

Grandes Desafios da Pesquisa em Computação no Brasil – 2006 – 2016

Relatório sobre o Seminário realizado em 8 e 9 de maio de
2006

Autores

- André C. Ponce de Leon F. de Carvalho, USP
- Angelo Brayner, UNIFOR
- Antonio Loureiro, UFMG
- Antonio L. Furtado, PUC-Rio
- Arndt von Staa, PUC-Rio
- Carlos José Pereira de Lucena, PUC-Rio
- Clarisse Sieckenius de Souza, PUC-Rio
- Claudia Maria Bauzer Medeiros, UNICAMP
- Cláudio Leonardo Lucchesi, UNICAMP
- Edmundo Souza e Silva, UFRJ
- Flávio Rech Wagner, UFRGS /
- Imrer' E'undo S #

vaxRGS F
-rW hR □ □
US #R!-ñ ñDÒP` & '#. €

Prefácio

Este relatório sintetiza os resultados do Seminário “Grandes Desafios de Pesquisa em Computação no Brasil: –2006 – 2016, realizado em São Paulo nos dias 8 e 9 de maio de 2006. O seminário, promovido pela Sociedade Brasileira de Computação (SBC), com o apoio da CAPES e da FAPESP, reuniu durante os dois dias 26 pesquisadores brasileiros da área de Computação, autores deste relatório. Os participantes foram selecionados por uma Comissão de Coordenação a partir de 47 propostas de desafios enviadas de todo o Brasil. Os critérios de seleção priorizaram a abrangência e visão de futuro em termos de pesquisa. O trabalho editorial da Comissão de Coordenação procurou preservar no relatório final apenas as propostas consideradas consensuais.

O programa do seminário foi dividido em duas etapas. Inicialmente, os trabalhos selecionados foram apresentados de forma resumida. A seguir, os participantes foram divididos em 6 grupos de trabalho, que discutiram as propostas, consolidando-as em 5 desafios, apresentados neste relatório. As duas etapas foram intercaladas por um debate de dois convidados - Profs Renato Janine Ribeiro e Gilberto Câmara - sobre suas visões relacionadas ao tema do encontro e à. Em seguida, na sessão final, os resultados da reunião foram discutidos com representantes de agências de fomento, Softex e setores do governo.

A SBC pretende, com esta iniciativa, dar início a uma série de seminários multidisciplinares que aprofundem estes e outros temas, visando a intensificar o planejamento e a pesquisa de longo prazo em Computação no Brasil e a cooperação com outros domínios científicos do conhecimento.

A Comissão de Coordenação do Seminário agradece a todos os que enviaram suas propostas e, em particular, aos participantes, pelo trabalho realizado. Agradece especialmente aos convidados externos, Carlos Nobre (CPTEC-INPE/ABC), Gilberto Câmara (diretor do INPE), Paulo Artaxo (USP/ABC) e Renato Janine Ribeiro (Diretor de Avaliação da CAPES) pela participação no evento e pelos valiosos comentários. Finalmente, agradece as sugestões e a presença, na sessão final, de Carlos Henrique Brito Cruz (Diretor Científico da FAPESP), Eratóstenes Ramalho de Araújo (SOFTEX), José Augusto Suruagy Monteiro (Coordenador do Forum de Coordenadores de Pós-graduação da SBC) e Antenor Correa (SEPIN).

Maio de 2006

Carlos José Pereira de u!ena" Claudia #auzer Medeiros" Cláudio eonardo u!!\$esi"
José Carlos Maldonado e %ir&'lio Au&usto (ernandes Almeida

1. Introdução

A Computação revolucionou a pesquisa científica, sendo hoje reconhecida como o “terceiro pilar” a sustentar tal pesquisa, junto com os pilares da teoria e da experimentação [1]. Desta forma, ela permeia os avanços em todas as áreas do conhecimento. Novas formas de interação entre as ciências, em vários níveis e escalas, são mediadas pela Tecnologia da Informação, que é a simbiose da Ciência da Computação com diferentes domínios do conhecimento. Muitas das grandes descobertas científicas recentes são resultados do trabalho de equipes multidisciplinares que envolvem cientistas da Computação. Finalmente, ela é um componente indispensável para a implementação e o fortalecimento dos objetivos econômicos, tecnológicos e sociais de um país. Desta forma, planejamento a longo termo para a área terá conseqüências que irão além da Computação propriamente dita. O Seminário Grandes Desafios em Computação no Brasil foi concebido a partir de iniciativas semelhantes realizadas em outras partes do mundo e empreendidas a partir dessas constatações.

.) * +ue é um Seminário de , randes - esa.ios de Pes+uisa/

Um seminário de Grandes Desafios de Pesquisa não é um evento científico tradicional. Seu objetivo não é apresentar resultados de pesquisas em andamento, mas sim definir questões de pesquisa que serão importantes para a ciência e para o país no longo prazo. É um evento em que o ambiente deve ser propício ao pensamento criativo, sem o excesso de críticas e necessidade de validações e provas que caracterizam as conferências tradicionais para a apresentação de resultados de pesquisa. Não é tampouco uma conferência de “defesa de idéias ou projetos pessoais de pesquisa”, mas sim um trabalho coletivo de identificação e caracterização de grandes problemas de pesquisa. É um esforço para o refinamento do pensamento mais orientado a problemas do que às disciplinas e às questões da própria área.

Grandes Desafios envolvem questões associadas a “problemas centrais” que não podem ser resolvidas por pesquisas que objetivam resultados de curto-prazo. São necessários múltiplos enfoques para atacar grandes desafios, e estes deverão ser pesquisados dentro de um horizonte de longo prazo. A identificação de grandes desafios de pesquisa contribui para a formulação de projetos denominados , , que têm o potencial de produzir avanços significativos no campo científico, com aplicações sociais e tecnológicas de grande valor. Alguns aspectos característicos dos Grandes Desafios em Pesquisa estão resumidos a seguir.

1. Os Grandes Desafios devem ser dirigidos a avanços significativos na ciência, ao invés de conquistas incrementais baseados em resultados existentes.
2. A visão de um Grande Desafio deve estar bem além daquilo que pode ser obtido em um projeto de pesquisa no período típico de financiamento de um “.....”.
3. Os Grandes Desafios devem ser passíveis de avaliação clara e objetiva que permita definir o seu sucesso.
4. Os Grandes Desafios devem ser decomponíveis e passíveis de diagnóstico incremental, de modo a permitir mudanças de curso durante sua realização..

5. Os Grandes Desafios devem ser ambiciosos e visionários, mas não irrealistas, viáveis dentro de um prazo predefinido – no caso desta iniciativa, 10 anos.
6. Os Grandes Desafios devem ser desafiadores e motivadores para a comunidade científica e motivadores para a sociedade
7. Muitos dos problemas que os constituem são multidisciplinares em sua natureza e nas possibilidades de solução.
8. Os tópicos dos Grandes Desafios emergem de um consenso da comunidade científica, para servir como um cenário de longo prazo para os pesquisadores, independentemente de políticas de financiamento ou questões conjunturais.

2.2 Proostas e e10eri2n!ias similares

Esforços para a definição de grandes desafios de pesquisa em diversas áreas da ciência vêm ocorrendo há vários anos em países de elevada tradição científica, como Estados Unidos e Inglaterra. Alguns exemplos são:

- realizada nos EUA pelo National Research Council (NRC) a pedido da National Science Foundation (NSF), para as áreas ambientais. Um dos princípios era buscar direções inter e multidisciplinares, tendo em vista que vários problemas de pesquisa na área ambiental transcendem barreiras de disciplinas científicas tradicionais.
- : os Grandes Desafios e necessidades de pesquisa determinados foram: química verde, substituição de solventes, melhoria de catalisadores, fontes de energia renovável e alfabetização digital em todos os níveis
- iniciativa em nível mundial tentando identificar 14 grandes desafios que, se resolvidos, podem levar a importantes avanços na prevenção e tratamento de doenças que ‘ nçar vaeáá E

Na Inglaterra a UK Computing Research Committee e a British Computer Society produziram a seguinte lista de desafios em 2005:

1. *In Vivo – In Silico*
2. *Ubiquitous Computing: experience, design and science*
3. *Memories of Life*
4. *The Architecture of Brain and Mind*
5. *Dependable Systems Evolution*
6. *Journeys in Nonclassical Computation*

) .3 A S#C e os , randes - esa.ios da Com0uta45o no #rasil6 20067
20)6

O evento brasileiro foi organizado pela Sociedade Brasileira de Computação, a partir do modelo dos eventos internacionais citados anteriormente. O objetivo foi gerar um conjunto de cinco Grandes Desafios para a Computação no Brasil, acompanhados de uma especificação clara e sucinta da visão para o tratamento do problema em foco. Além disso, a formulação de cada desafio contemplou debates sobre os seguintes pontos: i) explicitação dos benefícios da busca da solução para o problema ii) descrição da maneira de medir o sucesso do avanço das pesquisas sobre o problema em questão, iii) elaboração das dificuldades e barreiras para se alcançar sucesso nas pesquisas na área em questão e iv) proposta de ações que deveriam ser tomadas para se enfrentar o desafio em um período de 10 anos.

Os cinco desafios propostos, descritos nas próximas seções, foram:

1. Gestão da Informação em grandes volumes de dados multimídia distribuídos
2. Modelagem computacional de sistemas complexos artificiais, naturais e sócio-culturais e da interação homem-natureza
3. Impactos para a área da computação da transição do silício para novas tecnologias
4. Acesso participativo e universal do cidadão brasileiro ao conhecimento
5. Desenvolvimento tecnológico de qualidade: sistemas disponíveis, corretos, seguros, escaláveis, persistentes e ubíquos

2. Gestão da informação em grandes volumes de dados multimídia distribuídos

Quase tudo que vemos, lemos, ouvimos, escrevemos, medimos é coletado e disponibilizado em sistemas de informação computacionais. Para obter efetividade e eficiência, é fundamental criar soluções escaláveis que possam responder às necessidades de desenvolvimento de aplicações com esses dados. O objetivo deste desafio é, assim, desenvolver soluções para o tratamento, a recuperação e a disseminação de informação relevante, de natureza tanto narrativa quanto descritiva, a partir de volumes exponencialmente crescentes de dados multimídia.

Vários fatores contribuem para o crescimento explosivo de dados. O problema tornou-se mais visível com a disseminação da Internet, em que indivíduos, empresas, governos, instituições não governamentais são produtores de conteúdo em potencial, transformando o mundo em uma enorme base de dados que é atualizada em tempo real, por milhares de pessoas, a cada segundo. Outro fator crítico para a geração exponencial de dados é a aparição de dispositivos que capturam novos tipos de dados extremamente complexos – desde satélites, microssensores, ou telescópios, até câmeras de vídeo que capturam interações humanas, ou dispositivos que registram a atividade cerebral em face de estímulos diversos. Finalmente, dados secundários são gerados por cientistas e pesquisadores ao fazer experimentos e acessar serviços computacionais, em campos tão variados quanto ciências da terra, astronomia, bioinformática, medicina, ou ciências sociais. Esses dados são de vários tipos (por exemplo, som, vídeo, texto e dados analógicos posteriormente digitalizados) e usam distintas unidades de coleta e medida (por exemplo, bits, interações, coletas físicas de espécimes na natureza), em várias escalas temporais e espaciais.

Todo esse imenso conjunto heterogêneo de dados, distribuído por todo o planeta, precisa ser processado, armazenado e disponibilizado para tornar possível a extração de informação para os mais diferentes tipos de usuários, em soluções escaláveis. Isto exige, dentre outros enfoques, pesquisa em novas técnicas e métodos de gerenciamento, extração de conteúdo de som e imagem, integração, indexação e recuperação de dados e informação. Além disso, apresenta desafios de preservação e segurança. Mais ainda, tendo em vista a constante evolução tecnológica, é preciso garantir que dados armazenados continuem acessíveis com o passar dos anos – ou seja, garantia de durabilidade e acesso em longo prazo. Ainda, outros fatores de pesquisa a considerar são a modelagem de grandes volumes de dados multimídia, formas distintas de entrada e saída multimodal, algoritmos e estruturas para processamento otimizado de redução e consulta aos dados e informações e, cada vez mais, atendimento a diferentes perfis e necessidades de aplicações e usuários.

Para tratar esses grandes volumes de dados e informação distribuídos é também essencial a exploração eficiente de todos os níveis de paralelismo, do nível do chip, hoje com processadores (.), passando pelo nível de arquitetura dos nós integrados através de redes de alto desempenho em um cluster, até a formação de grades (. . . .) pela comunicação de clusters heterogêneos através de uma rede clássica, tipo internet.

O grande desafio é a integração de todas essas linhas e áreas para conduzir a aplicações que possam beneficiar o contexto sócio-econômico-cultural do País. Embora haja resultados de pesquisa em cada uma das áreas isoladas, alguns ainda incipientes, não existem propostas que consideram a sua integração.

O presente desafio é importante porque, além de estimular a pesquisa em áreas básicas em Computação, sua integração pode influir no desenvolvimento de inúmeras aplicações chave em vários setores da sociedade. Exemplos são: criação de conteúdo para atividades educacionais (.....), gestão eficiente da informação visando apoio a governo eletrônico (.....), extração de subconjuntos inter-relacionados de dados para apoio à pesquisa científica (.....), disponibilização de informações relevantes para diagnóstico médico à distância (telemedicina), bibliotecas digitais e entretenimento digital.

Das aplicações estratégicas acima mencionadas, o entretenimento digital é um bom exemplo dos problemas existentes, e vem ganhando crescente relevância pedagógica, econômica e social. Baseia-se, dentre outras, nas chamadas “informações narrativas”, modalidade bem menos tratada até hoje pela Computação do que a informação descritiva, tipicamente disponível em bancos de dados. Abrange tanto domínios práticos de sistemas gerenciais de informação, como gêneros de ficção. No primeiro caso, fornece apoio, por exemplo, para simulação, tomada de decisão e treinamento; no segundo, para a composição e narração de histórias (“.....”), jogos eletrônicos e geração de conteúdo para TV digital interativa. Soluções envolvem trabalhos em modelos conceituais de tais domínios ou gêneros, desenvolver métodos e implementar sistemas de manipulação de enredos, e ganhar experiência com os requisitos dos diferentes tipos de aplicação.

Resumidamente, alguns dos grandes problemas técnicos e científicos a abordar para fazer frente a este desafio são:

- Redução (abstração e sumarização) das massas de dados por meio de modelagem computacional, simulações e outros;
- Definição e uso da noção de contexto para a recuperação de informação, considerando fatores como localização do usuário, perfil de interesses, objetivos dentre outros;
- Projeto e implementação de descritores de conteúdo multimodal e algoritmos para extração e indexação desses descritores, permitindo buscas multimodais;
- Utilização de estruturas de indexação dinâmicas e distribuídas do tipo
- Estudos em modelos e mecanismos de conciliação e integração de dados altamente heterogêneos;
- Consideração, no armazenamento e recuperação, de fatores inerentes à heterogeneidade na aquisição de dados tais como fatores temporais e culturais, mas também tecnológicos, como sensores, celulares, PDAs (i.e., personal digital assistant), dentre outros;

- Estudo de formas alternativas de disponibilização da informação, incluindo pesquisa em novos tipos de interfaces;
- Tratamento da confiabilidade e validade dos dados e da propriedade intelectual;
- Formulação de modelos conceituais para especificar domínios ou gêneros envolvidos em aplicações de entretenimento digital, desenvolvimento de métodos e implementação de sistemas de manipulação de enredos de narrativas e seu uso experimental em aplicações diversas;
- Estudo de infra-estruturas adaptáveis e inteligentes para o processamento distribuído de informações;
- Estudo de técnicas e métodos que garantam a persistência de dados e informações por longos períodos de tempo, para fins de arquivamento histórico;
- Desenvolvimento de modelos, estruturas e algoritmos para construção de grandes bibliotecas digitais distribuídas, para gerenciamento da informação multimídia.

3. Modelagem computacional de sistemas complexos artificiais, naturais e sócio-culturais e da interação homem-natureza.

O termo *computacional* foi criado para contrastar com *analítico*, sendo usado em todo o mundo para designar modelos, algoritmos e ferramentas computacionais para solução de sistemas complexos de diferentes naturezas. Isto permite, por exemplo, estudos de desempenho de grandes sistemas computacionais, sem a necessidade de implementá-los (por exemplo, simulando o comportamento de redes de computadores com milhões de nós). Além disso, possibilita que cientistas em outros domínios do conhecimento possam investigar problemas que até recentemente não poderiam ser tratados – pelo excesso de dados manipulados, pela ausência de soluções analíticas, ou pela inviabilidade de estudá-los em laboratórios. Exemplos são estudos de genômica, de processos bioquímicos, da física de partículas, de mudanças climáticas, de fenômenos epidemiológicos ou de interações sociais com milhões de participantes (por exemplo, em comunidades digitais ou na Web). Modelagem e simulação computacionais permitem também redução de custos e avanços no setor industrial, executando experimentos virtuais em substituição à construção de uma infraestrutura física. O objetivo deste desafio é criar, avaliar, modificar, compor, gerenciar e explorar modelos computacionais para todos esses domínios e aplicações.

A modelagem matemática de fenômenos se baseia em princípios básicos (por exemplo, da física, da química, da estatística ou da matemática), em que a resolução das equações é o principal complicador. Já a modelagem computacional envolve um grande conjunto de algoritmos e técnicas de simulação, manipulação de dados, mineração de dados, dentre outros, em que o modelo é um dos produtos da própria pesquisa, sendo interpretado como um processo computacional que filtra, transforma, aglutina e gera dados. Isto envolve freqüentemente cooperação entre cientistas da Computação e de outros domínios. Normalmente, na modelagem computacional há incertezas sobre o próprio modelo, pois esses envolvem um grande número de parâmetros que devem ser explorados e ajustados.

A complexidade deste tipo de pesquisa aumenta à medida que crescem o volume de dados e/ou os parâmetros a serem considerados. Outro fator complicador é a necessidade freqüente de combinar vários domínios do conhecimento. Um exemplo típico disso, adaptado de [1], é o seguinte. “O estudo das bases biológica e ambiental de doenças respiratórias requer um esforço multidisciplinar complexo de modelagem, que combina dados de Ciências Sociais e de Saúde Pública a experimentos com modelos de dinâmica dos fluidos aplicados ao fluxo de ar e substâncias inalantes (fumaça, alergênicos). Requer, ainda, modelos das propriedades de materiais, modelos biofísicos de movimentos de *partículas* para ejetar corpos estranhos e modelos biológicos de suscetibilidade genética a doenças.” Já outro exemplo, mencionado por Câmara [2], na área de estudos globais do meio ambiente, mostra a necessidade de integrar diferentes modelos – mudanças climáticas, sistemas humanos e naturais, desenvolvimento sócio-econômico e emissões de concentração de gases e poluentes.

Esta área tem atraído atenção crescente, por causa dos benefícios econômicos e sociais a ela associados. Vários fatores vêm contribuindo para avanços neste tipo de pesquisa. Um deles é a crescente disponibilidade de dados coletados sobre fenômenos naturais ou artificiais, graças à disseminação do uso de sensores, cujas redes envolvem um sem-número de desafios de pesquisa. Outro fator é a evolução no hardware e em sistemas de software, que estão permitindo cálculos de grande complexidade, graças à adoção de processamento paralelo e grades computacionais – o que se costuma classificar como Avanços científicos em outras áreas do conhecimento, desde as ciências exatas e engenharias até as ciências humanas e sociais, vêm igualmente permitindo refinamento nos modelos e algoritmos.

Se todos estes fatores impulsionam a pesquisa em modelagem computacional de sistemas complexos, também trazem problemas para o seu desenvolvimento, pelo volume de dados gerados e pela complexidade crescente do processamento demandado. Por exemplo, o conjunto diário de dados recebidos de satélites e diferentes tipos de sensores já atinge petabytes (10^{15}), exigindo pesquisa em técnicas para redução de dados. Sistemas de busca na Web, como o Google, demandam algumas centenas de milhares de computadores trabalhando simultaneamente para conseguir responder a todas as consultas em um tempo aceitável, gerando assim demanda de pesquisa em simulação do desempenho de redes. Estudos sobre o funcionamento das células do corpo humano prevêem uma possibilidade de 10^{18} doenças distintas. Todos esses exemplos exigem pesquisa de ponta em arquiteturas avançadas de alto poder computacional. Desta forma, ao mesmo tempo que a permite estudos antes impossíveis em vários domínios, os avanços destes últimos geram desafios de pesquisa em Computação.

Vários níveis de desafios específicos relacionados com a modelagem computacional, podem ser citados, incluindo:

- Processamento em tempo real de altíssimas taxas de fluxos de dados provindos de milhares de sensores – por exemplo, em estudos de catástrofes naturais (por exemplo, enchentes) ou sistemas de evacuação urbanos em face de desastres artificiais (por exemplo, acidentes nucleares). Isto exige pesquisa em, por exemplo, distribuição do processamento, novas arquiteturas de bancos de dados, suporte de rede para comportar este fluxo;
- Criação de novos tipos de métodos de extração de requisitos, algoritmos e mecanismos de coleta e processamento de dados que capturem variáveis relativas a interações sociais e sócio-culturais;
- Definição de estruturas de armazenamento para registrar os modelos computacionais e os fatores associados à sua sintonia e execução em um ambiente paralelo e distribuído;
- Desenvolvimento de ferramentas para a construção colaborativa de modelos, apoio à sua execução e modificação adaptativa em tempo real, para que tal execução leve em conta mudanças ocorridas no mundo real enquanto o modelo é executado;

- Criação de novos algoritmos e técnicas em visualização científica, para permitir capturar de forma visual a complexidade dos objetos modelados e de suas interações – por exemplo, ajudando a entender a dinâmica de um tornado ou a evolução da erosão causada pela ocupação humana inadequada em uma região;
- Pesquisa sobre o impacto, em engenharia de software, resultante da necessidade da colaboração entre cientistas da Computação e cientistas dos demais domínios;
- Tratamento dos problemas de aumento na dimensionalidade (número crescente de variáveis em um problema) e escalabilidade, que contribuem para tempos de processamento exponencial;
- Utilização de processamento paralelo envolvendo recursos heterogêneos como nas grades (grid) computacionais;
- Estudos em interfaces multimodais extensíveis para permitir o entendimento dos fenômenos modelados e facilitar a configuração dinâmica dos modelos; e
- Integração de algoritmos, estruturas de dados e modelos criados por diferentes disciplinas usando metodologias próprias.

4. Impactos para a área de computação da transição do silício pa

computação quântica ou biológica, interagindo com processadores baseados em silício. Esta heterogeneidade de processadores e suas interações são outros exemplos de problemas a serem enfrentados.

Quais as consequências desta evolução no hardware? Quando falamos em computadores, geralmente associamos a noção de programa à máquina concreta na qual esses programas executam, os computadores baseados em chips de silício. Mas as bases teóricas da Ciência da Computação são independentes de máquinas concretas. As máquinas que chamamos de computadores são uma possível implementação de alguma máquina capaz de executar esse trabalho. Existem várias propostas de outras máquinas que podem executar esse trabalho, mas até agora elas não passaram de propostas acadêmicas. Os grandes avanços na área de Biologia Molecular da última década podem mudar esse panorama, dando à Ciência da Computação uma máquina completamente diferente: uma máquina biológica. Os organismos vivos são sistemas reativos: dependendo de estímulos (condições ambientais e presença ou ausência de certas substâncias), a célula produz uma reação (produzindo as proteínas/enzimas adequadas para tratar o estímulo). O DNA contido no núcleo de nossas células contém informações precisas sobre as seqüências de bases que geram cada proteína, e como e quando essas proteínas devem ser produzidas. Fazendo uma analogia, poderíamos ver o DNA como o software (englobando dados e programa), e a célula como o hardware biológico. Um exemplo

outro, está o projeto conjunto e otimizado do software e do hardware, necessário para obter a eficiência energética e compensar a baixa confiabilidade das novas tecnologias.

É evidente que esse novo cenário exige uma readequação completa dos modelos e técnicas de projeto em muitas sub-áreas tradicionais da Computação. Em especial, novos paradigmas deverão ser buscados em diferentes linhas de pesquisa. Na Engenharia de Software, por exemplo, será necessário por exemplo buscar abstrações adequadas à necessidade de otimização no nível físico, ou desenvolver novas técnicas de teste e verificação de software na presença de falhas. Em Sistemas Operacionais, será preciso incluir tarefas como o controle do consumo de energia, e permitir alternar serviços entre hardware e software para minimização energética. Novos compiladores irão requerer maior visibilidade de hardware não padronizados. O Processamento Paralelo exigirá pesquisa em combinar eficiência energética com os requisitos de alto desempenho, além de novas linguagens que explorem o paralelismo de forma eficiente nesse novo contexto. Neste cenário, todo sistema de computação deverá ser visto como um sistema embarcado, com potencial interconexão com os demais recursos num sistema pervasivo.

Ao mesmo tempo, a complexidade de plataformas paralelas e de sistemas altamente distribuídos está crescendo cada vez mais. Sistemas paralelos futuros terão centenas de milhares de processadores trabalhando em conjunto. Desta forma, um desafio associado é o desenvolvimento de software que seja escalável neste nível e que aproveite ao máximo a arquitetura disponível para garantir o nível de concorrência adequado. Adicionalmente, a maioria dos equipamentos para aplicações embarcadas e pervasivas será movida a bateria. A capacidade dessas não tem aumentado de maneira exponencial, provocando um gargalo na energia que pode ser consumida.

Resumidamente, alguns dos grandes problemas técnicos e científicos a abordar para fazer frente a este desafio são:

- Desenvolver pesquisa em Engenharia de Software para permitir capturar requisitos em todos os níveis de abstração, desde os níveis mais baixos de desempenho e consumo de energia, até as abstrações necessárias à automação crescente do processo de desenvolvimento de software;
- Definir conceitos mais abrangentes de desenvolvimento baseado em componentes, que levem em consideração também características não-funcionais, como desempenho, consumo de energia e confiabilidade;
- Desenvolver novas técnicas de teste e verificação de software na presença de falhas permanentes e transitórias, e que considerem as interações entre processadores não convencionais e processadores baseados em silício;
- Projetar mecanismos em que a implementação de tarefas do sistemas operacionais e middleware possam alternar entre software e hardware, minimizando e controlando o consumo de energia, num contexto de carga variável de trabalho;
- Projetar técnicas de compilação que permitam a compiladores mapear

automaticamente fragmentos de código para módulos altamente otimizados de hardware;

- Combinar eficiência energética com os requisitos de alto desempenho em sistemas paralelos e heterogêneos;
- Desenvolver novas linguagens que explorem o paralelismo de forma eficiente no contexto de sistemas massivamente paralelos e heterogêneos, onde o consumo de energia é uma restrição importante a ser considerada;
- Integrar os vários níveis de paralelismo, desde o paralelismo (.....), o paralelismo , até o paralelismo inter-nós (.....);
- Prover suporte ao desenvolvimento de sistemas escaláveis, através de monitoração e depuração;
- Investigar novos modelos de computação, bem como novas arquiteturas e máquinas capazes de implementá-los;
- Investigar a possibilidade e as perspectivas da utilização de mecanismos biológicos para a resolução de problemas;
- Criar linguagens e métodos para o desenvolvimento de programas em máquinas que implementam modelos não-convencionais de computação.

5. Acesso participativo e universal do cidadão brasileiro ao conhecimento.

A Tecnologia da Informação introduziu uma revolução na comunicação entre pessoas e sua forma de trabalho. Através de redes de computadores, todos podemos nos comunicar e compartilhar os mais diversos recursos (hardware, dados e software, e também informação visual e sonora), independentemente da localização ou de presença física simultânea. Esses novos tipos de interação são facilitados pela disponibilidade das redes de comunicação com banda larga e latência reduzida, associada aos dispositivos móveis e à computação ubíqua. Tal disponibilidade, no entanto, não é sinônimo de facilidade de uso e acesso universal.

Existem barreiras tecnológicas, educacionais, culturais, sociais e econômicas, que impedem o acesso e a interação. O objetivo deste desafio é, portanto, vencer essas barreiras, por meio da concepção de sistemas, ferramentas, modelos, métodos, procedimentos e teorias capazes de endereçar, de forma competente, a questão do acesso do cidadão brasileiro ao conhecimento. Este acesso deve ser universal e participativo, na medida em que o cidadão não é um usuário passivo, o qual recebe informações, mas também participa da geração do conhecimento. Apenas por meio da oportunidade de participação da construção do conhecimento é que o acesso vai poder levar a uma plena e consciente utilização do conhecimento disponibilizado.

O problema concerne, portanto, a extensão de sistemas computacionais ao cidadão comum, em sua diversidade, respeitando suas diferenças. Este desafio ganha novas proporções no cenário de uma população com a diversidade e os problemas da nossa, com enormes diferenças sócio-econômicas, culturais, regionais e de acesso à tecnologia e ao conhecimento. Além disso, vivemos um momento de convergência de mídias (como a Internet, TV, telefone celular) cujo galgalo será, sem dúvida, possibilitar o acesso.

O IBGE (Instituto Brasileiro de Geografia e Estatística) contabilizou, em 2003, 32,1 milhões de analfabetos funcionais, entendidos como a população com mais de 15 anos de idade e menos de 4 anos de escolaridade (26% da população). Ainda, segundo a mesma fonte, 24,5 milhões de pessoas têm algum tipo de deficiência (14,5% da população). Setores do governo, universidades e o setor privado podem buscar soluções tecnológicas de alcance social que procurem diminuir tais diferenças e resgatar valores de cidadania em nossa sociedade.

Para atacar este problema, é necessário, por exemplo, definir "acesso participativo" e "conhecimento" de forma bastante abrangente. Não se trata somente de capturar, organizar e divulgar informações e conhecimento, nem tampouco somente de reduzir as dificuldades de acesso físico a computadores e redes, seja por limitações de poder aquisitivo, seja por barreiras educacionais ou culturais, ou por incapacidade sensorio-motora. Trata-se também de produzir tecnologia de base computacional que permita e motive a participação dos usuários no processo de produção de conhecimento e decisão sobre seu uso. Além disso, deve-se ter em conta aspectos jurídicos, sociais e antropológicos

da cidadania brasileira, justamente para reduzir o risco de agravar problemas desta ordem, ou até mesmo de criar problemas novos resultantes de sua mera existência, exacerbando a "exclusão digital".

O problema é difícil porque é único; não havendo experiências nas quais possamos nos inspirar. O tratamento deste desafio exige competências multidisciplinares, em que a pesquisa em Computação possibilita estabelecer sistemas e métodos que sustentem a constituição de uma cultura digital para acesso do cidadão ao conhecimento, respeitando sua diversidade e diferenças. Acresça-se esta dificuldade à de viabilizar uma interlocução com outras áreas de conhecimento, como as Ciências Humanas e Sociais, cuja prática e discurso científicos diferem dos da Ciência da Computação. Esta cooperação deve permitir que qualquer cidadão brasileiro possa ter participação e decisão na cadeia de produção de conhecimento e cultura do país.

Várias áreas da Computação precisam contribuir para a solução deste problema. O acesso ao conhecimento começa com vencer os desafios das interfaces, em particular a interface de sistemas computacionais com o cidadão. Outros tópicos de pesquisa relacionados e em aberto envolvem o projeto e desenvolvimento de ambientes com mobilidade e consciência de contexto, com hardware de baixo custo e software abertos e adaptáveis a necessidades locais. A produção de conhecimento passa pela produção de conteúdo e mecanismos eficientes de armazená-lo e recuperá-lo, pela criação e disponibilização de bibliotecas digitais. Questões afetivas, sociais e culturais em computação são novas áreas que precisam também ser envolvidas nesta tarefa.

A ênfase no design de interfaces é uma característica deste desafio. O design para todos e as interfaces flexíveis e ajustáveis são alvo de pesquisa de ponta na própria área de IHC (Interação Humano-Computador), que aumentam à medida que se considera os variados contextos sócio-econômicos e a diversidade cultural do Brasil. Aspectos de interação multimodal também devem ser explorados, envolvendo corpo, olhar, áudio, fala ou gestos, e incluindo o uso de todo tipo de sensor. Psicologia, linguística, antropologia, sociologia e geografia são algumas das áreas necessárias para vencer os obstáculos existentes na pesquisa em interfaces.

Exemplos concretos de aplicação/domínio alvo que seriam beneficiados com a pesquisa neste desafio incluem, dentre outros, sistemas de governo eletrônico, sistemas para aprendizagem continuada, comunidades de prática relacionadas ao trabalho, apoio a comunidades em rede (.....) em áreas carentes. Governo eletrônico é entendido aqui não apenas como a disponibilização de serviços via Internet, mas também a possibilidade de participação do cidadão na geração de conhecimento a ser compartilhado, na discussão de assuntos que lhe afetem diretamente.

Os benefícios da pesquisa nesta área atingem a coletividade, que pode reduzir a distância que temos hoje do maior bem da sociedade contemporânea: a informação, o conhecimento. Ganha, também, a comunidade de Computação no Brasil, com novos instrumentos, artefatos e métodos para desenvolvimento de sistemas e suas interfaces. Ganham as demais disciplinas envolvidas nesse desafio, com uma compreensão maior do problema e dos limites de solução. Ainda, processos de socialização do conhecimento com

o compartilhamento e cooperação dos envolvidos na geração de conhecimento e a formação de agentes são também benefícios potenciais de ações neste desafio

Além da pesquisa em interfaces flexíveis e adaptáveis, alguns outros grandes problemas de pesquisa em Computação a abordar dentro deste desafio são:

- Projeto e desenvolvimento apropriado de novas infraestruturas de hardware e de comunicação;
- Experimentos com conteúdo centrado em aspectos humanos e modelagem de conteúdo, que incluam aspectos de dinâmica social e sistemas socialmente conscientes;
- Criação de sistemas de - infra-estrutura interna necessária para prover serviços aos cidadãos, que pode incluir processos de longa duração, que envolvam várias entidades e questões de interoperabilidade;
- Disponibilização de uma infra-estrutura necessária para a interação direta do cidadão e sua comunicação com o processo a ser realizado no
- Desenvolvimento de mecanismos ágeis para busca de dados e para armazenamento do conteúdo sendo gerado continuamente pelo cidadão;
- Projeto e implementação de estruturas de ontologia flexíveis e expansíveis, que possibilitem interoperabilidade entre áreas do conhecimento e interação entre pessoas de diferentes culturas e vocabulários;
- Criação de plataformas de aplicativos de que permitam a integração eficiente de ferramentas de comunicação para poderem ser usadas em conjunto com a finalidade de permitirem a prática do aprendizado eletrônico; e
- Definição de formas de garantir a administração adequada dos direitos autorais, em particular, e da propriedade intelectual, em geral, de modo a permitir uma ampla variedade de experimentos nas áreas de produção, administração e utilização do conhecimento.
- Projeto e construção de novos dispositivos para permitir acessibilidade universal, por exemplo ajudando usuários com deficiências a interagir com sistemas de software e hardware.

6. Desenvolvimento Tecnológico de Qualidade: sistemas disponíveis, corretos, seguros, escaláveis, persistentes e ubíquos.

A Tecnologia da Informação está cada vez mais presente em nosso cotidiano. Não é preciso ir muito longe para buscar exemplos. Vários eletrodomésticos e elevadores, por exemplo, têm controle via software; carros, tratores, aviões, celulares, sistemas de controle de tráfego e salas de cirurgia também dependem desta tecnologia. Enquanto alguns desses exemplos correspondem a sistemas relativamente simples – um forno de microondas –, outros custam muito caro e envolvem milhões de linhas de código e hardware sofisticado. Se esta ubiquidade traz conforto,

sistemas e eles precisam estar sempre

disponíveis

formar da forma

prevista

g o ponto de vista de

pesquisa em ambientes, técnicas, modelos, dispositivos e

etinos.

software

segurança

arquitetura

Qo para de sistemas e
sua aplicação para os. i atr, varizes. dos

software.

proliferação de tipos e usos desses dispositivos exige cada vez mais atenção ao desenvolvimento de sistemas para sua coordenação e funcionamento confiáveis. De fato, uma vez instalados, sensores deveriam sempre estar disponíveis, sua rede deveria ser escalável (funcionando de forma adequada independentemente do crescimento) e segura (impedindo ataques maliciosos). Finalmente, deveriam funcionar corretamente, ou seja, de acordo com sua especificação. A garantia dessas características ainda faz parte do futuro.

Redes de sensores foram usadas para ilustrar alguns dos problemas que motivam este desafio – elas são ubíquas e combinam problemas de hardware e software que põem em risco sua confiabilidade e disponibilidade. No entanto, tais redes são relativamente novas. Sistemas de software são também ubíquos e apresentam os mesmos problemas, embora já existam há dezenas de anos. Disponibilidade, corretude, segurança, escalabilidade e persistência de software são problemas tanto para aplicações críticas (como as que controlam aviões, centrais nucleares ou estações orbitais) quanto para as não críticas (como sistemas de pagamento de pessoal ou controle de estoque, em que estes problemas, embora complexos, podem ser eventualmente reparados sem por em risco vidas humanas). A necessidade de software que obedeça a todas essas propriedades está revolucionando a Engenharia de Software e disciplinas associadas.

A construção e manutenção de software robusto exigem profissionais qualificados e compromissos de longo prazo, que permitam manter, expandir e gerenciar o software. Praticar engenharia de software fidedigno é algo bem mais amplo do que incorporar tolerância a falhas e assegurar corretude, entre outras propriedades. O entendimento da noção de fidedignidade depende do serviço que o software presta. Em alguns serviços, por exemplo, comando e controle, uma falha pode ter conseqüências catastróficas. Já em outros, por exemplo busca de informação na Web, uma busca incorreta (falsos positivos ou ausência de respostas relevantes) é tolerável desde que isto ocorra com uma freqüência baixa. Entretanto, não é tolerável que um sistema cancele sua execução, invada a privacidade, ou represente um risco à segurança de quem o usa.

Em virtude da crescente participação de software na sociedade, torna-se cada vez mais necessário assegurar sua fidedignidade. Artigos recentes enfatizam os problemas da falta de qualidade em software. É sabido que cerca de 50% do software tornado disponível contém falhas não triviais. A evolução do software é muitas vezes considerada uma propriedade intrínseca, ou seja, virtualmente todo o software em uso tende a evoluir, sendo que isto constitui uma parcela muito significativa do seu custo total (.....). É preciso portanto desenvolver software que possa evoluir sem comprometer a sua qualidade. Finalmente, é sabido que, independentemente do rigor com que seja desenvolvido, todo software conterà falhas. Conseqüentemente, a ocorrência de falhas de uso, de hardware ou do próprio software é um fato com que se precisa conviver sem que, no entanto, estas falhas levem a um nível de danos inaceitável.

Há muito o Brasil, à similaridade com a Índia e a Irlanda, aspira ter uma pujante indústria de software e de serviços de desenvolvimento para exportação. Apesar de vários esforços neste sentido, os resultados ainda são muito pequenos. A melhoria da qualidade do nosso software certamente contribuirá para sua maior aceitação tanto no exterior como localmente. De fato, este é o caminho para garantir o próprio mercado interno.

Este problema na produção de software é difícil, mesmo quando se considera software com arquiteturas convencionais. A tendência é isto se agravar já que cada vez mais o desenvolvimento de software é feito de forma geograficamente distribuída (várias equipes desenvolvendo um mesmo software) e sua operação também é distribuída (várias CPUs contendo partes heterogêneas do software operando sobre uma única aplicação, por exemplo em sistemas . . .). Outra razão para que o problema se torne um desafio crescente é o fato de os sistemas de software tornarem-se cada vez mais complexos (volume e abrangência da funcionalidade e requisitos de qualidade mais restritivos), precisarem estar disponíveis por mais tempo (sistemas 24h / 7dias por semana) e serem utilizados por pessoas sem garantia de treinamento.

Há inúmeros benefícios advindos da pesquisa neste desafio. Vários deles são decorrentes das vantagens resultantes do desenvolvimento de sistemas confiáveis e seguros, contribuindo para a melhoria da qualidade e a redução dos custos de desenvolvimento e uso de software e sistemas computacionais em geral. O mau funcionamento de um sistema computacional pode ter sérias conseqüências. Por exemplo, falhas em sistemas de controle de processos de alto risco podem levar a perda de vidas, desastres ecológicos, quebra de empreendimentos. Em sistemas de comércio eletrônico, falhas podem levar a prejuízos na casa dos milhões de reais. Em sistemas de governo eletrônico, podem contribuir ainda mais para a exclusão social da população.

Falhas na segurança de dados, software ou sistemas acarretam também prejuízos financeiros e sociais, como aqueles causados por ataques a redes e vírus em todo o mundo. A violação da privacidade decorrente deste tipo de falha é outra conseqüência grave para o cidadão, com custos incalculáveis.

Em função dessas observações, alguns dos tópicos de pesquisa associados a este desafio são:

- Desenvolvimento e avaliação de modelos e ferramentas de modelagem de sistemas de software com base teórica sólida;
- Desenvolvimento e adaptação de tecnologias e instrumentos em geral de apoio à implementação e à avaliação de software fidedigno por construção;
- Desenvolvimento de ferramentas de apoio ao processo de implementação e evolução de software;
- Especificação e análise de novas técnicas e algoritmos de segurança de dados e sistemas, incluindo técnicas criptográficas e protocolos seguros de comunicação;
- Construção de mecanismos e ferramentas visando a tolerância a falhas e disponibilidade permanente; e
- Considerar a necessidade de ubiquidade no projeto e desenvolvimento de sistemas, incluindo fatores como ambientes distintos de trabalho e variações em requisitos.

7. Algumas ações propostas

Os participantes do Seminário propuseram algumas ações associadas aos Desafios especificados ao desenvolvimento da pesquisa em Computação no Brasil. Ressalte-se que várias outras ações podem ser derivadas dos desafios propostos.

As propostas principais envolveram os seguintes quesitos:

Multidisciplinariedade – a evolução da pesquisa e desenvolvimento no século XXI aponta para equipes multidisciplinares como sendo uma das formas mais comuns de obter resultados científicos. Desta forma, recomenda-se duas ações: (a) sensibilizar os pesquisadores em Computação para os problemas inerentes à pesquisa multidisciplinar, como estabelecimento de vocabulário comum e entendimento de diferenças metodológicas na pesquisa em cada campo; e (b) desenvolver modelos de ensino e pesquisa “.....” entre áreas, que visem a formação de profissionais e cientistas que possam trabalhar neste novo mundo, com ênfase em multi- e inter-disciplinariedade. Exemplos de aplicações multidisciplinares que poderiam ser usadas neste tipo de formação seriam meio ambiente, saúde pública, violência urbana, agropecuária,, entretenimento digital, telemedicina, história, dentre outros.

Tal multidisciplinariedade deve ocorrer não apenas entre a Computação e outros domínios científicos, mas também dentro da Computação. Por exemplo, especialistas em hardware precisam cooperar com especialistas em redes, em bancos de dados, em interação humano-computador. Todos, por sua vez, devem ter uma interlocução continuada com pesquisadores em engenharia de software e lançar mão dos conhecimentos de pesquisadores em computação gráfica, visualização científica, inteligência artificial e tantas outras áreas associadas à pesquisa necessária à solução dos desafios.

Integração com a indústria – a pesquisa de boa qualidade reverte em benefícios sociais e econômicos. Neste sentido, a principal ação considerada foi aproximação com a indústria para efeitos do desenvolvimento tecnológico de qualidade e a indicação de novas áreas com o potencial de se transformarem em mercados emergentes.

Transformação da fuga de cérebros em vantagem – vários pesquisadores brasileiros estão sendo atraídos por melhores condições de trabalho no exterior. Eles podem ser usados como contato para aumentar a cooperação em pesquisa do Brasil com o resto do mundo – por exemplo, auxiliando a promoção de congressos internacionais no Brasil, facilitando a vinda de professores visitantes estrangeiros, dentre outras possibilidades.

Estabelecimento de um centro para discussão de pesquisa – os presentes constataram que eventos de “brainstorming”, como o organizado, são muito raros. A maioria dos congressos e seminários em todo mundo segue um padrão básico de apresentação de artigos e discussões em torno de temas específicos. A ação proposta é a criação de um centro para encontros científicos à semelhança de Dagstuhl ou Oberwolfach na Alemanha. Estes centros foram criados para oferecer a cientistas da computação condições de desenvolver pesquisa de ponta. Uma das atividades mais conhecidas de Dagstuhl, por exemplo, são as conferências de até uma semana, em Computação, com no máximo 40 participantes, para

discutir e desenvolver ciência de ponta em nível internacional. Cada conferência é proposta com até 3 anos de antecedência a um comitê científico e os participantes pagam apenas os custos de transporte para o local. As acomodações em Oberwolfach servem propósito semelhante nas áreas de Matemática e Teoria da Computação. (ver por exemplo www.dagstuhl.de). .

8. Conclusões

A SBC espera que este evento seja o primeiro de uma série de encontros de Grandes Desafios em Computação no Brasil, visando ao planejamento a longo prazo da pesquisa em Computação no País.

Como mencionado na Introdução, a idéia de formular grandes desafios em pesquisa tem sido adotada por vários países, em diversas áreas do conhecimento. Constatou-se, nesses países, que a formulação de tais desafios tem levado ao estabelecimento de uma agenda de pesquisa de longo prazo, com conseqüências positivas não apenas do ponto de vista de avanços do conhecimento, mas também para a formação de novas gerações de pesquisadores. Desta forma, esta proposta da SBC poderá trazer contribuições significativas para a área de Computação e servir, igualmente, para dar subsídios à formulação das políticas de agências fomento à pesquisa no Brasil

Os desafios propostos exigem cooperação dentro da Computação e com várias outras áreas do conhecimento. Cada um deles engloba um grande conjunto de tópicos de pesquisa de ponta e está centrado em algum aspecto importante da Informática. Várias áreas de pesquisa em Computação aparecem na descrição de todos esses desafios – redes de computadores, gerenciamento de dados, desenvolvimento de software, projeto de algoritmos e estruturas de dados, design de interfaces e novos dispositivos e arquiteturas de computadores são algumas das questões a serem abordadas com maior ou menor ênfase em todos os itens.

Finalmente, a Computação permeia hoje todas as atividades científicas. A SBC espera, assim, que esta iniciativa contribua também para o avanço da ciência brasileira em todas as frentes, influenciando várias outras áreas de pesquisa no País. Relatórios produzidos em todo o mundo apontam para o papel fundamental que a Computação desempenha no progresso científico, tecnológico, econômico e social de uma nação. A busca pela excelência científica em Computação, com ênfase na visão em longo prazo, deverá desta forma ter impacto social e econômico duradouro.

Referências Bibliográficas

[1] Computational Science: Ensuring America's Competitiveness. PITAC Report to the President, EUA, Junho de 2005, Disponível em http://www.nitrd.gov/pitac/reports/20050609_computational/computational.pdf

[2] .G. Câmara. Grandes Desafios da Computação: A construção de uma terceira cultura. Apresentação no Seminário Grandes Desafios, maio de 2006

Programa do evento

Dia 1:

8:30-12:30 Apresentação dos trabalhos selecionados, com uma duração de até 8 minutos por proposta.

14:00-16:00 Mesa Redonda – Renato Janine Ribeiro e Gilberto Camara – A importância do planejamento estratégico de longo prazo, da pesquisa em Computação no Brasil

16:30 – 19:30 Organização de 5 grupos de trabalho, para discussão e consolidação de propostas de grandes desafios, com suas características e oportunidades. Os Grandes Desafios não necessariamente serão derivados dos trabalhos, mas sim da visão consensual de cada grupo sobre os problemas de futuros da pesquisa.

Apresentação dos resultados de cada grupo em cartazes.

Dia 2:

8:30-12:30 Apresentação dos grupos: sumário e consolidação dos grandes desafios propostos no dia anterior

14:00-15:00 Consolidação final dos desafios em discussão, apresentando novas visões e justificativas.

15:00- 17:00 Apresentação final da proposta de Grandes Desafios em Computação para o Brasil 2006-2016 para convidados externos.