Acta Universitatis Palackianae Olomucensis. Facultas Rerum Naturalium. Mathematica

Petr Lisoněk

Heuristic search algorithms on directed trees

Acta Universitatis Palackianae Olomucensis. Facultas Rerum Naturalium. Mathematica, Vol. 27 (1988), No. 1, 347--354

Persistent URL: `http://dml.cz/dmlcz/120204`

ACTA UNIVERSITATIS PALACKIANAE OLOMUCENSIS
FACULTAS RERUM NATURALIUM
1988 MATHEMATICA XXVII VOL. 91

Katedra výpočetní techniky
přírodovědecké fakulty Univerzity Palackého v Olomouci
Vedoucí katedry: RNDr. František Koliba, CSc.

HEURISTIC SEARCH ALGORITHMS ON DIRECTED TREES

PETR LISONĚK

(Received March 31, 1987)

1. Introduction

This paper deals with the problem of finding a minimal-cost path in a directed graph from a starting node to a set of goal nodes. We will examine the relationship between two well-known algorithms of the heuristic search - A^* (Hart, Nilsson and Raphael, 1968) and B' (Merö, 1982) both described in [2] (see also [3]).

Let G be a directed graph with a starting node s, a set of goal nodes $\underline{T}$ and a positive cost c(p,q) associated with every arc (p,q). We shall introduce the functions f(n), g(n), h(n) together with their estimates $\hat{f}(n)$, $\hat{g}(n)$, $\hat{h}(n)$ in their usual sense. Suppose the heuristic estimate $\hat{h}(n)$ to satisfy the conditions

$$\hat{h}(t) = 0 \qquad \text{for every } \underline{t} \in T \tag{1}$$

$0 \leq \hat{h}(n) \leq h(n)$ for every node $\underline{n}$ in the graph G (2)

sufficient for both A^* and B' to be admissible.

B' comes from A^*. It tries to improve the heuristic estimates for the nodes m,n where the consistency assumption

$$\hat{h}(m) - \hat{h}(n) \leq c(m,n)$$

does not hold. This is done by the following formulas:

a) For each son $\underline{m}$ of the recently selected (expanded) node $\underline{n}$, if $\hat{h}(m) < \hat{h}(n) - c(n,m)$ holds, then set $\hat{h}(m) \longleftarrow \hat{h}(n) - c(n,m)$.

b) Let $\underline{m}$ be the son of $\underline{n}$ for which $\hat{h}(m) + c(n,m) = \hat{h}_m(n)$ is minimal. If $\hat{h}_m(n) > \hat{h}(n)$, then set $\hat{h}(n) \longleftarrow \hat{h}_m(n)$.

It was shown in [2], that for every natural N there is a graph G_N with N nodes, at which A^* requires $O(2^N)$ node expansions. However, B' requires at most $\frac{1}{4}N^2 + O(N)$ node expansions at every graph with N nodes.

On the other hand, there was made no analysis of the graphs at which both A^* and B' require a smaller number of node expansions (say O(N)). We only know that at no graph does B' require more expansions than A^*. This led to the small notice formulated in paragraph 3.

2. Basic concepts

Definition 1. Let G be a directed graph. We shall say that G is of the type Tr, if it is a tree, the starting node being the root of this tree and the set of goal nodes being equal to the set of the leaves of this tree.

Definition 2. Let G be of the type Tr. Denote the minimal-cost path from $\underline{s}$ to a goal node n^* as an optimal path. Indeed, there can be more optimal paths with the same length.

Definition 3. Let G be of the type Tr. Denote by $a^*(G)$ the number of node expansions needed by A^* to find the optimal path in G. Similarly denote by $b'(G)$ the number of node expansions needed by B′ to find the optimal path in G.

Let N be the number of nodes in G. Note that $b'(G) \leq a^*(G) \leq N$ (if both A^* and B′ resolve ties in the same way), because every node in a graph of the type Tr is expanded at most once. Moreover, $\hat{g}(n) = g(n)$ for every node n since there is perfectly one path from s to n.

Definition 4. Let G be of the type Tr. Let $\hat{h}$ be the heuristic estimate satisfying (1), (2). Define the function $\bar{h}$ in continuity to description of B′:

1) $\bar{h}(s) \doteq \hat{h}(s)$ for starting node s
2) for $n \neq s$ let r be the father of n ; set $\bar{h}(n) = \max\left\{ \hat{h}(n), \bar{h}(r) - c(r,n)\right\}$.

The number $\bar{h}(p)$ defined this way is equal to $\hat{h}(p)$ after its (possible) modification according to a). We need not be concerned with the modification in b), since it can influence only the change of the heuristic estimate $\hat{h}(n)$ of the recently expanded node n. Since G is a tree, none of its nodes can be reopened and therefore the values $\hat{h}(n)$ and $\hat{f}(n)$ are no more significant. Furthermore, let us point out that definition 4 is unambiguous since for every node n the modification a) is made by B′ at most once.

Definition 5. For every node u in the graph G of the type Tr set $\bar{f}(u) = \hat{g}(u) + \bar{h}(u)$. Then the value $\bar{f}(u)$ is the estimate of $\hat{f}(u)$ after the modification of the heuristic estimate.

3. Results

Lemma 1. Let n be the son of r in the graph G of the type Tr. Then $\bar{f}(n) = \max\left\{\hat{f}(n), \bar{f}(r)\right\}$ holds.

Proof. (Recall that $\hat{g}(n) = g(n)$.) According to definition 4 we have $\bar{h}(n) = \max\left\{\hat{h}(n), \bar{h}(r) - c(r,n)\right\}$.
Since $g(r) = g(n) - c(r,n)$, then

$$\bar{f}(n) = g(n) + \bar{h}(n) = \max\left\{g(n) + \hat{h}(n), g(n) - c(r,n) + \bar{h}(r)\right\} = \max\left\{\hat{f}(n), \bar{f}(r)\right\} .$$

Lemma 2. Let G be a graph of the type Tr. For every node $\underline{m}$ let $(s = n_0, n_1, \dots, n_k = m)$ be the path in G from the starting node $\underline{s}$ to $\underline{m}$. Then

$$\bar{f}(m) = \max\left\{\hat{f}(n_0),\ \hat{f}(n_1),\ \dots,\ \hat{f}(n_k)\right\} \tag{3}$$

holds.

Proof by induction:

1) $k = 0$: Trivially $\bar{f}(s) = \hat{h}(s) = \hat{f}(s) = \max\left\{\hat{f}(s)\right\} = \max\left\{\hat{f}(n_0)\right\}$.

2) Let (3) be valid for a natural $\underline{k}$. Let $(s = n_0, n_1, \dots, n_k = m,\ n_{k+1} = m')$ be a path in G. According to Lemma 1 we have $\bar{f}(m') = \max\left\{\bar{f}(m),\ \hat{f}(m')\right\} = \max\left\{\max\left\{\hat{f}(n_0),\ \hat{f}(n_1),\ \dots,\ \hat{f}(n_k)\right\},\ \hat{f}(n_{k+1})\right\} = \max\left\{\hat{f}(n_0),\ \hat{f}(n_1),\ \dots,\ \hat{f}(n_{k+1})\right\}$ and the induction step is completed.

Theorem. Let G be a graph of the type Tr. If $b'(G) < a^*(G)$ holds, then on the optimal path found by A^* in G one can find a non-goal node $r \notin T$ such that $\hat{h}(r) = h(r)$. (For the goal nodes $t \in T$ the equality $\hat{h}(t) = h(t) = 0$ follows from the properties of $\hat{h}$.)

Proof. Recall that every node in G is expanded at most once. If $b'(G) < a^*(G)$, then there are some nodes in G expanded by A^* but not by B' - call them "A-nodes". There must be some non-goal A-nodes, since both A^* and B' expand perfectly one goal node. There are also nodes expanded by both algorithms, e.g. the starting node $\underline{s}$. Therefore on each path from $\underline{s}$ to a non-goal A-node one can find a non-goal node $\underline{m}$ such that $\underline{m}$ is expanded by A^* but it is in the OPEN-list of B' when a goal node is expanded by B'. Let $(s = n_0, n_1, \dots, n_k = m)$ be the path in G from $\underline{s}$ to this node. We know from Lemma 2 that $\bar{f}(m) = \max\left\{\hat{f}(n_0),\ \hat{f}(n_1), \dots,\ \hat{f}(n_k)\right\}$.

Since all $n_0, n_1, \ldots, n_k$ are expanded by $A^{\wedge}$, it holds (see [4], chapter 2.4.):

$$\hat{f}(n_i) \leqq f(n^*) = f(s) \qquad \text{for all } i = 0,1,\ldots,k \ .$$

Thus also $\bar{f}(m) \leqq f(n^*)$. On the other hand, $\bar{f}(m) \geqq \bar{f}(n^*) = f(n^*)$ (otherwise $\underline{m}$ should be expanded by B' instead of n^*). Indeed,

$$\bar{f}(m) = f(n^*) \quad . \tag{4}$$

But $\bar{f}(m) = \max\left\{\hat{f}(n_0), \hat{f}(n_1), \ldots, \hat{f}(n_k)\right\}$,
and so there must be $q \in \{0,1,\ldots,k\}$ such that

$$\hat{f}(n_q) = f(n^*) \quad . \tag{5}$$

Let us now distingush two cases:

1) n_q lies on the optimal path found by A^*. Then $f(n^*) = f(n_q)$ and with respect to (5) we have

$$\hat{f}(n_q) = f(n_q) \quad . \tag{6}$$

Hence $g(n_q) + \hat{h}(n_q) = g(n_q) + h(n_q)$ i.e. $\hat{h}(n_q) = h(n_q)$ for a non-goal node n_q, since there is a path $(n_q, n_{q+1}, \ldots, n_k = m)$ from n_q to a non-goal A-node $n_k = m$.

2) n_q does not lie on the optimal path found by A^*. Then let $\underline{v}$ be that node on the optimal path, which is in the OPEN-list of A^* when n_q is expanded by A^*. It must be $\hat{f}(v) \geqq \hat{f}(n_q) = f(n^*)$, on the other hand $\hat{f}(v) \leqq f(v) = f(n^*)$ (it is $\hat{g}(v) = g(v)$, $\hat{h}(v) \leqq h(v)$, i.e. $\hat{f}(v) \leqq f(v)$ and so $\hat{f}(n_q) = \hat{f}(v) = f(v)$. Note that $\underline{v}$ is a non-goal node on the optimal path - otherwise it should be expanded by A^* instead of n_q (it is $\hat{f}(n_q) = \hat{f}(v)$ from the last equality).

In both cases we have shown the existence of a non-goal node $\underline{r}$ (in $\underline{1}$ it was n_q, in $\underline{2}$ the node $\underline{v}$) which lies on the optimal path found by A^* and $\hat{f}(r) = f(r)$ i.e. $\hat{h}(r) = h(r)$. The proof of Theorem is now completed.

4. Conclusion

The only general information concerning $b'(G)$ and $a^*(G)$ we have is that $b'(G) \leq a^*(G)$ for an arbitrary graph G when both algorithms resolve ties in the same way. It was shown in [1], [2], [3] that for "complicated" graphs $a^*(G)$ can be much higher than $b'(G)$ (see paragraph 1).

Therefore it is somewhat interesting that $b'(G) = a^*(G)$ for nearly all trees G. (This follows from our Theorem, since the condition $\hat{h}(r) = h(r)$ for some non-goal node on the optimal path is very restricting when $\hat{h}$ is only assumed to be a non-negative lower bound of h.) It would be perhaps useful for computer practice to examine the relationship between $a^*(G)$ and $b'(G)$ for some other general classes of graphs.

REFERENCES

[1] M a r t e l l i, A.: On the complexity of admissible search algorithms. Artificial Intelligence, 8 (1977), 1-13.

[2] M e r ö, L.: A heuristic search algorithm with modifiable estimate. Artificial Intelligence, 23 (1984), 13-27.

[3] M e r ö, L.: On admissible heuristic search algorithms. Počítače a umelá inteligencia, 1 (1982), 395-408.

[4] N i l s s o n, N.J.: Principles of artificial intelligence. Palo Alto, Tioga Publishing Company 1980. In Russian, Moskva, Radio i svjaz 1985.

ALGORITMY HEURISTICKÉHO HLEDÁNÍ NA ORIENTOVANÝCH STROMECH

Souhrn

Článek se vztahuje k problému nalezení nejkratší cesty v orientovaném grafu pomocí heuristických algoritmů. Je studován vztah mezi algoritmy A^* a B' v případě jejich použití na orientovaných stromech. Výsledkem práce je věta obsahující nutnou podmínku pro to, aby B' vyžadoval k nalezení optimální cesty v orientovaném stromu méně iterací než A^*.

АЛГОРИТМЫ ЭВРИСТИЧЕСКОГО ПОИСКА НА ОРИЕНТИРОВАННЫХ ДЕРЕВЬЯХ

Резюме

Статья относится к проблеме раскрытия пути минимальной стоимости в ориентированном графе использованием эвристических алгоритмов. В ней исследовано отношение между алгоритмами A^* и B' в случае, когда они работают на ориентированных деревьях. В приведенной теореме дается необходимое условие для того, чтобы B' раскрыл оптимальную путь в ориентированном дереве за меньшее число итераций чем A^*.

Author's address:
RNDr. Petr Lisoněk
přírodovědecká fakulta
Univerzity Palackého
Gottwaldova 15
771 46 Olomouc
ČSSR /Czechoslovakia/

Acta UPO, Fac.rer.nat., Vol.91,Mathematica XXVII,1988,347-354.