Living in a Transparent Future: Search in a Wired World

Michael Wollowski
Computer Science and Software Engineering Department
Rose-Hulman Institute of Technology
Introduction

- Current state of search
- Goals for the web: The semantic web
- The essence of the semantic web
- Technologies for a wired future
- Three likely scenarios of a wired future
- Conclusions
Current Work in Search

- Material on the web is largely designed for human consumption
- Information is contained:
  - On web pages
  - In databases
- Two major clients for search engines:
  - Individual web-searcher
  - Corporations
Current Work in Search:
Individual web-searcher

- Wants a genie which answers questions
- Examples:
  - When was Harry S Truman born?
  - Find me a 1999-2004 Ford Lightning for $20,000
  - We’re looking for a family resort in the Caribbean with baby sitting, other activities for a family with a one and a three year old. Any Suggestions? [Liddy, 2002]
Current Work in Search: Individual web-searcher

Progress requires:
- Natural language processing techniques
- Artificial intelligence techniques for extracting information from documents
- Ontology
- Inference
- Semantic web
Current Work in Search: Corporate web-searcher

- Wants to locate documents which contain particular information
- Documents come in a variety of formats (jpg, pdf, ppt, xls, etc.)
- Example
  - 2000 Powerpoint presentation which introduced our Java-enabled car
- In a large corporation, there are likely 100s of documents which would match such a query
- Similar to the attempt of locating documents on desktop; except it is scaled to the company level
Current Work in Search: Corporate web-searcher

- Progress requires:
  - Locating distinguishing aspects of documents
  - Quick summary of documents
  - Artificial intelligence techniques
  - Semantic web
Old Goals and New Developments for the World-Wide Web

- The semantic web is a major current development for the web.
- The semantic web is designed to provide: “data integration across application and organization boundaries.” [TBL, 2003]
- This goal is even older than the web.
Old Goals and New Developments for the World-Wide Web

From TBL’s 1989 proposal for the “web.”
The Semantic Web

- Suppose we want to uniquely identify a person
  - Harry S Truman
  - Harry S. Truman
  - The 33rd president of the USA
- They all refer to the same person
- Traditionally: SSN
- Nowadays: Credit card numbers
The Semantic Web

- The semantic web would use a unique URI
- `<ID="http://www.whitehouse.gov/history/presidents/ht33.html">`<name>Harry S Truman</name> </ID>

- There is more than one URI about Truman
  - Ensure that everyone uses the same URI
  - Use a small set of URIs
  - Use inference to determine that they are about the same person
The Semantic Web

- Predicates are used to express attributes of objects and relationships between objects
- Example:
  - Harry loves Bess
- We know the meaning of “loves”
- Does a computer?
- It doesn’t matter
- For the purpose of the semantic web, we are interested in uniquely identifying predicates
The Semantic Web

- RDF (Resource Description Framework) gives “meaning” to predicates
- `<rdf:RDF xmlns="urn:predicates:">
  <rdf:loves>
    <rdf:person ID="…">Harry</rdf:person>
    <rdf:person ID="…">Bess</rdf:person>
  </rdf:loves>
</rdf:RDF>`
The Semantic Web

- Does the following refer to the same state of affairs?
- `<rdf:RDF xmlns="urn:predicates:">
  <rdf:liebt>
    <rdf:Person ID="…">Harry</rdf:Person>
    <rdf:Person ID="…">Bess</rdf:Person>
  </rdf:liebt>
</rdf:RDF>

- A computer can tell, if we add an inference engine
Towards a Wired Future

- The semantic web will facilitate progress towards the genie and for corporate search.
- In this talk, we want to focus on an area that is currently in its infancy: the wired world.
- In a wired world, devices are connected to the internet.
- Many devices will be connected in a wireless fashion.
Towards a Wired Future

- In principle, we can connect just about any device to the internet

Examples:
- Through X-10: anything for which we would traditionally use a switch, e.g. lights and coffee machines
- Refrigerators (Samsung)
- Soda machines (soda.cs.rose-hulman.edu)
Living in a Connected World

- How will life be different in a world in which “everything” is connected to the internet?
- Let’s assume a happy world in which information is used responsibly
Scenario 1: A connected car

Janice J. Heiss. The Network is the car.
http://java.sun.com/features/1999/06/concept_car.html,
Scenario 1: A connected car

- **Goal:** Color coded maps show traffic delays on the car's navigation system
- Interested? Buy an Acura RL or a Cadillac CTS
- **Requirements:**
  - Car sends its location to a computer
  - Car receives information
  - Car has electronic maps
- **Desirable:**
  - Navigation system plots alternate route
Scenario 1: A connected car

- **Goal 2**: Finding a parking space - in Boston
- Impossible!
- MIT’s Media Lab is thinking about *smart curbs*:
  - Parking spaces are wired
  - Report their availability
- Make them interact with E-Z Pass
- Can be used instead of parking meters
Scenario 1: A connected car

- Augment your car’s navigation system to direct you towards an available parking space
- It will likely be taken by the time you arrive
- Use AI to predict availability of parking spaces in an area
- Have your car’s computer talk to your PDA when it becomes obvious that you will be late for a meeting
- Have your PDA talk to your colleagues to move up the meeting for which you will be late
- Receive a note on your PDA that you are fired, because you missed yet another important client meeting
Scenario 1: A connected car

● **Requirements:**
  ● A GPS system
  ● Common communication protocols for car, curb, toll systems, PDA, colleagues’ schedules

● Semantic web ensures compatibility of data

● Java can ensure compatibility of systems

● AI and search for various tasks

● Does not have to be perfect
Scenario 2: A connected house

Scenario 2: A connected house

- Alzheimer’s Association, together with Intel sponsors research into connected houses
- **Goal**: To assist people with Alzheimer’s
- **In particular**: Reminding people of tasks such as cleaning the refrigerator or taking medicine
- **Benefits**:
  - Increased independence
  - Reduced health care cost
Scenario 2: A connected house

- Most of the devices shown in the drawing are in the future
- Samsung’s connected refrigerator is a refrigerator with a laptop attached to it
- It cannot report status of items in it or whether it needs to be cleaned
Scenario 2: A connected house

- Houses wired with X-10 and connected to the web are a reality
- The web-based soda machine has sensors detecting whether the door is open or whether the vending flap is open
- We count how many sodas were sold
Scenario 3: A java-enabled person

Wired Magazine, February 2000
Scenario 3: A java-enabled person

In Search of a Juicy Burger

[see Example 3 from paper]
Scenario 3: A java-enabled person

- Pace makers and defibrillators are implanted in people
- They collect data
- Data is analyzed on the device
- Data is downloaded to a separate device
- Data is sent to a server and analyzed by medical personnel
Scenario 3: A java-enabled person

- Microsoft’s Cambridge research center developed a prototype of a personal image and data recall system.
- McDonald’s operates a restaurant in Long Island which enables people to pay for their meals with the E-Z Pass system.
- Insurance companies (Progressive) offers discounts for drivers agreeing to install a black box.
- Life insurance policies have lower rates for non-smokers.
Conclusions

- **Claim**: The connected world may be here sooner than we think
- This is in contrast to the genie which will likely take more time
Conclusions

- More and more devices will be wired to the internet
- Much of the information necessary to make a connected world is already stored in electronic format:
  - Credit card transactions
  - PayPal transactions
  - Electronic toll systems
Conclusions

- Most of those systems can already uniquely identify users across systems.
- This is simply because in order to get an account, you have to divulge information such as credit card numbers and bank account numbers.
- The connected world relies on highly structured information.
- This information is meant to be processed by programs.
- It is not meant for human consumption.
Conclusions

- Ramifications of a connected future are hard to fathom
- Search goes on behind the scenes
- Search is not user initiated
- Search is initiated by the devices we use
- The internet “knows” us
- Individual systems may not know a lot about us, just enough to conduct necessary transactions
Conclusions

- Information, be it good or bad will be available quickly.
- Adds a lot of conveniences to our lives.
- One small misstep may have a big effect on people’s lives.
Future Work

- Develop acceptable norms for information use
- Build security into the systems we are about to build, so as to enforce information use policies