

RESTART: UN MÉTODO DE ACELERACIÓN DE SIMULACIONES DE SUCESOS INFRECIENTES

José Villén Altamirano
Universidad Politécnica de Madrid

ÍNDICE

- **Introducción**
- **Descripción de RESTART**
- **Eficiencia**
 - **Parámetros óptimos**
 - **Factores de Ineficiencia**
 - **Función de Importancia**
- **Ejemplos**
- **Conclusiones**

¿Qué es la simulación?

- **SIMULACIÓN:** Técnicas para imitar el funcionamiento de sistemas o procesos reales mediante programas de ordenador.
- **ÁREAS DE APLICACIÓN:**
 - Diseño y análisis de sistemas de producción
 - Evaluación de requerimientos hardware y software para un sistema de computación
 - Diseño de sistemas de comunicación
 - Diseño y operación de sistemas de transporte como aeropuertos, puertos o metro
 - Evaluación de diseños para la organización de servicios como hospitales u oficinas de correos
 - Análisis de sistemas económicos o financieros

SISTEMAS Y MODELOS

SISTEMA: Colección de entidades que interactúan entre sí para conseguir un determinado fin.

MODELO: Representación simplificada de un sistema que se utiliza para estudiar su comportamiento.

HIPÓTESIS SOBRE EL MODELO:

- Sencillas: posible obtener soluciones exactas o analíticas .
- Complejas: es lo más usual. Deben ser estudiados mediante simulación. Se obtienen soluciones aproximadas.

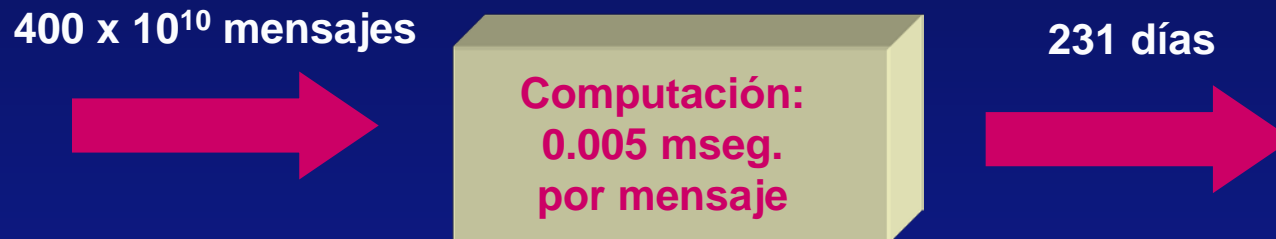


TIPOS DE MODELOS DE SIMULACIÓN

- Clasificación de los modelos de simulación:
 - Modelos estáticos frente a modelos dinámicos.
 - Modelos determinísticos frente a modelos estocásticos.
 - Modelos discretos frente a modelos continuos.
- Estudiaremos modelos de simulación
DINÁMICOS, ESTOCÁSTICOS Y DISCRETOS

EL PROBLEMA

- Estimar probabilidades de sucesos infrecuentes
- Ejemplo
 - Probabilidad de pérdida de un mensaje = 10^{-10}
 - Error relativo del intervalo de confianza : 10%



- Simulación normal es impracticable
- Se requieren técnicas de aceleración

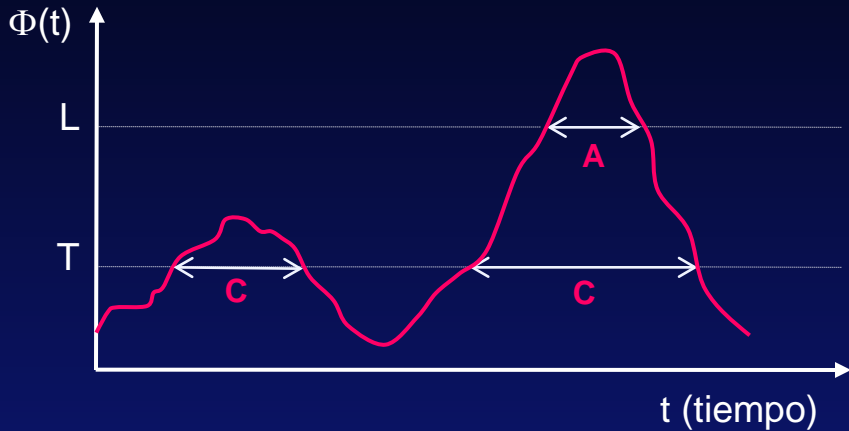
MUESTREO DE IMPORTANCIA

- Se quiere estimar $\theta = E[h(X)] = \int_{\Omega} h(x) p(x) dx$
- La idea es tomar muestras de otra v.a. X^* con función de densidad $p^*(x)$ de forma que el suceso raro ocurra más frecuentemente. Por ejemplo:
 - aumentando la tasa de llegadas en un sistema de colas
 - aumentando la tasa de fallo de cada componente en un estudio de fiabilidad.

$$\theta = \int_{\Omega} \frac{h(x) p(x)}{p^*(x)} p^*(x) dx = E[q(X^*)h(X^*)], \text{ siendo } q(x) = \frac{p(x)}{p^*(x)}$$

- Dificultad: elegir adecuadamente $p^*(x)$. Se requiere un estudio analítico para aplicarlo a cada caso particular.

FUNDAMENTOS DE RESTART (I)

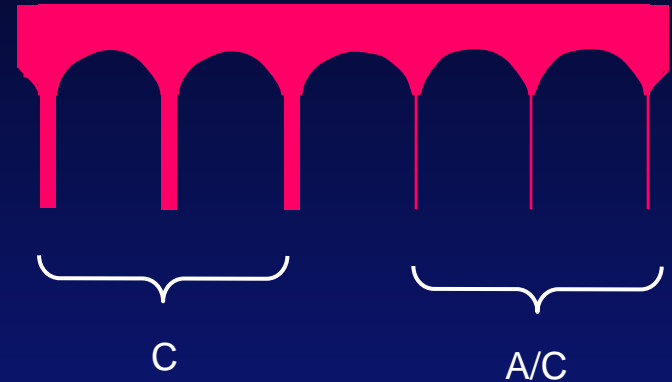


A : Suceso infrecuente

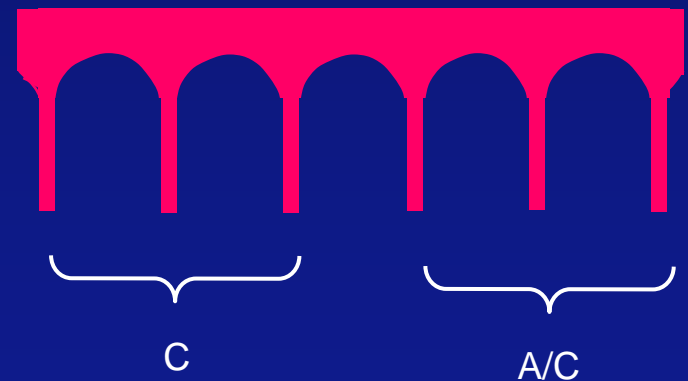
C : Estados con $\Phi \geq T$ $\begin{cases} C \supset A \\ 1 \gg P\{C\} \gg P\{A\} \end{cases}$

$$P\{A\} = P\{C\} \cdot P\{A/C\}$$

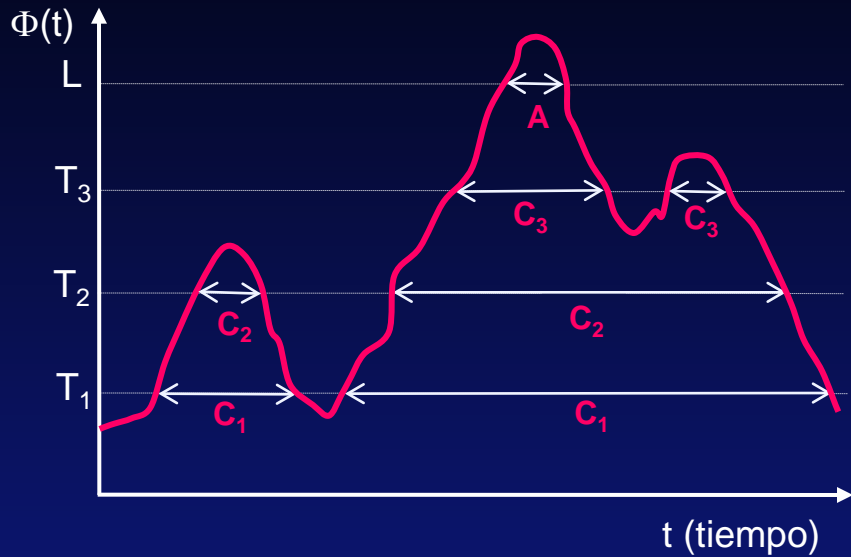
- Simulación normal:



- Con RESTART mono-umbral:



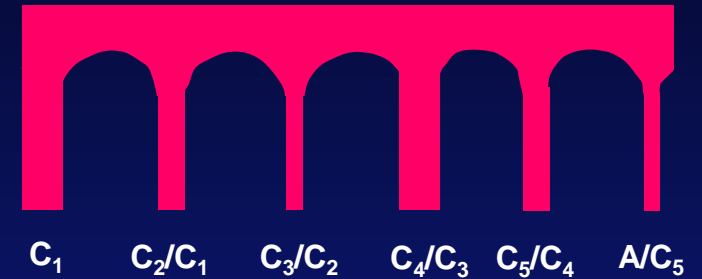
FUNDAMENTOS DE RESTART (II)



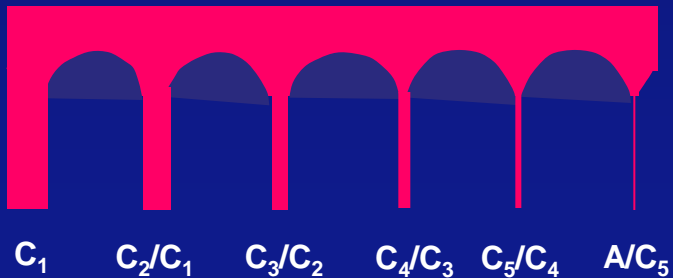
$$C_1 \supset C_2 \supset C_3 \supset \dots C_M \supset A$$

$$P\{A\} = P\{C_1\} \cdot P\{C_2/C_1\} \cdot \dots \cdot P\{A/C_M\}$$

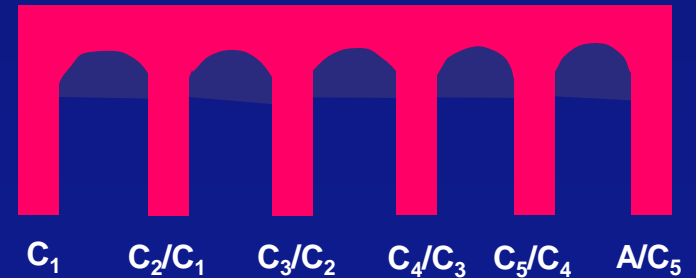
- Con RESTART mono-umbral:



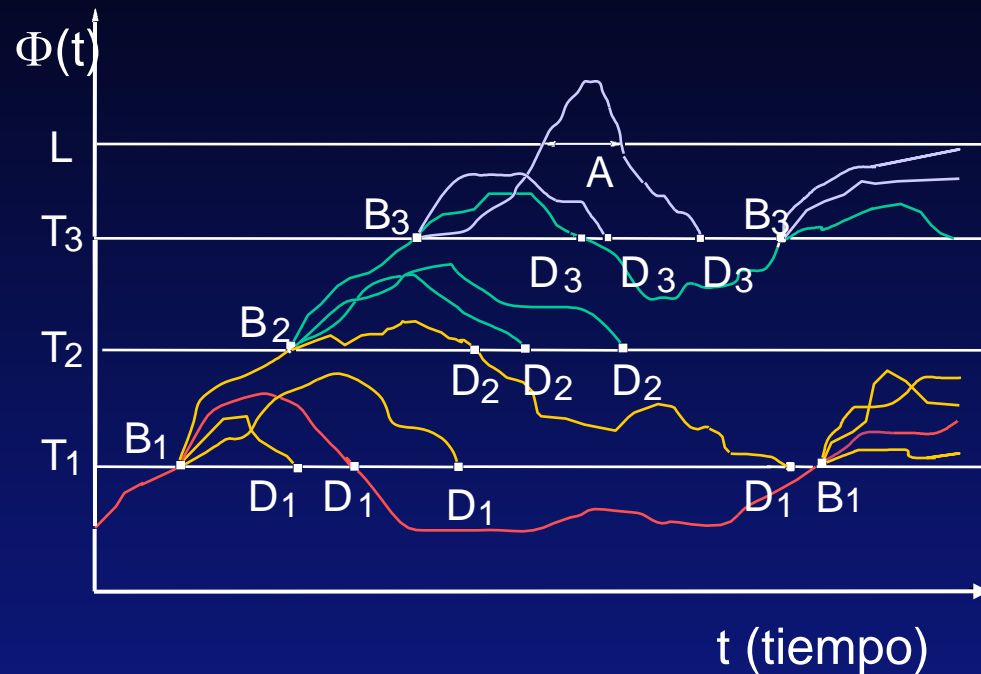
- Simulación normal:



- Con RESTART multi-umbral:



DESCRIPCIÓN DE RESTART (I)



Umbrales T_i definen: $P_i = \Pr\{\Phi \geq T_i\}$, $P_{i/j} = P_i / P_j$

Nº de reintentos en B_i : R_i

Multiplicador de muestreo en G_i : $r_i = \prod_{j=1}^i R_j$

$$\hat{P} = \frac{N_A}{N r_M}$$

RESTART: REpetitive **S**imulation **T**rials **A**fter **R**eaching **T**hresholds

Descripción de RESTART (II)



$$P = \Pr \{ A \} = \Pr \{ Q_2 \geq L \}$$

$$\Pr \{ C_i \} = \Pr \{ \Phi \geq T_i \}$$

$$\Phi = \frac{\ln \rho_2}{\ln \rho_1} Q_1 + Q_2$$

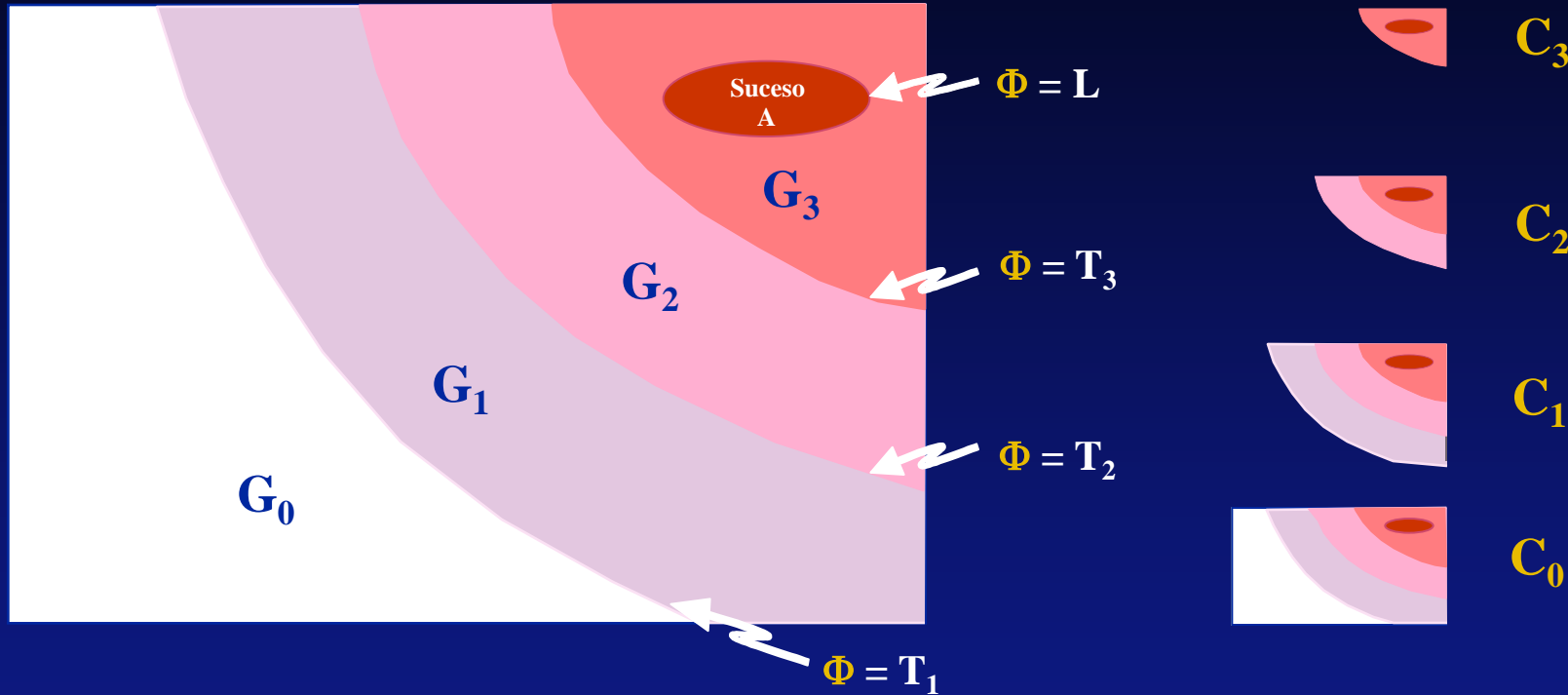
$$A \not\subset C_i \quad i = 1, \dots, M \quad \text{then: } P \{ A \} = \sum_{i=1}^M P \{ C_i \} P \{ A / C_i \}$$

$$\hat{P} = \sum_{i=1}^M \frac{N_{Ai}}{r_i \cdot N}$$

N : No. of simulated events (retrials not included)

N_{Ai} : No. of events A in retrials from sets C_i

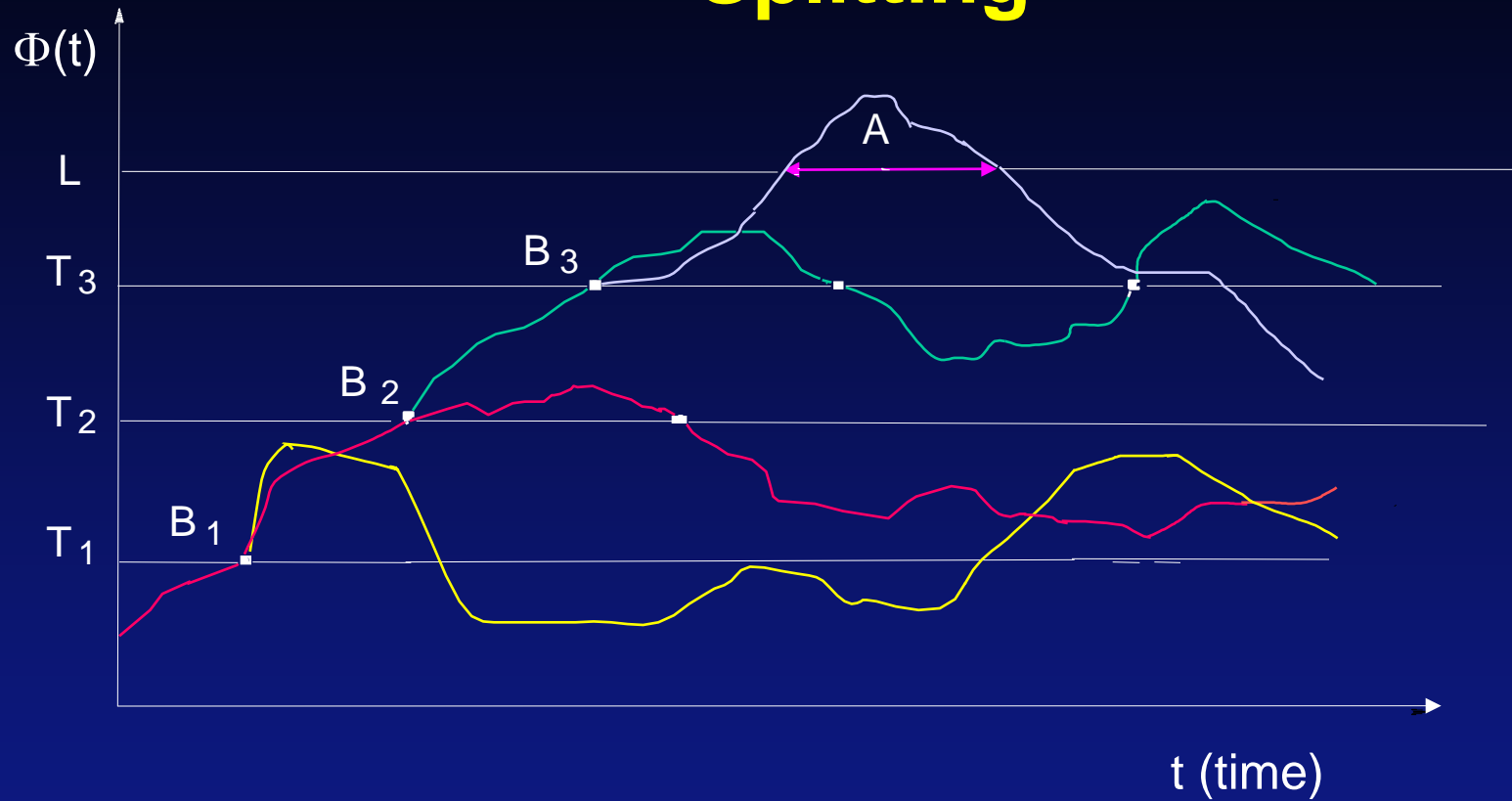
DESCRIPCIÓN DE RESTART (III)



Φ : Función de importancia
 T_i : umbrales
 $L = T_{M+1}$

G_i : Regiones de importancia
 C_i : Conjuntos de importancia
 $A = G_{M+1} = C_{M+1}$

Splitting



- Useful only for short samples. Inefficient for steady-state simulation

ESTIMADOR Y VARIANZA

El estimador de la probabilidad de que ocurra el suceso raro cuando sólo se puede alcanzar desde el último umbral es:

$$\hat{P} = \frac{N_A}{N r_M}$$

La varianza de este estimador es:

$$V(\hat{P}) = \frac{K_A P}{N} \left[\frac{1}{r_M} + \sum_{i=1}^M \frac{S_i P_{A/i} (R_i - 1)}{r_i} \right]$$

$$\text{Con } K_A = V(N_A^0) / E[N_A^0]$$

EFICIENCIA (I)

Coste de simulación para igual error relativo del intervalo de confianza:

Simulación normal

$$K \frac{1}{p}$$

RESTART

$$\frac{Ke^2}{4} \left(\ln \frac{1}{p} \right)^2 f_R f_T f_V f_O$$

$$Ganancia = \frac{4}{e^2} \frac{1/p}{\left(\ln \frac{1}{p} \right)^2} \frac{1}{f_R f_T f_V f_O}$$

Los factores $f \geq 1$ reflejan ineficiencia debida a:

f_R - R_i no óptimos

f_T - T_i no óptimos

f_V - varianza en B_i

f_O - consumo de tiempo del algoritmo

EFICIENCIA (II)

- Ganancia para $f_R = f_T = f_V = f_O = 1$

<u>P</u>	<u>Tiempo de computación</u>		<u>Ganancia</u>
	<u>Simulación normal</u>	<u>RESTART</u>	
10^{-4}	6 seg.	0.1 seg.	63
10^{-7}	1.7 horas	0.3 seg.	2×10^4
10^{-10}	2.4 meses	0.6 seg.	1×10^7
10^{-100}	2×10^{89} años	1 min.	1×10^{95}

- Suposición: 1 minuto para $P = 10^{-5}$ con simulación normal

REDUCCIÓN DE F_R

$$R_i = \frac{1}{\sqrt{P_{i+1/i-1}}} \Rightarrow r_i = \frac{1}{\sqrt{P_{i+1/1} \cdot P_{i/o}}} \Rightarrow f_R = 1$$

- Ajuste de r_i a su valor óptimo
- Elección de R_i para ajustar r_i

REDUCCIÓN DE F_T

Umbrales óptimos: $P_{i/i-1} = 1 \Rightarrow f_T = 1$

$$f_T = \frac{\left(\sum_{i=0}^M \frac{1 - P_{i+1/i}}{\sqrt{P_{i+1/i}}} + 1 \right)^2}{(-\ln P + 1)^2}, \text{ siendo } P_{i+1/i} = P(\Phi \geq T_{i+1}) / P(\Phi \geq T_i)$$

- Tomar el máximo número de umbrales posible, con la restricción de que R_i debe ser un número entero.
- No es crítico un ajuste preciso de $P_{i/i-1}$

P_{min}	Cota de f_T
0.5	1.03
0.1	1.53
0.01	4.62
0.001	20.91

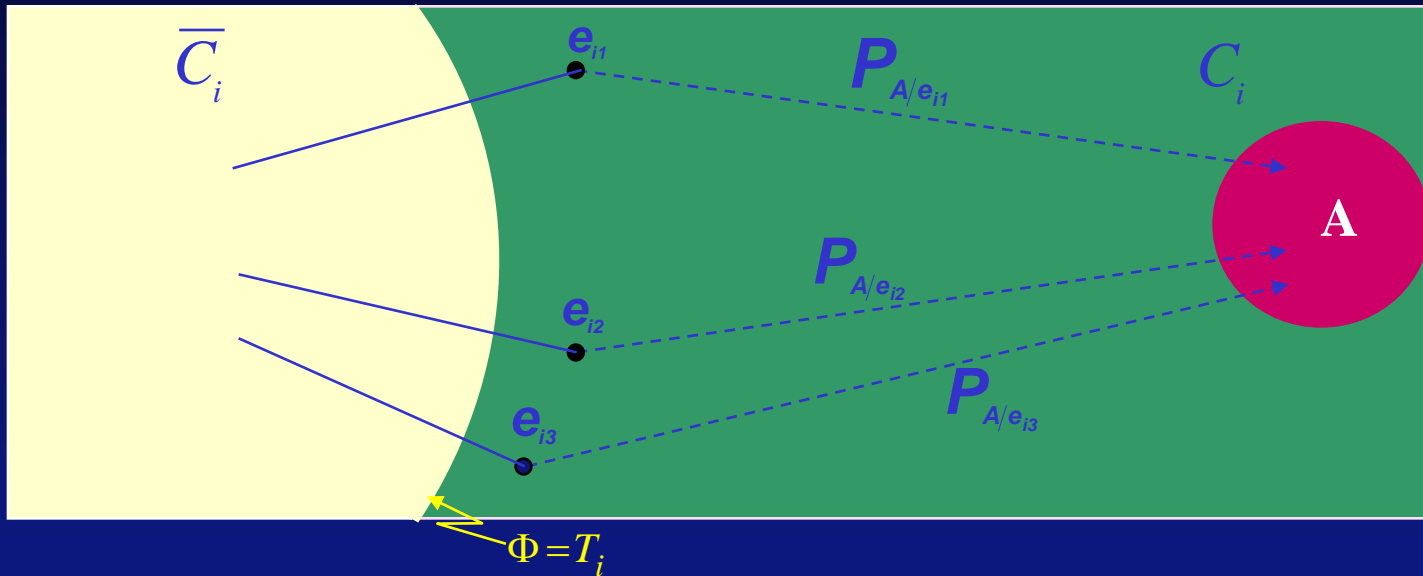
REDUCCIÓN DE F_0

- Afecta al tiempo de computación, no al n° de sucesos
- Debido al consumo de tiempo del algoritmo
 - **por suceso: evaluar Φ , comparar con T_i , ...**
 - **por reintento: salvar / restaurar el estado en B_i , revisar decisiones, ...**
- Directrices para reducir el tiempo consumido por reintento
 - **Volcado de memoria para salvar / restaurar el estado de B_i**
 - **Histéresis: reduce el factor a su raíz cuadrada**

REDUCCIÓN DE F_V (I)

$$f_v \leq \text{Max}(s_i)$$

$$s_i = \frac{a_i}{K_A} \left[K'_i + \frac{V(P_{A/e_{ij}}^*)}{(P_{A/i}^*)^2} \gamma_i \right]$$



- $e_{ij} \in \Omega_i$ (espacio de estados en B_i)
- $P_{A/e_{ij}}^*$ es el n° esperado de sucesos A en un reintento $[B_i, D_i)$ que empieza en e_{ij} . Se le llama importancia del estado e_{ij} .
- f_v es debido a la varianza entre las $P_{A/e_{ij}}$ de un mismo i

REDUCCIÓN DE F_V (III)

Ejemplo de la directriz 1



Suceso infrecuente : $q_2 \geq L$, por ej.: $L=40$

	Mala elección	Buena elección
Función de importancia	$\Phi = q_2$	$\Phi = aq_1 + q_2$, $a = 0.6$
Estados e_{ij} ($T_i = 20$)	$(0, 20), \dots, (18, 20), \dots, (33, 20), \dots$	$(0, 20), \dots, (18, 9), \dots, (33, 0)$
Valores de Eficiencia P_A / e_{ij}	<i>Muy diferentes</i> <i>Muy baja</i>	<i>Más similares</i> <i>Aceptable</i>

REDUCCIÓN DE F_V (IV)

Resultados

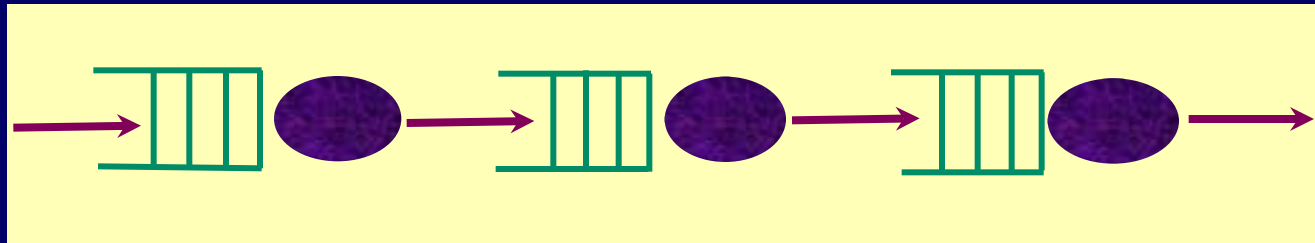
L	P	\hat{P}	Tiempo (minutos)	Ganancia (en sucesos)	f_V
20	2.8710^{-10}	$2.92 \cdot 10^{-10}$	0.25	$6 \cdot 10^6$	1.3
60	$2.36 \cdot 10^{-29}$	$2.39 \cdot 10^{-29}$	7	$5 \cdot 10^{23}$	3.1
100	$1.94 \cdot 10^{-48}$	$2.02 \cdot 10^{-48}$	59	$7 \cdot 10^{41}$	7.1

$$f_R \approx 1 \quad f_T \approx 1$$

$$f_v \approx \frac{\text{Ganancia teórica } (f_R = f_T = f_V = f_O = 1)}{\text{Ganancia real en sucesos}}$$

Importance Function (I)

- Three-queue Jackson tandem network



➤ Importance function (If $\rho_1 > \rho_2 > \rho_3$):

$$\Phi = Q_1 \frac{\ln \rho_1}{\ln \rho_3} + Q_2 \frac{\ln \rho_2}{\ln \rho_3} + Q_3$$

If $\rho_1 < \rho_2 < \rho_3$, or if $\rho_2 < \rho_1 < \rho_3$, $\Phi = Q_1 + Q_2 + Q_3$

If $\rho_2 < \rho_3 < \rho_1$, $\Phi = Q_1 \frac{\ln \rho_1}{\ln \rho_3} + Q_2 + Q_3$

If $\rho_1 < \rho_3 < \rho_2$, or if $\rho_3 < \rho_1 < \rho_2$, $\Phi = \frac{\ln \rho_2}{\ln \rho_3} (Q_1 + Q_2) + Q_3$

Importance Function (II)

General Jackson networks

$$P = \Pr \{ Q_{tg} \geq L \}$$

$$\Phi = \sum_{i=1}^H \alpha_{1i} \frac{\ln(\rho_{tg} / \rho_{tgi}^*)}{\ln \rho_{tg}} Q_{1i} + \sum_{j=1}^K \alpha_{2j} \frac{\ln(\rho_{tg} / \rho_{tgj}^\perp)}{\ln \rho_{tg}} Q_{2j} + Q_{tg}$$

$$\rho_{tg} = \frac{\gamma_{tg} + \sum_{j=1}^K \lambda_j p_{jtg}}{\mu_{tg}} = \frac{\lambda_{tg}}{\mu_{tg}} \quad \rho_{tgi}^* = \frac{\gamma_{tg} + \sum_{j=1}^K \text{Min}[\lambda_{2j} + (\mu_{1i} - \lambda_{1i}) p_{ij}, \mu_{2j}] p_{jtg} + \lambda_{tg} p_{tgtg}}{\mu_{tg}}$$

$$\rho_{tgj}^\perp = \frac{\gamma_{tg} + \mu_{2j} p_{jtg} + \sum_{l \neq j} \lambda_{2l} p_{ltg} + \lambda_{tg} p_{tgtg}}{\mu_{tg}}$$

$$\alpha_{1i} = 1 + \frac{\sum_{l \neq i} \gamma_{1l} \sum_{j=1}^K p_{lj} p_{jtg} + \sum_{j=1}^K \gamma_{2j} p_{jtg} + \gamma_{tg}}{\mu_{1i} \sum_{j=1}^K p_{ij} p_{jtg}}; \quad \alpha_{2j} = 1 + \frac{\sum_{i=1}^H \gamma_{1i} \sum_{l \neq j} p_{il} p_{ltg} + \sum_{l \neq j} \gamma_{2l} p_{ltg} + \gamma_{tg}}{\mu_j p_{jtg}}$$

Simulation Results (II)

Example 1: Jackson network with 7 nodes:

■ Rare event probability: $P(Q_t \geq 70) = 8.8 \cdot 10^{-35}$;

$$\rho_i = 0.3262; \quad \Phi = a \sum_{i=1}^4 Q_i + b \sum_{j=5}^6 Q_j + Q_t$$

ρ_1	ρ_2	a	b	Events (millions)	Time (minutes)	Gain (events)	f_V	Gain (time)	f_0
0.50	0.41	0.30	0.53	17	5.7	$8.9 \cdot 10^{29}$	1.8	$2.0 \cdot 10^{29}$	4.4
0.32	0.41	0.46	0.54	10	2.9	$1.5 \cdot 10^{30}$	1.1	$3.7 \cdot 10^{29}$	4.1
0.28	0.30	0.50	0.61	6	1.9	$2.5 \cdot 10^{30}$	0.7	$6.0 \cdot 10^{29}$	4.2

■ In the three cases the best results are obtained with the coefficients a and b obtained with the formulas of the paper.

● **Robustness:** Acceptable results are obtained for coefficient a and b up to 20% lower or greater than optimal ones.

Simulation Results (III)

Example 2: Large Jackson network with 15 nodes, four of them at distance 3 from the target node. The target node has the lowest load.

$$P(Q_t \geq 70) = 1.6 \cdot 10^{-33}; \quad \rho_t = 0.34; \quad \Phi = 0 \sum_{i=1}^4 Q_i + b \sum_{j=5}^9 Q_j + c \sum_{k=10}^{14} Q_k + Q_t$$

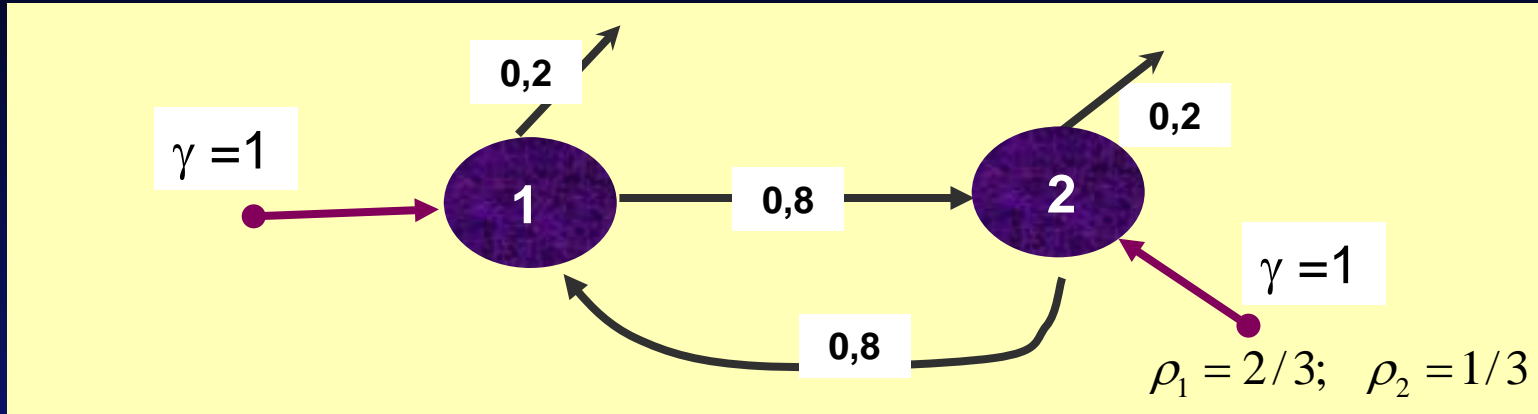
- 20 minutes of computational time for estimating this probability with a relative error of 10%.

$$P(Q_t \geq 30) = 8.8 \cdot 10^{-15}$$

- 4 minutes of computational time for estimating this probability with the same importance function and the same relative error.

Simulation Results (IV)

• *Example 3: Network with strong feedback: 2-node Jackson network.*



$$P(Q_2 \geq 70) = 4.0 \cdot 10^{-34}; \quad \Phi = 0.369Q_1 + Q_2$$

- 30 minutes of computational time for estimating this probability with a relative error of 10%.

$$P(Q_2 \geq 30) = 4.9 \cdot 10^{-15};$$

- 4 minutes of computational time for estimating this probability with the same importance function and the same relative error.

Importance Function Non-Jackson networks

- Would be fit for other networks the importance function derived for Jackson networks? Or at least, would be easy to modify it?
- We are studying networks with Erlang, Hyperexponential, ... interarrival and/or service times. For most of the cases studied (up to now), the importance function derived for Jackson networks is valid for non-Jackson networks with some modifications. Las fórmulas siguen siendo válidas si cambiamos la carga real por la carga “efectiva”.
- **REMARK:** the choice of the importance function affects to the efficiency of RESTART though it does not affect to the correctness of the estimates.

Importance function: highly-dependable system (I)

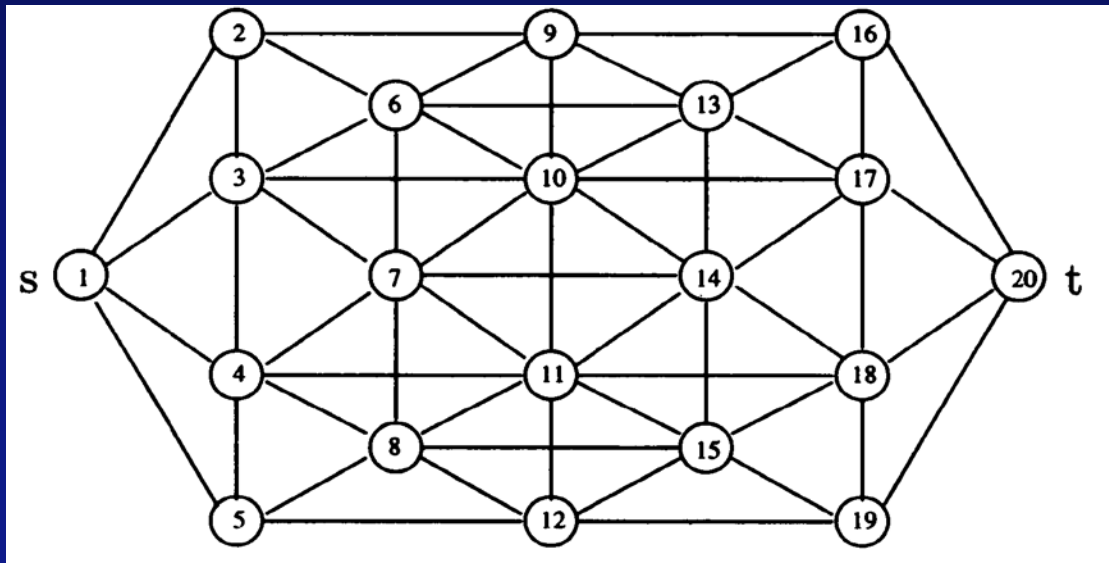
Objective: All the system states of the same threshold, X_i , must have similar importance.

- In a general system there are minimal cutsets with different cardinalities.
- It is more probable that system failure will be due to the failure of all the components of a minimal cutset with the lowest cardinality.
- The “distance” to the system failure is related with the number of components that remain operational in the cutset with the lowest number of operational components.
 - **Importance function:** $\Phi(t) = cl - oc(t)$

Importance function: highly-dependable system (II)

Network with redundancies

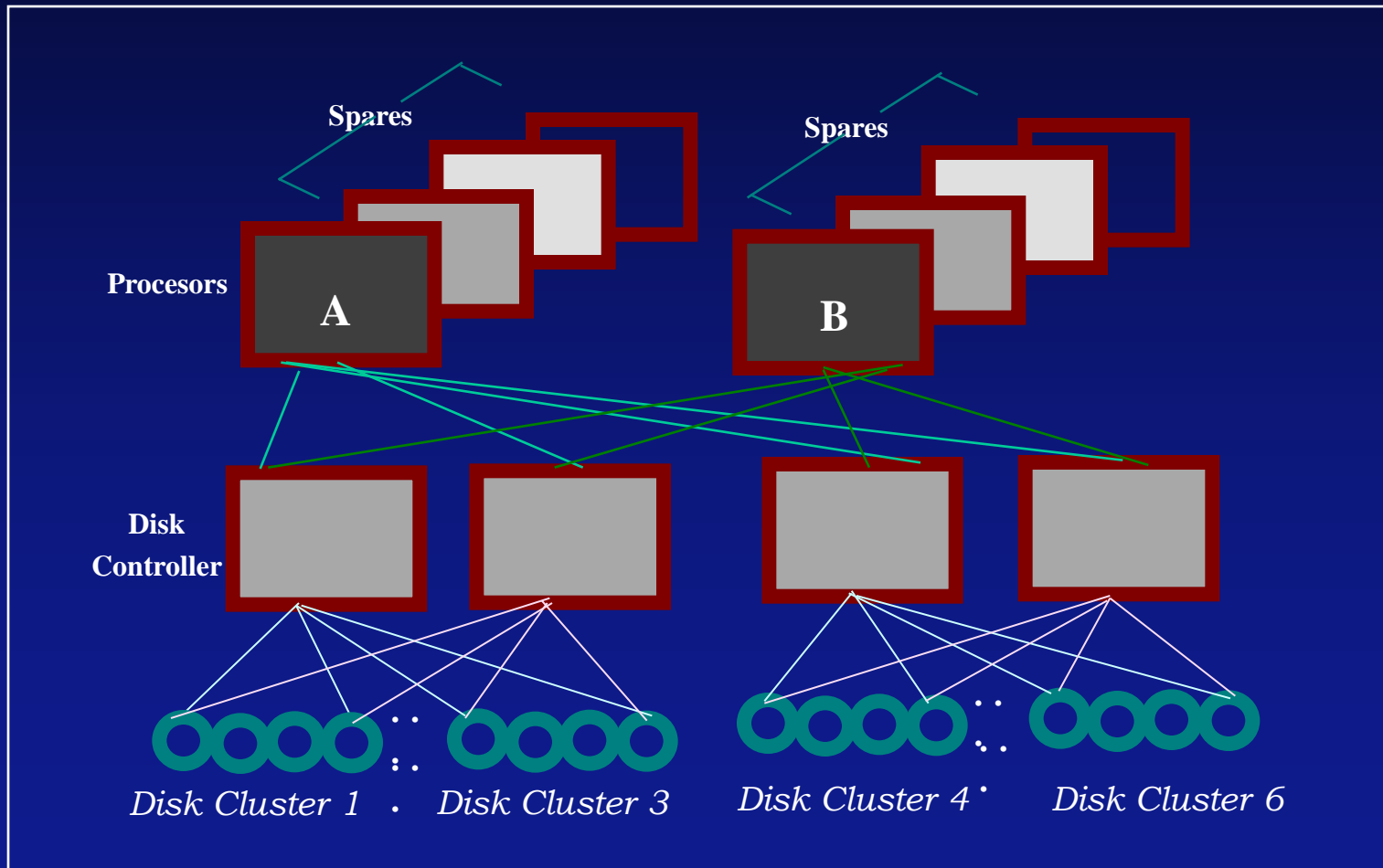
- **Importance function:** $\Phi(t) = cl - oc(t)$



Importance function: highly-dependable system (III)

- *Computing system*

- **Importance function: $\Phi(t) = cI - oc(t)$**



Model Features

Non-Markov $C(k, n: F)$ REPARABLE SYSTEM AND $(k-1)$ -STEP MARKOV DEPENDENCE

- The system fails if and only if k or more consecutive components have failed.
- The lifetime of components and the repair time have a general distribution.
- A failed component has the highest repair priority if the system failure risk after repair of that component is lower than after repair of any other failed component.
- If the system has failed, then no more components will fail.
- If there are h ($h < k$) consecutive failed components that precede the component i , the residual lifetime of component i will have increasing failure rate as h increases,

Simulation Results (I)

- **Table 1:** Unreliability estimates for C(4, 60: F) system. 95% confidence interval = $\pm 10\%$

Model	Interval (0, t_e)	<i>Unreliability</i>	Run-time (seconds)	Gain in time	Factor f_T	Factors $f_O \times f_V$
EL A	(0, 25)	3.4×10^{-6}	3	4.8×10^1	2.6	1.3
EL A	(0, 5)	3.8×10^{-8}	6	7.7×10^3	4.9	2.1
EL A	(0, 1)	1.2×10^{-10}	24	4.9×10^5	12.5	2.4
WL A	(0, 25)	8.6×10^{-13}	697	2.1×10^8	28.6	2.4
WL A'	(0, 25)	8.3×10^{-6}	11	1.6×10^1	2.4	2.0
WL A'	(0, 1)	2.0×10^{-10}	149	3.8×10^5	11.0	2.2

- **Repair times:** Lognormal in all the models.
- **Lifetimes:** Exponencial (models EL) or Weibull (models WL).
- **Model WL A':** Componets are 640 u.t. old

Simulation Results (II)

- **Table 2:** Unavailability and MTBF estimates for C(4, 60: F) system.
95% confidence interval = $\pm 10\%$

Model	Unavailability	MTBF	Run-time (seconds)	Factor f_T	Gain in time	Factors $f_V x f_O$
EL A	1.6×10^{-5}	5.8×10^5	4	2.6	38	4.4
EL B	4.6×10^{-8}	2.8×10^8	12	9.0	1267	5.9
EL C	4.2×10^{-10}	2.1×10^9	33	36.1	16650	7.8
WL A	2.2×10^{-5}	8.9×10^5	12	2.5	14	9.4
WL B	8.8×10^{-8}	1.4×10^8	45	8.2	263	17.7
WL C	5.6×10^{-10}	1.7×10^9	194	38.7	3159	29.4

- **Repair times:** Lognormal in all the models.
- **Lifetimes:** Same mean in Exponential and Weibull models.
- **Models B, C:** Mean lifetimes 3 and 10 times greater than model A

CONCLUSIONES

RESTART es una técnica de simulación de sucesos infrecuentes que aporta:

- Ahorro drástico del tiempo de computación
- Propósito general
- Aplicable si podemos definir regiones del espacio de estados anidadas que se vayan acercando a la región definida por el suceso raro

APLICACIONES (OTROS AUTORES)

- **Sistema de control de trenes europeo (ETCS)**

- **Control del tráfico aéreo**

El suceso raro es el choque de 2 trenes o 2 aviones

- **Simulación de ciber-ataques a sistemas críticos**

FUTURAS LÍNEAS DE TRABAJO

- **Obtener fórmulas cuando se llega al suceso raro desde varios umbrales**
- **Aplicaciones a diversos campos:**
 - Seguros y sistemas financieros**
 - Ciber-ataques a sistemas críticos**
 - Estimar probabilidad de pérdida en todos los nodos**
 - Estimación de fiabilidad de sistemas ultrafiabiles**
 - Problemas de Física, Química y Biología.**