# Part 3:a view from above

Keith Stenning
Human Communication Research Centre

Alan-Bundy's-Book-on-Edinburgh Talk

1 March 2006

# Thinking and doing

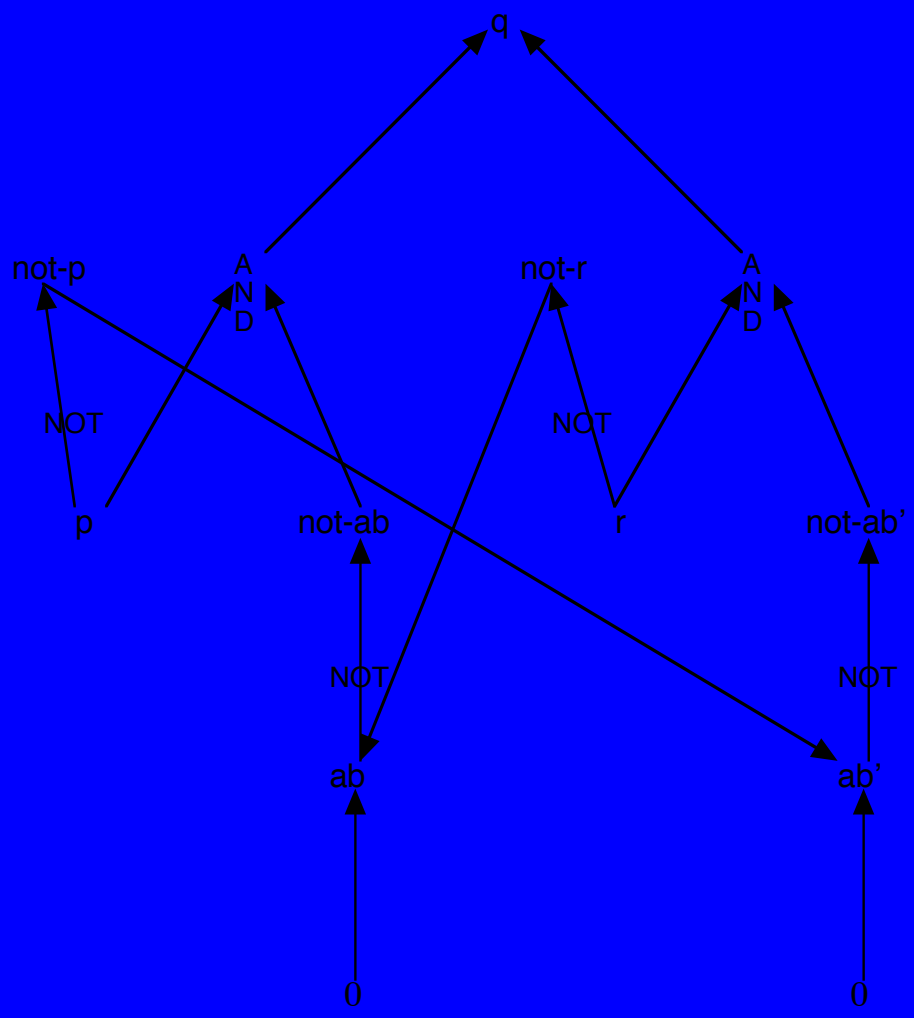
- human reasoning is most often defeasible, in a logic like this:

  $p \wedge \neg ab \rightarrow q, r \wedge \neg ab' \rightarrow q\ , \neg r \rightarrow, ab\ p \not\models q$

- in English it may sound something like this:

  *If she has an essay to write she will study late in the library.*
  *If the library is open, she will study late in the library.*
  *She has an essay to write.*
  *So she may not study late in the library because it may be shut.*

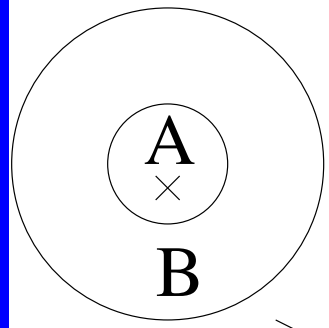
- implemented in a spreading activation network it looks like this:


q
not-p
AND
not-r
AND
NOT
NOT
p
not-ab
r
not-ab’
NOT
NOT
ab
ab’
0
0

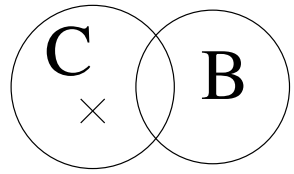
# Thinking and doing (continued)

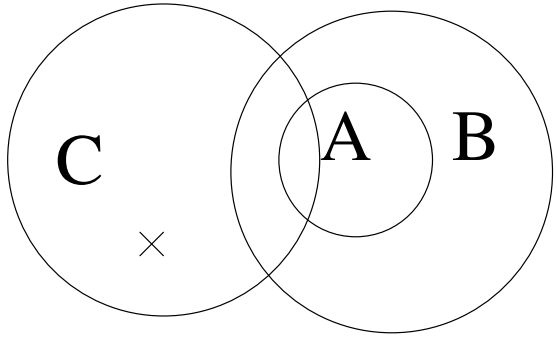
- looking at the network level, reasoning is just electrical signals
- plausibly the same logic of planning underlying motor control
- offers evolutionary continuity between action and communication
- and thought bears many traces of its origins in perception/action
- but the causal level alone doesn't decide if it's representational/symbolic/'higher'/...

All A are B




Some C are not B




Therefore, some C are not A

# Never mind the mapping—feel the analogy!

Karla, an old hawk, lived at the top of a tall oak tree. One afternoon, she saw a hunter on the ground with a bow and some crude arrows that had no feathers. The hunter took aim and shot at the hawk but missed. Karla knew the hunter wanted her feathers so she glided down to the hunter and offered to give him a few. The hunter was so grateful that he pledged never to shoot at a hawk again.

Once there was a small country called Zerdia that learned to make the world's smartest computer. One day Zerdia was attacked by its warlike neighbour, Gagrach. But the missiles were badly aimed and the attack failed. The Zerdian government realised that Gagrach wanted Zerdian computers so it offered to sell some of its computers to the country. The government of Gagrach was very pleased. It promised never to attack Zerdia again.

Once there was an eagle named Zerdia who nested on a rocky cliff. One day she saw a sportsman coming with some bird food that had no nuts in it. The sportsman offered the food but there was no nuts. Zerdia realised that the sportsman wanted her to have nuts so she flew down and donated a few of her own nuts to the sportsman. The sportsman was pleased. He promised never to come with nutless birdfood again.

# Tactics for 'representation': broad church or narrow?

- representational/computational processes constitute minds in two layers:
  - specifications of the functions computed by the (sub)organism of interest
  - implementations of those functions
- under this broad definition:
  'Behaviour X is a representational/computational process' is a very weak claim
- 'computation' for the cognitive scientist is a bit like 'life' for a biologist—all the work comes after
- "How is Sonic Hedgehog expressed in adult neurogenesis in the rat's frontal lobes?"
- "How do we construct discourse representations from WM text and LTM knowledge and how are they implemented?"

# Why keep computation/representation so general?

- my coauthors (I think) want to use 'representation' to subdivide the area
- what I do is representational: what Barbara does is not
- but ideas of computation are so dominated by symbol-shuffling: higher mental 'logical' processes = string shuffling: lower mental processes don't need computation
- and this view lets in the mongol hordes (more later)
- we share a goal (avoid a certain glib interpretation of computational models of mind) but we have different tactical views (inclusive/exclusive uses of rep/comp), perhaps tailored for different critics

# Turing and the computational mind

- "the mind is a universal Turing machine"—what could this mean?
- Turing computation is the right framework for specifying *what* is computed
- implementations of Turing computations are computational
- the mind is to be characterised by relations between the two
- so the mind is a computational phenomenon
- literally a 'computer', as long as we remember that the thing on the desk with Intel inside is only a computer (in this sense) because of its relation to an interpretation as an *abstract* machine

# Misguided objections to this 'Turing' model of mind

- 'rocks are computers'
- 'people can't compute all Turing-computable functions'
- 'Turing computation is symbolic but mind is subsymbolic'
- 'the failure of GOFAI showed that that architecture is hopeless'
- 'it can't do transducers'
- 'it can't do analogue computation'
- 'simple critters are just causal'
- 'it can't do timing, embodiment, . . . '
- 'it can't do quantum computation'
- . . .

# The theorist and the punter—computation is a theorist's toy?

- once we have the implementation, it's all causal, so what's the Turing specification got to do with anything?
- but the computational level *is* causally involved:
  - in humans one can often show that the punter has access
  - or there is a learning history of 'internalisation' of the function
  - in systems with ontogeny, functional specification of the maturation/learning 'goals'
  - in evolved systems, functional specification of the selection pressures

# The causalists (of Mongolian extraction)

- find a critter, freeze all but 3 parameters—a dynamical system
- we can write differential equations so it's science
- but unfortunately it's not biology (yet)
- biology is about the whole organism in its environment(s)
- self defense tips on meeting members of the Mongol hordes:
  - change the motivational state of the creature
  - change the epistemic history of the creature
- when close enough to see the whites of their eyes, they are, of course, a varied (and charming) lot

# Mind is a two-layered cake

- anyone who tells you he can do without either layer has forgotten the cream is the relation between the two layers
- he's often a physicist for whom two layers smacks of dualism
- but the Turing computation model defuses the penalities of dualism—by providing a criterion for causal implementation
- computation is a property of the whole system—not CPU, not brain, not individual without environment, not individual without social group, . . .
- 'hot cognition' as well as 'cold'—emotions are also computational phenomena

# The real problems with this classical model

- the hard work is just as much about getting an insightful specification as finding an (the right) implementation
- the real working problem with the classical model is its power
- it's so easy to give really bad specifications and it's hard to distinguish equivalent specifications, or implementations
- which is why it's important to realise that saying the mind is computational is a weak claim—leaving all the work still left to do