# Modelado y Simulación de Sistemas Complejos con Aplicaciones en Economía

Clase 3 Modelos Estadísticos

#### Contenido de la clase 3

- Introducción. Orden, desorden y entropía. Principio de Jaynes (máxima entropía).
- Equilibrio, distribución de Gibbs
- Métodos de Monte Carlo. Dinámica de Glauber.
- Modelo de Ising. Transiciones.
- Optimización combinatoria (recocido simulado)
- Ejemplos: Viajante de Comercio, Modelo de Ising.

### Introducción

- hemos visto ejemplos de sistemas complejos que muestran autoorganización y emergencia (hormigas, economías, sistemas sociales...)
- vimos cómo reglas sencillas e interacciones dan lugar a estos fenómenos y patrones
- necesitamos entender cómo se representa, comunica y procesa la información en sistemas complejos
- cómo puede cuantificarse la información

### Introducción

"Although [complex systems] differ widely in their physical attributes, they resemble one another in the way they handle information. That common feature is perhaps the best starting point for exploring how they operate."

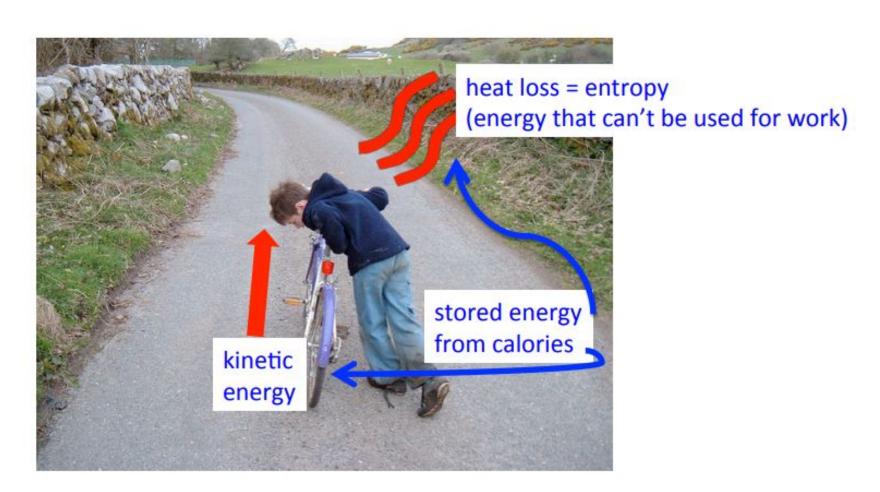
-- Murray Gell-Mann, The Quark and the Jaguar, 1995

#### Introducción

- Los ejemplos vistos hasta el momento involucran una gran cantidad de agentes y la dinámica de su interacción conduce a estados que es necesario caracterizar. Esto se hace con parámetros cuyo valor se obtienen como valores medios sobre distintas realizaciones.
- Los valores medios pueden ser definidos rigurosamente por medio de métodos estadísticos y rescatar por este medio parámetros que den una idea colectiva del conjunto de entes que integran un sistema.
- En sistemas físicos estos parámetros son la presión, la temperatura, la energía, la magnetización, etc.
- En el caso de sistemas sociales hay barreras que no existen en sistemas naturales:
  - o los actores individuales son mucho más complejos que átomos o moléculas
  - o los fenómenos involucran a pocos actores y los lapsos breves

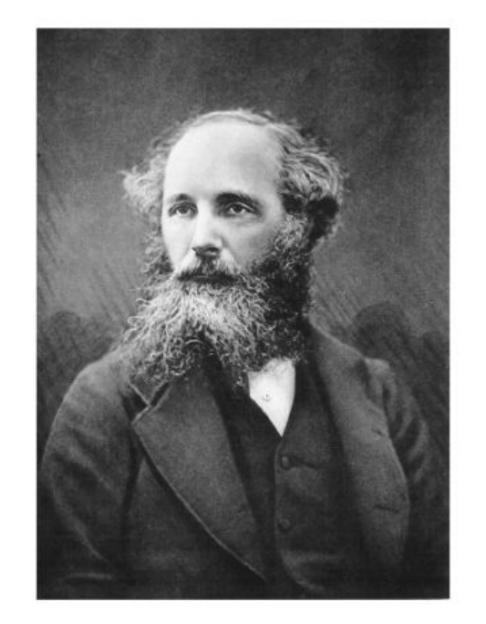
### Entropía, desorden e información

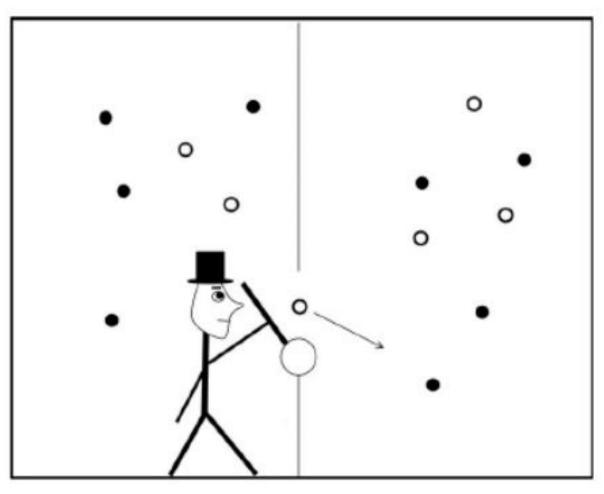
**Second law of thermodynamics:** In an isolated system, entropy always increases until it reaches a maximum value.



## Implicaciones de la Segunda Ley de la Termodinámica

- Los sistemas son desordenados por naturaleza. No pueden organizarse sin la aportación de trabajo
- El tiempo tiene una dirección: la dirección del aumento de la entropía





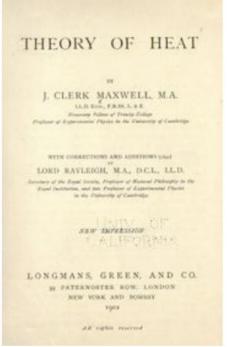
Maxwell's Demon

James Clerk Maxwell, 1831-1879

#### Maxwell's Demon

"The hot system [i.e., the right side] has gotten hotter and the cold [the left side] has gotten colder and yet no work has been done, only the intelligence of a very observant and neat-fingered being has been employed."

Maxwell: The second law of thermodynamics is "a statistical certainty".



## Leo Szilard (1898-1964): A "bit" of information

• Un bit de información es la cantidad de información necesaria para responder a una pregunta "rápido/lento", o a cualquier pregunta "sí/no".

#### ON THE DECREASE OF ENTROPY IN A THERMODYNAMIC SYSTEM BY THE INTERVENTION OF INTELLIGENT BEINGS

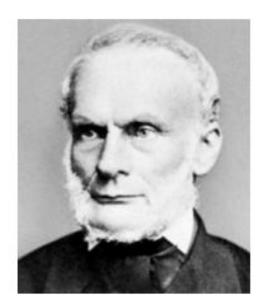
LEO SZILARD

Translated by Anatol Rapoport and Mechthilde Knoller from the original article "Über die Entropieverminderung in einem thermodynamischen System bei Eingriffen intelligenter Wesen." Zeitschrift für Physik, 1929, 53, 840-856.

Thermodynamic entropy measures the amount of heat loss when energy is transformed to work

Heat loss ≈ "disorder"

Theory is specific to heat



Rudolf Clausius, 1822-1888

Statistical mechanics entropy measures the number of possible microstates that lead to a macrostate

Number of microstates ≈ disorder

A more general theory



Ludwig Boltzmann, 1844-1906

#### A slight sidetrack to learn about microstates and macrostates



Microstates: specific state of the three slotmachine windows

Example microstate: {cherry, lemon, apple}

Note that a microstate here is a triple of fruit values, not a single fruit value. It is a description of the "state" of the slot machine.

Macrostate: Collection (or set) of microstates.

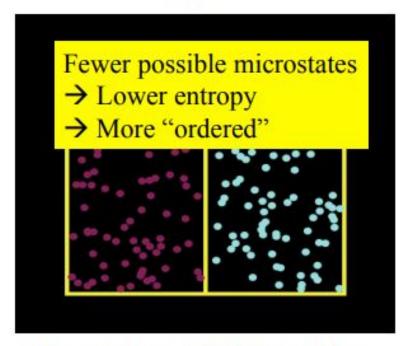
Example macrostate: Win (collection of microstates that have three of the same fruit showing).

Question 1: How many microstates give rise to the Win macrostate?

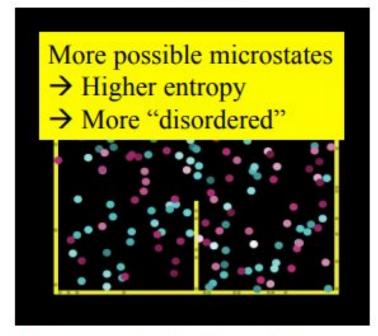
Question 2: How many microstates give rise to the Lose macrostate?

# Microstate: Position and velocity of of every particle

**Start** Finish



Macrostate: All fast particles are on the right, all slow particles are on the left.

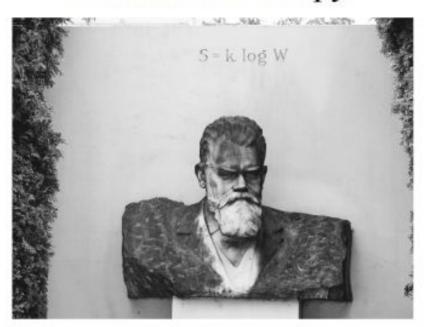


Macrostate: Fast and slow particles are completely mixed.

Second Law of Thermodynamics: In an isolated system, entropy will always increase until it reaches a maximum value.

Second Law of Thermodynamics (Statistical Mechanics Version): In an isolated system, the system will always progress to a macrostate that corresponds to the maximum number of microstates.

#### **Boltzmann Entropy**



Boltzmann's tomb, Vienna, Austria

The entropy S of a macrostate is k times the natural logarithm of the number W of microstates corresponding to that macrostate.

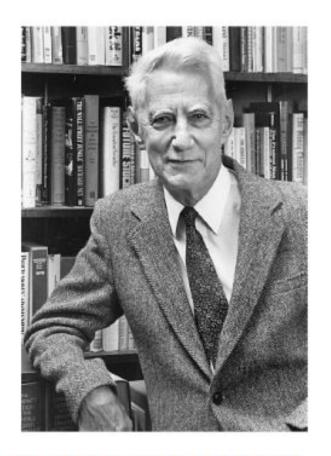
k is called "Boltzmann's constant". This constant and the logarithm are just for putting entropy into a particular units.

**General idea:** The more microstates that give rise to a macrostate, the more probable that macrostate is. Thus *high entropy = more probable macrostate*.

Second Law of Thermodynamics (Statistical Mechanics Version):

In an isolated system, the system will tend to progress to the **most probable macrostate**.

#### **Shannon Information**



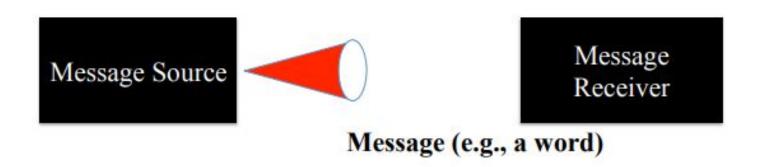
Claude Shannon, 1916-2001

Shannon worked at Bell Labs (part of AT&T)

Major question for telephone communication: How to transmit signals most efficiently and effectively across telephone wires?

Shannon adapted Boltzmann's statistical mechanics ideas to the field of communication.

#### **Shannon's Formulation of Communication**

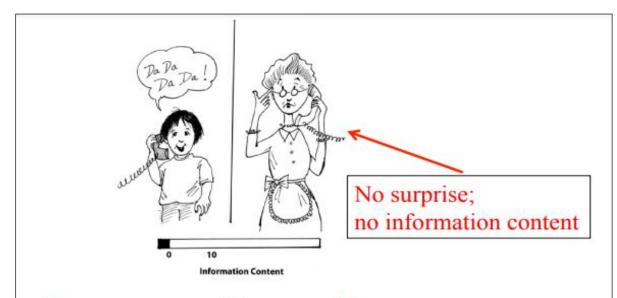


Message source: Set of all possible messages this source can send, each with its own probability of being sent next.

Message: E.g., symbol, number, or word

**Information content** *H* **of the message source:** A function of the number of possible messages, and their probabilities

**Informally:** The amount of "surprise" the receiver has upon receipt of each message

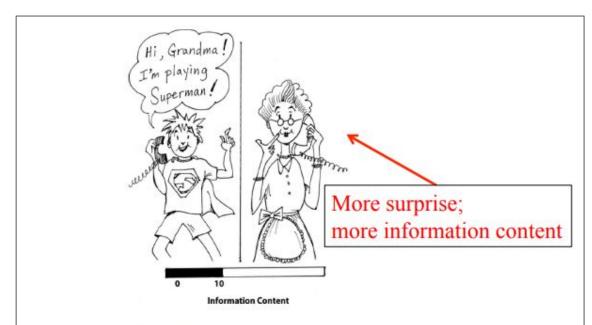


Message source: One-year-old

Messages:

"Da" Probability 1

**InformationContent** (one-year-old) = 0 bits



Message source: Three-year-old

**Messages:** 500 words  $(w_1, w_2, ..., w_{500})$ 

**Probabilities:**  $p_1, p_2, \dots, p_{500}$ 

**InformationContent** (three-year-old) > 0 bits

## Cómo computar la entropía de Shannon?

#### **Boltzmann Entropy**

#### **Shannon Information**

Microstate: Detailed configuration of system components (e.g., "apple pear cherry")

Macrostate: Collection of microstates (e.g., "all three the same" or "exactly one apple")

**Entropy** S: Assumes all microstates are equally probable

 $S(macrostate) = k \log W$ 

where *W* is the number of microstates corresponding to the macrostate. *S* is measured in units defined by *k* (often "Joules per Kelvin")

#### **Boltzmann Entropy**

Microstate: Detailed configuration of system components (e.g., "apple pear cherry")

Macrostate: Collection of microstates (e.g., "all three the same" or "exactly one apple")

**Entropy** S: Assumes all microstates are equally probable

 $S(macrostate) = k \log W$ 

where W is the number of microstates corresponding to the macrostate.

S is measured in units defined by k (often "Joules per Kelvin")

#### **Shannon Information**

**Message:** E.g., a symbol, number, or word.

Message source: A set of possible messages, with probabilities for sending each possible message

#### **Information content** *H*:

Let *M* be the number of possible messages. Assume all messages are equally probable.

 $H(message source) = \log_2 M$ 

H is measured in "bits per message"

## Entropía de Shannon: general

Sea M el número de mensajes posibles y p<sub>i</sub> la probabilidad del mensaje i.

$$H(message source) = -\sum_{i=1}^{M} p_i \log_2 p_i$$

$$H(message source) = -\sum_{i=1}^{N} p_i \log_2 p_i$$

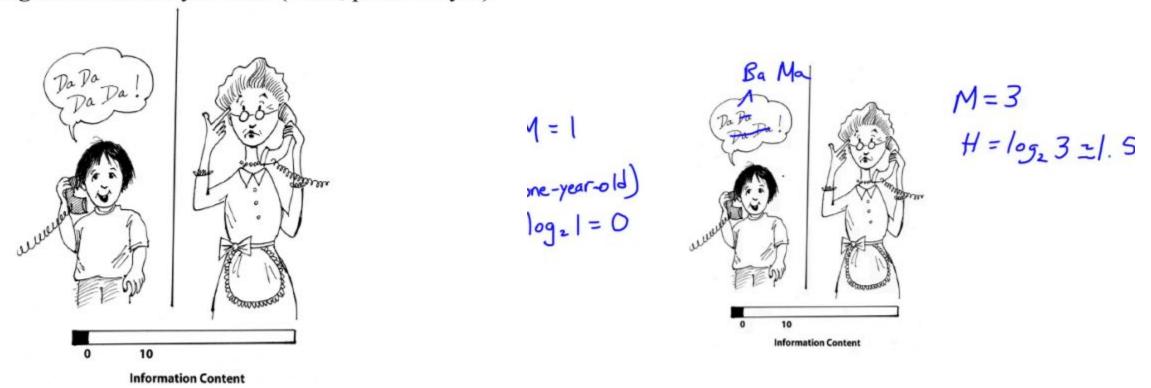
$$= -\sum_{i=1}^{N} \frac{1}{M} \log_2 \frac{1}{M}$$

$$= -\log_2 \frac{1}{M}$$

$$= -\log_2 M^{-1}$$

$$= \log_2 M$$

#### Message source: One-year-old: {"Da", probability 1}



$$H(\text{one-year-old}) = -\sum_{i=1}^{N} p_i \log_2 p_i = 1 * \log_2 1 = 0 \text{ bits}$$

## Entropía e información

• La idea es explorar conceptos que puedan aplicarse a sistemas con algún grado de organización interna o de información, como pueden ser los AC

#### **Definiciones:**

- Sea una fuente capaz de emitir símbolos  $X_i$  de un alfabeto A (la fuente puede ser un AC, la cotización del Merval, el precio de la soja en Chicago, etc.)
- Se supone que la ocurrencia de los caracteres  $X_i$  se ajustan a una distribución de probabilidades P(X)
- Se define entropía (o cantidad de información "a priori" (\*)) a

$$S(X) = -\sum_{i} P(X_{i}) \log(P(X_{i}))$$

- Si el log es en base 2 la entropía se mide en bits
- $\circ$  (\*) es la información que falta antes de recibir el mensaje conteniendo muchas realizaciones de X

## Entropía en casos simples

- Supongamos que la fuente de información en un AC de n células binarias evoluciona de acuerdo con alguna regla interna
- Los símbolos  $X_i$  que emite el AC son palabras de n bits
- Si el autómata llega a un punto fijo siempre emite un mismo símbolo
- $X_o$  con lo que la distribución de probabilidad P(X) será
- P(X=Xo) = 1  $y P(X \neq Xo) = 0$
- En este caso usando  $S(X) = -\sum_{i} P(X_i) \log(P(X_i))$  la entropía resulta S(X) = 0
- La llegada del mensaje no aporta ninguna información!

## Entropía en casos simples

- Si el AC recorre todos sus estados posibles emitirá palabras de un alfabeto con 2<sup>n</sup> símbolos
- Si todos ellos tienen la misma probabilidad,  $P(X_i) = \frac{1}{2^n}$ ,  $\forall X_i$ , con lo que

$$S(X) = -\sum_{i} P(X_{i}) \log_{2}(P(X_{i})) = -\sum_{i}^{2^{n}} 2^{-n} \log_{2}(2^{-n}) = n$$

- Cada mensaje aporta una información de n bits que son los necesarios para especificar el estado del AC en cada momento
- Por estas razones a la cantidad  $-\log(P(x)) = \log(1/P(x))$  se la suele llamar la sorpresa del evento x

## Entropía y desorden

- Si el AC emite cualquiera de los 2<sup>n</sup> símbolos de su alfabeto con igual probabilidad significa que el AC recorre todas sus posibles configuraciones sin seguir ningún orden preestablecido
  - Esta situación corresponde a una entropía máxima
- Por el contrario si el AC emite solamente símbolos que son una fracción de los 2<sup>n</sup> símbolos posible de su alfabeto indica que sólo recorre un subconjunto limitado de sus posibles configuraciones sugiriendo algún orden de preeminencia entre ellas
  - Esta situación corresponde a una entropía menor
- En el límite que sólo repita un solo símbolo la entropía es mínima

## Entropía y desorden

- Consideremos una urna conteniendo bolillas de dos posibles colores
- Si se supone que N es el número total de bolillas:  $N_0$  es el número de bolillas negras y  $N_1$  es el número de bolillas blancas, el número de situaciones posibles está dado por  $C = \frac{N!}{N_0! N_1!}$
- Si se le asigna a cada alternativa una probabilidad  $P(N_0N_1) = 1/C$
- y se utiliza la expresión asintótica  $\log(x!) \sim x \log x x$
- ullet es posible comprobar que la probabilidad se hace mínima cuando  $\,N_0=N_1=N\,/\,2\,$
- que es cuando el desorden es máximo o la mezcla de bolillas es más homogénea. En esas condiciones la sorpresa  $-\log(P) = \log(1/P)$

es máxima

## Aditividad de Entropía

- Se desea que el contenido de información de un mensaje tenga la propiedad de aditividad : la información que provenga de sistemas independientes debe ser la suma de las informaciones de cada uno
- Para ello se usa que la probabilidad conjunta de variables aleatorias independientes es el producto de las probabilidades:
- si  $x \in X$ ;  $y \in Y$  entonces P(x, y) = P(x)P(y)
- la entropía resulta  $S(x,y) = -\sum P(x,y)\log(P(x,y)) = -\sum P(x)P(y)\log(P(x)P(y))$

$$= -\sum P(x)P(y)[\log(P(x) + \log(P(y))] = S(x) + S(y)$$

• La entropía tomada como información resulta una variable *extensiva*: que sea aditiva hace que crezca con el tamaño del sistema

## Entropía y Equilibrio

- En la naturaleza hay ciertas transformaciones térmicas que se dan espontáneamente y otras que sólo son posibles si se actúa externamente y adrede sobre el sistema
- *Ejemplo*: ningún objeto aumentará espontáneamente su temperatura por encima de la del medio en que se encuentra.
- La evolución natural de un sistema físico tiene lugar solamente en una dirección, y es aquella en que su entropía o aumenta o permanece constante (teorema H de Boltzmann)
- Cuando la entropía permanece constante se dice que el sistema físico se encuentra en equilibrio.
- El concepto de equilibrio posee un alcance diverso en distintos contextos

## Entropía y Equilibrio

- Cuando se observa un sistema se dice que éste se encuentra en un estado de equilibrio cuando los valores medios de los parámetros que lo caracterizan no cambian y las condiciones no varían durante lapsos compatibles con el proceso de medición
  - equilibrio (en economía) = consistencia de planes
  - equilibrio = valores se mantienen constantes
  - equilibrio = estacionariedad
- Se puede buscar la distribución de probabilidad de las variables del sistema que corresponda a un equilibrio pero imponiendo además que ese equilibrio está caracterizado por la constancia del valor medio de algún parámetro determinado
- Esto se hace buscando una distribución de probabilidades de las variables del sistema que haga máxima su entropía e imponiendo la constancia del valor medio como una condición de vínculo

## Entropía y Equilibrio

- Por lo general se trata de averiguar cuál es la distribución de probabilidad de alguna variable aleatoria cuyas realizaciones pueden observarse(\*).
  - \*) Usualmente se evita este camino haciendo la hipótesis que la variable aleatoria sigue una distribución dada, por ejemplo, normal. Sin embargo existen casos en los que esa hipótesis puede conducir a errores importantes
- Conocer esa distribución de probabilidad permite calcular valores medios de cualquier función H(X).

$$\langle H \rangle = \sum_{i} P(X_{i}) H(X_{i})$$

• En realidad la pregunta que se debe hacer es: ¿Cuál es la distribución de probabilidad que sólo contiene la información de las observaciones y no tiene ningún sesgo adicional?

## Principio de máxima entropía

- La respuesta es el *Principio de Jaynes:* De todas las distribuciones de probabilidad que son compatibles con los vínculos que operan en el sistema, aquella que contiene un mínimo sesgo es la que maximiza la entropía
- Encontrar la distribución de probabilidad equivale a resolver el siguiente problema de maximización con vínculos

$$\max\{S(X_i)\} = \max\{\sum_i P(X_i) \log(P(X_i))\}$$
 s.a.  $\sum_i P(X_i) = 1$ 

- Si el único vínculo es la normalización de la probabilidad, la solución de la maximización con vínculos es P(X)=1/N
- ullet donde N es el el número de símbolos del alfabeto A de símbolos X

## Principio de máxima entropía

 Para imponer que un valor medio se mantiene constante se introduce un multiplicador de Lagrange

$$\max\{S(X_i)\} = \max\{\sum_i P(X_i) \log(P(X_i)\} \quad \text{s.a.} \quad \sum_i P(X_i) = 1$$
 y, además s.a. 
$$\langle E(X) \rangle = \sum_i E(X_i) P(X_i) = \text{constante} = \textbf{\textit{E}}$$

 La solución requiere introducir dos multiplicadores de Lagrange, uno de los cuales se elimina y el otro es β. Se obtiene:

Distribución de Gibbs 
$$P(X_i) = \frac{e^{-\beta E(X_i)}}{Z_{\beta}}$$

Función de partición 
$$Z_{\beta} = \sum_{i} e^{-i\beta E(X_{i})}$$

- En sistemas físicos si E corresponde a la energía, el  $parámetro \beta$  es la inversa de la temperatura T. Determina la escala en la que se deben medir la energía de los estados del sistema.
- Se puede suponer que cuando T es elevada, el sistema puede ser observado con probabilidad apreciable tanto en estados de baja como de alta energía mientras que si T => 0 el sistema sólo permanece y puede ser observado en su estado de mínima energía.
- Por esta razón se suele asociar *T* al nivel de "quietud" o de "agitación" interna. Para muy baja *T*, sólo son probables unos pocos estados de mínima energía. Para *T* elevada todos los estados del sistema tiene una probabilidad significativa de ser observados.
- Debería existir una "*temperatura*" asociada a cada variable que se conserva constante en valor medio en un sistema

# Modelo de Ising

### Modelo de Ising

- El modelo de Ising es un AC desarrollado para describir medios magnéticos en base a la interacción de dipolos magnéticos vecinos. Actualmente es la base de una gran diversidad de modelos en los que los agentes deben tomar una decisión binaria.
- La energía de dos spines en interacción es:  $E_{i,j} = \pm J s_i s_j \; \text{con} \; s_i = \pm 1 \; s_i = 1 \;$
- Se supone que sólo interactúan spines que son primeros vecinos. *J* es la intensidad de la interacción.
- El *estado de mínima energía* para esta interacción es que los dos spines se encuentren orientados en forma paralela o antiparalela, dependiendo del signo de la misma.
- Un medio magnético puede suponerse constituído por una gran cantidad de estos dipolos ubicados en los nodos de una grilla regular.
  - →La grilla puede tener una, 2 o 3 dimensiones.
  - →Para calcular la energía de N dipolos hay que establecer cuál es la vecindad de cada dipolo elemental.
  - →La orientación de los dipolos puede variar.

# Modelo de Ising

- El **Momento Magnético** de un material de este tipo es la suma de todos sus spines:  $M = \sum_{i=1}^{N} s_{j}$ 
  - puede alcanzar valores muy grandes aun cuando el momento magnético de cada spin sea muy pequeño (*M* es una variable extensiva).

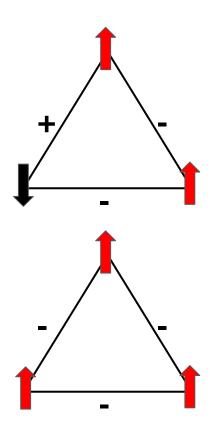
#### • Interacciones:

- $\circ$  Si la interacción es + J(J>0) los dipolos tienden a alinearse antiparalelamente
- $\circ$  Si la interacción es J(J>0) los dipolos tienden a alinearse paralelamente
- Si la interacción magnética es a veces con y otras con + se puede dar una situación de "frustración"

### Modelo de Ising

- Un **sistema frustrado** no tiene un único mínimo de la energía sino que tiene muchos mínimos secundarios que son equivalentes.
- Un sistema no frustrado en que todas las interacciones son – (atractivas) tiene un único mínimo de la energía (sistema ferromagnético)

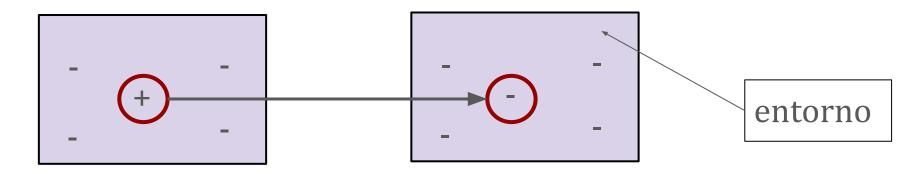
 Un sistema en que todas las interacciones son + es un antiferromagneto y es asimilable al juego de la minoría y puede ser frustrado



Si se fija una orientación inicial de todos los dipolos al azar, el medio magnético puede tratararse como un AC

La dinámica más sencilla es la de orientar en cada instante un dipolo seleccionado al azar, poniéndolo en forma paralela (o antiparalela) con la orientación de todo el resto del sistema.

Con cada actualización la energía tenderá a disminuir y el momento magnético a crecer en valor absoluto.



Se observa experimentalmente que un medio magnetizado (todos los dipolos orientados paralelamente) pierde ese orden tan particular si se lo calienta por encima de una temperatura crítica.

Por esta razón se suele asimilar el efecto de la temperatura a un desorden.

Desde el punto de vista de la dinámica del AC es equivalente a suponer que es un autómata probabilístico en el que la orientación de cada spin es aleatoria y sólo sigue a los demás con una probabilidad que depende de *T* 

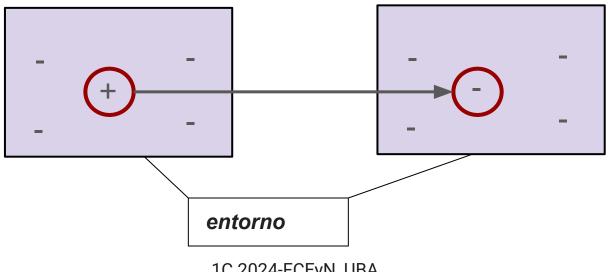
### Dinámica de Glauber

- Se supone que el relajamiento al equilibrio de un sistema a una dada temperatura se puede describir como la ocurrencia de numerosos cambios pero que, en cada oportunidad, se actualiza el valor sólo de una de las variables del sistema.
- Además la probabilidad de cada actualización depende solamente de su "costo" en energía a la temperatura que se encuentra el sistema.

- Si un AC de *N* celdas como el del modelo de Ising es perturbado poniéndolo en contacto con una fuente a temperatura *T*, todas las variables de las celdas del sistema se alterarán pero esas variables deberán alcanzar valores medios estables al cabo de cierto tiempo.
- Si en esas condiciones de equilibrio se practican mediciones sobre el sistema se lo encontrará solamente en los estados más probables para esa temperatura, o sea los de mínima energía.
- El proceso de relajamiento a ese estado de equilibrio se supone que se realiza siguiendo una dinámica particular denominada *Dinámica de Glauber*

#### Dinámica de Glauber

- El problema de determinar la configuración de equilibrio de un imán puesto en contacto con una fuente térmica consiste en determinar cómo se disponen todos los dipolos del sistema que hagan mínima la energía. Se hace utilizando la dinámica de Glauber
- Primero se fija al azar una orientación inicial de todos los dipolos y se reorienta cada uno, en cada paso de tiempo teniendo en cuenta que cambiar la orientación de un dipolo (el valor de una celda del autómata) tiene un costo en energía.



#### Dinámica de Glauber

- Para describir el proceso de relajamiento al equilibrio basta considerar el cambio de valor de **una** de sus celdas.
- Si los dos estados están caracterizados por las variables  $X_{+1}$  y  $X_{-1}$  y se les asocia las energías  $E_{+1}$  y  $E_{-1}$  se puede decir que las respectivas probabilidades son

$$P(X_{-1}) = \frac{e^{-\beta E_{-1}}}{Z}$$

$$P(X_{+1}) = \frac{e^{-\beta E_{+1}}}{Z}$$

$$Z = e^{-\beta E_{-1}} + e^{-\beta E_{+1}}$$

• con lo que se publicade de probabilidad asociada a un cambio de energía  $P(E_{+1}) = \frac{1}{1 + e^{-\beta \Delta E}}$ 

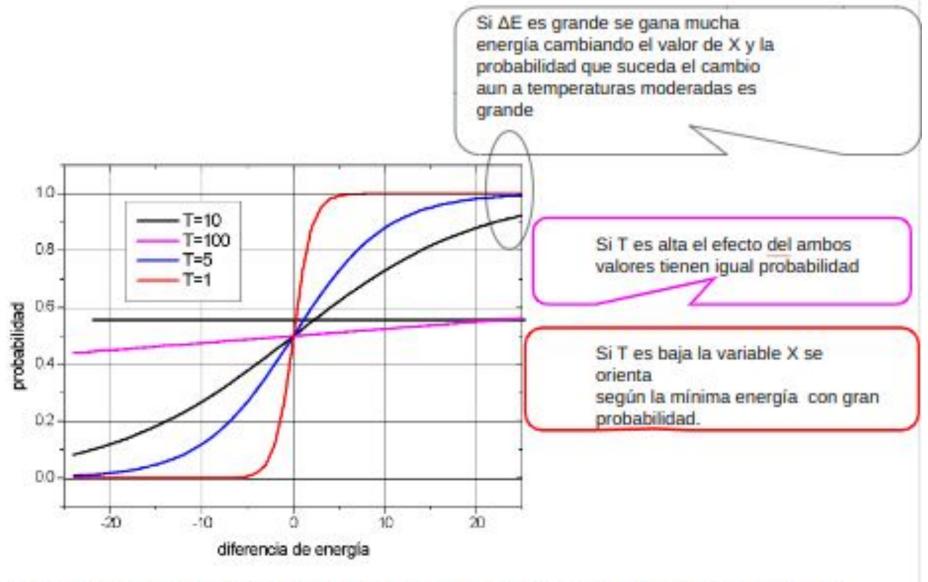
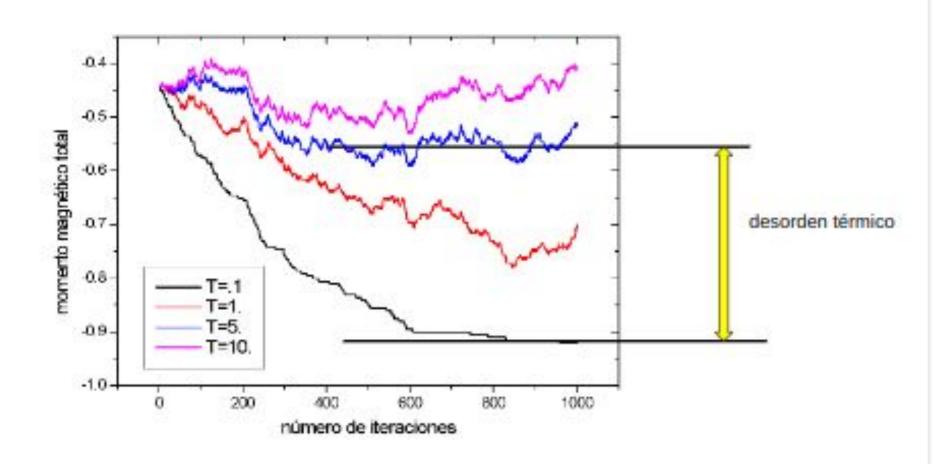


Gráfico de la probabilidad de transición  $P(X_{+1}) <=> P(X_{-1})$  en función de la diferencia de energía  $\Delta E$ .

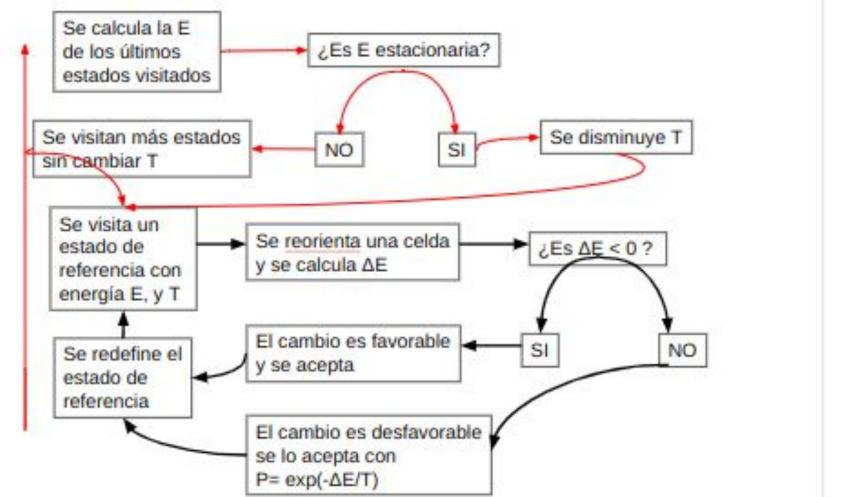


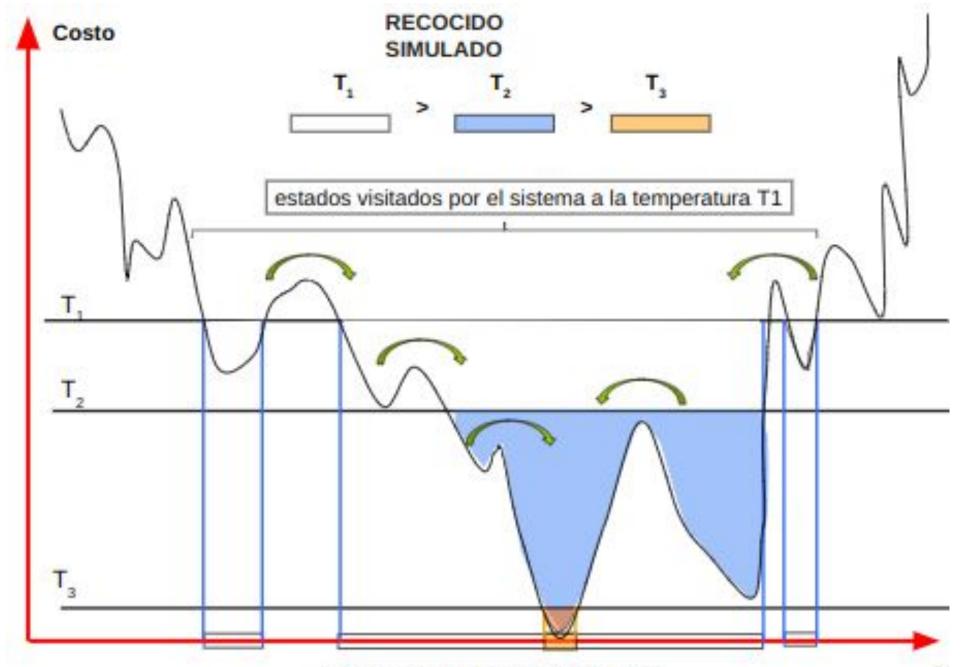
Dinámica de un modelo de Ising de 20 X 20 en el que cada spin se actualiza alineándose con el momento magnético de todo el sistema siguiendo una dinámica de Glauber en la que sólo se tiene en cuenta el campo total sobre cada dipolo o sea cada spin se orienta en la misma dirección que todo el resto. La condición inicial es una distribución de spines +1 (up) y -1 (down) con probabilidades .3 y .7 respectivamente.

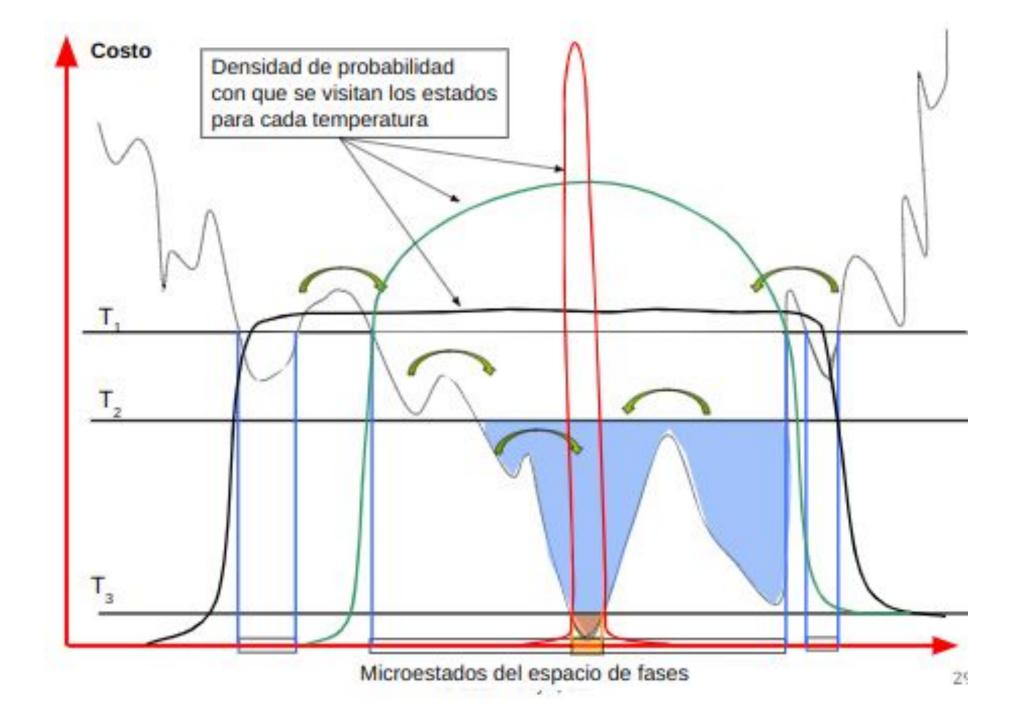
- El proceso de relajación al equilibrio equivalente al problema matemático de encontrar el extremo (mínimo) de una función de un gran número de variables que, por lo general no es monótona ni regular.
  - ej: las orientaciones de todos los dipolos del sistema)
- Esta semejanza ha sido explotada para construir heurísticas para resolver problemas de optimización combinatoria (el "Recocido simulado" de Kirkpatrik, Gelatt y Vecchi).
- Para ello el papel de la energía lo juega una función costo que se la que se debe minimizar y la temperatura es una variable de control que da la escala de los cambios de dicha energía.
  - (ej: en el problema del viajante de comercio, E es la distancia de cada itinerario)
- Se utiliza el algoritmo de Metropolis que aplica la dinámica de Glauber postulando una distribución de Gibbs para los estados del sistema y determinar cuáles son los estados que "visita" el sistema a medida que se disminuye progresivamente la temperatura.

### Algoritmo de Metropolis

• Se trata de hallar el mínimo de una función de las configuraciones de un sistema de *N* celdas (o agentes).

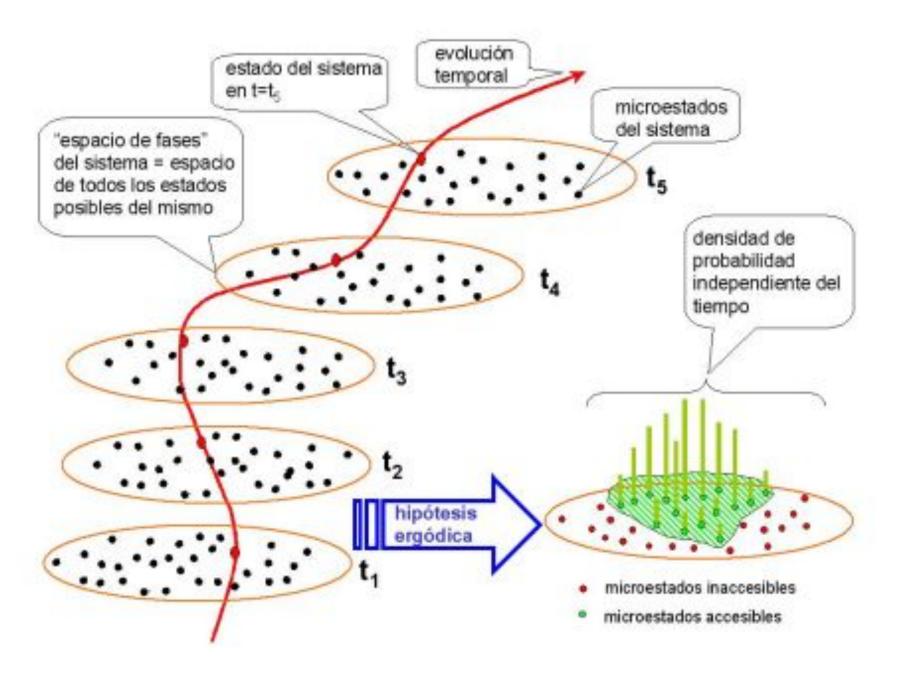






# Hipótesis ergódica

• Los valores medios temporales (a lo largo de la evolución) son equivalentes a valores medios sobre densidades de probabilidad independientes del tiempo que solo dependen de los estados internos del sistema.



# Hipótesis ergódica

• Se puede calcular promedios temporales observando la evolución del sistema y registrando los microestados que va ocupando y, por el otro, en un dado instante, se pueden estimar valores medios pesando cada uno de los microestados del volumen por su probabilidad de ocurrencia. La hipótesis ergódica dice ambos promedios son equivalentes.

¡Nos vemos en la próxima clase!