

Pokroky matematiky, fyziky a astronomie

Zdeněk Renc

Myšlení a umělá inteligence

Pokroky matematiky, fyziky a astronomie, Vol. 21 (1976), No. 4, 212--219

Persistent URL: <http://dml.cz/dmlcz/139508>

Terms of use:

© Jednota českých matematiků a fyziků, 1976

Institute of Mathematics of the Academy of Sciences of the Czech Republic provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.



This paper has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library* <http://project.dml.cz>

Myšlení a umělá inteligence

Zdeněk Renc, Praha

Věřím, že jsme na prahu epochy založené na aktivitě inteligentních strojů, které v ní budou pravděpodobně dominovat.

M. Minsky

1. Úvod

V tomto článku se budeme zabývat některými otázkami, vyvstávajícími v souvislosti s problematikou tzv. umělé inteligence. V úvodním paragrafu stručně vymezíme okruh zájmů tohoto oboru, jeho současný stav i možné přístupy k této problematice.

Otázky a úvahy o podstatě a zákonitostech lidského myšlení, o tom, zda je možno ho nějak formálně popsat a charakterizovat, jsou velice starého data. Nikterak nové nejsou ani úvahy o tom, může-li stroj – a když, tak v jaké míře – napodobit či dokonce nahradit člověka. Ocitujeme jako příklad dva důvody DESCARTOVY, kterými argumentuje pro rozdílnost člověka a stroje [2]. První důvod: „I kdyby stroj uměl pronášet slova, neuměl by je tak sestavovat, aby odpovídala všemu, co se řekne v jeho přítomnosti, jak to i nejtupější lidé mohou činit“. Druhý důvod: „I kdyby vykonávaly určité věci stejně dobře nebo snad i lépe než kdokoliv z nás, selhaly by nevyhnutelně v jiných, jimiž by vyšlo najevo, že nejednaly s vědomím, nýbrž toliko sestavením svých orgánů; neboť kdežto rozum je všeobecný nástroj, kterého lze využívat ve všech možných případech, musí tyto orgány mít nějaké zvláštní uzpůsobení pro každý úkon jednotlivý“. Takové úvahy se staly naléhavějšími poté, co CHARLES BABBAGE navrhl v roce 1835 první číslicový počítač. I když tento počítač nikdy nebyl úplně realizován, zdálo se, že to je již více méně otázka technického pokroku a že úvahy na téma vztahu člověka a stroje jsou zcela aktuální. Nejrůznější názory a spory kolem těchto věcí plynou často z rozdílného a někdy ne zcela ujasněného chápání základních pojmů jako myšlení, inteligence, stroj a podobně. Mezníkem ve vývoji studia umělé inteligence se stal TURINGŮV článek *Computing Machinery and Intelligence*, který byl napsán v roce 1950 (a od té doby mnohokrát přetištěn v různých sbornících – viz např. [7]) a který je považován za oficiální základní kámen Artificial Intelligence (AI) jakožto vědního oboru. V tomto článku Turing shrnuje v kostce celou problematiku AI do otázky „Mohou stroje myslet?“, kterou, chceme-li můžeme alternativně formulovat ve tvaru „Lze sestrojít inteligentní stroj?“.

Při úvahách o těchto otázkách budeme sledovat a pečlivě rozlišovat tyto čtyři pojmy.

1. *Imitaci* budeme rozumět napodobení vnějšího chování. Důraz je kladen na vnější, jevovou stránku, zajímavé je to, co stroj dělá a ne jak a proč to dělá. Na toto behavioristické stanovisko se postavil v citovaném článku Turing. K tomu, aby mohl říci, co rozumí pojmem „inteligentní stroj“ navrhl tzv. Turingův test. Je to imitační hra, kterou

Lze velmi stručně popsat takto: V jedné ze dvou uzavřených místností je muž, ve druhé žena. Mimo tyto místnosti je examinátor, který oběma klade otázky a jejich pomocí chce za předem stanovenou dobu zjistit, ve které místnosti je muž a ve které žena. Muž při tom odpovídá tak, aby experimentátora zmátl a znemožnil mu správnou identifikaci. Změňme nyní hru tak, že místo muže zaujme stroj (pokus je nutno ovšem předem upravit, aby ze způsobu formulací a předávání otázek a odpovědí nebylo možno rozhodnout, kdo je kdo – není tedy možno užít televize nebo mikrofonu, lze ale vydávat odpovědi např. pomocí psacího stroje nebo tiskárny). Podle Turingovy definice je stroj inteligentní, jestliže je schopen přesvědčit průměrného examinátora kladoucího své otázky po dobu pěti minut o tom, že je žena, lépe než by toho byl schopen průměrný muž. Námitkou proti takovému přístupu je to, že stroj je při něm *à priori* v nevýhodě. Lze do něho sice vložit nejrůznější vědomosti a vyzbrojit ho tak na pozdější otázky, ale nelze ve stroji uložit např. životní zkušenost člověka (tu stroj principiálně nemůže mít stejnou jako člověk), systém hodnot a podobně. Jinou závažnou námitkou proti extrémnímu behavioristickému stanovisku je to, že při řešení úlohy (jak lze chápat požadavek odpovědět na danou otázku) rozlišujeme jednak výsledek řešení (solution), jednak proces řešení (solving). Zdá se, že inteligence se nevztahuje ani tak k „solution“ jako k „solving“. Její mírou je „elegance“, zejména pak úspornost v procesu řešení. Z tohoto hlediska se zdá, že Turingův test není zcela adekvátní pro stanovení inteligence. Je ovšem třeba říci, že tento přístup k otázkám AI nebyl nikdy důsledně sledován.

2. *Simulaci* rozumíme kromě imitace navíc i napodobení metod rozhodování. Zde se dostáváme do těsné souvislosti s oblastí heuristiky a heuristických metod. Příkladem programů motivovaných touto oblastí je známý program nebo spíše obecná metoda zvaná GPS (ERNST, NEWELL a další, viz [3]), který např. dokazuje matematické věty, řeší heuristické úlohy typu hanojská věž a podobně. Sem lze též zařadit speciální programy na hraní šachů i jiných her. Je zajímavé, že i když kvalita a schopnosti např. šachových programů během krátké doby značně vzrostly (první šachový program, který pracoval se šachovnicí zredukovanou na 6krát 6 polí a hrál dost špatně, byl napsán v roce 1957, v roce 1967 byl již k dispozici program, jehož úroveň byla klasifikována stupněm „middle-amateur“), zdá se, že napsání skutečně vysoce výkonného šachového programu by znamenalo značný kvalitativní pokrok. Tak např. I. J. GOOD, který v této oblasti pracuje, se domnívá, že kdyby byl sestaven program na hraní šachů schopný porazit mistra světa byli bychom od zrodu „ultra-inteligentního“ stroje vzdáleni ne více než pět let.

3. *Modelování*. K imitaci a simulaci zde dále přistupuje i napodobení struktury a její funkce. Rozpracováním tohoto přístupu bychom se snažili alespoň do jisté míry odpovídat na otázku, jaká je vlastně podstata nebo chceme-li tajemství lidské inteligence. Z pokusů, které byly v tomto směru podniknuty, si povšimněme alespoň dvou. ROSENBLATT navrhl v roce 1962 tzv. perceptronový model mozku, což byla v podstatě síť umělých neuronů, založených na modelech neuronů MCCULLOCHA a PITTSE (1943). Studium perceptronů ovlivnilo práce v oblasti tzv. rozpoznávání obrazů, přineslo některé matematické výsledky, ale bohužel se ukázalo, že komplexní pojem inteligence je z různých důvodů mimo rámec možností perceptronového modelu. Jinou biologicky

motivovanou ideou byl pokus modelovat na počítači sám evoluční proces. FOGEL a další popsali v roce 1966 pokus umožňující produkovat mnoho generací tzv. konečných automatů použitím strategií mutace a selekce. I když tato technika byla schopna zvládnout několik prvních milionů let evolučního procesu během několika dnů strojového času, ukázalo se, že zajímavá střední a pozdější vývojová stadia vyžadovala již do té míry komplikovanou (i když ne ještě „inteligentní“) strukturu těchto konečných automatů, že se jejich další vývoj již nedal použitím počítače zrychlit.

4. *Konstrukce* znamená vytvoření totožného objektu, tedy (hovoříme-li o umělé inteligenci) homunkula. Tuto oblast ponecháme zcela stranou (autor nechápal předloženou stať jako pokus o science-fiction).

Něco jiného však máme na mysli, užíváme-li slova robot – kterým rozumíme integrovaný elektromechanický systém (tedy bez jakýchkoliv biologických prvků) řízený počítačem, sestrojovaný za účelem znásobení nejen fyzických, ale hlavně duševních schopností člověka (což ovšem v žádném případě neznamená, že by měl být člověku roven ve smyslu totožnosti – je to spíše náhražka člověka v určité vymezené oblasti). Robotika je speciálním vyústěním AI, které užívá poznatků z nejrůznějších disciplín do AI zahrnovaných a spojuje je s technickými prvky a postupy. Abychom měli konkrétnější představu o tom, co nazýváme robotem (a současně tedy také alespoň neúplný přehled disciplín zahrnovaných dnes do širokého rámce AI), popíšeme zcela stručně tři základní složky robota, totiž složku percepční, analytickou a rozhodovací a akční. Percepční složka zprostředkovává příjem vstupních dat a informací z relevantního (vzhledem k robotu) světa. Konkrétními čidly mohou být televize, mikrofony a jiná elektroakustická zařízení. V teoretické úrovni se této sféry týká problematika rozpoznávání obrazů, analýza jasového pole, analýza a porozumění řeči, obecná problematika signálních soustav a podobně. Analytická a rozhodovací složka tvoří vlastní jádro celého systému. Zahrnuje v sobě reprezentaci vnějšího světa jak co se týče souboru vědomostí o něm, tak i jeho zákonitostí; jinými slovy jde o vytvoření axiomatického modelu vnějšího světa. Oborem, který se dále významně prosazuje, je tzv. problem solving (řešení úloh) – hledání postupů, kterými lze mechanicky dojít k řešení netriviálních (tedy „inteligenci“ vyžadujících) problémů. Slovem problém se zde nemíní úloha typu „znásob dvě velká čísla“ (která se řeší na základě apriorní znalosti algoritmu pro násobení), nýbrž úloha typu hanojská věž (tedy úloha, v níž se algoritmus řešení teprve vytváří v jeho průběhu). V problem solvingu můžeme rozlišit dva základní přístupy; jednak kombinatoricko-heuristický, kde v podstatě jde o hledání cest ve stromu všech řešení, jednak přístup využívající výsledky matematické logiky (automatické dokazování vět a podobně). Dále je třeba studovat otázky hodnotících kritérií, různých strategií, použitelných při řešení úloh, jejich porovnávání a optimalizaci atd. Konečně třetí, tj. akční složka robota, zahrnuje pohybová a manipulační ústrojí, realizující učiněné závěry v podobě konkrétních akcí ve vnějším světě. Je zřejmé, že celý tento systém není strnulý, neboť činnost akční složky může měnit strukturu vnějšího světa, a to pak má za následek modifikaci jeho původního modelu, odtud případně modifikaci dílčích problémů, které musí stroj řešit.

Za současného stavu se dají rozlišit tři hlavní přístupy k problematice AI. Za první přístup řekněme psychologický, pro který je typická snaha porozumět lidské myšlenkové

činnosti tím, že se ji pokoušíme modelovat. Za druhé inženýrský, který bychom v rámci naší terminologie mohli též nazvat průmyslovou robotikou. Konečně pak přístup v určitém smyslu nejobecnější, který bychom snad mohli označit jako filozoficko-matematický. Nejde při něm o modelování lidského myšlení, ale o hledání a studium jiných metod, kterými lze dosáhnout stejných nebo analogických výsledků. Tohoto třetího přístupu se v naší práci přidržíme. V souladu s ním můžeme vymezit AI jako obor zabývající se imitováním a simulováním lidských myšlenkových procesů pomocí stroje (počítače).

2. Může stroj myslet?

Abychom se mohli pokusit odpovědět na tuto otázku, musíme – jak jsme již naznačili v úvodu – nejprve precizovat základní pojmy myšlení, inteligence a stroj. Nejjednodušší to bude u posledního z nich. Strojem budeme rozumět elektromechanický systém řízený počítačem, jinými slovy počítač spolu s přídavnými zařízeními, které mu umožňují oboustrannou komunikaci s vnějším světem (včetně examinátora). Počítačem při tom míníme číslicový počítač v podstatě současné úrovně, vybavený možná větší operační rychlostí a kapacitou paměti, ale neodlišující se od současných počítačů žádnými kvalitativně novými schopnostmi. Chceme-li, můžeme si pro názornost jako stroj představovat robot (tak jak jsme ho popsali v předchozím paragrafu), přičemž ale nemusíme klást zvláštní důraz na jeho akční složku. Podrobnějšího rozboru si vyžaduje pojem „myšlení“, respektive „inteligence“. GEORGE v [4] říká: „Myšlení je proces symbolizování událostí a manipulace s těmito symboly různými způsoby logického i ilogického vyvozování (inference), kde do zpracování (svého druhu zpracování dat) přicházejí představy. Představy se mohou chápat jako oživení vstupu v oslabené formě nějakou dobu poté, co byla přijata původní vstupní data. Představivost (imaginace) je cosi více než toto, neboť může vzájemně kombinovat různá předchozí vstupní data do konfigurace, která se nikdy před tím aktuálně nevyskytla“. Podobně MCCARTHY a HAYES [1] definují pojem inteligence takto: „Řekneme, že jsoucno je inteligentní, má-li adekvátní model světa (včetně intelektuálního světa matematiky, porozumění svým vlastním cílům a jiným duševním procesům), je-li dostatečně důmyslné, aby odpovídalo na širokou škálu otázek na základě tohoto modelu, je-li schopno získat v případě potřeby dodatečnou informaci z vnějšího světa a je-li schopno uskutečňovat takové úlohy ve vnějším světě, které si jeho cíle žádají a které jeho fyzické schopnosti připouštějí“.

Slovo „myslet“ budeme tedy v této práci používat ve smyslu „racionálně uvažovat“, tedy ve smyslu běžně užívaném v češtině ve spojeních přemýšlet o něčem, promýšlet něco, vymýšlet něco a odpovídajícím např. anglickému „to reason“. Nebudeme naopak do našeho „myslet“ zahrnovat význam uvědomovat si, tj. „cogito“ (myslím, tedy jsem) ani význam domnívat se, tj. „to think“ (myslím, že se ochladí). V souladu s tímto vymezením nazveme inteligencí schopnost racionálně uvažovat a řídit svoje jednání na základě tohoto uvažování (na rozdíl od instinktu, kde jednání je řízeno předem pevně vloženým programem).

Myšlení v našem slova smyslu je vždy symbolický proces, který jako své dílčí složky

obsahuje složku imaginační, dedukční a indukční. V procesu myšlení můžeme z pracovních důvodů vydělit dvě relativně oddělená stadia, která můžeme s jistou dávkou nepřesnosti charakterizovat jako statickou a dynamickou fázi tohoto procesu. První fáze představuje vytvoření symbolické reprezentace světa; nejde ovšem pouze o otisk okolní reality, o znalost světa, ale nutně i o poznání a zachycení jeho zákonitostí. Tato první fáze je nezbytně nutná k imaginaci, k tomu, abychom mohli ve svém mozku vytvářet nové modely světa, které jsou možné (abychom toto mohli rozhodnout, potřebujeme právě znát obecné zákonitosti), ale které nejsou nutně odrazem aktuálně existující skutečnosti (možnost vytváření takových modelů je podmíněna symbolickým charakterem naší reprezentace světa). Takovéto modely světa nám pak dále umožňují plánovat další myšlenkový postup případně rozhodovat mezi možnými pokračováními. Dynamickou fází procesu myšlení je vlastní inference, vlastní řešení úloh ve vnějším světě, které jsou vyžadovány z hlediska celkového cíle. Je patrné, že takovéto rozdělení procesu myšlení je pouze schematické a nelze mu přikládat absolutní platnost. Jednak totiž první fáze, kterou jsme označili jako statickou, ve skutečnosti tak zcela statická není (toto slovo bylo voleno spíše k vyznačení rozdílu mezi ní a druhou, vskutku dynamickou částí). Při vytváření modelu světa, zejména při kódování jeho zákonitostí naopak člověk atakuje vnější svět, aby tyto zákonitosti lépe a úplněji poznal. V důsledku toho se ovšem může zpětně pozměňovat model světa, který si až doposud vytvořil. Dá se říci, že toto vzájemné působení směřující k doplňování, pozměňování nebo i jen ověřování modelu světa v průběhu lidského života a jeho myšlenkové činnosti vlastně nikdy nekončí (jinak by se totiž mohlo stát, že usuzování prováděné na základě v jistém okamžiku uzavřeného modelu světa, povede někdy v budoucnu ke sporným závěrům, jinými slovy, že „starý“ model světa se ukáže co do struktury neadekvátní „novému“ světu). Za druhé pak nejsou naše dvě fáze od sebe nikterak striktně odděleny, ale prolínají se. Při hledání co nejvěrnější reprezentace světa je jistě třeba (např. v souvislosti s poznáváním a reprezentováním jeho zákonitostí), aby se zapojila inferenční složka. Nicméně však je užitečné si vzpomenuť dvou částí procesu myšlení výslovně povšimnout již proto, abychom mohli nyní přehledněji sledovat, jak se dá proces myšlení simulovat na počítači (stroji).

Abychom se nedostali příliš hluboko do vlastní odborné problematiky AI, popíšeme způsob této simulace pouze v hrubých rysech a navíc k tomu pro větší názornost použijeme příkladu z robotiky, který byl skutečně realizován na Stanford Research Institute v roce 1969. (Příklad je citován a popsán v [5].) Ten se týká robota (označovaného přezdívkou Shakey), který se pohyboval pouze na kolečkách (nemohl tedy vystupovat např. na schody ani sestupovat z nich dolů) dopředu a dozadu, mohl se libovolně otáčet a strkat do věcí a tím je posunovat (neměl ale žádná chapadla apod.); robot měl televizní kameru, kterou si mohl prohlédnout celé okolí a samozřejmě byl spojen s počítačem. Robot byl umístěn v místnosti, ve které bylo dále pódium (robot sám byl na podlaze mimo pódium), na němž ležel malý kvádr a mimo něj na podlaze nakloněná rovina. Úkolem robota bylo shodit kvádr z pódia. Řešení úlohy je zřejmě takové, že robot nalezne nakloněnou rovinu, přisune ji k pódiumu, vyjede po ní nahoru, dojede po pódiumu ke kvádru, odsune ho na kraj pódia a shodí dolů (tak by úlohu řešil člověk, byl-li by vybaven pouze těmi fyzickými schopnostmi, kterými jsme vybavili robot). První fáze

řešení spočívá v tom, že robot získá reprezentaci vnějšího světa (což je v tomto případě místnost s popsányými předměty a jejich konfigurací). Statický obraz může získat sám pomocí televizní kamery, on ovšem potřebuje více – znát zákony tohoto světa (musí se např. dozvědět, že nakloněnou rovinou lze pohybovat zatímco pódiem ne a podobně). Reprezentace světa se může realizovat například v tak zvaném situačním kalkulu. Jeho jazyk má dva druhy proměnných pro objekty a stavy (speciálně i konstanty druhu objekt a jednu konstantu druhu stav), dále funkční a predikátové symboly (formálně řečeno, pohybuje se v predikátovém počtu s funkcemi a dvěma druhy proměnných; při tom termy a formule tohoto jazyka se definují obvyklým způsobem). Při interpretaci jazyka jsou konstanty typu objekt konkrétní objekty nacházející se v místnosti, stavy chápeme jako konfigurace objektů (pozorovaných v daném okamžiku) spolu s relacemi mezi nimi (být vedle sebe, na sobě apod.), které interpretují predikátové symboly, vytčená stavová konstanta je interpretována počátečním stavem, ve kterém robot začíná svou činnost a konečně funkční symboly interpretujeme jako zobrazení, která daným objektům a stavu přiřazují nový stav. Axiómy jsou jednak tzv. rámcové, které popisují, co vše se během činnosti robota nemění, jednak ty, které vymezují základní fakta o počátečním stavu, predikátech a funkcích. V takovémto formálním jazyku má tedy robot zapsány poznatky o okolním světě a jeho vlastnostech i o svých možnostech. Aktivní podíl robota na získání, případně doplnění této reprezentace se může realizovat například tak, že mezi jeho schopnosti zařadíme tzv. verbální akci – zjistí-li robot během řešení úlohy, že nemůže pokračovat pro nedostatek znalostí, požádá examinatora, aby mu např. doplnil další axióm. Danou úlohu je nyní možno formulovat tak, že stroj má za úkol dokazovat (v situačním kalkulu) formuli tvaru „existuje situace, v níž je kvádr na podlaze“. Pro takové dokazování mohou být do počítače zabudovány různé postupy a strategie z oblasti problem solvingu (často se např. používá tzv. rezoluční metoda). Při nich stroj zkouší využívat svých přechodových funkcí k vytváření nových stavů, které ovšem fakticky nemusí realizovat, ale hodnotí, zda mu umožňují přiblížit se k cílovému stavu a podle toho je buď opouští, nebo z nich pokračuje dále. Takovým způsobem simuluje vlastní racionální uvažování, vytváření teorií a jejich ověřování atd. Způsob jeho dokazování je konstruktivní, takže po skončení důkazu vydá stroj jako výsledek term, jehož struktura ukazuje postup, kterým se k žádanému cílovému stavu skutečně dospěje (a tento postup potom realizuje).

Uvedený příklad ukazuje, že simulace lidského myšlení strojem je do jisté míry omezena jednak problémem, který byl stroji dán k řešení, jednak jeho schopnostmi, které do něho byly člověkem vloženy. Záleží ovšem na tom, jak a kolik se toho stroji řekne (např. kolik a jak obecných dokazovacích principů se do jeho paměti předem vloží a v jakém formálním jazyce budou vyjádřeny). Připomeňme si znovu, že pojem „myšlení“ jsme si vymezili ve smyslu „racionální inference“; to znamená, že jsme ponechali stranou některé aspekty, které při úvahách o procesu lidského myšlení a poznání vcelku k abstraktnímu myšlení obvykle automaticky připojujeme. I při našem omezení pojmu myšlení na tom však není stroj zase tak špatně. Může si např. ukládat do paměti postupy řešení dříve zadaných úloh i s tím, zda byly úspěšné, případně do jaké míry byly optimální, na základě toho se učit a provádět tzv. úsudky z analogie.

Vcelku se domníváme, že jsme prokázali, že myšlení v našem slova smyslu nebo

jinými slovy některé důležité stránky lidského myšlení lze (alespoň v jednodušších případech) simulovat strojem. Další vývoj spočívá patrně jak ve studiu dalších stránek procesu poznání a pokusech o jejich formální popis (i když víme nebo spíše doufáme, že se nám takto nepodaří zmapovat člověka a jeho proces poznávání vcelku), tak i ve zdokonalování a kvalitativním prohlubování současných postupů. Tak například velmi důležitá a zatím ne zcela uspokojivě řešená je otázka vhodné reprezentace vnějšího světa. Na našem příkladu robota jsme naznačili, jak se s ní lze vyrovnat v jednoduchém konkrétním případě. Obecně ovšem může být situace velmi složitá. Volba reprezentace světa závisí do značné míry na cílech, které sledujeme, na třídě problémů, které chceme řešit. Máme-li nějaký formální jazyk, ve kterém můžeme stroji předat jeho „usuzovací schopnosti“, je třeba také v tomto jazyce zformulovat problém (jehož řešení pak bude probíhat v podstatě volbou a aplikací patřičných transformací příslušných formálních výrazů). Třída problémů, které je možno v tomto jazyce formulovat udává míru obecnosti robotových schopností, ale jejich skutečnou sílu vyjadřuje ta podtřída problémů, které se opravdu vyřešit dají. Může se stát, že nějaký problém bude sice zformalizovatelný, ale jeho formální vyjádření bude do té míry složité a neohebné, že stroj nebude schopen jej fakticky vyřešit. Snahou samozřejmě je nalézt takovou reprezentaci světa, v níž se síla schopností stroje co nejvíce přiblíží jejich obecnosti. S jejím vytvářením jsou spojeny různé netriviální otázky, např. jak má vypadat reprezentace světa, aby byla schopna vyjádřit také nové přírodní zákony, které teprve budou objeveny, nebo které další stránky světa (mentální stavy poznání, cíle, hodnoty apod.) mají být formalizovány a jak. Po optimální reprezentaci vnějšího světa žádáme, aby byla tomuto světu adekvátní jak v tom smyslu, že není ve sporu s fakty (aby tak svět vzhledem k našim znalostem opravdu mohl vypadat), tak epistemologicky (aby byla vhodná pro vyjádření faktů) a konečně i heuristicky (aby v ní byl vyjadřitelný proces uvažování aktuálně probíhající při řešení problému, tj., aby dávala možnost introspekce). Všechny tyto problémy a požadavky jsou značně složité, a jak jsme již řekli, nejsou dosud v obecnosti vždy uspokojivě vyřešeny. Tím se dostáváme k poslednímu slovu, které se v otázce tvořící název tohoto paragrafu objevuje, totiž ke slovu „může“. Chápeme-li ho jako principiální možnost (což by sice z technického hlediska nemělo být příliš vzdáleno od možnosti aktuální, protože jsme řekli, že strojem rozumíme v podstatě současný počítač – pozor však na právě uvedené poznámky o problematice adekvátní reprezentace), lze na položenou otázku tedy s uvedenými výhradami odpovědět kladně. Pokud bychom chtěli být dotěrní a ptát se, kdy tato možnost bude vskutku reálná, nezbude ovšem než se uchýlit k odhadům a například opět citovat Turinga, který ve svém článku vyslovuje přesvědčení, že inteligentní stroje – inteligentní ve smyslu jeho pojetí, tj. ve smyslu Turingova testu – budou sestrojeny do konce století.

3. Závěr

Zabývali jsme se některými analogiemi v myšlení člověka a stroje. Zdůrazněme znovu, že termín „myšlení“ jsme vymezili jako „racionální inferenci“; ponechali jsme tedy do značné míry stranou afektivní stránku myšlení, smyslový aparát v jeho plně šíři, otázky

životních zkušeností a postojů, hierarchii hodnot apod. (které stroj buď principiálně nemůže mít stejně jako člověk, nebo v nejlepším případě nevíme ještě dost dobře, jak je kódovat a napodobit). Dále jsme předmět našeho zájmu, oblast Artificial Intelligence, vymezili jako imitování a simulování lidských myšlenkových procesů, nikoli tedy jako jejich modelování či dokonce konstrukci totožného objektu. Těmito omezeními předcházíme námitce, která se proti AI často uvádí a která hovoří o nebezpečí redukce člověka na stroj. Kromě toho se domníváme, že tato námitka není příliš pevná ani při pohledu z jiné strany. Snaha po symbolismu u člověka ho naopak (velmi zhruba řečeno) posunuje v jistém smyslu blíže k robotu. Kdybychom si chtěli znázornit jakési vícedimenzionální kontinuum inteligence tak, že bychom na jednu osu nanášeli rychlost a přesnost myšlení a na druhou jeho důmyslnost nebo podobně zachytili současně různé typy paměti (jednak repetiční, mechanickou, jednak asociativní), bylo by umístění člověka a stroje asi následující:



Cílem AI je posunout na těchto obrázcích stroj doprava (aniž by ovšem klesl níže), naopak u člověka se projevuje částečná snaha po posunu směrem vzhůru. Význam AI spočívá tedy v tom, že jednak studiem podstaty lidského myšlení a odhalováním jeho zákonitostí a simulováním myšlenkových procesů pomocí stroje můžeme podstatně znásobit možnosti a schopnosti strojů, na druhé straně ale současně zkoumáním struktury „inteligentních“ strojů a především struktury jejich činnosti můžeme zpětně dosáhnout lepšího porozumění principům lidského myšlení.

Literatura

- [1] J. MCCARTHY, P. HAYES: *Some Philosophical Problems from Standpoint of Artificial Intelligence*, *Machine Intelligence 4* (1969).
- [2] R. DESCARTES: *Rozprava o metodě*, Leichter, Praha 1947.
- [3] G. W. ERNST, A. NEWELL: *GPS: a Case Study in Generality and Problem Solving*, Academic Press, New York, 1969.
- [4] F. H. GEORGE: *Behavioural Cybernetics*, Survey of Cybernetics (J. Rose ed.).
- [5] I. M. HAVEL: *The Concept of Indirectness in Artificial Intelligence*, *Kybernetika 8* (1972), 2.
- [6] M. MINSKI: *Psychological Today 2* (1969).
- [7] A. M. TURING: *Computing Machinery and Intelligence*, *Computers and Thought* (Feigenbaum, Feldman eds.), McGraw-Hill, 1963.