

Module 2 - partie 2 chapitre 3 Codage de Huffman

Fiche d'identité

Résumé de l'activité	<i>Jouer avec la compression de textes</i>
Objectif(s)	<ul style="list-style-type: none"> Comprendre à quoi sert la compression de texte Expérimenter une technique de compression L'appliquer à un petit texte à compresser/transmettre/décompresser
Durée approximative de l'activité	<ul style="list-style-type: none"> 45 mn à 1h05 selon si on fait une extension ou pas
Participants	<ul style="list-style-type: none"> autant qu'on veut, à partir de 9-10 ans
Matériel nécessaire	<ul style="list-style-type: none"> des kits de marmottes, de préférence un pour 2 enfants pour accéder au kit
Préparation	<ul style="list-style-type: none"> découper les marmottes et les couloirs, mettre au dos de chaque marmotte un nombre et une lettre, correspondant aux lettres présentes dans un texte et leur fréquence d'apparition (attention à choisir un texte avec au plus 9 caractères sinon il vous faut plusieurs kits), recopier le nombre côté marmotte mettre au dos de chaque couloir un 1 (sur la branche de droite) et un 0 (sur la branche de gauche) si possible plastifier les différents éléments et mettre un moyen d'accrocher/décrocher ensemble couloirs et marmottes (scratch adhésif, patafix)
Notions liées	<i>compression de texte</i>
Lien éventuel avec le programme scolaire	<i>on travaille les additions et les multiplications, ainsi que la comparaison de nombres</i>

Déroulement

1	Donner la consigne et expérimenter – 15'	
	<p>Creuser son premier terrier :</p> <ul style="list-style-type: none"> on pose un premier couloir, le coude de ce couloir est l'entrée au bout d'un couloir on peut mettre soit une marmotte soit le coude d'un autre couloir, pas les deux 	<ul style="list-style-type: none"> <i>les laisser expérimenter</i> <i>il faut mettre toutes les marmottes (et on a besoin d'un couloir de moins que le nombre de marmottes)</i> <i>bien prendre les papiers côté marmottes, pas côté lettres</i>
	<p>Et si les marmottes ont le sommeil léger ?</p> <ul style="list-style-type: none"> pas de marmotte à un croisement sinon elle se fait réveiller par une autre qui se lève et qui est plus loin dans le couloir chaque marmotte si elle se réveille 3 fois et est à 4 couloirs de l'entrée va parcourir $3 \times 4 = 12$ couloirs (on ne compte pas les 	<ul style="list-style-type: none"> <i>Laisser les participants expérimenter</i> <i>voir les résultats des différents groupes et comparer le "bruit" de leurs terriers</i>

<Class'Code>

	retours) Comment faire pour minimiser le bruit total du groupe de marmottes (donc le total des déplacements) ?	
--	---	--

2 Trouver le meilleur terrier à tous les coups – 10'		
	Comment optimiser un terrier ? Réfléchir ensemble aux techniques permettant de bien choisir le terrier et la position des marmottes afin de minimiser le bruit dans un terrier.	<ul style="list-style-type: none"> ▪ <i>En général ils trouvent l'argument de "ceux qui se réveillent le plus souvent doivent être proches de l'entrée"</i> ▪ <i>ça nous dit qui mettre où dans un terrier déjà creusé, mais pas quelle forme donner au terrier : tous les couloirs de la même longueur ? Un terrier en forme de peigne ? Autre chose ? Ca dépend des nombres sur les marmottes.</i>
	Un algorithme pour trouver le meilleur terrier : <ul style="list-style-type: none"> ▪ <i>détacher tout</i> ▪ <i>relier ensemble les deux marmottes qui se réveillent le moins souvent</i> ▪ <i>noter sur le coude du couloir qui les relie la somme s de leur nombre de réveils puis les considérer comme une seule marmotte qui se réveille s fois.</i> ▪ <i>recommencer à chercher les deux marmottes (ou groupes de marmottes) qui se réveillent le moins souvent et les relier</i> ▪ <i>s'arrêter quand on n'a plus qu'un groupe avec toutes les marmottes reliées entre elles</i> 	<ul style="list-style-type: none"> ▪ <i>énoncer l'algorithme</i> ▪ <i>les laisser l'expérimenter (attention à la difficulté de considérer un groupe de marmottes reliées comme une seule)</i> ▪ <i>quand on a plus de deux (groupes de) marmottes qui se réveillent le même nombre de fois, on peut choisir celle qu'on veut</i> ▪ <i>voir le gain éventuel atteint par rapport à la solution par tâtonnements</i> ▪ <i>insister sur le fait que ça marcherait même avec 20, 100 ou 1000 marmottes</i>

3 Le lien avec l'informatique ? 5'		
	Cet algorithme est dû à David Huffman et sert à la compression de texte. Voir le dos des terriers (si on a mis des scratches on peut retourner) les fréquences d'apparition, les lettres et les 0 et 1 au dos des couloirs	<ul style="list-style-type: none"> ▪ <i>la suite de 0 et de 1 depuis l'entrée jusqu'à une lettre est son code binaire compressé</i> ▪ <i>réduire le déplacement des marmottes = réduire la taille du texte compressé</i>
	Encore mieux que ce que je pensais. Mettre tout le monde d'accord en leur révélant : <ul style="list-style-type: none"> ▪ <i>que si on donne un code binaire par lettre il n'y a pas mieux que le code de Huffman</i> ▪ <i>que cet algorithme est un des ingrédients de quasi tous les formats compressés qu'ils connaissent (mp3, mp4, zip,...)</i> 	<ul style="list-style-type: none"> ▪ <i>il est rare d'avoir un algorithme simple/explicable et encore utilisé, il faut le souligner.</i> ▪ <i>"un ingrédient" = d'abord les algorithmes de compression font une étape plus compliquée et dépendante du type de donnée à compresser (image, son,...) puis on fait une passe de l'algorithme de Huffman</i>

< Class'Code >

4 Et maintenant, compressez ! 10' à 30'	
<p>On utilise l'arbre pour compresser/décompresser un texte</p> <ul style="list-style-type: none">▪ Choisir un des arbres obtenus grâce à l'algorithme donner un texte ne contenant que des lettres figurant au dos des marmottes, et leur demander de compresser▪ inversement donner un texte/mot compressé en binaire et demander aux participants de le décompresser	<ul style="list-style-type: none">▪ ne pas être trop ambitieux sur la longueur des textes▪ si on le sent, on peut faire un concours : qui pense avoir décompressé le mot lève la main, nous le dit en secret et on voit qui trouve en premier-ère▪ recommencer quelques fois pour que tout le monde ait compris
<p>Si on a le temps : transmettre un texte compressé :</p> <ul style="list-style-type: none">▪ les participants écrivent un court texte▪ ils calculent les fréquences d'apparition des lettres▪ ils calculent l'arbre de Huffman▪ ils compressent le texte grâce à l'arbre <p>Et c'est à un autre groupe de décompresser le texte à l'aide de l'arbre</p>	<ul style="list-style-type: none">▪ encore une fois prendre des textes courts avec peu de caractères différents▪ on voit au passage qu'il est facile de se tromper en binaire :)

5 Conclusion 5'	
<ul style="list-style-type: none">▪ Qu'est-ce que j'ai appris ?▪ Qu'est-ce qui était difficile ?	

Pour aller plus loin	
Une fiche complète d'activité	https://members.loria.fr/MDuflot/files/med/marmottes.html
Une vidéo explicative	https://www.youtube.com/watch?v=oqMx1cuw6mo
Une autre activité de compression de texte	https://interstices.info/jcms/c_47072/enseigner-et-apprendre-les-sciences-informatiques-a-lecole page 23 et plus