Réseaux, information et communications (INFO-F303)
Partie "Théorie de l'Information et du codage"
Guide d'étude

Christophe Petit

October 7, 2022

La plupart des questions ci-dessous peuvent être répondues après lecture des sections correspondantes dans les transparents du cours et le syllabus du cours. Un certain nombre d'entre elles vous demandent quelques calculs, vous invitent à la réflexion, ou encore à la recherche d'information au-delà de ces sources.

Certains exercices ci-dessus (marqués $*$) proviennent de quizzes développés par Jean Cardinal pour ce cours. Je remercie chaleureusement Jean pour la mise à disposition de ce matériel.

0 Introduction

1. Quel est le but de ce cours ? Quelles sont les principales questions auxquelles on tentera de répondre ?
2. Quelle est la place de ce cours dans le diplôme de bachelier ? En quoi est-il important ?
3. Choisissez une de vos technologies multimedia ou de communication préférée, et renseignez-vous sur le ou les types de codage qu'elle utilise. Ce type de codage sera-t-il abordé dans le cadre de ce cours ?
4. Quels sont vos objectifs d'apprentissage personnels pour ce cours ?
5. Parcourez la table des matière des ouvrages de références, et les commentaires sur ceux-ci disponibles sur internet. Comment allez-vous les utiliser au mieux dans votre étude ?
6. Qu'entend-on par "flipped classroom"? comment allez-vous faire pour tirer le maximum de cette approche d'apprentissage ?
7. Essayer de programmer quelques lignes en Sage et/ou Magma. Explorez les fonctions disponibles pour les calculs de probabilités, les corps finis et les codes correcteurs d'erreurs.

8. Y a-t-il un sujet lié au cours sur lequel vous aimeriez en apprendre davantage ?

1 Notion de code

1. Révisez les principales notations et termes utilisés pour ce cours
2. Quel est l'intérêt d'un code univoque ?
3. Quels sont les avantages des codes en blocs et des codes à longeur variable ?
4. Qu'est-ce qu'un code sans préfixe ? Quel est l'intérêt d'un tel code ?
5. (*) Parmi les ensembles de mots de code binaires ci-dessous, lesquels sont sans préfixe ?
   (a) $\{0, 00, 01, 11\}$
   (b) $\{00, 01, 100, 101, 11\}$
   (c) $\{1, 00, 010, 011\}$
   (d) $\{01, 1, 00, 010, 011\}$
   (e) $\{000, 111, 001, 1110\}$
   (f) $\{0, 00, 000\}$
6. Qu'est-ce qu'un arbre de code ? Peut-on associer un tel arbre à tout code ?
7. Construisez l'arbre de code associé aux deux codes ci-dessous, ou expliquez pourquoi ceci n'est pas possible
   (a) $C = \{1, 01, 0001, 0010, 0011\}$
   (b) $C = \{1, 01, 0001, 010, 0011\}$
8. Sans regarder la preuve, essayez de vous convaincre que l'inégalité de Kraft et le théorème de McMillian sont vrais. (Vérifiez-les sur quelques exemples et tentez de trouver des contre-exemples.)
9. Etudiez maintenant les preuves de ces deux résultats. Quelles sont les étapes clés de ces preuves ? En quoi vous paraissent-elles naturelles ou astucieuses ?
10. Donnez un exemple de code pour lequel l'inégalité de Kraft est en fait une égalité, et un exemple pour lequel l'inégalité est stricte.
11. (*) Pour quelles longueurs de mots existe-t-il un code binaire sans préfixe ?
    (a) 1,2,3,3
    (b) 1,2,2,3
    (c) 1,2,3,3,4
    (d) 1,2,5,5
12. Ecrivez un petit programme dans le langage de votre choix réalisant le codage et le décodage pour un code sans préfixe de longueur variable de votre choix. Le programme doit pouvoir manipuler des chaînes de longueur supérieure à une.

2 Source aléatoire et codes efficaces

1. Que signifie l'hypothèse markovienne ? Quels propriétés du chapitre seront invalidées si cette hypothèse n'est plus vérifiée ?
2. Supposez une source aléatoire telle que pour tout i les symboles $2i$ et $2i+1$ émis sont corrélés entre eux mais indépendant des autres. Suggérez une approche pour se replacer dans l'hypothèse markovienne.
3. Donnez un ensemble de probabilités pour lequel le code de Shannon est optimal, et un ensemble de probabilités pour lequel il est sous-optimal.
4. Comparez les codes de Shannon et de Shannon-Fano-Elias.
5. Etudiez la preuve que le code de Shannon est sans préfixe. Quelles sont les étapes clés de cette preuve ? En quoi vous parait-elle naturelle ou astucieuse ? Pourrait-on reproduire cette preuve pour d'autres choix de code ?
6. Dans l'algorithme de Shannon-Fano, est-il important que les probabilités soient ordonnées ? Testez votre réponse sur un exemple.
7. Donnez un ensemble de probabilités produisant au moins trois codes de Shannon-Fano distincts.
8. Faites des recherches sur les formats JPEG et MP3 pour comprendre comment ils utilisent le codage de Huffman.
9. Reconstruisez le tableau II.2 en p16 du syllabus, pour les probabilités données en exemple puis un autre ensemble de probabilités de votre choix.
10. Etudiez la preuve que le code de Huffman est efficace. Quelles sont les étapes clés de cette preuve ? En quoi vous parait-elle naturelle ou astucieuse ?
11. En supposant l'existence de plusieurs code de Huffman pour un ensemble de probabilités donné, y en a-t-il un dont l'utilisation vous paraît a priori meilleure en pratique ?
12. (*) Quel sont les codes de Shannon et Huffman binaires associés à la distribution discrète $(1/2, 1/8, 1/8, 1/8, 1/8)$?
13. (*) Quel sont les codes de Shannon et Huffman binaires associés à la distribution discrète $(1/4, 1/8, 1/8, 1/8, 1/8, 1/8, 1/8)$?
14. Ecrivez des petits programmes dans le langage de votre choix construisant les codes de Shanon, Shannon-Fano, Shannon-Fano-Elias et Huffman pour un ensemble de probabilités données.

3 Entropie et codage efficace

1. Qu'est-ce qu'une mesure au sens mathématique du terme ?
2. Le syllabus de cours défini la quantité d'information de façon axiomatique (à partir de ses propriétés essentielles). Quel est l'intérêt d'une telle approche ? Les axiomes choisis vous paraissent-ils pertinents ? Proposez une définition alternative, non axiomatique.

3. Calculer la quantité d'information portée par chaque lettre de l'alphabet en français et dans une autre langue de votre choix (utilisez pour cela des statistiques comme `https://www.sttmedia.com/characterfrequency-french`). Comparez les moyennes de ces quantités dans les deux langues.
4. Donnez la définition de l'entropie, et redémontrez ses propriétés principales. Quelles sont les étapes clés des preuves que vous n'avez pas pu retrouver seuls ? En quoi vous paraissent-elles naturelles ou astucieuses ?
5. Donnez une définition axiomatique de l'entropie.
6. En quoi la quantité d'information est-elle liée à la "quantité de désordre" apparaissant en thermodynamique ? Les axiomes utilisés pour définir la quantité d'information (et l'entropie) vous paraissent-ils pertinents pour mesurer la "quantité de désordre"? Comparez les situations extrêmes d'entropie maximale et minimale dans les deux contextes.
7. Quelles sont les unités de mesure usuelles de l'entropie et comment sont-elles liées entre elles?
8. Spécialisez la fonction d'entropie au cas d'une source binaire ($r = 2$). Faites de même pour une source ternaire ($r = 3$).
9. (*) On considère une variable aléatoire X prenant les valeurs $\{2, 4, 6, 8, 10\}$ avec les probabilités respectives $1/2, 1/8, 1/8, 1/8, 1/8$. Quelle est son entropie (en bits) ?
10. (*) On se donne deux dés classiques à 6 faces, que l'on jette simultanément. On note X et Y les résultats respectifs des deux dés. Quelle est l'entropie de la variable aléatoire $Z = X + Y$?
11. (*) On se donne une variable aléatoire X qui prend ses valeurs dans les entiers positifs. On se donne la variable $Y = X^3$, le cube de X. Quelle est la relation entre les deux entropies $H(X)$ et $H(Y)$?
12. Comparez l'entropie d'une distribution de probabilités et la longueur moyenne du code de Shannon associé.
13. Exprimez le sens de l'inégalité de Gibbs en vos propres termes.
14. Qu'est-ce que la divergence de Kullback-Leibler et en quoi celle-ci est-elle liée à l'inégalité de Gibbs ?
15. Etudiez la preuve de l'inégalité de Gibbs. Quelles sont les étapes clés de cette preuve ? En quoi vous parait-elle naturelle ou astucieuse ?
16. Qu'entend-on par "extension d'une source" ? Quelle est l'entropie présente dans cette extension ?
17. Redémontrez la valeur de l'entropie d'une extension de source. Quelles sont les étapes clés de la preuve que vous n'avez pas pu retrouver seuls ? En quoi vous parait-elle naturelle ou astucieuse ?
18. Intuitivement, qu'est-ce que l'entropie peut nous apprendre sur "le code idéal" pour une source donnée ?
19. Etudiez la preuve du Théorème III.2.2. Quelles sont les étapes clés de cette preuve ? En quoi vous parait-elle naturelle ou astucieuse ?
20. Que dit le premier théorème de Shannon ?

21. Donnez des exemples de codes pour lequels les bornes inférieures et supérieures du premier théorème de Shannon sont atteintes.
22. Comment les longueurs moyennes des codes de Shannon, Shannon-Fano et Huffman sont-ils liés à l'entropie de la source?
23. (*) Quelles sont les affirmations correctes ?
    (a) Lorsque pour une variable aléatoire discrète X, toutes les probabilités ont la forme $(1/2)^c$ pour un entier c, alors la longueur moyenne du code de Huffman binaire pour X est exactement égale à $H(X)$.
    (b) Lorsque pour une variable aléatoire discrète X, toutes les probabilités ont la forme $(1/2)^c$ pour un entier c, alors la longueur moyenne du code de Shannon binaire pour X est exactement égale à $H(X)$.
    (c) La longueur moyenne d'un code binaire pour une variable aléatoire discrète X est égale à $H(X)$ si et seulement si toutes les probabilités des valeurs possibles de X ont la forme $(1/2)^c$ pour un entier c.
    (d) Lorsque pour une variable aléatoire discrète X, toutes les probabilités ont la forme $(1/2)^c$ pour un entier c, alors la longueur moyenne minimum d'un code de Huffman binaire pour X est exactement égale à $H(X) + 1$.
24. Etudiez la preuve du premier théorème de Shannon. Quelles sont les étapes clés de cette preuve ? En quoi vous parait-elle naturelle ou astucieuse ?
25. Recalculez le tableau III.1 (possiblement à l'aide d'un petit programme) pour $\mathbb{P}(0) = 0.6$.

4 Compression sans perte

1. Quelle est la différence entre compression avec ou sans perte ? Dans quels contextes la compression avec perte est-elle acceptable ?
2. Définissez taux de compression. Ce taux peut-il être négatif ?
3. Donnez les intuitions principales derrière les algorithmes de compression par dictionnaire et Huffman adaptatif.
4. Comparez les algorithmes de Lempel-Ziv et de Lempel-Ziv-Welch.
5. Quel est l'avantage d'une structure de données en arbre préfixe pour l'algorithme de Lempel-Ziv-Welch?
6. Quelles sont les garanties de qualité offertes par les algorithmes de compression par dictionnaire et Huffman adaptatif ?
7. Comparez le coût de la mise à jour du code de Huffman dans l'algorithme de Vitter avec le coût d'un recalcul complet de l'arbre.
8. Supposez un message alternant des passages dans des langues différentes. Comparez les performances a priori des différents algorithmes de compression discutés dans les notes.
9. (*) Donner le résultat de l'encodage par l'algorithme de Lempel-Ziv-Welch pour les chaînes suivantes définie sur l'alphabet $\{a, b\}$:

(a) *aaabbbaaabbb*
(b) *abbaabbbaaab*
(c) *abaaabaaabaa*

10. Ecrivez des petits programmes (dans le langage de votre choix) implémentant les algorithmes de Lempel-Ziv, Lempel-Ziv-Welch et Vitter.
11. L'invariant de Vitter est-il une condition nécessaire et suffisante pour obtenir un code optimal de hauteur minimale?
12. Etudiez la preuve de l'invariant de Gallager.
13. Quelles sont les garanties d'optimalité fournies par l'algorithme de Lempel-Ziv ? Etudiez la preuve de celles-ci dans l'article original.

5 Canal bruité

1. Décrivez un canal de communication bruité, comme modélisé par Shannon. Quels sont les paramètres d'un tel canal?
2. Rappelez la loi de Bayes.
3. Définissez entrope résiduelle, conditionnelle et croisée, et information mutuelle. Donnez une signification intuitive de ces concepts.
4. Quels sont les liens entre l'information mutuelle et l'information transmise à travers un canal bruité?
5. Reproduisez les calculs de l'exemple en section V.2.5 des notes de cours.
6. Qu'est-ce que la capacité d'un canal? En quoi ce concept diffère-t-il de l'information mutuelle?
7. Qu'est-ce qu'un canal symétrique (généralisé)? Quelles sont les propriétés d'un tel canal qui ne sont pas satisfaites en général ?
8. (*) On considère deux variables aléatoires X et Y prenant respectivement leurs valeurs dans les ensembles $\{A, B\}$ et $\{0, 1\}$. Les probabilités conjointes sont les suivantes :

	$Y=0$	$Y=1$
$X=A$	$1/2$	$1/4$
$X=B$	0	$1/4$

   Quelle est la valeur de l'entropie conditionnelle $H(X|Y)$ (en bits, à quatre chiffres près après la virgule) ?
9. (*) On considère deux variables aléatoires X et Y prenant respectivement leurs valeurs dans les ensembles $\{A, B, C\}$ et $\{0, 1\}$. Les probabilités conjointes sont les suivantes :

	$Y=0$	$Y=1$
$X=A$	$1/3$	0
$X=B$	0	$1/3$
$X=C$	$1/3$	0

   Quelle est la valeur de l'information mutuelle $I(X, Y)$ (en bits) ?

10. (*) On considère un canal symétrique à k symboles, pour lequel un symbole en entrée a une probabilité $1-p$ d'être transmis correctement, et une probabilité $p/(k-1)$ d'être transformé en chacun des $(k-1)$ autres symboles. Quelle est la capacité – en **bits** – d'un tel canal, pour $k=3$ et $p=1/2$ (avec 4 chiffres après la virgule) ?

6 Codes correcteurs d'erreurs

1. Qu'est-ce qu'un code correcteur d'erreur ? Quels sont les paramètres principaux d'un tel code ?
2. Choisissez un standard de stockage ou de communication de votre choix, et identifiez le ou les codes correcteurs d'erreur qu'il utilise. Quels sont les paramètres employés ?
3. Qu'est-ce qu'un code à répétition ? Quelles sont les propriétés de tels code ? Y a-t-il des utilisations pratiques d'un tel code ?
4. Qu'est-ce qu'un code à somme de contrôle ? Quelles sont les propriétés de tels code ? Y a-t-il des utilisations pratiques d'un tel code ?
5. (*) On considère un code à répétition de la forme $K : C \to C^3 : c \mapsto ccc$. Quelle est la probabilité d'erreur de décodage pour ce code sur un canal symétrique de paramètre $p = 1/4$?
6. Etudiez la preuve de la propriété VI.1.1. Quelles sont les étapes clés de cette preuve ? En quoi vous parait-elle naturelle ou astucieuse ?
7. Qu'est-ce que les capacités de détection et de correction d'un code? Comment ces quantités sont-elles liées ? Dans quels cas sont-elles égales ?
8. Qu'est-ce que la distance minimale d'un code, et comment les capacités de détection et correction sont-elles liées à cette distance ?
9. Vérifiez que la distance de Hamming est bien une distance au sens mathématique du terme. En quoi cette distance diffère-t-elle d'une distance Euclidienne ?
10. Qu'est-ce qu'une boule de Hamming et en quoi ce concept est-il similaire à une boule au sens usuel ?
11. Donnez la taille d'une boule de Hamming, et une approximation asymptotique de celle-ci.
12. Qu'est-ce que le rayon d'empilement et le rayon de recouvrement d'un code? Comment ces quantités sont-elles liées?
13. Qu'est-ce qu'un code parfait? Donnez un exemple.
14. Qu'est-ce qu'un code maximal? Donnez un exemple.
15. La théorie des codes correcteurs d'erreur considère généralement le cas d'une source émettant des messages équiprobables, alors que beaucoup de messages "naturels", comme de la voix ou des textes en français codés en ASCII, auront des distributions différentes. Expliquez comment se ramener au cas des sources uniformes pour ce genre de messages.

16. Qu'est-ce que le débit d'un code? En quoi cette quantité est-elle importante et que signifie-t-elle?
17. Que dit le deuxième théorème de Shannon ? Expliquez son sens en vos propres termes. Ce résultat vous paraît-il naturel ? Que signifie la borne $\mathcal{C}_r$ et pourquoi, intuitivement, intervient-elle ici? Cette borne est-elle optimale?
18. Etudiez la preuve du deuxième théorème de Shannon. Quelles sont les étapes clés de cette preuve ? En quoi vous parait-elle naturelle ou astucieuse ?
19. Citez et comparez les bornes de Singleton et de Hamming.
20. Etudiez les preuves de ces bornes. Quelles sont les étapes clés de ces preuves ? En quoi vous paraissent-elles naturelles ou astucieuses ?
21. En quoi la borne de Gilbert-Varshamov diffère-t-elle des bornes de Singleton et Hamming?
22. Etudiez la preuve de cette borne. La borne peut-elle être atteinte?
23. (*) Considérez le code binaire suivant:

$$\{00000000, 11110000, 00001111, 11111111\}.$$

Calculez son débit, sa distance minimale, sa capacité de détection d'erreurs, et sa capacité de correction d'erreurs.

24. (*) On considère un code binaire de débit 5/8 qui code tous les mots binaires de longueur 10. Quelle est la borne de Singleton sur la distance minimale d d'un tel code ?
25. (*) On considère un code binaire K de longueur 10 et de distance minimale 5. Quelle est la borne de Hamming sur la taille $|K|$ d'un tel code ?
26. Implémenter des fonctions d'encodage, d'ajout de bruit et de décodage pour un code de Hamming, dans le langage de votre choix.

7 Codes linéaires

1. Révisez tous les concepts d'algèbre discutés en section VII.1 du syllabus. Utilisez les calculateurs en ligne de Magma ou Sage pour définir et manipuler des groupes, espaces vectoriels, corps finis, anneaux entiers, modules, anneaux polynomiaux. En particulier, construisez un corps fini de cardinalité première, un polynôme irréductible sur ce corps, et construisez un corps d'extension à partir de ce polynôme. Vérifiez la caractéristique du corps. Définissez, additionnez et multipliez des éléments de ce corps finis, et comparez les résultats avec vos calculs à la main. Calculez un générateur et vérifiez son ordre multiplicatif.
2. Qu'est-ce qu'un code linéaire ? Quel est l'avantage d'un tel code par rapport à un code général ?
3. Qu'est-ce qu'une matrice génératrice d'un code linéaire ? Expliquez pourquoi on peut permuter les colonnes d'une telle matrice sans modifier le code. Quel est l'effet d'une permutation des lignes de G sur le code?

4. Spécialisez les bornes de Hamming et de Singleton au cas des codes linéaires. Ces bornes peuvent-elles êtres atteintes?
5. Donnez la définition d'une matrice de contrôle. Une telle matrice est-elle unique ?
6. Quels sont les liens entre les matrices génératrice et de contrôle d'un code linéaire?
7. Ecrivez un petit programme réalisant les opérations suivantes à partir de paramètres n (la longueur) et k (la dimension):
   (a) Choix d'une matrice génératrice pour un code (n, k).
   (b) Calcul de la forme canonique pour ce code.
   (c) Calcul de la distance minimale du code.
   (d) Calcul d'une matrice de contrôle pour ce code.
   (e) Calcul et stockage de la table de décodage.
   (f) Encodage, ajout d'erreur et décodage.
8. Qu'est-ce que le syndrôme d'un mot ? En quoi celui-ci fournit-il un algorithme de détection d'erreur? Quelle est l'efficacité de cet algorithme ?
9. Ecrivez un petit programme pour réaliser le codage et décodage d'un code de Hamming binaire $(15, 11)$.

8 Quelques familles de codes linéaires

1. Supposez un canal symmétrique binaire, avec une probabilité d'erreur $p = 0.3$. On veut envoyer des mots de $n = 32$ bits sur ce canal avec un débit maximal. Suggérez un code linéaire pour ce problème; évaluez son débit et son taux d'erreur résiduel. Le code choisi est-il proche de la capacité optimale ?
2. Comparez code parfait, code MDS et code approchant la capacité de Shannon.
3. Choisissez un outil ou un protocole de stockage ou transmission de l'information et recherchez quels sont le ou les codes correcteurs d'erreurs utilisés pour cette application. Quelles sont les raisons ayant justifié ce choix de code ?
4. (*) On considère un code binaire polynomial $(7, 5)$ avec le polynôme générateur $G(Z) = Z^2 + 1$ sur $\mathbb{F}_2[Z]$. Au mot d'entrée $x = [10110] \in \mathbb{F}_2^5$, on associe le polynôme $X(Z) = Z^4 + Z^2 + Z$. Quel est le mot du code correspondant à x ?
5. (*) On considère le code de Hamming binaire $(7, 4)$ avec la matrice généra-

trice canonique suivante :

$$\left[\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \hline 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{array}\right]$$

On reçoit le mot

$$[0111010]^T = \left[\begin{array}{c} 0 \\ 1 \\ 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{array}\right]$$

Un bit au plus a été transmis de manière erronée. Quel est le mot envoyé ?

6. Calculez le polynôme générateur pour un code BCH avec $n = 7$ et $\delta = 3$
7. Calculez le polynôme générateur pour un code BCH avec $n = 8$ et $\delta = 3$
8. Etudiez la preuve de la borne sur la distance minimale d'un code BCH. Quelles sont les étapes clés de cette preuve ? En quoi vous paraisse-elle naturelle ou astucieuse ?
9. Etudiez la preuve de la borne sur la distance minimale d'un code de Reed-Muller. Quelles sont les étapes clés de cette preuve ? En quoi vous paraisse-elle naturelle ou astucieuse ?
10. Dans un code BCH, les paramètres principaux du code dépendent-ils du choix de générateur α?
11. Ecrivez un programme en Sage pour construire le polynôme générateur d'un code BCH de paramètres donnés, et réaliser l'encodage et le décodage pour ce code.