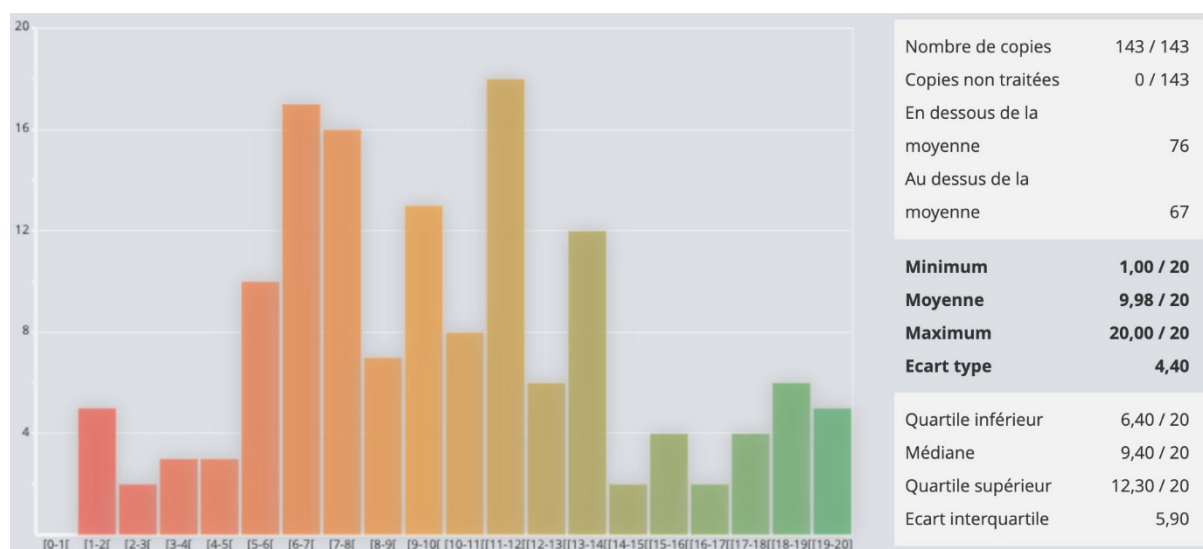




RAPPORT DE L'ÉPREUVE ÉCRITE D'ALGORITHMIQUE ET INFORMATIQUE



L'épreuve portait sur la notion de *Distance d'édition* (ou *Distance de Levenshtein*), et son calcul algorithmique.

La distance d'édition entre deux chaînes de caractères est définie comme le nombre minimal de transformations permettant de changer l'une des deux chaînes de caractère en l'autre ; les transformations pouvant être de l'un des 3 types :

- Suppression d'un caractère,
- Insertion d'un caractère,
- Modification d'un caractère.

Cette notion a de nombreuses applications, en phylogénétique, en linguistique, ou plus trivialement pour la conception de correcteurs orthographiques. Le sujet appliquait notamment la méthode pour concevoir un correcteur orthographique.

Après avoir remarqué l'impossibilité d'une résolution exacte en des temps raisonnables par des algorithmes dits de force brute, ou naïfs, le sujet construit une méthode efficace utilisant une approche par *programmation dynamique*. Il en déduit la réalisation d'un petit correcteur orthographique, puis dans une dernière partie, obtient outre la distance d'édition, la liste des transformations permettant de passer en un temps minimal d'une chaîne à l'autre. Cette dernière partie était estimée un peu hors-barème, destinée aux seuls candidats de niveau exceptionnel, s'il en est.

Dans l'esprit, le sujet ressemblait à des sujets déjà posés pour cette épreuve dans les années précédentes (bien qu'il soit, évidemment, entièrement original), car utilisant ce même paradigme de *programmation dynamique*. Les candidats qui se seraient entraînés sur des sujets des années précédentes auraient eu de bonnes chances de travailler des sujets assez semblables.

Le sujet était long ; il était clairement très difficile à traiter dans toute son intégralité durant les 45 minutes imparties. Certains candidats ont pu traiter toutes les questions, mais visiblement pour les dernières, dans la précipitation. C'était attendu, et le barème brut sur 28 estimait que 8 points seraient hors barème.

Sur 28, la moyenne se situait entre 6 et 7. Après harmonisation elle a été ramenée proche de 10, avec un écart-type de 4,40 et des notes s'étalant de 1 à 20.

Le déroulé de l'épreuve semble satisfaisant. Malgré sa longueur, le sujet était bien adapté à l'épreuve, et a permis de classer les candidats sur leur niveau informatique. Les candidats sont parvenus dans l'ensemble et en un temps très court (45') à s'accaparer une problématique complexe et à implémenter une résolution délicate.

Les candidats de cette année ont montré un niveau assez semblable aux années précédentes, même si aucun candidat n'est parvenu à composer parfaitement le sujet --ce que l'on peut interpréter par sa longueur. On peut regretter que certains candidats faibles (1 tiers) semblent répondre un peu au hasard ou 'au feeling' aux questions ; ce qui s'avère assez contre-productif, et ce dont les préparateurs devraient dissuader les candidats.