Question 1

$$
\begin{aligned}
f_1(y_{-1}, x_1 \ldots x_N, j, y) &= 1 \text{ if } j = 1 \text{ and } y \in \{A, B\} \\
&= 0 \text{ otherwise}
\end{aligned}
$$

$$
\begin{aligned}
f_2(y_{-1}, x_1 \ldots x_N, j, y) &= 1 \text{ if } j \geq 2 \text{ and} \\
&\quad \text{NOT}(y_{-1}, y \in \{A, B\} \text{ or } y_{-1}, y \in \{C, D\}) \\
&= 0 \text{ otherwise}
\end{aligned}
$$

$$
\begin{aligned}
f_3(y_{-1}, x_1 \ldots x_N, j, y) &= 1 \text{ if odd}(x_j) \text{ and } y \in \{A, C\} \\
&= 0 \text{ otherwise}
\end{aligned}
$$

$$
\begin{aligned}
f_4(y_{-1}, x_1 \ldots x_N, j, y) &= 1 \text{ if even}(x_j) \text{ and } y \in \{B, D\} \\
&= 0 \text{ otherwise}
\end{aligned}
$$

Here odd(x) returns true if x has an odd number of letters.
even(x) returns true if x has an even number of letters.
NOT$(TRUE) = FALSE$ and NOT$(FALSE) = TRUE$.

Define a feature-vector representation

$$f(t_1 \ldots t_m, j, y)$$

where:

- y is the label (`STARTS(S)`, `START(VP)`, `START(NP)`, etc.)
- $t_1 \ldots t_m$ are the sequence of subtrees in the current parser state
- j is the index of the left-most subtree that does not have `START(...)` or `JOIN(...)` as its root

Question 2 Continued

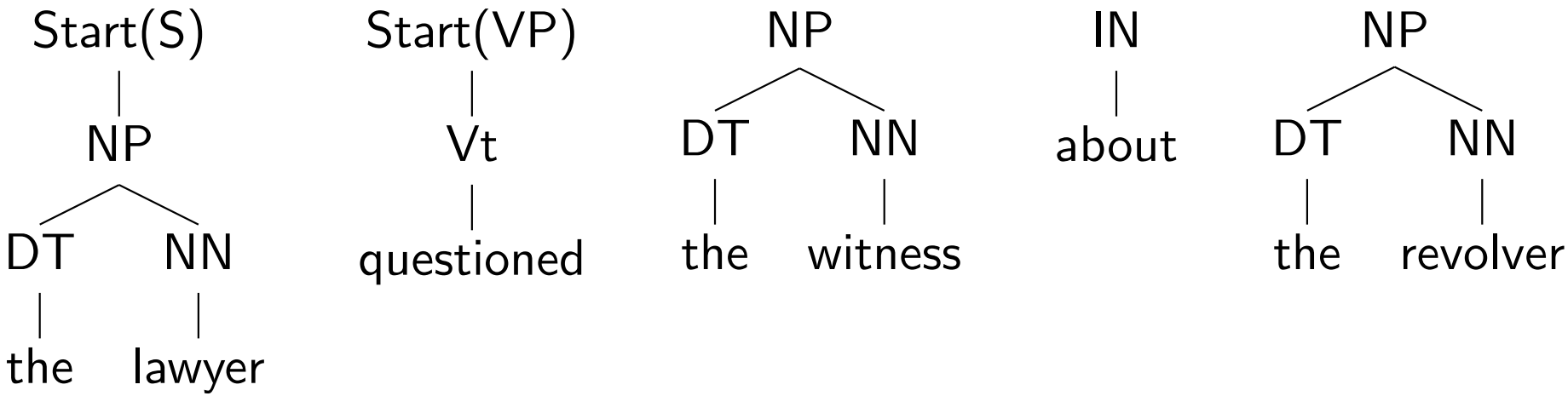
We can then define the following feature:

$$f(t_1 \dots t_m, j, y) = 1$$

if and only if:

- $y = \texttt{JOIN(VP)}$
- root label of $t_{j-1} = \texttt{START(VP)}$, and head-word of t_{j-1} is *questioned*
- root label of $t_j = \texttt{NP}$, and head-word of t_j is *witness*
- root label of $t_{j+1} = \texttt{IN}$, and head-word of t_{j+1} is *about*
- root label of $t_{j+2} = \texttt{NP}$, and head-word of t_{j+2} is *revolver*

Note: the feature is 0 if the above conditions are not met. We can introduce a similar feature for each possible value of the label y (e.g., `START(NP)` etc.)


Start(S)
NP
DT
NN
the
lawyer
Start(VP)
Vt
questioned
NP
DT
NN
the
witness
IN
about
NP
DT
NN
the
revolver