

Gymnázium Jana Keplera
Maturitní práce z Humanitních studií

2017/2018

ZVÍŘE – ČLOVĚK – STROJ

ANEB JAK NAJÍT MYSLÍCÍ STROJ

KATEŘINA KREJSOVÁ, R8.A

Vedoucí práce: Mgr. Jan Kolář

PROHLÁŠENÍ

Prohlašuji, že jsem tuto práci napsala samostatně pod vedením Mgr. Jana Koláře a s použitím uvedené literatury.

V Praze dne

.....

PODĚKOVÁNÍ

Chtěla bych poděkovat vedoucímu mé práce Janu Kolářovi za inspiraci a motivaci pro psaní práce, cenná doporučení literatury a čas strávený na konzultacích. Děkuji také Danilu Koževnikovi a Danu Maredovi za konzultaci a za vysvětlení problematiky z pohledu matematiků.

OBSAH

1	Úvod	2
2	Vymezení pojmů.....	4
2.1	Co je to mysl?	4
2.2	Co je to umělá inteligence?	5
2.3	Může stroj myslet?.....	6
3	Zvíře – člověk – stroj	8
3.1	Instinktivní přisuzování	8
3.2	Rozum a roboti	9
3.3	Zvíře – člověk – stroj.....	10
3.4	něco jako my.....	11
4	Evoluce	14
4.1	Aspekty mysli	14
4.2	Umělá mysl a život	15
4.3	Evoluce – zvíře	16
4.4	Evoluce – stroj.....	17
5	Jazyk.....	20
5.1	Úvod do jazyka – schopnost vyplnit povel.....	20
5.2	KTM	21
5.3	Determinismus.....	22
5.4	Čínský pokoj.....	23
5.5	Prosvětlený pokoj	24
6	Člověk – jazyk – stroj	26
6.1	Myslíme v jazyce?	26
6.2	Myslí stroj v jazyce?.....	28
6.3	Jak naučit stroj jazyku?.....	28
7	Závěr	30
	Použitá literatura.....	32

1 ÚVOD

Svou práci považuji za pout' za hledáním umělé mysli.

Prozkoumávám říši filosofie mysli s cílem objevit v ní své vlastní orientační body a také předat možný pohled na ni. Jak vlastně myslet o mysli? Hlavní motiv, s nímž jsem do této říše vstupovala, je hledání povahy mysli stroje. Jak by mohla vypadat? Je podobná té naší? Na jakém principu ji lze založit? Dokážeme si vůbec mysl stroje představit?

Snažím se být skromná; ptát se, „co je to mysl“ a „jestli stroj může myslet tak, jako nyní myslím já“, a zároveň nepředstírat, že poskytnu vyčerpávající odpovědi na všechny tyto otázky. Popravdě, za daleko důležitější pro filosofii považuji schopnost přistoupit k otázce, uchopit ji, obohatit ji a obohatit skrze ni sám sebe.

Pro splnění tohoto úkolu jsem se rozhodla zaměřit se na vztah mysli a zvířete. V současnosti existuje poměrně zažitá představa ohledně toho, zdali mají zvířata mysl, instinktivně na tuto otázku odpovídáme kladně. Nikde jsem však toto pojetí mysli nenalezla propojené s tím, jak uvažujeme o mysli stroje. Na následujících stránkách tedy aplikuji tento vztah zvíře – člověk na vztah člověk – stroj, a věřím, že právě tento vztah tvoří užitečný kontrast mezi tím, jaké druhy mysli si umíme představit, co za mysl považujeme a jak k otázce přítomnosti mysli přistoupit.

Na vztahu zvíře – člověk je mnoho intuitivního. Svou práci začínám prozkoumáním této perspektivy. Snažím se skrze ni nastínit přístup, který je pro člověka běžný či přirozený a zjistit, ve kterých situacích je nám užitečný a ve kterých selhává. Někdy se „instinktivní“, tedy automatický, bez složitější úvahy, přirozený reflexem, prolíná se slovem „intuitivní“, tedy na základě vnitřního pocitu, přirození emocí. Pro jasnost sdělení jsem zůstala u původního termínu „instinktivní přístup“, je však pravda, že mnohokrát znamená totéž jako intuitivní přístup.

Ze vztahu mezi myslí zvířecí, lidskou a tou u stroje (UI) vyplývají témata dalších dvou kapitol: nejprve se zaměřuji na evoluci, jakým způsobem ovlivňuje zvířata, jak se projevuje na jejich mysli a jak ji propojit se strojem. Poté se věnuji jazyku; stručnému představení počítačové teorie mysli (KTM), jakožto teorie zabývající se tématem člověk – stroj a argumentu Čínské pokoje (ČP), který se dotýká schopnosti stroje používat jazyk. V poslední kapitole pak spojuji člověka a stroj ve vztahu k jazyku, tak jako dřív ve vztahu k evoluci a zkoumám, jak skrze jazyk může stroj získat mysl.

Vstupme tedy do dalších kapitol s těmito otázkami: Jakým způsobem můžeme nějaké entitě přisoudit mysl? Co jsou potom ty nejdůležitější části mysli, kterou přisuzujeme? Jak by měla taková mysl vypadat?

2 VYMEZENÍ POJMŮ

2.1 CO JE TO MYSL?

Odpovím přímo: tak jako pro mnoho jiných pojmů ve filosofii na tuto otázku neexistuje jednoznačná odpověď. Problematika mysli je interdisciplinární, zabývají se jí mimo filosofů psychologové, neurovědci a v případě otázky umělé inteligence i informatici, a každý z nich nabízí perspektivy odpovídající svému oboru, nikdo z nich však nepřichází se společnou odpovědí. Mysl je pro nás speciální, je to jak něco, co nás všechny společně činí lidmi, tak něco, co je možné poznat jen v naprostém soukromí. Oba tyto aspekty ztěžují její zkoumání, jelikož se tohoto výsadního postavení podvědomě nechceme vzdát. Na druhou stranu, toto není problém, jímž by se tato práce měla zabývat. Přesto je nutno se k němu vyjádřit, jelikož s pojmem mysl pracuje.

Mysl si lze zjednodušit na věty jako „co není mysl“, „co je kupříkladu mysl“ či „co je podobné mysli“. Můžeme ji postavit proti pojmům podobným (či vedle nich), s nimiž se proplétá. Jak si mezi sebou stojí mysl (mind) a vědomí (consciousness)? Jaký vztah má k těmto slovům inteligence? Na druhou stranu se musíme zabývat otázkou, jak vlastně mysl samotná funguje, neboť se *naše* otázka týká něčeho úmyslně vytvořeného, tedy umělého.

„Vy například víte, že máte mysl, [...] Se svou myslí jste obeznámeni intimněji – tak intimně, že byste dokonce mohli říci, že jste svou myslí. Kniha nebo učitel vám může říci, co to mysl je, avšak to, že ji máte, vám nemusí říkat nikdo. Kdyby vás napadlo o tom, zda jste normální a máte mysl tak jako jiní lidé, pochybovat, okamžitě byste si uvědomili, jak na to upozornil Descartes, že samo vaše pochybování nade všechny pochybnosti demonstruje, že mysl skutečně máte.“¹

Tento přístup má ale velmi jednoduchý háček – rozumíme skrze něj tomu, že mysl máme, a tak nějak víme, co dělat se zvířaty, skrze intuici však upadáme do předsudků, které jsou u nových a nezvyklých témat nevídané. Přesto je podobné vymezení dobrá pojistka, neboť v hlubinách problematiky mysli se dá snadno ztratit; když se otázky zdají být neukotvitelné a kdy není jasné, o čem se tedy pojednává, můžeme si říct: „Jaký vztah má toto k mysli, kterou mám nyní já?“

¹ Dennet, D. C.: Druhy myslí, přel. J. Peregrin, Praha, Academia, 2004, s. 12-13

Mysl je to, co si myslím. Jsou to slova, skrze která tvořím myšlenky, moje vlastní řeč. Nebo vědomí? Vědomí je to, že jsem si něčeho vědoma, že si uvědomuji, že myslím, že si uvědomuji, že jsem. Abych mohla myslet, nemusím mít vědomí, ale to, že jsem, si mohu uvědomit jenom skrze to, že si to pomyslím, mysl tedy mít musím.

To je pouhé hraní se slovíčky. Stejně tak je ale součástí mé mysli to, že prožívám utrpení, že se něčeho bojím, což v ‚taxonomii‘ slova mysl snadno nenalézám. Avšak není už její součástí to, že pociťuji fyzickou bolest? Stejně tak je součástí mysli obraz, který vidím, není ale nutnou podmínkou její existence, vždyť jsou lidé, kteří nevidí.

Bereme-li jako mysl myšlení, což je pojem o něco méně vágní, a jako vědomí uvědomění (případně jako opačný stav než ztráta vědomí, ještě dále nevědomí, které je však poddruhem vědomí)², spatřujeme další pojmy, které do toho abstraktního balíčku také patří, třeba pocity či sny. Stejně tak tam budou patřit mentální schopnosti jako vcítění se do druhého, komunikace, představivost, kreativita, řešení problémů, vzpomínky.

2.2 CO JE TO UMĚLÁ INTELIGENCE?

(UI, AI: artificial intelligence)

Umělá inteligence sama o sobě není totéž jako umělá mysl či umělé vědomí, účelně není pojmenována přes tyto vágní pojmy, a ani s nimi obzvláště pro specialisty počítačových věd nemusí souviset. Slovo inteligence je podřazeno myslí, zastřešuje některé základní mentální úkony. Ačkoli ani pro informatiku není definice UI přesně daná, plně postačuje ‚uměle vytvořená entita vykazující vnější inteligentní chování‘. Cílem je jak nápodoba lidského rozhodování, tak práce s daty v procesech mimo lidskou úroveň (co se týče objemu dat či rychlosti jejich zpracování), jejich třídění na základě pravidel, schopnost se učit.

Filosof má s takovýmto tvrzením hned dva základní problémy. Zaprvé se hned začíná pít po umělé myslí, hledaje přítomnost dalších mentálních procesů. I takovou entitu nazýváme UI, vyskytuje se však častěji ve sci-fi a u technologických vizionářů než v učebnicích pro studenty informačních technologií. ‚Pozná tato entita, co je dobré? Cítí emoce či pocity? Je možné, že převezme iniciativu a vymkne se nám z rukou?‘ ptá se filosof.

² Pomocí těchto pojmů vymezuje mysl a vědomí Jay Ingram v první kapitole své knihy *Divadlo mysli*

Zadruhé, vnější inteligentní chování se zásadně liší od chování vnitřního. Filozofovi nemusí stačit, že UI jako inteligentní zvenku vypadá (behaviorismus)³, začne se zajímat, jak vypadají její mentální procesy zevnitř. Tento problém se má k tomu výše zmíněnému tak, že nezáleží, zdali je UI (navenek) ‚pouze‘ inteligentní či zdali má i jiné mentální schopnosti, podstatné je, jestli přítomnost mysli jen nepředstírá. Skutečně má ‚tu mysl, kterou mám nyní já‘? Jenže tehdy se filosof začíná topit, neboť s druhou bytostí kvůli subjektivitě mysli nikdy to, co touto otázkou myslí, nemůže sdílet.

2.3 MŮŽE STROJ MYSLET?

Nejlepší a nejkratší odpověď, která na tuto otázku existuje, zní: ‚toto je beznadějně vágní otázka, na kterou se kvůli špatné formulaci nedá odpovědět‘⁴. Přesto (možná právě proto) je nesmírně fascinující a zasluhuje si pozornost. Představuje jeden z hlavních podmětů této práce a nehledě na existenci relevantní odpovědi kolem ní bude celou dobu kroužit.

Bereme-li ji v úvahu, což není vždy přístup myslitelů, kteří se podobnou tematikou zabývají, je možno k ní přistoupit z několika úhlů:

- Umělá mysl je uměle vytvořená. Abychom mohli něco uměle vytvořit, měli bychom nejprve porozumět naší předloze. Nejdříve je nutné prozkoumat, jak myslí člověk, věnovat se výzkumu mozku a na základě získaných informací tvořit umělou entitu. Můžeme také říct, že porozumíme-li lidskému mozku a sestavíme-li mozek umělý, dosáhneme svého cíle. Odtud pochází část přitažlivosti neuronových sítí, neboť svou strukturou i způsobem, jakým fungují, připomínají naše vlastní neuronová propojení. (neurovědecký přístup)
- Pochopení lidské mysli či perspektivních technologií na poli umělé inteligence není nutné, postačí principiálně buďto dokázat, nebo zcela vyvrátit, kupříkladu skrze myšlenkové experimenty, že stroj nikdy nemůže/docela dobře může myslet. Poněkud kompromisnější tvrzení je, že budoucnost výzkumu z předchozího bodu lze tímto způsobem dopředu odhadnout. (argument čínského pokoje)
- Otázka je irelevantní. Mysl je fenomén čistě subjektivní a objektivně neprozkoumatelný u druhých bytostí. To jediné, co kdy můžeme objektivně, a tudíž

³ Viz Turingův test – kap. 1.3 (poznámka pod čarou), kap. 5.4

⁴ „Turing proposed that we abandon the question ‘Could a computer think?’ as hopelessly vague, replacing it with the question ‘Could a computer pass the Turing test?’.“
zdroj: <https://plato.stanford.edu/entries/computational-mind/> kap. 2

vědecky, posoudit, je vnější chování. Ačkoli my sami jsme si vnitřní přítomností své mysli jisti a ať již z jakýchkoli důvodů (indukčně, selským rozumem, zvykově, evolučně, nutně) předpokládáme, že ji i ostatní lidé mají, nikdy nemůžeme tuto zkušenost řádně sdílet a stejně tak ji s námi ani UI sdílet nedokáže.

Proč říkáme, že je mysl čistě subjektivní? Nelze ji sdílet přímo, jejími svědky jsme pouze my sami. Jakékoli sdílení je podřízeno médiu, které k němu používáme, povětšinou jazykem. Bereme-li jazyk jako soustavu naučených pojmů, vzniká jeho používáním výrazné zkreslení zprostředkování mysli – pro to, co sami cítíme, používáme ta slova, o kterých si myslíme, že těmto stavům odpovídají. Stejně tak máme sklon odpovídat na otázky ohledně mysli (zrcadlit) podle toho, jak mysl popisují lidé okolo nás. Neexistuje pak záruka, že ty samé pojmy nepoužíváme pro lehce odlišné mysli, aniž bychom si toho byli vědomi. Tento problém se výrazně rozrůstá, není-li tázaným člověk.

To neznamená, že nikdo další mysl nemá nebo že nemáme žádné důkazy o tom, že by ji ostatní měli. Například se do ostatních lidí dokážeme vcítit, a stejně tak se oni dokáží vcítit do nás. Tato schopnost je podmíněna přítomností mysli, neboť až když máme mysl a odpovídající mentální stavy, tak je dokážeme projektovat i do druhých bytostí.⁵ Nepředpokládáme, že by nějaká bytost jiným připisovala něco, co sama nemá, z čehož můžeme vyvodit, že schopnost vcítění ukazuje přítomnost mentálních stavů empatického tvora. Toto lze aplikovat v případě některých zvířat, u kterých pozorujeme vnější projevy tohoto procesu.

Proti této teorii se staví skutečnost, že počítač (neuronová síť) s dostatkem dat a na základě předchozích zkušeností s jinými lidmi zvládne do jisté míry předpovídat, co daný vlastník mysli udělá, aniž by mysl sám měl – stačí pouze ono vnější chování.

⁵ Ingram, Jay: Divadlo mysli, přel. H. Čížková (Praha, dubek, 2010), Teorie mysli, s. 155-157

3 ZVÍŘE – ČLOVĚK – STROJ

3.1 INSTINKTIVNÍ PŘISUZOVÁNÍ

Když se ptáme komu nebo čemu přisoudit mysl, je zcela přirozené obrátit pozornost k našim živým příbuzným. Navíc nám tento přístup může pomoci odpovědět na otázku, jak ji přisoudit. Můžeme ji hledat u tvorů, kteří jsou nám blízcí svou fyzickou přítomností a díky své povaze (nebo snad díky nám, jejich šlechtitelům)⁶ i našemu srdci, jako jsou domácí mazlíčci. Dále se můžeme zaměřit na tvory, kteří jsou našimi nejbližšími příbuznými či našimi přímými předchůdci, experimentovat na šimpanzích, s cílem přiblížit se jak našemu počátku, tak původu tohoto fenoménu zvaného mysl. A konečně nás zajímají také tvorové, kterým je připisována vlastnost pro lidskou mysl klíčová – inteligence. Zde se jedná o tvory jako delfini či sloni.

V těchto tvorech mysl instinktivně vidíme skrze jejich jednání, které nám připomíná vlastní svou vnější podobností či smysluplností a cílevědomostí, kterou jejich činům připisujeme. Neboť víme, že když my sami provádíme podobné činy, doprovázejí je také mentální procesy. Jako dobrý příklad jak podobného chování, tak velmi silného (instinktivního) pocitu přítomnosti mysli uveďme psy, jejichž činy jsou často interpretovány obdobně jako u člověka.

Některé jejich chování je charakteristické pro projev mentálních stavů. Můžeme uvést případ, kdy zvíře nakloní hlavu na stranu a zcela očividně poslouchá a „přemýšlí“ nad získanými vjemy. Nebo také situaci, kdy je z jeho pohledu patrná cílevědomost a úmysl získat to, co chce. Existují mnohé projevy radosti, strachu, bolesti či úmyslu, nebo příklady nikoli nahodilého, ale úmyslného chování, za kterým se skrývá motiv.

Obdobně můžeme sledovat vývoj inteligence mezi ostatními druhy. A nakonec, mnohá zvířata mají docela podobný mozek, přisoudíme-li tedy mozku výjimečné postavení a umístíme-li do něj naše centrum, povede to k předpokladu, že i obdobné orgány fungují obdobně. A tento pocit instinktivně zakotví v názorech i těch lidí, kteří při hlubším zkoumání s touto teorií nesouhlasí.

⁶ Zajímavou myšlenkou je, že za tuto podobnost si lidstvo může samo, a to skrze proces šlechtění a výběru. Nikoli však v tom kontextu, že by probíhal výběr nejbližší ideální představě, ale že instinktivně vybíráme také tvory, kteří jsou nám něčím blízcí i ve svém chování. Podobnou roli zastává výchova, automaticky vedená také k této blízkosti, nejen k disciplíně. Tvoříme si psi tak, abychom je měli rádi, tedy tvoříme i jejich případnou mysl.

Proto mluvím o instinktivním přisuzování myslí jiným tvorům (v kontrastu k přisuzování na základě konkrétní teorie). Nutno poznamenat, že může být v mnoha ohledech zavádějící, neboť není schopno poskytnout ani jasnou definici, ani odpovědi na případné konkretizující otázky. Na druhou stranu, přesně takto působí při prvním zkoumání samotný pojem mysl. Přesně víme, snad více než u mnohých jiných abstraktních pojmů, co to je – neboť vycházíme z pro nás (pro mě i pro tebe jakožto individuálního člověka) nevyvratitelného předpokladu, že sami mysl máme. Tvorba konkrétního popisu její přítomnosti u nás ani to, co pro nás samotné znamená, nemění. Neberme to jako důvod k jejímu zpochybnění, ale jako možnou inspiraci a cestu, jak se i skrze nejasné pojmy dopracovat k novým perspektivám. Koneckonců, přirozené mentální fenomény s sebou automaticky nenesou definici či vymezení, jako je tomu u nově zaváděných pojmů přírodních věd.

Zpět k našim adeptům na mysl: lze shrnout, že si jsme jisti tím, že jsou živí, stejně tak že prožívají nějaké základní emoce, základní potřeby – že mají hlad, že je něco bolí, ale i že něco chtějí. Respektive, jsme si instinktivně jisti. Právě protože v takové situaci chybí potřebné podkladné argumenty, nelze vyvrátit ty protichůdné – zde chceme ukázat, jakým způsobem člověk běžně nad tímto problémem uvažuje. Dále lze konstatovat, že jim nerozumíme – umíme možná rozklíčovovat jednoduché molekuly, můžeme prozkoumat jejich mozky, ale to, co je činí živými, stejně tak jako to, co nás činí námi, nenajdeme.

3.2 ROZUM A ROBOTI

Připíšme tedy zvířatům nám nejbližším cítění a zvažujme jejich mysl. Přesto je zde propast mezi jakoukoli živou bytostí a člověkem. Ačkoli na jedné straně nepochybujeme o některých mentálních procesech našich bližních, rozhodně nejsou úplně jako my. Onen faktor, který nejen člověka odlišuje, ale také vyzdvihuje (nebo právě proto, že jej odlišuje, jej vyzdvihuje) nad ostatní tvory, je jeho rozum. Za ním se skrývá racionalita, schopnost rozumně uvažovat, řešit problémy a přizpůsobit se i změně podmínek. Stejně tak rozum implikuje ‚vítězství ducha nad hmotou‘, moc být pánem svého těla, svých pudů, což zvířata dle našich předpokladů postrádají.

Vraťme se nyní, když vidíme rozdíl mezi zvířetem a člověkem, k předchozí části a pracujme dále s myšlenkou, že když už nic jiného, tak instinktivně přiděleným mentálním vlastnostem nerozumíme. Tím myslím, že je bez problémů můžeme přisoudit, aniž bychom je

uměli definovat či přesně vysvětlit, co představují. Zároveň je přisuzujeme instinktivně právě proto, že jim zcela nerozumíme.⁷ Toto vyplývá z instinktivního přístupu.

Přidejme do rovnice poslední člen – stroj. Uvažujme existenci, stvořenou ze dvou částí – člověkem navrženým hardwarem, který jí dodává schopnost reagovat v prostředí skrze umožnění aktivního pohybu a pasivního sběru dat. Druhou částí je software, člověkem do slov napsaný kód. Takovému robotovi na základní úrovni přesně rozumíme – toto tajemství má nepochybně vliv i na naše přisuzování myslí. Když vidím pod mikroskopem primitivní organismus, jak hledá potravu či jak sleduje světlo, je na tom něco fascinujícího. Nic takového však nepřipadá v úvahu u robota, ať už jej jeden naprogramuje sám či pokud to udělá specialista – jinými slovy, ať už jeho fyzické i abstraktní části rozumím či ne, a nehledě na vlastní schopnost tento proces replikovat.

Právě proto, že robotům rozumíme, můžeme je vyvíjet podle modelů, které kolem sebe vidíme v přírodě. Jenže to zároveň znamená, že to, čemu nerozumíme, jim dodat, a tudíž ani přisoudit, nemůžeme – neočekáváme, že v programu, který napsal ‚nějaký člověk‘ se náhle nad tím vším stvoří mysl, nehledě na to, jak moc nám blízké tvory napodobuje či jak *rozumným* se program stává.

Vidíme, že program může být ‚něco jako my, ale ne tak úplně‘, ačkoli úplně jiným způsobem než zvířata. Přestože u něj lze pozorovat vnější chování, nad kterým bychom se v případě živočicha pozastavili (a to i přes to, že v něm nebudeme spatřovat emoce, ale pouze logické chování), nic více mu intuitivně nepřipíšeme. Tento robot je ovšem v něčem daleko blíže člověku než zvíře. Má inteligentní chování, umí se chovat ve vnějším prostředí, pracuje s informacemi a obrovskou kapacitou. Umí učinit závěr, ačkoli jeho postupům rozumíme, což není případ našeho vlastního rozhodování. Má ještě něco více – má jazyk.

3.3 ZVÍŘE – ČLOVĚK – STROJ

Tím nám však množství používaných pojmů i množství jednotlivých vymezení roste. Načrtněme si tedy jednoduché schéma, kde uprostřed je člověk, k němuž se z jedné strany (zleva) blíží živí tvorové a z druhé strany (zprava) entity umělé. V nejbližším okolí člověka spatřujeme ‚mysl‘, část jí tedy připadá jak zvířatům, tak strojům, a každá tato část je z jiné strany a každá z nich má s člověkem společného něco radikálně odlišného.

⁷ Zde význam slova ‚rozumět‘ používám v inspiraci přírodními vědami, tedy jako schopnost popsat, aplikovat obecnou hypotézu, definovat, reprodukovat atd..

Zde se poprvé dostáváme k novému pojmu, ke kterému směřovaly stroje blízké člověku – umělá inteligence. Vedle definice a přístupů, skrze které jsme se jí přiblížili v úvodu, zde spatřujeme novou, instinktivnější a jinak vymezenou perspektivu. Jednoduše: považujeme za ni to, co je zprava hned nejbliže člověku. Sdílí s námi to, co tito roboti, obsahuje však navíc něco záhadného, nerozumíme jí. Pochází z druhé strany, je něčím, co nemá život (je to umělé), ale má rozumné uvažování (je to inteligence). Spatřujeme místo pro tuto entitu, zcela jinou otázkou však je, zdali může v našem světě existovat (je třeba ji stvořit, neboť je umělá), a pokud ano, tak jakých forem může nabývat.

To, co naopak takové entitě chybí, je něco, co nám, uvyklým na živočichy a na přírodu, připadalo od mysli i od vědomí neoddělitelné – pojmy jako život, smrt, bolest, strach, radost. Možná je lez shrnout pod střechu „životních potřeb“, jejich naplňování a boj o přežití. Umělá strana je zvláštní tím, že obsahuje něco člověkem vyzdviho vaného (tedy rozum), neobsahuje však jednoznačný zdroj základní motivace. Jeho zdrojem pro živé organismy může být příroda, evoluce, každopádně člověkem neovlivněný proces, jehož důsledkem je ustanovení přirozené rovnováhy. UI, postrádající tuto vlastnost, a přesto sdílející mnoho s člověkem, nás staví do nekomfortní situace.

Jestliže to, co mají člověk a jemu nejbližší zvířata společného, tedy když se ke člověku přiblížíme „zleva“, je skutečnost, že jsou naživu a že prošli evolucí, motivem, který spojuje člověka a umělou inteligenci, je nepochybně rozum. Není to však jen on – největším rozdílem mezi námi a cítícími zvířaty je naše schopnost ovládat jazyk, který ovšem spatřujeme i u strojů. Pokud jej používají, mají i schopnost mu rozumět? A je toto propojení vůbec na místě? Znamená pro ně to samé, co pro nás?

Rozeberme si tedy tyto dva pojmy – evoluci a jazyk o něco přesněji, jako něco, co zvířata i jiní živí tvorové mají a stroj nemá, a jako něco, co člověk má a stroj možná má, jako znak života i znak rozumu.

3.4 NĚCO JAKO MY

Ještě před tím je však třeba se zpětně ohlédnout za přístupem, skrze nějž jsme tyto dva pojmy (evoluci a jazyk) získali, a odhalit jeho případné nedostatky, aby nezpůsobily rozpory v následujícím zkoumání.

Všimněme si, že jsme lepší definici těchto pojmů zvíře – člověk – stroj našli až v jejich vzájemném propojení a vymezení. Tento přístup vyplývá z užitečnosti a úspornosti práce s již

známými představami při tvorbě a popisu představ nových. Jestliže se v již existující síti pojmů vyskytne mezera, nejjednodušší je do ní libovolný pojem umístit a získat jeho význam skrze významy těch ostatních a jeho propojení s nimi. V tomto je možné spolehnout se na intuitivnost, o to spíše je však potřeba se proti ní vymezit. Princip totiž funguje poměrně spolehlivě pro pojmy, které do našeho rámce zapadají, v případě zkoumání existence myslícího stroje se nevyhnutelně dostáváme do oblasti, která je mnohdy velmi neintuitivní, a jestliže byly v minulosti učiněny závěry, vyplývající z tohoto přístupu, že žádná taková mysl neexistuje, chceme-li alespoň nějakou nalézt, nesmíme postupovat stejným způsobem.

Kdybychom však jako důvod, proč je pro nás celé téma „intuitivního přisuzování mysli“ tak oblíbené, uvedli pouze lidský zvyk prozkoumávat okolní svět skrze sebe samého, nevysvětlili bychom ani zdaleka popularitu této myšlenky. Ta klíčová otázka spočívá v tom, zdali existuje někdo další „jako my“. Tento případ je o tolik odlišný od ostatního přirovnávání k věcem právě kvůli tomu, jakou mají naše mysl, naše schopnosti a náš rozum úlohu a důležitost v našich životech. Být „jako my“, totiž obsahuje i otázku „kdo jsme my“. Je velmi těžké komukoli jinému udělit tuto poctu, neboť přisouzení mysli, byť odlišné ve svém fungování, ale stejné v tom, jak jí rozumíme, znamená přisoudit také mnoho jiného. Stejně tak ale ovlivňuje naše vnímání pojmu „my“ či „lidství“ (býti člověkem): neboť není lidství definováno přes osobní identitu, kterou nejvíce pocítujeme vlastním „já“, našimi myšlenkami, tedy naší myslí? Řekneme-li, že máme umělou inteligenci s myslí, neříkáme tím také (neúmyslně, bokem), že má tato AI daleko více? (Duši? Pocity? Touhy? Utrpení?) Jelikož je však těžké jí toto všechno přisoudit a protože se tak děje kvůli propojení pojmů a ne záměrem, je daleko komplikovanější pronést onu původní frázi „máme umělou inteligenci s myslí“.

4 EVOLUCE

4.1 ASPEKTY MYSLI

Podívejme se tedy na první ze zvolených témat, na život a evoluci.

Mnohé aspekty lidské společnosti jsou umožněny myslí. Díky schopnosti myslet má člověk kulturu a umění, tvoří komplexní komunity s dělbou práce, dosahuje vědeckého pokroku. Tyto jevy však nejsou její podmínkou, ale důsledkem. Stejně tak se zdá, že mohou mít i jinou příčinu než naši mysl: včelí společenství spolu také nepotřebuje komunikovat slovy a vysvětlovat si úkoly, přesto je schopno spolupráce; bereme-li schopnost vytvářet takto složitou a fungující strukturu jako důkaz přítomnosti podobného mentálního fenoménu, jedná se o mysl výrazně odlišnou.

Abychom mohli získat umělou mysl, nesmíme od ní tyto důsledky požadovat. Ovšem vztah podmínka – důsledek není v tomto případě zcela jednoznačný: mysl se stejně dobře může vyvíjet společně s těmito schopnostmi. Včelí společenství má mysl právě díky své struktuře a díky společenskému vývoji lidí byl umožněn také vývoj jejich mysli. Nenacházíme se však na prahu věčného cyklu vejce-slepice. Nehledě na tyto vztahy si lze představit mysl s jinými schopnostmi, než je tato, a stále to mysl bude, implikace zde totiž neplatí obousměrně.

Není potom mysl jen kombinace náhodných aspektů? Příkladem takového konfliktu s definicí budiž diskuse ohledně definice náboženství. Jedná se o podobnou situaci, o vymezení pojmu skrze snahu přijít na to, co všechno pod něj spadá a co ještě je mysl/náboženství a co už ne. Například, součástí náboženství je rituál, posvátný text, božské bytosti, organizace atd. Jde o sadu pojmů (v případě náboženství, v případě mysli o schopnosti jako vypsané výše), jejichž kombinaci považujeme za náboženství. Problematické je, že snad vždycky lze najít výjimky. Po vyřazení pojmů, které neodpovídají těmto výjimkám, u kterých panuje obecné přesvědčení, že jsou náboženstvím, zůstaneme s pouze vágní definicí. V konkrétním případě to znamená, že nalezneme náboženství, které je jednoznačně za náboženství považováno, ovšem nemá žádný posvátný text; jsme nuceni zúžit naši definici pouze na božské bytosti, organizaci (atd). Tímto způsobem může dojít ke vzniku vágní kategorie, do které následně spadají i věci, které náboženstvím rozhodně nejsou.

Pokud by mysl byla obdobnou kombinací aspektů (nehledě na jejich problematiku v případě náboženství), tak to potvrzuje výchozí hypotézu, že nejde primárně o důsledky, ačkoli je vidíme jako naprosto zásadní pro naši vlastní mysl. Obdobně lze tvrdit, že ani u mysli

nedokážeme vybrat několik základních aspektů, které budou představovat každou mysl a nic jiného, ačkoli se myslí z takovýchto součástí skládají.

Proč tento přístup pro mysl funguje? Lze si představit mysl, které mají jen některé aspekty, ačkoli jejich existence je v některých případech kontroverzní. Příkladem mohou být již zmiňované včely, které odpovídají požadavku na komunikaci a sociální uspořádání, nemají ale patrný jazyk a naši úroveň komunikace mezi dvěma jedinci. Jsou však také příklady, kdy sice nacházíme jen některé aspekty, ale přítomnost myslí zpochybnit nelze. Blízké této představě jsou také lidské kultury, které vytváří společenskou strukturu odlišnou od té naší a s vnímáním světa radikálně odlišným od případného čtenáře. Mezi tyto aspekty patří vědecký pokrok, vnímání a měření času, jiná schopnost se vyjadřovat atd. Pro představitele takto odlišných skupin je velmi obtížné se dorozumět a myslí, které mají, jsou alespoň v něčem radikálně odlišné.

Toto potřebujeme rozlišit z toho důvodu, že jinak bychom mysl nepřipisovali perspektivním entitám, protože by nikdy nemohly kompletně splnit naše podmínky. Chceme se vyhnout situaci, kdy vytvoříme potenciální UI, která má například správné vnější chování, mysl jí však nepřipíšeme z nějakého odlišného, a hlavně irelevantního důvodu, třeba protože neumí být náležitě smutná. Tím myslím, hledáme jakoukoli mysl, neboť žádnou další stejnou mysl, jako máme my, neznáme. I kdybychom vzali v potaz včelí mysl, budou jí mnohé naše vlastnosti stále chybět. Existuje-li tedy v našem okolí, či bude-li existovat, bude jiná.

4.2 UMĚLÁ MYSL A ŽIVOT

Právě proto nemůžeme po strojích automaticky požadovat, aby byly živé. Pokud u nich chceme hledat mysl, je třeba připustit existenci takové myslí, pro kterou nebude život podmínkou. Přijmeme tedy, že představuje pouze aspekt, ačkoli klíčový, reprezentující obsáhlou kategorii, nikoli však princip samotný. O to důležitější je u myslí, která tento aspekt nemá, upustit také od všeho, co s ním souvisí, uvědomit si tuto skutečnost a mnohé projevy myslí nebrat při tomto hledání za samozřejmost.

Je pravda, že i tímto směrem se ubírá jedna cesta, ve které tvoříme umělou mysl tak, aby byla co nejblíže té živé, není však totožná s cestou naší (snaha o tvorbu živého stroje a snaha o tvorbu UI je odlišná). Představme si existenci UI, která má mnohé mentální schopnosti a splňuje některé kritéria živých organismů (reakce na podněty okolí – práce s informacemi, schopnost vývoje – schopnost se učit, změna), vymyká se však jiným (příjem a vylučování

potravy, aktivní pohyb, dědičnost a s ní spojené rozmnožování – existence druhů nad jedinci). Mysl mezi tato kritéria ani zdaleka nepatří, ačkoli je jimi silně ovlivněna (jakožto něco, co vzniká na základně živého organismu). Vezměme si způsob, jakým používáme slovo ‚robot‘ a ‚stroj‘ – něco slepě plnící příkazy, něco bez vlastní vůle, ale i něco bez citu. UI v současnosti postrádá touhu po životě, která představuje veliký motor pro vývoj mysli živých organismů.

To neznamena, že UI podobnou potřebu nemůže principiálně získat; odhaluje však důležitou směrdatnou sílu, kterou současná vývojová stadia UI postrádají, nemůže je tedy vést tak, jako vedla živočichy. Tato touha způsobuje další svazek pocitů, všechny ale představují pouhé aspekty mysli, ne její podstatu. Je-li příčinou bolesti, strachu, utrpení, potřeby síly, pak se tyto stavy o to složitěji přenáší do potenciálních myslících strojů. Tím pádem nikdy nemůžeme získat stejný typ entity uměle vytvořené a živé.

4.3 EVOLUCE – ZVÍŘE

Teoreticky může UI existovat ve zcela virtuálním světě bez jakékoli fyzické schránky (či fyzické podoby, která je schopna vnímat své okolí). Oproti tomu má život chování v prostředí, které navíc není přátelské. Představuje zdroj obživy, ale také nebezpečí. Toto prostředí nutí vyvíjející se organismus k inteligentnímu chování, chce-li přežít a zdokonalit se. Nacházejí se v něm také rovnocenní partneři, s nimiž jsou tvořeny vztahy jako lovec a kořist, matka a mláďata, smečka.

U inteligentního jednání zvířat můžeme experimentálně zkoušet, nakolik je prohlédnutelné, a tudíž podobné nemyslicímu stroji. Někdy vzniká pouhou kombinací jednoduchého činu (krab jde při hledání cesty zpět z kopce, případně se řídí směrem tekoucí vody) a vhodného prostředí (krab se nachází na skále u moře, kde jsou příhodné podmínky, bez kterých by byl ztracen). Nakolik je tento krab⁸, který se skoro vždy dokáže vrátit zpátky do moře, inteligentní? Tento ‚trik‘ ztratí na účinnosti při změně prostředí, kdy se zvíře i nadále chová zvykově a jeho opakované chyby naznačují, že za jeho činy nestojí pochopení situace, nýbrž že část jeho inteligence je externí, spočívající v jeho běžném okolí.

Tyto schopnosti se nadále rozvíjí a stávají se stále přesvědčivějšími, až se dostáváme ke člověku, jehož jednání bez pochyb představuje vrchol v současnosti existujícího řetězce. Vraťme se ale zpět ke dvojici lovec-kořist. Každý z nich obhájí svou primární potřebu zůstat

⁸ Zdroj příkladu – TED video o různých podobách inteligence a funkcionalismu, bohužel jsem jej zpětně nedokázala dohledat.

naživu, snaží se navzájem přečtyračit. Za pavoučí sítí si můžeme představovat pavoukovou analýzu: moucha prolétá vzduchem, když do něj postavíme překážku, může se v ní moucha zachytit. Za ptačí matkou simulující zlomené křídlo, aby ochránila mláďata, se toho může skrývat více: dravec útočí na tvory, kteří jsou zranění či oslabení. Zraněný pták se chová tak a tak. Tím, že se odláká pozornost dravce, zůstanou mláďata v bezpečí. Toto není obyčejná past, je to podvod a předstírání, což vyžaduje daleko složitější mentální schopnosti.⁹

Takové chování plyne z nutnosti. Ptačí matka nemá na výběr, zdali bude svá mláďata chránit; pokud něco ‚nevymyslí‘, zemře ona i její druh.¹⁰ UI sice může simulovat stejné chování, avšak z jiných důvodů – můžeme jí je nadiktovat, předepsat, tíhu reálného prostředí ale nepocítí. Není tedy přirozená cesta, jak by měla výše popsané inteligentní vlastnosti získat.

4.4 EVOLUCE – STROJ

Lze nějakým způsobem UI přeci jen do evoluce zařadit? V předchozí kapitole evoluce představovala způsob výběru přežívajících a úspěšných organismů a tlak na ně. Tento termín se dá také použít jako synonymum pro vývoj, pro posun vpřed, který si na rozdíl od prvního významu s UI spojujeme. V tomto kontextu ji můžeme považovat za nové evoluční stadium, tj. za entitu vyspělejší a vhodnější současným podmínkám než jiné entity.

Spojujeme si s ní pokrok, vylepšení životních podmínek a zdokonalení. Stejně tak patří dokonalá UI do pomyslné hierarchie nad člověka, díky jejím současným výpočetním a paměťovým kapacitám, i díky její stvořenosti, která by ji oprostila od mnohých lidských nedokonalostí. Jak již bylo zmíněno, UI nemá s běžným pojetím evoluce mnoho společného, zejména kvůli svému původu. Můžeme ji tam však zasadit ve dvou kontextech.

Samotná umělá inteligence do přírodní evoluce nezapadá. Stejně jako pro člověka byl velký krok používání nástrojů, neolitická revoluce či revoluce průmyslová a můžeme tedy tyto posuny označit za evoluční, bude, a částečně pro něj již je, podobným milníkem využívání jakékoli formy UI. Jestliže dnes ji používáme jako nástroj pomáhající při výkonu složitých úkonů (letecká doprava, medicína (chirurgie), věda¹¹, expertní stroje), vize sebevylepšení ji

⁹ Dennet, D. C.: Druhy myslí, přel. J. Peregrin, Praha, Academia, 2004, kap. 5 – nemyslicí přírodní psychologové

¹⁰ Mimořádně, zůstává otázkou, nakolik je klíčové, jak moc si živočich uvědomuje své vlastní chování, přestože schopnost pomyslet si, že si něco myslím hraje důležitou přelomovou roli v úrovních myslí.

¹¹ Současné experimentální fyzika se bez strojového zpracování dat neobejde, ze stejného důvodu jsou dnes bioinformatičtí tolik žádoucí.

staví skoro na roveň člověka, kterému konkrétně slouží. Takovýto člověk, jehož součástí je UI, by pak představoval vyšší vývojové stadium a začlenil ji tak do evolučního žebříčku.

Zajímavější a dřívějšímu problému bližší je představa, že evoluce nemusí probíhat samovolně, ale že vědomá evoluce je také možná, tedy další pojetí významu slova evoluce. UI celkem jasně splňuje podmínku vývoje směrem vpřed a funkce v současném světě, do kterého s vývojem informačních technologií UI stále lépe zapadá. Za nám známé historie života evoluce také neprobíhala konstantně; jsou tu období změn velmi pomalých (kupříkladu v počátku), změn náhlých či nečekaných (věčná otázka osudu dinosaurů). Už jen historie člověka je pozoruhodná a zdaleka ne snadno interpretovatelná. Je-li UI v současnosti sice na vzestupu, ale nikoli na vrcholku, nevylučuje to její jakoukoli budoucí úlohu.

Vedle toho můžeme uvést ještě jednu zajímavou myšlenku, byť je částečně v rozporu s tou předchozí. Není zřejmé, že by UI měla schopnost se (samostatně) vyvíjet, z časových důvodů ovšem nelze potvrdit, že by tomu bylo naopak. Jen díky naší dlouhé historii na zemi můžeme něco soudit o ostatních organismech, jejichž stav je v měřítku lidského života konstantní. Navíc je naše schopnost vidět do budoucnosti omezená. Při pohledu na organické molekuly nic nenaznačuje, že by mohly být základem pro lidský mozek vybavený myslí, jak tedy můžeme tvrdit, že komputacionalistický model je nedostatečný pro tvorbu něčeho obdobného jen na základě jeho primitivních stavebních jednotek?

5 JAZYK

5.1 ÚVOD DO JAZYKA – SCHOPNOST VYPLNIT POVEL

V minulé kapitole jsme došli až ke zvířatům, která jsou schopna poměrně komplexního chování, za kterým lze spatřit mentální procesy. Jejich schopnost cítit jsme nerozebírali dopodrobna, pouze jsme ji přímo propojili se skutečností, že jsou naživu, že se pohybují ve světě a že jsou hnáni touhou přežít. Stejně tak jsme zjistili, že je třeba nalézt platformu, na které založíme mysl stroje, když stroj evoluční nutnost mysli nemá.

Nyní se dostáváme zpět do bodu, kdy člověk zvířata překoná a kdy vzniká rozdíl mezi jejich myslí. Jedna ze schopností, jež činí člověka speciálním a která jej odlišuje od zvířat, je jeho dovednost řešit abstraktní problémy. Daniel Dennett ji ve své knize *Druhy mysli* ukazuje na jednoduchém příkladu při provádění experimentů týkajících se mysli: člověku, účastnícímu se experimentu, stačí vysvětlit, o co se jedná a jak se má chovat a on toto provede. To práci s ním činí výrazně snazší než s ostatními, byť inteligentními zvířaty, která je třeba spíše do požadovaných situací nalákat. Podstata tohoto rozdílu spočívá v lidské schopnosti vytvářet abstraktní pojmy a pak s nimi nadále pracovat.

Toto člověku umožňuje pouze jazyk, skrze něj získává obrovskou výhodu nad ostatními zvířaty, která tuto propast nikdy nepřekonají, nehledě na jejich jiné schopnosti či na to, zdali jim mysl přisoudíme, dokonce nehledě na jejich komunikační dovednosti. Lidské mysli se tak otevírají zcela nové obzory. Náhle je schopna nějakým způsobem uchopit vnější svět a přesunout jej skrze pojmy sama do sebe, kde s nimi nadále pracuje.

Jazyk jsme naposledy zmiňovali v souvislosti s jeho blízkostí člověku i stroji¹², jako jejich společnou schopnost, nevěděli jsme však, zdali to není pouhá podobnost v označení. Paradoxně je pro něj řešení příkladové situace typické; je to dokonce jeho klíčová vlastnost. Nadiktujeme, co je třeba učinit a stane se tak, obdobně jako u člověka. Pracuje tedy s abstraktní sadou pojmů a je schopen tyto příkazy vykonávat a dávat nám odpovídající řešení. Samozřejmě je tu na místě alarmující otázka, totiž, zdali těmto pojmům nějak rozumí. Nehledě na jakkoli složitou interpretaci vnímání stroje v této situaci porozumění vzniká, což lze ukázat na velmi jednoduchém příkladu, možná až příliš banálním. Zadám-li stroji pokyn, stroj mu nějak rozumět musí, neboť jej vykoná. Podobně, když dám psovi pokyn ‚sedni‘, pes je schopen tak učinit.

¹² Kap. 3.2

Nezáleží na konkrétní podobě tohoto symbolu, pes nemusí vědět, co pro mě znamená toto slovo, není ani důležité, co přesně slyší, ani jej neumí aplikovat v jiné situaci, jako to člověku umožňuje jazyk. Nějaké porozumění přesto vzniká, protože pes si skutečně sedne. Obdobně pak funguje i počítač.

5.2 KTM

Komputační teorie mysli (KTM) propojuje lidský mozek (lidskou mysl) a stroj (mysl UI, která vzniká na jeho základě). Podle ní tyto dva systémy fungují na stejném principu výpočtu – můžeme si představit, že náš mozek provádí operace obdobně, jako probíhají na počítači. Ačkoli tato teorie si klade za cíl vysvětlit, jakým způsobem funguje mysl, vyplývá z ní také, že není principiální důvod, proč by UI nemohla existovat.

Vedle toho KTM odráží jeden z klasických přístupů pro vztah jazyka a stroje; pracuje s modelem, který manipuluje symboly¹³. Převedení mozku na stroj manipulující se symboly představuje zdroj kritiky některých filosofů, jako je John Searle. Tento proces manipulace totiž probíhá na základě syntaxe. Zde se symbolem rozumí jakýkoli znak, jehož význam je navíc v tomto procesu irelevantní. Stroj se znaky dle jemu daných pravidel pohybuje, jejich interpretace je pak pouze na nás.

Absurdním příkladem, který Searle uvádí jakožto upozornění na rozdíl mezi komputační strukturou a skutečnou myslí, je zed¹⁴, jejíž struktura je uspořádaná tak, že se nám jeví jako jisté smysluplné vzorce. Argumentuje, že probíhající výpočet záleží na pozorovateli, vidí-li tedy pozorovatel ve zdi významy, je to jeho subjektivní názor. Stejně to funguje i s počítačem, ale nemělo by to tak být se skutečnou myslí. Zed' samotná pak nejen že nemá uvědomění, že by něco dělala, nic ani reálně nedělá. Avšak při řešení složitějších problémů musí mít naše entita vlastní systematičnost. Tehdy podobně absurdní případy odpadají. Pro vykonání úkonu musí UI znát vztahy mezi jednotlivými věcmi, ať už v jakékoli podobě. Tato spojení a podobný systém máme přeci i my, jen to nemusí být vědomý proces a tato síť je velmi komplikovaná.

¹³ Turingův stroj

¹⁴ Searle; uvádí jako součást argumentace ohledně syntaxe a sémantiky; pomocí zdi ukazuje, že syntax a komputace jsou jevy závislé na pozorovateli (observer-relative). Zdroj: <https://plato.stanford.edu/entries/chinese-room/>

5.3 DETERMINISMUS

Jestliže KTM odpovídá některým mentálním pochodům, je jen těžko představitelné, že by lidské myslí odpovídala kompletně. Vysvětlíme-li skrze ni schopnosti, které dnes odpovídají mozku i stroji, stále nám zbývá obrovská skupina schopností jiných, které takto napohled vysvětlitelné nejsou¹⁵. Můžeme si ovšem představit, že všechny lidské mentální vlastnosti budeme v budoucnu schopni takto převést. Z tohoto předpokladu vyplyne mnoho skutečností o naší současné myslí, KTM ji totiž představuje deterministicky. Žádná kreativita či originalita neexistuje, ve skutečnosti mysl funguje jako velmi složitá matematická funkce. Tím pádem lze jakýkoli mentální proces převést na funkci a ve formě funkce jej následně předat UI.

Případ této konkrétní vlastnosti podpoří myšlenka, že člověk neumí vymyslet nic za hranicí toho, co již zná¹⁶. Pouze kombinuje prvky, které již viděl, případně jej napadnou nové způsoby, jak je dát dohromady. Konkrétním příkladem jsou třeba středověké bestiáře a mýtická stvoření, které nepředstavují nic radikálně odlišného, pouze kombinaci již známých znaků. Jiným jsou lidské sny, kdy kolikrát neplatí mnohé principy, které při každodenním vnímání světa považujeme za samozřejmé (jako logická či prostorová návaznost událostí). Přesto pracují s námi známými prvky, pouze mění souvislosti mezi nimi.

Vracíme se k tomu, že teorie, kterou si zvolíme, ovlivní nejen to, jak uvažujeme o cizích myslích/co ještě považujeme za cizí mysl, ale také to, co je naše mysl a co jsme my jakožto lidé. Tato myšlenka se již jednou vyskytla na konci kapitoly Něco jako my. Přisoudíme-li jakékoli entitě mysl, ovlivní to naše vlastní vnímání lidství. A zároveň ‚přisouzení myslí znamená přisoudit také mnoho jiného‘, nehledě na již zmiňované nedostatky spoléhání se na tento¹⁷ proces. Tentokrát tento jev spatřujeme v opačném směru, kdy při aplikaci teorie získáváme mnoho dalších důsledků pro nás samé, tedy pro to, co známe ze všeho nejlépe.

Tento argument se přesto nezdá být postačující pro vyloučení výjimečnosti lidské schopnosti generovat umění a kulturu, případně náboženství. Zdá se, že v tomto všem je přeci jen něco více. Vedle toho, že se jeví nemožným něco takto komplexního a nepopsatelného¹⁸ uměle vytvořit, je nepředstavitelné, že by něco takového vznikalo na principu sčítání či podobně

¹⁵ Například kreativita (umění, vědy, tvorba hypotéz, víra) či morálka (spravedlnost, vztah k jinému člověku). Dále také práce s nejistotou (Nevím, jestli je nějaké informace pravdivá, nevím, jak někdo zareaguje, nevím, co se stane v budoucnosti. Naše životy jsou plné nevěděni.).

¹⁶ Zdroj nějaké hodiny

¹⁷ Kapitola 3.4

¹⁸ Vzpomeňme, že přirozené mentální fenomény s sebou automaticky neunesou definici či vymezení (kapitola Instinktivní přisuzování, 3.1).

jednoduché operace. To je však pouze instinktivní tvrzení, kterým nelze vyvrátit, že tomu tak pro mysl založenou na výpočtech být nemůže. Ačkoli jsme tak postupovali v případě zvířat, v tomto případě nelze automaticky použít instinktivní argumentaci jako součást neinstinktivního postupu.

KTM předpokládá principiální identitu mozku a stroje úplnou, nikoli pouze metaforickou.¹⁹ Biologické procesy mohou někdy fungovat obdobně, nefungují tak ale vždy. Informatika a biochemické reakce představují odlišné systémy, tímto způsobem vytvoříme odlišný druh mysli, která spadá do stejné kategorie, neobsahuje však stejné schopnosti. Lidská kreativita je tedy v rámci tohoto názoru nepřeveditelná, nejen proto, že věříme v existenci něčeho většího ve člověku (tedy instinktivně), ale také proto, že u nás nemusí vycházet z počítačového modelu. Ten má nepochybně své výhody a umožňuje pokrok v pochopení lidské mysli, představuje další platformu. Otázkou zůstává, zdali je tato platforma obdobná té, kterou tvoří život a zdali na ní obdobně může vznikat určitý typ mysli. Argumenty proti dostatečnosti manipulování symbolů cestu tímto směrem znesnadňují:

5.4 ČÍNSKÝ POKOJ

Existuje slavný myšlenkový experiment, propojující jazykové schopnosti UI a její počítačový model. Představil jej americký filosof John Searle v roce 1980, mimo jiné jako odpověď na KTM. Na tomto experimentu Searle simuluje vnitřní prostředí UI a ukazuje, že pouze vnější projev není dostačující. Agentem je UI, která splňuje vnější podmínky Turingova testu – je schopna s experimentátorem komunikovat tak, že nepozná, že se nejedná o člověka, ale o stroj²⁰. Searle argumentuje, že vnitřní pochody tohoto stroje nepřipomínají naše vlastní a že stroj slovům, které používá, nerozumí. Z toho vyplývá, že také nemá mysl (vzpomeňme, že náš model mysli stroje stojí daleko více na jazyce/na symbolech, než na jiných mentálních schopnostech).

Searle pracuje s modelem, který jednoduše manipuluje se slovy na základě jejich vztahů k jiným slovům, tedy syntaxe. Můžeme si tato slova představit ve formě přirozeného jazyka nebo jako soustavu matematických výrazů, které přenášejí souvislosti (nikoli však významy,

¹⁹ To znamená, že mozek skutečně provádí výpočty. Zdroj: <https://plato.stanford.edu/entries/computational-mind/> kap. 3. Vedle toho existují jiné verze KTM, propojené s dalšími teoriemi (funkcionalismem, RTM a jazykem mysli atd.), které k tomuto problému přistupují odlišně.

²⁰ Toto je Turingův způsob reformulace úvodní otázky – „Může stroj myslet?“.

podle některých názorů). Jeho tvrzením je, že takovýto program nerozumí slovům (symbolům, znakům), která používá, nerozumí tím pádem ani větám, které přijímá a které formuluje.

Při přeložení tohoto modelu do případu čínského pokoje, Searle vystupuje jako agent, který je schopen smysluplně odpovídat na otázky v čínštině. Sám jí ovšem nerozumí, neumí ani přečíst čínské znaky; požadovaného outputu dosahuje jednoduchým následováním pravidel a instrukcí, které lze považovat za počítačový program.

5.5 PROSVĚTLENÝ POKOJ

Searle skutečně ve svém myšlenkovém experimentu čínštině nerozumí. Na tuto situaci můžeme nejnázne odpovědět dvěma způsoby. Jednak zkusíme podmínky ČP změnit tak, aby Searle mohl čínštině porozumět, nebo řekneme, že v tomto systému Searle sice nerozumí, ale toto je dáno nikoli nefunkčností systému, ale Searlovou pozicí v něm. Pokud bychom ČP převedli na počítač, byl by Searle procesor, pouhá výpočetní jednotka. Entita, která čínštině rozumí, je od této jednotky odlišná, ať už ji tvoří pokoj jako celek, některá jiná jeho součást nebo vzniká-li v nějaké jiné, abstraktní rovině, jako nadstavba tohoto fyzického modelu.

První typ odpovědi se zabývá interakcí ČP (respektive obdobného systému) a vnějšího světa. Navrhuje pořídit mu vnější senzory a vyslat je do Číny. Porozumění pojmů je pak docíleno skrze interakci s nimi, skrze interakci se světem, aplikaci těchto slov v různých kontextech atd. Není třeba, aby toto vnímající tělo bylo fyzické, stejný typ vjemů mu může poskytnout prostor virtuální. V obou případech spočívá hlavní problém interakce v převodu informací z externí podoby do vnitřního prostředí robota, nezáleží, jestli byla tato data původně ve formě fyzických informací ve světě, nebo obdobných informací již převedených do virtuální podoby.

Můžeme také namítnout, že Searle zde zůstal uvězněn v dříve používaných intuitivních pojmech, aniž by uměl změnit perspektivu pro jiný typ mysli. Pro lidský mozek nemáme dostatečně zodpovězenou otázku, jak vlastně porozumění vzniká, jak tedy můžeme vědět, jestli v ČP toto porozumění není přítomno – jen v trochu jiné podobě, než by Searle předpokládal.

Abychom zde mohli ukázat, proč instinktivní přístup, uplatňovaný na zvířata, zde vede na scestí, je třeba objevit rozdíl mezi těmito obory. V případě mysli stroje či níže zmiňované fyziky aspirují tyto obory o posunutí hranic lidského vědění a chápání. Oproti tomu jsme se v případě zvířat pohybovali daleko více v oblasti, kde změna perspektivy a vnímání není nutná.

Velmi hezký příklad analogický ke kompozici Searlova argumentu prezentovali Paul a Patricia Churchlandovi v časopise *Scientific American* (1990), kde namísto čínského pokoje vytvořili prosvětlený pokoj (the Luminous room), který problematiku pochopení mysli připodobňuje k teoretické fyzice. Funguje na principu, že mnohé fyzikální teorie, které jsou obecně považovány za dokázané a platné, představují poměrně bizarní pohled na svět, ke kterému musíme přistupovat s jistým odstupem – věci jsou někdy ve své podstatě velmi zvláštní a těžko zařaditelné do našeho intuitivního světa (se kterým pracuje Searle). „*Searle is once more mistaking the limits of his (or the reader's) current imagination for the limits on objective reality.*“²¹

²¹ Churchland, P. a Churchland, P., 1990, 'Could a machine think?', *Scientific American*, 262(1) s. 37

6 ČLOVĚK – JAZYK – STROJ

6.1 MYSLÍME V JAZYCE?

Základní funkcí jazyka je mezilidská komunikace. Mám na mysli primitivní řeč, jak ji nazývá Ludwig Wittgenstein v úvodu do svých Filosofických zkoumání, ze kterých je také následující citát. Její podstatou je možnost naučit se jazyk pomocí toho, že ukazujeme věci, a představa, že slova zrcadlí či reprezentují vnější svět. Odtud vzniká jedna z verzí vysvětlení, jak naučit dítě/člověka jazyku.

„Když dospělí pojmenovali nějaký předmět a obrátili se přitom k němu, viděl jsem to a chápal jsem, že zvuky, které vydávali, onen předmět označují, protože na něj chtějí ukázat. Tento jejich úmysl vycházel však najevo z pohybu těla; ten je jakoby přirozenou řečí všech národů, spočívající v mimice a ve hře očí i v pohybu ostatních údů, přičemž zvuk hlasu naznačuje hnutí duše, když ta po něčem touží, nebo něco má, nebo něco odmítá, nebo před něčím prchá. Takto jsem se postupně učil chápat, jaké věci označují slova umístěná na svých místech v různých větvích a opakovaně slýchaná, a když si i moje ústa navykla na tyto znaky, vyjadřoval jsem jimi svoji vůli.“²²

Tento náš jazyk slouží evoluční potřebě se domluvit při získávání potravy, při péči o mladé či při organizaci společenství. Na druhou stranu, jazyk, který používáme v každodenním životě, je těmto možná původním funkcím svou složitostí i účelem velmi vzdálen.²³ Jazyk nevyužíváme pouze pro dorozumívání se s druhými, funguje také jako médium uchováající myšlenky (schopnost psát), nebo jako prostředek myšlenek samotných – to v okamžiku, kdy se naše vlastní myšlenky odehrávají ve slovech jazyka, kupříkladu kdy si doslova pomyslím ,to je ale pěkné počasí‘.

Mimo vnější svět se tedy jistá rovina mysli odehrává v jazyce. Zde jazyk funguje jako způsob uvažování nad světem, jako způsob uchopení věcí a významů. Skrze jednotlivé věty, které nemusí nijak souviset s čímkoli fyzickým či vnějším, uvažuje člověk sám o sobě. Otázkou je, zdali takto funguje jen určitá, byť velmi důležitá část mysli (jedna část vědomé složky tak funguje nepopíratelně), nebo jestli je pro ni jazyk ještě daleko důležitější.

²² citace Augustina z Confessiones I/8, Wittgenstein, L: Filosofická zkoumání, přel. J. Pechar (Praha, Filosofický ústav AV ČR, 1993) s. 13

²³ tamtéž

Existují totiž jednoznačně nejazykové vjemy, které ve svém vědomí člověk vnímá. Jedná se o celou vizuální oblast (člověku stačí obraz, ne proud slov), případně jiné informace získané ze světa. Třebaže si je skrze slova neuvědomíme, tak je skrze ně jakožto významy vnímáme.

Představme si, že vidím libovolný obraz. Ten vnímám jako vizuální informaci, avšak když se někdo zeptá, co je na tomto obraze, odpovím mu skrze pojmy jazyka. Tato slova nevznikají v okamžiku tázání, nemusím zpětně přemýšlet nad tím, co byly ony objekty na obraze (například hrnek, lampa a stůl), není třeba jej zpětně rozšifrovávat. Už od okamžiku vnímání je v něm vidím. Obdobně toto funguje i pro složitější pojmy.

Složitější je případ umění, které není tvořeno v jazyce (hudba, obrazy), případně jazykového umění, které není jednoznačně popsitelné jinými slovy, jako například poezie. Jedná se zde totiž o emoce, které se jazykem popisují daleko obtížněji, ať už jej přesahují, nebo vyžadují nejprve interpretaci pro možnost jejich převodu do běžnějších výrazů – slovy označený vjem se takto stává sekundárním, pouze interpretací své předlohy v mysli. Vystává obdobný problém jako konflikt v kapitole o determinismu, kdy existuje sféra nezapadající do výše nastíněného rámce. Můžeme jej však interpretovat jako další důvod, proč stroj, jehož případná mysl stojí na jazyce, tyto aspekty mít nebude, nikoli jako argument, proč tato teorie nefunguje.

Ani jeden z těchto případů však nevylučuje tvrzení, že jazyk skutečně tvoří hranice a tvary naší mysli. Nehledě na to, zdali člověk vnímá v něm, pohlíží na svět *skrze* něj. Obdobně jakákoli zpětná reflexe světa, důležitá mentální schopnost, probíhá v jazyce. Jazyk tvoří platformu, na které je založeno naše vnímání reality, tedy i my samotní. Dokonce, vzhledem ke kredibilitě důležitosti jazyka tak, jak byl právě popsán, můžeme schopnost jak vytvářet takovéto pojmy, tak jim porozumět, považovat za obdobu druhé dříve vyvozené platformy mysli, tedy života, na které se může založit mysl stroje.

Jiné jazyky, ať už ty podobné naší mateřštině (angličtina), jazyky velmi odlišné svou strukturou (japonština) či jazyky vědy (matematika, případně jiné vzorce a symboly) představují nové abstraktní pojmy, se kterými lze na této platformě manipulovat. Tvoří dohromady jeden jazyk jakožto systém významů, zcela nezávisle na konkrétním jazyce, dokud má přiřazený význam.

6.2 MYSLÍ STROJ V JAZYCE?

V kontrastu k tomuto přirozenému vztahu jazyka a lidské mysli používají stroje jazyk odlišně. Jeho základní funkcí samozřejmě nemůže být komunikace s jinými stroji; paradoxně za ni lze označit komunikaci s lidmi skrze programovací jazyk. Tím je matematika, do které jsou slova, která například zadává programátor, skrze proces, vyvíjející se již několik desetiletí, převáděny. Můžeme si tento proces představit tak, že programátor zadává pro něj smysluplná slova v libovolném programovacím jazyce, která dané rozhraní převádí do pro počítač jednodušších pojmů, se kterými se však programátorovi pracuje výrazně složitěji. Pokud je toto pochopení správné, znamená to, že vstupní jazyk stroje dokážeme vystopovat až do jeho vlastního jazyka (tedy až k symbolům, kterými manipuluje).

Je-li toto zjednodušení pravdivé, pak samotný symbol 1 nebo 0 význam nenesí; skrze matematické postupy však tyto symboly zůstávají v kontextu a ve vzájemných vztazích. Pakliže význam samotného symbolu neexistuje, něco takového tvrdit o síti významů vytvořené ze vztahů mezi symboly je daleko problematičtější.

6.3 JAK NAUČIT STROJ JAZYKU?

Pro stroj je tedy klíčové naučit se jazyku; je mu bližší než jakýkoli jiný aspekt mysli. Navíc nám postačuje pro určitou verzi mysli. Pokud by existovala možnost, jak tuto dovednost získat, jaká je možná cesta k ní? Obdobně, jako jsme v kapitole Úvod do jazyka²⁴ naučili psa povelu, obdobně, jako se jazyku může naučit ČP, obdobně, jako se jazyku učil Augustin – názorným ukázáním toho, co daný symbol znamená, ve světě. Stroj pro něco takového potřebuje jednak senzory (i vnitřní; jejich konstrukce na současné technologické úrovni nepředstavuje principiální problém) a jednak substrát²⁵, ve kterém budou tyto informace poskládané do významů moci zakořenit. Skutečnost, že jsme se strojem schopni komunikovat (buť skrze povelu) minimálně směřuje k primitivnímu jazyku. Tento jazyk nemusí být zdaleka tak komplexní jako ten náš, postačí určitý způsob práce s pojmy. Vzpomeňme, že se mysl vyvíjí společně s jazykem a že jeden je závislý na druhém.

Jakmile stroj získá základy, může z nich vybudovávat daleko složitější síť. Pochopí-li tedy pár pojmů, bude skrze souvislosti schopen pochopit i pojmy v primitivním jazyce nevysvětlitelné. A stroj s pamětí může mít skoro všechny souvislosti světa. Pochopení není v

²⁴ 5.1

²⁵ Substrát = prostředí, program, přednastavení, příhodný pro naučení se jazyka. V podobném stavu se nachází mozek malého dítěte, s věkem však jeho schopnost naučit se novému jazyku klesá.

tomto kontextu složité, ale je vzhledem k subjektivitě mysli a k nespolehlivosti svědectví ohledně ní neověřitelné, musíme v tuto skutečnost prostě ‚doufat‘, což je myšlenka, se kterou se těžce vyrovnává.

Ano, tyto informace sice umíme pospojovat, ale pořád by to byla jen skupina jedniček a nul, dokonce náhodná struktura na zdi. Buďto k tomuto problému můžeme přistoupit tak, že si této struktury časem stroj všimne a začne ji používat, nebo že i nehledě na tuto strukturu nějaké pochopení má, viz schopnost plnit příkazy. Druhým přístupem je nemožnost stroj takto redukovat; ačkoli jedničky a nuly jsou pravdivým popisem jeho mikrostruktury, nelze skrze ně vysvětlit všechny jevy makrostruktury.

Tento strukturální problém se opakuje na mnoha jiných případech ve světě. U evoluce jsme již zmiňovali základní stavební látku pro živou hmotu, organické sloučeniny, z nichž nakonec vznikl život tak komplexní, jako jsou tvorové dnes žijící na zemi. Dlouhá léta dokonce panoval mezi chemiky názor, že mezi organickou a anorganickou hmotou existuje principiální rozdíl a je nutno ke každé přistupovat zvlášť a s jinými pravidly. Tato vitalistická teorie byla prolomena při tvorbě močoviny z anorganické sloučeniny amoniaku v devatenáctém století.²⁶ Jako třetí příklad uveďme dvojici neurony – mozek. Aniž by bylo nutno přijmout teorii identity mozku a mysli, korelace a propojení mezi mozkovou hmotou a myslí nepochybně existují.

Stejně tak můžeme použít daleko obecnější příklad z fyziky, kde se vyskytuje bolestivá absence jednoznačné teorie spojující popis stavebních částic a světa, ve kterém žijeme, a který jsme schopni vnímat. Vzhledem k tomuto univerzálnímu rozporu mezi mikro a makro strukturou neznamenaají problémy zmiňované v první části této kapitoly slepou uličku pro mysl strojů. Jak se od tohoto liší jedničky a nuly? Skutečnost, že nerozumíme propojení mezi těmito dvojicemi nikdo nezpochybňuje, nesmíme však v jejím stínu principiálně vyloučit vznik mysli stroje na jejím základě.

²⁶ Syntéza močoviny z kyanatanu amonného provedená německým chemikem Friedrichem Wöhlerem (první polovina 19. století) představuje pouze jednu z více reakcí obdobného typu.

7 ZÁVĚR

Na těchto stránkách pokusila vymezit to, co považuji za mysl stroje a co mi jako mysl stroje postačuje. Klíčovou tezi představuje rozdělení mysli na živé a jazykové. Nelze vytvořit stejnou mysl pro zvíře a pro stroj, nelze tyto mysli volně prohazovat a očekávat od nich to samé. Skutečnost, že je tak složité najít mysl stroje, spojuji s tím, že hledáme jednu mysl. Základním sjednocujícím prvkem pro člověka, zvíře a jejich společnou mysl je život (evoluce jakožto hnací síla života) a s ním spojené cítění, chování, emoce či hodnoty. Tímto pojídlem se pro člověka a stroj zdá být jazyk, platforma umožňující myslícím entitám tvorbu a sdílení abstraktního světa.

Vstoupili jsme do říše filosofie mysli pomocí vztahu mysli a zvířete. Ten jsme popsali pomocí instinktivního přístupu (IP) – ukázalo se, že mysl je nejen intimní, ale také intuitivní téma. Tento přístup umožňuje intuitivní vymezení pojmů a novou perspektivu, jako je tomu v případě mysli či UI. Zároveň vymezuje vztah k porozumění. Protože nevíme, co to mysl je, tak ji můžeme přisoudit ve chvíli, kdy se v tvorovi vyskytuje něco nevysvětlitelného. Proto tak můžeme činit u zvířat, ale u strojů ne. Toto je jeden z okamžiků, kdy IP selhává v případě stroje a důvod, proč se na něj nemůžeme spoléhat.

Poté všechny tři pojmy stavím do rovnice. Právě skrze ni vznikají tato dvě hlavní témata, jimž se věnuje zbytek práce. Vystává také další problém IP, a to sklon hledat všude naši mysl. V kontrastu k němu nabízím teorii, že mysl se skládá z jednotlivých aspektů a že jako mysl postačuje i kombinace některých z nich.

Nejprve se zaměřuji na vztah mysli a aspektu života. Evoluci zde ukazují jako princip, který vede živé tvory k vývoji inteligentního chování, a tedy postupně ke vzniku mysli. Jelikož stroj tento tlak nepocítuje, nezíská stejný typ mentálních schopností. Vzniká tedy nutnost najít obdobný model pro mysl, která se na životě nezakládá. Je tedy třeba hlouběji rozebrat hypotézu, že se aspekt lidské mysli a celá mysl stroje zakládají na jazyce.

Takto znovu a lépe identifikuji rozdíl mezi zvířetem a člověkem, přesněji než jen jako fascinující ale nedostačující slovo „rozum“. Společná funkce jazyka pro člověka a pro stroj je práce s abstraktním světem, řešení abstraktních příkladů. Druhou společnou funkcí představuje komunikace.

V rámci popisu odbornější teorie vztahu stroje a jazyka stručně rozebírám KTM a argument čínského pokoje. KTM (daný model stroje) může být díky aspektům mysli zároveň platná pro stroj a umožní mu mysl a zároveň pouze částečně platná pro člověka, jak je ukázáno

na příkladu umění. Tato teze umožňuje tvrzení, že lidská kreativita je nedeterministická. Hlavním tématem argumentu čínského pokoje je schopnost stroje porozumět jazyku. Vracím se zde znovu k IP; v oboru, jako je tento, jej nelze uplatňovat tak stejně, jako jsme činili v první části. Jinou možnou odpovědí je teze, že jazyku může stroj porozumět skrze interakci se světem. Zároveň ve stroji existuje určité porozumění, neboť plní instrukce.

Nakonec představuji primitivní jazyk a teorii, jak se naučit primitivnímu jazyku. Jazyk, který máme v současnosti, je komplikovanější a umožňuje nám vytvořit nás vlastní abstraktní svět pojmů. Obhajuji, že jazyk je pro člověka vedle života a evoluce naprosto zásadní. Žijeme skrze jazyk, i nejazykové vjemy vnímáme jazykovými pojmy, jazyk nás odlišuje od jiných bytostí.

Nezbývá tedy než se vrátit k myslí stroje. Jakým způsobem přisoudit stroji mysl založenou na jazyce a proč je těžké mysl stroje najít? Docházím ke dvěma bodům. Prvním je, že tato mysl má skutečně jinou formu, než ta lidská či ta zvířecí. Pokud někde existuje mysl stroje, bude mít jinou formu než ta naše. Tímto neredukuji lidskou mysl ani mysl stroje, pouze proti instinktivnímu přístupu odděluji ty aspekty, které mysl stroje nemusí obsahovat.

Druhým problémem je struktura stroje. Zde se vrací problém IP, že této struktúře teoreticky rozumíme. Vedle toho však vystupuje Searle, jehož argument proti pochopení jazyka je odlišný, prezentuje nemožnost přeskočit od formální struktury k významům. Tento se objevil už v KTM, prostě jsou to jedničky a nuly. V tomto považuji za velmi důležitou myšlenku rozporu mezi mikro a makro strukturou. Existuje mezi nimi obecná nespojitost, která se projevuje také v mnoha jiných disciplínách. Proto, nehledě na jakékoli pocity, propojení mezi nimi v současnosti nedokážeme patřičně vysvětlit, a tudíž jej nemůžeme považovat za důvod k neexistenci myslí stroje. Právě naopak, I přes existenci tohoto rozporu můžeme s myslí stroje pracovat.

Na úplném konci se na okamžik vraťme ještě jednou k úvodní otázce: Proč je „může stroj myslet“ (podle Alana Turinga) tak vágní? Přítomnost myslí nemůžeme zatím ověřit – můžeme vidět vnější projevy, stavy v mozku, ale nikdy nevidíme myšlení samotné; můžeme o něm tedy přemýšlet, k takovýmto vnitřním důkazům však máme v současnosti daleko.

Kvůli tomu v tento okamžik nedokážeme říct, jak přesně to s myslí stroje a s umělou inteligencí dopadne, nemůžeme však vyloučit, že skrze postupný společný vývoj jazyka, myslí a učení časem nezíská mnohem více. Podle mého závěru, za myslící entitu jej lze považovat daleko dříve, než k takovému stavu dojde. Pro mou mysl stroje postačuje porozumění jazyku.

POUŽITÁ LITERATURA

Dennett, Daniel C.: *Druhy mysli* (Praha, Academia, 2004) přel. J. Peregrin,

Ingram, Jay: *Divadlo mysli* (Praha, dubbuk, 2010) přel. H. Čížková

Scruton, Roger: *Průvodce inteligentního člověka filosofií* (Barrister & Principal, 2003) přel. J. Ogrocký,

Churchland, P. a Churchland, P., 1990, 'Could a machine think?', *Scientific American*, 262(1): 32–37

Internetové zdroje: Stanford Encyclopedia of Philosophy

<https://plato.stanford.edu/entries/chinese-room/>

<https://plato.stanford.edu/entries/computational-mind/>

Chalmers, David J.: Introduction: Taking Consciousness Seriously, in *The Conscious Mind*, dostupné na: <http://consc.net/books/tcm/intro.html>

Ruisel, Imrich: *Základy psychologie inteligence* (Portál, 2000) (kapitola 10)

Wittgenstein, Ludwig: *Filosofická zkoumání* (Praha, Filosofický ústav AV ČR, 1993) přel. J. Pechar (strany něco a něco)

The Economist (January 6th-12th): The next frontier (články: The next frontier, Technology quarterly: Brain-computer interfaces)

