



Matteo Fischetti, DEI, University of Padova

COFIN – Matheon workshop, Villa Vigoni (Como), May 2006

MIP solvers for hard optimization problems

- **Mixed-integer linear programming** (MIP) plays a central role in modelling difficult-to-solve (NP-hard) combinatorial problems
- General-purpose (exact) MIP solvers are very sophisticated tools, but in some hard cases they are **not adequate** even after clever tuning
- One is therefore tempted **to quit the MIP framework** and to design ad-hoc heuristics for the specific problem at hand, thus loosing the advantage of working in a generic MIP framework
- As a matter of fact, too often a MIP model is developed only "to better describe the problem" or, in the best case, to compute bounds for **benchmarking** the proposed ad-hoc heuristics

Can we devise an alternative use of a general-purpose MIP solver, e.g., to address important steps in the solution process?

I MIP you

A neologism: To *MIP something* = translate into a MIP model and solve through a black-box solver



MIP-heuristic enslaved to an exact MIP solver

• **MIPping Ralph**: use a black-box (general-purpose) MIP heuristic for the separation of Chvàtal-Gomory cuts, so as to enhance the convergence of an exact MIP solver

(M. F., A. Lodi, "Optimizing over the first Chvàtal closure", IPCO'05, 2005)



MIPped !!!

 $P := \{ x \ge 0 : A x \le b \}$ $a^{T} x \le x_{0} + 0.999$ $valid for P, \qquad \Rightarrow valid for valid for P, \qquad valid for Valid for P, \qquad Yalid for Valid for Vali$ JUST MIP IT! max x + - xo at s ut A a.+0.999 ≥ utb (x, xo) intege

MIP-solver enslaved to a local-search metaheuristic

MIPping Fred: use a black-box (general-purpose) MIP solver to

- explore large solution neighbourhoods defined through invalid linear inequalities called local branching cuts;
- diversification is also modelled through MIP cuts

(M.F., A. Lodi, "Local Branching", Mathematical Programming B, 98, 23-47, 2003)



MIPped !!!

Given a feasible 0-1 solution x_H , define a MIP neighbourhood though the **local branching** constraint

$$\Delta(x, x^{H}) := \sum_{j \in B: x_{j}^{H} = 0} x_{j} + \sum_{j \in B: x_{j}^{H} = 1} (1 - x_{j}) \le k$$

MIPping critical sub-tasks in the design of specific algorithms

We teach engineers to use MIP models for solving **their** difficult problems (telecom, network design, scheduling, etc.)





Be smart as an engineer!

Model the most critical steps in the design of **your own** algorithm through MIP models, and solve them (even heuristically) through a general-purpose MIP solver...

A new heuristic algorithm for the *Vehicle Routing Problem*



Roberto De Franceschi, DEI, University of Padua

Matteo Fischetti, DEI, University of Padua

Paolo Toth, DEIS, University of Bologna

1

A method for the TSP (Sarvanov and Doroshko, 1981)



Capacitated Vehicle Routing Problem





K routes

not exceeding the given capacity

with minimum total cost

The SERR algorithm

	Steps
Initialization	generate, by any heuristic or metaheuristic, an initial solution
	Iteratively:
Selection	select the nodes to be extracted, according to suitable criteria (schemes)
Extraction	remove the selected nodes and generate a restricted solution (edges = potential insertion points)
Recombination	starting from extracted nodes, generate a (possibly large) number of potential node sequences
Re-insertion	re-insert a subset of the potential sequences into the restricted solution, in such a way that all the extracted nodes are covered again
Evaluation	verify a stopping condition and return, if it is the case, to the selection step

An example



An example



SERR Algorithm

Node re-insertion

Node re-insertion is done by solving the following *set-partitioning* model:

$$\begin{split} \min \sum_{s \in S} \sum_{i \in I} C_{si} x_{si} \\ \sum_{s \ni v} \sum_{i \in I} x_{si} &= 1 \quad \forall v \text{ extracted node} \\ \sum_{s \in S} x_{si} &\leq 1 \quad \forall i \in I \\ d(r) + \sum_{s \in S} \sum_{i \in r} d(s) x_{si} &\leq C \quad \forall r \in R \\ 0 &\leq x_{sj} \leq 1 \quad \text{integer} \quad \forall s \in S, \forall i \in I \end{split}$$

 $x_{si} = 1$ if and only if sequence *s* goes into the insertion point *i* C_{si} (best) insertion cost of sequence *s* into the insertion point *i* d(r) total demand of the restricted route *r*

d(s) total demand in the node sequence s

An example (cont.d)



11

An example (cont.d)



Initial Solution



Interesting solutions

Instance E-n101-k14 with rounded costs





Initial solution: cost 1076 Xu and Kelly, 1996 Final solution: cost 1067 14 *New best known solution*

Interesting solutions

Instance M-n151-k12 with rounded costs





Initial solution: cost 1023 Gendreau, Hertz and Laporte, 1996 Final solution: cost 1022 15 New best known solution

Some Computational Results

Instance	Optimal	SERR sol.	Gap	Time
P-n50-k8	631	631	0.00%	11:08
P-n55-k10	694	700	0.86%	16:50
P-n60-k10	744	744	0.00%	25:01
P-n60-k15	968	975	0.72%	12:27
P-n65-k10	792	796	0.51%	12:26
P-n70-k10	827	834	0.48%	50:08
B-n68-k9	1272	1275	0.24%	3:02:01
E-n51-k5	521	521	0.00%	4:30
E-n76-k7	682	682	0.00%	27:35
E-n76-k8	735	742	0.95%	30:39
E-n76-k10	830	835	0.60%	1:19:30
E-n76-k14	1021	1032	1.08%	2:45:20
E-n101-k8	815	820	0.61%	2:54:04
E051-05e	524.61	524.61	0.00%	4:51
E076-10e	835.26	835.32	< 0.01%	1:12:05
E101-08e	826.14	831.91	0.70%	2:30:55
E101-10c	819.56	819.56	0.00%	2:35:36
E-n101-k14	-	1076 -> 1067	-	1:36:05
M-n151-k12-a	-	1023 -> 1022	-	7:46:33

New best known solution

Optimal solution(*)

New best heuristic solution known

CPU times in the format [hh:]mm:ss

PC: Pentium M 1.6GHz

(*) Most optimal solutions have been found recently by Fukasawa, Poggi de Aragao, Reis, and Uchoa (September 2003)







Conclusions

Achieved goals

- 1. **Definition** of a new **neighborhood** with exponential cardinality and of an effective (non-polynomial) **search algorithm**
- 2. Simple implementation based on a general ILP solver
- **3. Evaluation** of the algorithm on a widely-used set of instances
- 4. Determination of the **new best solution** for two of the few instances not yet solved to optimality

Future directions of work

- 1. Adaptation of the method to more constrained versions of VRP, including VRP with precedence constraints
- 2. Use of an external metaheuristic scheme

Mipping Chvàtal-Gomory cuts

Matteo Fischetti University of Padova, Italy

Andrea Lodi University of Bologna, Italy IBM, T.J. Watson Research Center



Notation and definitions

• We consider an Integer Linear Program (ILP) of the form:

$$\min\{c^T x : Ax \le b, x \ge 0 \text{ integer}\}$$

and two associated polyhedra:

$$P := \{x \in \mathbb{R}^n_+ : Ax \le b\}$$
$$P_I := conv\{x \in \mathbb{Z}^n_+ : Ax \le b\} = conv(P \cap \mathbb{Z}^n)$$

• A Chvàtal-Gomory (CG) cut is a valid inequality for P_I of the form:

 $\lfloor u^T A \rfloor x \leq \lfloor u^T b \rfloor$ where $u \in R^m_+$ is called the CG multiplier vector, and $\lfloor \cdot \rfloor$ denotes lower integer part.

• The first Chvàtal closure of P is defined as:

$$P_1 := \{ x \ge 0 : Ax \le b, \lfloor u^T A \rfloor x \le \lfloor u^T b \rfloor \text{ for all } u \in I\!\!R^m_+ \}$$

• P_1 is indeed a polyhedron, i.e., a finite number of CG cuts suffice to define it. [Chvàtal 1973]

Notation and definitions (cont.d)

- Clearly, $P_I \subseteq P_1 \subseteq P$.
- Every fractional vertex x^* of P associated with a certain basis B (say) of (A, I) can be cut off by the CG cut in which u is chosen as the *i*-th row of B^{-1} , where *i* is the row associated with any fractional component of x^* . [Gomory 1958,1963]
- In some cases, one has that $P_I = P_1$ as, e.g., for matching problems where undominated CG cuts correspond to the famous Edmonds' blossom inequalities. [Edmonds 1965]
- By the well-known equivalence between optimization and separation, we will address the Chvàtal-Gomory separation problem (CG-SEP) of the form:

Given any point $x^* \in P$ find (if any) a CG cut that is violated by x^* , i.e., find $u \in \mathbb{R}^m_+$ such that $\lfloor u^T A \rfloor x^* > \lfloor u^T b \rfloor$, or prove that no such u exists.

• However, CG-SEP is NP-hard, so optimizing over P_1 also is. [Eisembrand 1999]

Some practical questions

- How difficult is, **in practice**, to optimize *exactly* over the first Chvàtal closure of a generic ILP?
- Which fraction of the integrality gap can be closed this way, e.g., for some hard problems in the MIPLIB library?
- Before affording the effort of designing and implementing sophisticated separation tools, we want to be sure the overall approach has some potentials...



www.Laerazilamaa.ila

MIPping CG separation

- Given the input point $x^* \ge 0$ to be separated, CG-SEP calls for a CG cut $\alpha^T x \le \alpha_0$ which is (maximally) violated by x^* , where $\alpha = \lfloor u^T A \rfloor$ and $\alpha_0 = \lfloor u^T b \rfloor$ for a certain $u \in \mathbb{R}_+$.
- Some properties:
 - 1. Any variable x_j such that $x_j^* = 0$ can be omitted. Indeed, it does not contribute to the violation and its coefficient can be recomputed a posteriori as $\alpha_j := u^T A_j$ (no time-consuming lifting operations being needed).
 - 2. The same holds for variables at their upper bound in x^* , which can be complemented.
 - 3. It is known that one can assume $u_i < 1$ in case the *i*-th row of (A, b) is integer.
- Avoiding weak cuts:
 - Several equivalent solutions of the separation problem (in its optimization version) typically exist, some of which produce very weak cuts.
 - In practice, finding stronger cuts corresponds to producing "minimal" CG multiplier vectors with as few nonzero entries as possible.
- Our approach is to model the rank-1 Chvàtal-Gomory separation problem, which is known to be NP-hard, through a MIP model, which is then solved (exactly/heuristically) through a general-purpose MIP solver.

Mipping CG separation (cont.d)

• We then propose the following MIP model for CG-SEP:

max

$$(\sum_{j\in J(x^*)}\alpha_j x_j^* - \alpha_0) - \sum_{i=1}^m w_i u_i$$
(1)

$$f_j = u^T A_j - \alpha_j, \quad \text{for } j \in J(x^*)$$
 (2)

$$f_0 = u^T b - \alpha_0 \tag{3}$$

$$0 \le f_j \le 1 - \delta, \quad \text{for } j \in J(x^*) \cup \{0\}$$
(4)

$$0 \le u_i \le 1 - \delta$$
, for $i = 1, \cdots, m$ (5)

$$\alpha_j \text{ integer}, \quad \text{for } j \in J(x^*) \cup \{0\}$$
 (6)

where $J(x^*) := \{j \in \{1, \dots, n\} : x_j^* > 0\}$ is the support of x^* (possibly after having complemented some variables and updated b accordingly).

- We chose the $\delta = 0.01$ so as to improve numerical stability.
- We also introduced the penalty term $-\sum_{i} w_{i}u_{i}$ in the objective function (1), where $w_{i} = 10^{-4}$ for all *i*, which is aimed at favoring the "minimality" of the CG multiplier vector *u*.

Solving the CG-separation MIP

- Preliminary experiments where the CG-separation MIP is solved through a commercial general-purpose MIP solver (ILOG-Cplex 9.0.2)
 - 1. When the LP relaxation of the original ILP model is solved, we take all the violated Gomory fractional cuts that can be read from the tableau, and skip CG separation.
 - 2. The MIP solver for CG separation is invoked with an initial lower bound of 0.01, meaning that we are only interested in CG cuts violated by more than 0.01.
 - 3. At each update of the MIP incumbent solution x^* , the corresponding CG cut is stored in a pool, and added at the end of the separation phase (among the cuts with the same violation, only the one with the sparsest support is added).
 - 4. The MIP execution for CG separation is stopped if:
 - either the optimal solution has been found,
 - or au branching nodes has been explored after the last x^* update.
 - $\tau=1000$ if the violation of the incumbent is less than 0.2, and $\tau=100$ otherwise.
- We keep generating violated CG cuts of rank 1 until either no such violated cut exists (in which case we have optimized over the first closure), or because a time-limit condition is met.

Can we solve matching problems?

ID	initial LB	Optimum	# iter.s	$\# \ cuts$	CPU time
eil101	619.0	623.0	26	43	9.01
gr120	6,662.5	6,694.0	33	45	10.47
pr124	50,164.0	51,477.0	124	320	555.54
gr137	66,643.5	67,009.0	11	31	1.68
pr144	32,776.0	33,652.0	39	78	9.57
ch150	6,281.0	6,337.0	59	141	71.19
rat195	2,272.5	2,297.0	85	237	202.87
kroA200	27,053.0	27,426.0	26	84	10.93
kroB200	27,347.0	27,768.0	189	558	2,249.55
ts225	115,605.0	121,261.0	323	857	4,906.48
pr226	55,247.5	57,177.0	401	901	4,077.66
gr229	127,411.0	128,353.0	78	224	219.00
gil262	2,222.5	2,248.0	105	266	372.10
a280	2,534.0	2,550.0	52	104	40.21
lin318	38,963.5	39,266.0	292	768	6,103.32

• We started in a "friendly" setting by addressing 2-matching problems

Can we solve matching problems? (cont.d)

- Some of these instances can be solved in a much shorter computing time by just applying ILOG-Cplex 9.0.2 MIP solver (by heavy branching), and obviously by considering the use of the special purpose separation of 2-matching inequalities. [Letchford, Reinelt and Theis 2004]
- However, for some hard instances a cut-and-branch approach in which we separate 100 rounds of rank-1 cuts, and we then switch to a commercial MIP solver for concluding the optimization gives promising results.

	ILOG-Cplex		cut-and-branch					
	% gap				% gap	separation		total
ID	closed	nodes	time	# cuts	closed	time	nodes	time
pr124	100.0	43,125	104.17	116	62.1	27.96	1,925	37.51
kroB200	100.0	330,913	2,748.24	129	64.1	49.34	4,113	76.30
ts225	47.1	230,115	1h	250	80.7	164.77	13,552	352.35
pr226	55.0	288,901	1h	179	62.9	61.13	19,977	281.89
gr229	100.0	15,005	180.79	126	82.8	9.65	155	60.94
gil262	100.0	117,506	2,094.77	110	84.0	12.24	217	36.78
lin318	53.3	117,100	1h	187	64.9	110.69	25,953	933.97

How tight is the first closure for MIPLIB instances?

• Instances from MIPLIB, time limit of 3 hours

				% gap	
ID	Optimum	# iter.s	# cuts	closed	time
air03	340,160.00	1	35	100.0	1.47
gt2	21,166.00	160	424	100.0	506.25
lseu	1,120.00	73	190	91.3	565.22
mitre	115,155.00	1,509	5,398	100.0	9,394.17
mod008	307.00	26	109	100.0	8.00
mod010	6,548.00	17	62	100.0	13.05
nw04	16,862.00	78	236	100.0	227.13
p0033	3,089.00	40	152	85.4	12.95
p0548	8,691.00	886	3,356	100.0	1,575.83
stein27	18.00	98	295	0.0	490.02

- A cut-and-branch approach on instance harp2 gave very interesting results:
 - 100 rounds of separation (211 rank-1 CG cuts, 53 tight at the end),
 - 1,500 CPU seconds and 400K nodes (including both cut generation and branching).
 - ILOG-Cplex alone required more than 15,000 CPU seconds and 7M nodes.

Beyond the first closure?

• We addressed the possibility of using our cutting plane method as a pre-processing tool, to be used to strengthen the user's formulation by exploiting cuts of Chvàtal rank larger than 1.

This idea was evaluated by comparing two different cut preprocessors, namely:

- *cpx*: Apply ILOG-Cplex 9.0.2 (with mip emphasis "move best bound") on the current ILP model, save the final root-node model (including the generated cuts) in a file, and repeat on the new model until a total time limit is exceeded.
- *cpx-cg*: Apply ILOG-Cplex 9.0.2 (with mip emphasis "move best bound") on the current ILP model, followed by 600 seconds of our CG separation procedure; then save in file the ILP model with all the cuts that are active in the last LP solution, and repeat on the new model until a total time limit is exceeded.
- For the first time we found a provable optimal solution of value 51,200.00 for the very hard instance nsrand-ipx.
- Precisely, *cpx-cg* ran for 4,800 CPU seconds obtaining a tightened formulation that brought the initial LP bound from 49,667.89 to 50,665.71.

Beyond the first closure? (cont.d)



Figure 1: Lower bounds provided by *cpx* and *cpx-cg* after each call of the separation procedures, for the hard MIPLIB instance *timtab1*.

Can we discover new classes of strong inequalities?

- As in the spirit of PORTA, we can use the framework for obtaining off-line the facial structure of a specific problem. Advantage: we are not restricted to instances of very small size.
- To illustrate a possible application to the Asymmetric Travelling Salesman Problem (ATSP), we took a partial ATSP formulation including out- and in-degree equations, plus the SECs on 2-node sets, i.e., the NP-hard Asymmetric Assignment Problem (AAP) relaxation. [Balas 1989]
- We applied the method to ry48p from ATSPLIB, and we stored the CG cuts along with the associated CG multipliers.
- Through a careful analysis of one returned cut we have that:

which is (by computational methods) facet-defining for ATSP. Using *clique lifting* we can then obtain a large class of ATSP facets, that to the best of our knowledge is new.

		Split closure	CG closure
% Gap closed	Average	71.71	62.59
% Gap closed	98-100	9 instances	9 instances
% Gap closed	75 - 98	4 instances	2 instances
% Gap closed	25-75	6 instances	7 instances
% Gap closed	< 25	6 instances	$7 \ instances$

Table 1: Results for 25 *pure* integer linear programs in the MIPlib 3.0.

		Split closure	pro-CG closure
% Gap closed	Average	84.34	36.38
% Gap closed	98-100	16 instances	3 instances
% Gap closed	75-98	10 instances	3 instances
% Gap closed	25-75	2 instances	11 instances
% Gap closed	< 25	5 instances	17 instances

Table 2: Results for 33 mixed integer linear programs in the MIPlib 3.0.

Conclusion and future work

- We have been able to show computationally the quality of the Chvàtal-Gomory cuts and to answer (at least partially) to several natural questions about their **practical** effectiveness.
- Although an NP-hard problem has to be solved to separate inequalities of rank 1, the issue of generating those cuts is affordable in practice and it definitely deserves attention.
- An obvious issue for future research is the design of more specific separation procedures for CG cuts, i.e., ad-hoc heuristics for the corresponding MIP model.





joint work with F. Glover (Univ. Colorado at Boulder, USA) and A. Lodi (DEIS, Univ. Bologna)

The basic scheme

• How do you define feasibility for a MIP problem of the form:

$$\min\{c^T x : Ax \ge b, x_j \text{ integer } \forall j \in \mathcal{I}\} ?$$

• We propose the following definition:

a feasible solution is a point $x^* \in P := \{x: Ax \geq b\}$ that is coincident with its rounding \widetilde{x}

where:

- 1. $[\cdot]$ represents scalar rounding to the nearest integer;
- 2. $\widetilde{x}_j := [x_j^*]$ if $j \in \mathcal{I}$; and
- 3. $\tilde{x}_j := x_j^*$ otherwise.
- Replacing coincident with as close as possible relatively to a suitable distance function $\Delta(x^*, \tilde{x})$ suggests an iterative heuristic for finding a feasible solution of a given MIP.

Definition of $\Delta(x^*, \widetilde{x})$

We consider the L₁-norm distance between a generic point x ∈ P and a given integer x̃, defined as:

$$\Delta(x, \tilde{x}) = \sum_{j \in \mathcal{I}} |x_j - \tilde{x}_j|$$

• Assuming (for the sake of notation) that all integer-constrained variables are binary, $\Delta(x^*, \tilde{x})$ attains the simple form:

$$\Delta(x,\tilde{x}) := \sum_{j \in \mathcal{I}: \tilde{x}_j = 0} x_j + \sum_{j \in \mathcal{I}: \tilde{x}_j = 1} (1 - x_j) \tag{1}$$

• Given an integer \tilde{x} , the closest point $x^* \in P$ can therefore be determined by solving the LP:

$$\min\{\Delta(x,\tilde{x}): Ax \ge b\}$$
(2)

- We start from any $x^* \in P$, and define its rounding \tilde{x} .
- At each iteration we look for a point x^{*} ∈ P which is as close as possible to the current x̃ by solving the problem:

$$\min\{\Delta(x,\tilde{x}): x \in P\}$$

Assuming $\Delta(x, \tilde{x})$ is chosen appropriately, is an easily solvable LP problem.

- If $\Delta(x^*, \tilde{x}) = 0$, then x^* is a feasible MIP solution and we are done.
- Otherwise, we replace \tilde{x} by the rounding of x^* , and repeat.

- From a geometric point of view, this simple heuristic generates two hopefully convergent trajectories of points x^* and \tilde{x} which satisfy feasibility in a complementary but partial way:
 - 1. one, x^* , satisfies the linear constraints,
 - 2. the other, \tilde{x} , the integer requirement.

A basic **FP** implementation

1. initialize nIT := 0 and $x^* := \operatorname{argmin} \{c^T x : Ax \ge b\};$ 2. if x^* is integer, return (x^*) ; 3. let $\tilde{x} := [x^*]$ (= rounding of x^*); 4. while (time < TL) do 5. let nIT := nIT + 1 and compute $x^* := \operatorname{argmin} \{\Delta(x, \tilde{x}) : Ax \ge b\};$ 6. if x^* is integer, return (x^*) ; 7. if $[x^*] \ne \tilde{x}$ then 8. $\tilde{x} := [x^*]$ else 9. flip the TT = rand(T/2,3T/2) entries \tilde{x}_j $(j \in \mathcal{I})$ with highest $|x_j^* - \tilde{x}_j|$ 10. endif 11. enddo

Plot of the infeasibility measure $\Delta(x^*,\widetilde{x})$ at each iteration



Summary of the computational results

- ILOG-Cplex 8.1 is run on default version but avoiding preprocessing (following the suggestion of Ed Rothberg).
- FP solves LPs by leaving ILOG-Cplex decide which is the best algorithm (CPXoptimize).
- Over 83 hard 0-1 MIP instances in the MIPLIB test-bed:

FP failed in finding a feasible solution only in 3 case, while

ILOG-Cplex 8.1 failed 19 times.

• The quality of the solutions obtained is generally comparable, as well as the computing times.