Banque MP inter ENS - Session 2016

Épreuve écrite d'informatique-mathématiques – Filière MP, spécialité Info

Écoles concernées : ENS de Cachan, ENS de Lyon, ENS de Paris, ENS de Rennes

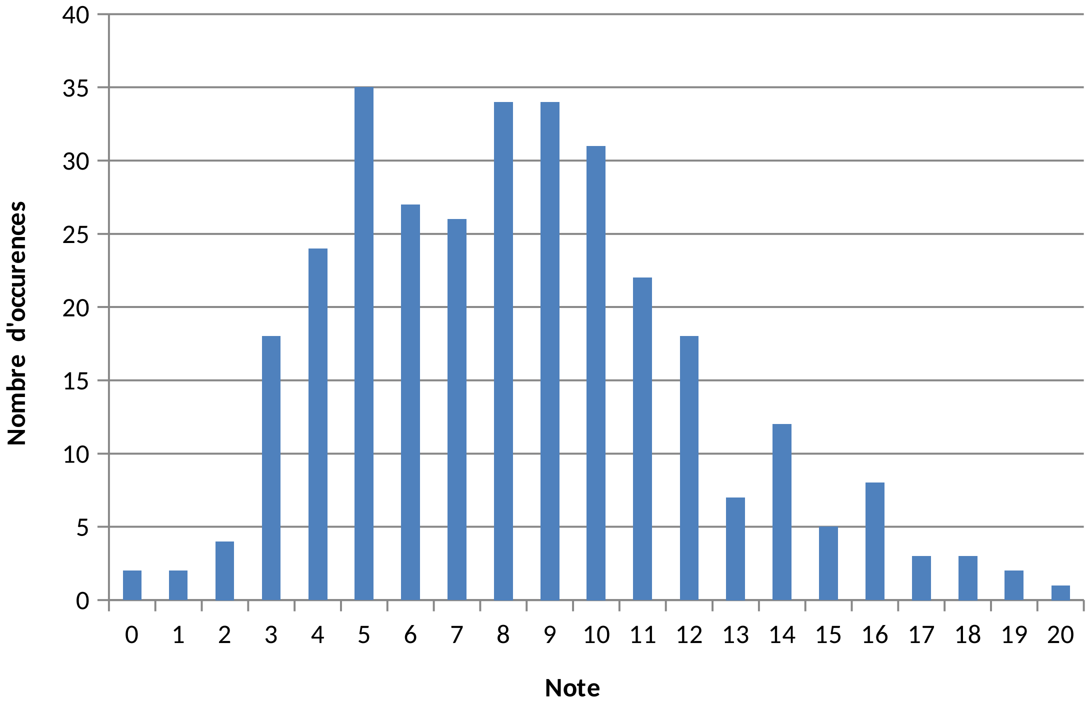
Coefficients (en % du total concours) :

ENS de Cachan :	13,16 %
ENS de Lyon :	12,70 % (pour l'option M et l'option P)
ENS de Paris :	13,33 %
ENS de Rennes :	08,57 %

Membres du jury :

Le nombre de candidats présents à l'épreuve est de 318.

Voici la répartition des notes. Elles allaient de 0,59 à 20.



La moyenne se situe à 8.84 et l'écart type à 3.77. Comparée aux années passées, la moyenne était un peu plus haute, et l'écart-type un peu plus bas. Globalement, il y a moins de très mauvaises notes.

Remarques globales :

Le barème a bien évidemment tenu compte de la longueur du sujet. Les correcteurs tiennent à souligner les points suivants, que nous répétons chaque année:

- Le manque de rigueur coûte généralement cher. Il est peut-être utile de souligner que les correcteurs accepteront une rédaction rapide d'autant plus facilement que le/la candidat(e) aura traité les questions similaires précédentes rigoureusement et identifie clairement les points auxquels prêter attention.
- Pour obtenir une bonne note il n'est pas nécessaire de faire beaucoup de questions, mais il faut affronter la difficulté. Le barème et la notation ont récompensé les efforts pour aborder les questions délicates. Traiter correctement la partie I (qui n'était pas la plus simple) apporte déjà la moyenne. Inversement, les copies qui ne traitaient aucune question difficile ont rarement reçu une note satisfaisante.
- Les tentatives d'escroquerie sont à proscrire. Il semble nécessaire de rappeler que l'honnêteté intellectuelle est une des qualités évaluées, étant la base de la démarche scientifique. Les escamotages de difficultés au fil de raisonnements inutilement tortueux, les affirmations sans explication de reformulations des questions posées, etc., sont donc très mal vues et rendent les correcteurs particulièrement stricts. Les bonnes idées et arguments partiels sont d'autant plus récompensés que la copie reconnaît les lacunes du raisonnement proposé.
- Il est important de bien lire les questions. Par exemple, pour les questions algorithmiques, il était écrit que « Pour les algorithmes, on demande une explication claire de leur fonctionnement, permettant de se convaincre que les programmes que vous écrirez sont corrects. ». Cela a été assez peu fait, la plupart des copies se bornant à écrire le code. Il suffit souvent d'une phrase utilisant le bon argument pour expliquer et justifier le code. Commenter le code aide mais ne remplace pas une phrase d'explications. Enfin, certains candidats apportent des réponses algorithmiques inventives, qui nécessiteraient quelques explications pour être mieux prises en compte.
- Une fraction non négligeable des copies se contente d'une explication haut-niveau des algorithmes sans produire de (pseudo)-code. Bien évidemment, cela escamote une partie des problèmes et est pénalisé assez lourdement. C'est également problématique pour dériver la complexité en $O(|E|^k)$ des algorithmes, pour lesquels les structures de données utilisées et le code importent. La complexité a d'ailleurs été assez mal traitée dans l'ensemble.

Remarques spécifiques :

Partie I

Question 1 : très bien traitée dans l'ensemble. Attention à penser à justifier que MX(i) est dans [0,1] pour tout i.

Question 2a) : une première question qui faisait la sélection des mauvaises copies. La notion de composante fortement connexe n'est pas toujours bien comprise. En particulier, il n'y a pas toujours d'arc direct entre 2 éléments d'une composante fortement connexe ! De plus, de trop nombreux candidats pensent qu'une composante connexe réduite à un élément et sans boucle sur cet élément « n'est pas une composante fortement connexe » (cela étant surtout un problème de terminologie, ça a été moins pénalisé).

Question 2b) : La question est bien traitée dans l'ensemble, le problème de mauvaise connaissance des composantes fortement connexes n'étant pas apparent ici.

Question 2c) : La question a été très mal traitée en général (90% des copies l'abordent, 75% des copies ont 0, une fraction non négligeable des copies a néanmoins le maximum de point à cette question à grand coefficient). On peut de toute manière avoir une excellente note finale sans traiter cette question. Les questions 2a) et 2b) servaient à guider les candidats en utilisant les bonnes techniques pour aider à résoudre cette question ouverte 2c). De très nombreux candidats pensent que I est caractérisé par exactement les cas qui ne satisfont pas 2a) et 2b). La grande majorité force une preuve incorrecte de ce fait. Certains trouvent des contre-exemples à cette conjecture, ce qui a été bien valorisé. La plupart de ceux-ci conclut par un résultat correct (attention néanmoins au « sens » des flèches du graphe).

Question 3)a). Les remarques sur la manière d'écrire et de décrire les algorithmes ont été faites plus haut. La plupart des copies prend des points non négligeables à cette question à fort coefficient en produisant un algorithme correct. Avoir l'intégralité des points nécessitait en revanche de produire un algorithme optimal, c'est-à-dire en complexité $O(|E|^2)$, ce qui a été peu fait. La solution la plus simple était de précalculer en temps quadratique à partir des (L_v) une matrice de booléens qui pour tout (i,j) dit si j est atteignable par i ou non. Sinon, il était nécessaire de parcourir des listes encore et encore, ce qui aboutissait à une complexité non optimale en $O(|E|^3)$.

Question 3b). Les remarques générales sur l'algorithmique et la complexité s'appliquent également à cette question. Même si la question 2c) avait été mal faite, la question de la complexité restait intéressante et importante dans cette question 3b).

Question 3c) De nombreux candidats ont évité cette question, qui était néanmoins plutôt facile et bien faite dans l'ensemble de ceux l'ayant abordée. Une inopportune erreur peut être visible dans l'algorithme : ligne 7 il manque un b(w) :=1. Les réponses argumentées que l'algorithme ne terminait pas (ce qui est exact vu l'erreur) ou que la complexité était linéaire (ce qui est correct en ignorant l'erreur) étaient toutes les deux acceptées. Corriger l'algorithme avec cette ligne manquante était de plus récompensé (plusieurs cas) par des points de bonus. Une réponse argumentée d'une complexité en temps quadratique, non optimale, était relativement pénalisée.

Partie II

La question 4), qui permettait d'aborder la notion de langage de chaine de Markov, nouvelle pour les candidats, a été bien traitée dans l'ensemble. Ces questions ouvertes laissent toutefois la place à quelques confusions.

Les questions 5)a)-c), très faciles et très bien faites, étaient faites pour aiguiller les candidats sur le fait que le spectre d'une matrice avait une grande influence sur son langage. Cela permettait également de rappeler que les valeurs propres d'une chaine de Markov pouvaient être négatives et même complexes, et que les vecteurs propres n'étaient en général pas des distributions. La question 5d) était un mal nécessaire au sujet – la donnée du résultat permettait aux candidats de ne pas perdre trop de temps à traiter cette question.

La question 6a) a été assez discriminante, une bonne moitié des candidats n'ayant pas de points ici, qu'il ne l'ait pas abordée ou mal traitée. Il fallait penser et justifier que le spectre d'une matrice et de sa transposée était le même, ce à quoi l'indication aidait.

La question 6b) a été mieux traitée que 6a), des preuves alternatives simples étant possibles.

Partie III

Question 7) et 8) la question 7) était facile, et la question 8) était une simple application de la question 7, qui pouvait être évitée en acceptant le résultat. On peut remarquer que les matrices de changement de base compliquaient beaucoup les calculs, qui était beaucoup plus simple à partir du moment où on exprimait X dans la base des vecteurs propres.

Question 9) : a été relativement bien faite, par ceux (une petite moitié des candidats) l'ayant traitée. La distinction entre ces copies s'est faite plutôt sur la qualité des justifications. C'était une question relativement importante, traitable indépendamment du reste du sujet, en utilisant le résultat donné dans la question 8.

Partie IV

Le but est d'établir dans le cas n=3 que les matrices dont les valeurs propres sont réelles positives ont des langages réguliers, c'est-à-dire la question 13 qui est très ouverte et à fort coefficient. Les questions 10 à 12 présentent le cas particulier d'une telle chaine de Markov et en fait calculer le langage, qui doit aider les candidats à généraliser ceci dans la question 13.

Les questions 10 à 12 ne présentaient pas de difficultés particulières. Les calculs se simplifiaient bien, mais de très nombreuses erreurs de calculs ont eu lieu. Elles n'ont eu que des conséquences limitées sur les questions suivantes – la notation s'intéressant plus à la méthodologie et l'argumentation employée plutôt qu'aux calculs eux-mêmes. Un raisonnement correct à partir d'un résultat incorrect obtenu précédemment a été largement valorisé. Comme à la question 8, l'utilisation de matrice de changement de bases complexifiait le calcul.

Faute de temps, peu de candidats se sont intéressés aux questions passé la 10a).Quelques réponses partielles intéressantes ont été données pour la question 13 (qui pouvait être faite sans traiter les questions précédentes). On s'attendait à une distinction de cas suivant les valeurs propres dominantes pour X_0 et X_1 (suivant laquelle est plus grande que l'autre, ou si elles sont identiques).

*   *   *