**PROYECTO FINAL**

En esta parte final se revisará si es posible descomponer un sistema en partes (subsistemas) sin que se pierda información. Esto implica que podemos tratar de descomponer el sistema en muchas formas diferentes (por ejemplo partir en 2 subsistemas, o partir en 3 subsistemas, o en 4 etc.). Para nuestro proyecto trabajaremos descomponiendo el sistema en dos partes; sin embargo, hay muchas formas de poder descomponer el sistema en dos subsistemas. El trabajo consistirá en revisar cuanta información se pierde al descomponer el sistema en dos subsistemas y encontrar cual es la forma en la que se genera la menor pérdida de información. Para esto se debe revisar la diferencia de información entre el sistema original y el sistema descompuesto.

Es de aclarar que es posible trabajar sobre el sistema completo o sobre una parte del sistema. Revisaremos varios ejemplos, para mayor claridad. En todo caso, estamos interesados en ver si los efectos y las causas de un sistema o subsistema van más allá de las causas y efectos de sus partes . En este ejemplo solo trabajaremos con los efectos (estado futuro).

**Ejemplo 1.**

Se trabajará sobre ACt presente en el estado actual (1 0), sobre ABCt+1 futuro . Para esto se va a descomponer  en dos partes, por ejemplo:  y .

Es decir, se quiere revisar si la descomposición del sistema inicial  en dos partes como por ejemplo y , hace alguna diferencia al sistema inicial.

El proceso consistira en lo siguiente:

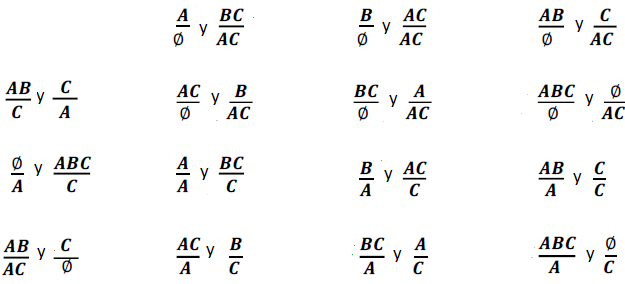
1. Se hará el cálculo de la distribución de probabilidades del sistema a tratar en el estado inicial dado. Para el ejemplo en cuestión se calcula que dará como resultado una distribución de probabilidades que llamaremos **Distribución Original**, que se hace como se explicó en el taller #3 (a partir de la matriz de probabilidades empezar a marginalizar en el fututo, luego en el tiempo actual etc.)
2. Luego se divide el sistema original en dos partes. Se trabajará sobre una de las partes en que se dividió el sistema, , que implica hacer el cálculo de , que dará como resultado una distribución de probabilidades que llamaremos **Distribución 1**  (que se calcula como ya se sabe y se explico en el taller #3).
3. Luego se trabajará sobre la otra parte, y se hace el cáculo de  que dará como resultado una distribución de probabilidades que llamaremos **Distribución 2** (que se calcula como ya se sabe y se explicó en el taller #3).
4. Luego se aplica el producto Tensor sobre las distribuciones resultantes en 2) y 3) y a esta nueva distribución se le denominará **Distribución Partida**:

Distribución Partida= Distribución 1  Distribución 2

1. Luego se compara la distribución de probabilidades resultante en 1, **Distribución Original,** con la distribución resultante en 4, **Distribución Partida,** y se obtiene la **Información Ganada**:

**Información Ganada = Comparar(Distribución Original, Distribución Partida).** Más adelante en el documento se menciona la medida Earth Mover’s Distance que deberán consultar para hacer esta comparación.

Este mismo proceso se deberá ejecutar para todas las posibles formas de descomponer con el fin de conseguir la descomposición que produzca la mínima información ganada. Es decir, para encontrar la información ganada mínima, se deben revisar todas las formas posibles en que el sistema puede descomponerse, como se presenta a continuación:



Siguiendo con el ejemplo, y de acuerdo con lo explicado anteriormente:

1. Empezaremos por calcular  en el estado actual (1 0), como se explicó en los talleres anteriores. Se obtiene la siguiente distribución:



1. Se requiere descomponer el sistema en dos partes,  y .
2. Entonces primero se trabajará sobre  . Para esto, se calcula  que da como resultado:

b) Luego se calcula 

Entonces en este caso no se marginaliza sobre las filas, porque no hay un estado definido en el tiempo presente. Dada esta situación lo que hacemos es calcular la probabilidad de C=0, y para eso nos ubicamos en la columna donde C=0 y sumamos todos los valores en todos los estados (filas) y se divide por la cantidad total de estados. Después hacemos el mismo procedimiento pero para C=1

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | E. futuro |  |  |  |
| E. Actual | | C | 0 | 1 |
| A | B | C |  |  |  |
| 0 | 0 | 0 |  | 1 | 0 |
| 1 | 0 | 0 |  | 0 | 1 |
| 0 | 1 | 0 |  | 0 | 1 |
| 1 | 1 | 0 |  | 1 | 0 |
| 0 | 0 | 1 |  | 1 | 0 |
| 1 | 0 | 1 |  | 0 | 1 |
| 0 | 1 | 1 |  | 0 | 1 |
| 1 | 1 | 1 |  | 1 | 0 |

|  |  |  |
| --- | --- | --- |
| C | 0 | 1 |
|  | 0,5 | 0,5 |



1. Ahora tomamos el producto tensor para conseguir un repertorio sobre el estado futuro formado por las dos partes en que se descompuso el sistema:





**COMBINACION DEL SISTEMA PARTICIONADO**

1. Calculamos la distancia entre la distribucion de probabilidades del sistema origina sin descomponer (realizada en 1) y la distribución del sistema partido



**SISTEMA ORIGINAL**



**COMBINACIÓN SISTEMA PARTIDO**



**Distancia ( Sistema Original , Combinación sistema partido) = 0.0**

En este caso no hay diferencia entre las dos distribuciones por tanto la información ganada es CERO.

Para medir la distancia entre las dos distribuciones de probabilidades se usará la **EARTH MOVER’S DISTANCE (EMD) (**pueden buscar el algoritmo para la solución analítica de la EMD) .

**Ejemplo 2.**

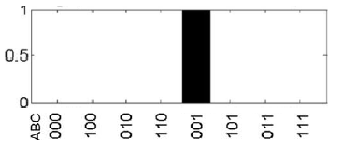
Se trabajará sobre presente, en (1 0 0), sobre futuro . Para esto se va a descomponer  en dos partes, por ejemplo:  y  .

Es decir, se quiere revisar si la descomposición del sistema original  en dos partes y , hace alguna diferencia al sistema original.

Siguiendo lo descrito anteriormente para el ejemplo 1:

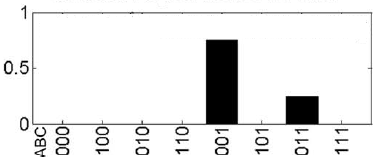
Se hará el cálculo de la distribución de probabilidades del sistema a tratar en el estado inicial dado. Para el ejemplo en cuestión se calcula  que dará como resultado una distribución de probabilidades que es la **Distribución Original** y se obtiene lo siguiente:

**SISTEMA ORIGINAL**

 = 

Luego se debe calcular cada uno de los dos subsistemas en que se dividio el sistema original y se obtiene:

**COMBINACIÓN SISTEMA PARTIDO**

 =

**Distancia ( Sistema Original , Combinacion sistema partido) = 0.25**

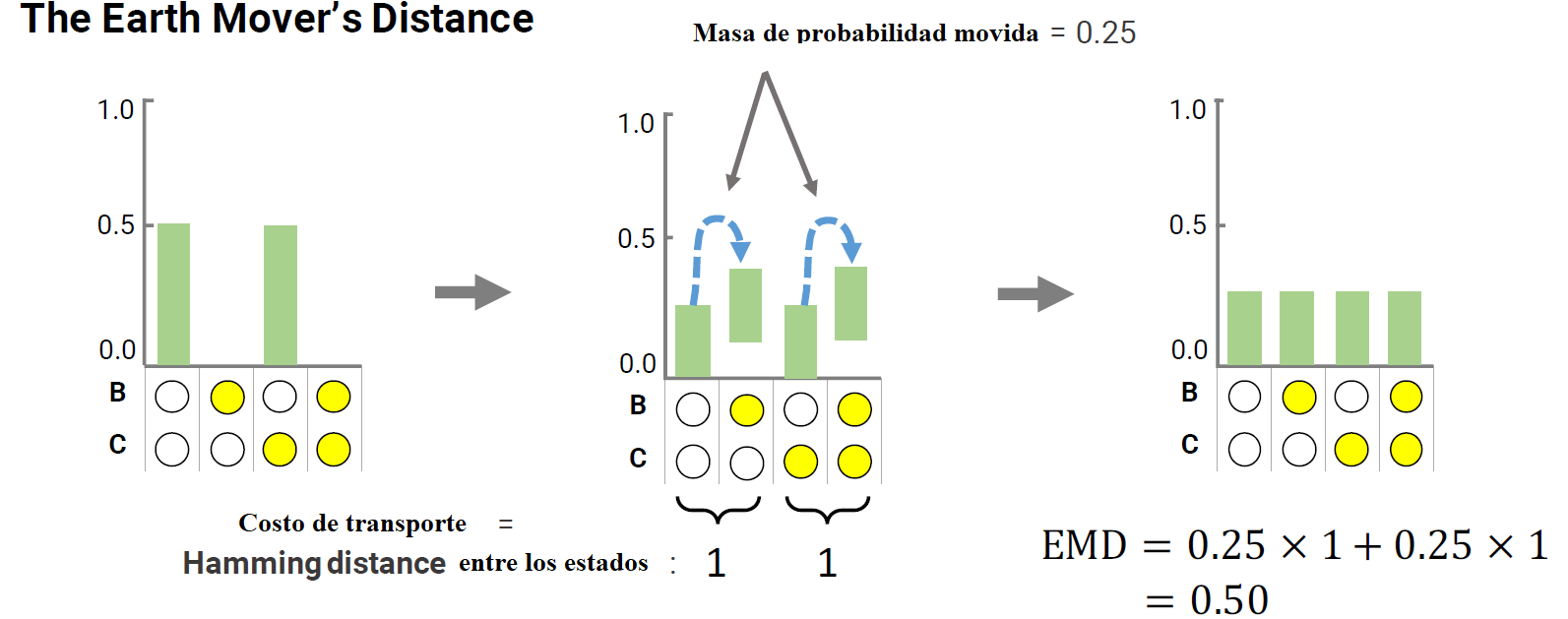
\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**El objetivo de final en este trabajo es buscar cual forma de descomponer el sistema proporciona la mínima información ganada.**

**ANEXO**

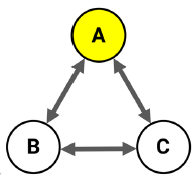
**Earth mover’s distance**

La EMD es el mínimo costo de transformar una pila de “tierra” en otra , donde el costo es la cantidad de “tierra” movida multiplicada por la distancia que ella viaja.

****Pueden emplear el algoritmo dela solución analítica de la EMD.

**Especificación de la propuesta alternativa**

Esta propuesta se basa en el uso de modelos gráficos causales. Estos modelos implican la representación del sistema como un grafo. En este sentido, usted tiene una serie de canales o red de nodos que se pueden representar gráficamente, donde cada nodo tiene su propia matriz de probabilidades de transición como se muestra en la siguiente figura:

Estas matrices pueden formar la matriz de probabilidades del sistema, como se muestra a continuación:



Como se mencionó anteriormente el sistema está compuesto de elementos interactuantes. Esos elementos están en uno de dos estados (1 , 0).Cada elemento recibe entradas y provee salidas. La función de entrada/salida de cada canal o elemento se caracteriza por su matriz de probabilidades. Este sistema se puede representar por un grafo dirigido de nodos interconectados (un nodo representa un canal), cada nodo, por tanto tiene una función que produce el estado del nodo en el tiempo t+1 dado el estado de sus padres en el tiempo previo t. El sistema esta completamente especificado por su matriz de probabilidades, que contiene las probabilidades de las transiciones de todos los estados de t a t+1.

Aprovechando esta representación gráfica suponga que el sistema es particionado en dos partes y los arcos que van de una parte a la otra son cortados , volviéndolos sin ninguna potencia causal, como el ejemplo que se muestra en la figura 2. Suponga que tiene un conjunto *Ecorte*que contiene los arcos del sistema que han de ser cortados, donde cada arco *e Є Ecorte* tiene un nodo fuente *a* y un nodo destino *b.* Para este arco se modifica la matriz de probabilidades del nodo *b*, marginalizando sobre los estados de *a* en *t*. La matriz resultante especifica la función implementada por *b* con la influencia desde *a* removida.

En esta estrategia el objetivo es quitar conexión por conexión y verificar si el sistema sin esa conexión no cambia (no hay perdida de información) respecto al sistema original.

Por ejemplo, supongamos que el sistema cuenta con 3 canales: A, B y C, entonces para determinar si las variables en el presente: AC ,BC y CC afectan o no a la variable A en el futuro AF, podemos buscar un subconjunto mínimo Z ⊆ { AC ,BC , CC } tal que

*P*(AF ∣ AC ,BC , CC) =  *P*(AF ∣ Z)

Entonces podemos eliminar del grafo cualquier arco V🡪 AF para V ∉ Z. Luego se repite este mismo procedimiento para cada uno de los canales BF y CF y al terminar la revisión se verifica si se generó una partición. Si no se genera una partición, se repite el proceso quitando las conexiones en orden de mínima perdida de información, y verificando cada vez si se ha generado una partición. Aqui ya depende de como quieran aterrizar la estrategia, yo solo les muestro la posibilidad de remover conexiones pero como lo buscan y como lo hacen ya depende de ustedes.

