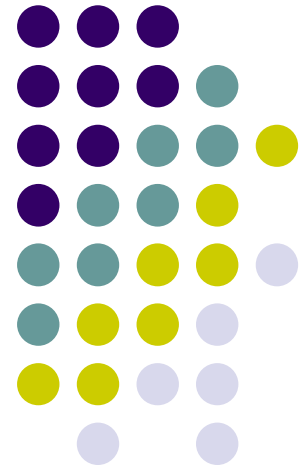


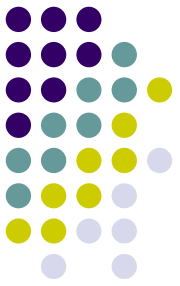
Structural Accuracy of Probabilistic Models in BOA

Claudio F. Lima
University of Algarve, Portugal

Work in collaboration with Martin Pelikan, David E. Goldberg,
Fernando G. Lobo, Kumara Sastry, and Mark Hauschild

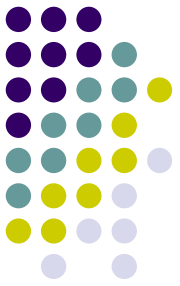


Outline

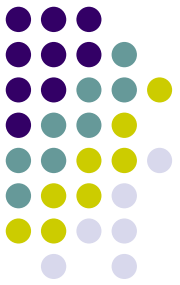


- BOA in one slide
- Motivation
- Measuring structural accuracy of PMs in BOA
- Influence of BOA's parameters
 - Population Size
 - Selection Strategy
 - Replacement Strategy
- Conclusions

BOA: Bayesian optimization algorithm (Pelikan *et al.*, 1999)

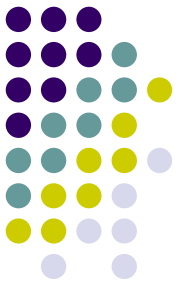


- EDA that builds and samples from a Bayesian network to guide search
- Similar to EBNA (Etxeberria *et al.*, 1999) and LFDA (Muhlenbein, 2000)
- BNs can model complex multivariate dependencies between variables
- Has been successfully applied in real-world problems (Pelikan; 2005, 2006)



Overfitting in EDAs

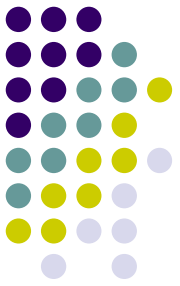
- EDAs can solve efficiently challenging optimization problems
 - Where exploiting problem structure is a plus
- But oftentimes their probabilistic models do not exactly reflect the problem structure
- Why?
 - PMs are learned from a sample of limited size (population)
 - Particular features of that specific sample are also encoded
 - Well-known problem in ML → **overfitting**
 - See Wu & Shapiro (2006)



A Closer Look at BOA (w/ DTs)

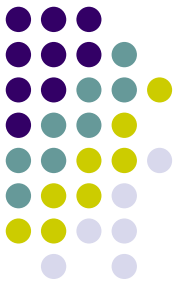
- Looking at the PMs in BOA, one can see:
 - Important dependencies are discovered
 - But with additional spurious linkages
- While BN structure captures such excessive complexity...
- ...the corresponding **conditional probabilities** nearly express **independency** between spurious and correct variables

When is Structural Accuracy Crucial?



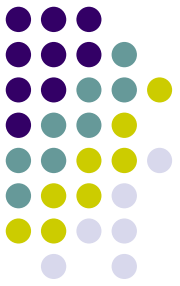
- **Model-based** efficiency enhancement techniques
 - Evaluation relaxation (Sastry *et al.*, 2004, Pelikan *et al.*, 2004, Sastry *et al.*, 2006)
 - Time continuation (Lima *et al.*, 2005; Lima *et al.*, 2006)
- **Off-line** usage of linkage information
 - Learn about the problem while trying to solve it
 - Build problem-specific operators with gained knowledge

What can affect model accuracy in BOA?



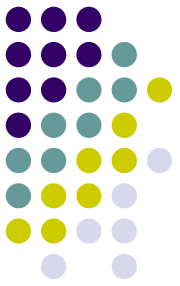
- Parameters
 - Population Size (Pelikan, 2005)
 - Selection Strategy (Johnson *et al.*, 2001; Santana *et al.*, 2005)
 - Replacement Strategy
- Problem Structure
 - Subfunction Size
 - Subfunction Overlapping Degree
 - Subfunction Fitness Distribution
 - Hierarchy
- Bayesian Network Learning
 - Search Procedure (Wu *et al.*, 2006; Echevoyen *et al.*, 2007)
 - Score Metric (Pelikan, 2005; Correa *et al.*, 2006)

Experimental Setup for Measuring Structural Accuracy of PMs

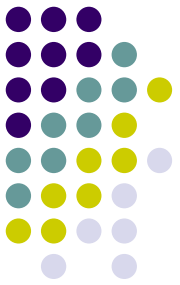


- Test problem(s) structure should be known
- Learning problem structure should be crucial to solve it efficiently
- Easy control of problem size and difficulty in order to test **scalability**
- *m-k* deceptive trap functions:
 - Concatenated **m** trap functions of size **k** each
 - Total problem size is $m \cdot k$ (in case of overlapping, $m(k-o)+o$)
 - **k**-order statistics are required to (efficient) success
 - Lower-order statistics lead the search away from the optimum

Analyzing Probabilistic Models in BOA

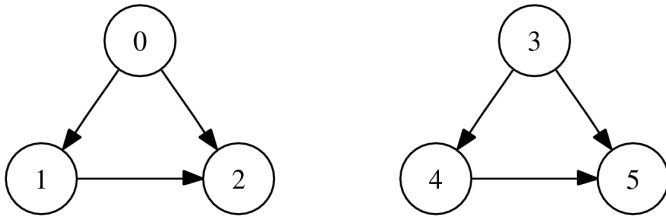


- The big question: When does the BN structure matches problem structure?
- Looking for answers:
 - Edges: The good, the bad, and the missing ones!
 - Mark Hauschild's talk on Monday, 15:15h
 - Problem substructure "represented" in the BN?
 - This talk
- Measure the proportion of substructures represented and their accuracy



A simple example for Trap3

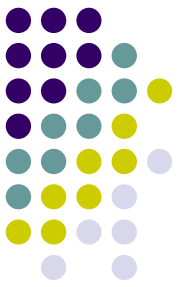
- Ideal BN for Trap3:



- Dependency relations

- $0 \leftarrow \{ \}$
- $1 \leftarrow \{ 0 \}$
- $2 \leftarrow \{ 0, 1 \}$
- $3 \leftarrow \{ \}$
- $4 \leftarrow \{ 3 \}$
- $5 \leftarrow \{ 2, 3 \}$

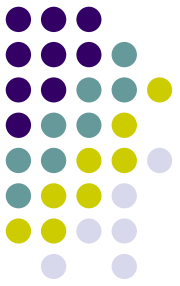
- Ideally, relations should respect a certain chain
- But the dependency that relates all k variables is crucial!
- And most difficult to learn
- In this way, k -order statistics are maintained



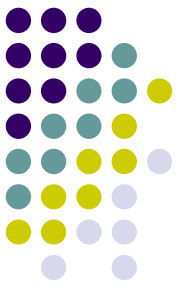
What are we going to measure?

- Proportion of subfunctions with **correct** linkage group
 - Example: $2 \leftarrow \{ 0,1 \}$
- Proportion of subfunctions with **spurious** linkage group
 - Example: $2 \leftarrow \{ 0,1,3,4,5 \}$
- Proportion of subfunctions with any of the above
 - Sum of the two previous proportions
- Average size of spurious linkage
 - Number of spurious variables

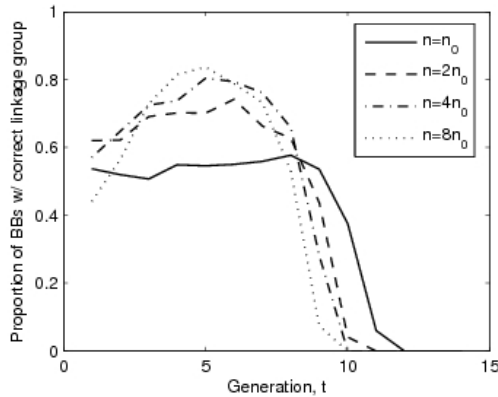
Influence of BOA's parameters



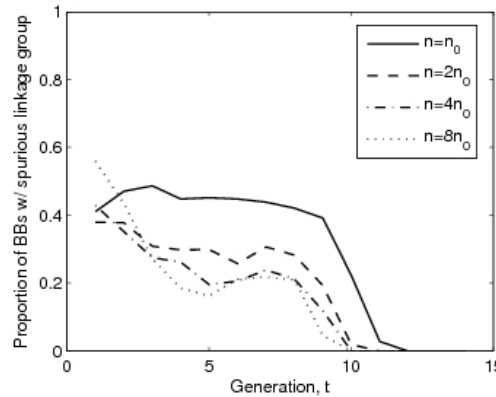
- Population Size
- Selection Strategy
 - Selection method
 - Selection intensity
- Replacement Strategy
 - Replacement method
 - Elitist intensity, diversity preservation



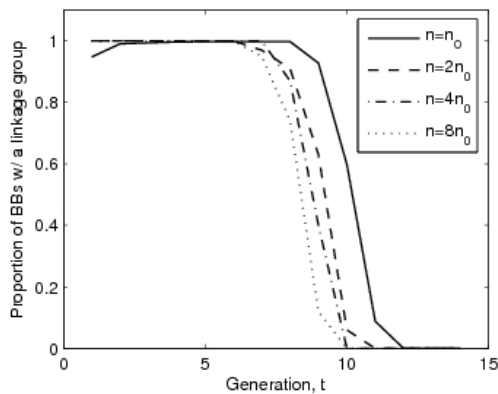
Influence of Population Size



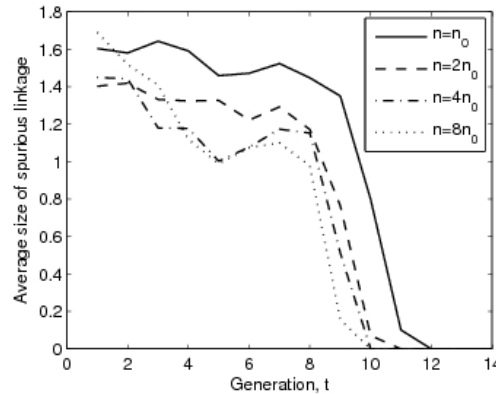
(a) Correct linkage groups



(b) Spurious linkage groups

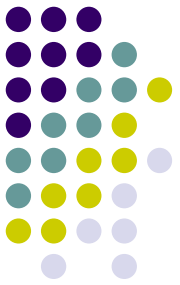


(c) All linkage groups



(d) Average size of spurious linkage

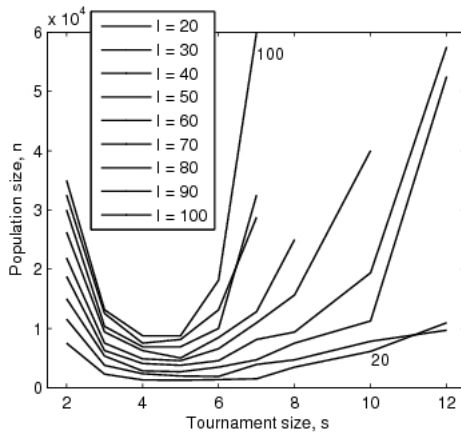
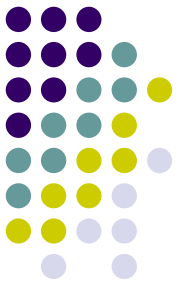
- $m = 10$ and $k = 5$
- n_0 is the minimal population size required to solve problem
- Exponentially increasing population sizes: n_0 , $2n_0$, $4n_0$, and $8n_0$
- Increasing population size slightly improves model accuracy



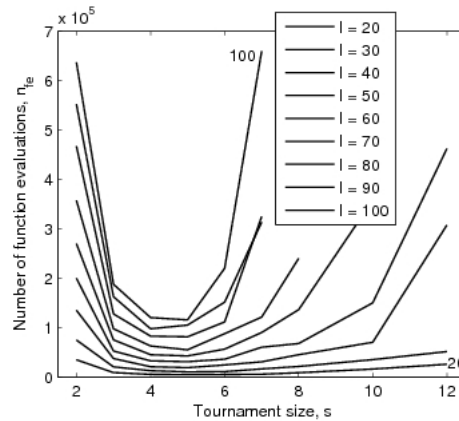
Influence of Selection Strategy

- Tournament Selection
 - Pick best solution from a tournament of s individuals
 - Repeat n times (w/ or wo/ replacement)
- Truncation Selection
 - Choose the best δ % proportion of the population
- Different selection dynamics
 - In terms of selection variance and lost of diversity (Blickle & Thiele, 1997)

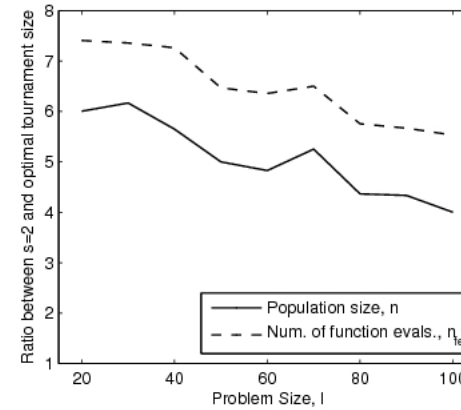
Influence of Selection Intensity on n , n_{fe} , and resulting Speedup



(a) Population size



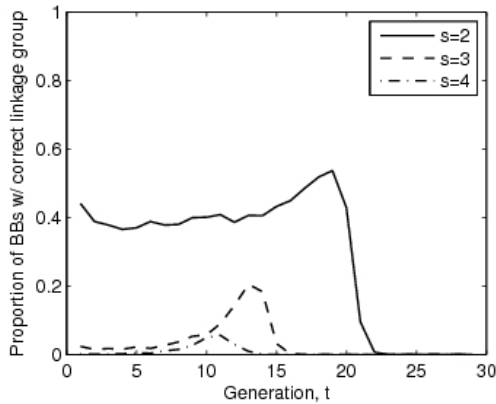
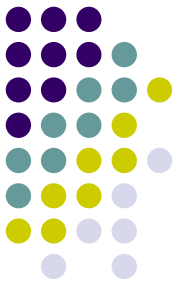
(b) Number of function evaluations



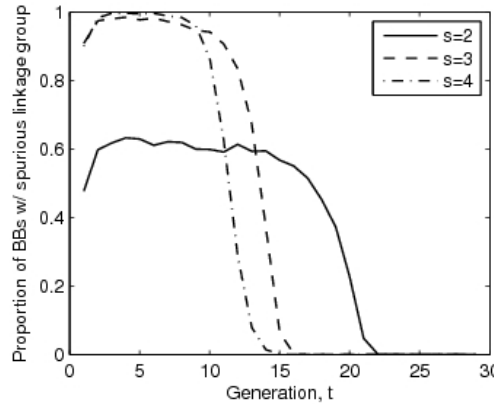
(c) Speedup

- **Good news:** Increasing selection pressure reduces the required population size and num. of evals
- **Not so good news:** the resulting speedup seems to decrease with problem size

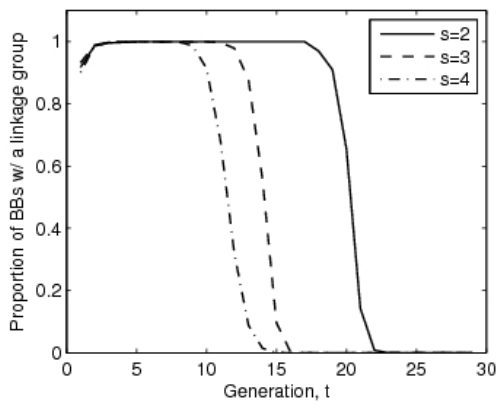
Tournament Selection



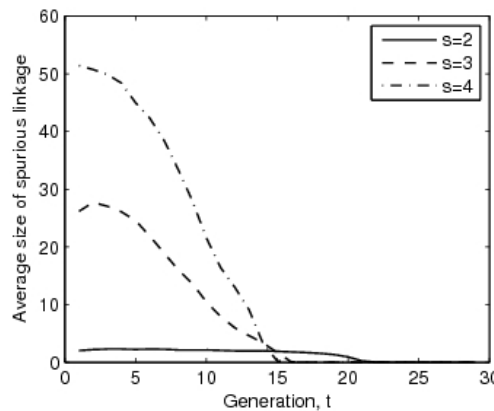
(a) Correct linkage groups



(b) Spurious linkage groups

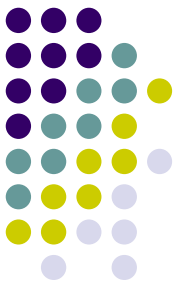


(c) All linkage groups

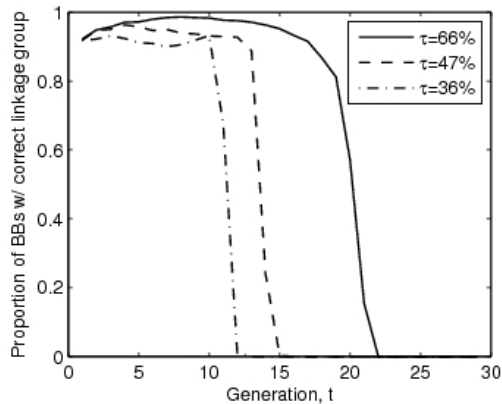


(d) Average size of spurious linkage

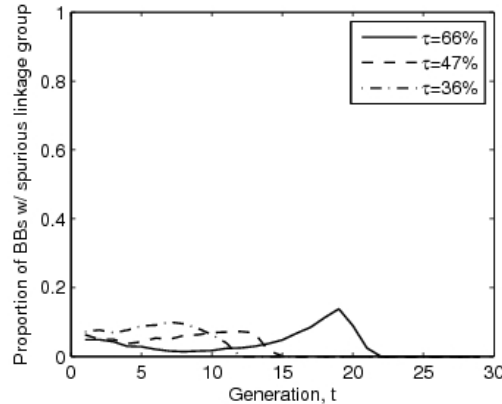
- $m = 24$ and $k = 5$ ($l = 120$)
- Tournament sizes $s = 2, 3, 4$
- Proportion of accurate structures is not significant (in particular for larger s)
- However, w/ or wo/ spurious linkage, all substructures are represented
- Higher selection intensity increases spurious linkage size



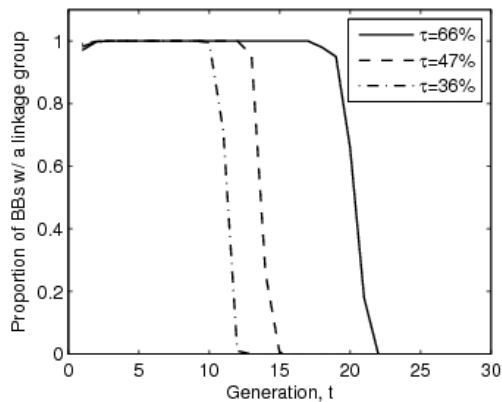
Truncation Selection



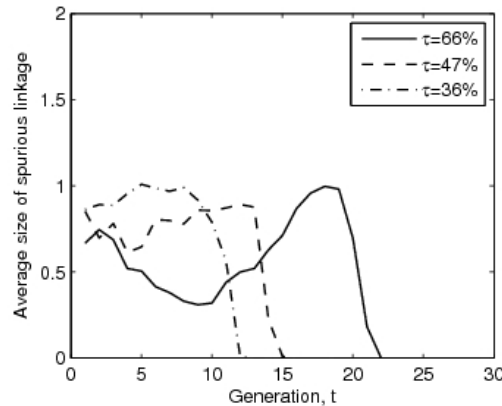
(a) Correct linkage groups



(b) Spurious linkage groups

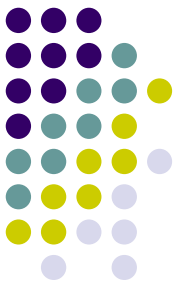


(c) All linkage groups

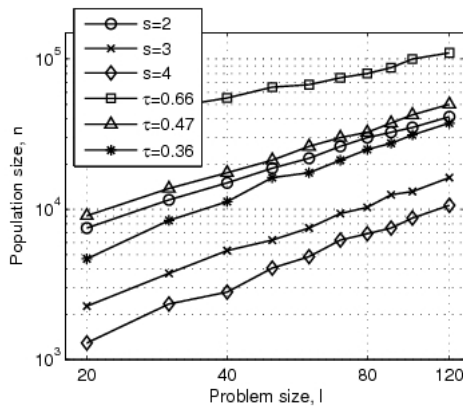


(d) Average size of spurious linkage

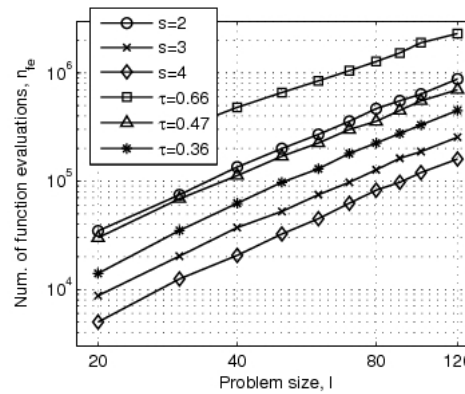
- $m = 24$ and $k = 5$ ($l = 120$)
- Truncation thresholds $\delta = 66\%$, 47% , 36%
- Same selection intensity as $s = 2, 3, 4$
- Proportion of accurate structures is now close to 100%
- Number of spurious variables is low
- Higher selection intensity has small impact on model structural accuracy



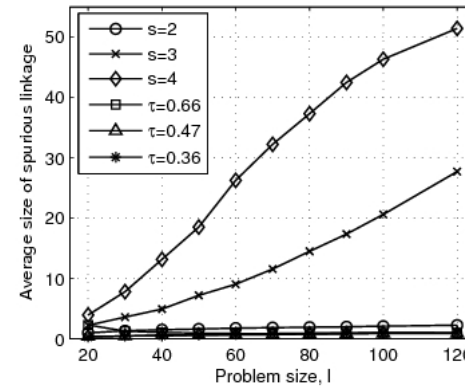
Influence of Selection Strategy



(a) Population size



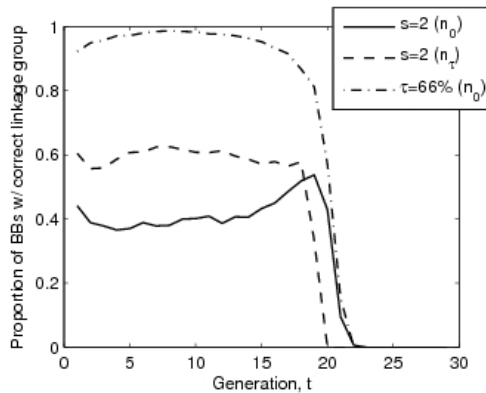
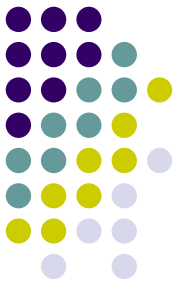
(b) Number of function evaluations



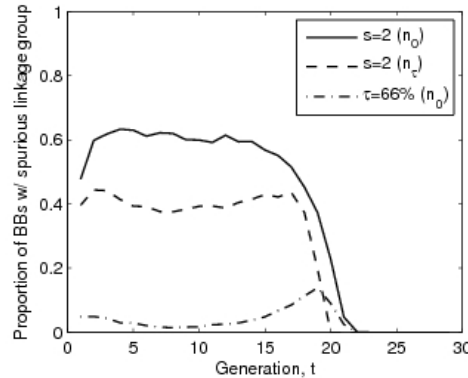
(c) Average size of spurious linkage

- n and n_{fe} required for truncation is higher than for tournament by a significant however constant factor (2-3 times more)

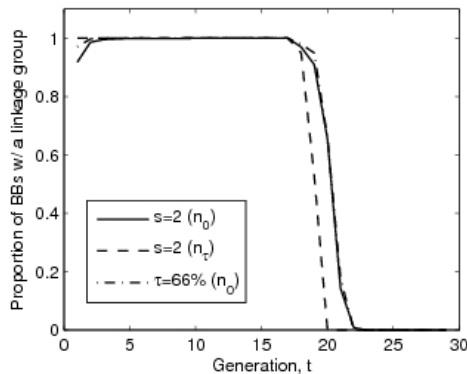
Tournament vs. Truncation: Is it a matter of population size?



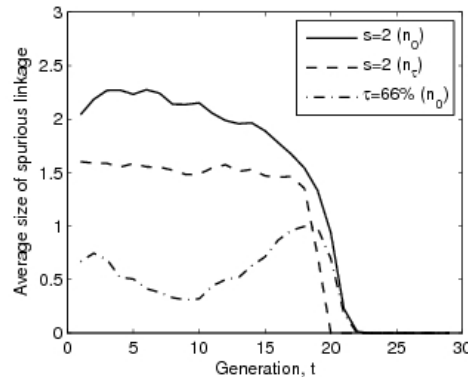
(a) Correct linkage groups



(b) Spurious linkage groups

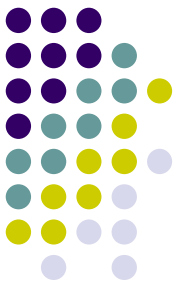


(c) All linkage groups



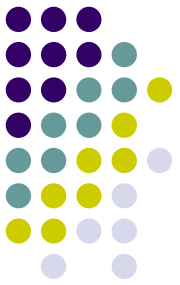
(d) Average size of spurious linkage

- $m = 24$ and $k = 5$ ($l = 120$)
- Comparing tournament ($s = 2$) and truncation ($\delta = 66\%$) with the same:
 - population size
 - selection intensity
- Tournament selection improves accuracy but is still not close to truncation



What is it then?

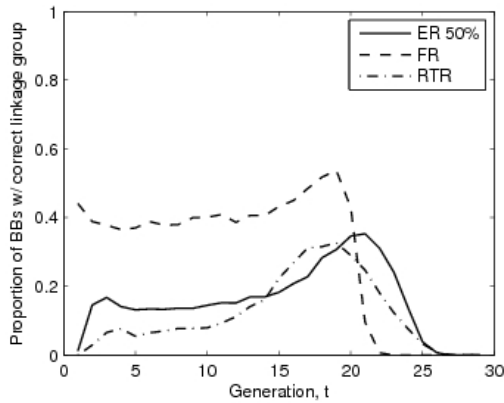
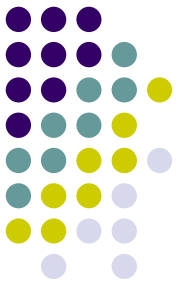
- For the same selection intensity:
 - Truncation has a higher loss of diversity...
 - ...and a lower selection variance (Blickle & Thiele, 1997)
- **But...** distribution of the number of copies in the selected population affects BN learning
 - In tournament the number of copies is somewhat proportional to its rank (best guy gets exactly s copies)
 - In truncation no particular relevance is given to top solutions
- Model overfitting of top individuals takes place (in particular for increasing s)



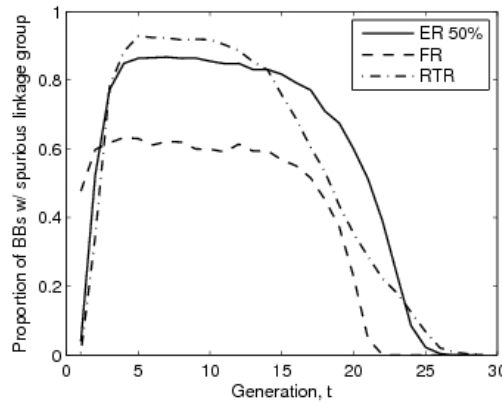
Influence of Replacement Strategy

- Full replacement (FR)
 - Offspring population fully replaces the parents
- Elitist replacement (ER)
 - Worst portion of the parents is replaced by offspring
- Restricted tournament replacement (RTR)
 - Offspring directly competes with similar parent for a place in the next population (diversity preservation)

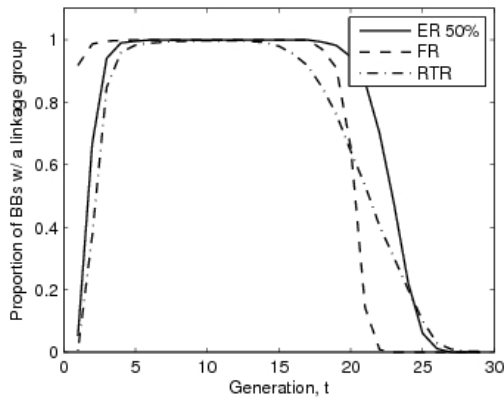
Influence of Replacement (Tournament Selection)



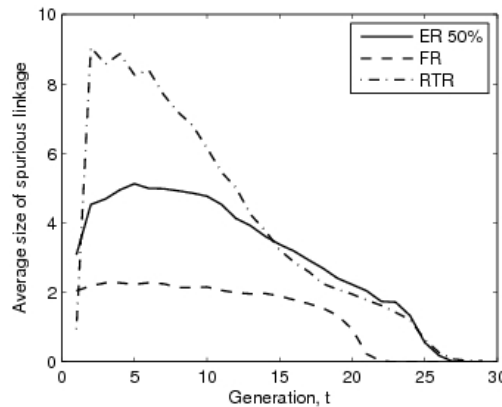
(a) Correct linkage groups



(b) Spurious linkage groups



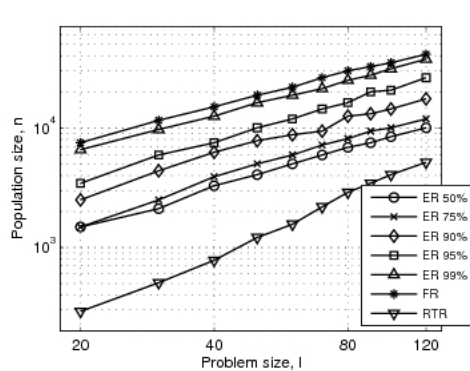
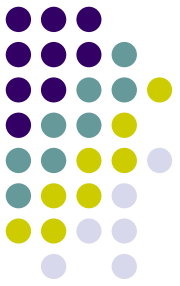
(c) All linkage groups



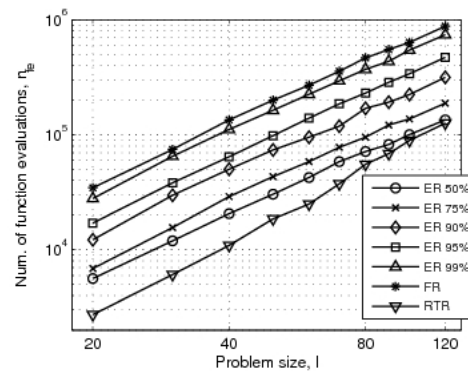
(d) Average size of spurious linkage

- $m = 24$ and $k = 5$ ($l = 120$)
- Comparing ER50%, FR, and RTR with tournament selection ($s = 2$)
- Structural accuracy is higher with FR
- Replacement strategy does not have the same impact as selection on the spurious linkage size
- Spurious linkage is more frequent with RTR because of the smaller pop. size

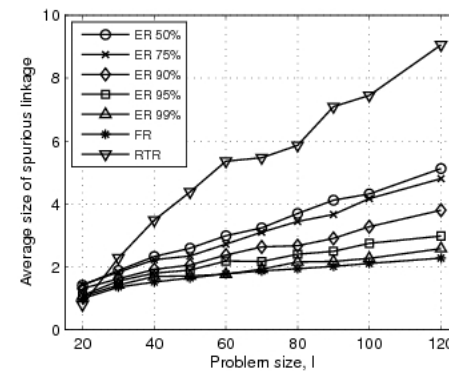
Influence of Replacement (Tournament Selection)



(a) Population size



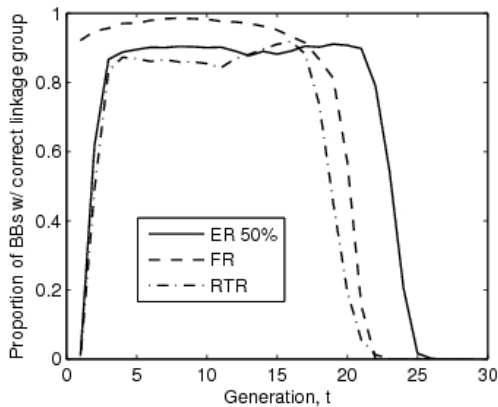
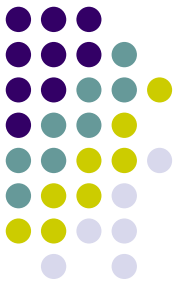
(b) Number of function evaluations



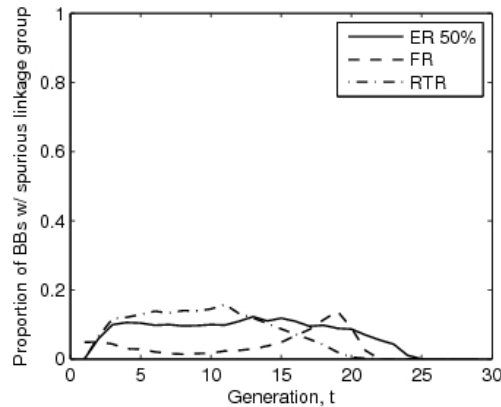
(c) Average size of spurious linkage

- Additional elitist proportions are shown (1-50%)
- RTR clearly requires less n and n_{fe}
- However, linkage information is not so accurate

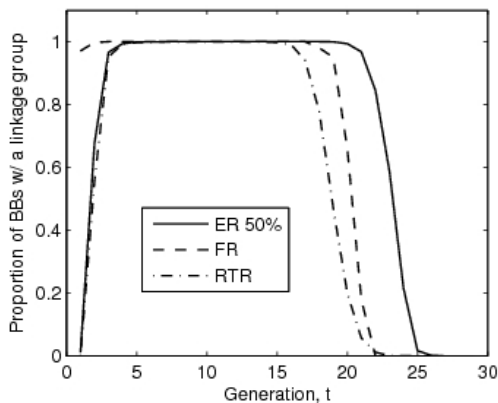
Influence of Replacement (Truncation selection)



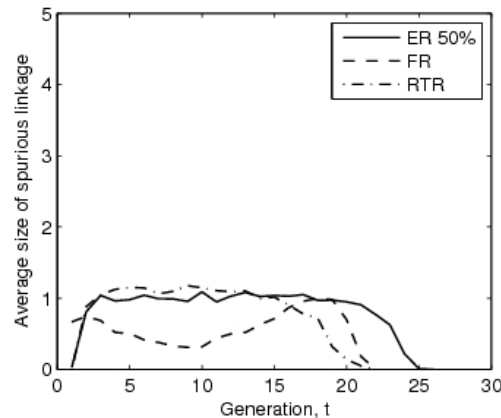
(a) Correct linkage groups



(b) Spurious linkage groups

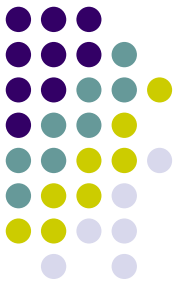


(c) All linkage groups



(d) Average size of spurious linkage

- $m = 24$ and $k = 5$ ($l = 120$)
- Same comparison with truncation selection ($\delta = 66\%$)
- Structural accuracy is close to 100% for all replacement strategies
- FR is still slightly better because it needs larger n
- Size of spurious linkage is quite low



Conclusions

- Trade-off between model accuracy and overall performance in BOA
- Truncation is better than tournament selection for accurate modelling
- For the same purpose, the replacement method is more relevant if tournament selection is used (in which case full replacement is better)
- If overall performance (num. of evals) is our main concern, tournament and RTR are the best options

Structural Accuracy of Probabilistic Models in BOA

Claudio F. Lima
University of Algarve, Portugal

Work in collaboration with Martin Pelikan, David E. Goldberg,
Fernando G. Lobo, Kumara Sastry, and Mark Hauschild

