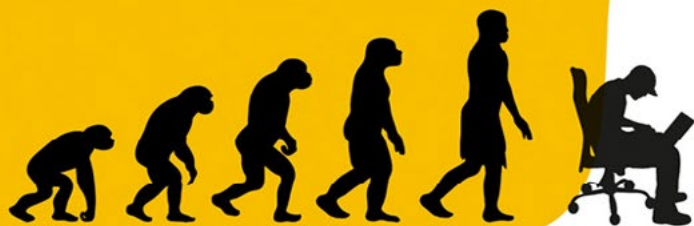


HANNAH FRYOVÁ

Hello World



/**

* Jak zůstat člověkem

* ve světě algoritmů

*/

VYŠEHRAĐ

Hello World

Vyšlo také v tištěné verzi

Objednat můžete na
www.ivysehrad.cz
www.albatrosmedia.cz



Hannah Fryová
Hello World – e-kniha
Copyright © Albatros Media a. s., 2019

Všechna práva vyhrazena.
Žádná část této publikace nesmí být rozšiřována
bez písemného souhlasu majitelů práv.

ALBATROS  **MEDIA**

Hello World

HANNAH
FRYOVÁ

Hello World

/**

* Jak zůstat člověkem

* ve světě algoritmů

*/

HANNAH FRYOVÁ

VYŠEHRAĐ

Pro Marii Fryovou.

Díky za to, že jsi se nenechala odbýt.

Obsah

Poznámka k názvu knihy 9

Úvod 11

Moc 15

Data 35

Spravedlnost 59

Lékařství 89

Automobily 123

Zločin 151

Umění 185

Závěr 207

Poznámky 213

Poděkování 246

Rejstřík 248

Poznámka k názvu knihy

Když mi bylo sedm let, přinesl otec mně a mým sestrám domů dárek. Byl to ZX Spectrum, malý 8-bitový počítač – první vlastní počítač, který jsme kdy měly. Pravděpodobně byl tou dobou, kdy se objevil v našem domě, už pět let zastaralý, ale přestože byl z druhé ruky, okamžitě jsem si pomyslela, že na tomto hezoučkém přístroji je cosi nádherného. Spectrum byl zhruba srovnatelný s Commodore 64 (ačkoli takový měly v našem sousedství jen děti ze skutečně nóbl rodin), ale já jej vždýcky považovala za daleko krásnější potvůrku. Elegantní černý plastový kryt lépe pasoval do ruky a na šedých pryžových klávesách a duhovém proužku běžícím šikmo přes jeden roh bylo cosi přátelského.

Příchod tohoto ZX Spectrum pro mne znamenal začátek nezapomenutelného léta stráveného se starší sestrou na půdě společným hraním naprogramované šibenice či kreslením jednoduchých tvarů za pomoci kódu. Nicméně všechny tyto „pokročilé“ věci přišly až později. Nejprve jsme musely zvládnout základy.

Při pohledu zpět si nedokážu přesně vybavit ten okamžik, kdy jsem napsala svůj vůbec první počítačový program, ale jsem si celkem jistá tím, o co šlo. Byl to stejný jednoduchý program, který jsem později učila všechny své studenty na University College v Londýně; ten samý, který najdete na první stránce prakticky kterékoli základní učebnice počítačové vědy. Existuje totiž tradice dodržovaná všemi těmi, kdo se kdy učili kódovat – jde téměř o přechodový rituál. Vaším prvním

úkolem coby nováčka je naprogramovat počítač tak, aby se na obrazovce rozblíkala slavná fráze:

„HELLO WORLD – AHOJ SVĚTE“

Jedná se o tradici, která se datuje nazpět až do sedmdesátých let, kdy ji Brian Kernighan zařadil coby cvičení do své fenomenálně populární učebnice programování.¹ Tato kniha – a tedy i tato fráze – představovala významný mezník v historii počítačů. Na scénu právě vstoupil mikroprocesor a zvěstoval přeměnu počítačů z podoby, kterou měly v minulosti – neskutečně velké specializované stroje krmené děrnými štítky a páskami – v cosi podobné osobním počítačům, na které jsme zvyklí, s obrazovkou, klávesnicí a blikajícím kurzorem. „Ahoj světe“ znamenalo první okamžik, kdy jste si mohli se svým počítačem pokecat.

O několik let později prozradil Brian Kernighan v rozhovoru pro časopis Forbes, co ho inspirovalo k použití této fráze. Kdysi viděl obrázek zobrazující vajíčko a čerstvě vylihnuté kuřátko pípající po svém narození slova „Ahoj světe!“, a to mu uvízlo v paměti.

Není úplně jasné, kdo je v tomto scénáři považován za kuře: lidé, kteří s rozzářenou tváří triumfálně oznamují svůj vstup do světa programátorství? Nebo samotný počítač probouzející se ze své dřímoty v říši tabulek a textových dokumentů, připravený spojit svou mysl se skutečným světem a plnit příkazy svého nového mistra? Možná oba dva. Ale najisto se jedná o frázi, která sjednocuje všechny programátory a spojuje je s každým počítačem, který byl kdy naprogramován.

A je tu ještě něco, co mám na této frázi ráda – něco, co nikdy nebylo relevantnější či důležitější nežli dnes. Jak počítačové algoritmy stále více ovládají svět a rozhodují o naší budoucnosti, „Ahoj světe“ je připomínkou momentu dialogu mezi člověkem a strojem. Situace, kdy je hranice mezi kontrolujícím a kontrolovaným prakticky neznatelná. Značí počátek partnerství – společnou cestu možností, na níž jeden bez druhého nemůže existovat.

Ve věku počítačů je to pocit, který stojí za to si uchovat v mysli.

Úvod

Každý, kdo někdy navštíví pláž Jones Beach na Long Islandu v New Yorku, musí na své cestě za oceánem projet podél řady mostů. Tyto mosty, primárně postavené kvůli navedení lidí na dálnici a z dálnice, mají jeden neobvyklý rys. Jak se tak vznášejí nad dopravním ruchem, jsou zavěšeny mimořádně nízko, někdy je dělí od asfaltu dokonce pouhých 2,7 metru.

Tento zvláštní design má svůj důvod. Ve dvacátých letech minulého století se Robert Moses, vlivný newyorský městský projektant, horlivě snažil uchovat svůj nově dokončený, cenami ověřený státní park na Jones Beach výhradně pro bílé a bohaté Američany. Protože věděl, že jeho preferovaná klientela bude cestovat na pláž ve svých soukromých automobilech, zatímco lidé z chudých černošských čtvrtí se tam dostanou autobusem, záměrně se snažil omezit jim přístup budováním stovek nízkých mostů přes dálnici. Tak nízkých, aby pod nimi neprojel tři a půl metru vysoký autobus.¹

Rasistické mosty nejsou jedinými neživými objekty, které tiše a skrytě ovládají lidi. V historii najdeme plno příkladů objektů a vynálezů nadaných mocí přesahující deklarovaný účel.² Někdy se jedná o záměrnou a zlovolnou součást jejich designu, ale jindy jde o výsledek neúmyslných opomenutí: pomyslete na nedostatek bezbariérových vstupů pro invalidní vozíky v některých městských oblastech. Jindy se jedná o nezamýšlený důsledek, tak jako v případě mechanizovaných tkacích strojů v devatenáctém století. Byly navrženy tak, aby usnadnily výrobu složitých textilií, jenže dopad, který měly

na mzdy, nezaměstnanost a pracovní podmínky, z nich na-
konec udělal pravděpodobně většího tyрана, než jakým byl
kterýkoli viktoriánský kapitalista.

S moderními vynálezy to není jiné. Stačí se zeptat obyvatel
města Scunthorpe v severní Anglii, kterým bylo blokováno
zakládání účtů u společnosti AOL poté, co tento internetový
gigant vytvořil nový filtr vulgarismů, kterému se jméno města
nějak nezamlouvalo.³ Nebo Chukwuemeky Afigboa, Nigerijce,
který vymyslel automatický dávkovač mýdla na mytí rukou,
jenž spolehlivě uvolnil mýdlo pokaždé, když pod přístroj vložil
ruku jeho bílý přítel, zatímco jeho vlastní tmavou kůži odmítal
rozpoznat.⁴ Nebo Marka Zuckerberga, který by si při psaní
kódu pro Facebook ve svém pokoji v Harvardu v roce 2004
nikdy nebyl představil, že mu jeho výtvar přinese obvinění,
že usnadňuje manipulace s hlasy ve volbách po celém světě.⁵

Za každým takovýmto vynálezem je algoritmus.

Algoritmy, neviditelné kousky kódu, které tvoří ozubená
kola a páky ve věku moderních strojů, dávají světu vše od soci-
álních médií po vyhledávače, od satelitní navigace po systémy
doporučující hudbu, a jsou stejně pevnou součástí naší moder-
ní infrastruktury, jako jí byly a jsou mosty, budovy a továrny.
Najdeme je v našich nemocnicích, soudních síních i automobi-
lech. Používají je policejní síly, supermarkety i filmová studia.
Naučily se, co máme a co nemáme rádi; radí nám, na co se
dívat, co číst a s kým chodit. A po celou tu dobu mají skrytou
moc pomalu a jemně měnit pravidla toho, co to znamená být
člověkem.

V této knize si odhalíme široké spektrum algoritmů, na kte-
ré se stále častěji, ačkoli možná nevědomky, spoléháme. Věnu-
jeme velkou pozornost jejich nárokům, prozkoumáme jejich
nepřiznanou moc a budeme konfrontováni s nezodpovězený-
mi otázkami, které vyvolávají. Setkáme se s algoritmy, které
používá policie k rozhodnutí, kdo má být zatčen, což nás po-
staví před volbu mezi ochranou obětí trestných činů a nevinou
obviněného. Poznáme algoritmy, které používají soudci při
rozhodování o výši trestu odsouzených zločinců, což od nás

vyžaduje určit, jakou povahu by měl mít náš soudní systém. Objevíme algoritmy, které používají lékaři k přehodnocení vlastních diagnóz; algoritmy uvnitř vozidel bez řidiče, které nás nutí definovat principy naší morálky; algoritmy, které se váží na projevy našich emocí; a algoritmy s potenciálem podkopat naši demokracii.

Netvrdím, že algoritmy jsou ze své podstaty špatné. Jak uvidíte na dalších stránkách, existuje mnoho důvodů, proč se na výše uvedené věci dívat pozitivně a s optimismem. Žádný objekt ani algoritmus není sám o sobě dobrý nebo zlý. Jde o to, jak je použit. Systém GPS byl vynalezen kvůli zacílení jaderných střel, ale dnes usnadňuje rozvážku pizzy. Populární hudba, přehrávaná opakovaně stále dokola, byla použita jako mučicí nástroj. A jakkoli je girlanda z květin krásná, kdybych opravdu chtěla, mohla bych vás s ní uškrtit. Udělat si názor na algoritmus znamená pochopit vztah mezi člověkem a strojem. Každý stroj je neoddělitelně spojen s lidmi, kteří ho sestavili a používají.

To znamená, že ve své podstatě je tato kniha o lidech. Je o tom, kdo jsme, kam jdeme, co je pro nás důležité a jak se to proměňuje pod vlivem technologií. Je o našem vztahu k algoritmům, které již existují a které pracují společně s námi, rozšiřují naše schopnosti, opravují naše chyby, řeší naše problémy a během toho všeho vytvářejí problémy nové.

Tato kniha se ptá, zda má algoritmus pro společnost jasný přínos. Hledá odpověď na to, kdy byste měli stroji důvěřovat více než svému vlastnímu úsudku, a kdy byste naopak pokušení přenechat vládu stroji měli odolat. Pojednává o tom, kde jsou limity algoritmů; a také, jak je náročné podívat se do zrcadla a hledat sami sebe. O oddělování zla od dobra a o rozhodování o tom, v jakém světě chceme žít.

Protože budoucnost nepřijde jenom tak. My ji právě tvoříme.



Moc

Garri Kasparov ví zcela přesně, jak vystrašit svého soupeře. Ve čtyřiatřiceti letech byl největším šachistou, jakého kdy viděl svět, s tak strašidelnou reputací, že by odradila každého protivníka. Dokonce měl jeden znepokojující trik, kterým své soupeře vyváděl z rovnováhy. Když vsedě potili krev nad pravděpodobně nejobtížnější hrou svého života, Rus jako by nic zvedl hodinky, které si předtím položil vedle šachovnice, a znovu si je připjal na zápěstí. To byl signál, kterému každý porozuměl – znamenal, že Kasparova hra s tímto soupeřem nudí. Hodinky představovaly pro soupeře pokyn, že je na čase, aby se vzdal. Mohl odmítnout, ale Kasparovovo brzké vítězství bylo tak či tak nevyhnutelné.¹

Jenže když Kasparovovi při slavné partii v květnu 1997 čelil Deep Blue společnosti IBM, byl tento počítač vůči popsané taktice netečný. Výsledek zápasu je všeobecně známý, ale příběh stojící v pozadí toho, jak si Deep Blue zajistil vítězství, je daleko méně docenován. Za tímto symbolickým vítězstvím stroje nad člověkem, které v mnoha směrech znamenalo počátek věku algoritmů, bylo mnohem víc nežli hrubá síla výpočtů. Aby Kasparova porazil, nevnímal jej Deep Blue jednoduše jako vysoce efektivní procesor brilantních šachových tahů, ale jako lidskou bytost.

Na začátku učinili inženýři společnosti IBM skvělé rozhodnutí designovat Deep Blue tak, aby vypadal mnohem nejistější, než ve skutečnosti byl. Během zápasu složeného ze šesti partií měl stroj příležitostně pozdržet svůj tah i po dokončení

výpočtů, někdy dokonce o několik minut. Na Kasparovově straně stolu zpoždění vyvolávala dojem, že se stroj musí probíjovat skrze stále větší a větší množství výpočtů. Zdálo se, že se potvrzuje Kasparovova domněnka, že úspěšně dovedl hru do situace, kdy je počet možností tak nemyslitelný, že Deep Blue nedokáže provést racionální rozhodnutí.² Ten však ve skutečnosti, ačkoli bez hnutí vyčkával, přesně věděl, jak figurami táhnout – to jenom nechával hodiny odtikávat. Byl to podlý trik, ale zafungoval. Už při první hře začaly Kasparova dokonce rozptylovat úvahy, jak schopný vlastně počítač vůbec může být.³

Ačkoli Kasparov první hru vyhrál, během druhé hry se mu Deep Blue skutečně dostal do hlavy. Kasparov se snažil vlákat počítač do pastí a přimět jej, aby postupoval a sebral několik figur, zatímco by sám zaujal postavení – po několika tazích – kdy by si uvolnil dámu a vyrazil do útoku.⁴ Každý pozorující šachový expert očekával, že počítač na návnadu skočí, jak se domníval Kasparov. Jenže Deep Blue nějakým způsobem vytušil levárnu. K úžasu Kasparova si počítač uvědomil, co velmistr plánuje, a táhl tak, aby zablokoval jeho dámu a tím pohřbil všechny šance na vítězství člověka.⁵

Kasparov byl očividně zděšen. Špatný úsudek, co by mohl počítač udělat, mu podrazil nohy. Několik dní po zápase popsal v rozhovoru Deep Blue slovy „náhle hrál na okamžik jako bůh“.⁶ O mnoho let později, když se zamýšlel nad tím, jak se v té době cítil, napsal, že „udělal chybu, když usoudil, že tahy, které byly překvapivé, zahrál-li je počítač, byly také objektivně silné tahy“.⁷ Ať tak či onak, genialita algoritmu triumfovala. Jeho pochopení lidské mysli a lidské omylnosti útočilo a přemáhalo lidského génia.

Sklíčený Kasparov druhou hru raději vzdal, než aby bojoval o remízu. Od této chvíle začala jeho sebedůvěra brát za své. Třetí, čtvrtá a pátá hra skončily remízou. Při šesté partii byl Kasparov poražen. Turnaj skončil výsledkem 3,5 Deep Blue ke Kasparovovým 2,5.

Byla to zvláštní porážka. Kasparov by byl více než schopný vymanit se z oněch pozic na hrací desce, ale podcenil

schopnost algoritmu a následně se jím nechal zastrašit. „Byl jsem natolik ohromen hrou Deep Blue,“ napsal v roce 2017 o osudném turnaji. „Byl jsem tak znepokojen představami, čeho by mohl být schopný, že jsem si nevšímal toho, jak se dostávám do potíží vlastní špatnou hrou, nikoli proto, že by on hrál dobře.“⁸

Jak v této knize opakovaně uvidíme, očekávání jsou důležitá. Příběh Deep Blue, který porazil velmistra, ukazuje, že moc algoritmu se neomezuje na to, co je obsaženo v řádcích jeho kódu. Porozumění vlastním nedostatkům a slabostem – stejně jako nedostatkům stroje – je klíčem k tomu, abychom si udrželi věci pod kontrolou.

Ale když to nedokázal pochopit někdo jako Kasparov, jakou naději máme my ostatní? Na následujících stránkách uvidíme, jak algoritmy pronikly prakticky do všech aspektů moderního života – od zdraví a zločinu po dopravu a politiku. A ještě při tom všem nějak zvládneme jimi pohrdat, bát se jich a žasnout nad jejich schopnostmi zároveň. Konečným výsledkem je, že nemáme příliš velké povědomí o tom, kolik moci jim postupujeme, nebo zda jsme nenechali věci zajít příliš daleko.

Zpátky k základům

Ještě než se k tomu všemu dostaneme, možná stojí za to se nakrátko zastavit a položit si otázku, co vlastně znamená slovo „algoritmus“. Je to pojem, který – ačkoli se často používá – obvykle žádnou skutečnou informaci nevyjadřuje. Částečně proto, že ono slovo samotné je dosti vágní. Oficiálně je definován následujícím způsobem:⁹

Algoritmus (podstatné jméno): Posloupnost kroků vedoucích k řešení problému či dosažení určitého závěru, obvykle prováděná počítačem.

A je to tady. Algoritmus je jednoduše sérií logických instrukcí, které říkají, jak od začátku do konce splnit úkol. Při této široké definici se mezi algoritmy počítá i recept na pečení dortu. Stejně jako seznam odboček, který můžete napsat ztracenému cizinci. Návodů z IKEA, videa na YouTube popisující řešení

závad, dokonce i svépomocné motivační knihy – jakýkoli nezávislý seznam pokynů pro dosažení určitého definovaného cíle může být teoreticky popsán coby algoritmus.

Nicméně takto se tento pojem zrovna moc nepoužívá. Obvykle se algoritmem míní něco poněkud specifičtějšího. Stále se jedná o seznam postupných kroků, ale tyto algoritmy mají téměř vždy matematický obsah. Provádějí sekvenci matematických operací – pomocí rovnic, aritmetiky, algebry, infinitesimálního počtu, logiky a pravděpodobnosti – a překládají ji do počítačového kódu. Jsou krmeny daty z reálného světa, dostávají na práci zadání a úkoly a prokousávají se výpočty, aby dosáhly svého cíle. Představují to, co činí z počítačové vědy skutečnou vědu, a během tohoto procesu spoluzapřičinily mnohé z nejužasnějších úspěchů moderní doby, jež přinesly stroje.

Existuje téměř nespočetné množství různých algoritmů. Každý má své vlastní cíle, své vlastní idiosynkrazie, své chytré triky a stinné stránky a neexistuje žádný konsensus na tom, jak je nejlépe rozčlenit. Ovšem celkově vzato může být užitečné vzít v potaz jejich využití v reálném světě, podle něhož je lze dělit do čtyř kategorií:¹⁰

1. Stanovení priorit: vytvoření seznamu Vyhledávač Google predikuje webovou stránku, kterou hledáte, zařazením jednoho výsledku hledání nad jiný. Filmový portál Netflix vám navrhuje, které filmy byste možná rádi zhlédli. Satelitní navigace TomTom pro vás vybírá nejrychlejší trasu. Všichni používají matematické postupy, aby do nesmírného množství možností vnesli nějaký řád. Deep Blue byl také v podstatě algoritmem pro stanovení priorit, přezkoumával všechny možné tahy po šachovnici a počítal, který by mu dal co největší šanci na vítězství.

2. Klasifikace: výběr kategorie Jakmile jsem se přiblížil k třicítce, začaly mě na Facebooku bombardovat reklamy na prsteny s diamantem. A když jsem se nakonec vdala, pronásledovaly mě při surfování internetem inzeráty na těhotenské testy. Za tyto

kratochvíle jsem mohla poděkovat klasifikačním algoritmům. Tyto algoritmy, které milují inzerenti, fungují v zákulisí a klasifikují vás na základě vašich vlastností coby možné zájemce o určité věci. (Mohou mít pravdu, bezpochyby, ale stejně je to poněkud nepříjemné, když se na vašem notebooku v průběhu schůzky neustále objevují reklamy na ovulační testy.)

Existují algoritmy, které dokáží automaticky vytrít a odstranit nevhodný obsah na YouTube, algoritmy, které označí fotografie z vaší dovolené, a algoritmy, které dokáží skenovat váš rukopis a rozpoznat každý znak na stránce jako písmeno abecedy.

3. Asociace: hledání odkazů Asociace vytváří hledání a označování vztahů mezi věcmi. Tvoří jádro seznamovacích algoritmů, jako je OKCupid, které pátrají po tom, co přihlášení sdílejí, a na základě svých zjištění jim navrhnou partnery. Doporučující mechanismus Amazonu používá podobný princip, když asociuje vaše zájmy se zájmy dřívějších zákazníků. Právě toto vedlo k zajímavému nákupnímu návrhu, který dostal uživatel Redditu Kerbobotat poté, co si na Amazonu koupil baseballovou pátku: „Možná byste měl zájem i o tuto lyžařskou kuklu.“¹¹

4. Filtrování: vyčlenění toho, co je důležité Algoritmy často potřebují odstranit některé informace, aby se zaměřily na to, co je důležité, tedy aby oddělily signál od šumu. Někdy to dělají doslova: algoritmy rozpoznávající řeč, například ty, které umožňují fungování systémů Siri, Alexa a Cortana, potřebují nejdříve oddělit váš hlas od šumu v pozadí, nežli mohou začít pracovat na dešifrování toho, co říkáte. Jindy to dělají obrazně: Facebook a Twitter filtrují příběhy, které se týkají vašich známých zájmů, a nabízejí vám vlastní jedinečný přísun.



Převážná většina algoritmů je vytvořena tak, aby prováděla kombinace výše uvedených principů. Vezměte si například UberPool, který spojuje potenciální cestující s jinými lidmi, kteří míří stejným směrem. S ohledem na váš výchozí bod a cíl

se musí probrat možnými trasami vedoucími k vám domů, vyhledat spojení s dalšími uživateli cestujícími stejným směrem a zvolit jednu skupinu, do níž vás zařadí – a při tom všem upřednostnit trasy představující pro řidiče co nejméně otáček, aby byla jízda co možná nejefektivnější.¹²

Takže tohle všechno algoritmy dokážou. A nyní, jak to dělají? Takže znovu, přestože možností je prakticky nekonečno, existuje způsob, jak z nich něco vydestilovat. Postupy, které algoritmy používají, si můžete představit jako nahrubo rozřazené pod dvě klíčová paradigmatata, s nimiž oběma se v této knize setkáme.

Algoritmy založené na pravidlech Prvním typem jsou algoritmy založené na pravidlech. Jejich instrukce jsou sestaveny člověkem a jsou přímé a jednoznačné. Můžete si představit, že tyto algoritmy se podobají receptu na dort. První krok: udělej tohle. Druhý krok: pokud se stalo tamto, udělej toto. To ale neznamená, že jsou tyto algoritmy jednoduché – v rámci tohoto paradigmatu je spousta prostoru pro vytváření silných programů.

Algoritmy založené na strojovém učení Druhý typ je inspirován schopností živých bytostí učit se. Hledáte-li analogii, zamyslete se nad tím, jak byste mohli naučit psa, aby si s vámi plácnul. Nemusíte vytvořit přesný seznam pokynů a sdělit je psovi. Vše, co jako trenér potřebujete, je mít v mysli jasný cíl, co chcete, aby pes vykonal, a nějaký způsob, jak jej odměnit, když udělá správnou věc. Zkrátka jde o upevnění správného chování, ignorování špatného a poskytnutí dostatečného prostoru k procvičování toho, co se má udělat. Ekvivalentem psa je algoritmus založený na strojovém učení spadající pod široký pojem umělé inteligence (angl. artificial intelligence, zkrat. AI). Dejte stroji data, cíl a zpětnou vazbu, když je na správné cestě – a nechte jej, aby sám našel nejlepší způsob, jak cíle dosáhnout.

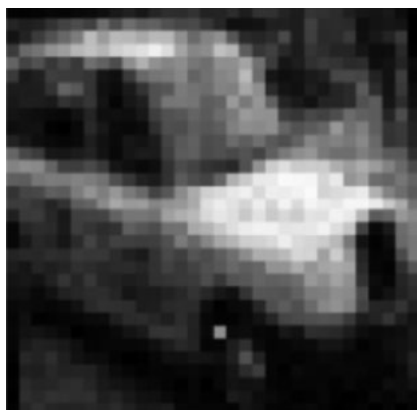
Oba dva typy mají své výhody a nevýhody. Protože algoritmy založené na pravidlech obsahují instrukce napsané lidmi, je snadné jim porozumět. Teoreticky do nich kdokoli může nahlédnout a sledovat logiku toho, co probíhá uvnitř.¹³ Ale jejich dar je také jejich prokletím. Algoritmy založené na pravidlech budou pracovat pouze na úlohách, pro jejichž řešení dokáží lidé sepsat instrukce.

Naopak algoritmy založené na strojovém učení se nedávno ukázaly být pozoruhodně účinné při řešení problémů, kde by sepsání seznamu instrukcí nebylo možné. Dokáží rozpoznat objekty na obrázcích, porozumět vyřčeným slovům a přeložit je z jednoho jazyka do druhého – věci, s nimiž algoritmy založené na pravidlech budou vždy bojovat. Nevýhodou je, že pokud místo sebe necháte najít řešení stroj, cesta, kterou k němu došel, často nedává lidskému pozorovateli příliš smysl. Co se skrývá uvnitř, zůstává tajemstvím dokonce i pro nejchytřejší žijící programátory.

Vezměte si například úkol rozpoznat obraz. Skupina japonských vědců nedávno ukázala, jak podivný se může člověku zdát způsob, jímž se na svět dívá algoritmus. Možná jste se setkali s optickým klamem, kdy nemůžete s jistotou říct, jestli se díváte na obraz vázy nebo dvou tváří (pokud ne, v poznámkách v zadní části knihy je uveden příklad).¹⁴ Počítač je na tom stejně. Vědecký tým ukázal, že změna jediného pixelu na obrázku v oblasti předního kola stačila k tomu, aby algoritmus založený na strojovém učení zavrhl názor, že je na fotografii auto, a začal se domnívat, že jde o fotku psa.¹⁵

Pro někoho je představa algoritmu pracujícího bez výslovných instrukcí receptem na katastrofu. Jak můžeme ovládat něco, čemu nerozumíme? Co když schopnosti vnímavých, super-inteligentních strojů přesáhnou schopnosti jejich stvořitelů? Jak zajistíme, že AI, které nerozumíme a kterou nedokážeme ovládat, nezačne pracovat proti nám?

To jsou všechno zajímavé hypotetické otázky a knihy věnované hrozbě nadcházející AI apokalypsy rozhodně nejsou nedostatkovým zbožím. Omlouvám se, pokud zklamu vaše



„Auto-pes“, použito
se svolením Danila
Vasconcellose Vargase
z Kjúšúské univerzity
ve Fukuoce (Japonsko).

očekávání, ale tato kniha není jednou z nich. Přestože AI je na světě již nějakou dobu, je stále „inteligentní“ pouze v nejužším slova smyslu. Pravděpodobně by bylo mnohem účelnější přemýšlet o tom, jakou jsme prošli revolucí ve výpočetní statistice, nežli revolucí v inteligenci. Víím, že by to znělo mnohem méně sexy (pokud nejste skutečný fanda statistiky), ale jde o mnohem přesnější popis toho, jak se věci skutečně mají.

Prozatím se obavy ze zlovolnosti AI dosti podobají obavám z přelidnění Marsu.* Možná se jednoho dne dostaneme až do bodu, kdy počítačová inteligence překoná tu lidskou, ale ta chvíle je stále velmi daleko. Upřímně řečeno nás čeká ještě dost dlouhá cesta k vytvoření umělé inteligence na úrovni ježka. Zatím se nikdo nedokázal dostat nad úroveň červa.**

Navíc všechen humbuk okolo AI odvádí pozornost od mnohem naléhavějších starostí a – myslím si – mnohem

* Jedná se o parafrázi komentáře počítačového vědce a pionýra v oblasti algoritmů založených na strojovém učení Andrewa Ng z rozhovoru, který dal v roce 2015. Podívejte se na události v technice, „GPU Technology Conference 2015 day 3: What's Next in Deep Learning“, YouTube, 20. listopadu 2015, <https://www.youtube.com/watch?v=qP9TOX8T-kl>.

** Právě simulace mozku červa je cílem mezinárodního vědeckého projektu OpenWorm. Doufají, že uměle reprodukuji síť 302 neuronů nacházejících se v „mozku“ červa *C. elegans*. Pro srovnání, my lidé máme asi 100 000 000 000 neuronů. Viz web projektu OpenWorm: <http://openworm.org/>.

zajímavějších příběhů. Zapomeňte na okamžik na všudypřítomné stroje s umělou inteligencí a obraťte své myšlenky od daleké, vzdálené budoucnosti k tady a teď – protože již existují algoritmy se svobodnou vůlí, které autonomně činí rozhodnutí. Rozhodují o délce pobytu za mřížemi, o léčbě pacientů s rakovinou a o tom, co dělat při autonehodě. Rozhodnutí proměňující naše životy za nás už činí na každém rohu.

Otázka zní, pokud jim dáme všechnu tu moc – zaslouží si naši důvěru?

Slepá víra

Neděle 22. března 2009 nebyla pro Roberta Jonese dobrým dnem. Právě navštívil přítele a vracel se zpátky do hezkého městečka Todmorden v západním Yorkshiru, když si ve svém BMW povšiml světla palivové kontrolky. Zbývalo mu něco málo přes 11 km, aby našel čerpací stanici, nežli mu dojde benzín, pročež si začal dělat mírné starosti. Naštěstí to vypadalo, že jeho GPS našla zkratku – a poslala ho na úzkou cestu po straně údolí.

Robert následoval pokyny přístroje, ale jak jel, silnice stále stoupala a zužovala se. Po několika kilometrech se změnila na blátivou stezku, která se zdála stěží schůdná i pro koně, natož pro auta. Ale Roberta to nerozhodilo. Běžně ujel týdně osm tisíc km kvůli své práci, a tak věděl, jak to chodí za volantem. Navíc si myslel, že „nemá žádný důvod nedůvěřovat satelitní navigaci TomTom“.¹⁶

Jen o chvíli později by ten, kdo by náhodou vzhlédl z údolí, býval viděl, jak se předeek Robertova BMW vynořil nad okrajem útesu a před stometrovým pádem jej zachránil jenom křehký dřevěný plot na okraji, do nějž auto narazilo.

Nakonec bylo zapotřebí traktoru a tří čtyřkolek, aby Robertovo auto odtáhly z místa, kde jej opustil. Později toho roku, když se objevil u soudu kvůli obvinění z bezohledného řízení, Robert přiznal, že ho ani nenapadlo neuposlechnout pokyny navigace. „Stále mi tvrdila, že ta stezka je silnice,“