

Computação Experimental

Prof. George Luiz Medeiros Teodoro

teodoro@unb.br

glmteodoro@gmail.com

Objetivo Geral

- Apresentar aos alunos de Ciência da Computação os conceitos e práticas relacionados ao método científico. Com isso, os alunos devem entender a Ciência da Computação como Ciência, e compreender o processo para o estabelecimento de teorias e a aplicação de métodos quantitativos como importantes ferramentas para fundamentar os resultados de uma investigação científica.

Programa

- M1: Ciência e o Método Científico:
 - O que é Ciência;
 - A revolução científica;
 - Tipos de inferência: dedutiva, indutiva;
 - Métodos de pesquisa;
- M2 Técnicas de Revisão de Literatura:
 - Estudo de mapeamento sistemático;
 - Revisão sistemática da literatura;
 - Como fazer a leitura (crítica) de trabalhos científicos;
- M3 Monitoração e Sumarização de Dados Coletados:
 - Tipos de cargas, modelos de cargas;
- M4 Revisão de Probabilidade e Sumarização de Dados Medidos:
 - Álgebra de eventos, espaços amostrais;
 - Probabilidade, probabilidade condicional;
 - Independência de eventos, regra de Bayes, teorema da probabilidade total;
 - Variáveis aleatórias, funções de probabilidade;
 - Média, mediana, moda, variância, desvio padrão, covariância;
 - Distribuições comuns de variáveis aleatórias discretas e contínuas;
 - Sumarização de dados: média aritmética, geométrica, harmônica;

Programa

- M5 Comparação de Sistemas Usando Dados de Amostragem e Intervalos de Confiança:
 - Intervalos de confiança para a média da amostra;
 - As distribuições z e t ;
 - Testes para média zero;
 - Testes para comparação de duas alternativas e de proporções;
 - Testes de hipóteses;
 - Intervalos de confiança de um lado e para proporções;
- M6 Projeto de Experimentos:
 - Terminologia, fatores, erros comuns na experimentação;
 - Tipos de projetos de experimentos: Projetos simples, com fatorial completo, com fatorial fracionado;
 - Projetos fatoriais 2^k ;

OBS: Tanto o programa, quanto diversos materiais utilizados nessa disciplina são baseados no curso de métodos quantitativos da Profa. Jussara e Prof. Virgílio da UFMG e Prof. Jaques da UFCG.

Metodologia

- Aulas expositivas e práticas.

Bibliografia

- **R. Jain. The Art Of Computer Systems Performance Analysis. Wiley, 1991.**
- Raul Sidnei Wazlawick. Metodologia de Pesquisa em Ciência da Computação. Ed. Campus, 2009.
- Peter Dalgaard. Introductory Statistics with R. Springer, January 2004.
- C. Wohlin. Experimentation in Software Engineering: An Introduction. The Kluwer International Series in Software Engineering. Kluwer Academic, 2000.
- Jacques Wainer. Métodos de pesquisa quantitativa e qualitativa para a ciência da Computação. In Tomasz Kowaltowski and Karin Breitman, editors, Atualização em informática. Sociedade Brasileira de Computação, 2007.
- George E. P. Box, William G. Hunter, J. Stuart Hunter, and William G. Hunter. Statistics for Experimenters: An Introduction to Design, Data Analysis, and Model Building. John Wiley & Sons, Second edition, June 2005.

Avaliação

- No plano de ensino.

Informações Gerais

- Prof. Dr. George Luiz Medeiros Teodoro (UnB, 2014-Atual)
 - Graduação em Ciência da Computação– UFMG - 2004
 - Mestre em Ciência da Computação - UFMG- 2006
 - Doutorado em Ciência da Computação, sistemas paralelos – UFMG 2010
 - Pós-Doutorado na University of Maryland at College Park, UMD, Estados Unidos – 2011
 - Pós-Doutorado na Georgia Institute of Technology, Georgia Tech, Estados Unidos - 2013
 - Pós-Doutorado na Emory University, Estados Unidos – 2013
 - Professor na Stony Brook University, 2015 – Atual

O que são Ciência e o Método Científico?

Jacques Sauvé
UFCG

Agradecimento

- Usei algum material de Virgilio Almeida da UFMG, o qual agradeço
- Várias partes deste material vêm de:
 - Métodos Estatísticos Aplicados à Engenharia de Software Experimental, ARAÚJO, BARROS, MURTA, TRAVASSOS, UFRJ

O que queremos responder

- No final, queremos discutir se a ciência da computação é uma ciência "normal"
- Antes de falar da ciência da computação, precisamos entender:
 - O que é a ciência?
 - Que métodos a ciência utiliza para chegar a seus resultados?
 - Como a ciência *realmente* é feita
 - versus a versão simplificada

O que é Ciência?

O que é Ciência? Várias definições

- 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
- 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, ..., **ciência da computação???**

A Pesquisa Científica

- Pesquisa é um processo sistemático de coleta, análise e interpretação de informação (dados) no sentido de aumentar nosso conhecimento de um fenômeno
- Isso parece meio vago, né?
 - Vamos ter que discutir mais a fundo

Características da pesquisa

- A pesquisa inicia com um ***problema***
- A pesquisa requer um ***objetivo*** claro
 - Resolver o problema? Estudar o problema? Caracterizar melhor o problema?
 - Objetivo = "research question"
- Pesquisa requer um ***plano***
 - Como atingir o objetivo?
- A pesquisa quebra um problema grande em ***subproblemas***
 - Divida para conquistar - análise

Exemplo

● Problema

- Estamos insatisfeitos com a recomendação de ofertas para os usuários de sites de compras coletivas

● Objetivo

- Deseja-se fazer um comparativo de eficácia e desempenho entre os algoritmos reconhecidamente mais eficazes em outros contextos de sistemas de recomendação

● Plano

- Escolha de hipóteses; design de experimento; execução e coleta de dados; análise de dados

● Subproblemas

- Se houver muitas hipóteses, cada uma é um subproblema

Tipos de Inferência

Métodos de raciocínio, base para
o método científico

Inferência

- Inferência = processo de chegar a conclusões a partir de premissas
- Tipos de inferência (ou de raciocínio)
 - Inferência dedutiva
 - Inferência indutiva
 - Inferência abdutiva
- Todos os tipos de inferência são utilizados no método científico

Raciocínio Dedutivo

- Inicia com premissas
 - Premissas são dadas como verdadeiras
- Deriva uma conclusão como consequência lógica das premissas
- Exemplo
 - Tulipas são plantas (premissa 1)
 - Plantas produzem energia com fotossíntese (premissa 2)
 - Então tulipas produzem energia com fotossíntese (conclusão)
- A veracidade da conclusão é garantida se as premissas forem verdadeiras

Raciocínio Indutivo

- Inicia com várias observações
 - Experiência sensorial (sentidos)
- Usam-se instâncias ou ocorrências específicas para chegar a conclusões sobre classes de objetos ou eventos
 - De alguns dados para conclusões genéricas
- Sempre requer evidência empírica
- A experiência dá suporte à conclusão mas não há garantia de sua veracidade

Raciocínio Indutivo: Exemplo

- Você vê o sol se levantando pela primeira vez uma certa manhã
 - Hmmm! Interessante!
- Você vê o sol se levantando novamente no dia seguinte
 - Hmmm! Muito interessante!
- No terceiro dia, você vê a mesma coisa e formula uma hipótese: "O sol se levanta todas as manhãs"
- O que vai acontecer no quarto dia?
 - Você tem **certeza**?

Raciocínio Abduutivo

- Como a indução pode ser um método lento, usa-se às vezes o “chute” para acelerar as coisas
 - Isto é Raciocínio Abduutivo
- Exemplo:
 - Se chover, a grama está molhada
 - Se a grama está molhada, então deve ter chovido
- Isso está formalmente errado pois estamos concluindo a partir do consequente
 - Se P então Q
 - Q
 - Então P
- Mas às vezes o chute funciona

O Método Científico

Sugestão de leitura

- Reis, Método Científico

- É simples e muito bom (apesar dos erros de português)

Por que um método especial?

- Como posso acreditar em alguma afirmação?
 - Pela autoridade de uma pessoa, de um grupo?
 - Pela autoridade de um livro sagrado?
 - Pelo senso comum?
 - Pelo processo democrático? (vote!)
- Precisamos de algo melhor

Caraterísticas desejáveis

- Minimizar o viés do pesquisador
- Admitir e corrigir erros (não há dogma)
- Permitir verificação por terceiros de conhecimento



Algo indiscutível, eg. doutrina religiosa

O Método Científico

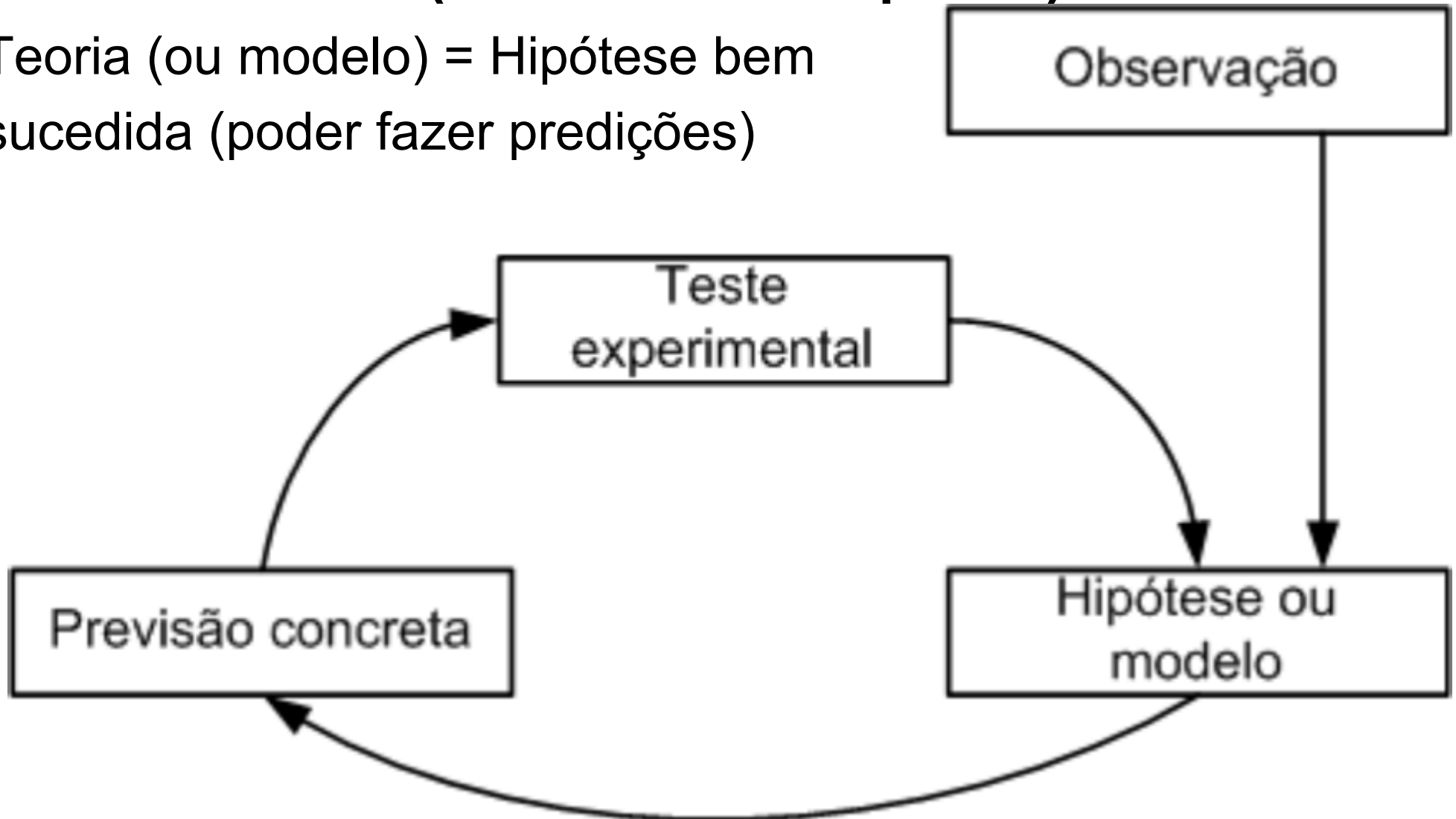
- O método científico é um método pelo qual novo **conhecimento** é obtido a partir de **observações**.
- As observações ...
 - levam a uma sequência de **hipóteses** que ...
 - permitem fazer **predições** que ...
 - podem ser **verificadas experimentalmente**, ...
 - **refutando** ou não as hipóteses
 - As hipóteses bem sucedidas levam a **teorias**
 - as quais podem um dia ser **refutadas**
- Teoria = explicação de um fenômeno
 - = hipótese bem sucedida
- Modelo = teoria de validade limitada
- O processo é cíclico

Exemplo

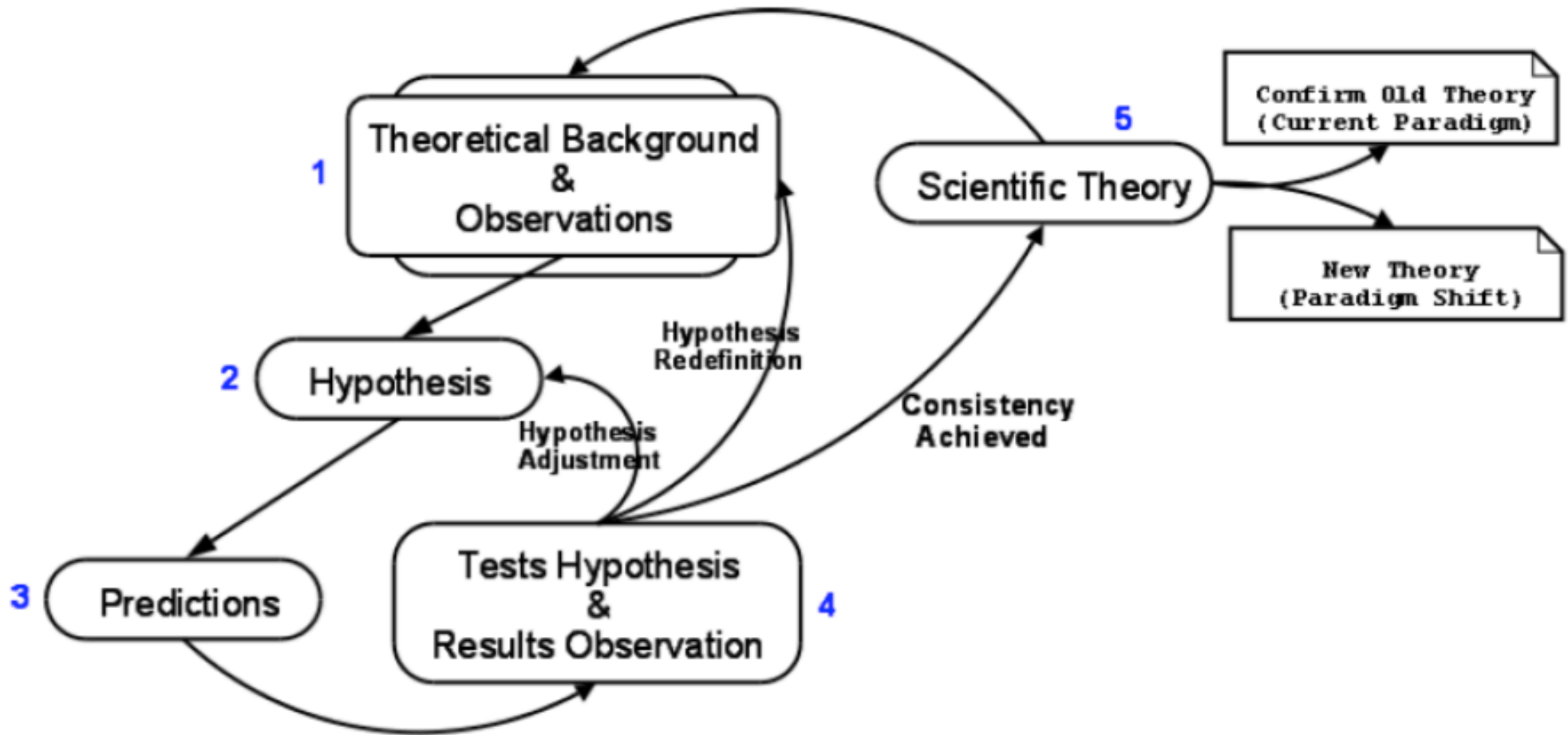
- Imagine um ET chegar à terra e descobrir as regras do futebol assistindo a algumas partidas
 - Ao longo do tempo, ele vai formular hipóteses e refutá-las (ou não), até chegar a algo bem próximo das regras verdadeiras
- A metáfora não é perfeita pois nós podemos interagir com o mundo através de experimentos

O Método Científico (versão simples)

Teoria (ou modelo) = Hipótese bem sucedida (poder fazer previsões)



Uma visão mais completa



Pontos importantes

- Hipóteses devem ser ***testáveis***
- Teorias podem ser ***refutadas*** por experimentos
 - Isso diferencia a ciência da religião
- Teorias não são "ideias sem comprovação"
 - Vejam [aqui](#)
- Experimentos podem ser ***reproduzidos*** por outros para verificar resultados
 - Mantém os pesquisadores honestos

Inferência e o método científico

Inferência e o Método Científico

- Pesquisadores podem desenvolver uma hipótese
 - Usando teoria (raciocínio dedutivo) ou
 - A partir de observações (raciocínio indutivo)
 - No chute, concluindo pelo consequente (raciocínio abdutivo)
- Então, usando raciocínio dedutivo, fazem previsões que serão vistas se a hipótese for verdadeira
- Também usando raciocínio dedutivo, podemos falsear uma hipótese
 - Se fatos X aparecerem, Deduzimos que a teoria está errada
- Usando raciocínio indutivo, eles podem generalizar a partir de dados de uma amostra para inferir características de uma população

O loop indutivo-dedutivo (Box & Hunter)

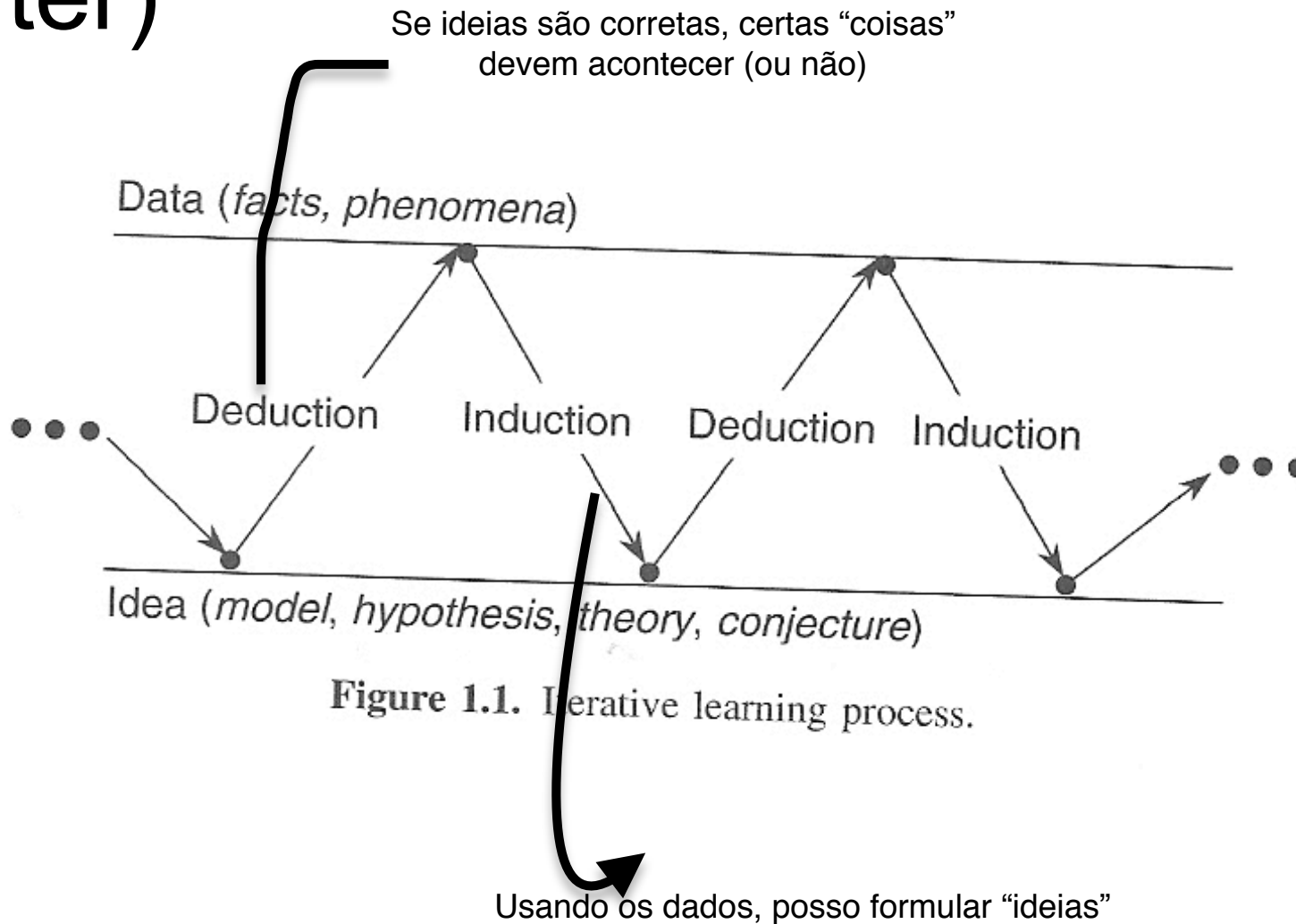


Figure 1.1. Iterative learning process.

Bacon, Descartes,
Popper, Khun

O Método Científico:

Francis Bacon (1561-1626)

- Escreveu Novum Organum (“Novo Método” ou “Novo Instrumento”)
 - Resposta ao “Organon” de Aristóteles sobre lógica e silogismo (inferência)
Inferência dedutiva, que apenas permite retirar informação das premissas, mas não criar conhecimento
- O método Baconiano:
 - É **indutivo**
 - Parte de observações e formula conclusões genéricas *tentativas*
 - O Organum de Aristóteles usa silogismo (conclusão *certa* a partir de premissas)
 - As conclusões iniciais podem estar erradas
 - A “verdade” vem de erros
 - Requer um procedimento planejado para investigar fenômenos naturais
 - É precursor do Método Científico

A Análise:

René Descartes (1596-1650)

- Escreveu o Discurso sobre o Método
- Quatro preceitos (regras de proceder)
 1. Duvide de tudo Penso, logo existo.
 - Exceto: “Cogito, ergo sum”
 - Portanto, analise sem noção preconcebida
 - Só aceita aquilo de que tem *certeza*
 2. Divida as dificuldades em partes menores (análise)
 3. Inicia com objetos mais simples e chega, passo a passo, a idéias mais complexas
 4. A cada passo, enumere tudo de forma a ter certeza de que não esqueceu nada

A Falseabilidade:

Karl Popper (1902-1994)

- Insatisfação com inducionismo
 - Fatos estão sendo usados para explicar qualquer conjunto de fatos (“teorias” de Marx, Freud, ...)
- Introduziu o conceito de **Falseabilidade**
 - Nenhum número de experimentos positivos pode provar uma teoria
 - Tem que haver fatos que levem a declarar a teoria falsa
 - Uma teoria pode ser refutada se um experimento resultar em “falso”
 - Popper: Algo só é científico se pode ser falseado

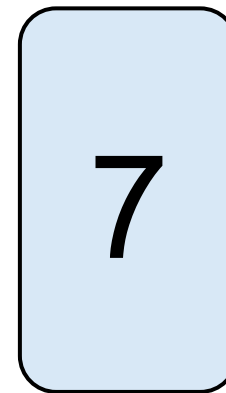
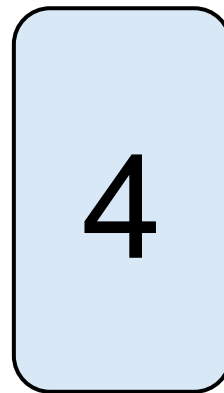
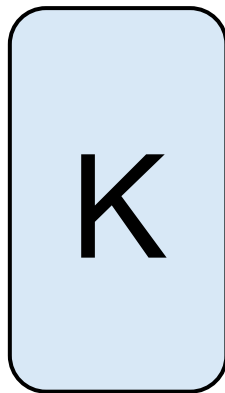
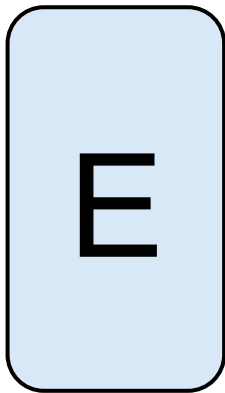


Falseabilidade - Exemplos

- Experimento de Knight e Leveson (1986) falsificou a suposição básica de que falhas em versões diferentes de um programa são independentes
- O tráfego de redes não é Poisson mas auto-similar
 - Artigo de Leland et al., 1994
- Contra-exemplo: Uma teoria só pode existir se houver algum experimento possível que, sendo realizado e dando errado, provará que a teoria está errada
 - Ex. Crítica de teoria de “Multi-universos”:
 - Não há experimentos possível (até agora) para a teoria, então ela não merece sequer ser chamada de “teoria” ou mesmo de “hipótese”

Exemplo: aprenda a falsear

- "Wason card puzzle"
- Cada carta tem um número numa face e uma letra na outra
- Qual é o número mínimo de cartas que devemos virar para verificar a regra: "cartas com vogal de um lado têm número par do outro"?



Resposta

- Duas cartas: E e 7
- A regra é do tipo "se p então q"
- Você deve testar a teoria tentando falseá-la, não tentando confirmá-la
- O falseamento significa testar "se p então não-q"
 - p (vogal) só ocorre com E
 - não-q (ímpar) só ocorre com 7
- Testar a carta 4 não adianta pois, se houver uma consoante do outro lado, isso não invalida a regra, que nada diz sobre consoantes

O "Paradigm Shift":

Thomas Khun (1922-1996)

- Escreveu *The Structure of Scientific Revolutions*
- A ciência não avança com acumulação linear de novo conhecimento
- Há avanços revolucionários (*paradigm shifts*) com transformação abrupta
 - Frequentemente não são baseados em observação
 - Ex.: Einstein usou “thought experiments”
 - Ex.: Dirac previu a anti-matéria apenas através de equações, sem dicas experimentais prévias
- Três estágios do avanço científico
 - Pré-ciência: não há paradigma central
 - Ciência normal: resolve problemas dentro de um paradigma central. Se não conseguir, não é o paradigma que está errado mas o cientista que errou
 - Ciência revolucionária: quando muitas anomalias já ocorreram, uma crise ocorre e um novo paradigma aparece para substituir o antigo

Métodos na pesquisa

Métodos na pesquisa

- A pesquisa usa vários métodos
- Experimentação
 - Controla uma variável e vê o resultado em outras
 - Exemplo: Testamos a vazão do protocolo TCP com vários valores de timeout
- Estudo de caso descritivo
 - Obtém dados de cenários reais
 - Inclui observação e medições de comportamentos
 - Exemplo: Observação do desempenho de uma equipe de programação usando um certo processo de desenvolvimento
 - Não há *controle*

Métodos na pesquisa

● Estudo de caso comparativo

- Para determinar e quantificar relacionamentos entre variáveis usando grupos expostos a tratamentos diferentes
- Exemplo: comparar dois processos de desenvolvimento (RUP e SCRUM) em ambiente real de desenvolvimento

● Modelagem

- Mímica aproximada de sistemas reais para conduzir experimentos e/ou fazer observações
- Exemplo: modelo de desempenho de datacenters

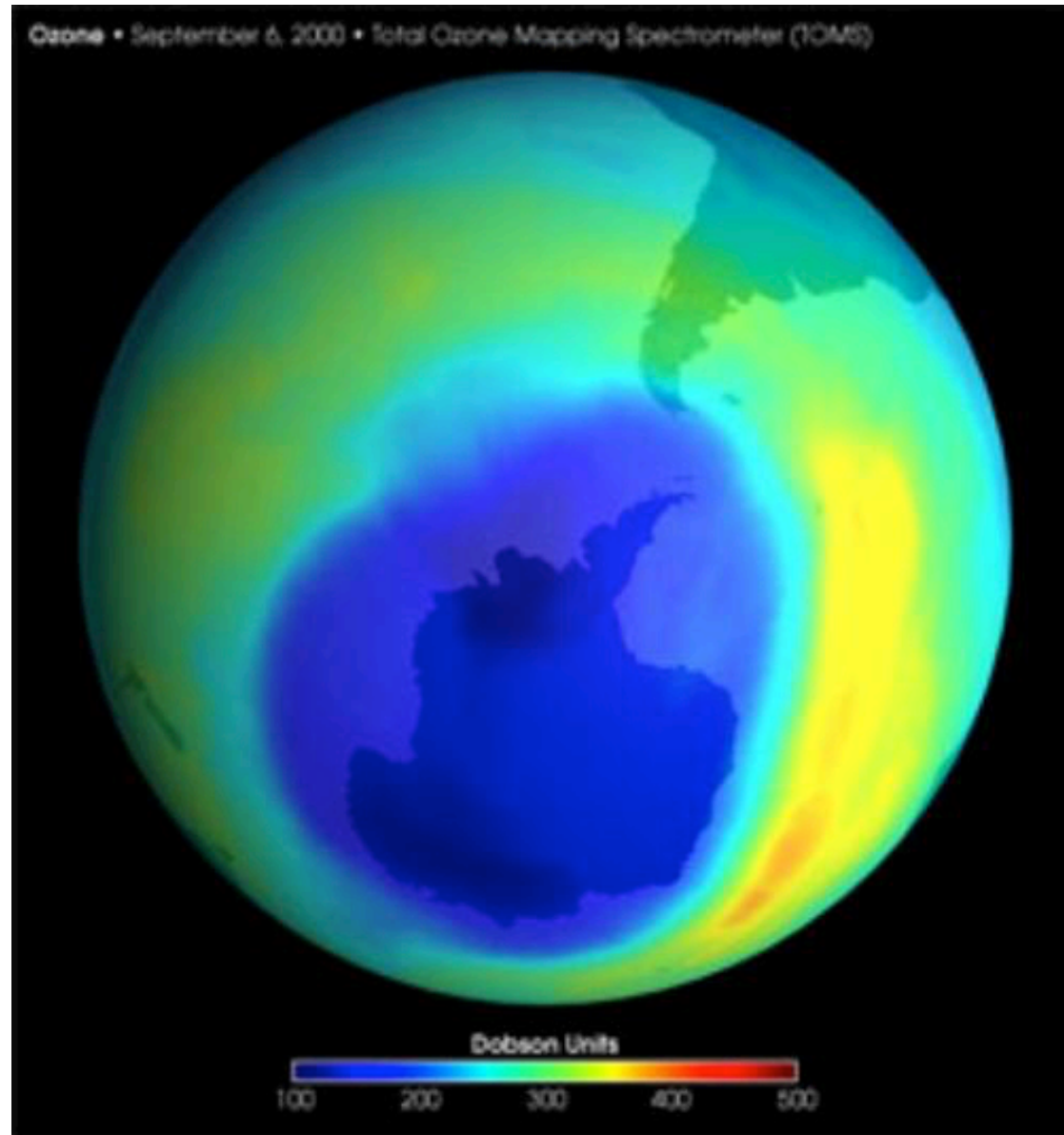
● Surveys

- Aquisição de dados da memória de pessoas sujeitas a um questionário ou entrevista
- Exemplo: obter a opinião de programadores sobre uma ferramenta

Métodos na pesquisa

- Os métodos são usados em combinação
- Modelagem e experimentação são usados para entender causalidade e prever eventos futuros
 - Dependem do conhecimento prévio de sistemas que pode ser obtido com estudos descritivos
- Estudos de caso descritivos e comparativos são usados para entender sistemas existentes em cenários reais
 - São usados para examinar a aplicação de resultados de experimentos e modelagem no mundo real
- Resultados de estudos descritivos e comparativos podem ser usados para confirmar relações causais identificadas por modelos e experimentos

Um exemplo: A Camada de Ozônio



Um exemplo: A Camada de Ozônio

- Estudos iniciais descritivos e comparativos apontam o problema (1957)
 - Décadas de obtenção de dados sobre a camada de ozônio
 - 1985: conclusão sobre um “buraco sazonal”
- Desenvolvimento de tecnologia abre novos caminhos de pesquisa (1970)
 - Detecção de pequenas concentrações de gases (incluindo clorofluorocarbonos - CFCs)

Um exemplo: A Camada de Ozônio

- Modelagem e pesquisa experimental são usadas para chegar a conexões causais (1972-1974)
 - Depois de ouvir resultados sobre CFCs, pesquisas adicionais revelaram que CFC pode destruir ozônio na atmosfera
- Pesquisa descritiva e pesquisa comparativa provêem confirmação do mundo real (1976-2002)
 - Medição com balões de alta altitude confirmando CFC
 - Ligação (correlação) entre perda de ozônio e câncer de pele

Um exemplo: A Camada de Ozônio

- A história inteira precisou de modelagem, experimentos, pesquisa comparativa e estudos descritivos para desenvolver uma teoria coerente sobre o ozônio na atmosfera e como humanos são afetados
- <http://www.beyonddiscovery.org/content/view/article.asp?a=73>

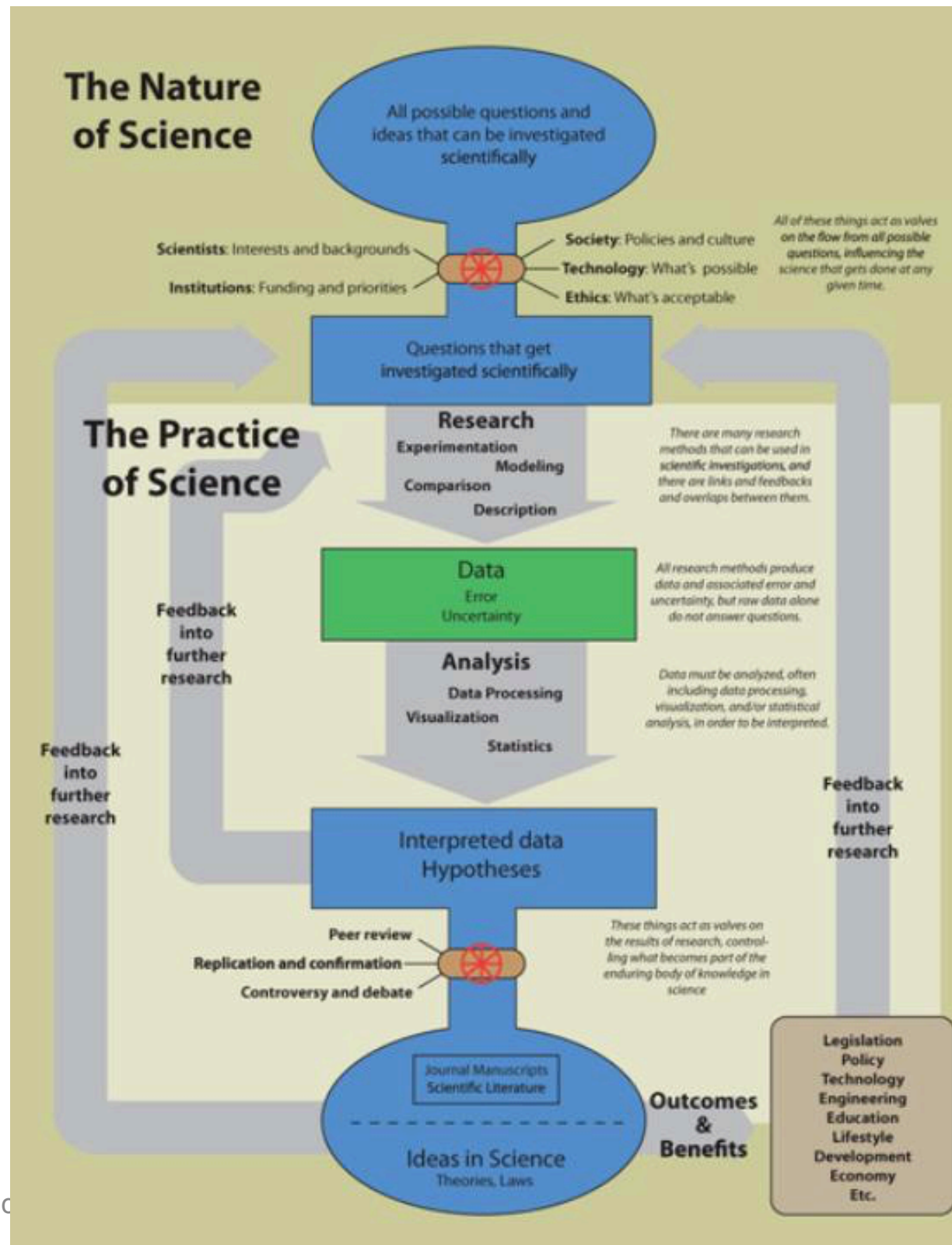
Ciência real: Comunicação,
comunidade, cultura

Comunicação, Comunidade, Cultura

- Há muita cooperação na ciência
- Como assegurar a integridade do método científico?
 - Publicação (conferências e revistas)
 - A importância de escrever bem!
 - “Peer review” (Revista por pares)
 - Documentação de replicação
 - Experimentos devem ser reproduzíveis
 - Arquivamento de fontes, dados, software, ...
 - Compartilhamento de dados

A Pesquisa na Prática

<http://www.visionlearning.com>



Erros Comuns ao Aplicar o Método Científico

- O cientista tem um viés (*bias*) e acredita tanto na hipótese que a toma como explicação verdadeira e
 - Não experimenta; ou
 - Ignorar dados que não suportam a hipótese
 - Sobre isto, leia o excelente ensaio de Gould: "[Measuring Heads: Paul Broca and the Heyday of Craniology](#)"
 - Você vai ficar **impressionado** sobre como bons dados podem ser usados para apoiar conclusões erradas ou suspeitas
 - Broca *parte das conclusões* e dá um jeito de apoiá-las com os dados
 - [Veja como o próprio Gould sofreu de viés!](#)
- Não tratar os erros experimentais (Incertezas/ruídos)
 - Falaremos disso à frente no curso

Bibliografia

- Reis, Método Científico,
http://www.projetoockham.org/ferramentas_metodo_1.html
 - simples e muito bom, apesar dos erros de português
- Fanelli,
[Do Pressures to Publish Increase Scientists' Bias? An Empirical Support from US States Data](#)
- Sagan (2006). *O mundo assombrado pelos demônios*, Companhia das Letras. Tradução do original de 1995.
- Métodos Estatísticos Aplicados à Engenharia de Software Experimental, ARAÚJO, BARROS, MURTA, TRAVASSOS, UFRJ
- Strong Inference
 - "[Strong Inference](#)" de Platt (1964)
 - http://en.wikipedia.org/wiki/Strong_inference
 - [Crítica de Strong Inference](#)

Bibliografia

- Research Methods: The Practice of Science,
http://www.visionlearning.com/library/module_viewer.php?mid=148
- Research Methods: Experimentation,
http://www.visionlearning.com/library/module_viewer.php?mid=150&l=
- Research Methods: Description,
http://www.visionlearning.com/library/module_viewer.php?mid=151&l=
- Research Methods: Comparison,
http://www.visionlearning.com/library/module_viewer.php?mid=152&l=
- [What kind of Scientist are you?](#)

Recursos

- Where good ideas come from?
 - Interagir é fundamental!
- Baloney detection kit

A Ciência e a Ciência da Computação

Jacques Sauvé

UFCG

Agradecimentos

- Usei algum material de Virgilio Almeida da UFMG, o qual agradeço
- Várias partes deste material vêm de:
 - Métodos Estatísticos Aplicados à Engenharia de Software Experimental, ARAÚJO, BARROS, MURTA, TRAVASSOS, UFRJ

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, 2004

Outra definição de Computação

- "The discipline of computing is the systematic study of algorithmic processes that describe and transform information: their theory, analysis, design, efficiency, implementation and application. The fundamental question underlying all of computing is, 'What can be (efficiently) automated?'"
 - Task Force on the Core of Computer Science, Computing as a Discipline, CACM 32/1 (Jan. 1989), p. 12.

Mais uma definição: Ciência da Computação

- [ACM Computing Curricula 2005](#)
- Define Ciência da Computação pelo que os profissionais fazem
 - Projetam e implementam software
 - Software desafiador
 - Supervisão de outros programadores
 - Inventam novas formas de usar computadores
 - Web, robótica, criação automática de conhecimento, aplicações em genética, ...
 - Desenvolvem formas eficazes de resolver problemas de computação
 - Melhor forma de: armazenar dados, enviar dados numa rede, mostrar imagens, ...
 - Foco em performance (velocidade, qualidade, ...)

Conclusão: confusão!

- As definições são muito diferentes
- **Será que sabemos o que é computação???**

Ciências, Objetos e Métodos

Ciência	Objetos	Método Dominante
	Simples	Análise
Lógica e Matemática	Abstrações, números, proposições, etc.	Dedução
Ciências Naturais	Objetos naturais: matéria, organismos vivos, etc.	Método indutivo hipotético
Ciências Sociais	Indivíduos, grupos, sociedade, etc.	Método indutivo hipotético, Hermenêutica(*)
Ciências Humanas	Idéias humanas, linguagens, culturas, etc.	Hermenêutica
	Complexo	Síntese

* Ramo da filosofia que se debate com a compreensão humana e a interpretação de textos escritos
Jacques Sauv  - UFCG - 2009

Ciências, Objetos e Métodos

- A Ciência da Computação lida com tudo isso!

Tipos de Pesquisa em Ciência da Computação

Tipos de Pesquisa em Ciência da Computação

- Ciência da Computação Teórica
- Ciência da Computação do tipo Engenharia
- Ciência da Computação Empírica

Ciência da Computação Teórica

- Tradições da lógica e matemática
- Procura entender os limites da computação
- Procura entender a capacidade de paradigmas computacionais
- Procura desenvolver soluções gerais para classes de problemas
- Uso de raciocínio dedutivo
 - Não é ciência, é matemática

Exemplo de pesquisa teórica

- Brewer's Conjecture and the Feasibility of Consistent, Available, Partition-Tolerant Web Services
 - Abstract: When designing distributed web services, there are three properties that are commonly desired: consistency, availability, and partition tolerance. It is impossible to achieve all three. In this note, we prove this conjecture in the asynchronous network model, and then discuss solutions to this dilemma in the partially synchronous model.

Ciência da Computação do tipo "Engenharia"

- Queremos fazer algo "melhor"
 - mais rápido, menor, mais barato, ...
- Procura construir algo novo (artefatos computacionais)
 - Criamos *novas* entidades
 - Elas devem ser avaliadas
- Cuidado! Se você vai "fazer software" na sua pesquisa, leia isto: "[Software is Not Science](#)"
- A avaliação dos artefatos computacionais frequentemente envolve observações
 - Ligação com Ciência da Computação Empírica

Ciência da Computação Empírica

- Baseada em observações
- Também chamada de Ciência da Computação Experimental (CCE)
 - Não é um nome muito legal porque podemos observar sem fazer experimentos Estudos de casos, surveys... não envolvem experimentos, mas são empíricos
 - Falaremos disso adiante
- Disciplina sintética (combinação de partes)
 - A ciência normalmente é mais analítica, reducionista
- Envolve a observação de artefatos computacionais ou de processos associados à computação
- Relação entre teoria e empiricismo não é necessariamente intensa em CC, como ocorre em física, química, etc.

Por que experimentamos na CC?

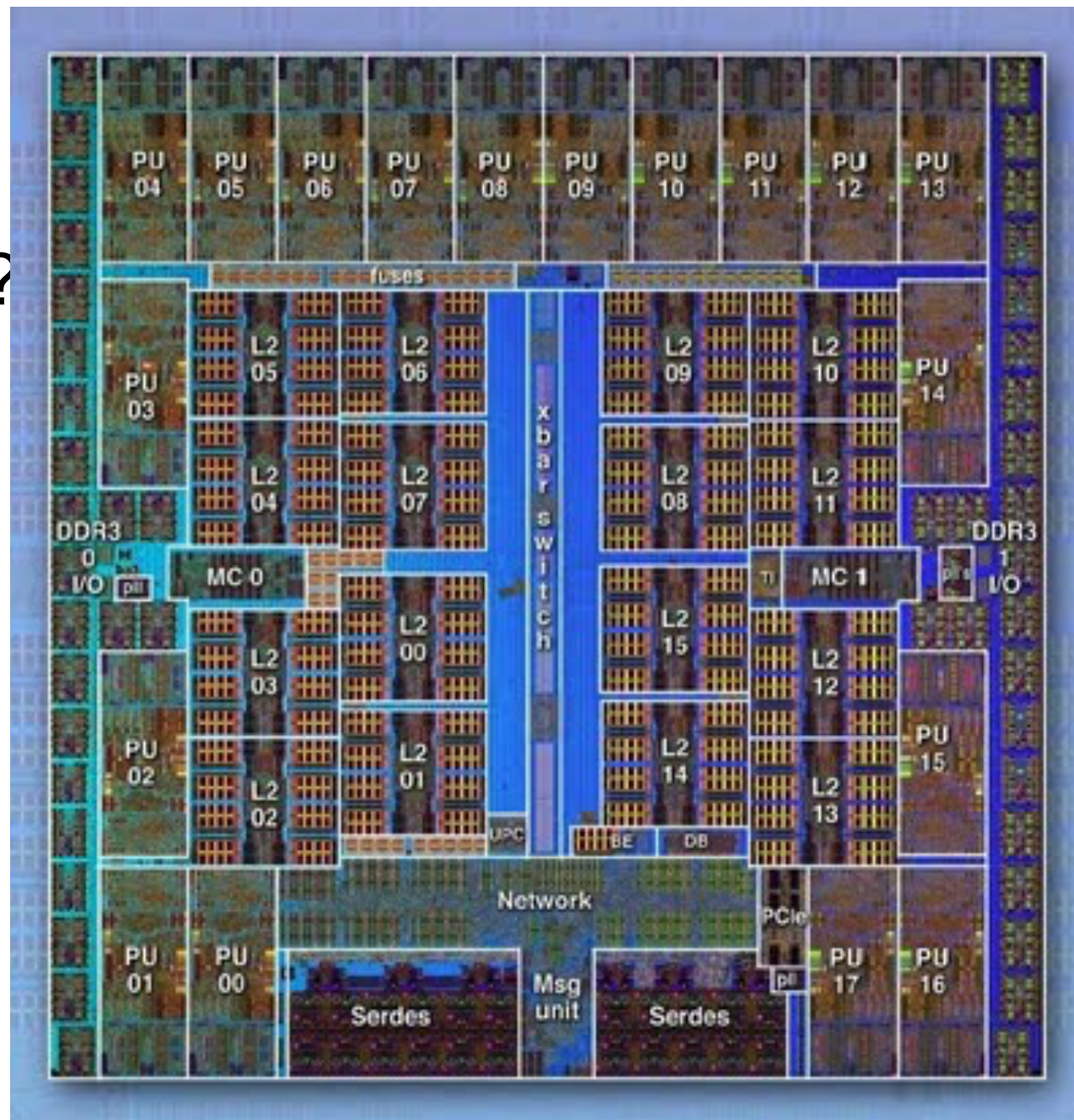
- Para entender o mundo que construímos
 - Nossos sistemas são muito complexos
 - Só porque os construímos não quer dizer que entendemos seu **comportamento emergente**
- Para complementar a teoria
 - Experimentar com resultados teóricos nos faz apreciar melhor como usá-los
- Para dar suporte ao projeto e à engenharia
 - Idéias que parecem boas nem sempre funcionam na prática
 - Devem ser testadas em condições realistas
 - Para provar que novas versões de fato são melhores

Você acredita isso
para
microprocessadores?

1 bi transistors

Você sabe onde
estão os hotspots?

IBM/BlueGene/Q
18 cores



Experimentação em Sistemas Computacionais

Por quê?:

“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.”

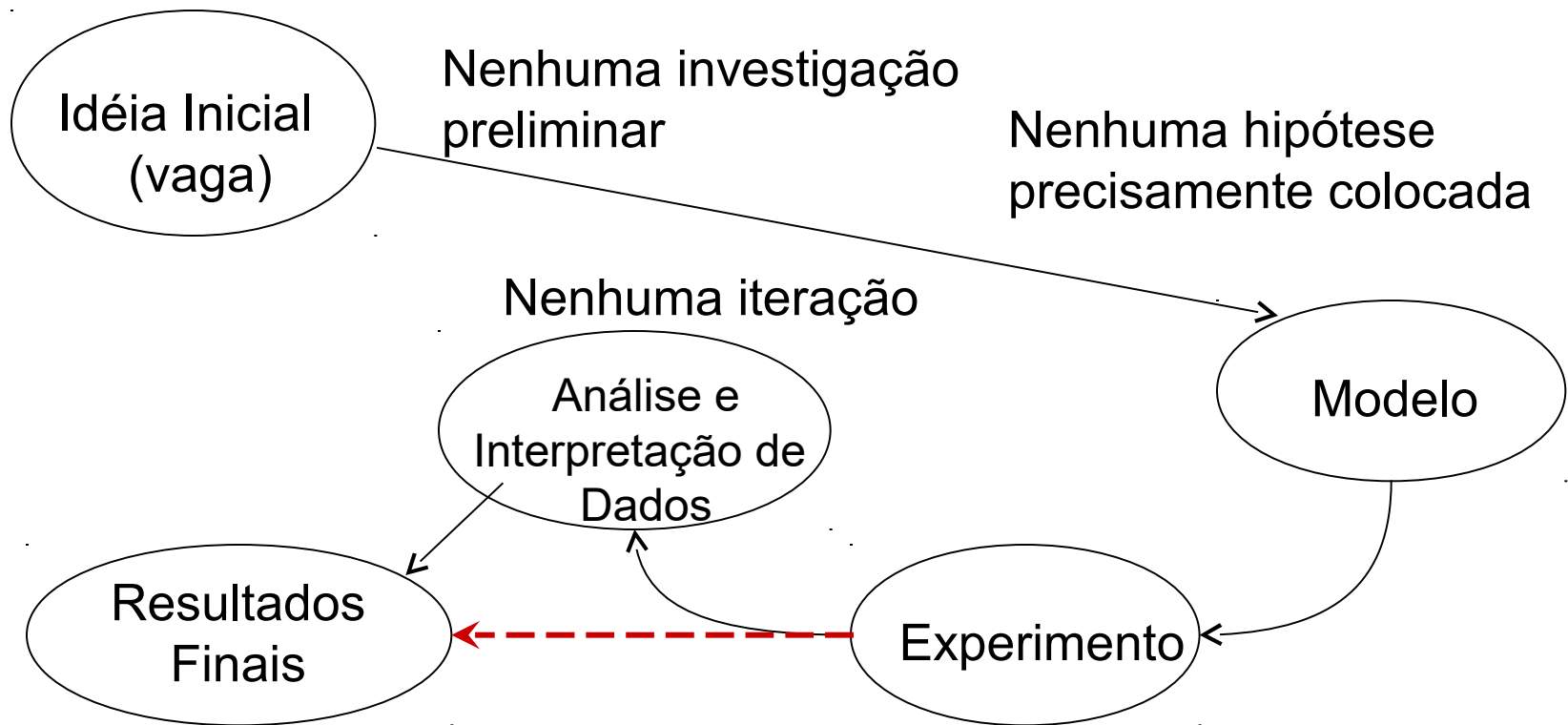
Richard Feynman, físico
Prêmio Nobel 1965

O Processo Experimental em Ciência da Computação e o Método Científico

Por que Ciência da Computação Experimental?

- Experimentação: Modelo-teste da teoria
 - Feynman: um experimento pode sempre identificar falhas na teoria (ex.: hipótese/suposições violadas pela realidade)
- Artigos para discutir:
 - D. Feitelson in “Experimental Computer Science: The Need for a Cultural Change”,
<http://www.cs.huji.ac.il/~feit/pub.html>
 - T. Mytkowicz et al., “Producing wrong data without doing anything obviously wrong”, Proc. ASPLOS 2010
 - W. Tichy in “Should Computer Scientists Experiment More?”, Computer 31(5), 1998

Prática Usual em Ciência da Computação



Ciclo de Vida Experimental



Várias maneiras questionáveis de atacar o processo experimental

- Específico, que se aplica somente ao “system under test”.
- Testes que não enxergam o futuro – motivações e observações baseadas apenas no passado.
- Ausência de cargas de trabalho representativas, falta de benchmarks confiáveis, etc.
- Ausência de uma cultura de replicação independente de outros experimentos.
- Dados reais, que são confusos e não confiáveis.

Aprenda o método correto

Processo Experimental Sistemático

1. Entenda o problema, estabeleça as perguntas e defina os objetivos
2. Selecione métricas
3. Identifique os parâmetros
4. Decida quais parâmetros serão estudados, i.e., serão variados (fatores)
5. Selecione a técnica
6. Selecione a carga de trabalho (workload)
7. Execute experimentos
8. Analise e interprete os resultados
9. Apresente os resultados e dados do experimento
10. Apresente conclusões

Processo Experimental Sistemático

1. Entenda o problema, estabeleça as perguntas e defina os objetivos: “*A problem well-stated is half-solved*”.
 - Deve-se ser objetivo
 - Seja capaz de responder “por que”, e também “como”
 - Delimite o escopo
 - Defina as perguntas que pretende responder
2. Selecione métricas que ajudarão analisar as perguntas.

Processo Experimental Sistemático

3. Identifique os parâmetros que afetam o comportamento
 - Parâmetros do sistema (ex.: configuração de hardware)
 - Parâmetros da carga (workload, ex.: padrões de chegada de requisições)
3. Decida quais parâmetros serão estudados, i.e., serão variados (fatores)
 - Normalização
 - Comece com lista pequena

Processo Experimental Sistemático

5. Selecione a técnica:

- Medição de uma implementação de protótipo
 - Quanto invasivo? Podemos quantificar o “overhead” da monitoração? Podemos medir o que desejamos?
- Simulação – quanto detalhada ? Como será a validação?
- Repetibilidade

6. Selecione a carga de trabalho (workload)

- Representativa?
- É aceita pela comunidade científica?
- Disponibilidade de dados?

Processo Experimental Sistemático

7. Execute experimentos

- Quantos testes devem ser rodados?
Quantas combinações dos parâmetros que formam o ambiente experimental?
- Análise da sensibilidade dos outros parâmetros.

8. Analise e interprete os resultados

- Use Estatística para analisar a variabilidade, “outliers”, etc.

Processo Experimental Sistemático

9. Apresente adequadamente os resultados e dados do experimento

- Gráficos: a questão da visualização dos resultados, distribuições estatísticas, etc.

10. Apresente conclusões

- Para onde os resultados nos levam?
- Quais os próximos passos?
- Novas hipóteses, novas questões, outros experimentos.

Processo Experimental Sistemático

1. Entenda o problema, estabeleça as perguntas e defina os objetivos
2. Selecione métricas
3. Identifique os parâmetros
4. Decida quais parâmetros serão estudados, i.e., serão variados (fatores)
5. Selecione a técnica
6. Selecione a carga de trabalho (workload)
7. Execute experimentos
8. Analise e interprete os resultados
9. Apresente os resultados e dados do experimento
10. Apresente conclusões

Processo Experimental Sistemático

1. Entenda o problema, estabeleça as perguntas e defina os objetivos
2. Selecione métricas
3. Identifique os parâmetros
4. Decida quais parâmetros serão estudados, i.e., serão variados (fatores)
5. **Selecione a técnica**
6. **Selecione a carga de trabalho (workload)**
7. Execute experimentos
8. Analise e interprete os resultados
9. Apresente os resultados e dados do experimento
10. Apresente conclusões

Visão Geral das Principais Técnicas que Iremos Estudar

1. Caracterização de Cargas
2. Sumarização de Dados
3. Comparando Quantitativamente Sistemas
4. Projeto Fatorial Completo:
 - a. Projetos fatoriais 2^k
 - b. Projetos fatoriais com replicação
 - c. Projetos fatoriais fracionários
 - d. Projeto de um fator
5. Regressão Linear Simples

Caracterização de Cargas

- Como descrever a carga do sistema alvo precisamente?
 - Qual importância?
 - Quais perguntas você deve tentar responder?
 - Qual propósito da caracterização?
 - Qual nível da descrição?
 - Quais componentes de carga queremos descrever?

Exemplo de um servidor Web

- Carga descrita por duplas (CPU time, I/O time)
- Dados disponíveis: logs de 10 requisições HTTP
- Caso 1: só 1 tamanho de documento (15KB)
 - 10 execuções ---> (0.013 sec, 0.09 sec)
- Caso 2 (mais realista):
 - documentos têm tamanhos diferentes.

Exemplo de um servidor Web

Request No.	CPU time	I/O time	Elapsed time
1	0.0095	0.04	0.071
2	0.0130	0.11	0.145
3	0.0155	0.12	0.156
4	0.0088	0.04	0.065
5	0.0111	0.09	0.114
6	0.0171	0.14	0.163
7	0.2170	1.20	4.380
8	0.0129	0.12	0.151
9	0.0091	0.05	0.063
10	0.0170	0.14	0.189
Average	0.0331	0.205	0.550

Como representar / sumarizar esta carga?

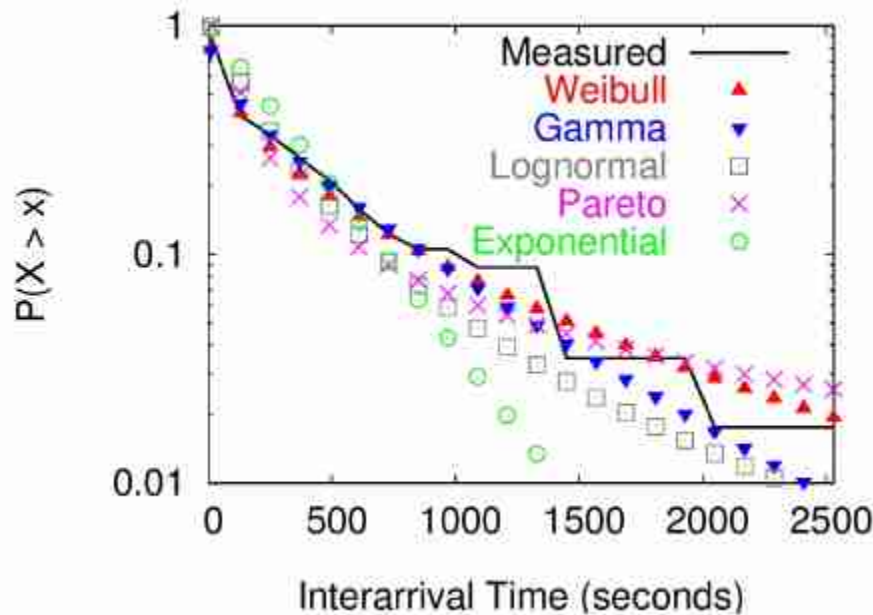
Somente uma média é suficiente (representativo)?

Caracterização de Cargas

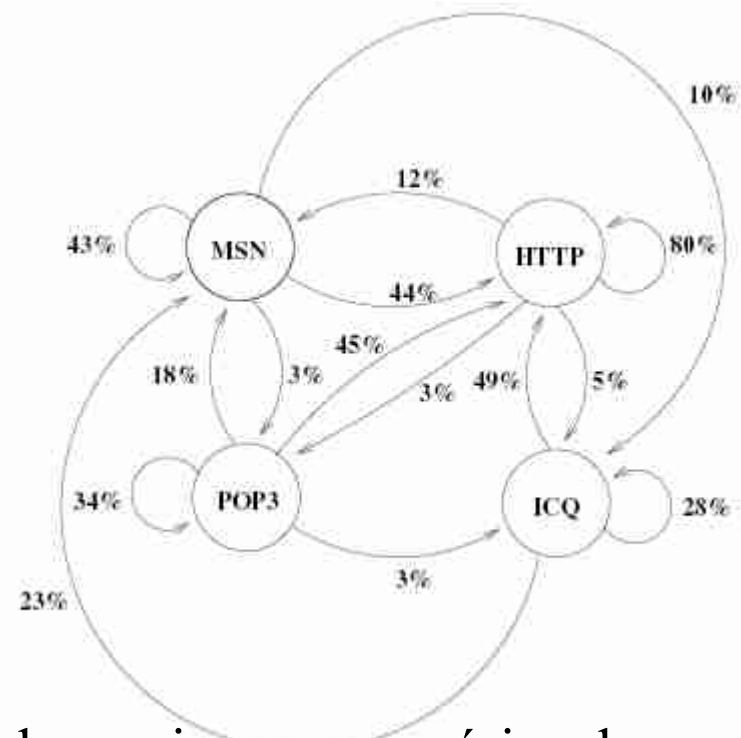
- Particionamento de cargas em sub-classes
 - Quando? Como?
- Representação de cargas em workload models
 - Distribuições estatísticas, geração de cargas sintéticas

Caracterização/Modelagem de Cargas

- Particionamento de cargas em sub-classes
 - Quando? Como?
- Representação de cargas em workload models
 - Distribuições estatísticas, geração de cargas sintéticas



eTeach

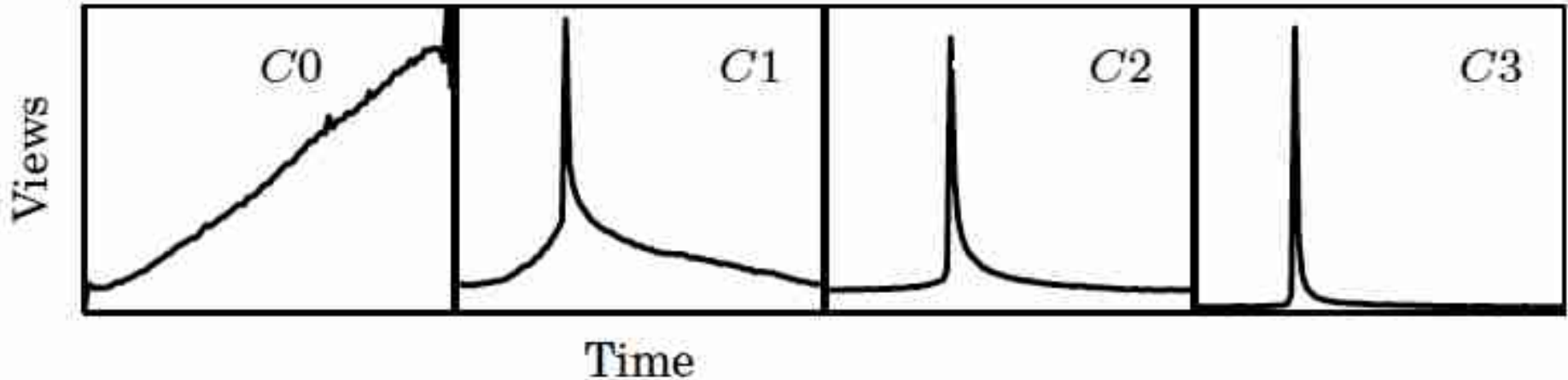


Tempos entre chegadas de requisições em um servidor de vídeo

Uso de serviços por usuários de banda larga

Caracterização/Modelagem de Cargas

- Particionamento de cargas em sub-classes
 - Quando? Como?
- Representação de cargas em workload models
 - Distribuições estatísticas, geração de cargas sintéticas



Perfis de popularidade em vídeos do YouTube

Sumarização de Dados

- Revisão:
 - probabilidade, independência, regra de Bayes
 - população X amostra
 - Média, variância, desvio padrão, CV, correlação, quantis
 - Distribuições estatísticas
 - Teorema Central do Limite:
“ a distribuição das médias amostrais será aproximadamente Normal”
- Estimando uma população a partir de uma amostra: tudo o que você quer fazer
 - Intervalo de Confiança : confiança X precisão

Exemplo

35 samples: 10 16 47 48 74 30 81 42 57 67 7
13 56 44 54 17 60 32 45 28 33 60 36 59 73
46 10 40 35 65 34 25 18 48 63

Sample mean $\bar{x} = 42.1$.

Standard deviation $s = 20.1$. $n = 35$

90% confidence interval is

$$42.1 \pm (1.645) \frac{20.1}{\sqrt{35}} = (36.5, 47.7)$$

De onde veio o 1.645?

Como interpretar o intervalo (36.5 , 47.7) para usá-lo na prática:
confiança X precisão da estimativa

Comparação de Sistemas

- Comparação Quantitativa de Sistemas: algoritmos, métodos, protótipos
 - Como usar Intervalos de Confiança
 - Teste com observações pareadas
 - Teste com observações não pareadas
 - Teste-t
 - Comparando valores médios X proporções
 - Escolha do tamanho da amostra

Exemplo 1

1. O tempo de processamento necessario para executar uma tarefa foi medido em dois sistemas.

Os tempos no sistema A foram:

$\{5.36, 16.57, 0.62, 1.41, 0.64\}$.

Os tempos no sistema B foram:

$\{19.12, 3.52, 3.38, 2.50, 3.60\}$.

Os dois sistemas são significativamente diferentes?

O sistema A é significativamente melhor/pior?

Exemplos 2 e 3

2. Um experimento foi repetido em dois sistemas 40 vezes. O sistema A foi superior ao B em 26 repeticoes. Podemos dizer que, com confianca de 99%, o sistema A e superior ao sistema B? E com uma confianca de 90%?

2. Cinco execuções independentes de um sistema levaram a tempos de execução iguais a 22.5, 19.8, 21.1, 26.7, 20.2 segundos
Quantas vezes devo repetir o experimento para obter um intervalo de confiança com precisão de $\pm 5\%$ e um nível de confiança de 90%?

Projetos Fatoriais 2^k

- Usados para determinar os efeitos de k fatores
 - Cada um com duas alternativas ou níveis
 - Cada fator medido é representado por seu nível máximo e pelo seu nível mínimo.
- Em geral, são usados de maneira preliminar, antes de estudos mais detalhados
 - Responde a pergunta: qual(is) dos k fatores tem maior impacto no sistema alvo?
 - Pode oferecer algum “insight” sobre as interações entre os vários fatores.
 - Existem interações com impacto significativo?
 - Se sim, qual a consequencia para o projeto experimental?

Exemplo de Projeto Fatorial 2²

- Uma arquitetura de máquina de busca, composta por N servidores.
 - 1º fator é o número de servidores N ($N = 8$ ou 64)
- Pode-se usar vários esquemas de distribuição ou escalonamento de *queries* para os servidores, por exemplo, *round-robin*, *gang*, *random*, *priority*, etc
 - 2º fator é o escalonamento (random ou round-robin)
- O objetivo é completar as *queries* no menor tempo possível.

Logo, a métrica de interesse é o tempo de execução da *query* em microsegundos.

Exemplo de Projeto Fatorial 2²

- Execução única de uma carga benchmark de *queries* nas duas configurações resultou nos seguintes tempos de execucao:

	8 Serv. (-1)	64 Serv. (+1)
Rand. (-1)	820	217
RR (+1)	776	197

- Fator A = número de servidores
- Fator B = escalonamento

Exemplo de Projeto Fatorial 2²

- Objetivo: produzir um modelo (não linear) para estimativa do tempo de execução de uma query

$$y = 502.5 - 295.5x_A - 16x_B + 6x_Ax_B$$

$q_0 = 502.5$ = tempo de execução médio

Fator A - número de servidores -
faz uma diferença de $\pm 295,5$ no tempo de resposta

De fato A explica 97% da variação nos dados

Fator B – escalonamento – baixo impacto (3%)

Interferência entre A e B também baixo impacto (0.4%)

Proceder: refinar avaliação do impacto de A: fator único

Projetos Fatoriais

- Projetos Fatoriais 2^k
 - 2^k experimentos
 - Objetivo: selecionar parâmetros de maior impacto para avaliação mais refinada a seguir
- Projetos Fatoriais com replicação
 - Replicação de experimentos independentes para cada uma das 2^k configuração
 - Permite estimar erros de experimentação e avaliar qualidade do modelo
- Projetos Fatoriais Fracionários
 - E se for muito caro rodar todos os 2^k (com/sem replicação) experimentos?
- Projeto de Um Fator

Exemplo: CEC 2010

Tuning Genetic Programming Parameters with Factorial Designs

Elisa Boari de Lima, Gisele L. Pappa, Jussara Marques de Almeida,
Marcos A. Gonçalves and Wagner Meira Jr.

Abstract—Parameter setting of Evolutionary Algorithms is a time consuming task with two main approaches: parameter tuning and parameter control. In this work we describe a new methodology for tuning parameters of Genetic Programming algorithms using factorial designs, one-factor designs and multiple linear regression. Our experiments show that factorial designs can be used to determine which parameters have the largest effect on the algorithm's performance. This way, parameter setting efforts can focus on them, largely reducing the parameter search space. Two classical GP problems were studied, with six parameters for the first problem and seven for the second. The results show the maximum tree depth as the parameter with the largest effect on both problems. A one-factor design was performed to fine-tune tree depth on the first problem and a multiple linear regression to fine-tune tree depth and number of generations on the second.

parameters for genetic algorithms, and search the space of parameter values using some sort of heuristic.

Many disciplines in Computer Science and Engineering require parameter setting. Factorial experimental designs [8] are one of the most broadly applied techniques for this purpose. The idea behind a factorial design is to obtain the maximum information about the factors (i.e., parameters) with the minimum number of experiments. Furthermore, it allows us to determine whether a parameter's effect on the algorithm's output is significant compared to the experimental error inherent to the measurement process.

The proposed methodology for parameter tuning relies on a 2^k factorial design to estimate the influence of the parameters on the GP runs. Once the parameters with the

Exemplo: CEC 2010

TWO PARAMETER CONFIGURATIONS OF THE FACTORIAL DESIGNS FOR SYMBOLIC REGRESSION, WITH LOWER (LB) AND UPPER BOUNDS (UB).

Parameter	Code	Config. 1		Config. 2	
		LB	UB	LB	UB
Population Size	A	20	10000	350	650
Generations	B	5	250	35	65
Tree Depth	C	2	20	2	20
Crossover Rate	D	0.6	0.99	0.6	0.99
Mutation Rate	E	0	0.4	0	0.4
Reproduction Rate	F	0	0.2	0	0.2

Exemplo: CEC 2010

TABLE III
RESULTS OF THE 2^6 30 FACTORIAL DESIGN FOR SYMBOLIC
REGRESSION. SIGNIFICANT EFFECTS AND CORRESPONDING
PERCENTAGE OF EXPLAINED VARIATION ARE PRESENTED.

Factor	Config. 1 %	Effect	Config. 2 %	Effect
A	35.8	-0.35	0.53	-0.03
B	15.11	-0.23	2.28	-0.07
C	16.38	-0.24	69.00	-0.37
E	0.63	-0.05	0.42	0.03
F	-	-	0.19	-0.02
AB	0.45	-0.04	-	-
AC	3.88	-0.12	0.15	-0.02
AE	0.61	0.05	-	-
BC	10.37	-0.19	1.91	-0.06
BE	0.56	-0.04	-	-
CE	-	-	0.95	0.04
CF	-	-	0.22	-0.02
ABC	1.47	-0.07	-	-
ABE	0.53	0.04	-	-
ABCE	0.05	0.01	-	-
% Explained	86.07%		75.95%	
Std Deviation	0.005		0.005	

The influence of A and B in the first but not in the second parameter configuration is closely related to the lower and upper bounds chosen for these parameters: since the total number of fitness evaluations varies in orders of magnitude from the lower to the upper bound in the first configuration, the impact of these parameters will certainly be large.

Regressão Linear Simples

- Fazer uma regressão é uma maneira de criar um modelo de predição para uma variável resposta y utilizando observações anteriores (x, y)
 - Modelo útil para previsão e também para analisar impacto de fator x em vários níveis.
- Como determinar se o modelo criado é adequado / bom?
- Como estimar os parâmetros do modelo?
- Como utilizar o modelo para realizar predições?
 - Respostas : precisão X confiança
 - Intervalos de confiança de parâmetros, predições, etc
- Quais premissas têm que ser respeitadas?

Exemplo

- O numero de disk I/Os e o tempo de CPU foram medidos para sete programas
(14, 2) (16, 5) (27, 7) (42, 9) (39, 10) (50, 13)
(83, 20).

Desenvolva um modelo de predicao de tempo de CPU em funcao do numero de operacoes de disco.

Este modelo e razoavel?

Resolução

- Modelo: $\text{CPU Time} = -0.0083 + 0.2438 \text{ DiskTime}$
- Modelo é bom?
Regressão explica 97.15 % da variação nos valores de CPU Time: muuuuuiitttoooo bom!!!
- Existe modelo melhor?
Podemos mostrar que o parâmetro constante (-0.0083) não é significativo com 90% de confiança

Logo um modelo mais simples e tão bom quanto (em termos de significância estatística) é:

$$\text{CPU Time} = 0.2438 \text{ DiskTime}$$