INTENSIDADE VERSUS DIVERSIDADE NA BUSCA LOCAL: UMA FALSA TROCA?

Alexandre Linhares
Escola Brasileira de Administração Pública e de Empresas, EBAPE/FGV
Praia de Botafogo 190/426, Rio de Janeiro 22257-970, Brasil; linhares@fgv.br

Horacio Hideki Yanasse
Instituto Nacional de Pesquisas Espaciais, LAC/CTE/INPE,
Av Astronautas 1758, S.J.Campos, SP12227-010, Brasil; horacio@lac.inpe.br

Resumo – Uma idéia implícita nas metaheurísticas de busca local é que há uma troca mutuamente exclusiva entre dois objetivos desejáveis, a diversidade de busca e a intensidade de busca. Neste trabalho apresentamos a tese de que esses objetivos não são mutuamente exclusivos e podem ser obtidos simultaneamente. Mostramos como métricas de distância e políticas de coordenação podem ser utilizadas em conjunto para integrar um arcabouço para o desenvolvimento de métodos avançados de busca local cooperativa que apresentam as aspirações de diversidade de busca e intensidade de busca em coexistência simultânea.

Abstract – One implicit tenet of local search metaheuristics is that there is a mutually exclusive balance between two desirable objectives, search diversity and search exploitation. We claim that these objectives are not mutually exclusive and can indeed be obtained simultaneously. We show how distance metrics and coordination policies can be used together to integrate a framework for the development of advanced collective search methods that present the desiderata of search intensity and search diversity co-existing simultaneously.

1. INTRODUÇÃO: INTENSIDADE VERSUS DIVERSIDADE?

Uma das abordagens mais bem sucedidas para a solução de problemas combinatórios de escala industrial massivamente multimodais é a utilização de metaheuristísticas de busca local. Existe um senso comum que tais métodos precisam exibir intensidade e diversidade de busca para que sejam efetivos. Procura-se com a intensidade determinar a melhor solução contida em uma região limitada relativamente pequena enquanto que com a diversidade procura-se amostrar um grande número de regiões diferentes para assegurar que o espaço de soluções seja explorado de maneira satisfatória e, desta forma, a região (as regiões) contendo o ótimo (os ótimos) global (globais). Sem uma intensidade adequada, um método pode passar próximo da solução ótima sem que consiga encontrá-la e sem uma diversidade adequada um método pode ficar concentrado em regiões relativamente pobres do espaço de solução sendo incapaz de determinar soluções de melhor qualidade pois estão localizadas em outras regiões.

As metaheurísticas existentes atualmente não exibem estes objetivos desejáveis coexistindo de maneira simultânea. Considere, por exemplo, o algoritmo de simulated annealing. Como se sabe, este algoritmo baseado na termodinâmica é baseado na distribuição de Gibbs para se obter configurações (soluções) a certas temperaturas. Em altas temperaturas, todas as configurações são igualmente prováveis enquanto que em baixas temperaturas (próximas de zero), somente as configurações de energia (custo) mínima são encontradas (sob algumas condições; veja, por exemplo, van Laarhoven e Aarts, 1988). Podemos então deduzir que simulated annealing apresenta diversidade e intensidade na busca, mas não de maneira simultânea. No algoritmo avança-se gradualmente da diversidade pura, na busca encontrada em altas temperaturas (onde muitas regiões são investigadas e poucos, se alguma, boas soluções são encontradas), para uma intensidade pura na busca, encontrada quando a temperatura atinge o zero (quando nenhuma solução pior pode ser aceita não havendo nenhuma diversidade a partir deste instante). No simulated annealing inicia-se com o máximo de diversidade (e intensidade mínima), e termina-se com o máximo de intensidade (e diversidade mínima). Durante o processo existe uma troca gradual de diversidade por intensidade.

Um caso análogo ocorre com algoritmos genéticos. Estas heurísticas populacionais (Beasley 2000) começam com um conjunto diversificado de indivíduos retirados de uma amostra aleatória do espaço populacional e, gradualmente, tende-se para uma intensidade maior quando o sistema converge na direção de um grupo de indivíduos que são similares entre si, perdendo ao longo do processo uma grande quantidade de diversidade genética.

Uma terceira meta-estratégia bem conhecida, a busca tabu, usa estruturas de memórias de longo prazo para disparar o modo de diversidade na busca ou, estruturas de memória de prazo intermediário para disparar o modo de intensidade na busca (Glover 1989; Glover e Laguna 1997). Assim, espera-se diversidade e intensidade nestes algoritmos, mas de maneira alternada. Alternamentos similares ocorrem na busca difusa (Glover, 1977; Cung et al., 1997), nas fases de inicialização e amostragem da otimização microcanônica (Linhares e Torreão 1998; Linhares et al; 1999) e nas fases de construção e melhora do GRASP (Feo e Resende, 1995; Argüello, Bard e Yu, 1997; Resende e Ribeiro, 1997).

Estes modelos clássicos de busca local apresentam sugestões tênues de uma troca mutuamente exclusiva entre intensidade e diversidade na busca. Numa revisão da literatura, entretanto, esta sugestão parece ficar mais evidente pois numerosos estudos admitem esta troca, alguns implicitamente, alguns explicitamente. No contexto de algoritmos genéticos (otimização aplicada), por exemplo, aprende-se que "*uma probabilidade alta de cruzamento encoraja uma subida regular para os valores ótimos do conjunto de parâmetros. Por outro lado, uma taxa alta de mutação resulta numa busca mais intensa de partes menos promissoras do espaço de soluções*" (Backhouse et. al, 1997); a palavra chave aqui é "por outro lado". Na citação de outro algoritmo genético, menciona-se que mecanismos "*projetados para aumentar a pressão para a melhoria pode ser feita em troca da diversidade da população. Esta estratégia pode melhorar os resultados em problemas pequenos e moderados, mas em problemas grandes podem não permitir uma exploração suficiente do espaço de soluções*" (Dowsland 1996), ou seja, quando se investe mais em intensidade, perde-se em diversidade.

No contexto de busca tabu, Laguna e Glover (1993) ressaltaram que "*procedimentos de busca tabu efetivos mantêm um equilíbrio entre intensificação e diversificação*". Em outros artigos usa-se o termo específico 'troca' como em "*troca de intensificação/diversificação*" (Glover 1990).

Comentários similares apareceram no contexto de busca difusa onde Cung et al. (1997) ressaltaram que "*...como indicado pelo seu nome, o método induz o desejo real de manter a coleção de pontos o mais difuso possível, desta forma tendo uma boa diversificação. Entretanto, intensificação também pode ser obtida...*"

Nossa maior reivindicação pode ser consolidada após a seguinte afirmação de Colorni et. al (1996):

"Duas características principais *precisam ser equilibradas na construção* de algoritmos heurísticos:

- grau de *aproveitamento*, isto é, a quantidade de esforço direcionado para a busca local na região corrente do espaço de soluções (se uma região é promissora, procure mais cuidadosamente);
- grau de *exploração*, isto é, a quantidade de esforço dispendido para procurar em regiões distantes do espaço (algumas vezes escolher uma solução numa região distante e/ou aceitar uma solução pior, para ter a possibilidade de descobrir novas soluções melhores).

Estas *duas possibilidades são conflitantes*: um bom *equilíbrio* entre elas é muito importante e precisa ser calibrada cuidadosamente em cada algoritmo." (Colorni et. al, 1996; nossa ênfase)

Após estas observações, a seguinte asserção é obvia:

Asserção 1. ***A crença da intensidade versus diversidade***. *Existe uma crença difundida de que intensidade e diversidade na busca local são mutuamente exclusivas; uma crença de que em metaheurísticas estes atributos desejáveis não podem ambos coexistir simultaneamente.*

As frases apresentadas anteriormente mostram que esta crença está amplamente espalhada realmente, mas será que ela é verdadeira? Será que é *impossível*, como se afirma na proposição, obter-se diversidade e intensidade na busca coexistindo perfeitamente de maneira simultânea? O principal objetivo neste trabalho é mostrar que *intensidade e diversidade na busca não são mutualmente exclusivas* – que estas aspirações podem coexistir simultaneamente – e, desta forma, mostrar que a crença de intensidade versus diversidade é falsa.

Observe que nossa proposição não é mostrar que ter intensidade e diversidade coexistindo simultaneamente é necessariamente vantajoso em termos de desempenho do algoritmo. Não sabemos se estes métodos avançados seriam superiores aos modelos de busca local existentes; entretanto, embora isto seja uma questão empírica certamente relevante (e interessante), não está no escopo deste trabalho. O foco aqui é demonstrar que diversidade e intensidade podem coexistir simultaneamente e que 'a crença de intensidade versus diversidade' mencionada anteriormente não é verdadeira. Este é o maior discernimento deste trabalho.
Nas seções que se seguem, apresentamos o arcabouço para se projetar novos modelos de busca local que incluem altas intensidade e diversidade nas buscas em coexistência simultânea. Esses novos modelos são baseados na aplicação de novos mecanismos referidos como *políticas de coordenação* (i.e., princípios de conduta para coordenar a busca cooperativa entre os processos) e também no uso de *métricas de distância* entre soluções.

2. MÉTRICAS DE DISTÂNCIA

Um relevo de otimização combinatória é definida por uma tripla (Ω, N, Z), onde Ω é um conjunto discreto e finito de soluções, N: $\Omega \rightarrow 2^{\Omega}$ é um operador de vizinhança e Z: $\Omega \rightarrow \Re$ é a função objetivo a ser minimizada. Alcançabilidade sob N pode ser admitida, i.e., para todo x, $y \in \Omega$, existe uma sequência $x = s_1, s_2, ..., s_d = y$, onde $s_{i+1} \in N(s_i)$ para todo $1 \leq i < d$. Se d é o número mínimo de elementos em qualquer sequência possível de x para y, então dizemos que $D_N(x,y) = d$, isto é, a distância de x para y sob o operador N é d.
Surpreendentemente, estas 'métricas de distância' não tem sido utilizadas em modelos de busca local (a única exceção é o algoritmo *bionômico* (Christofides, 1994; Maniezzo et al., 1998)). As distâncias podem revelar fatos importantes sobre um par de soluções. Elas fornecem uma idéia do esforço (e do tempo) requerido para se mover de uma solução à outra; elas estão correlacionadas com a probabilidade de que a segunda solução será visitada por um processo que passou pela primeira; elas dão uma medida precisa da similaridade entre as duas soluções. Observa-se a falta de uma medida de diversidade que seja aceita na maioria dos modelos heurísticos modernos e estas métricas de distância poderiam ser uma possível sugestão (obviamente considerando que elas estão relacionadas a um relevo específico, veja Jones, 1995).
Sem uma noção de distância, um processo de busca pode se tornar cego. Por exemplo, é possível que um processo de busca gravite em direção a um grande 'atrator' no espaço de busca, enquanto se mantém a uma distância relativamente pequena de uma solução particular, permanecendo 'ancorado' a esta região por um longo período. Um número grande de movimentos podem ser realizados com variações de distância negligíveis. Sem uma medida precisa de distância fica difícil perceber se este fenômeno está ocorrendo ou se o processo está explorando uma parte mais ampla do espaço de solução. Métricas de distância podem, portanto, ser úteis em metaheurísticas baseadas em buscas locais.
Infelizmente, termos como 'áreas' ou 'regiões' de uma espaço de busca tem um significado um tanto quanto nebuloso. Não existem limites ou demarcações[1] definidas no espaço de busca e toda solução tem aquele mesmo número de vizinhos diferentes. Uma área ao redor de p pode ser formalmente definida como $A(p,S)$:Ωx$\{1,2,\ldots,n\} \rightarrow 2^{\Omega}$, onde S é o 'ráio' da área. Assim $x \in A(p,S)$ se e somente se $x \in \Omega \wedge D_N(x,p) \leq S$ (para o operador de vizinhança N escolhido). Desta forma, referir-se a dois processos compartilhando 'a mesma área' é uma noção relativa a S (o tamanho da área) e de $D_N(x,p)$ (a distância entre os dois processos).

2.1. Um elo com a análise de relevo

[1] Este é um problema difícil de se compreender que aparece não somente em espaços de busca combinatoriais, mas também na metafísica da concepção de padrões (veja Linhares (2000)).

Métricas de distância possibilitaram o estudo recente de um novo tópico em buscas heurísticas: a análise de relevos em problemas de otimização combinatória. Muitos pesquisadores tem considerado a relação envolvendo a estrutura de custos e distâncias, isto é, o número mínimo de aplicações de um dado operador de vizinhança necessário para transformar uma solução em outra, entre soluções ótimas locais. Boese, Kahng e Muddu (1994), estudando o problema do caixeiro viajante e problemas de bisecção de grafos argumentam em direção a uma estrutura de *grande-vale* onde soluções ótimas locais tendem a ficar mais e mais próximas entre si quando se aproximam do ótimo global. Novos estudos se seguiram, com uma análise similar para o problema de flowshop $n/m/P/C_{max}$ e outros (Merz e Freisleben, 1999; Reeves, 1999). Em um estudo relacionado, Mak e Morton (1995) analisaram a relação entre as métricas k-opt e o 2-opt para o problema do caixeiro viajante.

Uma falha desses estudos é que a noção de distância foi tratada de maneira aproximada. A maioria dos autores reclama da falta de algoritmos de tempo polinomial conhecidos para se medir distâncias mínimas (em termos de movimentos) entre soluções para problemas de sequenciamento (representados por permutações) (Boese et al., 1994; Merz e Freisleben, 1999; Reeves, 1999) e, desta forma, recorrem a aproximações. No apêndice A, reduzimos esta falta introduzindo novos algoritmos de tempo polinomial para os operadores 2-troca e inserção para permutações.

Como um exemplo, considere um movimento de 2-troca pelo qual duas posições em uma permutação são trocadas. O cômputo da distância mínima 2-troca entre um par de permutações π e σ pode ser visto como um problema de caminho mínimo no grafo $G=(V, E)$, onde cada nó $v\in V$ está associado com uma permutação $\Pi(v)$, e $(u, v)\in E$ se e somente se $\Pi(u)$ pode ser obtido por um operação 2-troca em $\Pi(v)$. Em princípio, este problema poderia ser resolvido pela algoritmo bem conhecido de Dijkstra (1959). Entretanto, a aplicação do algoritmo de Dijkstra's requeriria uma complexidade de tempo polinomial no tamanho do grafo G o que torna esta abordagem intratável por causa do número exponencial de permutações Π. Apresentamos no apêndice A um algoritmo de tempo linear no tamanho da permutação que constrói o caminho 2-troca mínimo de π para σ; também apresentamos uma nova métrica para o operador de inserção.

Na próxima seção sugerimos um arcabouço para intensidade e diversidade simultânea baseado na introdução de mecanismos de controle denominados *políticas de coordenação*.

3. UM ARCABOUÇO PARA A COEXISTÊNCIA SIMULTÂNEA DE INTENSIDADE E DIVERSIDADE

Para defender a viabilidade de se empregar simultaneamente intensidade e diversidade na busca admite-se o seguinte:

1) que haja um conjunto concorrente de processos de busca, que cada um desses processos tende a explorar o espaço de busca com alta intensidade;
2) existe uma métrica de distância entre soluções que é computável em tempo polinomial (veja a seguir).

Analisemos a primeira suposição. Quando um processo de busca local intensiva (como aqueles considerados em Glover (1989) ou Linhares et. al (1999)) aproxima-se de uma área, ele tenta, durante um certo tempo, achar a melhor solução possível contida na área. Esta é uma suposição trivial; para todos os propósitos práticos, estes processos de busca podem ser simples exemplares de busca tabu *projetados para explorar o espaço de busca intensamente* (Glover 1989).

Esta suposição básica estabelecida de forma explícita seria:

Asserção 2. *O modelo proposto exibe alta intensidade ao longo de sua execução.*

O modelo pode exibir um comportamento intensivo notável procurando continuamente por soluções de alta qualidade dentro de certas regiões restritas do espaço de soluções. Pode-se admitir que isto ocorra durante toda a execução do algoritmo (Glover, 1989); intensidade (sob o modelo proposto) pode ser preservada durante a execução (ao invés de alternar com diversidade, como veremos adiante).

Precisamos mostrar que quando consideramos modelos cooperativos de busca, a diversidade de busca também pode existir durante toda a execução. Para se mostrar isto, considere que temos um conjunto de processos sendo executados em paralelo e que vamos implementar uma 'cerca' de acesso para cada

processo. Assim, enquanto múltiplos processos conduzem a busca, a distância entre processos é computada de modo que uma 'cerca' artificial protege a área de cada processador (a fim de garantir a preservação da diversidade de busca).
No caso de um processo particular tentar cruzar a cerca de outro processo, esta tentativa de movimento é rejeitada. O tamanho desta cerca pode ser variável e isto será discutido em detalhe mais adiante.
As *políticas de coordenação* sugeridas, i.e., princípios de direcionamento que coordenam a busca cooperativa (baseado nas distâncias medidas entre processos de busca), podem ser classificadas em duas categorias: políticas de coordenação de distribuição e políticas de coordenação de redistribuição. Nas políticas de coordenação de distribuição consideradas neste trabalho, certos pontos agem como *detratores* mantendo uma distância mínima específica entre processos, forçando uma distribuição ampla do sistema. Nas políticas de coordenação de redistribuição, tem-se alguns pontos agindo como *atratores* de outros processos trazendo estes processos para regiões potencialmente melhores do espaço de busca (veja (Linhares 2001)). Consideremos duas políticas de coordenação de distribuição.

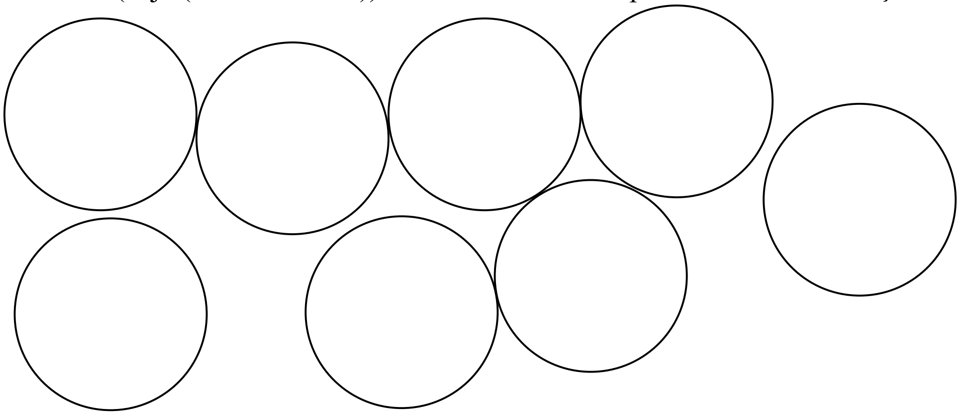

Fig. 1. A política de coordenação uniforme sem-sobreposição. Os processos devem manter uma distância mínima entre si (as cercanias de cada processo estão representadas por círculos). Estes mecanismos garantem altos níveis de diversidade na busca ao longo da execução.

3.1. Uniforme, sem-sobreposição

A primeira política que focalizaremos é a denominada política de coordenação *uniforme sem-sobreposição* (Fig 1). A idéia nesta política é atribuir uma área específica a cada processo de busca. Estas áreas ao redor das soluções e exploradas pelos processos são limitadas a um tamanho uniforme e não podem ser "invadidas" por outros processos. Sob esta política não pode haver sobreposição entre áreas e, assim, existe uma *distribuição garantida* dos processos ao longo do espaço de busca. É preciso mencionar que esta garantia formal não é atingida nos modelos clássicos de algoritmos genéticos (ou "espécies de algoritmo genético", como em Jelasitya e Dombib, 1998), simulated annealing ou mesmo busca tabu (apesar de que neste último pode-se incorporar estruturas de memória de longo-prazo desenvolvidas especificamente para diversificação na busca mas sem uma *garantia* formal, matemática, de que a procura será convenientemente distribuída ao longo do espaço de busca pois não se pode medir precisamente tal distribuição uma vez que não existe simplesmente uma função análoga a $D_N(x,p)$).
Suponha, por exemplo, que o processo P_2 tenta um movimento que o trará dentro da área demarcada pelo ráio do processo P_1. Para impedir isto o movimento pode ser rejeitado, ou, alternativamente, o processo P_1 pode ser movido para fora da área correspondente de movimentação do processo P_2. No primeiro caso, P_1 atua como um detrator enquanto que, no segundo caso, P_2 é um detrator. O processo 'vencedor' pode ser escolhido pela qualidade das melhores soluções encontradas por cada processo ou pela qualidade da solução corrente sendo explorada. O número de processos de busca interdependentes pode ser comparável ao número de processadores disponíveis possibilitando, desta forma, uma maior distribuição da procura.

Existe, entretanto, um alto custo computacional associado ao cômputo das distâncias entre pares de processos, especialmente quando as métricas de distância não são computáveis em tempo linear. Uma possibilidade de se contornar este problema é desenvolver métricas de distância incrementais, evitando o cálculo integral das distãncias a cada movimento, calculando as distâncias usando somente a informação incremental associada a cada operação. Reduções adicionais da complexidade computacional (para qualquer operador de vizinhança selecionado) podem ser obtidas se optarmor por reduzir o número de distâncias a serem computadas de $O(p^2)$ para $O(p)$, onde p é o número de processos. Isto pode ser feito amostrando-se aleatoriamente pares de processos ou adotando-se uma das políticas de coordenação 'mais rápidas' apresentadas em (Linhares 2001).

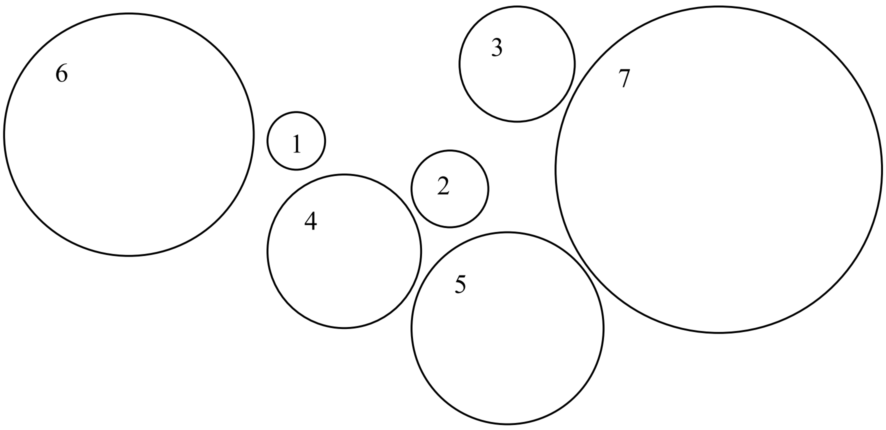

FIG. 2. A política de distribuição de custo mais próximo não-uniforme ordenada sem-sobreposição. Os processos de busca melhores tem áreas cada vez menores, permitindo uma exploração do espaço de solução que é proporcional a qualidade percebida de cada área, potencialmente tirando vantagem dos espaços com estrutura de grande-vale.

3.2. Políticas de distribuição não-uniforme

A mesma política pode também ser utilizada em uma região de tamanho não-uniforme (Fig. 2), ou seja, o tamanho das áreas atribuídas a cada processo pode variar de acordo com alguma variável específica. Um candidato para tal variável é a função custo obtida pelos processos. Pode-se usar a ordem dos processos como um guia para se estabelecer a área total de suas regiões.

Na proposta de políticas de coordenação sem-sobreposição pretende-se obter um processo de busca que seja intensivo e diverso ao mesmo tempo (desta forma mostrando que essas aspirações não são mutuamente exclusivas). Uma segunda proposta para políticas de coordenação cuja meta seria a exploração mais inteligente do espaço de busca é inspirada na hipótese previamente mencionada de 'grande-vale'. Uma correlação entre qualidade de solução e distância ao ótimo foi observada em alguns problemas de otimização combinatória (veja, por exemplo, Boese, Khang e Muddu, 1994). Esta correlação sugere que áreas nas quais soluções de melhor qualidade foram encontradas devem ser exploradas com maior intensidade que outras onde nenhuma solução (de melhor qualidade) foi encontrada.

A política de custo mais próximo não uniforme ordenada sem-sobreposição apresenta todas as vantagens da política uniforme ordenada sem-sobreposição e com a possibilidade de adaptabilidade aumentada. Aos melhores processos são atribuídos áreas menores do que aos piores processos. As regiões do espaço de busca onde se observa soluções de melhor qualidade são exploradas de maneira proporcional à qualidade percebida. A um processo que tenha encontrado soluções da mais alta qualidade é assegurado uma área menor permitindo assim que outros processos explorem a vizinhança

daquela área. Processos que ocupam áreas ruins tem em correspondência áreas de tamanhos maiores, não permitindo o acesso àquelas regiões supostamente de qualidade inferior.
Destas idéias, é fácil verificar que as asserções seguintes são válidas:
Asserção 3. *O modelo proposto exibe alta diversidade na busca durante a sua execução.*
A diversidade está garantida pelos cálculos de distâncias entre processos de busca. Esta asserção, junto com a asserção 2, conduz à seguinte conclusão:
Asserção 4. *Apesar do fato de que a 'crença intensidade versus diversidade' é largamente difundida, ela é falsa: intensidade não é mutuamente exclusiva da diversidade na busca local.*
Com este raciocínio podemos concluir que a crença intensidade versus diversidade, apesar de largamente difundida, não é verdadeira.
É preciso observar que as políticas de coordenação introduzidas são mecanismos para coordenação entre processos individuais de busca mas, como discutido, elas estão baseadas em *métricas* de distância claras entre duas soluções particulares. Como as tarefas de se obter estas distâncias são por si só problemas de otimização em espaços combinatórios, permanece para ser mostrado que algoritmos eficientes podem ser desenvolvidos para tais métricas. No apêndice A introduzimos métricas novas de sequenciamento (permutação).

4. CONCLUSÃO

Neste trabalho argumenta-se que intensidade e diversidade na busca não são mutuamente exclusivos, embora pesquisadores de renome tenham sugerido 'achar *um equilíbrio adequado* entre intensidade e diversidade' e, esta idéia de um 'equilíbrio' ou de uma suposta troca entre estas aspirações claramente pressupõe sua exclusão mútua. Introduzindo um arcabouço para modelos cooperativos nos quais a intensidade é atingida no nível de processamento enquanto que a diversidade é atingida pela cooperação ao nível de todo o sistema, demonstra-se que, para sistemas com múltiplos processos de busca, estas aspirações podem coexistir simultaneamente durante toda a execução do sistema (ao invés de alternadamente ou de uma movimentação gradual da diversidade para a intensidade).
Novos mecanismos de controle, referidos como *políticas de coordenação* foram introduzidas. Estas políticas são princípios gerais para guiar a distribuição (e também redistribuição) de processos no espaço de busca. Algumas políticas adicionais de coordenação para redistribuição de processos de busca estão esquematizados em (Linhares 2001).
Finalmente, introduzimos no Apêndice A novas métricas de distância para problemas de sequenciamento. Empregando-se estas métricas é possível computar, por exemplo, a distância entre a solução corrente sendo explorada e a melhor solução encontrada. É possível ver também com clareza se um processo de busca está explorando muitas regiões distintas do espaço solução ou se está gravitando ao redor de algum ponto. Uma outra possibilidade é o desenvolvimento de distâncias tabu: ao invés de utilizar listas de passos anteriores, pode existir uma "trilha" de soluções visitadas anteriormente (atuando como 'detratores') e uma distância correspondente ('ráio') a ser mantida com cada uma dessas soluções, garantindo assim uma alta diversidade ao longo da procura.
As métricas de distância permitem cálcular a dispersão dos processos no espaço de busca e, no caso de dois (ou mais) processos se aglomerarem em torno de uma pequena área, permitem direcionar um deles (ou mais) para fora desta área, a fim de eliminar a duplicação de esforços na busca.
Um leitor cético pode argumentar que 'neste trabalho não se apresenta um problema de otimização específico e uma estratégia de busca específica e, portanto, não demonstra que intensidade e diversidade simultânea é na verdade vantajosa em termos de desempenho do algoritmo'. Isto é verdade. Se um melhor desempenho será esperado é uma questão empírica significante mas não altera o fato de que *a idéia do equilíbrio entre diversidade e intensidade precisa ser revisada*. Mais ainda, como este arcabouço está separado das heurísticas inerentes ao processo de busca, das métricas de distância empregadas (que variam de operador para operador), e do problema de otimização sendo resolvido, ele merece ser discutido de maneira independente (algumas aplicações preliminares e experimentos foram feitos em Linhares, 2001).
Finalmente, a questão colocada no título pode ser respondida com um 'sim' conclusivo: é uma falsa troca. Na maioria dos modelos de busca local temos que em alguns momentos a intensidade é atingida

em troca da diversidade; enquanto que em outros, a diversidade é atingida em troca da intensidade. Apesar de que alguns pensadores argumentaram que intensidade e diversidade são “alternativas conflitantes”, foi demonstrado que ambos podem ser atingidos simultaneamente. Atrás desta sugestão esconde-se um grande território não explorado para pesquisa de heurísticas modernas.

APPENDICE A. – MÉTRICAS DE DISTÂNCIA

Uma das suposições básicas do modelo proposto é a existência de algoritmos eficientes em tempo polinomial para o cômputo de métricas de distâncias. Neste apêndice introduzimos novas métricas de distâncias desenvolvidas para avaliar em tempo polinomial a distância mínima entre quaiquer duas soluções (dado um operador de vizinhança padrão). Primeiramente observamos que como existem várias métricas bem conhecidas disponíveis na literatura para problemas 0-1 como a distância de *Hamming* (veja Mendes et. al, 2001), o foco aqui está concentrado em novas métricas para o caso de problemas de sequenciamento, nos quais soluções são representadas geralmente por permutações. Consideramos os seguintes operadores geralmente utilizados em problemas combinatórios envolvendo permutações: a 2-troca e o operador de inserção.

A.1. O OPERADOR 2-TROCA

Enquanto que o operador 2-opt é o mais usado no problema do caixeiro viajante, outros operadores são ainda mais populares em problemas diferentes de sequenciamento, como os operadores 2-troca e de inserção usados em muitas aplicações de programação. Nesta seção apresentamos um algoritmo de tempo linear que transforma π em σ usando o menor número de operações de 2-troca.
O leitor deve observar que, ao contrário do que o nome sugere, este problema é diferente da *ordenação de troca ótima* (Knuth, 1975, p.198), no qual a preocupação é com a árvore de comparação-troca (que é uma estrutura de dados utilizada para estudar o desempenho de algoritmos distintos de ordenação e troca, tais como quicksort, bubblesort ou mergesort). O leitor deve se referir a Knuth (1975) para detalhes.

A.1.1. Um algoritmo para a métrica de distância de 2-troca

O operador 2-troca troca duas posições i e j em uma permutação, ou seja, ele transforma a permutação $(\pi_1, \ldots, \pi_i, \ldots, \pi_j, \ldots, \pi_n)$ para $(\pi_1, \ldots, \pi_j, \ldots, \pi_i, \ldots, \pi_n)$. Apresentamos um algoritmo de tempo linear para computar a distância mínima 2-troca entre π e $\sigma = I$.

```
d = 0;
for i= 1 to n do
        if πi≠σi then begin
        2-EXCHANGE(i, σ^-1(πi) );
        d = d +1;
        end;
return d;
```

Este algoritmo é trivial e roda em tempo linear. O que não é trivial é mostrar que ele determina o número mínimo de movimentos 2-troca entre π e $\sigma = I$. Para provar a otimalidade do algoritmo, colocamos o problema na forma teórica de um grafo o qual denominaremos de *grafo de troca*. Seja $G = (V, E)$ um grafo direcionado com conjunto de vértices $V=(v_1, v_2, \ldots, v_n)$ e conjunto de arcos E. Associe cada vértice $v_i \in V$ com uma posição i na permutação π. Os arcos de G apontam da posição i de π para a posição em σ que o elemento π_i aparece, ou seja, o arco direcionado $v_iv_j \in E$ se e somente se $j = \sigma^{-1}(\pi_i)$. Na figura 3 o grafo de troca para as permutações π=(2,4,3,5,1,8,7,6) e $\sigma = I$ estão apresentados.
Observe que como cada vértice envia e recebe um arco direcionado, o grafo de troca consiste de um ou mais ciclos independentes. O tamanho destes ciclos varia de $\{1, 2, ..., n\}$ e o número de ciclos distintos pode ser no máximo n, no caso de n auto-ciclos (i.e., vértices apontando para eles mesmos). Observe que isto é exatamente a meta desejada, i.e., colcoar cada elemento em sua posição correta. Assim estamos preocupados em aumentar o número de ciclos com o número mínimo de operações possíveis. O problema pode ser colocado agora em termos de um grafo de troca. Seja B o número de ciclos no grafo de troca entre as permutações π e σ, e considere a proposição 1.

Proposição 1. *Dado $B < n$, é sempre possível aumentar o número de ciclos por uma unidade por um movimento de 2-troca.*
Prova. Dado que $B < n$, é sempre possível quebrar um ciclo C_{M+N} de tamanho $M+N$ em dois ciclos distintos C_M e C_N, para qualquer combinação de M □ 0 e N □ 0, em uma operação única de troca: selecione os vértices $v_i, v_j \in C_{M+N}$; e, após a operação de 2-troca, os vértices que apontavam para v_i e v_j agora apontam para v_j e v_i, respectivamente. Assim, o ciclo C_{M+N} é quebrado em dois ciclos distintos C_M e C_N, onde M é igual ao número de arcos no caminho de v_i para v_j em C_{M+N}, e N é igual ao número de arcos no caminho de v_j para v_i em C_{M+N}.

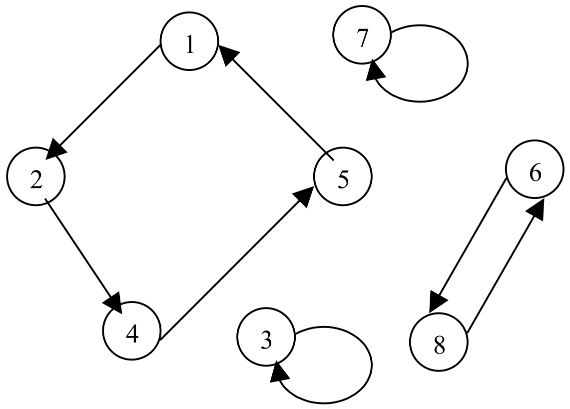

Fig. 3. Grafo de troca das permutações $\pi = (2,4,3,5,1,8,7,6)$ e I.

Proposição 2. *Um movimento 2-troca pode aumentar o número de ciclos de no máximo uma unidade.*
Prova. Quando dois elementos trocam de posição na permutação, 2 arcos são retirados do grafo e 2 arcos são inseridos. É necessário retirar 2 arcos para quebrar um ciclo; assim, cada movimento aumenta o número de ciclos de no máximo uma unidade.
Seja $T(C_S)$ o número mínimo de aplicações do operador 2-troca para transformar qualquer ciclo C_S de tamanho S em S auto-ciclos. Assim, temos

Lemma 1. $T(C_S) = S\text{-}1$.
Prova. Segue de maneira trivial das proposições 1 e 2.
Seja $D_{\text{2-exchange}}(\pi, \sigma)$ igual à metrica de distância 2-troca desejada entre as permutações π e σ. Então temos

Teorema 2. $D_{\text{2-exchange}}(\pi, \sigma) = n\text{-}B$.
Prova. Segue do lemma 1 que cada ciclo de tamanho S precisa de (S-1) movimentos para se transformar em S auto-ciclos. A métrica de distância desejada pode ser obtida diretamente pelo número de ciclos de tamanho S no grafo de troca, medido por

$D_{2-exchange}(\pi,\sigma) = \sum_{S=1}^{n}(S-1)\alpha_S = \sum_{S=1}^{n} S\alpha_S - \sum_{S=1}^{n}\alpha_S = n - B$, onde α_S é o número de ciclos de tamanho S entre as permutações π e σ.
Como o algoritmo apresentado coloca um elemento em sua posição correta, ele sempre quebra o ciclo em dois (com o novo ciclo tendo tamanho exatamente 1). Isto conduz à prova de corretude do algoritmo.

A.2. O OPERADOR DE INSERÇÃO

O operador de inserção para permutações elimina um elemento da permutação e o insere em uma nova posição, ou entre dois elementos, ou no começo, ou na posição final, mantendo a ordem dos demais intacta. Mais uma vez, a métrica desenvolvida aqui lida com permutações fixas. Entretanto, a extensão para problemas cíclicos pode ser feita de imediato pois para este caso, seria suficiente manter um elemento em uma posição fixa da permutação (por exemplo, mantendo o elemento 1 fixo na primeira posição da permutação). Observe que apesar do que o nome sugere, a métrica de distância de inserção considerada aqui não é aquela estudada anteriomente do *ordenamento por inserção* (Knuth, 1975, seção 5.2.1), no qual o número de movimentos se refere ao tempo de execução dos algoritmos. O leitor deve-se referir a Knuth (1975) para detalhes. No problema considerado aqui não é trivial ver que a permutação (8,5,6,9,4,2,1,7,3) requer 6 inserções enquanto a permutação (1,8,5,2,6,3,9,4,7) requer 4. Este problema é muito próximo daquele de Aigner e West (1987) para a distância de inserção do elemento líder, referido como inserções *cabeça*, i.e., considerando apenas inserções no primeiro elemento da permutação (o leitor deve referir-se a Aigner e West (1987) para detalhes). Eles mostraram que:

Teorema 3. *(Aigner e West* (1987)*). O número de inserções requerido para ordenar uma permutação σ por inserções cabeça é n-k, onde k é o maior inteiro tal que as últimas k entradas de σ formam uma subsequência crescente.*

Em nosso caso mais geral, para ordenar por inserção a permutação (1,8,5,2,6,3,9,4,7), por exemplo, a inspiração requerida é que a subsequência (1,2,3,4,7) já aparece na ordem correta, mas os demais elementos estão em posições incorretas. Referimo-nos a tal subsequência como uma *subsequência correta*. Uma *subsequência correta maximal* é uma que nenhum elemento da permutação pode ser adicionada a ela e a subsequência permanece em ordem correta. Uma subsequência correta maximal com o maior tamanho possível (i.e., número máximo de elementos) é denotado por uma *subsequência correta máxima*. Neste exemplo, a subsequência (1,5,6,9) é uma subsequência correta maximal, mas não máxima. O seguinte lemma capacita-nos estabelecer um elo entre a distância inserção e a subsequência correta máxima.

Lemma 2. *Se uma permutação contem uma subsequência correta maximal de tamanho K, é sempre possível ordená-la com n-K inserções.*

Prova. Relativo à subsequência correta maximal, existem *n-K* posições incorretas na permutação. Cada uma dessas posições pode ser corrigida inserindo-a em sua posição correta na subsequência correta maximal por uma aplicação do operador de inserção. Assim, *n-K* inserções são suficientes para ordenar esta permutação.

Desta forma, dada uma subsequência correta maximal de tamanho *K*, *n-K* inserções são suficientes para ordenar a permutação. (Não provamos, entretanto, que este é um número necessário de inserções para se ordenar a permutação).

É importante maximizar *K,* i.e., achar a subsequência correta máxima. A subsequência correta máxima pode ser obtida em tempo polinomial, como mostrado a seguir. Construiremos um grafo (*G, V*) que pode ser percorrido como um conjunto de árvores, onde cada árvore representa uma das subsequências corretas maximais. Primeiro, crie um conjunto de nós onde cada nó v_i corresponde a uma posição *i* na permutação π. A seguir, crie um conjunto *E* de arcos colocando um arco direcionado v_iv_j se e somente se (1) $i < j$ e, (2) $\pi_i < \pi_j$ e, (3) $\neg\exists k$: $i < k < j$ e $\pi_i < \pi_k < \pi_j$. As condições 1 e 2 asseguram que as posições *i* e *j* pertencem a subsequência correta enquanto que a condição 3 evita a transitividade. Observe que *G* é direcionado e não inclui ciclos e pode ser percorrido como um conjunto de árvores onde os nós que não recebem arcos são as raízes das árvores. Na Figura 4, a árvore-percorrimento associada ao exemplo permutação π = (1,8,5,2,6,3,9,4,7) é mostrada.

Lemma 3. *A subsequência correta máxima é igual à altura máxima H dentre todas as árvores-percorrimento contidas em G.*

Prova. Cada caminho dos nós às folhas representam uma subsequência correta maximal. Assim, a subsequência correta máxima será aquela dada pela maior altura das árvores contidas em *G*.

Lemma 4. *É necessário realizar no mínimo n-H inserções.*

Prova. Considere uma subsequência correta máxima. Se um elemento da permutação é retirado, o tamanho da subsequência correta máxima não pode ser maior do que a original. De fato, ela pode permanecer com o mesmo tamanho do original ou diminuir de uma unidade. Se um elemento é incluído na permutação, o tamanho da subsequência correta máxima não pode ser menor do que a original. Ela pode ser igual que a original ou ser acrescida de uma unidade. Um movimento de inserção é equivalente a estes dois passos sendo realizados consecutivamente. Assim, um movimento

de inserção pode aumentar a subsequência correta máxima de no máximo uma unidade. Como a subsequência correta máxima tem tamanho *H*, é necessário no mínimo *n-H* inserções para ordenar-se a permutação.
Como sabemos que *n-H* inserções são suficientes para se ordenar a permutação, isto conduz-nos à métrica de distância desejada:
Teorema 4. *A distância de inserção de uma permutação* π *de n inteiros é n-H.*
Prova. Segue dos lemmas 2, 3 e 4.

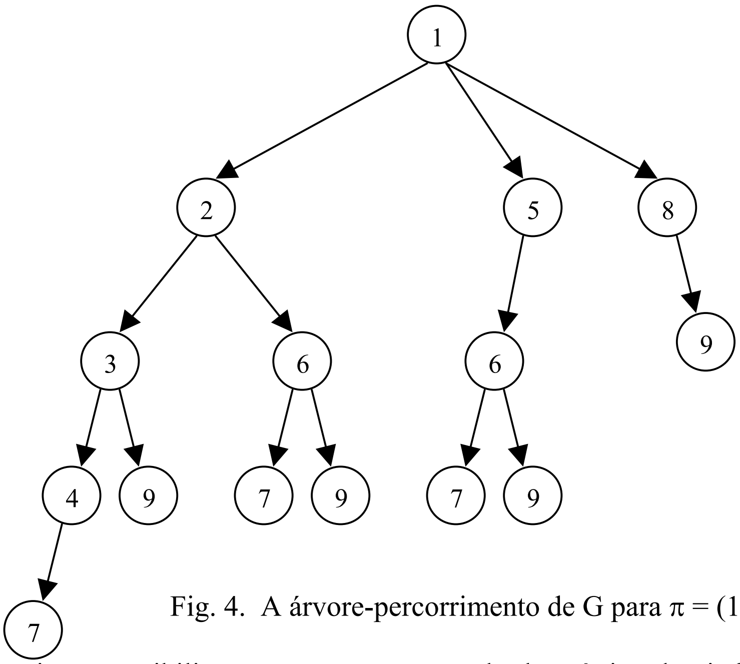

Fig. 4. A árvore-percorrimento de G para π = (1,8,5,2,6,3,9,4,7)

O algoritmo possibilita-nos provar a corretude da métrica desejada, mas precisamos observar que a criação do grafo *G* requer tempo $O(n^2)$ (e também o percorrimento das árvores), mas existem algoritmos mais rápidos disponíveis para "determinar a maior subsequência em uma sequência" (Orlowski e Pachter 1989). Nossa preocupação aqui é somente com a demonstração da corretude da métrica.

Agradecimentos –Agradecemos o suporte financeiro da FAPESP e CNPq.

Referências

M. Aigner, and D.B. West, Sorting by insertion of leading elements, *Journal of Combinatorial Theory Series A* 45 (1987) 306-309.
M.F. Argüello, J.F.Bard, and G.Yu, A GRASP for aircraft routing in response to groundings and delays, *Journal of Combinatorial Optimization* v. 1, p.211-228, 1997.
V. Bafna, and P.A. Pevzner, Genome rearrangements and sorting by reversals, *SIAM Journal on Computing* 25 (1996) 272-289.
V. Bafna, and P.A. Pevzner, Sorting by transpositions, *SIAM Journal on Discrete Mathematics* 11 (1998) 224-240.
J.E. Beasley, Population heuristics, *forthcoming in Handbook of applied optimization*, P.M. Pardalos and M.G.C. Resende (eds). Oxford University Press (2000).
K.D. Boese, A.B. Kahng, S. Muddu, A new adaptive multi-start technique for combinatorial global optimizations, *Operations Research Letters* 16 (1994) 101-113.
A. Caprara, Sorting permutations by reversals and Eulerian cycle decompositions, *SIAM Journal on Discrete Mathematics* 12 (1999) 91-110.
N. Christofides, The bionomic algorithm, *Proceedings of the AIRO'94 Conference*, Savona, Italy, 1994.

C.R. Reeves, Landscapes, operators, and heuristic search, *Annals of Operations Research* 86 (1999) 473-490.
M.G.C. Resende, and C.C. Ribeiro, A GRASP for graph planarization, *Networks* v.29, p.173-189, 1997.
N. Tran, An easy case of sorting by reversals, *Journal of Computational Biology* 5 (1998) 741-746.