Science de l'information : Série 1

L'entropie se définit comme $H(s) = -\sum p(s)\log_2(p(s)) = \sum p(s)\log_2(1/p(s))$, avec p(s) la densité de probabilité qu'un événement s survienne

Exercice 1.1 :

1. Soit $p(s) = p(M_1) = 1/3$ (Il tire au sort un message sur 3, donc la probabilité que chacun soit tiré est de 1/3). En appliquant la formule de l'entropie, nous trouvons $H(M_1) = 3*(1/3 * \log_2(3)) \approx$ **1.584**
   (Multiplication par 3 à cause de la somme. L'entropie de chaque message est égale, donc il faut multiplier l'entropie d'un des messages par 3 pour trouver l'entropie totale).
2. La probabilité que la lettre A soit tirée p(A) = 2/3, et celle de la lettre B p(B) est de 1/3. En appliquant la formule, nous trouvons $H(L_2) = 2/3*\log_2(3/2)+1/3*\log_2(3) \approx$ **0.9282**
3. Le total de lettre est de 18, avec 9 A et 9 B. La probabilité que la lettre A soit lue = lettre B = 9/18 = ½. L'entropie $H(L_3) = 2*(1/2*\log_2(2) = \log_2(2) =$ **1**
4. **La lettre et le message de Marge sont dépendants**. En effet, en déterminant la première lettre, les probabilités de chaque message changent. Si cette première lettre est un B, alors le message lu est forcément le second, alors que si elle lit un A, la probabilité du premier et 3ème message est de ½ alors que celle du second devient 0. La première lettre déterminée, tout peut changer !
5. **Les deux sont indépendants**. En effet, comme les probabilités de tirer un A ou un B sont égales (1/2), le fait de déterminer une lettre (sans savoir à quelle position) ne change absolument rien aux probabilités de chaque message.

Exercice 1.2

1. Dans cet exercice, il suffit d'appliquer la formule de l'entropie, en remplaçant p(s) par les valeurs associées à p(0), p(1), p(2) et p(3). Nous obtenons ainsi les entropies suivantes :
   H(A) = 2
   H(B) = 1
   H(C) = 1.75
   H(D) = H(C) = 1.75
   H(E) = 0
2. **La source de la plus grande entropie est A, et celle de la plus petite est E.**

Exercice 1.3

1.
   a. Soit p(s) la densité de probabilité de s. Alors :
      p(E) = 2/8 (elle apparaît 2x dans un total de 8 lettres)
      p(N) = p(T) = p(R) = p(O) = p(I) = 1/8.
   b. **Oui, elles sont indépendantes**. La probabilité de tirer une lettre L_1 est de 2/8 pour E, et de 1/8 pour les autres. Comme la lettre est remise en place après la lecture, les probabilités restent inchangées lors du tirage de L_2
   c. 6 lettres apparaissent une fois, et 1 lettre apparaît 2 fois. $H(L_1)$ est donc égale à $6*(1/8*\log_2(8))+2*(1/8*\log_2(8))$ = **2.75**
   d. $H(L_1) = H(L_2)$ = **2.75** (car les densités de probabilités ne changent pas)
   e. Comme L_1 et L_2 sont indépendants, $H(L) = H(L_1) + H(L_2) = 2.75 + 2.75$ = **5.5**
2.
   a. Comme Mr. Burns tire ses lettres en même temps, et en les choisissant, on peut dire que **les densités de probabilité de B_1 et de B_2 sont égales à 1**. En effet, il sait quelles lettres il va tirer. Donc les choisit mais dans un sens ou l'autre. Par exemple, s'il choisit le mot NO, en il pourra choisir $B_1 = N$ et $B_2 = O$, ou l'inverse. B_1 sera donc soit N, soit O, pareil pour B_2, donc les densités de probabilités sont de 1. Pour B en revanche, comme il aura forcément le mot voulu, **la densité de probabilité est de 1** (100% de chances d'obtenir le mot voulu).
   b. Comme les lettres sont tirées en même temps, elles sont **dépendantes**.
   c. H(B) se définit par le $1*\log_2(1)$, donc **0**. Cela semble logique, car comme le mot voulu sera forcément tiré, il n'y a aucune chance d'avoir un autre mot, donc une entropie de 0.
   d. Pareil qu'en c. L'entropie est de **0**, car la lettre voulue sera forcément tirée.
   e. Pareil qu'en c. et en d. L'entropie est de **0**, car la lettre voulue sera forcément tirée.