Department of Mathematics and Computer Science
University of Southern Denmark, Odense

March 6, 2011
Marco Chiarandini

DM826 - Modeling and Solving Constrained Optimization Problems

Obligatory Assignment 3, Spring 2011

Deadline: 31th March 2011 at noon.

Motivation One of the current threads of research in management is the attempt to understand human decision-making in search for new alternatives. Experimental work in this area use the *NK* model of rugged performance landscapes [2]. In these experiments, human subjects are tasked to search for high-performing product configurations. They combine several attributes — the *N* parameter in the *NK* model — to specify a product configuration. The value or payoff of a particular configuration is initially unknown and is only discovered after trying out the configuration. The complexity of alternatives — the *K* parameter — is captured by allowing for interactions among the attributes in the payoff function. As complexity increases and interactions among attributes proliferate, the problem of finding a high-performing configuration becomes more challenging and difficult. The subjects in the experiments have only a limited number of search trials, far fewer than the number of all possible configurations. The aim of the experiments is studying the opportunity cost of searching a new alternative varying the experimental treatment of performance landscape complexity *K*.

A dominant search strategy detected in the experiments is *neighborhood search*: the move to another position occurs by changing from 1 to 3 attributes. More specifically, for $N = 10$, experiments showed that out of 7539 active searches, 75.9% were within distance 3 from the current position with an average distance of 2.53, 84.5% within distance 4 and 90.3% within distance 5. In other terms, humans do not change radically the attributes available but perform local changes. Other observations show the importance of performance feedback [1].

Given these results there is interest to design an experiment in which the distance factor is removed. In other terms, researchers in management would like to have an *NK* model in which the user is given the possibility to only move to new search states that are in the neighborhood of the current one. A way to achieve this is to provide to the humans an easy graphical representation of the possible moves available. A chess board in which moves are only allowed between neighboring cells seems well suited for this goal. The problem is then to dispose the 2^N different search alternatives in the board in such a way that the maximum distance between any pair of adjacent cells is minimized.

Mathematical background The *NK* model defines a combinatorial search space, consisting of every string (chosen from a given alphabet) of length *N*. For each string in this search space, a scalar value (called the fitness) is defined. If a distance metric is defined between strings, the resulting structure is a landscape. Fitness values are defined according to the specific incarnation of the model, but the key feature of the *NK* model is that the fitness of a given string *S* is the sum of contributions from each locus S_i in

1111	0101	1100	0110	15	5	12	6
0111	1101	0100	1110	7	13	4	14
1011	0001	1000	0010	11	1	8	2
0011	1001	0000	1010	3	9	0	10

Figure 1: A solution to $N = 4$ and max distance $D = 2$. On the left the binary strings disposed on the board; on the right the representation of the solution in decimal numbers.

the string:

$$F(S) = \sum_{i=1}^{N} f(S_i)$$

and the contribution from each locus in general depends on the value of K other loci:

$$f(S_i) = f(S_i, S_1^i, \ldots, S_K^i)$$

where S_j^i are the other loci upon which the fitness of S_i depends. Hence, the fitness function is a mapping between strings of length $K + 1$ and scalar values. (Source: `http://en.wikipedia.org/wiki/NK_model`)

The problem formalized Given a board of $n \times n$ cells and n^2 distinct binary strings of $N = \lceil 2 \log_2 n \rceil$ bits find an assignment of strings to cells in such a way that the maximum Hamming distance between adjacent cells with wrap-around is minimized.

Note that in a $n \times n$ board a cell i, j has 8 neighboring cells.

A solution for $N = 4$ exists with max distance $D = 2$ but not with $D = 1$. Hence, $D = 2$ is the best possible maximal distance for this case. For higher N solutions are not known, in particular it would be interesting finding solutions for cases on $N = 8$ and $N = 10$.

Your tasks

1. Formulate and solve in Comet the decision version of the problem that asks whether there exists a solution if the maximum distance is fixed to D. Use set variables to formulate the problem.

2. Recognizing that the binary strings can be seen as encodings of integer numbers $[1..n^2]$, use this fact to implement an alternative formulation of the problem that uses the global constraint `alldiff` and the dual viewpoint of numbers in binary and decimal encoding. In particular, implement an ad hoc propagator to handle the constraint on the Hamming distance.

   Compare the two models on the basis of the size of problems they can solve in a reasonable time, say one hour. If their computational limits are on problems of the same size then compare the models on the basis of computing time, the number of failures and the number of branchings of the models from the following points.

3. In the previous two points you used the default labelling search strategy. Try to improve the performance by using different heuristics.

4. List all the symmetries that you can recognize and add constraints before the search to break some of these symmetries.

References

[1] Stephan Billinger, Nils Stieglitz, and Terry R. Schumacher. Search on rugged landscapes: An experimental study. *SSRN eLibrary*, 2010.

[2] Stuart A. Kauffman. *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford University Press, 1993.