

第 11 章

句法樹為本模型

教科書網站：www.statmt.org/book/

參考課程網站：mt-class.org/jhu/syllabus.html

Dec. 2, 2018

句法樹為本模型

- 統計式、類神經模型操作在詞序列之上
- 很多翻譯問題、現象用句法比較容易解釋、執行
 - 詞序重排, 例如, 英中翻譯中的介詞移動 (harmful to human 對人體有害)
 - 長距離的一致性 (如距離很遠的主詞、動詞的一致)

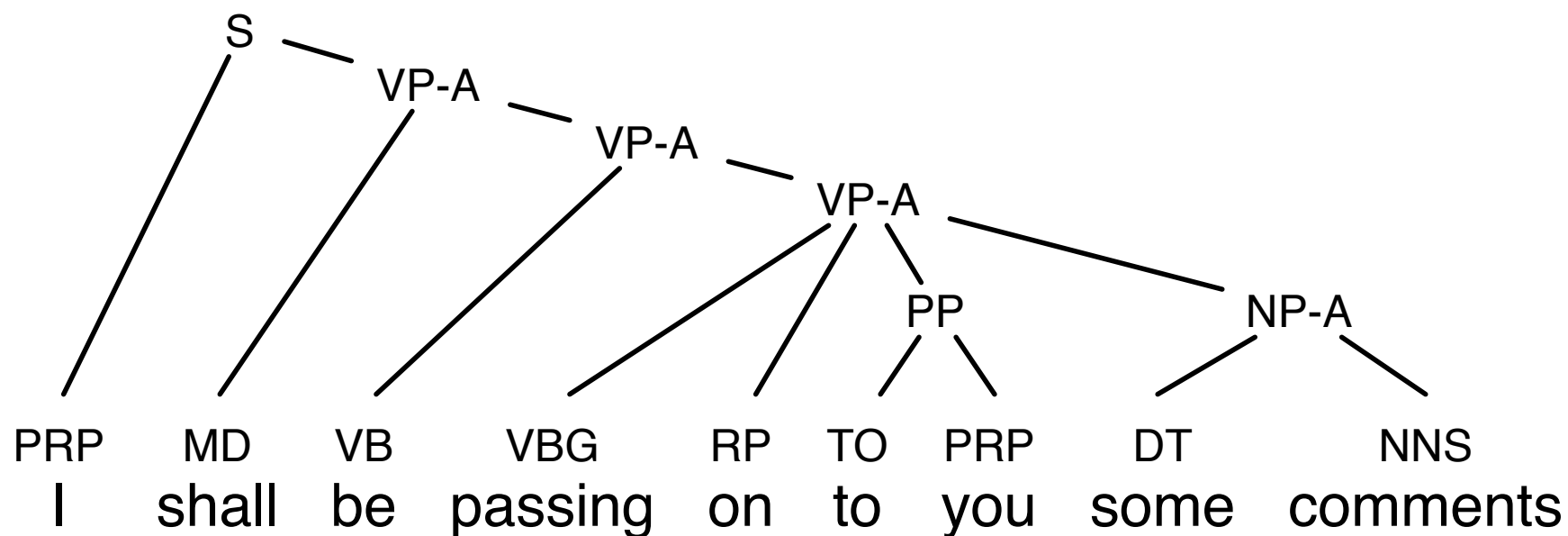
⇒ 用句法樹表達的翻譯模型

- 持續進行的研究
- 對某些語言配對還是最佳作法 (英德, 英中)

片語結構文法 Phrase Structure Grammar PSG

- 片語結構
 - 名詞片語: the big man, a house, ...
 - 介詞片語: at 5 o'clock, in Edinburgh, ...
 - verb phrases: going out of business, eat chicken, ...
 - 形容詞片語, 等等
- 上下文無關文法 Context-free Grammars (CFG)
 - 非終止符號 NT : 片語標籤, 詞性標籤 (名詞 NN 等等)
 - 終止符號 T : 詞
 - 生成規則 production rules: $NT \rightarrow [NT, T]^+$
 - 例子: $NP \rightarrow DET\ NN$
 - $DET \rightarrow the$
 - $NN \rightarrow dog$

片語結構文法 (2)



剖析器產生的句法樹 (Collins parser)

同步片語結構文法 Synchronous PSG

- 英語規則

$$\text{NP} \rightarrow \text{DET JJ NN}$$

- 法語規則

$$\text{NP} \rightarrow \text{DET NN JJ}$$

- 同步規則 (指出對應):

$$\text{NP} \rightarrow \text{DET}_1 \text{NN}_2 \text{JJ}_3 \mid \text{DET}_1 \text{JJ}_3 \text{NN}_2$$

同步文法規則

- 非終止規則

$$\text{NP} \rightarrow \text{DET}_1 \text{ NN}_2 \text{ JJ}_3 \mid \text{DET}_1 \text{ JJ}_3 \text{ NN}_2$$

- 終止規則

$$\text{N} \rightarrow \text{maison} \mid \text{house}$$
$$\text{NP} \rightarrow \text{la maison bleue} \mid \text{the blue house}$$

- 混合規則

$$\text{NP} \rightarrow \text{la maison JJ}_1 \mid \text{the JJ}_1 \text{ house}$$

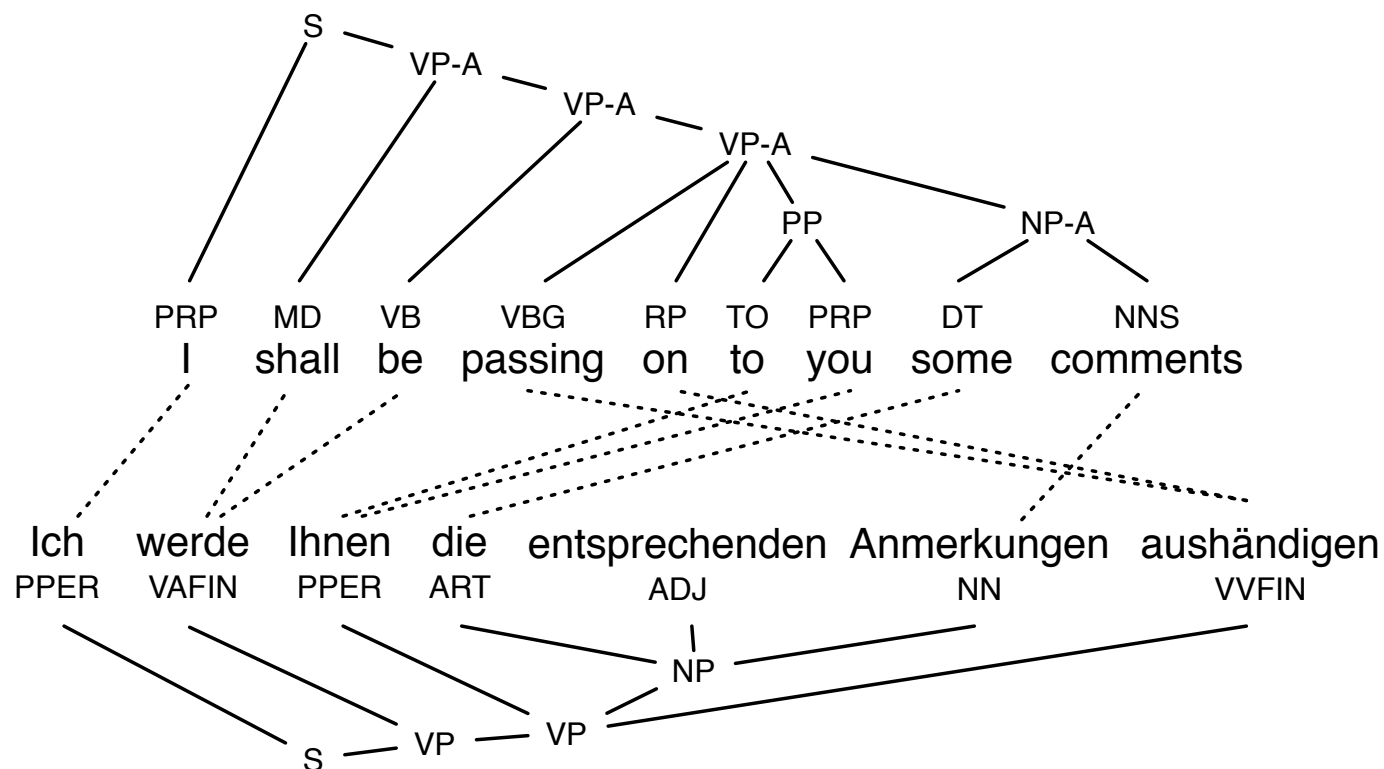
剖析樹為本翻譯模型

- 透過剖析的翻譯過程
 - 用同步文法剖析輸入句
 - 同時也產生了輸出句
 - 過程就是一連串的運用規則來比對輸入 (RHS) 簡化 (LHS)
- 翻譯機率：規則機率的乘積

$$\text{SCORE}(\text{TREE}, E, F) = \prod_i \text{RULE}_i$$

- 有各種方法來產生規則，指定規則機率

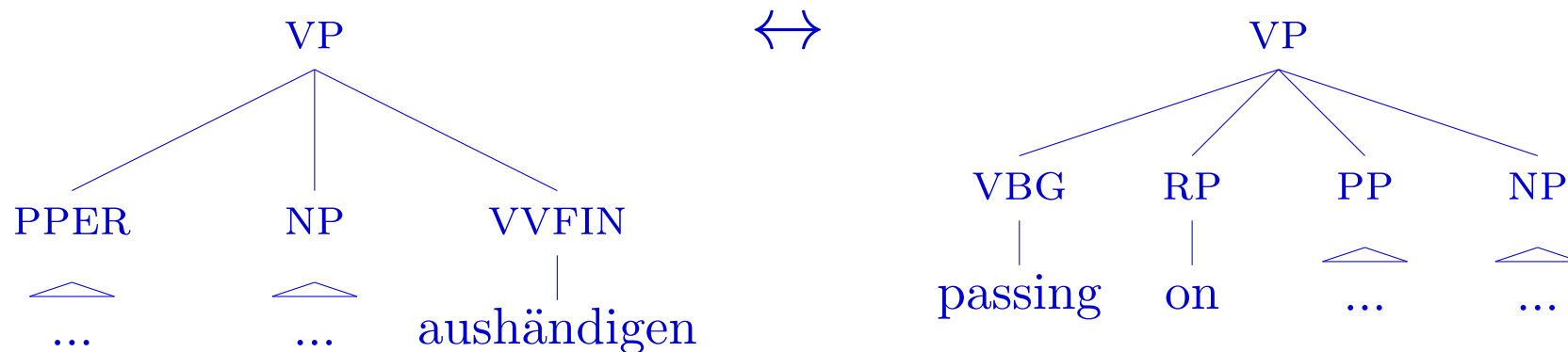
剖析樹對應



利用詞對應可以產生 PSG 剖析樹間的節點到節點對應
(德-英句配對的例子)

詞序重組規則 Reordering Rule

- 對應子樹到子樹 Subtree alignment



- 同步文法規則 Synchronous grammar rule

$$VP \rightarrow PPER_1 NP_2 \text{ aushändigen} \mid \text{passing on } PP_1 NP_2$$

- 註記：容許一對多的詞對應 (如 aushändigen 對到 passing on ok)
但是，對非終止符號不容許一對多 (非終止符號間有一對一的限制s)
詞彙化在不同語言間，常有差異，而文法概念有普遍性，所有語言都一致

更多同步規則的例子

- 子樹對應 Subtree alignment



- 同步文法規則 (跳過英語的內部結構，直接實現成片語)

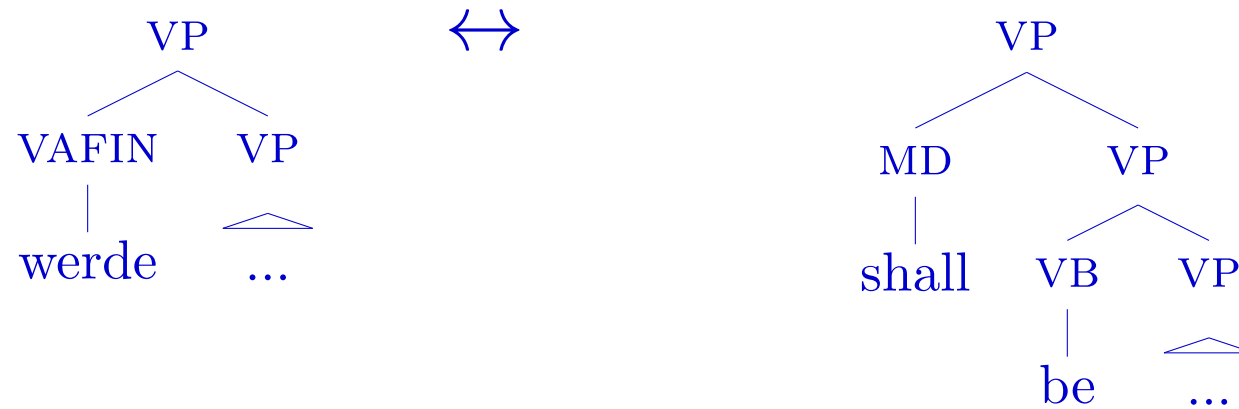
$\text{PRO/PP} \rightarrow \text{Ihnen} \mid \text{to you}$

- 有內部結構的規則 Rule with internal structure

$\text{PRO/PP} \rightarrow \text{Ihnen} \mid \begin{array}{cc} \text{TO} & \text{PRP} \\ | & | \\ \text{to} & \text{you} \end{array}$

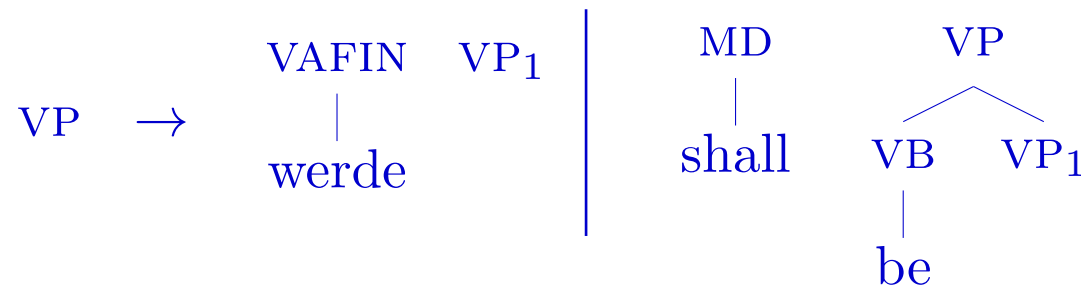
更多同步規則的例子

- 翻譯德語 *werde* 成為英語 *shall be*



- 翻譯規則必須指出動詞片語的對應 (VP_1)

\Rightarrow Complex rule



内部結構

- 去除内部結構

$VP \rightarrow \text{werde } VP_1 \mid \text{shall be } VP_1$

⇒ 同步 CFG

- 保留内部結構

$VP \rightarrow$

VAFIN	VP ₁		MD	VP
				/ \
werde			shall	VB VP ₁
				be

⇒ 同步樹替代文法 tree substitution grammar item 差別：機率的計算不同

學習同步文法

- 由做好詞對應平行語料庫
- 第一種：階層試片與模型 Hierarchical phrase-based model (有結構但不區隔動詞、名詞、形容詞的結構)
 - 只有一個非終止符號 x
 - 沒有語言學的句法，只是一種形式文法
- 第二種：同步片語結構模型 phrase structure model
 - 有非終止符號來代表不同片語、詞彙: NP, VP, PP, ADJ, ...
 - 語料庫需要先做句法剖析

回顧：擷取片語翻譯規則（準備轉成同步文法規則）

	Ich	werde	Ihnen	die	entsprechenden	Anmerkungen	aushändigen
I							
shall							
be							
passing							
on							
to							
you							
some							
comments							

→ shall be = werde

回顧：擷取片語翻譯規則 (2)

	Ich	werde	Ihnen	die	entsprechenden	Anmerkungen	aushändigen
I							
shall							
be							
passing							
on							
to							
you							
some							
comments							

some comments =
die entsprechenden Anmerkungen

擷取階層片語翻譯規則 (1) LHS

	Ich	werde	Ihnen	die	entsprechenden	Anmerkungen	aushändigen
I							
shall							
be							
passing							
on							
to							
you							
some							
comments							

.....▶ werde Ihnen die entsprechenden
Anmerkungen aushändigen
= shall be passing on to you
some comments

擷取階層片語翻譯規則 (2) RHS 分解

	Ich	werde	Ihnen	die	entsprechenden	Anmerkungen	aushändigen
I							
shall							
be							
passing							
on							
to							
you							
some							
comments							

subtracting subphrase

.....▶ werde X aushändigen
= shall be passing on X

正式定義—根據片語翻譯

- 回顧: 一致性區塊 consistent phrase pairs

(\bar{e}, \bar{f}) consistent with $A \Leftrightarrow$

$$\begin{aligned} & \forall e_i \in \bar{e} : (e_i, f_j) \in A \rightarrow f_j \in \bar{f} \\ \text{AND } & \forall f_j \in \bar{f} : (e_i, f_j) \in A \rightarrow e_i \in \bar{e} \\ \text{AND } & \exists e_i \in \bar{e}, f_j \in \bar{f} : (e_i, f_j) \in A \end{aligned}$$

- 令 P 為所有的一致性片語對應 (\bar{e}, \bar{f})

正式定義—產生階層片語翻譯規則

- 遞迴式定義 (操作):

$$\begin{aligned} &\text{if } (\bar{e}, \bar{f}) \in P \text{ AND } (\bar{e}_{\text{SUB}}, \bar{f}_{\text{SUB}}) \in P \\ &\quad \text{AND } \bar{e} = \bar{e}_{\text{PRE}} + \bar{e}_{\text{SUB}} + \bar{e}_{\text{POST}} \\ &\quad \text{AND } \bar{f} = \bar{f}_{\text{PRE}} + \bar{f}_{\text{SUB}} + \bar{f}_{\text{POST}} \\ &\quad \text{AND } \bar{e} \neq \bar{e}_{\text{SUB}} \text{ AND } \bar{f} \neq \bar{f}_{\text{SUB}} \\ &\text{add } (e_{\text{PRE}} + X + e_{\text{POST}}, f_{\text{PRE}} + X + f_{\text{POST}}) \text{ to } P \end{aligned}$$

(註記: e_{PRE} , e_{POST} , f_{PRE} , or f_{POST} 的每一個都可能是空字串)

- 階層片語配對有封閉性，不會無窮遞迴

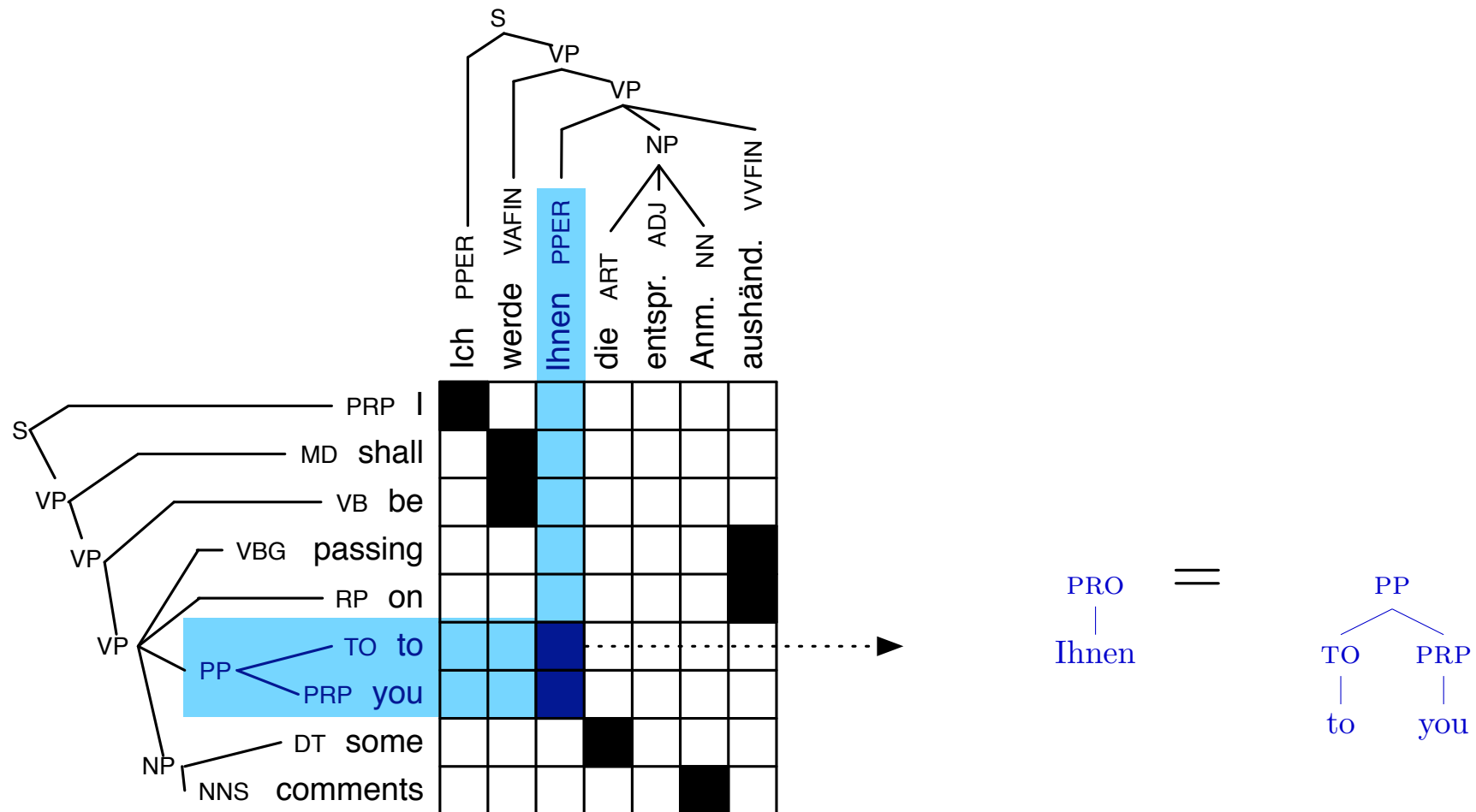
補充

- 去除有兩個非終止符號的規則，如：

$$Y \rightarrow X_1 X_2 \mid X_2 \text{ of } X_1$$

- 通常，需要一些限制來計算量 [Chiang, 2005]
 - 最多 2 個非終止符號
 - 一個到五詞 (在翻譯的兩邊的語言)
 - 最常涵蓋 15 詞 (包括 X 的空格 gaps)

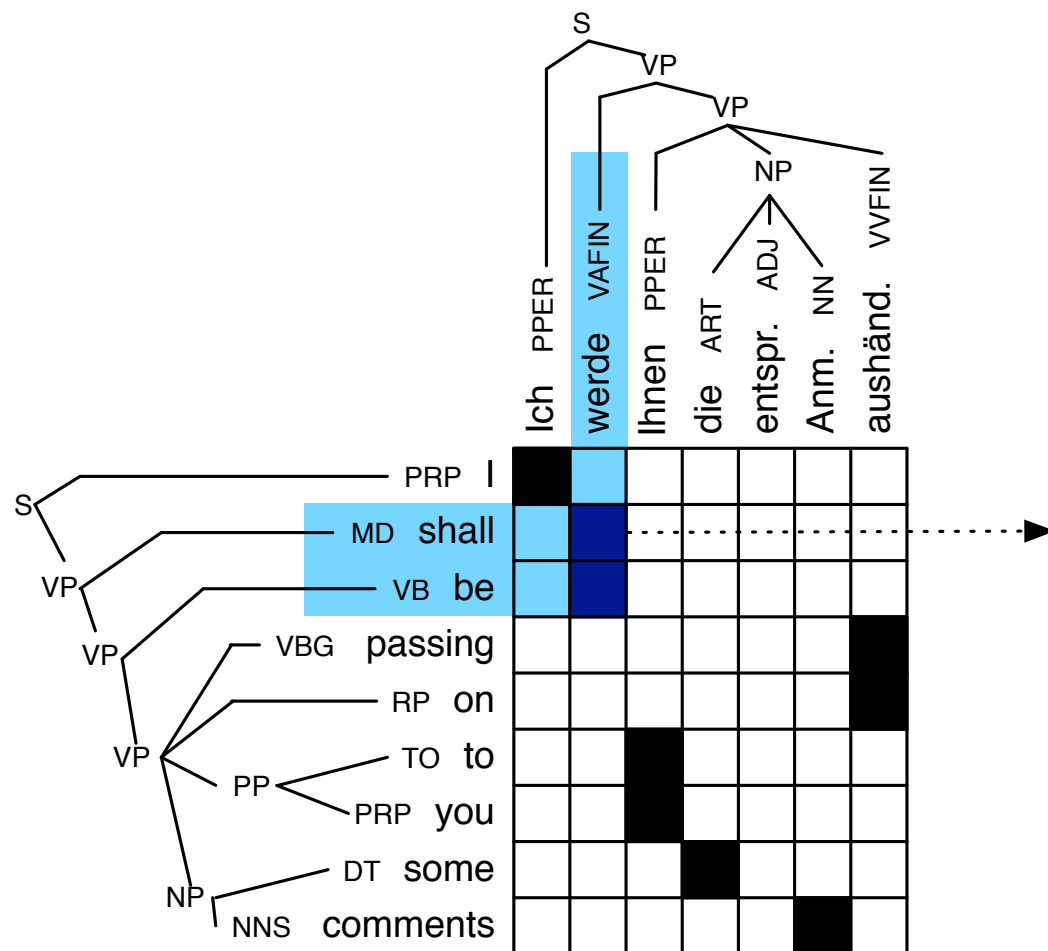
第二種：學習句法翻譯規則 (不只是階層)



句法規則的限制

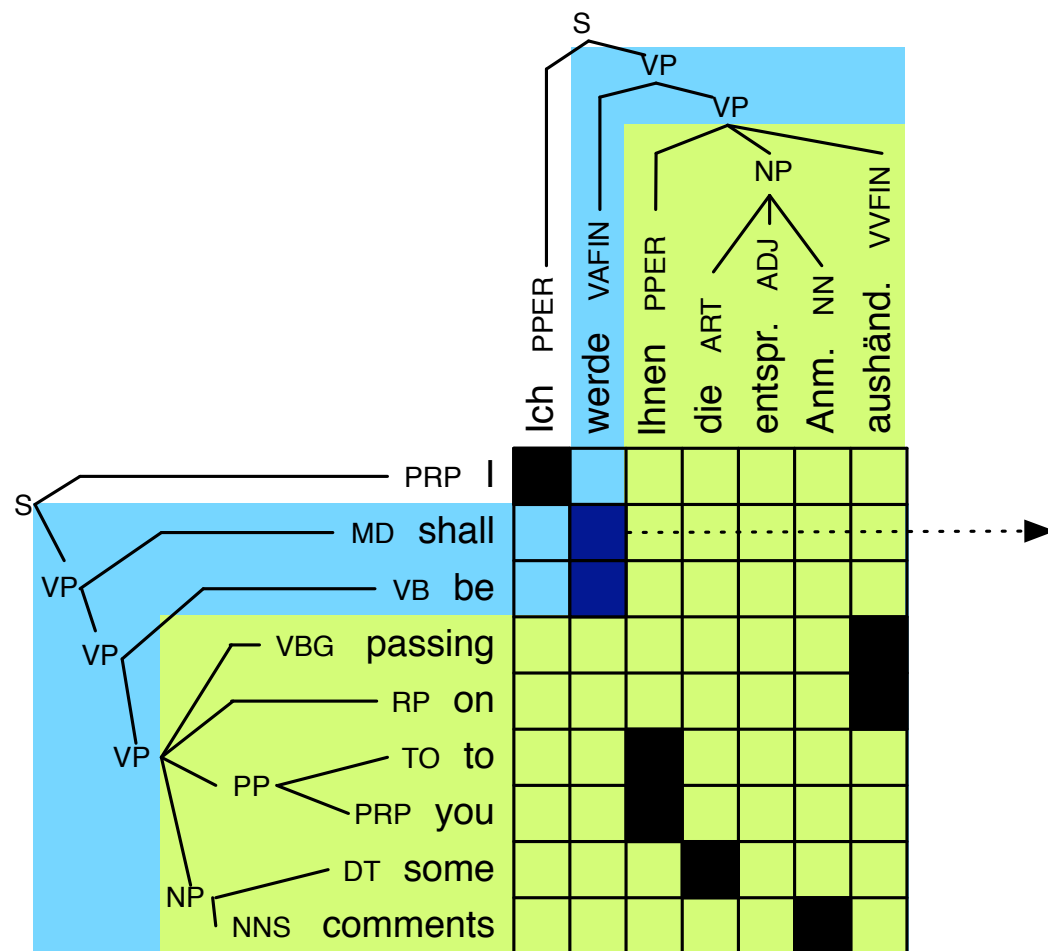
- 詞彙數量的限制和階層模型相同
- 階層模型: 規則可以跨越任何片段 `span`
⇔ 句法規則只能跨越剖析樹成分結構
- 階層模型: 次階層空格 `gap` 可為任意跨距
⇔ 句法規則的 只能跨越剖析樹成分結構
- 相對而言，只能擷取較少的句法規則 (相同的狀況下)
- 或許，規則少比較合理 (階層模型可有百萬規則)

違反規則的例子

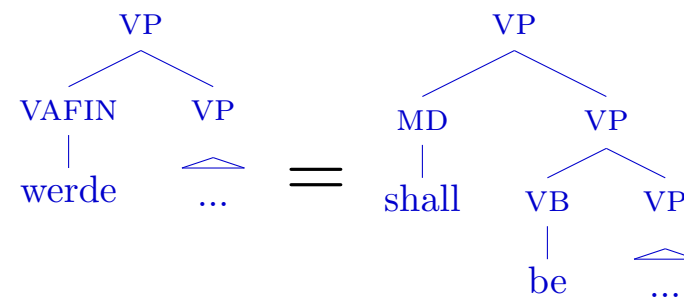


英語的片語不是一個文法成分
不能擷取規則

有上下文限制的規則



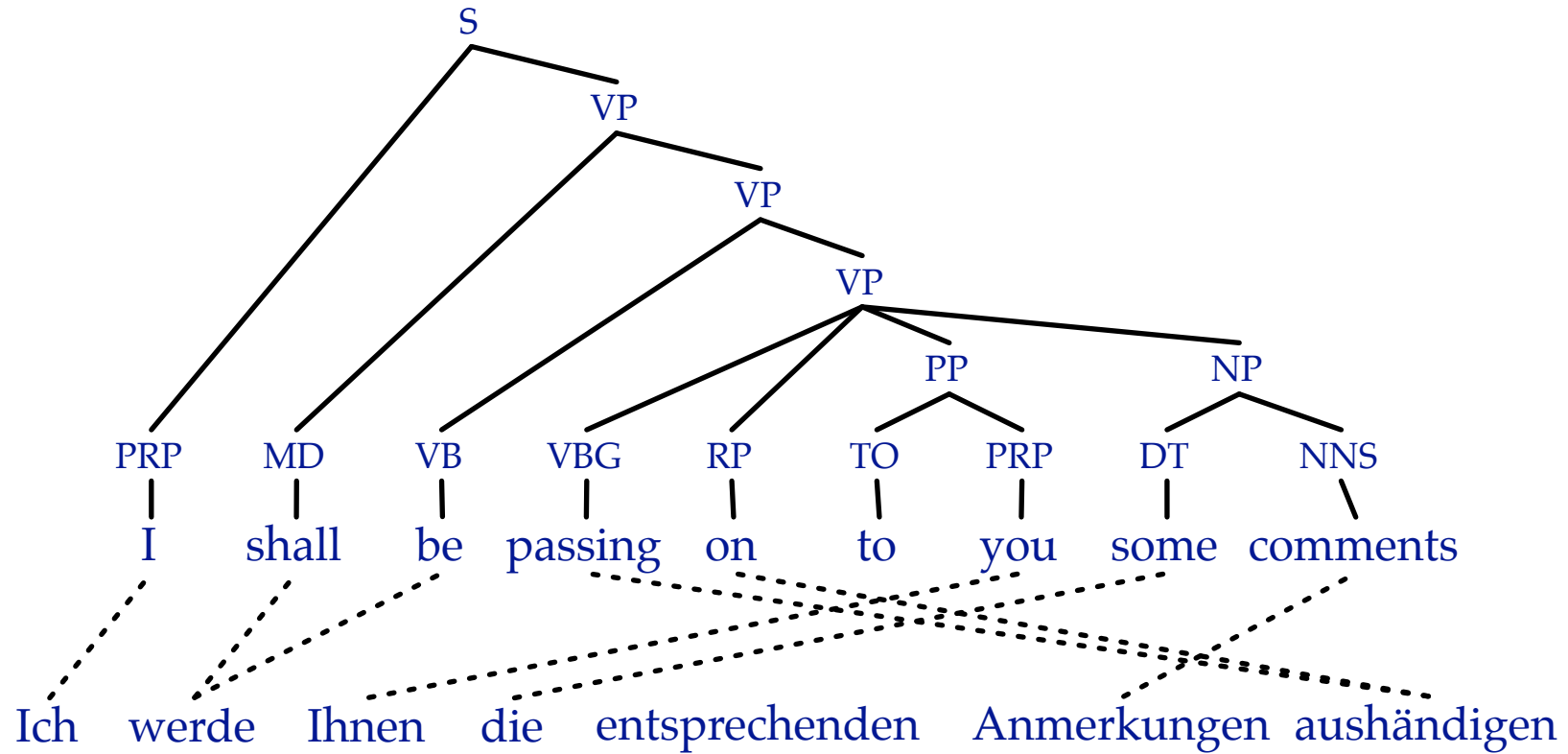
這組片語配對
需要句法的上下文(非CFG)



太多規則可以擷取

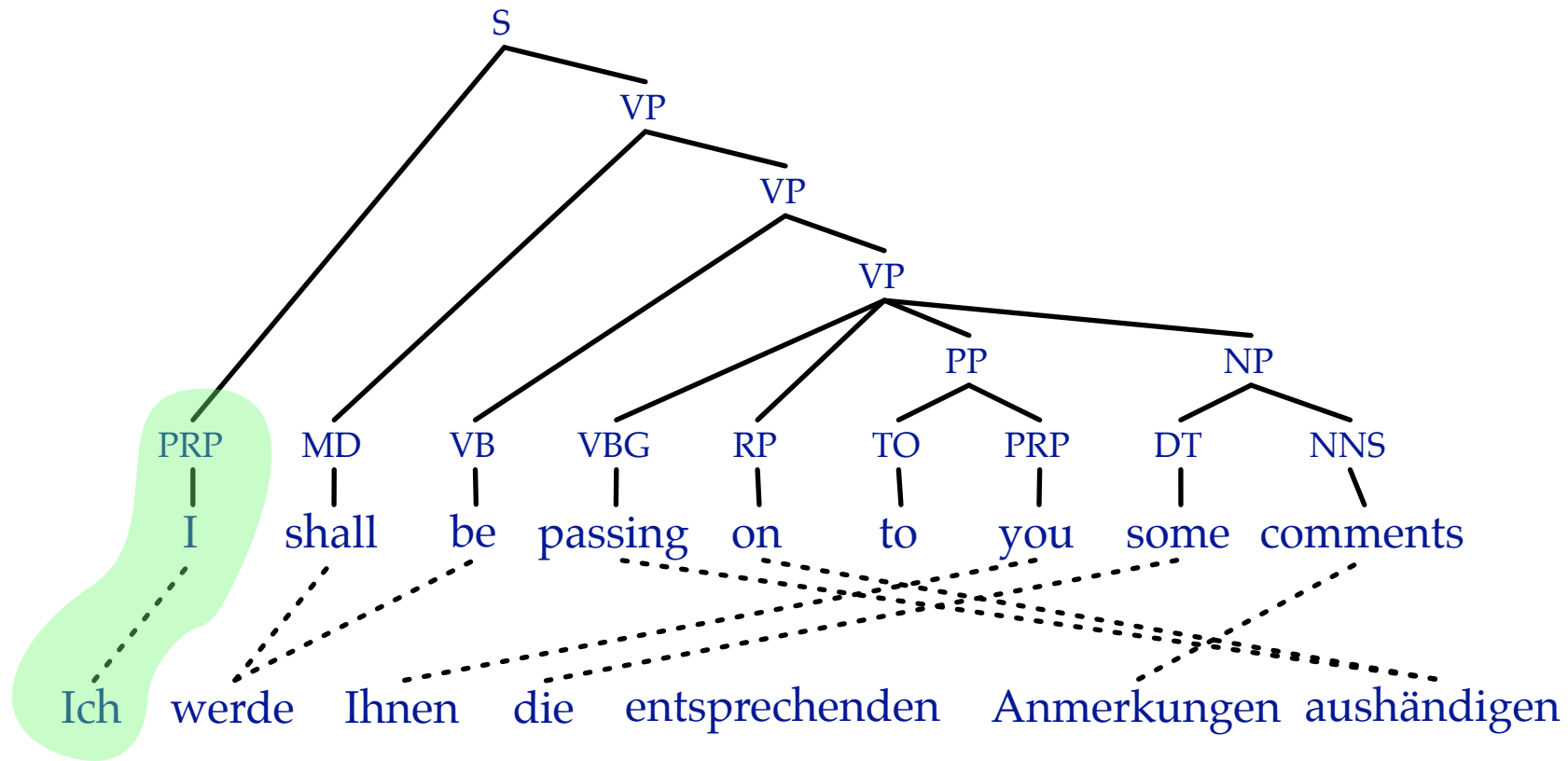
- 可以擷取大量的規則
(每組對應詞都可以產生很多規則 → 指數成長)
- 需要限制規則
- 選項 1: 和階層一樣的限制
(最長跨距、最多終止、非終止符號的限制等)
- 選項 2: 只擷取最小規則 minimal rules (**GHKM** 規則)
 - Michel **Galley**, Mark **Hopkins**, Kevin **Knight**, and Daniel **Marcu**. 2004. What's in a translation rule? In Proceedings of NAACL-04, 273–280.

Minimal Rules



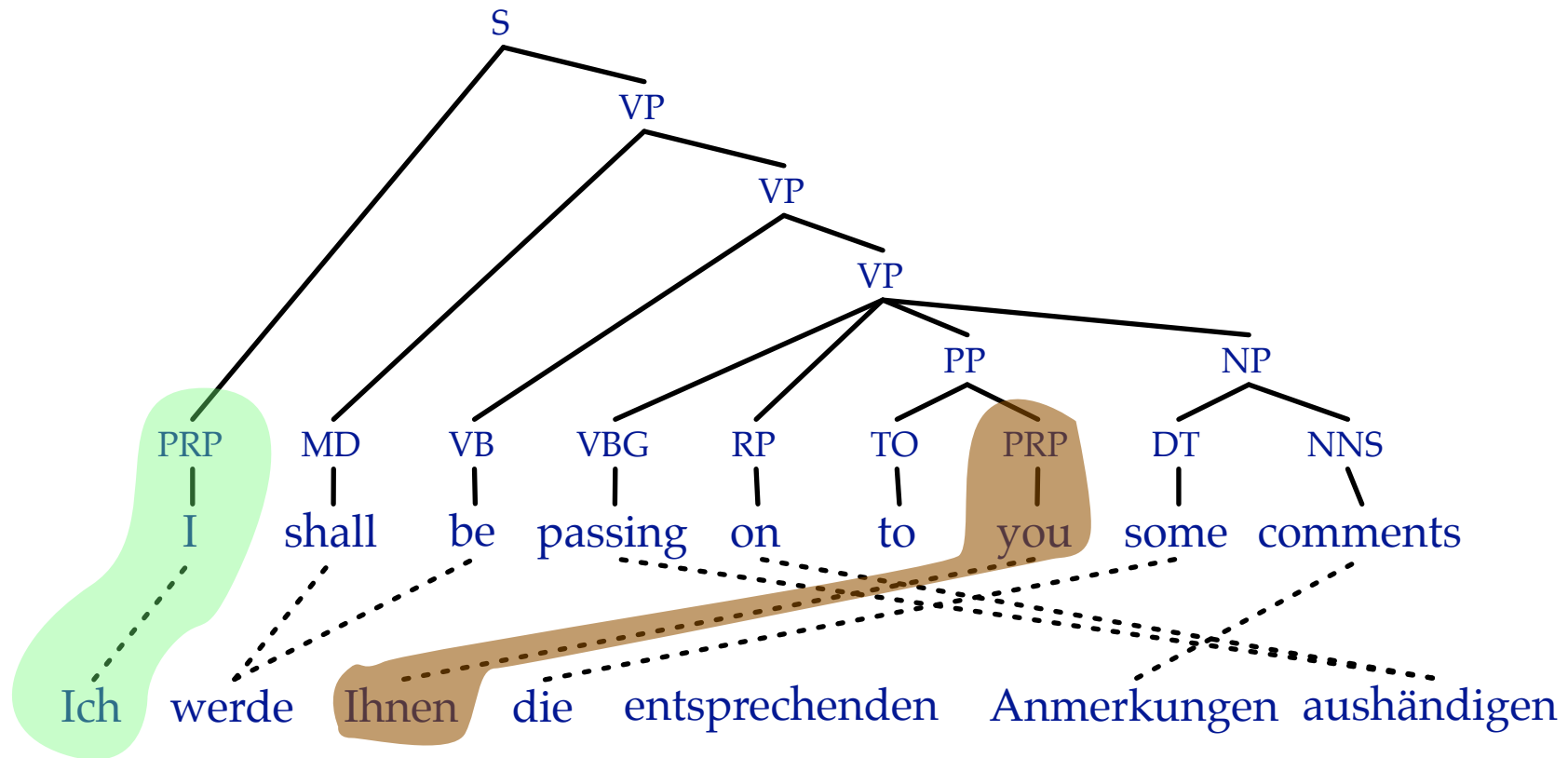
規則擷取: 可「解釋」這組句子配對的一組最少規則

最少規則 (1) 詞彙規則

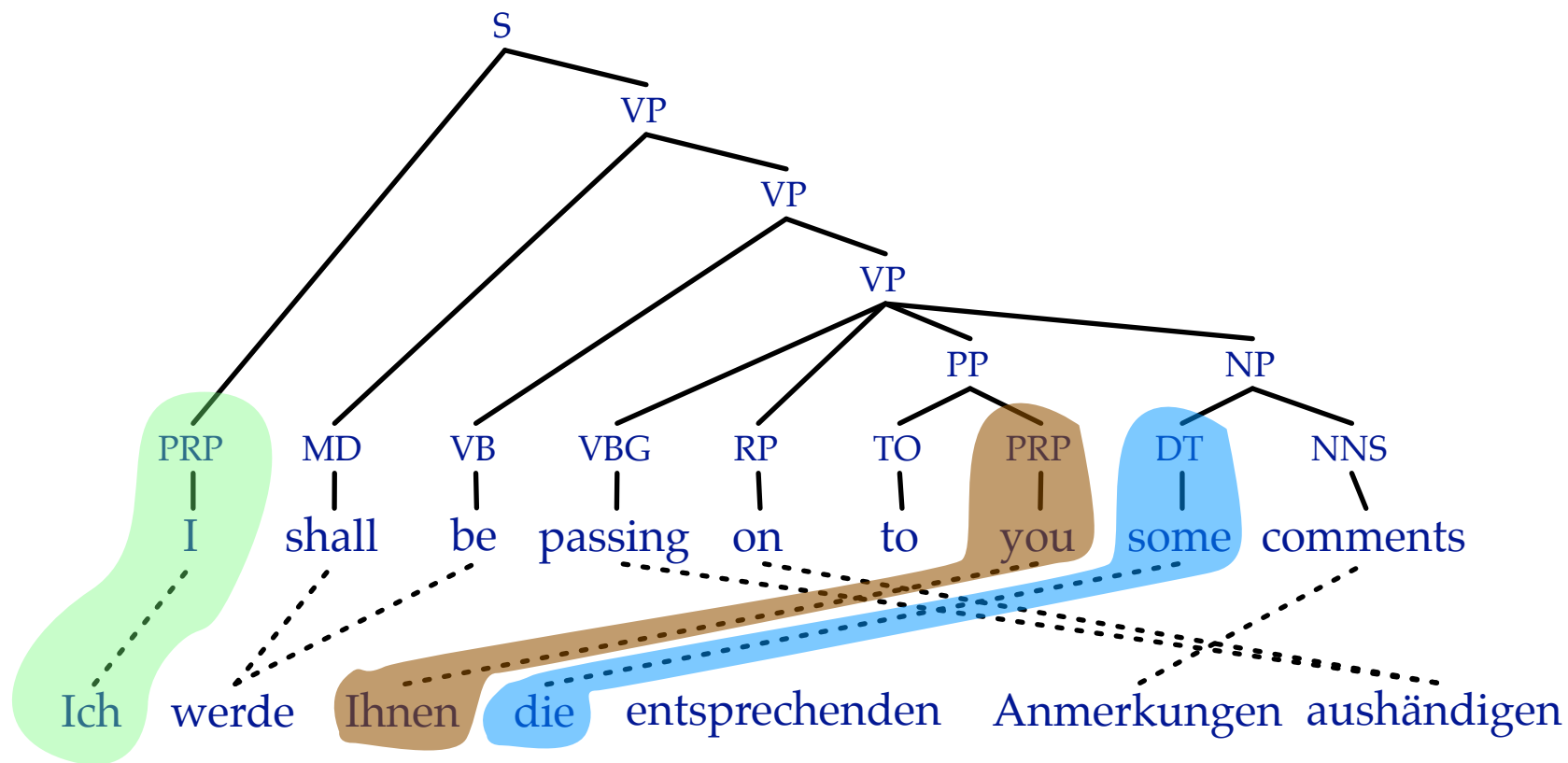


Extracted rule: $\text{PRP} \rightarrow \text{Ich} \mid \text{I}$

最少規則 (2) 詞彙規則

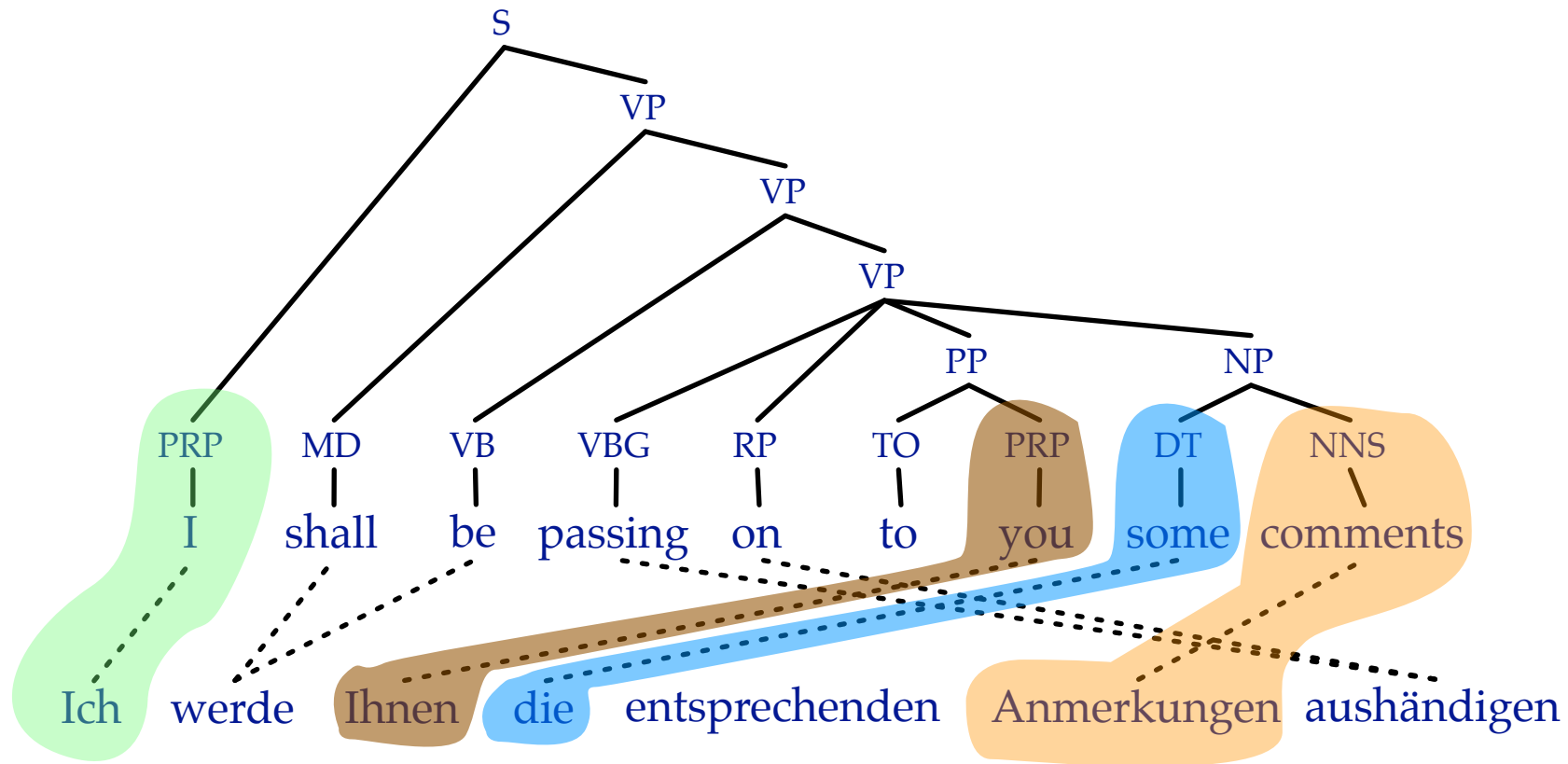


最少規則 (3) 詞彙規則



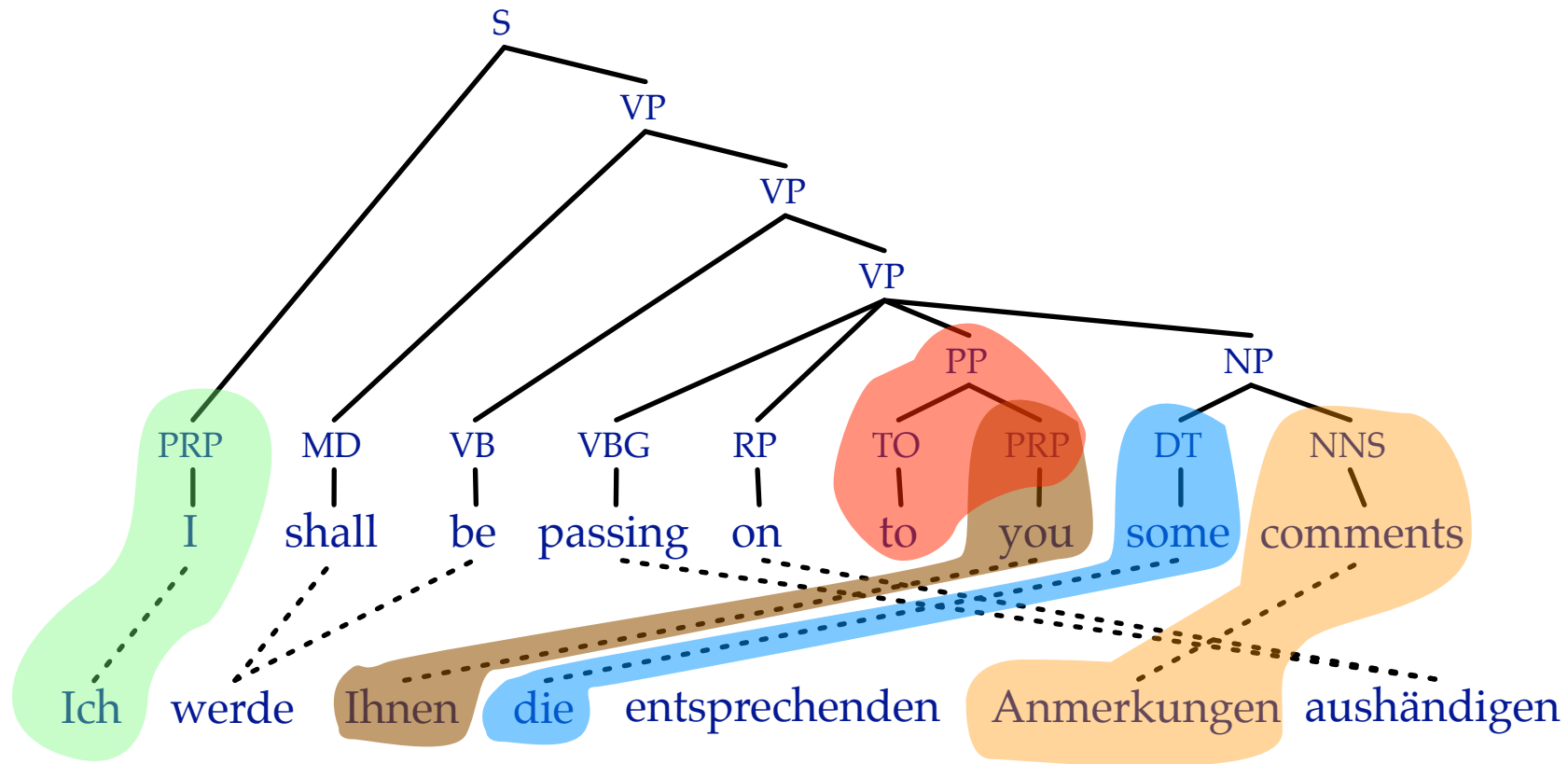
Extracted rule: $DT \rightarrow \text{die} \mid \text{some}$

最少規則 (4) 詞彙規則

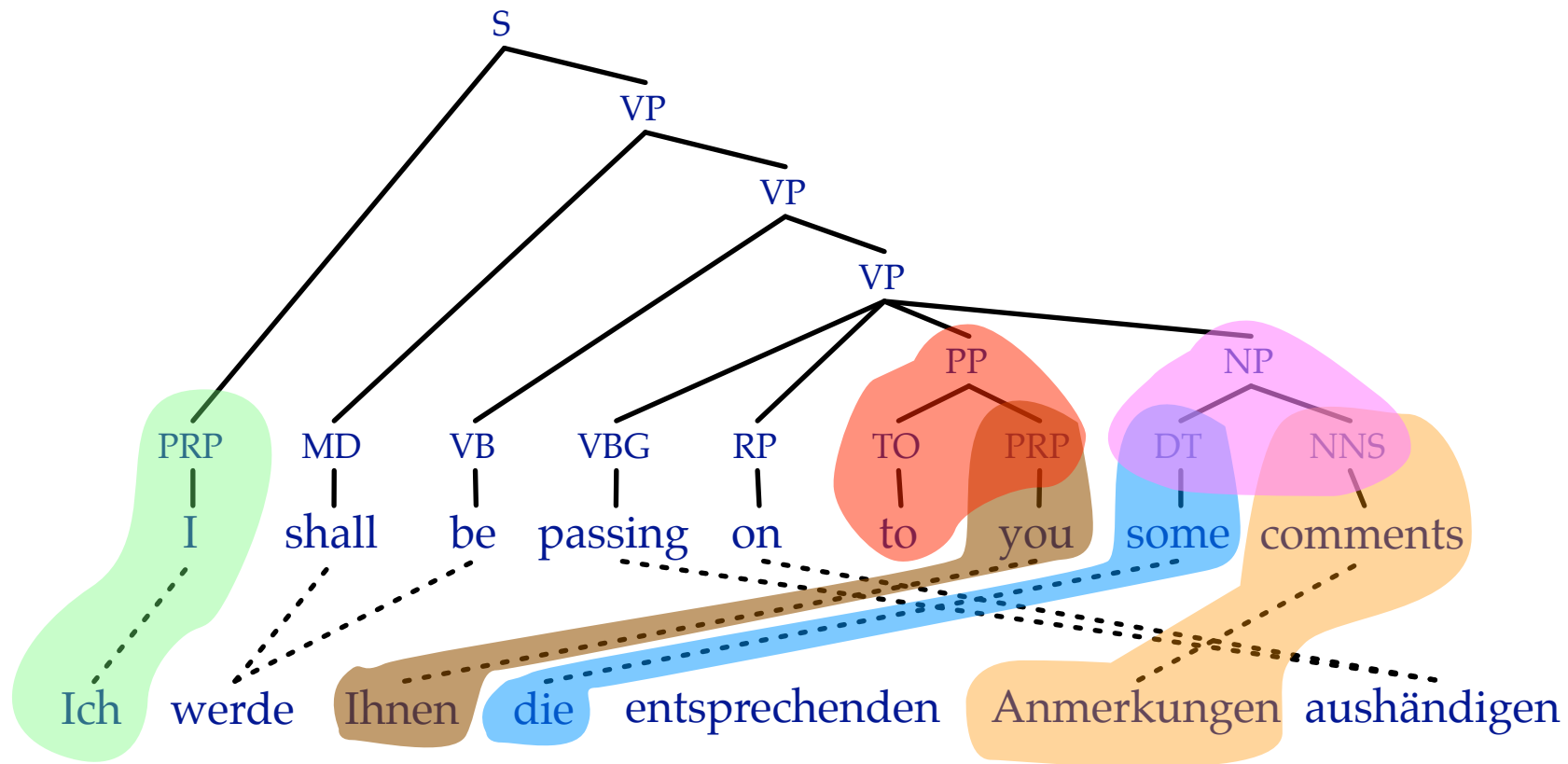


Extracted rule: $NNS \rightarrow \text{Anmerkungen} \mid \text{comments}$

最少規則 (5) 插入規則

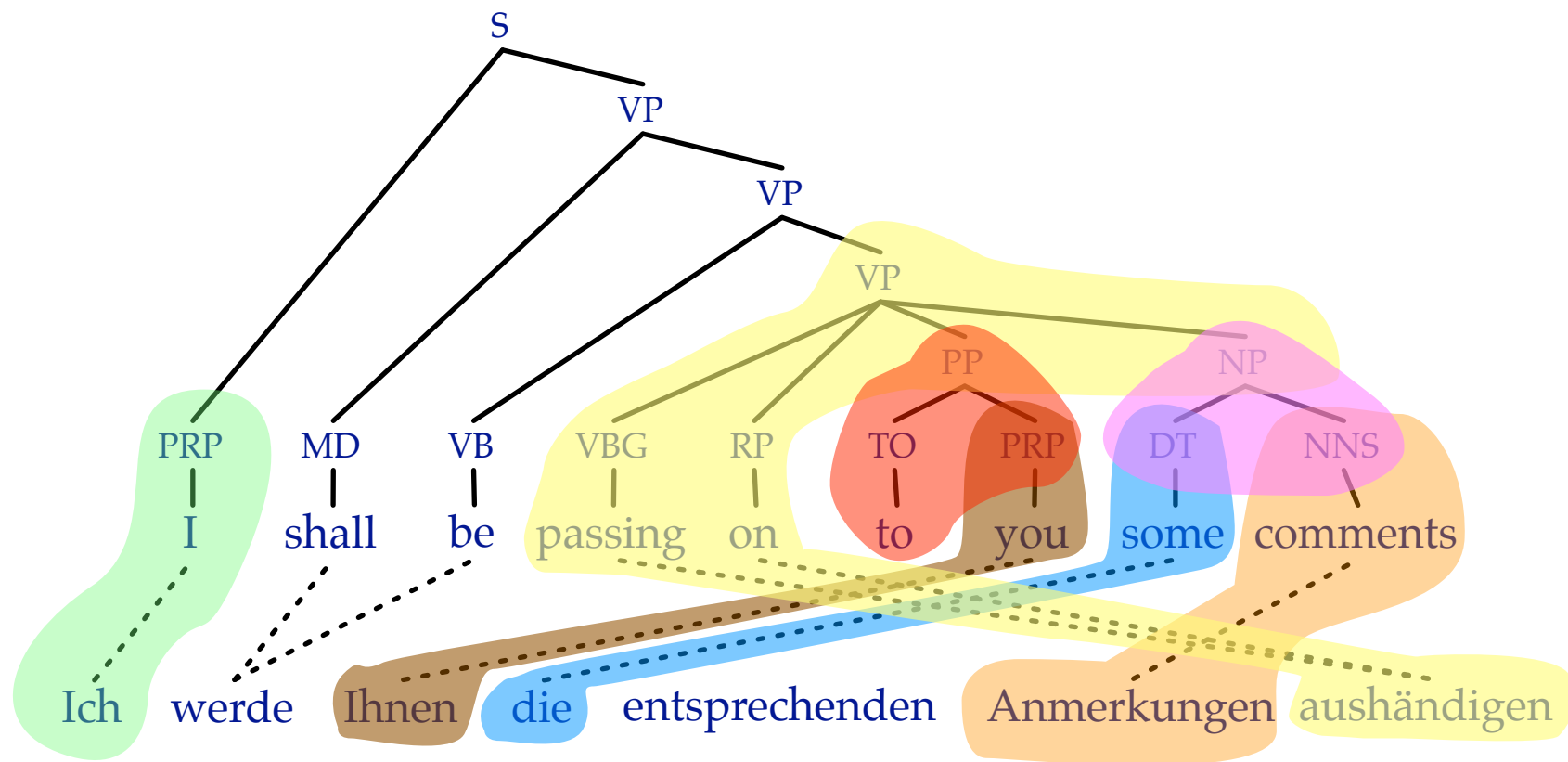


最少規則 (6) 内部規則



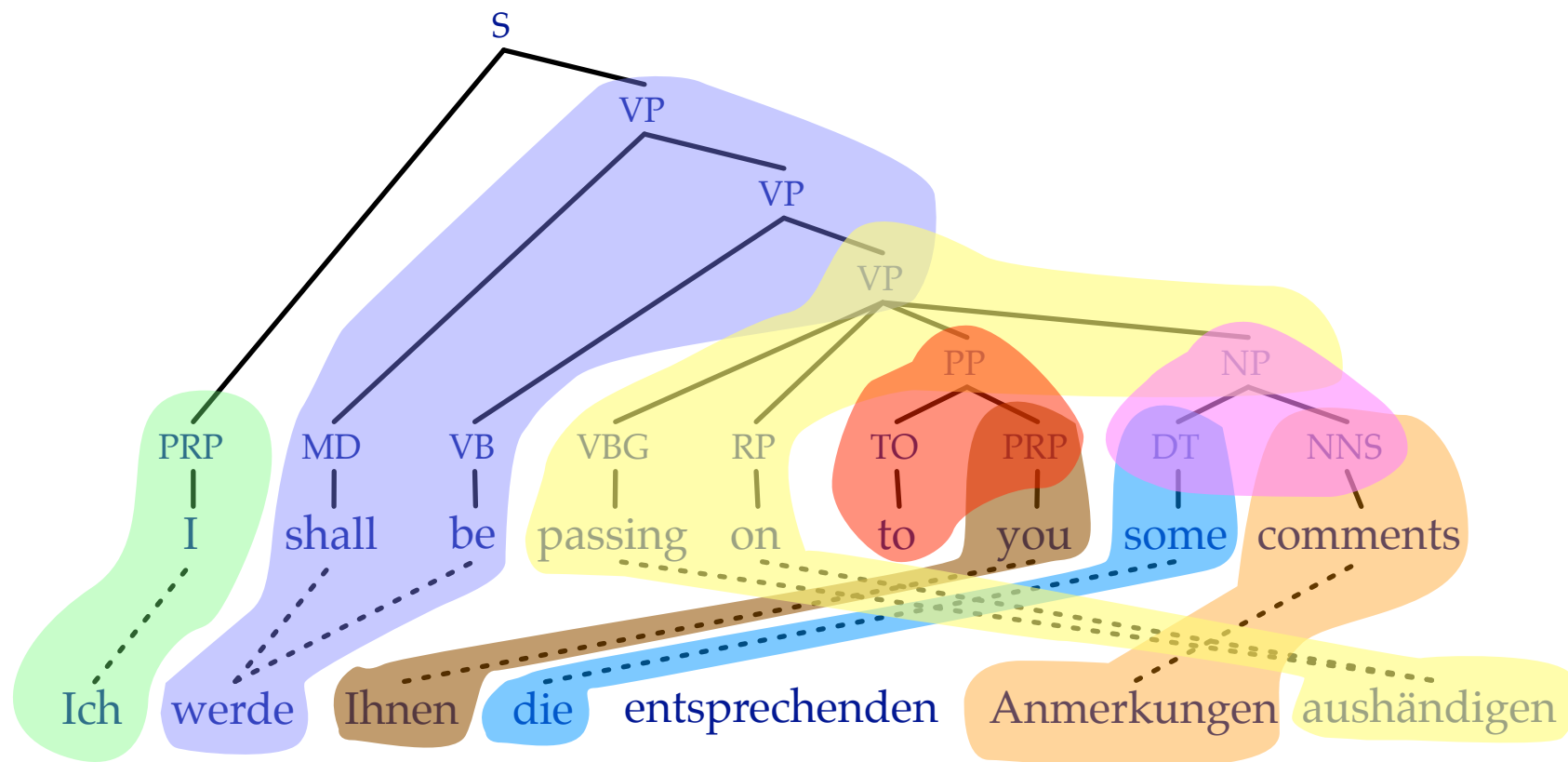
擷取規則: $NP \rightarrow X_1 X_2 \mid DT_1 NNS_2$

最少規則 (7) 有上下文的規則



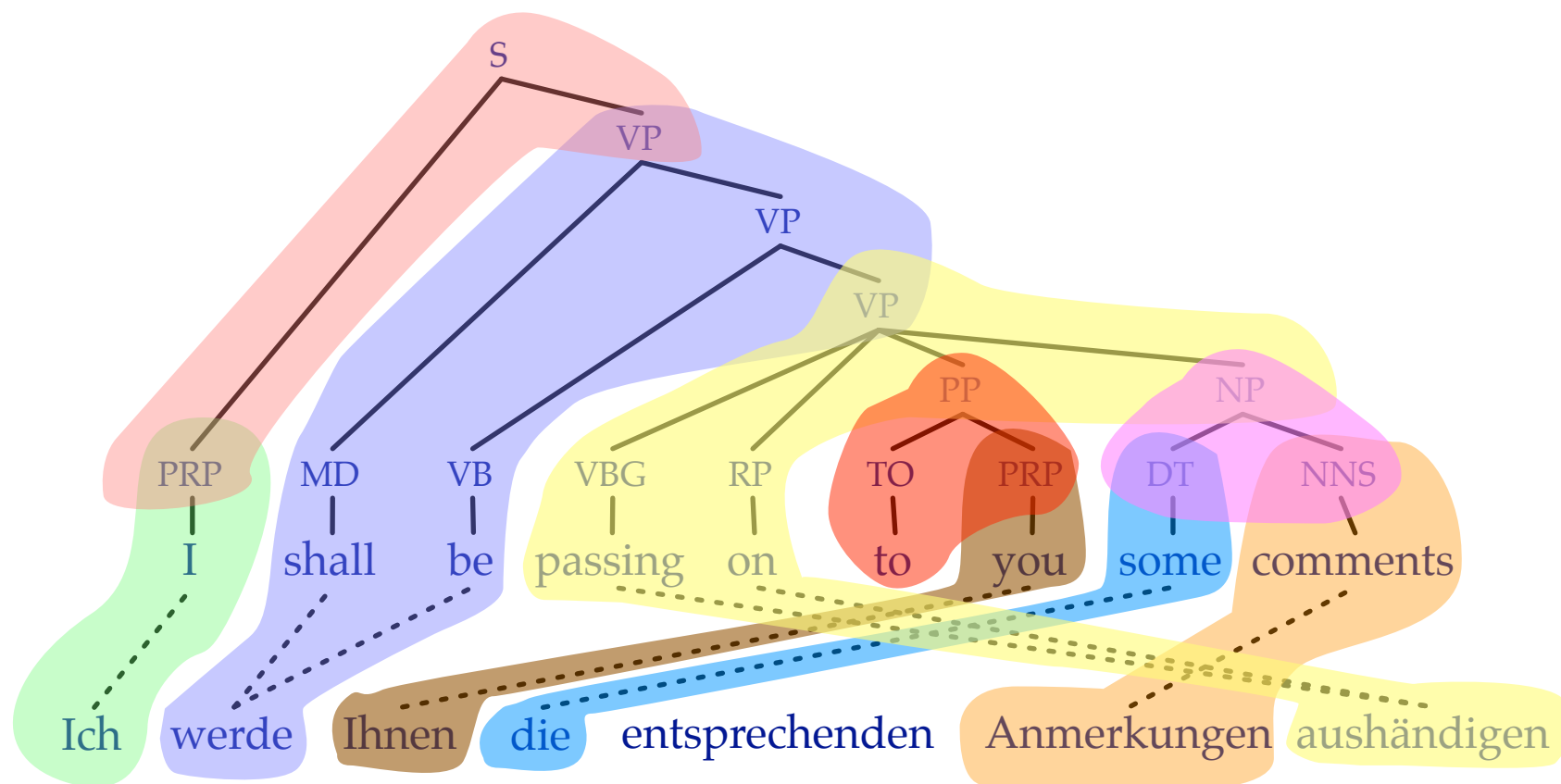
Extracted rule: $VP \rightarrow X_1 X_2 \text{ aushändigen} \mid \text{passing on } PP_1 NP_2$

最少規則 (8) 有上下文的規則



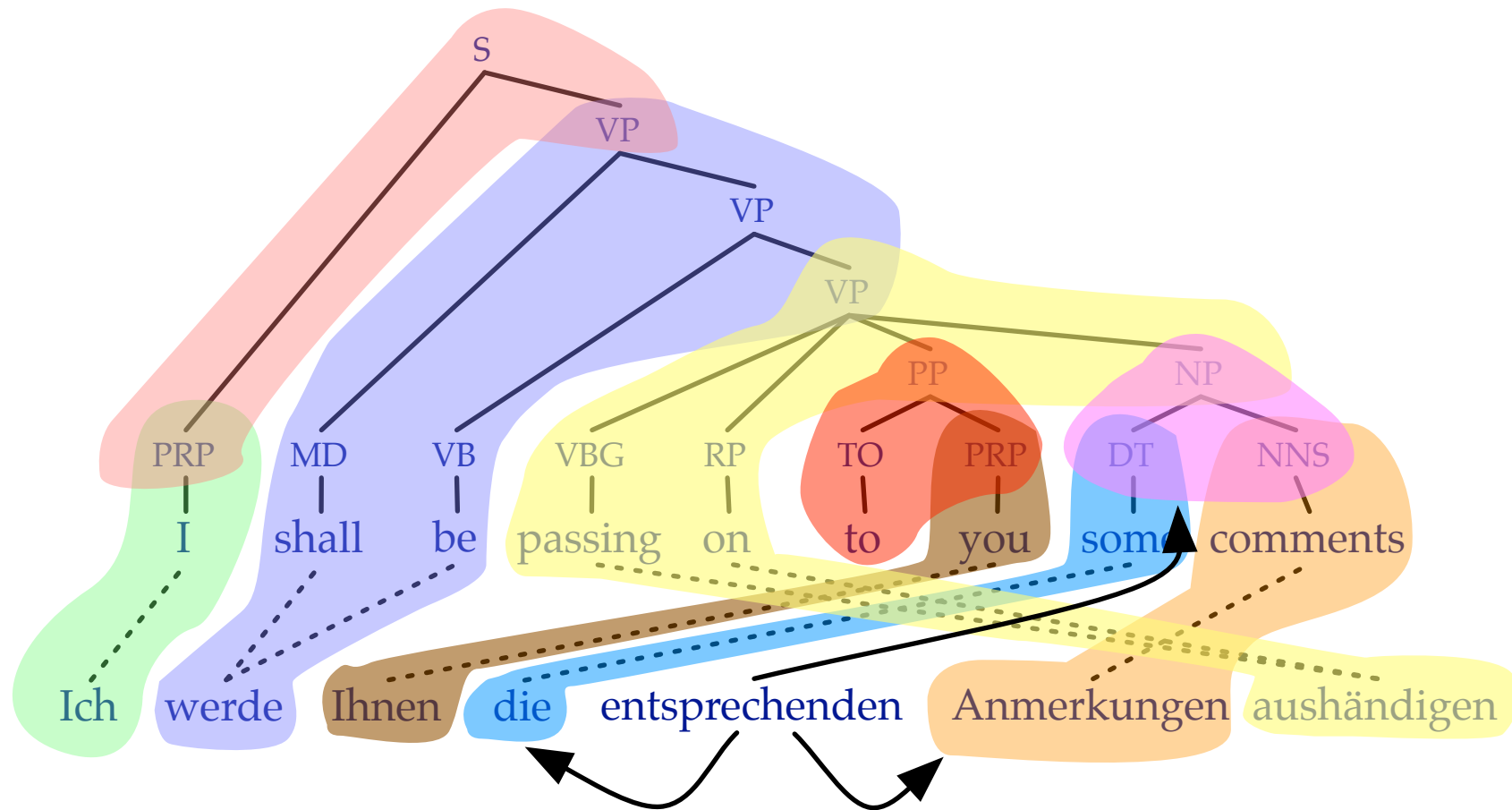
Extracted rule: $VP \rightarrow \text{werde } x \mid \text{shall be } VP$ (不管內部結構)

最少規則 (9) 內部規則



擷取規則: $S \rightarrow X_1 X_2 \mid PRP_1 VP_2$ 完成！ (註: 一個成分只有一條規則)

最少規則 (10) 未對應到的來源詞



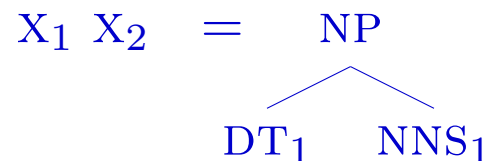
插入到鄰近詞或更高的節點 → 增加一條規則

太少片語規則嗎？

- 詞彙規則大部分都是1-對-1 (除非對應程式有不同的結果)
- 但是：統計式機器翻譯模型詞彙規則非常有利
- 解答：
 - 組合規則直到符號數量的上線 (如同在階層式模型中，記得 "選項 1")
 - 用最小的規則，涵蓋最多的內部節點 non-leaf nodes

複合規則

- 現有兩條規則



die = DT
|
some

entsprechenden Anmerkungen = NNS
|
comments

- 組合後規則

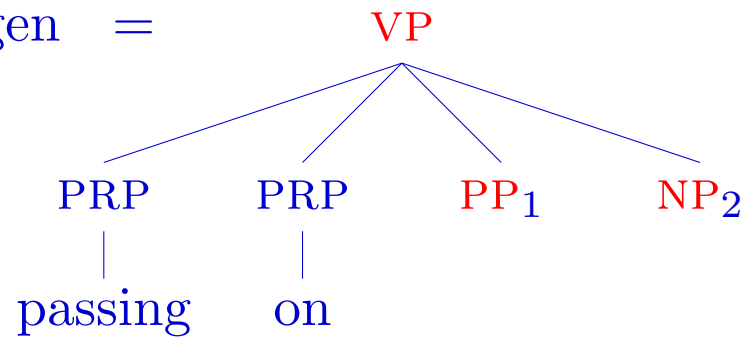


(1 non-leaf node: NP)

複合規則 (2)

- 最小規則:

$x_1 x_2$ aushändigen =

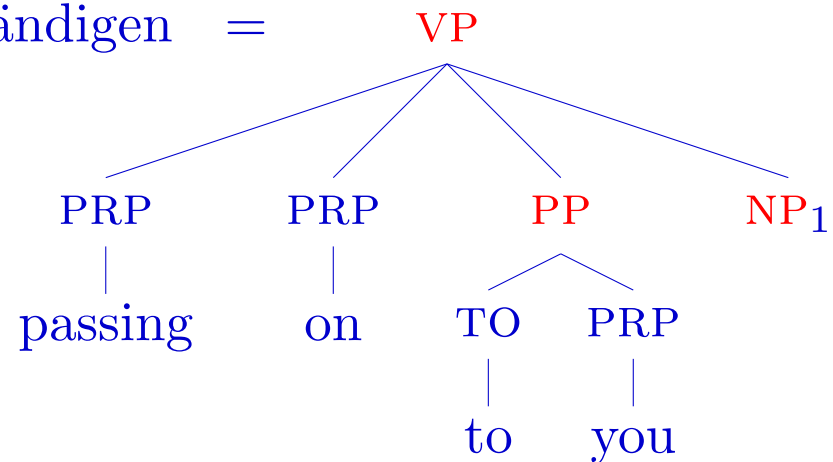


3 個内部節點:

VP, PP, NP

- 組合後規則:

Ihnen x_1 aushändigen =



3 個内部節點:

VP, PP and NP

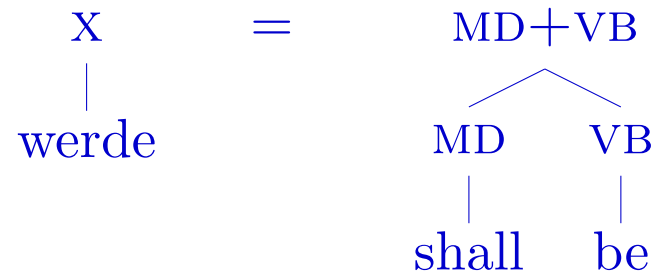
放鬆剖析樹限制

- 不可行的規則 impossible rule (LHS 不唯一)



- 產生新的非終止符號: MD+VB

⇒ 新規則



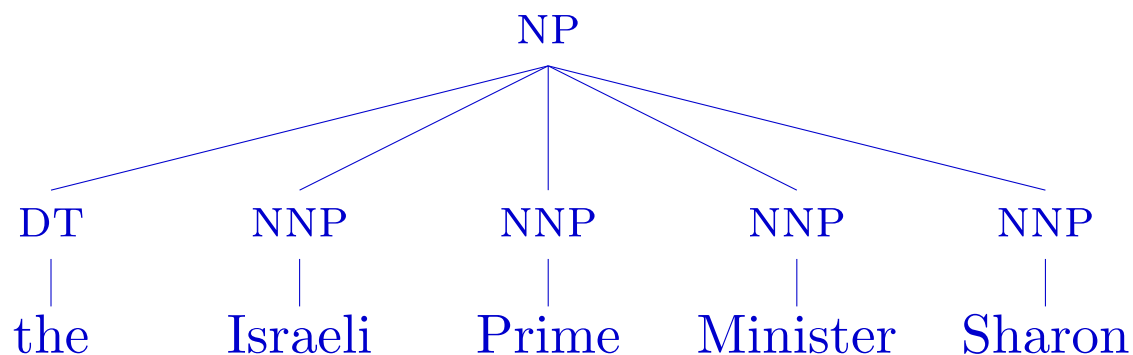
ZV 放鬆策略 Zollmann Venugopal Relaxation

- 跨越兩個成分，合成一個新符號 $x+y$
- 跨越三個成分，合成一個新符號 $x+y+z$
- 跨越的成分，屬於同一個親節點 x 而且包含
 - 包含最後節點 y 之外所有節點，標示為 x/y
 - 包含最前節點 y 之外所有節點，標示為 $x \setminus y$
- 其他情況，標示為 失敗 FAIL

⇒ 可以抽取更多規則，非終止符號暴增

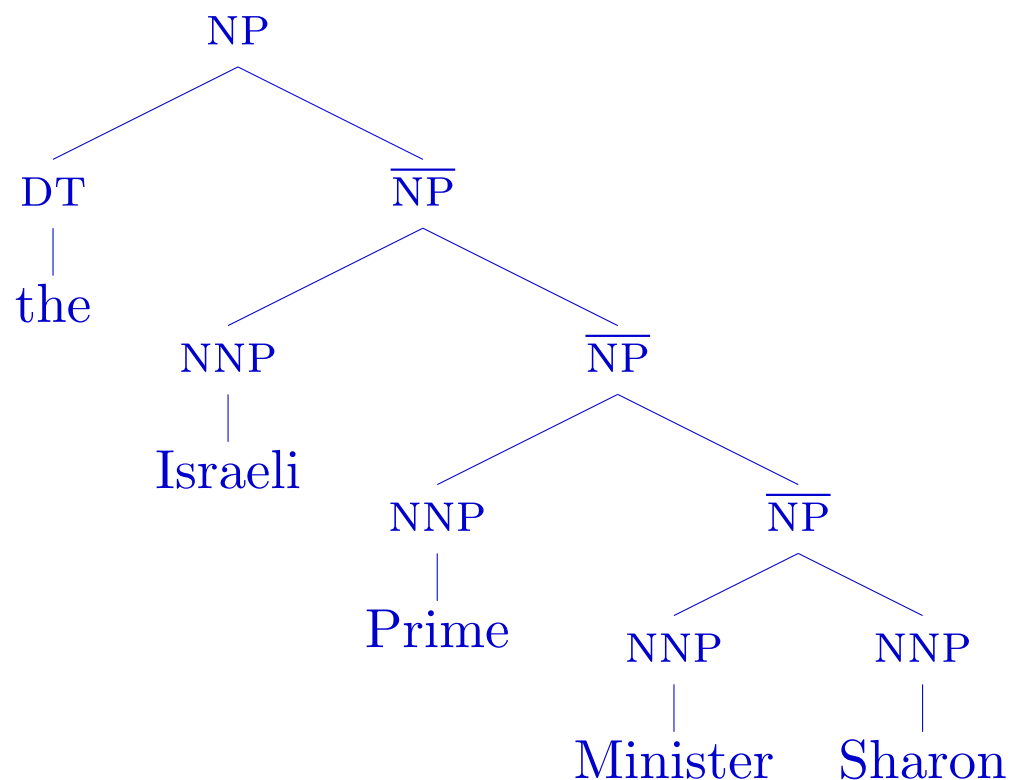
特殊問題：扁平結構

- 剖析器產生扁平的剖析樹＝有片語無結構，限制了規則擷取的效果



- 不是單一詞規則，就是整組片語規則——比原來的片語式模型更沒有彈性

用二元化剖析樹解決 Tree Binarization



的確可以抽取更多規則

向左展開的二元化，還是向右展開？(英語幾乎是向右展開)

翻譯規則的評分

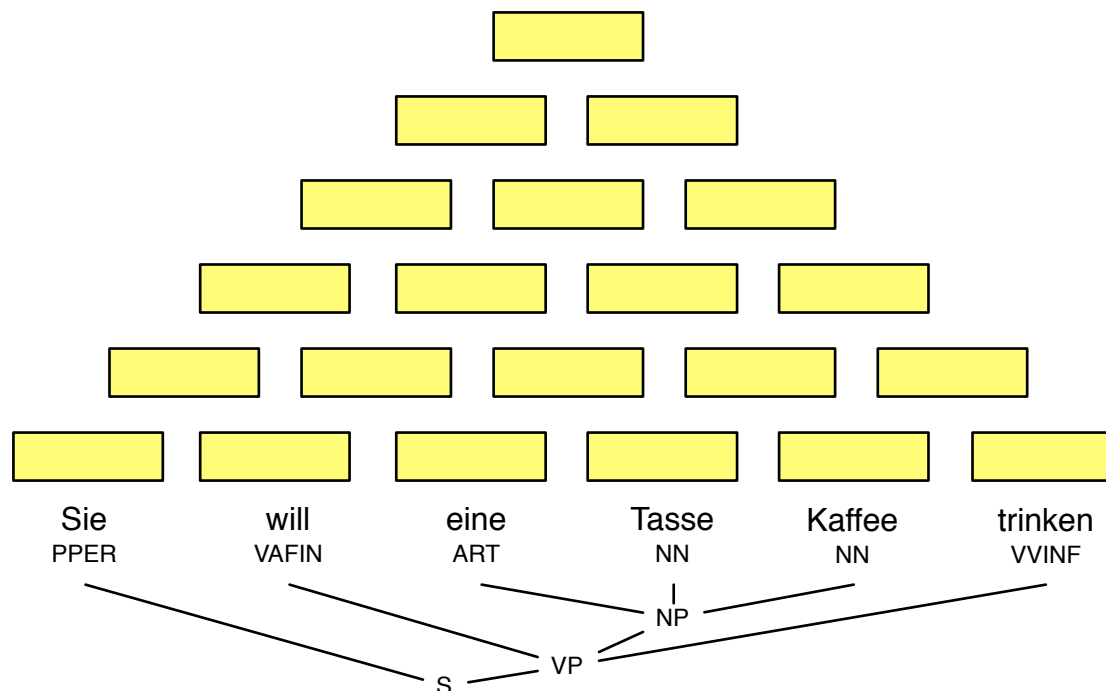
- 擷取語料庫的所有規則
- 用次數來評分
 - 聯合機率 (不以 LHS 為條件): $p(\text{LHS}, \text{RHS}_f, \text{RHS}_e)$
 - 條件機率 (以 LHS 為條件): $p(\text{RHS}_f, \text{RHS}_e | \text{LHS})$
 - 直接翻譯機率: $p(\text{RHS}_e | \text{RHS}_f, \text{LHS})$
 - 雜訊通道機率 (和翻譯方向 f 到 e 相反): $p(\text{RHS}_f | \text{RHS}_e, \text{LHS})$
 - 詞彙翻譯機率: $\prod_{e_i \in \text{RHS}_e} p(e_i | \text{RHS}_f, a)$

句法式解碼

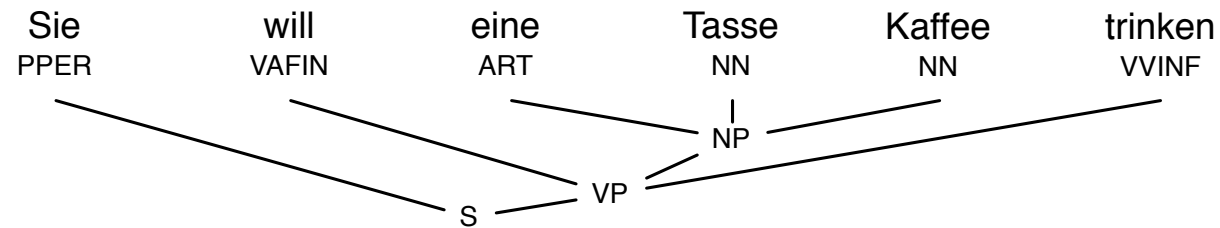
靈感來自單語的 chart parser 句法分析:

在解碼過程中，維持一份 $O(n^2)$ 跨距的表格可供填入片語標籤:

最後在頂端空格填入 S 就算成功完成剖析

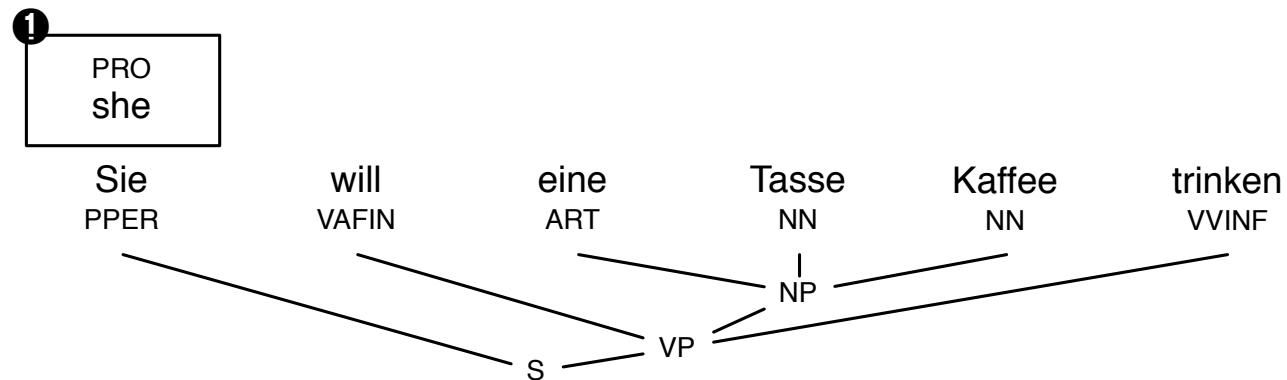


句法式解碼舉例 (1)



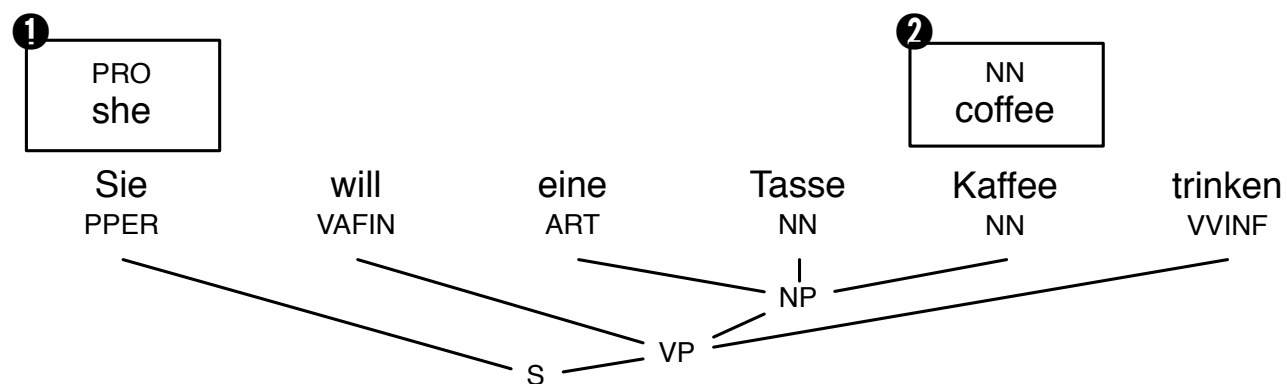
輸入德語句子和剖析樹

句法式解碼舉例 (2)



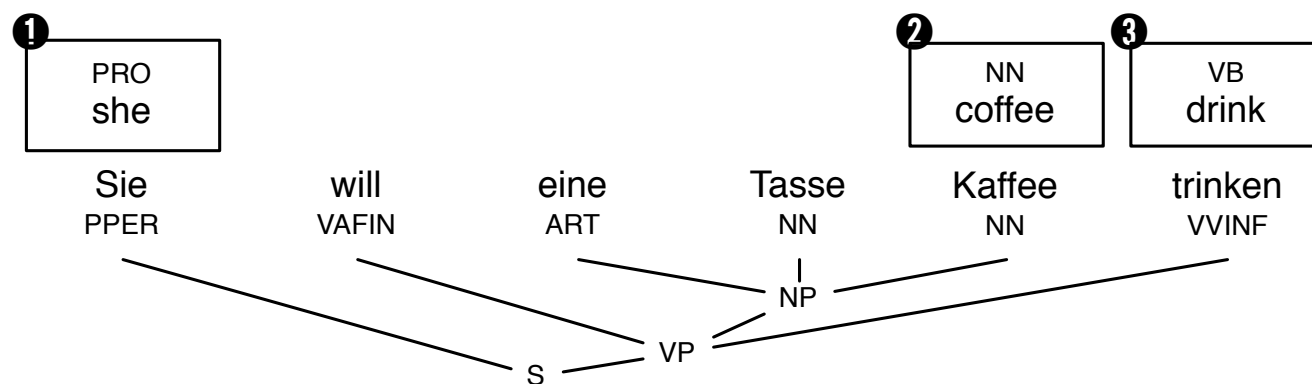
單純詞彙規則：涵蓋輸入詞 (Sie)，翻譯 (she)，填表 (PPER/PRO)

句法式解碼舉例 (3)



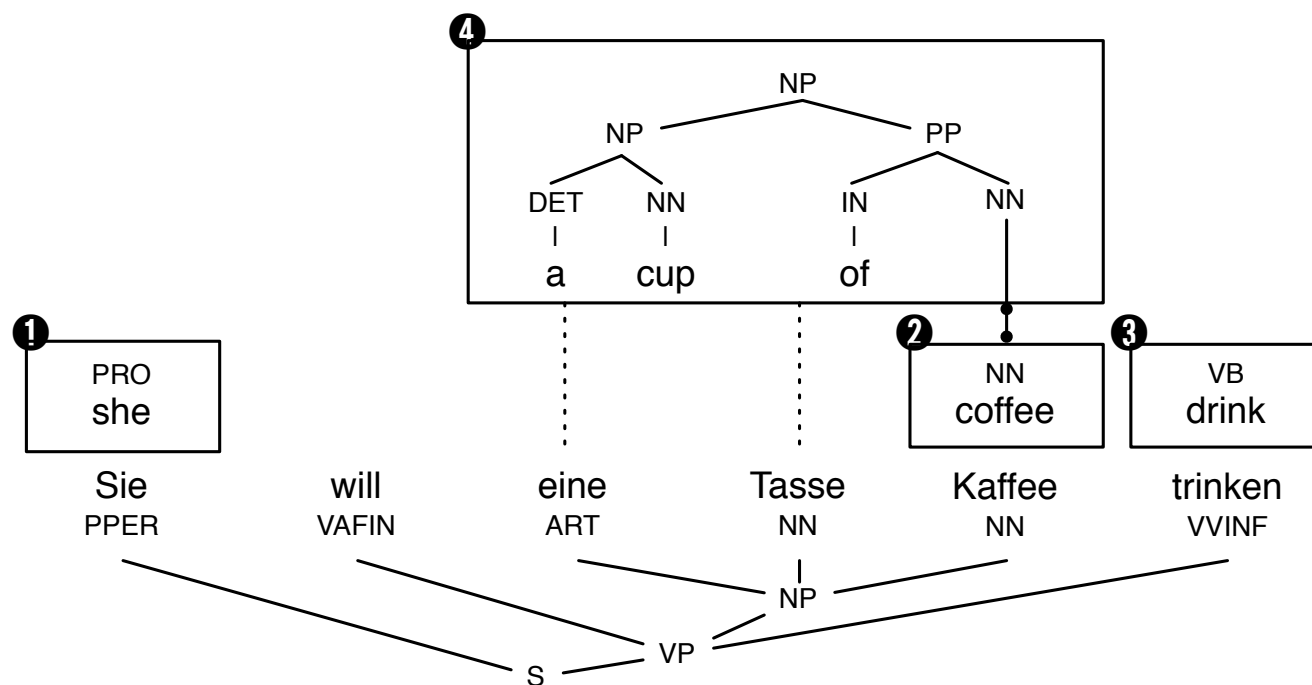
單純詞彙規則：涵蓋輸入詞 (Kaffe)，翻譯 (coffee)，填表 (NN/coffee)

句法式解碼舉例 (4)



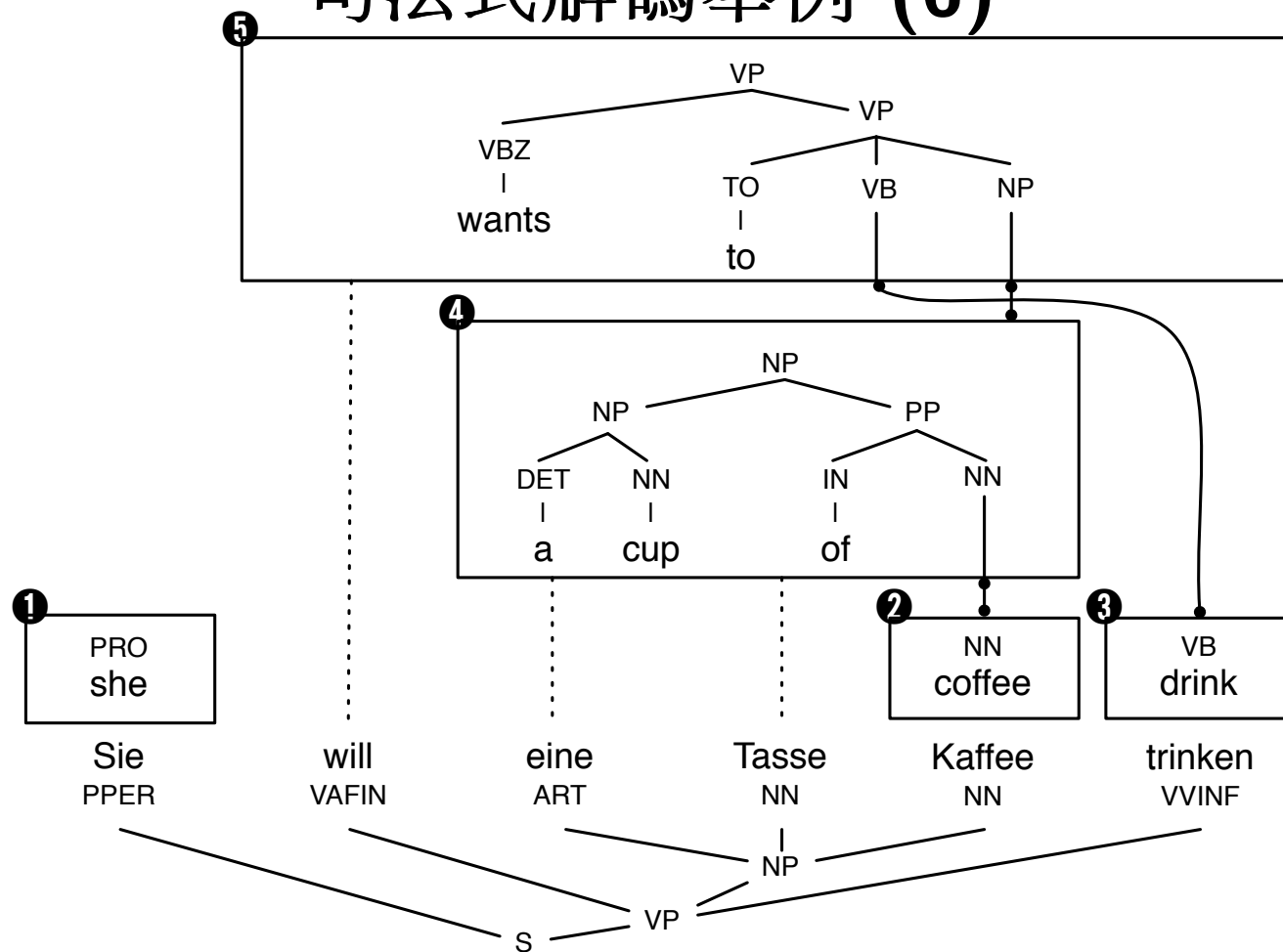
單純詞彙規則：涵蓋輸入詞 (trinken)，翻譯 (drink)，填表 (VB/drink)

句法式解碼舉例 (5)



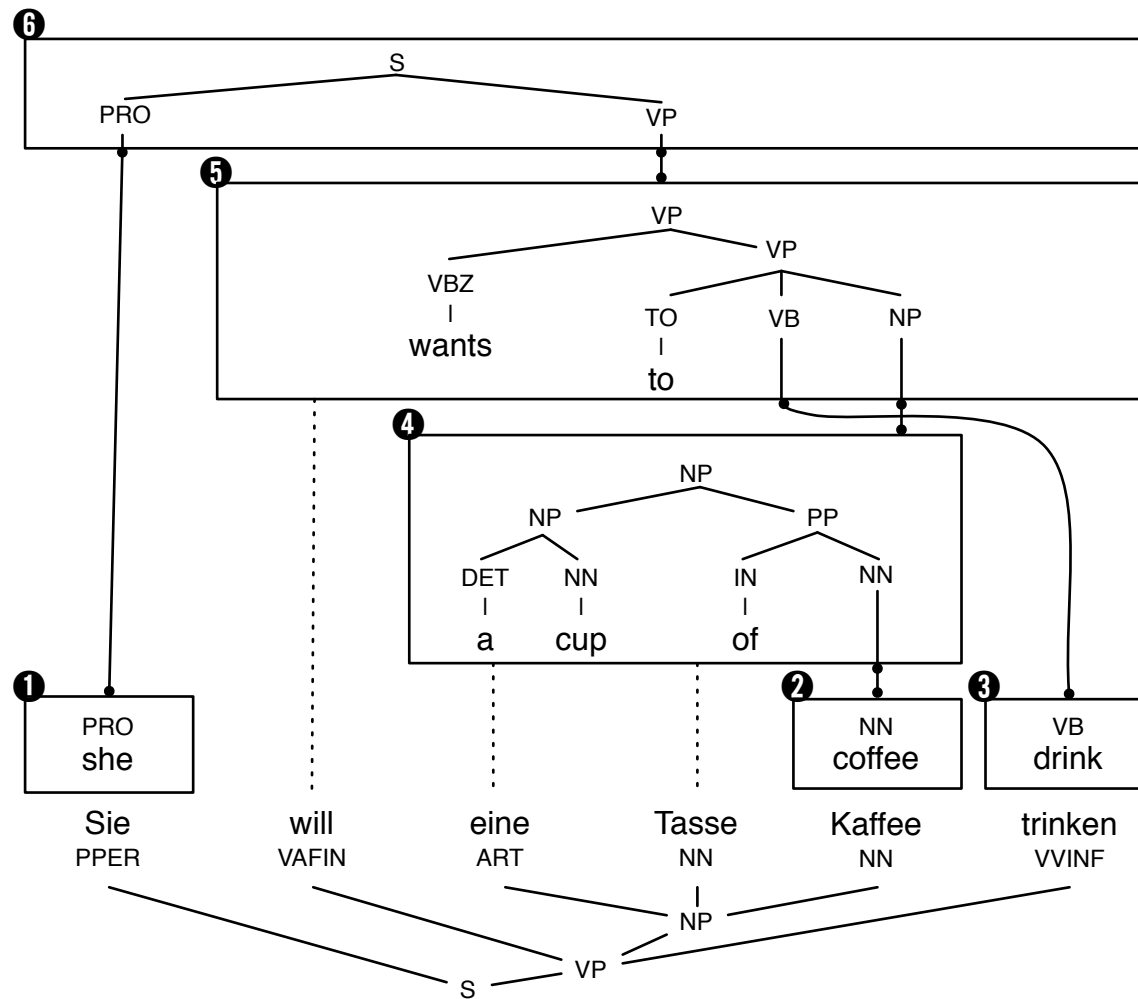
複雜規則：對齊下層成分 (NP) 並涵蓋一段文字 (eine Tasse/a cup of)

句法式解碼舉例 (6)



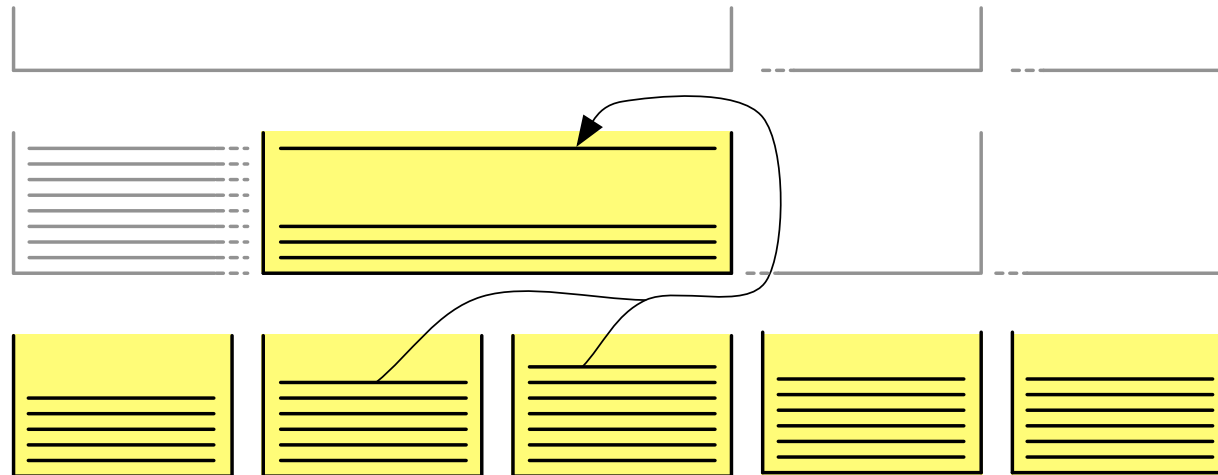
複雜規則：動詞詞序重組

句法式解碼舉例 (7)



由下而上的解碼

- 用階層式堆疊記錄每個片段 `span` 處理好的一組部分翻譯 (剖析)
- 用底層推疊，計算更長跨距的翻譯、剖析填入高階層堆疊
- 回顧：片語式模型是一系列推疊，而非階層堆疊



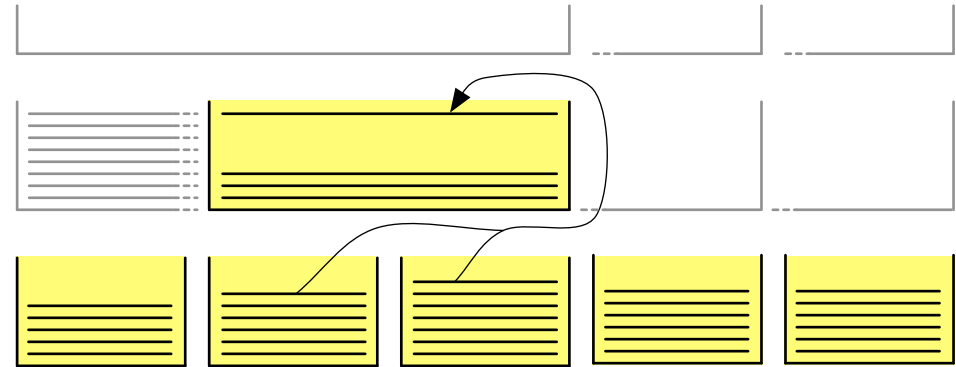
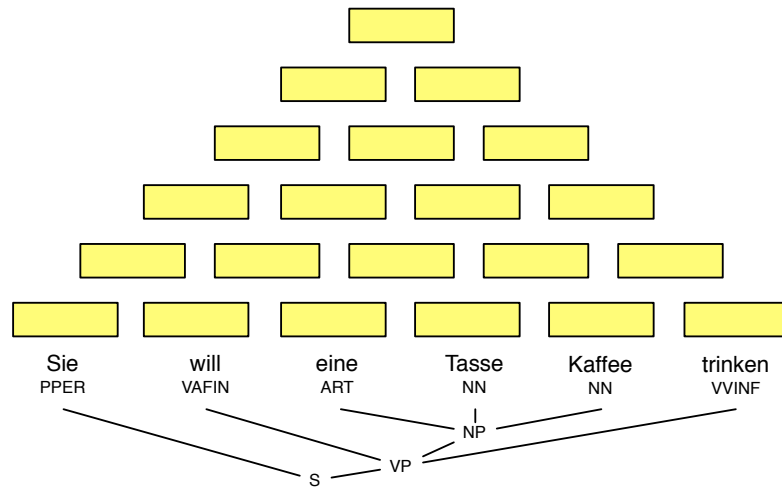
簡易演算法

Input: Foreign sentence $\mathbf{f} = f_1, \dots, f_{l_f}$, with syntax tree

Output: English translation \mathbf{e}

```
1: for all spans [start,end] (bottom up) do
2:   for all sequences  $s$  of hypotheses and words in span [start,end] do
3:     for all rules  $r$  do
4:       if rule  $r$  applies to chart sequence  $s$  then
5:         create new hypothesis  $c$ 
6:         add hypothesis  $c$  to chart
7:       end if
8:     end for
9:   end for
10: end for
11: return English translation  $\mathbf{e}$  from best hypothesis in span  $[0, l_f]$ 
```

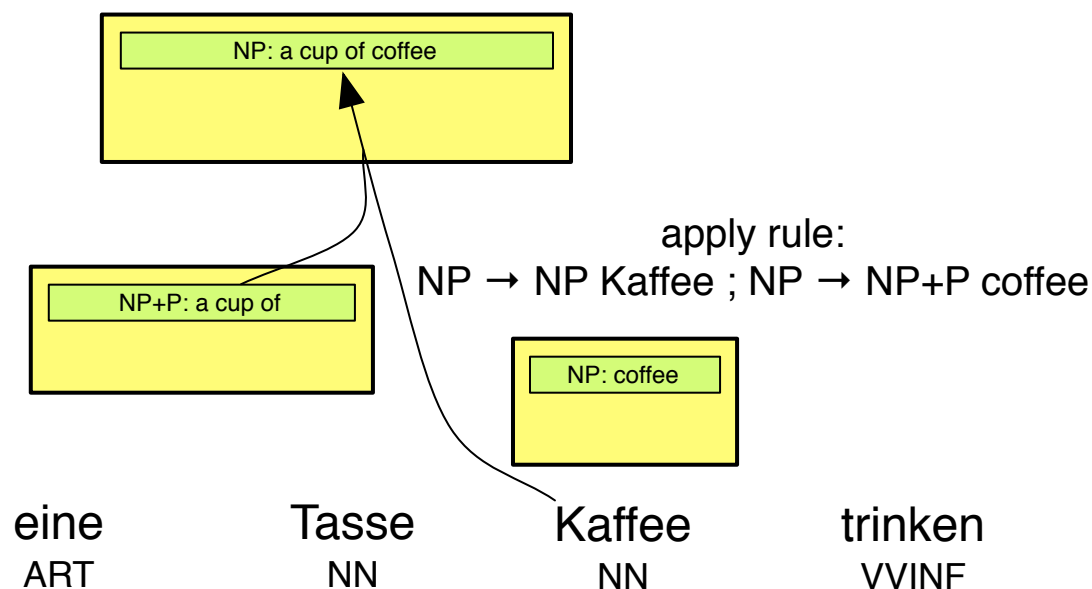
堆疊表的組織



- 剖析表 Chart 由格子組成，每格代表句子的跨距或連續片段
- 每個格子記錄一組「假設」（部分完成的剖析樹、翻譯）
- 假設 = 翻譯 + 木標語句法成分 (片語標籤)

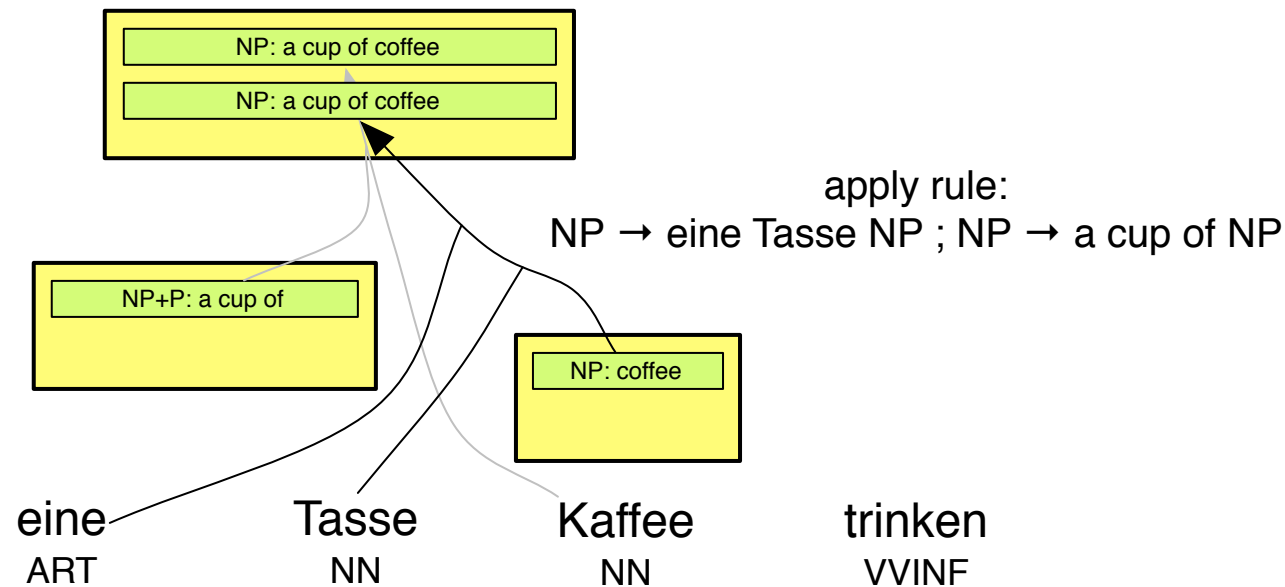
動態規劃

運用同步文法規則，把格子內容「縮減」更長的跨距成更高的句子成分 每個跨距只做一次，把結果存起來，不必反覆計算



動態規劃 (2)

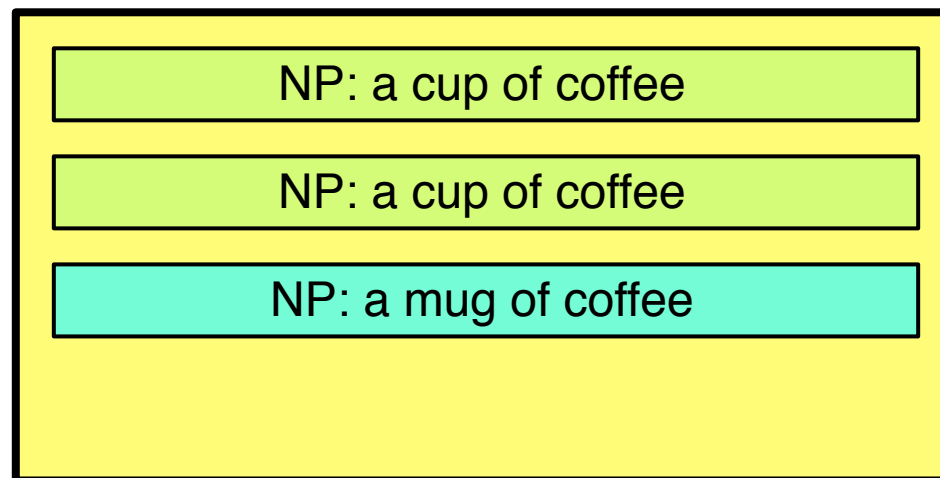
常常產生不只一個假設



如果這些假設的未來成本都是一樣
→ 可以合併 recombine，而不會導致「搜尋錯誤」

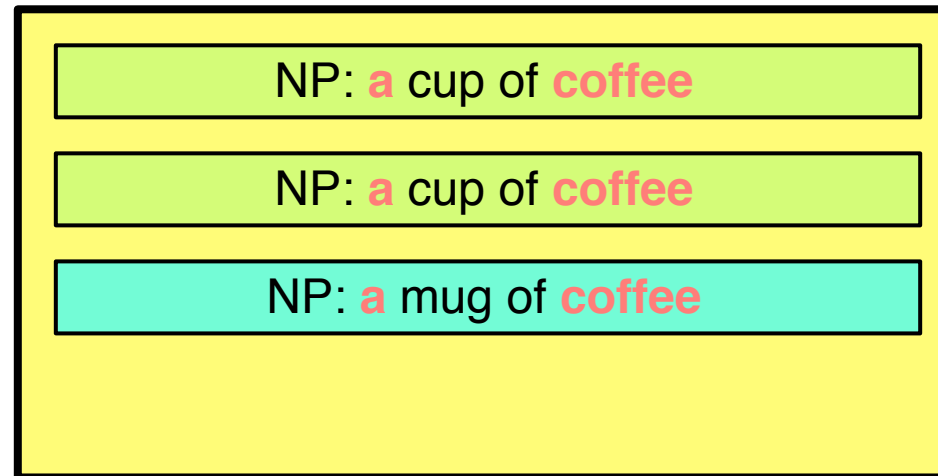
可以合併的假設或狀態

以下可以合併，留下最佳的一個？



可合併假設或狀態

是否可以合併?



可以。若且唯若：只使用 2-gram 語言模型 (最前、最後詞一樣就可以了)

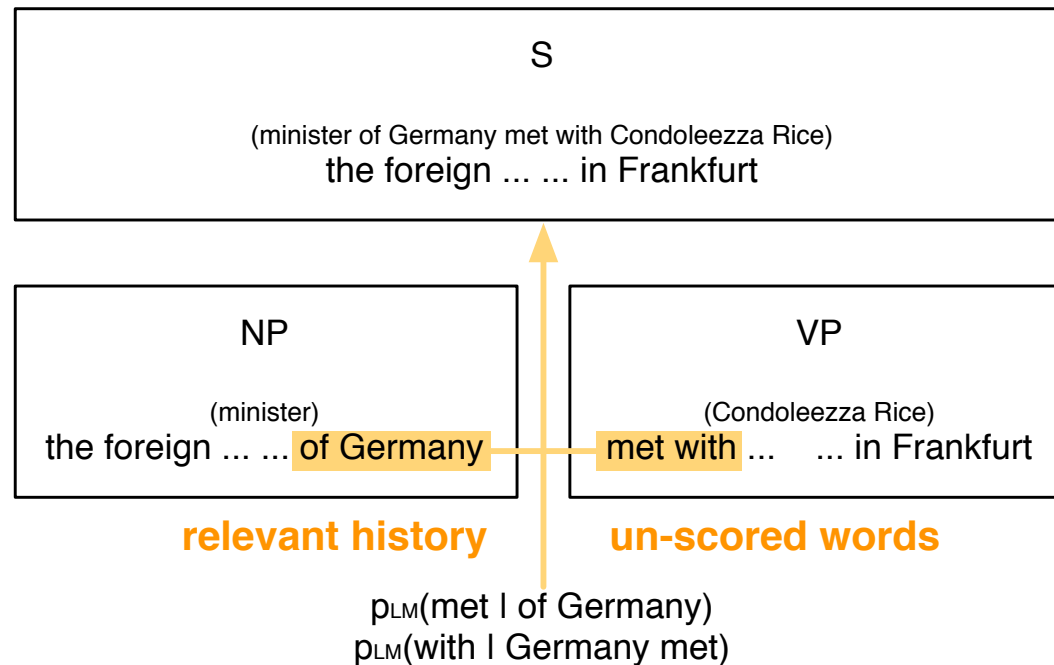
可合併假設或狀態

如果兩個「假設」合乎以下條件，就可以擇優合併

- 相同的跨距
- 相同的輸出成分標籤
- 前 $n-1$ 輸出詞一致
- 後 $n-1$ 輸出詞一致 (n 是語言模型的等級)

語言模型

當合併假設時，內部的語言模型已經被吸收，不再有影響



堆疊修剪 Stack Pruning

- 每個剖析表的格子內的「假設」數量暴增

⇒ 需要拋棄一些不好的假設
例如，保留最好的 100 個假設

- 也可以考慮輸出成分 (片語，如 NP) 不同，就安排各自的堆疊
- 修剪所需的成本估計
 - 翻譯模型的成本為已知
 - 內部詞彙的語言模型的成本也為已知
 - 外部成本，又如何估算?
(某一假設涵蓋 3–5 詞，輸出 NP 後來對整句的翻譯是否有用?)

簡易的演算法導致計算量過高

- 過多的跨距內的假設和詞彙的組合需要計算

for all sequences s of hypotheses and words in span $[start, end]$

- 太多規則，需要檢查是否適用

for all rules r

- 檢查規則是否適用，也並非簡單 (有太多組合，花太多時間—不像 CYK)

if rule r applies to chart sequence s

⇒ 不可行

解決之道

- 用前綴樹的資料結構來儲存規則
- 點規則 dotted rules (在 RHS 有一個點，表示 RHS 處理到的位置)
- 立方修剪 (cube pruning)–不組合，而是各取最佳組合，然後再取其鄰居-次加和最佳，然後次佳+次佳，等等)

如何儲存規則 Storing Rules

- 首先要關心：規則是否適用於此一跨距片段？
→ 必須考慮既有假設和輸入詞的組合，是否適用於此一規則
- 規則的例子

$NP \rightarrow X_1 \text{ des } X_2 \mid NP_1 \text{ of the } NN_2$

- 如何檢查
 - 有沒有一段標示為 **NP** 的假設？
 - 有沒有著詞 **des**？
 - 有沒皆這一段標示為 **NN** 的假設？
- 相關資訊的序列 Sequence of relevant information

NP • des • NN 或者 **NP₁ of the NN₂**

檢查規則是否適用 (1)

試試看 6 個詞是否適用以下規則

NP • des • NN → NP: NP of the NN

das Haus des Architekten Frank Gehry

檢查規則是否適用 (2)

首先: 檢查輸出標籤 NP

NP • des • NN → NP: NP of the NN



das

Haus

des

Architekten

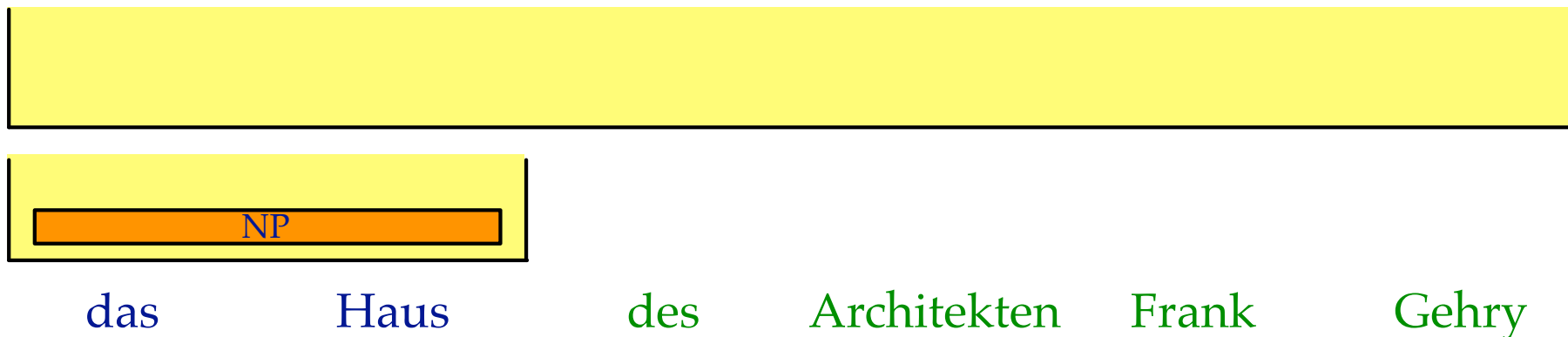
Frank

Gehry

檢查規則是否適用 (3)

找到了 NP 解設：對到規則的一個符號

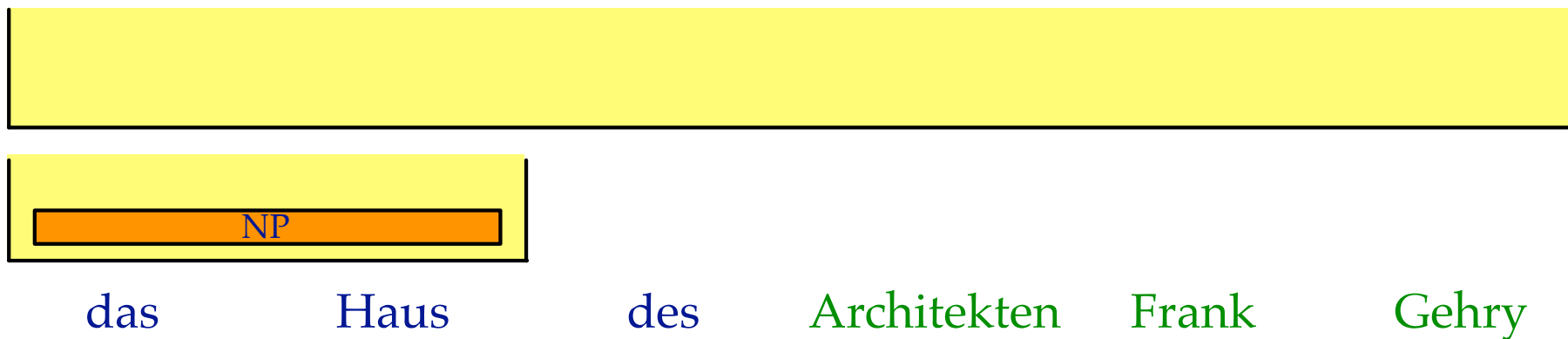
NP • des • NN → NP: NP of the NN



檢查規則是否適用 (4)

接著對到了規則的第二個符號 *des* 這個詞

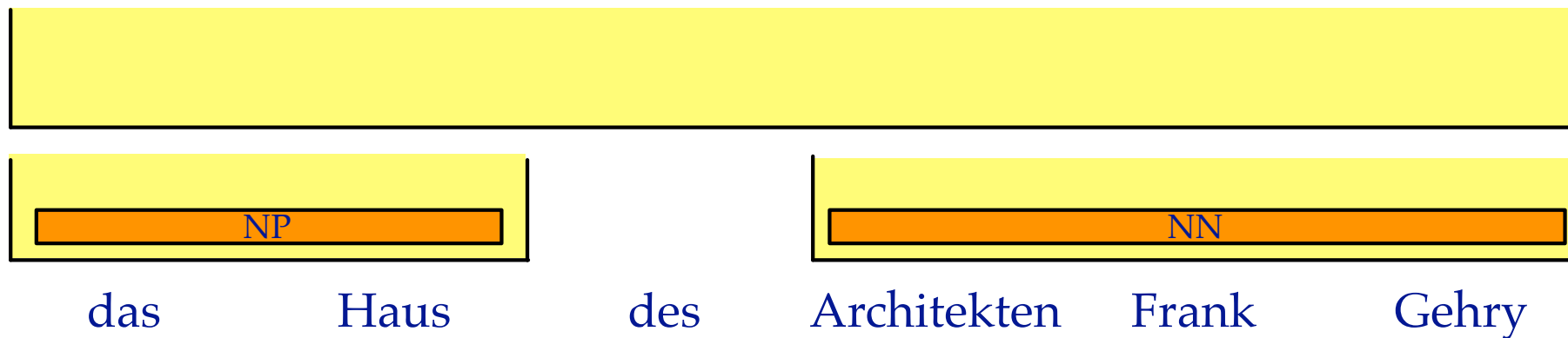
$NP \bullet des \bullet NN \rightarrow NP: NP \text{ of the } NN$



檢查規則是否適用 (5)

接著對到了規則的最後符號 **NN** 解設

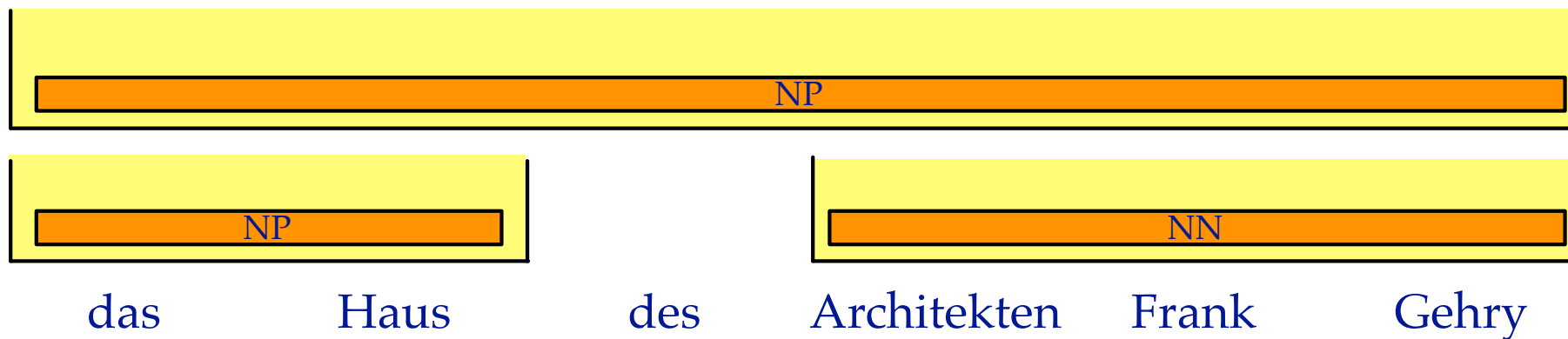
NP • des • NN → **NP: NP of the NN**



檢查規則是否適用 (6)

對到整條規則，所以就產生 LHS **NP** 的輸出標籤

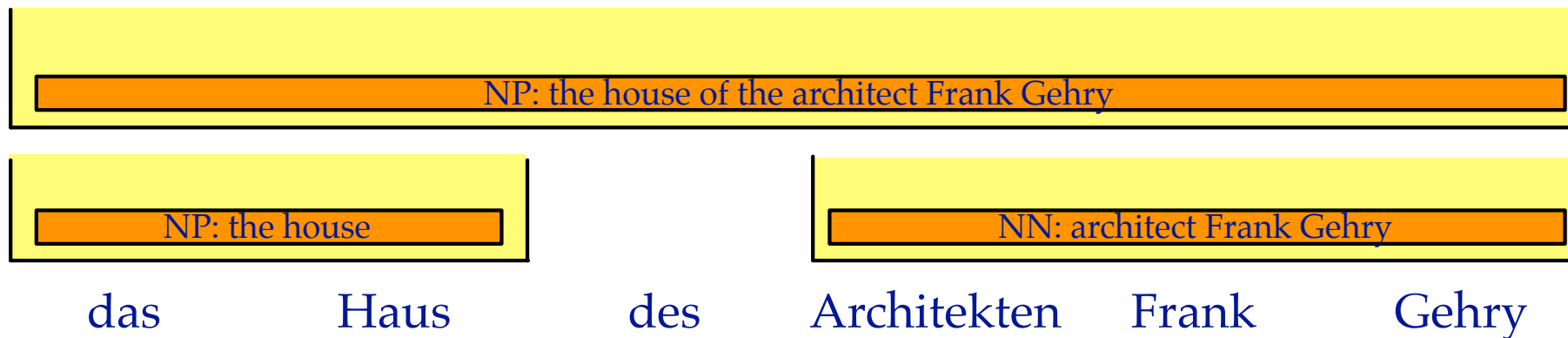
NP • des • NN → NP: NP of the NN



檢查規則是否適用 (7)

找到「輸出詞」和標籤合在一起來產生假設
(註記: 可能有不只一組 NP 和 NN 的下層假設)

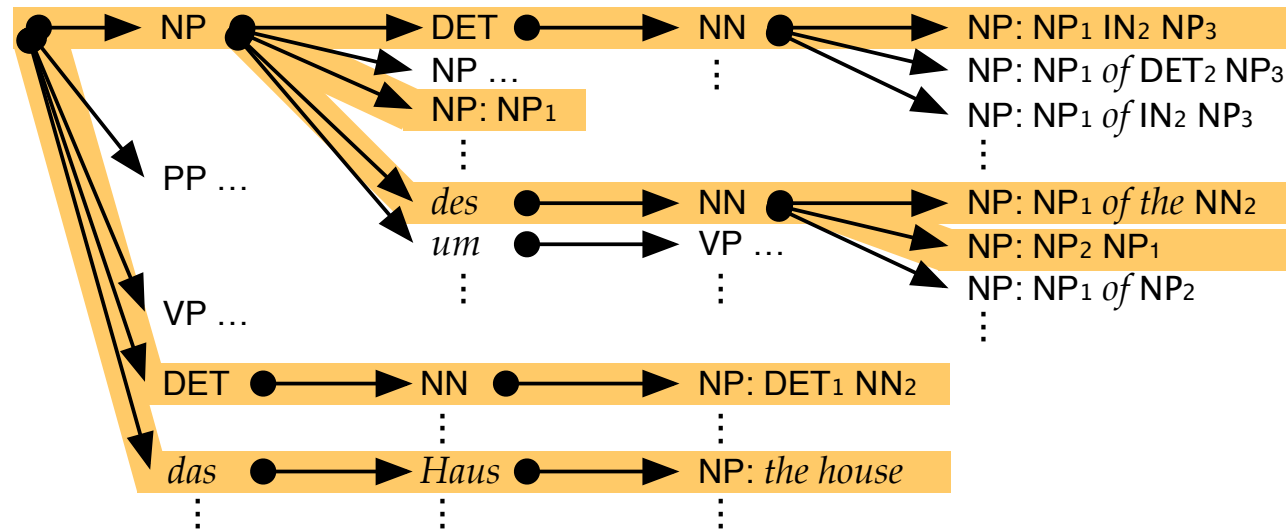
NP • des • NN → NP: NP of the NN



Checking Rules vs. Finding Rules

- What we showed:
 - given a rule
 - check if and how it can be applied
- But there are too many rules (millions) to check them all
- Instead:
 - given the underlying chart cells and input words
 - find which rules apply

規則的前綴樹



黃色重點的規則如下：

$NP \rightarrow NP_1 \text{ DET}_2 \text{ NN}_3 \mid NP_1 \text{ IN}_2 \text{ NN}_3$
 $NP \rightarrow NP_1 \mid NP_1$
 $NP \rightarrow NP_1 \text{ des } \text{NN}_2 \mid NP_1 \text{ of the } \text{NN}_2$
 $NP \rightarrow NP_1 \text{ des } \text{NN}_2 \mid NP_2 \text{ NP}_1$
 $NP \rightarrow \text{DET}_1 \text{ NN}_2 \mid \text{DET}_1 \text{ NN}_2$
 $NP \rightarrow \text{das Haus} \mid \text{the house}$

點規則的主要概念

- 如果如下的規則，可以適用於現在的剖析表的某一個點的跨距 chart

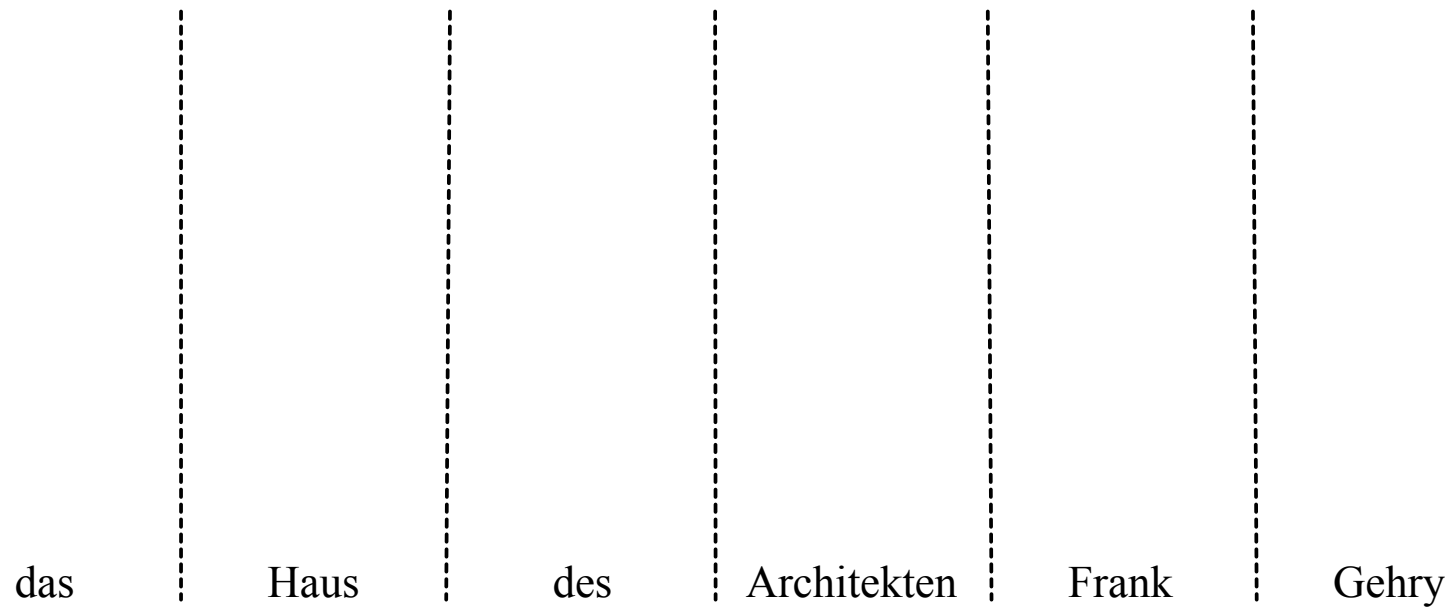
$$p \rightarrow A B C \mid x$$

- 那麼，有些規則也可能適用此一點開始的跨距（較小的跨距）如

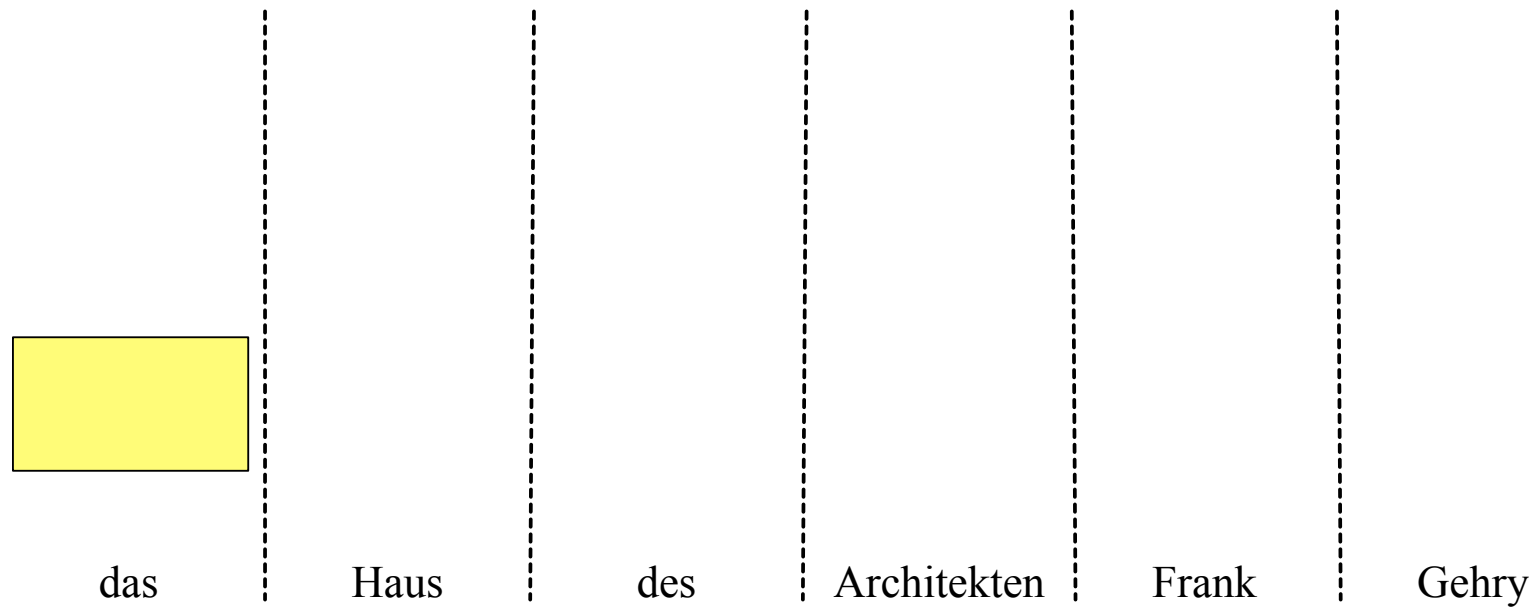
$$q \rightarrow A B \mid y$$

⇒ 所以，如果我們計算了以下的點規則，就可以節省查詢規則的時間（查一次得到所有適用規則，如 $A B \bullet$ (點規則)

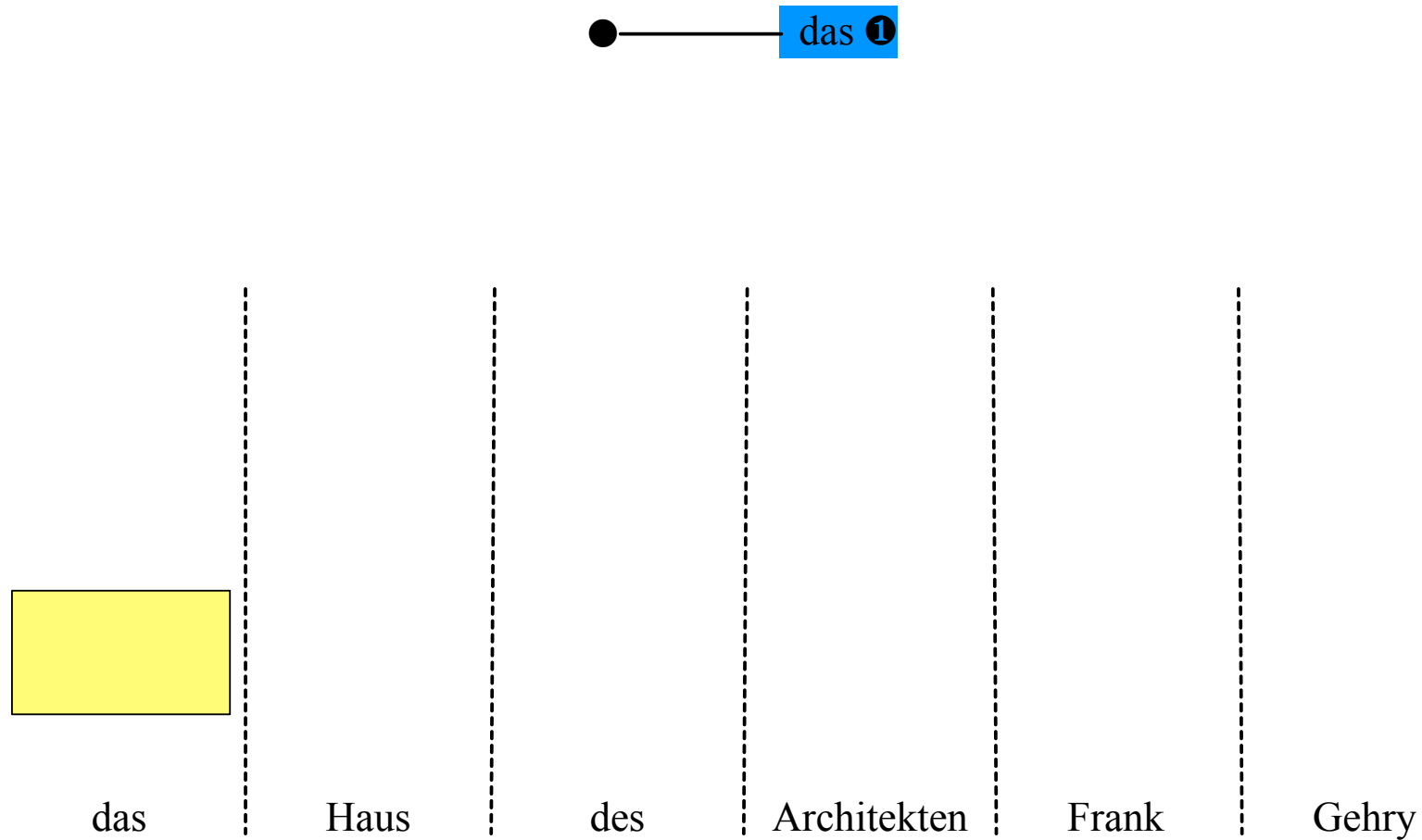
在前綴樹查詢適用規則



涵蓋第一個格子



查格子在前綴樹的位置



記錄點規則

● — das ①

das ①

das

Haus

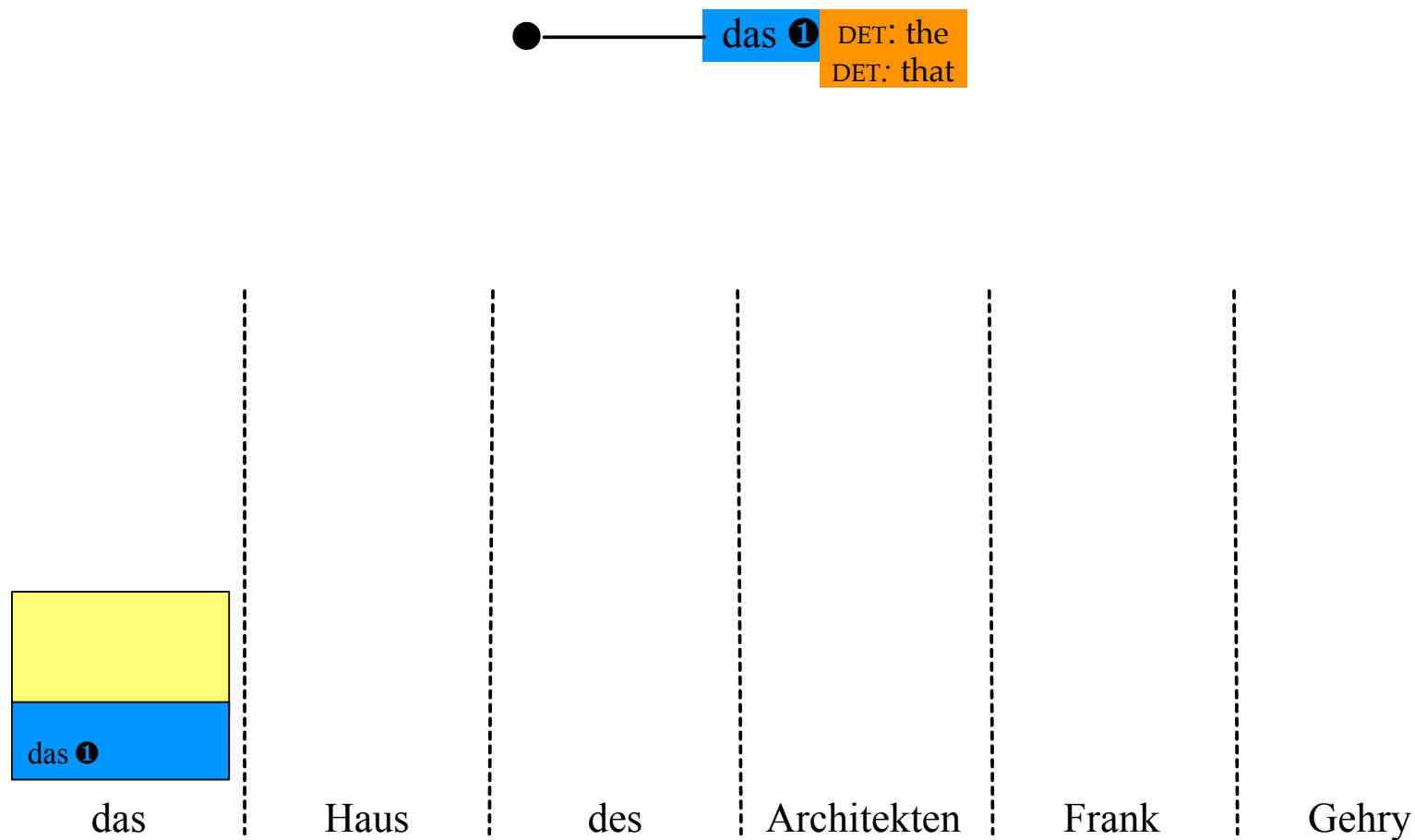
des

Architekten

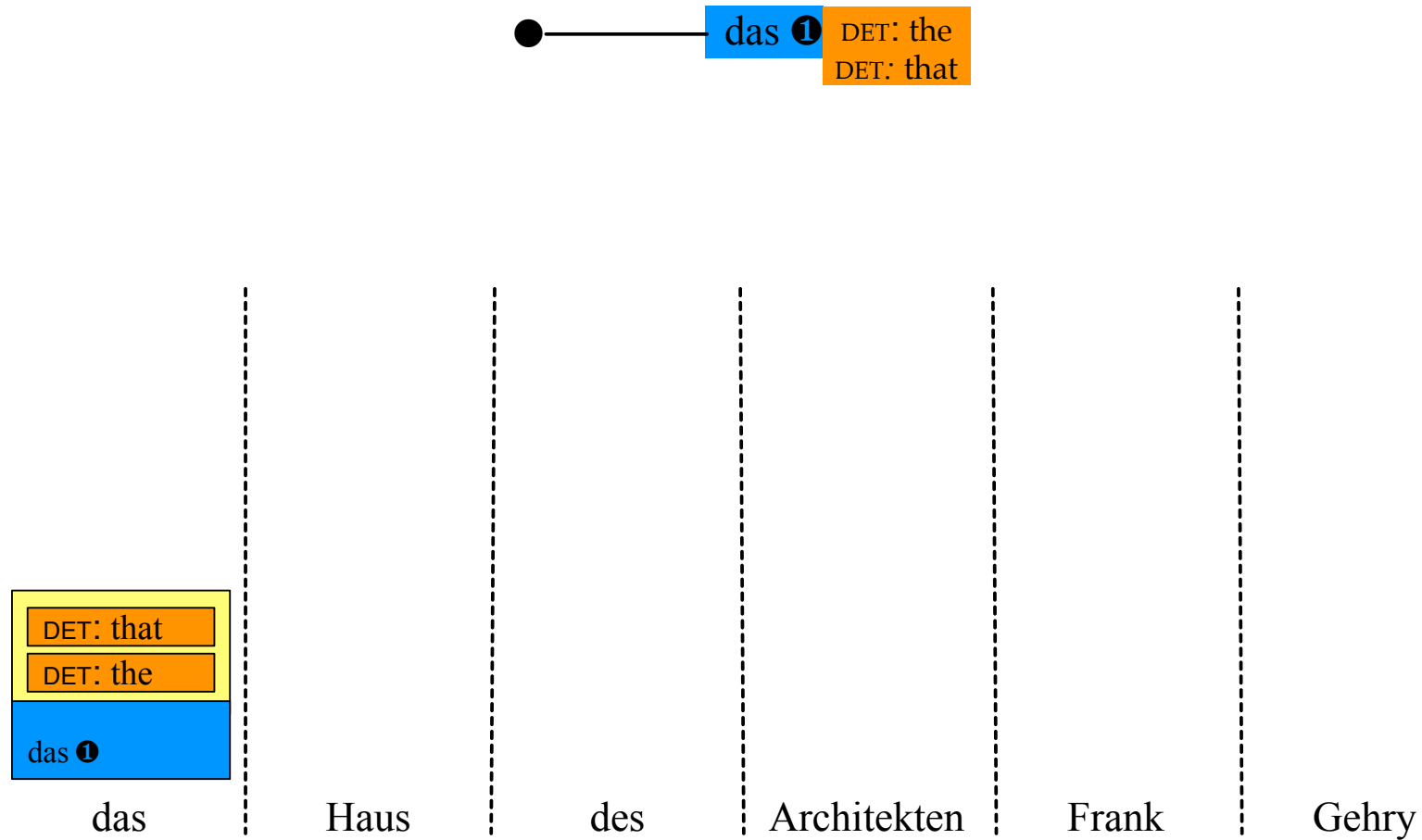
Frank

Gehry

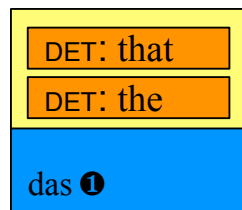
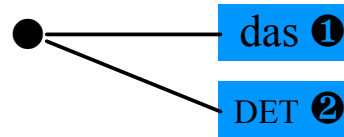
檢查點規則是否有翻譯的部份



實施翻譯規則



查詢成分標籤 (DET)



das

Haus

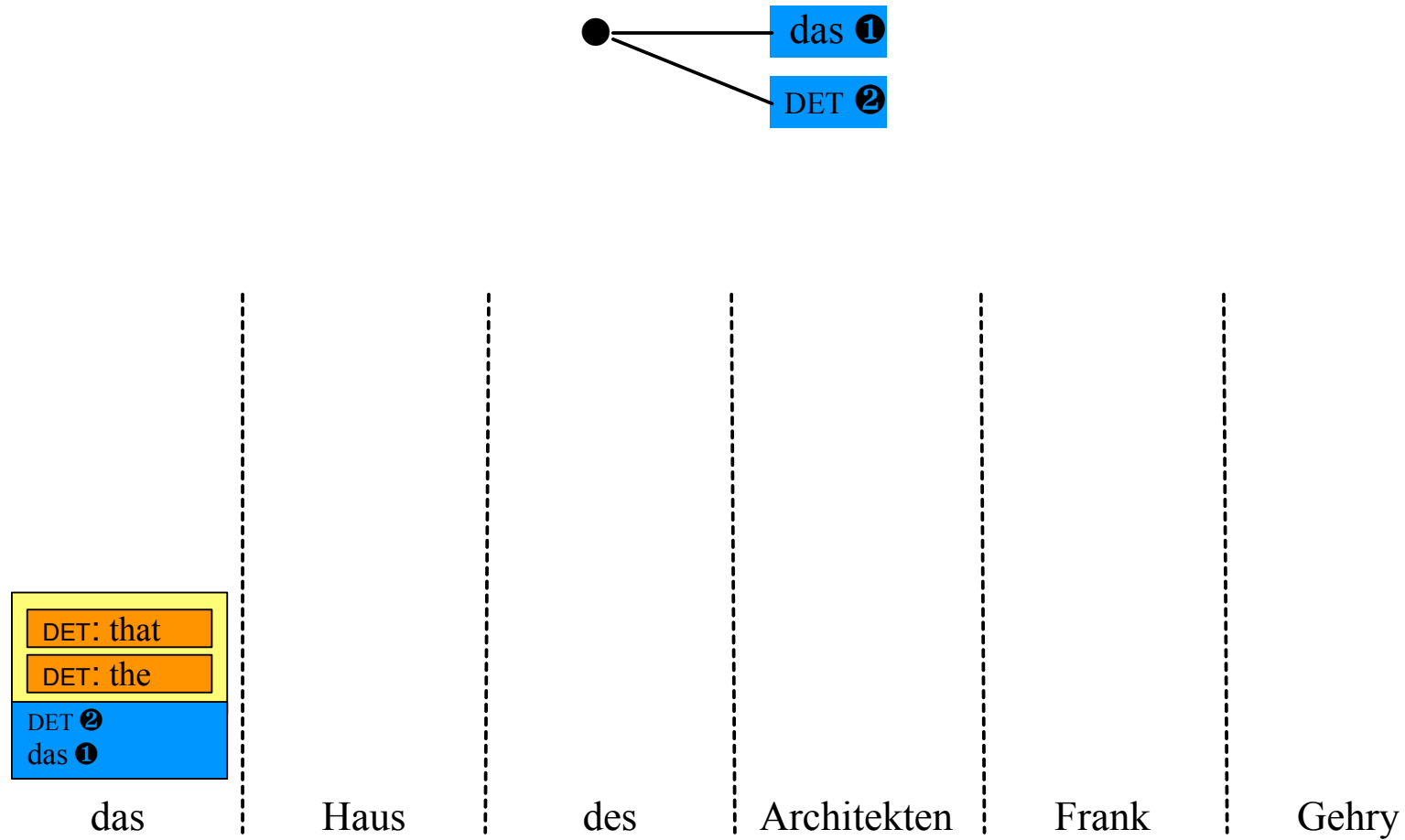
des

Architekten

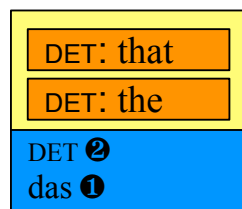
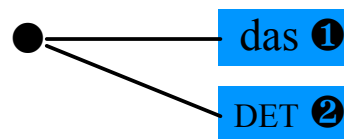
Frank

Gehry

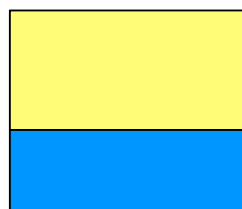
加到格子的點規則清單



移動到下一格



das



Haus

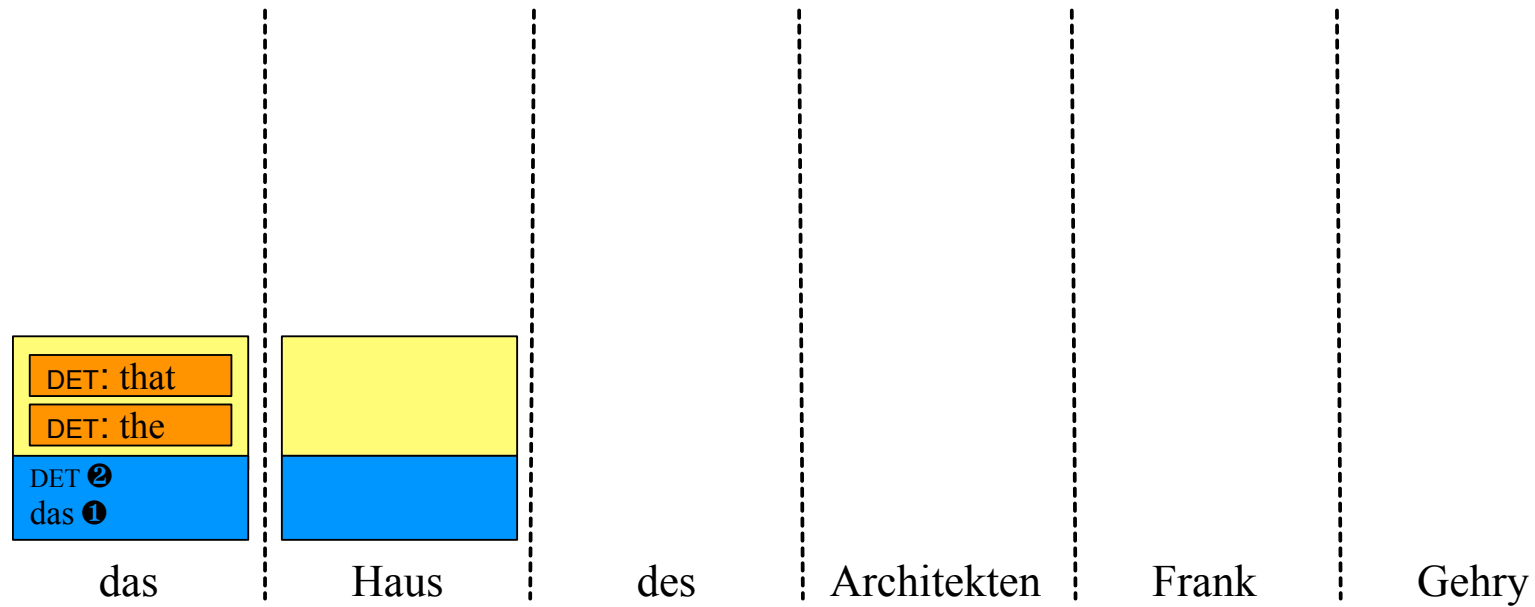
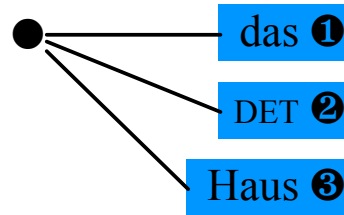
des

Architekten

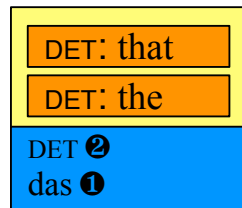
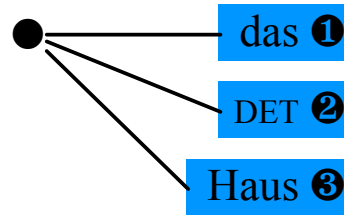
Frank

Gehry

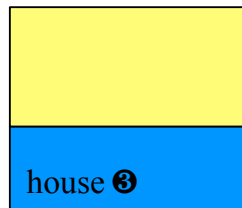
在前綴樹查詢規則



記錄點規則



das



Haus

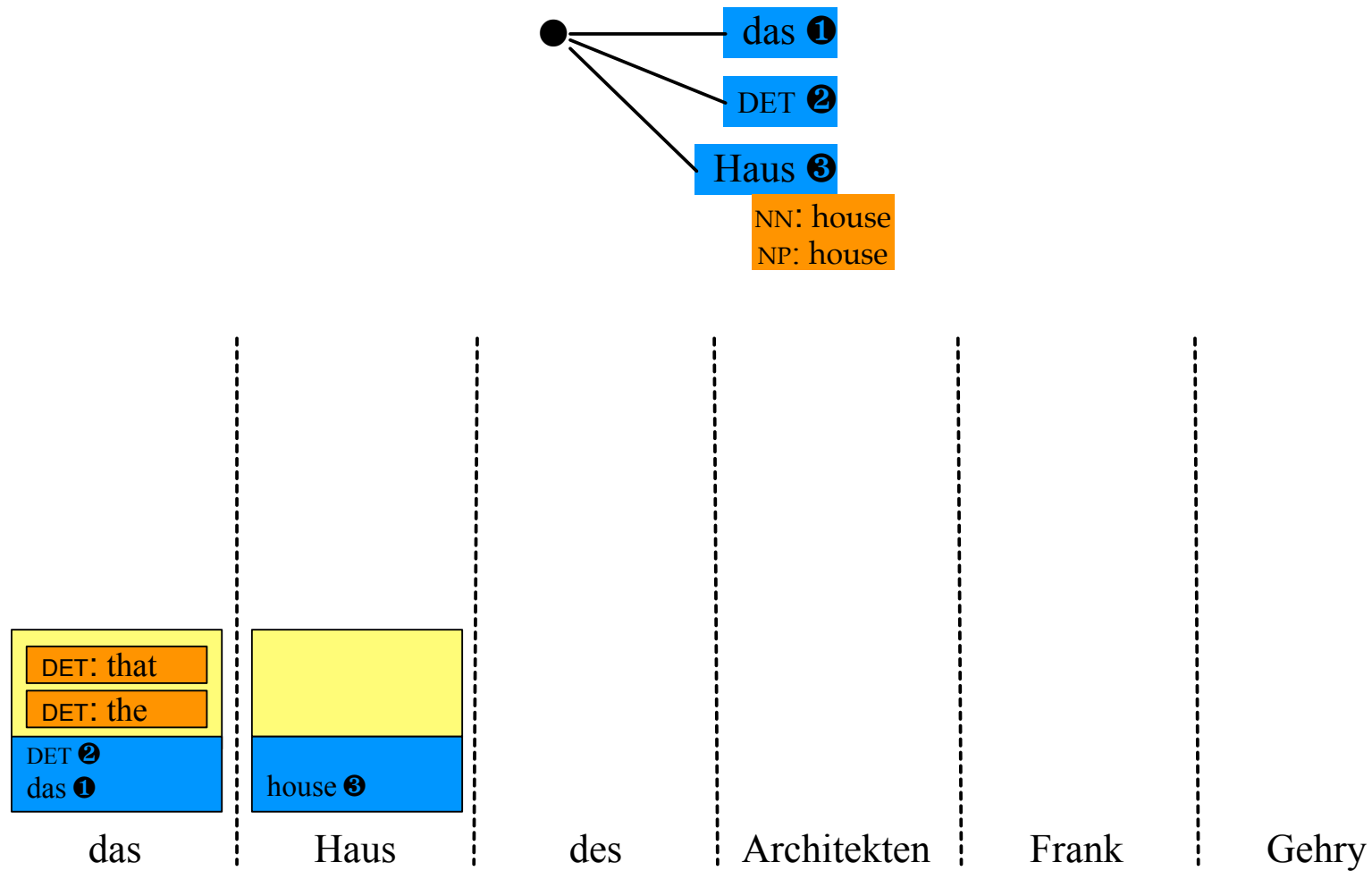
des

Architekten

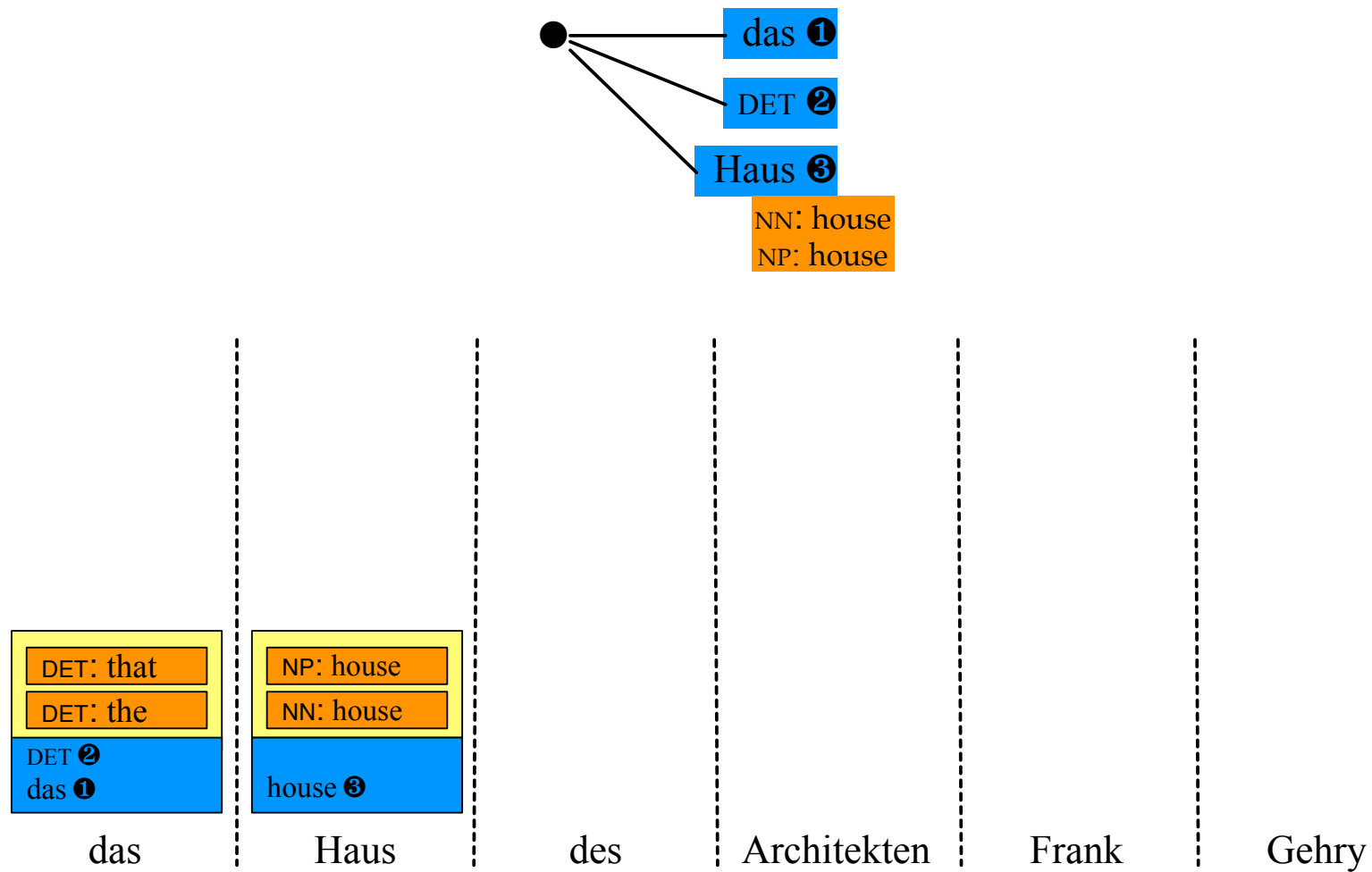
Frank

Gehry

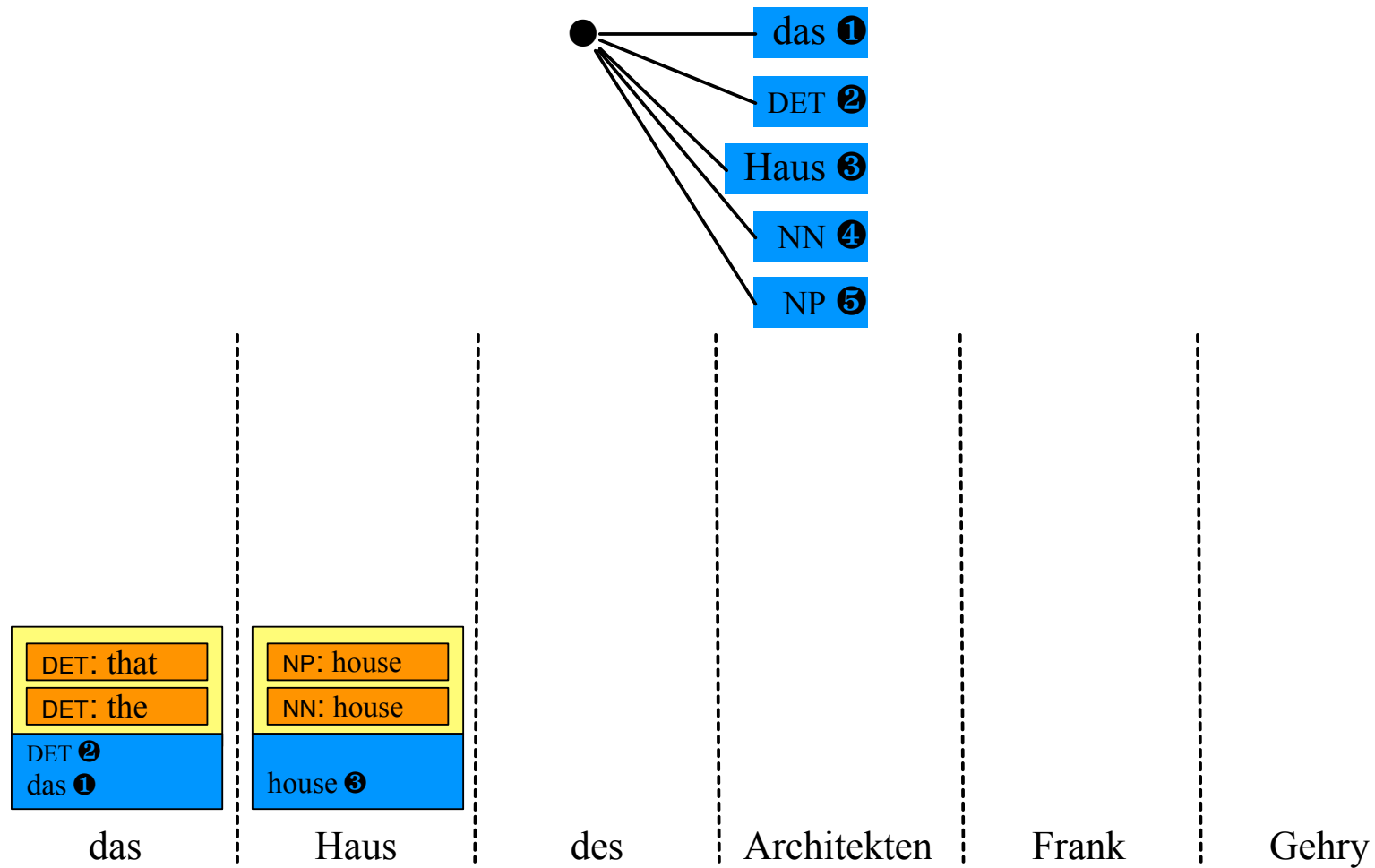
檢查點規則是否有翻譯



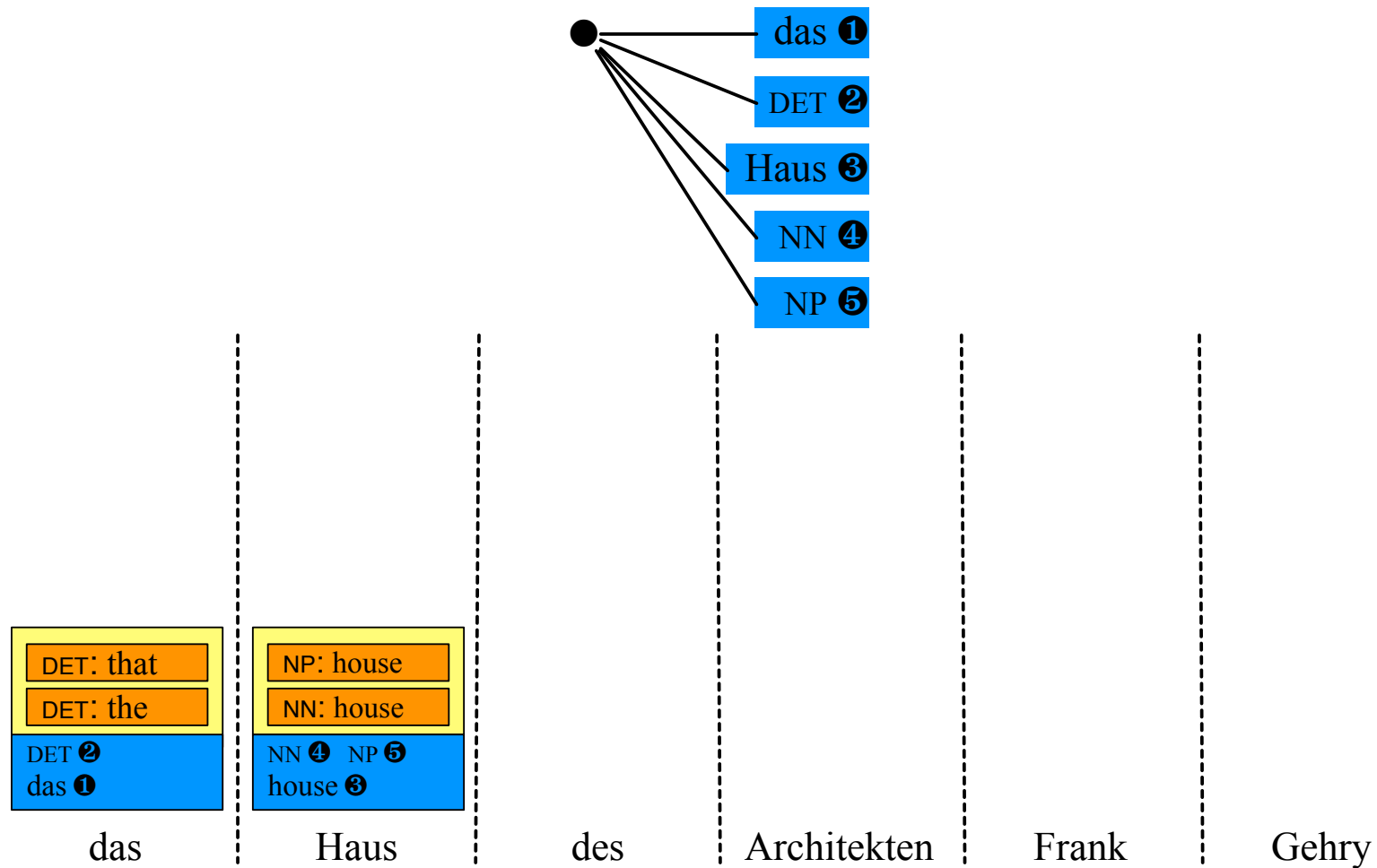
實施翻譯規則



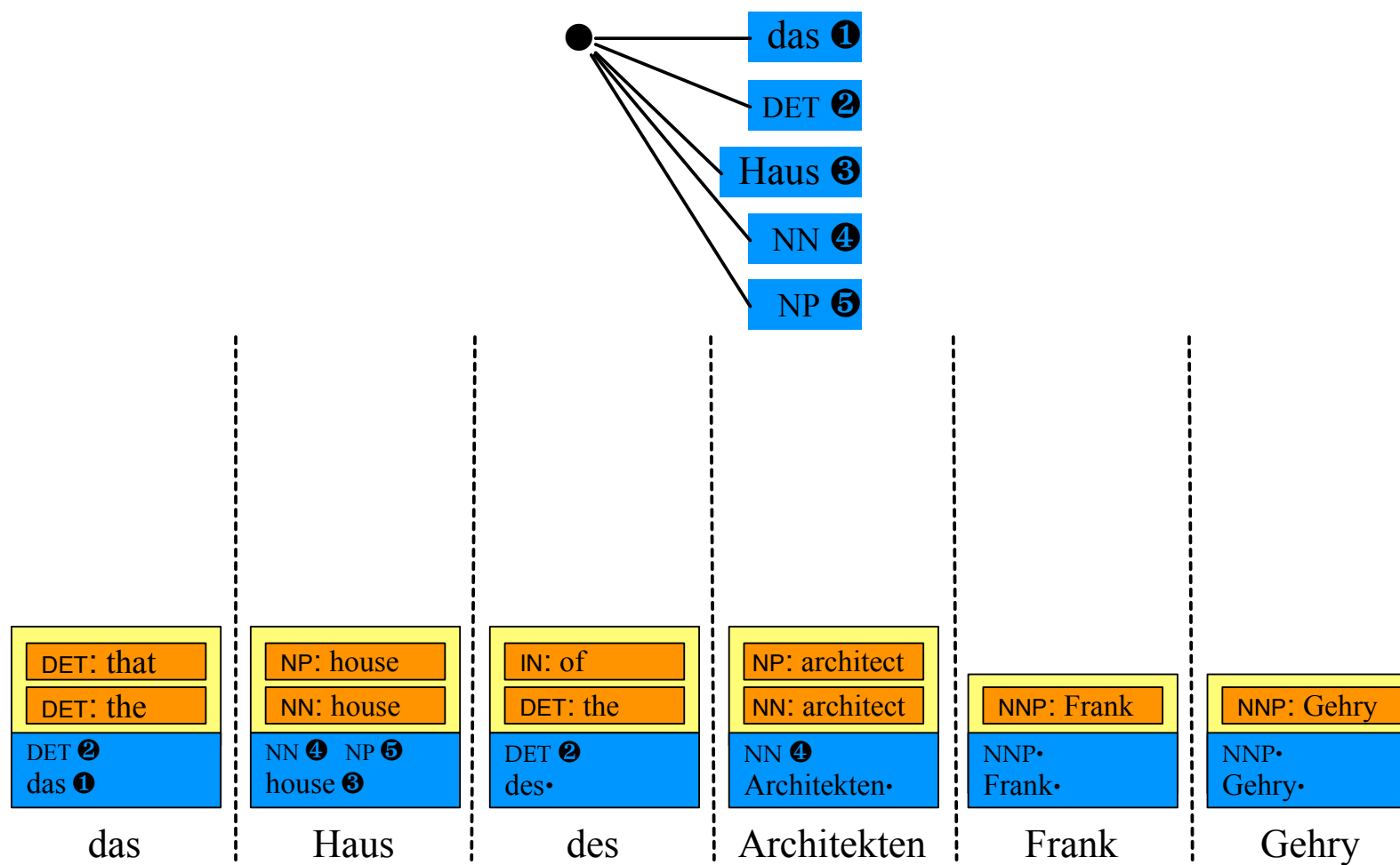
在前綴樹查詢規則



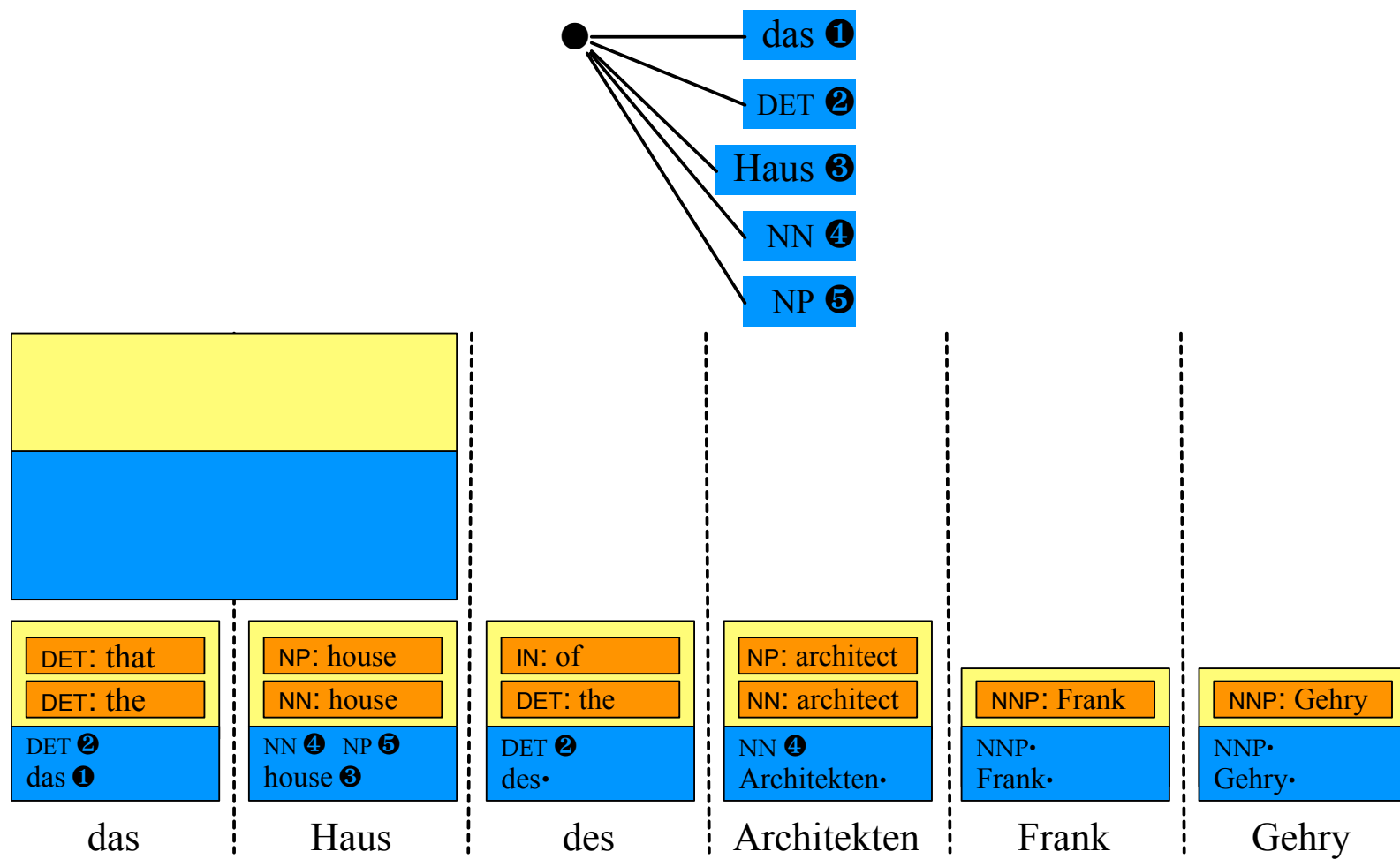
加到格子的點規則清單



大約如此，反覆操作

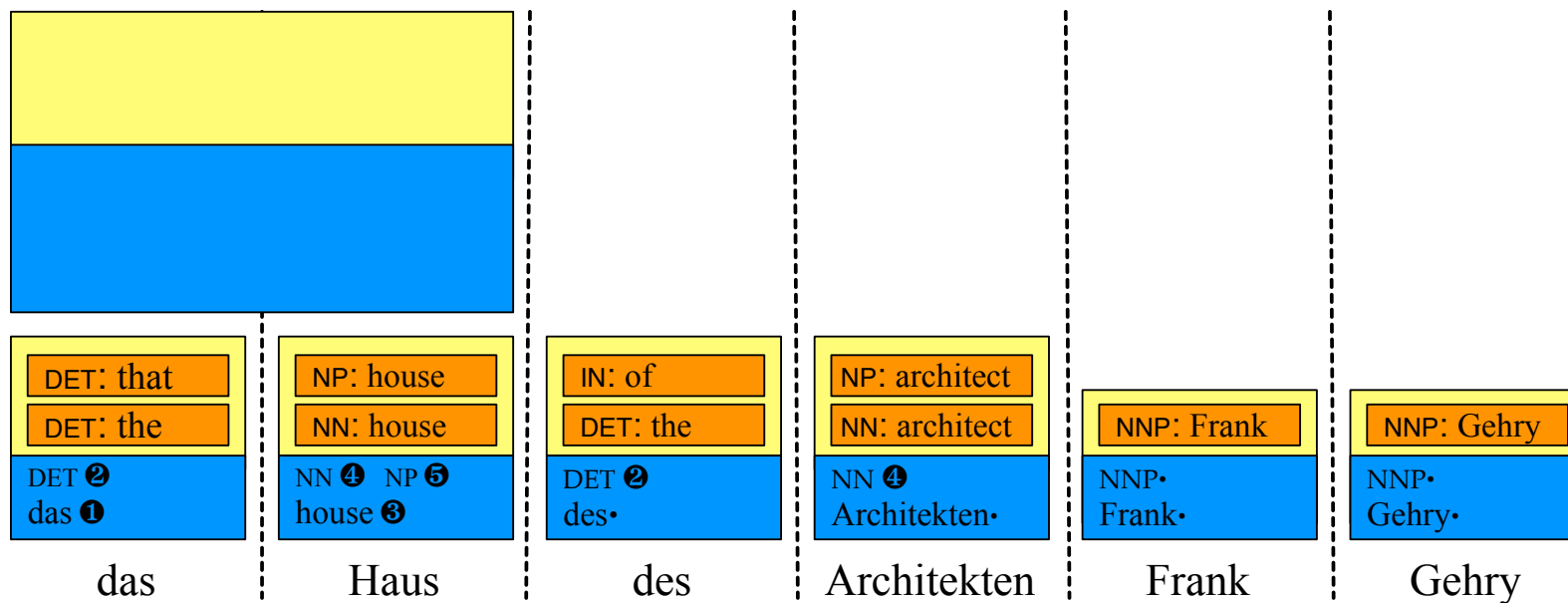


移動到下一格

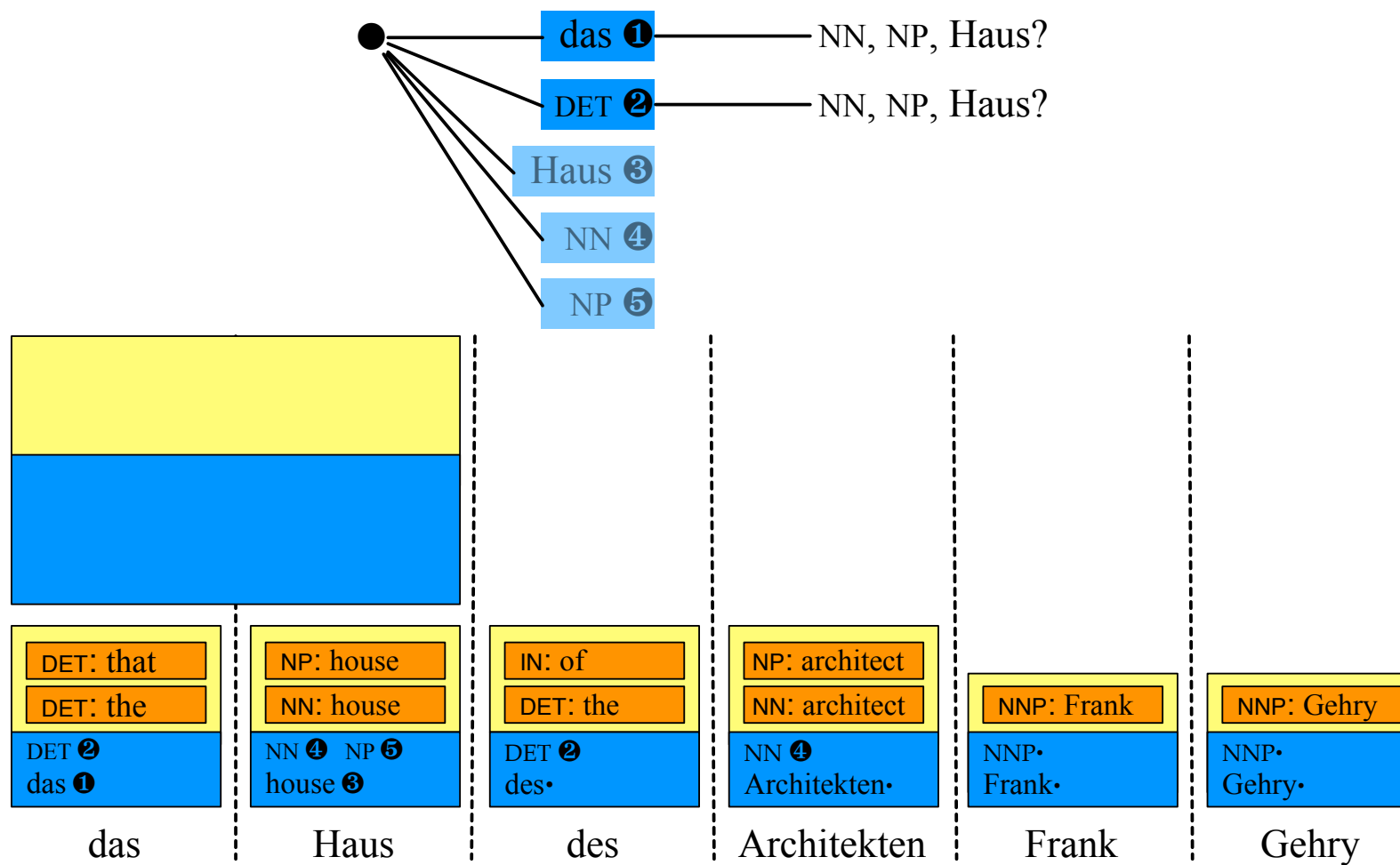


涵蓋更大的跨距

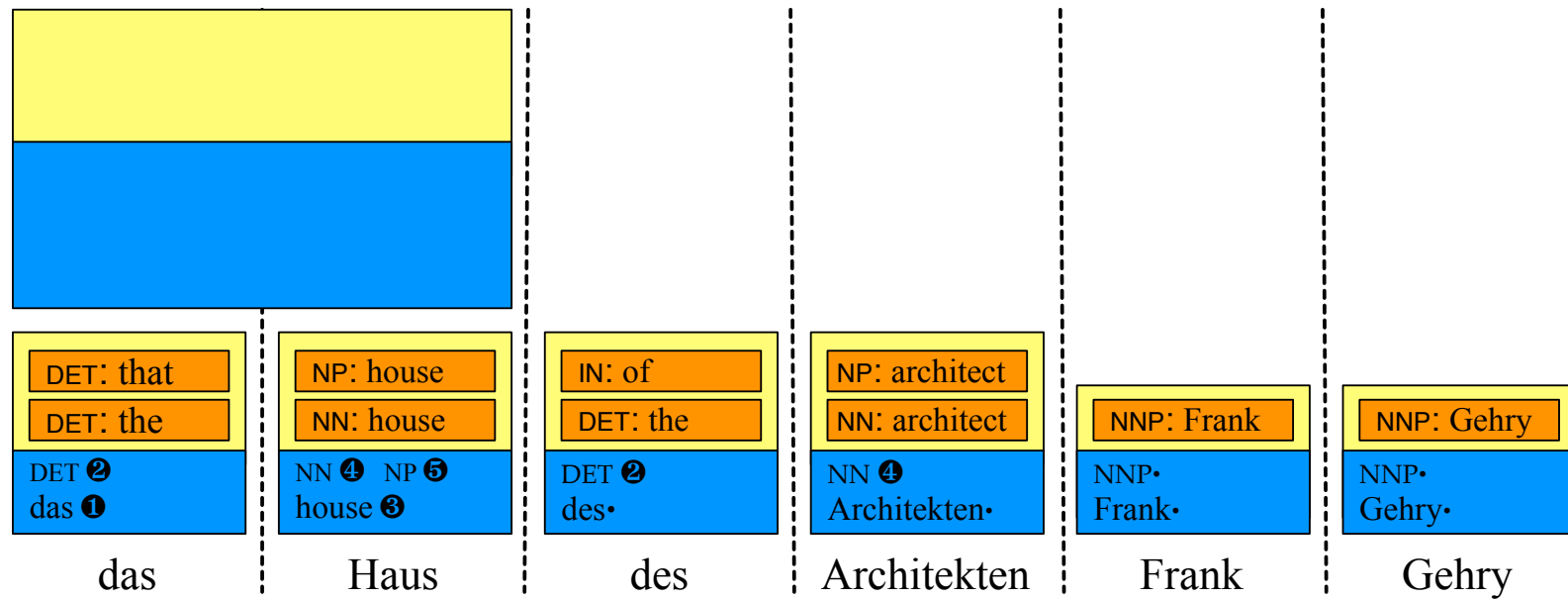
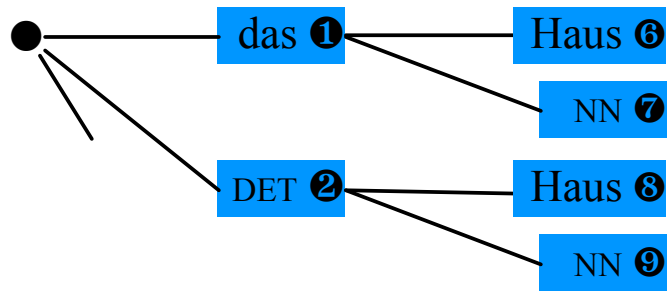
Cannot consume multiple words at once
All rules are extensions of existing dotted rules
Here: only extensions of span over **das** possible



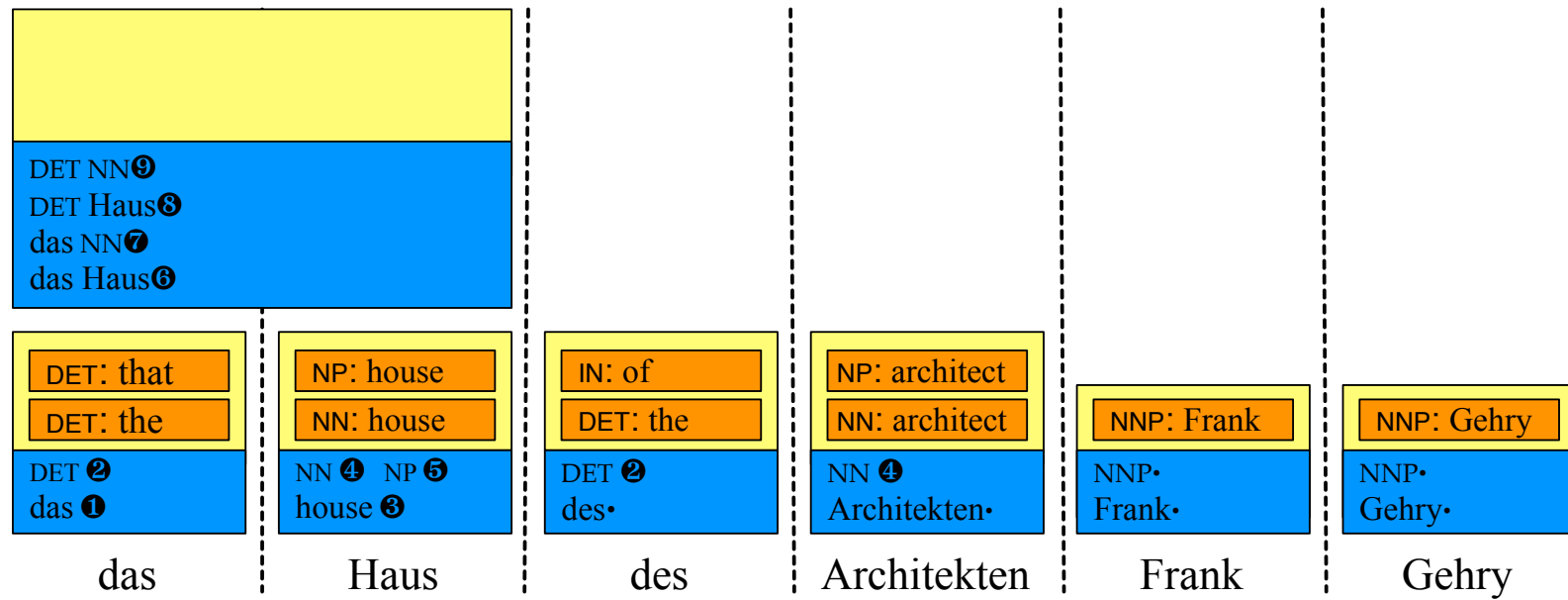
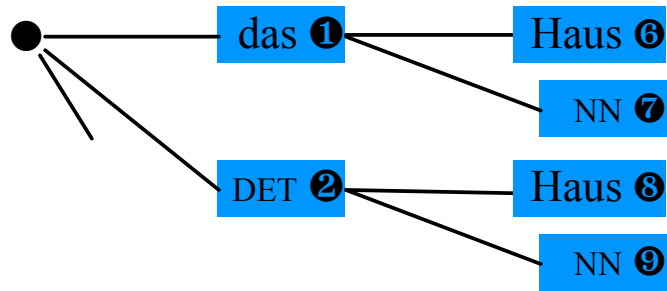
延長跨距以涵蓋 das



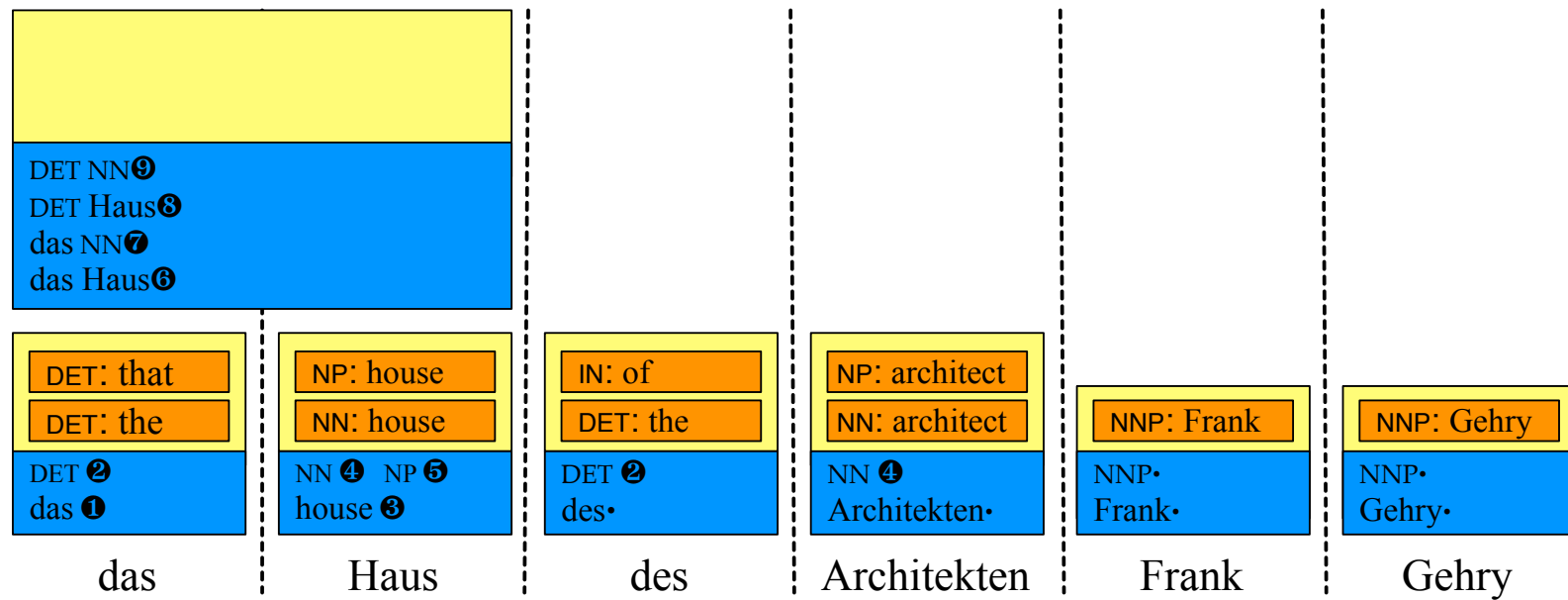
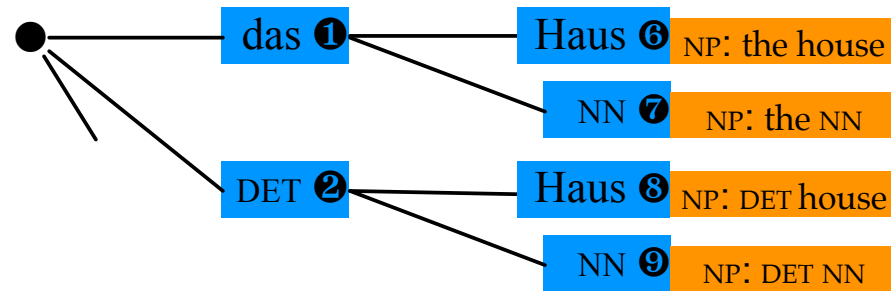
在前綴樹查詢規則



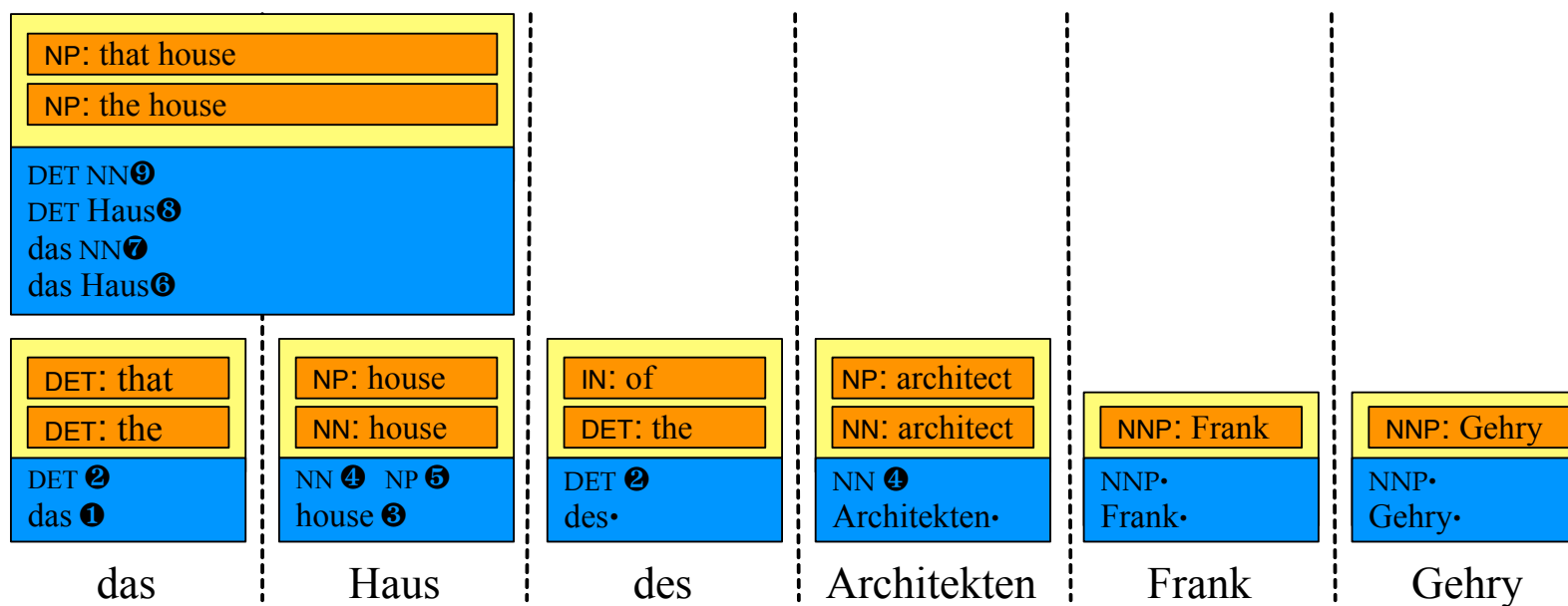
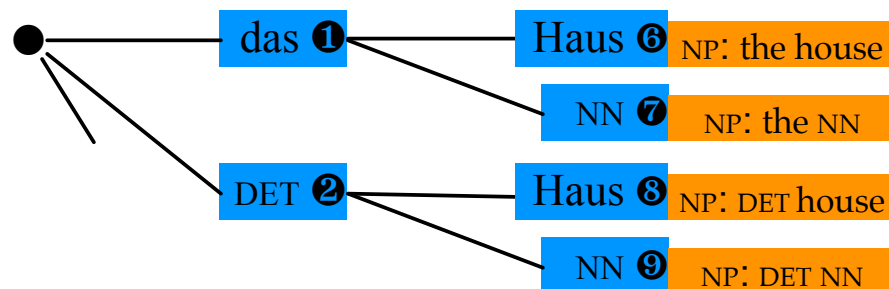
記錄點規則



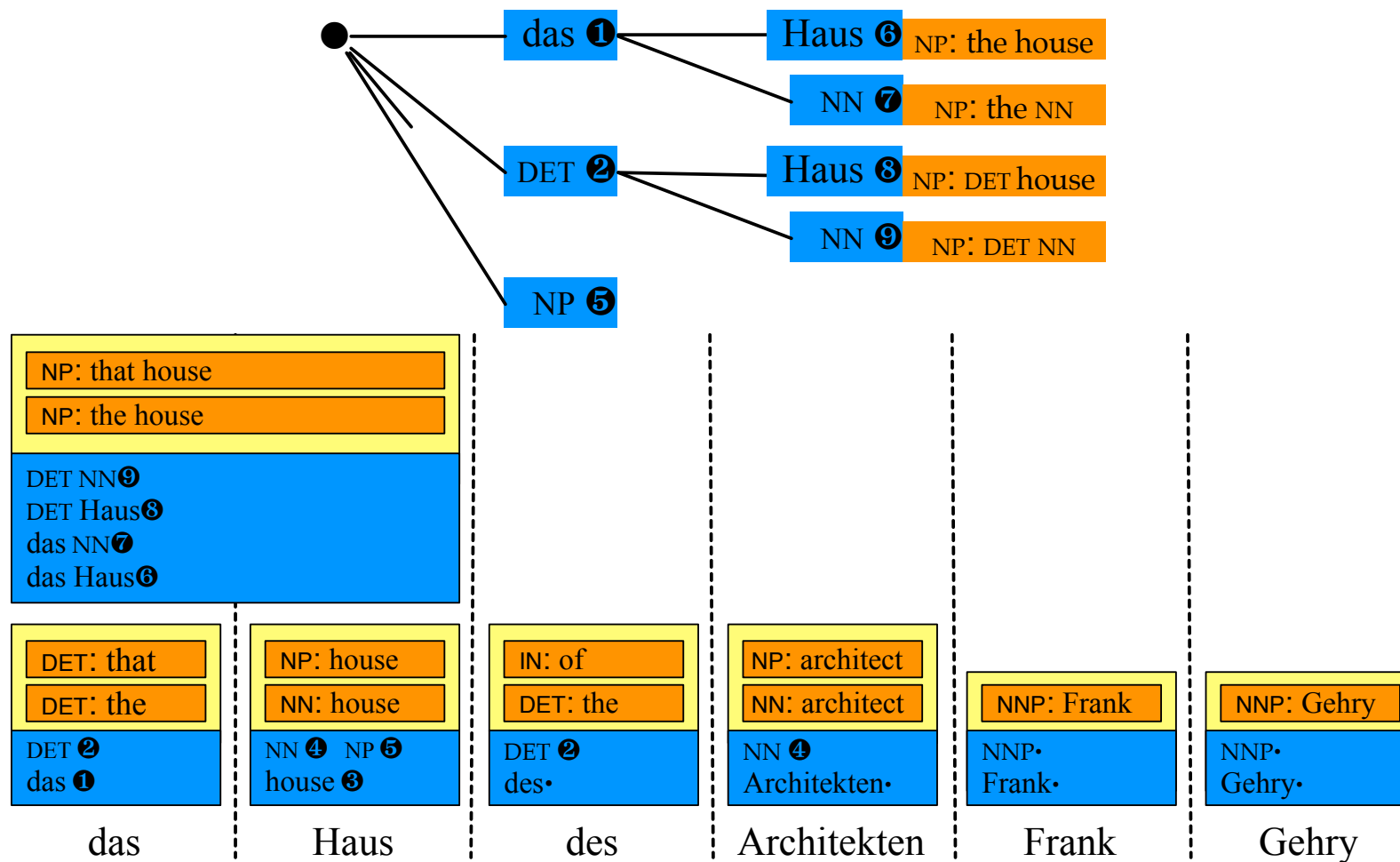
檢查點規則是否有翻譯



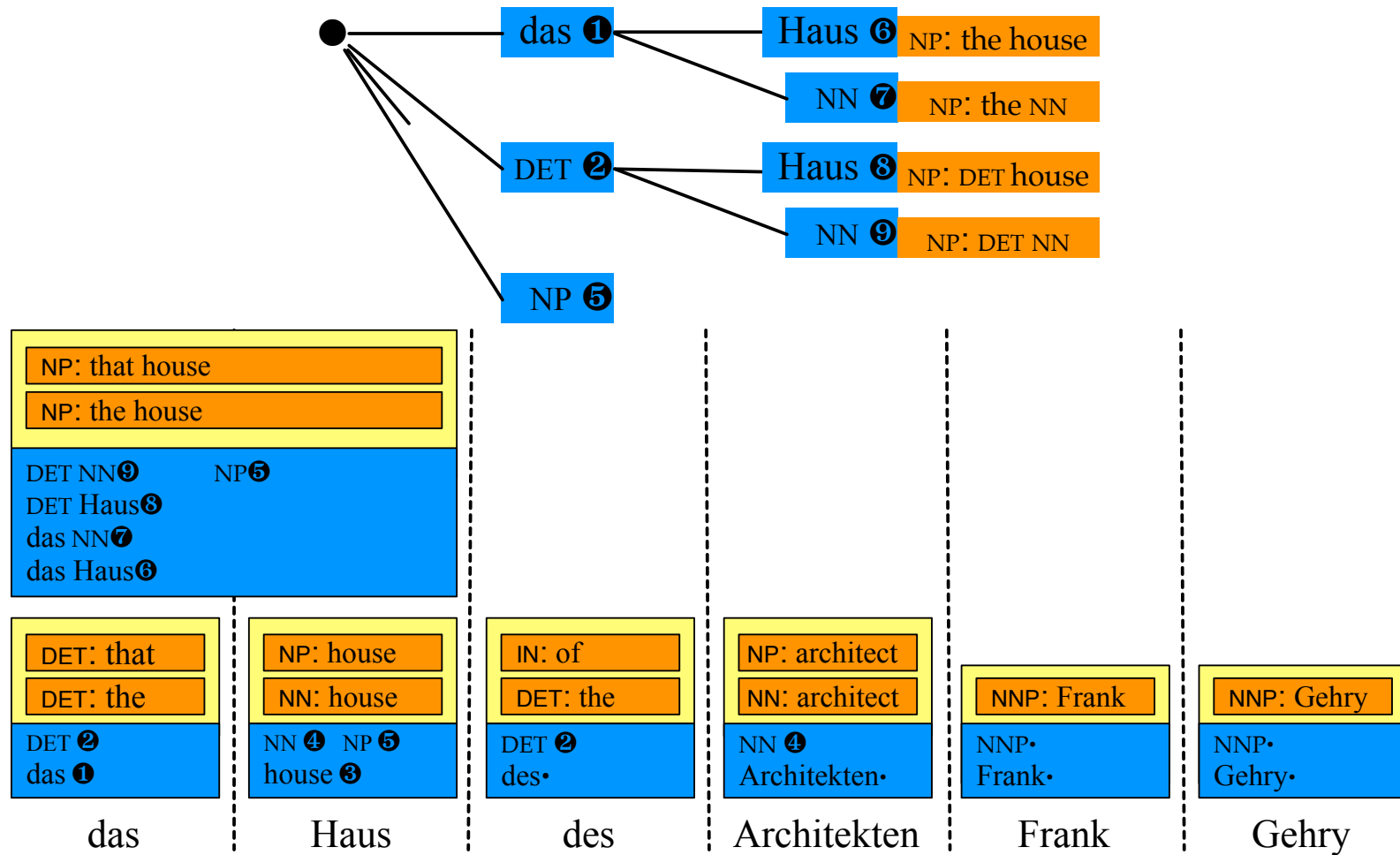
實施翻譯



在前綴樹查詢成分的標籤 (NN)



加到格子的點規則清單



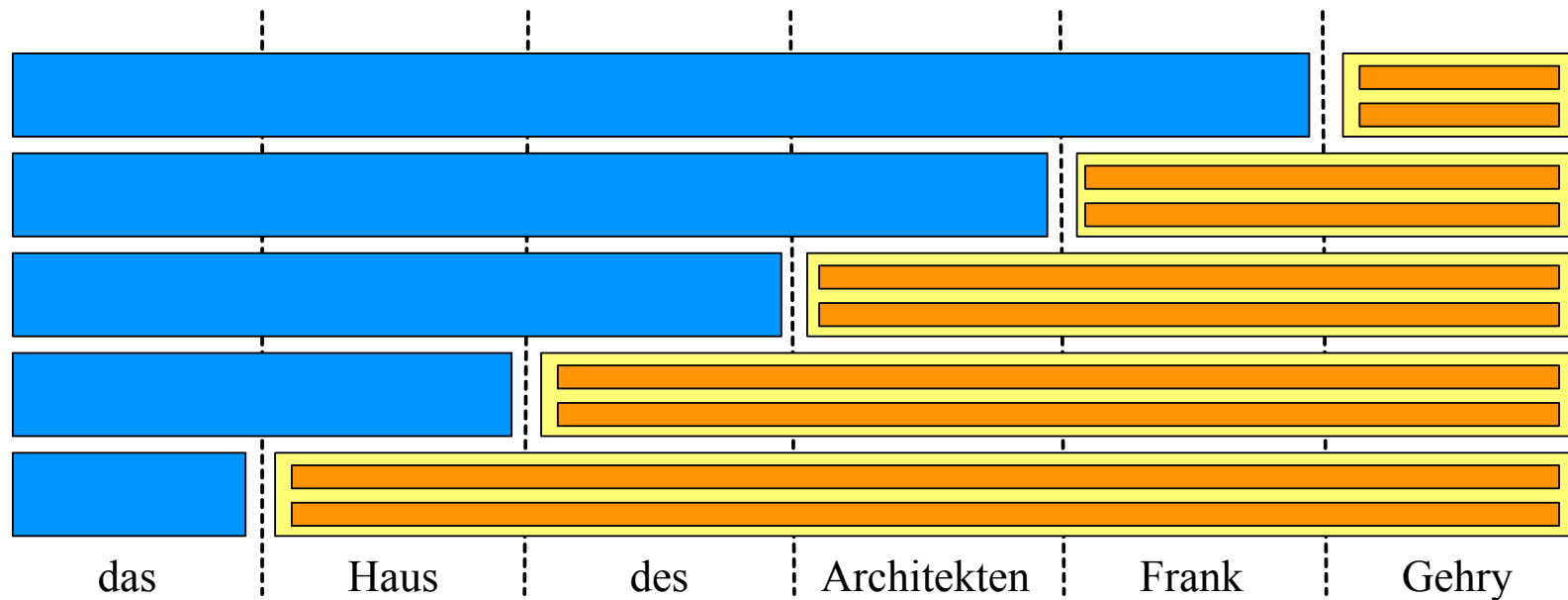
比以前更大的跨距

Extend lists of dotted rules with cell constituent labels

span's dotted rule list (with same start)

plus neighboring

span's constituent labels of hypotheses (with same end)



討論

- 複雜度為 $O(rn^3)$ 當句子長度為 n 而點規則數量為 r
 - 可以對跨距長度加以限制（不含開始的跨距）
 - 可以限制點規則的數量（很任意的）
- 點規則數量是否會暴增？
- 會，如果太多規則有相連的非終止符號
 - 這樣的規則到處都可以適用
 - 如果有詞彙在規則內，就會比較有限制性，不會到處適用

比較麻煩的規則

- 有些規則到處都可以適用（相連非終止符號），例如

$VP \rightarrow \text{gibt } X_1 \ X_2 \mid \text{gives } NP_2 \text{ to } NP_1$

- 相連非終止符號可以對到很多跨距的配對
- 特別是沒有區分片語的階層式模型（只有一個符號，適用所有狀況
- 對於句法式模型或許還好

- 三個相連非終止符號，那就更麻煩

$VP \rightarrow \text{trifft } X_1 \ X_2 \ X_3 \text{ heute } \mid \text{meets } NP_1 \text{ today } PP_2 \ PP_3$

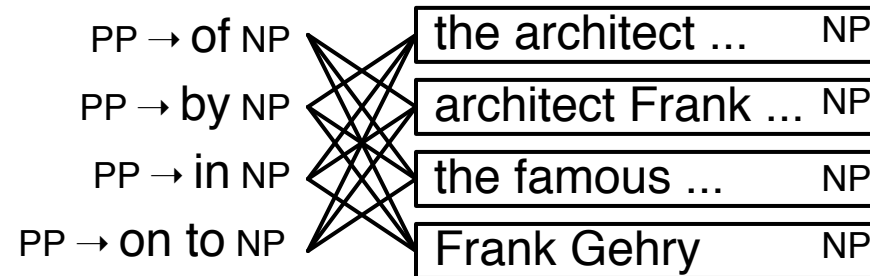
- 甚至句法式模型也無法應付
- 這時候就看到 Collins COBUILD grammar patterns 的優點
 - * 高度詞彙化，非終止符號之間都有詞彙（介詞，而非介詞片語）

小結

- 我們可以檢查規則是否適用
 - 我們知道適用跨距
 - 但是，還有有都的選擇與組合
 - 很多可能的翻譯
 - 非終止符號，可以對到不同跨距、不同標籤的假設
- 組合與選擇的數量，隨著非終止符號的指數函數成長

我們來看看一個非終止符號的例子

如果，有這些適用規則： $PP \rightarrow \text{des } X \mid \dots NP \dots$

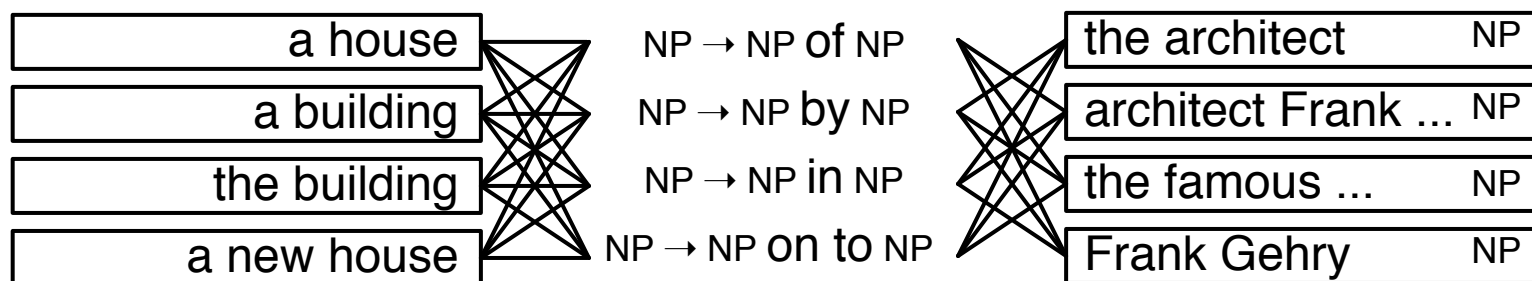


- 如果非終止符號可以對到 h 假設
 - 個別有 of t 詞彙翻譯
- \Rightarrow 複雜度為 $O(ht)$

(附註：我們可能不可以把規則以輸出標籤分群，
所以規則 rule $NP \rightarrow \text{des } X \mid \text{the } NP$ 也必須納入考慮)

兩個非終止符號的規則

如果，有這些適用規則： $NP \rightarrow X_1 \text{ des } X_2 \mid NP_1 \dots NP_2$



- 如果兩個非終止符號都可以對到 h 假設
 - 個別有 of t 詞彙翻譯
- \Rightarrow 複雜度為 $O(h^2t)$ — 適用三度空間的立方修剪 cubic pruning

(note: 規則還可以有不同的詞序重排—增加複雜度)

結語

- 同步上下文無關文法
- 用平行剖析樹語料庫，擷取規則
- 由下而上的解碼 (一面剖析一面解碼)
- 剖析表的組織很重要：動態規劃、堆疊、修剪
- 規則的前綴樹資料結構
- 點規則
- 立方修剪 cubic pruning