

V Congresso Anual
Sociedade Portuguesa de Estatística
Curia, 11 a 14 de Junho de 1997

Autómatos Estocásticos como Modelo de um Método de Optimização Combinatória

António José Batel Anjo
Departamento de Matemática
Universidade de Aveiro

Maria Rosália Dinis Rodrigues
Departamento de Matemática
Universidade de Coimbra

Resumo:

Durante a década de oitenta surgiu uma nova classe de métodos de aproximação para problemas de Optimização Combinatória, inspirados na observação de fenómenos físicos, habitualmente conhecidos por Meta-Heurísticas. Em particular o Método de Simulação da Têmpera que procura reproduzir o processo Termodinâmico da Recristalização. Neste trabalho mostramos como os Autómatos Estocásticos constituem um modelo adequado à formalização do Método de Simulação Têmpera.

1- Autómatos Estocásticos

O conceito de Autómato Estocástico (AE) surgiu nos anos 60 no âmbito do Controlo Adaptativo e, posteriormente, foi-se estendendo a outras áreas, entre as quais a Optimização Combinatória.

Um autómato é um sextuplo $\{Y, A, X, \Pi, \eta, a_1\}$ onde, Y é o conjunto de entrada, A é o conjunto de estado internos, X o conjunto de saída ou acções, $\Pi : A \times Y \rightarrow A$, $\eta : A \times Y \rightarrow X$ e a_1 é o estado inicial do autómato. Tendo em conta a natureza das aplicações envolvidas, podemos definir dois tipos de autómatos.

Definição 1:

Se ambas as aplicações forem determinísticas, o autómato resultante é determinístico.

Definição 2:

Se alguma destas aplicações for probabilística, então o autómato é estocástico.

Os Autómatos Estocásticos (AE) podem ser classificados segundo as características das aplicações envolvidas.

	Π	
η	Determinística	Probabilística Tipo-G
	Probabilística	Tipo-H Tipo-GH

Se a aplicação estocástica $\eta : A \rightarrow X$ for independente da acção Y então o autómato é chamado Tipo-GH⁺.

A natureza da entrada pode também caracterizar os AE. Assim se a entrada for binária, uma colecção de objectos distintos ou estiver contida em $[0,1]$ então o autómato será Modelo P, Modelo Q ou Modelo S, respectivamente.

Por último, se as matrizes de transição $[\Pi_{ij}]$ e de acção $[\eta_{ij}]$ forem invariantes no tempo, então diz-se que o autómato tem uma estrutura constante. Caso contrário, o autómato tem uma estrutura variável.

2- Algoritmos de Optimização e Autómatos Estocásticos

Um problema de Optimização Combinatória consiste essencialmente em, de entre um espaço finito de soluções, identificar um conjunto particular, para o qual a função objectivo apresenta um valor mínimo global.

De entre os Métodos de Pesquisa Local, denominamos Métodos com Capacidade de Tolerância (MCT) os que têm a particularidade de aceitar algumas soluções cujo valor da função objectivo é superior (no caso de estarmos a minimizar) ao valor da solução corrente. A decisão de "subir" é baseada numa determinada lei de probabilidade.

Algoritmo Geral dos MCT

{ i = solução inicial, escolhida no espaço das configurações }

{ T = parâmetro dependente do tempo }

repetir

Gerar j pertencente a uma vizinhança de i

se $\Delta_k = (f(j) - f(i)) < 0$

então Aceitar j

senão Aceitar j , com probabilidade $g(\Delta_k, T)$

até critério de paragem

Um autómato Modelo-P admite como variável de entrada Y dois valores possíveis: 0 e 1. O entrada $y=0$ é atribuída quando a diferença da função objectivo entre a solução vizinha e a solução corrente diminui, $\Delta_k < 0$. A entrada $y=1$ é atribuída no caso desta diferença aumentar. As alterações no estado do autómato como resposta à entrada Y são descritas por matrizes de probabilidade de transição. Cada elemento desta matriz é dado por $p_{ij} = c_i p_{ij}^1 + (1-c_i) p_{ij}^0$, a probabilidade c_i está associada à escolha da solução vizinha e as probabilidades p_{ij}^0 e p_{ij}^1 serão definidas em seguida.

Teorema 1:

Um MCT é um autómato estocástico Tipo-GH Modelo-S, com estrutura variável.

Demonstração:

Tomamos como estados internos do autómato $A=\{a_0,a_1,\dots,a_N\}$ as soluções do problema de optimização. O conjunto de acções é determinado pelo conjunto de vizinhos de cada solução a_i . A aplicação Π produz o estado associado a cada entrada e estado anterior. Esta aplicação é definida por duas matrizes de probabilidades de transição $P_0 = [p_{ij}^0]$ quando $\square_k \leq 0$ e $P_1 = [p_{ij}^1]$ quando $\square_k > 0$.

De acordo com a noção de MCT e de autómato, cada elemento da linha i da matriz P_0 é dado por:

$$p_{ij}^0 = \begin{cases} \frac{1}{|A_i|}, & j \in A_i \\ 0 & , j \notin A_i \end{cases}$$

A matriz P_1 é definida por:

$$p_{ij}^1 = \begin{cases} \frac{P(\text{aceitar } j)}{|A_i|}, & j \in A_j, i \neq j \\ 1 - \sum_{m=1, m \neq i}^{|A_i|} \frac{P(\text{aceitar } j)}{|A_i|}, & i = j \\ 0 & , j \notin A_i \end{cases}$$

Se a matriz P_0 for determinística, P_1 estocástica e $P(\text{aceitar } j)$ a probabilidade de aceitar transições de "subida" então a aplicação Π é estocástica.

Cada acção x_i do AE consiste na selecção de uma solução na vizinhança da solução corrente i , assim a aplicação η é também estocástica uma vez que as sucessivas acções são geradas aleatoriamente de acordo com uma lei uniforme. Deste modo temos um autómato Tipo-GH Modelo-S (o facto de ser Modelo-S, resulta apenas de uma transformação sobre a entrada).

Finalmente, como o parâmetro T varia com o tempo o MCT tem uma estrutura variável.

As cadeias de Markov são usadas como modelos das sucessivas transições inter-estados que ocorrem num AE Modelo-GH.

Teorema 2:

A cadeia de Markov induzida pelo MCT, com T constante, num espaço finito de objectos combinatorios é ergódica.

Demonstração:

-Para provar este teorema é necessário mostrar que a cadeia de Markov é irredutível, aperiódica e recorrente positiva ([Fel66], I Volume).

Uma cadeia de Markov é irredutível se e só se todos os estados comunicarem ($p_{jj}^n > 0$ e $p_{ji}^m > 0$, para todos os estados). Se existir uma métrica no espaço dos objectos combinatórios tal que a distância entre qualquer par de soluções seja finita e positiva e se o MCT satisfizer as duas condições seguintes:

- a função de geração constrói qualquer uma das soluções pertencentes a uma esfera de raio $\rho=1$, com probabilidade uniforme;
 - a função de aceitação da transição do estado i para o seu vizinho j ocorre com probabilidade $p_{ji} > 0$,
- então, existe um número mínimo de transições entre dois quaisquer estados consecutivos, $n = \rho(i,j) = \rho(j,i) > 0$. Assim, todos os estados comunicam e a cadeia diz-se irredutível. Numa cadeia de Markov irredutível e finita todos os estados são recorrentes.

Seja i uma solução para a qual $f(i)$ é mínimo. Então, para qualquer solução j , $f(i) \leq f(j)$ e, de acordo com as condições estabelecidas, $p_{ij}^1 > 0$. Portanto o m.d.c $\{n : p_{ii}^n > 0\} = 1$ e a cadeia de Markov aperiódica.

Nestas circunstâncias podemos calcular a distribuição estacionária da Cadeia de Markov.

A classe dos AE Tipo-GH Modelo-S engloba muitos outros métodos de resolução de problemas de optimização combinatória.

3- Simulação da Têmpera e Autómatos Estocásticos

Na secção anterior mostramos que o MCT, com o procedimento de Geração e Aceitação especificado segundo algumas restrições, é um autómato estocástico do Tipo-GH e que as transições entre estados são representadas numa cadeia de Markov com distribuição estacionária.

Em particular, dada a analogia ao fenómeno da Termodinâmica que esteve na origem, no Método de Simulação da Têmpera (MST) a função de aceitação é definida pela função distribuição de Boltzmann, $g(\Delta, T) = e^{-\Delta/T}$, onde o parâmetro T representa a temperatura, que é feita descer ao longo do tempo.

Teorema 3:

O MST é um autómato estocástico Tipo-GH Modelo-S com estrutura variável.

Demonstração:

Este método é um caso particular do método MCT.

4- Convergência do MST

Existem vários trabalhos publicados cujo objectivo fundamental é a prova de convergência do MST ([LM96], [AK90] e [LA87]). Todos eles garantem que sob certas condições, sobre as funções de Geração e de Aceitação, o MST converge assintoticamente para um conjunto de ótimos globais, com probabilidade 1.

Teorema 4:

O MST é um autómato estocástico assintoticamente óptimo Tipo-GH Modelo-S com estrutura variável.

Demonstração:

A prova de convergência resulta do que foi escrito na demonstração anterior e da definição de autómato assintoticamente óptimo.

Agradecimentos

Os autores agradecem ao *refree* as pertinentes críticas e sugestões que contribuíram para uma significativa melhoria deste trabalho.

Bibliografia

[AK90]E. H. Aarts and J. Korst.

Simulated Annealing and Boltzmann Machines - A Stochastic Approach to Combinatorial Optimization and Neural Computing. John Wiley & Sons, Inc., 1990.

[Fel66]W. Feller.

An Introduction to Probability Theory and its Applications, Volume I and II. John Wiley & Sons, Inc., 1966.

[IK96]H. O. Ibrahim and J.P. Kelly.

Meta-Heuristics: Theory & Applications. Kluwer Academic Publishers, United States of America, 1996.

[KGV83]S.Kirkpatrick, C.D. Gelatt, and M.P. Vecchi.

Optimization by simulated annealing. Science, 220(4598):671--680, 1983.

[LA87]P.J.M. Laarhoven and E.H. Aarts.

Simulated Annealing: Theory and Applications. D. Reidel Publishing Company, 1987.

[LM86] M.Lundy and A.Mess.

Convergence of an annealing algorithm. Mathematical Programming - North-Holand, 34:111--124, 1986.

[PS87] C.H. Papadimitriou and K. Steiglitz.
Combinatorial Optimization, Algorithms and Complexity. Prentice-Hall, Inc., New York, 1987.

[RA87] R. Rodrigues and A. Anjo.
On Simulating Thermodynamics, Springer Verlag, 396 (45-61). Lecture Notes in Economics and Mathematical Systems - Applied Simulated Annealing, 1993.

[SL90] E. Shragowitz and R. Lin.
Combinatorial Optimization by Stochastic Automata. Annals of Operations Research, 22 (293-324), 1990.