
MULTIMEDIA

Gérard Kuntz (g.kuntz@libertysurf.fr)

La rubrique multimédia de Repères-IREM est ouverte aux suggestions, aux critiques et aux contributions de ses lecteurs, qui pourront être envoyées à l'adresse ci-dessus. Elle donne aussi aux groupes de recherche des IREM la possibilité de faire connaître aux lecteurs un site ou des activités dont les technologies sont des outils essentiels pour apprendre les mathématiques. Les propositions sont bienvenues, à la même adresse !

La rubrique multimédia de ce numéro 125 a été prise en charge par le **groupe informatique de l'Irem de Strasbourg** (*), qui se présente en début d'article. Elle ne nécessite pas de version électronique, les deux liens utiles figurant en notes de bas de page.

Découvrir des concepts informatiques autour des arbres binaires Une série d'activités d'informatique débranchée

Basile SAUVAGE^{1,2}, Jean-Christophe GRIMONT^{1,3},
Régine HAMM-AUDONNET^{1,4}, Julien NARBOUX^{1,2}

RÉSUMÉ. Cet article relate une série d'activités d'informatique débranchée, construites ou adaptées par l'Irem de Strasbourg. Le concept commun est celui des arbres binaires, qui sont déclinés dans différentes activités, permettant d'aborder plusieurs concepts de la science informatique : algorithmes, complexité, codage de l'information. Les activités ont été éprouvées dans les classes de collège, même si la plupart s'adaptent du cycle 3 au lycée. Nous présentons les activités en soi, apportons une perspective scientifique, et discutons de la mise en œuvre pratique en classe. Les supports, pour les élèves et pour les enseignants, sont accessibles publiquement.

INTRODUCTION

Un groupe de travail sur le thème de l'informatique a été fondé à l'Irem de Strasbourg en 2018. Il regroupe essentiellement des enseignants-chercheurs en informatique à l'université, et des enseignants de collège et de lycée. L'informatique a une place singulière dans le secondaire, dans la mesure où elle est au programme d'autres disciplines (en particulier les mathématiques) au collège, et n'a une place de discipline à part entière au lycée que depuis 2019. En corollaire, la didactique de l'informatique est relativement peu développée en France, alors que les besoins de ressources pédagogiques et de formation vont croissant. Dans ce contexte, le groupe mène différentes actions : la construction de ressources pédagogiques ; une réflexion sur les programmes, la progression des apprentissages, et l'articulation entre disciplines ; la médiation scientifique, particulièrement via l'informatique débranchée ; la formation des enseignants et des animateurs scientifiques.

L'informatique débranchée, aussi appelée informatique déconnectée ou informatique sans ordinateur, désigne une manière de faire des activités de découverte et d'apprentissage de l'informatique, sans utiliser d'ordinateur, généralement en utilisant du petit matériel tangible. Cette approche, débarrassée des contraintes techniques numériques, a plusieurs avantages: le déploiement est facilité dans toutes les salles et pour tous les effectifs ; le côté ludique suscite l'intérêt ; l'aspect scientifique prime sur l'aspect technique.

1 Irem de Strasbourg
2 Université de Strasbourg

3 Collège Conrad Pfeffel, Colmar
4 Collège de Brumath

Dans le présent article, nous nous intéressons à une série d'activités d'informatique débranchée qui ont été testées dans les classes de collège. Le fil conducteur est celui des arbres binaires, illustrés figure 1. Si ce n'est qu'ils sont souvent représentés la tête en bas, on perçoit la parenté avec les arbres réels, dont ils reprennent le vocabulaire : racine, nœuds, feuilles, branches. Ils sont binaires en ce sens que chaque nœud donne naissance à deux branches. Cette structure d'arbre, très classique en informatique, est un support intéressant pour aborder différents concepts. Nous parlerons d'algorithmes et de complexité avec l'activité du « qui est-ce ? », de codage de l'information avec l'algorithme de Huffman⁵. Nous présenterons ensuite une déclinaison pour exercer l'arithmétique. Enfin, nous ferons le lien avec le codage binaire.

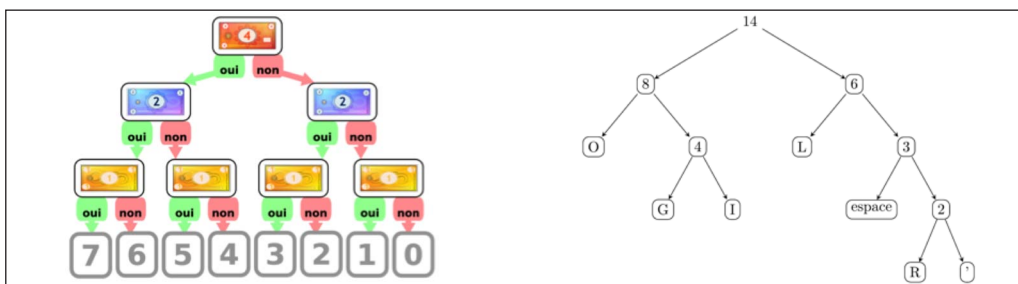


Figure 1. Les arbres en informatique sont représentés la « racine » en haut et les « feuilles » en bas. Il sont « binaires » car chaque « noeud » donne naissance à exactement deux « branches ». À gauche : arbre représentant le codage binaire des entiers. À droite : arbre représentant un codage binaire des lettres par l'algorithme de Huffman.

ALGORITHMES ET COMPLEXITÉ AVEC LE « QUI EST-CE ? »

L'activité illustrée figure 2 (voir page suivante) imite le jeu du « qui est-ce ? », qui consiste à deviner un personnage parmi n candidats, en posant des questions à réponse oui/non. Les quatre arbres de décision de la figure 3 sont distribués avec les personnages à part (non représentés ici pour des questions de droits d'auteurs). L'activité est animée en trois temps en petits groupes (2 à 4 élèves), entrecoupés de discussions en classe entière.

Dans un premier temps, les élèves placent les personnages dans les feuilles des arbres. Cela permet de comprendre le fonctionnement de l'arbre, à savoir que chaque feuille (personnage) est définie par une suite de questions (un chemin de la racine à la feuille). Dès le cycle 4, cela nécessite peu d'explications, et la plupart des élèves élaborent par eux-même une technique de remplissage en partant de la racine. Ici, l'arbre est vu comme une façon de ranger les personnages.

Dans un second temps, ils jouent en binôme, un élève faisant deviner à l'autre un personnage mystère, qui suit l'enchaînement de questions imposé par un des arbres. Ici, l'arbre est vu comme un algorithme : l'ordre des questions est imposé (notion de séquentialité), et dépend des réponses (notion de conditionnelle).

⁵ https://fr.wikipedia.org/wiki/Codage_de_Huffman

MULTIMEDIA

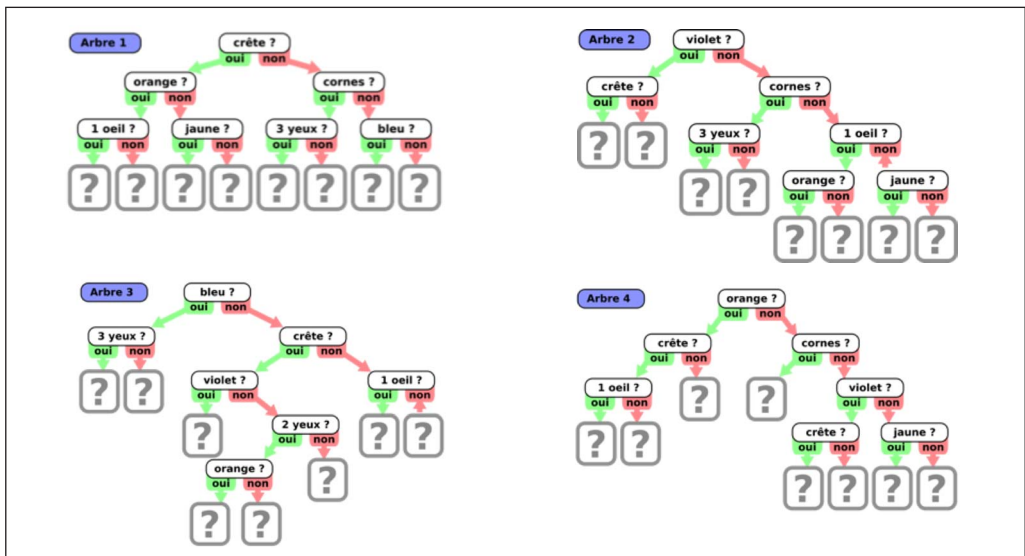


Figure 2. Arbres de décision pour le jeu du « qui est-ce ? ». Les personnages à deviner seront déposés dans les feuilles, ici représentés par avec un point d'interrogation. Une suite de nœuds en partant de la racine indique la suite de questions à poser. Les personnages (têtes de monstres avec un ou plusieurs yeux, une crête, des cornes, une couleur de peau, etc.) ne sont pas reproduits ici pour des raisons de droits d'auteurs. Les supports originaux pour cette activité sont disponibles auprès du groupe informatique à l'Irem de Strasbourg.⁶

Dans un troisième temps, les élèves déterminent le nombre de questions posées pour parvenir au but, exerçant leurs compétences en dénombrement et en calcul de moyennes. On peut faire calculer le nombre total de questions nécessaires pour faire deviner tous les personnages ou alors le nombre moyen de questions, selon le niveau des élèves. Ce nombre de questions est à rapprocher de la notion de complexité d'un algorithme qui mesure la quantité de ressources qu'il utilise : généralement le temps de calcul, comme ici, ou bien l'occupation mémoire. Elle permet de comparer l'efficacité des algorithmes entre eux. Elle est fonction de la taille du problème à résoudre (ici le nombre de personnages). Les élèves connaissent souvent le jeu Akinator, où l'ordinateur « devine » le personnage auquel le joueur pense grâce à un algorithme similaire. On peut poser la question aux élèves du nombre de personnages que l'on peut distinguer avec 2, 3, 4, ... 20 questions.

On observe que les arbres déséquilibrés permettent de trouver parfois en peu de questions, parfois en beaucoup de questions (2 à 5 sur l'arbre 3). Un arbre équilibré permet de trouver en un nombre fixe de questions (3 sur l'arbre 1). On peut vérifier qu'un arbre équilibré est plus rapi-

⁶ <https://mathinfo.unistra.fr/irem/groupe-irem/informatique/>

de en moyenne. Ces observations sont mises en parallèle avec les notions de complexité dites « au mieux », « au pire », et « en moyenne », lorsqu'un algorithme a une complexité qui varie en fonction des données. Certains élèves éprouvent une difficulté à faire le lien entre la complexité et le nombre de questions posées. En effet, il y a là une subtilité, si ce n'est une ambiguïté : dans le jeu réel, lorsqu'une question est posée, il faut parcourir tous les personnages restants ; ainsi une question engendre plusieurs opérations.

Pour les élèves de lycée, la réflexion peut être prolongée sur le logarithme en base 2 et l'exponentielle (Sauvage, 2019).

LE CODAGE DE L'INFORMATION ET L'ALGORITHME DE HUFFMAN

Cette seconde activité aborde le codage de l'information. L'articulation avec l'activité précédente peut s'appuyer sur deux observations.

D'une part, on montre qu'un chemin de la racine à une feuille peut être représenté par une suite de 0 (branche gauche) et de 1 (branche droite) définissant un code pour chaque personnage. Par exemple, le personnage orange à cornes avec deux yeux, dans l'arbre en haut à gauche figure 2, peut se coder "101", ce qui correspond aux réponses "non-oui-non". Si l'on remplace les personnages par des caractères (lettres de l'alphabet, chiffres et ponctuation), on peut définir un code binaire pour chaque caractère.

D'autre part, on montre que l'arbre équilibré est optimal à condition que les personnages soient choisis de manière équiprobable. En revanche, un arbre déséquilibré peut être plus performant pour un tirage déséquilibré : l'arbre 2 pourra être intéressant face à un joueur qui choisit plus fréquemment les personnages violets. On peut alors renverser la question : connaissant les fréquences de tirage, quel est le meilleur arbre ?

L'algorithme de Huffman répond à cette question. Étant donné un ensemble de symboles associés à une probabilité, l'algorithme construit incrémentalement un arbre optimal. Cet algorithme est utilisé pour compresser des données. Dans le cas de textes en langue naturelle, les symboles sont des caractères. En français par exemple, le 'E' est fréquent donc il vaut mieux le coder sur peu de bits (haut dans l'arbre), alors que le 'W' est rare donc ce n'est pas grave de le coder avec plus de bits (profond dans l'arbre). Ici, l'arbre est un code, dont la performance et l'optimalité se mesurent par la compacité du code pour un texte respectant les probabilités d'apparition.

Il est possible d'expliquer le principe de l'algorithme de Huffman à condition d'avoir du temps et des élèves suffisamment âgés. Nous renvoyons par exemple à l'activité « les marmottes au sommeil léger » (Dufлот-Kremer et Vincent, 2018). Au collège, nous nous contentons d'une activité de codage-décodage pour un arbre donné (figure 1, à droite), en trois temps. Dans un premier temps, les élèves associent un code à chaque caractère. Dans un second temps, les élèves codent une phrase. Nous utilisons la phrase « L'IGLOO RIGOLO », dont les fréquences (quatre 'O', deux 'I', etc.) donnent effectivement cet arbre par l'algorithme de Huffman. Dans un troisième temps, les élèves décodent une phrase. Cette activité est assez courte, et les élèves la compren-

MULTIMEDIA

nent rapidement. Pour qu'elle fonctionne, il faut faire attention à ce que les phrases ne soient ni trop courtes ni trop longues. Pour qu'elle soit instructive, il faut que l'arbre soit suffisamment déséquilibré.

Au-delà du concept de codage de l'information, cette activité est l'occasion de plusieurs développements. Premièrement, on observe que le décodage est plus difficile que le codage, car on ne sait pas a priori combien de bits il faut prendre, il dépend de la profondeur dans l'arbre. Pour autant, il n'y a pas d'ambiguïté : aucun code n'est le préfixe d'un autre code, car les caractères sont seulement dans les feuilles, jamais sur les nœuds internes. Deuxièmement, il est intéressant d'évoquer les formats zip et png, qui utilisent cet algorithme. Troisièmement, si l'arbre dépend du message (comme dans notre exemple, ou dans le cas du ZIP) et pas seulement du langage (e.g. le français), il faut prendre en compte la place mémoire occupée par l'arbre, qui doit être transmis lui aussi.

Enfin on peut poser la question de savoir s'il est possible de trouver un arbre dont toutes les feuilles seraient à une profondeur plus petite que celle de l'arbre équilibré : par un argument de cardinalité c'est impossible. Cela montre que tout système de compression ne peut pas être efficace sur toutes les entrées possibles, dans notre exemple un texte avec beaucoup de lettres rares.

LES DIVISEURS EN ARITHMÉTIQUE

Cette troisième activité se rapporte à la notion de divisibilité, étudiée en arithmétique au collège. Comme pour le « qui est-ce ? », elle utilise les arbres « à questions », illustrés dans la figure 3 : les feuilles sont des entiers naturels, et les nœuds sont des diviseurs. Nous proposons ici plusieurs variantes. Pour un exercice simple (en haut à gauche, en haut à droite), l'arbre contient les questions (nœuds internes), et les élèves doivent renseigner les feuilles. Pour un exercice différent (en bas à gauche), les feuilles sont renseignées, et l'élève doit renseigner les nœuds internes avec des questions discriminantes. Un arbre vide à remplir avec des feuilles et des nœuds (en bas à droite) est également proposé, laissant plus de place à la démarche d'investigation.

Cette activité exerce l'arithmétique en premier lieu. On s'aperçoit que la correction des arbres en classe ou le remplissage en petits groupes est une occasion assez riche de travailler la compétence « communiquer ». Enfin, la compétence « raisonner » est fortement mobilisée.

En effet, il ne s'agit pas seulement de tester une divisibilité, mais de comparer, regrouper, et répartir, à l'aide de raisonnements déductifs. Les élèves élaborent des raisonnements en se basant sur les caractéristiques des différents nombres. Cela leur permet de répartir les nombres proposés dans le premier arbre, de trouver les nombres manquants dans le deuxième arbre et de placer les questions dans le troisième arbre. Voici des exemples de stratégies utilisées : 45 est divisible par 5 et 9 mais pas par 6 ; 2 est le seul nombre premier proposé ; un nombre premier a un seul diviseur différent de 1 ; le seul nombre premier divisible par 7 est 7 ; il y a quatre nombres pairs proposés.

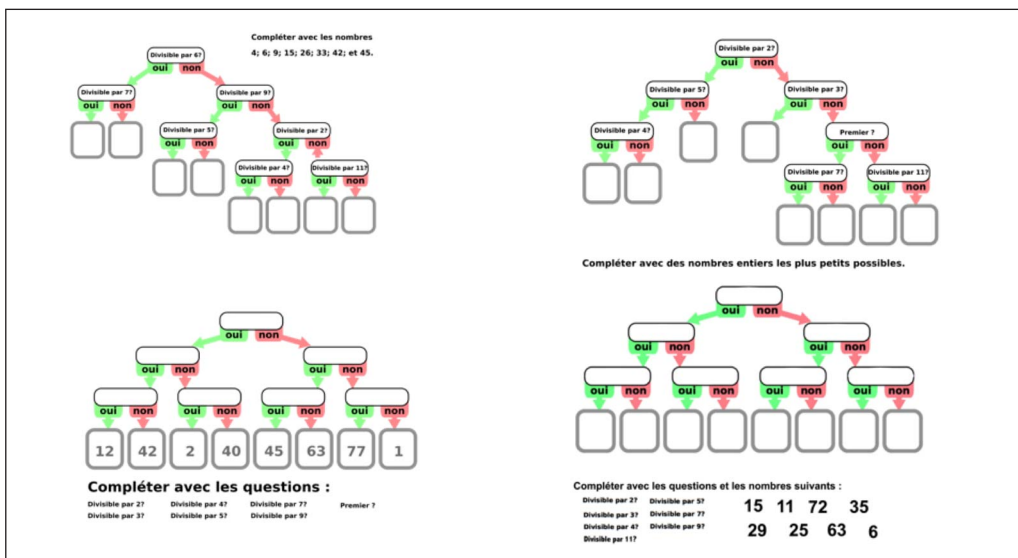


Figure 3. Arbres de décision arithmétiques. Les « questions » dans les nœuds sont des diviseurs. Les entiers naturels dans les feuilles doivent correspondre aux réponses menant de la racine à la feuille.

Cette activité est assez ludique et emporte l’adhésion de beaucoup d’élèves tout en s’inscrivant parfaitement dans le programme de cycle 4. On peut différencier le travail demandé grâce aux différents arbres proposés.

La variante avec l’arbre vide (figure 3, en bas à droite) s’apparente plus à une tâche avec prise d’initiative. En effet, cela demande plus de réflexion et implique des stratégies plus élaborées qui consistent par exemple à lister les diviseurs ou à décomposer en produits de facteurs premiers les entiers proposés. Pour limiter les combinaisons possibles, il est possible de jouer sur les choix proposés : ici, la racine de l’arbre est forcément « divisible par 3 ? » car c’est le seul diviseur ayant exactement 4 multiples dans la liste. On peut de plus se questionner sur l’unicité de la solution, ce qui n’est pas forcément habituel pour les collégiens.

LE CODAGE BINAIRE DES ENTIERS, ALGORITHME DU RENDU DE MONNAIE

Dans cette quatrième activité, nous abordons la question du codage binaire des nombres entiers. Ils sont usuellement représentés dans le système décimal, nous cherchons à les convertir dans le système binaire. Un algorithme simple à expliquer est l’algorithme glouton du rendu de monnaie. Imaginons que nous vivons au pays des puissances de 2, nous avons des billets de valeur 1€, 2€, 4€, 8€, etc. Le but est de payer une somme (disons $S = 5€$) avec le moins de billets

MULTIMEDIA

possibles. Pour cela, il suffit de prendre le billet le plus grand dont la valeur est inférieure ou égale à la somme, et de recommencer avec le reste. Par exemple, pour payer 5€ : utiliser un billet de 4€, reste 1€ ; un billet de 2€ serait trop grand ; utiliser un billet de 1€ ; fin.

Cet algorithme peut aussi être représenté par un arbre binaire. La première question est “Est-ce que l’on a besoin d’un billet de 4€ ?”, la deuxième est “Est-ce que l’on a besoin d’un billet de 2€ ?” (voir figure ci-dessous à gauche). On voit alors que pour atteindre 1, il faut répondre oui-non-oui, le codage de 5 en binaire est donc 101. Nous avons mis en œuvre cette activité avec des classes de collège. Chaque élève dispose d’un exemplaire de chaque billet jusqu’à 15€. Les élèves doivent “rendre la monnaie” pour des sommes entre 0 et 31 annoncées par l’animateur. Après quelques tours d’essai, de nombreux élèves élaborent par eux-même l’algorithme glouton ci-dessus. Il est alors partagé avec toute la classe, et le lien avec l’arbre binaire peut être montré.

Enfin pour les élèves qui auraient travaillé sur le codage binaire grâce à l’activité “Tour de magie binaire” produite par l’Irem de Clermont Ferrand (More et al., 2018), il est possible de faire un pont entre ces deux activités. Le figure 4 (à droite) montre que l’on peut présenter le déroulement du “tour de magie” comme un arbre (binaire !) de décision équilibré.

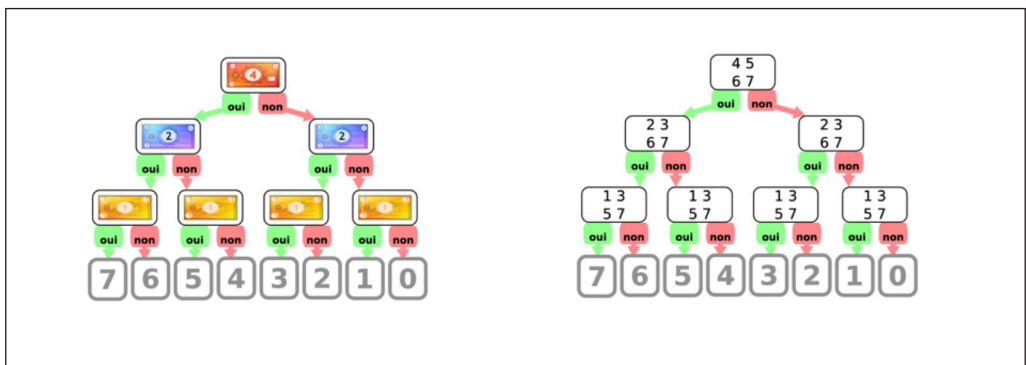


Figure 4. A gauche : rendu de monnaie au pays des puissances de 2. A droite : tour de magie binaire, représenté par un arbre binaire de décision.

CONCLUSION

Nous avons présenté quatre activités afférentes aux arbres binaires, un concept classique d’informatique. Ces activités sont particulièrement bien adaptées aux élèves de cycle 4, quoiqu’elles fonctionnent avant et après. Ces activités permettent d’aborder différents concepts d’informatique, et d’exercer des compétences mathématiques. Les versions présentées ici ont été testées et ajustées en classe. Elles utilisent des arbres dont la structure est donnée, les élèves devant renseigner les feuilles, les nœuds internes, ou simplement les parcourir. Une extension prometteuse, que nous avons commencé à expérimenter, consiste à faire construire l’arbre lui-même : à partir d’un ensemble

de personnages, d'un message à coder, ou d'un ensemble d'entiers. Cela permet de mobiliser des compétences de raisonnement plus poussées, et de comprendre plus en profondeur le fonctionnement des arbres binaires.

BIBLIOGRAPHIE

- Duflot-Kremer, M. et Vincent, J.-M. (2018) Les marmottes au sommeil léger. Dans *Tangente éducation*, 42-43, informatique débranchée.
- More, M., Drot-Delange, B., Fleury, S. et Lafourcade, P. (2018) Un tour de magie pour introduire la représentation binaire des nombres. *Bulletin Vert de l'APMEP*, 525-526.
- Sauvage, B. (2019). La complexité algorithmique : on débranche et on dénombre. *Colloque CORFEM 2019*.