

## Umělá inteligence a co má společného s psychologií

---

**Kristýna Študlarová, Šimon Kendra,  
Michal Kosina, Libor Sup**

Psychologický ústav FFMU, Brno

[info@psychologon.cz](mailto:info@psychologon.cz)



*Tato otázka se vlastně zdá velmi jednoduchá, protože psychologie a inteligence jsou prakticky příbuzné pojmy. Ale není umělá inteligence něco úplně jiného? Není to jen hraní si s počítačovými programy a sestavování algoritmů, kterým obyčejný smrtelník, IT studiem nepolíbený, nemůže nikdy porozumět? Pojďme se na to podívat podrobněji.*

*Článek se zabývá otázkou propojení psychologie s výpočetní technikou a především umělou inteligencí. Podává příklady paralel mezi oběma odvětvími, jejich vzájemného propojení, a dále klade otázky ohledně etické korektnosti výzkumu umělé inteligence a možných směrů jejího vývoje, přičemž čerpá jak z odborných publikací, tak jejich kritik z řad vědců i široké veřejnosti.*

*A po jeho dočtení třeba budete vědět, zda je umělá inteligence něco pro Vás.*

*This question seems to be very simple, because psychology and intelligence are practically related terms. But isn't AI something different? Isn't it just playing with computer programs and compilation algorithms, which ordinary mortal, uneducated in IT studies, can't understand? Let's look at it more detail.*

*This article deals with the connection between psychology and computing technologies, especially artificial intelligence. It offers examples of parallels between these two fields, their mutual connectedness, and further focuses on questions regarding ethical correctness of artificial*

*intelligence-based research and possible courses of its development, while draws information from both scientific publications and the critics among both experts and common public.*

*And after finish the article, you will maybe know if AI is something for you.*

*Klíčová slova: umělá inteligence, robot, psychologie, mozek, počítač, Alan Turing*

*Keywords: artificial intelligence, robot, psychology, brain, computer, Alan Turing*

---

## Válka, pokrok a Turing

Ačkoliv může mnoho lidí považovat psychologii za čistě humanitní vědu a informatiku za vědu ryze technickou, mají tyto dva obory společného víc, než se široké veřejnosti na první pohled mohlo dosud zdát. Kapitola vědecké psychologie se začala psát v roce 1879 v laboratoři Wilhelma Wundta v Lipsku (Plháková, 2006). Psychologie jako taková se však vyvíjela již od dob antiky. Rozvoj výpočetní techniky přišel mnohem později, lze datovat od dob druhé světové války, jež přinesla díky průsaku technologií z vojenské do civilní sféry obrovský technický pokrok. Jeden z prvních programovatelných počítačů MARK I tedy mohl sloužit k nastavení jaderné bomby v projektu Manhattan. Námořnictvo USA ho používalo až do roku 1959 k výpočtu balistických křivek. Násobení mu trvalo 3-5 sekund a sčítání zvládl za 0,3 sekundy. (Zelený & Mannová, 2006). Během padesáti let udělala počítačová technika neskutečně velký skok kupředu. Například už v roce 1997 se podařilo počítači od firmy IBM porazit v šachu mistra světa Garriho Kasparova.

Cesty psychologie a výpočetní techniky se setkaly mimo jiné v roce 1950, kdy se Alan Turing začal zabývat otázkami, jestli může být počítačový program inteligentní, v Turingově podání, zda může člověka obalamutit a namluvit mu, že je také člověkem (Sternberg, 2002). Toho roku Turing publikoval ve vědeckém časopise *Mind* článek s názvem *Computing Machinery and Intelligence*, ve kterém představil svoji dnes již dobře známou myšlenku imitační hry.

*„...hrají tři lidé, muž (A), žena (B) a moderátor (C), který může být jakéhokoliv pohlaví. Moderátor zůstává v místnosti oddělené od ostatních dvou hráčů. Úkolem moderátora je určit, kdo z obou hráčů je muž a kdo žena. Zná je pod označením X a Y a na konci hry řekne buď „X je A a Y je B“, nebo „X je B a Y je A“. Moderátor se může ptát A a B takto:*

*C: Mohlo by mi X sdělit délku svých vlasů?*

*Předpokládejme nyní, že X je A, takže A musí odpovědět. Úkolem A je mást C tak, aby provedl chybnou identifikaci. Může proto odpovědět:*

*„Mám dlouhé blond vlasy.“*

*Aby moderátorovi nepomáhal tón hlasu, musí být odpovědi napsány, nejlépe na psacím stroji. Ideální by bylo dálkopisné spojení mezi oběma místnostmi. Otázky a odpovědi mohou být případně předávány pomocí prostředníka. Úkolem třetího hráče (B) je napovídat moderátorovi. Pravděpodobně nejlepší strategií pro ženu je odpovídat správně. Může ke své odpovědi třeba dodat: „ Jsem žena, nevěř mu!“, ale ani to mnoho neznamená, protože muž může říct něco podobného. Nyní se můžeme zeptat, co se stane, když stroj převezme úlohu A v této hře? Rozhodne se pak moderátor stejně často špatně, jako kdyby hru hrál muž a žena? Tyto otázky nahradily původní otázku: „Mohou stroje myslet?“ (Leavitt, 2007)*

V tomto případě by pak měl moderátor za úkol rozlišit stroj a člověka. Pokud by moderátor udělal chybu a spletl by si stroj a člověka, prošel by stroj testem a mohl by se podle Turinga považovat za inteligentní.

Proti tomuto typu testu bylo vyneseno mnoho námitek, např. argument čínského pokoje pocházející od filosofa Johna Searla z roku 1980. Jeho námitka je založena na tezi, že schopnost odpovídat není dostatečný důkaz porozumění, což je to co očekáváme od umělé inteligence. Počítač podle Searla může své myšlení jen mechanicky napodobovat (Koval, 2008).

Tuto myšlenku vysvětlil smyšleným pokojem, ve kterém jsou knihy se všemi smysluplnými čínskými větami. Do tohoto pokoje je umístěn člověk, který čínsky neumí, ale má schopnost na základě zadaného textu najít odpověď. Zvenčí jsou mu posílány otázky a on díky své schopnosti najde příslušný text a pouhým opsáním části kontextu zvládne odpovědět. Člověk venku, který posílá otázky, by tedy mohl nabýt dojmu, že člověk uvnitř umí čínsky. Avšak ten pouze pracuje se symboly, jejichž význam nezná, tzn. jeho práci by mohl dělat i nemyslicí stroj (Lokvencová, 2014).

Nakonec i Turing sám přiznal: „...člověk si může být jist tím, že stroj myslí, pouze tehdy, jestliže se sám stane strojem a uvědomí si při tom své myšlení... Podobně je podle tohoto názoru potřeba být člověkem, aby někdo věděl, že člověk myslí“ (Leavitt, 2007).

V roce 2014 na univerzitě v Readingu byla Kevinem Wervickem pořádaná soutěž, kde počítačové programy podstupovaly Turingův test. Soutěž vyhrál chatterbot jménem Eugene Goostman. Tento program vytvořil tým z Ruska vedený Vladimírem Veselovem. Pokořil totiž bájnou hranici jedné třetiny popletených lidí. Celých 30% moderátorů považovala Eugena za člověka, čímž splnil Turingův test. Podle Kevina Warvicka tento výsledek jen upozorňuje na fakt, že počítače umějí stále lépe napodobit člověka, což se lehce dá zneužít v oblasti kyberzločinu (Kasík, 2014). Aplikace pod jménem Eugene Goostman je nyní v nabídce na Google Play, ale nemá moc dobrá hodnocení. Chat s Eugenem je také inzerován na stránkách chatbots.org, ty jsou zase pro změnu zpoplatněné.

## **Superpočítače jako realita**

Počítače a roboti zatím vlastnosti jako myšlení podobné lidskému a sebeuvědomění obsáhnout nedovedou. Ve fungování počítače a mozku lze však najít mnoho podobností. Podívejme se například na stavbu neuronu. Skládá se z těla (které bychom mohli chápat jako malý počítač), mnoha dendritů (tyto mohou v našem schématu představovat vstupy), které vedou signál do těla buňky, a jeden axon, který tvoří jediný výstup, obdobně jako u počítačů (Šmarda, 2007). Díky velmi složitému vzájemnému propojení vytvářejí neurony komplikovanou síť, která je schopna vést informaci obdobnou počítačové – 1 nebo 0 (neboli zapnuto – vypnuto, případně elektrický impuls – bez elektrického impulsu). Do těla neuronu tedy přijde mnoho signálů, které jsou zpracovány a opět ve formě 1 nebo 0 odeslány skrz axon dále.

O vytvoření superpočítače, který by se alespoň z části mohl vypadat jako lidský mozek, se snaží mnoho vědců např. projekt SyNAPSE- mikročip rozdělený do několika paměťových modulů podobně jako v mozku. Počítač, který by se chtěl mozku vyrovnat výkonem a počtem operací, by musel mít kapacitu operační paměti alespoň 3 584 terabytů a měl by být schopen zvládnout 38 000 trilionů operací za sekundu. Jeden z nejvýkonnějších superpočítačů současnosti, BlueGene od IBM, dokáže udělat 92 trilionů operací za sekundu a jeho operační paměť je 8 terabytů (Nývlt, 2009).

Gordon E. Moore – spoluzakladatel firmy Intel – zformuloval tzv. Mooreův zákon, podle něhož se každých 18 měsíců zdvojnásobí výkonnost výpočetní techniky (Moore, 1965), a bývá často vnímán jako měřítko inovace a pokroku. Podle předpokladu tohoto zákona/myšlenky, roste pokrok výpočetní techniky exponenciálním způsobem, což v praxi znamená, že rozdíl mezi

současnými technologiemi a těmi v roce 2025 bude podobný jako mezi dneškem a technickými výtobky roku 1965.

Další paralelou ve fungování lidského myšlení a funkcí počítače jsou počítačové jazyky, které mají za úkol překládat uživatelský záměr (požadavky člověka) do jazyka počítače. V programech lze mimo jiné pracovat s proměnnými, které si většinou musíme předem definovat. Definuji-li si proměnnou  $a$ , lze vytvořit pomyslnou paralelu s nějakým pojmem z lidského světa. Například pojem strom v sobě ukrývá mnoho vlastností (barva, velikost, náš vztah k němu, atd.), ale člověk ho má v paměti uložen jako strom. Obdobně jako program do své paměti uloží proměnnou  $a$ , která má určité vlastnosti, ale pro zjednodušení při další práci lze pracovat pouze s danou proměnnou, netřeba znovu zadávat její vlastnosti. Tato představa je velmi zjednodušená a slouží pouze k představě podobnosti.

Nicméně dnešní počítače jsou stále jen kusy železa, kterým člověk říká, například formou programovacích jazyků, jakou činnost mají vykonat. Snahou současných výzkumů je vytvořit program, který bude vykazovat určitou inteligenci. Inteligencí můžeme myslet, mimo Turingův model, schopnost adaptovat se prostředí nebo schopnost samostatně řešit komplexní problémy, které program nikdy předtím neřešil, a rozhodovat se spontánně.

Vědci z Rensselaer Artificial Intelligence and Reasoning (RAIR) Laboratory tvrdí, že jeden z jejich NAO Botů byl schopen sebeuvědomění. Uspěl v testu, kdy dva ze tří robotů dostali takzvanou „němou pilulku“, která jim měla zabránit v mluvě, a poslední z nich dostal pilulku, jež jej neměla nijak ovlivnit. Poté byli zapnuti a dotázáni, kterou pilulku dostali. Jeden z robotů poté odpověděl „nevím,“ a vzápětí odpověděl „promiňte, už vím. Byl jsem schopen dokázat, že jsem němou pilulku nedostal“ (Bringsjord a spol., 2015). Kritici z řad odborníků i laiků namítají, že naprogramovat robota k opravě sebe sama ještě nedosvědčuje jeho sebeuvědomění. Někteří dokonce tvrdí, že člověk nikdy nebude schopen stvořit robota, jenž by přijal emoce za své. Samotné umělé algoritmy nemohou obsáhnout tak komplexní procesy, jako náš organismus. (Maguire, Anathaswamy, 2014) V lidském chování hrají velkou úlohu hormony, chemické, biologické procesy. Podle James-Langovy teorie vznikají mnohé emoce na základně psychosomatických projevů, ne naopak (Dalglish, 2004). To znamená, že podle této teorie by měl umět robot s umělou inteligencí vyvolat emoce i na základě dějů uvnitř sebe samého, ne jen jako reakci na vnější podněty. Roboti budou jistě do několika let schopni nerozpoznatelně napodobit

lidskou nonverbální komunikaci, naši mimiku, gesta. Tento umělý projev emocí však není roven reálným emocím.

## **Máme se těšit, nebo obávat?**

Nemůžeme popřít, že technologie je velice prospěšná, pokud ji správně využíváme. Co kdyby ale s příchodem umělé inteligence, tak, jak si ji vyobrazuje například kinematografie, technologie fungovala nezávisle na nás? Představme si, že jednou budou sestaveny stroje jako z filmů Já, robot, kdy se umělá inteligence snažila lidi ochránit natolik, až omezila jejich svobodu, nebo Ex Machina, kde se umělá inteligence dokázala stát tak lidskou, že zabila svého stvořitele. Takové scénáře technologického pokroku nezní právě lákavě.

Je nesporným faktem, že technologický vývoj se v posledním desetiletí posunul o velký kus kupředu, a můžeme tak bez nadsázky hovořit o technologických skocích namísto kroků. Stáváme se například svědky toho, jak byl lidem za pomoci moderní technologie navrácen zrak či sluch, pomocí různých implantátů napodobujících funkce mozku a smyslových orgánů (Vrecková, 2016). Otázkou je, do jaké míry těmto lidem pomáhají užité technologie interpretovat okolní svět, resp. zdali tyto umělé součásti lidského těla nepropouštějí do oblasti lidské percepce jen ty části reality, které samy považují za „správné“.

Mimo spojení člověka a stroje na úrovni lékařské nás zanedlouho čeká například představení robota Roxxy firmy TrueCompanion (Ellyatt, 2015), který má potenciál v příštích letech radikálně změnit sexuální průmysl, a mohl by tak silně zasáhnout i oblast mezilidských vztahů. Jedná se v podstatě o robota zkonstruovaného tak, aby se co nejvíce podobal skutečné ženě, a mohl sloužit k sexuálnímu uspokojení. Má mluvit, má být dostupný v řadě charakterových provedení, a dokonce má tento robot simulovat i tlukot srdce. Podle expertky na robotickou etiku a signatářky petice za zákaz prodeje a vývoje těchto robotů doktorky Kathleen Richardsonové z DeMonthfortovy univerzity v britském Leicester tito roboti budou mít „škodlivý vliv na vztahy mezi muži a ženami, dospělými a dětmi, mezi muži navzájem, mezi ženami navzájem,“ jak uvedla 15. září v rozhovoru pro BBC.

S posunem vzhledu objektů, pomocí nichž dochází k sexuálnímu uspokojení, blíže lidské podobě, je nasnadě otázka, do jaké míry budou tito roboti narušovat lidské vztahy, vyvolávat

pocity žárlivosti ze stran partnerů, nebo naopak pomáhat špatně se socializujícím jedincům k vytvoření si jakéhosi placebo vztahu s živou osobou. Sexuální roboti by podle některých mohli pomoci i s redukcí trestných činů páchaných pedofily, a to prostřednictvím produkce dětských robotů, kdy by objekt zájmu zůstal stejný, přičemž by nedocházelo k újmě dětí.

„Neexistují žádné předpoklady pro to, že s sebou toto nutně přinese pozitivní výsledky – pouze věřím, že to stojí za kontrolované prozkoumání, aby bylo případně možné poskytnout společnosti lepší ochranu před recidivitou sexuálních delikventů,“ řekl pro časopis Forbes vedoucí laboratoře mobilních robotů v georgijském Technologickém institutu Ronald Arkin (Kashmir, 2014). Tato myšlenka s sebou přináší mnohé problémy, především etické, přičemž z druhého pohledu objekt uspokojení pedofila bude stále reprezentovat dítě. Je tedy nasnadě otázka, zda by toto řešení nakonec nepodněcovalo takového člověka ke zneužití opravdového dítěte ještě více.

## **Těžká otázka pro soubor algoritmů**

Existuje spousta odvětví, v kterých by takový robot mohl nalézt své uplatnění, jako například řadoví pracovníci v továrnách, ale i lékaři, kuchaři nebo policie, resp. robotičtí policejní příslušníci. Pokud už ale povolíme uzdu své fantasii, a představíme si situaci, ve které by i živý člověk zažíval dilema řešení dané situace, je na místě otázka, jak by stroj bez něčeho takového, jako je zakořeněná lidská etika a morálka tuto situaci vyřešil. Jako jednoduchý příklad nám může posloužit následující situace:

„Nákladní auto bez řidiče sjíždí z kopce po silnici, na níž si hraje skupina dětí, a to rychlostí, která by nepochybně zapříčinila mnohá vážná, ne-li smrtelná zranění. Stojíme u páky napojené na koleje, pomocí níž lze odklonit dráhu nákladního auta na silnici vedlejší, na které však pracuje skupina dělníků, a devastující výsledek situace bude pravděpodobně stejný.“ (Englert, Siebert, Ziegler, 2014)

Jedná se tedy o situaci, v které člověk musí volit mezi dvěma zly – tedy neexistuje způsob, jak rozhodnout správně, pouze méně špatně. Člověk uvržený do takového scénáře, by mohl zvolit možnost na základě svého charakteru, svobodné vůle a zažité morálce. Jak by řešil

nenaprogramovanou bezprecedentní situaci robot fungující pouze na základě souhry algoritmů, zůstává prozatím velkou neznámou.

Pokud se ale přece jenom podaří vytvořit stroj, který bude člověkem vedle vlastní autonomie vybaven i takovými schopnostmi, jako je morálka, či spontánní jednání, znamenalo by to, že je schopen na základě vlastní vůle měnit způsob, než jakým byl původně naprogramován. Na odpověď na otázku, jestli by pro člověka vytvoření takového stroje bylo pomocí či naopak předzvěstí zkázy, si nejspíš ještě pár let počkáme, nicméně – jak napovídá např. i Stephen Hawking (Cellan-Jones, 2014) – taková doba může přijít dříve, než bychom se byli schopni připravit.

#### Zdroje:

- Dalgleish, T. (2004). The emotional brain. *Nature Reviews Neuroscience*.
- Anathaswamy, A. (2014). Why robots can never be conscious. *New Scientist*. (222), 12.
- Bringsjord, S., Licato, J., Govindarajulu, N. S., Ghosh, R., & Sen, A. (2015). Real robots that pass human tests of self-consciousness. In *Robot and Human Interactive Communication (RO-MAN), 2015 24th IEEE International Symposium on* (pp. 498-504).
- Eysenck, M. W., & Keane, M. T. (2008). *Kognitivní psychologie*. Praha: Academia.
- Koval, D. (2008). Příležitosti a hrozby umělé inteligence. Brno: Masarykova univerzita: Filosofická fakulta, Ústav české literatury a knihovnictví, Kabinet informačních studií a knihovnictví.
- Leavitt, D. (2007). *Muž, který věděl příliš mnoho: Alan Turing a první počítač*. Praha: Dokořán.
- Lokvencová, I. (2014). Argument čínského pokoje-včera, dnes a zítra. Brno: Masarykova univerzita: Filosofická fakulta, Ústav Hudební vědy, Teorie interaktivních médií.
- Plháková, A. (2006). *Dějiny psychologie*. Praha: Grada Publishing.
- Sternberg, R. J. (2002). *Kognitivní psychologie*. Praha: Portál.
- Šmarda, J. (2007). *Biologie pro psychology a pedagogy*. Praha: Portál.
- Zelený, J., & Mannová, B. (2006). *Historie výpočetní techniky*. Praha: Scientia.

#### Internetové zdroje:



- Kasík, P. (6. 9. 2014). *Počítač přesvědčivě napodobil člověka. Hrál si na puberťáka*. Načteno z iDNES.cz ([http://technet.idnes.cz/uspesny-turinguv-test-07g-/software.aspx?c=A140609\\_181110\\_software\\_pka](http://technet.idnes.cz/uspesny-turinguv-test-07g-/software.aspx?c=A140609_181110_software_pka))
- Moore, G. (1965). *Gordon Moore, 1965, article*. Načteno z <http://web.eng.fiu.edu/> ([http://web.eng.fiu.edu/npala/eee6397ex/gordon\\_moore\\_1965\\_article.pdf](http://web.eng.fiu.edu/npala/eee6397ex/gordon_moore_1965_article.pdf))
- Nývlt, V. (13. 3. 2009). *Technet.idnes.cz*. Načteno z Technet.cz: [technet.idnes.cz](http://technet.idnes.cz) ([http://technet.idnes.cz/lidsky-mozek-je-mnohosetkrat-vykonnejsi-nez-nejlepsi-superpocitac-11o-/tec\\_technika.aspx?c=A090313\\_171006\\_tec\\_technika\\_NYV](http://technet.idnes.cz/lidsky-mozek-je-mnohosetkrat-vykonnejsi-nez-nejlepsi-superpocitac-11o-/tec_technika.aspx?c=A090313_171006_tec_technika_NYV))
- Vrecková, H. (16. 12. 2015). *Implantát dokáže nahradit funkci středního ucha. Mladá fronta E15*. Načteno z: [vtm.e15.cz](http://vtm.e15.cz) (<http://vtm.e15.cz/implantat-dokaze-nahradit-funkci-stredniho-ucha>)
- (20. 1. 2015) *Intelligent machines: Call for a ban on robots designed as sex toys. BBC*. Načteno z: [bbc.com](http://www.bbc.com) (<http://www.bbc.com/news/technology-34118482>)
- Ellyatt, H. (15.12. 2015). *Campaign launched against „harmful“ sex robots. CNBC*. Načteno z: [www.cnbc.com](http://www.cnbc.com) (<http://www.cnbc.com/2015/09/15/sex-robots-campaign.html>)
- Kashmir, H. (15.12. 2015). *Are child sex-robots inevitable? Forbes*. Načteno z: [forbes.com](http://www.forbes.com) (<http://www.forbes.com/sites/kashmirhill/2014/07/14/are-child-sex-robots-inevitable/#752518ab2ddf>)
- Cellan-Jones, R. (2014). *Stephen Hawking warns artificial intelligence could end mankind. BBC News, 2*. (<http://www.bbc.com/news/technology-30290540>)

Zdroj obrázku:

<http://www.micronobal.com/wp-content/uploads/2014/07/o-ARTIFICIAL-INTELLIGENCE-facebook-3-940x564.jpg> (14. 11. 2015)