Mixed-Integer Programming (MIP):

\[
\begin{align*}
\text{min} & \quad c^T x \\
\text{s.t.} & \quad Ax \leq b \\
& \quad x \in \mathbb{Z}_{\geq 0}^{n_l} \times \mathbb{R}_{\geq 0}^{n_c}
\end{align*}
\]

Primal heuristics:
- are incomplete methods which
- often find good solutions
- within a reasonable time
- without any warranty!

Inside an exact solver:
- they prove feasibility
- nearly optimal might be sufficient
- primal bound needed for pruning
- solutions guide remaining search
Categories of Heuristics

▷ Diving
  ▷ simulate DFS with special branching rule
  ▷ e.g., guided diving
  ▷ one LP resolve (dual simplex) per iteration

▷ Objective diving
  ▷ manipulate objective function
  ▷ e.g., feasibility pump
  ▷ one LP resolve (primal simplex) per iteration

▷ Large Neighborhood Search
  ▷ solve sub-MIP
  ▷ e.g., RINS, Local Branching
  ▷ 500 nodes of a MIP

▷ Rounding, Propagation
  ▷ no additional LPs or MIPs
How important are primal heuristics?

A major MIP software vendor says:

Our advanced MIP heuristics for quickly finding feasible solutions often produce good quality solutions where other solvers fall flat, leading to some of our biggest wins vs. the competition.
How important are primal heuristics?

A major MIP software vendor says:

Our advanced MIP heuristics for quickly finding feasible solutions often produce good quality solutions where other solvers fall flat, leading to some of our biggest wins vs. the competition.

Which heuristics?

▷ business secret
▷ few parameters to influence heuristics’ behavior
▷ output only tells you that solution found by some heuristic
How important are primal heuristics?

A major MIP software vendor says:

Our advanced MIP heuristics for quickly finding feasible solutions often produce good quality solutions where other solvers fall flat, leading to some of our biggest wins vs. the competition.

Which heuristics?

▷ business secret
▷ few parameters to influence heuristics’ behavior
▷ output only tells you that solution found by some heuristic

⇝ very important
Typical measure: Running time to prove optimality
Typical measure: Running time to prove optimality

- one vendor: 6% improvement
- other vendor: 9% improvement
- non-commercial solver: 15% improvement
Typical measure: Running time to prove optimality
▷ one vendor: 6% improvement
▷ other vendor: 9% improvement
▷ non-commercial solver: 15% improvement

⇝ not important at all
Typical measure: Running time to prove optimality

- one vendor: 6% improvement
- other vendor: 9% improvement
- non-commercial solver: 15% improvement

⇉ not important at all

So, what is wrong here?
Typical measure: Running time to prove optimality
▷ one vendor: 6% improvement
▷ other vendor: 9% improvement
▷ non-commercial solver: 15% improvement

→ not important at all

So, what is wrong here?
Goal of this talk: Introduce a new performance measure
How to measure the added value of a primal heuristic?

- time to optimality $t_{\text{solved}}$, number of branch-and-bound nodes
  - very much depends on dual bound

- time to best solution $t_{\text{opt}}$
  - nearly optimal solution might be found long before

- time to first solution $t_{1}$
  - disregards solution quality

- performance profiles
  - depend on $t_{\text{solved}}$, hence on dual bound
  - not an absolute number

- primal integral
  - favors finding good solutions early
  - considers each update of incumbent
  - $P(t_{\text{max}})/t_{\text{max}}$ “average solution quality”
  - expected quality of the incumbent, if stopped arbitrarily
Comparing performance

How to measure the added value of a primal heuristic?

- time to optimality $t_{\text{solved}}$, number of branch-and-bound nodes
  - very much depends on dual bound
- time to best solution $t_{\text{opt}}$
  - nearly optimal solution might be found long before

Performance profiles depend on $t_{\text{solved}}$, hence on dual bound.

Primal integral favors finding good solutions early.

$P(t_{\text{max}})/t_{\text{max}}$ "average solution quality" expected quality of the incumbent, if stopped arbitrarily.
How to measure the added value of a primal heuristic?

- time to optimality $t_{\text{solved}}$, number of branch-and-bound nodes
  - very much depends on dual bound
- time to best solution $t_{\text{opt}}$
  - nearly optimal solution might be found long before
- time to first solution $t_1$
  - disregards solution quality

Performance profiles depend on $t_{\text{solved}}$, hence on dual bound
- not an absolute number

Primal integral favors finding good solutions early
- considers each update of incumbent
  - $P(t_{\text{max}})/t_{\text{max}}$ "average solution quality"
  - expected quality of the incumbent, if stopped arbitrarily
Comparing performance

How to measure the added value of a primal heuristic?

- time to optimality $t_{\text{solved}}$, number of branch-and-bound nodes
  - very much depends on dual bound
- time to best solution $t_{\text{opt}}$
  - nearly optimal solution might be found long before
- time to first solution $t_1$
  - disregards solution quality
- performance profiles
  - depend on $t_{\text{solved}}$, hence on dual bound
  - not an absolute number
Comparing performance

How to measure the added value of a primal heuristic?

- time to optimality $t_{\text{solved}}$, number of branch-and-bound nodes
  - very much depends on dual bound
- time to best solution $t_{\text{opt}}$
  - nearly optimal solution might be found long before
- time to first solution $t_1$
  - disregards solution quality
- performance profiles
  - depend on $t_{\text{solved}}$, hence on dual bound
  - not an absolute number
- primal integral
3 steps we take on the next slides:

- primal gap
- primal gap function
- primal integral

3 pieces of information that we need:

- an optimal or best known solution $\tilde{x}_{opt}$
- development of incumbent solution (log file)
- the time limit $t_{max}$
Let $\tilde{x}$ be a solution, $\tilde{x}_{opt}$ be an optimum, $t_{max} \in \mathbb{R}_{\geq 0}$ be a timelimit.

Primal gap $\gamma \in [0, 1]$ of $\tilde{x}$:

$$
\gamma(\tilde{x}) :=
\begin{cases}
0, & \text{if } |c^T\tilde{x}_{opt}| = |c^T\tilde{x}| = 0, \\
1, & \text{if } c^T\tilde{x}_{opt} \cdot c^T\tilde{x} < 0, \\
\frac{|c^T\tilde{x}_{opt} - c^T\tilde{x}|}{\max\{|c^T\tilde{x}_{opt}|, |c^T\tilde{x}|\}}, & \text{else}.
\end{cases}
$$

Primal gap function $p : [0, t_{max}] \mapsto [0, 1]$:

$$
p(t) :=
\begin{cases}
1, & \text{if no incumbent until point } t, \\
\gamma(\tilde{x}(t)), & \text{with } \tilde{x}(t) \text{ incumbent at point } t.
\end{cases}
$$
Primal integral

- step function, changes at points $t_i$ when new incumbent found
- $p(0) = 1$, $p(t) = 0$ for all $t \geq t_{\text{opt}}$
- monotonously decreasing

Primal integral $P(T)$ of $T \in [0, t_{\text{max}}]$:

$$P(T) := \int_{t=0}^{T} p(t) \, dt = \sum_{i=1}^{l} p(t_{i-1}) \cdot (t_i - t_{i-1}),$$
How to measure the added value of a primal heuristic?

- time to optimality $t_{\text{solved}}$, number of branch-and-bound nodes
  - very much depends on dual bound
- time to best solution $t_{\text{opt}}$
  - nearly optimal solution might be found long before
- time to first solution $t_1$
  - disregards solution quality
- performance profiles
  - depend on $t_{\text{solved}}$, hence on dual bound
  - not an absolute number

- primal integral $P(t_{\text{max}})$
  - favors finding good solutions early
  - considers each update of incumbent
  - $P(t_{\text{max}})/t_{\text{max}}$ “average solution quality”
  - expected quality of the incumbent, if stopped arbitrarily
MIPLIB2010:

- 361 instances, benchmark set: 87
- 120–160k vars, 32–624k rows, 666–27M nz
- industry and academics
- diverse applications, combinatorics
- major vendors in committee
- http://miplib.zib.de
- + MIPLIB2003, MIPLIB 3.0
SCIP: Solving Constraint Integer Programs

- standalone solver / branch-cut-and-price-framework
- modular structure via plugins
- free for academic use: http://scip.zib.de
- very fast non-commercial MIP solver

![Bar chart showing performance comparison of different solvers.](chart.png)

results by H. Mittelmann (09/Aug/2012)
SCIP: Solving Constraint Integer Programs

- better support of MINLP
- new presolvers and propagators
- AMPL and MATLAB interface (beta)
- first releases of GCG and UG

![Graph showing performance improvements for different SCIP versions](image-url)
Primal Heuristics in SCIP

actcons diving

diving
cross over
cover
cover
under cover

cof diving

guided diving

diving

fracdiving

diving

intdiving
diving

veclen diving

pscost diving

objpscost diving

feaspump

cover
dins

intdiving

diving

linesearch diving

rootsol diving

of

Presolver

Branch

allfull

strong

full

strong

infer

eference

leastinf

mostinf

pscost

random

relps
cost

Conflict

Constraint

Handler

and

disjunc.
count

disjunctive

cumulative

disjunctive

indi
cator

integer

knap

sack

linear

linking

logicor

or

orbi
tope

quadr

setppc

soc

sos1

sos2

var

bound

xor

Cutpool

LP

clp
cpx

msk

none

qso

spx

xprs

Dialog

default

Display

default

Node

selector

bfs

dfs

estimate

hybrid

estim

restart

dfs

· · ·

Pricer

Separator

clique
cmir

flow

cover

gomory

implied

bounds

intobj

mcf

odd
cycle

rapid

learn

redcost

strong

cg

zero

half

Propagator

Relax

Primal Heuristics

Timo Berthold: Measuring the Impact of Primal Heuristics 14 / 24
Categories of Heuristics

▷ **Diving**
  ▷ simulate DFS with special branching rule
  ▷ e.g., guided diving
  ▷ one LP resolve (dual simplex) per iteration

▷ **Objective diving**
  ▷ manipulate objective function
  ▷ e.g., feasibility pump
  ▷ one LP resolve (primal simplex) per iteration

▷ **Large Neighborhood Search**
  ▷ solve sub-MIP
  ▷ e.g., RINS, Local Branching
  ▷ 500 nodes of a MIP

▷ **Rounding, Propagation**
  ▷ no additional LPs or MIPs
Solving process for n3seq24

SCIP default, $P(t_{\text{max}}) = 421$

no round & prop, $P(t_{\text{max}}) = 1073$

only round & prop, $P(t_{\text{max}}) = 797$

no heuristics, $P(t_{\text{max}}) = 1050$
### Computational Results

<table>
<thead>
<tr>
<th></th>
<th>def</th>
<th>noheur</th>
<th>nodive</th>
<th>noobj</th>
<th>nolns</th>
<th>noround</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi(t_1)$</td>
<td>3.8</td>
<td>15.4</td>
<td>4.1</td>
<td>3.9</td>
<td>4.0</td>
<td>6.1</td>
</tr>
<tr>
<td>$\phi(t_{opt})$</td>
<td>43.2</td>
<td>47.9</td>
<td>43.9</td>
<td>44.4</td>
<td>44.8</td>
<td>48.4</td>
</tr>
<tr>
<td>$\phi(t_{solved})$</td>
<td>107.5</td>
<td>114.7</td>
<td>109.7</td>
<td>114.8</td>
<td>110.2</td>
<td>105.9</td>
</tr>
<tr>
<td>$\phi(P(t_{max}))$</td>
<td>257</td>
<td>363</td>
<td>299</td>
<td>277</td>
<td>275</td>
<td>263</td>
</tr>
<tr>
<td>$\phi(P(t_{max}))/t_{max}$</td>
<td>7.1%</td>
<td>10.1%</td>
<td>8.3%</td>
<td>7.7%</td>
<td>7.6%</td>
<td>7.3%</td>
</tr>
</tbody>
</table>

- primal heuristics extremely important for first solution
- rounding heuristics: slight degradation for time to optimality
- $P(t_{max})$: def $\prec$ noround $\prec$ nolns $\approx$ noobj $\prec$ nodive $\prec$ noheur
- primal heuristics decrease average gap by more than 40%
## Computational Results

<table>
<thead>
<tr>
<th></th>
<th>def</th>
<th>noheur</th>
<th>dive</th>
<th>obj</th>
<th>Ins</th>
<th>round</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi(t_1)$</td>
<td>3.8</td>
<td>15.4</td>
<td>9.7</td>
<td>7.2</td>
<td>11.8</td>
<td>4.8</td>
</tr>
<tr>
<td>$\phi(t_{\text{opt}})$</td>
<td>43.2</td>
<td>47.9</td>
<td>54.3</td>
<td>53.2</td>
<td>44.6</td>
<td>43.6</td>
</tr>
<tr>
<td>$\phi(t_{\text{solved}})$</td>
<td>107.5</td>
<td>114.7</td>
<td>115.3</td>
<td>108.6</td>
<td>110.5</td>
<td>112.9</td>
</tr>
<tr>
<td>$\phi(P(t_{\text{max}}))$</td>
<td>257</td>
<td>363</td>
<td>329</td>
<td>309</td>
<td>355</td>
<td>349</td>
</tr>
<tr>
<td>$\phi(P(t_{\text{max}}))/t_{\text{max}}$</td>
<td>7.1%</td>
<td>10.1%</td>
<td>9.1%</td>
<td>8.6%</td>
<td>9.9%</td>
<td>9.7%</td>
</tr>
</tbody>
</table>

- again, hardly any change in $t_{\text{opt}}$ and $t_{\text{solved}}$
- rounding heuristics important for $t_1$
- $P(t_{\text{max}})$: single class cannot compensate the other
Average primal integral

$p(t)$ in %

t in sec.

no heuristics
default
variants of the primal integral:

- logarithmic time-axis (twice as early = twice as good)
- logarithmic gap-axis (twice as close to opt. = twice as good)
- consider dual gap (e.g. for cuts) or primal-dual gap
- consider other performance measures that change monotonously
variants of the primal integral:

- logarithmic time-axis (twice as early = twice as good)
- logarithmic gap-axis (twice as close to opt. = twice as good)
- consider dual gap (e.g. for cuts) or primal-dual gap
- consider other performance measures that change monotonously

future tests:

- test single primal heuristics
  - change SCIP defaults
  - which heuristics on which problems
- compare different solvers
variants of the primal integral:

- logarithmic time-axis (twice as early = twice as good)
- logarithmic gap-axis (twice as close to opt. = twice as good)
- consider dual gap (e.g. for cuts) or primal-dual gap
- consider other performance measures that change monotonously

future tests:

- test single primal heuristics
  - change SCIP defaults
  - which heuristics on which problems
- compare different solvers ...someone?
Conclusion

Primal integral:
- new performance measure
- captures overall solution process
- principle idea can be transferred to other measures

Measuring the impact:
- impact on time to optimality negligible
- overall impact (w.r.t. $P(t_{\text{max}})$) significant
- impact of single classes of heuristics limited
An exact rational mixed-integer programming solver  
Kati Wolter, Wed.2.H0110

A generic branch-price-and-cut solver  
Marco Lübbecke, Wed.3.H2032

Advances in linear programming  
Matthias Miltenberger, Thu.3.H2033

LNS and diving heuristics in column generation algorithms  
Christian Puchert, Thu.3.H2032

Approaches to solve mixed integer semidefinite programs  
Sonja Mars, Thu.3.H2033

ParaSCIP and FiberSCIP – Parallel extensions of SCIP  
Yuji Shinano, Fri.3.H1058

A computational comparison of symmetry handling methods in IP  
Marc Pfetsch, Fri.3.H2013
Measuring the Impact of Primal Heuristics

Timo Berthold
Zuse Institute Berlin

DFG Research Center MATHEON
Mathematics for key technologies

ISMP 2012, Berlin, 21/Aug/2012