Categorial Dependency Grammars

M. Dekhtyar
(Tver State University, Russia)
and
A. Dikovsky
(LINA CNRS FRE 2729, Université de Nantes)

Categorial Grammars
Grammaires Catégorielles
Montpellier, France
7-11 juin 2004

Abstract : *Categorial Dependency Grammars* are proposed as a model of word-driven (in contrast with head-driven) dependency syntax.

1. CDG : a subcommutative resource-sensitive calculus of dependency types,
2. enough expressive to naturally treat surface syntax phenomena including discontinuous constructions,
3. strongly more expressive than link grammars and many other projective DGs,
4. incomparable with multi-TAGs,
5. parsed in polynomial time (Earley-type parser),
6. disposing of an under-specified formal semantics,
7. learnable from strings, when rigid typed.

PLAN :

I. INTRODUCTION :
WORD-DRIVEN DEPENDENCIES

II. Categorial dependency grammars

III. Expressivity

IV. Complexity

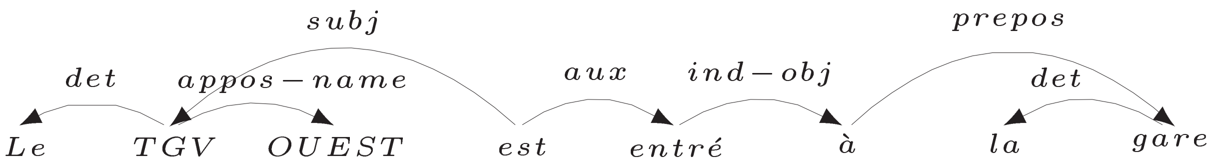
V. Problematic

I. Linguistic introduction

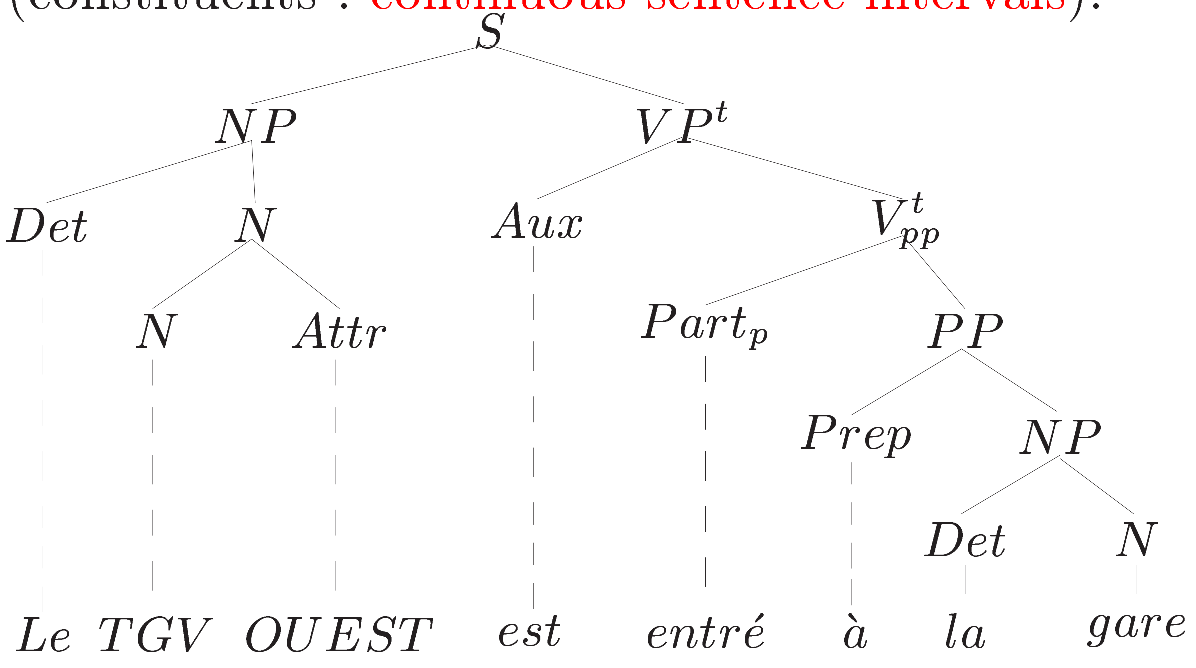
I.1 Syntactic dependencies

Two structural syntactic paradigms :

- dependency trees : (dependencies : irreflexive, anti-symmetric, anti-transitive binary relations on words).



- constituent (syntagmatic) trees : (constituents : continuous sentence intervals).



I.2. Head driven dependencies :
(Gladky'66, Robinson'70, Jackendoff'77)

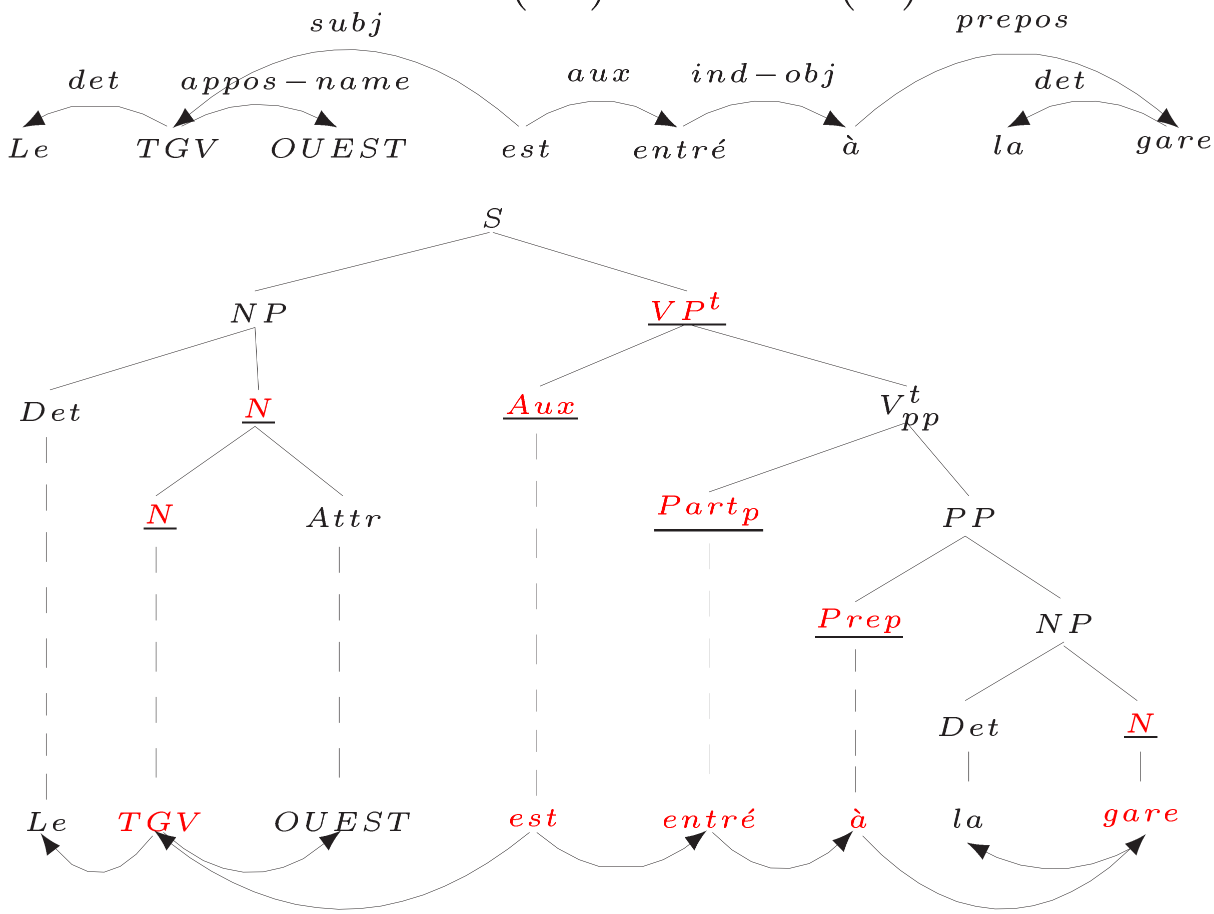
By-product of selection of a single head constituent in each non-unit constituent :

- a word w is the head of C :

$\forall C'(\{w\} \subseteq C' \subset C \Rightarrow selected(\underline{C'}))$

(C is projection of w).

- $C \subset C' \Rightarrow head(C') \Rightarrow^* head(C)$.

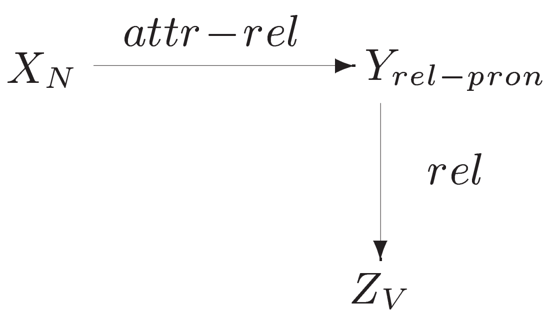


I.3 Word driven dependencies :

$$G\ (overnor) \xrightarrow{D} S\ (ubordinate) :$$

Binary relation D : ≪G licences S≫ verifying constraints on lexical and grammatical features, WO constraints, pronominalization constraints, etc. concerning only G and S.

EX1 : (Mel'čuk'88)



$$CO-GEND(X, Y)$$
$$CO-NUMB(X, Y)$$
$$CO-REF(X, Z.2)$$

EX2 :

$$est \xrightarrow{aux} entré$$

is an instance of :

$$aux : \begin{bmatrix} GOV & \begin{bmatrix} AUX \boxed{1} \\ NUM \\ GND \end{bmatrix} \\ SUB & \begin{bmatrix} V \boxed{2} & \\ MOD & participle \\ NUM & \\ GND & \end{bmatrix} \\ WO & precedes\{\boxed{1}, \boxed{2}\} \end{bmatrix} \text{ isa } dependency$$

Agreement :

$$\begin{bmatrix} aux & \\ GOV & [AUX[WORD\ \hat{e}tre]] \end{bmatrix} \Rightarrow \begin{bmatrix} aux & \\ GOV & [NUM[SUB\ NUM]] \\ GOV & [GND[SUB\ GND]] \end{bmatrix}$$

EX3 : Cf. (in the sentence *Le TGV …*) :

$$entr\acute{e} \overset{ind-obj}{\longrightarrow} \grave{a}$$

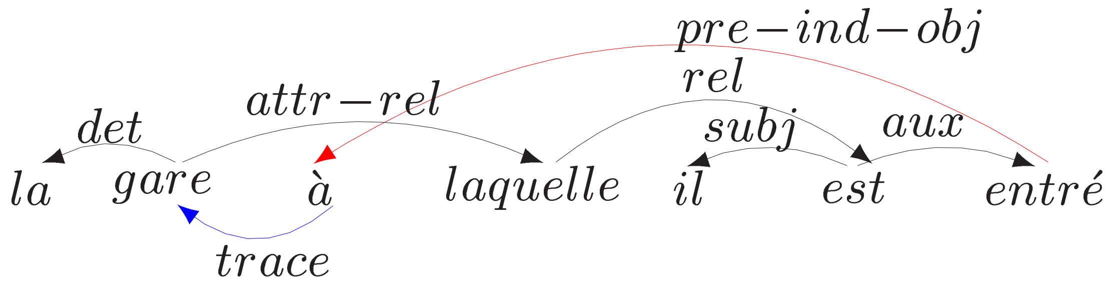
$$ind\text{–}obj : \left[\begin{array}{ll} GOV & \left[\begin{array}{l} V \\ VAL\ \{\boxed{1},\boxed{2}\} \\ SYN\ [\neg modal, \neg copul] \end{array}\right] \\ SUB & [WORD\ \boxed{3}\ preposition] \\ WO & precedes\{\boxed{1},\boxed{3}\} \end{array}\right] \text{ isa } dependency$$

But also :

$$\grave{a} \overset{pre-ind-obj}{\longleftarrow} entr\acute{e}$$

$$pre\text{–}ind\text{–}obj : \left[\begin{array}{ll} GOV & \left[\begin{array}{l} V \\ VAL\ \{\boxed{1},\boxed{2}\} \\ SYN\ [\neg modal, \neg copul] \end{array}\right] \\ SUB & [WORD\ \boxed{3}\ preposition] \\ WO & precedes\{\boxed{3},\boxed{1}\} \end{array}\right] \text{ isa } dependency$$

as in the sentence :

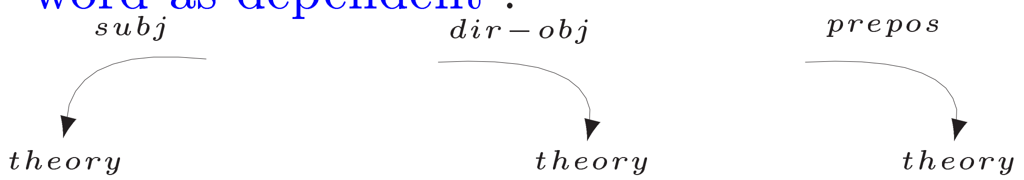


I.4. Types of word dependencies

Primitive types : dependency names (!!) (not constituent / constituent complement classes).

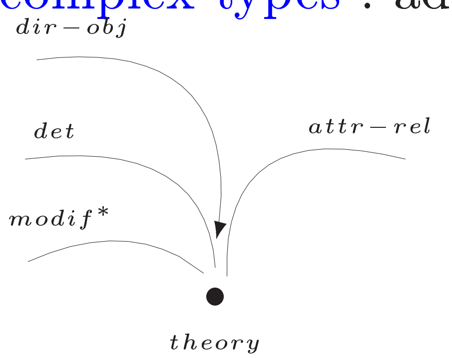
Complex types determine : oriented local valencies, polarized discontinuous valencies, order, adjacency :

- word as dependent :



- word as governor :

- complex types : admissible combinations :



$[modif^* \backslash det \backslash dir-obj / attr-rel]$

I.5. Definition of syntactic types (categories)

Polarities of distant dependencies :

Negative right : $\swarrow pre-indir-obj$: (expects right governor through the dependency $pre-indir-obj$),

Positive right : $\nearrow neg-compos$: (expects right subordinate through the dependency $neg-compos$).

Dual ($\breve{C}$) : ($\swarrow pre-indir-obj$) et ($\nwarrow pre-indir-obj$), ($\nearrow neg-compos$) et ($\searrow neg-compos$).

Negative anchored (on a category) : $\#(\swarrow clit-obj)$: clitic (e.g. anchored on a verb with category $[\#(\swarrow clit-obj)\backslash S/aux]$).

Negative not anchored are loose.

$\mathbf{C}$ being a set of primitive categories,
$V^+(\mathbf{C}) =_{df} \{(\nwarrow C), (\nearrow C) \mid C \in \mathbf{C}\}$: positive,
$V^-(\mathbf{C}) =_{df} \{(\swarrow C), (\searrow C) \mid C \in \mathbf{C}\}$: negative,
$Anc(\mathbf{C}) =_{df} \{\#(\alpha) \mid \alpha \in V^-(\mathbf{C})\}$: anchored,
$\mathbf{C}^* =_{df} \{C^* \mid C \in \mathbf{C}\}$: iterated.

Definition of complex categories (1^{st} order) :

$Cat(\mathbf{C})$: minimal set such that :

1. $\mathbf{C} \cup V^{-}(\mathbf{C}) \cup Anc(\mathbf{C}) \subset Cat(\mathbf{C})$.

2. For $C \in Cat(\mathbf{C})$,
$A_1 \in (\mathbf{C} \cup \mathbf{C}^* \cup Anc(\mathbf{C}) \cup \nwarrow \mathbf{C})$ et
$A_2 \in (\mathbf{C} \cup \mathbf{C}^* \cup Anc(\mathbf{C}) \cup \nearrow \mathbf{C})$

$[A_1 \backslash C], [C / A_2] \in Cat(\mathbf{C})$.

Constructors $\backslash, /$ are associative. So each category has the form :

$$[L_k \backslash \ldots L_1 \backslash C / R_1 \ldots / R_m].$$

EX : $[\#(\swarrow clit\!-\!dobj) \backslash subj \backslash S / aux]$

is one of the categories of the auxiliary *être*, which anchors a clitic and expects a subordinate subject on its left and expects a subordinate participle on its right (def. of aux).

II. Categorial dependency grammar (CDG)

$$G = (W, \mathbf{C}, S, \delta),$$

where W : words, $\mathbf{C}$: categories, $S \in \mathbf{C}$,
δ (lexicon) : finite substitution $W \to 2^{Cat(\mathbf{C})}$.

> EX 4 : $les \mapsto det$,
> $bordellais \mapsto modif$,
> $excellents \mapsto cop{-}adj$,
> $vins \mapsto [det\backslash subj/modif^*]$,
> $sont \mapsto [subj\backslash S/cop{-}adj^*]$

D-form (Δ, Γ) of a sentence $w = a_1 ... a_n$:

- Δ : a dependency graph (DG) on w,

- Γ : a string of positioned categories of DGs,
α^i : a category of a DG on w with the root a_i,

- $((V, \emptyset), C_1^1 \ldots C_n^n)$: initial D-form, if
$C_i \in \delta(a_i)$, $1 \leq i \leq n$,

> **EX** : $det^1[det\backslash subj/modif^*]^2 modif^3 [subj\backslash S/cop{-}adj^*]^4\ cop{-}adj^5$

- (Δ, S^j) : terminal D-form.

II.1. Sub-commutative dependency calculus

$\mathbf{L^l}$. $((V,E),\Gamma_1 C^j[C\backslash\beta]^l\Gamma_2)\vdash$
$((V,E\ \cup\{a_j \overset{C}{\longleftarrow}\ a_l\}),\Gamma_1[\beta]^l\Gamma_2),\ \ C\in\mathbf{C}.$

$\mathbf{I^l}$. $((V,E),\Gamma_1 C^j[C^*\backslash\beta]^l\Gamma_2)\vdash$
$((V,E\ \cup\{a_j \overset{C}{\longleftarrow}\ a_l\}),\Gamma_1[C^*\backslash\beta]^l\Gamma_2),\ \ C\in\mathbf{C}.$

$\mathbf{\Omega^l}$. $((V,E),\Gamma_1[C^*\backslash\beta]^l\Gamma_2)\vdash((V,E),\Gamma_1\beta^l\Gamma_2).$

$\mathbf{A^l}$. $((V,E),\Gamma_1\#(\alpha)^j[\#(\alpha)\backslash\beta]^l\Gamma_2)\vdash$
$((V,E),\Gamma_1\alpha^j\beta^l\Gamma_2),\quad \#(\alpha)\in Anc(\mathbf{C}).$

$\mathbf{P^l}$. $((V,E),\Gamma_1[(\nwarrow C)\backslash\alpha]^j\Gamma_2)\vdash$
$((V,E),\Gamma_1(\nwarrow C)^j\alpha^j\Gamma_2).$

$\mathbf{C^l}$. $((V,E),\Gamma_1\ C^l\alpha^j\ \Gamma_2)\vdash((V,E),\Gamma_1\ \alpha^j C^l\ \Gamma_2)$
if $\alpha\in(\nwarrow\mathbf{C}\cup\searrow\mathbf{C})$ and $C\in Cat(\mathbf{C})$ without occurrences of $\alpha,\#(\alpha)$ and $\breve{\alpha}$.

$\mathbf{D^l}$. $((V,E),\Gamma_1(\swarrow C)^j(\nwarrow C)^l\Gamma_2)\vdash$
$((V,E\ \cup\{a_j \overset{C}{\longleftarrow}\ a_l\}),\Gamma_1\Gamma_2).$

II.2. EX 5 : Example of a derivation :

$les \mapsto det,$

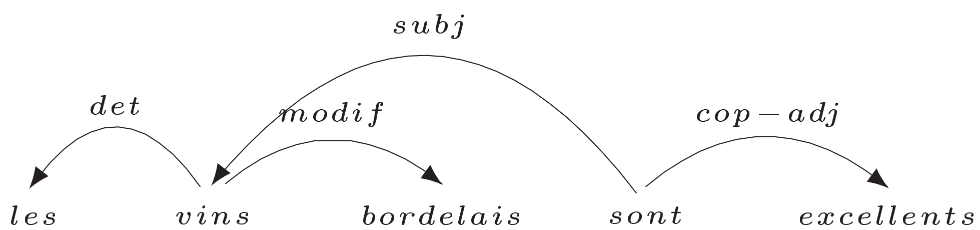
$bordellais \mapsto modif,$

$excellents \mapsto cop{-}adj,$

$vins \mapsto [det\backslash subj/modif^*],$

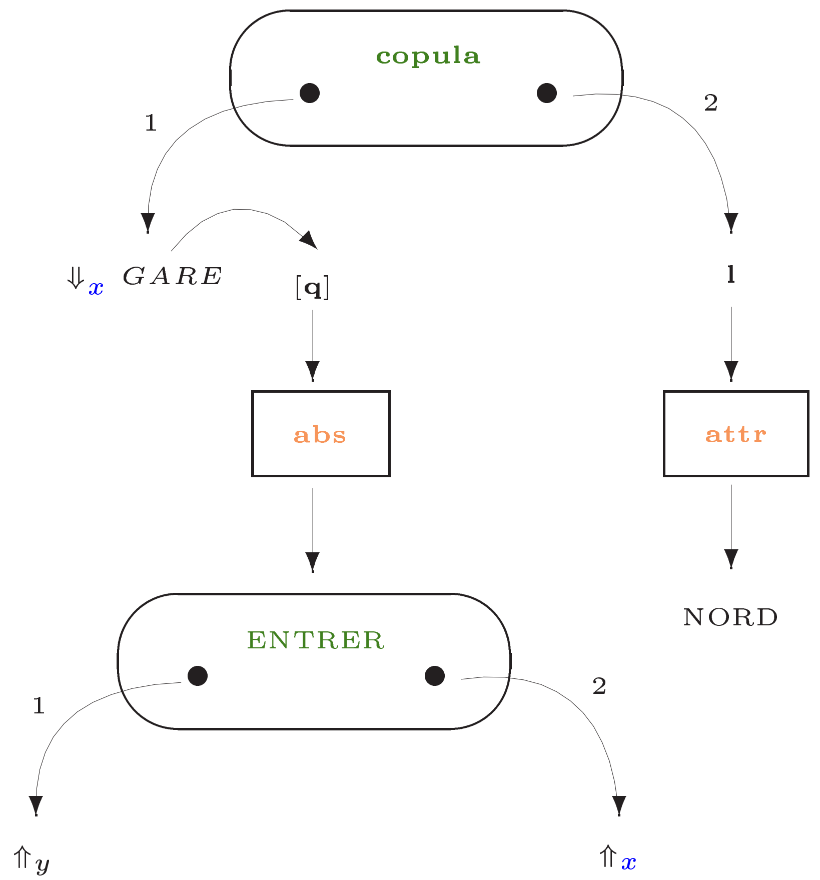
$sont \mapsto [subj\backslash S/cop{-}adj^*]$

$det^1[det\backslash subj/modif^*]^2 modif^3[subj\backslash S/cop{-}adj^*]^4$
$cop{-}adj^5 \vdash$
$det^1[det\backslash subj/modif^*]^2[subj\backslash S/cop{-}adj^*]^4 cop{-}adj^5 \vdash$
$det^1[det\backslash subj/modif^*]^2[subj\backslash S/cop{-}adj^*]^4 \vdash$
$det^1[det\backslash subj]^2[subj\backslash S/cop{-}adj^*]^4 \vdash$
$det^1[det\backslash subj]^2[subj\backslash S]^4 \vdash^* S^4$



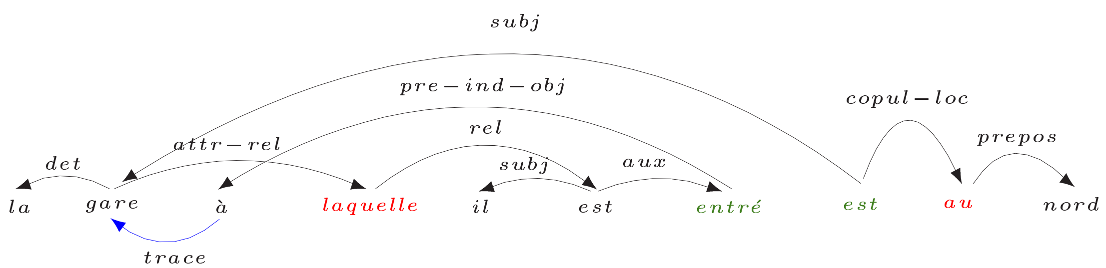
II.3. Types are justified by situational meaning

EX 6 :



See [A. Dikovsky, Formal Grammar'2003.]

The corresponding dependency structure :



II.4. Generated languages :

$G_C(D,w)$: $(\Delta_0,\Gamma_0) \vdash_C^* (D, S^j)$, for an initial D-form (Δ_0,Γ_0) of w and some j.
$(C \in \{sc, FA\})$.

D-language of G :

$D_C(G) =_{df} \{D \mid \exists w \; G_C(D,w)\}$.

Language of G :

$L_C(G) =_{df} \{w \mid \exists D \; G_C(D,w)\}$.

Families :

$$\mathcal{D}(CDG^{sc}), \; \mathcal{L}(CDG^{sc}).$$

III. Expressivity

III.1. Projectivity / non-projectivity

Head-driven DTs are projective. Word-driven DTs may be non-projective :

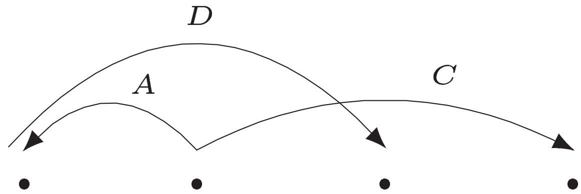
EX 7. Non-projective DTs :

$[A/(\nearrow D)]^1 [A\backslash S/C/\#(\searrow D)]^2 \#(\searrow D)^3 C^4 \vdash$

$[A/(\nearrow D)]^1 [A\backslash S/C]^2 (\searrow D)^3 C^4 \vdash^*$

$A^1 [A\backslash S/C]^2 (\nearrow D)^1 (\searrow D)^3 C^4 \vdash$

$A^1 [A\backslash S/C]^2 C^4 \vdash^* S^2$



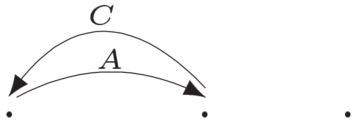
III.2. Cycles in dependency structures :

EX 8 :

$[(\swarrow C)/(\nearrow A)]^1 [(\nwarrow C)\backslash(\searrow A)]^2 S^3 \vdash$

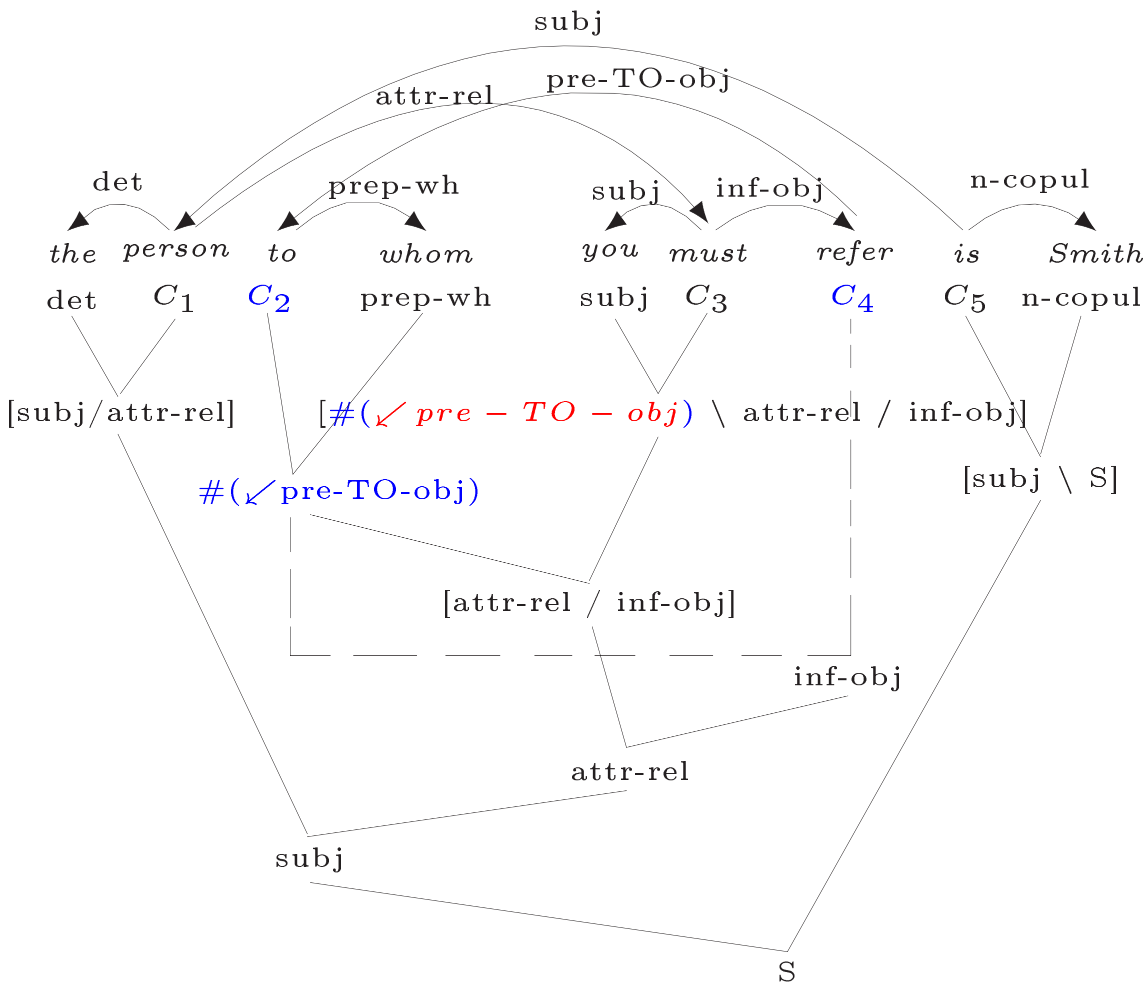
$(\swarrow C)^1 (\nearrow A)^1 (\nwarrow C)^2 (\searrow A)^2 S^3 \vdash$

$(\swarrow C)^1 (\nwarrow C)^2 (\nearrow A)^1 (\searrow A)^2 S^3 \vdash^* S^3$



III. Relative clauses

EX 9 :



$person \mapsto C_1 = [det \backslash subj/attr-rel]$,

$to \mapsto C_2 = [\#(\swarrow pre-TO-obj)/prep-wh]$,

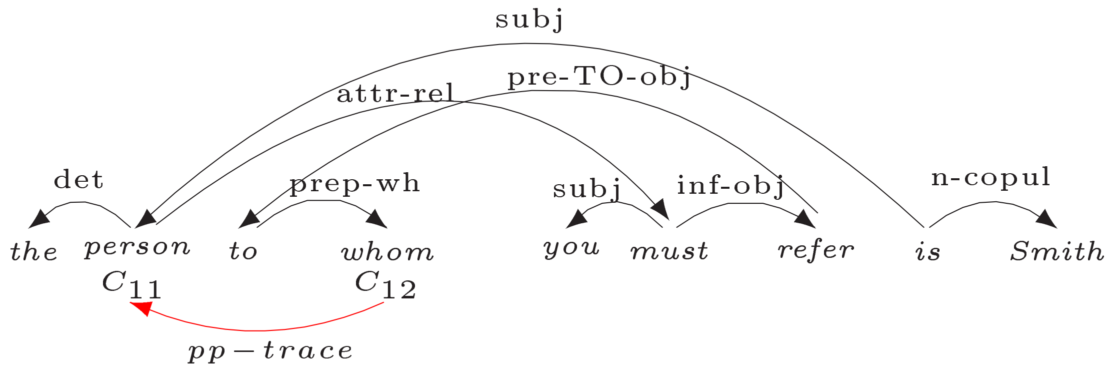
$must \mapsto C_3 = [subj \backslash \#(\swarrow pre-TO-obj) \backslash attr-rel/inf-obj]$,

$refer \mapsto C_4 = [(\nwarrow pre-TO-obj) \backslash inf-obj]$,

$is \mapsto C_5 = [subj \backslash S/n-copul]$

III.1. Co-reference through cycles

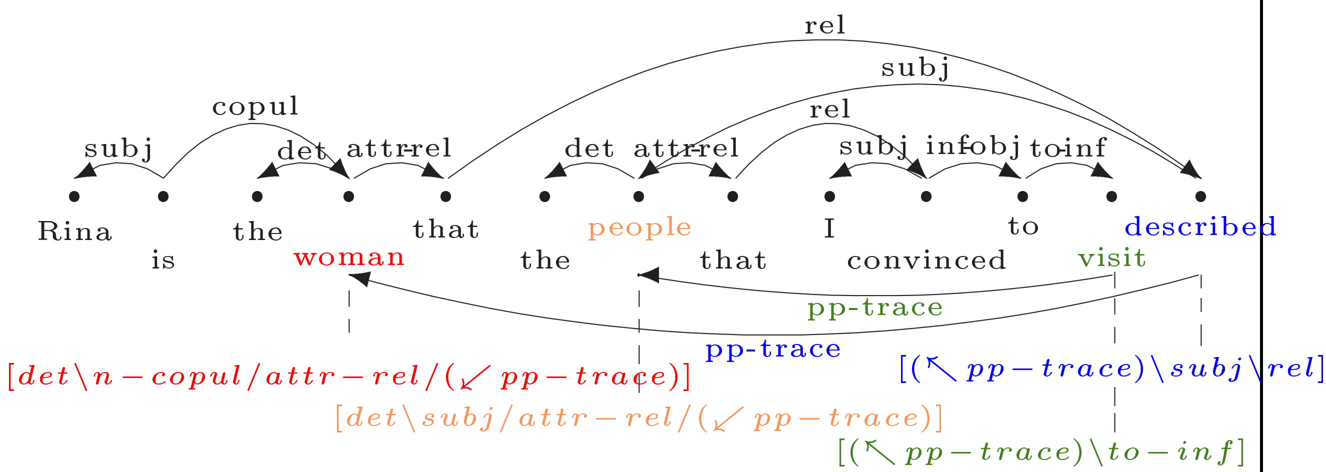
EX 10 :



$$C_{11} = [det\backslash subj/attr{-}rel/(\swarrow pp{-}trace)],$$
$$C_{12} = [(\nwarrow pp{-}trace)\backslash prep{-}wh]$$

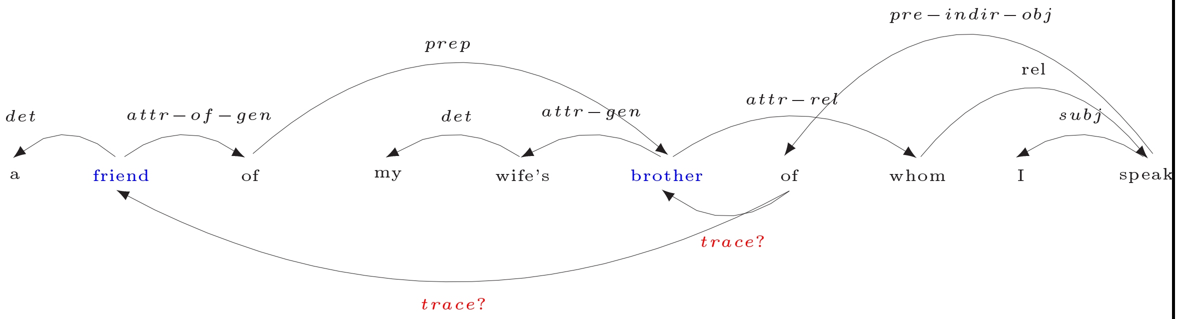
III.2. Pied-piping : still working

EX 11 (N. Vailette) :

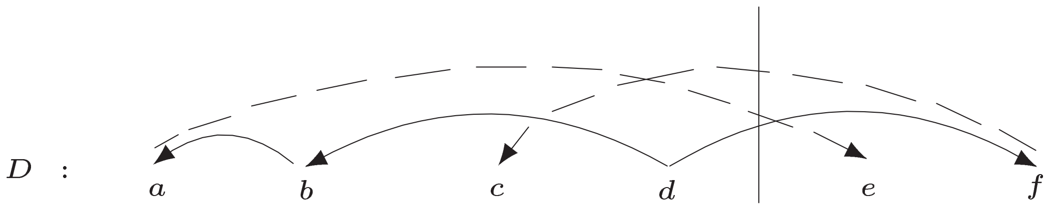


III.3. Where it is not working :

EX 12



III.4. Discontinuous dependencies' thickness (DDT)



$EPD(D) = 2\ (= max)$

TH If $DDT(G)$ is finite for $G \in GCD^{cs}$, then $L(G) \in \mathcal{L}(CF)$.

Note : In natural languages, DDT is bounded by a small constant (2-3).

III.5. CDG vs. CFG :

$G_0 = (\{a, b, c, d_1, d_2, d_3\}, \mathbf{C}_0, S, \delta_0) \in GCD^{FA}$:

$a \mapsto [\beta \backslash \alpha], [\alpha \backslash \alpha], \quad d_1 \mapsto \alpha,$

$b \mapsto [\alpha_1 \backslash D/A], \quad d_2 \mapsto [\alpha \backslash \beta_1 \backslash S/D],$

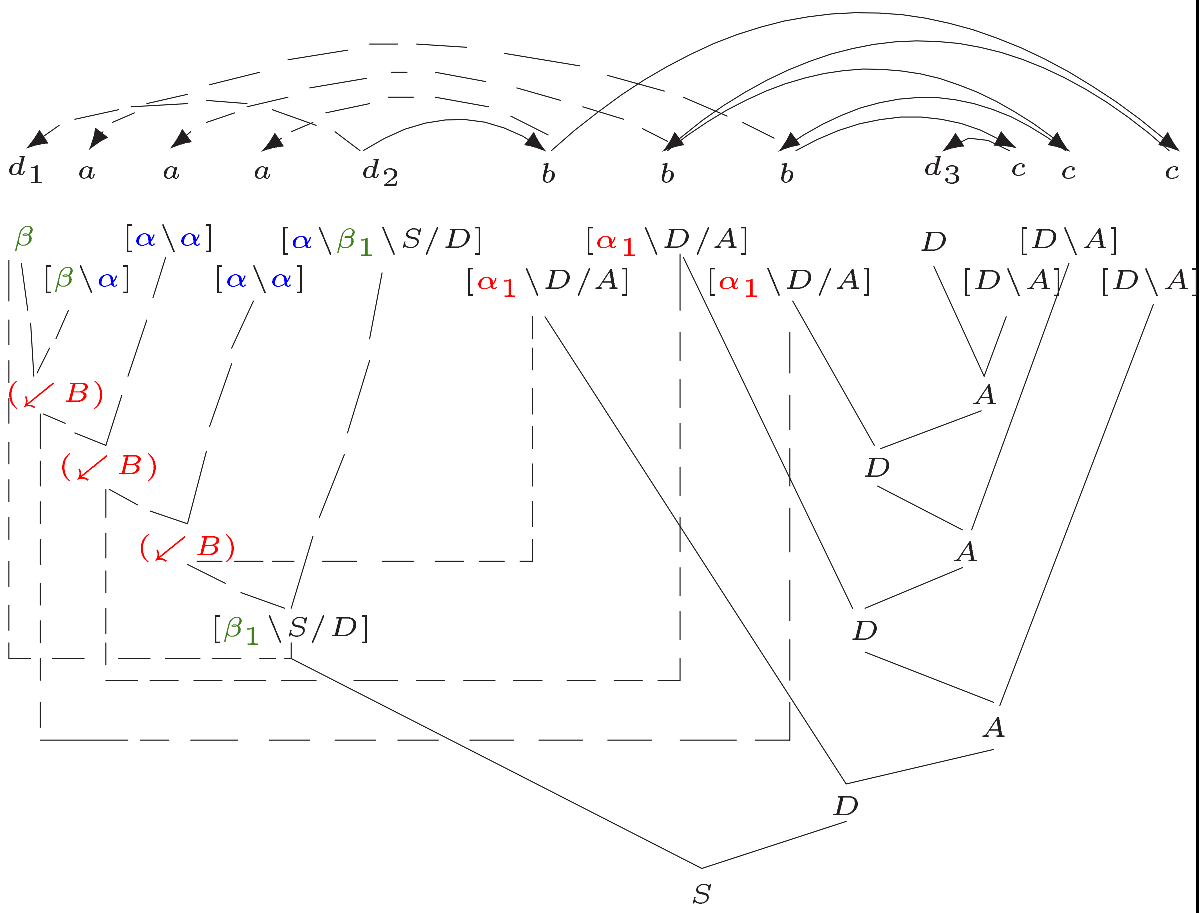
$c \mapsto [D \backslash A], \quad d_3 \mapsto D,$

$\alpha = \#(\swarrow B)$, $\alpha_1 = (\nwarrow B)$, $\beta = \#(\swarrow C)$,

$\beta_1 = (\nwarrow C)$

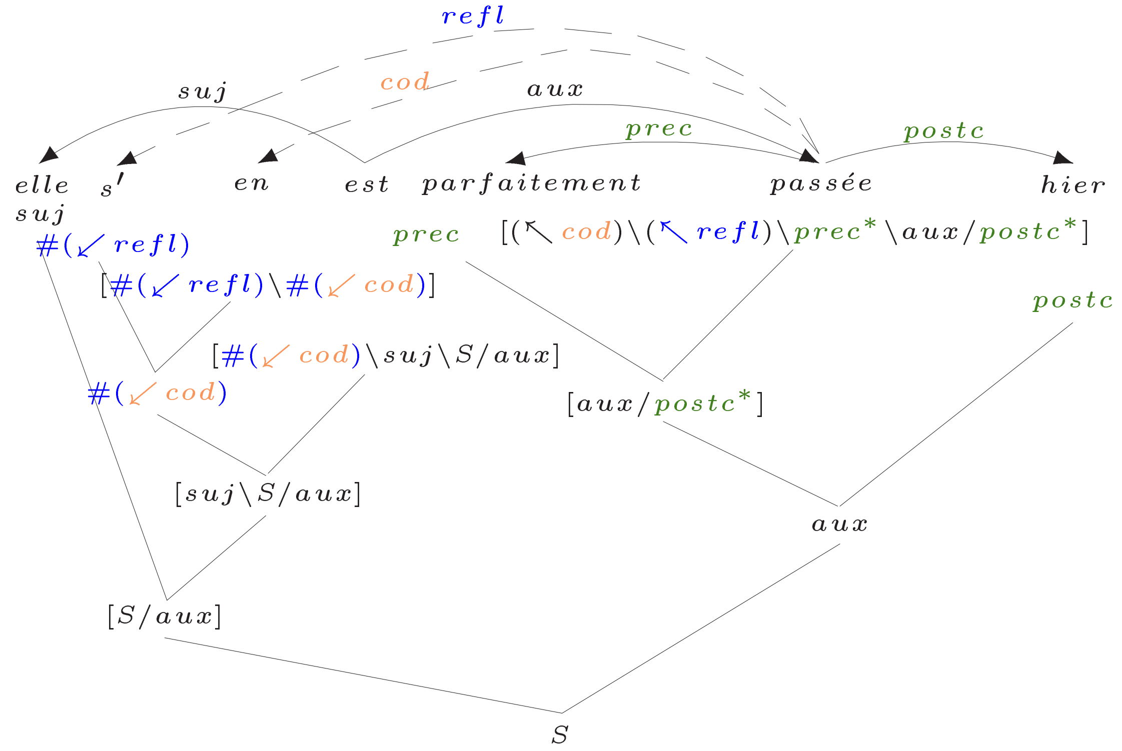
generates $L(G_0) = \{d_1 a^n d_2 b^n d_3 c^n \mid n > 0\}$.

A proof for $d_1 a^3 d_2 b^3 d_3 c^3$:



TH $\mathcal{L}(CF) \subsetneq \mathcal{L}(CDG)$.

EX 13 : Clitics through anchored types

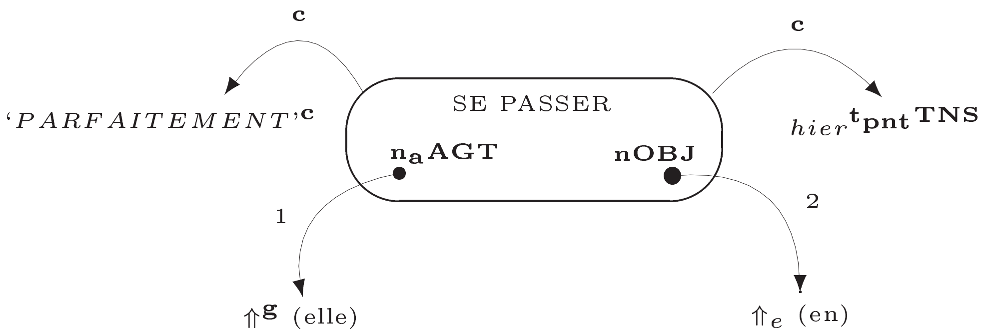


Types :

$\hat{e}tre \mapsto [\#(\swarrow cod)\backslash subj\backslash S/aux]$

$en \mapsto [\#(\swarrow refl)\backslash\#(\swarrow cod)]$

$se \mapsto \#(\swarrow refl)$



III.6. CDG vs. TAG

TH For all $m > 4$, the language

$\{d_0 a_0^n d_1 a_1^n \dots d_m a_m^n d_{m+1} | n \geq 0\} \notin \mathcal{L}(TAG)$

is generated by the CDG $G_{(m)}$:

$d_0 \mapsto [S/D_0]$ et $d_{m+1} \mapsto D_m$,

$a_0 \mapsto [D_0/D_0/(\nearrow A_m)/\dots/(\nearrow A_1)]$,

$d_i \mapsto [D_{i-1}/\#(\searrow A_i)]$,

$a_i \mapsto [\#(\searrow A_i)/\#(\searrow A_i)], [\#(\searrow A_i)/D_i]$

$(0 < i \leq m)$.

Hypothesis : The copy language

$$\{ww \mid w \in \{a, b\}^+\} \in \mathcal{L}(TAG)$$

is not generated by CDG.

III.7. CDG vs. multi-TAG

$$3_{mix} =_{df} \{w \in \{a, b, c\}^+ \mid |w|_a = |w|_b = |w|_c\}$$

Hypothesis : $3_{mix} \notin \mathcal{L}(multi - TAG)$.

TH (together with A. Foret) $3_{mix} \in \mathcal{L}(GCD)$.

IV. Complexity

TH Membership problem

$$\{(G, w) \mid G \in GCD^{sc}, w \in W^+\}$$

is NP-complete.

TH Parsing problem is polynomial ($\mathbf{O}(\mathsf{n}^4)$ if DDT is finite, $\mathbf{O}(\mathsf{n}^3)$ if G is projective).

Independent projections criterion :

Local projection : $\|\gamma\|_l$. Let $\gamma \in Cat(\mathbf{C})^*$.

l1. $\|\varepsilon\|_l = \varepsilon$; $\|C\gamma\|_l = \|C\|_l\|\gamma\|_l$ for $C \in Cat(\mathbf{C})$ and $\gamma \in Cat(\mathbf{C})^*$.

l2. $\|C\|_l = C$ for $C \in \mathbf{C} \cup \mathbf{C}^* \cup Anc(\mathbf{C})$.

l3. $\|C\|_l = \varepsilon$ for $C \in V^+(\mathbf{C}) \cup V^-(\mathbf{C})$.

l4. $\|[a\backslash\alpha]\|_l = [a\backslash\|\alpha\|_l]$ and $\|[\alpha/a]\|_l = [\|\alpha\|_l/a]$ for $a \in \mathbf{C} \cup \mathbf{C}^* \cup Anc(\mathbf{C})$ and $\alpha \in Cat(\mathbf{C})$.

l5. $\|[(\nwarrow a)\backslash\alpha]\|_l = \|[\alpha/(\nearrow a)]\|_l = \|\alpha\|_l$ for all $a \in \mathbf{C}$ and $\alpha \in Cat(\mathbf{C})$.

Valencies' projection : $\|\gamma\|_v$

v1. $\|\varepsilon\|_v = \varepsilon$; $\|C\gamma\|_v = \|C\|_v\|\gamma\|_v$ for $C \in Cat(\mathbf{C})$ and $\gamma \in Cat(\mathbf{C})^*$.

v2. $\|C\|_v = \varepsilon$ for $C \in \mathbf{C} \cup \mathbf{C}^*$.

v3. $\|C\|_v = C$ for $C \in V^+(\mathbf{C}) \cup V^-(\mathbf{C})$.

v4. $\|\#(C)\|_v = C$ for $C \in V^-(\mathbf{C})$.

v5. $\|[\alpha]\|_v = \|\alpha\|_v$ for all $[\alpha] \in Cat(\mathbf{C})$.

v6. $\|\#(C)\backslash\alpha\|_v = \|\alpha/\#(C)\|_v = \|\alpha\|_v$.

v7. $\|\alpha_1\backslash\alpha_2\|_v = \|\alpha_1\|_v\|\alpha_2\|_v$, if $\alpha_1 \notin Anc(\mathbf{C})$.

v8. $\|\alpha_1/\alpha_2\|_v = \|\alpha_1\|_v\|\alpha_2\|_v$, if $\alpha_2 \notin Anc(\mathbf{C})$.

TH Let $G = (W, \mathbf{C}, S, \delta)$ be a CDG.
$x \in L(G)$ iff there is a string of categories $\alpha \in \delta(x)$ and its realization $\gamma \in R(\alpha)$ such that :

1. $\|\gamma\|_l \vdash_p^* S$,
2. there is a correct pairing of dual categories $(\alpha, \breve{\alpha})$ in $\|\gamma\|_v$.

For $C = [\alpha D^*\backslash\beta]$, $[\alpha\beta]$, $[\alpha D\backslash\beta]$, $[\alpha D\backslash D\backslash\beta]$, etc. are its *realizations*.

$\vdash_p^*$: elimination-rule (**L**) provability.

V. Problematic :

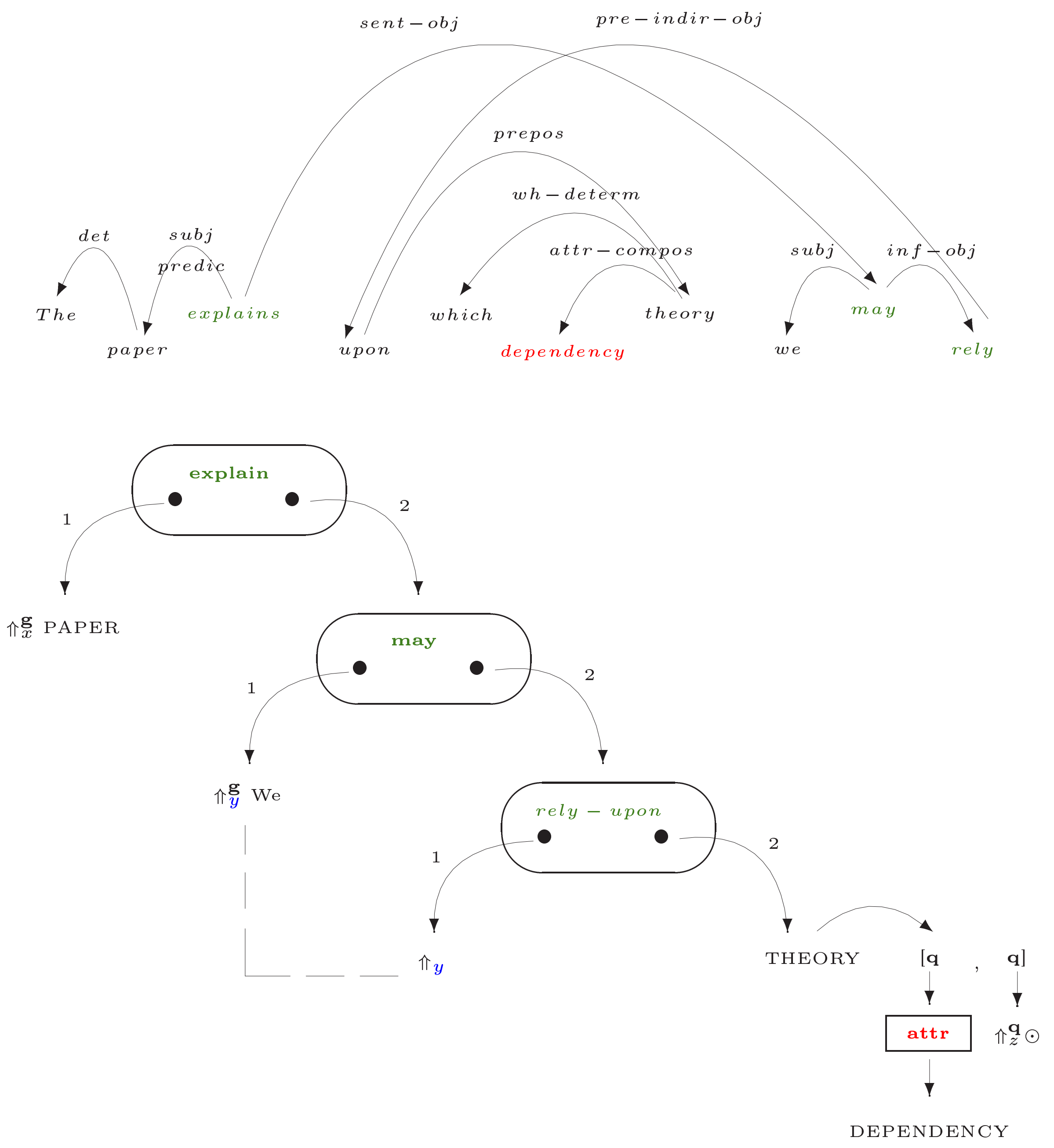
1. Translation of constraint dependency languages into CDG.

2. Sound type translation expansion to complete resource-logic semantics via situational semantics.

3. Robust (linear-time) incremental parser with a stochastic oracle.

4. Modular generation of CDG types from situational meanings.

EX 10 :


sent − obj
pre − indir − obj
prepos
wh − determ
det
subj
predic
attr − compos
subj
inf − obj
The
paper
explains
upon
which
dependency
theory
we
may
rely
explain
1
2
⇑ PAPER
may
1
2
We
rely − upon
1
2
THEORY
[q , q]
attr
DEPENDENCY

III.2. Comment éviter les cycles

1. Pour le calcul non-commutatif $\mathbf{CD^{FA}}$:

TH $G \in CDG^{FA} \wedge D \in D(G) \ \Rightarrow \ (D\ est\ AD)$.

2. Pour $\mathbf{CD^{sc}}$: contraintes sur les types.

GCD bien fondée :

$\gamma_1 \rightsquigarrow_l \gamma_2$: γ_2 "est dominée par" γ_1 (à travers les dépendances locales ou distantes) ;

$\gamma_1 \hookrightarrow \gamma_2$: $[\alpha\backslash\gamma_1/\beta_1\gamma'\beta_2] \in \delta(W)$ et $\gamma' \rightsquigarrow \gamma_2$.

BF : Il y a un préordre $\gtrsim$ sur Cat :

$\mathbf{w_1}$. If $\alpha_1 \rightsquigarrow \alpha_2$, then $\alpha_1 \gtrsim \alpha_2$.
$\mathbf{w_2}$. If $(\swarrow A) \hookrightarrow (\nearrow B) \rightsquigarrow \alpha$ and $(\nwarrow A) \hookleftarrow \alpha$, then $(\swarrow A) > (\nearrow B)$.

TH Si $G \in CDG^{sc}$ est WF et $D \in D(G)$, alors D est AD.