Lecture 11:

Consciousness:

What is it ?

How is it biologically implemented ?

How could it be implemented in a robot ?

Wolfgang Maass

Institut für Grundlagen der Informationsverarbeitung

Technische Universität Graz, Austria

Different ways of defining consciousness

Purves et al., p. 706: „Most definitions of consciousness refer to three different aspects of consciousness:"

- being awake

- being aware of the world

- being aware of oneself

# Different ways of defining consciousness: more precisely

- As opposite of being unconscious (clinical), e.g. the ability to respond to a stimulus *(any functioning robot whose actions depend on sensors is conscious in this sense)*

- Characterisation via the un-ambiguous nature of a sensory perception at any moment in time *(easy to implement in a robot with top-down control, but very difficult to achieve in a distributed network without central control units)*

- E.g. in vision: characterisation of conscious experiences as those that can be recalled *(could easily implemented in a robot; e.g. if the robot has a „ball detector" and a memory, it could record all ball detections).*

- Characterisation of a conscious experience through the sensation of particular features (e.g. the color „red", or more general „qualia") and associated emotions (C. Koch)

- Consciousness as a characteristic feature of the „mind" (relating it to the old mind-body problem)

- Characterisation through „self consciousness", e.g. the ability to reflect about some experience or action *(can be implemented in a robot with a suitable short-term memory structure)*

- Characterisation in the context of „free will" (although it is not clear from what the will is assumed to be „free")

Ambiguous images: We cannot maintain different interpretations simultaneously; rather our interpretation tends to flip between them

A
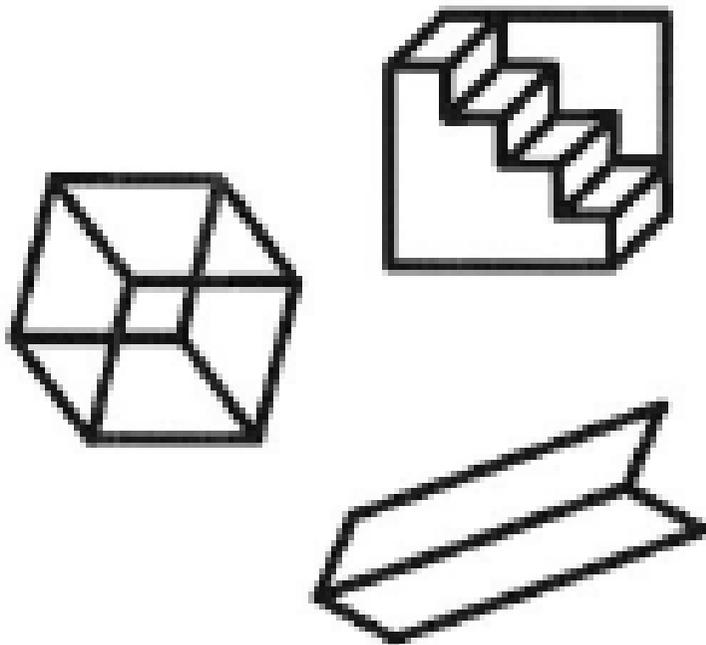
B
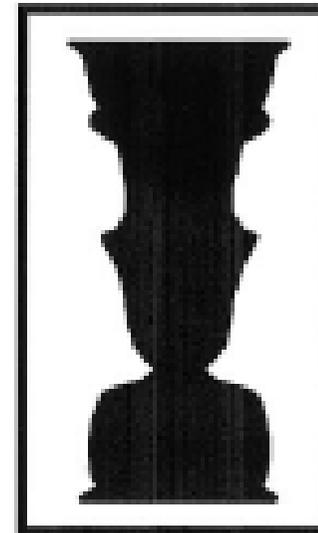
D. A. Leopold and N. K. Logothetis. Multistable phenomena: changing views in perception. *Trends Cogn. Science,* 3(7):254-264, 1999.

Percentages of neurons in different areas of visual cortex whose firing changes with the current interpretation of the image (while the visual stimulus remains unchanged)