Hofstadterian Architecture

Antti Halme

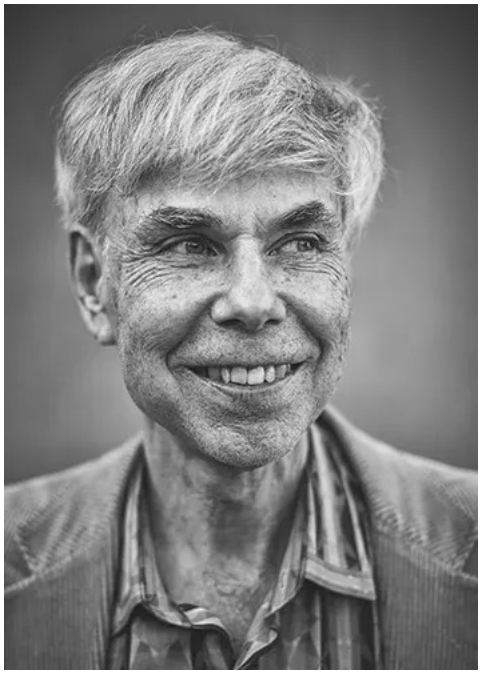
Online
02 December 2020

Agenda

- Hello (18:45)
- **Introduction** (19:00)
- **Hofstadterian Architecture**
- Terminology
- Principles of HA

  ~ BREAK ~

- **Components of HA**
- ***Discussion***
- Finish (21:00-21:30)



Douglas Hofstadter
***F**luid **A**nalogies **R**esearch **G**roup*

Introduction

- **Hofstadterian Architecture (HA)**
  - Family of related software systems
  - Replicate or simulate aspects of human cognition
  - **Analogies, Fluid Concepts**
  - Microdomains
  - Randomness and parallel processes → emergence

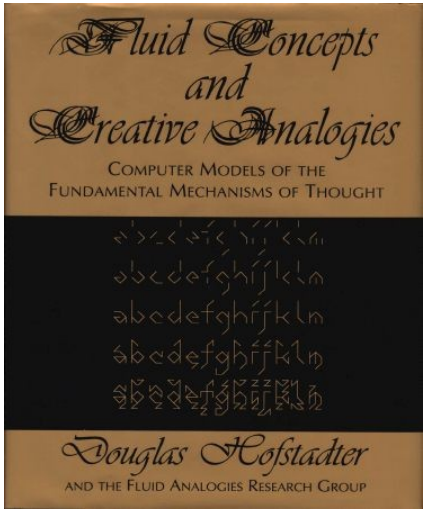
→ Model the way humans seem to think

→ High-level perception in controlled environments

Fluid Concepts and Creative Analogies (1995)

Computer Models of
the Fundamental Mechanisms of Thought

- A collection of papers with chapter commentaries
- The origin and development story of a “cognitive” software system architecture
- Span from late 1970s to early 1990s
- Details in the group’s PhD theses



***FARG**: Melanie Mitchell, David Chalmers, Robert M. French, ...*

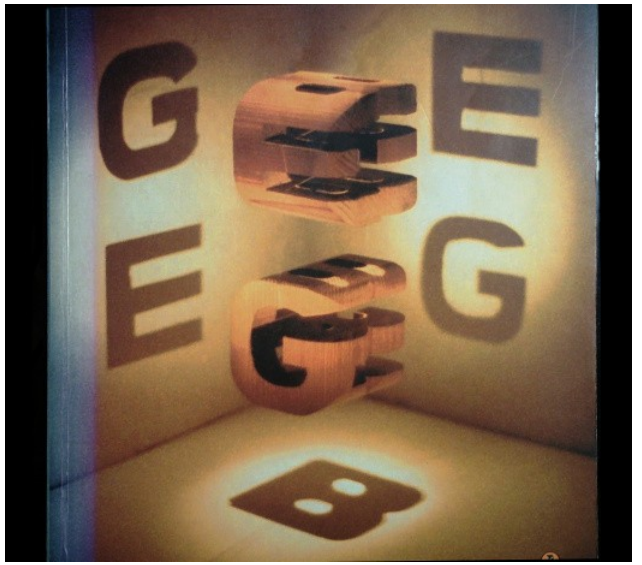
Hofstadterian Architecture

aka “active symbol architecture”, “FARGitecture”, “parallel terraced scan”, ...

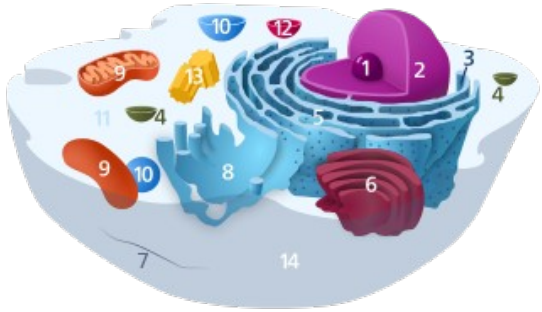
- **A stochastic meta-heuristic multi-agent search method**
  - “search”: find a match to a query, retrieve/compute data
  - “heuristic”: *informed* search, with selective prioritisation
  - “meta-heuristic”: higher levels of organisation
  - “multi-agent”: multiple “search heads” working in parallel
  - “stochastic”: randomness, “shades of grey”, baked in

Hofstadterian Architecture

- **GEB**
  - Analogy-making
  - Emergence, recurrence, meta



- **Cellular metabolism :: Cognition**
  - Cytoplasm :: “construction site”
  - DNA as strings, proteins as statements, …



- Human puzzle solving: a jumbled mess

Terminology

- **Perception**
  - Process of making sense of an environment
  - Finding the signal in the noise, the order in noise
- **Representation**
  - “The fruits of perception”
  - Raw data turned into useful structures by a process of filtering and organisation

NB: These are *my* definitions / formulations of FARG terms

Terminology

- **Concept**
  - Notionally the Platonic ideal
  - The essence of a thing as informed by available representations
  - Also: a point-like node embedded in a graph of its peers, complete with a *halo* of connectivity

→ **Fluid concepts** : concept graph dynamics lead to flexible behaviour on part of the concepts themselves

Terminology

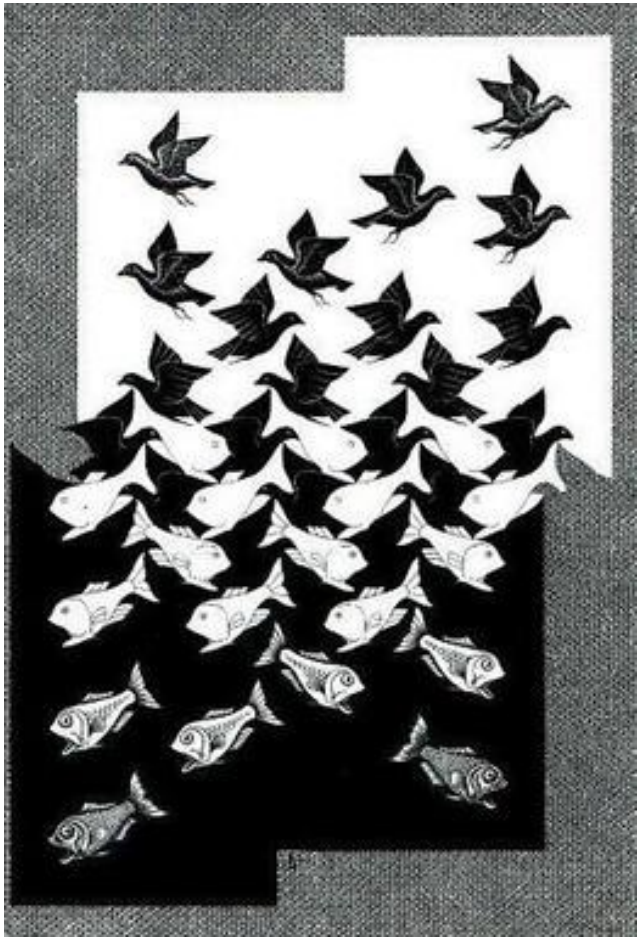
- **Analogy-making**
  - Process of making new connections between concepts
  - Concept expansion

- **Cognition**
  - Making use of the concept graph
  - Perception from the bottom up and analogy-making from the top down

HA in a nutshell

- ▼ **Analogy-making**
  Concept
  Representation
- ▲ **Perception**

COGNITION



Principles of HA

- **Inseparability of perception and high-level cognition**
  - Focusing on perception will help us make progress on the puzzle of cognition
- **Gradual build-up of representations**
  - Representations are generated and developed through a sequence of small, incremental changes powered by observations of the environment

Principles of HA

- **Easy reconfiguration of representations**
  - Cognitive representations operate at multiple levels simultaneously
  - Various kinds of links hold representations together
  - High-level perception can be thought of as the process of maintaining and upgrading these structures.
- **Influence**
  - Concepts have influence over other concepts, as does the broader context
  - This top-down control complements the bottom-up generative process

Principles of HA

- **Sub-cognitive pressures**
  - Concepts and representations have varying degrees of relevance / importance
  - Concepts are in constant competition with one another and some end up exerting greater influence than others
- **A blend of pressures**
  - Context-dependent and context-independent pressures at play at the same time
  - Co-existence facilitates simultaneous bottom-up and top-down processing

Principles of HA

- **Integration**
  - Perception and concept processing and analogy-making are all functionally integrated and intertwined, rather than separate modules in a linear pipeline
- **Concurrent execution**
  - Many search "heads" are in operation at the same time
  - Randomness baked into the system helps the search branch out
  - Agents distinguish themselves and may exhibit varying rates and modes of operation as the system evolves

Principles of HA

- **Centrality of analogy-making and variations on a theme**
  - High-level cognition is analogy-making
- **Varying degrees of sensitivity to pressure**
  - Some representations are deep and immune to context
  - Others are shallow and capable of changing shape under pressure

Principles of HA

- **Transience and versatility of representations**
  - Representations can take many shapes and influence concepts in many ways
  - Representations come and go, concepts tend to stick around
- **Crucial role of the inner structure of concepts**
  - "Conceptual neighbourhoods", the distance and overlap between concepts, and conceptual depth all contribute to the overall task
  - Concepts have dimensionality

Example: Copycat

*If I change the letter-string **abc** to **abd**, how would you change **ijk** in "the same way"?*

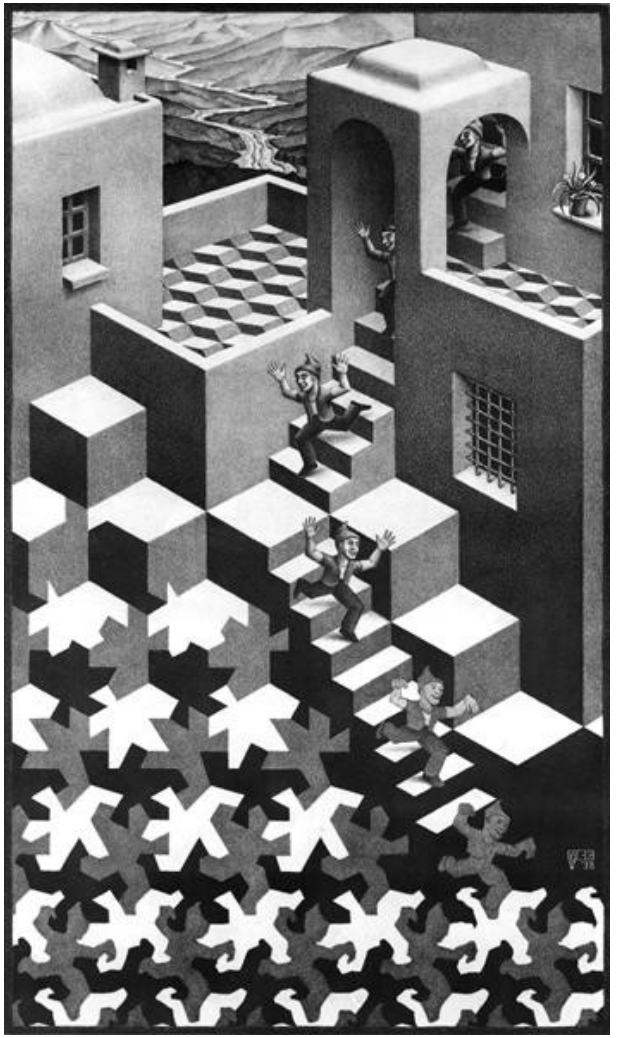
*If I change the letter-string **aabc** to **aabd**, how would you change **ijkk** in "the same way"?*

- **Analogies in the letters-strings microdomain**
- From rightmost *letter* to rightmost *group* is an example of conceptual fluidity
- The *letter* concept **slipped** into *group* under pressure
- Thanks to the accessibility of the microdomain, this is something that software systems might be able to learn to do

BREAK

Components of HA

- **Entities**
- **Slipnet**
- **Workspace**
- **Coderack**
- **Parallel Terraced Scan**
- **Pressures and Biases**
- **Temperature**



Entities

- FARG were primary interested in conceptual slippage
  - The particular micro-domains did not matter that much
- Grounding in what humans find interesting
  - We would like to *understand* what the system is doing, so we teach it *our* concepts
  - e.g., the fact that letter ‘t’ is elevent letter behind ‘i’ in the alphabet is not interesting to humans
- Innate concepts
  - No concepts were derived from data/environment
  - Copycat had some 60 concepts

Slipnet

- **The home** of all permanent concepts
- Long-term memory, repository of categories
- Types, not instances
- Not just storage, but a dynamic structure
  - A network/graph of nodes and their *halos*
  - *Links* connect nodes, measure of “conceptual distance”
  - Changes shape the graph to “fit” a given situation
- Node activation
- Conceptual depth (a priori)

Workspace

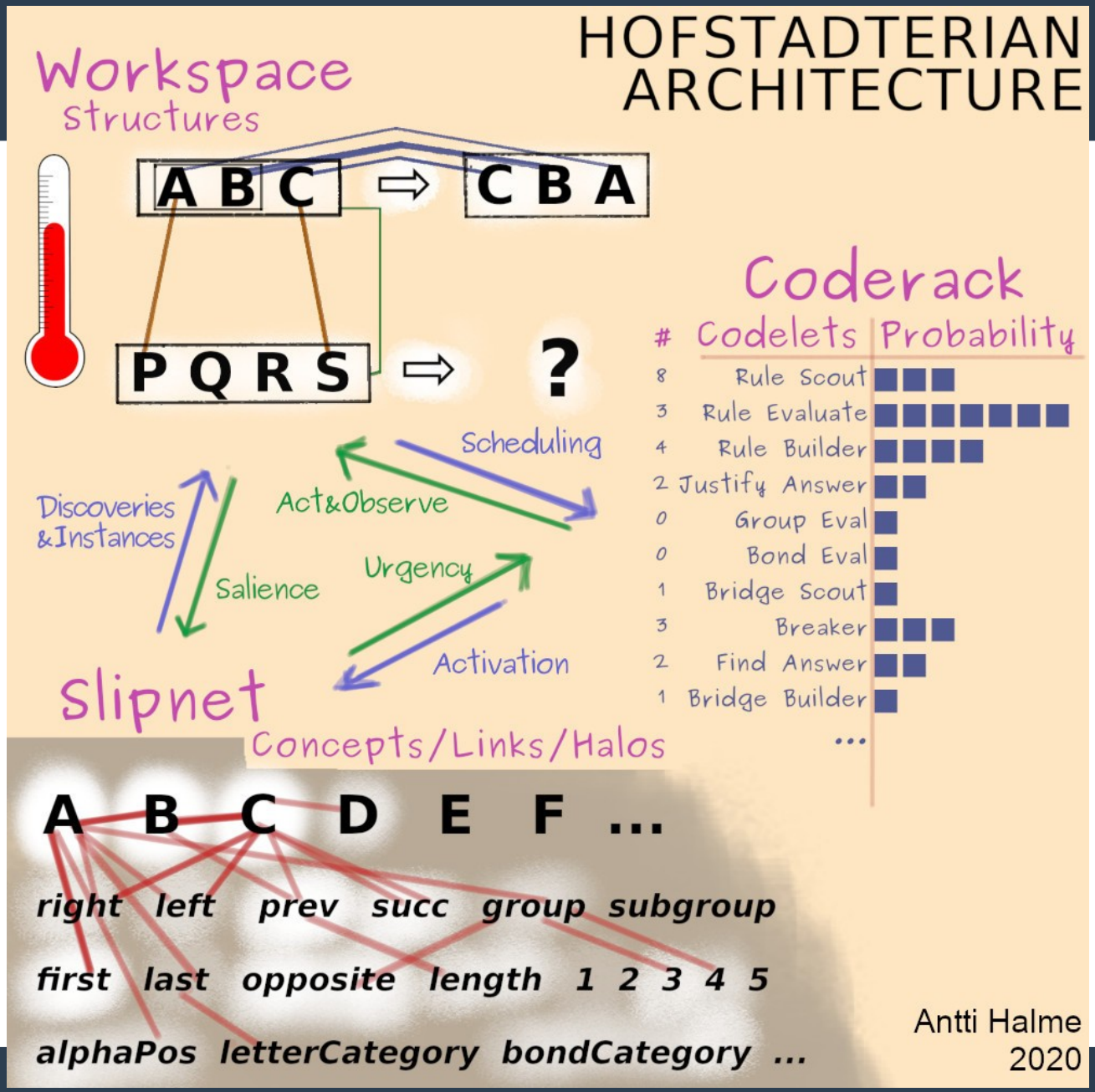
- **The construction site**
  - Where concepts are instantiated and put to work
  - The “locus of perceptual activity”
  - A busy, concurrent building environment ~ cytoplasm
- Temporary perceptual structures
  - Raw data from env + application of Slipnet concepts
- *Salience = importance + “unhappiness”*
  - Intrinsic and relational components determine what gets built in the construction site
- Bond-making, chunking, mapping
- Emergent *viewpoint*

Coderack

- **The waiting room** for “codelets”, Workspace operators
  - A repository for simple software agents that each contribute a small step to the overall process
- *Scouts* look for actions to do, *effectors* make it happen
- Codelet *urgency* establishes a stochastic priority order
- Bottom-up and top-down agents
  - Codelets are proxies of the system’s current pressures
  - No codelet makes a huge impact alone, pressures emerge
- Continual codelet replenishment

Big Picture

- **Complex feedback**
- **Stochastic processes**
- **Emergent pressures**



Parallel Terraced Scan

- **The overall algorithm**
- Many “fingers” carry out the search in parallel
  - Guided by aggregate pressures
  - Not actual objects, but manifestation of scout codelets
  - Varying rates of execution, a “commingling”
- Analogies
  - Mind: unconscious processes, unitary consciousness
  - Colony of ants: “feelers” on the fringes, latter exploration stages behind the vanguard

Pressures and Biases

- **Conceptual fluidity is an emergent phenomenon**
- Codelets transmit pressures, pushing and pulling
  - Pressures are probabilistic, the shape of evolving concepts
- Natural arc to the search
  - From general to situational, from local to global
  - Concepts pick up *themes* and become *biased* over time
  - Bias influences search directions
  - Early volatility settles down to more deterministic behaviour
- Nobody is in charge, no predetermined high-level process

Temperature

- **A measure of the system's "open-mindedness"**
  - How orderly is the Workspace?
  - Low temperature, high perceived structural quality
  - In a typical run the temperature goes both up and down
- Serves a feedback role, not just a randomness control
- The search for *paradigm shifts*
  - Temperature works a metric for evaluating program executions!
  - Final temperature ~ proxy for answer quality
- Sometimes sharp focus is needed to overcome snags
  - Open-mindedness needed to tackle conceptual blocks that prevent a situation from evolving in trend direction

Beyond Copycat

- **Copycat was “too unconscious”**
  - Kept hitting snags, or dead ends, repeatedly
- Need to develop **self-monitoring**
- Creative minds have special powers
  - A keen sense of what is interesting
  - Ability for recursion
  - Meta-level awareness
  - Ability to adjust accordingly

→ **METACAT** + creativity-focused FARG systems

FARG Systems (partial list)

- **Jumbo** by Douglas Hofstadter (1983): Unscramble a jumbled word, solve an anagram.
- **Numbo** by Daniel Defays (1988): Assemble a target number from given integers using arithmetic operations.
- **Seek-Whence** by Marsha Meredith (1986): Determine the next value or pattern in a number sequence.
- **Copycat** by Melanie Mitchell (1990): Pairs of letter strings.
- **Tabletop** by Robert French (1992): Looking down at objects on a virtual dining table, interaction analogies with everyday objects.
- **Letter Spirit** by Gary McGraw (1995) and John Rehling (2000): Creative typeface design.
- **Metacat** by James Marshall (1999): Copycat, but with introspection.
- **Phaeaco** by Harry Foundalis (2006): Bongard problems.
- **SeqSee** by Abhijit A. Mahabal (2009): Integer sequence extrapolation, Copycat meets Seek-Whence.
- **Musicat** by Eric Nichols (2012): Simulation of the process of listening to a simple melody.

FARGonautica

- **Software repository maintained by Alex Linhares**
  - https://github.com/Alex-Linhares/FARGonautica
- Attempt to collect and organise FARG systems and research into a single document repository
- Code samples, even full sources, for many FARG systems
  - Shared *as is*
  - Many systems written in ancient, defunct LISP dialects
  - GUIs...

Discussion

- **Summary**
  - Hofstadterian Architecture: A family of related software systems
  - Replicate or simulate aspects of human cognition
  - **Analogies, Fluid Concepts**
  - Microdomains
  - Randomness, parallel processes, feedback → Emergence
  - Perception, Representation, Concept, Analogy-making, Cognition
  - Principles
  - Slipnet, Workspace, Coderack
  - Temperature; Pressures and biases
- **Questions**
  - Is it all just convoluted meta-heuristic soup?

Images

- Douglas Hofstadter portrait by Massimiliano Sticca. https://www.sticca.it/
- Book covers, https://cogsci.indiana.edu/
- *Animal cell* -graphic, Wikimedia
- M. C. Escher, *Sky and Water II* (1938), woodcut
- M. C. Escher, *Cycle* (1938), lithograph