MINISTÉRIO DA CIÊNCIA E TECNOLOGIA
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

INPE-9375-PRE/5035

INDUSTRIAL PATTERN-SEQUENCING PROBLEMS: SOME COMPLEXITY RESULTS AND NEW LOCAL SEARCH MODELS

Alexandre Linhares
Horacio Hideki Yanasse

Trabalho apresentado no concurso de Teses e Dissertações – CDT 2002 , Florianópolis, 15 a 19 de julho de 2002.

INPE
São José dos Campos
2002

Industrial pattern-sequencing problems: some complexity results and new local search models

Alexandre Linhares[1,2], Horacio Hideki Yanasse (orientador)[1]

[1]Instituto Nacional de Pesquisas Espaciais, Laboratório Associado de Computação e Matemática Aplicada – C.P. 515– São José dos Campos – SP 12201-970– Brazil

[2]Fundação Getulio Vargas, EBAPE, Praia de Botafogo 190, Rio de Janeiro

linhares@fgv.br, horacio@lac.inpe.br

Abstract. *In this thesis we explore some industrial pattern sequencing problems arising in settings as distinct as the scheduling of flexible machines, the design of VLSI circuits, and the sequencing of cutting patterns. This latter setting presents us the* minimization of open stacks problem, *which is the main focus of our study. Some complexity results are presented for these sequencing problems, establishing a surprising connection between previously unrelated fields. New local search methods are also presented to deal with these problems, and their effectiveness is evaluated by comparisons with results previously obtained in the literature. The first method is derived from the simulated annealing algorithm, bringing new ideas from statistical physics. The second method advances these ideas, by proposing a collective search model based on two themes: (i) to explore the search space while simultaneously exhibiting search intensity and search diversity, and (ii) to explore the search space in proportion to the perceived quality of each region. Some preliminaries, given by coordination policies (to guide the search processes) and distance metrics, are introduced to support the model.*

Resumo. *Nesta tese nós exploramos alguns problemas industriais originários de tarefas tão distintas quanto a programação de uma máquina flexível, o projeto de circuitos integrados VLSI e o sequenciamento de padrões de corte. Desta última tarefa provêm o problema de minimização de pilhas em aberto que é o principal foco do estudo. Alguns resultados de complexidade são apresentados para estes problemas estabelecendo-se uma conexão entre áreas que previamente não pareciam estar relacionadas. Novos modelos de busca local também são apresentados para lidar com estes problemas e sua eficácia é medida por comparações com resultados previamente estabelecidos na literatura. O primeiro método é derivado do algoritmo de* simulated annealing *trazendo conceitos da física estatística. O segundo método expande estas idéias, propondo um modelo coletivo baseado em dois princípios: (i) a exploração do espaço de soluções exibindo uma busca simultaneamente intensiva e distribuída e, (ii) a concentração da busca em proporção com a qualidade percebida em cada região. São também introduzidos novos métodos essenciais para dar suporte ao modelo, como políticas de coordenação (dos processos de busca) e métricas de distâncias.*

1. Introduction

Yanasse (1997) identified an industrial problem arising in the sequencing of cutting patterns, termed *minimization of open stacks problem* – hereafter MOSP. Consider a production setting where J distinct patterns need to be cut, each one containing a combination of at most I piece types. We can define the piece-pattern relationship by a I x J binary matrix $P=\{p_{ij}\}$, with $p_{ij}=1$ if pattern j contains piece type i, and $p_{ij}=0$ otherwise. When a pattern is cut, the pieces are stored on stacks that remain fixed around the cutting saw until each stack is completed. Each stack holds only pieces of the same type and it remains open until the last pattern containing that piece type is cut. Difficulties in handling a larger number of distinct open stacks appear, for instance, if the number increases beyond the system capability, so some of the stacks must be temporarily removed, yielding additional costs, higher production time, and higher associated risks (as observed in glass-cutting settings). We are thus interested in the sequencing of the patterns to minimize the maximum number of open stacks during the cutting process. An open stacks versus cutting instants matrix $Q^{\pi}=\{q^{\pi}_{ij}\}$ is defined by:

$$q^{\pi}_{ij}=\begin{cases}1, & if \exists x, \exists y \mid \pi(x)\leq j\leq \pi(y), and\ p_{ix}=p_{iy}=1\\ 0, & otherwise\end{cases} \quad (1)$$

where π denotes a permutation of the $\{1,2,\ldots,J\}$ numbers, and defines a sequence on which the patterns are cut, such that $\pi(i)$ is the order (instant) in which the i-th pattern is cut. Note that matrix Q^{π} holds the consecutive-ones property for columns under permutation π: in each row, any zero between two ones will be changed to one. This is called a *fill-in*. Since we are interested in minimizing the number of simultaneous open stacks, we define the following cost functional:

$$Z^{\pi}_{MOSP}(P)=\max_{j\in\{1,\ldots,J\}}\sum_{i=1}^{I}q^{\pi}_{ij} \quad (2)$$

and define the *minimization of open stacks problem* (MOSP) as the problem of $\min_{\pi\in\Gamma} Z^{\pi}_{MOSP}(P)$, where Γ is the set of all possible one-to-one mappings $\pi:\{1,2,\ldots,J\}\rightarrow\{1,2,\ldots,J\}$. In Yanasse (1997), a mathematical model was proposed and some theoretical conjectures were raised, including one concerning the computational complexity of the MOSP: is MOSP NP-Hard? This question is the starting point to our research.

Proposition 1. *MOSP is NP-Hard.*

Proof. Reduction of Modified Cutwidth to MOSP.

MODIFIED CUTWIDTH (MCUT) [3]

INSTANCE: Graph $G=(V,E)$, positive integer K.

QUESTION:Is there a one-to-one function π: $V\rightarrow\{1,2,\ldots,|V|\}$ such that for all i, $1<i<|V|$,

$$|\ \{(u,v)\in E: \pi(u)< i<\pi(v)\}\ |\leq K \quad ?$$

This problem asks for a linear ordering of the vertices of G on a straight line such that no orthogonal plane that intersects the line will cross more than K edges of G. (It differs from the *minimum cut linear arrangem7ent problem,* also known as *cutwidth,* by the use of the "<" sign instead of the "≤") The optimization version of MCUT asks to minimize K over all possible linear orderings of the vertices of G, i.e., $\min_{\pi\in\Gamma} Z^{\pi}_{MCUT}(G)$, where $Z^{\pi}_{MCUT}(G) = \max_{1<i<|V|} |\{(u,v) \in E : \pi(u) < i < \pi(v)\}|$. We will show how to solve an instance $G=(V,E)$ of MCUT with an algorithm for MOSP. First, construct a corresponding I x J matrix P where $I=|E|$ and $J=|V|$ where we associate each edge in E to a row number and each vertex of V to a column number. Consider an edge $(u,v) \in E$, $u\in V$, $v\in V$. Let i be the row number assigned to this edge and let j and k be the number of the columns assigned to vertices u and v respectively. Make $p_{ij}=p_{ik}=1$, and $p_{il}=0$, for all $l\neq j$, $l\neq k$.

Lemma 1. *Each edge crossing in a linear layout of the vertices of G corresponds to one and only one "fill-in" to problem matrix Q^{π}, associated to a permutation of the columns of P with the consecutive-ones property for columns.*

Proof. Consider an ordering π of the vertices of G and an edge (u,v) crossing an orthogonal plane at vertex i. Let x be the row in matrix P associated with this edge. From the construction of P, we have $p_{x,\pi(u)}=p_{x,\pi(v)}=1$, and $p_{x,y}=0$ for all $y<\pi(u)$ and for all $y>\pi(v)$. At this row, following the column sequence that corresponds to the vertex ordering, there must be a fill-in at position (row x, column i), since the associated matrix Q^{π} holds the consecutive-ones property for columns under permutation π. Now that this basic relation is established, all that must be done is to ensure that the maximum column sum of the MOSP reflects only the fill-ins. In order to achieve this, we create a matrix P' as follows: first, let $p'_{ij} = p_{ij}$, $\forall i\in\{1,2,...,I\}$, $\forall j\in\{1,2,...,J\}$. Let $C = \max_{j}\sum_{i=1}^{I} p_{ij}$, and compute a gap $g = CJ - \sum_{j=1}^{J}\sum_{i=1}^{I} p_{ij}$. The rows I+1, I+2, ..., I+g of P' have only a single nonzero element equal to 1. They are distributed over the columns in rows I+1, I+2, ..., I+g in such a way that there are exactly $C - \sum_{i=1}^{I} p_{ij}$ ones on each column so that columns of P' have column sum equal to C. Now, for the corresponding MOSP instance P' [of dimension (I+g) x J], we have:

Lemma 2. $Z^{\pi}_{MOSP}(P') = Z^{\pi}_{MCUT}(G) + C$.

Proof. Observe that each column of P' has the same number of 1's. Thus, when solving MOSP what is being minimized is the maximum column sum of fill-ins. By lemma 1, the solution to this instance of MOSP solves the corresponding instance of MCUT.

Proposition 1, demonstrating that there should be no efficient exact algorithm for minimization of open stacks, closes the conjecture posed earlier [Yanasse, 1997] (it should be mentioned that the other conjectures placed by Yanasse [1997] have also been clarified, we refer the reader to chapter 2 of the thesis and also to [Linhares and Yanasse, 2002]).

2. New local search models

In light of this NP hardness result, it becomes worthwhile to invest in heuristic solution methods, which obtain good but not provably optimal results. We have resorted thus to developing a *microcanonical optimization algorithm* which outperformed the previous methods in the literature (we refer the reader to chapter 3 of the thesis, and to [Linhares et al., 1999]). Microcanonical optimization is based on two alternating modes of execution; we may say that one mode is concerned with intensification of the search (to the best solutions in an area), while another mode is concerned with the diversification of the search (to other areas) (Linhares et al., 1999). Many heuristics have these 'dual opposites' (Glover, 1990; Linhares, 1998). There is, in fact, in the local search literature, a widely held belief:

Claim 1. ***The intensity versus diversity belief.*** *There is a widespread belief that local search intensity and local search diversity are mutually exclusive; a belief that metaheuristic algorithms cannot achieve both of these desirable attributes under perfect simultaneous co-existence.*

This belief can be readily seen in several quotations from the literature. In the context of genetic algorithms (applied to optimization), for example, one learns that "*a high crossover probability will encourage steady hill-climbing towards the optimal set of parameter values. On the other hand, a high rate of mutation will result in more searching of the less promising parts of the search space*" (Backhouse et. al, 1997); the keyword here being "*on the other hand*". In another genetic algorithm quote, it is mentioned that mechanisms "*designed to increase the pressure for improvement may be at the expense of population diversity. While such strategies can improve the results on small to moderate problems, in large problems they may not allow sufficient exploration of the solution space*" (Dowsland 1996). In other words, as one gets more intensity, one loses precious diversity.

In the context of tabu search, Laguna and Glover (1993) have pointed out, for instance, that "*effective tabu search procedures keep a balance between intensification and diversification, that is, between reinforcing attributes associated with good solutions and driving the search into regions not yet visited*". Other articles use the explicit term 'tradeoff', as in "*intensification/diversification tradeoffs*" (Glover 1990). Similar remarks have appeared in the context of scatter search, where (Cung et al. 1996) point out that "*...as indicated by its name, the method induces a real willingness for maintaining the collection points as scattered as possible, hence to have a good diversification. However, intensification can also be achieved...*" (Cung et al. 1997).

Claim 1 can be consolidated after the following quote from Colorni et. al (1996): "Two main features *have to be balanced* in constructing heuristic algorithms: (i) The degree of *exploitation*, that is, the amount of effort directed to local search in the present region of the search space (if a region is promising, search more thoroughly); (ii) The degree of *exploration*, that is, the amount of effort spent to search in distant regions of the space (sometimes choose a solution in a far region and/or accept a worsening one, to gain the possibility of discovering new better solutions). These *two possibilities are conflicting*: a *good trade-off* between them is very important and must be carefully tuned in each algorithm." (Colorni et al. 1996; emphasis ours).

Thus, this *intensity versus diversity belief* is indeed pervasive in the literature. But we should place the question: is it true? Is it impossible to obtain search intensity and search diversity in perfect simultaneous co-existence? This is an implicit, unstated, hypothesis assumed in a large mass of literature, and it should deserve proper attention.

We may thus challenge this assumption. A trivial fact is that search intensity is clearly easy to achieve. Consider, for instance, a version of *tabu search* or of *microcanonical optimization* (Glover, 1990; Linhares and Torreão, 1998; Linhares et al., 1999): these search methods place increasing pressure on the vicinity of each search region, and can thus be classified as *search intensive*. However, these methods must alternate from the mode of search intensity whenever a locally minimum solution is obtained. Since further improvement from that point is simply not possible, the algorithms accept non-improving solutions and enter a phase designed to explore the space in a diverse manner. This alternation between search intensity and search diversity seems at first to give more credit to the *intensity versus diversity belief*.

Now consider a system that manages to have multiple search processes in parallel, each process following a simple tabu search algorithm. It is easy to claim the following:

Claim 2. *The proposed model exhibits high search intensity during the course of its execution.*

It is easy to demonstrate this fact because each individual process follows an intensive search, and thus (as a trivial extension) the collective system exhibits search intensity. Note, however, that one cannot guarantee the system to simultaneously display search diversity, as all search processes can in principle cluster around a particular region of the search space. We must introduce two new concepts: (i) distance metrics, and (ii) coordination policies.

Coordination policies are guiding principles for coordinating the collective search among distributed interdependent processes. These policies use information obtained from so-called *distance metrics*. Consider a collective system with two distributed search processes. At any point in time, these processes hold a *minimum number of steps* between them; this number will define the distance between the processes.

It is surprising that distance metrics have not been used extensively in local search. Distances can reveal important facts about a pair of solutions: they give an idea of the work and of the time required to move from one solution to the other; they are correlated with the probability that the second solution will be visited by a process that passed trough the first; and they also give a measure of similarity between any two solutions. In our work we will be interested in demonstrating that the *intensity versus diversity belief* is false (despite being widely held). In order to demonstrate so, we must introduce *non-overlapping coordination policies*.

Consider a set of distributed search processes that individually obey claim 2. With the use of distance metrics, it is possible to force these processes to occupy mutually exclusive areas of the search space, by keeping a minimum distance from the nearest search process. By managing this minimum distance, it is possible to effectively devise a system that displays the following property:

Claim 3. *The proposed model exhibits high search diversity during the course of its execution.*

The processes do not enter into each other's area, and thus the system displays high search diversity at the collective level. This is illustrated in Fig. 1.

Figure 1. In the proposed collective model, processes are required to maintain a minimum distance to each other (the boundaries of each process are represented by circles). Thus, if each process executes an intensive search, the collective system exhibits high search intensity and high search diversity in simultaneous co-existence, disproving the so-called "*intensity versus diversity*" belief (see claims 1 and 4).

Since diversity is enforced by the minimum required distances, and as each individual process executes an intensive search process, we come to the conclusion that

Claim 4. *Despite the fact that the 'intensity versus diversity belief' is widely held,* it is false*: local search intensity is not mutually exclusive with local search diversity.*

As we have seen, this crucial conclusion directly contradicts a large mass of literature (see for instance, Backhouse et al., 1997; Colorni et al., 1996; Cung et al., 1997; Dowsland, 1996; Glover, 1990; Laguna and Glover, 1993).

Finally, we must point out that other, more intricate, coordination policies have been proposed, and once again we must, for utter lack of space, refer the interested reader to chapter 4 of the thesis and also to (Linhares and Yanasse, 2002, under review).

3. Distance metrics

Of course these coordination policies require the use of underlying mechanisms of *distance metrics*. That is, given a pair of solutions *x* and *y*, what is the minimum number of local search movements between this pair of solutions? Since these metrics must be computed under polynomial time, a discouraging result comes from the NP-completeness of the 2-opt neighborhood for the traveling salesman problem (Linhares and Yanasse, 2002, under review).

Fortunately, for other widely used neighborhood operators, there are efficient algorithms. We have thus developed an efficient algorithm for the 2-exchange operator, which executes in linear time using an associated structure referred to as a *switch graph* (Linhares and Yanasse, 2002, under review). We have also developed a quadratic algorithm for the insertion operator for permutations, based on a data structure we refer to as the *traversal trees*. As required, we have also proved the optimality of the obtained metrics. These metrics are surprisingly related to important problems in the field of computational molecular biology, however, given the lack of space to present these algorithms, data structures, and associated theorems, we must refer the reader to chapter 5 of the thesis and to Linhares and Yanasse (2002, under review) for details.

Colorni, A., Dorigo, M., Maffioli, F., Manniezzo, V., Righini, G., and M. Trubian, (1996) Heuristics from nature for hard combinatorial optimization problems, *International Transactions on Operational Research*, 3, 1-21.

Cung, V.-D., Mautor, T., Michelon, P., and A. Tavares (1997) "A scatter search based approach for the quadratic assignment problem," In *Proceedings of the IEEE-ICEC'97 International Conference on Evolutionary Computation*, pages 165-170.

Dowsland, K. (1996) Genetic algorithms: a tool for OR?, *Journal of the Operational Research Society* 47, 550-561.

Glover, F. (1990) Tabu search: a tutorial, *INTERFACES* 20, 74-94.

Hong, Y.S., Park, K.H. and M. Kim (1989) A heuristic for ordering the columns in one-dimensional logic array. *IEEE Transactions on Computer-Aided Design* 8, p. 547-562.

Hu, Y.H. and S.J. Chen (1990) GM_Plan: a gate matrix layout algorithm based on artificial intelligence planning techniques. *IEEE Transactions on Computer-Aided Design* 9, p. 836-845.

Laguna, M., and F. Glover (1993) Bandwidth packing: a tabu search approach, *Management Science* 39, 492-500.

Linhares, A. (1998) State-space search strategies gleaned from animal behavior, *Biological Cybernetics* 78, 167-173.

Linhares, A. (1999) Synthesizing a predatory search strategy for VLSI Layouts, *IEEE Transactions on Evolutionary Computation* 3, 147-152.

Linhares, A. (2000) A glimpse at the metaphysics of Bongard problems, *Artificial Intelligence* 121, 251-270.

Linhares, A. (2001) *Industrial pattern sequencing problems: some complexity results and new local search methods*, Doctoral Dissertation, Brazilian Space Research Institute, Laboratory for Computing & Applied Mathematics, São José dos Campos, SP, Brazil. Available from http://www.pgrad.inpe.br/~linhares/DoctoralThesis.PDF.

Linhares, A., (2002) Minimization of open stacks: application of a re-distribution coordination policy. Manuscript in preparation.

Linhares, A and Torreão, JRA (1998) Microcanonical optimization applied to the traveling salesman problem, *International Journal Modern Physics* C9, 133-146.

Linhares, A., and H.H. Yanasse (2002) Local search intensity versus local search diversity: a false tradeoff?, under review in *Management Science.*

Linhares, A., and H.H. Yanasse (2002) Connections between cutting-pattern sequencing, VLSI design, and flexible machines, *Computers & Operations Research* 29, 1759-1772.

Linhares, A., H.H. Yanasse, and J.R.A. Torreão (1999) Linear gate assignment: a fast statistical mechanics approach, *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* 18, 1750-1758.

Yanasse, H.H. (1997) On a pattern sequencing problem to minimize the maximum number of open stacks. *European Journal of Operational Research*, v.100, p. 454-463.

MINISTÉRIO DA CIÊNCIA E TECNOLOGIA
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

AUTORIZAÇÃO PARA PUBLICAÇÃO

Número

LAC-001/2002

Titulo

Industrial patter-sequencing problems: some complexity results and new local search models.

Autor

LINHARES, A.; YANASSE, H.H.

Tradutor

Editor

Origem	Projeto	Série	No. de Páginas	No. de Fotos	No. de Mapas
LAC	ATLAC/MODOS		08		- 976

Tipo

☒ RPQ ☒ PRE ☐ NTC ☐ PRP ☐ MAN ☐ PUD ☐ TAE

Divulgação

☐ Externa ☐ Interna ☐ Reservada ☐ Lista de Distribuição Anexa

Periódico / Evento

Concurso de Teses e Dissertações - Teses de Doutorado, CTD 2002, 15 a 19 de julho de 2002, Anais do XXII Congresso da Sociedade Brasileira de Computação, p. 550-557, em CD-ROM.

Convênio

Projeto Temático FAPESP Projeto Integrado CNPq

Autorização Preliminar

__/__/__
Data

Haroldo Fraga de Campos Velho
Chefe
Laboratório Associado de Computação
e Matemática Aplicada - LAC
Titular da Unidade

Revisão Técnica

☐ Solicitada ☒ Dispensada

Titular de Nível "A"

Recebida __/__/__ Devolvida __/__/__

Assinatura do Revisor

Revisão de Linguagem

☐ Solicitada ☒ Dispensada

Titular de Nível "A"

Recebida __/__/__ Devolvida __/__/__

Assinatura do Revisor

Autorização Final

__/__/__
Data

Haroldo Fraga de Campos Velho
Chefe
Laboratório Associado de Computação
e Matemática Aplicada - LAC
Titular da Unidade

Palavras Chave

pattern sequencing; complexity; local search models

INPE-9375-PRE-5035

(continua na folha 2/2)