Knowledge-Based Systems

Knowledge Capture

Jeff Z. Pan

UNIVERSITY OF ABERDEEN

Lecture Outline

- Story so far
- Motivation
- Introduction to knowledge capture
- Some detailed discussions on knowledge capture
- Practical

UNIVERSITY OF ABERDEEN

Resource Description Framework

- Modern version of semantic network
- Basic building block: Subject-property-value triple
  - It is called a statement
  - E.g. Jeff teaches CS3025 is a statement

- RDF (Resource Description Framework) has been given a syntax in XML
  - This syntax inherits the benefits of XML
  - Other syntactic representations of RDF possible
    - such as Notation 3 (N3) syntax: [Subject property value .]
- RDF and database?

UNIVERSITY OF ABERDEEN

Database and RDF

- An RDF statement is a data unit with global and linkable IDs for data and schema

Student ID	Name	take-course
p001	John	cs3019
p002	Tom	cs3023

  - [csd:p001 rdf:type csd:Student .]
  - [csd:p002 rdf:type csd:Student .]
  - [csd:p001 csd:name “John” .]
  - [csd:p002 csd:name “Tom” .]
  - [csd:p001 csd:take-course csd:cs3019 .]
  - [csd:p002 csd:take-course csd:cs3023 .]

UNIVERSITY OF ABERDEEN

Schema Languages for RDF

- Basic schema language: RDF Schema
  - rdfs:subClassOf
  - rdfs:subPropertyOf
  - rdfs:domain
  - rdfs:range
- Advanced schema language: OWL
  - conjunction
  - disjunction
  - existential restriction
  - universal restriction
  - …

UNIVERSITY OF ABERDEEN

How many facts?

Schema-like

```
(deftemplate person
     (slot name)
     (slot age)
     (slot gender)
     (slot partner)
     (multislot children))
```

Data

```
(assert (person
                    (name Michael)
                    (gender male)
                    (age 35)
                    (partner Diana)
                    (children Philip
Julia)))
<fact-0>
```

UNIVERSITY OF ABERDEEN

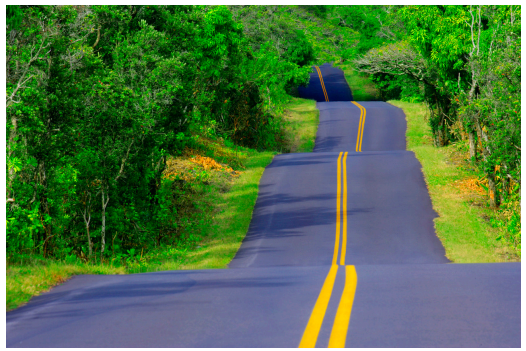
How many facts?

- Need to use bland nodes (or b-nodes) in RDF
  - a **node** (starting with _:) in an **RDF** graph representing a resource for which a URI or literal is not given
- Here are the RDF triples:
  - [_:m rdf:type Person .]
  - [_:m name “Michael” .]
  - [_:m gender male .]
  - [_:m age 35 .]
  - [_:m partner Diana .]
  - [_:m children Philip .]
  - [_:m children Julia .]

```
(assert (person
                    (name Michael)
                    (gender male)
                    (age 35)
                    (partner Diana)
                    (children Philip
Julia)))
```

UNIVERSITY OF ABERDEEN

Roadmap



- Foundation
  - KR, ontology and rule; set theory
- Knowledge capture
- Knowledge representation
  - Ontology: Semantic Web standards RDF and OWL, Description Logics
  - Rule: Jess
- Knowledge reasoning
  - Ontology: formal semantics, tableaux algorithm
  - Rule: forward chaining, backward chaining
- Knowledge reuse and evaluation
- Meeting the real world
  - Jess and Java, Uncertainty, Invited talk

UNIVERSITY OF ABERDEEN

Lecture Outline

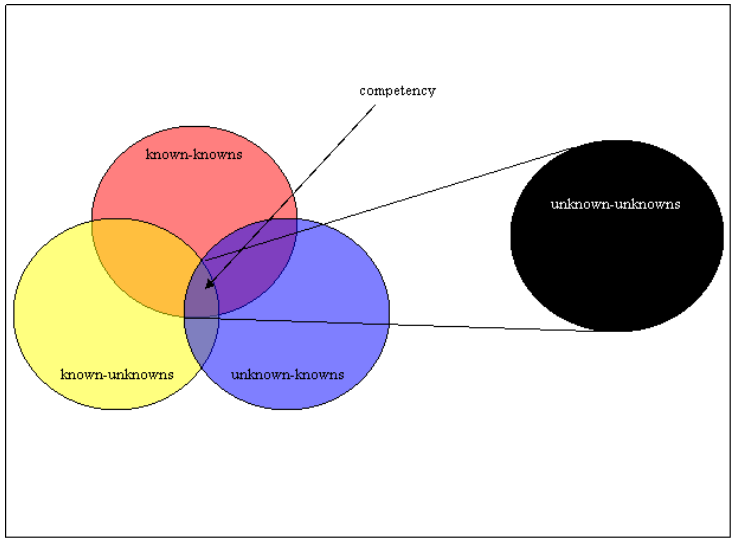
- Story so far
- Motivation
- Introduction to knowledge capture
- Some detailed discussions on knowledge capture
- Practical

Motivations

- Knowledge capture
  - not a trivial task
- The information is often locked away in the heads of domain experts
  - "*Oh, I know that, but you see I don't know how I do diagnosis, and yet I need to teach things to students. I create what I think of as plausible means for doing tasks and hope students will be able to convert them into effective ones.*" (Johnson, 1983, P81).
- We have to draq out and make explicit all the known knowns,known unknowns

UNIVERSITY OF ABERDEEN

Known & Unknowns

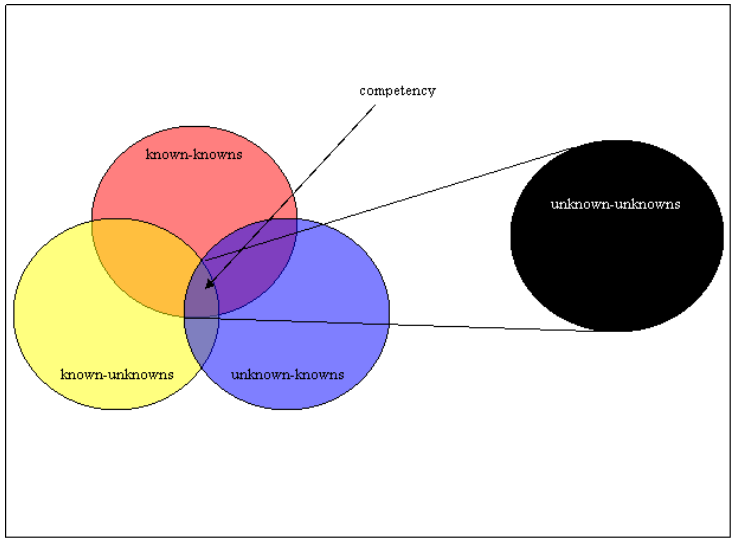


[Picture Credit: www.akkamsrazor.com]

- There are known knowns
  - there are things we know we know.
- We also know there are known unknowns
  - this is to say we know there are some things we do not know.
- There are also unknown knowns
  - these are things we do not know we know
- But there are also unknown unknowns
  - the ones we don't know we don't know.

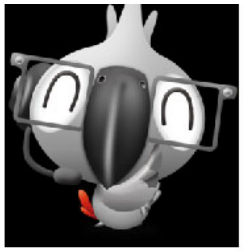
UNIVERSITY OF ABERDEEN

Known & Unknowns (II)

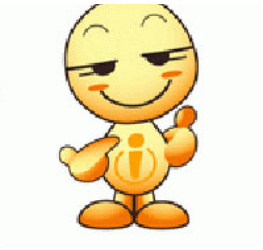


[Picture Credit: www.akkamsrazor.com]

- Experts have vast amounts of knowledge to share
  - yet they understand there are open questions in the domain
- They don't know all that they know and use
  - Experts have a lot of tacit knowledge
  - Tacit knowledge is hard (impossible) to describe.
- Each expert doesn't know everything



JIMI



小i

公子小白

Rokid

开心熊宝

在线客服

娱乐

教育

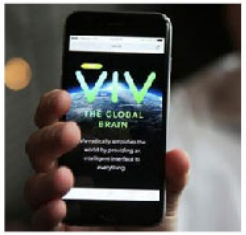
Online custom service

Entertainment

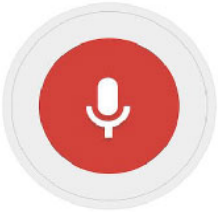
Education

Siri

小度



viv



Google Now

出门问问

IBM Waston

个人助理

智能问答

Personal assistant

Intelligent QA

IBM's Watson supercomputer destroys all humans in Jeopardy

[Video Credit: www.youtube.com]

UNIVERSITY OF ABERDEEN

Lecture Outline

- Story so far
- Motivation
- Introduction to knowledge capture
- Some detailed discussions on knowledge capture
- Practical

UNIVERSITY OF ABERDEEN

Nature of Expertise: Chess

EXPERIMENT 1 (Simon & Chase, 1973)

Experts & Novices were both shown a chess board of an actual game. Both groups were shown the board for the same period. Both groups were asked to recall what pieces they could remember.

Result: The Experts were able to recall significantly more pieces than the Novices.

EXPERIMENT 2

Both experts & novices were shown a Chess board where pieces had been placed randomly, for the same period.

Result: Both groups recalled similar numbers of pieces.

CONCLUSION: EXPERTS save their KNOWLEDGE as CHUNKS (ie meaningful patterns).

UNIVERSITY OF ABERDEEN

Nature of Expertise: Programming

- Expert & novice programmers shown:
  - Actual code {Experts recall much more than novices: that is they see structures like while loops whereas novices see reserved words like while}
  - Programs where the lines have been randomly presented {Experts & novices recall comparable amounts}

UNIVERSITY OF ABERDEEN

NATURE of EXPERTISE: SUMMARY

- Expert's knowledge is,
  - in a sense, "compiled" to enhance performance, and
  - this *may* impede explanation.
  - real experts can "extend" their knowledge to truly novel tasks.
- Expertise (both in intellectual and physical skills)
  - takes many years to acquire and
  - needs regular usage to maintain
- (Even) experts do not always/ often systematically investigate large solution spaces.
  - So it is possible, in some circumstances, for a KBS to outperform an "expert". (Based on a quote from Bruce Buchanan.)

UNIVERSITY OF ABERDEEN

Two Roles

- Experts
  - Academic
    - Regards domain as having a logical structure
  - Practitioner
    - Heavily into day-to-day problem solving
    - Implicit understanding of the domain
- Knowledge engineer
  - understand knowledge representation
  - but not the domain knowledge, so need to do some homework
  - should play an proactive role in the process of knowledge capture

UNIVERSITY OF ABERDEEN

Interview

- Helpful for creating a glossary of terms / approaches
- Before the interview
  - Establish goal of the session
  - Describe for yourself a profile of the expert
  - List relevant questions
- During the interview
  - Clarify goal and expectations
  - Indicate how the results will be used (trust and privacy)
  - Phrase questions in terms of probes: why, what, how
  - Avoid suggestive questions
    - How fast was the car going when it crashed into the wall?
    - How fast was the car going when it ran into the wall?

UNIVERSITY OF ABERDEEN

Interview

- During the interview (cont.)
  - Focus on glossary of terms / approaches :
    - **confirm basic concepts**
    - **relationships between key concepts**
    - **range of acceptable values for key concepts / properties**
    - **types of tasks covered**
  - Be aware of personal biases
  - Give summaries at intermediate points
- After the interview
  - Make appointments for the **next meetings**
  - Process interview results ASAP
  - Organize feedback round with expert

UNIVERSITY OF ABERDEEN

Unstructured Interview

- No detailed agenda
  - Few constraints
  - Delivers diverse, incomplete data
- Used in **early stages**: feasibility study, knowledge identification
- Useful to establish a common basis with expert
  - s/he can talk freely

UNIVERSITY OF ABERDEEN

Structured Interview

- Knowledge engineer plans and directs the session
  - Takes form of provider-elicitor dialogue
- Delivers more focused expertise data
- When to use:
  - Often used for “filling in the gaps” in the knowledge base
  - Also useful at start of knowledge representation
- Always create a transcript

UNIVERSITY OF ABERDEEN

Lecture Outline

- Story so far
- Motivation
- Introduction to knowledge capture
- Some detailed discussions on knowledge capture
- Practical

UNIVERSITY OF ABERDEEN

Techniques

- Interview
- Self report / protocol analysis
- Laddering
- Concept sorting
- Automated knowledge extraction techniques

UNIVERSITY OF ABERDEEN

Difficulties

[Picture Credit: www.stus.com]

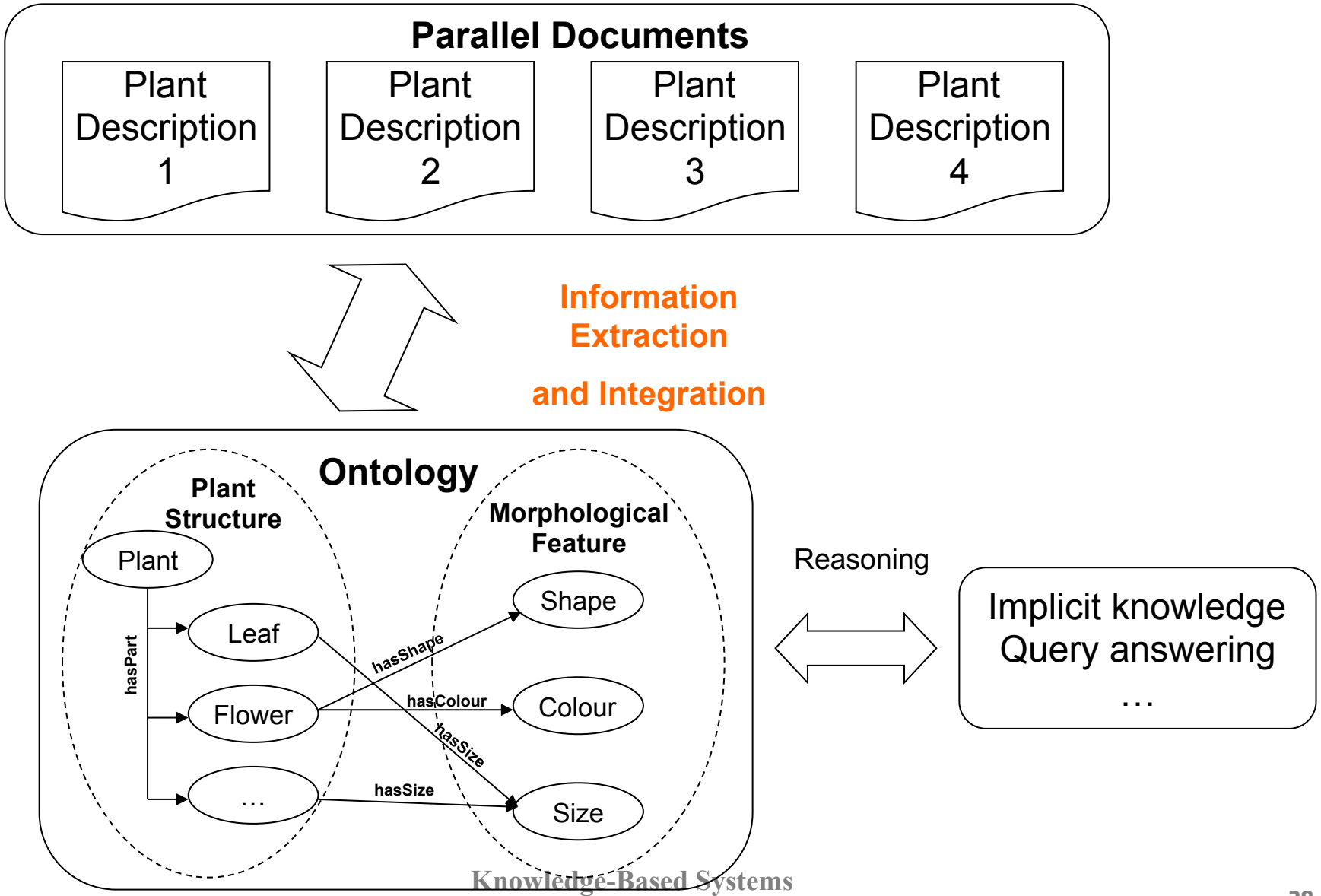
- Experts are very busy and valuable people
- Experts might have different opinions
- Problems with the expert
  - Not interested: would rather do own work
  - Uncooperative: feels threatened
  - Wrong level: too theoretical
  - Inarticulate: a 'techie'
- Relationship problems
  - Different interests
  - Status differences

UNIVERSITY OF ABERDEEN

Self Report

- Overcome some limitations of interviews
  - where experts can only produce what they can verbalise
  - Knowledge engineer must be sufficiently acquainted with the domain
- Expert performs a task while providing a running commentary
  - expert is “thinking aloud”
- Session protocol is always transcribed and analysed
  - Identify various types of knowledge such as relationships and definable concepts
  - Look out for “when”-knowledge (task control knowledge)

Automatic Knowledge Extraction



You may have heard of IBM's Watson…

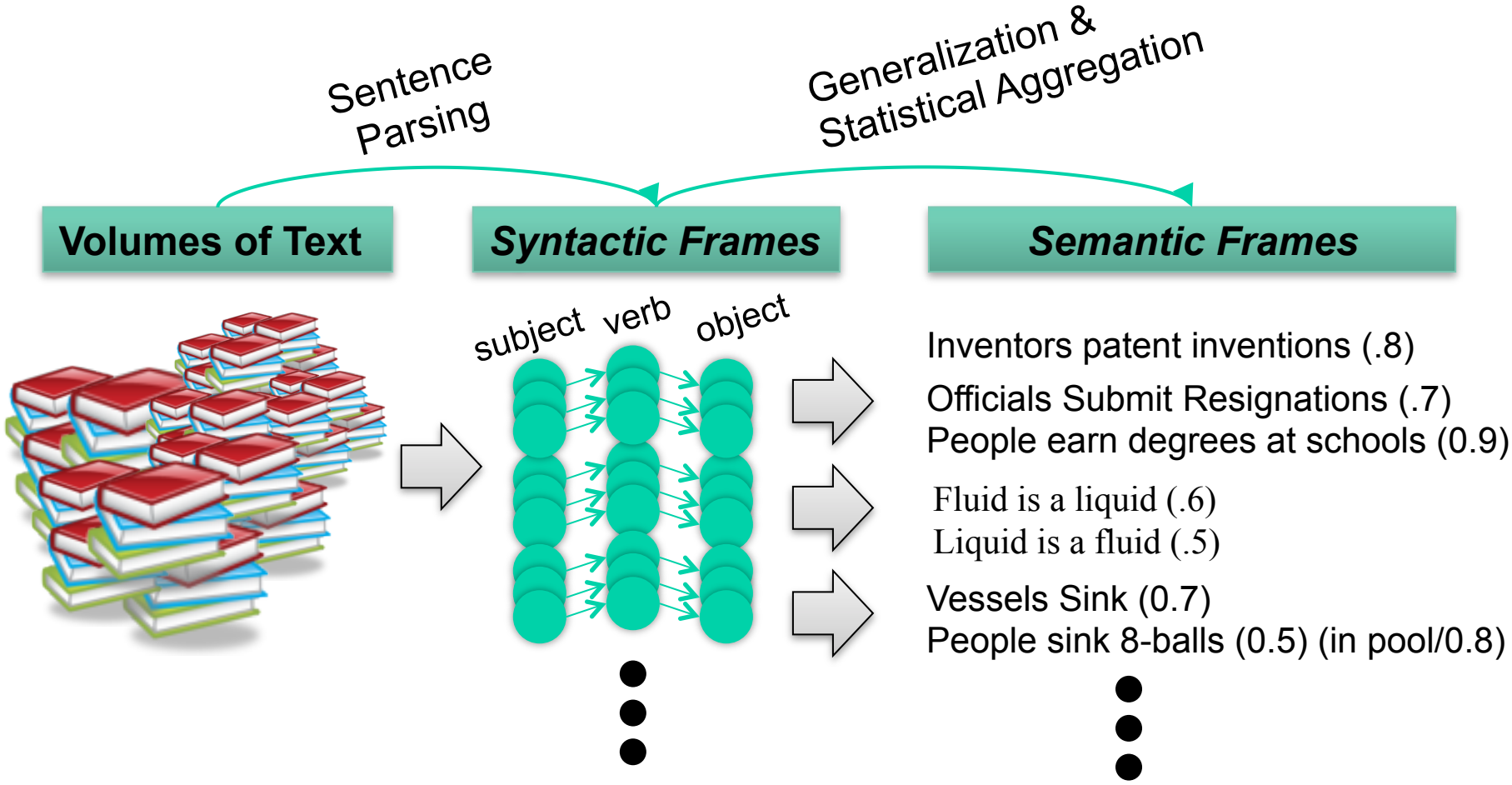
A. What is the computer system that played against human opponents on "Jeopardy"… and won.

Why Jeopardy?

The game of *Jeopardy!* makes great demands on its players – from the range of topical knowledge covered to the nuances in language employed in the clues. The question IBM had for itself was "is it possible to build a computer system that could process big data and come up with sensible answers in seconds—so well that it could compete with human opponents?"

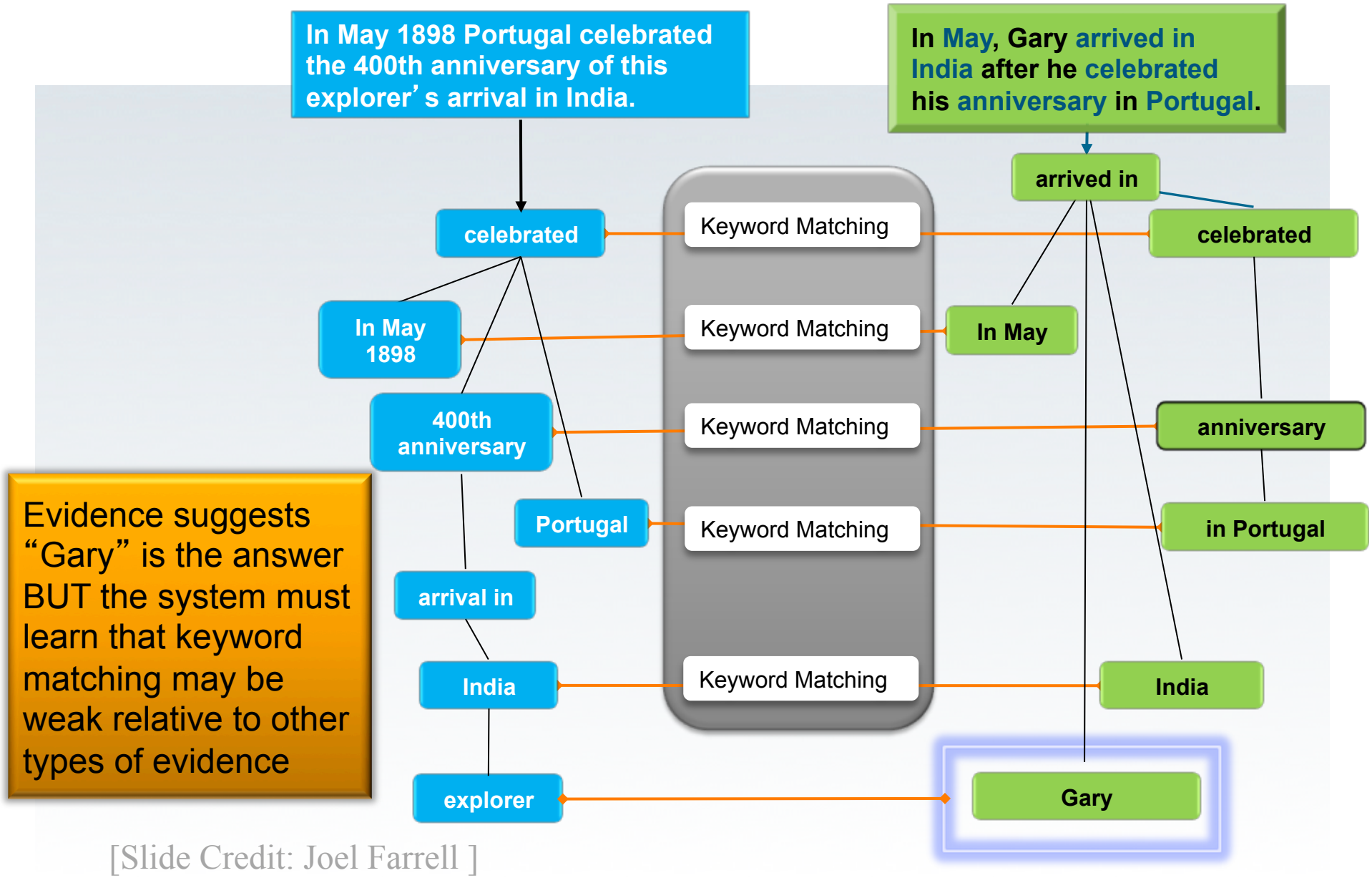
[Slide Credit: Dalal et al.]

Automatic Learning for "Reading"



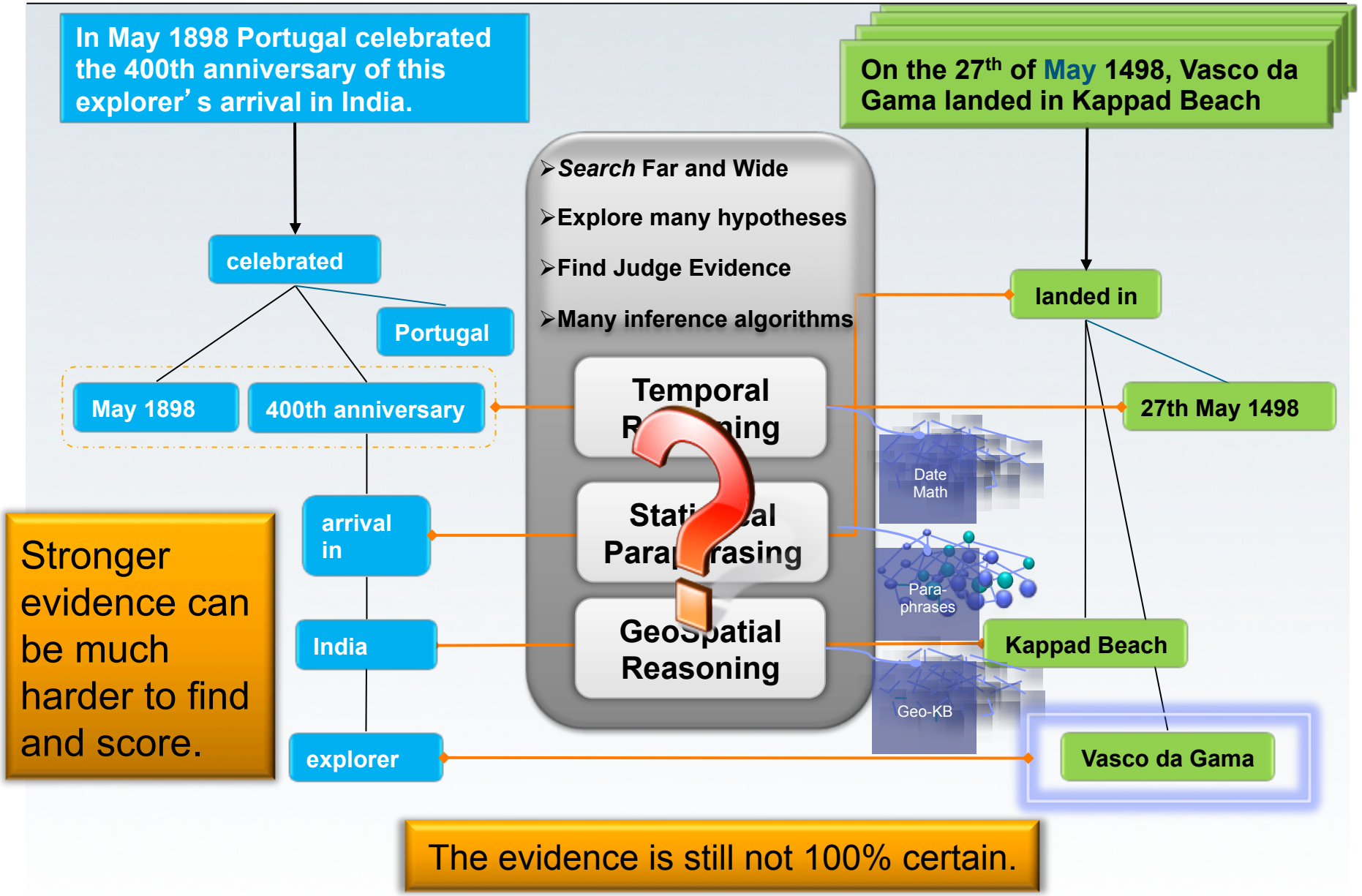
[Slide Credit: Joel Farrell]

Different Types of Evidence: Keyword Evidence



[Slide Credit: Joel Farrell]

Different Types of Evidence: Deeper Evidence



Combining Subjective Opinions and Logic

- Joint work with IBM for dealing with multiple evidences: **subjective description logic [AAMAS2013]**
- Subjective opinions (b,d,u) include information on belief, disbelief, and uncertainty
  - with b,d,u in the range of [0,1] and b+d+u=1
  - example: (0.5,0.3,0.2)
- Belief, disbelief, and uncertainty are calculated based on positive and negative evidiences
  - p=2*b/u, n=2*d/u
  - example: if p=10, n=2, then b=5/7, d=1/7, u=1/7

Java Tutorial

https://www.codecademy.com/learn/learn-java

- The TrOWL Award
  - for the best reasoner among the assessment submissions
  - to be announced at the revision lecture

UNIVERSITY OF ABERDEEN

Summary

- Knowledge capture
  - Expertise
- Two roles
  - Knowledge engineer
  - Expert
- Many techniques
  - Interview
  - Self report
  - automatic knowledge extraction