

# Tarea 8

Saul Ivan Rivas Vega

Diseño y análisis de algoritmos

Equipo Completo:

Yadira Fleitas Toranzo

Diego de Jesús Isla Lopez

Saul Ivan Rivas Vega

27 de noviembre de 2019

## 1. Ejercicio 1.

Considera el siguiente problema de *formato de textos*. Como entrada tenemos un arreglo  $W[1, \dots, n]$  tal que cada  $W[i]$  es una palabra; supondremos que el último símbolo en  $W[i]$  es un espacio en blanco. También recibimos un entero **long\_linea**  $> 0$  como parte de la entrada, tal que **long\_linea** es mayor o igual a la longitud de cada palabra  $W[i]$ , denotada  $|W[i]|$ . Lo que buscamos es dar *formato* al texto en  $W$  para esto hay que decidir cuales palabras van juntas en la misma línea. Si decidimos poner en una misma línea las palabras  $W[i], \dots, W[i']$ ,  $i \leq i'$ , lo que se denota  $l[i, i']$ , entonces la *penalización* de  $l[i, i']$  es:

- $\infty$  si **long\_linea**  $< |W[i]| + \dots + |W[i']|$  (es decir, las palabras no caben en una sola línea);
- de otra forma,  $(\mathbf{long\_linea} - (|W[i]| + \dots + |W[i']|))^3$

Entonces, la *penalización* de un *formato*  $l_1[1, i_1], l_2[i_1 + 1, i_2], \dots, l_j[i_{j-1} + 1, n]$  de  $W$  es la suma de las penalizaciones de sus líneas.

### 1.1. a) Demuestra que la siguiente estrategia *greedy* produce *formatos* arbitrariamente malos, es decir, con penalizaciones arbitrariamente grandes

Procesamos las palabras de  $W[1]$  a  $W[n]$ ; si aún hay espacio suficiente para meter  $W[i]$  en la línea actual, entonces la metemos; de otra forma iniciamos una nueva línea en la que ponemos a  $W[i]$ .

**Demostración.** Podemos demostrarlo dando un ejemplo donde esto sucede. Sea  $n = 3$ , **long\_linea** = 6 y con las longitudes  $|W[1]| = 3$ ,  $|W[2]| = 3$  y  $|W[3]| = 2$ .

- El algoritmo tomará la línea actual con espacio disponible de 6 y procesará la primera palabra,  $W[1]$ .
- Como el espacio disponible es 6 y  $|W[1]| = 3$  entonces la metemos en la línea actual.
- Ahora el espacio disponible es 3 y como  $|W[2]| = 3$  entonces la metemos en la línea actual.
- Como el espacio disponible es 0 y  $|W[3]| = 2$  creamos una nueva línea y metemos a  $W[3]$  en ella.

Al terminar el algoritmo terminamos con un formato  $f_1$  con las líneas  $l_1[1, 2]$  y la línea  $l_2[3, 3]$ . Veamos sus *penalizaciones*.

En el caso de  $l_1[1, 2]$ ,  $|W[1]| + |W[2]|$  es igual a *long\_linea* entonces entra en el segundo caso de la función de *penalización* entonces la calculamos con:

$$\begin{aligned}
penalizacion(l_1[1, 2]) &= (long\_linea - (|W_1| + |W_2|))^3 \\
&= (long\_linea - (3 + 3))^3 \\
&= (long\_linea - 6)^3 \\
&= (6 - 6)^3 \\
&= (0)^3 \\
&= 0
\end{aligned} \tag{1}$$

Para  $l_2[3, 3]$ ,  $|W[3]|$  es menor a *long\_linea* entonces entra en el segundo caso de la función de *penalización* entonces la calculamos con:

$$\begin{aligned}
penalizacion(l_2[3, 3]) &= (long\_linea - (|W_3|))^3 \\
&= (long\_linea - (2))^3 \\
&= (long\_linea - 2)^3 \\
&= (6 - 2)^3 \\
&= (4)^3 \\
&= 64
\end{aligned} \tag{2}$$

Ahora la *penalización* del formato  $f_1$  es la suma de las dos penalizaciones de sus líneas:

$$penalizacion\ f_1 = 64 + 0 = 64 \tag{3}$$

Sin embargo, tomemos el siguiente formato  $f_2$  con las líneas,  $l_1[1, 1]$  y  $l_2[2, 3]$ . Para  $l_1[1, 1]$  tenemos:

$$\begin{aligned}
penalizacion(l_1[1, 1]) &= (long\_linea - (|W_1|))^3 \\
&= (long\_linea - (3))^3 \\
&= (long\_linea - 3)^3 \\
&= (6 - 3)^3 \\
&= (3)^3 \\
&= 27
\end{aligned} \tag{4}$$

Y para  $l_2[2, 3]$  tenemos:

$$\begin{aligned}
penalizacion(l_2[2, 3]) &= (long\_linea - (|W_2| + |W_3|))^3 \\
&= (long\_linea - (3 + 2))^3 \\
&= (long\_linea - 5)^3 \\
&= (6 - 5)^3 \\
&= (1)^3 \\
&= 1
\end{aligned} \tag{5}$$

Por lo tanto la *penalización* total para el formato  $f_2$  es:

$$\text{penalización } f_2 = 27 + 1 = 28 \quad (6)$$

Finalmente como  $f_1$  es mayor que  $f_2$  y además podemos formar arreglos agregando las mismas 3 longitudes del ejemplo en ese orden decimos que el algoritmo produce *formatos* con penalizaciones arbitrariamente grandes.

**1.2. b) Usa la técnica de programación dinámica para diseñar un algoritmo que encuentre un formato con penalización mínima. Demuestra la correctitud de tu algoritmo y haz un análisis de tiempo y espacio.**

Tomemos la siguiente función para calcular el formato óptimo para las primeras  $i$  palabras:

$$OPT(i) = \begin{cases} 0, & i = 0 \\ \min_{j \in [1, \dots, i]} (OPT(j-1) + FuncionP(l[j, i])), & \text{en otro caso} \end{cases} \quad (7)$$

OPT toma la  $i$ -ésima palabra y busca la penalización mínima de colocarla en una línea que empieza en  $j$  donde  $j$  está en  $[1, \dots, i]$ , así se calcula la penalización de la línea  $l[j, i]$  y se suma al óptimo con las primeras  $(j-1)$  palabras, lo cual se calcula con  $OPT(j-1)$ .

Donde la función de penalización *FuncionP* está definida como en el enunciado del ejercicio:

$$FuncionP(l[i, j]) = \begin{cases} \infty, & \text{long\_linea} < |W[i]| + \dots + |W[j]| \\ (\text{long\_linea} - (|W[i]| + \dots + |W[j]|))^3, & \text{en otro caso} \end{cases} \quad (8)$$

Ahora con la función de optimización podemos escribir un algoritmo de programación dinámica como el siguiente:

**Data:** Arreglo de palabras  $W$ , entero positivo  $long\_linea$ .

**Result:** Tupla  $(minp, L)$  Que representa el valor de la penalización mínima y un arreglo de líneas  $L$  donde cada elemento es una tupla  $(i, j)$  que representa una línea que comienza con la palabra en  $i$  y termina con la palabra en  $j$ .

```
/* Inicializamos un arreglo para guardar las penalizaciones minimas de tamaño n,
   donde n = |W|. */
1  M = [n];
   /* Inicializamos el caso base. */
2  M[0] = 0;
   /* Recorremos todos los valores desde 1 hasta n. */
3  for i = 1, i ≤ n do
   /* Ahora para cada posición i revisaremos cual es la mejor línea que incluye a
      i, que es revisar las líneas que comienzan en una posición j ∈ [1, ..., i]. */
4  minp = ∞;
5  for j = 1, j ≤ i do
   /* Nos vamos quedando con el menor minp que veamos, en cada iteración
      obtenemos la penalización de la línea más el formato óptimo de las
      primeras j - 1 palabras. */
6  minp = min(minp, M[j - 1] + FuncionP(W, long_linea, j, i))
7  end
   /* Guardamos la menor penalización para las primeras i palabras. */
8  M[i] = minp;
9  end
   /* Inicializamos el arreglo respuesta L. */
10 L = [];
   /* La menor penalización se encuentra en M[n], ahora reconstruyamos la respuesta
      que es recorrer las posiciones que registraron el mínimo en cada i desde el
      final, Recorreremos el arreglo con un while por que no es necesario recorrer
      todas las posiciones i ∈ [1, ..., n], sino solamente las que pertenecen a la
      respuesta óptima */
11 i = n;
12 while i > 0 do
13   for j = 1, j ≤ i do
   /* revisamos si para la línea que termina en la palabra i empezar en j fue
      lo óptimo */
14   if M[j - 1] + FuncionP(W, long_linea, j, i) = M[i] then
   /* En ese caso agregamos la tupla a la respuesta y además pasamos a la
      siguiente posición i */
15     L ← (j, i);
16     i = j - 1;
   /* Interrumpimos el actual ciclo For */
17     break;
18   end
19 end
20 end
21 return (M[n], L);
```

**Algorithm 1:** Algoritmo FORMATODP.

**Data:** Arreglo de palabras  $W$ , entero positivo  $long\_linea$ , índice de inicio  $ini$ , índice de fin  $fin$ .

**Result:** Un valor de penalización de incluir a las palabras de  $W[ini, \dots, fin]$  en una sola línea.

```
/* Obtenemos la suma de las longitudes de las líneas. */
1 sumaW = 0;
2 for i = ini, i ≤ fin do
3   | sumaW = sumaW + |W[i]|;
4 end
/* Evaluamos la condición para la penalización con respecto a long_linea. */
5 if sumaW > long_linea then
6   | /* En este caso la suma de las longitudes es mayor a long_linea. */
7   | return ∞;
7 end
8 return (long_linea - sumaW)3;
```

**Algorithm 2:** Algoritmo FuncionP.

### 1.2.1. Demostración de correctitud

**El algoritmo FuncionP termina.** El algoritmo termina pues las únicas operaciones que realiza son evaluaciones y asignaciones que podemos considerar constantes, con algunas en un ciclo For, el cual realizara  $fin - ini + 1$  iteraciones, donde  $ini, fin \in [1, \dots, n]$ , por lo tanto la mayor cantidad de iteraciones que se realizan en una invocación del algoritmo es cuando  $ini = 1$  y  $fin = n$  y se harán  $n - 1 + 1 = n$  iteraciones. Finalmente decimos que el algoritmo FuncionP termina después de  $n$  iteraciones.

**El algoritmo FORMATODP termina.** El algoritmo en una primera parte realiza dos ciclos For anidados, como estos recorren  $n$  posiciones a lo mas podemos decir que realiza  $n^2$  iteraciones, sin embargo en cada iteración ademas de realizar comparaciones y asignaciones de tiempo constante el algoritmo además llama a FuncionP, y como vimos en el párrafo anterior este termina después de a lo más  $n$  iteraciones. Por lo tanto decimos que realiza  $n^3$  operaciones. En una segunda parte de igual forma tiene dos ciclos anidados y de igual forma se recorren a lo mas las  $n$  posiciones del arreglo pues  $i$  empieza igual a  $n$  y termina hasta llegar a 0, puede brincar en una sola iteración, sin embargo también puede ser el caso que se recorran todas las  $n$  posiciones, haciendo que los ciclos anidados de igual forma realicen  $n^2$  iteraciones. En cada iteración se llama también a FuncionP, agregando otro factor de  $n$ . Entonces decimos que este segundo ciclo anidado realiza  $n^3$  iteraciones. Finalmente el algoritmo después de recorrer los dos ciclos anidados tenemos que termina realizando  $2(n^3)$  iteraciones.

**El algoritmo FuncionP es correcto.** El algoritmo implementa la evaluación y los resultados de los dos casos que se presentaron en el enunciado del problema entonces corresponde a los valores que la función debe retornar.

**El algoritmo FORMATODP es correcto.** Por inducción.

**Invariante:** En el formato óptimo las líneas tienen como fin una palabra  $i$ , además tienen como inicio una palabra  $j \in [1, \dots, i]$ , y la  $(j - 1)$ -ésima palabra es la palabra al final de la siguiente línea.

**Caso base:** Hay 0 palabras en el arreglo, el formato óptimo no contempla ninguna línea, la penalización es 0. **Hipótesis de inducción:** El algoritmo obtiene el inicio  $j \in [1, \dots, i]$  con el que se obtiene el formato de penalización mínima que incluye a la línea  $l[j, i]$ . Suponer verdadero hasta la  $i$ -ésima palabra, esta  $i$ -ésima palabra debe pertenecer a alguna línea en el formato final, dicha línea puede empezar y terminar en  $i$ , siendo una línea con solo una palabra, formalmente la línea es de la forma  $l[j, i]$  donde  $j \in [1, \dots, i]$ , por nuestra suposición inicial si tomamos a algún  $j$  como nuestro inicio, ya sabemos cual es el formato de penalización mínima que incluye a la línea que termina en  $j - 1$ . Por lo tanto debemos obtener solamente aquel  $j$  tal que minimice la suma del formato óptimo con las primeras  $j - 1$  palabras mas la penalización de la línea  $l[j, i]$ . Que es recorrer todos los  $j \in [1, \dots, i]$  para obtener el mínimo de todos. Finalmente como es exactamente lo que realiza el algoritmo decimos que el algoritmo FORMATODP es correcto.

### 1.2.2. Análisis de tiempo.

**El algoritmo FuncionP tiene complejidad  $O(n)$ .** Como vimos en la sección anterior el algoritmo solo realiza un ciclo con  $n$  iteraciones ademas de operaciones como asignaciones y evaluaciones que podemos considerar de tiempo constante. Por lo tanto decimos que el algoritmo es de orden  $O(n)$ .

**El algoritmo FormatoDP tiene complejidad  $O(n^3)$ .** Como vimos en la sección anterior el algoritmo termina después de realizar  $2n^3$  iteraciones adicionalmente se realizan operaciones como asignaciones y evaluaciones que podemos considerar de tiempo  $O(1)$ . Por lo tanto decimos que el algoritmo es de orden  $O(2n^3)$  y tomando el factor dominante finalmente decimos que la complejidad es  $O(n^3)$ .

### 1.2.3. Análisis de espacio.

**El algoritmo FuncionP tiene espacio de orden  $O(1)$ .** Adicional a la entrada el algoritmo solo hace uso de una variable *sumaW* e  $i$  para el ciclo. Dichas variables no ocupan mas espacio conforme crece  $n$ , por lo tanto se mantiene constante y finalmente decimos que el algoritmo ocupa espacio de orden  $O(1)$ .

**El algoritmo FORMATODP ocupa espacio de orden  $O(n)$ .** Adicional a la entrada el algoritmo solo usa un arreglo de  $n$  posiciones para guardar los mínimos y un arreglo de tuplas  $L$  que en el caso que cada palabra este en su propia línea, este será de tamaño  $2n$ , pues se guarda una variable de inicio y una de fin. Por lo tanto el algoritmo ocupa  $3n$  mas unas variables constantes de asignación o para recorrer ciclos. Finalmente decimos que el algoritmo ocupa espacio de orden  $O(n)$ .

### 1.3. c) Modifica tu algoritmo para que funcione con una función de penalización de línea dada; la penalización de un formato sigue siendo la suma de las penalizaciones de sus líneas.

**Demostración de por que el algoritmo FORMATODP funciona con una función de penalización de línea dada.** Si la función con la que se calcula la penalización de una línea es modificada, esto repercute en la construcción y lo que devuelve el algoritmo FuncionP, sin embargo en FORMATODP esto solo afecta en el valor que devolverá en una  $i$ -ésima iteración y el formato óptimo sigue siendo la suma de las penalizaciones de las líneas, esa no ha cambiado y FORMATODP sigue minimizando esa suma.

Por contradicción.

Supongamos que existe un formato óptimo  $F'$  el cual no fue contemplado por el algoritmo y tiene una suma de penalizaciones estrictamente menor al formato  $F$  que devolvió el algoritmo. Tomemos las líneas desde la última hasta la primera. Si las líneas fueran las mismas entonces los formatos son los mismos y terminamos. En otro caso tomemos la primer línea desde el final en la que son distintas, y sea  $i$  la palabra que esta al final de esa línea. Tanto en el formato  $F'$  como en  $F$  la palabra  $i$  es la misma, pues fueron iguales en las líneas posteriores a  $i$  hasta el final. Entonces solo difieren en el inicio. Sea  $j' \in [1, \dots, i]$  el inicio de la línea en el formato  $F'$  y sea  $j \in [1, \dots, i]$  el inicio de la línea en el formato  $F$ , de tal forma que la suma de penalizaciones con las primeras  $j' - 1$  palabras mas la penalización de la línea  $l[j', i]$  es menor a la suma de penalizaciones con las primeras  $j - 1$  palabras mas la penalización de la línea  $l[j, i]$ . Eso significa que el algoritmo no reviso dicho  $j'$  puesto que de haberlo hecho lo hubiera tomado pues se queda con el mínimo. Sin embargo esto es una contradicción pues el algoritmo revisa todos los  $j \in [1, \dots, n]$  y no puede ser que lo haya ignorado y a la vez no. Finalmente decimos que el algoritmo FORMATODP funciona con una función de penalización de línea dada.

## 2. Ejercicio 2

Considera un arreglo  $A[1, \dots, n]$  con enteros positivos en sus entradas. Decimos que una pareja  $(i, j)$  es un *declive* si  $i \leq j$  y  $A[i] \geq A[j]$ ; la *longitud* de  $(i, j)$  es  $A[i] - A[j]$ . Diseña un algoritmo de tiempo y espacio  $o(n^2)$  que calcule un declive de  $A$  de longitud máxima (es o pequeña, investiga ese concepto). Demuestra que tu algoritmo es correcto y haz el análisis de tiempo y espacio.



**Data:** Arreglo  $A$ .

**Result:** Tupla  $(max\_long, r\_i, r\_j)$  Que representa el valor de la longitud máxima de un declive en  $A$  y los índices  $r\_i, r\_j$  de dicho declive  $(r\_i, r\_j)$ .

```
/* Inicializamos una cola de prioridad donde guardaremos tuplas  $(A[i], i)$  donde
   ordenaremos mayor a menor por  $A[i]$ . */
1  $Q = \{\}$ ;
   /* Inicializamos las respuestas actuales. */
2  $max\_long = -\infty$ ;
3  $r\_i = 0$ ;
4  $r\_j = 0$ ;
   /* Recorremos del final al inicio el arreglo  $A$ , pasando por todas las posiciones
   desde  $n$  hasta 1. */
5 for  $i = n, i \geq 1$  do
   /* Insertamos el elemento actual a la cola de prioridad. */
6    $(A[i], i) \rightarrow Q$ ;
   /* Obtenemos el elemento al frente de  $Q$ , es decir la tupla  $(A[i], i)$  con menor
    $A[i]$ , no lo removemos de la cola de prioridad solo obtenemos sus valores. */
7    $(valor, indice) \leftarrow Q$ ;
   /* Como procesamos de fin a inicio este elemento cumple que  $i \leq indice$ , ahora
   entonces revisamos si  $A[i] \geq valor$  lo cual sería un declive  $(i, indice)$ . */
8   if  $A[i] \geq valor$  then
   /* Calculamos la longitud del declive  $(i, indice)$ . */
9      $current\_long = A[i] - valor$ ;
   /* Actualizamos nuestra respuesta si es mayor a la que llevabamos. */
10    if  $current\_long > max\_long$  then
11       $max\_long = current\_long$ ;
12       $r\_i = i$ ;
13       $r\_j = indice$ ;
14    end
15  end
16 end
17 return  $(max\_long, r\_i, r\_j)$ ;
```

**Algorithm 3:** Algoritmo MAXDECLIVE.

## 2.1. Demostración correctitud.

**El algoritmo termina.** El algoritmo termina puesto que solo realiza un loop con  $n$  iteraciones en las cuales se realizan asignaciones y comparaciones que podemos asumir de tiempo constante, sin embargo también se realizan inserciones en una cola de prioridad que como hemos revisado anteriormente, son de tiempo  $\log n$ . Por lo tanto podemos decir que el algoritmo termina después de  $n$  iteraciones, posteriormente en el análisis temporal desarrollaremos la complejidad de cada iteración.

**El algoritmo devuelve un declive de longitud máxima.** Podemos partir de explicar cuál es un algoritmo que por fuerza bruta encuentra la respuesta. Es un algoritmo que solo intentará todas las posibles parejas  $i, j$  donde  $i \leq j$  y verificará si es un declive, conservando el declive con longitud máxima, lo cuál sería de tiempo  $O(n^2)$  lo cual no cumple con nuestra restricción temporal de  $o(n^2)$ . Sin embargo tomemos la siguiente observación.

**Observación 1** No es necesario revisar todos los  $j$  tales que  $j \geq i$  para encontrar un declive de longitud máxima que tiene a  $i$  como primer elemento, solo es necesario revisar al elemento  $j$  tal que  $j \in [i, \dots, n]$  y además tenga el menor valor en  $A[i, \dots, n]$ .

**Demostración de la Observación 1** Por contradicción.

Supongamos que el declive de longitud máxima que tiene a  $i$  como primer elemento no es con el  $j$  tal que  $j \in [i, \dots, n]$  y además tenga el menor  $A[j]$ , sino con un  $j'$  tal que por lo menos cumple que  $j' \in [i, \dots, n]$  y además  $A[j] \leq A[j']$  para ser un declive. Si  $j'$  tiene el menor valor en  $A[i, \dots, n]$  entonces significa que  $j' = j$ , en otro caso significa que  $A[j] < A[j']$ , y podemos desarrollar la siguiente ecuación:

$$\begin{aligned} A[j] &< A[j'] \\ -A[j] &> -A[j'] \\ A[i] - A[j] &> A[i] - A[j'] \end{aligned} \tag{9}$$

Lo cual significa que la longitud del declive tomando al elemento  $j$  es mayor al declive tomando al elemento  $j'$ , pero esto es una contradicción puesto que en la suposición inicial el declive  $(i, j')$  era el de longitud máxima. Finalmente como no puede existir dicho  $j'$  decimos que el declive de longitud máxima que tiene como primer elemento a  $i$  es  $(i, j)$  donde el elemento  $j$  cumple que  $j \in [i, \dots, n]$  y además tiene el menor valor en  $A[i, \dots, n]$ .

Ahora tomando la **Observación 1** podemos demostrar la correctitud del algoritmo por contradicción. Supongamos que el declive de longitud máxima  $(i', j')$  no fue devuelto por el algoritmo. Por la **Observación 1** para dicho declive significa que  $j'$  cumple que  $j' \in [i', \dots, n]$  y además tiene el menor valor en  $A[i', \dots, n]$ . El algoritmo debió haber ignorado a  $i'$  puesto que para cuando se encuentra en el índice  $i$  justamente revisa el elemento  $j$  tal que cumple que  $j \in [i, \dots, n]$  y además tiene el menor valor en  $A[i, \dots, n]$ . Sin embargo, el algoritmo si revisa todos los índices  $1 \leq i \leq n$  pero esto es una contradicción puesto que no puede haber revisado e ignorado al índice  $i'$ . Finalmente decimos que el algoritmo devuelve un declive de longitud máxima.

## 2.2. Análisis de tiempo

**El algoritmo es de complejidad  $O(n \log n)$  y por lo tanto  $o(n^2)$**  Como vimos en la sección donde se explicaba que el algoritmo termina, el algoritmo solo realiza un loop de fin a inicio por los  $n$  elementos en  $A$ , sin embargo en cada iteración además de las operaciones

de asignación y comparación que podemos asumir de tiempo  $O(1)$ , también tenemos operaciones de inserción y consulta a una cola de prioridad que como anteriormente hemos visto son de complejidad  $O(\log n)$ . Por lo tanto el algoritmo es de complejidad  $O(n \log n)$  y como para toda constante  $c > 0$  existe un  $n_0$  tal que para toda  $n \geq n_0$  se cumple que  $n \log n < n^2$  decimos que el algoritmo es  $o(n^2)$ .

## 2.3. Análisis de espacio

**El algoritmo ocupa espacio de orden  $O(n)$**  Adicional a la entrada el algoritmo ocupa algunas variables individuales para las respuestas y consultar a la cola de prioridad las cuales podemos considerar de orden  $O(1)$ . Sin embargo también hacemos uso de una cola de prioridad y adicionalmente no removemos los elementos que insertamos en ella y como cada elemento es una pareja, tendremos al final  $2n$  valores en la cola de prioridad. Como  $2n$  es el factor dominante decimos que el algoritmo tiene espacio de orden  $O(2n)$  y finalmente  $O(n)$ .

## 3. Ejercicio 3

Considera el problema de selección de centros visto en clase. Demuestra que el siguiente algoritmo devuelve un conjunto  $C$  con a lo más  $k$  centros tal que  $rc(C) \leq 2rc(C^*)$ , donde  $C^*$  es un conjunto con a lo más  $k$  centros óptimo, es decir con radio de cobertura mínimo. Puedes suponer como correctos todos los algoritmos y afirmaciones vistas en clase.

```
Alg. SelecCentros(S, k)
  If |S| <= k then
    return S
  Else
    Inicializa C con cualquier elemento s de S
    While |C| < k do
      Elegir cualquier elemento s en S que maximice dist(s,C)
      Agregar s a C
    endwhile
    return C
```

**Demostración.** Tomemos el algoritmo visto en clase y supongamos que es correcto:

```

Alg. SelecCentrosRadio(S, k, r)
  S' = S
  C = {}
  While S' != {} do
    Selecciona cualquier s en S' y agregalo a C
    Remover de S' toda ciudad t tal que dist(s,t) es a lo mas 2r
  Endwhile
  If |C| es a lo mas k then
    return C
  Else
    return NO hay k centros en S con radio de cobertura r

```

Este algoritmo *SelecCentrosRadio* como vimos en clase podemos suponer correctas las siguientes afirmaciones.

**Afirmación 1.** Si el algoritmo construye un  $C$  tal que  $|C| \leq k$ , entonces existen un conjunto con a lo más  $k$  centros de  $S$  cuyo radio de cobertura es a lo más  $2r$ .

**Afirmación 2.** Si el algoritmo construye un  $C$  tal que  $|C| > k$ , entonces no existe un conjunto con a lo más  $k$  centros de  $S$  cuyo radio de cobertura es a lo más  $r$ .

Revisemos que pasa cuando evaluamos en este algoritmo *SelecCentrosRadio* con el radio de cobertura del óptimo, es decir  $rc(C^*)$ .

Entonces el algoritmo no construye un  $C$  tal que  $|C| > k$  puesto que de haberlo hecho hubiéramos concluido que no es posible tener dicho radio de cobertura  $rc(C^*)$  ya que el algoritmo y la **Afirmación 2** son correctos.

Por lo tanto concluimos que el algoritmo forma un conjunto  $C$  tal que  $|C| \leq k$  y además que el radio de cobertura será a lo más de 2 veces  $rc(C^*)$ .

Una invariante de este algoritmo vista en clase, es que para cualesquiera dos centros de  $C$  se cumple que  $c, c' \in C$   $dist(c, c') > 2r$ . Esto es por que en efecto al seleccionar un elemento de  $S$  como centro estamos descartando a todos los que se encuentren a una distancia de a lo más  $2r$ . En caso de que en una  $i$ -ésima iteración todas las ciudades restantes se encuentren a esa distancia, serán descartadas como posibles centros.

Retomemos al primer algoritmo, a *SelecCentros*, en este algoritmo se selecciona como siguiente centro aquel que maximice su distancia a cualquier centro en  $C$ , si esta fuera una distancia de a lo más  $2r$  caeríamos en el caso discutido en el párrafo anterior, que en este punto no se requieren mas centros. Por lo tanto para seleccionar un siguiente centro para construir una solución como la de *SelecCentrosRadio* el siguiente centro debe estar a una distancia mayor que  $2r$  a cualquier centro en  $C$ , es decir este algoritmo selecciona centros

que cumplen con las mismas propiedades y formar el mismo  $C$  o alguno que cumpla lo mismo.

Por lo tanto como con el algoritmo visto en clase de *SelecCentrosRadio* forma un conjunto de centros cuya distancia entre ellos es mayor a  $2rc(C^*)$ , y con el algoritmo de este ejercicio *SelecCentros* vimos que en efecto selecciona centros que serán mayores a  $2rc(C^*)$  y en el momento que la distancia máxima sea menor o igual a  $2rc(C^*)$  significa que ya se cumple que todos los puntos están a lo más esa distancia de cualquier centro en  $C$ . Finalmente como los centros seleccionados en ambos algoritmos cumplen con las mismas propiedades de distancia a las demás ciudades decimos que el algoritmo es correcto.

## 4. Ejercicio 4

Escribe una versión recursiva del algoritmo probabilístico de corte mínimo visto en clase (basado en contracciones de aristas). Demuestra que el conjunto de aristas devuelto por el algoritmo es efectivamente un corte de la gráfica inicial.

**Data:** Gráfica  $G = (V, E)$ .

**Result:** Corte de  $G$ .

```

/* Validamos si es que ya tenemos un corte. */
1 if  $|V| = 2$  then
    /* Devolvemos el conjunto restante de aristas */
2     return  $E$ ;
3 end
/* Tomamos un par de vertices aleatoriamente de  $V$ . */
4  $(u, v) = \text{GetRandomV}(V)$ ;
/* Comprimos el grafo, es decir que eliminamos a  $u$  y a  $v$  de  $V$ , y en  $E$ 
   eliminamos las aristas de la forma  $(v - u)$  y la forma  $(u - v)$ , posteriormente
   agregamos un nuevo nodo  $w$  distinto a cualquier nodo en  $V$  y en  $E$  reemplazamos
   las aristas de la forma  $(u, x)$  y de la forma  $(v, y)$  para las aristas que
   conectaban a  $u$  y  $v$  con el resto de los nodos  $x, y \in V$  por aristas de la forma
    $(w, x)$ ,  $(w, y)$ . Teniendo como resultado un grafo compresado  $H$  */
5  $H = \text{ComprimeGrafo}(G)$ ;
/* Llamamos recursivamente al algoritmo, ahora sobre  $H$  para obtener un conjunto
    $E'$  de aristas de  $H$  que representan un corte de  $H$ . */
6  $E' = \text{GetCorte}(H)$ ;
/* Recuperamos de  $E'$  las aristas que son de  $E$  y que fueron alteradas para incluir
   a  $w$ . En este proceso tomamos las aristas de  $E'$  que tengan a  $w$  y buscando en  $E$ 
   determinar si eran de  $u$ , o  $v$ . Este proceso no recupera las aristas de  $(u, v)$  o
    $(v, u)$  puesto que no pertenecen al corte. */
7  $\text{CorteR} = \text{RecuperarAristas}(u, v, E, E')$ ;
8 return  $\text{CorteR}$ ;

```

**Algorithm 4:** Algoritmo GetCorte.

**Demostración.** Por contradicción.

Supongamos que el conjunto de aristas devuelto no son aristas que pertenecen al conjunto de aristas de la gráfica original. Eso significa que en el momento en que modificamos las aristas no fue posible revertir dicha modificación. Puesto que lo que hacemos al llegar al caso base solo regresamos el conjunto de aristas sin modificarlas como se realiza en las demás invocaciones del algoritmo que no son el caso base. Si no hubo dicha modificación, las aristas son las mismas que la gráfica original y terminamos. En otro caso, las modificaciones no pudieron ser revertidas. Sin embargo las modificaciones consisten en simplemente eliminar aristas que no pertenecen al corte y reemplazar tanto a  $u$  como  $v$  en las aristas originales, y si tenemos que  $w$  es distinto a cualquier otro nodo en  $V$  y también como referencia a  $E$  podemos entonces identificar las aristas que fueron modificadas y poder reemplazar a  $w$  por el valor original, ya sea  $u$  y  $v$ . Es decir que la modificación de reemplazar un nodo por otro tiene su operación inversa que es reemplazar dicho nodo por el nodo original. Y es precisamente el proceso que se realiza antes de devolver el Corte. Así pues las modificaciones son revertibles. Pero esto es una contradicción puesto que las operaciones no pueden ser revertibles y a la vez no. Finalmente decimos que el conjunto de aristas devuelto por el algoritmo es efectivamente un corte de la gráfica inicial.

## 5. Ejercicio 5

Tenemos  $n$  servidores que buscan coordinarse para ejecutar localmente la misma acción. De forma *abstracta*, los servidores llegan a un *consenso* sobre un bit, y ejecutan localmente la acción asociada. Por ejemplo, si el bit consensuado es 0, cada servidor hace *rollback* en su base de datos local, pero si el bit consensuado es 1, cada servidor hace *commit* localmente. Considera el siguiente algoritmo probabilístico para este problema:

```
Alg. ConsensoBinario
  r = 0
  While TRUE do
    1. r = r+1
    2. Cada servidor obtiene un bit aleatorio b_r con probabilidad uniforme
    3. Cada servidor comunica su bit b_r a todos los servidores
    4. Si todos los bits b_r de los n servidores son iguales, cada servidor
       ejecuta localmente la tarea correspondiente y termina
    5. De otra forma, continua
  endwhile
```

Responde a lo siguiente:

1. Demuestra que el número esperado de iteraciones para ejecutar la acción es  $O(2^n)$ .  
Tip: Modela el problema con una variable aleatoria con distribución geométrica.
2. ¿Cuál es el número esperado de iteraciones si en cada una de ellas los servidores obtienen su  $r$ -ésimo bit,  $b_r$ , llamando una función **shared\_random\_bit**( $r$ ) que devuelve el mismo  $r$ -ésimo bit a todos los servidores con probabilidad  $p$ , para alguna constante  $0 < p < 1$ ?

### 5.1. Demuestra que el número esperado de iteraciones para ejecutar la acción es $O(2^n)$ .

Modelemos el problema de la siguiente manera; Siguiendo el algoritmo vemos que termina en el momento en que todos los servidores obtienen el mismo bit. El bit lo obtiene cada servidor de manera aleatoria y con probabilidad, es decir que la probabilidad de que cierta combinación resultante de bits tiene la misma probabilidad de salir como cualquier otra.

Como ejemplo en 3 servidores las posibles combinaciones son: 000, 001, 010, 011, 100, 101, 110 y 111, son todas las posibles y cada una tienen la misma probabilidad de salir en cada iteración. Sin embargo en el momento que salga 000 o 111 los servidores el algoritmo termina.

En general para  $n$  servidores existen  $2^n$  combinaciones y solo con 2 el algoritmo termina.

Si en una iteración llamamos como éxito al evento de que salga alguna de estas 2 combinaciones, y como fracaso a que salga cualquier otra podemos modelar el problema con una variable aleatoria  $X$  con distribución geométrica.

$X$  es el número de iteraciones hasta el primer evento de éxito, es decir cuando sale alguna de las 2 combinaciones donde los bits son iguales. Esto es posible pues cada iteración es independiente, solo existe el evento de éxito y el de fracaso y la probabilidad de un evento de éxito es la misma en cada iteración. Dicha probabilidad se puede obtener de todos los resultados posibles en una iteración, los cuales son  $2^n$  y solo 2 son de éxito, como la probabilidad es uniforme podemos decir entonces que la probabilidad de que una iteración sea exitosa es de  $\frac{2}{2^n} = \frac{1}{2^{n-1}}$ .

Ahora la probabilidad de que la primera iteración del algoritmo sea exitosa es  $\frac{1}{2^{n-1}}$ , la probabilidad de éxito hasta la segunda iteración es igual a  $(1 - \frac{1}{2^{n-1}}) \times (\frac{1}{2^{n-1}})$  que es la probabilidad de fallar en la primera iteración y acertar en la segunda. La probabilidad de éxito hasta la tercera iteración es igual a  $(1 - \frac{1}{2^{n-1}})^2 \times (\frac{1}{2^{n-1}})$ , que es haber fallado en las primeras 2 iteraciones y acertar en la tercera.

En general la probabilidad de que el número de iteraciones hasta el primer evento de éxito, es decir  $X$ , sea igual a un entero positivo  $k$  es:

$$P(X = k) = (1 - \frac{1}{2^{n-1}})^{k-1} \times (\frac{1}{2^{n-1}}) \quad (10)$$

Ahora el valor esperado esta definido como la suma del producto de la probabilidad de cada evento por el valor de dicho evento. En nuestro caso si queremos el valor esperado de  $X$ , es decir el valor esperado de iteraciones hasta que los servidores coincidan y el algoritmo termine, es la suma:

$$\begin{aligned} E(X) &= \sum_{k=1}^{\infty} k \times P(X = k) \\ &= \sum_{k=1}^{\infty} k \times (1 - \frac{1}{2^{n-1}})^{k-1} \times (\frac{1}{2^{n-1}}) \end{aligned} \quad (11)$$



La ecuación 11 puede ser desarrollada de la siguiente forma equivalente:

$$\begin{aligned}
E(X) &= \sum_{k=1}^{\infty} k \times \left(1 - \frac{1}{2^{n-1}}\right)^{k-1} \times \left(\frac{1}{2^{n-1}}\right) \\
&= \left(1 \times \left(1 - \frac{1}{2^{n-1}}\right)^{1-1} \times \left(\frac{1}{2^{n-1}}\right)\right) + \sum_{k=2}^{\infty} k \times \left(1 - \frac{1}{2^{n-1}}\right)^{k-1} \times \left(\frac{1}{2^{n-1}}\right) \\
&= \left(1 \times \left(1 - \frac{1}{2^{n-1}}\right)^0 \times \left(\frac{1}{2^{n-1}}\right)\right) + \sum_{k=2}^{\infty} k \times \left(1 - \frac{1}{2^{n-1}}\right)^{k-1} \times \left(\frac{1}{2^{n-1}}\right) \\
&= \left(1 \times 1 \times \left(\frac{1}{2^{n-1}}\right)\right) + \sum_{k=2}^{\infty} k \times \left(1 - \frac{1}{2^{n-1}}\right)^{k-1} \times \left(\frac{1}{2^{n-1}}\right) \\
&= \left(\frac{1}{2^{n-1}}\right) + \sum_{k=2}^{\infty} k \times \left(1 - \frac{1}{2^{n-1}}\right)^{k-1} \times \left(\frac{1}{2^{n-1}}\right) \\
&= \left(\frac{1}{2^{n-1}}\right) + \sum_{k=1}^{\infty} (k+1) \times \left(1 - \frac{1}{2^{n-1}}\right)^k \times \left(\frac{1}{2^{n-1}}\right)
\end{aligned} \tag{12}$$

Ahora para simplificar la ecuación multipliquemos por  $\left(1 - \frac{1}{2^{n-1}}\right)$  a la ecuación 11.

$$\begin{aligned}
E(X) \times \left(1 - \frac{1}{2^{n-1}}\right) &= \left[\sum_{k=1}^{\infty} k \times \left(1 - \frac{1}{2^{n-1}}\right)^{k-1} \times \left(\frac{1}{2^{n-1}}\right)\right] \times \left(1 - \frac{1}{2^{n-1}}\right) \\
&= \sum_{k=1}^{\infty} k \times \left(1 - \frac{1}{2^{n-1}}\right)^k \times \left(\frac{1}{2^{n-1}}\right)
\end{aligned} \tag{13}$$

Podemos ahora restarle a la ecuación 12 la ecuación 13.

$$\begin{aligned}
E(X) - (E(X) \times (1 - \frac{1}{2^{n-1}})) &= [(\frac{1}{2^{n-1}}) + \sum_{k=1}^{\infty} (k+1) \times (1 - \frac{1}{2^{n-1}})^k \times (\frac{1}{2^{n-1}})] \\
&\quad - [\sum_{k=1}^{\infty} k \times (1 - \frac{1}{2^{n-1}})^k \times (\frac{1}{2^{n-1}})] \\
&= (\frac{1}{2^{n-1}}) + [(\sum_{k=1}^{\infty} (k+1) \times (1 - \frac{1}{2^{n-1}})^k \times (\frac{1}{2^{n-1}})) - \\
&\quad (\sum_{k=1}^{\infty} k \times (1 - \frac{1}{2^{n-1}})^k \times (\frac{1}{2^{n-1}}))] \\
&= (\frac{1}{2^{n-1}}) + \\
&\quad \sum_{k=1}^{\infty} [((k+1) \times (1 - \frac{1}{2^{n-1}})^k \times (\frac{1}{2^{n-1}})) - (k \times (1 - \frac{1}{2^{n-1}})^k \times (\frac{1}{2^{n-1}}))] \\
&= (\frac{1}{2^{n-1}}) + \sum_{k=1}^{\infty} (k+1 - k) \times (1 - \frac{1}{2^{n-1}})^k \times (\frac{1}{2^{n-1}}) \\
&= (\frac{1}{2^{n-1}}) + \sum_{k=1}^{\infty} (1 - \frac{1}{2^{n-1}})^k \times (\frac{1}{2^{n-1}}) \\
&= (\frac{1}{2^{n-1}}) + [(\frac{1}{2^{n-1}}) \times \sum_{k=1}^{\infty} (1 - \frac{1}{2^{n-1}})^k] \\
&= (\frac{1}{2^{n-1}}) \times [1 + \sum_{k=1}^{\infty} (1 - \frac{1}{2^{n-1}})^k] \\
E(X) - (E(X) \times (1 - \frac{1}{2^{n-1}})) &= (\frac{1}{2^{n-1}}) \times [1 + \sum_{k=1}^{\infty} (1 - \frac{1}{2^{n-1}})^k] \\
E(X) - E(X) + E(X) \times (\frac{1}{2^{n-1}}) &= (\frac{1}{2^{n-1}}) \times [1 + \sum_{k=1}^{\infty} (1 - \frac{1}{2^{n-1}})^k] \\
E(X) \times (\frac{1}{2^{n-1}}) &= (\frac{1}{2^{n-1}}) \times [1 + \sum_{k=1}^{\infty} (1 - \frac{1}{2^{n-1}})^k] \\
(2^{n-1}) \times E(X) \times (\frac{1}{2^{n-1}}) &= (2^{n-1}) \times (\frac{1}{2^{n-1}}) \times [1 + \sum_{k=1}^{\infty} (1 - \frac{1}{2^{n-1}})^k] \\
E(X) &= [1 + \sum_{k=1}^{\infty} (1 - \frac{1}{2^{n-1}})^k]
\end{aligned} \tag{14}$$

Como resultado de la ecuación 14 tenemos una serie geométrica la cual podemos desarrollar por separado:

$$\begin{aligned}
 S_n &= \sum_{k=1}^{\infty} \left(1 - \frac{1}{2^{n-1}}\right)^k \\
 &= \sum_{k=1}^{\infty} \left(\frac{2^{n-1}}{2^{n-1}} - \frac{1}{2^{n-1}}\right)^k \\
 &= \sum_{k=1}^{\infty} \left(\frac{2^{n-1} - 1}{2^{n-1}}\right)^k \\
 &= \sum_{k=1}^{\infty} \frac{(2^{n-1} - 1)^k}{(2^{n-1})^k}
 \end{aligned} \tag{15}$$

Ahora multipliquemos a 15 por  $2^{n-1}$

$$\begin{aligned}
Sn \times (2^{n-1}) &= \left[ \sum_{k=1}^{\infty} \frac{(2^{n-1} - 1)^k}{(2^{n-1})^k} \right] \times (2^{n-1}) \\
Sn \times (2^{n-1}) &= \left[ \sum_{k=1}^{\infty} \frac{(2^{n-1} - 1)^k \times (2^{n-1})}{(2^{n-1})^k} \right] \\
Sn \times (2^{n-1}) &= \left[ \sum_{k=1}^{\infty} \frac{(2^{n-1} - 1)^k}{(2^{n-1})^{k-1}} \right] \\
Sn \times (2^{n-1}) &= \left[ \sum_{k=0}^{\infty} \frac{(2^{n-1} - 1)^{k+1}}{(2^{n-1})^k} \right] \\
Sn \times (2^{n-1}) &= \left( \frac{(2^{n-1} - 1)^{0+1}}{(2^{n-1})^0} \right) + \left[ \sum_{k=1}^{\infty} \frac{(2^{n-1} - 1)^{k+1}}{(2^{n-1})^k} \right] \\
Sn \times (2^{n-1}) &= (2^{n-1} - 1) + \left[ \sum_{k=1}^{\infty} \frac{(2^{n-1} - 1)^{k+1}}{(2^{n-1})^k} \right] \\
Sn \times (2^{n-1}) &= (2^{n-1} - 1) + \left[ \left( \sum_{k=1}^{\infty} \frac{(2^{n-1} - 1)^k}{(2^{n-1})^k} \right) \times (2^{n-1} - 1) \right] \\
Sn \times (2^{n-1}) &= (2^{n-1} - 1) + [(Sn) \times (2^{n-1} - 1)] \\
Sn \times (2^{n-1}) &= (2^{n-1} - 1) \times [1 + (Sn)] \tag{16} \\
Sn \times \left[ \frac{(2^{n-1})}{(2^{n-1} - 1)} \right] &= 1 + (Sn) \\
Sn \times \left[ \frac{(2^{n-1})}{(2^{n-1} - 1)} \right] - (Sn) &= 1 \\
Sn \times \left[ \frac{(2^{n-1})}{(2^{n-1} - 1)} - 1 \right] &= 1 \\
Sn \times \left[ \frac{(2^{n-1}) - (2^{n-1} - 1)}{(2^{n-1} - 1)} \right] &= 1 \\
Sn &= 1 / \left[ \frac{(2^{n-1}) - (2^{n-1} - 1)}{(2^{n-1} - 1)} \right] \\
Sn &= \left[ \frac{(2^{n-1} - 1)}{(2^{n-1}) - (2^{n-1} - 1)} \right] \\
Sn &= \left[ \frac{(2^{n-1} - 1)}{(2^{n-1} - 2^{n-1} + 1)} \right] \\
Sn &= \left[ \frac{(2^{n-1} - 1)}{1} \right] \\
Sn &= 2^{n-1} - 1
\end{aligned}$$

Ahora sustituimos el valor resultante de la ecuación 16 en la ecuación 14.

$$\begin{aligned}
E(X) &= [1 + Sn] \\
E(X) &= [1 + (2^{n-1} - 1)] \\
E(X) &= [1 + 2^{n-1} - 1] \\
E(X) &= 2^{n-1}
\end{aligned} \tag{17}$$

Por lo tanto el valor esperado de  $X$ , que es el que buscamos, es de orden  $O(2^{n-1})$ , lo que también es  $O((1/2) \times 2^n)$  y finalmente  $O(2^n)$ .

## 5.2. ¿Cuál es el número esperado de iteraciones si en cada una de ellas los servidores obtienen su $r$ -ésimo bit, $b_r$ , llamando una función `shared_random_bit( $r$ )` que devuelve el mismo $r$ -ésimo bit a todos los servidores con probabilidad $p$ , para alguna constante $0 < p < 1$ ?

En este caso la probabilidad de éxito en una iteración es constante con respecto a  $n$ , es decir que es independiente de  $n$ , a diferencia que en el caso anterior donde con mas servidores la probabilidad es menor pues era de  $\frac{1}{2^{n-1}}$ .

Ahora con respecto a nuestra modelación anterior con la variable aleatoria  $X$  los cambios se ven reflejados en la probabilidad de que el número de iteraciones hasta el primer evento de éxito, es decir  $X$ , sea igual a un entero positivo  $k$  es:

$$P(X = k) = (1 - p)^{k-1} \times p \tag{18}$$

Y a su vez el valor esperado ahora es de la forma:

$$\begin{aligned}
E(X) &= \sum_{k=1}^{\infty} k \times P(X = k) \\
&= \sum_{k=1}^{\infty} k \times (1 - p)^{k-1} \times p
\end{aligned} \tag{19}$$

En general el valor esperado para una variable aleatoria con distribución geométrica como es el caso de  $X$  es:

$$E(X) = \frac{1}{p} \tag{20}$$

Lo cual es la respuesta para el ejercicio. Finalmente realizaremos la demostración de dicha afirmación.

**Demostración.** Partamos de la ecuación 19 desarrollandola como en el ejercicio anterior.

$$\begin{aligned}
E(X) &= \sum_{k=1}^{\infty} k \times (1-p)^{k-1} \times p \\
&= \sum_{k=0}^{\infty} (k+1) \times (1-p)^k \times p \\
&= (0+1) \times (1-p)^0 \times p + \sum_{k=1}^{\infty} (k+1) \times (1-p)^k \times p \\
&= (1) \times 1 \times p + \sum_{k=1}^{\infty} (k+1) \times (1-p)^k \times p \\
&= p + \sum_{k=1}^{\infty} (k+1) \times (1-p)^k \times p
\end{aligned} \tag{21}$$

Ahora si multiplicamos la ecuación 19 por  $(1-p)$ , tenemos:

$$\begin{aligned}
E(X) \times (1-p) &= \left[ \sum_{k=1}^{\infty} k \times (1-p)^{k-1} \times p \right] \times (1-p) \\
&= \left[ \sum_{k=1}^{\infty} k \times (1-p)^{k-1} \times p \right] \times (1-p) \\
&= \sum_{k=1}^{\infty} k \times (1-p)^{k-1} \times p \times (1-p) \\
&= \sum_{k=1}^{\infty} k \times (1-p)^k \times p
\end{aligned} \tag{22}$$

Ahora restemos la ecuación 22 a la ecuación 21.

$$\begin{aligned}
E(X) - (E(X) \times (1 - p)) &= [p + \sum_{k=1}^{\infty} (k + 1) \times (1 - p)^k \times p] - [\sum_{k=1}^{\infty} k \times (1 - p)^k \times p] \\
&= p + [\sum_{k=1}^{\infty} (k + 1) \times (1 - p)^k \times p - \sum_{k=1}^{\infty} k \times (1 - p)^k \times p] \\
&= p + [\sum_{k=1}^{\infty} (k + 1 - k) \times (1 - p)^k \times p] \\
&= p + [\sum_{k=1}^{\infty} (1) \times (1 - p)^k \times p] \\
&= p + [\sum_{k=1}^{\infty} (1 - p)^k \times p] \\
&= p + p \times [\sum_{k=1}^{\infty} (1 - p)^k] \\
&= p \times (1 + [\sum_{k=1}^{\infty} (1 - p)^k]) \\
E(X) - (E(X) \times (1 - p)) &= p \times (1 + [\sum_{k=1}^{\infty} (1 - p)^k]) \\
E(X) - (E(X) - E(X) \times p) &= p \times (1 + [\sum_{k=1}^{\infty} (1 - p)^k]) \\
E(X) - E(X) + E(X) \times p &= p \times (1 + [\sum_{k=1}^{\infty} (1 - p)^k]) \\
\left(\frac{1}{p}\right) \times E(X) \times p &= \left(\frac{1}{p}\right) \times p \times (1 + [\sum_{k=1}^{\infty} (1 - p)^k]) \\
E(X) &= (1 + [\sum_{k=1}^{\infty} (1 - p)^k])
\end{aligned} \tag{23}$$

Como resultado de la ecuación 23 tenemos una serie geométrica la cual podemos desarrollar por separado:

$$\begin{aligned}
Sn &= \sum_{k=1}^{\infty} (1-p)^k \\
&= \sum_{k=1}^{\infty} \left( \frac{p}{p} - \frac{p^2}{p} \right)^k \\
&= \sum_{k=1}^{\infty} \left( \frac{p-p^2}{p} \right)^k
\end{aligned} \tag{24}$$



Ahora multipliquemos la ecuación 24 por  $p$ .

$$\begin{aligned}
Sn \times p &= \left[ \sum_{k=1}^{\infty} \left( \frac{p-p^2}{p} \right)^k \right] \times p \\
&= \left[ \sum_{k=1}^{\infty} \frac{(p-p^2)^k}{(p)^k} \right] \times p \\
&= \sum_{k=1}^{\infty} \frac{(p-p^2)^k \times p}{(p)^k} \\
&= \sum_{k=1}^{\infty} \frac{(p-p^2)^k}{(p)^{k-1}} \\
&= \left[ \sum_{k=1}^{\infty} \frac{(p-p^2)^{k-1}}{(p)^{k-1}} \right] \times (p-p^2) \\
&= \left[ \sum_{k=0}^{\infty} \frac{(p-p^2)^k}{(p)^k} \right] \times (p-p^2) \\
&= \left[ \frac{(p-p^2)^0}{(p)^0} + \sum_{k=1}^{\infty} \frac{(p-p^2)^k}{(p)^k} \right] \times (p-p^2) \\
&= \left[ \frac{1}{1} + \sum_{k=1}^{\infty} \frac{(p-p^2)^k}{(p)^k} \right] \times (p-p^2) \\
&= \left[ 1 + \sum_{k=1}^{\infty} \frac{(p-p^2)^k}{(p)^k} \right] \times (p-p^2) \\
&= [1 + Sn] \times (p-p^2) \\
Sn \times p &= [1 + Sn] \times (p-p^2) \\
Sn \times p \times \frac{1}{(p-p^2)} &= 1 + Sn \\
Sn \times \frac{p}{(p-p^2)} - Sn &= 1 \\
Sn \times \left( \frac{p}{(p-p^2)} - 1 \right) &= 1 \\
Sn \times \left( \frac{p}{(p-p^2)} - \frac{(p-p^2)}{(p-p^2)} \right) &= 1 \\
Sn \times \left( \frac{p - (p-p^2)}{(p-p^2)} \right) &= 1 \\
Sn \times \left( \frac{p - p + p^2}{(p-p^2)} \right) &= 1 \\
Sn \times \left( \frac{p^2}{(p-p^2)} \right) &= 1 \\
Sn &= \frac{(p-p^2)}{p^2}
\end{aligned} \tag{25}$$

Continuemos desarrollando la Serie de la ecuación 25.

$$\begin{aligned}
 Sn &= \frac{(p - p^2)}{p^2} \\
 Sn &= \frac{(p)}{p^2} - \frac{(p^2)}{p^2} \\
 Sn &= \frac{1}{p} - 1
 \end{aligned} \tag{26}$$

Y sustituyendo el resultado de la ecuación 26 en la ecuación 23 Tenemos:

$$\begin{aligned}
 E(X) &= (1 + [\sum_{k=1}^{\infty} (1 - p)^k]) \\
 E(X) &= (1 + [Sn]) \\
 E(X) &= (1 + [\frac{1}{p} - 1]) \\
 E(X) &= (1 + \frac{1}{p} - 1) \\
 E(X) &= \frac{1}{p}
 \end{aligned} \tag{27}$$