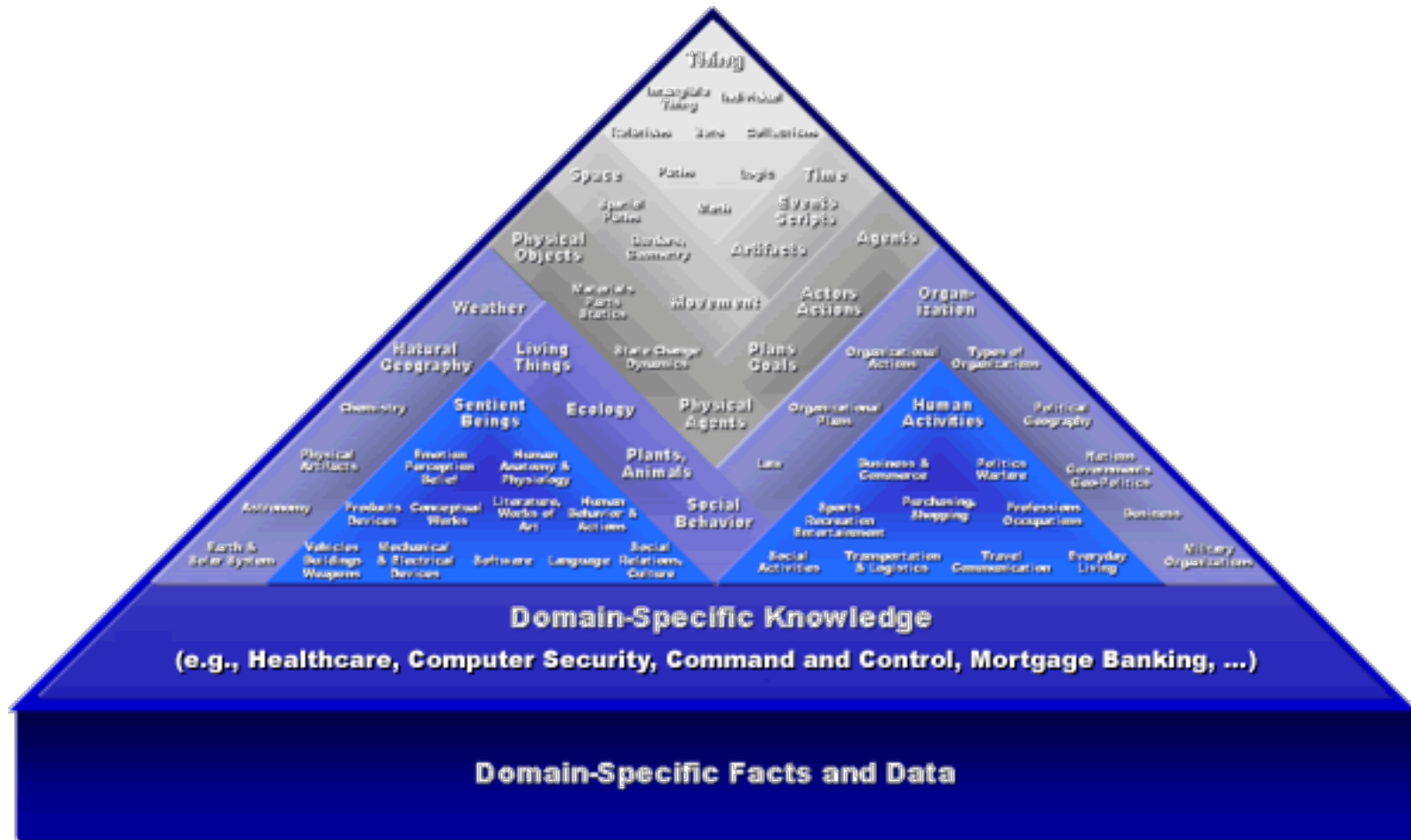


What does Cyc know?



Outline

- Overview of Cyc
- Use cases
- Tools & licensing
- Re: Soar

Big Picture

- Lobotomizing the agent for performance reasons will not lead to AI
- AI will need lots of knowledge, but no individual application will achieve ROI for developing a common sense KB
- Goals
 - AI via efficient inference over full KB
 - Keep others from reinventing this decades-old wheel



What is Cyc?

- CycL Language
- Knowledge base
 - >500K concepts
 - >5M facts
 - >1M rules
- Inference engine

CycL (> FOL)

- Constants
 - # $\$$ DougPearson
 - # $\$$ isa
 - # $\$$ Action
- Sentences
 - (# $\$$ isa # $\$$ DougPearson # $\$$ Person)
 - (# $\$$ prettyString # $\$$ Action "acts")
 - (# $\$$ implies
 - (# $\$$ knows # $\$$ Muffet # $\$$ Patches-Cat)
 - (# $\$$ likesAsFriend # $\$$ Muffet # $\$$ Patches-Cat))

CycL -> Ontology + Facts

- Processing over uniform substrate induces organization
 - Collections via generalization (`#$genls`) of `#$Collection`
 - Instances via `#$isa` relationship with any collection
 - Inheritance via logical implication (`#$implies`)
- Organization for multiple [possibly contradictory] contexts (microtheories), temporal knowledge, action sequences, lexicon, hypotheticals, beliefs, meta data, system parameters, etc.

#\$DougPearson

The screenshot shows a web browser window titled "OpenCyc Browser (ashima.local)". The address bar contains the URL "http://localhost:3602/cgi-bin/cyccgi/cg?cb-start". The page displays the profile for "DougPearson".

Navigation and User Info:

- Search bar with "person" entered.
- Buttons: Search, Clear.
- Navigation icons: Assert, Compose, Create, Doc, History, Query.
- User info: You are: [CycAdministrator](#) [Logout] Server: ashima.local:3600 Preferences Tools

Left Sidebar:

- DougPearson**
- Icons: [Create Similar] [Rename] [Merge] [Kill]
- Documentation
 - [Definitional Info](#)
 - [Internal Data](#)
 - [Assertions History](#)
- All Asserted Knowledge (6)
 - [Bookkeeping Info](#) (2)
 - [Inferred Index](#)
- All KB Assertions (4)
 - All GAFs (4)
 - Arg_1 (4)
 - isa (2)
 - [prettyString](#)
 - [prettyString-Canonical](#)

Main Content:

- Individual : [DougPearson](#)**
- Bookkeeping Assertions :**
 - [\(myCreationTime DougPearson 20040818\) in BookkeepingMt](#)
 - [\(myCreationSecond DougPearson 164533\) in BookkeepingMt](#)
- GAF Arg : 1**
- Mt : [UniversalVocabularyMt](#)**
 - isa : [Individual](#)
- Mt : [PeopleDataMt](#)**
 - isa : [Person](#)
- Mt : [EnglishMt](#)**
 - [prettyString](#) : "Pearson"
 - [prettyString-Canonical](#) : "Doug Pearson"

Copyright © 1995 - 2009 Cycorp. All rights reserved.

Footer:

- Update Comm: Storing Only Agenda: Idle KB: 5018 System: 10.128401
- Learn about [ResearchCyc](#)

#Person

The screenshot shows the OpenCyc Browser interface. The browser window title is "OpenCyc Browser (ashima.local)". The address bar shows the URL "http://localhost:3602/cgi-bin/cyccgi/cg?cb-start|cb-hier-default&c284". The search bar contains "Google". The main content area displays a hierarchical tree structure of concepts:

- TemporalThing
 - Thing
 - owl:Thing
 - Person**
 - [AllPeopleAreHumanMt -> (HomoSapiens)]
 - [BaseKB -> ((CollectionUnionFn (TheSet Person Animal)), (CollectionUnionFn (TheSet Person (FictionalFn Person))), (CollectionUnionFn (TheSet Person (GroupFn Person))))]
 - [HumanSocialLifeMt -> (LegalAgent)]
 - [UniversalVocabularyMt -> ((CollectionDifferenceFn SomethingExisting Organization), Sentient, SocialBeing)]
 - [WebSearchEnhancementMt -> (SocialBeing)]
 - (BelieverFn Abolitionism) [HumanActivitiesMt] <-
 - (BelieverFn Anarchism) [BaseKB] <-
 - (BelieverFn AnticomunistIdeology) [BaseKB] <-
 - (BelieverFn AntiSemitism) [BaseKB] <-
 - (BelieverFn Atheism) [BaseKB] ...
 - (BelieverFn Capitalism) [BaseKB] ...
 - (BelieverFn CentristPoliticalIdeology) [BaseKB] <-
 - (BelieverFn Communism) [BaseKB] <-
 - (BelieverFn Conservatism) [BaseKB] <-
 - (BelieverFn Environmentalism) [BaseKB] <-
 - (BelieverFn Fascism) [BaseKB] <-

At the bottom of the browser window, there is a status bar with the text: "Update Comm: Storing Only Agenda: Idle KB: 5018 System: 10.128401" and a link "Learn about ResearchCyc".

Cyc Inference

- SubL language
 - Superset of CLisp
- Generalized search over proof space guided by >1000 competing heuristic modules
 - Bounded by time, # answers, “steps” (node expansions), forward rule transformations, etc.
 - Iterative, continuable, justifiable

Sample Query

The screenshot displays the OpenCyc Browser interface in a web browser window. The browser's address bar shows the URL `http://localhost:3602/cgi-bin/cyccgi/cg?cb-start`. The page title is "OpenCyc Browser (ashima.local)".

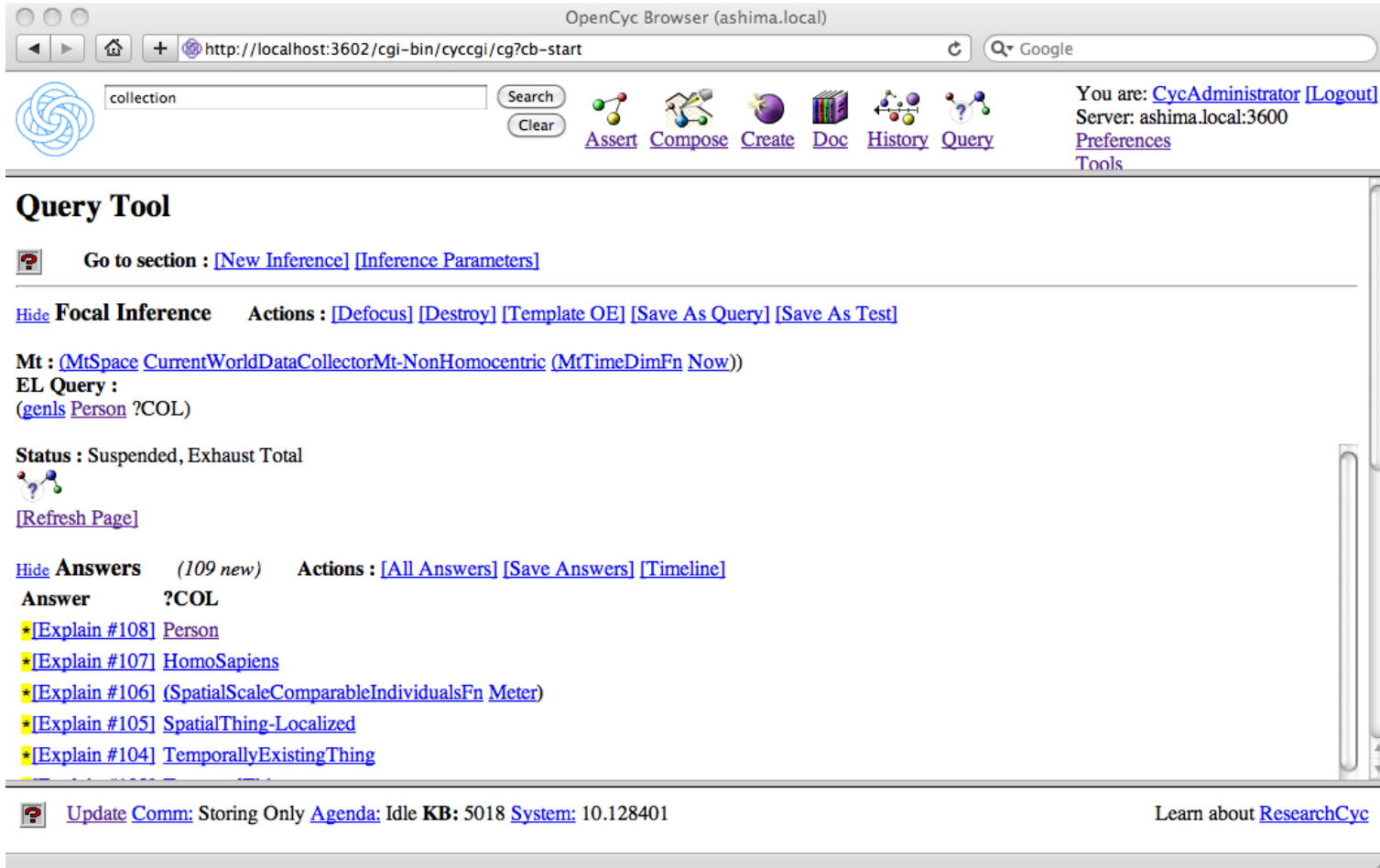
At the top, there is a search bar with the text "collection" and buttons for "Search" and "Clear". To the right of the search bar are several icons representing different Cyc functions: [Assert](#), [Compose](#), [Create](#), [Doc](#), [History](#), and [Query](#). Further right, the user is identified as [CycAdministrator](#) with a [Logout](#) link. Below this, the server information is shown: "Server: ashima.local:3600" and links for [Preferences](#) and [Tools](#).

The main content area is divided into several sections:

- New Inference:** Includes a "Reset All Fields" button and a "Start Inference" button.
- Query Context:** Shows a "Complete" button, an "EL Sentence:" field containing "(genls Person ?COL)", and a "Cyclify" button. There is also a "Clear Sentence" button.
- Pragmatics:** A section with a "Show Pragmatics" link.
- Inference Parameters:** Contains buttons for "Save Parameters", "Load Parameters", and "Reset Parameters". It also features four groups of radio buttons:
 - Number cutoff:** answers, Get everything
 - Time cutoff:** 30 seconds, Run forever
 - Step cutoff:** steps, No limit
 - Inference engine mode:** Minimal, Shallow, Extended, Maximal

At the bottom of the page, there is a status bar with a help icon, the text "Update Comm: Storing Only Agenda: Idle KB: 5018 System: 10.128401", and a link to "Learn about ResearchCyc".

Sample Results



OpenCyc Browser (ashima.local)

http://localhost:3602/cgi-bin/cyccgi/cg?cb-start

collection

[Assert](#) [Compose](#) [Create](#) [Doc](#) [History](#) [Query](#)

You are: [CycAdministrator](#) [Logout](#)
Server: ashima.local:3600
[Preferences](#)
[Tools](#)

Query Tool

Go to section : [\[New Inference\]](#) [\[Inference Parameters\]](#)

[Hide Focal Inference](#) Actions : [\[Defocus\]](#) [\[Destroy\]](#) [\[Template OE\]](#) [\[Save As Query\]](#) [\[Save As Test\]](#)

Mt : [\(MtSpace CurrentWorldDataCollectorMt-NonHomocentric \(MtTimeDimFn Now\)\)](#)
EL Query :
([gens Person](#) ?COL)

Status : Suspended, Exhaust Total

[Hide Answers](#) (109 new) Actions : [\[All Answers\]](#) [\[Save Answers\]](#) [\[Timeline\]](#)

Answer ?COL

- [\[Explain #108\]](#) [Person](#)
- [\[Explain #107\]](#) [HomoSapiens](#)
- [\[Explain #106\]](#) [\(SpatialScaleComparableIndividualsFn Meter\)](#)
- [\[Explain #105\]](#) [SpatialThing-Localized](#)
- [\[Explain #104\]](#) [TemporallyExistingThing](#)

[Update Comm:](#) Storing Only [Agenda:](#) Idle **KB:** 5018 [System:](#) 10.128401

Learn about [ResearchCyc](#)

Sample Justification



OpenCyc Browser (ashima.local)

http://localhost:3602/cgi-bin/cyccgi/cg?cb-start

collection Search

[Assert](#) [Compose](#) [Create](#) [Doc](#) [History](#) [Query](#)

You are: [CycAdministrator](#) [Logout](#)
Server: ashima.local:3600
[Preferences](#)
[Tools](#)

Inference Answer [\[Inference Answer 18.0.107\]](#) [\[Socratic Query\]](#)

Mt : [\(MtSpace CurrentWorldDataCollectorMt-NonHomocentric \(MtTimeDimFn Now\)\)](#)
EL Query :
[\(gens Person ?COL\)](#)

Creation Time : after 0.049 seconds (at 10:51:29 today)
Steps to This Answer : 111
Answer Bindings :
?COL → [HomoSapiens](#)

Justifications :
[:GENLS \(gens Person HomoSapiens\)](#) in [\(MtSpace CurrentWorldDataCollectorMt-NonHomocentric \(MtTimeDimFn Now\)\)](#)

Copyright © 1995 - 2009 [Cycorp](#). All rights reserved.

 [Update Comm:](#) Storing Only [Agenda:](#) Idle **KB:** 5018 [System:](#) 10.128401 [Learn about ResearchCyc](#)

Inference Notes (1)

- Multi-layer implementation
 - Strategist, tactician, worker
 - Combines forward/backward inference
 - Combines breadth, depth
- Multiple representations for possibly exploiting efficient computation via modules
 - Eg. Rule macros
- Parameter sweep -> “modes”
- Lazy rule matching via LEAPS
 - Integrate conflict resolution in matching
 - eliminates beta network

Inference Notes (2)

- Massive “common sense” regression test suite
- Search state and intermediate results accessible to inference
 - Learning via storage of intermediate results and heuristic module utility
- Can efficiently access external data sources
 - Multi-level data source description in KB

Interesting Use Cases

- Planning
 - Known software exploits -> intrusion plans
- Codify legal regulations -> compliance proof
- World model -> assisted NLP (NCAA)
 - Final % NOT in the input
- Terrorism
 - Plans for attacking points of interest
 - Determine likely culprits for an attack

Relevant Licensing

- OpenCyc
 - Free, no restriction on use
 - Full ontology, limited assertions
 - No NLP
- ResearchCyc
 - Free for research/development
 - All non-proprietary Cyc content/tools

Tools

- Server: java, devours RAM (~2.4GB for OpenCyc)
 - Telnet (SubL)
 - Web tool (SubL over HTTP via Tomcat)
 - Java API (SubL over TCP)
 - Synchronous and asynchronous query handling
 - Infrastructure for synchronizing KBs across peer instances
- Web services
 - Input: URL, Output: xml

Re: Soar (1)

- Source of symbols
 - SML apps provide Cyc constants on input/output
 - Supports integration with Cyc
 - Supports limited NLP input/output
- Outsource SMem
 - Efficient, scalable, lots of data
 - But how to distribute?
 - Possible to explore Cyc as an integration of EpMem/SMem (via event representation and temporal reasoning)

Re: Soar (2)

- Model for specialized processing
 - Independent, efficient modules implement/outsource specialized processing (like database queries, GIS, etc)
- Rete improvements
 - LEAPS trades potential re-computation over partial matches for improved memory and dynamic re-ordering (could help with MVAs?)
 - JIT-style dynamic rete re-computation (in a separate thread) with usage statistics

Re: Soar (3)

- Model for regression testing
 - Explicit representation of inputs/outputs
 - Lots of problems, data sources

Concerns

- How can an agent learn to effectively/efficiently query Cyc?
 - Seems as hard as automating acquisition
 - How to write rules to react to unknown symbols
- Planning
 - Due to inference model, essentially need to declaratively implement an algorithm, productions/operators, etc.
 - Despite lots of “common sense,” Cyc can’t solve blocks world without lots of help
 - Helicopters might be involved in the solution...

#\$finish

