

Neinformované metody prohledávání stavového prostoru

Michal Pěchouček, Milan Rollo

Department of Cybernetics
Czech Technical University in Prague



<http://cw.felk.cvut.cz/doku.php/courses/a3b33kui/start>



Úvodní poznámka o umělé inteligenci

Symbolický funkcionismus je založen na modelování dvou základních aspektů inteligentního uvažování:

- znalostí
- uvažování

Obojí lze modelovat na různých úrovních obecnosti. **Silné metody** umožňují obecné modely uvažování, zatímco **slabé metody** jsou specifické.

:: Zjednodušená úloha umělé inteligence podle symbolického funkcionalismu tedy zní: jak reprezentovat ty správné znalosti a naprogramovat takové uvažovací mechanismy, které obohatí soubor znalostí o nové hypotézy.





Úvodní poznámka o umělé inteligenci

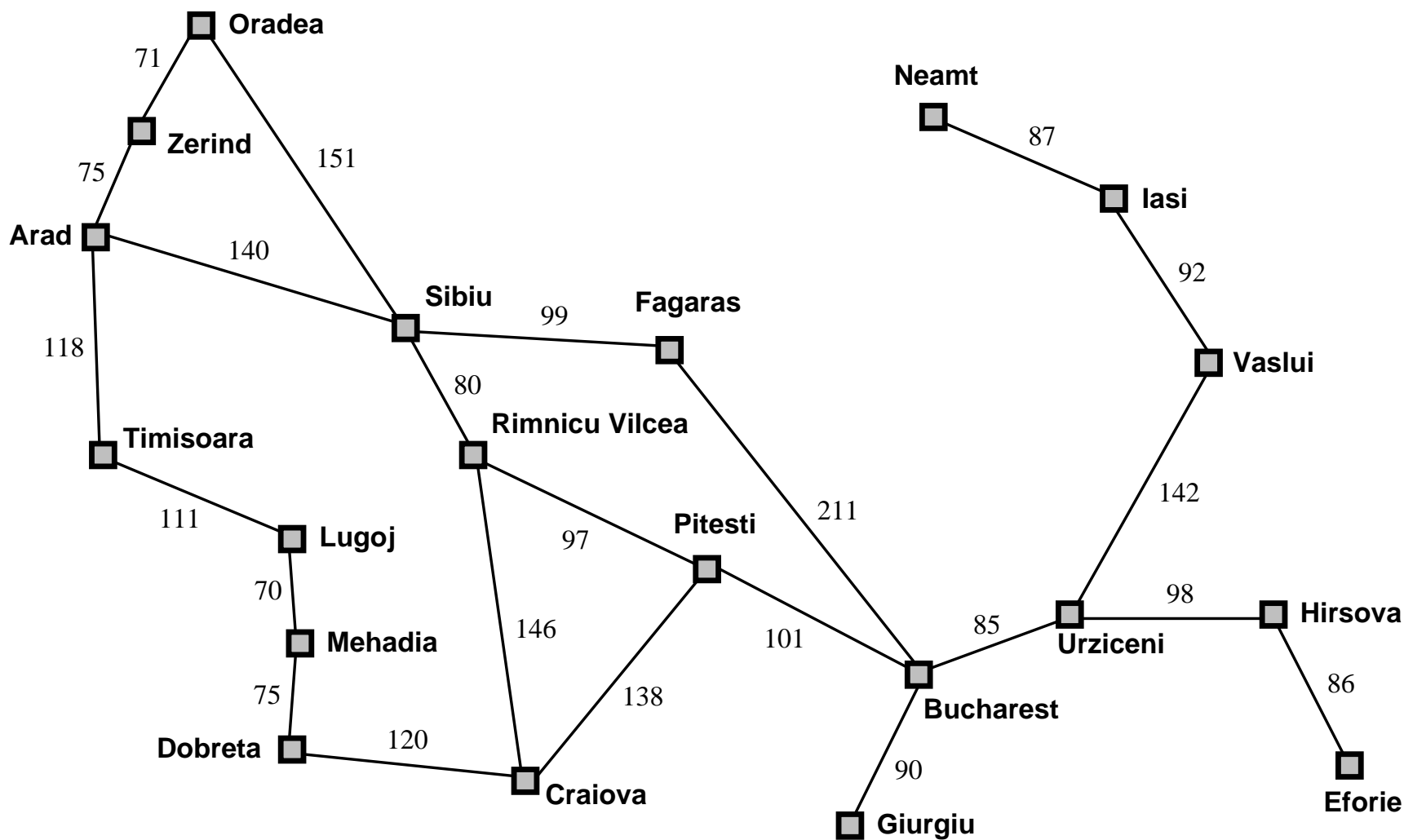
Mějme dva extrémní případy implementace umělé inteligence podle symbolického funkcionalismu:

- **silný** – znalosti jsou reprezentovány pomocí systému predikátové logiky a uvažování je reprezentováno výpočetním modelem dedukce
- **slabý** – znalosti jsou reprezentovány jako množina výroků a uvažování je reprezentováno pomocí souboru `if-then` pravidel

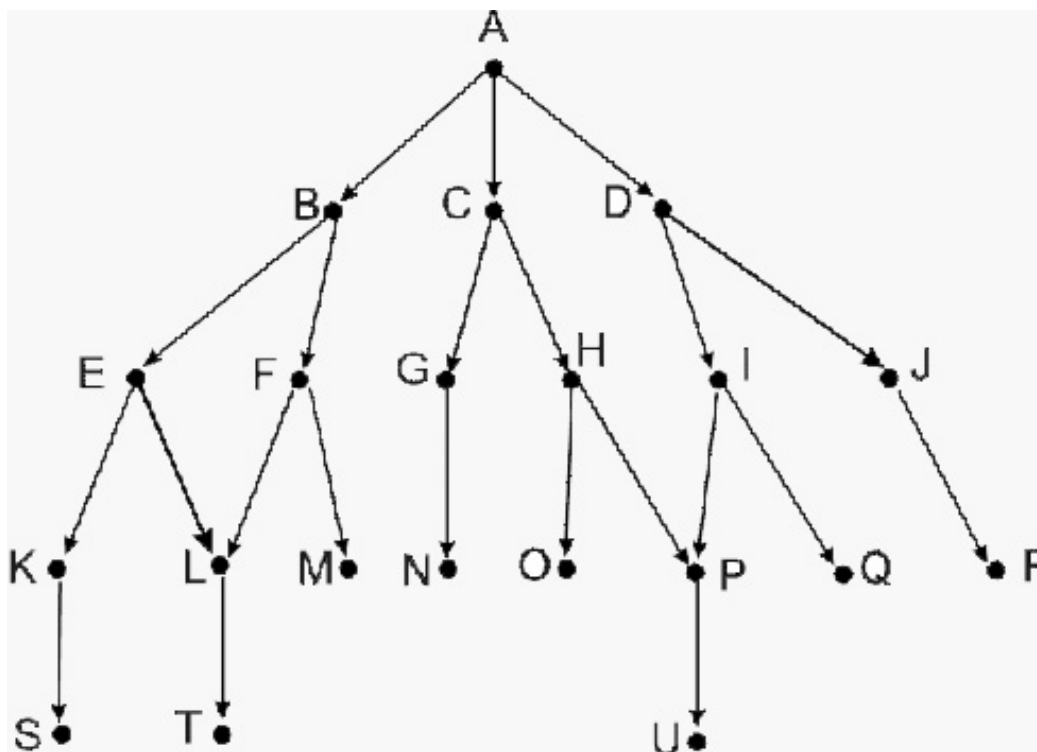
V obou případech je množina nových znalostí (ať už vytvořených nebo hypotetických) veliká a je potřeba ji inteligentně vytvářet a prohledávat. Strategie prohledávání je součástí modelu uvažování. Prostor nových znalostí se nazývá **stavový prostor**. Při řešení problémů se skládá stavový prostor z meziřešení či pomocných hypotéz. Některá meziřešení jsou klasifikovaná jako cílová řešení.



Příklad: Obchodní cestující



Prohledávání Stavového Prostoru



Prohledávání do šířky – Breadth First Search (BFS)



```
begin
  open := [Start]
  while (open <> []) do begin
    X := first(_open)
    open := open - [X]
    if X = GOAL then return(SUCCESS)
    else begin
      E := expand(X)
      open := open + E
    end
  end
  return(failure)
end.
```

znak <> znamená nerovno,

operátor - znamená odebrání prvku/prvků ze seznamu a

operátor + znamená přiřazení prvku/prvků na konec seznamu





Mějme b *maximální faktor větvení* (největší počet hran jdoucích z libovolného uzlu) daného stromu, d - nejmenší hloubka stromu, kde se nachází řešení a m maximální hloubka stromu - může být ∞ .

- **úplné:** ANO (je-li b konečné)



Mějme b maximální faktor větvení (největší počet hran jdoucích z libovolného uzlu) daného stromu, d - nejmenší hloubka stromu, kde se nachází řešení a m maximální hloubka stromu - může být ∞ .

- úplné: ANO (je-li b konečné)
- čas ?



Vlastnosti BFS

Mějme b *maximální faktor větvení* (největší počet hran jdoucích z libovolného uzlu) daného stromu, d - nejmenší hloubka stromu, kde se nachází řešení a m maximální hloubka stromu - může být ∞ .

- **úplné:** ANO (je-li b konečné)
- **čas:** $1 + b + b^2 + b^3 + \dots + b^d + b(b^d - 1) = O(b^{d+1})$ – podle algoritmu uvedeného na slidu 25, počítá se počet expandovaných uzlů (maximální počet uzlů na open seznamu), platí jen v případě, že $m > d$ (jinak je $O(b^d)$).





Vlastnosti BFS

Mějme b *maximální faktor větvení* (největší počet hran jdoucích z libovolného uzlu) daného stromu, d - nejmenší hloubka stromu, kde se nachází řešení a m maximální hloubka stromu - může být ∞ .

- **úplné:** ANO (je-li b konečné)
- **čas:** $1 + b + b^2 + b^3 + \dots + b^d = O(b^d)$
- **paměť:** $O(b^d)$





Vlastnosti BFS

Mějme b maximální faktor větvení (největší počet hran jdoucích z libovolného uzlu) daného stromu, d - nejmenší hloubka stromu, kde se nachází řešení a m maximální hloubka stromu - může být ∞ .

- úplné: ANO (je-li b konečné)
- čas: $1 + b + b^2 + b^3 + \dots + b^d = O(b^d)$
- paměť: $O(b^d)$
- optimální ?





Vlastnosti BFS

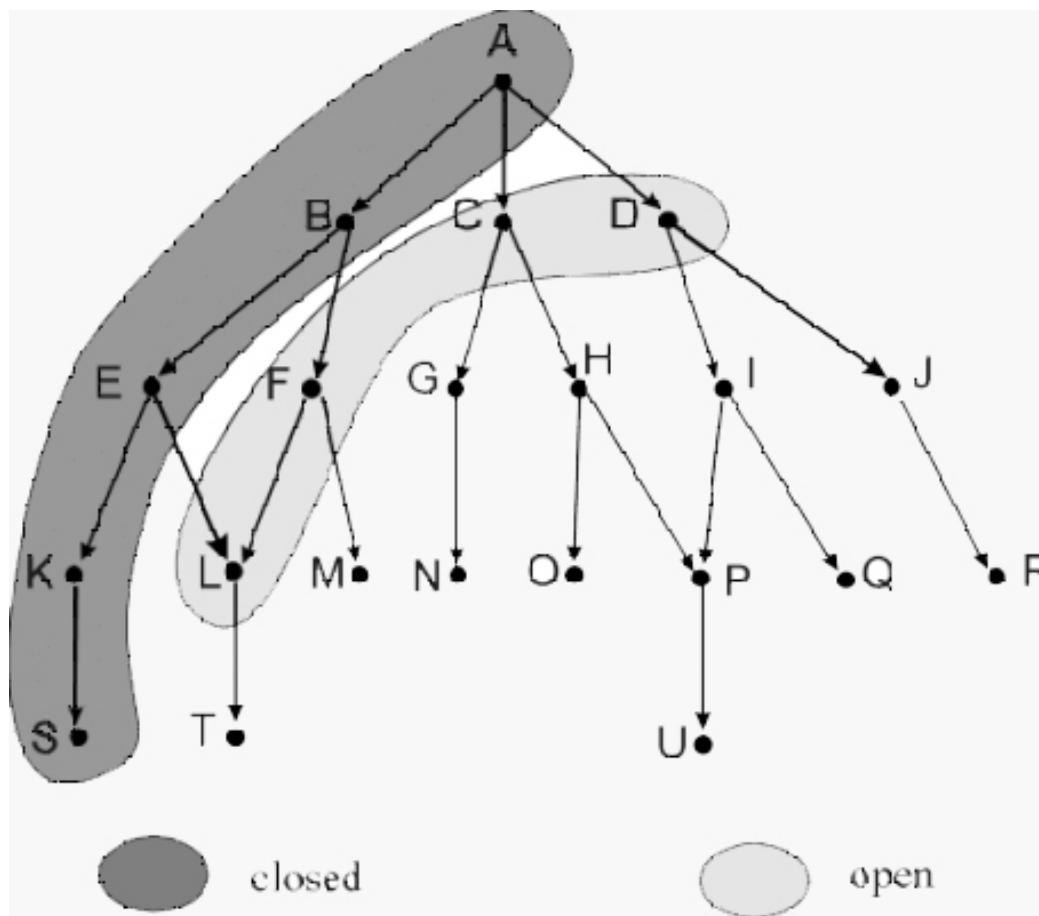
Mějme b *maximální faktor větvení* (největší počet hran jdoucích z libovolného uzlu) daného stromu, d - nejmenší hloubka stromu, kde se nachází řešení a m maximální hloubka stromu - může být ∞ .

- **úplné**: ANO (je-li b konečné)
- **čas**: $1 + b + b^2 + b^3 + \dots + b^d = O(b^d)$
- **paměť**: $O(b^d)$
- **optimální**: ano, optimalizuje-li se hloubka

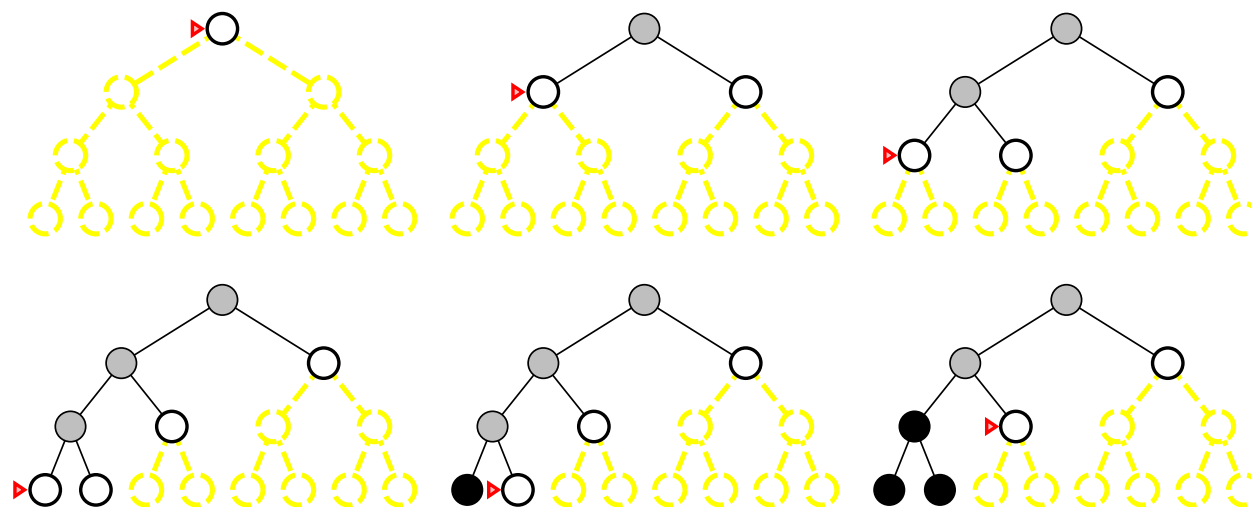
Prostor je největší problém - lehce lze generovat 100MB/sec



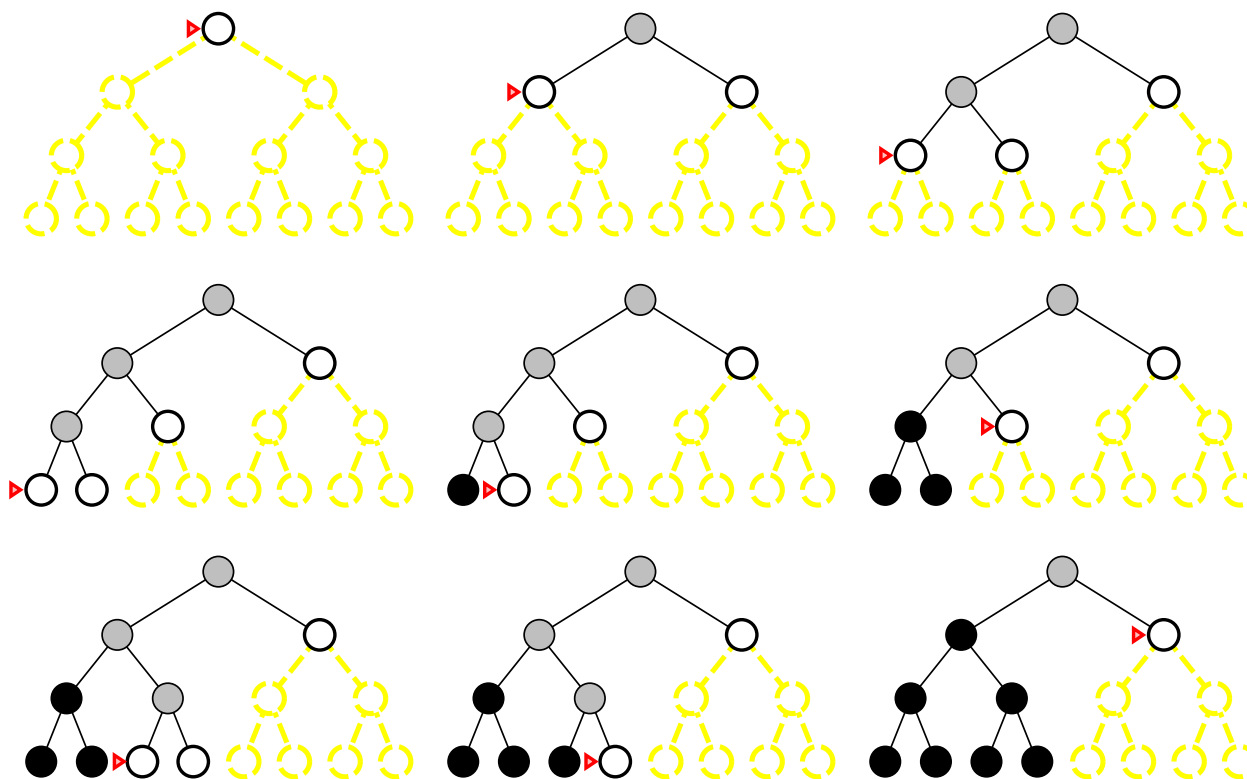
Prohledávání do hloubky – Depth First Search (DFS)



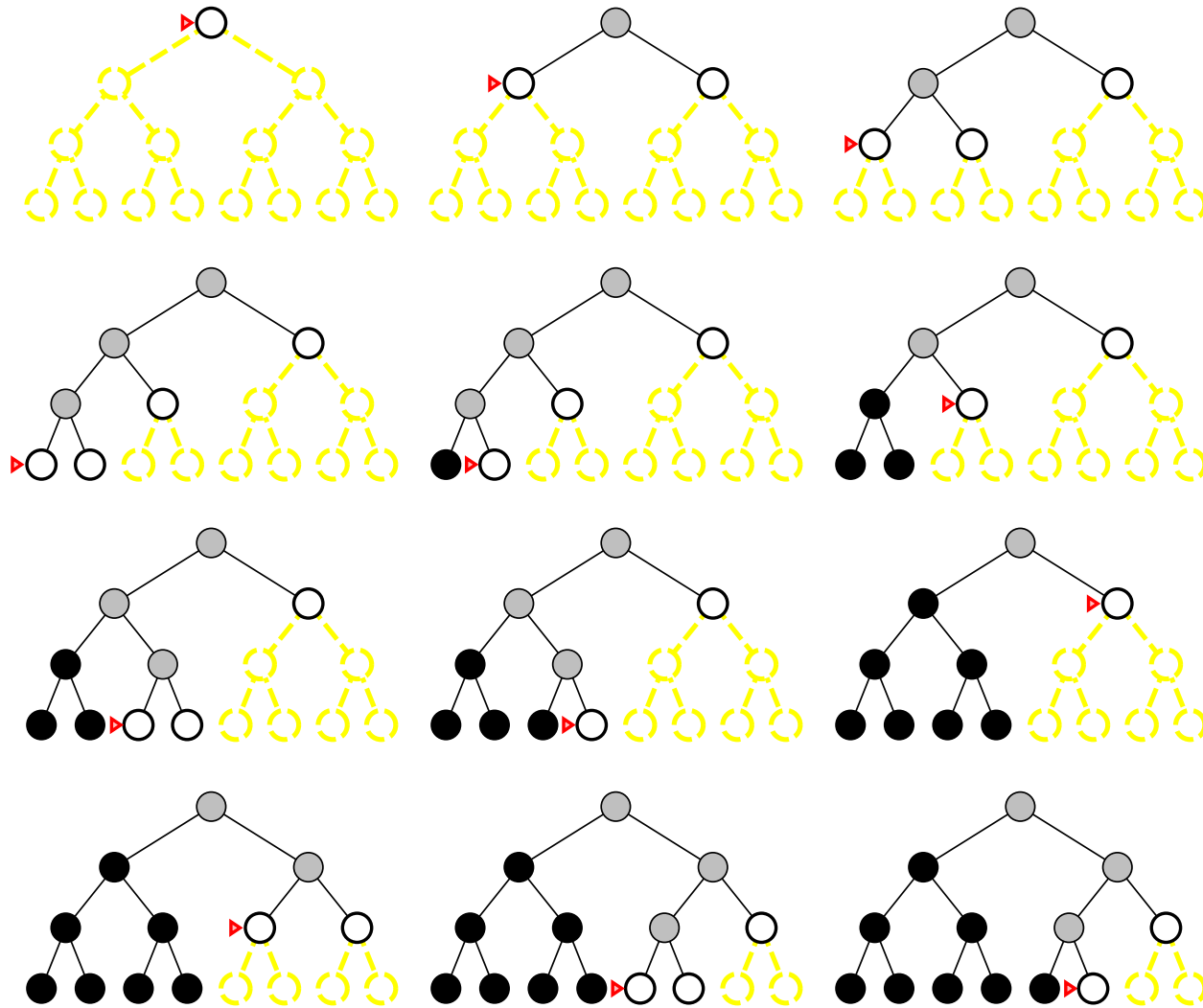
Prohledávání do hloubky – Depth First Search (DFS)



Prohledávání do hloubky – Depth First Search (DFS)



Prohledávání do hloubky – Depth First Search (DFS)





Algoritmus, který neřeší možnost zacyklení:

```
begin
  open := [Start]
  while (open <> []) do begin
    X := first(open)
    open := open - [X]
    if X = GOAL then return(SUCCESS)
    else begin
      E := expand(X)
      open := E + open
    end
  end
  return(failure)
end.
```


Vlastnosti DFS



Algoritmus, který zabraňuje zacyklení za použití seznamu closed:

```
begin
    open := [Start], closed := []
    while (open <> []) do begin
        X := first(open)
        closed := closed + [X], open := open - [X]
        if X = GOAL then return(SUCCESS)
        else begin
            E := expand(X)
            E := E - closed
            open := E + open
        end
    end
    return(failure)
end.
```





Vlastnosti DFS

Mějme b *maximální faktor větvení* (největší počet hran jdoucích z libovolného uzlu) daného stromu, d - nejmenší hloubka stromu, kde se nachází řešení a m maximální hloubka stromu - může být ∞ .

- úplné ?





Vlastnosti DFS

Mějme b maximální faktor větvení (největší počet hran jdoucích z libovolného uzlu) daného stromu, d - nejmenší hloubka stromu, kde se nachází řešení a m maximální hloubka stromu - může být ∞ .

- úplné: NE (i když je b konečné, z důvodu možné existence smyček)
- čas ?



Vlastnosti DFS

Mějme b *maximální faktor větvení* (největší počet hran jdoucích z libovolného uzlu) daného stromu, d - nejmenší hloubka stromu, kde se nachází řešení a m maximální hloubka stromu - může být ∞ .

- **úplné:** NE (i když je b konečné, z důvodu možné existence smyček)
- **čas:** b^m – tzn. exponenciálně podle m , problémy, je-li m výrazně větší než d .





Vlastnosti DFS

Mějme b *maximální faktor větvení* (největší počet hran jdoucích z libovolného uzlu) daného stromu, d - nejmenší hloubka stromu, kde se nachází řešení a m maximální hloubka stromu - může být ∞ .

- **úplné:** NE (i když je b konečné, z důvodu možné existence smyček)
- **čas:** b^m – tzn. exponenciálně podle m , problémy, je-li m výrazně větší než d .
- **paměť ?**





Vlastnosti DFS

Mějme b *maximální faktor větvení* (největší počet hran jdoucích z libovolného uzlu) daného stromu, d - nejmenší hloubka stromu, kde se nachází řešení a m maximální hloubka stromu - může být ∞ .

- **úplné:** NE (i když je b konečné, z důvodu možné existence smyček)
- **čas:** b^m – tzn. exponenciálně podle m , problémy, je-li m výrazně větší než d .
- **paměť:** $O(bm)$





Vlastnosti DFS

Mějme b *maximální faktor větvení* (největší počet hran jdoucích z libovolného uzlu) daného stromu, d - nejmenší hloubka stromu, kde se nachází řešení a m maximální hloubka stromu - může být ∞ .

- **úplné:** NE (i když je b konečné, z důvodu možné existence smyček)
- **čas:** b^m – tzn. exponenciálně podle m , problémy, je-li m výrazně větší než d .
- **paměť:** $O(bm)$
- **optimální ?**





Vlastnosti DFS

Mějme b *maximální faktor větvení* (největší počet hran jdoucích z libovolného uzlu) daného stromu, d - nejmenší hloubka stromu, kde se nachází řešení a m maximální hloubka stromu - může být ∞ .

- **úplné:** NE (i když je b konečné, z důvodu možné existence smyček)
- **čas:** b^m – tzn. exponenciálně podle m , problémy, je-li m výrazně větší než d .
- **paměť:** $O(bm)$
- **optimální:** ne





Alternativní strategie

DL-DFS (Depth-limited) search:

prohledávání do hloubky s omezenou hloubkou prohledávání l

ID-DFS (Iterative deepening) search:

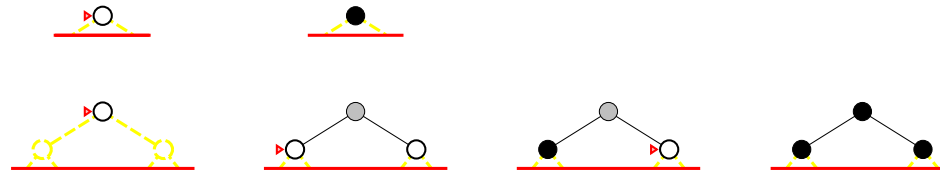
prohledávání do hloubky s iterativní se zvyšující hloubkou prohledávání l

Algoritmus:

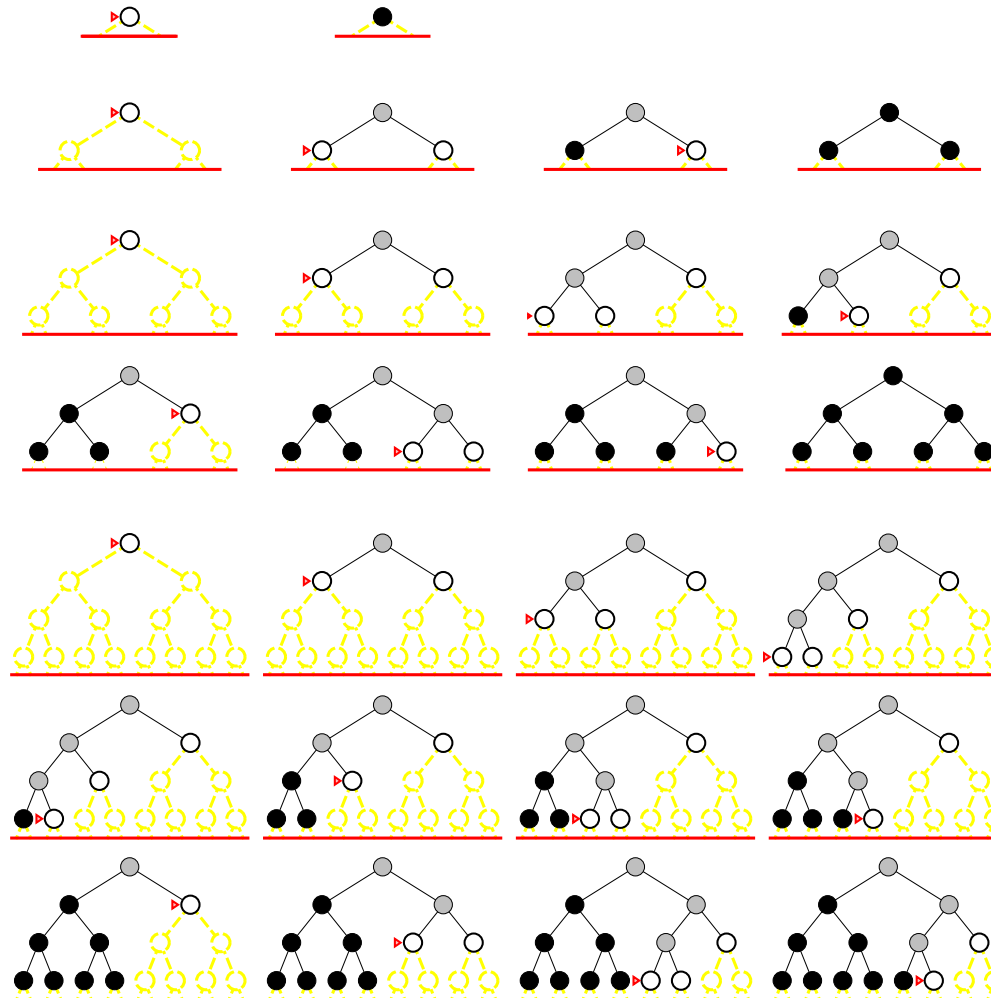
1. $l = 1$
2. proveď DL-DFS s hloubkou l
3. if řešení nalezeno konec
jinak $l = l + 1$ a jdi na 2



Algoritmus IDDFS prohledávání



Algoritmus IDDFS prohledávání



Algoritmus ID-DFS prohledávání



- úplné: ANO (je-li b konečné)



Algoritmus ID-DFS prohledávání



- úplné: ANO (je-li b konečné)
- čas ?





Algoritmus ID-DFS prohledávání

- **úplné:** ANO (je-li b konečné)
- **čas:** $d + 1 + (d)b + (d - 1)b^2 + (d - 2)b^3 + \dots + b^d < db^d = O(b^d)$
předpokládáme-li, že každé jedno prohledávání je realizováno algoritmem o komplexitě $O(b^l)$





Algoritmus ID-DFS prohledávání

- **úplné:** ANO (je-li b konečné)
- **čas:** $d + 1 + (d)b + (d - 1)b^2 + (d - 2)b^3 + \dots + b^d < db^d = O(b^d)$
- **paměť ?**





- **úplné:** ANO (je-li b konečné)
- **čas:** $d + 1 + (d)b + (d - 1)b^2 + (d - 2)b^3 + \dots + b^d < db^d = O(b^d)$
- **paměť:** $O(bd)$





Algoritmus ID-DFS prohledávání

- **úplné:** ANO (je-li b konečné)
- **čas:** $d + 1 + (d)b + (d - 1)b^2 + (d - 2)b^3 + \dots + b^d < db^d = O(b^d)$
- **paměť:** $O(bd)$
- **optimální ?**





Algoritmus ID-DFS prohledávání

- **úplné:** ANO (je-li b konečné)
- **čas:** $d + 1 + (d)b + (d - 1)b^2 + (d - 2)b^3 + \dots + b^d < db^d = O(b^d)$
- **paměť:** $O(bd)$
- **optimální:** ano, optimalizuje-li se hloubka





Algoritmus ID-DFS prohledávání

- **úplné:** ANO (je-li b konečné)
 - **čas:** $d + 1 + (d)b + (d - 1)b^2 + (d - 2)b^3 + \dots + b^d < db^d = O(b^d)$
 - **paměť:** $O(bd)$
 - **optimální:** ano, optimalizuje-li se hloubka
- :: **porovnání:** pro $b = 10$ a $d = 5$ v nejhorším případě:





Algoritmus ID-DFS prohledávání

- **úplné:** ANO (je-li b konečné)
- **čas:** $d + 1 + (d)b + (d - 1)b^2 + (d - 2)b^3 + \dots + b^d < db^d = O(b^d)$
- **paměť:** $O(bd)$
- **optimální:** ano, optimalizuje-li se hloubka

:: **porovnání:** pro $b = 10$ a $d = 5$ je počet expandovaných uzlů nejhorším případě:

- $N(\text{id-dfs}) = 6 + 50 + 400 + 3,000 + 20,000 + 100,000 = 123,456$
- $N(\text{bfs}) = 1 + 10 + 100 + 1,000 + 10,000 + 100,000 = 111,111$

ID-DFS expanduje pouze o cca 11% uzlů, což se díky výrazným úsporám paměti vyplatí.





kritérium/algorithmus	BFS	DFS	DL-DFS	ID-DFS	BiDir
čas	b^d	b^m	b^l	b^d	$b^{\frac{d}{2}}$
paměť	b^d	bm	bl	bd	$b^{\frac{d}{2}}$
optimalita	ano	ne	ne	ano	ano
úplnost	ano	ne	ano (pro $l \geq d$)	ano	ano

kde b je faktor větvení, d je hloubka ve které se nachází nejmělkčí řešení, m je maximální hloubka stromu, l je mez prohledávání.

