NSF EPSCoR 21st National Conference 2009

Sustainability, Innovation, and EPSCoR

October 19–20, 2009

Cyberinfrastucture and Computing:
Developing the Metal And the MeNtal Tools

Deborah L. Crawford, Ph.D.
Deputy, Directorate for Computer and Information Science and Engineering

We’ll restore science to its rightful place, and wield technology’s wonders…. [President Obama, Inaugural Address, January 20, 2009]

Science is more essential for our prosperity, our security, our health, our environment, and our quality of life than it has ever been before. And if there was ever a day that reminded us of our shared stake in science and research, it’s today. [President Obama, speech at the National Academy of Sciences (NAS), April 27, 2009]

21st Century Science & Engineering

Traditional

- Discipline-based
- Local focus
- Single sector activity
- Building-block formal education
- Research vs. education

Emerging

- Topic-based
- Global reach
- Robust partnerships
- Holistic lifelong learning
- Integration of research & education
- Expanding role of IT – democratizing change agent

Gales of Creative Destruction

– Schumpeter

- On Innovation
  - equilibrium destruction
  - innovators as social leaders
- New for Old
  - technical
  - sociological
- Two-wave Cycle
  - development
  - exploitation

Opportunities & Challenges

- Broadening Participation
- Community-Building
- Institutional & Infrastructural Ecology
  - Technological change more rapid than institutional change
- Seamless Integration of New and Old
  - Balancing upgrades of existing and creation of new resources
  - Legacy data/models
- Providing Sustainable Support

Developing The Metal Tools

- Computing: PDAs, PCs, Workstations, Supercomputers
- Networking: Wireless, Wired, Last Mile, Campus, National Backbone

Developing The MeNtal Tools

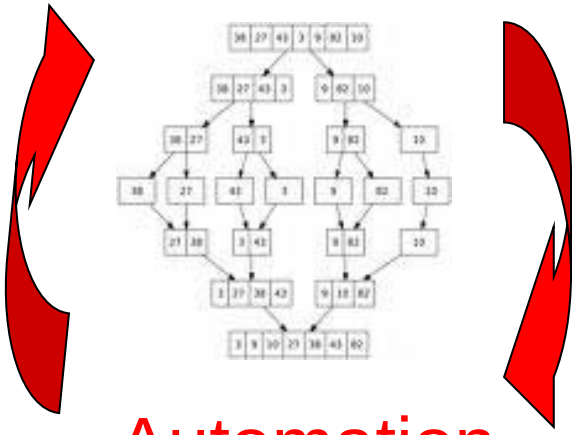
Vision: Computational thinking will be a fundamental skill used by everyone in the world by the middle of the 21st Century

Just like reading, writing, and arithmetic
Incestuous: Computing and computers will enable the spread of computational thinking.
In research: scientists, engineers, …, historians, artists
In education: K-12 students and teachers, undergrads, …

J.M. Wing, "Computational Thinking," *CACM* Viewpoint, March 2006, pp. 33-35.
Paper off http://www.cs.cmu.edu/~wing/

Computing is the Automation of Abstractions

Abstractions



Automation

1. Machine
2. Human
3. Human + Machine
4. Networks of 1, 2, or 3

Computational Thinking is the <u>process of abstraction</u>

- choosing the right abstractions
- operating in terms of multiple layers of abstraction simultaneously
- defining the relationships the between layers

Measures of a “Good” Abstraction in Computational Thinking

- Efficiency
  - How fast?
  - How much space?
  - How much power?
- Correctness
  - Does it do the right thing?
    - Does the program compute the right answer?
  - Does it do anything?
    - Does the program eventually produce an answer?
- -ilities
  - Usability
  - Modifiability
  - Maintainability
  - Simplicity and elegance
  - Cost
  - …

Being More Specific

- Focus on Classes of Abstractions/Concepts
  - Complexity: computability, intractability, undecidability,
  - Algorithms: space/time performance, approximation, randomization, heuristics, optimization
  - Data: data structures
  - Abstract machines: automata, state machines
  - Architecture/Design: decomposition/composition, modularity, layers of abstraction
  - Linguistic: syntax, semantics, grammars
  - Reasoning: correctness, logics, invariants, types, verification, debugging, local vs. global
  - Control: recursion, iteration, conditional, nondeterminism, parallelism, distribution
  - Communication: synchronous/asynchronous, broadcast/P2P, client-server, shared memory/message-passing
  - Physical world constraints: fault-tolerance, reliability, power
  - etc.

- Not
  - Computer literacy, i.e., how to use Word and Excel or even Google
  - Computer programming, i.e., beyond Java Programming 101
  - Potpourri of concepts (I hope)

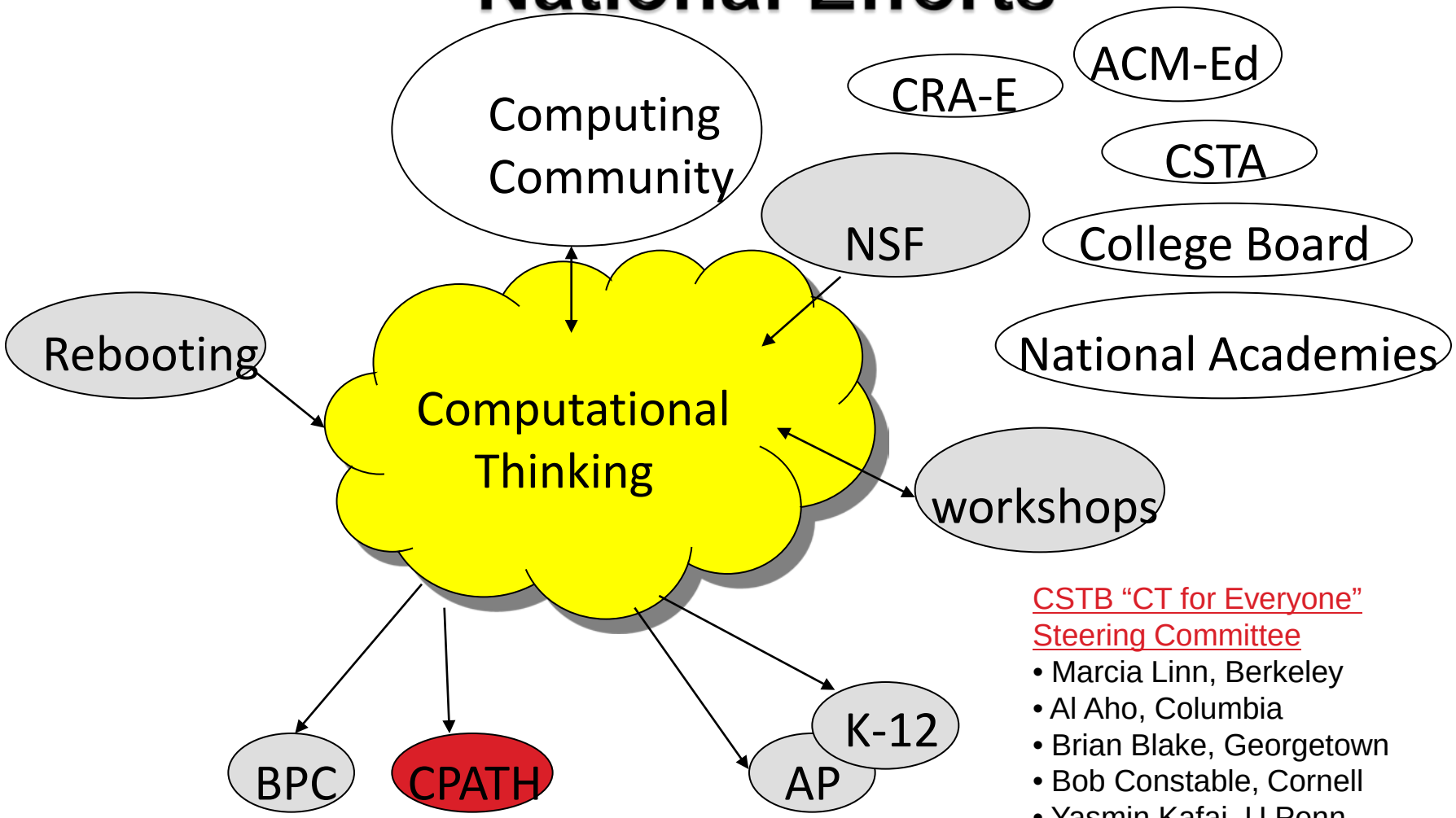
CPATH – Pathways to Revitalized Undergraduate Education in Computing

- Develop Computational Thinking competencies in the future workforce
- Infuse Computational Thinking learning opportunities into core computing and other disciplines
- Demonstrate Computational Thinking education models replicable across institutions

Deadline: April 22, 2010

www.nsf.gov/pubs/2009/nsf09528/nsf09528.html

Computational Thinking in Education: National Efforts



CSTB "CT for Everyone" Steering Committee

- Marcia Linn, Berkeley
- Al Aho, Columbia
- Brian Blake, Georgetown
- Bob Constable, Cornell
- Yasmin Kafai, U Penn
- Janet Kolodner, Georgia Tech
- Larry Snyder, U Washington
- Uri Wilensky, Northwestern

Education Implications for K-12

Question and Challenge for the Academic Community

What is an effective way of learning (teaching) computational thinking by (to) K-12?

What concepts can students (educators) best learn (teach) when?
What is our analogy to numbers in K, algebra in 7, and calculus in 12?

We uniquely also should ask how best to integrate The Computer with teaching the concepts.

Computer scientists are now working with educators and cognitive learning scientists to address these questions.

Cyber-Enabled Discovery and Innovation (CDI)

Computational Thinking for Science and Engineering

- Paradigm shift
  - Not just computing's metal tools (transistors and wires) but also our mental tools (abstractions and methods)
- It's about partnerships and transformative research.
  - To innovate in/innovatively use *computational thinking*;
  - To advance more than one science/engineering discipline.
- Investments by all directorates and offices
  - FY08: $48M, 1800 Letters of Intent, 1300 Preliminary Proposals, 200 Full Proposals, 36 Awards
  - **FY09: $63M+, 830 Preliminary Proposals, 283 Full Proposals, 53+ Awards**

Range of Disciplines in CDI Awards

- Aerospace engineering
- Astrophysics and cosmology
- Atmospheric sciences
- Biochemistry
- Biomaterials
- Biophysics
- Chemical engineering
- Civil engineering
- Communications science and engineering
- Computer science
- Cosmology
- Ecosystems
- Genomics
- Geosciences
- Linguistics
- Materials engineering
- Mathematics
- Mechanical engineering
- Molecular biology
- Nanocomputing
- Neuroscience
- Proteomics
- Robotics
- Social sciences
- Statistics
- Statistical physics
- Sustainability
- …

… advances via Computational Thinking

Data……. Data……. Data…….. Data

Imagine scientific progress if . . .

To get a book, you must find and ask the author

Books are “corrected”; title and author don’t change

Titles are changed with no link to the previous title

Cited articles have often moved, changed, or perished and can’t be found

The organization of text varies arbitrarily among books with no table of contents

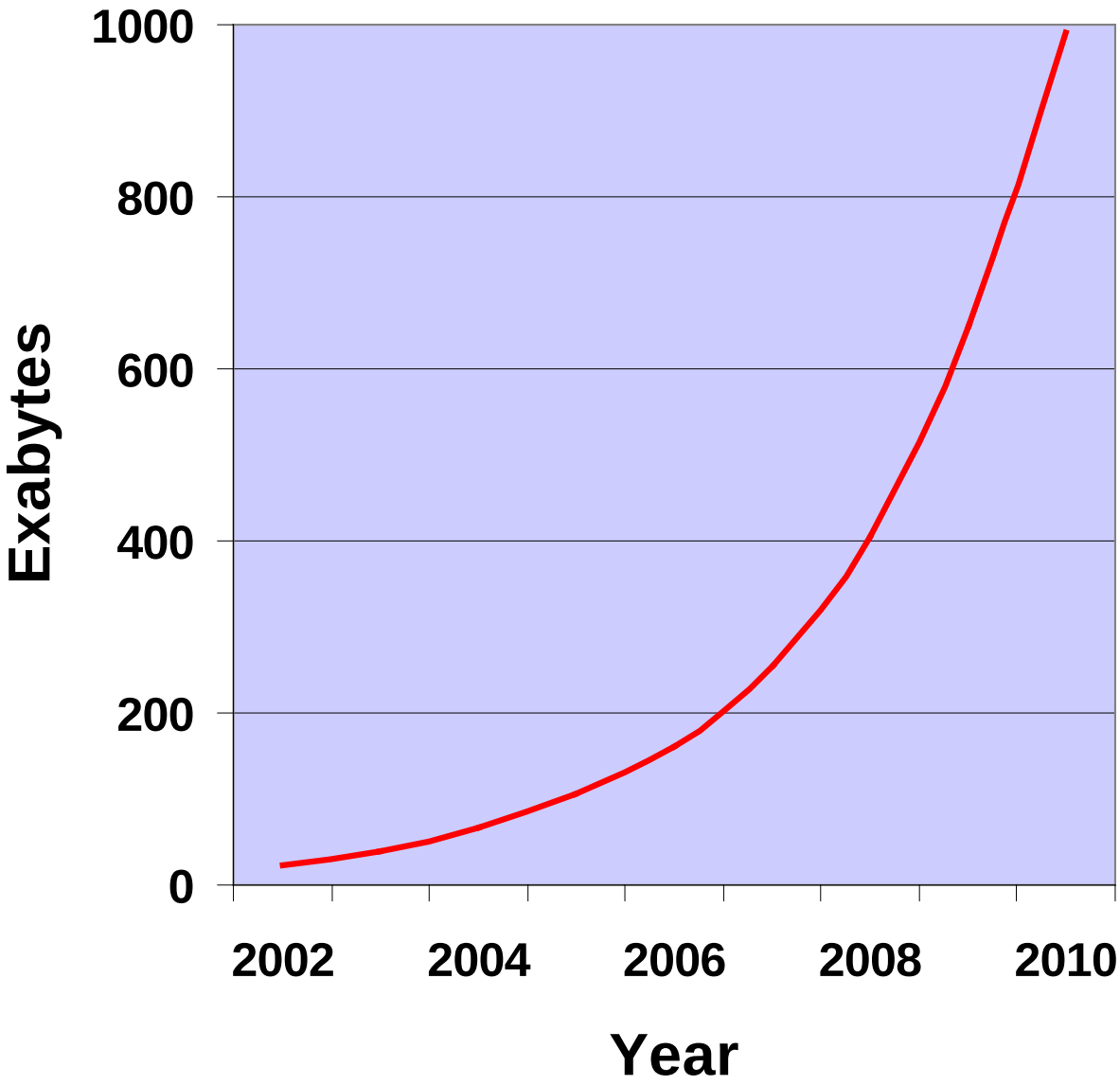
Periodically books are translated to new languages; original is destroyed

For articles and books in the print media, this is fiction

For scientific data, this is often fact

Source: Derived from comments by Gary King, Harvard University

Estimated Annual Digital Information Totals



Sources:

•The Expanding Digital Universe, March 2007, IDC White Paper sponsored by EMC Corporation, www.emc.com/about/destination/digital_universe/

How Much Information? 2003, Peter Lyman and Hal Varian, Berkeley School of Information, http://www2.sims.berkeley.edu/research/projects/how-much-info-2003/

LSST

Large Synoptic Survey Telescope

Digital data are any information that can be stored in digital form and accessed electronically

Examples include numeric and textual information, images, video and audio streams, algorithms, software, operating systems, models and simulations, sensor streams, synthetic data, 3- and 4-dimensional renderings (including virtual environments), etc.

Vision

Science and engineering digital data are routinely deposited in a well-documented form, are regularly and easily consulted and analyzed by specialists and non-specialists alike, are openly accessible while suitably protected, and are reliably preserved.

NSF Cyberinfrastructure Vision for 21st Century Discovery

Finding Your Voice

What should be saved, for how long, and who decides

What means can be used to assess the value of data whose most important use may lie years in the future for unforeseen purposes?

For preservation of software (including models and simulations): When is emulation adequate and when is preservation of the original hardware and operating systems required?

What information (metadata) about data processing, filtering, transformations, workflows, and other manipulations should be saved and how can it be linked reliably to the original data?

What physical media are the best choice for digital archiving?

DataNet Program (OCI/CISE)

- Catalyze development of a system of science and engineering data collections that is open, extensible and evolvable

- Develop new methods, management structures and technologies to manage the diversity, size, and complexity of current and future data sets and data streams

Each DataNet awardee:

- Combines expertise in library and archival sciences; computer, computational and information sciences; cyberinfrastructure; and domain sciences and engineering
- Develops models for economic and technological sustainability over multiple decades
- Engages at the frontiers of science and engineering research and education
- Works cooperatively and in coordination to create a functional data network with revolutionary new capabilities for information access, use, and integration without regard to conventional barriers such as data type and format, discipline or subject area, and time and place/institution.

The Bottom Line

- The Winds of Change are a-Blowing
- Metal and Mental Tools Critical to Participation
- Computational Thinking competencies essential to engage effectively in brave new world
- Your Participation in Defining the Future is Imperative

Revolutionizing Science and Engineering

"[Science is] a series of peaceful interludes punctuated by intellectually violent revolutions . . .[in which] . . . one conceptual world view is replaced by another."

--Thomas Kuhn

From *The Structure of Scientific Revolutions*