The Travelling Salesman Problem and its Applications

CO@W Berlin

Martin Grötschel

21.9.2009

16:00 – 17:30

Martin Grötschel
- Institut für Mathematik, Technische Universität Berlin (TUB)
- DFG-Forschungszentrum "Mathematik für Schlüsseltechnologien" (MATHEON)
- Konrad-Zuse-Zentrum für Informationstechnik Berlin (ZIB)

groetschel@zib.de
http://www.zib.de/groetschel

Contents

1. Introduction
2. The TSP and some of its history
3. The TSP and some of its variants
4. Some applications
5. Heuristics
6. How combinatorial optimizers do it

Contents

1. Introduction
2. The TSP and some of its history
3. The TSP and some of its variants
4. Some applications
5. Heuristics
6. How combinatorial optimizers do it

Combinatorial optimization

Given a finite set E and a subset $\mathcal{I}$ of the power set of E (the set of feasible solutions). Given, moreover, a value (cost, length,…) c(e) for all elements e of E. Find, among all sets in $\mathcal{I}$, a set I such that its total value c(I) (= sum of the values of all elements in I) is as small (or as large) as possible.

The parameters of a combinatorial optimization problem are: (E, $\mathcal{I}$, c).

$$\min\left\{c(\mathrm{I}) = \sum_{e\in \mathrm{I}} c(e) \mid \mathrm{I} \in \mathcal{I}\right\},\ where\ \mathcal{I} \subseteq 2^{E} and\ E\ \ finite$$

Important issues:

- How is $\mathcal{I}$ given?
- What is the encoding length of an instance?
- How do we measure running time?

Encoding and Running Times

Important issues:

- How is I given?
- What is the encoding length of an instance?
- How do we measure running time?

Special „simple“ combinatorial optimization problems

Finding a

- minimum spanning tree in a graph
- shortest path in a directed graph
- maximum matching in a graph
- a minimum capacity cut separating two given nodes of a graph or digraph
- cost-minimal flow through a network with capacities and costs on all edges
- …

These problems are solvable in polynomial time.

Is the number of feasible solutions relevant?

Special „hard“ combinatorial optimization problems

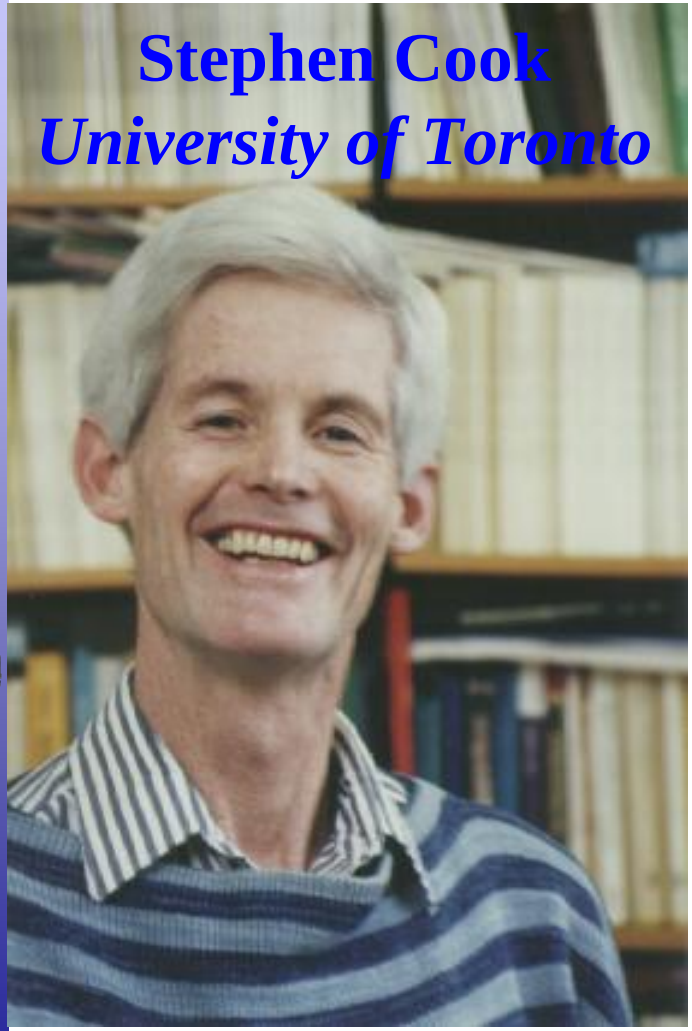
- travelling salesman problem (the prototype problem)
- location und routing
- set-packing, partitioning, -covering
- max-cut
- linear ordering
- scheduling (with a few exceptions)
- node and edge colouring
- …

These problems are NP-hard
(in the sense of complexity theory).

Complexity Theory

- Complexity theory came formally into being in the years 1965 – 1972 with the work of Cobham (1965), Edmonds(1965), Cook (1971), Karp(1972) and many others
- Of course, there were many forerunners (Gauss has written about the number of elementary steps in a computation, von Neumann, Gödel, Turing, Post,…).
- But modern complexity theory is a the result of the combined research efforts of many, in particular, of many computer scientists and mathematical programmers trying to understand the structures underlying computational processes.

Complexity Theory



1965 Polynomial time
Class P
Nondeterministic polynomial time
Class NP
Edmonds, Cobham

1971 Cook "The Complexity of Theorem Proving Procedures"
introduced the theory of NP completeness

Hierarchies of complexity classes...

The most important open problem:

$$P = NP\ ?$$

Clay Mathematics Institute

dedicated to increasing and disseminating mathematical knowledge

Millennium Prize Problems

Announcement

Rules for the CMI Millennium Prize Problems

Publication Guidelines

Historical Context

Press Statement

Press Reaction

P versus NP

The Hodge Conjecture

The Poincaré Conjecture

The Riemann Hypothesis

Yang-Mills Existence and Mass Gap

Navier-Stokes Existence and Smoothness

The Birch and Swinnerton-Dyer Conjecture

Announced 16:00, on Wednesday, May 24, 2000
Collège de France

The first NP-complete Problem

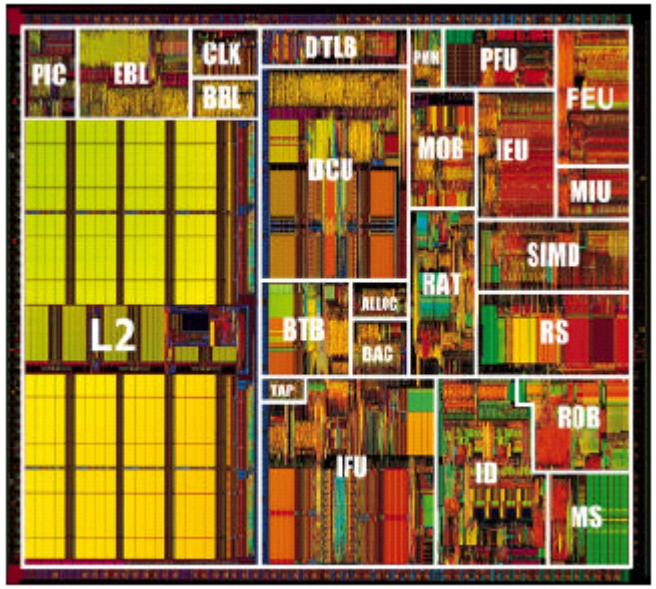
Satisfiability: Is there a truth assignmeent to the following formula:

$$(\neg x_1 \vee x_2) \wedge (x_1 \vee x_2 \vee x_3) \wedge (x_1 \vee \neg x_2) \wedge (x_1 \vee x_2 \vee \neg x_3) \wedge (\neg x_1 \vee \neg x_2)$$

Truly important Application:

Verification of computer chips and "systems on chips"

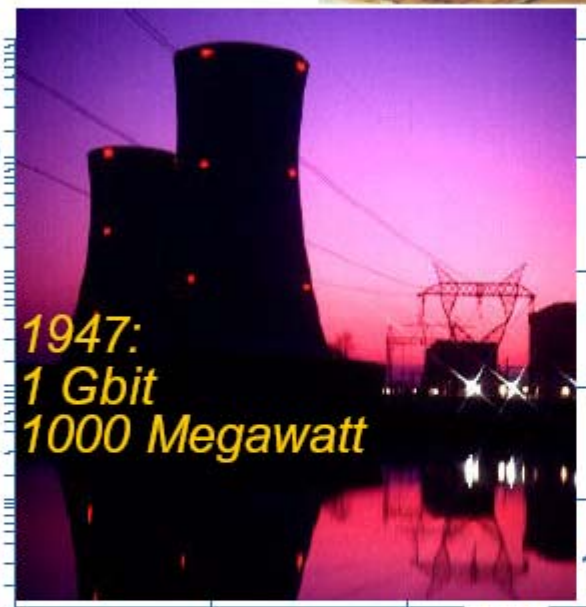
A design is correct iff a certain SAT formula associated with the chip has no truth assignment.



Martin Grötschel

Moore's Law

productivity gap



Transistors per Chip

10^{11} 10^{10} 10^{9} 10^{8} 10^{7} 10^{6} 10^{5} 10^{4} 10^{3}

1970 1980 1990 2000 2010

DRAMs

1k 4k 16k 64k 256k 1M 4M 16M 64M 256M 1G 4G 16G 64G

Logic Module

4004 8086 80286 80386 i486 Pentium Pentium Pro PPC 620 Pentium II Pentium III

6 days music
4 h video
2.000.000 pages

Source: ICN

Intel
Motorola

© Siemens AG, Mobile Solutions Marketing Georg Berner

The travelling salesman problem

Given n „cities" and „distances" between them. Find a tour (roundtrip) through all cities visiting every city exactly once such that the sum of all distances travelled is as small as possible. (TSP)

The TSP is called symmetric (STSP) if, for every pair of cities i and j, the distance from i to j is the same as the one from j to i, otherwise the problem is called aysmmetric (ATSP).

The travelling salesman problem

Two mathematical formulations of the TSP

1. *Version*:

Let $K_n = (V, E)$ *be the complete graph (or digraph) with* n *nodes and let* c_e *be the length of* $e \in E$. *Let* H *be the set of all hamiltonian cycles (tours) in* K_n. *Find*

$$\min\{c(T) \mid T \in H\}.$$

2. *Version*:

Find a cyclic permutation π *of* $\{1,\ldots,n\}$ *such that*

$$\sum_{i=1}^{n} c_{i\pi(i)}$$

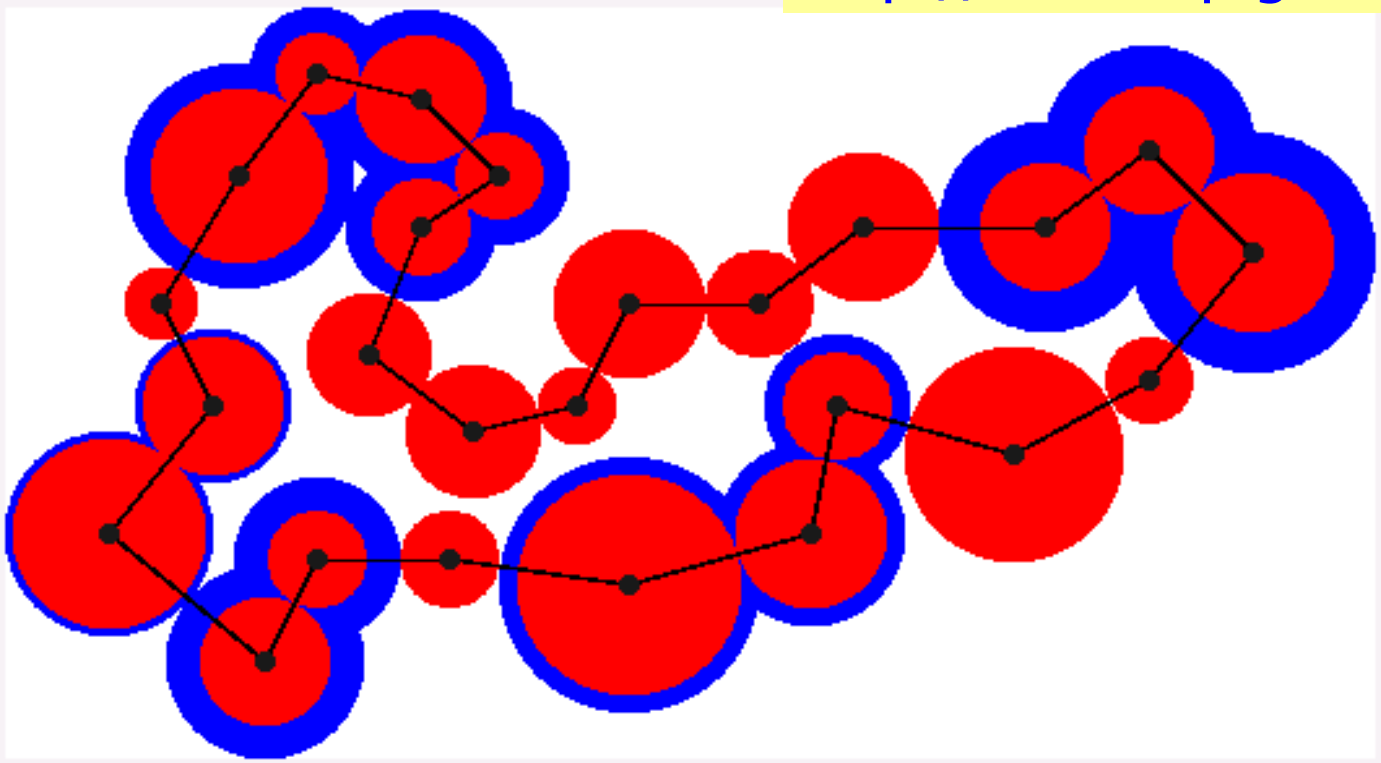
is as small as possible.

- Does that help solve the TSP?

The Problem

Given a collection of cities and the cost of travel between each pair of them, the **traveling salesman problem**, or **TSP** for short, is to find the cheapest way of visiting all of the cities and returning to your starting point. In the case we study, the travel costs are symmetric in the sense that traveling from city X to city Y costs just as much as traveling from Y to X.

http://www.tsp.gatech.edu/



The simplicity of the statement of the problem is deceptive -- the TSP is one of the most intensely studied problems in computational mathematics and yet no effective solution method is known for the general case. Indeed, the resolution of the TSP would settle the P versus NP problem and fetch a $1,000,000 prize from the Clay Mathematics Institute.

Contents

1. Introduction
2. The TSP and some of its history
3. The TSP and some of its variants
4. Some applications
5. Heuristics
6. How combinatorial optimizers do it

History of the TSP

Mathematical problems related to the traveling salesman problem were treated in the 1800s by the Irish mathematician Sir William Rowan Hamilton and by the British mathematician Thomas Penyngton Kirkman. The picture below is a photograph of Hamilton's Icosian Game that requires players to complete tours through the 20 points using only the specified connections. A nice discussion of the early work of Hamilton and Kirkman can be found in the book *Graph Theory 1736-1936* by N. L. Biggs, E. K. LLoyd, and R. J. Wilson, Clarendon Press, Oxford, 1976.

Usually quoted as the forerunner of the TSP →

Usually quoted as the origin of the TSP ↓

The general form of the TSP appears to be have been first studied by mathematicians starting in the 1930s by Karl Menger in Vienna and Harvard. The problem was later promoted by Hassler Whitney and Merrill Flood at Princeton. A detailed treatment of the connection between Menger and Whitney, and the growth of the TSP as a topic of study can be found in Alexander Schrijver's paper ``On the history of combinatorial optimization (till 1960)''.

about 100 years earlier

Der Handlungsreisende wie er sein soll

und was er zu thun hat, um Aufträge zu erhalten und eines glücklichen Erfolgs in seinen Geschäften gewiß zu sein.

Von

einem alten Commis - Voyageur.

Ilmenau 1832,

Druck und Verlag von B. Fr. Voigt.

From the Commis-Voyageur

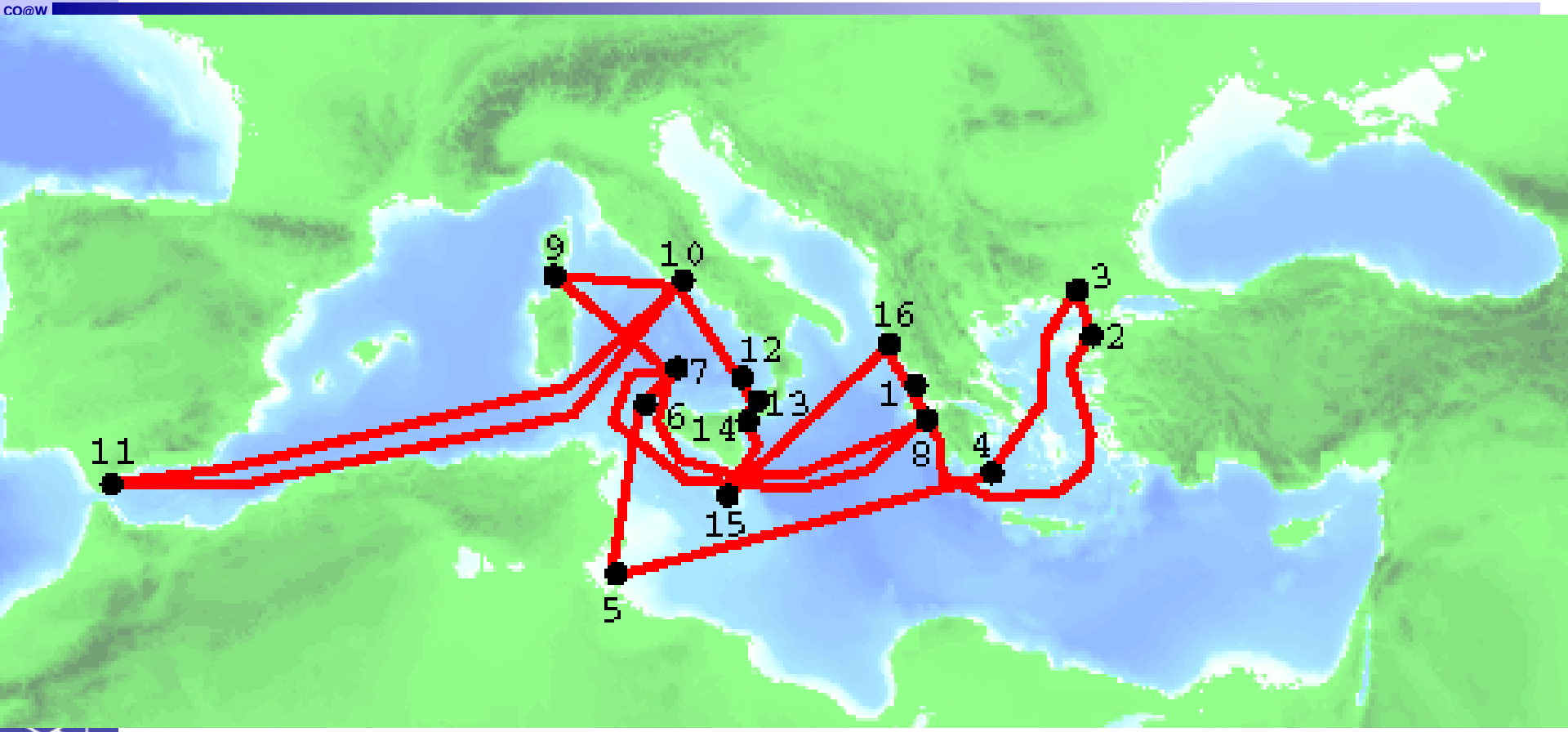
aber es kann durch eine zweckmäßige Wahl und Eintheilung der Tour, manchmal so viel Zeit gewonnen werden, daß wir es nicht glauben umgehen zu dürfen, auch hierüber einige Vorschriften zu geben.

By a proper choice and scheduling of the tour one can gain so much time that we have to make some suggestions

worauf der Reisende hauptsächlich zu sehen hat, des Hin- und Herreisens, mit mehr Oekonomie einzurichten. Die Hauptsache besteht immer darin: so viele Orte wie möglich mitzunehmen, ohne den nämlichen Ort zweimal berühren zu müssen.

The most important aspect is to cover as many locations as possible without visiting a location twice

Ulysses roundtrip (an even older TSP ?)



Ulysses

	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	509	501	312	1019	736	656	60	1039	726	2314	479	448	479	619	150
2		126	474	1526	1226	1133	532	1449	1122	2789	958	941	978	1127	542
3			541	1516	1184	1084	536	1371	1045	2728	913	904	946	1115	499
4				1157	980	919	271	1333	1029	2553	751	704	720	783	455
5					478	583	996	858	855	1504	677	651	600	401	1033
6						115	740	470	379	1581	271	289	261	308	687
7							667	455	288	1661	177	216	207	343	592
8								1066	759	2320	493	454	479	598	206
9									328	1387	591	650	656	776	933
10										1697	333	400	427	622	610
11											1838	1868	1841	1789	2248
12												68	105	336	417
13													52	287	406
14														237	449
15															636

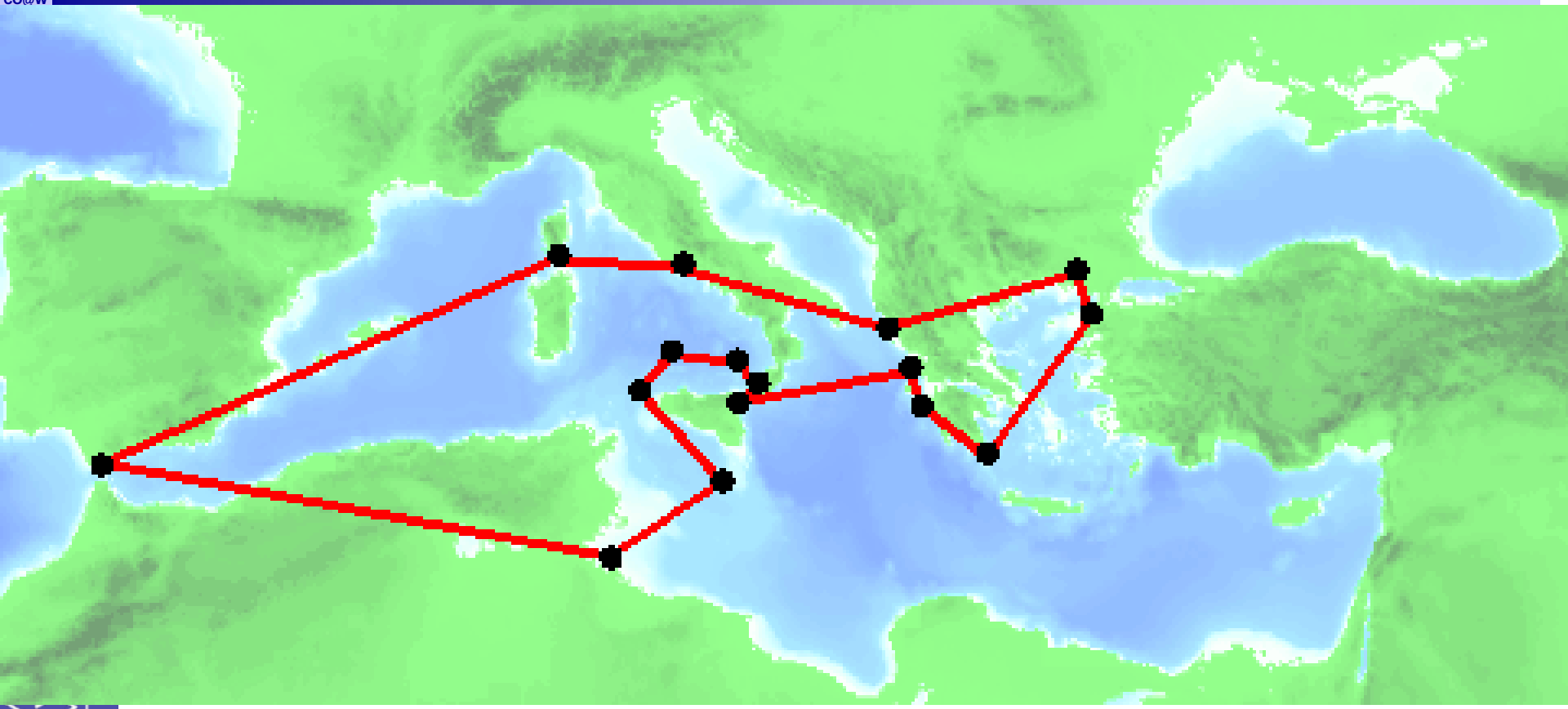
Table 2. The distance table for Ulysses 2000.

The distance table

1	Ithaca	38.24N	20.42E
2	Troy	39.57N	26.15E
3	Maronia	40.56N	25.32E
4	Malea	36.26N	23.12E
5	Djerba	33.48N	10.54E
6	Favignana	37.56N	12.19E
7	Ustica	38.42N	13.11E
8	Zakinthos	37.52N	20.44E
9	Bonifaccio	41.23N	9.10E
10	Circeo	41.17N	13.05E
11	Gibraltar	36.08N	5.21W
12	Stromboli	38.47N	15.13E
13	Messina	38.15N	15.35E
14	Taormina	37.51N	15.17E
15	Birzebbuga	35.49N	14.32E
16	Corfu	39.36N	19.56E

Table 1. Polar coordinates of the 16 locations in the Mediterranean.

Ulysses roundtrip

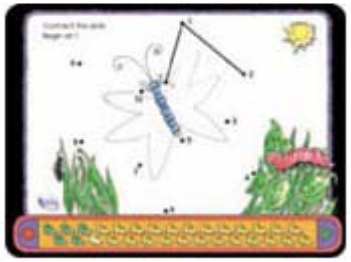


optimal „Ulysses tour“

ZIB
Martin Grötschel

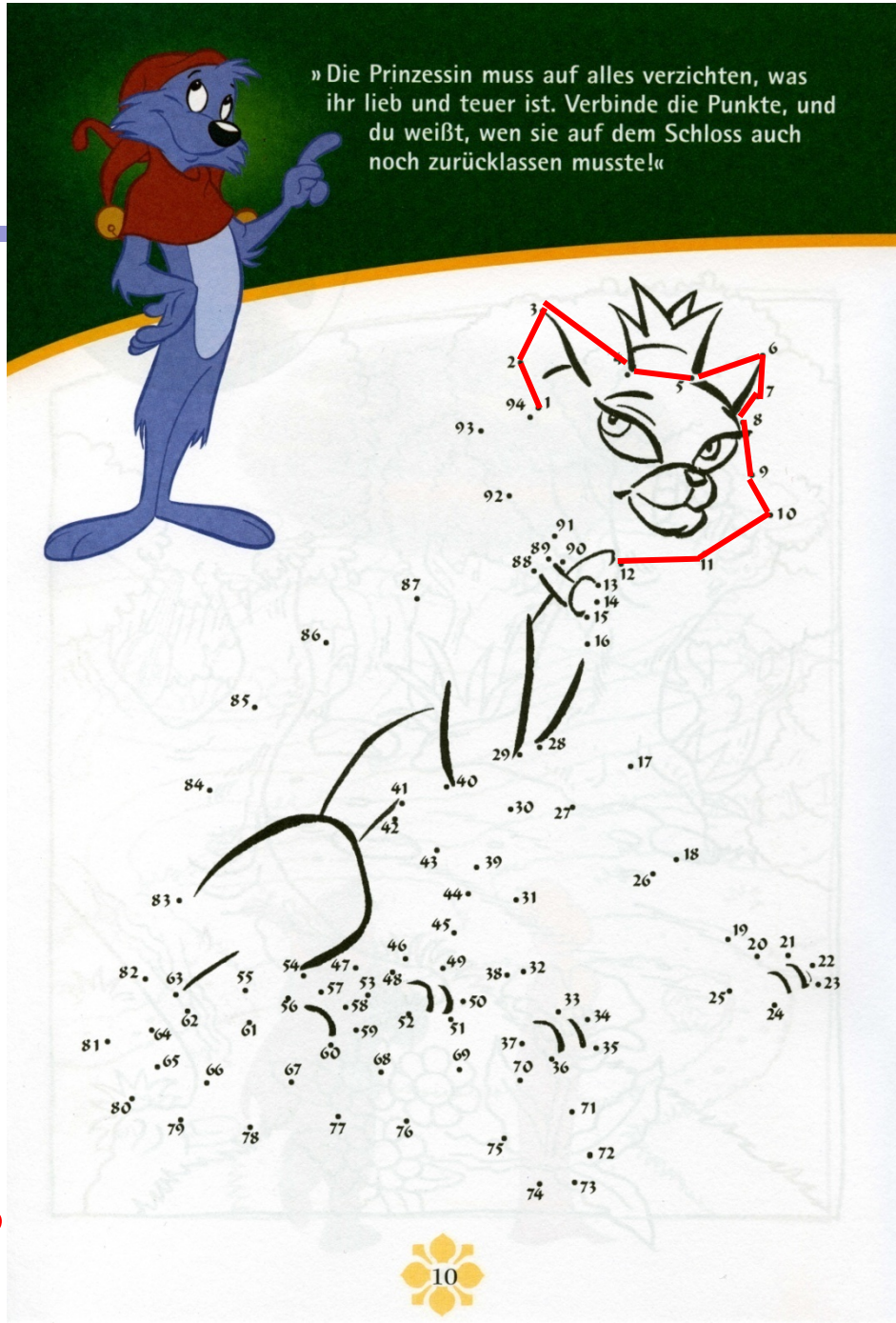
Malen nach Zahlen TSP in art ?

Number Sequence



Drawing By Numbers

- When was this invented?



Martin Grötschel

http://www.touregypt.net/featurestories/flinders.htm

The TSP in archeology

- Flinders Petrie (1853-1942) and the Luxor graves
- In the words of James Baikie, author of the book *A Century of Excavation in the Land of the Pharaohs,* "if the name of any one man must be associated with modern excavation as that of the chief begetter of its principles and methods, it must be the name of Professor Sir W.M. Flinders Petrie. It was he…who first called the attention of modern excavators to the importance of "unconsidered trifles" as means for the construction of the past…the broken earthenware of a people may be of far greater value than its most gigantic monuments."
- Petrie began to analyze the grave goods methodically. Grave A might contain certain types of pot in common with Grave B; Grave B also contained a later style of pot, the only type to be found in Grave C. By writing cards for each grave and filing them in logical order, Petrie established a full sequence for the cemetery, concluding that the last graves were probably contemporary with the First Dynasty. The development of life along the Nile thus was revealed, from early settlers to farmers to political stratification.

The TSP in archeology: Flinders Petrie's contribution

- Introduction of the "Hamming distance of graves", before Richard Wesley Hamming (1915 –1998) introduced it in mathematics.

  (The Hamming distance is used in telecommunication to count the number of flipped bits in a fixed-length binary word, an estimate of error. Hamming weight analysis of bits is used in several disciplines including information theory, coding theory, and cryptography.)

- Definition of the hamiltonian path problem through "graves".

The Seminal DFJ-Paper of 1954

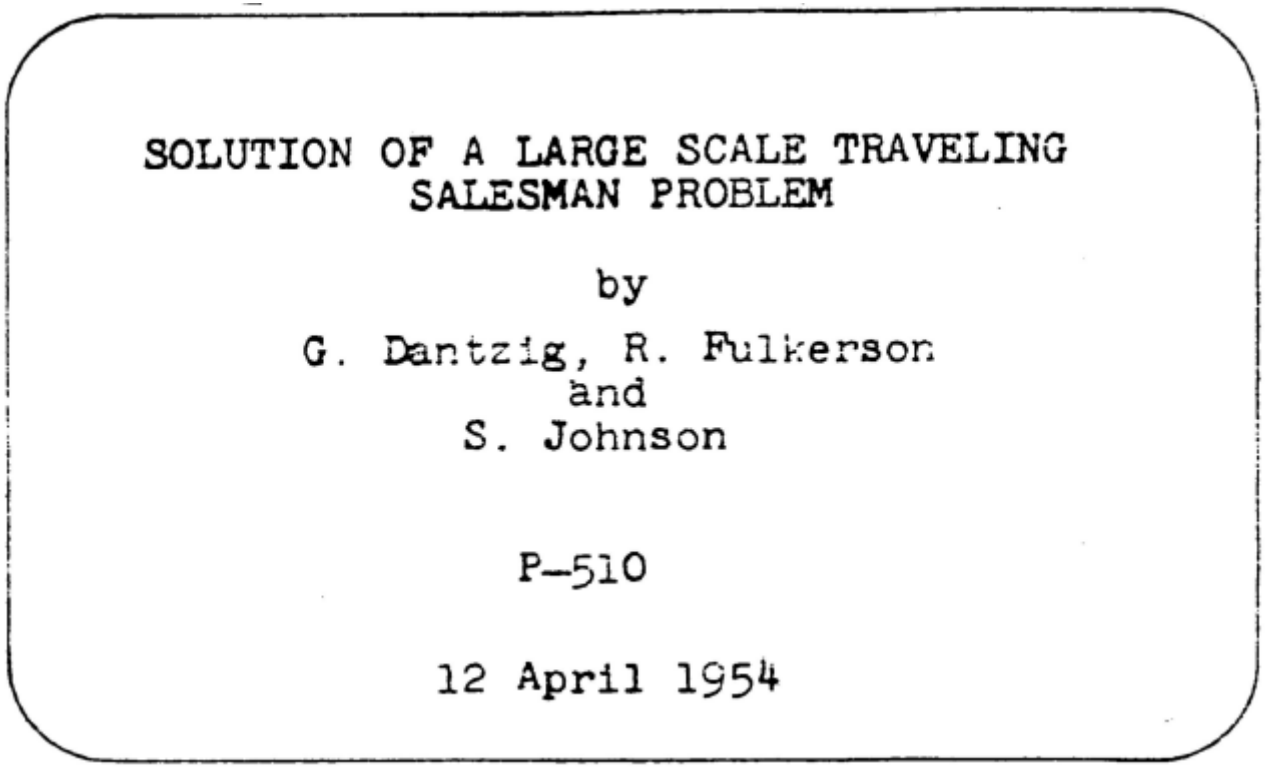
SOLUTION OF A LARGE-SCALE TRAVELING-SALESMAN PROBLEM*

G. DANTZIG, R. FULKERSON, AND S. JOHNSON
The Rand Corporation, Santa Monica, California

(Received August 9, 1954)

It is shown that a certain tour of 49 cities, one in each of the 48 states and Washington, D. C., has the shortest road distance.

The Seminal DFJ-Paper of 1954 preprint



SOLUTION OF A LARGE SCALE TRAVELING SALESMAN PROBLEM

by

G. Dantzig, R. Fulkerson
and
S. Johnson

P–510

12 April 1954

G. Dantzig, R. Fulkerson, S. Johnson, *Solution of a Large Scale Traveling Salesman Problem*, Paper P-510, The RAND Corporation, Santa Monica, California, [12 April] 1954. [53, 984, 997, 999, 1003]

G. Dantzig, R. Fulkerson, S. Johnson, Solution of a large-scale traveling-salesman problem, *Journal of the Operations Research Society of America* 2 (1954) 393–410. [6, 53, 984, 995]

Some Quotes from DFJ 1954

Since there are only a finite number of possibilities (at most $\frac{1}{2}(n-1)!$) to consider, the problem is to devise a method of picking out the optimal arrangement which is reasonably efficient for fairly large values of n.

For undirected tours, the symbol x_{IJ} will be treated identically with x_{JI} so that we may rewrite (1) as

$$\sum_{J=1}^{n} x_{IJ}=2. \qquad (x_{IJ}\geq 0;\ I=1,2,\cdots,n;\ I\neq J;\ x_{IJ}\equiv x_{JI}) \quad (2)$$

IP Formulation

The problem is to find the minimum of the linear form

$$D(x)=\sum_{I>J} d_{IJ}\, x_{IJ}, \qquad (3)$$

where the $x_{IJ}=0$ or 1 and the $x_{IJ}=1$ form a tour, and where the summation in (3) extends over all indices (I,J) such that $I>J$.

To make a linear programming problem out of this (see ref. 2) one

Polyhedral Approach

needs, as we have observed, a way to describe tours by more linear restraints than that given by (2). This is extremely difficult to do as illustrated by work of I. Heller[4] and H. Kuhn.[6] They point out that such relations always exist. However, there seems to be no simple way to characterize them and for moderate size n the number of such restraints appears to be astronomical. In spite of these difficulties, this paper will

An important class of conditions that tours satisfy, which excludes many non-tour cases satisfying (2), are the 'loop conditions.' These are linear inequality restraints that exclude subcycles or loops. Consider a non-tour solution to (2) which has a subtour of $n_1 < n$ cities; we note that the sum of the x_{IJ} for those links (I,J) in the subtour is n_1. Hence we can

eliminate this type of solution by imposing the condition that the sum of x_{IJ} over all links (I,J) connecting cities in the subset S of n_1 cities be less than n_1, i.e.,

$$\sum_{S} x_{IJ} \leq n_1 - 1 \tag{4}$$

where the summation extends over all (I,J) with I and J in the n_1 cities S. From (2) we note that two other conditions, each equivalent to (4), are

$$\sum_{\bar{S}} x_{IJ} \leq n - n_1 - 1, \tag{5}$$

where $\bar{S}$ means the summation extends over all (I,J) such that neither I nor J is in S, and

$$\sum_{S\bar{S}} x_{IJ} \geq 2, \tag{6}$$

where $S\bar{S}$ means that the summation extends over all (I,J) such that I is in S and J not in S.

There are, however, other more complicated types of restraints which sometimes must be added to (2) in addition to an assortment of loop conditions in order to exclude solutions involving fractional weights x_{IJ}. In the 49-city case we needed two such conditions. However, later when

Subtour Elimination Constraints in several forms

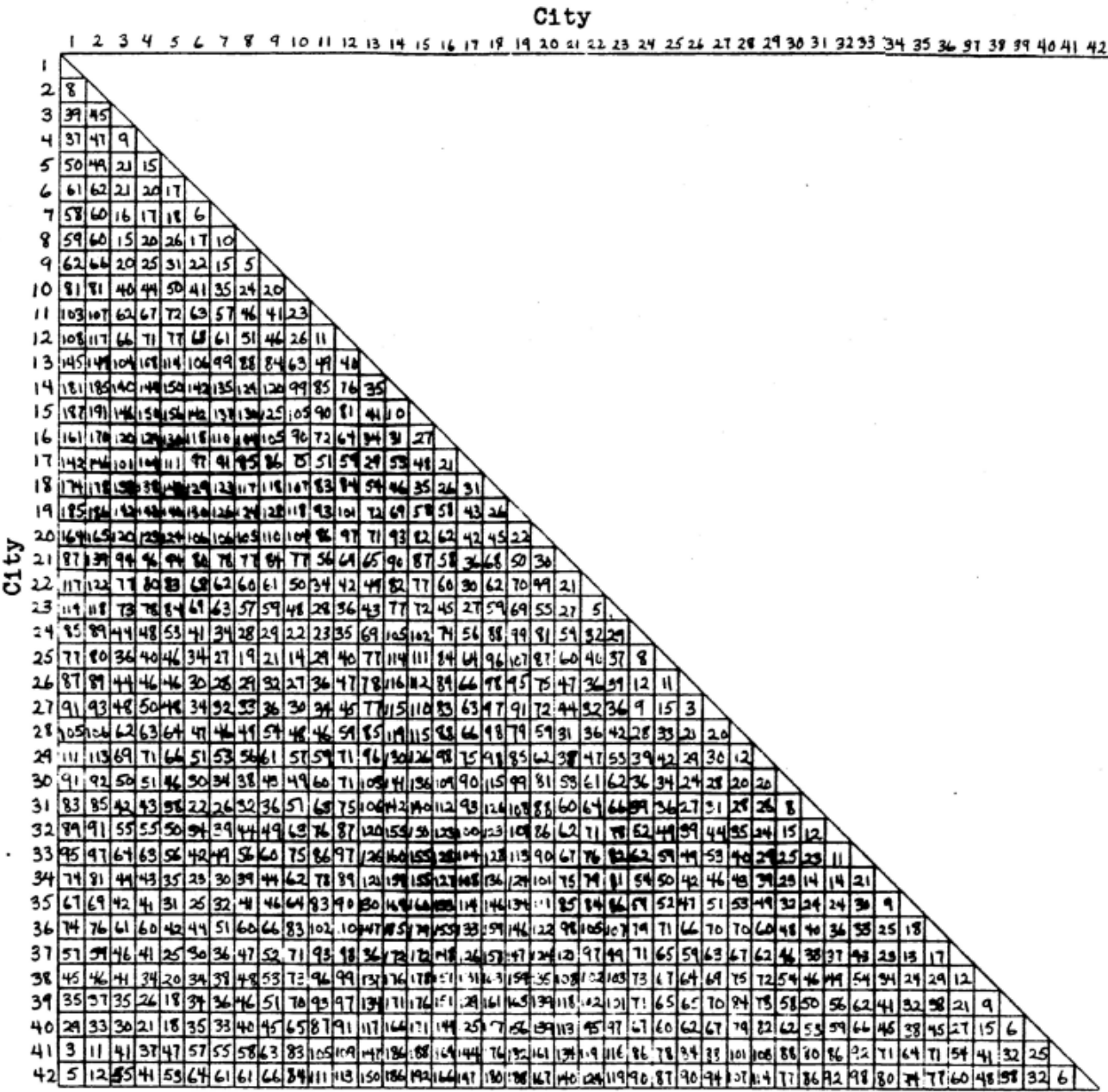
Remarks

- The preprint version is much clearer than the published paper. The editors have replaced abstract insight by a sequence of examples and thus almost destroyed the "real" contents of the paper.
- The authors outline the branch and bound technique.
- They explain the cutting plane methodology and observe clearly where the difficulties and chances of this method are.
- They mention the importance of heuristics.
- They are modest:

CONCLUDING REMARK

It is clear that we have left unanswered practically any question one might pose of a theoretical nature concerning the traveling-salesman problem; however, we hope that the feasibility of attacking problems involving a moderate number of points has been successfully demonstrated, and that perhaps some of the ideas can be used in problems of similar nature.

Table of Road Distances between Cities
in Adjusted Units



The Authors provide data

Distance table
hand-written
by D. R. Fulkerson
(from the preprint
Bob Bland owns)

The Optimal Solution

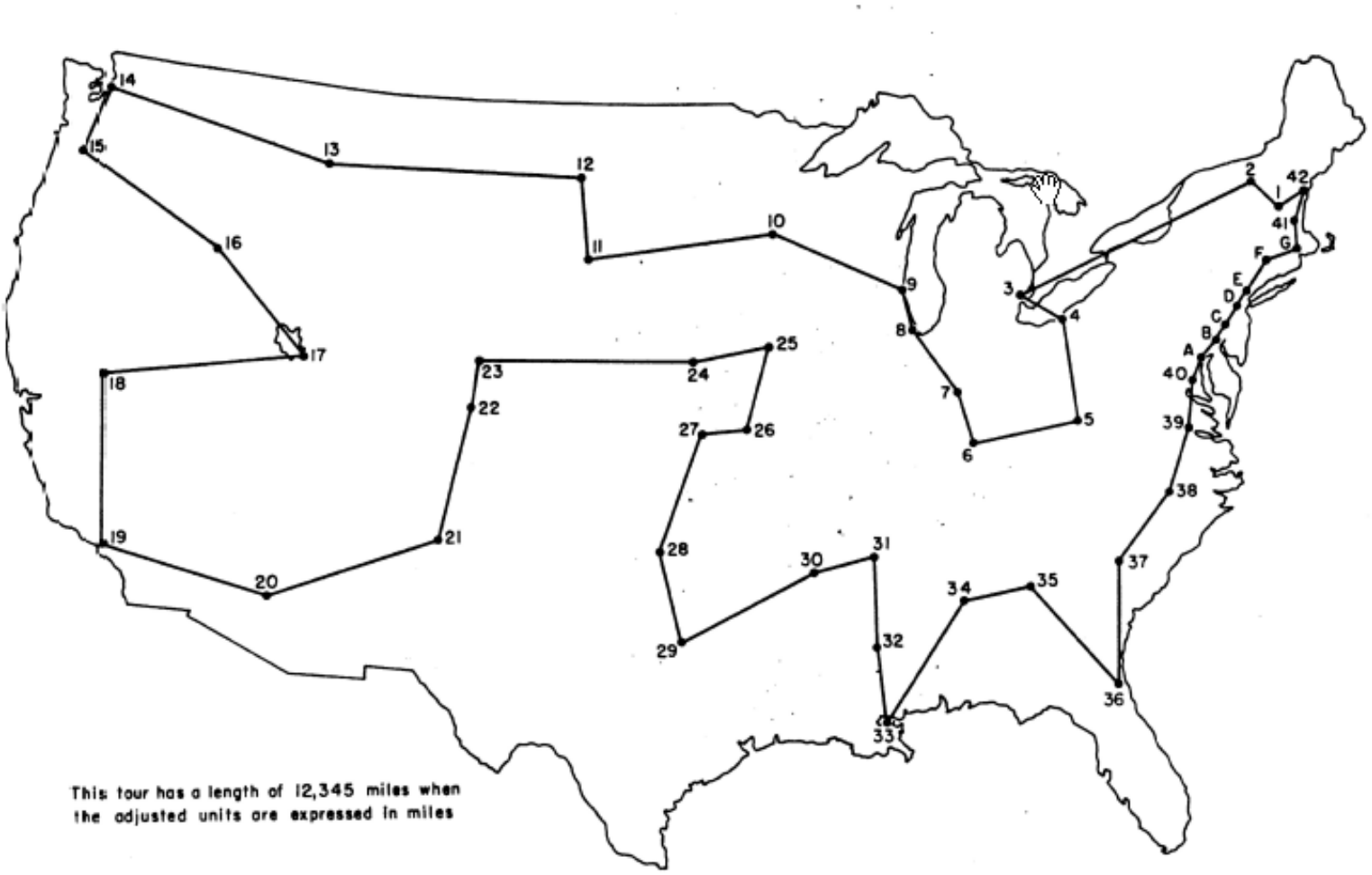


FIG. 16. The optimal tour of 49 cities.

Martin Grötschel

Survey Books

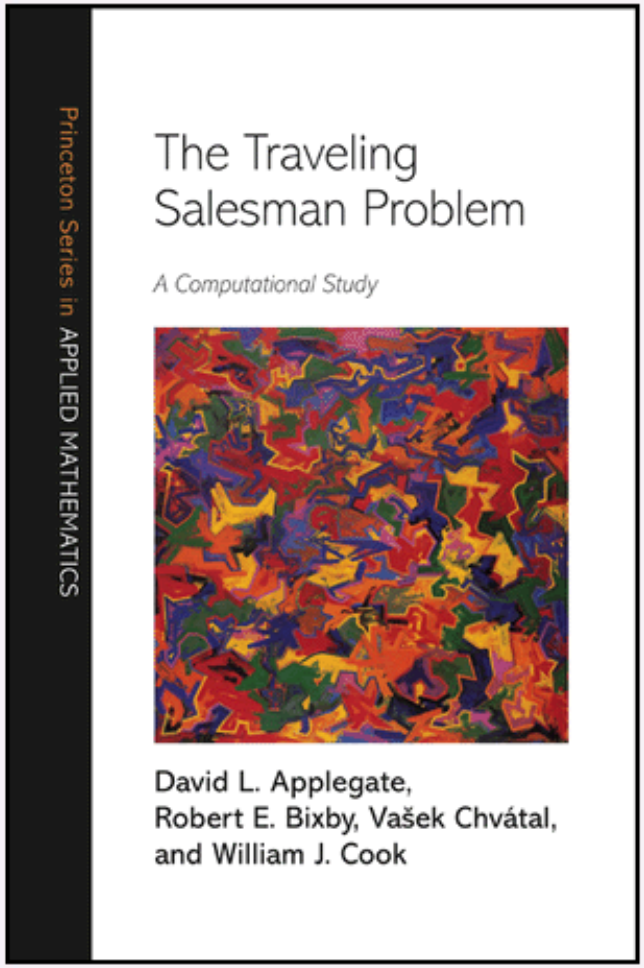
Literature: more than 1300 entries in Zentralblatt/Math

Zbl 0562.00014 **Lawler, E.L.(ed.); Lenstra, J.K.(ed.); Rinnooy Kan, A.H.G.(ed.); Shmoys, D.B.(ed.)**
The traveling salesman problem. A guided tour of combinatorial optimization. Wiley-Interscience Series in Discrete Mathematics. A Wiley-Interscience publication. Chichester etc.: John Wiley \& Sons. X, 465 p. (1985). *MSC 2000:* *00Bxx 90-06

Zbl 0996.00026 **Gutin, Gregory (ed.); Punnen, Abraham P.(ed.)**
The traveling salesman problem and its variations. Combinatorial Optimization. 12. Dordrecht: Kluwer Academic Publishers. xviii, 830 p. (2002). *MSC 2000:* *00B15 90-06 90Cxx

THE TSP book

TSP Book



The Traveling Salesman Problem: A Computational Study

David L. Applegate, Robert E. Bixby, Vasek Chvátal & William J. Cook

Winner of the 2007 Lanchester Prize, INFORMS

Princeton University Press 2006, 606 pp.

Our primary concern in this book is to describe a method and computer code that have succeeded in solving a wide range of large-scale instances of the TSP. Along the way we cover the interplay of applied mathematics and increasingly more powerful computing platforms, using the solution of the TSP as a general model in computational science.

- Table of Contents
- Links to Bookstores

Cover illustration by Julian Lethbridge, *Traveling Salesman 4*, 1995, oil on linen, 72 x 72 inches, The Robert and Jane Meyerhoff Collection, photograph by Adam Reich.

THE TSP Web page
http://www.tsp.gatech.edu/index.html

CO@W

TSP

Home

The Traveling Salesman Problem

The Traveling Salesman Problem is one of the most intensively studied problems in computational mathematics. These pages are devoted to the history, applications, and current research of this challenge of finding the shortest route visiting each member of a collection of locations and returning to your starting point.

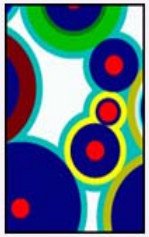
New!	Mona Lisa TSP	A 100,000-city challenge problem ($100 Prize).
New!	Google Maps	Plot an optimal TSP tour with a Google interface.
New!	pla85900	Solution of a 85,900-city TSP.
New!	Ron Schreck's Flight	All 109 public airports in North Carolina in a single day!

- > Home
- The Problem
- History
- Applications
- Solving a TSP
- World Records
- Gallery
- TSP Games
- Google Maps
- Concorde
- Test Data
- News
- TSP Book
- Search Site

TSP Book

The Traveling Salesman Problem: A Computational Study by Applegate, Bixby, Chvatal, and Cook.

Lower Bounds



Description of the techniques we use to compute lower bounds on the lengths of all TSP tours.

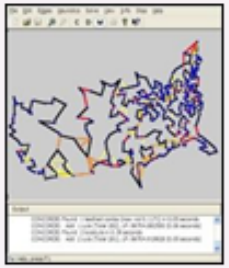
Sweden TSP

Optimal solution for visiting all 24,978 cities in Sweden. Tour has length approximately 72,500 kilometers.

Car 54

The TSP was featured in a contest run by Proctor and Gamble in 1962. The challenge problem had 33 cities.

Concorde for Windows



A graphical user interface available for Concorde on Windows' platforms.

Contents

1. Introduction
2. The TSP and some of its history
3. The TSP and some of its variants
4. Some applications
5. Heuristics
6. How combinatorial optimizers do it

The Travelling Salesman Problem and Some of its Variants

- The symmetric TSP
- The asymmetric TSP
- The TSP with precedences or time windows
- The online TSP
- The symmetric and asymmetric m-TSP
- The price collecting TSP
- The Chinese postman problem (undirected, directed, mixed)
- Bus, truck, vehicle routing
- Edge/arc & node routing with capacities
- Combinations of these and more

http://www.densis.fee.unicamp.br/~moscato/TSPBIB_home.html

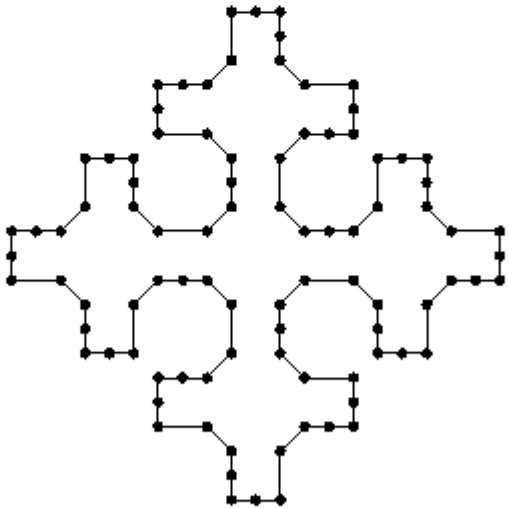
TSPBIB Home Page

This page intends to be a comprehensive listing of papers, source code, preprints, technical reports, etc, available on the Internet about the Traveling Salesman Problem (TSP) and some associated problems.

Please send us information about any other work you consider it should be included in this page.

Pablo Moscato

email: moscato@cacr.caltech.edu
email: moscato@densis.fee.unicamp.br



The picture above shows an instance of the Euclidean, Planar TSP
and the optimal curve among the set of cities.
This instance has been named MNPeano Order 2.

Contents

1. Introduction
2. The TSP and some of its history
3. The TSP and some of its variants
4. Some applications
5. Modeling issues
6. Heuristics
7. How combinatorial optimizers do it
8. Art, Astronomy & Astrology

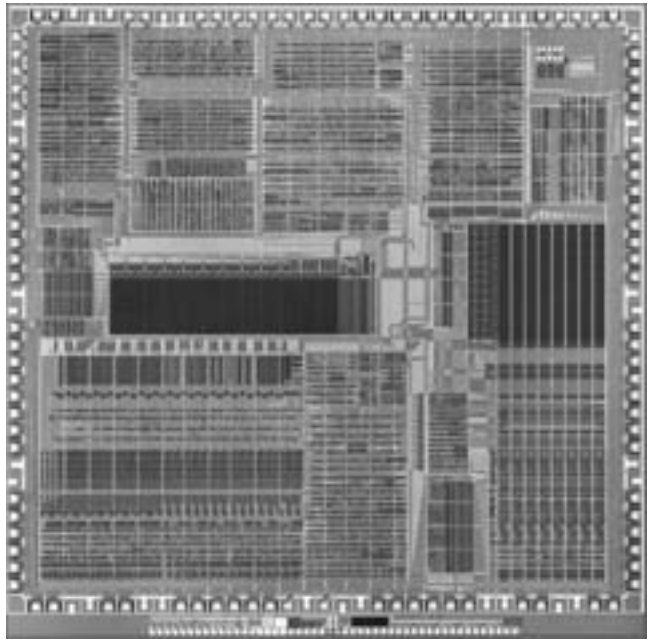
Application list from http://www.tsp.gatech.edu/index.html

Applications

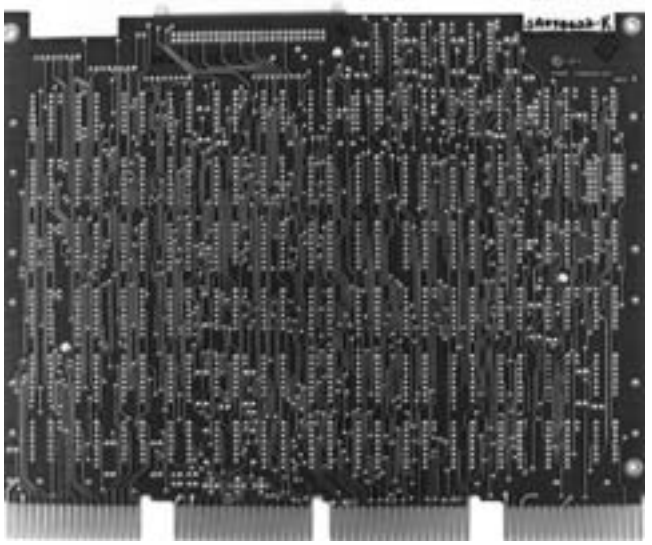
- Genome
- Starlight
- Scan Chains
- DNA
- Whizzkids
- Baseball
- Coin Collection
- Airport Tours
- USA Trip
- Sonet Rings
- Power Cables

We will see many TSP applications.

Production of ICs and PCBs



Integrated Circuit (IC)

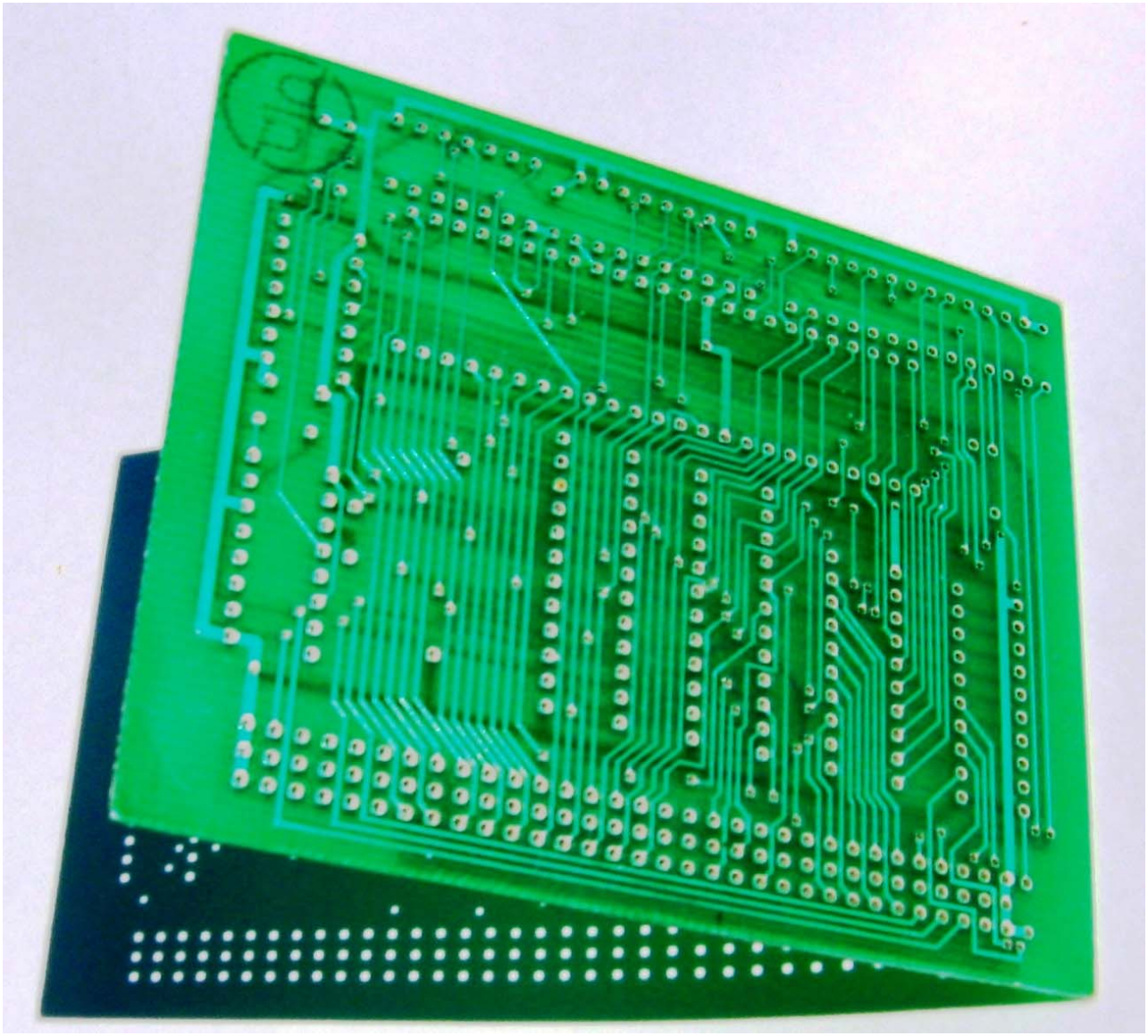


Printed Circuit Board (PCB)

Problems: Logical Design, Physical Design Correctness, Simulation, Placement of Components, Routing, Drilling,...

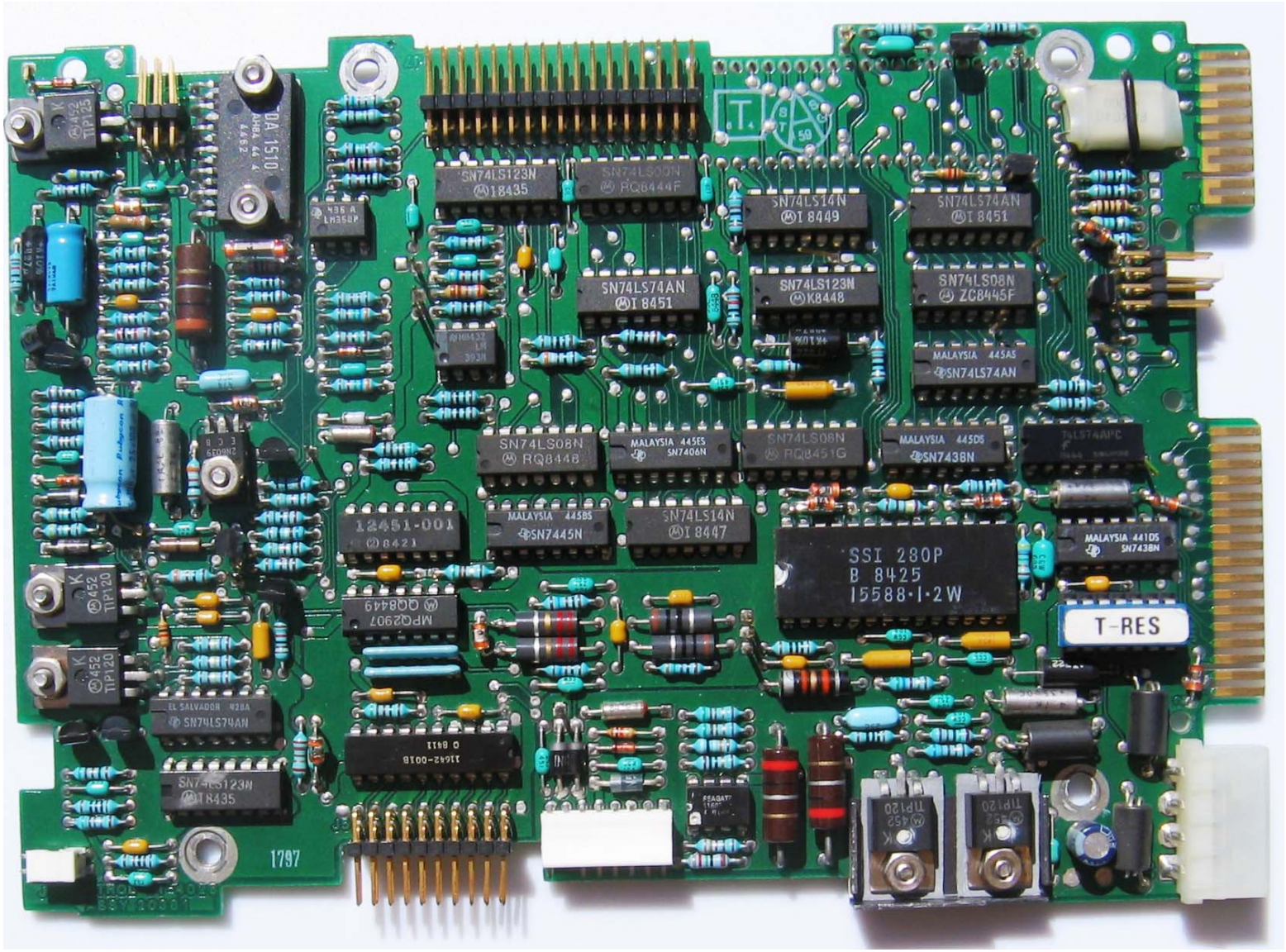
Martin Grötschel

442 holes to drill



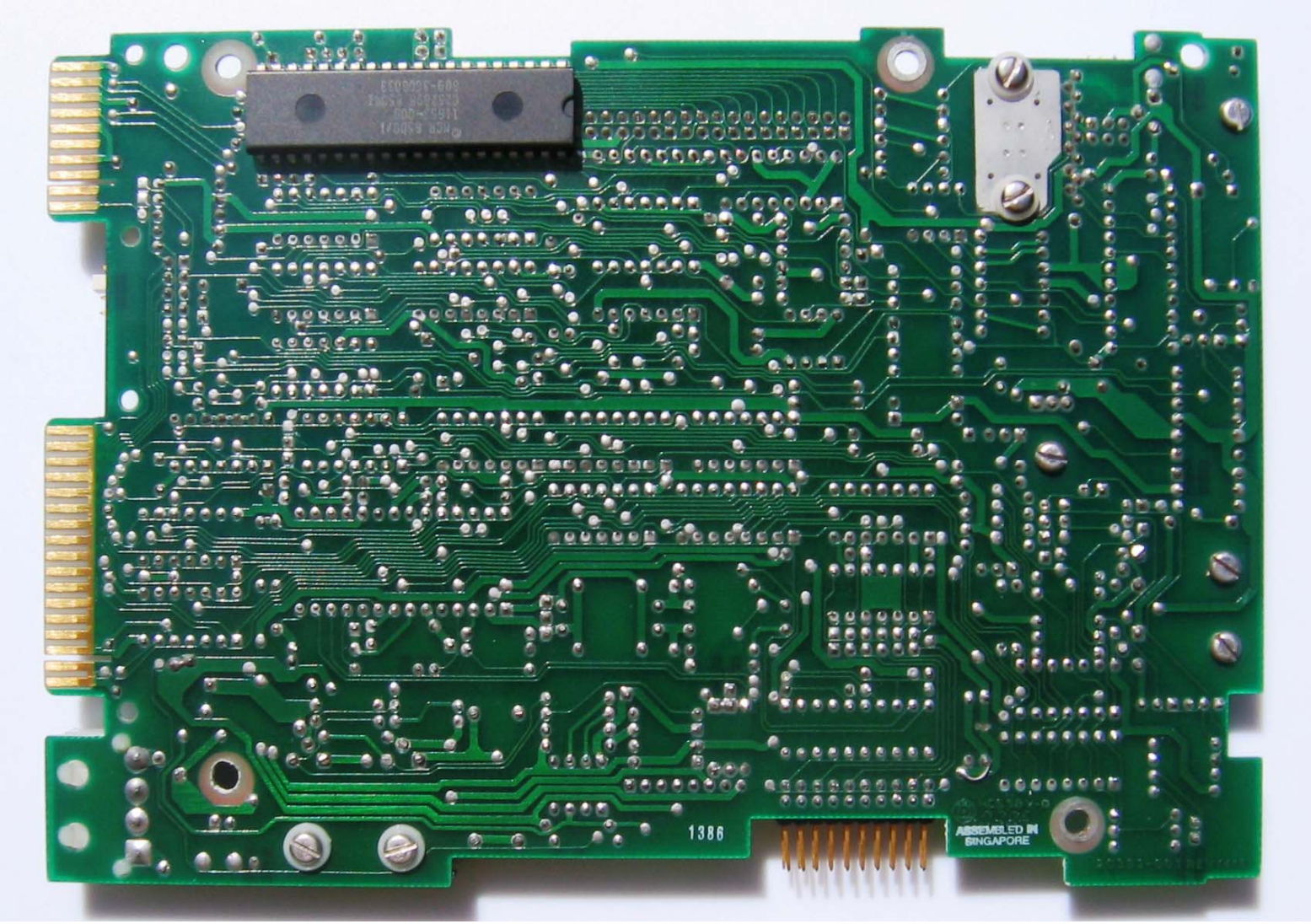
Martin Grötschel

Printed Circuit Board (front)



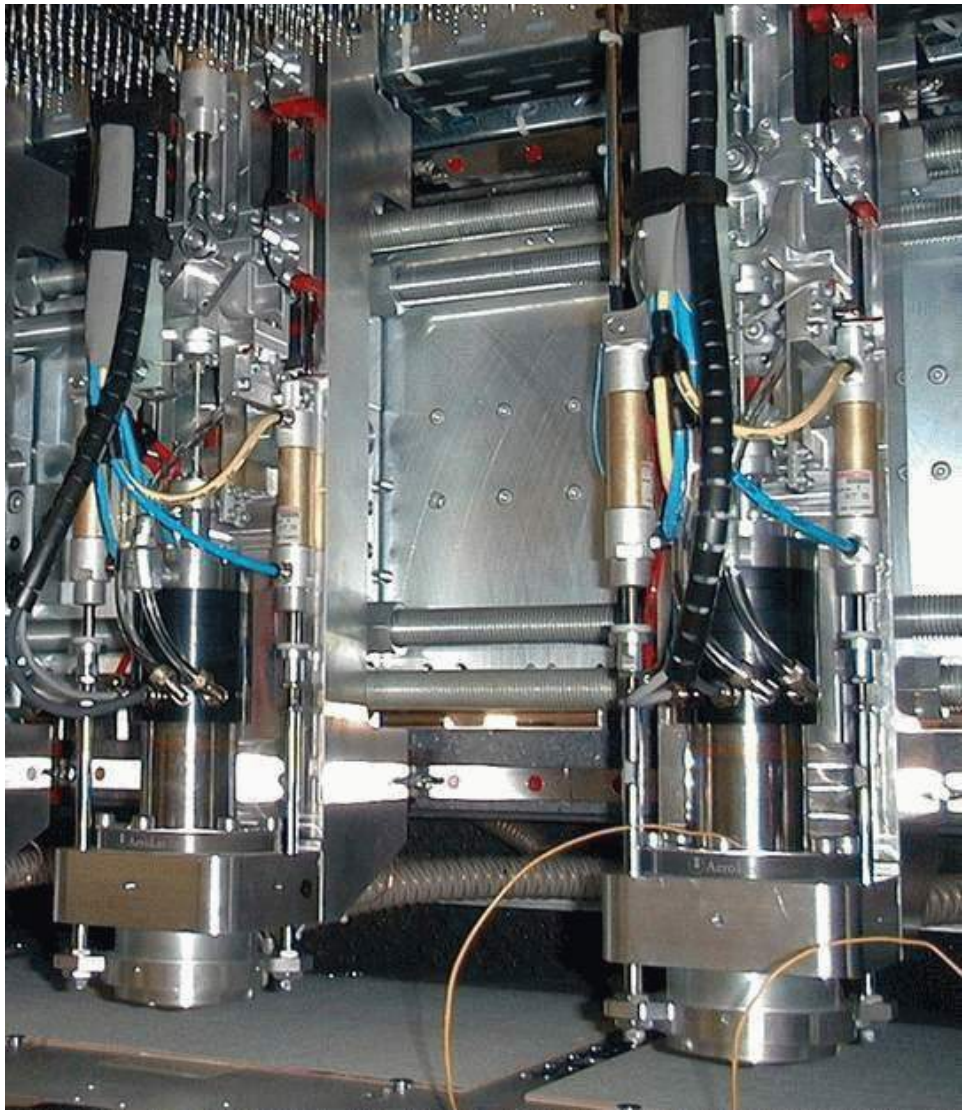
Martin
Grötschel

Printed Circuit Board (back)

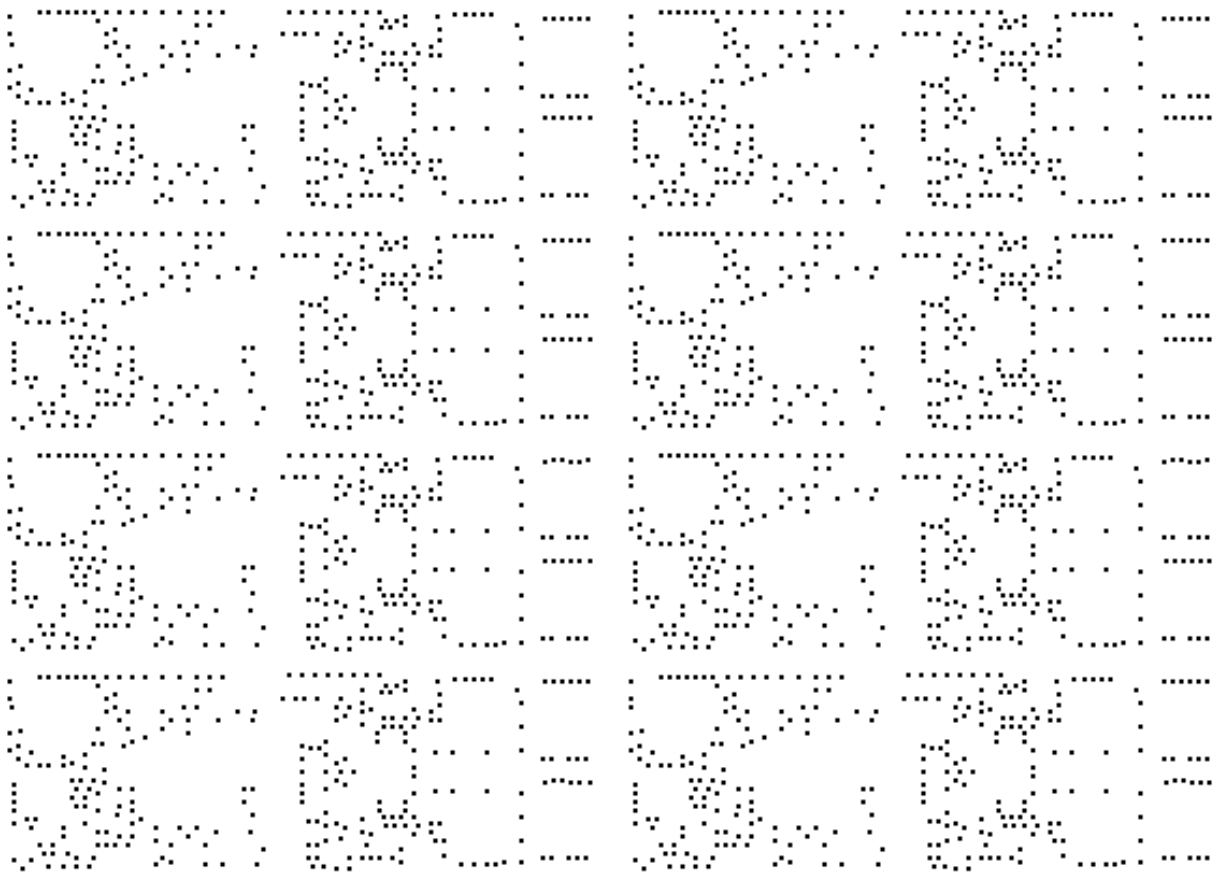


Martin
Grötschel

Drilling Machine



Correct modelling of a printed circuit board drilling problem



length of a move of the drilling head:

Euclidean norm,

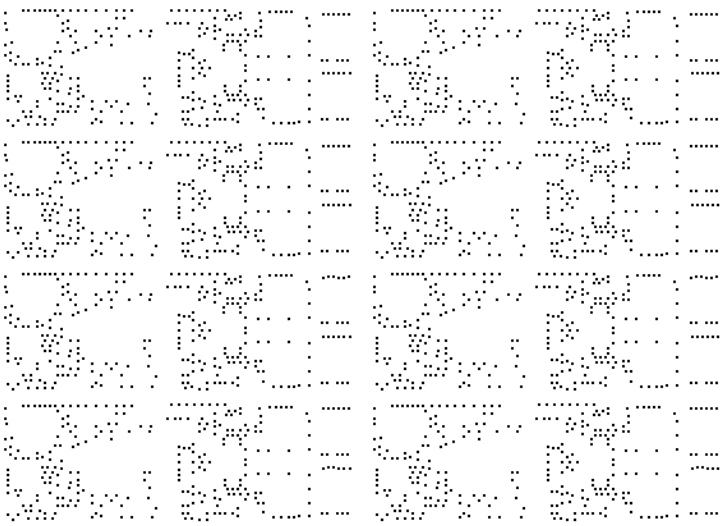
Max norm,

Manhatten norm?

2103 holes to be drilled

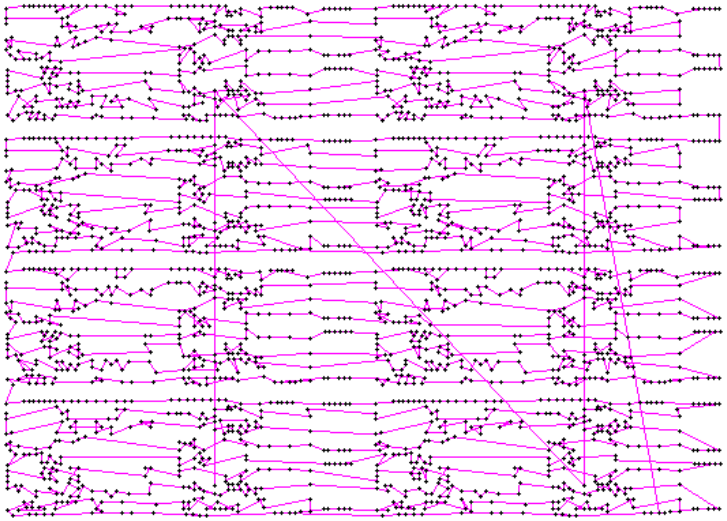
Martin Grötschel

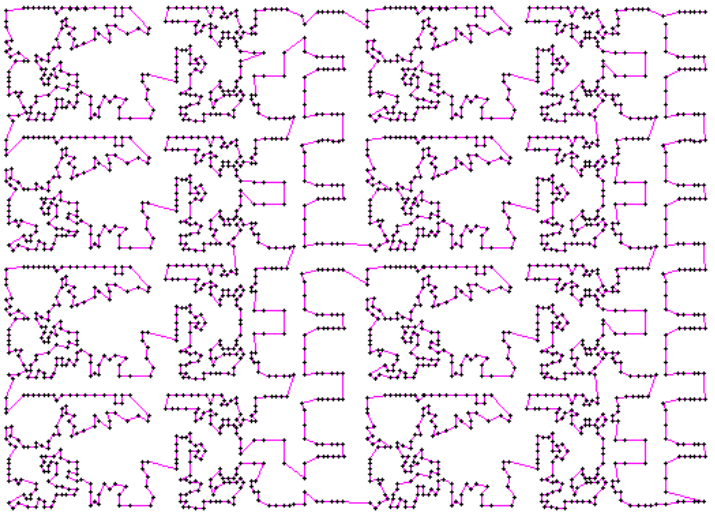
Drilling 2103 holes into a PCB



Significant Improvements
via TSP

(Padberg & Rinaldi)





industry solution

optimal solution

Martin
Grötschel

Siemens-Problem PCB da1

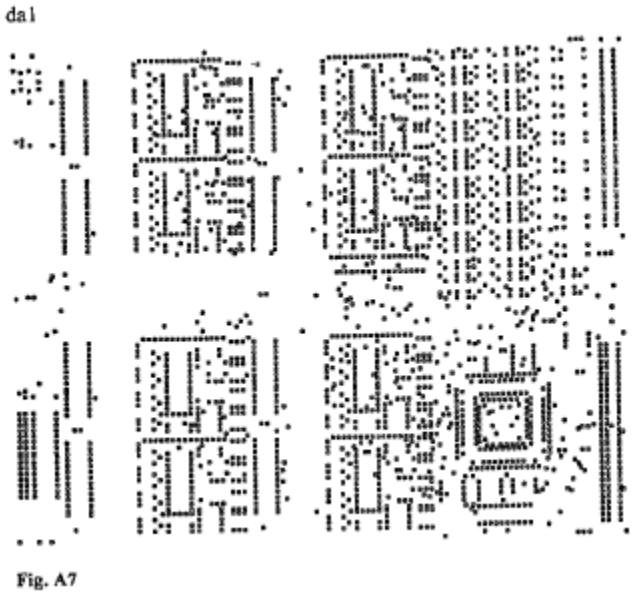
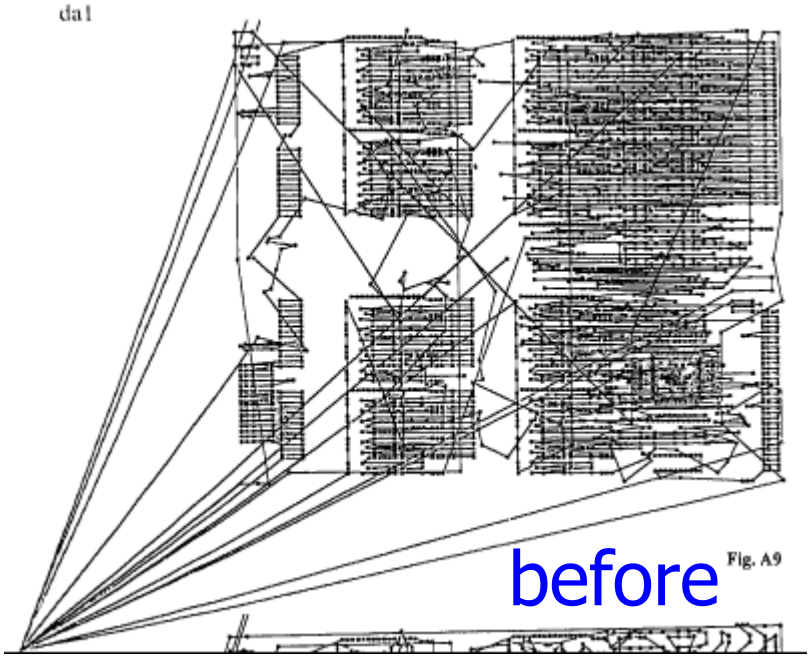
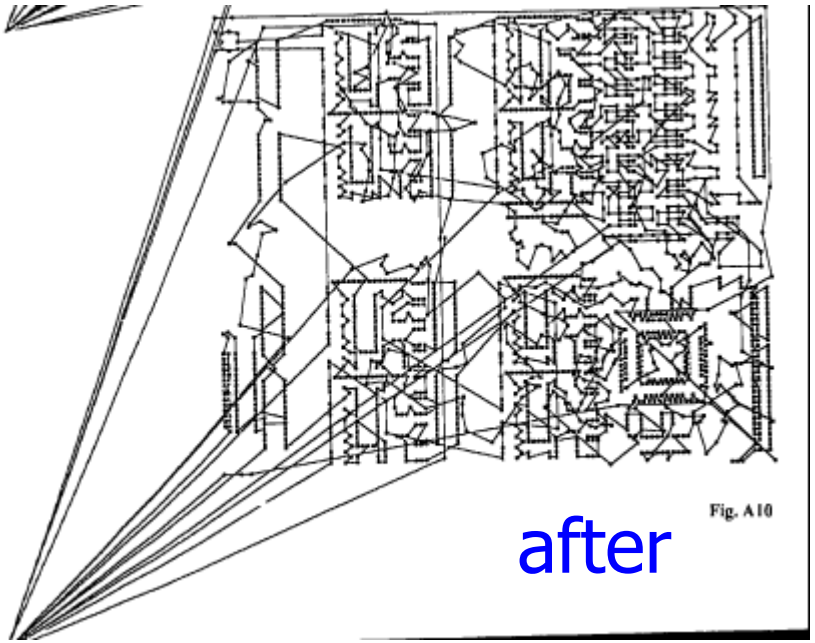


Fig. A7

Grötschel, Jünger, Reinelt





Siemens-Problem PCB da4

Fig. A8

Grötschel, Jünger, Reinelt

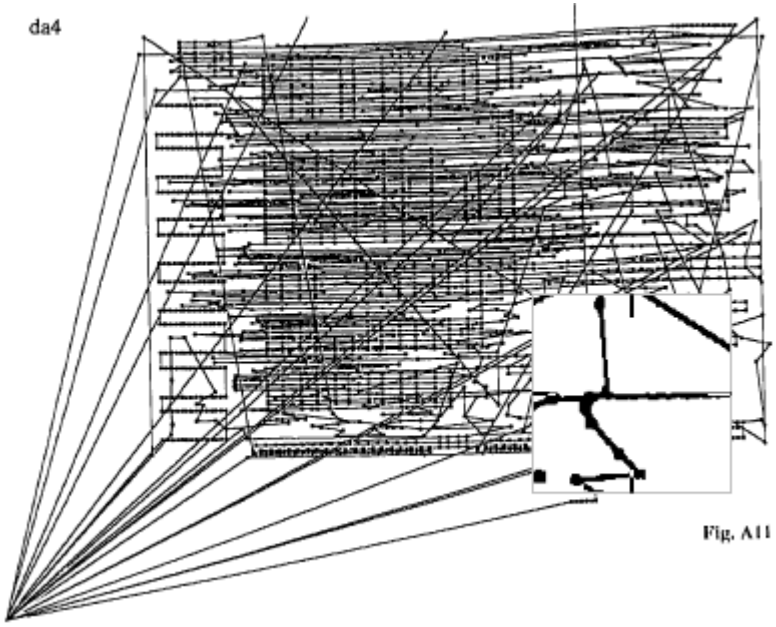


Fig. A11

before

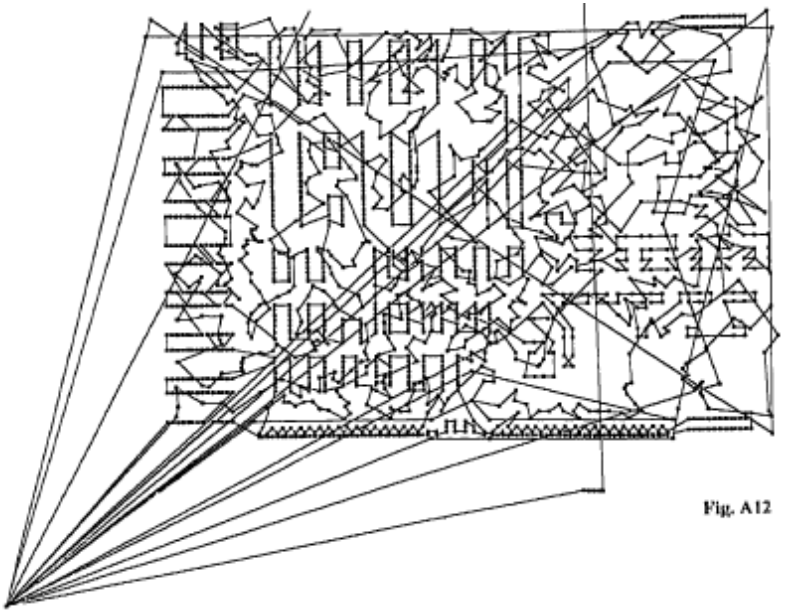


Fig. A12

after

Example: Control of the stacker cranes in a Herlitz warehouse

Martin Grötschel

Contents

1. Introduction
2. The TSP and some of its history
3. The TSP and some of its variants
4. Some applications
5. Heuristics
6. How combinatorial optimizers do it

Need for Heuristics

- Many real-world instances of hard combinatorial optimization problems are (still) too large for exact algorithms.
- Or the time limit stipulated by the customer for the solution is too small.
- Therefore, we need heuristics!
- Exact algorithms usually also employ heuristics.
- What is urgently needed is a decision guide:

  Which heuristic will most likely work well on what problem ?

Primal and Dual Heuristics

- Primal Heuristic: Finds a (hopefully) good feasible solution.
- Dual Heuristic: Finds a bound on the optimum solution value (e.g., by finding a feasible solution of the LP-dual of an LP-relaxation of a combinatorial optimization problem).

Minimization:

dual heuristic value $\leq$ optimum value $\leq$ primal heuristic value

(In maximization the inequalities are the other way around.)

quality guarantee
in practice and theory

Primal and Dual Heuristics

Primal and Dual Heuristics give rise to worst-case guarantee:

Minimization:

optimum value ≤ primal heuristic value
≤ (1+ε) optimum value

dual heuristic value ≤ primal heuristic value
≤ (1+ε) dual heuristic value

(In maximization the inequalities are the other way around.)

quality guarantee
in practice and theory

Heuristics: A Survey

- Greedy Algorithms
- Exchange & Insertion Algorithms
- Neighborhood/Local Search
- Variable Neighborhood Search, Iterated Local Search
- Random sampling
- Simulated Annealing
- Taboo search
- Great Deluge Algorithms
- Simulated Tunneling
- Neural Networks
- Scatter Search
- Greedy Randomized Adaptive Search Procedures

Heuristics: A Survey

- Genetic, Evolutionary, and similar Methods
- DNA-Technology
- Ant and Swarm Systems
- (Multi-) Agents
- Population Heuristics
- Memetic Algorithms (Meme are the "missing links" gens and mind)
- Space Filling Curves
- Fuzzy Logic Based…
- Fuzzy Genetics-Based Machine Learning
- Fast and Frugal Method (Psychology)
- Ecologically rational heuristic (Sociology)
- Method of Devine Intuition (Psychologist Thorndike)
- …..

An Unfortunate Development

- There is a marketing battle going on with unrealistic, or even ideological, claims about the quality of heuristics – just to catch attention
- Linguistic Overkill:
- Vodoo Approach

A Quote

Quote:
Genetic Programming is an evolutionary computation technique which searches for those computer programs that best solve a given problem.

(Will this also solve P = NP?)

Kalyanmoy Deb:
„Multi-objective optimization using evolutionary algorithms“ (Wiley, 2001)

from the Preface

- Optimization is a procedure of finding and comparing feasible solutions until no better solution can be found.

- Evolutionary algorithms (EAs), on the other hand, can find multiple optimal solutions in one single simulation run due to their population-approach. Thus, EAs are ideal candidates for solving…

?

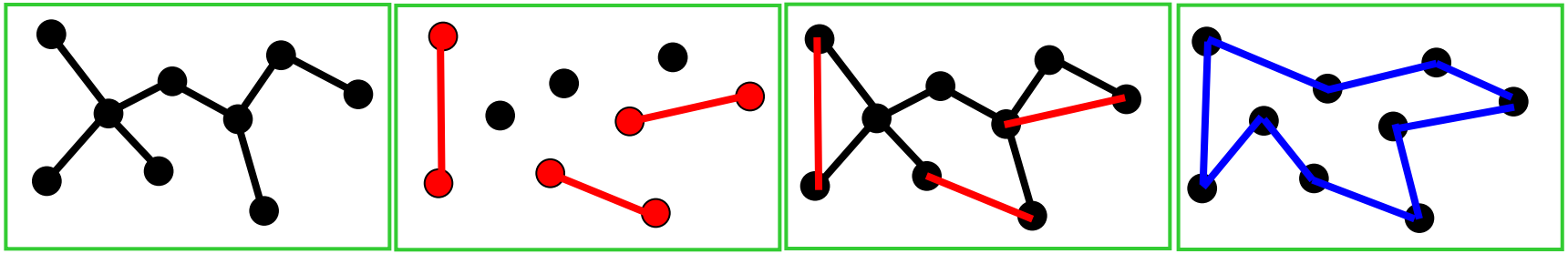
Heuristics: A Survey

Currently best heuristic with respect to worst-case guarantee:

Christofides heuristic

- compute a shortest spanning tree
- compute a minimum perfect 1-matching of the graph induced by the odd nodes of the minimum spanning tree
- the union of these edge sets is a connected Eulerian graph
- turn this graph into a TSP-tour by making short-cuts.

For distance functions satisfying the triangle inequality, the resulting tour is at most 50% above the optimum value



Understanding Heuristics, Approximation Algorithms

- worst case analysis
  - There is no polynomial time approx. algorithm for STSP/ATSP.
  - Christofides algorithm for the STSP with triangle inequality
- average case analysis
  - Karp's analysis of the patching algorithm for the ATSP
- probabilistic problem analysis
  - for Euclidean STSP in unit square: TSP constant 1.714..$n^{1/2}$
- polynomial time approximation schemes (PAS)
  - Arora's polynomial-time approximation schemes for Euclidean STSPs
- fully-polynomial time approximation schemes (FPAS)
  - not for TSP/ATSP but, e.g., for knapsack (Ibarra&Kim)

- These concepts – unfortunately – often do not really help to guide practice.
- experimental evaluation
  - Lin-Kernighan for STSP (DIMACS challenges))

Contents

1. Introduction
2. The TSP and some of its history
3. The TSP and some of its variants
4. Some applications
5. Heuristics
6. How combinatorial optimizers do it

Polyhedral Theory (of the TSP)

STSP-, ATSP-,TSP-with-side-constraints-

Polytope:= Convex hull of all incidence vectors of feasible tours

$$Q_T^n := conv\{\chi^T \in \mathbf{Z}^E \mid T \text{ tour in } K_n\} \qquad (\chi_{ij}^T = 1 \text{ if } ij \in T, \text{ else } = 0)$$

To be investigated:

- Dimension
- Equation system defining the affine hull
- Facets
- Separation algorithms

The symmetric travelling salesman polytope

$$Q_T^n := conv\{\chi^T \in \mathbf{Z}^E \mid T \text{ tour in } K_n\} \qquad (\chi_{ij}^T = 1 \text{ if } ij \in T, \text{ else } = 0)$$

$$\subseteq \{x \in \mathbf{R}^E \mid x(\delta(i)) = 2 \qquad \forall i \in V$$

$$x(E(W)) \leq |W| - 1 \quad \forall W \subset V \setminus \{1\}, 3 \leq |W| \leq n-3$$

$$0 \leq x_{ij} \leq 1 \qquad \forall ij \in E\}$$

- IP formulation

$$\min c^T x$$

$$x(\delta(i)) = 2 \qquad \forall i \in V$$

$$x(E(W)) \leq |W| - 1 \quad \forall W \subset V \setminus \{1\}, 3 \leq |W| \leq n-3$$

$$x_{ij} \in \{0,1\} \qquad \forall ij \in E$$

- The LP relaxation is solvable in polynomial time.

Dimension of the sym TSP polytope

Theorem. $\dim Q_T^n = \frac{1}{2}n(n-3) = |E| - |V| = \binom{n}{2} - n$

(a) The matrix A defined by $x(\delta(i)) = 2 \ \forall i \in V$ has full row rank.

(b) Q_T^n contains $|E| - |V| + 1$ linearly independent points.

Proof : (a) The submatrix B of A consisting of all rows and a set of columns corresponding to a spanning tree T plus an edge that forms one odd cycle with T is a nonsingular matrix.

(b) Use Lucas' result that a complete graph with 2k+1 nodes can be partitioned into k Hamiltonian cyles, and a complete graph with 2k nodes into k-1 Hamiltonian cycles and one perfect matching to construct the desired points (incidence vectors of tours).

Show cases n=4,5,6 on the blackboard.

Relation between IP and LP-relaxation

Open Problem:

- If costs satisfy the triangle inequality, then

  IP-OPT <= 4/3 LP-SEC

  IP-OPT <= 3/2 LP-SEC (Wolsey)

Facets of the TSP polytope

- Finding facets!
- Proving that an inequality defines a facet!
- Finding exact or heuristic separation algorithms to be used in a cutting plane algorithm!

Why are facets important?

An integer programming formulation from a textbook:

$$\min c^T x$$

$$x(\delta(i)) = 2 \qquad \forall i \in V$$

$$x(E) = n$$

$$x(C) \leq |C| - 1 \quad \forall C \subset E, C \text{ a nonhamiltonian cycle}$$

$$x_{ij} \in \{0,1\} \qquad \forall ij \in E$$

What would you say?

Martin Grötschel

Subtour elimination constraints: equivalent versions

- SEC constraints
- cut constraints

General cutting plane theory: Gomory Cut (the „rounding trick“)

Let $P = \{x \in \mathbf{R}^n \mid Ax \le b\}$ be a polyhedron, and we suppose that A and b are integral.

We would like to describe the convex hull P_I of all integral points in P.

Observation: For any $y \in \Re^m$

$y^T Ax \le y^T b$

is a valid inequality for P.

Observation: For any $y \in \Re^m$

$\lfloor y^T Ax \rfloor \le \lfloor y^T b \rfloor$

is a valid inequality for P_I.

$$\lfloor y^T Ax \rfloor = \sum_{j=}^{n} \sum_{i=1}^{m} \lfloor y_i a_{ij} \rfloor x_j \le \left\lfloor \sum_{i=1}^{m} y_i b_j \right\rfloor$$

Choose y so that $y_i a_{ij}$ is integral: $\lfloor y^T Ax \rfloor = \sum_{j=}^{n} \sum_{i=1}^{m} y_i a_{ij} x_j \le \left\lfloor \sum_{i=1}^{m} y_i b_j \right\rfloor$

Chvátal-Gomory Procedure

- Does the rounding procedure deliver P_I?
- How many rounds of rounding do we need?
- Other better methods?

General cutting plane theory: Gomory Mixed-Integer Cut

- Given $y, x_j \in \mathbb{Q}_+$, and

  $$y + \sum a_{ij} x_j = d = \lfloor d \rfloor + f,\ f > 0$$

- Rounding: Where $a_{ij} = \lfloor a_{ij} \rfloor + f_j$, define

  $$t = y + \sum \left(\lfloor a_{ij} \rfloor x_j : f_j \le f \right) + \sum \left(\lceil a_{ij} \rceil x_j : f_j > f \right) \in \mathbb{Q}$$

- Then

  $$\sum \left(f_j x_j : f_j \le f \right) + \sum \left(f_j - 1 \right) x_j : f_j > f = d - t$$

- Disjunction:

  $$t \le \lfloor d \rfloor \Rightarrow \sum \left(f_j x_j : f_j \le f \right) \ge f$$

  $$t \ge \lceil d \rceil \Rightarrow \sum \left(\left(1 - f_j \right) x_j : f_j > f \right) \ge 1 - f$$

- Combining

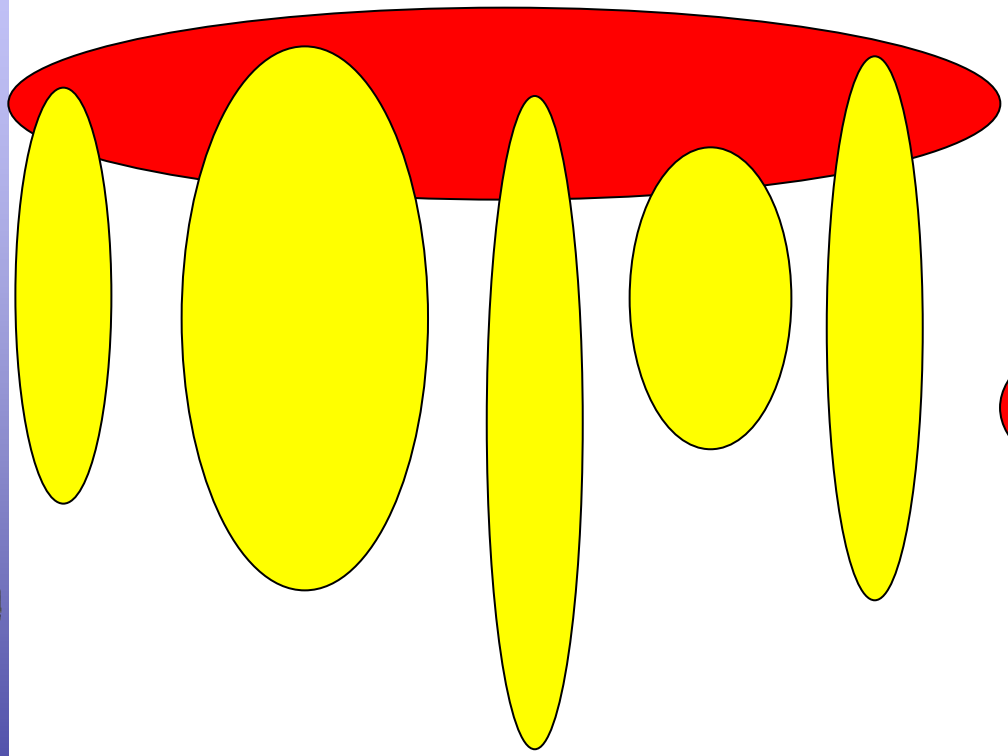
$$\sum \left(\left(f_j / f \right) x_j : f_j \le f \right) + \sum \left(\left[\left(1 - f_j \right) / \left(1 - f \right) \right] x_j : f_j > f \right) \ge 1$$

From SECs to

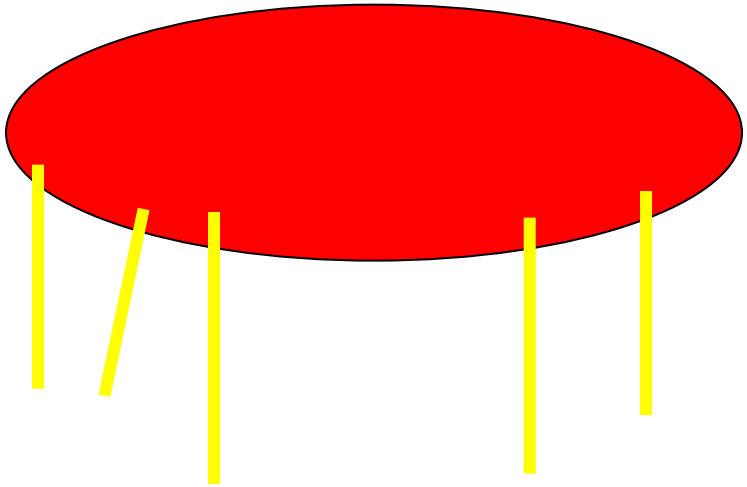
- 2-matching constraints
- combs
- clique tree inequalities
- etc.

Polyhedral Theory of the TSP

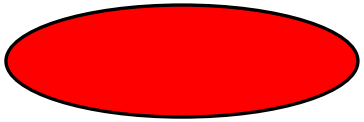
Comb inequality



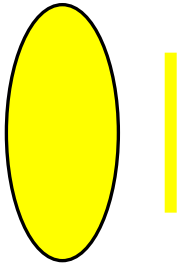
2-matching constraint



handle

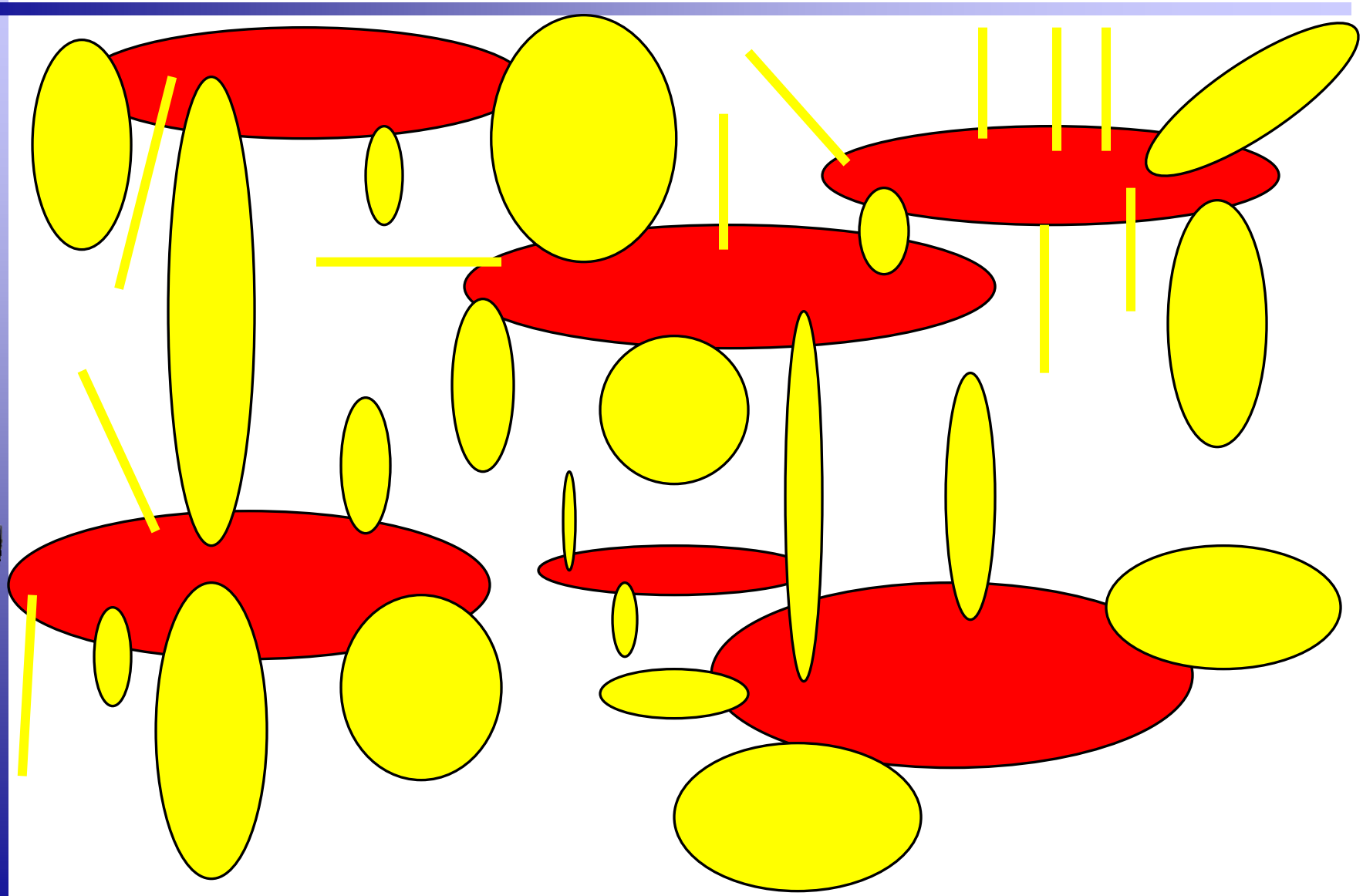


tooth



Martin Grötschel

Clique Tree Inequalities

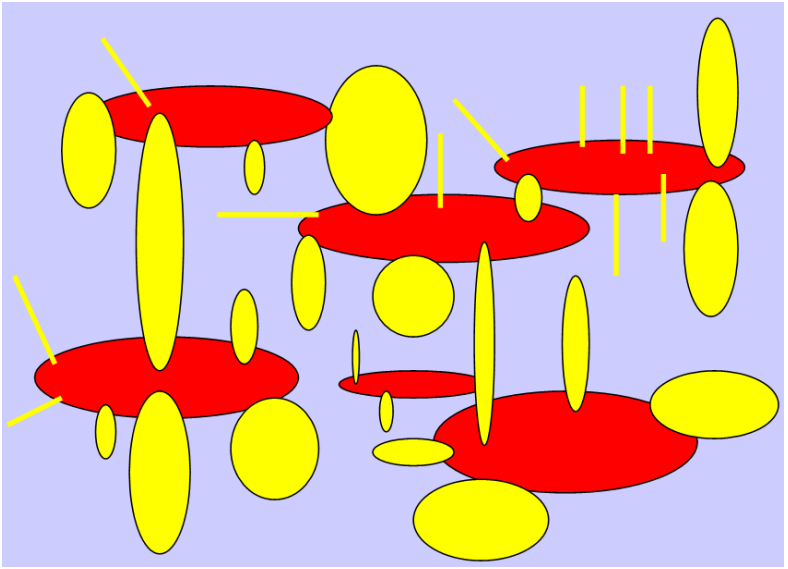


Martin Grötschel

Clique Tree Inequalities

$$\sum_{i=1}^{h} x(\partial(H_i)) + \sum_{j=1}^{t} x(\partial(T_j)) \geq \sum_{i=1}^{h} |H_i| + h + 2t$$

$$\sum_{i=1}^{h} x(E(H_i)) + \sum_{j=1}^{t} x(E(T_j)) \leq \sum_{i=1}^{h} |H_i| + \sum_{i=1}^{t} (|T_j| - t_j) - \frac{t+1}{2}$$

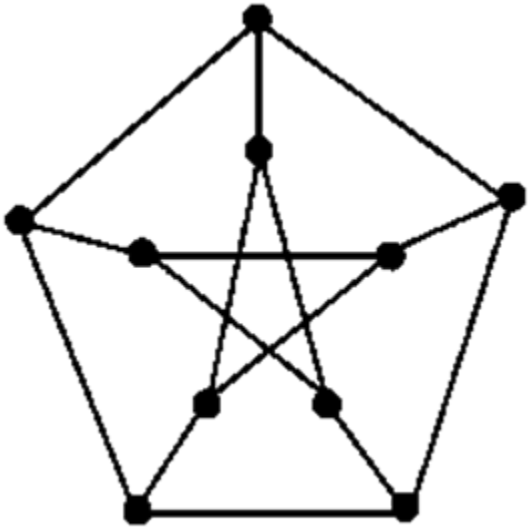


Martin Grötschel

Valid Inequalities for STSP

- Trivial inequalities
- Degree constraints
- Subtour elimination constraints
- 2-matching constraints, comb inequalities
- Clique tree inequalities (comb)
- Bipartition inequalities (clique tree)
- Path inequalities (comb)
- Star inequalities (path)
- Binested Inequalities (star, clique tree)
- Ladder inequalities (2 handles, even # of teeth)
- Domino inequalities
- Hypohamiltonian, hypotraceable inequalities
- etc.

A very special case



Petersen graph, G = (V, F),
the smallest hypohamiltonian graph

$$x(F) \leq 9 \;\; defines \; a \;\; facet \; of \; Q_T^{10}$$

$$but \; not \; a \;\; facet \; of \;\; Q_T^n, n \geq 11$$

M. Grötschel & Y. Wakabayashi

Valid and facet defining inequalities for STSP: Survey articles

- M. Grötschel, M. W. Padberg (1985 a, b)
- M. Jünger, G. Reinelt, G. Rinaldi (1995)
- D. Naddef (2002)

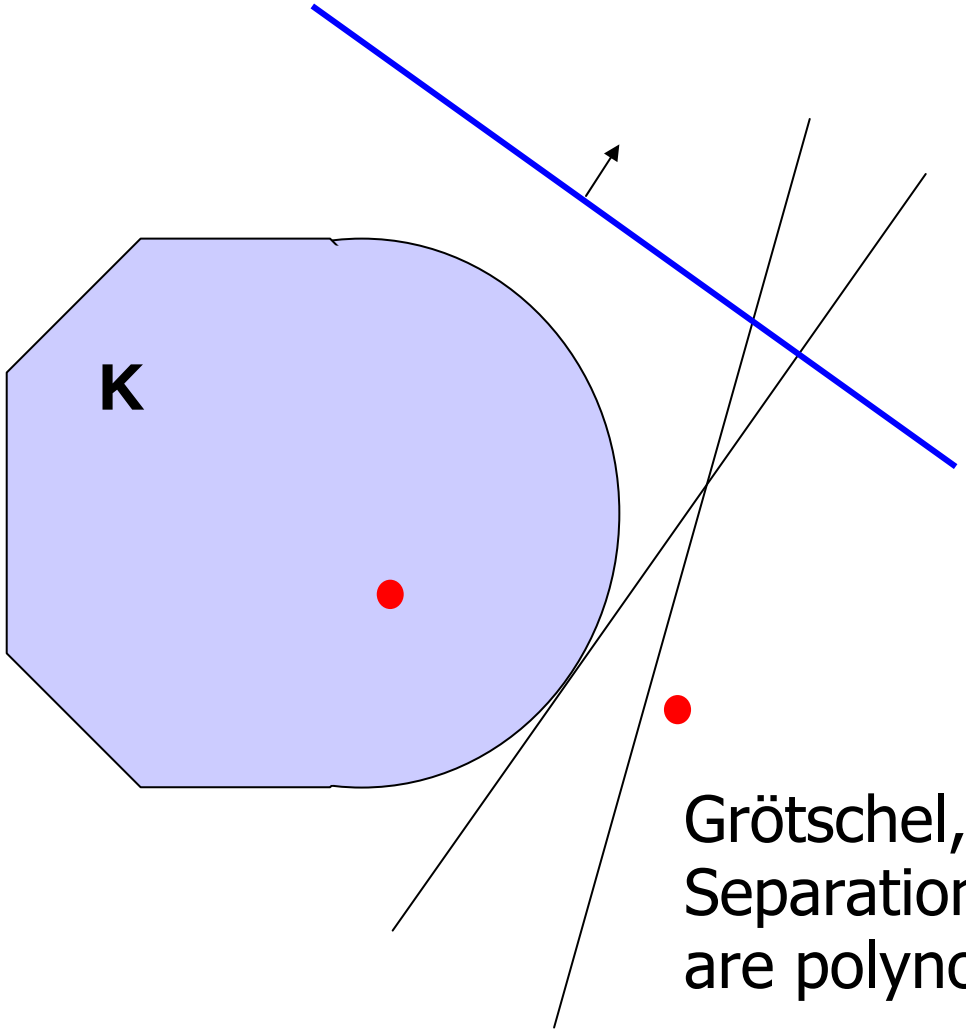
Counting Tours and Facets

n	# tours	# different facets	# facet classes
3	1	0	0
4	3	3	1
5	12	20	2
6	60	100	4
7	360	3,437	6
8	2520	194,187	24
9	20,160	42,104,442	192
10	181,440	>= 52,043.900.866	>=15,379

Separation Algorithms

- Given a system of valid inequalities (possibly of exponential size).
- Is there a polynomial time algorithm (or a good heuristic) that,
  - given a point,
  - checks whether the point satisfies all inequalities of the system, and
  - if not, finds an inequality violated by the given point?

Separation



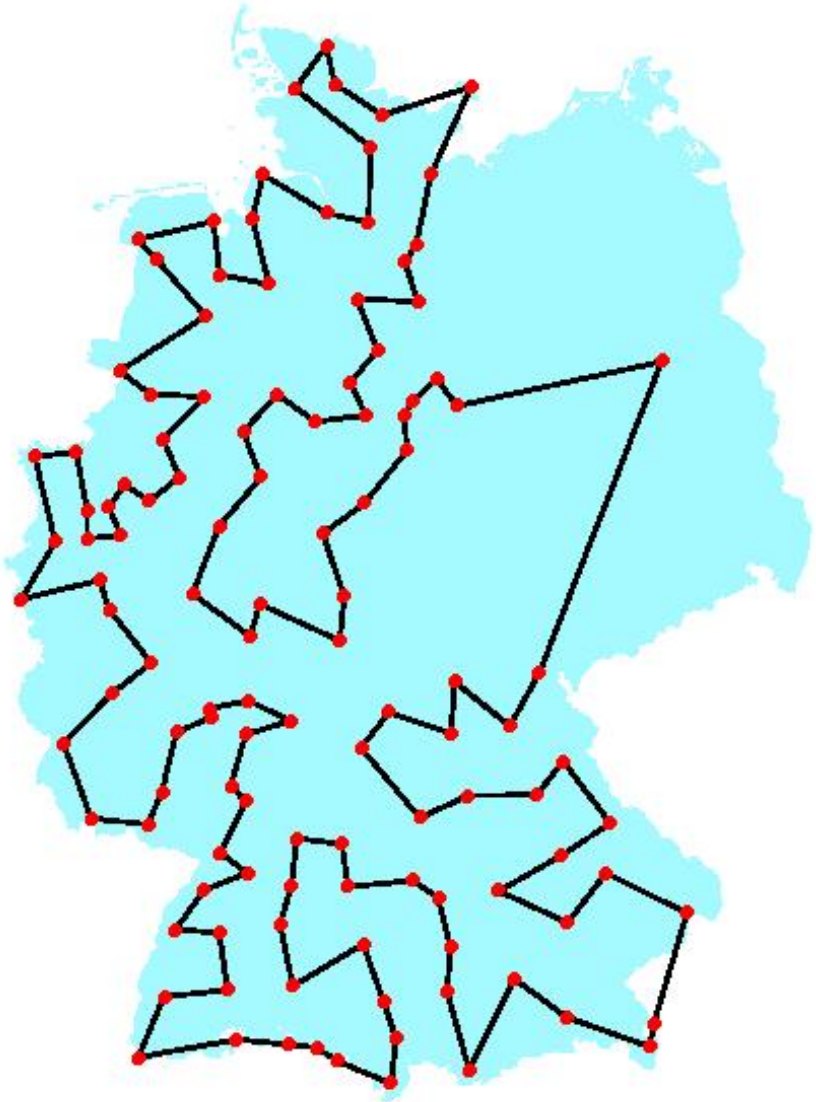
Grötschel, Lovász, Schrijver:
Separation and optimization
are polynomial time equivalent.

Separation Algorithms

- There has been great success in finding exact polynomial time separation algorithms, e.g.,
  - for subtour-elimination constraints
  - for 2-matching constraints (Padberg&Rao, 1982)
- or fast heuristic separation algorithms, e.g.,
  - for comb constraints
  - for clique tree inequalities
- and these algorithms are practically efficient

SEC Separation

West-Deutschland und Berlin



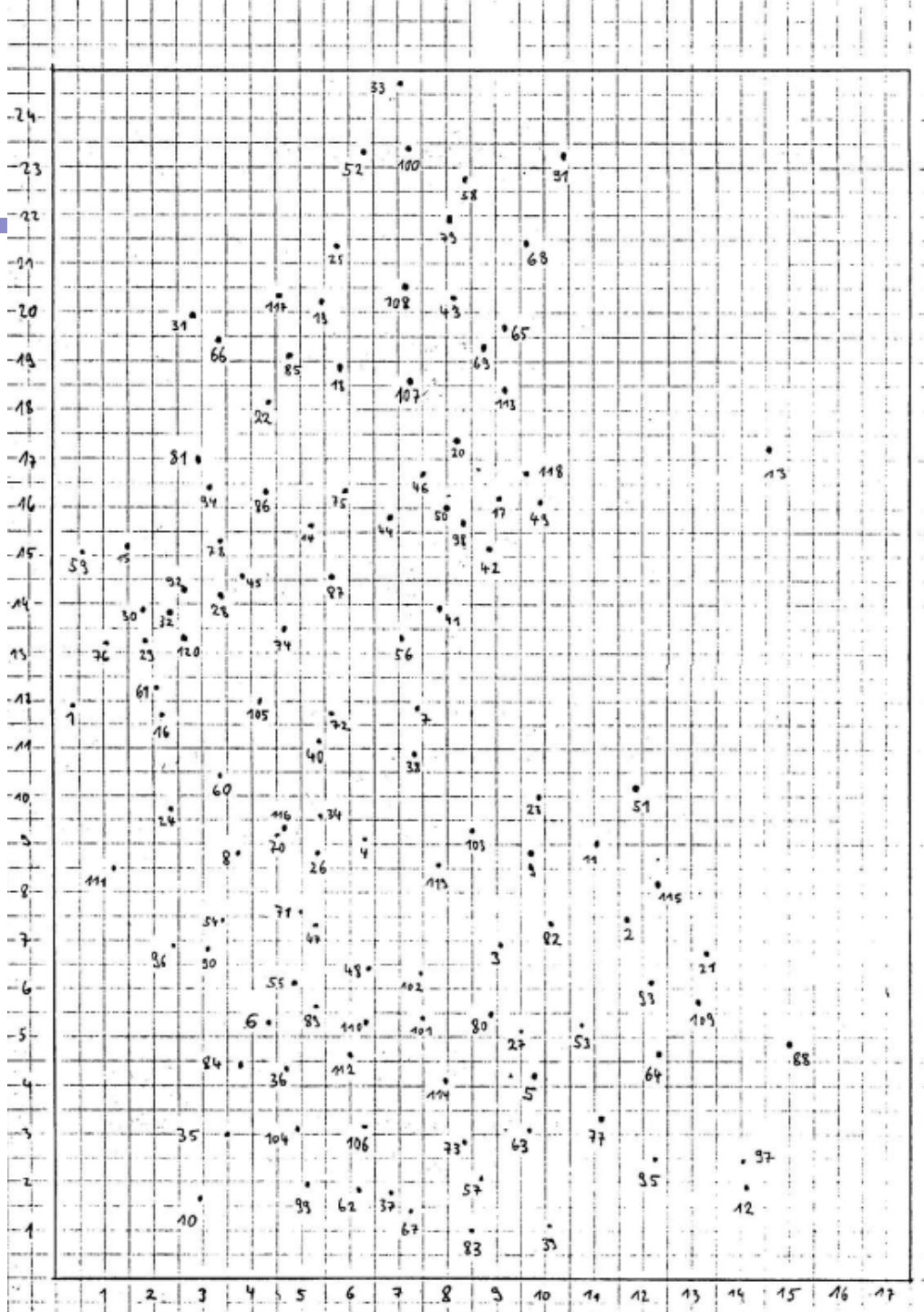
120 Städte
7140 Variable

1975/1977/1980

M. Grötschel

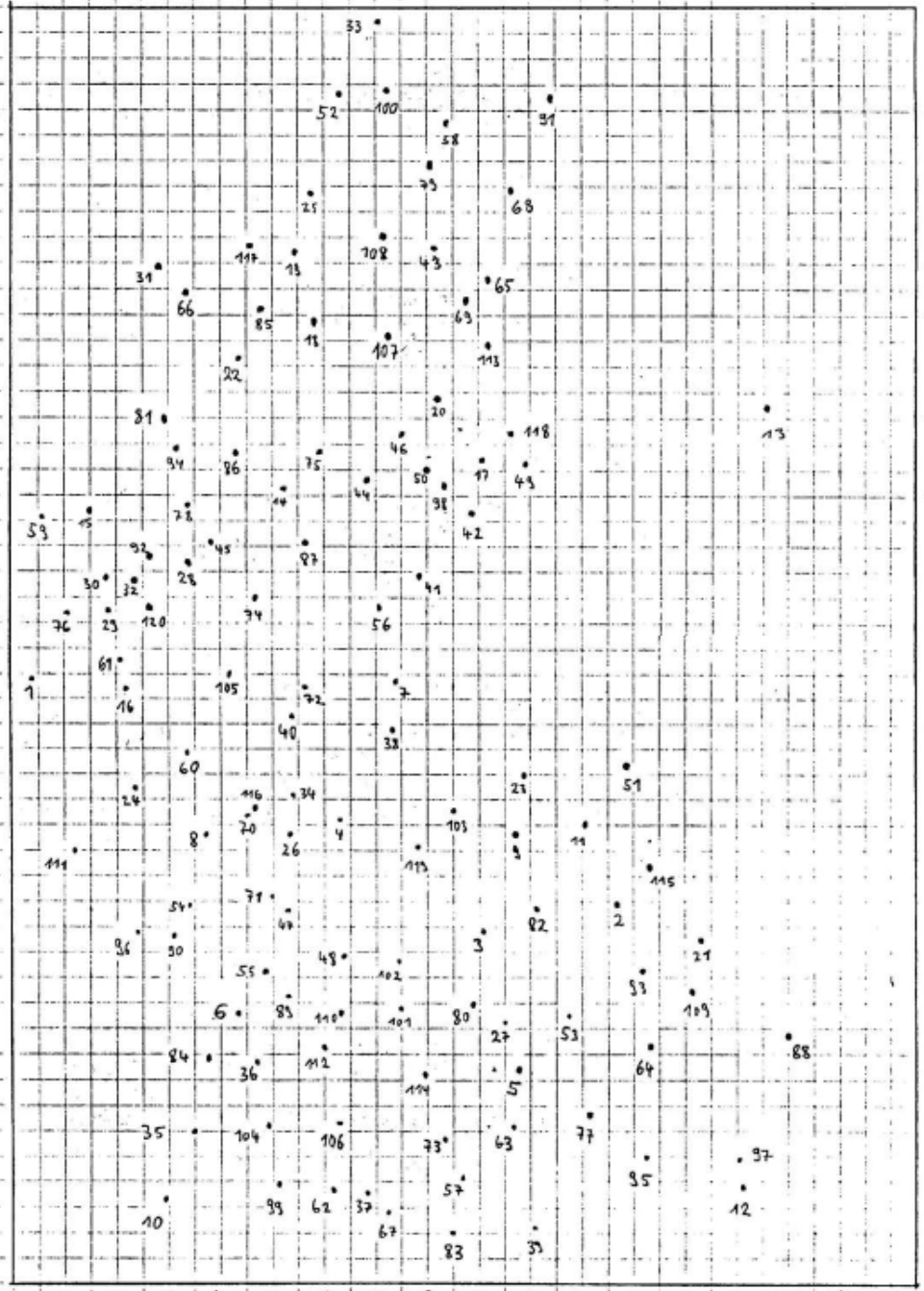
In the old days: 1975, TSP 120

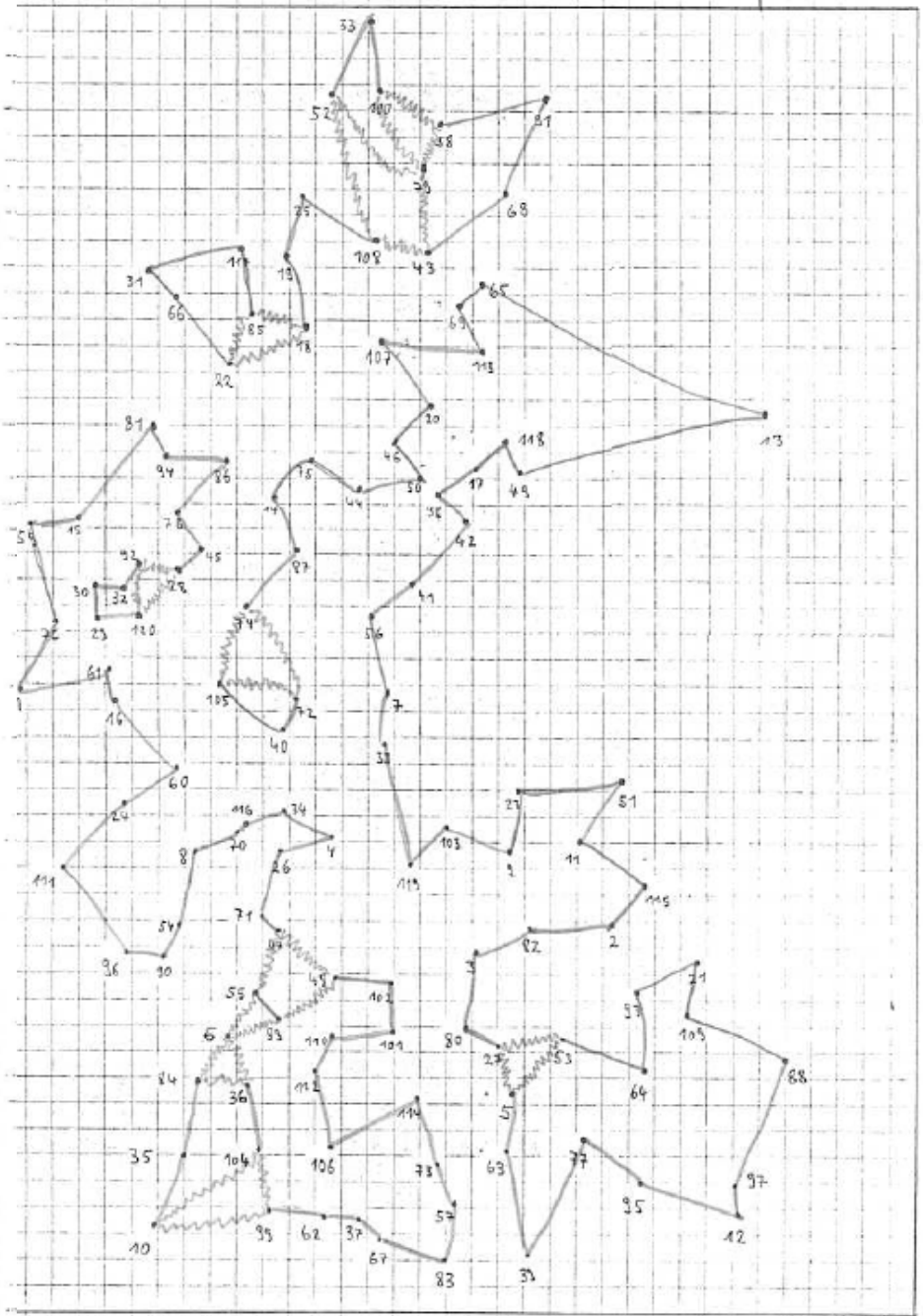
- my drawing of Germany



Martin Grötschel

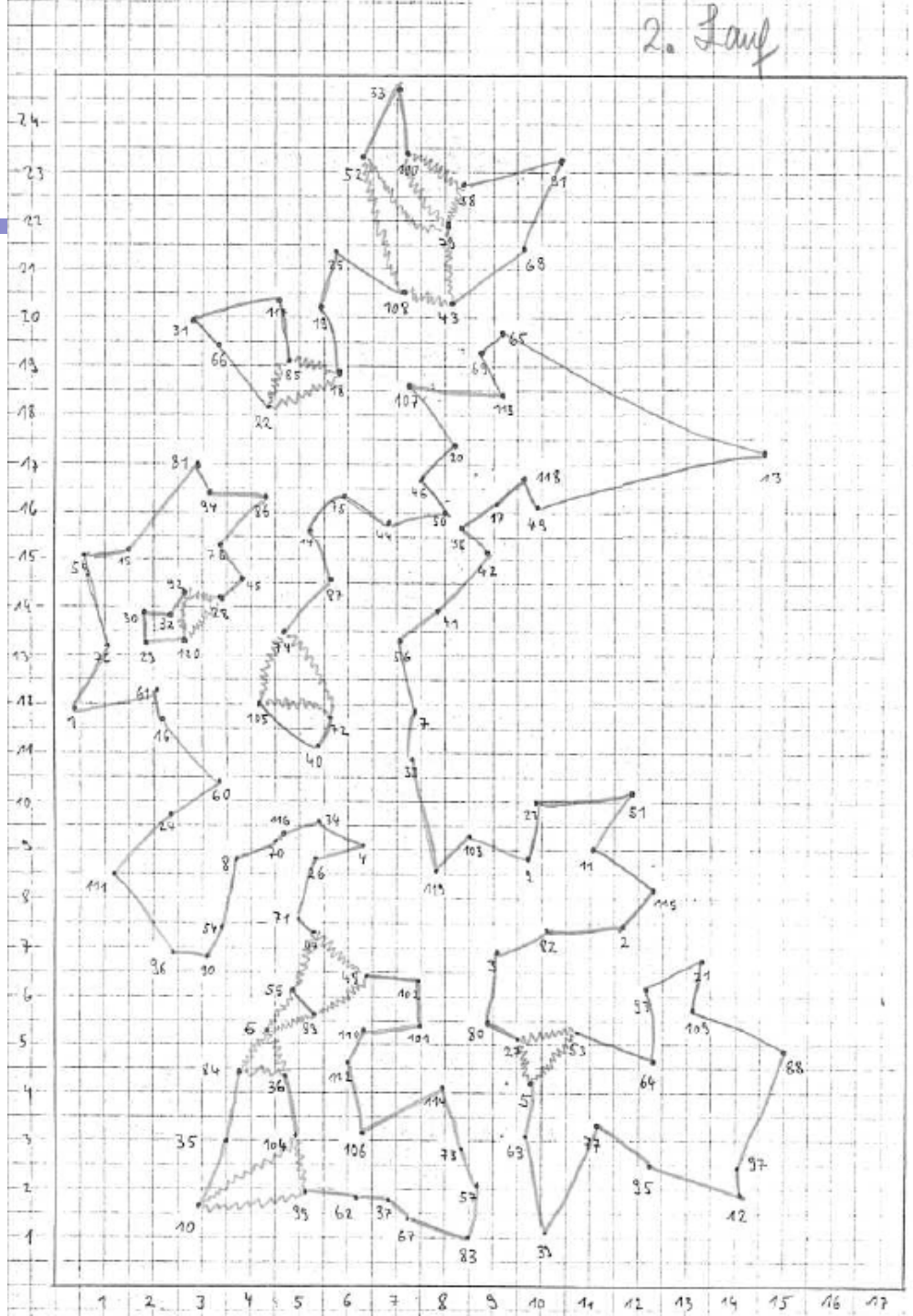
2. Lauf





In the old days: 1975 TSP 120

- optimal LP solution after second run

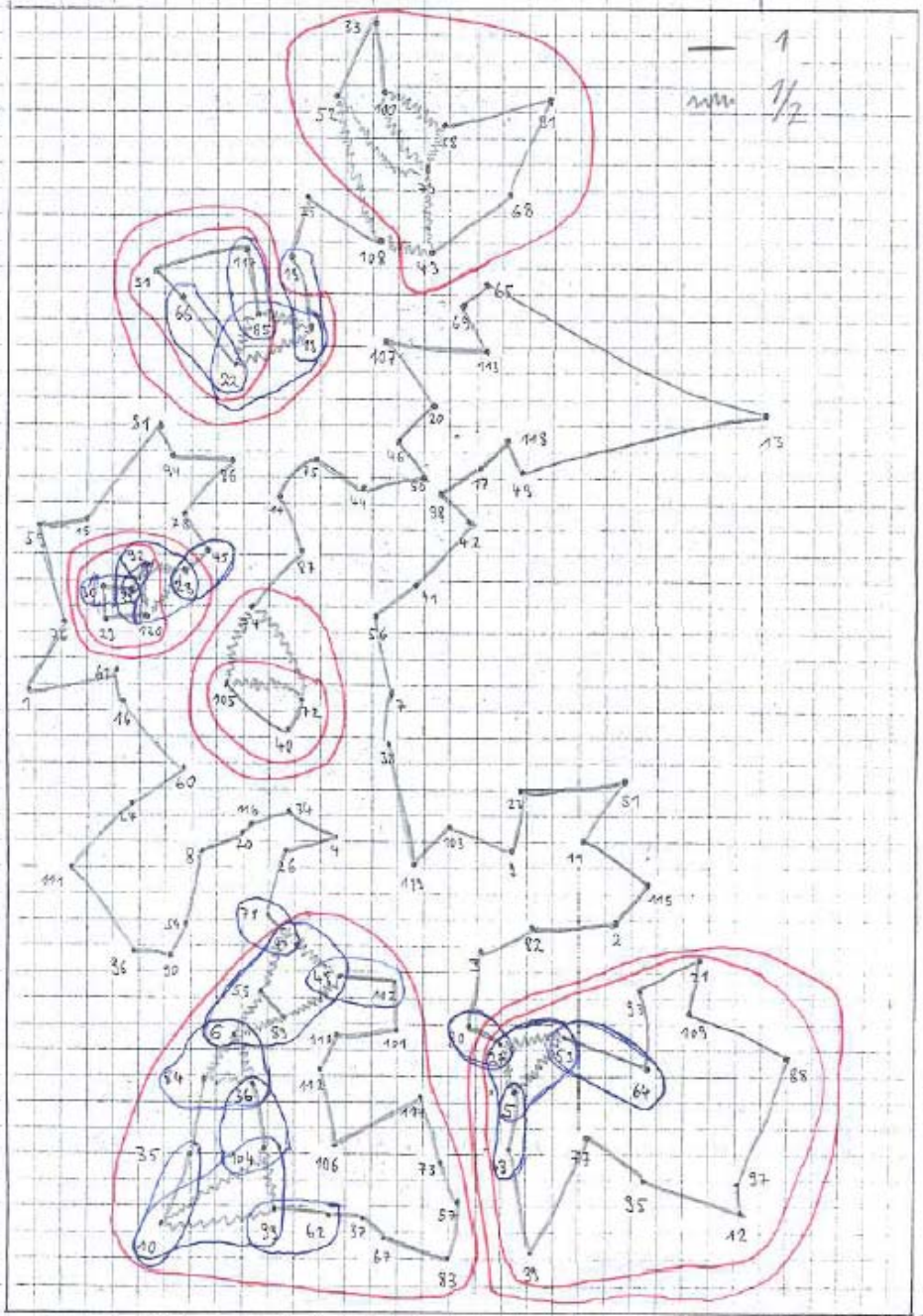


Martin Grötschel

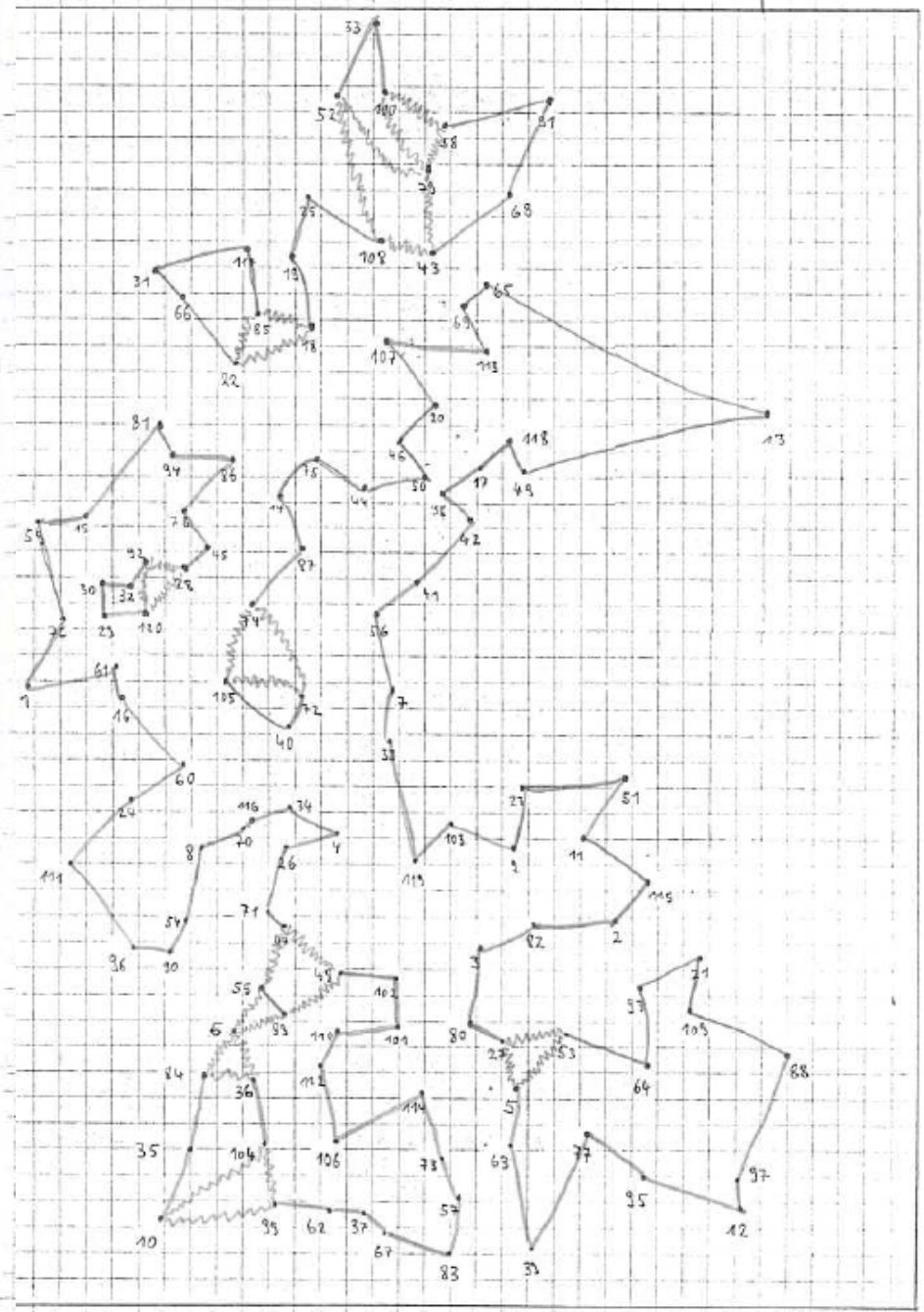
2. Lauf

— 1

≈ 1/2



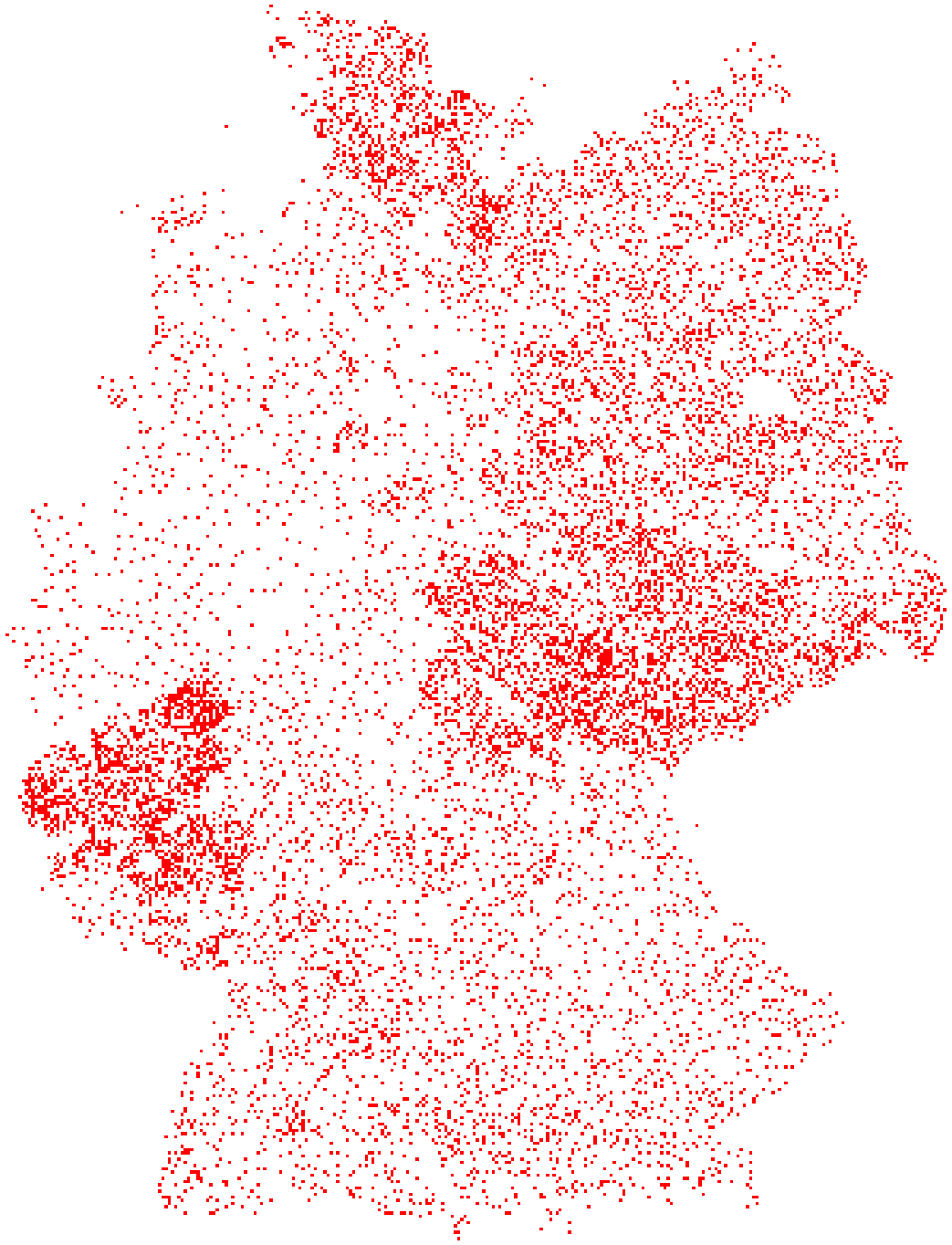
2. Lauf



Polyhedral Combinatorics

This line of research has resulted in powerful cutting plane algorithms for combinatorial optimization problems.

They are used in practice to solve exactly or approximately (including branch & bound) large-scale real-world instances.



Deutschland 15,112

D. Applegate, R.Bixby,
V. Chvatal, W. Cook

15,112

cities

114,178,716

variables

2001

How do we solve a TSP like this?

- **Upper bound:**

  Heuristic search

  - Chained Lin-Kernighan

- **Lower bound:**
  - Linear programming
  - Divide-and-conquer
  - Polyhedral combinatorics
  - Parallel computation
  - Algorithms & data structures

The **LOWER BOUND** is the mathematically and algorithmically hard part of the work

Work on LP relaxations of the symmetric travelling salesman polytope

$$Q_T^n := conv\{\chi^T \in \mathbf{Z}^E \mid T \text{ tour in } K_n\}$$

$$\min c^T x$$

$$x(\delta(i)) = 2 \qquad \forall i \in V$$

$$x(E(W)) \leq |W| - 1 \qquad \forall W \subset V \setminus \{1\}, 3 \leq |W| \leq n-3$$

$$0 \leq x_{ij} \leq 1 \qquad \forall ij \in E$$

~~$x_{ij} \in \{0,1\} \qquad \forall ij \in E$~~

- Integer Programming Approach

cutting plane technique for integer and mixed-integer programming

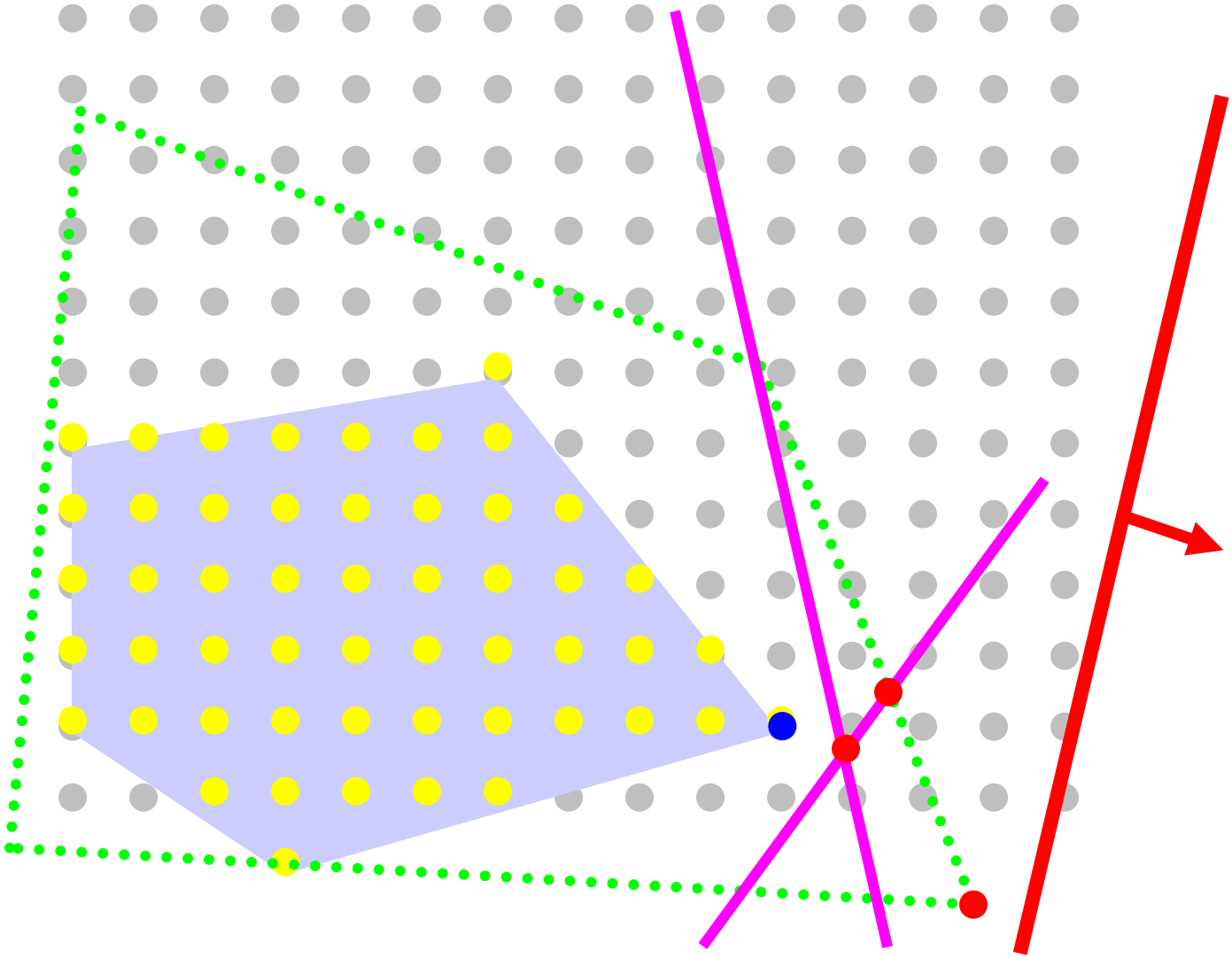
Feasible integer solutions

Objective function

Convex hull

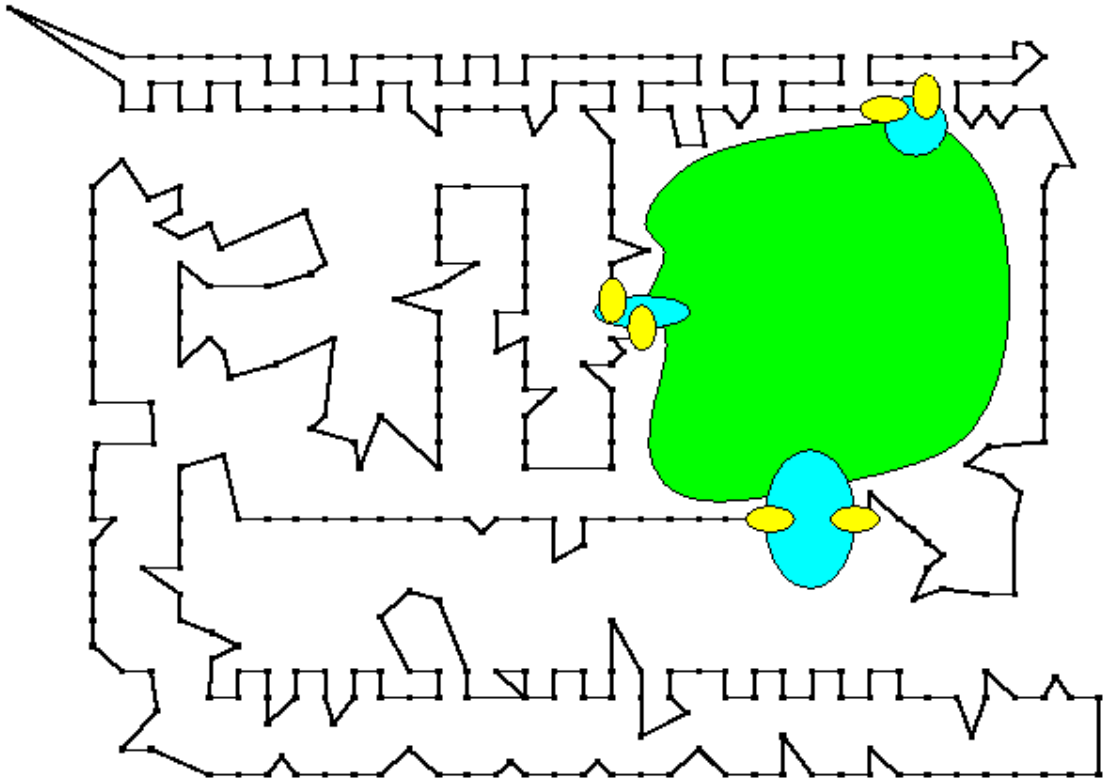
LP-based relaxation

Cutting planes



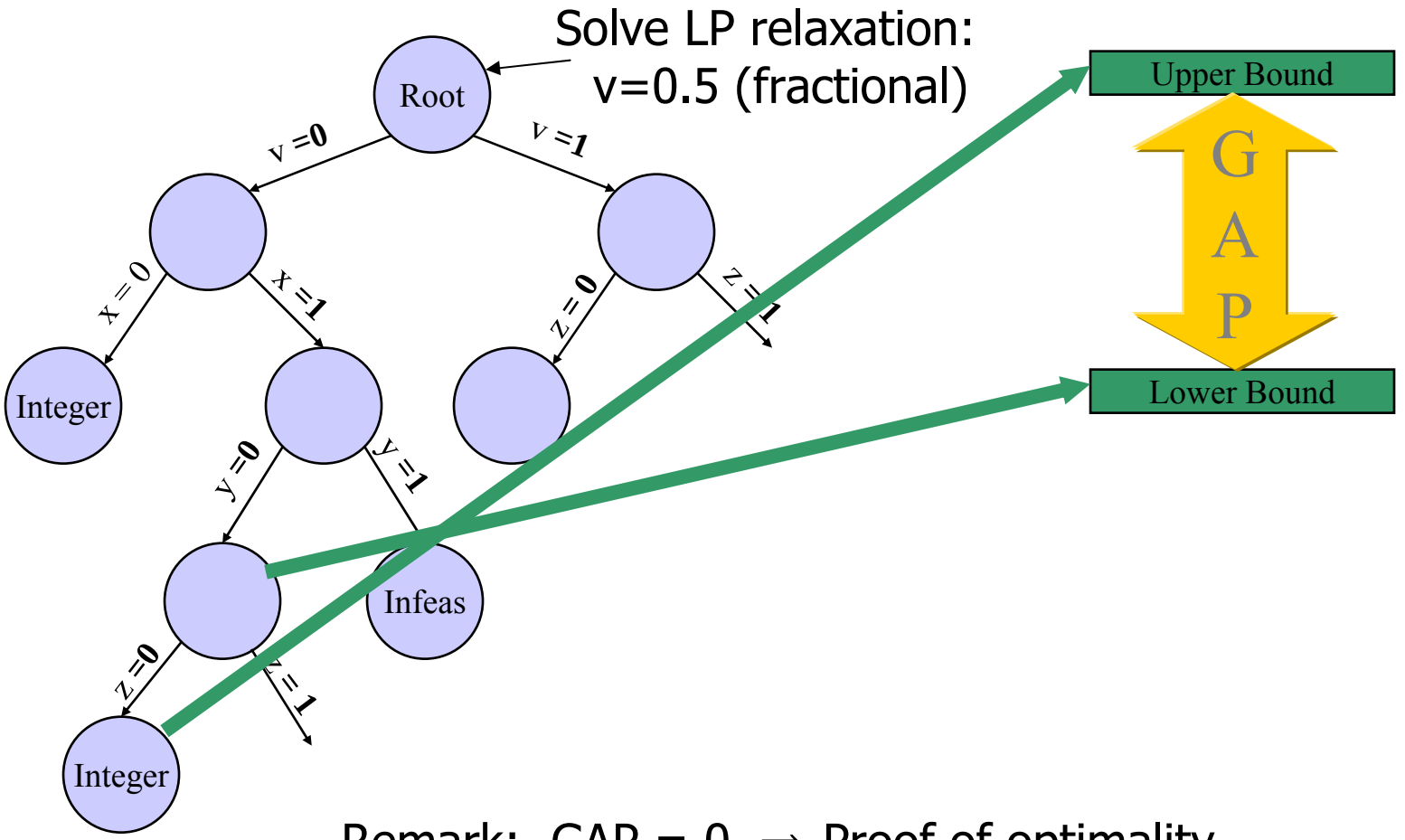
Martin Grötschel

Clique-tree cut for pcb442



Martin Grötschel

LP-based Branch & Bound

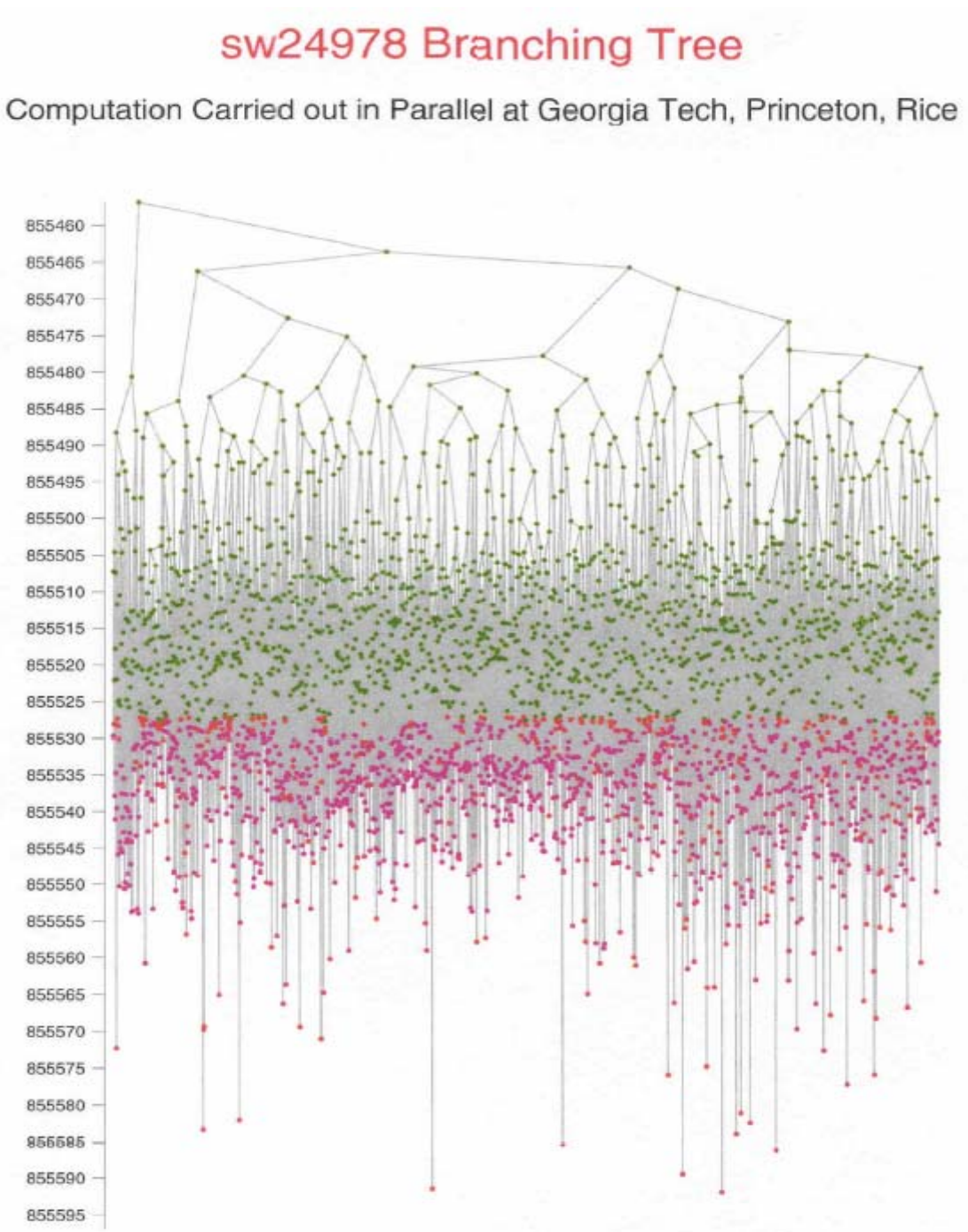


Remark: GAP = 0 $\Rightarrow$ Proof of optimality

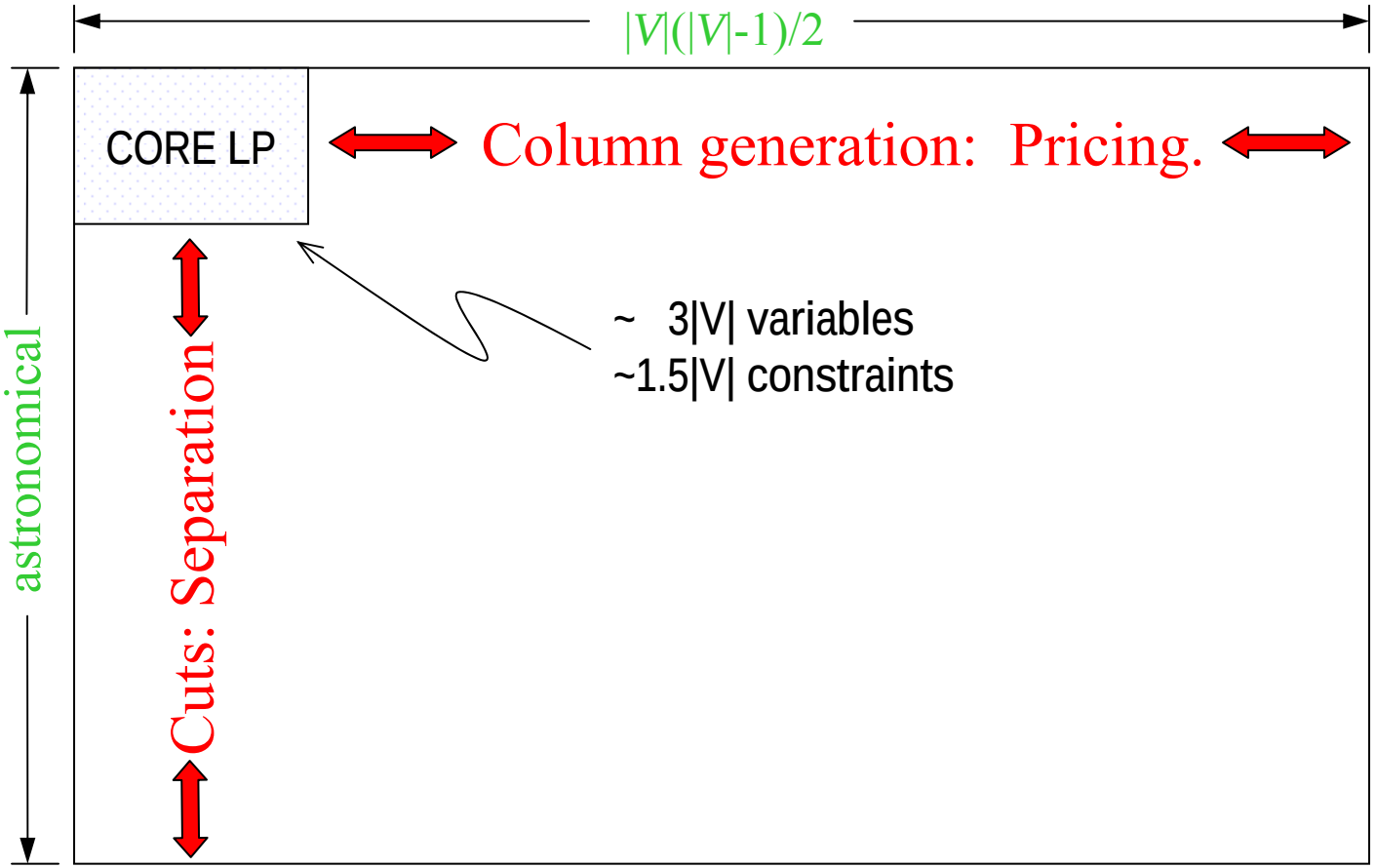
Martin Grötschel

A Branching Tree

- Applegate
- Bixby
- Chvátal
- Cook



Managing the LPs of the TSP



Martin Grötschel

A Pictorial History of Some TSP World Records

Some TSP World Records

2006
pla 85,900
solved

3,646,412,050
variables

number of cities
2000x
increase

4,000,000
times
problem size
increase

in 52
years

year	authors	# cities	# variables
1954	DFJ	42/49	820/1146
1977	G	120	7140
1987	PR	532	141,246
1988	GH	666	221,445
1991	PR	2,392	2,859,636
1992	ABCC	3,038	4,613,203
1994	ABCC	7,397	27,354,106
1998	ABCC	13,509	91,239,786
2001	ABCC	15,112	114,178,716
2004	ABCC	24,978	311,937,753
2005 W. Cook, D. Epsinoza, M. Goycoolea		33,810	571,541,145

The current champions

http://www.tsp.gatech.edu/concorde/index.html

ABCC stands for

D. Applegate, B. Bixby, W. Cook, V. Chváta (plus students)

- almost 15 years of code development
- have made their code CONCORDE available in the Internet

Concorde TSP Solver

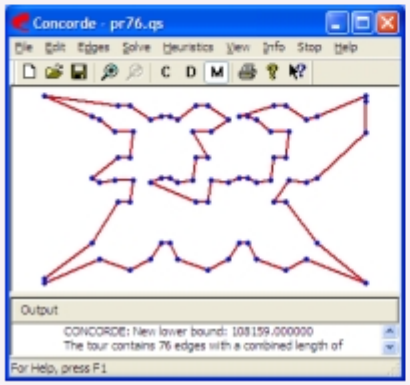
Concorde is a computer code for the symmetric traveling salesman problem (TSP) and some related network optimization problems. The code is written in the ANSI C programming language and it is available for academic research use; for other uses, contact William Cook for licensing options.

Concorde's TSP solver has been used to obtain the optimal solutions to 107 of the 110 TSPLIB instances; the largest having 15,112 cities.

The Concorde callable library includes over 700 functions permitting users to create specialized codes for TSP-like problems. All Concorde functions are thread-safe for programming in shared-memory parallel environments; the main TSP solver includes code for running over networks of UNIX workstations.

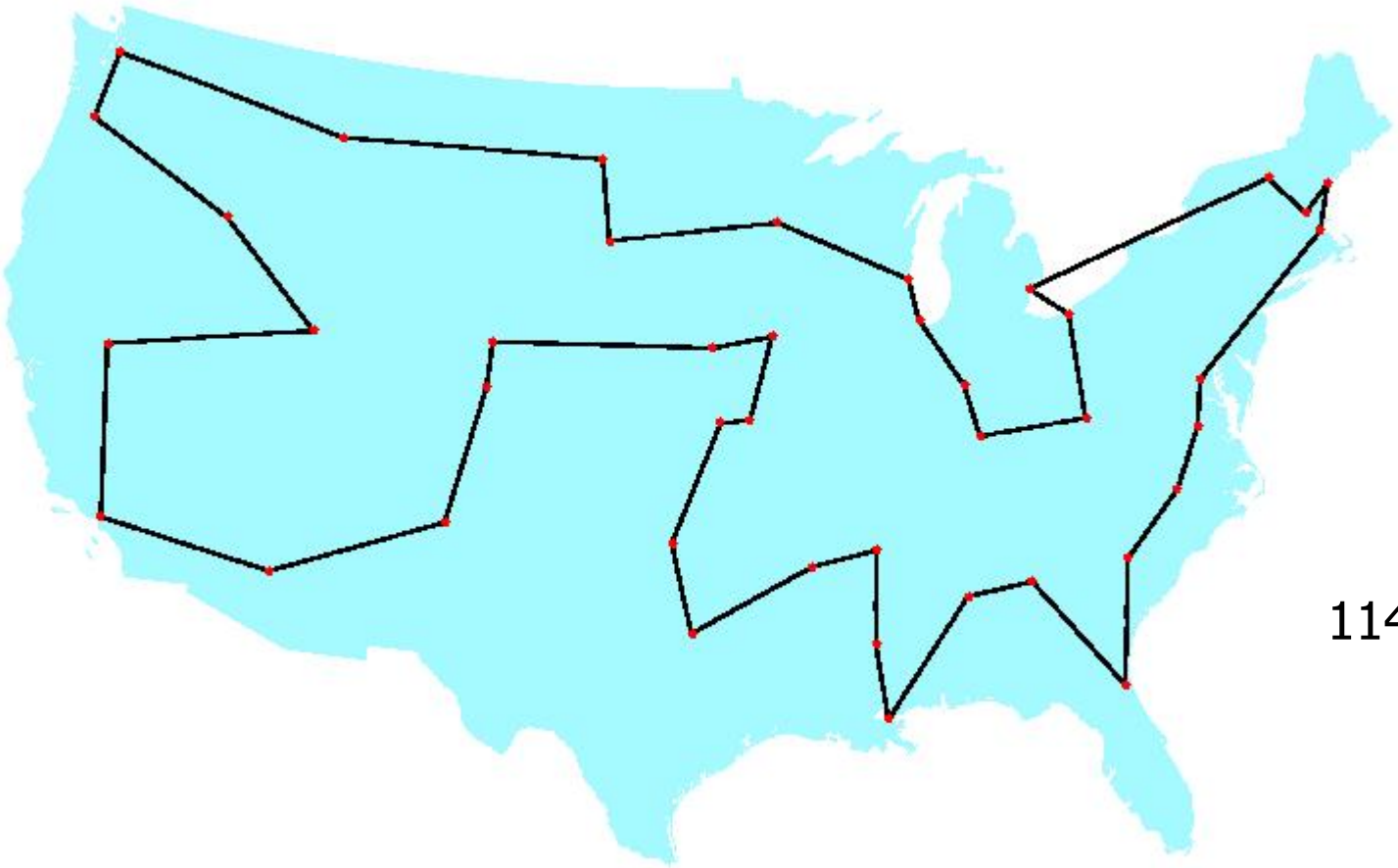
Concorde now supports the QSopt linear programming solver. Executable versions of concorde with qsopt for Linux and Solaris are now available

Hans Mittelmann has created a NEOS Server for Concorde, allowing users to solve TSP instances online.



- Concorde for Windows
- Benchmark Results
- User Documentation for Source Code
- Downloading Concorde

USA 49



49 cities
1146 variables

1954

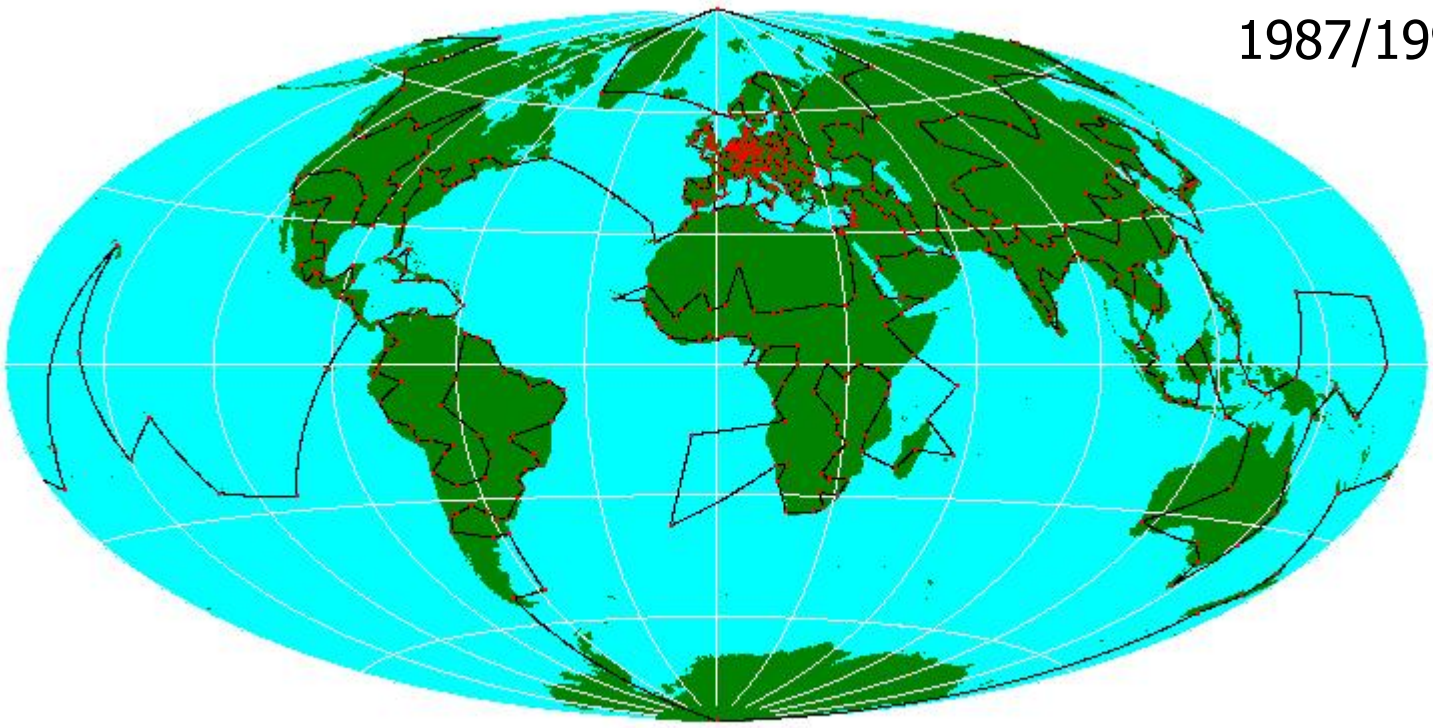
G. Dantzig, D.R. Fulkerson, S. Johnson

The journey around the world

666 cities
221.445 variables

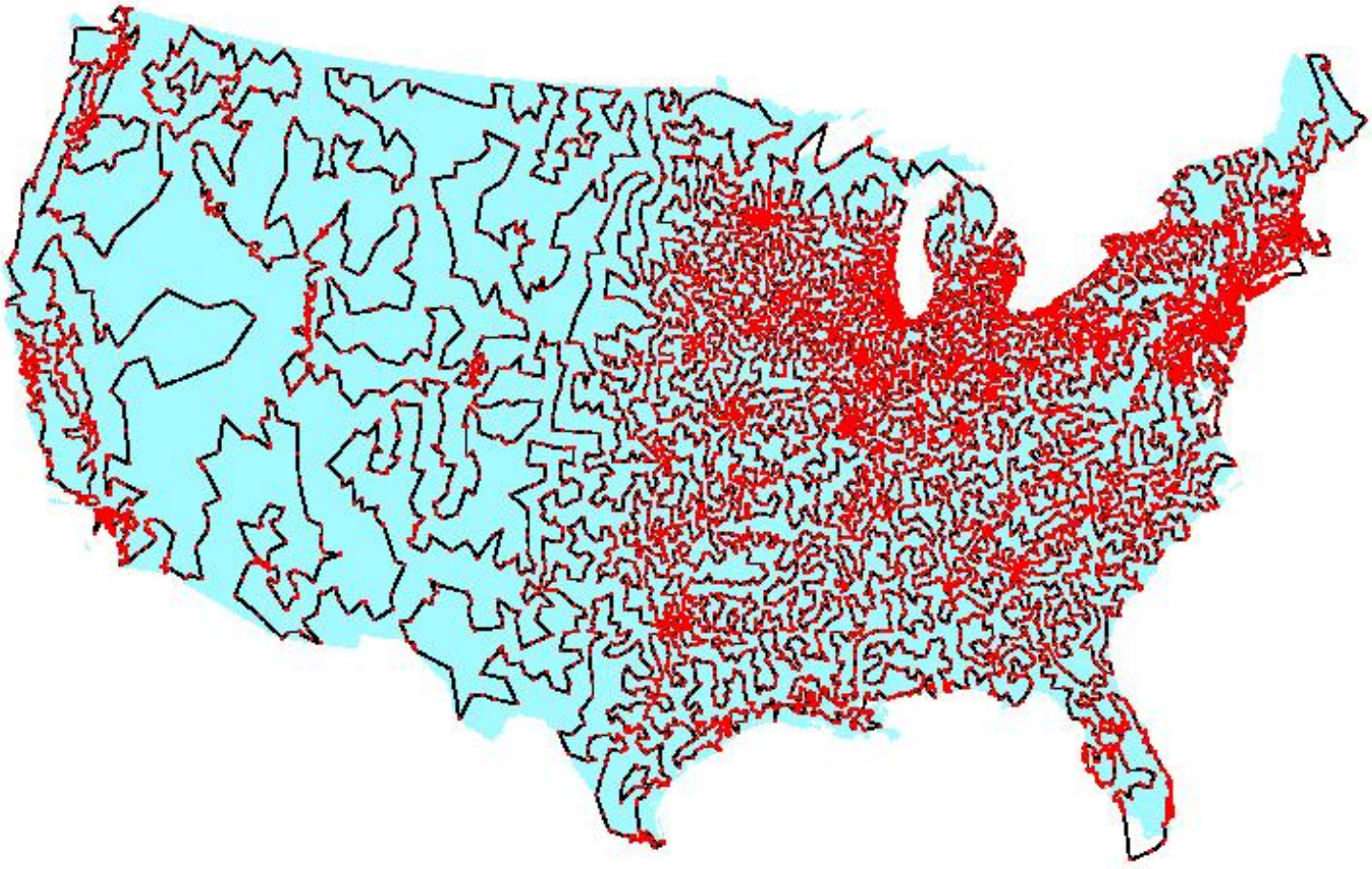
city list

1987/1991



M. Grötschel, O. Holland

USA cities with population >500

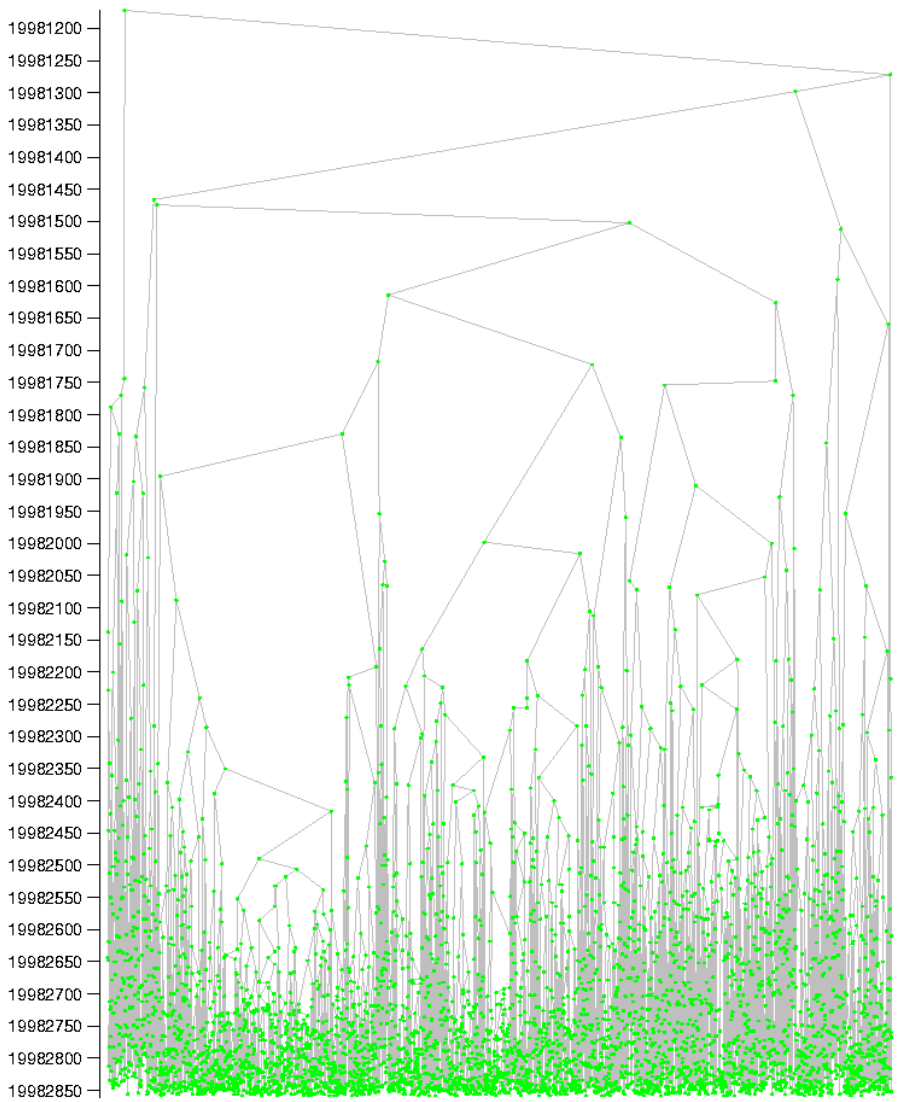


13,509
cities
91,239,786
Variables

1998

D. Applegate, R.Bixby, V. Chvátal, W. Cook

usa13509: The branching tree

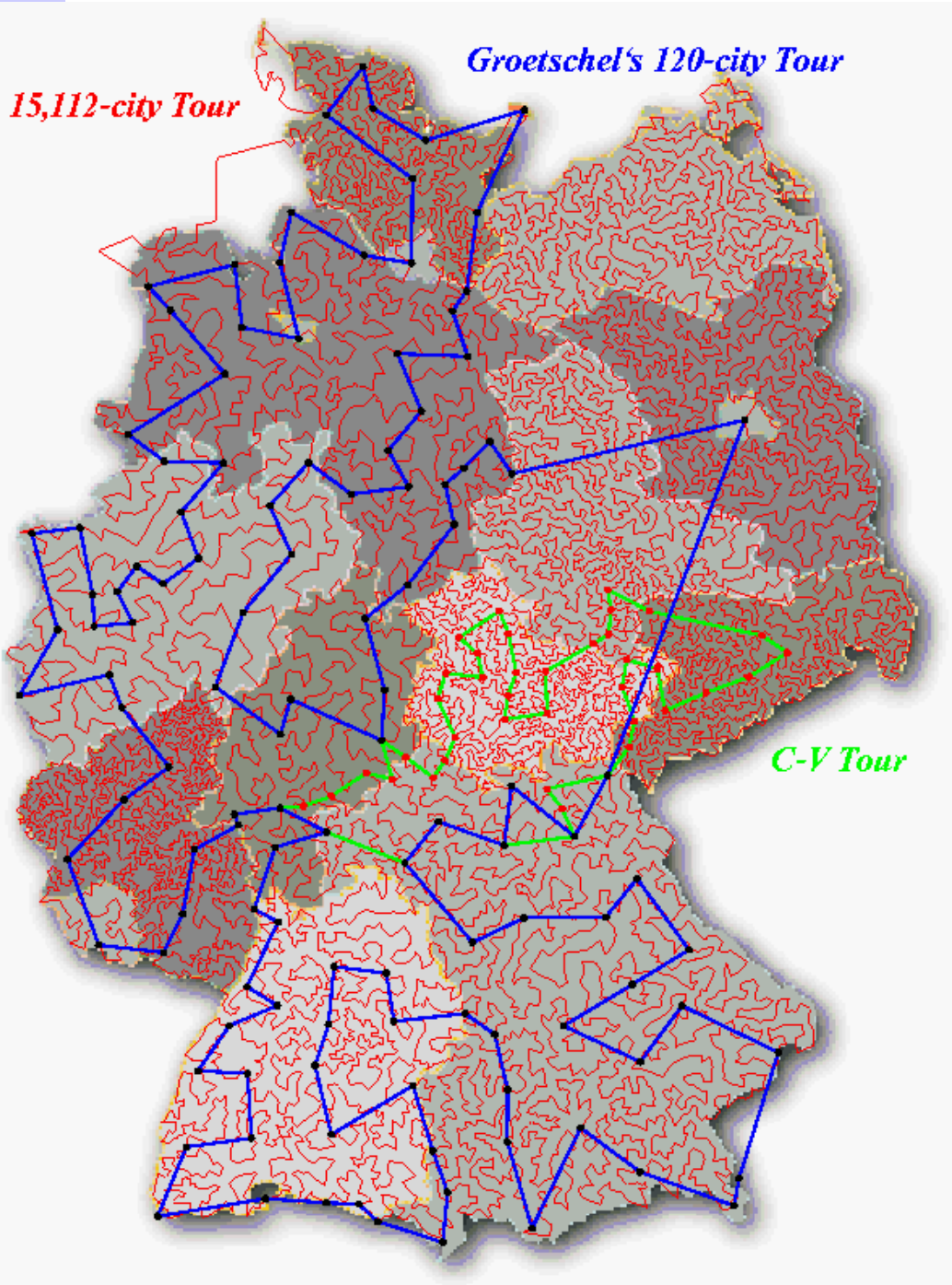


0.01%
initial gap

Martin
Grötschel

Summary: usa13509

- 9539 nodes branching tree
- 48 workstations (Digital Alphas, Intel Pentium IIs, Pentium Pros, Sun UntraSparcs)
- Total CPU time: **4 cpu years**

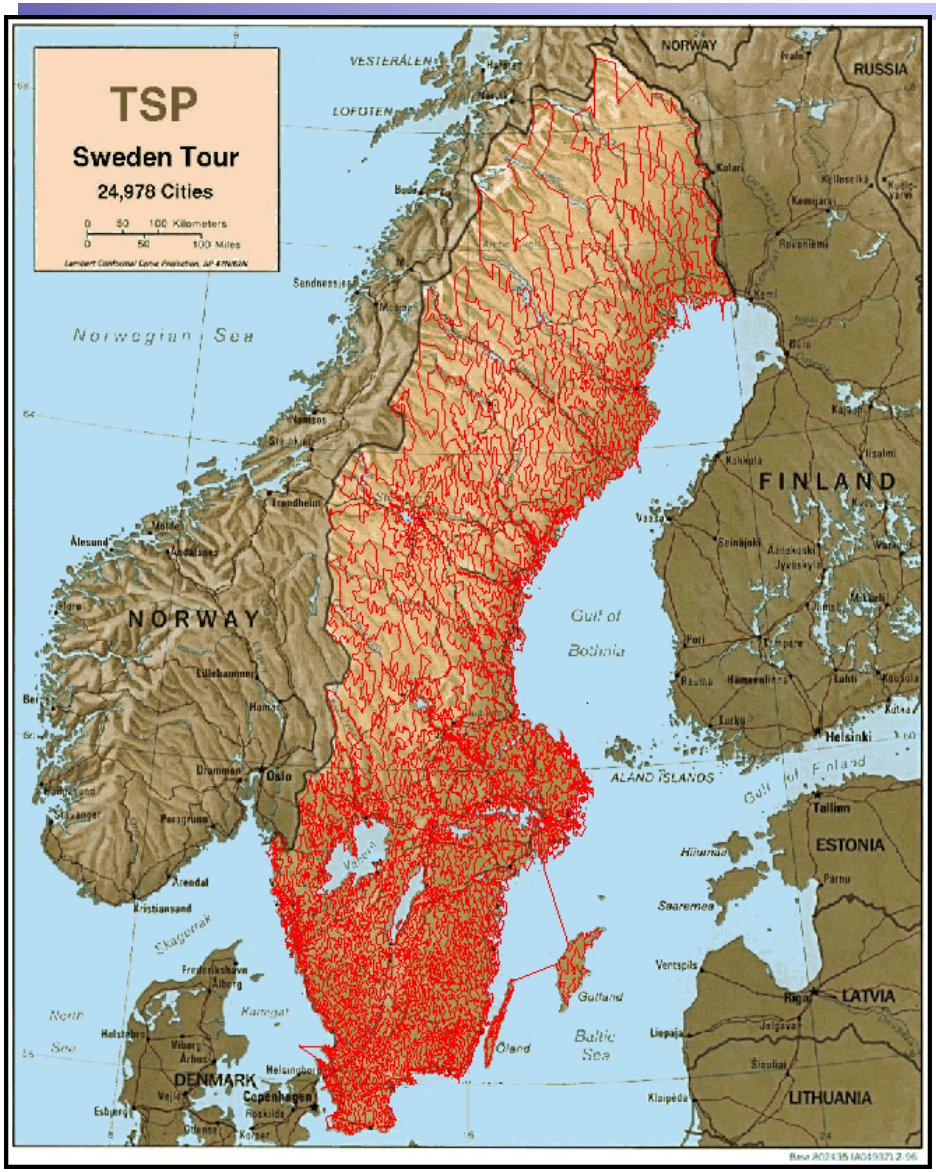


Overlay of 3 Optimal Germany tours

from
ABCC 2001

http://www.math.princeton.edu/tsp/d15sol/dhistory.html

Optimal Tour of Sweden

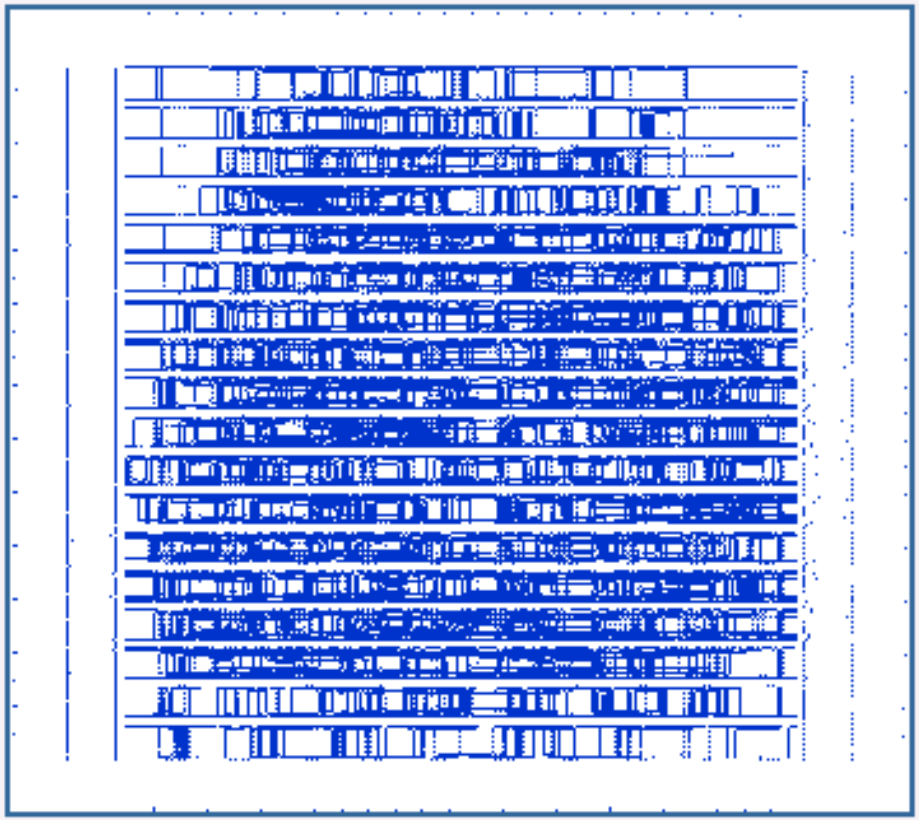


311,937,753 variables

ABCC
plus
Keld Helsgaun
Roskilde Univ.
Denmark.

Martin Grötschel

The current world record



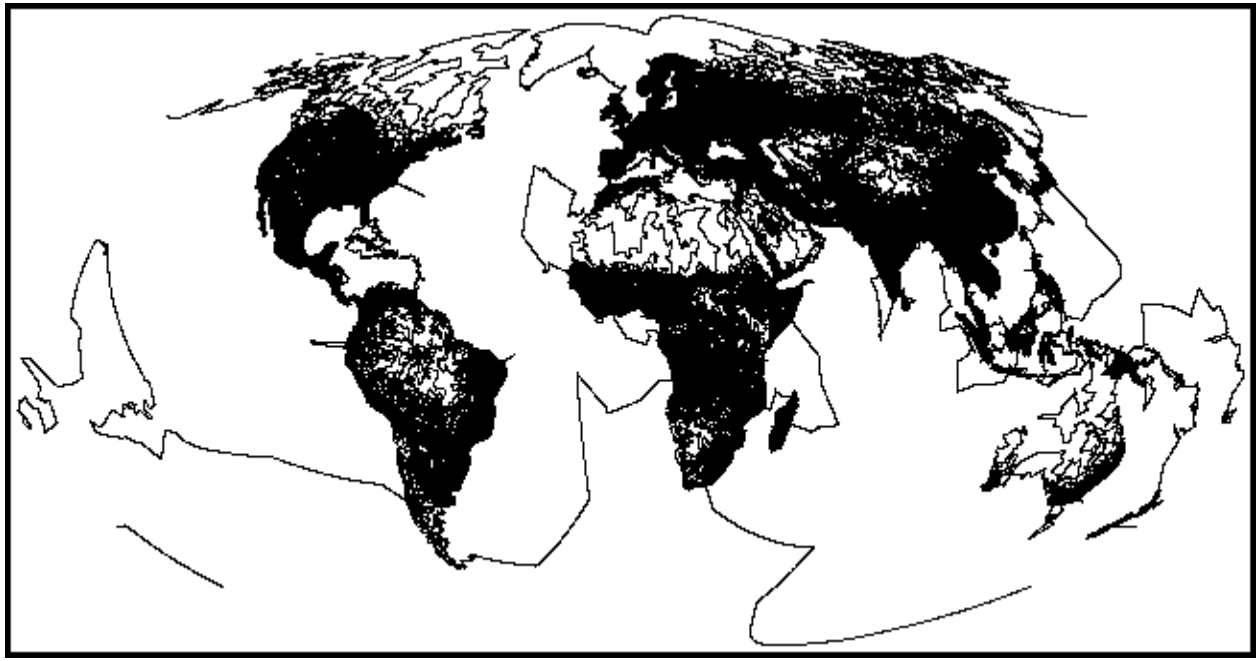
85,900 Locations in a VLSI Application
Solved in 2006

http://www.tsp.gatech.edu/optimal/index.html

http://www.tsp.gatech.edu/pla85900/index.html

World Tour, current status

http://www.tsp.gatech.edu/world/



We give links to several images of the World TSP tour of length **7,516,353,779** found by Keld Helsgaun in December 2003. A lower bound provided by the Concorde TSP code shows that this tour is at most **0.076%** longer than an optimal tour through the 1,904,711 cities.

The Travelling Salesman Problem and its Applications

CO@W Berlin

The End

Martin Grötschel
- Institut für Mathematik, Technische Universität Berlin (TUB)
- DFG-Forschungszentrum "Mathematik für Schlüsseltechnologien" (MATHEON)
- Konrad-Zuse-Zentrum für Informationstechnik Berlin (ZIB)

groetschel@zib.de http://www.zib.de/groetschel