

Nepředpokládáme, že čtenář zná hru go, ale snad někdy hrál šachy nebo dámu. Pokud tyto hry nehrajeme impulzivně, vybíráme své tahy z několika možností a při výběru se řídíme možnými odpověďmi soupeře a pozicemi, které po těchto soupeřových tazích vzniknou. Podle situace a svých schopností takto pozici „pročítáme“ o více tahů dopředu. Při tvorbě robota vývojáři postupují podobně. Buďto zvažují všechny možné tahy v pozici, nebo vytvoří a naprogramují metodu, která pomůže vybrat jen některé z nich a vytvoří takto pozice o několik tahů dopředu. Pak je třeba navrhnout a naprogramovat metodu, která ohodnotí šanci na výhru každé z těchto pozic. Následuje pro programátora robota rutinní práce – ohodnocení pozic propaguje od závěrečných k předchozím tak, že vybírá pro hráče tah pro něj nejvýhodnější.

Šachový robot Deep Blue, který porazil v roce 1997 světového šampiona Garryho Kasparova, v každé pozici probíral prakticky všechny tahy, kterých je několik desítek. Ve hře go je situace odlišná, jednak je možných tahů v pozici několik set, dále při hodnocení závěrečné pozice hráči do značné míry využívají intuici založenou na svých zkušenostech. Proto roboti hrající go dlouho zaostávali dokonce za silnějšími amatéry, a když v roce 2016 robot AlphaGo porazil nejsilnějšího hráče světa, Korejce I Se-tola, byla komunita hráčů go ohromena. Jaký průlom nastal v oblasti umělé inteligence, že vedl k takovému úspěchu?

Vývojáři britské firmy DeepMind použili ke hře dvě umělé neuronové sítě,

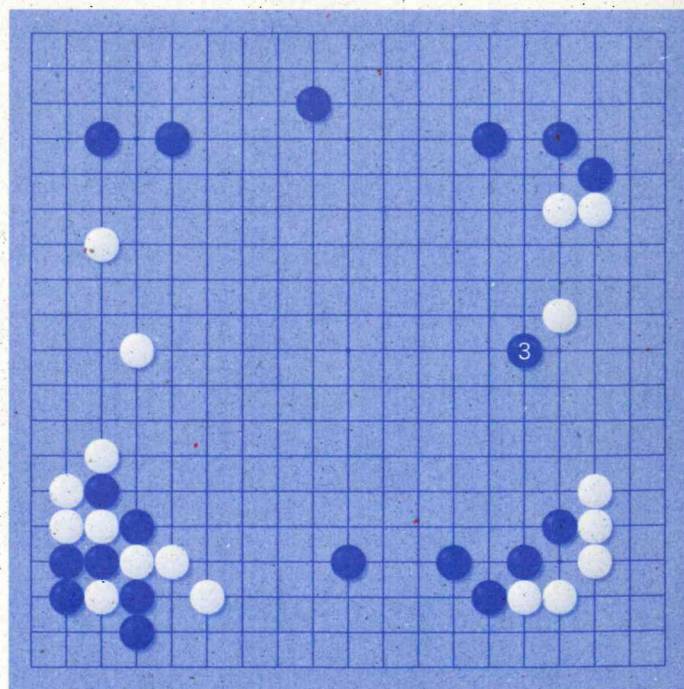
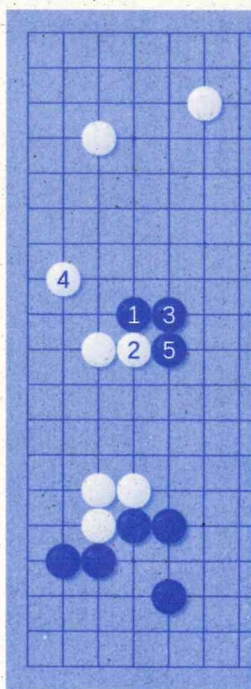
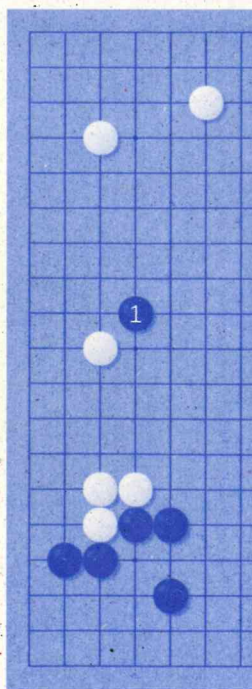
první doporučovala tahy a druhá určila, s jakou pravděpodobností povede daná pozice k vítězství. Umělá neuronová síť je software navržený podle vzoru lidského mozku a obsahuje na softwarové úrovni nervové buňky a jejich propojení, přitom jejich intenzita je na začátku daná náhodně a mění se během tréninkového procesu. Jeden způsob trénování je učení s učitelem (*supervised learning*), při něm je na vstupu otázka (například na tah v dané pozici ve hře go) a učitelem daná odpověď na otázku. Neuronová síť nejdříve vygeneruje svoji odpověď a dále na základě této a učitelovy odpovědi drobně přenastaví intenzitu svých propojení s cílem co nejvíce přiblížit svoji odpověď učitelově. První verze AlphaGo byla trénována právě tímto způsobem.

Stejný princip se používá k jiným úkolům, zmiňme jako příklad rozpoznávání obrazů. Na vstupu předkládáme obrázky a například odpověď na otázku: je na daném obrázku lidská tvář? Dalším příkladem mohou být data o žadateli o půjčku v bance a otázka: bude tento žadatel rád ně splácet? Tréninkem získané neuronové sítě pak v prvním případě pomohou při zpracování a analýze obrazů i videí a v druhém případě poradí při rozhodování o poskytnutí půjčky.

Jinou možností trénování neuronových sítí je tzv. posilující učení (*reinforcement learning*), které se podobá metodě pokus-omyl, kterou s oblibou používají malé děti. Zde robot interaguje s prostředím (například sehraje několik partií go), dostává zpětnou vazbu (na konci partie se dozví,

zda vyhrál, nebo prohrál) a tuto zkušenost použije k přenastavení spojení mezi svými neurony. Intenzita těchto spojení je reprezentována číslem v paměti počítače a přenastavení spojení znamená změna tohoto čísla, pochopitelně jsou za tímto procesem výpočty na základě matematické teorie. Způsob, jakým se propojení mění, je tedy v současné době naprogramován programátorem z masa a kostí.

První robot firmy DeepMind AlphaGo byl nejdříve učen daty z partií silných amatérských hráčů získaných z veřejně dostupných dat na herních serverech. Bylo použito učení s učitelem a neuronová síť byla trénována na uhodnutí dalšího tahu v pozici. Další trénink probíhal pomocí posilujícího učení a hry AlphaGo proti sobě. V památném roce 2016 hrála AlphaGo s I Se-tolem pět partií a AlphaGo vyhrála čtyři z nich. Ukážeme na jednom z tahů z těchto partií, že stroj, na rozdíl od člověka, netrpí předsudky. Na diagramu 1 má bílý vybudovanou pozici na straně desky, kterou černý svým tahem 1 redukuje. Toto je běžná technika používaná silnějšími hráči a zisk po možné sekvenci na diagramu 2 je považován za vyrovnaný. Jiná situace je na diagramu 3, pozice je zde o jednu linku dále od kraje a zde je zisk bílého považován za příliš velký. Když AlphaGo zahrála tah 3, byli přítomní profesionální hráči šokováni. Po dalších dvaceti tazích ale ocenili, jak dobře tah 3 doplňuje pozici černého.





Rok poté, v roce 2017, AlphaGoMaster další generace sehrál tři partie s čínským hráčem Kche Ťiem a tentokrát mu nedal žádnou šanci. Zajímavý výsledek přinesla hned první hra, kdy AlphaGoMaster vyhrál nejtěsnějším možným rozdílem, Go, na rozdíl od šachů, kdy hra končí matem, patem nebo vzdáním jednoho z hráčů, se hraje na body. Při hře si hráči budují pozice a v koncovce získávají území, které pak spočítají. AlphaGoMaster první partii vyhrál o půl bodu, což je nejtěsnější možný rozdíl a takový výsledek u dvou lidských hráčů signalizuje velmi vyrovnanou hru, kdy stačilo málo a vítěz se mohl změnit. Experti ale říkají, že o vítězství AlphaGoMastera nebylo pochyb, a jeho hru vysvětlují tím, že při hře optimalizuje pravděpodobnost výhry a ne velikost výsledku. Člověk také někdy hru zjednoduší a nevádí mu ztráta, pokud má značný náskok, přesto způsob dovedení partie k půlbodovému vítězství silné hráče překvapil. Možná bychom to mohli vysvětlit absencí strachu, že se může v koncovce ještě něco pokazit.

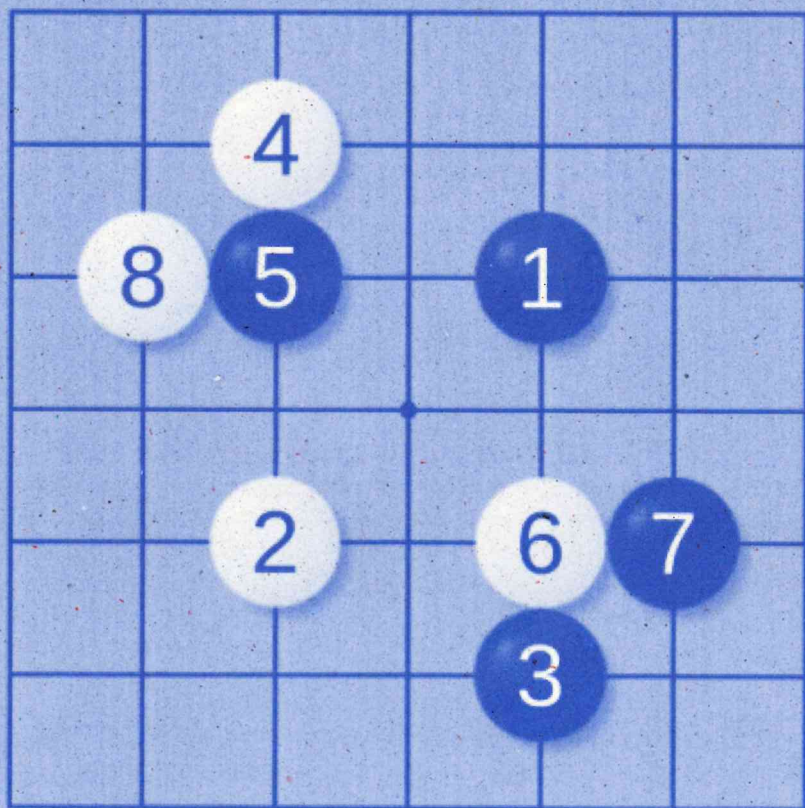
Britskou firmu Deep Mind koupil v roce 2014 Google a AlphaGo se stala reklamním článkem. DeepMind i Googlu, který ukazuje možnosti umělé inteligence způsobem pochopitelným pro širší

veřejnost. Je otázka, zda by bez peněz Googlu mohla AlphaGo vzniknout – máme na mysli hardware potřebný k vytrénování neuronových sítí. I přesto AlphaGo inspirovala mnoho hráčů k naprogramování dalších robotů hrajících go na poměrně slušné úrovni. Na hře těchto robotů ukážeme dva příklady, kdy lidský způsob uvažování vítězí nad strojovým.

První souvisí s konkrétní konstelací, kterou hráči go pro tvarovou podobnost nazývají schody. Zde je možné přecházet pozici mnoho tahů dopředu, v podstatě neomezeně, v praxi několik desítek. Je to tím, že máte jasný cíl (a je tedy snadné ohodnotit závěrečné pozice; viz úvod eseje) a po vašem tahu se soupeř může bránit jen jedním způsobem. Tyto schody někteří roboti nedokážou zahrát správně, přičemž záleží, jak byly trénované jejich neuronové sítě. Po tréninku pomocí pozic lidských hráčů nemají se schody problém, ale tréninkem výhradně hrou robota samotného proti sobě často zahrají schody špatně, přestože jinak hrají velmi silně.

Další situace souvisí s tzv. zrcadlovou hrou. Go se hraje na čtvercové síti 19 na 19, na začátku je deska prázdná a hráči střídavě pokládají kameny na křížové body sítě.

Na diagramu jsme desku v zájmu úspory místa zmenšili na 7 na 7; ukazujeme zrcadlovou hru, bílý kopíruje tahy černého na protější straně desky. Tuto strategii je možné snadno překazit položením kamene na prostřední průsečík. Stroj ale takto neuvazuje, pokaždé hledá ideální tah v dané pozici a střed obsadí jen v případě, že ho považuje za ideální místo. Pokud se bílému povede dotáhnout zrcadlovou hru až do konce partie, vyhraje, protože mu pravidla hry přiřknou několik bodů navíc jako kompenzaci za to, že jeho soupeř pokládal kámen jako první.





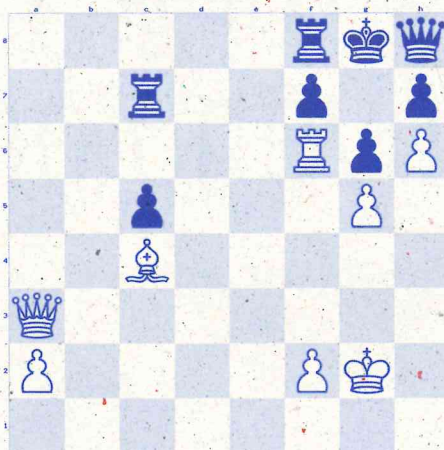
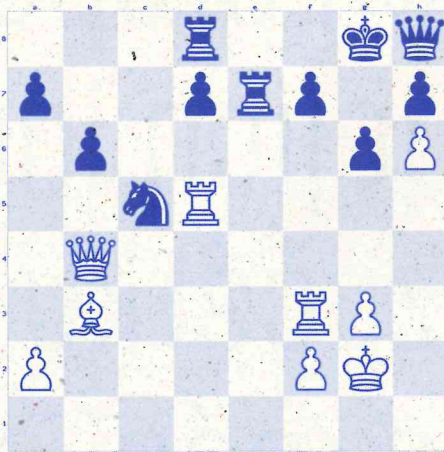
Pro podrobnější vysvětlení zmiňovaných dvou situací srovnáme práci člověka a stroje. Člověk při hře go používá podobné dvě neuronové sítě jako robot, jen je má hůře vytrénované. Používá při hře ale i další síť: někdy ke své škodě, kdy ho ovlivní předsudky nebo strach, a někdy ku prospěchu, kdy dokáže pochopit koncept schodů a zrcadlové hry.

Zápasy AlphaGo s I Se-tolem a Kche Ťiem zásadně změnilu hru go, ale nevzaly hráčům radost z ní. Roboti různých úrovní se dnes používají k analýze partií a též pro hráče objevují nové tahy a jejich sekvence. Tyto sekvence se pak často stávají mezi hráči velmi populárními.

Dalším krokem firmy DeepMind byl vývoj robota AlphaGoZero, při kterém byla přeskočena první fáze učení z dat silných amatérů. Robotu byla pouze předložena pravidla, začal náhodnou hrou sám proti sobě podle těchto pravidel a postupně podle výsledků partií zdokonaloval své neuronové síť. Posledním stupněm ve vývoji byla AlphaZero, které mohou být předložena pravidla jakékoliv hry dvou hráčů, ve které se střídají v tazích a mají úplnou informaci o stavu hry – tedy například šachů, go nebo dámy na rozdíl např. od pokeru, kde hráči neznají karty soupeřů, nebo člověče nezlob se, kde nevědí, jak dopadne hod kostkou.

Vývojáři DeepMind vytrénovali AlphaZero ke hře v šachy a s napětím čekali, zda AlphaZero odkryje další prostor pro pokrok za hranicemi možností stávajících robotů, kteří v té době již dvacet let poráželi nejsilnější světové hráče. Zároveň byli zvědaví, zda je výraznější průlom vůbec možný, a dočkali se překvapení: silní hráči označují hru AlphaZero za krásnou a stylem více podobnou lidské hře než hře silných robotů, které AlphaZero současně poráží. Během hry za větší dosah svých zbylých figur, případně za snížení dosahu figur soupeřových mnohem ochotněji obětuje figuru než stávající roboti. Na diagramu je ukázka z partie AlphaZero s robotem Stockfish. AlphaZero hraje bílými a první diagram je před tahem, ve kterém AlphaZero vymění věž za koně na c5, aby o deset tahů později v pozici na dalším diagramu neměl černý dobrý tah, byl nucen vyměnit dámu za věž a po dalších třech tazích černý partii vzdává.

Člověk při hře go používá podobné dvě neuronové sítě jako robot, jen je má hůře vytrénované. Používá při hře ale i další síť: někdy ke své škodě, kdy ho ovlivní předsudky nebo strach, a někdy ku prospěchu, kdy dokáže pochopit koncept schodů a zrcadlové hry.



Závěrem poznamenejme, že s vytrénovanými neuronovými sítěmi se v aplikacích, které často používáme, setkáváme častěji, než bychom očekávali. Dříve, než se začaly neuronové sítě používat, byly v aplikacích naprogramovány akce pro každou myslitelnou situaci. Nazvěme tento postup procedurální. Taktó byli naprogramováni šachoví roboti Deep Blue i Stockfish. Stejně tak mohly být naprogramovány systémy na poskytnutí úvěru. Zde použití neuronové sítě zvýšilo výkon aplikací. V případě úloh jako hraní go, analýza obrazu nebo jazykový překlad bychom namísto zvýšení výkonu mohli mluvit o průlomu.

Výsledkem je aplikace, které předáváte vstupy v digitální podobě a získáte od ní výstupy, zpravidla též digitální. Dále stojí za zmínku, že energeticky je lidská tvorba se strojovou nesrovnatelná, I Se-tol jak při hře, tak při tréninku spotřeboval podstatně méně zdrojů než AlphaGo.

Úplným závěrem dovolte malou provokaci vůči všem, kteří někdy někoho učili na jakémkoli vzdělávacím stupni. Zamyslete se, ke kterému ze způsobů, jež jsem nastínila v článku – procedurálnímu, učení s učitelem (*supervised learning*), posilujícímu učení (*reinforcement learning*) – má vaše působení na svěřence nejbliž.