

# Databáze koncoviek

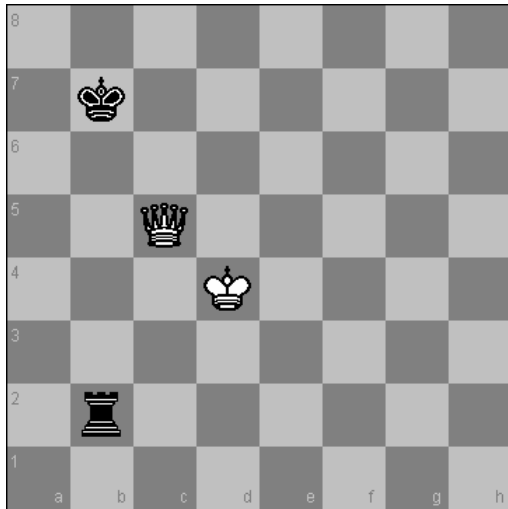
# Databáza vs. minimax

- na riešenie väčšiny deterministických stolových hier s perfektnými znalosťami a rozumným faktorom vetvenia používame prehľadávanie *minimax* spolu a *alfa-beta odsekávaním* a inými pomocnými algoritmami
- tieto algoritmy poskytujú dobré výsledky pre väčšinu situácií, ktoré môžu v takýchto hrách nastať
- aj napriek použitiu alfa-beta sa však málokedy dočkáme úplného prepočtu pozície (väčšinou až tesne pred matom)
- hra sa niekedy môže dostať do „koncovky“ – to je tá fáza hry, v ktorej na hracom poli ostane len niekoľko kameňov
- aj v koncovke môže byť (a často býva) pozícia minimaxovou metódou neriešiteľná v reálnom časovom horizonte
- pre koncovky existuje lepší spôsob riešenia ako algoritmus minimax – databáze koncoviek

# Čo je databáza koncovky?

- pre jednoznačnosť výkladu budeme ďalej uvažovať len šachové koncovky
- je to zoznam všetkých pozícií, ktoré sa môžu vyskytnúť s určitým počtom kameňov na šachovnici
- každá pozícia má priradené dve čísla – jedno je ohodnotenie, pokiaľ je na ťahu biely, druhé je ohodnotenie, pokiaľ je na ťahu čierny
- čísla určujú počet polťahov do matu pri optimálnej hre oboch strán
- mat má ohodnotenie 0

# Použitie – príklad



- na ťahu je biely
  - nájdeme v databáze túto pozíciu (v logaritmickom čase – prvky sú zoradené)
  - ohodnotenie pre bieleho je +33, čo znamená, že biely vyhrá pri optimálnej hre za 33 polťahov, teda 17 ťahov
  - ak čierny urobí chybu, výhra nastane skôr
  - získali sme informáciu o tom, koľko polťahov zostáva do matu, ale **ako máme nájsť najlepší ťah bieleho?**
- urobíme zoznam všetkých pozícií, ktoré vzniknú vykonaním jedného ťahu
  - tieto pozície vyhľadáme v databáze a bude nás pri nich zaujímať ohodnotenie pre opačnú farbu – v tomto prípade čiernu
  - dostaneme nasledujúci zoznam:

# Použitie – príklad

Poradové číslo ťahu	Ťah	Ohodnotenie
<b>1.</b>	<b>Dc5-f5</b>	<b>32</b>
2.	Kd4-d5	36
3.	Dc5-d5	36
4.	Kd4-e5	38
5.	Dc5-a5	38
6.	Dc5-e5	38
7.	Dc5-c4	38
8.	Dc5-a3	38
9.	Dc5-e7	38
10.	Dc5-f8	38
11.	Kd4-c3	40
12.	Dc5-g5	40
13.	Dc5-c3	40
14.	Dc5-c1	40

Poradové číslo ťahu	Ťah	Ohodnotenie
15.	Kd4-e3	42
16.	Dc5-h5	42
17.	Kd4-e4	44
18.	Kd4-d3	44
19.	Kd4-c4	remíza
20.	Dc5-d6	remíza
21.	Dc5-c2	31
22.	Dc5-c8	31
23.	Dc5-a7	31
24.	Dc5-b5	29
25.	Dc5-c7	29
26.	Dc5-b4	29
27.	Dc5-b6	29
28.	Dc5-c6	27

# Použitie – príklad

- ťahy sú zoradené podľa toho, ako rýchlo vedú k matu
- párne (sudé) čísla znamenajú prehru v  $n$  polťahoch pre hráča, ktorý je na ťahu
- nepárne (liché) čísla znamenajú výhru
- čierny najrýchlejšie prehrá z pozície, ktorá vznikne po ťahu Dc5-f5
- rozloženie ťahov býva väčšinou také, ako v našom príklade:
  - jeden (dva, tri) ťahy sú najlepšie
  - nasleduje séria horších, ale stále vyhrávajúcich ťahov
  - remízové ťahy
  - prehrávajúce ťahy
- použitie databáze je triviálnou úlohou
  - počítač dokáže v priebehu veľmi krátkeho času nájsť najlepší ťah, aj keď pracuje s pevným diskom
- zložitejšie je jej vytvorenie

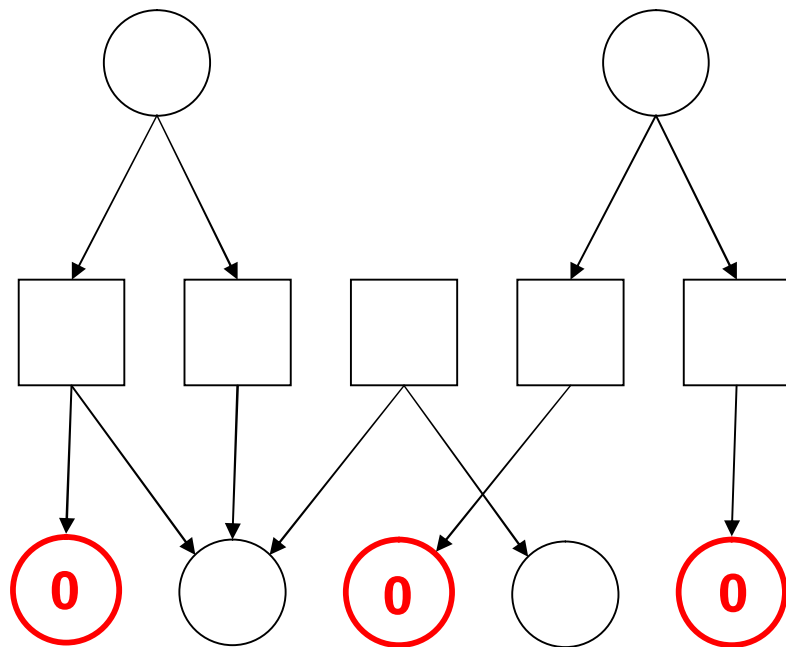
# Vytvorenie databáze

- generujeme zoznam všetkých možných pozícií
- vyradíme pozície, ktoré sú
  - nemožné (viac kameňov na jednom poli)
  - nelegálne (obaja králi v šachu, nemožnosť ťahu pre obe strany)
  - symetrické – kôli úspore miesta
- ohodnotíme všetky pozície ako nehodnotené
- pri ohodnocovaní pozícií postupujeme retroanalyticky (induktívne, od konca)

Problém ohodnocovania si rozdelíme na dva podproblémy. Najprv budeme hľadať a ohodnocovať len pozície víťazné pre bieleho. Pre čierneho aplikujeme ten istý postup, len s prehodením farieb.

# Vytvorenie databáze – diagram

1. ohodnotíme všetky matové pozície



Lukáš Mojžiš

Databáze koncoviek

Legenda:

štvorec – biely na ťahu

kruh – čierny na ťahu

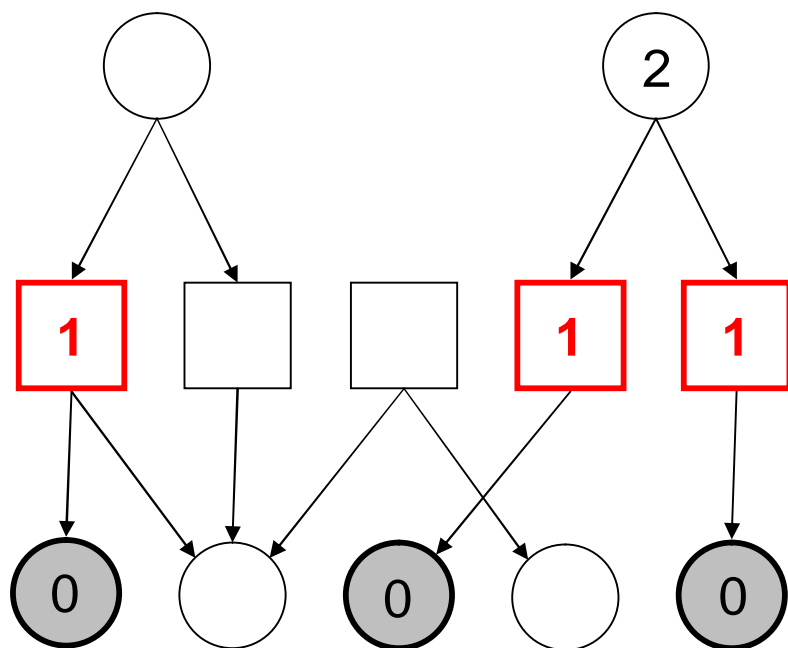
číslo – ohodnotenie pozície

bez čísla – neohodnotená pozícia



# Vytvorenie databáze – diagram

2. ohodnotíme pozície, v ktorých biely môže dať jedným poľahom mat



Legenda:

štvorec – biely na ťahu

kruh – čierny na ťahu

číslo – ohodnotenie pozície

bez čísla – neohodnotená pozícia

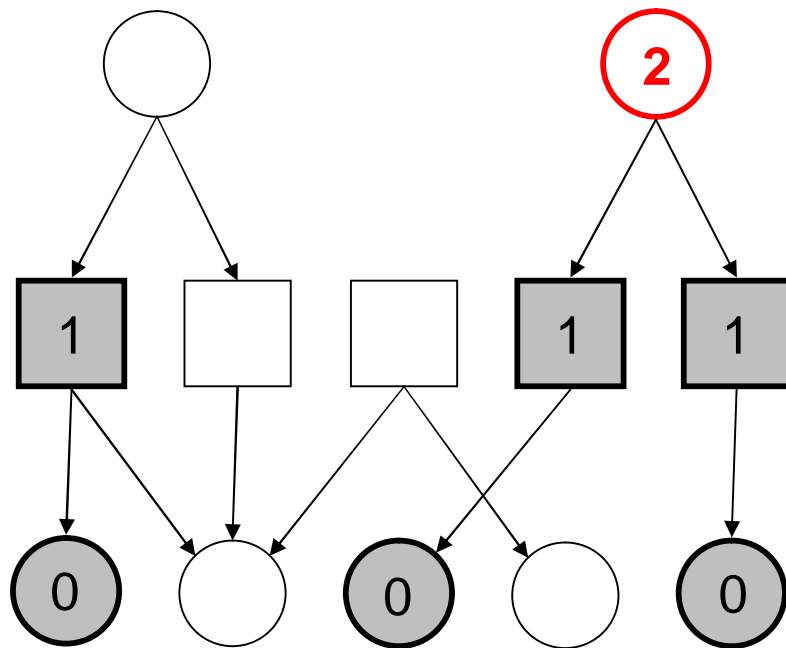
Lukáš Mojžiš

Databáze koncoviek

9

# Vytvorenie databáze – diagram

3. ohodnotíme pozície, v ktorých sa čierny nemôže vyhnúť matu druhým poľahom



Legenda:

štvorec – biely na ťahu

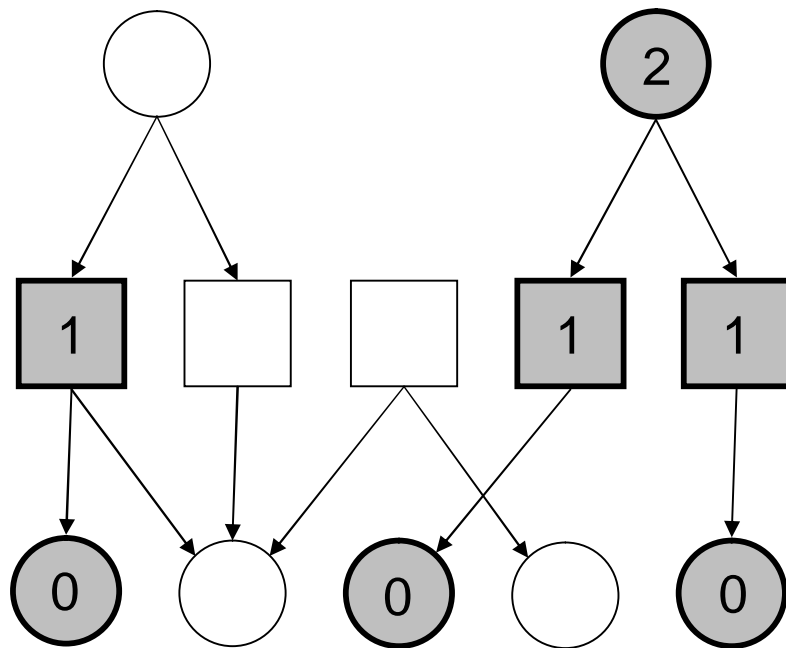
kruh – čierny na ťahu

číslo – ohodnotenie pozície

bez čísla – neohodnotená pozícia

# Vytvorenie databáze – diagram

- opakujeme kroky 2 a 3  
pokiaľ nemáme čo ohodnocovať, skončíme



Legenda:

štvorec – biely na ťahu

kruh – čierny na ťahu

číslo – ohodnotenie pozície

bez čísla – neohodnotená pozícia

# Vytvorenie databáze

1. ohodnotíme nulou všetky pozície, v ktorých čierny dostal mat
2. nech  $n=1$
3. hľadáme pozície, v ktorých je na ťahu biely a vyhrá  $n$ -tým poťahom
  - z každej pre bieleho zatiaľ neohodnotenej pozície **P** v databáze generujeme všetky ťahy
  - ak aspoň jeden ťah vedie do pozície s ohodnotením  $n-1$  pre čierneho, ohodnotíme **P** číslom  $n$
4. hľadáme pozície, v ktorých je na ťahu čierny a prehrá  $(n+1)$  poťahom
  - z každej pre čierneho zatiaľ neohodnotenej pozície **Q** v databáze generujeme všetky ťahy
  - ak všetky ťahy vedú do pozícií s ohodnotením najviac  $n$  pre bieleho, ohodnotíme **Q** číslom  $n+1$
5. induktívne aplikujeme pravidlá 3 a 4 ( $n+=2$ ), až pokiaľ neohodnotíme ani jednu pozíciu

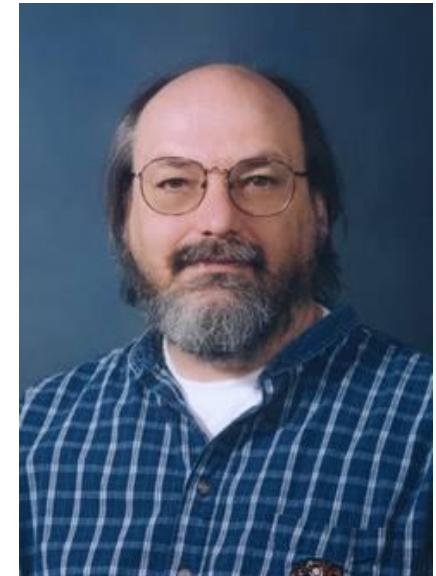
# Vytvorenie databáze

Vo výsledku nám ostanú aj neohodnotené pozície. Tieto pozície sú remízové a môžeme ich vypustiť.

- alternatívny (asi efektívnejší) spôsob je generovať predchodcov:
- namiesto toho, aby sme prehľadávali celú databázu a robili test pre každú neohodnotenú pozíciu, vytvoríme z každej už ohodnotenej pozície **P** jej predchodcov – všetky pozície, z ktorých sa jedným ťahom dostaneme do **P**
- ďalšie testy už robíme len s takto generovanými pozíciami, neprehľadávame databázu

# História

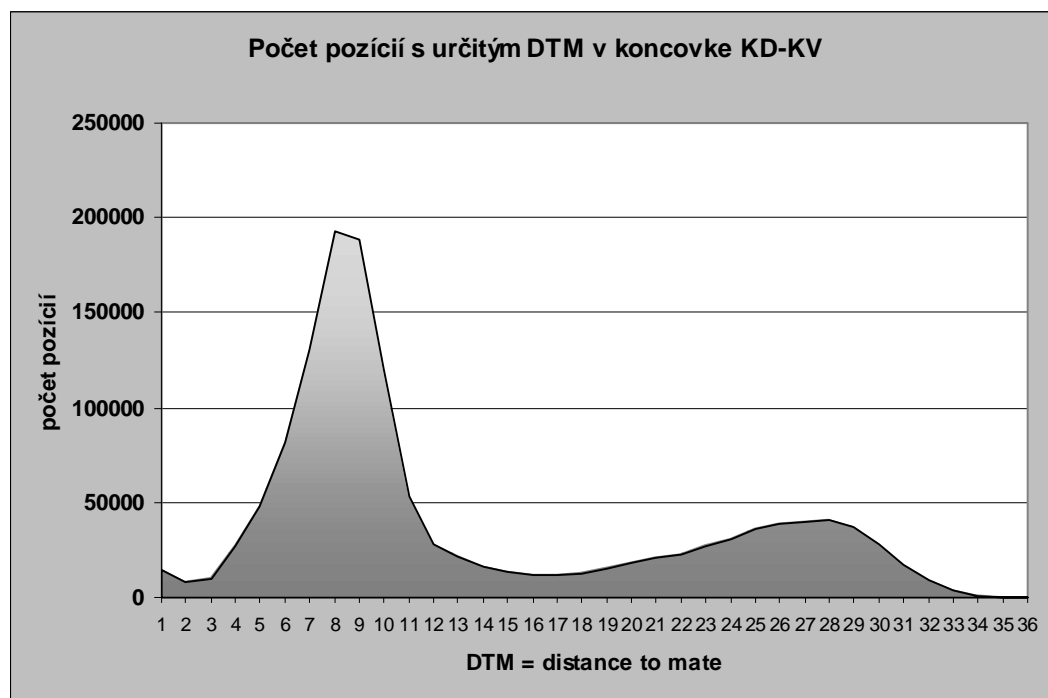
Prvý kto skonštruoval koncovkovú databázu bol Ken Thompson. Bola to koncovka kráľ a dáma proti kráľovi a veži. Táto koncovka sa dovtedy považovala za ľahkú výhru pre stranu s dámou. Preto boli veľmajstri na šachovom turnaji, ktorý sa konal začiatkom 80. rokov v USA prekvapení, keď Thompson ponúkol 100 dolárov tomu, kto nastúpi s dámou v tejto koncovke proti jeho novému programu. Väčšina prítomných veľmajstrov považovala túto výzvu za žart, a trojnásobný šampión USA Walter Browne stávkou ihneď prijal. Odohralo sa historické stretnutie, v ktorom veľmajster zistil, že poraziť počítač nie je až také jednoduché. Stávkou prehral. Potom nasledovala odveta, v ktorej sa mu tesne podarilo zvíťaziť. Počítaču trvalo vytvorenie databázy pol dňa. Bolo v nej 1,7 milión záznamov. Dnešné PC to zvládne za 1 minútu.



# Príklad koncovky: KD-KV

Nasledujúci graf zhŕňa počet pozícií, ktoré majú rovnaké ohodnotenie pre bieleho v koncovke KD-KV. Zaujímavé je na ňom to, že zodpovedá obtiažnosti tejto koncovky v jednotlivých jej fázach. Ak začneme napr. z pozície, z ktorej je do matu 35 ťahov, spočiatku sa nám bude hrať ľahko. Keď sa

však dostaneme k asi 22 ťahom do matu, začne sa počet možností, ktoré vedú k výhre zmenšovať. Preto sa v tejto fáze dá ľahko urobiť chyba. Ak sa dostaneme k číslu DTM = 12, ďalej sa už bude opäť hrať ľahko, až do matu. Aj veľmajster Browne pri svojej prvej partii urobil najviac chýb pri DTM 13 až 22.



# Súčasný stav teórie koncoviek

Dnes sú všetky dôležité 4, 5 a 6 kameňové koncovky hotové. Bolo analyzovaných už aj niekoľko sedemkameňových koncoviek. Za zmienku stojí ukázať, ako sú tieto databáze asi veľké – napríklad koncovka KVS-KJJ má asi 100 miliárd pozícií a najväčšia vzdialenosť do matu je 223 ťahov. Bola vytvorená v roku 1991 na počítači, ktorý mal 65536 procesorov. V počte ťahov do matu drží dnes rekord sedemkameňová koncovka.

Zatiaľčo štvor a päťkameňové koncovky sú ešte sčasti pochopiteľné, zo šesťkameňových sa nijaká pre ľudí využiteľná stratégia nedá odvodiť.

Na stránke <http://www.k4it.de/index.php?topic=egtb&lang=en> sú online prístupné šesťkameňové koncovky.



# Dámové koncovky

V dámových koncovkových databázach sa pri pozícii neukladá vzdialenosť do matu, ale iba informácia win/draw/lose. Preto môže byť koncovka lepšie komprimovaná za tú cenu, že pri jej používaní treba robiť dodatočné výpočty. V dáme boli nedávno analyzované všetky desaťkameňové koncovky. Spolu zaberajú 250 GB po komprimácii. Výpočet bežal na niekoľkých paralelne spojených PC 4,2 roka. Pri takomto dlhom čase výpočtu nie je pravdepodobnosť, že sa počítač pomýli malá, a preto treba dodatočne overiť integritu databáze.