Nokia Workshop on Machine Consciousness 2008
Päivä 1

Antonio Chellan puhe "Robotanic: an externalist outlook of a robotic architecture".

Panormo-robotti, viihdyttää vieraita botanisessa puutarhassa.

Mallin sydämessä havainto-toiminta -silmukka, jota muokkaavat valvojajärjestelmä ja emotionaalinen järjestelmä.

Malissa on kaksi oppimistapaa, match-mismatch ja vapaa oppiminen, jossa robotti tutkii robottia itsenäisesti. Siinä oli emotionaalinen järjestelmä antamassa eri toiminnoille kontekstia: eri "emotionaalisissa tiloissa", robotti toimii eri tavoilla. Emootioiden emulointi tekee myös helpommaksi ihmisille sopeutua robottiin. Toinen arkkitehtuurin osa oli supervisor-järjestelmä, jonka tehtävä on muokata robotin toimintaa havainto-toimintasilmukan perusteella ja suunnatta robotin huomiota.

Verrattuna Haikosen kognitiiviseen arkkitehtuuriin, robotin arkkitehtuuri on paljon suuremassa roolissa antamassa kontekstia kaikelle toiminnolle, siinä missä Haikosen mallissa emotionaalinen arviointi on vain yksi osa muiden joukossa.

Baarsin (Franklin LIDA) arkkitehtuuriin verrattuna on samankaltaisuus attentionaalisen spotlightin ja supervisorsysteemin kanssa.

Emotionaalinen järjestelmä perustaa Morénin ja Balkeniuksen amygdalaa simuloivaan neuronimall

Robotin emootiot: Joy sad anger suprise disgust fear, vaikuttaa myös robotin äänensävyihin. Emootiot esitetään kolmeulotteisena avaruutena: activation (A), evaluation (E), power (P).

Supervisorin rooli on allokoida havaintosilmukan resurssit. Kahneman (1973): tehtävän luonne vaikuttaa resurssien allokointiin.

DramaTour-metodologia: Informaation esittäminen on tehokkaampaa jos siihen liittyy dramaattinen narraatio. Robotti esittää draamaa ja pyrkii antamaan miellyttävän kokemuksen vierailijoille. Tämä saa vierailijat kokemaan robotin oikeana entiteettinä. Tarinan mukaan robotti tulee planeetalta Ziz-Balarm, ja putosi botaniseen puutarhaan vuonna 1795, juttelee kasveille, ja pelkää Manlio-Superbikeä, puutarhan kuraattoria, koska Manlio haluaa tuhota hänen virtapiirinsä. Robotti kiertää tiettyä aluetta puutarhasta.

Garden basinin lähellä on kriittistä että supervisor allokoi tarpeeksi resursseja, koska robotin havainnointiin käyttämä laser ei huomaa järveä. Basinin lähellä robotti allokoi lisää resursseja videonäköön ja korjaa kartan paremmaksi.

Seuraavana oli Will Brownen puhe Emotional Cognitive Steps Towards Consciousness.

Onko emootioista hyötyä roboteille?

Learning Classifier Systems. Kokeillaan erilaisia sääntöjä datan käsittelemiseen ja sovelletaan geneettisiä algoritmeja parempien sääntöjen käsittelemiseen.

Toimiko tämä? Iso sääntötietokanta, hidas käyttää. Joutui usein samanlaisiin tilanteisiin, mutta niissä oli aina pieniä eroja, yleistämistä ei tapahtunut. Saattoi tehdä asioita, mutta paljon muuttui.

Emootioiden määritelmästä ei oikein ole yksimielisyyttä. Siitä on, että emootioit eivät ole yksinkertaisia kuvauksia tapahtumasta taikka "jos tila, niin emootio..." -produktiosääntöjä.

How to make benefits of production rule systems:

* Reinforcement learning
* Global search
* Transparency

Environment <--> Needs and desires

Ominaisuuksia:
* Kun tarve täytetään, emootio alkaa hiipua (kun olet kylläinen, nälkä katoaa)
* Emootiot eivät riippumattomia eivätkä lineaarisia
* Emotiot ovat linkitettyjä, dynaamisia (multiple causes interacting over nested timescales), rajoitettua, attraktoritiloja (pienet tönäykset eivät saa pois attraktiotilasta).

Tila on eri olosuhteiden vektori. State s - Action a - Reward r.

What makes emotions useful?

* Behaviour emerges rather than hard coded
* Generalize across known and unknown situations
- Tietyissä tunnetiloissa muistaa helpommin samaan tunnetilaan liittyviä muistoja: tehokkaampaa yleistämistä
* Episoidic and temporal
* Fast reposne
* Non-linear, nond-dterministic stochastic

Experimental
Robot-emotions useful for successfully completing real-world tasks
* Emotions can set goals; balance explore vs. exploit
* Emotions can modifty existing behaviours
* Emotions facilitate action in unnown domains

Carlos: Systems, models and self-awarneess: Towards a SysML model of consciousness

Konetietoisuuden ongelmat

1) Mitä on tietoisuus?
2) Emme tiedä miten rakentaa tietoisuuden? (osittain koska 1)

Kaksi eri strategiaa konetietoisuuden rakentamiseen:

* Direct approahces: try to build a machine following a particular theory of consciousness (eg. GWT)

* Incremental: Try to add competencies one-at-a-time up to reaching the level of consciousness

* No matter what is the approahc, MC will suffer a receding-horizon phonemenon similar to AI (there is not a ultimate test procedure).

Towards a positive theory of consc.

* Too many basic understandings of what consc is
* Can be traceed back to architectural reverse-engineering problem: the extractin of function from non-formal description of architectures
* A rapid scan of the literature on the topic leaves the following impressions:
 * Most theories that target the whle thing seem just literature (or palin rubbish)
 * Positive theories are mostly seen as naive

* We need aunified theory of consciousness
 * Targets the whole thing (even qualia)
 * That is widely agreed across disciplines)
 * That is expressible in different abstratcion levels to be at the same time
  * general and
  * precise and
  * verifiable

* We must express the theory in a formal enough language to minimize the problem of multiple interpretations
* Some approaches to "formal theories" are vailabe are available but far from targetting the whole

Modes of expression

* Verbal-Linguistic
  * Why are we stuck here?
* Logical-Mathematical
 * Equations, MatLab simulations...
* Graphical
* Others?

Rationale ofr a Systems Approach to Consciousness

* The reconition of the enormous complexity of building consc. machines points into a direction that can help solve both problems

* Objective: Reference Models of Consciousness
* Cnsolidate a unified vision on consciousness functions and mechanisms
* MEthod:
** Define an ontology for describ consc. cmponents and systems
** Define reference models as standard elements that include ..

Can both address problem 1) and 2)
Unifies the model-based theory of consciousness and the model-based practice of embedded systems construction

Minds are model-based controllers!

Three classes of global research objectives for MC:
* Building machiens like us (the C3PO drive)
* Understanding biological consciousness
* Building better machines

The key is using enineering models as self-models
* Autonomous control based on self-models enables an increased awareness that can lead to more robust performance

SysML model will enable the expression of a consciousness models through formal models.

Ruokatauko.

Seuraavana vuorossa oli Ben Goertzel, esityksen nimenä Achieving advanced machine consciousness via Artificial General Intelligence in Virtual Worlds.

A useful philosophical perspective on consciousness.
Metaphysical foundation: Charles Peircean/Jungian categories

First: raw, unprocessed being... e.g. qualia
Second: reaction ... e.g. pure physical reaction
Third: relationship

(beyond Peirce... 'Fourth: synergy', etc.)

First person: firstness of X... the world as directly experienced... the stream of qualia...

Third person: thirdness of X... the world as an objective, relational structure, a network of patterns

Fourth person (normally called "second person"): fourthness of X... the synergy of relationships... the Buber-ian I-You

The real second person: secondness of X... experiencing the world as an automaton

Example of a hypothesis spanning perspectives:
The more intense qualia experienced by a system, correspond to the more infomrationally singnificant detectable in that system by an intelligent, well-informed observer.

Hypothesis:
Among the more informationally significant patterns in gneerally intelligent systems are:

* The phenomenal self
Metzingerin "sisäinen VR"
* Reflective consciousness
* The illusion of will

AGI vs. kapea AI

Teoria: AGI emergoituu noista kolmesta patternista

Hypothesis: integrative design can allow multiple AI algorithms to quell each other's combinatorial explosions

---

Tätä seurasi Yasuo Kinouchi: A logical model of consciousness on a neural network system with a simple abstract brain-like structure.

Objective: To develop an information processing system that works in a way similar to primitive thinking, or very simple brain, we investigate a logical model of consciousness in an autonomous adaptive system wiht a simple brain-like structure.

Functional hierarchy: A method that functions of a cmplex system are described hierarchically in various layers. Usually, upper layer functions are simply and understandable, and lower layer function specify system operation

Invisible computing: A method that a user or an operator of a system aware of necessary matter only, invisble details are neglatcted.

1st step: Lower layer is designed. Restraint conditions and basic structures similar to the brain are set to the system.

2nd step: Applying locial operation based on "what we feel or not" to the lower layer, "self and consciousness" are defined as locial substances that control the system simply and efficiently.

Basic assumpitons
1. System must adapt autonmously o a complex environemnt without a teacher. Large amounts of information are input to the system fromt he environemnt. Main functin of the syste, is to decide action.

2. System has basic functions correspondg to the instinctual body-control systems of animals. For example, pleasant and unpleasant awards. System composed of an adaptable and inadaptable part.

3. The system must decide upon an action ASAP in accordance with situation and experiences, though neural nodes that operate very slowly are used.

4. The system must use tehse as effectively as opossible.

5. Only a minimum amount of resources and energy can be used by the system to adapt.

Information representation: To process large amoutns of infromation speedy and efficiently, the system has three levels of information represenation. Mciro features, concept and scene

Micro feature: minor features of object (red, blue, round, square...)
Concepts: A set of micro features
Scene: Combination of concepts

Basic processing of lower layer

Adapatation executed by 3 stages
1st stage:
* Encoding (concept recognition), pre-selection (a kind of attention)
* Reduction of info by encoding (concept recognition) and pre-selection (kind of an attention), information interact with one another in each local area

2nd stage:
Re-configuration of an object and location
Information are buffered

3rd stage:
Top-level selection of effective scene
* reward prediction, action decision
* Memorizing and recollection
Various information scattered in wide are interact with another

Main learning method of the system
The system tries to take repeat actions that resulted in a pleasant feeling and avoid repeating actions that resulted in an unpleasent feeling. The system performs these actions faster via learning.

Environment ->
Sensor ->
Concept module (stage 1)
Binding Module (Stage 2)
Declaration memory module
2nd stage: Binding (re-configuration and buffering, Binding images for top-level selection) -

Concept module with concept formation functions operates bi-directionally as an encoder of large amount of information expressed by micro-features, and a decoder of recollect group of concepts in the system. Pre-selects important information.

Binding module connects information of an object and its location so that the voting and decision in 3rd stage can be processed quickly and appropriately.

Declarative memory module memorizes combination of selected scence, decided action, and reward spntanoeusly and sequentially. Extract of generalized appropraite rules of actin from sequence scenes, rules and rewards.

Integration and core module selects the most important comibnation of concepts or information in the system at that time on the basis of a mutual voted executed by nodes representing concepts or emotional states. The core part only has an macitve control system and decided whetther to act as a systems based on emotional states. Emtion is thought to be information sent automatically from the automated part of the system.

To decide upon an appropriate action as quickly as possible at the time, processing is executed mainly in 3 paths.

1: The quickes porcess, bypasses the BM and ICM.
2: An impulsive rpocess in which an action can be quick, but sometimes inappropriate because declarative memory was not accessed.

3: A reflective process in which an action can be slowly and more appropriately taken because of the use past experiences with assosciative memory access.

Self and consciousness as upper layer function

Upper layer: We can feel activities in this layer. Scenes or imageries are associatively recollected.

/\
|
| Correspondence related to "what we can feel or not"
| "The information not selected is invisible in the upper layer
|
\/

Lower layer: Each information are processed in detail. We can not feel activities directly.

Self and consciousness in upper layer
The "self" recollects past experiences from declarative memory as if it is in the world expressed compactly by decoded infromation such as scenes (combination of concepts), and decided actions, though it does not directly feel these decided actions. It feels only a kind of images to act, or a change in the body and environment by this action.

Discussion: Consc. and decision flexibility

Conscious works effecitvely when the system flexibly decides on an action by considering the overall situation

Path 1: To act ASAP, the sysmte is unconscious of the action. Actions become a patterned system without flexibility.
Path 2: Concscious to the action without recollection.
Path 3: The system is conscious to the action with recollection. Actions are more appropriate and rich in variety using past experiences.

As "self" cannot be directly aware of its own state with the exception of emotional memory, the contents of "sefl's state" has to go around CM and BM to be conscious. The "self" can only see thoruhg itself: for example , a miorrr, and has a porperty of a "recursive self" based on the basic confifguration. We suppose this is basically comtability with findings of Libet.

Harri Valpola: The engine of thought - a bio-inspired mechanism for distributed selection of useful information.

Tietoisuuden neuraaliset korrelaatit melkein mahdottomia: hyppäävät suoraan neuronien tasolta korkeimmalle tasolle, skipaten algoritmien tason välillä. Sen sijaan:

Käytöksen taso
/\
| Neural correlates of infomration processing
\/
Laskennan taso
/\

| Information processing correlates of consciousness
\/
Biologian taso

Marr: Algoritmi on helpompi ymmärtää sen ratkomien ongelmien kuin tutkimalla sitä mekanismia, joka sitä pyörittää.

Problems to be solved by the brain:

Make decisions
* The ultiamte purpose of the brain: control movement
Perception
* Extraction information about the world
* Needed for control, decision making
* Find regularities, develop representations
Simulate the future
* Needed for decision making

Bayesian theory says:
* Decisions are based on
1. Beliefs (measured by probability)
2. Utilities
* The recipe:
1. Evaluate the probabilities of all possible states of the world (probabilistic inference)
2. Evaluate the probabilities of all outcomes for each and every potential action (probabilistic inference)
3. Choose the action which maximises the expected utility
* This is optimal if there are no restrictions on the available computational resources

How to select information?
* But... computational resources are restricted -->
* It is impossible to consider all states and actions -->
* It is necessary to select information in order to make decisions

* Selection is a type of decision, in other words:
* In order to decide we need to decide... infinite regress!

Let's study the solution adopted by the brain

Hierarchy of areas
* The cerebral cortex is connected as a hierarchy of areas
* The representations get more abstract on higher levels

Korkeampien visuaalisten ja auditooristen alueiden välillä yhteyksiä, sekä ylhäältä-alas yhteyksiä alueiden sisällä.

Brain's solution: distributed selection
* Each cortical area selects information to be represented
* Biased-competition model of attention: attention emerges from local selection and global communication

Menee vaikeaksi, kun joutuu sekoittamaan eri abstraktiotasoja.

Attention and learning: selection on different timescales
* Our work: biased competition + learning
* Within the Bayesian framework, the only difference between perceptual inference and learning is the timescale
* Attention and learning in the cortex are intimately coupled
* Both are a form of selection, only timescales differ

The value of information
* Motor areas may be able to rely (at least partly) on global reward signals (reinforcement learning)
* Sensory areas or a large part: credit assignment problem
* More specific but locally available information: predictive power or "are the others listening?"
- Tiedemiehet haluavat muiden viittavaan heidän papereihinsa: kunnianhimo aivojensa käyttämä tapa mitata toimintojen tehokkuutta

what about consciousness...?

* We are beginning to understand the information processing in the cortex (perception, attention, learning, imagination, decision making...)
* Now we can ask how it relates to consciousness

* My own answer: information processing (the list above) relates to consciousness as metabolism and replication relate to life
* But: most people need to meet a sufficiently intelligent machine to get convinced (a nice robot would do fine)

Uma Ramamurthy: Might a LIDA Controlled Robot be Phenomenally Conscious?

LIDA model
* Model of cognition
* Implements Global Workspace Theory
* Learning mechanisms
* Control structure for software agents and autonomous robots

Mechanisms/modules in LIDA
* Perception module - Copycat (Hosftader)
* Global Workspace Theory (Baars)
* Memory Systems - Sparse Distributed Memory (Kanerva)
* Action Selection - Behavior Net (Maers)

Enhancements in LIDA
* Various Learning Mechanisms
* Feelings and Emotions as primary motivators and learning facilitators
* Expectation and Automatization
* Sensori-Motor Automation

Cognitive Cycle - Step 1
* Sensory stimuli received and interpreted by perception
* Perceptual associative memory identifiers relevant emotions connected with the objects and relations in the percept

Cognitive Cycle - Step 2
* The percept with is interpretation stored in working memory

Cognitive Cycle - Step 3
* With incoming percept and residual contents of preconscious buffer as cues, local associte retrieved from longer term memories
* Emotions form part of the cues

Cognitive Cycle - Step 4
* Atteniton codelets (processors) look at LTWM, gather information, form coalitions and compete for access to consciousness
* Present and past emotions influence the competition for consciousness

Cognitive Cycle - Step 5
* The winning coalition of processors gains access to the global workspace and its contents broadcast

Cognitive Cycle - Step 6
Relevant behavior codelets respond to the broadcast, thus consciousness aids in recruiting resources
* Emotions assist in attracting relevant resources to deal with current situation
* Recurrent processors instantiate new goal context hiearchies

Cognitive Cycle - Step 8 & 9
* Behavior net chosoes a single behavior from a newly instantiated or previously active behavior stream

Cognitive cycles
* Every autonomous agent operates by frequent iteration of sense-process-act -cycles.
* Steps in a cycle overlap to enable parallelism between adjacent cycles
* Consciousness maintains serial order, the illusion of continuity
* Cycle may start with action selection

Learning mechanisms
* Perceptual learning
** Identify, classify, relate

* Epsiodic learning
** What, where, when

* Procedural learning
** New or improved skills
* Automatize

* Attentional learning
** To what to attend

LIDA - How Conscious?
* LIDA implements GWT (Baars)
** Implements only part of consciousness
** _Functionally_ conscious in the sense of GWT

* Does LIDA need mechanisms to produce a stable, coherent perceptual field (Merker 2005)?
** Separate target-related guidance of behavior from own movement produced sensory information
** Might be significant step toward phenomenal consciousness

Merker's Proposal
* Logistics problem created by self-motion
* Make targets available for decision making
** Present targets as part of a coherent, stable world-space

Providing a stable arena
* Spatially sensitive sensory mechanisms -
** Movement of sense organ produces apparent motion
** Independent of actions in the environment

* A mobile, autonomous robot
** Requires spatially sensitive sensory mechanisms to move appropriately in its world
** Ability to distinguish motion in the domain from self-produced motion

Mechanisms needed in LIDA
* To differentiate between real movements and apparent movements (self-produced) in its perception module (a feedback loop?)
* To shield the robot's action selection from apparent motion produced by its own movement of its sense organs
* To effect a coherent, stable perceptual world condition for the robot
* Other mechanisms... a Self system

Conclusion
* Limitations and gaps in the LIDA model
* Our continuing goals:
** To understand how the mind works
** To model the cognitive processes
** To build smart systems (Software agents, robots)
* Do these systems need phenomenal consciousness?
* Mechanism that will take LIDA model in that direction

ccrg.cs.memphis.edu

Ryhmäkeskustelu

(Paljon puheenvuoroja joita ei ole kirjattu ylös, koska en ehtinyt tai puhujalla oli liian vahva aksentti jotta olisin pysynyt kärryillä)

Kysytään yleisöltä mitä tahdotaan käsitellä ja koetetaan sitten käsitellä sitä.

Eri arkkitehtuurien vertailu

Ricardo: Got the impression that most architectures presented here handle even things like perception in such different ways that no comparison of the models is possible. They reflect each one engineer's concept of the most basic cognitive details too much.

Somebody: Agree that there are fundamental differences, but also similarities which can be compared. You have to understand the systems in more detail to do that, however. There are different goals such as biological plausibility.

Valpola: Biological plausibility depends on the level of abstraction - none are plausible on the level of molecules, but going up, each of them are somewhat plausible.

Goertzel: Agreed that each of the systems are based on different perceptions, but look at transportation: people use helicopters, hovercraft etc., but it's still transportation. One problem is that not many researchers really understand the other's ideas, even reading the papers isn't enough as there isn't a shared vocabulary. Ten years from now, you could see ten totally different approaches with shared vocabularies.

...

Goertzel: What are the basic principles behind different forms of transportation? You have to go down to physics to get the basic principles that allow *any* movement. It may be the same thing in intelligence: you may take different approaches, each with different stregnths and weaknesses.

Keskustelua teoriasta, jonka mukaan yksilö ei voi olla tietoinen, vaan tietoisuus kehittyy yhteisöstä. Haikonen epäili tätä ("lyö vasaralla sormeasi mahdollisimman lujaa, niin saat todisteen siitä että olet tietoinen vaikka olisitkin yksin"), Ricardo taas ajatteli että kasvatus saattaa tuoda meidät siihen pisteeseen että voimme olla tietoisia vaikka emme olekaan seurasssa.

Q: How necessary is embodiment?

Ricardo: The concept of embodiment is rubbish. There are no disembodied things (aside ghosts and spirits, which don't exist). Everything is embodied. I have been asking people whether they have a test of when a system is sufficiently embodied, and they never do.

Aleksander: Maybe people, when talking about embodiment, aren't speaking about embodiment as such, but rather a connection with the world.

Ricardo: Microsoft Word is connected with the world!

Yleisönjäsen: It's not just about having a physical system, it's more related to the body being a part of the physical (and other) activities we are doing.

Ricardo: You can't separate the mind and the body. You can't separate the program and the processors executing it.

Yleisönjäsen: Can a mind develop consciousness without interacting with an entertainment? Is there a minimum bandwith? How badly can we separate a mind and an environment?

Goertzel: You could create a very intelligent system without any body in the normal sense and which did nothing more than proving mathematical theorems. Embodiment is clearly bread and parcel of human-like intelligence, and human-like intelligences certainly require embodiment, but that doesn't mean embodiment is a certain requisite of general intelligence.

Valpola: Agreed with Ben, and I think embodiment is a technical necessity: if we don't know how to train an intelligence, we can create embodiment and experiment, but I don't see anything in theory against it. It's like asking "can you have

Somebody: We don't need consciousness for an AI that proves math theories. What kind of an AI needs consciousness?

Aleksander: Penrose would disagree with you, saying you need "insight" for intelligence, including mathematical theorem proving.

Goertzel: I don't agree that math theorems wouldn't require reflective consciousness.

Somebody: I mean that do you need qualia-like feelings to prove math theorems?

Goertzel: I would say so. I'd say that to achieve high-level generality in AI systems will lead to the emergence of qualia-like structures.

Tests of consciousness

Kinouchi: I think consciousness is a control system for something. If something has that control system, then it is conscious.

Aleksander: But there are many control systems, in airplanes for instance. Are they conscious?

Valpola: I think if you can define what consciousness is good for, then by testing whether something is good at something, then that's a test of consciousness. In principle you could build a rule-based system for doing the same things as for conscious machines, but in practice not.

Joku / Ricardo: There are different definitions for consciousness, technical functional ones and

Sleep

Aleksander: On the role of sleep - you may want to Google Sidarta Rai(jotain), he's done wonderful work to prove that all the levels of sleep are necessary to modify memory. It's a physiological paper, and I think explains a lot of stuff about why we sleep and such, Nobel prize winning stuff.

Scaling and complexity

Aleksander: That's a big issue, as any demonstrations are going to be on a much smaller scale than reality. Is it true that if we build a system well below the scale of even a small animal we believe to be conscious, that the results will be valid?

Goertzel: I think that conscious structures and such may not emerge below a certain scale. The system in order to build a model of itself, based on its own interactions with the world, needs a lot of data. Human babies gets a lot of data about interaction between itself and the world, and then there's this pattern mining about all this mess of a data to build a model, like we're doing with our virtual dogs.

Yleisönjäsen: Are there any sharp limits, above which there is consciousness, or is it a question of defintion?

Aleksander: No, there are no sharp limits.

Valpola: Back to the scaling issue, in a very simple world evolution wouldn't build a conscious system to deal with it, though we might be able to build a conscious system to deal with it. I agree with Ben that you need enough representational capacity in order to model yourself.

Nokia Workshop on Machine Consciousness 2008
Päivä 2

(Nukuin pommiin ja missasin osan alun luennot, kuten myös osan tämän alusta.)

Markku Åberg: Neuron microchips

* R Efficient Prciese Computation with Noisty omcomponents: Extrapolating from an elextronic cochlea

* With low S/ ratio, analog computation is both more powerful and more area efficieint

* When wiring is cheap, minim cost of computing is achieved with a low S/N ratio and high paralellism (favoring analog system)
* this is the solution biological nerves have

* In microelectronic wiring is rather expensice and high S/N ratio, can easily be achieved, therefore digital systems often favored

* But if the system is inherentyl highly paralell, analog can be favored

Consciousness?
* It is not in the physical platform
** But needs some physical platform to exist

* It can be realized on many different platforms
From electronics to biology to...
* But some platforms may be vastly more effective than others

Can it be realized on any type o fplatform?
* When made large and complex enough
* Or does it need in principle certain types of hardware architecture?

Example: Convulotional decoder

* The decoding problem is to find a bit sequence B, which minmizes the function: (function)
* The output for kth neuron for L = 3 rate 1/2 decoder can be expressed with (functions)
* The RNN decoder consists of the followingn blocks, multiplier, a CSD-block (D/A-conversion, summing decision) and AWGN-noise generator
* The core of the neuron is a CSD-block, which D/A-converts the terms, sums them togehter with a AWGN-signal and finally makes a decision

Benefits of the approach: The algorithm needs a lot of (low bit) summings, which are much faster and more power effective in analog domain.

Associative neuron Microchip (1)
The operation of the associative neuron

Conclusion

Several types of neural circuits implemented:
* RNN decoders
* Associative neuron circuits
** Several different versions
** With analog, digital or hybrid synapses
* Standard CMOS processes used
** With some tailored devices
* All circuits with intended functionality and performance
* Howeever, their sizes have been rather limited
** More proof-of-concept

ConsScale: A Plausible Test for Machine Consciousness? Raúl Arrabales.
www.conscious-robots.com

Problem: Accurately testing for consciousness
* Biological organismis (humans, other mammals, bacteria...)
* Artificial agents (autonomous robots, control systems...)

We need a measure of consciousness
* Defined levels help evaluation

Why?
* To better understand consciousness
* To find out if we can effectively test for consciousness
* To assess the state of the art in Machine Consciousness

This is a controversial and unsolved question.

Biological criteria cannot be used directly:

- Different underlying machinery
 - NCCS vs. CCC
- Different behavioral patterns
 - Biological body vs. artificial body
 - Cultural environment vs. artificial ecosystem
- Lack of accurate verbal report
 - Communication skills

Testing methods:

Determining the presence of consicousness:

- Turing Test
- Heterephenomenology (Dennett 1991)
- NCC (Crick & Koch 1999)
- NCC + Behavioral patterns (Seth, Baars & Edelman, 2005)

Determining the level of consciousness:
- Glasgow Coma Scale
- Simplified Motor Scale
- Private Self-Consciousness Scale (Fenigstein et al. 1975)

But what about artificial agents?

Criteria for the presence of consciousness:
- Minimal set of axioms (Aleksander & Dunmal 2003; Aleksander & Morton 2007)
- Perception, Imagination, Attention, Planning, Emotion

- Haikonen (2007):
- Reportability
  - Ability to report conscious consent (to others and to oneself)
  - etc.

Determining the level of consciousness:
- ConsScale (Arrabales et al. 2008) offers a functional classification of agents

- Levels of development of consciousness:
- Agent's architectural components
- Agent's cognitive skills
- Agent's characteristic behavior

Computational approach:

Abstract model of agents, based on Wooldbridge (1999)
B - Body
E - Environment
S - Sensory Machinery
A - Action Machinery
R - Sensorimotor Coordination Machinery
M - Memory

Cognitive functions
Theory of Mind (Lewis 2003)
- "I know"
- "I know I know"
- "I know you know"
- "I know you know I know"

Executive Function (Perner & Lang 1999)

Emotional Functions (Damasio 1999)
- Emotions
- As physical reactions of body triggered by an object
- Feelings

- As representations of the physical reactions to objects
- Feelings of Feeling

ConsScale:

-1 Disembodied
Behavior: not a situated agent
Phylogeny: amino acid

+0 Isolated
Behavior: not a situated agent
Phylogeny: isolated chromosome

+1 Decontrolled
Behavior: not a situated agent
Phylogeny: dead bacteria

+2 Reactive
Behavior: reflexes
Phylogeny: virus

+3 Adapative
Behavior: basic ability to learn new reflexes
Phylogeny: earthworm

+4 Attentional
Behavior: attack and escape. Attention + emotion
Phylogeny: fish

+5 Executive
Behavior: set shifting. Emotional learning.
Phylogeny: quadruped mammal.

+6 Emotional
Behavior: feelings influence behavior
Phylogeny: monkey. ToM stage 1: "I know".

+7 Self-Conscious
Behavior: advanced planning. Usage of tools. Mirror test.
Phylogeny: monkey. ToM stage 2: "I know I know".

+8 Empathic
Behavior: making of tools. Social behavior.
Phylogeny: chimpanzee. ToM stage 3: "I know you know".

+9 Social
Behavior: linguistic capabilities. Ability for culture.
Phylogeny: human. ToM stage 4: "I know you know I know"

+10 Human-Like
Behavior: accurate verbal report. Culture. Technology.
Phylogeny: human. Adapted environment.

+11 Super-Conscious
Behavior: several streams of consciousness.
Phylogeny: n/a.

Theoretical foundations:
* ConsScale
- Functional and bio-inspired approach
- Abstract architectural components
- Minimum required behavioral patterns

* An agent comply with a given level if and only if
- It shows all the behavioral patterns of all lower levels
- ..

* ConsScale follows a particular discretized path within the map of possible implemantations
- Identify the most signifcant levels of phylogenetic development that led to human-like cosciousness

Future work:
* Redefine levels as required
* Concrete tests have to be defined for each level.
 - Learning capabilities
* Classify existing implementations according to the proposed scale

---

Emmanuel Lesser
A cognitive approahch to silent speech detection
www.avai-poject.net

* Does a conscious AI tasks better than an intelligent system?
* Is it possible to recognizie multilingual speech by analyzing EMG signals?
* What changes in architecture are necessary to make an existing system conscious?

Silent Speech Recognition
- First developed by NASA

- Goals
 - Communication in harsh environments
 - MEdical applications
 - Subvocal speech

- Extract speech information from sEMG signals in the vocal chord region
- Simple system architecture: data acquisition -> data processing -> recognition

- One differentially amplified EMG channel
- Bandlimited signal: 10 - 2000 Mhz
- Via EMG device
- 1500 useful samples acquired
- English, Dutch, Chinese

- We want to obtain a training set that can be used with supervised learning techniques

- Steps
Activity detection -> Feature extraction -> Reduce dimensionality

- Feature extraction
 - Full-wave rectification
 - Kingsbury's Complex Dual-Tree Wavelet Transform

(teknisiä detaljeja)

What is consciousness?
- Thinking and imagination, the immaterial feeling of being here, self, mind & free will...

The associative model of consciousness
- Consciousness is invoked by associative processes in the brain
- Processes in the brain are nonnumeric
- Introspection is an intrinsic requirement for machines consciousness architectures
- Distributed representations and cognitive architectures
- Haikonen Associative Neural Network (HANN)

"AI is definitely artificial has somehow managed to exclude intelligence"

(teknisiä detaljeja)

Results: Vertical grounding circuit (word meaning)

Increase in accuracy
 - For multilingual speech
 - 85% correct as CGNN for normal speech

Random number generation
 - decrease in accuracy

Bistable perception
 - Significant increase in accuracy for simulation results

- IEEEI 2008

Conclusions:

- A conscious system always performs AI tasks more accurately than an intelligent system, but decrease in results can occurr if consciousness (or introspection) is not wanted
- Mutlilingual speech can be recognized via EMG analysis. Consciousness helps to distinguish between different languages (not necessarily recognize).
- Changing the NN part in the system architecture is often sufficient to obtain better results.

Toward an expectation-based robotic model of affective experience
Ron Chrisley, COGS/Informatics, University of Sussex

Part 1. General issues

Explanatory MC research:
* Modelling vs. implementing
* Science vs. engineering
* Emphasis on modelling does not diminish the importance of embodied, working implementations

Take Home Message 1: Not all MC research need have as its goal the creation of a conscious artifact.

Synthetic phenomenology:
* An example of explanatory MC
* Using the states of a robotic model of experience to specify the content of experience (similar to, but distinct from Gamez' usage)
* More precise than nautral language; allows specification of non-conceptual aspects of experience
* Discriminative vs. constitutive accounts
* Applications: modelling of user experience (NB Prof. Alinikula!)

Take Home Message 2: Not all MC research need be an attempt to solve the "hard problem".

Interactive empiricism:
* Consciousness: impasse is conceptual
* Some conceptual change can only come about as the result of activity: intervention is the subject matter
* So the activity of doing MC is well poised to help us develop our concepts of consciousness to resolve apparent paradoxes

Take Home Message 3: MC research is as much about changing us as it is about machine functionality.

MC and the explanation of qualia:
* An example of MC-drive conceptual drive
* Mismatch between our concept of qualia (as defined by, e.g., Dennett) and physiological facts
* Typically, two responses: dualism (Chalmers; Jackson); or eliminativism (Dennett)
* But just because nothing meets our concept of qualia doesn't mean qualia don't exist (cf. ancient concept of gold - "a compound", "a gift from the gods", etc.)
* Qualia will be the actual facts about a system that make it useful/inevitable to think of itself as having prviate, intrinsic, ineffable, etc. qualitative states

Take Home Message 4: MC research can provide a non-dualistic but non-eliminative account of qualia.

Part 2. Expectation-based architecture.

EBA: An implementable framework for modelling perceptual experience

Expectation-based theory of visual experience:
C = Part of the non-conceptual content of the visual experience of a subject at a time t
C = conjunction, for all actions a, of E(a, t): the foveal input the agent would expect to have at t + ot were it to perform action a at t

(Tässä kohtaa iski ajanpuute, luennoitsija skippailli hirveästi kalvoja ja antoi vain THM:t)

Take Home Message 5: Your visual experience at t is determined by your expectations at t, not your input at t.

Take Home Message 6: Representations whose contents are part of experience at t need not be tokened (active) at t.

Take Home Message 7: EBA promises to explain and specify precisely the content of visual experience.

Take Home Message 8: EBA can extend representationalism from dispassionate to "hot" cognition.

Paneelikeskustelu

Milloin on konetietoisuutta?

"very soon, ~25 years", "20 years + 20 to verify that we really have consciousness", "10 years", "10 + 10", "for some kind of consciousness, rather close: 10-20 years, human-like consciousness: I don't know if this is a goal we should aim for, we should aim for machine-like consciousness", "I hope in my lifetime, for it would be nice to see", "tempted to say I have no idea, but if I modify the question to 'when will we develop objects such that it'll be hard for people to avoid attributing phenomenal states for them', then very soon because people are easy to fool, 50-100 years, then for human-level, that's going to be a long time", "if I got enough money and we implemented the architecture in my second book into a robot, then I will not be modest - it would have qualia and everything in the best possible way".

Is consciousness a state or a process? Can it be quantitatively measured?

"It's a starting process for our science, not something that can be reduced to anything" "We constantly measure how conscious we are, so it probably can be measured. On the other hand, we constantly attribute consciousness to things that don't have it." "It's a process. There's a deep categorical mistake between 'conscious states' and 'mental processess, they're completely different - process refers to physical occurrances, state is an epistemical, complex term that we use to refer to certain values of possible situations." "As for measuring - we know very well how computer programs work, yet we can't deassemble the source code of Windows to improve it without Microsoft's permission. I think measuring the consciousness of the brain is even more difficult."

Has the concept of qualia enhanced or detracted from philosophy of mind?

"This term is heavily loaded with different premises, many of which would not be accepted by the people using the term. Yet the term has spread to many different fields. It was a useful term, which raised the consciousness of different people to the issues, so it was both positive and negative."

(Tässä kohtaa jouduin lähtemään.)