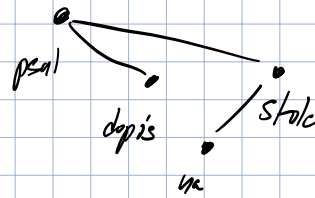


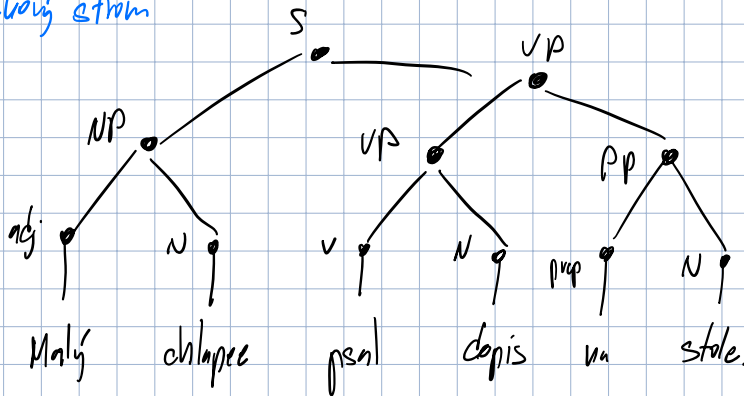
Syntaktická analýza

Závislostní strom

- reflektuje významovou závislost slov:
 - nodální vázba, jak ho ašknat
 - ne vždy je možné určit závislý a řídicí člen
- ((Malý chlapce) ((psal dopis) (na stole)))



Složkový strom



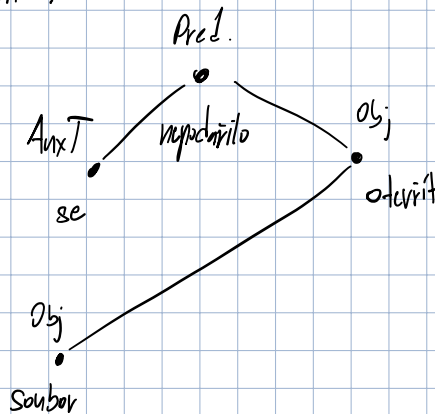
Proč to děláme:

jarní louka

- jarní má smys o sobě 27 značek
- společně ale louka říká jarní a nečinná daleko méně významových značek pro jarní.

Neprojektivní projekce:

"Soubor se nepovedlo otevřít":



?? (Soubor (se nepovedlo) otevřít) ??

→ potřebujeme nějakou vyjádření
"se nepovedlo", aby byl
soubor u otevřít.

Kláňavý slovesný: přechází větu na: "Nepovedlo se otevřít soubor".

- Syntaktická analýza je daleko těžší úkol než morfologická analýza.

Deskriptivismus:

- pouze klasifikuje a registruje, nicméně nelystologe.

Logický přístup:

- význam konceptů povrchové a hluboké syntaktické struktury.
- jeden povrchový může mít více hlubokých...

(Soubor se nepochybně otevřít a nepodařilo se otevřít soubor.)

Transformační grammatika:

Noam Chomsky

3 základní komponenty:

- báze: soubor bezkontextových pravidel, generující složovací stromy
- transformační komponenta: transformační pravidla operující na
odůlých frázových ukazatelích
- fonologická komponenta: určuje se tím, jak se slovo vyslovuje.

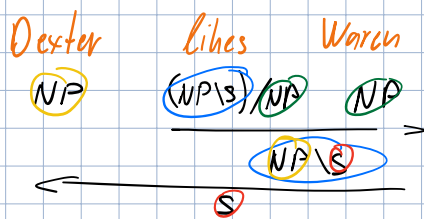
Tree Adjoining grammars

Elementární strukturou je strom. Každý neterminál může nahradit novou kmenou se stromem. Proces končí, když nemáme žádný uzol, který bych mohl nadále rozšiřovat

> ane to smrdí jako bezkontextová grammatika, ve skutečnosti je to ještě silnější.

Kategorická grammatika

- každému vstupnímu slovu se přiřadí kategorie
- grammatika pak definuje transformace slov podle kategorií



→ to co mám na stejné straně, už vyjádřím.

→ S zná celý celou větu zkonstruovanou.

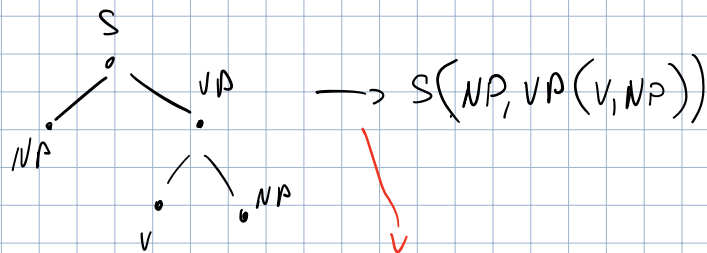
Unifikční gramatika

- objekt je reprezentovaný množinou vlastností ($\langle \text{množ} \rangle : \langle \text{hodnota} \rangle$)
- sjednocuje se do té doby, než dojde k neshodám, mezi atributy.
 - vyžaduje tedy gramatická pravidla definující neshodné vlastnosti
- problém je, že lze unifikovat vlastnosti, které spolu vůbec nesouvisí

Nástroje pro syntaktickou analýzu:

Q - systém:

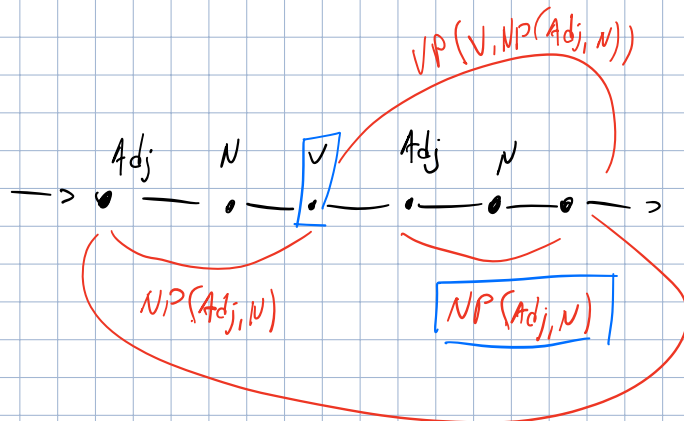
- gram. jsou linearnizované



drhý směr nelze,
jelikož existuje více repr. stromů

- systém se rozdělí na více přechodů

Máme vstupní gram.



A pravidla:

$$Adj + N \Rightarrow NP(Adj, N)$$

$$V + NP(u^*) \Rightarrow VP(V, NP(u^*))$$

$$NP(u^*) + VP(v^*) \Rightarrow S(NP(u^*), VP(v^*))$$

Q-systém jsou ale hodně komplikované, velká a výpočetně náročná.

Funkční gramatický popis

- teorie valenčních vztahů (většina řízených řídicích sloh)
- navrhuje na přechodu lingv. šlach

Valence:

- sada pravidel, kterými nám de facto uvěříme význam slov

- dva základní druhy závislých členů na TG rovině:

- aktanty: kanoniz., pacient, adresát, origin., efekt

- všechno může být zastoupeno ve větě jen jednou.

- volní doplňky: mohou se vyskytnout neomezeně krát

obligatorní - ten tam vždy být musí (tutelně nepřítom v kontextu)

Kontrola gramatické správnosti:

- problémy specifické pro češtinu.

- shoda podmětu s přísudkem, ↗

- interpunkce,

- neprojelativní konstrukt,

- zájmen

často nemusí být podnět vyjádřen,

tudíž bez analýzy nemáme, k čemu se více přisluhuje.

Tatínek šel do práce.

↳ křivka.

→ proto také nemůžeme automaticky označit jako chybu.

Ude co mám dělat? ↗

Jak kontrolovat?

- chybové vzorby: vhodné pro jazyky s pevným slovosledem

- například že musím mít podnět vždy před slovesem a mají stejné číslo

- zkouška stavby syntaktického stromu → když to nejde, vědecky máme chybu.

- redukční analýza:

- odstráníme bezvýznamná slova

- při odstránění se nesmí žádný chyb „opomít“ a

někdo namítl se nesmí žádný nový chyb objevit.

Problém homonymie: → hráči na cello

- čelisti klopaly / čelisti klopali

- jukmile existuje jediná správná varianta, nesmím to označit ani chybně,
protože nesmím, který význam je použit ve skutečnosti.