Thinking Big: Web Scale AI

Thinking Big: Web Scale AI

Michael Witbrock
Cycorp Europe
witbrock@cycorp.eu
March 13th 2009

Michael Witbrock
Cycorp Europe
witbrock@cycorp.eu
March 13th 2009

The Syntax of CycL

Terms, Predicates, Functions, Connectives, Constraints, Sentences, Rules, Rule Macros and Knowledge and all that

Cycorp.eu

CycL Constants

- … denote specific individuals or collections
  - relations, people, computer programs, types of cars…
- Examples of constants:
  - Collections:
    - #$Dog, #$SnowSkiing, #$RedColor
  - Individuals:
    - #$BillClinton, #$Rover, #$UnitedStatesOfAmerica
  - Relations
    - #$likesAsFriend, #$bordersOn, #$objectHasColor, #$and, #$not, #$implies, #$forAll

Truth Functions

- … are relations that can be used to form sentences.
- … begin with a lower-case letter.

- Types of Truth Functions:
  - Predicates:
    #$likesAsFriend, #$bordersOn, #$objectHasColor, #$isa
  - Logical Connectives:
    - #$and, #$or, #$not, #$implies
  - Quantifiers:
    - #$forAll, #$thereExists

CycL Sentences

- **CycL Formula**:
  - a relation applied to some arguments, enclosed in parentheses
- Examples:
  - (#$isa #$BarakObama #$Person)
  - (#$likesAsFriend #$BarakObama #$GeorgeWBush)
  - (#$BirthFn #$GeorgeWBush)
- **CycL Sentence** is a
  well-formed CycL Formula
  with a Truth Function in the *arg0* position
- CycL Sentences
  … have truth values.
  … are used to form assertions and queries.

Non-atomic Terms

- **Functional Denotation** can be applied to *some arguments* to denote *something*
  - Usually ends in "Fn"
- **CycL Functional Term** is a
  … well-formed CycL Formula
  … with a Function-Denotational in the *arg0* position.
- Examples of functional denotations:
  - #$BirthFn, #$GovernmentFn, #$BorderBetweenFn
- Examples of CycL Non-atomic Terms:
  - (#$BirthFn #$JacquelineKennedyOnassis)
  - (#$GovernmentFn #$France)
  - (#$BorderBetweenFn #$Poland #$Russia)
- CycL Non-atomic Terms can be used like any other, as in:
  - (#$residenceOfOrganization (#$GovernmentFn #$Poland) #$CityOfWarsawPoland)

Important Relations - #$genls

- Relates a given *collection* to those *collections* that subsume it
  - (#$genls SUBCOL SUPERCOL) means that SUPERCOL is a super-collection of SUBCOL
  - anything that is an instance of SUBCOL is also an instance of SUPERCOL

- Examples:
  - (#$genls #$Dog #$Mammal)
  - (#$genls #$UnitedStatesPresident #$HeadOfGovernment)
  - (#$genls #$FamousHuman #$HomoSapiens)
  - (#$genls #$UnitedStatesPresident
    (#$CitizenFn #$UnitedStatesOfAmerica))

Important Relations - #$isa

- Relates things of any kind to *collections* of which they are *instances*.
  - (#$isa THING COL) means that THING is an instance of the collection COL.
- transfers through #$genls relation
  - (#$isa THING COL) and (#$genls COL SUPERCOL) jointly imply (#$isa THING SUPERCOL)
- Examples:
  - (#$isa #$BillClinton #$MaleHuman)
  - (#$isa #$BillClinton #$LeftHandedHuman)
  - (#$isa #$BillClinton #$FamousHuman)
  - (#$isa #$BillClinton
    (#$PastOrPresentPresidentFn #$UnitesStatesOfAmerica))

Important Relations - #$prettyString

- (#$prettyString TERM STRING) means that the STRING names the thing denoted by TERM
- Normally, STRING corresponds to a proper name

- Examples:
  - (#$prettyString #$BillClinton "William Clinton")
  - (#$prettyString #$BillClinton "President Bill Clinton")
  - (#$prettyString #$Dog "dogs")
  - (#$prettyString #$Dog "hound")

Essentially the only NL content in OpenCyc. Research Cyc has far more, as we'll see in May

Well-formedness - #$arity

- Predicate for representing arity constraints
- Examples:
  - (#$arity #$performedBy 2)
    - Represents the fact that #$performedBy takes two arguments
      e.g. (#$performedBy
      #$AssassinationOfPresidentLincoln
      #$JohnWilkesBooth)
  - (#$arity #$JokeAboutFn 1)
    - Represents the fact that #$JokeAboutFn takes one argument
      e.g. (#$JokeAboutFn #$BillClinton)
  - (#$arity #$between 3)
    - Represents the fact that #$between takes three arguments
      e.g. (#$between #$PlanetMars #$PlanetJupiter #$AsteroidBelt)

Arity of #$arity

What's the arity of #$arity?

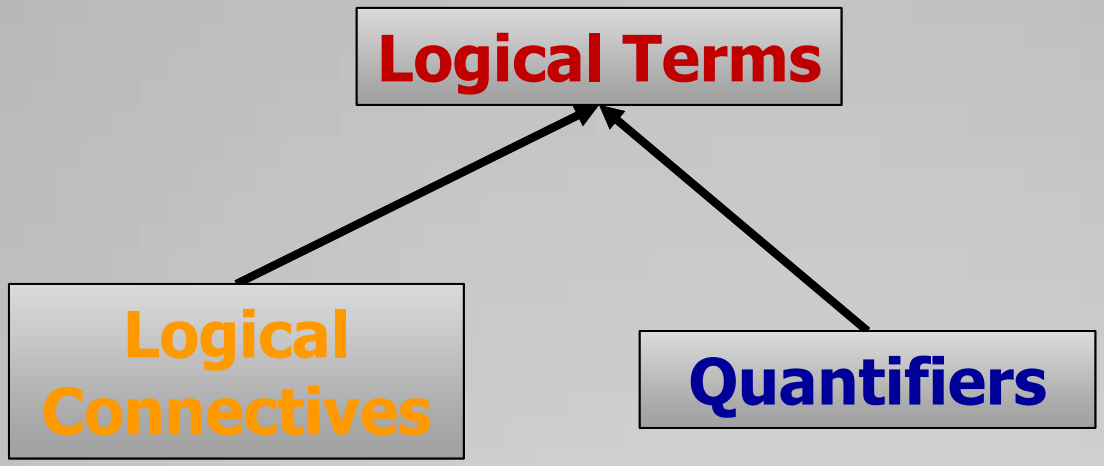
```
(#$arity  #$arity  2)
```

Well-Formedness - Argument Type Constraints

- #$argIsa
  - Tells individual of which collection is the argument
  - (#$argIsa #$performedBy 1 #$Action)
    - the first argument of #$performedBy must be an individual #$Action
- #$argGenl
  - Tells which collection subsumes the argument
  - (#$argGenl #$skillCapableOf 2 #$SkilledActivity)
    - the second argument of #$skillCapableOf must be a type of #$SkilledActivity

Complex Formulas

- CycL includes logical terms to allow us to
  - connect formulas
  - quantify into them



Logical Connectives

#$And, #$**Or**, #$**Not**, #$**Implies**

- Are Truth Functions
  - the truth value of the whole sentence is determined by truth value(s) of constituent sentences
- Take sentences as their arguments

Examples with Logical Connectives

- Examples:
  - (#$**and**
    (#$performedBy #$GettysburgAddress #$Lincoln)
    (#$objectHasColor #$Rover #$TanColor))
  - (#$**or**
    (#$objectHasColor #$Rover #$TanColor)
    (#$objectHasColor #$Rover #$BlackColor))
  - (#$**not**
    (#$performedBy #$GettysburgAddress #$BillClinton))
  - (#$**implies**
    (#mainColorOfObject #$Rover #$TanColor)
    (#$**not** (#$mainColorOfObject #$Rover #$RedColor)))

Quantification

- Universal quantification
  - E.g. All dogs have ears.
- Existential quantification
  - E.g. Everybody is loved by someone.

Universal Quantification

- Corresponds to English expressions like:
  - Every, All, Always, Everyone, Anything
- Examples:
  - All dogs have ears.
    - $\forall x\ (\text{Dog}(x) \rightarrow \text{HasEars}(x))$
  - Every person in this room is alive.
    - $\forall x\ ((\text{Person}(x) \ \&\ \text{InThisRoom}\ (x)) \rightarrow \text{Alive}\ (x))$
  - Anything which is in my house is mine.
    - $\forall x\ (\text{LocatedIn}(x, \textit{HouseOfWitbrock}) \rightarrow$

      $\text{BelongsTo}(x, \textit{Witbrock}))$

Rules using Universal Quantification

- English:
  - All dogs have ears.
- CycL:
  - (#$forAll ?DOG
    (#$implies
      (#$isa ?DOG #$Dog)
      (#$anatomicalBodyParts ?DOG #$Ear)))

Rules using Universal Quantification

- English:
  - Every person in this room is alive.
- CycL:

```
(#$forAll ?PERSON
   (#$implies
      (#$and
         (#$isa ?PERSON #$Person)
         (#$objectFoundInLocation
                        ?PERSON $Room))
      (#$isa ?PERSON #$Alive)))
```

Unbound Variables in Rules

- Unbound variables are treated as implicitly universally quantified
- Example:
  - (#$forAll ?DOG
    (#$implies
      (#$isa ?DOG #$Dog)
      (#$anatomicalBodyParts ?DOG #$Ear)))
- can be written as
  - (#$implies
    (#$isa ?DOG #$Dog)
    (#$anatomicalBodyParts ?DOG #$Ear))

Existential Quantification

- Corresponds to English expressions like:
  - There is (a/an)…, Some, Someone, Something, Somewhere
- Examples:
  - Someone is sitting in Bill's chair.
    - $\exists$ x (Person(x) & SittingIn(x, BillsChair))
  - Bill left his keys somewhere.
    - $\exists$ x (Place(x) & LeftObjectAt (Bill, BillsKeys, x))
  - Everybody is loved by someone.
    - $\forall x$ (Person(x) $\rightarrow$ $\exists y$ (Person(y) & Loves (y, x)))

Example of Existential Quantification

- English:
  - Someone is sitting in Bill's chair.
- CycL:

```
(#$thereExists ?PERSON
   (#$and
      (#$isa ?PERSON #$Person)
      (#$postureOfObject
                ?PERSON #$SittingPosture)
      (#$objectFoundInLocation
                ?PERSON #$BillsChair))))
```

Rule using Existential Quantification

- English:
  - Everybody is loved by someone.
- CycL:

```
(#$forAll ?PERSON
   (#$implies
      (#$isa ?PERSON #$Person)
      (#$thereExists ?LOVER
         (#$and
            (#$isa ?LOVER #$Person)
            (#$loves ?LOVER  ?PERSON))))
```

Other Quantification

- (#$thereExistExactly 12 ?ZOS
  (#$isa ?ZOS #$ZodiacSign))

- (#$thereExistAtLeast 8 ?PLNT
  (#$isa ?PLNT #$Planet))

Overview of OpenCyc Content

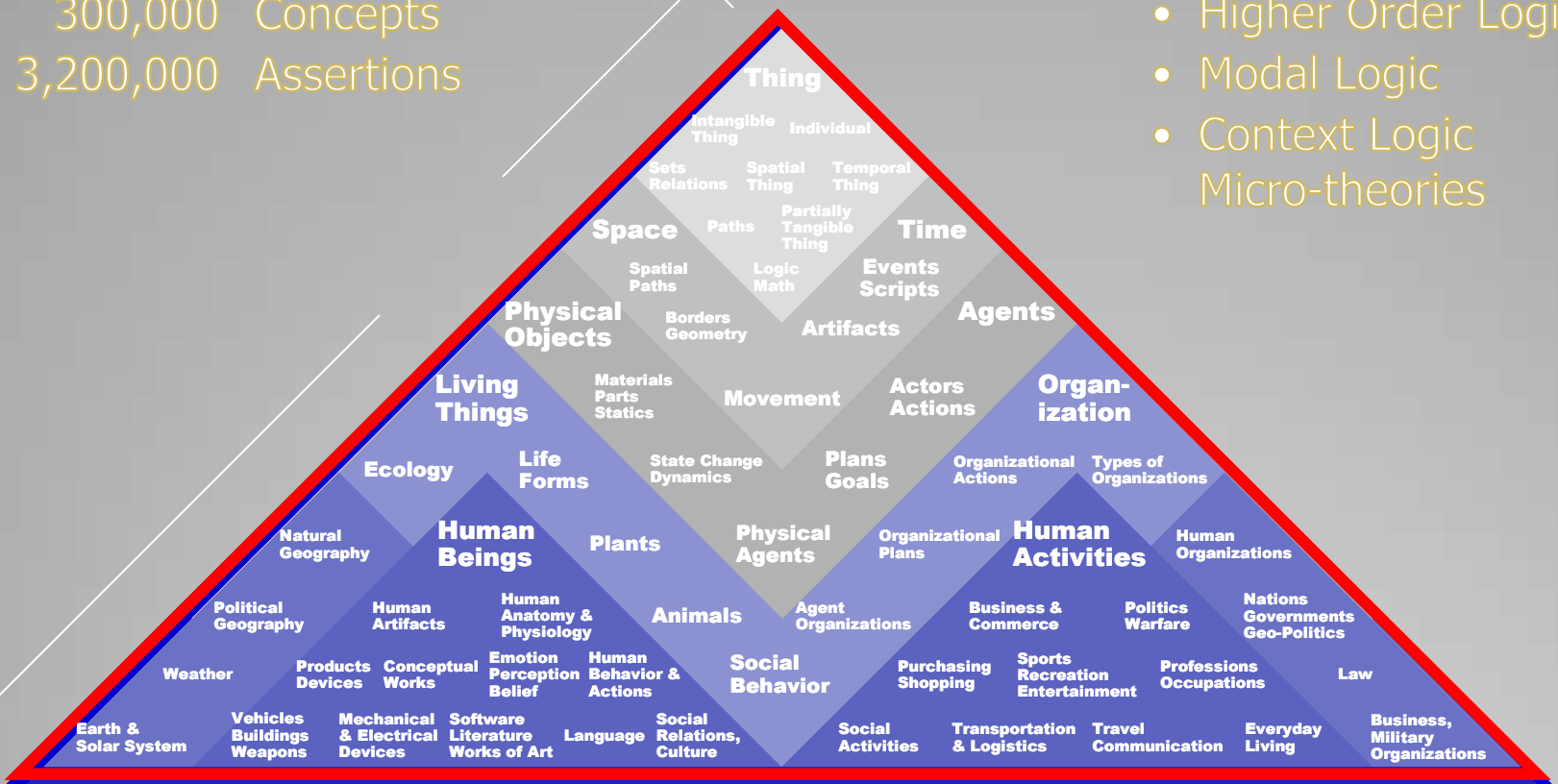
What's there, and what's elsewhere

Cyc contains:

15,000 Predicates
300,000 Concepts
3,200,000 Assertions

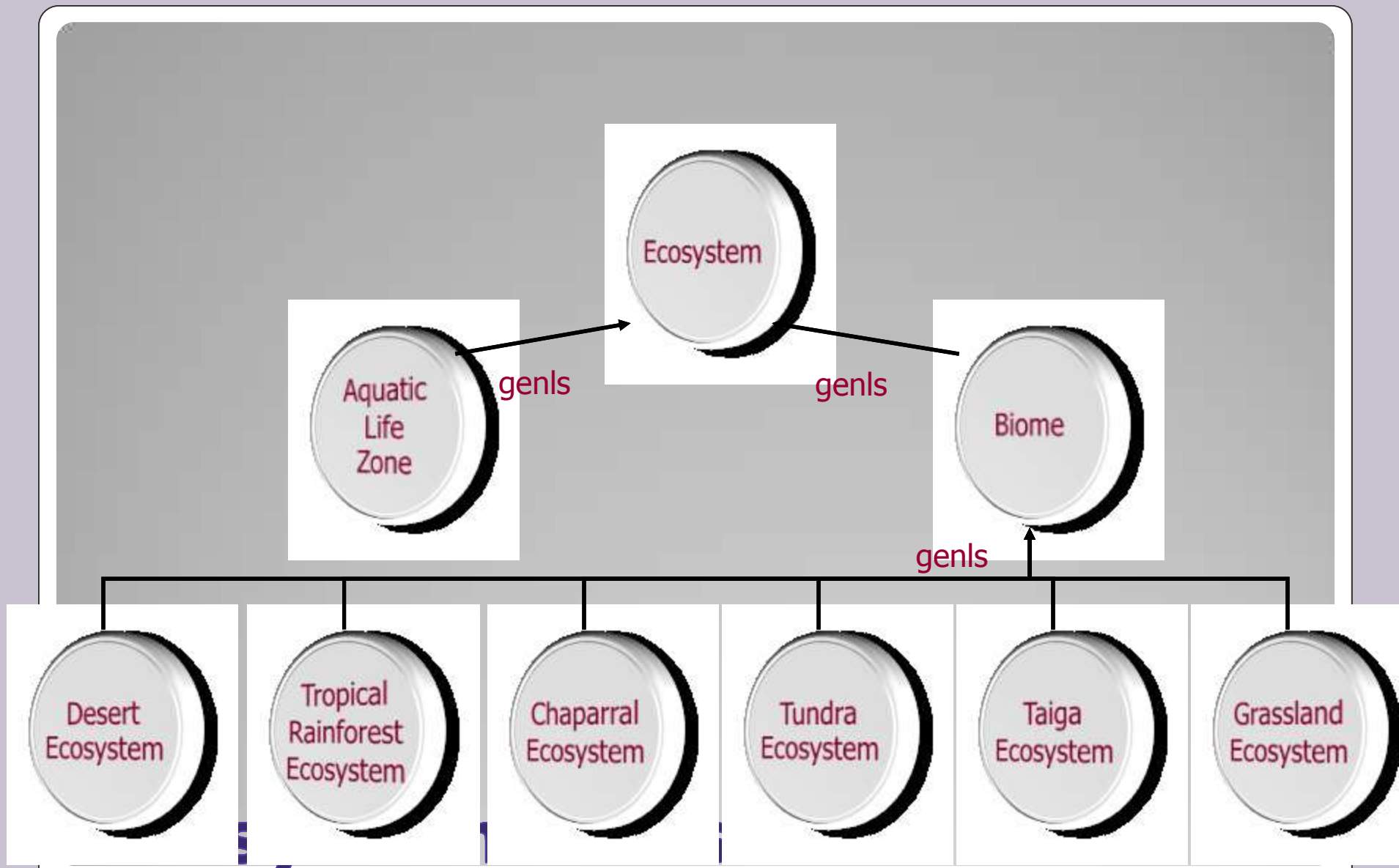
Represented in:

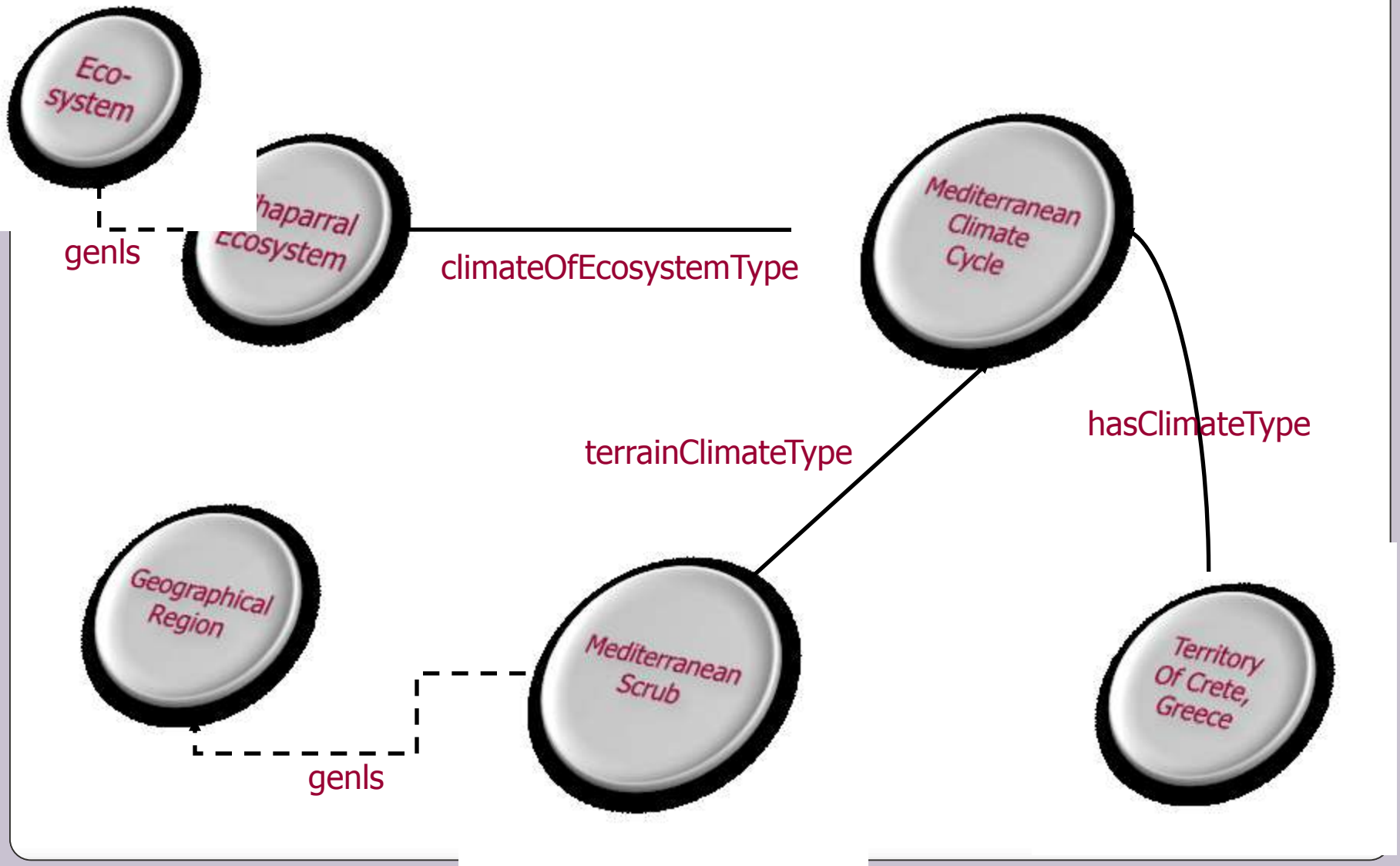
- First Order Logic
- Higher Order Logic
- Modal Logic
- Context Logic
  Micro-theories



General Knowledge about Various Domains

Specific data, facts, and observations


Ecosystem
Aquatic Life Zone
genls
genls
Biome
genls
Desert Ecosystem
Tropical Rainforest Ecosystem
Chaparral Ecosystem
Tundra Ecosystem
Taiga Ecosystem
Grassland Ecosystem


Eco-
system
genls
haparral
Ecosystem
climateOfEcosystemType
Mediterranean
Climate
Cycle
hasClimateType
terrainClimateType
Geographical
Region
Mediterranean
Scrub
genls
Territory
Of Crete,
Greece

- #$TransportationEvent
- #$ControllingATransportationDevice
- #$TransportWithMotorizedLandVehicle
- (#$SteeringFn #$RoadVehicle)
- #$TransporterCrashEvent
- #$VehicleAccident
- #$CarAccident
- #$Colliding
- #$IncurringDamage
- #$TippingOver
- #$Navigating
- #$EnteringAVehicle …

Some Transportation Event Types

Relating Events and Participants

- #$performedBy
- #$causes-EventEvent
- #$objectPlaced
- #$objectOfStateChange
- #$outputsCreated
- #$inputsDestroyed
- #$assistingAgent
- #$beneficiary
- #$fromLocation
- #$toLocation
- #$deviceUsed
- #$driverActor
- #$damages
- #$vehicle
- #$providerOfMotiveForce
- #$transportees …

Over 400 more.

Event Types

- PhysicalStateChangeEvent
- TemperatureChangingProcess
- BiologicalDevelopmentEvent
- ShapeChangeEvent
- MovementEvent
- ChangingDeviceState
- GivingSomething
- DiscoveryEvent
- Cracking
- Carving
- Buying
- Thinking
- Mixing
- Singing
- CuttingNails
- PumpingFluid

- **over 11,000 more**

- governingBody
- WholeOrganizationFn
- parentCompany
- subOrgs-Command
- subOrgs-Permanent
- subOrgs-Temporary
- subOrgs-OnlyDuringOperation
- physicalQuarters
- hasHeadquartersInCountry
- officeInCountry
- memberTypes
- organizationHead
- PolicyFn

Organizational Relations

Emotions

- Types of Emotions:
  - **Adulation**
  - **Abhorrence**
  - **Relaxed-Feeling**
  - **Gratitude**
  - **Anticipation-Feeling**
  - Over 120 of these
- Predicates for Defining and Attributing Emotions:
  - **contraryFeelings**
  - **appropriateEmotion**
  - **actionExpressesFeeling**
  - **feelsTowardsObject**
  - **feelsTowardsPersonType**

Relations between Agents and Propositions

- goals
- intends
- desires
- hopes
- expects
- beliefs
- opinions
- knows
- rememberedProp
- perceivesThat
- seesThat
- tastesThat

Propositional Attitudes

Biology

- Organisms classified by:
  - Taxon
  - Habitat
  - Source of Nutrients
- Organism Anatomy
  - Gross Anatomy
  - Cell biology
  - Physiological Processes
- Medicine
  - Cardio-thoracic surgery
  - Respiratory system
  - …

- Common Substances
- Attributes of Materials
- States Of Matter
  - SolidStateOfMatter
  - LiquidStateOfMatter
  - GaseousStateOfMatter
- Solutions
- Electrical Conductivity
- Thermal Conductivity
- Structural Attributes
- Tangible Attributes
  - SolidTangibleThing
  - LiquidTangibleThing
  - GaseousTangibleThing

Materials

Devices

- Over 4000 Specializations of PhysicalDevice
  - **ClothesWasher**
  - **NuclearAircraftCarrier**
- Vocabulary for Describing device functions
  - **primaryFunction-DeviceType**

- Device Specific Predicates
  - **gunCaliber**
  - **maximumSpeedOf**
- Device States (40+)
  - **DeviceOn**
  - **CockedState**

Weather

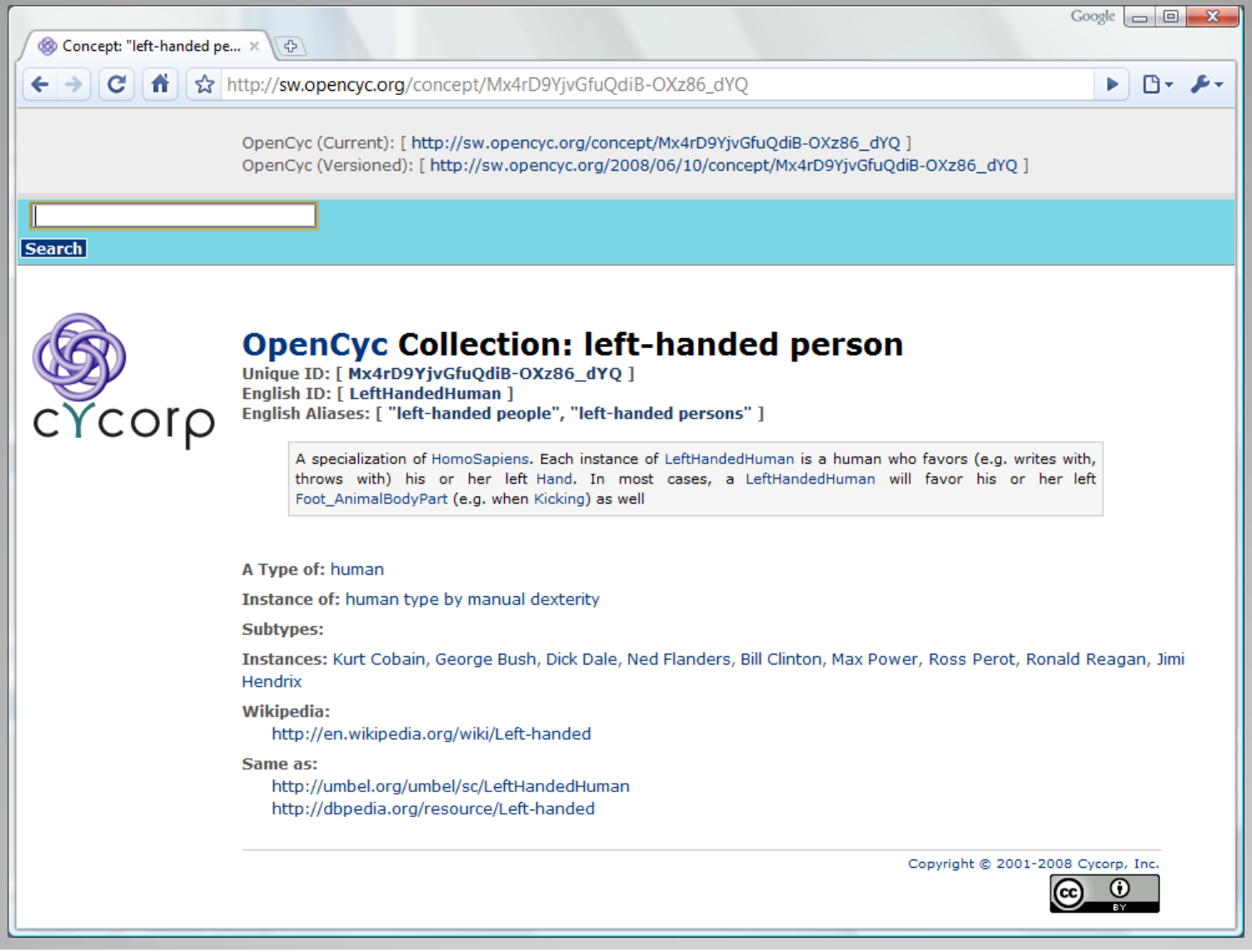
- Weather Objects
  - CloudInSky
  - SnowMob
- Weather Events
  - TornadoAsEvent
  - SnowProcess
- Weather Attributes
  - ClearWeather
  - (LowAmountFn Raininess)

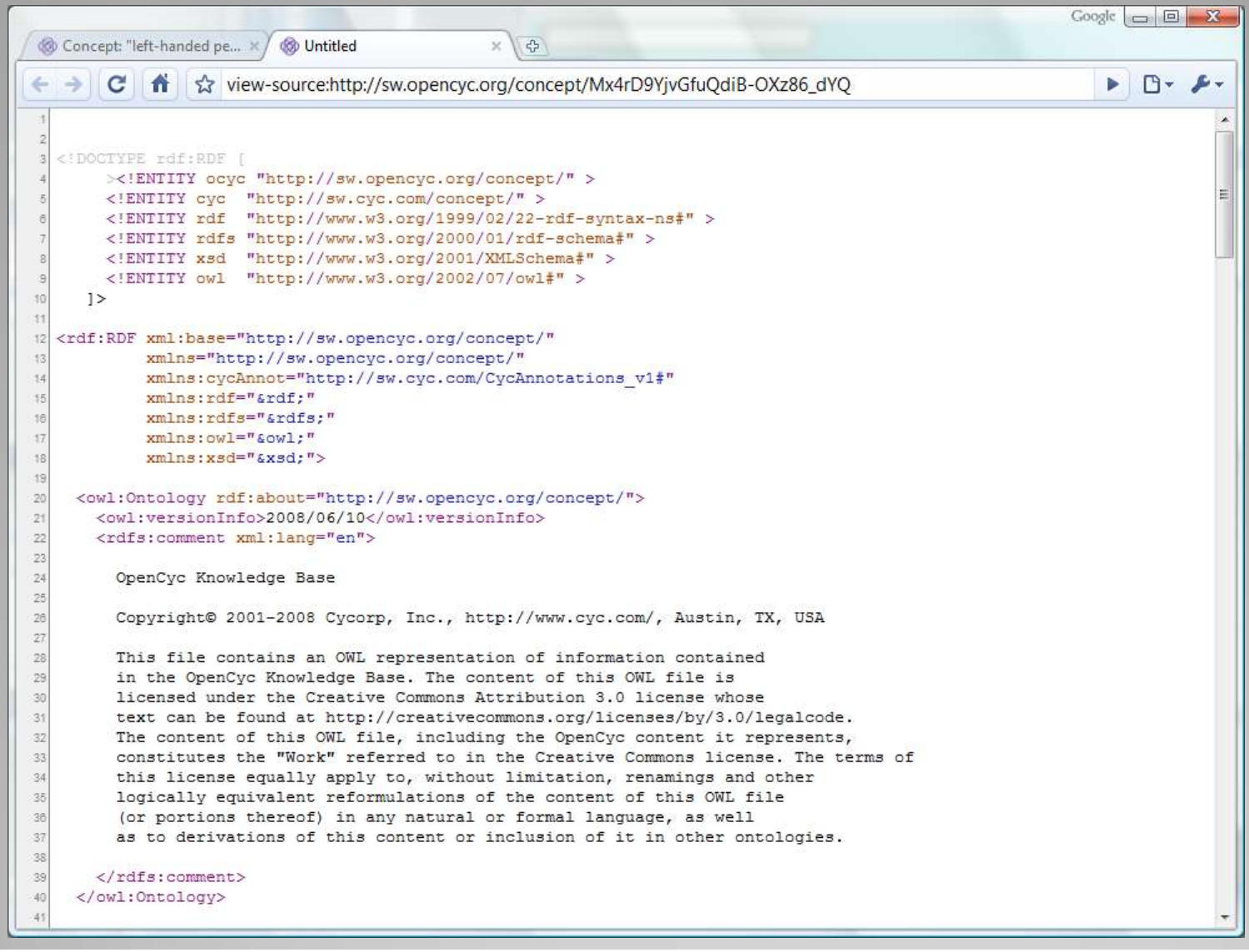
As of Feb (#$February). 24 (24), Air Force (#$UnitedStatesAirForce) officials (#$PublicOfficial #$OrganizationRepresentative) reported (#$RegisteringAComplaint #$Reporting) that personnel (#$Employee) in the area (#$Area 0 #$FieldOfStudy #$Region-Underspecified) numbered (#$Counting) close to 8,000 (8000). The 100 (100) aircraft (#$AirTransportationDevice) based (#$Base-Support #$MilitaryBase-Grounds #$BaseOfLandProtrusion #$NitrogenBase #$ChemicallyBasicSubstance) in Saudi Arabia (#$SaudiArabia) for patrols (#$Patrolling) over southern Iraq ((#$SouthernRegionFn #$Iraq)) has (#$possesses) seen (#$VisualPerception #$MeetingSomeone #$sees) the addition (#$DoingAddition) of two (2) dozen (12) F-15 (#$FighterPlane-F15) and F-16 fighter jets (#$FighterPlane-F16) to Bahrain (#$Bahrain-TheIsland #$Bahrain (#$CityNamedFn Bahrain #$Bahrain)). The Air Force (#$UnitedStatesAirForce) has (#$possesses) also authorized (#$GrantingPermission) the dispatch (#$SendingSomething) of 12 (12) F-117 (#$FighterPlane-F117) stealth (#$DodgeStealthCar) fighter jets (#$JetOfFluid #$JetPropelledAircraft) to Kuwait (#$CityOfKuwaitKuwait (#$ProperSubcollectionNamedFn-Ternary kuwait #$Individual 34057665-f4ed-11d9-9bea-0002b3a85b0b) #$Kuwait), three (3) B-1 bombers (#$B-1-Bomber) to Bahrain (#$Bahrain-TheIsland #$Bahrain (#$CityNamedFn Bahrain #$Bahrain)) and 14 (14) B-52 (#$B-52-Bomber) bombers (#$SubmarineSandwich #$BomberPlane #$Bomber) to the island (#$Island) of Diego Garcia. It also has (#$possesses) diverted (#$AmusingSomeone #$DivertingSomething) dozens (#$Dozens-Quant 12) of support (#$SupportingSomething #$ShowingSupportForSomeone (#$SubcollectionOfWithRelationFromTypeFn #$PartiallyTangible #$supportingObject #$SupportingSomething)) aircraft (#$AirTransportationDevice) to the region (#$TheRegion) for refueling (#$Refueling (#$MakingAvailableFn #$CombustibleFuelSubstance)),

Entire Cyc KB

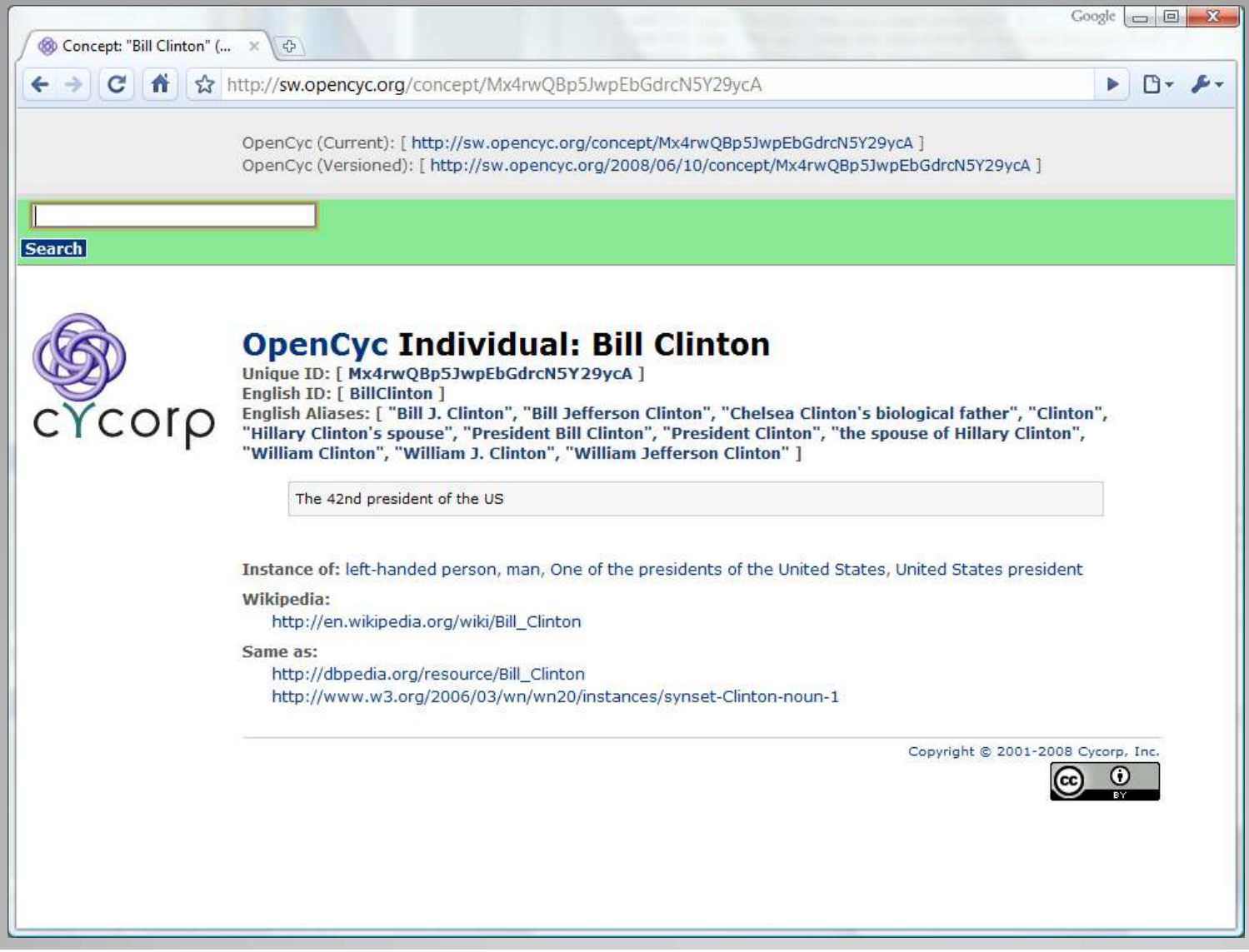
Cyc Content in RDFS and OWL

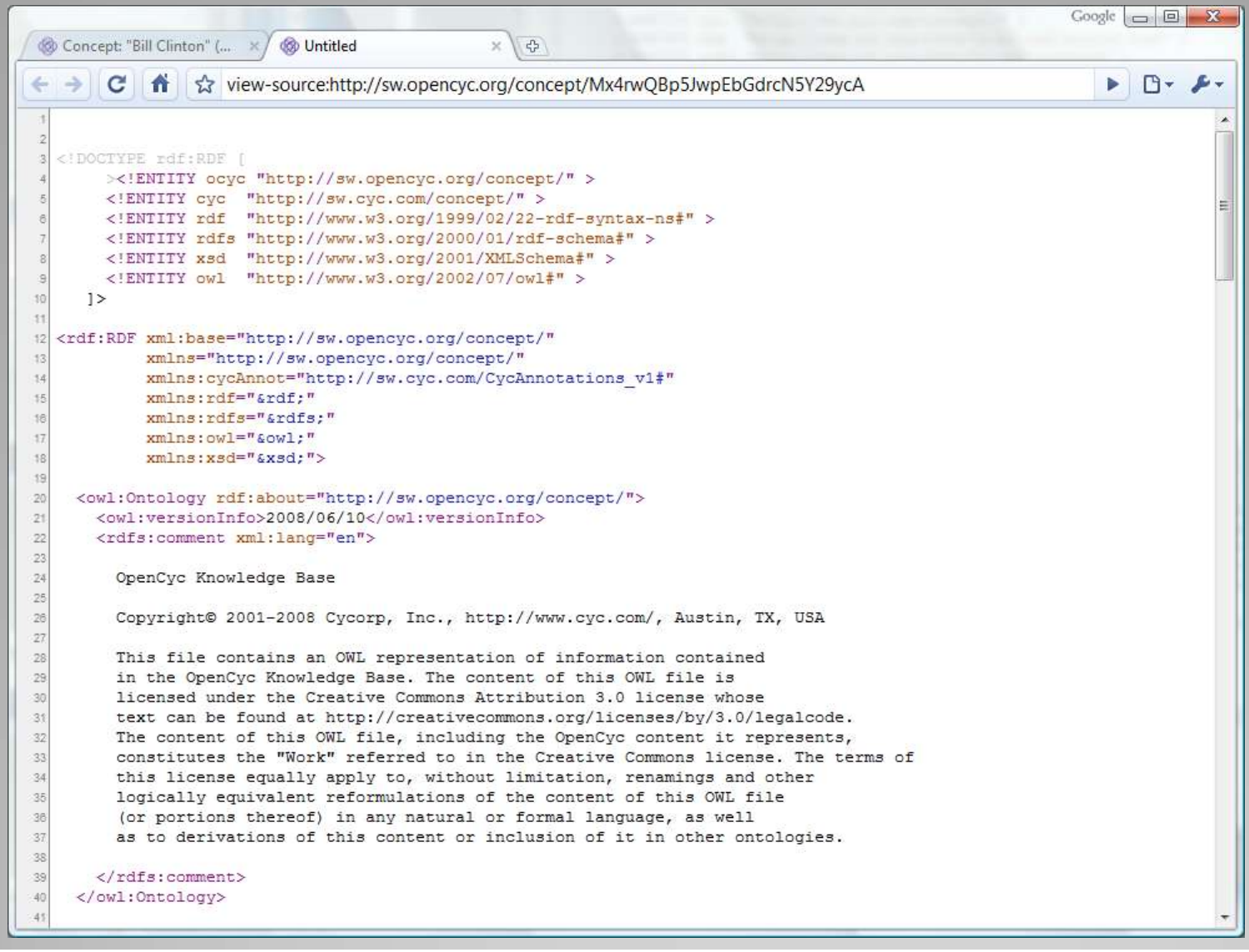
The Syntax of OpenCyc


Google
Concept: "left-handed pe...
http://sw.opencyc.org/concept/Mx4rD9YjvGfuQdiB-OXz86_dYQ
OpenCyc (Current): [http://sw.opencyc.org/concept/Mx4rD9YjvGfuQdiB-OXz86_dYQ]
OpenCyc (Versioned): [http://sw.opencyc.org/2008/06/10/concept/Mx4rD9YjvGfuQdiB-OXz86_dYQ]
Search
cYcorp
OpenCyc Collection: left-handed person
Unique ID: [Mx4rD9YjvGfuQdiB-OXz86_dYQ]
English ID: [LeftHandedHuman]
English Aliases: ["left-handed people", "left-handed persons"]
A specialization of HomoSapiens. Each instance of LeftHandedHuman is a human who favors (e.g. writes with, throws with) his or her left Hand. In most cases, a LeftHandedHuman will favor his or her left Foot_AnimalBodyPart (e.g. when Kicking) as well
A Type of: human
Instance of: human type by manual dexterity
Subtypes:
Instances: Kurt Cobain, George Bush, Dick Dale, Ned Flanders, Bill Clinton, Max Power, Ross Perot, Ronald Reagan, Jimi Hendrix
Wikipedia:
http://en.wikipedia.org/wiki/Left-handed
Same as:
http://umbel.org/umbel/sc/LeftHandedHuman
http://dbpedia.org/resource/Left-handed
Copyright © 2001-2008 Cycorp, Inc.

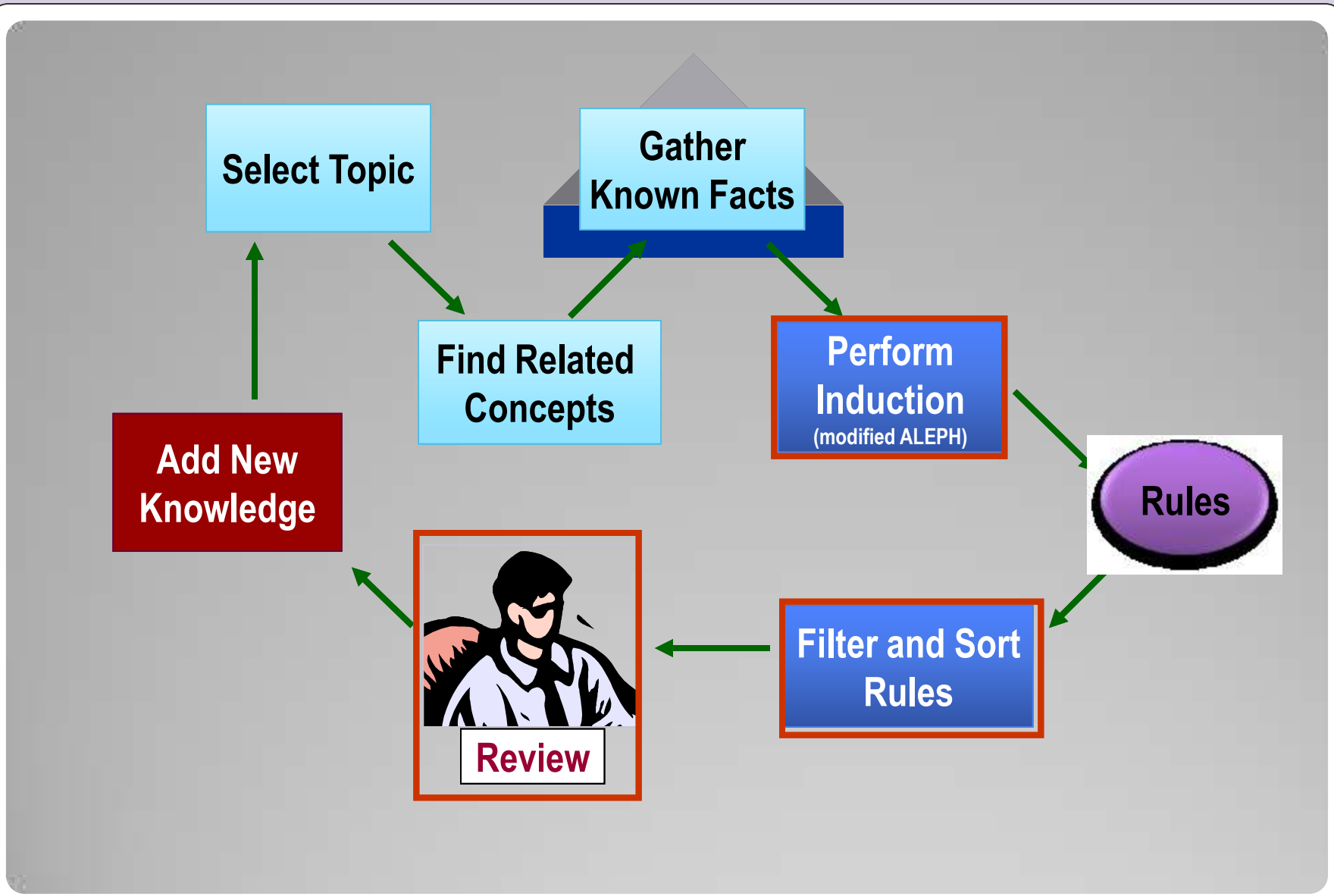



```
<!DOCTYPE rdf:RDF [
     ><!ENTITY ocyc "http://sw.opencyc.org/concept/" >
     <!ENTITY cyc  "http://sw.cyc.com/concept/" >
     <!ENTITY rdf  "http://www.w3.org/1999/02/22-rdf-syntax-ns#" >
     <!ENTITY rdfs "http://www.w3.org/2000/01/rdf-schema#" >
     <!ENTITY xsd  "http://www.w3.org/2001/XMLSchema#" >
     <!ENTITY owl  "http://www.w3.org/2002/07/owl#" >
   ]>

<rdf:RDF xml:base="http://sw.opencyc.org/concept/"
         xmlns="http://sw.opencyc.org/concept/"
         xmlns:cycAnnot="http://sw.cyc.com/CycAnnotations_v1#"
         xmlns:rdf="&rdf;"
         xmlns:rdfs="&rdfs;"
         xmlns:owl="&owl;"
         xmlns:xsd="&xsd;">

  <owl:Ontology rdf:about="http://sw.opencyc.org/concept/">
    <owl:versionInfo>2008/06/10</owl:versionInfo>
    <rdfs:comment xml:lang="en">

      OpenCyc Knowledge Base

      Copyright© 2001-2008 Cycorp, Inc., http://www.cyc.com/, Austin, TX, USA

      This file contains an OWL representation of information contained
      in the OpenCyc Knowledge Base. The content of this OWL file is
      licensed under the Creative Commons Attribution 3.0 license whose
      text can be found at http://creativecommons.org/licenses/by/3.0/legalcode.
      The content of this OWL file, including the OpenCyc content it represents,
      constitutes the "Work" referred to in the Creative Commons license. The terms of
      this license equally apply to, without limitation, renamings and other
      logically equivalent reformulations of the content of this OWL file
      (or portions thereof) in any natural or formal language, as well
      as to derivations of this content or inclusion of it in other ontologies.

    </rdfs:comment>
  </owl:Ontology>
```

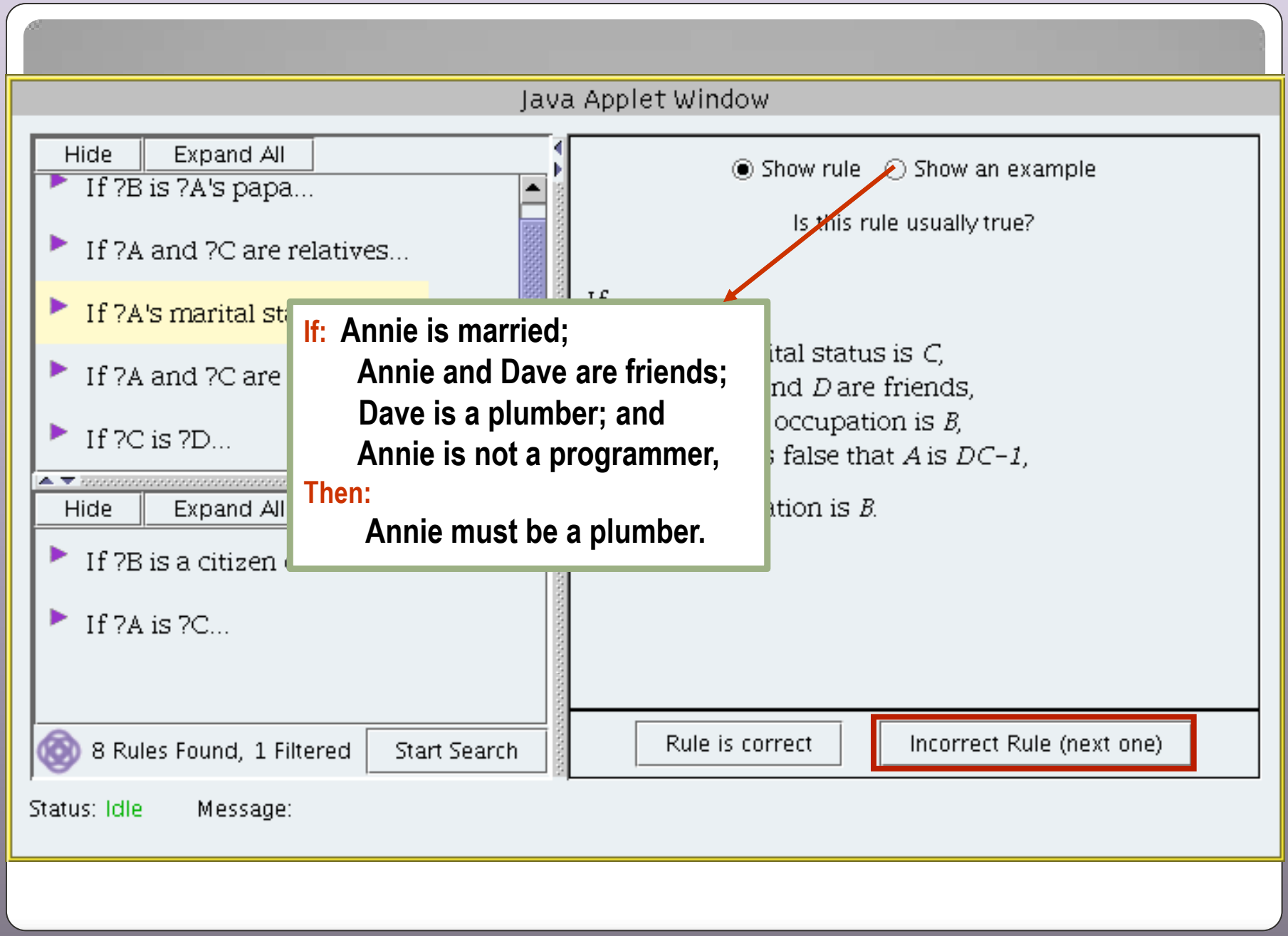

Google
Concept: "Bill Clinton" (...
http://sw.opencyc.org/concept/Mx4rwQBp5JwpEbGdrcN5Y29ycA
OpenCyc (Current): [http://sw.opencyc.org/concept/Mx4rwQBp5JwpEbGdrcN5Y29ycA]
OpenCyc (Versioned): [http://sw.opencyc.org/2008/06/10/concept/Mx4rwQBp5JwpEbGdrcN5Y29ycA]
Search
cYcorp
OpenCyc Individual: Bill Clinton
Unique ID: [Mx4rwQBp5JwpEbGdrcN5Y29ycA]
English ID: [BillClinton]
English Aliases: ["Bill J. Clinton", "Bill Jefferson Clinton", "Chelsea Clinton's biological father", "Clinton", "Hillary Clinton's spouse", "President Bill Clinton", "President Clinton", "the spouse of Hillary Clinton", "William Clinton", "William J. Clinton", "William Jefferson Clinton"]
The 42nd president of the US
Instance of: left-handed person, man, One of the presidents of the United States, United States president
Wikipedia:
http://en.wikipedia.org/wiki/Bill_Clinton
Same as:
http://dbpedia.org/resource/Bill_Clinton
http://www.w3.org/2006/03/wn/wn20/instances/synset-Clinton-noun-1
Copyright © 2001-2008 Cycorp, Inc.

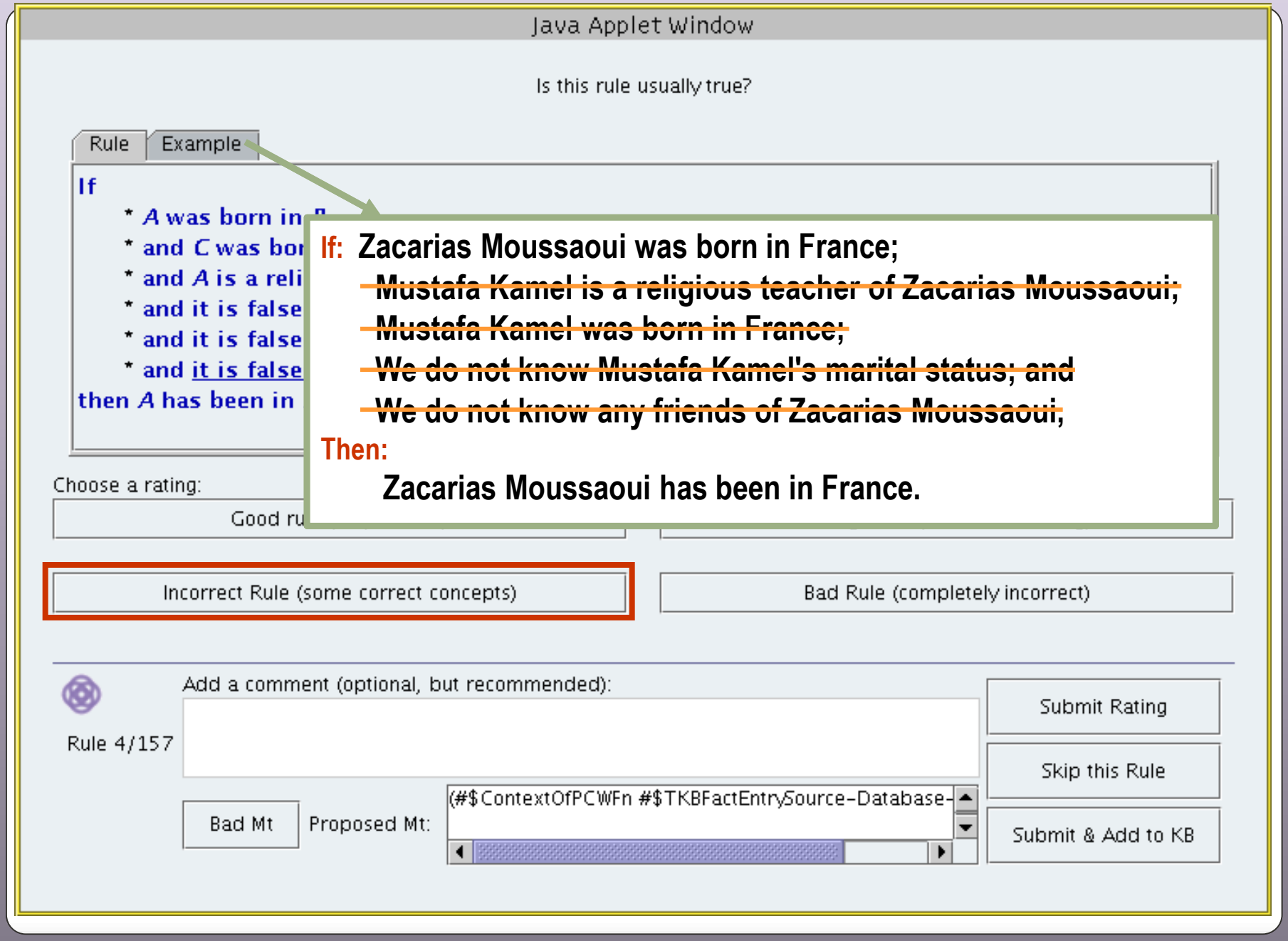



```
<!DOCTYPE rdf:RDF [
     ><!ENTITY ocyc "http://sw.opencyc.org/concept/" >
     <!ENTITY cyc  "http://sw.cyc.com/concept/" >
     <!ENTITY rdf  "http://www.w3.org/1999/02/22-rdf-syntax-ns#" >
     <!ENTITY rdfs "http://www.w3.org/2000/01/rdf-schema#" >
     <!ENTITY xsd  "http://www.w3.org/2001/XMLSchema#" >
     <!ENTITY owl  "http://www.w3.org/2002/07/owl#" >
   ]>

<rdf:RDF xml:base="http://sw.opencyc.org/concept/"
         xmlns="http://sw.opencyc.org/concept/"
         xmlns:cycAnnot="http://sw.cyc.com/CycAnnotations_v1#"
         xmlns:rdf="&rdf;"
         xmlns:rdfs="&rdfs;"
         xmlns:owl="&owl;"
         xmlns:xsd="&xsd;">

  <owl:Ontology rdf:about="http://sw.opencyc.org/concept/">
    <owl:versionInfo>2008/06/10</owl:versionInfo>
    <rdfs:comment xml:lang="en">

      OpenCyc Knowledge Base

      Copyright© 2001-2008 Cycorp, Inc., http://www.cyc.com/, Austin, TX, USA

      This file contains an OWL representation of information contained
      in the OpenCyc Knowledge Base. The content of this OWL file is
      licensed under the Creative Commons Attribution 3.0 license whose
      text can be found at http://creativecommons.org/licenses/by/3.0/legalcode.
      The content of this OWL file, including the OpenCyc content it represents,
      constitutes the "Work" referred to in the Creative Commons license. The terms of
      this license equally apply to, without limitation, renamings and other
      logically equivalent reformulations of the content of this OWL file
      (or portions thereof) in any natural or formal language, as well
      as to derivations of this content or inclusion of it in other ontologies.

    </rdfs:comment>
  </owl:Ontology>
```



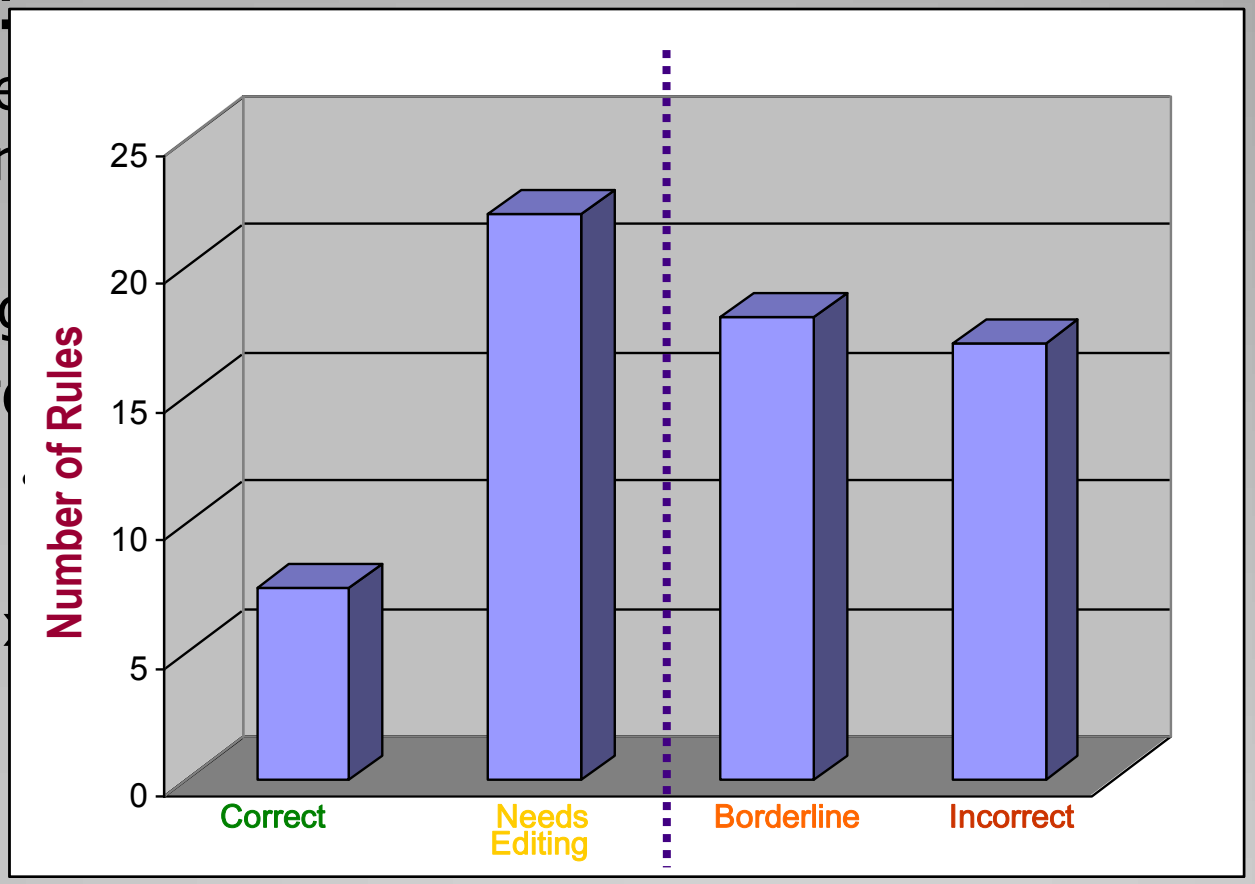
Rule Induction at Scale


Java Applet Window
Hide
Expand All
If ?B is ?A's papa...
If ?A and ?C are relatives...
If ?A's marital st
If ?A and ?C are
If ?C is ?D...
Hide
Expand All
If ?B is a citizen
If ?A is ?C...
8 Rules Found, 1 Filtered
Start Search
Show rule
Show an example
Is this rule usually true?
If: Annie is married;
Annie and Dave are friends;
Dave is a plumber; and
Annie is not a programmer,
Then:
Annie must be a plumber.
ital status is C,
nd D are friends,
occupation is B,
false that A is DC-1,
ation is B.
Rule is correct
Incorrect Rule (next one)
Status: Idle
Message:


Java Applet Window
Is this rule usually true?
Rule
Example
If
* A was born in
* and C was bor
* and A is a reli
* and it is false
* and it is false
* and it is false
then A has been in
If: Zacarias Moussaoui was born in France;
Mustafa Kamel is a religious teacher of Zacarias Moussaoui;
Mustafa Kamel was born in France;
We do not know Mustafa Kamel's marital status; and
We do not know any friends of Zacarias Moussaoui;
Then:
Zacarias Moussaoui has been in France.
Choose a rating:
Good ru
Incorrect Rule (some correct concepts)
Bad Rule (completely incorrect)
Add a comment (optional, but recommended):
Rule 4/157
Submit Rating
Skip this Rule
Bad Mt
Proposed Mt:
(#$ContextOfPCWFn #$TKBFactEntrySource-Database-
Submit & Add to KB

Early Results

- ~150 rules created in a variety of domains:
  - Family re
    affiliation
    …
  - Changing
- Results r
  - Overall,
    into KB;
    (by an e



- If someone's time has been requested for a task by that person's primary project, the time will be assigned.

```
(implies
      (and
         (cyclistPrimaryProject ?KE ?PROJECT)
         (projectTasks ?PROJECT ?TASK)
         (requestedEffortPercent ?TASK ?KE ?X))
      (assignedEffortPercent ?TASK ?KE ?X))
```

- People participate in the projects they manage.

```
(implies
      (and
         (projectManagers ?PROJECT ?AGENT))
```

- People are assigned to tasks requested of them for projects managed by that person's direct supervisor.

```
(implies
      (and
         (primarySupervisor ?AGENT AGENT-1)
         (requestedEffortPercent ?TASK ?AGENT ?X)
         (projectManagers ?PROJECT ?AGENT-1)
         (projectTasks ?PROJECT ?TASK))
      (assignedEffortPercent ?TASK ?AGENT ?X))
```

Sample Rules Produced

Competition

Removing

"cut"

Incision

Injury

Running a Foot Race

"race"

Candidate

Future: Complex texts with rich background models:

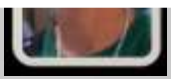
- Patent matching
- Myth understanding: Norse c.f. Greek c.f. Japanese
- Event recognition (piracy vs …)

'08 Re Pri

thesis Of: 2007

"Nov."

"Gov. Romney



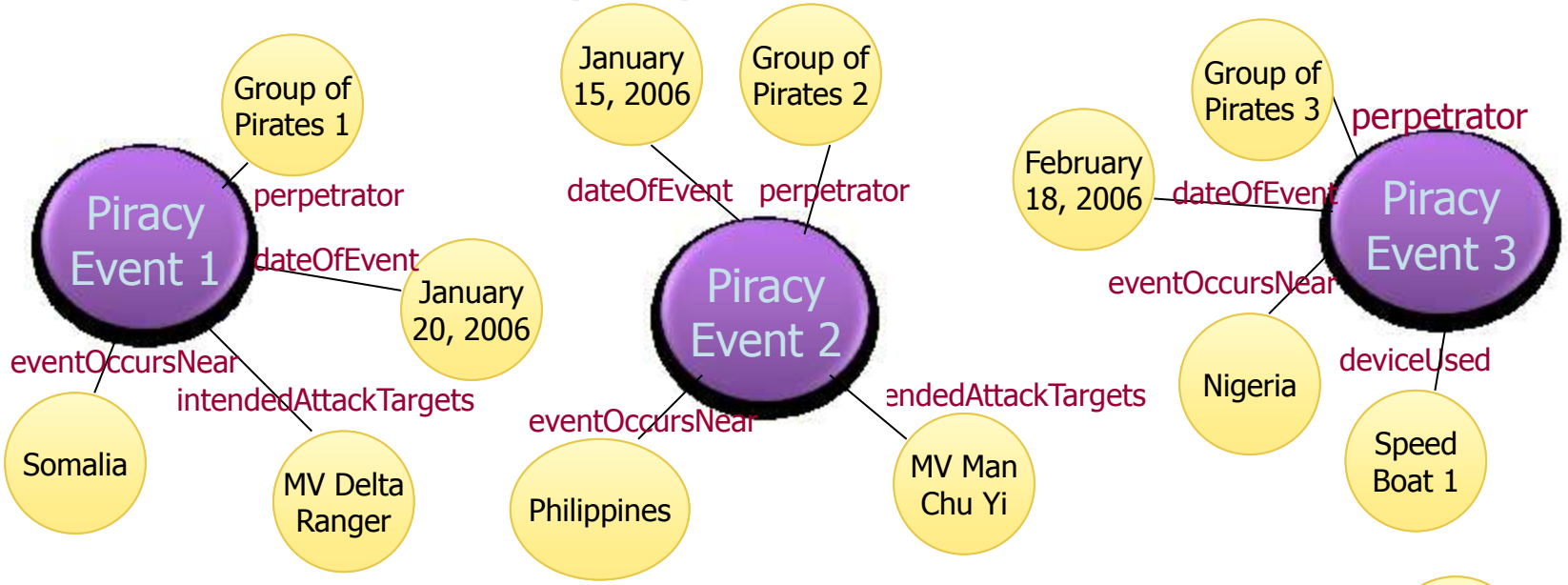
"Wariner"

Gov. Romney cut in Nov. race.

Wariner cut in Nov. race.

Willard Romney Picture by Ann Marie Curling http://blog.electromneyin2008.com/; Jeremy Wariner Picture is public domain

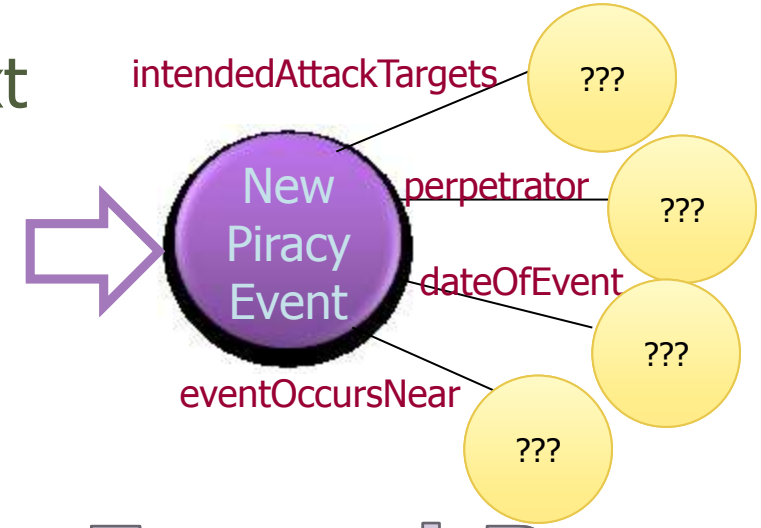
Given a set of formally represented events…



…recognize new instances in text

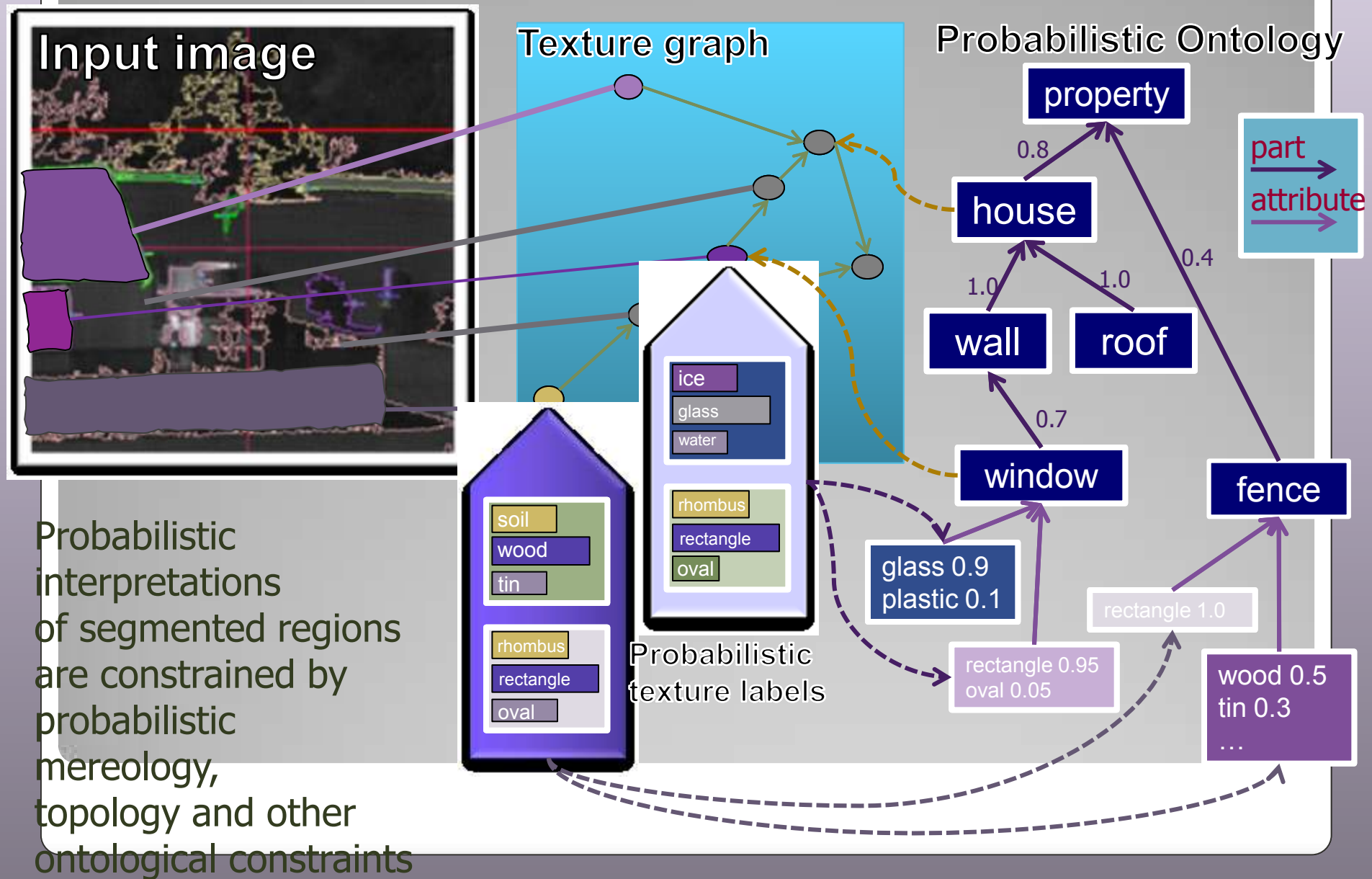
Malacca Straits:

On 17 April 2006, a Malaysian fishing vessel was attacked by armed pirates at approximately nine nautical miles off Parit Haji Baki coast in the Malacca Straits at about 0200 Hrs LT. Six pirates armed with guns in a speedboat closed in rapidly and opened fire at the fishing vessel underway. Several shots hit the side of the vessel but the crew escaped injuries. The fishing vessel crew lodged a police report.

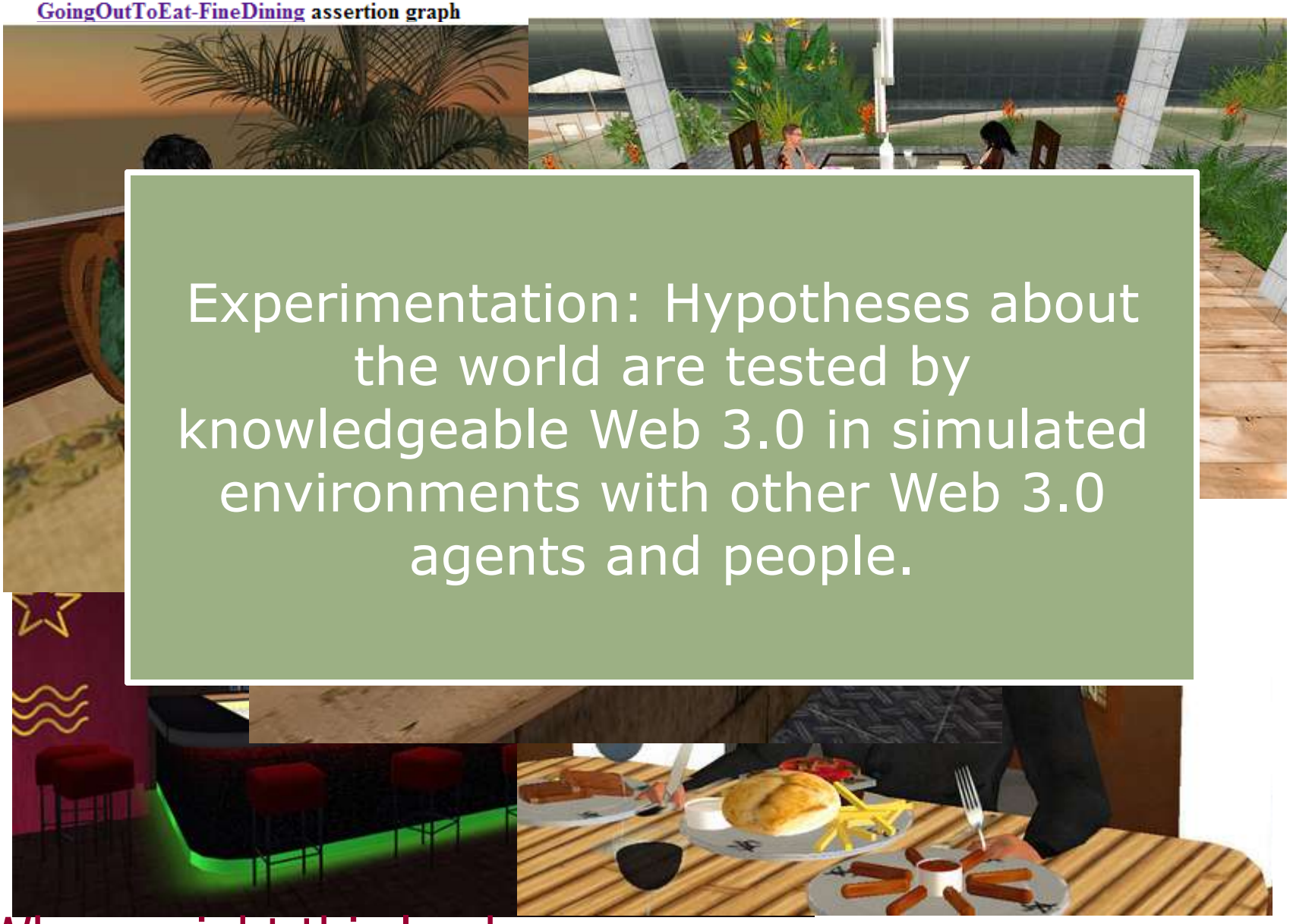


Use of Probabilistic Formal Rep.

Background Knowledge for Visual Situations



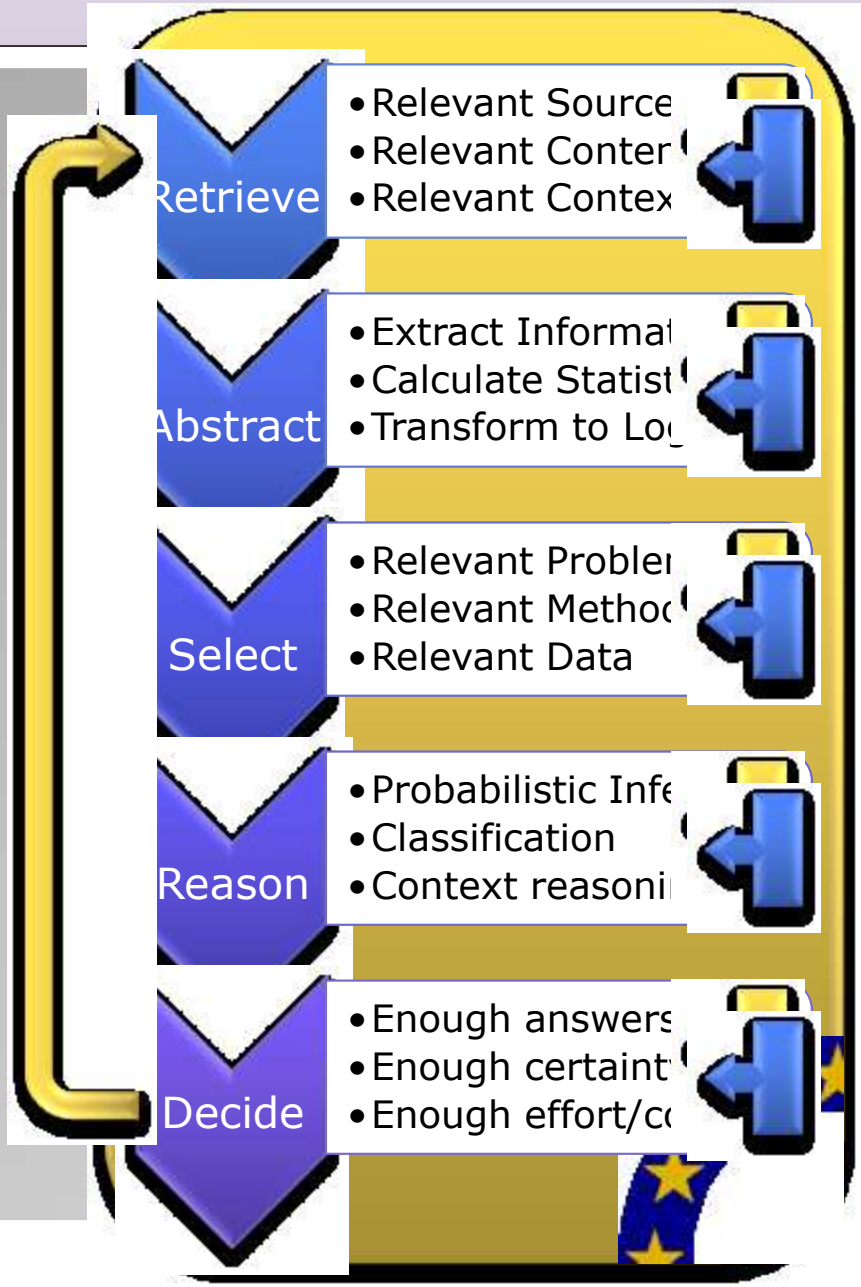
Probabilistic interpretations of segmented regions are constrained by probabilistic mereology, topology and other ontological constraints



Experimentation: Hypotheses about the world are tested by knowledgeable Web 3.0 in simulated environments with other Web 3.0 agents and people.

Where might this lead…


LARKC


Retrieve
•Relevant Source
•Relevant Conter
•Relevant Contex
Abstract
•Extract Informat
•Calculate Statist
•Transform to Lo
Select
•Relevant Proble
•Relevant Metho
•Relevant Data
Reason
•Probabilistic Infe
•Classification
•Context reasoni
Decide
•Enough answers
•Enough certaint
•Enough effort/c

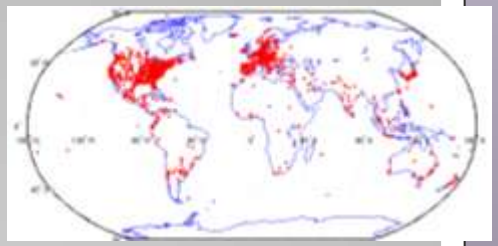
Infinite scalability?

parallelisation

- cluster and high-performance computing

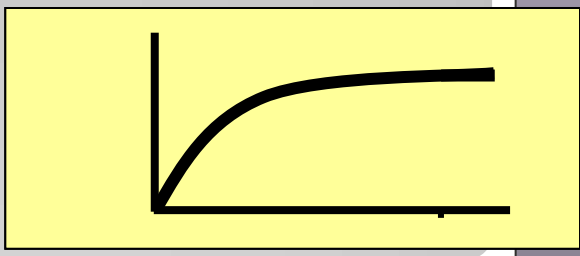
distribution

- "Thinking@home", "self-computing semantic Web"

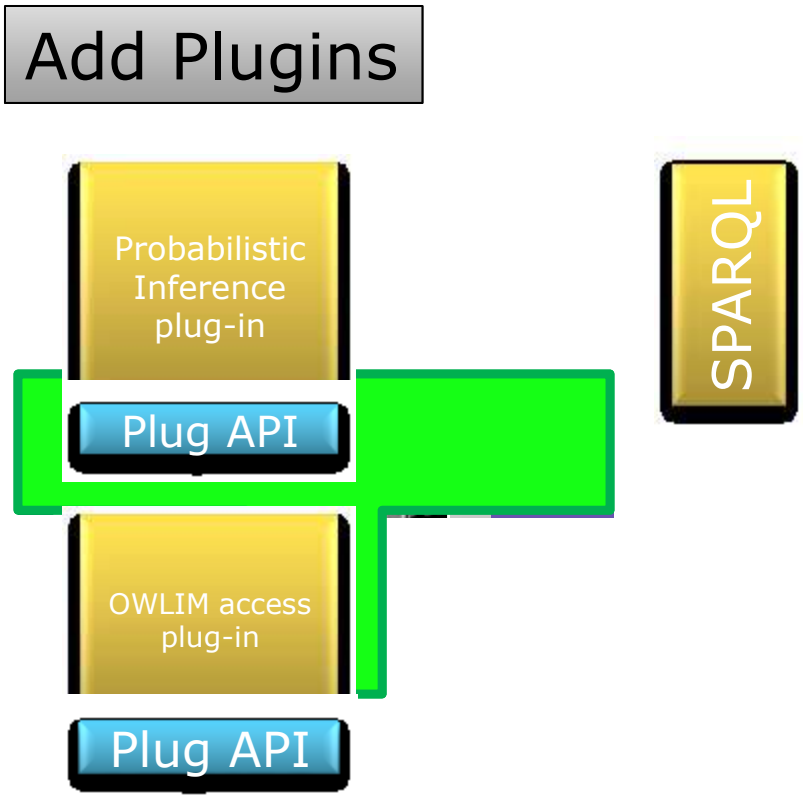


approximation

- gets better with more resources
- "almost" is often good enough



Initial LarKC platform



Description of plug-ins for the platform:

used by meta-reasoner and tactical memory

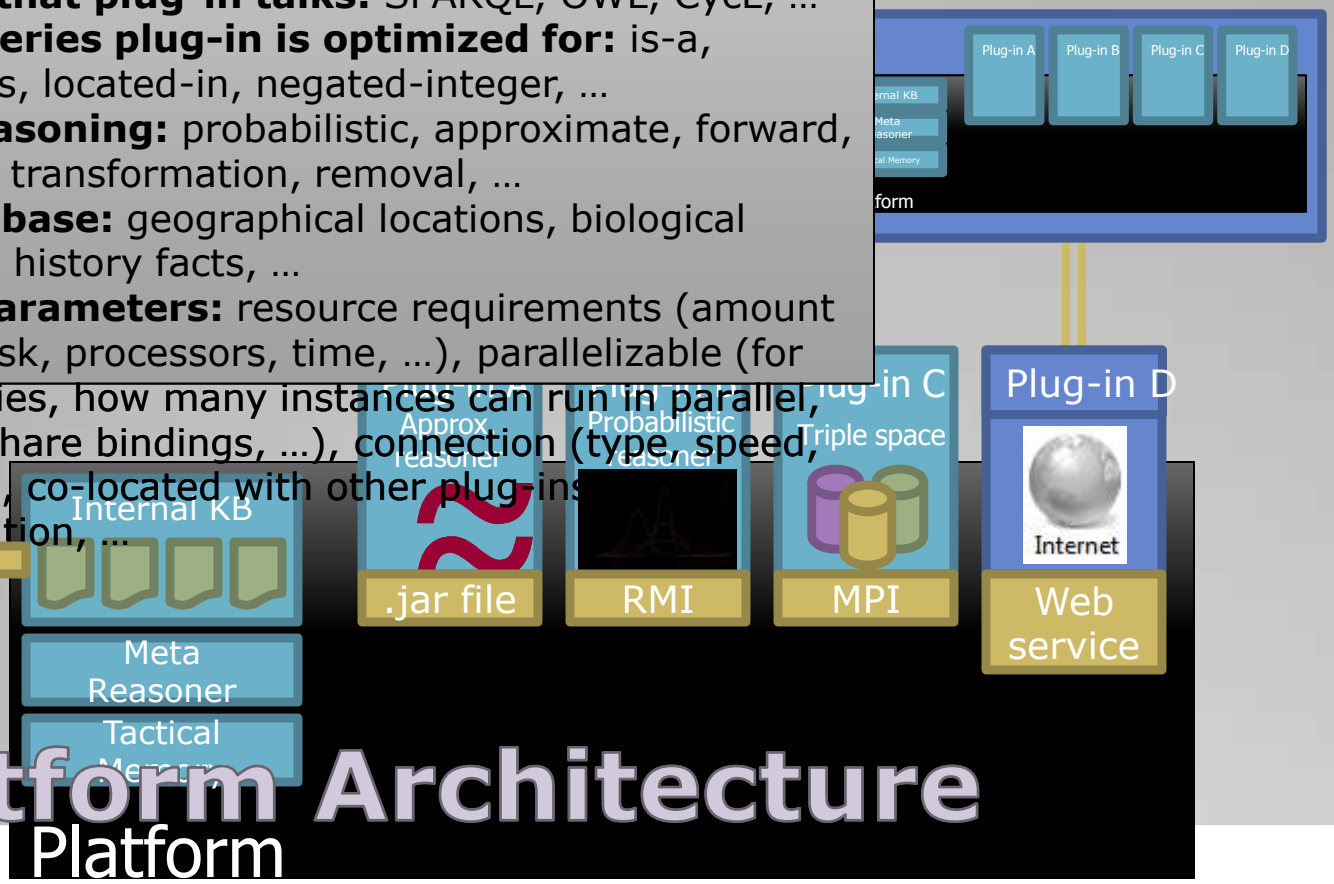
Languages that plug-in talks: SPARQL, OWL, CycL, …

Types of queries plug-in is optimized for: is-a, generalizes, located-in, negated-integer, …

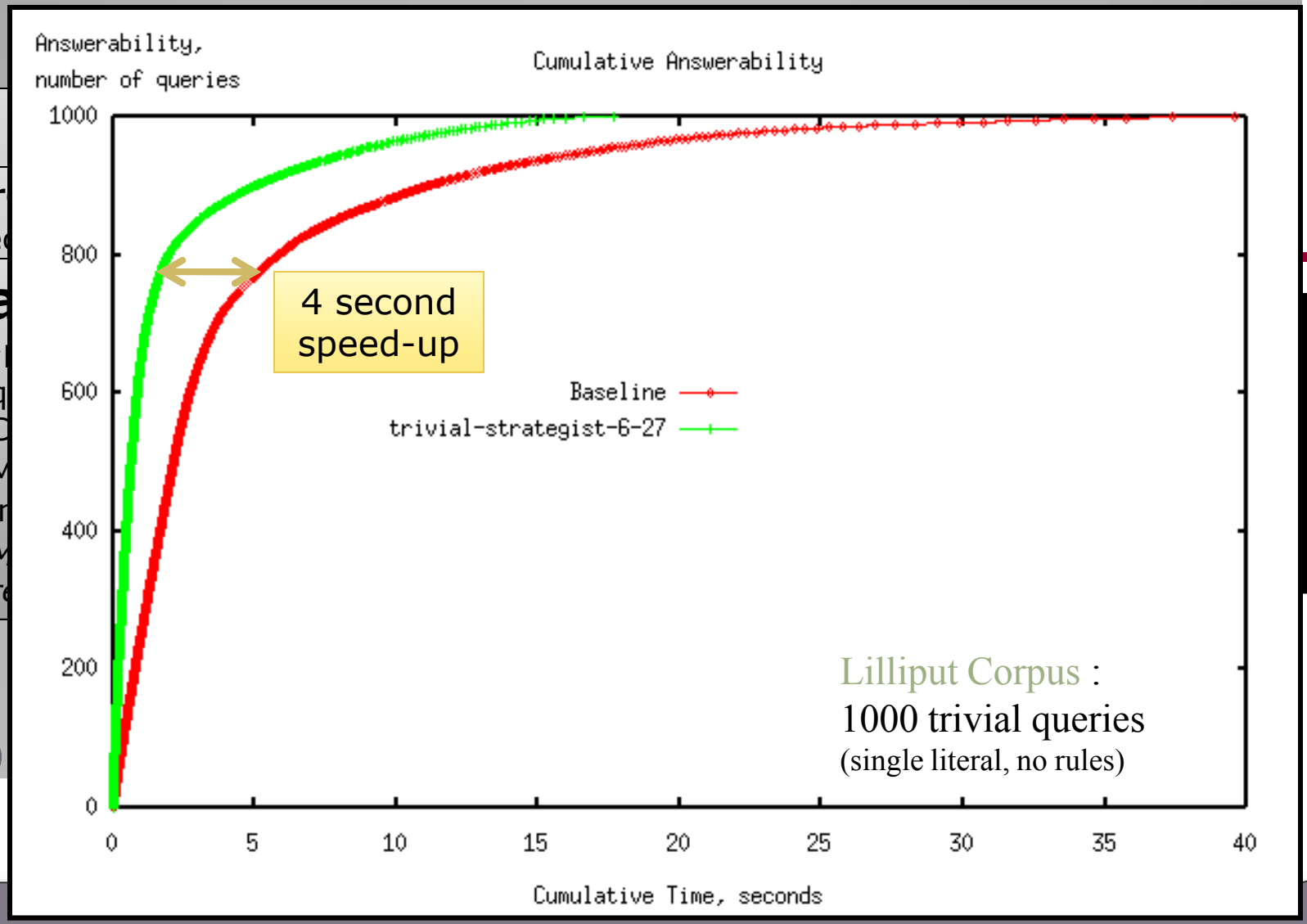
Types of reasoning: probabilistic, approximate, forward, backward, transformation, removal, …

Knowledge base: geographical locations, biological taxonomy, history facts, …

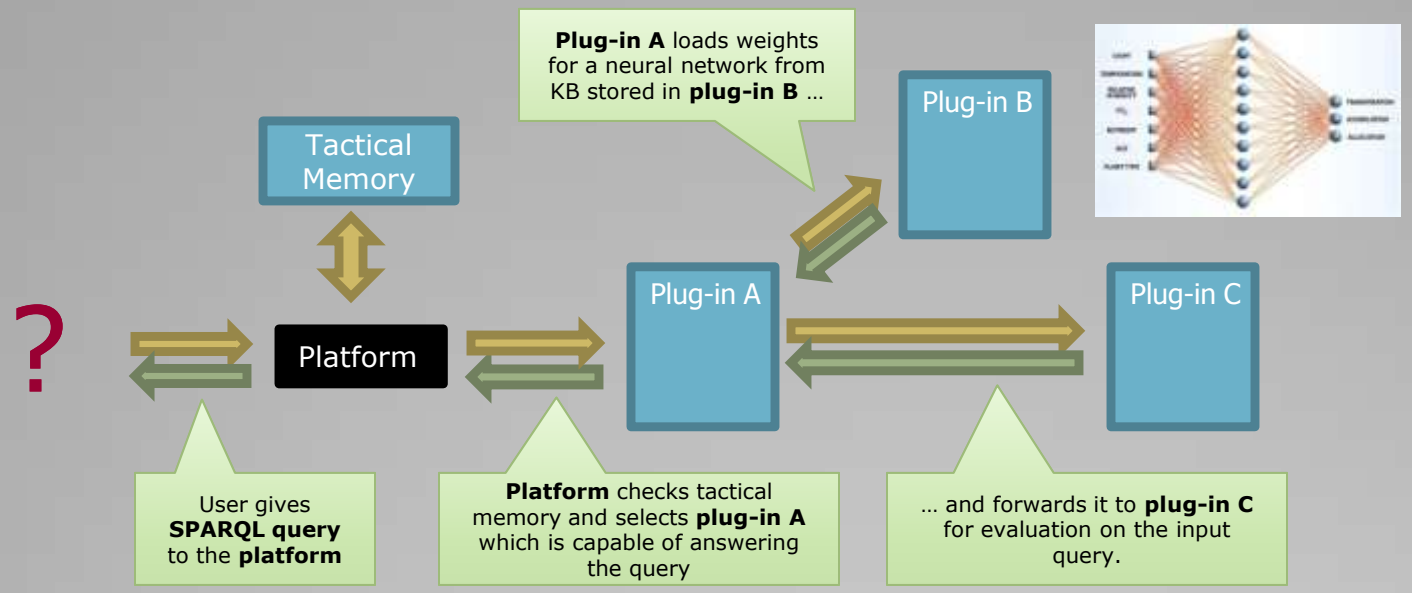
Technical parameters: resource requirements (amount of RAM, disk, processors, time, …), parallelizable (for what queries, how many instances can run in parallel, can they share bindings, …), connection (type, speed, bandwidth, co-located with other plug-ins, authentication, …

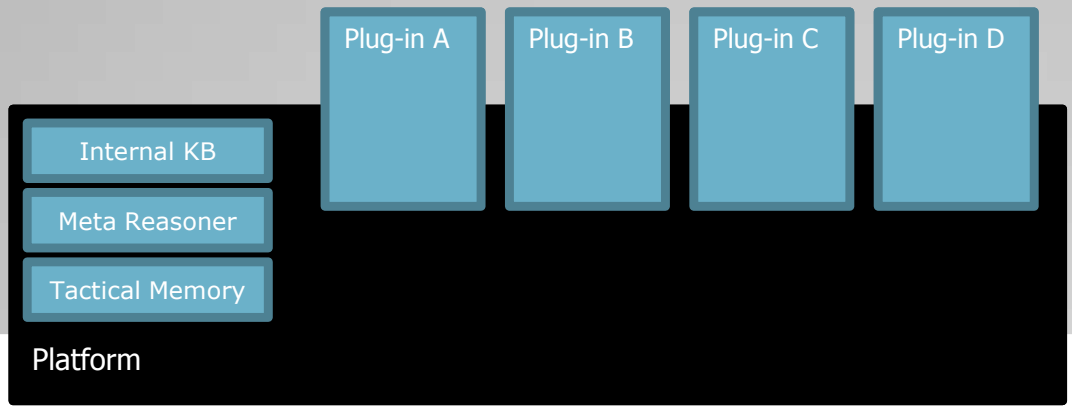


Platform Architecture


Answerability,
number of queries
Cumulative Answerability
1000
800
600
400
200
0
0
5
10
15
20
25
30
35
40
Cumulative Time, seconds
4 second
speed-up
Baseline
trivial-strategist-6-27
Lilliput Corpus :
1000 trivial queries
(single literal, no rules)

Inference driven dynamic plug-in composition



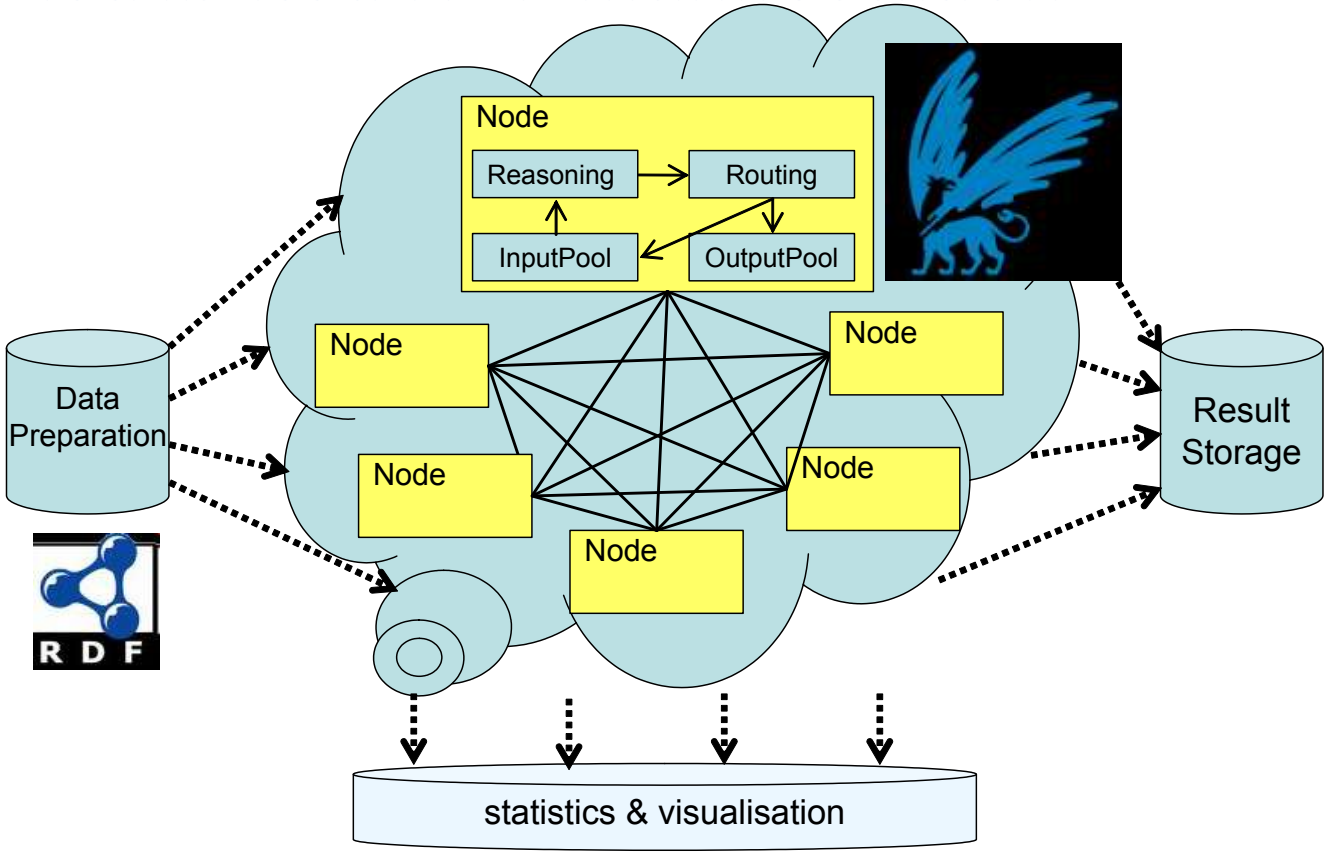


Other potential plug-ins

- Leverage other larger-scale reasoners
  - (within their domain of applicability)
- General purpose inference :
  - Vampire, DPLL, SAT solvers, LOOM, etc.
- Special purpose inference :
  - symbolic arithmetic => *Mathematica*
  - linear algebra => *Matlab, LAPack*
  - machine learning => *SVM, Neural Networks*, *Reinforcement learning*
  - planning, linear programming => *iLog, constraint solvers*
  - Humans => *mechanical turk*
  - etc.

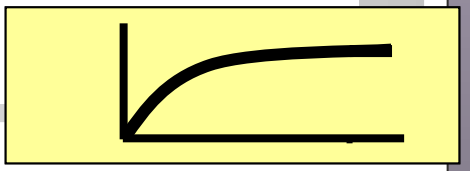
LarKC Experiment: MaRVIN

Distributed RDF manifestation



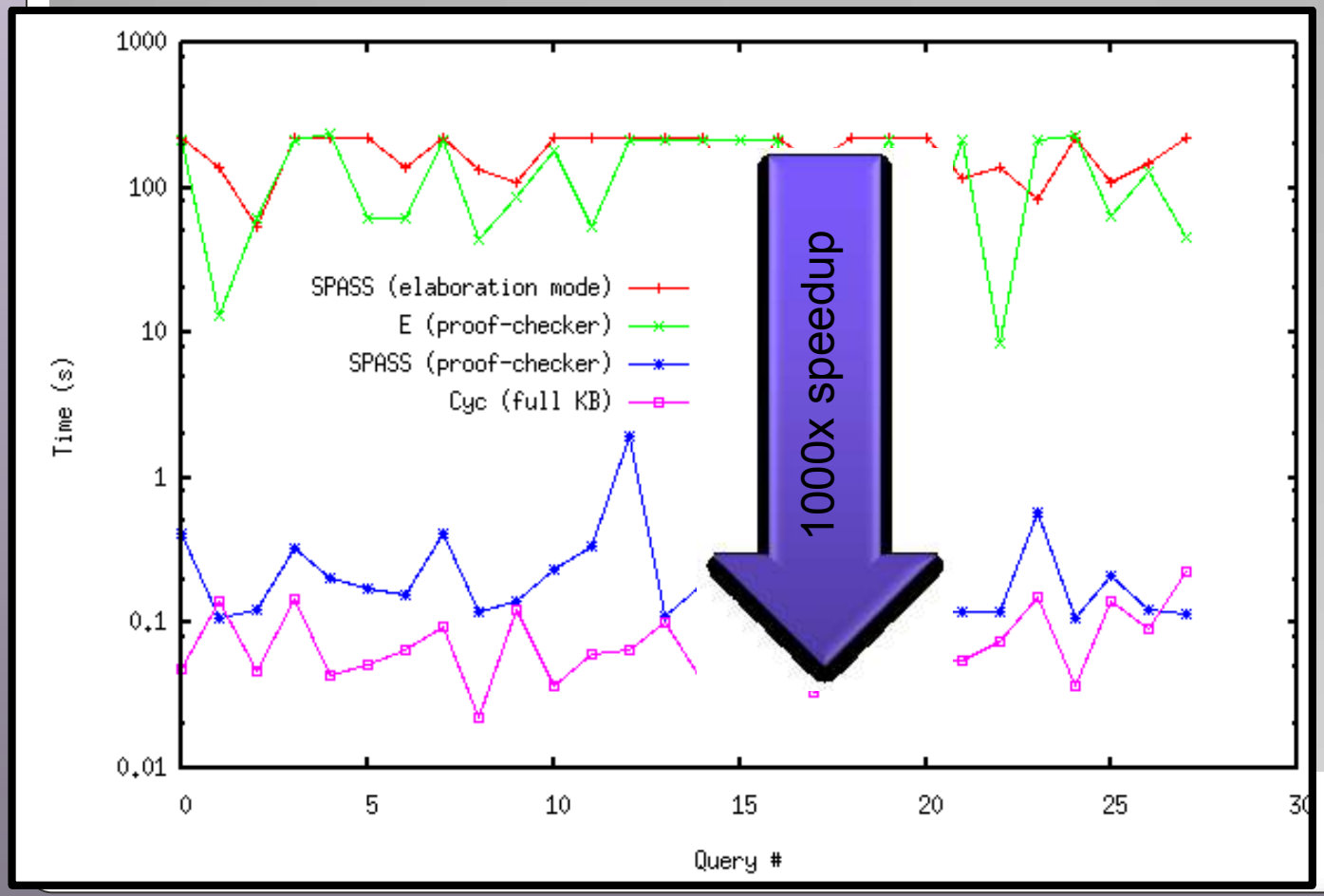
MaRVIN scales by:

- distribution (over many nodes)
- approximation (sound but incomplete)
- anytime convergence (more complete over time)



Plugin theory reasoning: disjointWith

e.g. `(disjointWith Doctor-Medical HumanInfant)`



Proof checker:
<100 relevant axioms

Elaboration Mode:
1600 relevant axioms

Cyc KB:
4 million axioms relevant & irrelevant

Reinforcemen

Acyclic Cyc

Microtheories should not have C

```
(#$thereExists ?OTHER-M
   (#$and
      (#$isa ?MT #$Micr
      (#$genlMt ?MT ?OT
      (#$genlMt ?OTHER-
      (#$different ?MT
   (#$unknownSentence
      (#$coGenlMts ?MT
```

Handcoded search: **4774 infere**

Learned search policy: **5 infere**

All In The Family

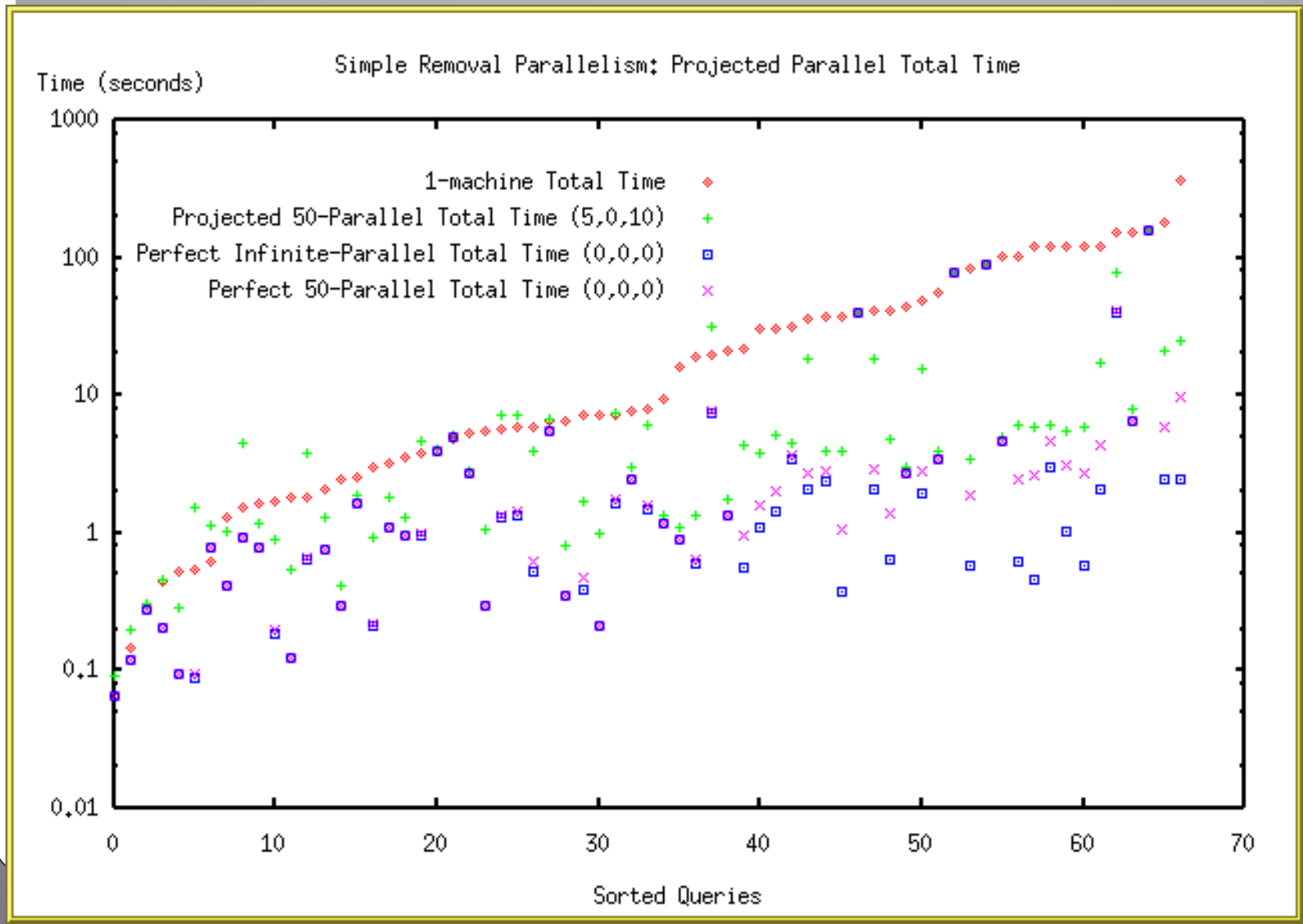
List terrorists with shared name and group affiliation or alias:

```
(#$thereExists ?WHO1-1
   (#$thereExists ?WHO2-1
      (#$and
      (#$isa ?WHO1-1 #$Terrorist)
      (#$isa ?WHO2-1 #$Agent-Generic)
      (#$familyName ?WHO1-1 ?FAMILYNAME)
      (#$familyName ?WHO2-1 ?FAMILYNAME)
      (#$givenNames ?WHO1-1 ?GIVENNAME)
      (#$givenNames ?WHO2-1 ?GIVENNAME)
      (#$different ?WHO1-1 ?WHO2-1)
      (#$extentCardinality
         (#$TheSetOf ?ALIAS-1
            (#$and
               (#$alias ?WHO1-1 ?ALIAS-1)
               (#$alias ?WHO2-1 ?ALIAS-1))) ?M)
      (#$extentCardinality
         (#$TheSetOf ?GROUP-1
            (#$and
               (#$hasMembers ?GROUP-1 ?WHO1-1)
               (#$hasMembers ?GROUP-1 ?WHO2-1))) ?N)
               (#$evaluate ?SUM
                  (#$PlusFn ?M ?N))
               (#$greaterThan ?SUM 0))))
```

Handcoded search: **14,678 inference steps**

Learned search policy: **11 inference steps**

Parallel Inference Prediction




CYC