

Approche bayésienne de la composition sémantique dans les systèmes de dialogue oral

Marie-Jean Meurs



Séminaire UNIGE

1 juin 2010

Plan

Composition sémantique stochastique dans les SDS

Introduction

Corpus MEDIA

Frames
Sémantiques,
apprentissage

Génération
des fragments
sémantiques

Composition
des fragments
sémantiques

Conclusions
Perspectives

- 1 Introduction
- 2 Corpus MEDIA
- 3 Frames Sémantiques, apprentissage
- 4 Génération des fragments sémantiques
- 5 Composition des fragments sémantiques
- 6 Conclusions Perspectives

Introduction

Composition
sémantique
stochastique
dans les SDS

Introduction

Corpus MEDIA

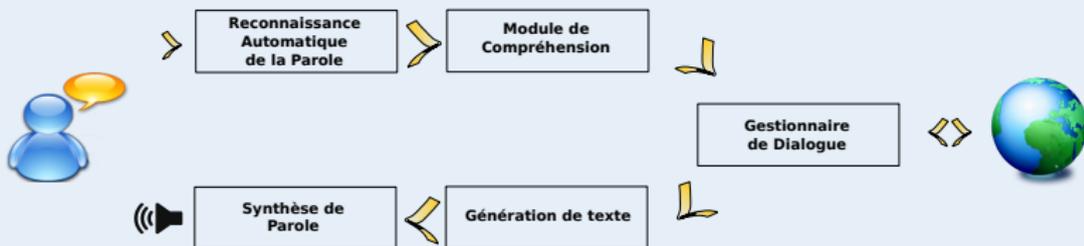
Frames
Sémantiques,
apprentissage

Génération
des fragments
sémantiques

Composition
des fragments
sémantiques

Conclusions
Perspectives

Structure d'un Système de Dialogue Oral



Système de Dialogue Oral - Applications

Échange d'informations

ATIS (Price, 1990) “How may I help you ?” (Gorin et al., 1997)

ARISE (Lamel et al., 2000)

“Le 3000” (Damnati et al., 2007)

Tutorat, apprentissage

SCoT (Pon-Barry et al, 2006) ITSPOKE (Litman et al., 2004)

Aide au handicap, aide aux personnes âgées

Smart WheelChair (Hockey&Miller, 2007)

LET'S GO (Raux et al., 2003)

Problématique

Complexité $\blacktriangleleft\blacktriangleright$ Tâche visée

- **routage, formulaires** 😊
 - représentation des connaissances sémantiques simple
- **demandes multiples, négociations** ☹️
 - représentation des connaissances sémantiques de haut niveau
 - ▷ interactions avec états courants du dialogue
 - ▷ robustesse aux erreurs de reconnaissance et compréhension

Système de Dialogue

- ✓ remise en cause
- ✓ propositions alternatives
- ✓ adaptation aux négociations

⇒ Représentation sémantique

- ✓ composition intra tour de parole
- ✓ composition inter tours de parole

Information sémantique

- ✓ riche
- ✓ pertinente
- ✓ généralisable

⇒ **Etape de composition supplémentaire**

- ✓ capturer abstractions et relations sous-tendues par la représentation conceptuelle de base

Système robuste et adapté à l'incertitude

⇒ Composants stochastiques

- ✓ amélioration de la robustesse
- ✓ production d'hypothèses multiples
- ✓ validation d'hypothèses
- ✓ limitation des besoins d'expertise
- ✓ limitation des coûts de développement

Matériaux et Constructions

- Corpus de dialogues MEDIA
- Constructions des structures sémantiques adaptées
- Annotation manuelle des dialogues de test
- Production des données d'apprentissage (système d'annotation déterministe)
- Entraînement des composants stochastiques
- Génération stochastique des fragments sémantiques sur les dialogues de test
- Composition des fragments sémantiques sur les dialogues de test

Corpus MEDIA

Production

- dialogues en français
- serveur téléphonique [tourisme, réservation d'hôtel].
- enregistrement selon le protocole de *Magicien d'Oz*.
- 1257 dialogues, 70 heures de parole, 250 locuteurs.
- 15.000 tours de parole locuteur, vocabulaire locuteur : 2.715 mots

Transcription et Annotation

- transcription manuelle
- annotation conceptuelle manuelle
- 2 annotateurs ELDA
- IAg \geq 80%

- ✓ Bonneau-Maynard *et al*,
"Semantic annotation of the MEDIA corpus for spoken dialog",
ISCA Eurospeech 2005

Le corpus MEDIA

Composition sémantique stochastique dans les SDS

Introduction

Corpus MEDIA

Frames
Sémantiques,
apprentissage

Génération
des fragments
sémantiques

Composition
des fragments
sémantiques

Conclusions
Perspectives

W^c	<i>concept c</i>	<i>valeur</i>
<i>Je voudrais réserver une chambre pour deux nuits à Marseille</i>	commande nb-chambre nb-nuit localisation-ville	réservation 1 2 Marseille

Exemple d'annotations sémantiques MEDIA.

- 83 concepts
- paire *concept-valeur* – segment de mots

► *nécessité de structures plus complexes pour une représentation complète de la composition sémantique*

Frames sémantiques, apprentissage

Frames sémantiques

Frame sémantique :

- situation abstraite ou concrète.
 - modèle computationnel : une entité sémantique \leftrightarrow ses propriétés.
 - rôles sémantiques : frame éléments (FEs).
 - représentation des dialogues de négociation.
 - adaptation aux actions complexes du gestionnaire de dialogue.
-
- ✓ J.B. Lowe, C.F. Baker, C.J. Fillmore,
"A frame-semantic approach to semantic annotation",
SIGLEX Workshop, 1997
 - ✓ C.J. Fillmore and C.R. Johnson and M.R.L. Petruck,
"Background to FrameNet",
International Journal of Lexicography, 2003

Base de connaissances

Composition sémantique stochastique dans les SDS

Introduction

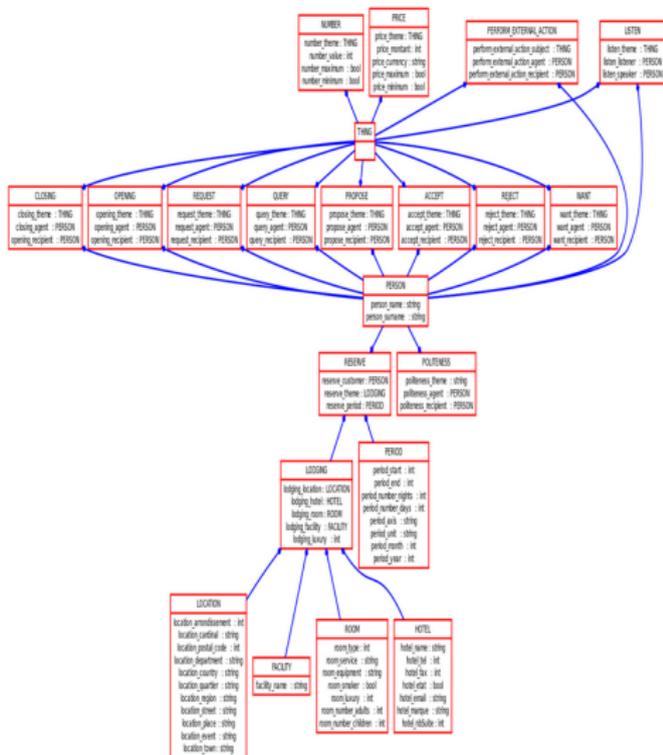
Corpus MEDIA

Frames Sémantiques, apprentissage

Génération des fragments sémantiques

Composition des fragments sémantiques

Conclusions Perspectives



Comparaisons FrameNet - Base MEDIA

Composition sémantique stochastique dans les SDS

Introduction

Corpus MEDIA

Frames
Sémantiques,
apprentissage

Génération
des fragments
sémantiques

Composition
des fragments
sémantiques

Conclusions
Perspectives

	Base MEDIA	FrameNet français	FrameNet
Frames	21	138	963
FES	86	1371	6800
Frame	REQUEST	REQUEST	REQUEST
FES associés	<i>FE_agent</i> <i>FE_recipient</i> <i>FE_theme</i>	<i>FE_speaker</i> <i>FE_adresse</i> <i>FE_message</i> <i>FE_topic</i> <i>FE_medium</i> <i>FE_manner</i> <i>FE_means</i>	<i>FE_speaker</i> <i>FE_adresse</i> <i>FE_message</i> <i>FE_topic</i> <i>FE_time</i> <i>FE_medium</i> <i>FE_manner</i> <i>FE_means</i> <i>FE_beneficiary</i>

- ✓ S. Pado and G. Pitel,
"Annotation précise du français en sémantique de rôles
par projection cross-linguistique", TALN, 1997

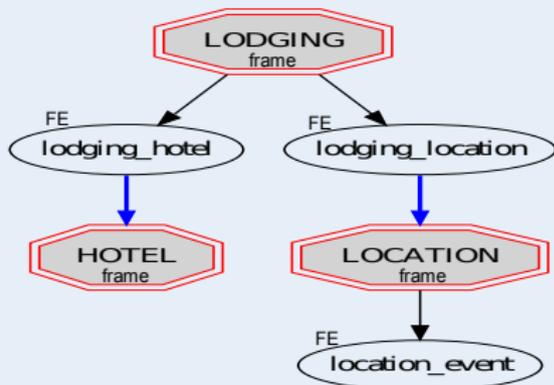
Définition des Frames

- 21 frames et 86 FEs dédiés au domaine MEDIA.
- frames et FEs décrits par un ensemble de patrons manuellement définis.
- des patrons combinant unités lexicales (LU) et conceptuelles (CU).

```
<frame fname="LOCATION">
  <concept value="localisation" />
  <lexical_units value="lieu,endroit" />
  <framelement fname="location_town">
    <concept value="localisation_ville" />
    <generic_lexical_units value="ville,cité,village" />
    <specific_lexical_units value="Paris,Marseille..." />
  </framelement>
  ...
</frame>
```

Extrait de la définition de la frame MEDIA LOCATION

Visualisation des annotations : exemple



Annotation de *“séjourner dans un hôtel proche du Festival de Cannes”*

► construction hiérarchique

Une annotation de référence en frames et FEs

Un processus en deux étapes, à base de règles :

- 1 **reconnaissance de patrons** : génération
 - 2 **règles logiques** : améliorations et liaisons
- Composition = création, modification ou suppression d'instances de frames et FEs
 - 100 règles environ
 - mots et ordre d'instanciation non considérés
 - processus progressivement enrichi

Une annotation de référence en frames et FEs

Composition
sémantique
stochastique
dans les SDS

Introduction

Corpus MEDIA

Frames
Sémantiques,
apprentissage

Génération
des fragments
sémantiques

Composition
des fragments
sémantiques

Conclusions
Perspectives

Évaluation de l'annotation de référence :

	Frames	FE	FE{Frames}	Liens	Liens{Frames}
\bar{p}					
\bar{r}					
$\overline{F-m}$					

*Précision, Rappel, F-mesure (valeurs moyennes) obtenus par le système d'annotation à base de règles sur les **3005** tours de parole de l'ensemble de test MEDIA.*

Une annotation de référence en frames et FEs

Évaluation de l'annotation de référence :

	Frames	FE	FE{Frames}	Liens	Liens{Frames}
\bar{p}	0.99	0.99	1.00	0.99	1.00
\bar{r}	0.99	0.97	0.97	0.94	0.95
$\overline{F-m}$	0.99	0.97	0.98	0.95	0.96

*F-mesure obtenu par le système d'annotation à base de règles sur les **3005** tours de parole de l'ensemble de test MEDIA.*

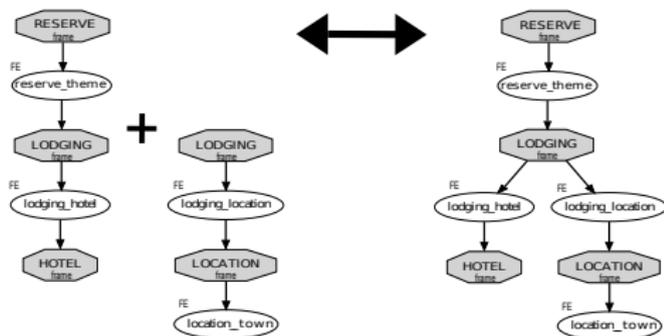
► *production d'une annotation de référence sur le corpus d'entraînement sur lequel les modèles stochastiques sont appris*

Génération des fragments sémantiques

Fragments sémantiques

Fragments

- éléments composés de frames et FE
- définition locale pour permettre un décodage séquentiel
- composition pour retrouver les relations globales



Réseaux Bayésiens Dynamiques (DBN)

Avantages des modèles probabilistes

- amélioration de la robustesse
- hypothèses multiples
- validation

DBN - Motivations

- flexibilité, modularité
- généricité
- bonnes performances
- cadre homogène

- ✓ F. Lefèvre,
"A DBN-based multi-level stochastic spoken language understanding system"
IEEE Workshop on SLT, 2006

DBN : Modèle compact - Modèle factorisé

Composition
sémantique
stochastique
dans les SDS

Introduction

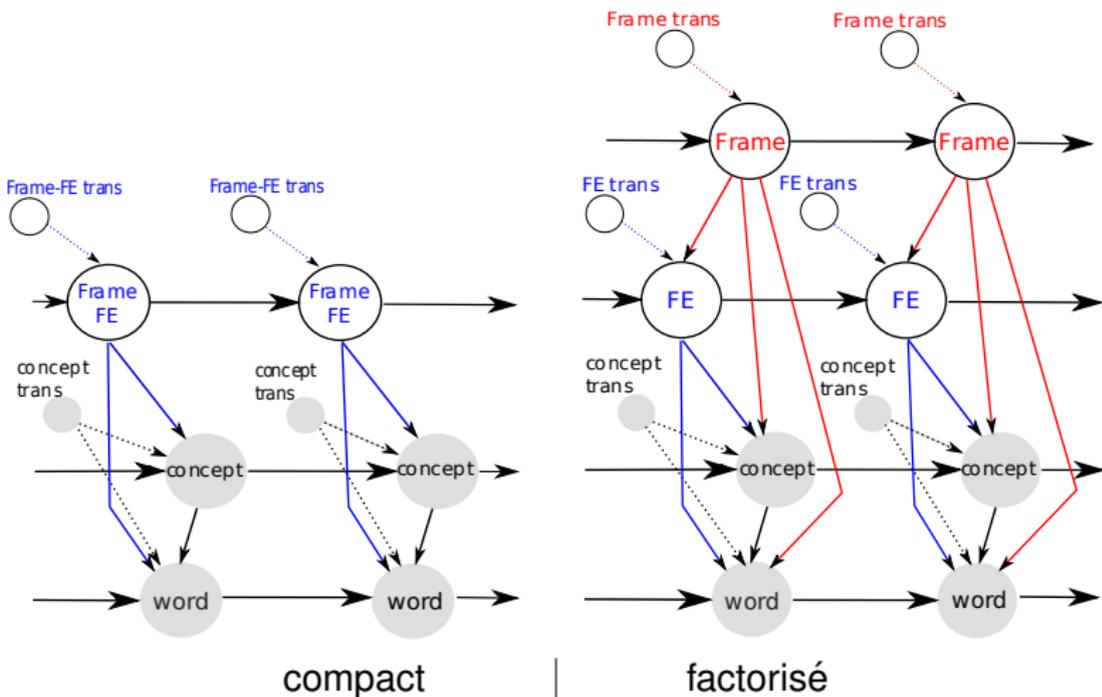
Corpus MEDIA

Frames
Sémantiques,
apprentissage

Génération
des fragments
sémantiques

Composition
des fragments
sémantiques

Conclusions
Perspectives



DBN : Modèle à deux niveaux

Composition
sémantique
stochastique
dans les SDS

Introduction

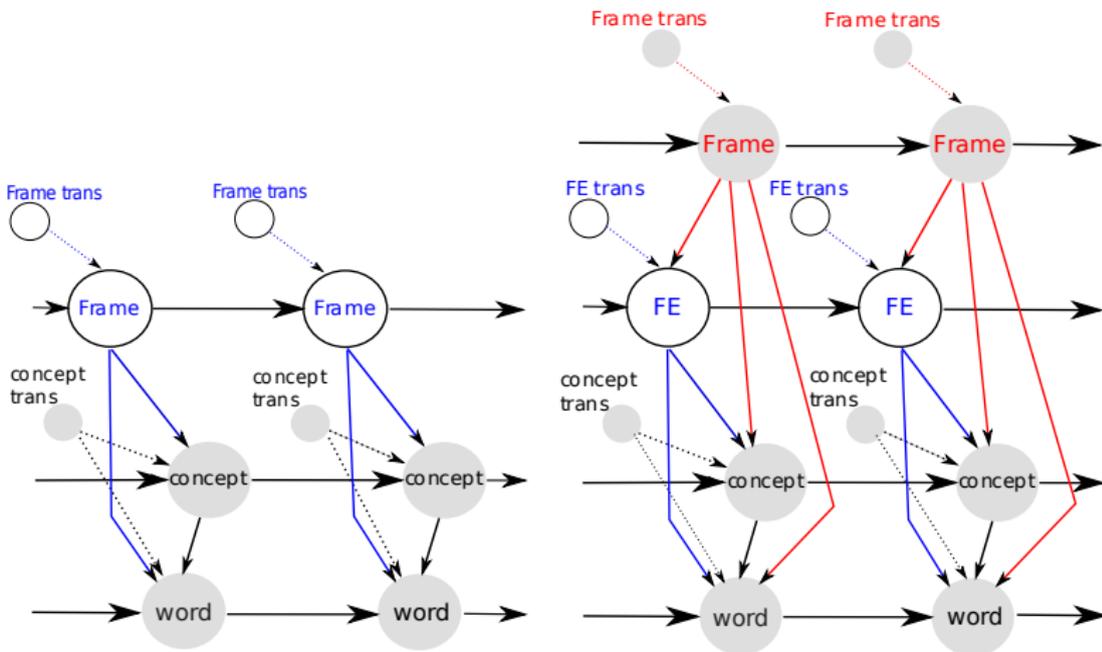
Corpus MEDIA

Frames
Sémantiques,
apprentissage

Génération
des fragments
sémantiques

Composition
des fragments
sémantiques

Conclusions
Perspectives



2 niveaux

DBN - Implémentation

Déterminer F qui maximise

$$P(W, C, F) = P(W|C, F)P(C|F)P(F).$$

Dans le contexte de l'interprétation du dialogue oral :

- variables toutes observées pendant l'entraînement
- tables de probabilités conditionnelles directement issues des observations
- amélioration des estimations :
 - modèles de langages factorisés FLM (\neq éléments)
 - repli parallèle généralisé GPB

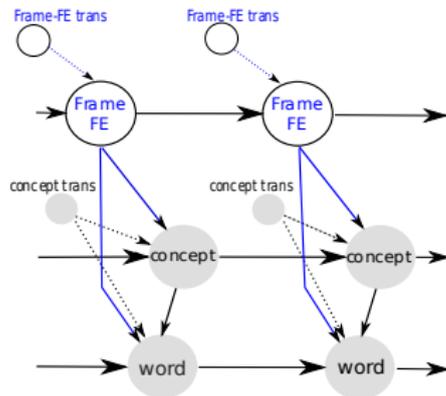
Paramètres des modèles

Modèle compact

$$P(FFE) \simeq \prod_{k=0}^n P(ffe_k | ffe_{k-1}) ;$$

$$P(C|FFE) \simeq \prod_{k=0}^n P(c_k | c_{k-1}, ffe_k) ;$$

$$P(W|C, FFE) \simeq \prod_{k=0}^n P(w_k | w_{k-1}, c_k, ffe_k).$$



Paramètres des modèles

Composition
sémantique
stochastique
dans les SDS

Introduction

Corpus MEDIA

Frames
Sémantiques,
apprentissage

Génération
des fragments
sémantiques

Composition
des fragments
sémantiques

Conclusions
Perspectives

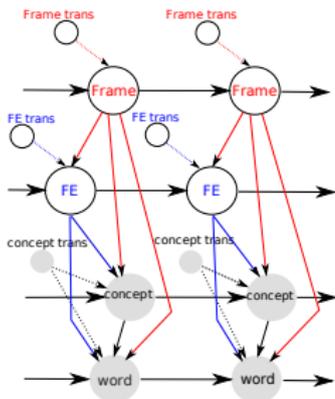
Modèle factorisé

$$P(F) \simeq \prod_{k=0}^n P(f_k | f_{k-1});$$

$$P(FE|F) \simeq \prod_{k=0}^n P(fe_k | fe_{k-1}, f_k);$$

$$P(C|FE, F) \simeq \prod_{k=0}^n P(c_k | c_{k-1}, fe_k, f_k);$$

$$P(W|C, FE, F) \simeq \prod_{k=0}^n P(w_k | w_{k-1}, c_k, fe_k, f_k).$$



Paramètres des modèles

Composition
sémantique
stochastique
dans les SDS

Introduction

Corpus MEDIA

Frames
Sémantiques,
apprentissage

Génération
des fragments
sémantiques

Composition
des fragments
sémantiques

Conclusions
Perspectives

Modèle à 2 niveaux

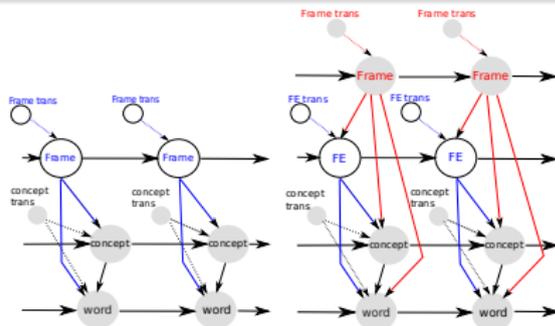
- *Premier niveau* : compact, frames seules
- *Second niveau* : factorisé, frames observées

$$P(\hat{F}) \simeq \prod_{k=0}^n P(\hat{f}_k | \hat{f}_{k-1});$$

$$P(FE | \hat{F}) \simeq \prod_{k=0}^n P(fe_k | fe_{k-1}, \hat{f}_k);$$

$$P(C | \hat{F}, FE) \simeq \prod_{k=0}^n P(c_k | c_{k-1}, \hat{f}_k, fe_k);$$

$$P(W | C, \hat{F}, FE) \simeq \prod_{k=0}^n P(w_k | w_{k-1}, c_k, \hat{f}_k, fe_k).$$



Algorithme de projection d'arbres

Composition sémantique stochastique dans les SDS

Introduction

Corpus MEDIA

Frames
Sémantiques,
apprentissage

Génération
des fragments
sémantiques

Composition
des fragments
sémantiques

Conclusions
Perspectives

- Représentation hiérarchique \Rightarrow recouvrements lors de la détermination des **fragments** de frames et FEs associés à un concept
 - algorithme de projection appliqué à l'arbre de la phrase complète
 - extraction des fragments associées à un concept
 - opérations : séparation, duplication

séquence :

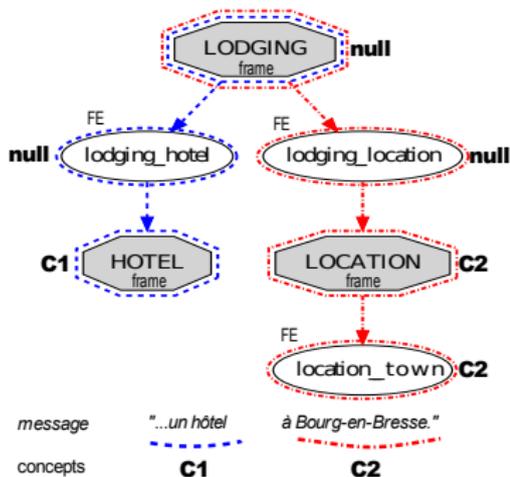
"un hôtel à Bourg-en-Bresse"

fragments projetées :

HOTEL-lodging_hotel-LODGING et

location_town-LOCATION

-lodging_location-LODGING.



Ensembles d'entraînement et de test :

- entraînement des DBN : transcriptions manuelles, annotations conceptuelles manuelles et annotations en frames semi-manuelles.
- fiabilité du système à base de règles : $F-m \geq 0,95$ pour l'identification des frames, FEs et liens sur le test.
- ensemble de test MEDIA : 3005 tours de parole utilisateur

Dimensionnement des modèles DBN

Composition sémantique stochastique dans les SDS

Introduction

Corpus MEDIA

Frames
Sémantiques,
apprentissage

Génération
des fragments
sémantiques

Composition
des fragments
sémantiques

Conclusions
Perspectives

DBN	Mots	Concepts	frames-FE	frames	FE
compact	2201	78	636	x	x
factorisé	2201	78	x	234	339
2-niveaux	2201	78	x	234	339

Données initiales :

- MAN : tours de parole manuellement transcrits et annotés ;
- SLU : concepts décodés à partir des transcriptions manuelles en utilisant un modèle à base de DBN ;
- ASR+SLU : concepts décodés à partir de la 1-best hypothèse de séquences de mots générée par le système ASR.

Type de données	SLU	ASR + SLU
Taux d'erreurs mots (%)	0,0	31,4
Taux d'erreurs concepts (%)	18,3	33,0

Résultats

Composition sémantique stochastique dans les SDS

Introduction

Corpus MEDIA

Frames
Sémantiques,
apprentissage

Génération
des fragments
sémantiques

Composition
des fragments
sémantiques

Conclusions
Perspectives

Modèles DBN	Données	MAN	
		Frames	FE Frames-FE
F/FE (compact)	$\overline{F-m}$		
F et FE (factorisé)	$\overline{F-m}$		
F puis FE (2-niveaux)	$\overline{F-m}$		
		SLU	
F/FE (compact)	$\overline{F-m}$		
F et FE (factorisé)	$\overline{F-m}$		
F puis FE (2-niveaux)	$\overline{F-m}$		
		ASR + SLU	
F/FE (compact)	$\overline{F-m}$		
F et FE (factorisé)	$\overline{F-m}$		
F puis FE (2-niveaux)	$\overline{F-m}$		

Résultats

Composition sémantique stochastique dans les SDS

Introduction

Corpus MEDIA

Frames
Sémantiques,
apprentissage

Génération
des fragments
sémantiques

Composition
des fragments
sémantiques

Conclusions
Perspectives

Modèles DBN	Données	Frames	MAN	
			FE	Frames-FE
F/FE (compact)	$\overline{F-m}$	0.92	0.92	0.89
F et FE (factorisé)	$\overline{F-m}$	0.91	0.88	0.87
F puis FE (2-niveaux)	$\overline{F-m}$	0.91	0.87	0.85
SLU				
F/FE (compact)	$\overline{F-m}$	0.90	0.88	0.87
F et FE (factorisé)	$\overline{F-m}$	0.89	0.85	0.84
F puis FE (2-niveaux)	$\overline{F-m}$	0.90	0.84	0.83
ASR + SLU				
F/FE (compact)	$\overline{F-m}$	0.84	0.84	0.81
F et FE (factorisé)	$\overline{F-m}$	0.83	0.82	0.78
F puis FE (2-niveaux)	$\overline{F-m}$	0.84	0.81	0.77

Résultats

Composition sémantique stochastique dans les SDS

Introduction

Corpus MEDIA

Frames
Sémantiques,
apprentissage

Génération
des fragments
sémantiques

Composition
des fragments
sémantiques

Conclusions
Perspectives

- trois modèles aptes à générer des fragments sémantiques consistants.
- modèle compact avantage par la structure et le dimensionnement de la base de connaissances sémantique

**Pas de prise en compte des dépendances
“longue-distances”**

▶ **étape de composition complémentaire**

Composition des fragments sémantiques

Recomposition d'arbres sémantiques

Composition
sémantique
stochastique
dans les SDS

Introduction

Corpus MEDIA

Frames
Sémantiques,
apprentissage

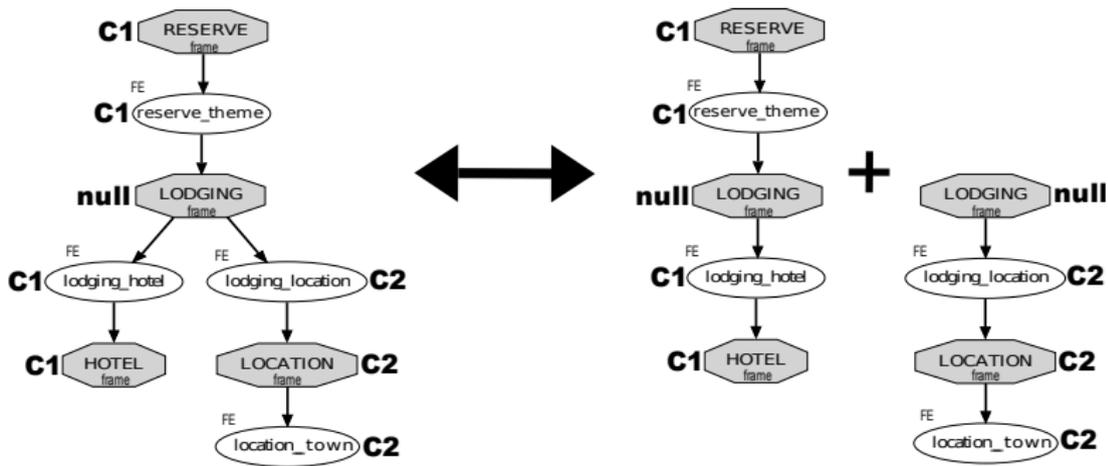
Génération
des fragments
sémantiques

Composition
des fragments
sémantiques

Conclusions
Perspectives

Rétablir l'arbre sémantique du message complet

- opérations de liaison entre frames et FE
- opérations de regroupement entre frames ou FE



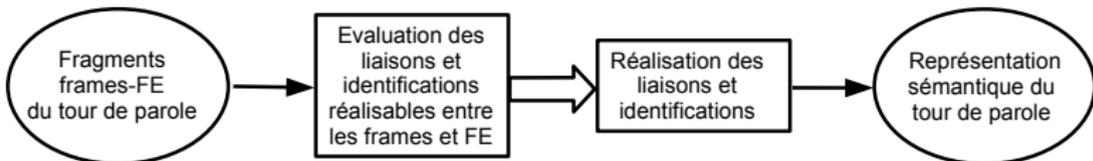
message : "réserver un hôtel à Bourg-en-Bresse"

concepts : **C1**

C2 

Méthode de connexion forte

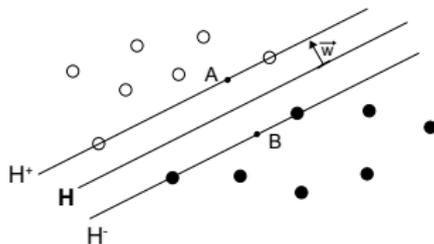
- ▶ toute opération de liaison ou d'identification possible selon l'ontologie est réalisée



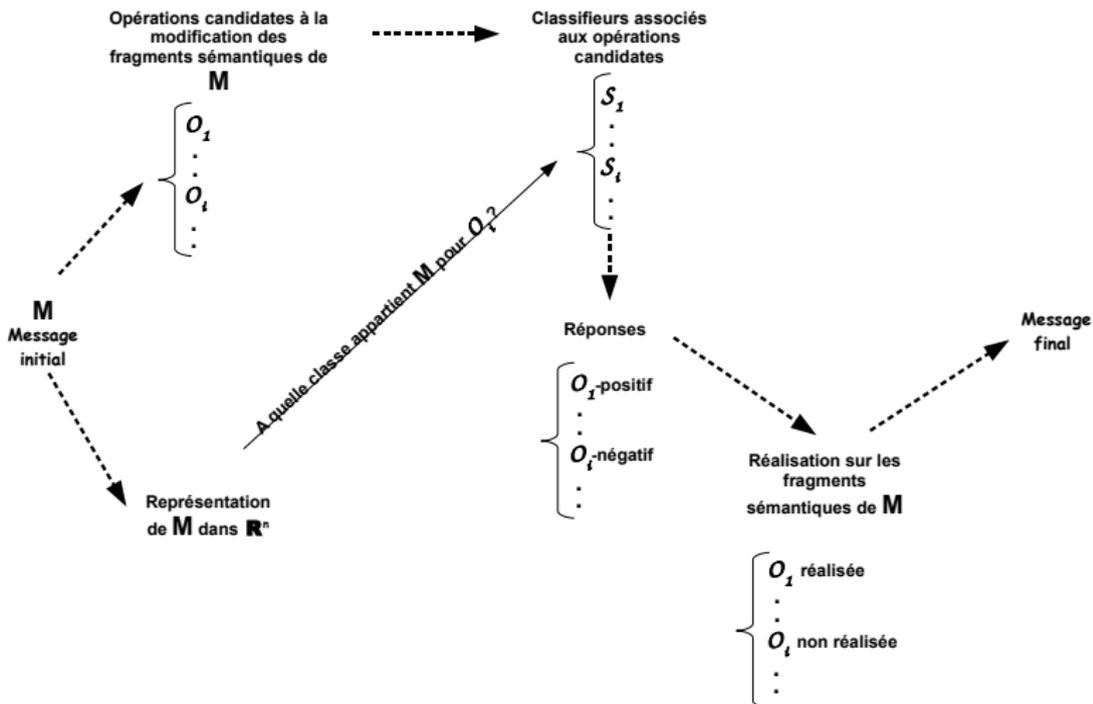
Méthode de connexion SVM

► Apprentissage des classifieurs

- recensement dans l'ensemble d'apprentissage des opérations de séparation et duplication réalisées lors de la projection des arbres
- pour chaque opération $\mathcal{O} = f_1 \mathcal{R} f_2$ (f_1 et f_2 frames et/ou FE) :
 - extraction des messages contenant f_1 et f_2
 - séparation des messages en deux classes selon qu'ils ont ou non déclenché \mathcal{O}
- apprentissage d'un classifieur SVM par opération



Méthode de connexion SVM



Données initiales :

- MAN : tours de parole manuellement transcrits et annotés ;
- SLU : concepts décodés à partir des transcriptions manuelles en utilisant un modèle à base de DBN ;
- ASR+SLU : concepts décodés à partir de la 1-best hypothèse de séquences de mots générée par le système ASR.

Type de données	SLU	ASR + SLU
Taux d'erreurs mots (%)	0,0	31,4
Taux d'erreurs concepts (%)	18,3	33,0

- Évaluation : fragments issus du modèle DBN compact

Résultats

Composition sémantique stochastique dans les SDS

Introduction

Corpus MEDIA

Frames
Sémantiques,
apprentissage

Génération
des fragments
sémantiques

Composition
des fragments
sémantiques

Conclusions
Perspectives

CONNEXION FORTE

CONNEXION SVM

	\bar{p}	\bar{r}	$\overline{F-m}$	\bar{p}	\bar{r}	$\overline{F-m}$
MAN						
Frames	0.95	0.93	0.93	0.94	0.95	0.93
FE	0.94	0.86	0.88	0.94	0.87	0.88
FE{Fr}	0.99	0.91	0.94	0.99	0.91	0.94
Liens	0.91	0.82	0.82	0.91	0.82	0.82
Liens{Fr}	0.98	0.88	0.91	0.98	0.88	0.91
SLU						
Frames	0.92	0.90	0.89	0.92	0.91	0.89
FE	0.91	0.83	0.84	0.91	0.84	0.84
FE{Fr}	0.98	0.91	0.93	0.98	0.91	0.93
Liens	0.90	0.80	0.79	0.89	0.80	0.79
Liens{Fr}	0.98	0.88	0.90	0.98	0.88	0.91
ASR+SLU						
Frames	0.86	0.82	0.80	0.86	0.83	0.80
FE	0.86	0.79	0.78	0.86	0.80	0.78
FE{Fr}	0.97	0.90	0.92	0.97	0.90	0.92
Liens	0.88	0.77	0.75	0.87	0.77	0.75
Liens{Fr}	0.97	0.87	0.90	0.97	0.87	0.90

Résultats

Composition sémantique stochastique dans les SDS

Introduction

Corpus MEDIA

Frames Sémantiques, apprentissage

Génération des fragments sémantiques

Composition des fragments sémantiques

Conclusions Perspectives

MAN

<i>Frames</i>	<i>Total REF</i>	CONNEXION FORTE		CONNEXION SVM	
		<i>Insertions</i>	<i>Suppressions</i>	<i>Insertion</i>	<i>Suppressions</i>
FACILITY	299	1	59	10	5
HOTEL	604	5	25	97	19
LODGING	859	61	286	148	278
LOCATION	574	0	43	7	34
RESERVE	478	65	4	76	2

SLU

<i>Frames</i>	<i>Total REF</i>	CONNEXION FORTE		CONNEXION SVM	
		<i>Insertions</i>	<i>Suppressions</i>	<i>Insertion</i>	<i>Suppressions</i>
FACILITY	299	10	78	24	30
HOTEL	604	22	42	95	36
LODGING	859	75	302	184	293
LOCATION	574	21	58	27	53
RESERVE	478	78	13	86	11

ASR+SLU

<i>Frames</i>	<i>Total REF</i>	CONNEXION FORTE		CONNEXION SVM	
		<i>Insertions</i>	<i>Suppressions</i>	<i>Insertion</i>	<i>Suppressions</i>
FACILITY	299	25	95	39	57
HOTEL	604	46	88	124	85
LODGING	859	84	324	182	318
LOCATION	574	84	111	87	108
RESERVE	478	91	39	98	39

- ▶ *connexion SVM à tester en dimension plus élevée*
- ▶ *évaluation des deux stratégies sur un sous-ensemble de messages complexes*

Conclusions et Perspectives

Conclusion

Composition
sémantique
stochastique
dans les SDS

Introduction

Corpus MEDIA

Frames
Sémantiques,
apprentissage

Génération
des fragments
sémantiques

Composition
des fragments
sémantiques

Conclusions
Perspectives

Représentation sémantique, données d'apprentissage

- Réalisation d'une base de données sémantique
- Développement d'un système à base de règles
- Annotation de l'ensemble d'apprentissage
- Annotation manuelle de l'ensemble de test

Génération des fragments sémantiques

- Modèles DBN
- Algorithme de décomposition d'arbres pour l'apprentissage des paramètres
- Production séquentielle de fragments consistants
- Robustesse à l'incertitude

Composition des fragments sémantiques

- Algorithmes de composition d'arbres : Connexion forte - Connexion SVM
- Traitement global des messages
- Recomposition pertinente des fragments

Module de compréhension

Atouts

- Système ouvert
- Base de connaissances sémantique évolutive
- Modèles DBN flexibles
- Cadre homogène pour les étapes de décodage séquentiel
- Collaboration DBN - SVM (généricité - discrimination)
- Proposition d'hypothèses multiples

Limites

- Corpus d'apprentissage indispensable
- Coûts de mise en oeuvre

Perspectives

Composition sémantique stochastique dans les SDS

Introduction

Corpus MEDIA

Frames
Sémantiques,
apprentissage

Génération
des fragments
sémantiques

Composition
des fragments
sémantiques

Conclusions
Perspectives

- Production de listes d'hypothèses scorées
- Prise en compte du contexte de dialogue
- Homogénéisation des décodages DBN

- Intégration dans un système de dialogue complet
- Développement d'un gestionnaire de dialogue apte à utiliser toutes les informations proposées
- ▶ *Évaluation de la pertinence et de l'influence de ces informations*

...

Composition sémantique stochastique dans les SDS

Introduction

Corpus MEDIA

Frames
Sémantiques,
apprentissage

Génération
des fragments
sémantiques

Composition
des fragments
sémantiques

Conclusions
Perspectives



Processus de construction de la base de connaissances

Composition sémantique stochastique dans les SDS

Introduction

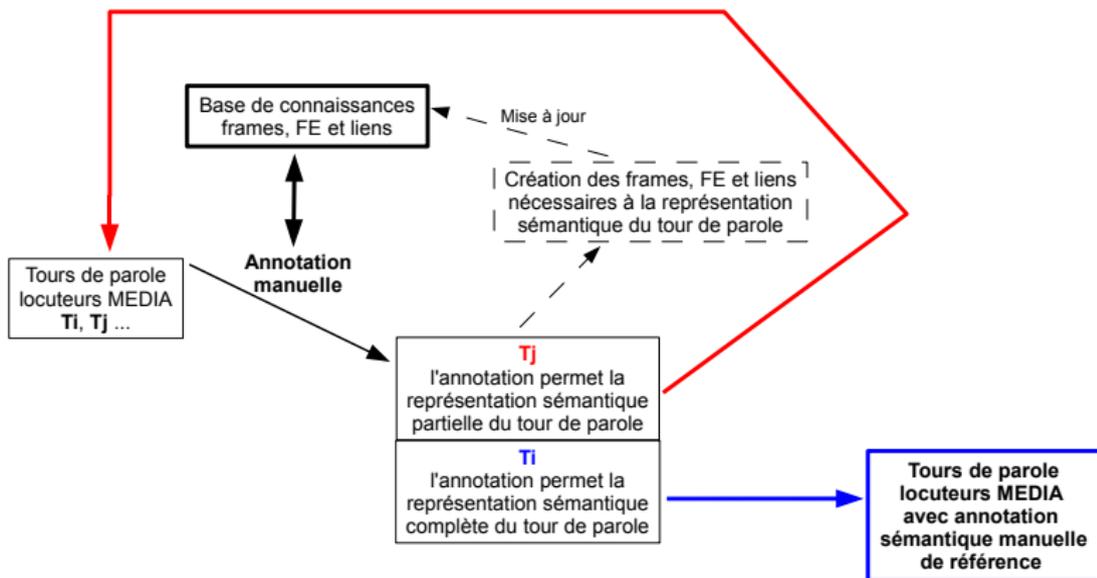
Corpus MEDIA

Frames
Sémantiques,
apprentissage

Génération
des fragments
sémantiques

Composition
des fragments
sémantiques

Conclusions
Perspectives



Base de connaissances

Composition sémantique stochastique dans les SDS

Introduction

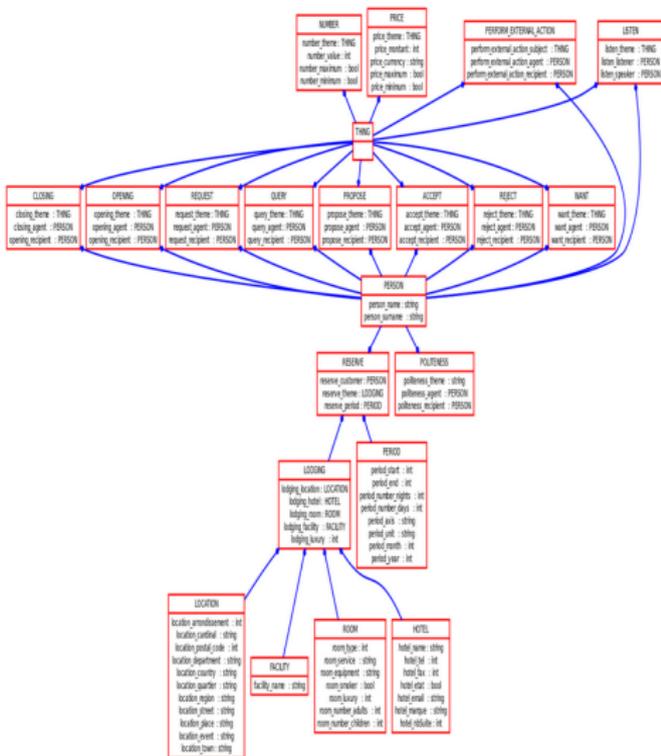
Corpus MEDIA

Frames Sémantiques, apprentissage

Génération des fragments sémantiques

Composition des fragments sémantiques

Conclusions Perspectives



Dimensionnement des modèles DBN

Composition sémantique stochastique dans les SDS

Introduction

Corpus MEDIA

Frames
Sémantiques,
apprentissage

Génération
des fragments
sémantiques

Composition
des fragments
sémantiques

Conclusions
Perspectives

DBN	Mots	Concepts	frames-FE	frames	FE
compact	2201	78	636	x	x
factorisé	2201	78	x	234	339
2-niveaux	2201	78	x	234	339

Résultats Fragments

Composition sémantique stochastique dans les SDS

Introduction

Corpus MEDIA

Frames
Sémantiques,
apprentissage

Génération
des fragments
sémantiques

Composition
des fragments
sémantiques

Conclusions
Perspectives

Modèles DBN	Données		MAN	
		Frames	FE	Frames-FE
F/FE (compact)	\bar{p}/\bar{r}	0.94/0.89	0.93/0.90	0.92/0.87
	$\bar{F}\text{-}m$	0.92	0.92	0.89
F et FE (factorisé)	\bar{p}/\bar{r}	0.92/0.89	0.89/0.88	0.88/0.86
	$\bar{F}\text{-}m$	0.91	0.88	0.87
F puis FE (2-niveaux)	\bar{p}/\bar{r}	0.91/0.90	0.88/0.86	0.85/0.84
	$\bar{F}\text{-}m$	0.91	0.87	0.85
SLU				
F/FE (compact)	\bar{p}/\bar{r}	0.93/0.87	0.92/0.83	0.91/0.84
	$\bar{F}\text{-}m$	0.90	0.88	0.87
F et FE (factorisé)	\bar{p}/\bar{r}	0.92/0.88	0.89/0.81	0.86/0.82
	$\bar{F}\text{-}m$	0.89	0.85	0.84
F puis FE (2-niveaux)	\bar{p}/\bar{r}	0.91/0.89	0.88/0.80	0.85/0.82
	$\bar{F}\text{-}m$	0.90	0.84	0.83
ASR + SLU				
F/FE (compact)	\bar{p}/\bar{r}	0.87/0.80	0.88/0.80	0.85/0.77
	$\bar{F}\text{-}m$	0.84	0.84	0.81
F et FE (factorisé)	\bar{p}/\bar{r}	0.85/0.82	0.84/0.78	0.80/0.76
	$\bar{F}\text{-}m$	0.83	0.82	0.78
F puis FE (2-niveaux)	\bar{p}/\bar{r}	0.86/0.82	0.84/0.77	0.80/0.75
	$\bar{F}\text{-}m$	0.84	0.81	0.77

Classifieurs SVM

Composition sémantique stochastique dans les SDS

Introduction

Corpus MEDIA

Frames
Sémantiques,
apprentissage

Génération
des fragments
sémantiques

Composition
des fragments
sémantiques

Conclusions
Perspectives

- 105 classifieurs
- identifications : 44 classifieurs (18 F et 26 FE)
- liaisons F-FE : 61 classifieurs