

Otimização por Enxame de Partículas

Alan Godoy Souza Mello
godoy@dca.fee.unicamp.br

Laboratório de Bioinformática e Computação Bio-Inspirada
Departamento de Engenharia de Computação e Automação Industrial
Faculdade de Engenharia Elétrica e de Computação
Universidade Estadual de Campinas

7 de maio de 2012



- 1 Cognição social
- 2 Auto-organização
- 3 Otimização por Enxame de Partículas

Cognição social

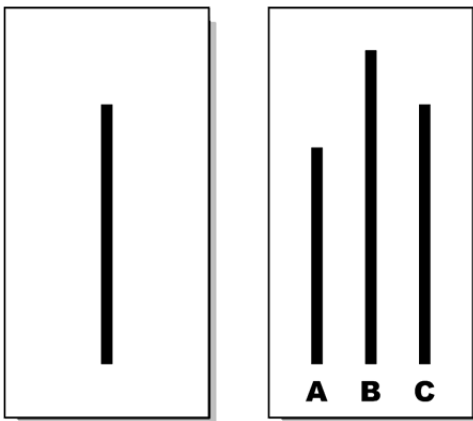


Figura: Experimento de conformidade de Asch

- Pensar é um ato social:

“Although some might claim that the brain as the physical site of mental processing requires that we treat cognition as a fundamentally individual and even private activity, we are prepared to argue that all mental activity – from perceptual recognition to memory to problem solving – involves either representations of other people or the use of artifacts and cultural forms that have a social history” – Levine, Resnick e Higgins (1993)



- Crenças, memórias, atitudes e processos de pensamento são fortemente influenciados pelo meio em que o indivíduo se encontra.
- Experimentos indicam que memórias formadas por sugestão são indistinguíveis das memórias de fatos vivenciados pelos indivíduos através de suas percepções (Loftus, 1994).
 - Memória como construção social.
- Assim, a cultura tem um papel fundamental para a compreensão das capacidades de pensamento individuais.

Como se formam as culturas?

- Modelo cultural de Axelrod (1997)
 - Indivíduos são formados por vetores de características.
 - A organização dos indivíduos em uma grade, tendo vizinhos acima, abaixo e aos lados.

27217	74924	31157	53671	22660	37316	07959	57666	33206	92725
66219	08226	26707	45600	48767	39481	62784	89859	27792	35492
37262	66163	89178	60968	91098	19937	62103	07562	03500	13864
87746	66209	94122	72784	03593	16647	19776	87819	22160	48185
16880	09713	76057	30843	92125	41152	74156	98801	64760	00144
86287	66161	23271	46773	53014	44442	25424	98309	32553	16678
90624	65685	68785	32385	90770	24676	68806	25347	16640	30602
98681	11402	57304	68003	16943	01041	44693	63237	76040	61075
52249	30617	91425	92780	82342	30467	19721	84117	96595	55215
79949	70851	29089	89311	19176	67653	95954	64805	51332	74301

Tabela: Estado inicial

- Modelo cultural de Axelrod

- Probabilidade de interação com vizinho é proporcional à quantidade de características iguais.
- Na ocorrência de uma interação, o indivíduo “imita” uma das características do vizinho.
- Passos: avaliar, comparar e imitar

27217	74924	31157	53671	22660	37316	07959	57666	33206	92725
66219	08226	26707	45600	48767	39481	62784	89859	27792	35492
37262	66163	89178	60968	91098	19937	62103	07562	03500	13864
87746	66209	94122	72784	03593	16647	19776	87819	22160	48185
16880	09713	76057	30843	92125	41152	74156	98801	64760	00144
86287	66161	23271	46773	53014	44442	25424	98309	32553	16678
90624	65685	68785	32385	90770	24676	68806	25347	16640	30602
98681	11402	57304	68003	16943	01041	44693	63237	76040	61075
52249	30617	91425	92780	82342	30467	19721	84117	96595	55215
79949	70851	29089	89311	19176	67653	95954	64805	51332	74301

22233	22233	22233	22233	22233	22233	22233	22233	22233	22233
22233	22233	22233	22233	22233	22233	22233	22233	22233	22233
22233	65955	65955	22233	22233	22233	22233	33588	33588	33588
22233	22233	22233	22233	22233	22233	22233	33588	33588	33588
22233	22233	22233	22233	22233	22233	22233	22233	22233	22233
22233	22233	22233	22233	22233	22233	22233	22233	22233	22233
22233	22233	22233	22233	22233	22233	22233	13157	22233	22233
22233	22233	22233	22233	22233	22233	22233	13157	22233	22233
22233	22233	22233	22233	22233	22233	22233	22233	22233	22233
22233	22233	22233	22233	22233	22233	22233	22233	22233	22233

Tabela: Estado final

- Assim como ocorre com a cultura, formam-se agrupamentos locais que compartilham características comuns.

- E se alterarmos o algoritmo para que os agentes busquem resolver algum problema, imitando seu vizinho não mais quando ele é similar, mas quando ele é mais bem-sucedido?

[illegible]

Tabela: Resolução do problema do caixeiro viajante (Kennedy, 1997)

- Apesar de nenhum agente individualmente saber como resolver o problema, eles foram capazes coletivamente de resolver um problema relativamente complexo.
- A cognição move-se da mente individual para as relações interpessoais.
- Observa-se, também, que a cultura otimiza a cognição: as interações são locais, mas as inovações são transmitidas para toda a sociedade através da cultura (*no-trial learning*).
- Conforme as habilidades vão se espalhando pela sociedade, a população converge para processos ótimos (Bandura, 1965).
- A sociedade pode ser vista, portanto, como uma entidade distribuída de processamento de informação.



- Verifica-se em sociedades humanas de forma similar a capacidade de resolver eficientemente problemas de forma descentralizada:
 - Organização em tribos: alguns responsáveis pela caça, outros pelo cultivo e outros pelo cuidado com as crianças.
 - Sociedade moderna e o trabalho especializado: aumento de eficiência mas também aumento na interdependência.
 - Economia de livre-mercado: a composição de decisões individuais é capaz de prover uma alocação eficiente de recursos, sendo raras as ocorrências de desabastecimentos, mesmo que nenhum indivíduo tenha conhecimento completo (ou mesmo suficiente) das necessidades de cada pessoa.
- O funcionamento de sociedades humanas é, portanto, um processo auto-organizado.

Auto-organização

- Conjunto de mecanismos dinâmicos pelos quais múltiplos elementos conseguem estabelecer ordem sem imposição por quaisquer forças externas.
- Ingredientes principais: realimentação positiva e negativa, amplificação das perturbações e múltiplas interações (Bonabeau et al, 1999).
- Aumento da capacidade coletiva:
 - O todo é mais do que a mera soma dos indivíduos.
- Regras muito simples: economia para codificação.
- Forte interação dos elementos com seus vizinhos.
- Uso apenas de informações locais, sem referência ao padrão global almejado.

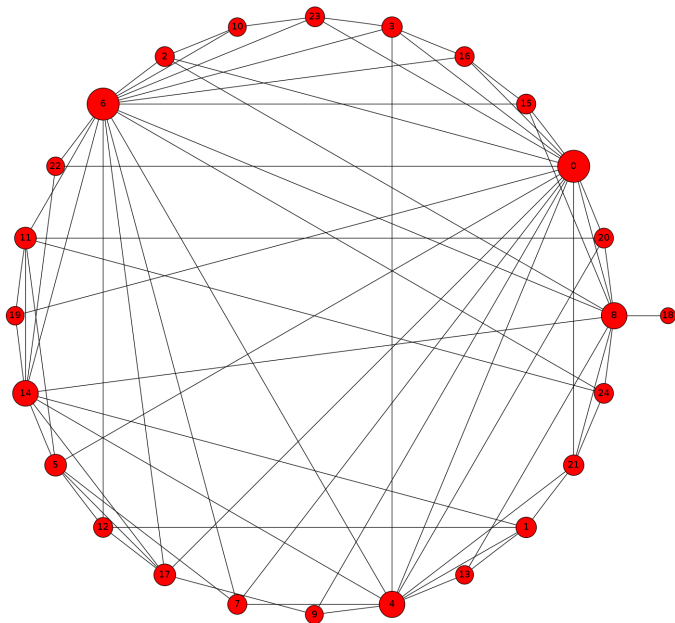


- O conjunto como um todo é a principal estrutura de controle, e não um líder centralizado.
- Demanda poucos recursos dos componentes do sistema.
- Baixas exigências no que tange a estrutura de comunicação.
- Ordem robusta e flexível:
 - Permite respostas diferentes a cenários diferentes, mesmo sem a codificação explícita nas regras do sistema.

Otimização por Enxame de Partículas

- Metaheurística populacional para a otimização de funções contínuas baseada em mecanismos de simulação do comportamento social de animais, como revoadas de pássaros.
- Utiliza como metáfora o modo de influência em formas de pensar e agir entre seres humanos.
- Conjunto de partículas navegando pelo espaço de soluções.
- Cada partícula estabelece sua trajetória combinando suas experiências passadas com as experiências de seus vizinhos (outras partículas com as quais elas se comunicam).





- No PSO, a “crença” de um indivíduo é representada por um vetor x_i (posição da partícula no espaço).
- A cada passo, as partículas têm suas posições atualizadas de acordo com suas velocidades v_i :

$$x_i \leftarrow x_i + v_i$$

- Após calculada a nova posição a partícula i avalia a qualidade de seu x_i , comparando com a melhor posição que a partícula já visitou, p_i . Caso x_i seja melhor que p_i , a melhor posição é atualizada para a posição atual.



- A velocidade de cada partícula no passo atual é calculada considerando-se a velocidade no passo anterior e a posição de seus dois atratores: p_i , a melhor posição já visitada pela partícula, e p_g , a melhor posição já visitada por algum dos vizinhos da partícula:

$$v_j \leftarrow \omega v_j + c_1 r_1 (p_j - x_j) + c_2 r_2 (p_g - x_j)$$

- As constantes ω , c_1 e c_2 , são chamadas, respectivamente, de fator de inércia, fator cognitivo e fator social, e r_1 e r_2 são vetores aleatórios, cujos elementos assumem valores no intervalo $[0, 1]$.
 - ω : introduz uma espécie de “atrito” no movimento de cada partícula, reduzindo a velocidade inercial.
 - c_1 : “nostalgia” da partícula.
 - c_2 : indica o quanto uma partícula será influenciada por suas vizinhas.

Algoritmo (versão para minimização)

Inicializar a posição x_i , a velocidade v_i e a melhor posição pessoal p_i das N partículas;

enquanto o critério de parada não é satisfeito **faça**

para $j \leftarrow 1$ até N **faça**

$g \leftarrow \arg \min_{i \in \text{vizinhos}(j)} f(p_i)$;

$v_j \leftarrow \omega v_j + c_1 r_1 (p_j - x_j) + c_2 r_2 (p_g - x_j)$;

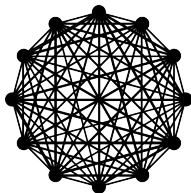
$x_j \leftarrow x_j + v_j$;

se x_j está dentro do espaço de busca e $f(x_j) < f(p_j)$ **então**

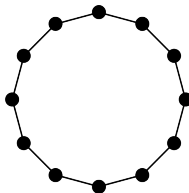
$p_j \leftarrow x_j$;

Vizinhança

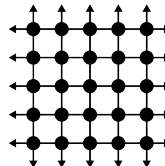
- A vizinhança normalmente é topológica, não espacial.



Global



Anel



von Neumann

Outras vizinhanças

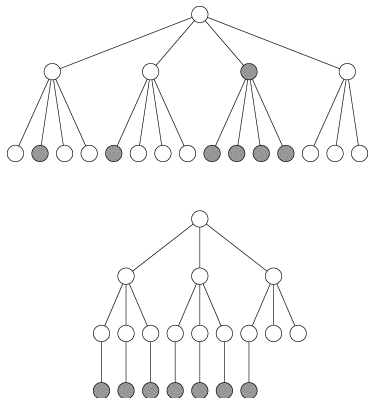


Figura: H-PSO

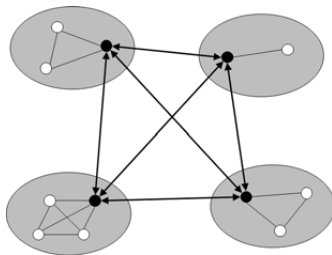


Figura: TRIBES

Aplicações

- Engenharia: ajuste de sistemas de recomendação, modelagem de antenas, projeto de sistemas de controle e de circuitos VLSI.
 - Resultados muito bons no treinamento de redes neurais e de redes neuro-fuzzy.
- Biologia: Inferência de redes gênicas, reconstrução de árvores filogenéticas, otimização de modelos biomecânicos.
- Artes: improvisação musical.



- Exemplo de processo totalmente auto-organizado:
 - Unidades simples de processamento, através de intensa comunicação entre si e interação com o ambiente, cooperando para a obtenção de um objetivo comum.
 - Partículas isoladas tem limitado poder de otimização: a capacidade do enxame é uma qualidade emergente.
 - Ausência de um controle centralizado, como a evolução natural.
- Pontos positivos: técnica simples, robusta, flexível, altamente distribuída, com baixas exigências de memória e poder de processamento e de rápida convergência a um ótimo.
- Ponto negativo: rápida perda de diversidade e convergência prematura a ótimos locais.



Comparação com Computação Evolutiva

- A Otimização por Enxame de Partículas é considerada como uma técnica de Computação Evolutiva, apesar de trazer uma metáfora diferente da evolução das espécies.
- As influências pessoais e sociais guardam semelhança com operadores de recombinação e de *crossover*, enquanto que a conservação de movimento da partícula atua como uma espécie de mutação direcional.
- Há diferenças, no entanto, sendo a maior delas o papel da seleção natural: enquanto que métodos evolutivos têm como parte essencial a morte dos indivíduos menos aptos, em um processo social os indivíduos são preservados durante a execução do algoritmo, de forma que o próprio indivíduo se adapta no decorrer do tempo.



Modificações

- Modelo de influência (exemplos: *Fully Informed Particle Swarm* – FIPS).

$$v_j \leftarrow \omega v_j + \frac{\sum_{i \in \text{vizinhos}(j)} W(i) c_i (p_i - x_j)}{\sum_{i \in \text{vizinhos}(j)} W(i) c_i}$$

- Regra de atualização (exemplos: *Bare Bones Particle Swarm*, AR-PSO).
- Topologias dinâmicas, como o DMS-PSO, o TRIBES e o FDR-PSO.

Problemas multiobjetivo

- Modificações: não há mais uma solução que pode ser considerada ótima, mas um conjunto na fronteira de Pareto – como escolher o atrator?
 - Convergência prematura torna-se um problema maior: atração rápida a solução única leva a menor cobertura do Pareto.
- Arquivo de soluções não-dominadas.



PSO multiobjetivo (Reyes-Sierra E Coello-Coello, 2006)

Inicializar a posição x_i , a velocidade v_i e a melhor posição pessoal p_i das N partículas;

Escolher os líderes em um arquivo externo;

enquanto o critério de parada não é satisfeito **faça**

para $j \leftarrow 1$ até N **faça**

 Selecione o líder g do arquivo segundo algum critério;

$v_j \leftarrow \omega v_j + c_1 r_1 (p_j - x_j) + c_2 r_2 (p_g - x_j)$;

$x_j \leftarrow x_j + v_j$;

$x_j \leftarrow mutacao(x_j)$;

se x_j está dentro do espaço de busca e $Q(x_j) < Q(p_j)$

então

$p_j \leftarrow x_j$;

Atualizar o arquivo de soluções não-dominadas;

- Otimizando os objetivos:

- Agregação de todos os objetivos em um único valor:

$$F(x) = \sum_i w_i f_i(x)$$

- Ordenação lexicográfica: um objetivo é otimizado por vez.
- Sub-enxames: cada enxame otimiza um objetivo (mantendo a troca de informação entre os sub-enxames).
- Baseado em Pareto: mais usual, todos os objetivos são otimizados ao mesmo tempo; seleção de um líder dentre soluções em um arquivo.



- Líder: escolha aleatória dentre soluções não-dominadas ou usa-se um critério adicional (cobertura da fronteira de Pareto, por exemplo)?
 - Escolha usualmente se baseia no número de soluções similares ou com resultados semelhantes (o líder seria a solução não-dominada com menos soluções parecidas com ele no arquivo).
 - Possibilidade de três grupos de arquivos: ótimos pessoais, ótimos locais e ótimos gerais (os melhores valores encontrados durante a busca por todas as partículas); na prática costuma ser usado apenas um arquivo geral.
- Mutação: é necessário manter a diversidade do enxame, então usa-se ω para garantir uma variação maior nos resultados ou é aplicado um fator de turbulência.



Referências I



T. Blackwell and J. Branke.

Multi-swarm optimization in dynamic environments.

Lecture Notes in Computer Science, 3005:489–500, 2004.



M. Clerc and J. Kennedy.

The particle swarm - explosion, stability, and convergence in a multidimensional complex space.

Evolutionary Computation, IEEE Transactions on, 6(1):58–73, 2002.



R. Eberhart and J. Kennedy.

A new optimizer using particle swarm theory.

In *Micro Machine and Human Science*, 1995. *MHS '95., Proceedings of the Sixth International Symposium on*, pages 39–43, 1995.



Referências II



R. C. Eberhart and Y. Shi.

Comparing inertia weights and constriction factors in particle swarm optimization.

In *Evolutionary Computation, 2000. Proceedings of the 2000 Congress on*, volume 1, pages 84–88 vol.1, 2000.



A. Godoy.

Aplicação de redes complexas para a definição de vizinhança na otimização por enxame de partículas.

Master's thesis, Universidade Estadual de Campinas, Campinas, 2010.





Referências III



J. Kennedy.

Thinking is social: Experiments with the adaptive culture model.

The Journal of Conflict Resolution, 42(1):56–76, 1998.



J. Kennedy.

Small worlds and mega-minds: effects of neighborhood topology on particle swarm performance.

In *Congress on Evolutionary Computation*, volume 3, pages 1931–1938, 1999.



J. Kennedy and R. Eberhart.

Particle swarm optimization.

In *Neural Networks, 1995. Proceedings., IEEE International Conference on*, volume 4, pages 1942–1948 vol.4, 1995.

Referências IV



R. Poli.

Analysis of the publications on the applications of particle swarm optimisation.

Journal of Artificial Evolution and Applications, 2008(1):1–10, Jan. 2008.



R. Poli, J. Kennedy, and T. Blackwell.

Particle swarm optimization.

Swarm Intelligence, 1(1):33–57, 2007.



M. Reyes-Sierra and C. A. C. Coello.

Multi-objective particle swarm optimizers: A survey of the state-of-the-art.

International Journal of Computational Intelligence Research 2(3):287–308, 2006.

