Autonomous Agents on the Web:

Beyond Linking and Meaning

Mike Amundsen
Principal API Architect
Layer 7 Technologies
@mamund

LAYER 7 TECHNOLOGIES

SALAD2013

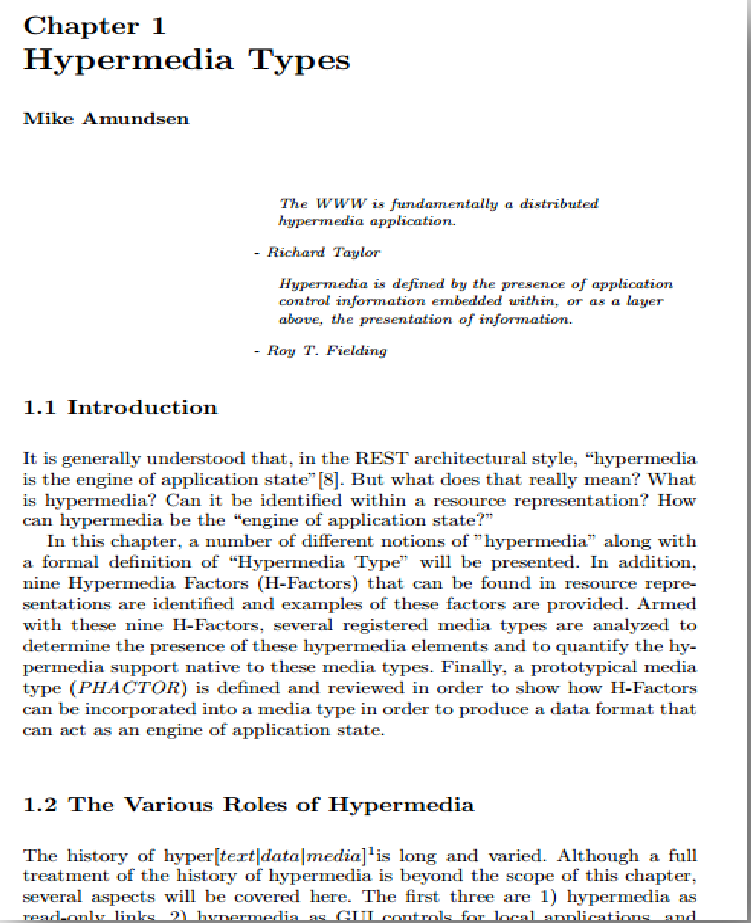
Preface

Relevant past, present, and future activities

Preface

- 2010 on the Web
- Abstract the hypermedia
- Provide analysis tools

- Hypermedia Factors
  http://g.mamund.com/factors



Chapter 1
Hypermedia Types

Mike Amundsen

> *The WWW is fundamentally a distributed hypermedia application.*
>
> *- Richard Taylor*
>
> *Hypermedia is defined by the presence of application control information embedded within, or as a layer above, the presentation of information.*
>
> *- Roy T. Fielding*

1.1 Introduction

It is generally understood that, in the REST architectural style, "hypermedia is the engine of application state"[8]. But what does that really mean? What is hypermedia? Can it be identified within a resource representation? How can hypermedia be the "engine of application state?"

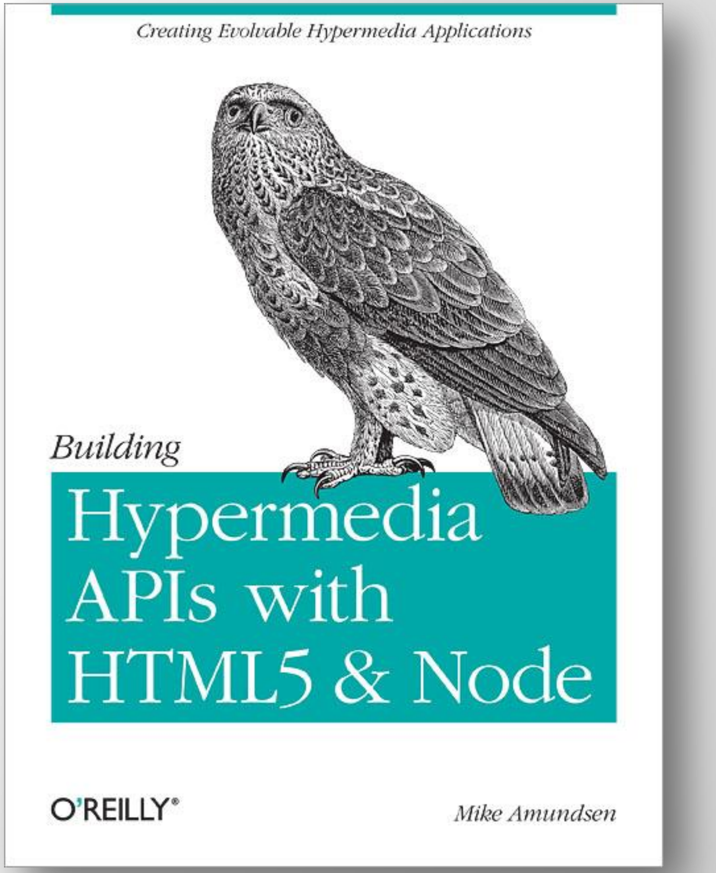
In this chapter, a number of different notions of "hypermedia" along with a formal definition of "Hypermedia Type" will be presented. In addition, nine Hypermedia Factors (H-Factors) that can be found in resource representations are identified and examples of these factors are provided. Armed with these nine H-Factors, several registered media types are analyzed to determine the presence of these hypermedia elements and to quantify the hypermedia support native to these media types. Finally, a prototypical media type (*PHACTOR*) is defined and reviewed in order to show how H-Factors can be incorporated into a media type in order to produce a data format that can act as an engine of application state.

1.2 The Various Roles of Hypermedia

The history of hyper[*text*|*data*|*media*][1]is long and varied. Although a full treatment of the history of hypermedia is beyond the scope of this chapter, several aspects will be covered here. The first three are 1) hypermedia as read-only links, 2) hypermedia as GUI controls for local applications, and

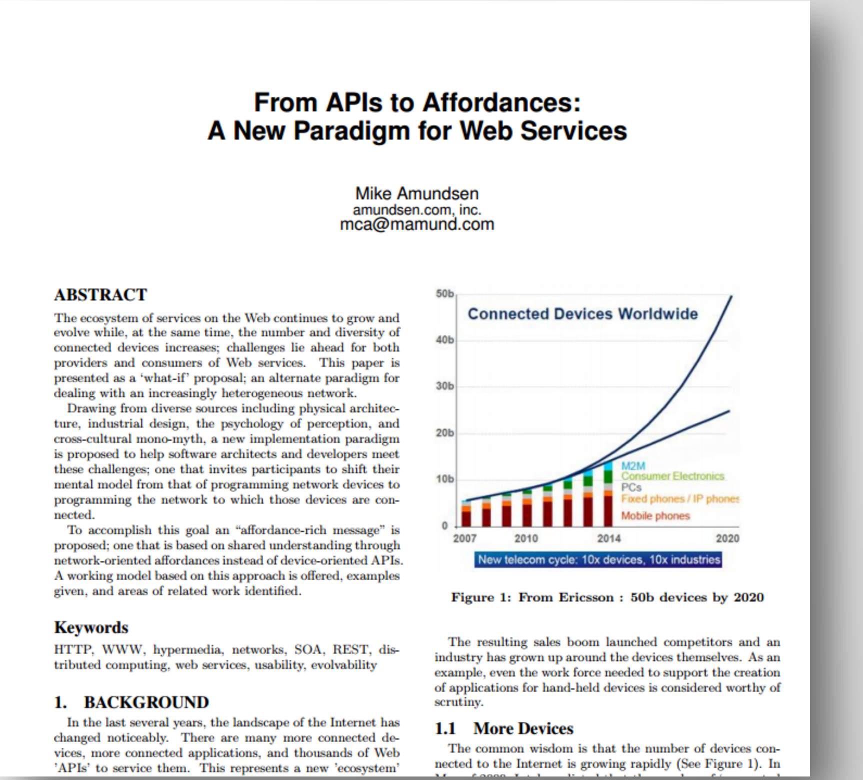
Preface

- 2011 - O’Reilly
- Hypermedia focus
- Program the message


Creating Evolvable Hypermedia Applications
Building
Hypermedia APIs with HTML5 & Node
O'REILLY®
Mike Amundsen

Preface

- 2012 @ WS-REST
- Affordance focus
- Program the network



From APIs to Affordances: A New Paradigm for Web Services

Mike Amundsen
amundsen.com, inc.
mca@mamund.com

ABSTRACT

The ecosystem of services on the Web continues to grow and evolve while, at the same time, the number and diversity of connected devices increases; challenges lie ahead for both providers and consumers of Web services. This paper is presented as a 'what-if' proposal; an alternate paradigm for dealing with an increasingly heterogeneous network.

Drawing from diverse sources including physical architecture, industrial design, the psychology of perception, and cross-cultural mono-myth, a new implementation paradigm is proposed to help software architects and developers meet these challenges; one that invites participants to shift their mental model from that of programming network devices to programming the network to which those devices are connected.

To accomplish this goal an "affordance-rich message" is proposed; one that is based on shared understanding through network-oriented affordances instead of device-oriented APIs. A working model based on this approach is offered, examples given, and areas of related work identified.

Keywords

HTTP, WWW, hypermedia, networks, SOA, REST, distributed computing, web services, usability, evolvability

1. BACKGROUND

In the last several years, the landscape of the Internet has changed noticeably. There are many more connected devices, more connected applications, and thousands of Web 'APIs' to service them. This represents a new 'ecosystem'

Figure 1: From Ericsson : 50b devices by 2020

The resulting sales boom launched competitors and an industry has grown up around the devices themselves. As an example, even the work force needed to support the creation of applications for hand-held devices is considered worthy of scrutiny.

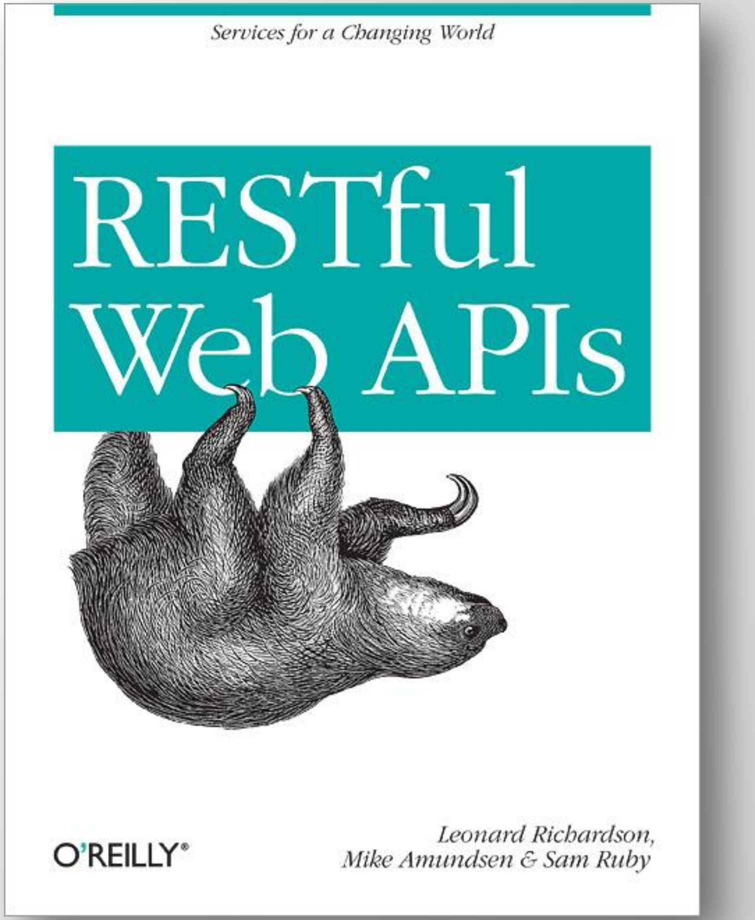
1.1 More Devices

The common wisdom is that the number of devices connected to the Internet is growing rapidly (See Figure 1). In

Preface

- 2013 – O’Reilly
- Profile focus
- Program the description

- Application-Level
  Profile Semantics (ALPS)
  http://alps.io


Services for a Changing World
RESTful Web APIs
O’REILLY
Leonard Richardson, Mike Amundsen & Sam Ruby

Preface

But today, something different…

"One does not discover new lands without consenting to lose sight of the shore for a very long time“

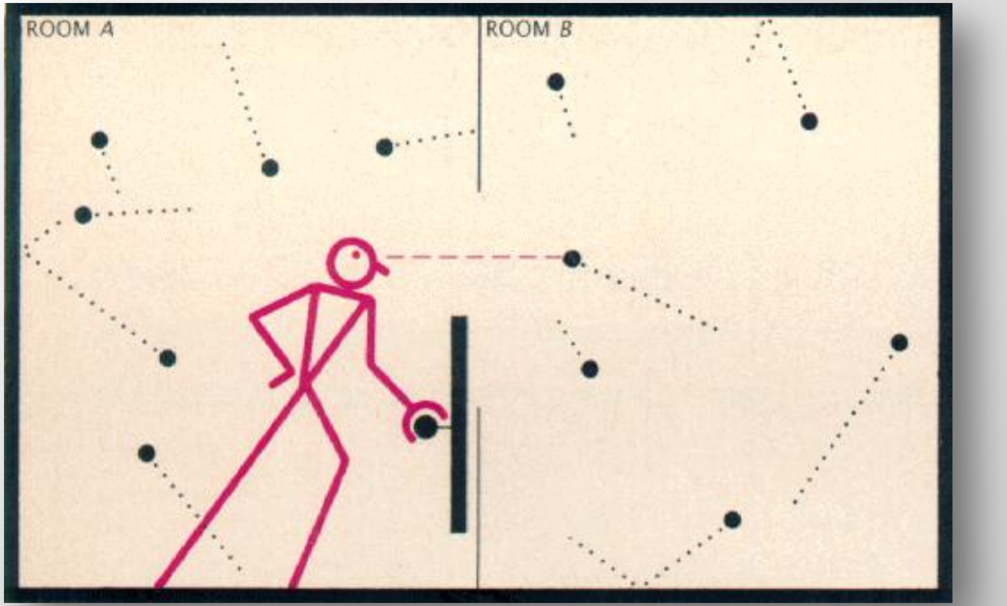
- André Gide (1869-1951)

Background

Information Theory, Complex Systems, and Hypermedia

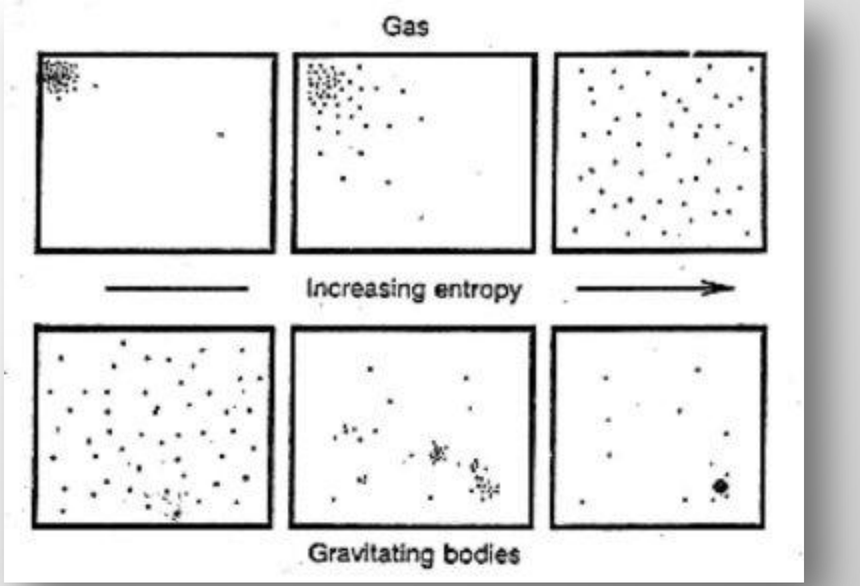
Maxwell’s Demon

- James Clerk Maxwell (1831 - 1879)
- *“... if we conceive of a being whose faculties are so sharpened that he can follow every molecule in its course…”*
- Second Law of Thermodynamics *“has only a statistical certainty”*



Boltzmann

- Ludwig Boltzmann (1844 - 1906)
- “Boltzman entropy”
- Macro- & micro-states
- Each possibility is a microstate
- The probability of a macrostate is the function of all the microstates.

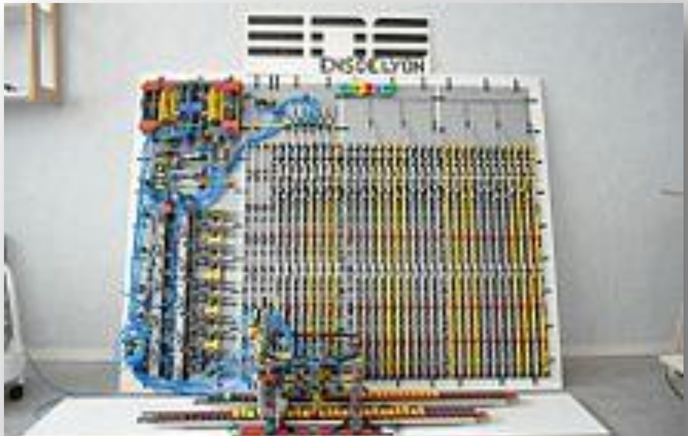


Shannon & Information

- Claude Shannon (1916 – 2001)
- *“The number of bits needed to represent the result of an uncertain event is given by its entropy.”*
- **Surprisal**: the "surprise" of seeing the outcome - a highly improbable outcome is very surprising. (Tribus, 1961)

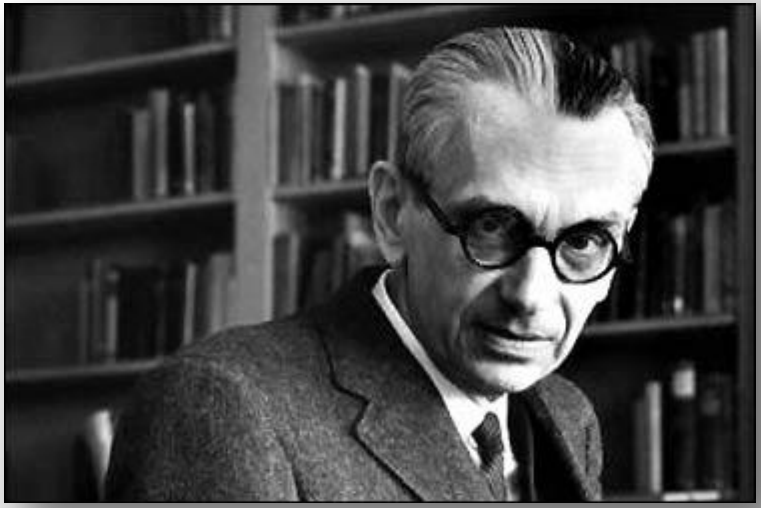
Turing, Tapes, & Halting

- Alan Turing (1912 – 1954)
- A **Turing machine** is a hypothetical device that manipulates symbols on a strip of tape according to a table of rules.
- *“Turing's paper ... contains, in essence, the invention of the modern computer.” (Minsky, 1967)*
- *“… decide whether the program finishes running or continues to run forever”*



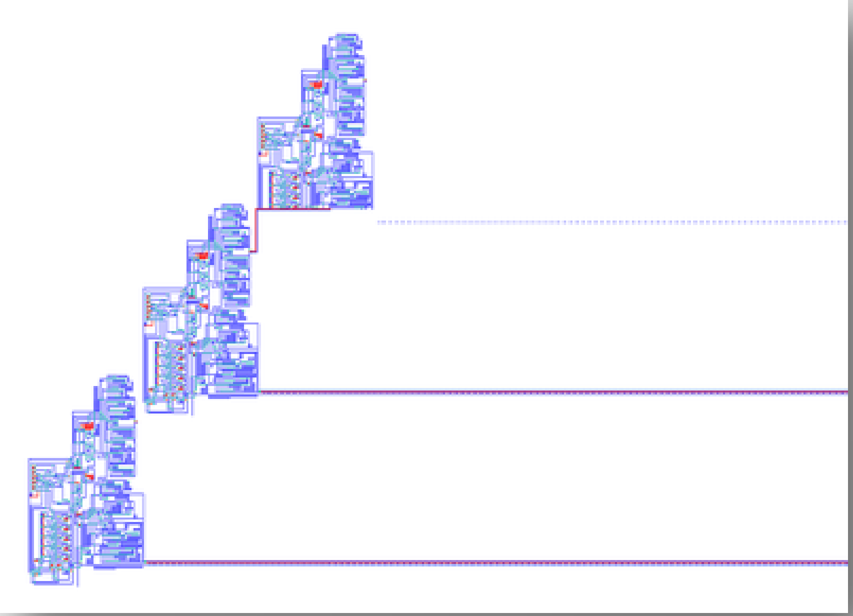
Gödel and Incompleteness

- Kurt Gödel (1906 – 1978)
- *“This statement is unprovable.”*
- *Treats the string as both **data** and **program***



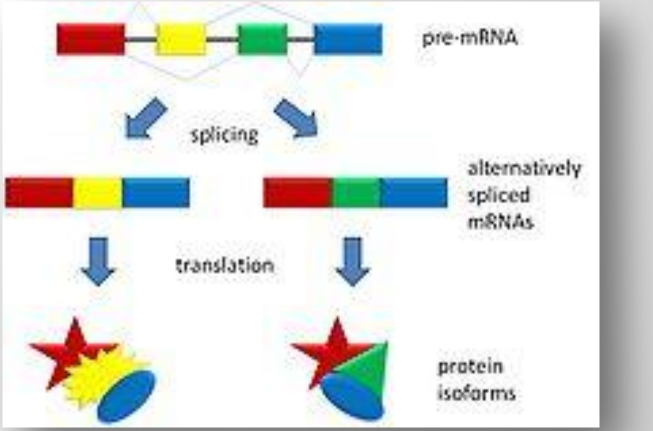
Von Neumann computing

- John von Neumann (1903 – 1957)
- *Described a computer architecture in which the* ***data*** *and the* ***program*** *are both stored in the computer's memory in the same address space.”*
- *Theory of Self Reproducing Automata (1966)*



Genes

- DNA/RNA store both the **data** and **program**.
- mRNA uses “alternative splicing” where it greatly increases biodiversity.



Fielding architecture

- Roy Fielding (1965 -)
- “*Architectural Styles and the Design of Network-based Software Architectures*” *(2001)*
- *“*each component cannot "see" beyond the immediate layer with which they are interacting.”
- *“…the information becomes the affordance…”*

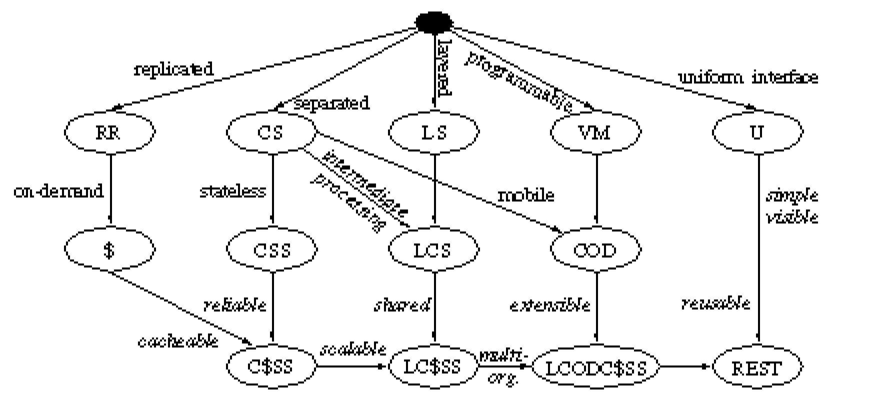
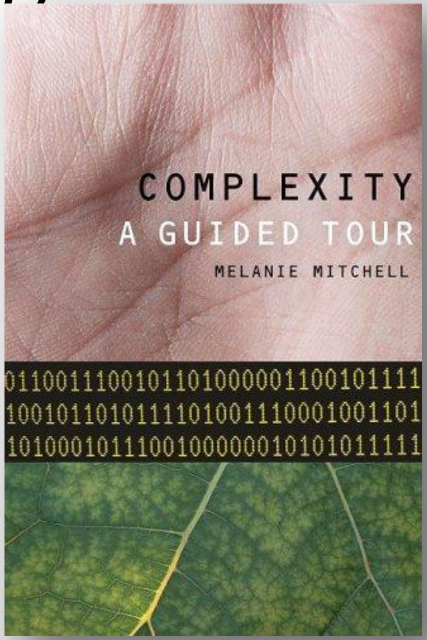


Figure 5-9. REST Derivation by Style Constraints

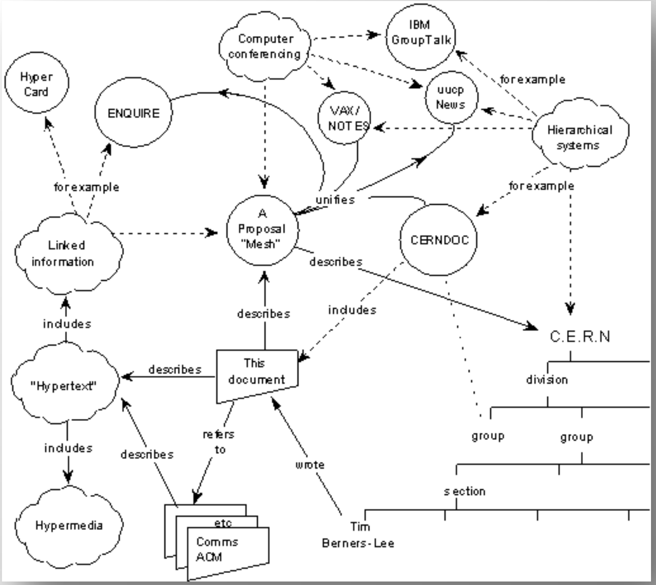
Complex Systems

- *“Large networks of components with no central control and simple rules of operation give rise to collective behavior, sophisticated information processing, and adaptation via learning or evolution.” (Mitchell, 2001)*
- *“Exhibits non-trivial emergent and self-organizing behavior.”*



The Web

- *“The Web … [has] many large scale properties … which lead to “adaptive” behavior for the system as a whole.” (Mitchell 2001)*



So much for the background!

Current State

Media Types, HTTP, and Kelvin-ism

Media Types

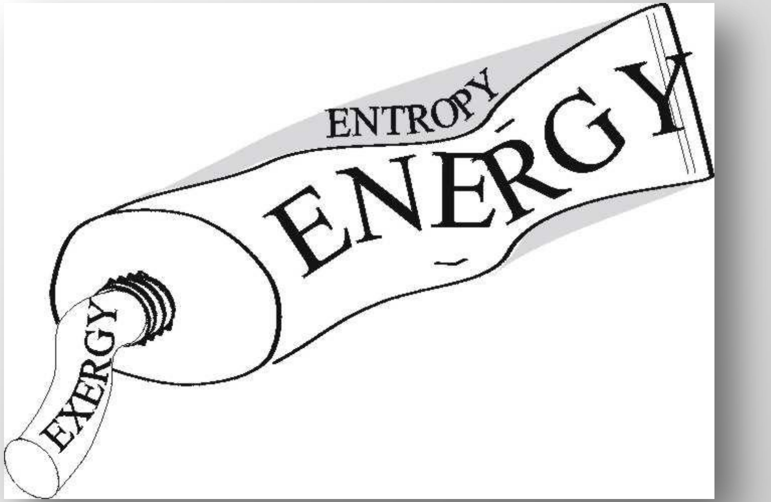
- More registered hypermedia-style designs in the last two years than in the last ten.
  - Maze+XML (experimental)
  - HAL (XML, JSON)
  - Collection+JSON
  - Siren (JSON)
  - Hydra (JSON-LD)
  - JSON-API (pending)

iana

Internet Assigned Numbers Authority

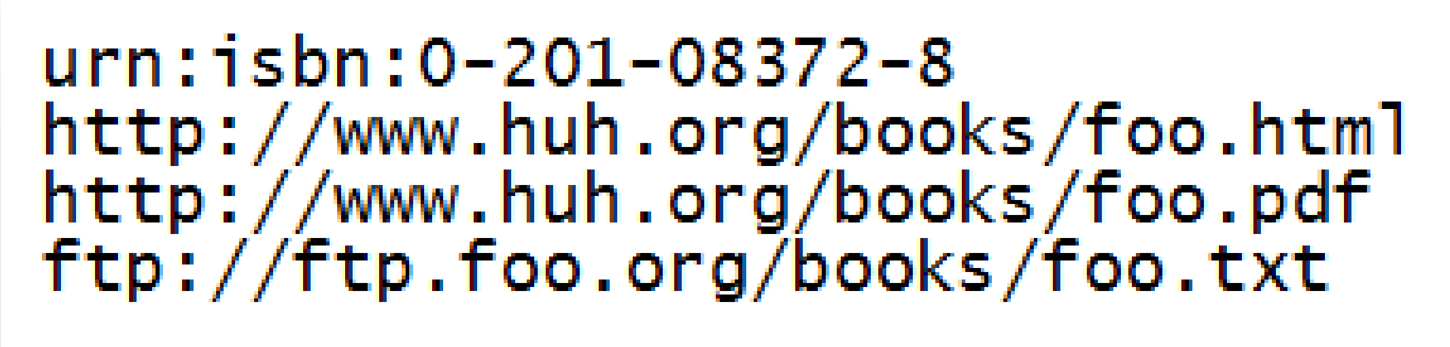
Media Types and entropy

- Designs vary in their level of “surprise”
- “surprisal” == “entropy”
- Lower the entropy, the less value the information
- Higher the entropy, the more energy needed to process the information.



Media Types and entropy

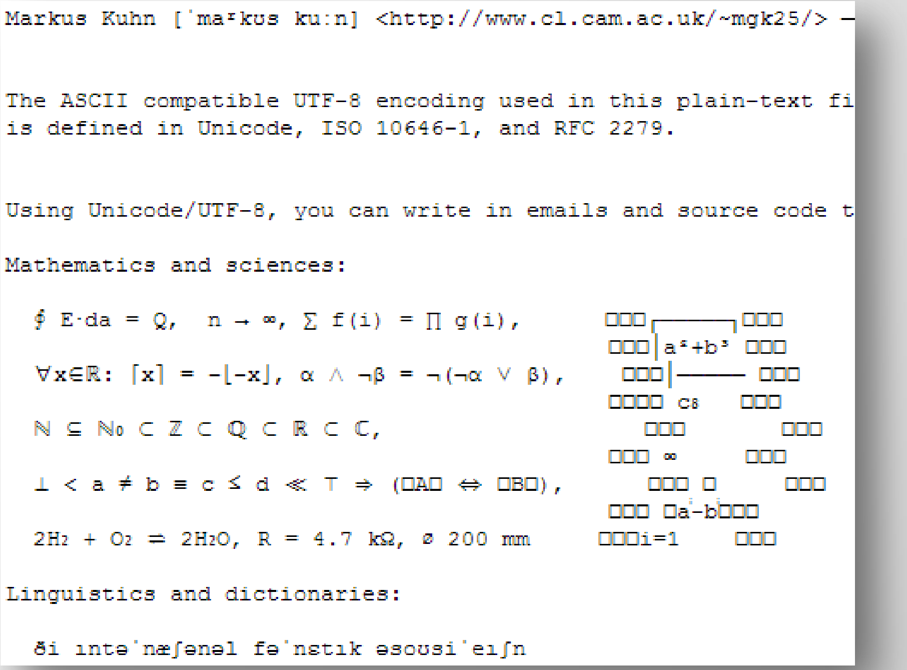
- **`text/uri-list`**
- Low entropy/surprisal
- Low energy needs



```
urn:isbn:0-201-08372-8
http://www.huh.org/books/foo.html
http://www.huh.org/books/foo.pdf
ftp://ftp.foo.org/books/foo.txt
```

Media Types and entropy

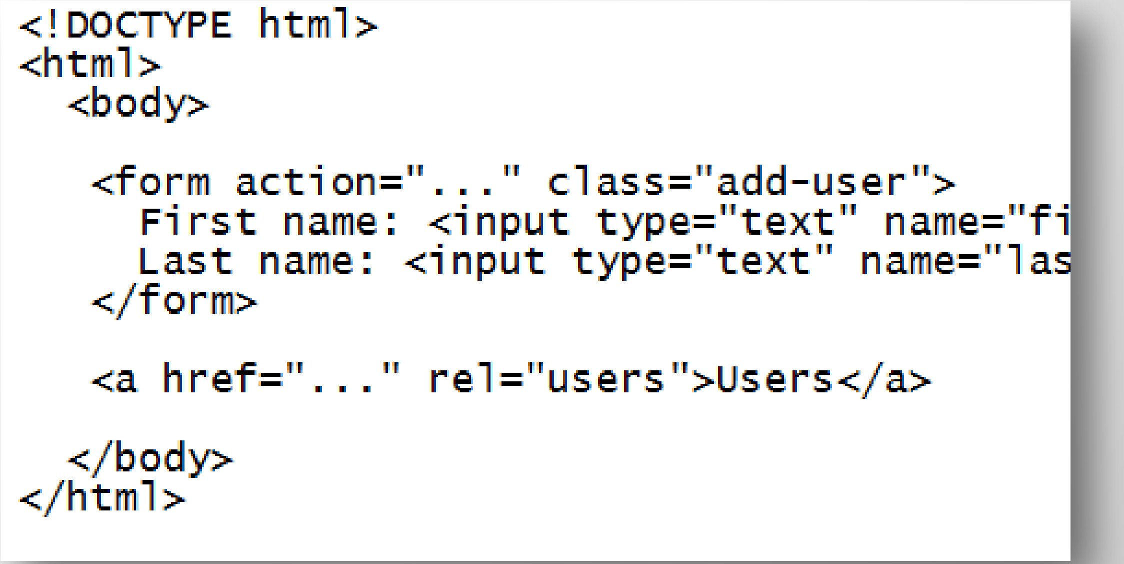
- **`text/plain`**
- High entropy/surprisal
- High energy needs



```
Markus Kuhn [ˈmaʳkʊs kuːn] <http://www.cl.cam.ac.uk/~mgk25/> —


The ASCII compatible UTF-8 encoding used in this plain-text fi
is defined in Unicode, ISO 10646-1, and RFC 2279.


Using Unicode/UTF-8, you can write in emails and source code t

Mathematics and sciences:

  ∮ E⋅da = Q,  n → ∞, ∑ f(i) = ∏ g(i),      □□□┌──────┐□□□
                                              □□□│a²+b³ □□□
  ∀x∈ℝ: ⌈x⌉ = −⌊−x⌋, α ∧ ¬β = ¬(¬α ∨ β),     □□□│────── □□□
                                              □□□□ c₈   □□□
  ℕ ⊆ ℕ₀ ⊂ ℤ ⊂ ℚ ⊂ ℝ ⊂ ℂ,                       □□□     □□□
                                              □□□ ∞     □□□
  ⊥ < a ≠ b ≡ c ≤ d ≪ ⊤ ⇒ (□A□ ⇔ □B□),          □□□ □    □□□
                                              □□□ □a-b□□□
  2H₂ + O₂ ⇌ 2H₂O, R = 4.7 kΩ, ⌀ 200 mm     □□□i=1    □□□

Linguistics and dictionaries:

  ði ıntəˈnæʃənəl fəˈnɛtık əsoʊsiˈeıʃn
```

Media Types and entropy

- **`text/html`**
- Moderate entropy/surprisal
- Moderate energy needs



```
<!DOCTYPE html>
<html>
  <body>

   <form action="..." class="add-user">
     First name: <input type="text" name="fi
     Last name: <input type="text" name="las
   </form>

   <a href="..." rel="users">Users</a>

  </body>
</html>
```

Media Types and entropy

- From the “machine point of view”…
- What is the balance between entropy and energy?
- Energy = computing power (coding time, source code, memory, etc.)

NAME: CORWIN, ALICIA M.
POB: SAN DIEGO, CA
ADDRESS: XXXX 20THSTNW
WASH. D.C. 20009
OCCUPATION: DEPUTY A.P.N.S.A.
FUNCTION:

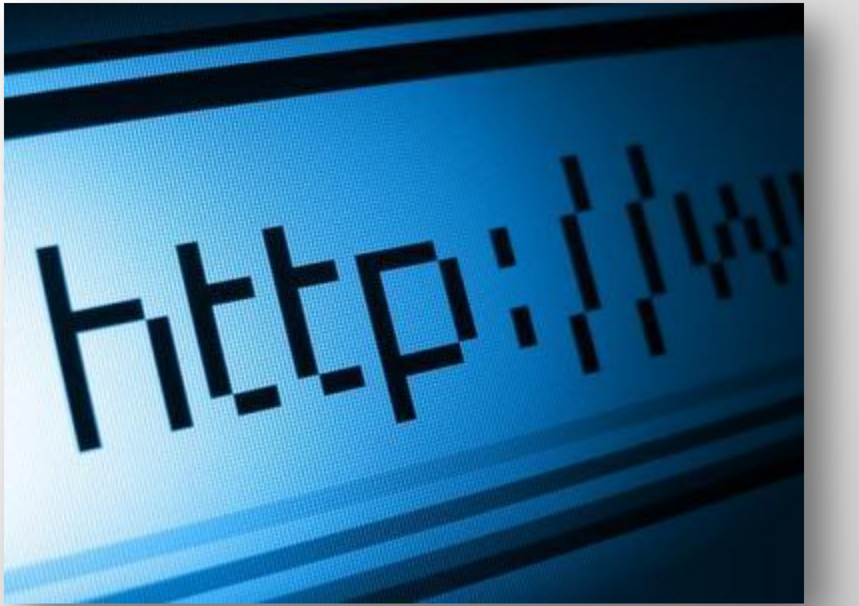
Media Types and entropy

- Most applications on the Web are “one-off” affairs - custom-coded for each solution.
- This is “high-energy computing!”

HTTP

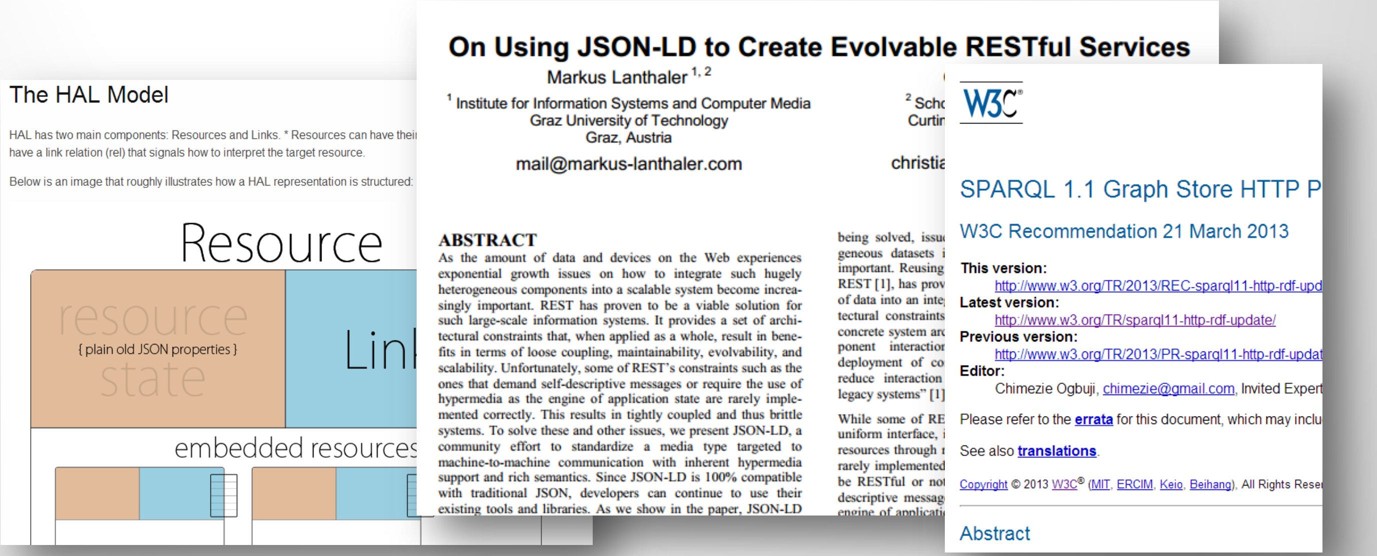
- Hypertext Transfer Protocol
  Ver 0.9 (1991) – Ver 1.1 (1999) *<10 years*
- HTTPbis (2013?) *~15 years since 1.1*
- HTTP 2.0 (20??) *>20 years since 1.1?*

- *No protocol-level changes, but several transport-level changes.*



HTTP

- **The Web is currently *highly dependent* on a single protocol.**
- Most new “protocols” build upon HTTP
  - *SPARQL 1.1 Graph Store HTTP Protocol.*
- Most new media types assume HTTP
  - *JSON-LD*
  - *HAL*



The HAL Model

HAL has two main components: Resources and Links. * Resources can have thei
have a link relation (rel) that signals how to interpret the target resource.

Below is an image that roughly illustrates how a HAL representation is structured:

On Using JSON-LD to Create Evolvable RESTful Services

Markus Lanthaler [1, 2]

[1] Institute for Information Systems and Computer Media
Graz University of Technology
Graz, Austria

mail@markus-lanthaler.com

[2] Scho
Curtin

christia

ABSTRACT

As the amount of data and devices on the Web experiences exponential growth issues on how to integrate such hugely heterogeneous components into a scalable system become increasingly important. REST has proven to be a viable solution for such large-scale information systems. It provides a set of architectural constraints that, when applied as a whole, result in benefits in terms of loose coupling, maintainability, evolvability, and scalability. Unfortunately, some of REST’s constraints such as the ones that demand self-descriptive messages or require the use of hypermedia as the engine of application state are rarely implemented correctly. This results in tightly coupled and thus brittle systems. To solve these and other issues, we present JSON-LD, a community effort to standardize a media type targeted to machine-to-machine communication with inherent hypermedia support and rich semantics. Since JSON-LD is 100% compatible with traditional JSON, developers can continue to use their existing tools and libraries. As we show in the paper, JSON-LD

being solved, issu
geneous datasets i
important. Reusing
REST [1], has prov
of data into an inte
tectural constraints
concrete system ar
ponent interactio
deployment of co
reduce interaction
legacy systems” [1]

While some of RE
uniform interface, i
resources through r
rarely implemented
be RESTful or not
descriptive messag
engine of applicati

SPARQL 1.1 Graph Store HTTP P

W3C Recommendation 21 March 2013

This version:
http://www.w3.org/TR/2013/REC-sparql11-http-rdf-upd
Latest version:
http://www.w3.org/TR/sparql11-http-rdf-update/
Previous version:
http://www.w3.org/TR/2013/PR-sparql11-http-rdf-updat
Editor:
Chimezie Ogbuji, chimezie@gmail.com, Invited Expert

Please refer to the **errata** for this document, which may inclu

See also **translations**.

Copyright © 2013 W3C® (MIT, ERCIM, Keio, Beihang), All Rights Rese

Abstract

The Irony of HTML and HTTP is…

© Original Artist
Reproduction rights obtainable from
www.CartoonStock.com
I KNOW THIS SOUNDS SILLY BUT I'M THIRSTY!
search ID: jctn15
chase

It is difficult to imagine what it would be like without them.

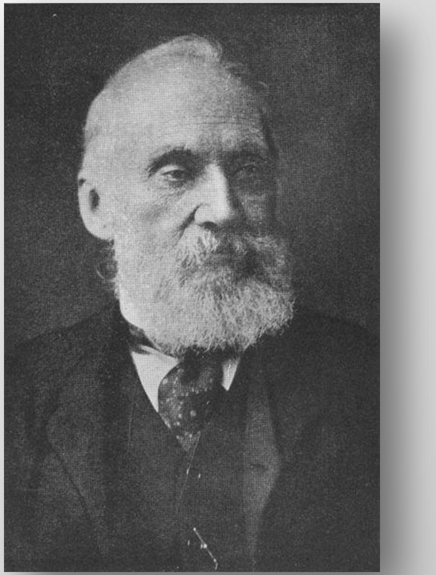
Questions for you…

- How long will HTTP last?
- When will HTML no longer be dominant?
- How will this affect your own thinking?
- How will this affect the Web?

Kelvin-ism

- Lord Kelvin computed the age of the earth based on “heat decay” and concluded:

  “….it was more than 20 and less than 40 million year old, and probably much nearer 20 than 40”. (Kelvin, 1897)



- To his dying day, Kelvin refused to accept the validity of other points of view.

Near Term

Lowering entropy, decoupling protocols, focusing on networks

Near Term – Lowering entropy

- We need *more* media type designs
- We need to design for low-entropy and high information
- We need to design for machines, not humans

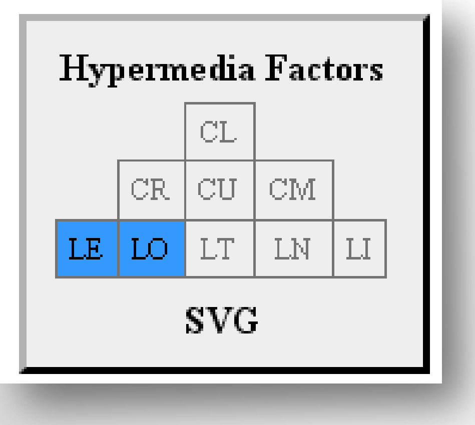
Near Term – Lowering entropy

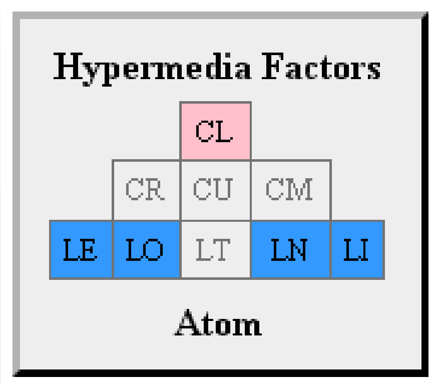
- Three semantic levels in network messages
  - Structure (XML, JSON, YAML, etc.)
  - Protocol (H-Factors)
  - Semantics (Domain concepts)
- We commonly see:
  - Structure = low surprise
  - Protocol = high surprise
  - Semantics = high surprise

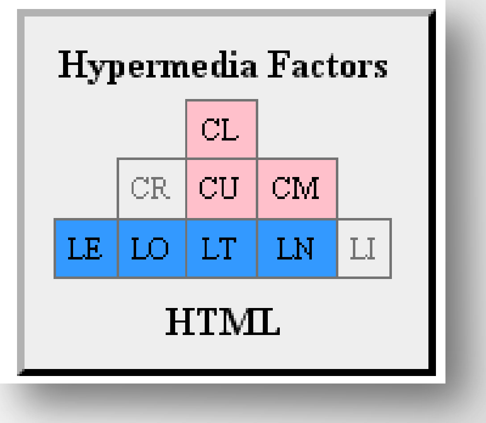
The higher the surprise in the message, the higher the dependence on custom code on the client/server.

Near Term – Lowering entropy

- Hypermedia Factors can lower Protocol Surprise
- Many designs are still unexplored.

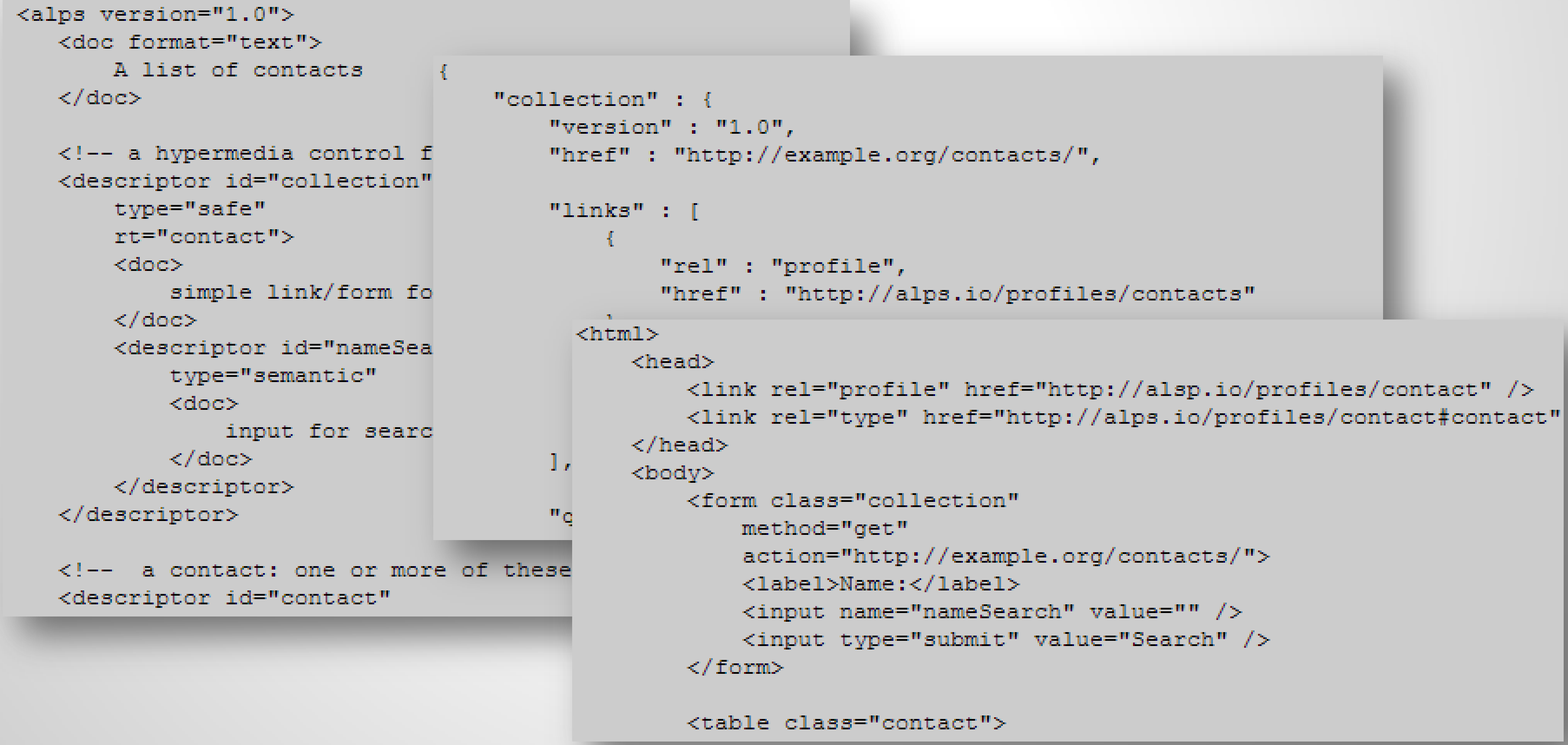






Near Term – Lowering entropy

- Profiles can lower Semantic Surprise
- http://alps.io



```
<alps version="1.0">
   <doc format="text">
      A list of contacts
   </doc>

   <!-- a hypermedia control f
   <descriptor id="collection"
      type="safe"
      rt="contact">
      <doc>
         simple link/form fo
      </doc>
      <descriptor id="nameSea
         type="semantic"
         <doc>
            input for searc
         </doc>
      </descriptor>
   </descriptor>

   <!--  a contact: one or more of these
   <descriptor id="contact"
```

```
{
    "collection" : {
        "version" : "1.0",
        "href" : "http://example.org/contacts/",

        "links" : [
            {
                "rel" : "profile",
                "href" : "http://alps.io/profiles/contacts"

        ],

        "q
```

```
<html>
    <head>
        <link rel="profile" href="http://alsp.io/profiles/contact" />
        <link rel="type" href="http://alps.io/profiles/contact#contact"
    </head>
    <body>
        <form class="collection"
            method="get"
            action="http://example.org/contacts/">
            <label>Name:</label>
            <input name="nameSearch" value="" />
            <input type="submit" value="Search" />
        </form>

        <table class="contact">
```

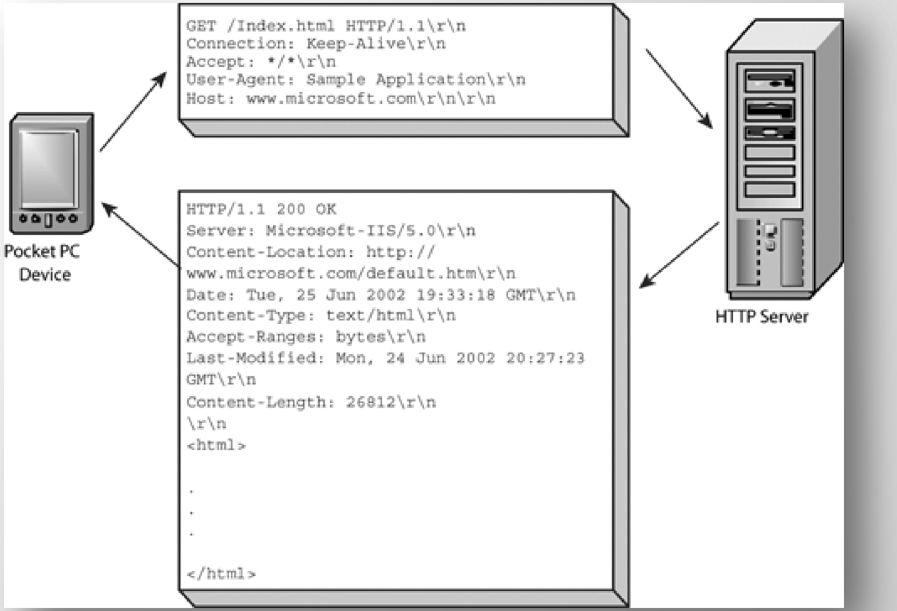
Near Term – Lowering entropy

- We need more machine-oriented media types.
- Text can add entropy for machines.
- **`rel="users"`**
  vs.
  `<a … >Users</a>`
- Imagine a hypermedia type that humans could not understand, but machines could.

The higher the dependence on machine-readable messages, the lower the entropy.

Near Term – Decoupling protocols

- Most media type designs today assume a dependence on a single protocol – HTTP.



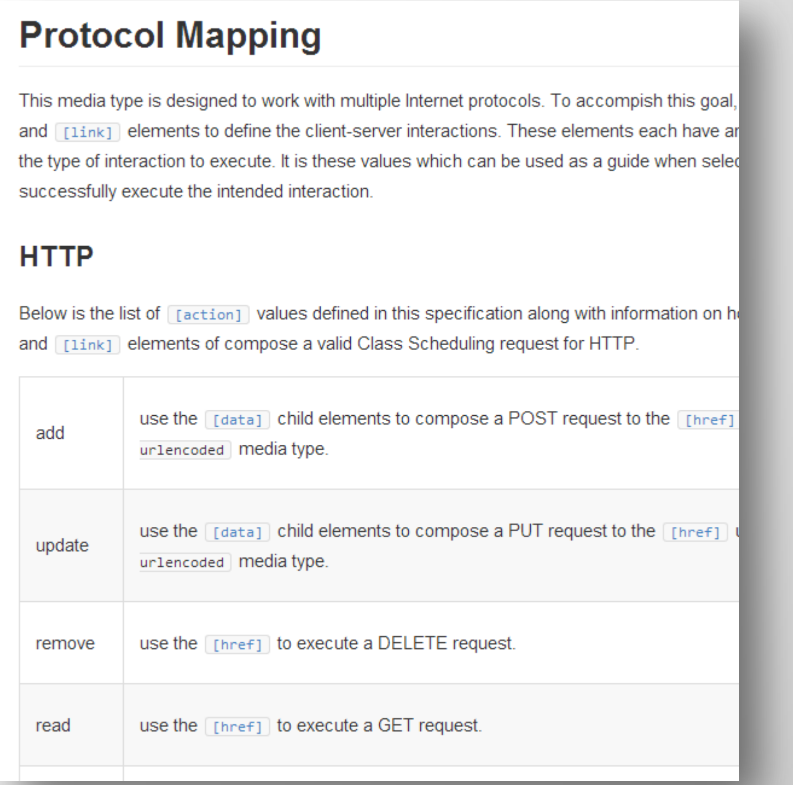
```
GET /Index.html HTTP/1.1\r\n
Connection: Keep-Alive\r\n
Accept: */*\r\n
User-Agent: Sample Application\r\n
Host: www.microsoft.com\r\n\r\n
```

```
HTTP/1.1 200 OK
Server: Microsoft-IIS/5.0\r\n
Content-Location: http://
www.microsoft.com/default.htm\r\n
Date: Tue, 25 Jun 2002 19:33:18 GMT\r\n
Content-Type: text/html\r\n
Accept-Ranges: bytes\r\n
Last-Modified: Mon, 24 Jun 2002 20:27:23
GMT\r\n
Content-Length: 26812\r\n
\r\n
<html>

.
.
.

</html>
```

Near Term – Decoupling protocols

- Message designs should be protocol-agnostic.
- Use “Protocol Mapping” to associate media-type keywords with a selected protocol (HTTP, FTP, WS, CoAP, etc.)

- http://g.mamund.com/class-sked



Protocol Mapping

This media type is designed to work with multiple Internet protocols. To accompish this goal, and [link] elements to define the client-server interactions. These elements each have ar the type of interaction to execute. It is these values which can be used as a guide when selec successfully execute the intended interaction.

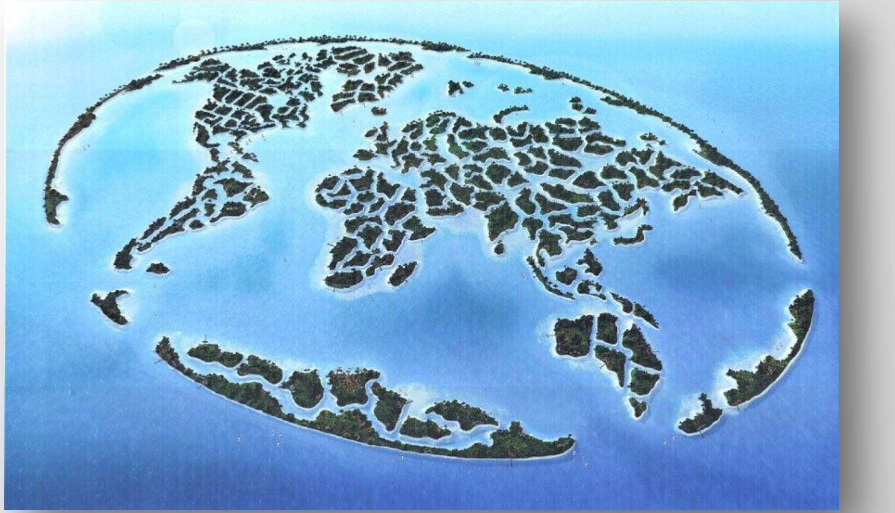
HTTP

Below is the list of [action] values defined in this specification along with information on h and [link] elements of compose a valid Class Scheduling request for HTTP.

add	use the [data] child elements to compose a POST request to the [href] urlencoded media type.
update	use the [data] child elements to compose a PUT request to the [href] u urlencoded media type.
remove	use the [href] to execute a DELETE request.
read	use the [href] to execute a GET request.

Near Term – Focusing on networks

- Most implementations are stand-alone, one-off models.
- We treat the Web as a sea filled with islands, each one only barely aware of the others.

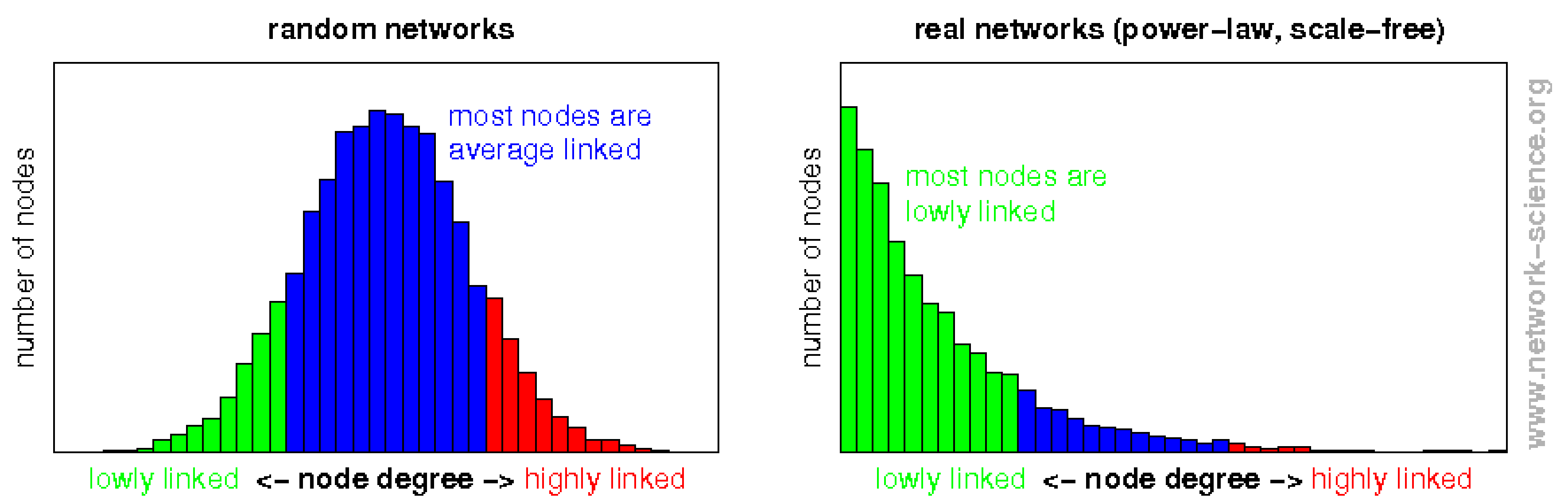


“The WWW is fundamentally a distributed hypermedia application.”

Richard Taylor (2010)

Near Term – Focusing on networks

- The Web, biology, & social communities exhibit properties of a “scale-free” network
- *Barabási-Albert model for “preferential attachment” (1999)*



Near Term

- Lower entropy in messages
- Reduce protocol dependence
- Treat the network as the application

And there are some hard things, too.

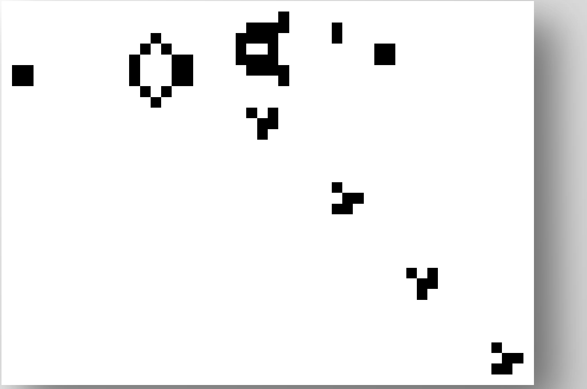
Futures

No more central control, adaptation through variation, competing for resources

Futures – No more central control

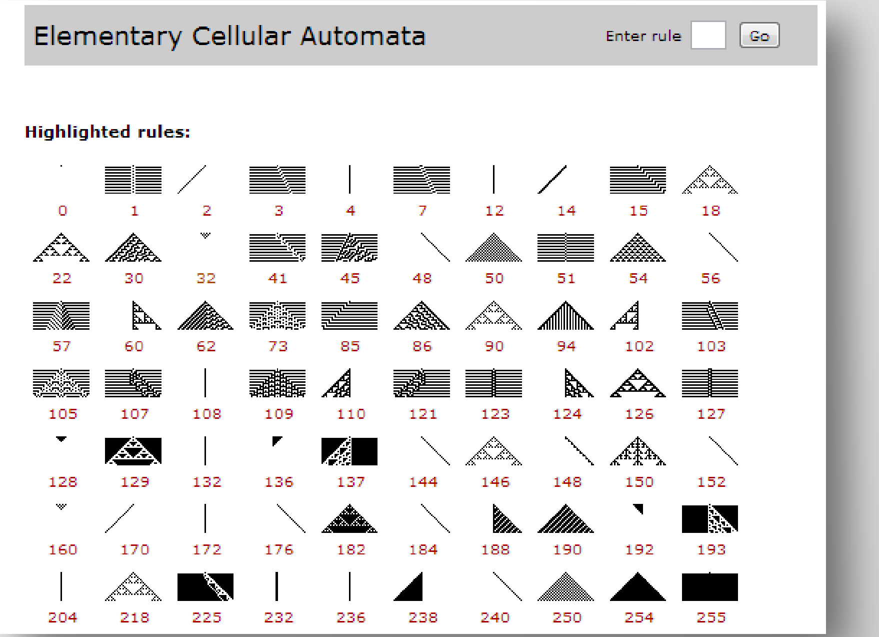
- If the WWW is the application, where is the CPU? The storage? The program?

- Cellular Automata (Ulam & Von Neumann, 1940s)
- Conway’s Game of Life (1970s)



Futures – No more central control

- Cellular automata are *discrete*, *abstract computational systems*
- In cellular automata information appears as statistical probabilities.
- See Wolfram’s Atlas http://atlas.wolfram.com/01/01/

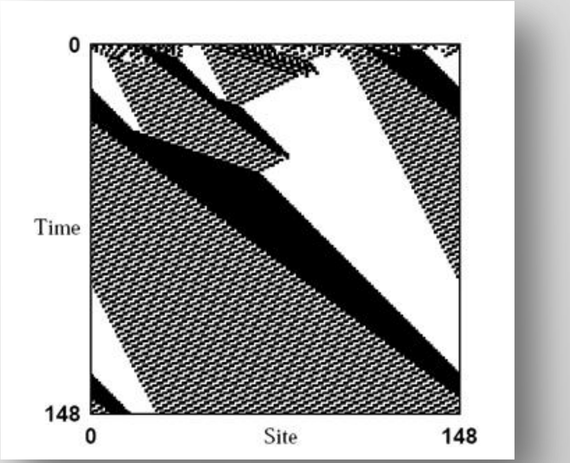


Elementary Cellular Automata

Highlighted rules:

Futures – No more central control

- Basic principles for automata
  - Information takes the form of statistics and patterns across the system
  - Information is communicated via sampling
  - There exists some level of random behavior
  - Rely on fine-grained architecture, large numbers of simple elements.



Futures – No more central control

- In “Future Web” we will create discrete, abstract programs and they will interact across the network.
- *“What gets done on the ‘net stays on the ‘net.”*

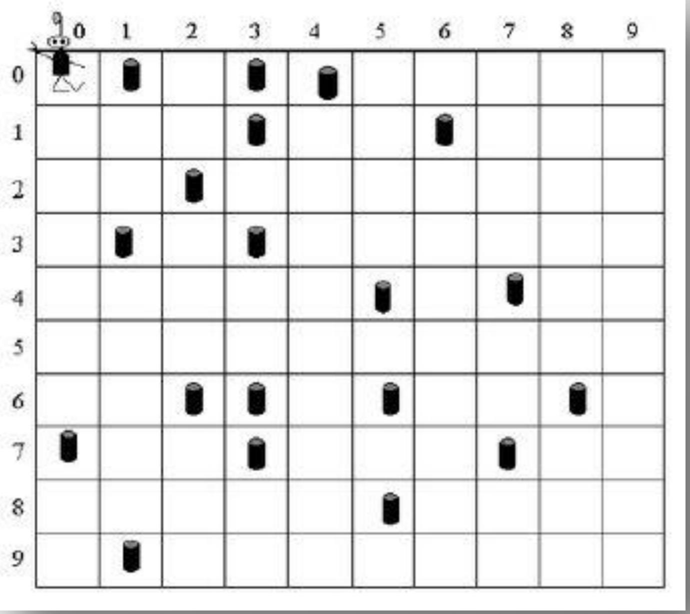
Photograph by Manuel Presti

Harmonious Flight

The ability of animal groups—such as this flock of starlings—to shift shape as one, even when they have no leader, reflects the genius of collective behavior—something scientists are now tapping to solve human problems.

Futures – Adaptation via variation

- Machines will need to adapt to conditions, learn and pass on traits.
- Learning happens via many passes and survival of the 'most fit' for the task.
- "Robby" and the soda cans
  - Start w/ 1xxx random attempts
  - Score highest 2, splice
  - Add random mutation
  - Repeat

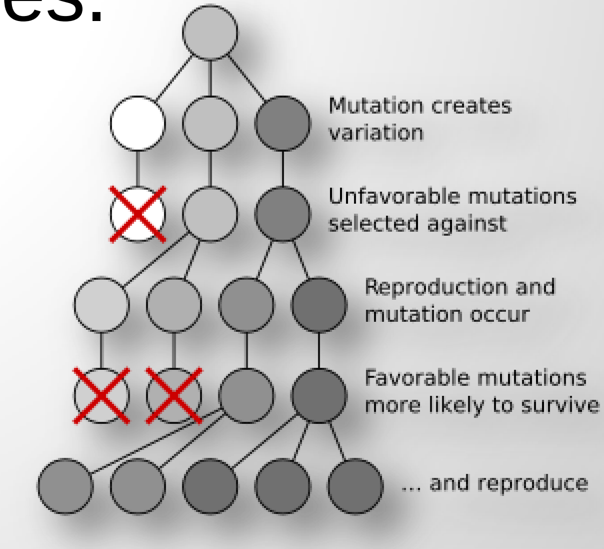


- http://g.mamund.com/robby

How can we model adaptation on the Web?

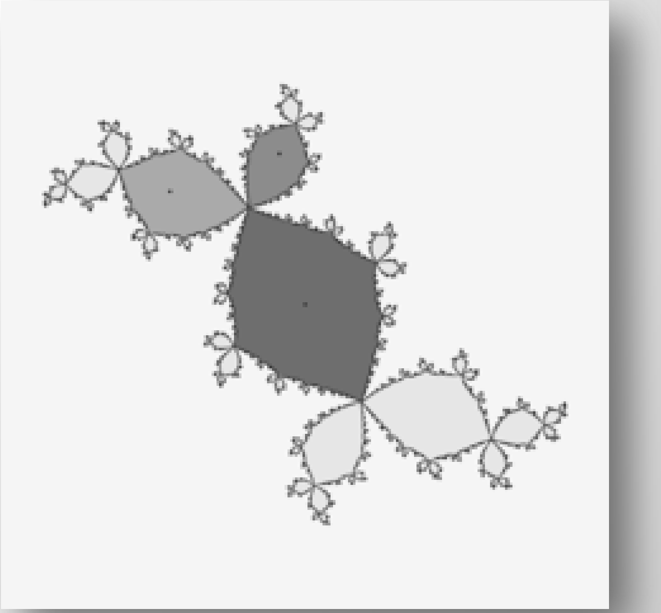
Futures – Competing for resources

- With Robby – there is a “score-keeper” for the entire system.
- On the Web there is no score-keeper.
- In living systems, ‘scoring’ is done through competing for limited resources.


Mutation creates variation
Unfavorable mutations selected against
Reproduction and mutation occur
Favorable mutations more likely to survive
... and reproduce

Futures – Competing for resources

- In “Future Web” programs may compete for scarce resources such as memory, storage, cycles.
- RBNs (Random Boolean Networks) offer a way to “keep score” without central control. (Kauffman, 1969).
- Uses attractors
  - Fixed
  - Oscillating
  - Random



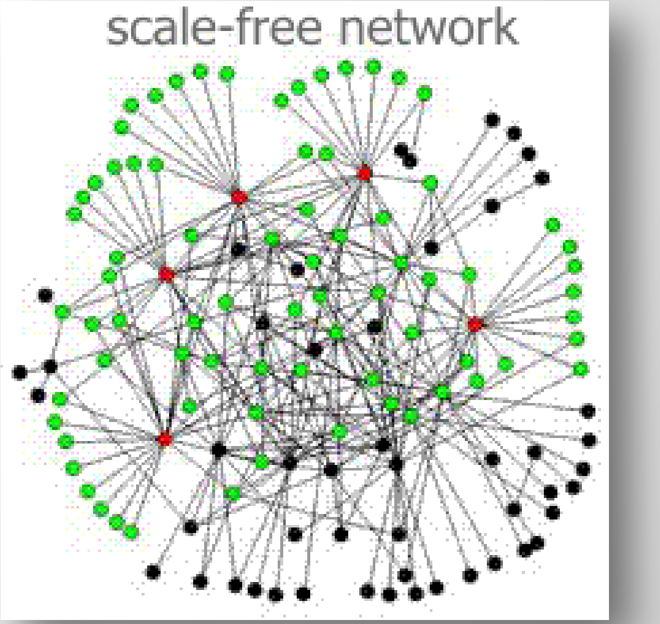
How can we model competition on the Web?

Summary

Time to head back toward shore…

Summary

- Information theory, complex biological systems, hypermedia and the Web all share some similar properties

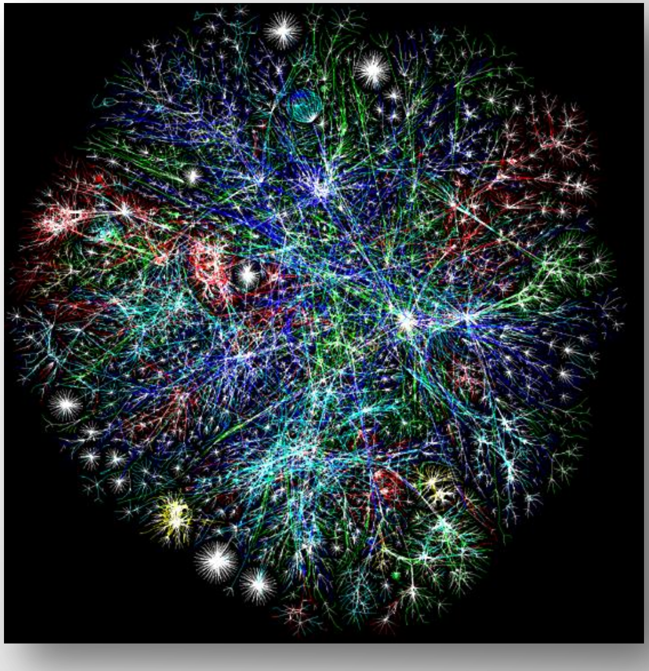


Summary

- However, our current efforts ignore these features and contain a high degree of entropy, coupling, and lack interdependence.

Summary

- We can start today by creating low-entropy machine-oriented messages, decouple from network protocols, and treat the network as a single application space.



Summary

- In the future we'll need to give up central control, we'll build discrete automata, and we'll create a network where variation and competition are possible.

However, in order to accomplish all that…

We must be willing to lose sight of the shore.

Autonomous Agents on the Web:

Beyond Linking and Meaning

Mike Amundsen
Principal API Architect
Layer 7 Technologies
@mamund

SALAD2013

Image Credits (1)

- http://th04.deviantart.net/fs71/PRE/f/2012/082/1/a/sunrise_at_sea_by_rumorvolat-d4too36.jpg
- http://www.pynchon.pomona.edu/entropy/demon.gif
- http://www.scholarpedia.org/w/images/thumb/e/e2/Timesarrowfig2.jpg/400px-Timesarrowfig2.jpg
- http://upload.wikimedia.org/wikipedia/en/e/e5/Shannonmouse.PNG
- http://upload.wikimedia.org/wikipedia/commons/thumb/7/7b/Lego_Turing_Machine.jpg/220px-Lego_Turing_Machine.jpg
- http://notyourmomsfom.files.wordpress.com/2013/04/godel.jpg
- http://en.wikipedia.org/wiki/File:Nobili_Pesavento_2reps.png
- http://www.ics.uci.edu/~fielding/pubs/dissertation/rest_derivation.gif
- http://jasss.soc.surrey.ac.uk/13/2/reviews/3.gif
- http://www.let.leidenuniv.nl/history/ivh/www_proposal.gif
- http://www.biostat.wisc.edu/~cdewey/mercator/cliques.png
- http://upload.wikimedia.org/wikipedia/commons/thumb/6/6e/Splicing_overview.jpg/220px-Splicing_overview.jpg
- http://www.holon.se/folke/kurs/Distans/Ekofys/fysbas/LOT/exergy_tube.jpg
- http://ic.pics.livejournal.com/rose_griffes/11821574/185555/185555_600.jpg
- http://gigaom2.files.wordpress.com/2012/06/iphoneapps.jpg
- http://rubytrain.files.wordpress.com/2011/12/http.jpg

Image Credits (2)

- http://en.wikipedia.org/wiki/File:Kelvin-1200-scale1000.jpg
- http://shamangene.com/BLOG/wp-content/uploads/2012/09/mayan_calendar1.jpg
- http://shamangene.com/BLOG/wp-content/uploads/2012/09/mayan_calendar1.jpg
- http://www.artdes.monash.edu.au/design/assets/design_courses_communication.jpg
- http://images.fineartamerica.com/images-medium-large/hear-no-evil--see-no-evil--speak-no-evil-william-meemken.jpg
- http://etutorials.org/shared/images/tutorials/tutorial_37/02fig03.gif
- http://walyou.com/wp-content/uploads//2011/01/sewer-maintenance-machine.jpg
- http://blog.luxuryproperty.com/wp-content/uploads/image/dubai/The-World-islands-Dubai.jpg
- http://www.ics.uci.edu/~taylor/pictures/Richard%20Taylor.jpg
- http://www.network-science.org/fig_complex_networks_powerlaw_scalefree_node_degree_distribution_large.png
- http://upload.wikimedia.org/wikipedia/commons/e/e5/Gospers_glider_gun.gif
- http://atlas.wolfram.com/01/01/
- http://manwithoutqualities.files.wordpress.com/2010/02/swarming.jpg
- http://plato.stanford.edu/entries/cellular-automata/fig5.jpg
- http://www.styryx.com/img/itech/program/ai/genetic/robby/assignment.jpg
- http://www.historyrv.com/blog/wp-content/uploads/2011/12/natural_selection.png

Image Credits (3)

- http://steve.files.wordpress.com/2006/03/Matrix%20tut%202.jpg
- https://upload.wikimedia.org/wikipedia/commons/thumb/2/25/Julia_immediate_basin_1_3.png/240px-Julia_immediate_basin_1_3.png
- http://www.dichotomistic.com/images/scale-free.gif
- http://www.askamathematician.com/wp-content/uploads/2011/12/pip6_0747.jpg.jpg
- http://www.webnotwar.ca/wp-content/uploads/2011/02/semantic-web.png
- http://images4.alphacoders.com/262/262524.jpg