

# *Chatboti – historie, nejznámější chatboti a jak pracují*



PB016  
Jakub Konečný

# *Osnova*

- ◆ Historie
- ◆ Nejstarší chatboti
- ◆ Nejúspěšnější chatboti
- ◆ Jak fungují chatboti?

# *Pojem chatbot*

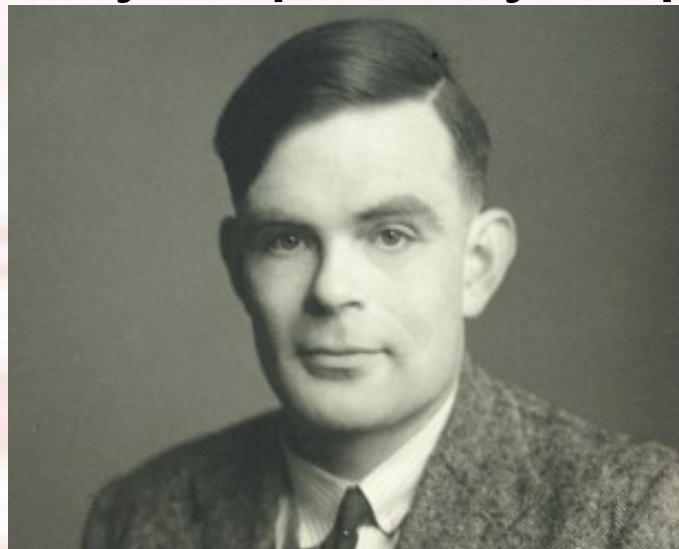
Termínem chatterbot nebo také chatbot označujeme počítačový program, jenž simuluje psanou konverzaci s člověkem, a to v přirozeném jazyce.

Praktické využití:

- Virtuální asistenti
- Výzkumný nástroj

# *Alan Turing*

- ◆ Autor myšlenky tzv. Turingova testu
- ◆ Myšlenka: za inteligentní můžeme stroj považovat tehdy, když nejsme schopni odlišit jeho výstup od výstupu člověka.



# *Loebner Prize*

- ◆ Založena v roce 1990
- ◆ Hugh Loebner
- ◆ Hlavní cenou 100 000 dolarů pro program nerozlišitelný od člověka
- ◆ 6 000 dolarů pro nejlidštější program



# *Nejstarší chatboti*

## Eliza

- ◆ 1966, Joseph Weizenbaum
- ◆ Rogerovská terapie, předstírá psychologa
- ◆ Jazyk Basic
- ◆ Eliza efekt

## Parry

- ◆ 1972, Kenneth Colby
- ◆ Pacient s paranoiou
- ◆ Napojení na mafii, přístup policie

# *Nejúspěšnější chatboti*

Albert One

- ◆ Robby Garner, vítěz testu z let 1998 a 1999
- ◆ V roce 2001 zapsán do Guinnessovy knihy rekordů

A.L.I.C.E.

- ◆ Richard Wallace, vítěz testu z let 2000, 2001 a 2004
- ◆ Psán v jazyce JAVA (postupně přepisován do různých jazyků)



# *Rosette (Suzette)*

- ◆ Bruce Wilcox, vítěz z let 2010 a 2011
- ◆ Databáze stovek koncepcí
- ◆ Téma identifikuje pomocí klíčového slova
- ◆ SQL
- ◆ Pokud nezná odpověď, zahraje rozhovor do autu, nebo předstírá psychické narušení





# ***Jak fungují chatboti?***

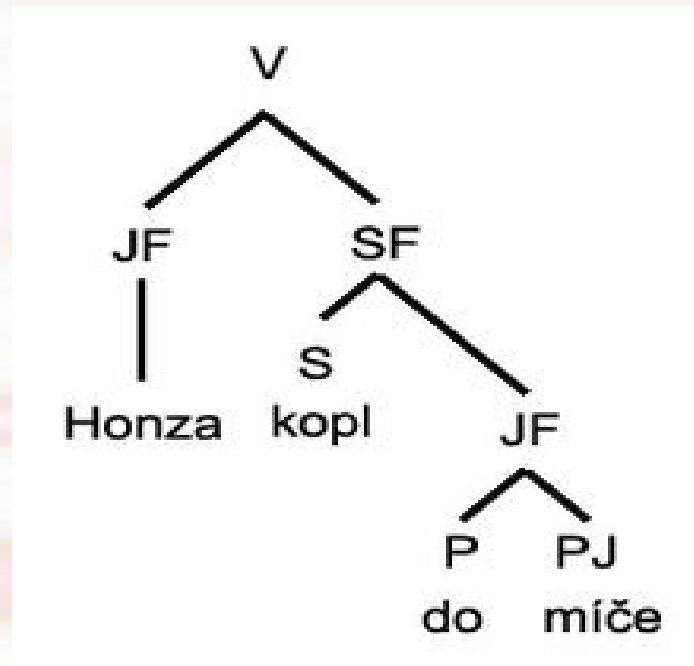
Několik možností, jak zpracovat vstup:

- ◆ Tzv. krmení
- ◆ Hledá klíčová slova
- ◆ Zpracování pomocí regulárních výrazů
- ◆ Bere v úvahu i předchozí konverzaci
- ◆ Provede syntaktickou analýzu

Obvykle se používá kombinace možností.

# Syntaktická analýza

- ◆ Podoba derivačního stromu
- ◆ Velmi omezené využití



# ***Učení, Chatboti bez umělé intelligence***

- ◆ Schopnost učit se (chatbot si sám zvětšuje databázi)

Chatboti bez umělé intelligence se připojí k internetu a použijí vyhledávač

- ◆ Pokud klíčové slovo nerozpozná, odpoví, že nerozumí
- ◆ Pokud máme více klíčových slov, musíme rozhodnout, podle jakého vygenerovat odpověď = keyword ranking

# *Keyword ranking*

- ◆ Musela by být přesná shoda klíčového slova se slovem v databázi
- ◆ Je žádoucí si text předpřipravit
- ◆ Český jazyk velmi těžký na přípravu

**KEYWORD  
RANKING**

# Levenshteinova vzdálenost

- ◆ Podobnostní vyhledávání
- ◆ Určuje podobnost mezi 2 řetězci znaků
- ◆ Důležité si určit číslo
- ◆ Metodu využívá i Google



Děkuji za pozornost.

