

Evaluation automatique de productions écrites d'étudiants

Philippe Dessus
LSE & IUFM, Grenoble

Exposé au SEIGAD-REG (UJF) du 31 mars 2004

1. Problème
2. Présentation générale de LSA
3. Présentation d'Apex
4. Comment ça marche ?
5. Projets

1 2 Evaluer automatiquement des productions d'étudiants

- Questions à choix multiple
 - Réponses à la fois plausibles et fausses ?
- Niveau de surface : lisibilité
 - Nb de virgules, connecteurs, longueur moyenne des mots...
- Niveau lexical
 - fréquence des mots, cooccurrences, hapax
- Niveau sémantique
 - accès à la signification des mots. Représentation intermédiaire des mots et textes (*e.g.*, vectorielle)

2 3 Présentation générale de l'Analyse de la sémantique latente

- Méthode statistique proche d'une analyse factorielle
- Mise au point vers 1990 pour améliorer les requêtes de recherche d'information
- Devenue depuis un modèle de traitement de l'information humaine à partir de grands corpus de textes

2 4 Champs d'application de LSA

- Recherche d'informations documentaires
- Traduction automatique : langues, multimodalité
- Modélisation de l'apprentissage
- Mesure de la cohérence textuelle
- Compréhension de métaphores
- Tuteurs intelligents, EAD
- Apprentissage de jeux simples
- Décisions dans environnement dynamique : gestion de feux de forêt, enseignement

3 5 Apex, Assistance à la Préparation des EXamens (1)

- Cours hiérarchisé à 2 niveaux : thème (3-4 pages) et notions (1/2 page)

#T Recherches à propos de l'enseignant

#N Présentation générale des approches : prescriptive/descriptive et rationnel/arationnel
La littérature de sciences de l'éducation à propos de l'enseignant utilise beaucoup d'images (voir également Briciaux, 1997), parfois contradictoires, ...

#N L'approche prescriptive vs descriptive

Bru (1994), en distinguant les modèles de l'enseignement et les modèles pour l'enseignement, nous permet d'opérer une distinction claire entre deux sortes d'images. Celles ...

#N L'enseignant rationnel vs arationnel

Le deuxième critère est donc d'observer une (a-) rationalité chez l'enseignant (Tochon, 1993). Il s'agit de tester si la pensée de l'enseignant est prédictible, reproductible, ...

#T Modèles de planification

#N Modèles de planification, vue générale (tyleriens, décisionnels, de résolution de problèmes)

L'enseignant travaille sur des contenus, qu'il doit mettre en oeuvre pour des élèves et dans un contexte donné. Ce double ordre du jour qu'il doit gérer (Leinhardt) l'amène à faire des choix, anticiper des contenus, prévoir...

3 6 Apex (2) Traitements

- Ajout au cours d'une grande base textuelle (*Le Monde* 1999, 24 M mots) et traitement
- Etudiant écrit dissertation, à tout moment,
 - *contenu*, en appariant le texte avec chacune des notions du thème
 - *plan*, en appariant chaque paragraphe de la copie avec chaque notion du thème choisi
 - *cohérence*, proximité entre phrases contiguës

3 7 **Apex (3) copie d'écran**

SEIGAD-REG 04

3 8 **Apex (4) Liaison entre note humaine et note d'Apex**

- 21 copies d'examen de licence SE ont été évaluées par Apex
- Assez forte corrélation ($r = .59$) entre les notes données par l'enseignant et celles données par Apex

SEIGAD-REG 04

3 9 **Versions suivantes...**

- Apex 1.2 Portage sur PHP (Interface web), gestion des cours, enseignants, étudiants
- Apex 1.5 Analyse plus fine du résumé
 - segmentation automatique de la copie, puis comparaison de chaque segment à chaque partie de cours
 - Résultat : corrélation humains/Apex passe à 0,62
- Apex 2.0 Prise en compte du jugement de compréhension de l'étudiant

SEIGAD-REG 04

SEIGAD-REG 04

3 11 **Vue générale d'Apex 2.0**

Requête en langage naturel

Sélection texte plus proche

Boucle lecture

1. Afficher texte #n
2. Texte Résumable ?
3. Afficher prochain texte

Boucle écriture

1. L'étudiant écrit un résumé
2. Eval. du système: Comparaison avec tous les textes Résumables
3. Afficher les textes évalués différemment par l'étudiant et le système
4. Les textes évalués de la même manière deviennent non Présentables

#Txt	Lu	Ré	Pr
#1	1	1	0
#2	1	0	1
#3	1	1	0

SEIGAD-REG 04

3 12 **Apex 2.0 > Lecture**

Txt #1 Résumable

« I think I could sum it up. » « I don't think I could. »

Txt #2 Non Résumable

Txt #3 Résumable

SEIGAD-REG 04

