

# Métodos Quantitativos para Ciência da Computação Experimental

-Bloco #1-

---

Virgílio A. F. Almeida  
Março 2008



Departamento de Ciência da Computação  
Universidade Federal de Minas Gerais

# Sobre a Organização do Curso

- Cada aluno deve seguir rigorosamente o programa de leituras, que está no programa do curso. Deve ainda acompanhar o material teórico do livro texto.
- Nota Final: participação nas aulas (10%) projeto (30%) e provas (60%)
- Projeto do Curso
  - Enfoque sistemático, com decisões de projeto explicitamente justificadas dentro da metodologia científica.
- 2 provas (30% cada) em sala com problemas sobre a teoria
- Aulas: 50 minutos de apresentação do material teórico e o restante da aula será de discussão com os alunos.
  - Seleção de um voluntário por aula sobre o tema da aula e/ou artigo marcado para aula.
- Material do curso: [www.dcc.ufmg.br/~virgilio](http://www.dcc.ufmg.br/~virgilio)

# Programa: tópicos do curso

- O que é Ciência?
- O que é o Método Científico?
- O que é Ciência da Computação?
- Ciência da Computação Experimental e o Método Científico
- Monitoração
- Revisão dos Conceitos de Probabilidade, Estatística e Sumarização de Dados Medidos
- Comparação de Sistemas usando dados de amostragem e intervalos de confiança
- Modelos de Regressão Linear Simples e outros modelos
- Projeto Experimental (Design of Experimentais)
- Análise de Componentes Principais (PCA)
- Apresentação Gráfica de Dados Experimentais
- Análise de Estudos de Casos em todos os tópicos

# O que é Ciência?

- Quais as visões de vocês sobre o significado de ciência?

# O que é Ciência?

- Ciência é o estudo sistemático das propriedades do mundo físico, através de medições e experimentos replicáveis e do desenvolvimento de teorias universais que são capazes de descrever e prever observações.
- As afirmações em ciência devem ser precisamente formuladas de modo que outros possam testá-las.
- Ciência é:
  - A observação sistemática de eventos naturais e condições com o propósito de descobrir fatos sobre os eventos e formular leis e princípios baseados nos fatos.
  - É um conjunto organizado de conhecimento, derivado de tais observações e que **pode ser verificado ou testado** para novas pesquisas
  - Um ramo específico do conjunto de conhecimentos, como biologia, física, geologia, astronomia e ... (CC???)

# Métodos de Pesquisa Científica

- O Método Científico é o esquema lógico usado por cientistas na procura de respostas às questões colocadas dentro ciência, como também para formular teorias e assegurar os meios para produzi-las (instrumentos, ferramentas, algoritmos)
- Passos Fundamentais do Método Científico:
  1. Formular um problema: colocar uma questão
  2. Hipóteses: sugerir respostas plausíveis a serem testadas
  3. Teste: construir e realizar experimentos para verificar as proposições hipotéticas
  4. Aceitar (ou refutar) provisoriamente as hipóteses e fazer novos testes para verificar as teorias associadas as hipóteses
  5. Analisar e repetir, se for o caso.

# Ciências, Objetos e Métodos

<b>Ciência</b>	<b>Objetos</b>	<b>Método Dominante</b>
	<b>Simples</b>	<b>Análise</b>
Lógica e Mat.	Abstrações, números, proposições, etc	Dedução
Ciências Naturais	Objetos naturais: matéria, organismos vivos, etc	Método dedutivo hipotético
Ciências Sociais	Indivíduos, grupos, sociedade, etc	Método dedutivo hipotético + hermenêutica(*)
Ciências Humanas	Idéias humanas, linguagens, cultura, etc	Hermenêutica
	<b>Complexo</b>	<b>Síntese</b>

(\*) um ramo da filosofia que se debate com a compreensão humana e a interpretação de textos escritos

# Ciência da Computação

- Que tipo de ciência?
- Quais são os objetos?
- Quais são os métodos de pesquisa?



# O que é ciência da computação?

- “Computer Science is the study of computers and what they can do – the inherent powers and limitations of abstract computers, the design and characteristics of real computers, and the innumerable applications of computers to solving problems.” (\*)

\* National Research Council of the National Academies, USA

# Computação, TI e outras áreas do conhecimento



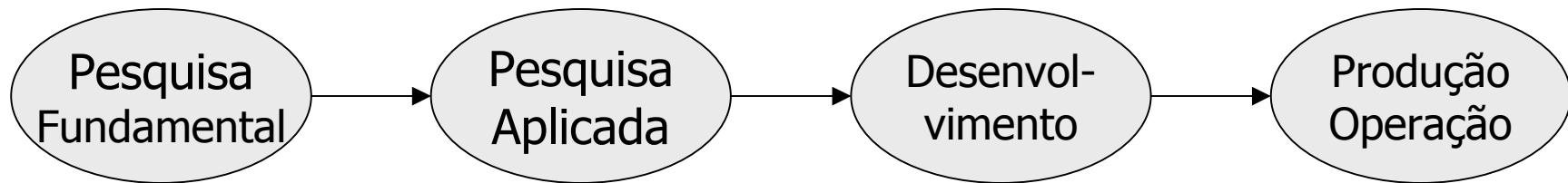
# Núcleo-chave (“core”) da área:

- Algumas disciplinas típicas:
  - Algoritmos e complexidade computacional
  - Sistemas e arquiteturas
  - Linguagens de programação, compiladores e eng. software
  - Inteligência artificial
  - Interfaces e computação gráfica
  - Banco de dados
  - ....

# Natureza e Forma da Pesquisa em Computação

- Classificação da Pesquisa
  - Básica x Aplicada
  - Modelo do Quadrante de Stoke
- Formas da Pesquisa:
  - Teórica x Experimental

# Pesquisa – dicotomia: básica x aplicada



# Modelo de Pesquisa: Matriz de Stoke\*

## Preocupação com os Fundamentos

Preocupação Com Uso

	Não	Sim
Sim	Pesquisa Aplicada Edison	Pesquisa Básica Inspirada pelo Uso Pasteur <b>Computação</b>
Não		Pesquisa Básica Bohr

\***Pasteur's Quadrant, D. Stokes, 1997**

Virgílio Almeida, DCC-UFMG 2006

# Forma da Pesquisa em Computação

- Forma da pesquisa:
  - teórica e experimental
- Ciência da computação teórica – tradições da lógica/matemática
  - Procura entender os limites da computação e a capacidade de paradigmas computacionais, bem como desenvolver soluções gerais para classes de problemas
- Ciência da computação experimental (CCE) é uma disciplina sintética, que envolve a criação e/ou experimentação e análise de artefatos computacionais.
  - Relação entre teoria e experimentação não é necessariamente intensa em ciência da computação, como ocorre em física, química, etc.

# Communication of ACM – Nov 2007

By Dror G. Feitelson, *Guest Editor*

## EXPERIMENTAL COMPUTER SCIENCE

*Only  
experimentation  
can confirm  
or refute  
the relevance  
of theory and  
established  
practices.*

In many fields of science (such as physics and chemistry) theory and experimentation complement and challenge one another. Theory may make predictions that can be confirmed or refuted only through experimentation. Experimentation may find new phenomena that must be explained theoretically. In computer science, however, theory seems to play a more dominant role, typically with little if any direct connection to experimentation. But experimentation is increasingly recognized as necessary and beneficial both as a complement to theory and as an element in constructing systems.

ILLUSTRATION BY PETER HOEV



# Communication of ACM – Nov 2007

Viewpoint | Peter A. Freeman

## Back to Experimentation

Three forward-looking projects depend on experimentation under real-world conditions.



Some of us in the computing field have been around long enough to start to see the adage “History repeats itself” come true in the way we produce major advances in computing. Here, I want to note the recent emergence of serious experimentation on a scale not seen in years<sup>1</sup> and relate it to what some of us participated in when the field was young.

Since the late 1990s, a number of proposals and initial attempts have sought to develop and experiment with new technology under real-world, but nonproduction, conditions, on a scale relative<sup>2</sup> to the desired result not seen since CACM was young. Three such efforts point the way toward a renewed and very valuable trend of experimentation.

The most visible is the Defense Advanced Research

helping us all set our sights on what can be achieved through experimentation at scale.

The second, just starting to do some preliminary prototyping after extensive planning, is the Global Environment for Network Innovations Project ([www.geni.net](http://www.geni.net)) begun in 2004 by the National Science Foundation’s Directorate for Computer & Information Science & Engineering. GENI intends to refocus networking research on new architectures and mechanisms for future networks, not just on developing patches for our current networks. The project’s Web site, which describes GENI and provides pointers to related information, is maintained by the GENI Project Office operated by BBN Technologies under agreement with NSF. GENI will support such research with a large-scale, experimental network that will be the largest experimental piece of “equipment” built solely for computer science research. It is not yet well known outside the computing research commu-

# Ciência da Computação Experimental: o que é?

- Baseado em observações, dados, experimentos;
- Pesquisa experimental deve ser reproduzível!!!
- *Trabalho experimental* deve complementar a pesquisa teórica
  - Teorias podem ter incertezas (ex.: Qual a ordem de grandeza da constante  $k$ ? É um problema difícil?).
  - Teorias podem resultar de observações.
  - Teorias podem ser testadas por observações.
  - De forma oposta, teorias podem guiar a pesquisa experimental.
  - Nem toda pesquisa em ciência da computação pode ser resolvida teoricamente (ex.: CHI, uso malicioso)

# Pesquisa Experimental e artefatos computacionais

- Ciência da computação experimental (CCE) refere-se a criação de artefatos computacionais, que podem ser definidos como a implementação de um ou mais fenômenos computacionais.
- Podem ser extraordinariamente complexos;
- Papel dos artefatos na pesquisa em Ciência Computação:
  - Prova de desempenho (performance)
  - Prova de existência
  - Prova de conceito
- CCE envolve a criação e/ou experimentação com artefatos computacionais.
- Proximidade da CCE e engenharias, no que se refere à criação de artefatos (sistemas de computação, aviões, geradores de energia, etc)

# Prova de Performance

- Um artefato funcionando como “proof-of-performance” provê:
- Um aparato para medição direta e experimentação
  - O artefato existe ou pode ser construído
  - Os resultados da prova são geralmente quantitativos
  - O artefato mais frequente nas pesquisas experimentais de computação e engenharia
- Exemplos: processador com arquitetura RISC, um novo algoritmo de busca

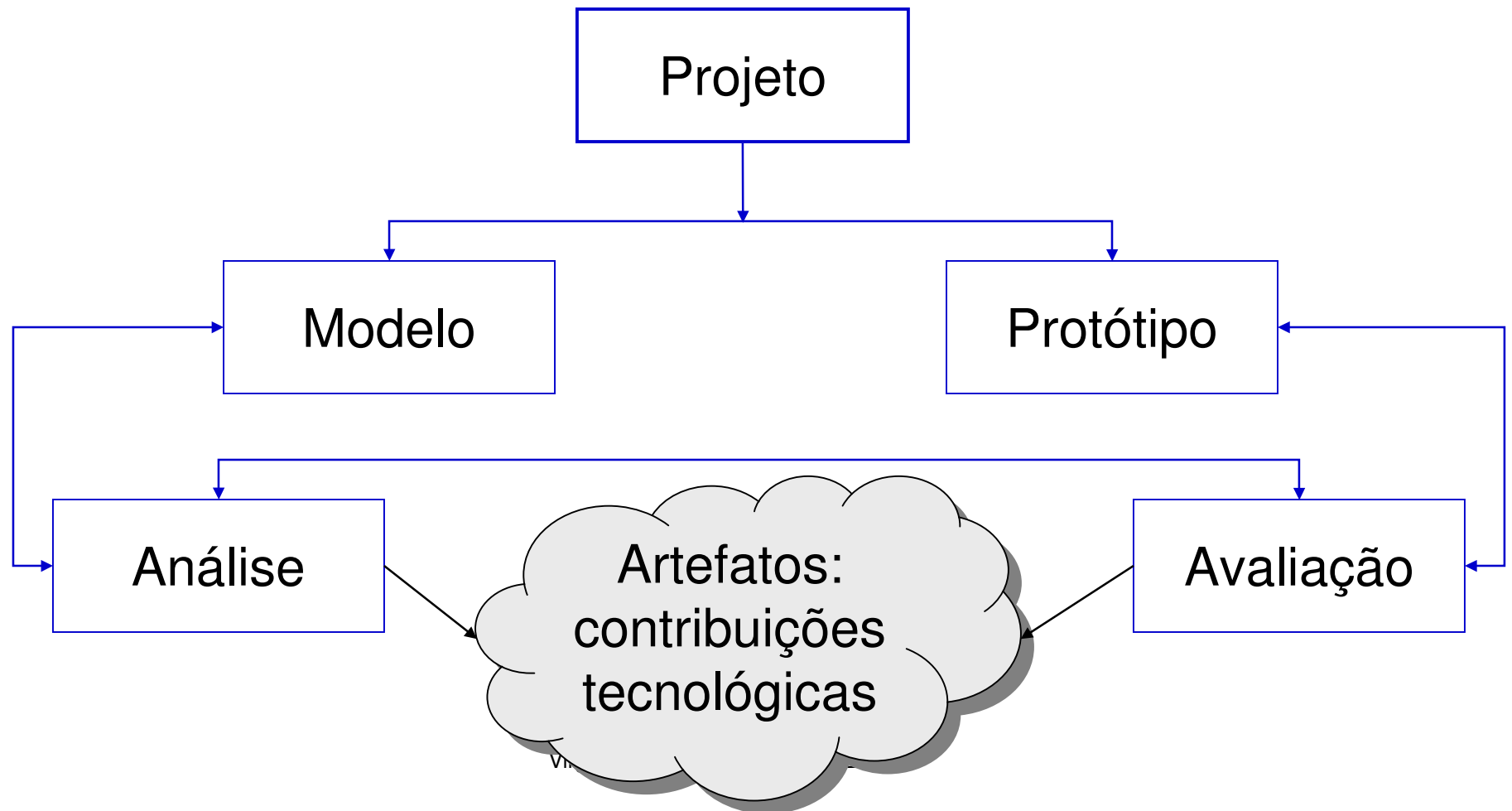
# Prova de Conceito

- Um artefato funcionando como “proof-of-concept” demonstra pelo seu comportamento que um sistema complexo de componentes pode executar um determinado conjunto de atividades.
- O comportamento não poderia ser inferido ou determinado a partir de uma argumentação lógica ou uma argumentação baseada em alguns princípios básicos.
- O sistema em operação, isto é o artefato, é uma “testemunha” de prova que os conceitos estão corretos, pelo menos para aquela configuração.
- Exemplo: computadores experimentais implementando sistema de memória virtual ou uma CDN (content distribution network)

# Prova de Existência

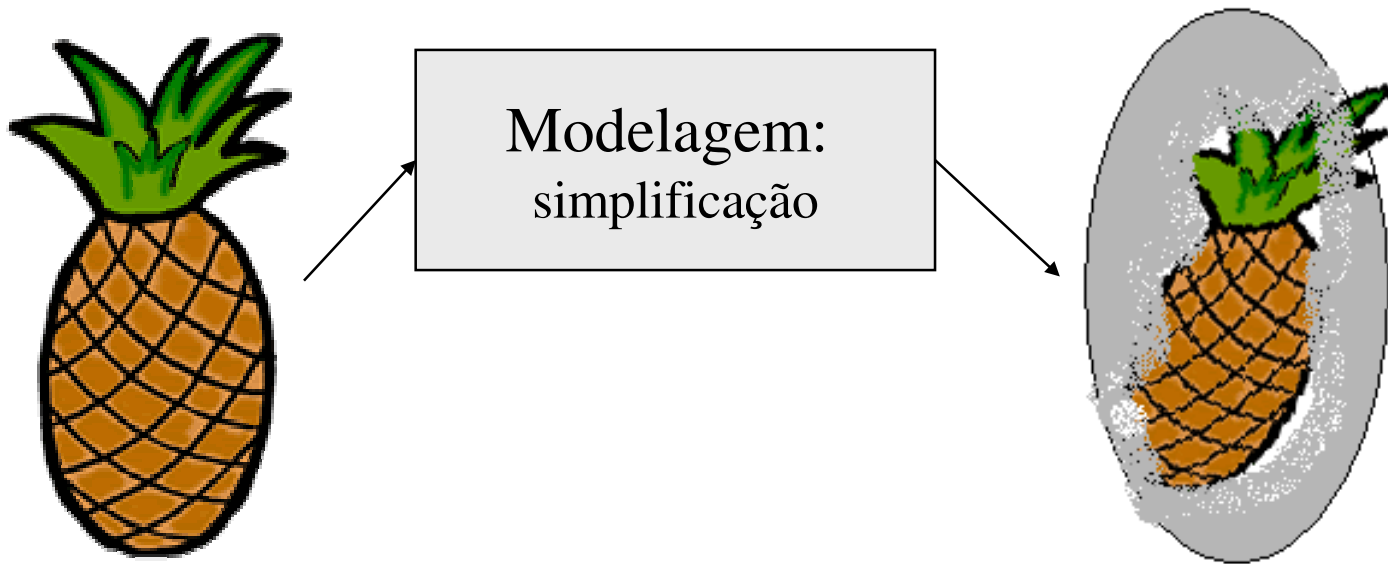
- Um artefato funcionando com o papel de “proof-of-existence” apresenta a essência de um fenômeno inteiramente novo.
- Exemplo: mouse, iphone, outros???
- Menos comum das três provas.

# Ciência da Computação Experimental



# Método Científico em Ciência da Computação

- Por que modelos?
- Modelos → abstrações, simplificações...





# Experimentação em Sistemas Computacionais

Por que?:

“It doesn’t matter how beautiful your theory is, it doesn’t matter how smart you are – if it doesn’t agree with the experiment, it’s wrong.”

R. Feynman, físico - Prêmio Nobel  
1965

# Alguns tipos de “papers” em Ciência da Computação

- Três tipos de artigos que descrevem a implementação de um algoritmo:
  - Artigo de aplicação  
*“Aqui está um bom algoritmo para a solução X”*
  - Artigo de “marketing”  
*“Aqui está um algoritmo novo e interessante”*
  - Artigo Experimental  
*“Aqui está como o algoritmo comporta-se com dados reais”*
- O método experimental deve se aplicar aos três tipos de artigos.

# Discussão

- Apresentação dos alunos. Cada aluno deverá indicar na segunda aula, qual artigo da lista de leituras irá apresentar.
- Quais são interesses de pesquisa?
  - Quais conferências representam os problemas de interesse de sua comunidade de pesquisa?
- Você tem algum projeto de pesquisa em andamento?
- Quais as razões que o levaram a fazer este curso?  
(ex: *Meu orientador obrigou-me!*)
- Você tem alguma “expertise” anterior no tema que pode ser útil para a turma?

# Exemplo de questões

- Considere que seu trabalho é comparar o desempenho de dois algoritmos (A e B) de computação gráfica, que usam métodos diferentes para geração de faces humanas realísticas.
- São sistemas complexos cuja execução leva tempos longos para geração das faces. O sistema A foi testado 8 vezes e o sistema B apenas 5, onde em cada experimento utilizou-se o mesmo padrão de resultado a obter.
- Os tempos de teste dos algoritmos estão na tabela a seguir. Com base nesses resultados, pede-se que se determine qual algoritmo teve melhor desempenho?

# Tempos de execução

<b>Experimento</b>	<b>Algoritmo A (seg)</b>	<b>Algoritmo B (seg)</b>
1	1011	894
2	998	963
3	1113	1098
4	1008	982
5	1100	1046
6	1039	-
7	1003	-
8	1098	-

# Por que Ciência da Computação Experimental?

- Artigo básico:
  - W. Tichy in “Should Computer Scientists Experiment More?”
  - [merlot.usc.edu/cs551-f03/papers/Tichy.pdf](http://merlot.usc.edu/cs551-f03/papers/Tichy.pdf)
- Experimentação: Modelo-teste da teoria
  - Feynman: um experimento pode sempre identificar falhas na teoria (ex.: hipótese/suposições violadas pela realidade)

# Should Computer Scientists Experiment More?

---

By Walter Tichy

Presented by Virgilio Almeida

# Outline

- Is Computer Science a Science?
- Why should we experiment?
  - Eight fallacies exposed
- Why substitutes won't work
- Inherent problems with experimentation



# Is Computer Science a Science?

- No, an engineering discipline (F Brooks)
- Yes, much more than synthetic results
  - Study of informacional structures & processes
  - Synthetic results (computers & programs) are models
  - Difference: work with information – neither energy or matter

# Is Computer Science a Science? (P. Denning, CACM 5/2005)

- Science of information processes and their interactions with the world.
- Science or technology (man-made objects)?
- Science means explaining, modeling and predicting phenomena in the world.
- Pure and applied science.
- In a sample of 400 papers before 1995, Tichy found that about 50% (in particular in software engineering) of those proposing models or hypothesis did not test them. In other fields of science the fraction of papers with untested hypothesis was about 10%.

# Science versus Art

- Science
  - Principles
  - Fundamental laws
  - Explanation
  - Discovery
  - Analysis
  - Dissection
- Art
  - Practice
  - Skilled performance
  - Action
  - Invention
  - Synthesis
  - construction

Art refers to useful practices of a field.

# Why should we experiment?

- Theory testing and exploration
  - Theory falsification: **falsifiability** (or **refutability** or **testability**) is the logical possibility that an assertion can be shown false by an observation or a physical experiment.
  - Example: Knight and Leveson experiments (\*)

(\*) The experiment was concerned with the failure probabilities of multi-version programs. Conventional theory predicted that the failure probability of a multi-version program was the product of the failure probabilities of the individual versions. However, Knight and Leveson observed in an experiment that the failure probabilities of real multi-version programs were significantly higher. In essence, the experiment falsified the basic assumption of conventional theory, namely that faults in program versions are statistically independent.

# Eight Fallacies Exposed (#1)

Traditional scientific method isn't applicable

- *Rebuttal: To understand information processes, computer scientists must observe phenomena, formulate explanations, and test them. This is the scientific method.*
- Repeatability

# Repeatability

## **Internet Measurement Conference (IMC) 2008** Sponsored by [ACM SIGCOMM](#) and in cooperation with [USENIX](#) October 20-22, 2008 Vouliagmeni, Greece

The Eighth Internet Measurement Conference is a two and a half day event focusing on Internet measurement and analysis, building on the success of past [IMCs](#). We invite submissions of papers that contribute to our understanding of how to collect or analyze Internet measurements, or give insight into Internet structure and behavior. Examples of relevant topics are:

- Internet traffic analysis
- Internet structure and topology characteristics
- Internet performance measurements
- Measurement-based network management such as traffic engineering
- Inter-domain and intra-domain routing
- Network applications such as WWW, multimedia streaming, and gaming
- Measurements of content distribution, peer-to-peer, overlay, and social networks

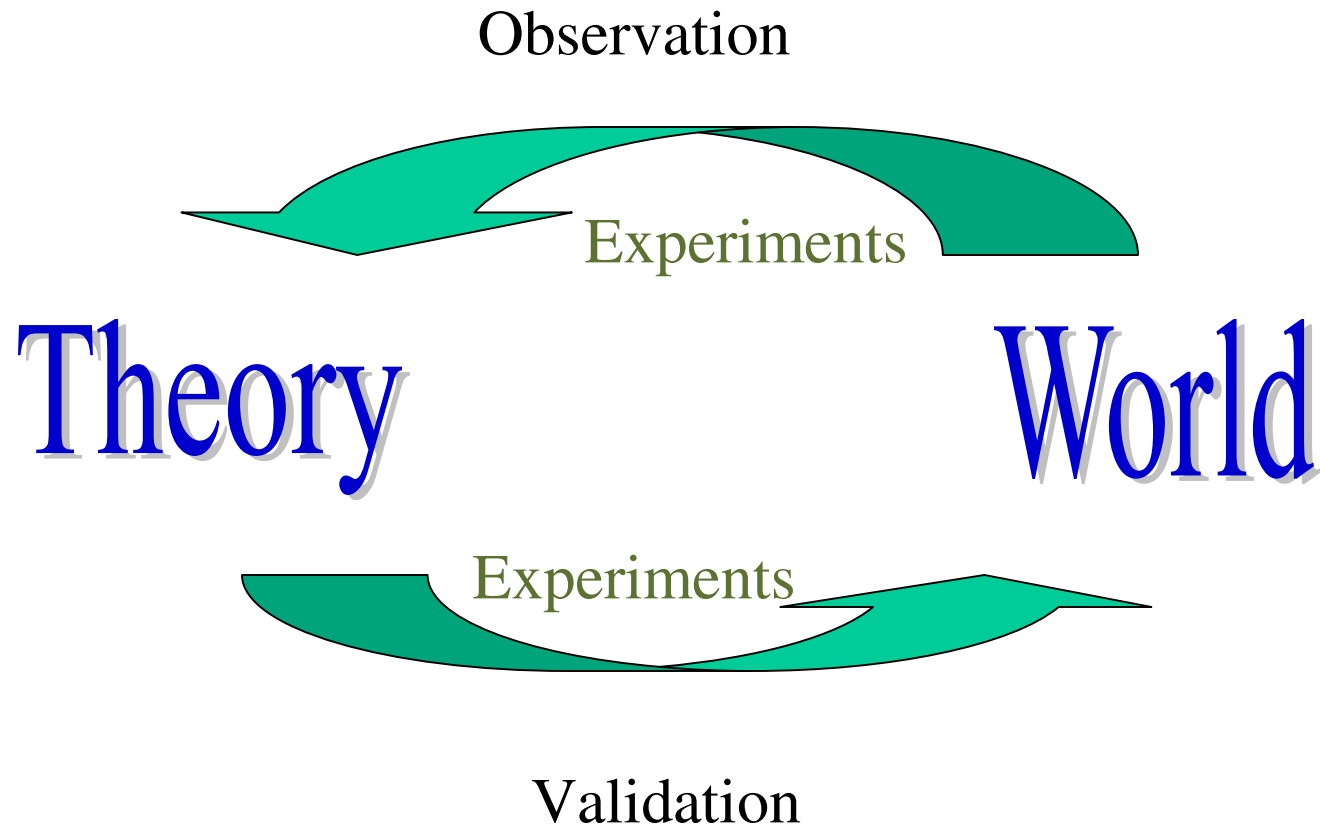
### **Important dates**

- **May 2, 2008: 10PM EDT:** Registration of title and 250-word abstract
- **May 9, 2008: 10PM EDT** *HARD* submission deadline
- **July 18, 2008:** Notification
- **August 27, 2008:** Camera Ready Copy due
- **October 20-22, 2008:** Conference held in Vouliagmeni, Greece

To encourage broader data sharing in the community, the conference will present a best paper award for the top paper that makes its data sets publically available by the time of camera ready. For example, wireless-network data sets may be published through [CRAWDAD](#). Authors that would like their paper to be considered for this award should add a footnote on the first page.

A limited number of travel grants may be available to students who are unable to secure funding from their advisors.

# Eight Fallacies Exposed (#1)



## Eight Fallacies Exposed (#2)

The current level of experimentation is good enough

- *Rebuttal: Relative to other sciences, the data shows that computer scientists validate a smaller percentage of their claims*
- Balancing theory and eng. with experiment
  - Build reliable base & reduce uncertainties (eg: TREC)
  - Leads to new areas of investigation where engr progress is slow
  - Accelerate progress by pruning fruitless approaches



## Eight Fallacies Exposed (#3)

### Experiments cost too much

- *Rebuttal: Meaningful experiments can fit into small budgets; expensive experiments can be worth more than their cost*
- Constrained by cost
  - Plan appropriate research programs
  - Look for affordable experimental techniques
  - Intermediate steps with partial results
- Experiments in the industry
- Experiments in other areas
  - Pharmaceuticals, aeronautics, biology

## Eight Fallacies Exposed (#4)

### Demonstrations will suffice

- *Rebuttal: Demos can provide incentives to study a question further. Too often, however, these demos merely illustrate a potential*
- Proof of concept
- No solid evidence
- Require clear question, experimental apparatus to test the question, data collection, interpretation, sharing of results

## Eight Fallacies Exposed (#5)

There's too much noise in the way

- *Rebuttal: Fortunately, benchmarking can be used to simplify variables and answer questions*
- Benchmarks
  - allow repeatable and objective comparisons
  - aids in identifying promising approaches and discarding poor ones
- Experiments involving humans also repeatable

## Eight Fallacies Exposed (#6)

Experimentation will slow progress

- *Rebuttal: Increasing the ratio of papers with meaningful validation has a good chance of actually accelerating progress*
- Good conceptual papers will continue to be published
- Need to get beyond assertion

## Eight Fallacies Exposed (#7)

### Technology changes too fast

- *Rebuttal: If a question becomes irrelevant quickly, it is too narrowly defined and not worth spending a lot of effort on.*
- Probe for fundamental and not the ephemeral

## Eight Fallacies Exposed (#8)

You'll never get it published

- *Rebuttal: Smaller steps are still worth publishing because they improve our understanding and raise new questions*
- Non-theoretical journals and conferences accept papers on solid experimentation
- Respectable experimentalists articulate how their systems contribute to our knowledge

# Why Substitutes Won't Work

- Feature comparisons vs formulation of models, hypothesis and tests (experimentation)
- Intuition\*\*\* – good starting point but need solid evidence as backup
- Experts – don't always present evidence; maintain healthy skepticism

# Inherent Problems with Experimentation

- Competing Theories
  - Weak reasoning gives way or combines with other theory
  - Rarely produce falsifiable theories
- Unbiased Results – involve managers and other decision-makers



# Discussion

- Do you agree that we should experiment more?
- Examples????
- Are there weaknesses in these arguments?
- How can we empirically collect data and ensure that variables are controlled?
- Why does Computer Science have the word "Science"?