Truly tight bounds for TSP heuristics

Operations Research Letters 51 (2023) 425–431

Contents lists available at ScienceDirect

Operations Research Letters

journal homepage: www.elsevier.com/locate/orl

Truly tight bounds for TSP heuristics

Cor Hurkens

Eindhoven University of Technology, the Netherlands

ARTICLE INFO

Article history:
Available online 21 June 2023

Keywords:
TSP-heuristics
Worst-case analysis
Matching bounds

ABSTRACT

We present an improved performance analysis of select-and-extend heuristics for the metric traveling salesman problem. Our main contributions concern the Arbitrary Addition and Farthest Addition methods.

1. Introduction

The *traveling salesman problem* (TSP) is a fundamental and well-known problem in combinatorial optimization [10]. An instance of the TSP consists of n cities $1, 2, \ldots, n$ together with non-negative distances $d(i, j)$ for $1 \le i, j \le n$. A *partial tour* is a path that visits each of the cities at most once. A *tour* visits each of the n cities exactly once, and in the end returns to its starting point. A *sub-tour* is a tour on a subset of the cities. The objective in the TSP is to find a tour of minimum length $\sum_i d(i, j_i)$, where j_i is the direct successor of i in the tour.

Throughout this note, we assume that the distances are symmetric and hence satisfy $d(i, j) = d(j, i)$ for all $1 \le i, j \le n$. Moreover, we assume that the distances satisfy the triangle inequality: $d(i, k) + d(k, j) \ge d(i, j)$ for all $1 \le i, j, k \le n$. We refer to this setting as the *metric* TSP. Special cases are the *Euclidean* TSP, where city i has coordinates (x_i, y_i) and $d(i, j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$, and the *graphical* TSP, where cities are the nodes of a graph G, and $d(i, j)$ is the number of edges in a shortest i-j-path in G. Finally we have the *network* TSP, where cities correspond to nodes in some *weighted* graph and distance $d(i, j)$ is the minimum length of an i-j-path.

There is a host of tour constructing heuristics for the metric TSP, ranging from greedy and myopic strategies to more sophisticated approaches that try to capture the overall distance distribution. The seminal paper by Rosenkrantz et al. [12] discussed a wide range of such heuristics and provided a performance analysis of several of them. We first give a brief overview of several methods, and summarize in a table the known results on their worst case performance. We refer to the literature for a more extensive treatment of the methods, and their worst-case analysis. Note that without the assumption of the metric property, no approximation algorithm with finite worst-case ratio exists, unless $P = NP$. Such approximation would be able to distinguish between graphs that have or have not a Hamiltonian tour.

The *nearest neighbor rule* (NNR) is a greedy heuristic. It starts with an arbitrarily chosen city x_1 as partial tour. Then NNR repeats the following step for $k = 1, \ldots, n-1$: if the current partial tour is $x_1, \ldots, x_k$, then let x_{k+1} be the city not yet contained in the partial tour that is closest to x_k; ties are broken arbitrarily. In the end, the NNR tour returns from city x_n to city x_1.

The Greedy-*edge* construction of a tour amounts to sorting all edges by increasing length and then selecting edges in this order provided the resulting edge set is a subset of some Hamiltonian tour.

A wide range of heuristics discussed in [12] use the so-called *select-and-extend* paradigm. These algorithms start with an arbitrarily chosen city x_1 as a sub-tour on one node. Next the following steps for $k = 1, \ldots, n-1$ are repeated: if the current sub-tour is $x_1, \ldots, x_k, x_1$, then let x_{k+1} be the next city *selected* from the remaining cities. The sub-tour is then *extended* by inserting the selected city between some pair of consecutive sub-tour vertices. Different implementations of the *selection* and the *extension* procedure lead to a plethora of heuristics. In this setting, also NNR can be seen as a select-and-extend type of heuristic, where the selected vertex is inserted between the last vertex added and the starting vertex.

For *extending* a sub-tour with a selected vertex we consider two options, *Addition* and *Insertion*. For *Addition* of a selected vertex i into the intermediate sub-tour T, a closest sub-tour vertex j is selected, i.e., $d(i, j) = \min_{k \in V(T)} d(i, k)$. Then vertex i is inserted in the sub-tour next to vertex j, by selecting arbitrarily one of the edges $(j, k) \in E(T)$ and replacing it with $\{(j, i), (i, k)\}$. It is easily verified that the increment in tour costs is bounded by $2d(i, j)$ by the triangle inequality. By abuse of notation, in this note we let both (i, j) and (j, i) represent the edge $\{i, j\}$. For *Insertion* of a

E-mail address: c.a.j.hurkens@tue.nl.

https://doi.org/10.1016/j.orl.2023.06.003
0167-6377/

Table 1
Worst case ratios for TSP heuristics.

Overview worst case ratios TSP-heuristics			
heuristic	upper bound	lower bound	reference
Nearest Neighbor	$\frac{1}{2}(1+\lg^*(n))$	$(1-\epsilon)\frac{1}{2}(1+\lg(n))$	here §2, [12], [7], [5], [11]
Greedy (edge)	$\frac{11}{6}+\frac{1}{3}\lceil {}^{1.5}\log(\lceil \frac{1}{2}n\rceil-2)\rceil$	$\frac{1}{3}\frac{\log(n)}{\log\log(n)}$	[3]
Arbitrary Addition	$\lg^*(n-1)$	$\lg^*(n-1)$	here §3
Farthest Addition	$\approx \lg^*(n)-\frac{1}{2}$	$\approx \lg^*(n)-\frac{1}{2}$	here §4
Nearest Addition	$2(n-1)/n$	$2(n-1)/n$	[9], [12]
Cheapest Addition	$2(n-1)/n$	$2(n-1)/n$	[9], [12]
Arbitrary Insertion	$\lg^*(n-1)$	$\frac{\log(n)}{0.5+\log\log(n)}$	here §3, [1]
Farthest Insertion	$\approx \lg^*(n)-\frac{1}{2}$	6.5	here §4, [6]
Nearest Insertion	$2(n-1)/n$	$2(n-1)/n$	[9], [12]
Cheapest Insertion	$2(n-1)/n$	$2(n-1)/n$	[9], [12]
Nearest Merger	$2(n-1)/n$	$2(n-1)/n$	[9], [12]
Doubled MST	$2(n-1)/n$	$2(n-1)/n$	[9], [12]
Christofides-Serdyukov	$1.5-\frac{1}{2}/\lfloor\frac{1}{2}n\rfloor$	$1.5-\frac{1}{2}/\lfloor\frac{1}{2}n\rfloor$	[2], [9]

selected vertex i into the intermediate sub-tour T, all edges (j,k) in the sub-tour are considered and a detour cost $d(i,j)+d(i,k)-d(j,k)$ is computed. The edge (j,k) with minimum detour cost is then selected, and i is then inserted between j and k. Ties are broken arbitrarily.

In order to *select* the next vertex, we consider four options. As a first, we take any *arbitrary* vertex as the next one. This yields the methods *Arbitrary Addition* and *Arbitrary Insertion*. Alternatively, we can select the node that has *largest* distance to the set of nodes in the current sub-tour. It seems counter-intuitive to extend a sub-tour by a node furthest away, but in this way the whole “area” gets “covered” rather quickly. Outliers get caught first, and less important decisions are postponed. This choice yields *Farthest Addition* and *Farthest Insertion*. The third option would be to select the node that has *smallest* distance to the set of nodes in the current sub-tour. This yields methods called *Nearest Addition* and *Nearest Insertion*. Finally we may select that node for which subsequent Addition or Insertion will lead to the lowest increase in tour costs. Ties are broken arbitrarily. The resulting methods are called *Cheapest Addition* and *Cheapest Insertion*. The latter two methods bear resemblance to building minimum weight spanning trees using Prim-Dijkstra.

In addition we mention some methods that are more focused on the global result. *Nearest Merger* starts with n single node sub-tours. In each step it selects two nearest sub-tours, and merges them into a new sub-tour by exchanging two sub-tour edges against two new edges, at a minimum increase in tour costs. The method relates to the Kruskal method for obtaining a minimum weight spanning tree. *Doubled MST* simply starts from a minimum weight spanning tree, doubles all the edges, finds a Eulerian tour along these edges. After short-cutting paths of length 2 that visit a vertex for a second time, we end up with a tour, the length of which is not more than twice the length of the minimum weight spanning tree we started with. This is again because of the triangle inequality. Finally the method by *Christofides-Serdyukov* actually fine-tunes the Doubled MST method, by realizing that only the odd-degree nodes in the minimum weight spanning tree need to be paired up and connected by a set of extra edges, so as to arrive at a Eulerian graph.

Table 1 gives an overview of known bounds on the worst case performance of these TSP heuristics. The upper bounds are based on tailored analysis, the lower bounds are based on constructions of families of instances, either in the plane, or based on an underlying graph metric. We also give references where proofs can be found. Here the function $\lg(x)$ denotes the base 2 logarithm.

The main result in [12] is a worst-case analysis of NNR, which shows that the worst-case ratio of any reasonable select-and-extend heuristic is at most $O(\lg(n))$. Moreover, it provides a family of TSP-instances for which the length of some NNR tour is a factor $\frac{1}{3}\lg(n)$ times the optimal tour length. This settles the worst-case-ratio of NNR, up to a constant factor.

In this note, we try to fix the constant in the leading term of the worst-case ratio. We tighten the analysis of the worst case performance for a range of select-and-extend heuristics, and provide an upper bound using a subtle modification $\lg^*$ of the base 2 logarithm lg. For *Arbitrary Addition* we provide a lower bound on the worst case ratio that comes arbitrarily close to the improved upper bound, for all $n \geq 3$. The lower bound consists of a family of network TSP instances. For *Farthest Addition* we provide an even tighter analysis of the upper bound on the worst case ratio. Next we provide a matching lower bound on the worst case ratio, again for all $n \geq 3$. The matching bounds come from two distinct families of network TSP instances.

2. Upper bounds on TSP heuristics

Rosenkrantz et al. [12] actually prove that for the metric TSP, the worst case ratio of many heuristics is bounded from above by $\lceil\lg(n)\rceil+1$. This proof is based on a lemma that can slightly be sharpened. We reformulate the lemma and give a useful extension. Our proofs are actually less intricate than the ones proposed by [12].

Definition 1. Let $\lg^* : (0,\infty) \to \mathbb{R}$ be given by: $\lg^*(x) := \lfloor\lg(x)\rfloor - 1 + x \cdot 2^{-\lfloor\lg(x)\rfloor}$, for $x > 0$.

The function $\lg^*$ mimics the standard base 2 logarithm lg. It is concave, continuous and piecewise linear, and coincides with $\lg(x)$ in its breakpoints $x = 2^q$, for $q \in \mathbb{Z}$, only. It satisfies, for $a, b > 0$:

$$\lg^*(a+b) \;=\; 1+\lg^*(\frac{a+b}{2}) \;\geq\; 1+\tfrac{1}{2}\left[\,\lg^*(a)+\lg^*(b)\,\right], \qquad (1)$$

with equality for $a \leq b$ with $\lceil\lg(b)\rceil \leq \lfloor\lg(a)\rfloor + 1$.

Lemma 1. *Let $V = \{1,\ldots,n\}$ be a vertex set, with symmetric distance function $d : V \times V \to \mathbb{R}_+$, satisfying the triangle inequality and a mapping $\beta : V \to \mathbb{R}$ satisfying: $\beta(j) \leq 2 \cdot d(i,j), \forall i < j$. Let Ω denote a cyclic ordering of V (a Hamiltonian cycle), with length $\ell(\Omega)$. Then*

$$\beta(V) - \beta(1) \;\leq\; \lg^*(n) \cdot \ell(\Omega).$$

Here we use the convention that $\beta(V)=\sum_{v\in V}\beta(v)$, for any set V and any mapping $\beta: V\to\mathbb{R}$.

Proof. We will prove, more generally, that for W, $\{1\}\subseteq W\subseteq V$:

$$\beta(W)-\beta(1)\le \lg^*(|W|)\cdot \ell(\Omega_W).$$

Here Ω_W is the cyclic ordering on W defined by the restriction of Ω to W. The proof is by induction on $|W|$. For $|W|\le 2$ the proposition is trivial. The proof is continued for $|W|>2$ by finding some convenient partition of the edge set E_W of Ω_W. For $i\in W$, let s_i denote its successor and p_i its predecessor with respect to Ω_W. Then $E_W=\{(i,s_i)|i\in W\}$. Define $W_1:=\{i\in W|i>s_i\}$ and $W_2:=\{i\in W|i>p_i\}$. Then edge sets $E_1:=\{(i,s_i)|i\in W_1\}$ and $E_2:=\{(p_i,i)|i\in W_2\}$ partition E_W. We then have

$$\beta(W)=\beta(W_1)+\beta(W\setminus W_1)\le 2\cdot d(E_1)+\beta(W\setminus W_1); \qquad (2)$$

$$\beta(W)=\beta(W_2)+\beta(W\setminus W_2)\le 2\cdot d(E_2)+\beta(W\setminus W_2). \qquad (3)$$

Note that W_1 and W_2 are non-empty, that neither of them contains vertex 1, and that $|W\setminus W_1|+|W\setminus W_2|=|W|$. Hence, we can combine (2) and (3), and proceed by induction to find:

$$\begin{aligned}
&\beta(W)-\beta(1)\le d(E_1)+d(E_2)\\
&\quad+\tfrac{1}{2}\cdot[\,\beta(W\setminus W_1)-\beta(1)+\beta(W\setminus W_2)-\beta(1)\,]\\
&\le \ell(\Omega_W)+\tfrac{1}{2}\cdot\big[\lg^*(|W\setminus W_1|)\cdot\ell(\Omega_{W\setminus W_1})\\
&\quad+\lg^*(|W\setminus W_2|)\cdot\ell(\Omega_{W\setminus W_2})\big]\\
&\le (1+\tfrac{1}{2}\cdot\big[\lg^*(|W\setminus W_1|)+\lg^*(|W\setminus W_2|)\big])\cdot\ell(\Omega_W)\\
&\le \lg^*(|W|)\cdot\ell(\Omega_W).
\end{aligned}$$

The last two inequalities are based on the triangle inequality for the distance function, and the concavity of $\lg^*(x)$, respectively, cf. (1). □

If more is known about the function β, the first lemma can be strengthened.

Lemma 2. *Let V, d, β, and Ω satisfy the requirements for Lemma 1 and let, furthermore, $n\ge 3$ and $\beta(2)+\beta(3)\le d(1,2)+d(2,3)+d(3,1)$. Then*

$$\beta(V)-\beta(1)\;\le\;\lg^*(n-1)\cdot\ell(\Omega).$$

Proof. Again, we will prove a slightly more general statement. We show that, for W, $\{1,2,3\}\subseteq W\subseteq V$:

$$\beta(W)-\beta(1)\le\lg^*(|W|-1)\cdot\ell(\Omega_W).$$

The proof is by induction on $|W|$ and by direct use of Lemma 1. For $|W|=3$ we have $\beta(W)-\beta(1)=\beta(2)+\beta(3)\le d(1,2)+d(2,3)+d(3,1)=\ell(\Omega_W)$.

If $E_W\cap\{(1,2),(2,3),(3,1)\}\neq\emptyset$, let $\{a,b,c\}=\{1,2,3\}$, with $(a,b)\in E_W$. We define a set $W':=W\setminus\{1,2,3\}\cup\{1',2'\}$ and a distance function $d': W'\times W'\to\mathbb{R}_+$, with $d'(i,j):=d(i,j)$ for $4\le i,j\le n$; $d'(i,2'):=d(i,c)$ for $4\le i\le n$; $d'(1',2'):=\frac{1}{2}(d(1,2)+d(2,3)+d(3,1))$ and $d'(i,1'):=\min\{d(i,b)+\frac{1}{2}(d(a,b)-d(b,c)+d(c,a)),\ d(i,a)+\frac{1}{2}(d(a,b)+d(b,c)-d(c,a))\}$, for $4\le i\le n$.

Let $\beta'(i)=\beta(i)$ for $i\ge 4$, $\beta'(1')=\beta(1)$, and $\beta'(2')=\beta(2)+\beta(3)$, then clearly $\beta'(j)\le 2\cdot d'(i,j)$ for $i<j$, $i,j\in W'$. For the cyclic ordering Ω' on W', derived from Ω_W by identifying c with $2'$, and $\{a,b\}$ with $1'$ we have: $\ell(\Omega')\le\ell(\Omega_W)$. Applying Lemma 1 we find:

$$\begin{aligned}
\beta(W)-\beta(1)&=\beta'(W')-\beta'(1')\le\lg^*(|W'|)\cdot\ell(\Omega')\\
&\le\lg^*(|W|-1)\cdot\ell(\Omega_W).
\end{aligned}$$

If $E_W\cap\{(1,2),(2,3),(3,1)\}=\emptyset$, then there exist, as in the proof of Lemma 1, edge sets E_1,E_2, and non-empty vertex sets W_1,W_2, neither of them containing vertex 1, 2, or 3, such that

$$\begin{aligned}
&\beta(W)-\beta(1)\le d(E_1)+d(E_2)\\
&\quad+\tfrac{1}{2}\cdot[\,\beta(W\setminus W_1)-\beta(1)+\beta(W\setminus W_2)-\beta(1)\,]\\
&\le\ell(\Omega_W)+\tfrac{1}{2}\cdot\big[\lg^*(|W\setminus W_1|-1)\cdot\ell(\Omega_{W\setminus W_1})\\
&\quad+\lg^*(|W\setminus W_2|-1)\cdot\ell(\Omega_{W\setminus W_2})\big]\\
&\le(1+\tfrac{1}{2}\cdot\big[\lg^*(|W\setminus W_1|-1)+\lg^*(|W\setminus W_2|-1)\big])\cdot\ell(\Omega_W)\\
&\le\lg^*(|W|-2)\cdot\ell(\Omega_W)\le\lg^*(|W|-1)\cdot\ell(\Omega_W).
\end{aligned}$$

Induction can be applied, as both $W\setminus W_1$ and $W\setminus W_2$ contain $\{1,2,3\}$. □

Lemma 2 is used as follows. For any construction heuristic H that works by selection and extension, let the vertices be labeled according to the order in which the heuristic selects them. Let Ω denote an optimal Hamiltonian tour. Take $\beta(1)=0$ and let, for $i>1$, $\beta(i)$ denote the detour cost to insert vertex i into the intermediate tour on vertices $\{1,\dots,i-1\}$. Then, by definition, $\beta(2)=2\cdot d(1,2)$, and $\beta(3)=d(2,3)+d(3,1)-d(1,2)$, so $\beta(2)+\beta(3)=d(1,2)+d(2,3)+d(3,1)$. If furthermore the extension mechanism is such that detour cost $\beta(j)\le 2\cdot d(i,j)$ for $i<j$, we find that the length of the constructed tour T_H satisfies:

$$\ell(T_H)\;=\;\beta(V)\;\le\;\lg^*(|V|-1)\cdot\ell(\Omega)=\lg^*(n-1)\cdot\ell(\Omega). \qquad (4)$$

For NNR we apply Lemma 1. Let the vertices be labeled in *reverse order* of selection, and for $i>1$, let $\beta(i)$ denote twice the cost of adding vertex $i-1$ to the path (i.e., $\beta(i)=2\cdot d(i,i-1)$). Let $\beta(1)$ denote twice the cost of closing the final Hamiltonian path to a tour, i.e., let $\beta(1):=2\cdot d(1,n)$. Then $\beta(1)\le\ell(\Omega)$, $\beta(j)\le 2\cdot d(i,j)$ for $i<j$, hence the Nearest Neighbor tour T_{NN} satisfies:

$$\ell(T_{NN})\;=\;\tfrac{1}{2}\beta(V)\;\le\;\tfrac{1}{2}\cdot(1+\lg^*(n))\cdot\ell(\Omega).$$

It has been shown in [11] that the factor $\frac{1}{2}$ cannot be lowered.

3. Tightness for Arbitrary Addition

We now consider *Arbitrary Addition*. Let the cities be labeled in the order the heuristic selects them. It is easily verified that, if $\beta(j)$ denotes the increment in tour costs by inserting j in the intermediate tour, then β satisfies the requirements for Lemma 2 and thus upper bound (4) applies. The same is true for Arbitrary Insertion. We describe a family of instances of TSP for which Arbitrary Addition yields a tour with a length that is arbitrarily close to $\lg^*(n-1)$ times optimum. Thus we prove that the upper bound is tight.

The instance involves $N=2^M$ vertices lying on the N-cycle C_N with edges of length 1, and one auxiliary vertex, lying at distance $\epsilon<1$ from some vertex on the cycle. Let the vertices be labeled u_i, for $i=0,1,\dots,2^M$, with auxiliary vertex u_0 adjacent to vertex u_{2^M} by an edge of length ϵ, and vertex u_i adjacent to vertex u_{i+1} by an edge of length 1, for $1\le i<2^M$. We now have a graph on $n=2^M+1$ vertices. The instance is finally described by defining for each pair of vertices i,j: $d(i,j)=$ length of the shortest (i,j)-path along the weighted graph. It is obvious that an optimal tour has length $OPT=2^M+2\epsilon$.

We next describe a scenario for the Arbitrary Addition algorithm which will yield a very long tour T_{AA}. The algorithm may

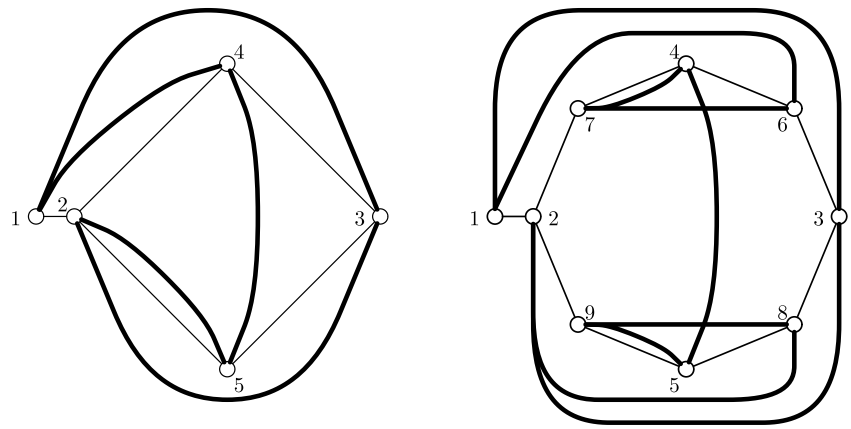

Fig. 1. Worst possible AA tours; nodes indexed by order of selection.

start by selecting vertices u_0, u_{2^M}, and $u_{2^{M-1}}$, yielding a tour of length $2^M + 2\epsilon$. Next, vertices $u_{2^{M-2}}$ and $u_{3.2^{M-2}}$ are selected and inserted. The resulting tour may be

$$(u_0, u_{2^{M-2}}, u_{3.2^{M-2}}, u_{2^M}, u_{2^{M-1}}),$$

of length $2 \cdot 2^M + 2\epsilon$. The remainder of the process is best described as a series of stages, labeled K, $1 \le K \le M-2$. In stage K, the algorithm selects and inserts vertices at distance 2^{M-2-K} from the tour. Let stage $K = 0$ denote the part of the process in which vertices $u_{2^{M-2}}$ and $u_{3.2^{M-2}}$ have been selected.

Lemma 3. *At the start of stage $K \ge 1$, the following invariant holds: If a cycle vertex u_j is contained in the intermediate sub-tour and has been selected in the previous stage $K-1$, it is adjacent to one of its sub-tour neighbors by an edge of length at most $\epsilon + 2^{M-2}$.*

Proof. The proof is by induction on K. For stage $K = 0$ we have that selected nodes $u_{2^{M-2}}, u_{3.2^{M-2}}$ have distance at most $\epsilon + 2^{M-2}$ from their neighbors. This establishes the base of our induction, for $K = 1$. In stage $K \ge 1$, the algorithm selects vertices u_i, at distance 2^{M-2-K} from the tour. Each selected vertex has two nearest neighbors at that distance, one of which has been selected in the previous round. Breaking ties unfortunately, the algorithm may select vertex $u_{j(i)}$, with $u_{j(i)}$ selected in the previous stage, as tour vertex closest to u_i. Let $j(i) = j(i')$, with $i < j(i) < i'$. By induction, vertex $u_{j(i)}$ lies on a "short" tour edge $(u_{j(i)}, u_k)$. Without loss of generality, we may assume that $k < j(i)$. Inserting $u_{i'}$ between u_k and $u_{j(i)}$, and u_i between $u_{i'}$ and $u_{j(i)}$ yields an increase of the tour length with $2 \cdot 2^{M-2-K}$, for each insertion, while the invariant will hold for selected vertices u_i and $u_{i'}$, as they now have distance $2^{M-1-K} \le \epsilon + 2^{M-2}$ from each other. If $k > j(i)$, we first insert u_i and next $u_{i'}$. □

The final tour has length:

$$\begin{aligned} AA =& 2\epsilon + (1 \cdot 2^M + 2 \cdot 2^{M-1} + 4 \cdot 2^{M-2} + \cdots + 2^{M-1} \cdot 2) \\ =& 2\epsilon + M \cdot 2^M. \end{aligned}$$

Hence, the ratio $\frac{AA}{OPT} = \frac{M \cdot 2^M + 2\epsilon}{2^M + 2\epsilon} \to M = \lg^*(n-1)$, for $\epsilon \to 0$.

The above example works for $n = 2^M + 1$, $M \ge 1$. When m vertices are to be added to the instance, with $m < 2^M$, we can do so by subdividing edges of length 1 into two edges of length $\frac{1}{2}$. We do so for m edges adjacent to $\lceil \frac{m}{2} \rceil$ nodes selected in the final stage. The optimal tour length stays the same, but the heuristic will add cost 1 for each additional vertex. Hereby, the ratio increases to $\frac{AA}{OPT} = \frac{m + M \cdot 2^M + 2\epsilon}{2^M + 2\epsilon} \to M + m \cdot 2^{-M} = \lg^*(2^M + 1 + m - 1)$, for $\epsilon \to 0$. Hence, for each $n \ge 3$, a family of instances exists with $AA/OPT \to \lg^*(n-1)$.

For $n = 2^M + 1$, $M = 2, 3$, we show worst-case Arbitrary Addition tours (in bold) in Fig. 1.

4. A tight upper bound for Farthest Addition

We now turn to *Farthest Addition*. It differs from Arbitrary Addition in the way nodes are selected. After selection of a first node, each consecutive node is one that has *largest* distance to the set of nodes in the current sub-tour. It seems counter-intuitive to extend a sub-tour by a node furthest away, but in practice, this selection method seems to work well, and often better than greedy selection mechanisms that try to stay close to the current tour (cf. [4],[8]).

We have shown earlier that a general class of tour constructing heuristics has a worst case ratio of at most $\lg^*(n-1)$. In order to prove a stronger upper bound on the worst case ratio for Farthest Addition and Farthest Insertion, we use a property specific to the farthest vertex selection mechanism. Let the vertices be labeled $1, \dots, n$ in order of selection by the heuristic. Then the extra information about the detour cost $\beta(j)$ for adding vertex j to the intermediate tour is as follows:

$$\text{for } 1 \le i < k \le j \le n: \quad \beta(j) \le 2 \cdot d(i, k). \tag{5}$$

This condition is indeed satisfied by Farthest Addition and Farthest Insertion, as $\frac{1}{2} \cdot \beta(j) \le \min\{d(i', j) | i' < j\} \le \min\{d(i', j) | i' < k\} = d(j, \{1, 2, \dots, k-1\}) \le d(k, \{1, 2, \dots, k-1\}) = \min\{d(i', k) | i' < k\} \le d(i, k)$, for $i < k \le j$. Using (5) we can more easily achieve the bound of Lemma 1.

Lemma 4. *Let $V = \{1, \dots, n\}$ be a vertex set, with symmetric distance function $d : V \times V \to \mathbb{R}_+$, satisfying the triangle inequality and a mapping $\beta : V \to \mathbb{R}$ satisfying: $\beta(k) \le 2 \cdot d(i, j), \forall i < j \le k$. Let Ω denote a cyclic ordering of V (Hamiltonian cycle), with length $\ell(\Omega)$. Then*

$$\beta(V) - \beta(1) \le \lg^*(|V|) \cdot \ell(\Omega).$$

Proof. The statement is trivial for $n \le 2$. For $n \ge 3$, observe that if the edges $\{i, j\} \in E(\Omega)$, with $i > j$, are ordered such that

$$(i_1, j_1, i_2, j_2, \cdots, i_n, j_n)$$

is lexicographically maximal, then

$$n + 1 - \lceil \frac{k}{2} \rceil \ \ge \ i_k, \text{ and } \tfrac{1}{2} \cdot \beta(n + 1 - \lceil \tfrac{k}{2} \rceil) \le d(i_k, j_k), \quad \text{for } 1 \le k \le n. \tag{6}$$

Hence,

$$\begin{aligned}\beta(V)-\beta(1) &= \tfrac{1}{2}\cdot\beta(\{\lfloor\tfrac{n}{2}\rfloor+1,\dots,n\})\\ &+\tfrac{1}{2}\cdot\beta(\{\lceil\tfrac{n}{2}\rceil+1,\dots,n\})\\ &+\tfrac{1}{2}\cdot(\beta(\{1,2,\dots,\lfloor\tfrac{n}{2}\rfloor\})-\beta(1))\\ &+\tfrac{1}{2}\cdot(\beta(\{1,2,\dots,\lceil\tfrac{n}{2}\rceil\})-\beta(1))\\ &\le \ell(\Omega)+\tfrac{1}{2}\cdot(\lg^*(\lfloor\tfrac{n}{2}\rfloor)+\lg^*(\lceil\tfrac{n}{2}\rceil))\cdot\ell(\Omega)\\ &=\lg^*(n)\cdot\ell(\Omega).\end{aligned} \tag{7}$$

The inequality is based on induction on n, the triangle inequality, and the observation in (6) that $\frac{1}{2}\beta(n)\le d(i_1,j_1)$, $\frac{1}{2}\beta(n)\le d(i_2,j_2)$, $\frac{1}{2}\beta(n-1)\le d(i_3,j_3)$, $\frac{1}{2}\beta(n-1)\le d(i_4,j_4)$, etcetera. □

Lemma 5. *Let V, d, β, and Ω satisfy the requirements for Lemma 4, and let, furthermore, $n\ge 3$ and*

$$\beta(\{1,2,3\})\le d(1,2)+d(2,3)+d(3,1), \tag{8}$$

then

$$\frac{\beta(V)}{\ell(\Omega)}\le\begin{cases}(a)\ \lg^*(n)-0.50,\\ \quad\text{if } n=3,4 \text{ or } \exists q\in\mathbb{Z}_+\left[6\cdot 2^q\le n\le 9\tfrac{1}{3}\cdot 2^q\right];\\ (b)\ \lg^*(n)-0.45-0.05\cdot m\cdot 2^{-q},\\ \quad\text{if } n=5\cdot 2^q+m,\ 0\le m\le 2^q,\text{ for } q=\lfloor\lg(\tfrac{n}{5})\rfloor;\\ (c)\ \lg^*(n)-0.45-0.075\cdot m\cdot 2^{-q},\\ \quad\text{if } n=5\cdot 2^{q+1}-m,\ 0\le m\le\tfrac{2}{3}\cdot 2^q,\text{ for } q=\lceil\lg(\tfrac{n}{10})\rceil.\end{cases} \tag{9}$$

These bounds are tight, which will be shown by constructing two families of instances, $\mathcal{A}=(A_n)_{n\ge 3}$ and $\mathcal{B}=(B_n)_{n\ge 5}$, with matching Farthest Addition results satisfying:

$$FA/OPT(A_n)=\lg^*(n)-0.5,\quad\text{for } n\ge 3; \tag{10}$$

$$FA/OPT(B_n)=\lg^*(n/5)+1.8 \tag{11}$$

$$=q+1.8+0.2\cdot m\cdot 2^{-q},\quad\text{for } n=5\cdot 2^q+m,\ 0\le m\le 5\cdot 2^q,\ q\ge 0. \tag{12}$$

The right hand side in (10) coincides with bound (9)(a), and the right hand side in (11) coincides with bounds (9)(b) and (c). Note that $\lg(5)\approx 2.3219$, which implies that $(\lg(n)-0.5)-(\lg(n/5)+1.8)\approx 0.0219$. The graphs of $(\lg^*(n)-0.5)$ and $(\lg^*(n/5)+1.8)$ do intersect infinitely often due to the linear interpolations. For $n=6\cdot 2^q$, with $q\in\mathbb{Z}$, both attain value $q+2$; for $n=\frac{28}{3}\cdot 2^q$ they both attain value $q+\frac{8}{3}$. The upper bound (9) coincides with the maximum of the two functions.

Proof. We will now prove property (9) for any $\beta: V\to\mathbb{R}_+$ satisfying inequality (5) and property (8), and for any cyclic ordering Ω of the node set V. Let Ω_k denote the Hamiltonian cycle obtained by restricting Ω to vertex set $\{1,2,\dots,k\}$. By the triangle inequality we have that $\ell(\Omega_k)\le\ell(\Omega_{k+1})$, for all k. The proof breaks down in a number of small reasoning steps. Each reasoning step starts with a single statement, and is followed by a short argument or proof. Steps ⟨1⟩–⟨8⟩prove the main part of bound (9)(a). In particular ⟨3⟩ already covers it for n with $6\le n\cdot 2^{-q}\le 8$. Case distinctions ⟨5⟩,⟨6⟩ are needed for the special case $n=9$. Step ⟨9⟩ covers upper bound (9)(b). Case distinctions ⟨10⟩,⟨11⟩ enable to seek the boundary between bounds (9)(c) and (9)(a), as laid out in ⟨12⟩ and ⟨13⟩.

⟨1⟩ $\beta(\{1,2,3\})\le d(1,2)+d(2,3)+d(3,1)=\ell(\Omega_3)$.

⟨2⟩ $\beta(\{1,2,3,4\})\le 1.5\cdot\ell(\Omega_4)$, since $\beta(4)\le\frac{1}{2}\cdot\Sigma_{e\in E(\Omega_4)}\ d(e)=0.5\cdot\ell(\Omega_4)$.

⟨3⟩ For n with $3\cdot 2^q\le n\le 4\cdot 2^q$, $q\ge 0$: $\beta(\{1,2,\dots,n\})\le(\lg^*(n)-0.5)\cdot\ell(\Omega)$.
This is true for $n=3,4$, by ⟨1⟩,⟨2⟩; it follows for $q>0$ by induction. If $3\cdot 2^q\le n\le 4\cdot 2^q$, then $3\cdot 2^{q-1}\le\lfloor\frac{n}{2}\rfloor\le\lceil\frac{n}{2}\rceil\le 4\cdot 2^{q-1}$, and so, analogous to (7):

$$\begin{aligned}\beta(\{1,2,\dots,n\}) &\le \ell(\Omega)+\tfrac{1}{2}\cdot\beta(\{1,\dots,\lfloor\tfrac{n}{2}\rfloor\})\\ &\quad+\tfrac{1}{2}\cdot\beta(\{1,\dots,\lceil\tfrac{n}{2}\rceil\})\\ &\le \ell(\Omega)+\tfrac{1}{2}\cdot(\lg^*(\lfloor\tfrac{n}{2}\rfloor)-0.5)\cdot\ell(\Omega)\\ &\quad+\tfrac{1}{2}\cdot(\lg^*(\lceil\tfrac{n}{2}\rceil)-0.5)\cdot\ell(\Omega)\\ &=(\lg^*(n)-0.5)\cdot\ell(\Omega).\end{aligned} \tag{13}$$

⟨4⟩ $\beta(\{1,2,3,4,5\})\le 1.80\cdot\ell(\Omega_5)$, since property (5) yields: $5\cdot\beta(\{4,5\})\le 4\cdot\ell(\Omega_5)$.

⟨5⟩ If $\beta(\{1,2,3,4,5\})\le 1.75\cdot\ell(\Omega_5)$, then $\beta(\{1,2,\dots,n\})\le(\lg^*(n)-0.5)\cdot\ell(\Omega_n)$, for $n\ge 3$.
The statement is true for $n=3,4$ by ⟨1⟩ and ⟨2⟩; it is true for $n=5$, by assumption; and follows for $n\ge 6$ by induction, as $n\ge 6$ implies $n>\lceil\frac{n}{2}\rceil\ge\lfloor\frac{n}{2}\rfloor\ge 3$, so (13) applies.

⟨6⟩ If $\beta(\{1,2,3,4,5\})>1.75\cdot\ell(\Omega_5)$, then $d(i,j)>0.125\cdot\ell(\Omega_5)$, for all $1\le i<j\le 5$.
Proof: if $d(i,j)\le 0.125\cdot\ell(\Omega_5)$, then $\beta(5)\le 2\cdot d(i,j)\le 0.25\cdot\ell(\Omega_5)$, contradicting ⟨2⟩.

⟨7⟩ $\beta(\{1,\dots,9\})\le 2.625\cdot\ell(\Omega_9)=(\lg^*(9)-0.5)\cdot\ell(\Omega_9)$.
If $\beta(\{1,2,3,4,5\})\le 1.75\cdot\ell(\Omega_5)$, then the statement follows directly from ⟨5⟩. Suppose, to the contrary, that $\beta(\{1,2,3,4,5\})>1.75\cdot\ell(\Omega_5)$. Let $e\in E(\Omega_5)\cap E(\Omega_9)\neq\emptyset$. Obviously, $\beta(\{4,5\})\le\frac{2}{3}\cdot(\ell(\Omega_5)+d(e))$. By ⟨6⟩, $d(e)>\frac{1}{8}\cdot\ell(\Omega_5)$. With $\beta(\{6,7,8,9\})\le\ell(\Omega_9)-d(e)$, $\beta(\{1,2,3\})\le\ell(\Omega_5)$, we find: $\beta(\{1,\dots,9\})\le\ell(\Omega_9)+\frac{5}{3}\cdot\ell(\Omega_5)-\frac{1}{3}\cdot d(e)<\ell(\Omega_9)+(\frac{5}{3}-\frac{1}{3\cdot 8})\cdot\ell(\Omega_5)\le\frac{21}{8}\cdot\ell(\Omega_9)$.

⟨8⟩ For n with $6\cdot 2^q\le n\le 9\cdot 2^q$, $q\ge 0$: $\beta(\{1,2,\dots,n\})\le(\lg^*(n)-0.5)\cdot\ell(\Omega_n)$.
It has been shown to be true for $n=6,7,8,9$ ($q=0$), by ⟨3⟩ and ⟨7⟩, and follows for $q>0$ by induction.
Remark: this bound is tight, cf. (10).

⟨9⟩ For $n=5\cdot 2^q+m$, with $0\le m\le 2^q$, $q\ge 0$: $\beta(\{1,\dots,n\})\le(\lg^*(n)-0.45-0.05\cdot m\cdot 2^{-q})\cdot\ell(\Omega_n)$.
This is evident for $n=5,6$ ($q=0,m=0,1$), by ⟨3⟩ and ⟨4⟩, while it follows for $q>0$ by induction.
Remark: this bound is tight, since we have: $2^{q+2}<n=2^{q+2}+2^q+m<2^{q+3}$, and so $\lg^*(n)-0.45-0.05\cdot m\cdot 2^{-q}=q+1.8+0.2\cdot m\cdot 2^{-q}$, cf. (12).

⟨10⟩ If $\beta(\{1,\dots,10\})\le 2.75\cdot\ell(\Omega_{10})$, then for $n=5\cdot 2^{q+1}-m$, with $0\le m\le 2^q$, $q\ge 0$: $\beta(\{1,\dots,n\})\le(\lg^*(n)-0.50)\cdot\ell(\Omega_n)$.
The statement is true for $n=9$ by ⟨7⟩, it is true for $n=10$ by assumption, so it is true for $q=0$. It follows for $q>0$ by induction.

⟨11⟩ If $\beta(\{1,\dots,10\})>2.75\cdot\ell(\Omega_{10})$, then for $n=5\cdot 2^{q+1}-m$, with $0\le m\le 2^q$, $q\ge 0$: $\beta(\{1,\dots,n\})\le(q+1)\cdot\ell(\Omega_n)+\frac{5}{3}\cdot\ell(\Omega_5)+(\frac{2}{3}-m\cdot 2^{-q})\cdot d(e_-)$, where $d(e_-):=\min\{d(e)|e\in E(\Omega_5)\}$.
We prove the statement for $q=0$. For $q>0$ it follows by induction.
Let $e^*\in E(\Omega_5)\cap E(\Omega_9)\neq\emptyset$. We have: $\beta(\{6,\dots,10\})\le\ell(\Omega_{10})$; $\beta(\{6,7,8,9\})\le\ell(\Omega_9)-d(e^*)$; $\beta(\{4,5\})\le\frac{2}{3}\cdot(\ell(\Omega_5)+d(e_-))$; $\beta(\{1,2,3\})\le\ell(\Omega_5)$. So, $\beta(\{1,\dots,9\})\le\ell(\Omega_9)+\frac{5}{3}\cdot\ell(\Omega_5)+\frac{2}{3}\cdot d(e_-)-d(e^*)\le\ell(\Omega_9)+\frac{5}{3}\cdot\ell(\Omega_5)+(\frac{2}{3}-1)\cdot d(e_-)$, while $\beta(\{1,\dots,10\})\le\ell(\Omega_{10})+\frac{5}{3}\cdot\ell(\Omega_5)+(\frac{2}{3}-0)\cdot d(e_-)$.

⟨12⟩ For $n=5\cdot 2^{q+1}-m$, $0\le m\le\frac{2}{3}\cdot 2^q$, $q\ge 0$: $\beta(\{1,\dots,n\})\le(\lg^*(n)-0.45-0.075\cdot m\cdot 2^{-q})\cdot\ell(\Omega_n)$.
If $\beta(\{1,\dots,10\})\le 2.75\cdot\ell(\Omega_{10})$, then by ⟨10⟩: $\beta(\{1,\dots,n\})\le(\lg^*(n)-0.5)\cdot\ell(\Omega_n)\le(\lg^*(n)-0.45-0.075\cdot m\cdot$

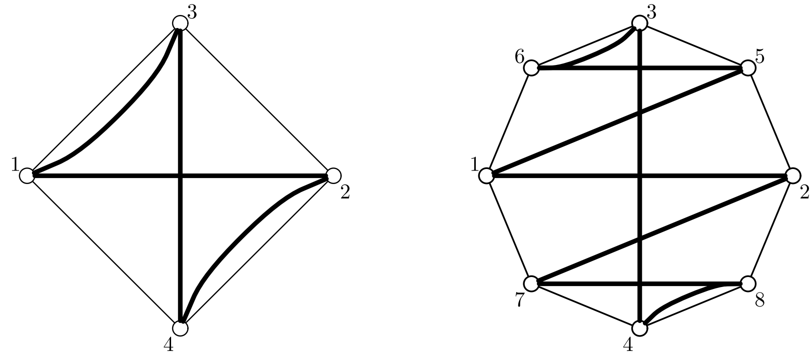

Fig. 2. FA results on A_4 and A_8.

$2^{-q}) \cdot \ell(\Omega_n)$. If $\beta(\{1,\dots,10\}) > 2.75 \cdot \ell(\Omega_{10})$, then by $\langle 11\rangle$: $\beta(\{1,\dots,n\}) \le (q+\frac{8}{3}) \cdot \ell(\Omega_n) + (\frac{2}{3} - m \cdot 2^{-q}) \cdot d(e_-) \le (q+\frac{8}{3}) \cdot \ell(\Omega_n) + (\frac{2}{3} - m \cdot 2^{-q}) \cdot 0.2 \cdot \ell(\Omega_5) \le (q + 2.8 - 0.2 \cdot m \cdot 2^{-q}) \cdot \ell(\Omega_n) = (q + 3 + (2^{q+1} - m) \cdot 2^{-q-3} - 0.45 - 0.075 \cdot m \cdot 2^{-q}) \cdot \ell(\Omega_n) = (\lg^*(n) - 0.45 - 0.075 \cdot m \cdot 2^{-q}) \cdot \ell(\Omega_n)$.

$\langle 13\rangle$ For $n = 5 \cdot 2^{q+1} - m$, with $\frac{2}{3} \cdot 2^q \le m \le 2^q$, $q \ge 0$: $\beta(\{1,\dots,n\}) \le (\lg^*(n) - 0.5) \cdot \ell(\Omega_n)$.
If $\beta(\{1,\dots,10\}) \le 2.75 \cdot \ell(\Omega_{10})$, then by $\langle 10\rangle$: $\beta(\{1,\dots,n\}) \le (\lg^*(n) - 0.5) \cdot \ell(\Omega_n)$. If $\beta(\{1,\dots,10\}) > 2.75 \cdot \ell(\Omega_{10})$, then by $\langle 6\rangle$,$\langle 11\rangle$: $\beta(\{1,\dots,n\}) \le (q+1) \cdot \ell(\Omega_n) + \frac{5}{3} \cdot \ell(\Omega_5) + (\frac{2}{3} - m \cdot 2^{-q}) \cdot d(e_-) \le (q+1) \cdot \ell(\Omega_n) + \frac{5}{3} \cdot \ell(\Omega_5) + (\frac{2}{3} - m \cdot 2^{-q}) \cdot \frac{1}{8} \cdot \ell(\Omega_5) \le (q + 2.75 - m \cdot 2^{-q-3}) \cdot \ell(\Omega_n) = (\lg^*(n) - 0.5) \cdot \ell(\Omega_n)$. Note that $\langle 6\rangle$ applies as $\beta(\{6,\dots,10\}) \le \ell(\Omega_{10})$, so $\beta(\{1,\dots,5\}) > 1.75 \cdot \ell(\Omega_{10}) \ge 1.75 \cdot \ell(\Omega_5)$. □

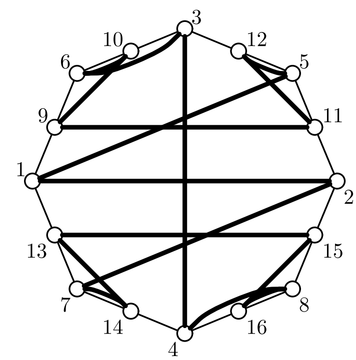

Fig. 3. FA results on A_{16}.

We conclude by describing the two families of instances $\mathcal{A}$ and $\mathcal{B}$, which contain worst case examples for Farthest Addition.

Family $\mathcal{A}$ resembles the family of worst case examples for Arbitrary Addition. For $n \ge 3$, instance $A_n \in \mathcal{A}$ consists of n selected nodes on the N-cycle with edges of length 1, where N is a power of 2, such that $\frac{1}{2}N < n \le N$. Nodes are labeled in order of selection by the FA algorithm. The cost of an edge (u, v) is the length of the shortest uv-path along the cycle. It is easily verified that the nodes can indeed be selected by FA in index order. The tour edges are such that it can easily be verified (working backwards) that a node k is inserted between nodes i and j with $i, j < k$ and such that at least one of these is nearest to k among all nodes indexed lower than k. The description of $\mathcal{A}$ centers around the instance A_8. We give A_4 and A_8 explicitly as in Fig. 2.

For $n < 8$, instance A_n is derived from A_8 by unlabeling the vertices $n+1,\dots,8$, and deleting them from the tour. For $\frac{1}{2}N = 2^q < n \le N$, $q \ge 3$, instance A_n is derived from $A_{\frac{1}{2}N}$ by subdividing each edge into two edges of length 1, labeling some of the new vertices $\frac{1}{2}N+1,\dots,n$, and inserting these vertices in the Farthest Addition tour in between vertices with labels higher than $\frac{1}{4}N$, so that insertion has cost 2. By sensible insertion this is always possible. The optimal tour length is $OPT = N$, while the Farthest Addition tour has length $FA = (q - \frac{1}{2})N + 2(n - \frac{1}{2}N) = N(\lg^*(n) - 0.5)$. As a final example we give A_{16}, in Fig. 3.

Family $\mathcal{B}$ is slightly more complicated, although it resembles $\mathcal{A}$, in a sense. For $n \ge 5$, instance $B_n \in \mathcal{B}$ consists of n selected nodes on the N-cycle with edges of length 1, and chords of length $0.2N$ and $0.4N$. Here $N = 5 \cdot 2^q$, for some $q \ge 0$, such that $\frac{1}{2}N < n \le N$. The description of $\mathcal{B}$ centers around the instance B_{10}. We give B_5 and B_{10} explicitly in Figs. 4 and 5. Instances B_n with $n < 10$ are derived from B_{10} by unlabeling vertices $n+1,\dots,10$. Instances B_n with $\frac{1}{2}N = 5 \cdot 2^{q-1} \le n \le N$ are derived from $B_{\frac{1}{2}N}$ by subdividing each cycle edge into two edges of length 1, and labeling and inserting some of these vertices in such a way, that insertion of a vertex has cost 2. This is always possible, if one chooses for a "sensible" insertion.

The N-cycles have the nodes 1,4,2,3,5, in this cyclic order, at distance $0.2N$ along the cycle. Further they have chords of length $0.2N$ between vertex 2 and all vertices between vertex 5 and vertex 3, and chords of length $0.4N$ between vertex 1 and all vertices between vertex 2 and vertex 3. The distance function $d(u, v)$ is the length of the shortest uv-path in the constructed graph. This definition "enables" Farthest Addition to make bad choices, in the early stage of the process. Once vertices 1,2,3,4,5 have been selected and inserted, so as to form a cyclic tour $(1, 2, 5, 4, 3)$ of length $1.8N$, the chords do not play a role anymore (Fig. 6). The instance families $\mathcal{A}$ and $\mathcal{B}$ and the family of worst case instances for Arbitrary Addition have one important feature in common. After an initial setup phase, during a *stage* (labeled k) nodes are added to the tour that all have distance 2^k to the intermediate tour. Moreover they all have a neighbor at distance 2^k that has been added to the tour in the previous stage. And finally these older nodes have a relative short edge. Therefore each new node will be added to the tour at a cost of 2^{k+1} and after this stage they will be incident to an edge of length 2^{k+1}. The last stage is labeled $k = 0$.

Notice that the upper bound (9) also applies to Farthest Insertion, but the lower bound instances A_n and B_n are optimally solved by FI.

Data availability

No data was used for the research described in the article.

Acknowledgement

I dedicate this work to the memory of Gerhard Woeginger, who encouraged me to publish these results.

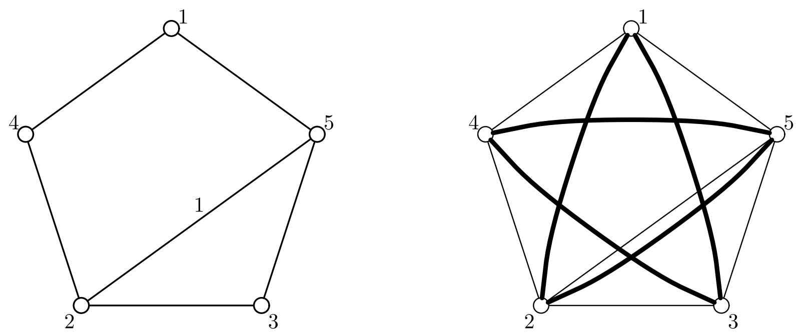

Fig. 4. FA results on B_5.

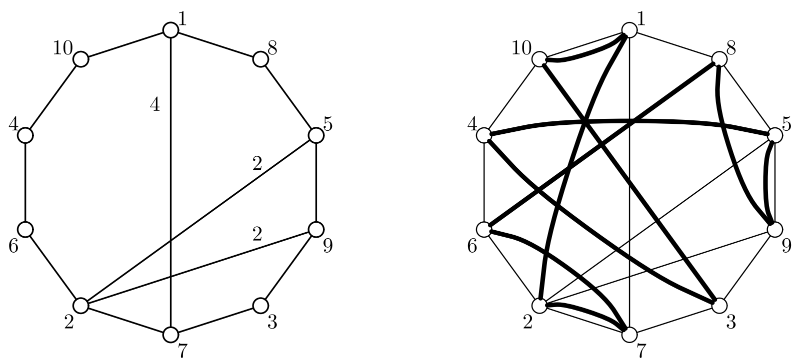

Fig. 5. FA results on B_{10}.

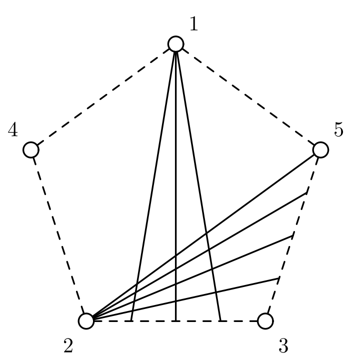

Fig. 6. Chords in construction of B_{5N} instance.

References

[1] Y. Azar, Lower bounds for insertion methods for TSP, Comb. Probab. Comput. 3 (1994) 285–292.

[2] G. Cornuéjols, G.L. Nemhauser, Tight bounds for Christofides' traveling salesman heuristic, Math. Program. 14 (1978) 116–121.

[3] A.M. Frieze, Worst-case analysis of algorithms for travelling salesman problems, Methods Oper. Res. 32 (1979) 97–112.

[4] B.L. Golden, W.R. Stewart, Empirical analysis of heuristics, in: E.L. Lawler, J.K. Lenstra, A.H.G. Rinnooy Kan, D.B. Shmoys (Eds.), The Traveling Salesman Problem, John Wiley & Sons Ltd., 1985, pp. 207–249.

[5] S. Hougardy, M. Wilde, On the nearest neighbor rule for the metric traveling salesman problem, Discrete Appl. Math. 195 (2015) 101–103.

[6] C.A.J. Hurkens, Nasty TSP instances for farthest insertion, in: Proceedings of IPCO 2, Pittsburgh, 1992, pp. 346–352.

[7] C.A.J. Hurkens, G.J. Woeginger, On the nearest neighbor rule for the traveling salesman problem, Oper. Res. Lett. 32 (1) (2004) 1–4.

[8] D.S. Johnson, Local optimization and the traveling salesman problem, in: Proceedings of the 17th Colloquium on Automata, Languages & Programming, Springer Verlag, Berlin, 1990, pp. 446–461.

[9] D.S. Johnson, C.H. Papadimitriou, Performance guarantees for heuristics, in: E.L. Lawler, J.K. Lenstra, A.H.G. Rinnooy Kan, D.B. Shmoys (Eds.), The Traveling Salesman Problem: A Guided Tour of Combinatorial Optimization, Wiley, Chichester, 1985, pp. 145–180 (Chapter 5).

[10] E.L. Lawler, J.K. Lenstra, A.H.G. Rinnooy Kan, D.B. Shmoys (Eds.), The Traveling Salesman Problem: A Guided Tour of Combinatorial Optimization, Wiley, Chichester, 1985.

[11] D. Pritchard, Nearest neighbor network traversal, arXiv:cs/0702114v1 [cs.DC], 2018.

[12] D.J. Rosenkrantz, R.E. Stearns, P.M. Lewis II, An analysis of several heuristics for the traveling salesman problem, SIAM J. Comput. 6 (1977) 563–581.