

Activité 5

Vingt devinettes - *Théorie de l'information*

Résumé

Combien d'informations un livre de 1 000 pages contient-il? Y a-t-il plus d'informations dans un annuaire de 1 000 pages, dans une rame de 1 000 feuilles de papier blanc ou dans *Le Seigneur des Anneaux* de Tolkien? Si nous pouvons mesurer cela, alors nous pouvons estimer l'espace nécessaire au stockage des informations. Par exemple, pouvez-vous lire la phrase suivante?

Ls vlls d ctt phrs nt dspr.

Vous le pouvez sûrement, car les voyelles ne véhiculent pas beaucoup d'« informations ». Cette activité est une introduction à un moyen de mesurer le contenu des informations.

Liens pédagogiques

- ✓ Mathématiques : nombres. Supérieur à, inférieur à, moyennes.
- ✓ Algèbre. Séquences et suites
- ✓ Français

Compétences

- ✓ Comparer les nombres et travailler avec les moyennes
- ✓ Dédutions
- ✓ Poser des questions

Âge

- ✓ 10 ans et plus

Matériels

- ✓ Aucun support n'est nécessaire pour la première activité
- Pour l'activité supplémentaire, les enfants ont besoin de :
- ✓ L'exercice : Arbres de décision (page 40)

Vingt devinettes

Questions pour lancer la discussion

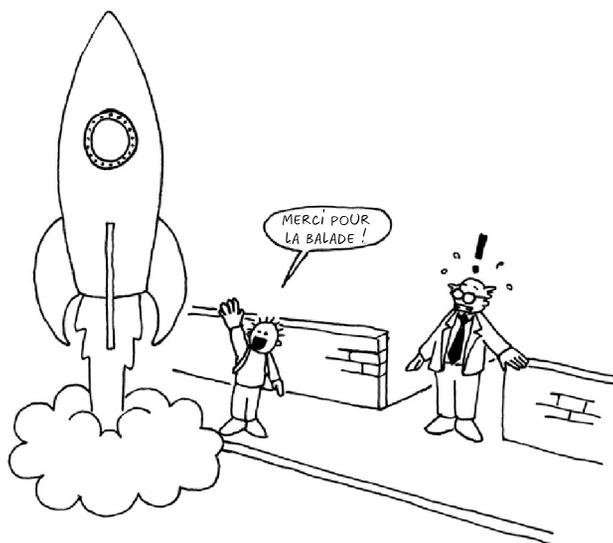
1. Demandez aux enfants comment ils définissent une information.
2. Comment peut-on mesurer la quantité d'informations contenues dans un livre? Le nombre de pages ou de mots est-il important? Un livre contient-il plus d'informations qu'un autre? Qu'en est-il si le livre est assommant ou s'il est particulièrement intéressant? Un livre de 400 pages contenant la phrase « bla bla bla » aurait-il plus ou moins d'informations qu'un annuaire téléphonique par exemple?

Expliquez que les informaticiens mesurent les informations d'après le degré de nouveauté du message (ou du livre!). Si vous entendez quelque chose que vous connaissez déjà (par exemple, un ami qui se rend toujours à l'école à pied dit « Je suis venu à l'école à pied aujourd'hui »); le message ne vous apporte pas de nouvelles informations, il ne comporte rien de nouveau. Si votre ami avait dit « Je suis venu à l'école en hélicoptère aujourd'hui », alors oui, ça *aurait* été nouveau, et nous aurions eu dans ce message beaucoup d'informations.

Comment le degré de nouveauté d'un message peut-il se mesurer?

L'une des méthodes consiste à observer la difficulté avec laquelle on devine l'information. Si votre ami dit « Devine comment je suis venu à l'école aujourd'hui? » et qu'il est venu à pied, vous devineriez sûrement tout de suite. Mais il vous faudrait plusieurs essais avant de deviner que c'est en hélicoptère et encore plus si c'était en navette spatiale.

La quantité d'informations contenue dans un message se mesure d'après la facilité ou la difficulté à les deviner. Le jeu suivant donne une idée de la complexité à deviner l'information.



Activité : vingt questions

Voici une adaptation du jeu de 20 questions. Choisissez un enfant. Les autres enfants doivent lui poser des questions auxquelles il ne peut répondre que par oui ou par non jusqu'à ce qu'ils trouvent la bonne réponse. Toutes les questions peuvent être posées, à condition que la réponse soit « oui » ou « non ».

Suggestions :

Je pense à :

- ✓ un nombre compris entre 1 et 100
- ✓ un nombre compris entre 1 et 1 000
- ✓ un nombre compris entre 1 et 1 000 000.
- ✓ un nombre quelconque.
- ✓ une suite de 6 nombres qui respecte une séquence (adaptée au groupe). Devinez les nombres dans l'ordre du premier au dernier. (par ex. 2, 4, 6, 8, 10, 12)

Comptez le nombre de questions posées. C'est la mesure de la valeur des « informations ».

Sujets de discussion

Quelles stratégies avez-vous appliquées ? Quelles étaient les plus efficaces ?

Notez qu'il ne faut que 7 propositions pour deviner un nombre compris entre 1 et 100 si vous divisez l'intervalle par deux à chaque question. Par exemple :

Le nombre est-il inférieur ou égal à 50 ?	Oui.
Le nombre est-il inférieur ou égal à 25 ?	Non.
Le nombre est-il inférieur ou égal à 37 ?	Non.
Le nombre est-il inférieur ou égal à 43 ?	Oui.
Le nombre est-il inférieur ou égal à 40 ?	Non.
Le nombre est-il inférieur ou égal à 41 ?	Non.
Ce ne peut être que 42 !	Oui !

Il est intéressant de remarquer que si on étend l'intervalle jusqu'à 1000, il ne faut pas 10 fois plus de propositions, mais 3 seulement. Chaque fois que l'intervalle double, une seule question supplémentaire suffit à trouver la réponse.

À la suite de cet exercice, il serait intéressant de faire jouer les enfants à un jeu *Mastermind*.

Activité supplémentaire : Combien d'informations un message contient-il ?

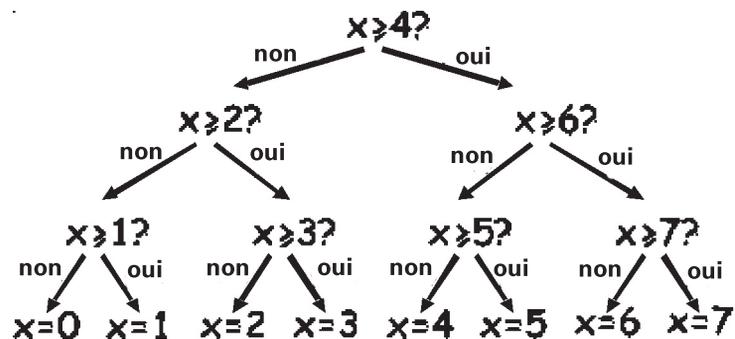
Les informaticiens ne cherchent pas uniquement à deviner des nombres, ils peuvent aussi deviner la lettre qui a le plus de chance de compléter un mot ou une phrase.

Essayez de deviner une phrase courte de 4 à 6 mots. Les lettres doivent être devinées dans l'ordre, de la première à la dernière. Quelqu'un écrit les lettres trouvées et note le nombre de propositions qui ont été faites pour chacune d'entre elles. Vous pouvez poser toutes les questions dont la réponse est oui ou non, comme « Est-ce que c'est un *t* ? » « Est-ce que c'est une voyelle ? » « Est-ce que c'est avant le *m* dans l'alphabet ? » Un espace entre les mots compte également comme une « lettre » et doit être deviné. Prenez la parole chacun à votre tour et trouvez les parties du message les plus faciles à déchiffrer.

Exercice : Arbres de décision

Si vous connaissez déjà la stratégie pour poser les questions, vous pouvez transmettre un message sans poser une seule question.

Voici un schéma appelé « arbre de décision » pour deviner un nombre entre 0 et 7 :



Quelles décisions oui/non sont nécessaires pour « deviner » le nombre 5 ?

Combien de décisions oui/non sont nécessaires pour trouver un nombre quelconque ?

Observe à présent ce qui est fascinant : Sous les nombres 0, 1, 2, 3 ... du dernier niveau de l'arbre, écris les nombres en binaire (voir Activité 1).

Regarde l'arbre attentivement. Si non = 0 et oui = 1, que constates-tu ?

Dans le jeu de devinette des nombres, nous essayons de choisir les questions pour que la suite des réponses permette de représenter le nombre de cette façon.

Construis ton propre arbre de décision pour deviner les nombres compris entre 0 et 15.

Pour les plus forts :

Quel type d'arbre utiliserais-tu pour deviner l'âge de quelqu'un ?

Quel arbre permettrait de deviner la prochaine lettre d'une phrase ?

Ce qu'il faut retenir

Claude Shannon, célèbre mathématicien américain (pratiquant également la jonglerie et le monocycle), a fait de nombreuses expériences avec ce jeu. Il a mesuré la quantité d'informations en bits (chaque oui/non étant équivalent à un bit **1 / 0**). Il a trouvé que la quantité d'« informations » contenues dans un message dépend de ce que l'on sait déjà. Parfois, nous pouvons poser une question qui élimine le besoin d'en poser de nombreuses autres. Dans ce cas, le contenu informatif du message est restreint. Par exemple, lancer une pièce représente normalement une information d'un bit : pile ou face. Mais s'il se trouve que la pièce est faussée et qu'elle tombe neuf fois sur dix sur face, alors l'information n'est plus un bit mais moins, aussi incroyable que cela paraisse. Comment connaître le résultat d'un lancer en posant moins d'une question oui/non ? C'est simple : avec une seule question comme « les *deux* prochains lancers tomberont-ils sur face ? » Pour une suite de lancers avec une pièce faussée, la réponse serait « oui » à 80 %. Dans les 20 % des cas restant, lorsque la réponse est « non », vous devrez poser deux autres questions. Mais, en moyenne, vous aurez posé moins d'une question par lancer de pièce !



Claude Shannon appelait ce contenu informatif d'un message « l'entropie ». L'entropie dépend non seulement du *nombre* de résultats possibles (dans le cas du lancer de pièce : deux), mais également de la probabilité de l'événement. Les événements improbables, ou informations nouvelles, nécessitent beaucoup plus de questions pour deviner le message car ils transmettent plus d'informations encore inconnues, tout comme dans l'exemple de l'enfant qui va à l'école en hélicoptère.

L'entropie d'un message est très importante pour les informaticiens. Il est impossible de compresser un message pour qu'il occupe moins de place que son entropie et les meilleurs systèmes de compression sont semblables à des devinettes. Puisqu'un programme informatique résout les devinettes, la liste des questions peut être reproduite plus tard, et tant que les réponses (bits) sont stockées, les informations peuvent être reconstituées ! Les meilleurs systèmes de compression peuvent réduire les fichiers texte à un quart environ de leur taille initiale (une grande économie sur l'espace de stockage !)

La méthode des devinettes peut également servir à construire une interface informatique qui prédit quel texte l'utilisateur s'apprête à saisir ! Cela peut être très utile pour les personnes présentant un handicap physique qui ont des difficultés à taper. L'ordinateur suggère ce qu'il pense que l'utilisateur va taper ensuite et ce dernier n'a qu'à indiquer si c'est bien ce qu'il souhaite. Un bon système nécessite en moyenne seulement deux questions oui/non par caractère et peut être d'une grande utilité pour quelqu'un qui peine à réaliser les petits mouvements précis nécessaires pour manipuler la souris ou le clavier. Ce type de système est également utilisé sous différentes formes pour « saisir » du texte sur les téléphones mobiles.

Solutions et astuces

La réponse à une seule question par oui ou par non correspond exactement à un bit d'information - que ce soit une question simple comme « Le nombre est-il inférieur à 50 ? » ou plus complexe comme « Le nombre est-il compris entre 20 et 60 ? ».

Dans le jeu de devinette des nombres, si les questions sont choisies d'une certaine manière, la suite des réponses est simplement codée par la représentation binaire d'un nombre (où on aurait gardé les 0 à gauche du nombre qui sont habituellement supprimés). Trois s'écrit **011** en binaire et code les réponses « non, oui, oui » dans l'arbre de décision, ce qui revient à écrire non pour **0** et oui pour **1**.

L'arbre que nous utiliserions pour l'âge d'une personne pourrait être décalé vers des nombres plus petits.

La décision concernant les lettres dans une phrase pourrait dépendre de la lettre précédente.