

# Information Sharing in MCTS

## Diplomová práce

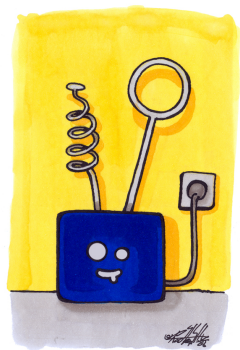
Petr Baudiš <pasky@ucw.cz>

MFF UK

Září 2011

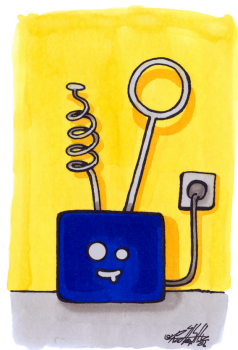
# Téma práce

- Umělá inteligence pro deskovou hru Go

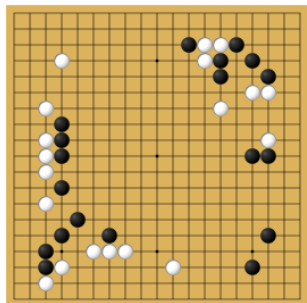


# Téma práce

- Umělá inteligence pro deskovou hru Go
- Ucelený úvod do Go z pohledu UI a do algoritmu MCTS
- Podrobný popis programu Pachi
- Tři nové výsledky: Dynamické komi, kritičnost tahů, mapy svobod
- Návrh dalších směrů vývoje

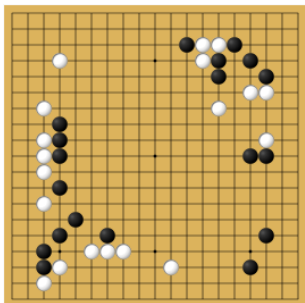


# Hra Go a umělá inteligence



- Desková hra s jednoduchými pravidly
- Často považována za jednu z “výzev umělé inteligence”

# Hra Go a umělá inteligence



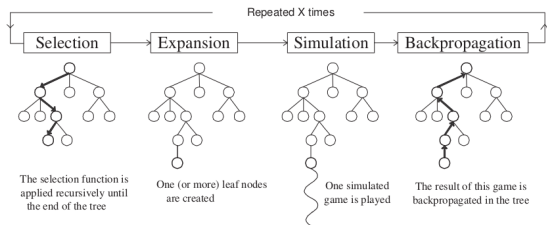
- Desková hra s jednoduchými pravidly
- Často považována za jednu z “výzev umělé inteligence”
- Velký stavový prostor
- Není známa dobrá ohodnocovací funkce ani prořezávací heuristiky

# Monte Carlo Tree Search

- Pravděpodobnostní minimaxový strom, každý uzel má očekávanou pravděpodobnost výhry  $\mu$
- Multi-armed Bandit: exploration vs. exploitation

# Monte Carlo Tree Search

- Pravděpodobnostní minimaxový strom, každý uzel má očekávanou pravděpodobnost výhry  $\mu$
- Multi-armed Bandit: exploration vs. exploitation
- Na začátku pouze kořen; iterujeme:
  - Sestupujeme do listu, syny vybírá Multi-armed Bandit
  - Navštívíme-li list  $n$ -krát, vytvoříme všechny následníky
  - V listu spustíme Monte Carlo simulaci
  - Výsledek hry zpropagujeme zpět do cesty ke kořeni



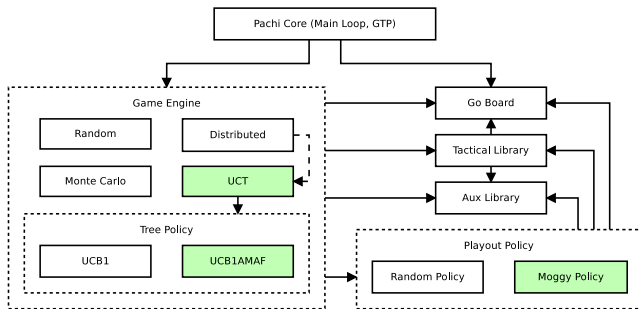
# Monte Carlo Tree Search

- Pravděpodobnostní minimaxový strom, každý uzel má očekávanou pravděpodobnost výhry  $\mu$
- Multi-armed Bandit: exploration vs. exploitation
- Na začátku pouze kořen; iterujeme:
  - Sestupujeme do listu, syny vybírá Multi-armed Bandit
  - Navštívíme-li list  $n$ -krát, vytvoříme všechny následníky
  - V listu spustíme Monte Carlo simulaci
  - Výsledek hry zpropagujeme zpět do cesty ke kořeni
- Strom roste dynamicky podle směru prohledávání
- Sdílení informací?
- RAVE Multi-armed Bandit: Prohledáváme tahy, které se osvědčily v simulacích



# Program Pachi

- Počítačový hráč Go
- Standardní MCTS: algoritmus RAVE a sada heuristik (detailně vyladěné parametry)
- Modulární, malý, optimalizované C, otevřený zdrojový kód



# Program Pachi

- Počítačový hráč Go
- Standardní MCTS: algoritmus RAVE a sada heuristik (detailně vyladěné parametry)
- Modulární, malý, optimalizované C, otevřený zdrojový kód
- Pravidelně hraje na internetu, vysoká třída (3 dan), pravidelná umístění v soutěžích programů
- Ve spolupráci s Jean-Loupem Gaillym: Algoritmus pro efektivní rozvrhování času
- Testovací infrastruktura “autotest”

## Nové výsledky ve sdílení informací

- Dynamické komi — zvýšení rozlišení v extrémních situacích (např. hendikepové partie), posunujeme bodový práh výhry
- Kritičnost — souvislost mezi vlastnictvím bodu a výhrou; starší myšlenka, my ji integrujeme do RAVE
- Mapy svobod — průběžné “samoučení heuristik” v konkrétních situacích definovaných okolím dotčených kamenů

# Závěr

- Silný program Pachi s otevřeným zdrojovým kódem, bude dále vyvíjen v rámci PhD studia
- Sdílení informací — povedlo se dosáhnout několika méně zásadních, ale zajímavých zlepšení
- Plán do budoucnosti — intenzivní sdílení informací pomocí “lokálních stromů”

# Závěr

- Silný program Pachi s otevřeným zdrojovým kódem, bude dále vyvíjen v rámci PhD studia
- Sdílení informací — povedlo se dosáhnout několika méně zásadních, ale zajímavých zlepšení
- Plán do budoucnosti — intenzivní sdílení informací pomocí “lokálních stromů”

**Děkuji za pozornost!**