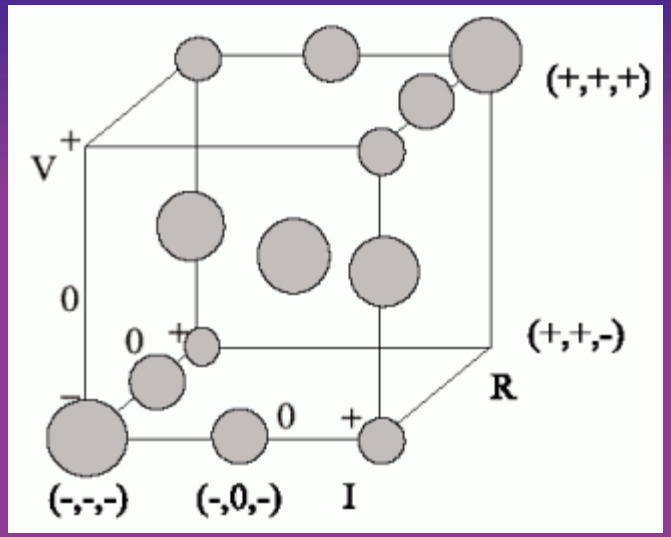
ANNs evaluate intuitively? Yes, although intuition is used also in reasoning.

Intuition in chess has been studied in details (Newell, Simon 1975).

Intuition may result from implicit learning of complex similarity-based evaluation that are difficult to express in symbolic (logical) way.

Intuitive thinking





Question in qualitative physics (PDP book): if R_2 increases, R_1 and V_t are constant, what will happen with current and V_1, V_2 ?

Learning from partial observations:

Ohm's law $V=I\times R$; Kirhoff's $V=V_1+V_2$.

Geometric representation of qualitative facts:

+ increasing, 0 constant, - decreasing.

True (I_-,V_-,R_0), (I_+,V_+,R_0), false (I_+,V_-,R_0).

5 laws: 3 Ohm's 2 Kirhoff's laws.

All laws $A=B+C$, $A=B\times C$, $A^{-1}=B^{-1}+C^{-1}$, have identical geometric interpretation!

13 true, 14 false facts; simple P-space, but complex neurodynamics.

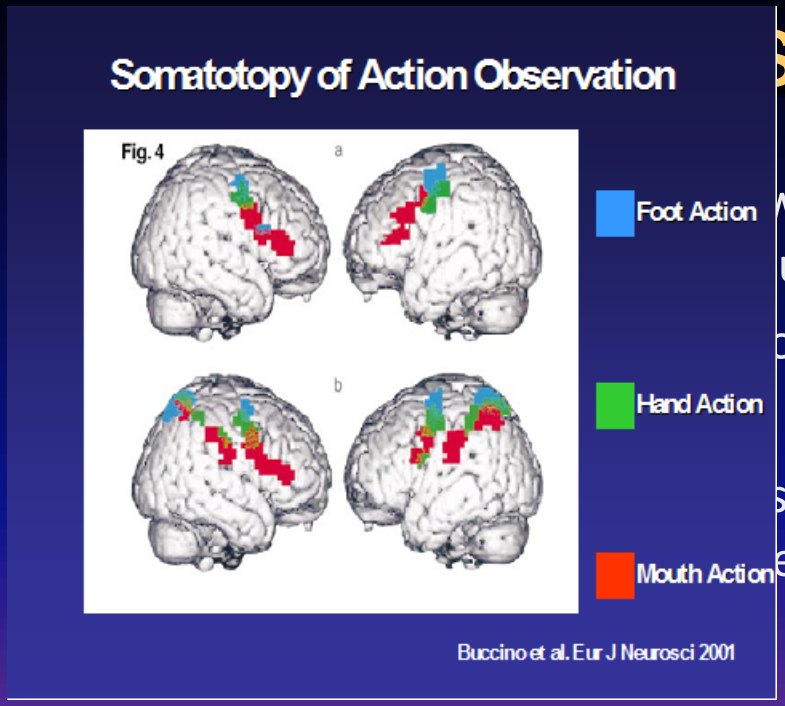
s in the brain

w that most likely categorical,
used, not the acoustic input.

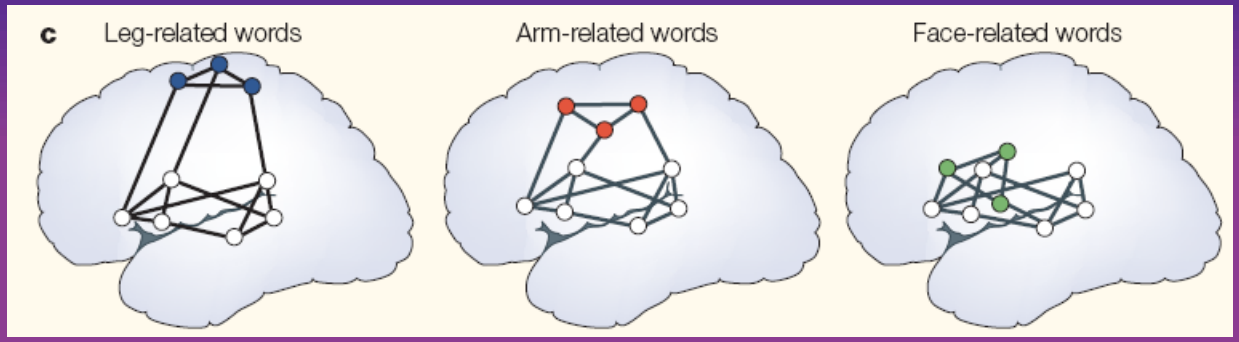
ords => semantic concepts.

semantic by 90 ms (from N200 ERPs).

science of Language. On Brain Circuits of Words
ersity Press.



Action-perception networks inferred from ERP and fMRI

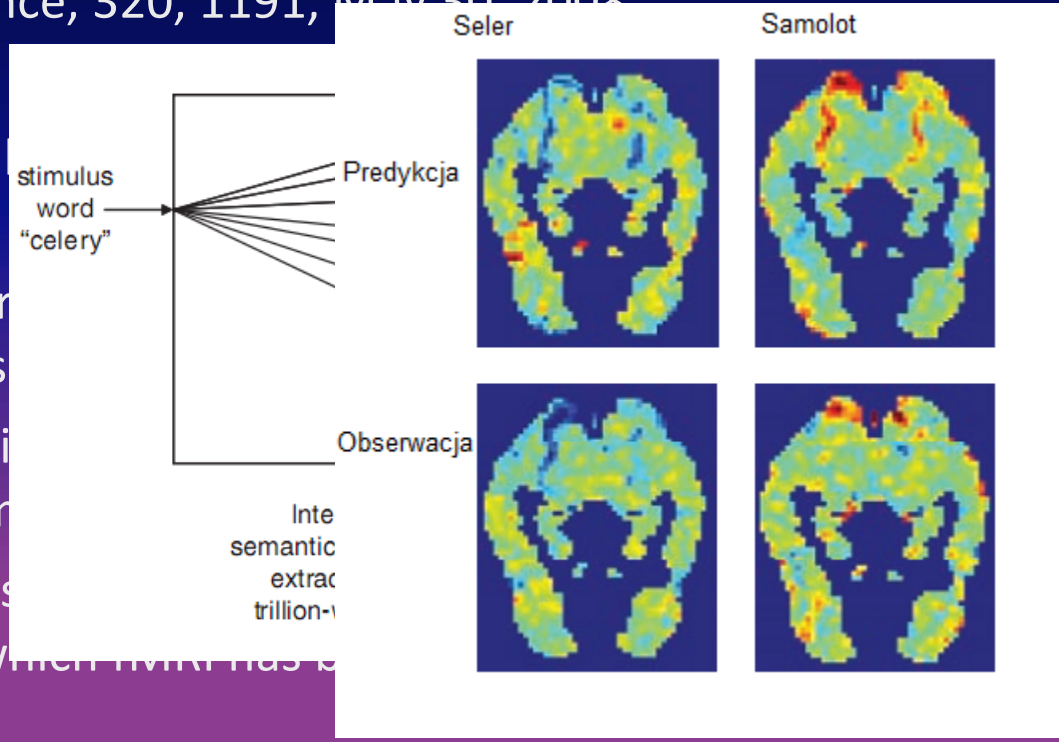


Left hemisphere: precise representations of symbols, including phonological components; right hemisphere? Sees clusters of concepts.

Neuroimaging words

Predicting Human Brain Activity Associated with the Meanings of Nouns," T. M. Mitchell et al, Science, 320, 1191, May 30, 2008

- Clear differences between fMRI
  about different nouns.
- Reading words and seeing the dr
  presumably reflecting semantics
- Although individual variance is si
  of different people, a classifier m
- Model trained on ~10 fMRI scans
  activity for over 100 nouns for w

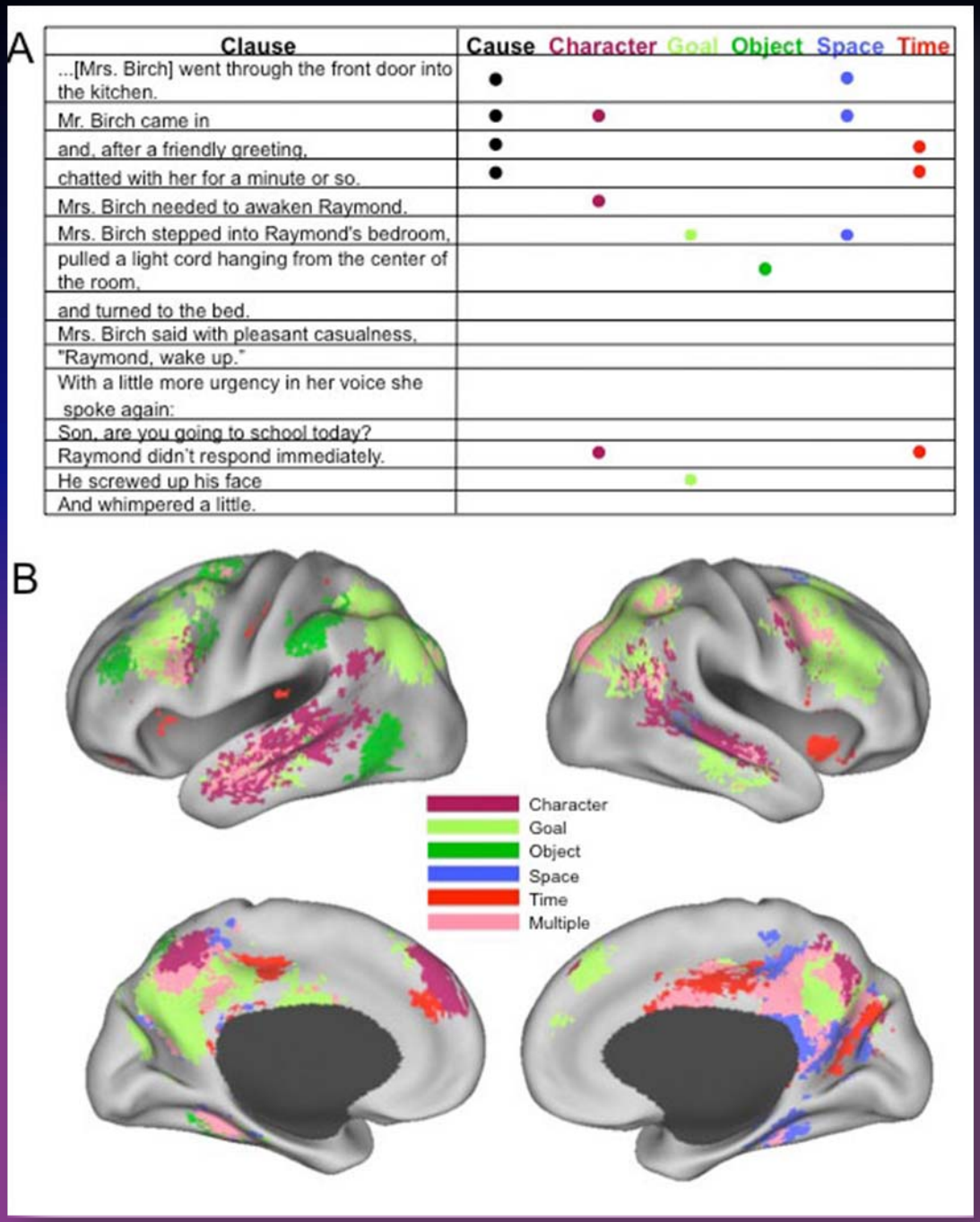


Overlaps between activation of the brain for different words may serve as expansion coefficients for word-activation basis set.

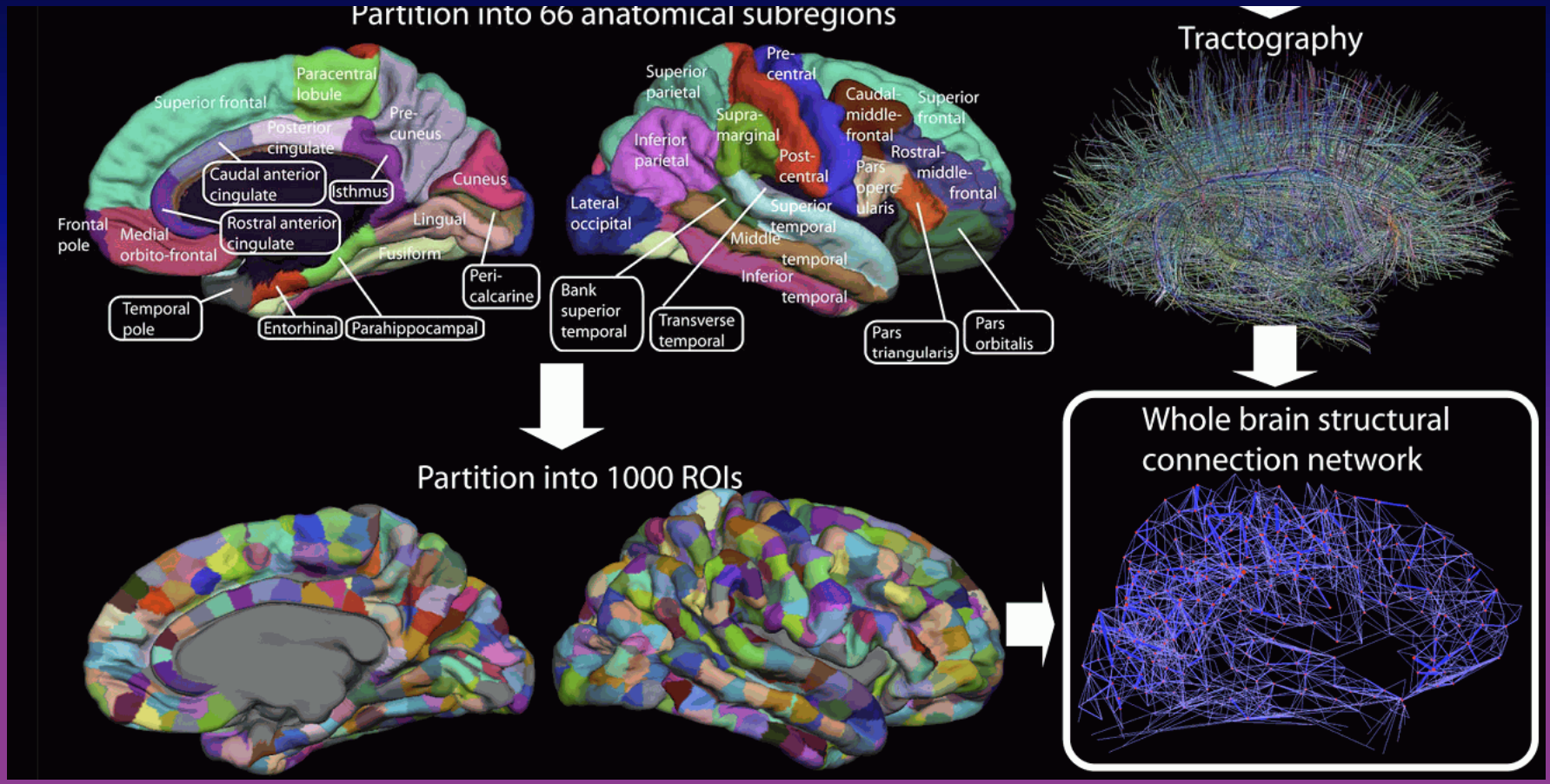
In future: I may know what you'll think before you will know it yourself! Intentions may be known seconds before they become conscious!

Nicole Speer et al. Reading Stories Activates Neural Representations of Visual and Motor Experiences. *Psychological Science* (2010, in print).

Meaning: always slightly different, depending on the context, but still may be clusterized into relatively samll number of distinct meanings.

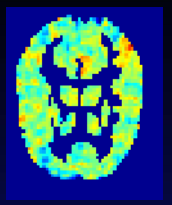


Connectome Project



Brain-based representations of concepts should be possible.

Hidden concepts



- **Language**, symbols in the brain: phonological labels associated with protypes of distributed activations of the brain.

Helps to structure the flow of brain states in the thinking process.

Do we have conscious access to all brain states that influence thinking?

Right hemisphere activations just give us the feeling wrong something here.

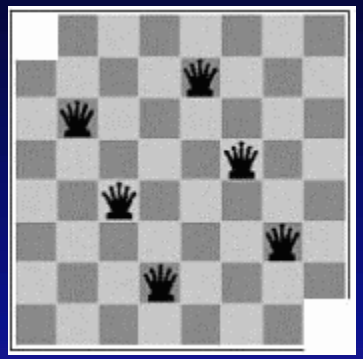
- Right hemisphere is as busy as left – concepts without verbal labels?
- Evidence: insight phenomena, intuitive understanding of grammar, etc.

Can we describe verbally natural categories?

- Yes, if they are rather distinct: see 20 question game.
- Is object description in terms of properties sufficient and necessary?
- Not always. Example: different animals and dog breeds.
- 20Q-game: weak question (seemingly unrelated to the answer) may lead to precise identification! RH may contribute to activation enabling associations

Problems requiring insights

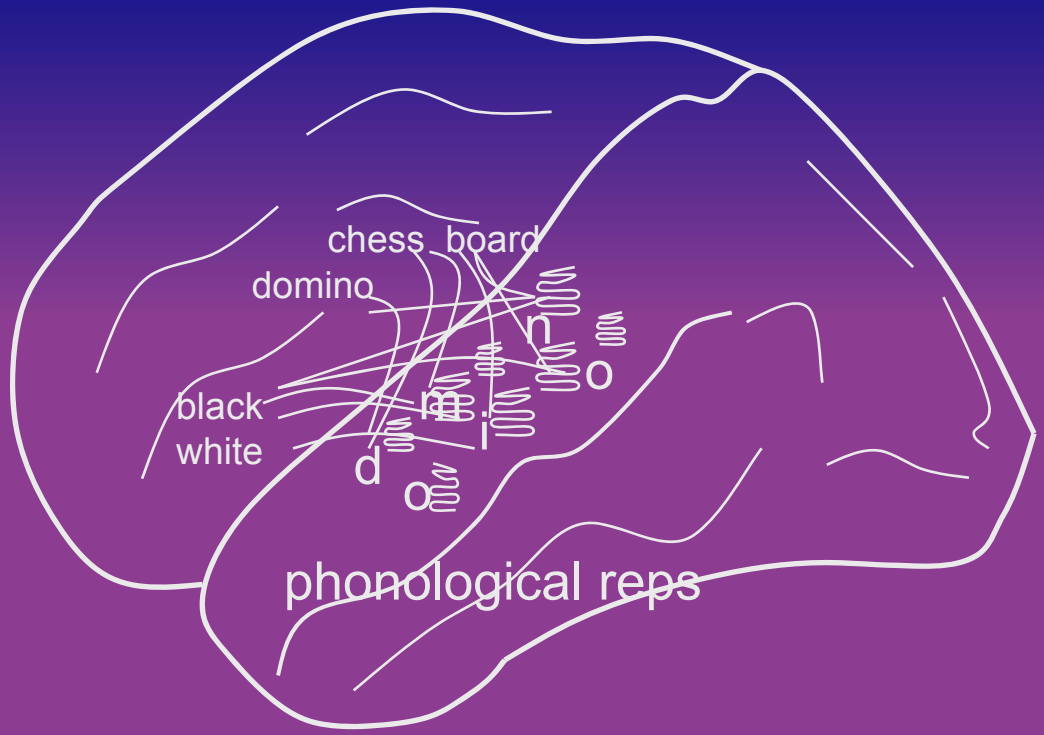
Given 31 dominos ■■ and a chessboard with 2 corners removed, can you cover all board with dominos?



Analytical solution: try all combinations.

Does not work … to many combinations to try.

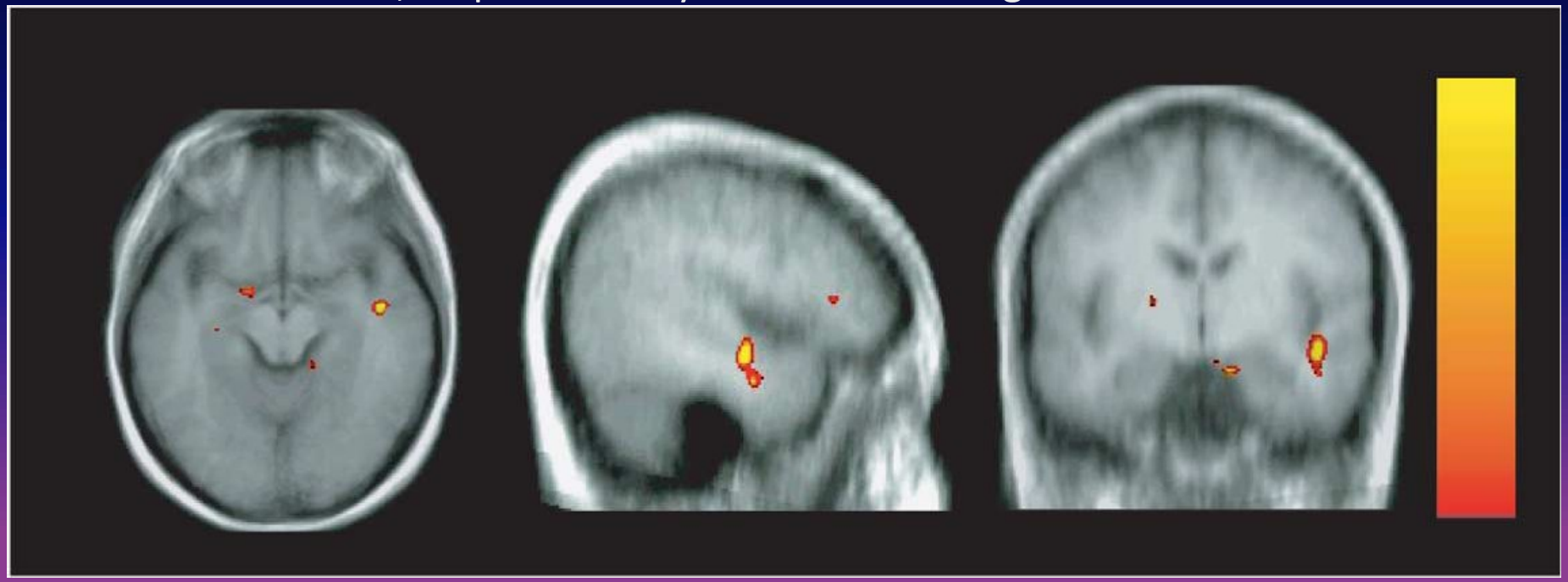
Logical, symbolic approach has little chance to create proper activations in the brain, linking new ideas: otherwise there will be too many associations, making thinking difficult.



Insight <= right hemisphere, meta-level representations without phonological (symbolic) components ... counting?

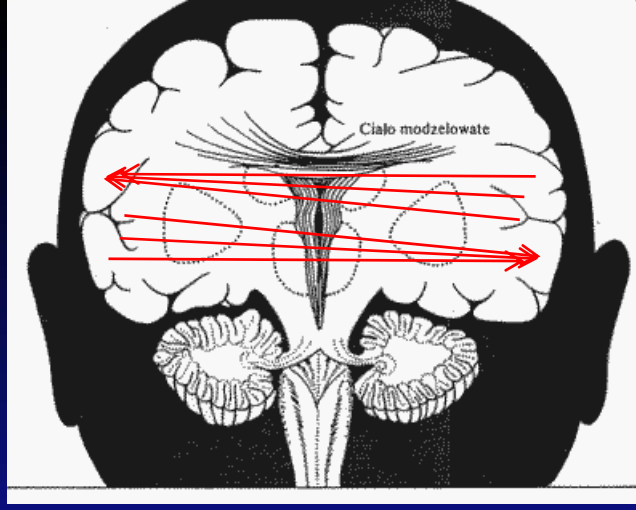
Insights and brains

Activity of the brain while solving problems that required insight and that could be solved in schematic, sequential way has been investigated.



An increased activity of the right hemisphere anterior superior temporal gyrus (RH-aSTG) was observed during initial solving efforts and insights. About 300 ms before insight a burst of gamma activity was observed, interpreted by the authors as „making connections across distantly related information during comprehension ... that allow them to see connections that previously eluded them”.

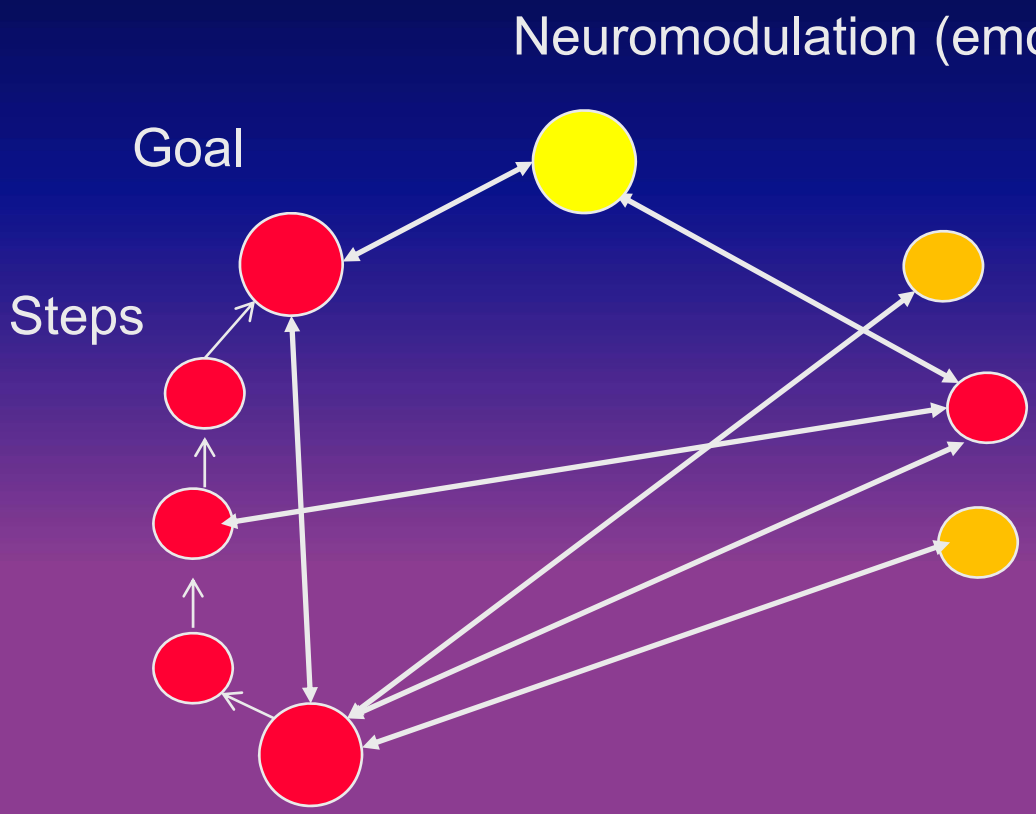
Insight interpreted



What really happens? My interpretation:

- LH-STG represents concepts, S=Start, F=final
- understanding, solving = transition, step by step, from S to F
- if no connection (transition) is found this leads to an impasse;
- RH-STG ‘sees’ LH activity on meta-level, clustering concepts into abstract categories (cosets, or constrained sets);
- connection between S to F is found in RH, leading to a feeling of vague understanding;
- gamma burst increases the activity of LH representations for S, F and intermediate configurations; feeling of imminent solution arises;
- stepwise transition between S and F is found;
- finding solution is rewarded by emotions during Aha! experience; they are necessary to increase plasticity and create permanent links.

Solving problems with insight



Dog breeds

329 breeds in 10 categories:

Sheepdogs and Cattle Dogs; Pinscher and Schnauzer; Spitz and Primitive; Scenthounds; Pointing Dogs; Retrievers, Flushing Dogs and Water Dogs; Companion and Toy Dogs; Sighthounds

Write down properties and try to use them in the 20-question game to recognize the breed … fails!

Visually each category is quite different,
all traditional categorizations are based on behavious and features that are not easy to observe.

- Ontologies do not agree with visual similarity.
- Brains discover it easily => not all brain states have linguistic labels.

Mental models

Kenneth Craik, 1943 book "The Nature of Explanation", G-H Luquet attributed mental models to children in 1927.

P. Johnson-Laird, 1983 book and papers.

Imagination: mental rotation, time ~ angle, about 60^{o}/sec.

Internal models of relations between objects, hypothesized to play a major role in cognition and decision-making.

AI: direct representations are very useful, direct in some aspects only!

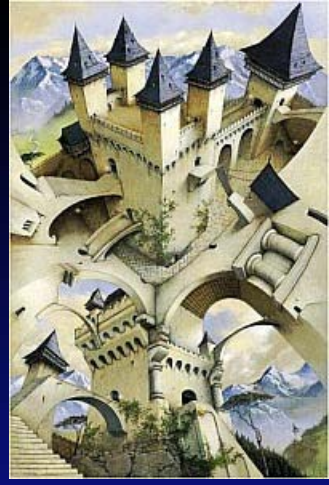
Reasoning: imaging relations, "seeing" mental picture, semantic?

Systematic fallacies: a sort of cognitive illusions.

- If the test is to continue then the turbine must be rotating fast enough to generate emergency electricity.
- The turbine is not rotating fast enough to generate this electricity.
- What, if anything, follows? Chernobyl disaster …

If A=>B; then ~B => ~A, but only about 2/3 students answer correctly..

Mental models summary



The mental model theory is an alternative to the view that deduction depends on formal rules of inference.

1. MM represent explicitly what is true, but not what is false; this may lead naive reasoner into systematic error
2. Large number of complex models => poo
3. Tendency to focus on a few possible moc irrational decisions.

Cognitive illusions are just like visual illusions.

M. Piattelli-Palmarini, Inevitable Illusions: Ho Minds (1996)

R. Pohl, Cognitive Illusions: A Handbook on Fa Judgement and Memory (2005)

Amazing, but mental models theory ignores e

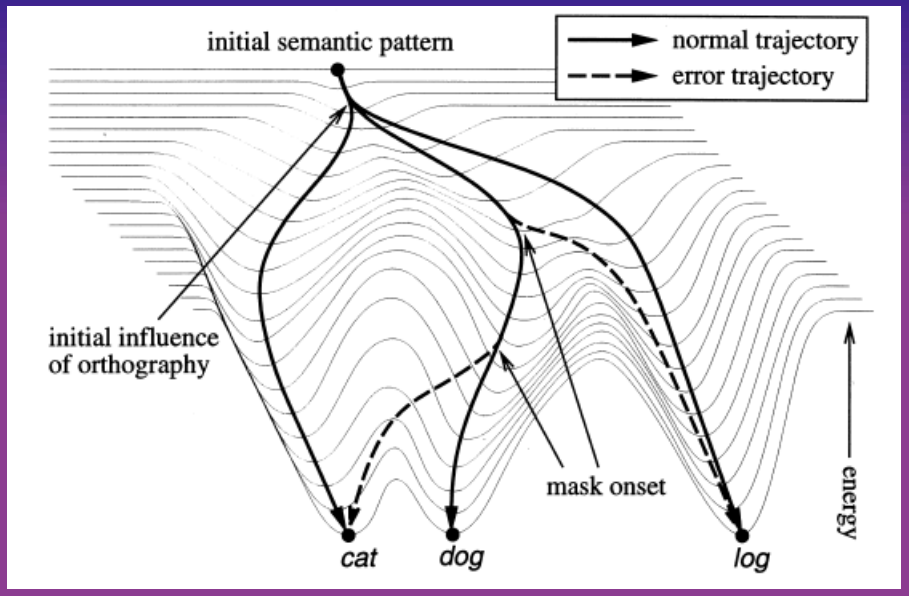
learning in any form! How and why do we reason the way we do?

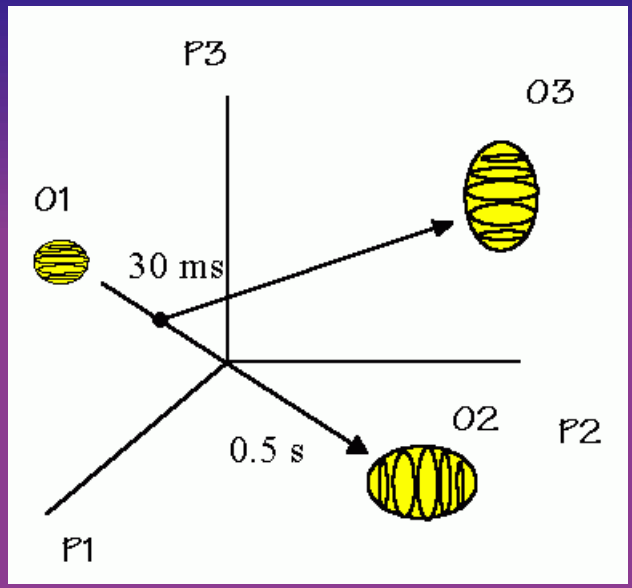
I'm innocent! My brain made me do it!

Energies of trajectories

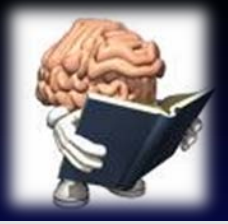
P.McLeod, T. Shallice, D.C. Plaut,
Attractor dynamics in word recognition: converging evidence from errors by normal subjects, dyslexic patients and a connectionist model.
Cognition 74 (2000) 91-113.

New area in psycholinguistics: investigation of dynamical cognition, influence of masking on semantic and phonological errors.



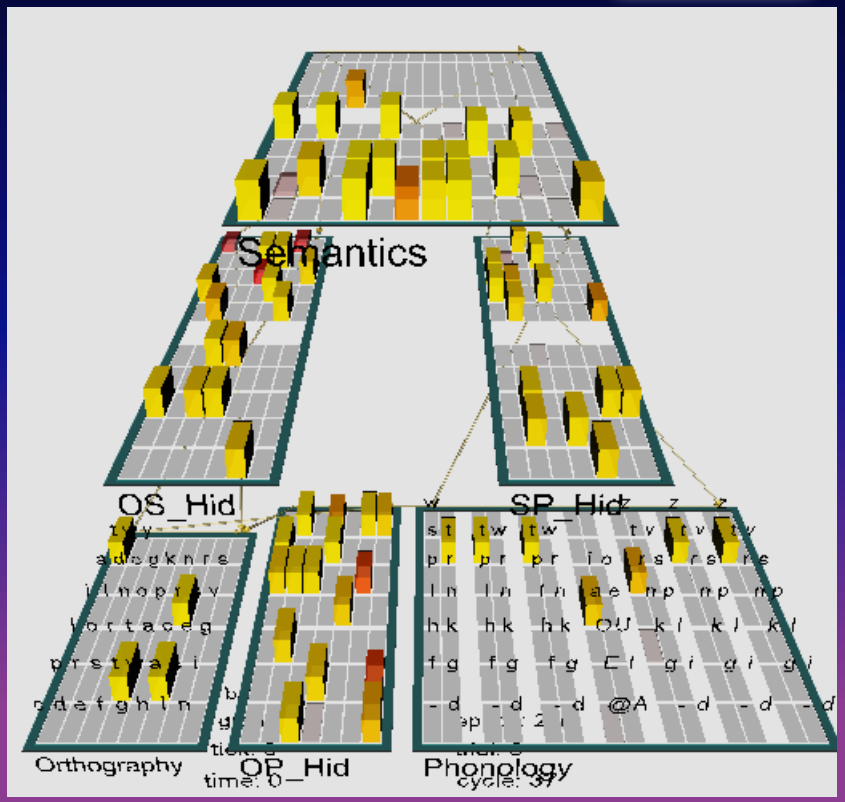


Model of reading & dyslexia



Emergent neural simulator:

Aisa, B., Mingus, B., and O'Reilly, R. The emergent neural modeling system. Neural Networks, 21, 1045-1212, 2008.



3-layer model of reading:

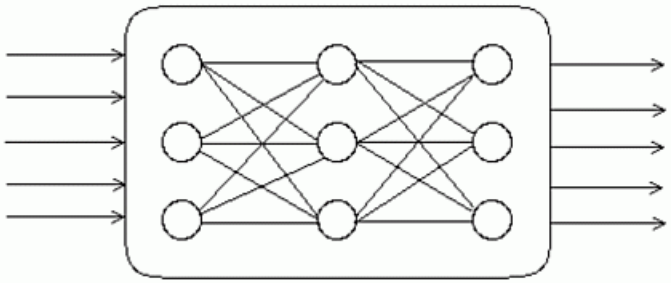
orthography, phonology, semantics, or distribution of activity over 140 microfeatures of concepts.

Hidden layers in between.

Learning: mapping one of the 3 layers to the other two.

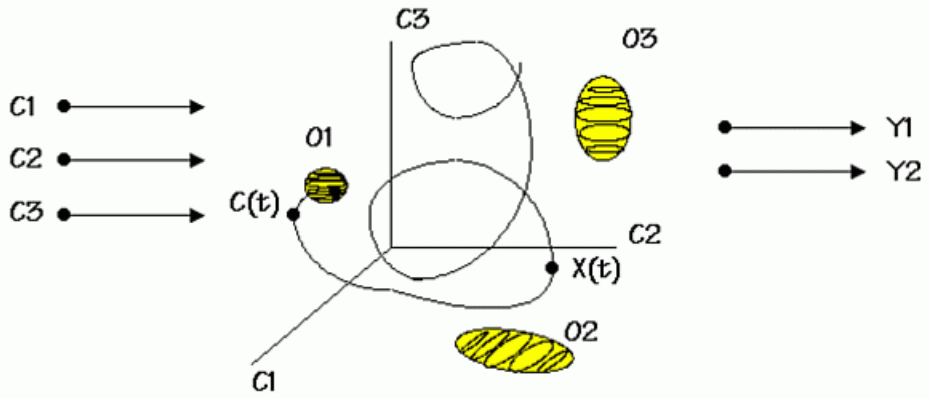
Fluctuations around final configuration = attractors representing concepts.

How to see properties of their basins, their relations?



Neurodynamics

Psychological space



P-spaces

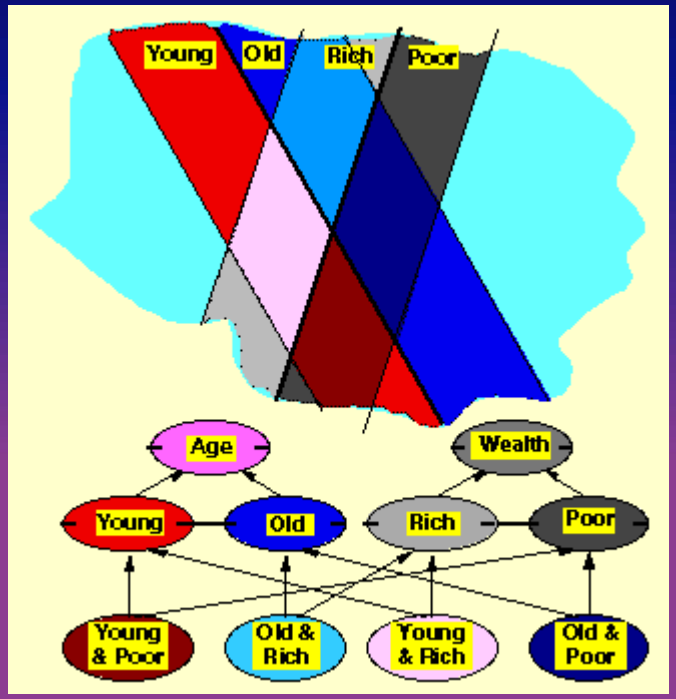
Psychological spaces: how to visualize inner life?

K. Lewin, The conceptual representation and the measurement of psychological forces (1938), cognitive dynamic movement in phenomenological space.

George Kelly (1955):
personal construct psychology (PCP),
geometry of psychological spaces as alternative to logic.

A complete theory of cognition, action, learning and intention.

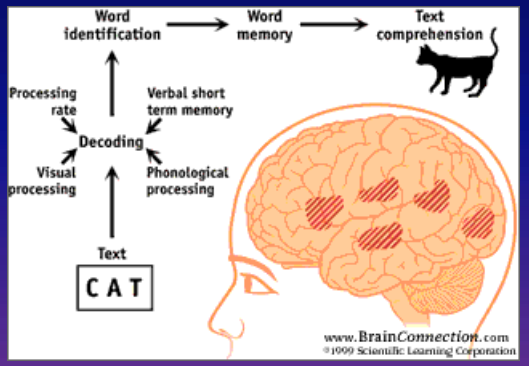
PCP network, society, journal, software …
quite active group.



Many things in philosophy, dynamics, neuroscience and psychology, searching for new ways of understanding cognition, are relevant here.

P-space definition

P-space: region in which we may place and classify elements of our experience, constructed and evolving,
„a space without distance”, divided by dichotomies.



P-spaces should have (Shepard 1957-2001):

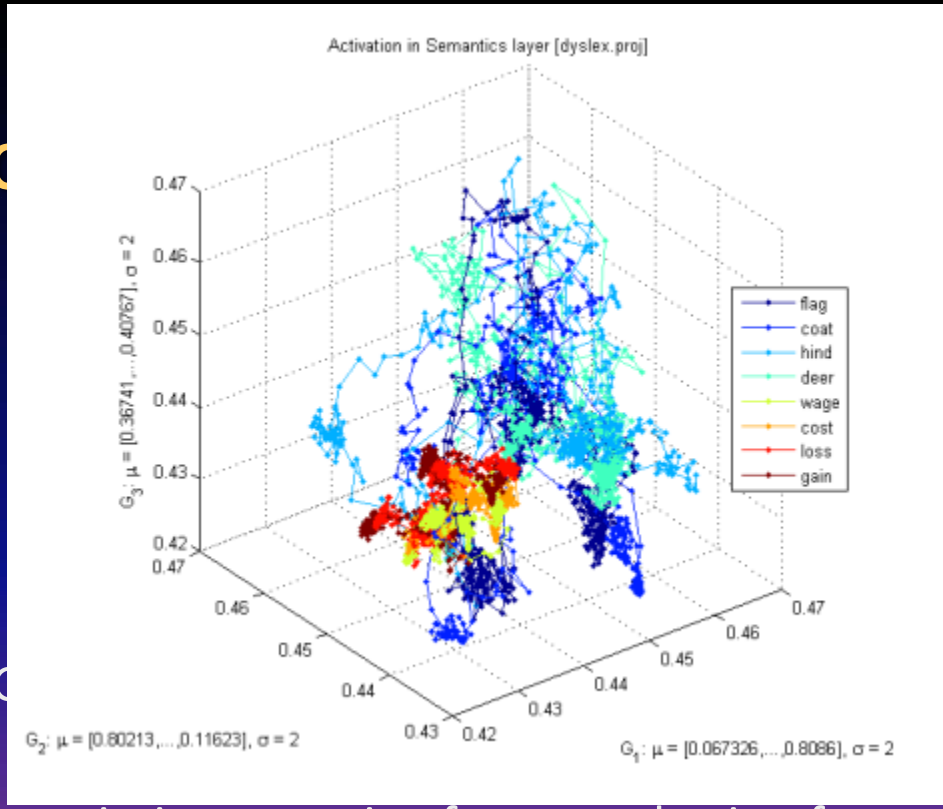
- minimal dimensionality;
- distances that monotonically decrease with increasing similarity.

This may be achieved using multi-dimensional non-metric scaling (MDS), reproducing similarity relations in low-dimensional spaces.

Many Object Recognition and Perceptual Categorization models assume that objects are represented in a multidimensional psychological space; similarity between objects ~ 1/distance in this space.

Can one describe the state of mind in similar way?

Attrac



Attention results from:

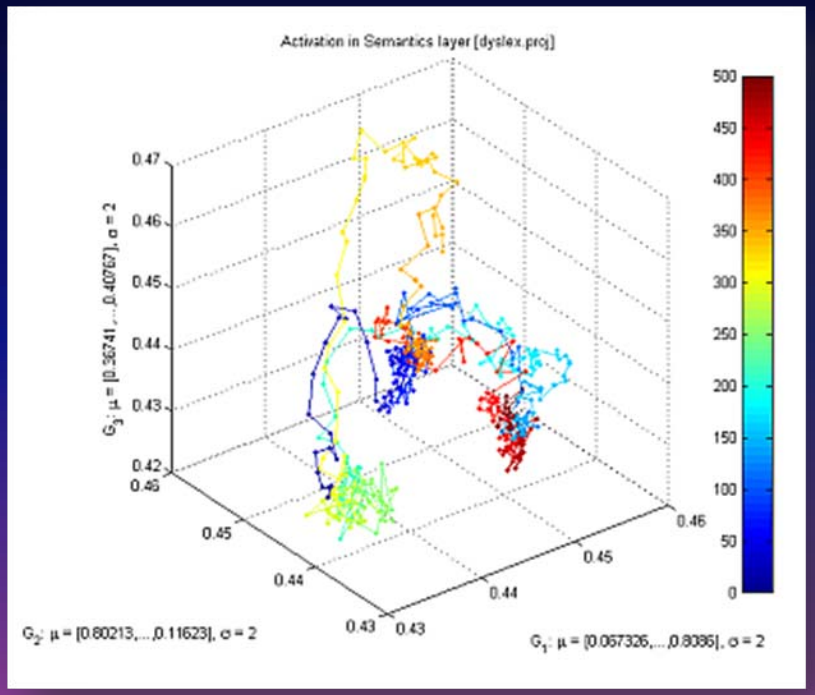
- inhibitory competition,
- bidirectional interactive processing,
- multiple constraint satisfaction.

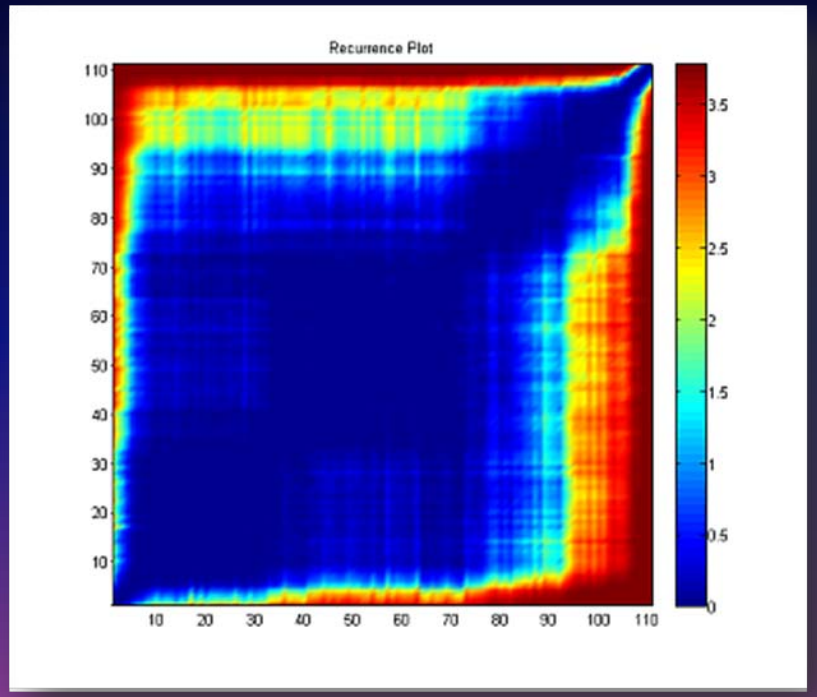
Basins of attractors: input activations {LC

- Normal case: relatively large, easy associations, moving from one basin of attraction to another, exploring the activation space.
- Without accommodation (voltage-dependent K^+ channels): deep, narrow basins, hard to move out of the basin, associations are weak.

Accommodation: basins of attractors shrink and vanish because neurons desynchronize due to the fatigue; this allows other neurons to synchronize, leading to quite unrelated concepts (thoughts).

Recurrence plots

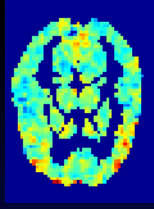




Starting from the word “flag”, with small synaptic noise (var=0.02), the network starts from reaching an attractor and moves to another one (frequently quite distant), creating a “chain of thoughts”.

Same trajectories displayed with recurrence plots, showing roughly 5 larger basins of attractors and some transient points.

Neurocognitive reps.



How to approach modeling of word (concept) *w* representations in the brain? Word $w = (w_f, w_s)$ has

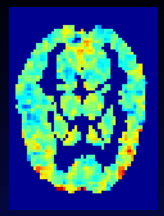
- phonological (+visual) component w_f, word form;
- extended semantic representation w_s, word meaning;
- is always defined in some context *Cont* (enactive approach).

$\Psi(w,Cont,t)$ evolving prob. distribution (pdf) of brain activations.
Hearing or thinking a word *w* , or seeing an object labeled as *w* adds to the overall brain activation in a non-linear way.

How? Spreading activation in neural spaces, maximizing overall self-consistency, mutual activations, meanings that don't fit to current context are automatically inhibited. **Result:** almost continuous variation of this meaning.

This process is rather difficult to approximate using typical knowledge representation techniques, such as connectionist models, semantic networks, frames or probabilistic networks.

Approximate reps.



States $\Psi(w,Cont)$ ⇔ lexicographical meanings:

- clusterize $\Psi(w,Cont)$ for all contexts;
- define prototypes $\Psi(w_k,Cont)$ for different meanings w_k.

A1: use spreading activation in semantic networks to define Ψ.
A2: take a snapshot of activation Ψ in discrete space (vector approach).

Meaning of the word is a result of priming, spreading activation to speech, motor and associative brain areas, creating affordances.

$\Psi(w,Cont)$ ~ quasi-stationary wave, with phonological/visual core activations w_f and variable extended representation w_s selected by *Cont*.

$\Psi(w,Cont)$ state into components, because the semantic representation

E. Schrödinger (1935): best possible knowledge of a whole does not include the best possible knowledge of its parts! Not only in quantum case. Left semantic network *LH* contains w_f coupled with the *RH*.

QM-like formalism is useful for any probability waves.

Semantic => vector reps

Some associations are subjective, some are universal.

How to find the activation pathways in the brain? Try this algorithm:

- Perform text pre-processing steps: stemming, stop-list, spell-checking ...
- Map text to some ontology to discover concepts (ex. UMLS ontology).
- Use relations (Wordnet, ULMS), selecting those types only that help to distinguish between concepts.
- Create first-order cosets (terms + all new terms from included relations), expanding the space – acts like a set of filters that evaluate various aspects of concepts.
- Use feature ranking to reduce dimensionality of the first-order coset space, leave all original features.
- Repeat last two steps iteratively to create second- and higher-order enhanced spaces, first expanding, then shrinking the space.

Result: a set of **X** vectors representing concepts in enhanced spaces, partially including effects of spreading activation.

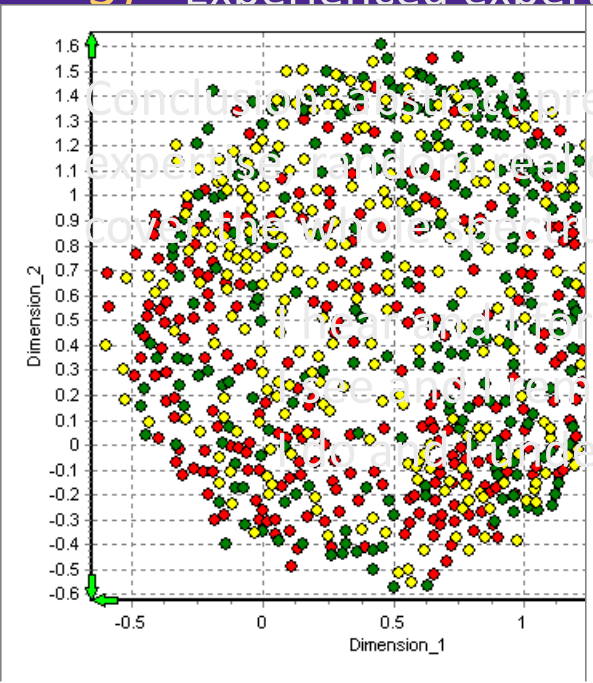
How to become an expert?

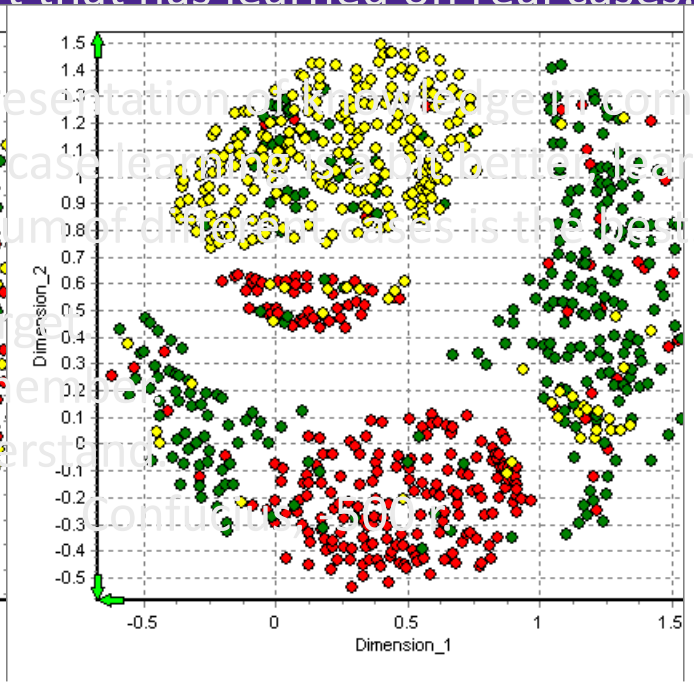
Textbook knowledge in medicine: detailed description of all possibilities.

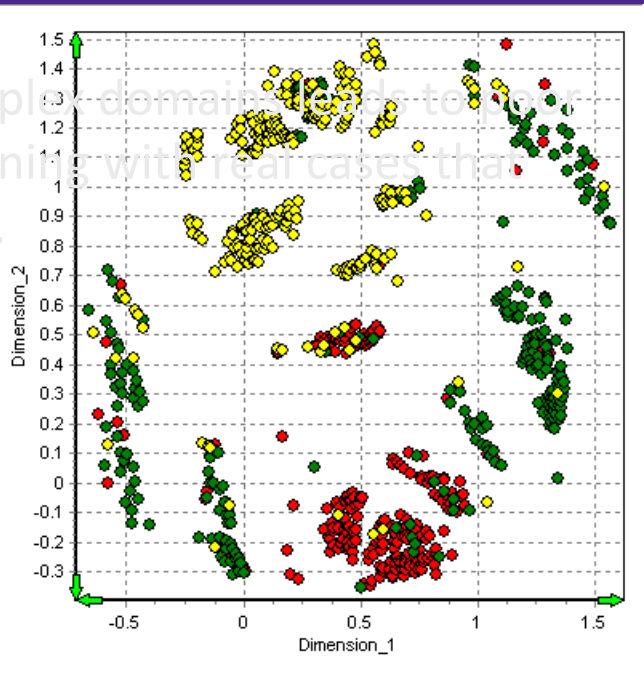
Effect: neural activation flows everywhere and correct diagnosis is impossible.
Correlations between observations forming prototypes are not firmly established.
Expert has correct associations.

Example: 3 diseases, clinical case description, MDS description.

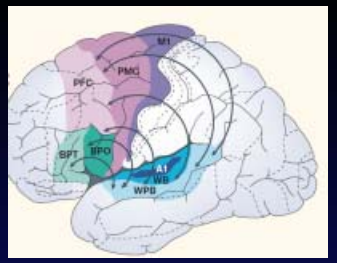
1) System that has been trained on textbook knowledge.
2) Same system that has learned on real cases.
3) Experienced expert that has learned on real cases.

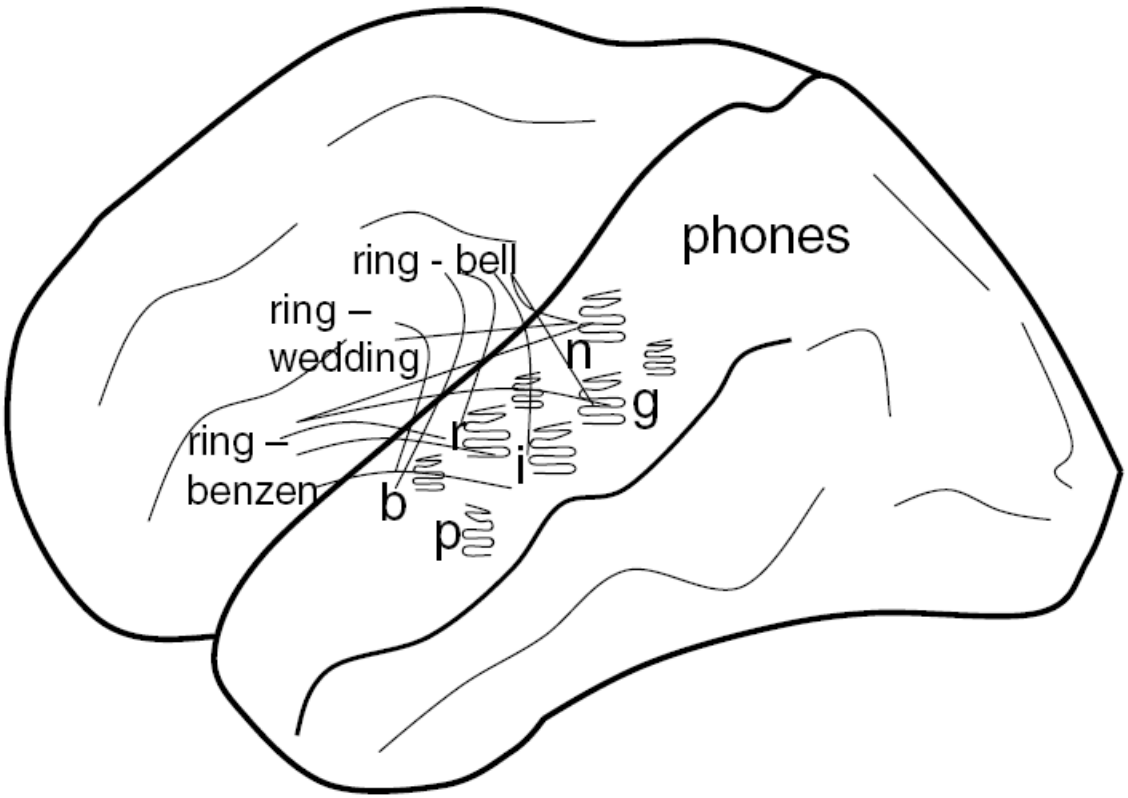






ds





eatures of products;

dictionary.

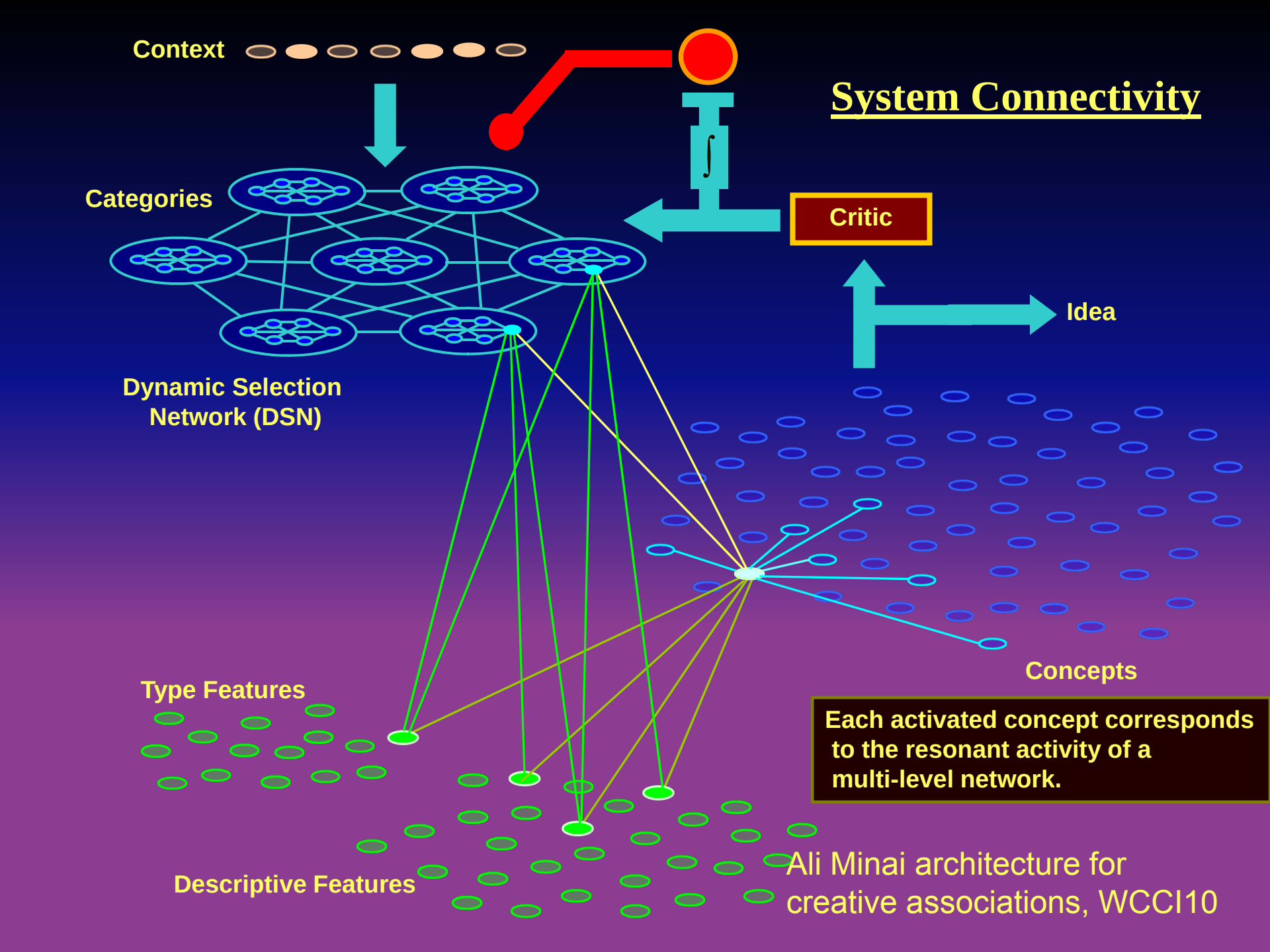
new words are being
y cortex.

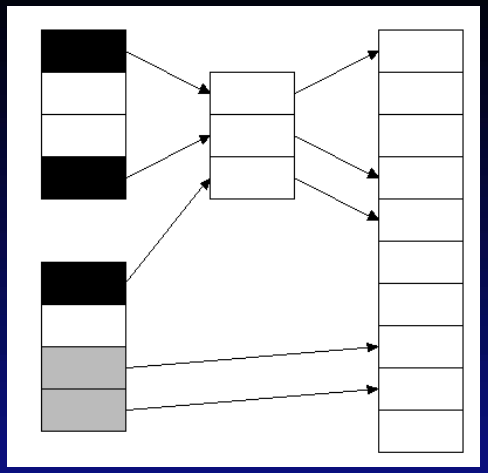
ation of phonemes will
ns; context + inhibition in
rds.

iltering (competition)

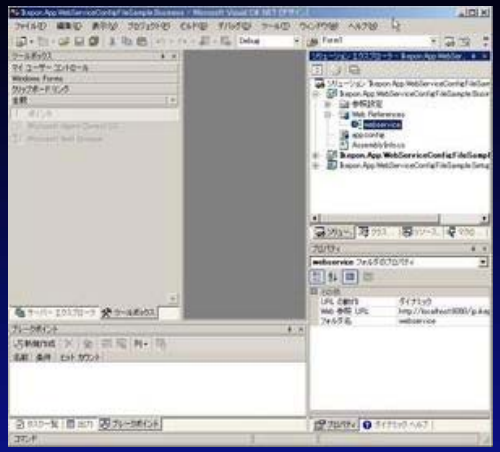
Imagination: chains of phonemes activate both word and non-word representations, depending on the strength of the synaptic connections. **Filtering**: based on associations, emotions, phonological/semantic density.

discoverity = {disc, disco, discover, verity} (discovery, creativity, verity)
digventure ={dig, digital, venture, adventure} new!
Server: http://www-users.mat.uni.torun.pl/~macias/mambo/index.php


Context
System Connectivity
Categories
Critic
Idea
Dynamic Selection Network (DSN)
Concepts
Type Features
Each activated concept corresponds to the resonant activity of a multi-level network.
Descriptive Features
Ali Minai architecture for creative associations, WCCI10



Query



Semantic memory

Applications, search,
20 questions game.

Humanized interface

Store

concepts
relations

Part of speech tagger
& phrase extractor

verification

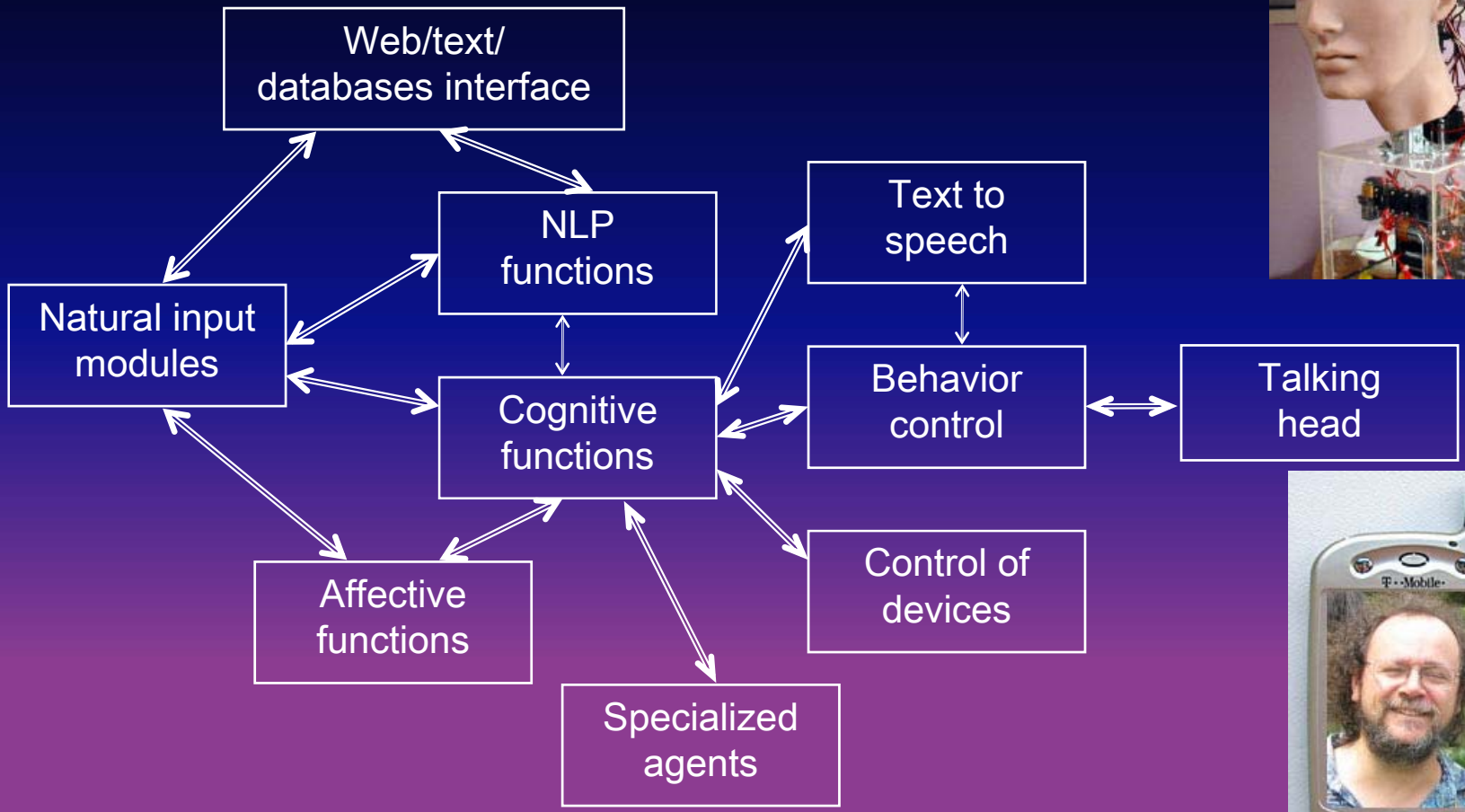
Manual

Parser

Internet

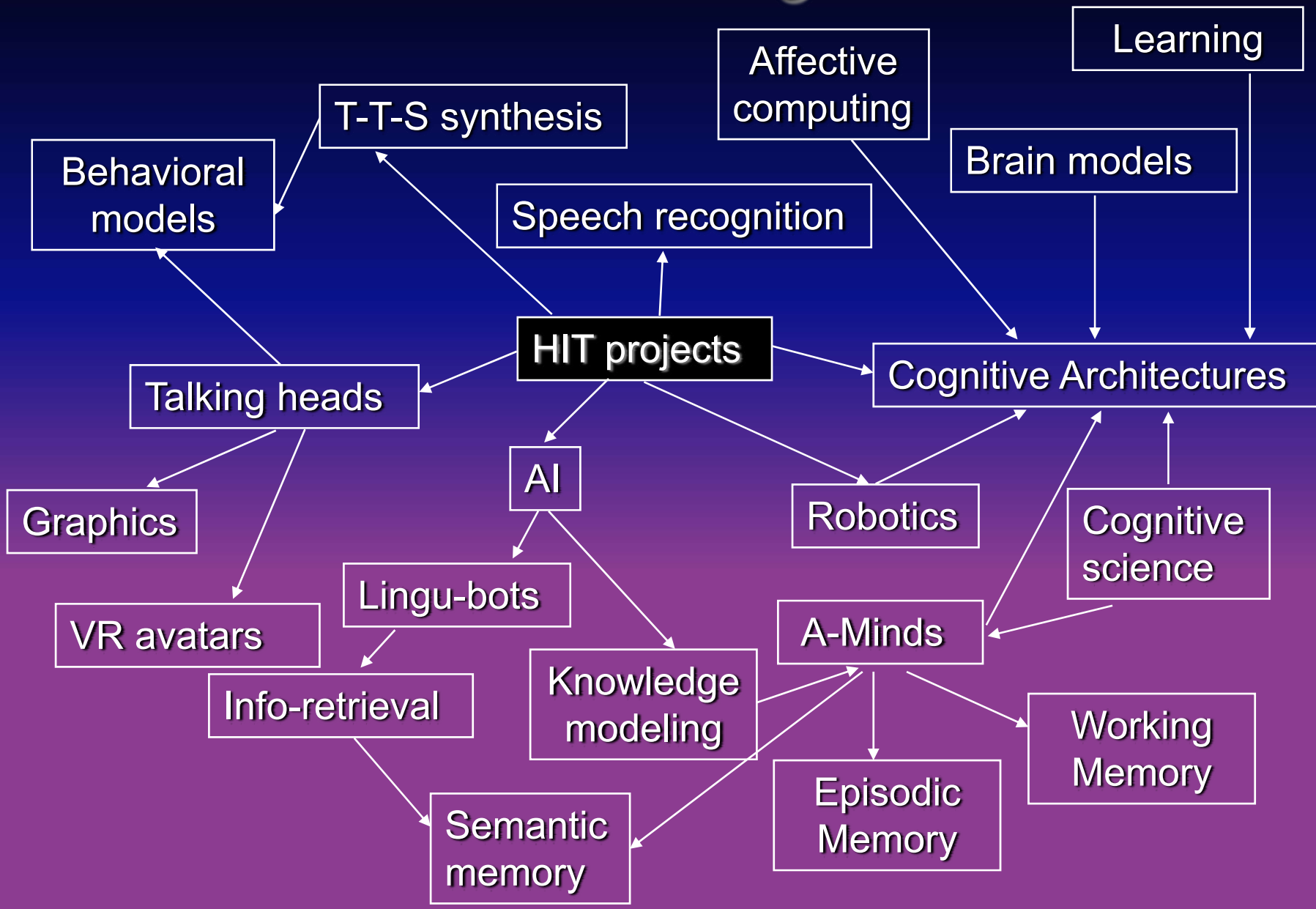
On line dictionaries
Active search and
dialogues with users

DREAM top-level architecture



DREAM project is focused on perception (visual, auditory, text inputs), cognitive functions (reasoning based on perceptions), natural language communication in well defined contexts, real time control of the simulated/physical head.

DREAM/HIT – larger view …

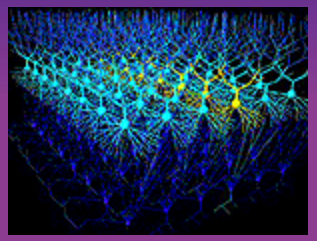


Connections

Geometric/dynamical ideas related to mind may be found in many fields:

Neuroscience:
D. Marr (1970) “probabilistic landscape”.
C.H. Anderson, D.C. van Essen (1994): Superior Colliculus PDF maps
S. Edelman: “neural spaces”, object recognition, global representation space approximates the Cartesian product of spaces that code object fragments, representation of similarities is sufficient.

Psychology:
K. Levin, psychological forces.
G. Kelly, Personal Construct Psychology.
R. Shepard, universal invariant laws.
P. Johnson-Laird, mind models.



Folk psychology: to put in mind, to have in mind, to keep in mind (mindmap), to make up one's mind, be of one mind ... (space).

More connections

AI: problem spaces - reasoning, problem solving, SOAR, ACT-R, little work on continuous mappings (MacLennan) instead of symbols.

Engineering: system identification, internal models inferred from input/output observations – this may be done without any parametric assumptions if a number of identical neural modules are used!

Philosophy:

P. Gärdenfors, Conceptual spaces

R.F. Port, T. van Gelder, ed. Mind as motion (MIT Press 1995)

Linguistics:

G. Fauconnier, Mental Spaces (Cambridge U.P. 1994).
Mental spaces and non-classical feature spaces.

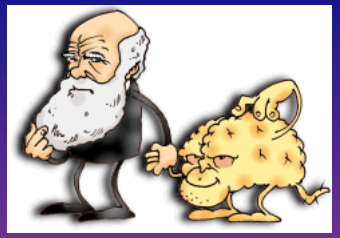
J. Elman, Language as a dynamical system; J. Feldman neural basis;
Stream of thoughts, sentence as a trajectory in P-space.

Psycholinguistics: T. Landauer, S. Dumais, Latent Semantic Analysis, Psych. Rev. (1997) Semantic for 60 k words corpus requires about 300 dim.

Conclusions

Robots and avatars will make a steady progress towards realistic human-like behavior – think about progress in computer graphics.

- Artificial minds of brain-like systems will claim qualia; they will be as real in artificial systems as they are in our brains.
- There are no good arguments against convergence of the neural modeling process to conscious artifacts.
- Achieving human-level competence in perception, language and problem-solving may take longer than creation of basic consciousness.

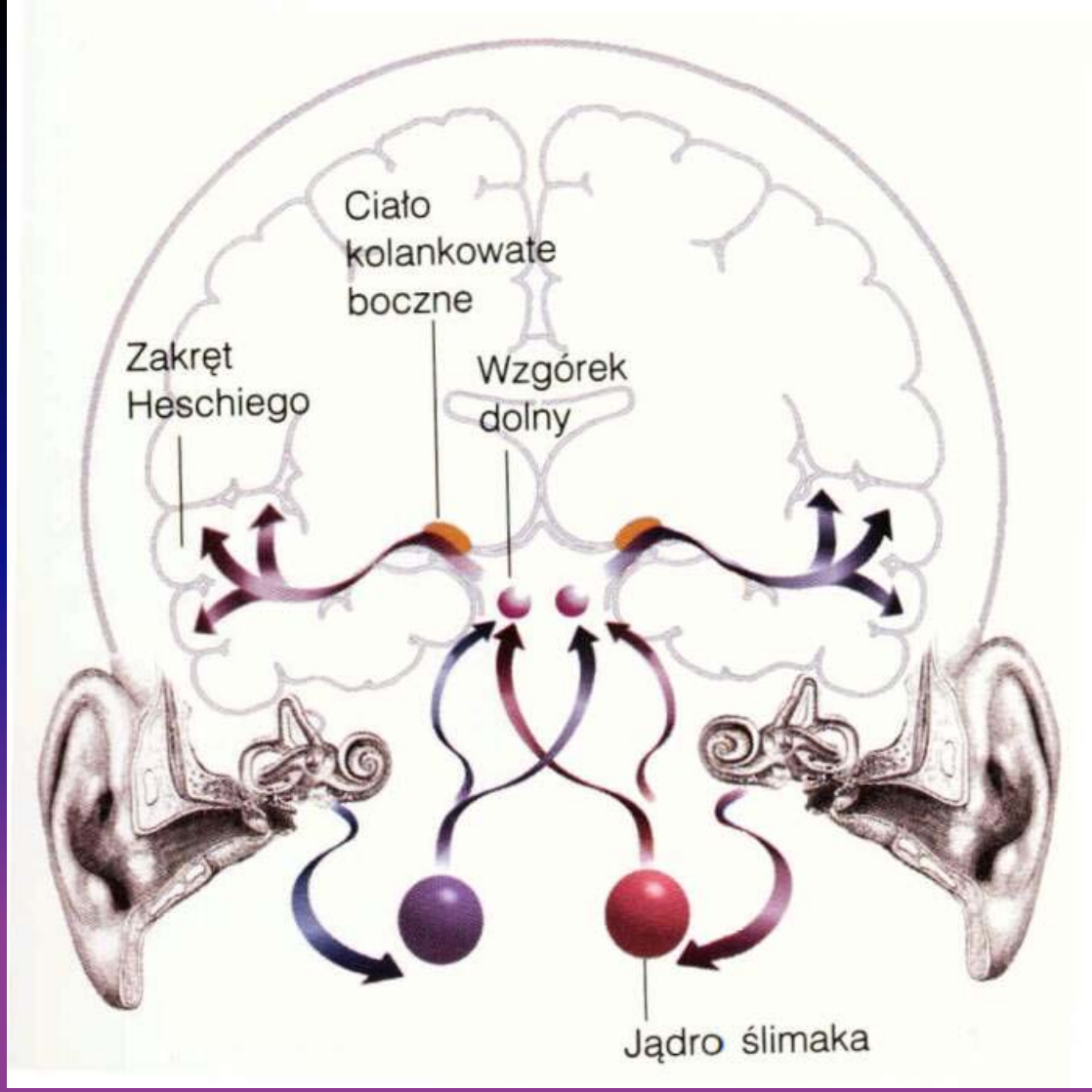


Creation of conscious artilects will open Pandora's box

What should be their status?
Will it degrade our own dignity?
Is switching off a conscious robot a form of killing?
...
Will they ever turn against us ... or is the governor of California already one of them ?

Exciting times are coming!

Thank you for lending your ears



Google: W Duch => Papers