

*Semestrální práce
z předmětu KKY/HKUI*

Turingův test a problémy s porozuměním umělé inteligence

“Problém je v tom, že zatím nemůžeme obecně určit, které výpočetní postupy lze nazvat inteligentní. Rozumíme některým principům inteligence, ale velká část této disciplíny je nám stále skryta. Inteligence v rámci této vědy se proto chápe pouze jako výpočetní složka schopnosti dosáhnout cílů ve světě.”

— John McCarthy, konference na dartmouthské univerzitě, 1956



Dodnes zůstává jednou z nejzáhadnějších otázek v oblasti počítačových systémů a algoritmů otázka: “Jsou stroje schopné myslet?”. Předpokládá se, že zakladatelem této úvahy je slavný anglický matematik, specialista v oblasti informatiky a kryptografie, Alan Turing.

“A computer would deserve to be called intelligent if it could deceive a human into believing that it was human.”

– Alan Turing (1912 – 1954)



V roce 1950 byl zformulován a publikován empirický (založený na zkušenostech, pozorování, měření, experimentu) test ve článku “Computing Machinery and Intelligence” v časopisu “Mind”. Nazvali ho Turingův test. Téměř 70 let po prvním zveřejnění prošel postup drobnými úpravami, ale podstata testu zůstává stejná.

Turingův test

Standardní interpretace tohoto testu je: “Člověk komunikuje s jedním počítačem a jednou osobou. Na základě odpovědí na otázky musí určit, s kým mluví: s osobou nebo s počítačovým programem. Úkolem počítačového programu je uvést člověka v omyl a nutit ho, aby se rozhodl špatně.”

Jinými slovy, myšlenkou testu je, aby byl počítač schopný přesvědčit člověka o své lidskosti. Pokud člověk nemůže s jistotou říci, kdo je kdo, má se za to, že stroj prošel zkouškou. Zdá se, že jedním z jeho hlavních nedostatků je, že ve skutečnosti je úkolem stroje zmást a oklamat člověka. Teoreticky by mělo platit, že robot nebo program, který prošel Turingovým testem, musí dobře napodobovat, opakovat jednání osoby v komunikaci. V současné době se věří, že přísnou verzí tohoto testu zatím žádný stroj neprošel. Někdy však programy dokážou přesvědčit účastníka o tom, že mluví s osobou. Výstižný příklad: program který se jmenuje Eugene Goostman (“13letý chlapec z Oděsy”), který v roce 2014 přesvědčil 33% soudců, že není stroj, který s nimi komunikuje a odpovědi, které dává, patří osobě.

Jeho výsledky se však nepovažují za úplný test, protože to nebyl dospělý člověk a soudci poukázali na všechny podezřelé chyby a na skutečnost, že Eugene Goostman byl ještě dítě.

Kritika a experiment “Chinese Room”

Existují dva pohledy na otázku přítomnosti rozumu ve strojích, které se nazývají hypotézy Silné a Slabé umělé inteligence.

- Pojem Silná umělá inteligence představil John Searle a vyznačuje se tím, že stroje nebudou jen vzorem lidského rozumu, ale skutečností, že oni sami budou mít svůj vlastní intelekt na stejné úrovni jako lidé.
- Naopak zastánci Slabé UI upřednostňují pohlížení na program pouze jako na nástroj umožňující řešit určité problémy, které nevyžadují celé spektrum lidských kognitivních schopností.

Ve svém myšlenkovém experimentu “Chinese Room”, John Searle ukazuje, že ani procházení Turingového testu nemusí být dostatečným kritériem pro to, aby stroj měl skutečný proces myšlení. Jinými slovy, cílem tohoto experimentu je vyvrácení hypotézy tzv. Silné umělé inteligence a kritiku Turingova testu. Tento filosofický argument je stále jedním z nejvíce diskutovaných v oblasti kognitivní vědy.

Popis experimentu *Chinese Room*: “Představte si, že osoba, která nezná čínštinu, by byla umístěna v oddělené místnosti, ve které jsou kostky s hieroglyfy. Člověk nezná hieroglyfy, ale má instrukci v angličtině říkající, která krychle by měla být odebrána a která krychle by měla být vydána v reakci na určité požadavky přicházející zvnějšku. Pokud člověk použije kostku správně, dostaneme na výstupu odpověď v čínštině. To znamená, že Číňané, kteří tento experiment pozorují, uvidí, že odpovědi jsou obecně smysluplné. Ale sám člověk v této místnosti nikdy neví, co řekl a co udělal.” V důsledku je instrukce druh počítačového algoritmu a osoba v místnosti provádí algoritmus stejným způsobem jako počítač.

John Searle dochází k závěru, že ačkoli program může projít Turingovým testem, nemusí docházet k pochopení jazyka uvnitř systému, což znamená, že Turingův test není adekvátním testem mentálních schopností.

Tento myšlenkový experiment vyvolal v akademické obci intenzivní kritiku. Proti Searlovmu argumentu byly předloženy různé důvody. Experiment byl publikován v roce 1980 ve článku “Minds, Brains, and Programs” časopisu “The Behavioral and Brain Sciences”. Ještě před publikací experiment vyvolal diskuzi, proto článek obsahoval nejen Searleho argument, ale také námitky 27 výzkumníků v oblasti kognitivních věd.

Mnoho vědců se shoduje, že jediná odpověď na otázku, co je umělá inteligence neexistuje a ve filozofii je otázka povahy a postavení lidské inteligence stále nevyřešena.

“Je tedy možné vytvořit něco, co ještě nevíme, jak by to mělo fungovat?”

Další úspěchy Alana Turinga

- V roce 1936 Alanem Turingem byl navržen abstraktní výpočetní stroj pro formalizování konceptu algoritmu. “Turing machine”, jak se jmenoval ten stroj, lze považovat za univerzální počítačový model, který je také považován za zakladatele počítačového zpracování a je stále používán v mnoha teoretických a praktických studiích.
- Na počátku roku 1940 vypracoval dešifrovací stroj “Bombe”, který umožňoval čtení Luftwaffe zpráv kódu “Enigma”. Principem “Bombe” bylo třídění podle možných možností šifrovacího klíče a pokusů o dešifrování textu, pokud byla známá část textu nebo struktury dešifrované zprávy. Bombe vydávala během operace tikající zvuk, za co dostala své jméno. Turingova “bomba” byla nastartována 18. března 1940 v Bletchley.
O šest měsíců později se podařilo prolomit více odolnější šifru Kriegsmarine kódu „Lorenz“. Později, v roce 1943, Turing hmatatelně přispěl k vytvoření lepšího dešifrovacího počítačového stroje “Coloss” – jeden z prvních DPS na světě, který se používá pro stejné účely.
- V roce 1952 Alan Turing publikoval práci nazvanou “Chemické základy morfogeneze”, kde byl poprvé matematicky popsán proces samoorganizace hmoty.

Zdroje

<http://www.raai.org/library/books/sirl/ai.htm>

<https://habr.com/ru/post/69758/>

https://aif.ru/dontknows/eternal/cto_takoe_test_tyuringa_i_pochemu_ego_tak_slozhno_proyti

<http://cogprints.org/7150/1/10.1.1.83.5248.pdf>

https://en.wikipedia.org/wiki/Chinese_room

http://chinapads.ru/c/s/tyuring_alan_-_nauchnyie_dostizheniya_i_otkryitiya

<https://vikent.ru/author/941/>

https://en.wikipedia.org/wiki/Turing_machine

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

https://ru.wikipedia.org/wiki/%D0%A2%D1%8C%D1%8E%D1%80%D0%B8%D0%BD%D0%B3_%D0%90%D0%BB%D0%B0%D0%BD

Alan Turing, «Computing Machinery and Intelligence», *Mind*, vol. LIX, no. 236, October 1950, pp. 433—460