Algunas Aplicaciones de la Optimización Matemática

Juan José Salazar González
DEIOC – Facultad de Matemáticas
Universidad de La Laguna

Sociedad, Ciencia, Tecnología y Matemáticas
15 abril 2004

Introducción

Pretende ayudar en los procesos de toma de decisiones donde hay limitaciones de recursos y se desea optimizar una función objetivo:

maximizar $f(\mathbf{x})$
sujeto a: $g(\mathbf{x}) \leq \mathbf{0}$

Destacan:

Programación Lineal Continua

$\max\ \mathbf{b}^T \cdot \mathbf{x}$
$s.a.:\ A \cdot \mathbf{x} \leq \mathbf{c}$

$\max\ \mathbf{b}^T \cdot \mathbf{x}$
$s.a.:\ A \cdot \mathbf{x} \leq \mathbf{c}$
$\mathbf{x}$ *entero*

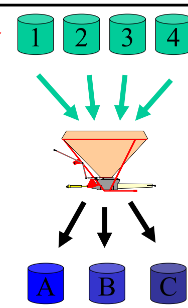
Programación Lineal Entera

Enumeración:

El número de ordenaciones diferentes de n elementos viene dado por $1.2.3.4.5...(n-1).n = n!$

n	*n!*	*tiempo*
12	479001600	1 minuto
13	6227020800	13 minutos
14	87178291200	3 horas
15	1307674368000	45 horas
16	20922789888000	30 días
17	355687428096000	1 año
18	6402373705728000	18 años
19	121645100408832000	3 siglos
20	2432902008176640000	60 siglos

Optimizar una Mezcla



Problema:

M.Prima	Disponibilidad	Costo/barril
1	3000 barriles	300 €
2	2000 barriles	600 €
3	4000 barriles	400 €
4	1000 barriles	500 €

Tipo	Octanos	1	2	3	4	Beneficio/barril
A	98	≤30%	≥40%	≤50%	-------	5,5 mil €
B	95	≤50%	≥10%	-------	-------	4,5 mil €
C	97	≥70%	-------	-------	-------	3,5 mil €

Optimizar una Mezcla

Modelo Matemático:

x_{ij} = cantidad de producto i que se destina a gasolina j, para cada $i \in \{1,2,3,4\}$ y cada $j \in \{A,B,C\}$

$$\max\ 5500(x_{1A}+x_{2A}+x_{3A}+x_{4A})+4500(x_{1B}+x_{2B}+x_{3B}+x_{4B})+3500(x_{1C}+x_{2C}+x_{3C}+x_{4C})$$
$$-300(x_{1A}+x_{1B}+x_{1C})-600(x_{2A}+x_{2B}+x_{2C})-400(x_{3A}+x_{3B}+x_{3C})-500(x_{4A}+x_{4B}+x_{4C})$$

Sujeto a:

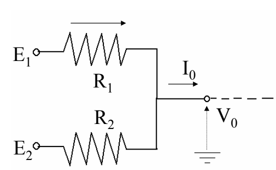
$$x_{1A}+x_{1B}+x_{1C} \leq 3000$$
$$x_{2A}+x_{2B}+x_{2C} \leq 2000$$
$$x_{3A}+x_{3B}+x_{3C} \leq 4000$$
$$x_{4A}+x_{4B}+x_{4C} \leq 1000$$

$$x_{1A} \leq 0,3(x_{1A}+x_{2A}+x_{3A}+x_{4A})$$
$$x_{2A} \geq 0,4(x_{1A}+x_{2A}+x_{3A}+x_{4A})$$
$$x_{3A} \leq 0,5(x_{1A}+x_{2A}+x_{3A}+x_{4A})$$
$$x_{1B} \leq 0,5(x_{1B}+x_{2B}+x_{3B}+x_{4B})$$
$$x_{2B} \geq 0,1(x_{1B}+x_{2B}+x_{3B}+x_{4B})$$
$$x_{1C} \geq 0,7(x_{1C}+x_{2C}+x_{3C}+x_{4C})$$

$$x_{1A} \geq 0,\ x_{2A} \geq 0,\ x_{3A} \geq 0,\ x_{4A} \geq 0, \ldots,\ x_{1C} \geq 0,\ x_{2C} \geq 0,\ x_{3C} \geq 0,\ x_{4C} \geq 0.$$

Optimizar un Divisor de Tensión

Problema:



Dos generadores de tensión con fuerzas electromotrices $E_1 > E_2$ con valores centrados ($\underline{E}_1$ y $\underline{E}_2$) y tolerancias (E_1^-, E_1^+, E_2^- y E_2^+) conocidos. Se desea determinar las resistencias centradas $\underline{R}_1$ y $\underline{R}_2$ de manera que la impedancia resistiva del divisor sea mínima y el potencial V_0 se mantenga en $[V_0^{min}, V_0^{max}]$ cuando la corriente I_0 se mueve entre $[I_0^{min}, I_0^{max}]$. Se conocen las tolerancias de las resistencias y que $E_1^+ \geq V_0^{max} \geq V_0^{min} \geq E_2^+$.

Optimizar un Divisor de Tensión

Modelo Matemático:

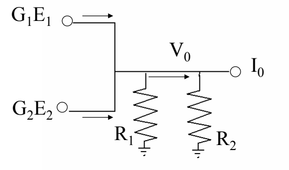
impedancia: $R_0=R_1R_2/(R_1+R_2)$ $\equiv$ admitancia: $G_0=G_1+G_2$
minimizar R_0^+ $\equiv$ maximizar $G_0^-=G_1^-+G_2^-$
Sean ε_1 y ε_2 tales que: $G_1^- := (1-\varepsilon_1)\,\underline{G}_1$ y $G_2^- := (1-\varepsilon_2)\,\underline{G}_2$

$\underline{G}_1$ = admitancia centrada de la resistencia R_1
$\underline{G}_2$ = admitancia centrada de la resistencia R_2

max $(1-\varepsilon_1)\,\underline{G}_1 + (1-\varepsilon_2)\,\underline{G}_2$
s.a:

$$(1-\varepsilon_1)(E_1^- - V_0^{min})\,\underline{G}_1 + (1+\varepsilon_2)(E_2^- - V_0^{min})\,\underline{G}_2 \geq I_0^{max}$$
$$(1+\varepsilon_1)(E_1^+ - V_0^{max})\,\underline{G}_1 + (1-\varepsilon_2)(E_2^+ - V_0^{max})\,\underline{G}_2 \leq I_0^{min}$$
$$\underline{G}_1 \geq 0,\ \underline{G}_2 \geq 0.$$

Circuito equivalente como fuentes de intensidad:



$$V_0 = \frac{E_1G_1+E_2G_2-I_0}{G_1+G_2}$$

Para R_2 fija:

$$\frac{\delta V_0}{\delta G_1} = \frac{G_2(E_1-E_2)}{(G_1+G_2)^2} > 0$$

Es decir, para R_2 fijo, V_0 aumenta cuando G_1 aumenta.
Y de igual forma, para R_1 fija, V_0 aumenta cuando G_2 disminuye.
Por tanto, V_0 toma su mayor valor con G_1^+, G_2^- y I_0^{min}, y así:

$$V_0 \leq V_0^{max} \equiv \frac{E_1^+G_1^+ + E_2^+G_2^- - I_0^{min}}{G_1^+ + G_2^-} \leq V_0^{max} \equiv \frac{E_1^+(1+\varepsilon_1)\underline{G}_1 + E_2^+(1-\varepsilon_2)\underline{G}_2 - I_0^{min}}{(1+\varepsilon_1)\,\underline{G}_1 + (1-\varepsilon_2)\,\underline{G}_2} \leq V_0^{max}$$

Optimizar Encendidos

Problema:

Se tiene una máquina durante los próximos n días. Para cada día i conocemos el beneficio b_i ($\geq\leq 0$) si la máquina está encendida. Si nos imponen un límite máximo k de encendidos, ¿qué días debemos encender la máquina para maximizar el beneficio?

Optimizar Encendidos

Modelo Matemático:

x_i = 1 si la máquina se enciende el día i; =0 otro caso
y_i = 1 si la máquina funciona el día i; =0 otro caso.

max $b_1y_1+\ldots+b_ny_n$
s.a: $x_1+\ldots+x_n \leq k$
$x_i \geq y_i - y_{i-1}$ para todo $i=1,\ldots,n$
$0 \leq x_i \leq y_i \leq 1$ para todo $i=1,\ldots,n$

donde se asume un día 0 en el que la máquina no funcionó:
$y_0 := 0$.

Observación: no es necesario exigir "variables enteras".

Optimizar una Banda de Músicos

Problema:

9 músicos (A,...,I) y 8 sinfonías (1,...,8)

musi	A	B	C	D	E	F	G	H	I
Sinfo	1,3,8	2,5,8	1,2,6,8	1,4,6	2,4,6,7	1,2,5,7,8	4,6,8	5,7	1,2,4
coste	2	3	3	2	1	2	2	1	2

¿Cómo debe ordenar el director las sinfonías para que un ensayo le cueste lo menos posible?

Optimizar una Banda de Músicos

Una columna por cada músico

$$A = \begin{vmatrix} 1&0&1&1&0&1&0&0&1 \\ 0&1&1&0&1&1&0&0&1 \\ 1&0&0&0&0&0&0&0&0 \\ 0&0&0&1&1&0&1&0&1 \\ 0&1&0&0&0&1&0&1&0 \\ 0&0&1&1&1&0&1&0&0 \\ 0&0&0&0&1&1&0&1&0 \\ 1&1&1&0&0&1&1&0&0 \end{vmatrix}$$

Una fila por cada sinfonía

a_{ik}=1 sii la sinfonía i necesita al músico k

Optimizar una Banda de Músicos

Para cada sinfonía i y cada posición j en el orden:

x_{ij}=1 si la sinfonía se toca en la posición j; =0 otro caso

Para cada músico k:

u_k = primera posición donde se necesita el músico k

w_k = última posición donde se necesita el músico k

min $\sum_k c_k \cdot (w_k+1-u_k)$

s.a: $\sum_j x_{ij} = 1$ $\quad\forall$ sinfonía i

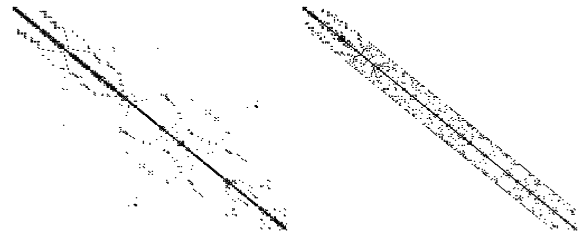
$\sum_i x_{ij} = 1$ $\quad\forall$ posición j

$x_{ij} \in \{0,1\}$ $\quad\forall$ sinfonía i y $\forall$ posición j

$\sum_j j \cdot x_{ij} \geq u_k$ $\quad\forall$ sinfonía i y $\forall$ músico k : $a_{ik} = 1$

$\sum_j j \cdot x_{ij} \leq w_k$ $\quad\forall$ sinfonía i y $\forall$ músico k : $a_{ik} = 1$

Minimizando la banda de una matriz



(a) Initial layout: $\psi = 115$ (b) Optimal layout: $\phi(G) = 21$

Minimizando la banda de una matriz

Para cada fila/columna i y cada posición j en el orden:

x_{ij}=1 si i va a la posición j; =0 otro caso

La posición donde va la fila/columna i viene dada por:

$p_i = \sum_j j\, x_{ij}$

Y la banda máxima b deberá cumplir:

$b = \max\{p_i - p_j : a_{ij}=1\}$

min b

s.a: $\sum_j x_{ij} = 1$ $\quad\forall$ fila/columna i

$\sum_i x_{ij} = 1$ $\quad\forall$ posición j

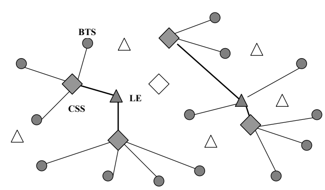
$x_{ij} \in \{0,1\}$ $\quad\forall$ fila/columna i y $\forall$ posición j

$p_i = \sum_j j \cdot x_{ij}$ $\quad\forall$ fila/columna i

$p_i - p_j \leq b$ $\quad\forall$ filas/columnas i y j : $a_{ij} = 1$

Problema de Telecomunicaciones

- Problema:



Universal Mobile Telecommunication System (UMTS):

- *Mobile Terminal (MT)*
- *Base Transceiver Station (BTS)*
- *Cell Site Switch (CSS)*
- *Local Exchange (LE)*

Proyecto Europeo STORMS, bajo *CSELT* (*Telecom*, Italia)

Problema de Telecomunicaciones

Modelo Matemático:

- $y_j^{CSS\text{-}h}$ =1 si un CSS de tipo h se instala en j
- y_k^{LE} =1 si un LE se instala en k
- $x_{ij}^{BTS\text{-}CSS}$ =1 si el BTS en i se asigna a CSS en j
- $x_{ik}^{BTS\text{-}LE}$ =1 si el BTS en i se asigna a LE en k
- $x_{jk}^{CSS\text{-}LE}$ =1 si el CSS en j se asigna a LE en k
- $z_{jk}^{CSS\text{-}LE} \equiv$ número de módulos de CSS en j a LE en k
- $w_{jk}^{CSS\text{-}LE} \equiv$ tráfico de CSS en j a LE en k.

Giorgio Romanin Jacur, J.J.S.G.,"Optimisation of the Interconnecting Network of a UMTS Radio Mobile Telephone System", **European Journal of Operational Research**, 2001.

Problema de Telecomunicaciones

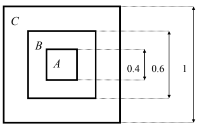
$$\min \sum_{j=1}^{m}\sum_{h=1,2} f_j^{CSS-h} y_j^{CSS-h} + \sum_{j=1}^{p} f_k^{LE} y_k^{LE} + \sum_{i=1}^{n}\sum_{j=1}^{m}(c_{ij}^{BTS-CSS} e_i^{BTS} + f_i^{BTS-CSS}) x_{ij}^{BTS-CSS} +$$
$$\sum_{i=1}^{n}\sum_{k=1}^{p}(c_{ik}^{BTS-LE} e_i^{BTS} + f_i^{BTS-LE}) x_{ik}^{BTS-LE} + \sum_{j=1}^{m}\sum_{k=1}^{p} c_{jk}^{CSS-LE} z_{jk}^{CSS-LE}$$

s.a:

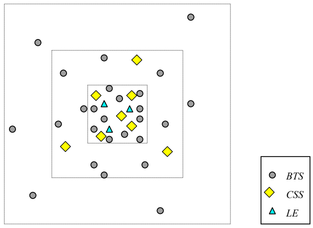
$$\sum_{j=1}^{m} x_{ij}^{BTS-CSS} + \sum_{k=1}^{p} x_{ik}^{BTS-LE} = 1 \qquad \forall i$$
$$\sum_{i=1}^{n} T_i^{BTS} x_{ij}^{BTS-CSS} \leq \sum_{h=1,2} T_j^{CSS-h} y_j^{CSS-h} \qquad \forall j$$
$$\sum_{i=1}^{n} x_{ij}^{BTS-CSS} \leq \sum_{h=1,2} N_j^{CSS-h} y_j^{CSS-h} \qquad \forall j$$
$$\sum_{i=1}^{n} e_i^{BTS} x_{ij}^{BTS-CSS} \leq \sum_{h=1,2} E_j^{CSS-h} y_j^{CSS-h} \qquad \forall j$$
$$\sum_{i=1}^{n} d_i^{BTS} x_{ij}^{BTS-CSS} \leq Q \sum_{k=1}^{p} z_{jk}^{CSS-LE} \qquad \forall j$$
$$z_{jk}^{CSS\text{-}LE} \leq M_{jk} x_{jk}^{CSS\text{-}LE} \qquad \forall j,k$$

$$\sum_{j=1}^{m} w_{jk}^{CSS-LE} + \sum_{i=1}^{n} T_i^{BTS} x_{ik}^{BTS-CSS} \leq T_k^{LE} y_k^{LE} \qquad \forall k$$
$$\sum_{i=1}^{n} T_i^{BTS} x_{ij}^{BTS-CSS} = \sum_{k=1}^{p} w_{jk}^{CSS-LE} \qquad \forall j$$
$$w_{jk}^{CSS-LE} \leq F_{jk} x_{jk}^{CSS-LE} \qquad \forall j,k$$
$$\sum_{j=1}^{m} z_{jk}^{CSS-LE} + \sum_{i=1}^{n} e_i^{BTS} x_{ik}^{BTS-CSS} \leq E_k^{LE} y_k^{LE} \qquad \forall k$$
$$\sum_{h=1,2} y_j^{CSS-h} \leq 1 \qquad \forall j$$
$$\sum_{k=1}^{p} x_{jk}^{CSS-LE} = \sum_{h=1,2} y_j^{CSS-h} \qquad \forall j$$

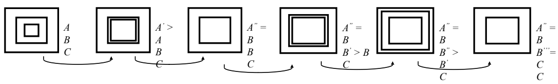
Problema de Telecomunicaciones



División del área geográfica:
•densidad de clientes
•tráfico en Erlangs por cliente
•densidad de elementos
•radio de las celdas.



Distribución final de los elementos



Problema de Telecomunicaciones

	Number of subscribers per km²	*Traffic per subscriber (mErl)*	*Radius of the cell (km)*
Region A	750	70	0.4
Region B	500	50	0.8
Region C	250	35	1.2

Device	***Cost (€)***
Basic *BTS*	100,000
Isolated *BTS*	150,000
Simple CSS	450,000
Complex CSS	600,000
LE	1,400,000

Term	***Cost (€)***
Connection fee per two ends	1,454
Fixed monthly charge per two ends	764
Variable monthly charge per 100m	2.9

Capacity	***Distance***	***Cost (€)***
4×2 Mbps	0-15 km	24,243
16×2 Mbps	0-15 km	29,092
64×2 Mbps	15-70 km	334,000

Term	***Distance***	***Cost (€)***
Connection fee per two ends	-	2,254
Fixed monthly charge per two ends	-	764
Variable monthly charge depending on the distance	0-35 km	52 (€/km)
– – – – – – –	35-70 km	40 (€/km)
– – – – – – –	70-150 km	12 (€/km)
– – – – – – –	>150 km	7.2 (€/km)

Problema de Telecomunicaciones

	BTS	CSS	LE	Tabu UB	Best UB	Final LB	gap
A	100	12	4	19.850.255	19.850.255	19.606.797'0	1'23%
B	95	9	4	18.917.721	18.915.544	18.687.073'3	1'21%
C	110	14	4	23.215.028	23.214.196	21.560.353'6	7'12%
D	96	10	5	19.088.121	19.087.437	18.847.882'4	1'26%
E	105	10	5	20.683.960	20.680.389	20.523.362'4	0'76%
F	115	15	5	23.975.503	23.967.148	22.508.426'1	6'09%
G	100	14	5	19.840.342	19.840.342	19.580.270'7	1'31%
H	110	16	5	23.220.740	23.220.740	21.573.970'3	7'09%
I	100	25	5	19.838.083	19.835.722	19.592.028'3	1'23%
L	120	12	4	24.927.101	24.925.856	23.559.843'3	5'48%
M	90	9	3	18.179.351	18.178.546	17.804.722'9	2'06%
N	85	10	4	16.981.990	16.981.213	16.863.167'4	0'70%
O	100	10	3	19.850.259	19.849.892	19.603.163'5	1'24%
P	85	6	3	16.624.947	16.624.227	16.510.956'5	0'68%

Problema de Telecomunicaciones

"Gestión Óptima de Redes de Telecomunicaciones"

Aprobado en la XXII Comisión Mixta Permanente en Aplicación del Acuerdo Cultural entre España y Bélgica (Comunidad Francesa) para el período 2000-2002

Universidad de La Laguna
Ministerio de Asuntos Exteriores

Universidad Libre de Bruselas
Commissariat général aux Relations Internationales de la Communauté Française de Belgique

Optimizar en Transporte

- Problema:

Optimizar los recursos en problemas de transporte de mercancías y pasajeros.

Optimizar en TITSA

Proyecto PI2000/116 financiado por el "Gobiernos de Canarias" entre 2001-2003.

TITSA: 500 guaguas
1500 trabajadores
50 millones viajeros/año

Desarrollar un sistema informático para la ayuda a la Toma de Decisiones en TITSA.

Creación de Turnos Óptimos

- Problema:

  Dado un reglamento que regula lo que se considera un turno (o servicio) válido para un conductor de TITSA, y teniendo presente los viajes que la empresa desea realizar cada día, elaborar una asignación de sus conductores a los viajes de manera que se minimice el número de servicios.

LINEA 019

LA LAGUNA - El Cardonal - Taco - SAN MATIAS

LA LAGUNA	SAN MATIAS
06'45	07'15
07'45	08'15
09'00	09'40
10'25	11'00
11'40	12'20
13'00	13'40
14'25	15'00
15'40	16'20
17'00	17'40
18'25	19'00
19'40	20'20
21'00	21'40

Creación de Turnos Óptimos

Convenio colectivo urbano:

Máximo de 312 horas de trabajo cada bloque de 8 semanas.

- Jornada continua
- máximo de 8 horas
- mínimo de 5 horas
- si "mañana", no terminar después de las 15:00 (lunes-viernes) y 15:15 (fin semana y festivos); si "tarde", no comenzar antes de las 12:00.
- al menos 25 minutos entre 3ª y 4ª hora de descanso
- Jornada partida (no más del 25%)
- complemento de 5 €/día
- nunca más de dos porciones
- no más de 4 horas entre las porciones (incluído el "toma" y "deje" del vehículo en el depósito) ni menos de 2 horas.
- Jornada "corre-turnos" (no más del 10%)

Creación de Turnos Óptimos

Convenio colectivo interurbano:

Máximo de 39 horas de trabajo cada bloque de 1 semana.

- Jornada continua
- máximo de 9 horas
- mínimo de 6 horas
- si "mañana", no terminar después de las 15:00 (lunes-viernes) y 15:15 (fin semana y festivos); si "tarde", no comenzar antes de las 12:00.
- al menos 25 minutos entre 3ª y 4ª hora de descanso
- Jornada partida (no más del 10%)
- complemento del 12% adicional
- nunca más de dos porciones
- no más de 8 horas entre las porciones (incluído el "toma" y "deje" del vehículo en el depósito) ni menos de 2 horas.
- Jornada nocturna (duración máxima 7:35 horas/día)

Creación de Turnos Óptimos

Normas de la empresa:

- Cada turno tiene un costo para la empresa que se obtiene de combinar:
  - tiempo de inactividad (5)
  - tiempo extra por descanso en jornadas partidas (2.5)
  - tiempo extra por descanso en jornadas continuas (1)
  - números de cambios de línea (8)
  - tiempo por debajo de la media (10)
  - tiempo por encima de la media (2).
- En jornadas continuas no mas de 30 minutos de descanso.
- No más de 5 minutos de inactividad entre dos viajes.

Creación de Turnos Óptimos

- Modelo Matemático:

Dados n viajes, generamos todos los m turnos válidos.

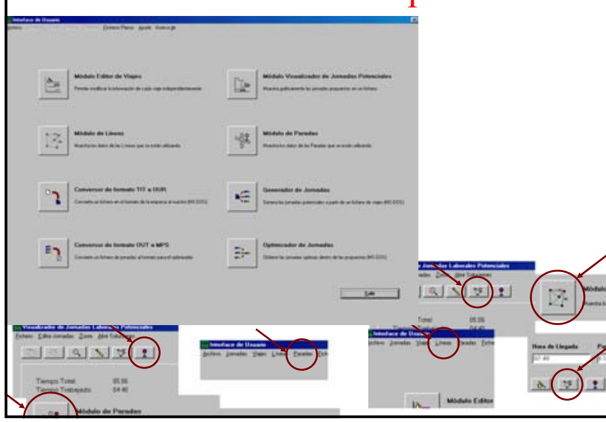
Sea $A=[a_{ij} : i=1...n , j=1...m]$ donde $a_{ij}=1$ si el turno j debe realizar el viaje i, y sea c_j el costo de j.

Consideremos $x_j=1$ si debemos elegir j, $=0$ otro caso.

$$\begin{aligned} \min\ & \Sigma_j\, c_j \cdot x_j \\ \text{s.a:}\ & \Sigma_j\, a_{ij} \cdot x_{ij} = 1 && \text{para todo viaje } i \\ & x_j \in \{0,1\} && \text{para todo turno } j \end{aligned}$$

Inconveniente: aún con n pequeño, m puede ser grande

Creación de Turnos Óptimos



Creación de Turnos Óptimos

Para evaluar nuestro algoritmo lo hemos aplicado a las líneas de microbuses en Santa Cruz:

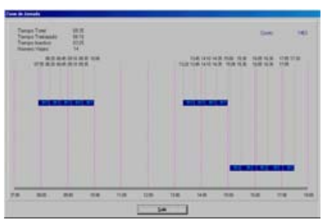
Línea	Descripción
902	Plaza España-Barrio Nuevo
912	Estación-Los campitos
913	Circunvalación
914	Circunvalación
915	Murillo-Tío Pino-Camino del Hierro
916	Estación-Los Valles
917	Estación-Valleseco

Creación de Turnos Óptimos

Hemos creado 10 ejemplos para analizar:

Fichero	Líneas	Partes	Viajes	Total Jornadas
A1	913-914	1	139	1.559
A2	913-914	1 + 2	139	35.697
B1	916	1	16	0
B2	916	1 + 2	16	0
C1	912-917	1	67	278
C2	912-917	1 + 2	67	3.325
D1	902-915	1	115	1.187
D2	902-915	1 + 2	115	134.589
E1	913-914-912-917	1	206	21.679
E2	913-914-912-917	1 + 2	206	839.752

Ejemplo	J posibles	J favorables	Dur. Media	Viajes	cálculo (seg)	costo
A1	1.559	49	7:17	136	0,5	20.857
A2	35.697	53	7:38	139	13	17.352
C1	278	6	7:03	51	0,02	3.091
C2	3.325	6	7:47	67	1,49	1.822
D1	1.187	14	7:01	97	0,35	8.140
D2	134.589	11	7:25	113	108,18	6.141
E1	21.679	55	7:25	206	9,19	20.937
E2	839.752	55	7:44	206	930,3	15.407



Otros problemas en TITSA

Las nuevas balizas recogen información precisa sobre las demandas de los clientes que permitirán optimizar las frecuencias de las guaguas en cada línea.

Se están instalando nuevos sistemas que dan más información, y por tanto que permiten optimizar mejor los recursos si se estudian.

Otros problemas en Transporte

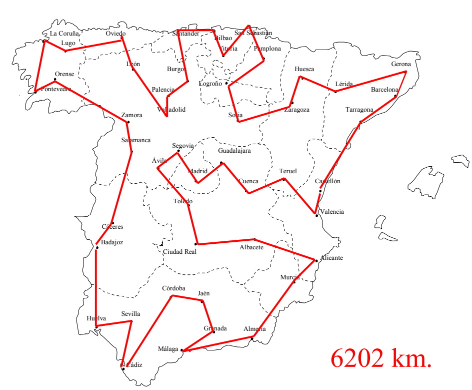
- Problema del Vendedor (TSP)
  (clásico)
- Problema del Vendedor Generalizado
  (con Paolo Toth y Matteo Fischetti)
- Problema del Comprador
  (con Jorge Riera Ledesma y Gilbert Laporte)
- Problema del Vendedor con Recogidas y Entregas
  (con Hipólito Hernández Pérez)
- Problema de Localización de Ciclos
  (con Inmaculada Rodríguez Martín y Martine Labbé)

Proyecto Nacional TIC2000-1750-C06

(Ministerio de Ciencia y Tecnología, enero 2001-diciembre 2003)

Valencia, ULL, Murcia, Elche, Madrid, Barcelona

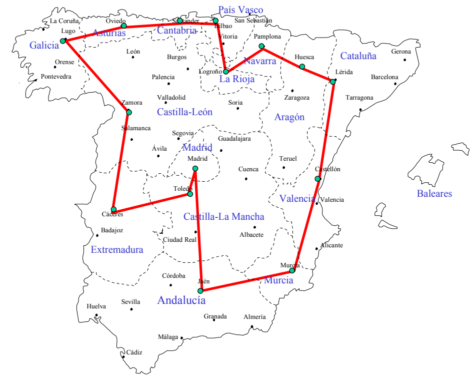
Problema del Vendedor (TSP)



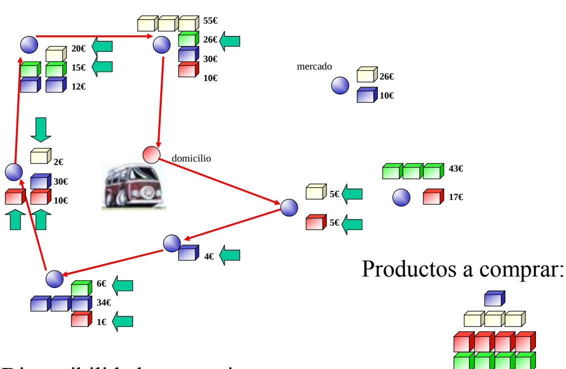
Problema del Vendedor (TSP)

#ciudades	Año	Autores
42	1954	Dantzig, Fulkerson, Johnson
48	1962	Held, Karp
57	1964	Karg, Thompson
64	1971	Held, Karp
80	1974	Helbig, Hanon, Krarup
100	1975	Camerini, Fratta, Maffioli
120	1980	Grötschel
318	1980	Crowder, Padberg, Hong
532	1987	Padberg, Rinaldi
666	1991	Grötschel, Holland
2392	1991	Padberg, Rinaldi
7397	1994	Applegate, Bixby, Chvátal, Cook
13509	1998	Applegate, Bixby, Chvátal, Cook
15112	2001	Applegate, Bixby, Chvátal, Cook

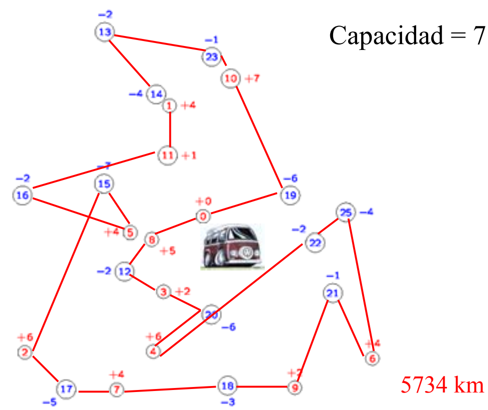
Problema del Vendedor **Generalizado (GTSP)**



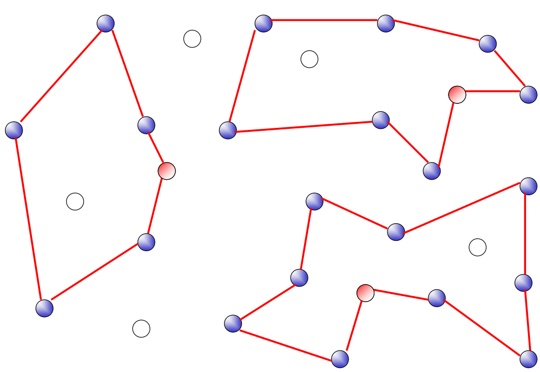
Problema del Comprador (TPP)



Problema con Recogidas-Entregas (PDTSP)



Problema de Localización de Ciclos



Optimización en Institutos de Estadística

Instituto Nationale di Statistica (Italia)
Statistics Netherlands (Holanda)
Office for National Statistics (UK)
University of Plymouth (UK)
University of Southampton (UK)
The Victoria University of Manchester (UK)
Statistisches Bundesambt (Alemania)
Universidad de La Laguna (España)
Institut d´Estadistica de Catalunya (España)
Institut National de Estadística (España)

IST-2000-25069 ; 1 enero 2001 – 31 diciembre 2003
"Computational Aspects of Statistical Confidentiality"
Costo total: 2.792.061 euros (=465 millones de ptas.)

Confidencialidad Estadística

* Microdatos
  - recodificación
  - perturbación
* Tablas Estadísticas
  - supresión de celdas
  - redondeo controlado

Supresión de Celdas

	Región A	Región B	Región C	TOTAL
Actividad I	20	50	10	80
Actividad II	~~8~~	19	~~22~~	49
Actividad III	~~17~~	32	~~12~~	61
TOTAL	45	101	44	190

Típicamente hacen falta supresiones secundarias, pero minimizando la pérdida de información.

Pérdida=37

menor valor = 5
mayor valor = 29

PrInf = 17
PrSup=7
PrAnc=24

					Mínimo	máximo	Protección inferior	Protección superior	Ancho de protección	Pérdida de información
Patrón 1	**Región A**	**Región B**	**Región C**	**TOTAL**						
Actividad I	20	50	10	80						
Actividad II	*	19	*	49						
Actividad III	*	32	*	61						
TOTAL	45	101	44	190	5	29	17	7	24	37
Patrón 2	**Región A**	**Región B**	**Región C**	**TOTAL**						
Actividad I	*	50	*	80						
Actividad II	*	19	*	49						
Actividad III	17	32	12	61						
TOTAL	45	101	44	190	2	30	20	8	28	38
Patrón 3	**Región A**	**Región B**	**Región C**	**TOTAL**						
Actividad I	20	*	*	80						
Actividad II	8	*	*	49						
Actividad III	17	32	12	61						
TOTAL	45	101	44	190	0	32	22	10	32	79
Patrón 4	**Región A**	**Región B**	**Región C**	**TOTAL**						
Actividad I	20	*	*	80						
Actividad II	*	*	*	49						
Actividad III	*	32	*	61						
TOTAL	45	101	44	190	0	40	22	18	40	116
Patrón 5	**Región A**	**Región B**	**Región C**	**TOTAL**						
Actividad I	20	50	10	80						
Actividad II	8	19	*	*						
Actividad III	17	32	*	*						
TOTAL	45	101	44	190	0	34	22	12	34	122
Patrón 6	**Región A**	**Región B**	**Región C**	**TOTAL**						
Actividad I	20	50	10	80						
Actividad II	8	19	*	*						
Actividad III	17	32	12	61						
TOTAL	45	101	*	*	0	+ infinito	22	+ infinito	+ infinito	283

Supresión de Celdas

- INPUT: el instituto de estadística nos da:
  * valores nominales: $y=[y_i:i=1,\ldots,n]$
  * Ecuaciones: $Ay=b$
  * pérdida de información: $w=[w_i:i=1,\ldots,n]$
  * cotas: $lb=[lb_i:i=1,\ldots,n]$ $ub=[ub_i:i=1,\ldots,n]$
  * celdas sensibles: PS y niveles de protección.
- OUTPUT: nosotros debemos devolver un conjunto de celdas $SS\subseteq\{1,\ldots,n\}$:
  * se garantizan los niveles de protección
  * se minimiza la pérdida de información

Computational results on a PC Pentium 266 Mhz

name	type	cells	links	p	pl	pl_0	t_0	HEU_1	t_1
CBS1	41 × 31	1271	72	3	6	6	0.3	73.79	0.4
CBS2	183 × 61	11163	244	2467	4934	2	6.1	0.00	0.2
CCSR	359 × 46	16514	405	4923	9846	54	36.3	0.00	24.6
CBS3	6 × 8 × 13	624	230	17	34	26	0.1	5.91	0.3
CBS4	6 × 33 × 8	1584	510	146	292	201	1.3	1.68	7.5
CBS5	6 × 8 × 8 × 13	4992	2464	517	1034	947	119.9	90.38	1533.8
USDE1	linked	1254	1148	165	330	320	0.5	30.21	16.8
USDE2	linked	1141	1000	310	620	572	8.9	33.18	29.8
USDE1a	linked	1254	1148	165	330	322	0.8	26.43	17.1
USDE1b	linked	1254	1148	165	330	322	0.8	27.31	17.3

name	r-HEU	r-LB	r-time	optimum	sup	node	iter	time	cap	bri	cov
CBS1	29.13	2.91	5.1	103	5	5	75	9.4	80	184	92
CBS2	0.00	0.00	6.4	403	2	1	1	6.4	2	0	0
CCSR	0.00	0.00	61.3	256	27	1	1	61.3	27	0	0
CBS3	1.88	1.00	8.4	22590362	27	32	416	66.0	251	504	523
CBS4	1.25	0.20	40.9	186433	51	19	70	123.7	215	52	69
CBS5	0.00	0.00	4924.1	6312	261	1	76	4924.1	2139	0	0
USDE1	3.42	0.55	626.6	2228523	254	22	202	1187.0	1195	781	86
USDE2	1.38	2.68	702.0	4643198	181	46	238	2397.2	1266	364	96
USDE1a	1.43	1.92	689.1	2325788	273	97	473	2614.6	1517	923	196
USDE1b	1.25	1.21	670.6	2157	274	16	240	1311.5	1587	849	137

Mathematical Programming (1999)
Journal of American Statistical Association (2000)
Management Science (2001)

Redondeo Controlado

original	Región A	Región B	Región C	TOTAL
Actividad I	20	50	10	80
Actividad II	8	19	*22*	49
Actividad III	17	32	12	61
TOTAL	45	101	44	190

Patrón 1	Región A	Región B	Región C	TOTAL
Actividad I	20	50	10	80
Actividad II	10	20	*20*	50
Actividad III	20	30	10	60
TOTAL	50	100	40	190

Patrón 2	Región A	Región B	Región C	TOTAL
Actividad I	20	50	10	80
Actividad II	10	10	*20*	40
Actividad III	10	40	20	70
TOTAL	40	100	50	190

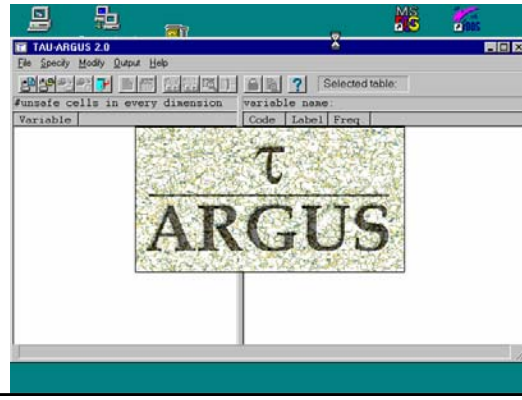
Supresión parcial: NUEVO METODO

	Región A	Región B	Región C	TOTAL
Actividad I	20	50	10	80
Actividad II	8	19	22	49
Actividad III	17	32	12	61
TOTAL	45	101	44	190

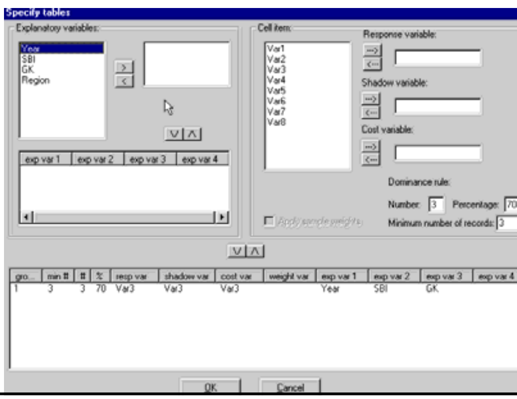
Patrón 1	Región A	Región B	Región C	TOTAL
Actividad I	[18...24]	50	[6...12]	80
Actividad II	[4...10]	19	*[20...26]*	49
Actividad III	17	32	12	61
TOTAL	45	101	44	190

Patrón 2	Región A	Región B	Región C	TOTAL
Actividad I	20	50	10	80
Actividad II	8	[22...15]	*[19...26]*	49
Actividad III	17	[29...36]	[8...15]	61
TOTAL	45	101	44	190

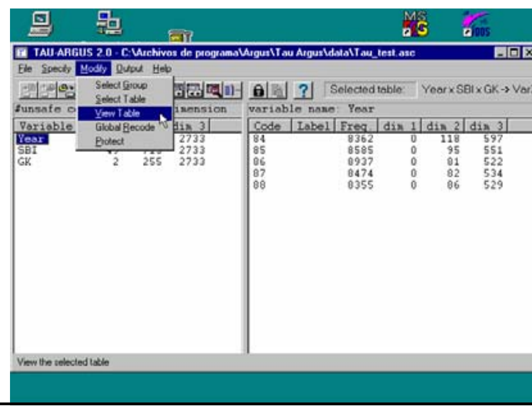
Supresión de Celdas y Redondeo Controlado



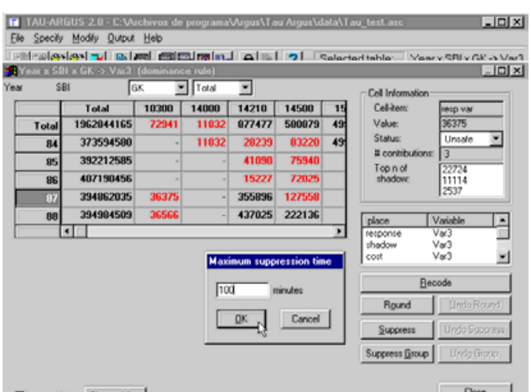
Supresión de Celdas y Redondeo Controlado



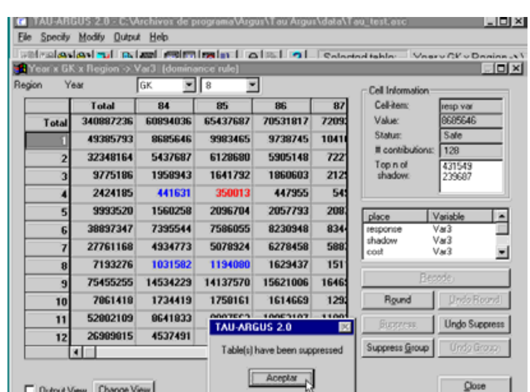
Supresión de Celdas y Redondeo Controlado



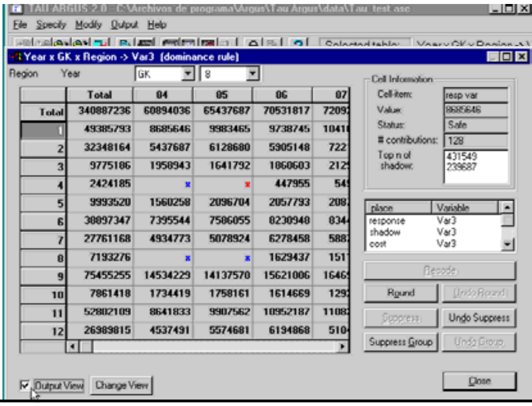
Supresión de Celdas y Redondeo Controlado



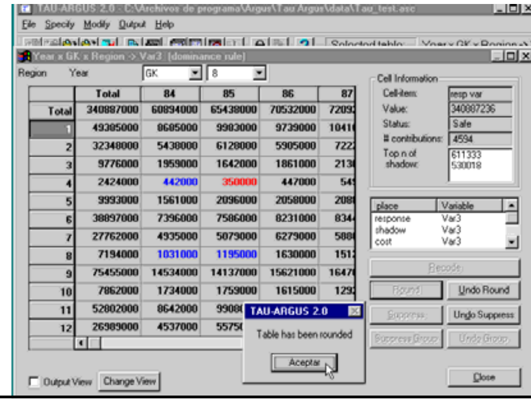
Supresión de Celdas y Redondeo Controlado



Supresión de Celdas y Redondeo Controlado



Supresión de Celdas y Redondeo Controlado



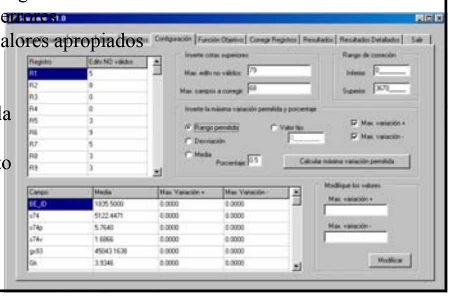
Técnicas de Edición e Imputación de Datos Estadísticos (TEIDE)

Ministerio de Ciencia y Tecnología: TIC 2002-00895
(Enero 2003 – Diciembre 2005)

Dado un conjunto de registros (ejemplo, Censo 2001)
y un conjunto de "reglas":

- localizar errores
- estimar valores apropiados

de manera que el nuevo conjunto de registros cumpla todas las reglas y difiera del conjunto original en el menor número de registros.



Técnicas de Edición e Imputación de Datos Estadísticos (TEIDE)

Modelo Matemático:

Para cada registro i y para cada campo j:

x_{ij}=1 si cambiamos i en j

z^-_{ij}= disminución respecto a a_{ij}

z^+_{ij}= aumento respecto a a_{ij}

Para cada registro i:

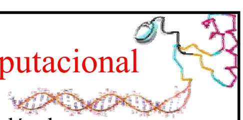
$$\min \quad \sum_j c_{ij}\, x_{ij} + \sum_j c^+_{ij}\, z^+_{ij} + \sum_j c^-_{ij}\, z^-_{ij}$$

$$\text{sujeto a: } A^i\,(a + z^+ - z^-) = b^i$$

$$0 \le z^+_{ij} \le M^+_{ij}\, x_{ij} \qquad \forall j$$

$$0 \le z^-_{ij} \le M^-_{ij}\, x_{ij} \qquad \forall j$$

Biología Molecular Computacional



* Nace de problemas analizando biomoléculas

* Biólogos en los 60s y 70s; Físicos y Matemáticos en los 70s y 80s; Informáticos en los 90s.

* Acelerado por los proyectos GENOMA

* Algunos libros:

- M.S.Waterman (1995) *Introduction to Computational Biology*, Chapman & Hall.
- J. Setubal, J. Meidanis (1997) *Introduction to Computational Molecular Biology*, PWS Publishing.
- P.A. Pevzner (2000) *Computational Molecular Biology: An Algorithmic Approach*, MIT Press.

Biología Molecular Computacional

- Problemas de alinear una secuencia:

molécula DNA ≡ secuencia de alfabeto con 4 letras

proteína (aminoácidos) ≡ secuencia de alfabeto con 20 letras

Se pretende comparar diferentes DNAs o proteínas.

Para ello se coloca una sobre otra y se rellenan con espacios para que tengan igual longitud. Dos reorganizaciones se diferencian mediante una función de *distancia*: en DNA, el número de letras distintas menos el número de espacios; en proteínas se le dan pesos cuando dos letras coinciden y penalizaciones a los bloques de espacios.

Biología Molecular Computacional

- Modelo Matemático para alinear una secuencia:

Sea A alfabeto finito y $S=\{s_1,\ldots,s_k\}$ secuencias finitas.

Sea $A'=A\cup\{-\}$. Una alineación $S'=\{s'_1,\ldots,s'_k\}$:

* las cadenas en S' tienen todas igual longitud

* ignorando la letra -, cada s'_i coincide con s_i.

Por tanto una "alineación" es una matriz con *k* filas.

Dada una función de distancia $d(s'_i, s'_j)$, se busca la alineación con menor suma/máximo de distancias.

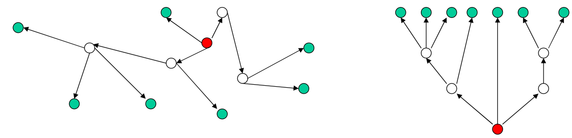
k=2 puede fácilmente resolverse en orden cuadrático, pero estudios en 1994 encontraron órdenes lineales.

k>2 suele ser NP-difícil y se usa Teoría de Grafos

Biología Molecular Computacional

- Problemas de árboles de Evolución:

Se pretende reconstruir la Evolución asumiendo que sigue una estructura en árbol con un ancestro como "raíz" y con los datos contemporáneos en "hojas".



Los datos contemporáneos pueden venir:

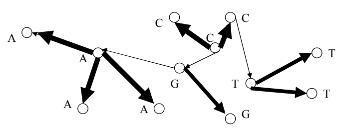
- alineaciones, donde cada secuencia es una hoja
- matriz de distancias $d(s'_i, s'_j)$ entre hojas

Biología Molecular Computacional

- Modelo Matemático: (modelo del filogén perfecto)

Sea *k* el número de especies, cada una caracterizada por *m* caracteres (una de *r* posibles letras: nucleótidos A,C,G,T).

Un árbol es compatible con un carácter cuando los nodos con igual letra en dicho carácter forman un subárbol.



Un árbol se dice filogén perfecto cuando todos los *m* caracteres son compatibles con él.

Biología Molecular Computacional

El "Perfect Phylogeny Problem" consiste en saber si existe un árbol filogén perfecto y en caso afirmativo construirlo.

El "Parsimony Problem" consiste en buscar un árbol de que contenga la menor cantidad de árboles en los bosques inducidos por los distintos caracteres. (=mínimo número de mutaciones para explicar las especies contemporáneas).

La idea de que ha pasado la misma cantidad de tiempo (debido a la existencia del ancestro) lleva al estudio de problemas de árboles donde todas las hojas distan igual de su raíz (árboles ultrametricos)

Dado un árbol con *k* hojas y *m* nodos internos, cada hoja etiquetada con una secuencia s_i, encontrar una alineación S' de *k*+*m* cadenas válida para S y que induzca en el árbol la menor suma de distancias entre hojas y ancestro.

Biología Molecular Computacional

- Problema: Genoma

El orden de los genes en el DNA es resultado de la Evolución.

Deshaciendo el orden se podría derivar el DNA de ancestros.

Dado una secuencia ordenada de *n* genes, "invertir" es la operación de seleccionar un subconjunto de *m* continuos, invertir su orden y colocarlos en la posición inicial.

La "distancia de inversión" entre dos secuencias es el menor número de inversiones que transformar una en la otra.

Puesto que la dirección de los genes es relevante en el DNA, conviene pensar en secuencias de genes con signo: invertir el orden de un trozo conlleva cambiar el signo de ellos.

El problema es calcular el menor número de inversiones con signo para pasar de una permutación con signo a otra dada.

Biología Molecular Computacional

Problema: Proyecto Genoma Humano

Determinar la localización de los genes en el DNA y la secuencia de los nucleótidos del genoma humano: sobre 3 billones de letras en 23 cromosomas.

Las técnicas Bioquímicas dan información sobre fragmentos (≈500 fragmentos) que deben alinearse (mapping).

El problema de reconstruir la secuencia es:

dada una colección de fragmentos Γ y una razón de error $0 \le \varepsilon < 1$, encontrar la secuencia S más corta tal que cada fragmento en Γ, o su inverso, esté en S con error $\le \varepsilon|\Gamma|$

Asumiendo que los fragmentos están en el orden adecuado y que no hay error, entonces el problema es NP-difícil pero hay muchos algoritmos heurísticos bastante eficientes.

CONCLUSION

La Optimización Matemática es una Ciencia que ofrece herramientas (modelos y algoritmos) útiles para ayudar a resolver problemas reales en Economía, Ingeniería, Biología, Física, Química, etc.

http://webpages.ull.es/users/jjsalaza
jjsalaza@ull.es

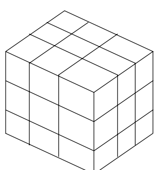
Índice:

- Introducción
- Optimizar una mezcla química
- Optimizar un divisor de tensión
- Optimizar encendidos de una máquina
- Optimizar una banda de música
- Algo sobre Telecomunicaciones
- Algo sobre Transporte
- Algo sobre Datos Estadísticos
- Algo sobre Biología Molecular

Optimizar Trazado de Líneas

Problema:

Supongamos las caras del cubo:



Línea = 3 celdas en sección vertical, horizontal, diagonal de una misma cara del cubo. Hay 48.

Pintar cada celda de "blanco" o "negro" de manera que se minimice el número de líneas con sus celdas de igual color.

Optimizar Trazado de Líneas

Modelo Matemático:

$x_i=1$ si la celda i se pinta de "blanco" y $=0$ si de "negro" para $i\in\{1,\ldots,54\}$,

$y_j=1$ si la línea j tiene sus tres celdas con igual color para $j\in\{1,\ldots,48\}$.

min	$y_1+\ldots+y_{48}$	
s.a:	$x_{i1}+x_{i2}+x_{i3}-y_j\leq 2$	para cada línea j compuesta
	$x_{i1}+x_{i2}+x_{i3}+y_j\geq 1$	por $\{i_1,i_2,i_3\}$
	$x_i\in\{0,1\}$	para cada celda $i\in\{1,\ldots,54\}$
	$y_j\in\{0,1\}$	para cada línea $j\in\{1,\ldots,48\}$.

Optimizar una Banda de Músicos

Modelo Matemático:

x_{ij}= 1 si la sinfonía i se coloca en la posicion j ; =0 otro caso
y_k= posición en la que comienza el músico k
z_k= posición en la que finaliza el músico k

min	$\sum_k c_k\cdot(z_k-y_k+1)$	
s.a:	$\sum_j x_{ij}=1$	para toda sinfonía i
	$\sum_i x_{ij}=1$	para toda posición j
	$x_{ij}\in\{0,1\}$	para toda sinfonía i y posición j
	$\sum_j j\cdot x_{ij}\geq y_k$	para toda i y músico k : $m_{ik}=1$
	$\sum_j j\cdot x_{ij}\leq z_k$	para toda i y músico k : $m_{ik}=1$