

**Rétrospective travail
effectué durant la
pause pédagogique**

Mise en place d'un filtre : Estetical Restructuration Rejector

Définition : filtre ayant pour but la détection de paraphrase

Exemple : triste mais néanmoins \longrightarrow tristement

Fonctionnement du filtre EsteticalRestructurationRejector

Définition :

Utilisation de la distance de Levenshtein:

- mot < 4 caractères :
 - distance L. $<$ taille plus grand mot
- mot ≥ 4 caractères :
 - distance L. acceptée $> 1/2$ taille plus petit mot

Fonctionnement du filtre EsteticalRestructurationRejector

Exemple de modifications rejetées:

- sa -> leur
- L'implémentation -> La mise en oeuvre

Restructuration globale du code

- **Restructuration du code (création d'interfaces et de classes abstraites)**
- **Possibilité d'obtenir les rejets de chaque filtres dans un fichier .csv**
- **Possibilité d'afficher les statistiques de chaque filtre**

Statistiques pour chaque filtre

RollbackFilter

Définition : Rejette les transformations suivantes : **A** -> **B** puis **B** -> **A**

exemple : La sorbonne -> La faculté des Lettres
↓
La faculté des Lettres -> La sorbonne

Statistiques : Pour 200 cas

	Nombre de cas rejetés	Taux de refus
RollbackFilter	42 cas	21 %

EsteticalRestructurationRejector

Définition : Rejette les paraphrases

exemple : triste mais néanmoins \longrightarrow tristement

Statistiques : Pour 200 cas

	Nombre de cas rejetés	Taux de refus
EsteticalRestructuration Rejector	89 cas	44,5 %

NumberRejector

Définition : Rejette les modifications concernant les chiffres

exemple :

1990 → 2012

Statistiques :

Pour 200 cas

	Nombre de cas rejetés	Taux de refus
NumberRejector	16 cas	8 %

SentencePurifier

Définition : Ne garde que la phrase qui doit être modifiée

exemple : Il fait beau. Les oiseaux chantent \longrightarrow Il fait beau.

Statistiques : Pour 200 cas

	Nombre de caractères traités/ supprimés	Taux de caractères supprimés
NumberRejector	66626/38809 caractères	58,24 %

SpecialCharacterPurifier

Définition : Supprime les caractères spéciaux en début de phrase

exemple : *Le lion est un animal —————> Le lion est un animal

Statistiques : Pour 200 cas

	Nombre de caractères supprimés	Taux de caractères supprimés par cas
NumberRejector	42 caractères	58,24 %

Couplage des filtres

Statistiques :

Pour 200 cas

	Nombre de cas traités	Nombre de cas rejetés	Taux de cas rejetés
RollbackFilter	200	42	21 %
NumberRejector	158	12	7,59 %
EsteticalRestructuration Rejector	146	62	42,46 %