NWERC 2014 Presentation of solutions

The Jury

2014-11-30

NWERC 2014 Jury

- Per Austrin (KTH Royal Institute of Technology)
- Thomas Beuman (Leiden University)
- Jeroen Bransen (Utrecht University)
- Tommy Färnqvist (LiU)
- Jan Kuipers (AppTornado)
- Lukáš Poláček (KTH and Spotify)
- Alexander Rass (FAU)
- Fredrik Svensson (Autoliv Electronics)
- Tobias Werth (FAU)

Problem

Given two lists of strings, match them up so that as many as possible are identical.

Problem

Given two lists of strings, match them up so that as many as possible are identical.

Solution

Count how many times each word appears in each list.

Problem

Given two lists of strings, match them up so that as many as possible are identical.

- Count how many times each word appears in each list.
- Number of matches for each word is minimum of #occurrences in first list and #occurrences in second list.

Problem

Given two lists of strings, match them up so that as many as possible are identical.

- Count how many times each word appears in each list.
- Number of matches for each word is minimum of #occurrences in first list and #occurrences in second list.
- Sum over all words.

Problem

Given two lists of strings, match them up so that as many as possible are identical.

Solution

- Count how many times each word appears in each list.
- Number of matches for each word is minimum of #occurrences in first list and #occurrences in second list.
- Sum over all words.

Statistics: 135 submissions, 92 accepted

Problem

Find minimum of the following function:

$$f(c) = \frac{n \cdot (\log_2 n)^{c\sqrt{2}}}{10^9 p} + (1 + \frac{1}{c}) \cdot \frac{s}{v}$$

Problem

Find minimum of the following function:

$$f(c) = \frac{n \cdot (\log_2 n)^{c\sqrt{2}}}{10^9 p} + (1 + \frac{1}{c}) \cdot \frac{s}{v}$$

Solution

• This is a convex function, so use ternary search:

Problem

Find minimum of the following function:

$$f(c) = \frac{n \cdot (\log_2 n)^{c\sqrt{2}}}{10^9 p} + (1 + \frac{1}{c}) \cdot \frac{s}{v}$$

- This is a convex function, so use ternary search:
- Calculate value at some small and large c (e.g. 1 and 100)

Problem

Find minimum of the following function:

$$f(c) = \frac{n \cdot (\log_2 n)^{c\sqrt{2}}}{10^9 p} + (1 + \frac{1}{c}) \cdot \frac{s}{v}$$

- This is a convex function, so use ternary search:
- Calculate value at some small and large c (e.g. 1 and 100)
- Calculate value at 1/3 and 2/3 of this interval (i.e. 34 and 67)

Problem

Find minimum of the following function:

$$f(c) = \frac{n \cdot (\log_2 n)^{c\sqrt{2}}}{10^9 p} + (1 + \frac{1}{c}) \cdot \frac{s}{v}$$

- This is a convex function, so use ternary search:
- Calculate value at some small and large c (e.g. 1 and 100)
- Calculate value at 1/3 and 2/3 of this interval (i.e. 34 and 67)
- If f(34) < f(68), the minimum is in [1,68], otherwise in [34, 100].

Problem

Find minimum of the following function:

$$f(c) = \frac{n \cdot (\log_2 n)^{c\sqrt{2}}}{10^9 p} + (1 + \frac{1}{c}) \cdot \frac{s}{v}$$

- This is a convex function, so use ternary search:
- Calculate value at some small and large c (e.g. 1 and 100)
- Calculate value at 1/3 and 2/3 of this interval (i.e. 34 and 67)
- If f(34) < f(68), the minimum is in [1,68], otherwise in [34, 100].
- Iterate this a few dozen times to narrow down the interval.

Problem

Find minimum of the following function:

$$f(c) = \frac{n \cdot (\log_2 n)^{c\sqrt{2}}}{10^9 p} + (1 + \frac{1}{c}) \cdot \frac{s}{v}$$

Solution

- This is a convex function, so use ternary search:
- Calculate value at some small and large c (e.g. 1 and 100)
- Calculate value at 1/3 and 2/3 of this interval (i.e. 34 and 67)
- If f(34) < f(68), the minimum is in [1,68], otherwise in [34, 100].
- Iterate this a few dozen times to narrow down the interval.

Statistics: 91 submissions, 61 accepted

Problem

Split list of n item prices into d+1 groups, minimize sum of rounded costs.

Problem

Split list of n item prices into d+1 groups, minimize sum of rounded costs.

Solution

• Let C(i,j) be minimum cost of first i items using j groups.

Problem

Split list of n item prices into d+1 groups, minimize sum of rounded costs.

- Let C(i,j) be minimum cost of first i items using j groups.
- $C(i,j) = \min_{i' < i} \{ C(i',j-1) + \text{Round}(x_{i'+1} + \ldots + x_i) \}$

Problem

Split list of n item prices into d+1 groups, minimize sum of rounded costs.

- Let C(i,j) be minimum cost of first i items using j groups.
- $C(i,j) = \min_{i' < i} \{ C(i',j-1) + \text{Round}(x_{i'+1} + \ldots + x_i) \}$
- Compute with dynamic programming.

Problem

Split list of n item prices into d+1 groups, minimize sum of rounded costs.

- Let C(i,j) be minimum cost of first i items using j groups.
- $C(i,j) = \min_{i' < i} \{ C(i',j-1) + \text{Round}(x_{i'+1} + \ldots + x_i) \}$
- Compute with dynamic programming.
- As written $O(n^3d)$ too slow

Problem

Split list of n item prices into d+1 groups, minimize sum of rounded costs.

- Let C(i,j) be minimum cost of first i items using j groups.
- $C(i,j) = \min_{i' < i} \{ C(i',j-1) + \text{Round}(x_{i'+1} + \ldots + x_i) \}$
- Compute with dynamic programming.
- As written $O(n^3d)$ too slow
- Don't recompute sums $O(n^2d)$ fast enough



Problem

Split list of n item prices into d+1 groups, minimize sum of rounded costs.

- Let C(i,j) be minimum cost of first i items using j groups.
- $C(i,j) = \min_{i' < i} \{ C(i',j-1) + \text{Round}(x_{i'+1} + \ldots + x_i) \}$
- Compute with dynamic programming.
- As written $O(n^3d)$ too slow
- Don't recompute sums $O(n^2d)$ fast enough
- Be more clever O(nd) even faster



Problem

Split list of n item prices into d+1 groups, minimize sum of rounded costs.

Solution

- Let C(i,j) be minimum cost of first i items using j groups.
- $C(i,j) = \min_{i' < i} \{ C(i',j-1) + \text{Round}(x_{i'+1} + \ldots + x_i) \}$
- Compute with dynamic programming.
- As written $O(n^3d)$ too slow
- Don't recompute sums $O(n^2d)$ fast enough
- Be more clever O(nd) even faster

Statistics: 216 submissions, 62 accepted

H – Hyacinth

Problem

Given a tree, assign a number (the frequency) to each edge, such that every node is adjacent to at most 2 different numbers (its NIC's frequencies), and as many numbers as possible are used.

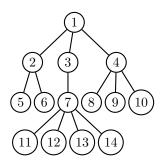
H – Hyacinth

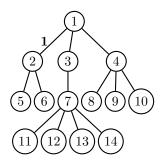
Problem

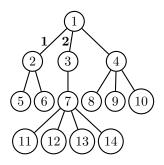
Given a tree, assign a number (the frequency) to each edge, such that every node is adjacent to at most 2 different numbers (its NIC's frequencies), and as many numbers as possible are used.

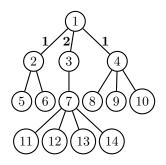
Solution

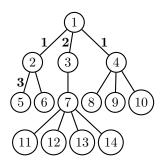
Greedy: start somewhere, assign a new number to an edge if possible, otherwise reuse one, and then recurse.

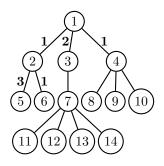


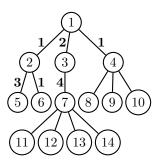


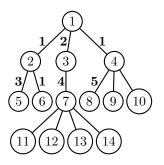


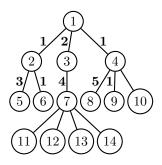


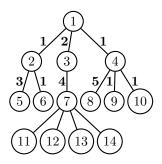


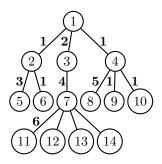


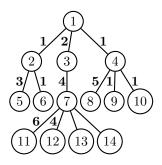


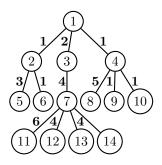


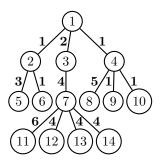


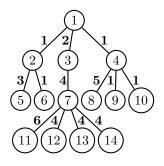




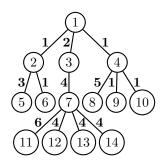








Note: the number of different used channels equals the number of internal nodes $+\ 1$.



Note: the number of different used channels equals the number of internal nodes + 1.

Statistics: 213 submissions, 70 accepted

Problem

Given n points in \mathbb{R}^2 , is there a straight line passing through at least a p fraction of them?

Problem

Given n points in \mathbb{R}^2 , is there a straight line passing through at least a p fraction of them?

Solution

• If line exists it is uniquely determined by any two of its points.

Problem

Given n points in \mathbb{R}^2 , is there a straight line passing through at least a p fraction of them?

- If line exists it is uniquely determined by any two of its points.
- Probability that a random point lies on the line is $\geq p$.

Problem

Given n points in \mathbb{R}^2 , is there a straight line passing through at least a p fraction of them?

- If line exists it is uniquely determined by any two of its points.
- Probability that a random point lies on the line is $\geq p$.
- Repeat 250 times:
 - Pick two distinct points uniformly at random
 - 2 Check if line defined by the points has enough points.

Problem

Given n points in \mathbb{R}^2 , is there a straight line passing through at least a p fraction of them?

- If line exists it is uniquely determined by any two of its points.
- Probability that a random point lies on the line is $\geq p$.
- Repeat 250 times:
 - 1 Pick two distinct points uniformly at random
 - ② Check if line defined by the points has enough points.
- Pr[false negative] $pprox (1-p^2)^{250} \le (1-1/25)^{250} < 5\cdot 10^{-5}$ (cheated slightly it's a bit worse)

Problem

Given n points in \mathbb{R}^2 , is there a straight line passing through at least a p fraction of them?

Solution

- If line exists it is uniquely determined by any two of its points.
- Probability that a random point lies on the line is $\geq p$.
- Repeat 250 times:
 - 1 Pick two distinct points uniformly at random
 - ② Check if line defined by the points has enough points.
- Pr[false negative] $pprox (1-p^2)^{250} \le (1-1/25)^{250} < 5\cdot 10^{-5}$ (cheated slightly it's a bit worse)

Many other ways e.g. divide and conquer

Problem

Given n points in \mathbb{R}^2 , is there a straight line passing through at least a p fraction of them?

Solution

- If line exists it is uniquely determined by any two of its points.
- Probability that a random point lies on the line is $\geq p$.
- Repeat 250 times:
 - Pick two distinct points uniformly at random
 - ② Check if line defined by the points has enough points.
- Pr[false negative] $pprox (1-p^2)^{250} \le (1-1/25)^{250} < 5\cdot 10^{-5}$ (cheated slightly it's a bit worse)

Many other ways e.g. divide and conquer

Statistics: 222 submissions, 23 accepted

Problem

Given a directed acyclic graph, where the nodes are switches that send balls to the left and right in turns, what is its end state.

Problem

Given a directed acyclic graph, where the nodes are switches that send balls to the left and right in turns, what is its end state.

Solution

• Don't simulate 10¹⁸ balls one by one!

Problem

Given a directed acyclic graph, where the nodes are switches that send balls to the left and right in turns, what is its end state.

- Don't simulate 10¹⁸ balls one by one!
- Observation: if N balls arrive at a node, $\lceil N/2 \rceil$ go to one side, and $\lfloor N/2 \rfloor$ to the other side, and its state is flipped iff N is odd.

Problem

Given a directed acyclic graph, where the nodes are switches that send balls to the left and right in turns, what is its end state.

- Don't simulate 10¹⁸ balls one by one!
- Observation: if N balls arrive at a node, $\lceil N/2 \rceil$ go to one side, and $\lfloor N/2 \rfloor$ to the other side, and its state is flipped iff N is odd.
- Process nodes one by one

Problem

Given a directed acyclic graph, where the nodes are switches that send balls to the left and right in turns, what is its end state.

- Don't simulate 10¹⁸ balls one by one!
- Observation: if N balls arrive at a node, $\lceil N/2 \rceil$ go to one side, and $\lfloor N/2 \rfloor$ to the other side, and its state is flipped iff N is odd.
- Process nodes one by one
- Do it in topological sort order.

Problem

Given a directed acyclic graph, where the nodes are switches that send balls to the left and right in turns, what is its end state.

Solution

- Don't simulate 10¹⁸ balls one by one!
- Observation: if N balls arrive at a node, $\lceil N/2 \rceil$ go to one side, and $\lfloor N/2 \rfloor$ to the other side, and its state is flipped iff N is odd.
- Process nodes one by one
- Do it in topological sort order.

Statistics: 327 submissions, 43 accepted

Problem

How long does it take to pick up all n knapsacks at the luggage carousel?

Problem

How long does it take to pick up all n knapsacks at the luggage carousel?

- Given starting position, can simulate the process in $O(n \log n)$
 - Keep positions of knapsacks in a set for fast lookup of when the next knapsack will arrive.

Problem

How long does it take to pick up all n knapsacks at the luggage carousel?

- Given starting position, can simulate the process in $O(n \log n)$
 - Keep positions of knapsacks in a set for fast lookup of when the next knapsack will arrive.
- The interesting starting positions are the $\leq n$ slots where there are knapsacks.

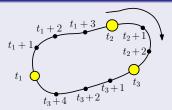
Problem

How long does it take to pick up all n knapsacks at the luggage carousel?

- Given starting position, can simulate the process in $O(n \log n)$
 - Keep positions of knapsacks in a set for fast lookup of when the next knapsack will arrive.
- The interesting starting positions are the $\leq n$ slots where there are knapsacks.
- Time from others grow arithmetically

Problem

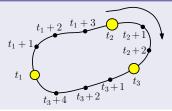
How long does it take to pick up all n knapsacks at the luggage carousel?



Problem

How long does it take to pick up all n knapsacks at the luggage carousel?

Solution



Statistics: 68 submissions, 22 accepted

Problem

Given a weighted graph, determine whether there exists a cycle with length L that visits every vertex once.

Problem

Given a weighted graph, determine whether there exists a cycle with length L that visits every vertex once.

Solution

• First of all, $\mathcal{O}(N!)$ is too slow $(14! = 87 \cdot 10^9)$.

Problem

Given a weighted graph, determine whether there exists a cycle with length L that visits every vertex once.

- First of all, $\mathcal{O}(N!)$ is too slow $(14! = 87 \cdot 10^9)$.
- Take 0 as starting point and loop over middle point: $\mathcal{O}(N)$.

Problem

Given a weighted graph, determine whether there exists a cycle with length L that visits every vertex once.

- First of all, $\mathcal{O}(N!)$ is too slow $(14! = 87 \cdot 10^9)$.
- Take 0 as starting point and loop over middle point: $\mathcal{O}(N)$.
- Partition remaining points into 1st and 2nd half: $\mathcal{O}(2^N)$.

Problem

Given a weighted graph, determine whether there exists a cycle with length L that visits every vertex once.

- First of all, $\mathcal{O}(N!)$ is too slow $(14! = 87 \cdot 10^9)$.
- Take 0 as starting point and loop over middle point: $\mathcal{O}(N)$.
- Partition remaining points into 1st and 2nd half: $\mathcal{O}(2^N)$.
- Calculate possible lengths of half-cycles: $\mathcal{O}((N/2)!)$.

Problem

Given a weighted graph, determine whether there exists a cycle with length L that visits every vertex once.

- First of all, $\mathcal{O}(N!)$ is too slow $(14! = 87 \cdot 10^9)$.
- Take 0 as starting point and loop over middle point: $\mathcal{O}(N)$.
- Partition remaining points into 1st and 2nd half: $\mathcal{O}(2^N)$.
- Calculate possible lengths of half-cycles: $\mathcal{O}((N/2)!)$.
- For every half-cycle length K, check whether L-K exists in $\mathcal{O}(1)$ with hashset or so.

Problem

Given a weighted graph, determine whether there exists a cycle with length L that visits every vertex once.

Solution

- First of all, $\mathcal{O}(N!)$ is too slow $(14! = 87 \cdot 10^9)$.
- Take 0 as starting point and loop over middle point: $\mathcal{O}(N)$.
- Partition remaining points into 1st and 2nd half: $\mathcal{O}(2^N)$.
- Calculate possible lengths of half-cycles: $\mathcal{O}((N/2)!)$.
- For every half-cycle length K, check whether L-K exists in $\mathcal{O}(1)$ with hashset or so.

Statistics: 79 submissions, 7 accepted

Problem

Given points $(x_1, y_1), \ldots, (x_n, y_n)$, find point (x^*, y^*) minimizing sum of Manhattan distances to all points, while being within Manhattan distance d of all points.

Problem

Given points $(x_1, y_1), \ldots, (x_n, y_n)$, find point (x^*, y^*) minimizing sum of Manhattan distances to all points, while being within Manhattan distance d of all points.

Solution

• The "ideal" position is $x^* = \text{median}(x_1, \dots, x_n)$ and $y^* = \text{median}(y_1, \dots, y_n)$

Problem

Given points $(x_1, y_1), \ldots, (x_n, y_n)$, find point (x^*, y^*) minimizing sum of Manhattan distances to all points, while being within Manhattan distance d of all points.

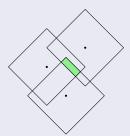
- The "ideal" position is $x^* = \text{median}(x_1, ..., x_n)$ and $y^* = \text{median}(y_1, ..., y_n)$
- But might be more than distance d away from some point...

Problem

Given points $(x_1, y_1), \ldots, (x_n, y_n)$, find point (x^*, y^*) minimizing sum of Manhattan distances to all points, while being within Manhattan distance d of all points.

Solution

Feasible region is intersection of diamonds, forms a "skewed diamond"



Problem

Given points $(x_1, y_1), \ldots, (x_n, y_n)$, find point (x^*, y^*) minimizing sum of Manhattan distances to all points, while being within Manhattan distance d of all points.

- Key observation: there is an optimal point which either:
 - 1 Is a corner of the feasible region, or
 - 2 Shares x-value with an input point, or
 - 3 Shares y-value with an input point.

Problem

Given points $(x_1, y_1), \ldots, (x_n, y_n)$, find point (x^*, y^*) minimizing sum of Manhattan distances to all points, while being within Manhattan distance d of all points.

- Key observation: there is an optimal point which either:
 - 1 Is a corner of the feasible region, or
 - 2 Shares x-value with an input point, or
 - Shares y-value with an input point.
- Given a candidate value for x^* , we get a range of allowed y values $y_{lo}...y_{hi}$. Best choice is median y (if in range) or y_{hi} .

G - Gathering

Problem

Given points $(x_1, y_1), \ldots, (x_n, y_n)$, find point (x^*, y^*) minimizing sum of Manhattan distances to all points, while being within Manhattan distance d of all points.

- Key observation: there is an optimal point which either:
 - 1 Is a corner of the feasible region, or
 - 2 Shares x-value with an input point, or
 - Shares y-value with an input point.
- Given a candidate value for x^* , we get a range of allowed y values $y_{lo}...y_{hi}$. Best choice is median y (if in range) or y_{hi} .
- (Similarly for best x value given candindate value of y^*)

G - Gathering

Problem

Given points $(x_1, y_1), \ldots, (x_n, y_n)$, find point (x^*, y^*) minimizing sum of Manhattan distances to all points, while being within Manhattan distance d of all points.

Solution

- Key observation: there is an optimal point which either:
 - 1 Is a corner of the feasible region, or
 - 2 Shares x-value with an input point, or
 - Shares y-value with an input point.
- Given a candidate value for x^* , we get a range of allowed y values $y_{lo}...y_{hi}$. Best choice is median y (if in range) or y_{hi} .
- (Similarly for best x value given candindate value of y^*)

Statistics: 20 submissions, ? accepted

Problem

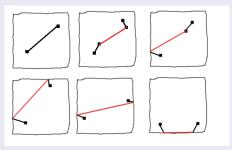
Find shortest path, possibly via some of the given bike stations or the (infinitely many possible) stations the boundary.

Problem

Find shortest path, possibly via some of the given bike stations or the (infinitely many possible) stations the boundary.

Solution

Possible solutions use 0 or 2 stations:



Problem

Find shortest path, possibly via some of the given bike stations or the (infinitely many possible) stations the boundary.

Solution

• Loop over station 1 and 2 (existing, or NSEW boundary).

Problem

Find shortest path, possibly via some of the given bike stations or the (infinitely many possible) stations the boundary.

- Loop over station 1 and 2 (existing, or NSEW boundary).
- Calculate distance.

Problem

Find shortest path, possibly via some of the given bike stations or the (infinitely many possible) stations the boundary.

- Loop over station 1 and 2 (existing, or NSEW boundary).
- Calculate distance.
- To determine correct location on the boundary use ternary search, since distance as function of position is a convex function.

Problem

Find shortest path, possibly via some of the given bike stations or the (infinitely many possible) stations the boundary.

- Loop over station 1 and 2 (existing, or NSEW boundary).
- Calculate distance.
- To determine correct location on the boundary use ternary search, since distance as function of position is a convex function.
- For 2 boundary stations, use nested ternary searches.

Problem

Find shortest path, possibly via some of the given bike stations or the (infinitely many possible) stations the boundary.

- Loop over station 1 and 2 (existing, or NSEW boundary).
- Calculate distance.
- To determine correct location on the boundary use ternary search, since distance as function of position is a convex function.
- For 2 boundary stations, use nested ternary searches.
- Doing it with math becomes a terrible mess.

Problem

Find shortest path, possibly via some of the given bike stations or the (infinitely many possible) stations the boundary.

Solution

- Loop over station 1 and 2 (existing, or NSEW boundary).
- Calculate distance.
- To determine correct location on the boundary use ternary search, since distance as function of position is a convex function.
- For 2 boundary stations, use nested ternary searches.
- Doing it with math becomes a terrible mess.

Statistics: 7 submissions, ? accepted

Problem

Problem

Find shortest tour around a track defined by two polygons.

Solution

Without the outer boundary: convex hull

Problem

Find shortest tour around a track defined by two polygons.

- Without the outer boundary: convex hull
- A somewhat unusual convex hull algorithm.
 - Find a point where we make a right-turn, shortcut if possible
 - Repeat until done

Problem

Find shortest tour around a track defined by two polygons.

- Without the outer boundary: convex hull
- A somewhat unusual convex hull algorithm.
 - Find a point where we make a right-turn, shortcut if possible
 - 2 Repeat until done
- With outer boundary: when shortcutting, can't go straight, need to wrap around the outer boundary
 - Convex hull computation

Problem

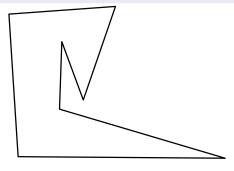
Find shortest tour around a track defined by two polygons.

Solution

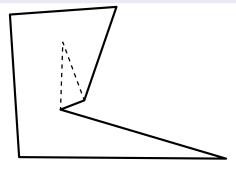
- Without the outer boundary: convex hull
- A somewhat unusual convex hull algorithm.
 - Find a point where we make a right-turn, shortcut if possible
 - Repeat until done
- With outer boundary: when shortcutting, can't go straight, need to wrap around the outer boundary
 - Convex hull computation

Can also use Dijkstra, need to be careful to make a full lap around the inner polygon.

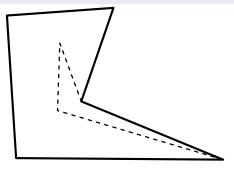
Problem



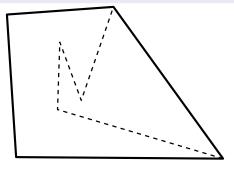
Problem



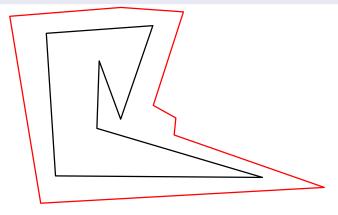
Problem



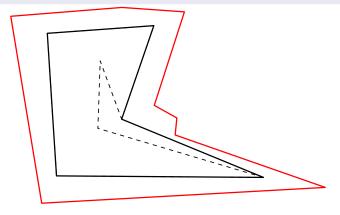
Problem



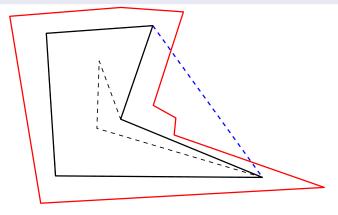
Problem



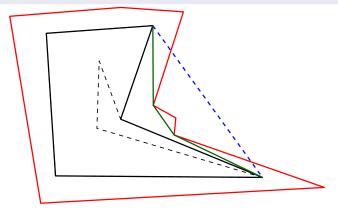
Problem



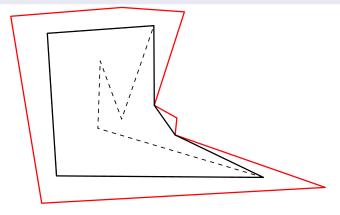
Problem



Problem

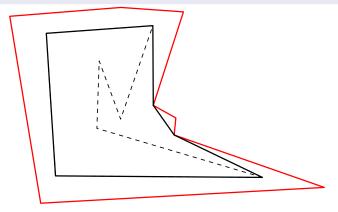


Problem

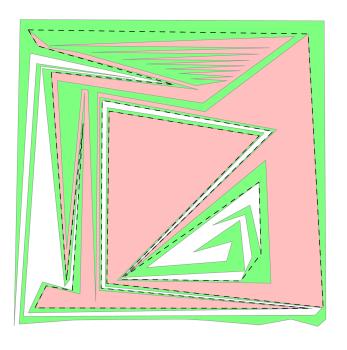


Problem

Find shortest tour around a track defined by two polygons.



Statistics: 12 submissions, ? accepted



Random numbers produced by the jury

- 1087 number of posts made in the jury's forum.
 - 671 commits made to the problem set repository.
 - 380 number of lines of code used in total by the shortest judge solutions to solve the entire problem set.
- 16.7 average number of jury solutions per problem (including incorrect ones)
 - 1 number of beers Jan lost to Per over bets on the problems

Random numbers produced by the jury

- 1087 number of posts made in the jury's forum.
 - 671 commits made to the problem set repository.
 - 380 number of lines of code used in total by the shortest judge solutions to solve the entire problem set.
- 16.7 average number of jury solutions per problem (including incorrect ones)
 - 1 number of beers Jan lost to Per over bets on the problems
 - 1 number of beers Per lost to Jan over bets on the contest results