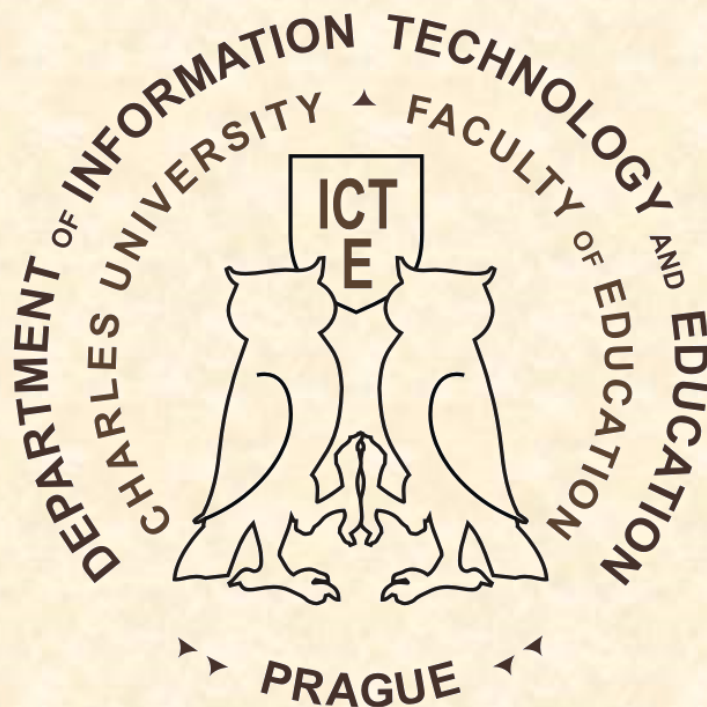




Univerzita Karlova, Pedagogická fakulta
Katedra informačních technologií a technické výchovy

Umělá inteligence (UI/AI)



Intelligence

- Počet neuronů/synapsí v mozku? (mozek jako nejsložitější struktura ve Vesmíru?).
- Inteligentní systém? - učení, logické/algorithmické myšlení, sebereflexe, vědomí, symbolické myšlení, sebeuvědomování, schopnost abstrakce, intuice, altruismus, empatie, emoce, rozum, fantazie, kreativita, puzení, p(r)ožitek,...
- Tvořivá činnost, umění - namalovat obraz, složit symfonii, napsat román/báseň, představivost,...?
- Poznání sebe sama (slon, delfín, lidoop, havran,...), použití nástrojů?
- Individua vyššího řádu (=obligátní společenstva) – včely, mravenci, termity...?
- Společný pohyb hejna ryb, ptáků, hmyzu...? (=> např. skupina dronů)
- Dispozice pro myšlení, učení a adaptaci (přizpůsobení se prostředí)?
- Převážně dědičná.
- Sociální (dá se naučit), emoční,...
- Lze testovat (IQ).

Projevy inteligentních systémů

- Vnímání okolí (reálného světa) s porozuměním, rozpoznávání objektů a situací.
- Vnímání lze zpracovávat.
- Učení.
- Zobecnění poznatků.
- Vytváření nových pojmů.
- Využívání analogií.
- Stanovování nových cílů.
- Uvědomění si sebe sama/prožívání; nutkání.

Projevy inteligentních systémů

- Ne každý systém schopný rozpoznávání je možno označit za inteligentní.
- Triviální rozpoznávání na základě např. měření hran těles nevyžaduje inteligenci.
- UI se vyznačuje:
 - Vysokou měrou začlenění zkoumaného objektu do okolí.
 - Zkoumáním možných variant interpretace.

Umělá inteligence (UI)

- Artificial intelligence (AI); původně např. „výpočtová racionalita (1956)“ (odlišení od kybernetiky).
- Obor informatiky zabývající se tvorbou strojů vykazujících známky „inteligentního chování“.
- Problematika - uvažování, znalostí, plánování, učení, zpracovávání přirozeného jazyka/komunikace, vnímání a schopnost se pohybovat či manipulovat s předměty, rozpoznávání, analýza vztahů, tvorba modelů, účelná rozhodnutí, předvídání důsledků,...
- Snaha o dosažení obecné/univerzální umělé inteligence (x „jednoúčelová“ inteligence); nároky na UI.
- Turingův test (stačí 30%?, 2018 - Google Duplex (kadeřník))/reverzní Turingův test, Argument čínského pokoje; CleverBot.

Umělá inteligence

- Silná UI („univerzální“) – uvědomění si sebe sama, schopnost porozumění; počítač - simulace x (lidské) vědomí? - [Mariin pokoj](#) (v místnosti počítač, venku člověk).
- Slabá UI („jednorázová“, „jednoúčelová“) – např. řízení automobilu.
- UI se neunaví, rychlá, přesná, nezapomíná...
- Problémy
 - Ironie, vágnost (filtr poznání).
 - Neumí vysvětlovat/zdůvodnit rozhodnutí.
(provede selekci a předloží člověku alternativy nebo zdůvodnění, na základě kterých člověk rozhodne X nebude člověk schopen reálně proces rozhodnutí ověřit)
 - Chápání kontextu; souvztažnost vět.
 - Kulturní kontext lidí.
 - Kreativita/tvořivá činnost ([Midjourney](#), [DALL-E](#)), tvůrčí činnost (výzkum), humor?
 - [Tramvajové dilema](#) (Trolley problem); člověk - předsudky, diskriminace, rasismus, pudy, (ne)odpovědnost, pomsta, (po)city, záměr,...

Co zkoumat?

- Systémy UI nepoužívají pro nalezení řešení hrubou sílu (Brute Force, Backtracking); např. prohledávání alternativ.
- UI pojme a posoudí najednou mnohem více faktorů, mnohem rychleji a ve větších souvislostech než člověk.
- Zkoumají pouze řešení, která „mají smysl“.
 - Na základě algoritmu.
 - Na základě zkušenosti.
 - Šachista nezkoumá pohyby všech figur ve hře.
- Ideálně systém hodnotí pouze varianty vedoucí k řešení.
 - Vybírá podle „efektivity“ řešení (cena, poč. kroků, energie, čas,...).

Základ UI

- Většina systémů UI je založena na využití znalostí o úloze (znalostní báze).
- Znalosti musí být správně pochopeny a použity.

REPREZENTACE ZNALOSTÍ
je základním problémem
systémů umělé inteligence

UI dnes

- Technologie, zabývá se návrhem **racionálních agentů** (agent - vnímá prostředí pomocí senzorů a zpětně mění prostředí pomocí akčních prvků); racionální agent maximalizuje očekávanou míru užitku. Peníze jako univerzální míra výkonu...? Člověk = predikovatelně iracionální (víme, jakou provede hloupost).
- **Slabá umělá intelligence (Weak AI)** (současnost) - vypadají, že dělají, co mají.
- **Silná umělá intelligence (Strong AI)** - myšlení stroje.
- **Všeobecná umělá intelligence (General AI)** - stroj řeší libovolný problém, univerzální stroj ("řešič problémů"); člověk nevybírá algoritmus, kdyžtak se ho AI naučí.
- **Úzká umělá intelligence (Narrow AI)** - oblast, kde je AI (většinou) lepší než člověk (hry, rozpoznávání obrázků,...), dnes běžné.
 - Generativní umělá intelligence - vytváří něco nového a jedinečného, např. text (jazykový model), obrázky, videa,...

Problémy řešené UI

- Řešení úloh.
- Dokazování matematických vět.
- Porozumění přirozenému jazyku.
- Expertní systémy.
- Vnímání prostředí.
- Lidmi nezvládnutelné objemy dat.
- ...

Př. Návrh rotoru pro vrtulník.

Přístupy k řešení

- Neuronové sítě („křehké“).
- Evoluční programování.
 - Evoluční/Genetické algoritmy.
- Expertní systémy – pomocí příznaků zužování místa, kde by mohla být chyba (např. diagnostika u automobilu, medicína).
- Dobývání znalostí (Data Mining).
- Strojové učení.

Realizace

- SW systémy.
- Automaty.
- Adaptivní systémy (Neuronové sítě).
- Simulace - důležité eliminovat chyby a nepřesnosti v simulovaném prostředí => při tréninku vzniká/povzbuzuje nežádoucí chování – využívá chyby => v praxi poté nepoužitelné (např. procházení zdí).

Řešení úloh

- Úloha je popsána:
 - Výchozím stavem.
 - Cílovým stavem.
 - Přípustnými kroky.
- Cílem je přejít z výchozího stavu do cílového pomocí přípustných kroků.
- Programování robotů/automatů.

Porozumění přirozenému jazyku

- Problematika „souvislostí“.
- Lidé si rozumí, pokud mají podobný model světa.
 - Věta je fragmentem modelu osoby sdělující.
 - Informace zapadá do modelu osoby přijímající.
- Krátká věta může díky inferenci (odvozování) nést velkou informaci.
- „ASK SGT STAR“; Facebook, Google; asistentky Apple Siri, MS Cortana, Amazon Alexa, Google Assistant,...; chatovací boti (ChatGPT), helpdesk, call centrum, banky apod.

Expertní (znalostní) systémy

- Využívají znalostí reálných osob.
 - Znalosti lidského experta do formálních pravidel if - then, $A \Rightarrow B, \dots$
- Řeší úzce specializované úlohy (jednouúčelová AI).
- Analytické úlohy.
 - Diagnostika (medicína, technika).
- Syntetizující úlohy.
 - Plánování postupů, výběr konstrukčních prvků.
- Hledání souvislostí (Data Mining) – např. banky.
 - např. zjišťování spolehlivosti klientů.

Robotika

➤ 3 zákony robotiky (Isaac Asimov, 1942).

- 1) Robot nesmí ublížit člověku nebo svou nečinností dopustit, aby bylo člověku ublíženo.
- 2) Robot musí uposlechnout příkazů člověka, kromě případů, kdy jsou tyto příkazy v rozporu s prvním zákonem.
- 3) Robot musí chránit sám sebe před poškozením, kromě případů, kdy je tato ochrana v rozporu s prvním nebo druhým zákonem.

Postupem času další modifikace... (0. zákon apod.).

Robot(y/i) univerzální x jednoúčelové, humanoidní (android).

Např. [Boston Dynamics](#), [DARPA Challenge](#),...; problematika chůze.

Úrovně autonomního řízení

➤ 6 (5) úrovní

- 0 - jakýkoliv vůz, který ovládá stoprocentně řidič („hloupé auto“; bezpečnost - ABS, ESP, ASR, indikace námrazy...).
- 1 – asistenti aktivně zasahují do řízení (adaptivní tempomat – zrychlení/zpomalení/odstup, aktivní udržení v jízdním pruhu), parkování.
- 2 – sám zastavit a rozjet, projet zatáčky i křižovatky, analyzovat situaci kolem sebe, číst značky a předvídat chování ostatních účastníků, parkování. Řidič ruce na volant, pozornost (např. TESLA).

Úrovně autonomního řízení

- 3 – prvky umělé inteligence, vyžadována řidičova pozornost; za určitých podmínek autopilot (např. TESLA).
- 4 – vysoká automatizace (např. Google car), „plná“ autonomie, bez řidičova zásahu, mohou chybět tradiční ovládací prvky; lze řídit; výjimečné situace.
- 5 – plná automatizace, bez volantu.

Problémy – počasí (hustý déšť, sníh, mlha/smog, oslnění), čtení dopravního značení, nezpevněný povrch/náledí/břečka, zanesené/zamrznuté senzory; UI je potřeba natrénovat = naučit; zpracování dat např. grafické karty (NVIDIA); hybridní provoz (člověk x UI). Steven Mahan - slepý "řidič" samořiditelného Google auta.

UI dnes - problémy

- Hackerské útoky (ovládnutí, zneužití,...).
- Zabezpečení komunikace/datových přenosů (odposlech).
- Zodpovědnost za chybové chování automatu s UI - návrhář, výrobce, programátor, prodejce, majitel,...? Chyba SW, aktualizace?
- Řešení reklamací, pojištění, zodpovědnost při nehodě?
- UI jako sluha/nadvláda nebo partner/spolupráce?
- Zvýšení nezaměstnanosti (Průmysl 4.0, IoT,...)?
- Možné otázky filozofické, teologické, etické/mravní, morální,...?
- Problematika využití ve válečných konfliktech?
- „Selský“ rozum, instinkt, odhad chování ostatních, riskování, nutkání k činnostem/uspokojení, improvizace (lidé – strach ze smrti),...?
- Není nutné být neomylný, jen lepší než člověk? Nestandardní situace?

UI dnes - problémy

- Můžeme vůbec zjistit, jestli má počítač vědomí?; a je to vůbec důležité?
- Měli bychom vyvíjet AI (všeobecnou, silnou; etické otázky)? - fake informace (obrana => princip Occamovy břitvy - nejjednodušší řešení je většinou správné). Dnes: "lež = alternativní pravda".
- Stroje by neměly být černé skříňky, měly by zdůvodnit svoje rozhodnutí (Proč?).
- Křehkost systémů (mírná změna na vstupu, systém zklame, protože nemyslí; např. rozpoznávání značek – překreslená dopravní značka).
- Bias - systém bude něco preferovat, zaujatost.
- Počítač „papouškuje“ to, co do něj vloží lidé - jaká data se vloží, takové budou výsledky.
- Rozhodování samostatně nebo jenom doporučení pro člověka? (obchodní přihazovací systémy, lékař,...)?
- Vojenské využití - autonomní zbraně?
- Trolley problem (v různých částech světa jiný přístup!, např. zabít krávu v Indii).