AURA

Representing Structure, Function and their Relationship

Vinay K. Chaudhri
Nikhil Dinesh
Deborah Frazier

PROJECT HALO

Acknowledgment

- This work is funded by Vulcan Inc.
  - See http://www.vulcan.com and www.projecthalo.com

Outline

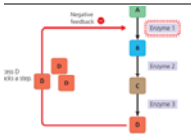
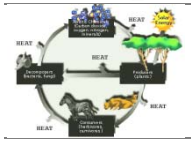
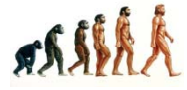
AURA

- Project Overview
- Representing Structure and Function
  - Representation and Reasoning needs
  - Upper ontology and the KR language
  - Representing structure
  - Representing Function
  - Representing structure function relationship
  - Answering questions
- Some open research problems

PROJECT HALO

Core Themes in Biology

AURA

	Challenge
 Structure and Function	Relating structure and function
 Regulation	Qualitative reasoning about dynamic processes
 Energy Transfer	Representing energy production, consumption
 Continuity and Change	Representing genetic change across generations
 Evolution	Models of population dynamics
 Science as a Process	Experimentation and hypothesis testing
 Interdependence in Nature	Represent large inter-related complex systems
 Science, Technology, Society	Represent technological and social forces

PROJECT HALO

Core Theme: Structure & Function

AURA

Structure and function are correlated at all levels of biological organization: *The form fits the function*

(a) A bird's wings have an aerodynamically efficient shape.

(b) Wing bones have a honeycombed internal structure that is strong but lightweight.

PROJECT HALO

Computational Meaning of a Core Theme

AURA

- Identify the requirements in terms of a set of questions
  - Diagnostic questions
    - Help assess the basics of KR&R
  - Educationally useful questions
    - The question must be of interest to teachers and students
    - The question must be ``Google hard”
    - The question should not require solving an open-ended research problem

PROJECT HALO

Diagnostic S&F Questions

AURA

- What is the structure of X?
- What is the function of X?

PROJECT HALO

Educationally Useful Questions

AURA

- Relate Structures to Functions
  - What structure of Biomembrane facilitates a function of biomembrane, namely phagocytosis?
- Qualitative Comparisons
  - If the Loop of Henle gets longer, how will its function be impacted?
- Detailed Comparisons
  - What is the functional similarity between prions and viroids?
- Similarity Reasoning
  - Glucose is to Glycogen as ATP is to what?
- Negatively Modified Structures Impacting Functions
  - If hydrogen is removed from a saturated fatty acid, then how is its function impacted?

PROJECT HALO

Starting Point(s)

AURA

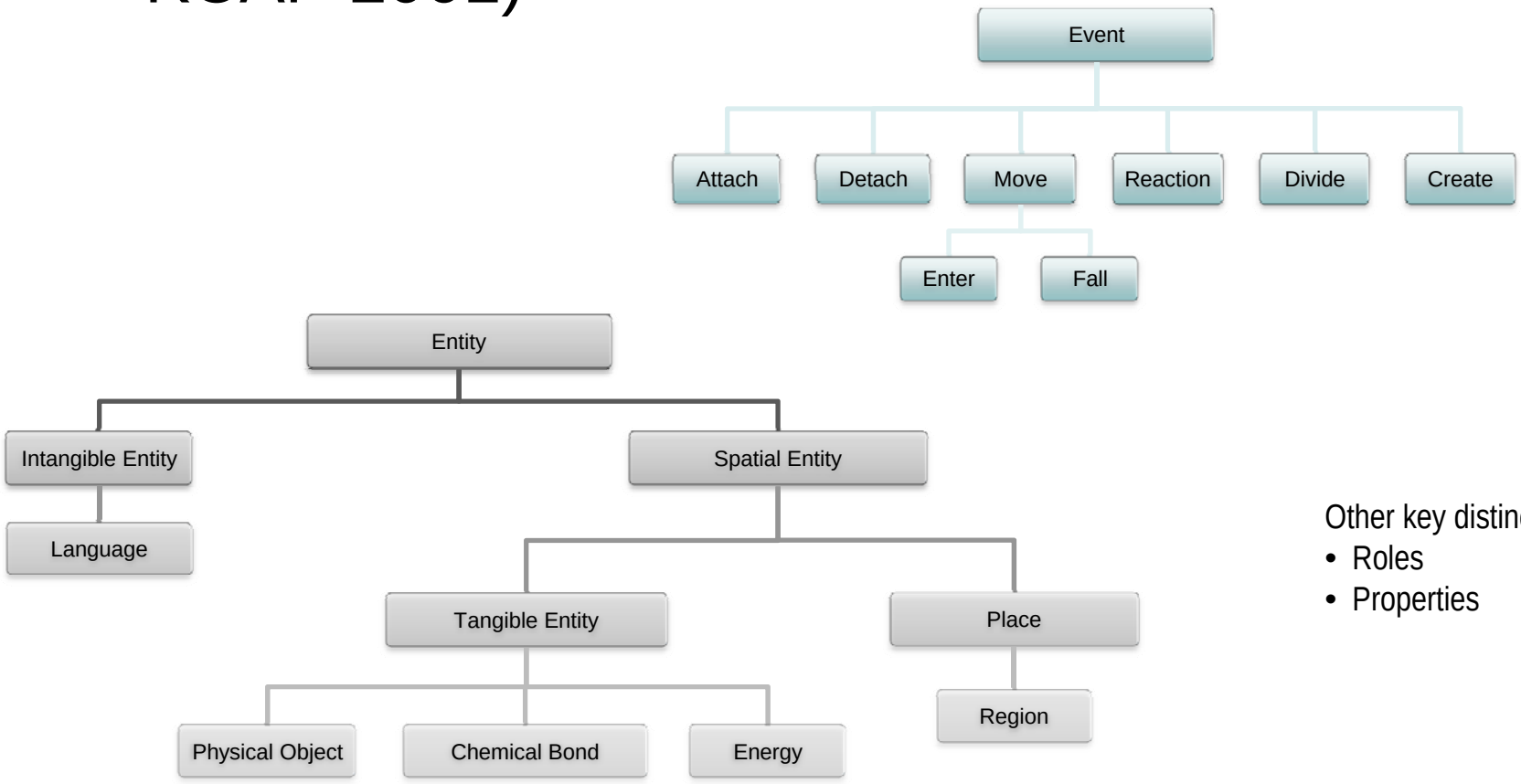
- Component Library
  - An upper ontology of generic events and entities
- Prototypes: A KR language for making universally true statements
  - A tool accessible to biologists to build conceptual models

PROJECT HALO

Component Library

AURA

- A simple upper ontology designed to be accessible to domain experts (Barker et. al, KCAP 2001)



Other key distinctions:
- Roles
- Properties

PROJECT HALO

Component Library

AURA

- A vocabulary of relations to describe events

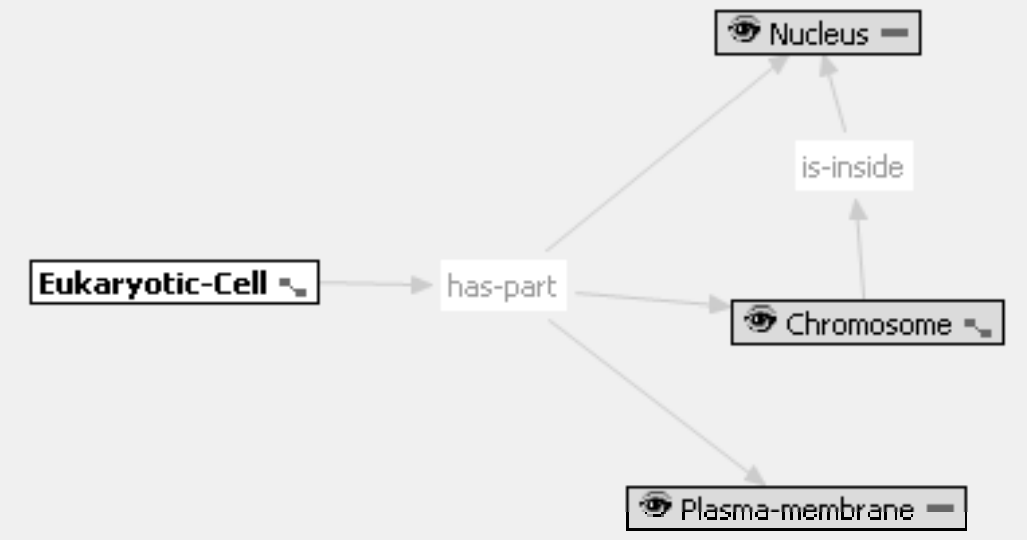
Event to Entity	Event to Event	Event to Value
agent	first-subevent	direction
object	next-event	distance
instrument	causes	duration
raw-material	enables	frequency
result	prevents	intensity
site	inhibits	rate
origin	by-means-of	

PROJECT HALO

Prototypes

AURA

- A prototype is an axiom that follows the forall/exists quantification pattern (Clark, et. al, KCAP 2001)




```
(forall ?c
    (=> (instance-of ?c Eucaryotic-Cell)
        (exists ?x ?y ?z
        (and
        (instance-of ?x Nucleus)
        (instance-of ?y Chromosome)
        (instance-of ?z Plasma-Membrane)
        (has-part ?c ?x) (has-part ?c ?y)
        (has-part ?c ?z) (is-inside ?y ?x)))))
```

Prototypes are meant to be used by biologists
Ontologists can use full first order logic

Cannot be expressed in OWL
Require at least RIF-BLD

PROJECT HALO

Representing Structure

AURA

- Structure of an entity represents its parts, their spatial arrangements and sizes

Meronymic	Spatial	Properties
has-part	is-at	length
has-region	is-inside	diameter
material	is-outside	height
possesses	abuts	area
element	is-between	depth
	is-along	volume

PROJECT HALO

Choosing Structural Slots

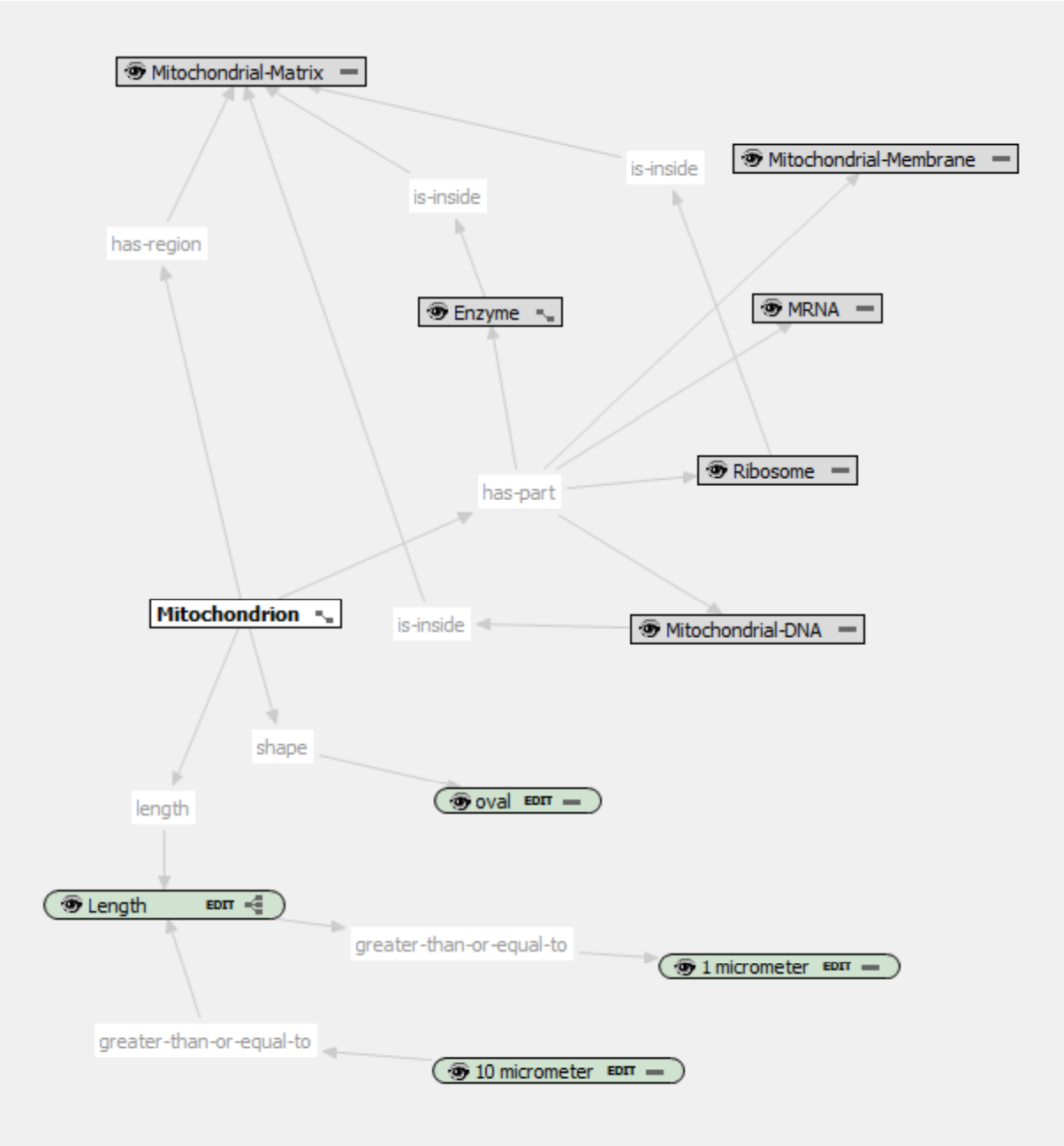
AURA

- Inspired by work of Maria Keet, but simplified for use by biologists:
  - It must make sense to say ``X has Y" in English
  - X has-region Y if
    - Y is a region of space defined in relation to X
    - It does not make sense to associate Y with properties such as mass or density, but can be associated with measures such as length, area, or volume
  - X has material Y only if
    - Y is tangible and pervasive in X
  - X has element Y if
    - X is a set of entities of the same type (or sibling types) that Y is an instance of
  - X possesses Y only if
    - Y is Energy, bond or gradient
  - Otherwise X has part Y

PROJECT HALO

Example Structure Representation

AURA

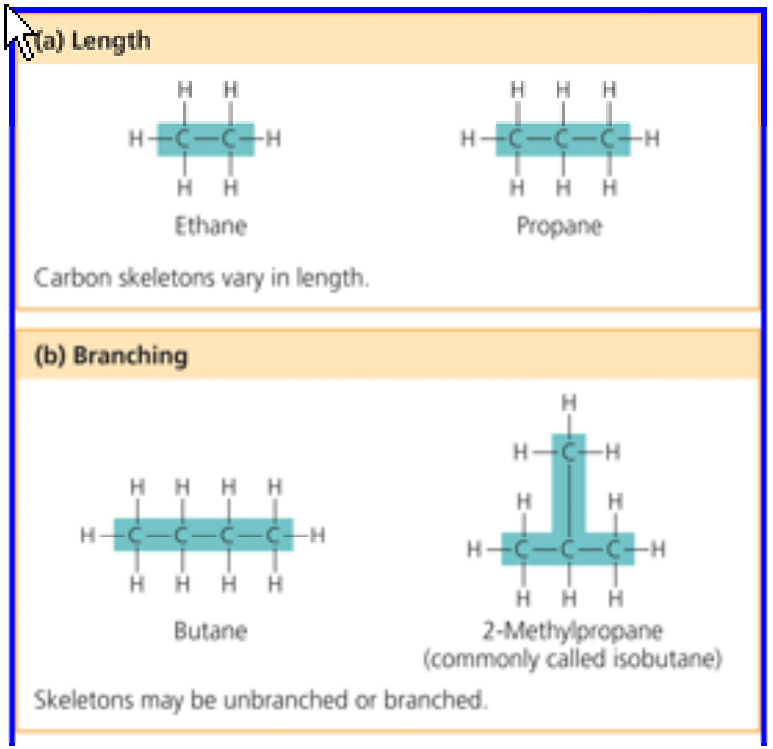


PROJECT HALO

A difficult example: Carbon Skeleton

AURA

What should be the relationship between an organic molecule and a skeleton?



It is more than simply a set of entities

-Can have length and shape

Is not an entity in its own right

- Biologists do not associate mass with it

The remaining choice is has-region

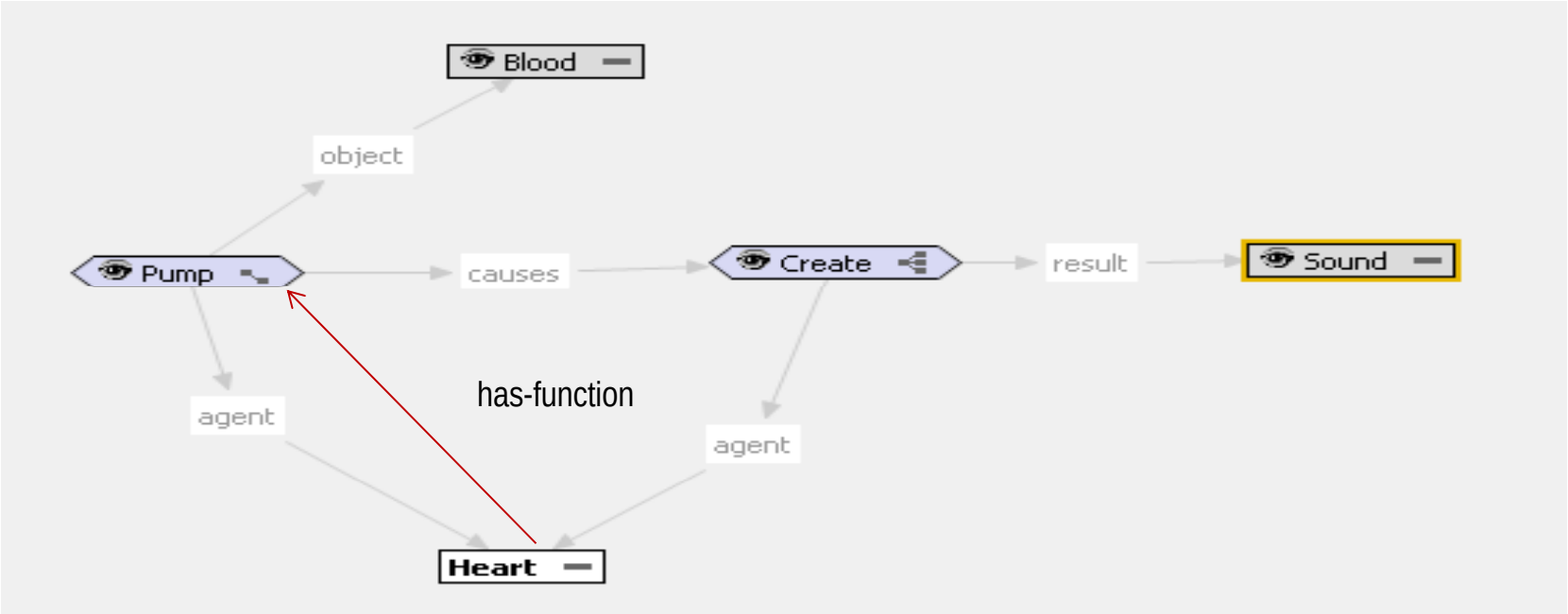
-behaves differently than a human skeleton

PROJECT HALO

Representing Functions

AURA

- Is function a primitive or a computed notion?



- It is a primitive notion and should be encoded by a biologist

PROJECT HALO

Representing Functions

AURA

- What is a function?
  - We understand functions as “special” events in which an entity participates
  - Alternatively, a function is an event which is a reason for an entity’s existence
  - The “special” nature of functions will be indicated by using a new slot called **has-function**

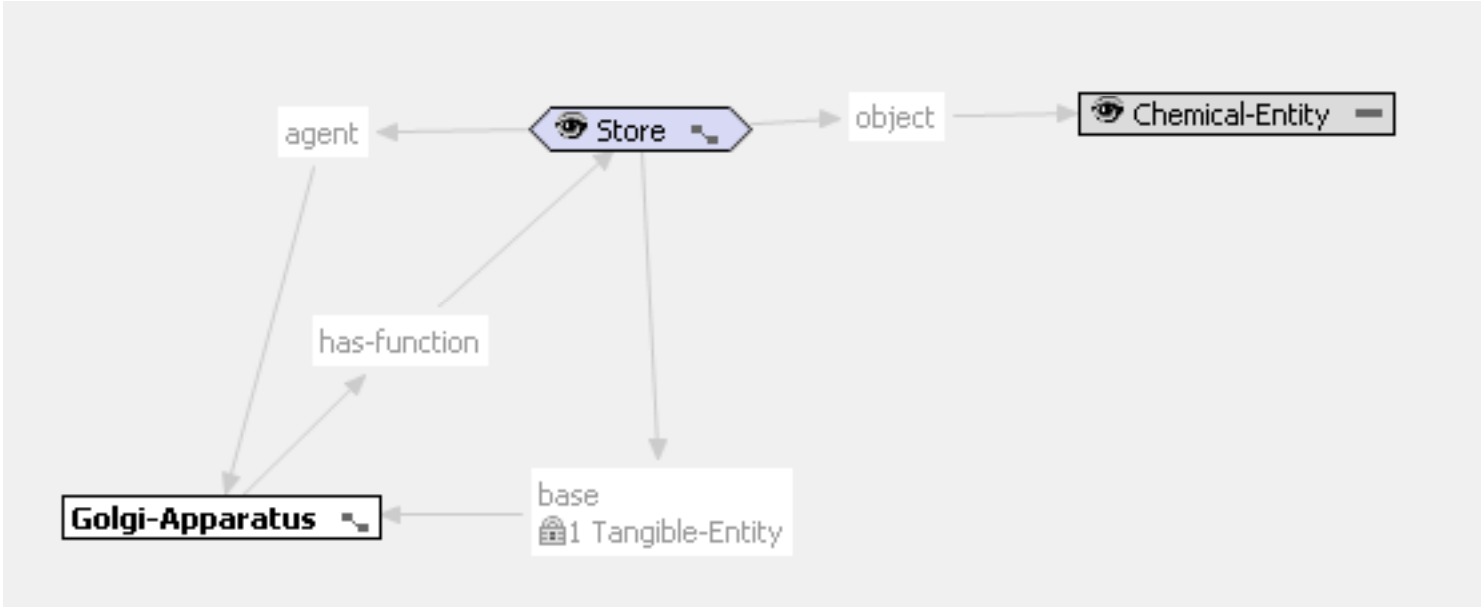
- Types of functions
  - Inherent functions of an entity
    - These will appear on the entity’s concept graph
  - Contextual functions of an entity
    - These will appear on *another* entity or event’s concept graph

PROJECT HALO

Example of an Inherent Function

AURA

- An inherent function of a Golgi Apparatus is to store chemicals
  - This is true regardless of which specific type of cell it is a part of
  - Inherent functions are placed on the Entity Cmap, using the has-function slot

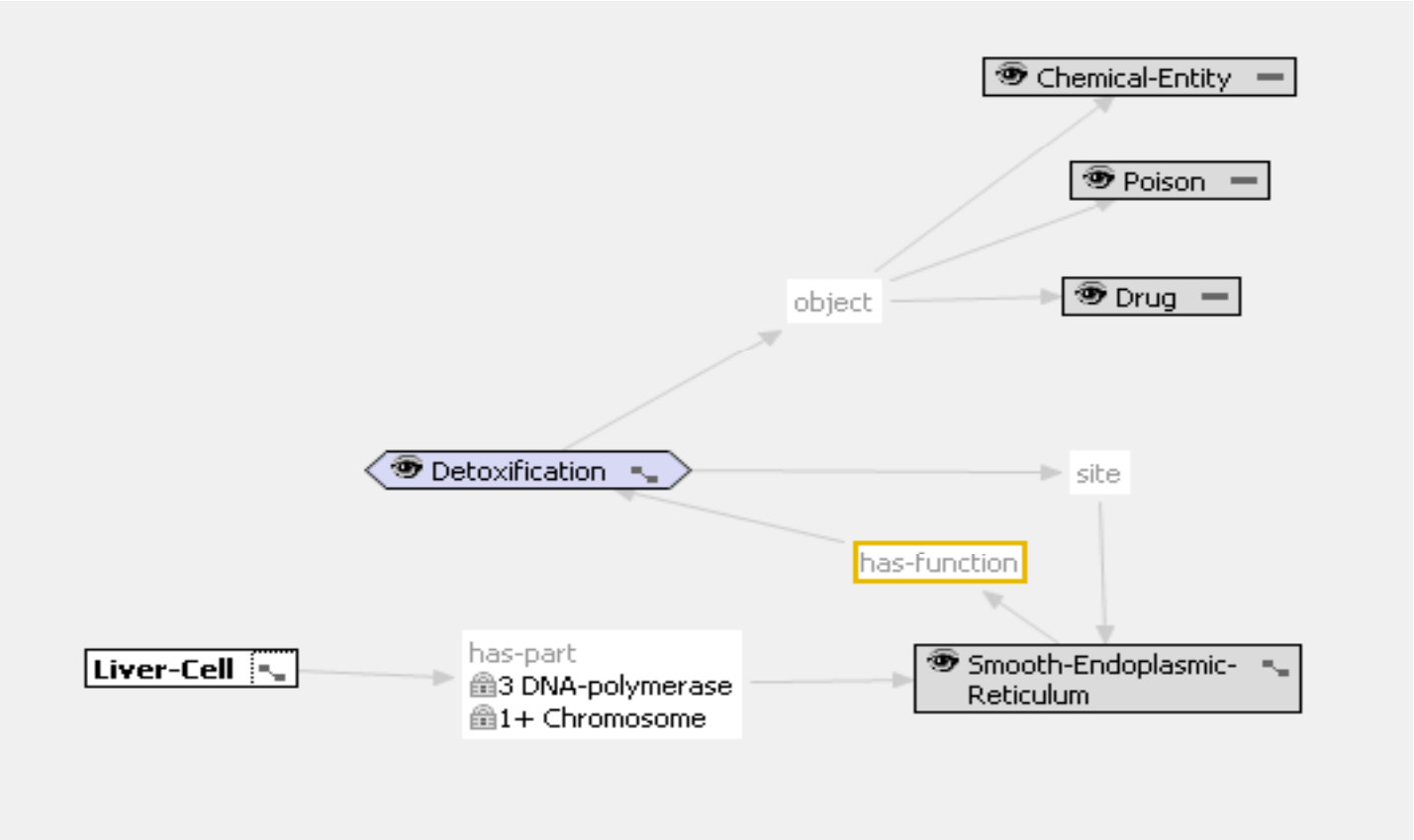


PROJECT HALO

Example of Function in an Environment

AURA

- Not every smooth ER detoxifies drugs
- However, drug detoxification is the function of a smooth ER in a liver cell

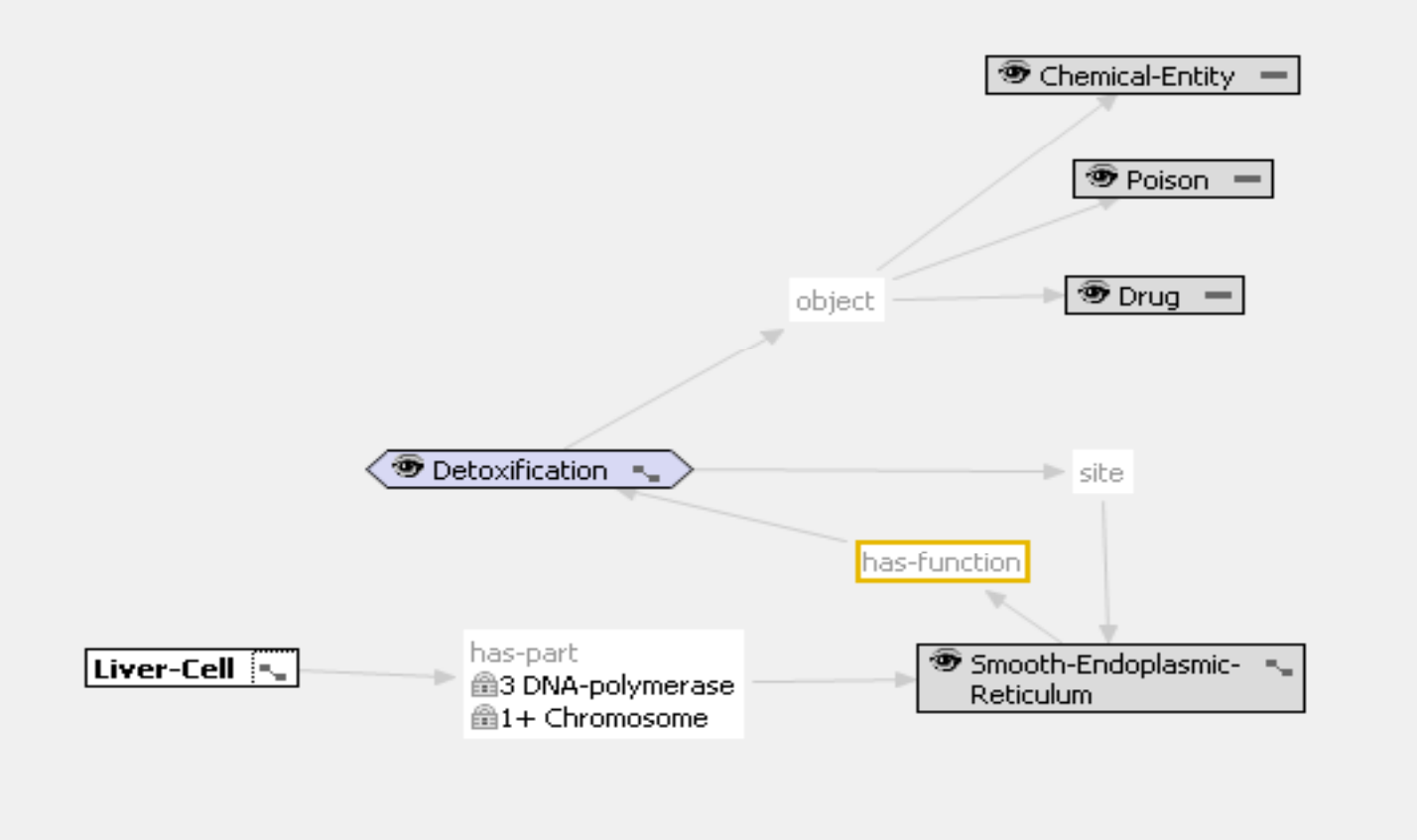


PROJECT HALO

Structure Function Relationship

AURA

- We know how an entity participates in a function

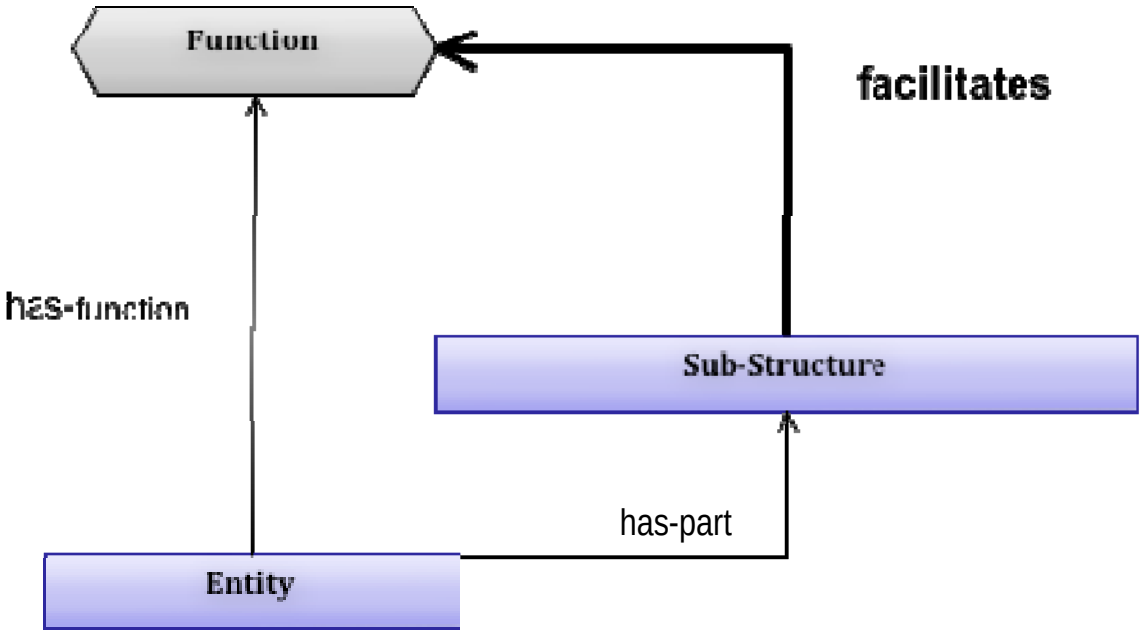


PROJECT HALO

Structure Function Relationship

AURA

- We do not know how an entity participates in a function



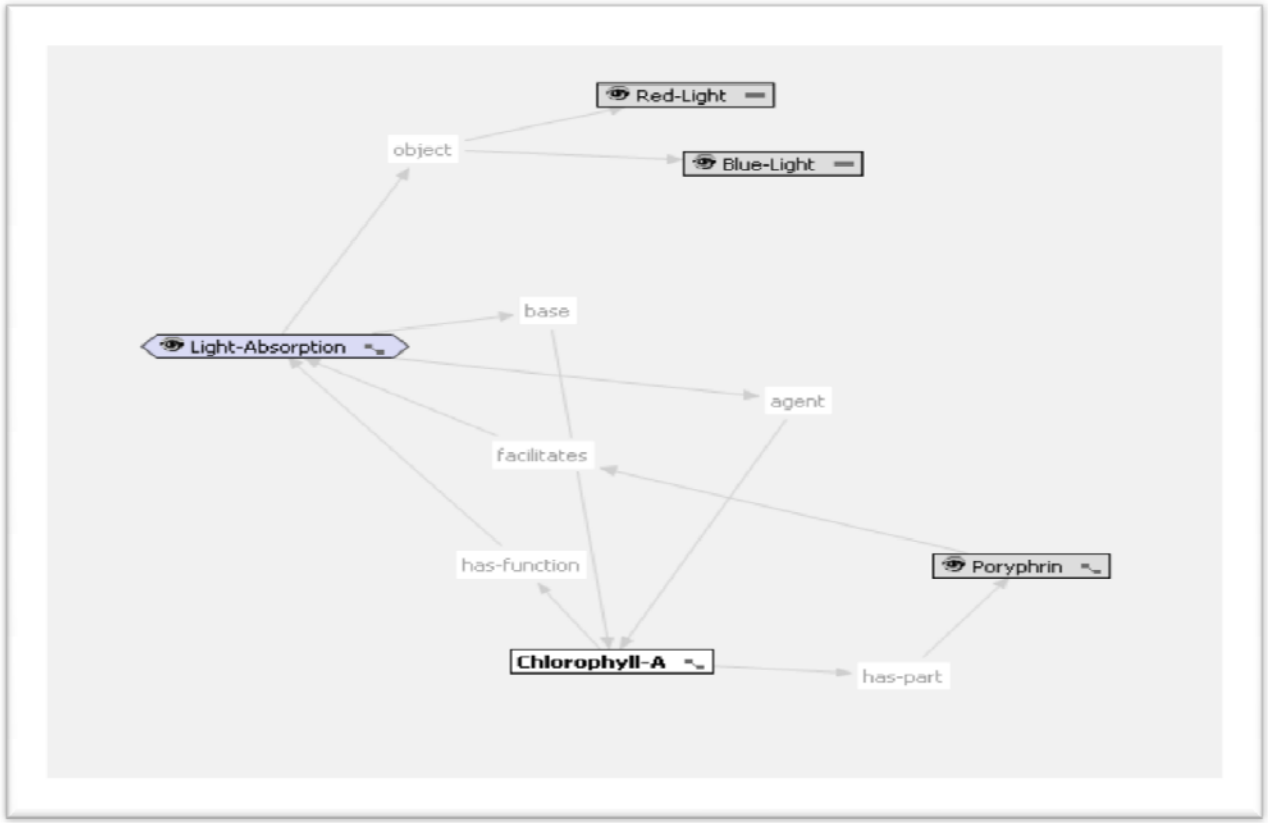
For example, Chlorophyll-A contains the functional group CH3. The text says that CH3 facilitates Chlorophyll-A's function of absorbing violet-blue light, but does not say how.

PROJECT HALO

Structure Function Relationship

AURA

- We do not know how an entity participates in a function



PROJECT HALO

Answering Questions

AURA

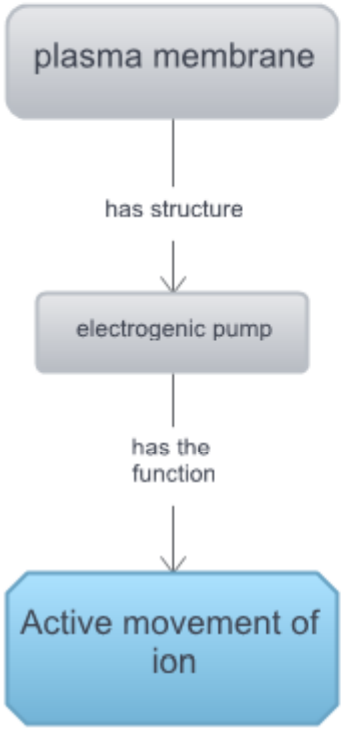
- Create an ABOX
  - Instantiate every concept in the knowledge base and compute the individuals it is related to up to depth three
- Path finding
  - Find all possible paths between two individuals

PROJECT HALO

Relate Structure to Function

AURA

What structures of a plasma membrane facilitate a function of the plasma membrane, namely active movement of ions?



The membrane at the boundary of every cell that acts as a selective barrier, regulating the cell's chemical composition.

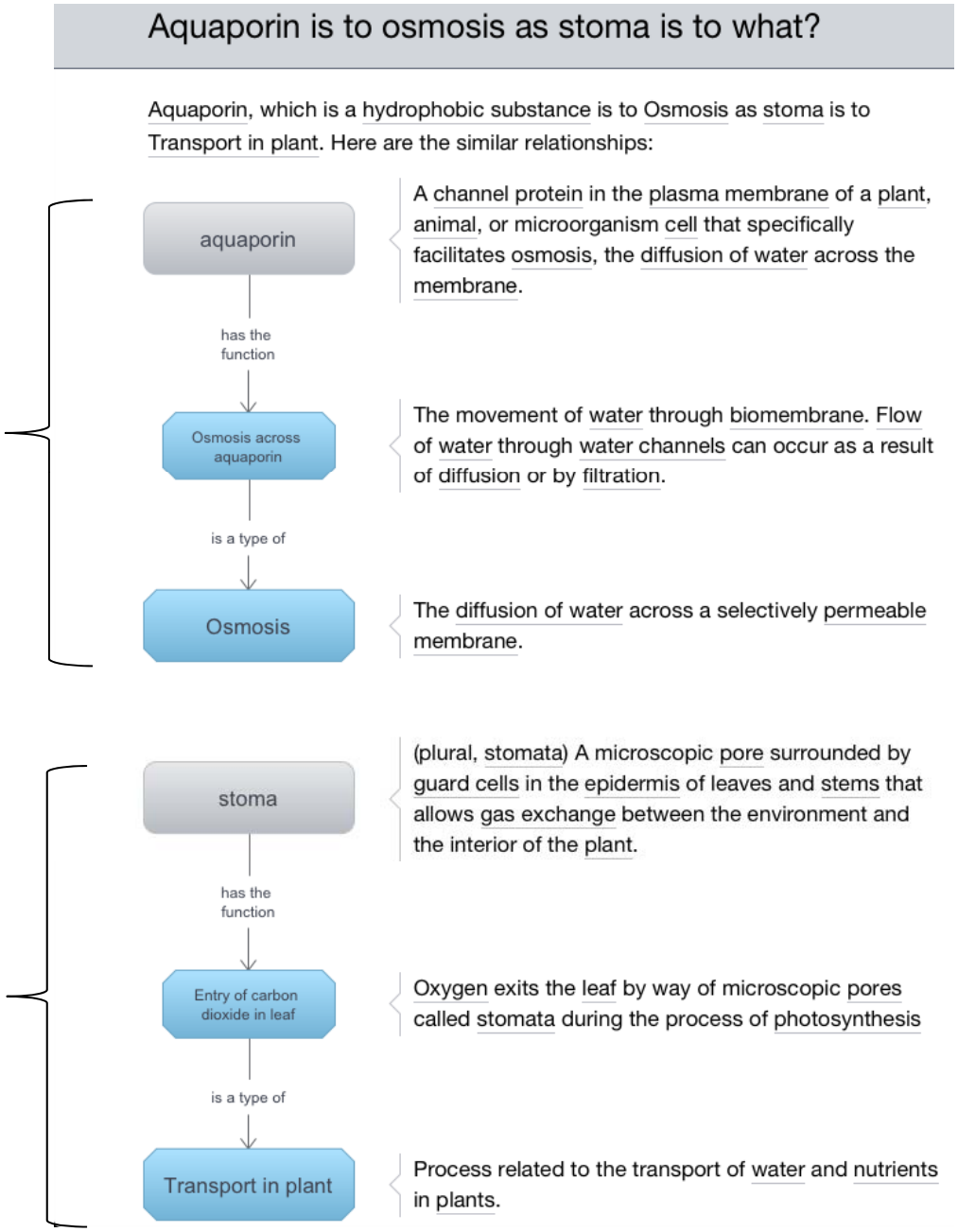
An ion transport protein that generates voltage across a membrane.

The proteins mediate the movement of ions to move across a membrane.

PROJECT HALO

Path-Based Similarity Reasoning

AURA



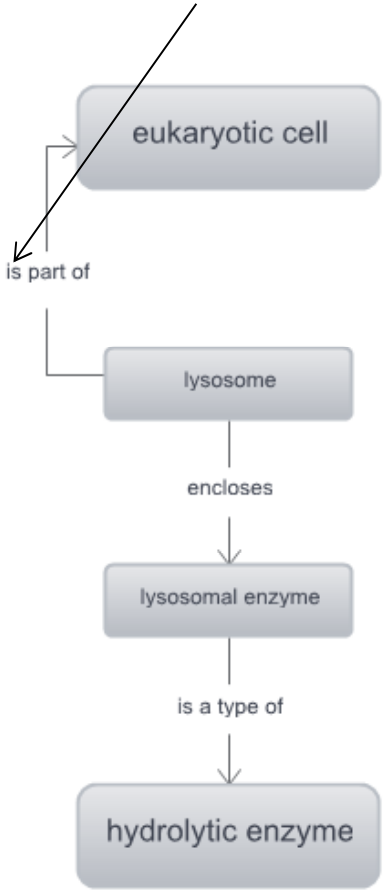
PROJECT HALO

Structural and Functional Relationships

AURA

What is the structural relationship between hydrolytic enzymes and eukaryotic cells?

Only structural slots



A type of cell with a membrane-enclosed nucleus and membrane-enclosed organelles. Organisms with eukaryotic cells (protists, plants, fungi, and animals) are called eukaryotes.

A membrane-enclosed sac of hydrolytic enzymes found in the cytoplasm of animal cells and some protists.

Lysosomal enzymes are synthesized in the cytosol and the endoplasmic reticulum. Some important lysosomal enzymes include lipase, carbohydrase, proteases, nucleases and phosphoric acid

Hydrolytic enzymes are enzymes which are involved in break-down of biological molecules

PROJECT HALO

Structure Function Relationship

AURA

- What are some longer-term research problems?
  - Specifying the structure at multiple levels of detail and from multiple perspectives
  - Defining spatial slots for the whole book

PROJECT HALO

AURA

Deep Knowledge Representation Challenge Workshop

Search this site

- Call For Participation
- Example Topics
- Submissions
- Key Dates
- Challenges
  - Hard Sentences

Co-located with K-CAP 2011
Date: June 26, 2011
Location: Banff, Alberta, Canada
Format: Full-Day Workshop

Allen Institute (ai)² for artificial intelligence

This event is sponsored by AI²

- Welcome
- Challenge
  - Questions
- Problem Features
- Criteria for Solution Assessment
- Committees
- Important Dates
- Submission

Welcome

2nd Deep Knowledge Representation Challenge (DKRC2012)

Co-located with QR 2012

Date: July 17, 2012
Location: Playa Vista, California, USA
Format: Half-Day Workshop

IAOA

This is an IAOA Supported Event

- Specify practical modeling problems in a formal and research context
- Generate interest in experimental approach to modeling
- Provide topics for student projects/theses

PROJECT HALO

Problem Features

The challenge has a number of features whose handling in the solutions seems of key importance, including:

- Qualitative descriptions of quantities
- Qualitative distinct states of behaviour and behaviour changing landmarks
- Multiple levels of detail (e.g. main process versus decomposed sub-steps)
- Perspective (e.g. chemical reactions/molecules versus amount/flow)
- Advanced qualitative inferences, including: handling ambiguity, feedback, comparative analysis, and processes at different time-scale.
- Multiple problem-solving tasks
  - Question answering
  - Answer explanation
  - Readiness for dialogue

Question 1

At one point in the process of glycolysis, both dihydroxyacetone phosphate (DHAP) and glyceraldehyde 3-phosphate (G3P) are produced. Isomerase catalyzes the reversible conversion between these two isomers. The conversion of DHAP to G3P never reaches equilibrium and G3P is used in the next step of glycolysis. What would happen to the rate of glycolysis if DHAP were removed from the process of glycolysis as quickly as it was produced?

Answer: Glycolysis is likely to stop, or at least slow it down. The conversion of the two isomers is reversible, and the removal of DHAP will cause the reaction to shift in that direction so more G3P is converted to DHAP. If less (or no) G3P were available, the conversion of G3P into DHAP would slow down (or be unable to occur).

Prior Work

AURA

- Structure, Behavior & Function (Chandrasekran, 2000)
- Basic Foundational Ontology (Arp & Smith, 2008)
- General Formal Ontology (Herre, et. al., 2006)
- DOLCE (Borgo et. al. 2010)

PROJECT HALO

Summary

AURA

- Our goal is to capture the content of a complete biology textbook by December 2013
- We are addressing core ontology design problems in an experimental context
- First cut detailed design of structure/function completed
- Designs for more core themes to come!

PROJECT HALO