Elements for the description of fitness landscapes associated with local operators for layered drawings of directed graphs

Pascale Kuntz, Bruno Pinaud, Rémi Lehn

Chapter 1

ELEMENTS FOR THE DESCRIPTION OF FITNESS LANDSCAPES ASSOCIATED WITH LOCAL OPERATORS FOR LAYERED DRAWINGS OF DIRECTED GRAPHS

Pascale Kuntz

pascale.kuntz@irin.univ-nantes.fr

Bruno Pinaud

bruno.pinaud@irin.univ-nantes.fr

Rémi Lehn

remi.lehn@irin.univ-nantes.fr

Ecole polytechnique de l'université de Nantes (Polytech'Nantes)
Institut de recherche en informatique de Nantes (IRIN)
France

Abstract   Minimizing arc crossings for drawing acyclic digraphs is a well-known *NP*-complete problem for which several local-search approaches based on local transformations (switching, median ,...) have been proposed. Their adaptations have been recently included in different metaheuristics. As an attempt to better understand the dynamic of the search processes, we study the fitness landscapes associated with these transformations. We first resort to a set of multi-start descents to sample the search space for three hundred medium-sized graphs. Then, we investigate complete fitness landscapes for a set of 1875 smaller graphs, this aims at showing some instance characteristics that influence search strategies. The underlying idea is to consider a fitness landscape as a graph whose vertices are drawings and arcs representing a transformation of a drawing into another. We confirm that the properties of basins of attraction closely depend on the instances. Also, we show that the probability of being stuck on a local optimum is linked to the specific shapes of the basins of attraction of global optima which may be very different from the regular image of the continuous case generally used as a reference.

Keywords: layered digraph drawing, fitness landscape, metaheuristics, local search

1.1   Introduction

Layered drawing of directed acyclic graphs is a classical problem of the huge graph drawing literature (Di-Battista et al. 2000; 1999). For the applications, this problem has known renewed interest in the past years in particular in the field of "information visualization " (Web browsing, cartography, etc.) (Herman et al. 2000). It consists in intelligibly arranging vertices on vertical layers and representing arcs by oriented line segments which flow in the same direction. Experimental observations have recently confirmed that one of the most important criteria for readability is arc crossing minimization (Purchase 2000).

Minimizing arc crossing for a layer digraph could a priori seem easier than the general problem of minimizing arc crossing on a plane since the choice of geometric coordinates is here transformed into a choice of vertex ordering on each layer. Yet, it remains *NP*-complete even if there are only two layers (Garey and Johnson 1983, Eades and Wormald 1994). Several deterministic local transformations have been proposed. The easiest ones, which have originally been defined for 2-layered drawings (Eades and Kelly 1986, Jünger and Mutzel 1997), are based on simple permutations on each layer. For instance, the greedy-switch heuristic iteratively switches consecutive pairs of vertices if this decreases the crossing number, and the split heuristic assigns the vertices above or below a given pivot vertex according to the same criterion. But probably, the most popular transformations are the so-called "averaging heuristics" including the barycenter heuristic (Sugiyama et al. 1981) and the median heuristic (Eades and Wormald 1994). The basic underlying idea is that arc crossings tend to be minimized when connected vertices are placed facing each other. Roughly speaking, these approaches compute the average positions, i.e. the barycenter or median of their neighbors, for the vertices on each layer and sort them according to these values. Several variants and implementations of these operators have been defined (a chronological review is provided in Laguna et al. (1997)).

More recently, these transformations have been combined for different meta-heuristics, in particular Tabu search (Laguna et al. 1997), GRASP (Marti 2001) and Evolutionary algorithms (Utech et al. 1998). These approaches have been essentially validated by numerical experiments on test sets constructed with the graph size and the density criteria. However, as far as we know, the limits of the influence of these criteria on the search processes have not been precisely studied. Moreover, some hypotheses on the properties of the search spaces -on the multimodality notably- have been voiced without proof. As an attempt to better characterize these laters, we here propose a descriptive analysis of fitness landscapes associated with the classical operators introduced above. In

particular, we aim at precising the distribution of the optima, and defining the shapes of their basins of attraction.

In the first part of the paper, we resort to a set of multi-start descents to explore the search space for 300 medium-sized graphs which are characteristic of the graphs we used for statistical rule visualization in data mining (Kuntz et al. 2000). Previous numerical experiments with one thousand descents on another graph family have underlined the great variability of the results (Lehn and Kuntz 2001); the number of descents converging on an "optimal" solution - i.e. the best solution among the computed solution set- followed a distribution close to a Poissonian one. For validating the robustness of the first tests we here extend the experiments with a set of 5000 descents. The results confirm the dispersion previously observed with differences in the distribution, but they show that the graph density criterium can partly explain the observed features only.

In the second part, we complete these statistical results by computing the exact fitness landscapes of a graph family from the graph-based definition introduced by Jones and Forrest (Jones 1995, Jones and Forrest 1995b). Due to the combinatorial explosion of problems of that nature, we were obliged to restrict ourselves to small instances, but, we have preserved a variability (1875 graphs) representative of the general problem. Besides the classical fitness-distance correlation, we compute the probability of being stuck on a local optimum with a descent and show that this result is closely linked to the specific shapes of the basins of attraction which are very different from the regular image of the continuous case generally used as a reference.

1.2   Problem formulation

Hereafter we consider an acyclic digraph $G = (V, A)$ with a set V of n vertices and a set A of m arcs, a set $L = \{l_1, l_2, ..., l_k\}$ of K layers, and a given distribution $V_1, V_2, ..., V_K$ of V on L with respectively $n_1, n_2, ..., n_K$ vertices. The vertex ordering on l_k is defined by $\sigma_k : V_K \rightarrow \{1, 2, ..., n_K\}$: $\sigma_k(u) = i$ which means that the vertex $u \in l_k$ is on the ith position on l_k. In a layered drawing, every arc $(u, v) \in A$ flows in the same direction: if $u \in V_i$ then $v \in V_j$ where $i < j$. Moreover, we suppose that the graph is proper i.e. each arc $(u, v) \in A$ is connected to vertices on consecutive layers: $u \in V_i$ and $v \in V_{i+1}$. We reach this hypothesis by replacing an arc whose length λ is greater than one by a path of $\lambda - 1$ dummy vertices on consecutive layers.

The problem is rewritten as defining a vertex ordering σ_k on each layer l_k so that the associated drawing minimizes the arc crossing number. We denote by Ω^G the set of all layered drawings of G and $f(D_i)$ the arc crossing number for a drawing $D_i \in \Omega^G$.
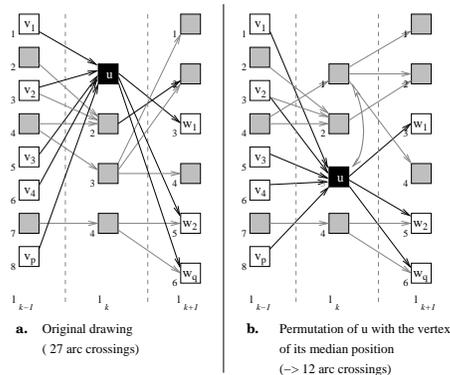
Figure 1.1. Permutation with the median operator ($O2$).

A neighbor of a drawing $D_i \in \Omega^G$ is a drawing $D_j \in \Omega^G$ deduced from D_i by the application of a local operator O which acts on one layer at a time. The set of neighbors of D_i is denoted by $N_O(D_i)$. For considerations discussed in section 3, we here retain two local operators: the classical switching and a variant of the median.

Switching ($O1$). $D_j = O1(D_i)$ means that D_j is deduced from D_i by swapping two adjacent vertices in the same layer $l_k \in L$ if $f(D_j) < f(D_i)$.

Median ($O2$). We consider a definition inspired by the barycenter definition introduced by Sugiyama et al. (1981) in a variant of their well-known method: the median of a vertex u on a layer l_k depends on the connected vertices in both layers l_{k-1} and l_{k+1}. Let us suppose that the connected vertices of u on l_{k-1} (resp. l_{k+1}) are $v_1, ..., v_p$ (resp. $w_1, ..., w_q$). The median position $m(u)$ is the median of the following sorted set of the normalized position of each neighbor of u in l_{k-1} and l_{k+1} :

$$\left\{ \frac{\sigma_{k-1}(v_1)}{n_{k-1}}, ..., \frac{\sigma_{k-1}(v_p)}{n_{k-1}}, \frac{\sigma_{k+1}(w_1)}{n_{k+1}}, ..., \frac{\sigma_{k+1}(w_q)}{n_{k+1}} \right\}$$

The median is computed for each vertex on l_k and the transformation $O2$ is defined by the new vertex ordering σ'_k on l_k obtained after sorting the median values in increasing order: $\sigma'_k(u) > \sigma'_k(v)$ if $m(u) > m(v)$ (see figure 1.1). This transformation is applied if it improves f.

1.3 Exploration with a multi-start descent

Let us first remark that the search space is intrinsically multimodal in that there may exist some local optima. For instance, the drawings D_4 and D_5 of figure 1.2 are local optima for $O1$; different local permutations (e.g. c and d on D_2) necessary to reach the global optimum D_3 lead separately to an arc
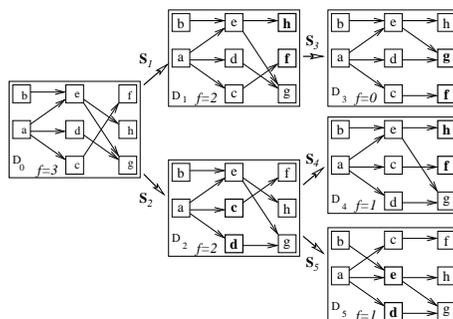
Figure 1.2. Series of transformations of a drawing D_0 on the left with the switch operator ($O1$). The drawing D_3 is a global optimum.

crossing increasing each of them, and are consequently not selected. Moreover, the global optimum is not necessarily unique (mainly because of global or local symmetries in the layout).

1.3.1 Experimental conditions

In order to precise the characteristics of this multimodality, we use an extended multi-start on a graph family. Let Δ be a set of acyclic digraphs which are not trees s.t.[1] $10^6 \leq \Pi_{k=1}^{K} n_k! \leq 10^{14}$. We have randomly generated a set $\Delta' \subset \Delta$ of 300 graphs with $4 \leq K \leq 40$ and $n_k \leq 15$. For each random graph G, a set of 5000 layouts is randomly generated, and each layout is improved by an iterative application of an operator O: O is applied on each layer taken one after the other and this loop goes on until the objective function f stabilizes.

As the objective here is not to develop a new heuristic but to better know search spaces associated with operators close to those used in the literature, we have just tested different simple variants of O easy to compute in the above multi-start descent. We first restricted ourselves to $O = O1+O2$: for a layer l_k the ordering is first improved by the switching operator $O1$ then by the median operator $O2$. And we have experimentally observed that the obtained results can be slightly improved when a barycenter operator is applied after the median. As for the median, we have considered that the barycenter of a vertex u on a layer l_k depends on the connected vertices in both layers l_{k-1} and l_{k+1}. This definition has proved less performing in the Sygiyama's heuristic when vertices whose orders have not yet been improved are considered in the computation (Sugiyama et al. 1981). But, since the operator is not applied alone

---

[1]$\Pi_{k=1}^{K} n_k!$, where n_k is the number of vertices in layer k of a particular instance of the problem, represents the number of vertices of the fitness landscape associated with the instance i.e. the size of the associated search space.
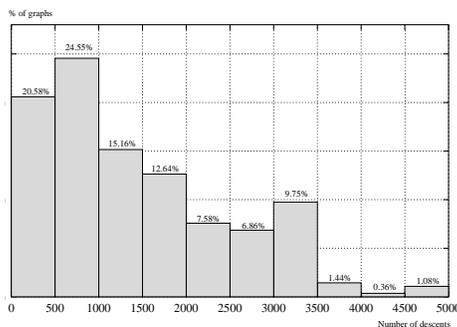
% of graphs

Figure 1.3. Distribution of the number of descents converging on $\widehat{f_G}$ solutions.

here, the situation is different. The improvement is probably a combination of two effects. It is well-known that for some subgraphs, the barycenter is better adapted than the median and vice versa (Di-Battista et al. 1999). And there may also be a side effect due to the chosen median definition which does not explicitly take the parity into account. The following results include improvements caused by the barycenter operator.

For each graph G of Δ', the minimum arc crossing reached on the 5000 runs is denoted by $\widehat{f_G}$. We cannot affirm that this value is actually a global optimum on Ω^G but the application of a genetic algorithm described in Lehn (2000) confirms that no better optimum is reached in 92% of the cases.

1.3.2   Results

The percentage of descents converging on $\widehat{f_G}$ solutions for the graph set Δ' is given on figure 1.3. The average rate of descents converging on a $\widehat{f_G}$ solution is quite low (28.6%) but there are important differences on Δ' (the standard deviation is 0.21): for 45% of the graphs more than 80% of the descents lead to a degenerated solution whereas for 19% of the graphs more than half of the descents converge to a best solution.

For graph drawing, a density coefficient which is function of the arc and vertex cardinalities is often introduced for graph class discrimation. The usual definition of the density $d(G)$ of a graph G is given by $d(G) = m/n(n-1)$ where the denominator is the arc number in a complete graph of size n. However, this definition is not well adapted to layered digraphs where the maximal graph has $m_{max} = \sum_{k=2}^{K} n_{k-1} \times n_k$ arcs. Hence, we here consider the following density function $d'(G) = m/m_{max}$.

For graphs very close to the two extrema of $d'(G)$ (which correspond respectively to a tree and a maximally connected digraph), the drawing problem
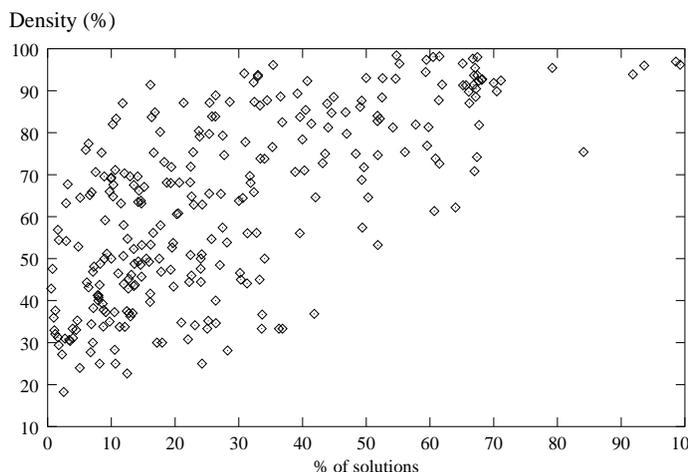
Density (%)

Figure 1.4.   Comparison between density and the percentage of $\widehat{f_G}$ solutions in a multistart descent. Each print corresponds to one graph of Δ'.

is quite simple as few permutations are possible. But, the situation is more confused for intermediate values. The relationship between the percentage of descents leading to a $\widehat{f_G}$ solution (x-axis) and the density of the graphs (y-axis) is described on figure 1.4. The correlation between these two variables is non-null ($\rho = 0.35$). The results confirm that the density is an important factor of discrimination but, as the dispersion is high, it should be completed by other criteria to explain the results.

1.4   Fitness landscape structure

The concept of fitness landscape, generally attributed to Wright (1932) for his studies on evolution, is an efficient model for understanding the dynamic of various heuristics. The definition proposed by Jones and Forrest (1995b) in the context of Genetic Algorithms is based on three components: a solution coding, a fitness function which associates each solution with a numerical value, and an operator which defines a neighborhood on the solution set. The fitness landscape can be represented by a graph LG_O -called *landscape-graph* thereafter- whose vertices are solutions and where arcs describe an O-transformation of a solution into another: there exists an arc between two vertices s and s' if $s' = O(s)$.

1.4.1   Landscape-graph for the drawing problem

In our case, vertices of LG_O are drawings of Ω^G and arcs represent transformations of a drawing into another by operators O described in Section 2. A

local optimum of a landscape is a vertex $D_i \in \Omega^G$ s.t. $f(D_i) \leq f(D_j)$ for all $D_j \in N_O(D_i)$ and a *global optimum* is a vertex $\widehat{D} \in \Omega^G$ s.t. $f(\widehat{D}) \leq f(D_i)$ for all layouts $D_i \in \Omega^G$. The *basin of attraction* of a vertex D_n is the set of vertices

$$B_O(D_n) = \Big\{ D_0 \in \Omega^G; \exists D_1, ..., D_{n-1} \in \Omega^G$$
$$\text{with } D_{i+1} \in N_O(D_i) \text{ and } f(D_{i+1}) < f(D_i), \forall i = 0, n-1 \}$$

For instance, figure 1.5 represents the landscape-graph of the graph represented on figure 1.7 for the operator $O1$. In this representation, vertices are arranged on layers: left vertices are drawings with numerous arc crossings whereas right vertices are drawings with fewer arc crossings. Let us remark that graph-landscapes here are not necessarily connected.

Local optima are represented by ovals and global optima by lozenges. It is very important to note that, according to the above definition adopted by many authors, the basins of attraction of the local and global optima may have here a non-null intersection. For instance, the gray vertex $[[a, b][c, e, d][h, g, f]]$ on figure 1.5 with four arc crossings belongs to the basin of attraction of the local optimum $[[a, b][c, e, d][f, h, g]]$ and to the basin of attraction of the global optimum $[[b, a][e, d, c][h, g, f]]$ (See figure 1.6). It is obviously possible to restrict the previous definition by considering that a drawing D is in the basin of attraction $B_O(D_n)$ if local search starting from D leads to D_n only. However, this formal change must not make us forget that the situation is here completely different from the continuous case generally taken as reference. Hereafter, we study the consequence of this characteristic on the search process.

1.4.2   Exhaustive description of a family of small graphs

In order to precise the characteristics of the optima, we have computed the complete fitness landscapes for a large family of small graphs. The graph size limitation is here imposed by the computation time required by an exhaustive exploration. For that reason, we restrict ourselves to a class Δ'' of 1875 non trivial small graphs built like in the previous section with $\prod_{k=1}^{K} n_k! \leq 2000$, and we consider the greedy operator $O1$. Note that the average vertex number of the graph-landscape is equal to 925 in this case, which remains practicable. Different statistical approaches have been proposed to analyze the distribution of the peaks and measure landscape ruggedness. The fitness distance correlation (Jones and Forrest 1995a) is one of the most popular measurements. It relies on a preliminary choice of a metric on the landscape. In our case, there are several possibilities for defining a distance between permutations, and we have chosen one of the simplest based on changes on the permutation. Let D_i and D_j be two drawings of Ω^G and σ_{ki} and σ_{kj} their respective vertex orderings on each layer $l_k \in L$. The distance between D_i and D_j is defined as
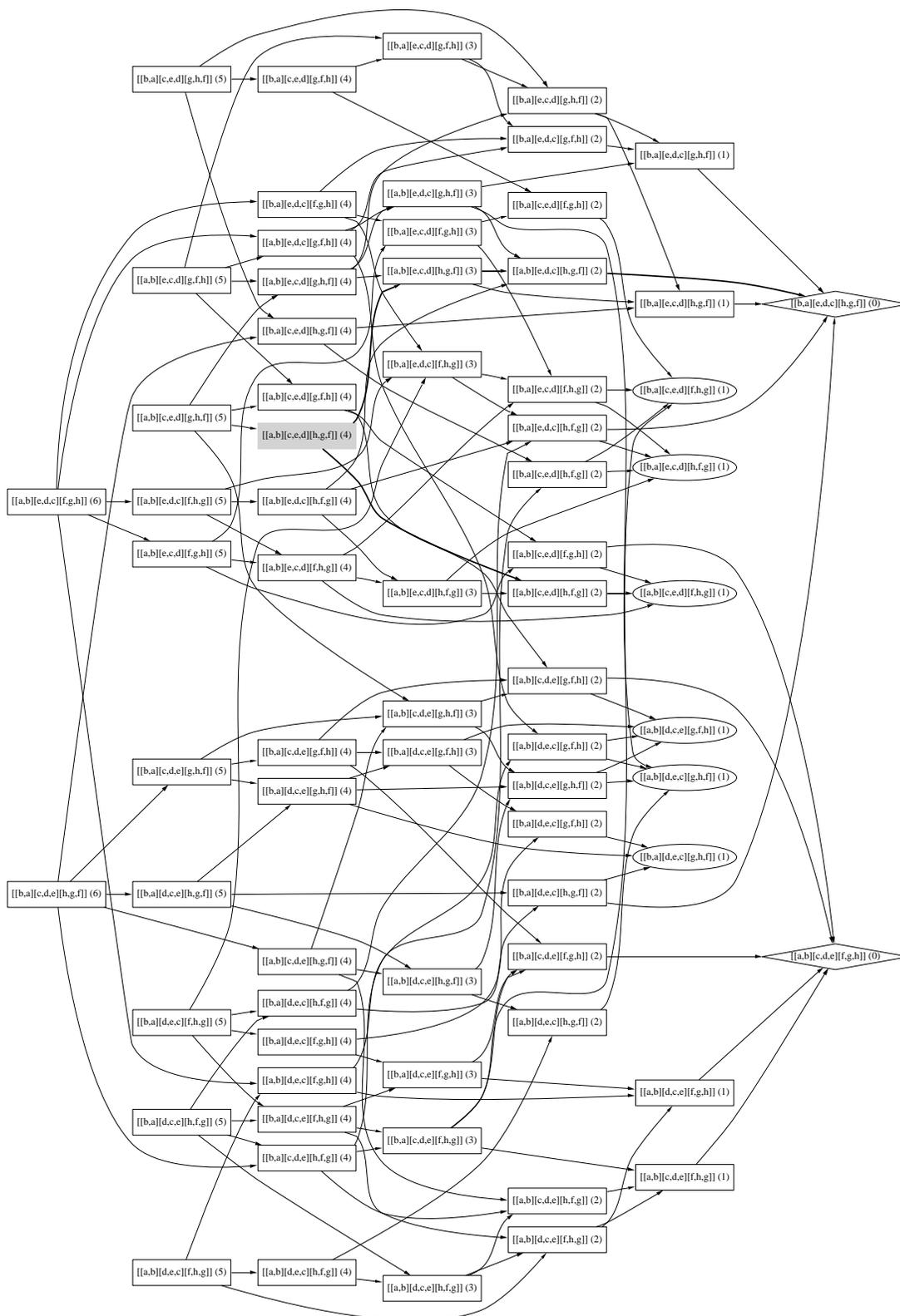
Figure 1.5.   Representation of a landscape-graph (operator O1). Vertices represent different
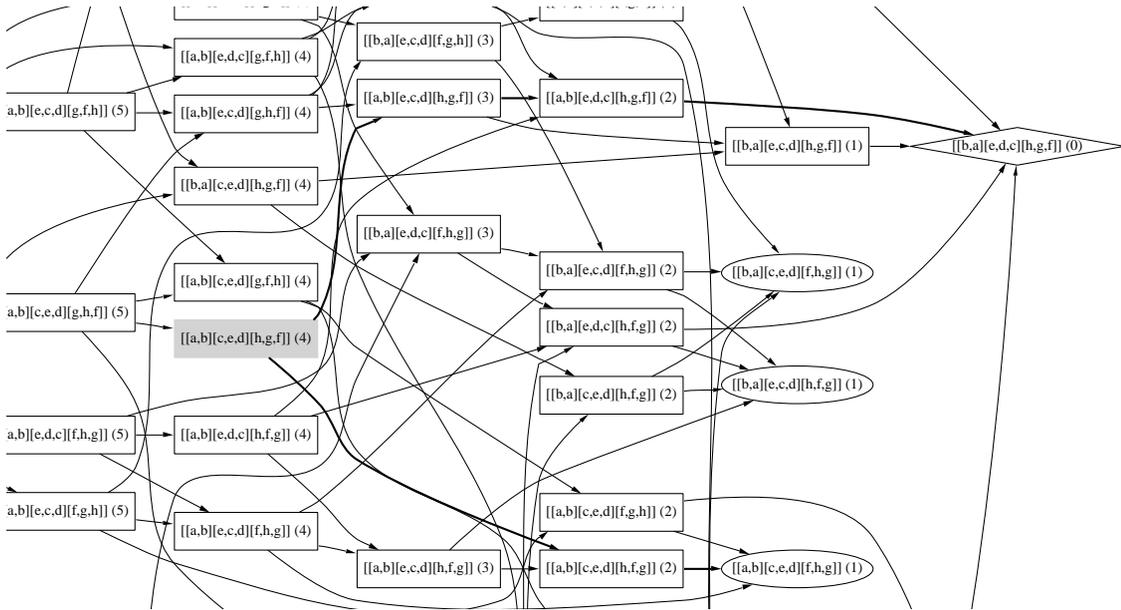drawings of a graph with 3 layers and 8 vertices (2 on L_1, 3 on L_2 and 3 on L_3). For instance,

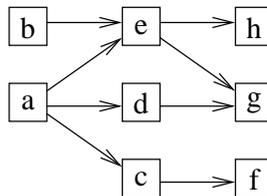*Figure 1.6.*   Zoom on the gray vertex of figure 1.5

Figure 1.7.   A vertex of the landscape-graph of figure 1.5 which is a drawing of a global optimum.
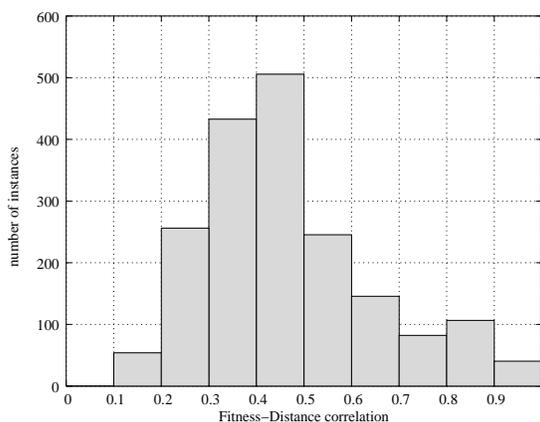
Figure 1.8.   Distribution of the fitness distance correlation

$$d\left(D_i, D_j\right) = \sum_{k \in K} |C(\sigma_{ki}, \sigma_{kj})| \text{ where } C\left(\sigma_{ki}, \sigma_{kj}\right) = \{u; \sigma_{ki}(u) \neq \sigma_{kj}(u)\})$$

The mean fitness distance correlation is equal to 0.47. This value is quite high which is partly explained by the size of the instances. Nevertheless, even for these small instances, the variation may be important (figure 1.8). Our exhaustive computation of the landscapes allows to complete the information measured by this global index by an analysis of the exact distribution of the optima. There exists at least one local optima in 76% of the cases, and most of the landscapes are strongly multimodal: the mean number of local optima is equal to 34.9 and that of the global optima is equal to 36.1, which can be partly explained by the numerous drawing symmetries. The distribution of the optima greatly varies from graph to graph: the standard deviation of the number of local optima (resp. global optima) is equal to 37.6 (resp 79.8).

In an optimization process, it is important to know whether the local optima values are close or not to the global one. Hence, we introduce the relative height $h(D_i)$ of a local optimum D_i : $h(D_i) = 1 - \frac{f(D_i) - \widehat{f}}{f_w - \widehat{f}}$ where $\widehat{f}$ (resp. f_w) is the optimal (resp. the worst) f value on the landscape. If $h(D_i)$ is very close to 1 then the local optimum D_i can be considered as an acceptable solution for some applications. The distribution of the relative heights on Δ'' is given on figure 1.9.

1.4.3   Convergence on local optima

In order to better understand the influence of the shapes of the basins of attraction in the discrete case, we study the probability distribution of being stuck on a local optima. Here, we simply consider a greedy heuristic: *1)* select
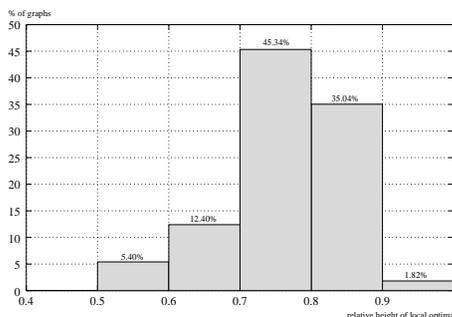
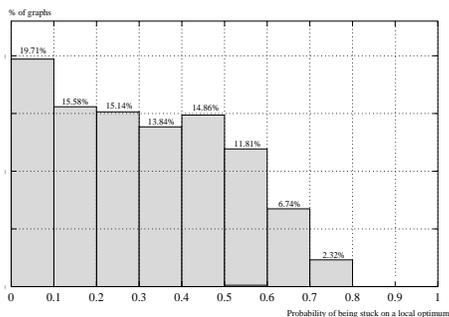*Figure 1.9.* Relative height of local optima

Figure 1.10. Probability of being stuck on a local optimum

a vertex of LG_O different from an optimum; *2)* choose the best neighbor. If two neighbors have the same f value, select one of them randomly; *3)* iterate the process until it converges. This heuristic is applied for each vertex of LG_O different from an optimum. Let n_{no} be their number and n_l be the number of runs leading to a local optimum. The probability p_l of being stuck on a local optimum is given by n_l/n_{no}. The average probability on Δ'' is equal to 0.31. Since this calculation considers small graphs only and that all neighbors are known for each vertex, this value is far from being negligible. Again, we find important differences between graphs (figure 1.10).

In particular, two classes are opposite: an easy one where 19.7% of the graphs have a probability p_l less than 0.1 and a difficult one (20.9% of graphs with $p_l > 0.5$) which requires more adapted search processes.

Due to the shape of the basins of attraction presented in Section 1.4.1, three different situations may be considered for the n_{no} starting vertices. If they belong to the basin of attraction of a local (resp. global) optimum only, the process obviously converges on this local (resp. global) optimum. But, if they belong to the intersection of two basins of attraction, the result can vary depending on the random selection at step *2)*. In order to measure the conse-

quence of this characteristic, we compute the probability to converge on a local optimum when the starting vertex is only in the intersection. The new average probability is equal to 0.14 and can exceed 0.2 for 32% of the graphs. These non-null values clearly highlight the complexity of the search space for this kind of problem.

1.5   Conclusion

The main conclusions which can be drawn from the numerical experiments presented in this paper are concerned with two points: the variability of the solution quality depending on the instance characteristics, and the properties of the basins of attraction of the optima.

Our experiments with multi-start descents have confirmed that the heuristic performances closely depend on the processed instances. Due to the huge size of the search spaces, the conclusions of such a statistical study must obviously be cautiously considered. However, they allow to show that the debate recently opened on the respective performances of Evolutionary Algorithms and Stochastic Hill Climbing for graph drawing (Rosete-Suarez et al. 1999) cannot dispense with a preliminary study of the instance characteristics. In this perspective, we have studied the influence of the graph density on the descent convergence. The correlation between the density and the descent performance is not null but it is too low to discriminate different graph categories. A manual perusal of a part of the results shows that symmetries seem to play a non negligible role for this problem. However, the detection of symmetries in graphs is generally a difficult problem (Manning 1990) and the definition of a global measure is not easy. Nevertheless, we believe that determining a typology based on a small set of discriminant quantitative criteria is a necessary preliminary work to seriously compare the different meta-heuristics mentioned in the introduction.

Our exhaustive study of the graph-landscapes has highlighted the complexity of the basins of attraction, and it contributes to justify the use of a " highly " stochastic approach. We have explicitely showed that, contrary to the continuous image where all basins of attraction are clearly separated, they are here interwoven. Consequently, for some drawings, a simple descent can lead either to a global optimum or a local one depending on a random choice of the neighbors in case of equality of the fitness. For the paradigmatic continuous case, most strategies focus on the search of the global optimum attraction set, and also sometimes, on the reduction of the number of moves necessary to reach the optimum (Johnson et al. 1988). However, in addition, it is here necessary to resort to a strategy able to move away from neighbors of a local optimum "late" in the search process.

Putting this property into practice depends on the meta-heuristic. For instance, for Tabu Search, the evolution of the tabu list during the search process - which has not still been applied to this problem- may be promising: the list size or the criterium for selecting forbidden directions could vary to get out basins of attraction of local optima very close to basins of attraction of global optima. In the perspective of the development of a genetic algorithm, which has initially stimulated the work presented here, it is clear that the adjustment of the mutation criterium could play a particularly important part. Nevertheless, a similar study should be carried for the crossover landscapes. Indeed, it is easy to find examples where the combination of two sub-graphs (for instance exchange right and left parts of a graph) can lead to a better drawing. Consequently, for the layered drawing problem, GA could additionally fully exploit the property of reconstruction by recombination.

Bibliography

G. Di-Battista, P. Eades, R. Tamassia, and I.-G. Tollis. *Graph drawing – Algorithms for the visualization of graphs*. Prentice-Hall, 1999.

G. Di-Battista, P. Eades, R. Tamassia, and I.-G. Tollis. Algorithms for drawing graphs: an annotated bibliography. *IEEE Trans. on Visualization and Computer Graphics*, 6(11):24–43, 2000.

P. Eades and D. Kelly. Heuristics for reducing crossings in 2-layered networks. *Ars Combinatorics*, 21:89–98, 1986.

P. Eades and N. Wormald. Edge crossings in drawings of bipartite graphs. *Algorithmica*, 11:379–403, 1994.

M.-R. Garey and S. Johnson. Crossing number is *NP*-complete. *SIAM J. of Algebraic and Discrete Methods*, 4(3):312–316, 1983.

I. Herman, G. Melançon, and M.S. Marshall. Graph visualization and navigation in information visualization: a survey. *IEEE Trans. on Visualization and Computer Graphics*, 6(11):24–43, 2000.

D.S. Johnson, C.H. Papadimitriou, and M. Yannakakis. How easy is local search ? *J. of Computer and System Sciences*, 17:79–100, 1988.

T. Jones and S. Forrest. Fitness distance correlation as a measure of problem difficulty for genetic algorithms. In *L. Eshelman (ed.): Proc. of the Sixth Int. Conf. on Genetic Algorithms*, pages 184–192. Morgan Kaufmann, 1995a.

T. Jones and S. Forrest. Genetic algorithms and heuristic search. In *Santa Fe Institute Tech. Report 95-02-021*. Santa Fe Institute, 1995b.

T.-C. Jones. *Evolutionary Algorithms, Fitness Landscapes and Search*. PhD thesis, University of New Mexico, Alburquerque, 1995.

M. Jünger and P. Mutzel. 2-layer straightline crossing minimization: performance of exact and heuristic algorithms. *J. of Graph Algorithms and Applications*, 1(1):1–25, 1997.

Pascale Kuntz, Fabrice Guillet, Rémi Lehn, and Henri Briand. A user-driven process for mining association rules. In *Principles of data mining and knowledge discovery*, pages 483–489. LNCS 1910, 2000.

M. Laguna, R. Marti, and V. Valls. Arc crossing minimization in hierarchical design with Tabu Search. *Computers and Operations Res.*, 24(12):1175–1186, 1997.

R. Lehn. *An interactive system of visualization and rule mining for knowledge discovery in databases (in french)*. PhD thesis, University of Nantes, France, 2000.

R. Lehn and P. Kuntz. A contribution to the study of the fitness landscape for a graph drawing problem. In *Proc. of EvoCop2001 - Applications of Evolutionary Computing*, pages 172–181. LNCS 2037, 2001.

J. Manning. *Geometry symmetry in graphs*. PhD thesis, University of Purdue, New-York, 1990.

R. Marti. Arc crossing minimization in graphs with GRASP. *IIE Trans.*, 33 (10):913–919, 2001.

H. Purchase. Effective information visualisation: a study of graph drawing aesthetics and algorithms. *Interacting with computers*, 13(2):477–506, 2000.

A. Rosete-Suarez, A. Ochoa-Rodrigez, and M. Sebag. Automatic graph drawing and stochastic hill-climbing. In *Proc. of the Genetic and Evolutionary Conf.*, pages 1699–1706. Morgan Kaufmann, 1999.

K. Sugiyama, S. Tagawa, and M. Toda. Methods for visual understanding of hierarchical systems. *IEEE Trans. Sys. Man and Cybernetics*, 1981.

J. Utech, J. Branke, H. Schmeck, and P. Eades. An evolutionary algorithm for drawing directed graphs. In *Proc. of the Int. Conf. on Imaging Science, Systems and Technology*, pages 154–160. CSREA Press, 1998.

S. Wright. The roles of mutation, inbreeding, crossbreeding and selection in evolution. In *Proc. of the Sixth Int. Conf. on Genetics*, volume 1, pages 356–366, 1932.