Chapter 4

Berners-Lee: What is Solvable on the Web?

Content in the chapter

- Overview
- The World Wide Web
- Tim Berners-Lee
- The Semantic Web Roadmap
- Logic on the Semantic Web
- Conclusion

The World Wide Web

- Three major events accelerated the development of the Information Revolution
  - The introduction of the World Wide Web (WWW)
  - Increased availability of the graphical browser
  - The unleashing of commercialization on the Internet
- An important man: Tim Berners-Lee
  - Hyper Text Transfer Protocol (HTTP)
  - Hypertext Markup Language (HTML)
  - Universal Resource Locator (URL)

Tim Berners-Lee

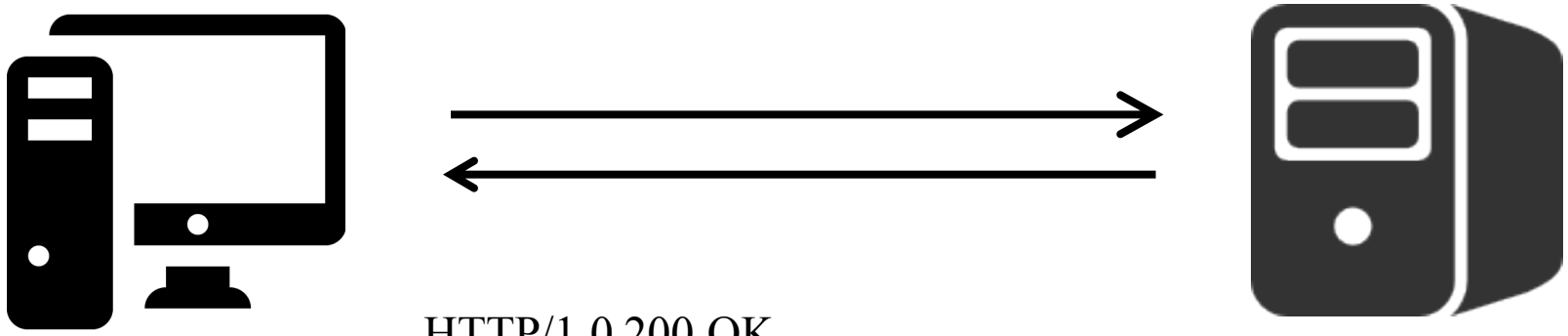
- Known as the inventor of the World Wide Web
- Today he is the head of the World Wide Web Consortium (W3C)
- Looked at ways to simplify the use of the Internet
- Chose to continue as an academic instead of an Internet entrepreneur

HTTP

- Network protocol used to deliver files and data on the Web
- Takes place (usually) through TCP/IP sockets
- Client-Server model
  - An HTTP client opens a connection and sends a request message to an HTTP server
  - The server returns an response message containing the requested resource
  - After delivering the resource the server closes the connection

HTTP

GET/path/file.html HTTP/1.0
From: someuser@somehost.com
User-Agent: HTTPTool/1.0

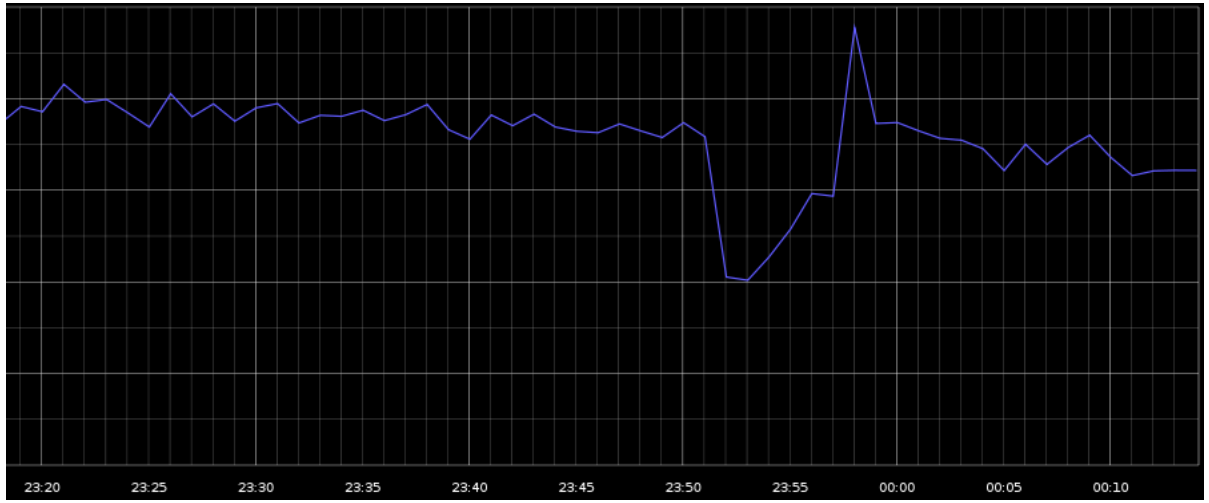


HTTP/1.0 200 OK
Date: Fri, 31 Dec 1999 23:59:59 GMT
Content-Type: text/html
Content-Length: 1354

```
<html>
 <body>
   Hello World
</body>
</html>
```

Limits of Today's Web

- The Web has changed over the years
- Google accounts for about 40% of total Internet traffic in 2013 (https://engineering.gosquared.com/googles-downtime-40-drop-in-traffic)
- Expect an increase in average bandwidth



Limits of Today's Web

- Developed for humans without a focus on automatic processing
- Hard to do information retrieval from HTML as it is not directly capable of being directly exploited for this
- We are stuck with keyword searches
- Today's web does not have meaningful relations between content
- This is where the Semantic Web comes in

The Semantic Web roadmap

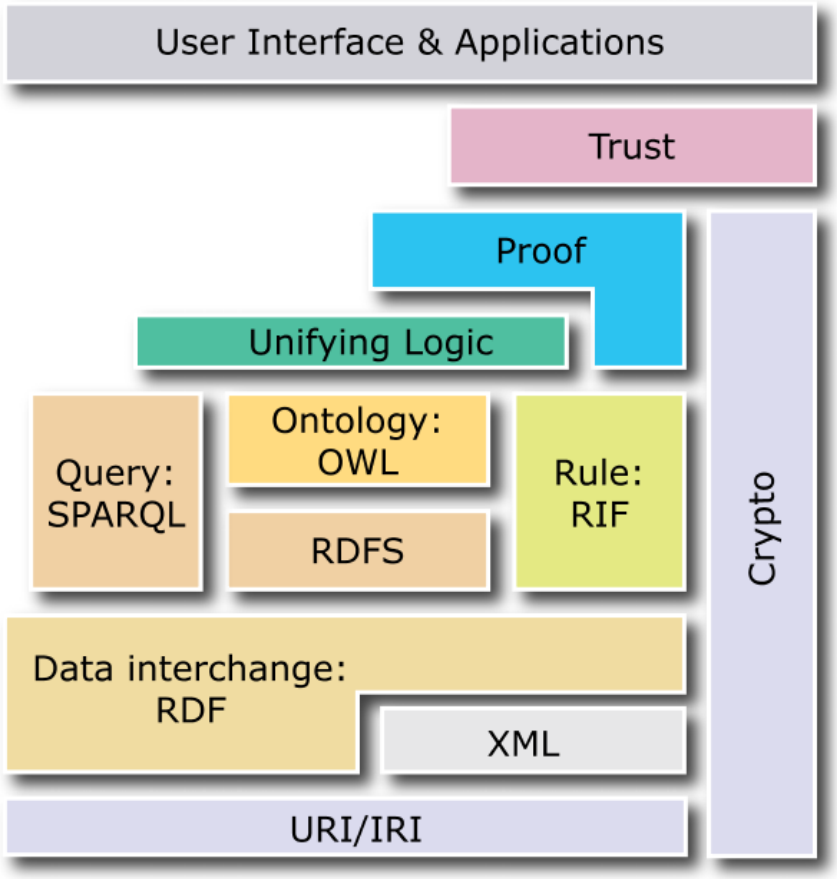
- “On the Semantic Web we will be able to express ourselves in terms that our computers can interpret and exchange. By doing so, we will enable them to solve problems that we find tedious, to help us find quickly what we're looking for: medical information, a movie review, a book purchase order, etc. The W3C languages RDF, XML, XML Schema and XML signatures are the building blocks of the Semantic Web”

  –W3C

The Semantic Web roadmap

- Meaningful content
- The Semantic Web should provide a framework
- Powerful logic, but at the same time not too powerful
  - Can be tricked into considering a paradox
    - A **paradox** is a statement that apparently contradicts itself and yet might be true (or wrong at the same time).

The Semantic Web roadmap



Resource Description Framework (RDF)

- Contains the concept of an assertion and allows assertions about assertions.

- Example:
  - “The book [subject] has the title [predicate] *Gödel, Escher, Bach: An Eternal Golden Braid* [object].”

Ontology

- An ontology is an agreement between agents that exchange information
- Enables agents to exchange vast quantities of data
- Web Ontology Language (OWL)
  - Uses RDF
  - Facilitates greater machine readability of Web content

Making an Inference

- Machines cannot approach the level of human intelligence in terms of general purpose cognitive capability
- The Semantic Web will not be possible until software agents have the means to figure out some things by themselves
- The use of Knowledge Representation and inference will help us in the process
  - Combined gives us an inference engine, which process the knowledge we have to derive resulting conclusions

Computational Complexity for Large Networks

- Computational-complexity theory establishes how much of a resource is required to solve certain classes of computations
- Some problems might be too hard to solve
- Figure out a balance
  - Too powerful: might be tricked
  - Too little power: not sufficient for performing desired tasks
- One solution is to accept an approximate instead of an exact solution

SWRL

- The Semantic Web Rule Language (SWRL)
  - XML-based
  - Play an important role in facilitating business rules and interactions over the Web

Rule-Based Systems

- Rule-based systems use symbolic representation and inference mechanisms to reach conclusions from facts and rules
- Expert systems are applications that make decisions in real-life situations, which would otherwise be performed by a human expert
  - Medicine etc.
  - Inference engine + knowledge = expert system

Inference Engines

- Controls overall execution of a set of rules to process the knowledge available
- Forward chaining vs. backward chaining

Semantic Web Services

- OWL-S (an ontology for Web Services)
  - Discovery
  - Invocation
  - Composition and interoperation
  - Execution monitoring

Logic on the Semantic Web

- The goal is to create a unifying system
- The Semantic Web does not actually define a reasoning engine
- The Semantic Web cannot find answers, it cannot even check that an answer is correct, but it can follow a simple explanation that an answer is correct

Intractability and Undecidability

- Operating with a mass of untrustworthy data
- Finding a self-contradictory statement, what to do?
- Challenges
  - Semantics are complex, and it will not be easy to use
  - Practices like metatag spamming, and even trademark hijacking, are subject to abuse
  - Because of the diversity in developers and development tools, Semantic Web technology will have to be technically open for implementation and use.

Summary

- Machine-understandable
- Formal logic introduced to the Web architecture
- The architecture of the Semantic Web
- Balance is key to make it work