# Série 10: Solutions

## 1 L'entropie d'un jeu d'échecs

**a)** Voici un exemple de stratégie optimale (il y en a plusieurs équivalentes):

Q1: Est-ce une case vide? Si oui, on sait que la case est vide (avec 1 question posée). Si non, on continue:

Q2: La pièce est-elle noire? Q3: La pièce est-elle un pion?

Avec ces 2 questions supplémentaires, si la réponse à la question 3 est oui, on sait que la case contient un pion (blanc ou noir), avec donc 3 questions posées en tout. Si non, on continue:

Q4: Est-ce que la pièce est un cavalier ou un fou?

Si oui: Q5: Est-ce que la pièce est un cavalier? Avec les réponses à ces 2 dernières questions, on peut identfier la pièce (donc avec en 5 questions en tout).

Si non: Q5: Est-ce que la pièce est une tour? Si oui, on peut à nouveau identifier la pièce avec 5 questions en tout. Sinon, on doit encore poser une question: Q6: est-ce que la pièce est une dame? Et on obtient donc dans ce cas la réponse en 6 questions.

Au total, en tenant compte des probabilités d'apparition de chaque pièce sur l'échiquier, on trouve que le nombre moyen de questions à poser est:

$$\frac{32}{64} \times 1 + \frac{16}{64} \times 3 + \frac{12}{64} \times 5 + \frac{4}{64} \times 6 = \frac{1}{2} + \frac{3}{4} + \frac{15}{16} + \frac{3}{8} = \frac{41}{16} = 2.5625$$

**b)** Dans le diagramme de gauche, le nombre et la nature des pièces n'a pas changé par rapport à la situation de départ, et donc l'entropie reste la même, de la même façon que l'ordre des lettres dans une séquence de lettres n'influence pas l'entropie de la séquence: seules les probabilités d'apparition comptent dans le calcul de l'entropie.

Dans le diagramme de droite, un pion blanc a disparu. On pourrait faire ici un calcul long et compliqué pour recalculer l'entropie du jeu avec la formule de l'énoncé, mais un argument simple et intuitif permet de conclure que l'entropie diminue: avec un pion de moins sur l'échiquier, le nombre de positions différentes qu'on peut représenter est moindre que celui qu'on peut représenter au départ. Si vous n'êtes pas convaincus, pensez maintenant à la situation d'une fin de partie où il n'y a plus que les deux rois sur l'échiquier et un pion, par exemple. Dans ce cas, l'entropie du jeu est plus petite qu'au départ, car le nombre de positions différentes qu'on peut former avec ces 3 éléments sur l'échiquier est clairement bien moindre qu'avec toutes les pièces.

*Note:* Si vous avez déjà entendu parler du second principe de la thermodynamique, qui dit que l'entropie d'un système fermé augmente au cours du temps, vous avez peut-être été surpris par le résultat ci-dessus. Remarquez cependant qu'un jeu d'échecs n'est pas un système fermé, car au fur et à mesure du jeu, on pose les pièces capturées sur le bord de l'échiquier.

## 2 Calcul d'entropie

Les probabilités d'apparition des lettres sont les suivantes:

| A | espace | S, T, B | H, L, V, I , Y, ! |
|---|---|---|---|
| 5/20 | 3/20 | 2/20 | 1/20 |

L'entropie vaut donc:

$$H = \frac{5}{20} \log_2\left(\frac{20}{5}\right) + \frac{3}{20} \log_2\left(\frac{20}{3}\right) + \frac{6}{20} \log_2\left(\frac{20}{2}\right) + \frac{6}{20} \log_2(20) = \frac{17}{10} - \frac{3}{20} \log_2(3) + \frac{3}{4} \log_2(5) \quad \simeq 3.2037$$

## 3 Comparaisons d'entropies

**a)** On indique ci-dessous en **gras** le mot avec la plus grande entropie:

1. EPFL et EEPPFFLL: Les entropies sont égales (et valent 2) car dans les deux cas, les 4 lettres E, P, F et L apparaissent toutes avec probabilité 1/4.

2. AAAH et **HAHA**: L'entropie du premier mot vaut $(3/4)\log_2(4/3) + (1/4)\log_2(4) \simeq 0.81$, tandis celle du deuxième vaut 1. On peut aussi voir qu'il y a moins de possibilités de former des mots différents avec les 4 lettres AAAH qu'avec les 4 lettres HAHA, donc moins d'entropie dans le premier mot.

3. **MEDITERRANNEE** et MEDETERRENNEE: L'entropie du premier mot est clairement plus grande, car le nombre de lettres dans chaque mot est le même, mais toutes les voyelles sont remplacées par des E dans le second mot.

4. ABB et **ABBA**: L'entropie du premier mot vaut $(2/3)\log_2(3/2) + (1/3)\log_2(3) \simeq 0.91$, tandis que celle du second vaut 1.

5. **ACD** et ACDC: L'entropie du premier mot vaut $\log_2(3) \simeq 1.58$, tandis que celle du second vaut 1.5.

6. ABR et **ABRI**: L'entropie du premier mot vaut $\log_2(3)$, tandis que celle du second vaut $\log_2(4)$.

7. CALC et **CALCUL**: L'entropie du premier mot vaut 1.5, tandis que celle du second vaut $(2/3)\log_2(3) + (1/3)\log_2(6) \simeq 1.92$.

**b)** En comptant simplement le nombre d'apparitions des lettres, on déduit facilement l'ordre suivant pour les entropies:

$$1. < 2. = 4. < 3.$$

La place de la séquence N°5 est quant à elle plus difficile à déterminer. En calculant les entropies explicitement (à l'aide d'une machine à calculer ou de `http://www.shannonentropy.netmark.pl/`), on trouve que

$$1. < 2. = 4. < 5. < 3.$$

Voici ceci dit une façon de comparer "à la main" les entropies des séquences N°3 et N°5. Les nombres d'apparition des lettres sont (dans l'ordre décroissant):

| | |
|---|---|
| dans la séquence N°3: | 3 2 2 1 1 1 1 1 1 1 1 1 |
| dans la séquence N°5: | 4 2 1 1 1 1 1 1 1 1 1 1 |

La seule différence entre ces deux séquences est que dans la séquence N°3, on a 2 lettres (la 1ère et la 3ème) qui apparaissent respectivement 3 et 2 fois, tandis que dans la séquence N°5, on a deux lettres qui apparaissent respectivement 4 et 1 fois, ce qui permet de générer moins de mots différents, donc la séquence a une plus faible entropie. La comparaison des entropies des séquences N°4 et N°5 est quant à elle plus délicate...

## 4 Est-ce que l'entropie augmente ou diminue?

**a)** La réponse est non. Il se peut que l'entropie augmente (par exemple, avec ABB et ABBA) ou diminue (par exemple, avec ACD et ACDC).

**b)** Ici, la réponse est oui (par exemple, avec ABR et ABRI). La démonstration formelle de ce fait dépasse le cadre de ce cours, mais une intuition simple est la suivante: rajouter une lettre qui ne fait pas partie de la séquence amène quelque chose de *nouveau* à celle-ci et augmente donc le nombre de possibilités de former des mots différents, ce qui implique une augmentation de l'entropie.

## 5 Pour le plaisir: minimiser le nombre de pesées*

**a)** Observer tout d'abord la ressemblance entre ce problème et le jeu des questions vu au cours. Ici, à la place d'un "oracle" qui nous donne des réponses "oui" ou "non", on a maintenant une balance qui nous indique "penche à gauche", "penche à droite", ou "reste stable". Chaque pesée nous permet donc théoriquement de diviser par 3 (et non par 2) l'ensemble des possibilités. Il s'agit d'effectuer les bonnes pesées pour être le plus efficace possible.

"Numérotons" les 9 pièces par des lettres: ABCDEFGHI. Une d'entre elles est plus légère: il y a donc 9 possibilités en tout. Pour diviser par 3 l'ensemble des possibilités, nous effectuons la pesée suivante:

ABC-DEF

(signifiant qu'on place ABC à gauche et DEF à droite de la balance, les autres pièces GHI restant sur la table.)

- Si la balance penche à gauche (ce qu'on note ABC>DEF: le poids de ABC est plus grand que celui de DEF), alors on sait que la pièce plus légère est dans DEF.

- Si la balance penche à droite (ce qu'on note ABC<DEF), alors on sait que la pièce plus légère est dans ABC.

- Si la balance reste stable (ce qu'on note ABC=DEF), alors on sait que la pièce plus légère est dans GHI.

Ainsi, on a bien réduit par 3 l'ensemble des possibilités. Il suffit ensuite de répéter l'opération avce les trois pièces restantes XYZ. Plus précisément, on effectue la pesée:

X-Y

Si X>Y, la pièce plus légère est Y; si X<Y, la pièce plus légère est X; si X=Y, alors forcément, $Z$ est la pièce plus légère, toutes les autres ayant le même poids.

En conclusion, 2 pesées suffisent à trouver la pièce défectueuse. Il est à noter que $2 = \log_3(9)$. C'est la définition de l'*entropie ternaire* (alors que dans le cours et les exercices précédents, nous avons affaire à l'entropie binaire).

**b)** Pour ce problème, "numérotons" les 4 pièces par des lettres: ABCD et appelons X la pièce qu'on a dans notre poche et qu'on sait avoir le bon poids. On a de nouveau 9 possibilités:

- ou bien une des 4 pièces est plus légère que les autres;

- ou bien une des 4 pièces est plus lourde que les autres;

- ou bien toutes les pièces ont le même poids.

Pour diviser cet ensemble de 9 possibilités en 3 parties égales, voici la première pesée à effectuer:

AB-CX

(avec donc D qui reste sur la table).

1) Si AB=CX, alors il nous reste bien 3 possibiliés: soit la pièce D est plus lourde, soit elle est plus légère, soit toutes les pièces ont le même poids. On effectue donc une deuxième pesée D-X qui nous donne la réponse.

2) Si AB<CX, alors il reste également 3 possibilités: soit l'une des pièces A ou B est plus légère, soit la pièce C est plus lourde. On effectue alors la pesée A-B: si A=B, c'est C qui est plus lourde; si A<B, c'est A qui est plus légère; si A>B, c'est B qui est plus légère.

3) Si AB>CX, alors il reste également 3 possibilités: soit l'une des pièces A ou B est plus lourde, soit la pièce C est plus légère. On effectue alors la pesée A-B: si A=B, c'est C qui est plus légère; si A<B, c'est B qui est plus lourde; si A>B, c'est A qui est plus lourde.

Ainsi, au total, 2 pesées suffisent à répondre aux questions de l'énoncé. Ce n'est à nouveau pas un hasard que $2 = \log_3(9)$.