

REMARQUES D'ORDRE GÉNÉRAL

1. Erreurs courantes

Les correcteurs ont noté que les candidats ne semblent pas toujours lire le sujet avec suffisamment d'attention. Plus précisément, certains candidats utilisent les tables de vérité au lieu des équations et réciproquement ou exploitent des références, des répétitions, etc. ; malgré des indications explicites très claires sur les techniques de résolution attendues en logique, question par question et de contraintes très précises sur les constructions autorisées dans les langages CaML et PASCAL. Sur ce second point, notons que ces contraintes, issues du programme de l'option d'Informatique en classe préparatoire, sont les mêmes depuis la création de l'épreuve d'option d'Informatique au sein du Concours Commun Polytechnique. Il semble que ce ne soit pas seulement le choix de la technique que les candidats maîtrisent le mieux indépendamment des indications données par le sujet, car certains d'entre eux ont utilisé les tables de vérité en lieu et place des équations puis les équations en lieu et place des tables de vérité.

Quelques rares candidats donnent des résultats sans aucun détail. Parfois, il s'agit de démonstration dont le résultat est énoncé dans le sujet et pour lequel seule une paraphrase est fournie au lieu d'une preuve. Mais, il s'agit également parfois de résultats de calcul que les candidats ont effectué au brouillon et n'ont pas recopié. Il n'est pas forcément nécessaire de détailler dans tous les cas toutes les étapes de calcul, surtout si celles-ci sont évidentes. Par contre, il est toujours nécessaire de donner les éléments essentiels qui permettent de juger de la qualité de la réponse.

Quelques candidats ne numérotent pas les réponses qu'ils donnent dans leurs copies. Lorsque celles-ci ne sont pas bien rédigées, les correcteurs risquent de ne pas interpréter correctement à quelle question s'adresse une réponse et de ne pas appliquer le barème de manière satisfaisante.

Certains candidats ne numérotent pas leurs copies ou se trompent dans cette numérotation. Ceci est particulièrement gênant pour des copies "à trous", c'est-à-dire dans lesquelles les candidats laissent de la place pour traiter plus tard certaines questions. Le risque devient non négligeable de ne pas voir une question figurant sur une page mal numérotée sans rien indiquer avant ou après.

Enfin, de rares candidats raturent ou barrent une partie de leurs réponses, qui sont parfois partiellement correctes et ne fournissent pas de réponses de substitution. Les éléments corrects qui ont été barrés ou raturés sont considérés comme ne figurant pas sur la copie. Il est donc préférable de n'enlever quelque chose d'une copie que pour le remplacer par autre chose. Sinon, il vaut mieux laisser une réponse partiellement correcte plutôt que de tout éliminer.

2. Remarques sur le texte, sa compréhension

Le sujet de l'épreuve d'option d'Informatique pour la session 2012 du Concours Commun Polytechnique est composé de trois parties indépendantes qui recouvrent les aspects principaux du programme : logique des propositions, automates et langages, algorithmiques et programmation. L'approche proposée est proche des sujets des années précédentes : un respect strict et une couverture large des différents aspects du programme à travers des questions de difficulté croissante.

La volonté, introduite lors de la session précédente, de décorréler les différentes questions, pour éviter la propagation des erreurs, est maintenue. La première partie propose aux candidats des exercices de modélisation en calcul des propositions, puis de résolution en utilisant les tables de vérité, puis les formules de De Morgan. La deuxième partie consiste à étudier un opérateur de composition d'automates : l'union disjointe ; d'abord sur un exemple ; puis par la preuve à base de récurrence sur la longueur des mots (ou induction sur la structure des mots). La troisième partie est composée de deux exercices et d'un problème. Le premier exercice traite de la programmation pour une structure de liste d'entiers ; le second exercice concerne l'interprétation et la preuve de correction d'un algorithme sur les arbres binaires de recherche : l'élimination d'un élément dans un arbre ; le problème vise la manipulation d'un B-Arbre représenté par un nœud contenant le sous-arbre gauche et une liste de paires : étiquette, sous-arbre droit.

Les candidats ne semblent pas avoir rencontré de difficultés particulières de compréhension du texte.

La partie Algorithmique et Programmation en langage PASCAL comportait une erreur qui ne figurait pas dans la partie quasi-identique en langage CaML. L'arbre binaire "exemple" défini avant la question III.4 n'était pas un arbre binaire de recherche. Les candidats qui ont abordé cette question ont soit lu la partie CaML et repris l'exemple correct; soit proposé un autre exemple. Les deux approches ont été notées de la même manière.

3. Réactions (prévues ou non)

Les candidats ont globalement bien traité l'exercice de Logique et de Calcul des propositions (de l'ordre de 70 %) ainsi que l'exercice d'Automates et langages (de l'ordre de 50 %). Ces parties du programme sont bien maîtrisées mais les résultats devraient être encore supérieurs car il s'agit vraiment d'applications directes du contenu des enseignements.

Les candidats ont ensuite traité convenablement l'exercice sur les séquences croissantes d'entiers (de l'ordre de 50 %).

Puis, seules quelques questions de lecture et écriture de programmes ont été abordées par approximativement 30 % de candidats (Q III.4, III.6, III.9, III.10 et III.11).

Le reste du sujet est très peu abordé (moins de 10 % des candidats). En particulier, les candidats ont très peu traité les questions de preuve de correction des programmes.

Certains candidats s'obstinent à proposer des réponses récursives terminales plus complexes sans que ceci soit justifié. Les solutions proposées sont alors plus complexes et plus souvent erronées.

4. Bilan

Comme les années précédentes, le sujet se révèle trop long sans que les copies ne permettent de déterminer si les candidates n'ont pas traité certaines questions par manque de temps, ou par incapacité face aux difficultés posées par ces questions.

Les correcteurs ont indiqué une évolution positive dans la qualité de la rédaction des copies (89 % de satisfaction des correcteurs).

Le nombre de candidats qui utilisent le langage PASCAL est toujours en baisse constante et semble se situer maintenant entre 5 et 10 %. Les correcteurs insistent encore sur le fait que la note moyenne des candidats utilisant le langage PASCAL est inférieur à celle de ceux qui exploitent le langage CaML. Cette différence est toujours de l'ordre d'un à deux points. Une modification du programme

devrait être effectuée pour imposer un langage unique dans l'apprentissage de la programmation et éliminer le risque que le choix du langage par les enseignants désavantage les candidats qui le savaient.

Les correcteurs ont remarqué que certains candidats lisent encore mal le sujet et n'utilisent pas les techniques imposées en logique ou en programmation. Ils estiment que 10 % des candidats n'utilisent pas les techniques imposées. Plus précisément, l'exploitation des constructions interdites de CaML (références, séquences, affectations, répétitions) et PASCAL (répétitions) semble être en baisse significative par rapport aux années précédentes (moins de 10 % des candidats).

Les correcteurs signalent également un progrès dans la rédaction des preuves concernant les automates et langages. Plus précisément, l'utilisation de "..." pour indiquer qu'il faut appliquer une étape élémentaire sur un caractère pour tous les caractères du mot, en lieu et place d'une récurrence, est en nette régression. Par contre, les questions ne sont pas plus traitées que les années précédentes.

Le contenu du sujet ne semble pas avoir posé de problèmes techniques particuliers aux candidats. Comme les sujets précédents, celui-ci s'est révélé long mais permet ainsi de bien distinguer les candidats. La moyenne de l'épreuve est de 9,74 et l'écart type de 3,23. Ce résultat est satisfaisant et de meilleure qualité que les années précédentes. La première partie et le premier exercice de la troisième partie sont en général bien traités. La seconde partie et le second exercice de la troisième partie sont abordés partiellement par la plupart des candidats et se sont révélés très discriminants. Le problème de la troisième partie est très peu abordé.

RAPPORT DÉTAILLÉ

(Erreurs rencontrées, rédaction de certains candidats, questions sans problème)

Le % indiqué pour chaque question correspond à la note moyenne obtenue par les candidats par rapport au barème attribué à la question.

Logique et calcul des propositions

- Modélisation sous la forme d'une formule du calcul des propositions (questions I.1 (72 %), I.2 (75 %), I.3 (79 %), I.4 (74 %), I.6 (71 %) et I.7 (72 %)).

La grande majorité des candidats maîtrisent bien la modélisation (entre 70 et 80 % de succès).

Quelques uns ont indiqué que trois déclarations étaient nécessaires, souvent sans se justifier; parfois en ne prenant pas en compte le fait que la valeur de vérité de chaque formule est connue.

Certains ont considéré qu'un « changeant » ne changeait qu'une seule fois.

D'autres ont mal interprété : "Le cercle n'est visible que si le losange est visible." en proposant la formule " $L \Rightarrow C$ " au lieu de " $C \Rightarrow L$ ". Cette erreur est habituelle lorsque cette forme de phrase doit être modélisée et permet de distinguer les candidats qui font une analyse plus profonde de la signification de la phrase de ceux qui se contentent d'identifier lexicalement une conditionnelle.

Enfin, certains interprètent le "soit" comme un "ou" classique au lieu du "ou exclusif".

Quelques candidats ont considéré que les mêmes orateurs intervenaient dans les deux parties contrairement aux indications du sujet et ont donc changé les noms attribués aux variables.

- Résolution avec les tables de vérité et interprétation (question I.5 (61 %)).

Certains candidats utilisent un ordre apparemment aléatoire pour les lignes des tables, oublient certains lignes et en dupliquent d'autres.

Quelques uns font des erreurs au niveau des implications.

- Résolution avec les formules de De Morgan et interprétation (question I.8 (38 %)).
Les résultats proposés pour ce type de question sont significativement plus faibles que les années précédentes car un nombre important de candidats ont fait de petites erreurs de modélisation qui entraînent des erreurs importantes lors de la résolution. Ceci est moins sensible pour les tables de vérité pour lesquelles les étapes intermédiaires apparaissent plus systématiquement, ce qui limite la portée des erreurs de modélisation.

Automates et langages

- Construction d'automates et d'expressions régulières (questions II.1 (66 %), II.2 (62 %), II.3 (33 %)).
Les premières expressions sont en général correctes, la dernière ne décrit qu'une partie du langage. Il faut préciser que celui-ci était très complexe. Il faut quand même noter qu'un nombre important de candidat ne sont pas capables de donner une expression correcte pour le cas le plus simple. L'automate construit est souvent partiel. Certains états ou transitions font défaut. En particulier, quelques candidats se sont arrêtés dès le premier état terminal obtenu au lieu de poursuivre la construction de l'automate.
- Preuve par récurrence/induction et bilan (questions II.4 (56 %), II.5 (42 %), II.6 (39 %) et II.7 (46 %)).
Les candidats qui abordent ces questions les traitent relativement correctement. Certains oublient les cas d'arrêt. D'autres parodient la récurrence en utilisant "...". Mais, globalement, les réponses proposés sont plus satisfaisantes que les années précédentes. Plus précisément, le nombre de réponses est plus faible, mais celles-ci sont plus souvent correctes. Un nombre non négligeable de candidats dont la preuve est correcte, oublie de citer un raisonnement par récurrence ou induction. Certains candidats donnent les éléments nécessaires pour les preuves mais oublient de conclure.

Algorithmique et programmation en CaML/PASCAL

- **Question préliminaire : Séquence croissante d'entiers**
 - programmes (questions III.1 (88 %) et III.3 (48 %))
Les candidats qui traitent ces questions le font en général correctement mais pas de la manière la plus efficace. Par exemple, de nombreuses réponses exploitent un fonction "rev" au lieu de calculer celle-ci à la volée.
Certains candidats essaient d'exploiter en CaML un filtrage sur la concaténation de listes "@".
 - calcul de complexité (question III.2 (36 %))
Les candidats ne maîtrisent absolument pas la rédaction formelle du calcul de la complexité d'un algorithme. Ils donnent ici le résultat directement dans la plupart des cas sans expliciter son origine.
- **Question préliminaire : Arbre binaire de recherche**
 - évaluation symbolique (question III.4 (49 %))
Les candidats qui abordent cette question donnent en général une réponse correcte. Le faible score correspond à une minorité importante qui n'a pas abordé la question.

- preuve de correction (questions III.5 (5 %))
Très peu de candidats ont traité cette question. La forme de la plupart des réponses est satisfaisante. Par contre, celles-ci sont en général incomplètes.
- preuve de terminaison (question III.6 (28 %))
Les arguments proposés par les candidats sont en général satisfaisants, mais ceux-ci oublient toujours certains cas, car ils n'accordent pas assez d'importance à la rigueur de la forme de leur réponse.
- calcul de complexité (question III.7 (9 %) et question III.8 (10 %))
Les exemples fournis sont rarement complètement satisfaisants. Les complexités proposées sont rarement correctes. Et, même dans ce cas, leur calcul n'est pas formalisé.

- Problème : B-Arbre

- programmes (questions III.9 (33 %), III.10 (32 %), III.11 (22 %), III.12 (4 %) et III.16 (1 %))
De nombreux candidats qui traitent ces questions n'exploitent pas l'aspect ordonné de la structure et parcourent celle-ci entièrement, avec une complexité linéaire au lieu de logarithmique. Étonnamment, certains correspondent aux candidats qui se sont obstinés à mauvais escient à faire des fonctions récursives terminales. Ceci laisse penser que cette mauvaise habitude leur a été inculquée. La constante "ordre" n'est pas exploitée par tous les candidats qui lui substitue sa valeur.
- évaluation symbolique (question III.13 (3 %))
Les rares réponses fournies sont en général correctes.
- preuve de terminaison (question III.14 (1 %))
Cette question a été traitée par un très petit nombre de candidats et toujours de manière partielle et informelle.
- preuve de correction (questions III.15 (1 %))
Cette question a été traitée par un très petit nombre de candidats et toujours de manière partielle et informelle.

CONCLUSION

Les parties traitées par la majorité des candidats sont simples et souvent proches de questions de cours. Celles-ci semblent occuper la majorité du temps des candidats, y compris les plus performants.