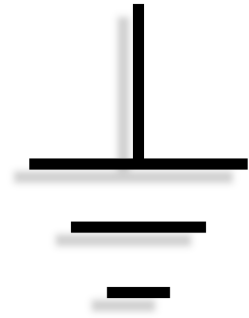


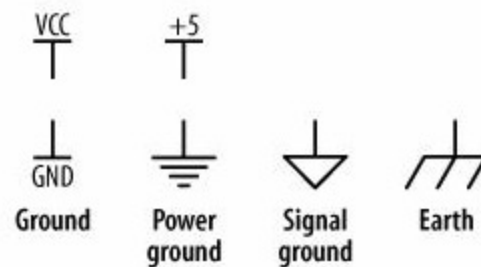
# The Grounding of Symbols

---



Reading Club *Kognitive Systeme*

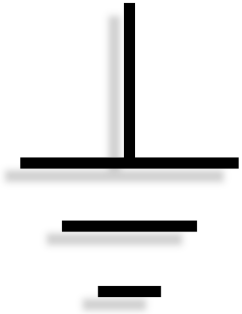
Sommersemester 2010



International Ground Symbols

# Motivation

---



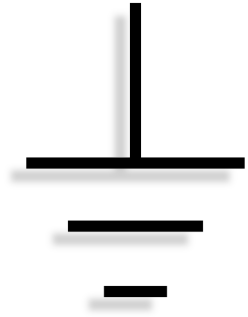
„How can the semantic interpretation of a formal symbol system be made intrinsic to the system, rather than just parasitic on the meanings in our heads?“

„The problem is analogous to trying to learn Chinese from a Chinese/Chinese dictionary alone.“

„How is symbol meaning to be grounded in something other than just more meaningless symbols?“

# Agenda

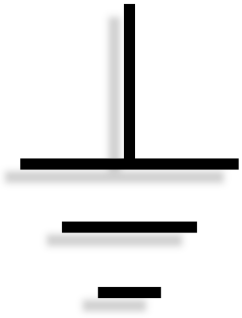
---



- Harnad's starting point
- The meaning of *meaning*
- What *is* a symbol system?
- What *does* a symbol system?
- 15 years later ...
- Example: Sun and Heidegger
- In the end ..?

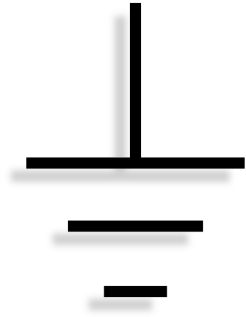
# Harnad's starting point

---




- The (in-) famous *symbol grounding problem*
- Formulated by **Stevan Harnad** 1990  
(Physica D 42, 1990, p. 335-346)
- How can „semantic interpretation [...] be made *intrinsic* to a system“?  
  
→ Meaning has to arise within the system

# The *meaning* of meaning



- Webster:

Main Entry: **mean·ing** 

Pronunciation: \ˈmē-nɪŋ\

Function: *noun*

Date: 14th century

**1 a** : the thing one intends to convey especially by language : **PURPORT** **b** : the thing that is conveyed especially by language : **IMPORT**

**2** : something meant or intended : **AIM** <a mischievous meaning was apparent>

**3** : significant quality; *especially* : implication of a hidden or special significance <a glance full of meaning>

**4 a** : the logical connotation of a word or phrase **b** : the logical denotation or extension of a word or phrase

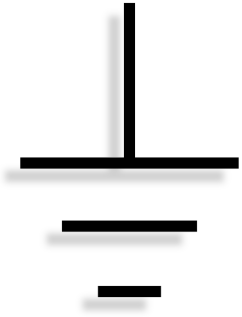
— **meaning** *adjective*

— **mean·ing·ly**  \-nɪŋ-lē\ *adverb*

→ What does this *mean*?

# The *meaning* of meaning

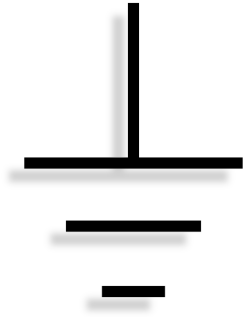
---



- Frege: reference to an object (Gegenstand)
  - Wittgenstein: „Die Bedeutung eines Wortes ist sein Gebrauch in der Sprache“
  - Aristoteles: „Das Seiende wird in mehrfacher Bedeutung ausgesagt, aber immer in Beziehung auf Eines und auf eine einzige Natur und nicht nach bloßer Namensgleichheit (homonym)“
  - C.G. Jung: rooted in unconcious archetypical forces
  - Peirce; de Saussure;... and many more
- no common definition

# The *meaning* of meaning

---

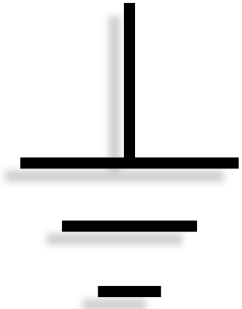


- A pragmatic view<sup>1</sup>:
- A *sign* is clearly *pointing* to some object (denotative), an unambiguous reference
- A symbol (think of religious or archetypical symbols) goes far beyond:
  - Denotative and representative
  - Not fully explainable
  - Linking many objects
  - Might be highly ideosyncratic

<sup>1</sup> This definition is my own, and according to a WWW search, some might share this point, while others might completely disagree. However, some common **ground** is necessary when talking about *meaning*.

# The *meaning* of meaning

---



- So we might rephrase Harnad's question:

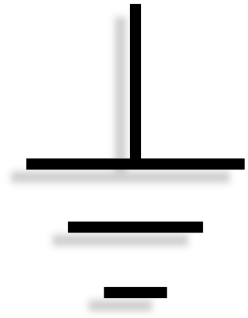
„How can a system form connections between objects in the world (and inner states) which have the power to qualitatively grasp key features of these objects and which reflect the system's experience with these objects?“

→ reference to Heidegger: *being-in-the-world* as ideosyncratic experience, inextricable connection between subject and world



# What is a symbol system?

---

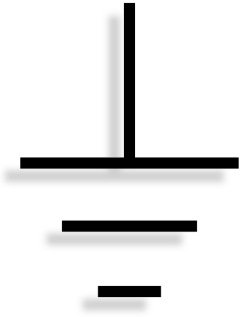


- arbitrary „physical tokens“
- „explicit rules“ for manipulating
- that are also tokens/ strings of tokens
- purely syntactic (shape)
- rulefully combining
- primitive atomic tokens
- composite symbol-token-strings
- semantically interpretable

→ Harnad, based on Newell & Simon (1976)

# What is a symbol system?

---



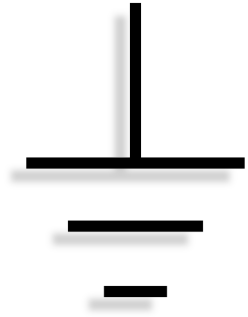
- Core assumption:

*The symbolic (mental) level is completely independent from its physical representation*

- Only makes sense as a system quality
- Often criticized:
  - Penrose: mind relies on quantum computations
  - Winograd/Flores: structural coupling

# What does a symbol system?

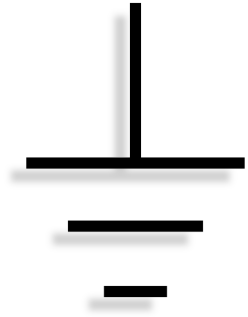
---



- 1) iconic representations (sensory)
- 2) categorial (feature detection)
- 3) symbolic representation

# What does a symbol system?

---

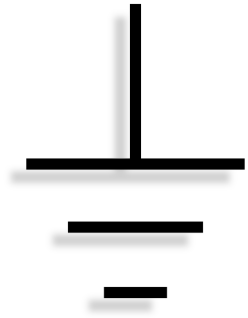


## 1) iconic representations (sensory)

- Closely related to input
- e.g. retinal activation pattern
- cf. definition of *sign*

# What does a symbol system?

---



2) categorial (feature detection)

→ cf. *machine learning*

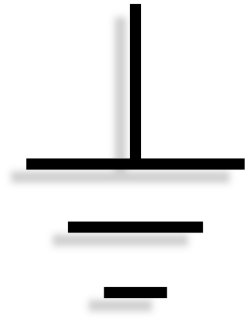
→ only needs some sort of bias/ restriction

- decision tree
- k-nearest-neighbors
- neural network

→ Harnad: still no semantics necessary

# What does a symbol system?

---



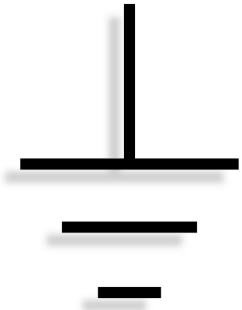
## 3) symbolic representation

→ rulefully combination

- grounded names „strung together into propositions about further category membership relations“
- „Zebra“ = „horse“ & „stripes“

# What does a symbol system?

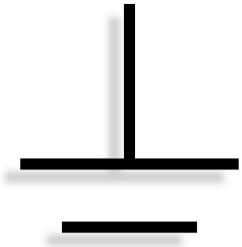
---



- Harnad's proposal – hybrid architecture:
    - „the `connection´ between the names and the objects that give rise to their sensory projections and their icons would be provided by connectionist networks“
    - „intrinsically dedicated symbol system“
  - „there is really only one viable route from sense to symbols: from the ground up“ (cf. Dörner)
- problem solved?

# 15 years later ...

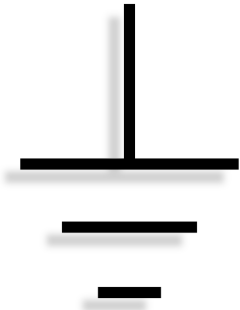
---



- Taddeo & Floridi: Solving the Symbol Grounding Problem: —  
a Critical Review of Fifteen Years of Research  
(IEG – Research Report, 2005, Oxford University and University of Bari)
- „Zero Semantical Commitment Condition“
  - no form of innatism
  - no form of externalism



# 15 years later ...



- Pro

Innatismn/ externalism always bears the risk of *parasitic* meaning that was not generated by the system itself

ZSCC avoids this risk

→ *Discussion*

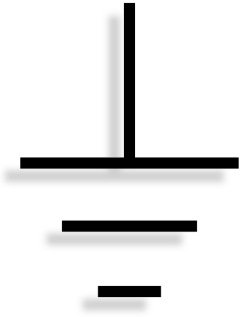
- Contra

Natural systems have evolved over millions of years and accumulated information – even genetic information is a mirror of phylogenetic experience (cf. Bischof 2009) and thus a form of *nurture*. Culture passes high-level symbolic information (→ *meme* theory)

When a living system can't be regarded without its environmental coupling (Dörner, Winograd/Flores, Heidegger, Bischof), an artificial system somehow must compensate for this lack of experience

# 15 years later ...

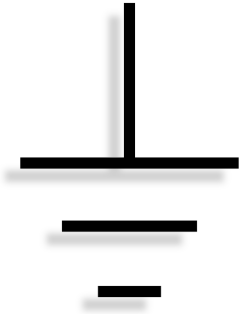
---



- However: Good overview of existing approaches for solving the SGP
  - Hybrid model (Harnad): NN + symbol system
  - Functional semantics for categorization (Mayo)
  - Intentional grounding (Sun) → *details follow*
  - Learning by observation and examples
  - Grounding in sensorimotor activities (Vogt)
  - Communicative validation
  - ... and mixtures

# Example: Sun and Heidegger

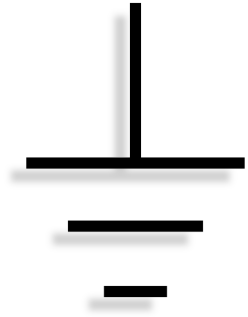
---



- **Symbol Grounding: A New Look At An Old Idea**  
(Sun, *Philosophical Psychology*, Vol.13, No.2, pp.149-172. 2000)
- Symbols grounded in subsymbolic activities *and* „the direct interaction between the agent and the world“
- Linked to goals, needs and actions
- No internal/external dichotomy (NB: different from the notion of embodiment!)

# Example: Sun and Heidegger

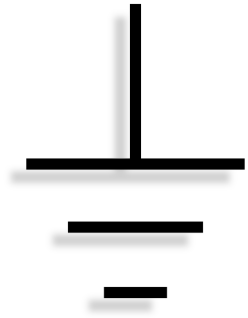
---



- Dual-level-architecture
  - conceptual and sub-conceptual
    - trial-and-error + reinforcement
    - intermediate neural network
    - rule learning algorithms
      - inference: forward and backward chaining
  - context-dependent and task-oriented concept formation (model: CLARION)

# In the end ..?

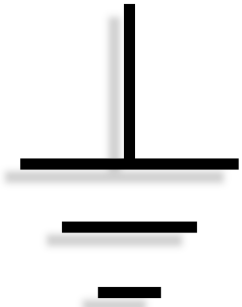
---



- Is there a solution to the SGP?
- Think of Cartesian dualism:
  - a mind/body dichotomy is only a problem if the dichotomy is taken for granted
  - Dreyfus, Winograd & Flores, Bischof, Dörner, Heidegger, Sun, ...
  - problem or solution?

# In the end ..?

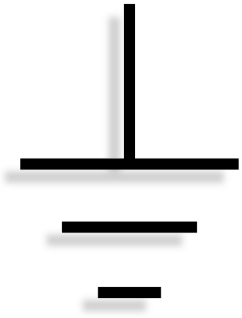
---



- These criticisms point out that a system must be regarded as coupled with its environment
- They stress the notion that beliefs and intentions are important features of a grounded system
- They tell us that millions of years of phylogenetic evolution and thousands of years of cultural evolution provide living systems with essential grounding mechanisms

# In the end ..?

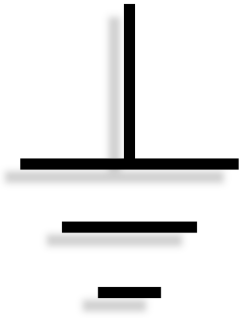
---



- No learning is possible without input
  - No (machine) learning is possible without biases and restrictions
  - No machine learning is possible without a machine
- inputs, restrictions and machines must be carefully designed

# In the end ..?

---

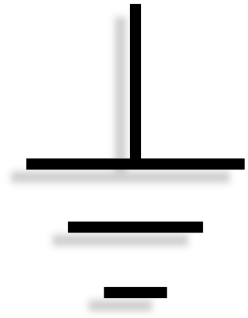


- What is *carefully*?
  - in accordance with psychological theories
  - in accordance with philosophical insights
  - in accordance with biological features
  - *always reflecting these assumptions*



# In the end ..?

---



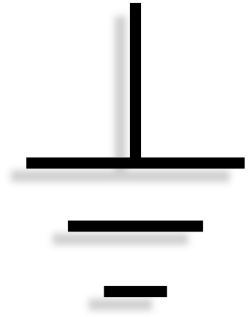
- The next step:

In my bachelor thesis I will (together with Mark Wernsdorfer) design a system that, although in many respects rudimentary, will be *carefully designed*

On top, the inductive learning system IGOR will be used to provide this autonomous agent (AA) with the ability to detect regularities in its sensory inputs and internal states

# In the end ..?

---



- So ...

→ by utilizing IGOR our AA will generate rules that mirror its interaction with the environment

→ rules might be considered as high-level connections (i.e. abstract symbols)

→ the plausibility of IGOR as a cognitive device will be critically evaluated

– *finis* –