

Ciência da computação é ciência?

Jacques Wainer

Instituto de Computação
UNICAMP

Plano

- I) O que é ciência - algumas definições
- II) O que é ciência - minhas ideias
- III) Ciência da computação é ciência? - revisão bibliográfica
- IV) Ciência da computação é ciência? - minha opinião
- V) O que tem sido feito em publicações de computação?

Caveat

Minhas opiniões sobre as diferentes áreas da computação e das outras ciências é derivado de um conhecimento desigual de cada uma delas. Eu classifico meu conhecimento das práticas de pesquisa nas áreas:

- IA (CC), sistemas colaborativos(CC), medicina, e física

como sendo os melhores, e nas áreas:

- desempenho (CC), geologia, e química

como sendo os piores.

I) O que é ciência?

O que é ciência?

Wikipédia: *Investigação racional ou estudo da natureza, direcionado à descoberta da verdade. Tal investigação é normalmente metódica, ou de acordo com o método científico - um processo de avaliar o conhecimento empírico.*

O que é ciência? (2)

Popper (na Wikipédia): *Popper argumentou que a teoria científica será sempre conjectural e provisória. Não é possível confirmar a veracidade de uma teoria pela simples constatação de que os resultados de uma previsão efetuada com base naquela teoria se verificaram.*

O que a experiência e as observações do mundo real podem e devem tentar fazer é encontrar provas da falsidade daquela teoria. Este processo de confronto da teoria com as observações poderá provar a falsidade (falsify) da teoria em análise. Nesse caso há que eliminar essa teoria que se provou falsa e procurar uma outra teoria para explicar o fenômeno em análise. (Ver Falseabilidade).

Ciência = método científico?

Wikipédia: O método científico é um conjunto de regras básicas para um cientista desenvolver uma experiência a fim de produzir conhecimento, bem como corrigir e integrar conhecimentos pré-existentes. É baseado em juntar evidências observáveis, empíricas, e mensuráveis, baseadas no uso da razão.

Embora procedimentos variem de uma área da ciência para outra, consegue-se determinar certos elementos que diferenciam o método científico de outros métodos. Primeiramente os pesquisadores propõem hipóteses para explicar certos fenômenos, e então desenvolvem experimentos que testam essas previsões. Então teorias são formadas juntando-se hipóteses de uma certa área em uma estrutura coerente de conhecimento.

II) Ciência - minhas opiniões

O que é ciência - minhas opiniões

Ciência tem a ver com **descobertas empíricas** (ou experimentos)

Ciência tem a ver com fazer **teorias** que explicam essas descobertas

Algumas ciências fazem teorias **profundas**. Determinar as conseqüências da teoria ou mesmo se ela explica um fenômeno ou não é difícil - usa-se/inventa-se a matemática necessária (física, química?)

Minhas opiniões(2)

Algumas ciências criam teorias **pouco profundas**. As teorias não usam matemática, e as previsões são poucas e pouco detalhadas. (geologia?, biologia?)

Algumas ciências criam teorias **rasas**. As teorias nessas ciências são “leis” ou “princípios”, que apenas generalizam o fenômeno mas não permite novas previsões. (zoologia?, antropologia?)

Algumas ciências quase **não propõem** teorias (ainda). A ciência é centralmente sobre as descobertas empíricas (medicina)

Minhas opiniões (3)

Teorias pouco profundas ou rasas apenas “explicam” um fenômeno, mas normalmente não permitem previsões - e portanto não são falseáveis no sentido de Popper

Algumas teorias são estatísticas - elas refletem um caso “normal”, mas não atribuem probabilidade aos eventos. Outras teorias são determinísticas ou são probabilísticas (atribuem probabilidades calculáveis para os eventos).

Matemática não é ciência

Matemática e lógica **não são ciência.**

Embora a qualidade (grau de “verdade”) de uma afirmação em matemática seja muita alta (como se espera em ciência), a matemática e a lógica não são empíricas.

A qualidade das afirmações é uma consequência e não um pré-requisito das ciências!

Matemática não é ciência (2)

do Wikipédia (formal science): *However, many scholars oppose including formal science as a branch of science. They admit that formal science is a very powerful tool to natural and social science, but it does not mean formal science is science. Most importantly, they define science as the discipline using scientific method which bases on observation and empirical study. As knowledge in formal science is a priori and always constructed by rules of deduction from axioms and definition without any empirical study, they refuse to classify formal science as a branch of science.*

Ciência básica x aplicada

Ciência básica: *busca explorar as fronteiras do entendimento fundamental da natureza sem preocupações com utilidade e praticidade* - ex: interação entre partículas sub-nucleares

Ciência aplicada: *busca conhecimentos úteis para problemas práticos* - ex: vacina para malária

Ciência básica x aplicada (2)

Ciência básica - trabalha mais no nível de teorias - desenvolvimento de teorias, desenvolvimento de experimentos críticos que distinguem teorias, etc

Ciência aplicada - trabalha mais no empírico - o que dá certo para atingir um certo objetivo prático.

A distinção é um contínuo. Biologia, por exemplo, parece ser tanto básica como aplicada.

Até onde eu entendo Ciência aplicada = Engenharia

Níveis de ciência aplicada

1) **Pré Arte** (não-ciência):

fazer A deu certo para o objetivo O

Não existe pesquisa científica nesse nível(?).

Níveis de ciência aplicada (2)

2) Arte (não-ciência):

fazer A (normalmente) dá certo para o objetivo O

exemplos: pintura, administração, etc

Não existe pesquisa científica nesse nível. A técnica A é ensinada mas não é explorada.

Níveis de ciência aplicada (3)

3) Ciência empírica comparada

Fazer A é melhor que B para o objetivo O porque há experimentos que comparam as duas alternativas e a A se saiu melhor, embora não se saiba explicar porque.

exemplos: medicina, metalurgia(?)

A ciência está em fazer os experimentos de comparação entre as diferentes técnicas.

Níveis de ciência aplicada (4)

4) Empírica embasada teoricamente

Fazer A é melhor que B para o objetivo O porque há experimentos que comparam as duas alternativas e a A se saiu melhor, e a primeira técnica é derivável/explicável na teoria T (mais outras suposições).

exemplos: engenharia elétrica, aeronáutica,

Ciência está em fazer os experimentos comparativos, e fazer as deduções que comprovam os resultados.

III) Computação é ciência? - a bibliografia

As respostas da bibliografia

- Computer science as empirical inquiry: symbols and search - Newel and Simon, Comm da ACM, 1976 - Turing award lecture
- The computer scientist as a toolsmith II - Fedrerick Brooks - Comm da ACM 1996 - Alan Newell award lecture
- Is computer science science? - Peter Denning - Comm da ACM, 2005.

Newell e Simon

(Ambos são criadores da IA)

Ciência da computação é uma ciência natural que modela o fenômeno de cognição. Inteligência artificial é uma disciplina empírica (com experimentos computacionais e em humanos) que tenta modelar a cognição.

Newell e Simon inauguram a posição de ciências naturais da CC - que (acho) outros seguiram, especialmente em vida artificial e biologia computacional.

Crítica à Newell e Simon

Em 76 havia alguma relação entre IA e ciências cognitivas, e algumas pessoas achavam que estavam fazendo modelagem cognitiva (Newell e Simon entre eles). Hoje não há ninguém em IA que faça qualquer argumentação de “plausibilidade psicológica” ou de modelagem da cognição.

Há o mesmo argumento sendo aplicado à Biologia Computacional e hoje há pesquisas interdisciplinares com as duas áreas. Mas mesmo que programas de simulação estejam sendo usados para determinar algumas consequências de uma teoria (de outra área), isso garante cientificidade à CC?

Brooks

(pesquisador de computação gráfica)

Ciência da computação como criador de ferramentas:

- novidade da ferramenta não é importante
- utilidade da ferramenta é que é importante
- quem especifica a ferramenta é o usuário não o desenvolvedor
- usuário e desenvolvedor tem que trabalhar junto
- uma das “coisa importantes” para o desenvolvedor é usar bons algoritmos para resolver o problema do usuário.

Denning

Três aspectos da computação:

- ciência: “*experimental algorithms, experimental computer science and computational science*”
- engenharia: “*design, development, software engineering, computer engineering*”
- matemática: “*computational complexity, mathematical software, numerical analysis*”

Crítica à Denning

As coisas que Denning cita como centrais na perspectiva científica quase não existem!

O que é “*experimental computer science*”? Existe “*experimental software engeneering*”!

“*Experimental algorithmicists study the performance of real algorithms on real data sets and formulate models to predict their time and storage requirements; they may one day produce a more accurate theory than Big-O-Calculus and include a theory of locality.*”

Crítica à Denning (2)

Acho que isso é “*computational science*:”

“Bioinformaticians explain DNA as encoded biological information and study how transcription enzymes read and act on it; computer models of these processes help customize therapies to individual patients. Pharmaceutical and materials labs create man-made molecules through computer simulations of the information processes underlying chemical compositions.”

Crítica à Denning (3)

Simuladores (e coisas afim) servem um papel similar à matemática em teorias profundas - demonstrar que uma previsão deriva da teoria é um processo difícil onde simuladores são uma ferramenta apropriada (biologia, sociologia).

Mas usar simuladores em outras áreas científicas confere cientificidade aos simuladores? E à toda Ciência da Computação?

IV) Computação é ciência? - minha opinião

CC é ciência? - minha visão

A parte da computação referente à provas (complexidade, correção) é parte de Matemática e não é ciência!

Computação não parece ter teorias - e portanto não é uma ciência básica

Algumas partes da computação são ciência aplicada: a maioria no nível de pré-arte e arte, e algumas no nível de empírico comparativo, e pouca no nível teoricamente embasado.

CC é ciência? - minha visão (2)

Muito do que se faz não é nem ciência aplicada (pré-arte) - desenvolve-se **novos** artefatos, de vez em quando para resolver problemas conhecidos, de vez em quando para problemas novos.

Computação (simulação) é usada como ferramenta em outras ciências - para derivar previsões das teorias - isso não confere cientificidade à CC.

CC gera teorias no máximo rasas e estatísticas, algumas são teorias limitadas no tempo.

Arte em CC

exemplos do nível arte em CC como ciência aplicada:

- OO é bom para reuso
- abstração e hierarquização são bons para desenvolvimento de programas
- código estruturado é mais fácil de entender
- RBAC é um bom modelo para controle de acesso

Empírico comparativo em CC

Na prática:

- 3SAT é fácil
- redes neurais são bons aproximadores de funções para problemas reais
- não vale apenas altas taxas de compressão de código, se o objetivo é a redução da área de silício (Guido, Rodolfo)
- quicksort é mais rápido que mergesort (teoricamente embasado - complexidade média mas equiprobabilidade das permutações)

Princípios/leis/teorias da CC

Se tirarmos complexidade assintótica, as teorias ou leis da CC são rasas e estatísticas (e datadas), sobre coisas como custos de criar software, como pessoas trabalham com computadores ou com suporte de computadores, etc

- 90% do custo de um software é na fase de manutenção
- o custo de corrigir um erro multiplica por 10 a cada fase de desenvolvimento do software

Classificação de sub-áreas da CC

- pré-arte: várias (partes de IA)
- arte: Eng de software, programação, linguagens
- empírica comparada: algumas partes de IA, otimização, algoritmos empíricos, HCI
- teoricamente embasadas: desempenho(?), otimização(?)

Modelo de maturidade em CC

- imaturas - pré-arte: criam algo novo para mostrar que é possível
- semi-maduras - arte: algumas avaliações não comparativas, alguma repetição no uso
- maduras - comparativas e/ou embasadas: experimentos de comparação “metodologicamente bem feitos,” com ou sem argumentação teórica.

E se matemática for ciência??

Então a situação é pior ainda do ponto de vista científico. Complexidade assintótica é uma teoria (profunda) de desempenho de programas que é repetidamente demonstrada como ERRADA e mesmo assim a comunidade como um todo se mantém nesse *paradigma* (no sentido de Kuhn) aceitando essa *anomalia* (Kuhn) brutal, e ao que parece sem um paradigma competitivo! De qualquer forma, complexidade assintótica é uma teoria para desempenho de programas, e a CC muito mais que isso.

Valores de pesquisa em CC

em todo artigo, em toda tese de doutorado:

- deve-se criar um artefato novo. Por que essa fixação com o novo? Talvez derivado da matemática e das ciências profundas que só aceitam a publicação de “coisas novas”
- deve-se implementar o artefato proposto. Implementação não prova nada - tudo que pode ser descrito pode ser implementado - se não há avaliação, a implementação é irrelevante.
- e o aluno é que deve implementar o artefato - não se pode contratar alguém para fazê-lo!!

O que é/será a boa CC

A computação pode no máximo aspirar a ser uma ciência aplicada no nível empírico comparado ou teoricamente embasado.

A boa CC será derivada do conhecimento metodológico de fazer comparações entre as várias alternativas.

Há sempre espaço para o novo: criar uma nova técnica, uma nova solução de um problema existente, mas o cientificamente importante é a comparação com as técnicas já existentes, e mostrar que a técnica nova é **melhor** que as anteriores.

O que é/será a boa CC (2)

Um CC nesses moldes será mais cara, publicará muito menos.

O cientista da computação dependerá da sorte (que sua nova solução é de fato a melhor para o problema), precisará de melhores conhecimentos do estado da arte, precisará do apoio de indexadores da produção científica, precisará ter acesso a “dados reais”, etc

E será mais útil aos praticantes da computação.

V) O que se publica em CC

O que se publica em CC?

Tichy e colegas analisou 50 artigos selecionados aleatoriamente dos publicados pela ACM em 1993 e classificou-os nas seguintes categorias:

- **teoria** - teoremas e provas
- **projeto e modelagem** - sistemas, técnicas e modelos cujas características não podem ser provadas formalmente
- **trabalho empírico** - usa desenhos/projetos de outros para avaliação.
- **teste de hipóteses** - define uma hipótese e experimentos para testá-la
- **outros** - surveys, etc

Tichy 93

teoria	12%
empírico	2%
hipótese	2%
outros	14%
projeto	70%
0% avaliação	30%
0-10 %	6%
10-20 %	12%
20-50 %	22 %
>50 %	0%

Tichy repetido 2006

Claudia Barsottini, Danilo Lacerda e Leandro Marco analisaram 145 artigos selecionados aleatoriamente da ACM publicados em 2005 (1% do total) - 2 revisores por artigo.

Classificaram nas mesmas classes que Tichy - treinaram analisando os 50 artigos do trabalho original

Eu servi de voto de minerva no caso de discordância.

Tichy 2006

	1993	2005
teoria	12%	4%
empírico	2%	18%
hipótese	2%	5%
outros	14%	3%
projeto	70%	70%
0% avaliação	30%	23%
0-10 %	6%	7%
10-20 %	12%	15%
20-50 %	22 %	21 %
>50 %	0%	3%

Discussão

23% (projeto com 0% de avaliação) do publicado não é ciência (apenas novo)

4% não é ciência pois é teoria

3% não classificado

70% (projeto $>0\%$ + empírico + hipótese) parece estar no nível de arte ou acima.

Na minha experiência, a maioria da avaliação não é comparativa!!

Outras referências

- Experimental evaluation in computer science: A quantitative study - Walter F. Tichy, Paul Lukowicz, Lutz Prechelt and Ernst A. Heinz - Journal of Systems and Software Volume 28, Issue 1, January 1995, Pages 9-18
- Medical Informatics: art or science - Jan H Bemmell - Methods of information in Medicine 1996
- Special issue on The Philosophy of Computer Science - Mind and machines - July 2007.