



Ontologie e locuzioni: modellizzazione dei termini complessi nell'ambito del progetto DIACOM

Paolo Frassi

Marco Rospocher

Dipartimento di Lingue e Letterature Straniere
Università degli Studi di Verona

XXX CONVEGNO ASS.I.TERM.

Risorse e strumenti per l'elaborazione e la diffusione della terminologia

EURAC BOLZANO

15-16 OTTOBRE 2020





- Progetto di eccellenza *Le digital humanities applicate alle lingue e letterature straniere* (2018-2022)
 - Sottoprogetto DIACOM-fr (banca dati terminologica in diacronia relativa al dominio del Commercio internazionale)
 - Rappresentazione dei dati terminologici,
 - unità multilesemiche (termini complessi)
 - Legami con altre entità lessicali
 - Ontologia OWL (W3C Web Ontology Language) -> Lemon

Progetto DIACOM-fr



- Progetto di eccellenza *Le digital humanities applicate alle lingue e letterature straniere* (2018-2019)
- Banca dati terminologica relativa al dominio del Commercio internazionale (francese)
- Costituzione di un corpus
 - Circa 530 testi (di tipologia diversa), 10.000.000 tokens, 3 periodi (1850-1914; 1945-1970; 1990-2018), riguardanti tematiche di varie ospitati su una piattaforma dedicata
 - Corpus disponibile per interrogazione tramite kontext
 - Estrazione terminologica (TermoStat) per realizzare banca dati



UNI VR Corpora Dashboard 2020.08 DIACOM FR								
Drafts	Documents	Trash	Macroéconomie et économie internationale					
+ New Draft	Refresh		Titre	Auteur	Texte intégral public	Date de publication	Date de consultation	Mots-clés
			9 aout 1873					macroéconomie et économie internationale, archives commerciales des consulats
			L'Alliance des peuples_10 octobre 1873	AA.VV.	No	1873	2018	comme internationnal, archives commerciales des consulats
			L'Alliance des peuples_12 juillet 1873	AA.VV.	No	1873	2018	comme internationnal, archives commerciales des consulats



kontext Query Corpora Save Concordance Filter Frequency Collocations View Help

Corpus: diacom_FR | Query: commerce, électronique (188 hits)

Hits: 188 | i.p.m.: 80.29 (related to the whole corpus) | ARF: 18.22 | Result is sorted

Line selection: simple

doc#0	, par exemple , à la manière de régir le commerce électronique , qui joue un rôle si grand dans l'économie
doc#6	l'économie et de la politique mondiales . RÉSUMÉ Le commerce électronique n'est plus un simple support promotionnel de la distribut
doc#6	, les rapports humains suscitent inévitablement des conflits et le commerce électronique ne fait pas exception à la règle . Ce nouveau
doc#6	de leurs caractéristiques , l'arbitrage fait ainsi écho au commerce électronique . En d'autres termes la gestion de la procédure
doc#6	répondre modestement au travers cet exposé . Mots clés : commerce électronique - résolution des litiges en ligne - arbitrage-arbitrage en lic
doc#6	un système de circulation de l'information et avec le commerce électronique , un réseau de circulation de biens offerts ou demandés
doc#6	vecteur puissant et efficace pour le commerce . Ainsi le commerce électronique est-il devenu l'un des grands enjeux actuels -
doc#6	et sans nul doute futur de l'internet . Le commerce électronique n'est plus un simple support promotionnel de la distribut
doc#6	effet , des entreprises de tout bord . Pour le commerce électronique , le Réseau est un outil dont le coût d'
doc#6	à une telle évolution . Cependant la problématique juridique du commerce électronique ne se limite pas aux difficultés soulevées par l'immédiat
doc#6	, les rapports humains suscitent inévitablement des conflits et le commerce électronique ne fait pas exception à la règle . Ce nouveau
doc#6	de leurs caractéristiques , l'arbitrage fait ainsi écho au commerce électronique . En réplique à la fulgurance des échanges , il
doc#6	I : Différences quant à la forme des conflits du commerce électronique Dans le cyberspace , la partie émergée de l'iceberg
doc#6	entrer dans une énumération fastidieuse41 , les différends liés au commerce électronique , peuvent être scindés en deux catégories : les litiges
doc#6	II : Différences quant au fond du droit applicable au commerce électronique « Internet is « a disputation arena , a forum
doc#6	d'ors et déjà une place de choix dans le commerce électronique mais qu'ils seront amener inévitablement à y jouer un
doc#6	va pas nécessairement de soi , notamment en matière de commerce électronique . Ainsi , le professeur Vincent Gautrais souligne qu'il excelle

Pagina 1 di 21

Unità multilesemiche



- ❖ Generalmente considerate termini complessi - es. **vente à distance** - in opposizione ai termini semplici - es. **vente**
- ❖ Quali proprietà sintattico-semantiche? Quali tipi di entità lessicali?
 - ❖ Locuzioni (es. **vente à distance**) e collocazioni (es.: **client fidèle**)
 - ❖ Proprietà di queste entità lessicali così come identificate per la lingua generale (Gross 1996; Mel'čuk 2013)
 - ❖ Locuzioni: esistono nelle tre tipologie proposte da Mel'čuk (2013):
 - ❖ Forti: *marge arrière*
 - ❖ Semi-deboli: *marketing viral*
 - ❖ Deboli: *vente à distance*

Unità multilesseemiche

- Studio pilota condotto relativamente alle locuzioni
 - Due ambiti tematici specifici (commercio elettronico; politica commerciale e relazioni internazionali)
 - Estrazione effettuata con Termostat →
 - Filtraggio manuale (eliminazione del rumore)
 - Risultati:
 - Convalidata la presenza dei tre tipi di locuzioni (forti 5-6%, semi-deboli 21-31%, deboli 65-73%);
 - Individuazione di sotto-categorie per le locuzioni deboli

Corpus >> marketing_txt

Résultats

Liste des termes | Nuage | Statistiques | Structuration | Bigrammes

Candidat de regroupement	Score	Fréquence (Spécificité)	Variantes orthographiques	Matrice
point de vente	468	108.08	point de vente points de vente points de ventes	Nom Préposition Nom
intelligence émotionnelle	364	105.19	intelligence émotionnelle	Nom Adjectif
contenu de marque	276	91.55	contenu de marque contenus de marque contenus de marques	Nom Préposition Nom
intention d' achat	268	89.86	intention d' achat intentions d' achat intentions d' achats	Nom Préposition Nom
brand content	222	82.07	brand content	Nom Adjectif
histoire racontée	226	81.29	histoire racontée histoires racontées	Nom PPAdj
client fidèle	215	79.4	client fidèle client fidèles clients fidèles	Nom Adjectif
stratégie de contenu	178	73.45	stratégie de contenu stratégie de contenus	Nom Préposition Nom
auto-efficacité perçue	133	63.43	auto-efficacité perçue	Nom PPAdj
grille de lecture	138	57.42	grille de lecture grilles de lecture	Nom Préposition Nom
marketing relationnel	100	54.93	marketing relationnel	Nom Adjectif
coopération commerciale	108	53.5	coopération commerciale coopération commerciales coopérations commerciales	Nom Adjectif
stratégie de contenu de marque	92	52.66	stratégie de contenu de marque stratégies de contenu de marque	Nom Préposition Nom Préposition Nom
marque de luxe	103	51.38	marque de luxe marques de luxe	Nom Préposition Nom



CONCERNER

'qui concerne...'

Accusé de réception

S'OCCUPER DE

'qui s'occupe de..
? 'qui concerne...'

Établissement commercial

OBJECTIF/FONCTION

'qui a comme
objectif/fonction...'
? 'qui concerne...'

Pays exportateur

ORIGINE

'qui a son origine dans...'

Marque de fabrique

REPOSER SUR

'qui repose sur...'
?'qui a son origine dans...'

Publicité comparative

MANIFESTATION

'qui se manifeste par...'

Argent électronique

LOCATIF

'qui a lieu, se trouve...'

Commerce de proximité

MOYEN

'qui se fait/a lieu par...'

Paiement électronique

PHASES/ÉCHELONS

'qui se situe au stade x'

Client final

TEMPS

'qui a lieu à un moment x'

Service après-vente



Legami con altre entità lessicali e relazioni paradigmatiche



- **Prix de vente**

- Legame formale con *vente* e con *prix*
- *Vente* è un termine polisemico:
 - *Vente1*: ‘Vente par X de Y à Z contre la somme W’ (es.: La limitation de responsabilité qui précède est une condition de vente de la marchandise au prix indiqué ou aux prix indiqués et s’appliquera [...] nonobstant tout défaut ou défaillance, incluant la défaillance complète de toute marchandise)
 - *Vente2*: ‘Vente par X de Y’ (es.: ‘La publicité de ces ventes se fait habituellement dans les journaux locaux et par des annonces sur papier affichées dans toute la ville’)



Prix de
vente

vente1

Vente2



Prix de
vente

Locution faible
Concerner

vente1

co-polysémie

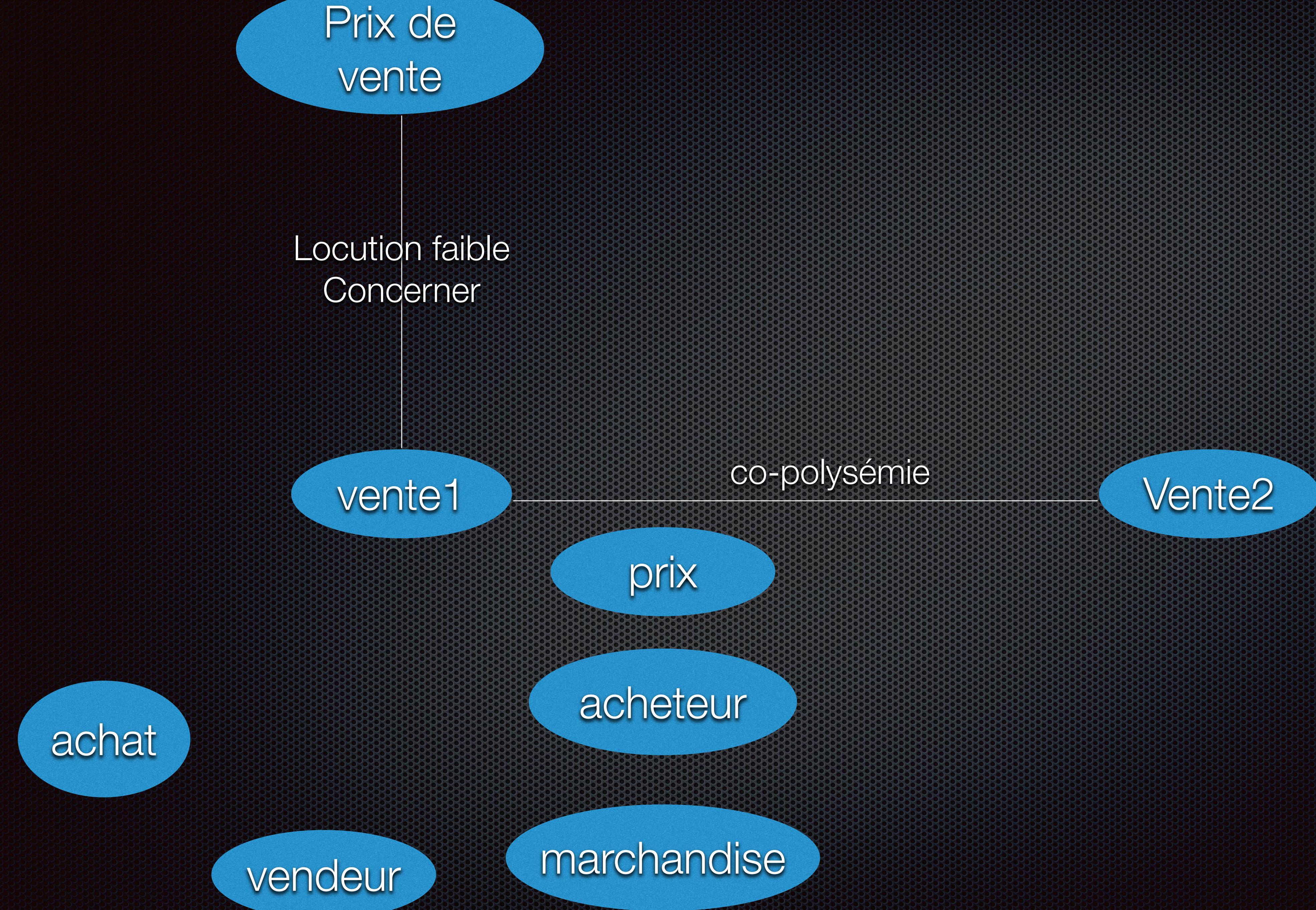
Vente2

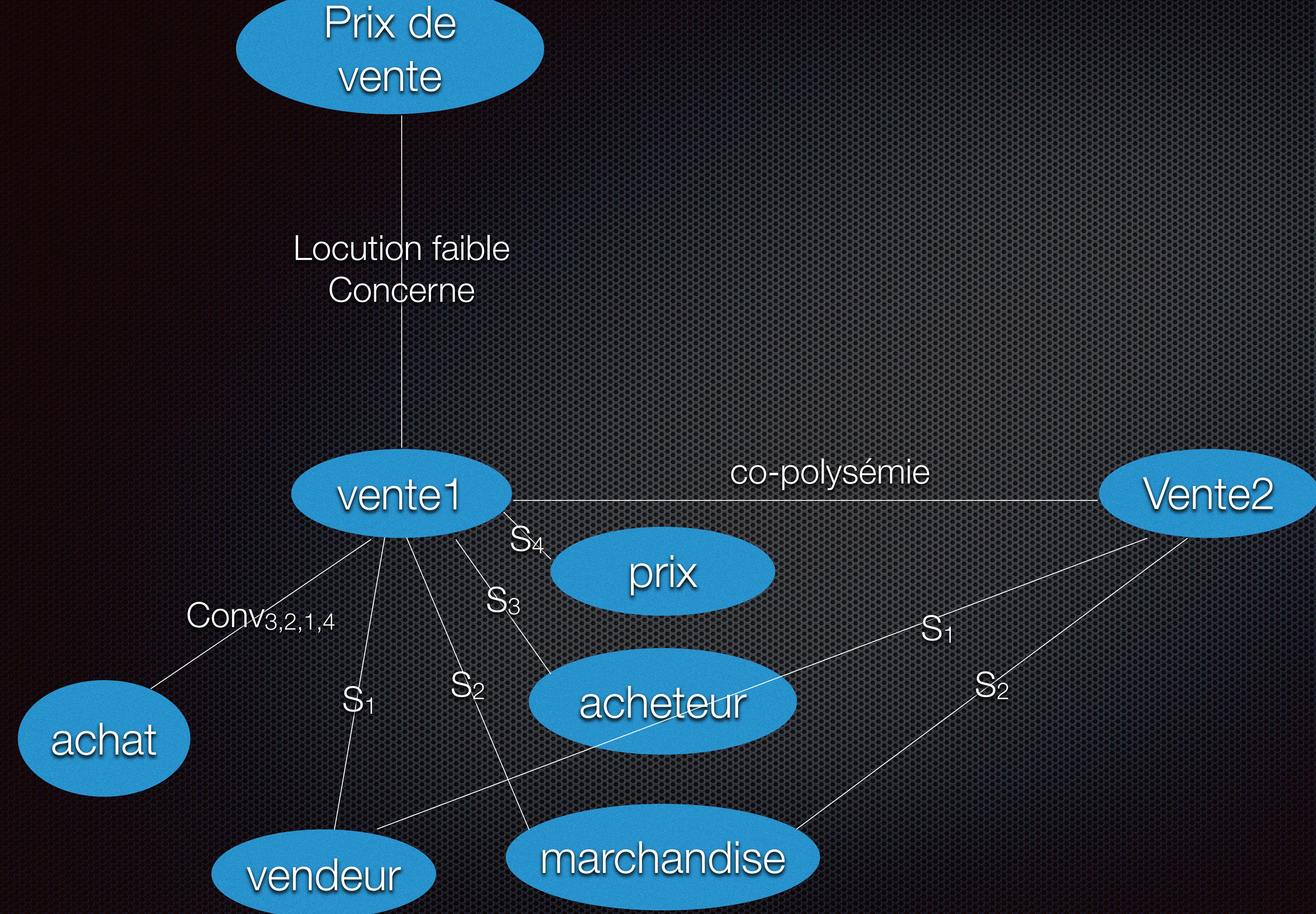
Legami con altre entità lessicali e relazioni paradigmatiche

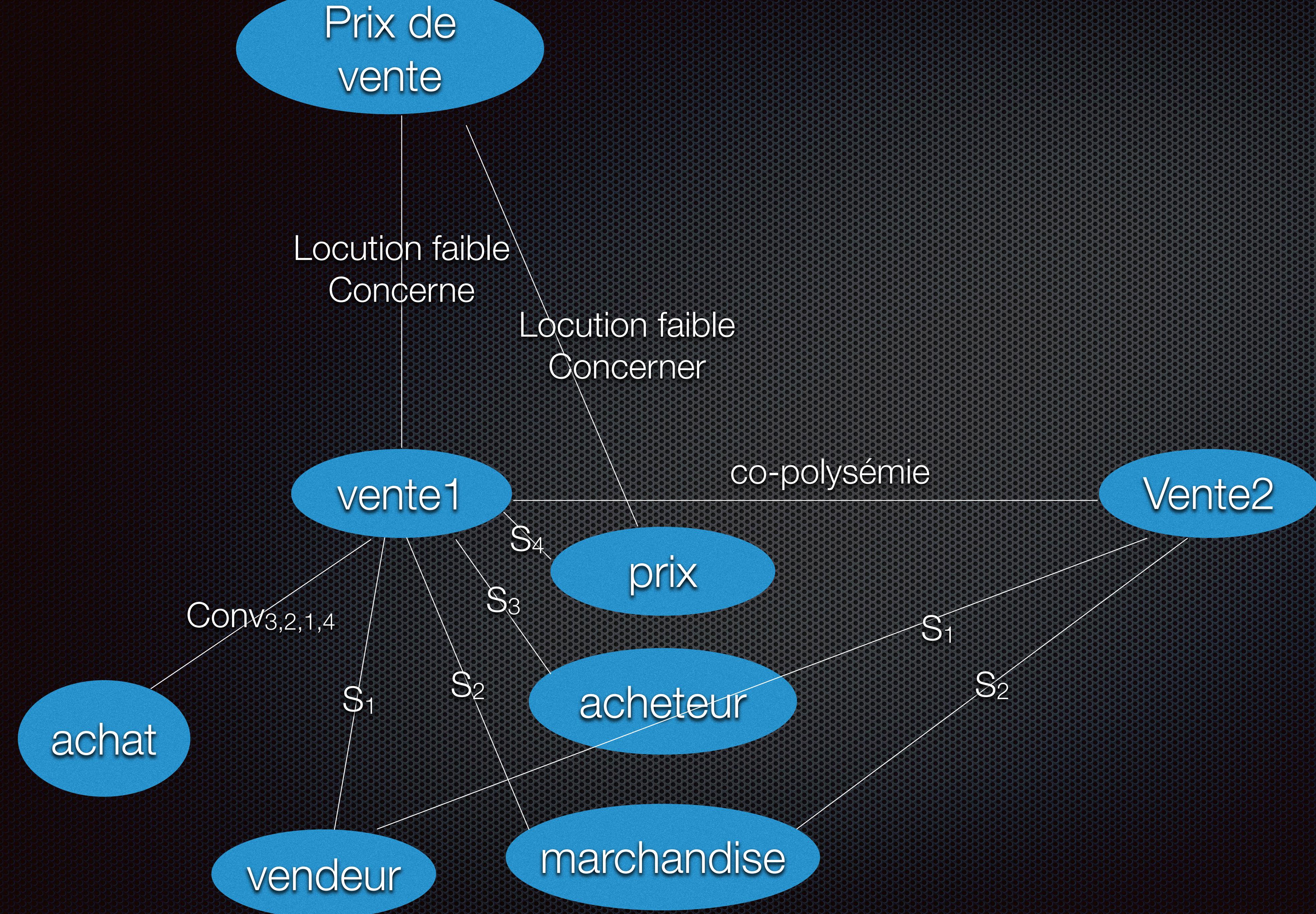


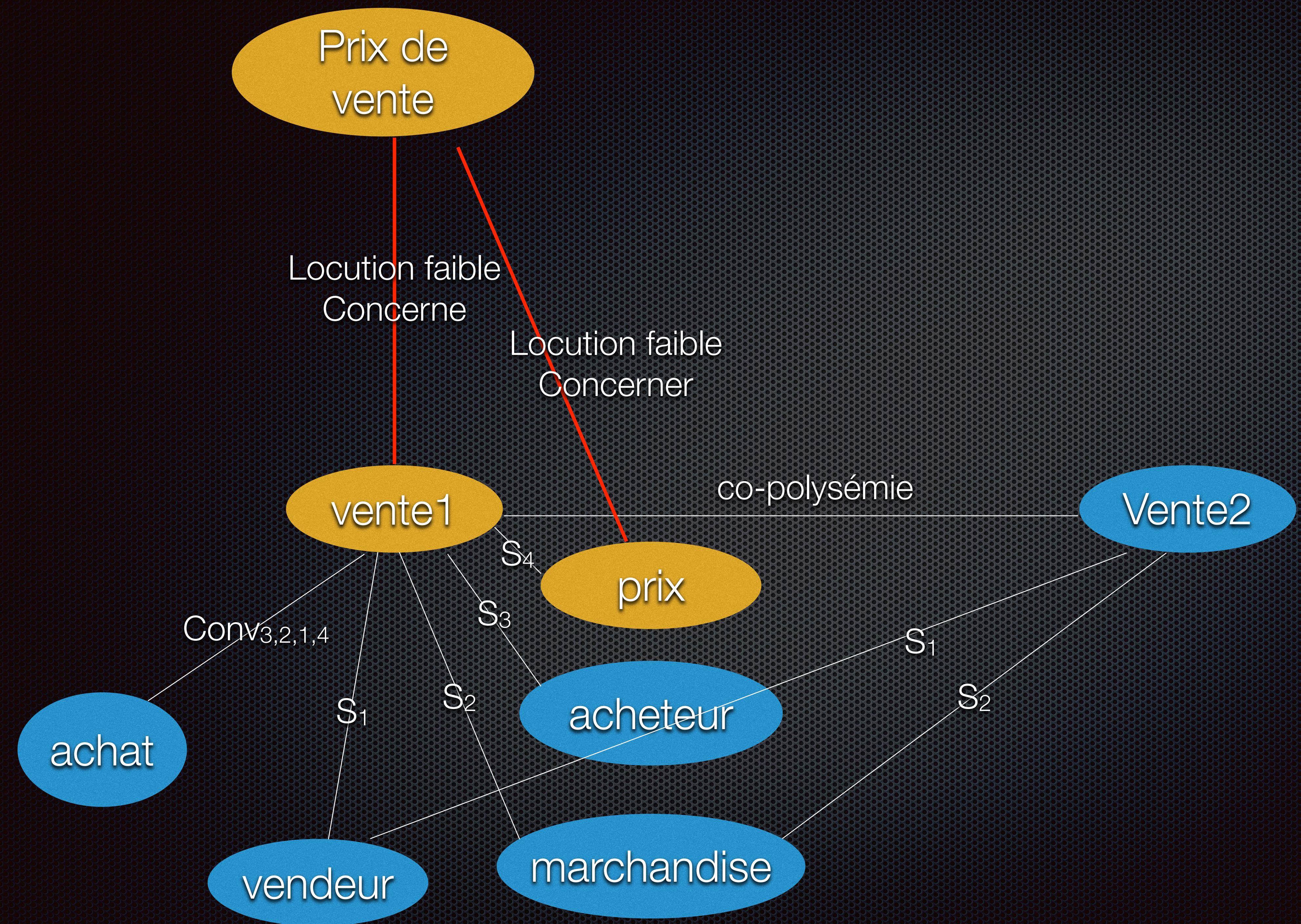
▪ **Prix de vente**

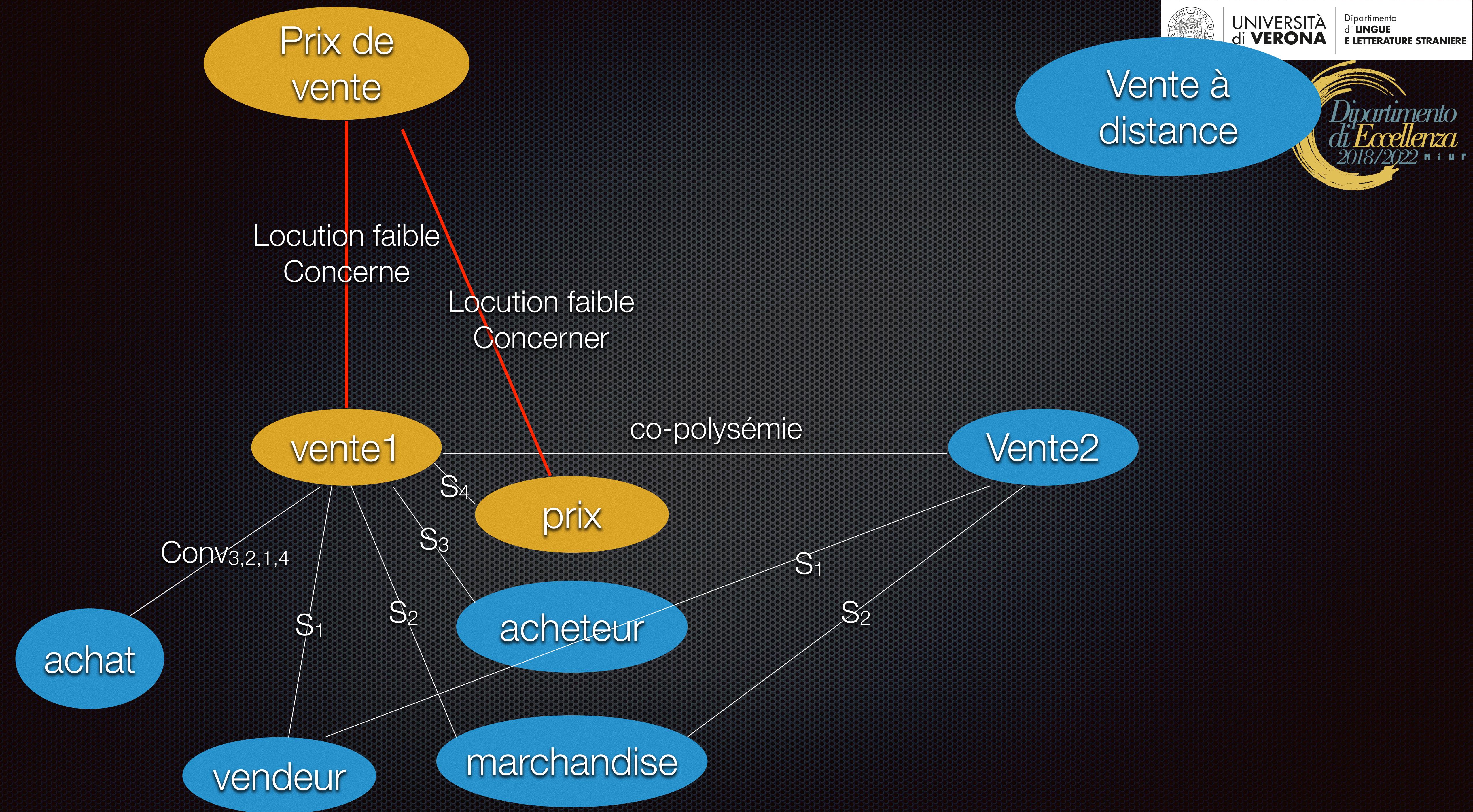
- Legame formale con *vente* e con *prix*
- *Vente1* e *Vente2* intrattengono legami paradigmatici
 - *Vente1*:
 - *Contr achat*
 - *Syn échange*
 - *S₁ vendeur*
 - *S₂ marchandise*
 - *S₃ acheteur*
 - *Vente2*:
 - *Syn échange*
 - *S₁ vendeur*
 - *S₂ marchandise*

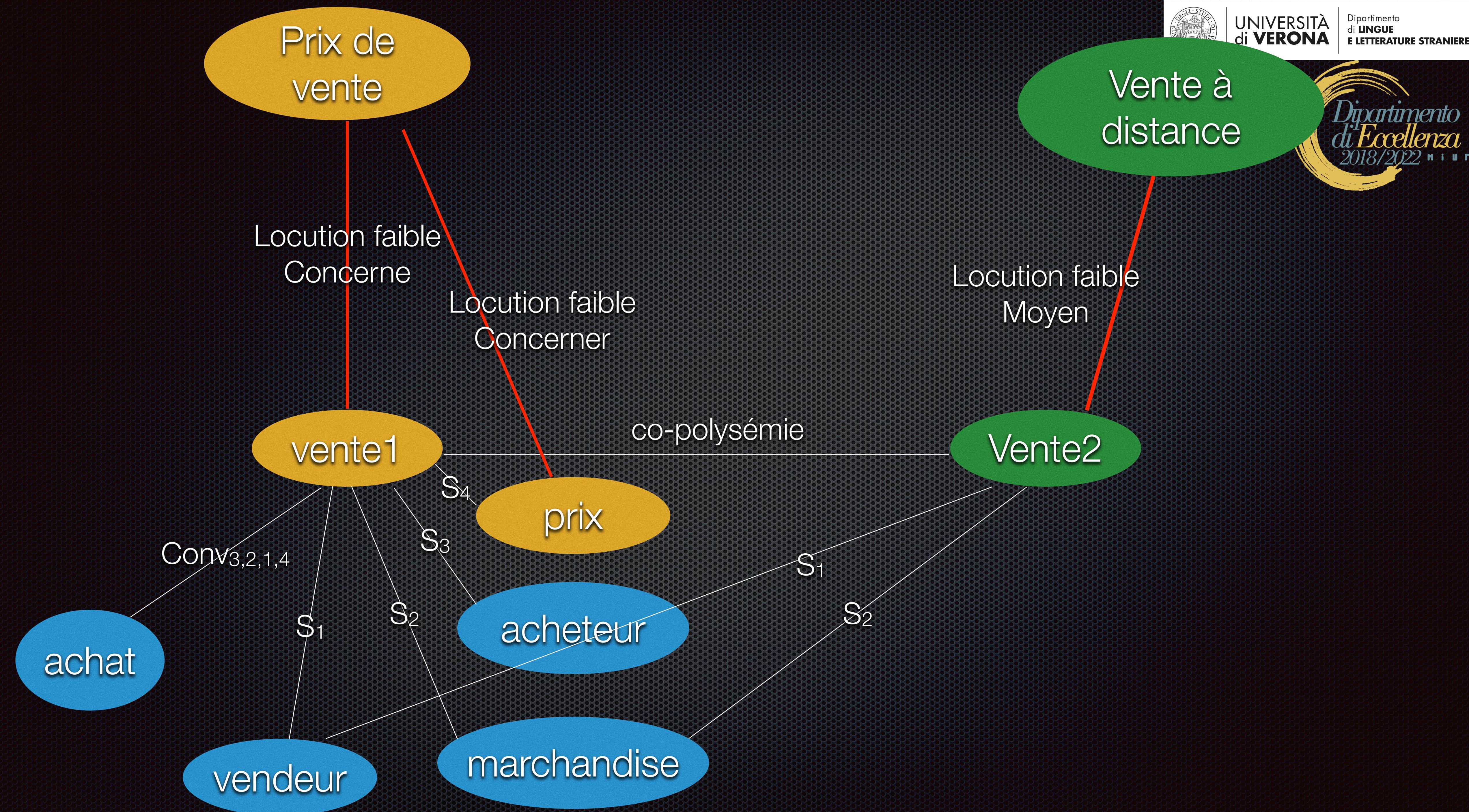














Elementi da codificare



- Tipi di locuzioni (deboli)
- Tipi di legami sintattico-semantici che possono intercorrere fra diversi tipi di unità terminologiche a seguito dell'estrazione



Risorsa Lessicale DiaCom

- Creazione di una risorsa secondo le best-practice del Semantic Web e Linked Data e le relative tecnologie (e.g., OWL, RDF)
- Questa scelta permette:
 - il collegamento semantico della risorsa con altre risorse linguistiche (e.g., Linguistic Linked Open Data);
 - l'interrogazione, in maniera flessibile, di tutte le informazioni contenute;
 - la visualizzazione delle locuzioni e i loro legami.

Risorsa Lessicale DiaCom



- Riuso di ontologie e vocabolari esistenti:
 - Lemon (Lexicon Model for Ontologies - W3C Ontolex) e LexInfo
 - Framework di base per la modellazione delle entità lessicali e sensi
- Lexfom (A lexical functions ontology model - Fonseca et al. 2016)
 - Modellazione delle funzioni lessicali (e.g., attanti semantici - Mel'čuk, 2004)

Tassonomia delle locuzioni deboli



Class hierarchy: S'occuper de

Annotations: S'occuper de

Description: S'occuper de

Asserted

owl:Thing

concept (skos:Concept)

concept scheme

entrée lexicale

expression

foaf:Organization

foaf:Person

forme

Lexicon

Locution

Locution faible

Concerner

Objectif/Fonction

S'occuper de

Locatif

Manifestation

Moyen

Origine

Reposer sur

Phases/Échelons

Temps

Locution forte

Locution sémi-compositionnelle

meaning

OntoMap

ParadigmaticLFSenseRelation

rdfs:Resource

Vocabulary

rdfs:label [language: fr]

S'occuper de

rdfs:comment [language: fr]

Qui s'occupe de...

exemple [language: fr]

Établissement commercial

Equivalent To +

SubClass Of +

Concerner

General class axioms +

SubClass Of (Anonymous Ancestor)

Instances +

Target for Key +

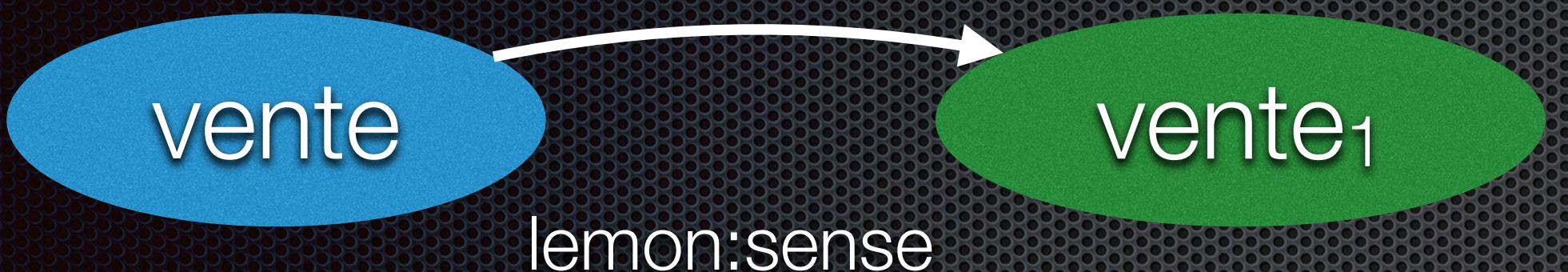
Disjoint With +

Disjoint Union Of +

Modellazione delle Locuzioni

- Rappresentate come individui, classificati secondo le rispettive tipologie

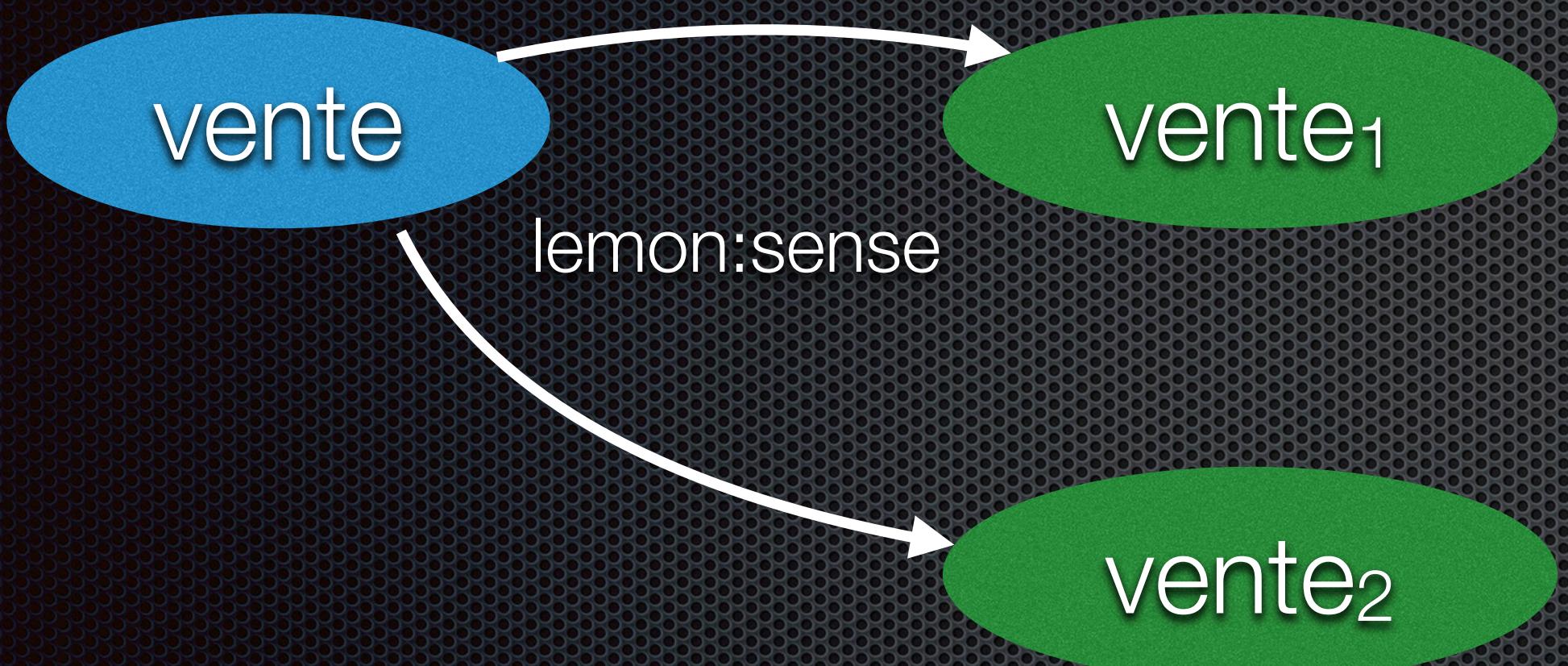
lemon:LexicalEntry
lemon:Word



Modellazione delle Locuzioni

- Rappresentate come individui, classificati secondo le rispettive tipologie

lemon:LexicalEntry
lemon:Word

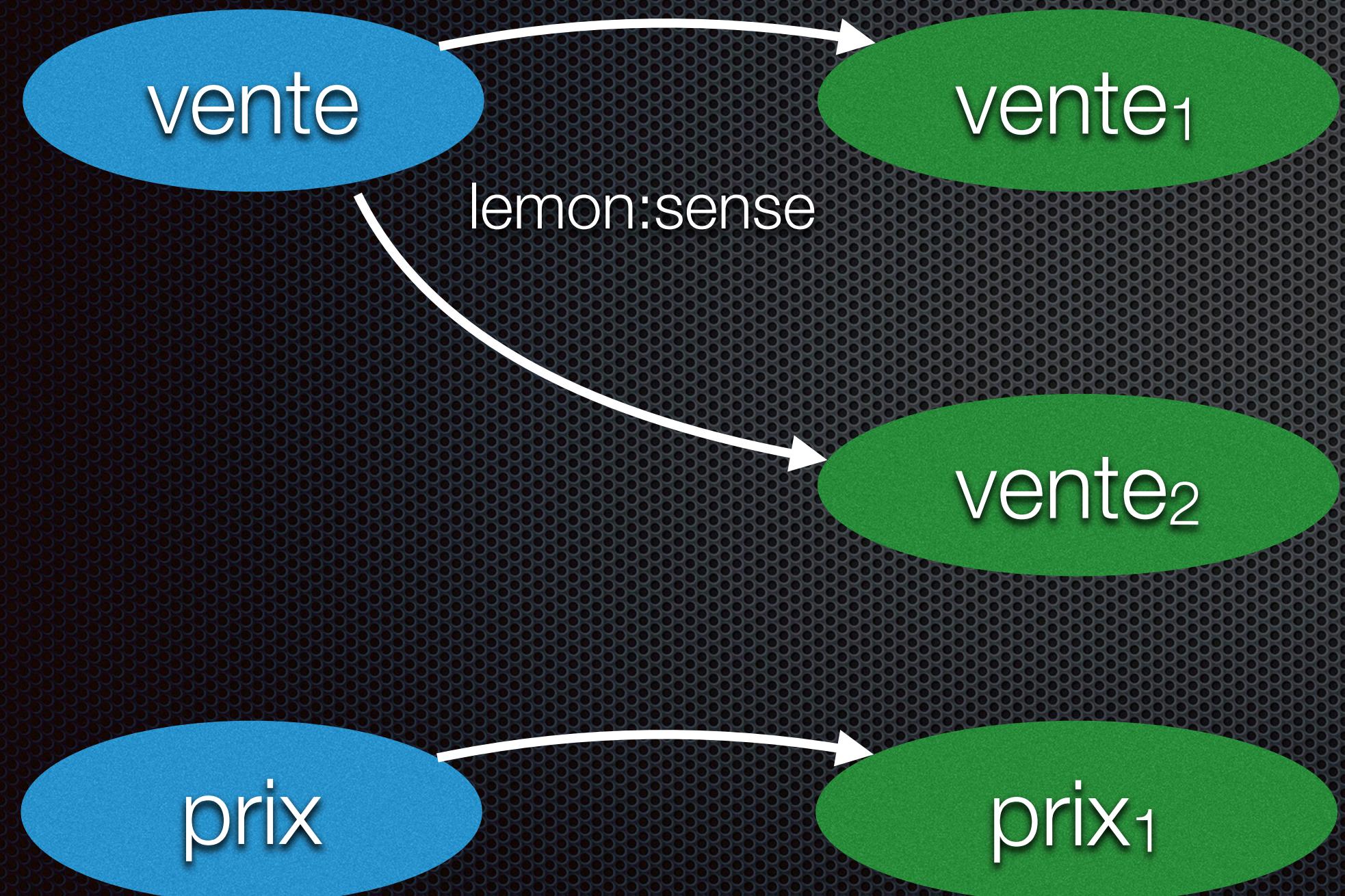


lemon:LexicalSense

Modellazione delle Locuzioni

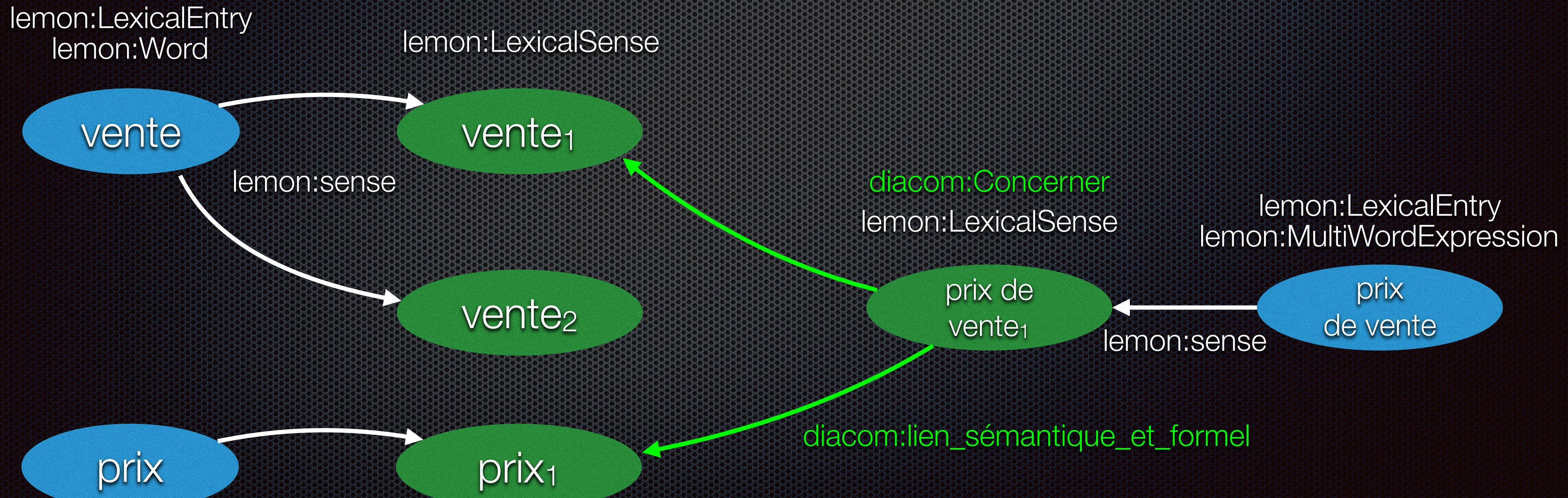
- Rappresentate come individui, classificati secondo le rispettive tipologie

lemon:LexicalEntry
lemon:Word

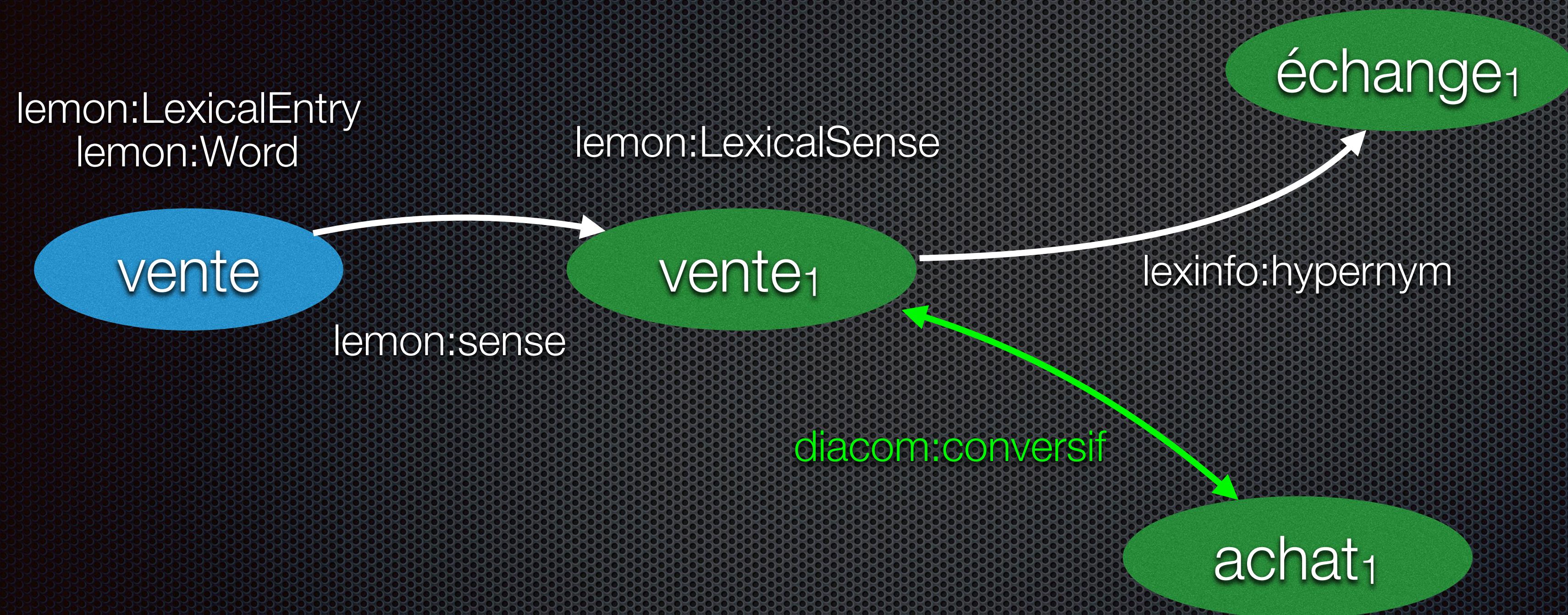


Modellazione delle Locuzioni

- Rappresentate come individui, classificati secondo le rispettive tipologie

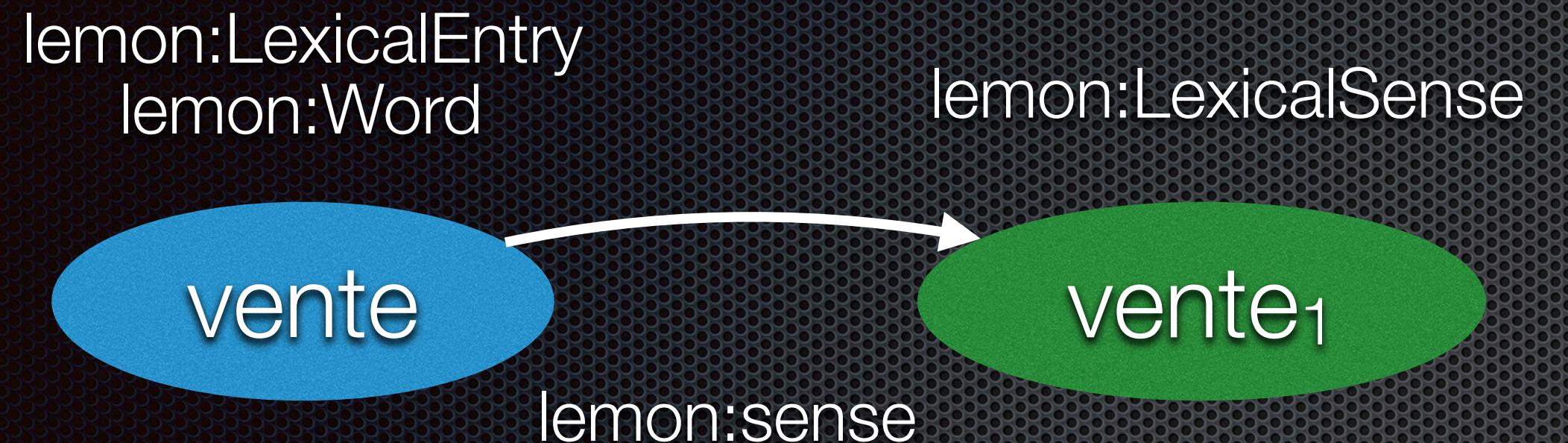


Modellazione dei legami paradigmatici



Modellazione dei legami paradigmatici

- Attanti semantici



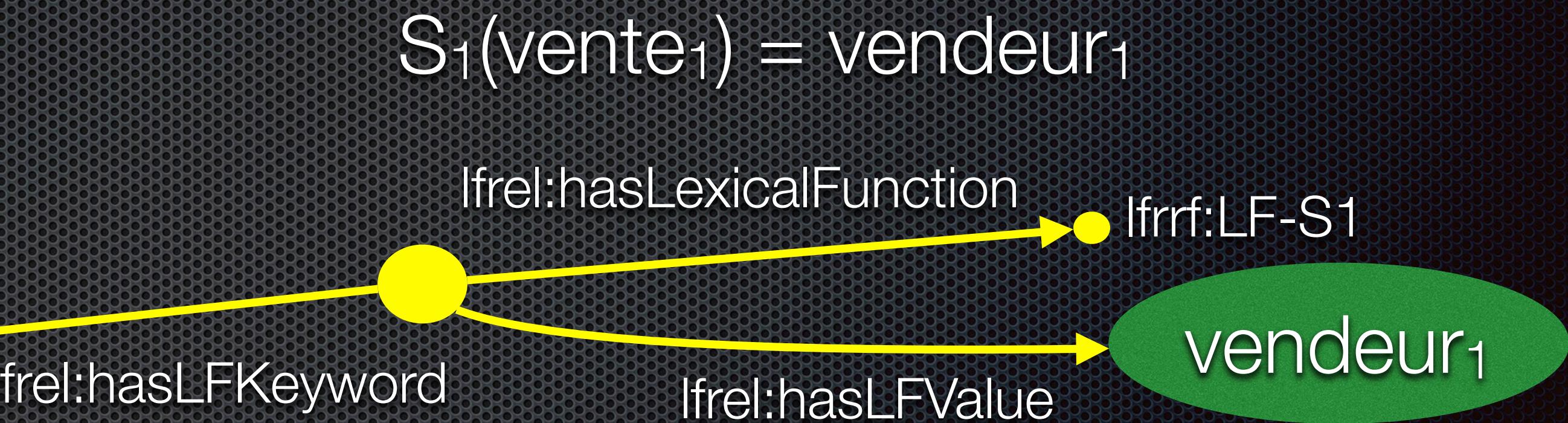
Modellazione dei legami paradigmatici

- Attanti semantici

lemon:LexicalEntry
lemon:Word



lemon:LexicalSense

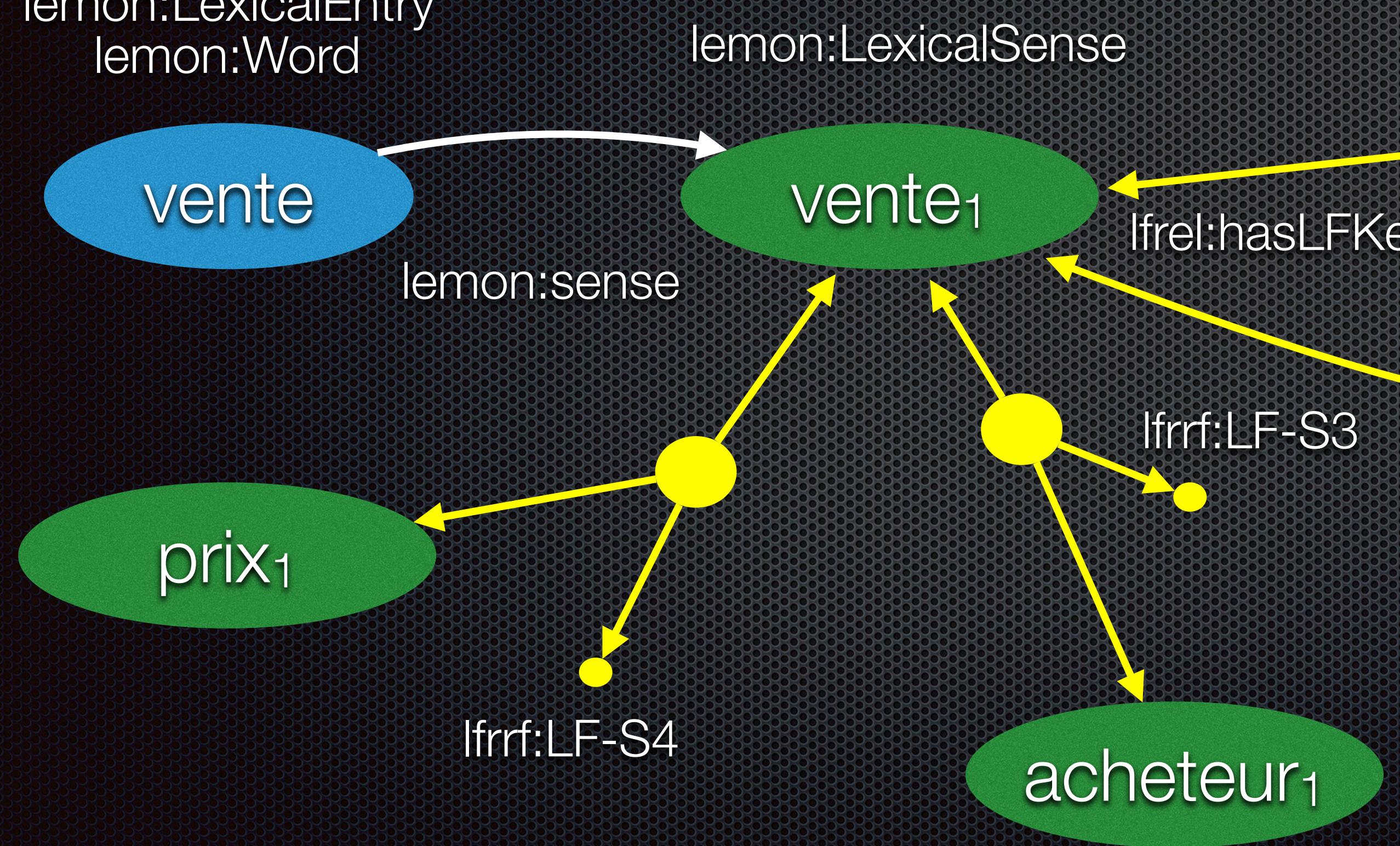


$S_1(vente_1) = \text{vendeur}_1$

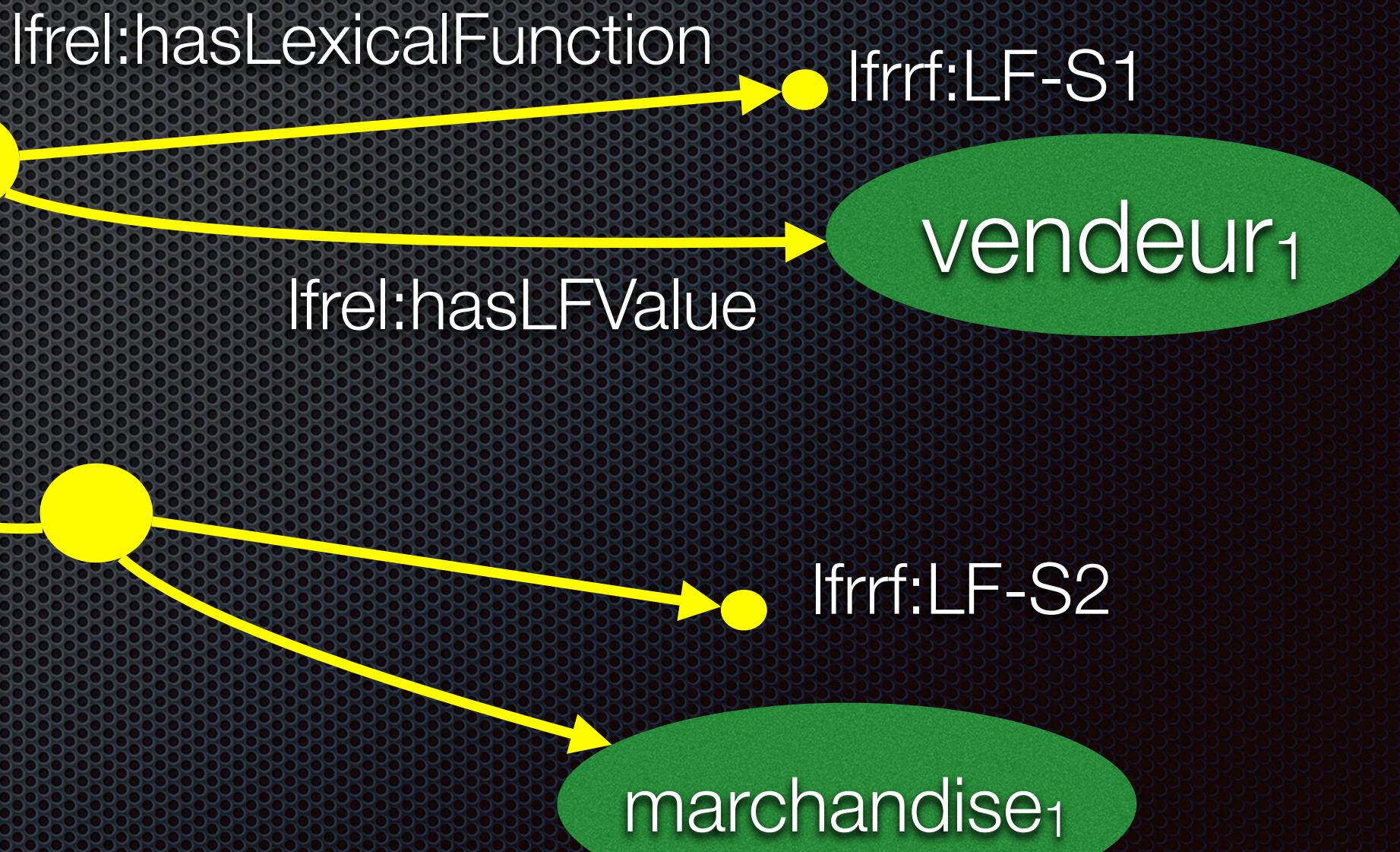
Modellazione dei legami paradigmatici

- Attanti semanticici

lemon:LexicalEntry
lemon:Word



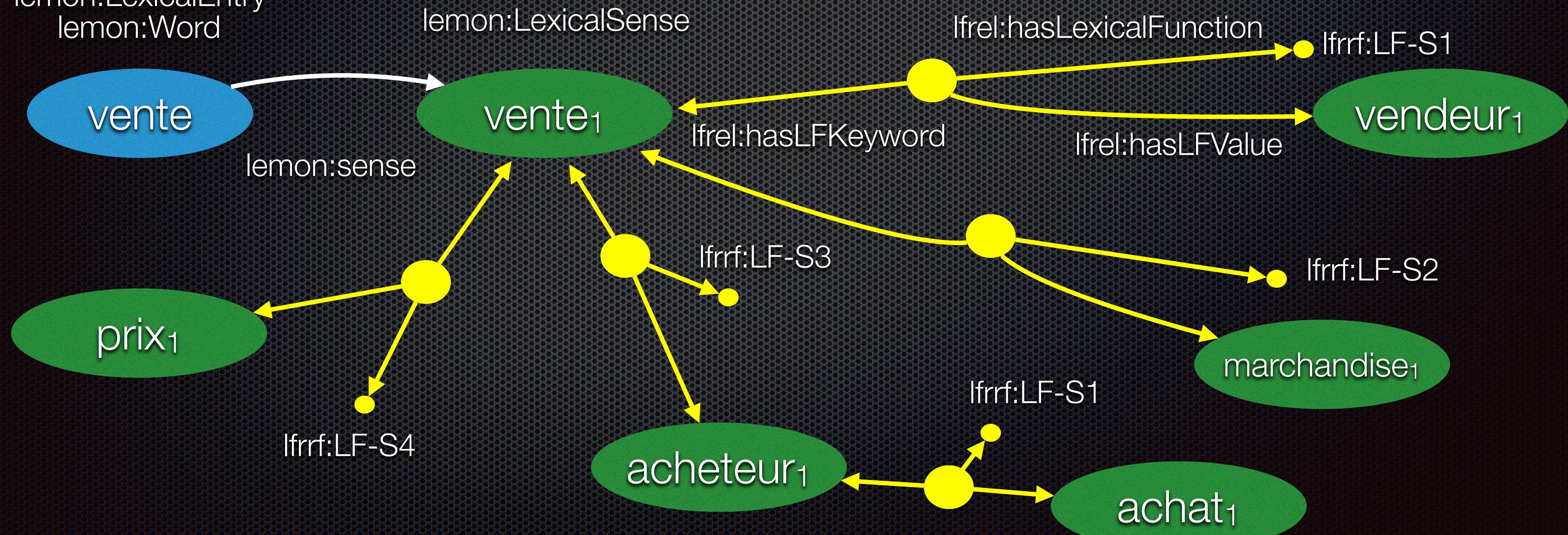
$S_1(vente_1) = \text{vendeur}_1$



Modellazione dei legami paradigmatici

- Attanti semantici

lemon:LexicalEntry
lemon:Word



$S_1(\text{vente}_1) = \text{vendeur}_1$

Esempio di modellazione

- ▼ ♦ **Prix de vente**
 - ♦ 'Prix de vente' signification **prix_de_vente_1**
- ▼ ♦ **prix_de_vente_1**
 - ♦ **prix_de_vente_1** 'lien sémantique et formel' **vente_1**
 - ♦ **Individual:** **prix_de_vente_1**
 - ♦ **prix_de_vente_1** exemple "Dans ce contexte de grande morosité, les hausses significatives des prix de vente ont permis de réaliser une année convenable"@fr
 - ♦ **prix_de_vente_1** **Type** Concerter
 - ♦ **prix_de_vente_1** 'lien sémantique et formel' **prix_1**
- ▼ ♦ **Vente**
 - ♦ Vente signification **vente_1**
- ▼ ♦ **vente_1**
 - ♦ **vente_1** exemple "La limitation de responsabilité qui précède est une condition de vente de la marchandise au prix indiqué ou aux prix indiqués et s'applique à la vente."@fr
 - ♦ **Individual:** **vente_1**
 - ♦ **vente_1** rdfs:comment "Échange de bien/marchandise contre argent (Vente par X de Y à Z pour la somme W)"@fr
 - ♦ **vente_1** gener échange_1
 - ♦ **vente_1** conversif achat_1
 - ♦ **vente_1** **Type** 'signification lexicale'
- ▼ ♦ **vente_1_s1_vendeur_1**
 - ♦ **vente_1_s1_vendeur_1** hasLFKeyword **vente_1**
- ▼ ♦ **vente_1_s2_marchandise_1**
 - ♦ **vente_1_s2_marchandise_1** hasLFKeyword **vente_1**
- ▼ ♦ **vente_1_s3_acheteur_1**
 - ♦ **vente_1_s3_acheteur_1** hasLFKeyword **vente_1**
- ▼ ♦ **vente_1_s4_prix_1**
 - ♦ **vente_1_s4_prix_1** hasLFKeyword **vente_1**



Conclusioni e Prossimi Passi

- Completamento dell'estrazione dei dati terminologici
- Rappresentazione di quanto estratto secondo il data model individuato
- Rilascio della risorsa secondo il paradigma dei Linked Open Data