HALOBOOK and Progress Towards Digital Aristotle

David Gunning
Vulcan Inc.
daveg@vulcan.com

IAAI
August 11, 2011
WHAT IS VULCAN?

■ **Vulcan manages Paul Allen’s assets**
  - Vulcan creates a variety of world-class endeavors and high impact initiatives that change and improve the way we live, learn, do business, and experience the world.

■ **Vulcan Operating Groups**
  - Vulcan Sports and Entertainment
  - Vulcan Productions
  - Vulcan Real Estate
  - Paul G. Allen Family Foundation
  - Allen Institute for Brain Sciences
  - Vulcan Capital
  - Vulcan Technology
VIDEO CLIPS OF PAUL’S INTERESTS
PROJECT HALO TEAM

- Mark Greaves
- Peter Clark
- Benjamin Grosof
- Dave Gunning
- Rob Jasper
- Wil Smith
- Jesse Wang
- Tommy Lu
- Phil Harrison (not shown)
PROJECT HALO CREATION

- Motivation
  - Inspired by Dickson’s *Final Encyclopedia*, HAL-9000, and the broad science fiction vision of Big AI
  - The volume of scientific knowledge has outpaced our ability to manage it
  - Need for systems to reason and answer questions, rather than simply retrieve relevant documents

- Project Formation
  - Workshop conducted in 2001
  - Digital Aristotle vision formulated
  - Project Halo created in 2002
Digital Aristotle – a reasoning system capable of answering novel questions and solving advanced problems in a broad range of scientific disciplines.

The project envisions two primary classes of application:
- A tutor capable of instructing and assessing students in those subjects
- A research assistant with broad, interdisciplinary skills to help scientists and others in their work

**Project Halo** is a staged, long-range research effort by Vulcan Inc. towards the development of a Digital Aristotle.
WHERE TO BEGIN?

- **AI Challenge**
  - Read a chapter and answer questions in the back of the book (Reddy, 2003)

- **Focus on Science**
  - Focus on science where knowledge is explicitly stated

- **Advanced Placement Exam**
  - Use this widely accepted test of human competence as a metric

- **Build an Integrated System**
  - Include the knowledge, knowledge acquisition, reasoning, and question answering
<table>
<thead>
<tr>
<th>Authors</th>
<th>Pilot</th>
<th>Phase II</th>
<th>HaloBook</th>
<th>DA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single Domain Expert</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logic Queries</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
HALOBOOK

Electronic Textbook

iPad application

AURA Knowledge Base

Web service
INQUIRE VIDEO DEMO
THE HALO PROJECT TODAY – THREE EFFORTS IN SYSTEMS AI

- HaloBook
  - Scaling-up AURA question answering to an entire textbook

- SMW+
  - Crowdsourcing knowledge acquisition and use via Semantic Wikis

- SILK
  - Deeper knowledge representation in a scalable inference system
SEMANTIC MEDIA WIKI (SMW)

**MediaWiki**
- Powerful Wiki engine
- Basic CMS feature set

**Semantic MediaWiki**
- Semantic Wiki engine
- Authoring explicit knowledge within content
- Basic reasoning capabilities

**Halo Extension**
- User-centric extension to Semantic MediaWiki
- Increases Consensus
- Increases exploitation of semantics

**SMW+**
- Shrink wrap suite of open source software products
- Comes with ready to use ontology
- Easy to procure and install
- Standard support contract available
SMW+ TEAM

Ontoprise

Free University Berlin
SILK

- Advanced KR language and system
  - Defaults and robust conflict handling
  - Higher-order and flexible meta-reasoning
  - Actions, events, and process models
- Radically extends KR power of current systems
  - W3C RIF, OWL, SPARQL, SQL
  - Web scalability with strong semantics
- Integrated system
  - Includes reasoner, UI with KB editor, and APIs
  - Knowledge interchange (including AURA and Cyc)
SILK TEAM

SUNY Stony Brook

Cycorp

BBN

XSBers

Paul Haley
# PROJECT HALO DEVELOPMENT STRATEGY

<table>
<thead>
<tr>
<th>Authors</th>
<th>Pilot</th>
<th>Phase II</th>
<th>HaloBook</th>
<th>DA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single Domain Expert</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Uses</th>
<th>Logic Queries</th>
<th>AP Question Answering</th>
<th>AP QA General QA Education</th>
<th>AP QA General QA Education Research</th>
</tr>
</thead>
</table>
AURA

CONCEPT OF OPERATIONS

Domain experts enter knowledge

Users ask questions

Knowledge Base
SRI AURA TEAM
AURA

Automated User-Centered Reasoning and Acquisition System (AURA)
**Fig 7-1.** A prokaryotic cell. Lacking a true nucleus and the other membrane-enclosed organelles of the eukaryotic cell, the prokaryotic cell is much simpler in structure. Only organisms of the domains Bacteria and Archaea have prokaryotic cells.

**Fig 7-7.** Overview of an animal cell. This drawing of an animal cell incorporates the most common structures of animal cells (no cell actually looks just like this). The cell has a variety of organelles ("little organs"), many of which are bounded by membranes. The most prominent organelle in an animal cell is usually the nucleus. Most of the cell's metabolic activities occur in the cytoplasm, the entire region between the nucleus and the plasma membrane. The cytoplasm contains many organelles suspended in a semifluid medium, the cytosol. Permeating much of the cytoplasm is a labyrinth of membranes called the endoplasmic reticulum (ER).
AURA ARCHITECTURE

- **KM Knowledge Base**
  - Component Library (CLIB) ontology
  - Domain specific knowledge

- **Inference**
  - General logical inference
  - Unification Mapping (UMAP) composition
  - Specialized solvers and answering methods

- **User Interaction**
  - Document viewer
  - Concept map editor

- **Question Answering**
  - Controlled English (CPL)
AURA’S BALANCING ACT

Ease of Use ⟷ Expressiveness

- Medium-sized CLIB ontology
  - 800 concepts
  - 4,000 axioms
- Fixed set of CLIB relations
  - 200 relations
- Prototypes + Composition
CLIB ONTOLOGY

- Simple Organization
  - Entities
  - Events
- Medium-sized
  - 800 concepts
  - 4,000 axioms
- Fixed Relations
  - 200 fixed relations available to authors
PROTOTYPES

- Concept Map (user view)

- Logical Axioms (machine view)

\[
∀x \text{ isa}(x, \text{Eukaryotic-cell}) \rightarrow \\
∃p,n,d \text{ isa}(p, \text{Plasma-membrane}) \land \\
\text{isa}(n, \text{Nucleus}) \land \text{isa}(d, \text{DNA}) \land \text{has-part}(x, p) \land \\
\text{has-part}(x, n) \land \text{has-part}(x, d) \land \text{is-inside}(d, n)
\]
UMAP COMPOSITION

Eukaryotic Cell

Plant Cell

Has-part:
- Plasma membrane
- Nucleus
- DNA

Has-part:
- Plasma membrane
- Cell wall
- Chloroplast

Has-part:
- Plasma membrane
- Cell wall
- Chloroplast
- Nucleus
- DNA
A boulder is dropped.
The initial speed of the boulder is 0 m/s.
The duration of the drop is 23 seconds.
The acceleration of the drop is 7.9 m/s^2.
What is the distance of the drop?
19) E19. A train has a constant velocity of 180 km/h (50 m/s). It starts to break with a constant acceleration 500 m before the next station where it stops. Calculate the necessary acceleration.

CPL:

A train is the object of a move. The speed of the move is 12 m/s. The duration of the move is 12 s. What is the distance of the move?

Enter CPL: [help][test]

A train is the object of a move.
The speed of the move is 12 m/s.
The duration of the move is 12 s.
What is the distance of the move?

---

**Diagram:**

- **Move**
  - **Distance**
  - **12 second**
  - **12 meter-per-second**

- **Train**
19) E19. A train has a constant velocity of 180km/h (50m/s). It starts to break with a constant acceleration 500m before the next station where it stops. Calculate the necessary acceleration.

CPL:
- A train is the object of a move. The speed of the move is 12 m/s. The duration of the move is 12 s. What is the distance of the move?

Enter CPL: [help][test]
A train is the object of a move.
The move has a speed of 12 m/s.
The move has a duration of 12 s.
The move has a distance of unknown.
What is the distance?

**Answer**

The distance of the move = 144 m

**Explanation**

\[ \text{motion-with-constant-velocity} : \]

Given the values
- \( v = 12 \text{ m/s} \)
- \( t = 12 \text{ s} \)

Given the equations
- \( t = t_f - t_i \)
- \( x = p_f - p_i \)
- \( v_{\text{avg}} = v \)
- \( v = x / t \)

Solved equation \( v = x / t \) for \( x \)
giving \( x = v \times t \)

Evaluated \( x = v \times t \) giving \( x = 144 \text{ m} \)
- \( x = 144 \text{ m} \)
AURA TEST RESULTS (2008)  
QUESTION ANSWERING PERFORMANCE

<table>
<thead>
<tr>
<th></th>
<th>Biology</th>
<th>Chemistry</th>
<th>Physics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reference</td>
<td>71</td>
<td>73</td>
<td>79</td>
</tr>
<tr>
<td>Selected</td>
<td>76</td>
<td>44</td>
<td>70</td>
</tr>
<tr>
<td>All Novel</td>
<td>47</td>
<td>18</td>
<td>36</td>
</tr>
<tr>
<td>Novel AP</td>
<td>33</td>
<td>9</td>
<td>36</td>
</tr>
</tbody>
</table>
## PROJECT HALO DEVELOPMENT STRATEGY

<table>
<thead>
<tr>
<th>Authors</th>
<th>Pilot</th>
<th>Phase II</th>
<th>HaloBook</th>
<th>DA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single Domain Expert</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uses</td>
<td>Logic Queries</td>
<td>AP Question Answering</td>
<td>AP QA General QA Education</td>
<td>AP QA General QA Education Research</td>
</tr>
</tbody>
</table>
HaloBook is a concept for a new kind of electronic textbook—
that contains an underlying knowledge-base of the book’s contents—
that answers the reader’s questions and provides tailored instruction.

This briefing presents a sketch of how a HaloBook might look in 2015—
assuming continued Halo development and modest success in solving research challenges—
An initial HaloBook prototype will be developed in 2010.
A HaloBook user can:

- **READ** the contents dynamically, interactively
- **ASK** questions and get explanations on any subject in the book
- **LEARN** to master the subject through individualized tutoring
- **CREATE** and explore their own conceptualizations of the material
We are developing a scalable knowledge entry process using a hybrid team of SME and KR experts.
A cell without rigid walls can tolerate neither excessive uptake nor excessive loss of water. This problem of water balance is automatically solved if such a cell lives in isotonic surroundings. Seawater is isotonic to many marine invertebrates. The cells of most terrestrial (land-dwelling) animals are bathed in an extracellular fluid that is isotonic to the cells. In hypertonic or hypotonic environments, however, organisms that lack rigid cell walls must have other adaptations for osmoregulation, the control of solute concentrations and water balance. For example, the unicellular protist Paramecium caudatum lives in pond water, which is hypotonic to the cell. The ciliate has a plasma membrane that is much less permeable to water than the membranes of most other cells, but this only slows the uptake of water, which continually enters the cell. The P. caudatum cell doesn’t burst because it is also equipped with a contractile vacuole, an organule that functions as a bilge pump to force water out of the cell as fast as it enters by osmosis (Figure 7.16a). We will examine other evolutionary adaptations for osmoregulation in Chapter 44.

### Water Balance of Cells with Walls

The cells of plants, prokaryotes, fungi, and some protists are surrounded by walls (see Figure 6.28). When such a cell is immersed in a hypotonic solution—bathed in rainwater, for example—the wall helps maintain the cell’s water balance. Consider a plant cell. Like an animal cell, the plant cell swells as water enters by osmosis (Figure 7.16b). However, the relatively inelastic wall will expand only so much before it exerts a back pressure on the cell, called turgor pressure, that opposes further water uptake. At this point, the cell is turgid (very firm), which is the healthy state for most plant cells. Plants that are not woody, such as most houseplants, depend for mechanical support on cells kept turgid by a surrounding hypotonic solution. If a plant’s cells and their surroundings are isotonic, there is no net tendency for water to enter, and the cells become flaccid (limp).

However, a wall is of no advantage if the cell is immersed in a hypertonic environment. In this case, a plant cell, like an animal cell, will lose water to its surroundings and shrink. As the plant cell shrinks, its plasma membrane pulls away from the wall. This phenomenon, called plasmolysis, causes the plant to wilt and can lead to plant death. The walled cells of bacteria and fungi also plasmolyze in hypertonic environments.

### Facilitated Diffusion: Passive Transport Aided by Proteins

Let’s look more closely at how water and certain hydrophilic solutes cross a membrane. As mentioned earlier, many polar molecules and ions impeded by the lipid bilayer of the membrane diffuse passively with the help of transport proteins that span the membrane. This phenomenon is called facilitated diffusion. Cell biologists are still trying to learn exactly how various transport proteins facilitate diffusion. Most transport proteins are very specific. They transport some substances but not others.

![Facilitated Diffusion Diagram](attachment:image.png)
A cell without rigid walls can tolerate neither excessive uptake nor excessive loss of water. This problem of water balance is automatically solved if such a cell lives in isotonic surroundings. Seawater is isotonic to many marine invertebrates. The cells of most terrestrial (land-dwelling) animals are bathed in an extracellular fluid that is isotonic to the cells. In hypertonic or hypotonic environments, however, organisms that lack rigid cell walls must have other adaptations for osmoregulation, the control of solute concentrations and water balance. For example, the unicellular protist Paramecium caudatum lives in pond water, which is hypotonic to the cell. P. caudatum has a plasma membrane that is much less permeable to water than the membranes of most other cells, but this only slows the uptake of water, which continually enters the cell. The P. caudatum cell doesn’t burst because it is also equipped with a contractile vacuole, an organelle that functions as a bilge pump to force water out of the cell as fast as it enters by osmosis (Figure 7.16). We will examine other evolutionary adaptations for osmoregulation in Chapter 44.

Water Balance of Cells with Walls

The cells of plants, prokaryotes, fungi, and some protists are surrounded by walls (see Figure 6.28). When such a cell is immersed in a hypotonic solution—bathed in rainwater, for example—the cell helps maintain the cell’s water balance. Consider a plant cell. Like an animal cell, the plant cell swells as water enters by osmosis (Figure 7.15b). However, the relatively inelastic wall will expand only so much before it exerts a back pressure on the cell, called turgor pressure, that opposes further water uptake. At this point, the cell is turgid (very firm), which is the healthy state for most plant cells. Plants that are not woody, such as houseplants, depend for mechanical support on cells kept turgid by a surrounding hypotonic solution. If a plant’s cells and their surroundings are isotonic, there is no net tendency for water to enter, and the cells become flaccid (limp).

However, a wall is of no advantage if the cell is immersed in a hypertonic environment. In this case, a plant cell, like an animal cell, will lose water to its surroundings and shrink. As the plant cell shrivels, its plasma membrane pulls away from the wall. This phenomenon, called plasmolysis, causes the plant to wilt and can lead to plant death. The walled cells of bacteria and fungi also plasmolyze in hypertonic environments.

Facilitated Diffusion: Passive Transport Aided by Proteins

Let’s look more closely at how water and certain hydrophilic solutes cross a membrane. As mentioned earlier, many polar molecules and ions impeded by the lipid bilayer of the membrane diffuse passively with the help of transport proteins that span the membrane. This phenomenon is called facilitated diffusion. Cell biologists are still trying to learn exactly how various transport proteins facilitate diffusion. Most transport proteins are very specific: They transport some substances but not others.

Figure 7.17 Two types of transport proteins that carry out facilitated diffusion.

In both cases, the protein can transport the solute in either direction, but the net movement is down the concentration gradient of the solute.
COVERAGE ANALYSIS
(DIFFICULT TO REPRESENT SENTENCES)

All Sentences

Not relevant (~50%)

Not relevant (~50%)

Relevant (~50%)

Relevant (~50%)

Encoded (~25%)

Encoded (~25%)

KR Issue (~25%)

KR Issue (~25%)

KR Issue

Core Ontology 62%

Process/Causality 47%

Defaults/Negation 17%

Uncertain/Approximate 10%

Other 10%
HALOBOOK 1.0

- HaloBook iPad Application
  - eBook reader (standard features)
  - Knowledge-based features
    - Glossary pages (from KB)
    - Suggested questions and answers
    - Free-form QA dialog

- AURA Knowledge Server
  - Knowledge base (20 chapters)
    - Concept taxonomy (full book)
  - Suggested question generation
  - Language understanding
  - Reasoning and QA methods
12.2 The mitotic phase alternates with interphase in the cell cycle
down into five stages: **prophase, prometaphase, metaphase, anaphase, and telophase**. Overlapping with the latter stages of mitosis, cytokinesis completes the **mitotic phase**. Figure 12.7, on the next two pages, describes these stages in an animal cell. Study this figure thoroughly before progressing to the next two sections, which examine mitosis and cytokinesis more closely.

---

**The Mitotic Spindle: A Closer Look**

Many of the events of mitosis depend on the **mitotic spindle**, which begins to form in the cytoplasm during prophase. This structure consists of fibers made of microtubules and associated proteins. While the mitotic spindle assembles, the other microtubules of the cytoskeleton partially disassemble, providing the material used to construct the spindle. The spindle microtubules elongate (polymerize) by incorporating more subunits of the protein tubulin (see Table 6.1) and shorten (depolymerize) by losing subunits.

In animal cells, the assembly of spindle microtubules starts at the
12.2 The mitotic phase alternates with interphase in the cell cycle

down into five stages: prophase, prometaphase, metaphase, anaphase, and telophase. Overlapping with the latter stages of mitosis, cytokinesis completes the mitotic phase. Figure 12.7, on the next two pages, describes these stages in an animal cell. Study this figure thoroughly before progressing to the next two sections, which examine mitosis and cytokinesis more closely.

**The Mitotic Spindle**

Many of the events of mitosis depend on a mitochondrion, which begins to form in the cytoplasm during prophase. This structure consists of fibers made of microtubules and associated proteins. While the mitotic spindle assembles, the other microtubules of the cytoskeleton partially disassemble, providing the material used to construct the spindle. The spindle microtubules elongate (polymerize) by incorporating more subunits of the protein tubulin (see Table 6.1) and shorten (depolymerize) by losing subunits.

In animal cells, the assembly of spindle microtubules starts at the...
12.2 The mitotic phase alternates with interphase in the cell cycle.

The mitotic phase is divided into five stages: prophase, prometaphase, metaphase, anaphase, and telophase. Overlapping with the latter stages of mitosis, cytokinesis completes the mitotic phase. Figure 12.7 on the next two pages, describes these stages in an animal cell. Study this figure thoroughly before progressing to the next two sections, which examine mitosis and cytokinesis more closely.

**The Mitotic Spindle**

Many of the events of mitosis depend on a mitotic spindle, which begins to form in the cytoplasm during prophase. This structure consists of fibers made of microtubules and associated proteins. While the mitotic spindle assembles, the other microtubules of the cytoskeleton partially disassemble, providing the material used to construct the spindle. The spindle microtubules elongate (polymerize) by incorporating more subunits of the protein tubulin (see Table 6.1) and shorten (depolymerize) by losing subunits.

In animal cells, the assembly of spindle microtubules starts at the...
Prophase

The first stage of mitosis, in which the chromatin condenses, the mitotic spindle begins to form, and the nucleolus disappears, but the nucleus remains intact.

Prophase is a type of mitotic phase.

Contents
- Participants, Location
- Steps
- Related questions

![Figure 12.11](Mitosis in a plant cell.)

Participants:
- cell
- chromatin
- chromosome
- activation energy
- reactant substrate
- free-energy
- mitotic spindle
- centrosome

Location:
- nucleus
- eukaryotic cell
- organelle
- cytoplasm

Steps of prophase

- Coming together of a chromatin
- A chromatin is contracted
- A chromosome is modified
- synthesis
- Growth of a mitotic spindle
- Moving of a centrosome and a centrosome

Related questions

- prophase is a step of what process?
- During prophase, what is located at a cell pole?
- What is the relationship between a prophase and a mitotic spindle?
can lead to cancer. The agent of such change can be random spontaneous mutation. However, it is likely that many cancer-causing mutations result from environmental influences, such as chemical carcinogens, X-rays and other high-energy radiation, and some viruses.

Cancer research led to the discovery of cancer-causing genes called oncogenes (from the Greek onco, tumor) in certain types of viruses (see Chapter 19). Subsequently, close counterparts of viral oncogenes were found in the genomes of humans and other animals. The normal versions of the cellular genes, called proto-oncogenes, code for proteins that stimulate normal cell growth and division.

How might a proto-oncogene—a gene that has an essential function in normal cells—become an oncogene, a cancer-causing gene? In general, an oncogene arises from a genetic change that leads to an increase either in the amount of the proto-oncogene’s protein product or in the intrinsic activity of each protein molecule. The genetic changes that convert proto-oncogenes to oncogenes fall into three main categories: movement of DNA within the genome, amplification of a proto-oncogene, and point mutations in a control element or in the proto-oncogene itself (Figure 18.23, on the next page).
18.5 Cancer results from genetic changes that affect cell cycle control

can lead to cancer. The agent of such change can be random spontaneous mutation. However, it is likely that many cancer-causing mutations result from environmental influences, such as chemical carcinogens, X-rays and other high-energy radiation, and some viruses.

Cancer research led to the discovery of cancer-causing genes called oncogenes (from the Greek onco, tumor) in certain types of viruses (see Chapter 19). Subsequently, close counterparts of viral oncogenes were found in the genomes of humans and other animals. The normal versions of the cellular genes, called proto-oncogenes, code for proteins that stimulate normal cell growth and division.

How might a proto-oncogene—a gene that has an essential function in normal cells—become an oncogene, a cancer-causing gene? In general, an oncogene arises from a genetic change that leads to an increase either in the amount of the proto-oncogene’s protein product or in the intrinsic activity of each protein molecule. The genetic changes that convert proto-oncogenes to oncogenes fall into three main categories: movement of DNA within the genome, amplification of a proto-oncogene, and point mutations in a control element or in the proto-oncogene itself (Figure 18.23, on the next page).
What is the difference between an oncogene and a proto-oncogene?

How is an oncogene created, in the context of a proto oncogene?

What happens to a proto oncogene, in the context of a oncogene?

What is a proto oncogene?

What is the structure of a proto oncogene?

What is the structure of a oncogene?

What are the types of oncogene?
What is the difference between an oncogene and a proto-oncogene?

An oncogene and a proto-oncogene are both types of genes; a proto-oncogene is also a type of a regulatory gene.

<table>
<thead>
<tr>
<th>an oncogene</th>
<th>a proto-oncogene</th>
</tr>
</thead>
<tbody>
<tr>
<td>an agent of:</td>
<td>a change</td>
</tr>
<tr>
<td>an object of:</td>
<td>—</td>
</tr>
<tr>
<td>a result of: DNA transfer gene amplification point mutation translocation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>a regulation of cell cycle mutation</td>
</tr>
</tbody>
</table>

**FIGURE 18.23** Genetic changes that can turn proto-oncogenes into oncogenes.

**Related questions**

- How is a oncogene created, in the context of a proto oncogene?
Q. Are lysosomes involved in ATP synthesis?
A. No, it is not true that a lysosome is related to synthesis of ATP.

Q. Are lysosomes involved in intracellular digestion?
A. Yes, a lysosome performs the intracellular digestion.

Q. Are lysosomes involved in lipid transport?
A. No, it is not true that a lysosome is related to transmission.

Q. Do lysosomes store carbohydrate?
A. No, it is not true that a lysosome performs the storage and a carbohydrate participates in the storage.

Q. What organelles are involved in protein synthesis?
A. The site of synthesis of protein is Ribosome.

Q. Does guanine pair with cytosine in RNA transcription?
A. Yes, the complement of a guanine is a cytosine.
FROM AP TO HALOBOOK

AP Questions
AP Questions

Questions In the Wild (QIWs)

Students’ Questions

Questions AURA can answer

FROM AP TO HALOBOOK
FROM AP TO HALOBOOK

AP Questions

Students’ Questions
Questions In the Wild (QIWs)

AURA can answer

Google can answer

Friday, August 12, 11
TARGET QUESTION PATTERNS

- **Relationships**
  - How is X related to Y?
  - How is X similar to/different from Y?
  - What is the functional similarity between X and Y?
  - What is the qualitative relationship between X and Y?

- **Structure and function**
  - How is entity X involved in function Y?
  - If entity X is damaged, what functions will be effected?
  - If entity X no longer has relationship R with entity Y, what events will be effected?

- **Regulation**
  - How does the A of B regulate X of Y?
  - Compare what happens to X of Y, in situation A vs situation B.
2011 DEVELOPMENT PLAN

Kickoff

Structure & Function

Similarity & Differences
Regulation

How & Why
Energy Transfer

Educationally
Useful KB

Inquire
Release 1.0

First “Double
Helix”

Second “Double Helix”

Third “Double
Helix”

After Class
Experiment

Inquire
Release 2.0

Feb  Mar  Apr  May  June  July  Aug  Sept  Oct  Nov  Dec  Jan  Feb
ASSESSMENT

- **Successes**
  - Basic QA performance (for a subset of questions)
  - Scalable KA (for a range of KR)
  - Exciting HaloBook/Inquire platform

- **Challenges**
  - Missing 50% of the knowledge
  - Brittleness for novel questions
  - Salience
  - Need to expand to a more hybrid approach
PROJECT HALO
DEVELOPMENT STRATEGY

<table>
<thead>
<tr>
<th>Authors</th>
<th>Pilot</th>
<th>Phase II</th>
<th>HaloBook</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single Domain Expert</td>
<td></td>
<td></td>
<td>Community of Scientists, Teachers, and KR Experts</td>
</tr>
<tr>
<td>Logic Queries</td>
<td>AP Question Answering</td>
<td>AP QA General QA Education</td>
<td>AP QA General QA Education Research</td>
</tr>
</tbody>
</table>
HARD PROBLEMS IN AI
(FROM THE APPENDIX OF IDEA MAN)

- Full Range of Human Language
- Visual/Spatial Reasoning
- Actions, Causality, and Simulation
- Pervasive Uncertainty and Vagueness

- Implicit Knowledge in Language
- Contradictions and Messy Knowledge
- Commonsense Reasoning
- Applying Knowledge in New Contexts
NEWLY FORMED (AI)^2

- **Mission**
  - Supports long-range research that has the potential to accelerate progress in AI
  - Focused on solving issues of scale and brittleness that have traditionally limited progress in AI

- **Activities**
  - Supports individual researchers and research groups
  - Sponsors conferences, prizes, competitions, and the construction of large public knowledge bases and evaluation frameworks

- **Process**
  - (AI)^2 is not soliciting applications for funding at this time
AI CHALLENGE PROBLEMS

- History of AI Challenges
  - Chess
  - Read a chapter (Reddy, 2003; Feigenbaum, 2003)
  - DARPA Grand Challenges

- DARPA Workshop on Cognitive Grand Challenges
  - Search Engine for the Real World: *Robotic Treasure Hunt*
  - Test Taker: *Taking the SAT*
  - Report Generator: *Handy Andy, the DARPA Essayist*
  - Reading to Learn: *Learn by reading a textbook*

- Many many others

KCAP 2011
DEEP KR CHALLENGE

- **Challenge:**
  - Given the difficult-to-represent sentences from Halobook
  - Participants were invited to submit papers addressing the KRR challengers

- **Prize Winners:**
  - “Using Answer Set Programming for Representing and Reasoning with Preferences and Uncertainty in Dynamic Domains” by Ravi Palla, Dan Tecuci, Vinay Shet and Mathaeus Dejori
  - “How to model the shapes of molecules?” by Janna Hastings, Oliver Kutz and Till Mossakowski
  - “Why is Surface Area Important to Normal Cell Function?” by Geoff Sutcliffe and Adam Pease
(AI)2 CHALLENGES

Options
- Integrated vs. individual challenges
- Standing prize vs. annual competition
- What infrastructure to provide

Issues
- Selecting the right metric
- Incentives for participation
- How to integrate results
Thank You