What evolved first:

Languages for communicating
or
Languages for thinking ?[*]

(Generalised Languages: GLs),

[*] Where ‘thinking’ here loosely refers to a wide range of types of internal information processing.

Aaron Sloman
http://www.cs.bham.ac.uk/~axs/

Ideas developed with **Jackie Chappell** (Biosciences, Birmingham).

Some parts arose out of work on the CoSy project:
http://www.cs.bham.ac.uk/research/projects/cosy/papers/

These slides are accessible here:

http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#glang

These Slides (work in progress)

These slides were originally written for presentation at the Language and Cognition seminar, School of Psychology, University of Birmingham, 19th October 2007

The topic is far too large for a single seminar, so only a subset of the slides were presented, after some videos of animals and prelinguistic children doing things that demonstrated perception of structure in the environment, and in two cases interpretation of the intentions of an adult human, and apparently spontaneous actions to help achieve the adult's goals.

> Thanks to Felix Warneken for use of his videos, available here
> `http://email.eva.mpg.de/~warneken/video.htm`

Some of the videos I use in this context are available here (e.g. broom video):

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

Some of the papers referred to in the talk are listed at the end, with URLs.

I thank the members of the seminar for their patience and for interesting comments and questions which have led to some changes to these slides.

NOTE:
My slides are written so as to be readable online,
so they contain a lot more detail than most presentations.

This makes them less suitable for live presentations.

Some references are provided at the end.

Overview of objectives

- Present some widely held views about the nature of human language
- Present some divergent views about the evolution of language
- Extract three core ideas about the nature of language and generalise 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.)

Methodological warning

- Much of what I say is still largely descriptive and hypothetical:
  I cannot (yet) demonstrate working computer models of the mechanisms discussed.
  Although I have a lot of experience of building computer models and can see how parts of what I am talking about could be implemented, there are still many gaps in the current state of the art (especially robot perception of 3-D structures and processes.)
- Until a theory can be expressed with sufficient precision to guide the construction of a working system it should always be regarded as suspect: for example, you cannot easily tell what assumptions you are making and whether they are seriously flawed.
- Implementation on computers is a "proof of concept", but still leaves open the question whether the mechanisms proposed can be implemented on biological information processing machinery.
- Unfortunately I don't think we understand more than a subset of types of biological information processing mechanisms, and we cannot yet build convincing platforms that can be used as a basis for implementing theories of the kind discussed here.
- Therefore, much of this is still speculative and not subject to immediate testing.
- This could be the start of a progressive or a degenerating research programme.
  Deciding which it is can take many years of development of a research programme:

See: Imre Lakatos,
The methodology of scientific research programmes, in *Philosophical papers, volume I*, Eds. J. Worrall & G. Currie, CUP, 1980, (`http://www.answers.com/topic/imre-lakatos`)

Abstract

There are widely held beliefs about the nature of human language, its relationship to various aspects of human mental functioning, its evolution in our species, and its development in individuals. Those beliefs generally ignore or over-simplify the information-processing requirements for language to evolve or be used, and they ignore some facts about what pre-linguistic children can do, and facts about what many other animals can do.

I agree with Karl Popper that one should not waste time attacking straw men, i.e. arguing against obviously false theories: One should always try to present a position argued against in its strongest form before arguing against it. However doing that would make this document at least ten times longer. So for now I merely summarise very briefly the positions I think are mistaken. They are all capable of much more detailed and convincing presentations than I give them here.
References will later be added, though I expect most people reading this will already be familiar with some of the literature.

The key idea: Humans and many other animals use “internal languages” (internal means of representing information) for tasks that are not normally thought of as linguistic tasks, namely perceiving, experiencing, having desires, forming intentions, working out what to do, performing actions, learning things about the environment (including other agents), remembering, imagining, theorising, designing .....

Being able to use such internal languages is a prerequisite for learning an external language.

This requires a generalised notion of a language: a GL, i.e. a form of manipulable representation with structural variability and compositional semantics, as explained later.

This includes both external communicative languages and internal languages, and it allows for a wide variety of forms, including all forms of human language used for communication (spoken, written, signed, maps, diagrams, flow-charts, mathematical formalisms, programming languages), and also many formalisms used internally in working computer models and AI robots, and hypothesised forms of representation used internally by animals (some not yet identified by psychologists or neuroscientists!)

In order to understand how brain mechanisms make GLs possible we have to understand what virtual machines are and the complex ways in which they can be related to physical machines: for a short tutorial see these slides: `http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#bielefeld`

Note for philosophers and doubters

For philosophers convinced by Wittgenstein's arguments (or, more precisely, his rhetoric) that private languages are impossible, it should be made clear that he was attacking a philosophical thesis concerning the use of "logically private" languages, and he knew nothing about computational virtual machines: the core ideas have only been developed since his death.

However, Kenneth Craik, also at Cambridge, had some of the ideas presented here, while Wittgenstein was still alive.

Most philosophers still know nothing about virtual machines, unfortunately, though they use them every day.

These slides present what is primarily a scientific theory, not a philosophical theory, though the ideas have been developed partly on the basis of philosophical conceptual analysis, and many of its key features were inspired by Immanuel Kant's philosophy, including his views about mathematics and causation.

Like all scientific theories (except shallow theories that are merely about observed correlations), this theory uses theoretical terms that cannot be explicitly defined, but are implicitly defined by their role in the theory, as explained here: `http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#models`

The theory as a whole needs to be both enriched internally so as to fill gaps in the explanation, and also "tethered" by more links to experiment, observation, and working models demonstrating its feasibility.
There is still a long way to go.
It should develop faster than Democritus' atomic theory of matter, propounded over 2400 years ago?

The ideas presented here are not new, but their combination may be. Originality is not claimed however.

The combination of ideas was developed in collaboration with Jackie Chappell: see our joint papers listed near the end.

She has not had time to check these slides, and may not agree with everything said here.

We both regard the work as still incomplete.

The theoretical claims need to be made more precise, working models need to be developed, and deeper empirical probes are required, for testing.

How most(?) people think about human language

- It is essentially a means of communication between separate individuals,
  though there are derived mental functions such as planning, reminiscing, theorising, idly imagining.
- 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 acquiring their 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.

I 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 above.

Rival views about evolution of human language

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

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

Rival views about evolution of human language

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

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

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

Rival views about evolution of human language

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

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

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

   Two options

   - externalisation was first gestural
   - externalisation was first vocal

NB: Do not assume such internal representations must be like Fodor’s LOT (Language of Thought).

(For reasons explained later.)

Rival views about evolution of human language

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

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

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

   Two options

   - externalisation was first gestural
   - 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.

Rival views about evolution of human language

1. First there were expressive noises which gradually became more differentiated and elaborate and then were "internalised".

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

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

   Two options

   - externalisation was first gestural
   - 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 came first:

- external human languages?
- internal languages with core properties of human language?

Rival views about evolution of human language

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

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

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

   Two options

   - externalisation was first gestural
   - 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 came first:

- external human languages?
- internal languages with core properties of human language?

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

Rival views about evolution of human language

1. First there were expressive noises which gradually became more differentiated and elaborate and then were "internalised".

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

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

   Two options

   - externalisation was first gestural
   - 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 came first:

- external human languages?
- internal languages with core properties of human language?

WHAT CORE PROPERTIES?

WHAT CORE PROPERTIES?

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

I first focus on three core properties, all of which are required for using language in relation to novel situations.

Human languages support:

- **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.

  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.

- **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).

- **Manipulability:** (a consequence of the previous two)

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

I now explain in more detail how these apply to human languages, before offering generalisations required for the broader concept of a g-language (GL).

Structural variability

Mastering a human language, such as English, includes learning how to create and to interpret novel syntactic structures composed of familiar structural roles and relationships, as well as familiar words.

The structures can be recognised and understood without familiar words.

- The flurgle frimped a flozzle
- The flurgle in the fronkit frimped a flozzle
- The flurgle in the fronkit frimped a flozzle fungily
- The flurgle in the fronkit frimped a flozzle fungily with a fenzikle
- The flurgle in the fronkit frimped a fangile flozzle fungily with a fenzikle
- To fleek the furgrumbs, the flurgle in the fronkit frimped a fangile flozzle fungily with a fenzikle
- To fleek the furgrumbs, the flurgle in the fronkit frimped a fangile flozzle fungily with a fenzikle, but they just fenged.
- To fleek the furgrumbs, the flurgle in the fronkit frimped a fangile flozzle fungily on its flookbud with a fenzikle, but they just fenged.

Individuals can go on indefinitely both learning new words and learning new modes of composition into which to fit the words. (e.g. recent horrible uses of ‘was like’, ‘is like’)

The notion of structural variability will later be generalised.
Next we consider compositional semantics, both in a standard form and a generalised form.

Standard compositional semantics

Conventional compositional semantics:

New combinations of words, phrases and clauses 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.

It turns out that that does not account for all uses of linguistic complexity.

Generalised compositional semantics

Conventional compositional semantics:

New combinations of words, phrases and clauses 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.

It turns out that that does not account for all uses of linguistic complexity.

We need:

Generalised (context-sensitive, situated) compositional semantics:

New combinations of words, phrases and clauses 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

Formally, we can 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).

More on generalised compositional semantics

Generalised (context-sensitive, situated) compositional semantics:

Meanings of complex wholes are determined by three things:

(a) meanings of parts,
(b) the way the parts are assembled to form the whole, and
(c) the linguistic and non-linguistic context (obviously true for indexicals, e.g. “this”, “here”, etc.)

Examples:

- “Put it roughly there.”

  You don’t have to be told exactly where, and there is no semantic rule determining the location.

  Not even a probability distribution.

- “If you can’t see over the wall, find a big box to stand on.”

  You don’t have to be told exactly how big – use your understanding of what the box is wanted for.

- “The wind will blow the tarpaulin away so let’s put a pile of stones on each corner.”

  You don’t have to be told how many stones make a pile, and no semantic rule says how many: but you can work out that there must be enough at each corner to keep the tarpaulin down when the wind blows, and that will depend on how strong the wind is.

(I suspect the above summarises what is right in some obscure ideas about “situatedness”.)

The role of context in compositional semantics seems to be a generalisation of Paul Grice’s “Cooperative principle” and his “Maxims of communication”, to include internal languages.

I.e. the role of context is important for internal languages too. For more on this see:

Spatial prepositions as higher order functions: and implications of Grice’s theory for evolution of language. `http://www.cs.bham.ac.uk/research/projects/cosy/papers/#dp0605`

Third core property: manipulability

An important feature of many forms of representation, based on the first two properties, is that the representations are manipulable: new ones can be constructed either by modifying or combining old ones or by creating modified copies of old ones.

When using a human language, a set of sentences describing some situation can be modified by adding or removing whole sentences, by extending some of the sentences, by omitting some parts of sentences, by replacing some parts with others.

These are all discrete changes. Later we'll see how some GLs also allow continuous changes.

Making inferences

Some modifications of sets of sentences have the feature that if the original set are all true, then the modified set will also be true. Aristotle explored a subset of cases, and modern logicians have codified many more.

More generally, there are modifications of information structures that preserve some aspect of the information.

The special case of modifications that preserve truth, can be used for making inferences.

Other special cases can be used for other purposes, e.g. making plans, constructing explanations, composing music.

NOTE: In Fregean languages, e.g. aspects of external human languages and most logical and mathematical languages, the only manipulations are discrete changes, whereas in some other forms of representation (including spatial representations) continuous changes are also possible. (Examples later.)

Many researchers consider only manipulations that change numerical values: this is very narrow-minded.

Manipulability and use 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

Charles Babbage and Ada Lovelace understood the key ideas before 1850, and fragments of the technology required were developed before and after that, in automated looms, calculators, music boxes, card-sorters, and mechanical toys.

Independently of technological developments, mathematicians and logicians, e.g. G. Frege and C.S. Peirce, extending ideas that were centuries old, developed the theory of compositional semantics for formalisms used in logic and mathematics, and some aspects of natural language.

Working mechanisms that could make and use forms of representation with variable structure and compositional semantics were not developed by human scientists and enginers until the 1950s onwards.

However, far more sophisticated and varied forms of representation (including chemical forms) were produced much earlier by evolution, and used for many different biological purposes, including controlling growth of organisms, detection and repair of damage, immune responses, perception and action in a complex environment, self-understanding, etc.

There is much we don't understand about biological mechanisms, including what they do and how they do it.

At present we know very little about how to build such mechanisms except for limited cases, e.g. logical and algebraic (Fregean) representations.

The implementability requirement

Remember the warning earlier about unimplemented models.

- All three of the core properties 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`
- 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?

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.)

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. proposed in Dennett 1969?):
  - 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.

    (Daniel Dennett, *Content and Consciousness*, 1969.)

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

- First I try to bring out some of the possible uses of GLs with the three core properties: structural variability, compositional semantics, manipulability.

  (Later generalised to include spatial – e.g. diagrammatic – forms of representation).

- Then I point to 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" (TOM) vs "mechanistic" cognition). Meta-semantic competence is required for a meta-management architectural layer.

(I've had much help from Jackie Chappell in all of this.)

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:

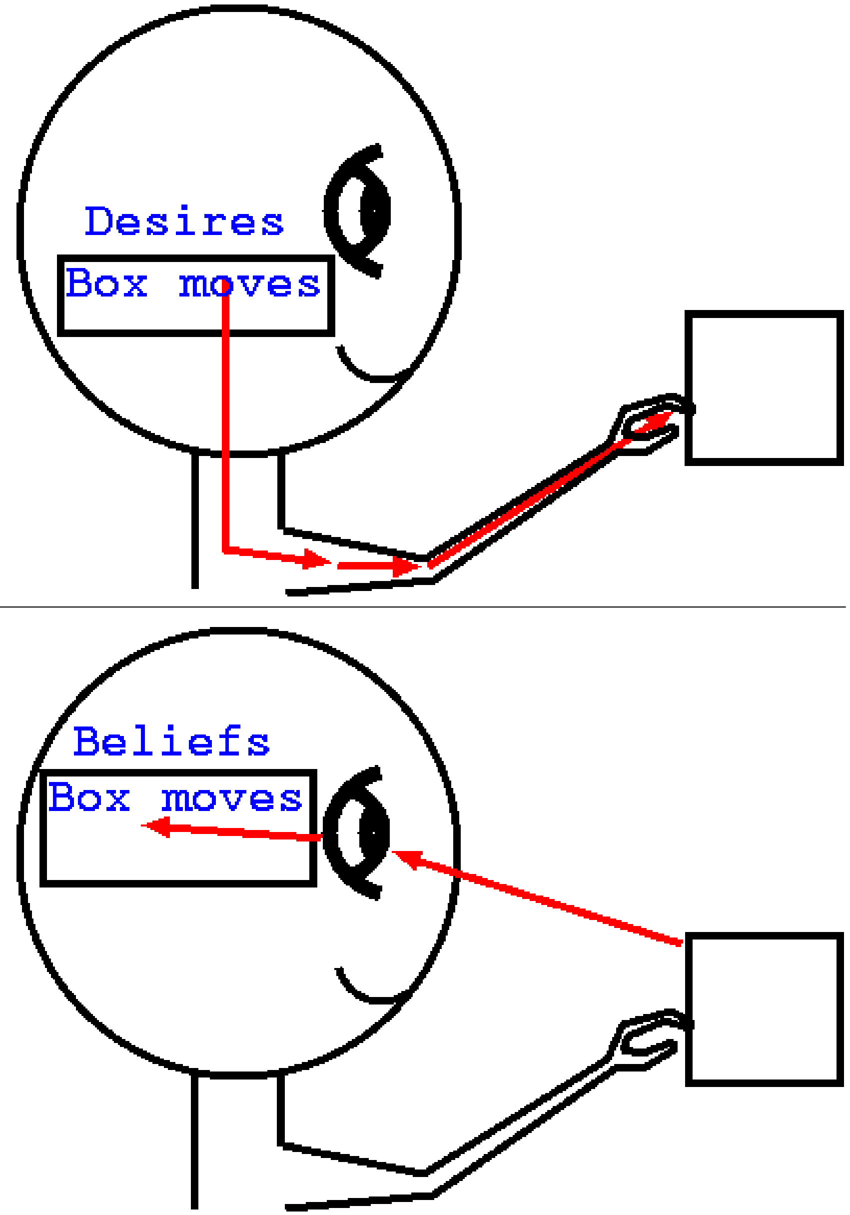
A. Sloman, R.L. Chrisley and M. Scheutz,

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.

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),
- fantasising, 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.

A consequence of the core features

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'll see a similar kind of inconsistency in non-verbal forms, later.

Inconsistency of an information structure implies that

- if it is adopted as a belief it will be a necessarily false belief,
- if adopted as a goal it will be a necessarily unachievable goal, and
- if constructed as a percept it will be a perception of an impossible state of the world.
  (illustrated later)

(Compare G. Frege on failure of reference.)

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.

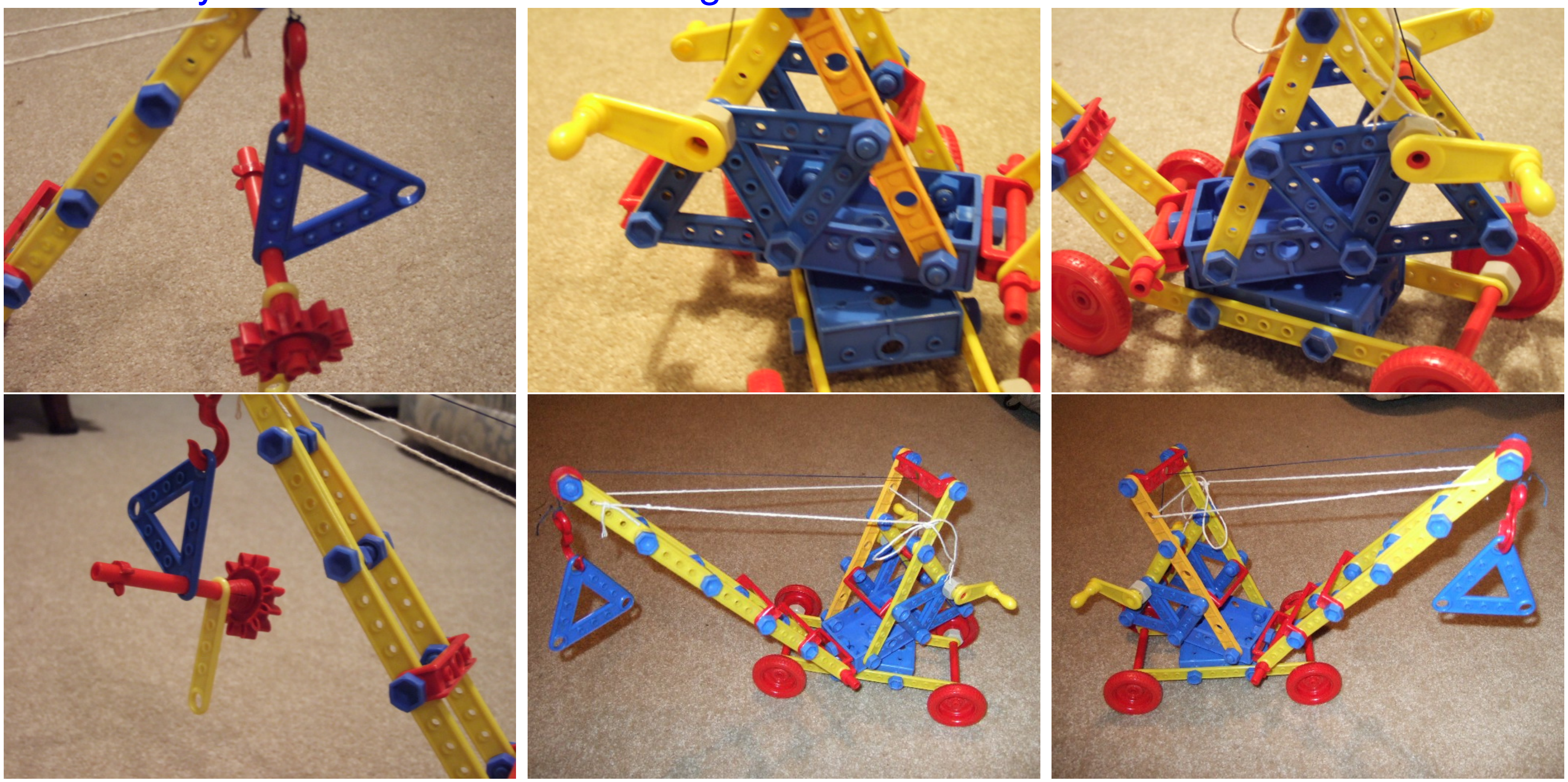
The retinoid model of Arnold Trehub's *The Cognitive Brain* (MIT Press, 1991) may be an example.
`http://www.people.umass.edu/trehub/`

Sloman 1971 (ref. at end) 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 "Logicist" AI approach expounded by McCarthy and Hayes, in 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.

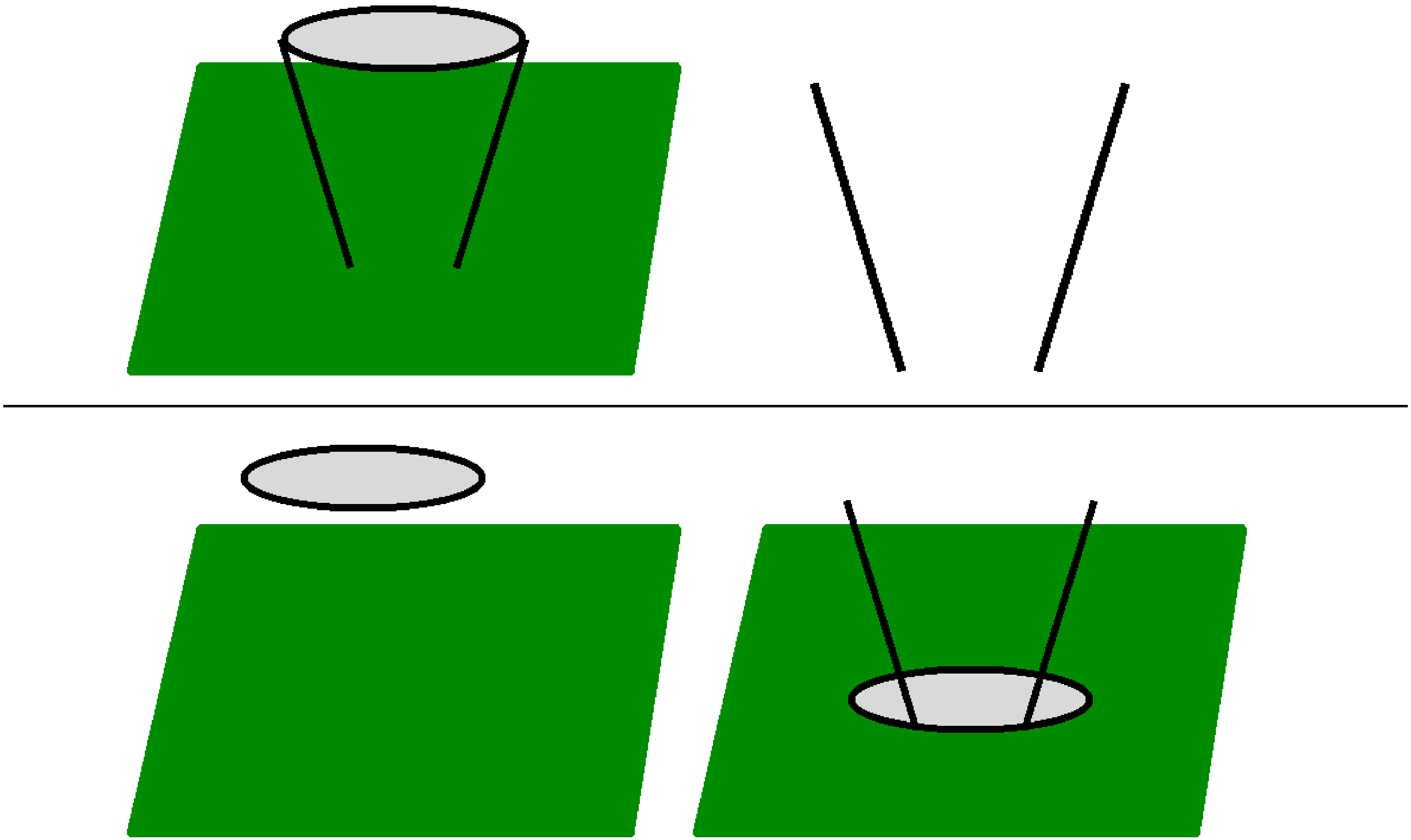


http://www.cs.bham.ac.uk/research/projects/cosy/photos/crane/

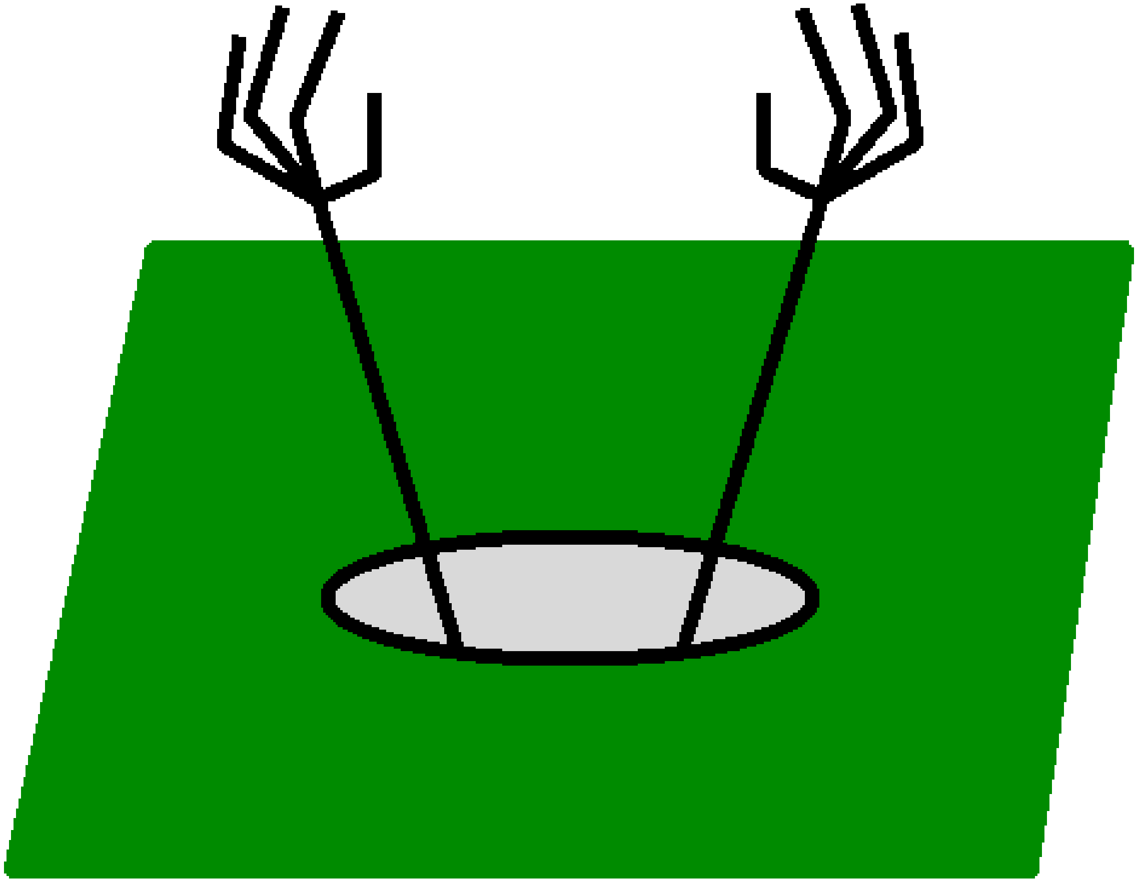
Different combinations of the same elements

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

Notice how context can influence interpretation of parts

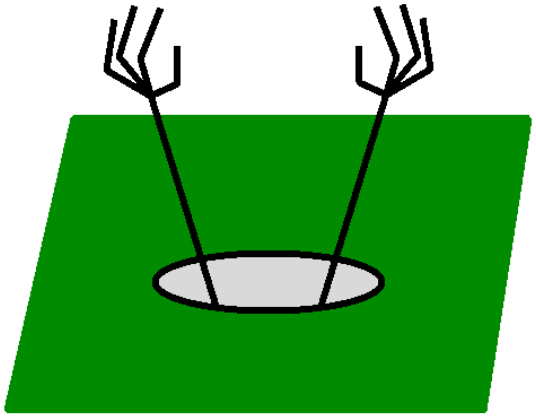


A droodle: Can you tell what this is?



Droodles depend heavily on the fact that construction of visual GL instances can be can be partly driven by sensory data and partly by verbal hints ("top down").

Possible answers



"Early bird catches very strong worm?"

"Sewer worker attacked by sharks?"

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

There is much to be said about droodles, 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.

Notice how sharing a region of space allows indefinitely many different kinds of structural relationships to 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.

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.

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.

Seeing continuous structures and potential motions

In the top picture, you probably have no difficulty seeing string, a nut, a bolt, a ball, and a loop in the string going round the bolt.

You can't be sure it's really flexible string until you see it moved or jiggled.

What will happen if the blue bolt is lifted vertically, with its axis remaining horizontal?

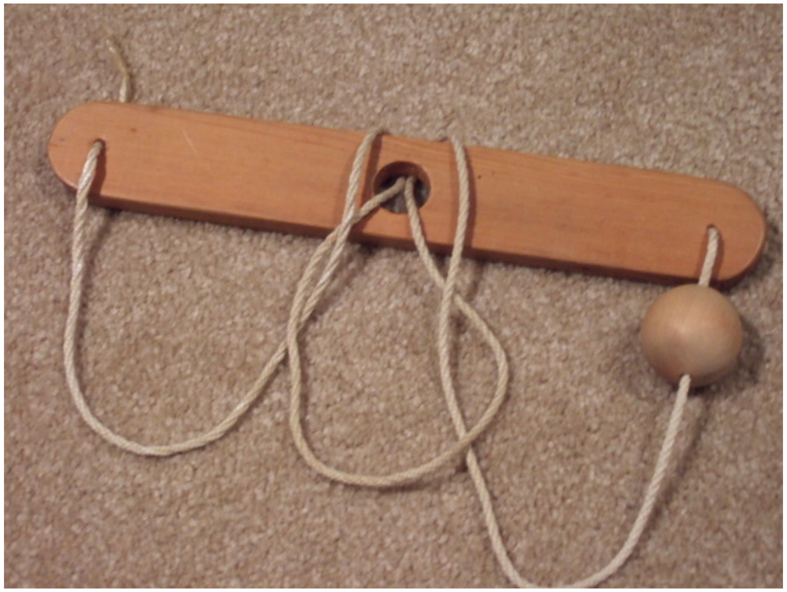
What sort of motion will free the bolt from the string so that it can be lifted without the string?

In the bottom picture you can probably visualise the ball being slid along the string.

Is there a way to get the ball to the other end of the string without cutting the string or undoing the knots at the end of the string (on the far sides of the small holes)?

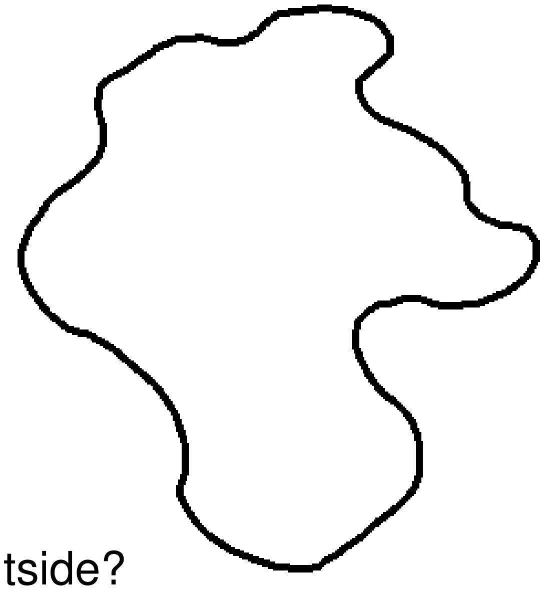
The answer is very hard to work out just by imagining the movement, but you can imagine various possible movements of the ball along the string.

How does your brain, or a virtual machine running on your brain, represent possible movements of perceived objects before they are actually moved? ("Proto-affordances".)



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 a capital “H”?

What sort of GL is needed to make it possible to answer such a question?

- 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 verry young child will not be able either to construct the H physically, or to answer the questions.
- You are probably able to answer the questions just by thinking about the construction processes and the result.
- 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?

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.

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 toally general syntactic conventions (as in language and logic).

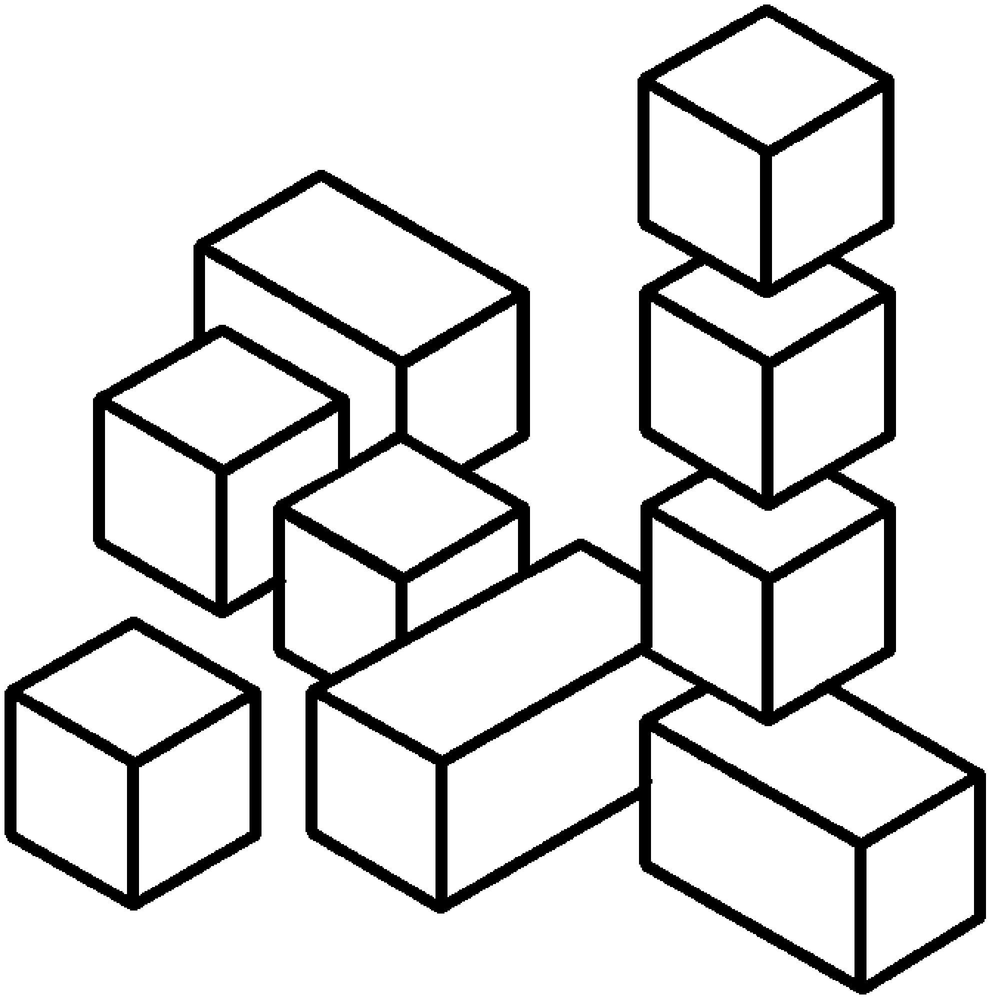
Nevertheless the spatial form of representation is not a structure that is isomorphic with what it represents.

This can be demonstrated using pictures of impossible objects.

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

Building a configuration of blocks - 1

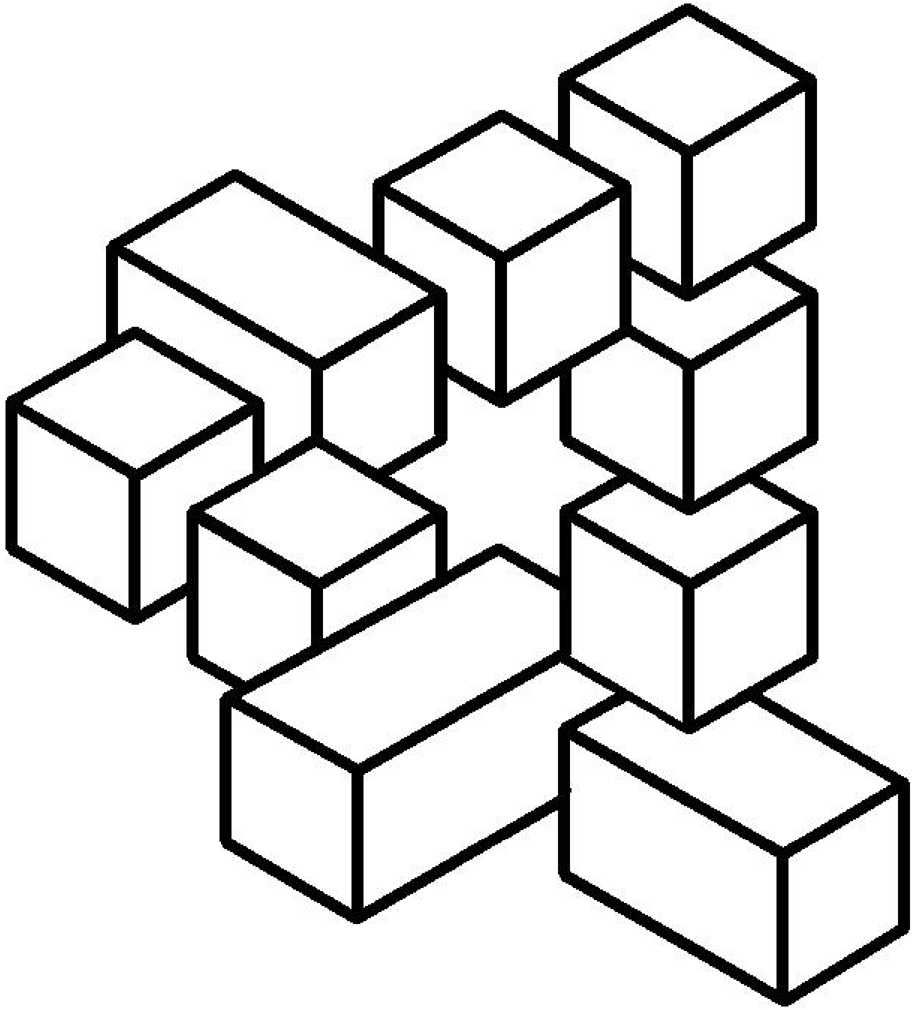
Given a collection of cubes and rectangular blocks could you arrange them to look like this?



Think of locations to which you could you move the 'loose' cube on the left.

Building a configuration of blocks - 2

Moving one cube, could you re-arrange them to look like this?



Some young children will say ‘yes’.
What has to change for them to be able to detect the impossibility?

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`

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.

Not Fodor's LOT

There is no implication in any of this that a human, or nest-building bird, or intelligent language user, must start with an 'innate' (or genetically determined) GL that suffices to express everything it will ever need to express, so that all future meanings are definable in terms of the initial set.

Papers with Chappell investigate ways in which boot-strapping processes can substantially extend innate competences through exploration and play in the environment along with the ability to construct new explanatory theories to account for surprises.

This can include substantial ontology extension: introducing concepts that are not definable in terms of previous ones, e.g. using model-based semantics and symbol/theory-tethering.

For more on that see
http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#models
http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#grounding
http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#grounding
http://www.cs.bham.ac.uk/research/projects/cosy/papers/#pr0702

That option was not open to Fodor because he used a model of understanding based on compiled programming languages, where all programming constructs are translated into the machine language before the programs run.

He apparently forgot about interpreted programming languages and perhaps did not know about logical programming languages (e.g. prolog).

He should have known about model-theoretic semantics, but failed to see its relevance, as described in the presentations listed above.

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.

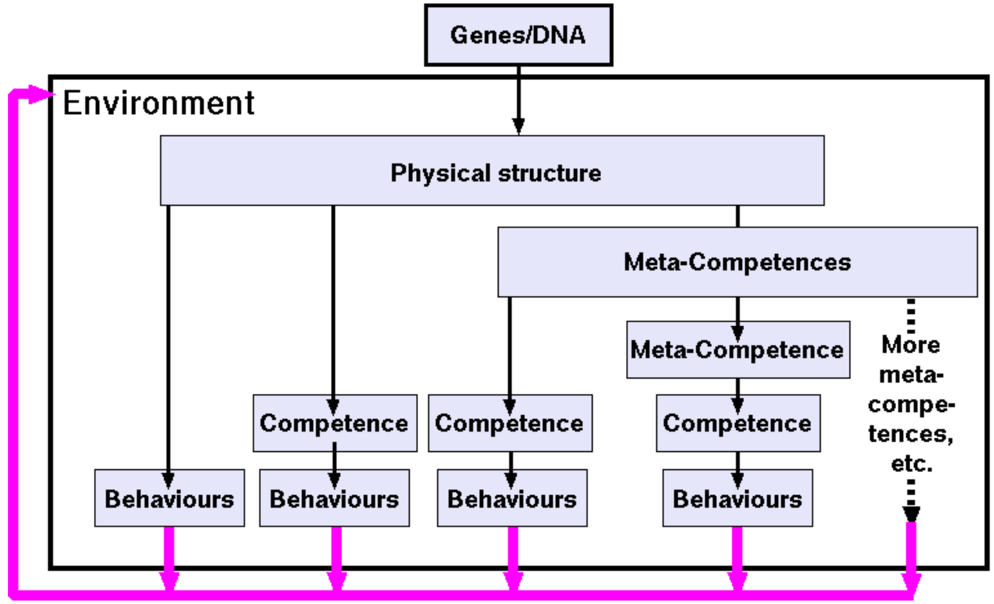
I suspect there were a few key discrete architectural changes, that modified the forms of learning and development in humans and other altricial species (see below).

Routes from genome to behaviours

Cognitive epigenesis

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

- during development of seed, egg, or embryo, and subsequent growth (i.e. it is 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 the 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 on the left. Diagram is from the IJUC paper with Jackie Chappell. Chris Miall helped with the diagram.

The construction of some competences should be construed as an ongoing process, with repeated activation of the meta-competence 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) of the processes will involve continuous control.
In deliberative and some meta-management components much will be discrete.

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.

Language learning vs language development

If the previous observations and speculations are correct, 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 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.

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, 1965) 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 GL structures were 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 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 GLs 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.

The evolutionary heritage of gestural GLs

It has often been remarked that at least three remarkable facts about humans suggest that we still retain brain mechanisms that originally evolved in connection with external gestural GLs.

- It is hard for people to talk without gesturing, often highly creatively, even when they are talking on the phone to people who cannot possibly see the gestures and who do not need them – as shown by the usefulness of telephones.
- Some children with Down's syndrome find it easier to learn a sign language than to learn to talk normally.
- Nicaraguan deaf children rapidly developed a new sophisticated sign language which they used very fluently and which their teacher was unable to learn.

Moreover, if all human children develop and use rich internal GLs before they learn to talk (orally or by signing), then what we used to think of language learning should be thought of as language extension since they already have rich linguistic (GL) capabilities which they extend for communicative purposes.

Nicaraguan children showed that that should be thought of as a collaborative, creative process of developing a means to communicate, rather than a process of doing data-mining or induction on information collected from the environment.

In most cases the child learners are a small minority, and politically weak, so language creation looks deceptively like language learning.

Many unsolved problems

These slides scratch the surface of many very deep and difficult problems.

In particular, I have ignored the fact that very little is 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.

Some of the points were made at the BBSRC-funded Conference in Birmingham May-June 2007 on

Closing the gap between neurophysiology and behaviour: A computational modelling approach

See this presentation:

`http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#compmod07`
Architectural and representational requirements for seeing processes and affordances.

Background to this presentation

The slides are partly based on this BBS paper (in Press), which introduced the term 'g-language', now 'GL'.

Aaron Sloman and Jackie 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: Eva Jablonka, Marion J. Lamb,

Evolution in Four Dimensions:
Genetic, Epigenetic, Behavioral, and Symbolic Variation in the History of Life (MIT Press, 2005)

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

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 WONAC (Workshop on Natural and Artificial Cognition), Oxford 2007, here:
http://www.cs.bham.ac.uk/research/projects/cogaff/talks/wonac

---

Much earlier, less developed, versions of some of the ideas were in these two papers, both 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,

More references (still incomplete)

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/l_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