

# Umělá inteligence z krabiček od sirek

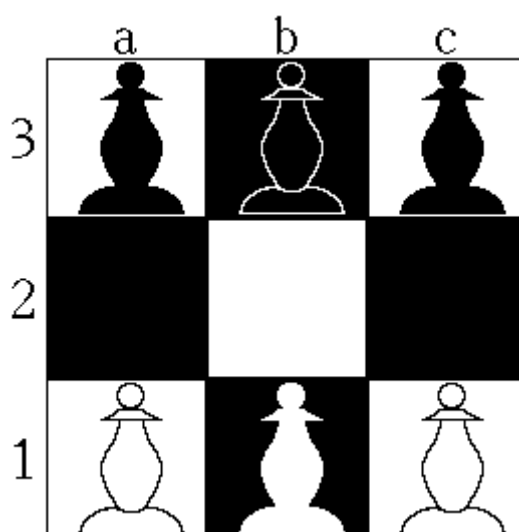
Tomáš Tayari

## Materiál a metoda

Cílem práce bylo sestavit jednoduchý počítač, který se umí učit, a sledovat účinnost učebních technik. Základem a inspirací pro tuto práci byl článek od Martina Gartnera (dostupné online: <http://cs.williams.edu/~freund/cs136-073/GardnerHexapawn.pdf>). Ten ve svém článku navrhuje jednoduchého robota HER (Hexapawn Educational Robot) z krabiček od zápalek a z korálků, schopného s člověkem hrát velmi zjednodušenou verzi šachů, tzv. hexapawn.

## Hexapawn

Hexapawn se hraje na 3x3 šachovnici pouze se šesti pěšci (trojicí od každé barvy) a s upravenými výherními kondicemi. Bílá zahájí hru. Hráči se střídají v tazích. Pěšci se pohybují dopředu o jedno pole, nebo berou nepřátelského pěšce křížem, stejně jako v šachách. Narozdíl od šachů pěšec i v prvním tahu táhne jen o jedno pole. Vyhrát se dá třemi způsoby: 1) hráč vyhraje, pokud jeho pěšec dosáhne druhé strany šachovnice, 2) hráč vyhraje, pokud protivníkovi sebere všechny figurky, 3) hráč prohraje, pokud ve svém tahu nemá možnost žádného legálního tahu.



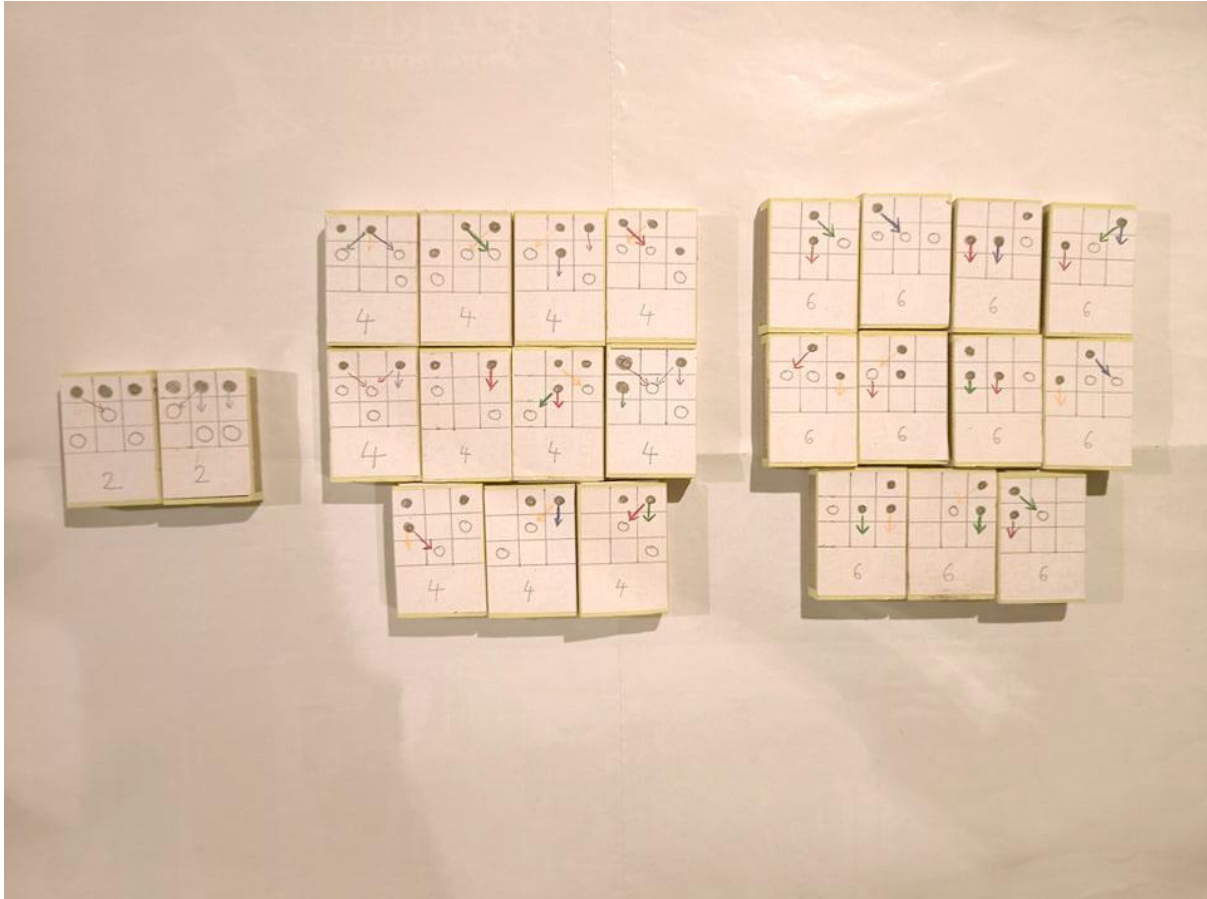
Obr 1: Rozložení při hře hexapawn  
(přejato z <https://en.wikipedia.org/wiki/Hexapawn>)

## Princip počítače

Počítač HER je zkonstruován z krabiček od sirek. HER vždy hraje s černými figurkami, tudíž začíná vždy lidský hráč. HER tedy stačí ke konstrukci pouze 24 krabiček.

Každému možnému rozložení figurek na hrací ploše odpovídá jedna krabička od sirek. Toto rozložení je znázorněno na přelepce na vrchu krabičky. Počítač potřebuje lidskou asistenci při operaci. Když je HER na tahu, lidský hráč vybere příslušnou krabičku s

odpovídajícím stavem hry. Na krabičce jsou pomocí šipek barevně vyznačeny všechny legální tahy, které může HER udělat. Každý tah je znázorněn šipkou s jinou barvou. Uvnitř krabiček jsou korálky ve stejných barvách jako možné tahy počítače. Lidský hráč zatřese krabičkou a naslepo krabičku otevře a vyjme z ní jeden korálek. Lidský hráč poté provede tah, který je znázorněn na vrchu krabičky stejnou barvou jako vytažený korálek. Korálek je následně vrácen zpět do krabičky. Krabičky byly přehledně očíslovány podle jednotlivých tahů, tedy 2, 4 a 6., protože HER hraje vždy jako druhá.



Obr 2: Počítač HER sestavený v rámci soustředění Hraštica

## Učební metody počítače

Martin Gartner ve svém článku navrhuje dvě možnosti učení počítače. Pomocí odměn a pomocí trestu. Po vyhrané partii lze počítač odměnit přidáním korálku do poslední krabičky v barvě tahu, na který počítač vyhrál. Případně lze počítač potrestat pro prohrané partii, kdy je mu vyndán korálek v barvě tahu, na který počítač prohrál.



Obr 3: Počítač HER při operaci

## Vlastní výzkum

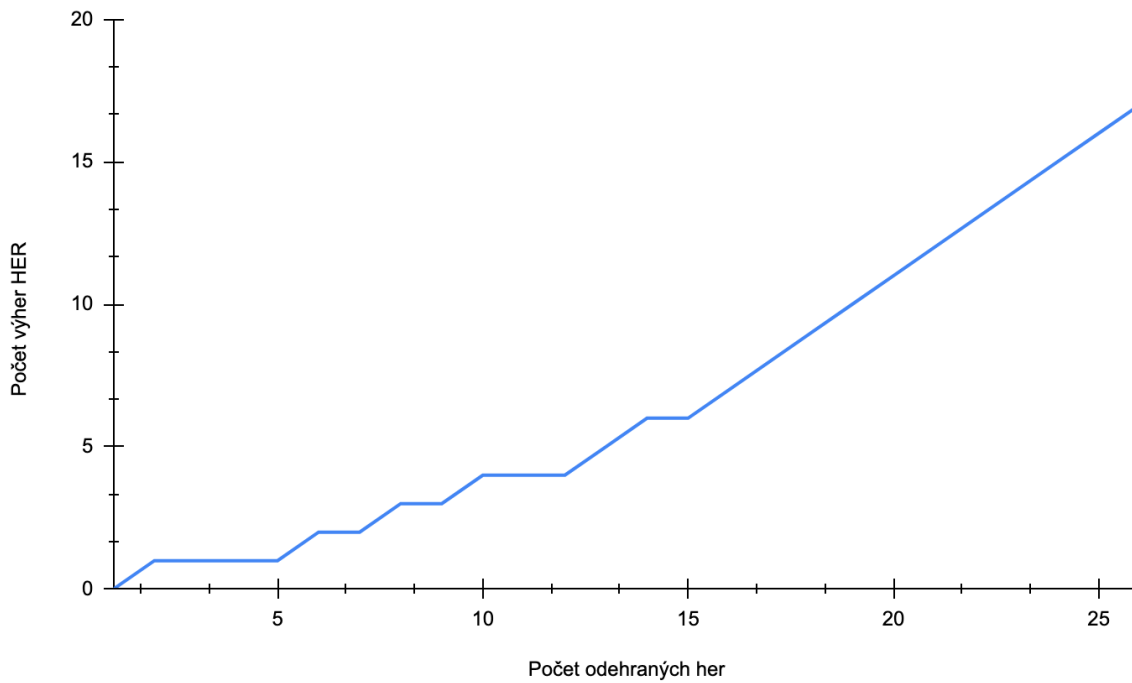
V práci byly realizovány oba způsoby učení. V první části byl počítač pouze trestán. Když dosáhl přijatelného počtu výher, byl počítač resetován do výchozího nastavení (v každé krabičce jen jeden korálek od každé barvy) a učení bylo pouze pomocí odměn. Cílem bylo srovnat metodu učení jen pomocí trestů a jen pomocí odměn.

### Učení pomocí trestu

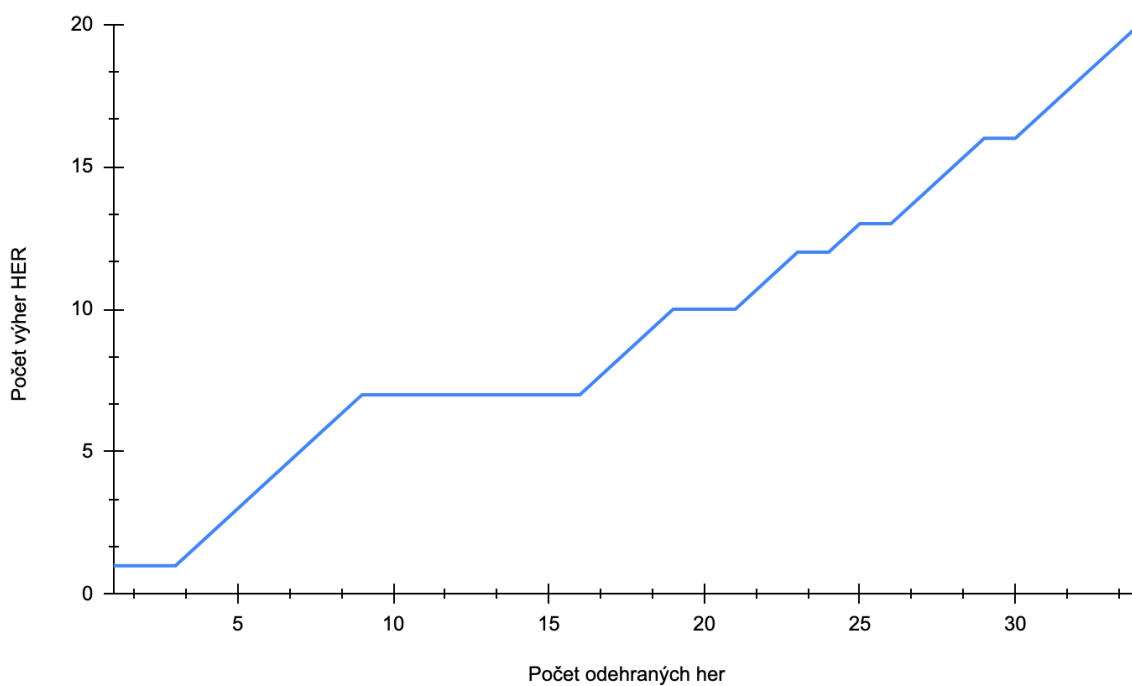
V prvním cyklu byl počítač HER učen pomocí trestu. V grafu 1 můžeme vidět závislost počtu vyhraných partií a počtu odehraných partií. Jak je vidět, od 15. hry byla HER již natolik zkušená, že pouze vyhrávala. Analýzou hry bylo zjištěno, že hráč, který hraje jako druhý, pokud hraje optimálně, nemůže prohrát. Experiment byl proto po 26 odehraných partiích zastaven. HER v momentě optimalizace (15. tah) vyhrála 40% her.

### Učení pomocí odměny

V druhém cyklu byl počítač HER učen pomocí metody odměn. Narozdíl od trestu se ani po 34 hrách nepodařilo dojít k optimalizaci, respektive k tomu, aby počítač vyhrál pokaždé. Na druhou stranu počítač po 15 odehraných partiích, které stačily při učení pomocí trestu k optimalizaci, měl úspěšnost asi 47%, což je o 7% vyšší než po 15 hrách při trestu. V momentě ukončení experimentu měla HER úspěšnost 59%. Ze směrnice grafu lze očekávat, že by se úspěšnost stále zvyšovala, ale z povahy experimentu lze usoudit že optimalizace nebude nikdy dosaženo.



Graf 1: Závislost výher počítače HER na počtu odehraných her při učební metodě trestu



Graf 2: Závislost výher počítače HER na počtu odehraných her při učební metodě odměny

## Závěr

V rámci podzimního soustředění Hraštice byl sestaven počítač HER schopný hrát s živým člověkem zjednodušené šachy Hexapawn.

Počítač byl schopen se učit a odděleně byly otestovány dvě metody učení, pomocí odměn a pomocí trestu. Ukázalo se, že počítač při trestu dosáhne optimalizace výrazně rychleji než při odměnách, na druhou stranu hra je v tom případě velmi přímočará a nezábavná. Při odměně má průměrný počet výher vyšší, a mimo jiné jsou hry s ním zábavnější vzhledem k tomu, že se počítač nebojí experimentovat, bohužel počítač nikdy nedojde úplné optimalizace.

Výsledky experimentu jsou zohledněny během kariéry středoškolského učitele, kterou autor práce v současné době vykonává.

Na psychický stav počítače při experimentu nebyl brán zřetel.