

This assignment will be closed on May 16, 2023 (23:59:59).

On May 02, 2023 (22:20:24), we had submissions for the following questions: 1 — [Details](#) ➔
(/agns/INF371/TD03/2022/uploads/:my/).

Le 36^e chapitre de «Madame Bovary»

Jean-Marc Notin, Luc Maranget

- Description de l'algorithme
 - Construction de la table associative
 - Génération de texte
 - Structures de données
- Les listes chaînées
 - Problème 1: premières fonctions sur les maillons
 - Problème 2 : maillons vs liste chaînée
 - Problème 3 : manipulation des préfixes
- Problème 4 : Tables de hachage
 - Définition de la table associative
- Problème 5 : Madame Bovary
 - Construction de la table associative
 - Génération de texte
 - Le 36^e chapitre de Mme Bovary.

Avant de commencer, une remarque sur les tests automatiques : le TD est découpé en 5 problèmes de plusieurs questions. Il y a un formulaire par problème, mais il y a plusieurs tests par problème. Chaque test est lié à une méthode (ou un ensemble de méthodes) : si la méthode n'existe pas, le test est ignoré — ceci est indiqué par une pastille jaune dans les résultats. Cela vous permet de soumettre vos solutions à un problème alors que vous n'avez pas encore répondu à toutes les questions.

Les problèmes sont dépendants (E.g. le problème 2 utilise la classe Node du problème 1). Cependant, il est inutile de soumettre les fichiers des problèmes précédents lors du téléversement, ils seront automatiquement copiés pour vous. Cependant, si vous le faites, ces derniers seront alors redéposés dans leur formulaire respectif et les tests relancés. (Ainsi, à la limite, si vous soumettez tous vos fichiers dans le dernier formulaire, tous vos fichiers seront déposés dans leur problème respectif et tous les tests seront repassés)

Le but de ce TD est d'écrire un programme Java qui va générer un pseudo texte à partir des 35 chapitres du livre de Gustave Flaubert. La méthode employée pour la génération du texte est basée sur les chaînes de Markov.

Le texte généré aura les propriétés suivantes : Étant donné un entier $n \geq 1$

- Toutes les sous-suites de $n + 1$ mots du texte généré sont des sous-suites de $n + 1$ mots du texte de *Madame Bovary*.
- Les n premiers mots du texte sont les n premiers mots d'un chapitre de *Madame Bovary*.
- Les n derniers mots du texte sont aussi les n derniers mots d'un chapitre de *Madame Bovary*.

L'entier n est donné en argument de la ligne de commande (ou, dans Eclipse, dans le menu *Run > Run Configurations...* > *Java Application*, onglet *Arguments*). Ainsi, la commande

```
% java Bovary 3
```

produit un texte plutôt étrange, par exemple :

Un soir que la fenêtre était ouverte, et que, assise au bord, elle venait de s’échapper dans le jardin. Emma, tout exprès, avait retiré la clef de sol, et s’occupait volontiers de littérature après son dîner, quand il ne jouait pas aux cartes. M. Homais le considérait pour son instruction ; madame Homais l’affectionnait pour sa complaisance, car souvent il accompagnait au jardin les petits Homais, marmots toujours barbouillés, fort mal élevés et quelque peu lymphatiques, comme leur mère. Ils avaient pour les soigner, outre la bonne, Justin, l’élève en pharmacie, un arrière-cousin de M. Homais et peut-être les pesanteurs du déjeuner, restait indécis et comme sous la fascination du pharmacien qui répétait : [...]

Le texte ci-dessus est de toute évidence dérivé de l’œuvre de Flaubert. « Dérivé » est d’ailleurs le mot juste, on perçoit un certain flottement dans l’écriture.

On notera par exemple, que le chapitre numéro 28 commence par «*Un soir que [Charles l’écoutait,...]*». La phrase «*Emma, tout exprès, avait retiré la clef de sol, [...]*» n’est pas de Flaubert, mais Flaubert a écrit :

- «*Emma, tout exprès, avait retiré la clef de la barrière, que Charles crut perdue.*» Une manœuvre d’Emma pour faciliter ses rencontres avec Rodolphe, chap. 19.
- «*Puis il possédait des talents, il peignait à l’aquarelle, savait lire la clef de sol, et s’occupait volontiers de littérature après son dîner, quand il ne jouait pas aux cartes.*» Une description de Léon, le clerc de notaire, chap. 12.

On remarquera que «*retiré la clef de*» et «*la clef de sol,*» sont deux suites de $n + 1 = 4$ mots qui permettent de relier la clef de la barrière à la clef de sol.

Ce petit exercice pseudo-littéraire sera le prétexte de découvrir en Java quelques unes des structures de données classiques utilisées en informatique.

Description de l’algorithme

L’algorithme emploie une table associative dont les clés sont des séquences de n mots et les valeurs associées des multi-ensembles (ensembles avec répétitions possibles des éléments) de mots. La première phase de l’algorithme construit la table à partir du texte de *Madame Bovary*; la seconde phase utilise la table pour produire un nouveau texte.

Construction de la table associative

Le principe de cette première phase est de construire une table associant à chaque sous-séquence P de n mots du texte, le multi-ensemble des mots qui suivent immédiatement P . L’algorithme doit donc parcourir les 35 chapitres de *Mme Bovary*, lire chaque mot et l’associer à la séquence des n mots précédemment lue. Une telle séquence est appelée *préfixe* et est notée $w_1, w_2, \dots, w_{n-1}, w_n$.

Pour construite la table associative, nous utiliserons l’algorithme suivant :

Pour chaque chapitre :

- Le préfixe P est initialisé à <START>, <START>, ... , <START> .
- Pour chaque mot w du chapitre en cours :
 - on ajoute w aux mots associés à P,
 - le préfixe P devient w_2, w_3, \dots, w_n, w .
- À la fin, on ajoute <END> aux mots associés à P.

Les chaînes de caractères <START> et <END> sont des marqueurs spéciaux servant à représenter respectivement le debut et la fin de chaque chapitre. Ils n’apparaissent explicitement pas dans le texte.

Génération de texte

Une fois la table associative construite, nous utiliserons l’algorithme suivant pour générer un nouveau texte :

- Le préfixe P est initialisé à `<START>`, `<START>`, `...`, `<START>`.
- Répéter :
 - choisir w parmi les mots associés à P , selon une loi uniforme,
 - si w est `<END>`, alors terminer,
 - sinon :
 - on affiche w ,
 - le préfixe P devient w_2, w_3, \dots, w_n, w .

Structures de données

Dans cet algorithme, il est nécessaire de manipuler plusieurs structures de données.

- D’une part, nous devons représenter des multiensembles (<https://fr.wikipedia.org/wiki/Multiensemble>) de mots (c’est-à-dire des ensembles de mots dans lesquels il peut y avoir plusieurs fois le même mot). La taille de ceux-ci ne pouvant être connue à l’avance, il est judicieux de choisir une structure de données facilement extensible. Nous utiliserons donc une structure de liste chaînée (https://fr.wikipedia.org/wiki/Liste_cha%C3%AEn%C3%A9e) pour implémenter un multi-ensemble de mots. *De fait, les premières questions sont classiques et destinées à vous familiariser avec la manière dont les listes chaînées sont implémentées en Java.*
- D’autre part, nous utilisons une table associative qui, à la manière d’un dictionnaire, associe une *valeur* à une *clé* donnée. Pour une telle structure de données, il est essentiel de garantir une recherche rapide de la clé. Nous implémenterons donc la table associative à l’aide d’une table de hachage (https://fr.wikipedia.org/wiki/Table_de_hachage).

Si vous travaillez sous Eclipse, commencez par créer un nouveau projet Java pour le TD.

Les listes chaînées

Une liste chaînée est implémentée par une suite de maillons chaînés les uns aux autres : chaque maillon contient la référence du maillon suivant. Ainsi, avec la référence vers le premier maillon, on peut accéder successivement à tous les maillons de la chaîne. À la fin de la chaîne, on met la référence du maillon suivant à `null`.

Concrètement, les maillons d’une liste chaînée sont déclarés de la manière suivante :

```
class Node {
    String head;
    Node next;

    Node(String head, Node next) {
        this.head = head;
        this.next = next;
    }
}
```

Le champ `head` permet de stocker la valeur associée à un maillon et le champ `next` permet d’accéder au maillon suivant dans la chaîne. La classe `Node` contient également la définition d’un constructeur permettant la création d’un nouveau maillon à partir d’une valeur et d’une référence à un maillon.

Pour créer en Java la chaîne correspondant à la liste `["foo", "bar", "baz"]`, on écrira le code suivant :

```
Node foobar = new Node("foo", new Node("bar", new Node("baz", null)));
```

Rappel : `null` est une valeur spécifique en Java qui représente une référence ne correspondant à aucun objet en mémoire. C’est la valeur par défaut de toute variable de type non primitif.

Problème 1: premières fonctions sur les maillons

Créez un nouveau fichier `Node.java` qui implémente les liste chaînées telles que décrites ci-dessus.

Ce formulaire est pour le problème 1. Vous devez soumettre le fichier `Node.java`.

Last submission: May 02, 2023 (17:31:37)

Choose

Choose one or more files...

Submit

☒ Reuse files from previous submissions ?

Expected files: Node.java

Question 1 : longueur d’une chaîne

Dans la classe `Node`, définissez la fonction récursive (statique) `lengthRec`, qui permet de calculer la longueur d’une chaîne. Le profil de la fonction `lengthRec` sera le suivant :

```
static int lengthRec(Node l) {  
    ...  
}
```

Vous pourrez tester votre fonction `lengthRec` en ajoutant une fonction `main` dans la classe `Node` et affichant la longueur de la chaîne `foobar` donnée en exemple.

Question 2 : longueur d’une chaîne (bis)

Dans les langages impératifs tels que Java, on évite de recourir systématiquement aux fonctions récursives. Comme vu en cours, l’appel de fonction est coûteux en temps et en espace mémoire. La pile ayant une taille limitée, un trop grand nombre d’appels imbriqués se terminent par une erreur de *stack overflow* :

```
Exception in thread "main" java.lang.StackOverflowError
```

Nous allons donc réécrire le calcul de la longueur d’une chaîne en utilisant un parcours itératif :

```
for (Node cur = l; cur != null; cur = cur.next) {  
    ...  
}
```

Dans la classe `Node`, ajoutez une fonction `length` permettant de calculer la longueur d’une chaîne de manière itérative.

Question 3 : affichage d’une chaîne

Toujours dans la classe `Node`, définissez la fonction :

```
static String makeString(Node l) {  
    ...  
}
```

qui renvoie la chaîne désignée par `l`. Sur la chaîne `foobar` définie précédemment, la fonction `makeString` doit renvoyer la chaîne suivante :

```
"[foo, bar, baz]"
```

Question 4 : ajout dans une chaîne

Définissez la fonction `static void addLast(String s, Node l)` qui ajoute `s` à la fin de la chaîne `l`. Pour cette fonction, nous supposons que `l` ne vaut pas `null`.

Testez votre fonction `addLast` en ajoutant la chaîne `"qux"` à la suite de la chaîne `foobar` donnée en exemple.

Vous remarquerez que l’argument `l` de la fonction `addLast` n’est pas modifié directement, mais que le résultat (attendu) de cette fonction, est de modifier, dans la mémoire, le chaînage entre les différents enregistrements de type `Node`.

Question 5 : chaînes triées

Écrivez une fonction `static Node copy(Node the)` qui renvoie une copie d’une chaîne.

Il est possible de comparer deux `String s1` et `s2` pour voir laquelle est avant l'autre dans l'ordre alphabétique. On utilise pour cela la méthode `compareTo`. Plus précisément, `s1.compareTo(s2)` rend :

- une valeur strictement positive si `s2` est avant `s1` dans l'ordre lexicographique,
- une valeur strictement négative si `s1` est avant `s2`,
- la valeur 0 si les deux sont identiques (donc si `s1.equals(s2)` vaut `true`).

Écrivez une fonction `static Node insert(String s, Node l)` qui suppose que la chaîne `l` est triée pour fabriquer la chaîne triée dans laquelle on aura inséré `s` (on a le droit de modifier la chaîne `l`).

Utilisez cette fonction pour écrire une fonction `static Node insertionSort(Node l)` qui fabrique la version triée de `l`. Cette fonction doit être simple, peut avoir une complexité quadratique, et a le droit de modifier la chaîne `l`.

Problème 2 : maillons vs liste chaînée

Telle quelle, la classe `Node` permet bien d'implémenter la structure de liste chaînée. Néanmoins, comme le montre la question précédente, des problèmes se posent pour l'écriture des opérations sur les listes, notamment l'ajout et la suppression des éléments. Par exemple, dans la question précédente, comment implémenter l'ajout d'un élément dans une chaîne vide ?

Pour écrire facilement les opérations sur les listes, une solution est d'envelopper cette structure de maillons dans un nouvel enregistrement `WordList`, défini de la manière suivante :

```
class WordList {
    Node content;

    WordList() {
        content = null;
    }
}
```

Avec cette implémentation, comment est représentée la liste vide ? Comparez cette représentation avec celle de la chaîne de `Node`.

Ce formulaire est pour le problème 2. Vous devez soumettre le fichier `WordList.java`.

You will have access to the files you submitted to Pb.1. (You do not need to resubmit them or to copy their content in the submitted files). If you submit files that belong to these previous questions, they will be propagated up and the automated tests will be triggered.

Choose

Choose one or more files...

Submit

☒ Reuse files from previous submissions ?

Expected files: WordList.java

Question 6

Dans la classe `WordList`, écrivez la définition d'une variable (statique) `foobar`, de type `WordList`, correspondant à la liste `["foo", "bar", "baz"]` (vous pourrez vous aider d'une fonction ou d'un constructeur auxiliaire).

Opérations sur les listes chaînées

Contrairement à la partie précédente, nous implémenterons les opérations sur les listes de mots directement dans la classe `WordList`. Nous abandonnerons donc l'emploi du mot clé `static` et parlerons de *méthode* au lieu de fonction. Une méthode s'utilise sur un objet existant en mémoire et a accès directement aux champs de données.

Dans la mesure du possible, il est souhaitable d'appeler les fonctions de la partie précédente plutôt que de copier-coller leur code.

Question 7

Ajoutez dans la classe `WordList` une méthode `int length()` calculant la longueur de la liste.

Question 8

Vous ajouterez également la méthode `toString()`, qui retourne la chaîne de caractère correspondante. Cette méthode respectera le même format que la fonction `makeString` écrite précédemment. Pour des raisons qu'on verra dans un autre cours, il faut préciser `public` devant la déclaration de cette fonction.

Question 9

Implémentez dans la classe `WordList` les opérations suivantes :

- `void addFirst(String w)`, qui ajoute un mot en tête de la liste;
- `void addLast(String w)`, qui ajoute un mot à la fin de la liste;
- `String removeFirst()`, qui enlève le premier mot de la liste, et le rend en résultat;
- `String removeLast()`, qui enlève le dernier mot de la liste, et le rend en résultat.

Dans le cas où la liste est vide, les fonctions `removeFirst` et `removeLast` retourneront la valeur `null`.

Question 10

Ajoutez à la classe `WordList` les méthodes `void insert(String s)` et `void insertionSort()` correspondant aux fonctions statiques homonymes de la classe `Node`.

Question 11 (cette question est facultative)

Ajoutez à la classe `Node` une fonction `static Node merge(Node l1, Node l2)` qui fusionne deux chaînes triées en une seule chaîne triée.

Utilisez cette fonction pour ajouter à `WordList` une méthode de tri efficace `void mergeSort()`.

Question 12

Ajoutez dans la classe `WordList` un constructeur permettant de créer une liste à partir d'un tableau de `String`. Ce constructeur devra avoir le profil suivant :

```
WordList(String[] t) {  
    ...  
}
```

Ce constructeur devra respecter l'ordre des éléments contenus dans `t`.

Question 13

Toujours dans la classe `WordList`, écrivez une méthode `String[] toArray()` qui permet de construire un tableau de chaînes de caractères à partir d'une liste chaînée.

Problème 3 : manipulation des préfixes

Les clés de notre table associative sont des séquences de n mots. Nous allons donc avoir besoin de créer une classe spécifique `Prefix` permettant de manipuler ces séquences :

```
class Prefix {  
    String[] t;  
  
    final static String start = "<START>", end = "<END>", par = "<PAR>";  
}
```

Ce formulaire est pour le problème 3. Vous devez soumettre le fichier `Prefix.java`.

You will have access to the files you submitted to Pb.1, Pb.2. (You do not need to resubmit them or to copy their content in the submitted files). If you submit files that belong to these previous questions, they will be

propagated up and the automated tests will be triggered.

Choose

Choose one or more files...

Submit

☒ Reuse files from previous submissions ?

Expected files: Prefix.java

Question 14

Ajouter dans la classe `Prefix` un constructeur prenant en argument un entier n et construisant le préfixe `<START>, ..., <START>` de taille n . Ce préfixe servira à amorcer la construction de la table associative et la génération de texte.

Égalités

Considérons le programme suivant :

```
class EqString {
    public static void main(String[] args) {
        String foo1 = new String("foo");
        String foo2 = new String("foo");
        System.out.println(foo1 == foo2);
    }
}
```

Étonnamment (ou pas), le résultat affiché est `false` . En effet, l'opérateur `==` compare les valeurs des variables, ce qui correspond, dans le cas des enregistrements, à comparer les références en mémoire. Or, dans l'exemple `foo1` et `foo2` sont des références à deux enregistrements distincts en mémoire.

Heureusement, la classe `String` fournit une méthode `equals` renvoyant `true` si et seulement si les valeurs des deux enregistrements de type `String` sont égales. Ainsi, l'instruction suivante :

```
System.out.println(foo1.equals(foo2));
```

affiche bien `true` sur la sortie standard.

Question 15

Dans la classe `Prefix` , ajoutez une méthode statique `eq(Prefix p1, Prefix p2)` qui renvoie `true` si et seulement si les préfixes `p1` et `p2` sont identiques (au sens de `equals`).

Question 16

Dans la classe `Prefix` , ajoutez une méthode `Prefix addShift(String w)` qui à partir du préfixe w_1, w_2, \dots, w_n va construire le préfixe w_2, \dots, w_n, w .

Attention, cette méthode ne doit pas modifier l'objet `this` mais construire un nouveau tableau.

Problème 4 : Tables de hachage

Ce formulaire est pour le problème 4. Vous devez soumettre les fichiers `HMap.java` , `Entry.java` & `EntryList.java` .

You will have access to the files you submitted to Pb.1, Pb.2, Pb.3. (You do not need to resubmit them or to copy their content in the submitted files). If you submit files that belong to these previous questions, they will be propagated up and the automated tests will be triggered.

Choose

Choose one or more files...

Submit

☒ Reuse files from previous submissions ?

Une table de hachage est une collection de données où chaque entrée est un couple (*clef*, *valeur*). L'intérêt d'une table de hachage, par rapport à une simple liste, est que si l'on connaît la clef, on peut savoir si la collection contient cette clef et récupérer la valeur associée en temps moyen constant. Par exemple, si l'on déclare un type enregistrement `Eleve` avec comme champs : `nom` , `prénom` , `matricule` , etc. et que l'on range des objets de ce type dans une table de hachage en prenant le matricule comme clef, il sera alors possible d'obtenir la fiche d'un élève à l'aide du matricule. Si l'on ajoute une seconde table pour laquelle la clef est le nom, il sera aussi possible de chercher la fiche d'un élève par son nom.

Une seule valeur est associée à une clef et associer une nouvelle valeur à une clef déjà présente dans la table écrase l'ancienne valeur.

Implémenter une table de hachage est trivial dans le cas où les clefs sont des entiers pris dans un intervalle $[0, n]$ quand n n'est pas trop grand. Dans ce cas la table de hachage est un tableau et la case i contient la valeur associée à la clef i et `null` si aucune valeur ne lui est associée.

Dans le cas général, les tables de hachage se ramènent au cas précédent : on utilise une fonction de hachage qui calcule un entier à partir de la clef. La fonction de hachage doit être telle que si l'on prend deux valeurs égales (au sens sémantique, mais cela peut être deux références différentes), leurs valeurs de hachage doivent aussi être égales. En pratique, pour que la structure soit efficace, on demande beaucoup plus à la fonction de hachage et en particulier, on veut que les valeurs qu'elle donne pour les valeurs rencontrées pendant l'exécution du programme soient les plus éloignées possible les unes des autres.

Une fois cette bonne fonction de hachage h trouvée, on procède de la manière suivante : On commence par créer un tableau à n cases et si l'on veut ajouter l'entrée (*clef*, *valeur*) , on stocke le couple (*clef*, *valeur*) dans la case $h(\text{clef}) \bmod n$ du tableau. Comme il peut y avoir plusieurs valeurs dans la même case, chaque case contient une liste de couples (*clef*, *valeur*) .

La structure est illustrée par le dessin suivant :

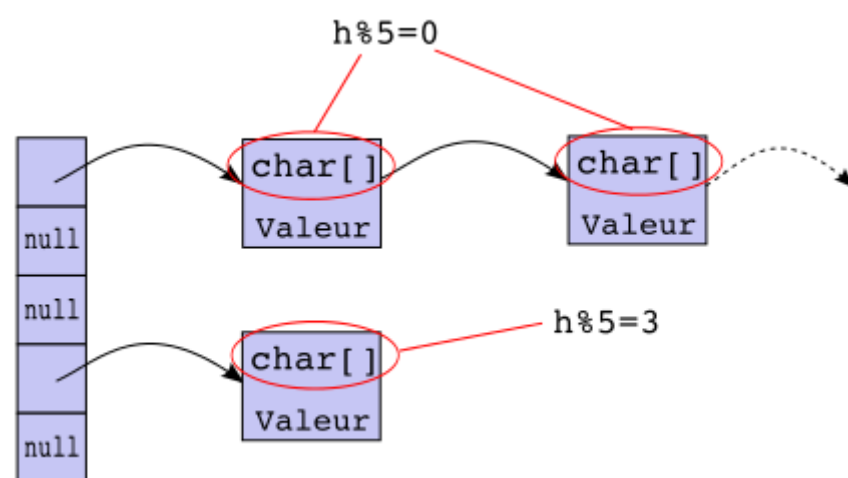


Table de hachage

En général, les opérations sur la table de hachage sont donc les suivantes :

- `add(key, value)` : considérer la liste présente dans la case $h(\text{key}) \bmod n$, chercher si `key` est dans la liste et dans ce cas lui associer la nouvelle valeur; sinon ajouter le couple (`key` , `value`) à la liste.
- `find(key)` : considérer la liste présente dans la case $h(\text{key}) \bmod n$, chercher si `key` est dans la liste et dans ce cas retourner la valeur associée; sinon retourner `null` .

On notera que dans ce cas, plus la table est pleine, plus les listes se remplissent et moins le temps de recherche est constant, même avec un bon hachage. Ainsi, si le tableau a pour taille n , mais que la table contient n^2 entrées, chaque liste a, au mieux, n éléments et les opérations ci-dessus sont en temps n . Ainsi, quand la table commence à être trop remplie, il est nécessaire de faire un « rehachage » des entrées : on agrandit le tableau et on répartit à nouveau les données à l'aide de la fonction de hachage.

Fonction de hachage

La classe `String` de la librairie standard de Java fournit une fonction de hachage sur les chaînes de caractères. Nous allons utiliser celle-ci pour écrire une fonction de hachage pour les préfixes.

Question 17 : Fonction de hachage

Ajoutez dans la classe `Prefix` une fonction `public int hashCode()` qui calcule la valeur de hachage pour un préfixe donné. Cette fonction procédera au mélange des valeurs de hachage de chacune des chaînes de caractères qui constituent le préfixe de la manière suivante :

```
h = 37*h + t[i].hashCode();
```

avec `i` parcourant le tableau des mots du préfixe, et en partant de `h=0` pour le tableau vide.

Pour travailler avec un tableau de taille `n` il faut ramener la valeur de hachage à un entier entre `0` et `n - 1` . Pour cela, ajoutez également une méthode `int hashCode(int n)` qui fait cela. On pourra utiliser le reste de la division entière de `a` et `b` qui est noté `a % b` en Java.

Attention ! La méthode `hashCode()` est susceptible de rendre des valeurs négatives (pour être précis, `%` n’implémente pas le modulo habituel mais le reste de la division entière). Du coup `a % b` est également susceptible d’être négatif. Il faut en tenir compte en écrivant `int hashCode(int n)` . Il existe maintenant une fonction `floorMod()` qui fait ce que l’on veut, à moins que l’on préfère `remainderUnsigned()` , mais on ne les utilisera pas pour ce TD.

Définition de la table associative

Gestion des entrées dans la table

Les entrées dans la table de hachage sont des couples $(clef, valeur)$ où *clef* est un objet de type `Prefix` et *valeur* un objet de type `WordList` .

Créez un nouvel enregistrement `Entry` permettant d’implémenter ces couples $(clef, valeur)$.

```
class Entry {
    Prefix key;
    WordList value;

    Entry (Prefix key, WordList value) {
        this.key = key;
        this.value = value;
    }
}
```

On a vu que la table de hachage utilise un tableau de taille fixe *N* pour stocker les couples $(clef, valeur)$. Dans le cas général, il peut arriver que 2 entrées dans la table aient le même indice dans la table. Pour implémenter la table de hachage, il faut prendre en compte ces collisions.

Pour cela, nous allons ajouter une classe `EntryList` , implémentant une chaîne d’objets de type `Entry` :

```
class EntryList {
    Entry head;
    EntryList next;

    EntryList(Entry head, EntryList next) {
        this.head = head;
        this.next = next;
    }
}
```

Question 18 : définition de la table

Créez une nouvelle classe `HMap` contenant la définition de la table de hachage. Cette classe doit contenir 2 attributs :

- `t` : le tableau de listes d’entrées;
- `nbEntries` : le nombre d’entrées stockées dans la table.

Vous ajouterez dans cette classe deux constructeurs :

- `HMap(int n)` permettant d’initialiser une table de hachage avec un tableau de listes d’entrées de taille *n*;
- `HMap()` faisant la même chose, mais pour une taille par défaut (par exemple 20).

Question 19 : recherche d’une entrée dans la table

Écrivez dans la classe `HMap` la fonction `WordList find(Prefix key)` qui renvoie la liste de mots associée à un préfixe donné. Si `key` ne peut être trouvé dans la table, `find` renverra la valeur `null`.

Question 20 : ajout d’une entrée dans la table

Ajoutez dans la classe `HMap`, une fonction `void addSimple(Prefix key, String w)` qui va ajouter le mot `w` à la liste de mots associés à `key` dans la table. S’il n’y a pas encore d’entrée associée à `key` alors on crée cette entrée en associant à `key` la liste contenant uniquement `w`.

Remarquez que :

- On a ici une version des tables de hachage spécifique à notre problème. On n’efface jamais une entrée et on peut juste modifier les entrées existantes en ajoutant des mots à la valeur.
- Vous devrez faire attention à bien faire évoluer le champ `nbEntries`. Il augmente lorsqu’on crée une nouvelle entrée, mais pas lorsqu’on ajoute un mot à une entrée existante.
- Il peut être utile d’ajouter une fonction `HMap.print` pour vous aider à débbugger votre code, mais ce n’est pas obligatoire.

Redimensionnement du tableau d’indice

Question 21 : rehachage

Comme expliqué précédemment, l’efficacité de la table de hachage dépend du choix de la taille du tableau d’entrées. Il est donc nécessaire d’implémenter une méthode `void rehash(int n)` qui fixe le tableau d’indice à la taille n et re-répartit l’ensemble des entrées dans la nouvelle table.

Question 22 : ajout amélioré

Ecrivez une méthode `add` similaire à `addSimple` mais améliorée pour provoquer un rehachage automatique quand la table a plus d’entrées que trois-quarts de la taille du tableau. On choisira dans ce cas de doubler la taille du tableau d’entrées.

Problème 5 : Madame Bovary

Lecture du texte

Le texte de Madame Bovary est donné sous la forme de 35 fichiers (un par chapitre de `01.txt` à `35.txt`) dans l’archive [bovary.zip](#) ([resources/bovary/bovary.zip](#)). Après avoir téléchargé l’archive, vous pouvez l’importer dans votre projet dans eclipse en commençant par cliquer avec le bouton droit de la souris sur votre projet java, dans la barre de gauche, puis en sélectionnant `import`, `Archive File`. Sélectionner le fichier `bovary.zip`, et `Finish`.

Vous disposez maintenant d’un sous-répertoire `bovary` qui contient les 35 chapitres. Le texte est soumis à [cette licence](#) (http://abu.cnam.fr/cgi-bin/donner_licence).

Vous devez également récupérer le fichier [WordReader.java](#) ([resources/src/WordReader.java](#)) contenant la classe `WordReader` permettant de lire les chapitres. Vous pouvez similairement l’importer avec un click droit sur `src` ou `(default package)`, `import`, `File System`.

Cette classe comprend entre autres, un constructeur `WordReader (String filename)` qui prend un nom de fichier en argument; ainsi qu’une méthode `main` d’essai. Pour passer un argument à la fonction `main`, aller dans `Run`, `Run Configurations...`, sélectionner `WordReader` dans `Java Application` à gauche, aller dans l’onglet `Arguments` à droite, et entrer `bovary/01.txt` (par exemple) dans `Program arguments`. C’est l’équivalent de la ligne de commande

```
% java WordReader bovary/01.txt
```

L’exécution affiche alors

```
[Nous]
[étions]
[à]
[l'Etude,]
[quand]
[le]
[Proviseur]
[entra]
[suivi]
[d'un]
...
[<PAR>]
...
```

Outre les mots du premier chapitre écrit par Flaubert, vous voyez apparaître des mots `<PAR>` qui indiquent simplement les paragraphes. Ces mots `<PAR>` sont à traiter comme des mots ordinaires.

La classe `WordReader` définit une méthode `String read()` qui lit un mot dans le chapitre en cours. Si le chapitre a été entièrement lu, la méthode `read()` renvoie `null`.

Ce formulaire est pour le problème 5. Vous devez soumettre le fichier `Bovary.java`.

You will have access to the files you submitted to Pb.1, Pb.2, Pb.3, Pb.4. (You do not need to resubmit them or to copy their content in the submitted files). If you submit files that belong to these previous questions, they will be propagated up and the automated tests will be triggered.

Choose

Choose one or more files...

Submit

☒ Reuse files from previous submissions ?

Expected files: Bovary.java

Construction de la table associative

Question 23 : `buildTable`

Créez une nouvelle classe `Bovary`, dans laquelle vous définirez la fonction `static HMap buildTable(String[] files, int n)` qui construit la table associative à partir des fichiers dont le nom se trouve dans le tableau `files`, avec des préfixes de taille n .

Génération de texte

Question 24 : `generate`

Ajoutez dans la classe `Bovary` une fonction `static void generate(HMap t, int n)` qui affiche sur la sortie standard un pseudo texte à partir de la table `t`, en utilisant des préfixes de taille n .

Sans chercher à produire une présentation parfaite, l’affichage devra être lisible (espaces entre les mots, saut de ligne à la place du mot `<PAR>`, retour à la ligne à la fin du texte).

Un click droit dans la console eclipse, là où le texte apparaît, permet de sélectionner `Word Wrap` pour une meilleure lisibilité.

Le 36^e chapitre de Mme Bovary

Question 25

Ajoutez une fonction `main` qui combine les fonctions `buildTable` et `generate` pour produire sur la sortie standard le 36^e chapitre de *Mme Bovary*. Elle pourra coder « en dur » la longueur 3 pour les préfixes et les noms des fichiers.