3.2.13 CYK. Označme $\mathcal{G} = (N, \Sigma, S, P)$ a $w = a_1 a_2 \dots a_k$. Postupně vytváříme množiny $X_{i,j}$ pro $1 \le i \le j \le k$, kde

$$X_{i,j} = \{A \in N \mid A \Rightarrow^\star_{\mathcal{G}} a_i a_{i+1} \dots a_j\}, \quad \text{tj. z } A \text{ vygenerujeme slovo } a_i \dots a_j.$$

Platí

$$A \in X_{i,i} \quad \text{právě tehdy, když} \quad A \to a_i \in P.$$

Navíc

$$X_{1,k} = \{A \in N \mid A \Rightarrow^\star_{\mathcal{G}} a_1 a_2 \dots a_k = w\}.$$

Dále si uvědomte, že $A \Rightarrow^\star_{\mathcal{G}} a_i a_{i+1} \dots a_j$ právě tehdy, když existují neterminály B, C takové, že

$$\begin{array}{llll}
A \to BC \in P, \text{ kde} & \text{buď} & B \Rightarrow^\star_{\mathcal{G}} a_i & \text{a} \quad C \Rightarrow^\star_{\mathcal{G}} a_{i+1} \dots a_j \\
 & \text{nebo} & B \Rightarrow^\star_{\mathcal{G}} a_i a_{i+1} & \text{a} \quad C \Rightarrow^\star_{\mathcal{G}} a_{i+2} \dots a_j \\
 & \text{nebo} & B \Rightarrow^\star_{\mathcal{G}} a_i a_{i+1} a_{i+2} & \text{a} \quad C \Rightarrow^\star_{\mathcal{G}} a_{i+3} \dots a_j \\
 & & \dots & \qquad \dots \\
 & \text{nebo} & B \Rightarrow^\star_{\mathcal{G}} a_i \dots a_{j-1} & \text{a} \quad C \Rightarrow^\star_{\mathcal{G}} a_j
\end{array}$$

Předpokládejme, že máme zkonstruovány všechny množiny $X_{p,q}$, kde $q - p < n$. Pak množiny $X_{i,j}$ pro $j - i = n$ utvoříme takto:

$$\begin{array}{llll}
A \in X_{i,j} \text{ iff } \exists A \to BC \in P \text{ tak, že} & \text{buď} & B \in X_{i,i} & \text{a} \quad C \in X_{i+1,j} \\
 & \text{nebo} & B \in X_{i,i+1} & \text{a} \quad C \in X_{i+2,j} \\
 & \text{nebo} & B \in X_{i,i+2} & \text{a} \quad C \in X_{i+3,j} \\
 & & \dots & \qquad \dots \\
 & \text{nebo} & B \in X_{i,j-1} & \text{a} \quad C \in X_{j,j}
\end{array}$$

Začínáme tedy konstrukcí k množin $X_{i,i}$, $i = 1, 2, \dots, k$, následuje pak $k - 1$ množin $X_{i,i+1}$, $i = 1, 2, \dots, k - 1$, $k - 2$ množin $X_{i,i+2}$, $i = 1, 2, \dots, k - 2$, atd. dvě množiny $X_{1,k-1}$, $X_{2,k}$ a nakonec jedna množina $X_{1,k}$ a to podle následujícího postupu:

$$X_{i,j} = \{A \in N \mid \exists A \to BC \in P \text{ tak, že } B \in X_{i,i+m}, C \in X_{i+m+1,j}\}.$$

Platí $w \in L(\mathcal{G})$ právě tehdy, když $S \in X_{1,k}$.

3.2.14 Příklad. Je dána gramatika $\mathcal{G}$ pravidly

$$\begin{array}{l}
S \to AB \mid BC \\
A \to BA \mid a \\
B \to CC \mid b \\
C \to AB \mid a
\end{array}$$

Pomocí algoritmu CYK rozhodněte, zda slovo $aabab$ je generováno bezkontextovou gramatikou $\mathcal{G}$.

Řešení: Konstrukci množin $X_{i,j}$ pro $1 \le i \le j \le 5$ si znázorníme do tabulky. Tabulka bude mít 5 řádků a 5 sloupců, kde vyplněných bude jen ta část, která se nenachází „nad diagonálou". Poslední řádek obsahuje pět množin $X_{1,1}$, $X_{2,2}$, $X_{3,3}$, $X_{4,4}$ a $X_{5,5}$. Předposlední řádek obsahuje čtyři množiny $X_{1,2}$, $X_{2,3}$, $X_{3,4}$ a $X_{4,5}$. Řádek, který je třetí od zdola (a také shora) obsahuje tři množiny $X_{1,3}$, $X_{2,4}$ a $X_{3,5}$. Řádek, který je čtvrtý od odzdola (a druhý shora) obsahuje dvě množiny $X_{1,4}$ a $X_{2,5}$. Nejvyšší řádek obsahuje jednu množinu $X_{1,5}$. V našem případě jsou neterminály množin X_{ij} v následující tabulce:

S, C				
S, A, C	B			
B	B	S, C		
B	S, C	S, A	S, C	
A, C	A, C	B	A, C	B
a	a	b	a	b

Z předchozí tabulky také můžeme odvodit derivace slova *aabab* gramatikou $\mathcal{G}$. Jedna z takových levých derivací je např. tato:

$$S \Rightarrow AB \Rightarrow aB \Rightarrow aCC \Rightarrow aaC \Rightarrow aaAB \Rightarrow aaBAB \Rightarrow aabAB \Rightarrow aabaB \Rightarrow aabab.$$

3.2.15 Poznámka. Algoritmus CYK pracuje v čase úměrném k^3, kde k je délka terminálního slova w. Velikost gramatiky v Chomského normálním tvaru považujeme za konstantu.

Jedná se o algoritmus ve stylu dynamického programování. V případě speciálních bezkontextových gramatik je možné sestrojit algoritmus pracující v lineárním čase.

3.2.16 Zásobníkový automat. Zhruba řečeno, zásobníkový automat se skládá z řídící jednotky, která je v jednom z konečně mnoha možných stavů, ze vstupní pásky se čtecí hlavou a ze zásobníku. Na základě toho, v jakém stavu se automat nachází, co hlava čte na vstupní pásce a jaký symbol je na vrcholu zásobníku, automat udělá akci: přejde do nového stavu, posune čtecí hlavu o jedno políčko doprava nebo stojí (to v případě, že automat reagoval na prázdné slovo) a zásobníkový symbol, který je na vrcholu, nahradí zásobníkovým slovem daným přechodovou funkcí. Formálně:

Definice. *Zásobníkový automat* je sedmice $(Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$, kde

- Q je konečná množina stavů,
- Σ je konečná množina vstupních symbolů,
- Γ je konečná množina zásobníkových symbolů,
- δ přiřazuje každé trojici (q, a, X), $q \in Q$, $a \in \Sigma \cup \{\varepsilon\}$, $X \in \Gamma$, konečnou množinu dvojic (p, α), kde $p \in Q$ a $\alpha \in \Gamma^\star$. Formálně:

  $$\delta: Q \times (\Sigma \cup \{\varepsilon\}) \times \Gamma \to \mathcal{P}_f(Q \times \Gamma^\star).$$

  ($\mathcal{P}_f(A)$ značí množinu všech konečných podmnožin množiny A.)
- $q_0 \in Q$ je počáteční stav,
- $Z_0 \in \Gamma$ je počáteční zásobníkový symbol a
- $F \subseteq Q$ je množina koncových stavů.

$\square$

Uvědomte si, že zásobníkový automat tak, jak byl definován, je nedeterministický.

Situace zásobníkového automatu. Je dán zásobníkový automat $(Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$. Jestliže je zásobníkový automat ve stavu q, na pásce je ještě nepřečtené slovo u a obsah zásobníku je slovo γ (první symbol γ je vrchol zásobníku), pak popisujeme tuto *situaci* trojicí (q, u, γ), kde první symbol slova γ je vrchol zásobníku. $\square$

Jeden krok práce zásobníkového automatu – relace $\vdash_A$. Je dán zásobníkový automat $A = (Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$, který je v situaci $(q, au, X\gamma)$, kde $a \in \Sigma \cup \{\varepsilon\}$, $X \in \Gamma$. Jestliže existuje $(p, \alpha) \in \delta(q, a, X)$, pak A přejde do situace $(p, u, \alpha\gamma)$. Tento fakt značíme $(q, au, X\gamma) \vdash_A (p, u, \alpha\gamma)$. Píšeme též $(q, au, X\gamma) \vdash_A (p, u, \alpha\gamma)$ právě tehdy, když $(p, \alpha) \in \delta(q, a, X)$. $\square$

Relace $\vdash_A^\star$. Jeden krok zásobníkového automatu rozšíříme na konečný počet kroků. Automat A *přejde ze situace S do situace S'*, píšeme $S \vdash_A^\star S'$, právě tehdy, když buď $S = S'$ nebo existuje konečný počet situací $S_1, S_2, \ldots, S_n$ takových, že

$$S \vdash_A S_1,\ S_1 \vdash_A S_2\ \ldots,\ S_n \vdash_A S'.$$

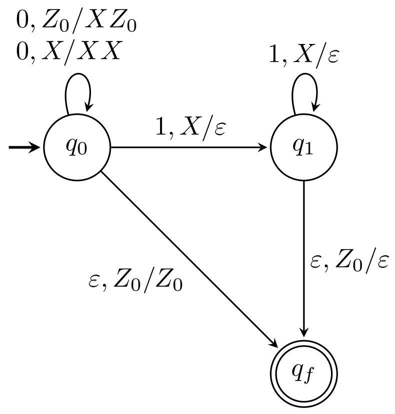
O $S \vdash_A^\star S'$ také mluvíme jako o *výpočtu* nad situací S. $\square$

3.2.17 Stavový diagram zásobníkového automatu. Obdobně jako u konečných automatů můžeme přechodovou funkci zásobníkového automatu zadat tabulkou nebo stavovým diagramem. V tabulce máme tolik sloupců, kolik je dvojic stav, zásobníkový symbol. U stavového diagramu máme hranu ze stavu q do stavu p ohodnocenou $a, X/\gamma$ právě tehdy, když $(p, \gamma) \in \delta(q, a, X)$, kde $a \in \Sigma \cup \{\varepsilon\}$, $X \in \Gamma$. Ukážeme si oba způsoby na následujícím příkladě.

Příklad. Je dán zásobníkový automat $A = (Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$, kde $Q = \{q_0, q_1, q_f\}$, $\Sigma = \{0, 1\}$, $\Gamma = \{Z_0, X\}$, jehož přechodová funkce je daná tabulkou

	$(0, Z_0)$	$(0, X)$	$(1, Z_0)$	$(1, X)$	(ε, Z_0)	(ε, X)
$\rightarrow q_0$	(q_0, XZ_0)	(q_0, XX)	$-$	(q_1, ε)	(q_f, Z_0)	$-$
q_1	$-$	$-$	$-$	(q_1, ε)	(q_f, ε)	$-$
$\leftarrow q_f$	$-$	$-$	$-$	$-$	$-$	$-$

a stavový diagram je:



3.2.18 Jazyky přijímané zásobníkovým automatem.
Pro daný zásobníkový automat rozlišujeme dva jazyky jím přijímané. Kromě jazyka přijímaného koncovým stavem (jedná se o obdobu jazyka přijímaného konečným automatem) je to ještě jazyk přijímaný prázdným zásobníkem. Zde jsou definice těchto pojmů.

Definice. *Jazyk přijímaný koncovým stavem* $L(A)$, kde $A = (Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ je zásobníkový automat, je definován takto:

$$L(A) = \{u \in \Sigma^\star \mid (q_0, u, Z_0) \vdash_A^\star (p, \varepsilon, \gamma), p \in F\}.$$

□

Zhruba řečeno, zásobníkový automat začne v počáteční situaci, tj. v počátečním stavu q_0, na vstupní pásce má slovo u a na zásobníku pouze počáteční zásobníkový symbol Z_0. Slovo je přijato právě, když existuje výpočet, který přečte u a automat je v některém koncovém stavu. To, zda je současně vyprázdněn zásobník nebo není, nehraje roli.

Definice. *Jazyk přijímaný prázdným zásobníkem* $N(A)$, kde $A = (Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ je zásobníkový automat, je definován takto:

$$N(A) = \{u \in \Sigma^\star \mid (q_0, u, Z_0) \vdash_A^\star (p, \varepsilon, \varepsilon), p \in Q\}.$$

□

Zhruba řečeno, zásobníkový automat začne v počátečním stavu q_0, na vstupní pásce má slovo u a na zásobníku pouze počáteční zásobníkový symbol Z_0 (počáteční situace). Slovo je přijato právě tehdy, když najdeme výpočet, který po přečtení slova u má prázdný zásobník.

Protože při přijímání prázdným zásobníkem nezáleží na množině koncových stavů (nepotřebujeme je), budeme množinu koncových stavů vynechávat a zásobníkový automat v takovém případě bude pouze šestice $(Q, \Sigma, \Gamma, \delta, q_0, Z_0)$.

3.2.19 Příklad z 3.4.2 – pokračování. Ukažme si práci zásobníkového automatu nad slovem $0^3 1^3$. Zásobníkový automat začíná v situaci $(q_), 0^3 1^3, Z_0)$ a pokračuje

$$(q_0, 0^3 1^3, Z_0) \vdash (q_0, 0^2 1^3, XZ_0) \vdash^2 (q_0, 1^3, X^3 Z_0) \vdash (q_1, 1^2, X^2 Z_0) \vdash^2 (q_1, \varepsilon, Z_0) \vdash (q_f, \varepsilon, \varepsilon).$$

To znamená, že slovo $0^3 1^3$ je přijato jak koncovým stavem, tak prázdným zásobníkem.

Není těžké ukázat, že jazyk přijímaný koncovým stavem je

$$L(A) = \{0^n 1^n \mid n \geq 0\},$$

kdežto jazyk přijímaný prázdným zásobníkem je

$$N(A) = \{0^n 1^n \mid n \geq 1\}.$$

Ano, prázdné slovo ε není přijato prázdným zásobmíkem.

3.2.20 Tvrzení. Je-li L jazyk přijímaný zásobníkovým automatem A prázdným zásobníkem, pak existuje zásobníkový automat B, který přijímá jazyk L koncovým stavem; tj.

$$N(A) = L(B).$$

□

Nástin důkazu: Máme zásobníkový automat $A = (Q_A, \Sigma, \Gamma_A, \delta_A, q_A, Z_A)$. K němu sestrojíme zásobníkový automat $B = (Q_B, \Sigma, \Gamma_B, \delta_B, q_B, Z_B, F_B)$ takto:

- K množině stavů Q_A přidáme dva nové stavy — nový počáteční stav q_B a koncový stav q_f; tj. $Q_B = Q_A \cup \{q_B, q_f\}$, $q_B, q_f \notin Q_A$.
- K množině zásobníkových symbolů Γ_A přidáme nový počáteční zásobníkový symbol; tj. $\Gamma_B = \Gamma_A \cup \{Z_B\}$, $Z_B \notin \Gamma_A$.
- δ_B dostaneme tak, že k δ_A přidáme následující přechody:
  - $\delta_B(q_B, \varepsilon, Z_B) = \{(q_A, Z_A Z_B)\}$, (to znamená, že na začátku práce aniž bychom četli vstupní symbol přejdeme do původního počátečního stavu a zásobník změníme tak, že na vrcholu je „starý“ počáteční zásobníkový symbol a pod ním „nový“ počáteční zásobníkový symbol).
  - Pro každé $p \in Q_A$ položíme $\delta_B(p, \varepsilon, Z_B) = \{(q_f, Z_B)\}$, (to znamená, že vyprázdní-li „starý“ zásobníkový automat svůj zásobník, přejde „nový“ zásobníkový automat do koncového stavu).

Dá se dokázat, že zásobníkový automat A přijme slovo u prázdným zásobníkem právě tehdy, když ho přijme zásobníkový automat B koncovým stavem.

3.2.21 Tvrzení. Je-li L jazyk přijímaný zásobníkovým automatem A koncovým stavem, pak existuje zásobníkový automat B, který přijímá jazyk L prázdným zásobníkem; tj.

$$L(A) = N(B).$$

□

Nástin důkazu: Označme zásobníkový automat $A = (Q_A, \Sigma, \Gamma_A, \delta_A, q_A, Z_A, F_A)$. K němu sestrojíme zásobníkový automat $B = (Q_B, \Sigma, \Gamma_B, \delta_B, q_B, Z_B)$ takto:

- K množině stavů Q_A přidáme dva nové stavy — nový počáteční stav q_B a stav q_M (který bude sloužit k vyprázdnění zásobníku automatu B); tj. $Q_B = Q_A \cup \{q_B, q_M\}$, $q_B, q_M \notin Q_A$.
- K zásobníkovým symbolům přidáme nový počáteční zásobníkový symbol; tj. $\Gamma_B = \Gamma_A \cup \{Z_B\}$, $Z_B \notin \Gamma_A$;
- δ_B dostaneme tak, že k δ_A přidáme následující přechody:
  - $\delta_B(q_B, \varepsilon, Z_B) = \{(q_A, Z_A Z_B)\}$, (význam je stejný jako v konstrukci důkazu 3.2.23).
  - pro každé $q \in F_A$, $Y \in \Gamma_B$ přidáme $\delta_B(q, \varepsilon, Y) = \{(q_M, \varepsilon)\}$ a $\delta_B(q_M, \varepsilon, Y) = \{(q_M, \varepsilon)\}$, (to znamená, že přečteme-li celé vstupní slovo a jsme v koncovém stavu zásobníkového automatu A, zásobníkový automat B vymaže obsah svého zásobníku).

Dá se dokázat, že zásobníkový automat A přijme slovo u koncovým stavem právě tehdy, když ho přijme zásobníkový automat B prázdným zásobníkem.