

## Composition d'Informatique (4 heures), Filière MP

Rapport de MM. Frédéric MAGNIEZ et Benjamin WERNER, correcteurs.

## L'épreuve

Il s'agissait d'un problème de représentation de dictionnaires. Deux représentations étaient étudiées reposant respectivement sur une structure tabulée (Partie II) et sur une structure binaire (Partie III). Des fonctions de manipulation de ces dictionnaires étaient à programmer relatif à l'impression, l'appartenance, et l'insertion. La partie IV demandait d'écrire des fonctions de conversion permettant de passer d'un dictionnaire à l'autre. Ces dictionnaires étaient enfin utilisés pour résoudre des problèmes de mot le plus long (Partie V) et d'anagramme (Partie VI).

## Remarques générales

La plupart des fonctions demandées faisaient environ 10 lignes et 20 dans le pire cas. Les candidats doivent essayer de rendre des programmes simples. En général, lorsqu'un candidat part sur une solution compliquée, il s'agit d'une mauvaise réponse. Les candidats devraient avoir cette remarque en tête lorsqu'ils écrivent leurs programmes.

Parmi les candidats français, 12% ont choisi d'écrire leurs programmes en langage Pascal. La moyenne des copies Pascal est inférieure à celle des copies CAML, mais de très bonnes copies ont tout de même été écrites en Pascal. Il est donc comme d'habitude difficile d'attribuer cette différence aux langages.

Pour obtenir la note maximale il était nécessaire de traiter le problème en entier.

Les notes des candidats français se répartissent selon le tableau suivant :

$0 \leq N < 4$	116	15,8%
$4 \leq N < 8$	270	36,8%
$8 \leq N < 12$	197	26,8%
$12 \leq N < 16$	116	15,8%
$16 \leq N \leq 20$	35	4,8%
Total	734	100 %
Nombre de copies : 734		
Note moyenne : 8,24		
Écart-type : 4,18		

## Commentaire détaillé

Pour chaque groupement de questions sont indiqués entre crochets le pourcentage de candidats ayant obtenu au moins la moitié des points correspondants, et le pourcentage de candidats ayant obtenu la note 0 (que la question ait été traitée ou non).

### Partie I

**Question 1 [95% - 3%].** Peu de problèmes ici, si ce n'est quelques confusions sur le sens dans lequel imprimer le mot, et le traitement du retour à la ligne. Quelques réponses bizarres utilisant des tableaux.

**Question 2 [94% - 5%].** On a légèrement sanctionné l'utilisation de `rev` ainsi que les fonctions quadratiques au lieu de linéaires.

### Partie II

Certains candidats ont considéré que les mots passés en arguments se terminaient par le caractère `$`. Cela simplifiait très légèrement certaines réponses ; pour obtenir la note maximale, il fallait préciser ce point. Sinon, il était alors difficile de juger si la question du caractère de terminaison, qui était une des clés de cette partie, était comprise.

**Question 3 [86% - 2%].** Il s'agissait essentiellement de montrer que l'on avait compris la gestion des fins de mots.

**Question 4 [47% - 44%].** A cause, justement, de la difficulté de la gestion du caractère de terminaison, de nombreux programmes affichaient en fait tous les préfixes, ce qui explique les statistiques des notes sur cette question.

Une autre difficulté était d'imprimer les mots au bon moment tout en gardant le préfixe courant pour la suite du parcours d'arbre. Certains candidats ont géré ce problème de manière inutilement compliquée, aboutissant le plus souvent à un programme erroné. La solution était en fait courte une fois la structure comprise.

**Question 5 [85% - 7%].** Même problème qu'à la question précédente, mais de nombreux candidats n'ont pas fait la même erreur ici.

**Question 6 [72% - 15%].** La difficulté était de rajouter le suffixe manquant au dictionnaire lorsque c'était nécessaire. Cette modification a parfois été délicate à programmer.

### Partie III

Bien que les questions reprenaient celles de la partie II, on a pu observer une certaine décorrélation et ce ne sont pas toujours les mêmes candidats qui ont bien traité l'une et l'autre. Sans doute à cause de la structure moins visuelle du dictionnaire, la partie III a été un peu moins bien traitée. En particulier, le fait que la  $k$ -ième lettre d'un mot n'apparaisse pas à la profondeur  $k$  dans le dictionnaire rendait plus difficile de tester la présence d'un préfixe dans le dictionnaire.

Certains candidats ont considéré que les lettres étaient rangées par ordre alphabétique. Dans ce cas, pour obtenir la note maximale, il fallait que ce soit précisé et fait de manière cohérente entre les différentes questions.

**Question 7 [82% - 7%].** Pour justifier la disparition de la racine il fallait deux arguments : elle ne porte pas d'information et n'a qu'un seul fils. Souvent un de ces arguments manquait.

**Question 8 [44% - 48%].** Même difficulté qu'à la question 4. Étonnamment ce ne sont pas toujours les mêmes candidats qui ont échoué à ces deux questions.

**Question 9 [67% - 27%].** Plus d'échecs qu'à la question 5 correspondante, sans doute à cause de la structure du dictionnaire plus complexe et moins visuelle. Il fallait parcourir tous les fils droits d'un nœud avant de s'apercevoir qu'un mot n'était pas dans le dictionnaire.

**Question 10 [50% - 37%].** Souvent des préfixes déjà existants ont été rajoutés (cf ci-dessus), ce qui était impossible avec la structure tabulée (question 6). Cela explique l'écart de statistiques.

### Partie IV

**Question 11 [16% - 57%].** Cette question, assez longue, n'a pas été correctement traitée dans son ensemble. Beaucoup de manque de rigueur. Cette question de complexité nécessitait de la méthode. Il fallait identifier les dictionnaires correspondant au pire et au meilleur cas, puis en tirer les formules de l'encadrement et enfin justifier que tous les

autres cas étaient intermédiaires. Les deux premiers points suffisaient à obtenir une note quasi maximale.

**Question 12 [37% - 55%].** Cette question et la suivante ont été sélectives, mais plutôt bien traitées par les candidats y ayant répondu. Certaines des solutions très élégantes ont montré qu'il était possible de répondre simplement.

Les solutions consistant à utiliser en cascade les fonctions imprimer et ajout ont été peu valorisées.

**Question 13 [28% - 66%].** Mêmes remarques que ci-dessus, la difficulté étant sensiblement supérieure.

## Partie V

**Question 14 [18% - 78%].** Certaines solutions élégantes, beaucoup d'autres très embrouillées et compliquées car la récursion était plus délicate à réaliser.

Les solutions ayant recours à l'énumération brutale (à l'aide de imprimer) ont été très peu récompensées.

## Partie VI

**Question 15 [7% - 90%].** La solution était très proche de la précédente, mais finalement très peu de candidats l'ont réalisé. Il fallait cette fois relancer le programme précédent jusqu'à épuisement complet des lettres du mot passé en argument ; l'impression ne devant se faire que si exactement toutes les lettres avaient été consommées par les mots sélectionnés du dictionnaire.