Defeasible Reasoning with Knowledge Graphs

13 November, KGSWC-2023, Zaragoza

Dave Raggett dsr@w3.org

W3C/ERCIM

European Research Consortium for Informatics and Mathematics ERCIM

Defeasible reasoning has been studied since the days of Ancient Greece

This work is supported by the European Union's Horizon research and innovation programme under grant agreement No. 957406 for project TERMINET on next generation smart interconnected IoT.

Why Logic is Inadequate in the Real World

- **Logic** is based upon **deductive proof** and **assumes perfect knowledge**
- Logic isn't applicable for knowledge that is uncertain, imprecise, incomplete, inconsistent and changing, i.e. imperfect knowledge
- That however is typically the case for everyday knowledge
- Defeasible reasoning is the basis for legal arguments, ethics, political arguments and everyday discussion
  - *Defeasible reasoning is more general than logic covering deduction, induction, abduction, analogy and fallacies*

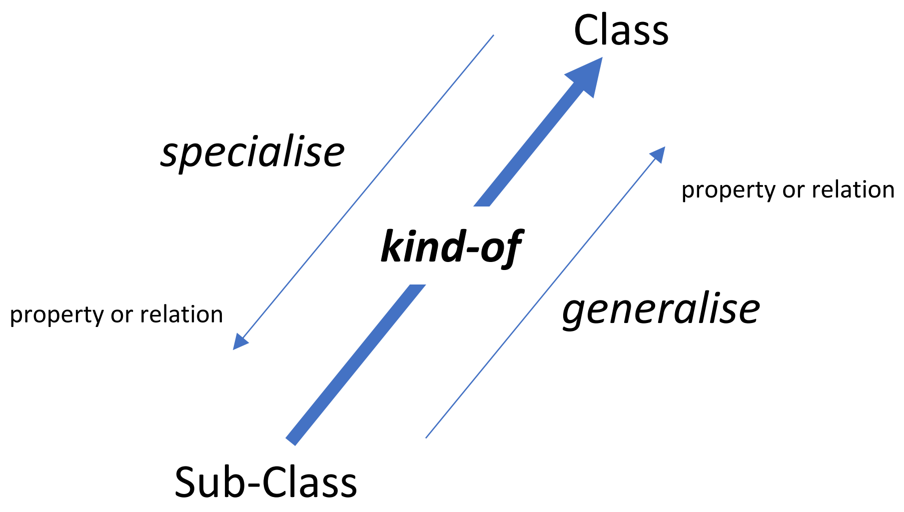
- **Defeasible reasoning** deals with ***plausible arguments***
  - *Exploiting prior knowledge for inferences*
- Arguments in *support* of, or *counter* to the supposition in question
- Conclusions may need to be withdrawn in the light of new information
- Arguments are *combative* where the parties try to beat each other down, or *collaborative* where the parties work towards a better mutual understanding

Previous Work on Argumentation Theory

- The Stanford Encyclopaedia of Philosophy lists five types of arguments: *deduction*, *induction*, *abduction*, *analogy* and *fallacies*
- Studies of argumentation have been made by a long line of philosophers dating back to Ancient Greece, e.g., Carneades and Aristotle
- More recently, logicians such as Frege, Hilbert and Russell were primarily interested in mathematical reasoning and argumentation
- Stephen Toulmin subsequently criticized the presumption that arguments should be formulated in purely formal deductive terms
- Douglas Walton extended tools from formal logic to cover a wider range of arguments – a set of argument schemes
- Ulrike Hahn, Mike Oaksford and others applied Bayesian techniques to reasoning and argumentation
- AIF is an ontology intended to serve as the basis for an interlingua between different argumentation formats
- Alan Collins applied a more intuitive approach to plausible reasoning that takes sub-symbolic knowledge into account to model rough notions of metadata in lieu of statistics
  - Collins inspired my work on the **Plausible Knowledge Notation**
- Arguments in support of, or counter to, some supposition, build upon the facts in the knowledge graph or the conclusions of previous arguments
- Preferences between arguments are derived from preferences between rules with additional considerations in respect to consistency
- Counter arguments can be classified into three groups
  - **undermining** another argument when the conclusions of the former contradict premises of the latter.
  - **undercutting** another argument by casting doubt on the link between the premises and conclusions of the latter argument.
  - **rebutting** another argument when their respective conclusions can be shown to be contradictory.

Plausible Inferences using Prior Knowledge

- Inferring likely properties and relations across other relations



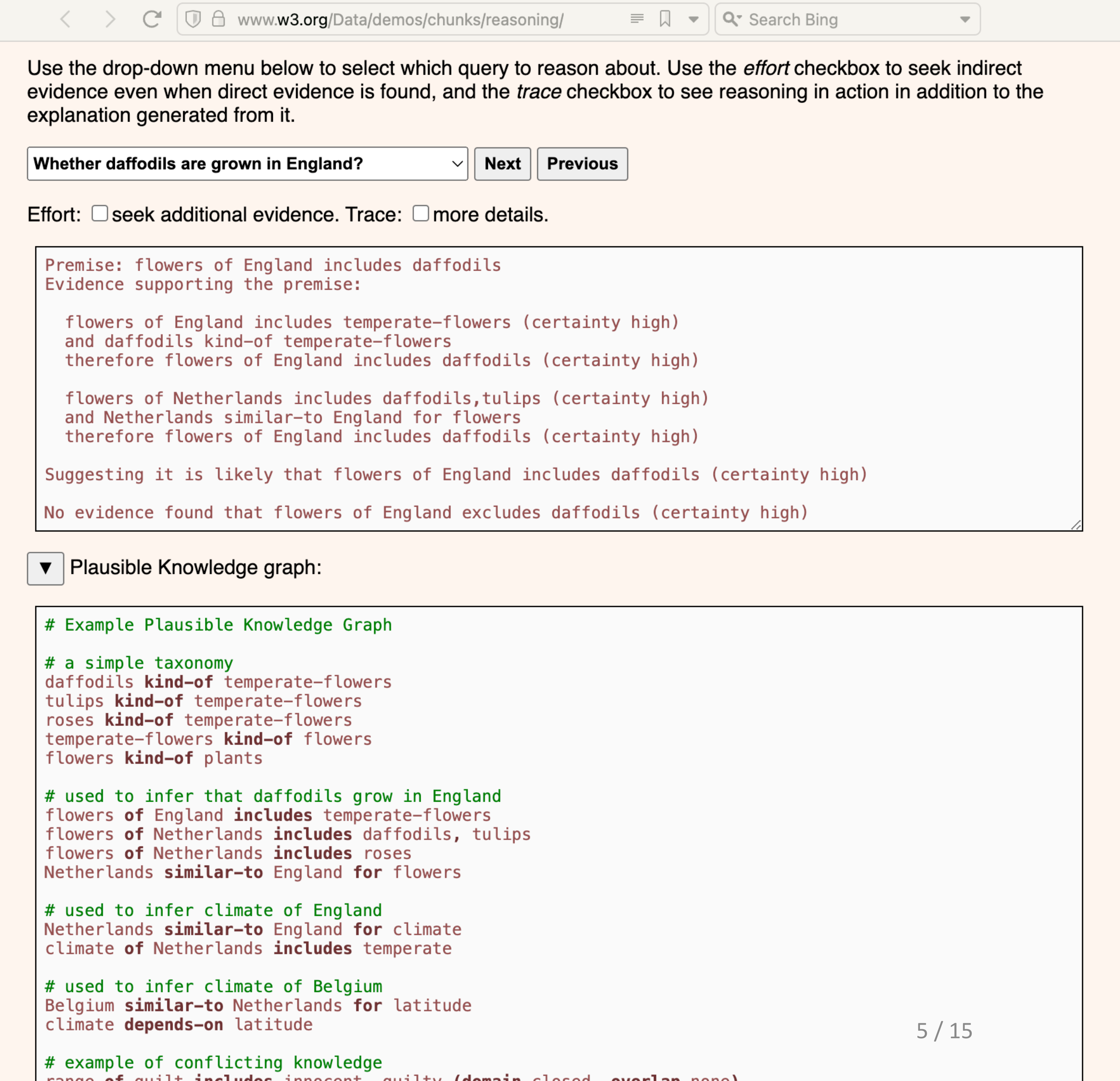
- Expected certainty influenced by qualitative metadata
  - e.g. typicality, similarity, strength, dominance, multiplicity, scope, …

- Forward and backward inferences using implications
  - If it is raining then it is cloudy
  - If it is cloudy it may be rainy
- Inferences based upon analogies
  - matching structural relationships
- Scalar ranges (fuzzy logic)
  - fuzzy terms, e.g. cold, warm and hot
  - fuzzy modifiers, e.g. *very* old
  - fuzzy quantifiers, e.g. *few, many*
- Multiple lines of argument for and against the premise in question
  - Just as in the courtroom
- Plausible knowledge Notation
  - W3C Cognitive AI Community Group

PKN Demonstrator

- Proof of concept implementation in JavaScript as a web page
- Large collection of examples
- Works back from the supposition towards the supporting facts
- Avoids circular arguments
- Explanation generated in forward pass through trace of execution

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



Plausible Knowledge Notation (PKN)

*The Plausible Knowledge Notation (**PKN**) includes enriched semantics and an **easier to use** notation **relative to RDF/turtle***

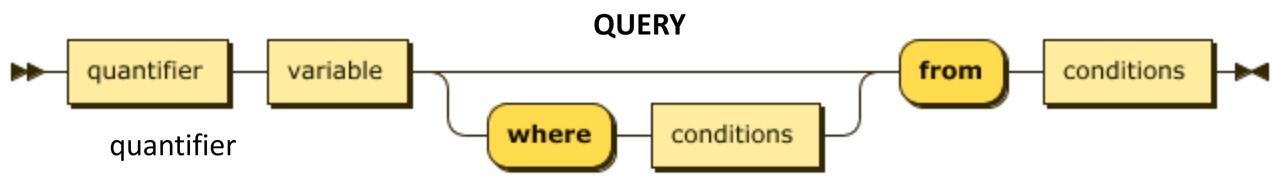
properties, relationships, contextual scope, implication rules, fuzzy ranges, fuzzy modifiers, fuzzy quantifiers, analogies, parameters denoting gut feelings, statements about statements

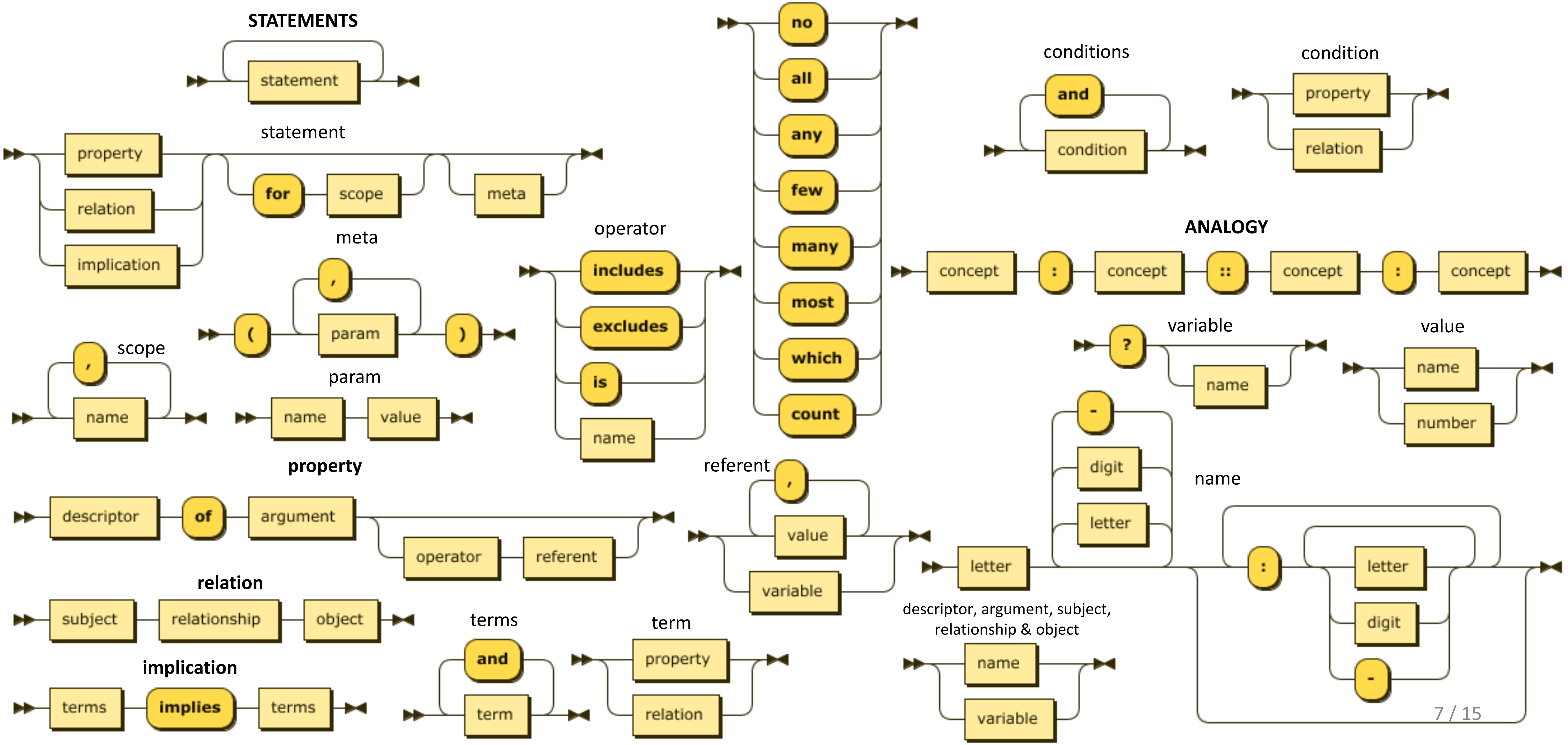
See: [W3C Cognitive AI Community Group specification for PKN]()

climate of Belgium includes temperate
guilt of accused excludes guilty
roses kind-of temperate-flowers
circuit analogous-to plumbing
flow increases-with pressure for plumbing
current increases-with voltage for circuit
flow:current::pressure:voltage
dog:puppy::cat:?
weather of ?place includes rainy
  implies weather of ?place includes cloudy (strength high, inverse low)
up opposite-to down
Mary younger-than Jenny
younger-than equivalent-to less-than for age
range of age is infant, child, adult for person
age of infant is birth, 4 for person
John loves chess
subject of loves includes person
object of loves includes hobby (strength medium)
which ?x where ?x is-a person and age of ?x is very:old
count ?x where age of ?x greater-than 20 from ?x is-a person
few ?x where color of ?x includes yellow from ?x kind-of rose
Mary believes {{John says {John loves Joan}} is-a lie}

PKN grammar

statements, queries, analogies





Strategies and Tactics for Argumentation

- Further work is now needed on an intuitive syntax for reasoning strategies and tactics, as well as ways to model the role of feelings and emotions as part of compelling arguments
- Building upon well established **principles** for **effective arguments**, e.g. classical rhetorical guidelines dating back to Aristotle
  - **Ethos**: establishing credibility to engender trust
  - **Pathos**: using emotion to stir people's feelings
  - **Logos**: using logic to emphasise rational support
  - **Kairos**: opportune, i.e. timely and topical in nature

How does PKN relate to RDF and Artificial Neural Networks?

- W3C's Resource Description Framework (RDF) is based upon triples, i.e. labelled directed graph edges

  <subject, predicate, object>
- PKN statements can be modelled as a collection of triples
- Consider: flowers of England includes roses (certainty high)
- This can be mapped to an RDF blank node as the subject for 5 triples*
- PKN uses comma separated lists for collections – something that is harder to express in RDF
- PKN uses curly braces for graphs
  Mary believes {John loves Sarah}

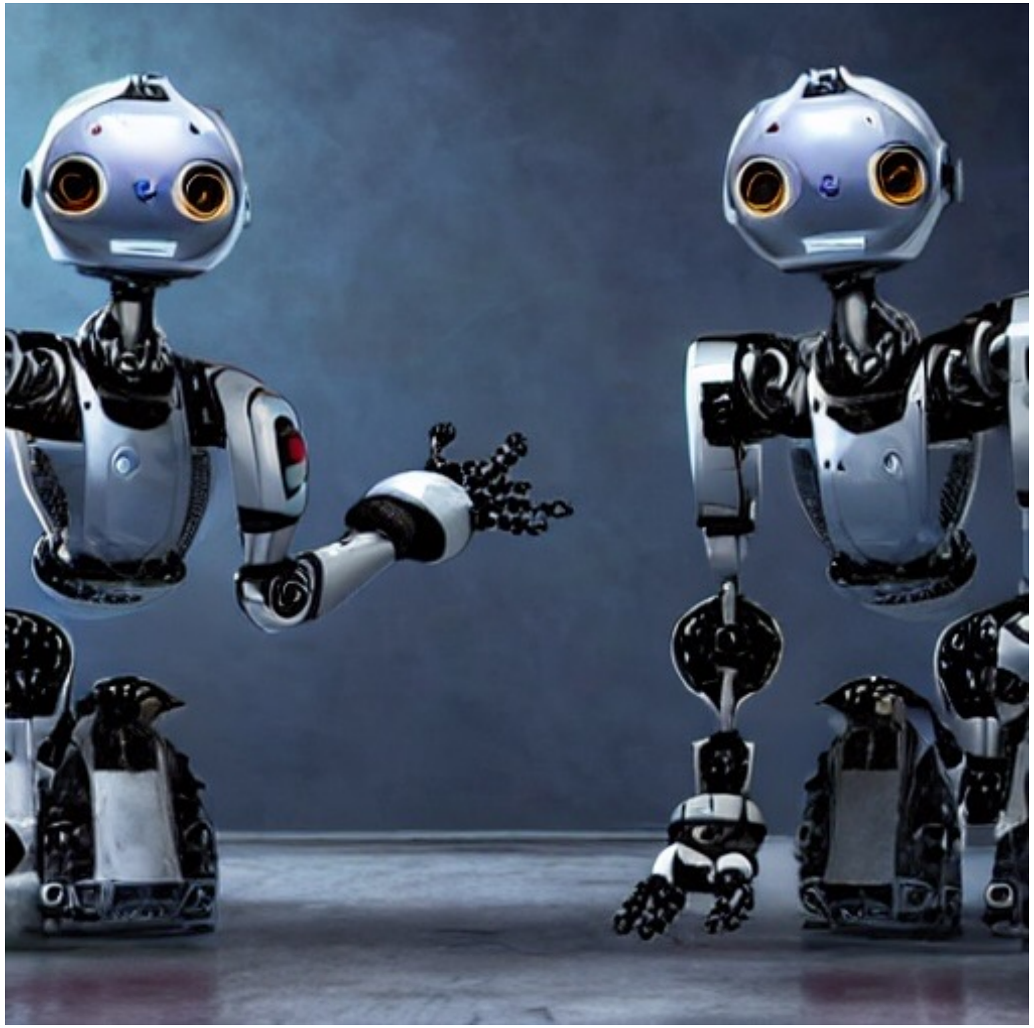
* using context files analogous to JSON-LD
  to define the mapping from names to URIs

- PKN can be considered as a member of a family of notations that provide richer semantics compared to RDF
- PKN can also be mapped to cognitive chunks – sets of name/value pairs
  - e.g. as used by CMU's cognitive architecture: ACT-R ‡
- Large language models internally represent chunks as vectors of activation values
  - Distributed statistical relationships
  - Multi-head transformers for attention
  - Opaque and lacking transparency
- This explains how artificial neural networks can manipulate semantic graphs in their working memory

‡ See also chunks and rules from the W3C Cognitive AI CG

Semantic Interoperability

Knowing that we understand each other by using a shared language and vocabulary

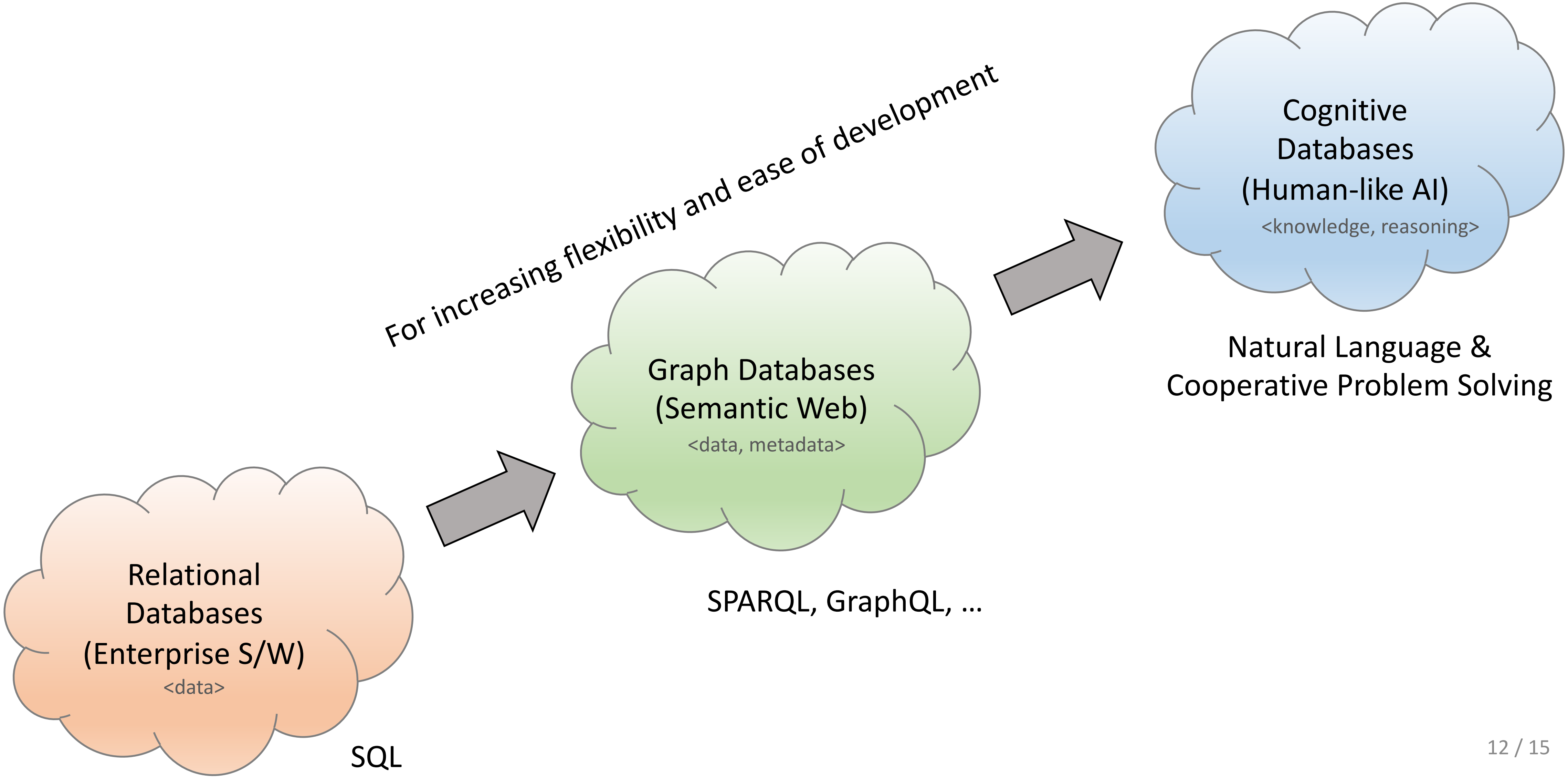


Generative AI lacks semantic consistency, as shown by the lack of support for the robot's body

Ensuring Mutual Understanding

- People keep written records when they don't want to rely on fallible memory
- The same applies to businesses
- Everyday language isn't good enough when we need to be sure of a mutual understanding
  - Business contract between a supplier and a consumer
    - Use of standardised terms and legal language for contracts
- For technical exchanges we use structured data with agreed data models and semantics
- This relies on symbolic representations
- We will continue to need this as we make greater use of AI
- Knowledge Graphs as an evolution of databases
- Standardised vocabularies

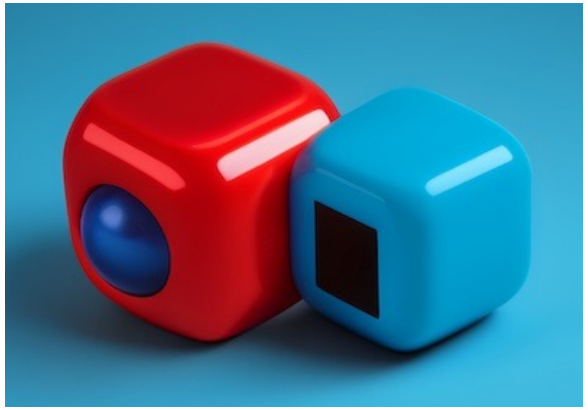
Evolution in ICT Systems



AGI and Defeasible Reasoning?

- Generative AI, e.g. GPT-4, is surprisingly effective at understanding and reasoning with human language, but …
- Prone to distractions and hallucinations
- Weak on logical reasoning and semantic consistency
- Lack of continual learning and episodic memory for past, present and future
- Work is needed on new neural network architectures inspired by what we know about human cognition
- This will enable practical AGI solutions for human-computer collaboration
- For more details, see my talk at: http://www.w3.org/2023/10/10-Raggett-AI.pdf

Hmm, how many fingers do humans have?



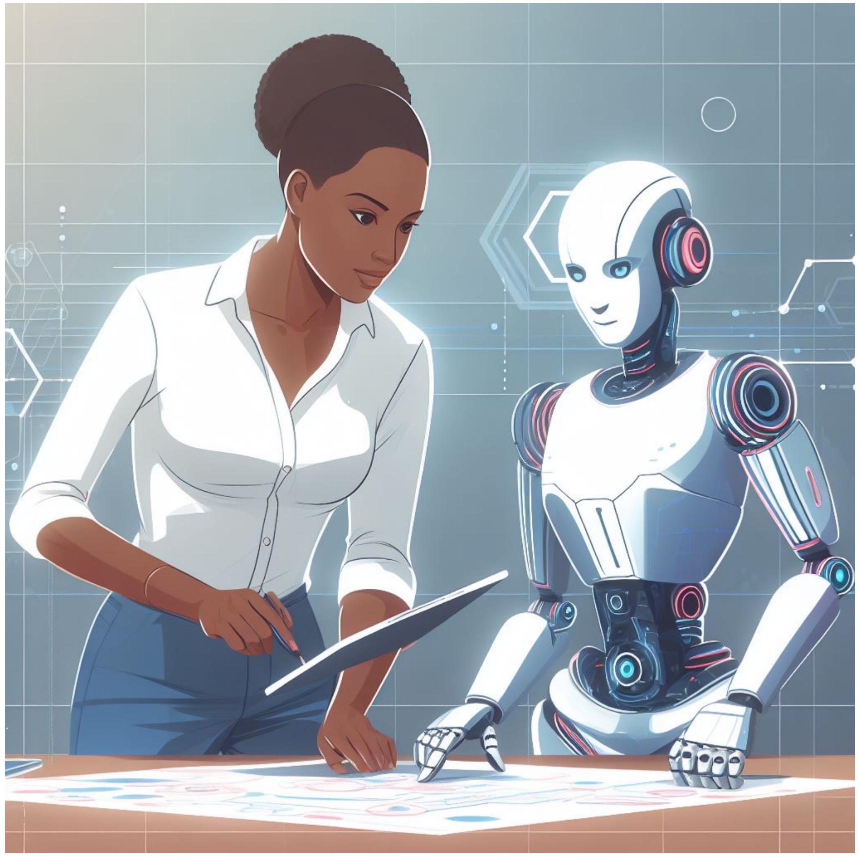
"3 red balls and 2 blue cubes on a wooden floor", **really???**

Is 1 kg heavier than 2 kg: **no** ✓

Is 1 kg of lead heavier than 2 kg of feathers: **yes** ✗

Collaborative Knowledge Engineering

- Hand crafting knowledge graphs + rule sets is difficult and time consuming – this makes it hard to scale up
- Self-guided machine learning with neural networks is very much easier to scale up, but suffers from a lack of transparency
  - Knowledge is buried in the network parameters
- How can we use AI for collaborative knowledge engineering?
  - Human partner working together with an artificial agent
  - Agent operates on knowledge graphs + rule sets guided by human partner
  - Curating datasets, e.g. for new or updated use cases
  - Automated updates to rules as ontologies are revised
  - Versioning to support old and new applications



Note unsupported tablet floating in the air!

Questions and Comments?

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

European Research Consortium for Informatics and Mathematics
ERCIM

This work is supported by the European Union's Horizon research and innovation programme under grant agreement No. 957406 for project [TERMINET](#) on next generation smart interconnected IoT.