

Projet

Mise en page de texte

1 Objectifs du projet

Le but du projet est de mettre en oeuvre les techniques d'algorithmique vues en cours et de les comparer.

2 Modalités du projet

Le travail est à réaliser en binôme.

Le langage de programmation est libre, à choisir parmi les langages suivants : C, C++, Java, Python. Le programme doit pouvoir s'installer et s'exécuter sur les machines du département, sans nécessiter l'installation de bibliothèques ou packages supplémentaires. Pour tout autre langage de programmation, consultez votre chargé de TD.

Vous devez rendre votre code et un rapport sur DiDel. Le rapport contiendra un résumé de ce que vous avez implémenté, un guide d'installation (comment compiler, comment lancer l'exécution), un guide de fonctionnement (comment fonctionne le programme). Vous devez également décrire les algorithmes, et analyser leur complexité théorique. Prenez bien soin de détailler et justifier les récurrences que vous utilisez.

Une partie importante du projet consiste à évaluer les performances de vos algorithmes. Vous devez détailler dans le rapport les points suivants. Pour chaque test, donnez l'objectif du test, la méthodologie utilisée (comment avez-vous réalisé les tests ? Quelles étaient les données ?), les résultats, sous forme d'exemples d'exécution ou de courbe, selon le cas. Dans tous les cas, donnez vos conclusions. Celles-ci doivent se rapporter à l'objectif du test.

Une partie de la note portera sur l'originalité de votre travail.

Le tout est à rendre avant le 29 novembre 2013 à 18h.

3 Description du problème : La mise en page et l'impression équilibrée d'un texte

On dispose d'un fichier texte sans passage à la ligne (il s'agit d'un seul paragraphe) que l'on veut imprimer sur une feuille de taille donnée. Chaque caractère imprimé occupe la même largeur unitaire, y compris les espaces, comme sur une machine à écrire mécanique.

Le problème est de découper le texte en lignes, pour distribuer les espaces qui restent à la fin de chaque ligne (sauf la dernière) de sorte à obtenir la mise en page la plus agréable possible du point de vue visuel (on parle d'*impression équilibrée*). On considère que les espaces qui restent à la fin de la dernière ligne sont "inévitables", en conséquence ils seront ignorés pour l'évaluation du déséquilibre de la mise en page.

Par exemple, voici deux découpages possibles pour le texte “la programmation dynamique est une stratégie algorithmique” sur des lignes de longueur 17.

-----		-----
la programmation		la programmation
dynamique est		dynamique est une
une stratégie		stratégie
algorithmique		algorithmique

Formellement, le texte est une suite de n mots de longueurs l_1, l_2, \dots, l_n , mesurées en nombre de caractères. Chaque ligne de la feuille contient M caractères, blancs compris. Si une ligne contient les mots i à j (inclus), où $i \leq j$, et qu'on laisse exactement un espace entre deux mots, le nombre de caractères blancs à la fin de la ligne est

$$M - j + i - \sum_{k=i}^j l_k$$

Une approche possible serait de prendre comme mesure du déséquilibre la somme des espaces blancs apparaissant en fin de chaque ligne et essayer de minimiser cette quantité. Implémentez cette approche avec toutes stratégies algorithmiques vues en cours qui sont susceptibles d'être appliquées à ce problème. Un algorithme de type diviser-pour-régner ? Programmation dynamique avec une approche “top-down” ? “Bottom-up” ? Un algorithme glouton ?

Pourquoi cette approche ne peut pas distinguer entre les déséquilibres des différents découpages et garantir ainsi la construction du découpage qui donne l'impression la plus équilibrée possible ?

Proposez d'autres approches pour afficher le texte de manière équilibrée, et les implémenter avec des stratégies algorithmiques différentes. Vous devez évaluer chacune des implémentations, en termes de temps d'exécution, espace mémoire nécessaire, ainsi que l'évaluation subjective de la qualité visuelle du résultat.