17

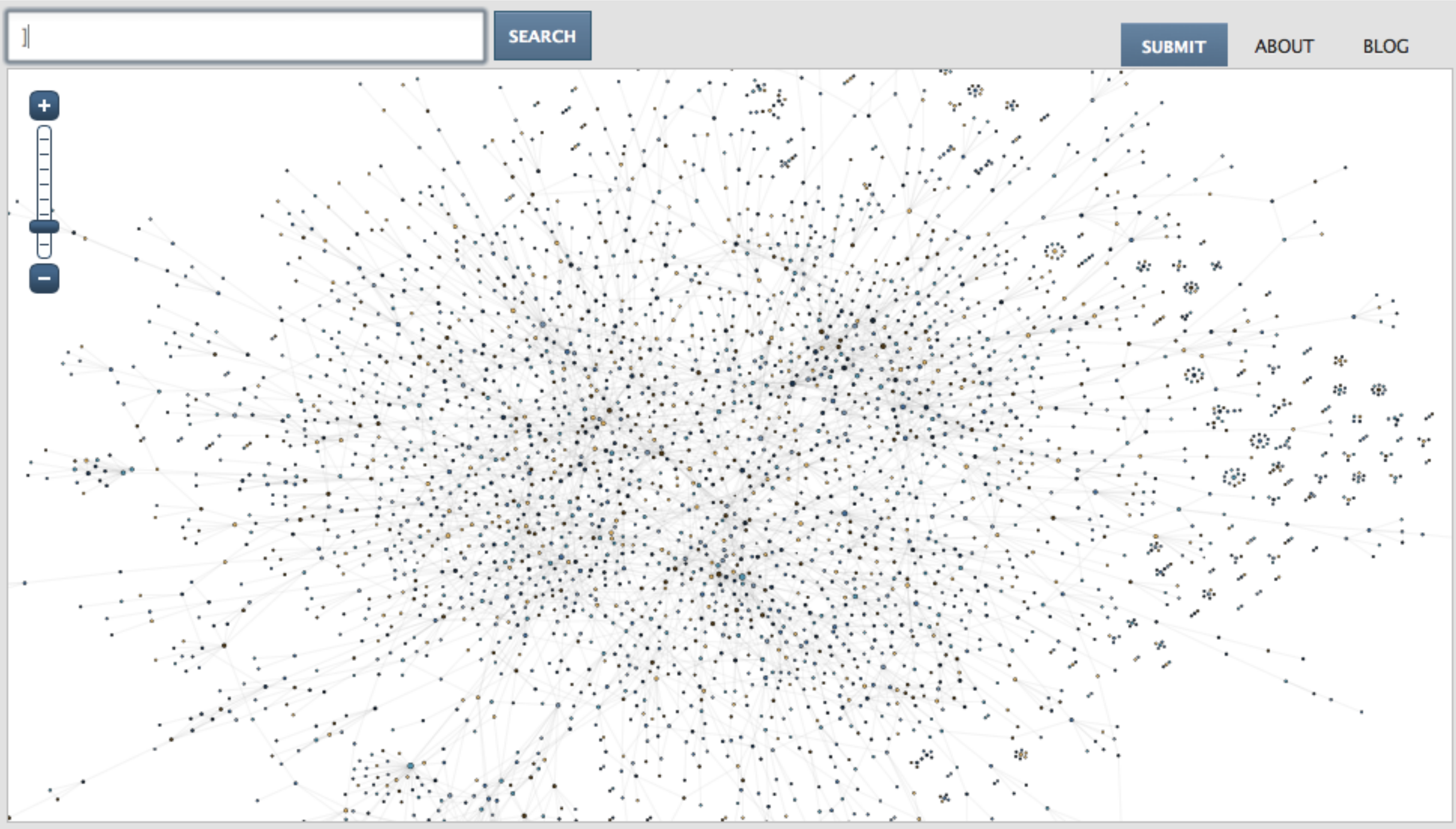
VISUALIZAÇÃO DE ESTRUTURAS RELACIONAIS

Profa. Raquel C. de Melo Minardi

Um modelo gerado por computador produzido com o supercomputador IBM Blue Gene, parte do projeto Blue Gene, mostra as 30.000.000 conexões entre 10.000 neurônios em uma única coluna neocortical - a parte mais complexa do cérebro de um mamífero. As cores diferentes indicam níveis de atividade elétrica

VISUALIZAÇÃO DE ESTRUTURAS RELACIONAIS

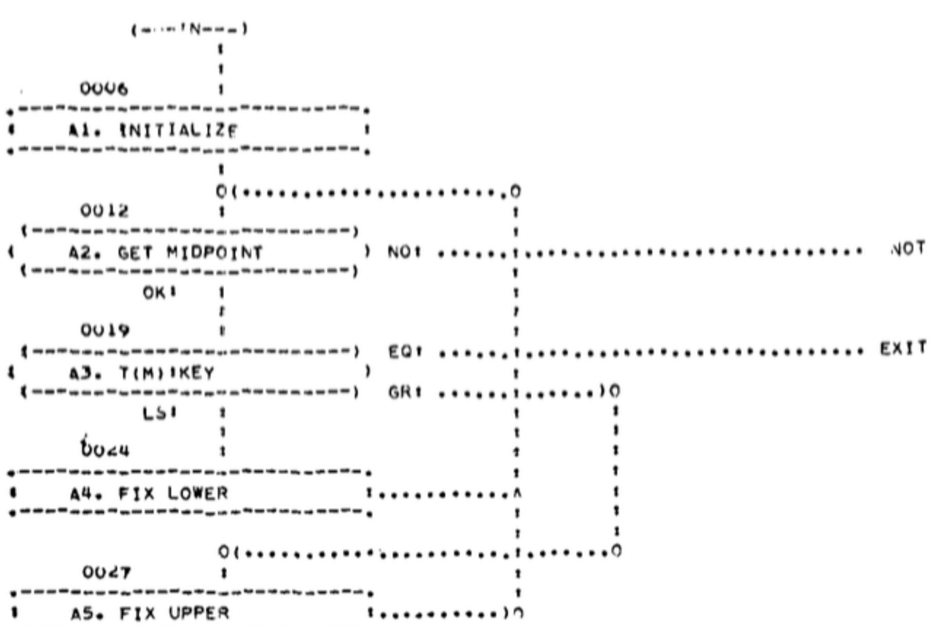
- Só são úteis na medida em que os diagramas associados transmitem informações efetivamente às pessoas que as utilizam


SEARCH
SUBMIT
ABOUT
BLOG

PRIMEIRO ALGORITMO PARA DESENHO DE GRAFOS

```
0000                                                       FL.O                  A.  SERCH.
0001                                          T0001        9I.R   1000    1999
0002                                          LOWER        EQU    B01A
0003                                          UPPER        EQU    B02A
0004                                          KEY          EQU    B03A
0 105                                                      HHH         H
0006  4006  BBB 0 60  B03A  4010              SERCH        STA    KEY                A1. INITIALIZE
0007  4010  B9B 0 25  4012  4014                           LDA#   00000   10000
0008  4014  B9B 0 60  B01A  4018                           STA    LOWER
0009  4018  B9B 0 37  0300  4024                           SHL     0300   1F
0010  4024  B9B 0 60  B02A  4028              1            STA    UPPER
0011  4028  B9B 0 70  B01A  4033                           ADD    LOWER   3F
0012  4033  B9B 0 77  4033  4036              3            ATL                       A2. GET MIDPOINT
0013  4036  B9B 0 85  4039  4015                           MUL#   00000   000A5
0014  4015  B9B 0 05  000A  4019                           LDX    RA
0015  4019  B9B 0 30  B01A  4023                           LDL    LOWER
0016  4023  B9B 0 87  4026  4226                           TGR    2F
0017  4226  B9B 0 82  4026  4029                           TEQ    2F      NOT
0018  4026  B9B 0 70  4229  000A              2            ADD            RA
0019  4228  B9B 0 25  0999  4001                           LDA    T0000                A3. T(M):KEY
0020  4001  B9B 0 30  B03A  4005                           LDL    KEY
0021  4005  BBB 1 82  0000  4009                           TEQ2    0000
0022  4009  B9B 0 87  4212  4412                           TGR    2F
0023  4412  B9B 0 25  000C  4016                           LDA    RX
0024  4016  B9B 0 70  4219  4021                           ADD#   00000   10000        A4. FIX LOWER
0025  4021  B9B 0 60  B01A  4025                           STA    LOWER
0026  4025  B9B 0 70  B02A  4033                           ADD    UPPER   3B
0027  4212  B9B 0 25  000C  4216              2            LDA    RX                   A5. FIX UPPER
0028  4216  B9B 0 75  4419  4024                           SUB            1B
0029  4418  B9B 0 00  0001  0000                           CON    00000   10000
0030
00..                                                       HHH
```

Fig. 2. Assembly language corresponding to Figure 1




```
*    A.   SERCH.
*         THIS SUBROUTINE SEARCHES THROUGH TABLE T
*         TO SEE IF IT CAN FIND AN ENTRY MATCHING
*         A GIVEN KEY.
*    A1.  INITIALIZE
*         START OUT BY SETTING 'LOWER' TO 1,
*         'UPPER'TO 1000.
*         THE TABLE IS T0001 THROUGH T1000 AND IS IN
*         ASCENDING SEQUENCE.
*    A2.  GET MIDPOINT
*         SET 'M' TO (LOWER+UPPER)/2.  'M' WILL THUS
*         APPROXIMATE THE MIDPOINT OF THE INTERVAL
*         WHERE WE HAVE PINPOINTED THE SEARCH.
*         IF 'UPPER' IS LESS THAN 'LOWER', THE KEY
*         IS NOT IN THE TABLE.
*    A3.  T(M):KEY
*         COMPARE T(M) WITH THE SEARCH KEY.
*         IF EQUAL, WE EXIT.
*         IF GREATER, TO A5.
*    A4.  FIX LOWER
*         SET 'LOWER' TO M+1,AS T(M) IS TOO SMALL.
*         TO A2.
*    A5.  FIX UPPER
*         SET 'UPPER' TO M-1, AS T(M) IS TOO BIG.
*         TO A2.
*    CODING DETAILS:  AT ENTRY RB2 CONTAINS THE EXIT
*         LOCATION AND RA CONTAINS THE KEYWORD.
*         IF FOUND, THE PLACE FOUND IS IN RX.
*         IF NOT IN TABLE, EXIT OCCURS TO LOCATION'NOT'
*
*
*
```

Fig. 1. Flow chart and flow outline for binary search

- Knuth’s 1963 paper on drawing flowcharts was perhaps the first paper to present an algorithm for drawing a graph for visualization purposes

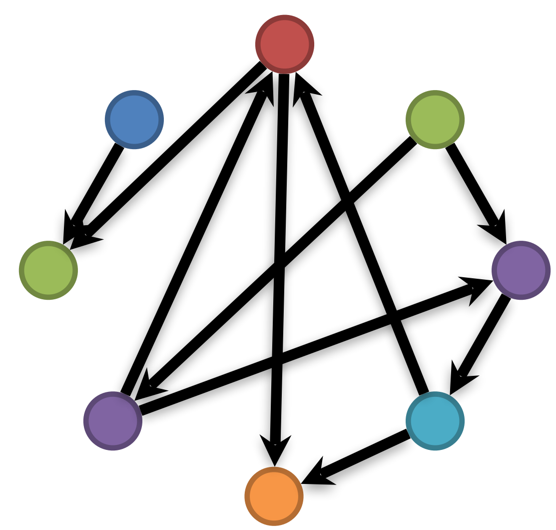
Knuth, Donald E. "Computer-drawn flowcharts." *Communications of the ACM* 6.9 (1963): 555-563.

O PROBLEMA DO DESENHO DE GRAFOS

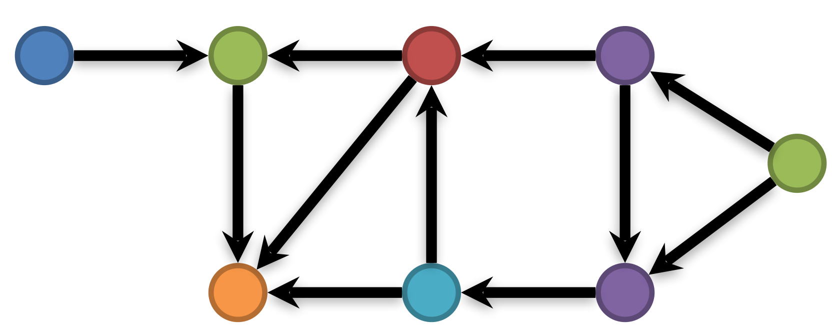
- Um grafo $G=(V,E)$ é um conjunto V de vértices conectados por um conjunto E de arestas
  - Grafos são desenhados como pontos espalhados pelo espaço e conectados por segmentos retos ou curvos representando as arestas

DESENHO DE GRAFOS

- Um diagrama pobre pode ser confuso e enganador



- Um bom diagrama ajuda o leitor a entender o sistema



CRITÉRIOS ESTÉTICOS

- Especificam as propriedades do desenho que gostaríamos de aplicar, na medida do possível, para alcançar a legibilidade

CRUZAMENTOS

- Minimizar o número de cruzamentos entre arestas

ÁREA

- Minimização da área do desenho.

- A área pode ser formalmente definida de diferentes maneiras:
  - a área do menor polígono que cobre o desenho
  - a área do menor retângulo com lados horizontais e verticais cobrindo o desenho.

ARESTAS

- Comprimento total das arestas:
  - minimização da soma dos comprimentos das arestas.

- Comprimento máximo das arestas:
  - minimização do comprimento máximo de uma aresta.

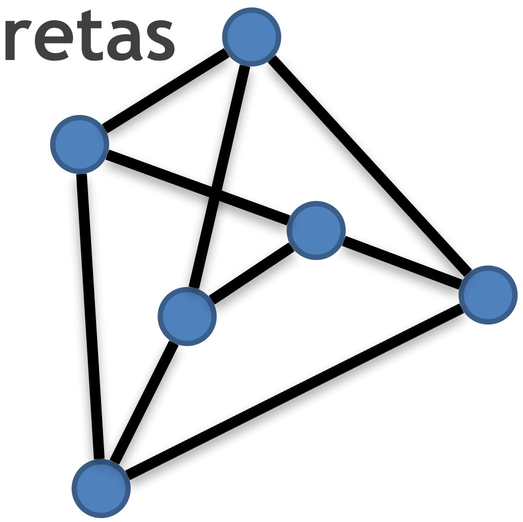
- Comprimento de arestas uniforme:
  - minimização da variância dos comprimentos das arestas.

DOBRAS DE ARESTAS

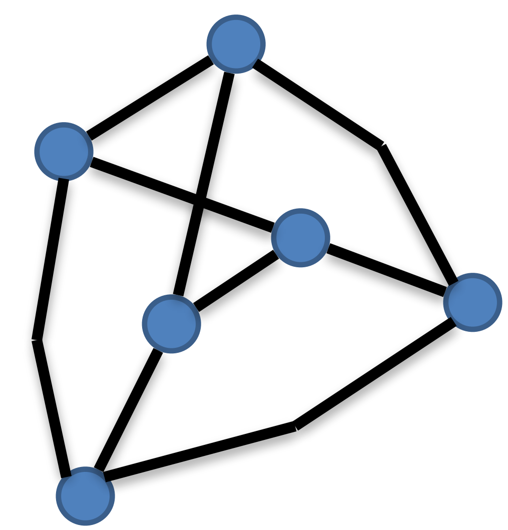
- Dobras de arestas:
  - minimização do número total de dobras nas arestas

- Este critério é especialmente importante para desenhos ortogonais, enquanto é trivialmente satisfeito por desenhos em linha reta.

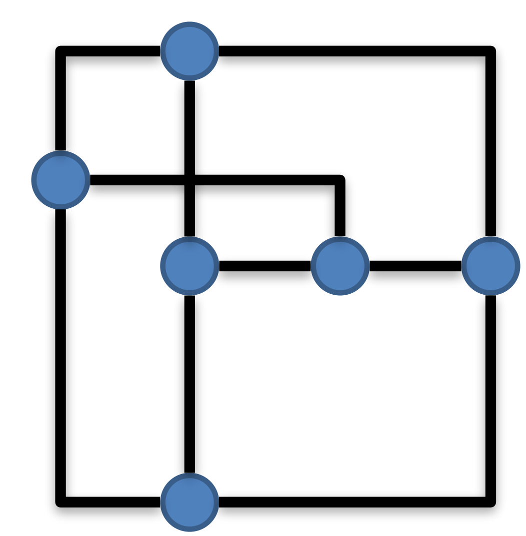
Linhas retas



Poli-linhas



Ortogonal

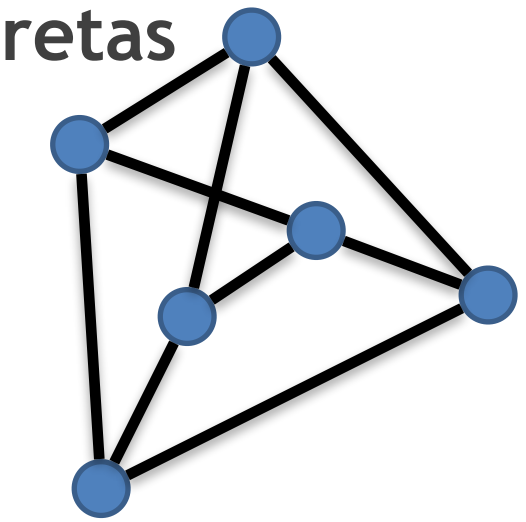


DOBRAS DE ARESTAS

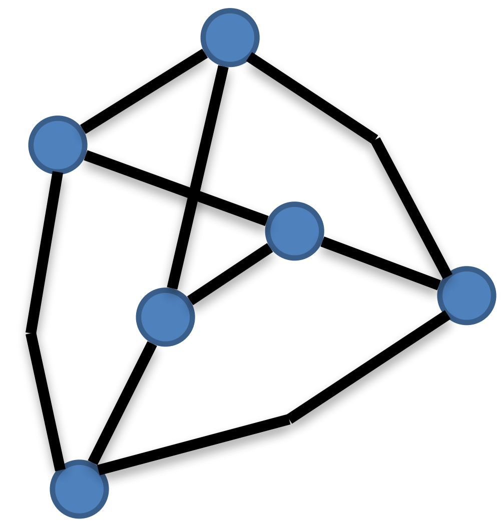
- Número máximo de dobras:
  - minimização do número de dobras em uma aresta

- Dobras uniformes:
  - minimização da variância do número de dobras nas arestas

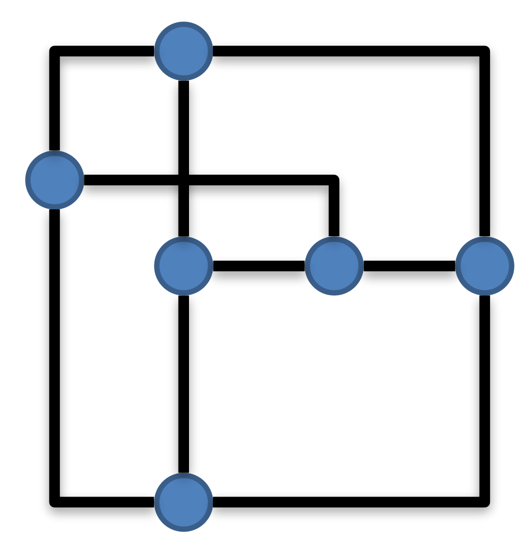
Linhas retas



Poli-linhas



Ortogonal

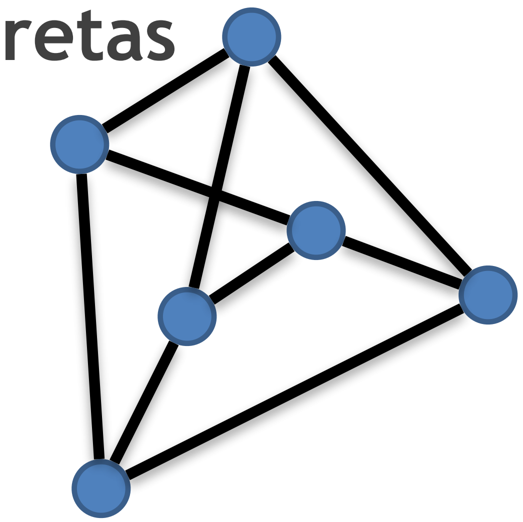


RESOLUÇÃO ANGULAR

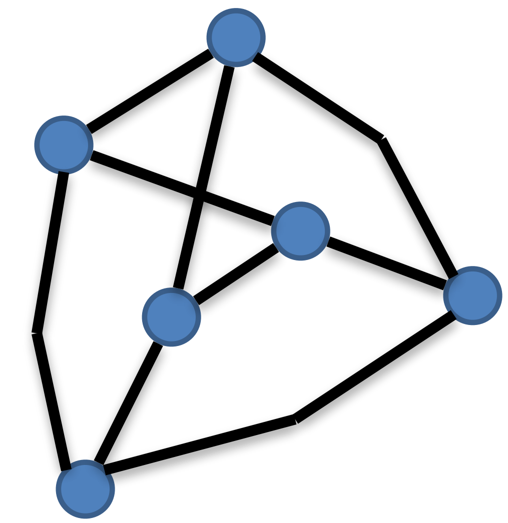
- Ângulo entre duas arestas que incidem no mesmo vértice.

- Esta estética é especialmente relevante para desenhos em linha reta.

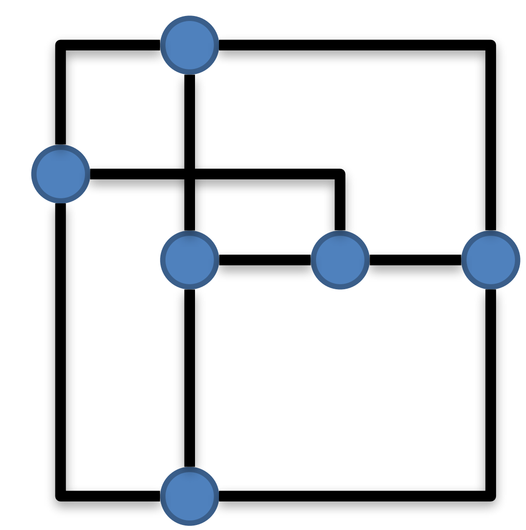
Linhas retas



Poli-linhas

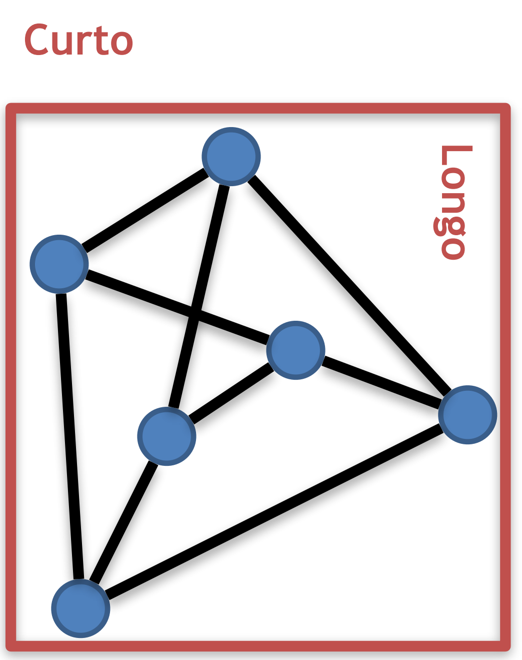


Ortogonal



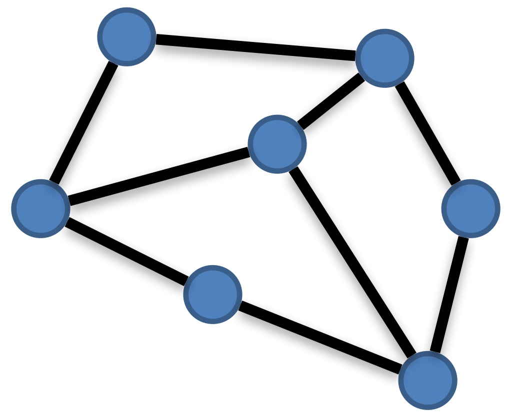
RAZÃO DO ASPECTO

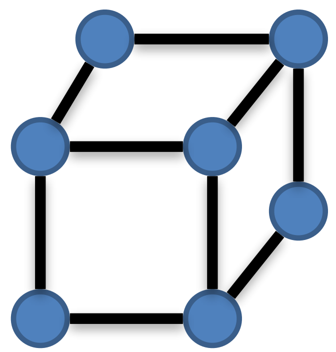
- Minimização da razão do aspecto do desenho
  - Definida como a relação entre o comprimento do lado mais longo e o comprimento do lado mais curto do menor retângulo que envolve o desenho



SIMETRIA

- Ilustra a simetria topológica do grafo



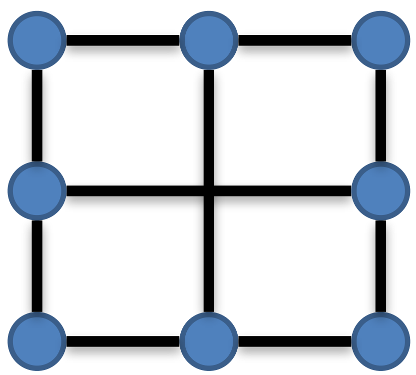


CRITÉRIOS ESTÉTICOS

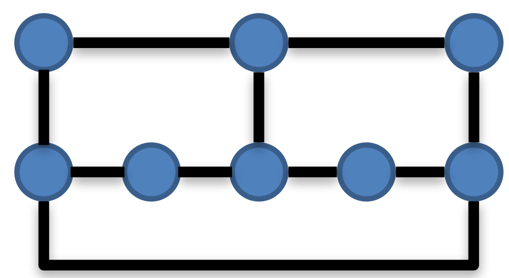
- Estão naturalmente associados a problemas de otimização
  - Computacionalmente difíceis
  - Estratégias de aproximação e heurísticas

CRITÉRIOS ESTÉTICOS

- Muitas vezes conflitam uns com os outros



(a)



(b)

Dois desenhos de grafos ortogonais:

(a) com o número mínimo de dobras;

(b) (b) com o número mínimo de cruzamentos

- Mesmo não sejam conflitantes, muitas vezes é difícil algoritmicamente lidar com todos eles ao mesmo tempo

PRECEDÊNCIA ENTRE CRITÉRIOS ESTÉTICOS

- A maioria das metodologias de desenho de grafos estabelecem uma relação de precedência entre estética

- As abordagens geralmente dividem o processo em uma seqüência de passos, cada um destinado a satisfazer um certo critério

AN ALGORITHM FOR DRAWING GENERAL UNDIRECTED GRAPHS

T. Kamada e S. Kawai
Information Processing Letters
1989

LEI DE HOOKE

- Elasticidade dos corpos
- Calcular a deformação sofrida por um corpo elástico quando aplicada uma determinada força
  - $F = - k \Delta x$

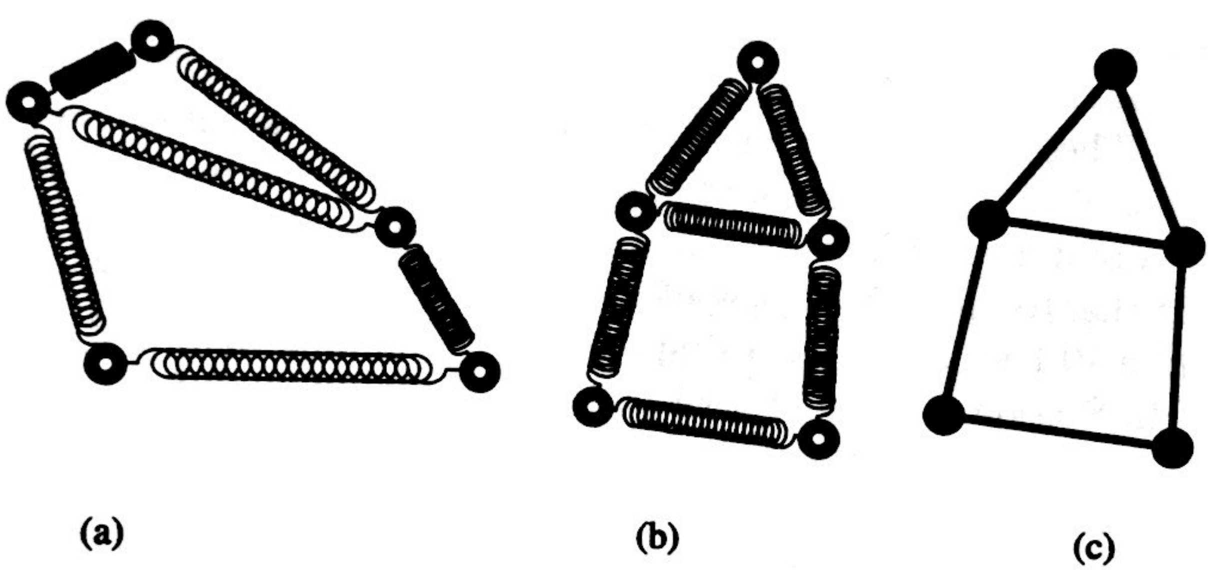


Figure 10.1: A spring algorithm.

- **Metáfora**: cada grafo de um nó é representado por um anel e as arestas são substituídas por molas formando um sistema mecânico
- Os vértices são colocados em uma posição inicial e as molas se movimentarão levando o sistema ao mínimo de energia (potencial elástica)
  - $E = \frac{1}{2}k \Delta x^2$

KAMADA-KAWAI

- Sistema de $n = |V|$ partículas (p_i) conectadas mutuamente por molas cuja energia total é dada por

$$E = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \frac{1}{2} k_{ij} \left(|p_i - p_j| - l_{ij} \right)^2.$$

- onde l_{ij} é o comprimento da mola (espaço desejado entre p_i e p_j)

KAMADA-KAWAI

$$l_{ij} = L \times d_{ij} \qquad L = L_0 / \max_{i<j} \mathrm{d}_{ij}$$

- sendo L a largura desejada para as arestas do desenho; L_o é o tamanho do lado do display e d_{ij} o caminho mais curto entre v_i e v_j
- O parâmetro k_{ij} é a constante da mola entre p_i e p_j e K é uma constante determinada empiricamente

$$k_{ij} = K / d_{ij}^2$$

- l_{ij} e k_{ij} são simétricas

KAMADA-KAWAI

- Como estamos trabalhando no espaço 2D

$$E = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \tfrac{1}{2} k_{ij} \left(| p_i - p_j | - l_{ij} \right)^2 .$$

- pode ser escrita como

$$E = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \tfrac{1}{2} k_{ij} \left\{ (x_i - x_j)^2 + (y_i - y_j)^2 + l_{ij}^2 - 2 l_{ij} \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \right\} .$$

- Deseja-se obter os valores das variáveis x e y que minimizem E

KAMADA-KAWAI

- Busca de mínimos locais

- É preciso que

$$\frac{\partial E}{\partial x_m} = \frac{\partial E}{\partial y_m} = 0 \quad \text{for } 1 \leqslant m \leqslant n.$$

- Neste estado, todas as molas estão balanceadas

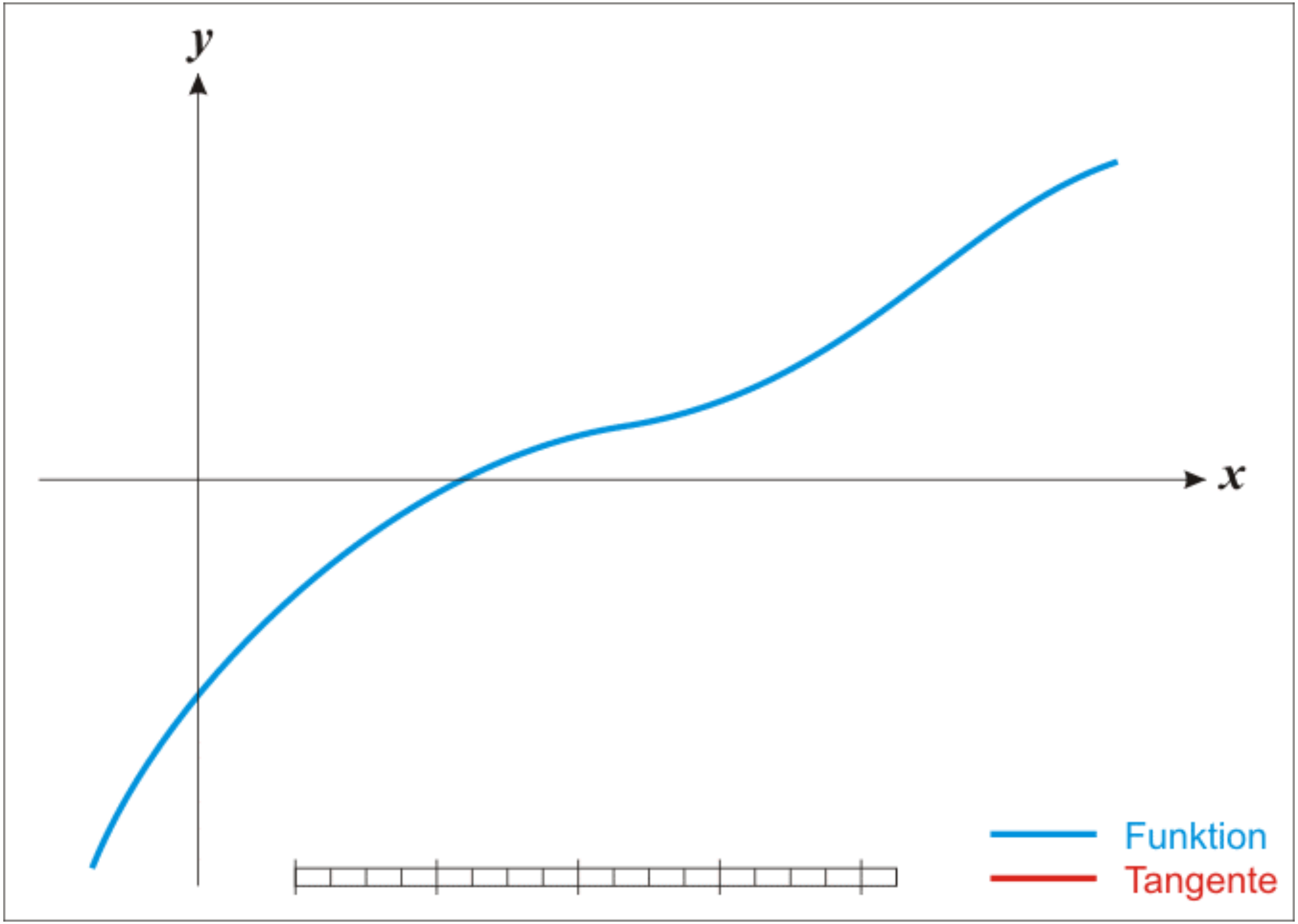
KAMADA-KAWAI

$$E = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \tfrac{1}{2} k_{ij} \left\{ (x_i - x_j)^2 + (y_i - y_j)^2 + l_{ij}^2 - 2 l_{ij} \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \right\}.$$

$$\frac{\partial E}{\partial x_m} = \sum_{i \neq m} k_{mi} \left\{ (x_m - x_i) - \frac{l_{mi}(x_m - x_i)}{\{(x_m - x_i)^2 + (y_m - y_i)^2\}^{1/2}} \right\},$$

$$\frac{\partial E}{\partial y_m} = \sum_{i \neq m} k_{mi} \left\{ (y_m - y_i) - \frac{l_{mi}(y_m - y_i)}{\{(x_m - x_i)^2 + (y_m - y_i)^2\}^{1/2}} \right\}.$$

KAMADA-KAWAI



$$x_{n+1} = x_n + f(x_n)/f'(x_n)$$

KAMADA-KAWAI

- Este sistema não pode ser resolvido pelo método de *Newton-Raphson*

- A opção adotada pelos autores é a de movimentar apenas uma partícula por iteração mantendo as outras fixas

- A partícula escolhida é a que maximiza o valor de

$$\Delta_m = \sqrt{\left\{ \frac{\partial E}{\partial x_m} \right\}^2 + \left\{ \frac{\partial E}{\partial y_m} \right\}^2}.$$

KAMADA-KAWAI

- Aplica-se a seguir o método de Newton-Raphson até que *Δm* seja pequeno o bastante

$$x_m^{(t+1)} = x_m^{(t)} + \delta x, \qquad y_m^{(t+1)} = y_m^{(t)} + \delta y \qquad \text{for } t = 0, 1, 2, \ldots$$

NEWTON-RAPHSON 2D

- Desejamos resolver o sistema não linear

$$\vec{\mathbf{F}}\ (\vec{\mathbf{X}}) = \vec{\mathbf{0}}$$

- dada a aproximação inicial $\vec{\mathbf{P}}_0$ e gerando a sequência $\{\vec{\mathbf{P}}_k\}$ que converge para a solução
- ou seja

$$\vec{\mathbf{F}}\ (\vec{\mathbf{P}}) = \vec{\mathbf{0}}$$

- Suponha que $\vec{\mathbf{P}}_k$ tenha sido obtido, use os seguintes passos para obter $\vec{\mathbf{P}}_{k+1}$

NEWTON-RAPHSON 2D

- Avalie a função $\vec{\mathbf{F}}(\vec{\mathbf{P}}_k) = \begin{pmatrix} f_1(p_k, q_k) \\ f_2(p_k, q_k) \end{pmatrix}$

- Avalie o Jacobiano $\mathbf{J}(\vec{\mathbf{P}}_k) = \begin{pmatrix} \frac{\partial}{\partial x} f_1(p_k, q_k) & \frac{\partial}{\partial y} f_1(p_k, q_k) \\ \frac{\partial}{\partial x} f_2(p_k, q_k) & \frac{\partial}{\partial y} f_2(p_k, q_k) \end{pmatrix}$

- Resolva o sistema $\mathbf{J}(\vec{\mathbf{P}}_k)\,\Delta\vec{\mathbf{P}} = -\vec{\mathbf{F}}(\vec{\mathbf{P}}_k)$

- Compute a próxima aproximação $\vec{\mathbf{P}}_{k+1} = \vec{\mathbf{P}}_k + \Delta\vec{\mathbf{P}}$

KAMADA-KAWAI

$$\frac{\partial^2 E}{\partial x_m^2} = \sum_{i \neq m} k_{mi} \left\{ 1 - \frac{l_{mi}(y_m - y_i)^2}{\{(x_m - x_i)^2 + (y_m - y_i)^2\}^{3/2}} \right\},$$

$$\frac{\partial^2 E}{\partial x_m \partial y_m} = \sum_{i \neq m} k_{mi} \frac{l_{mi}(x_m - x_i)(y_m - y_i)}{\{(x_m - x_i)^2 + (y_m - y_i)^2\}^{3/2}},$$

$$\frac{\partial^2 E}{\partial y_m \partial x_m} = \sum_{i \neq m} k_{mi} \frac{l_{mi}(x_m - x_i)(y_m - y_i)}{\{(x_m - x_i)^2 + (y_m - y_i)^2\}^{3/2}},$$

$$\frac{\partial^2 E}{\partial y_m^2} = \sum_{i \neq m} k_{mi} \left\{ 1 - \frac{l_{mi}(x_m - x_i)^2}{\{(x_m - x_i)^2 + (y_m - y_i)^2\}^{3/2}} \right\}.$$

KAMADA-KAWAI

$$\frac{\partial^2 E}{\partial x_m^2}\left(x_m^{(t)},\ y_m^{(t)}\right)\delta x + \frac{\partial^2 E}{\partial x_m \partial y_m}\left(x_m^{(t)},\ y_m^{(t)}\right)\delta y = -\frac{\partial E}{\partial x_m}\left(x_m^{(t)},\ y_m^{(t)}\right),$$

$$\frac{\partial^2 E}{\partial y_m \partial x_m}\left(x_m^{(t)},\ y_m^{(t)}\right)\delta x + \frac{\partial^2 E}{\partial y_m^2}\left(x_m^{(t)},\ y_m^{(t)}\right)\delta y = -\frac{\partial E}{\partial y_m}\left(x_m^{(t)},\ y_m^{(t)}\right).$$

KAMADA-KAWAI

compute d_{ij} for $1 \leqslant i \neq j \leqslant n$;
compute l_{ij} for $1 \leqslant i \neq j \leqslant n$;
compute k_{ij} for $1 \leqslant i \neq j \leqslant n$;
initialize $p_1, p_2, \ldots, p_n$;
while ($\max_i \Delta_i > \epsilon$){
    let p_m be the particle satisfying $\Delta_m = \max_i \Delta_i$;
    while ($\Delta_m > \epsilon$){
        compute δx and δy by solving (11) and (12);
        $x_m := x_m + \delta x$;
        $y_m := y_m + \delta y$;
    }
}

KAMADA-KAWAI

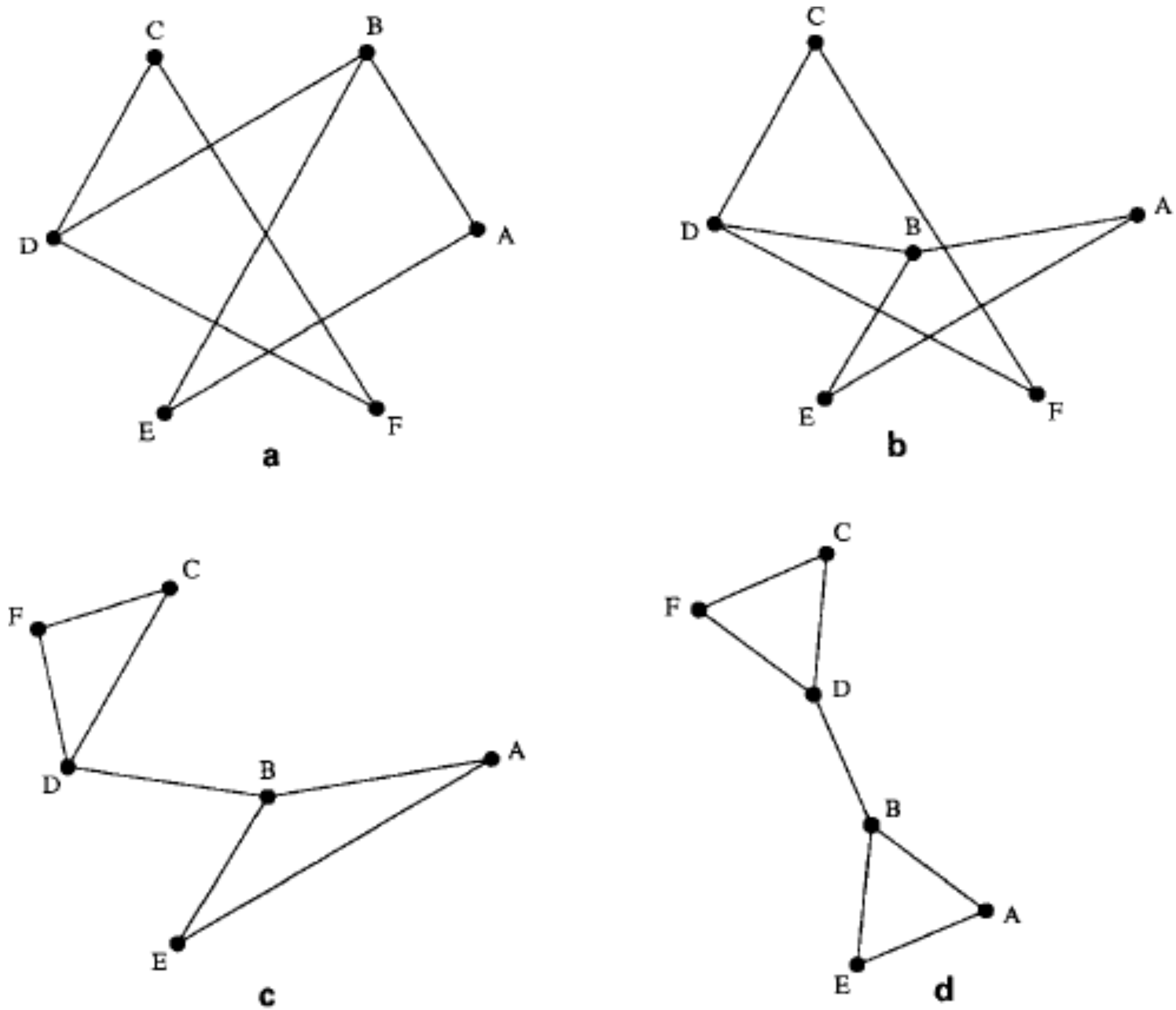


Fig. 2. The energy minimization process.

KAMADA-KAWAI

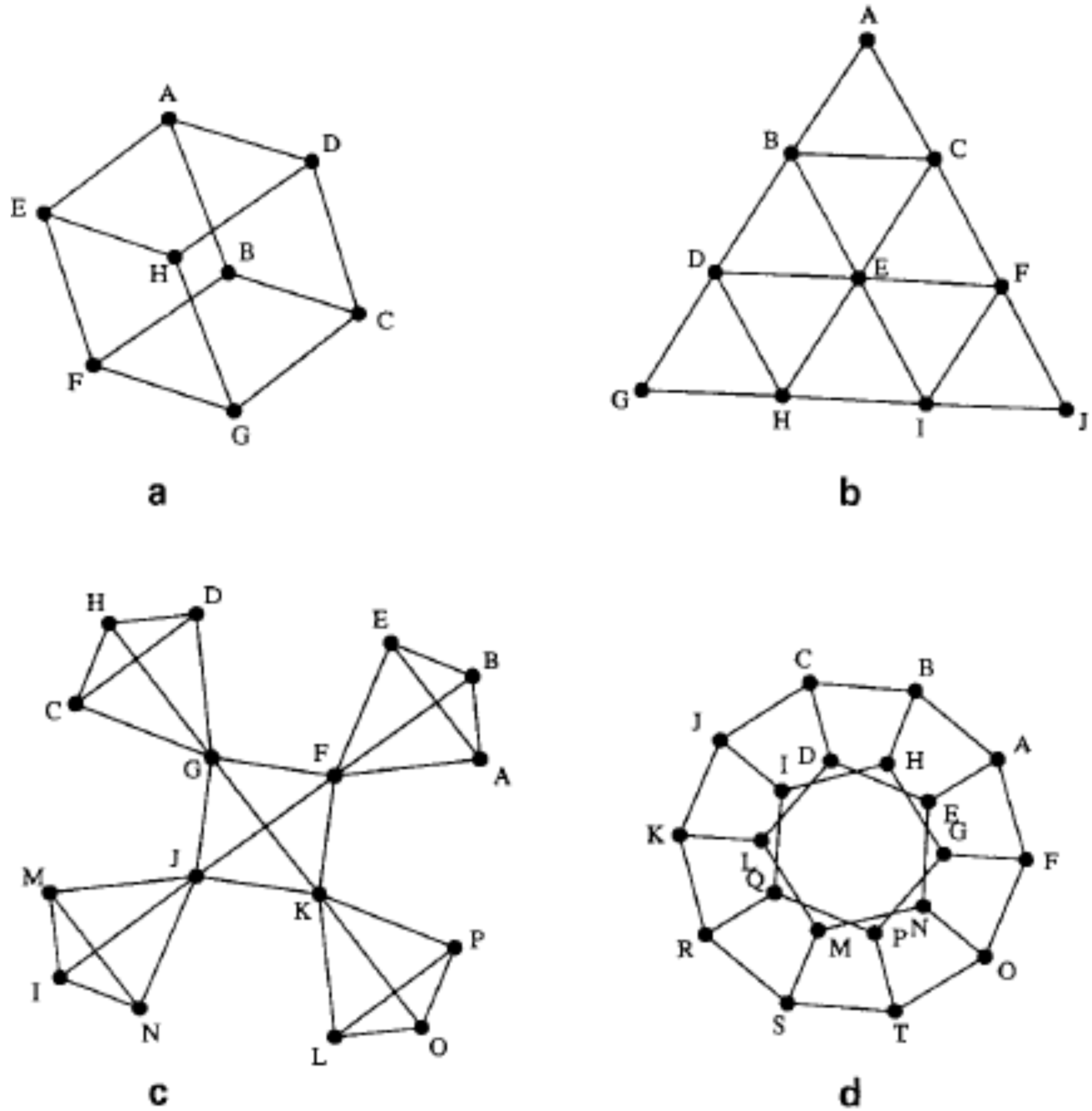


Fig. 3.

KAMADA-KAWAI

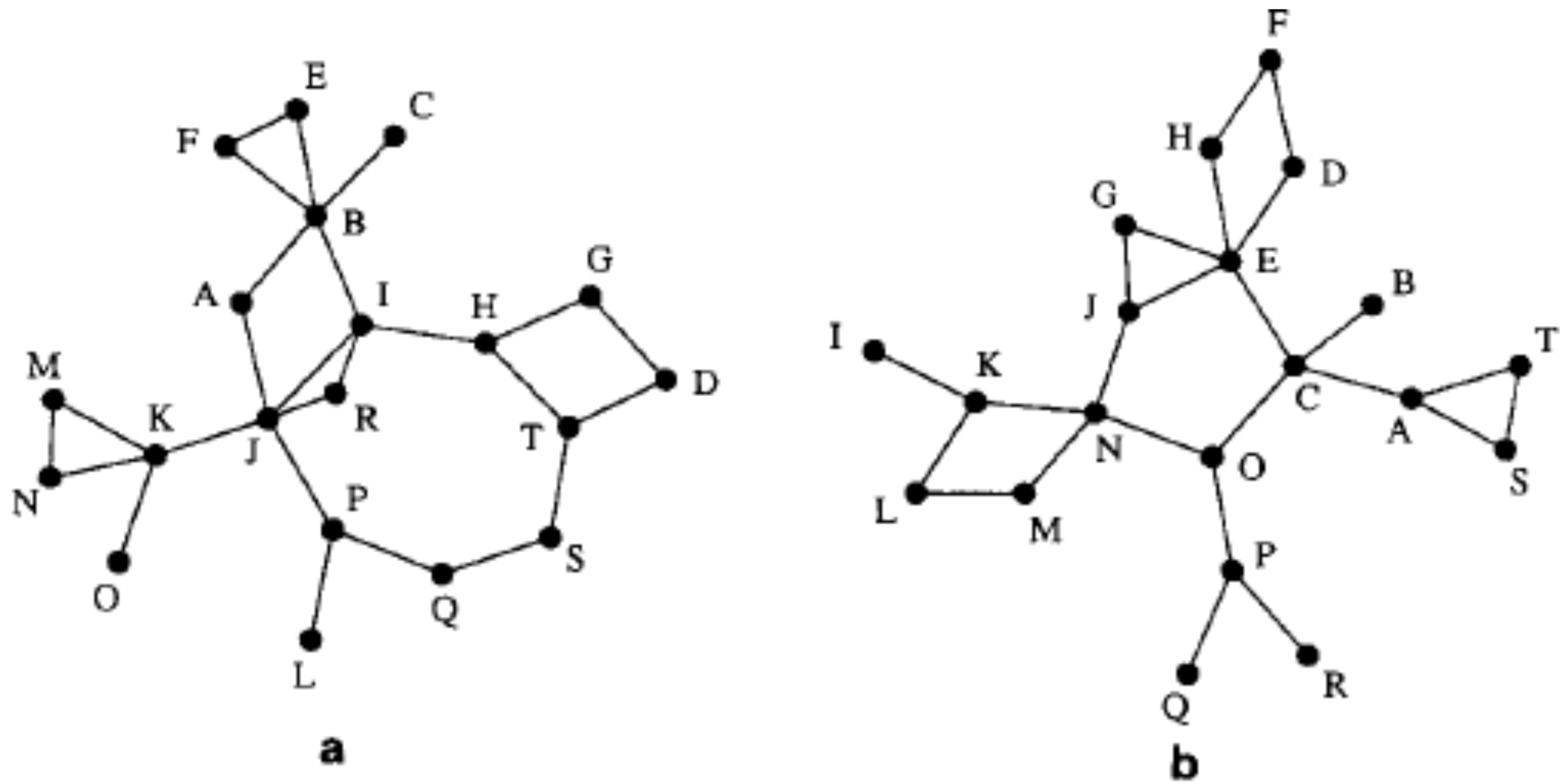


Fig. 4. Pictures of asymmetric graphs.

KAMADA-KAWAI

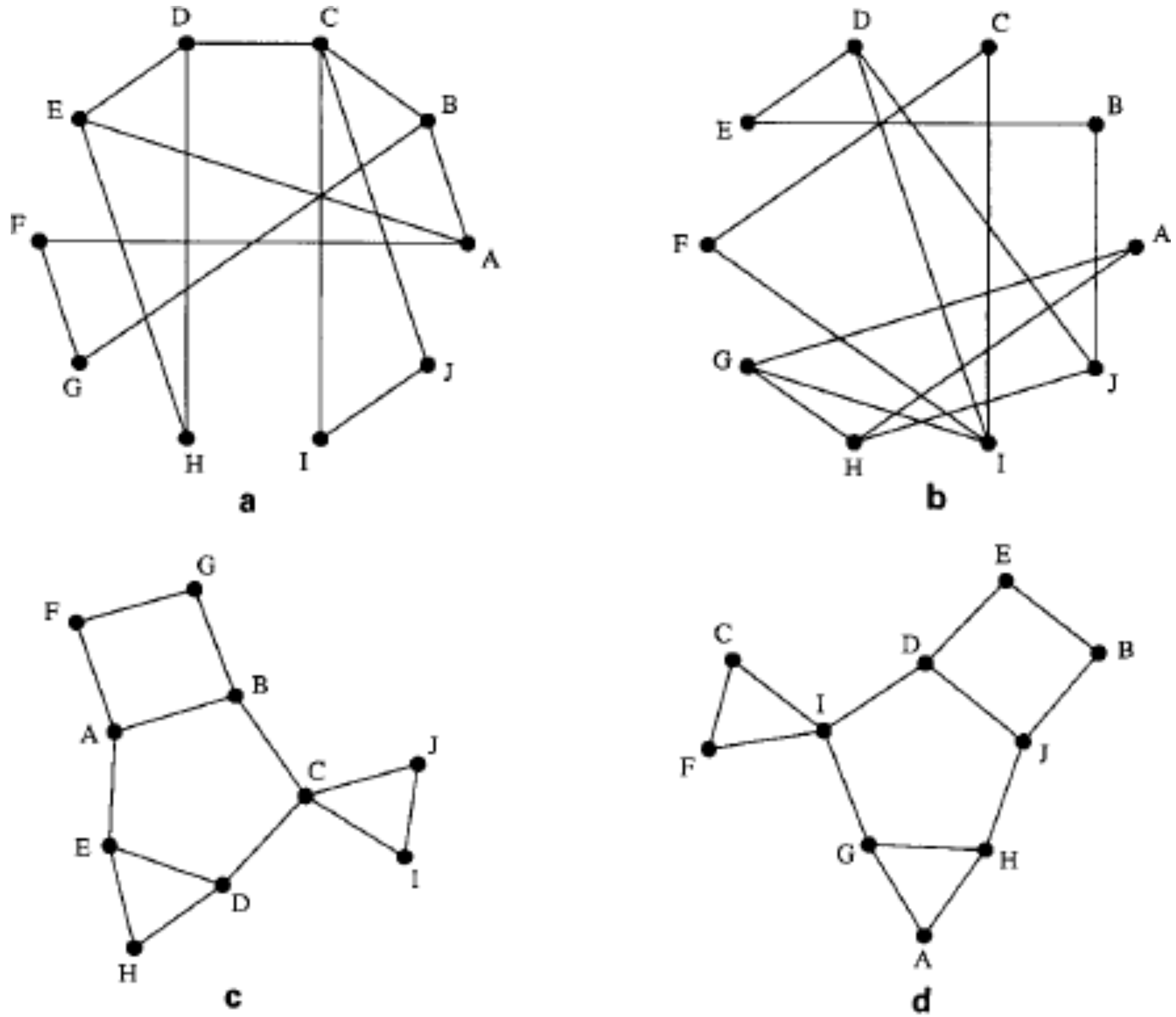


Fig. 5. Pictures of isomorphic graphs.

KAMADA-KAWAI

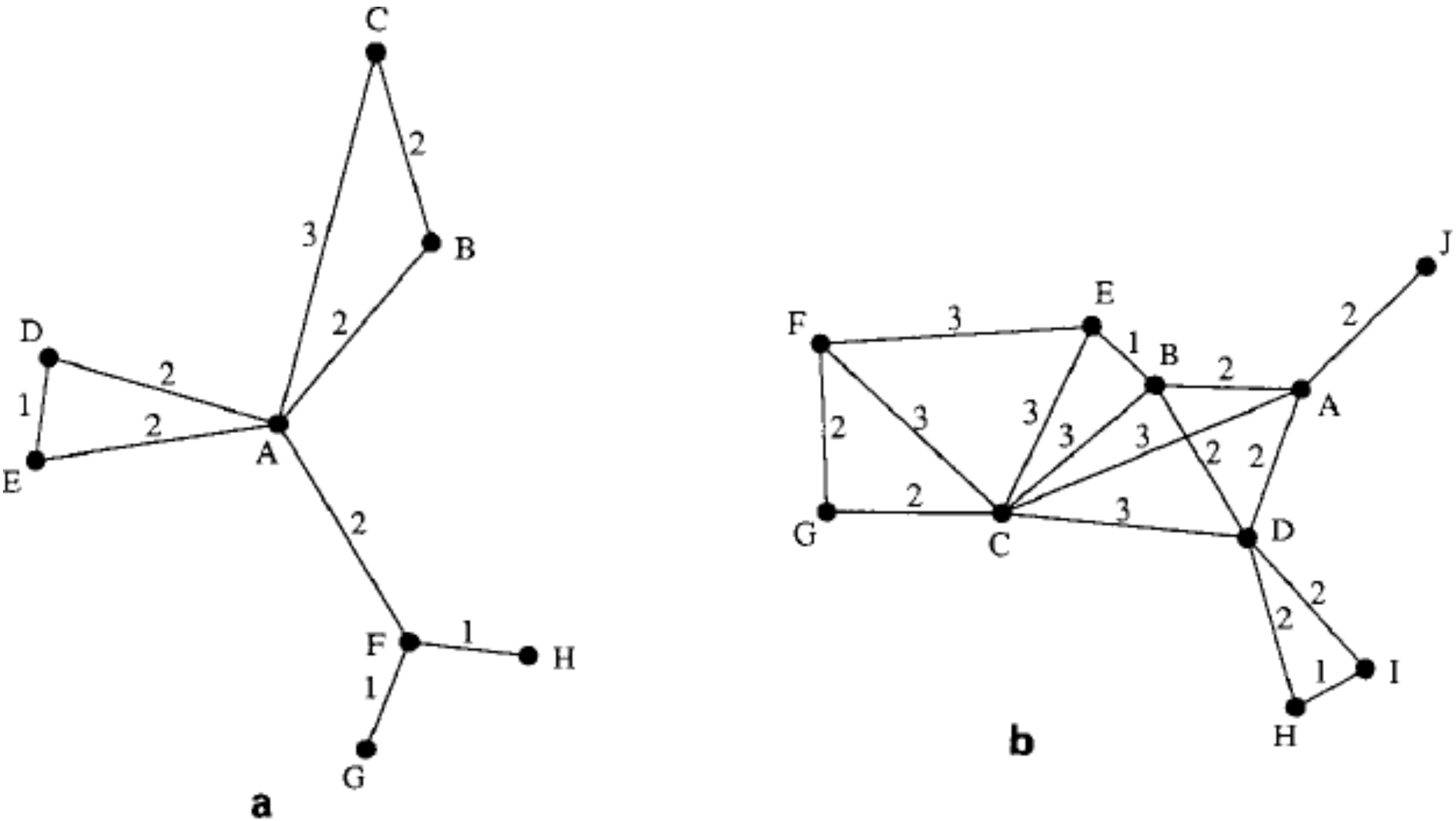


Fig. 6. Pictures of weighted graphs.

GRAPH-DRAWING BY FORCE-DIRECTED PLACEMENT

T.M.J. Fruchterman e E.M. Reingold
Software – Practice and Experience
1991

FRUCHTERMAN-GOLDMAN

- Partículas atômicas ou corpos celestes exercendo forças atrativas e repulsivas uns nos outros, induzindo movimento

- Não é necessário ser fiel às forças, mas sim aos critérios estéticos

- Aplicação de forças não realistas de maneira também não realista

- Forças não implicam em aceleração

FRUCHTERMAN-GOLDMAN

$$k = C\sqrt{\left(\frac{\text{area}}{\text{number of vertices}}\right)}$$

- *k* é a distância ótima entre os vértices para que eles fiquem uniformemente distribuídos no espaço

- *C* é obtido empiricamente visando melhorar o grafo esteticamente

FRUCHTERMAN-GOLDMAN

$$f_a(d) = d^2/k$$

$$f_r(d) = -k^2/d$$

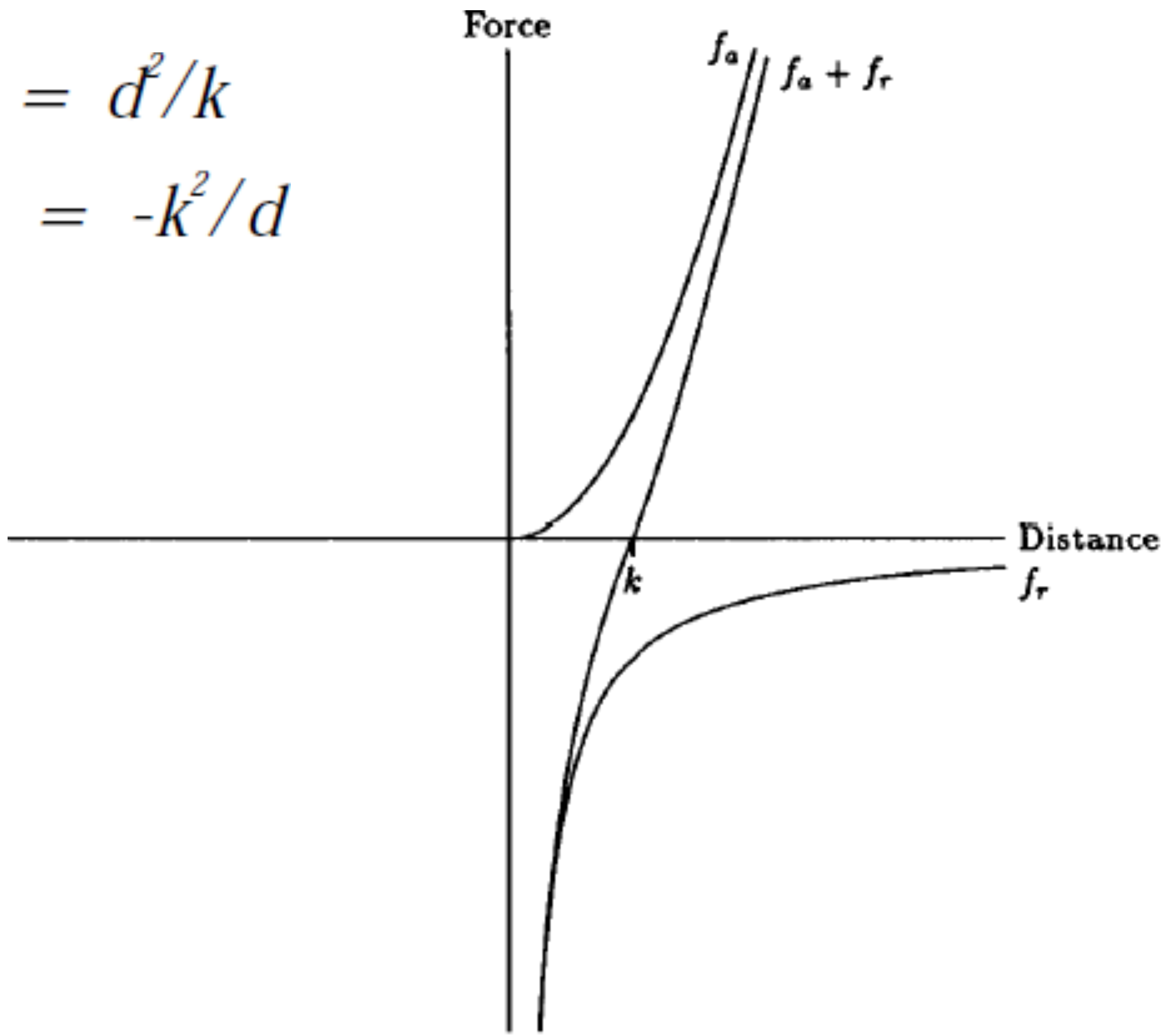


Figure 2. Forces versus distance

FRUCHTERMAN-GOLDMAN

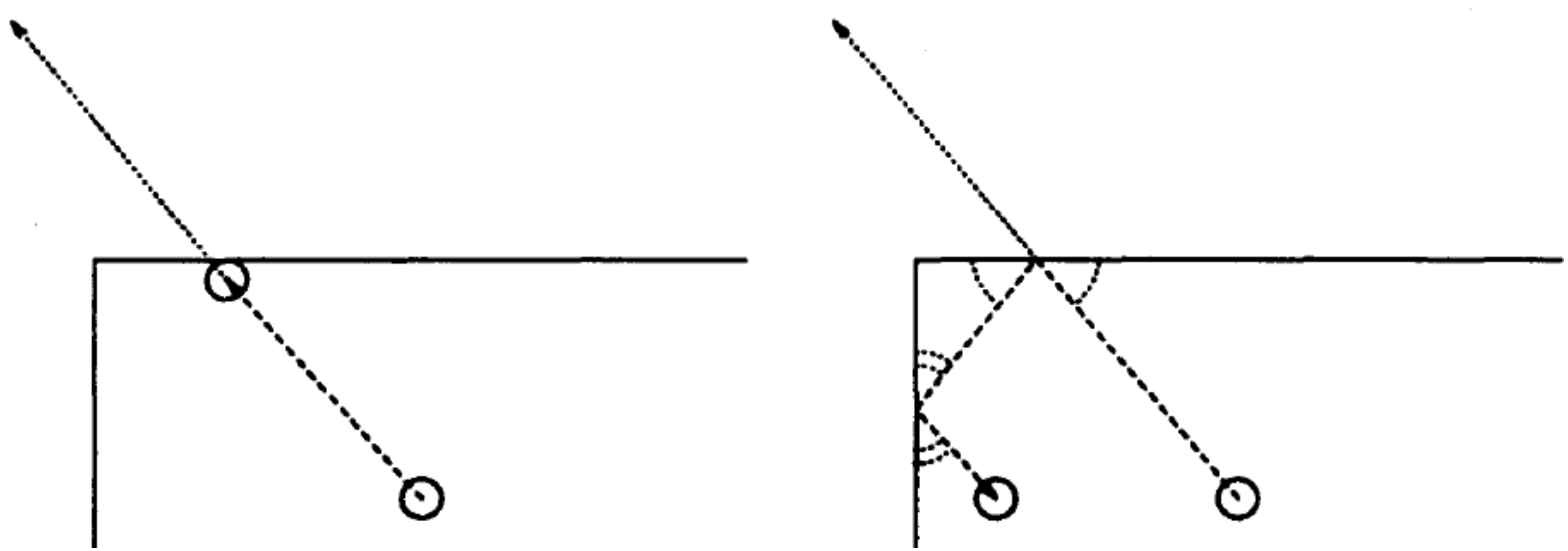
```
area  := W * L; { W and L are the width and length of the frame }
G := (V, E); { the vertices are assigned random initial positions }
k  := √(area/|V|);
function f_a(z) := begin return x²/k end;
function f_r(z) := begin return k²/z end;

for i := 1 to iterations do begin
    { calculate repulsive forces}
    for v in V do begin
        { each vertex has two vectors: .pos and .disp }
        v.disp := 0;
        for u in V do
            if (u # v) then begin
                { Δ is short hand for the difference}
                { vector between the positions of the two vertices )
                Δ := v.pos - u.pos;
                v.disp := v.disp + ( Δ /| Δ |) * fr (| Δ |)
            end
    end
```

FRUCHTERMAN-GOLDMAN

{ calculate attractive forces }
for e **in** E **do begin**
  { each edge is an ordered pair of vertices $.v$ and $.u$ }
  $\Delta := e.v.pos - e.u.pos$
  $e.v.disp := e.v.disp - (\Delta / |\Delta|) * f_a(|\Delta|)$;
  $e.u.disp := e.u.disp + (\Delta / |\Delta|) * f_a(|\Delta|)$
end

FRUCHTERMAN-GOLDMAN

```
    { limit the maximum displacement to the temperature t }
    { and then prevent from being displaced outside frame}
    for v in V do begin
        v.pos := v.pos + ( v. disp/ |v.disp|) * min ( v.disp, t );
        v.pos.x  :=  min(W/2,  max(-W/2,  v.pos.x));
        v.pos.y  :=  min(L/2,  max(–L/2,  v.pos.y))
    end
    { reduce the temperature as the layout approaches a better configuration }
    t := cool(t)
end
```



DRAWING GRAPHS NICELY USING SIMULATED ANNEALING

R. Davidson e D. Harel
ACM Transactions on Graphics
1996

SIMULATED ANNEALING

- É um método de otimização flexível usado para solução de problemas de otimização combinatória

- Espaços de busca enormes onde a busca exaustiva é inviável e se tem uma função de custo que se quer maximizar ou minimizar

- Em geral, inicia-se de uma solução aleatória que é refinada iterativamente terminando comumente em um mínimo local

SIMULATED ANNEALING

- Tentativa de escapar dos mínimos locais usando o processo de resfriamento (annealing)
- É conhecido o fato de que líquidos quando resfriados de forma lenta formam uma estrutura cristalina extremamente organizada, chamada cristal, que tem energia mínima
- A cada intervalo de tempo, os átomos tem tempo de se organizar atingindo o equilíbrio térmico

SIMULATED ANNEALING

- Entidades a serem determinadas
  - O conjunto de configurações que o sistema pode ter inclusive a inicial (aleatória)
  - As regras para geração de novas configurações (vizinhança escolhida aleatoriamente)
  - A função objetivo a ser minimizada (energia)
  - O esquema de resfriamento
  - Condição de terminação (tempo, valor da função objetivo)

SIMULATED ANNEALING - ALGORITMO

1. Escolha a configuração inicial do sistema σ e uma temperatura T
2. Repita
   A. Escolha uma nova configuração σ' para o sistema a partir da vizinhança de σ
   B. Seja E e E' as energias nos estados σ e σ', se $E' < E$, então $\sigma = \sigma'$
   C. Decremente a temperatura
3. Se o critério de parada for alcançado, pare. Se não, vá para o passo 2

DAVIDSON-HAREL

- **Configuração:** um desenho candidato, ou seja, posições no grid para cada um dos nós
- **Vizinhança:** configurações vizinhas são aquelas que diferem pela posição de apenas um nó (dentro de um raio que vai decrescendo com o andamento do processo)
- **Distribuição dos nós:** $a_{ij} = \lambda_1 / d_{ij}^2$
- **Bordas:** $m_i = \lambda_2 (1/r_i^2 + 1/l_i^2 + 1/t_i^2 + 1/b_i^2)$
- Aumentar o valor λ_2 de em relação a λ_1 puxa os nós para o centro, enquanto o contrário provoca o uso dos espaços próximos às bordas

DAVIDSON-HAREL

- **Comprimento das arestas:** $c_k = \lambda_3 d_k^2$
- **Cruzamento de arestas:** λ_4
- **Distâncias nó-aresta:** $h_{kl} = \lambda_5 / g_{kl}^2$, onde g_{kl} é a menor distância entre o vértice k e a aresta l (usado apenas no estágio de refinamento)

DAVIDSON-HAREL

- **Esquema de resfriamento:** esquema geométrico
- **Temperatura inicial:** se configuração inicial aleatória, temperatura deve ser grande o bastante para permitir qualquer movimentação. Se configuração inicial próxima de algum mínimo local, escolher temperatura mais baixa para manter proximidade com solução

- **Resfriamento:** $T_p+1 = \lambda\, T_p\ (0.6 < \lambda < 0.95)$
- **Muito rápido:** não escapa de mínimos locais
- **Muito lento:** melhores resultados às custas de longo tempo de execução

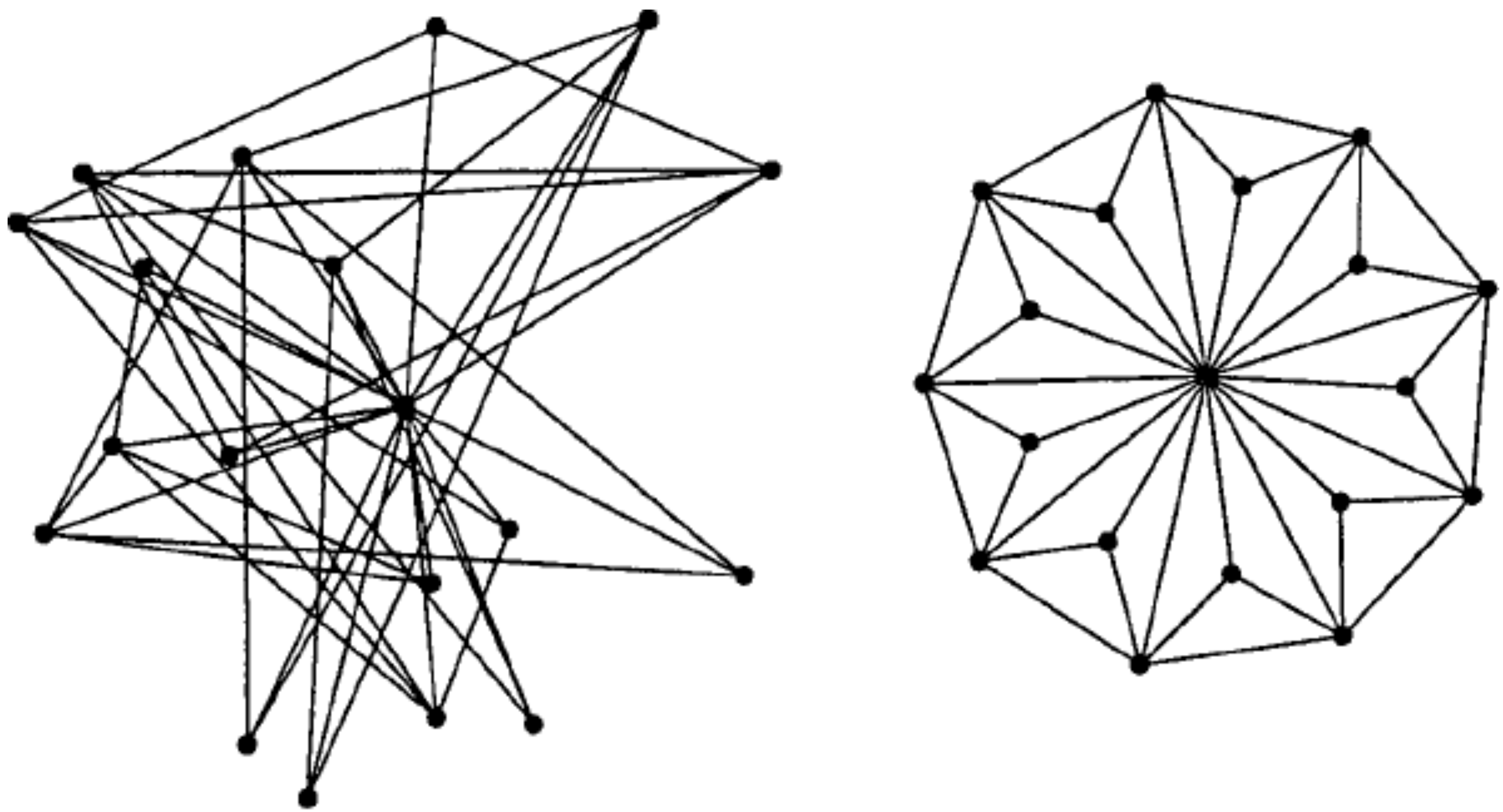
DAVIDSON-HAREL

- **Condição de parada:** usam 10 estágios (em termos de valores de temperatura), mas relatam que melhor seria usar critério de convergência

DAVIDSON-HAREL

- **Refinamento:** executado apenas ao término da execução do algoritmo original
  - Função de energia ganha o componente de distância nó-aresta
  - Permite movimentação de nós em grau extremamente pequeno (alguns pixels) por três estágios

DAVIDSON-HAREL



(a) (b)

Figure 1

DAVIDSON-HAREL

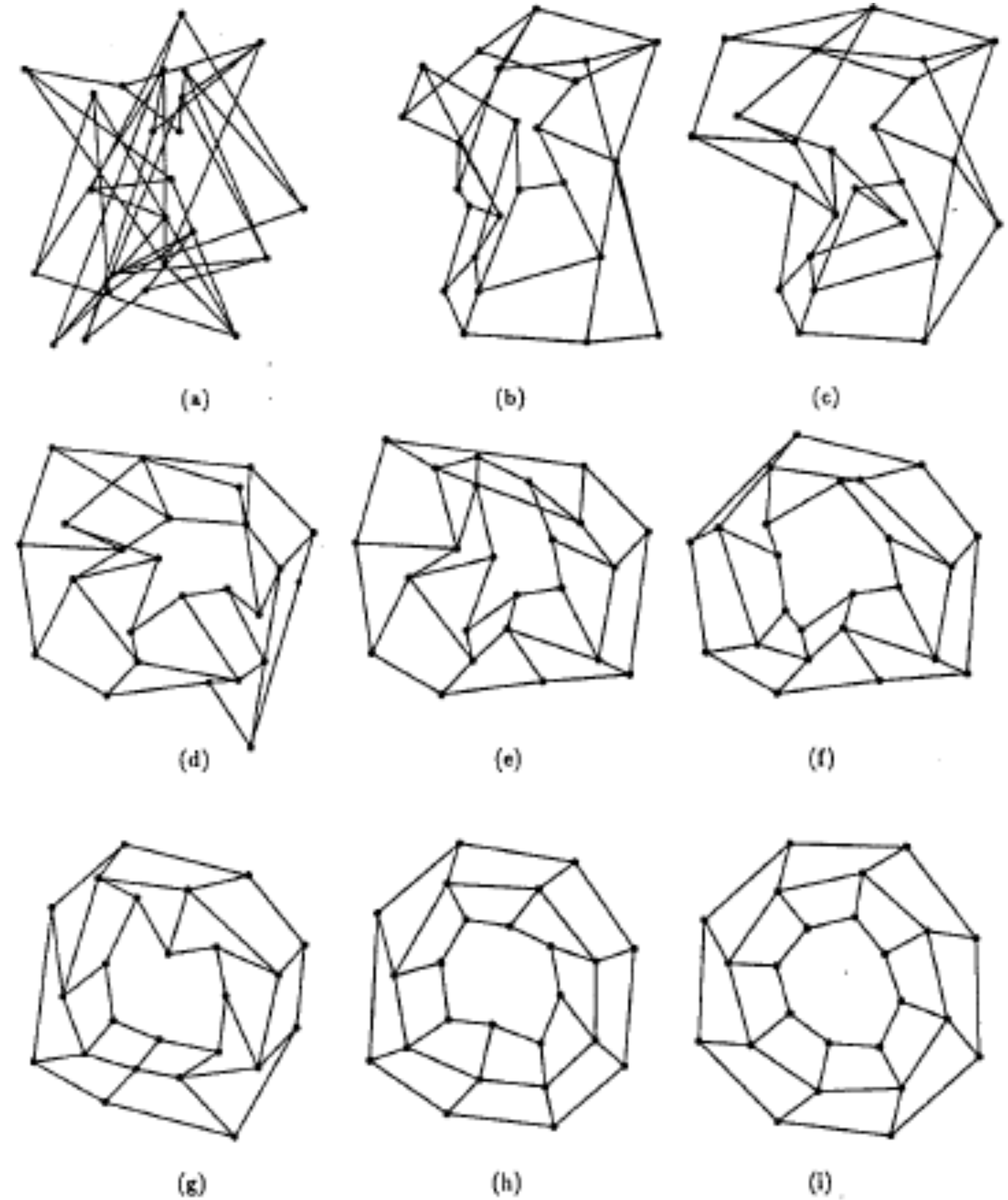
(a) entrada

(b-h) passos intermediários

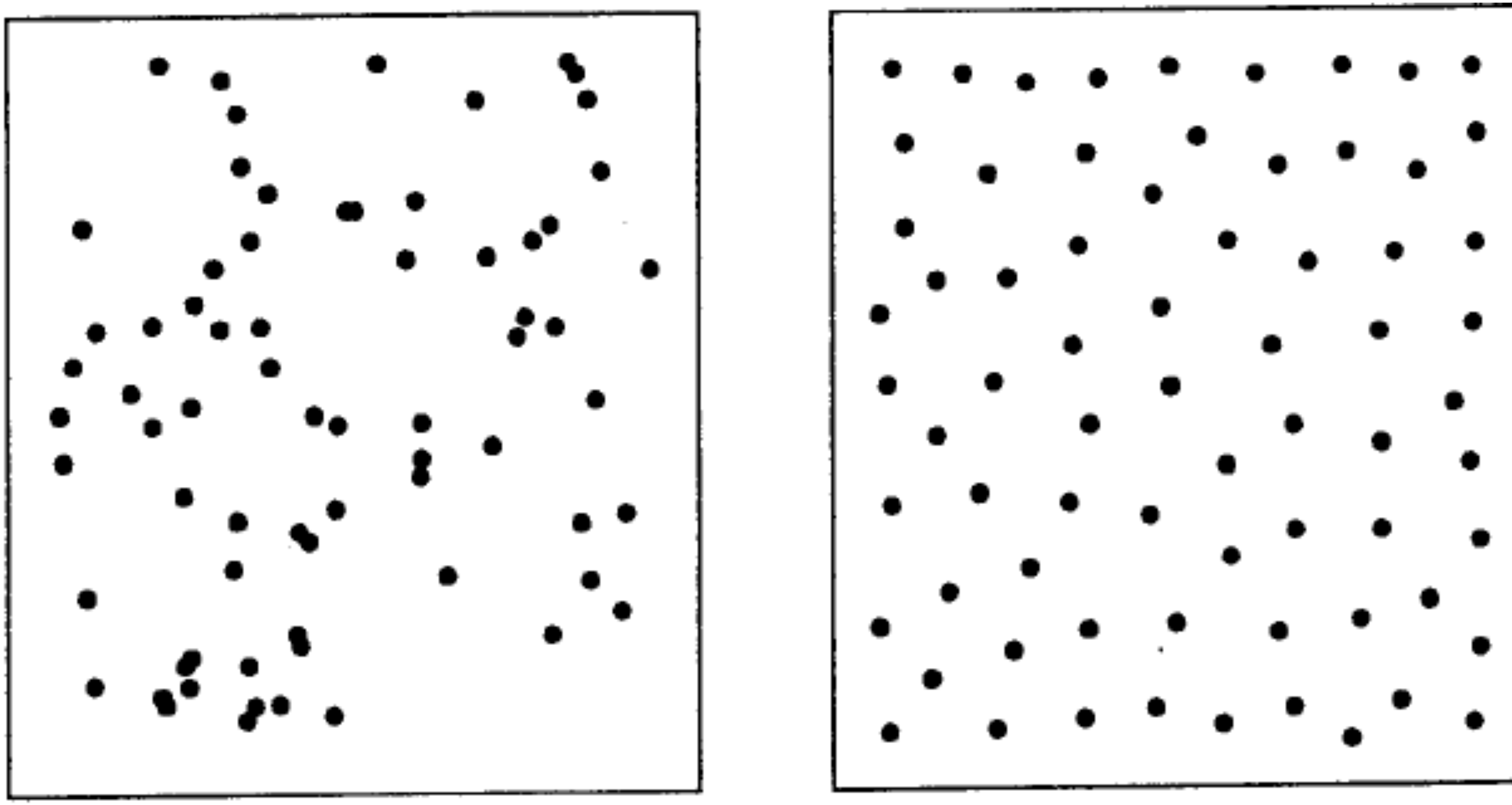
(g) fim do annealing

(h) fine tunning

(i) resultado final



DAVIDSON-HAREL



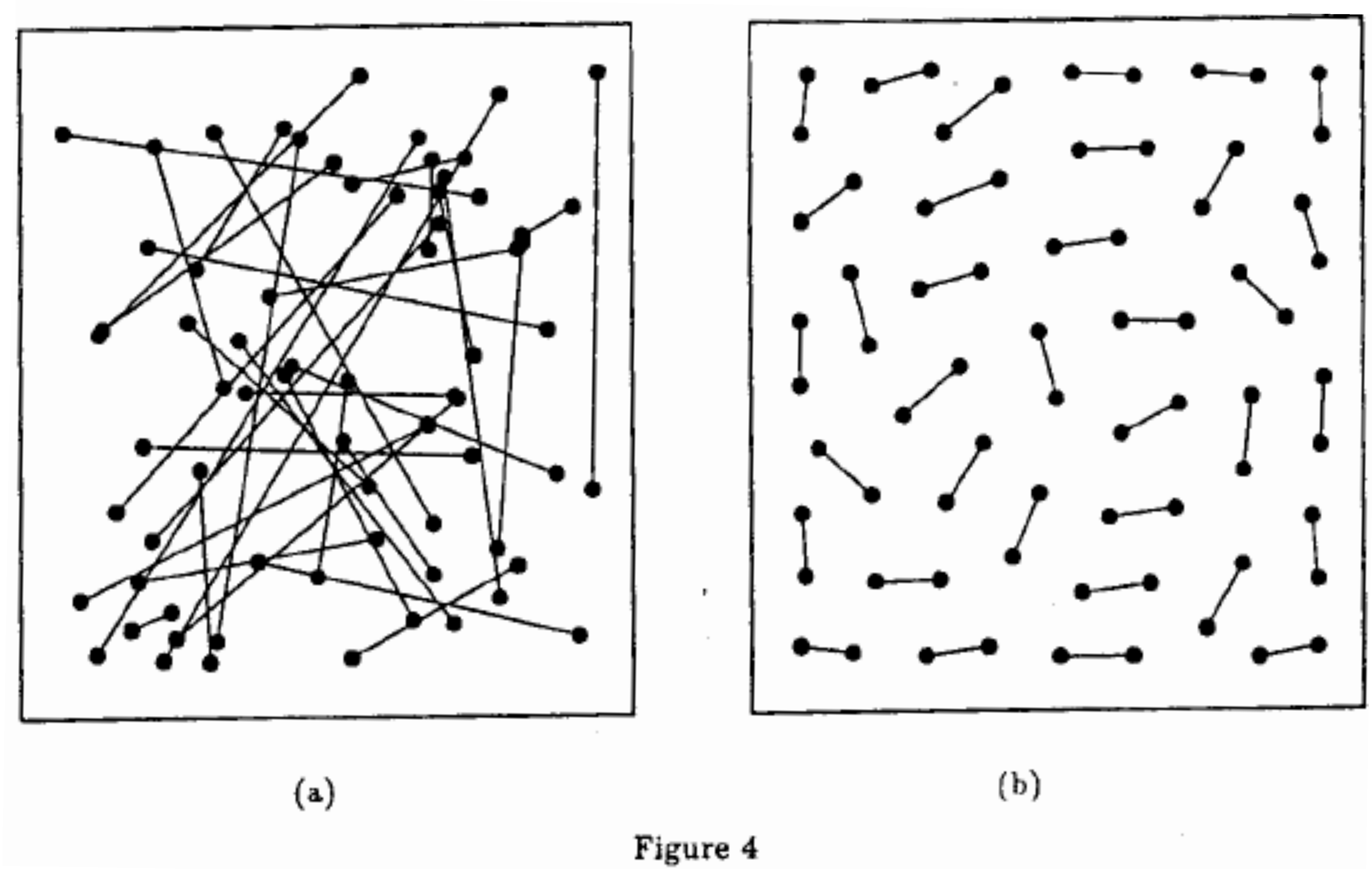
(a)

(b)

Figure 3

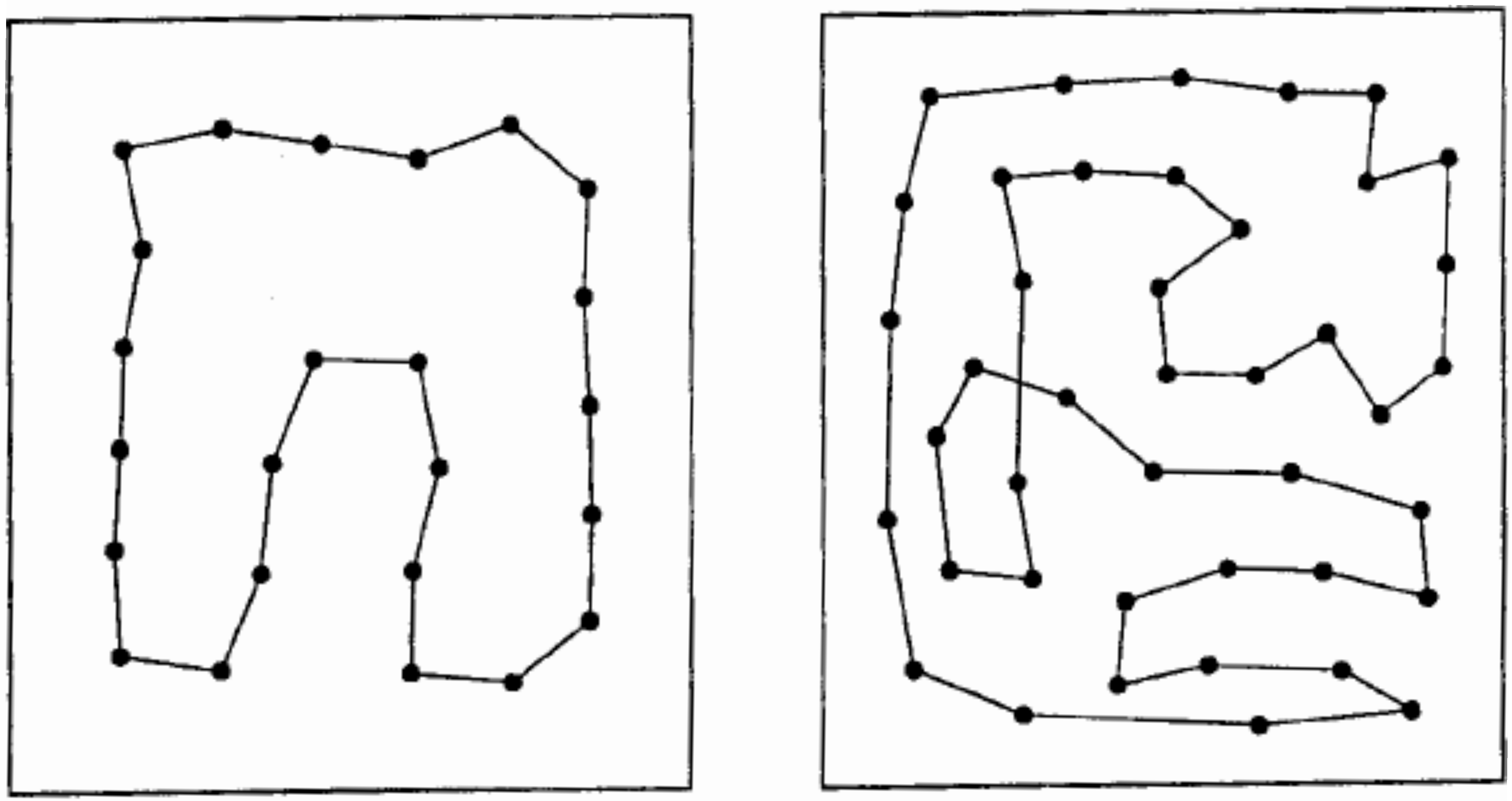
Efeito do algoritmo com 70 pontos desconectados: pontos bem distribuídos

DAVIDSON-HAREL



Efeito do algoritmo com 25 arestas desconectadas

DAVIDSON-HAREL



(a) (b)

Figure 5

Ciclos com 24 e 40 nós

DAVIDSON-HAREL

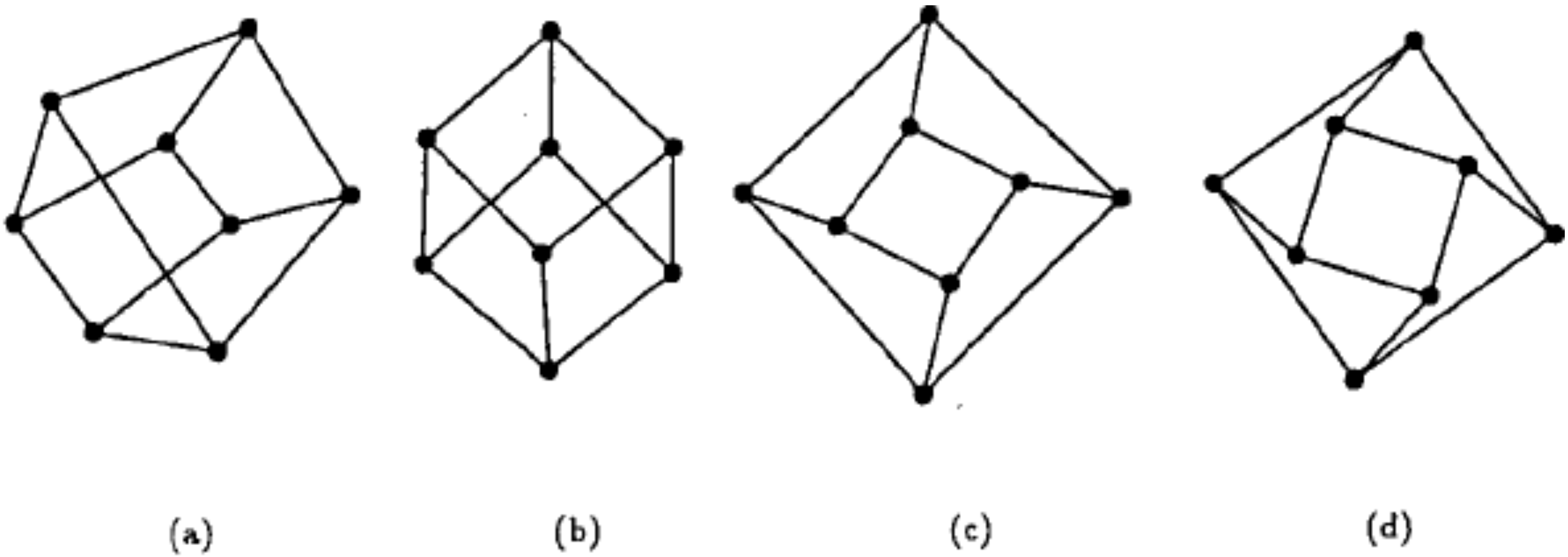


Figure 6

(a) e (b) diferentes resultados com o mesmo grafo (não determinismo)

(c) Dobrou o peso do componente cruzamento de arestas

(d) Reduziu o componente distância entre nós e arestas

OBS.: Algoritmos de Fruchterman-Reingold e o de Kamada-Kawai desenham como em (b)

DAVIDSON-HAREL

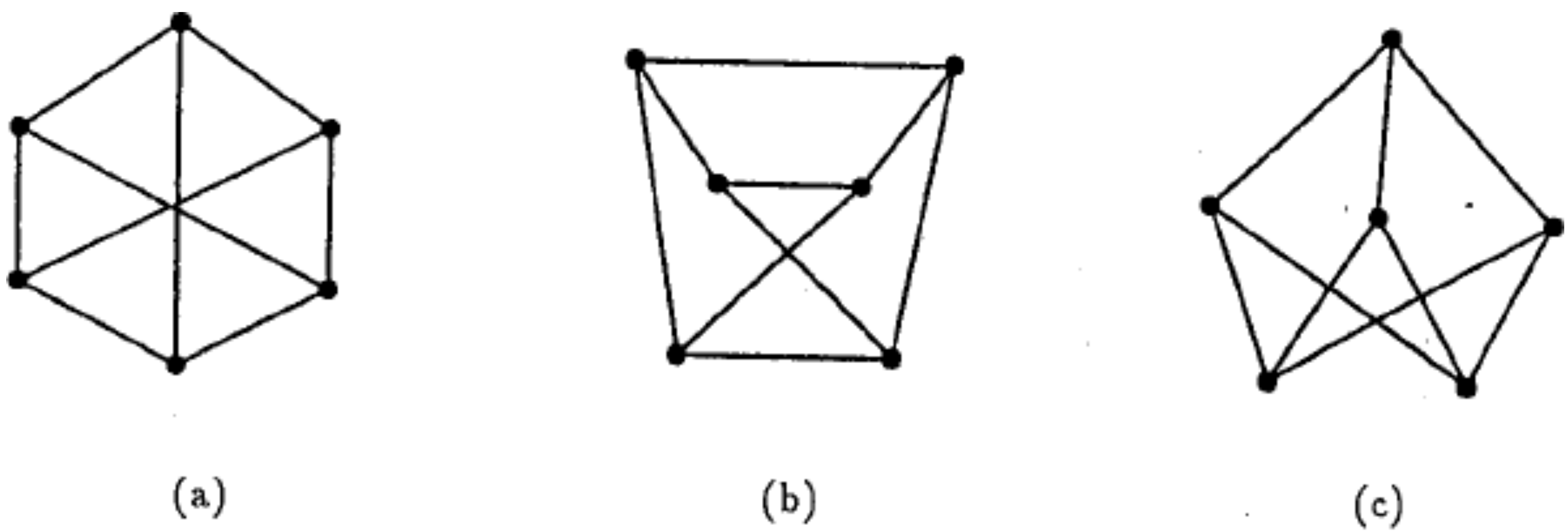


Figure 7

Famoso grafo K33

(b) Obtida com os parâmetros padrão

(c) Obtida eliminando o componente de cruzamento de arestas

DAVIDSON-HAREL

(a) Grafo original

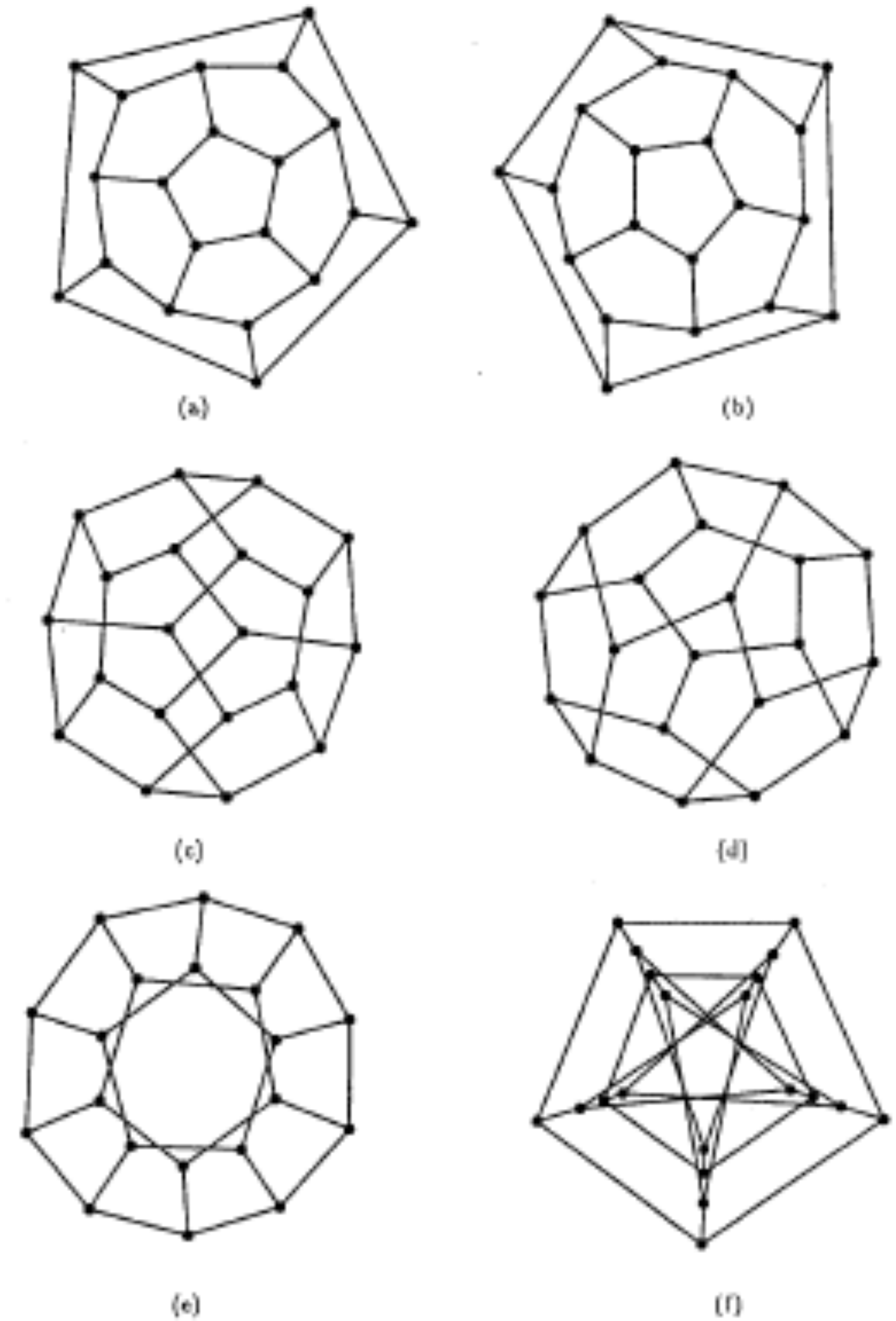
(b) Grafo desenhado pelo algoritmo

(c) Grafo desenhado pelo algoritmo

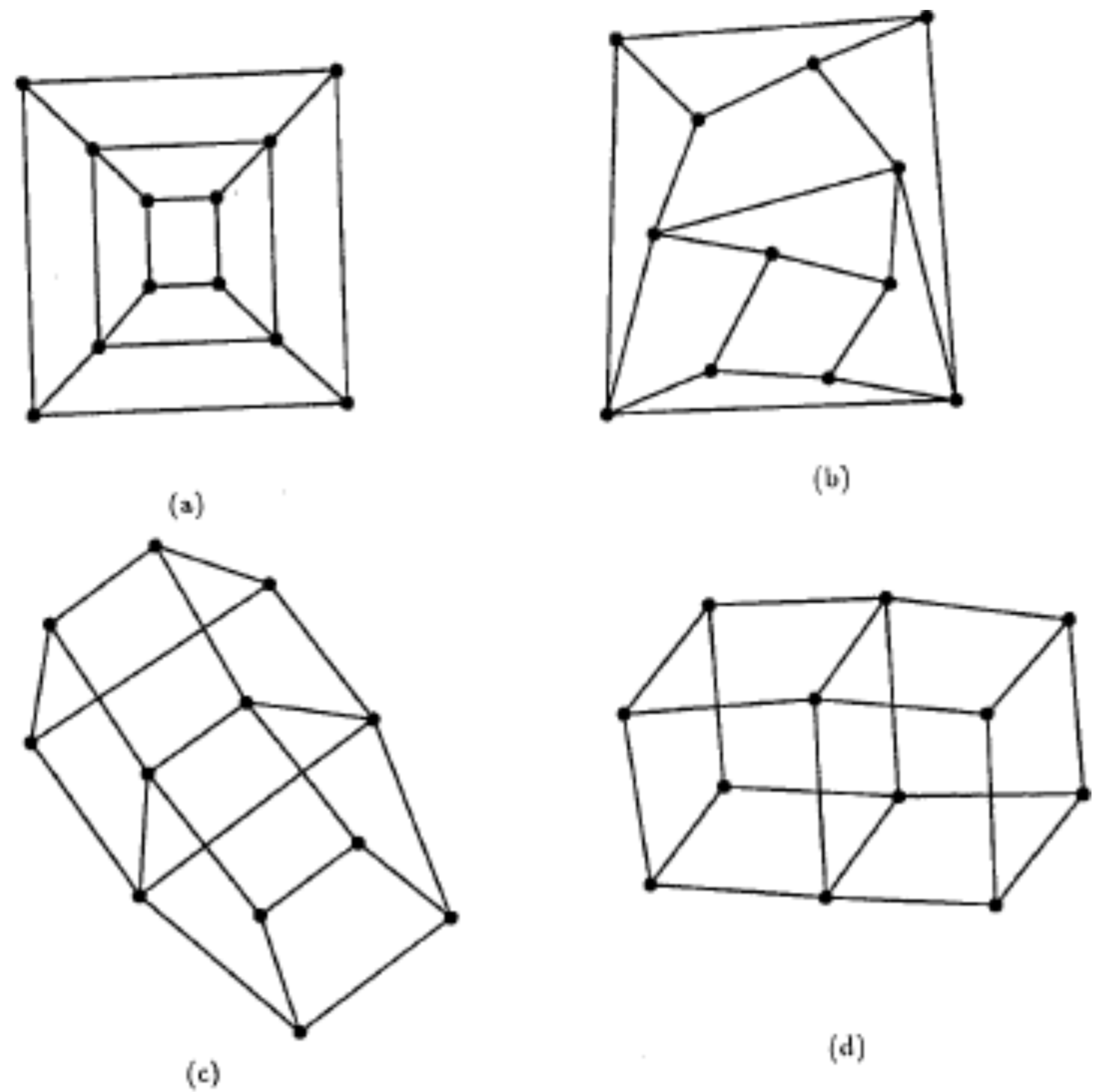
(d) Maior peso para arestas uniformes (Fruchterman e Reingold)

(e) (Fruchterman e Reingold) e (Kamada e Kawai)

(f) Lipton et al., 1985 mostra que nem sempre o critério de simetria produz critérios interessantes



DAVIDSON-HAREL



DAVIDSON-HAREL

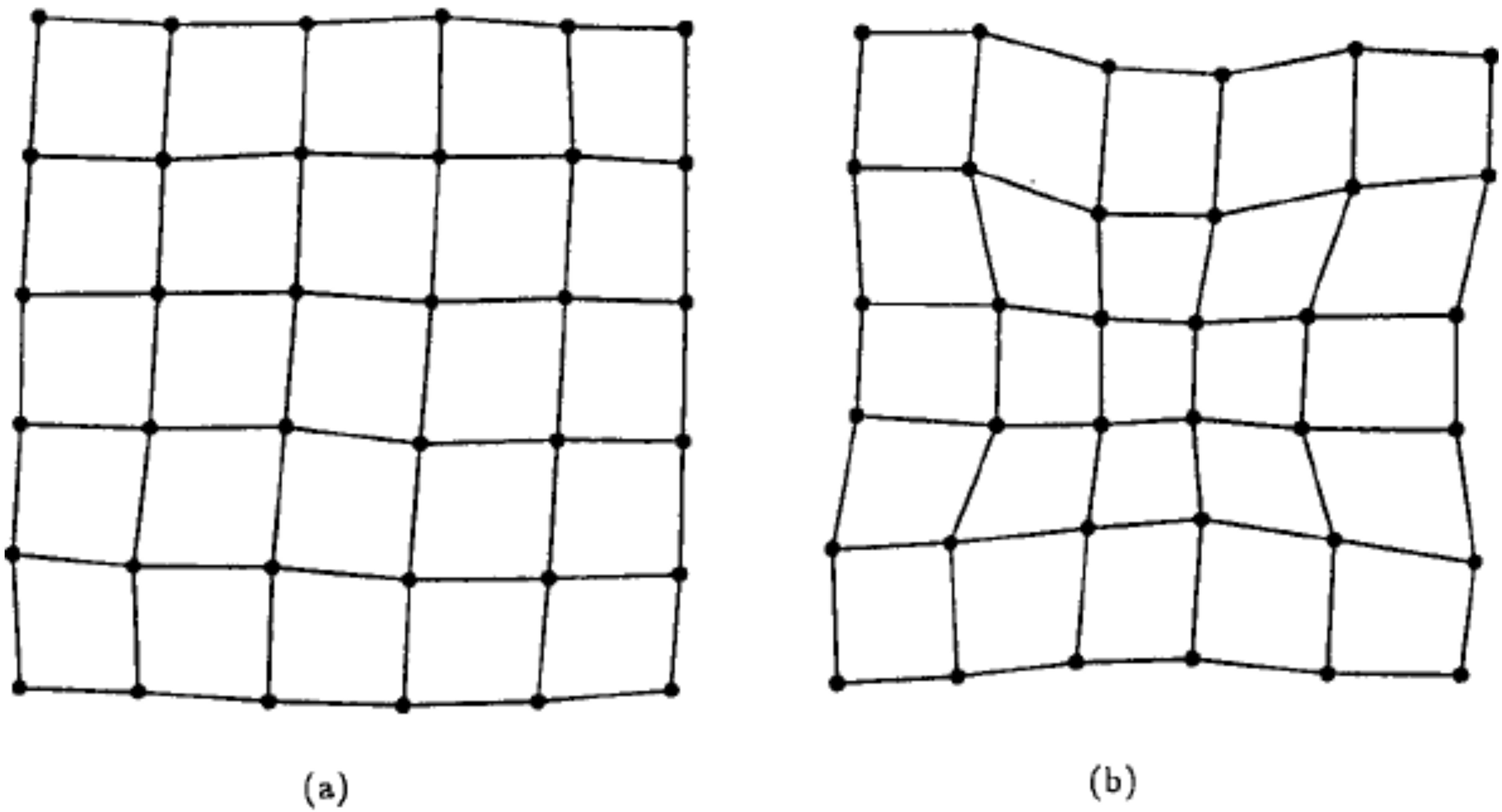


Figure 10

Quatro nós do centro tem diferentes pesos (o dobro) para o componente comprimento das arestas

DAVIDSON-HAREL

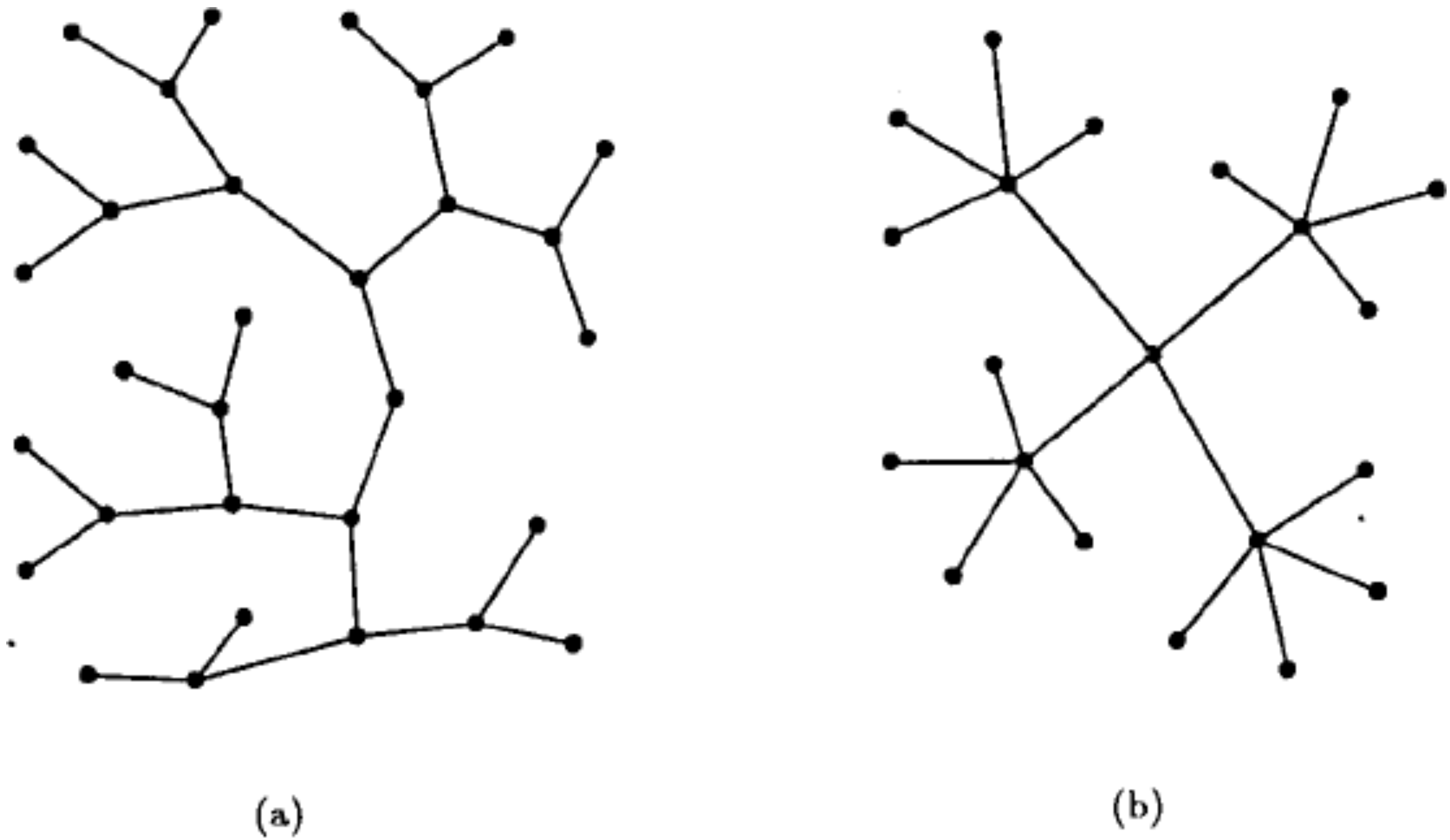


Figure 11

Árvores binárias (nenhum conhecimento sobre a topologia é incluído)

DAVIDSON-HAREL

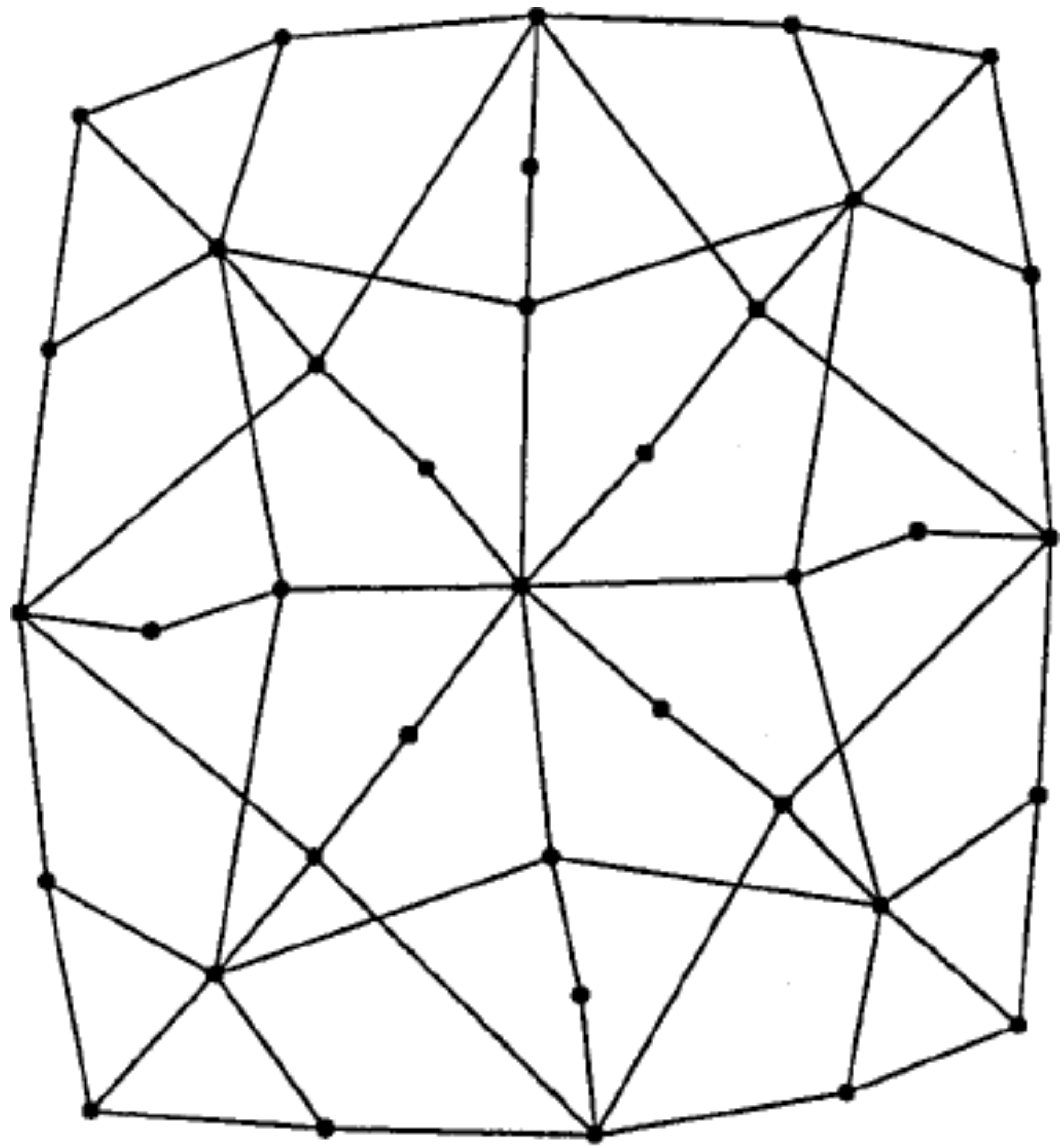


Figure 12

DAVIDSON-HAREL

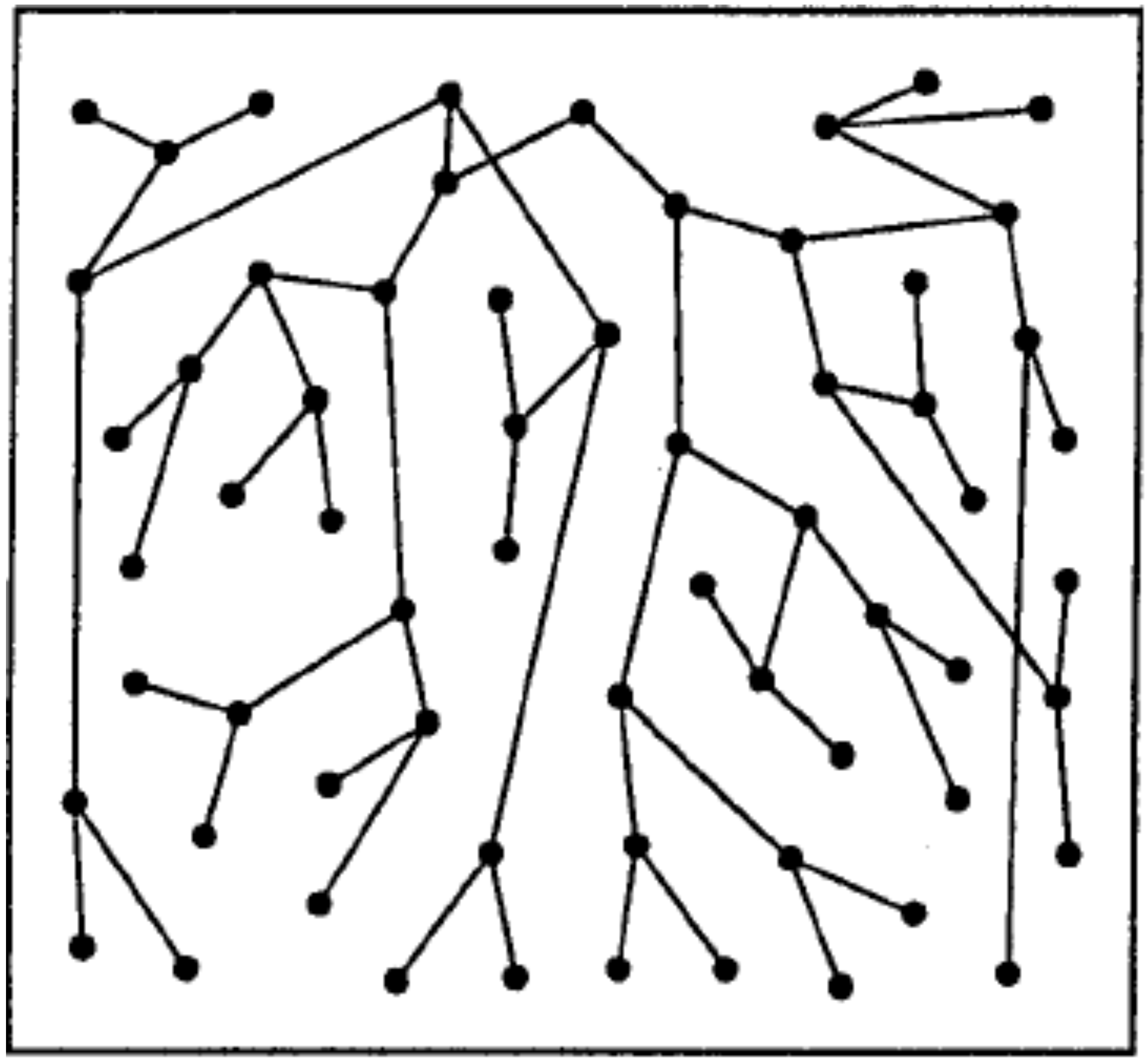


Figure 13

A PRACTICAL APPROACH TO DRAWING UNDIRECTED GRAPHS

D. Tunkeland
1994

TUNKELAND

- Comprimentos de arestas devem ser uniformes
- Nodos não adjacentes deve ficar distantes
- Cruzamentos de arestas devem ser minimizados

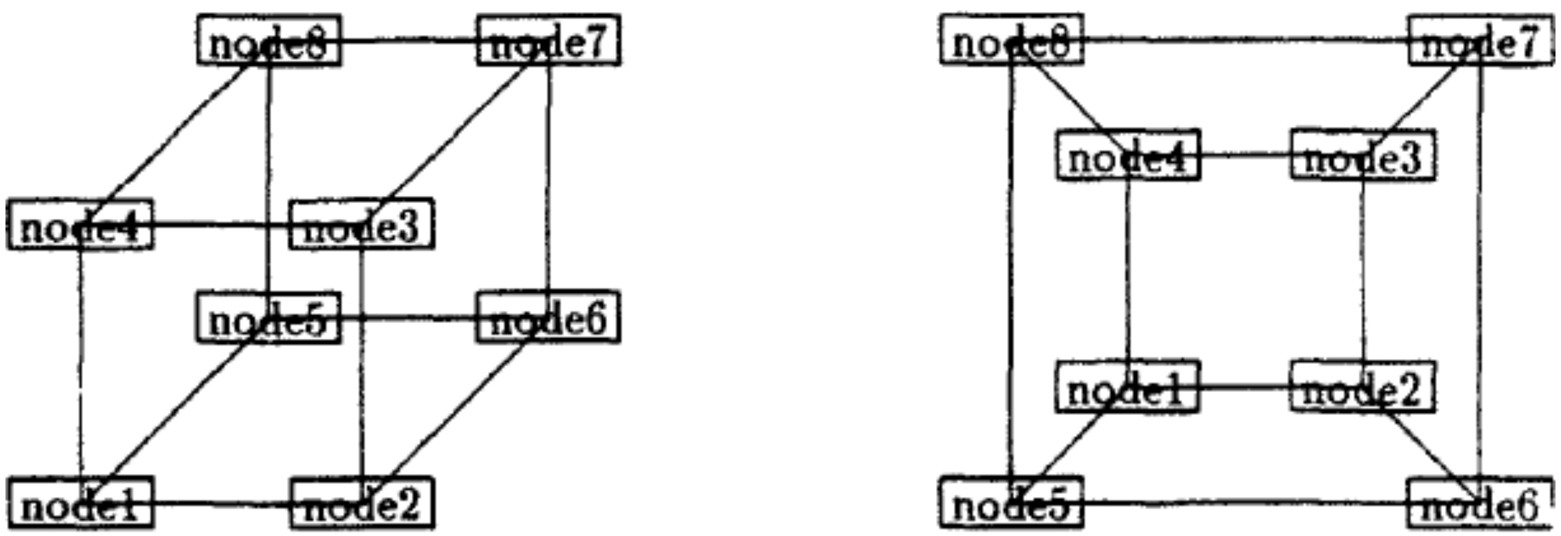


Figure 1

- Estes critérios são, muitas vezes, conflitantes

  (a) Comprimentos de arestas uniformes

  (b) Sem cruzamento de arestas

- Qual critério é mais estético?

CRITÉRIOS ESTÉTICOS X OBJETIVOS

- Circuitos digitais precisam de arestas na horizontal ou vertical
- No desenho de uma hierarquia, membros de um mesmo nível devem ser desenhados como tal
- Em grafos dirigidos, pode haver uma preferência de orientação para as arestas

PROBLEMA

- Os critérios estéticos fornecem diretrizes mas não especificam o peso relativo que deve ser aplicado entre eles

- O cuidado com a estética não é importante apenas pelo apelo visual do resultado, mas porque facilita a compreensão dos fenômenos representados pelo grafo

- Um algoritmo para desenho de grafos de propósito geral deve lidar com os critérios tradicionais de estética mas deve ser capaz de lidar com critérios dependentes de cada contexto

TUNKELAND

1. Leia o grafo (lista de arestas e nós)
2. Compute a ordenação dos nós
3. Para cada nó *N*, na ordem especificada pelo passo 2
   A. Amostre o espaço de desenho para encontrar a posição inicial de *N*
   B. Otimização local de *N* e seus vizinhos
4. Otimização local em todos os nós

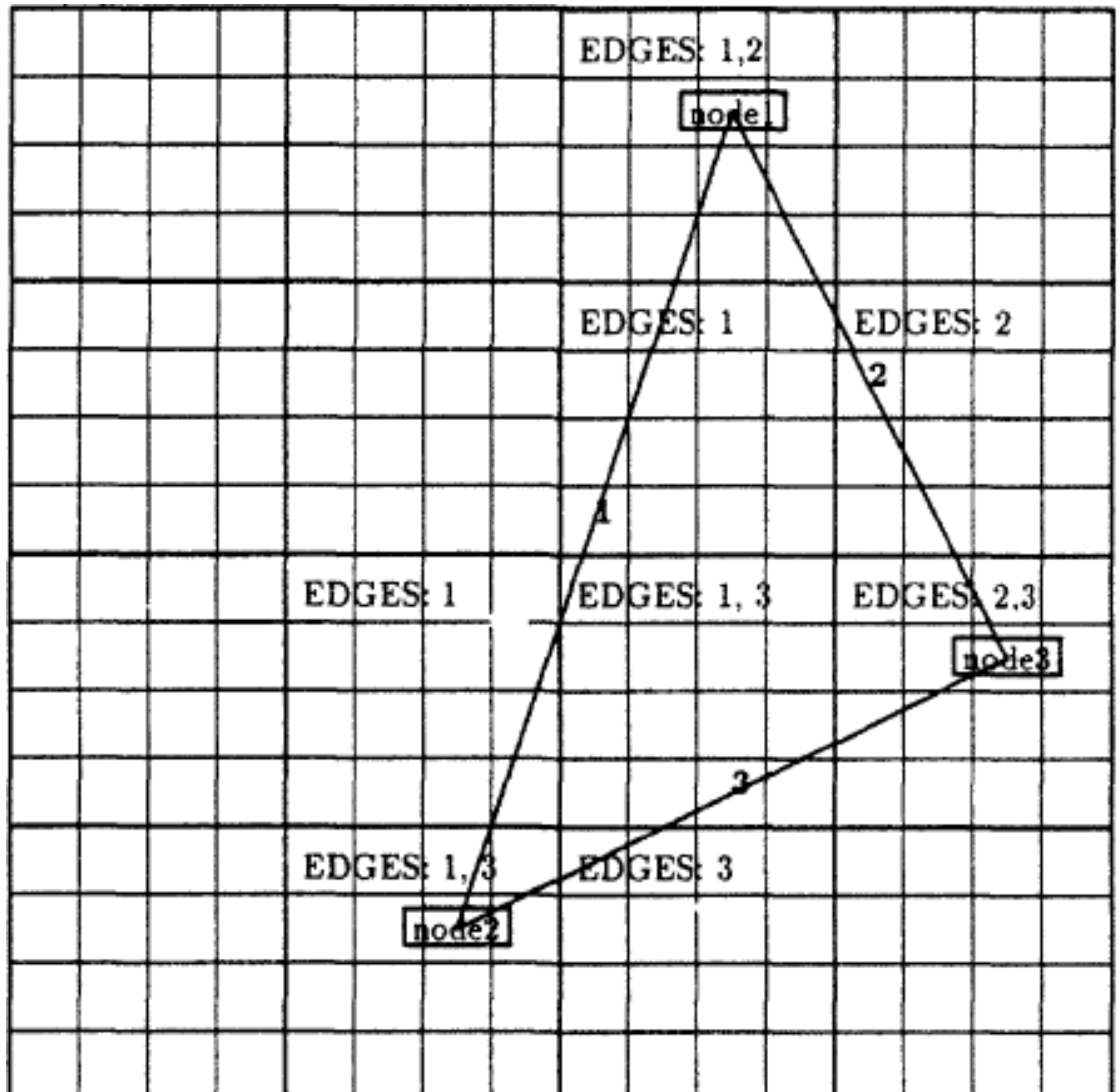
CRITÉRIOS ESTÉTICOS

- Soma de três componentes ponderados, cada um referente a um dos três critérios estéticos considerados
  - **Atração entre nós vizinhos:** $f_a(d) = w_a d^2$
  - **Repulsão entre todos os pares de nós:** $f_r(d) = w_r / d^2$
    - onde d é o comprimento da aresta
  - **Cruzamento de arestas:** $f_c(n) = w_c n$
    - onde n é o número de cruzamentos entre arestas

TUNKELAND

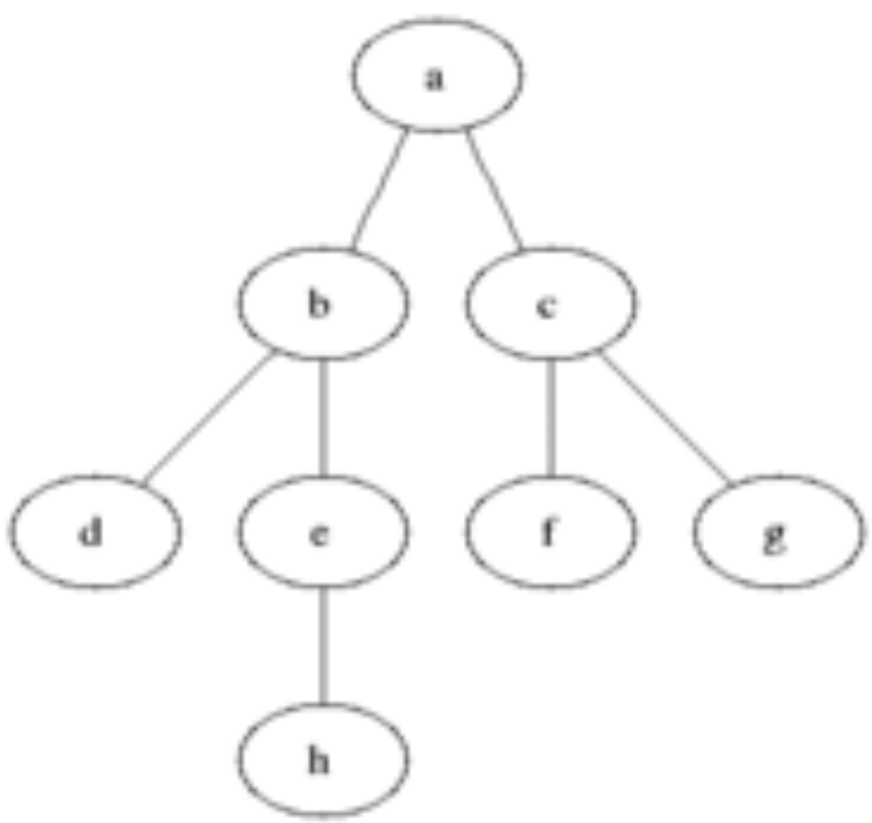
- Usa um grid uniforme
- Cada célula mantém duas listas:

- Nós
- Arestas



ORDEM DA COLOCAÇÃO DOS NÓS

- Inicia-se do centro do grafo
- Percorre o grafo em largura iniciando do nó central



- Nó central é o nó c que minimiza
  - $max_{v \epsilon V}\, d(c,v)$, onde $d(c,v)$ é o comprimento do menor caminho conectando c e v

TENTATIVAS DE POSICIONAMENTO

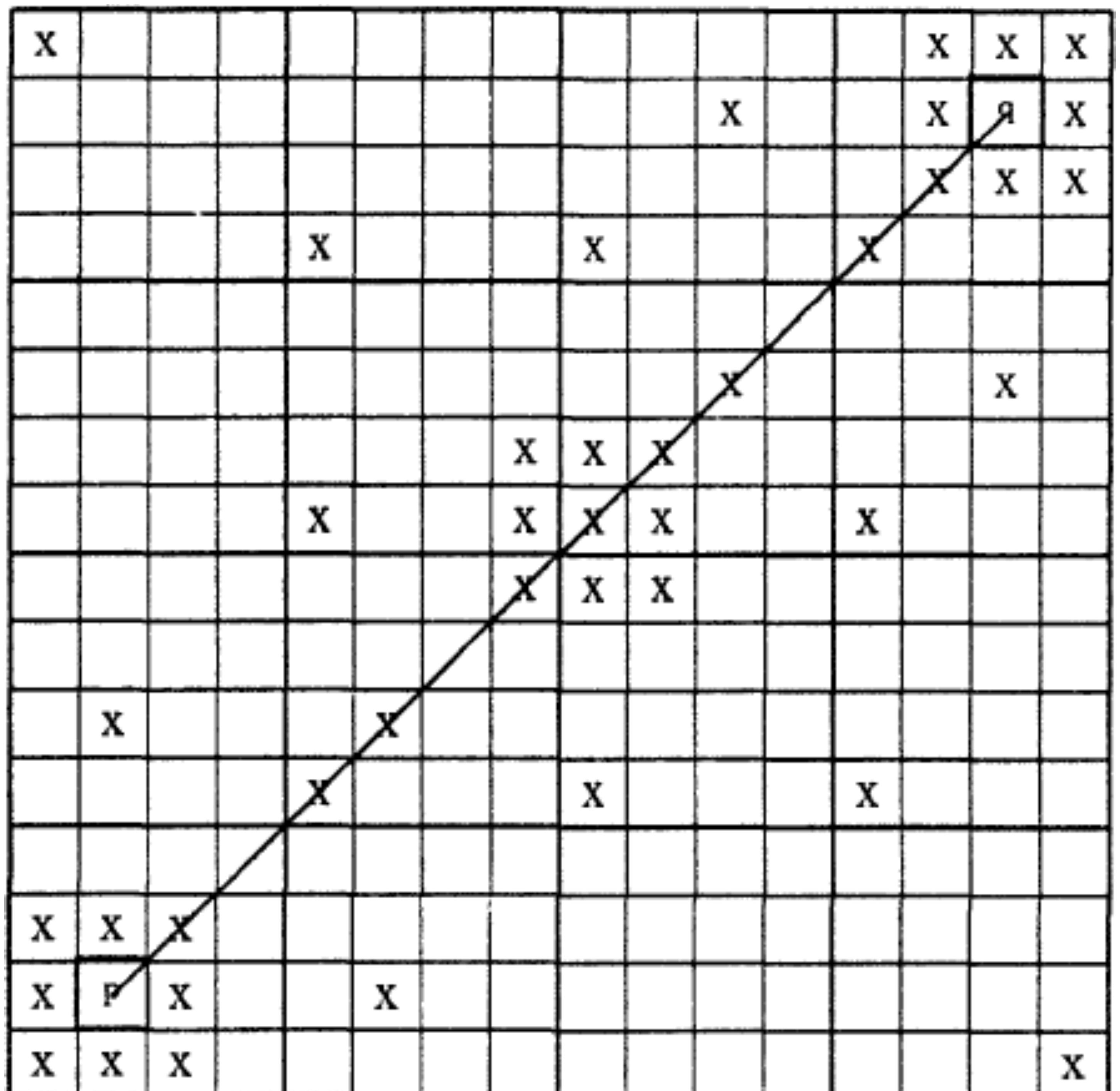


Figure 7

REFINAMENTO

- Calculo da função de custo para cada uma das 16 células vizinhas no mesmo esquema da amostragem da posição inicial

- Se houver melhoria, movimenta o nó e chama este mesmo procedimento recursivamente para todos os nós vizinhos

- Uma melhoria desencadeia uma série de melhorias nos nós vizinhos

MÉTRICAS DE QUALIDADE

- **Q1: variância nos tamanhos das arestas** no grafo normalizado. Melhor grafo teria variância zero ou seja todas as arestas do mesmo tamanho

- **Q2:** soma sobre todos os pares de nós dos **custos de repulsão** no grafo normalizado (proporcional ao quadrado das distâncias entre os nós)

- **Q3:** número de **cruzamentos** de arestas

GRAFO PEQUENO ESPARSO

#	Description	Nodes	Edges
1	Path	16	15
2	Cycle	16	16
3	Complete Binary Tree	15	14
4	$K_{3,3}$	6	9
5	Dodecahedron	20	30
6	Square Mesh	16	24
7	Random Tree	15	14
8	Random Graph	16	20

#	Proposed Algorithm			[FR91]			[DH91]		
	Q_1	Q_2	Q_3	Q_1	Q_2	Q_3	Q_1	Q_2	Q_3
1	0	.0858	0	.0031	.2393	1	.0081	.2079	0
2	.0012	.1124	0	.0009	.1114	0	.0055	.2222	0
3	.0005	.1297	0	.0055	.1561	0	.0073	.3070	3
4	.0082	.5090	1	.0058	.4670	3	.0093	.4981	1
5	.0037	.2643	5	.0009	.2684	6	.0034	.3659	14
6	.0007	.1841	0	.0023	.3416	7	.0066	.2827	5
7	.0001	.1231	0	.0048	.1677	1	.0108	.3126	0
8	.0011	.1629	1	.0031	.1640	1	.0028	.2356	2

GRAFO PEQUENO DENSO

#	Description	Nodes	Edges
9	Wheel	13	24
10	Triangular Mesh	15	30
11	Hypercube	16	32
12	K_6	6	15
13	Icosahedron	12	30
14	K_{12}	12	66
15	Random Graph	16	40
16	Random Graph	16	64

#	Proposed Algorithm			[FR91]			[DH91]		
	Q_1	Q_2	Q_3	Q_1	Q_2	Q_3	Q_1	Q_2	Q_3
9	.0036	.2925	0	.0036	.3431	5	.0064	.3327	0
10	.0001	.2015	0	.0006	.2029	0	.0045	.3505	12
11	.0089	.5125	10	.0001	.4670	24	.0072	.4798	35
12	.0119	.8257	3	.0051	.5418	15	.0130	.9182	4
13	.0143	.5481	6	.0021	.4151	23	.0060	.4765	20
14	.0040	1.461	185	.0021	.7418	391	.0041	1.503	214
15	.0061	.6249	22	.0017	.3810	36	.0062	.6187	42
16	.0042	.8606	117	.0017	.5445	160	.0039	.8617	164

GRAFO GRANDE ESPARSO

#	Description	Nodes	Edges
17	Path	48	47
18	Cycle	48	48
19	Complete Binary Tree	63	62
20	Fibonacci Tree	54	53
21	Hexagonal Mesh	54	72
22	Square Mesh	49	84
23	Random Tree	63	62
24	Random Graph	60	80

#	Proposed Algorithm			[FR91]			[DH91]		
	Q_1	Q_2	Q_3	Q_1	Q_2	Q_3	Q_1	Q_2	Q_3
17	0	.0329	0	.0016	.1672	6	.0071	.3474	56
18	.0021	.0474	0	.0021	.1719	10	.0068	.3983	41
19	.0005	.0677	0	.0018	.1316	18	.0071	.3395	80
20	.0005	.0661	0	.0025	.1743	12	.0049	.3465	71
21	.0008	.0939	5	.0011	.2170	26	.0060	.4599	165
22	.0004	.0921	0	.0009	.1622	22	.0039	.3591	130
23	.0009	.0638	0	.0024	.1410	13	.0065	.4124	87
24	.0029	.1642	20	.0011	.2095	44	.0030	.3842	135

GRAFO GRANDE DENSO

#	Description	Nodes	Edges
25	Wheel	61	120
26	Torus	64	128
27	Triangular Mesh	55	135
28	Random Graph	60	120
29	Random Graph	60	150
30	Random Graph	60	180

#	Proposed Algorithm			[FR91]		
	Q_1	Q_2	Q_3	Q_1	Q_2	Q_3
25	.0053	.4415	57	.0020	.4270	143
26	.0037	.2591	116	.0007	.2818	168
27	.0000	.0906	0	.0005	.1400	41
28	.0033	.4707	163	.0009	.3333	291
29	.0038	.7621	404	.0006	.4216	495
30	.0029	.8944	828	.0007	.5341	992

FORCEATLAS2, A CONTINUOUS GRAPH LAYOUT ALGORITHM FOR HANDY NETWORK VISUALIZATION DESIGNED FOR THE GEPHI SOFTWARE

Mathieu Jacomy,, et al., PloS One, 2014

FORCEATLAS2

- Foi desenvolvido para implementação no pacote Gephi
  - Uso típico dos usuários: redes livres de escala com 10 a 10.000 nós
- Leiaute de força
- Foi desenvolvido para a experiência do usuário e para tal seu algoritmo é contínuo
- Também é um algoritmo que permite a disposição geográfica do grafo

FORCEATLAS2

Energy Model

Every force-directed algorithm relies on a certain formula for the attraction force and a certain formula for the repulsion force. The "spring-electric" layout [16] is a simulation inspired by real life. It uses the repulsion formula of electrically charged particles ($F_r = k/d^2$) and the attraction formula of springs ($F_a = -k \cdot d$) involving the geometric distance d between two nodes. Fruchterman and Rheingold [17] created an efficient algorithm using custom forces (attraction $F_a = d^2/k$ and repulsion $F_r = -k^2/d$, with k adjusting the scaling of the network). Note that actually, non-realistic forces have been used since the beginning, noticeably by Eades [16] in his pioneer algorithm. Fruchterman and Rheingold were inspired by Eades' work, and they noticed that despite using the spring metaphor to explain his algorithm, the attraction force is not that of a spring.

Sixteen years later, Noack [11] explained that the most important difference among force-directed algorithms is the role played by distance in graph spatialization. In physical systems, forces depend on the distance between the interacting entities:

FORCEATLAS2 - FORÇA DE ATRAÇÃO "CLÁSSICA"

A classical attraction force. The attraction force F_a between two connected nodes n_1 and n_2 is nothing remarkable. It depends linearly on the distance $d(n_1,n_2)$. We will explain later why there is no constant adjusting of this force.

$$F_a(n_1,n_2) = d(n_1,n_2) \tag{1}$$

FORCEATLAS2 – FORÇA DE REPULSÃO PELO GRAU

$$F_r(n_1,n_2) = k_r \frac{(deg(n_1)+1)(deg(n_2)+1)}{d(n_1,n_2)}$$

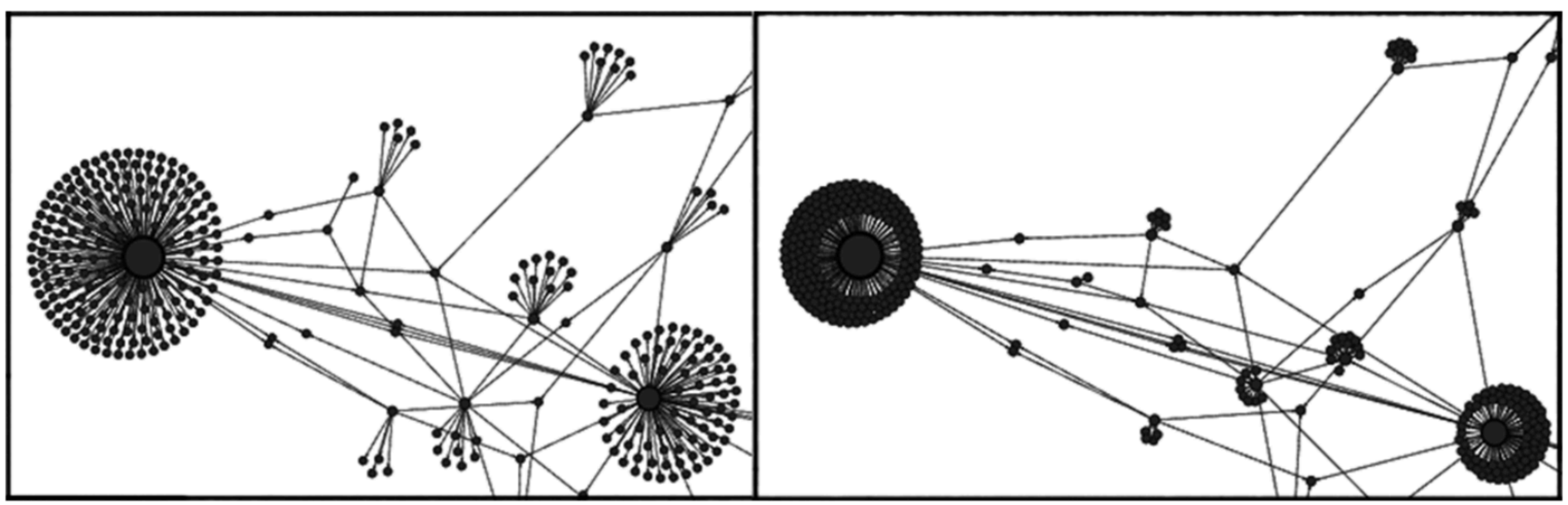


Figure 2. Regular repulsion vs. repulsion by degree. Fruchterman-Rheingold layout on the left (regular repulsion) and ForceAtlas2 on the right (repulsion by degree). While the global scheme remains, poorly connected nodes are closer to highly connected nodes. ($a-r=1$).
doi:10.1371/journal.pone.0098679.g002

FORCEATLAS2 – CONFIGURAÇÕES – LINLOG

LinLog mode. Andreas Noack produced an excellent work on placement quality measures [18]. His LinLog energy model arguably provides the most readable placements, since it results in a placement that corresponds to Newman's modularity [14], a widely used measure of community structure. The LinLog mode just uses a logarithmic attraction force.

$$F_a(n_1,n_2) = log(1 + d(n_1,n_2)) \tag{3}$$

FORCEATLAS2 – CONFIGURAÇÕES – FORÇA DA GRAVIDADE

Gravity. Gravity is a common improvement of force-directed layouts. This force $F_g(n)$ prevents disconnected components (islands) from drifting away, as pictured in Figure 3. It attracts nodes to the center of the spatialization space. Its main purpose is to compensate repulsion for nodes that are far away from the center. In our case it needs to be weighted like the repulsion:

$$F_g(n) = k_g(deg(n)+1) \tag{4}$$

k_g is set by the user.

The "Strong gravity" option sets a force that attracts the nodes that are distant from the center more ($d(n)$ is this distance). This force has the drawback of being so strong that it is sometimes stronger than the other forces. It may result in a biased placement of the nodes. However, its advantage is to force a very compact layout, which may be useful for certain purposes.

$$F'_g(n) = k_g(deg(n)+1)d(n) \tag{5}$$

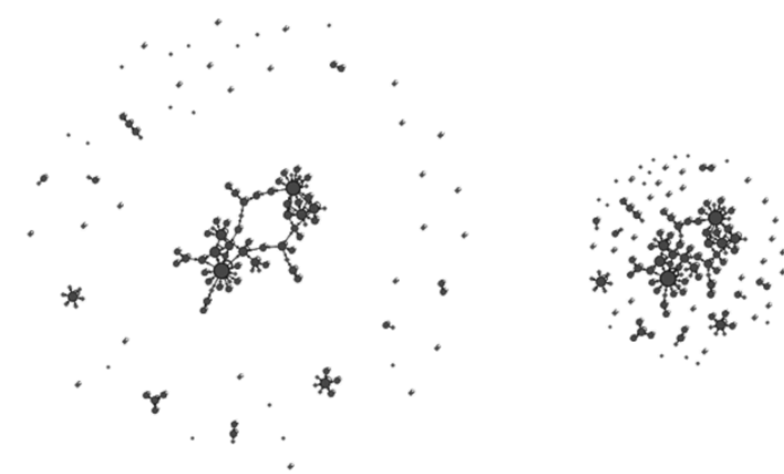


Figure 3. Effects of the gravity. ForceAtlas2 with gravity at 2 and 5. Gravity brings disconnected components closer to the center (and slightly affects the shape of the components as a side-effect).
doi:10.1371/journal.pone.0098679.g003

FORCEATLAS2 - CONFIGURAÇÕES - ESCALA

Scaling. A force-directed layout may contain a couple of constants k_a and k_r playing an opposite role in the spatialization of the graph. The attraction constant k_a adjusts the attraction force, and k_r the repulsion force. Increasing k_a reduces the size of the graph while increasing k_r expands it. In the first version of ForceAtlas, the user could modify the value of both variables. For practical purposes, however, it is better to have only one single scaling parameter. In ForceAtlas2, the scaling is k_r while there is no k_a. The higher k_r, the larger the graph will be, as you can see in Figure 4.

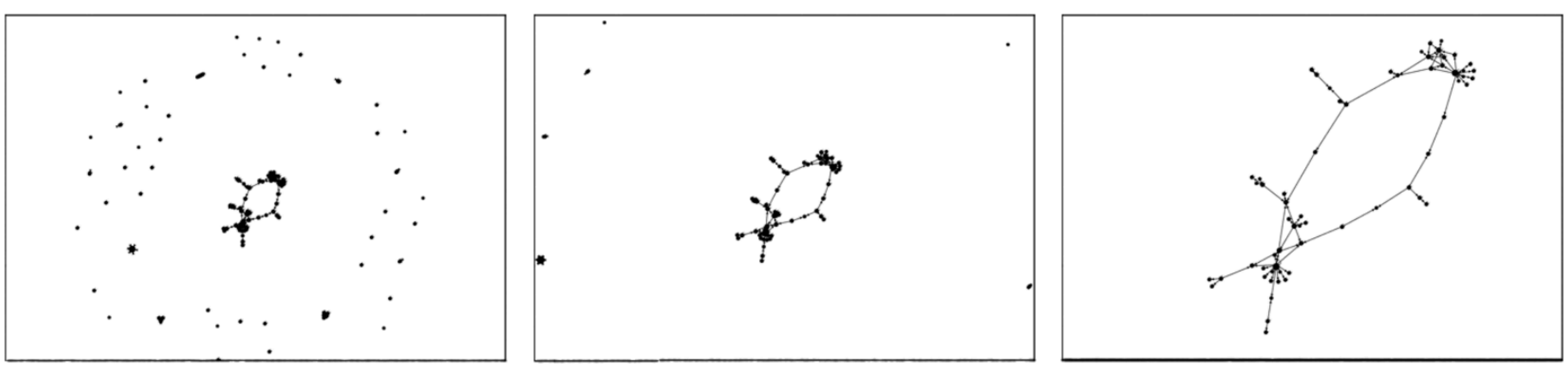


Figure 4. Effects of the scaling. ForceAtlas2 with scaling at 1, 2 and 10. The whole graph expands as scaling affects the distance between components as well as their size. Note that the size of the nodes remains the same; scaling is not zooming.
doi:10.1371/journal.pone.0098679.g004

FORCEATLAS2 – CONFIGURAÇÕES – PESOS DAS ARESTAS

Edge weight. If the edges are weighted, this weight will be taken into consideration in the computation of the attraction force. This can have a dramatic impact on the result, as pictured in Figure 5. If the setting "Edge Weight Influence" δ is set to 0, the weights are ignored. If it is set to 1, then the attraction is proportional to the weight. Values above 1 emphasize the weight effects. This parameter is used to modify the attraction force according to the weight $w(e)$ of the edge e:

$$F_a = w(e)^{\delta} d(n_1, n_2) \tag{6}$$

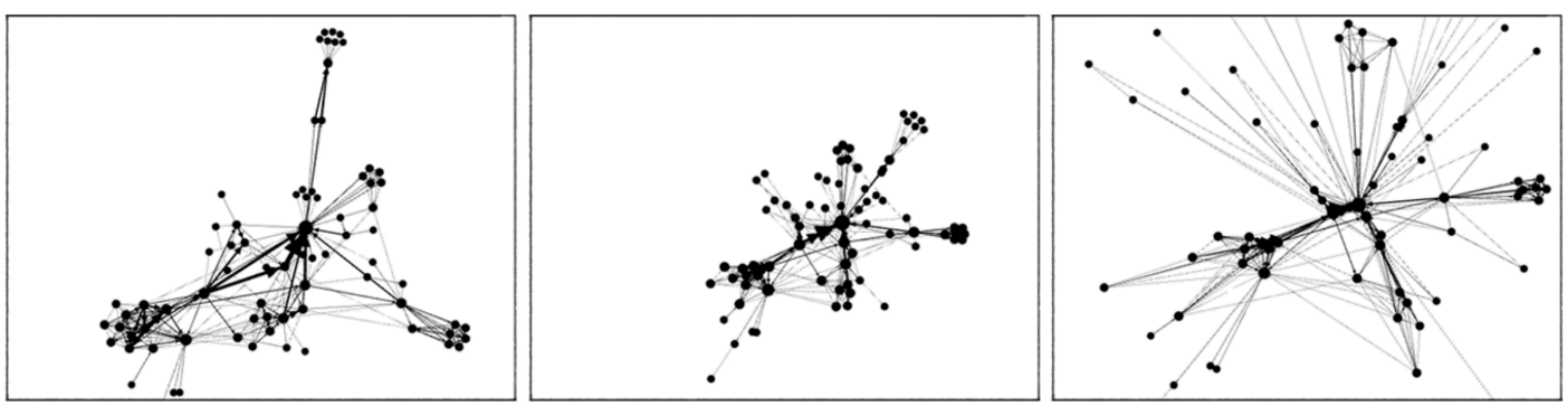


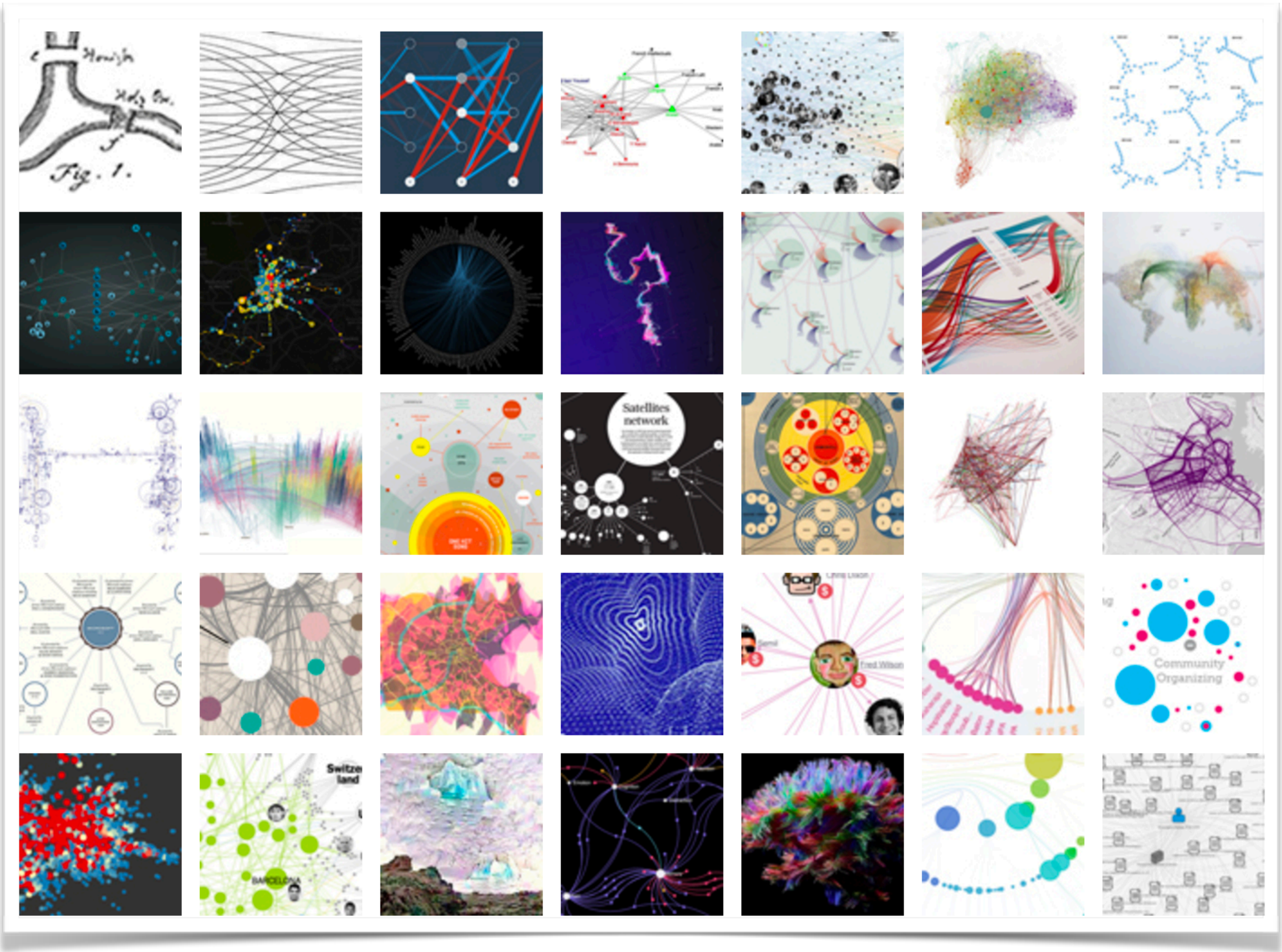
Figure 5. Effects of the edge weight influence. ForceAtlas2 with Edge Weight Influence at 0, 1 and 2 on a graph with weighted edges. It has a strong impact on the shape of the network.
doi:10.1371/journal.pone.0098679.g005

FORCEATLAS2 - CONFIGURAÇÕES - SOBREPOSIÇÃO

- Há ainda um parâmetro para evitar a sobreposição de nós

VISUAL COMPLEXITY

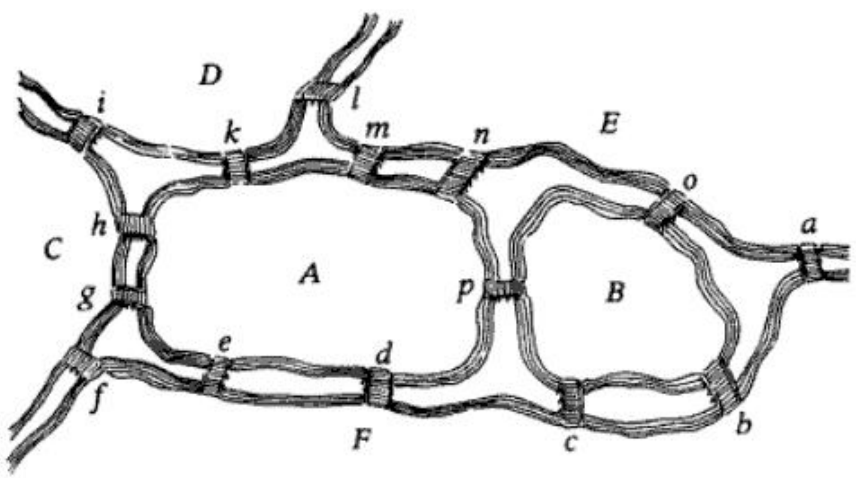
Manuel Lima



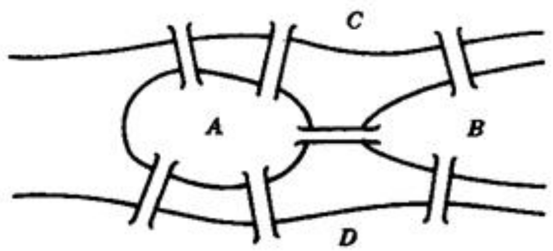
The birth of network science

Although human beings have previously envisioned models of network-like structures, the first documented mathematical analysis of the process occurred in 1736 by Leonhard Euler

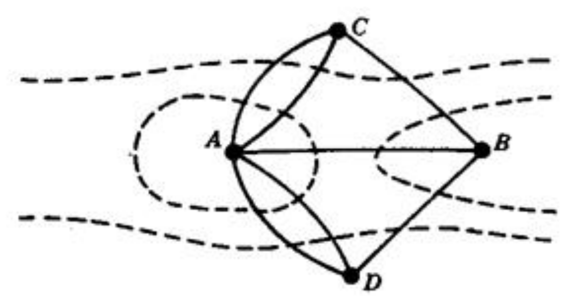
In his article on the Königsberg bridges, Euler considered another bridge problem, which is illustrated below.*



Two islands, A and B, are surrounded by water that leads to four rivers. Fifteen bridges cross the rivers and the water surrounding the islands. Is it possible to make a trip that crosses each bridge exactly once?



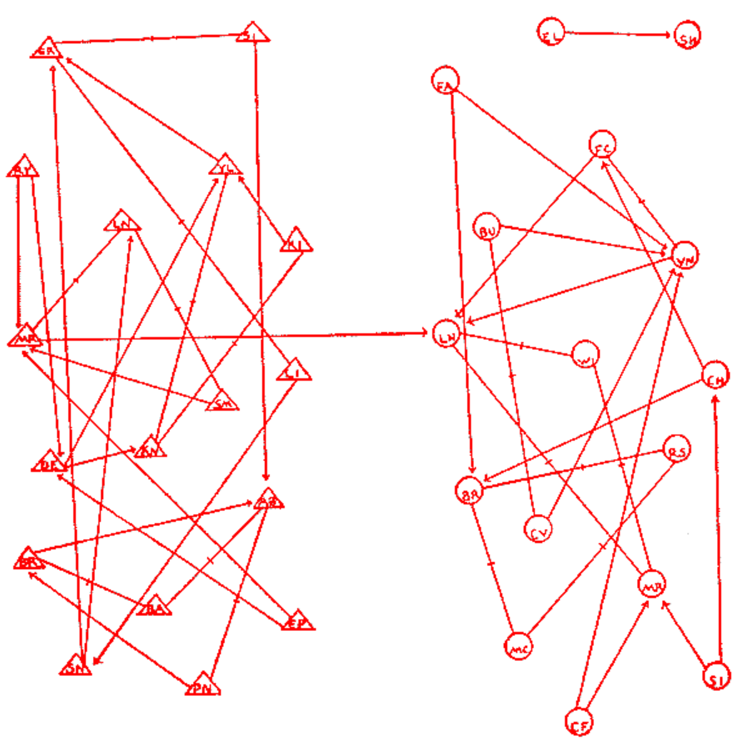
(a) Königsberg in 1736



(b) Euler's graphical representation

Psycological geography

Jacob Moreno created the *sociogram*: a graphic representation of social ties between a group of boys and girls from one elementary school, making the beginning of sociometry, which later came to be know as social network analysis



Two independent groups, boys (triangles) and girls (circles), and links within each group and between groups

New York Times, April 3, 1933

Purpose of network visualization

So what are the purposes of network visualization?

Purpose of network visualization

So what are the purposes of network visualization?

A potential visual decoder of complexity

Purpose of network visualization

Commonly driven by five key functions:

- document,
- clarify,
- reveal,
- expand,
- and abstract

Purpose of network visualization

Document

- To map a system that has **never been depicted** before
- To portray a new **unfamiliar** territory
- To **stimulate interest** and awareness of a subject matter
- To **open the door** for further discoveries and interpretations

Purpose of network visualization

Clarify

- To make the system more **understandable**, intelligible, transparent
- **Simplification**
- To **explain** important aspects and clarify given areas of the system
- To **communicate** in a simple, effective way becoming means for information processing and understanding

Purpose of network visualization

Reveal

- To find a **hidden pattern** in or to explicit new insights
- Revealing should concentrate on **causality** by leading the disclosure of unidentified relationships and correlations
- To **check** initial **assumptions** and central questions

Purpose of network visualization

Expand

To serve as a vehicle for **other uses** and set the stage for further exploration

- Portrayal of **multidimensional** behaviours

- Underlying layer of **additional visualizations**

Purpose of network visualization

Abstract

- To explore the networked schema as a platform for abstract representation

- Network visualization can be a vehicle for hypothetical and metaphorical expression, depicting intangible concepts that might not even rely on an existing data set

Principles of network visualization

Manuel Lima lists eight principles of network visualization

- The first four are larger universal considerations that, due to their broad assessment, can be applied in a variety of graphical representations

- The subsequent four encompass detailed principles, tackling explicit challenges in the depiction of networks

Principles of network visualization

1 - Start with a question

Every project should start with an inquiry that leads to further insights about the system

Principles of network visualization

1 - Start with a question

Every project should start with an inquiry that leads to further insights about the system

and perhaps answer questions that were not originally asked

Principles of network visualization

2 - Look for relevancy

Human cognition is relevance oriented

Something is **relevant** if it serves as an effective means to a particular purpose or, more specifically, if it increases the likelihood of achieving an underlying goal

Principles of network visualization

2 - Look for relevancy

In the context of visualization, relevancy comes into place when selecting two central elements:

- Supporting data set (content)
- Visualization techniques (method)

Principles of network visualization

3 – Enable multivariate analysis

In many cases the depiction of networks is seen as a binary system, where connections are simply turned on and off

Ties among elements in a network are rich and detailed, and the inclusion of additional information can be fundamental

Principles of network visualization

3 – Enable multivariate analysis

The exploration and analysis of large multivariate networks is still a challenge

Current methods are focused on either the **structural aspect** of the multi-variated network or the **multidimensional data** attached to the nodes and links

Van den Elzen, Stef and van Wijk, Jarke J.. "Multivariate network exploration and presentation: from detail to overview via selections and aggregations." IEEE Transactions of Visualization and Computer Graphics (2014): 20(12): 2310-2319.

Principles of network visualization

3 – Enable multivariate analysis

DOSA (Detail to Overview via Selections and Aggregations):

- An exploration method that enables users to explore and analyze both network structure and multivariate data associated with the nodes and links simultaneously

Van den Elzen, Stef and van Wijk, Jarke J.. "Multivariate network exploration and presentation: from detail to overview via selections and aggregations." IEEE Transactions of Visualization and Computer Graphics (2014): 20(12): 2310-2319.

Principles of network visualization

3 – Enable multivariate analysis

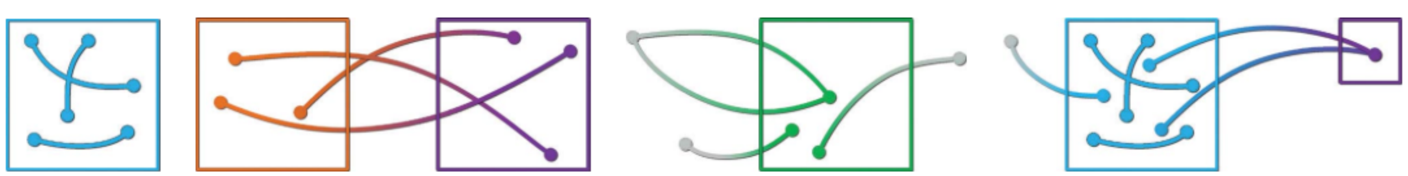
- Intuitive creation and modification of selections of interest, and
- A juxtaposed detail and high-level overview, for
- Production of high-level, infographic style overviews, focusing on the non-expert users

Van den Elzen, Stef and van Wijk, Jarke J.. "Multivariate network exploration and presentation: from detail to overview via selections and aggregations." IEEE Transactions of Visualization and Computer Graphics (2014): 20(12): 2310-2319.

Principles of network visualization

3 – Enable multivariate analysis

Fig. 1. Multivariate network exploration using selections of interest, detail view (left) and high-level infographic-style overview (right).



(a) Within. (b) Between. (c) Background. (d) Combined.

Fig. 4. Different types of edges involved in a node selection. a) Within edges showing all internal connections of a selection; both source and target node are contained in the selection. b) Between edges show all connections between two selections; both source and target node are contained in different selections. c) Background edges show all connections from a selection to the background selection. d) Combined, showing all involved edges for a selection, within, between and background.

Van den Elzen, Stef and van Wijk, Jarke J.. "Multivariate network exploration and presentation: from detail to overview via selections and aggregations." IEEE Transactions of Visualization and Computer Graphics (2014): 20(12): 2310-2319.

Principles of network visualization

3 – Enable multivariate analysis

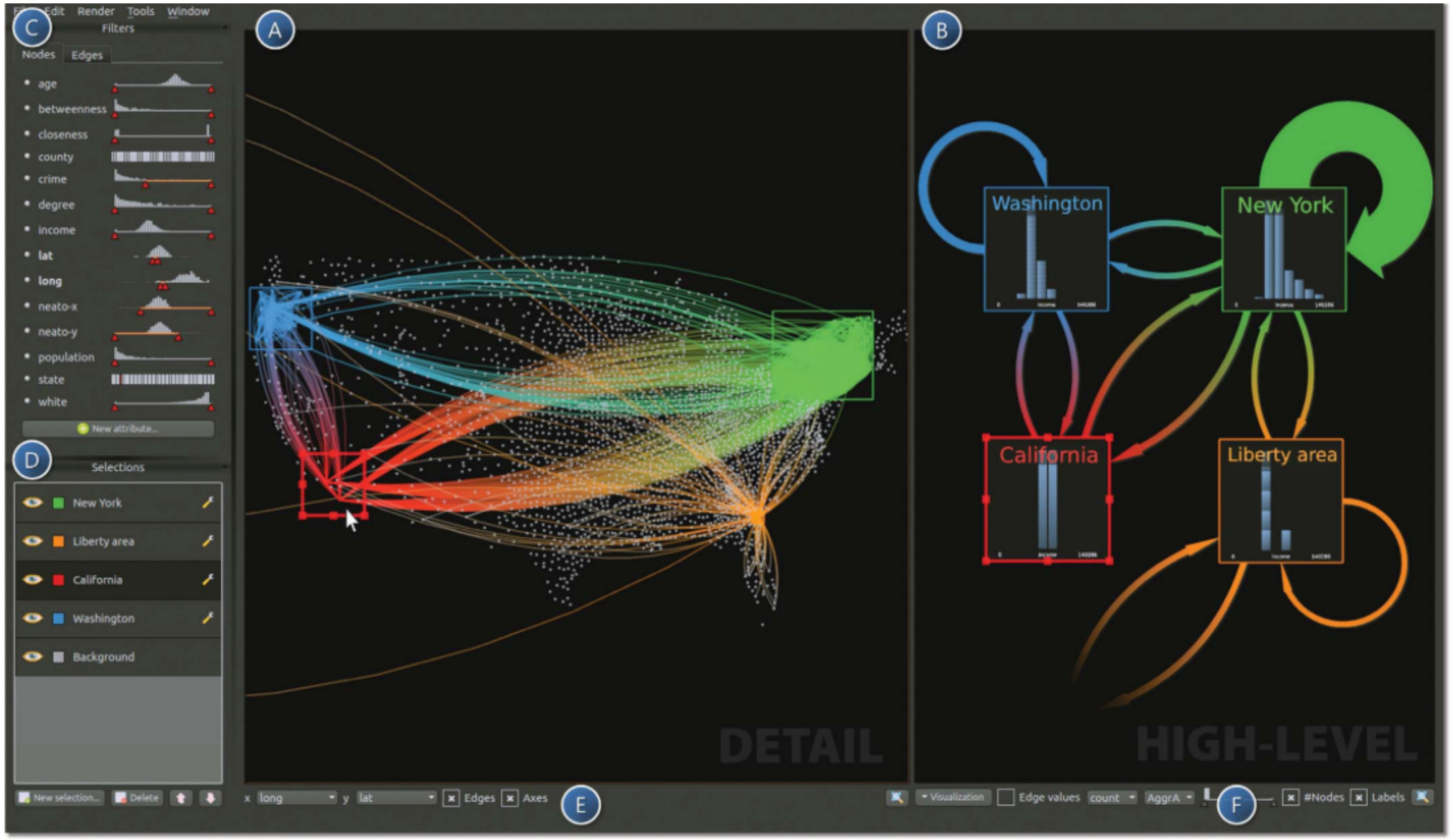


Fig. 3. Graphical user interface of the implemented prototype showing all coherent components: a) Low-level detail view showing a two-dimensional projection of the nodes based on available attributes. The projection and other visual attributes can be set using controls at the bottom (e). Four selections of interest are shown in the detail view, visualized using boxes for direct manipulation. All selections of interest show *between* edges, the green, blue, and, orange selections also show *within* edges. The orange selection additionally shows edges with the *background* selection. b) High-level overview showing aggregations of the selections of interest including associated aggregated edges. For each of the selections an interactive histogram visualization is shown. Visual representation and attribute mapping are configurable to users needs with controls at the bottom (f). c) Attribute component showing all available attributes with according Scented Widgets for the nodes and links in different tab-pages. The Scented Widgets provide information on the distribution of attributes and can be used to directly control the ranges of the multidimensional selections of interest. d) Selection component containing a list of all selections. Selection priority (order) is controlled via drag and drop operations. Additionally, selections can be hidden or locked here.

Van den Elzen, Stef and van Wijk, Jarke J.. "Multivariate network exploration and presentation: from detail to overview via selections and aggregations." IEEE Transactions of Visualization and Computer Graphics (2014): 20(12): 2310-2319.

Principles of network visualization

3 – Enable multivariate analysis

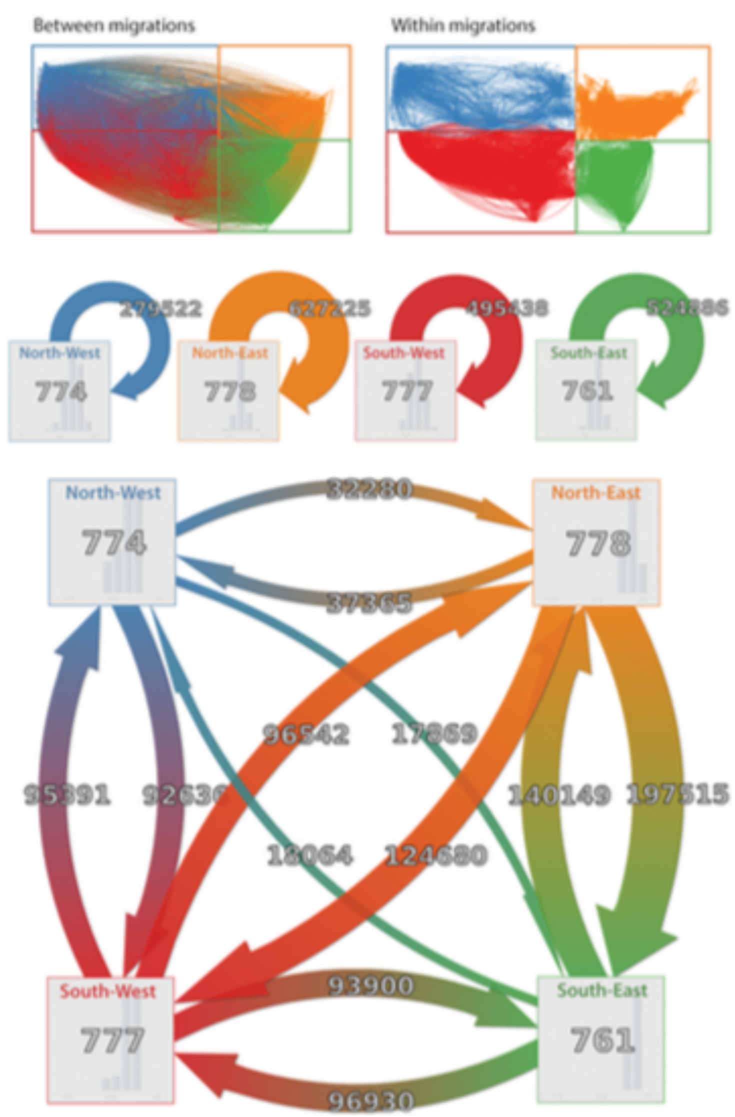


Fig. 7. United States migration data exploration testing for predominantly inbound or outbound regions with detail (top) and overview (bottom). Number of counties in each selection, shown on the boxes, is approximately equal to achieve fair comparisons. The North-East region is outbound with migrations mainly going to the South-East and South-West regions. The rest of the migrations are fairly balanced. The North-West has the least internal migrations and North-East the most (middle).

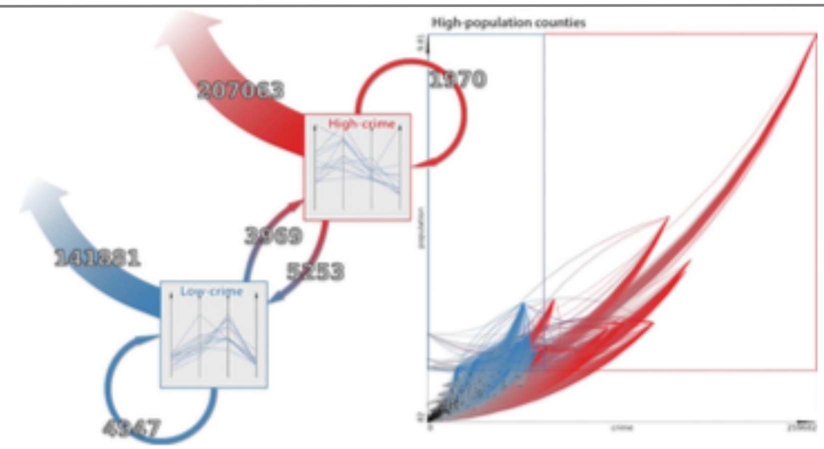


Fig. 8. Migration of highly populated low and high crime regions.

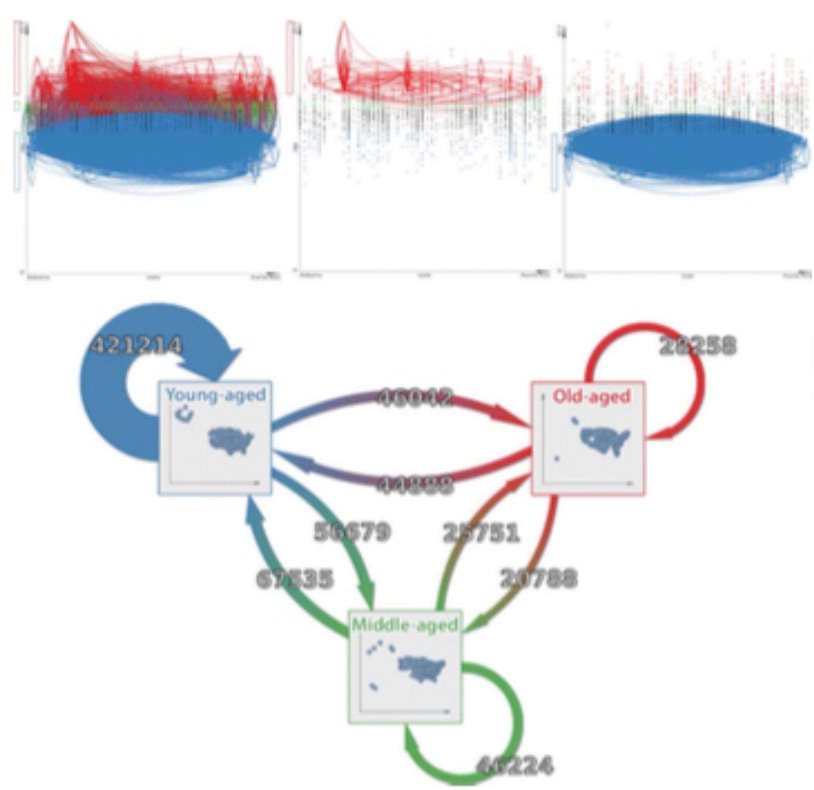


Fig. 9. Testing for correlation between age and migration.

Van den Elzen, Stef and van Wijk, Jarke J.. "Multivariate network exploration and presentation: from detail to overview via selections and aggregations." IEEE Transactions of Visualization and Computer Graphics (2014): 20(12): 2310-2319.

Principles of network visualization

3 – Enable multivariate analysis

Fig. 10. Network path exploration, finding indirect routes from New York to Washington. Aggregated visual representations show number of nodes (counties) contained in them. Aggregated edges show number of edges. From New York there are 102 possible routes to 17 counties in the Florida region, from these 17 counties another 48 routes lead to Washington. From top to bottom we display: (top) Direct links (migrations) between New York and Washington, (middle) counties in New York that have exactly distance two to the counties selected for Washington, and (bottom) all outgoing links from the counties in the New York selection and a selection in the Florida region containing counties that connect New York with Washington.

Van den Elzen, Stef and van Wijk, Jarke J.. "Multivariate network exploration and presentation: from detail to overview via selections and aggregations." IEEE Transactions of Visualization and Computer Graphics (2014): 20(12): 2310-2319.

Principles of network visualization

3 – Enable multivariate analysis

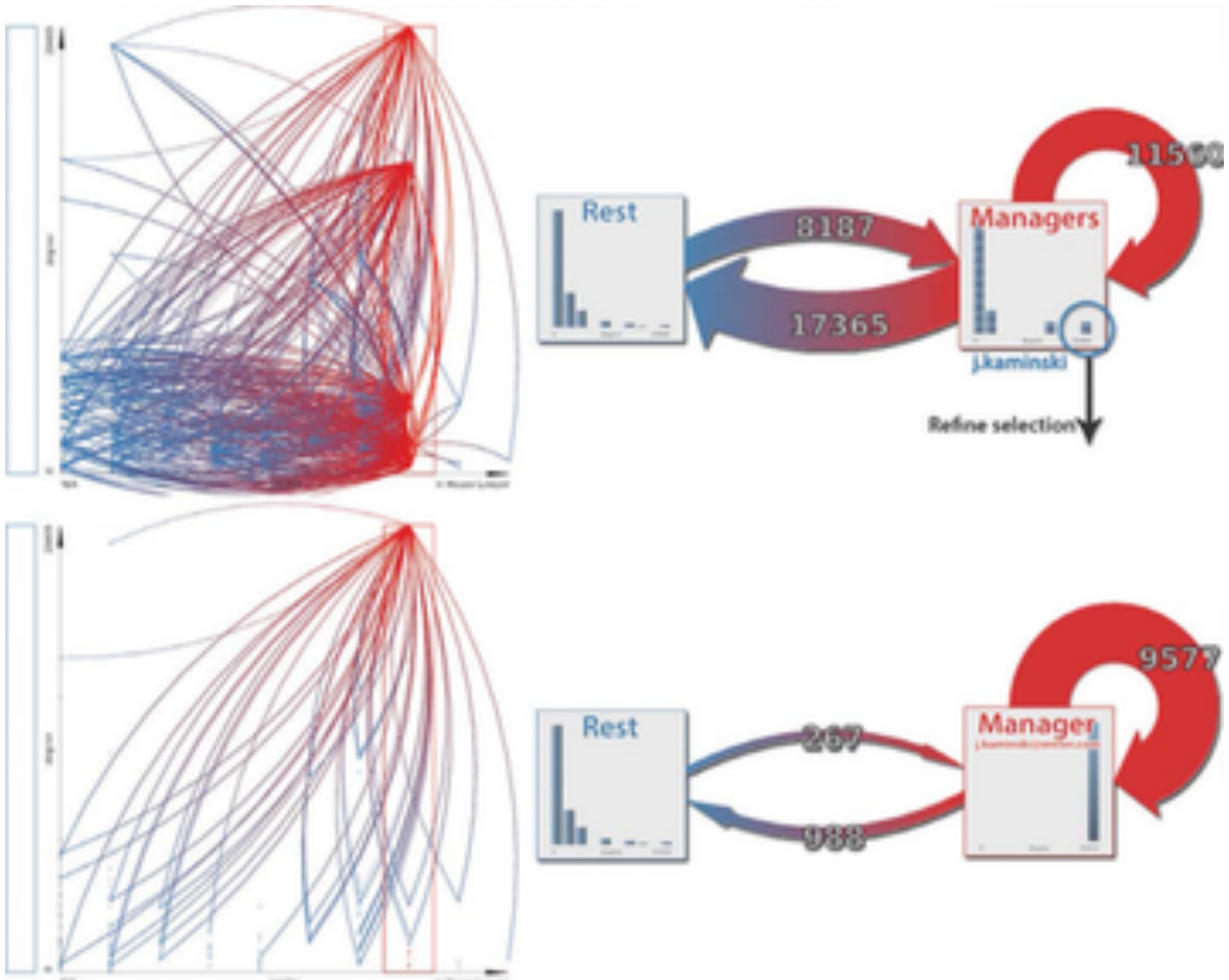


Fig. 11. Enron email communication exploration using two selections of interest: one representing the managers, the other the rest of the employees. Managers stand out due to a large self-loop (11,560 emails). After refinement the cause appears to be a single manager emailing himself all the time (9,577 emails).

Van den Elzen, Stef and van Wijk, Jarke J.. "Multivariate network exploration and presentation: from detail to overview via selections and aggregations." IEEE Transactions of Visualization and Computer Graphics (2014): 20(12): 2310-2319.

Principles of network visualization

3 – Enable multivariate analysis

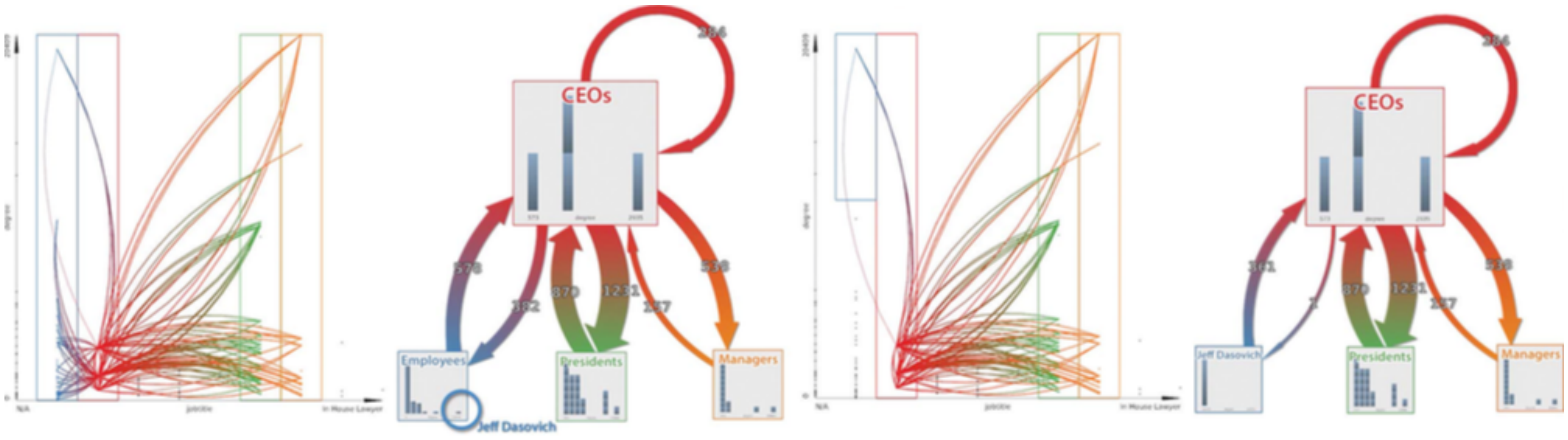


Fig. 12. Typical C-level communication: CEOs are heavily communicating with Vice presidents and managers. However, also communication is present between CEOs and regular employees, which turns out is only one person heavily broadcasting to the CEOs (right).

Van den Elzen, Stef and van Wijk, Jarke J.. "Multivariate network exploration and presentation: from detail to overview via selections and aggregations." IEEE Transactions of Visualization and Computer Graphics (2014): 20(12): 2310-2319.

Principles of network visualization

4 – Embrace time

Time is one of the hardest variables to map in any complex system

It is also one of the richest

Principles of network visualization

4 – Embrace time

This is an economics class composed of 11th and 12th graders at a high school. On this day, economics has two teachers. The first is the usual teacher and the second is a businessman who donates his time a couple days a month to assist the class with their lesson on stock investments. After a minute of undefined class time, the two teachers prescribe collaborative group work and assist students in conducting it. The students are assigned groups within which they are to study the stock market and make mock investments they can follow over time.

Principles of network visualization

4 – Embrace time

Principles of network visualization

4 – Embrace time

Principles of network visualization

4 – Embrace time

Principles of network visualization

4 – Embrace time

Principles of network visualization

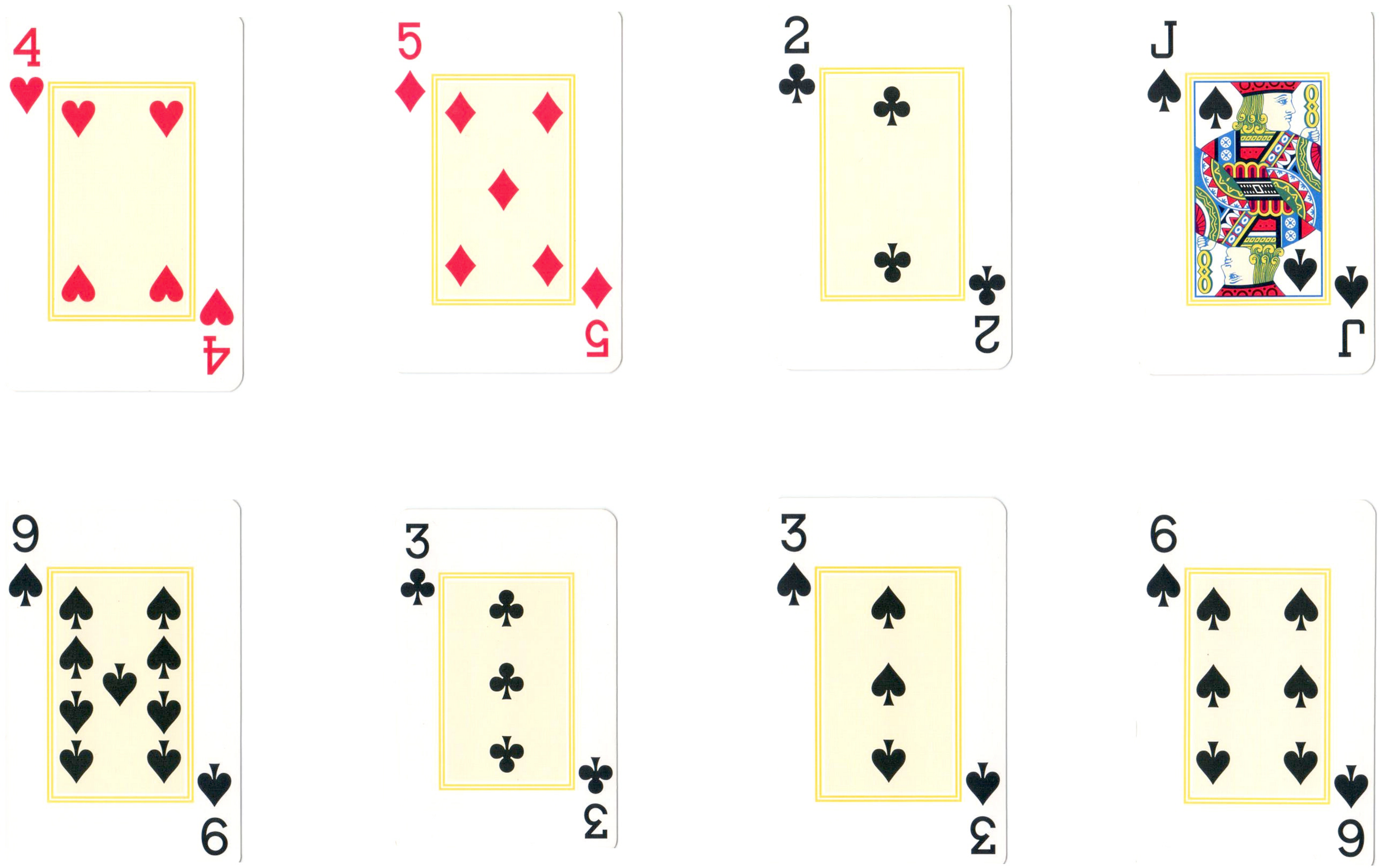
4 – Embrace time

Principles of network visualization

4 – Embrace time

Principles of network visualization

4 – Embrace time



Principles of network visualization

4 – Embrace time

Principles of network visualization

5 – Enrich your vocabulary

Whenever considering the representation of a network, there are two vital elements to consider

Principles of network visualization

5 – Enrich your vocabulary

Whenever considering the representation of a network, there are two vital elements to consider

- Nodes
- Edges

Principles of network visualization

5 – Enrich your vocabulary

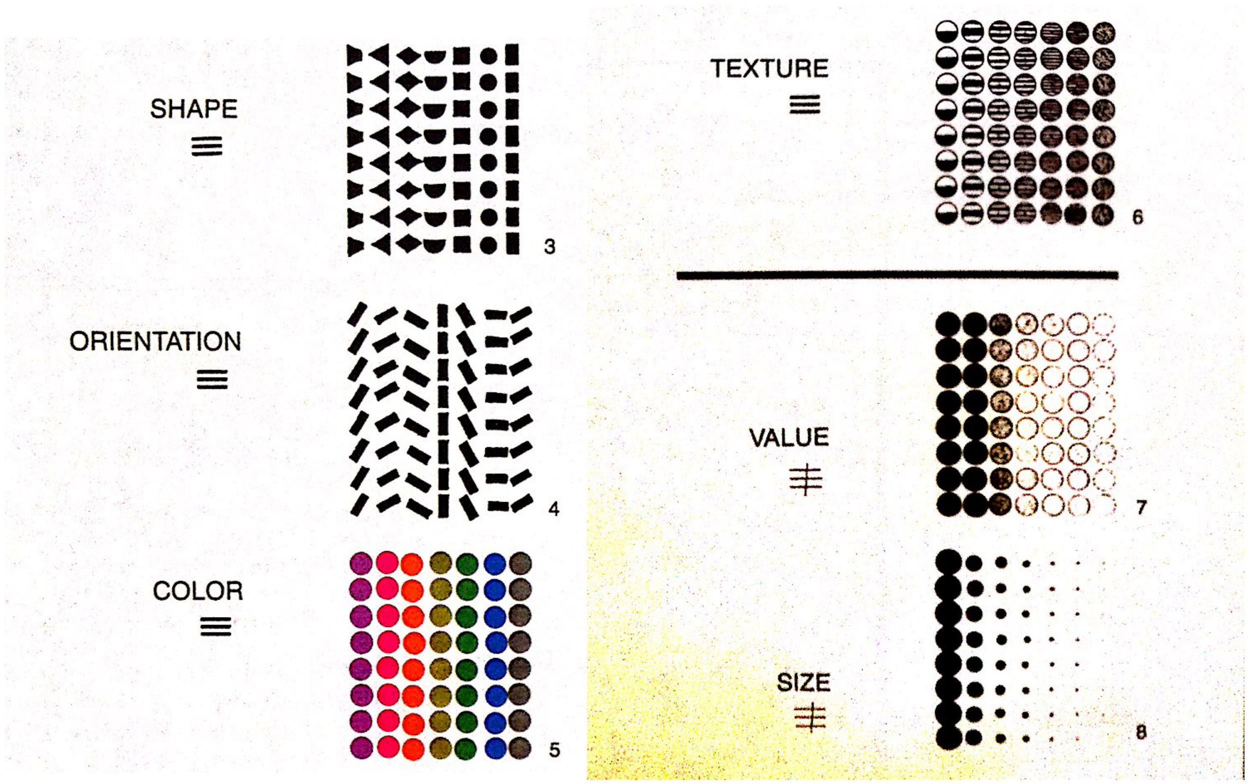
While the recipe is simple enough these two essential ingredients are rarely used to their fullest potential

They are represented by mere circles or squares and undistinguishable connecting lines

Principles of network visualization

5 – Enrich your vocabulary

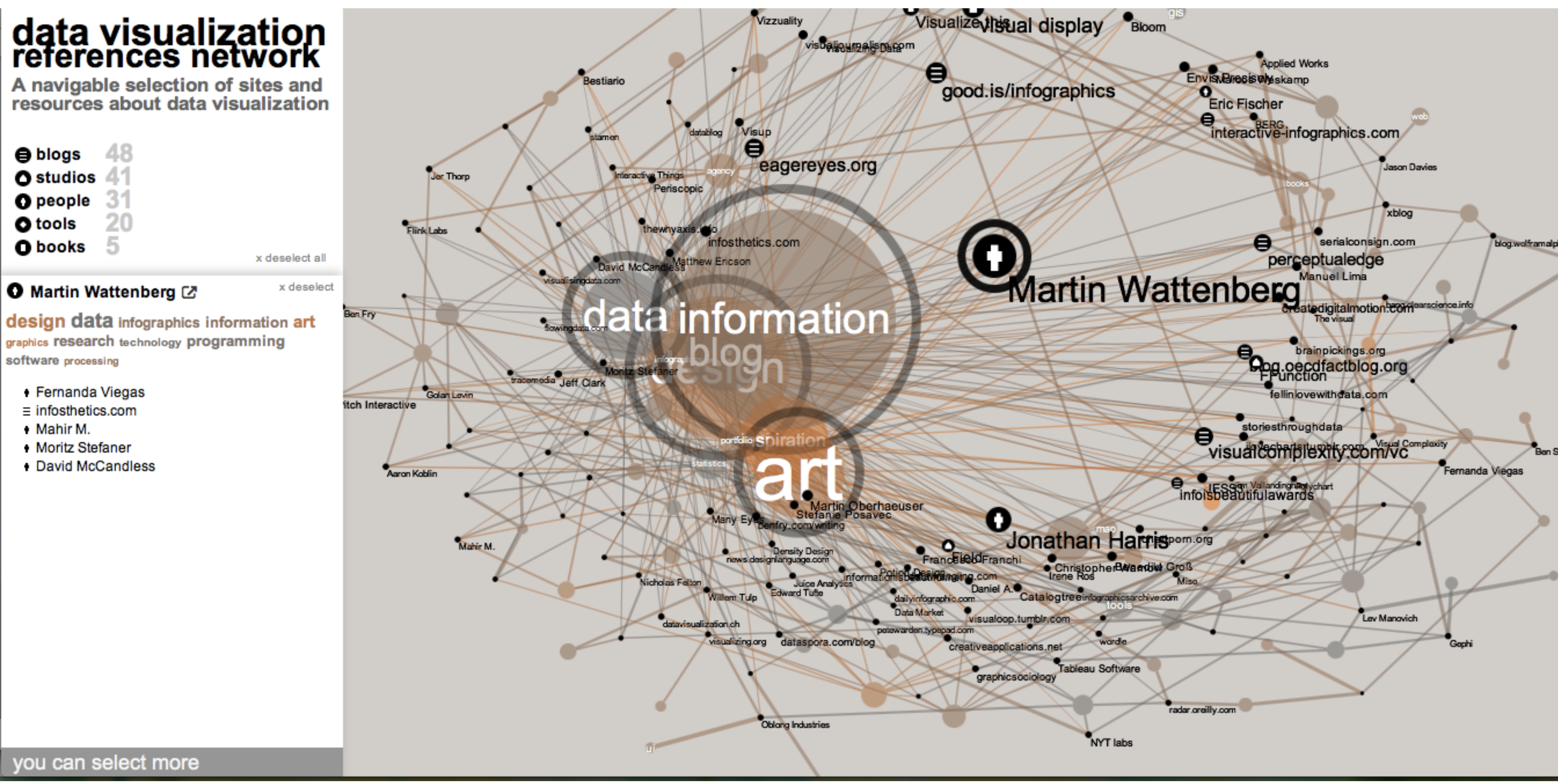
A consideration of the full spectrum of visual properties – color, shape, size, orientation, texture, value, and position – can and should be used



Bertin, Jacques. "Semiology of graphics: diagrams, networks, maps." (1983).

Principles of network visualization

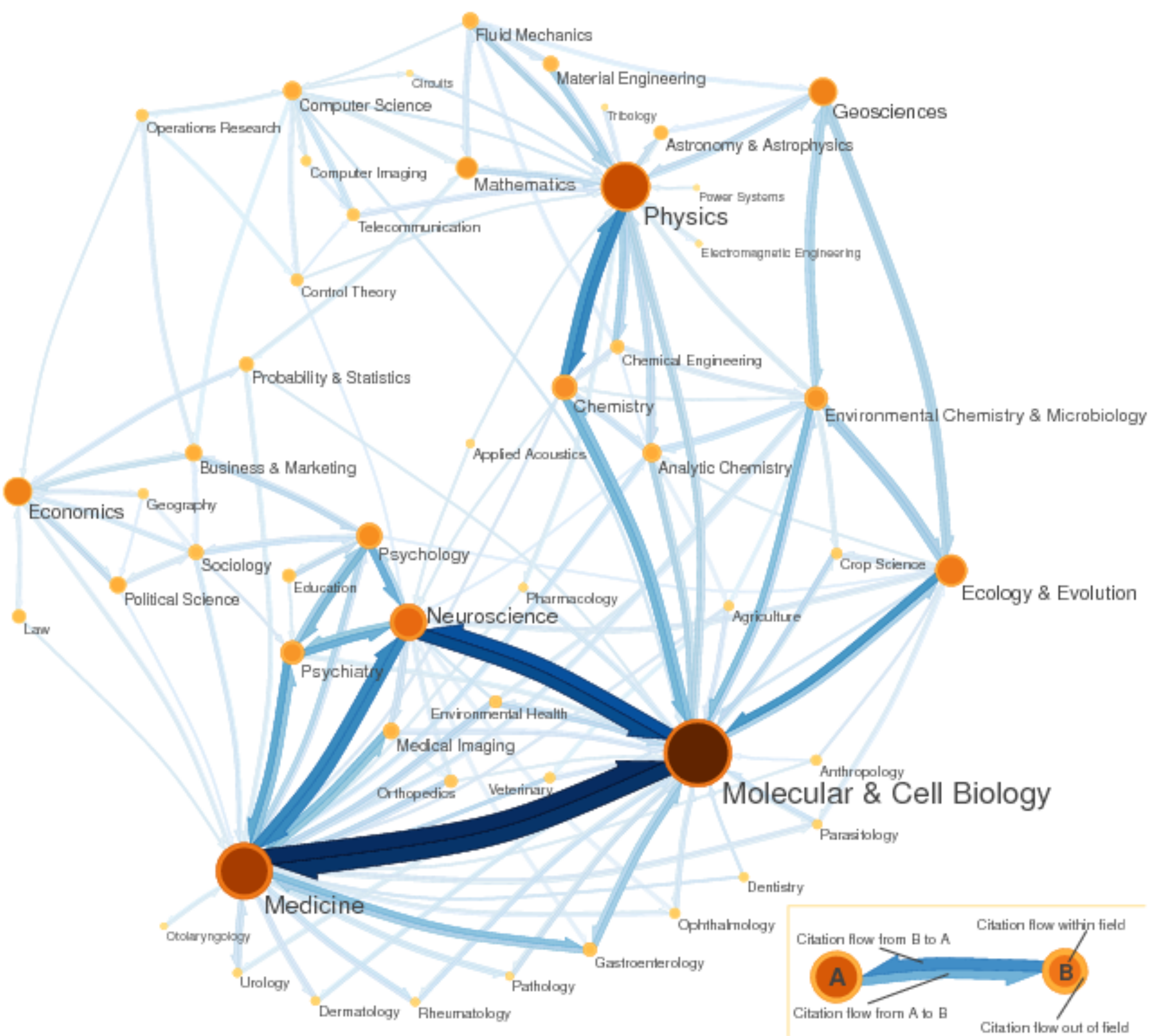
5 – Enrich your vocabulary



http://moebio.com/datavisnetwork/

Principles of network visualization

5 – Enrich your vocabulary



http://www.eigenfactor.org

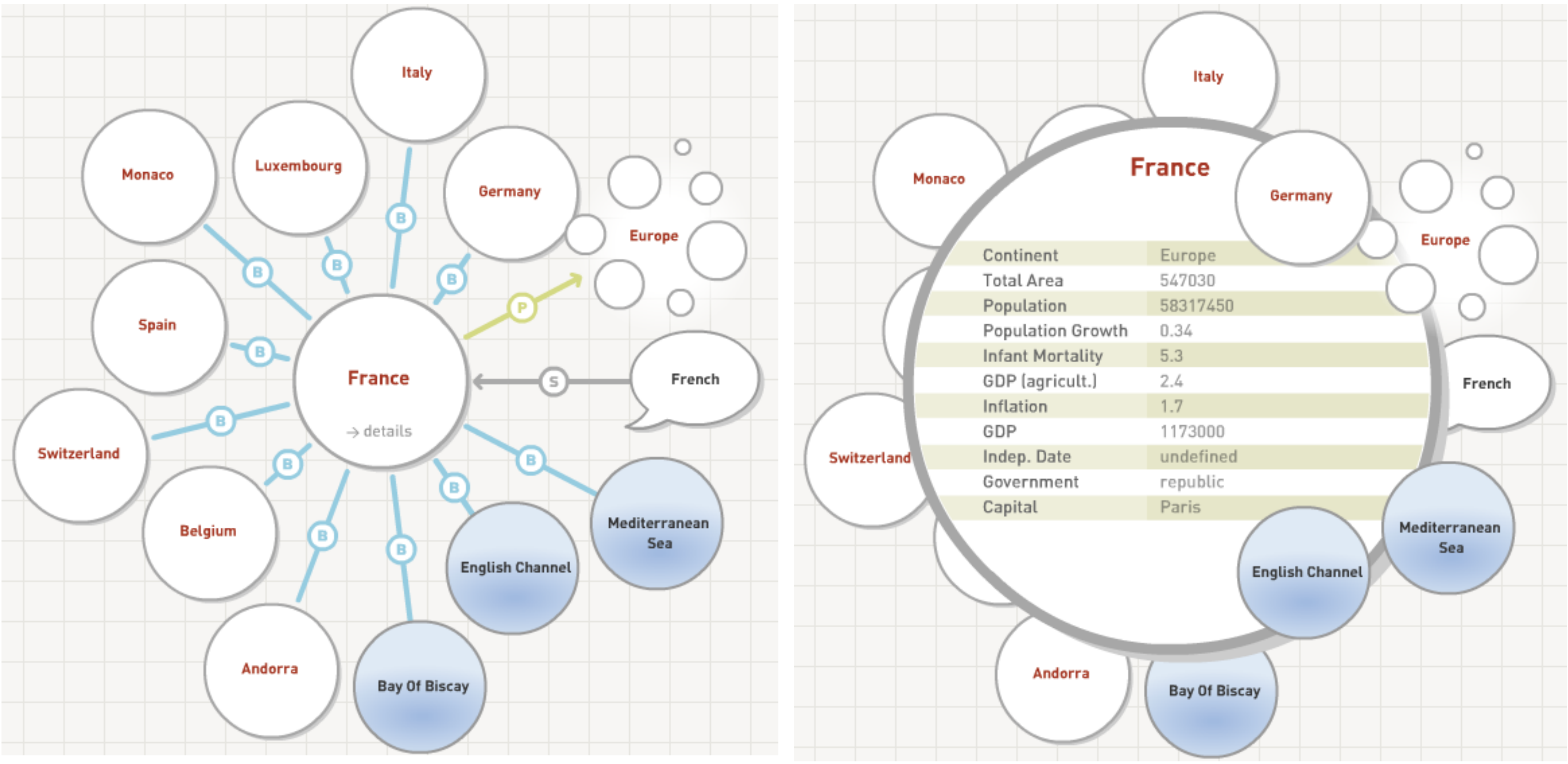
Principles of network visualization

Richer nodes

- They can be made more intelligible with an appropriate use of color and graphical features
- They can become responsive and provide important contextual information
- They can expand or shrink, show or hide relevant information, and morph

Principles of network visualization

Richer nodes



http://archive.stefaner.eu/projects/relation-browser/

Principles of network visualization

Expressive edges

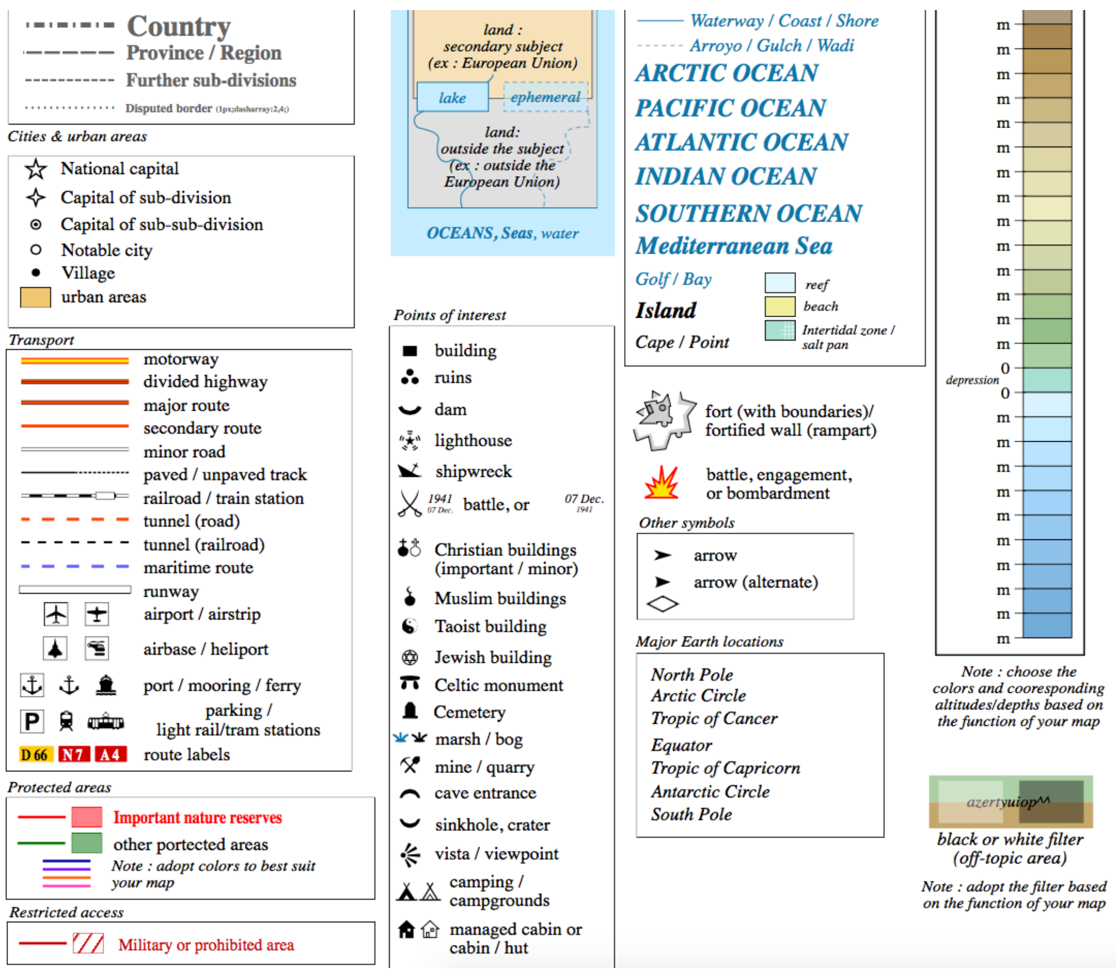
Edges can express much more than a simple connection between entities

For every relationship between nodes, there are countless layers of quantitative and qualitative information pertaining to the nature of the connection

Principles of network visualization

Expressive edges

Cartography is a great source for inspiration when examining the portrayal of edges



http://commons.wikimedia.org/wiki/File:Maps_template-en.svg

Principles of network visualization

Expressive edges

A similar process could be applied for network visualization

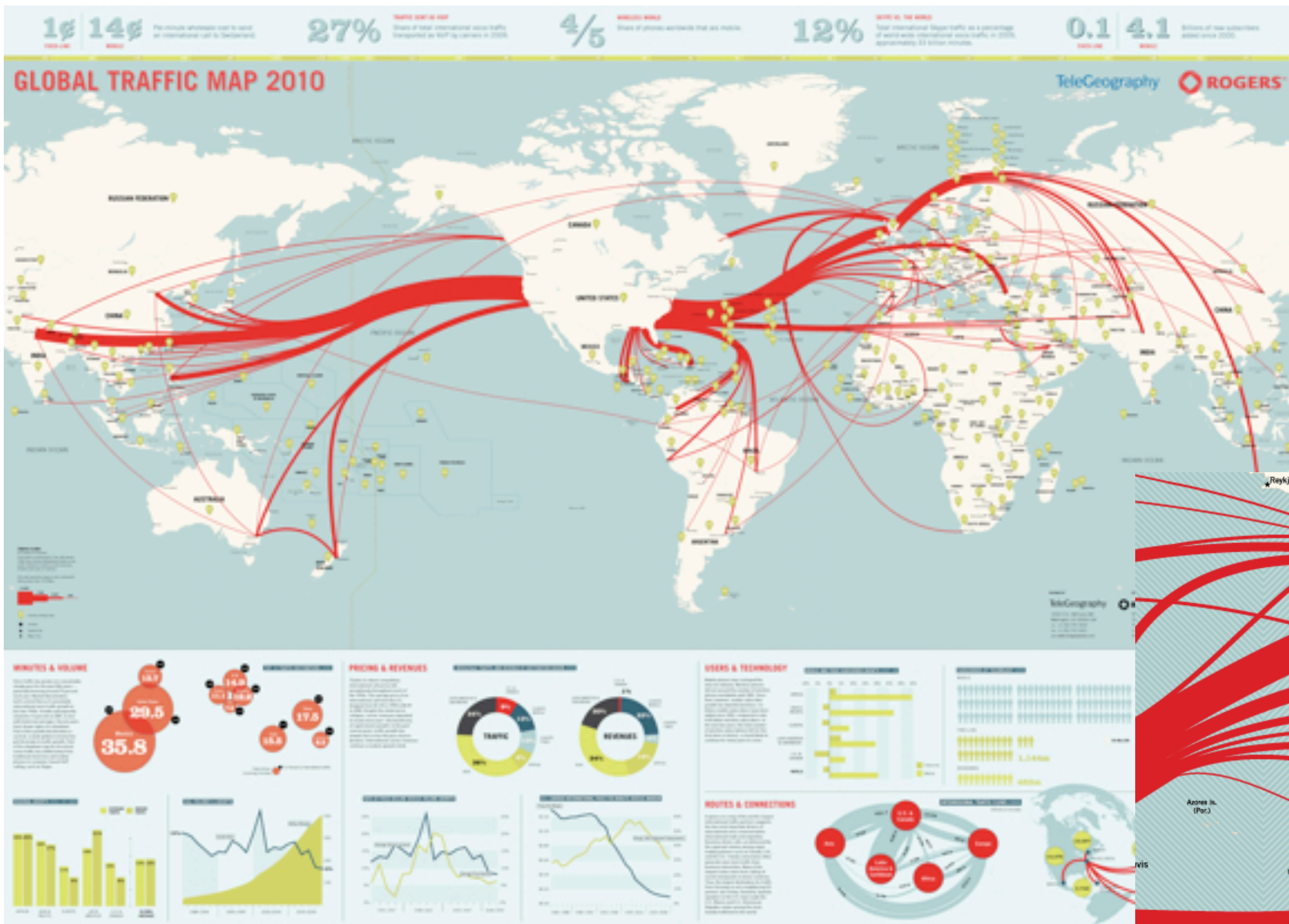
Principles of network visualization

Expressive edges

- **Length** to suggest a gradation of values, such as physical proximity, degree of relationship, strength, similarity, or relatedness;
- **Width** to express density or intensity flow, or an alternative gradation of values;
- **Color** to differentiate or highlight particular groups, categories and clusters, or alternatively, singular connections;
- **Shape** to communicate the type of relationship

Principles of network visualization

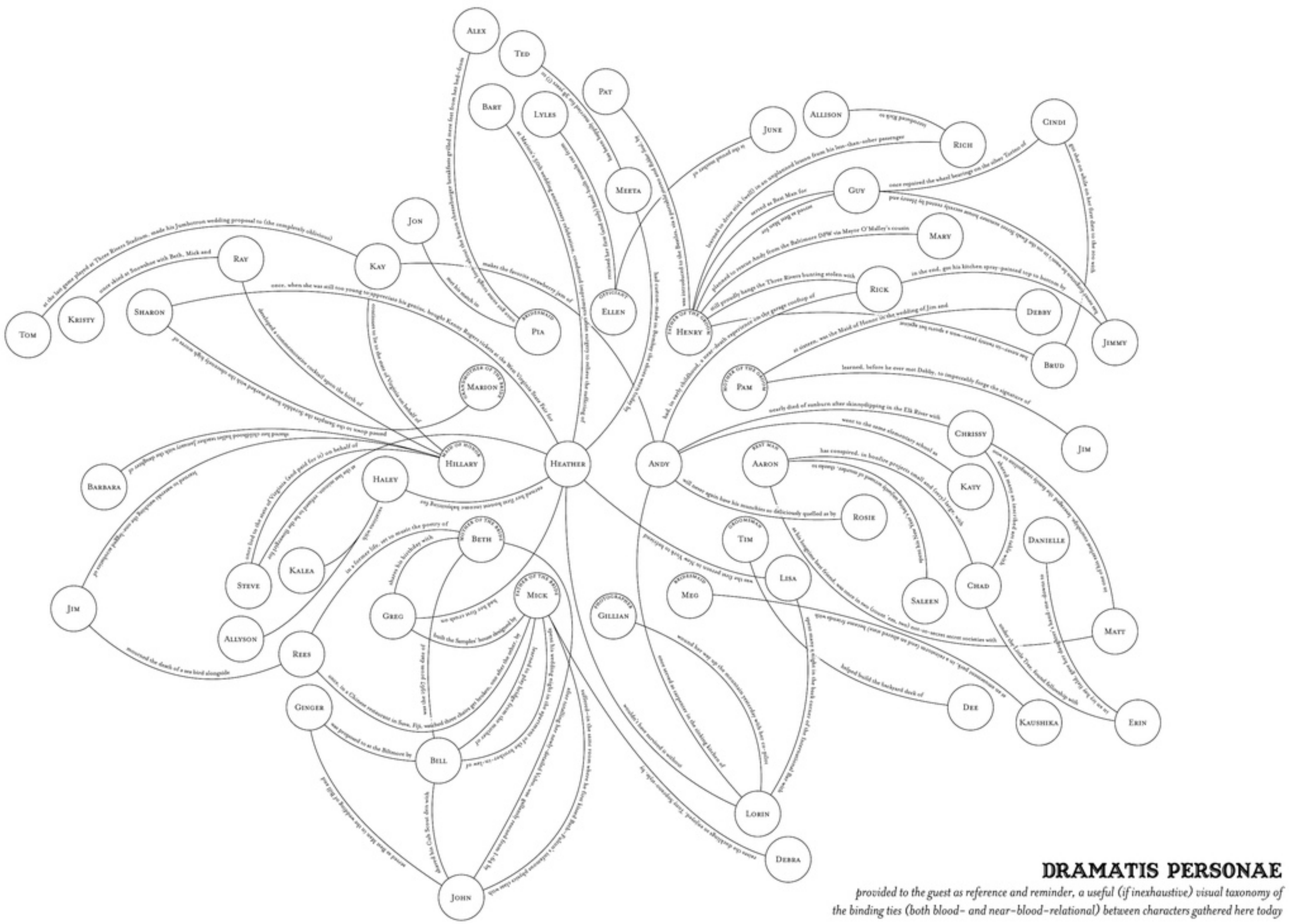
Expressive edges



https://www.telegeography.com/telecom-maps/global-traffic-map/

Principles of network visualization

Expressive edges

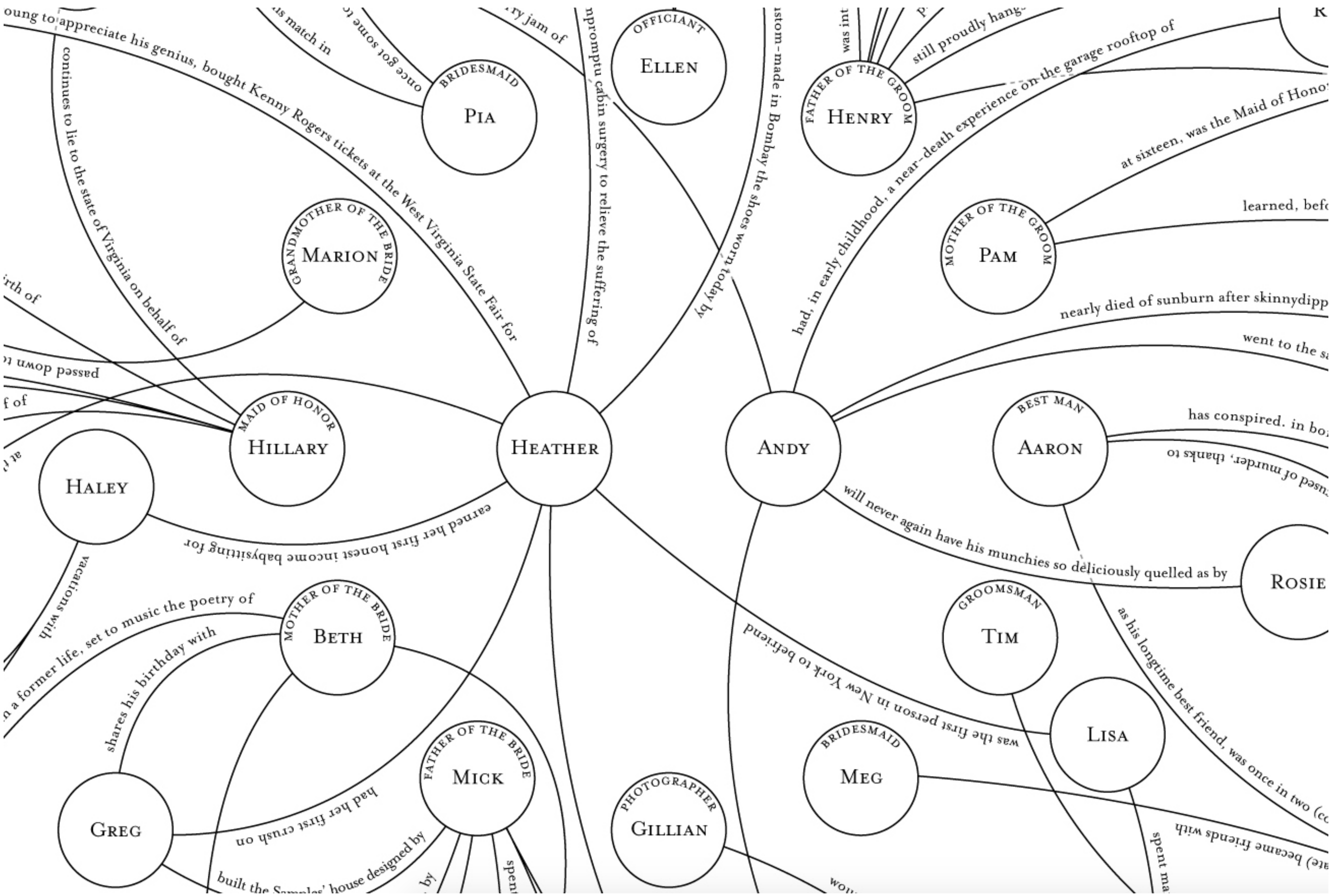


A map of relationships between the guests at Andrew and Heather's wedding. With the goal of helping start conversations, the couple produced a chart of the tightly knit group of family and friends at the wedding, connecting guests based on favorite shared stories, which they have collected and solicited from their parents.

https://www.flickr.com/photos/andrewcoulterenright

Principles of network visualization

Expressive edges

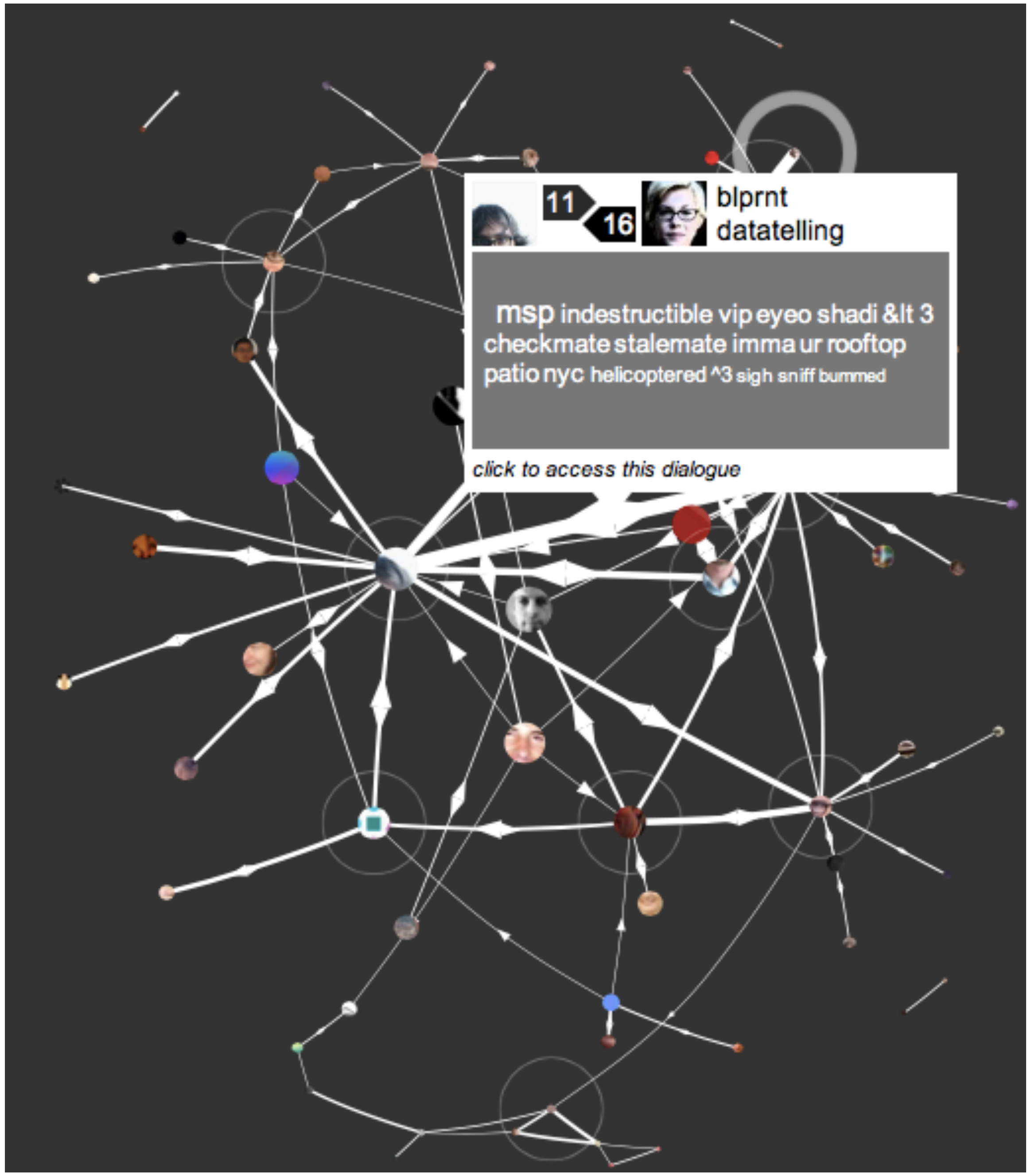


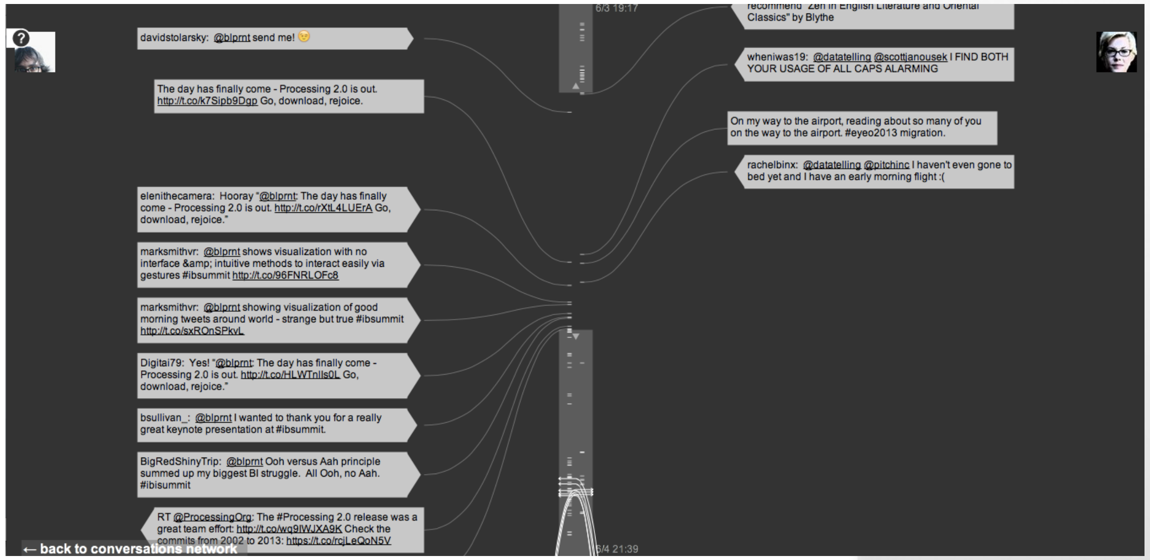
A map of relationships between the guests at Andrew and Heather’s wedding. With the goal of helping start conversations, the couple produced a chart of the tightly knit group of family and friends at the wedding, connecting guests based on favorite shared stories, which they have collected and solicited from their parents.

https://www.flickr.com/photos/andrewcoulterenright

Principles of network visualization

Expressive edges





Principles of network visualization

Clear visual language

One of the caveats behind the implementation of diverse graphical attributes is to beware of **creating a visual language that might not be immediately recognized by everyone**

A **legend** is simple, yet vital, allowing for a quick interpretation of the various graphic components

Principles of network visualization

6 – Expose grouping

The ability to showcase variation in a depicted system is a central attribute of network visualization

This can be explored by enriching the **visual vocabulary** but also by exploring the potentialities of **spatial arrangements**

Principles of network visualization

6 – Expose grouping

Grouping allows the apprehension of clusters, islands, prominent patterns, and the general distribution of nodes and links

The idea is to combine several units of information into related chunks in order to reinforce relationships, reduce complexity, an improve cognition

Principles of network visualization

6 – Expose grouping

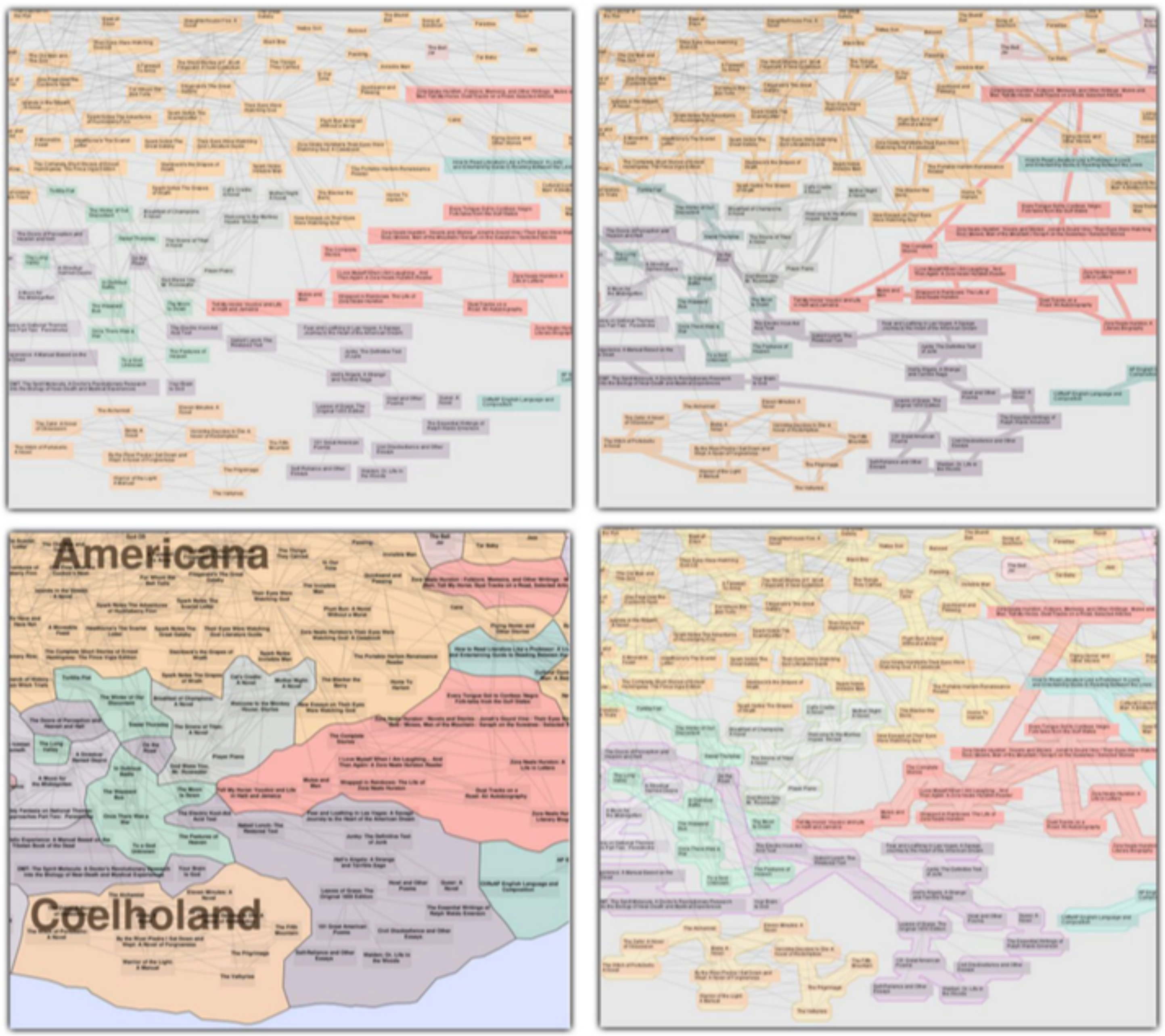


Fig. 1. Four visualizations for viewing group information over node-link diagrams: colored node-link diagram (top left), LineSets (top right), GMap (bottom left), and BubbleSets (bottom right).

Jianu, Radu, et al. "How to display group information on node-link diagrams: an evaluation." IEEE Transactions of Visualization and Computer Graphics (2014): 20(11): 1530-1541.

Principles of network visualization

Gestalt psychology

- An attempt at comprehending our perception of visual patterns
- Results of immense importance to most forms of visual communication
- Rules of perceptual organization, also known as Gestalt laws of grouping (similarity, proximity, an common fate) are important rules in exposing groups in networks visualization

Principles of network visualization

Law of similarity (graphical treatment)

Asserts that elements that are similar (either in terms of color, shape, or size) are perceived to be more related than elements that are dissimilar

Highlights the need for a differentiated graphical vocabulary in the depiction of nodes, as a critical measure for spotting similarities and differences

Principles of network visualization

Law of proximity (spatial arrangement)

States that elements that are close together are perceived as being more related than elements that are farther apart

The mere placement of homologous nodes closer to each other suggests inherent relationships not solely manifested by edges

Principles of network visualization

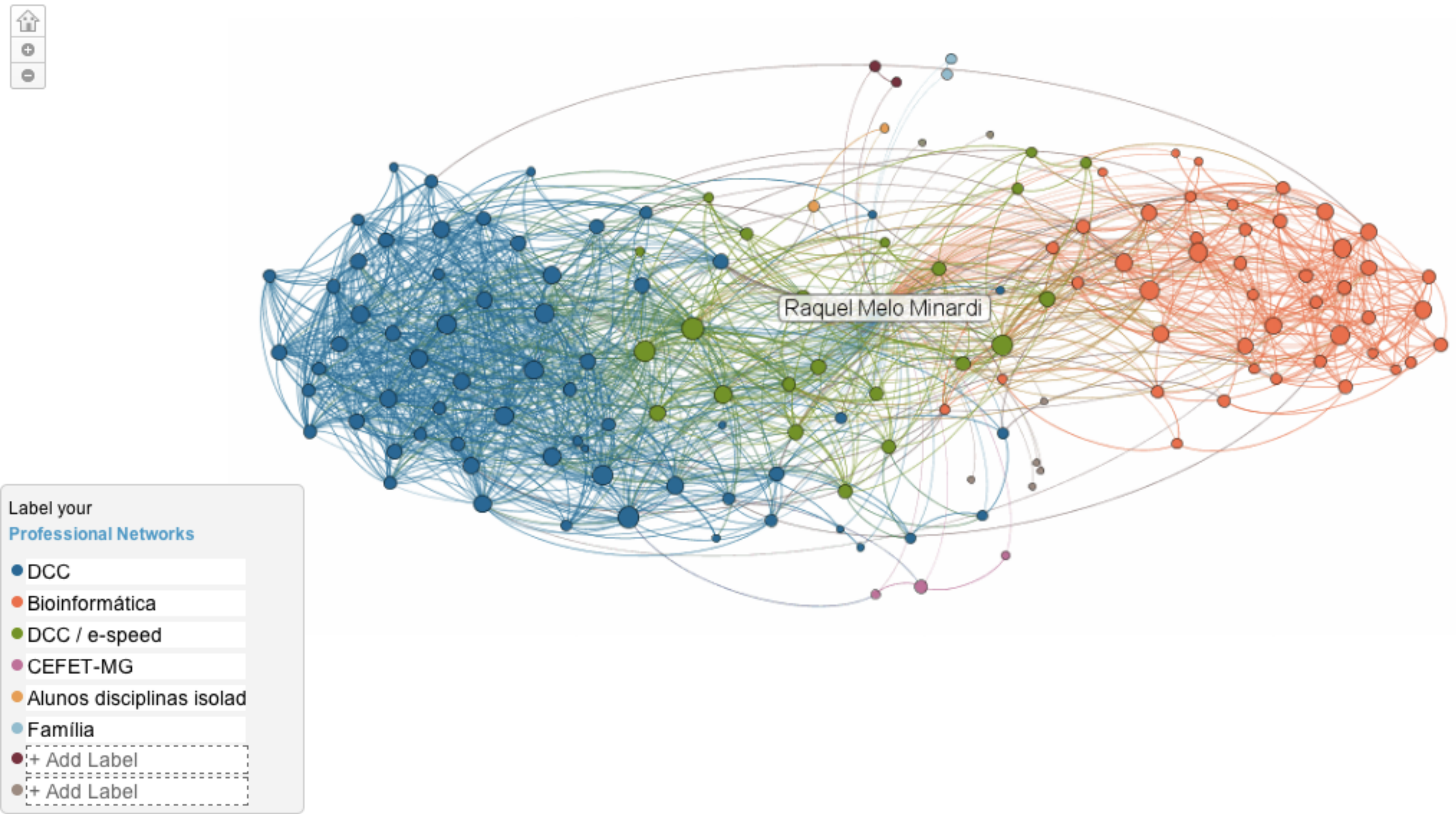
Law of common fate (motion)

Proclaims that elements that move simultaneously in the same direction and at the same speed are perceived as being more related that elements that are stationary or that move in different directions

Important to highlight contrast through animation

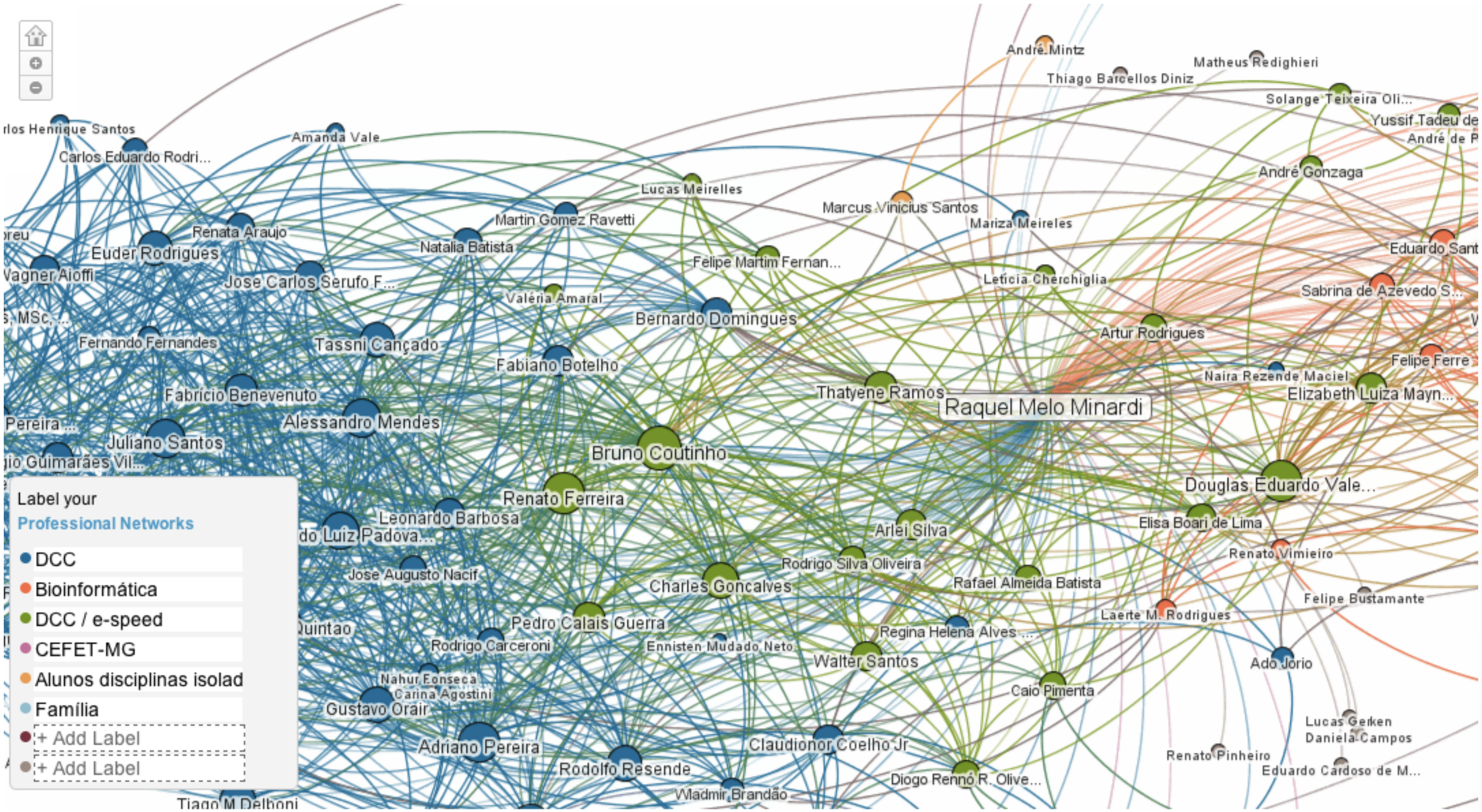
Principles of network visualization

6 – Expose grouping



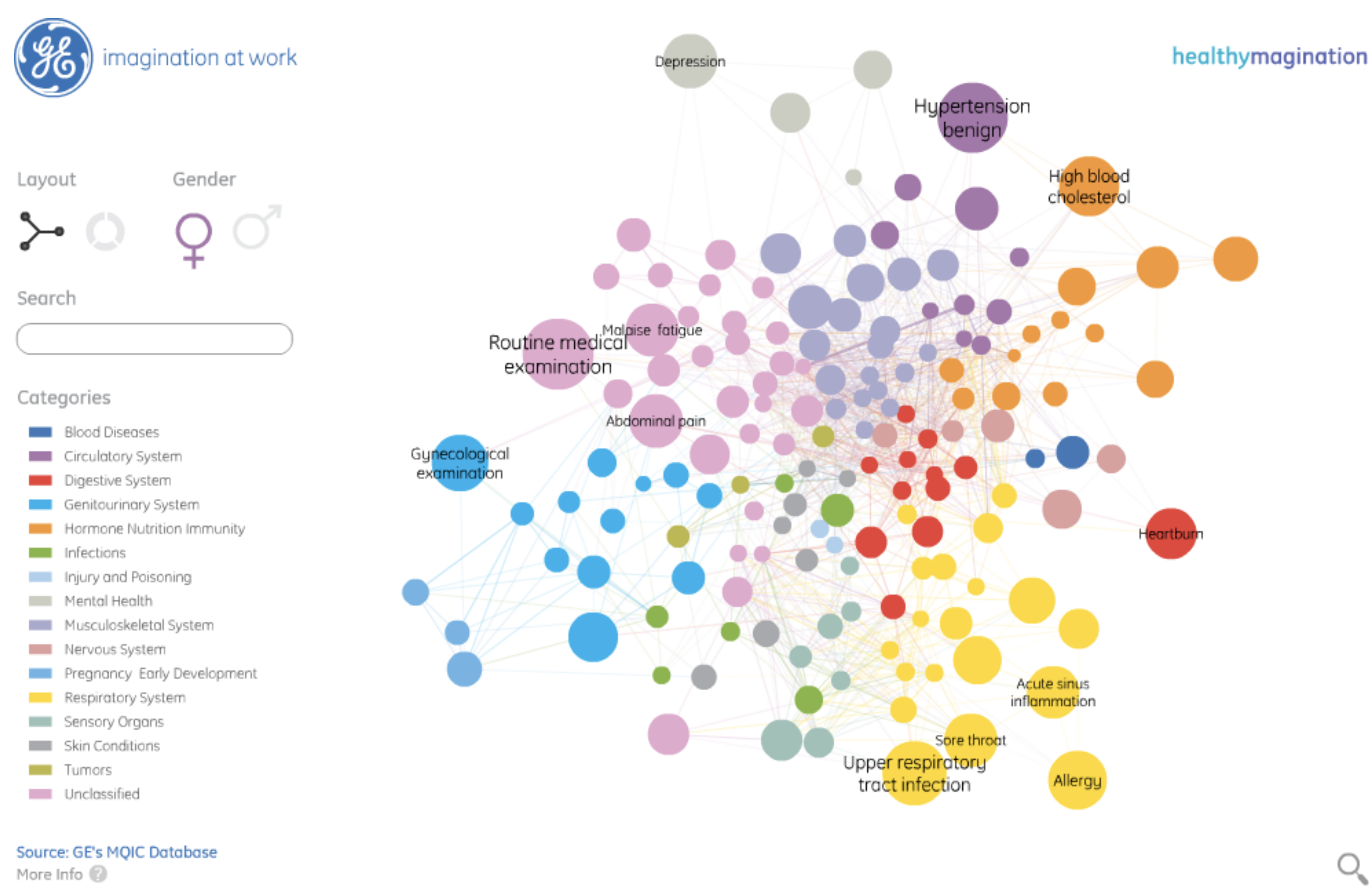
Principles of network visualization

6 – Expose grouping



Principles of network visualization

6 – Expose grouping



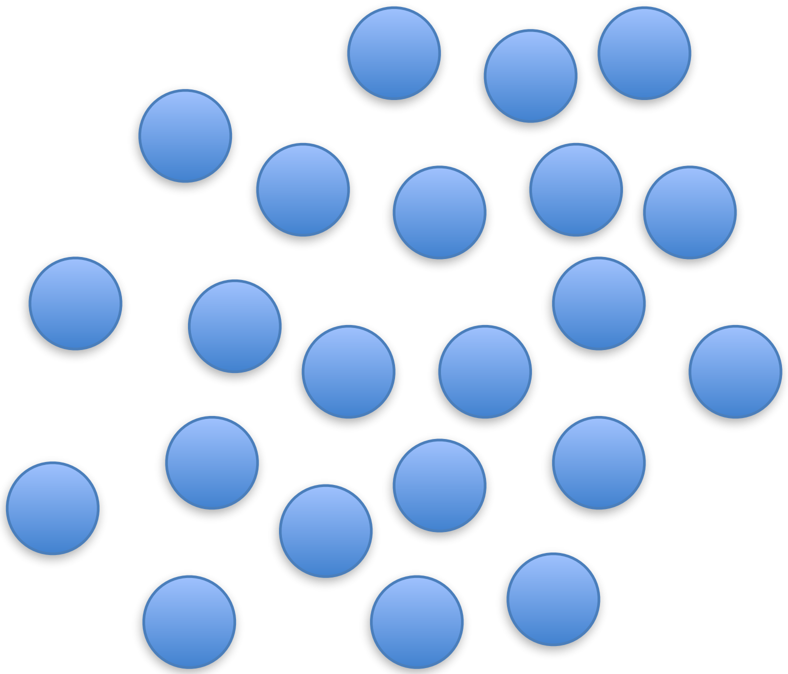
Principles of network visualization

7 – Maximizing scale

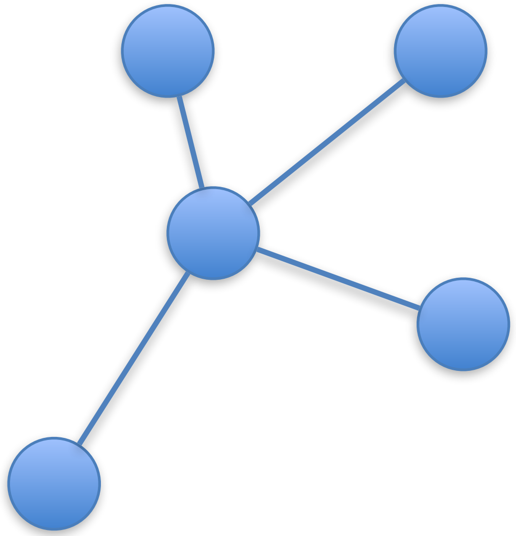
One of the biggest misconceptions in networks visualization is the notion that a representation that works at one scale will also work at a larger or smaller scale

Principles of network visualization

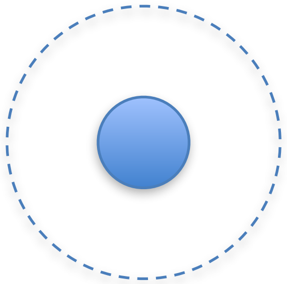
7 – Maximizing scale



Macro analysis
(patterns)



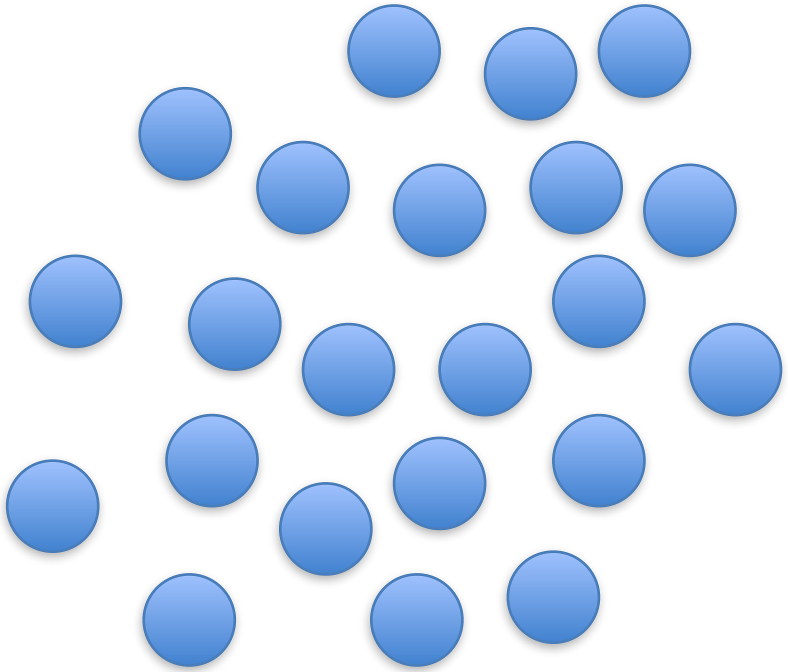
Relationship analysis
(connectivity)



Micro analysis
(entities)

Principles of network visualization

7 – Maximizing scale



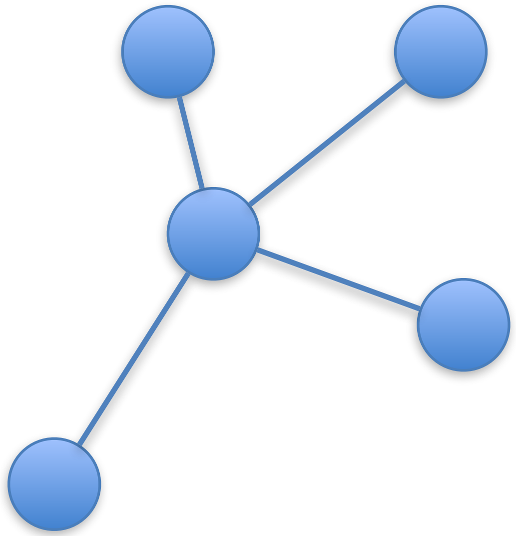
Macro analysis
(patterns)

The visualization does not have to provide a detailed understanding of individual links, and less so of individual nodes

It should provide a general view into the networks and highlight certain clusters, as well as isolated groups, within its structure

Principles of network visualization

7 – Maximizing scale

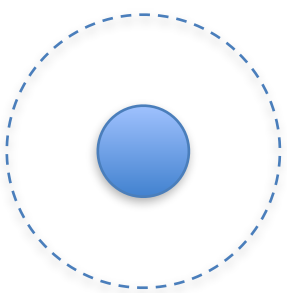


The visualization is concerned with an effective analysis at the types of relationships among the mapped entities (nodes)

Relationship analysis
(connectivity)

Principles of network visualization

7 – Maximizing scale



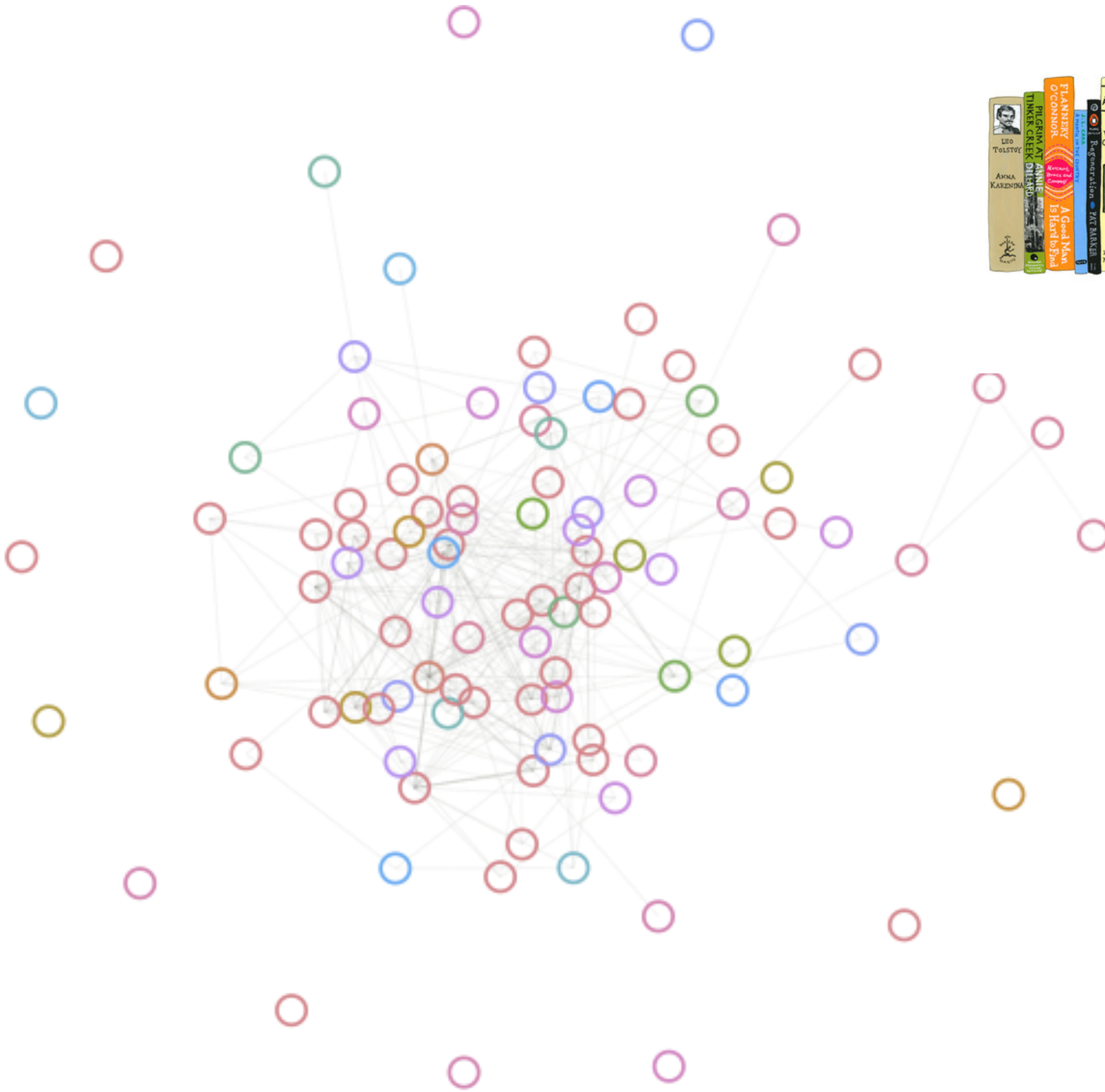
Micro analysis
(entities)

The last layer of insight provided by an efficient network visualization relates to the disclosure of an individual node's qualitative attributes

Should be comprehensive and explicit, providing detailed information, facts, and characteristics on a single-node entity

Principles of network visualization

7 – Maximizing scale



SUMMARY STATS

- Number of shelves: **101**
- Number of Books Chosen: **1,645**
- Unique Books According to Google's API: **1,431**
- Average number of books chosen:**16.28**
- Average Pagecount: **381.2**
- Average Year of Publication: **1992**
- Top 5 Chosen Books:

  1. **Lolita** chosen by 8 contributors
  2. **Moby Dick** (chosen by 7)
  3. **Jesus' Son** (chosen by 5 contributors)
  4. **The Wind-Up Bird Chronicle** (chosen by 5 contributors)
  5. **Ulysses** (chosen by 5 contributors)

- Top 5 Authors:

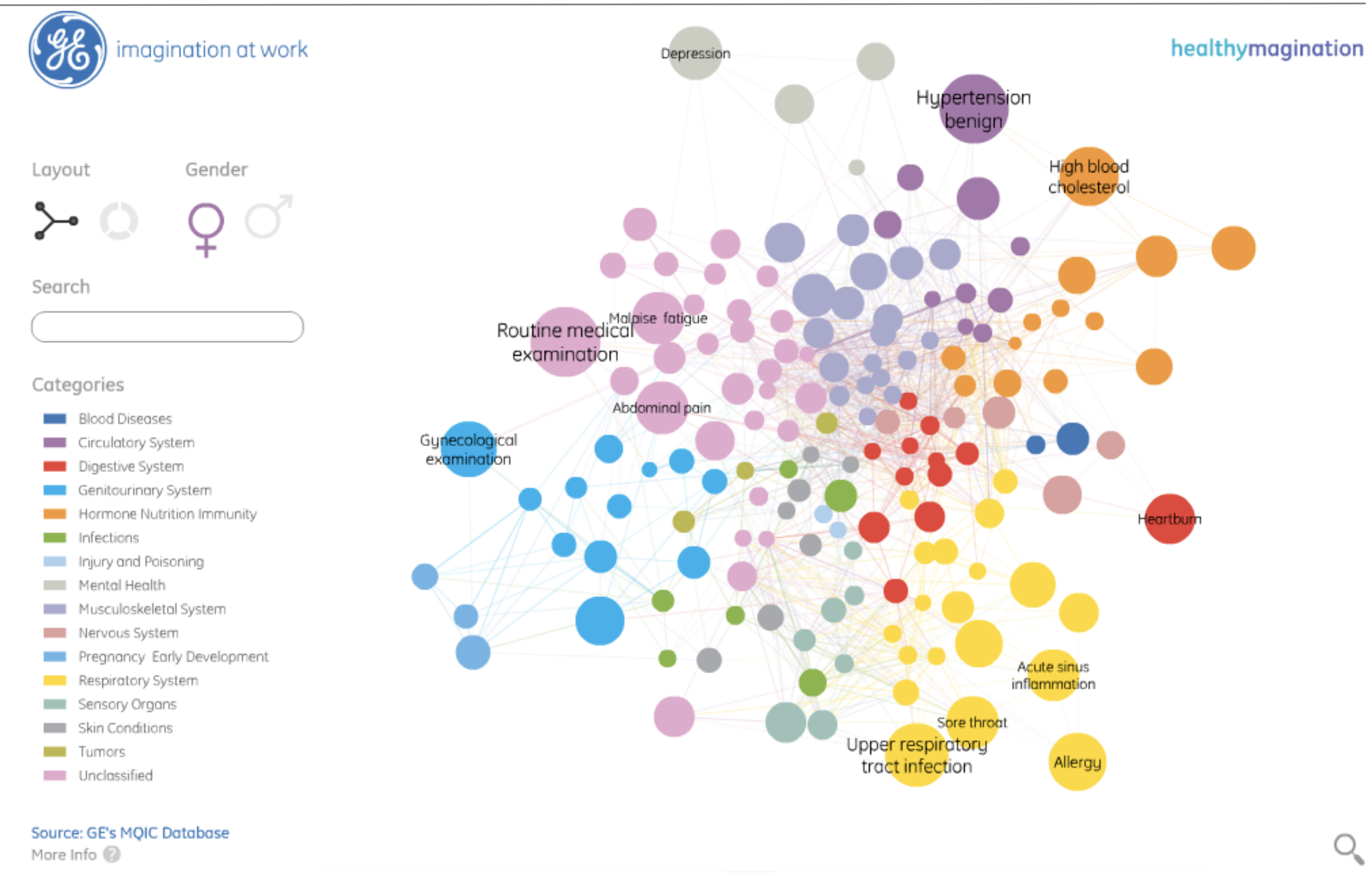
  1. **William Shakespeare** (10 different books)
  2. **Ernest Hemingway** (7 different books)
  3. **Graham Green** (7 different books)
  4. **Anton Chekov** (6 different books)
  5. **Edith Wharton** (6 different books)

- Contributor with the most number of books: **James Franco**
- Contributor with the most number of shared books: **James Franco**
- Longest Book: **The Oxford English Dictionary, Second Edition: Volume XX*** chosen by Stephin Merritt, 22,000 pages.
- Shortest Book: **Pac-Mastery: Observations and Critical Discourse by C.F. Gordon** chosen by Tauba Auerbach, 12 pages

** Jane was only able to paint one volume of Stephen's OED for his shelf, but the authors agreed it could stand as a synedoche for his choice of the entire edition.*

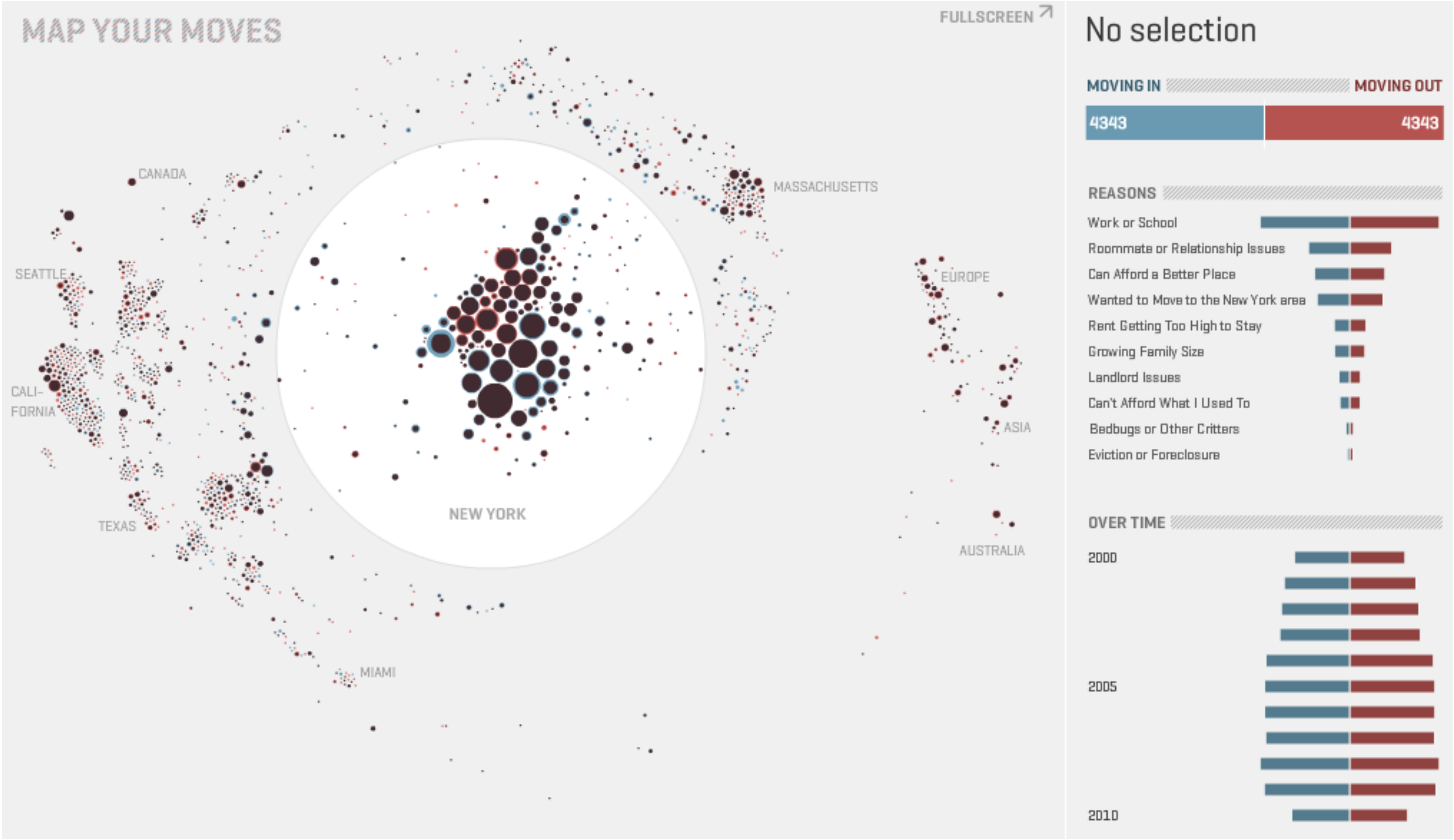
Principles of network visualization

7 – Maximizing scale



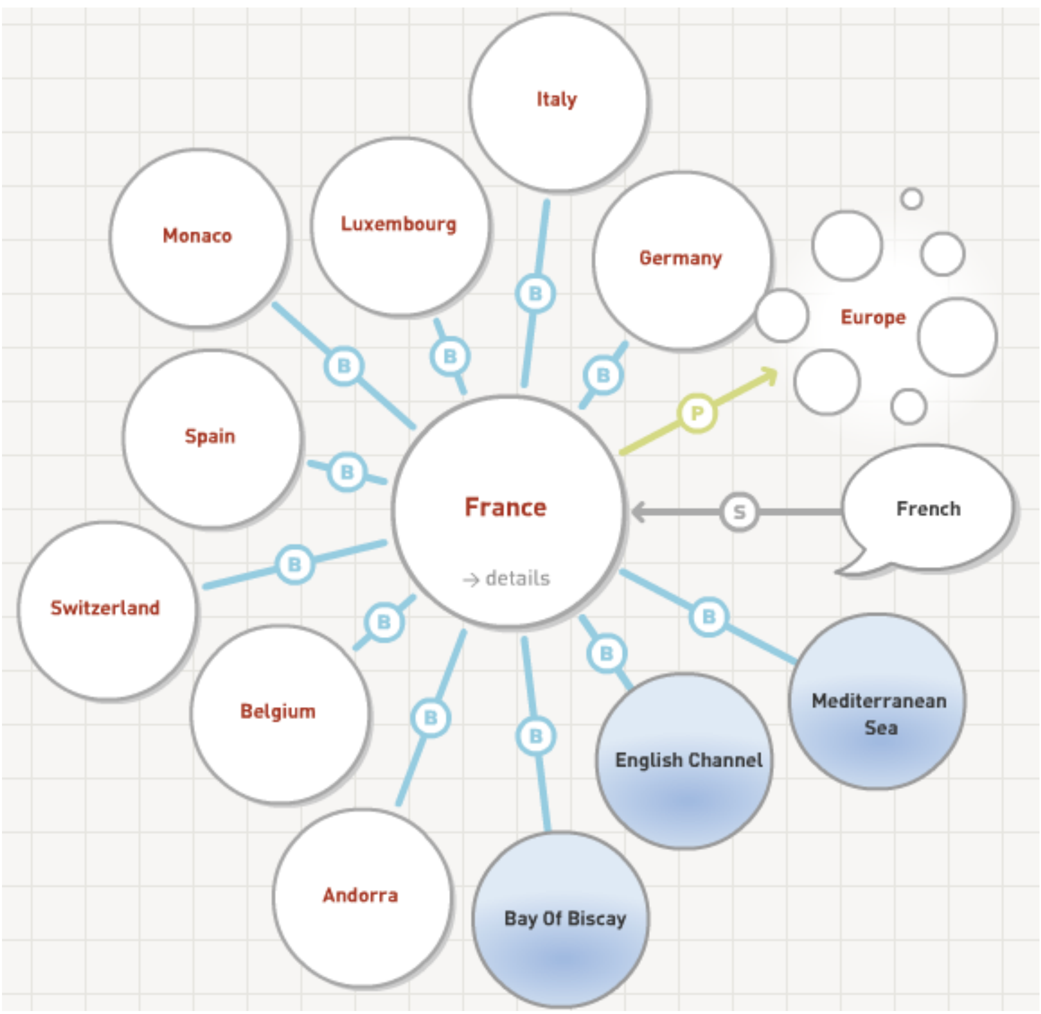
Principles of network visualization

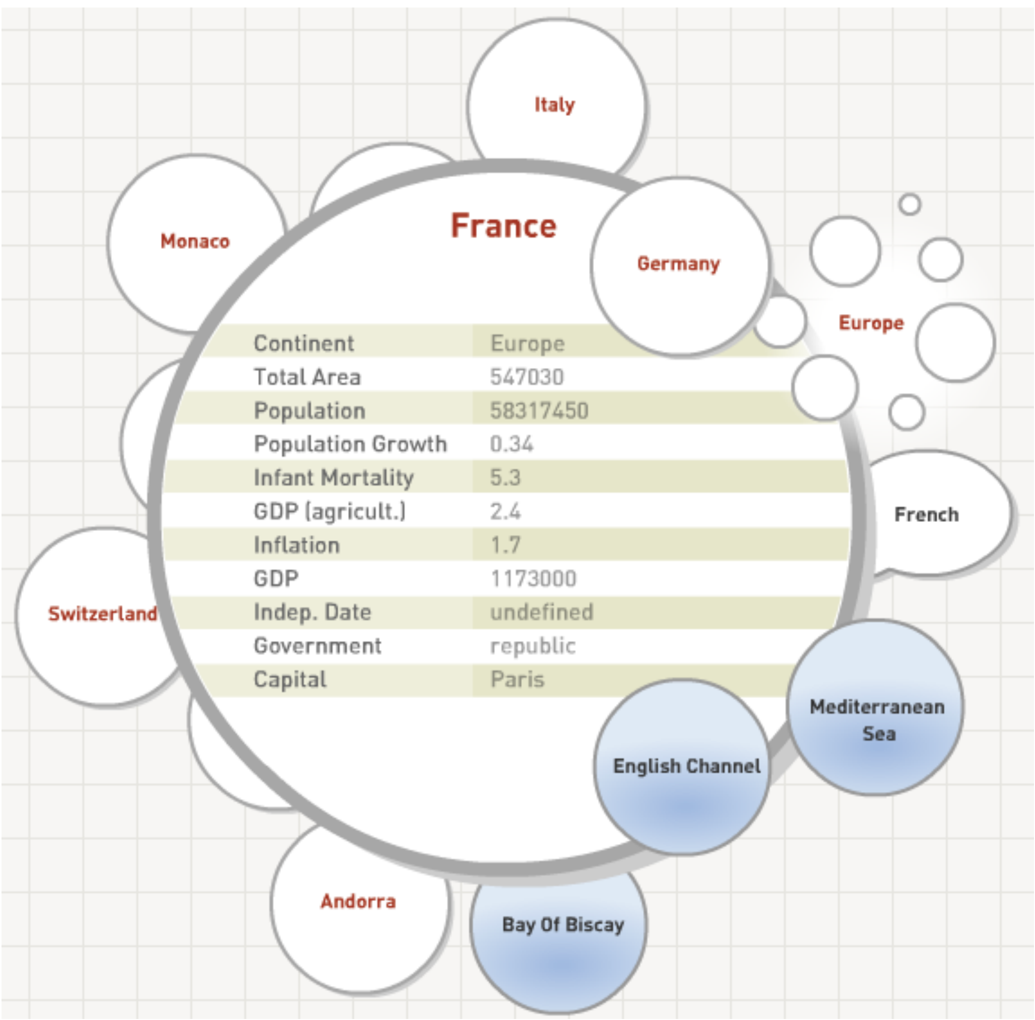
7 – Maximizing scale



Principles of network visualization

7 – Maximizing scale





http://archive.stefaner.eu/projects/relation-browser/

Principles of network visualization

8 – Manage intrincacy

Ben Schneiderman's renowned visual-information-seeking mantra proposed in his seminar paper "The eyes have it: a task by data type taxonomy for information visualizations":

> "Overview first, zoom and filter, then details on demand"