Packing with biased random-key genetic algorithms

Mauricio G. C. Resende AT&T Labs Research Florham Park, New Jersey

mgcr@research.att.com



Lecture given at XV ONPCE S. J. do Rio Preto, Brazil ❖ March 21-22, 2013

Summary

- Metaheuristics and basic concepts of genetic algorithms
- Random-key genetic algorithm of Bean (1994)
- Biased random-key genetic algorithms (BRKGA)
 - Encoding / Decoding
 - Initial population
 - Evolutionary mechanisms
 - Problem independent / problem dependent components
 - Multi-start strategy
 - Specifying a BRKGA
 - Application programming interface (API) for BRKGA
- BRKGA for 2-dim and 3-dim packing
- BRKGA for 3-dim bin packing



Metaheuristics

Metaheuristics are heuristics to devise heuristics.



Metaheuristics

Metaheuristics are high level procedures that coordinate simple heuristics, such as local search, to find solutions that are of better quality than those found by the simple heuristics alone.



Metaheuristics

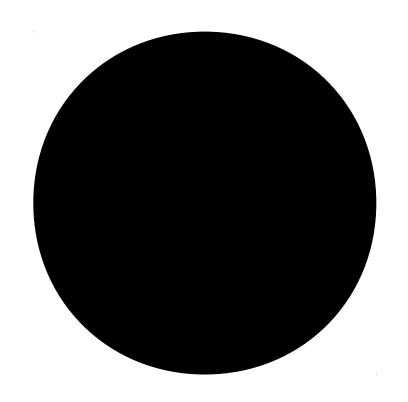
Metaheuristics are high level procedures that coordinate simple heuristics, such as local search, to find solutions that are of better quality than those found by the simple heuristics alone.

Examples: GRASP and C-GRASP, simulated annealing, genetic algorithms, tabu search, scatter search, ant colony optimization, variable neighborhood search, and biased random-key genetic algorithms (BRKGA).





Holland (1975)

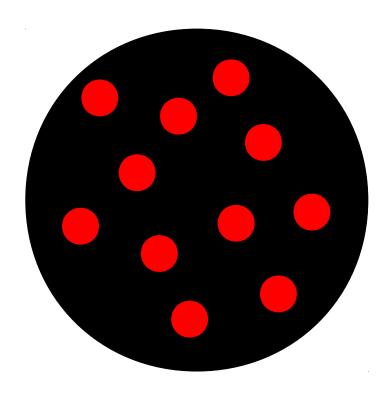


Adaptive methods that are used to solve search and optimization problems.

Individual: solution



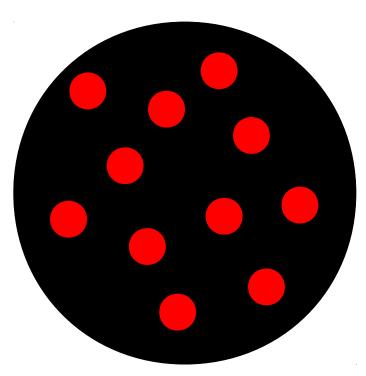




Individual: solution (chromosome = string of genes)

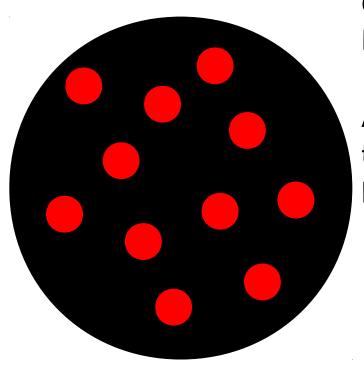
Population: set of fixed number of individuals





Genetic algorithms evolve population applying Darwin's principle of survival of the fittest.

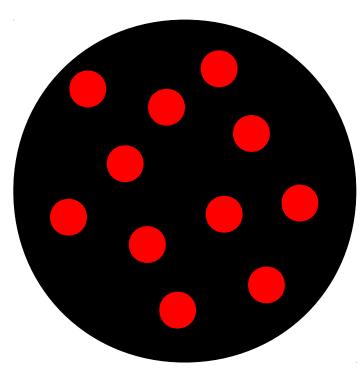




Genetic algorithms evolve population applying Darwin's principle of survival of the fittest.

A series of generations are produced by the algorithm. The most fit individual of the last generation is the solution.



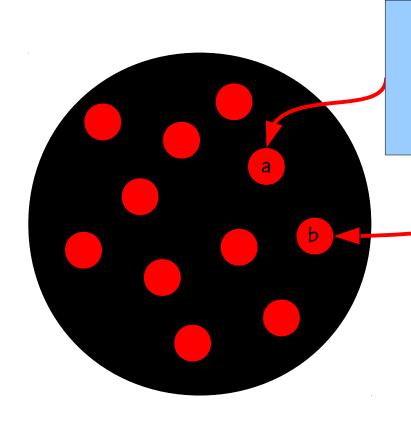


Genetic algorithms evolve population applying Darwin's principle of survival of the fittest.

A series of generations are produced by the algorithm. The most fit individual of the last generation is the solution.

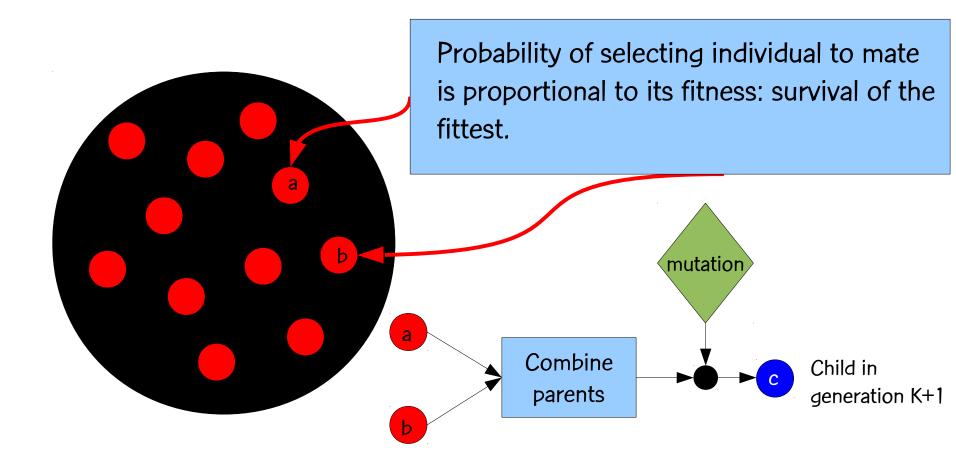
Individuals from one generation are combined to produce offspring that make up next generation.





Probability of selecting individual to mate is proportional to its fitness: survival of the fittest.

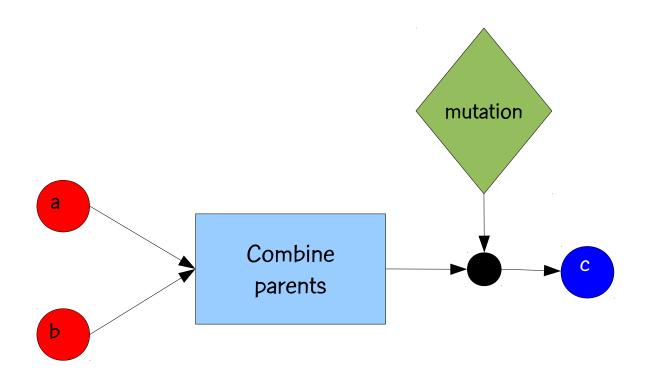




Parents drawn from generation K

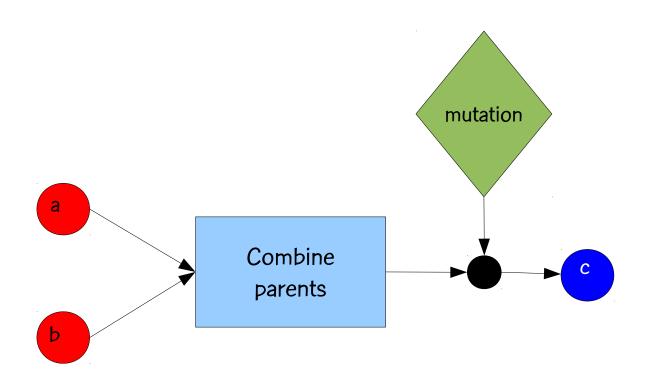


Crossover and mutation





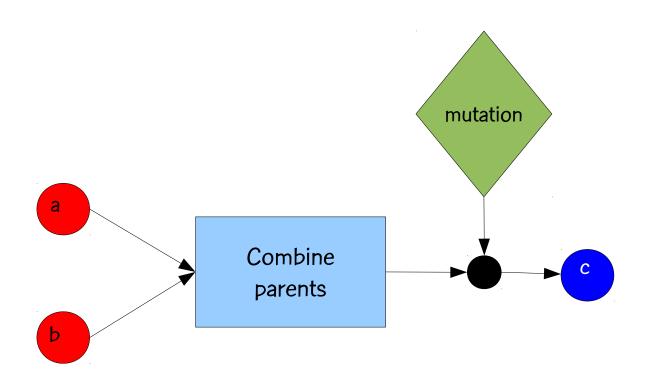
Crossover and mutation



Crossover: Combines parents ... passing along to offspring characteristics of each parent ...

Intensification of search

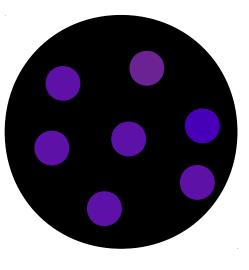
Crossover and mutation



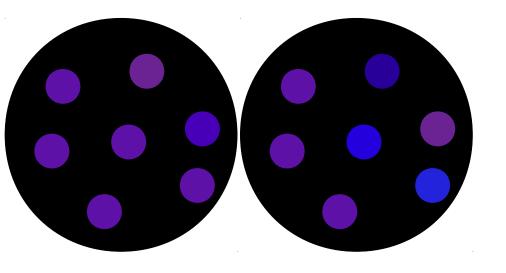
Mutation: Randomly changes chromosome of offspring ...

Driver of evolutionary process ...

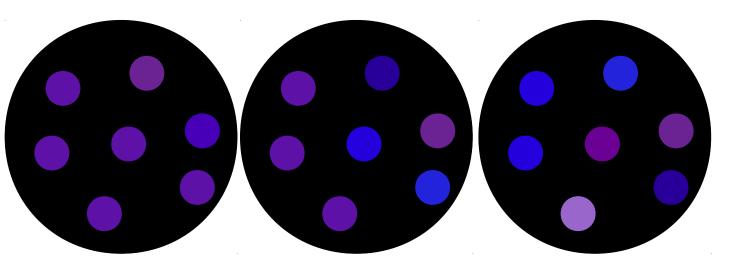
Diversification of search



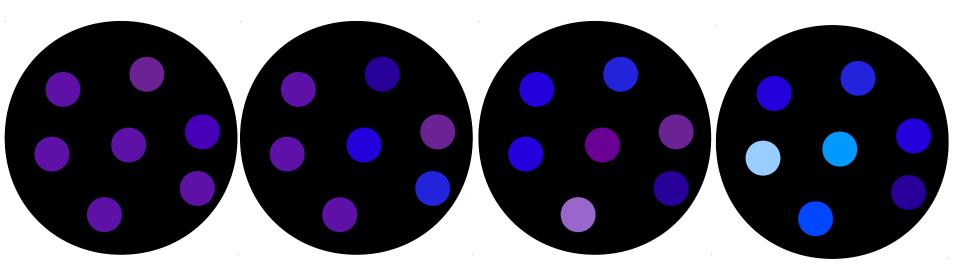




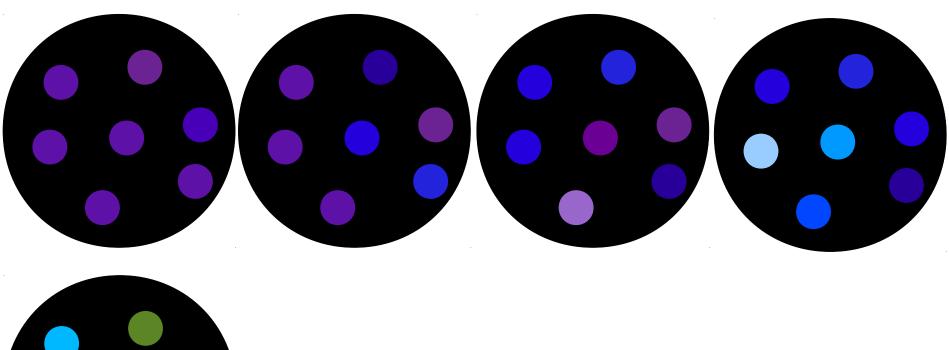


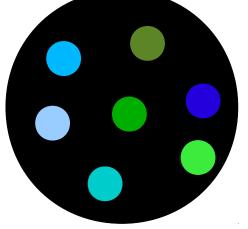




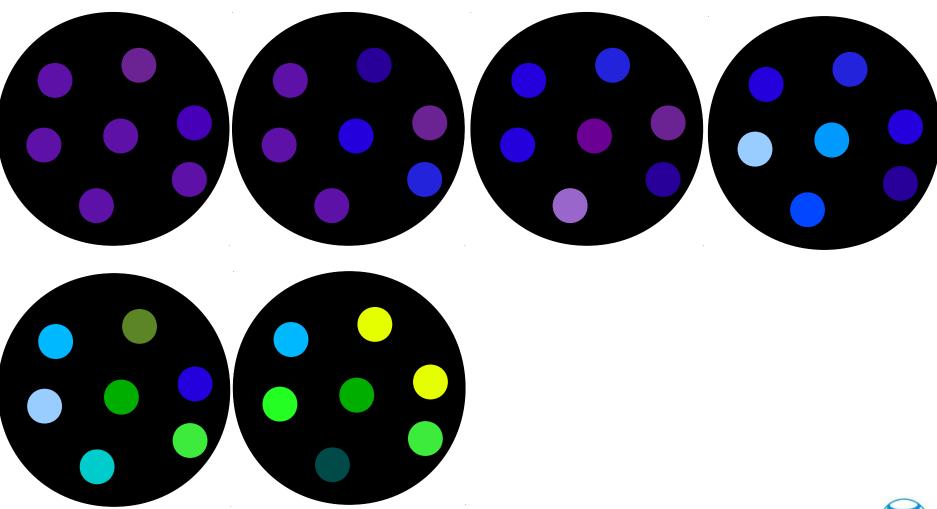




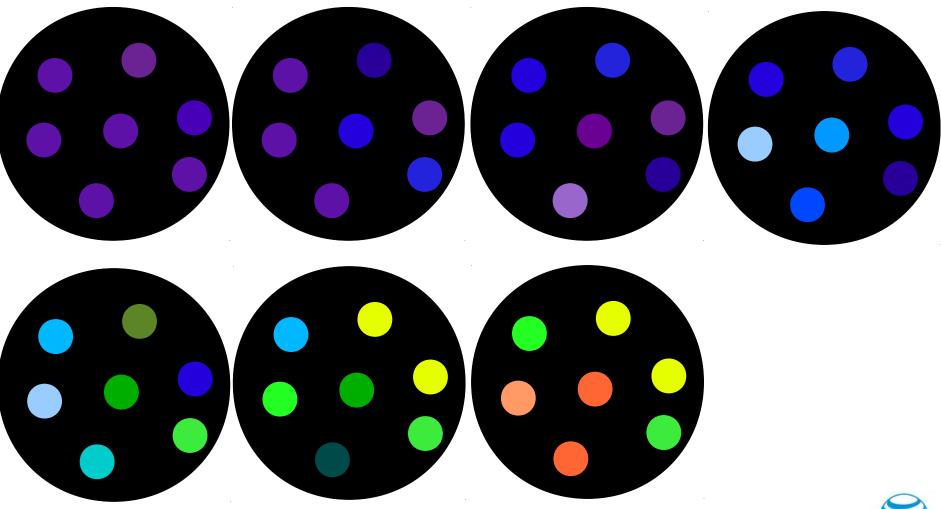


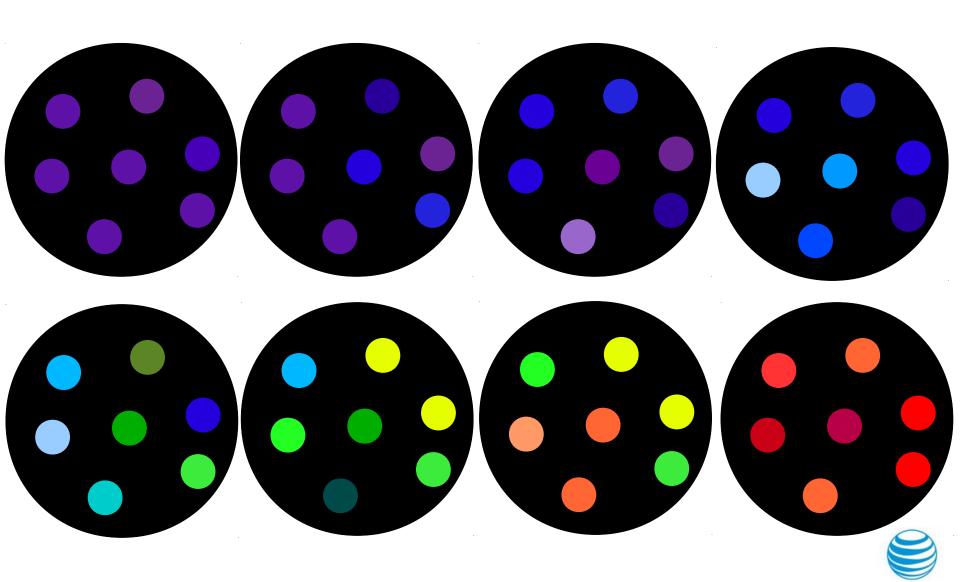




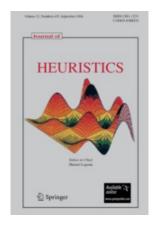








Reference



J.F. Gonçalves and M.G.C.R., "Biased random-key genetic algorithms for combinatorial optimization," J. of Heuristics, vol.17, pp. 487-525, 2011.

Tech report version:

http://www.research.att.com/~mgcr/doc/srkga.pdf



Encoding solutions with random keys



 A random key is a real random number in the continuous interval [0,1).



- A random key is a real random number in the continuous interval [0,1).
- A vector X of random keys, or simply random keys, is an array of n random keys.



- A random key is a real random number in the continuous interval [0,1).
- A vector X of random keys, or simply random keys, is an array of n random keys.
- Solutions of optimization problems can be encoded by random keys.



- A random key is a real random number in the continuous interval [0,1).
- A vector X of random keys, or simply random keys, is an array of n random keys.
- Solutions of optimization problems can be encoded by random keys.
- A decoder is a deterministic algorithm that takes a vector of random keys as input and outputs a feasible solution of the optimization problem.



Encoding with random keys: Sequencing

Encoding

```
[ 1, 2, 3, 4, 5]
```

X = [0.099, 0.216, 0.802, 0.368, 0.658]



Encoding with random keys: Sequencing

Encoding

```
[ 1, 2, 3, 4, 5]
```

X = [0.099, 0.216, 0.802, 0.368, 0.658]

Decode by sorting vector of random keys

```
[ 1, 2, 4, 5, 3]
```

X = [0.099, 0.216, 0.368, 0.658, 0.802]



Encoding with random keys: Sequencing

Therefore, the vector of random keys:

X = [0.099, 0.216, 0.802, 0.368, 0.658]

encodes the sequence: 1-2-4-5-3



Encoding with random keys: Subset selection (select 3 of 5 elements)

Encoding

```
[ 1, 2, 3, 4, 5]
X = [ 0.099, 0.216, 0.802, 0.368, 0.658 ]
```



Encoding with random keys: Subset selection (select 3 of 5 elements)

```
Encoding
```

```
[ 1, 2, 3, 4, 5]
```

```
X = [0.099, 0.216, 0.802, 0.368, 0.658]
```

Decode by sorting vector of random keys

```
[ 1, 2, 4, 5, 3]
```

$$X = [0.099, 0.216, 0.368, 0.658, 0.802]$$



Encoding with random keys: Subset selection (select 3 of 5 elements)

Therefore, the vector of random keys:

X = [0.099, 0.216, 0.802, 0.368, 0.658]

encodes the subset: {1, 2, 4}



Encoding with random keys: Assigning integer weights $\in [0,10]$ to a subset of 3 of 5 elements

Encoding

```
[ 1, 2, 3, 4, 5 | 1, 2, 3, 4, 5]
```

 $X = [0.099, 0.216, 0.802, 0.368, 0.658 \mid 0.4634, 0.5611, 0.2752, 0.4874, 0.0348]$



Encoding with random keys: Assigning integer weights $\in [0,10]$ to a subset of 3 of 5 elements

Encoding

```
[ 1, 2, 3, 4, 5 | 1, 2, 3, 4, 5]
```

 $X = [0.099, 0.216, 0.802, 0.368, 0.658 \mid 0.4634, 0.5611, 0.2752, 0.4874, 0.0348]$

Decode by sorting the first 5 keys and assign as the weight the value $W_i = floor [10 X_{5+i}] + 1$ to the 3 elements with smallest keys X_i , for i = 1,...,5.



Encoding with random keys: Assigning integer weights $\in [0,10]$ to a subset of 3 of 5 elements

Therefore, the vector of random keys:

X = [0.099, 0.216, 0.802, 0.368, 0.658 | 0.4634, 0.5611, 0.2752, 0.4874, 0.0348]

encodes the weight vector W = (5,6,-,5,-)



Genetic algorithms and random keys



 Introduced by Bean (1994) for sequencing problems.



- Introduced by Bean (1994) for sequencing problems.
- Individuals are strings of real-valued numbers (random keys) in the interval [0,1).

$$S = (0.25, 0.19, 0.67, 0.05, 0.89)$$

 $s(1) s(2) s(3) s(4) s(5)$



- Introduced by Bean (1994) for sequencing problems.
- Individuals are strings of real-valued numbers (random keys) in the interval [0,1).
- Sorting random keys results in a sequencing order.

$$S = (0.25, 0.19, 0.67, 0.05, 0.89)$$

 $s(1)$ $s(2)$ $s(3)$ $s(4)$ $s(5)$

$$S' = (0.05, 0.19, 0.25, 0.67, 0.89)$$

 $s(4) s(2) s(1) s(3) s(5)$

Sequence:
$$4 - 2 - 1 - 3 - 5$$



 Mating is done using parametrized uniform
 Crossover (Spears & DeJong, 1990)

```
a = (0.25, 0.19, 0.67, 0.05, 0.89)
b = (0.63, 0.90, 0.76, 0.93, 0.08)
```



- Mating is done using parametrized uniform
 Crossover (Spears & DeJong, 1990)
- For each gene, flip a biased coin to choose which parent passes the allele (key, or value of gene) to the child.

```
a = (0.25, 0.19, 0.67, 0.05, 0.89)
b = (0.63, 0.90, 0.76, 0.93, 0.08)
```



- Mating is done using parametrized uniform
 Crossover (Spears & DeJong, 1990)
- For each gene, flip a biased coin to choose which parent passes the allele (key, or value of gene) to the child.

```
a = (0.25, 0.19, 0.67, 0.05, 0.89)

b = (0.63, 0.90, 0.76, 0.93, 0.08)

c = (
```



- Mating is done using parametrized uniform
 Crossover (Spears & DeJong, 1990)
- For each gene, flip a biased coin to choose which parent passes the allele (key, or value of gene) to the child.

```
a = (0.25, 0.19, 0.67, 0.05, 0.89)

b = (0.63, 0.90, 0.76, 0.93, 0.08)

c = (0.25)
```



- Mating is done using parametrized uniform
 Crossover (Spears & DeJong, 1990)
- For each gene, flip a biased coin to choose which parent passes the allele (key, or value of gene) to the child.

```
a = (0.25, 0.19, 0.67, 0.05, 0.89)

b = (0.63, 0.90, 0.76, 0.93, 0.08)

c = (0.25, 0.90)
```



- Mating is done using parametrized uniform
 Crossover (Spears & DeJong, 1990)
- For each gene, flip a biased coin to choose which parent passes the allele (key, or value of gene) to the child.

```
a = (0.25, 0.19, 0.67, 0.05, 0.89)

b = (0.63, 0.90, 0.76, 0.93, 0.08)

c = (0.25, 0.90, 0.76)
```



- Mating is done using parametrized uniform
 Crossover (Spears & DeJong, 1990)
- For each gene, flip a biased coin to choose which parent passes the allele (key, or value of gene) to the child.

```
a = (0.25, 0.19, 0.67, 0.05, 0.89)

b = (0.63, 0.90, 0.76, 0.93, 0.08)

c = (0.25, 0.90, 0.76, 0.05)
```



- Mating is done using parametrized uniform
 Crossover (Spears & DeJong, 1990)
- For each gene, flip a biased coin to choose which parent passes the allele (key, or value of gene) to the child.

```
a = (0.25, 0.19, 0.67, 0.05, 0.89)

b = (0.63, 0.90, 0.76, 0.93, 0.08)

c = (0.25, 0.90, 0.76, 0.05, 0.89)
```



- Mating is done using parametrized uniform
 Crossover (Spears & DeJong, 1990)
- For each gene, flip a biased coin to choose which parent passes the allele (key, or value of gene) to the child.

```
a = (0.25, 0.19, 0.67, 0.05, 0.89)

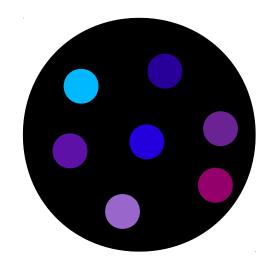
b = (0.63, 0.90, 0.76, 0.93, 0.08)

c = (0.25, 0.90, 0.76, 0.05, 0.89)
```

If every random-key array corresponds to a feasible solution: Mating always produces feasible offspring.

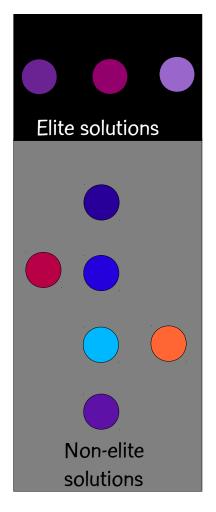


Initial population is made up of P random-key vectors, each with N keys, each having a value generated uniformly at random in the interval [0,1).



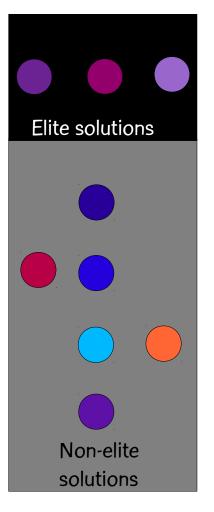


At the K-th generation, compute the cost of each solution ...



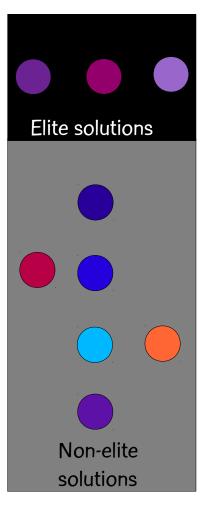


At the K-th generation, compute the cost of each solution and partition the solutions into two sets:



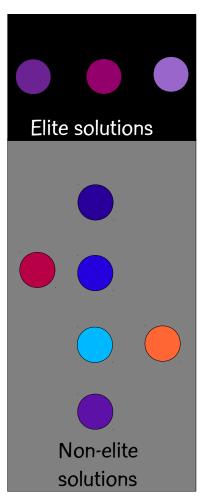


At the K-th generation, compute the cost of each solution and partition the solutions into two sets: elite solutions and non-elite solutions.





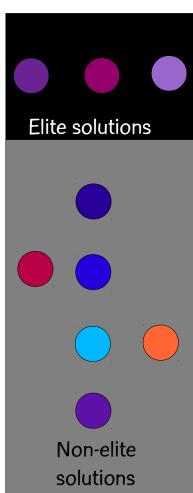
At the K-th generation, compute the cost of each solution and partition the solutions into two sets: elite solutions and non-elite solutions. Elite set should be smaller of the two sets and contain best solutions.





Evolutionary dynamics

Population K

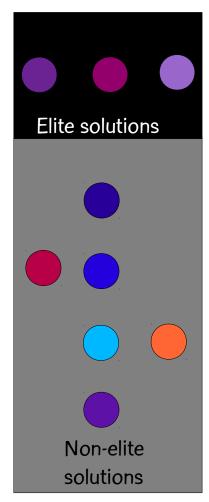


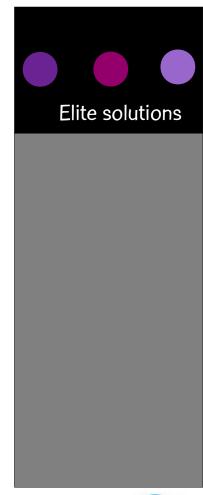


Evolutionary dynamics

Copy elite solutions from population
 K to population K+1



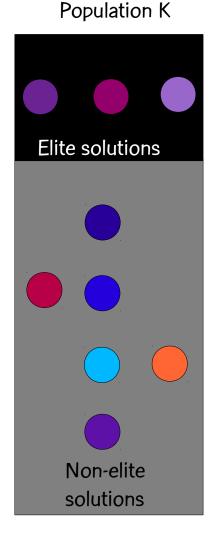


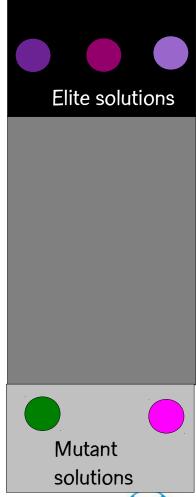




Evolutionary dynamics

- Copy elite solutions from population
 K to population K+1
- Add R random solutions (mutants)
 to population K+1

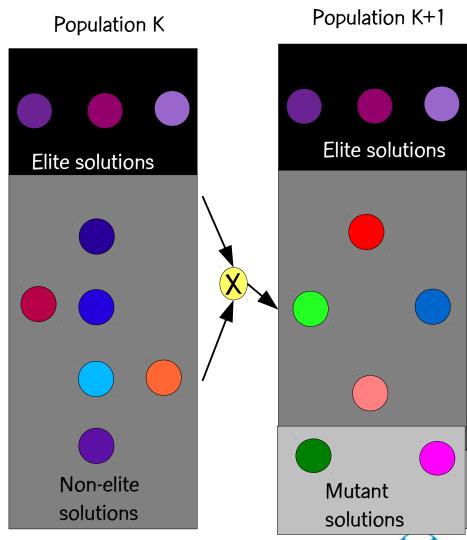






Evolutionary dynamics

- Copy elite solutions from population
 K to population K+1
- Add R random solutions (mutants)
 to population K+1
- While K+1-th population < P
 - RANDOM-KEY GA: Use any two solutions in population K to produce child in population K+1. Mates are chosen at random.



Biased random key genetic algorithm

• A biased random key genetic algorithm (BRKGA) is a random key genetic algorithm (RKGA).



Biased random key genetic algorithm

- A biased random key genetic algorithm (BRKGA)
 is a random key genetic algorithm (RKGA).
- BRKGA and RKGA differ in how mates are chosen for crossover and how parametrized uniform crossover is applied.



RKGA BRKGA

both parents chosen at random from entire population



RKGA

both parents chosen at random from entire population

BRKGA

both parents chosen at random but one parent chosen from population of elite solutions



RKGA

both parents chosen at random from entire population

BRKGA

both parents chosen at random but one parent chosen from population of elite solutions

either parent can be parent A in parametrized uniform crossover



RKGA

both parents chosen at random from entire population

BRKGA

both parents chosen at random but one parent chosen from population of elite solutions

either parent can be parent A in parametrized uniform crossover

best fit parent is parent A in parametrized uniform crossover



Biased random key GA

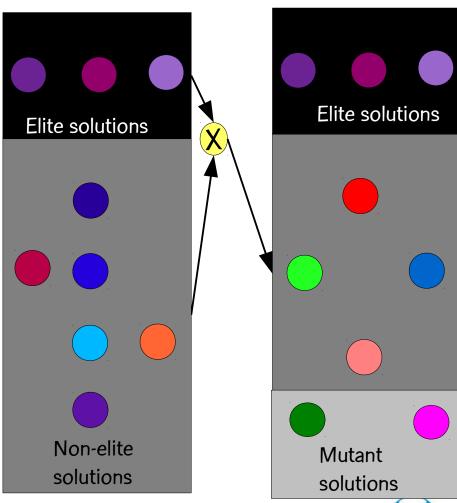
BRKGA: Probability child inherits key of elite parent > 0.5 Po

Population K

Population K+1

Evolutionary dynamics

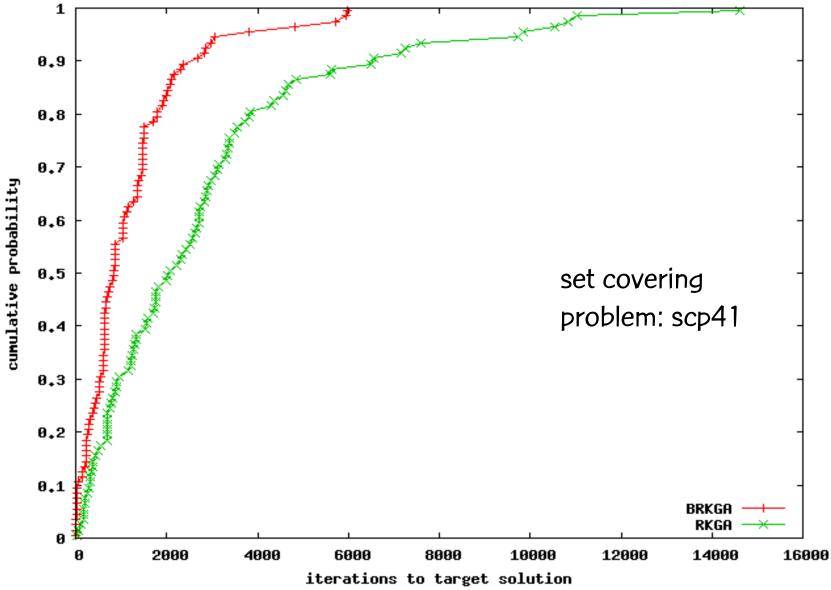
- Copy elite solutions from population
 K to population K+1
- Add R random solutions (mutants)
 to population K+1
- While K+1-th population < P
 - RANDOM-KEY GA: Use any two solutions in population K to produce child in population K+1. Mates are chosen at random.
 - BIASED RANDOM-KEY GA: Mate elite solution with other solution of population K to produce child in population K+1. Mates are chosen at random.

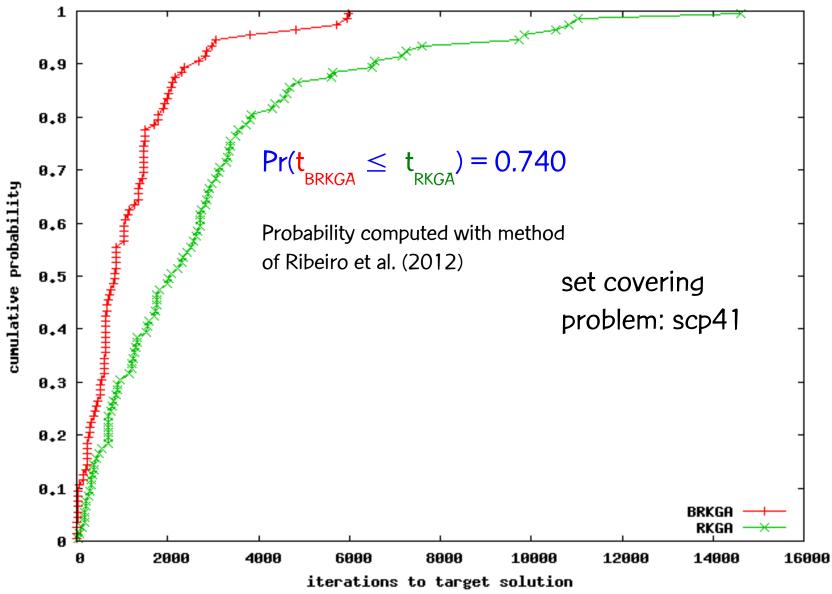


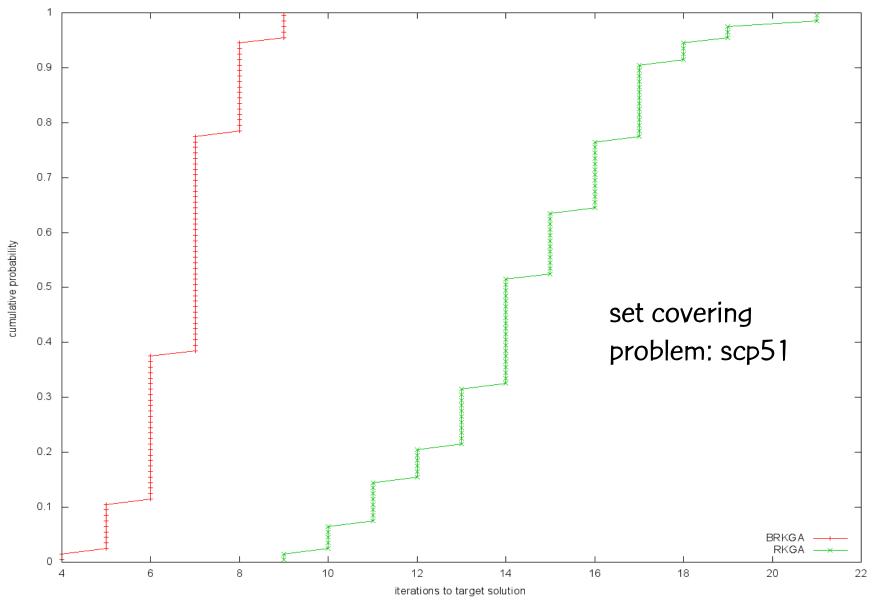
Paper comparing BRKGA and Bean's Method

Gonçalves, R., and Toso, "Biased and unbiased random-key genetic algorithms: An experimental analysis", AT&T Labs Research Technical Report, Florham Park, December 2012.

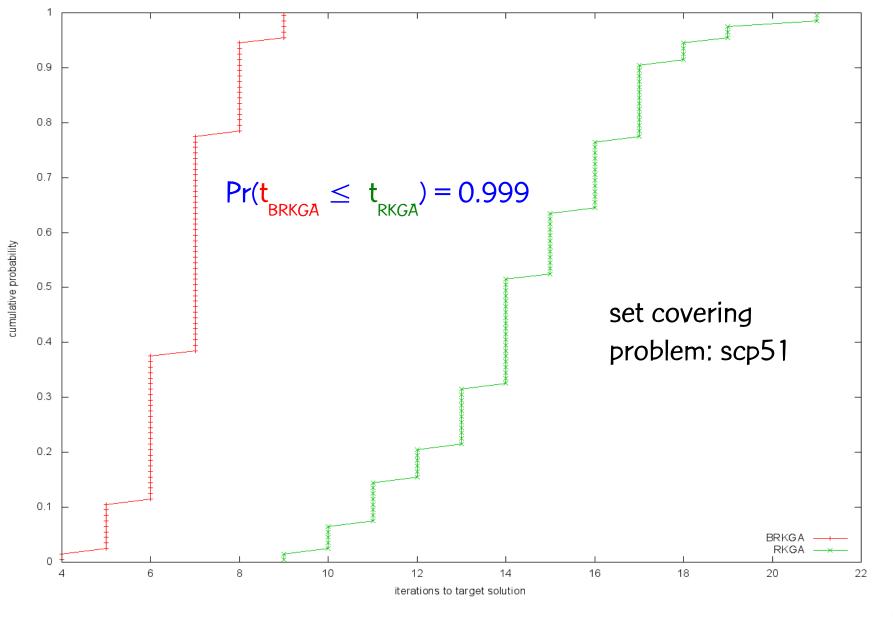




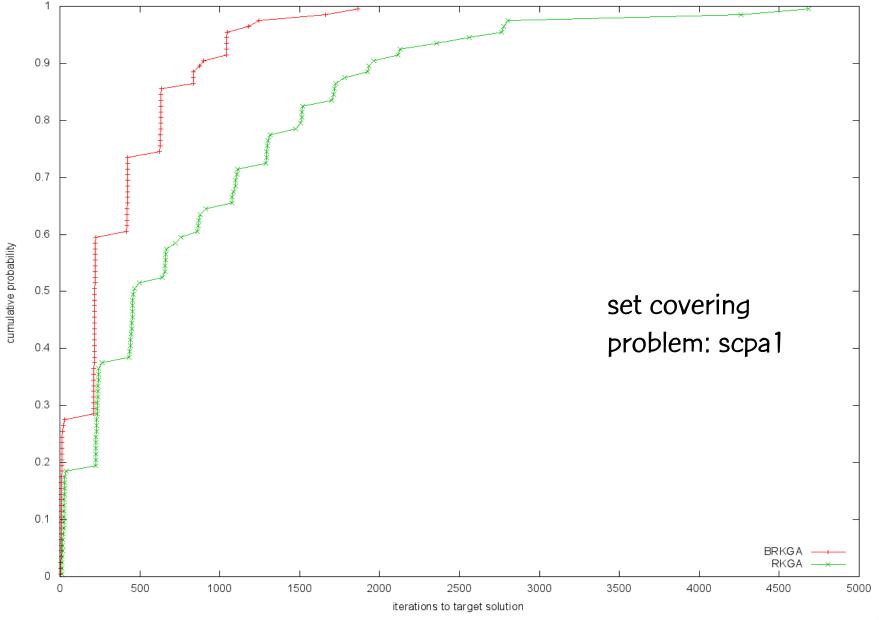




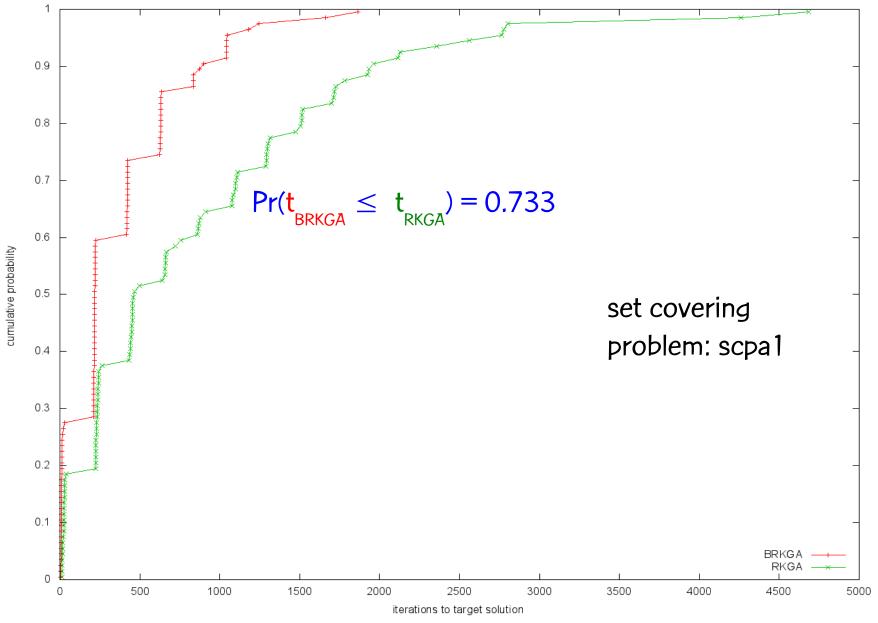




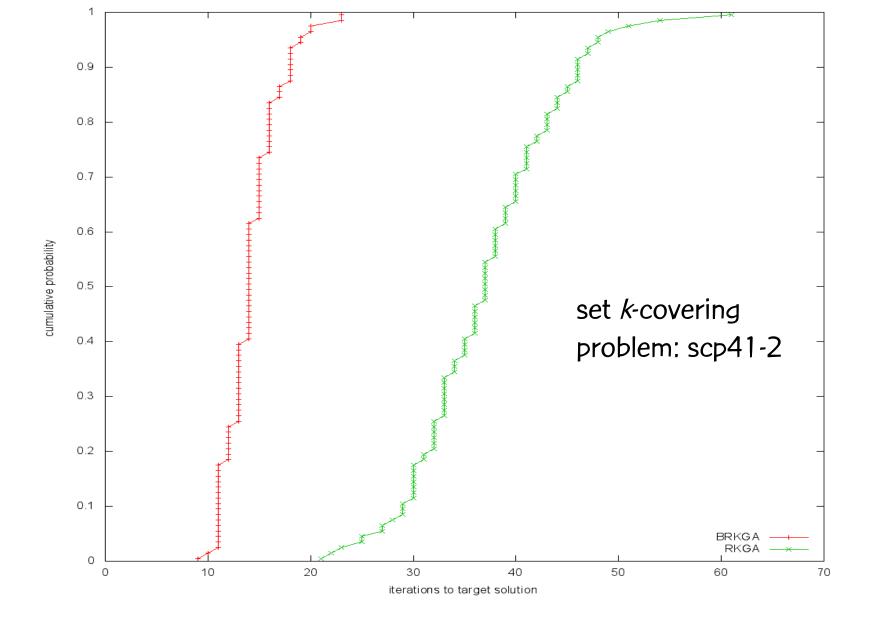




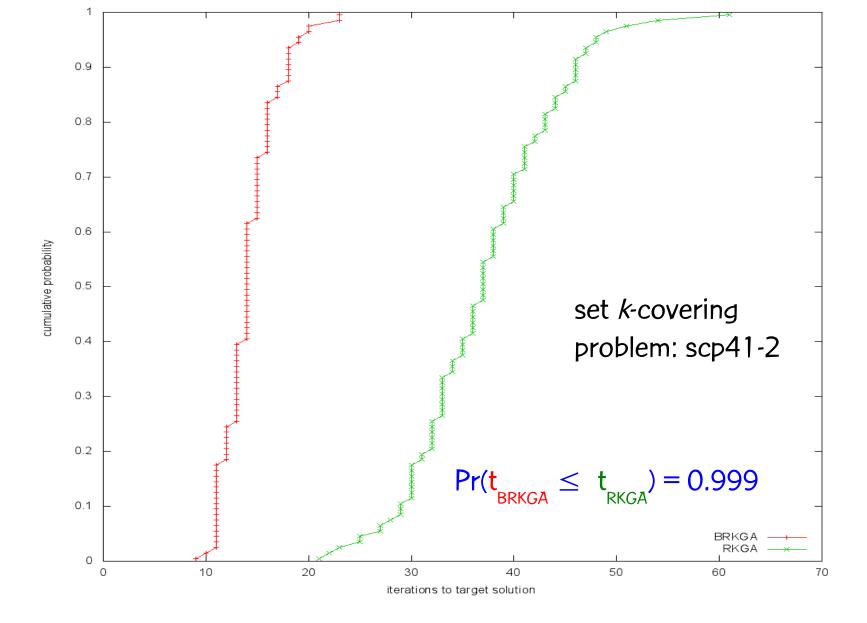




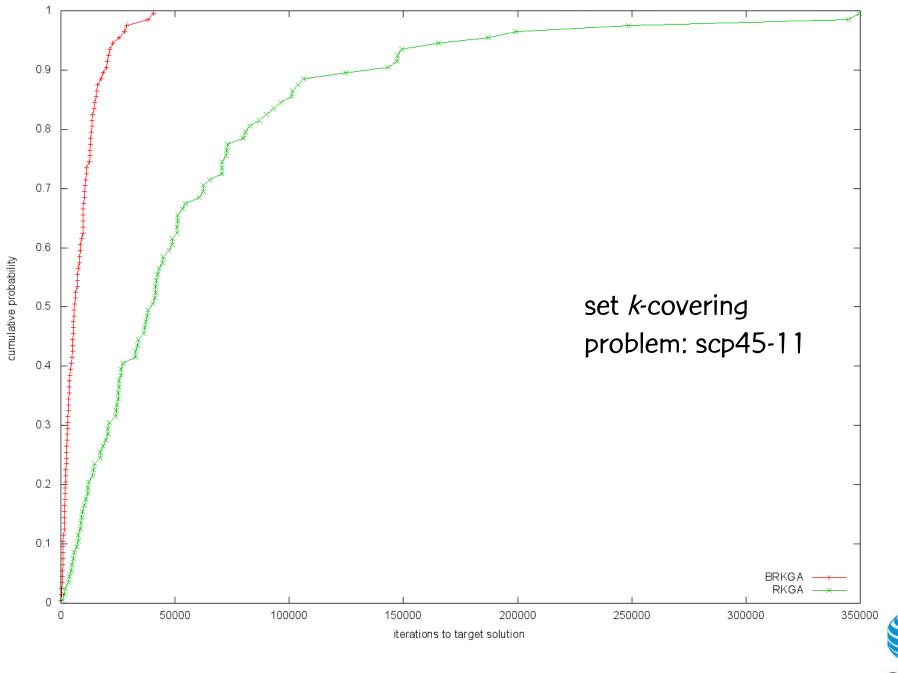


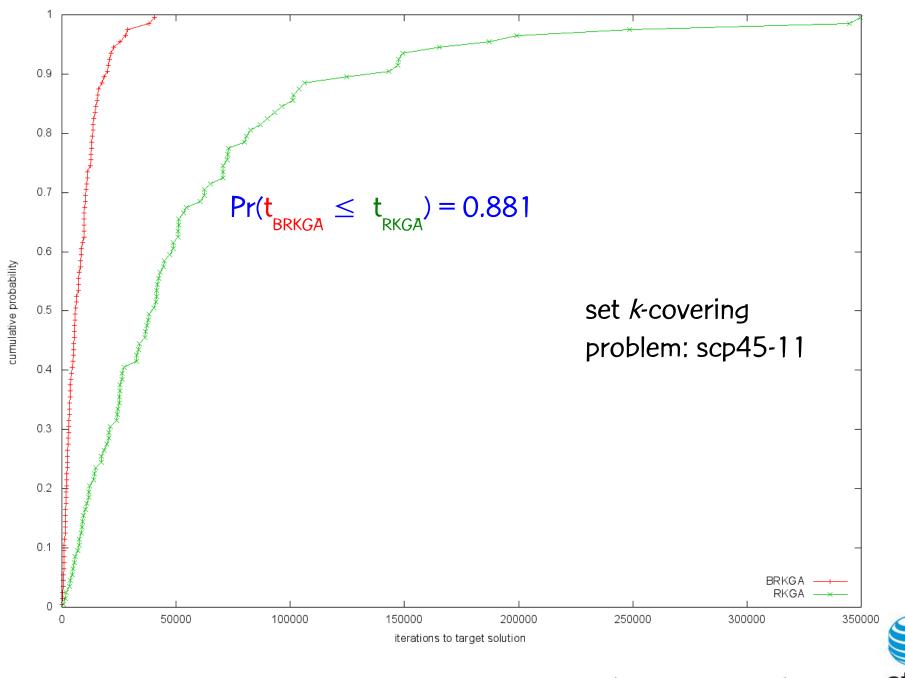


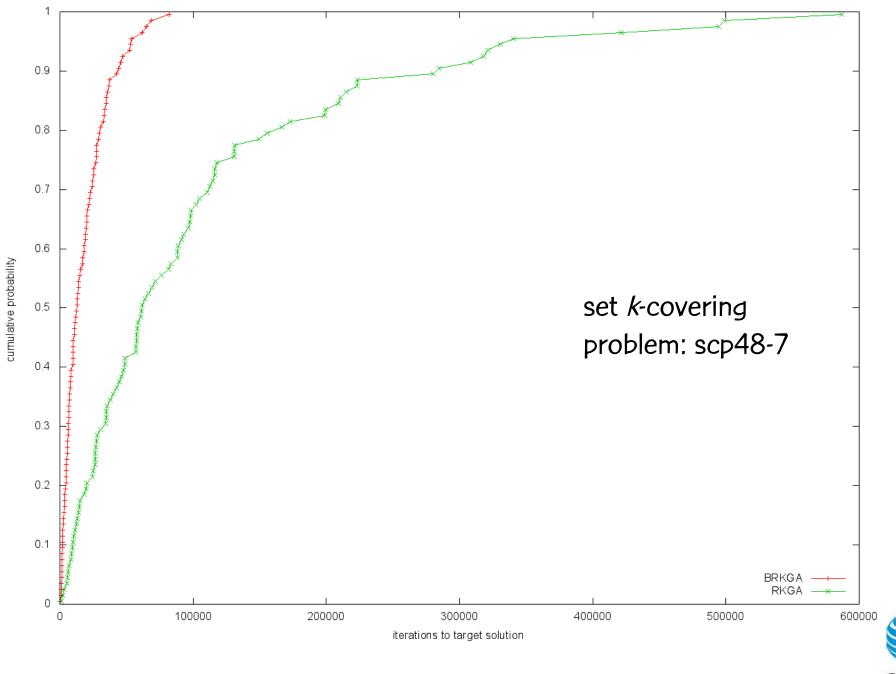


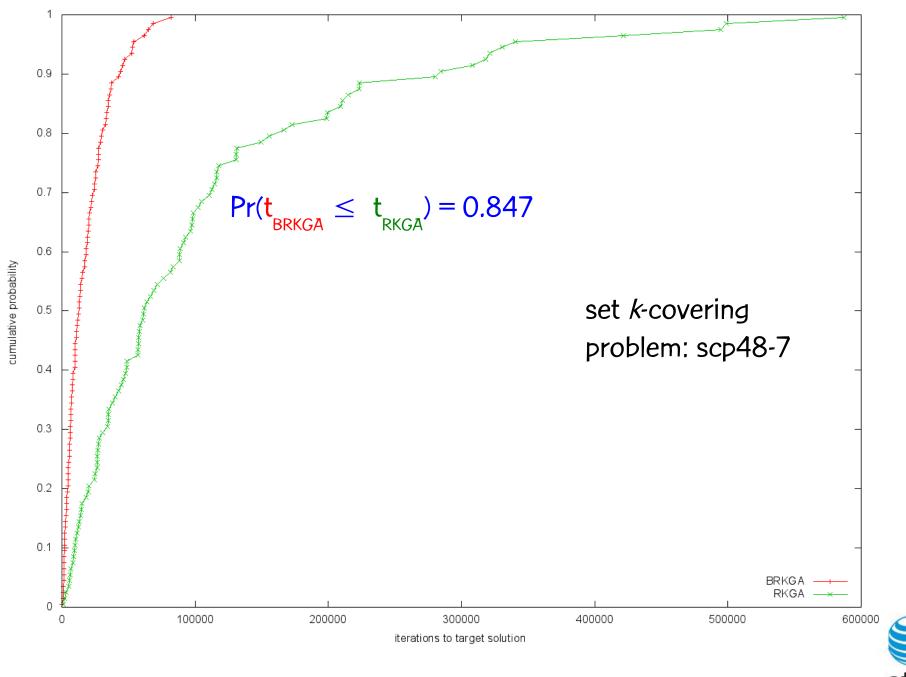












 Random method: keys are randomly generated so solutions are always vectors of random keys



- Random method: keys are randomly generated so solutions are always vectors of random keys
- Elitist strategy: best solutions are passed without change from one generation to the next (incumbent is kept)



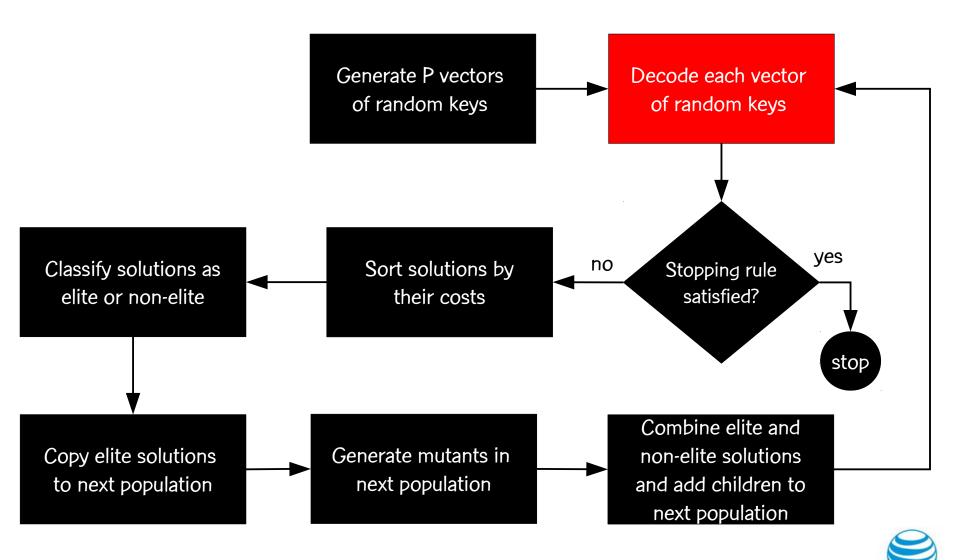
- Random method: keys are randomly generated so solutions are always vectors of random keys
- Elitist strategy: best solutions are passed without change from one generation to the next (incumbent is kept)
- Child inherits more characteristics of elite parent:
 one parent is always selected (with replacement) from the
 small elite set and probability that child inherits key of elite
 parent > 0.5 Not so in the RKGA of Bean.



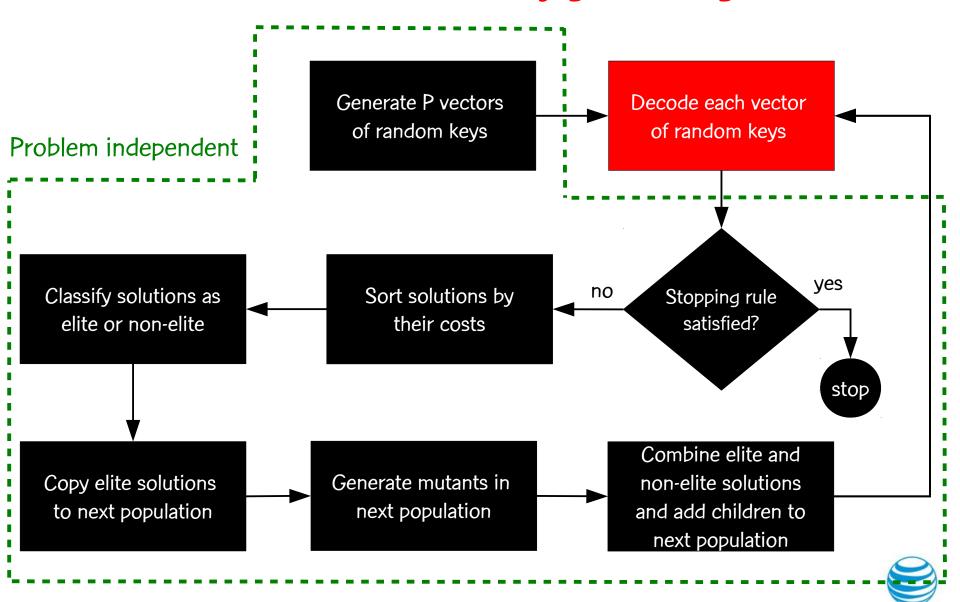
- Random method: keys are randomly generated so solutions are always vectors of random keys
- Elitist strategy: best solutions are passed without change from one generation to the next (incumbent is kept)
- Child inherits more characteristics of elite parent:
 one parent is always selected (with replacement) from the
 small elite set and probability that child inherits key of elite
 parent > 0.5 Not so in the RKGA of Bean.
- No mutation in crossover: mutants are used instead (they play same role as mutation in GAs ... help escape local optima)



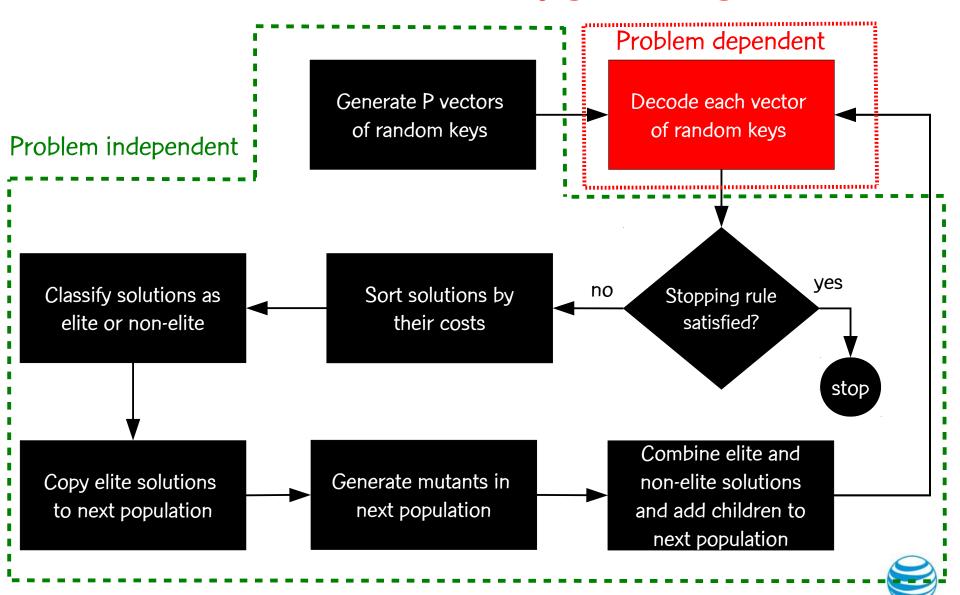
Framework for biased random-key genetic algorithms



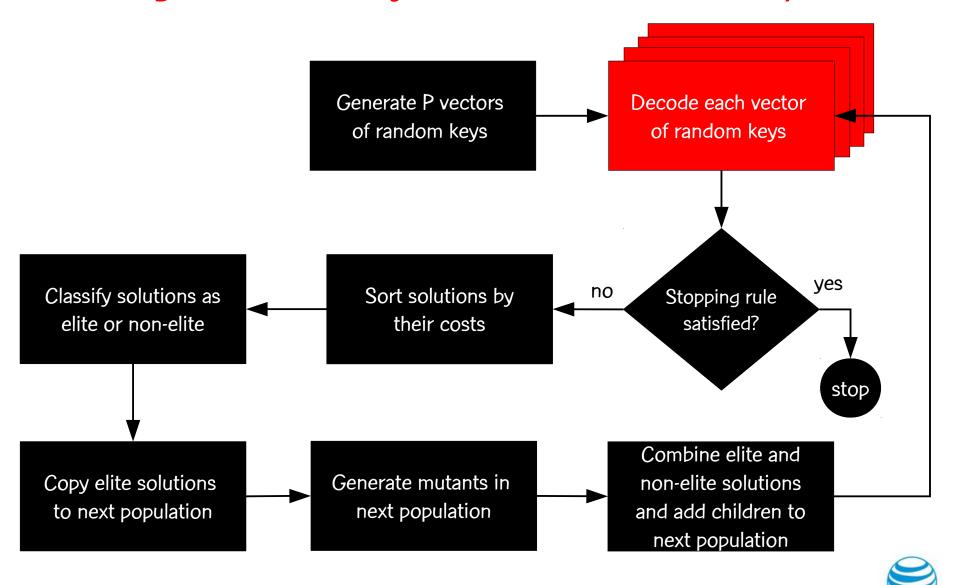
Framework for biased random-key genetic algorithms



Framework for biased random-key genetic algorithms



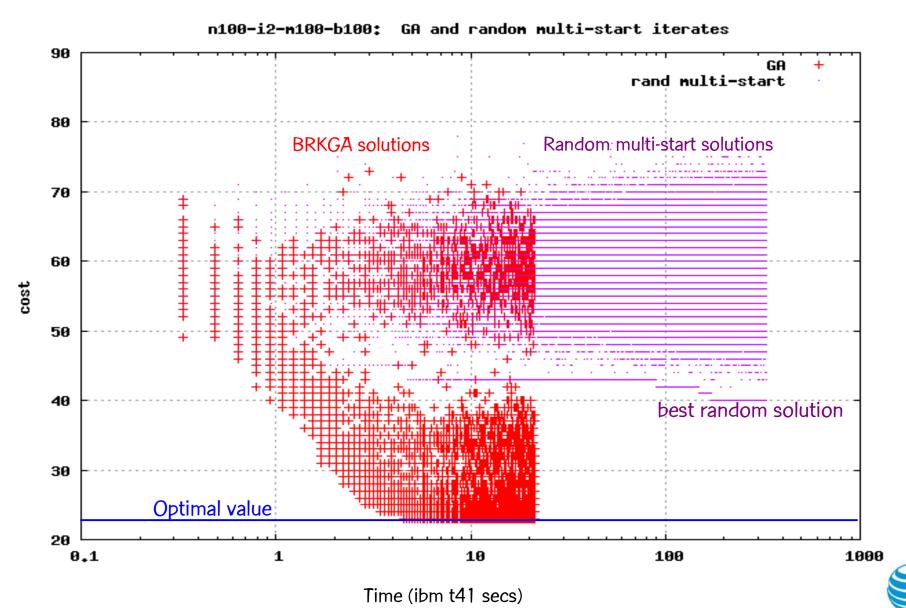
Decoding of random key vectors can be done in parallel



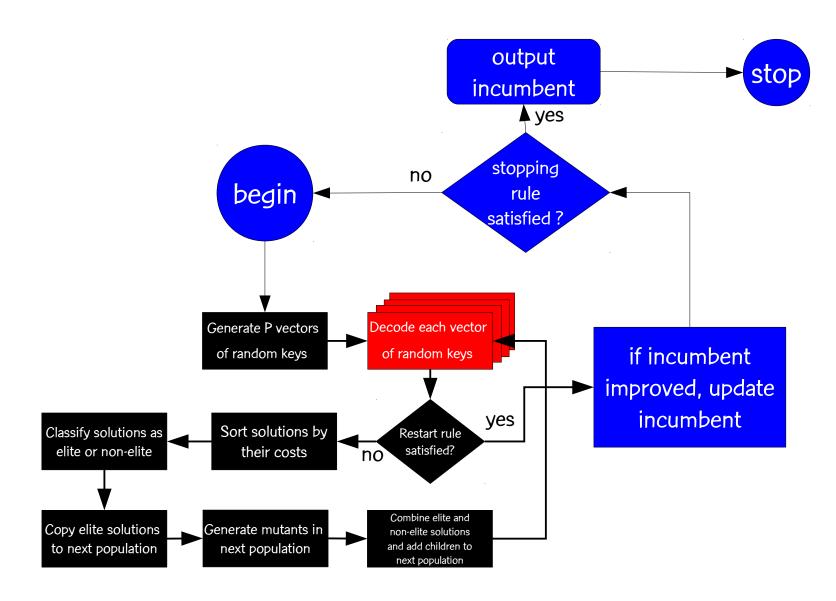
Is a BRKGA any different from applying the decoder to random keys?

- Simulate a random multi-start decoding method with a BRKGA by setting size of elite partition to 1 and number of mutants to P-1
- Each iteration, best solution is maintained in elite set and P—1 random key vectors are generated as mutants ... no mating is done since population already has P individuals

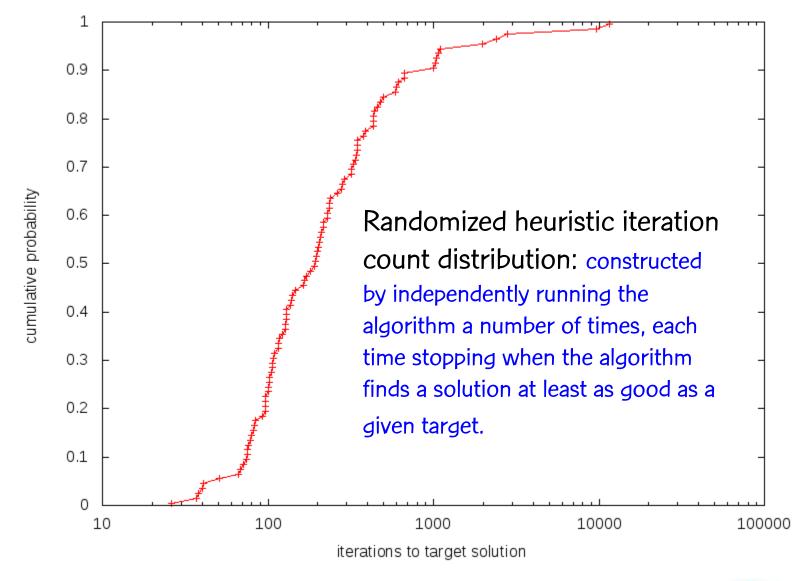




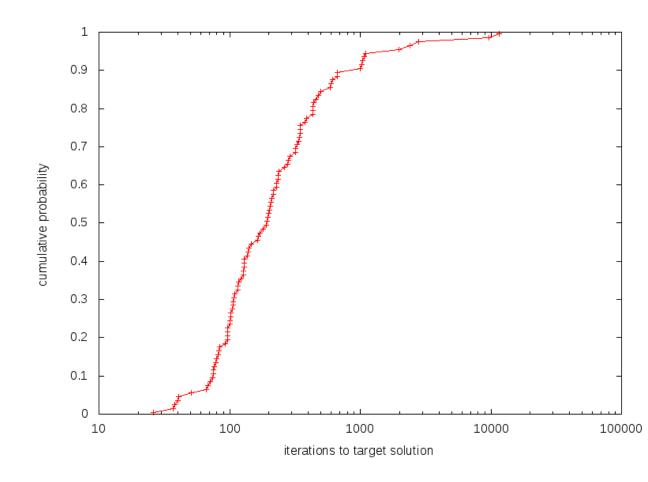
BRKGA in multi-start strategy





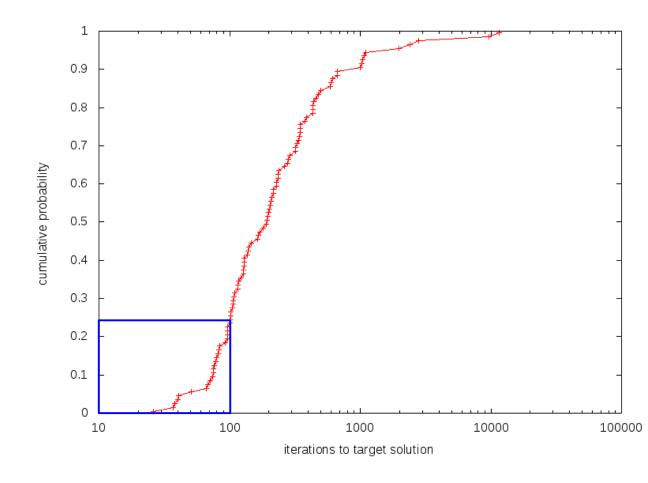






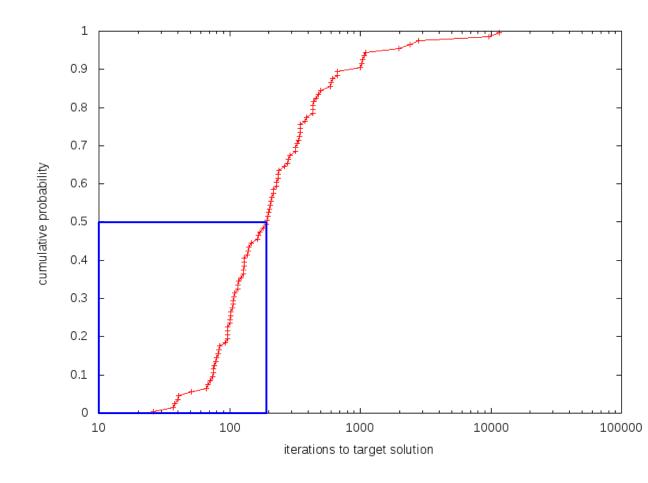
In most of the independent runs, the algorithm finds the target solution in relatively few iterations:





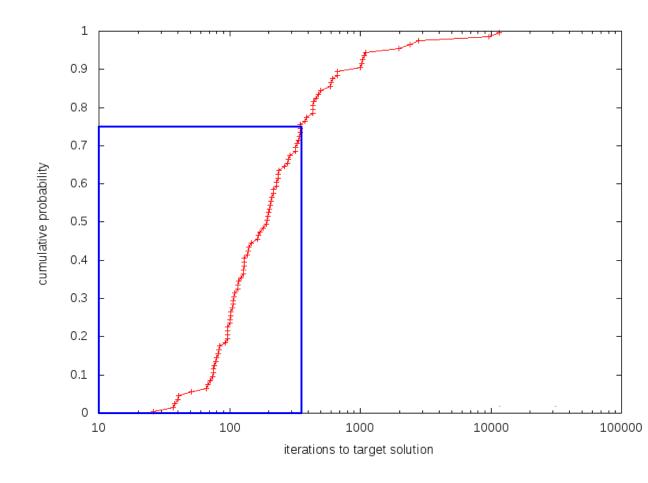
In most of the independent runs, the algorithm finds the target solution in relatively few iterations: 25% of the runs take fewer than 101 iterations





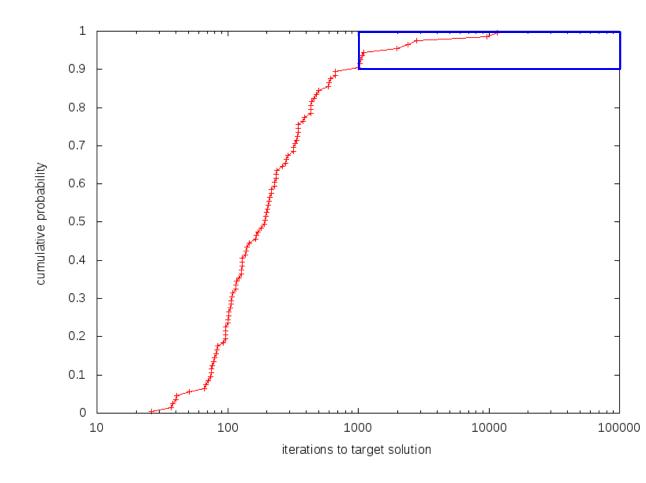
In most of the independent runs, the algorithm finds the target solution in relatively few iterations: 50% of the runs take fewer than 192 iterations





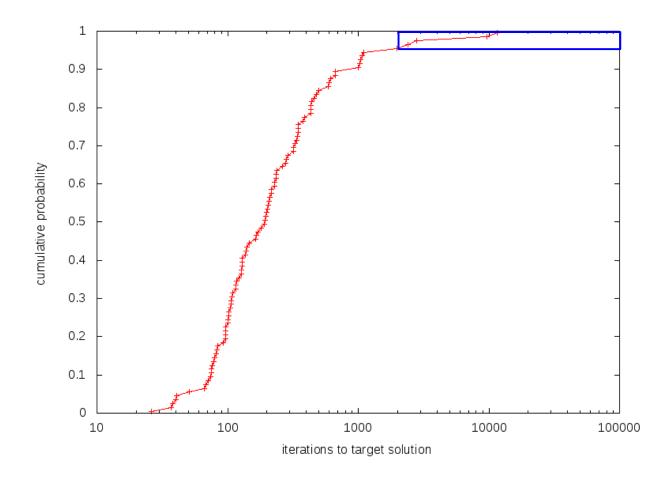
In most of the independent runs, the algorithm finds the target solution in relatively few iterations: 75% of the runs take fewer than 345 iterations





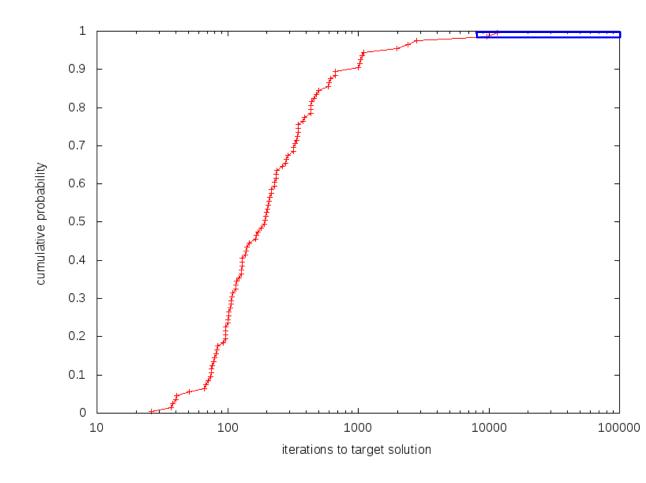
However, some runs take much longer: 10% of the runs take over 1000 iterations





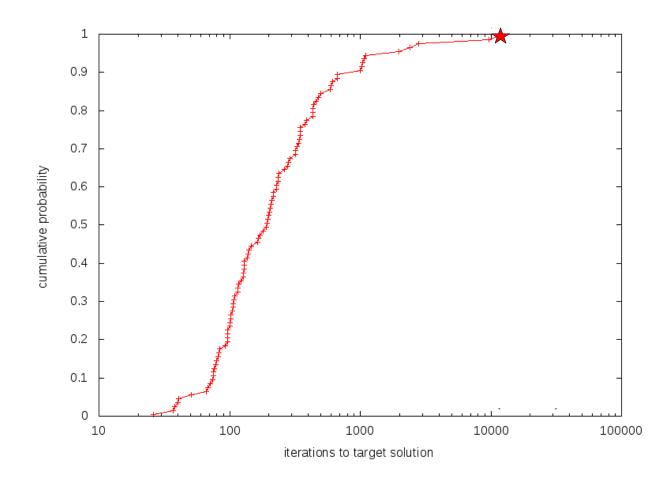
However, some runs take much longer: 5% of the runs take over 2000 iterations





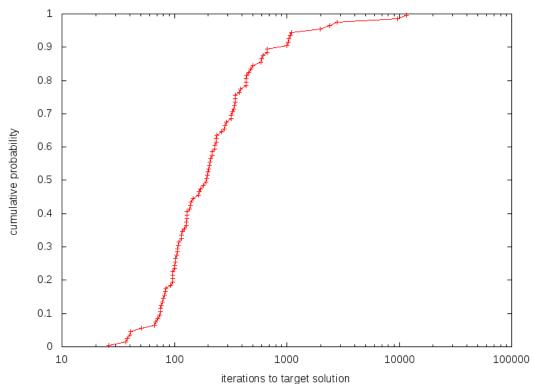
However, some runs take much longer: 2% of the runs take over 9715 iterations





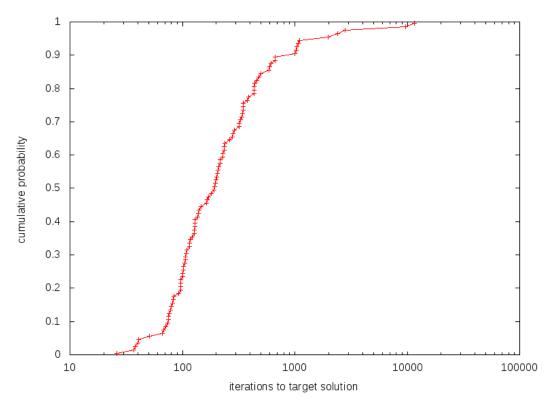
However, some runs take much longer: the longest run took 11607 iterations





Probability that algorithm will take over 345 iterations: 25% = 1/4



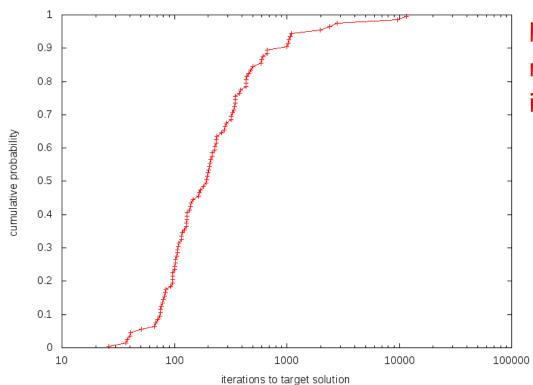


Probability that algorithm will take over 345 iterations: 25% = 1/4

By restarting algorithm after 345 iterations, probability that new run will take over 690 iterations: 25% = 1/4

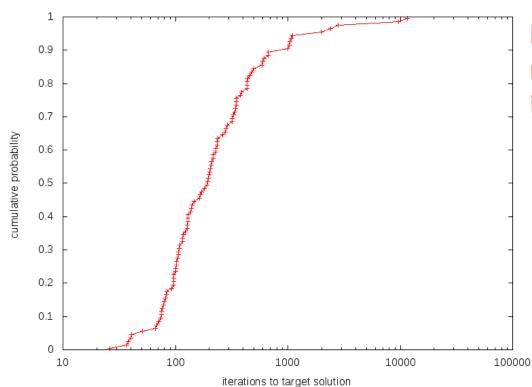
Probability that algorithm with restart will take over 690 iterations: probability of taking over 345 X probability of taking over 690 iterations given it took over 345 = $\frac{1}{4} \times \frac{1}{4} = \frac{1}{4^2}$





Probability that algorithm will still be running after K periods of 345 iterations: 1/4^K

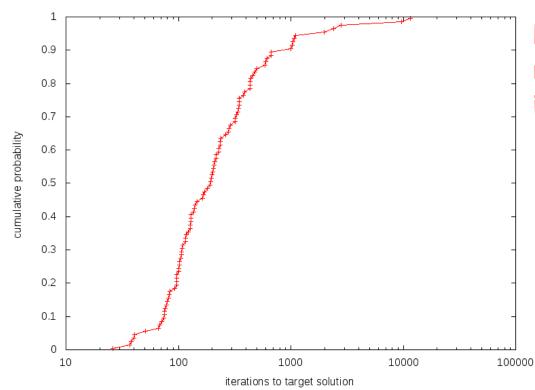




Probability that algorithm will still be running after K periods of 345 iterations: 1/4^K

For example, probability that algorithm with restart will still be running after 1725 iterations (5 periods of 345 iterations): $1/4^5 \approx 0.0977\%$





Probability that algorithm will still be running after K periods of 345 iterations: 1/4^K

For example, probability that algorithm with restart will still be running after 1725 iterations (5 periods of 345 iterations): $1/4^5 \approx 0.0977\%$

This is much less than the 5% probability that the algorithm without restart will take over 2000 iterations.



Restart strategies

- First proposed by Luby et al. (1993)
- They define a restart strategy as a finite sequence of time intervals $S = \{\tau_1, \tau_2, \tau_3, ...\}$ which define epochs $\tau_1, \tau_1 + \tau_2, \tau_1 + \tau_2 + \tau_3, ...$ when the algorithm is restarted from scratch.
- Luby et al. (1993) prove that the optimal restart strategy uses $\tau_1 = \tau_2 = \tau_3 = \cdots = \tau^*$, where τ^* is a constant.



Restart strategies

- Luby et al. (1993)
- Kautz et al. (2002)
- Palubeckis (2004)
- Sergienko et al. (2004)
- Nowicki & Smutnicki (2005)
- D'Apuzzo et al. (2006)
- Shylo et al. (2011a)
- Shylo et al. (2011b)
- Resende & Ribeiro (2011)



Restart strategy for BRKGA

- Recall the restart strategy of Luby et al. where equal time intervals $\tau_1 = \tau_2 = \tau_3 = \dots = \tau^*$ pass between restarts.
- Strategy requires τ* as input.
- Since we have no prior information as to the runtime distribution of the heuristic, we run the risk of:
 - choosing τ* too small: restart variant may take long to converge
 - choosing τ* too big: restart variant may become like norestart variant



Restart strategy for BRKGA

- We conjecture that number of iterations between improvement of the incumbent (best so far) solution varies less w.r.t. heuristic/ instance/ target than run times.
- We propose the following restart strategy: Keep track of the last generation when the incumbent improved and restart BRKGA if K generations have gone by without improvement.
- We call this strategy restart(K)

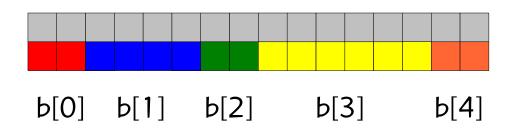


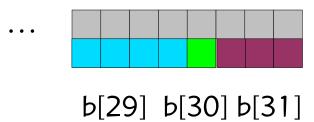
Given an unordered sequence of 1024 integers p[0], p[1], ..., p[1023]





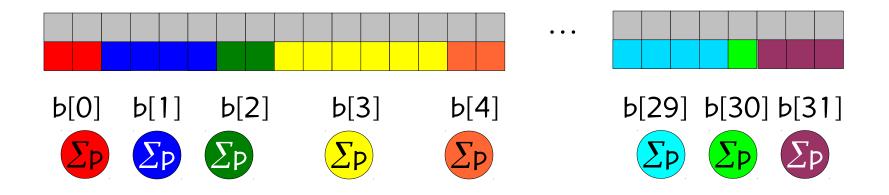
Place consecutive numbers in 32 buckets b[0], b[1], ..., b[31]





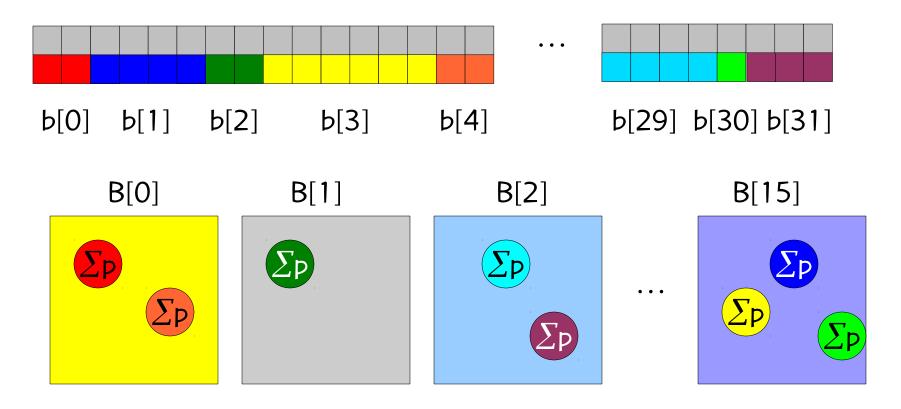


Add the numbers in each bucket b[0], b[1], ..., b[31]



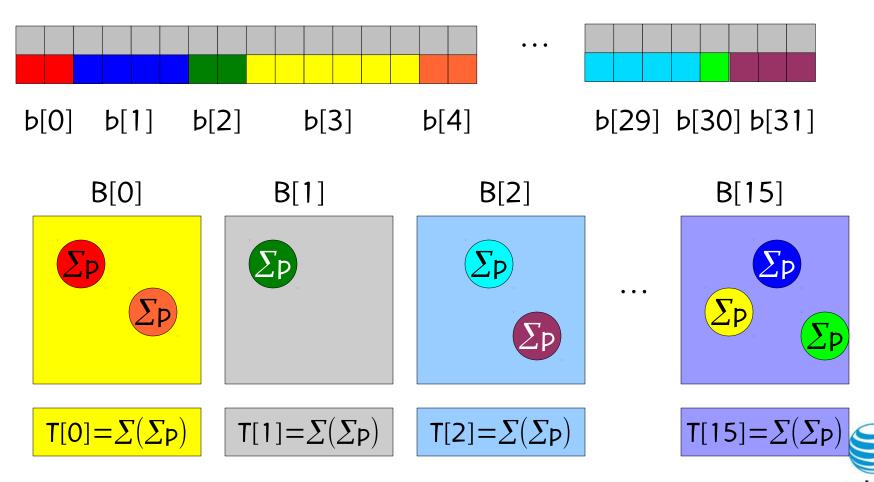


Place the buckets in 16 bins B[0], B[1], ..., B[15]

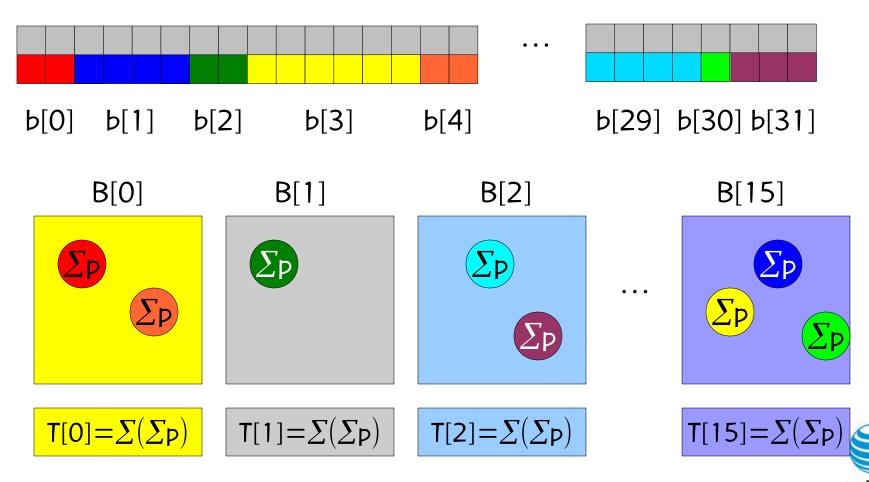




Add up the numbers in each bin B[0], B[1], ..., B[15]



OBJECTIVE: Minimize { Maximum (T[0], T[1], ..., T[15]) }



Encoding

Decoding

x[1], x[2], ..., x[32] are used to define break points for buckets

x[32+1], x[32+2], ..., x[32+16] are used to determine to which bins the buckets are assigned



Encoding

$$X = [x[1], x[2], ..., x[32] x[32+1], x[32+2], ..., x[32+16]]$$

Decoding

x[1], x[2], ..., x[32] are used to define break points for buckets

Size of bucket i = floor
$$(1024 \times x[i]/(x[1]+x[2]+\cdots+x[32]))$$
, i=1,...,15

Size of bucket 16 = 1024 - sum of sizes of first 15 buckets



Encoding

Decoding

x[1], x[2], ..., x[32] are used to define break points for buckets

x[32+1], x[32+2], ..., x[32+16] are used to determine to which bins the buckets are assigned

Bin that bucket i is assigned to = floor $(16 \times x[32+i]) + 1$

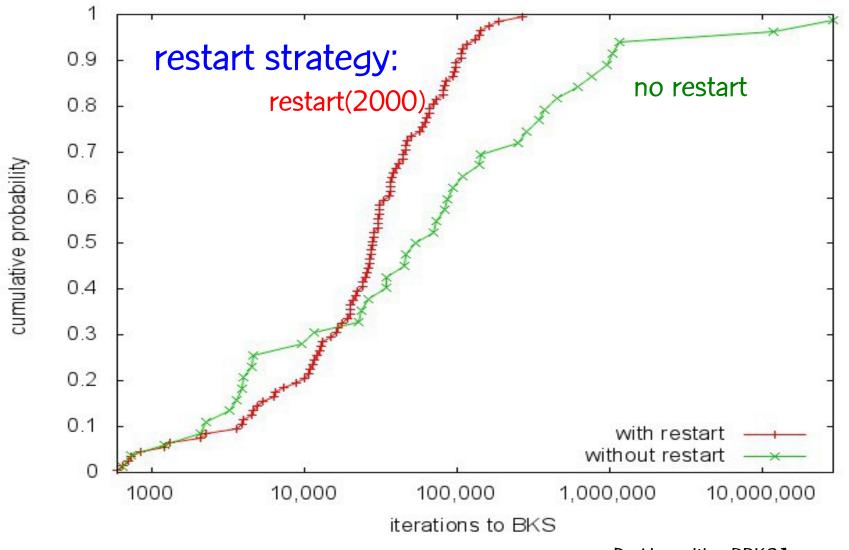


Decoding (Local search phase)

- while (there exists a bucket in the most loaded bin that can be moved to another bin and not increase the maximum load) then
 - move that bucket to that bin
- end while

Make necessary chromosome adjustments to last 16 random keys of vector of random keys to reflect changes made in local search phase: Add or subtract an integer value from chromosome of bucket that moved to new bin.







Specifying a BRKGA



 Encoding is always done the same way, i.e. with a vector of N random-keys (parameter N must be specified)



- Encoding is always done the same way, i.e. with a vector of N random-keys (parameter N must be specified)
- Decoder that takes as input a vector of N random-keys and outputs the corresponding solution of the combinatorial optimization problem and its cost (this is usually a heuristic)



- Encoding is always done the same way, i.e. with a vector of N random-keys (parameter N must be specified)
- Decoder that takes as input a vector of N random-keys and outputs the corresponding solution of the combinatorial optimization problem and its cost (this is usually a heuristic)
- Parameters



- Size of population
- Parallel population parameters
- Size of elite partition
- Size of mutant set
- Child inheritance probability
- Restart strategy parameter
- Stopping criterion



- Size of population: a function of N, say N or 2N
- Parallel population parameters
- Size of elite partition
- Size of mutant set
- Child inheritance probability
- Restart strategy parameter
- Stopping criterion



- Size of population: a function of N, say N or 2N
- Parallel population parameters: say, p = 3, v = 2, and x = 200
- Size of elite partition
- Size of mutant set
- Child inheritance probability
- Restart strategy parameter
- Stopping criterion



- Size of population: a function of N, say N or 2N
- Parallel population parameters: say, p = 3, v = 2, and x = 200
- Size of elite partition: 15-25% of population
- Size of mutant set
- Child inheritance probability
- Restart strategy parameter
- Stopping criterion



- Size of population: a function of N, say N or 2N
- Parallel population parameters: say, p = 3, v = 2, and x = 200
- Size of elite partition: 15-25% of population
- Size of mutant set: 5-15% of population
- Child inheritance probability
- Restart strategy parameter
- Stopping criterion



- Size of population: a function of N, say N or 2N
- Parallel population parameters: say, p = 3, v = 2, and x = 200
- Size of elite partition: 15-25% of population
- Size of mutant set: 5-15% of population
- Child inheritance probability: > 0.5, say 0.7
- Restart strategy parameter
- Stopping criterion



- Size of population: a function of N, say N or 2N
- Parallel population parameters: say, p = 3, v = 2, and x = 200
- Size of elite partition: 15-25% of population
- Size of mutant set: 5-15% of population
- Child inheritance probability: > 0.5, say 0.7
- Restart strategy parameter: a function of N, say 2N or 10N
- Stopping criterion



- Size of population: a function of N, say N or 2N
- Parallel population parameters: say, p = 3, v = 2, and x = 200
- Size of elite partition: 15-25% of population
- Size of mutant set: 5-15% of population
- Child inheritance probability: > 0.5, say 0.7
- Restart strategy parameter: a function of N, say 2N or 10N
- Stopping criterion: e.g. time, # generations, solution quality,# generations without improvement



 Efficient and easy-to-use object oriented application programming interface (API) for the algorithmic framework of BRKGA.



- Efficient and easy-to-use object oriented application programming interface (API) for the algorithmic framework of BRKGA.
- Cross-platform library handles large portion of problem independent modules that make up the framework, e.g.
 - population management
 - evolutionary dynamics



- Efficient and easy-to-use object oriented application programming interface (API) for the algorithmic framework of BRKGA.
- Cross-platform library handles large portion of problem independent modules that make up the framework, e.g.
 - population management
 - evolutionary dynamics
- Implemented in C++ and may benefit from shared-memory parallelism if available.



- Efficient and easy-to-use object oriented application programming interface (API) for the algorithmic framework of BRKGA.
- Cross-platform library handles large portion of problem independent modules that make up the framework, e.g.
 - population management
 - evolutionary dynamics
- Implemented in C++ and may benefit from shared-memory parallelism if available.
- User only needs to implement problem-dependent decoder.



Paper: Rodrigo F. Toso and M.G.C.R., "A C++
Application Programming Interface for
Biased Random-Key Genetic Algorithms,"
AT&T Labs Technical Report, Florham Park, August 2011.

Software: http://www.research.att.com/~mgcr/src/brkgaAPI



An example BRKGA: Packing weighted rectangles



Reference



J.F. Gonçalves and M.G.C.R., "A parallel multipopulation genetic algorithm for a constrained two-dimensional orthogonal packing problem," Journal of Combinatorial Optimization, vol. 22, pp. 180-201, 2011.

Tech report:

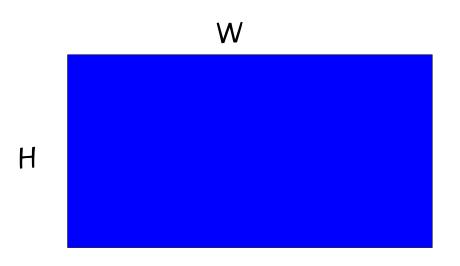
http://www.research.att.com/~mgcr/doc/pack2d.pdf



 Given a large planar stock rectangle (W, H) of width W and height H;

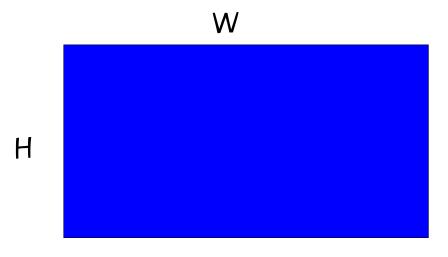


 Given a large planar stock rectangle (W, H) of width W and height H;



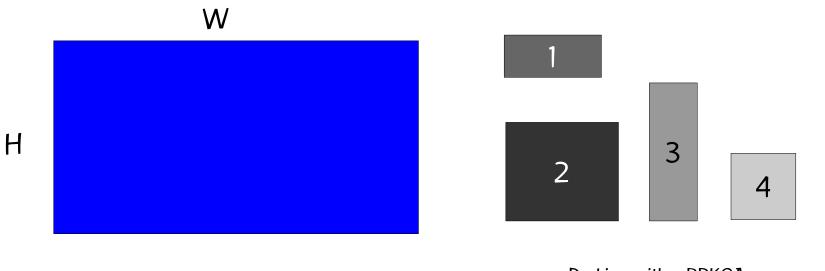


- Given a large planar stock rectangle (W, H) of width W and height H;
- Given N smaller rectangle types (w[i], h[i]),
 i = 1,...,N, each of width w[i], height h[i], and value v[i];



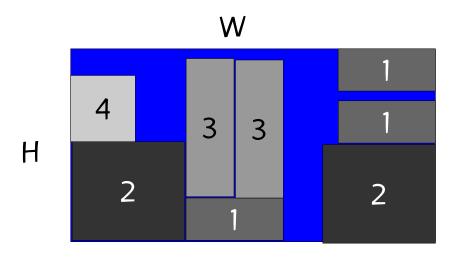


- Given a large planar stock rectangle (W, H) of width W and height H;
- Given N smaller rectangle types (w[i], h[i]),
 i = 1,...,N, each of width w[i], height h[i], and value v[i];





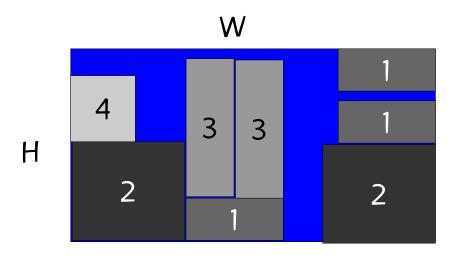
 r[i] rectangles of type i = 1, ..., N are to be packed in the large rectangle without overlap and such that their edges are parallel to the edges of the large rectangle;





- r[i] rectangles of type i = 1, ..., N are to be packed in the large rectangle without overlap and such that their edges are parallel to the edges of the large rectangle;
- For i = 1, ..., N, we require that:

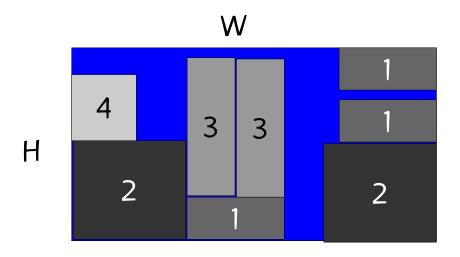
$$0 \le P[i] \le r[i] \le Q[i]$$





- r[i] rectangles of type i = 1, ..., N are to be packed in the large rectangle without overlap and such that their edges are parallel to the edges of the large rectangle;
- For i = 1, ..., N, we require that:

$$0 \le P[i] \le r[i] \le Q[i]$$

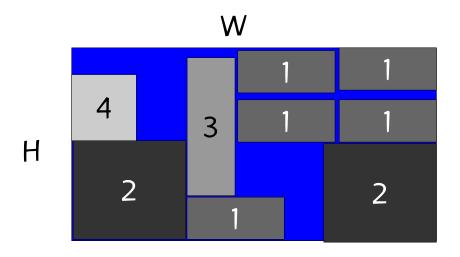


Suppose $5 \le r[1] \le 12$



- r[i] rectangles of type i = 1, ..., N are to be packed in the large rectangle without overlap and such that their edges are parallel to the edges of the large rectangle;
- For i = 1, ..., N, we require that:

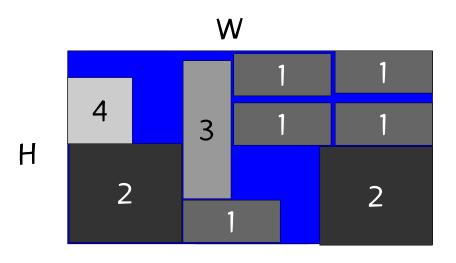
$$0 \le P[i] \le r[i] \le Q[i]$$



Suppose $5 \le r[1] \le 12$

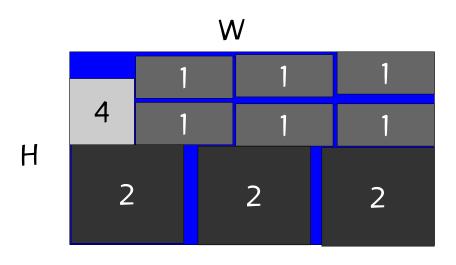


$$v[1] r[1] + v[2] r[2] + \cdots + v[N] r[N]$$



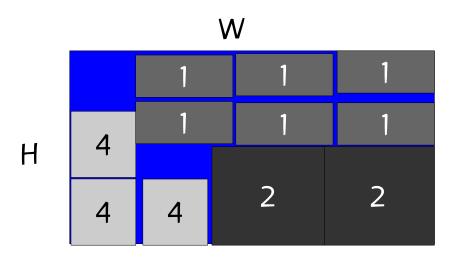


$$v[1] r[1] + v[2] r[2] + \cdots + v[N] r[N]$$



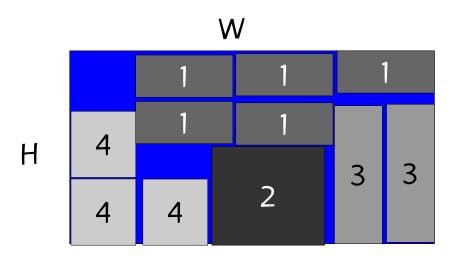


$$v[1] r[1] + v[2] r[2] + \cdots + v[N] r[N]$$





$$v[1] r[1] + v[2] r[2] + \cdots + v[N] r[N]$$





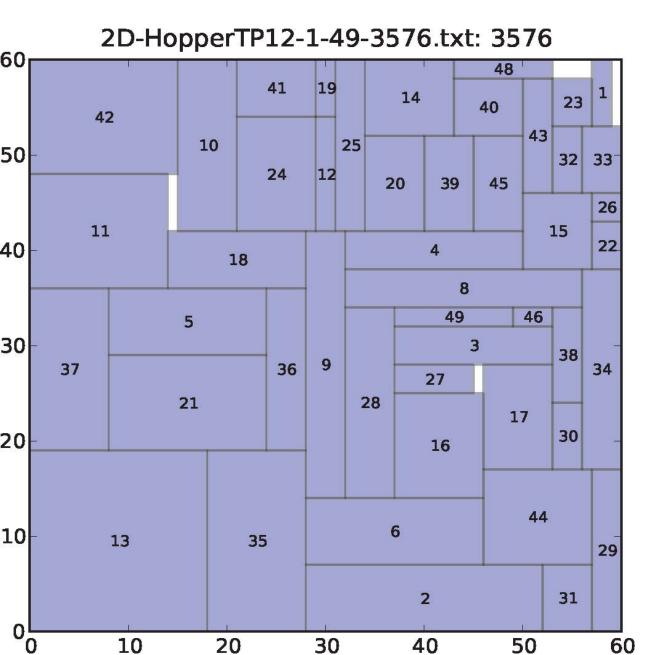
Applications

Problem arises in several production processes, e.g.

- Textile
- Glass
- Wood
- Paper

where rectangular figures are cut from large rectangular sheets of materials.

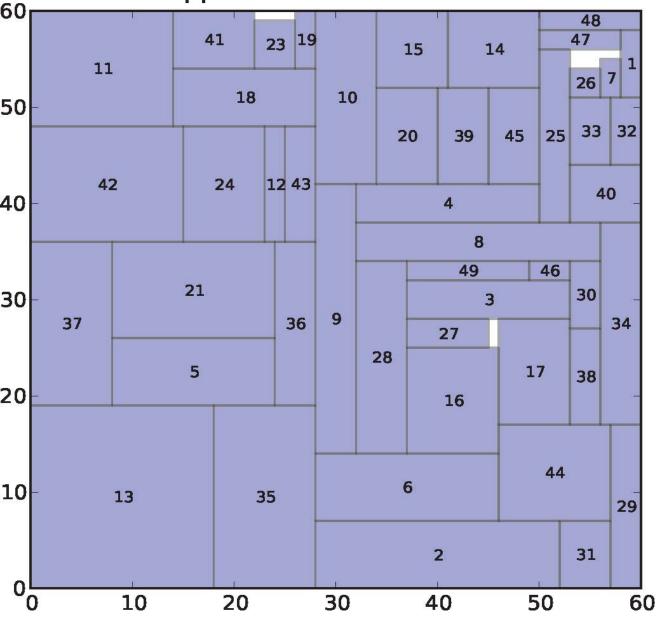




Hopper & Turton, 2001 Instance 4-1 60 x 60 Value: 3576



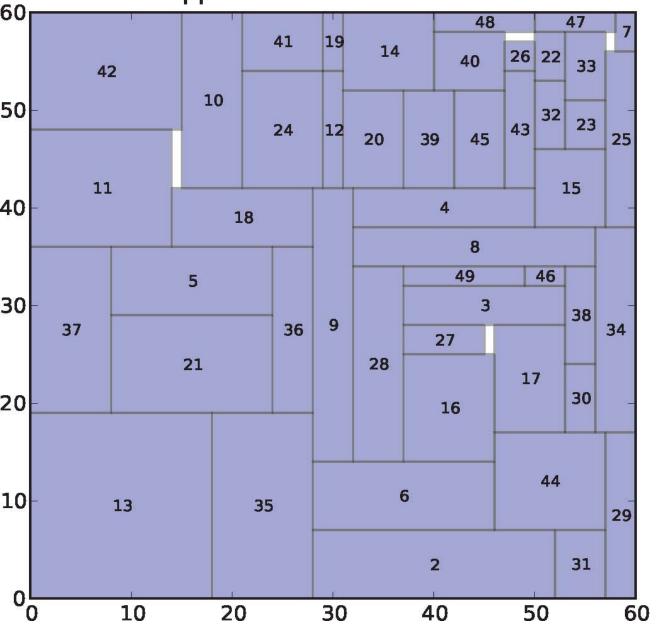
2D-HopperTP12-1-49-3585.txt: 3585



Hopper & Turton, 2001 Instance 4-2 60 x 60 Value: 3585



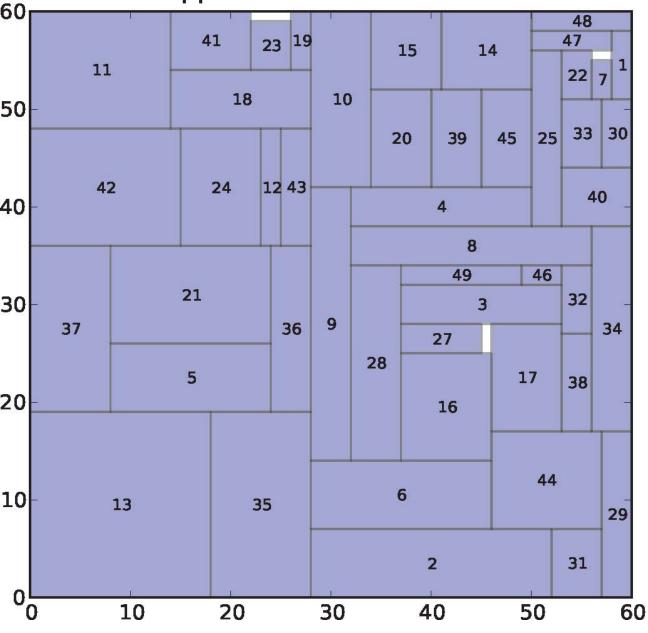
2D-HopperTP12-1-49-3586.txt: 3586



Hopper & Turton, 2001 Instance 4-2 60 x 60 Value: 3586



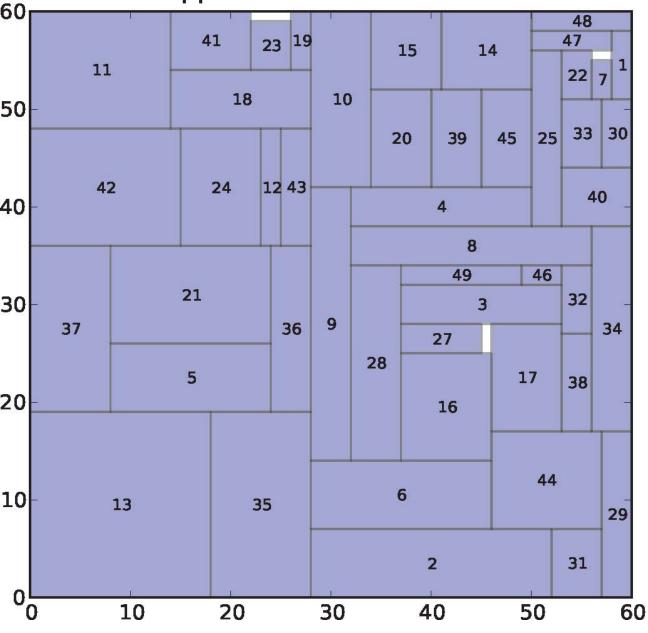
2D-HopperTP12-1-49-3591.txt: 3591



Hopper & Turton, 2001 Instance 4-2 60 x 60 Value: 3591



2D-HopperTP12-1-49-3591.txt: 3591



Hopper & Turton, 2001 Instance 4-2 60 x 60 Value: 3591 New best known solution! Previous best: 3580 by a

Tabu Search heuristic

(Alvarez-Valdes et al., 2007)



BRKGA for constrained 2-dim orthogonal packing



Encoding

- Solutions are encoded as vectors K of
 2N' = 2 { Q[1] + Q[2] + ··· + Q[N] }
 random keys, where Q[i] is the maximum number
 of rectangles of type i (for i = 1, ..., N) that can be
 packed.
- K = (k[1], ..., k[N'], k[N'+1], ..., k[2N'])



Encoding

Solutions are encoded as vectors K of

$$2N' = 2 \{ Q[1] + Q[2] + \cdots + Q[N] \}$$

random keys, where Q[i] is the maximum number of rectangles of type i (for i = 1, ..., N) that can be packed.

•
$$K = (k[1], ..., k[N'], k[N'+1], ..., k[2N'])$$

Rectangle type packing sequence (RTPS)



Encoding

Solutions are encoded as vectors K of

$$2N' = 2 \{ Q[1] + Q[2] + \cdots + Q[N] \}$$

random keys, where Q[i] is the maximum number of rectangles of type i (for i = 1, ..., N) that can be packed.

• K = (k[1], ..., k[N'],

k[N'+1], ..., k[2N'])

Rectangle type packing sequence (RTPS)

Vector of placement procedures (VPP)



- Simple heuristic to pack rectangles:
 - Make Q[i] copies of rectangle i, for i = 1, ..., N.
 - Order the N' = Q[1] + Q[2] + \cdots + Q[N] rectangles in some way.
 - Process the rectangles in the above order. Place the rectangle in the stock rectangle according to one of the following heuristics: bottom-left (BL) or leftbottom (LB). If rectangle cannot be positioned, discard it and go on to the next rectangle in the order.



- Simple heuristic to pack rectangles:
 - Make Q[i] copies of rectangle i, for i = 1, ..., N.
 - Order the N' = Q[1] + Q[2] + \cdots + Q[N] rectangles in some way. Sort first N' keys to obtain order.
 - Process the rectangles in the above order. Place the rectangle in the stock rectangle according to one of the following heuristics: bottom-left (BL) or leftbottom (LB). If rectangle cannot be positioned, discard it and go on to the next rectangle in the order.



- Simple heuristic to pack rectangles:
 - Make Q[i] copies of rectangle i, for i = 1, ..., N.
 - Order the N' = Q[1] + Q[2] + \cdots + Q[N] rectangles in some way. Sort first N' keys to obtain order.
 - Process the rectangles in the above order. Place the rectangle in the stock rectangle according to one of the following heuristics: bottom-left (BL) or left-bottom (LB). If rectangle cannot be positioned, discard it and go on to the next rectangle in the order. Use the last N' keys to determine which heuristic to use. If k[N'+i] > 0.5 use LB, else use BL.



- A maximal empty rectangular space (ERS) is an empty rectangular space not contained in any other ERS.
- ERSs are generated and updated using the Difference Process of Lai and Chan (1997).
- When placing a rectangle, we limit ourselves only to maximal ERSs. We order all the maximal ERSs and place the rectangle in the first maximal ERS in which it fits.
- Let (x[i], y[i]) be the coordinates of the bottom left corner of the i-th ERS.



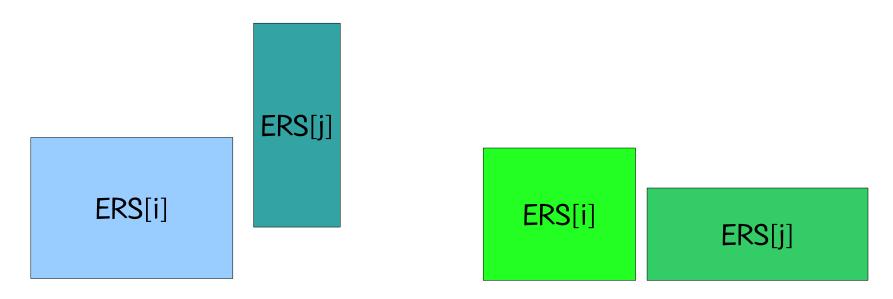
- A maximal empty rectangular space (ERS) is an empty rectangular space not contained in any other ERS.
- ERSs are generated and updated using the Difference Process of Lai and Chan (1997).
- When placing a rectangle, we limit ourselves only to maximal ERSs. We order all the maximal ERSs and place the rectangle in the first maximal ERS in which it fits.

ERS

• Let (x[i], y[i]) be the coordinates of the bottom left corner of the i-th ERS.

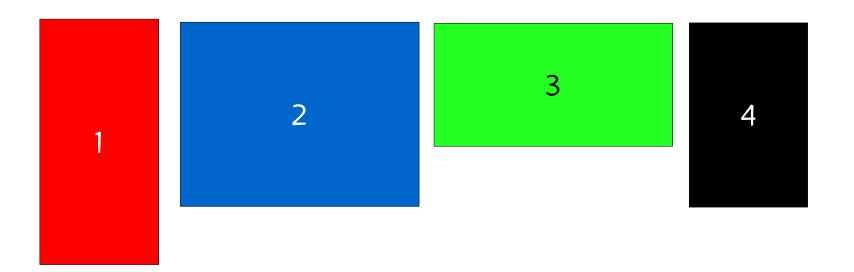


 If BL is used, ERSs are ordered such that ERS[i] < ERS[j] if y[i] < y[j] or y[i] = y[j] and x[i] < x[j].



ERS[i] < ERS[j]



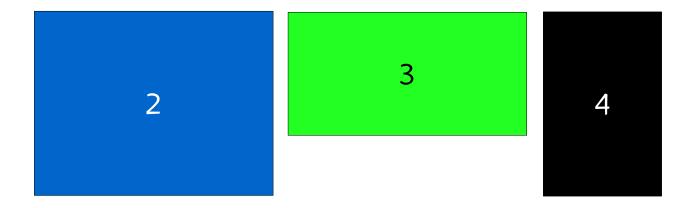


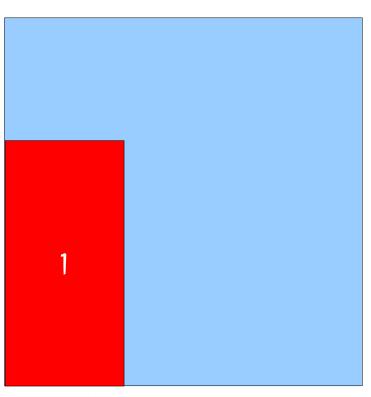
BL can run into problems even on small instances (Liu & Teng, 1999).

Consider this instance with 4 rectangles.

BL cannot find the optimal solution for any RTPS.

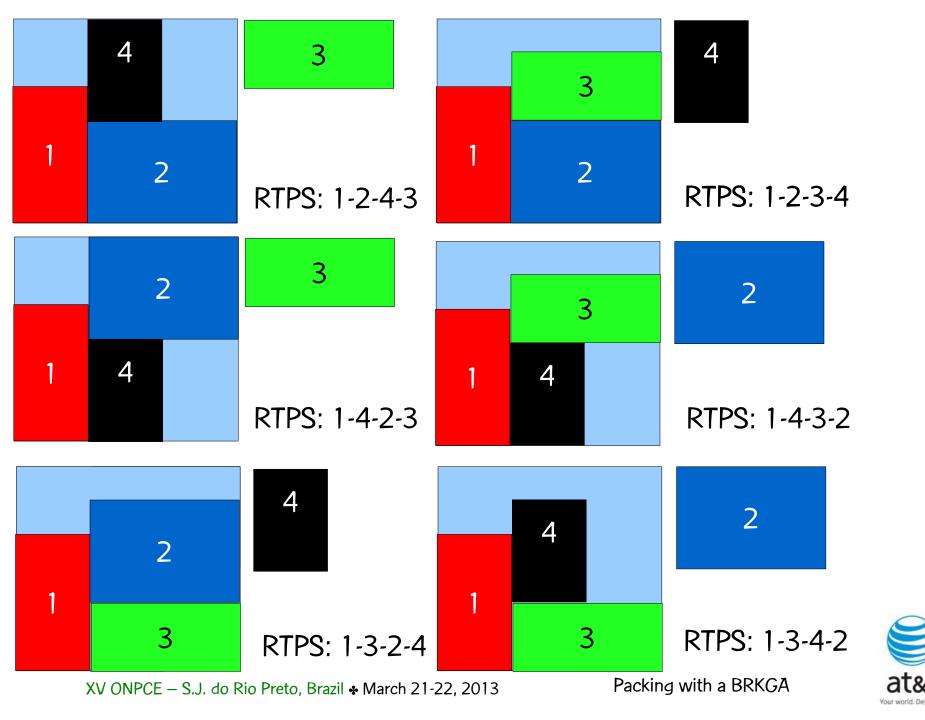


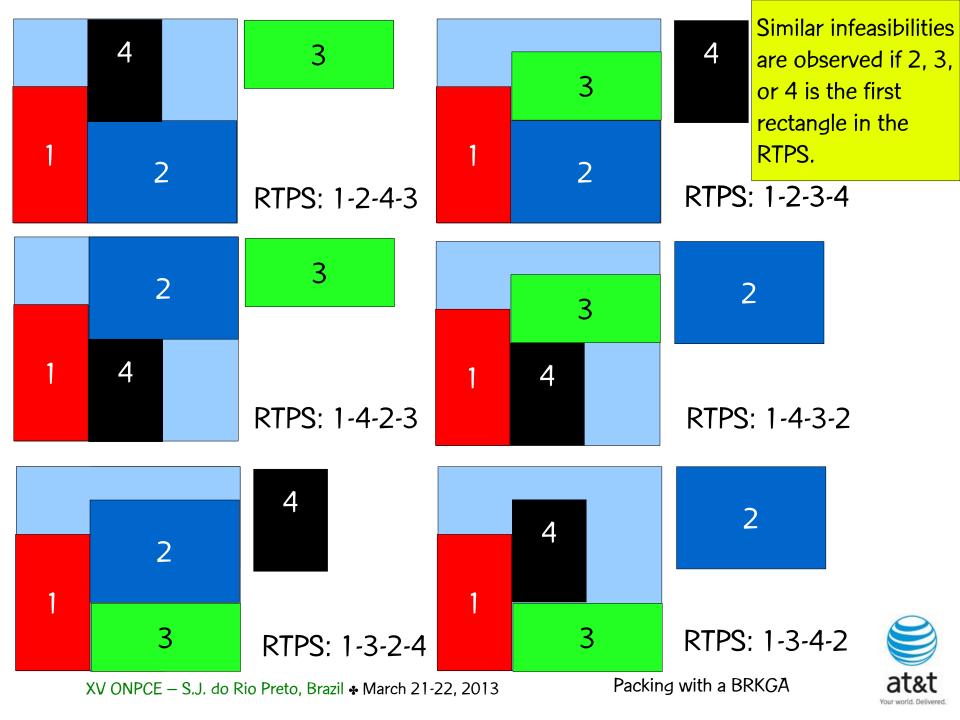




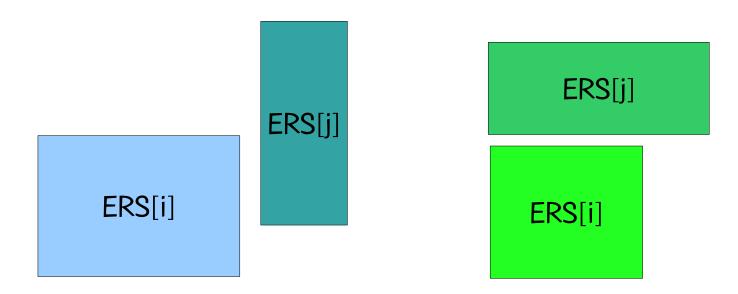
We show 6 rectangle type packing sequences (RTPS's) where we fix rectangle 1 in the first position.





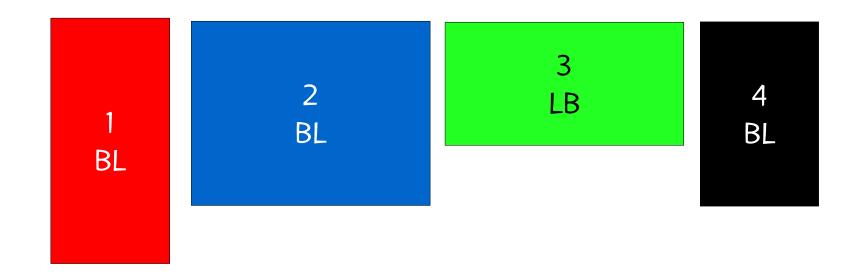


 If LB is used, ERSs are ordered such that ERS[i] < ERS[j] if x[i] < x[j] or x[i] = x[j] and y[i] < y[j].

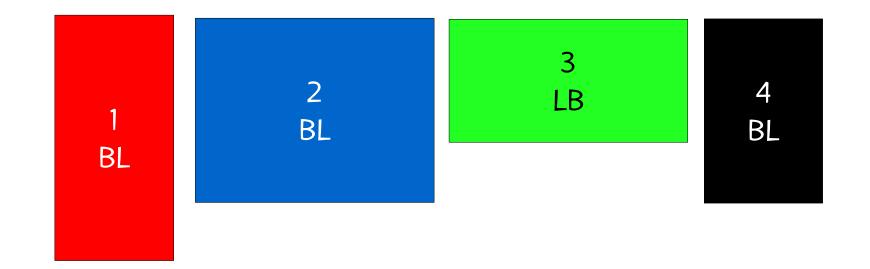


ERS[i] < ERS[j]



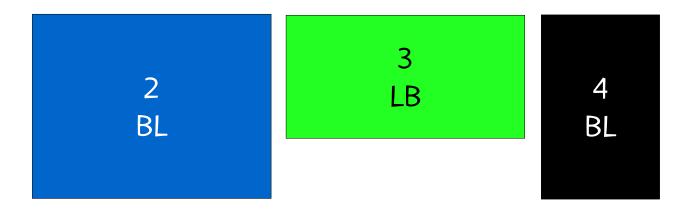


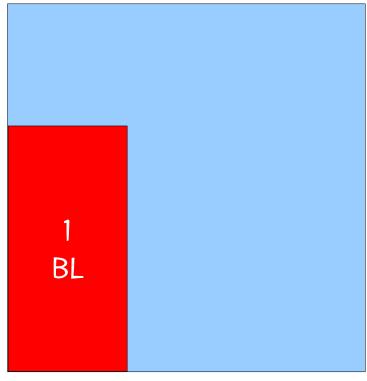






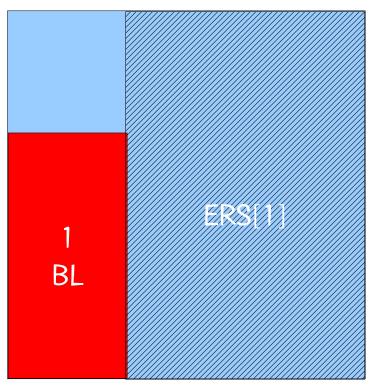




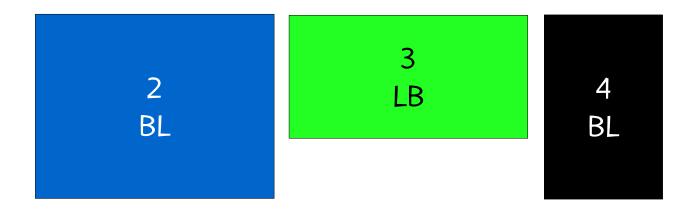


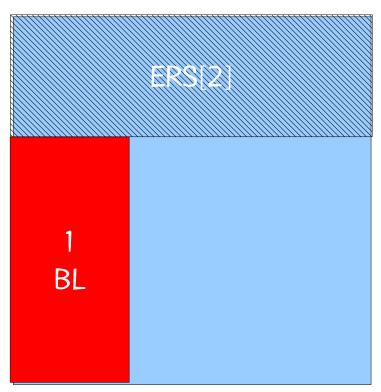






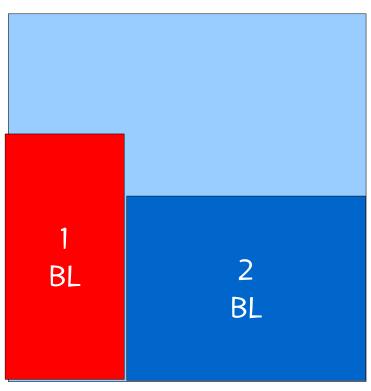






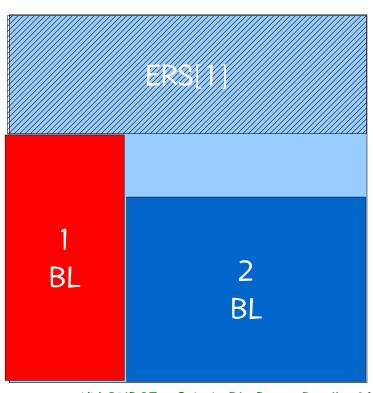




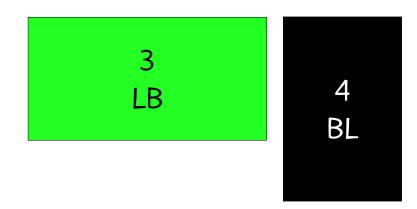


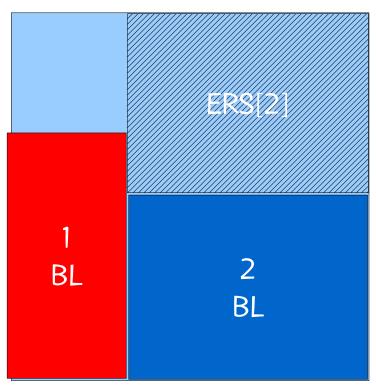




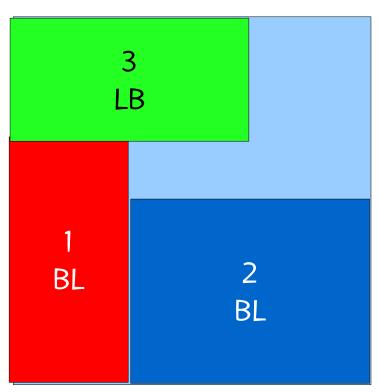




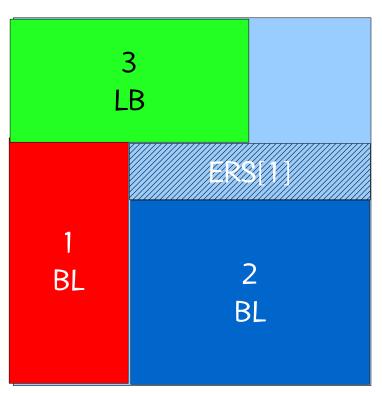








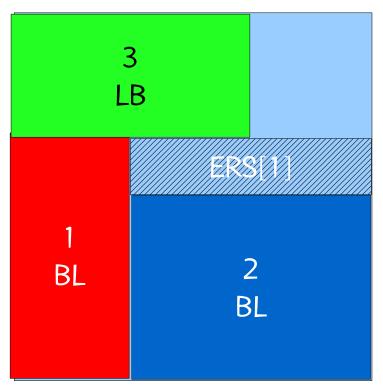








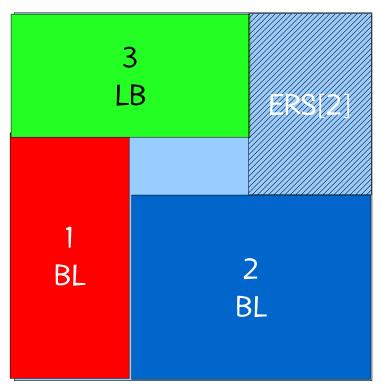
4 does not fit in ERS[1].



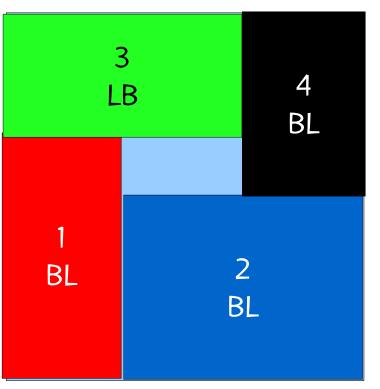




4 does fit in ERS[2].







Optimal solution!



Experimental results



 We compare solution values obtained by the parallel multi-population BRKGA with solutions obtained by the heuristics that produced the best computational results to date:



- We compare solution values obtained by the parallel multi-population BRKGA with solutions obtained by the heuristics that produced the best computational results to date:
 - PH: population-based heuristic of Beasley (2004)



- We compare solution values obtained by the parallel multi-population BRKGA with solutions obtained by the heuristics that produced the best computational results to date:
 - PH: population-based heuristic of Beasley (2004)
 - GA: genetic algorithm of Hadjiconsantinou & Iori
 (2007)



- We compare solution values obtained by the parallel multi-population BRKGA with solutions obtained by the heuristics that produced the best computational results to date:
 - PH: population-based heuristic of Beasley (2004)
 - GA: genetic algorithm of Hadjiconsantinou & Iori
 (2007)
 - GRASP: greedy randomized adaptive search procedure of Alvarez-Valdes et al. (2005)



- We compare solution values obtained by the parallel multi-population BRKGA with solutions obtained by the heuristics that produced the best computational results to date:
 - PH: population-based heuristic of Beasley (2004)
 - GA: genetic algorithm of Hadjiconsantinou & Iori
 (2007)
 - GRASP: greedy randomized adaptive search procedure of Alvarez-Valdes et al. (2005)
 - TABU: tabu search of Alvarez-Valdes et al. (2007)



Number of best solutions / total instances

Problem	PH	GA	GRASP	TABU	BRKGA BL-LB-L-4NR
From literature (optimal)	13/21	21/21	18/21	21/21	21/21
Large random [*]	0/21	0/21	5/21	8/21	20/21
Zero-waste			5/31	17/31	30/31
Doubly constrained	11/21		12/21	17/21	19/21

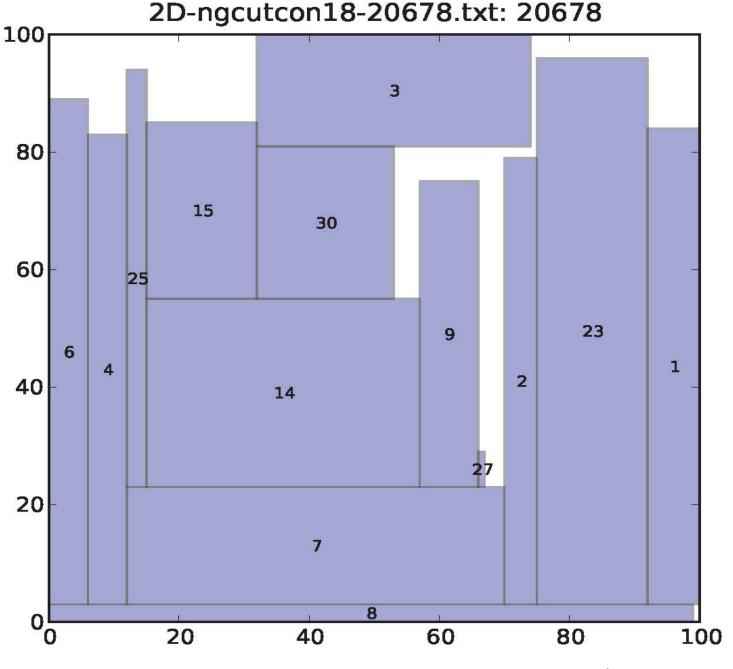
^{*} For large random: number of best average solutions / total instance classes



Minimum, average, and maximum solution times (secs) for BRKGA (BL-LB-L-4NR)

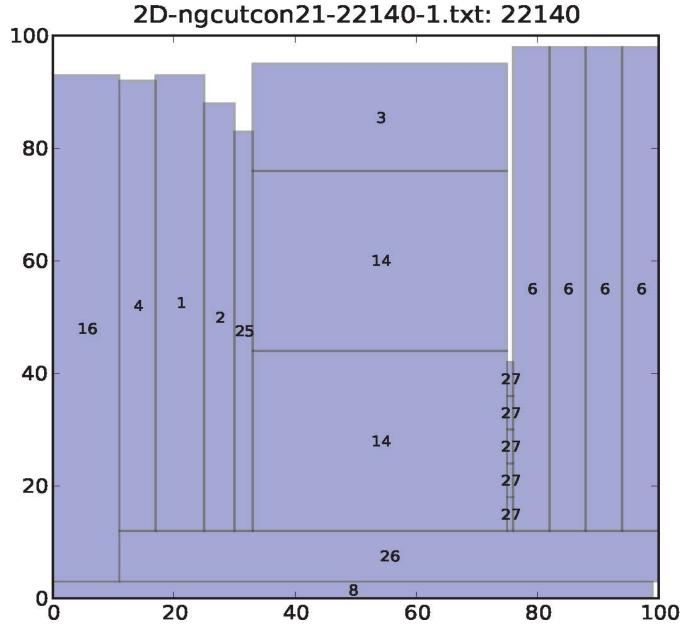
Problem	Min solution time (secs)	Avg solution time (secs)	Max solution time (secs)
From literature (optimal)	0.00	0.05	0.55
Large random	1.78	23.85	72.70
Zero-waste	0.01	82.21	808.03
Doubly constrained	0.00	1.16	16.87





New BKS for a 100 x100 doubly constrained instance of Fekete & Schepers (1997) of value 20678. Previous best was **19657** by tabu search of Alvarez-Valdes et al., (2007).

30 types30 rectangles



New BKS for a 100 x 100 doubly constrained instance Fekete & Schepers (1997) of value **22140**.

Previous BKS was **22011** by tabu search of Alvarez-Valdes et al. (2007).

29 types97 rectangles

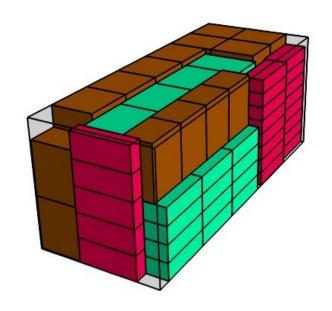


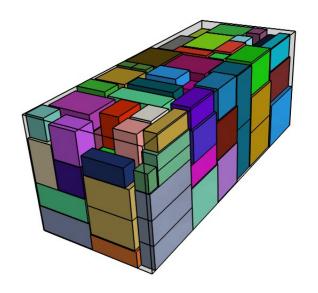
Some remarks

We have extended this to 3D packing:

J.F. Gonçalves and M.G.C.R., "A parallel multi-population biased random-key genetic algorithm for a container loading problem," Computers & Operations Research, vol. 29, pp. 179-190, 2012.

Tech report: http://www.research.att.com/~mgcr/doc/brkga-pack3d.pdf







3D bin packing

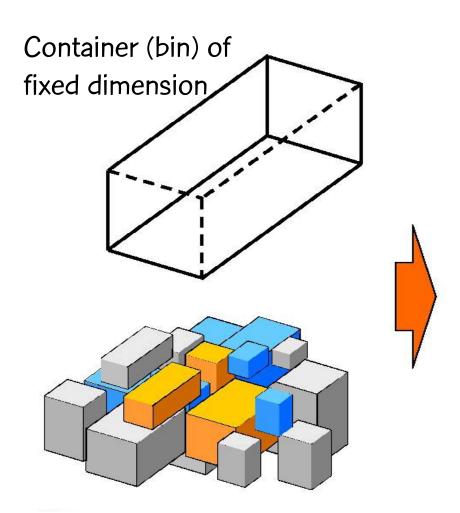


J.F. Gonçalves and M.G.C.R., "A biased random-key genetic algorithm for a 2D and 3D bin packing problem," AT&T Labs Research Technical Report, 2012 (submitted).

http://www.research.att.com/~mgcr/doc/brkga-binpacking.pdf

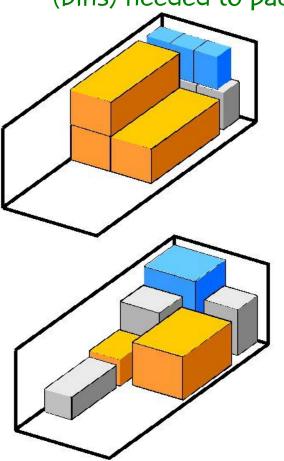


3D bin packing problem



Boxes of different dimensions

Minimize number of containers (bins) needed to pack all boxes



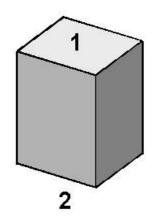


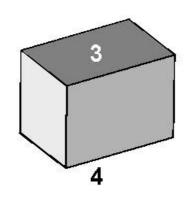
3D bin packing constraints

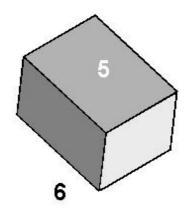
- Each box is placed completely within container
- Boxes do not overlap with each other
- Each box is placed parallel to the side walls of bin
- In some instances, only certain box orientations are allowed (there are at most six possible orientations)

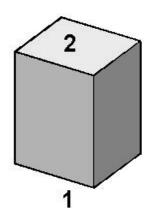


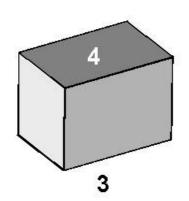
Six possible orientations for each box

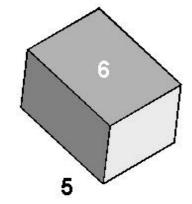








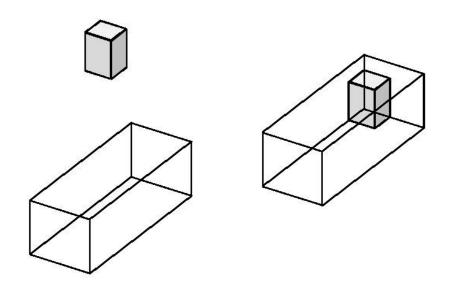




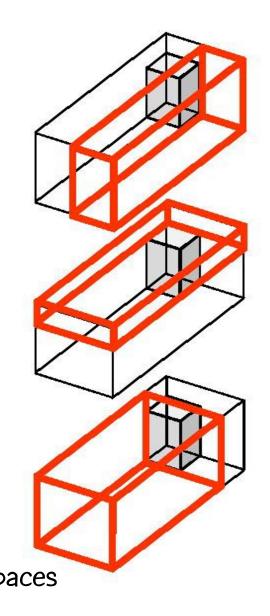


Difference process - DP

(Lai & Chan, 1997)



When box is placed in container ... use DP to keep track of maximal free spaces





Encoding

Solutions are encoded as vectors of 3n random keys, where n is the number of boxed to be packed.

$$X = (x_1, x_2, ..., x_n, x_{n+1}, x_{n+2}, ..., x_{2n}, x_{2n+1}, x_{2n+2}, ..., x_{3n})$$

Box packing sequence

Placement heuristic

Box orientation



Decoding

- 1) Sort first n keys of X to produce sequence boxes will be packed;
- 2) Use second n keys of X to determine which placement heuristic to use (back-bottom-left or back-left-bottom):
 - if $x_{n+1} < \frac{1}{2}$ then use back-bottom-left to pack i-th box
 - if $x_{n+1} \ge \frac{1}{2}$ then use back-left-bottom to pack i-th box
- 3) Use third n keys of X to determine which of six orientations to use when packing box:
 - $x_{2n+i} \in [0,1/6)$: orientation 1;
 - $x_{2n+i} \in [1/6,2/6)$: orientation 2; ...
 - $\cdot x_{2n+i} \in [5/6,1]$: orientation 6.



Decoding

For each box

- scan containers in order they were opened
- use placement heuristic to place box in first container in which box fits with its specified orientation
- if box does not fit in any open container, open new container and place box using placement heuristic with its specified orientation



Fitness function

Instead of using as fitness measure the number of bins (NB)

- use adjusted fitness: aNB
- aNB = NB + (LeastLoad / BinVolume), where
 - * LeastLoad is load on least loaded bin
 - * BinVolume is volume of bin: H x W x L



Experiment

Parameters:

- population size: p = 30n
- $_{-}$ size of elite partition: $p_{_{e}} = .10p$
- $_{-}$ number of of mutans: $p_{_{m}} = .15p$
- crossover probability: 0.7
- stopping criterion: 300 generations



Experiment

Instances:

- 320 instances of Martello et al. (2000)
- generator is available at http://www.diku.dk/~pisinger/codes/html
- 8 classes
- 40 instances per class
- 10 instances for each value of n ∈ {50, 100, 150, 200}



Experiment

- We compare BRKGA with:
 - TS3, the tabu search of Lodi et al. (2002)
 - GLS, the guided local search of Faroe et al. (2003)
 - TS2PACK, the tabu search of Crainic et al. (2009)
 - GRASP, the greedy randomized adaptive search procedure of Parreno et al. (2010)



Class 1 - Bin size: 100 x 100 x 100

Boxes	LB	BRKGA	GRASP	TS3	TS2PACK	GLS
50	12.5	13.4	13.4	13.4	13.4	13.4
100	25.1	26.6	26.6	26.6	26.7	26.7
150	34.7	36.4	36.4	36.7	37.0	37.0
200	48.4	50.9	50.9	51.2	51.1	51.2
Sum:	120.7	127.3	127.3	127.9	128.2	128.3



Class 2 - Bin size: 100 x 100 x 100

Boxes	LB	BRKGA	GRASP	TS3	TS2PACK	GLS
50	12.7	13.8	13.8	13.8		
100	24.1	25.6	25.7	25.7	Didn	ot rup
150	35.1	36.7	36.9	37.2	ו טוע וו	ot run
200	47.5	49.4	49.4	50.1		
Sum:	119.4	125.5	125.8	126.8		



Class 3 - Bin size: 100 x 100 x 100

Boxes	LB	BRKGA	GRASP	TS3	TS2PACK	GLS
50	12.3	13.3	13.3	13.3		
100	24.7	25.9	26.0	26.0	Didn	ot run
150	36.0	37.5	37.6	37.7	טומ וו	ot run
200	47.8	49.8	50.0	50.5		
Sum:	120.8	126.5	126.9	127.5		



Class 4 - Bin size: 100 x 100 x 100

Boxes	LB	BRKGA	GRASP	TS3	TS2PACK	GLS
50	28.7	29.4	29.4	29.4	29.4	29.4
100	57.6	59.0	59.0	59.0	58.9	59.0
150	85.2	86.8	86.8	86.8	86.8	86.8
200	116.3	118.8	118.8	118.8	118.8	119.0
Sum:	287.8	294.0	294.0	294.0	293.9	294.2



Class 5 - Bin size: 100 x 100 x 100

Boxes	LB	BRKGA	GRASP	TS3	TS2PACK	GLS
50	7.3	8.3	8.3	8.4	8.3	8.3
100	12.9	15.0	15.0	15.0	15.2	15.1
150	17.4	20.0	20.1	20.4	20.1	20.2
200	24.4	27.1	27.1	27.6	27.4	27.2
Sum:	62.0	70.4	70.5	71.4	71.0	70.8



Class 6 - Bin size: 10 x 10 x 10

Boxes	LB	BRKGA	GRASP	TS3	TS2PACK	GLS
50	8.7	9.8	9.8	9.9	9.8	9.8
100	17.5	18.8	19.0	19.1	19.1	19.1
150	26.9	29.2	29.2	29.4	29.2	29.4
200	35.0	37.2	37.4	37.7	37.7	37.7
Sum:	88.1	95.0	95.4	96.1	95.8	96.0



Class 7 - Bin size: 40 x 40 x 40

Boxes	LB	BRKGA	GRASP	TS3	TS2PACK	GLS
50	6.3	7.4	7.4	7.5	7.4	7.4
100	10.9	12.2	12.5	12.5	12.3	12.3
150	13.7	15.2	16.0	16.1	15.8	15.8
200	21.0	23.4	23.5	23.9	23.5	23.5
Sum:	51.9	58.2	59.4	60.0	59.0	59.0



Class 8 - Bin size: 100 x 100 x 100

Boxes	LB	BRKGA	GRASP	TS3	TS2PACK	GLS
50	0.8	9.2	9.2	9.3	9.2	9.2
100	17.5	18.9	18.9	18.9	18.8	18.9
150	21.3	23.5	24.1	24.1	23.9	23.9
200	26.7	29.3	29.8	30.3	30.0	29.9
Sum:	66.3	80.9	82.0	82.6	81.9	81.9



Summary

Class	Bin size	BRKGA	GRASP	TS3	TS2PACK	GLS
1	100 ³	127.3	127.3	127.9	128.2	128.3
2	100 ³	125.5	125.8	126.8		
3	100 ³	126.5	126.9	127.5		
4	100 ³	294.0	294.0	294.0	293.9	294.2
5	100 ³	70.4	70.5	71.4	71.0	70.8
6	10 ³	95.0	95.4	96.1	95.8	96.0
7	40 ³	58.2	59.4	60.0	59.0	59.0
8	100 ³	80.9	82.0	82.6	81.9	81.9
Sum(ro	ows 1, 4-8):	725.8	728.6	732.0	729.8	730.2
Sur	n(rows 1-8):	977.8	981.3	986.3		



Concluding remarks

- Reviewed BRKGA framework
- Applied framework to
 - 2D/3D packing to maximize value packed
 - 2D/3D bin packing to minimize number of bins
- All decoders were simple heuristics
- BRKGA "learned" how to "operate" the heuristics
- In all cases, several new best known solutions were produced



hanks

These slides and all of the papers cited in this lecture can be downloaded from my homepage:

http://www.research.att.com/~mgcr

