

# Computação Evolucionária

**Prof. Heitor Silvério Lopes**

*hslopes@utfpr.edu.br*



# Algumas áreas de aplicação de CE

## # Computação

- Mineração de dados, bancos de dados, engenharia de software, robótica, processamento paralelo...

## # Engenharias

- Eletro-eletrônica, civil, mecânica, desenho industrial

## # Matemática

- Pesquisa operacional, álgebra, séries temporais

## # Biologia

- Bioquímica, ecologia

## # Economia

- Econometria, modelagem

## # Física

- Nuclear, óptica

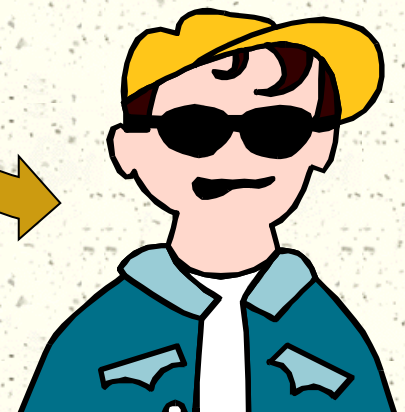
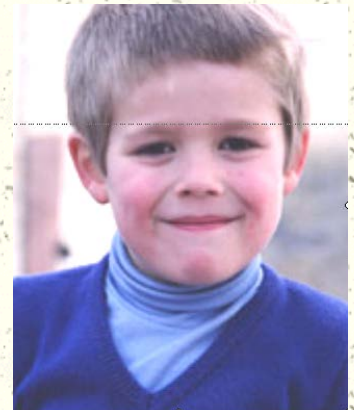
## # Artes

- Música, pintura

# Modelagem de problemas com CE

## # Nível de sofisticação do modelo:

- Muito complexo:
  - # É mais próximo da
  - # Difícil modelagem
  - # Computacionalmente intratável
- Muito simples:
  - # Não retrata a realidade
  - # Facilmente modelável
  - # Computacionalmente tratável



# Alguns projetos desenvolvidos no CPGEI que eu participei

- Navegação de robôs móveis (Fabro, Calvo)
- Equilíbrio de carga na rede secundária de distribuição (Augusto)
- Projeto de amplificadores e filtros analógicos (Kaminski)
- Projeto e sintonia de filtros digitais /eletrônica evolutiva (André, Daniel)
- Decodificação de códigos corretores de erros para transmissão de dados (Fidelis)
- Projeto de enlaces de redes IP (Yabzinski)
- Posicionamento de ERBs em redes wireless (Talau)
- Locação otimizada de equipes de manutenção em redes de distribuição de energia elétrica (Bobel)
- Planejamento da locação de subestações de distribuição (Geraldo)
- Otimização do envio de produtos por poliduto (Jonas)
- Cinemática inversa de manipuladores robóticos (Munif)
- Identificação de processos petroquímicos (Evangelista)
- Otimização de corte na indústria de embalagens (Selow)
- Otimização da linha de produção de veículos (Solivan, Malinowski)
- Identificação de sistemas / modelagem de séries temporais (Wagner)
- Problema do empacotamento múltiplo (Fernanda, Jonas)
- Roteamento múltiplo de veículos (Dalle Molle)
- Mineração de dados para diagnóstico clínico (Célia, Parpinelli, Noda)
- Reconhecimento de padrões em imagens de raios-X (Felizberto)
- Alinhamento de sequências de proteínas (Moritz)
- Dobramento de proteínas (Scapin, Perretto, Bitello, Betito, Benitez, Parpinelli, Marlon)
- Criação de árvores filogenéticas (Perretto)
- Busca e reconhecimento de *motifs* para classificação de proteínas (Denise)
- Indução de autômatos celulares e linguagens formais livres de contexto (Ernesto)
- Registro de imagens/reconhecimento de objetos/faces em imagens (Rafael, Chidambaram, Hugo, Manassés)
- Otimização de estruturas mecânicas (Rafael, Marlon)
- Otimização do roteamento dinâmico em redes veiculares (Rodrigo)
- Inferência de redes gênicas (Leandro)
- ..... não tem mais espaço....



# Alocação de canais em sistemas de comunicações

## # *Frequency Assignment Problem*

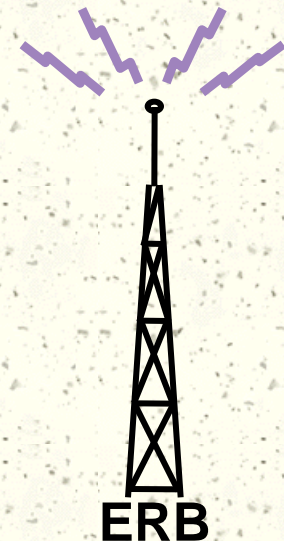
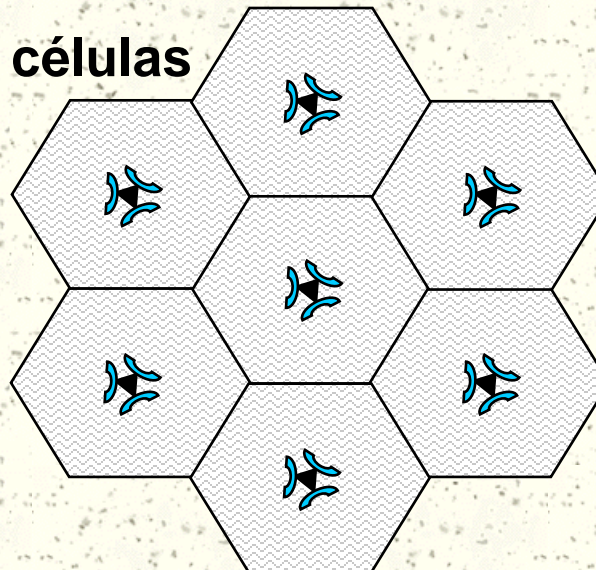
- Problema NP-completo

- Restrições:

- número limitado canais disponíveis no espectro
- tráfego dinâmico / número mínimo de canais por célula
- interferências mínimas inter-células e intra-célula



assinante

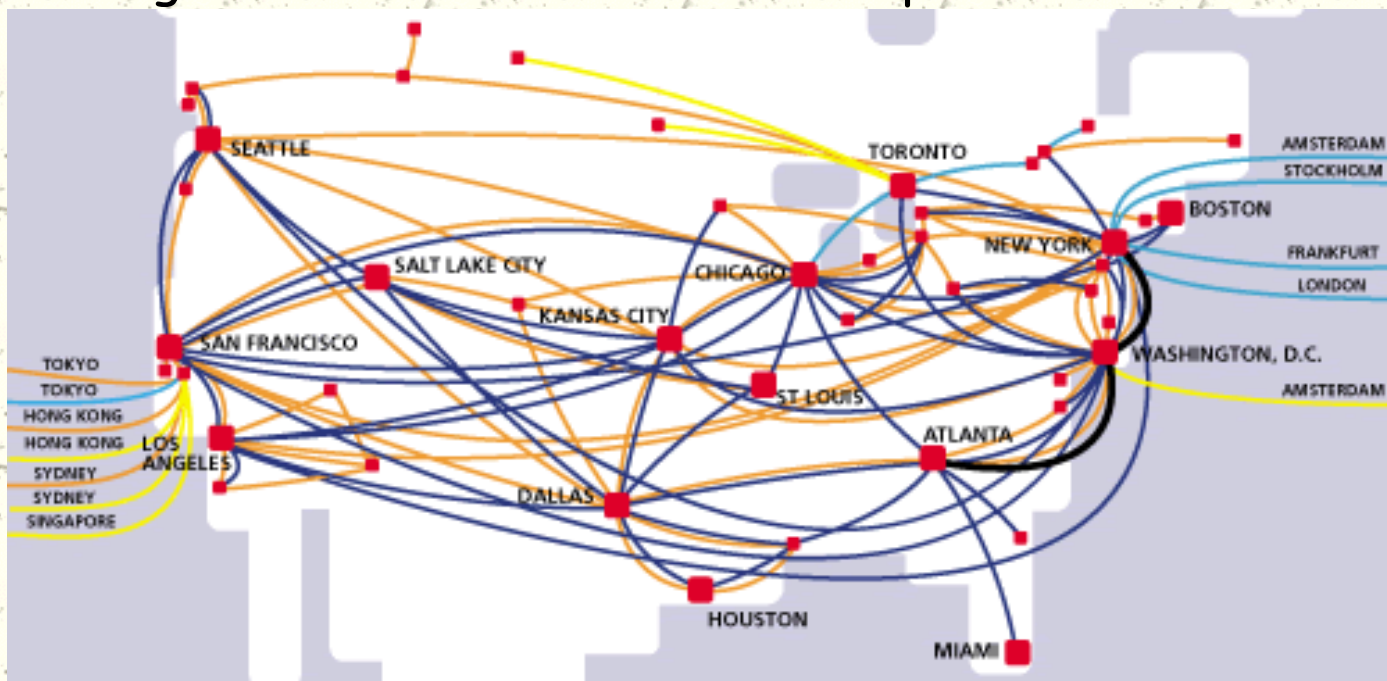




# Roteamento dinâmico em redes de computadores

## \* WANs:

- Altamente interconectadas (PIRs, NAPs)
- Alta confiabilidade; redundância
- Tráfego flutuante; decisões em tempo-real

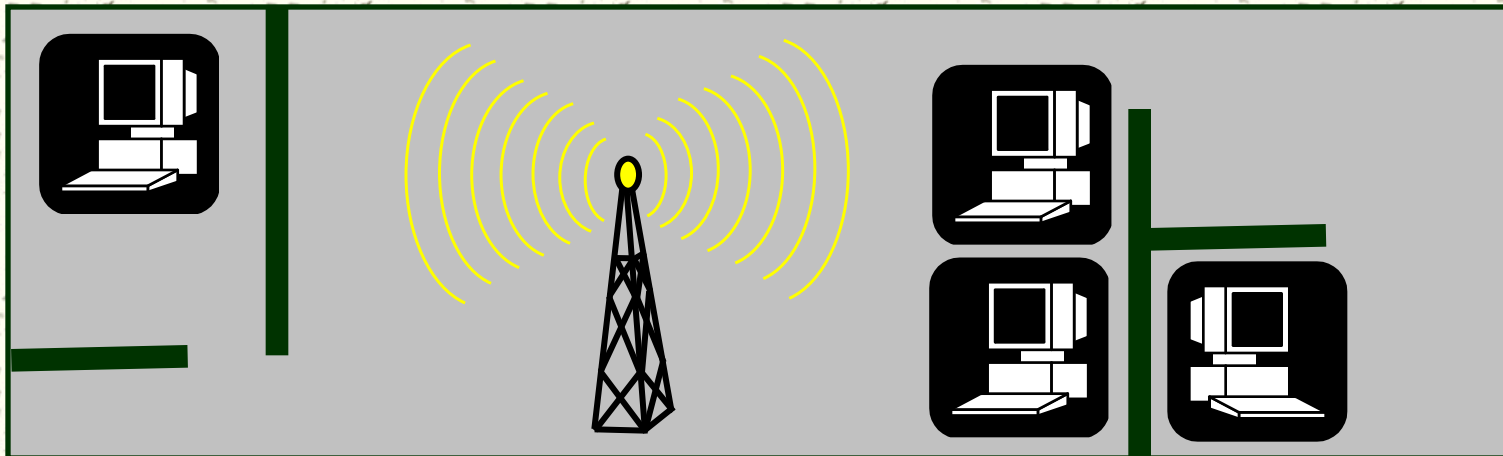




# Planejamento de redes WLAN

## # Wireless LAN:

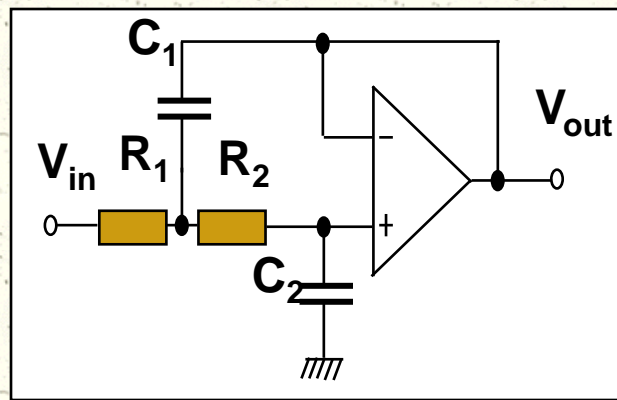
- Substitui o cabeamento físico
- Ambientes industriais; estações móveis...
- Problema: posicionamento das ERBs
- Fatores de projeto:
  - Localização física, densidade de tráfego, propagação do sinal (atenuação e interferências)
- Conjunto de Pareto (multiobjetivos)





# Projeto de circuitos eletrônicos

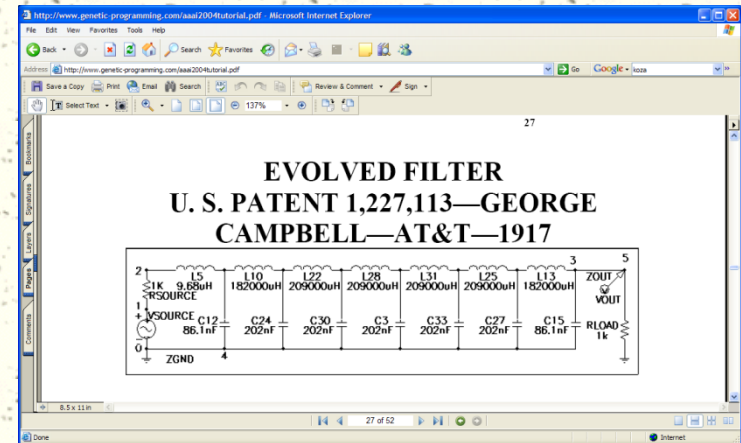
- ✦ Projeto eletrônico:
  - Busca de um conjunto de valores de componentes eletrônicos discretos que compõem um circuito
- ✦ Otimização multiobjetivo:
  - Resposta em frequência, ganho, distorção harmônica, temporização, etc...
- ✦ AG acoplado a um simulador de circuitos
- ✦ Modificação dos valores e da estrutura:
  - Programação Genética



# Projeto de circuitos eletrônicos

## ✦ Programação Genética:

- Muita “força-bruta”, pouca inteligência especializada (minha opinião)
- Foram obtidas várias patentes com este método
- Muitos circuitos conhecidos foram “reinventados”



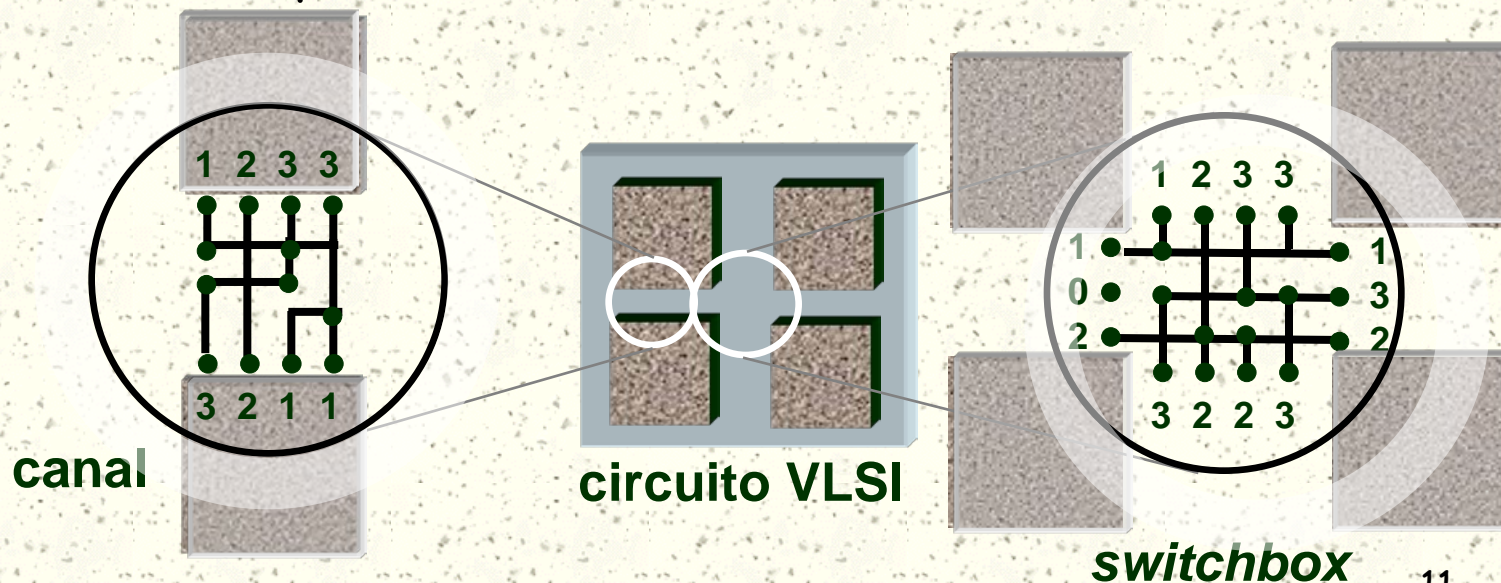
The screenshot shows a web browser window displaying a table titled "2 PATENTABLE INVENTIONS CREATED BY GENETIC PROGRAMMING". The table lists two inventions created by genetic programming, their patent application dates, and the inventors.

	Claimed invention	Date of patent application	Inventors
1	Improved general-purpose tuning rules for a PID controller	July 12, 2002	Martin A. Keane, John R. Koza, and Matthew J. Streeter
2	Improved general-purpose non-PID	July 12, 2002	Martin A. Keane, John R. Koza, and Matthew J. Streeter

The browser window shows the address bar with the URL "http://www.genetic-programming.com/aaa2004tutorial.pdf" and the page number 30 of 52.

# Projeto de circuitos integrados

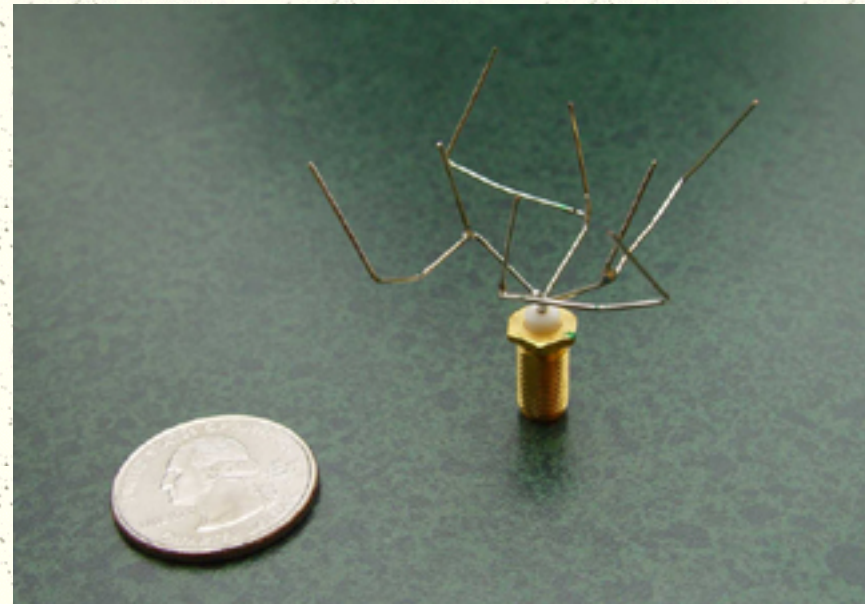
- # VLSI - *very large scale integration*
  - *Softwares e workstations* de alto custo
- # Restrições:
  - Elétricas: *Crosstalk* entre trilhas, tempo de atraso
  - Físicas: comprimento de rotas, número de vias e camadas





# Projeto de antenas especiais

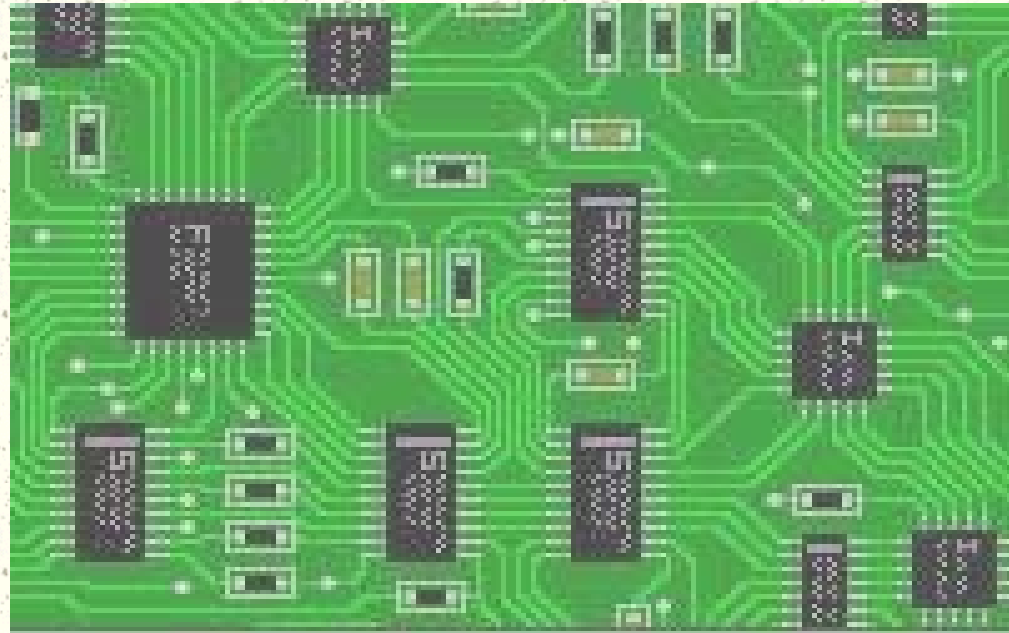
- # Projeto da NASA com *evolvable hardware*
- # A antena funciona melhor do que aquela projetada por engenheiros
- # Ninguém entende completamente por que funciona



<http://ic.arc.nasa.gov/projects/esg/research/antenna.htm>

# Projeto de fiação impressa

- # Semelhante ao projeto de VLSI
  - Demanda processamento intensivo e *software* específico
- # Tem muitos graus de liberdade
  - Várias camadas, vários caminhos entre pinos
  - Busca-se soluções satisfatórias e não ótimas
- # Fatores
  - Estéticos
  - Elétricos



# Planejamento, previsão e despacho de carga

## ✦ Planejamento:

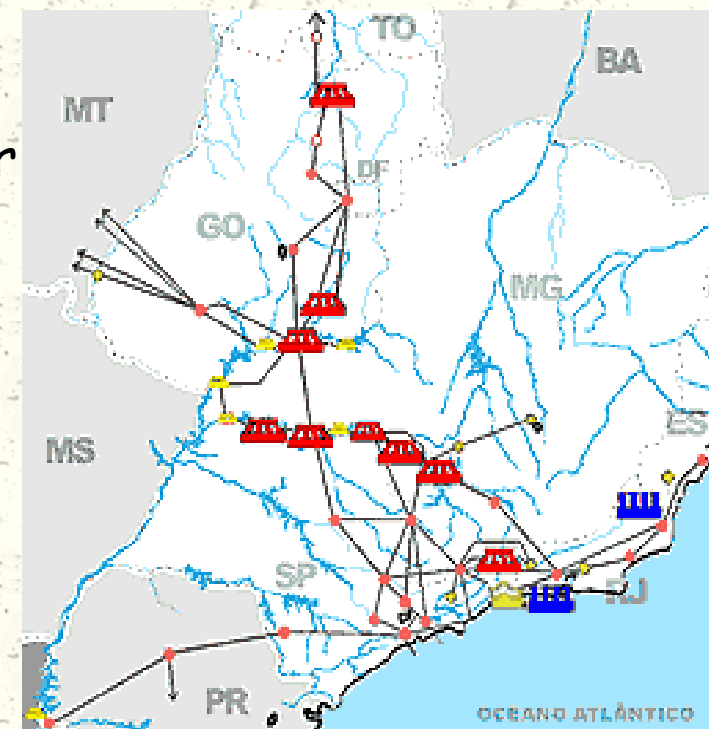
- Estruturar o crescimento da oferta de energia elétrica para o futuro (longo prazo)
- AGs com abordagem determinística

## ✦ Previsão:

- Estimativa antecipada da energia a ser consumida numa região (curto prazo)

## ✦ Despacho de carga:

- Execução da previsão de carga
- Problema dinâmico





# Despacho econômico de carga

- # Problema dinâmico:
  - A demanda flutua com o horário do dia, dia da semana, mês do ano, etc...
- # Limitações:
  - Capacidade instalada de geração e de transmissão
- # Variáveis adicionais:
  - Compra e venda de energia entre concessionárias
  - Variações climáticas inesperadas
  - Manutenção / desligamentos planejados
- # Objetivo final:
  - Satisfazer a demanda com operação segura e perdas mínimas.

# Outras aplicações em sistemas elétricos de potência

- # Comissionamento de turbinas
  - Subproblema do despacho de carga
- # Planejamento da manutenção
  - Manutenção programada de linhas de transmissão e de unidades de geração
- # Localização otimizada de equipamentos
  - Seccionadores, subestações, relés de proteção...
- # Planejamento de redes de distribuição
  - Em baixa e alta tensão
- # Otimização da potência reativa

# Planejamento de trajetórias para manipuladores robóticos

## # Manipuladores

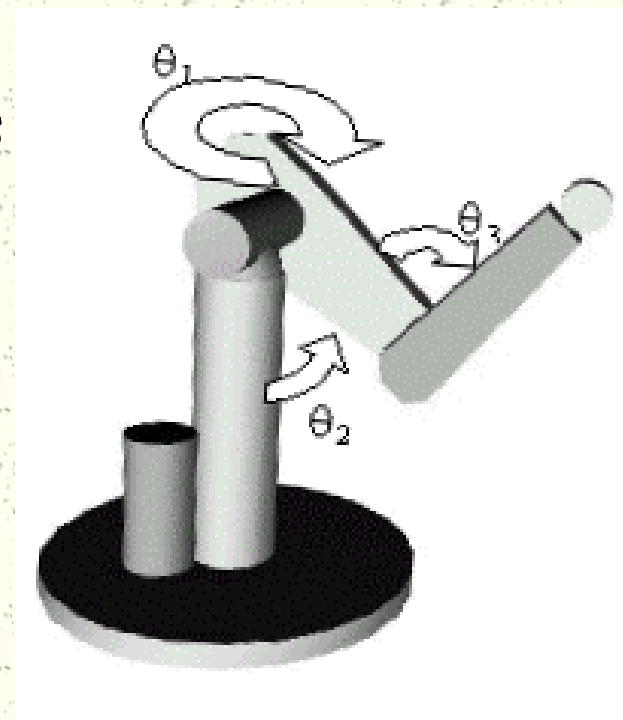
- Redundantes e não-redundantes
- Cinemática direta / inversa

## # Cinemática inversa:

- Dada uma posição e orientação do efetuador, calcular os ângulos das juntas
- Soluções: 0, 1, muitas ou  $\infty$
- Envolve funções não-lineares

## # Para trajetórias:

- AGs otimizam ponto-a-ponto num ambiente dinâmico



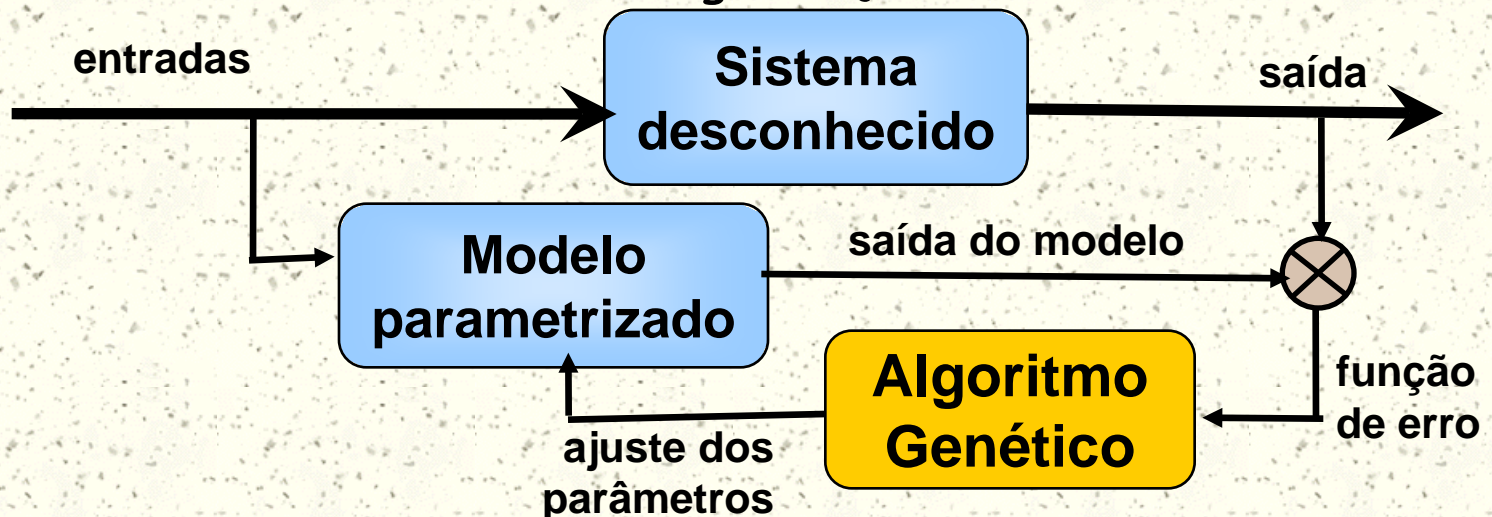


# Identificação de sistemas

## # Identificação:

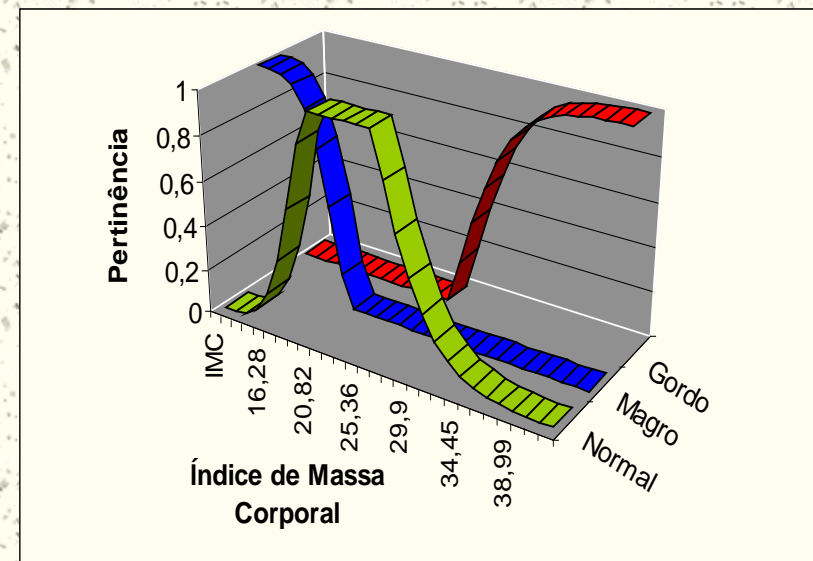
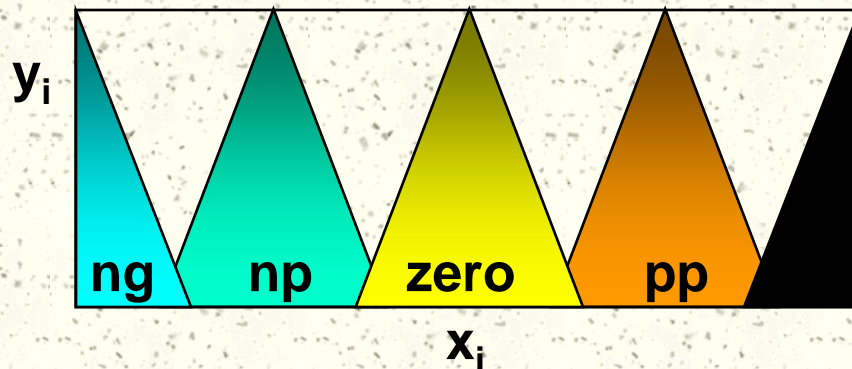
- A partir de um modelo conhecido e de um conjunto de entradas/saídas, identificar os parâmetros do modelo
- Extensa aplicação em engenharia de controle
- Particularmente interessante para sistemas não-lineares e caóticos, contínuos ou discretos

## # Modelo desconhecido: Programação Genética

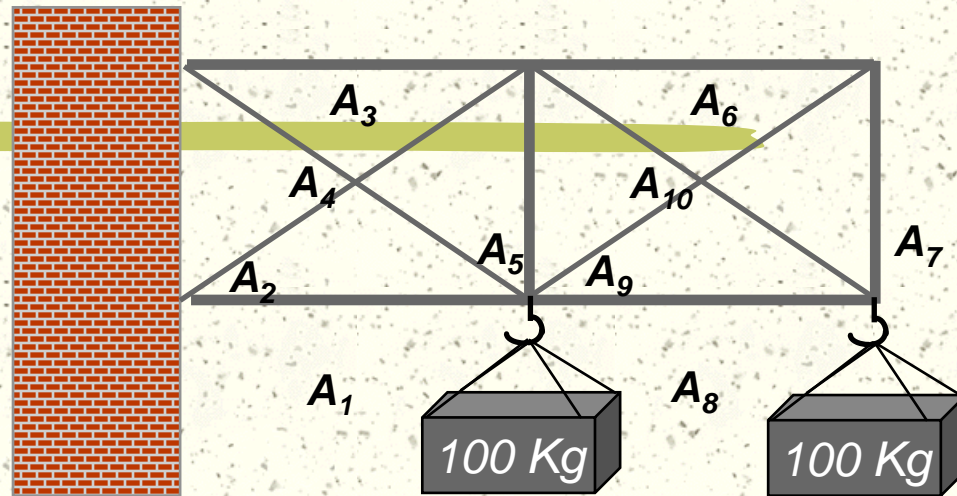


# Otimização de sistemas *fuzzy*

- # Sistemas de controle fuzzy (FLC)
  - Conjunto de variáveis do problema ( $n$ )
  - Conjunto de funções de pertinência ( $m$ )
  - Dimensão da base de regras:  $m^n$
  - Ajuste empírico
- # AGs para otimização da:
  - Forma e número de FPs
  - Base de regras
  - De ambos



# Eng.Civil: otimização estrutural



## # objetivo:

- minimizar o custo do material utilizado, minimizando a seção transversal de cada parte da treliça
- restrição: limites de segurança

## # codificação:

- binário, 4 bits/variável, valores discretos (bitolas comerciais) - espaço de busca:  $2^{40} \approx 10^{12}$

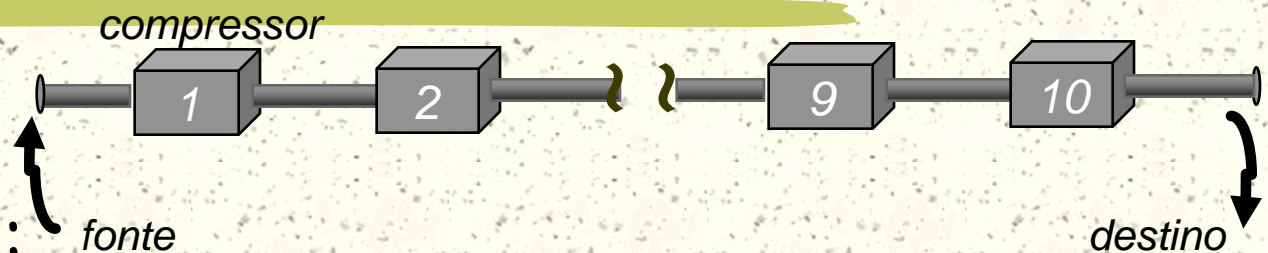
## # resultados:

- pop=200, ger=40 : 8000 avaliações
- desvio de 1% do ótimo conhecido



# Otimização de linha de distribuição de gás

## Goldberg (1983)



### # objetivo:

- minimizar a potência necessária para satisfazer a demanda de pressão de gás nas tubulações
- restrição:  $P_{\max}$  e  $P_{\min}$  por tubulação
- variável controlada: elevação de pressão através de cada compressor

### # codificação:

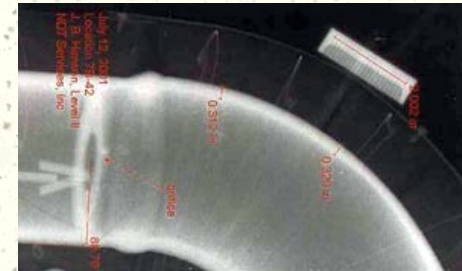
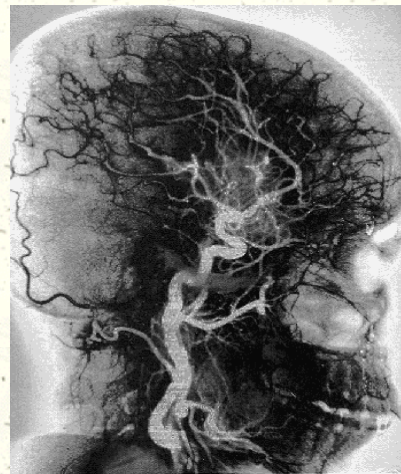
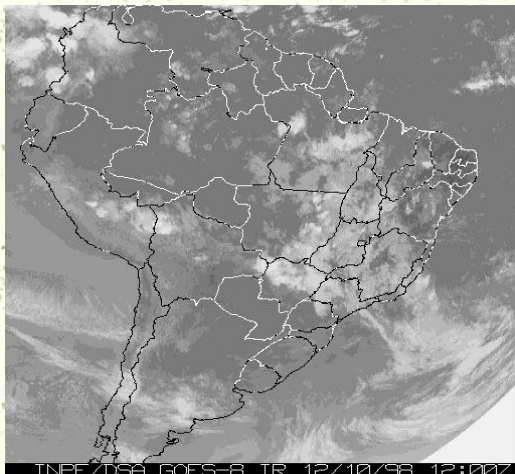
- binário, 4 bits/variável
- espaço de busca:  $2^{40} \approx 10^{12}$

### # resultados:

- pop=50, ger=60 : 3000 avaliações
- erro menor que 2% do ótimo

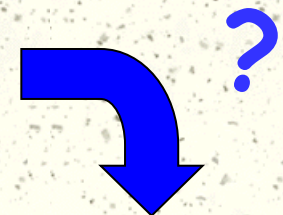
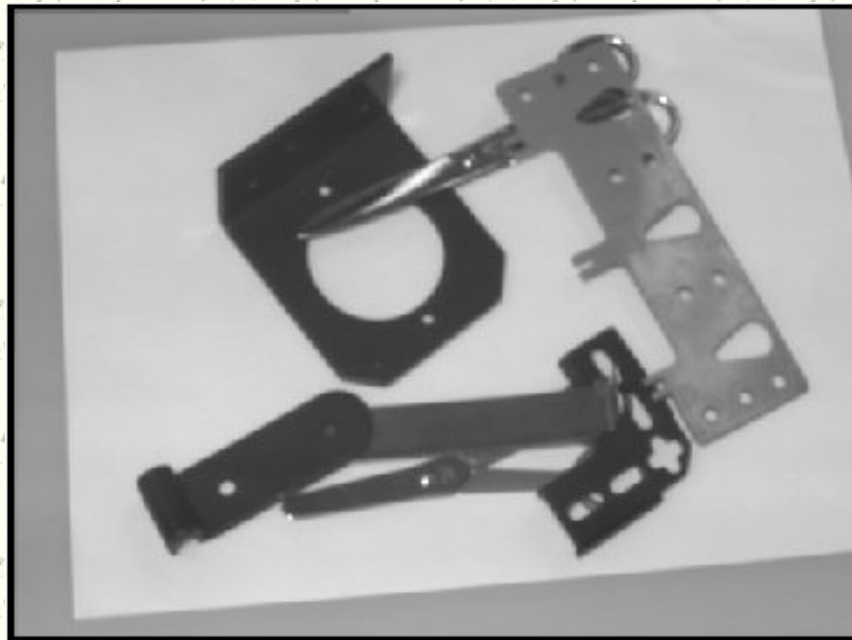
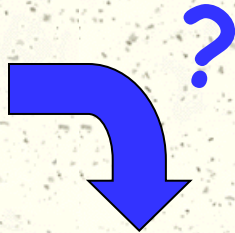
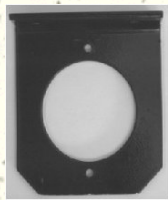
# Registro de imagens

- Encontrar um mapeamento bilinear para transformar as coordenadas de uma imagem alinhando com outra
- Aplicação em angiografia por subtração digital (DAS)
- AG busca os coeficientes da transformação para minimizar a diferença entre imagens
- Aplicação em outros tipos de imagens:
  - imagens de satélites, radiografias, assinaturas, inspeção de tubos



# Identificação de objetos em imagens (*template matching*)

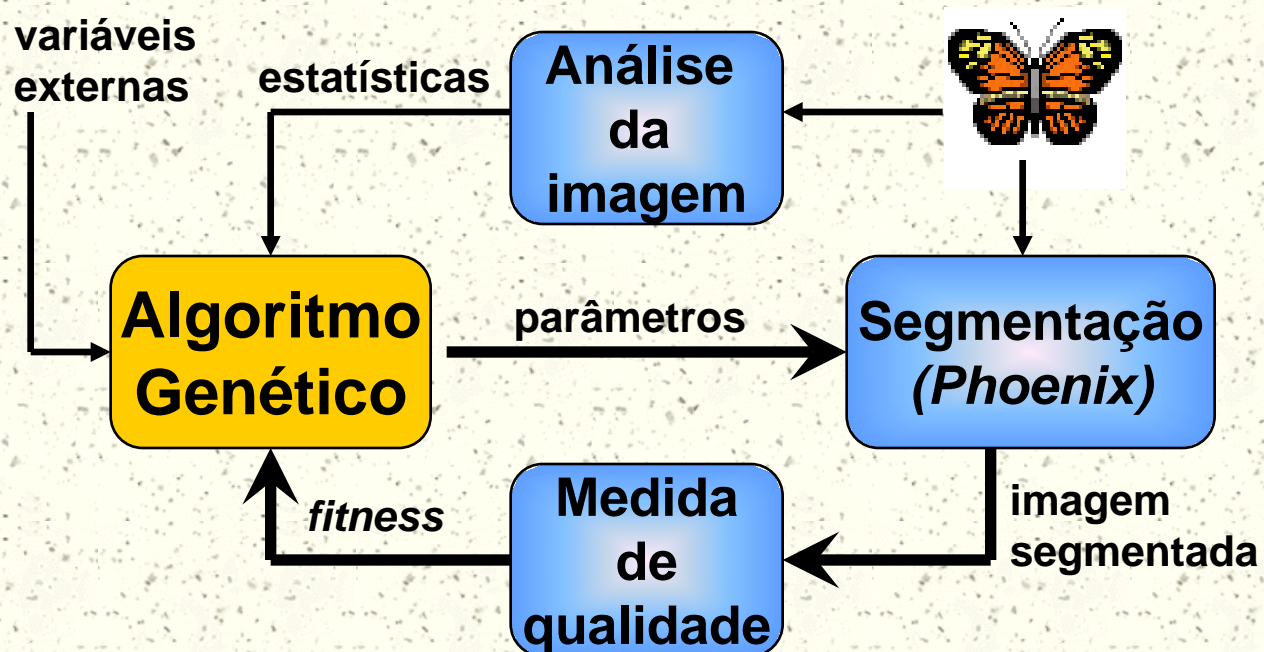
# É uma extensão do registro de imagens





# Segmentação adaptativa de imagens

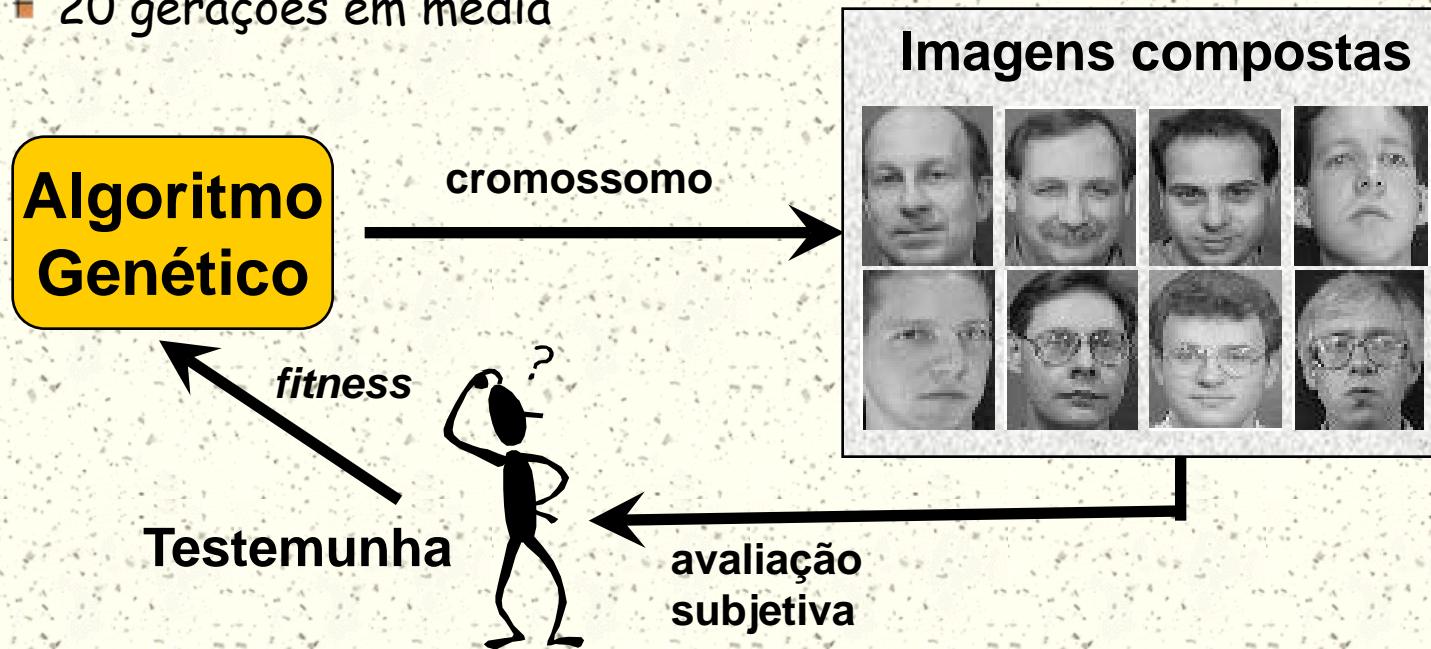
- Aplicação em visão computacional e reconhecimento de imagens em tempo-real
- *Fitness*: função ponderada de 5 medidas tradicionais de qualidade da segmentação:
  - ✦ Coincidência *edge-border*, Consistência limítrofe, Classificação de pixels, Sobreposição de objetos, Contraste de objetos





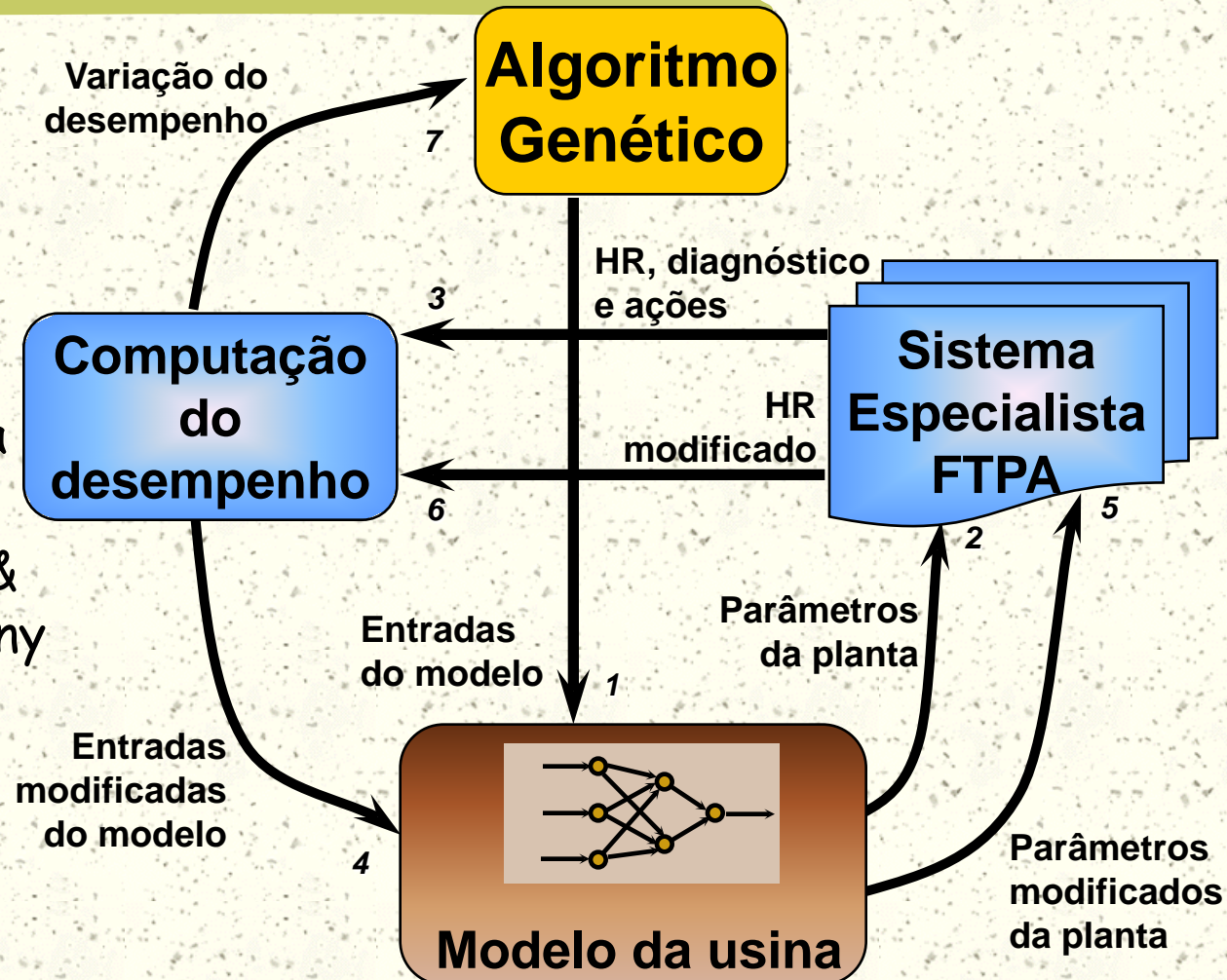
# Identificação interativa de imagens faciais

- Técnicas tradicionais de "retrato falado" (Photofit)
  - 5 grupos de características faciais
  - >15 bilhões de combinações
- GA:
  - Pop. Inicial: 20 imagens aleatórias → Espaço de busca: >34 bilhões
  - *Fitness*: estimativa subjetiva de semelhança por uma testemunha
  - 20 gerações em média



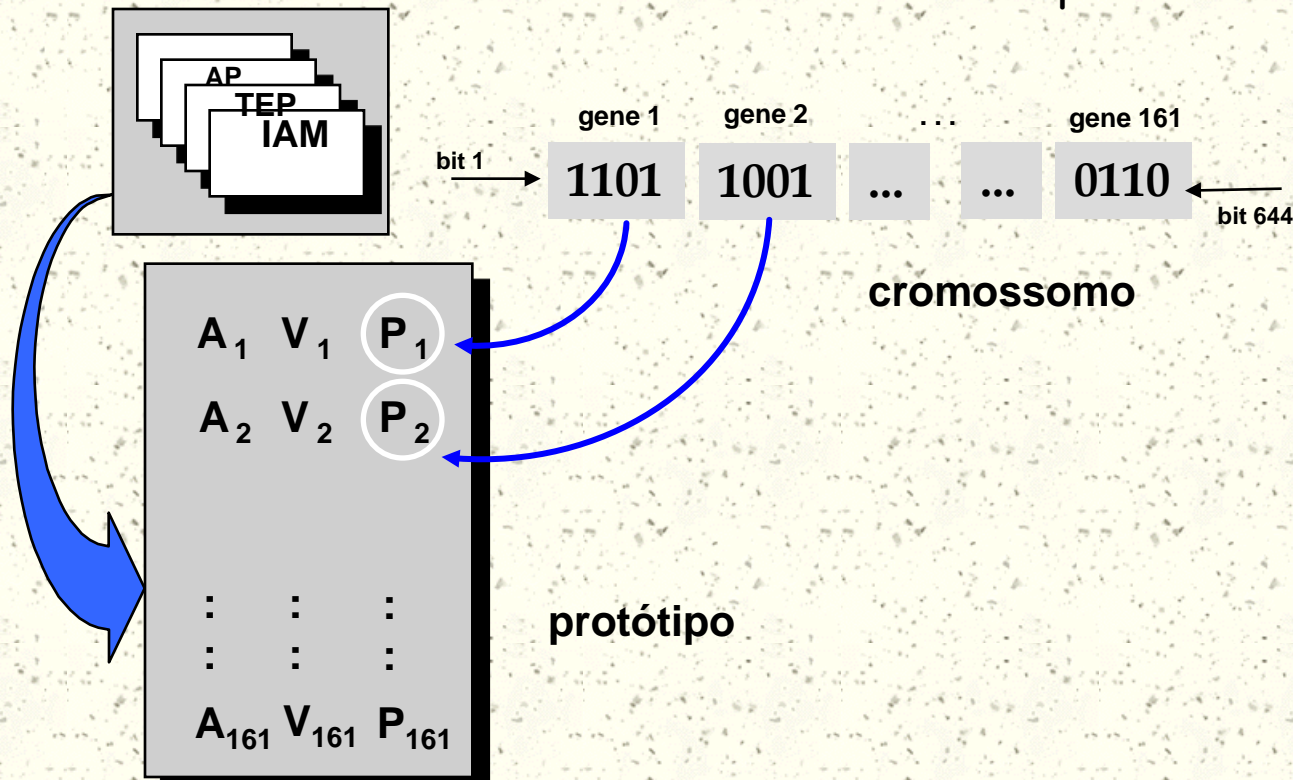
# Validação de Sistemas Especialistas

- ✦ Sistema Especialista FTPA
- ✦ monitora a operação de usina termoeletrica a carvão
- ✦ New York Gas & Electric Company
- ✦ Simula o "advogado do diabo"



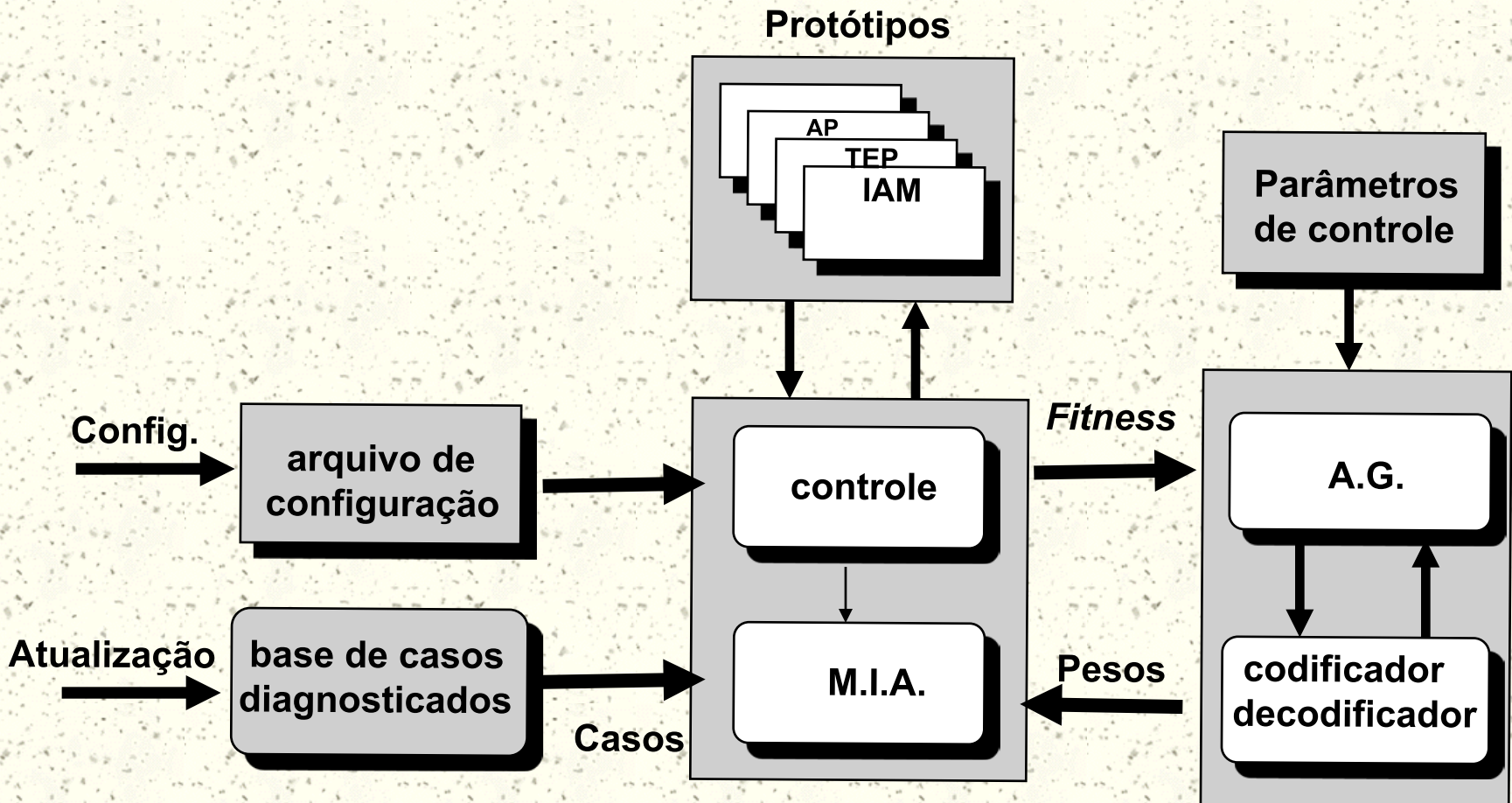
# Otimização de sistema especialista baseado em analogia (DoToR)

- Raciocínio por analogia
  - Analogia superficial com protótipos
  - Aplicação em diagnóstico clínico de dor torácica (12 doenças)
- Dimensionalidade
  - 161 variáveis x 4 bits  $2^{161 \times 4} = 2^{644} \approx 10^{194}$  para cada doença



# Sistema GADoToR

- Otimização dos conjuntos de pesos utilizando AG paralelo



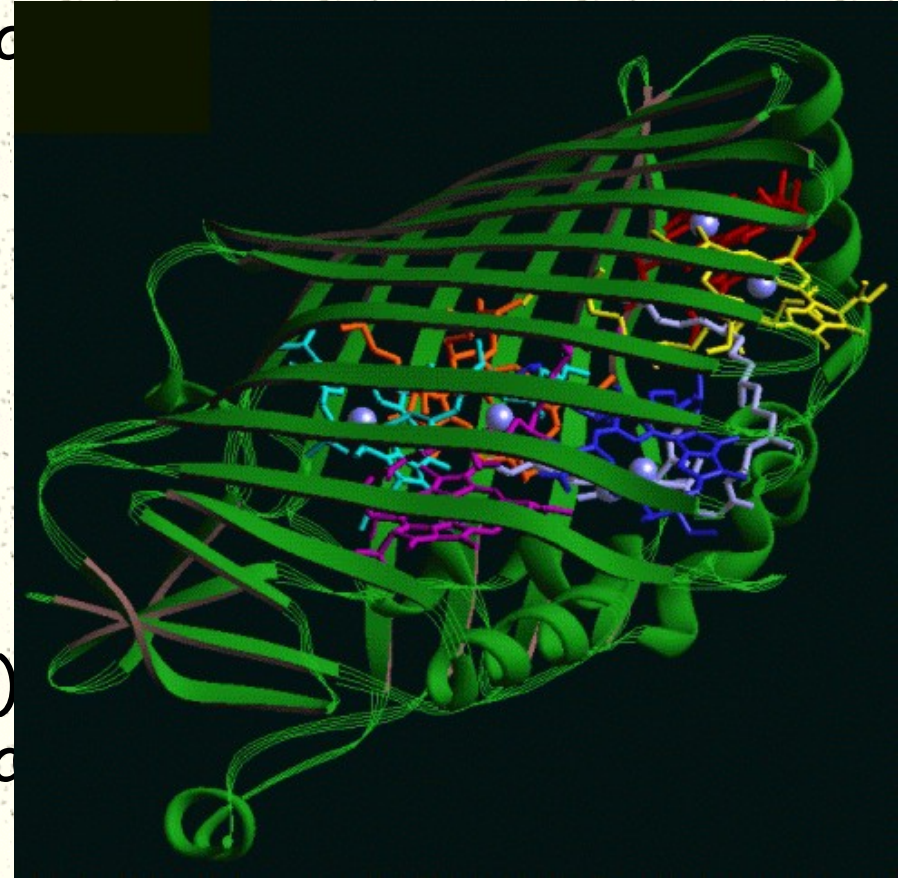


# Aplicações em Bioinformática

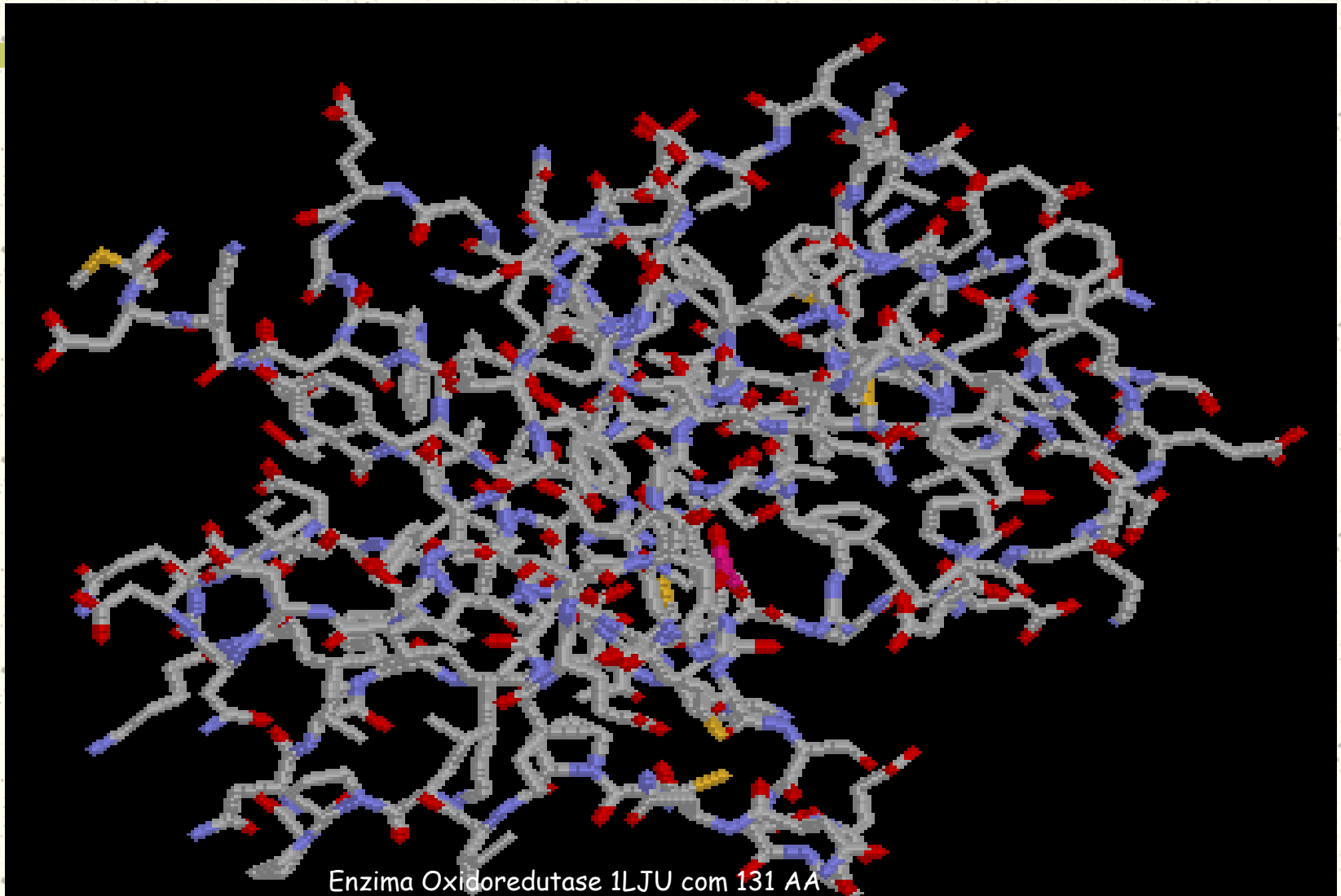
- # Determinação da estrutura de proteínas: estrutura 3D ou com modelos treliça
- # Busca de *motifs* para a identificação de famílias de proteínas
- # Busca de genes e sinais em sequências de DNA
- # Remontagem de fragmentos de DNA
- # Construção de árvores filogenéticas
- # Descoberta de interações entre proteínas
- # Inferência de redes gênicas

# Dobramento de proteínas

- # Proteínas são essenciais para a vida e são formadas por cadeias de aminoácidos
- # Após a síntese, dobram-se numa forma tridimensional única que dá a sua funcionalidade
- # O problema da predição da estrutura de proteínas (PSP) é de suma importância para a Biologia/Bioinformática

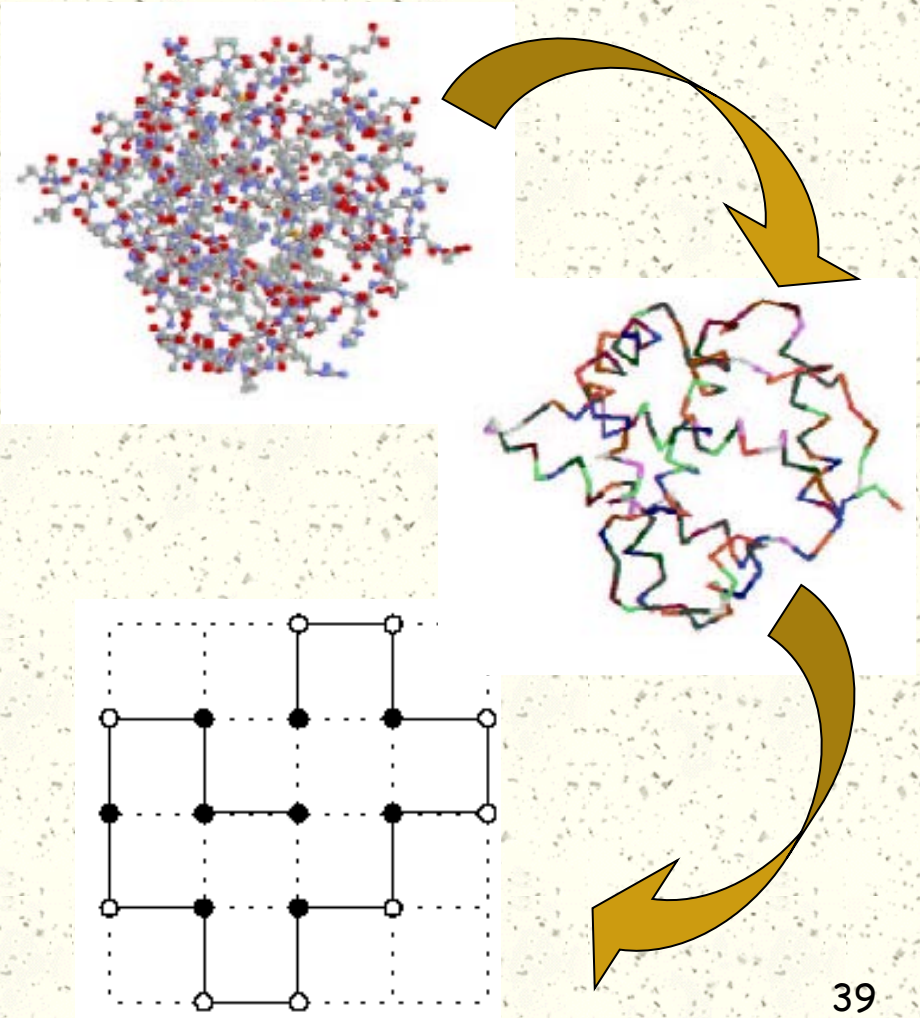


# Estrutura 3D de uma proteína



# Dobramento de proteínas

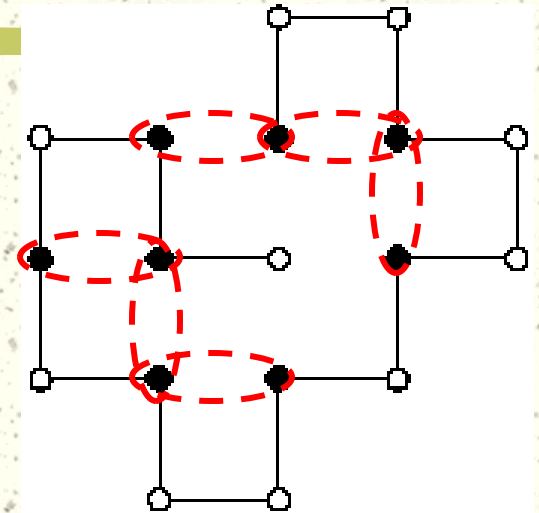
- # A busca exaustiva do espaço conformacional é impossível
- # Utiliza-se modelos ultra-simplificados de treliças 2D e 3D
- # O modelo 2D-HP é o mais usual, mas a sua complexidade é **NP-completo**



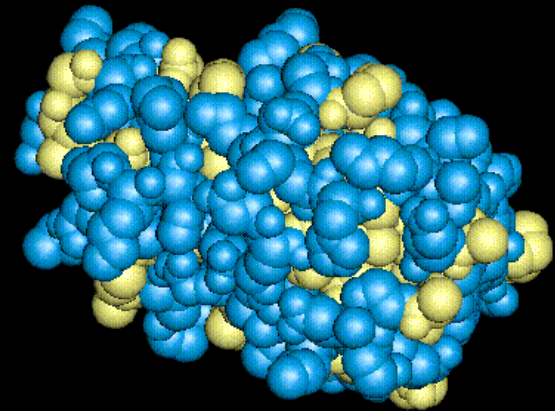


# Dobramento de proteínas

- ✦ A conformação nativa é a de mínima energia:
- ✦ A energia livre é inversamente proporcional ao número de ligações hidrofóbicas não-locais (contatos H-H)
  - Resíduos Hidrofóbicos (**H**) formam o núcleo da molécula
  - Resíduos Hidrofílicos (**P**) formam a interface com o ambiente



● *Hydrophobic Residues*  
● *Hydrophilic Residues*

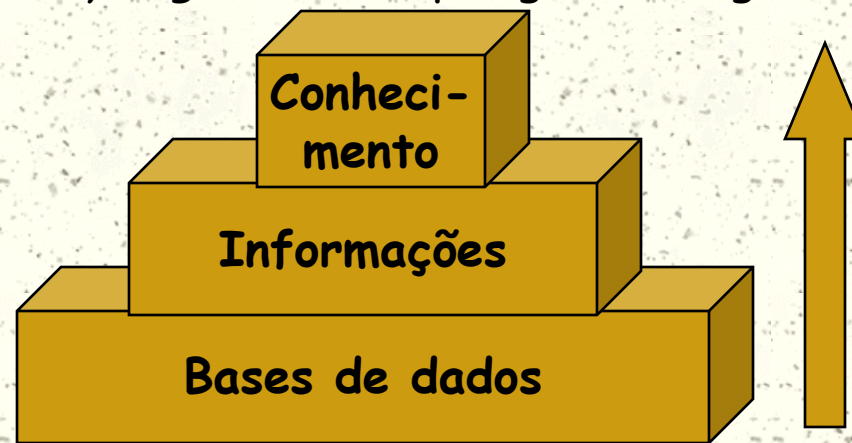


# Mineração de dados - *Data Mining*

- *Data Mining* - Definição
  - processo de busca de conhecimento correto, útil, compreensível e não-trivial em grandes bancos de dados
- Utiliza muitos métodos
  - Indução, Métodos estatísticos, Redes Neurais, Sistemas Fuzzy, Sistemas baseados em regras, Raciocínio baseado em casos, Computação Evolucionária
- Tarefas mais comuns:
  - Classificação, regressão, *clustering*, sumarização, dependência, análise de relacionamento entre atributos, análise de sequência.
- Multidisciplinariedade:
  - IA / IC; Aprendizado de máquina; Estatística; Bancos de Dados
- Principais aplicações de AG e PG:
  - Classificação
  - Seleção de atributos

# Aplicações de CE em *Data Mining*

- Predição: classificação e previsão de séries temporais
- Descoberta de regras (*fuzzy* ou não) para diagnóstico clínico
- Seleção de atributos relevantes e ponderação
- Segmentação de dados / agrupamentos
- Regras de associação
- Sumarização & visualização
- Ver: A.A.Freitas - *Data Mining and Knowledge Discovery with Evolutionary Algorithms*. Springer-Verlag, 2002.



# Sistemas Classificadores em Diagnóstico Médico

## ✦ Sistemas Classificadores

- ✦ Sistema baseado em regras que evolui através de AGs
  - ✦ Utiliza banco de casos em vez de um especialista para obter as regras
  - ✦ Cada indivíduo é uma regra (Michigan) ou um conjunto de regras (Pittsburgh) que objetivam classificar uma dada informação
- ## ✦ Inúmeras aplicações em diagnóstico e tomada de decisão clínica:
- ✦ Pneumologia, Oncologia, etc.



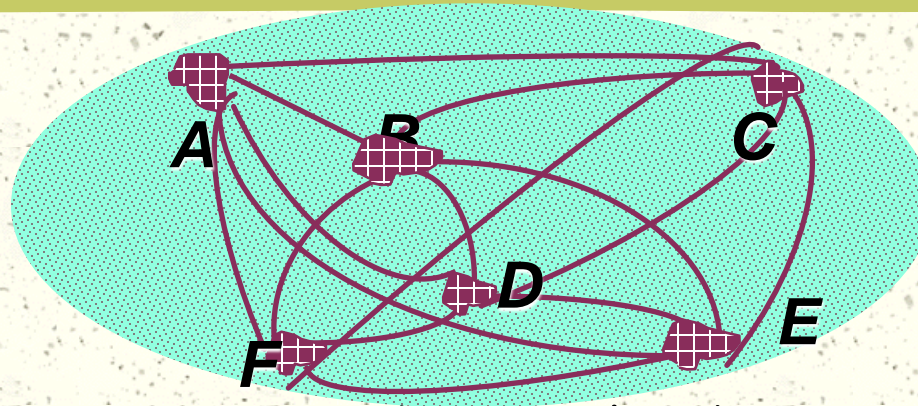
# Algumas classes de aplicações em pesquisa operacional

- # Problema do empacotamento 1D, 2D, 3D...
- # Balanceamento de linhas de produção (*job shop/ flow shop*).
- # Alocação de recursos/pessoas (*timetable/crew problems*).
- # Problemas de percurso mínimo e de transporte (logística).
- # Seleção de *portfolios*.
- # Problemas de programação linear, inteira, estocástica...
- # Problemas de locação de facilidades

# Problema da "Mochila"

- Também conhecido como problema de empacotamento.
  - Dados  $n$  objetos que podem ser colocados num recipiente (mochila, caixa, container...), onde cada um tem uma **limitação dimensional** (peso, volume, área...) e uma **utilidade** (valor, lucro...); deve-se escolher quais serão colocados no recipiente, dentro de suas limitações (volume, área, capacidade...) de modo que o somatório das utilidades seja máximo.
- Dimensionalidade:
  - 1D: corte de barras metálicas / canos, etc
  - 2D: corte de chapas, embalagens, tecidos, etc
  - 3D: carregamento de container, depósito, etc

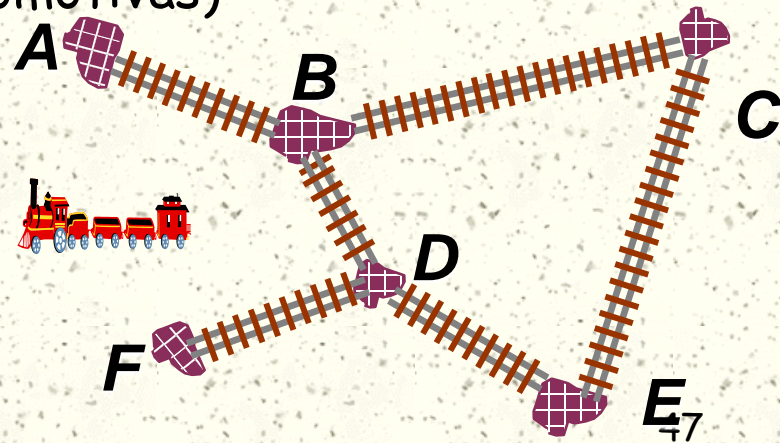
# Problema do caixeiro viajante (TSP)



- Dados um conjunto de locais ligados por caminhos, onde cada um tem um custo (distância, tempo...), achar a rota que liga todos os locais, passando uma única vez por cada um, de modo que o custo total seja mínimo.
- Restrições: otimizar o "custo" associado a cada trecho, pontos de "reabastecimento", tempos mínimos
- Detalhes:
  - ✦ representação literal, binária, por percurso, por borda
  - ✦ operadores especiais (*crossover*!)
  - ✦ muitas variações

# Problemas de transporte generalizado

- Multiplicidade de origens, destinos, rotas, tráfego, etc... leva a problemas multiobjetivos (MO)
- Distribuição e transporte de produtos e pessoas
- *Scheduling* de frota e pessoal
- Exemplo: transporte ferroviário
  - min. a distância percorrida
  - min. tempos de manobra e espera
  - min. recursos físicos (vagões, locomotivas)
  - min. tempo de entrega e trânsito
  - max. num. de cidades servidas





# Alocação de pessoal (*Crew scheduling*)

## # Problema:

- Escala de plantão médico mensal para a maternidade do HU/UFSC (24 horas por dia) com 12 médicos e turnos de tamanho variável

## # Restrições:

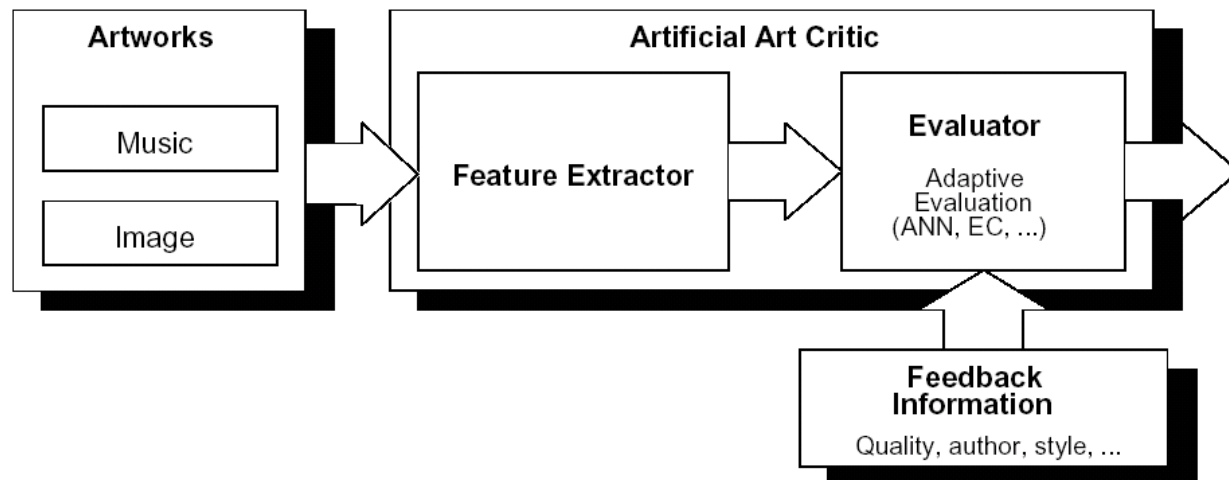
- Restrições gerais: carga horária, turno, feriados, finais de semana, horas consecutivas.
- Restrições individuais: períodos, número de horas, horários fixos, carga total. Exemplo: Médico "Fulano de Tal":
  - # Máximo de 32 h em finais de semana
  - # Máximo de 12 h de plantão noturno
  - # Máximo de 20 h à tarde
  - # Não pode trabalhar de manhã

## # Resultados satisfatórios:

- Fácil adaptação para as variações mensais
- Independe de fatores humanos

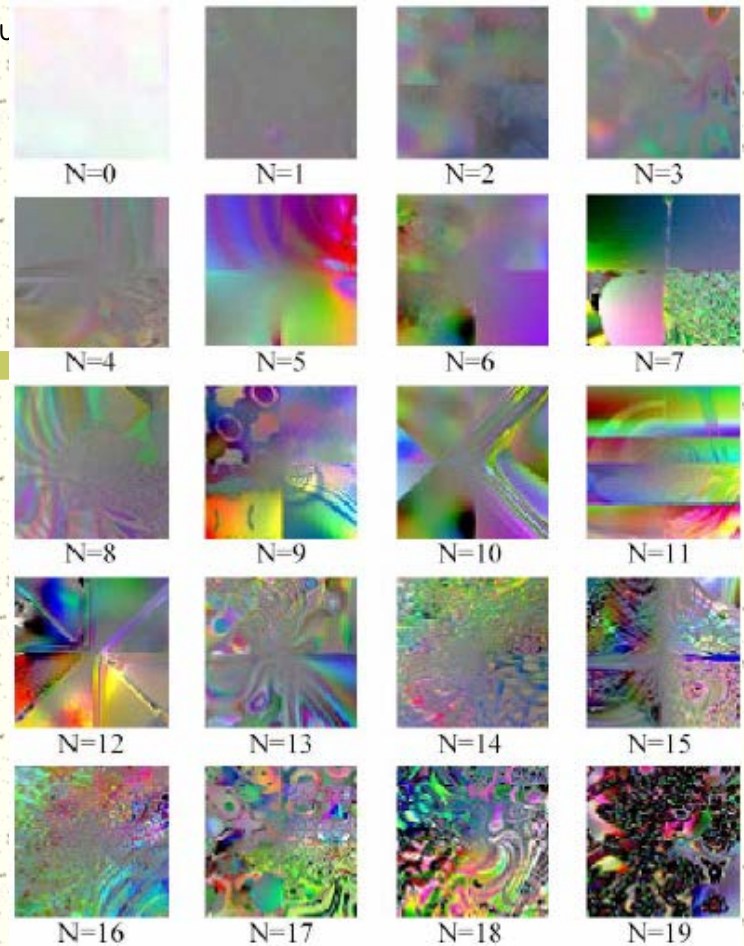
# Arte evolutiva

- ✦ Inspiração para a criatividade computacional:
  - Criatividade humana
  - Natureza: a evolução é criativa
- ✦ Possibilidades:
  - O usuário gera a idéia e o computador evolui
  - A idéia surge da iteração entre usuário e computador
  - O computador gera a idéia e o usuário é passivo
- ✦ Abordagem mais popular para música e imagem: criatividade assistida por computador: o computador gera idéias e o usuário as avalia





# Arte evolutiva



- # PG para gerar imagens a partir de primitivas
- # AG acoplado a RN para gerar música

## # Alguns links:

- # <http://www.biota.org/ksims>
- # <http://helios.hampshire.edu/l spectator>
- # <http://graphics.stanford.edu/~bjohanso/gp-music>
- # <http://www.dei.uc.pt/~machado/NEvAr>





# Arte evolutiva

