Kyushu University 100th 2011
知の新世紀を拓く

形式言語理論

Formal Language Theory

11.　プッシュダウンオートマトン

九州大学
KYUSHU UNIVERSITY

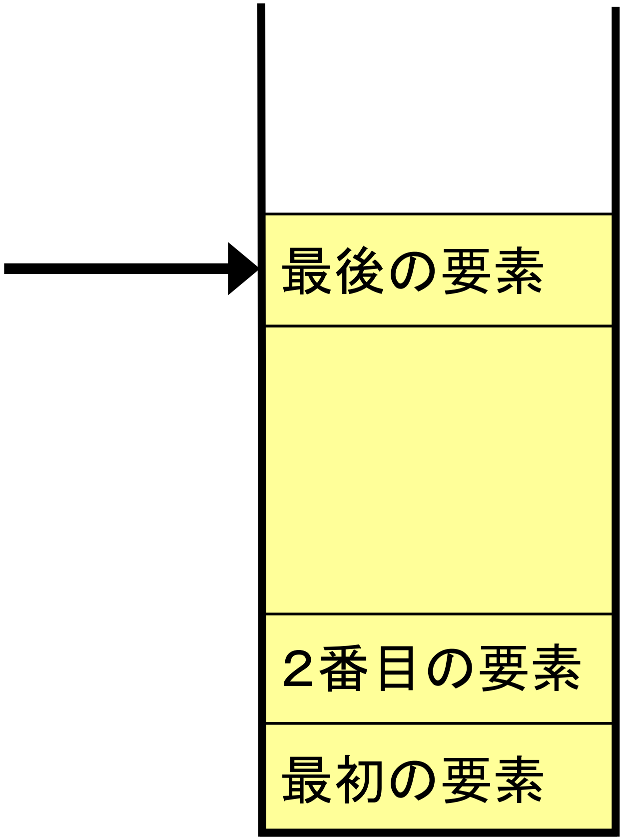
本日の内容

- 準備
  - スタック
- プッシュダウンオートマトン(PDA)
- PDAの動作例
- 演習
- PDAの受理に関するその他の定義
- PDAの標準形
- PDAと文脈自由言語

準備

スタック

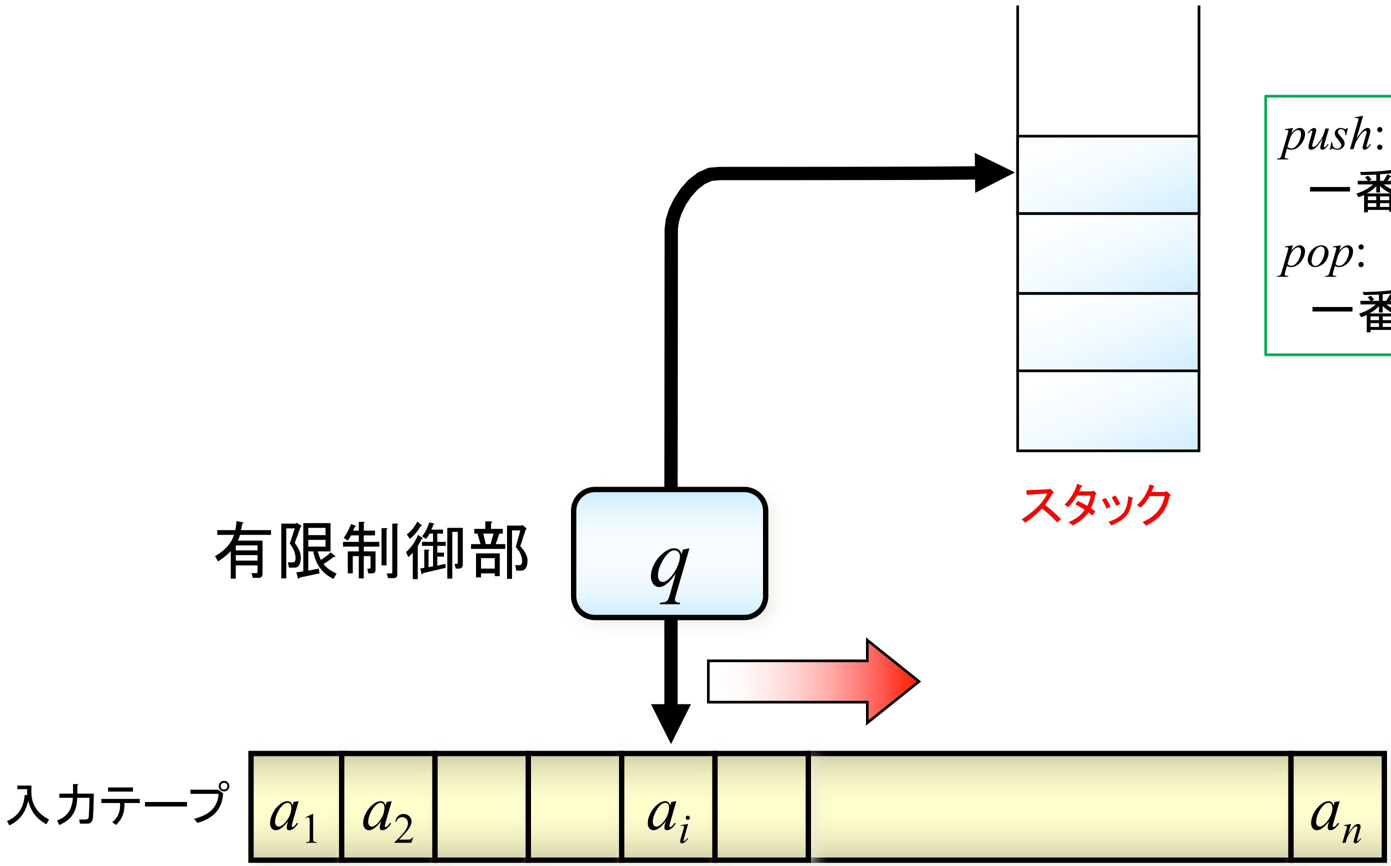
- 特別な制限のあるリスト.
  - リストの末尾に対してのみ要素の挿入・削除が可能.
  - プッシュダウンリスト, LIFO (Last In First Out) などともよばれる.
- モデル.
  - 積み重ねた本
  - ポーカーチップの山
  - …



プッシュダウンオートマトン

プッシュダウンオートマトン(PDA)

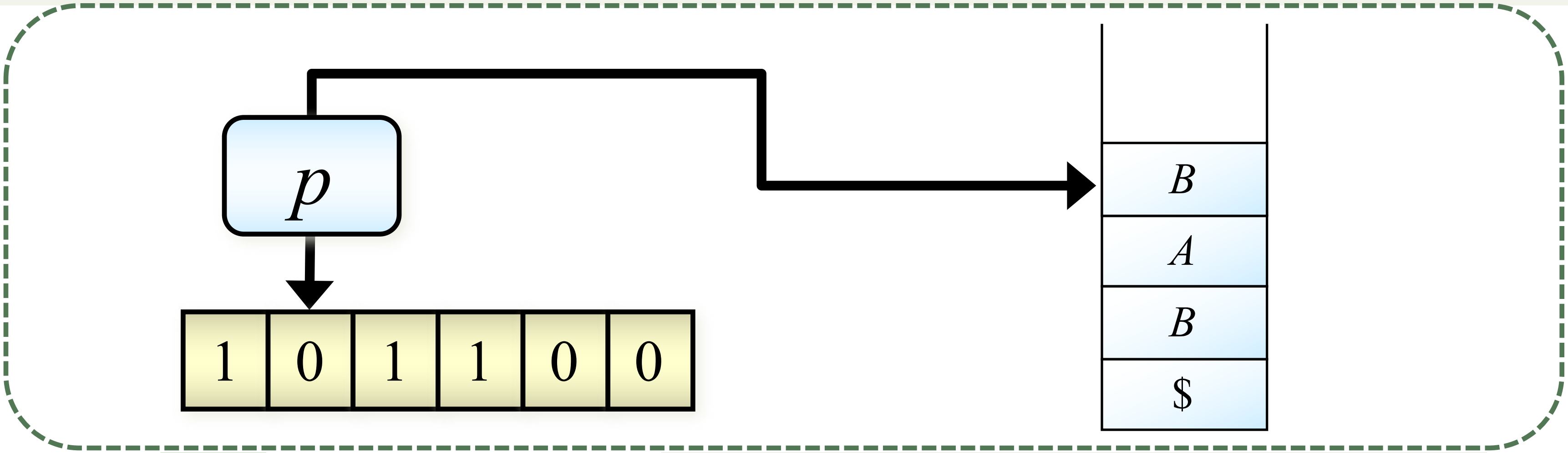
- PDA = ε-動作付 NFA＋スタック（プッシュダウンスタック）



push:
　一番上に記号を積む.

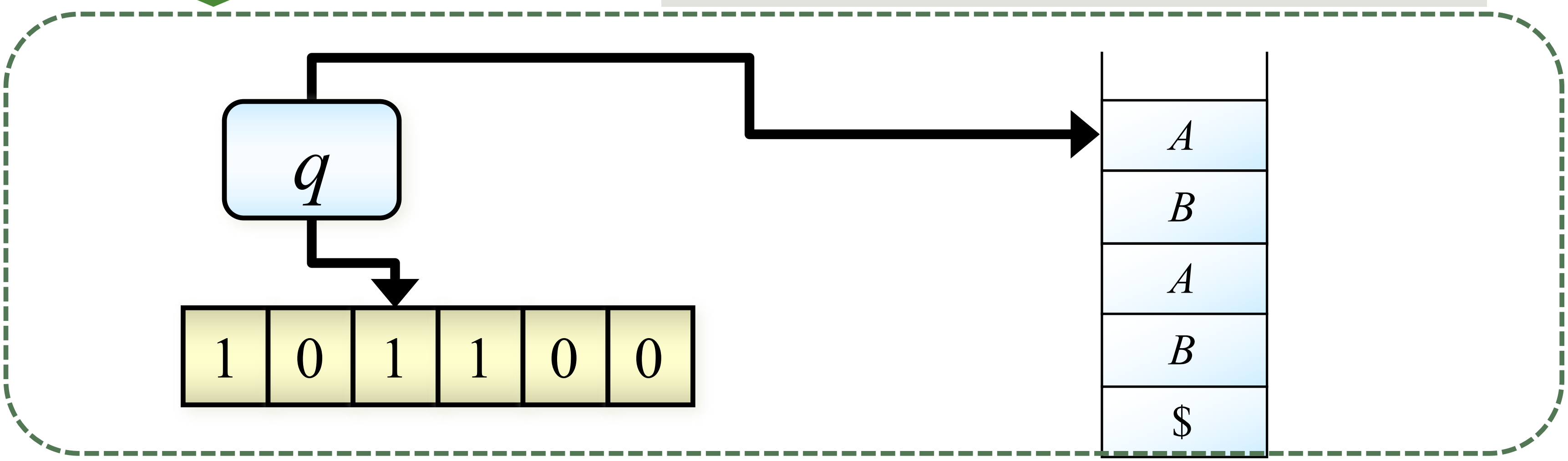
pop:
　一番上の記号を削除する.

PDAの動作の1ステップ



$\delta(p, 0, B) \ni (q, AB)$

- 状態をqに変更.
- 入力ヘッドを1つ右へずらす.
- スタックのトップの記号BをABで置換する.



PDAの動作の1ステップ

- PDAの動作は，状態 q とスタック記号の列 α から成る組 (q, α) で表される．これは以下の意味である．
  - 状態を q へ変え，
  - スタックのトップ記号を記号列 α で置換する．
- (q, α) は，次の3つ組 (p, a, Z) に依存して非決定的に決まる．
  - 状態 p.
  - 入力記号 a (ε-動作のときは a=ε).
  - スタックの トップ記号 Z.
- すなわち，$(q, \alpha) \in \delta(p, a, Z)$ である．

プッシュダウンオートマトン(pushdown automaton; PDA)

【定義】プッシュダウンオートマトンとは，次のような7つ組 $M=(Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ をいう.

- Q は状態の空でない有限集合.
- Σ は記号の空でない有限集合(入力アルファベット).
- Γ は記号の空でない有限集合(スタックアルファベット).
- $\delta: Q\times(\Sigma\cup\{\varepsilon\})\times\Gamma\rightarrow 2^{Q\times\Gamma^*}$ は状態遷移関数.
- $q_0\in Q$ は開始状態.
- $Z_0\in\Gamma$ は初期スタック記号.
- $F\subseteq Q$ は最終状態の集合.

例 言語 $L=\{\ a^n b^n \mid n>0\ \}$を受理する PDA

状態集合	$\{q_1, q_2, q_3\}$
入力アルファベット	$\{a, b\}$
スタックアルファベット	$\{\$, A\}$
状態遷移関数	（下表のとおり）
初期状態	q_1
初期スタック記号	$\$$
最終状態集合	$\{q_3\}$

q	a	Z	$\delta(q, a, Z)$
q_1	a	$\$$	$\{(q_1, A\$)\}$
q_1	a	A	$\{(q_1, AA)\}$
q_1	b	$\$$	$\varnothing$
q_1	b	A	$\{(q_2, \varepsilon)\}$
q_1	ε	$\$$	$\varnothing$
q_1	ε	A	$\varnothing$

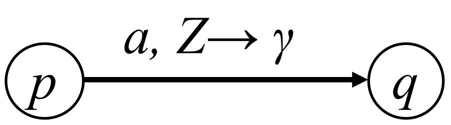
q	a	Z	$\delta(q, a, Z)$
q_2	a	$\$$	$\varnothing$
q_2	a	A	$\varnothing$
q_2	b	$\$$	$\varnothing$
q_2	b	A	$\{(q_2, \varepsilon)\}$
q_2	ε	$\$$	$\{(q_3, \varepsilon)\}$
q_2	ε	A	$\varnothing$

q	a	Z	$\delta(q, a, Z)$
q_3	a	$\$$	$\varnothing$
q_3	a	A	$\varnothing$
q_3	b	$\$$	$\varnothing$
q_3	b	A	$\varnothing$
q_3	ε	$\$$	$\varnothing$
q_3	ε	A	$\varnothing$

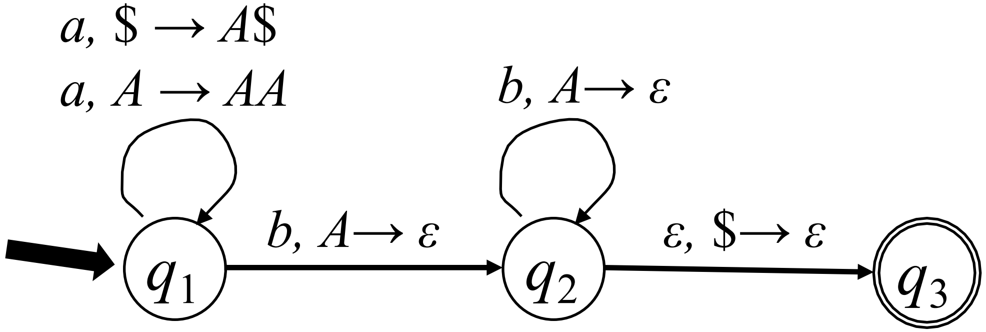
例 言語 $L=\{\, a^n b^n \mid n>0 \,\}$を受理する PDA

状態集合	$\{q_1, q_2, q_3\}$
入力アルファベット	$\{a, b\}$
スタックアルファベット	$\{\$, A\}$
状態遷移関数	（下図のとおり）
初期状態	q_1
初期スタック記号	$\$$
最終状態集合	$\{q_3\}$

遷移 $(q, \gamma) \in \delta(p, a, Z)$ を次のように表す

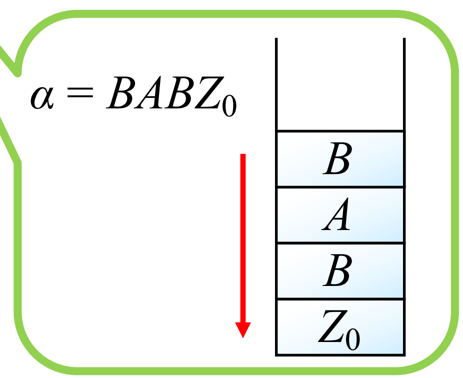


状態遷移図



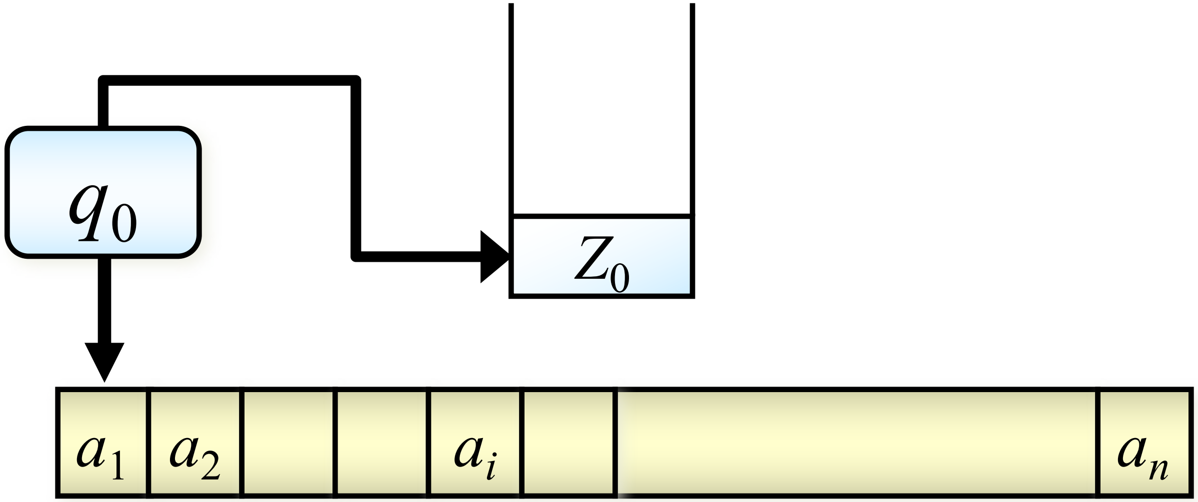
計算状況

- PDA $M=(Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ に対し，次の3つ組 (q, w, α) を**計算状況**という.
  - 状態 $q \in Q$.
  - 入力文字列の未読部分 $w \in \Sigma^*$.
  - スタックの内容 $\alpha \in \Gamma^*$.



初期計算状況

【定義】 PDA M =($Q, \Sigma, \Gamma, \delta, q_0, Z_0, F$) への入力を w とする．このとき，計算状況 (q_0, w, Z_0) を**初期計算状況**という．



$w = a_1 \ldots a_n$

PDAの動作

- PDA $M=(Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ に対し，
  集合 $Q \times \Sigma^* \times \Gamma^*$ 上の関係 $\vdash_M$ を次で定義.

$$(p, ax, Z\alpha) \vdash_M (q, x, \beta\alpha)$$

$$\Leftrightarrow$$

$$(q, \beta) \in \delta(p, a, Z)$$

$(a \in \Sigma \cup \{\varepsilon\}, x \in \Sigma^*, Z \in \Gamma, \alpha, \beta \in \Gamma^*)$

PDAは，初期計算状況からスタートし，δ に沿って計算状況を繰り返し変化させてゆくことによって，計算を行う.

最終状態による受理

【定義】 PDA $M=(Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ が文字列 $w \in \Sigma^*$を（最終状態によって）受理するとは，

$\exists q \in F,\ \exists \gamma \in \Gamma^*$ に対し

$$(q_0, w, Z_0) \vdash_M^* (q, \varepsilon, \gamma)$$

となるときをいう． (q, ε, γ) を（最終状態による）受理計算状況という．

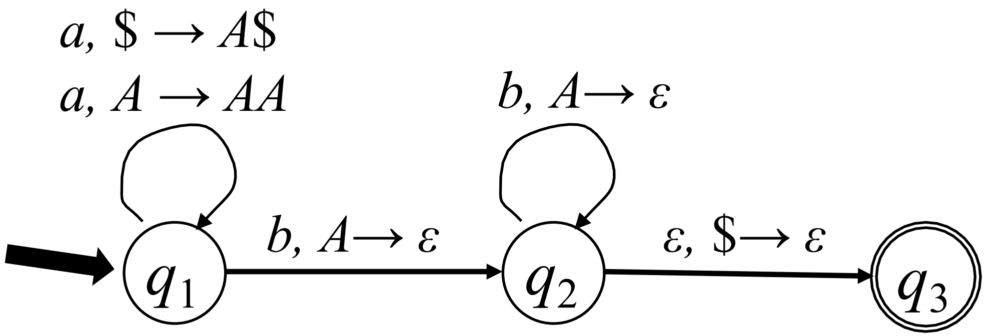
【定義】 PDA $M=(Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ が（最終状態によって）受理する文字列の集合を

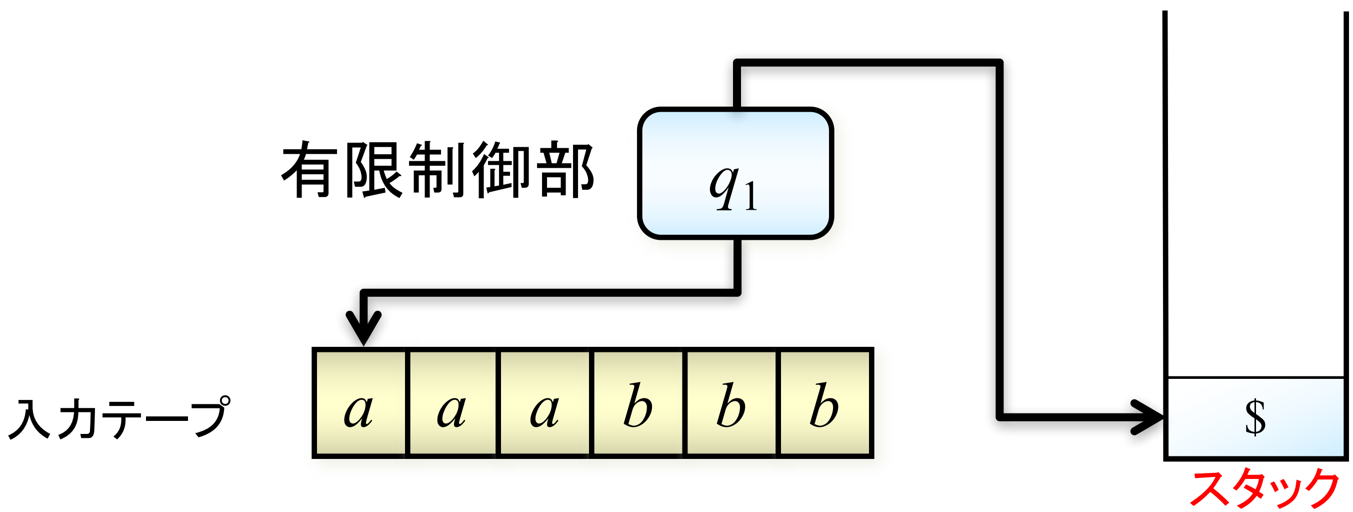
$L(M)=\{w \in \Sigma^* \mid \exists q \in F,\ \exists \gamma \in \Gamma^*, (q_0, w, Z_0) \vdash_M^* (q, \varepsilon, \gamma)\}$

とする．

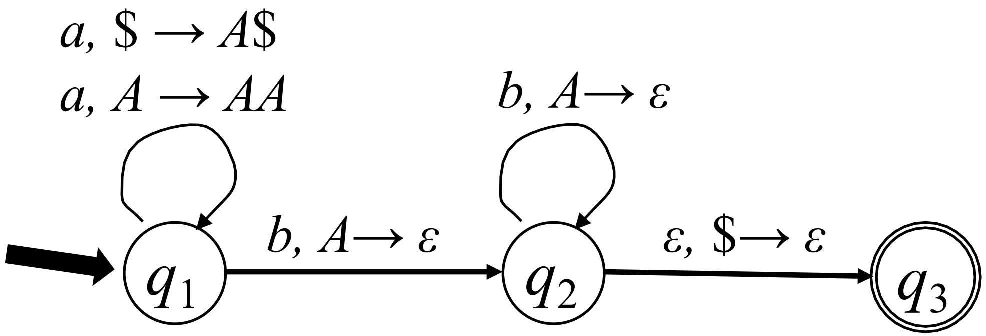
PDAの動作例

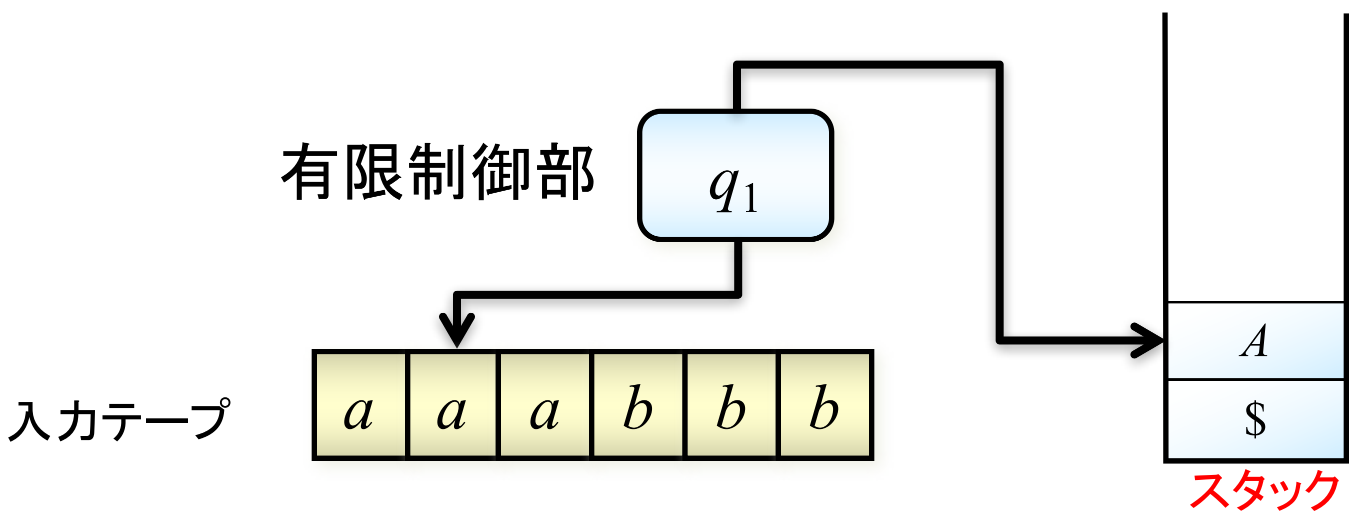
入力 w=$aaabbb$ に対する動作例



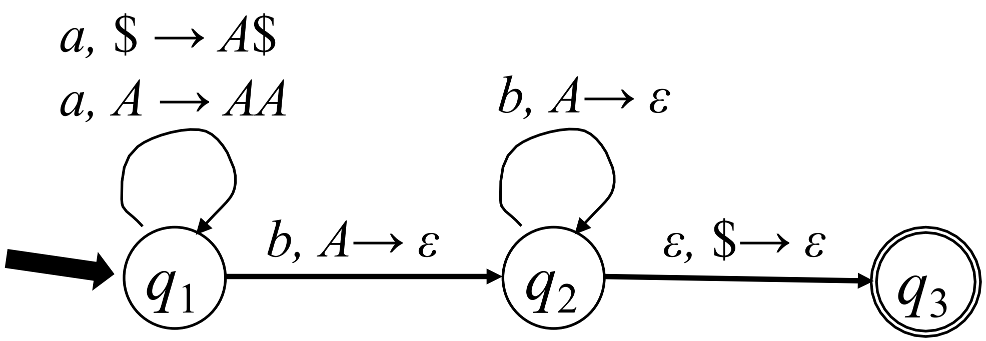


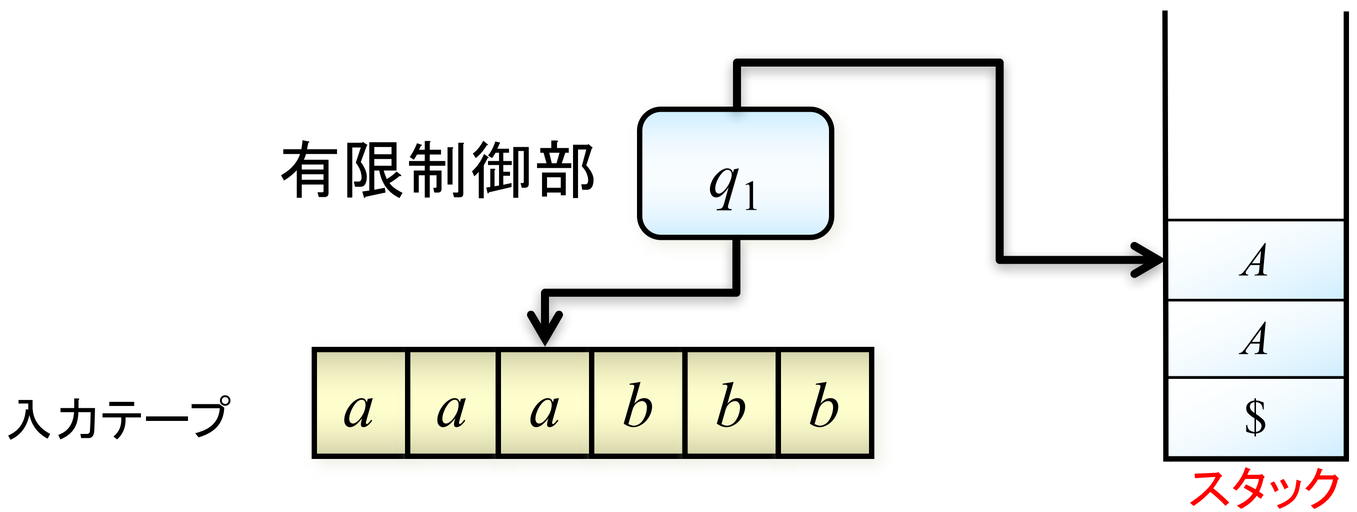
入力 $w=aaabbb$ に対する動作例



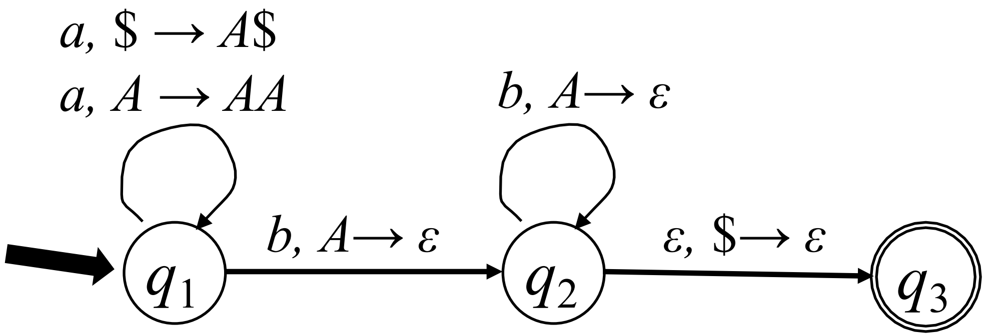


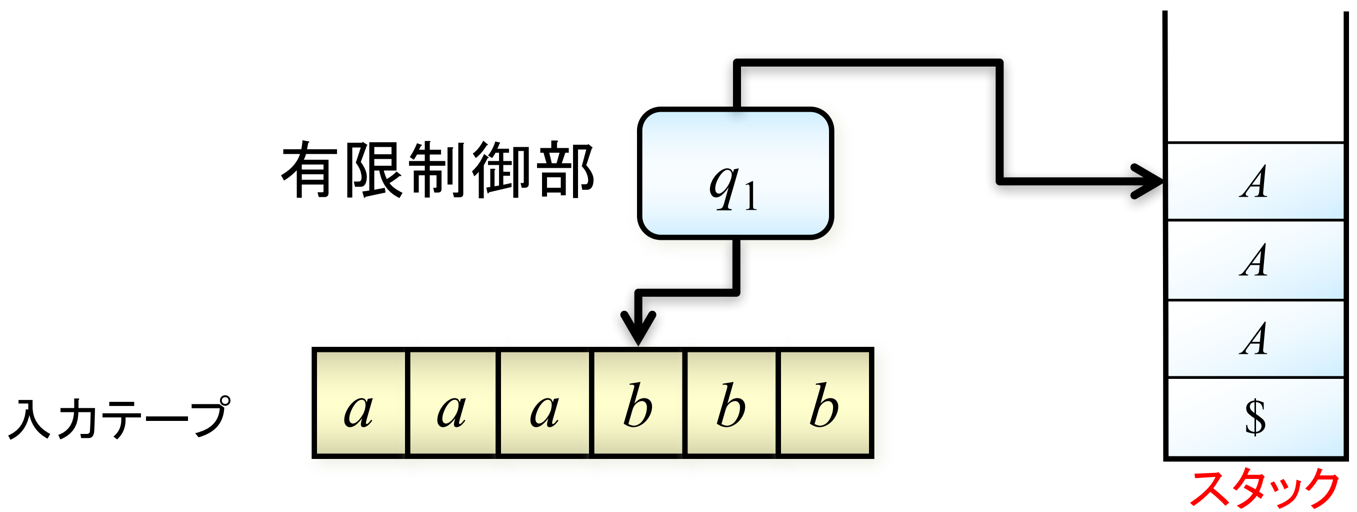
入力 w=*aaabbb* に対する動作例



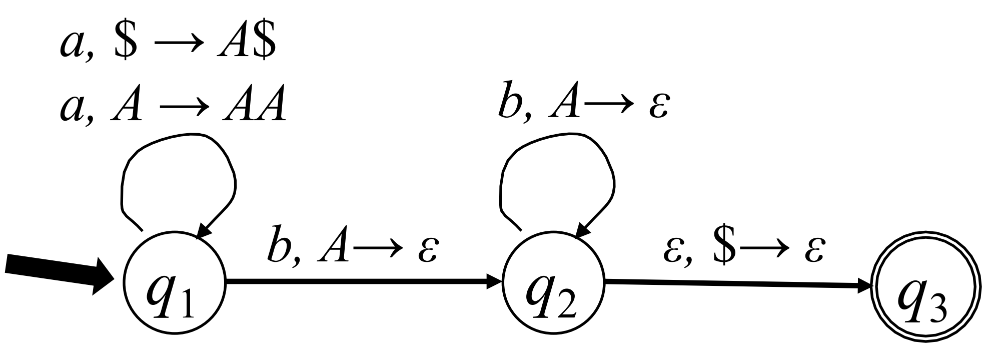


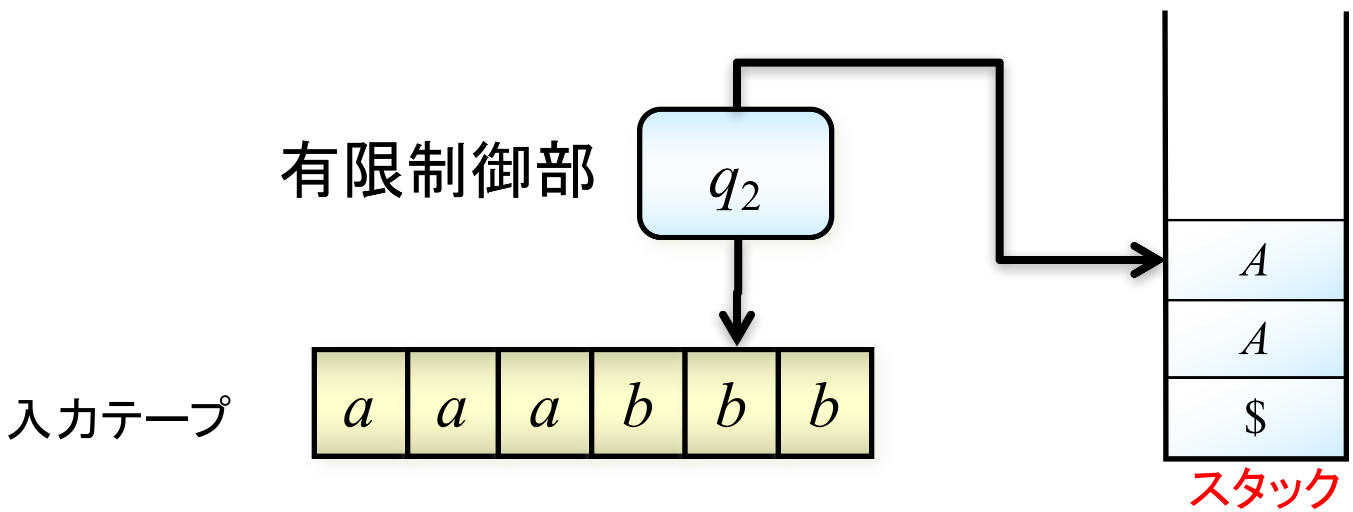
入力 $w=aaabbb$ に対する動作例



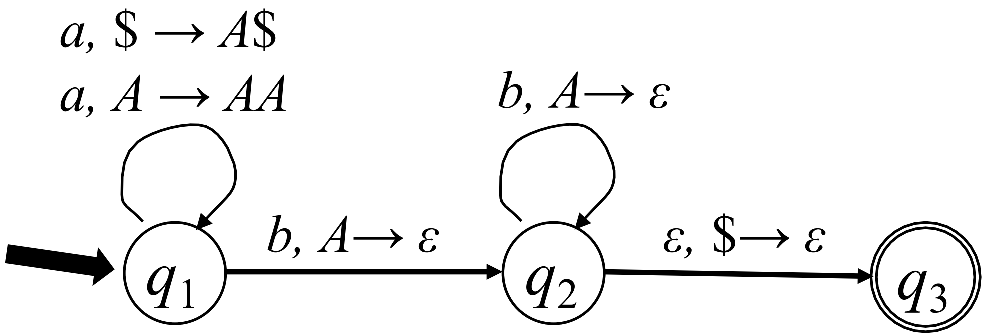


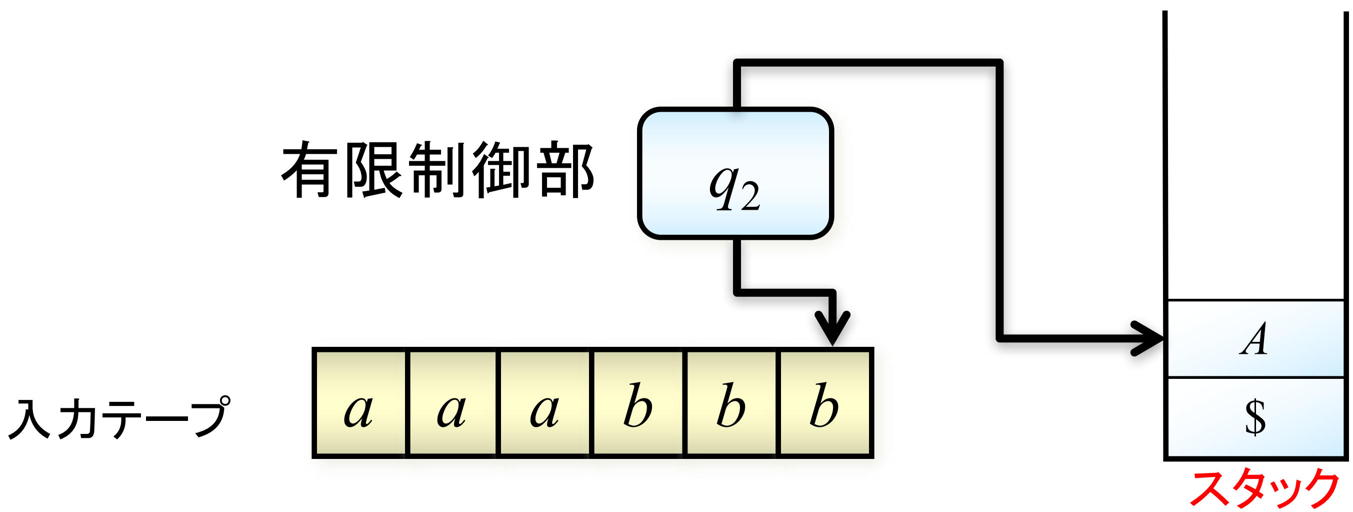
入力 w=$aaabbb$ に対する動作例



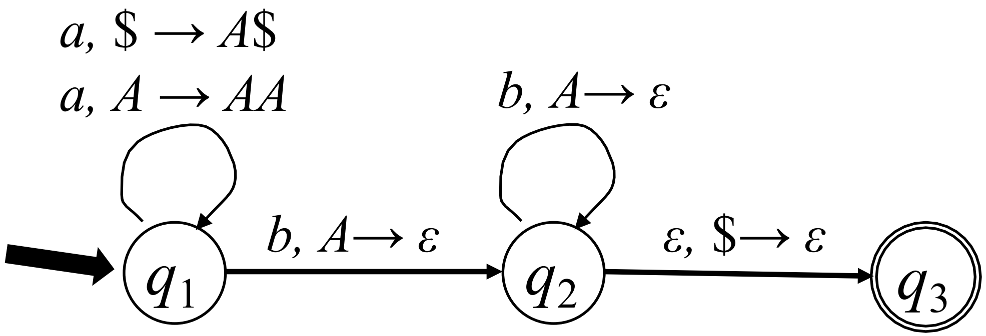


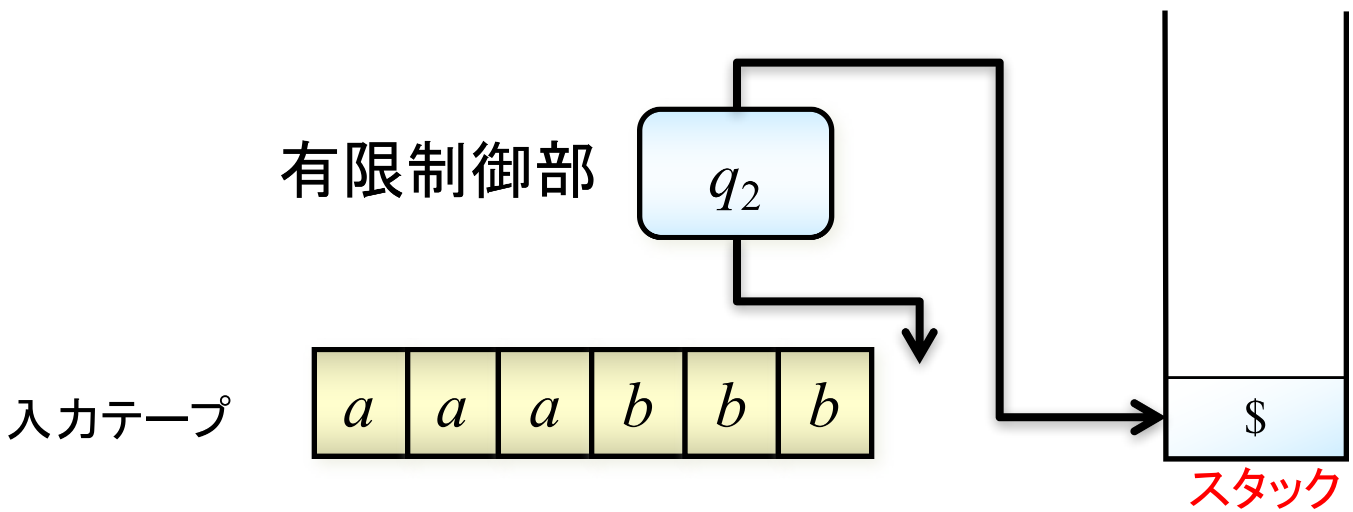
入力 w=*aaabbb* に対する動作例



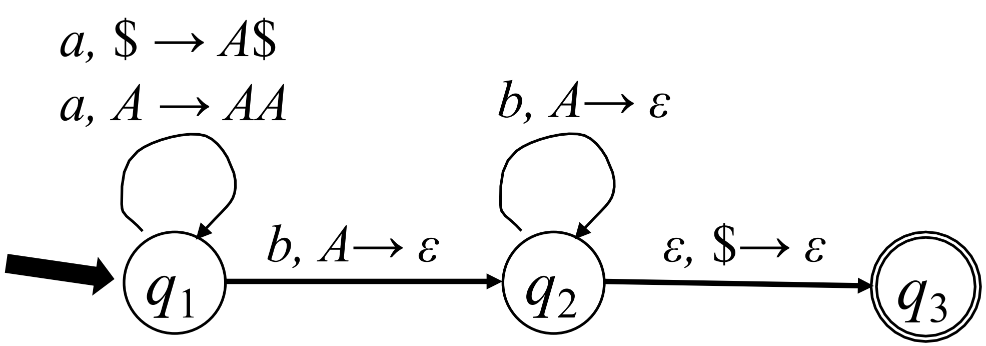


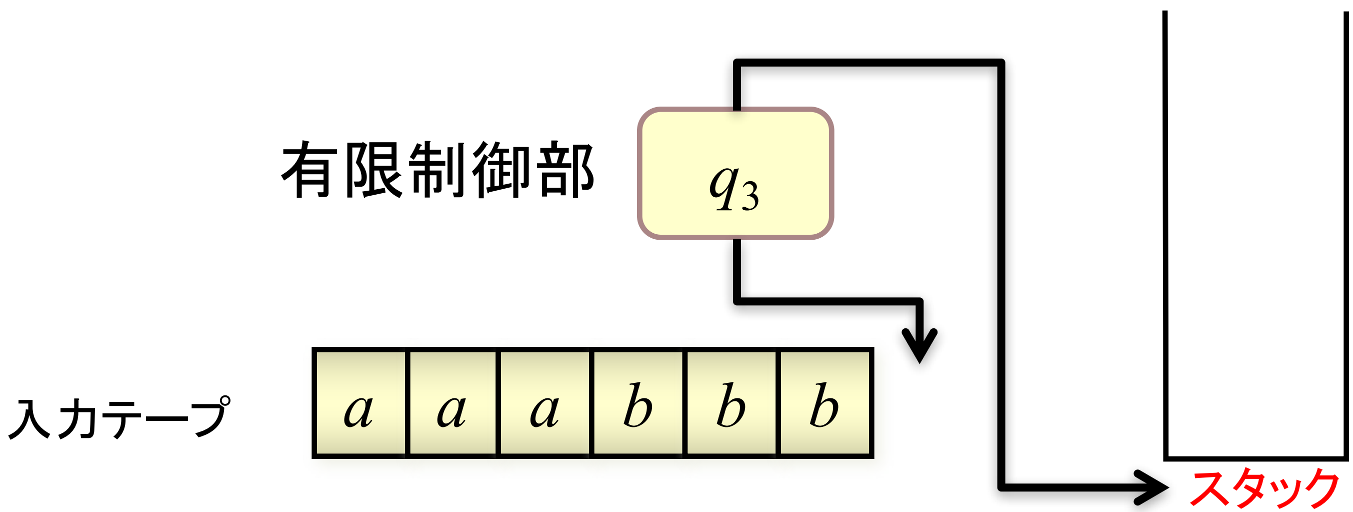
入力 *w*=*aaabbb* に対する動作例





入力 w=$aaabbb$ に対する動作例





演習

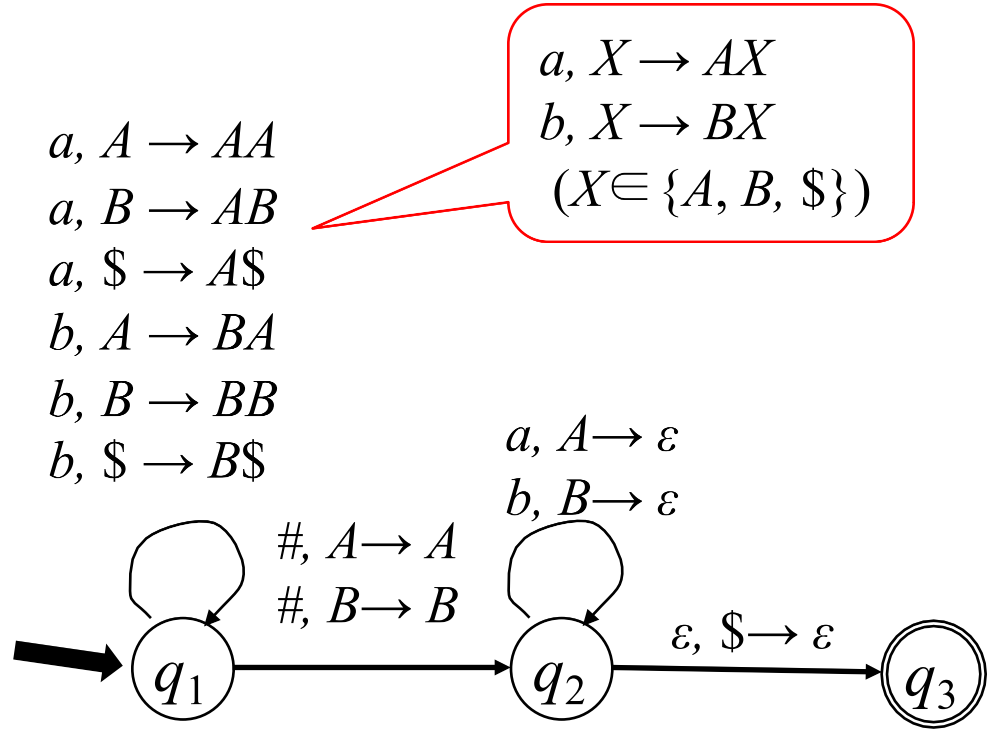
例題5.1

【例題5.1】 $\Sigma=\{a, b, \#\}$とする. Σ 上の言語

$$L=\{u\#u^R \mid u \in \{a, b\}^+\}$$

を最終状態で受理するPDA を与えよ.

状態集合	$\{q_1, q_2, q_3\}$
入力アルファベット	$\{a, b, \#\}$
スタックアルファベット	$\{A, B\}$
状態遷移関数	(右図のとおり)
初期状態	q_1
初期スタック記号	\$
最終状態集合	$\{q_3\}$



演習問題

【演習問題5.1】 $\Sigma=\{a, b\}$とする. Σ 上の言語

$$L=\{uu^R \mid u\in\Sigma^+\}$$

を最終状態で受理するPDA を与えよ.
ただし, 状態数は 3以下 とする.

PDAの受理に関する
その他の定義

受理に関する他の定義

- 有限オートマトンは，入力を読み終わった時点で最終状態に在れば受理する．
- PDAの受理についても同様に定義した．これを最終状態による受理と呼んでいる．
- PDAの受理については，他の定義として次の2つが知られている．
  - 空スタックによる受理
  - 最終状態と空スタックによる受理

空スタックによる受理

【定義】PDA $M=(Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ が文字列 $w \in \Sigma^*$ を空スタックによって受理するとは,

$\exists q \in Q$ に対し

$$(q_0, w, Z_0) \vdash_M^* (q, \varepsilon, \varepsilon)$$

となるときをいう. $(q, \varepsilon, \varepsilon)$ を空スタックによる受理計算状況という.

【定義】PDA $M=(Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ が空スタックによって受理する文字列の集合を

$$L_S(M)=\{w \in \Sigma^* \mid \exists q \in Q, (q_0, w, Z_0) \vdash_M^* (q, \varepsilon, \varepsilon)\}$$

とする.

最終状態と空スタックによる受理

【定義】PDA $M=(Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ が文字列 $w \in \Sigma^*$ を最終状態と空スタックによって受理するとは,
$\exists q \in F$ に対し

$$(q_0, w, Z_0) \vdash_M^* (q, \varepsilon, \varepsilon)$$

となるときをいう. $(q, \varepsilon, \varepsilon)$ を最終状態と空スタックによる受理計算状況という.

【定義】PDA $M=(Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ が最終状態と空スタックによって受理する文字列の集合を

$$L_{\mathrm{FS}}(M)=\{w \in \Sigma^* \mid \exists q \in F, (q_0, w, Z_0) \vdash_M^* (q, \varepsilon, \varepsilon)\}$$

とする.

受理の定義による本質的な違いはない

【定理5.1】 任意の言語 $L \subseteq \Sigma^*$ に対して次の(1)-(3)は等価である.

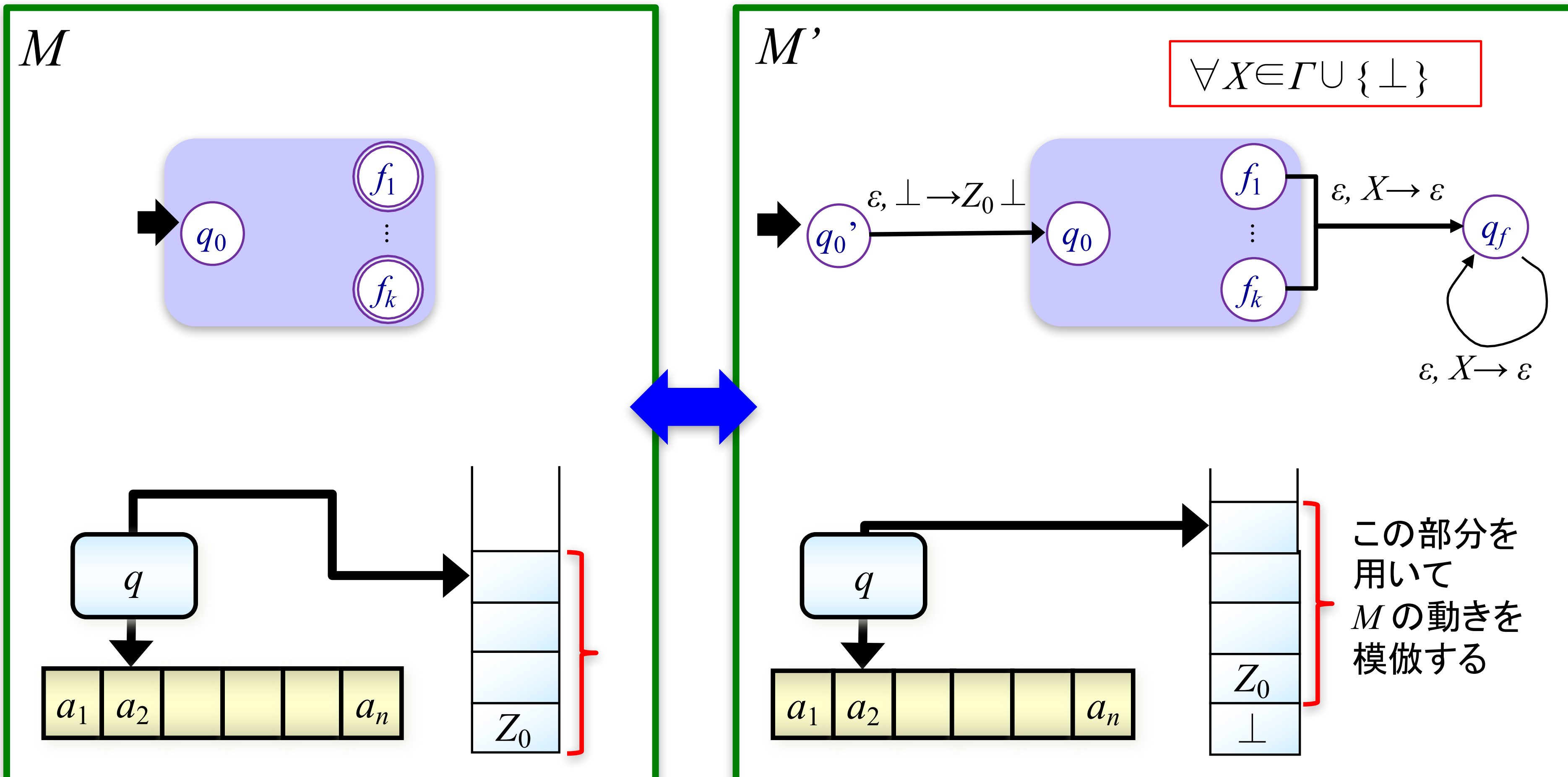
(1) PDA M が存在して $L=L(M)$. 最終状態による受理

(2) PDA M が存在して $L=L_{\mathrm{S}}(M)$. 空スタックによる受理

(3) PDA M が存在して $L=L_{\mathrm{FS}}(M)$. 最終状態と空スタックによる受理

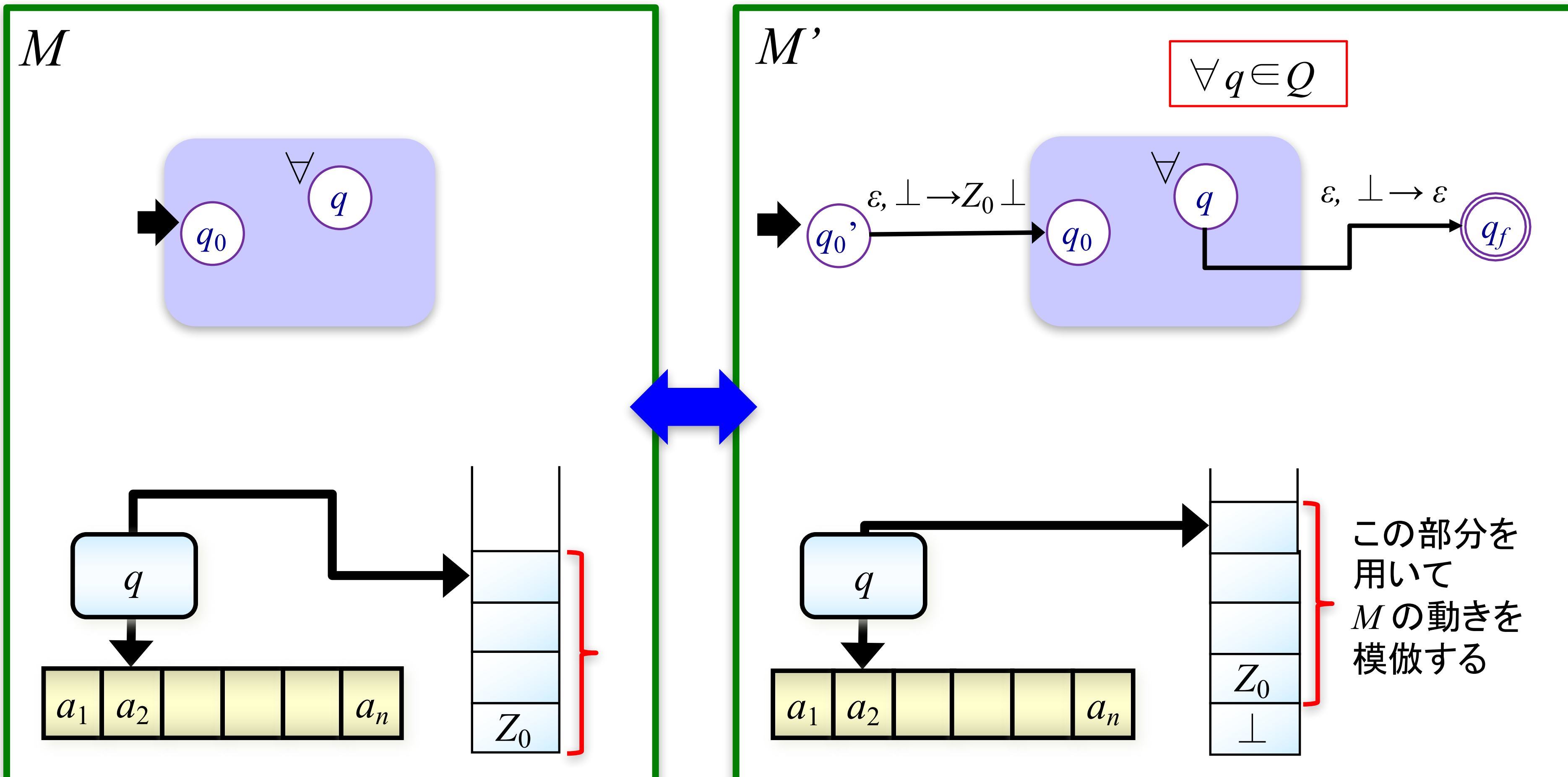
定理5.1の証明:(1)⇒(2)のアイデア

- M'を以下のように定めれば $L(M)=L_S(M')$となる.



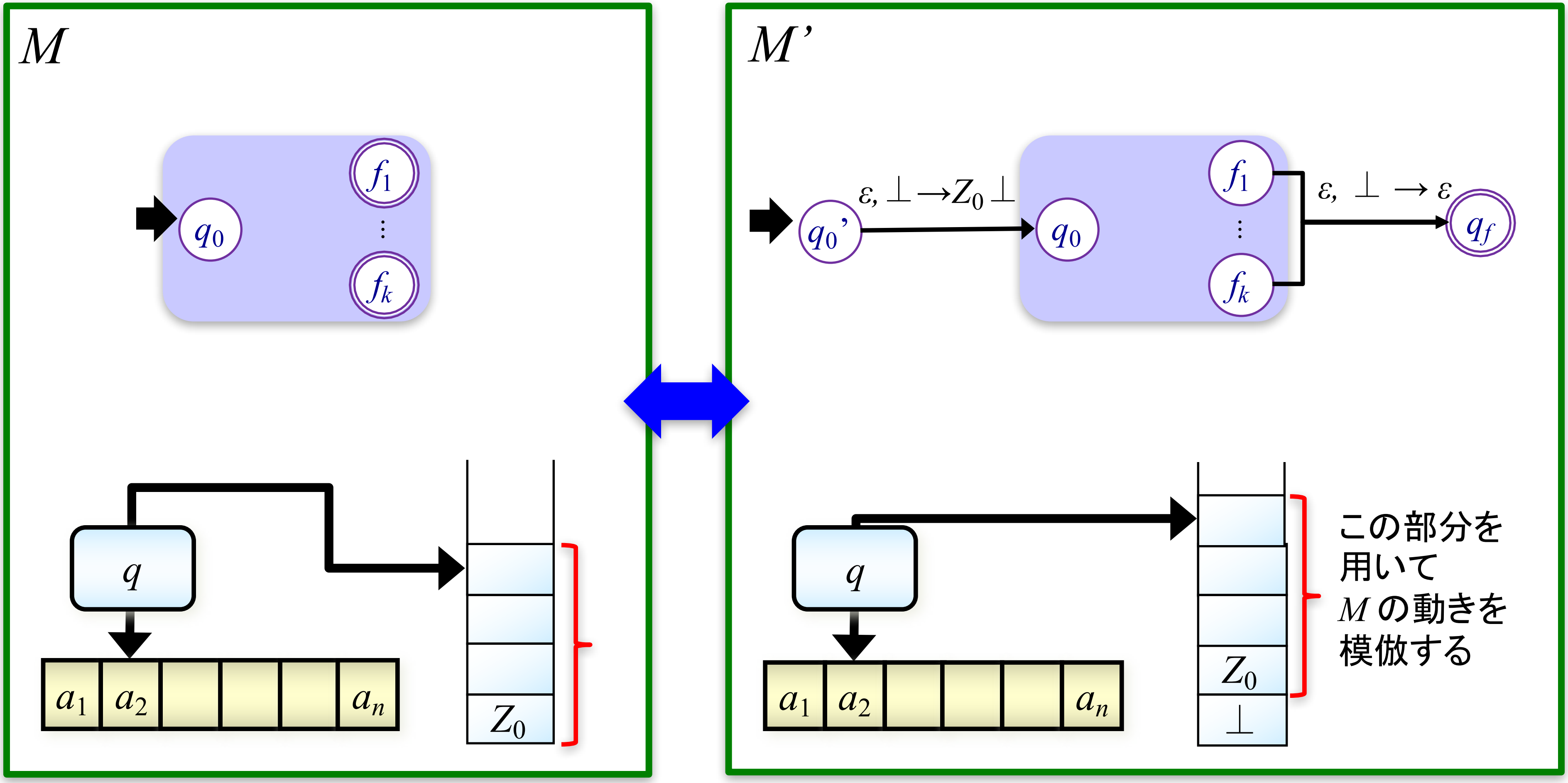
定理5.1の証明：(2)⇒(3)のアイデア

- M' を以下のように定めれば $L_S(M)=L_{FS}(M')$ となる.



定理5.1の証明：(3)⇒(1)のアイデア

- M'を以下のように定めれば $L_{FS}(M)=L(M')$ となる.



PDAの標準形

スタックの記号の置換に関する制限

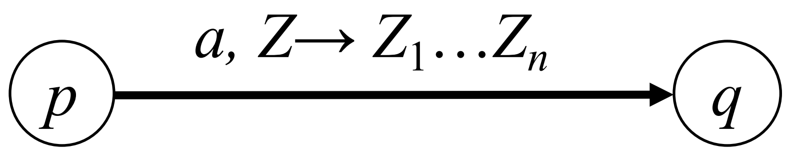
【補題5.1】 任意の PDA $M=(Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$に対し，次の①②を満たす PDA $M'=(Q', \Sigma, \Gamma', \delta', q_0', Z_0', F')$ が存在する．

① $L(M)=L(M')$

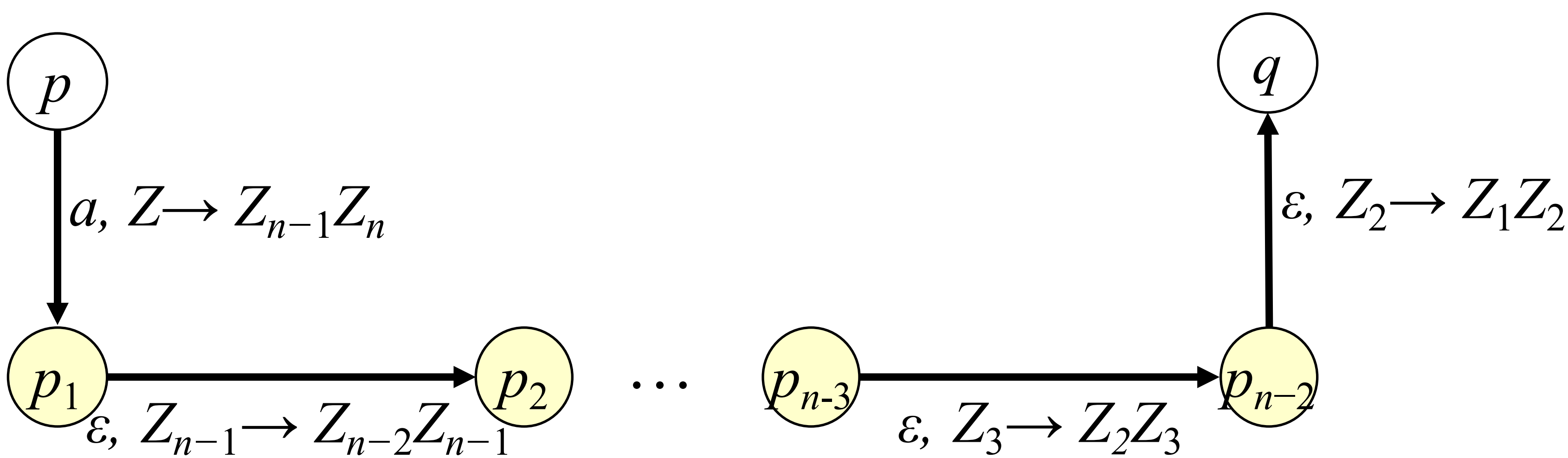
② 任意の $(p, a, Z)\in Q'\times(\Sigma\cup\{\varepsilon\})\times\Gamma'$ に対して，

$$(q, \gamma)\in\delta'(p, a, Z) \Rightarrow |\gamma| \le 2.$$

補題5.1の証明:アイデアのみ



新しい状態 $p_1, \ldots, p_{n-2}$ を導入し
1回の動作を
$n-1$回の動作で置き換える.



PDAの標準形

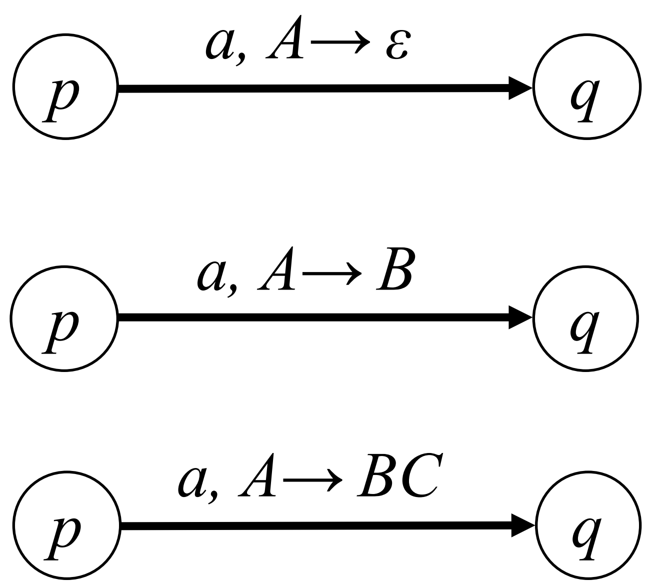
【定義】 PDA $M = (Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ が標準形であるとは，任意の $(p, a, Z) \in Q \times (\Sigma \cup \{\varepsilon\}) \times \Gamma$ と任意の $(q, \gamma) \in \delta(p, a, Z)$ に対して，γ が次のいずれかを満たすときをいう．

$\gamma=\varepsilon$	POPする
$\gamma=Z$	スタック不変
$\gamma=YZ \ (Y \in \Gamma)$	YをPUSHする

【定理5.2】 任意の PDA M に対し，$L(M)=L(M')$ となる標準形の PDA M' が存在する．

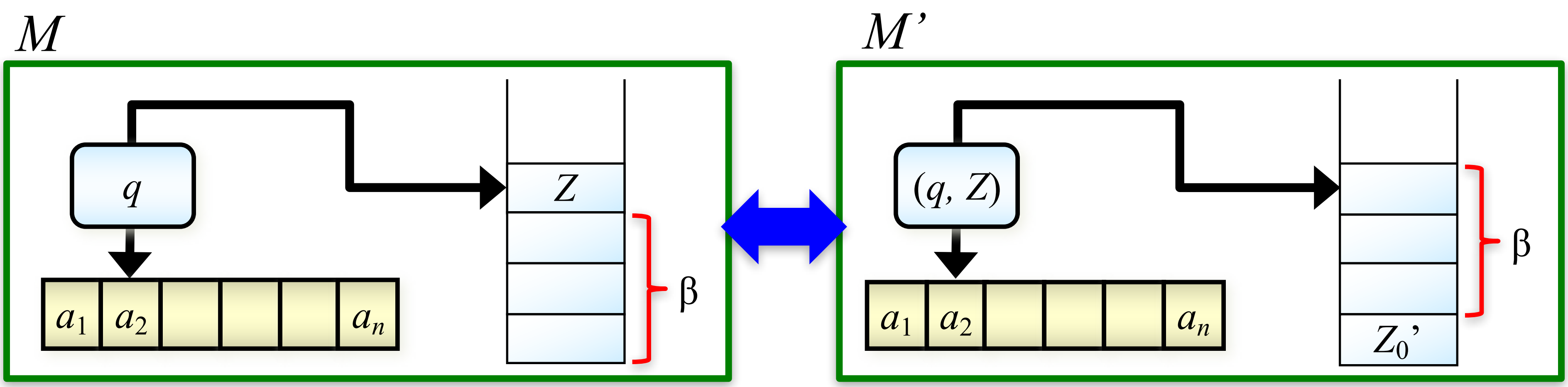
定理5.2の証明：アイデア（1／3）

- $M=(Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ とする．補題5.1より，遷移規則は次の3通りと仮定してよい．ここに，$p, q \in Q$, $a \in \Sigma \cup \{\varepsilon\}$，$A,B,C \in \Gamma$ である．



定理5.2の証明：アイデア（2／3）

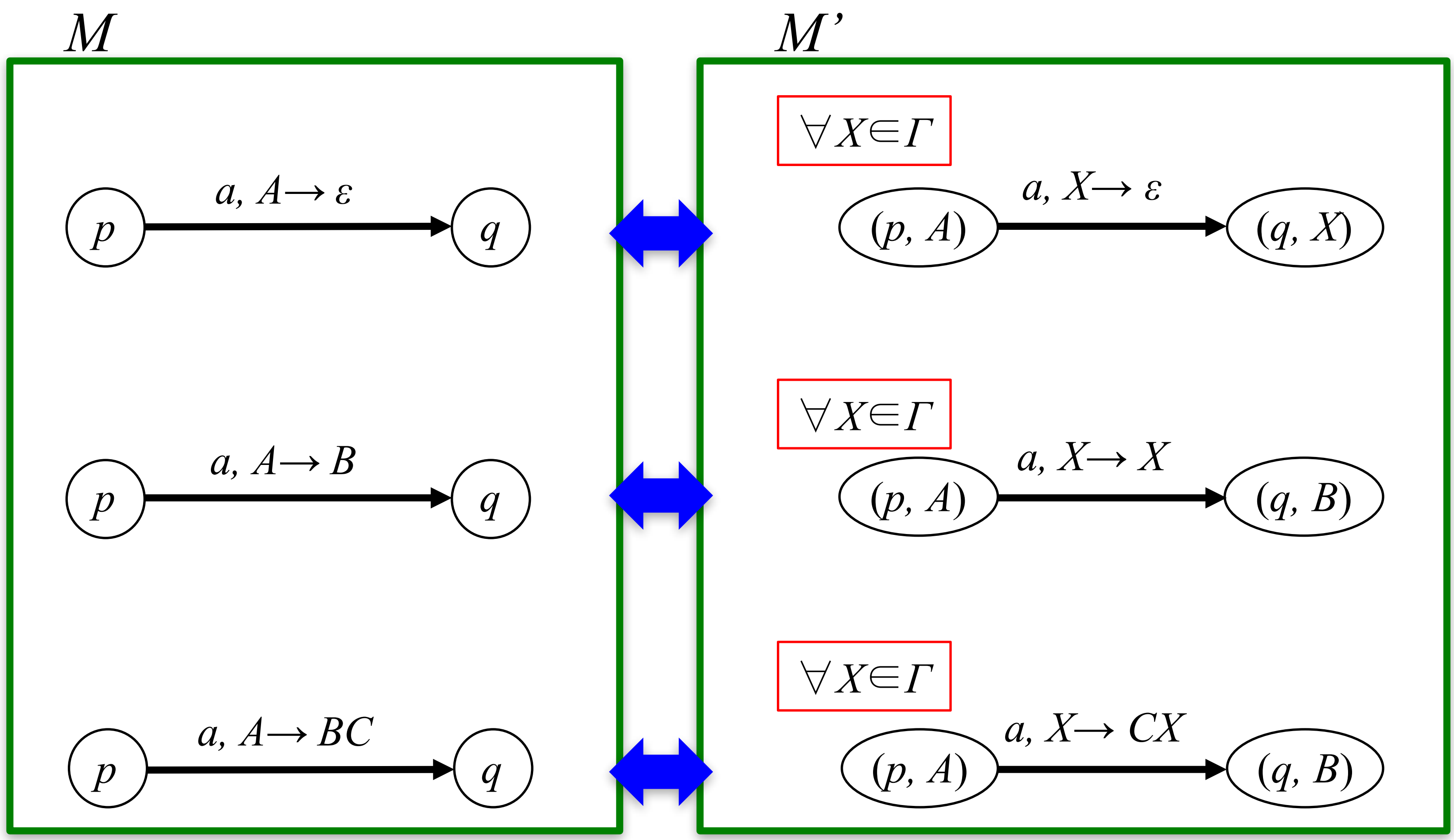
- M と M' の計算状況を以下のように対応させる.



Mのスタックトップ記号を，M'では有限状態制御部に記憶

定理5.2の証明：アイデア(3／3)

- 遷移規則を次のように対応させる.



定理5.2の証明

- $M=(Q,\Sigma,\Gamma,\delta,q_0,Z_0,F)$ とする．補題5.1より，$(q,\gamma)\in\delta(p,a,A) \Rightarrow \gamma\in\{\varepsilon\}\cup\Gamma\cup\Gamma\Gamma$.
- $M'=(Q',\Sigma,\Gamma',\delta',q_0',Z_0',F')$を以下で定めると M' は標準形である.
  - $Q'=Q\times\Gamma$, $\Gamma'=\Gamma\cup\{Z_0'\}$, $q_0'=(q_0,Z_0)$, $F'=F\times\Gamma$.
  - δ'は以下のとおり．ここで，$a\in\Sigma\cup\{\varepsilon\}$, $A,B,C\in\Gamma$.
    - $(q,\varepsilon)\in\delta(p,a,A) \Leftrightarrow \forall X\in\Gamma, ((q,X),\varepsilon)\in\delta'((p,A),a,X)$
    - $(q,B)\in\delta(p,a,A) \Leftrightarrow \forall X\in\Gamma, ((q,B),X)\in\delta'((p,A),a,X)$
    - $(q,BC)\in\delta(p,a,A) \Leftrightarrow \forall X\in\Gamma, ((q,B),CX)\in\delta'((p,A),a,X)$
- $L(M)=L(M')$が成り立つ.

QED

PDAと文脈自由言語

PDAと文脈自由言語

【補題5.2】 任意の CFL L に対して, $L=L_S(M)$ となる PDA M が存在する.

【補題5.3】 任意のPDA M に対して, $L_S(M)$ は CFL である.

補題5.2の証明：アイデア

- $L=L(G)$ となる（広い意味の）Greibach標準形のCFG を $G=(N, \Sigma, P, S)$とする.
- G の生成規則は以下のいずれかの形式である.

$$S \rightarrow \varepsilon, \qquad A \rightarrow a\gamma \ \ (a \in \Sigma, \gamma \in N^*)$$

- G における最左導出

$$S \underset{G}{\Rightarrow}^* u A\beta \underset{G}{\Rightarrow} ua\gamma\beta \underset{G}{\Rightarrow}^* uax = w$$

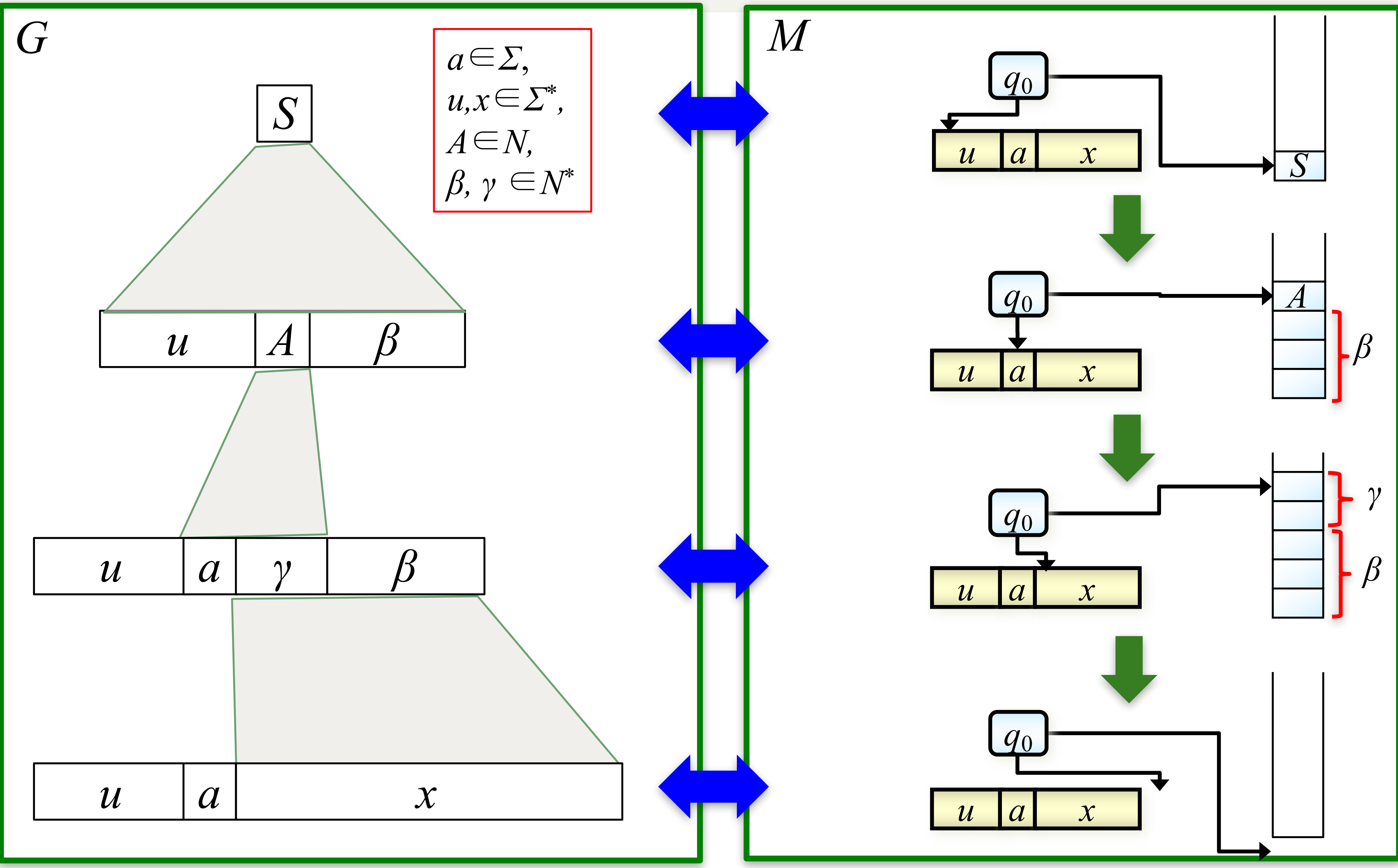
と PDA Mの動作

$$(q_0, uax, S) \vdash_M^* (q_0, ax, A\beta) \vdash_M (q_0, x, \gamma\beta) \vdash_M^* (q_0, \varepsilon, \varepsilon)$$

を次スライドのように対応づける. ここに,

$a \in \Sigma, u, x \in \Sigma^*, A \in N, \beta, \gamma \in N^*$ である.

補題5.2の証明：アイデア



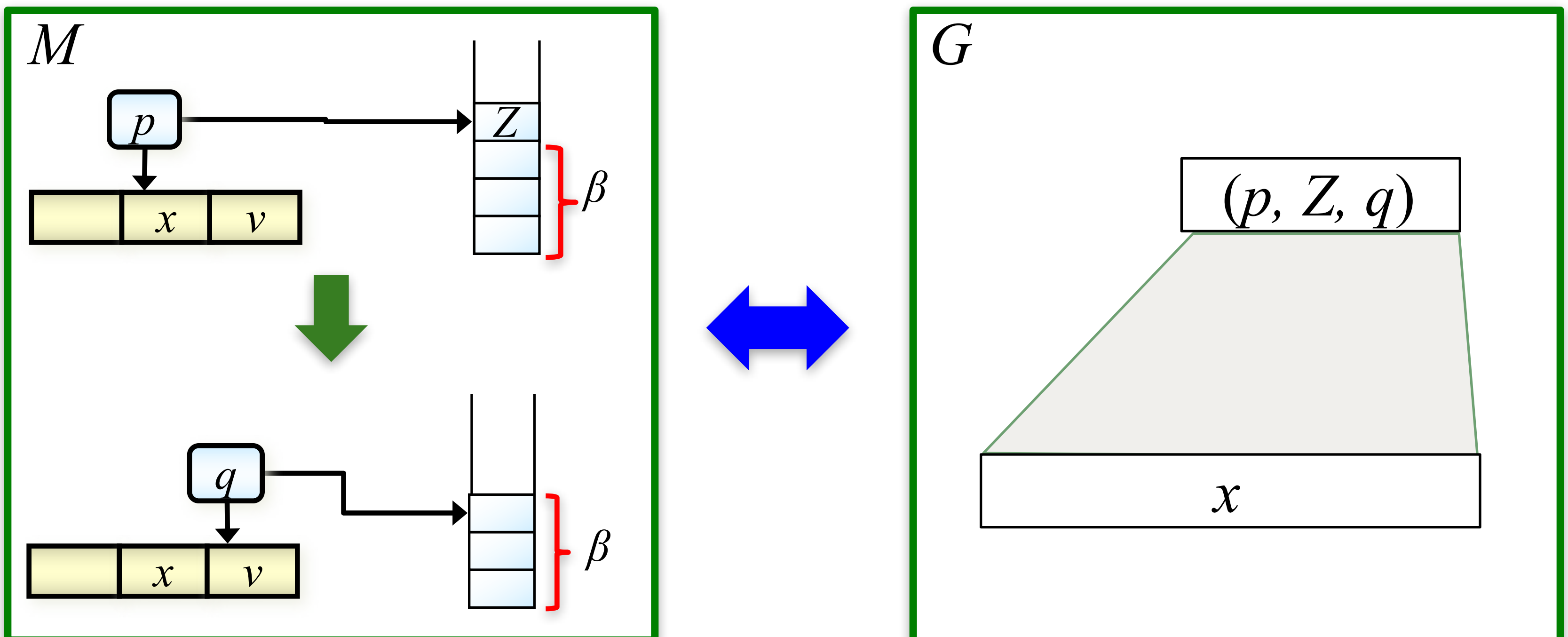
補題5.2の証明

- $L=L(G)$ となる（広い意味の）Greibach標準形のCFGを $G=(N, \Sigma, P, S)$とする.
- PDA $M=(Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ を次のように定める.
  - $Q=\{q_0\}, \Gamma = N, Z_0=S, F=\varnothing$.
  - δ は次のとおり.
    - $\delta(q_0, a, A)=\{(q_0, \gamma) \mid A \rightarrow a\gamma \in P\}\quad (a \in \Sigma, A \in N)$
    - $\delta(q_0, \varepsilon, S)=\{(q_0, \varepsilon) \mid S \rightarrow \varepsilon \in P\}$
- $L=L_S(M)$が成り立つ.

QED

補題5.3の証明：アイデア

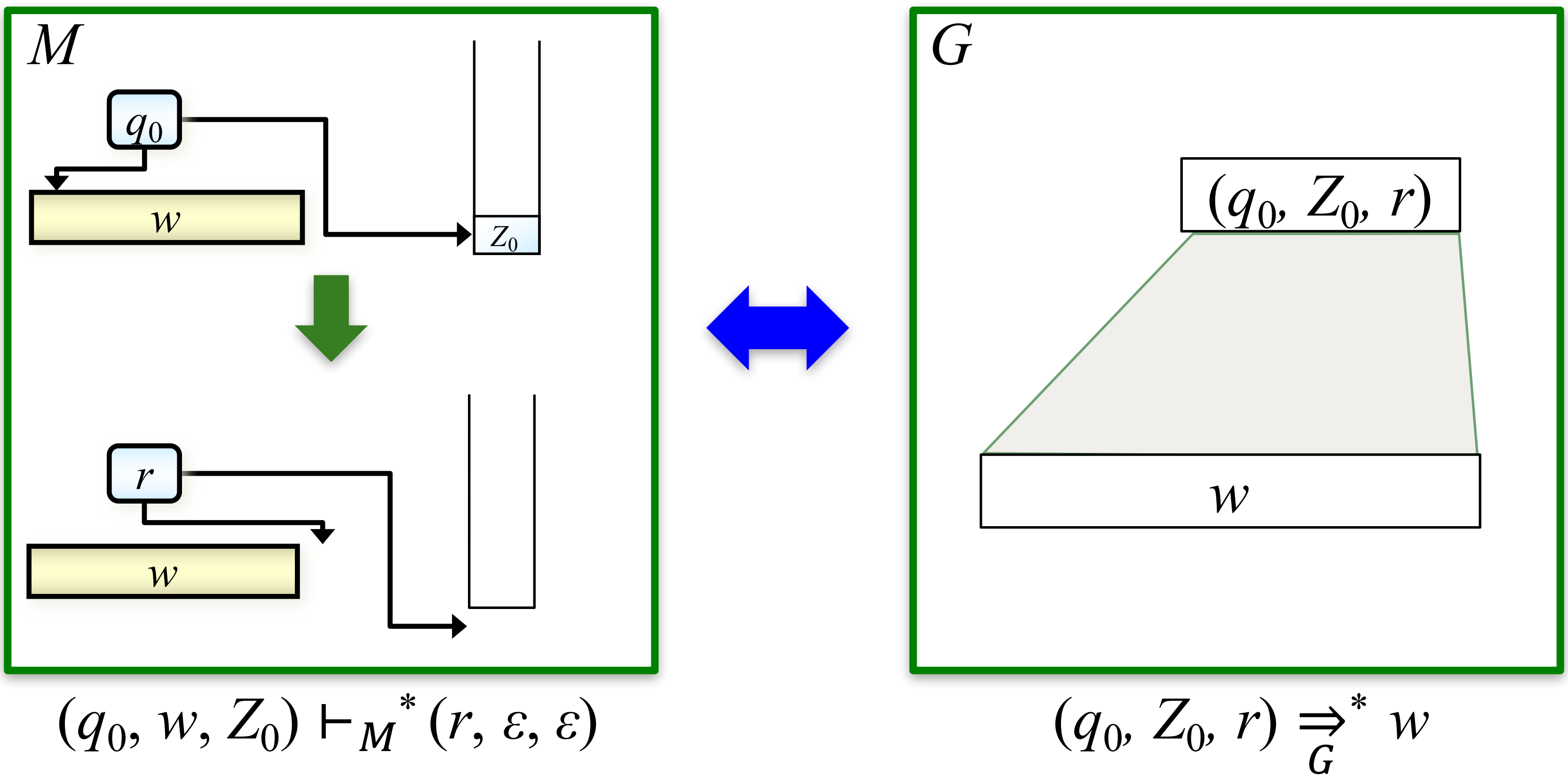
- $(p, Z, q) \in Q \times \Gamma \times Q$ を非終端記号とするCFG G により，M の動きを以下のように模倣する．



$$(p, xv, Z\beta) \vdash_M^* (q, v, \beta) \qquad (p, Z, q) \underset{G}{\Rightarrow}^* x$$

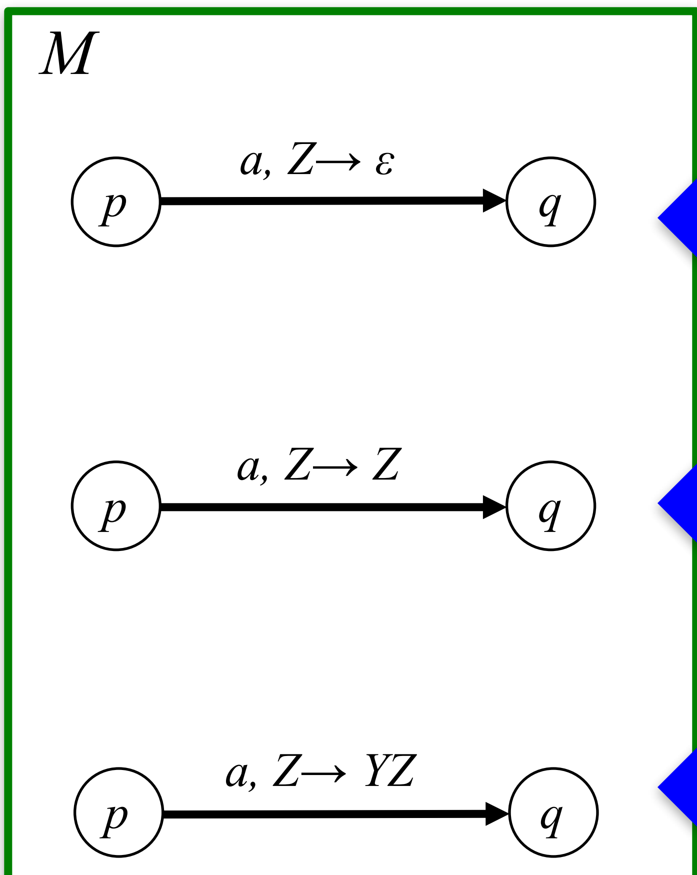
補題5.3の証明:アイデア

- すると, $(q_0, w, Z_0) \vdash_M^* (r, \varepsilon, \varepsilon) \Leftrightarrow (q_0, Z_0, r) \underset{G}{\Rightarrow}^* w$
すなわち, $w \in L_S(M) \Leftrightarrow w \in L(G)$ となる.



補題5.3の証明：アイデア

- PDA $M=(Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ は標準形と仮定し，次のような生成規則を作成すればよい．



G

$(p, Z, q) \to a$

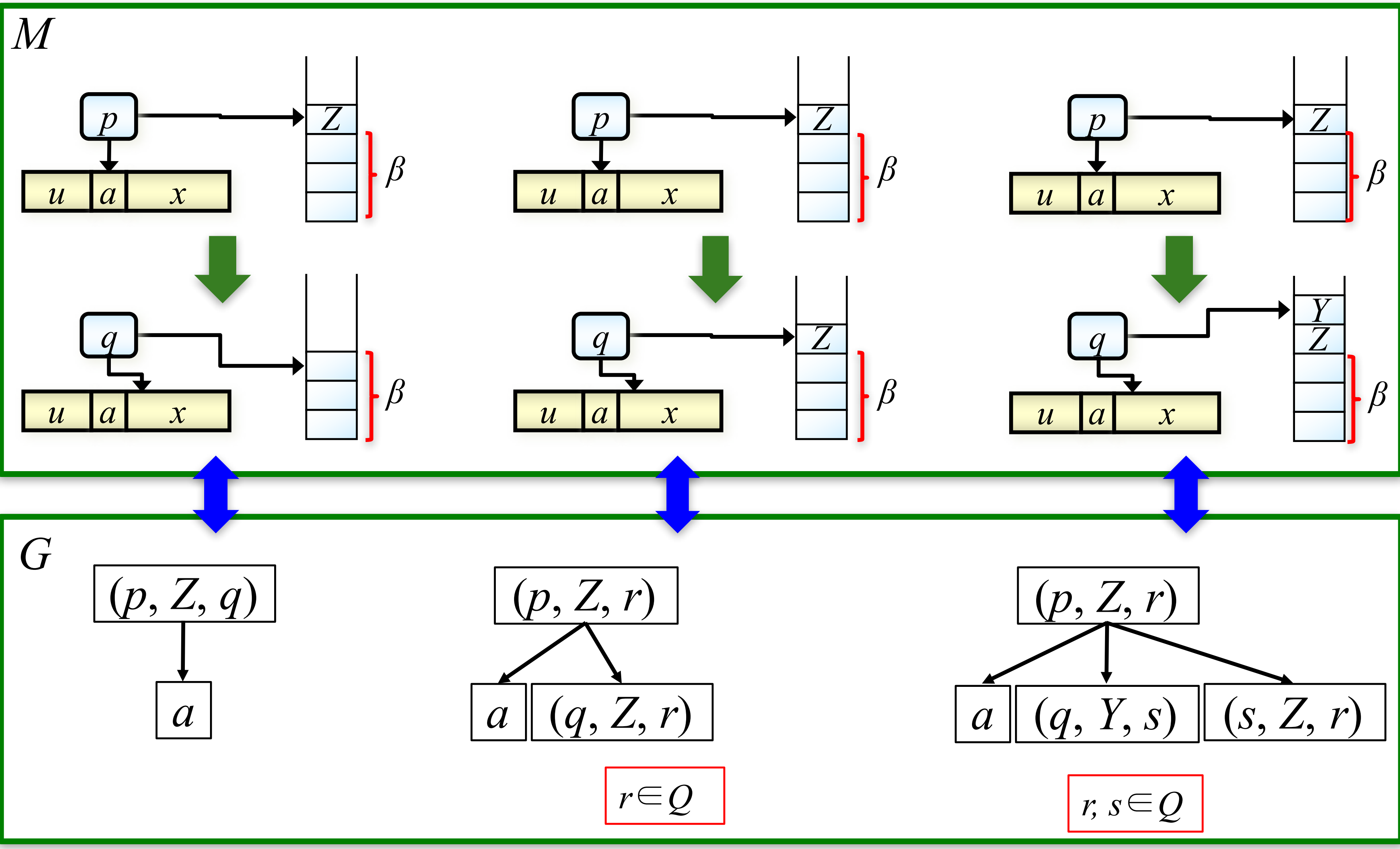
$\boxed{\forall r \in Q}$

$(p, Z, r) \to a\,(q, Z, r)$

$\boxed{\forall r, s \in Q}$

$(p, Z, r) \to a\,(q, Y, s)\,(s, Z, r)$

補題5.3の証明:アイデア



補題5.3の証明

- 定理5.2より，PDA $M=(Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ は標準形であると仮定してよい.
  - $(q, \gamma)\in\delta(p, a, Z)$ ならば，$\gamma = \varepsilon, \gamma = Z, \gamma = YZ \ (Y, Z\in N)$.
- CFG $G=(N, \Sigma, P, S)$を次で定義する.
  - $N = Q\times\Gamma\times Q \cup \{S\}$.
  - $P = \{S \to (q_0, Z_0, r) \mid r\in Q\}$
    $\cup \{ (p, Z, q) \to a \mid (q, \varepsilon)\in\delta(p, a, Z)\}$
    $\cup \{ (p, Z, r) \to a\,(q, Z, r) \mid (q, Z)\in\delta(p, a, Z), r\in Q\}$
    $\cup \{ (p, Z, r) \to a\,(q, Y, s)\,(s, Z, r) \mid (q, YZ)\in\delta(p, a, Z), r, s\in Q\}$
- $L_S(M)=L(G)$が成り立つ.

QED

PDAと文脈自由言語

- 定理5.1と補題5.2，5.3により，次の定理を得る.

【定理5.3】 任意の言語 $L \subseteq \Sigma^*$ に対して次の(1)-(4)は等価である.

(1) L は CFL である.

(2) PDA M が存在して $L=L(M)$. 最終状態による受理

(3) PDA M が存在して $L=L_{\mathrm{S}}(M)$. 空スタックによる受理

(4) PDA M が存在して $L=L_{\mathrm{FS}}(M)$. 最終状態と空スタックによる受理