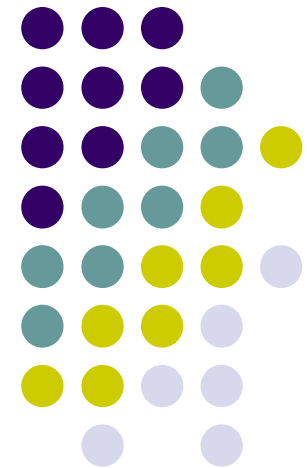# Intelligent Systems for the Future

What About the Internet?
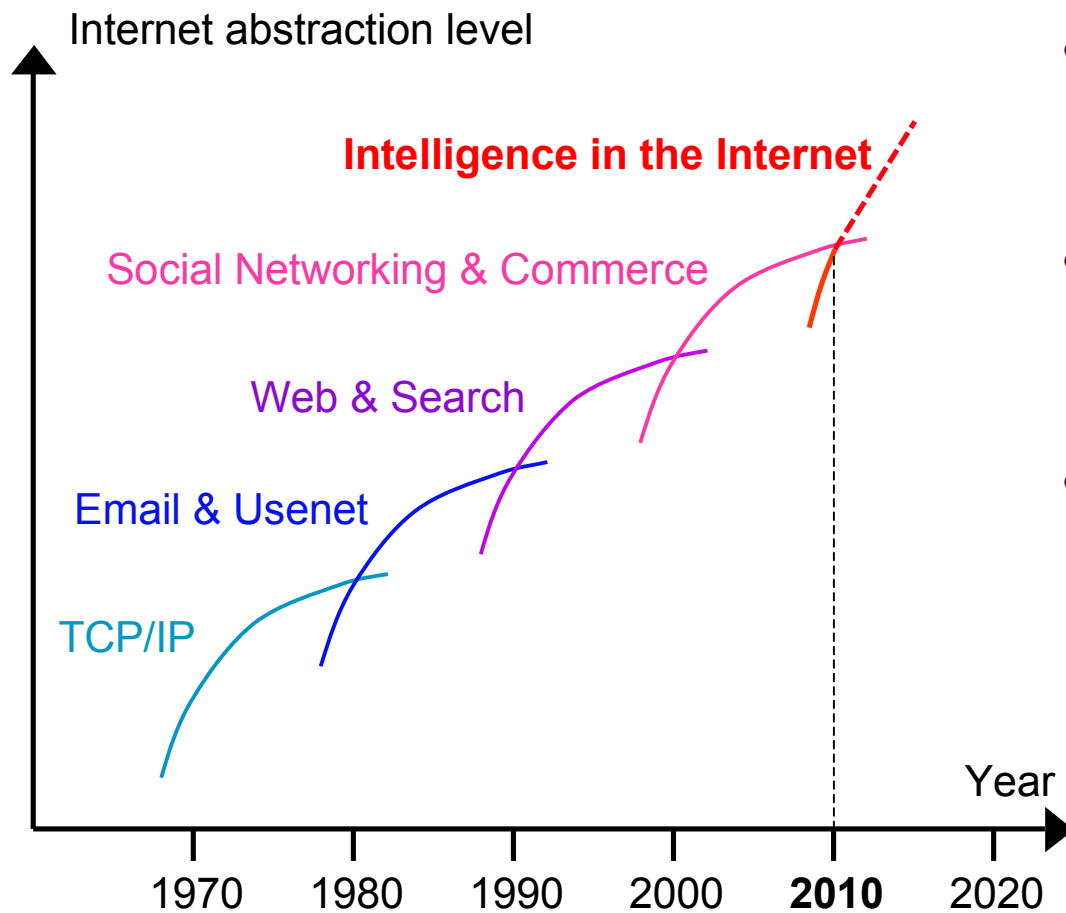
Feb. 8, 2011

Peter Van Roy

ICTEAM Institute

Université catholique de Louvain

Louvain-la-Neuve, Belgium

History of the Internet…

Internet abstraction level

Intelligence in the Internet

Social Networking & Commerce

Web & Search

Email & Usenet

TCP/IP

Year

1970   1980   1990   2000   **2010**   2020

- The Internet has gone through four revolutions since its inception
  - Old timers like me saw most of them (I started using it in 1983)
- But the past revolutions do not do much complex computation
  - Mostly transmission and storage
  - Computations on single machines
- A fifth revolution is now starting…
  - Combine the power of many machines: computing power plus complex algorithms
  - This enables a new kind of ability that will transform the Internet yet again
  - Let's see what's happening…
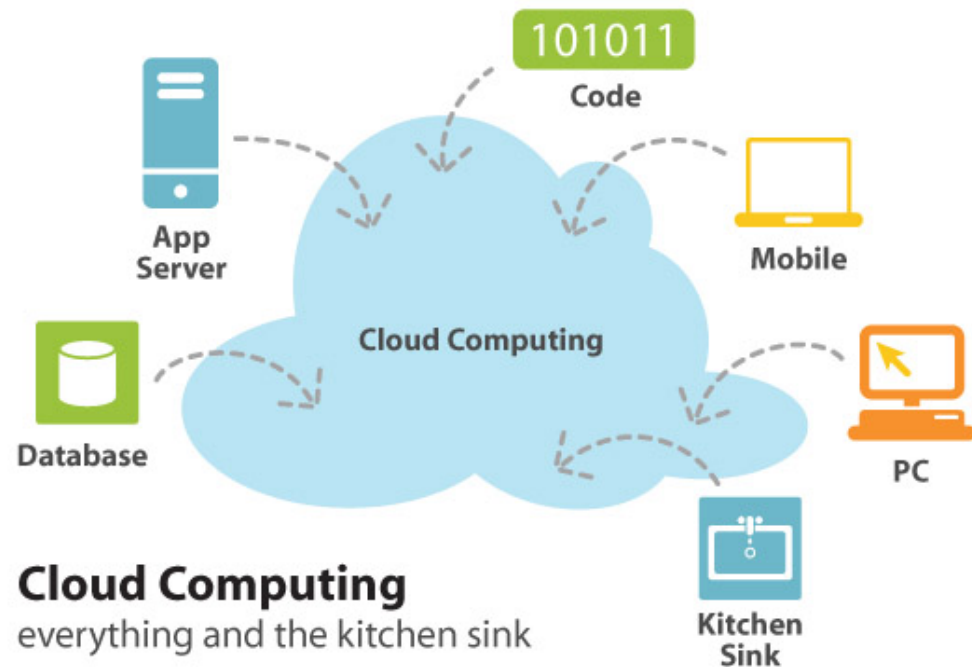
2

© 2011 Peter Van Roy

Intelligence in the Internet

- Intelligence requires enormous computing resources and highly complex algorithms
  - Much more than what can be done on one node
- We now have both!
  - **Internet** (800,000,000 nodes, but dispersed) and **clouds** (>100,000 per cloud, **concentrated**)
  - **Large-scale machine learning** (computing: Map/Reduce, storage: "NoSQL" key/value stores)
- We are just starting to see the applications…
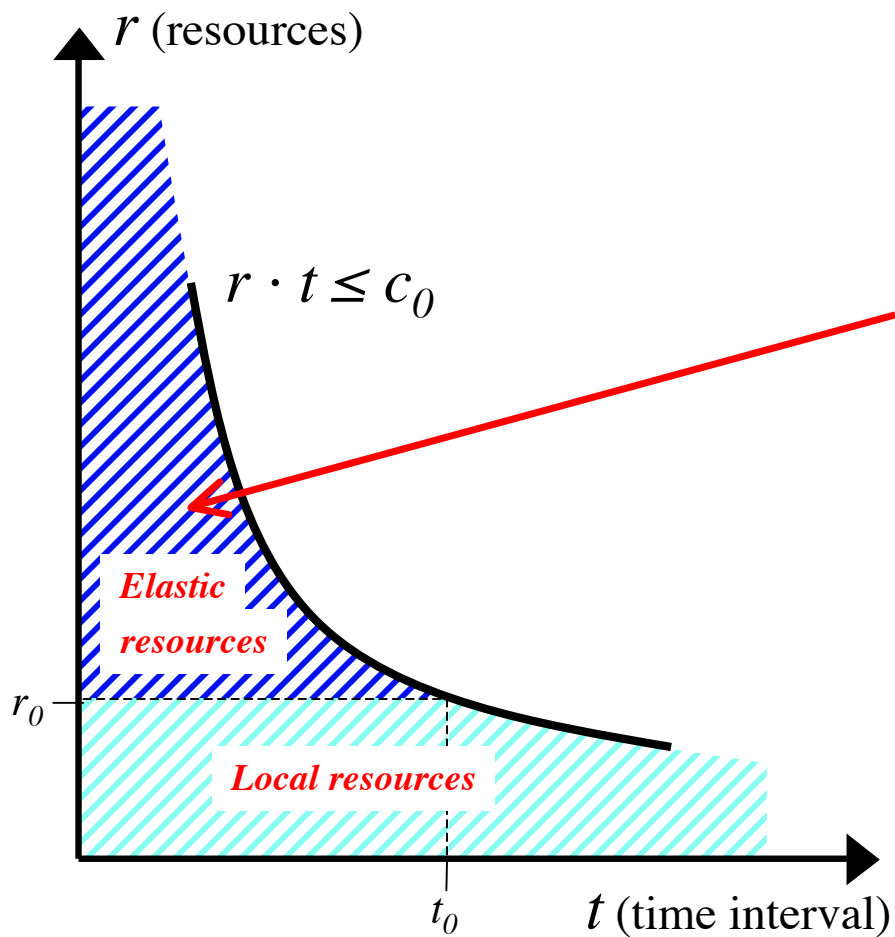 - Search, recommendation, translation, recognition

Clouds are the First Key: Much More Than Meets the Eye!

- **Cloud computing** is a form of **client-server** where the "server" is a *dynamically scalable network* of loosely coupled heterogeneous nodes that are *owned by a single institution*

- It allows enterprises to offload their computing infrastructure

- It gives mobile devices an easy way to manage data

- Is that all that cloud computing offers?
  - No! This is just the tip of the iceberg!
  - Cloud computing is the beginning of a much more profound change

Cloud Computing
everything and the kitchen sink

Clouds are Elastic!

r (resources)

$$r \cdot t \le c_0$$

Elastic resources

r_0

Local resources

t_0 t (time interval)

- Elasticity is the ability to ramp up resources quickly to meet demand
  - Like electric power distribution
- With elastic clouds the enormous **dark blue area** becomes available
- Applications that need enormous resources for short times can get them for low cost!
  - Like electric power distribution, pay only for the volume (cost is product of time and number of machines)
  - This is exactly what intelligent applications need!

5

© 2011 Peter Van Roy

Machine Learning is the Second Key

- Machine learning is the discipline that studies how to program computers to evolve behaviors based on example data or past experience

- Machine learning can solve complex problems that we cannot solve in any other way
    - It has many successes in practical applications both big and small, e.g., speech recognition, computer vision (face and handwriting, etc.), social prediction (epidemics, economics, retail, etc.), robot control (drones, cars, etc.), data mining, aiding natural sciences (biology, astronomy, neurology, etc.)
    - It is a major force on the Internet in big companies (Google, Amazon, Netflix, Facebook, etc.) as well as in startups (e.g., RecordedFuture)

- Machine learning will (eventually) transform programming!
    - Programmers will not work on raw data any more; instead they will build machine learning systems

6

© 2011 Peter Van Roy

Space of Intelligent Applications

XL — Tomorrow's applications

Real-time audio language translation

Learning/setup phase
elastic resource requirements
(learning time constraints)

L

Advanced applications

Real-time expert guidance

Creative problem solving

Intelligent augmented reality

Weather forecasting

Recorded Future

Standard applications

M

Google Search
Google Translate
Recommendation sys.
Speech recognition
Skype connection
Social networks
Media translation

Wolfram Alpha
Image recognition

Championship chess (Deep Fritz)

IBM Jeopardy (Watson)

S

BitTorrent
WIMP GUI
Microsoft Office

Computer algebra
Peer-to-peer CDN
Google Earth
JIT Compiler

MMORPG
Role-playing games
Chess programs

One-shot   S   M   L   XL

One-way stream

Conversation   Interactivity (learning + query)

Query/use phase
elastic resource requirements
(response time constraints)

© 2011 Peter Van Roy

Some Intelligent Applications

- Real-time audio language translation
  - Google is already working on this (announced Feb. 2010)!
  - Full media interchangeability (text, audio, image, video)
- Knowledge extraction from raw data
  - A huge amount of raw data already exists in digital form: 1.2×10^{21} bytes (2010)
  - Learning algorithms based on large corpora, inferencing and canonical forms
- Expert guidance (a form of augmented reality)
  - Guiding humans interactively in real time to perform expert tasks
  - For example, anyone can become an expert car mechanic
- Creative problem solving (tamed brute force search)
  - Combining information to provide useful solutions to human-specified problems
  - The exponential search is tamed by learning algorithms
- Continuous fluid interaction
  - No detours through WIMP GUIs; direct interaction with detailed immersive reality
  - Not programmed, but learned by example and user feedback

8

Executive Summary

- Intelligence requires enormous computing resources and highly complex algorithms (combining many machines)
  - We now have both: elastic clouds and machine learning
- Elastic clouds are key
  - Clouds can be used for much more than just enterprise computing and mobile data storage
- Large-scale machine learning is key
  - Uses scalable computing (map/reduce) and scalable storage (key/value stores) which are already available today (especially as open source!)
- Applications (*you* will build them – the "garage era" is back!)
  - Today: search, recommendation, translation, recognition
  - Tomorrow (just a few examples): real-time audio language translation, knowledge extraction, creative problem solving, expert guidance in augmented reality

Some Local People…

- ## Machine learning

  - Pierre Dupont, Michel Verleysen, Marco Saerens, …

- ## Cloud and peer-to-peer computing

  - Peter Van Roy, Olivier Bonaventure

- ## Tamed brute force search

  - Yves Deville, Peter Van Roy (constraint programming)

- ## Human-computer interfaces

  - Jean Vanderdonckt

- ## Many related disciplines

  - Many people in the ICTEAM institute