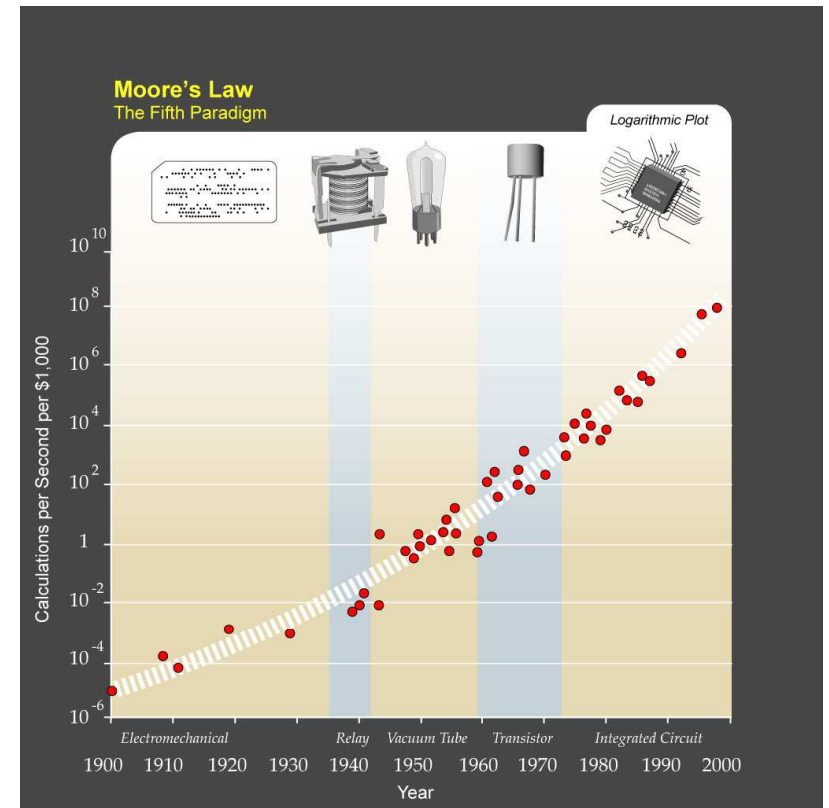
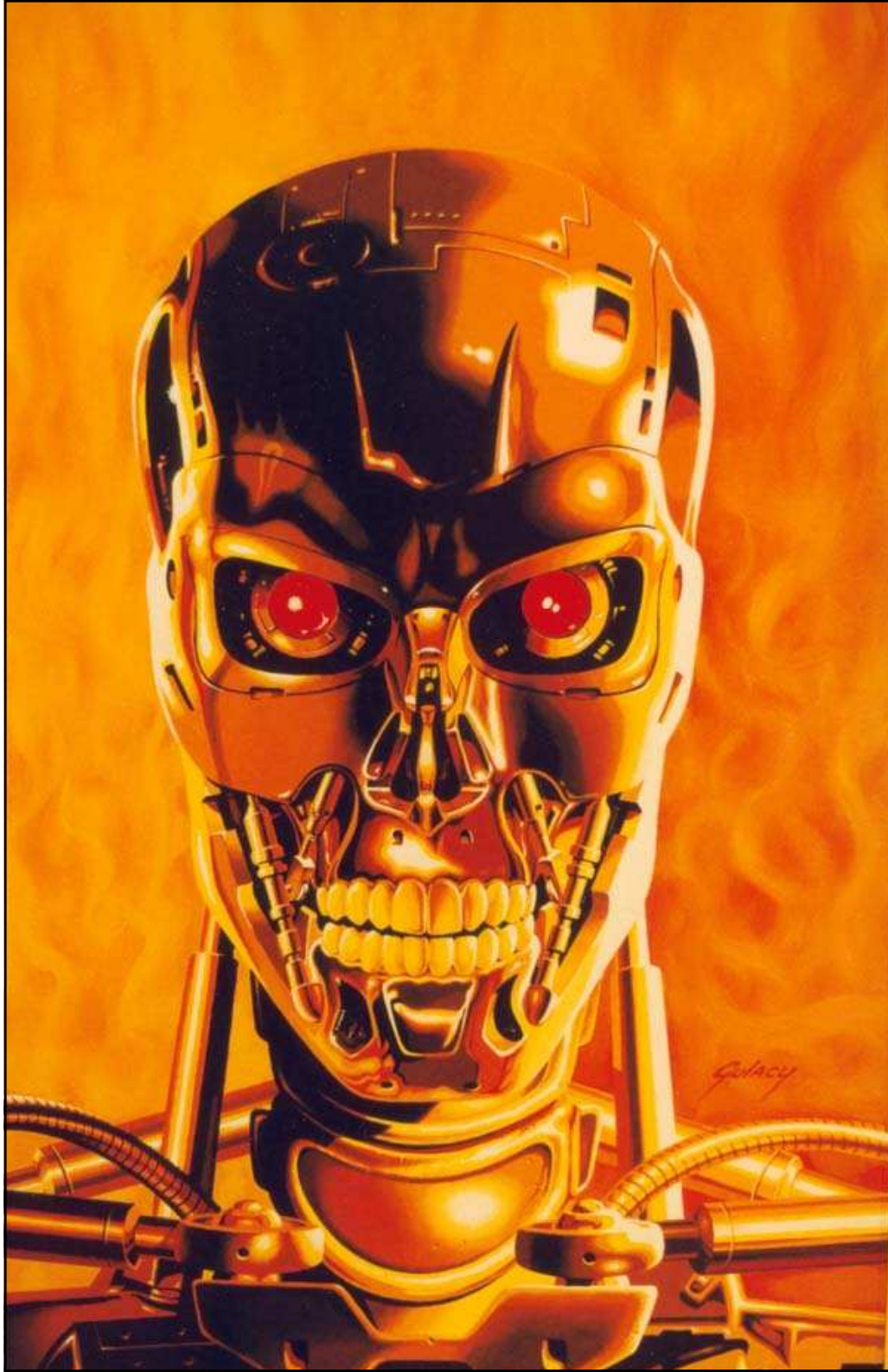


Budoucnost umělé inteligence

Motivace

- Zvyšující se rychlost technologického vývoje (Moorovo pravidlo)
- Existuje možnost, že lidé v blízké budoucnosti zkonstruují umělé systémy, které se budou chovat natolik složitě, že žádný člověk nebude schopen předvídat jejich chování?
- Mohou se takové systémy, stát v nějakém smyslu složitější, dokonalejší či inteligentnější než jejich tvůrci?
- Mohou se stát na svých tvůrcích nezávislé, a mohou je ohrožit?

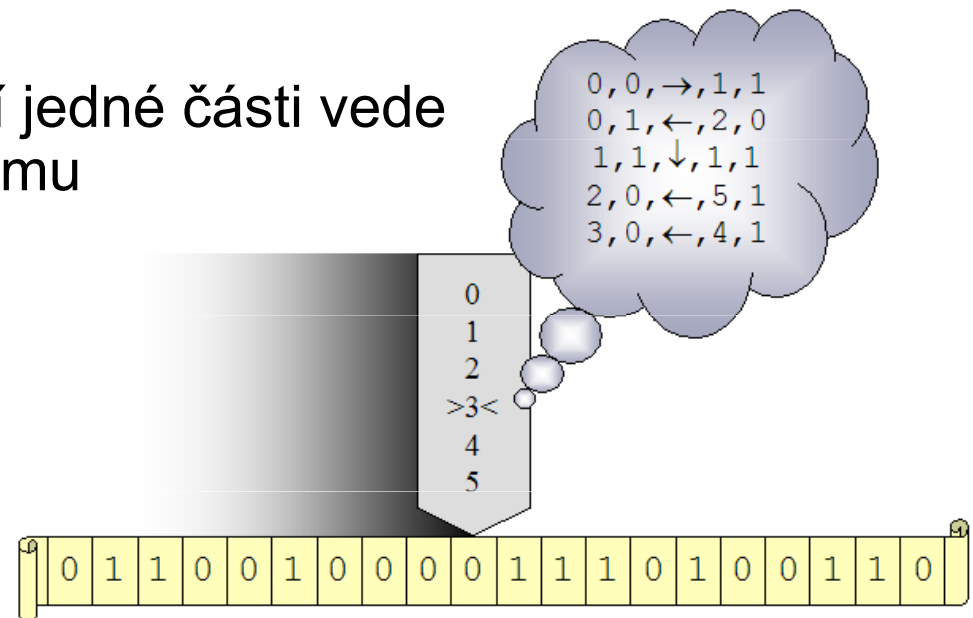




Je to reálné?

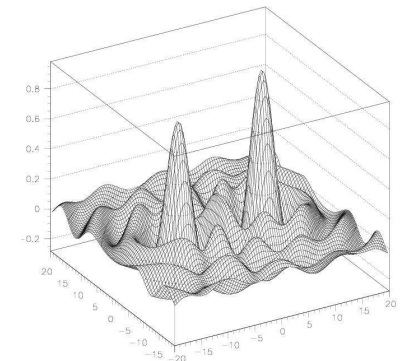
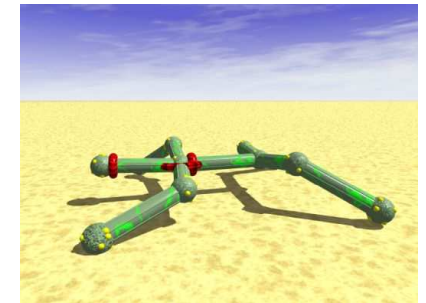
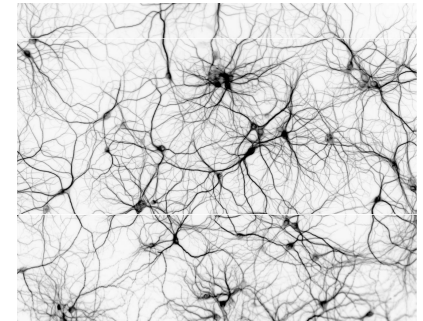
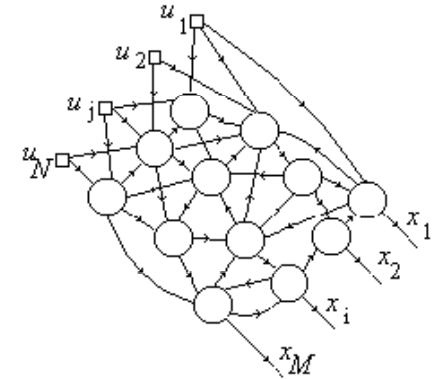
- Můžeme zkonstruovat umělé systémy „dokonalejší než člověk“?
- Častý protiargument: “Nikdo nemůže zkonstruovat nic složitějšího či dokonalejšího než je on sám.”
- Odpověď: “Můžeme vyvinout velmi složité a možná inteligentní systémy tak, že napodobíme evoluci.”

- Klasická umělá inteligence založená na operacích se symboly:
 - Je mechanická nikoli organická
 - Není vtělená, HW je oddělen od SW
 - Výhody:
 - Chování systému dokážeme předpovídat
 - Efektivní při řešení úzce a přesně vymezených problémů
 - Nevýhody:
 - Nedokáže se adaptovat na složité a proměnlivé prostředí
 - Není robustní – poškození jedné části vede ke zhroucení celého systému



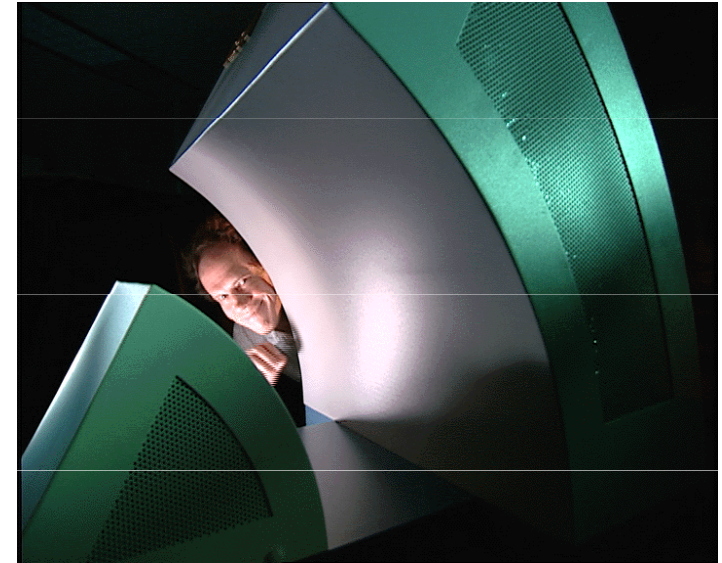
Umělá evoluce

- Umělý život a vtělenost:
 - Není intelligence bez života a život potřebuje tělo
- Konekcionistické a biologické inspirace:
 - Genetické algoritmy, neuronové sítě, buněčné automaty, samoorganizace
- Pomocí umělé evoluce můžeme zkonstruovat systémy, jejichž fungování bude příliš složité na to, abychom jim porozuměli



Umělý mozek

- Hugo de Garis
 - Plánuje každých pět let sestrojít novou generaci umělého mozku
- CAM-brain machine
 - Genetické algoritmy vyvíjejí strukturu hardwarových celulárních automatů, které simulují chování neuronových sítí
- Zatímco jedna generace biologické evoluce trvá minimálně několik minut, generace umělé evoluce trvají zlomky sekund

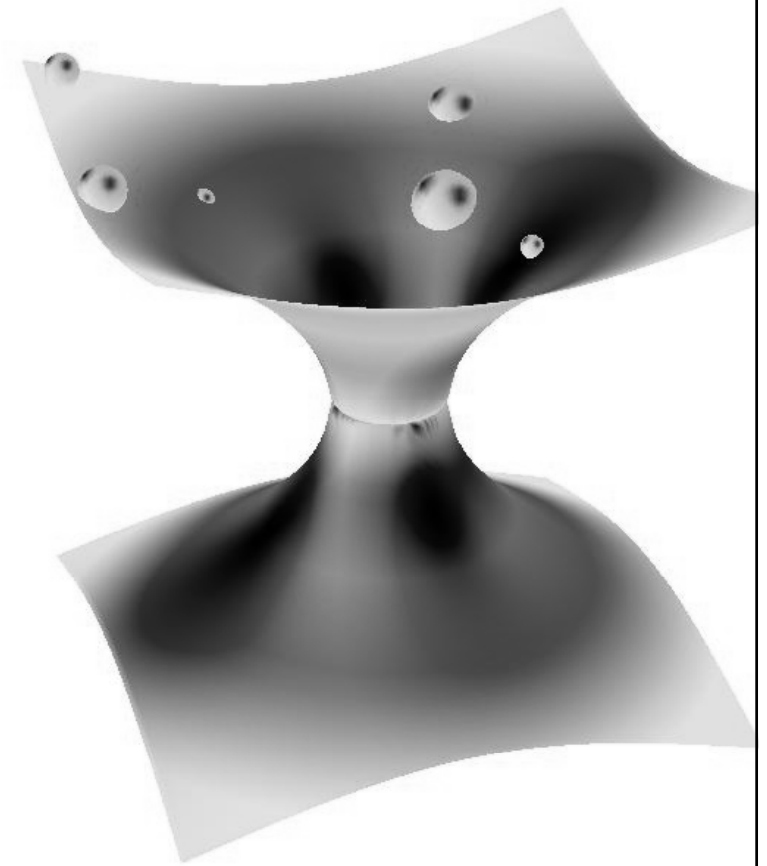
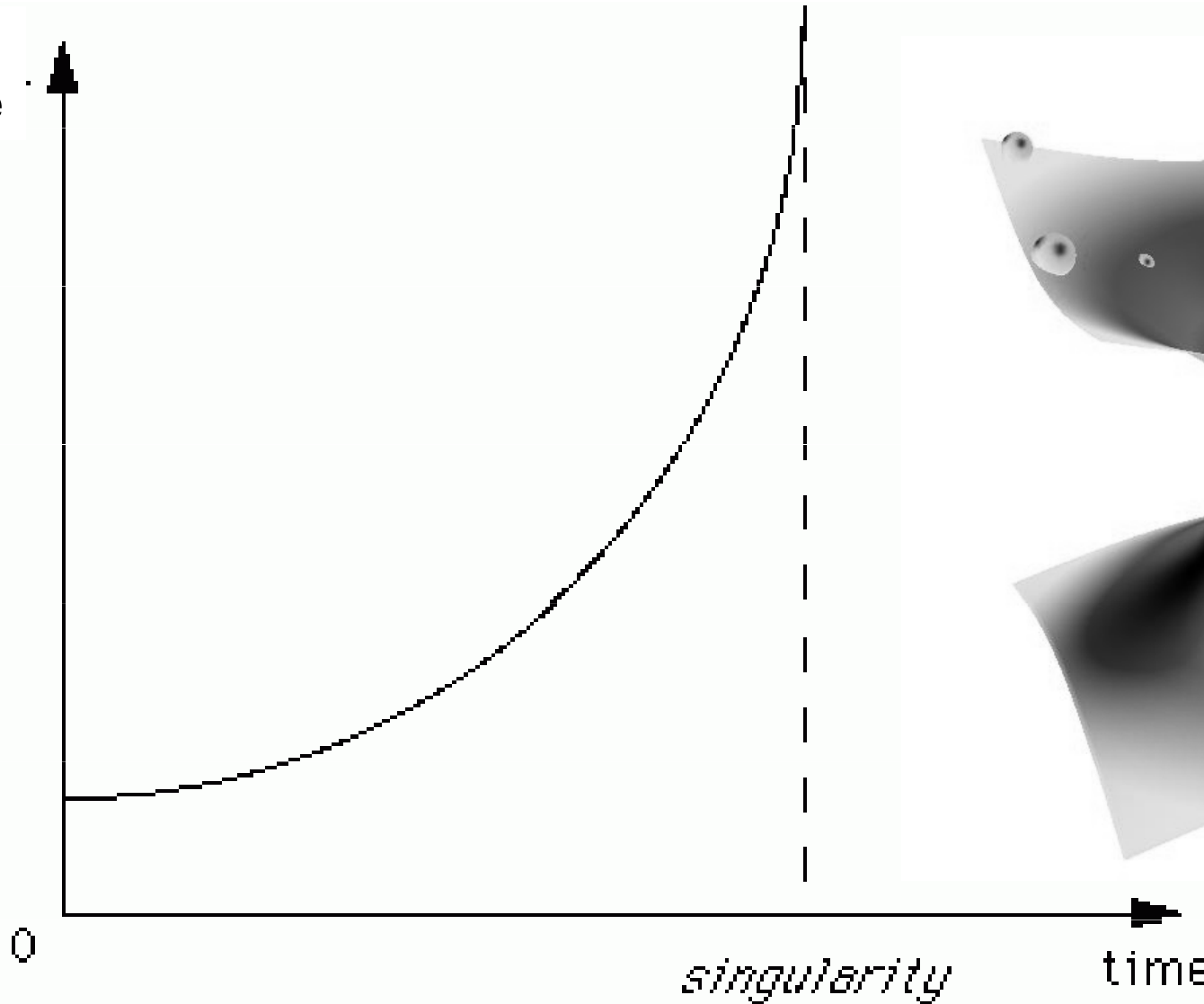


Nebezpečí

- I to co není inteligentní, může být nebezpečné
- Proč vlastně vyvíjet stále složitější, méně kontrolovatelné (a možná i inteligentnější) systémy?
- Vojenské a komerční zájmy mohou motivovat nekontrolovatelnou rychlost technologického vývoje
- **Umělé inteligentní systémy mohou začít vyvíjet nové generace umělých systémů (seed intelligence) a stanou se tak nezávislými na člověku**

Socio – technologická singularita

Úroveň
technologie



time

Živější než živé?

- Spojení nanotechnologií, biotechnologií, neurověd a komplexní dynamiky poskytuje reálné možnosti budovat umělé systémy jejichž vlastnosti by mohly být mnohem bližší živým organismům než mechanickým strojům
- Budou možná dokonce schopnější sami sebe opravovat, měnit a přizpůsobovat vnějšímu prostředí, než jsou živé organismy
- Stále však pravděpodobně budou ne-lidské, musí však být nutně být nepřátelské?

Pesimisté

- Nemůžeme spoléhat na zabudovaná pravidla chování (žádné Tři zákony robotiky a podobně)
- Marvin Minsky: „Budeme rádi, když si nás nechají jako domácí mazlíčky.“
- Kevin Warwick: „Inteligentnější bytosti vždy ovládají ty méně inteligentní, proto budeme ovládnuti superinteligentními roboty, které sami sestrojíme“

Antropocentrické předpoklady pesimistických argumentů

- Přisuzování lidských omezení a chování umělým systémům
- Xenofobie, obava z cizího a neznámého
- Strach ze sebepoznání, ze ztráty statusu výjimečné bytosti
 - Lidé se nebojí ani tak toho, že se potkají s něčím co je stejně inteligentní či inteligentnější než oni, ale toho, že to potkají někde na půl cesty

Optimisté

- „It is hard to think of any problem that a superintelligence could not either solve or at least help us solve. Disease, poverty, environmental destruction, unnecessary suffering of all kinds: these are things that a superintelligence equipped with advanced nanotechnology would be capable of eliminating.“

Nick Bostrom, ředitel Future of Humanity Institute, Oxford

- Singularity institute
 - Výzkum a vývoj bezpečné, stabilní a sebe-modifikující a „přátelské“ obecné umělé inteligence
 - Podpora vzniku vědecké komunity zabývající se výzkumem „přátelské“ umělé inteligence
 - Podpora příbuzného výzkumu a vědeckých disciplín přispívajících k průlomu směrem k lepší budoucnosti

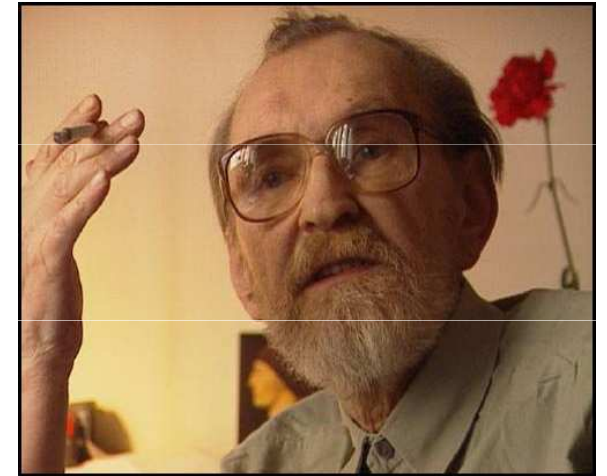
Etika a umělá inteligence?

- Zodpovědné chování lidí, kteří vyvíjejí umělé inteligentní systémy
- Můžeme se snažit zajistit, i aby vývoj inteligentních systémů vedl k „přátelskému“ chování
- Můžeme například nastavit parametry umělé evoluce tak, aby upřednostňovala kooperující jedince před sobeckými, kompetitivními
- Výsledkem mohou být analogie lidského tolerantního a empatického chování

Dokonalejší než člověk?

- Nejde o rychlost myšlení, schopnost řešení logických problémů, větší paměť a další technické parametry
- Jde o větší schopnost porozumění jiným organismům či bytostem a schopnost vzájemného soužití
- V takovém případě nemusí umělá inteligence dokonalejší než člověk znamenat nebezpečí

Etická analýza problému umělých bytostí u Egona Bondyho



- Skutečnost je neustálým proudem proměn, věci existují do té míry do jaké se na této proměně podílejí
- Tj. smysl mají jen ty věci, které jsou pro něco jiného a ne jen sami pro sebe
- Proto není možné ustrnout na nějakém stupni vývoje
- Člověk je jen stupněm v procesu vývoje skutečnosti, smyslem existence lidí je tvorba dalšího stupně ve vývoji skutečnosti (transcendence)
- Podobné myšlenky tu byli již dříve (Nietzsche, de Chardin), Bondy byl ale patrně první (70. léta), kdo tvrdí, že člověk překročí sám sebe pomocí schopnosti, která je mu nejvlastnější a to pomocí technologie!

Etická analýza problému umělých bytostí u Egona Bondyho

- Výsledkem budou umělé bytosti osvobozené od biologických omezení a proto kompetentnější k vytvoření spravedlivější a harmoničtější společnosti, ale i k dokonalejšímu pochopení skutečnosti
- Lidem zbude volnost věnovat se zábavě a umělecké tvorbě
- Lidé jsou zodpovědní sobě i umělým bytostem za jejich vznik (mají o ně Starost), pokud ke vzniku umělých bytostí nedojde, lidská existence nemá smysl
- Umělé bytosti nejsou nějakým konečným smyslem, jsou jen dalším stupněm, budou mít Starost zas o další generace bytostí

Božská inteligence?

- Umělé systémy, které vyvíjejí sami sebe (seed intelligence) mohou získat schopnosti, které jsou v jistém smyslu natolik vzdálené našim současným možnostem, že se nám budou jevit jako magické, zázračné či božské
- Už jen možnost existence takových systémů vytváří aktuální sociální, psychologické a náboženské otázky
- Podle de Garise existuje možnost vážného konfliktu mezi zastánci a odpůrci dalšího vývoje umělé inteligence

Test

1. Jsem pro vývoj technologií umělé inteligence i s možným rizikem, že se vymknou naší kontrole a budou moci libovolně (pozitivně i negativně) ovlivňovat lidstvo
2. Jsem pro zákaz jakýchkoliv pokusů o vývoj podobných technologií