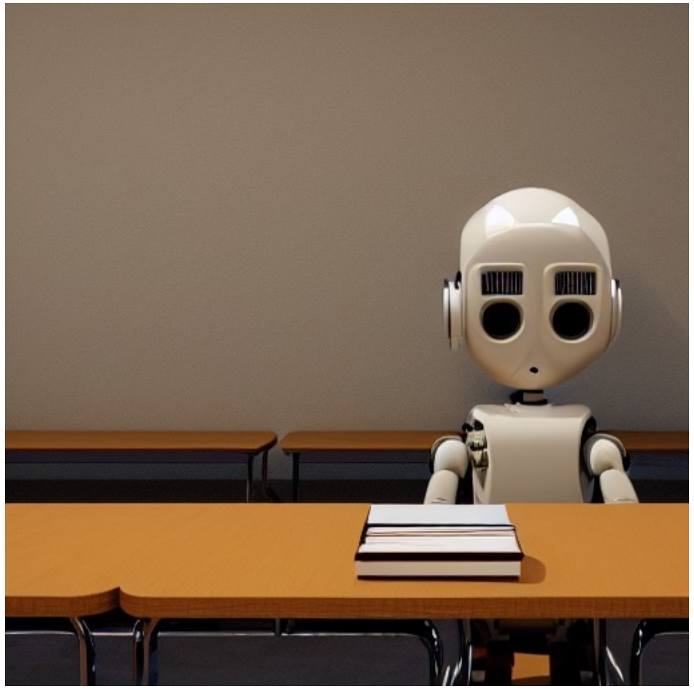
Workshop on Computing across the Continuum

Cognitive Agents, Plausible Reasoning and Scalable Knowledge Engineering

Dave Raggett <dsr@w3.org> W3C/ERCIM



*Learning is so hard, sigh!**

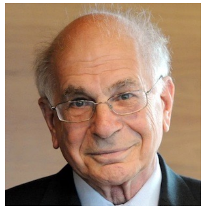
14 November 2022

This work has been supported through funding for the TERMINET project from the European Union's Horizon 2020 research and innovation programme under grant agreement No 957406.

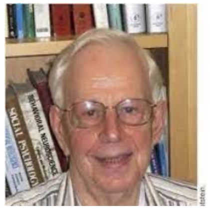
* image generated by an AI asked to imagine a confused robot in a school room

Some acknowledgements

We owe these pioneers a huge debt for their hard won insights



Daniel Kahneman, Nobel prize winning psychologist who studied System 1 & 2 thinking, along with cognitive biases in "Thinking Fast and Slow".



Philip Johnson-Laird, cognitive scientist renowned for his work on how humans reason in terms of mental models rather than using logic and statistics.

Dedre Gentner, cognitive and developmental psychologist renowned for her work on analogical reasoning.

Allen Newell, researcher in computer science and cognitive psychology renowned for his work on AI and the Soar cognitive architecture.

John R. Anderson, cognitive scientist renowned for his work on the ACT-R cognitive architecture for sequential cognition (System 2).

Allan M. Collins, cognitive scientist renowned for his work on plausible reasoning and intelligent tutoring systems.

George Lackoff, cognitive linguist and philosopher renowned for his work on conceptual metaphors in language and cognition.

Lotfi Zadeh, mathematician and computer scientist renowned for his work on fuzzy reasoning and control.

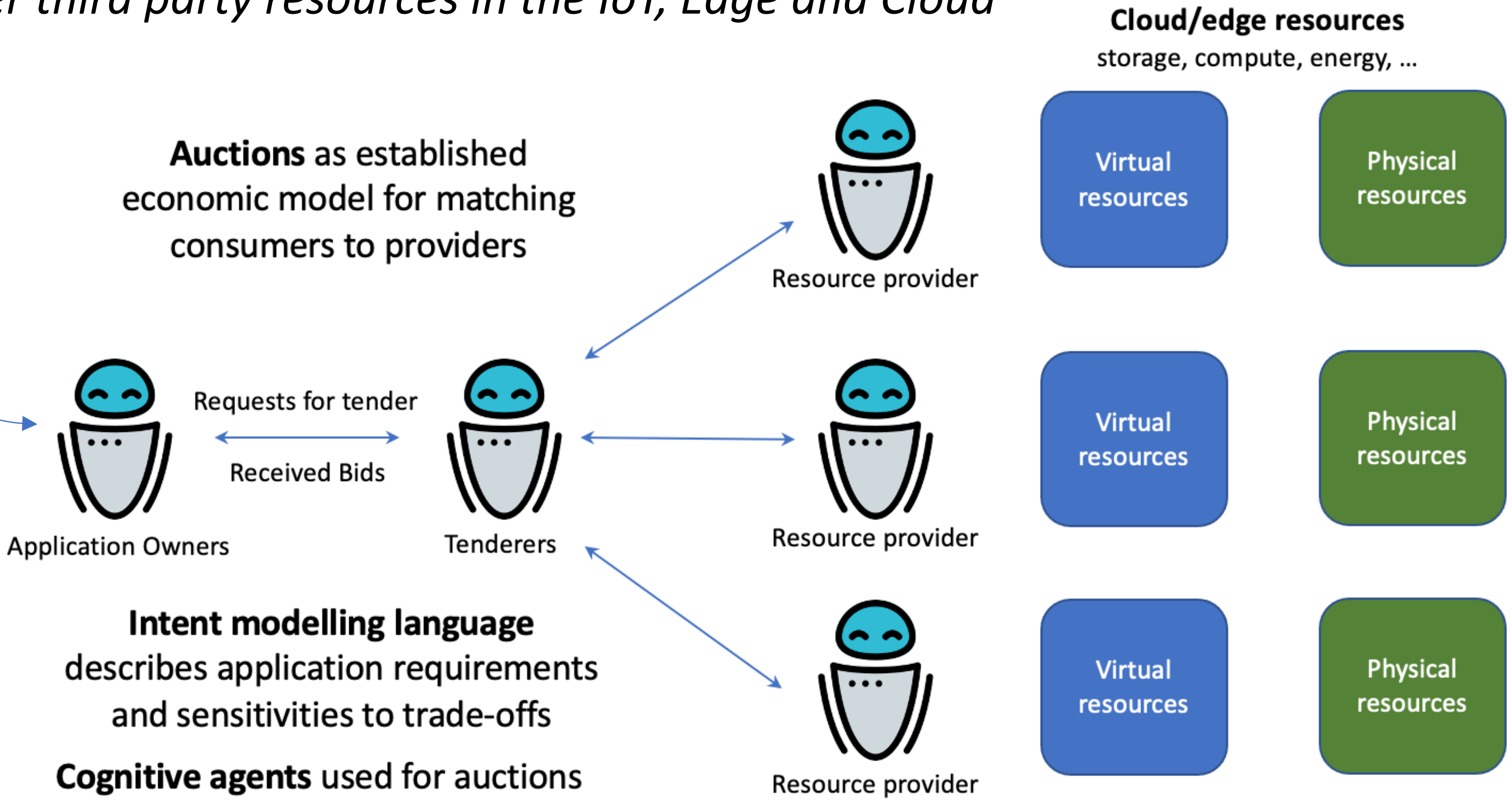
Applying Cognitive Agents to Efficient Resource Management & Orchestration

Application owners seeking to run their applications over third party resources in the IoT, Edge and Cloud

Cognitive agents combine perception, reasoning and action, loosely akin to Amazon Alexa and Apple Siri

Auctions as established economic model for matching consumers to providers

Cognitive agents working on behalf of their users



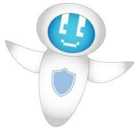
Intent modelling language describes application requirements and sensitivities to trade-offs

Cognitive agents used for auctions and throughout service life cycle

Robot icon courtesy of Pixel Perfect

Cognitive agents also assist with cyber-defence, hand in hand with Soft Defined Networking

Privacy Centred Ecosystems of Services



Digital Guardian Angels as Personal Assistants

- Today's Consumer Web is dominated by advertising-based business models
- Strong focus on gathering personal information for targeting adverts through live auctions
- Consumers are habituated to click away annoying permission requests for enabling tracking
- It is time to make privacy a central part of ecosystems of services
- Personal Assistants that act on their user's behalf in respect to providing services using ecosystems of third party providers
- A privacy centred evolution away from dominance by Web search engines
- Personal Assistants apply their user's values, as learned from observing their behaviour via privacy protecting federated machine learning
- Personal Assistants select matching services using service metadata plus independent trust attestations* and live auctions
- Personal Assistants share pertinent personal information, e.g. their user's travel plans and preferences when seeking proposals for flights, hotels, local travel, restaurants, museums, etc.
- Smart notifications for services that take significant time to fulfil
- Personal Assistants are involved in downstream checks on use of personal data subject to the agreed terms and conditions
- Gets more complicated as data is progressively transformed and merged with other sources of personal information
- Reliant on advances in human-like reasoning with everyday knowledge and natural language understanding and generation

* Open questions around business models, e.g. for personal assistants and for attestations based upon aggregating feedback from users and personal assistants

The huge strides in AI of the last decade have been in machine perception*, but *not* in machine reasoning

Bradley Allen (2018)

- Whilst today's knowledge graphs claim to capture knowledge there is very little attention currently to machine reasoning
  - One exception is inheritance down class hierarchies
  - Conventional logic has limited use
- Application behaviour is instead embedded in application code
  - This makes it hard to understand and costly to update
- *But why do we put up with this?*

- Knowledge presumes reasoning and is otherwise just information!
  - Information is structured labelled data, such as column names for tabular data
  - Knowledge is *understanding* how to *reason* with information
  - In practice, knowledge is often imperfect and imprecise
  - It is high time to focus on machine reasoning for human-machine cooperative work
  - Boosting productivity and compensating for skills shortages

* and its inverse – *machine generation*

Plausible Reasoning with Imperfect Knowledge

Everyday knowledge is subject to uncertainty, incompleteness and inconsistencies: we're learning all the time, and consequently our current knowledge is imperfect

During the 80's Alan Collins and co-workers developed a theory of plausible reasoning* based upon recordings of how people reasoned. They found that:

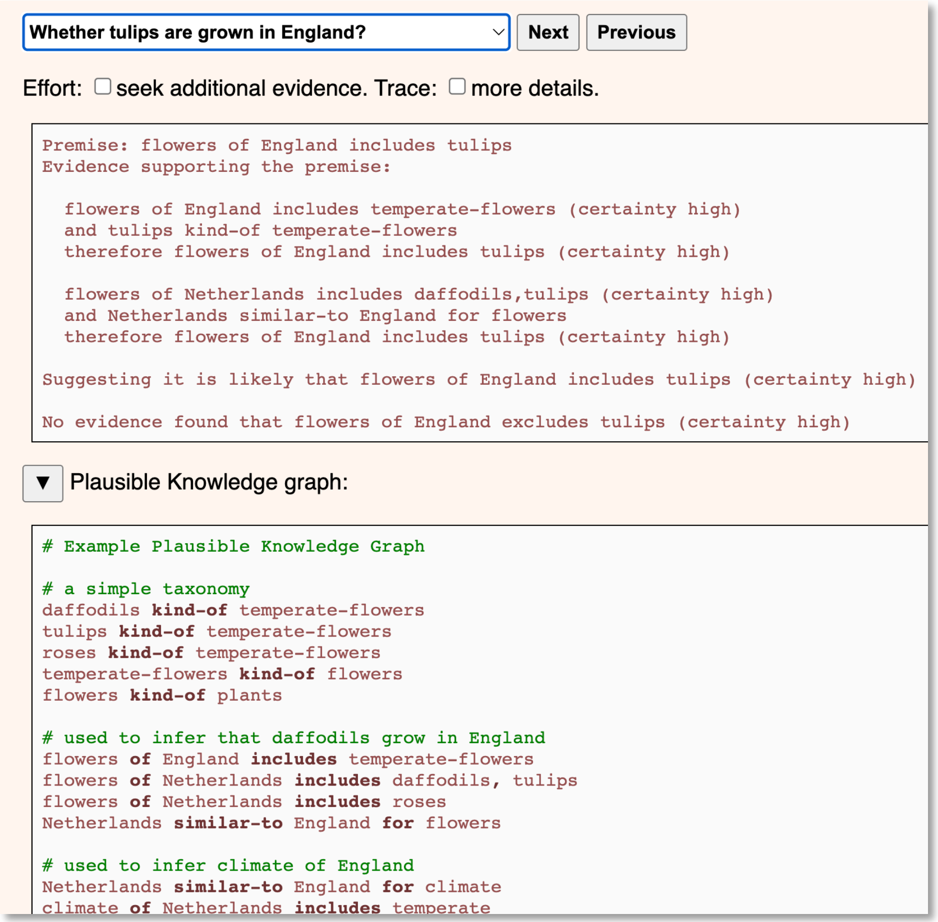
- There are several categories of inference rules that people commonly use to answer questions
- People weigh the evidence that bears on a question, both for and against, rather like in court proceedings
- People are more or less certain depending on the certainty of the premises, the certainty of the inferences and whether different inferences lead to the same or opposite conclusions
- Facing a question for which there is an absence of directly applicable knowledge, people search for other knowledge that could help given potential inferences

** See: Collins & Michalski (1988) and subsequent extensions by Burstein, Collins and Baker (1991)*

Web-based Proof of Concept Demo

The demo includes a variety of examples along with analogical reasoning and fuzzy quantifiers

- Implementations are invaluable for testing understanding of previous work and for identifying challenges for new work
- I developed a web-based demo using a novel simple notation* and an inference engine inspired by the work of Allan Collins et al.
  - See: https://www.w3.org/Data/demos/chunks/reasoning/
- Collins distinguishes four kinds of plausible assertions
  - properties, relationships, implications and dependencies
- Inference involves qualitative parameters‡ as metadata
  - certainty, typicality, similarity, frequency, dominance, conditional likelihood
- A collection of static reasoning strategies …
  - future work is planned on metacognition and continuous learning, including syntagmatic learning, paradigmatic learning and skill compilation



* Plausible knowledge notation (PKN) for cognitive databases

‡ Qualitative due to lack of statistical precision

Everyday Reasoning Isn't Simple

Consider the PKN query for *who is very old*

which ?x where ?x is-a person and age of ?x is very:old

Which is to be evaluated using some facts, e.g.

John is-a person
age of John is 63
Pamela is-a person
age of Pamela is 82

Your age increases until you die

?person is-a person and age of ?person is ?age
implies ?age less-or-equal age-at-death

A person's age is a scalar and associated with terms, e.g.

range of age is infant, child, adult for person
age of infant is 0, 4 for person
age of child is 5, 17 for person
age of adult is 18, age-at-death for person

Another set of terms can be applied to adults, e.g.

range of age is young, middle-age, old, geriatric for adult
age of young is 18, 44 for adult
age of middle-age is 45, 65 for adult
age of old is 66, age-at-death for adult
age of geriatric is 78, age-at-death for adult

We may choose to define very old as geriatric

very:old equivalent-to geriatric

- Solve the query by considering people's ages
  - Over persons, e.g. John is 63 and Pamela is 82

  John is-a person
  age of John is 63
- Infer that John and Pamela are adults
  - Compare each person's age with the ranges, knowing that their age ≤ their age at death

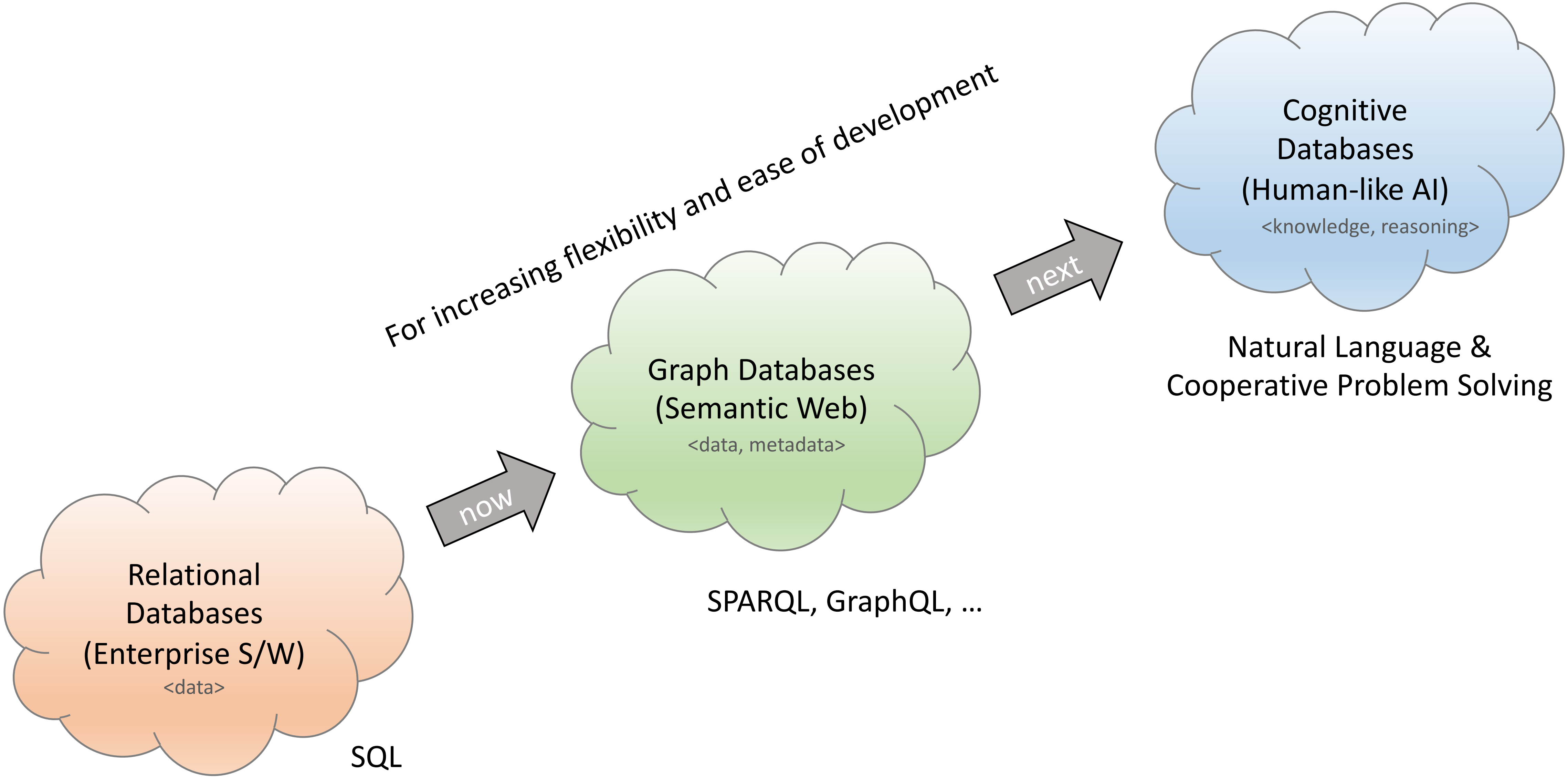
  age of adult is 18, age-at-death for person
- Infer that John is middle-aged, and Pamela is old and geriatric

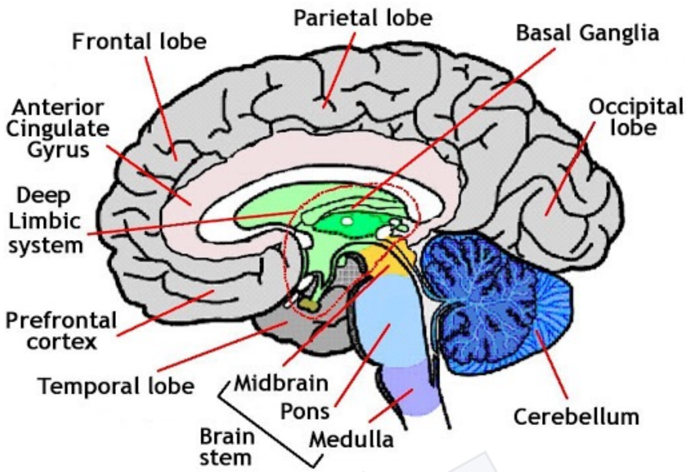
  age of middle-age is 45, 65 for adult
  age of old is 66, age-at-death for adult
  age of geriatric is 78, age-at-death for adult
- Finally, infer that Pamela is very old

  very:old equivalent-to geriatric
- And add her to the query results

The definitions are debatable, and depend on your age and experience – a given person may be considered old by a child, and young by an adult!
Note: ranges are related to Zadeh's fuzzy sets, and terms may overlap, e.g. old subsumes geriatric in the upper part of its range.

Evolution in action



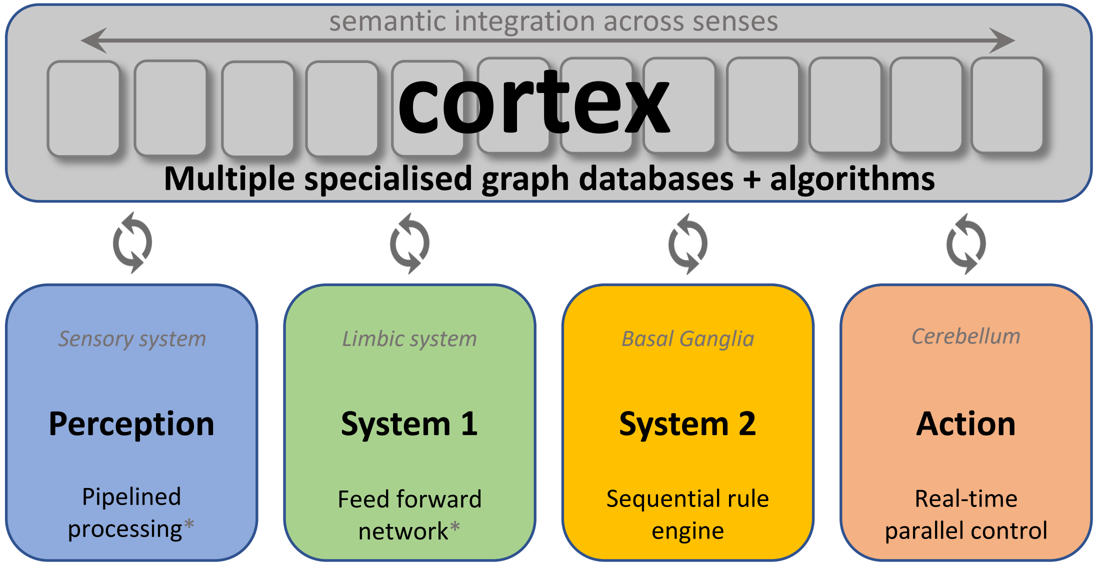


Courtesy of Clipart Library

Cognitive Architecture for artificial minds

Anterior temporal lobe as hub for integration across senses

Multiple cognitive circuits loosely equivalent to shared blackboard



- **Cortex** supports memory and parallel computation. Recall is stochastic, reflecting which memories have been found to be useful in past experience. Spreading activation and activation decay mimics human memory with semantic priming, forgetting curve and spacing effect.
- **Perception** interprets sensory data and places the resulting models into the cortex. Cognitive rules can set the context for perception, and direct attention as needed. Events are signalled by queuing chunks to cognitive buffers to trigger rules describing the appropriate behaviour. A prioritised first-in first-out queue is used to avoid missing closely spaced events.
- **System 1** covers intuitive/emotional thought, cognitive control and prioritising what's important. The limbic system provides rapid assessment of past, present and imagined situations. Emotions are perceived as positive or negative, and associated with passive or active responses, involving actual and perceived threats, goal-directed drives and soothing/nurturing behaviours.
- **System 2** is slower and more deliberate thought, involving sequential execution of rules to carry out particular tasks, including the means to invoke graph algorithms in the cortex, and to invoke operations involving other cognitive systems. Thought can be expressed at many different levels of abstraction, and is subject to control through metacognition, emotional drives, internal and external threats.
- **Action** is about carrying out actions initiated under conscious control, leaving the mind free to work on other things. An example is playing a musical instrument where muscle memory is needed to control your finger placements as thinking explicitly about each finger would be far too slow. The cerebellum provides real-time coordination of muscle activation guided by perception.

*e.g. based upon deep learning and artificial neural networks

System 1 and 2

Popularised by Daniel Kahneman's "Thinking fast and slow"

- System 1 is fast, and apparently effortless, yet opaque
  - we aren't aware how we came to a conclusion
- System 1 is subject to many cognitive biases and sometimes wrong
- Natural language is largely handled via System 1
  - We understand what people are saying, and construct a coherent explanation in real-time that hides the ambiguity of language
  - How do we do this?*
- System 1 & 2 work in cooperation
  - Playing complementary roles

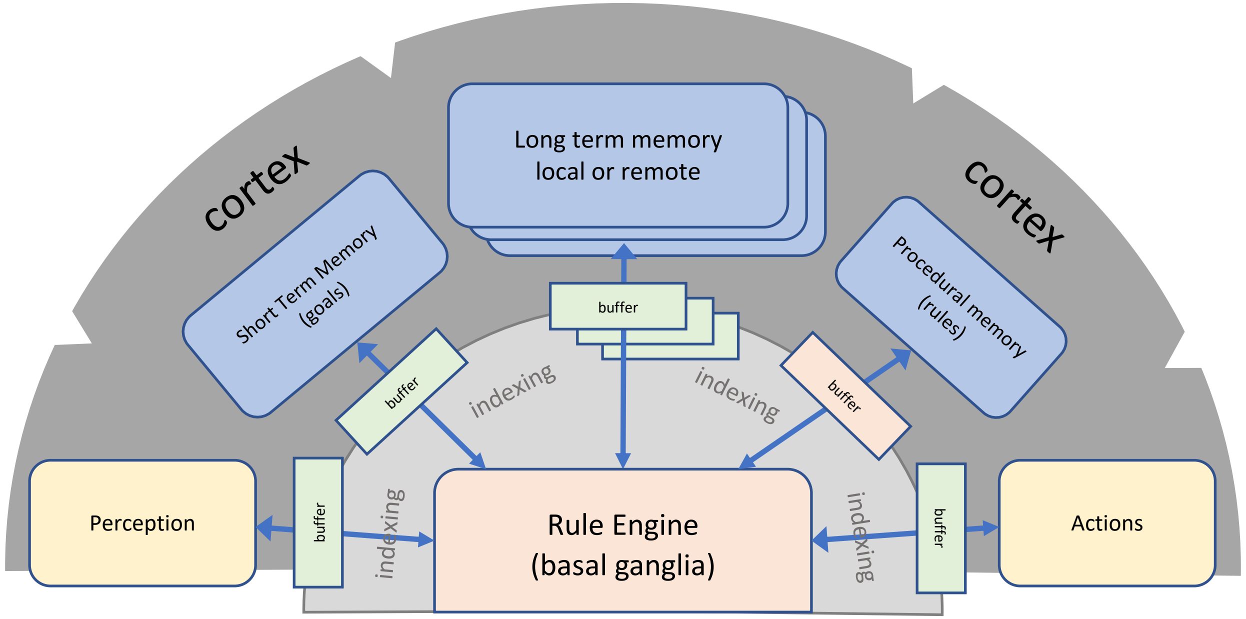
- System 2 is accessible to introspection, and is much slower
  - Overriding System 1 as needed
- System 2 is effortful – making thinking hard work and quite exhausting!
- John Anderson's work on ACT-R
  - Sequential rule engine working on cognitive buffers that hold single chunks (sets of name/value pairs)
  - My reworking of ACT-R in JavaScript
- What about high level cognition?
  - I've started work on integrating natural language and everyday reasoning, along with combining System 1 and 2

* Large Language Models already do this, but in an opaque way

Cortico-Basal Ganglia Circuit

a functional model of System 2

Cognition – Sequential Rule Engine



Cognitive Buffers hold single chunks
Analogy with HTTP client-server model

- Inspired by John R. Anderson's ACT-R
  - Novel simple notation for Chunk graphs together with Condition-Action rules
  - Implemented in JavaScript
- Chunks as a collection of properties for literals and references to other chunks
  - Each chunk buffer is equivalent to the concurrent activity of a bundle of nerve fibres connecting to a given cortical region, see Chris Eliasmith's work on semantic pointers for pulsed neural networks
  - Rules operate over chunk buffers and invoke asynchronous operations on cortical modules that update the buffers
  - Stochastic selection from matching rules whenever buffer contents are updated

See W3C Cognitive AI Community Group for demos and specs

Smart Factory Demo of a Cyber Physical System

https://www.w3.org/Data/demos/chunks/robot/

- **Cognitive AI demo that runs in a web page**
- Live simulation of wine bottling plant with robot, conveyor belts, filling and capping stations
- Control by a cognitive agent using chunks and rules with delegated real-time control over robot movement

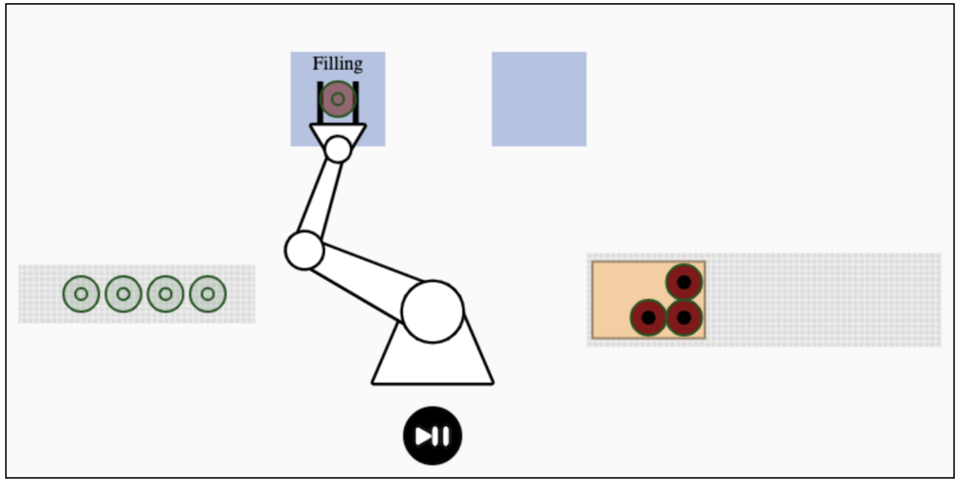
Cognitive rules

```
# add bottle when belt1 has space and wait afresh
space {thing belt1} =>
     action {@do addBottle; thing belt1},
     space {@do wait; thing belt1; space 30}

# stop belt1 when it is full and move arm
full {thing belt1} =>
     action {@do stop; thing belt1},
     action {@do move; x -120; y -75; angle -180; gap 40; step 1}

# move robot arm into position to grasp empty bottle
after {step 1} => robot {@do move; x -170; y -75; angle -180; gap 30; step 2}

# grasp bottle and move it to the filling station
after {step 2} =>  goal {@do clear}, robot {@do grasp},
     robot {@do move; x -80; y -240; angle -90; gap 30; step 3}
```



Log:

```
set goal to: after _:54 {step 1}
executed rule _:27 move
set goal to: after _:55 {step 2}
executed rule _:30 grasp
set goal to: after _:56 {step 3}
starting belt1
wait on filled
executed rule _:34 start
```

Scalable Knowledge Engineering

- Hand crafted knowledge doesn't scale and is brittle when it comes to the unexpected
- Deep learning scales, but is similarly brittle, and requires vast datasets for training
- Humans are much better at generalising from few examples by seeking causal explanations based upon prior knowledge
- Humans are good at reasoning using mental models and chains of plausible inferences, supported by metacognition
- We need research focussed on extending artificial neural networks to support human-like learning and reasoning
- At the same time, we should also explore scalability of machine learning for symbolic representations of knowledge
- Hand authoring for small scale experiments can help illustrate what's needed from more scalable approaches

Image generated with Stable Diffusion, where the internal model fails to understand the anatomy of the human hand

How can we combine deep learning from large corpora with learning taxonomic and causal knowledge? This would provide cognitive agents with the means to learn from fewer examples, by understanding them at a much deeper level

Continuous Learning

- Different ways to learn
  - Syntagmatic learning – patterns in co-occurrence statistics
  - Paradigmatic learning – taxonomic abstractions
  - Skill compilation – speeding reasoning with previous solutions, and use of analogies
  - Reasoning about potential causal explanations of behaviour
  - Metacognition with strategies and tactics for how to reason
- Learning from direct experience
  - Interacting with the world
  - Issues around safety and cost
- Learning from observation and asking questions
  - Children are incredibly good at this
  - Huge opportunity to learn from large corpora of texts, images and videos
- Lessons and assessments
  - Courses designed for AI agents

Manual knowledge engineering is too hard and too slow, so a key research goal will be to allow cognitive agents to discover how to represent and reason for themselves by building upon a core of manually constructed prior knowledge. This could involve machine learning from a corpus of assertions, questions and answers, where the agent seeks to improve its understanding, e.g. by developing causal models as a way to explain the training texts. This could involve a generative-adversarial approach in which one agent seeks to fool another agent into thinking it is human! Metacognition is then a further elaboration where agents are tasked with solving problems, and given some hints, where the aim is to learn how to reason with different strategies as appropriate to the problem in hand.


Courtesy of Dave Lebow

Human-like AI

falling down the rabbit hole into a new world!

General purpose Human-like AI will dramatically change how we work, how we communicate, and how we see and understand ourselves!

Let's seize the opportunity and encourage research on realising human-like learning and reasoning!

See also: HiPEAC vision for the creation of a "Guardian Angels" moonshot programme to create a "next web" that intertwines the cyber and physical worlds for industrial and personal use, overcoming the fragmentation of vertically-oriented closed systems, heterogeneity and the lack of interoperability. It should demonstrate self-configuration and self-management in a dynamic plug-and-play environment, while also coping with security and privacy of personal and corporate data and offering natural interfaces for their users.