

# Two Related Themes

What are the functions of vision?

How did human language evolve?

(Languages are needed for internal information processing[\*])

**Aaron Sloman**

<http://www.cs.bham.ac.uk/~axs/>

[\*] Including learning, perceiving, thinking, wanting, deciding, intending, planning, wondering, ....

Khalid Khattak (a student on the course) kindly recorded a presentation in March 2015, related to these slides.

<http://www.cs.bham.ac.uk/research/projects/cogaff/movies/#ailect2-2015> (158MB)

These slides are available at

<http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#talk111>

Also available on my slideshare.net page:

<http://www.slideshare.net/asloman>

These slides are based on older sets of online slides in my talks directory:

<http://www.cs.bham.ac.uk/research/projects/cogaff/talks/>

This is also a sequel to a talk on AI and Philosophy:

<http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#talk109>

(All work in progress)

Part of the Meta-Morphogenesis project:

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/meta-morphogenesis.html>

# Evolution of language & Functions of vision

---

Most people think language is essentially concerned with **communication between individuals**.

So they ask the wrong questions about evolution of language, and give limited answers – concerned only with forms of communication.

A different view of language opens up more questions, requiring more complex and varied answers.

**A language is primarily a means by which information can be represented, for any purpose, including internal purposes such as learning, reasoning, formation of intentions and control of actions. That includes perceptual information, e.g. visual information.**

Instead of asking: **how did communication using language evolve?**

We can ask:

- For what purposes do organisms use information?  
Learning about the environment (e.g. through visual perception), control of actions, selection of goals, formation of plans, execution of plans, making predictions, asking questions, finding answers, communication with other individuals, social teaching and learning....(add your own ideas).
- What types of information do organisms need to acquire and use?
- In what forms (languages) can the information usefully be represented?
- What mechanisms are required for acquisition, storage and use of information?
- A special case: **How did languages also come to be used for communication?**

# Some key ideas

---

Animals need “internal languages” (**internal** representations/encodings of information) for purposes that are not normally thought of as linguistic.

E.g. perceiving, experiencing, having desires, forming questions, forming intentions, working out what to do, initiating and controlling actions, learning things about the environment (including other agents), remembering, imagining, theorising, designing .....

Without the use of richly structured internal languages, human vision, thought, learning, planning would be impossible.

There would be no need for communicative languages if individuals had nothing to communicate, and had no internal means of storing and using information communicated or acquired by perception or learning.

**So, having one or more internal languages is a prerequisite for using an external language.**(Sloman, 1978b, 1979)

**Internal languages (forms of representation) must therefore have evolved first, and must develop first in individuals:** later both can develop in parallel.

This requires a “generalised” notion of a language: a GL, and both internal GLs and external languages (ELs) require forms of representation that are **manipulable**, with additional properties, including:

- structural variability,
- varying complexity (e.g. for information about objects/events of varying complexity),
- compositional semantics (allowing new meanings to be assembled from simpler ones)
- used in perceiving, learning, recalling, reasoning, planning, controlling actions, .....

# Some important features of a language

---

What features of a language (external or internal) that give it its power?

- **Structural variability** in what is expressed and in the form of expression.
  - So that different sorts of things, of varying complexity can be described or represented, e.g.
    - a dot on a blank surface
    - a collection of dots
    - a collection of dots and lines moving in a plane
    - a plan of the furniture in a room
    - a plan for building a house
    - a generalisation about houses
    - a question or intention concerning houses
    - the reason why something will or will not have its intended effect
- **Compositional semantics** – complex meanings can be built up from simpler meanings. Linguistic expressions are “combinatorial” (composed of parts), NOT “wholistic” (unanalysable wholes). Meaningful components can be combined in different ways to express different things, including totally new things, e.g.: **I ate a yellow crocodile with red spots for breakfast yesterday.**
- **Use for expressing motives, preferences, goals, values, questions, hopes, etc.**
- **Use for reasoning** (predicting, explaining, hypothesising, planning),
  - by manipulating and recombining parts of complex representing structures.
  - So that you can derive new predictions, plans, summaries, generalisations, explanations, hypotheses, designs, maps, computer programs, etc. from old information.

**Illustrate with SHRDLU demo.** (Winograd, 1972)

The original: <http://hci.stanford.edu/winograd/shrdlu/>

For a non-interactive video of a simple Pop11 program based on Winograd’s SHRDLU running see demo

12 here: <http://www.cs.bham.ac.uk/research/projects/poplog/figs/simagent>

# Standard and context-sensitive compositional semantics

## Conventional compositional semantics:

New combinations of words, phrases, pictures, and other components of meaningful structures, are understood because the meaning of a whole is determined by **two things**:

- the meanings of the parts
- the way the parts are assembled to form the whole.

## However, that does not account for all uses of linguistic complexity.

The **non-linguistic context** in which information structures are combined also often helps to determine what the combination means for the user.

- Often it is not just **the internal structure**: the components of the representation and their relationships do not suffice to determine the semantics of a complex whole.
- Aspects of **the context both inside the user of the representation and in the external environment** are often important in determining what is expressed.

(This is obvious with indexicals such as “now”, “here”, “this”, “you”.)

- So the standard notion of compositional semantics does not account for all uses of representational complexity, unless we think of every syntactic construct as having an extra argument: the current context (which may or may not all be shared between speaker and hearer). (Compare the notion of a “rogator” in (Sloman, 1962, 1965).)

We’ll see later that this requirement for compositional semantics to use context applies to both linguistic complexity and visual complexity.

The idea can be generalised to include multiple uses of creative compositionality in biological evolution:

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/compositionality.html>

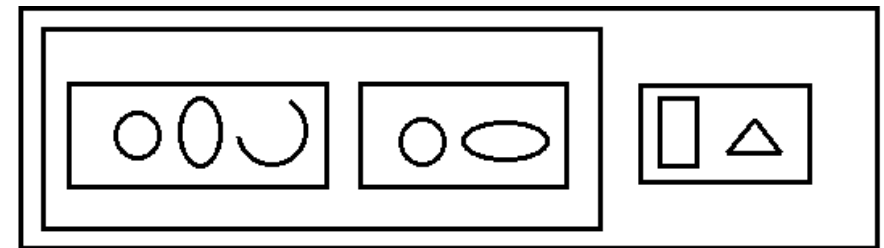
# A diagrammatic version

**Generalised (context-sensitive, situated)** compositional semantics can be depicted diagrammatically:

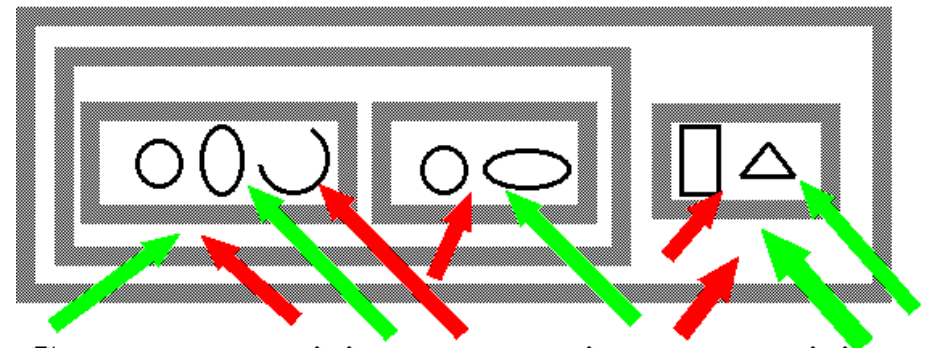
New combinations of words, phrases and other components are understood because the meaning of a whole is determined by

**three things:**

- the meanings of the parts
- the way the parts are assembled to form the whole
- linguistic and non-linguistic aspects of the context, including
  - the physical environment
  - the goals of the speaker and hearer
  - current tasks in progress ... and other things



**Context free semantic composition**



**Context sensitive semantic composition**

Formally, we can think of every syntactic construct (every box) as having extra arguments that enrich the interpretation: the current context and current goals (which may or may not all be shared between speaker and hearer). The contexts that enrich the semantics (green and red arrows) may come from inside the symbol user (including memory, motives, attitudes, etc.), or from the external physical or social environment.

These ideas extend Frege's *Sinn/Bedeutung* (sense-reference) distinction (Sloman, 1962, 1965) – making *Sinn* partly context sensitive.

# Generative grammars for information structures

Human languages have some features that are intricately related to the notion of a **generative grammar** – a grammar that allows novel sentences to be constructed.

Languages have meaningful units that can be combined in various ways,

e.g. words, phrases, particles (e.g. ‘ing’ (in ‘walking’), ‘ed’ (in ‘lifted’), ‘s’ (in ‘runs’)), and many more.

Languages use grammatical rules for combining meaningful units to form more complex meaningful units.

Some of the rules allow larger units to be joined to form still larger units:

He ran / down the road / toward the station / avoiding other people / nearly falling over / several times, / and thought / about being too late / to stop Jane / marrying / the station master / .....and so on, and so on...!

In principle, the rules of English allow arbitrarily long phrases and sentences, though human brains, like computers, have “performance” limits (Chomsky [http://en.wikipedia.org/wiki/Linguistic\\_competence](http://en.wikipedia.org/wiki/Linguistic_competence) ).

Intelligent agents need **internal meaningful units**, and something like **grammatical rules**, for expressing novel information contents of visual and other forms of perception, beliefs, intentions, wants, fears, and many more, even if they don’t use language to communicate.

For more about grammars (syntax), grammatical structures, structural ambiguities, varieties of semantic content, there is much online material, e.g.

[http://en.wikipedia.org/wiki/Syntax\\_%28logic%29](http://en.wikipedia.org/wiki/Syntax_%28logic%29)

Elementary demos using Pop-11 teaching libraries are here:

<http://www.cs.bham.ac.uk/research/projects/poplog/teach/grammar>

(Introductory overview.)

<http://www.cs.bham.ac.uk/research/projects/poplog/figs/simagent/#gblocks>

(A demo loosely based on the work of Winograd at MIT in 1971)

<http://www.cs.bham.ac.uk/research/projects/poplog/cas-ai/video-tutorials.html#haikus>

(A demo of a grammar used to generate and analyse Haikus.)

**“Grammars” for non-sentential structures are needed for visual contents.** (Kanef, 1970)

# Some background assumptions

---

Things we have learnt from AI research over the last 60 years or so include the following:

- An animal or machine acting in the world needs an information-processing **architecture** including different components capable of performing different sorts of tasks concurrently
- This is a **virtual-machine architecture**, not a physical architecture
- The various components need **information** of different sorts, which has to be **represented** or **encoded** in an appropriate way
- There is no one right form of representation: different tasks have different requirements, e.g.
  - for collections of symbolic facts (particular and general)
  - for structures representing spatial relationships e.g. maps, 3-D models
  - for algorithms that can be executed to produce internal or external behaviours
  - for doing statistical processing (e.g. building and using histograms)
  - for doing fast pattern recognition (using statistical or symbolic mechanisms)
  - for representing control information, including goals, preferences, partially executed plans, future intentions, etc.
- **SOME** information processing requirements **BUT NOT ALL** can be very dependent on the contents of the environment and on the body of the robot or animal (its structure, the materials used, etc.).
- Some animal information-processing architectures are mostly genetically determined, allowing only minor adaptations, whereas others are grown during individual development and are strongly influenced by interactions with the environment



# NOTE ON INFORMATION

---

The concept of “information” I am using is not Claude Shannon’s concept ([Shannon, 1948](#)), but the much older familiar concept of information used in ordinary everyday conversation, thoughts, reasoning, intending, questioning, planning, controlling, etc., explicitly referenced repeatedly in Jane Austen’s novels e.g. **Pride and Prejudice**. Her use of “information” refers not to the **medium** of thought or language but to the semantic **content** of thoughts, beliefs, intentions, questions, percepts, plans, hypotheses and many more.

For an introduction to Jane Austen’s (widely shared) *implicit* theory of information see

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/austen-info.html>

(By using the word “information” for his theory (perhaps under pressure from his employer The Bell Telephonen Company??), Shannon unintentionally confused a whole generation of researchers, although he was not confused himself.)

# How can information be used?

---

For an animal or robot to **use** information there are various requirements that need to be satisfied

- There must be ways of **acquiring** information e.g. through the genes, through the senses, by reasoning or hypothesis formation, or some combination  
NOTE: acquisition through senses is not a simple matter of recording sensory signals: A great deal of analysis, interpretation, abstraction, and combination with prior knowledge may be involved.
- There must be some way in which the information can be **encoded or represented** so that it can be used. This may be
  - **transient and used once**  
(e.g. information used in continuous control – in “online intelligence”)
  - **enduring and reusable**  
for short or long periods (offline intelligence) e.g.
    - \* percepts and immediate environment
    - \* generalisations,
    - \* geographical (extended spatial) information,
    - \* motives, preferences, values, ....
    - \* intended or predicted future events/actions, ..... etc.
- There must be mechanisms for **selecting** relevant information from large stores.
- The forms of representation used must allow **information-manipulations** that derive new information from old or construct new hypotheses or goals
- Some animals (e.g. humans) and some robots need ways of representing novel information about things never previously encountered.

# Generalising features of language

We can generalise the three features commonly thought to be core features of human language, as follows:

## A language with structural variability, compositional semantics and means of making inferences

(a) need not be composed of things we would recognise as words:

e.g. musical notations, circuit diagrams, maps, graphs, stick-figure drawings, chemical formulae, 3D models of molecules, computer programs, computer flow charts, interactive graphical design tools, and sub-neural chemistry (Grant, 2018);

(b) need not be used for communication:

e.g. it may be used entirely **inside** a perceiver, thinker, planner problem-solver, including uses for formulating goals, questions, hypotheses, plans, percepts, etc.

(c) need not be composed only of discrete units

(e.g. in maps and graphs).

Let's use the label "Generalised Language" (GL (Sloman & Chappell, 2007)) to refer to a form of expression or representation that has

- structural variability,
- compositional semantics
- means of making inferences,

and which is capable of being used for one or more information-processing purposes, communicative or non-communicative (e.g. control of action, or plan formation).

[You should now try to think about other examples you know about and invent some possible examples and check that they satisfy those conditions.]

# Evolution of language and language learning

---

## Common assumptions about language and its evolution

### It is commonly assumed that:

1. The primary or sole function of language is communication between individuals though there are derived mental functions such as planning, reminiscing, theorising, idly imagining...
2. Language initially evolved through primitive forms of communication, e.g.:
  - **Vocal communication** (grunts, roars, hisses, etc.) according to some theories
  - **Gestural communication** according to other theories (E.g.(Fadiga & Craighero, 2007));
3. Only after a rich **external** language had evolved did **internal** uses of language evolve; E.g. after language was used to communicate, it came to be used for self-communication, and then evolution produced short-cuts between brain mechanisms so that people could talk to themselves silently, instead of having to think aloud.  
Compare [http://en.wikipedia.org/wiki/Julian\\_Jaynes](http://en.wikipedia.org/wiki/Julian_Jaynes)

### Questions most theories of evolution of language don't answer:

What goes on inside individuals when they understand their external languages?

What goes on inside speakers who have something to communicate?

What goes on inside hearers when they understand or misunderstand, what is communicated?

# What makes it possible for a machine to understand?

**The designer stance** raises questions about how to design working perceivers, learners, actors, and communicators. (Compare Dennett & McCarthy.)

It challenges popular views of functions of language and evolution of language, raising new more subtle and complex questions – about how to make things work!

- What are the information-processing requirements for competences of pre-verbal children and intelligent animals that cannot use a human language?
- What are the the information processing requirements of language learning and understanding: Can heard sentences be understood without use of some internal means of representing information – internal language? Is there an infinite regress?
- Compare: How do computers understand

## (a) machine languages?

They refer to entities and operations in the computer, e.g. operations on ‘bit patterns’, represented by sets of on-off switches (usually transistors nowadays) interpreted as addresses, numbers, operations.

## (b) high-level languages?

- **Compiled languages:** are pre-translated into machine-languages before they run (batch vs incremental compiler)
- **JIT: Just in time compiler:** translates to machine-language at the last moment.
- **Interpreted languages:** instructions/rules invoke pre-built (possibly complex) programs.
- **Run-time extensions:** new programs created automatically in response to changing needs. (Downloading or creating new programs to use new hardware, e.g. camera, or remote service.)
- **Use of layered and concurrent virtual machines and external interfaces.** (See next slide)

How can a computer understand an instruction to send email to another machine, or to look for something on a remote machine? (An email address, or web-site address, refers beyond the boundaries of the machine, and can't be translated into machine code.)

Does any of this give us ideas about how brains might use internal languages?

# Use of virtual machines can extend semantic powers

States and processes exist at different “levels” in computers.

Since the 1940s new designs have allowed new kinds of states, processes and interactions.

- At very low levels there are physical (electronic) components: switches on or off, circuits connected or disconnected, voltages high or low, currents flowing or not flowing.
- Cleverly designed circuitry introduced **digital virtual machines**, manipulating bit patterns, on which bit operations can be performed (e.g. copying, modifying particular bits) and which themselves can be interpreted (by the machine) as “instructions” that cause operations to be performed (or control transferred) or as “addresses” that specify at which locations the operations should be performed, or as “binary numerals” representing numbers that can be added, multiplied, etc.
- Later, more and more abstract and diverse layers of information processing were added: multiple operating systems and virtual machines (some distributed across networks), including manipulation of text strings, list structures (binary trees), graphs, images, word-processors, game-players, proof checkers, theorem provers, virus checkers, networked file systems and user accounts, email systems, databases, internet services, human interface device managers, printers, cameras, robots, traffic lights, distributed financial services, flight controllers, chemical plant controllers, and many more.  
<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/vm-functionalism.html>
- Causal embedding of control mechanisms gave them primitive semantic powers, on which increasingly complex semantic contents were built, e.g. referring to remote mail servers, using information about remote customers and clients, checking and responding to status of many sub-systems. **["Symbol grounding" is not necessary for symbols used in “tethered” theories. (Sloman, 2007) (Also Kant.)]**
- **Human engineers took a handful of decades: evolution had billions of years.** There seem to be far more complex and diverse layers of virtual machinery built on physics and chemistry in brains – most of it still not understood, yet clearly acquiring, storing, manipulating, deriving, and using information of many kinds – and in some cases communicating information, e.g. to potential mates.

**High level semantics cannot be translated into machine semantics, e.g. when control loops involve environments.** (A fact that's often ignored by AI researchers and their critics, e.g. (Searle, 1984).)

# Evolution: the great AI designer/engineer

---

Close observation of achievements of pre-verbal children and intelligent animals who lack **external** languages used for communication suggests that **internal** languages express structured information used in perceiving, learning, wanting, controlling, reasoning, being puzzled, inferring, wondering why, and understanding some social interactions.

If these internal information structures are available before **external** languages in young humans, perhaps the internal mechanisms evolved before the external languages with generative grammars and compositional semantics used for communication.

What many animals achieve would be impossible if they started with no abilities to store and manipulate structured information.

Likewise children would not learn to communicate using highly structured languages.

Even some insects seem to use structured semantic contents in controlling their actions, e.g. finding routes, coping with obstacles when transporting food to the nest, fighting, etc.

## Warning:

Some AI theorists, philosophers, and others, claim that there's no need to postulate any semantically rich **internal** information structures to explain human and animal competences because all the behaviours somehow emerge out of ways in which complex physical bodies and their environments interact – like a ball finding its way to the bottom of a helter-skelter. Some say: “The world is the best representation of itself”.

But their demonstrations using robots don't come near the intelligence of human toddlers or nest-building by corvids, or carnivorous mammals hunting and eating larger grazing mammals – or even ant intelligence.

Neither, so far, do robots using internal languages! But they do have more types of intelligence than purely reactive robots using gravity, compliance, and physical constraints, like “passive walker” robots.

See also: <http://www.cs.bham.ac.uk/research/projects/cogaff/misc/chewing-test.html>

# Three views about evolution of human language

**Theory 1** First there were expressive noises (e.g. grunts) which gradually became more differentiated and elaborate and then were “internalised”.

Only after that did thinking, planning, reasoning, hypothesising, goal formation, become possible.



# Three views about evolution of human language

**Theory 1** First there were expressive noises (e.g. grunts) which gradually became more differentiated and elaborate and then were “internalised”.

**Theory 2** First there were expressive **gestures**, then noises, then as in 1.

Only after that did thinking, planning, reasoning, hypothesising, goal formation, become possible.

# Three views about evolution of human language

**Theory 1** First there were expressive noises (e.g. grunts) which gradually became more differentiated and elaborate and then were “internalised”.

**Theory 2** First there were expressive gestures, then noises, then as in 1.

**Theory 3** First there were **internal representations** used for perceiving, thinking, forming goals, forming questions, planning, controlling actions; then later, external forms developed for communicating meanings.

Two rival sub-theories:

3(a) Externalisation was first **gestural**

3(b) Externalisation was first **vocal**

**NB: Do not assume such internal representations must be some fixed, innate, genetically determined form of representation – a “language of thought” (LOT):  
New forms can emerge through learning and development.**

# Three views about evolution of human language

**Theory 1** First there were expressive noises (e.g. grunts) which gradually became more differentiated and elaborate and then were “internalised”.

**Theory 2** First there were expressive gestures, then noises, then as in 1.

**Theory 3** First there were internal representations used for perceiving, thinking, forming goals, forming questions, planning, controlling actions; then later, external forms developed for communicating meanings.

Two rival sub-theories:

3(a) Externalisation was first **gestural**

3(b) Externalisation was first **vocal**

All the above options allow for the possibility that the existence of **external** languages and cultures produced evolutionary and developmental pressures that caused **internal** languages to acquire new functions and more complex forms.

# Three views about evolution of human language

**Theory 1** First there were expressive noises which gradually became more differentiated and elaborate and then were “internalised”.

**Theory 2** First there were expressive gestures, then noises, then as in 1.

**Theory 3** First there were internal representations used for perceiving, thinking, forming goals, forming questions, planning, controlling actions; then later, external forms developed for communicating meanings.

Two rival sub-theories:

3(a) Externalisation was first **gestural**

3(b) Externalisation was first **vocal**

All the above options allow for the possibility that the existence of external languages and cultures produced evolutionary and developmental pressures that caused internal languages to acquire new functions and more complex forms.

Our question is: what evolved first:

- **external human languages**? (Theory 1/Theory 2)
- **internal languages** with core properties of human language? (Theory 3)

# Three views about evolution of human language

**Theory 1** First there were expressive noises which gradually became more differentiated and elaborate and then were “internalised”.

**Theory 2** First there were expressive gestures, then noises, then as in 1.

**Theory 3** First there were internal representations used for perceiving, thinking, forming goals, forming questions, planning, controlling actions; then later, external forms developed for communicating meanings.

Two rival sub-theories:

3(a) Externalisation was first **gestural**

3(b) Externalisation was first **vocal**

All the above options allow for the possibility that the existence of external languages and cultures produced evolutionary and developmental pressures that caused internal languages to acquire new functions and more complex forms.

Our question is: what evolved first:

- **external human languages**? (Theory 1/Theory 2)
- **internal languages** with core properties of human language? (Theory 3)

A similar question about what comes first can be asked about **individual** development.

# Three views about evolution of human language

**Theory 1** First there were expressive noises which gradually became more differentiated and elaborate and then were “internalised”.

**Theory 2** First there were expressive gestures, then noises, then as in 1.

**Theory 3** First there were internal representations used for perceiving, thinking, forming goals, forming questions, planning, controlling actions; then later, external forms developed for communicating meanings.

Two rival sub-theories:

3(a) Externalisation was first **gestural**

3(b) Externalisation was first **vocal**

All the above options allow for the possibility that the existence of external languages and cultures produced evolutionary and developmental pressures that caused internal languages to acquire new functions and more complex forms.

Our question is: what evolved first:

- **external human languages**? (Theory 1/Theory 2)
- **internal languages** with core properties of human language? (Theory 3)

**WHAT CORE PROPERTIES?**

# How most(?) people think about human language

- It is essentially a means of communication between separate individuals, though there are “derived solo functions” such as planning, reminiscing, wondering why, theorising, learning, idly imagining, keeping a diary, dreaming, and many others.
- It is essentially vocal, though there are secondary means of expression including sign languages, writing, specialised signalling systems (e.g. morse code, semaphor), ...
- It (mostly) uses a discrete linear medium, though it can encode non-linear information-structures, e.g. trees, graphs.
- Each language has a syntax with unbounded generative power, and compositional semantics (plus exceptions and special cases).
- It evolved from primitive to complex communication, and was later “internalised”.
- Individual humans acquire linguistic competence by finding out what languages are used in their environment and somehow extracting information about the rules, vocabulary, and ontology, in a usable form.

The acquisition process

- EITHER uses a specialised innate “language acquisition device” LAD (Chomsky),
- OR uses general learning mechanisms and general intelligence (the current majority view??)

- Only humans have linguistic abilities naturally, though there are some other animals that can, under very special circumstances, be trained to use a tiny restricted subset.

## **We introduce a more general concept: Generalised-Language (GL).**

Human communicative language is a special subset.

Pre-existing internal GLs are required before human languages can exist.

## **This challenges most of the points above.**

# An alternative theory of how language evolved

Languages first evolved for **internal** use, not in the form in which we now know language, but with some of the key features of human languages, e.g. compositionality.

- As a side-effect of this internal use, some complex actions based on complex intentions and plans, naturally became means of communication, intentional and unintentional.

## **Observable actions can communicate intentions and plans unintentionally.**

- Based on this, humans gradually evolved a sophisticated sign language capability  
Unintended communicative effects become intended effects, e.g. “showing how”.  
(Compare what intelligent carnivorous animals do to train their offspring?)
- Later, spoken (and written) language took over for most people, but not all.  
Many use only sign languages (and written languages), e.g. communities of deaf people.
- **But the evolutionary heritage of gestural language remains with all of us.**  
Most human children are capable of acquiring a sign language, even if they are not deaf.  
Some Down Syndrome children struggle learning to talk, but learn a sign language more easily.  
Most people can't help gesturing while talking, even talking to someone out of sight, e.g. using a phone.
- After languages for communication developed, that helped to accelerate the development of many human competences, partly through cultural and social evolution (e.g. new educational practices) and partly through development of brain mechanisms suited to learning from other humans.  
What was previously learnt slowly is learnt much faster from a teacher.  
This accelerates evolution!

**What are the main features of languages that give them their powers?**

This presentation gives only a partial answer.



# Temporary change of topic: to vision

---

The questions about evolution of human language need to be related to questions about evolution of vision and the sorts of information processing mechanisms required to support the functions of vision, both in organisms and in future intelligent robots.

This includes identifying some of the forms of internal language required for visual processing.

# Some questions about vision

---

**Some questions to think about (not all answered in these slides).**

What is vision?

What are the uses/functions of vision?

What sorts of information contents can vision have?

What forms of representation are required to express that information?

How can visual capabilities and mechanisms be products of evolution?

How many different sorts of vision are there, in animals and machines?

How do the contents and representational requirements change during individual development?

How do the powers and uses of vision change during individual development?

How do environments, cultures, educational procedures, affect those changes?

# Vision research problems

---

- Many researchers assume that it is obvious what vision is for, i.e. what functions it has, leaving only the problem of explaining how those functions are fulfilled.
- So they postulate mechanisms and try to show how those mechanisms can produce the required effects, and also, in some cases, try to show that those postulated mechanisms exist in humans and other animals and perform the postulated functions.
- James Gibson's main achievement was to draw attention to functions of perception that most other researchers had ignored.
- I'll present some of the earlier work, show how Gibson extended and improved it, and then point out how much more there is to the functions of vision and other forms of perception than even Gibson had noticed.
- Many vision researchers, unlike Gibson, ignore vision's functions in **on-line** control, and perception of continuous processes
- Most, including Gibson, ignore perception of **possibilities** and **constraints on possibilities** required for use of vision in reasoning and mathematical discovery, i.e. **off-line** intelligence. (But I may have missed something in Gibson's writings!)
- Many ignore or over-simplify meta-cognitive perception: perception of other minds.

**The functions of vision require use of internal languages. WHY?  
What does this imply for research in AI/Intelligent robotics?**

# High level outline 1

---

(If you are not interested in this history, jump to the slide headed: “Marr 3: functions of vision”)

## A short, incomplete, illustrative, history of theories of functions and mechanisms of vision

- Previous millennia, E.g. Aristotle
- Previous centuries: lots of philosophers, including Hume, Berkeley, Kant, Russell, and non-philosophers, e.g. (non-philosopher? polymath?) von Helmholtz.

*Helmholtz proposes a “sign” theory, according to which sensations symbolize their stimuli, but are not direct copies of those stimuli. While Müller explains the correspondence between sensation and object by means of an innate configuration of sense nerves, Helmholtz argues that we construct that correspondence by means of a series of learned, “unconscious inferences.”*

<http://plato.stanford.edu/entries/hermann-helmholtz>

- AI Vision work – 1960s on

Lots of work on finding and representing structure in images (e.g. lines, junctions, and also higher order, possibly non-continuous structures, e.g. views of occluded objects). (E.g. A. Guzman)

Also work on trying to find 3-D interpretations, using 3-D models (e.g. Roberts, Grape), or using constraint propagation from image details, e.g. Huffman, Clowes, Winston, Waltz, ...)

Barrow and Tenenbaum, on “intrinsic images” ([Barrow & Tenenbaum, 1978](#)) (image fragments (syntax) denote fragments of scenes (semantics) in a systematic way: an idea revived recently).

Use of soft constraints/relaxation, and neural net inspired mechanisms e.g. ([Hinton, 1976](#)).

Many ideological battles between researchers whose systems were **all** very limited, compared with animal vision.

- Nearly all researchers assumed that the **functions** of vision were obvious and only mechanisms/explanations were in dispute. Gibson shattered that.

# High level outline 2: pre-Marr – 1960s onwards

## Lots of image analysis routines

e.g. Azriel Rosenfeld: Many of the algorithms just transformed images, e.g. showing edges.

## Ideas about pictures having structure, often inspired by Chomsky's work on language

E.g. **S. Kaneff** (editor) *Picture language machines* (1970), and

**Max Clowes** <http://www.cs.bham.ac.uk/research/projects/cogaff/sloman-clowestribute.html#bio>

## Analysis by synthesis/Hierarchical synthesis

Ulric Neisser *Cognitive Psychology*, 1967 (parallel top-down and bottom up processing using models)

Oliver Selfridge, PANDEMONIUM (partly neurally inspired?)

## Model-based vision research on polyhedra and other 3-D structures

Roberts, Grape, (particular polyhedral models, e.g. wedge, block, etc.);

Clowes, Huffman, (model fragments: faces, edges, vertices);

Later work - Marr/Nishihara generalised cylinders; Biederman geons

**NB** More recent, more systematic work on polyhedra by Ralph Martin's group (Cardiff).

<http://ralph.cs.cf.ac.uk/Data/Sketch.html>

## Use of “expert systems” techniques to analyse pictures Hanson and Riseman (UMASS)

## Finding structure in images from:

**Stereo** (many people – badly influenced by results from random dot stereograms (Julesz))

**Motion** (Longuet-Higgins, Clocksin, Ullman, Spacek ....)

**Intensity/shading** (Horn, ...)

**Texture and optical flow** (Gibson)

## More general relations between image fragments and scene fragments

Barrow and Tennenbaum “Recovering intrinsic scene characteristics from images” (1978)

## Parallel work on 2D pattern recognition (disparaged by AI people who emphasised 3D)

# High level outline 3: Marr (a)

---

David Marr: papers at MIT in late 1970s, died 1981, posthumous book (Marr, 1982)

Unfortunately, convinced many people that vision is (merely?):

a process of producing descriptions of 3-D shape, size, distance, orientation, etc, from 2-D data  
i.e. “reversing” the production of an image by projection.

Marr stressed three levels of theory (causing much confusion, in my view):

(1) Computational, (2) Algorithmic, (3) Implementational.

I suggest this is mainly a confused way of introducing the old engineering distinctions:

(1) Requirements analysis (What’s X for?)

(2) Design (What are the high level design features for systems that meet the requirements?)

(3) Implementation

(How to implement the high level design in low level mechanisms,  
physical, electronic, physiological, or computational.)

**NB:** All of those can have different levels – e.g. implementation in virtual machines,  
implemented in lower level virtual machines, .... implemented in physical machines.

I.e. the same three levels can recur for each level of implementation, e.g. for transistors.

Marr’s Levels were badly named and far too widely accepted as important, by people  
without engineering experience.

For a critique of the three levels by McClamrock (1991) see

<http://www.albany.edu/~ron/papers/marrlev1.html>

# High level outline 3: Marr (b)

---

## Some of Marr's main influential points:

- Reject artificial images, e.g. line drawings – for use as test images (too impoverished):
  - use natural images (actual photographs of real 3-D objects)
  - rich in data, so making tasks easier (??) [Illustrated using teddy bear photo]  
He discounted the possibility of informed selection of artificial images to study well-defined problems.
- Processing pipeline: primal sketch  $\Rightarrow$  2.5D sketch  $\Rightarrow$  3-D interpretations
- Use of generalised cylinders (Compare Biederman's geons)  
(But generalised cylinders proved unsuitable as models for many objects.)
- The function of vision is to produce descriptions/representations of **what's out there**:  
3-D geometry, distance, surface orientations, curvature, textures, spatial relationships, colours(?).
- **He shared the common assumption that metrical (Euclidean) coordinate frames are required.**

But he allowed that different frames of reference can be used for scene descriptions

- Scene centred (Use a global coordinate system for everything visible in the scene)
- Object centred (Attach coordinate systems to objects, or object parts)
- **Viewer centred**
  - o **Egocentric** (Represent scene objects in terms of relationships with perceiver).
  - o **Allocentric** (Represent scene objects in terms of relationships with another viewer).  
(I am not sure Marr made this distinction. Others have.)

# High level outline 3: Marr (c): functions of vision

Marr wrote: “... the quintessential fact of human vision – that it tells about shape and space and spatial arrangement”.

## Comments:

(a) This ignores the functions of vision in on-line control of behaviours:

e.g. visual servoing while grasping, moving, avoiding obstacles –  
or assumes (wrongly) that these control functions can all be subsumed under the descriptive functions.

Compare: (J. J. Gibson, 1966) (Sloman, 1983)

(b) This ignores many of the social (meta-cognitive) functions of vision, e.g. discovering what other people are attending to, how they feel about it, what they intend to do, and many more.

(c) It ignores many of the functions of vision concerned with discovering what is and is not possible in a situation, which links up closely with mathematical discoveries in geometry and topology (discussed later).

(d) The emphasis on spatial structure seems to ignore perception of colour, or material (seeing something as liquid, or fragile), and perception of causation: seeing X break Y, support Y, lift Y, etc. (Michotte, 1962))

(e) Marr’s view fits the common idea of vision as “reversing” the projection process – using information in the optic array (or image) to construct a 3-D model of what’s visible in the scene.

But that common idea is mistaken: visual systems do not represent information about 3-D structure in a 3-D model (information structure isomorphic with things represented – like computer 3D stereo models)

but in **a collection of information fragments, all giving partial information.**

**A collection of partial descriptions of a structure can be inconsistent, whereas a 3-D model cannot be inconsistent.**

This point is illustrated below, with pictures of impossible 3-D objects.

The way we see such pictures would be impossible if Marr-like theories of 3-D perception were correct: such theories are very tempting and very popular – but wrong!



# Moving beyond Marr and predecessors

---

Marr:

“... the quintessential fact of human vision – that it tells about shape and space and spatial arrangement”.

**What else could there be, besides shape and space and spatial arrangement?**

**Lots!**

Gibson noted some of it.

(Starting in the 1960s ([J. J. Gibson, 1966](#)).)

There is a useful but brief discussion of Gibson’s ideas and how they relate to AI in section 7.v of ([Boden, 2006](#)), pp. 465–472.

Some philosophers who are ignorant of that work by Gibson seem to think the ideas came up much later, when philosophers and others started discussing embodied cognition, partly inspired by the work of Brooks, some time after Gibson’s first book, e.g. ([Brooks, 1990, 1991](#))

Some of the ideas about embodied cognition were also in ([Simon, 1969](#)), though he (rightly) treated those examples as special cases.

## Marr (d): A hint of a move towards Gibson's ideas.

On p.31 of **Vision**, Marr wrote “Vision is a process that produces from images of the external world a description that is useful to the viewer and not cluttered with irrelevant information.”

Not cluttered? My visual contents often are!

(That's Gibsonian – except that Gibson would not talk of descriptions or representations.)

Gibson believed in a kind of “resonance” produced by the environment.

But for information to be usable it must be encoded or represented in some structured form in some medium – an internal language, a GL.

In computers the medium is often a non-physical virtual machine: and brains probably use those too, but far more sophisticated forms. (Sloman, 2013)

Marr also wrote:

p. 32 “Vision is used in such a bewildering variety of ways that the visual systems of animals must differ significantly from one another”.

“For a fly the information obtained is mainly subjective...”

(Mainly concerned with image contents?? Or something like affordances for the fly??)

If what different animals see when looking at the same things constitute a “bewildering variety”, then perhaps their visual systems are not all attempting to acquire the same information, namely information about the structure of the environment.

**Evolution produced something far more subtle, and more useful.**

Moreover, animals need expressive internal languages to store the information acquired through vision and other senses: not just labelling or 3-D model-building mechanisms.

# James Gibson started a revolution

---

Gibson noticed other abstract features of visual contents: especially features related to possible actions by the perceiver

*The Senses Considered as Perceptual Systems*, (J. J. Gibson, 1966)

*The Ecological Approach to Visual Perception*, (J. J. Gibson, 1979).

He claimed that for organisms, the function of vision (more generally perception)

- is **not** to describe some objective external reality, and **not** to attach labels to sensory data
- but to **serve biological needs**

in different ways for different organisms.

The different functions of vision arose at different stages in biological evolution, as (a) the physical environment became more complex and more structured, and (b) the organisms and their needs, shelters, predators, food, mates, and offspring became more complex.

Some of the consequences were physical (and physiological) –

e.g. development of independently movable manipulators: jaws, hands, tongue, etc.

That had implications for information processing requirements.

Not just “What am I experiencing?” or “What’s out there?”, but

“What can I, can’t I grasp, pick up, move through, climb over, push out of the way, eat, drink,” ....etc.?

(Using an **exosomatic** rather than a **somatic** ontology, referring to **environment**, not **sensory signals**.)

Somatic = Marr’s ‘subjective’ ?? (see previous slide)

Exo-somatic = Marr’s ‘objective’ ??

Gibson’s view is often incorrectly summarised by the claim that seeing affordances is seeing **what objects are for** – as if affordances came from intentions of object designers.

But most affordances are not intended: even designed objects can have unintended affordances.

# Gibson's Revolution (continued)

---

For organisms the function of vision (more generally perception) is not to describe some objective external reality but to serve biological needs

- Providing information about positive and negative affordances (what the animal can and cannot do in a situation, given its body, motor capabilities, and possible needs).
- Using invariants in static and changing optic arrays  
texture gradients, optical flow patterns, various rates of change, contrast edges, “common fate”.
- Using actions to probe the environment, e.g. so as to change contents of the optic array  
The sensors and effectors work together to form “perceptual systems” (J. J. Gibson, 1966)  
(compare “active vision”)

Some of Gibson's ideas are well substantiated, but not all.

## **A brilliant idea:**

Visual processing does not start with retinal image information, but with an optic array, whose changes are systematically coupled to various kinds of actions:

**I.e. the retina is mainly a device for sampling optic arrays: cones of information converging on viewpoints.**

(In primates, it does that in close cooperation with brain area V1, at rear of brain.)

## **The implausible claims in Gibson's theories:**

- There are no internal representations involved.
- Perception is immediate and direct, there is no reasoning, or inference.  
 (“pickup”, not “interpretation” – no computation, no representation: suggests a kind of magic. I suspect he merely intended to deny bad philosophical theories about reasoning from “sense-data”.)

# Moving Beyond Marr and Gibson!

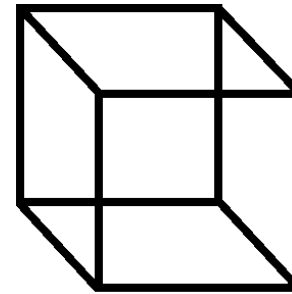
Are visual contents spatial? Yes – but only **some** of them.

Contrasting two sorts of ambiguous image makes this clear:

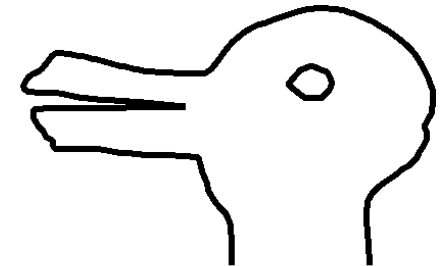
What changes in what you see when an ambiguous image “flips”?

In **some cases** it is shape, and space and spatial arrangement, as Marr claimed.

But not in all cases (Sloman, 1983, 2001):



Necker Cube



Duck-rabbit

Everything that changes when the Necker cube “flips” is spatial:

orderings (what’s nearer, further, higher, lower), directions (sloping away up, sloping away down), relative distances and orientations e.g. of the top and bottom surfaces.

In the duck-rabbit, however, there is no geometric flip:

The changes are much more abstract and involve changes in how parts are identified (e.g. **ears** vs **bill**) and also more abstract changes like “**facing this way**”, “**facing that way**”, which presupposes that *other organisms can be seen as perceivers (information-users) and potential movers*.

That’s the basis of **vicarious** affordances (e.g. seeing what your children should not do) and **social** affordances (e.g. seeing an opportunity to win friends, or help someone, included in (J. J. Gibson, 1979).)

We see some things as more than physical objects: e.g. we can see a rabbit **looking, trying, wanting**, etc.

That’s **perception**, not just **inferred cognition**, because perceptual details are represented in registration with the optic array, e.g. **seeing happiness in someone’s eyes**. (What sort of inner language is needed for that?)

The fact that human visual systems can learn to read text, music, mathematical formulae, blueprints, ..., shows how far **visual contents** extend beyond the physical environment.

# Vision and mathematics

---

Most vision researchers, including Gibson, ignore the roles of vision in mathematical discovery, e.g. the discoveries thousands of years ago that eventually led to Euclid's Elements.

<http://www.gutenberg.org/ebooks/21076>

Some examples are provided in:

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/triangle-theorem.html>

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/trisect.html>

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/torus.html>

It can be argued that those mathematical visual functions are closely related to abilities to perceive affordances involving possibilities for change and constraints on possibilities for change, in the environment.

# Problems with Gibson's theory

Although he continually emphasised **information**, e.g. the information available in static and changing optic arrays, Gibson denied the need for representations or information processing (computation), proposing a mysterious concept of “direct pickup” instead.

He provided many important insights regarding interactions between vision and action, and the immediately usable information in vision, but ignored other roles for vision e.g.

- multi-step planning,

I can see that if I move the pile of bricks to the right, and push that chair against the bookcase, and stand on it, I'll be able to reach the top shelf. (Here seeing feeds information into reasoning processes.)

- seeking explanations

Can marks on the road and features of the impact suggest why the car crashed into the lamp post?

- understanding causal linkages that could be used

IF this string is pulled down, THEN that pulley will rotate clockwise, causing that gear to turn, and ....

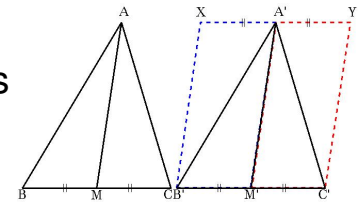
(Contrast *immediate* perception of causation, e.g. X bumps into Y causing it to move (Michotte, 1962))

- geometric reasoning

The line from any vertex of any triangle to the middle of the opposite side produces two smaller triangles of the same area, even when the shapes are different. Why?

- Design of new machines, tools, functional artifacts (e.g. door-handles).

- Perceiving intentional actions. Fred is looking for something.



Contrast Eleanor J. Gibson and Anne D. Pick,

*An Ecological Approach to Perceptual Learning and Development*, OUP, 2000.

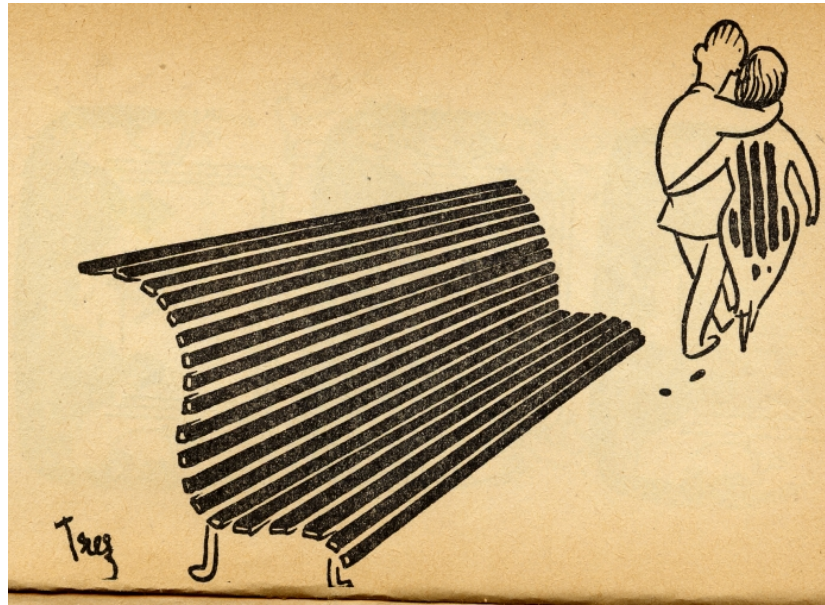
They allow for the development of a wider range of cognitive competences using vision.

But even they don't allow for learning by **working things out**. (E. J. Gibson & Pick, 2000)



# Example: Perceiving causation

Often our ability to perceive causal connections is used in humour:



From a book of “French Cartoons”.

Some cartoons depend on our ability to see processes of various sorts.

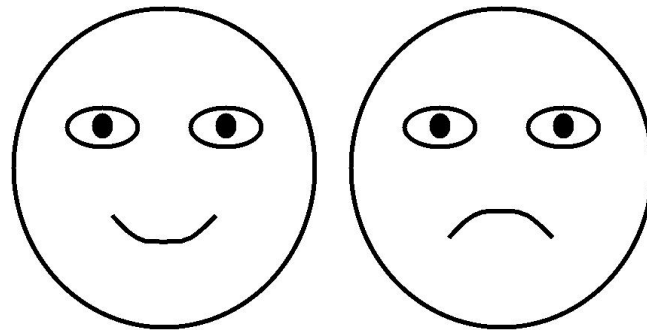
In the above it's a **past** process reaching into the present.

Some cartoons present a **future** process extending from the present – e.g. a pompous pontificator heading unwittingly for a fall, collision, come-uppance, etc.

Help needed: I can't track down the publisher or cartoonist to ask for permission.



# Visual contents can be very abstract, non-physical



Look at each face in turn for a few seconds  
Do the eyes look different? Why? How?  
Compare illusory contours – Kanizsa

[http://en.wikipedia.org/wiki/Illusory\\_contours](http://en.wikipedia.org/wiki/Illusory_contours)

The eyes above are geometrically identical. Some people (not all) see them as different: happy and sad.

## Problems

- What sorts of non-shape information can a visual system get from the environment? (intention, mood, emotion, effort... compare Johansson movies ([Johansson, 1973](#)))
- Why are these kinds of information useful – to what sorts of agents?
- How is the information represented in the perceiver? What information, exactly?
- What mechanisms derive it from what information?
- What are the roles of [evolution](#) and [development/learning](#) in the creation/modification of such mechanisms?
- How do these mechanisms and forms of representation interact with others?
- How do the interactions help a perceiver, or a group of individuals?

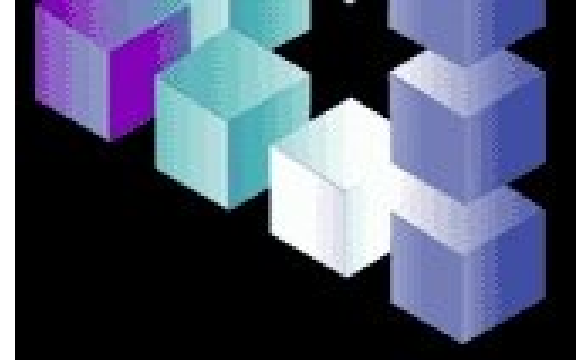
# 3-D structures – and possible actions

On the top right is part of a picture by Swedish artist, Oscar Reutersvärd (1934) which you probably see as a configuration of coloured cubes.

As with the Necker cube you have experiences of both 2-D lines, regions, colours, relationships and also 3-D surfaces, edges, corners, and spatial relationships.

You probably also perceive various affordances: places you could touch the surfaces, ways you could grasp and move the various cubes (perhaps some are held floating in place by magnetic fields).

E.g. you can probably imagine swapping two of them, thinking about how you would have to grasp them in the process – e.g. using left and right hands to swap the white cube with the one on its left, or the one on its right.



## 3-D structures – and possible actions

On the top right is part of a picture by Swedish artist, Oscar Reutersvärd (1934) which you probably see as a configuration of coloured cubes.



As with the Necker cube you have experiences of both 2-D lines, regions, colours, relationships and also 3-D surfaces, edges, corners, and spatial relationships.

You probably also perceive various affordances: places you could touch the surfaces, ways you could grasp and move the various cubes (perhaps some are held floating in place by magnetic fields).

E.g. you can probably imagine swapping two of them, thinking about how you would have to grasp them in the process – e.g. using left and right hands to swap the white cube with the one on its left, or the one on its right.

The second picture on the right (from which the first one was extracted) has a richer set of 2-D and 3-D contents.

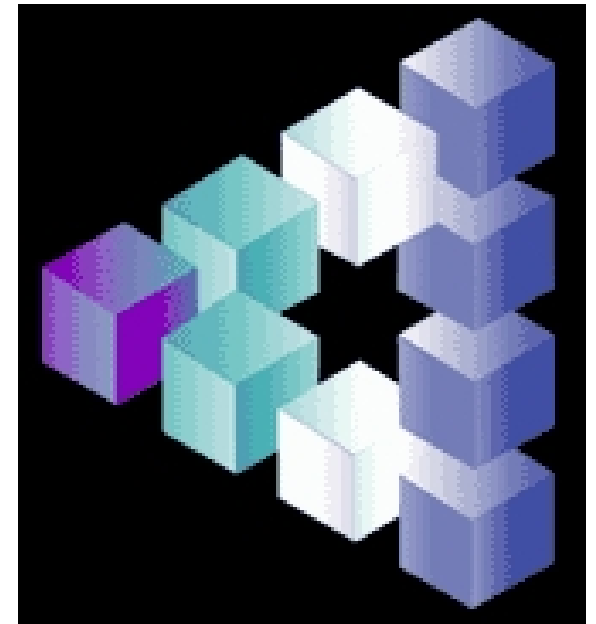
Again there is a collection of 2-D contents (e.g. a star in the middle), plus experience of 3-D structures, relationships and affordances: with new possibilities for touching surfaces, grasping cubes, moving cubes.

The picture is outside you, as would the cubes be if it were not a picture. But the contents of your experience are represented in you: a multi-layered set of visual “qualia”: experienced 2-D and 3-D structures and also process possibilities and constraints.

The scene depicted in the lower picture is geometrically impossible, even though the 2-D configuration is possible and exists, on the screen or on paper, if printed. **The cubes, however, could not exist as shown.**

**Your visual experience can be of something impossible**

though the experience is not impossible. Can we give machines such experiences? How?



# A consequence of the core features of internal/external languages

---

A consequence of the core features is that it is possible to produce well-formed linguistic expressions for which the compositional semantics will produce an impossible (internally inconsistent) interpretation.

E.g. Consider this conjunction

Tom is taller than Mary  
and Mary is taller than Jane  
and Jane is taller than Dick  
and Dick is taller than Tom

If

- (a) 'Taller than' has its normal meaning
- (b) Each repeated occurrence of the same name refers to only one individual

then

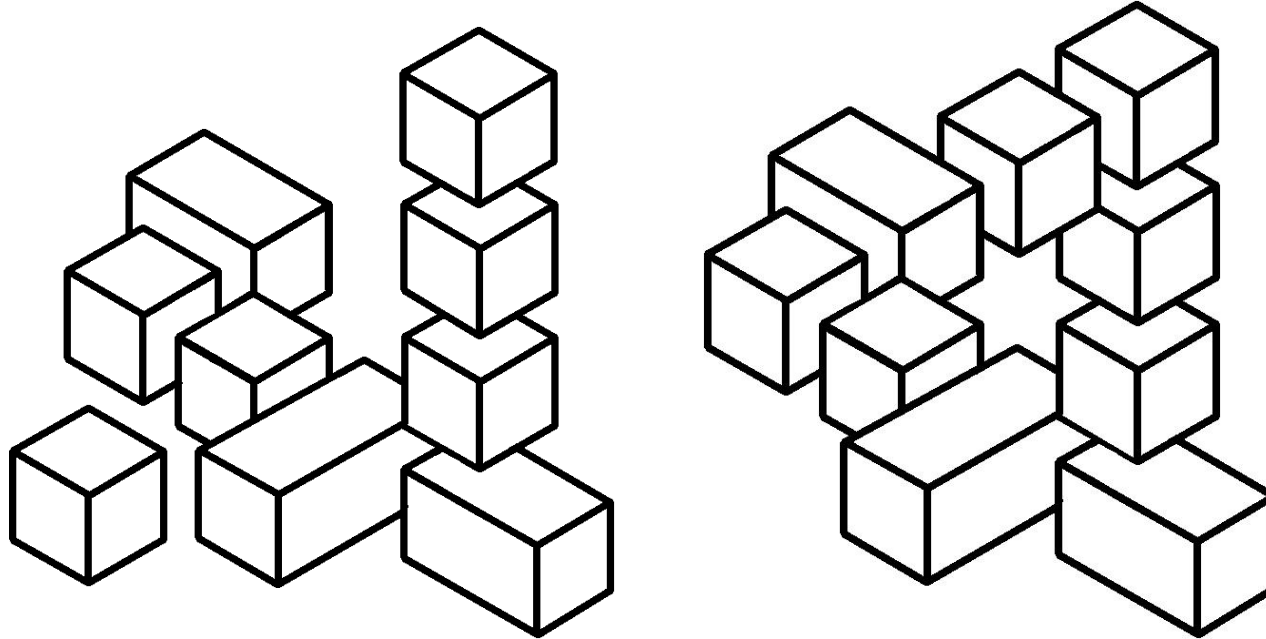
That conjunction is inconsistent:  
not all conjuncts can be true simultaneously.

We have also seen a similar kind of inconsistency in non-verbal forms, in the previous slide: Some well formed pictures made by assembling well formed parts cannot depict something that exists, even though all the parts can, and all the subsets obtained by removing one block can. **What does that tell us about the language used by the visual system?**

Can you create a new picture with that property? Could a future robot? How?

# Building a configuration of blocks - 1

Could a collection of cubes and rectangular blocks be arranged as in the group on the left?



What has changed between the scene on the left and the scene on the right?

Could both scenes be constructed from rectangular blocks?

Some young children will say 'yes'.

What has to change in a child to enable her to see which configuration is impossible?

Could the impossibility be detected by a child who has not yet learnt to talk?

# Disentangling complex scenes

Example used by Max Clowes in 1973:  
Image from the movie “Sunday Bloody Sunday”.

Can you see four hands?

Can you tell which hand belongs to whom?

How?

Knowledge-based image interpretation differs from use of general principles relating 2-D and 3-D structures.

Much perception is knowledge-based.

(Knowledge of human anatomy here.)

(Abercrombie, 1960; Clowes, 1973)



What sort of internal language is required to express what you see in a picture like that?  
How much of it would be shared with other animals, or with 2 month old humans?

Can Gibson’s theories accommodate perception of scenes like this?

If understanding the scene requires perceiving invisible connections between body-parts does that refute Marr’s claims about “reversing the projection process” ?

**Exercise:** Is there a change that would make this a picture of an impossible scene?

# Implications of pictures of impossible objects

---

The impossible pictures rule out the assumption that seeing involves building a structure that is isomorphic with what is seen: for it is impossible to build a structure that is isomorphic with an impossible structure.

What we (and other animals?) do must be much more subtle, general and powerful, and connected with manipulability, structural variation, and compositional semantics, all of which are important in seeing affordances.

The example of logic shows that it is possible to assemble coherent fragments of information into an incoherent whole: this seems also to be what happens when we see pictures of impossible objects, though in that case we do not seem to be using a logical formalism.

Exactly what sort of GL suffices for the purpose requires further research,

- We need to analyse requirements for GLs, including both being usable for representing what exists and being usable for representing and reasoning about changes that are possible

- We seem to use those features of GLs in understanding many examples of causation.

Fortunately we don't normally need to check for consistency because the 3-D environment cannot be inconsistent.

See also <http://www.cs.bham.ac.uk/research/projects/cogaff/challenge-penrose.pdf>

NOTE: all of this leaves many questions unanswered about what other animals can and cannot see and what sorts of internal languages their visual systems need.

# There are many sorts of things humans can see besides geometrical properties:

---

- that one object is supported by another,
  - that one object constrains motion of another (e.g. a window catch),
  - that something is flexible or fragile,
  - which parts of an organism are ears, eyes, mouth, bill, etc.,
  - which way something is facing,
  - what action some person or animal is about to perform (throw, jump, run, etc.),
  - whether an action is dangerous,
  - whether someone is happy, sad, angry, etc.,
  - whether a painting is in the style of Picasso...
  - what information is available about X
  - which changes in the scene or changes of viewpoint will alter available information about X.
- and other [epistemic](#) affordances.

All of these percepts require some sort of internal language to express them, record them, reason with them, use them in planning or controlling actions. For what other purposes can we (or detectives?) use perceptual information?



# Attaching affordance information to scene fragments

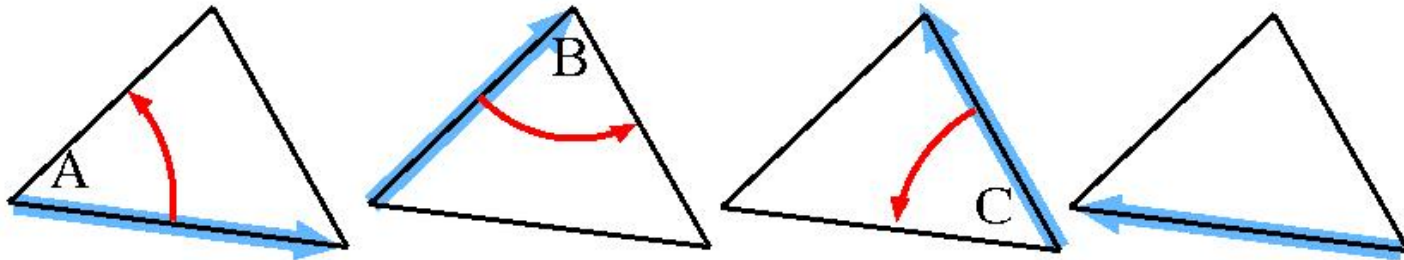
Information gained from visual perception (in intelligent animals) includes information about what is and is not possible (**possibilities** and **constraints**) in the environment – and about realised possibilities when processes are seen (e.g. objects or surfaces moving, rotating, deforming, interacting).

- The Reutersvard “impossible triangle” image, above, shows how rich collections of possibilities can be “attached” to various scene fragments.
- Some of the information is about possibilities and conditionals (proto-affordances):
  - X is possible, Y is possible, Z is possible
  - if X occurs then, ...; if Y occurs then, ...; if Z occurs then, ...  
(X, Y and Z need not be actions of the perceiver: generalising Gibson’s affordances)
- including “epistemic affordances”  
(if motion **M** occurs, then information **I** will become available/or become inaccessible).
- This generalises **aspect graphs**, which encode information about how changes of viewpoint (or rotations of a perceived object) will change what is and is not visible. (Faugeras et al., 1992)  
[http://people.csail.mit.edu/bmcutler/6.838/project/aspect\\_graph.html#1](http://people.csail.mit.edu/bmcutler/6.838/project/aspect_graph.html#1)
- An important special case of this idea concerns the effects of different kinds of **material** on what would happen if various things were to happen:
  - e.g. material being: rigid, plastic, elastic, liquid, viscous, sticky, etc.
  - <http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#babystuff>

The idea of generalising the concept of “aspect graph” to include effects of actions other than viewpoint change, arose in a discussion with Jeremy Wyatt around 2003. (It has probably been thought of by others.)

# Mathematical perception

Mary Pardoe's proof of the triangle sum theorem:



<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/triangle-sum.html>

Extraordinary achievements of our ancestors were recorded in Euclid's *Elements* about 2,500 years ago. <http://www.gutenberg.org/ebooks/21076>,

These must have been the outcome of millions of years of change in many abilities:

abilities to perceive, notice relationships, use relationships in solving practical problems, understand WHY the solutions worked, discuss with other individuals, teach other individuals, and eventually record the discussions and conclusions.

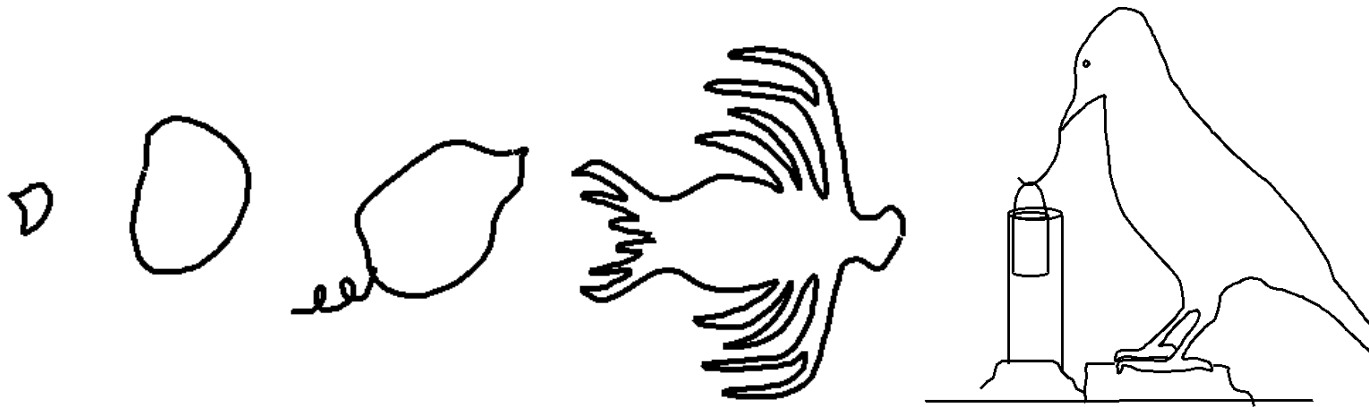
I believe very little is known about that history: one challenge for AI theories of cognition is to show how those discovery processes might be replicated in intelligent machines.

Although many (not all) modern mathematicians assume that mathematical reasoning must make use of explicit axioms and logical inference procedures, these abilities depend on logical discoveries made only in the last two centuries, e.g. by Boole, Peano, Frege, Russell and others.

**Euclid could not have used modern logic. What alternatives are possible?**

# All organisms are information-processors but the information to be processed has changed and so have the means

---



From microbes to hook-making crows:  
How many transitions in information-processing powers were required?

Contrast these transitions:

- transitions in physical shape, size and sensory motor subsystems
- transitions in information processing capabilities.

Fossil records don't necessarily provide clues.

---

Photographs and videos of Betty the famous hook-making crow are available at the Oxford ecology web site:

[http://users.ox.ac.uk/~kgroup/tools/crow\\_photos.shtml](http://users.ox.ac.uk/~kgroup/tools/crow_photos.shtml)

<http://users.ox.ac.uk/~kgroup/tools/movies.shtml>

# Environments have agent-relative structure

---

Environments in which animals evolve, develop, compete, and reproduce, vary widely in their information-processing requirements and opportunities.

If we ignore that environmental richness and diversity, our theories will be shallow and of limited use.

In simple environments everything can in principle be represented numerically, e.g. using numbers for location coordinates, orientations, velocity, size, distances, etc.

In practice the optic cone may lack the required clarity and detail; but for many purposes it may suffice to perceive partial orderings and topological relationships, as illustrated in

<http://www.cs.bham.ac.uk/research/projects/cosy/papers/changing-affordances.html>

In more complex environment things to be represented include:

- Structures and structural relationships, e.g. what is inside, adjacent to, connected with, flush with, in line with, obstructing, supporting...
- Different sorts of processes, e.g. bending, twisting, flowing, pouring, scratching, rubbing, stretching, being compressed.
- Plans for future actions in which locations and arrangements and combinations of things are altered (e.g. while building a shelter).
- Intentions and actions of others.
- Past and future events and generalisations.

**How can all those be represented in brains, or computers?**

**How can the information be used?**

# Varied environments produce varied demands

## Types of environment with different information-processing requirements

- Chemical soup
- Soup with detectable gradients
- Soup plus some stable structures (places with good stuff, bad stuff, obstacles, supports, shelters – requiring enduring location maps.)
- Environments that record who moved where (chemical trails, footprints, permitting stigmergy).  
<http://en.wikipedia.org/wiki/Stigmergy>
- Things that have to be manipulated to be eaten (disassembled)
- Controllable manipulators
- Things that try to eat you
- Food that tries to escape
- Mates with preferences
- Competitors for food and mates
- Collaborators that need, or can supply, information.
- and so on .....

## How do the information-processing requirements change across these cases?

For more on evolutionary transitions see

<http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#talk89>

Genomes for self-constructing, self-modifying information-processing architectures

# Beyond affordances and invariants

---

- Vision does not just have one function, but many, and the functions are extendable through learning and development – building extensions to the architecture.  
E.g. in humans: reading text, music, logic, computer programs, seeing functional relations, understanding other minds, ....
- Vision deals with multiple ontologies
- Vision is not just about what's there but (as Gibson says) also about **what can happen**
- But what can happen need not be caused by or relevant to the viewer's goals or actions  
Trees waving in the breeze, clouds moving in the sky, shadows moving on the ground, leaves blown by the wind – can all be seen, as can their possibilities when they are not happening.
- Besides **action affordances** there are also **epistemic affordances** concerned with availability of information. (Compare (Natsoulas, 1984))
- Besides affordances for the viewer some animals can see **vicarious affordances**, i.e. affordances for **others** (J. J. Gibson, 1979, p.141)  
including affordances for predators, prey, potential mates, infants who need protection, etc.
- Seeing structures, relationships, processes, and causal interactions (or fragments thereof) not relevant to the goals, needs, actions, etc. of the viewer can make it possible to do novel things in future, by combining old competences.  
Great economies and power introduced by using an ontology that is exosomatic, amodal, viewer-neutral. (Still missing from current robots?)
- Functions of vision for other organisms may not be obvious:  
e.g. Portia Spider. (Tarsitano, 2006) <http://dx.doi.org/10.1016/j.anbehav.2006.05.007>

# Multi-strand relationships and processes

Some representational problems arising from complexities of perception of processes were discussed at a conference in 1982 (Sloman, 1983).

## In particular:

- Perceived objects whose parts move continuously can change their shapes, increasing or decreasing the numbers of features – e.g. a line growing a blob, a blob growing an appendage, two shapes merging, a moving string acquiring new changes of curvature, new inflection points, new contact points, or new junctions or knots.
- When two complex objects with parts are perceived together (e.g. two hands), not only are the wholes related but also the parts: they are related both within and across objects, e.g. a part of one object aligning with a part of another, and the relations can be of different sorts – metrical, topological, causal, etc.: “Multi-strand relationships”.

Compare the figure in the earlier slide on “Disentangling complex scenes” requiring shapes to be disentangled (from the film “Sunday Bloody Sunday”).

- When objects exhibiting multi-strand relationships move or change shape, several of the pre-existing relationships can change in parallel: “Multi-strand processes”.
- Perceiving and understanding multi-strand relationships and processes is required for many physical actions, e.g. washing up cups, saucers, bowls and cutlery, in a bowl of water using a dishmop, dressing or un-dressing a child, reversing a car into a narrow parking space, and many more.
- Some art forms depend on the possibility of perception of multi-strand processes, e.g. Ballet performances.

What sorts of internal “language” can serve to record those processes, or allow reasoning about them?

- Description logics can be used to express static multi-strand relationships, but it is not clear whether they are useful for perception of complex multi-strand processes in which collections of relationships alter continuously or go in and out of existence (e.g. in a ballet performance, or rapids in a river).

# Beyond Gibson: Generalised Gibsonianism (GG)

Considering the many functions of vision,

including roles in mathematical reasoning (geometrical, topological, logical, algebraic), and various roles in a robot capable of seeing and manipulating 3-d structures,

leads to an extension of Gibson's theories,

while accepting his rejection of the naive view (e.g. Marr) that the function of vision is only to provide information about what objectively exists in the environment.

In particular, don't expect **one** set of functions to be common to all animals that use vision.

Many species use vision only for the **online** control of behaviour,

using many changes in features of optic array, and correlations of those changes with actions, to provide information about what can be or should be done immediately (e.g. the need to decelerate to avoid hard impact, the need to swerve to avoid an obstacle, the possibility of reaching forward to grasp something).

In contrast, humans (though not necessarily newborn infants) and possibly some other species, use vision for other functions that go beyond Gibson's functions.

Moreover, in order to cope with novel structures, processes, goals and actions, some animals need vision to provide **lower level** information than affordances, information that is potentially shareable between different affordances: "proto-affordances":

**I.e. Information about what changes are possible in a situation, and what constraints there are relating changes – independently of relevance to particular actions. (Sloman, 1996)**

E.g. these two surface fragments could move closer; this surface fragment will obstruct that motion, that rolling ball cannot fit through that gap....

See also (Sloman, 1996, 2008a) ((Siegel, 2014) proposes this label for a very similar concept.)



# Beyond behavioural expertise

---

Annette Karmiloff-Smith has drawn attention to the fact that once behavioural expertise in some domain has been achieved (e.g. ability to move at varying speeds, keeping upright on rough ground, avoiding obstacles, going through gaps, keeping up with others, etc.) there are sometimes new competences to be acquired, which she describes as involving “Representational Redescription”.

Example: after learning how to run with the herd some animals may develop the ability to remember what they did and did not do, and why they failed to do certain things. They might also be able not only to behave but also to reason about possible behaviours and their consequences prior to producing those behaviours (often a necessary prerequisite for producing sensible behaviour).

See ([Karmiloff-Smith, 1992](#)) for more on varieties of representational redescription and their consequences.

I suspect that in some cases it's a change in architecture rather than a change in representation that's important.

I have tried to bring out some of the implications of the book here

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/beyond-modularity.html>

(Work in progress.)

Many researchers in robotics and cognition, especially **embodied** cognition, ignore differences between behavioural expertise and additional requirements such as knowing why you did **not** do something, and what would have happened if you had done it.

# Connections with evolution of mathematical competences

Long before there were mathematics teachers, our ancestors must have begun to notice facts about geometrical shapes that were later codified in Euclidean geometry.

Long before forms of logical reasoning powerful enough to serve the purposes of mathematicians were discovered/created by Frege, Russell and others in the 19th century, mathematicians were making discoveries and proving them, using their ability to notice possibilities for change, and invariants across change, in geometrical configurations.

I've discussed examples in (Sloman, 1968/9, 1971, 1978a, 1996, 2010)

If my generalisations of Gibson's notion of "perception of affordance" to cover a very much wider variety of perceptions of possibilities for change and constraints on change than he mentioned, are correct, then we can see how some of the roots of mathematical cognition as demonstrated in the discovery and proof of theorems in Euclidean geometry and topology may have developed from ancient animal abilities to perceive and reason about affordances required for selecting complex goals and plans.

Some fragments of Euclidean geometry concerned with possibilities are discussed here:

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/triangle-theorem.html>

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/triangle-sum.html>

Such animal capabilities were discussed by Kenneth Craik in (Craik, 1943).

I have tried to illustrate these capabilities in the processes of discovery of "toddler theorems" in young children, in (Sloman, 2008b).

For more information about toddler theorems and their relationships to Annette Karmiloff-Smith's ideas about "Representational Redescription" (Karmiloff-Smith, 1992) see

<http://tinyurl.com/BhamCog/misc/toddler-theorems.html>

# Returning to language

---

- Earlier slides presented some widely held views about the nature of human language.
- And some divergent views about the evolution of language.
- We extracted three core ideas about the nature of language and generalised them to define the concept of a Generalised Language (GL) that includes “internal” languages:
  - Extend the notion of compositional semantics to allow for richer context-dependence in GLs (analogous to Gricean principles for communication).
  - Extend the generalisation to include non-verbal languages using diagrams and other spatial structures to combine information.
- Show that a GL used internally (i.e. not for communication) can be useful for an intelligent animal or robot. (But there are deep unsolved problems about the details.)
- Show that some competences of both prelinguistic humans and some other animals seem to require use of **internal GLs**, representing structures, processes, intentions, ...
  - See research on cognitive and social competences of infants and toddlers, e.g. E. Gibson & A. Pike *An Ecological Approach to Perceptual Learning and Development*, OUP, 2000
- Conclude that internal GLs **evolved before** human external languages, and that in individual humans they **develop before** an external language is learnt.
- Point out some implications for theories of evolution of human language.
- Point out some implications for theories of language learning in humans
  - Supported by the example of Nicaraguan deaf children, and Down’s syndrome children. Nicaraguan deaf children brought together and taught a sign language, rapidly developed a new more sophisticated sign language which they used very fluently and which their teacher was unable to learn. This was **creation**, not **learning**. I suggest that’s also true of normal language development.

[http://en.wikipedia.org/wiki/Nicaraguan\\_Sign\\_Language](http://en.wikipedia.org/wiki/Nicaraguan_Sign_Language)

# What representations are NOT

---

It is often stated, e.g. in AI text-books, and in research publications, that representations “stand for” or “stand in for” or “resemble” the things they represent: **this is a serious confusion.**

- Information about  $X$  and  $X$  itself will generally be used for quite different purposes.
  - A recipe for a type of cake gives information about ingredients and actions to be performed to make an instance.
    - If you mix and cook the ingredients properly (eggs, flour, sugar, etc.) you may get a cake.
    - Compare trying to mix and cook bits of the recipe (e.g. the printed words for the ingredients)
- A 2-D representation of a 3-D object cannot be used as a replacement for the object.
- If  $X$  is some type of physical object, then information about  $X$  might be used to work out how to make  $X$ , to decide whether to make  $X$ , to reason about the cost of making  $X$ , to work out how to destroy  $X$ , how to produce a better  $X$ , ...
- If  $X$  is a type of action, then information about  $X$  can be used to decide whether to perform  $X$ , to work out how long  $X$  will take, to work out risks in doing  $X$ , to decide how to perform  $X$ , to produce a performance of  $X$ , to modulate the performance of  $X$ , to evaluate the performance of  $X$ , to teach someone else how to perform  $X$ .
- If  $X$  represents a generalisation, e.g. “All unsupported objects near the surface of the earth fall”, then there is no object  $X$  refers to that can be used, manipulated, modified etc.

# Some requirements for human and animal competences

## A few reminders:

- Humans and other animals can take in information about objects and situations of widely varying complexity – which can change **while perceived**.
- We can notice and reason about portions of a situation that can change or move in relation to others, producing new situations.
- We can think about some these things even when we have never seen them happen and what we are thinking about is not happening.
- We can use these abilities in coping with dangers and opportunities in the environment, and in planning and controlling actions so as to use opportunities and avoid or solve problems in the environment.
- All those competences involve abilities to acquire, manipulate and use information about things that exist or could exist.
- **That means we need mechanisms for creating, manipulating, storing, using, deriving new internal structures that encode information.**  
**I.e. there are mechanisms for using internal languages.**
- How to do that is a major topic in research in AI – there has been some progress, but there are still many unsolved problems.

# Reasoning with spatial structures

Will the pen hit the rim of the mug if moved downwards?

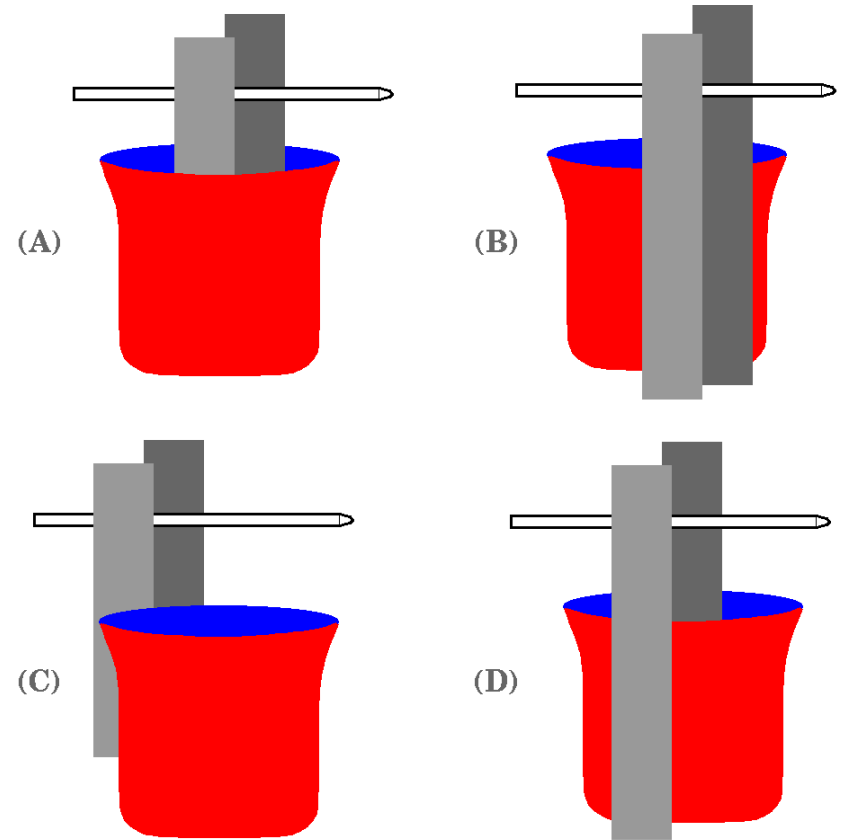
In the scenes depicted, you can probably distinguish cases where the answer is clear from cases where the answer cannot be determined.

Where the answer is clear you can find the answer by imagining the pen moving down between the rectangular sheets, and working out whether it will hit the rim or not.

This is a simple illustration of a general point: we often reason spatially by visualising a view of some configuration and imagining parts moving around and seeing what happens.

Where the answer is uncertain, because of some ambiguity in what you see, you can probably imagine a way of moving left or right, or up or down, so as to remove, or reduce the uncertainty.

I argued in (Sloman, 1971) that visualisation can provide valid inferences, just as logical reasoning can, and that AI researchers need to investigate such modes of inference.



# Main Theses

---

## Main Theses

(a) GLs evolved in biological organisms **for internal uses**, before human languages developed for external communication

where internal uses included perception of complex situations and formation and execution of complex plans for action,

(b) GLs develop **in pre-verbal human infants** before they learn to use a human language for communication.

For examples of infant competences see

E. Gibson & A. Pike

*An Ecological Approach to Perceptual Learning and Development*, OUP, 2000

The main evidence for (a) is the fact that many non-human animals that do not communicate in anything recognisable as a human language, nevertheless have competences, which, from an AI standpoint, seem to require the use of internal GLs.

SHOW SOME VIDEOS, OF CHILDREN AND ANIMALS.

# Conjecture: Gestural languages came first

---

If one of the uses of GL's was formulation of executable plans for action, then observing someone's action could provide a basis for inferring intentions: so actions could communicate meanings that had previously been expressed in internal GLs.

- So perhaps **involuntary, unintended** communication of plans by executing actions came first?
- The usefulness of such communication could have led to **voluntary** gestural communication, e.g. during performance of cooperative tasks.
- Since there was already a rich internal GL used for perceiving, thinking, planning, acting, etc. there could be both motive and opportunity to enrich actions to extend their **voluntary** communicative functions.

The fact that there are already rich and complex meanings (including plan-structures) to be communicated, and benefits to be gained by communicating them (e.g. better cooperation) makes the evolution of rich forms of communication more likely.

- There are many explanations of the pressure to switch from gestural language (sign language) to spoken language, but that required complex evolution of the physiology of breathing, swallowing, and also control of vocalisations.
- Empirical evidence of the primacy of sign languages:
  - The example of Nicaraguan deaf children, and Down's syndrome children (mentioned previously).
  - Also the fact that most people can't help gesturing while they talk.
  - Some even gesture when on their own, thinking.



# MORE ON CORE PROPERTIES OF GLs

---

Human languages (including sign languages) use many formats and have many features.

Earlier, I described three core properties required for using language in relation to novel situations, for multiple uses, all found in both external human languages and internal GLs.

- **Structural variability:**

Linguistic utterances can include varying numbers of distinct components and are not restricted to flat vectors but can have deeply nested substructures, with pronouns, other forms of anaphora and repeated elements providing cross-links.

- **Context-sensitive compositional semantics:**

Novel structures can be given a meaning in a systematic way on the basis of the meanings of the components and the mode of composition (i.e. structural or syntactic relationships between the components), taking linguistic and non-linguistic context into account when necessary.

Familiar labels for this property include: 'generative' and 'productive'.

An implication is that not everything that can be communicated has to be learnt, or previously agreed.

- **Manipulability:** (closely related to the previous two)

Meaningful structures can be extended, modified or combined for various purposes, discussed later.

The idea of “compositional semantics” needs to be generalised to meet the requirements for internal GLs: compositional semantics is not restricted to human languages used for communication.

# Manipulation of information requires mechanisms

The mere fact that a form of representation supports manipulability as explained above does not in itself explain how actual manipulation occurs in any machine or animal.

That requires **mechanisms** to be available that can construct, modify, combine, store, compare, and derive new instances of representations.

E.g. new phrases, new sentences, new stories, new plans, new diagrams, new working models

If an animal or machine has a large repertoire of information and mechanisms, selecting the appropriate ones to use can itself require additional mechanisms and additional information about how to use the resources.

AI systems typically have powerful abilities, but current systems don't know that they have them; nor can they choose which ones would be best to use: except by following simple pre-programmed rules, which they don't learn, and don't modify.

That will need to be changed.

At present we still have a lot to learn about how to build mechanisms that grow themselves in a machine with human-like competences.

# What does “internalising language” mean?

---

What does the blue part of this common assumption mean:

External human language evolved from primitive to complex communication,  
and was later internalised. (NB: I am not defending this claim: I think it is wrong!)

The reference to **being internalised** could mean something like this:

- Evolution several times extended brain functions so that mechanisms that originally evolved for **peripheral** modules become available for **purely internal** uses  
e.g. visual mechanisms later used for imagining?
- Modules evolved for linguistic communication were later modified for internal use, in something like this sequence of steps (e.g. (Dennett, 1969) ?? (check)):
  - After external languages evolved for communication, humans discovered that it could sometimes be useful to talk to themselves, e.g. when making plans, solving problems, formulating questions ...
  - Subsequent evolutionary changes enabled **talking silently**: i.e. brain mechanisms became able to provide inputs directly to the speech input portions of the brain, instead of having to route them externally.
  - This made it possible to construct internal meaningful, manipulable linguistic structures that could be used to think, plan, reason, invent stories, solve problems, construct explanations, remember what has happened, etc.

However, such theories of “internalisation” ignore the internal representational (GL) mechanisms required for external language use in the first place. (Sloman, 1978b, 1979)

# Biological relevance

---

**THESIS:** Some animal competences and some competences of pre-linguistic children need richly structured **internal**, manipulable forms of representation with context-sensitive compositional semantics, which are constructed and used for perception, reasoning, planning and generation and achievement of goals related to complex features of the environment.

- **I have tried to bring out some of the possible uses of GLs with the three core properties:** structural variability, context-sensitive compositional semantics, manipulability.  
(Later generalised to include spatial – e.g. diagrammatic – forms of representation).
- **We can point to many competences displayed by prelinguistic children and some other species** that are hard to explain without the use of GLs  
Examples include nest-building, hunting, dismembering a carcass in order to eat it, playing with toys, using tools, making tools, fighting with others, collaborating with others.  
In particular both Humean and Kantian causal reasoning require use of GLs, though in different ways.
- An important point I shall not have time to go into is the need for specific forms of GL that provide meta-semantic competences, e.g. the ability to represent and reason about one's own or others' goals, beliefs, thought processes, preferences, planning strategies, etc. (So-called "**mentalistic**" vs "**mechanistic**" cognition).

For more on that see

<http://www.cs.bham.ac.uk/research/projects/cosy/papers/#dp0604>

Requirements for "fully deliberative" architectures.

# Another generalisation: Non-verbal forms

---

The three core features of human languages **structural variability**, **generalised compositional semantics** and **manipulability** are also features of many non-verbal forms of representation.

Given a map, a flow-chart for an algorithm, a circuit diagram, or a picture of an object to be constructed, more components can go on being added, sometimes indefinitely.

If we use paper, or some other markable surface, it is possible to

- expand a picture or diagram **outwards**,
- add more **internal** details (e.g. extra lines),  
but eventually there is ‘clutter limit’ because the structure is not stretchable.  
(Other kinds of limits relate to short-term memory constraints.)

Structural variability of such spatial forms of representation has recently been enhanced by the use of film or computing techniques that allow zooming in and out to reveal more or less of the ‘nested’ detail.

It is possible that virtual machines evolved in brains allow such ‘zooming’ in and out, though precise requirements for such a facility to be useful still need to be specified.

The retinoid model of Arnold Trehub’s *The Cognitive Brain* (MIT Press, 1991) may be an example.

<http://www.people.umass.edu/trehub/>

---

(Sloman, 1971) describes more precisely a distinction between “Fregean” and “analogical” forms of representation, claiming that both can be used for reasoning, planning, and proofs.

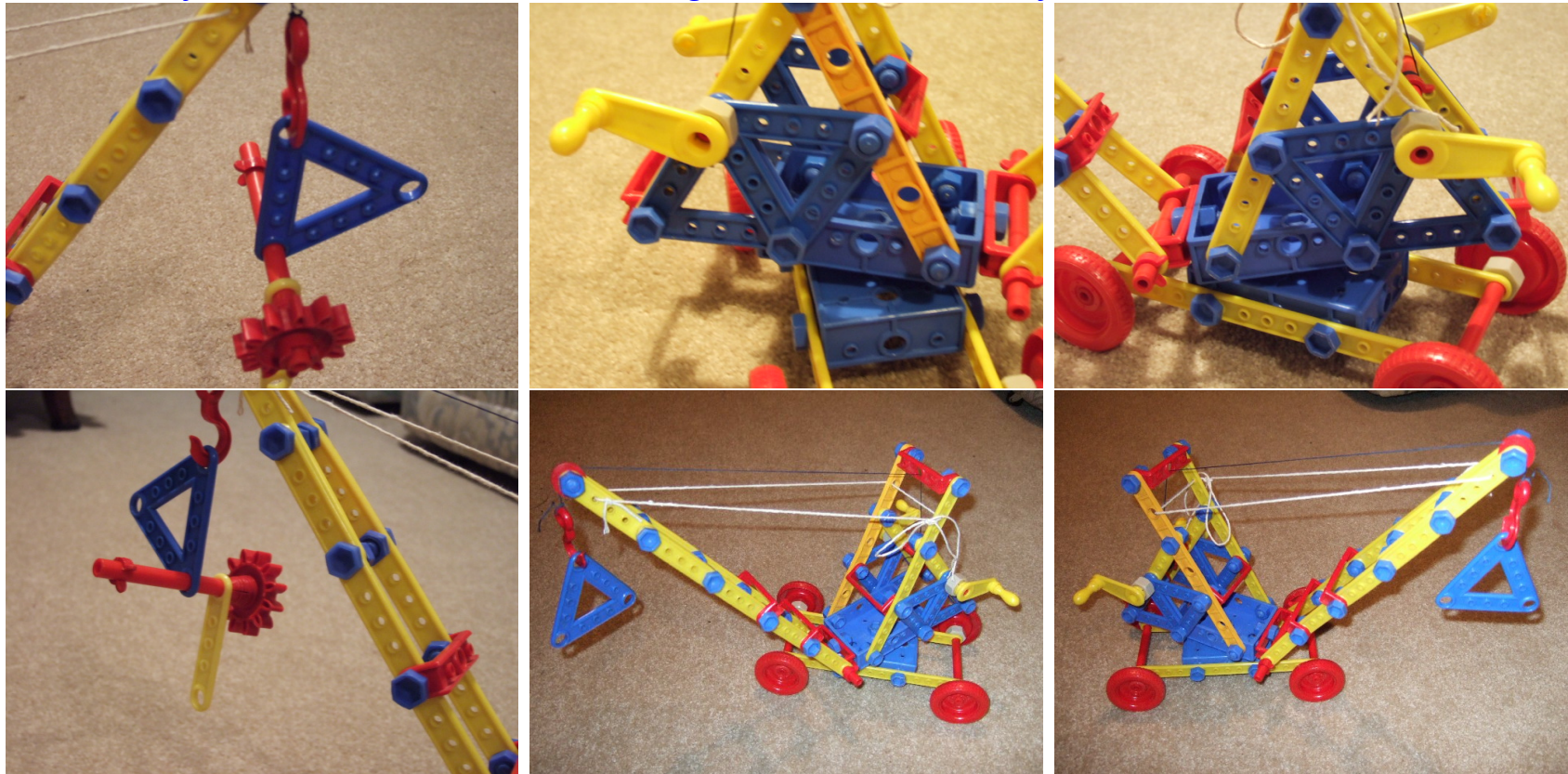
This was a criticism of the purely “Logician” approach to AI expounded by McCarthy and Hayes (McCarthy & Hayes, 1969).



# Compositional semantics and structural variability in vision

Your familiarity with the role of low level pictorial cues in representing features like edges, orientation, curvature of surfaces, joins between two objects or surfaces, etc., allows you to use compositional semantics to see the 3-D structure, and some causal and functional relationships, in pictures you have never previously seen.

No AI vision system comes close to being able to do that – yet.

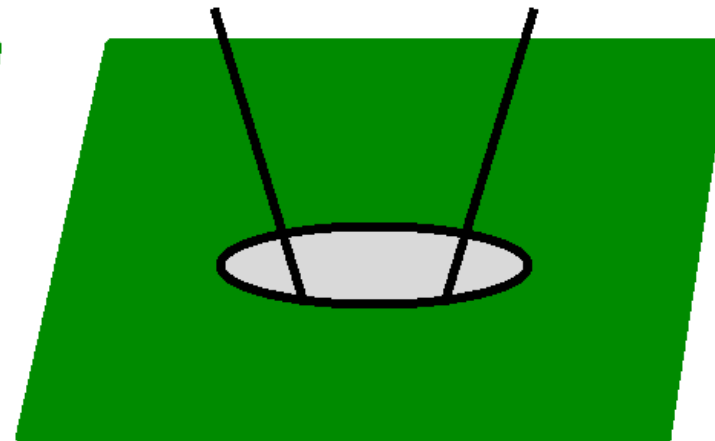
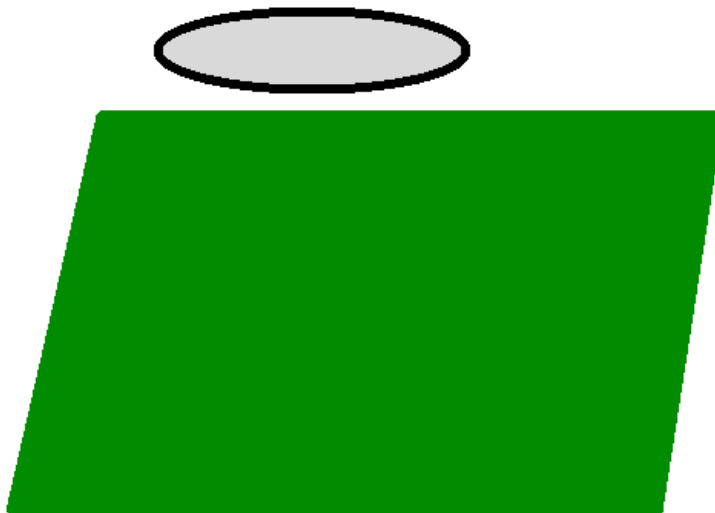
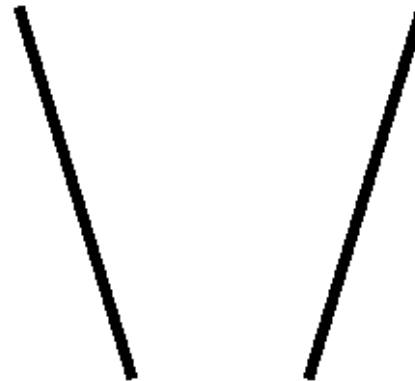
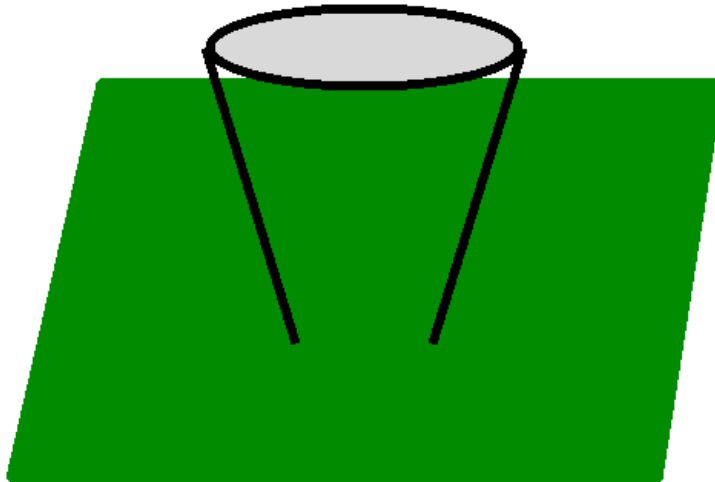


<http://www.cs.bham.ac.uk/research/projects/cosy/photos/crane/>

# Different combinations of the same elements

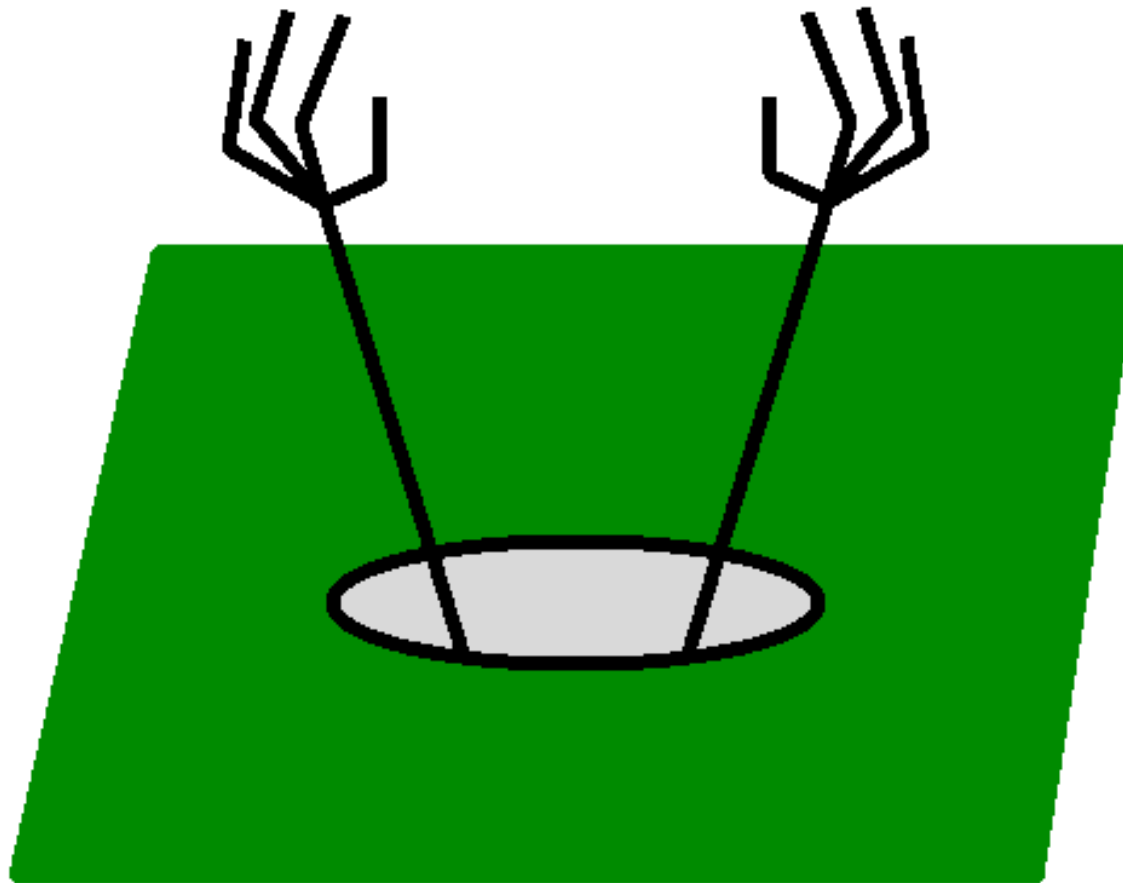
What do you see in these pictures? Only 2-D configurations?

Notice how context can influence 3-D interpretation of parts



# A doodle: Can you tell what this is?

---



Doodles depend heavily on the fact that interpretation of visual GL instances can be partly driven by sensory data and partly by verbal hints (“top down”).



# Possible answers to doodle question

“Early bird catches very strong worm?”

“Sewer worker attacked by a shark?”

Interpretation of visual scenes can include perception of **causal relationships**, as in both the above doodle interpretations.

There is much to be said about doodles, but no time today.

Perceptual combination of spatial and causal relationships is also needed in use or construction of tools: e.g. shape of a spanner’s head.

When objects share a region of space, indefinitely many different kinds of structural and causal relationships can be perceived and interpreted: in contrast with the constrained, rule-based, use of syntactic relations in human formal and informal languages.

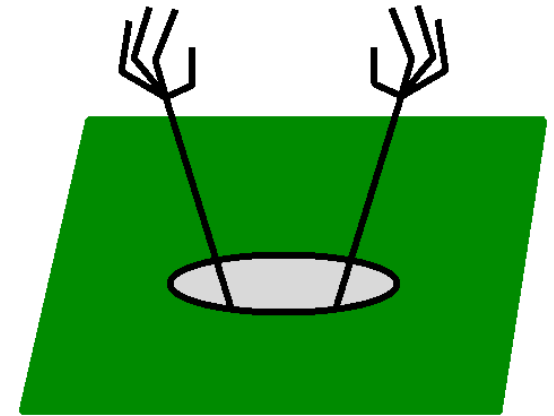
Show broom video, available here (with others)

<http://www.cs.bham.ac.uk/research/projects/cosy/conferences/mofm-paris-07/sloman/vid/>

Long before children can talk, they can take in and make use of structural relationships in the environment in order to produce and control actions.

That’s in addition to their ability to manipulate continuously changing dynamical systems, e.g. maintaining balance while walking, reaching, etc.

Likewise many other animals.



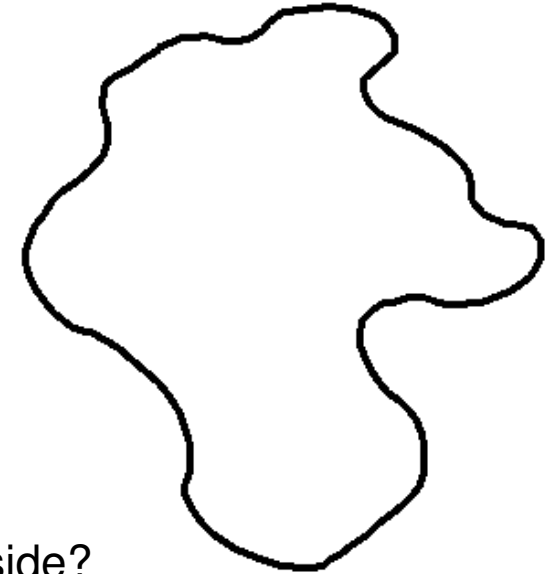
# Example of reasoning with a GL: Making an “H”

## Making a capital “H” using an elastic band and pins

Suppose you had an elastic band and a pile of pins: could you use the pins to hold the stretched rubber band in the form of an outline capital “H”? What about an “A” ?

What sort of GL is needed to make it possible to answer such a question? Compare these questions:

- How many pins would you need?
- Could you do it using only one hand?
- In what order would you insert the pins?
- How many pins would be inside the band and how many outside?
- Could you do it if the pins were replaced with marbles?



You can probably answer the questions in two ways: by trying physically and examining what happens, and by merely thinking about it and examining what happens.

- A very young child will not be able either to construct the H physically, or to answer the questions. What changes as the child becomes able to do both?
- You are probably able to answer the questions just by thinking about the construction processes and the result. How?
- What is your brain doing while you visualise the process of creating the final configuration?
- Do you first visualise the final configuration, and then make a plan for constructing it, or do you get to the final configuration by making a plan, or visualising the construction process?
- What is your brain doing while you count the imagined pins, inside or outside the band?

## Direction of fit of GL structures to the world

---

Many information structures (**but not all!**) are used to refer to some portion of the world and represent that portion as having certain features, possibly quite complex features:

in principle such things can be true or false, or in some cases more or less accurate or inaccurate, more or less close to being true, etc. all depending on how the world is.

Various philosophers (e.g. Anscombe, Austin, Searle) have pointed out that two major kinds of use of such structures can be distinguished:

- where the information-user tends to construct or modify **the representation** so as to make it true or keep it true (**belief-like** uses)
- where the user tends to monitor and alter **the world** so as to make or keep the information structure true (**desire-like** uses).

Sometimes referred to as a difference in “direction of fit” between beliefs and desires.

The distinction also has a clear role from the standpoint of designers of robots or other intelligent systems, though, as I’ve shown elsewhere, there are more intermediate cases to consider in complex, multi-functional machines (e.g. animals).

---

These ideas about belief-like and desire-like states of an organism or machine are developed further in:

A. Sloman, R.L. Chrisley and M. Scheutz,

The architectural basis of affective states and processes, in  
*Who Needs Emotions?: The Brain Meets the Robot*, Eds. M. Arbib & J-M. Fellous, OUP, 2005, pp. 203–244,  
<http://www.cs.bham.ac.uk/research/cogaff/03.html#200305>

# Desires, beliefs and direction of fit

## Content vs function of mental states

Both beliefs and desires can be checked against current perceptual input, but the consequences of mismatches are different.

What makes something a desire, or belief, or fear, or idle thought depends not on the **form** of the information structure, nor its **medium**, but on its **causal role** in the whole architecture.

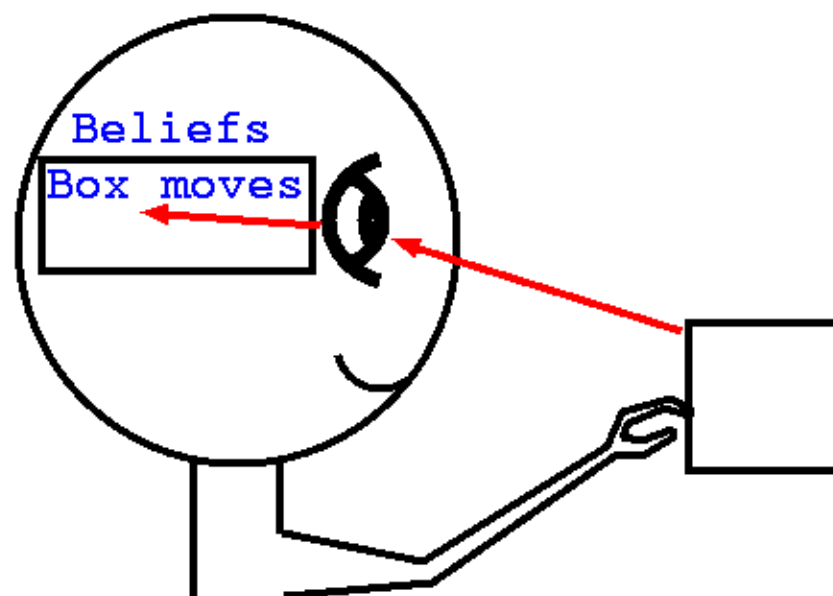
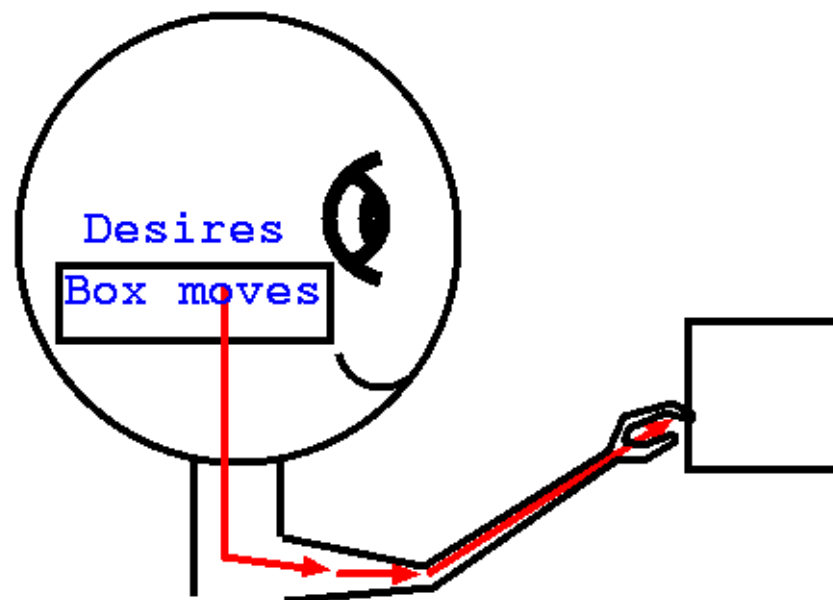
Simple architectures allow for only simple causal roles, whereas more sophisticated architectures allow information structures to have very varied causal roles.

To understand fully the variety of functions served by GLs in a particular type of animal (or machine) we would need to have a detailed specification of the information-processing architecture, and available biological mechanisms.

**We are not ready for that yet!**

See the presentations on architectures here

<http://www.cs.bham.ac.uk/research/projects/cogaff/talks/>



# Varieties of uses of internal GLs

---

Within an organism or robot, a GL structure may have many different kinds of use: depending on the conditions under which it is created, how it is used, what sorts of things modify it and when, and what effects it has and what sorts of things can affect it. For example,

- The use of representations in **perceptual** subsystems is related to one direction of fit (produce information structures that represent how things are)
- Their role in **motivational** subsystems is clearly related to the other direction of fit (change the world so that an information structure represents how things are.)
- An organism's or robot's ability to have very diverse beliefs, desires and competences is connected with the structural variability and compositional semantics of its GLs.
- GLs can be substantially extended during development: they are not innately given.
- Some representations need to endure and be usable in different contexts (e.g. facts, values, competences), whereas others are needed only transiently (e.g. feedback).
- The conditions for a GL to be used **for planning several steps ahead** are different from the conditions for using information **for online control** of continuous actions.

The former requires more complex virtual machines that evolved much later and in relatively few animals, and benefits from an animal's ability to represent states of affairs and processes independently of the sensory and motor signals involved in perceiving or producing them, using an amodal, exosomatic ontology.

I suspect confusion about so-called mirror neurones can arise from a failure to understand that point. (Should they have been called 'abstraction neurones'?)

# Other uses of GL structures in humans

---

Besides expressing semantic contents for desire-like and belief-like states, GL structures can have a wide variety of causal roles, depending not only on their location in the architecture, but also on their form and the mechanisms available for manipulating them. E.g.

- Comparing and evaluating things, states of affairs, possible actions, goals, policies, ...
- creating more or less complex plans for future actions
- using a plan to control actions (either continuously, as in visual servoing, or step-by-step)
- synchronising concurrent processes, or modulating ongoing processes
- expressing a question,
  - i.e. constructing a GL structure that directs a search to determine whether it is true or false, or how it needs to be modified or expanded to make it true.  
<http://www.cs.bham.ac.uk/research/projects/cosy/papers/#dp0502>
- considering unobserved possibilities to explain what has been observed,
- predicting things that have not yet happened
  - (e.g. Humean or Kantian causal reasoning),
- fantasizing, e.g. wondering what would have happened if,
- inventing stories
- day-dreaming
- meta-management functions (making use of meta-semantic competences).

**Most animals, and current robots, have much simpler information processing competences.**

# Perceiving spatial structure vs creating images

---

Information structures in a spatial GL should not be confused with images

An image is a very large collection of small image features,

which may include colour, brightness, texture, edge-features, optical flow, and various kinds of gradient, and various metrical and qualitative 'low-level' relationships such as brighter, same colour, coarser textured, so many degrees apart, etc.

For pictorial or spatial GLs to be useful in the ways described, they must be composed of **larger** structures with **more global relationships** not restricted to simple metrical comparisons.

This topic was much discussed by AI vision researchers in the 1960s. See S. Kanefff (ed) *Picture Language Machines*, Academic Press 1970.

and <http://hop1.murdoch.edu.au/showlanguage.prx?exp=7352&language=Clowes>

## The larger structures

- may be image components like lines, regions, polygons, with relationships like touching, enclosing overlapping, being collinear, approaching, etc., OR
- they may be representations of 3-D or other objects and processes represented by the 2-D structures, e.g. fingers, pools, planks, rocks, rivers, trees, trains going into tunnels, etc., with static and changing 3-D and causal relationships, e.g. supporting, penetrating, grasping, pushing, going behind, etc.

For the user of the GL to be able to perform manipulations and transformations that are useful for tasks like predicting, planning, explaining, formulating questions, it is necessary to do something like **parsing** of the representations, i.e. segmenting them into relatively large components with relationships, so that either components or relationships can be changed.

This is quite unlike what is called "image processing", e.g. enhancing images or applying global transformations to them, such as convolution.

# Major problems for vision researchers

---

Relationships between **static** complex 3-D objects involve many relationships between parts, some metrical, some topological, and some causal/functional. I.e. relationships between complex, structured, objects are **multi-strand relationships**.

When **processes** occur involving changing or moving 3-D objects, many relationships can change at the same time:  
they are **multi-strand processes**.

- The changes are not just geometrical.  
They can include changing causal and functional relationships (e.g. supporting, compressing, obstructing, etc.).
- Perception of processes can include perception of changing affordances.
- I.e. perceived changes can involve several ontological layers.

We can perceive multi-strand processes in which complex 3-D objects change many relationships at the same time. What forms of representation and what mechanisms make that possible? As far as I know, neuroscientists have no explanations and AI vision researchers have no working models. I'll be glad to be proved wrong.

For more on that see

<http://www.cs.bham.ac.uk/research/projects/cosy/papers/#pr0505>  
A (Possibly) New Theory of Vision (October 2005)

<http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#compmod07>  
Architectural and representational requirements for seeing processes and affordances.  
(31 May 2007, BBSRC Workshop)



# Partial summary so far

---

Many familiar kinds of competence involving

- perception of 3-D structures and processes,
- planning and control of actions in a 3-D environment,
- predicting and explaining processes in the environment

require the use of structured, manipulable **internal** forms of representation with context-sensitive compositional semantics.

Those forms of representation, GLs, have some (but not all) features of human language, but use additional mechanisms and are used internally for information processing.

Some of the manipulations that are possible are **discrete** (e.g. adding or removing an object, or a contact or containment relation), others **continuous** e.g. sliding something, distorting a shape.

In some forms of GL, the structural and functional relationships in the interpretation arise from **spatial embedding** of different parts of the same information structure: rather than use of arbitrary or totally general syntactic conventions (as in language and logic).

Nevertheless the spatial form of representation is not a structure that is **isomorphic** with what it represents.

This can be demonstrated using pictures of impossible objects.

---

Some of these points were made in Sloman 1971 and in Sloman 1979

# Examples: To be expanded

---

Show Felix Warneken movies showing prelinguistic children and chimps apparently spontaneously determining and responding to goals of an adult human.

This requires them not only to use GLs without being able to talk but also possessing some meta-semantic competence.

<http://email.eva.mpg.de/~warneken/>

Warneken was mainly concerned with evidence for altruism.

I am mainly concerned with the cognitive mechanisms presupposed by the performances, whatever the motives.

Nest building birds, e.g. corvids.

Could you build a rigid nest using only one hand (or hand and mouth), bringing one twig at a time?

Betty making hooks in different ways and using them for a common task.

Search using google for

betty crow hook

Humans can solve many problems about spatial structures and processes in their heads, illustrated in previous slides.

# Implications of the examples

---

GLs are needed for many capabilities shown by other animals and capabilities shown by pre-linguistic children.

So they cannot be a by-product of evolution of human language.

Since GLs can express plans that can be used to control actions, and since actions can reveal intentions, they are already well suited as a basis for generating communicative language

Implication: sign-languages evolved first, but previous theories about how that happened must be wrong

E.g. theories claiming that simple external gestures arose first, then increasing complexity, then vocalisation and finally internalisation must be back to front.

# Unanswered questions

---

Despite the evolutionary continuities between humans and some other species it is clear that there are many spectacular discontinuities

(e.g. only humans make tools to make tools to make tools .... to build things, and it seems to be the case that only humans prove mathematical theorems, enjoy thinking about infinite sets, tell stories to one another, etc.).

## What explains these discontinuities?

We need to consider various possibilities:

- Was there some change in degree that went past a threshold whose effects were then amplified? (E.g. some memory structure increased in size?)
- Was there a particular crucial discontinuous change in architecture, or some mechanism, or some form of representation, after which effects cascaded?
- Were there several different changes, with independent causes, which combined to produce new spectacularly different effects?
- other possibilities???

We don't know enough to answer, but I suspect the first answer (a quantitative change passed a threshold) is unlikely.

I suspect there were a few key discrete architectural changes, that modified the forms of learning and development in humans and other altricial species (see below).

# Language learning vs language development

If those observations and speculations are correct, some previous theories about language learning must be wrong!

- Previous theories imply that children do not acquire a way of representing information that supports structural variability, compositional semantics and useful manipulability **until they have learnt an external human language**, which they do by some sort of data-mining of perceived uses of language by others.
- If our speculations are correct, the process of language learning is primarily one of **creative and collaborative problem solving** in which new ways of expressing pre-existing meanings are devised collaboratively.
- This is a process of development of internal GLs (able to cope with structural variability and compositional semantics) along with their extension to an external mode of expression.
- The fact that learners are normally in a minority and can have little influence on the outcome makes it **look as if** they are absorbing rather than creating.

But the Nicaraguan case shows that must be wrong. Nicaraguan deaf children rapidly developed a new sophisticated sign language which they used very fluently and which their teacher was unable to learn.

[http://en.wikipedia.org/wiki/Nicaraguan\\_Sign\\_Language](http://en.wikipedia.org/wiki/Nicaraguan_Sign_Language)

Once humans had acquired the ability to communicate rich and complex meanings, cultural evolution, including development of new linguistic forms and functions, could enormously speed up transmission of information from one generation to another and that might produce evolutionary opportunities to extend the internal GL-engines.

# Implications for Chomskyan theories

---

Does all the above imply that humans have anything like the kind of innate (genetically determined) **Language Acquisition Device** (LAD) postulated by Chomsky

(E.g. in **Aspects of the Theory of Syntax**, (Chomsky, 1965) – or his later work?)

or is the learning of human language completely explained by general purpose learning mechanisms at the basis of all human intelligence?

**Our theories imply that the answer is somewhere in between and back to front.**

The discussion of the need for GLs in humans and other animals implies that evolution produced something used **internally** with the three core properties, thereby supporting intelligent perception and manipulation of objects in the environment.

The use of GLs also supports the development of communicative language:

a pre-verbal child has things to communicate about

and has motives that can be served by such communication.

But details of such linguistic development have not yet been understood or modelled.

They may turn out to fit the sort of theory presented in (Chappell & Sloman, 2007) (see diagram later), related to (Karmiloff-Smith, 1992).

# A different view of language development

---

The GL structures were at first not overtly communicated and did not use the grammars of later human languages.

Insofar as internal GLs are partly acquired through interaction with the environment, instead of being wholly innate, it follows that the genome of some species provides one or more [GL acquisition devices \(GLADS\)](#), though they are better viewed not as completely innate devices, but as self-extending mechanisms, whose self-extending capabilities are themselves extended by things derived from the environment.

When evolution of communicative uses of GLs began they would have built most naturally on the role of GLs in controlling behaviour (e.g. executing a plan), since what you do often communicates your intentions.

That probably involved many evolutionary steps that will be hard to find evidence for.

Only later would new pressures cause vocal information structures to take over.

The additional constraints of that impoverished medium (compared with the earlier gestural GL) may have driven both human languages and the brain mechanisms down narrow channels, further constraining the permitted structural variability and modes of composition.

But that's a topic for another time.

# Routes from genome to behaviours

Cognitive epigenesis (Chappell & Sloman, 2007)

The meta-configured genome

The diagram shows different stages at which the environment influences gene expression, e.g.:

- during development of seed, egg, or embryo, and later growth (i.e. not all controlled by DNA)
- triggering meta-competences to produce new competences or new meta-competences (e.g. after previous competences have produced exploratory and learning processes)
- during the triggering and deployment of the competences to produce behaviours

Insofar as behaviours influence the environment there can be complex developmental feedback loops. Competences and behaviours further to the right may use several 'simpler' competences and behaviours developed earlier on the left.

A video explaining some of this, is here:

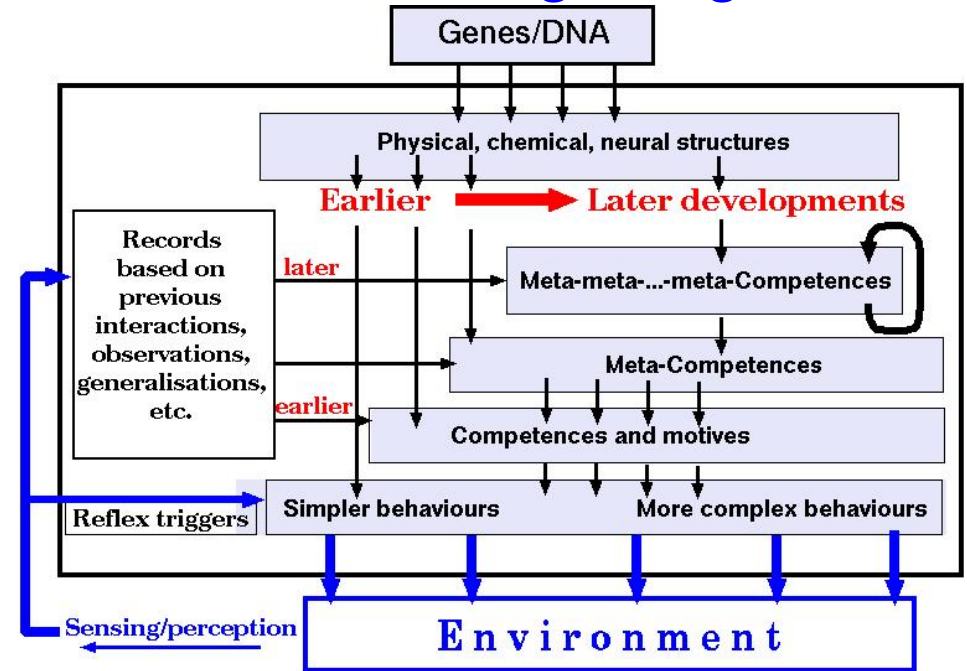
<https://www.youtube.com/watch?v=G8jNdBCAxVQ>

The construction of some competences should be construed as an ongoing process, with repeated activation of some meta-competences (e.g. de-bugging) over time.

These schematic specifications may have different sorts of instantiations in different parts of a multi-functional architecture, e.g. in reactive and deliberative components.

In reactive components many (but not all) processes will involve **continuous** control.

In deliberative and meta-management components many processes will be **discrete**.



Chris Miall helped with the original version of the diagram. This version gives more detail.

Diagram is from the IJUC paper with Jackie Chappell. (Chappell & Sloman, 2007)



# Cascaded development and learning

---

If learning has to go through partially-ordered competences, where each competence builds on what has been built in previous stages, and that involves building new layers of brain mechanism, then that might explain why each new GL extension can only happen at a certain stage of development.

A particular GL cannot be added too early because it needs prior resources to provide

- the representing structures,
- the ability to manipulate them, and
- the contents that they represent.

It can't happen too late because lots of other things are normally delayed until the appropriate GL has got going, and if that doesn't happen they may develop anyway, but in inferior forms and they cannot be disassembled and reassembled later.

There may also be facts related to the sequence in which portions of brains develop. (e.g. myelinization??)

But the stages may be only **partially** ordered – allowing different learners to traverse different developmental trajectories in a network of possible trajectories.

(Compare Waddington's epigenetic landscape.)

All this still needs to be made a lot more precise – preferably in working models, if current forms of computation are sufficiently powerful!

# The implementability requirement

---

We need to be very cautious about unimplemented models.

- All three of the core properties (structural variability, compositional semantics and manipulability) have implications for mechanisms, and architectures in which they can be combined.
- Some mechanisms cannot support structural variability, e.g. many of those that deal only with vectors of numerical values.
- Some mechanisms have no use for compositional semantics because they do not do any significant **interpretation** of the structures they operate on.
- **The three core properties should be regarded as properties of virtual machines implemented in brains not as properties of physical mechanisms:**

E.g. your brain does not get rewired when you see a new scene, make an inference, create and compare a set of plans, compose a poem in your head, ..., but a virtual network might be rewired.

For a short introduction to virtual machines and supervenience see

<http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#bielefeld>

A longer, more detailed, introduction: (Sloman, 2013)

- Current computer-based models support only a small subset of the types of manipulability discussed here.  
Current biologically-inspired mechanisms (e.g. existing neural models) are so far inadequate for these purposes.
- **Perhaps animal brains run virtual machines no modellers have thought of yet?**
- E.g. there's growing evidence that sub-neural chemistry plays enormously important roles (Grant, 2018) – perhaps using a mixture of continuous and discrete computation.

# Many unsolved problems

---

These slides scratch the surface of many deep and difficult problems. In particular, I have ignored the fact that there's much that is still not understood about what the varied functions of visual perception are, how they work, and what forms of representation (GLs) they use.

It does not seem to me that anyone in psychology, neuroscience, or AI/Robotics is near answering the questions.

In particular, as far as I know, there are no models of neural mechanisms that can support the abilities required to manipulate, interpret, and reason about complex structures and processes that involve geometry and topology. E.g. statistics based probabilities cannot represent or demonstrate **impossibilities** and **necessary** connections characteristic of ancient spatial reasoning (Kant, 1781; Sloman, 1962). Some examples are here:

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/impossible.html>

<http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#compmod07>

“Architectural and representational requirements for seeing processes and affordances.”

It goes into more detail on many of the topics mentioned here.

In particular it brings out with more examples some of the human abilities to perceive both possibilities and impossibilities (constraints on possibilities), abilities that are closely related to the ability to make mathematical discoveries.

**An important consequence of this line of enquiry seems to be that, at least in humans, the abilities to perceive, to develop linguistic competences and to make mathematical discoveries are all deeply connected in ways that have not yet been studied adequately, and are still far from being explained or modelled.**

# Background to this presentation

---

The slides overlap with these two papers, the first of which introduced the term ‘G-language’, now ‘GL’.

(Sloman & Chappell, 2007),

‘Computational Cognitive Epigenetics’, in *Behavioral and Brain Sciences Journal*, 30(4).

<http://www.cs.bham.ac.uk/research/projects/cosy/papers/#tr0703>

Commentary on: (Jablonka & Lamb, 2005)

Precis of book: <http://www.bbsonline.org/Preprints/Jablonka-10132006/Referees/>

Jackie Chappell and Aaron Sloman, (2007)

‘Natural and artificial meta-configured altricial information-processing systems’,

in *International Journal of Unconventional Computing*, 3, 3, pp. 211–239,

<http://www.cs.bham.ac.uk/research/projects/cosy/papers/#tr0609>

Much earlier, less developed, versions of some of the ideas were in these three papers, all now online.

Sloman71 <http://www.cs.bham.ac.uk/research/cogaff/04.html#200407>

Interactions between philosophy and AI: The role of intuition and non-logical reasoning in intelligence,  
*Proc 2nd IJCAI London*, pp. 209–226, Reprinted in *Artificial Intelligence Journal* 1971.

Describes a distinction between “Fregean” and “analogical” forms of representation, claiming that both can be used for reasoning, planning, and proofs.

Sloman79 <http://www.cs.bham.ac.uk/research/projects/cogaff/81-95.html#43>

The primacy of non-communicative language,

in *The analysis of Meaning: Informatics 5 Proceedings ASLIB/BCS Conference*, Oxford, March 1979, Eds. M. MacCafferty and K. Gray, Aslib, London, pp. 1–15,

Sloman78

<http://www.cs.bham.ac.uk/research/projects/cogaff/07.html#713>

What About Their Internal Languages? 1978,

Commentary on articles by Premack, D., Woodruff, G., Griffin, D.R., Savage-Rumbaugh, E.S., Rumbaugh, D.R., Boysen, S.  
in *Behavioral and Brain Sciences Journal* 1978, 1 (4), pp. 515,

There are several other closely related joint papers by Chappell and Sloman (2005 to 2007) on the CoSy project web site:

<http://www.cs.bham.ac.uk/research/projects/cosy/papers/>

We also have some slide presentations on kinds of causal reasoning in animals and robots prepared for

The Workshop on Natural and Artificial Cognition(WONAC), Oxford 2007, here:

<http://www.cs.bham.ac.uk/research/projects/cogaff/talks/wonac>

# NOTES

---

For people who are not familiar with the story of the Nicaraguan deaf children, there are various summaries on the web including

Brief history and some links <http://www.signwriting.org/nicaragua/nicaragua.html>

PBS documentary including video [http://www.pbs.org/wgbh/evolution/library/07/2/1\\_072\\_04.html](http://www.pbs.org/wgbh/evolution/library/07/2/1_072_04.html)

BBC summary <http://news.bbc.co.uk/2/hi/science/nature/3662928.stm>

Wikipedia summary [http://en.wikipedia.org/wiki/Nicaraguan\\_Sign\\_Language](http://en.wikipedia.org/wiki/Nicaraguan_Sign_Language)

Ann Senghas, Language Emergence: Clues from a New Bedouin Sign Language, (Senghas, 2005)

Bruce Bridgeman has a theory that overlaps significantly with the one presented here:

On the Evolution of Consciousness and Language. *Psychology*: 3(15) Consciousness (1) (1992)

<http://www.cogsci.ecs.soton.ac.uk/cgi/psyc/newpsy?consciousness.1>

References involving mirror neurons and the gestural theory of evolution of language

Fadiga L., Craighero L. Cues on the origin of language. From electrophysiological data on mirror neurons and motor representations, in In S. Breten (Ed.), *On Being Moved: From mirror neurons to empathy*. Amsterdam, John Benjamins. 2007

[http://web.unife.it/progetti/neurolab/pdf/2007\\_1\\_Fadiga-Craighero\\_chapter6.pdf](http://web.unife.it/progetti/neurolab/pdf/2007_1_Fadiga-Craighero_chapter6.pdf)

There are strong connections with the work of Annette Karmiloff-Smith on “Representational Redescription”, outlined in her 1992 book *Beyond Modularity*, reviewed from our viewpoint here

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/beyond-modularity.html>

---

## References

- Abercrombie, M. (1960). *The Anatomy of Judgement*. New York: Basic Books.
- Barrow, H., & Tenenbaum, J. (1978). Recovering intrinsic scene characteristics from images. In A. Hanson & E. Riseman (Eds.), *Computer vision systems* (pp. 3–26). New York: Academic Press.
- Boden, M. A. (2006). *Mind As Machine: A history of Cognitive Science (Vols 1–2)*. Oxford: Oxford University Press.
- Brooks, R. A. (1990). Elephants Don't Play Chess. *Robotics and Autonomous Systems*, 6, 3–15. (<http://people.csail.mit.edu/brooks/papers/elephants.pdf>)
- Brooks, R. A. (1991). Intelligence without representation. *Artificial Intelligence*, 139–159. Retrieved from <http://people.csail.mit.edu/brooks/papers/representation.pdf>
- Chappell, J., & Sloman, A. (2007). Natural and artificial meta-configured altricial information-processing systems. *International Journal of Unconventional Computing*, 3(3), 211–239. Retrieved from <http://www.cs.bham.ac.uk/research/projects/cogaff/07.html#717>
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
- Clowes, M. (1973). Man the creative machine: A perspective from Artificial Intelligence research. In J. Benthall (Ed.), *The Limits of Human Nature*. London: Allen Lane.
- Craik, K. (1943). *The nature of explanation*. London, New York: Cambridge University Press.
- Dennett, D. C. (1969). *Content and Consciousness*. New York: Humanities Press.
- Fadiga, L., & Craighero, L. (2007). Cues on the origin of language. From electrophysiological data on mirror neurons and motor representations,. In S. Breten (Ed.), *On Being Moved: From mirror neurons to empathy*. Amsterdam: John Benjamins. Retrieved from [http://web.unife.it/progetti/neurolab/pdf/2007\\_1.Fadiga-Craighero\\_chapter6.pdf](http://web.unife.it/progetti/neurolab/pdf/2007_1.Fadiga-Craighero_chapter6.pdf)
- Faugeras, O., Mundy, J., Ahuja, N., Dyer, C., Pentland, A., Jain, R., & Ikeuchi, K. (1992, March). Why aspect graphs are not (yet) practical for computer vision. In *Cvgip image understanding workshop* (Vol. 55, pp. 212–218). Retrieved from [https://www.ri.cmu.edu/publication\\_view.html?pub\\_id=4280](https://www.ri.cmu.edu/publication_view.html?pub_id=4280)
- Gibson, E. J., & Pick, A. D. (2000). *An Ecological Approach to Perceptual Learning and Development*. New York: Oxford University Press.
- Gibson, J. J. (1966). *The Senses Considered as Perceptual Systems*. Boston: Houghton Mifflin.
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston, MA: Houghton Mifflin.
- Grant, S. G. N. (2018). Synapse molecular complexity and the plasticity behaviour problem. *Brain and Neuroscience Advances*, 2, 1–7. Retrieved from <https://doi.org/10.1177/2398212818810685>
- Hinton, G. E. (1976). Using relaxation to find a puppet. In *Proceeding of the A.I.S.B. Summer Conference*. University of Edinburgh. Retrieved from <http://www.cs.toronto.edu/~hinton/absps/puppet.pdf>
- Jablonka, E., & Lamb, M. J. (2005). *Evolution in Four Dimensions: Genetic, Epigenetic, Behavioral, and Symbolic Variation in the History of Life*. Cambridge MA: MIT Press.
- Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception and Psychophysics*, 14, 201–211.
- Kanef, S. (Ed.). (1970). *Picture language machines*. New York: Academic Press.
- Kant, I. (1781). *Critique of pure reason*. London: Macmillan. (Translated (1929) by Norman Kemp Smith)
- Karmiloff-Smith, A. (1992). *Beyond Modularity: A Developmental Perspective on Cognitive Science*. Cambridge, MA: MIT Press.
- Marr, D. (1982). *Vision*. San Francisco: W.H.Freeman.
- McCarthy, J., & Hayes, P. (1969). Some philosophical problems from the standpoint of AI. In B. Meltzer & D. Michie (Eds.), *Machine Intelligence 4* (pp. 463–502). Edinburgh, Scotland: Edinburgh University Press. Retrieved from <http://jmc.stanford.edu/articles/mcchay69.html>
- Michotte, A. (1962). *The perception of causality*. Andover, MA: Methuen.



- Natsoulas, T. (1984, July). Towards the Improvement of Gibsonian Perception Theory. *J. for the Theory of Social Behaviour*, 14(2), 231–258. Retrieved from [DOI:10.1111/j.1468-5914.1984.tb00496.x](https://doi.org/10.1111/j.1468-5914.1984.tb00496.x)
- Searle, J. (1984). *Minds brains and science*. BBC Publications. ((The Reith lectures))
- Senghas, A. (2005). Language Emergence: Clues from a New Bedouin Sign Language. *Current Biology*, 15(12), R463–R465. Retrieved from <http://dx.doi.org/10.1016/j.cub.2005.06.018>
- Shannon, C. (1948, July and October). A mathematical theory of communication. *Bell System Technical Journal*, 27, 379–423 and 623–656.
- Siegel, S. (2014). Affordances and the Contents of Perception. In B. Brogaard (Ed.), *Does Perception Have Content?* (pp. 39–76). USA: OUP. Retrieved from <http://philpapers.org/rec/SIEAAT>
- Simon, H. A. (1969). *The sciences of the artificial*. Cambridge, MA: MIT Press. ((Second edition 1981))
- Sloman, A. (1962). *Knowing and Understanding: Relations between meaning and truth, meaning and necessary truth, meaning and synthetic necessary truth* (DPhil Thesis) (Doctoral dissertation, Oxford University). Retrieved from <http://www.cs.bham.ac.uk/research/projects/cogaff/62-80.html#1962>
- Sloman, A. (1965). Functions and Rogators. In J. N. Crossley & M. A. E. Dummett (Eds.), *Formal Systems and Recursive Functions: Proceedings of the Eighth Logic Colloquium Oxford, July 1963* (pp. 156–175). Amsterdam: North-Holland Publishing Co. Retrieved from <http://www.cs.bham.ac.uk/research/projects/cogaff/62-80.html#rog>
- Sloman, A. (1968/9). Explaining Logical Necessity. *Proceedings of the Aristotelian Society*, 69, 33–50. Retrieved from <http://www.cs.bham.ac.uk/research/projects/cogaff/62-80.html#1968-01>
- Sloman, A. (1971). Interactions between philosophy and AI: The role of intuition and non-logical reasoning in intelligence. In *Proc 2nd ijcai* (pp. 209–226). London: William Kaufmann. Retrieved from <http://www.cs.bham.ac.uk/research/cogaff/62-80.html#1971-02> (Reprinted in *Artificial Intelligence*, vol 2, 3-4, pp 209-225, 1971)
- Sloman, A. (1978a). *The computer revolution in philosophy*. Hassocks, Sussex: Harvester Press (and Humanities Press). Retrieved from <http://www.cs.bham.ac.uk/research/cogaff/62-80.html#crp> ((Revised 2018))
- Sloman, A. (1978b). What About Their Internal Languages? Commentary on three articles by Premack, D., Woodruff, G., by Griffin, D.R., and by Savage-Rumbaugh, E.S., Rumbaugh, D.R., Boysen, S. in *Behavioral and Brain Sciences Journal* 1978, 1 (4). *Behavioral and Brain Sciences*, 1(4), 515. Retrieved from <http://www.cs.bham.ac.uk/research/projects/cogaff/62-80.html#1978-02>
- Sloman, A. (1979). The primacy of non-communicative language. In M. MacCafferty & K. Gray (Eds.), *The analysis of Meaning: Informatics 5 Proceedings ASLIB/BCS Conference, Oxford, March 1979* (pp. 1–15). London: Aslib. Retrieved from <http://www.cs.bham.ac.uk/research/projects/cogaff/62-80.html#1979-01>
- Sloman, A. (1983). Image interpretation: The way ahead? In O. Braddick & A. Sleigh (Eds.), *Physical and Biological Processing of Images (Proceedings of an international symposium organised by The Rank Prize Funds, London, 1982.)* (pp. 380–401). Berlin: Springer-Verlag. Retrieved from <http://www.cs.bham.ac.uk/research/projects/cogaff/81-95.html#57>
- Sloman, A. (1996). Actual possibilities. In L. Aiello & S. Shapiro (Eds.), *Principles of knowledge representation and reasoning: Proc. 5th int. conf. (KR '96)* (pp. 627–638). Boston, MA: Morgan Kaufmann Publishers. Retrieved from <http://www.cs.bham.ac.uk/research/cogaff/96-99.html#15>
- Sloman, A. (2001). Evolvable biologically plausible visual architectures. In T. Cootes & C. Taylor (Eds.), *Proceedings of British Machine Vision Conference* (pp. 313–322). Manchester: BMVA. Retrieved from <http://www.cs.bham.ac.uk/research/projects/cogaff/00-02.html#76> doi: <http://dx.doi.org/10.5244/C.15.33>
- Sloman, A. (2007). *Why symbol-grounding is both impossible and unnecessary, and why theory-tethering is more powerful anyway* (Research Note No. COSY-PR-0705). Birmingham, UK. Retrieved from <http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#models>
- Sloman, A. (2008a). Architectural and representational requirements for seeing processes, proto-affordances and affordances. In A. G. Cohn, D. C. Hogg, R. Möller, & B. Neumann (Eds.), *Logic and probability for scene interpretation*. Dagstuhl, Germany: Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, Germany. Retrieved from <http://drops.dagstuhl.de/opus/volltexte/2008/1656>
- Sloman, A. (2008b). *A New Approach to Philosophy of Mathematics: Design a young explorer, able to discover “toddler theorems”*. Retrieved from <http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#toddler>
- Sloman, A. (2010). If Learning Maths Requires a Teacher, Where did the First Teachers Come From? In A. Pease, M. Guhe, & A. Smaill (Eds.), *Proc. Int. Symp. on Mathematical Practice and Cognition, AISB 2010 Convention* (pp. 30–39). De Montfort University, Leicester. Retrieved from

<http://www.cs.bham.ac.uk/research/projects/cogaff/10.html#1001>

Sloman, A. (2013). *Virtual Machine Functionalism (The only form of functionalism worth taking seriously in Philosophy of Mind and theories of Consciousness)* (Research Note). School of Computer Science, The University of Birmingham. Retrieved from

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/vm-functionalism.html>

Sloman, A., & Chappell, J. (2007). Computational Cognitive Epigenetics (Commentary on (Jablonka & Lamb, 2005)). *Behavioral and Brain Sciences*, 30(4), 375–6. Retrieved from <http://www.cs.bham.ac.uk/research/projects/cosy/papers/#tr0703>

Tarsitano, M. (2006, December). Route selection by a jumping spider (*Portia labiata*) during the locomotory phase of a detour. *Animal Behaviour*, 72, Issue 6, 1437–1442. Retrieved from <http://dx.doi.org/10.1016/j.anbehav.2006.05.007>

Winograd, T. (1972). Procedures as a Representation for Data in a Computer Program for Understanding Natural Language. *Cognitive Psychology*, 3(1).