

## O Que é Singularidade

Vernor Vinge on Singularity  
Departamento de Ciências Matemáticas  
Universidade de Estado de San Diego

(c) 1993 por Vernor Vinge (Este artigo pode ser reproduzido para finalidades não comerciais se for reproduzido em sua totalidade, incluindo esta observação) A versão original deste artigo foi apresentada no Symposium VISION-21 patrocinado pelo Centro Lewis de Pesquisa da NASA em Ohio e pelo Instituto Aeroespacial de Ohio, 30 e 31 de março de 1993. Uma versão ligeiramente alterada apareceu na edição de Inverno da Whole Earth Review de 1993. TRADUÇÃO PARA O PORTUGUÊS: Mário Porto Nota de TR Sumário Dentro de trinta anos, teremos os meios tecnológicos para criar uma inteligência super-humana. Logo após, a era humana se extinguirá. Seria tal progresso evitável? Se não pode ser evitado, podem os eventos ser orientados de modo a que possamos sobreviver? Estas são as perguntas investigadas. Algumas respostas possíveis (e alguns perigos decorrentes) são apresentadas. O Que é a Singularidade?

A aceleração do progresso tecnológico foi a característica central deste século. Eu discuto neste trabalho que estamos no limiar de uma mudança comparável à ascensão da vida humana na terra. A causa precisa desta mudança é a criação eminente pela tecnologia das entidades com inteligência maior do que a humana. Há diversos meios pelos quais a ciência pode conseguir esta descoberta (e esta é uma outra razão para se ter a confiança que o evento ocorrerá):

- Podem ser desenvolvidos computadores "conscientes" e super-humanamente inteligentes. (Atualmente existem muitas controvérsias a respeito sobre se podemos criar a equivalência humana em uma máquina. Mas, se a resposta for "sim, nós podemos", então restariam poucas dúvidas de que seres mais inteligentes podem ser construídos logo depois disso.)

- As grandes redes de computador (e seus usuários associados) podem "despertar" como uma entidade super-humanamente inteligente.

- As interfaces computador/homem podem tornar-se tão íntimas que os usuários podem razoavelmente ser considerados super-humanamente inteligentes.

- A ciência biológica pode fornecer meios para melhorar o intelecto natural humano As primeiras três possibilidades dependem em grande parte das melhorias no hardware dos computadores. O progresso no hardware dos computadores seguiu uma curva, surpreendentemente, constante nas últimas décadas [17]. Bastante baseado nesta tendência, eu acredito que a criação de uma inteligência maior do que a humana ocorrerá durante os próximos trinta anos. (Charles Platt [20] afirmou que os entusiastas da AI têm feito reivindicações como esta nos últimos trinta anos. Desta maneira, não sou culpado de uma ambigüidade temporal relativa, deixe-me ser mais específico: Ficarei surpreso se este evento ocorrer antes de 2005 ou após 2030.) Quais são as conseqüências deste evento? Quando as inteligências maiores do que a humana impulsionarem este progresso ele será muito mais rápido. Na verdade, não parece existir nenhuma razão para o próprio progresso não envolver a criação de entidades mais inteligentes -- em uma escala de tempo ainda mais curta. A melhor analogia que eu visualizo é com o passado evolutivo: Os animais podem adaptar-se aos problemas e fazer invenções, mas usualmente mais devagar do que o trabalho da seleção natural -- o mundo age como seu próprio simulador no exemplo da seleção natural. Nós seres humanos temos a habilidade de internalizar o mundo e conduzir "what if's" em nossas cabeças; podemos resolver muitos problemas milhares de vezes mais rápido do que a seleção natural. Agora, criando os meios para executar aquelas simulações em uma velocidade muito mais elevada, nós estamos entrando em um regime tão radicalmente diferente de nosso passado humano quanto nós humanos somos dos animais mais elementares. Do ponto humano da vista esta mudança representa desprezar todas as regras anteriores, talvez num piscar de olhos, uma escalada exponencial além de qualquer esperança de controle. Desenvolvimentos que antes eram pensados para somente acontecer em "milhões de anos" (se acontecessem) acontecerão provavelmente no século seguinte. (Em [5], Greg Bear pinta um retrato das mudanças importantes que acontecem numa questão de horas.) Eu penso que é justo classificar este evento como uma singularidade ("A Singularidade" para as finalidades deste trabalho). Este é um ponto onde nossos modelos velhos devam ser rejeitados e uma nova realidade tome lugar. À medida que nos movemos para mais perto deste ponto, ele se tornará mais amplo e mais ilimitado sobre as atividades humanas até que a noção se torne um lugar comum. Contudo quando finalmente acontecer pode ainda ser uma grande surpresa e um grande desconhecido Nos anos 50 havia muito poucos que o viam: Stan Ulam [28] parafraseou John von Neumann dizendo: Uma conversa centrada no sempre acelerado progresso da tecnologia e das mudanças no modo da vida humana, que demonstra aproximarmos-nos de alguma singularidade essencial na história da raça além da qual as atividades humanas, como nós a conhecemos, não poderiam continuar. Von Neumann usa o termo singularidade, embora pareça que está pensando no progresso normal e não na criação do intelecto super-humano. (Para mim, a super-humanidade é a essência da Singularidade. Sem aquela nós chegaríamos à sociedade da riqueza técnica, sem nunca absorvê-la corretamente. (ver [25]).) Nos anos 60 havia um reconhecimento de algumas das implicações da inteligência super-humana. I. J. Good escreveu [11]: Defina-se uma máquina ultra-inteligente como uma máquina que possa superar de longe todas as atividades intelectuais de todo e qualquer homem embora inteligente. Uma vez que o projeto de máquinas é uma destas atividades intelectuais, uma máquina ultra-inteligente poderia projetar até mesmo máquinas melhores; então haveria uma inquestionável "explosão de inteligência," e a inteligência do homem seria deixada bem para trás. Assim a primeira máquina ultra-inteligente é a última invenção que o homem necessitará realizar, contanto que a máquina seja dócil o bastante para nos dizer como mantê-la sob o controle. É mais provável que uma máquina ultra-inteligente será construída dentro do vigésimo século do que não seja, e isso será a última invenção que o homem necessita realizar. Good capturou a essência da escalada, mas não traçou suas conseqüências perturbadoras. Qualquer máquina inteligente do tipo que descreve não seria "ferramenta" da humanidade -- mais do

que seres humanos são as ferramentas dos coelhos, tordos ou chimpanzés. Através dos anos 60 dos 70 e dos 80, o reconhecimento do cataclismo se espalhou [29] [1] [31] [5]. Talvez tenham sido os escritores de ficção científica que sentiram primeiro o impacto concreto. Apesar de tudo, os escritores "duros" de ficção científica são aqueles que tentam escrever histórias específicas sobre o que toda essa tecnologia pode fazer por nós. Cada vez mais, estes escritores sentiam uma parede opaca através do futuro. Uma vez que, poderiam pôr tais de fantasias milhões de anos no futuro [24]. Agora viram que suas extrapolações mais diligentes resultaram num desconhecido... próximo. Uma vez que, os impérios galácticos poderiam parecer um domínio Pós-Humano. Agora, certamente, até mesmo os interplanetários são.

Que dizer sobre os anos "90" os "00" e os "10", à medida que deslizamos para seus limites? Como a aproximação da Singularidade se espalhará através da visão humana do mundo? Por enquanto, todavia, os críticos gerais da sapiência das máquinas exercerão uma boa pressão. Afinal, até que nós tenhamos hardware tão poderoso quanto o cérebro humano é provavelmente bobagem pensar que seremos capazes de criar uma inteligência equivalente à do ser humano (ou até mesmo maior). (Existe a possibilidade remota de que poderíamos fazer um equivalente humano a partir de um hardware menos poderoso, caso estejamos dispostos a abdicar da velocidade, aceitando desenvolver um ser artificial que seria literalmente lento. [30]. Mas, certamente, planejar o software será um processo complicado, envolvendo muitos caminhos e experimentos falsos. Assim sendo, a chegada de máquinas com autoconsciência não acontecerá antes do desenvolvimento do hardware que seja, substancialmente, mais poderoso do que equipamento natural dos seres humanos"). Mas enquanto o tempo passa, devemos observar mais sintomas. O dilema sentido por escritores de ficção científica será percebido em outros esforços criativos. (Eu tenho escutado consagrados autores de estórias em quadrinhos preocupados sobre como obter efeitos espetaculares quando tudo visível possa ser produzido por efeitos tecnológicos comuns). Veremos a automatização substituir mais e mais os níveis de trabalho. Nós já temos as ferramentas (programas de matemática, CAD/CAM) isto nos liberará do trabalho pesado grosseiro. Ou posto de uma outra maneira: O trabalho verdadeiramente produtivo é domínio de uma constantemente menor e mais elitista fração da humanidade. Na chegada da Singularidade, estamos vendo que as previsões do verdadeiro desemprego tecnológico finalmente se tornarão verdadeiras. Um outro sintoma do progresso na direção da Singularidade: as idéias propriamente ditas devem espalhar-se sempre mais rapidamente, e mesmo o desejo o mais radical torna-se rapidamente lugar comum. Quando eu comecei a escrever ficção científica em meados dos anos 60, parecia muito fácil encontrar as idéias que levaram anos para percolar na consciência cultural; agora o tempo de maturação parece algo como dezoito meses. (Naturalmente, isto poderia apenas ser minha perda de imaginação à medida que envelheço, mas eu vejo o mesmo efeito no outros também). Como o choque em um fluxo compressivo, a Singularidade se aproxima quando aceleramos na direção da velocidade crítica. E sobre a chegada da Singularidade propriamente? O que pode ser dito de sua aparência real? Uma vez que envolve uma escalada intelectual, ela ocorrerá, provavelmente, muito mais rapidamente do que qualquer revolução técnica percebida até agora. O evento desencadeador será provavelmente inesperado -- talvez até mesmo para os pesquisadores envolvidos. ("Mas todos nossos modelos precedentes eram catatônicos! Nós apenas ajustávamos alguns parâmetros...") Se as redes estiverem suficientemente espalhadas (em sistemas fortemente universais), pode parecer como se nossos artefatos, como um todo, tenham subitamente enfraquecido. E que acontece um mês ou dois (ou um dia ou dois) em seguida? Eu tenho somente analogias para indicar: A ascensão da humanidade. Nós estaremos na era Pós-Humana. E apesar de todo o meu agressivo otimismo tecnológico, às vezes eu penso que estaria mais tranqüilo se estivesse considerando estes eventos transcendentais mil anos distantes... em vez de vinte. Pode a Singularidade ser evitada? Bem, talvez não aconteça nunca: Às vezes eu tento imaginar os sintomas que devemos esperar ver se a Singularidade não se desenvolver. Existem os argumentos enormemente respeitados de Penrose [19] e de Searle [22] contra a praticabilidade da máquina inteligente. Em agosto de 1992, a Thinking Machines Corporation desenvolveu um workshop para investigar a pergunta "How We Will Build a Machine that Thinks" [27]. Como se pode supor do título do workshop, os participantes não eram especialmente defensores dos argumentos contra a inteligência das máquinas. De fato, havia um acordo geral que as mentes podem existir em carcaças não-biológicas e que os algoritmos são de importância central para a existência das mentes. Entretanto, havia muito debate sobre o poder básico do hardware que está presente no cérebro humano. Uma minoria sentia que os 1992 maiores computadores estavam dentro de três ordens de grandeza do poder do cérebro humano. A maioria dos participantes concordou com a estimativa de Moravec [17] que nós estávamos de dez a quarenta anos distantes da paridade no hardware. No entanto, havia uma outra minoria que apontava [7] [21], e conjecturava que a competência computacional de um simples neurônio pode ser muito mais alta do que geralmente se acreditava. Se assim for, os hardwares dos computadores atuais pôde ser algo como dez ordens de grandeza menores do que o valor do equipamento que nós carregamos por aí em nossas cabeças. Se isto for verdadeiro (ou quanto a isto, se a crítica de Penrose ou de Searle for válida), poderemos nunca presenciar uma Singularidade. Ao invés disto, no início dos anos "00" nossas curvas de desempenho do hardware começariam a nivelar-se devido à nossa inabilidade em automatizar o trabalho de projeto necessário para impulsionar melhorias adicionais no hardware. Acabariamos com algum computador bastante potente, mas sem capacidade de melhorá-lo. O processamento comercial de sinais digitais pode ser fantástico dando uma aparência analógica até mesmo às operações digitais, mas nada "acordaria" e nunca haveria a escalada intelectual que é a essência da Singularidade. Provavelmente, isto seria visto como uma idade dourada e seria também o fim do progresso. Isto combina muito como o futuro predito por Gunther Stent. Na verdade, na página 137 de [25], Stent cita, explicitamente, o desenvolvimento da inteligência transhumana como uma condição suficiente para quebrar suas projeções. Mas se a Singularidade tecnológica puder acontecer, ela virá. Mesmo se todos os governos do mundo compreenderem a "ameaça" e estiverem com medo mortal dela, o progresso na direção do objetivo continuaria. Na ficção, surgiram muitas histórias de leis sendo aprovadas proibindo a construção "de uma máquina com aparência da mente humana" [13]. De fato, a vantagem competitiva econômica, militar, e mesmo artística de cada avanço na automação é tão significativa que aprovar leis ou se estabelecerem costumes que proibam tais coisas, meramente, assegura que alguma outra pessoa as iniciará

pioneiramente. Eric Drexler [8] forneceu introspecções espetaculares sobre como quão distante o avanço técnico pode alcançar. Ele concorda que as inteligências super-humanas estarão disponíveis no futuro próximo e que tais entidades representam uma ameaça ao "status quo" humano. Mas Drexler discute que nós podemos confinar tais dispositivos transhumanos de modo que seus resultados possam ser examinados e usados com segurança. Esta é a máquina ultra-inteligente de I.J.Good, com uma dose de cautela. Eu argumento que o confinamento é intrinsecamente impraticável. Para o caso do confinamento físico: Imagine-se trancado em sua casa com somente acesso limitado aos dados externos, a seus padrões. Se aqueles padrões pensarem em uma taxa digamos, um milhão de vezes mais lenta do que você, existe poucas dúvidas de que em um período de anos (seu tempo) você poderia desenvolver uma "sugestão útil" que eventualmente o libertaria. (Eu chamo esta forma "pensante rápida" de superinteligência de "super-humanidade fraca". Uma entidade tão "fracamente super-humana" queimar-se-ia provavelmente em algumas semanas do tempo exterior. A "super-humanidade forte" seria mais do que alavancar a velocidade do relógio de uma mente humana equivalente. É difícil dizer precisamente como seria "a super-humanidade forte", mas a diferença parece ser profunda. Imagine fazer funcionar uma mente de um cão com velocidade elevada. Mil anos da vida do cão proporcionariam alguma introspecção humana? (Agora se a mente do cão for inteligentemente religada e então posta a funcionar na alta velocidade, poderemos ver algo diferente...) Muitas especulações sobre a superinteligência parecem ser baseadas no modelo super-humano fraco. Eu acredito que nossas melhores suposições sobre o mundo Pós-Singularidade podem ser obtidas pensando na natureza da super-humanidade forte. Eu retornarei a este ponto mais tarde neste trabalho.) Uma outra aproximação ao confinamento é construir regras na mente da entidade super-humana criada (por exemplo, leis de Asimov [3]). Eu penso de que qualquer regra rígida o bastante para ser eficaz vai também produzir um dispositivo cuja habilidade seria claramente inferior às versões sem limitações (e assim a competição humana favoreceria o desenvolvimento daqueles modelos mais perigosos). Ainda assim, o sonho de Asimov é maravilhoso: Imaginar um escravo disponível, que tenha 1000 vezes suas potencialidades em todos os sentidos. Imaginar uma criatura que poderia satisfazer a cada desejo seguro seu (o que quer que isto signifique) e ainda ter 99.9% de seu tempo livre para outras atividades. Haveria um novo universo que nós nunca realmente compreenderemos, mas repleto de deuses benevolentes (embora um dos meus desejos seja poder me transformar em um deles). Se a Singularidade não puder ser evitada ou confinada, quanto pode ser ruim a era Pós-Humana? Bem...consideravelmente ruim. A extinção física da espécie humana é uma possibilidade. (Ou como Eric Drexler a pôs em termos de nanotecnologia: Considerando tudo que tal tecnologia pode fazer, talvez os governos simplesmente decidam que já não necessitam cidadãos!). Contudo a extinção física pode não ser a possibilidade a mais tenebrosa. Outra vez, analogias: Pensemos nas maneiras diferentes que nós nos relacionamos com os animais. Alguns dos abusos físicos cruéis são implausíveis, contudo... Em um mundo Pós-Humano haveria ainda uma abundância de nichos onde a automatização equivalente do ser humano seria desejável: sistemas embutidos nos dispositivos autônomos, programas demos autoconscientes nas funções inferiores das consciências maiores. (Uma inteligência fortemente super-humana seria, provavelmente, uma Sociedade da Mente [16] com alguns componentes muito competentes.) Alguns destes equivalentes humanos poderiam ser usados para nada mais do que processar sinais digitais. Seriam mais como baleias do que seres humanos. Outros poderiam ter uma aparência muito humana, contudo com um lado, uma faceta, que os colocaria em um hospital mental em nossa era. Embora nenhuma destas criaturas pudessem ser considerados seres humanos de carne e osso, seriam as coisas as mais próximas no novo ambiente ao que nós chamamos de ser humano hoje. (I.J. Good teve algo a dizer sobre isto, embora nesta data tardia o conselho possa ser duvidoso: Good [12] propôs "uma Regra Dourada", que pode ser parafraseada como "ameace seus inferiores como você seria ameaçado por seus superiores." É uma idéia maravilhosa, paradoxal (e a maioria de meus amigos não acredita) uma vez que o payoff do jogo teórico é tão duro de articular. Contudo se nós pudermos segui-la, em algum sentido, ela pode dizer algo sobre a plausibilidade de tal gentileza neste universo). Eu argumentei acima que nós não podemos impedir a Singularidade, que sua vinda é uma consequência inevitável da competitividade natural humana e das possibilidades inerentes na tecnologia. No entanto... nós somos os iniciadores. Mesmo a maior avalanche é provocada por coisas pequenas. Nós temos a liberdade para estabelecer circunstâncias iniciais, fazer as coisas acontecerem nas formas que são menos prejudiciais do que outras. Naturalmente (como no início das avalanches), pode não estar claro qual realmente é o exato começo:

Outros caminhos da Singularidade: Amplificação da Inteligência (IA) Quando as pessoas falam em criar inteligências seres super-humanamente inteligentes, estão geralmente imaginando um projeto de AI. Mas como frisei no começo deste trabalho, há outros caminhos para a super-humanidade. As redes de computador e as interfaces homem-computador parecem mais mundanas do que a AI, no entanto podem conduzir à Singularidade. Eu chamo esta abordagem contrastante de Amplificação da Inteligência (IA). IA é algo que está acontecendo muito naturalmente, na maioria dos casos nem mesmo está sendo reconhecido por seus desenvolvedores. Mas cada vez que nossa habilidade de acessar informação e de comunicar isto com outros são melhorados, em certo sentido nós conseguimos um aumento sobre a inteligência natural. Mesmo hoje, a equipe de uma estação de trabalho operada por PHD humanos (mesmo uma estação de trabalho fora de rede!) poderia provavelmente maximizar qualquer teste de inteligência existente. E é muito provável que IA seja uma estrada muito mais fácil para o alcance da super-humanidade do que a AI pura. Nos seres humanos, os problemas de desenvolvimento mais difíceis já foram resolvidos. Construir de dentro de nos mesmos deverá ser bem mais fácil do que descobrir antes o que nós realmente somos e então construir máquinas que representem aquilo. E há pelo menos um precedente conjectural para esta abordagem. Cairns-Smith [6] especulou que a vida biológica pode ter começado como um acessório para vidas ainda mais primitivas baseadas no crescimento cristalino. Lynn Margulis (em [15] e em seguida) produziu argumentos fortes que o mutualismo é uma grande força orientadora da evolução. Note-se que eu não estou propondo que a pesquisa em AI seja ignorada ou menos financiada. O que vai para AI terá frequentemente aplicações em IA, e vice-versa. Eu estou sugerindo que nós reconheçamos que na pesquisa das redes e das interfaces existe algo tão profundo (e potencialmente intocado) quanto a inteligência artificial

(AI). Com essa introspecção, nós podemos enxergar projetos que não são tão diretamente aplicáveis como interfaces convencionais e como projeto de redes, mas que servem para que avancemos na direção da Singularidade ao longo do caminho da IA. Eis aqui alguns projetos possíveis que assumem uma significância especial, do ponto de vista da IA:

- Equipe de automação homem/computador: Imagine os problemas que são considerados normalmente de solução puramente mecânica (como problemas "hill-climbing"), e problemas de projeto e interface que utilizam a vantagem da intuição humana e o hardware disponível. Considerando todo o bizarrismo dos maiores problemas dimensionais tipo "hill-climbing" (e dos algoritmos puros que tenham sido planejados para sua solução), poderiam existir ferramentas muito interessantes de controle e apresentação disponíveis ao membro humano da equipe. (NT-Hill climbing é um método de busca (local) que utiliza um procedimento de melhora iterativa (iterative improvement)).
- Desenvolver s simbiose homem/computador na arte: Combinar a potencialidade gráfica da geração de máquinas modernas e a sensibilidade estética dos seres humanos. Naturalmente, tem havido uma quantidade enorme de pesquisa no projeto de periféricos de computadores para artistas, como ferramentas para economia de trabalho. Eu estou sugerindo que nós no orientemos explicitamente para uma maior fusão de competência, reconheçamos explicitamente a abordagem cooperativa que é possível. Karl Sims [23] fez um trabalho maravilhoso nesta direção.
- Permitir equipes humanas/computador em torneios de xadrez. Já temos os programas que podem jogar melhor do que quase a maioria dos seres humanos. Mas quanto trabalho foi feito na direção desta facilidade poder ser usada por um ser humano, para obter resultados ainda melhores? Se fosse permitido a tais equipes participarem em pelo menos alguns torneios de xadrez, isto poderia ter o efeito positivo na pesquisa de IA que a permissão de computadores nos torneios teve para o nicho correspondente na AI.
- Desenvolver interfaces que permitam o acesso do computador e de rede sem requerer que o ser humano esteja amarrado a um ponto, sentando-se na frente de um computador. (Este é um aspecto de IA que se ajusta muito bem com as vantagens econômicas conhecidas, que vários esforços já estão sendo gastos nele).
- Desenvolver uns sistemas de suportes à decisão mais simétricos. Uma área pesquisa/produto popular em anos recentes têm sido os sistemas de sustentação da decisão. Esta é uma forma de IA, mas pode estar demasiadamente focada em sistemas que são dogmáticos. Tanto quanto o programa dando informação ao usuário, deve existir a idéia do usuário dando orientação ao programa.
- Usar redes locais para criar equipes humanas que realmente trabalham (i.e. são mais eficazes do que seus membros componentes). Esta é geralmente a área de "groupware", que já é um objetivo comercial muito popular. A mudança no ponto de vista aqui seria considerar a atividade do grupo como um organismo combinação. Em um sentido, esta sugestão pôde ser considerada como o objetivo de inventar "regras de ordem" para tais operações da combinação. Por exemplo, o foco do grupo pode mais facilmente ser mantido do que em reuniões clássicas. A expertise de membros humanos individuais poderia ser isolada das manifestações do ego de maneira que a contribuição de membros diferentes esteja focada no projeto da equipe. E, naturalmente, bases de dados compartilhadas poderiam ser usadas muito mais convenientemente do que em operações convencionais de comitê. (Note-se que esta sugestão é orientada visando mais operações da equipe do que reuniões políticas. Em um ajuste político, a automatização descrita acima reforçaria simplesmente o poder das pessoas que fazem as regras!).
- Explorar a Internet mundialmente como uma ferramenta de combinação homem/máquina. De todos os artigos na lista, o progresso neste item está fluindo mais rápido e pode nos levar a Singularidade antes de qualquer outra coisa. O poder e a influência da Internet, mesmo nos dias presentes, é amplamente subestimada. Por exemplo, eu penso que nossos sistemas computadorizados contemporâneos quebrariam sob o peso de sua própria complexidade se não fosse pela vanguarda que a USENET "group mind" dá à administração do sistema e ao suporte das pessoas. A enorme anarquia do desenvolvimento da Net mundialmente é evidência de seu potencial. À medida que a conectividade, largura de banda, tamanho dos arquivos e velocidade dos computadores, tudo isto aumentar, nós estaremos vendo algo como a visão de Lynn Margulis [15] da biosfera como processador de dados recapitulada, mas em uma velocidade um milhão vezes maior e com milhões de agentes humanos inteligentes (nós mesmos). Os exemplos acima ilustram a pesquisa que pode ser feita dentro do contexto de departamentos contemporâneos de informática. Há outros paradigmas. Por exemplo, muito do trabalho na inteligência artificial e em redes neurais beneficiar-se-iam de uma conexão mais próxima com a vida biológica. Em vez simplesmente de tentar modelar e compreender a vida biológica com computadores, a pesquisa poderia ser dirigida para a criação dos sistemas compostos que confiam na vida biológica para a orientação ou para fornecimento de características as quais não compreendemos bem o bastante ainda para implementar nos hardwares. Um sonho de longo prazo da ficção científica têm siso a ligação direta do cérebro com interfaces de computador [2] [29]. Na verdade, há trabalho concreto que pode ser feito (e está sendo feito) nesta área:
  - amputados protetizados (Limb prosthetics) é um tópico da aplicabilidade comercial direta. Do nervo aos transdutores do silicone pode ser feito [14]. Esta é uma excitante etapa quase-final na direção de uma comunicação direta.
  - As ligações diretas em cérebros parecem praticáveis, se a taxa de bits for baixa: dada a flexibilidade de aprendizagem humana, o poder real dos alvos do neurônio do cérebro pode não ter sido selecionada precisamente. Mesmo 100 bits por segundo seriam de grande uso para atacar vítimas que seriam de outra maneira confinadas às relações direcionadas por menu.
  - Ligações ao tronco ótico têm o potencial para larguras de banda de 1 Mbit/segundo mais ou menos. Mas para isto, nós necessitamos saber que a arquitetura em escala fina da visão, e nós necessitamos colocar uma rede enorme dos elétrodos com precisão acurada. Se nós quisermos elevar a largura de banda da nossa conexão somando-se aos trajetos já que estão presentes no cérebro, o problema torna-se vastamente mais difícil. Apenas colocar uma grade de receptores de elevada largura de banda no cérebro, certamente, não funcionará. Mas suponhamos que a grade de largura de banda esteja presente enquanto a estrutura do cérebro realmente se forma à medida que o embrião se desenvolve. Isso sugere:

- Experiências com embrião de animais. Eu não esperaria nenhum sucesso de IA nos primeiros anos de tal pesquisa, mas dar o acesso aos cérebros em desenvolvimento a estruturas neurais simuladas complexas pode ser muito interessante para as pessoas que estudam como o cérebro embrionário se torna. No longo prazo, tais experiências poderão produzir animais com os caminhos adicionais dos sentidos e habilidades intelectuais interessantes.

Originalmente, eu esperava que esta discussão de IA rendesse algumas aproximações mais seguras para a Singularidade. (Apesar de tudo, a IA permite nossa participação em um tipo de transcendência.) Infelizmente, recordando sobre estas propostas de IA, apesar de tudo, estou certo do que deva ser considerado e do que pode nos possa oferecer mais opções. Mas como segurança... também, algumas das sugestões são um pequeno susto em sua face. Um de meus revisores informais anotou que IA para seres humanos individuais cria uma elite bastante sinistra. Nós seres humanos temos milhões dos anos da bagagem evolucionária que nos faz considerar a competição como uma visão mortal. Muito dessa fatalidade não pode ser necessária no mundo de hoje, um mundo no qual os perdedores assumem os truques dos vencedores e são cooptados nos empreendimentos vencedores. Uma criatura que fosse construída de novo pôde possivelmente ser uma entidade muito mais benigna do que uma com uma semente baseada em dentes e garras. Mesmo a visão igualitária de uma Internet que acorde junto com toda a humanidade pode ser vista como um pesadelo [26]. O problema não é simplesmente que a Singularidade representa a passagem da humanidade de um estágio central, mas que ela contradiz nossas noções mais profundas do que é um ser. Eu penso que um olhar mais próximo na noção da super-humanidade forte pode mostrar porque é assim. Super-humanidade forte e o melhor que nós podemos esperar. Suponhamos que poderíamos costurar a Singularidade. Suponhamos que poderíamos alcançar nossas mais extravagantes esperanças. O que então pediríamos: Que os seres humanos propriamente se transformassem em seus próprios sucessores, que qualquer injustiça que ocorresse seria moderada pelo conhecimento de nossas raízes. Para aqueles que remanesceram inalterados, a meta seria um tratamento benigno (talvez mesmo dar aos retardatários a aparência de serem mestres de bons escravos à semelhança de Deus). Seria uma era dourada que também envolve o progresso (overleaping Stent's barrier). A Imortalidade (ou pelo menos uma vida tão longa quanto possamos fazer o universo sobreviver [10] [4]) seria alcançável. Mas neste mundo brilhante e mais amável, os problemas propriamente filosóficos tornar-se-iam intimidantes. Uma mente que permaneça com a mesma capacidade não pode viver para sempre; depois que alguns milhares anos pareceria mais como uma fita repetindo em "loop" do que uma pessoa. (O retrato mais frio eu vi disto está em [18].) Para viver indefinidamente, a mente propriamente deve crescer... e quando se torna grande bastante, e olha para trás... que sentimento pode ele ter com a alma original? Certamente o ser posterior seria tudo que o original era, mas muito mais amplo. E assim mesmo para o indivíduo, a noção de Cairns-Smith ou de Lynn Margulis da vida nova crescer incrementalmente fora da antiga seja ainda válida. Este "problema" sobre a imortalidade aparece em modos mais diretos. A noção do ego e da autoconsciência foi o alicerce do racionalismo obstinado dos últimos séculos. Contudo agora a noção da autoconsciência está sob o ataque do pessoal da inteligência artificial ("autoconsciência e outras ilusões"). A Amplificação da Inteligência diminui nosso conceito do ego de uma outra direção. O mundo da pós-Singularidade envolverá redes de elevada largura de banda. Uma característica central de entidades fortemente super-humana será provavelmente sua habilidade de comunicar-se em larguras de faixa variáveis, incluindo algumas bem mais rápidas do que o discurso falado ou mensagens escritas. O que acontece quando as partes de ego podem ser copiadas e mescladas, quando o nível da autoconsciência pode crescer ou encolher para se ajustar à natureza dos problemas analisados? Estas são características essenciais da super-humanidade forte e da Singularidade. Pensando sobre elas, começa-se a sentir como, essencialmente, estranha e diferente será a era Pós-Humana - não importa quão inteligente e benigna ela venha a ser. De um ângulo, a visão se encaixa em muitos de nossos sonhos mais felizes: um momento interminável, onde nós podemos verdadeiramente conhecer um outro e compreender os mistérios mais profundos. De um outro ângulo, se parece muito com o pior cenário que imaginei no início deste trabalho. Qual é o ponto de vista válido? Na verdade, eu penso que a era nova é simplesmente demasiado diferente para se encaixar nas molduras de bom e mau. Que a moldura está baseado na idéia de mentes isoladas e imutáveis conectadas pelas tênues ligações de banda-baixa. Mas o mundo Pós-Singularidade se ajusta com a enorme tradição de mudança e cooperação que se iniciou à bastante tempo (talvez mesmo antes da ascensão da vida biológica). Eu penso que existem noções da ética que se aplicariam em tal era. A pesquisa em IA e em comunicações em elevadas larguras de banda deve melhorar esta compreensão. Eu vejo agora apenas os relances disto [32]. Há uma boa Meta Regra de Ouro; talvez existam regras para distinguir o ego tendo como base a largura de banda da conexão. E quando a mente e o ego serão vastamente mais transiente do que no passado, muito do que damos valor (conhecimento, memória, pensamento) não precisa nunca ser perdido. Eu penso que Freeman Dyson está certo quando diz [9]: "Deus é no que a mente se transforma quando tiver passado além da escala de nossa compreensão." [Desejo agradecer John Carroll da Universidade de Estado do San Diego e Howard Davidson de Sun Microsystems por discutir o esboço deste trabalho comigo.] Fontes anotadas [e uma ocasional ajuda bibliográfica mantida no original em inglês] [1] Alfvén, Hannes, writing as Olof Johanneson, *The End of Man?*, Award Books, 1969 earlier published as "The Tale of the Big Computer", Coward-McCann, translated from a book copyright 1966 Albert Bonniers Forlag AB with English translation copyright 1966 by Victor Gollanz, Ltd. [2] Anderson, Poul, "Kings Who Die", *If*, March 1962, p8-36. Reprinted in *Seven Conquests*, Poul Anderson, MacMillan Co., 1969. [3] Asimov, Isaac, "Runaround", *Astounding Science Fiction*, March 1942, p94. Reprinted in *Robot Visions*, Isaac Asimov, ROC, 1990. Asimov describes the development of his robotics stories in this book. [4] Barrow, John D. and Frank J. Tipler, *The Anthropic Cosmological Principle*, Oxford University Press, 1986. [5] Bear, Greg, "Blood Music", *Analog Science Fiction-Science Fact*, June, 1983. Expanded into the novel *Blood Music*, Morrow, 1985. [6] Cairns-Smith, A. G., *Seven Clues to the Origin of Life*, Cambridge University Press, 1985. [7] Conrad, Michael et al., "Towards an Artificial Brain", *BioSystems*, vol 23, pp175-218, 1989. [8] Drexler, K. Eric, *Engines of Creation*, Anchor Press/Doubleday, 1986. [9] Dyson, Freeman, *Infinite in All Directions*, Harper & Row, 1988. [10] Dyson, Freeman, "Physics and Biology in an Open Universe",

\_Review of Modern Physics\_, vol 51, pp447-460, 1979. [11] Good, I. J., "Speculations Concerning the First Ultraintelligent Machine", in \_Advances in Computers\_, vol 6, Franz L. Alt and Morris Rubinoff, eds, pp31-88, 1965, Academic Press. [12] Good, I. J., [Help! I can't find the source of Good's Meta-Golden Rule, though I have the clear recollection of hearing about it sometime in the 1960s. Through the help of the net, I have found pointers to a number of related items. G. Harry Stine and Andrew Haley have written about metalaw as it might relate to extraterrestrials: G. Harry Stine, "How to Get along with Extraterrestrials ... or Your Neighbor", \_Analog Science Fact- Science Fiction\_, February, 1980, p39-47.] [13] Herbert, Frank, \_Dune\_, Berkley Books, 1985. However, this novel was serialized in \_Analog Science Fiction-Science Fact\_ in the 1960s. [14] Kovacs, G. T. A. \_et al.\_, "Regeneration Microelectrode Array for Peripheral Nerve Recording and Stimulation", \_IEEE Transactions on Biomedical Engineering\_, v 39, n 9, pp 893-902. [15] Margulis, Lynn and Dorion Sagan, \_Microcosmos, Four Billion Years of Evolution from Our Microbial Ancestors\_, Summit Books, 1986. [16] Minsky, Marvin, \_Society of Mind\_, Simon and Schuster, 1985. [17] Moravec, Hans, \_Mind Children\_, Harvard University Press, 1988. [18] Niven, Larry, "The Ethics of Madness", \_If\_, April 1967, pp82-108. Reprinted in \_Neutron Star\_, Larry Niven, Ballantine Books, 1968. [19] Penrose, Roger, \_The Emperor's New Mind\_, Oxford University Press, 1989. [20] Platt, Charles, Private Communication. [21] Rasmussen, S. \_et al.\_, "Computational Connectionism within Neurons: a Model of Cytoskeletal Automata Subservient Neural Networks", in \_Emergent Computation\_, Stephanie Forrest, ed., pp428-449, MIT Press, 1991. [22] Searle, John R., "Minds, Brains, and Programs", in \_The Behavioral and Brain Sciences\_, vol 3, Cambridge University Press, 1980. The essay is reprinted in \_The Mind's I\_, edited by Douglas R. Hofstadter and Daniel C. Dennett, Basic Books, 1981 (my source for this reference). This reprinting contains an excellent critique of the Searle essay. [23] Sims, Karl, "Interactive Evolution of Dynamical Systems", Thinking Machines Corporation, Technical Report Series (published in \_Toward a Practice of Autonomous Systems: Proceedings of the First European Conference on Artificial Life\_, Paris, MIT Press, December 1991. [24] Stapledon, Olaf, \_The Starmaker\_, Berkley Books, 1961 (but from the date on forward, probably written before 1937). [25] Stent, Gunther S., \_The Coming of the Golden Age: A View of the End of Progress\_, The Natural History Press, 1969. [26] Swanwick Michael, \_Vacuum Flowers\_, serialized in \_Isaac Asimov's Science Fiction Magazine\_, December(?) 1986 - February 1987. Republished by Ace Books, 1988. [27] Thearling, Kurt, "How We Will Build a Machine that Thinks", a workshop at Thinking Machines Corporation, August 24-26, 1992. Personal Communication. [28] Ulam, S., Tribute to John von Neumann, \_Bulletin of the American Mathematical Society\_, vol 64, nr 3, part 2, May 1958, pp1-49. [29] Vinge, Vernor, "Bookworm, Run!", \_Analog\_, March 1966, pp8-40. Reprinted in \_True Names and Other Dangers\_, Vernor Vinge, Baen Books, 1987. [30] Vinge, Vernor, "True Names", \_Binary Star Number 5\_, Dell, 1981. Reprinted in \_True Names and Other Dangers\_, Vernor Vinge, Baen Books, 1987. [31] Vinge, Vernor, First Word, \_Omni\_, January 1983, p10. [32] Vinge, Vernor, To Appear [ :-) ]. [ Nota 1] Alguns termos foram por opção deixada no original por falta de um termo adequado em português. Outros foram de certa forma criados como "super-humanamente" objetivando retratar mais fielmente a idéia original. Um trabalho de vanguarda como este encerra muitas dificuldades para o tradutor na medida em que no próprio idioma de origem o autor cunha termos novos.