Topologia de Problemas de Otimização por meio de Modelos em Grafos:
Evidências da Hipótese de *Building-Blocks*

Jean Paulo Martins

jean@icmc.usp.br

Universidade de São Paulo

Instituto de Ciências Matemáticas e Computação,

São Carlos, 13566-590, SP, Brazil

Resumo

Neste artigo investiga-se a possibilidade de confirmação da hipótese de *building-blocks* para alguns problemas de otimização irrestritos. Primeiramente, constrói-se uma rede Bayesiana utilizando para isso uma amostra do espaço de soluções do problema, a rede criada terá um nó para cada variável do problema e as arestas representarão relações de dependência entre as variáveis. Supõe-se que a obtenção de uma rede de alta modularidade através desse procedimento tenha relação direta com a hipótese de *Building-blocks*.

1 Introdução

Redes complexas descrevem uma ampla variedade de sistemas, desde escalas microscópicas como reações químicas dentro das células até escalas macroscópicas como redes sociais e a Internet. O porque de estruturas similares emergirem em contextos tão diferentes é um dos principais tópicos de estudo em redes complexas [Albert and Barabási, 2002].

Neste trabalho, propõem-se a modelagem de problemas de otimização em redes complexas, de modo a possibilitar a caracterização da modularidade dessas redes. A construção dos modelos de redes utilizará como base redes Bayesianas [Jensen, 1996] que serão criadas a partir de uma amostra do espaço de soluções do problema em questão, a hipótese assumida para o desenvolvimento deste trabalho afirma que uma possível alta modularidade das redes obtidas por esse procedimento seriam evidências para a comprovação da hipótese de *Building-blocks*.

Nas próximas seções conceitos importantes ao desenvolvimento deste trabalho são revisados. Inicia-se, na Seção 1.1, com a definição de problemas de otimização, seguida, na Seção 1.2, com a descrição da hipótese de *Building-blocks* e dos experimentos e conclusões na Seção 4.

1.1 Problemas de Otimização

Problemas de Otimização podem ser definidos pela necessidade de se obter uma melhor configuração para um conjunto de variáveis que atendam um conjunto de restrições e minimizem ou maximizem uma função objetivo f. Formalmente, uma solução para um problema de otimização pode ser representada por um vetor x com l variáveis de decisão (ou seja, $x \in R^l$), sendo que, a partir de uma função objetivo f, é possível avaliar a qualidade da solução representada por x [Michiels, Aarts, and Korst, 2007; Papadimitriou and Steiglitz, 1998].

$$\begin{array}{lrcll} \text{Minimizar} & f(x); & & & \\ \text{sujeito a} & g_j(x) & \geq & 0, & j = 0, \ldots, J; \\ & h_k(x) & = & 0, & k = 0, \ldots, K; \end{array} \tag{1}$$

Uma solução x que não satisfaça todas as $J + K$ restrições é considerada uma solução não-factível, por outro lado, soluções que atendam as restrições são chamadas factíveis e definidas como constituintes do conjunto S. A partir disso, pode-se definir que o objetivo em um Problema de Otimização é encontrar uma dentre as melhores soluções factíveis x, a quais são chamadas ótimos globais (ver Definição 1.1) [Papadimitriou and Steiglitz, 1998].

Definição 1.1 (Ótimo global) *Um ótimo global x^*, em um problema de otimização na forma de minimização deve satisfazer: $f(x^*) \leq f(x)$, $\forall x \in S$.*

Dependendo do formato do espaço de soluções S, podem existir soluções que aparentem ser ótimos globais, mas que no entanto, representam apenas a melhor solução de um subespaço $N \subseteq S$, tais soluções são chamadas ótimos locais (ver Definição 1.2) [Michiels et al., 2007; Papadimitriou and Steiglitz, 1998]. A Figura 1 descreve a relação entre entre um ótimo local x' e um ótimo global x^*, onde Δf indica a diferença entre os valores da função objetivo para x' e para x^*.

Definição 1.2 (Ótimo local) *Um Ótimo Local x', em relação a $N \subseteq S$, em um problema de otimização na forma de minimização deve satisfazer: $f(x') \leq f(x), \ \forall x \in N(x')$.*

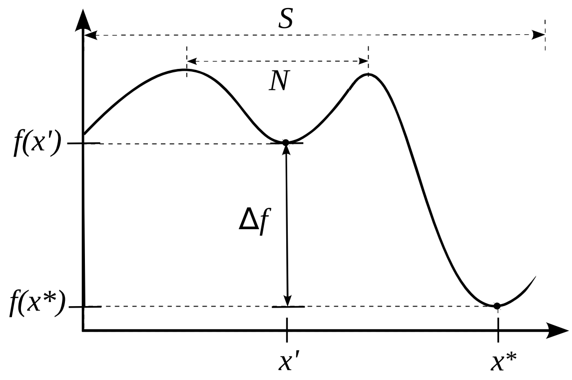


Figura 1: *Espaço de busca contendo um ótimo local x' e um ótimo global x^* em um problema de minimização.*

Por fim, os problemas de otimização podem ser qualificados quanto ao tipo de variáveis de decisão utilizadas na representação de uma solução x, tem-se os Problemas de Otimização Contínua, representados por um espaço de busca S contínuo, e os Problemas de Otimização Discreta (Combinatória), representados por um espaço de busca S discreto. Neste trabalho os experimentos serão avaliados sobre problemas de otimização combinatória.

1.2 Algoritmos Genéticos e a Hipótese de *Building-Blocks*

A partir das primeiras tentativas de explicar teoricamente o funcionamento dos GAs algumas hipóteses foram levantadas quanto aos pontos cruciais para a eficiência dos GAs. Considerando os efeitos de operadores genéticos como *crossover* e mutação, algumas conclusões foram obtidas, sendo a principal o Teorema dos Schemas e a Hipótese dos *Building-Blocks* [Goldberg, 1989; Holland, 1992].

De forma direta, a Hipótese dos *Building-Blocks* (ver Hipótese 1.1) define que existem dependências entre as variáveis de um problema, de modo que só é possível aos GAs conseguir um desempenho próximo ao ótimo caso tais grupos de variáveis Blocos Construtivos (BBs[1]) sejam combinados de modo a formar soluções completas ao decorrer do processamento. Ao processo de identificação de BBs foi dado o nome *Linkage Learning* [Harik, Lobo, and Sastry, 2006; Bäck, De Graaf, Kok, and Kosters, 2001].

Hipótese 1.1 (Hipótese dos *Building-Blocks*) *Um Algoritmo Genético obtém desempenho próximo ao ótimo através da justaposição de Building-Blocks.*

[1] *Building-Blocks*

Uma consequência direta da hipótese refere-se ao desempenho dos GAs caso os BBs não sejam efetivamente considerados. Supondo a hipótese verdadeira, um problema de otimização seria mais facilmente solucionado por um Algoritmo Genético em situações em que, iterativamente, BBs sejam identificados e combinados em soluções completas. O ponto em questão é que, os operadores genéticos responsáveis pela combinação de BBs (*crossover*) podem destruir blocos criados em iterações anteriores do algoritmo, nesses casos, a convergência rápida a boas soluções ficaria comprometida, visto que a identificação de alguns BBs teria que ser refeita nas demais gerações.

Nesse contexto, com o objetivo de contornar essas dificuldades surgiram os Algoritmos Genéticos Competentes, algoritmos que utilizam a informação sobre os BBs como forma de melhorar o desempenho dos GAs. Tais algoritmos foram aplicados com sucesso em diversos contextos, apresentando na maioria dos casos convergência mais rápida a soluções de boa qualidade que os GAs convencionais, contudo, seu ponto central é baseia-se na veracidade da hipótese de *Building-blocks*, o que pode criar , em alguns casos, certo ceticismo da comunidade científica quanto a real efetividade de tais algoritmos.

1.3 Objetivos e Justificativa

Partindo do contexto descrito nas seções anteriores, relacionado ao sucesso, que vem sendo obtido por diversos trabalhos na área de redes complexas, quanto à modelagem e caracterização de redes nos mais diversos tipos de situações, este trabalho se direciona ao desenvolvimento de uma abordagem que permita a modelagem de problemas de otimização em redes, de modo que cada vértice represente uma variável do problema e as arestas descrevam as interdependências entre tais variáveis. É estritamente importante neste ponto, que a rede obtida represente fielmente a estrutura do problema em questão.

Inicialmente, serão geradas redes para alguns problemas de *benchmark* baseados em funções deceptivas [Goldberg, 1989], tais problemas apresentam uma modularidade conhecida, dessa forma, serão utilizados como base para a escolha do método de geração do modelo em redes.

2 Metodologia

Nesta seção descreve-se a abordagem utilizada para modelagem problema descrito na Seção 2.1 em uma rede. O procedimento se divide em duas etapas, a amostragem do espaço de soluções (Seção 2.2) e a construção da rede (Seção 2.2.1).

2.1 Problemas Abordados

Como forma de validação dois tipos de problemas serão abordados neste trabalho, um problemas deceptivo de alta modularidade, e um problema chamado `one-max` onde não existe relações entre variáveis e desta forma a modularidade deve ser pequena.

As funções deceptivas são problemas de otimização geralmente utilizados para *benchmark* de Algoritmos de Estimação de Distribuição [Larrañaga and Lozano, 2002]. A Equação 2 descreve uma função deceptiva chamada *trap4*, nesse problema de otimização, cada conjunto de 4 variáveis consecutivas possui alta dependência (definida pela Equação 3).

Como pode-se perceber, esta função é um típico caso de função em que a hipótese de *Building-blocks* é verdadeira, de fato, em uma análise superficial tem-se que o espaço de soluções descrito por ela contém (para variáveis binárias) 2^l soluções, enquanto que, conhecidas as variáveis contidas em cada um dos m blocos de tamanho k, a busca se reduz a $2^k m$ combinações, uma diferença de várias ordens de magnitude. A Figura 2 representa a variação dos valores da função objetivo de acordo com a função da Equação 3.

$$\max \ f_{trap4}(x), \ x \in \{0,1\}^{mk}, \tag{2}$$

$$\begin{aligned} f_{trap4}(x) &= \sum_{i=0}^{m-1} trap4(x_{ki} + x_{ki+1} + ... + x_{ki+k-1}), \\ trap4(u) &= \begin{cases} 4 & \text{if } u = 4, \\ 3-u & \text{if } u < 4\,. \end{cases} \end{aligned} \tag{3}$$

O problema `one-max` é um problema simples onde o objetivo é maximizar o número de variáveis com o valor 1. Esse tipo de função é utilizada aqui como forma de avaliar a rede Bayesiana obtida em um problema onde não existe relação entre as variáveis, o que espera-se seja mostrado é que o padrão de conexões entre as variáveis neste contexto não representa informações importantes.

Ao encontrarmos uma modelagem deste problema em redes que seja representativa, ou seja, mantenha características reais do problema, provavelmente teremos uma rede com alta modularidade, sendo cada módulo uma representação dos grupos de variáveis relacionados pela Equação 2. A partir desse ponto será possível utilizar tal procedimento de modelagem em problemas reais, o que é o objetivo principal deste trabalho.

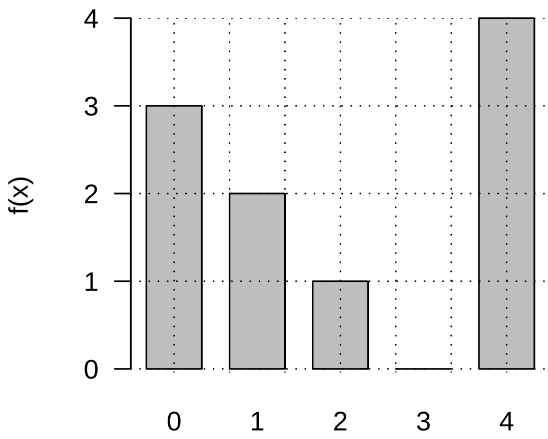


Figura 2: *O conjunto de variáveis é dividido em m subconjuntos de $k = 4$ elementos, a cada um desses subconjuntos a Equação 3 é aplicada e a soma total representa o valor da função objetivo. Esta figura representa a variação dos valores para um desses subconjuntos.*

2.2 Modelagem do Espaço de Soluções

As redes Bayesianas já são a algum tempo utilizadas para a modelagem de problemas de otimização no contexto de Algoritmos de Estimação de Distribuição (EDAs[2]) [Ahn, Ramakrishna, and Goldberg, 2004; Pelikan, Sastry, and Goldberg, 2005; Pelikan, Goldberg, and Cantu-Paz, 1999]. Nesses algoritmos, uma amostra do espaço de soluções é utilizada para estimar a distribuição e as relações de dependência entre as variáveis, a partir desses modelos probabilísticos novas soluções são geradas para o problema, de modo que, quanto melhor a amostra, melhor a qualidade das soluções que serão geradas. No contexto deste trabalho, no entanto, o objetivo é apenas a modelagem do problema e não a solução do mesmo, desta forma, o foco desta seção é a amostragem e não a geração de soluções.

Foi utilizada uma amostragem do espaço de soluções bem simples, baseada no método de seleção por torneio, intensamente utilizado em algoritmos genéticos. A seleção por torneio recebe como entrada um conjunto de soluções geradas aleatoriamente, S, de tamanho $|S|$, para cada uma dessas soluções o valor da função objetivo f é calculado, a partir desses valores um novo conjunto de soluções S_{novo}, de mesmo tamanho do conjunto original é gerado pelo procedimento a seguir:

1. Selecione aleatoriamente t soluções do conjunto S;
2. Insira a melhor das t soluções no conjunto S_{novo};

[2] *Estimation of Distribution Algorithms*

3. Repita enquanto $|S_{novo}| < |S|$.

O valor t é um parâmetro que pode ser escolhido experimentalmente, neste trabalho assume-se $t = 16$. A seleção por torneio extrai soluções do conjunto S que levam a melhores valores para a função objetivo. Por meio desse procedimento, espera-se que a estrutura do problema também esteja sendo decodificada, de modo que os *Building-blocks* possam ser identificados.

2.2.1 Construção da Rede Bayesiana

As redes Bayesianas são construídas por meio de algoritmos baseados no teorema de Bayes [Jensen, 1996], de acordo com o qual a probabilidade de um evento A ocorrer dado que um evento B ocorreu depende não somente da relação entre A e B, mas também da probabilidade de A medida independentemente. Analogamente, a criação de uma rede Bayesiana para um problema de otimização descreverá a influência que um valor assumido por uma variável x_b terá no valor que deverá ser assumido por outra variável x_a, para que e obtenha soluções de boa qualidade, quanto mais fiel o modelo probabilístico descrito pela rede Bayesiana maior a probabilidade de que boas soluções sejam representadas.

3 Experimentos

Para cada uma das funções (deceptiva e `one-max`) foram avaliados problemas de tamanho $l_i = \{8, 12\}$, enquanto o número de amostras utilizadas foram $|S_i| = \{1500, 2000\}$ para a função deceptiva e $|S_i| = \{750, 1000\}$ para `one-max`. Os gráficos nas seções a seguir descrevem as etapas de construção da rede Bayesiana para cada um dos problemas.

3.1 Função trap4

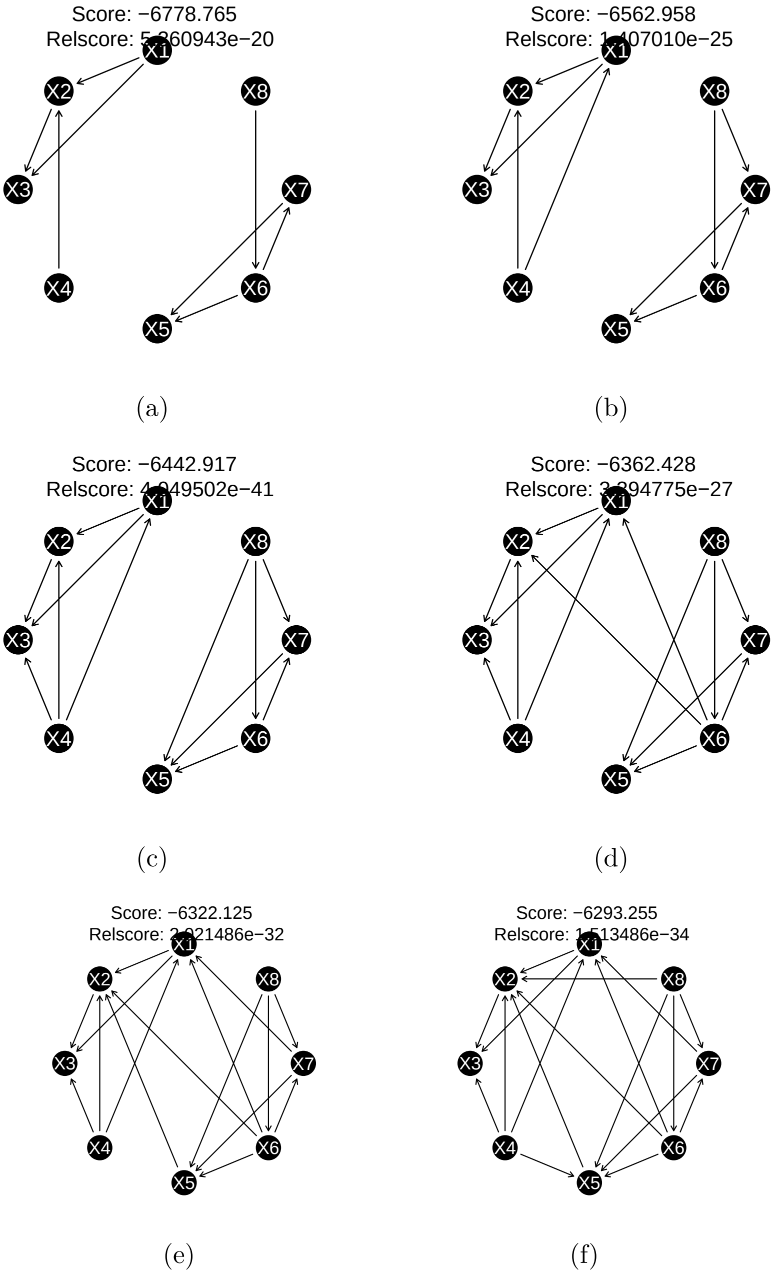
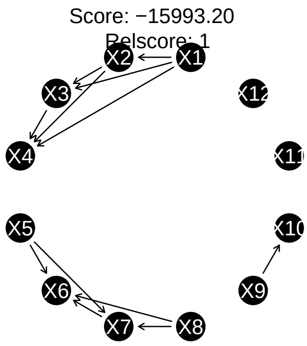
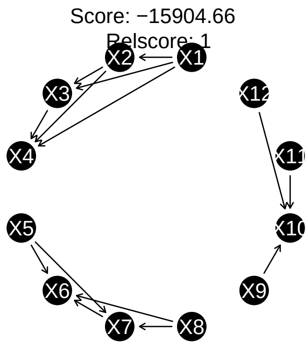


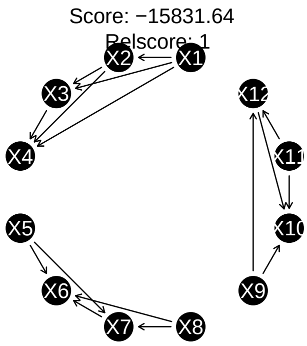
Figura 3: *Etapas de construção da rede Bayesiana para um problema com $l = 8$ e $k = 4$.*



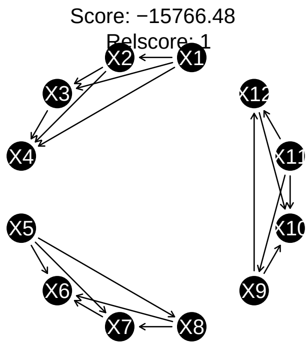
(a)



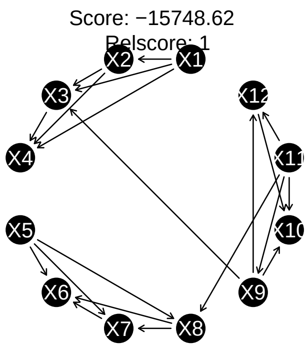
(b)



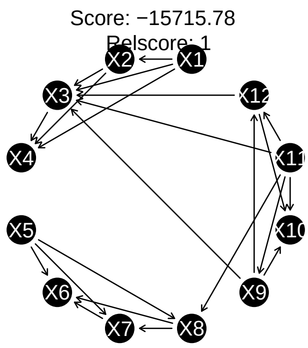
(c)



(d)



(e)



(f)

Figura 4: *Etapas de construção da rede Bayesiana para um problema com $l = 12$ e $k = 4$.*

3.2 One-max

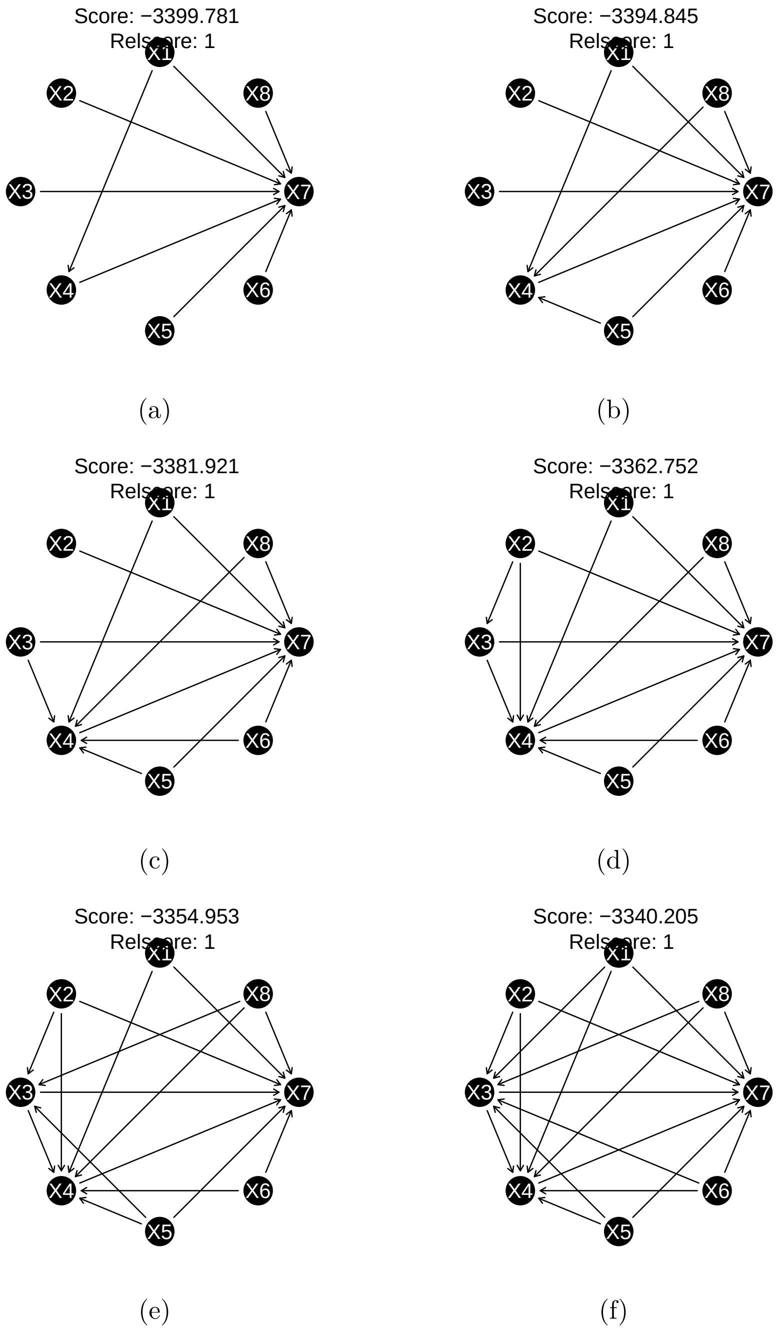


Figura 5: *Etapas de construção da rede Bayesiana para um problema com* $l = 8$.

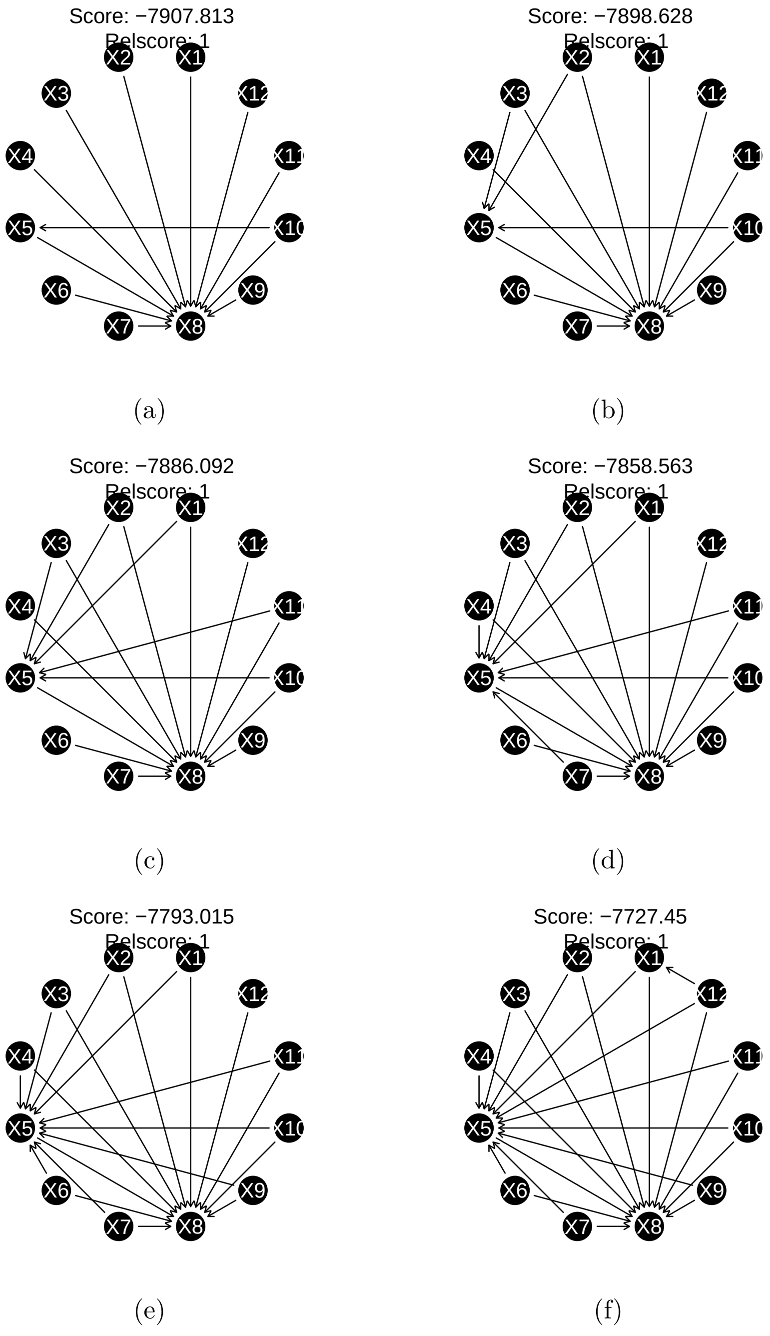


Figura 6: *Etapas de construção da rede Bayesiana para um problema com* $l = 12$.

4 Conclusões

Através das figuras apresentadas nas seções anteriores, evidencia-se que o procedimento de modelagem da estrutura dos problemas foi eficaz, para a função *trap*4 observa-se ao longo da inserção de arestas a evolução da rede em direção a uma característica de alta modularidade, a qual representa as relações entre variáveis e de fato descreve o procedimento de avaliação e o espaço de soluções.

Para que evidências mais fortes quanto a hipótese de *Building-blocks* sejam consideradas é necessária uma etapa posterior ao trabalho aqui apresentado, nesta outra etapa seriam avaliados diversos problemas de otimização com aplicação real como o *Knapsack Problem* e o *Traveling Salesman Problem*. A avaliação desses tipo de problemas permitiria a obtenção de conclusões quanto a aptidão de certos métodos para resolução de determinados problemas, por exemplo, seria esperado que Algoritmos de Estimação de Distribuição obtivessem bons resultados em problemas cuja a estrutura em rede apresente alta modularidade.

Referências

Réka Albert and Albert-László Barabási. Statistical mechanics of complex networks. *Rev. Mod. Phys.*, 74(1):47–97, Jan 2002. doi: 10.1103/RevModPhys.74.47.

Finn V. Jensen. *Introduction to Bayesian Networks*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 1st edition, 1996. ISBN 0387915028.

W. Michiels, E. Aarts, and J. Korst. *Theoretical aspects of local search*. Springer-Verlag New York Inc, 2007.

C.H. Papadimitriou and K. Steiglitz. *Combinatorial optimization: algorithms and complexity*. Dover Pubns, 1998.

D.E. Goldberg. *Genetic algorithms in search, optimization, and machine learning*. Addison-wesley, 1989. ISBN 0201157675.

J.H. Holland. *Adaptation in natural and artificial systems*. MIT Press Cambridge, MA, USA, 1992. ISBN 0262581116.

G. Harik, F. Lobo, and K. Sastry. Linkage learning via probabilistic modeling in the extended compact genetic algorithm (ecga). *Scalable Optimization via Probabilistic Modeling*, pages 39–61, 2006.

T. Bäck, J.M. De Graaf, J.N. Kok, and W.A. Kosters. Theory of genetic algorithms. In *Current trends in theoretical computer science*, pages 546–578. World Scientific Publishing Co., Inc., 2001. ISBN 9810244738.

P. Larrañaga and J.A. Lozano. *Estimation of distribution algorithms: a new tool for evolutionary computation*. Springer Netherlands, 2002. ISBN 0792374665.

C.W. Ahn, RS Ramakrishna, and D.E. Goldberg. Real-coded Bayesian optimization algorithm: Bringing the strength of BOA into the continuous world. In *Genetic and Evolutionary Computation–GECCO 2004*, pages 840–851. Springer, 2004.

M. Pelikan, K. Sastry, and D.E. Goldberg. Multiobjective hBOA, clustering, and scalability. In *Proceedings of the 2005 conference on Genetic and evolutionary computation*, pages 663–670. ACM, 2005. ISBN 1595930108.

Martin Pelikan, David E. Goldberg, and Erick Cantu-Paz. Boa: The bayesian optimization algorithm. pages 525–532. Morgan Kaufmann, 1999.