

Univerzita Hradec Králové
Filozofická fakulta

BAKALÁŘSKÁ PRÁCE

2019

Jakub Vlček

Univerzita Hradec Králové
Filozofická fakulta
Katedra filozofie a společenských věd

Etické aspekty umělé inteligence
Bakalářská práce

Autor: Jakub Vlk

Studijní program: B 6101, Filozofie

Studijní obor: Filozofie a společenské vědy

Forma studia: prezenční

Vedoucí práce: prof. RNDr. Jaroslav Peregrin, CSc.

Hradec Králové, 2019

Prohlášení

Prohlašuji, že jsem tuto bakalářskou práci vypracoval (pod vedením vedoucího bakalářské práce) samostatně a uvedl jsem všechny použité prameny a literaturu.

V Hradci Králové dne 28.11. 2019

.....
Jakub Vlk

Poděkování

Rád bych poděkoval prof. RNDr. Jaroslavu Peregrinovi, CSc. za odborné vedení, vstřícnost, cenné poznámky a rady při zpracování této práce.

Jakub Vlk

Bibliografický záznam

VLK, Jakub. *Etické aspekty umělé inteligence*. Hradec Králové: Filozofická fakulta, Univerzity Hradec Králové, 2019. 53 s. Bakalářská práce

Anotace

Cílem práce je zpracovat přehled etických problémů, před které nás může v dohledné době postavit rozvoj umělé inteligence a robotiky, a pokusit se zhodnotit jejich závažnost a dosah. Zvážena by měla být i obecná otázka pozitivních a negativních aspektů toho, co může umělá inteligence lidstvu přinést. Za tímto účelem je nutné provést analýzu současného stavu věcí a výhledů do blízké budoucnosti, tedy toho, kam nás umělá inteligence může v nejbližší době vést a jaké etické problémy s tím mohou být spojeny.

Klíčová slova

Etická otázka, umělá inteligence, pokrok, technologie, strojové myšlení, hrozby, příležitosti

Annotation

The aim of this work is to summarize ethical problems we can face due to development of artificial intelligence and robotics. It also tries to evaluate seriousness and reach of these problems. We should even consider general question of positive and negative features of what the artificial intelligence is able to bring to the mankind. It is necessary to present a summary of current situation and future prospects. Important questions we should ask are where the artificial intelligence is able to lead us and what ethical problems it can bring.

Keywords

Ethical question, artificial intelligence, progress, technology, machine thinking, threats, opportunities

Obsah

Úvod.....	7
1. Co je to život?	8
1.1 Technologie uhání mílovými kroky	10
1.2 Autonomní vozidla.....	10
1.3 Technologičtí rivalové	12
2. Umělý mozek?	14
2.1 Neuronové sítě	14
2.2 Fuzzy logika	16
2.3 Umělý život.....	19
3. Ohlédnutí do blízké minulosti aneb První milník Alana Turinga.....	20
3.1 Zákony robotiky – definice umělé inteligence.....	21
3.2 Asimovovy zákony a Moravcův paradox	22
3.3 Moorův zákon	24
4. Následky hojnosti a rozvoje.....	25
4.1 Stroj jako pracovní síla.....	26
4.2 Experimenty s androidy	27
4.2.1 Mohou být stroje morální?.....	29
5. Etická otázka.....	32
5.1 Praktické využití.....	33
5.1.1 Algoritmické předsudky	35
5.1.2 Netransparentnost	36
5.1.3 Soukromí.....	37

5.1.4	Právo	38
5.1.5	Etická problematika v různých odvětvích	38
6.	Technologie a predikce budoucna	43
6.1	Názory	44
6.1.1	Optimisté.....	44
6.1.2	Pesimisté	45
6.1.3	Moje stanovisko	46
	Závěr	48
	Použitá literatura:	49

Úvod

Masivní rozvoj moderních technologií ve světě vyvolává v každém z nás přinejmenším určitou míru pozornosti. Myslím si, že by toto téma související s naší budoucností nemělo být nikomu lhostejné. Je to možná aktuálnější záležitost, než si myslíme. Otázkou zůstává, jestli je to správný směr pro lidstvo a jestli si svými kroky nepodepisujeme pomyslnou úmluvu s „děblem“, kterou může být právě naše budoucnost poznamenána. Jestli si vědomě nevytváříme mlžný opar, který nám zabraňuje dohlédnout dál. Neměli bychom být opatrnější v tom, co vytváříme? Jsme si opravdu vědomi toho, jaké to pro nás samé a obecně pro lidstvo může mít následky?

Ve své práci se budu zabývat vývojem prvních strojů a factory, které vedly k rychlému rozvoji technologií. Dále zhodnotím nynější situaci se zřetelem k blízké budoucnosti a otevřu otázku, jestli je šance, že by lidstvo mohl předstihnout technický pokrok, který se neustále rozvíjí. Myslím si, že tato otázka je aktuální a úzce souvisí s momentálním světem kolem nás. Jsou všechny výhody a pohodlí, které nám technologické „zázraky“ přinášejí, bezpodmínečně dobré? Nepovede to k etickému konfliktu, nerovnováznému trhu či nadbytku nebo k ještě horším scénářům? Dále bych chtěl zvážit a porovnat názory všech stran, ať už pozitivní, nebo negativní a na základě toho se pokusím zformulovat vlastní stanovisko k nastíněné problematice.

Proto si kladu otázku, kam nás může rozvoj moderních technologií, potažmo umělé inteligence přivést? Bude to k našemu prospěchu, či nikoliv? Je možné, že by stroje začaly myslet, a k jakým problémům by to mohlo vést? Jaké jsou etické aspekty umělé inteligence?

1. Co je to život?

Život, pro někoho samozřejmost, pro druhého nesmírná záhada. Můžeme o stroji říci, že žije, když dělá věci jako živí tvorové? Pro uvedení do problematiky své práce bych se chtěl pokusit popsat, co je za život považováno a co není. Ačkoliv jistě existuje nespočet způsobů, jak tuto otázku zodpovědět, zvolil jsem definici, kterou uvádí Kevin Warwick ve své knize „*Úsvit robotů, soumrak lidstva*“. Podle jeho názoru existují znaky, které jsou všem živým tvorům společné, jako je růst, dýchání, reprodukce, výživa a v neposlední řadě vyprazdňování. Nesmíme opomenout pohyb a dráždivost ve smyslu reakce na vnější podněty.¹ Čeho jsem si ale hned povšiml a autor na to odkazuje též, je fakt, že mezi hlavními znaky není inteligence, ačkoliv známe mnoho živočichů na bakteriální úrovni, o kterých lze říci, že mají vyvinutou důmyslnou chemickou komunikaci. Tuto feromonovou či chemickou komunikaci bychom mohli považovat za jakýsi „inteligentní systém“. Pokračuje tím, že tyto vlastnosti mají všechny živé organismy na naší planetě. Nás by ale hned nemusely napadnout stroje, naše výtvary, v otázce života. Když bychom na problém nahlédli z jiného úhlu pohledu, teoreticky bychom se mohli ptát, co by mimozemšťané považovali za zdroj života? Jak autor píše, více než 70% naší země pokrývá voda. Je pravděpodobné, že by setrvali na Zemi právě kvůli zdroji vody a na co by narazili, by byly bezpochyby lodě. Jak autor sám říká, pro člověka je samozřejmé, že lodě nejsou formy života, ale co když by se našim návštěvníkům tak mohly jevit? Podle výše zmíněných kritérií je hned viditelný pohyb, výživa v podobě přikládání do kotle, kouř jako odpad, dráždivost – při potencionálním kontaktu s lodí by zaznamenali i reakci při nárazu. Taktéž spalování paliva by mohlo být považováno za dýchání. Bez toho, že by zaznamenali růst nebo reprodukci, by se mohli mimozemšťané domnívat, že na

¹ WARWICK, Kevin. *Úsvit robotů - soumrak lidstva*. Praha: Vesmír, 1999. Medusa (Vesmír). ISBN 80-85977-16-8. Str. 25.

planetě Zemi, jak autor popisuje, žádný život není. Nanejvýš by se mohli domnívat, že lodě jsou pouze jakýmsi zárodkem života.²

Kdybychom ale vynechali v otázce života možnost setkání se s mimozemšťany, stačí se ohlédnout zpět do naší historie. Zmínky o prvních „robotech“ sahají, pro mě vcelku nečekaně, hluboko do naší minulosti. Například samotná pověst o Golemovi je příkladem toho, že onen Golem mohl být v podstatě robot, který měl chránit židovské město. To, co ho přivádělo k „životu“, byl takzvaný šém vkládaný rabínem, jeho stvořitelem, Golemovi pod jazyk. Jak popisuje autor, šém mu dával obrovskou sílu, přímo nadlidskou. Bez něj to byl jen obyčejný hliněný panák. Pověst praví, že se jednou Golem vymkl kontrole a naopak ohrožoval město do té doby, dokud rabín šém nevyjmul. Je tedy možné, že se dal díky rabínovi zapínat nebo vypínat. To, co vyplývá z této a podobných legend, je to, že je pro nás velmi obtížné představit si bytost nebo věc inteligentnější než my sami.³

Člověk není nejsilnější ani nerychlejší bytostí, ale to, co nám poskytuje převahu nad ostatními tvory, je právě naše inteligence a věci s ní spojené, jako je abstraktní myšlení nebo vědomí.⁴

² WARWICK, Kevin. *Úsvit robotů - soumrak lidstva*. Praha: Vesmír, 1999. Medusa (Vesmír). ISBN 80-85977-16-8. Str. 26

³ Tamtéž, str. 29

⁴ Tamtéž, str. 29

1.1 Technologie uhání mílovými kroky

V knize *Druhý věk strojů* od autora Erika Brynjolfssona jsem objevil citát od Arthura C. Clarka, který zní: „*Žádnou dostatečně pokročilou technologii nelze odlišit od magie*“.⁵ Vidím v tom souvislost s dobou, ve které byl Golem sestrojen, a nedivím se, že Golema tehdejší lidé spojovali s magií vzhledem k tomu, že neměli možnost jiného vysvětlení.

Nejrůznější technologie naší současnosti leckdy přesahují možnosti chápání běžného člověka – laika, který může být mnoha artefakty a technologiemi fascinován, stejně jako zděšen. Dnes se zdá být možné to, co našim předchůdcům mohlo připadat jako sci-fi.

Výrazný zlom v dějinách techniky byl způsoben průmyslovou revolucí, jejíž dopad můžeme vnímat v oblasti každodenního lidského života, ať už máme na mysli pohodlí, zkrácení vzdálenosti, zlepšení hygieny, pozitivní demografický vývoj, ale zároveň znamenala i vzdálení se přírodnímu způsobu života (téma kupříkladu fenomenologů či existencionalistů nebo i marxistů), stres, nové tzv. civilizační choroby, radikální změny v chování ve sféře soukromé i profesní. Nástup moderny v podstatě lze chápat jako tzv. první věk strojů, jemuž dominovaly objevy v oboru strojírenství, chemie apod.⁶

1.2 Autonomní vozidla

Nyní v roce 2019 pro nás již nejsou takovým překvapením, už v roce 2012, jak výše uvedený autor Erik Brynjolfsson popisuje, bylo možné si díky společnosti Google vyzkoušet jízdu ve zcela autonomním voze projektu Chauffeur. Jednalo se o auto, které po stisknutí tlačítka přešlo do plně automatizovaného řízení. V souvislosti s tímto jsem si ihned vybavil funkci

⁵ BRYNJOLFSSON, Erik. *Druhý věk strojů: práce, pokrok a prosperita v éře špičkových technologií*. Přeložil Filip DRLÍK. V Brně: Jan Melvil Publishing, 2015. Pod povrchem. ISBN 978-80-87270-71-4. Str. 22

⁶ Tamtéž, str. 15

autopilota v letadlech, s tím rozdílem, že ve vzduchu není takový provoz. Také bych se nejspíše zarazil, možná vylekal, kdybych viděl auto bez řidiče, takže předpokládám, že to ještě nějaký čas bude trvat, než si i okolí zvykne na takovéto „podivnosti“. Kromě toho se oba autoři shodli na tom, že jízda, ačkoliv se mohla zdát jako adrenalinový zážitek, byla bez řidiče vcelku nudná. Vše zaznamenávaly snímače po celém autě a jako by předpovídaly chování ostatních aut a i dopravní situace.⁷ Být skeptický je jistě namístě, ale překvapivý je fakt, že společnost Google tato auta již nějakou dobu testuje, ta najela tisíce kilometrů bez lidského zásahu, při tom došlo jen ke dvěma chybám, za obě mohl v dané situaci člověk.⁸ Dle mého názoru je rozvoj autonomních vozidel jedním z mnoha klíčových technologických „zlepšováků“, ale zároveň také potencionální hrozbou. Tato vozidla budou šetrnější k životnímu prostředí, protože budou s největší pravděpodobností využívat elektromotory. Ostatně Tesla Motors je dnes už běžnou realitou. Autonomní vozidla se budou jistě dále vyvíjet a jejich prosazení bude klíčové. Otázkou je, jak dokážeme posoudit, jestli jsou opravdu bezpečnější než vozidla řízená lidmi. Když bychom vzali v potaz, že by byla nadále bezpečnější, jaká kritéria bychom měli požadovat? V nejzazším případě bychom se mohli tázat, pokud by byla autonomní vozidla bezpečnější než ta s lidským řidičem, jestli bychom neměli zcela zakázat automobily, které funkci autonomního řízení nedisponují. Alternativou by též bylo vytvoření speciálních koridorů nebo jízdnic pruhů pro určité typy vozidel. Tyto otázky jsou zcela na místě a odpovědi na ně ovlivní nespočet lidských životů v následujících letech.⁹

⁷ BRYNJOLFSSON, Erik. *Druhý věk strojů: práce, pokrok a prosperita v éře špičkových technologií*. Přeložil Filip DRLÍK. V Brně: Jan Melvil Publishing, 2015. Pod povrchem. ISBN 978-80-87270-71-4. Str. 22 – 23

⁸ Tamtéž, str. 27

⁹ The key ethical question for self-driving cars: are they safe? - Vox. *Vox - Understand the News* [online]. Copyright © 2019 [cit. 18.09.2019]. Dostupné z: <https://www.vox.com/future-perfect/2018/11/9/18072678/self-driving-cars-philosophy-safety-trolley-problem-mit>

1.3 Technologičtí rivalové

Dalším příkladem pokroku technologií je známá americká soutěž „Jeopardy!“, jejíž českou obdobou byla například soutěž „Riskuj!“. Tato televizní soutěž má již od roku 1964 dlouholetou tradici a objevuje se na televizních obrazovkách dodnes. Důvodem toho, proč je tato soutěž stále tak atraktivní, je, že je snadno pochopitelná, ale je velmi těžké ji hrát dobře. Jedná se o otázky z širokého spektra témat, na které daný soutěžící musí odpovědět jako první a správně, aby nasbíral co nejvíce peněz a zajistil si výhru. Za každou špatně zodpovězenou otázku se mu část nasbíraných peněz odečte, takže musí disponovat určitou přesností. Všestrannost hráčů vystihuje i to, že dostávají nápovědy v hádankách. Proto je pro výhru nutné, aby hráči s nejvyšší přesností porovnávali vzory a ovládali komplexní komunikaci okamžitě a stále dokola.¹⁰

Hra „Jeopardy!“ má historicky dva nejúspěšnější hráče, kteří disponují z velké části takovými schopnostmi. Prvním a historicky nejlepším hráčem je Ken Jennings, který v roce 2004 vyhrál po sedmdesáté čtvrté řadě a odnesl si více než 3 miliony dolarů. Druhým je Brad Rutter, který Jenningse porazil v obdobném pořadu a odnesl si rovněž horentní sumu.

Na základě toho byl týmem Charlese Lickela, vedoucího výzkumu umělé inteligence v IBM, sestroyen superpočítač Watson, který měl porazit lidi v jejich vlastní hře. Na začátku se testy ani zdaleka nepřibližovaly cíli. Po delší době se však ukázalo, že superpočítač se rychle učí, a v roce 2010 se dokázal se správným nastavením přihlásit o otázku v 70% jako první a následně v 85% dokázal odpovědět správně. Nicméně se ale stále nemohl rovnat komplexní lidské bytosti. Na počátku roku 2011 Watson oba soupeře překvapivě porazil. Byl rychlý a odpovídal převážně správně a velice komplexně. Domnívám se, že

¹⁰ BRYNJOLFSSON, Erik. *Druhý věk strojů: práce, pokrok a prosperita v éře špičkových technologií*. Přeložil Filip DRLÍK. V Brně: Jan Melvil Publishing, 2015. Pod povrchem. ISBN 978-80-87270-71-4. Str. 32 – 33

to bylo v důsledku strojového učení, které inteligentní stroje vykazují. Jak autoři zmiňují, Watson nasbíral více než trojnásobek bodů oproti svým lidským soupeřům.¹¹ Za velmi výstižnou považuji tezi, kterou zformuloval Jennings a kterou uvádějí ve své knize výše zmiňovaní autoři: „jsme první pracující ve vědomostním průmyslu, kteří kvůli této nové generaci ‚myslících‘ strojů přišli o práci. ‚Soutěžící ve vědomostním kvízu‘ je možná první práce, kterou Watson nahradil, ale jsem si jistý, že nebude jediná“¹²

¹¹ BRYNJOLFSSON, Erik. *Druhý věk strojů: práce, pokrok a prosperita v éře špičkových technologií*. Přeložil Filip DRLÍK. V Brně: Jan Melvil Publishing, 2015. Pod povrchem. ISBN 978-80-87270-71-4. Str. 34

¹² Tamtéž, str. 35

2. Umělý mozek?

Dalo by se tedy o stroji, který provádí takovéto úkony a člověka při tom „nechává za sebou“, říci, že nějakým způsobem myslí, nebo je to jen věc vysokého stupně naprogramování? Mnohým odborníkům v dnešním světě hrozí riziko, že se stanou oběťmi svých oborů. Kde je ta hranice, když se oprostíme od toho, že myslet může jen biologická bytost, tedy člověk? K nalezení, nebo aspoň přiblížení se odpovědi je zapotřebí znát určitou funkci našeho mozku a porozumět našim neuronům.

Prameny, které se zmiňují o něčem takovém, jako je mozek, sahají až do starověkého Egypta. Přesněji se jedná o jakýsi papyrový dokument, kde bylo toto slovo poprvé zmíněno. Výzkum mozku ale nejintenzivněji začíná až v tomto století, kdy o něm víme na fyziologické úrovni poměrně dost. Víme, z čeho se skládá, které části řídí jaké funkce, ale co spíše nevíme, je například to, jak si vlastně uvědomíme sami sebe, tedy nevíme, jak přesně funguje naše psychologická část.¹³

2.1 Neuronové síť

Pro zjišťování funkcí mozku existují dnes přístroje, jako je magnetická rezonance, nebo tomografie. Tyto metody slouží k mapování jeho struktury, a to velmi podrobně. Díky tomu máme velmi detailně popsany mozek, ale žádná nám neumožňuje zjistit a podchytit aktivity, které mají vliv na následné činnosti jedince. Pojem inteligence zde tedy nehraje – zatím – žádnou roli. Již Aristoteles prohlásil, že „*celek je více nežli pouhý souhrn jeho částí*“, od čehož se odvíjí to, že něco jako inteligence je komplexní proces, který zahrnuje racionální i emotivní funkce.¹⁴ Podle Radkina Honzáka se například emoční stránka mozku

¹³ ZELINKA, Ivan. *Umělá inteligence: hrozba nebo naděje?*. Praha: BEN - technická literatura, 2003. ISBN 80-7300-068-7. Str. 30

¹⁴ ZELINKA, Ivan. *Umělá inteligence: hrozba nebo naděje?*. Praha: BEN - technická literatura, 2003. ISBN 80-7300-068-7. Str. 31

vyvinula již v době dinosaurů. Strach a vztek byly prvotními emočními stavy. Evolučním vývojem mozek dospěl do stádia, jak ho známe dnes.¹⁵

První zmínka o neuronech pochází z 20. století, kdy W. S. Culloch se svým studentem W. Pittsem vypracoval model neuronu, který se používá dodnes. Dal by se popsat jako „černá skříňka“ s několika vstupy a se stejným počtem výstupů. Uvnitř se nachází algoritmus (přenosová funkce), který určuje, jak bude neuron reagovat na podněty a co z nich vyvodí. Bez ohledu na stáří této ideje se tento model, i když obohacený, používá dodnes. Na tomto základě roku 1958 F. Rosenblatt vytvořil první funkční tzv. „perceptronovou síť“, která měla velký nedostatek, který vedl k jejímu zapomení téměř na 20 let. Síť byla schopna řešit problémy pouze na lineární úrovni. V 80. letech došlo k renesanci problému a D. Rumelhart, G. Hinton a R. Williams v práci „*Learning Internal Representation by Error Propagation*“ popsali síť, která dokáže problémy řešit mezi více vrstvami – v různých úrovních. Stačilo přitom přidat pouze jednu vrstvu navíc.¹⁶

Mozek se skládá ze vzájemně prolínajících se funkcí. Neurony, jako základní stavební prvek, vytvářejí neuronové sítě, ve kterých vznikají vzruchy, a to dává našemu mozku možnost masivní výpočetní schopnosti. 25 miliard neuronů dokáže vyprodukovat až 10 triliónů vzájemných propojení, čili už jen to, že by se takové množství dalo mechanicky napodobit, zní jako pohádka, protože u výkonnosti mozku záleží na kvalitě a počtu spojů. Největším háčkem v detailním napodobení mozku je jeho funkční stránka. Ačkoliv lidský mozek poznáváme čím dál více a topologie neuronových sítí nám jsou více známé, stále zcela neumíme neurony kvantifikovat a rozlišit jednotlivé spoje a vzruchy.¹⁷ Jen pro představu, pokud bychom chtěli napodobit činnosti a propojení mozku

¹⁵ SaVIO: Emoce a jak s nimi pracovat (MUDr. Radkin Honzák) - YouTube. YouTube [online]. Dostupné z: <https://www.youtube.com/watch?v=aedhxvooyPg&t=378s>

¹⁶ ZELINKA, Ivan. *Umělá inteligence: hrozba nebo naděje?*. Praha: BEN - technická literatura, 2003. ISBN 80-7300-068-7. Str. 38

¹⁷ Tamtéž, str. 32 – 33

v lidském měřítku, čili jeden člověk by se rovnal jednomu neuronu, jak popisuje ve své knize Ivan Zelinka: „*tak by každý člověk na této planetě musel hovořit zároveň osmnácti telefony...navíc v každém z nich na jiné téma.*“¹⁸

Činnosti neuronových sítí, jako základ myšlení a inteligence, může vypadat jako velice komplikovaná věc, avšak je třeba si uvědomit, že právě tyto procesy využívá každý z nás automaticky, každý den a komplexně. Zde je vidět srovnání toho, jak banální se to každému z nás může jevit, naproti tomu jak nesmírně obtížné je to napodobit, či vytvořit. Neurony a jejich sítě jsou pořád jen algoritmy, ale jejich jedinečnost tkví v učení se. Jejich samotné využití bude tedy spíše ve formě klasifikátorů, tedy typů algoritmů, které budou řídit stroje pro různá odvětví lidských činností.¹⁹

2.2 Fuzzy logika

Ruku v ruce s rozvojem umělých neuronů a neuronových sítí bezpochyby kráčí i tzv. fuzzy logika. Další z mnoha, podle mého názoru velmi významných a nutných částek toho, čemu bychom později mohli říkat umělý život či inteligence. Klasická logika jako taková pracuje s dvěma protipóly, které přísluší do dvou skupin. (Pravda – nepravda, ano či ne, černá – bílá).²⁰ Vývoj logiky postupně dospíval k tomu, že nestačilo pouze rozhodnout mezi 1 nebo 0, ale bylo nutné vzít v úvahu i takové hodnoty jako třeba $\frac{1}{2}$. Z antického pohledu to byl nesmysl. Elektrotechnik L. Zadeh přišel s teorií fuzzy množin, která tento problém vysvětlovala. Laicky řečeno, mezi 1 a 0 se vyskytuje více hodnotových možností, kam určitý prvek, věc zařadit, tedy daný prvek, věc patří do dané množiny více, středně či méně. Fuzzy logika se dnes využívá např. v procesorech v pračkách, ale i v komplikovaných systémech, jako jsou spalovací

¹⁸ ZELINKA, Ivan. *Umělá inteligence: hrozba nebo naděje?*. Praha: BEN - technická literatura, 2003. ISBN 80-7300-068-7. Str. 41

¹⁹ ZELINKA, Ivan. *Umělá inteligence: hrozba nebo naděje?*. Praha: BEN - technická literatura, 2003. ISBN 80-7300-068-7. Str. 54

²⁰ Tamtéž, str. 55

motory, nebo dokonce ve Vesmíru, kde byl použit fuzzy regulátor ramene raketoplánu, který držel Hubbelův teleskop. Stále nachází širší využití.²¹

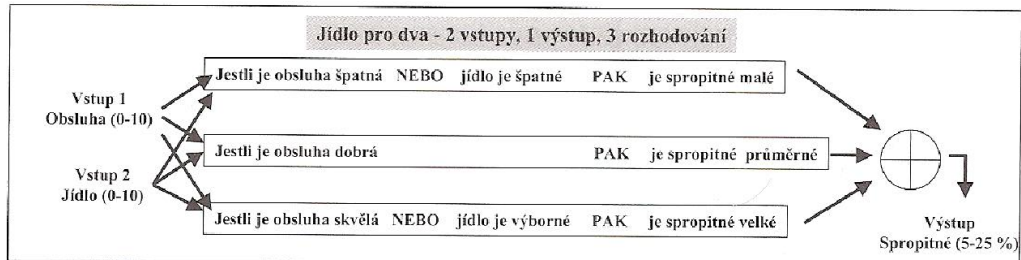
Aby mohla fuzzy logika platit, je zapotřebí, aby nasbírané hodnoty byly převedeny na množiny, poté se provede fuzzyfikace – čímž se každé hodnotě množiny přiřadí stupeň příslušnosti (dejme tomu 1 – 10). Proces pokračuje řešením operací a tzv. defuzzyfikací, která vyhodnotí, do jaké míry daný prvek náleží do dané množiny. Ačkoliv se tento proces může zdát komplikovaný, náš mozek ho využívá častěji, než bychom čekali. Ilustrujme tento proces na příkladu: Představme si, že nakupujeme zboží v obchodě. Běžně zboží podvědomě hodnotíme podle toho, jestli se nám více, či méně líbí nebo chutná a z toho vyvozujeme, jestli si ho koupíme. Právě na toto hodnocení se dá nahlížet jako na fuzzy, protože v našem mozku, stejně jako u počítače, probíhají operace, z nichž vyplývá závěr, zda si produkt koupíme, či ne. Z toho tedy lze vyvodit, že fuzzy logika dokáže jistým, ačkoliv datově velmi objemným způsobem, napodobovat lidské chování či rozhodování. Představme si funkčního robota vybaveného fuzzy logickým programem, je zřejmé, že nákup v obchodě by zvládl a to není zdaleka všechno.²²

Veškeré lidské aktivity vycházejí z našeho uvažování, to se však dá logicky rozložit na operace, které lze pomocí právě fuzzy logiky vysvětlit či počítačově aplikovat. Jak to funguje? „Jestliže prší, vezmu si deštník.“ – je na první pohled syntakticky běžná věta. Po bližší analýze je zde zřejmé pravidlo (Jestliže – Tak), které používáme dnes a denně. Na tomto základě je pak stroj vybaven tzv. pravidlem If–Then.

²¹ Tamtéž, str. 56

²² ZELINKA, Ivan. *Umělá inteligence: hrozba nebo naděje?*. Praha: BEN - technická literatura, 2003. ISBN 80-7300-068-7. Str. 55 – 58

Ilustrace je vhodná opět na příkladu spropitného v restauraci. Jak nám tedy fuzzy logika může říci, v jaké výši má být ono spropitné za kvalitu jídla a služeb číšníka?²³ Viz obrázek 25.



Obr. 25 Pravidla If-Then pro ohodnocení výše spropitného

Jak tedy popisuje Ivan Zelinka ve své knize, fuzzy proces probíhá tím způsobem, že z daných vstupů a jejich hodnot (viz obrázek 25) se provede zmíněná fuzzyfikace a vybere se příslušná množina – fuzzy množiny „skvělá obsluha“ a „jídlo je výborné“ tvarují výstupní fuzzy množinu „velké spropitné“. Poté se jednotlivé fuzzy množiny sloučí a za pomoci poslední operace se vyvodí konečná hodnota.²⁴

Podle mého názoru je fuzzy logika obtížně pochopitelná záležitost. Je předvídatelné, že není bezchybná, ať už se jedná o množství nebo tvar množin, ze kterých má být závěr vyvozen. Snažil jsem se nastítnit, jak problematická záležitost to je oproti tomu, jak blízko od ní stojíme. Může to být jen další krok na cestě ke vzniku umělých životů?

²³ Tamtéž, str. 59

²⁴ ZELINKA, Ivan. *Umělá inteligence: hrozba nebo naděje?*. Praha: BEN - technická literatura, 2003. ISBN 80-7300-068-7. Str. 61

2.3 Umělý život

Seznámili jsme tedy se dvěma významnými aspekty „inteligence“ v robotice i na reálné úrovni. Avšak věda se snaží neúprosně směřovat kupředu i v oblasti počítačové – tedy „in silico“. Umělý život lze definovat do tří směrů:

- in hardware – již zmíněná robotika
- in software – simulace života v prostředí počítačů
- in wetware – jakýsi hybridní směr, ve kterém se replikují a vyvíjejí makromolekuly

V zásadě jde o to, že jsou to stále dělicí se a vyvíjející se makromolekuly. Je to jakási virtuální alchymie, pomocí které se vědci snaží vytvořit *homunkula*²⁵ v softwarovém světě. Jako u definic života najdeme i zde definice, co je to život umělý. Nejvěrohodnější z nich je ta, která předkládá tezi, že k životu je nutná určitá složitost. Zavádí tedy pojem „emergence“, což znamená „zjevit se“. Nutno podotknout, že to není myšleno ve smyslu spirituality. Všechny struktury života se vyznačují svou komplexitou, jejímž výsledkem je, že struktury mohou nějakým způsobem inteligentně reagovat nebo se reprodukovat. Život jako takový je spíše podmíněn způsobem, jakým dokáže reagovat na určité typy informací. „*Není důležité, zda je nositelem neuronová síť, nebo křemíkový procesor (...) Důležitá je komplexita informačního procesu a schopnost paměti*“²⁶

²⁵ Homunkulus = latinský pojem, který dostal v raném novověku nový význam v souvislosti se spekulacemi o možném vytvoření umělého člověka

²⁶ ZELINKA, Ivan. *Umělá inteligence: hrozba nebo naděje?*. Praha: BEN - technická literatura, 2003. ISBN 80-7300-068-7. Str. 102

3. Ohlédnutí do blízké minulosti aneb První milník Alana Turinga

V roce 1950 Alan Turing, logik a průkopník počítačů, přišel s něčím, čemu se říká „imitační hra“. Jeho článek „*Computing Machinery and Intelligence*“ se stal významný hlavně proto, že si zde položil zásadní otázku, zda mohou stroje myslet. Chtěl tedy stanovit obecné kritérium, díky kterému by tuto otázku bylo možné zodpovědět.²⁷

Toto kritérium je realizováno v podobě testu. Představme si místnost předělenou zástěnou s malým otvorem, kde na jedné straně vybraný člověk – porotce, komunikuje s entitou za zástěnou, o níž neví, zda se jedná o člověka, nebo počítač. Porotce předkládá druhé straně písemně otázky a ta na ně odpovídá. Na základě odpovědí má poté určit, zda se jedná o lidskou bytost, nebo počítač. Počítač by tedy prošel testem tehdy, kdyby jeho odpovědi byly natolik přesvědčivé, že by ho porotce – člověk od lidské bytosti nerozeznal.²⁸

Jaroslav Peregrin ve své knize uvádí, že velmi kritický ohlas vyvolal John Searle, který přišel s experimentem zvaným „čínský pokoj“. Searle říká, představme si člověka sedícího v pokoji, který má k dispozici spoustu příruček a slovníků v čínštině. Má za úkol odpovídat na otázky také v čínštině, díky všemožným poučkám a příručkám odpovědi sestaví. Je tedy zřejmé, že tento člověk je schopen projít Turingovým testem v čínštině, aniž by zde byla nějaká známka porozumění. Searle tímto dokládá, jak popisuje Peregrin, že na jedné straně máme jakousi manipulaci se symboly a na druhé porozumění a myšlení. Tedy v závěru dvě rozdílné věci. „*Počítač může podle zadaných pravidel operovat se symboly, nemůže však nikdy vědět, co tyto symboly znamenají.*“²⁹

²⁷ PEREGRIN, Jaroslav. *Kapitoly z analytické filosofie*. 2. vyd. Praha: Filosofia, 2014. ISBN 9788070074206. Str. 231 – 232

²⁸ Tamtéž, str. 232

²⁹ PEREGRIN, Jaroslav. *Kapitoly z analytické filosofie*. 2. vyd. Praha: Filosofia, 2014. ISBN 9788070074206. Str. 237

V té době bylo velmi těžké si představit stroj, který by toto dokázal. Ani v současné době nemáme úplnou jistotu, že nejmodernější technologie jsou schopny skloubit obě výše uvedené stránky. Ovšem technologie opravdu uhání mílovými kroky vpřed a křivka pokroku neustále stoupá vzhůru.

3.1 Zákony robotiky – definice umělé inteligence

Toto vše vede k otázce umělé inteligence a problémům s ní spojeným. Existuje velké množství definic, kterými je umělá inteligence vykládána. Už i některé stroje vykazují jakousi jednoduchou inteligenci. Navzdory tomu, že v podstatě dodnes zcela přesně nevíme, co to samotná inteligence je. Zelinka uvádí: „*K tomu, aby něco mohlo spolehlivě fungovat, je nutné vědět, co to je a jak to funguje.*“ Lépe řečeno, člověk jako takový má potřebu kontrolovat věci kolem sebe. Dostává se mu jakési jistoty, možná i pocitu bezpečí. Ale co když by existovalo něco, co by člověk pochopit nedokázal? Stejně tak je to i s inteligencí. Abychom jí mohli důvěřovat, měli bychom vědět, jak se chová a co od ní můžeme očekávat. Jak jsem uvedl, definic je mnoho, například jedna z nich zní: „*Umělá inteligence se zabývá problematikou postupu zpracování poznatku – osvojováním a způsobem použitím poznatku při řešení problému.*“³⁰

Alan Turing v podstatě definoval umělou inteligenci operačně, tzn. pomocí experimentu, jak je uvedeno v předchozí kapitole. Pojem „umělá inteligence“ byl však poprvé použit v roce 1956 na MIT ve Spojených státech a definoval ho John McCarthy. Podle jeho tvrzení by stroj v sobě skrývající umělou inteligenci měl umět hrát složité hry, jako například šachy. Měl by být schopen konat zásadní rozhodnutí v oblastech jako lékařství, vojenství atd., s čímž souvisí i potřeba zpracování hlasu. Stroj by měl umět zaznamenat hlas, porozumět a správně zareagovat. Také by měl, jak již bylo zmíněno, umět napodobit biologickou schopnost neuronových sítí ve smyslu konání

³⁰ ZELINKA, Ivan. *Umělá inteligence: hrozba nebo naděje?*. Praha: BEN - technická literatura, 2003. ISBN 80-7300-068-7. Str. 26

specifických úkolů a simulování inteligence. V neposlední řadě stroj ve formě autonomní jednotky by měl být interaktivní v kontaktu se svým okolím.³¹

V dnešní době existují roboty, které jsou vytvořeny pro určitý druh činnosti, jako již zmíněný Watson ve hře Jeopardy! (viz první kapitola), ale stále nemáme natolik všestranného robota, který by dokázal nahradit či zastoupit lidskou činnost ve všeobecném konání.

3.2 Asimovovy zákony a Moravcův paradox

V dnešní době, kdy se vývoj robotiky výrazně zrychluje, se pro člověka stává aktuální otázka: „Co kdybychom nad vším, co tvoříme, ztratili kontrolu?“ Již dříve bylo jasné, že budou muset existovat nějaké obecné zákony pro kontrolu potencionálních strojů, aby se nevymkly kontrole. Spisovatel sci-fi žánru Isaac Asimov, jeden z prvních lidí zabývajících se otázkou umělé inteligence, v roce 1941 stanovil tři pravidla, která by měla zajišťovat, aby se robot nikdy nepostavil proti svému tvůrci:

- 1. Robot nesmí ublížit člověku, ani svou nečinností dopustit, aby člověku bylo ublíženo*
- 2. Robot musí uposlechnout příkazů člověka, kromě případů, kdy by uposlechnutí rozkazu bylo ve sporu se zákonem prvním*
- 3. Robot musí chránit vlastní existenci, pokud tato ochrana není v rozporu s prvním a druhým zákonem³²*

Tento logický komplex doplnil ještě o nultý zákon, který říká, že „robot nesmí ublížit lidstvu, ani svou nečinností dopustit, aby lidstvu bylo ublíženo.“

³¹ ZELINKA, Ivan. *Umělá inteligence: hrozba nebo naděje?*. Praha: BEN - technická literatura, 2003. ISBN 80-7300-068-7. Str. 27

³² Tamtéž, str. 28

Je však zřejmé, že stroje, které by byly vytvořeny, aby se chovaly inteligentně a autonomně, tedy stroje, jenž se učí překonávat překážky, by mohly mít i své vlastní cíle.³³

Asimovovy zákony mají na robotiku vliv již přes sedmdesát let. Jeden z neznámějších zkonstruovaných robotů na počátku 21. století byl ASIMO (Advanced Step in Innovative Mobility) od firmy Honda. Paradoxně se proslavil nechvalně známou předváděcí akcí v roce 2006 v Tokiu, kde byl představen. Při výstupu na pódium se robotovi podlomila kolena a rozbil si hlavu o podlahu. Tímto incidentem ještě před samotnou prezentací prokázal, že se nedokáže řídit třetím Asimovým zákonem.³⁴ ASIMO od té doby prošel řadou vylepšení a dnes již dokáže běhat, tančit, nosit lidem kávu, kopat do míče i chodit po schodech. Co považují za udivující a svým způsobem zvláštní, je to, že něco tak propracovaného, jako je autonomní robot, i přes veškerá vylepšení stále nedokáže spoustu věcí. To odkazuje na tzv. Moravcův paradox. Stejně tak, jak autoři popisují, odborník na robotiku Hans Moravec tvrdí, že: „*je poměrně snadné vyrobit počítače, které podávají výkony dospělých lidí při testech inteligence nebo hraní dámy, ale těžké či nemožné je jim dát dovednosti jednoročního dítěte, co se vnímání a pohyblivosti týče.*“ Je to zapříčiněno hlavně tím, že vyšší logické myšlení vyžaduje paradoxně méně výpočtů než senzomotorické dovednosti.³⁵ Nynějším strojům chybí lidská všestrannost. Často býváme uchvázeni, když vidíme automatické stroje v továrnách, jak jim jde bez sebemenšího zaváhání či únavy vše skvěle od ruky. Důležité je si uvědomit, že bez lidského aspektu by tyto stroje nefungovaly. Člověk je stále musí k lince odvést, naprogramovat a zapojit.³⁶

³³ ZELINKA, Ivan. *Umělá inteligence: hrozba nebo naděje?*. Praha: BEN - technická literatura, 2003. ISBN 80-7300-068-7. Str. 29

³⁴ BRYNJOLFSSON, Erik. *Druhý věk strojů: práce, pokrok a prosperita v éře špičkových technologií*. Přeložil Filip DRLÍK. V Brně: Jan Melvil Publishing, 2015. Pod povrchem. ISBN 978-80-87270-71-4. Str. 36

³⁵ Tamtéž, str. 36

³⁶ Tamtéž, str. 37

3.3 Moorův zákon

V této podkapitole pomínu fyziologický vývoj humanoidních robotů, kteří jsou schopni chodit po schodech a autonomně se učit, a zaměřím se na jejich digitální stránku. Gorgon Moore, spoluzakladatel Intelu, tvrdil, že díky integrovaným obvodům budeme svědky nevídaného rozvoje. Říkal, že klíč je právě v nich. Jeho vize, jak Brynjolfsson uvádí, byla: „*Hustota tranzistorů na integrovaném obvodu při minimální ceně komponent se každý rok zvýší zhruba dvojnásobně ... v dohledné době se tato rychlost nezmění, možná se i zvýší. Z dlouhodobého hlediska je tento růst nejistý, ale není důvod pochybovat o tom, že tato rychlost bude nejméně po 10 let konstantní*“³⁷

V podstatě to znamená, že hustota tranzistorů, coby výpočetní síla se každým rokem zdvojnásobuje, čili v roce 1963 jste si mohli koupit jednu tolik výpočetního výkonu za dolar, než v roce 1962. Tento výrazný pokrok je masivní a tak výrazný hlavně proto, že v digitálním světě jsou limity výrazně volnějšího charakteru. Omezení v tomto odvětví se týkají například toho, kolik elektronů může projít integrovaným obvodem za sekundu nebo jak rychle mohou světelné paprsky cestovat optickým kabelem. Je to tedy otázka vývoje. Jen pro představu, superpočítač Cray – 2 z roku 1985 a iPad 2 od společnosti Apple měly takřka stejnou výpočetní rychlost a to je rozdíl pouhých pár desetiletí. Navíc iPad disponoval zařízeními jako kamera, wi-fi, GPS navigace a nespočet jiných funkcí, o kterých tvůrci superpočítačů tehdejší doby netušili.

Moorův zákon má mnoho odpůrců i příznivců, nicméně jeho platnost, a to hlavně díky „chytrým“ modifikacím tomto průmyslu, trvá.³⁸

³⁷ BRYNJOLFSSON, Erik. *Druhý věk strojů: práce, pokrok a prosperita v éře špičkových technologií*. Přeložil Filip DRLÍK. V Brně: Jan Melvil Publishing, 2015. Pod povrchem. ISBN 978-80-87270-71-4. Str. 47

³⁸ Tamtéž, str. 56 – 57

4. Následky hojnosti a rozvoje

Případ s ASIMEM naznačil, že co se týče flexibility, lidé mají stále navrch i přesto, že problém stability ASIMO nakonec dokázal eliminovat. Samozřejmě jen díky skupině odborníků. Nicméně by bylo v tuto chvíli naivní psát o robotech, kteří převezmou nadvládu nad světem. Naopak jsem přesvědčen, že díky veškerým vymoženostem v digitálním světě, nevyjímaje počátky umělé inteligence, se nám život výrazně ulehčil. Důležité jsou v tuto chvíli stroje kognitivních úkonů.

Příkladem může být i počín firmy OrCam z roku 2013, jejíž odborníci do brýlí zakomponovali malý počítač a digitální snímače, které byly propojeny algoritmy za účelem poskytnout nevidomým zrakové počítky. Dalším příkladem mohou být i invalidní vozíky ovládané myšlenkami pacienta, zkrátka technologie už dnes umí pomoci těm, kterým lidé jako takoví pomoci nedokážou.³⁹ V některých oborech zdravotnictví se již UI používá. Andrew Beck vyvinul systém zvaný C-Path, který dokáže diagnostikovat rakovinu prsu. Je zajímavé, že odborníci – lidé zkoumají vždy stejnou skupinu několika vlastností rakovinných buněk, kdežto C-Path prohlíží vzorek vždy „čerstvým“ okem. Navíc určil tři skupiny, podle kterých se dá posoudit, jakou má pacient šanci na přežití, o tom patologové doposud nevěděli.⁴⁰

To vše nám poskytuje neuvěřitelné možnosti a blahobyty. Je toto všechno ale ze své podstaty dobré, nebo je to jakýsi klam, který vědomě tvoříme? Vzhledem k preciznosti a spolehlivosti strojů je etický aspekt spojen i s otázkou pracovních míst pro lidi, na kterou by se nemělo zapomínat.

³⁹ BRYNJOLFSSON, Erik. *Druhý věk strojů: práce, pokrok a prosperita v éře špičkových technologií*. Přeložil Filip DRLÍK. V Brně: Jan Melvil Publishing, 2015. Pod povrchem. ISBN 978-80-87270-71-4. Str. 92

⁴⁰ Tamtéž, str. 94

UI má samozřejmě i stinné stránky, které bych chtěl v této práci také zmínit. Každá epocha přináší nové inovace vzbuzující strach, nebo ohromení, působící dojemem mocného nástroje, což vyvolává velké množství etických otázek očekávaných i překvapivých.

4.1 Stroj jako pracovní síla

Takto velký rozmach za tak krátkou dobu nemá vždy jen pozitivní dopad. Podle autora již po roce 1800 v anglických textilních továrnách nahrazovaly práci lidí automatické tkalcovské stavy. Tehdejší pracovníci, též „luddité“, pojmenováni po hlavním iniciátorovi Nedu Luddovi, se bouřili a ničili stroje i továrny. Byla to jakási první předzvěst toho, že automatizace může nepříznivě ovlivnit mzdy i pracovní příležitosti lidí.⁴¹ V tom se rozcházely dva pohledy na věc. Jedna skupina odborníků označila tuto situaci jako stav, kdy je nezaměstnanost jen dočasná. A následně pak vznikne více pracovních míst, než kolik jich zaniklo. Politolog William Leiserson připodobnil nezaměstnanost k hasičům, kteří čekají ve zbrojnici na poplach. Také by se dalo říci, že momentálně v té chvíli vlastně nepracují. John Bates Clark a jeho příznivci nahlíželi na tuto situaci dynamicky – některá místa technologie vezme, jiná stvoří.⁴²

Naopak John Keynes, anglický ekonom, tvrdil, že automatizace může lidi o práci připravit trvale. Ve své eseji popsal tzv. technologickou nezaměstnanost, což znamená, že úspora pracovní síly je rychlejší, než je nacházení nového využití pro tuto pracovní sílu. Domníval se, že klesající ceny nemusí vždy nutně vést k tomu, aby bylo konzumováno větší množství služeb či zboží. Také tvrdil, že neudržíme krok s technologií a lidé si budou muset hledat novou práci, čili na nějakou dobu nastane období nezaměstnanosti. Tehdejší situaci v otázce

⁴¹ BRYNJOLFSSON, Erik. *Druhý věk strojů: práce, pokrok a prosperita v éře špičkových technologií*. Přeložil Filip DRLÍK. V Brně: Jan Melvil Publishing, 2015. Pod povrchem. ISBN 978-80-87270-71-4. Str. 169

⁴² Tamtéž, str. 169

pracovních míst změnila druhá světová válka, po které tato otázka nabyla opět na aktuálnosti.⁴³

Čemu tedy věřit? Obecně vzato ovládl ekonomiku první proud s názorem, že technický pokrok vytvoří více pracovních míst než těch, které stihne zastoupit.⁴⁴ Nyní jsme však ve fázi, kdy se pokrok rozvíjí čím dál rychleji. Co by se tedy stalo, kdyby práci člověka zcela nahradil stroj? Musel by si práci hledat jinde, ale zaměstnal by ho někdo, když snazší je „zaměstnávat“ stroj? Pokud by stroj dokázal zastoupit člověka v jeho práci, zůstal by tak člověk natrvalo nezaměstnaný?

4.2 Experimenty s androidy

Dnes si již dokážeme představit roboty jako plně automatizované pracující dělníky v továrnách. Například otočná ramena jsou naprogramována k úkonům, které zvládají s přesností a harmonií. Každý robot v továrně zastupuje funkci jednotlivce a dohromady tvoří symfonii pracovního procesu, kterou můžeme s úžasem sledovat. Jistě zde musí figurovat člověk jako vyšší moc nad těmito „pracovníky“, která jim dává „život“ v podobě naprogramování a zdroje energie.

Obecně vzato se nabízí možnost následující situace. Vývoj by dospěl do bodu, kdy by lidskou práci zastoupili androidi se schopnostmi jako lidé. Roboti by disponovali výjimečnými přednostmi, jako je nízká cena a schopnost pracovat nepřetržitě. Zřejmě by to vedlo k narušení pracovního trhu, protože firmy by upřednostnily právě androidy vzhledem k jejich finančním i pracovním výhodám. Následkem by byla vysoká nezaměstnanost. Čili by se ekonomická hodnota vztahovala jen na majitele androidů a zároveň by byly jedinými

⁴³ BRYNJOLFSSON, Erik. *Druhý věk strojů: práce, pokrok a prosperita v éře špičkových technologií*. Přeložil Filip DRLÍK. V Brně: Jan Melvil Publishing, 2015. Pod povrchem. ISBN 978-80-87270-71-4. Str. 173

⁴⁴ Tamtéž, str. 171

konzumenty. Práce ostatních lidí by byla *de facto* bezcenná. Výše uvedená fakta upozorňují na to, že pokrok není vždy spojen s tvorbou pracovních míst. Respektive neexistuje předpis, který by myslel v ohledu pracovním na pracovníky samotné. Jsem přesvědčen, že lidé práci potřebují. Nemyslím tím pracovat za každou cenu až na pokraj sil, ale nějakým způsobem se realizovat a vytvářet hodnoty. I stroje mají slabé stránky. Ačkoliv např. svou silou dokážou to, co člověk ne, pravděpodobně nebudou lidi nahrazovat, ale doplňovat. To by mělo za následek vytváření nových produktů, které by stroje nebo lidé samostatně vyrábět nemohli a naopak.⁴⁵ Toto je jedna z mnoha etických otázek, které by se v dalších krocích vývoje měla věnovat pozornost. Umělá inteligence by měla být vyvíjena tak, aby bylo předcházeno fenoménům, jako je právě nezaměstnanost nebo společenská nerovnováha.⁴⁶

Je to samozřejmě krásná myšlenka, že by se stroje a lidé navzájem doplňovali. Avšak vývoj probíhá nepřetržitě a všude a jsem přesvědčen, že svět není zcela jednotný ve svých cílech. Je těžké, až bych řekl, nemožné, sjednotit priority všech k dosažení obecného blaha. Technologie nám slouží k pobavení, usnadnění práce i k pomoci, ale cítím, že stojíme na jakémsi prahu odpovědnosti za naši budoucnost. Nejsem pesimista, ale pokládám si otázku, zda si uvědomujeme i potenciaální nebezpečí v případě pokročilých technologií.

⁴⁵ BRYNJOLFSSON, Erik. *Druhý věk strojů: práce, pokrok a prosperita v éře špičkových technologií*. Přeložil Filip DRLÍK. V Brně: Jan Melvil Publishing, 2015. Pod povrchem. ISBN 978-80-87270-71-4. Str. 176 – 177

⁴⁶ Úvodní stránka | Vláda ČR [online]. Copyright © [cit. 15.09.2019]. Dostupné z: https://www.vlada.cz/assets/evropske-zalezitosti/aktualne/AI-pravne-eticka-zprava-2018_final.pdf. Str. 14

4.2.1 Mohou být stroje morální?

Morálka, z latinského *moralis* nebo *mores*, tedy mravy, definuje vzory chování, které se týkají vzájemných vztahů mezi lidmi. Není pochyb, že je to záležitost ryze lidská, ale co stroje? Je možné uvažovat o stroji, který by se díky svému inteligentnímu chování mohl zařadit do společnosti? Co je tím rozhodujícím činitelem?

Naše lidství je samo o sobě specifické a jeho projevy lze jen stěží přesně interpretovat. Podle některých odborníků právě to křemíkové obvody nedokážou, proto panuje v tomto směru spíše negativní názor. Ovšem nelze opominout takzvanou „wet artificial intelligence“, čili „mokrou“ umělou inteligenci, což znamená, že robotický vývoj by se přesunul na organickou úroveň pod mikroskopy. Změnilo by to potom pohled na věc? Tento proces je stále na prebakteriální úrovni a zatím nejsme schopni ho programovat, nicméně lidstvo dokáže překonat nemalé problémy.⁴⁷

Dle mého úsudku by bylo příliš jednoduché říci, že co není živé a není to člověk či zvíře, nemůže morálně jednat, a tak dávám šanci i stroji, který má mravní potenciál. V knize *Moral machines* její autoři zmiňují, že A. Newell a H. Simon z univerzity Carnegie Mellon manifestovali umělou inteligenci následovně: „*Systém fyzických symbolů má obecně nezbytné a dostatečné prostředky k inteligentnímu jednání.*“⁴⁸ Jedná se o velice odvážný výrok a bezpochyby by s ním řada lidí nesouhlasila, nicméně se v tomto případě polemika stává realitou. Jde o jakousi predikci? Mohla by UI nějakým způsobem jednat sama za sebe? Je to problematické vzhledem k tomu, že nelze zatím přesně stanovit univerzální pravidla, kterými by se stroje za každých okolností měly řídit. Další otázkou je, jestli bychom jim dali vůbec možnost. Možná v sobě

⁴⁷ WALLACH, Wendell a Colin. ALLEN. *Moral machines: teaching robots right from wrong*. New York: Oxford University Press, 2009. ISBN 0195374045. Str. 55

⁴⁸ Tamtéž, str. 57

umělá inteligence skrývá určitý potenciál, ale nyní obecně panuje odmítavý postoj. Ale co kdyby právě umělá inteligence byla schopna přijít na řešení, která by člověka nenapadla? Proto autoři studie „Výzkum potenciálu rozvoje umělé inteligence v České republice“ pro Úřad České republiky doporučují hlubší analýzu a následné vyvození možných rizik a výhod.⁴⁹

Lidský organismus se vyvinul z biochemické látky a to, co nazýváme myšlením nebo rozumem, je jakousi přirozenou funkcí našeho mozku. Dohromady vše tvoří složitý, těžko uchopitelný systém. Umělá inteligence se vyvíjí na logické platformě, což dává počítačům určité výhody, ať už eliminování emocí, či širší škálu výpočetních úkonů, není to úsudek vytvořený lidským způsobem, jak jsme zvyklí. Jak jsem zmínil, toto by mohlo vést k lepším výsledkům než případě lidského faktoru.⁵⁰

Uvedl jsem, že lidský mozek má nezpochybnitelnou masivní výpočetní sílu díky neuronům a možnosti komplexně zpracovávat informace. Technologie ho stále napodobit v jeho všestrannosti nedokáže, ale díky exponenciálnímu růstu tohoto oboru může být brzy vše jinak. Ostatně jak Ray Kurzweil předpovídá, ekvivalentní kapacita jednoho lidského mozku bude k dispozici v stolních počítačích do roku 2020. To však není vše, do roku 2029 budou vědci a odborníci schopni konstruovat, modelovat a simulovat všechny oblasti mozku.

To poskytne lidem software a algoritmické metody simulující všechny schopnosti lidského mozku, včetně emoční inteligence. Pak už bude jen na nás, jestli budeme ochotni přijmout takovou výzvu.⁵¹

⁴⁹ Úvodní stránka | Vláda ČR [online]. Copyright © [cit. 15.09.2019]. Dostupné z: https://www.vlada.cz/assets/evropske-zalezitosti/aktualne/AI-pravne-eticka-zprava-2018_final.pdf

⁵⁰ WALLACH, Wendell a Colin. ALLEN. *Moral machines: teaching robots right from wrong*. New York: Oxford University Press, 2009. ISBN 0195374045. Str. 71

⁵¹ Neuralink: How the Human Brain Will Download Directly from a Computer. Interesting Engineering [online]. Copyright © Copyright 2019 [cit. 15.08.2019]. Dostupné z: <https://interestingengineering.com/neuralink-how-the-human-brain-will-download-directly-from-a-computer>

Dnešní technologie a lidstvo obecně dokážou mnohé. Výše zmíněné predikce a objevy jsou fascinující. Ale já bych do budoucna koncipoval otázku místo „Mohou být stroje morální?“ jako „Chceme, aby byly stroje morální“?

5. Etická otázka

Kromě potencionální nezaměstnanosti a otázky rovnosti jsou tu samozřejmě i další etické problémy, které je potřeba řešit současně s vývojem umělé inteligence. Odpovědnost za autonomní stroj je však fenoménem, který vyžaduje majoritní responzi. Dle odborníků, jakým je například Virginia Dignumová, je nutné tuto otázku jasně specifikovat a zodpovědět pro další vývoj autonomního jednání.⁵²

Vývoj autonomie za pomoci strojového učení umožňuje UI systémům jednat bez lidské pomoci a přítomnosti. Zajištění odpovědných systémů, jak říká Dignumová, přispěje k větší důvěře v ně. To by mělo být klíčové pro budoucnost. Je zapotřebí začít chápat umělou inteligenci jako součást sociálně technických vztahů. Nový přístup tak zajistí, aby umělá inteligence přispívala k obecnému dobru. Pochopení je nejdůležitějším aspektem. Dignumová dále představuje tři principy, podle kterých by se měla etická umělá inteligence řídit:

1. Odpovědnost

Dignumová tvrdí, že v tomto ohledu se odpovědnost týká samotných inteligentních strojů. Měly by být schopny zdůvodnit kroky svého jednání ostatním participantům, se kterými interagují. To musí být zřejmé z předešlých kroků, které UI použila. Jak říká, tato operace zahrnuje potřebu zastoupení mravních a společenských norem, které využívá UI pro své jednání. Odpovědnost se tedy týká jak vytváření přesvědčení a přijímání rozhodnutí, tak i funkce vysvětlení.

2. Zodpovědnost

Neméně důležitá je však i zodpovědnost samotných lidí za UI, kteří musí koordinovat její rozhodování a kontrolovat chyby. Tím, jak roste řetězec

⁵² The ART of AI Design — Accountability, Responsibility, Transparency. Design for Values & Responsible Innovation at TU Delft - Portal [online]. Copyright © Copyright 2013 [cit. 02.11.2019]. Dostupné z: <https://www.delftdesignforvalues.nl/2018/the-art-of-ai-accountability-responsibility-transparency/>

odpovědnosti, je zapotřebí prostředků k propojení rozhodnutí UI se správným uplatňováním opatření zúčastněných stran v rozhodování. Výzkumní pracovníci by si tak měli být vědomi také své vlastní odpovědnosti.

3. Transparentnost

Transparentnost sama o sobě se týká hlavně možnosti a potřeby popisovat a kontrolovat mechanismy, se kterými UI pracuje. Je také zapotřebí mít možnost tyto systémy upravovat formou otevřeného kódu. Pozorovatelé i zúčastnění požadují za každých okolností vysvětlení jednotlivých operací. Dignumová dodává, že jsou potřeba metody pro kontrolu algoritmů i jejich výsledků pro správu dat a jejich dynamiku.⁵³

Jak je to tedy se samotnou odpovědností? Kdo ji ponese? Dignumová říká, že je potřeba uvědomit si, jak lidé technologie chápou a jak s nimi pracují, aby bylo možné vyvinout rámec pro odpovědnou umělou inteligenci. Odpověďmi na tyto otázky musí být informace o předpisech k odpovědnému používání umělé inteligence, které zavádějí společnosti.⁵⁴

5.1 Praktické využití

Jako ilustrace mohou posloužit již zmíněná autonomní vozidla – trend moderní doby, který je na vzestupu. Každý rok na silnici zahynou statisíce lidí. Podle odborníků z Centra Karla Čapka pro studium hodnot ve vědě a technice by zavedení autonomních vozů mohlo předejít až 90% dopravních nehod. Dále by usnadnilo řízení starším či handicapovaným lidem, zajistilo plynulejší provoz a přineslo mnoho dalších výhod. Ovšem pokud by se jednalo o centrální systém, který by řídil veškerou jejich logistiku, kdo by nad ním měl kontrolu?

⁵³ The ART of AI Design — Accountability, Responsibility, Transparency. Design for Values & Responsible Innovation at TU Delft - Portal [online]. Copyright © Copyright 2013 [cit. 02.11.2019]. Dostupné z: <https://www.delftdesignforvalues.nl/2018/the-art-of-ai-accountability-responsibility-transparency/>

⁵⁴ Tamtéž

Co by se stalo, kdyby přece jen autonomní vozidlo havarovalo a ublížilo člověku? Kdo by byl vinen?⁵⁵ Mnozí odborníci k ilustraci tohoto problému využívají tzv. tramvajové dilema.⁵⁶

Patrick Lin popisuje situaci následovně. Představme si, že se ocitáme při jízdě na dálnici v situaci, kdy před sebou máme naložený jedoucí kamion a po obou stranách jiné auto a motorkáře. Z korby kamionu se uvolní těžký náklad a my nebudeme schopni předejít kolizi. Jak se rozhodneme? Narazíme do nákladu, abychom neohrozili ostatní účastníky provozu, čímž ohrozíme sebe? Uhneme vpravo, abychom minimalizovali naše ztráty, ale ohrozíme na životě motorkáře? Nebo uhneme vlevo a předejdeme tak teoretickým ztrátám na našich životech i na životech v druhém autě?⁵⁷ Možných kombinací existuje nespočet. Člověk za volantem by zareagoval reflexivně. Autonomní vozidlo, které by bylo zproštěno smyslů a pocitů by muselo udělat rozhodnutí.⁵⁸ Jaké ale?

Výrazem odpovědnosti vozidla by bylo přesné zdůvodnění, které by mělo vycházet z mravních a společenských norem. Bezpochyby v této situaci, pokud by nebylo jiné východisko, by se snažilo především ochránit život pasažérů a minimalizovat škody na okolí. S tím souvisí i transparentnost, díky které lze zkoumat příčiny rozhodnutí, což dává možnost potencionálních úprav či vylepšení. Nakonec zůstává zodpovědnost samotných konstruktérů, či společností, kterou by měli nést.

⁵⁵ Centrum Karla Čapka | Novinky [Etika a autonomní vozidla]. Karel Čapek Center [online]. Copyright © 2018, Centrum Karla Čapka pro studium hodnot ve vědě a technice [cit. 01. 11. 2019]. Dostupné z: <https://www.cevast.org/cz/news/28-etika-a-autonomni-vozidla>

⁵⁶ Tramvajové dilema – Wikipedie. [online]. Dostupné z: https://cs.wikipedia.org/wiki/Tramvajov%C3%A9_dilema

⁵⁷ Patrick Lin: The ethical dilemma of self-driving cars | TED Talk Subtitles and Transcript | TED. TED: Ideas worth spreading [online]. Copyright © TED Conferences, LLC. All rights reserved. [cit. 03.12.2019]. Dostupné z: https://www.ted.com/talks/patrick_lin_the_ethical_dilemma_of_self_driving_cars/transcript#t-31642

⁵⁸ Centrum Karla Čapka | Novinky [Etika a autonomní vozidla]. Karel Čapek Center [online]. Copyright © 2018, Centrum Karla Čapka pro studium hodnot ve vědě a technice [cit. 01. 11. 2019]. Dostupné z: <https://www.cevast.org/cz/news/28-etika-a-autonomni-vozidla>

Zároveň by ale toto automatizované rozhodování mohlo být příčinou favorizování, či diskriminování, protože by bylo naprogramováno předem. Mohlo by být přece jen rozhodnutí založené na reakci v takovém případě lepší, než predeterminovaný úsudek autonomního vozu za účelem snížení škod? Nebo by bylo prospěšnější co nejdříve uvést do provozu právě bezpilotní vozidla s předpokladem snížení počtu dopravních kolizí?⁵⁹

Nejedná se však pouze o problematiku autonomních vozidel. Nyní se pokusím vytyčit určité formy nedostatků, které umělá inteligence má. Vývoj takové inteligence můžou ohrozit předsudky vývojářů či společností. Důležitou součástí je již zmíněná transparentnost systémů právě z hlediska odpovědnosti a další. Umělá inteligence je pro oblast etiky další výzvou jednak proto, že je založena na autonomním rozhodování, ale také proto, že je nutné pokusit se vytvořit jakýsi etický rámec, který by fungoval pružně jako živý proces, pro ukotvení tohoto fenoménu.

5.1.1 Algoritmické předsudky

Vzhledem k tomu, že se s předsudky vůči ostatním vyrovnává každý jedinec, o to větší důraz by měl být kladen na samotné firmy či vývojáře inteligentních strojů. Jedním z velkých etických problémů je předcházení zakotvení tzv. algoritmických předsudků do inteligentních modelů. Čili „otisk“ každého vývojáře do stroje by mohl mít vliv na jeho pozdější chování. Konstrukteři by se tedy měli řídit určitými etickými pravidly. Pokud připustíme automatizovanou budoucnost, mohl by pak vzhled robota nebo jeho úsudky, které odrážejí chápání světa samotného vývojáře, vést například k větší diskriminaci osob nebo skupin.⁶⁰ Tato problematika už začíná u

⁵⁹ The trolley dilemma in AI. Strongbytes [online]. Dostupné z: <https://strongbytes.ai/the-trolley-dilemma-in-ai/>

⁶⁰ Úvodní stránka | Vláda ČR [online]. Copyright © [cit. 15.09.2019]. Dostupné z: https://www.vlada.cz/assets/evropske-zalezitosti/aktualne/AI-pravne-eticka-zprava-2018_final.pdf Str. 11 – 12

tréninkových dat. V zásadě jde o „elektronickou učebnici“, se kterou vývojář a umělá inteligence pracuje. Je to typ vstupních dat určený ke správnému naprogramování za pomoci různých metod v závislosti na proveditelnosti a požadavcích. Přesněji řečeno, jde o důležitá data, jež pomáhají strojům se učit a předvídat. Pokud by bylo nakládáno s nežádoucími daty, ve stroji by se uzamkl algoritmus, který by byl ovlivněn jakýmsi předsudkem.⁶¹ Souvisí s tím jakási klasifikace osob, kterým jsou na základě jejich identifikace strojem přiřazeny určité vlastnosti, kterými však daná osoba nemusí vůbec disponovat. Ty spadají do určitých skupin, na jejichž základě stroj vyhodnocuje svůj závěr. Algoritmus pak zjištěné informace zaznamená a zařadí člověka do určité kategorie, ze které se již nemůže vymanit. Doposud, jak autoři studie pro Úřad České republiky uvádí, byly zaznamenány chybové případy upřednostňování mužů před ženami v náborové selekci, či byly chybně stanoveny automatickými systémy vyšší tresty pro Afroameričany. Nyní již existuje výzkum, který má za cíl eliminovat takové případy, a jsou vyvíjeny jiné algoritmy pro identifikaci potenciální diskriminace.⁶²

5.1.2 Netransparentnost

Netransparentnost v souvislosti s předchozím tématem hraje též významnou roli a masové využití UI může vést k negativním důsledkům ve společnosti. Inteligentní systémy jsou často považovány za nezávislé v řešení problémů, tudíž se od nich očekává nezájatost a přesnost. Avšak kdyby obsahovaly zmíněné chyby, mohlo by to vést k negativním důsledkům. Jestliže by stroj chyboval, je nutné zjistit v čem a mít možnost v jeho systému tuto chybu přezkoumat. Jinak by docházelo nepřesným a chybovým verdiktům. Jelikož

⁶¹ What is AI Training Data? - Cogito Tech LLC - Medium. Medium – a place to read and write big ideas and important stories [online]. Dostupné z:

<https://medium.com/@cogitotech/what-is-ai-training-data-f26c27dcc617>

⁶² Úvodní stránka | Vláda ČR [online]. Copyright © [cit. 15.09.2019]. Dostupné z:

https://www.vlada.cz/assets/evropske-zalezitosti/aktualne/AI-pravne-eticka-zprava-2018_final.pdf Str. 11 – 12

osoba touto možností nedisponuje, často se tak stává méně věrohodnou než daný systém, což může, podle autorů studie, vést k prohlubování společenských rozdílů a snižování autonomie jedince. V praxi mají bohatí lidé často privilegia ve formě osobního přístupu, ale o valné většině osob dnes rozhodují právě automatické systémy. Otázkou tedy zůstává, jak s přesností určit, v čem daný stroj dělá chyby a jak přezkoumat jeho fungování. S nadcházejícím dobou se budou s největší pravděpodobností zvyšovat možnosti řešení, jak tomuto problému předcházet.⁶³

5.1.3 Soukromí

Vzhledem ke komplexitě automatických systémů, které dokážou zpracovávat informace o mase lidí formou přirozeného jazyka či komunikace, objevuje se i otázka ochrany soukromí. Soukromí a ochrana osobních údajů jsou pro mnohé jedince klíčové a umělá inteligence poskytuje nástroj k monitorování lidí a získávání informací o nich. Příkladem pozitivního užití je kupříkladu, jak autoři popisují, monitorování pozornosti dětí v jedné z čínských škol, což nejspíše vede k lepším známám. Ovšem hrozí i nežádoucí ovlivňování automatickými systémy. Takovéto vlivy by mohly v člověku vyvolávat pocit, že je sledován, a následně by podle toho měnil své chování. Nemluvě o častějších závislostech na technologiích, čímž se vytváří nepoměr a snižuje se tak podíl osobní komunikace nebo dovednosti jedince.⁶⁴

Přináší to i svá pozitiva. Nové studie polemizují i o jiných vlivech na člověka, než jsou například chytré aplikace, které trénují kognitivní funkce. Stimulace mozku k lepším výkonům nebo úprava genetického kódu nejsou zdaleka hudbou budoucnosti. Souvisí s tím však problém, že lidé se k otázce „vylepšování“ člověka staví různě. Jedni tvrdí, že by mělo být naší morální

⁶³ Úvodní stránka | Vláda ČR [online]. Copyright © [cit. 15.09.2019]. Dostupné z: https://www.vlada.cz/assets/evropske-zalezitosti/aktualne/AI-pravne-eticka-zprava-2018_final.pdf Str. 12

⁶⁴ Tamtéž, str. 13

povinností, aby naši potomci byli nositeli lepšího genomu, druhá strana zastává názor, že by se neměla upravovat, ani jinak vylepšovat naše genetická stopa.⁶⁵

5.1.4 Právo

Etická problematika, která se týká udělování práv robotům, je jedním z kontroverzních témat. Kdyby si robot začal nějakým způsobem uvědomovat svou vlastní existenci, měla by mu být přidělena určitá práva? Pokud ano, bylo by v pořádku, aby měl stejné nároky a práva jako jeho tvůrce? Na toto téma jsou vedeny bouřlivé diskuze. Pokud bychom chtěli mít umělou inteligenci pod dohledem, je zapotřebí ji nějakým způsobem kontrolovat. Ovšem je třeba brát v potaz i názor, že takové jednání by bylo otrocké, protože lidé inteligentní systémy antropomorfizují. Kvůli jejich inteligenci jim jsou připisovány lidské vlastnosti a emoce a dokážou proto navazovat hluboké vztahy.⁶⁶

5.1.5 Etická problematika v různých odvětvích

Dalším odvětvím, do kterého zasahuje umělá inteligence, může být již zmíněná a též kontroverzní sexuální otázka. O tomto ožehavém tématu se diskutuje už po staletí. Pro to, abychom mohli vůbec přemýšlet o sexu s robotem, je třeba se seznámit s důvody, proč lidé platí sexuálním pracovníkům za jejich služby. Bezpochyby to vede k různým benefitům, jedním z nich je „různorodost“. Robot takto bude schopen poskytovat rozmanitost, co se vzhledu, hlasu, proporcí a mnohého dalšího týče. Pochopitelně se nejedná jen o vzhled, často lidé vyhledávají to, čeho se jim jinak nedostává. Souvisí to i s dalšími oblastmi sexuálního světa, se kterými by „sexboti“ neměli problém. Stručně řečeno, robotické prostitutky nebo prostituti dokáží uspokojit touhy žen i mužů snadno, kdykoliv a bez odporu. Navíc mohou tak předejít různým pohlavním

⁶⁵ Úvodní stránka | Vláda ČR [online]. Copyright © [cit. 15.09.2019]. Dostupné z: https://www.vlada.cz/assets/evropske-zalezitosti/aktualne/AI-pravne-eticka-zprava-2018_final.pdf Str. 13

⁶⁶ Tamtéž, str. 14

chorobám či infekcím. V tom tkví jejich úspěch.⁶⁷ Hlavním účelem takovýchto pracovníků je dosahování uspokojení klienta bez nutnosti další osoby. Samozřejmě tento fenomén vyvolává i řadu etických otázek.

Je zřejmé, že určitou výhodou by to bylo pro jedince, kteří si nemohou dlouhodobě najít partnera. Sexboti mají prakticky stejný účel jako ženské vibrátory, které se mnohdy prodávají i na poličkách lékáren. Proč by měl být tedy problém se sexboty? Zdroj benevolentního etického postoje vůči sexbotům spatřuji v sexuální asistenci vážně handicapovaným lidem. Stejně tak lze nafukovací panny, coby erotické pomůcky, považovat za primitivní předchůdce sexuálních robotů. David Levy argumentuje, že vzhledem k rozšířenému prodeji vibrátorů, by bylo problematické označit jejich použití za nemorální, stejně jako by bylo obtížné tvrdit, že design, vývoj, výroba a prodej sexbotů jsou neetické. V samotném vztahu mezi dvěma lidmi mohou být názory rozdílné, jedni budou považovat použití sexbotů za nevěru, druhí je přivítají. Etika využívání robotické prostitutky bude velmi záviset na vnímání tohoto fenoménu společností. Tomu nahrává skutečnost, že mnoho najímaných žen-prostitutek je úzce spojeno s drogami a pohlavními nemocemi. Samozřejmě se tomu dá předejít použitím prezervativů, ačkoliv ty mnohdy také nezajišťují svou funkci či pohodlí. Navíc společnost se na prostituci samotnou ne dívá zcela přívětivě. V tomto případě by mělo použití sexbotů pozitivní dopad. Ovšem může být i zdrojem řady problémů. Americký deník *The Guardian* popsal situaci tak, že newyorské prostitutky cítí obavy z takovýchto technologií. Bojí se, že by je mohly zcela zbavit předmětu podnikání. Dle mého zřejmě oprávněně. Zároveň s touto situací je spojen další etický problém. Jak píše Levy, mnoho z nich se obrací k

⁶⁷ LIN, Patrick., Keith ABNEY a George A. BEKEY. *Robot ethics: the ethical and social implications of robotics*. Cambridge, Mass.: MIT Press, c2012. Intelligent robotics and autonomous agents. ISBN 9780262016667. Str. 223 – 226

prostituci, protože nemají žádný jiný způsob, jak si vydělat peníze.⁶⁸ Nástup sexuálních robotů tedy vyvolává mnoho neznámých. Je spojen s pozitivou i negativou, jak Levy uvádí. Zároveň bude potřeba zvážit jeho etický dopad.⁶⁹

Mezi další etické otázky ve sféře umělé inteligence patří oblasti, jako je například zdravotnictví, kde bude třeba počítat s otázkami predikce nemoci, či přežití. Lékařská praxe je velmi široká a jak jsem popisoval, umělá inteligence je velkým přínosem. Již v 80. letech minulého století byla UI, ačkoliv v primitivnějším stádiu, součástí lékařské praxe. V roce 2005 vyšel v lékařském časopisu *Clinical Window* článek doktora Paula Kligfielda z Cornellovy univerzity týkající se EKG zařízení, která jsou mimochodem dosud používána a která jsou schopna poskytovat automatickou diagnostiku, což značně ulehčuje lékařům práci. Ovšem ani v tomto odvětví nemůže vše a za všech podmínek fungovat bezchybně. Jak autor Nils J. Nilsson ze Stanfordovy univerzity popisuje: „*při hodnocení Kligfield uvedl, že u 3954 pacientů bez kardiostimulátoru 7,8% interpretací srdečního rytmu vyžadovalo revizi a kombinovaný názor dvou odborných kardiologů.*“ Bez ohledu na úspěch diagnostických systémů a UI v lékařském světě odpovědí na otázku, kdo by měl nakonec rozhodnout, zůstává vždy lékař, který má zodpovědnost, protože stroj jako takový tuto zodpovědnost nést nemůže.⁷⁰

⁶⁸ LIN, Patrick., Keith ABNEY a George A. BEKEY. *Robot ethics: the ethical and social implications of robotics*. Cambridge, Mass.: MIT Press, c2012. Intelligent robotics and autonomous agents. ISBN 9780262016667. Str. 226 – 228

⁶⁹ Tamtéž, str. 229 – 230

⁷⁰ NILSSON, Nils J. *The quest for artificial intelligence: a history of ideas and achievements*. New York: Cambridge University Press, 2010. ISBN 0521122937. Str. 623 – 625

Předmětem mnoha současných etických diskuzí kolem použití UI je i oblast vojenství a zbrojení. UI ve vojenství v mnoha případech pomáhá v boji v potencionálně nebezpečných zónách, kam jsou vysíláni roboti namísto lidských vojáků, čímž se předchází ztrátám na lidských životech. Avšak jejich využití vyvolává určité etické otázky.⁷¹

Jedním, z hlavních protagonistů je Noel Sharkey, který situaci popisuje následovně. Velkým trendem v moderním vojenství jsou bezpilotní letouny. Vojáci sedící za joystickem tisíce kilometrů daleko od svých cílů navádějí bezpilotní letouny a rozhodují tak o životě jiných ze své „komfortní zóny“. Jedná se o jakýsi „letecký simulátor“, který umožňuje zabíjet. Navíc se mohou letouny řídit autonomně, což podle některých vytváří ostrý kontrast mezi normálním životem a životem na bitevním poli. Mezi takové patří třeba bezpilotní letouny Predators, které jsou vybaveny 14ti raketami Hellfire, či kombinací raket a bomb, které od roku 2007 až do roku 2010 nalétaly bezmála 1 milion leteckých hodin. Tato menší, ačkoliv markantně ničivá část amerického letectva, jak píše Sharkey, podle tisku v Pákistánu zabila díky dronům 14 vůdců Al–Kaidy. Může to však ospravedlnit fakt, že kvůli útokům bylo též zabito přes 600 civilistů?⁷² Jistě tato metoda dokáže uchránit spoustu vojáků, kteří operují za obrazovkou, nicméně i přes to, že neexistuje nic jako bezrizikové válčení, by mohl právě tento způsob boje vést k více povstaleckým útokům na základě nerovnosti boje.⁷³

⁷¹ LIN, Patrick., Keith ABNEY a George A. BEKEY. *Robot ethics: the ethical and social implications of robotics*. Cambridge, Mass.: MIT Press, c2012. Intelligent robotics and autonomous agents. ISBN 9780262016667. Str 109

⁷² Tamtéž, str. 115

⁷³ Tamtéž, str. 125

Další etické hranice bude třeba vymezit například v oblasti komunikace, bude nutno definovat formy nevhodného vyjadřování či falešného sdělování. Dále v oblasti etiky a obchodu,⁷⁴ kdy společnosti musí vědět, jaký dopad mohou mít nové technologie, jako je UI, na podnikání.⁷⁵

Okolo tohoto tématu prozatím vzniká možná více otázek než odpovědí. Faktem je, že UI představuje potencionální rizika, ale i výhody. Podle studie pro Úřad vlády České republiky je nutností formovat etické kodexy pro určené problémy, které by měly propagovat ponechání kontroly nad inteligentními systémy člověku – „*human-in-command*“. Dále tvrdí, stejně jako Dignumová, že je nutné, aby systémy byly transparentní z důvodu zpětné vazby. Je třeba včas podchytit to, jak umělá inteligence funguje, aby se dalo předejít potencionálním hrozbám.⁷⁶

⁷⁴ Úvodní stránka | Vláda ČR [online]. Copyright © [cit. 15.09.2019]. Dostupné z: https://www.vlada.cz/assets/evropske-zalezitosti/aktualne/AI-pravne-eticka-zprava-2018_final.pdf Str. 15

⁷⁵ The business ethics of artificial intelligence. Object moved [online]. Copyright © 2019 RICS. All rights reserved. [cit. 16.09.2019]. Dostupné z: <https://www.rics.org/uk/wbef/data-technology/the-business-ethics-of-artificial-intelligence/>

⁷⁶ Úvodní stránka | Vláda ČR [online]. Copyright © [cit. 15.09.2019]. Dostupné z: https://www.vlada.cz/assets/evropske-zalezitosti/aktualne/AI-pravne-eticka-zprava-2018_final.pdf

6. Technologie a predikce budoucna

Jak už jsem předjímal, lidé už od pradávna sní, že budou moci naplnit své materiální potřeby bez úmorné práce a budou se moci věnovat svým zálibám a prožívat štěstí. Řekl bych, že je to až jakýmsi skrytým ideálem lidství, ačkoliv se to nemusí týkat všech. K tomu by nám moderní technologie a zejména umělá inteligence mohly dopomoci. Spolu s ní je spojeno velké množství etických dilemat, která mají vliv na okolní společnost, a je nutné je pojmenovat a vyřešit. Jak popisuje Brynjolfsson, v příštích čtyřadvaceti letech se výkon digitálních technologií znásobí více než tisíckrát. Také dodává, že to jistě změní ekonomiku a technologie by pro nás mohly „pracovat“.⁷⁷

Ovšem je tu i opačná stránka. Více by nás mohly začít trápit otázky existenčních rizik či konflikt svobody jako potencionální vedlejší účinky technologie. Dále i nevyhnutelná spletnost a technologická infrastruktura by měla svá rizika. Počáteční chyby v systémech by mohly způsobit nedozírné následky například v podobě superinteligence.⁷⁸ Kvůli komplexnosti by se tento prostor mohl stát cílem hackerů. Proto by v dalších krocích měly být zváženy možné dopady, aby tato oblast byla pod kontrolou a aby bylo zamezeno potencionálním rizikům.

⁷⁷ BRYNJOLFSSON, Erik. *Druhý věk strojů: práce, pokrok a prosperita v éře špičkových technologií*. Přeložil Filip DRLÍK. V Brně: Jan Melvil Publishing, 2015. Pod povrchem. ISBN 978-80-87270-71-4. Str 237 – 238

⁷⁸ Superinteligence je podle Nicka Bostroma lidmi stvořená vyšší forma inteligence, která by nám mohla být nadřazena tak, jako jsme například my nadřazeni lidoopům.

Náš druh doposud neměl možnost sebedestrukce, dnes už ji podle Brynjolfssona má: „*Naše schopnost vědět více je v konfliktu s naší schopností zabránit ostatním, aby o nás věděli.*“⁷⁹ Z budoucích technologií bychom mohli těžit, nebo na nich ztroskotat. Závěry nejsou zcela jasné, ale je zřejmé, že růst je pravděpodobně nezastavitelný.

6.1 Názory

Názory odborníků na budoucnost technologického světa jsou různé. Jedni souhlasí a přitakávají jeho rozmachu, jiní jsou spíše skeptičtí vůči tomu, co může přijít. Nyní se seznámíme s některými z nich.

6.1.1 Optimisté

Známým optimistou v tématu umělé inteligence je futurolog a spisovatel Martin Ford. V rozhovoru pro Tampa Bay Times uvedl, že vynález umělé inteligence se dá srovnat s vynálezem elektřiny, ale s více nevýhodami. Podle něj má umělá inteligence potenciál a pomohla by řešit globální chudobu či by dokázala léčit závažnější onemocnění, ale zároveň připouští, že je možné, že vytvoří větší nerovnost.⁸⁰

Podle interview pro Great Thinkers Ford uvedl 3 sekce pracovních míst, které budou pro blízkou budoucnost robotizace v bezpečí. Jednak ty, ve kterých nelze předvídat, ty, které vyžadují kreativitu. Dále ty, které zahrnují sofistikované vztahy mezi lidmi, čili například zdravotní sestřičky, lékaři nebo obchodníci. Ford také říká: „*Mezilidské sofistikované interakce s lidmi jsou něčím, co stroje nebudou v blízké době obzvláště dobré.*“ Poslední kategorií,

⁷⁹ BRYNJOLFSSON, Erik. *Druhý věk strojů: práce, pokrok a prosperita v éře špičkových technologií*. Přeložil Filip DRLÍK. V Brně: Jan Melvil Publishing, 2015. Pod povrchem. ISBN 978-80-87270-71-4. Str. 240

⁸⁰ Artificial intelligence might not threaten the human race, but it could be coming for your job. Tampa Bay, Florida news | Tampa Bay Times/St. Pete Times [online]. Copyright © 2019 All Rights Reserved [cit. 24.08.2019]. Dostupné z: https://www.tampabay.com/news/business/Artificial-intelligence-might-not-threaten-the-human-race-but-it-could-be-coming-for-your-job_171963256

podle Forda, jsou povolání, která vyžadují mobilitu, jako jsou nepostradatelní jako elektrikáři, či instalatéri. Nicméně v otázce umělé inteligence je značným zastáncem pozitivismu.⁸¹ Druhým významným pozitivistou je Ray Kurzweil, vynálezce, futurolog a vývojář společnosti Google. Kurzweil tvrdí, že nejlepším způsobem, jak předejít konfliktu s moderní technikou, je zdokonalování sebe samých. Dále poukazuje na to, že za celou dobu vývoje jsme se prakticky nesetkali s fatálním problémem, nýbrž nám UI představila mnoho užitečného, např. nalézá léky, poskytuje vzdělání či pomáhá lidem s postižením. Uvádí, že UI bude klíčovým faktorem v budování budoucnosti. Kurzweil a další optimisté věří, že možnosti umělé inteligence přinesou lidské společnosti bohatství.⁸²

6.1.2 Pesimisté

Pesimisté jsou jiného názoru. Domnívají se, to povede k nevyhnutelné katastrofě. Nick Bostrom naznačuje, že pokud by nám UI nějakým způsobem přerostla přes hlavu a snažili bychom se vytvořit superinteligenci, o které pojednává mimochodem ve své knize, vytvořili bychom tak silný systém, který by se lehce mohl vymknout kontrole. Mohli bychom jí tak vnuknout cíle, které by mohly vést v krajním případě například ke zničení lidstva. Každopádně kdyby existovala taková eventualita, Bostrom věří, že by se tak stalo.⁸³ Považuji tento názor za poněkud přehnaný, nicméně je to výzva, která by mohla motivovat lidstvo k větší opatrnosti.

Dalším je kupříkladu Elon Musk. Musk je v dnešní době velmi populární osobou ve sféře technologií, proto mě udivilo jeho jméno v řadách pesimistů.

⁸¹ Great Thinkers: Interview With Martin Ford, Futurist, Speaker, A.I. Expert. The Medical Futurist [online]. Copyright © 2019 [cit. 24.08.2019]. Dostupné z: <https://medicalfuturist.com/martin-ford-great-thinkers>

⁸² Don't fear artificial intelligence | by Ray Kurzweil | Kurzweil. Kurzweil Accelerating Intelligence [online]. Copyright © 2019 Kurzweil Network [cit. 24.08.2019]. Dostupné z: <https://www.kurzweilai.net/dont-fear-artificial-intelligence-by-ray-kurzweil>

⁸³ AI, Optimists vs Pessimists and Why The Singularity Isn't Near. Medium – a place to read and write big ideas and important stories [online]. Dostupné z: <https://medium.com/ml-everything/ai-optimists-vs-pessimists-and-why-the-singularity-isnt-near-5d3a614dbd45>

Jeho počiny a plány v oblasti elektromobilů, autonomních vozidel nebo kolonií na Marsu netřeba zmiňovat. Na první pohled je zřejmé, že technologiím je nakloněn, ale pokud jde o umělou inteligenci, jeho názor je poněkud odlišný. Dokonce ji v roce 2014 přirovnal k vyvolávání démona. V rozhovoru pro Friday with Recode's Kara Swisher poznamenal: „*pokud bude UI chytřejší než člověk, poměr relativní inteligence bude pravděpodobně podobný poměru mezi osobou a kočkou.*“ Dále dodává, že bychom měli být velice opatrní. S Muskem souhlasilo mnoho vědců, dokonce jeho názory reflektovaly i názory Stephena Hawkinga. Snažíme se o nasazení výkonných systémů UI, ale možná si neuvědomujeme, co by mohlo následovat, kdybychom se zmýlili či pochybili.⁸⁴

6.1.3 Moje stanovisko

Můj osobní názor v problematice etických aspektů umělé inteligence je stále nejasný. Balancuji na obou stranách, proto zastávám spíše pozici skeptika. Těším se z nových objevů a vymožeností, ale zároveň jsem si vědom nebezpečí, které může UI přinést. Domnívám se, že naše generace bude již tou, kterou tento trend možná okrajově zasáhne. Nemyslím si, že by se jednalo o nějaký převrat, nicméně další generace by se měly mít na pozoru. Co se týče jisté automatizace v mém budoucím pracovním oboru, tak obavy nemám vzhledem k tomu, že bych rád pracoval s digitálními technologiemi právě v oblasti kreativní, proto si naopak myslím, že by mohla být prospěšná. Budoucnost je nejasná, ale věřím, že ještě nějakou dobu budeme těžit z nových objevů a budeme se těšit z toho, co naše technologie dovedou. Souhlasím také s Brynjolfssonem, že přes všechny problémy, které může umělá inteligence přinést v ohledu ekonomickém nebo existenčním, jsme doposud byli svědky růstu blahobytu, bohatství či svobody. Samozřejmě jsem si vědom toho, že za takovými odměnami se skrývá velká zodpovědnost a potencionální následky. To, že se stroje budou čím dál více

⁸⁴ Why Elon Musk fears artificial intelligence - Vox. Vox - Understand the News [online]. Copyright © 2019 [cit. 24.08.2019]. Dostupné z: <https://www.vox.com/future-perfect/2018/11/2/18053418/elon-musk-artificial-intelligence-google-deepmind-openai>

zapojovat do pracovního procesu, by mohlo vést nejen k většímu blahobytu a sebenaplnění, ale po čase i k větší nuditosti, nadbytku a potencionálnímu ohrožení z něj. Proto připouštím i možnost ohrožení lidí. V neposlední řadě souhlasím s Dignumovou, že bychom měli co nejdříve nastavit platný rámec, v němž by měla být zohledněna právě etika pro fungování umělé inteligence v praxi. Co je patrné již nyní, je to, že technologie má na lidstvo ohromný vliv. Řekl bych, že je vesměs pozitivní, což je skvělé, počet omezení klesá, ale zároveň, jak jsem zmínil, se naskýtá velké množství nejednoznačně zodpovězených etických otázek. V dalších dekádách, jak se domnívám, budeme rozhodovat o tom, jakým směrem se svět začne ubírat. Sdílím s Brynjolfssonem tezi, že bychom se měli do budoucna více zabývat tím, co chceme jako jednotlivci a jako společnost. Faktem je, jak popisuje, že naše generace zdědila více příležitostí ke změně než kterákoliv jiná, což je důvod k optimismu, ale o to víc bychom měli být na pozoru a naše rozhodnutí činit rozvážně.⁸⁵

„Technologie není osud. Osud si utváříme samy.“

– Erik Brynjolfsson –

⁸⁵ BRYNJOLFSSON, Erik. *Druhý věk strojů: práce, pokrok a prosperita v éře špičkových technologií*. Přeložil Filip DRLÍK. V Brně: Jan Melvil Publishing, 2015. Pod povrchem. ISBN 978-80-87270-71-4. Str. 244

Závěr

Co říci závěrem? Nastínili jsme, co to je samotný život a jak ho definovat. Zjistili jsme, že vědomí je to, co z nás dělá „pány tvorstva“. Ohromily nás velké příběhy, podívali jsme se, obrazně řečeno, pod kapotu autonomních vozidel společnosti Google, poznali jsme Watsona, počítač s velkým potenciálem společnosti IBM, který porazil lidi – své tvůrce ve vědomostních hrách. Při bližším zamyšlení jsme byli svědky toho, za jak krátkou dobu technologie poznala světlo světa a pokračuje dál. Polemizovali jsme o možném umělém mozku, ale zjistili jsme, že náš lidský mozek je sofistikovaný a velice komplikovaný proces v závislosti na neuronech, že ani v tuto chvíli není zcela možné utvořit kopii k naprogramování. Algoritmy a křemíkové obvody jsou potřebné přísady k dalšímu pokroku technologií. Zákoutí Fuzzy logiky nám podkryly taje algoritmů a procesů, které přispívají k pokroku. Vrátili jsme se zpátky do minulosti a pomocí vět jsme poznali Alana Turinga, který stál za zrodem nové éry strojů, která by se vyvíjela podstatně jinak, kdyby lidé jako Isaac Asimov nestanovili zákony robotiky. V minulosti jsme ještě chvíli zůstali a viděli jsme, jak to bylo s nástupem prvních automatických strojů. Etická otázka otevřela polemiku mé práce a začali jsme smýšlet kupředu. Je morálně správný technologický vývoj? Můžeme u strojů vůbec mluvit o něčem takovém, jako je morálka? Nastavení správného etického rámce pro UI začíná již na samotném začátku. Měli bychom být odpovědni za naši budoucnost. Poté jsme přesunuli pozornost k aktuální otázce, jak technologie napomáhá lidskému bytí. Skýtá mnoho výhod, ale jsou všechny v etické symbióze? Měly by stroje zasahovat do něčeho, co je lidské? Bude budoucnost s roboty obohacenými umělou inteligencí správná? Nakonec jsme se seznámili s názory odborníků a nebylo obtížné zjistit, že v otázce technologické budoucnosti lidstva nejsou zajedno. Podle mého názoru, jak už jsem uvedl, nejde o to, co přinese budoucnost, ale o to, co přineseme budoucnosti my sami.

Použitá literatura:

BRYNJOLFSSON, Erik. *Druhý věk strojů: práce, pokrok a prosperita v éře špičkových technologií*. Přeložil Filip DRLÍK. V Brně: Jan Melvil Publishing, 2015. Pod povrchem. ISBN 978-80-87270-71-4.

LIN, Patrick, Keith ABNEY a George A. BEKEY. *Robot ethics: the ethical and social implications of robotics*. Cambridge, Mass.: MIT Press, c2012. Intelligent robotics and autonomous agents. ISBN 9780262016667.

NILSSON, Nils J. *The quest for artificial intelligence: a history of ideas and achievements*. New York: Cambridge University Press, 2010. ISBN 0521122937.

PEREGRIN, Jaroslav. *Kapitoly z analytické filosofie*. 2. vyd. Praha: Filosofia, 2014. ISBN 9788070074206.

WALLACH, Wendell a Colin ALLEN. *Moral machines: teaching robots right from wrong*. New York: Oxford University Press, 2009. ISBN 0195374045.

WARWICK, Kevin. *Úsvit robotů - soumrak lidstva*. Praha: Vesmír, 1999. Medusa (Vesmír). ISBN 80-85977-16-8.

ZELINKA, Ivan. *Umělá inteligence: hrozba nebo naděje?*. Praha: BEN - technická literatura, 2003. ISBN 80-7300-068-7.

Neuralink: How the Human Brain Will Download Directly from a Computer. Interesting Engineering [online]. Copyright © Copyright 2019 [cit. 15.08.2019]. Dostupné z: <https://interestingengineering.com/neuralink-how-the-human-brain-will-download-directly-from-a-computer>

Artificial intelligence might not threaten the human race, but it could be coming for your job. Tampa Bay, Florida news | Tampa Bay Times/St. Pete Times [online]. Copyright © 2019 All Rights Reserved [cit. 24.08.2019]. Dostupné z: https://www.tampabay.com/news/business/Artificial-intelligence-might-not-threaten-the-human-race-but-it-could-be-coming-for-your-job_171963256

Great Thinkers: Interview With Martin Ford, Futurist, Speaker, A.I. Expert. The Medical Futurist [online]. Copyright © 2019 [cit. 24.08.2019]. Dostupné z: <https://medicalfuturist.com/martin-ford-great-thinkers>

Don't fear artificial intelligence | by Ray Kurzweil | Kurzweil. Kurzweil Accelerating Intelligence [online]. Copyright © 2019 Kurzweil Network [cit. 24.08.2019]. Dostupné z: <https://www.kurzweilai.net/dont-fear-artificial-intelligence-by-ray-kurzweil>

Úvodní stránka | Vláda ČR [online]. Copyright © [cit. 15.09.2019]. Dostupné z: https://www.vlada.cz/assets/evropske-zalezitosti/aktualne/AI-pravne-eticka-zprava-2018_final.pdf

What is AI Training Data? - Cogito Tech LLC - Medium. Medium – a place to read and write big ideas and important stories [online]. Dostupné z: <https://medium.com/@cogitotech/what-is-ai-training-data-f26c27dcc617>

The key ethical question for self-driving cars: are they safe? - Vox. Vox - Understand the News [online]. Copyright © 2019 [cit. 18.09.2019]. Dostupné z: <https://www.vox.com/future-perfect/2018/11/9/18072678/self-driving-cars-philosophy-safety-trolley-problem-mit>

SaVIO: Emoce a jak s nimi pracovat (MUDr. Radkin Honzák) - YouTube. YouTube [online]. Dostupné z: <https://www.youtube.com/watch?v=aedhxvooyPg&t=378s>

Centrum Karla Čapka | Novinky [Etika a autonomní vozidla]. Karel Čapek Center [online]. Copyright © 2018, Centrum Karla Čapka pro studium hodnot ve vědě a technice [cit. 01.11.2019]. Dostupné z: <https://www.cevast.org/cz/news/28-etika-a-autonomni-vozidla>

Tramvajové dilema – Wikipedie. [online]. Dostupné z: https://cs.wikipedia.org/wiki/Tramvajov%C3%A9_dilema

The trolley dilemma in AI. Strongbytes [online]. Dostupné z: <https://strongbytes.ai/the-trolley-dilemma-in-ai/>

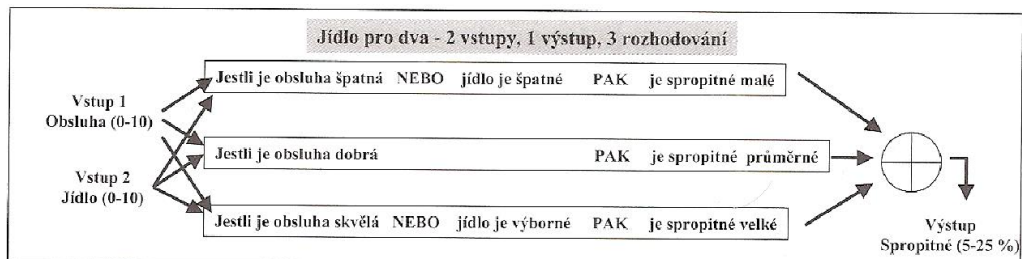
The ART of AI Design —Accountability, Responsibility, Transparency. Design for Values & Responsible Innovation at TU Delft - Portal [online]. Copyright © Copyright 2013 [cit. 02.11.2019]. Dostupné z: <https://www.delftdesignforvalues.nl/2018/the-art-of-ai-accountability-responsibility-transparency/>

Seznam příloh:

Příloha č. 1: Pravidla If–Then pro ohodnocení výše spropitného (obrázek)

Přílohy

Příloha č. 1: Pravidla If-Then pro ohodnocení výše spropitného (obrázek)⁸⁶



Obr. 25 Pravidla If-Then pro ohodnocení výše spropitného

⁸⁶ ZELINKA, Ivan. *Umělá inteligence: hrozba nebo naděje?*. Praha: BEN - technická literatura, 2003. ISBN 80-7300-068-7. Str. 55 – 58