Sztuczna inteligencja. A* (cd). Problemy więzowe

Paweł Rychlikowski

Instytut Informatyki UWr

21 marca 2024

Część 1

Przeszukiwanie z z wiedzą o problemie (informed search). Dokończenie

Algorytm A*. Przypomnienie

Definicje

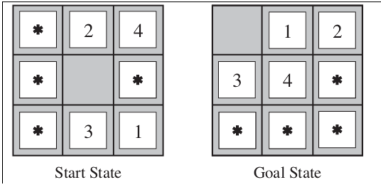
- $g(n)$ – koszt dotarcia do węzła n
- $h(n)$ – szacowany koszt dotarcia od n do (najbliższego) punktu docelowego ($h(s) \geq 0$)
- $\mathrm{f(n)} = \mathrm{g(n)} + \mathrm{h(n)}$

Algorytm

Przeprowadź przeszukanie, wykorzystując $f(n)$ jako priorytet węzła (czyli rozwijamy węzły od tego, który ma najmniejszy f).

Heurystyki możemy budować korzystając z baz wzorców, zapamiętujących koszty rozwiązań podproblemów danego zadania.

Przykład:



Działanie bazy wzorców

- Znajdujemy wszystkie podproblemy dla danego stanu (które mamy w bazie)
- A następnie bierzemy maksimum kosztów jako wartość heurystyki
- Możemy do tego maksimum dołożyć jakieś proste heurystyki (typu h_2).

Pytanie

A czy nie moglibyśmy użyć sumowania, zamiast maksimum?

Działanie bazy wzorców (2)

- Niestety suma daje niedopuszczalne heurystyki (bo pewne ruchy liczymy wielokrotnie, gwiazdki w jednym wzorcu są istotnymi kafelkami w innym)
- **Pytanie:** Jak temu zapobiec?
- **Odpowiedź:** stosując „rozłączne” wzorce (nic się nie powtarza) i w każdym wzorcu liczyć tylko ruchy kafelków z liczbami.

  To to są te najefektywniejsze heurystyki dla 8-ki

Wyszukiwanie dróg (Google Maps, etc)

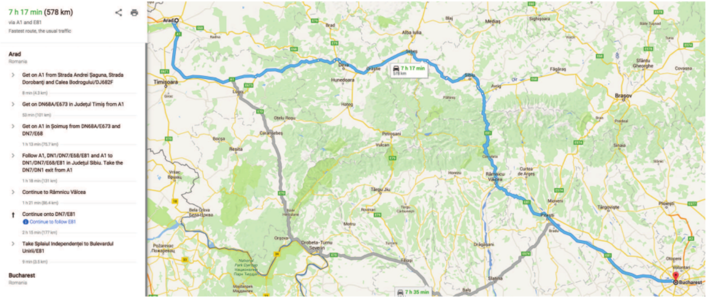


Figure 3.28 A Web service providing driving directions, computed by a search algorithm.

Charakterystyka:

- Duży graf, dany explicite
- Wynik potrzebny od zaraz
- Możliwy preprocessing (bo zapytania mogą być podobne do siebie)

Punkty orientacyjne

Landmarks

Zakładamy istnienie pewnej liczby **punktów orientacyjnych**, będziemy się do nich odwoływać przy liczeniu heurystycznej odległości między węzłami

Zakładamy, że policzyliśmy odległości między **każdym** **p.o.** a **każdym** węzłem.

Heurystyka „naturalna"

$$h(n) = \min_{L \in \text{Landmarks}}(C^*(n, L) + C^*(L, \text{goal}))$$

Ogólnie nie jest optymistyczna, chyba że ścieżka przechodzi przez *jakiś* punkt orientacyjny!

Differential heuristic

Differential heuristic

$h(n) = \max_{L \in \text{Landmarks}}(C^*(n, L) - C^*(L, \text{goal}))$
(jak wyjdzie ujemna, to dajemy 0)

- Myślimy, że punkt orientacyjny jest **za celem** (jadąc do L mijamy cel po drodze)
- Jak cel jest trochę z boku drogi, to tracimy dokładność, ale nie optymizm.

Heurystyki niedopuszczalne

- Heurystyki mogą być niedopuszczalne (w szczególności, jeżeli są wynikiem uczenia się heurystyk)
- Oczywiście tracimy wówczas (w teorii i praktyce) gwarancje optymalności.
- Ale można otrzymać istotnie szybsze wyszukiwanie (o czym, mam nadzieję, przekonamy się na pracowni 2)

Rezygnacja z optymizmu

Weighted A^* search

$f(n) = g(n) + W \times h(n)$, dla $W > 1$

Różne warianty W, dla podsumowania

- A^*: $g(n) + h(n)$, czyli $W = 1$
- **Uniform-cost search**: $g(n)$, czyli $W = 0$
- **Greedy best-first search**: $h(n)$, czyli $W = \infty$
- **Weighted A^* search**

Przykładowe wyniki Weighted A*

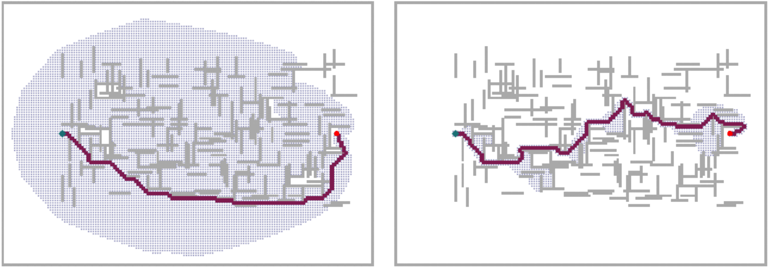

(a) (b)

Figure 3.21 Two searches on the same grid: (a) an A* search and (b) a weighted A* search with weight $W = 2$. The gray bars are obstacles, the purple line is the path from the green start to red goal, and the small dots are states that were reached by each search. On this particular problem, weighted A* explores 7 times fewer states and finds a path that is 5% more costly.

Własności A^*

Plan

Spróbujemy dowieść następujących rzeczy:

1. A^* zwraca najkrótszą drogę
2. A^* jest zupełny.

Dowód optymalności

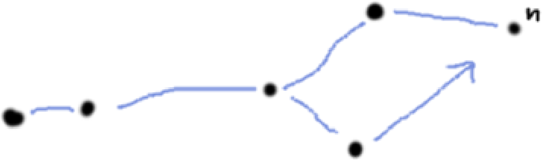
Potrzebujemy dwóch faktów:

F1. Jeżeli h jest spójna, wówczas na każdej ścieżce wartości f są niemalejące.
D-d (n' jest następnikiem n):

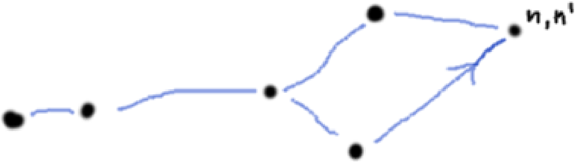
$$f(n') = g(n') + h(n') = g(n) + c(n, n') + h(n') \geq g(n) + h(n) = f(n)$$

F2. Zawsze, gdy algorytm bierze węzeł do rozwinięcia, to koszt dotarcia do tego węzła jest optymalny (najmniejszy możliwy).

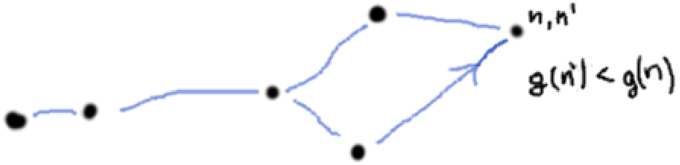
Dowód F2



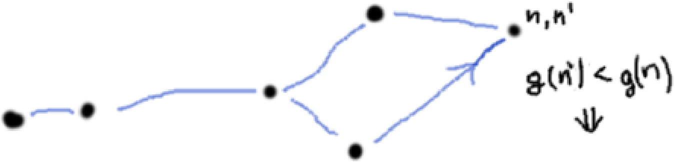
Dowód F2



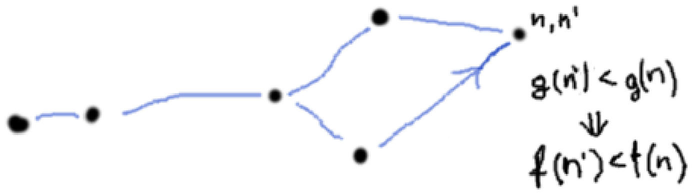
Dowód F2



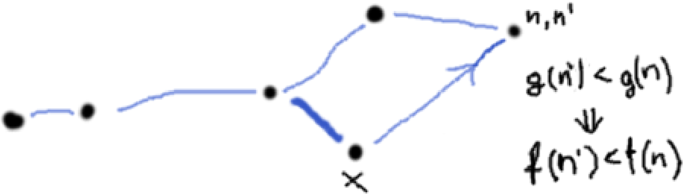
Dowód F2



Dowód F2

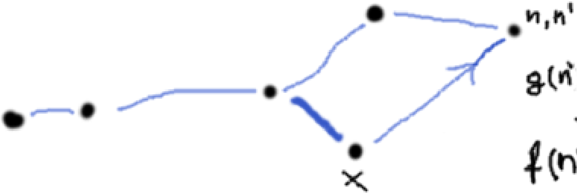


Dowód F2



z F1: $f(x) \leq f(n') < f(n)$

Dowód F2



$g(n') < g(n)$

$\Downarrow$

$f(n') < f(n)$

z F1: $f(x) \leq f(n') < f(n)$

SPRZECZNOŚĆ

Dowód F2

- Bierzemy nie wprost węzeł n, do którego kolejne dotarcie daje mniejszy koszt niż dotarcie pierwsze.
- Wartość f dla tego węzła drugi raz jest mniejsza (h takie same, g mniejsze)
- Na ścieżce od początku do n' drugiego mamy widziany w pierwszym przebiegu węzeł x (tuż po rozgałęzieniu)
- Z własności F1 mamy: $f(x) < f(n)$

Zatem algorytm powinien wybrać $x = x_1$ przed n.
Podobnie x_2, x_3, ... **Sprzeczność.**

Optymalność

Popatrzmy na pierwszy znaleziony węzeł docelowy (n_{end})

- $f(n_{\text{end}}) = g(n_{\text{end}}) + h(n_{\text{end}}) = g(n_{\text{end}})$ (bo h jest rozsądna)
- Każdy kolejny węzeł docelowy jest nielepszy, bo dla niego $f(n) \geq f(n_{\text{end}})$
  (czyli $g(n) \geq g(n_{\text{end}})$, bo h dla końcowych jest 0)

Zupełność

Niech C^* będzie kosztem najtańszego rozwiązania ($g(n_{\text{end}})$)

- Algorytm ogląda wszystkie węzły, dla których $f(n) < C^*$
- Być może oglądnie również pewne węzły z konturu $f(n) = C^*$, przed wybraniem docelowego n, t.że $f(n) = g(n) + 0 = C^*$

Uwaga

Skończona liczba węzłów o $g(n) \leq C^*$ gwarantuje to, że algorytm się skończy.

Do skończoności z kolei wystarczy na przykład założyć, że

- istnieje $\varepsilon > 0$, t.że wszystkie koszty są od niego większe bądź równe,
- węzły mają skończoną liczbę dzieci.

Lepsze heurystyki

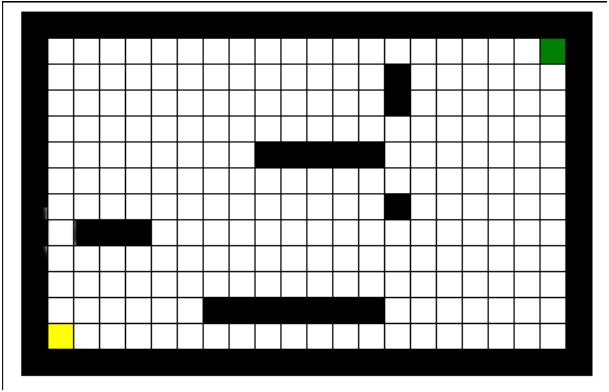
Uwaga

A^* nie rozwija węzłów t.że $f(n) > C^*$. Zatem im większa h (przy założeniu spełniania warunków dobrej heurystyki), tym lepsza.

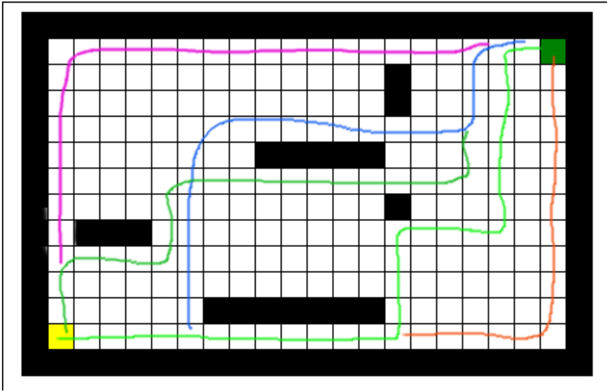
Pytanie

Co się będzie działo, jeżeli nasza funkcja h będzie liczyła dokładną odległość od celu?

Bierzemy heurystykę Manhatańską (przyjmijmy, że cel jest jeden), czyli $h(n) = |g_x - n_x| + |g_y - n_y|$



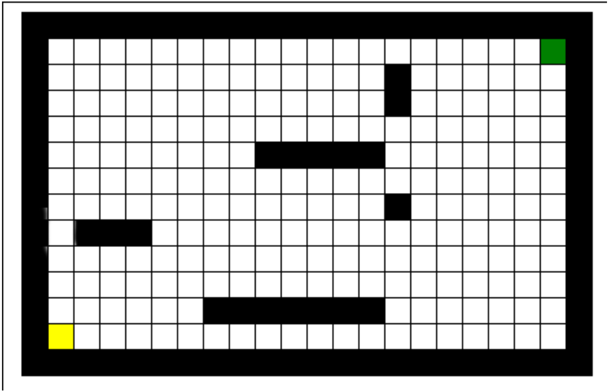
Płaska funkcja f



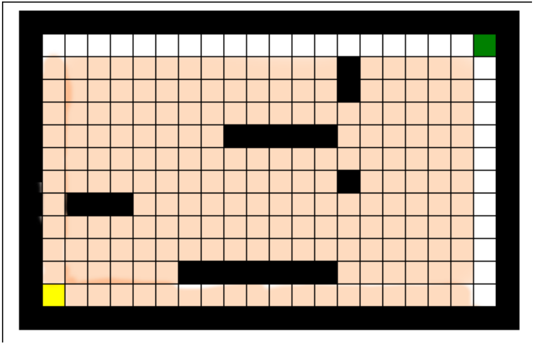
Wszystkie ścieżki idące w prawo i do góry są optymalne. Funkcja f jest stała (równa 30).

Porównanie heurystyk

Porównajmy na przykładzie heurystykę manhatańską i euklidesową. Które węzły **na pewno** musi obejrzeć A^* z heurystyką Euklidesową?



Uzasadnienie lepszości heurystyki manhattańskiej



- Funkcja f z heurystyką h_M ma wszędzie wartość 30
- Funkcja f z heurystyką h_E osiąga wartość 30 jedynie na brzegach.

Algorytm

Przeprowadź przeszukanie, wykorzystując $f(n)$ jako priorytet węzła (czyli rozwijamy węzły od tego, który ma najmniejszy f).

Kluczowa właściwość

1. A^* rozwija wszystkie węzły t.że $f(n) < C^*$.
2. A^* rozwija niektóre węzły t.że $f(n) = C^*$
3. A^* nie rozwija węzłów t.że $f(n) > C^*$.

Część 2

Problemy więzowe (CSP = Constraint Satisfaction Problem)

Problem spełnialności więzów

Uwaga

Między ósemką a hetmanami jest istotna różnica (mimo, że oba można przedstawiać jako problemy przeszukiwania).

- W ósemce interesuje nast droga dotarcia do celu, który jest dobrze znany (i tym samym mało ciekawy)
- W hetmanach interesuje nas, jak wygląda cel – droga do niego może być dość trywialna (dostawianie po kolei poprawnych hetmanów, przestawianie hetmanów z losowego ustawienia).

Problem spełnialności więzów (2)

Problemy takie jak hetmany są:

a) Bardzo istotne (ze względu na ich występowanie w rzeczywistym świecie)

b) Na tyle specyficzne, że warto dla nich rozważać specjalne metody.

Problemy spełnialności więzów. Definicja

Definicja

Problem spełnialności więzów ma 3 komponenty:

1. Zbiór zmiennych $X_1, \ldots, X_n$
2. Zbiór dziedzin (przypisanych zmiennym)
3. Zbiór więzów, opisujących dozwolone kombinacje wartości jakie mogą przyjmować zmienne.

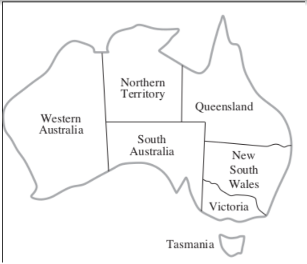
Przykład

Zmienne: X, Y, Z
Dziedziny: $X \in \{1, 2, 3, 4\}, Y \in \{1, 2\}, Z \in \{4, 5, 6, 7\}$
Więzy: $X + Y \geq Z, X \neq Y$

Komentarz do definicji

- Powyższy przykład to były więzy na dziedzinach skończonych, jeden z najważniejszych przypadków więzów.
- Ale można rozważać inne dziedziny:
  a) liczby naturalne, (trochę boli, że to nierozstrzygalny problem)
  b) liczby wymierne,
  c) ciągi elementów, napisy
  d) krotki
- Więzy określają relacje, często da się je wyrazić wzorem, ale nie jest to wymagane.

Kolorowanie Australii (inny przykład zadania więzowego)

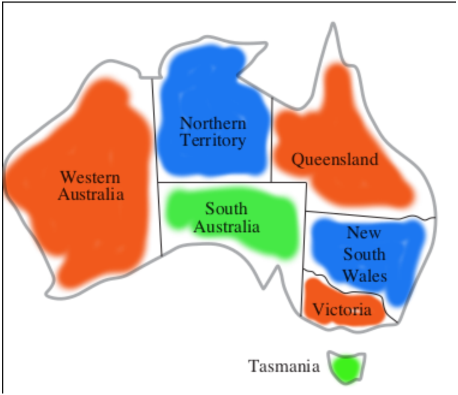


- Mamy pokolorować mapę Australii, za pomocą 3 kolorów: (R, G, B)
- Sąsiadujące prowincje muszą mieć różne kolory.

Kolorowanie jako problem więzowy

Zmienne: WA, NT, Q, NSW, V, SA, T
Dziedziny: {R,G,B}
Więzy: $SA \neq WA, SA \neq NT, SA \neq Q, SA \neq NSW, SA \neq V, WA \neq NT, NT \neq Q, Q \neq NSW, NSW \neq V$

Przykładowe kolorowanie

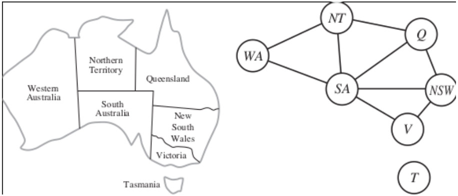


Arność więzów

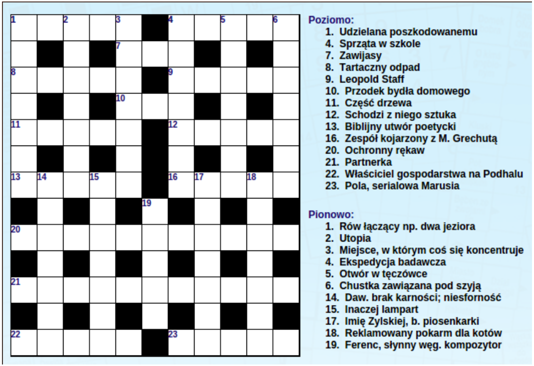
- Więzy (jako relacje) mogą mieć różną arność.
- Unarne – można potraktować jako modyfikację dziedziny (Tasmania nie jest czerwona) i zapomnieć.
- Binarne – jak w naszym przykładzie z kolorowaniem
- Mogą mieć też inną arność, w zasadzie dowolną (w praktyce spotyka się więzy o arności np. kilkaset)

Graf więzów

- Dla więzów binarnych możemy stworzyć graf, w którym krawędź oznacza, że dwie zmienne są powiązane więzem.
- Więzy binarne są istotną klasą więzów, wiele algorytmów działa przy założeniu binarności więzów.



Problemy dualne



Pytanie

Co powinno być zmienną w zadaniu rozwiązywania krzyżówki?

Krzyżówka (podstawowa)

Pomijamy (chwilowo?) kwestie zgodności hasła z definicją.

- Zmienne odpowiadają kratkom, dziedziną są znaki
- Więzy (fragment):
  jest-słowem-7(A,B,C,D,E,F,G), jest-słowem-3(B,H,I), ...

Krzyżówka (dualna)

- Zmienne to słowa (dziedziną jest słownik przycięty do określonej długości)
- Mamy więz dla każdej pary krzyżujących się słów. Jaki?

Więz dla słów

Przykładowy więz $C(w_1, w_2)$:

- w_1 ma długość 6 (dziedzina)
- w_2 ma długość 10 (dziedzina)
- trzeci znak w_1 jest taki sam, jak piąty znak w_2

Problemy dualne (2)

- Zwróćmy uwagę, że to, co zrobiliśmy z krzyżówką stosuje się do dowolnych więzów.
- Więzy w problemie prymarnym zmieniają się na zmienne w problemie dualnym (z dziedziną będącą dozwolonym zbiorem krotek)
- Dodatkowo potrzebujemu więzów, które mówią, że i-ty element jednej krotki jest j-tym elementem drugiej (te więzy są binarne!)

Uwaga 1

Tym, co odróżnia CSP od zadania przeszukiwania, jest możliwość wykorzystania dodatkowej wiedzy o charakterze problemu do przeprowadzenia wnioskowania.

Uwaga 2

Podstawowym celem wnioskowania jest zmniejszenie rozmiaru dziedzin (a tym samym zmniejszenie przestrzeni przeszukiwań).

Przykład

Zmienne: X, Y, Z
Dziedziny: $X \in \{1, 2, 3, 4\}, Y \in \{1, 2\}, Z \in \{5, 6, 7, 8\}$
Więzy: $X + Y \geq Z, X \neq Y$

Czy możemy nie tracąc żadnego rozwiązania skreślić jakieś wartości z dziedzin?

Możemy wywnioskować, że: $X \in \{3, 4\}, Y \in \{1, 2\}, Z \in \{5, 6\}$

Jak widać, możemy też skreślić więz $X \neq Y$

Spójność więzów

- Spójność (intuicyjnie) rozumiemy jako niemożność wykreślenia żadnej wartości z dziedziny.
- Mamy różne rodzaje spójności:
  1. Węzłowa (każda wartość z dziedziny spełnia więzy unarne dla zmiennych)
  2. Łukowa: jak dwie zmienne są połączone więzem, to dla każdej wartości z dziedziny X jest wartość w dziedzinie Y, t.że dla tych wartości więz jest spełniony.

Uwaga

Są jeszcze inne rodzaje spójności. Więcej na ćwiczeniach.

Spójność więzów. Przykład

Więz: $X < Y$
Dziedzina X: $\{4, 6, 7, 10, 20\}$
Dziedzina Y: $\{1, 2, 4, 6, 7, 10\}$
Brak spójności

- Jeżeli weźmiemy X, to możemy wykreślić wartośći 10, 20
- Jeżeli weźmiemy Y, to możemy wykreślić wartości 1,2, 4

Po wykreśleniu tych wartości warto przyjrzeć się innym więzom z X i Y.