Kellerautomaten

Eduard Mehofer

Fakultät für Informatik

Währinger Straße 29

Universität Wien

Kellerautomat (engl. Pushdown Automaton)

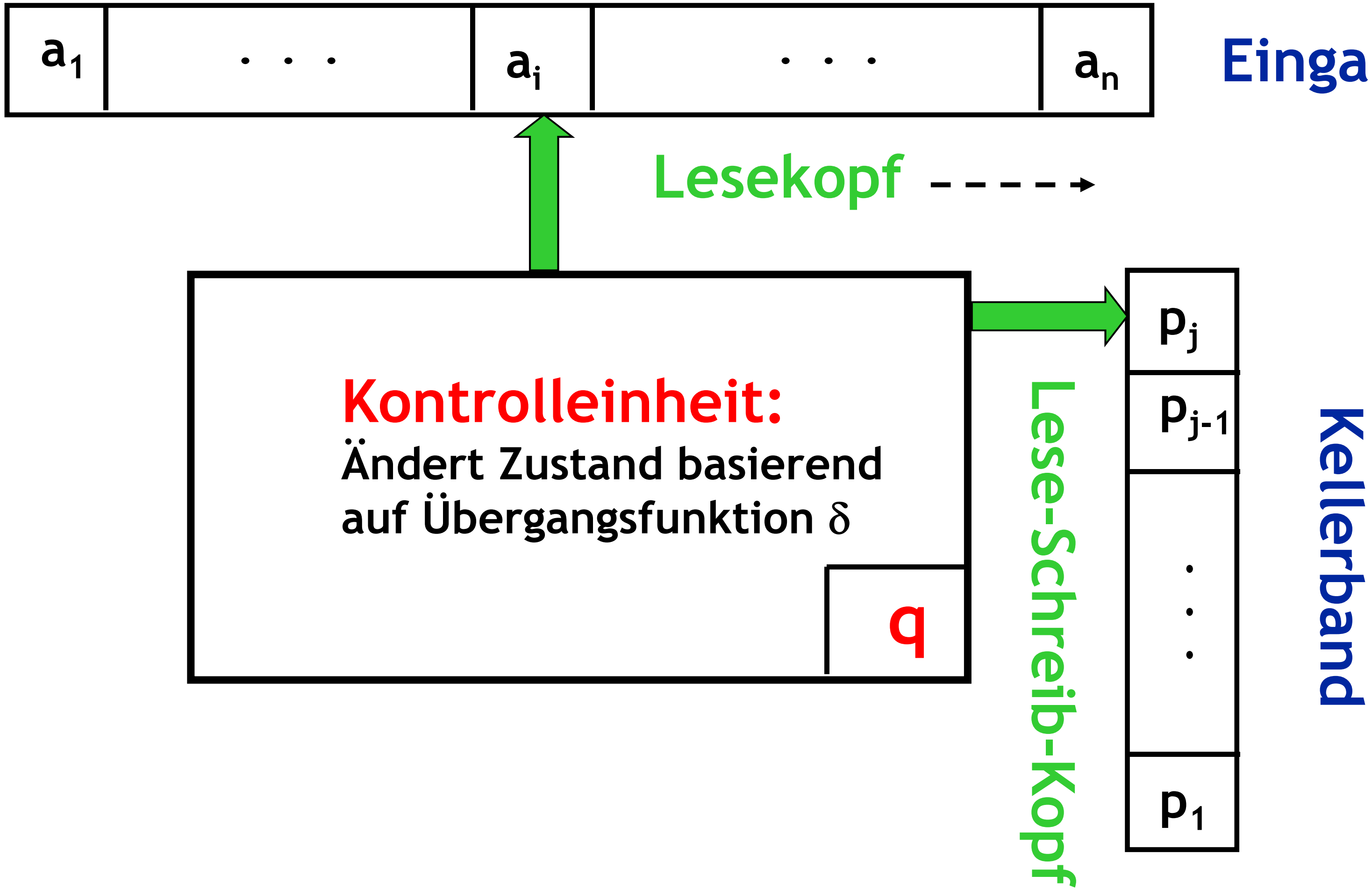
Ein **Kellerautomat** ist eine Erweiterung des ε-NEA um einen **Keller**.

Keller: Speicher bei dem nur das oberste Symbol gelesen oder verändert werden darf, i.e. ein Stack.

Ein Kellerautomat besteht aus folgenden Komponenten:

- **Eingabeband** mit Lesekopf, der die Eingabesymbole von links nach rechts liest
- **Kellerband** mit Lese-Schreib-Kopf, der sich über dem obersten Kellersymbol befindet
- **Kontrolleinheit** mit endlich vielen Zuständen

Graphische Darstellung der Komponenten



Arbeitsweise eines Kellerautomaten (1)

Ein Kellerautomat startet in einem vorgegebenen Anfangszustand (q_0), hat den Lesekopf über dem ersten Symbol des Eingabewortes und hat den Lese-Schreib-Kopf über dem Startsymbol (Z_0) des ansonsten leeren Kellerbandes.

Die Aktionen eines Kellerautomaten werden durch die Überführungsfunktion δ definiert.

Der Inhalt des Kellerbandes ist ein Wort, wobei geschrieben das oberste Symbol des Kellerbandes am weitesten links steht, i.e. $p_j\, p_{j-1} \ldots p_1$.

Arbeitsweise eines Kellerautomaten (2)

In Abhängigkeit vom aktuellen Zustand und den beiden gelesenen Symbolen werden folgende Aktionen durchgeführt:

1. Der Lesekopf rückt um ein Symbol nach rechts oder bleibt über dem aktuellen Symbol (ε-Übergang).
2. Das oberste Kellersymbol wird gelöscht.
3. Ein Wort über dem Kelleralphabet oder ε wird auf das Kellerband geschrieben und der Lese-Schreib-Kopf wird auf das neue oberste Symbol des Kellerbandes gesetzt.
   - schreiben von ε: bewirkt eine pop-Operation
   - schreiben vom gelöschten obersten Kellersymbol: Keller bleibt unverändert
   - schreiben von anderen Symbolen: Keller wird geändert
4. Der Kellerautomat geht in einen neuen Zustand über.

Kellerautomat KA

Definition: Ein (nichtdeterministischer) Kellerautomat ist ein 7-Tupel

$$KA = (Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$$

wobei

- **Q: Menge der Zustände (endlich)**
- **Σ: Eingabealphabet (endlich)**
- **Γ: Kelleralphabet (endlich)**
- **$q_0 \in Q$: Startzustand**
- **$Z_0 \in \Gamma$: Startsymbol auf dem Kellerband**
- **$F \subseteq Q$: Menge der Endzustände (endlich)**
- **$\delta : Q \times (\Sigma \cup \{\varepsilon\}) \times \Gamma \rightarrow$ Menge aus $P(Q \times \Gamma^*)$: Überführungsfunktion**
  (**P** … *Potenzmenge*)

<u>Man beachte</u>: Auf den Stack darf ε (Γ^*) geschrieben werden, gelesen wird ε jedoch vom Stack nie (Γ).

Interpretation von δ

1. $\delta(q, a, Z) = \{ (q_1, \gamma_1), \dots , (q_n, \gamma_n) \}$ mit
   $q_i \in Q$, $a \in \Sigma$, $Z \in \Gamma$, $\gamma_i \in \Gamma^*$

   Der Kellerautomat wählt nichtdeterministisch ein Paar (q_i, γ_i), aus, geht in den Zustand q_i über, löscht Z vom Kellerband, schreibt auf das Kellerband γ_i, rückt den Lesekopf um eins nach rechts und setzt den Lese-Schreib-Kopf über das erste Symbol von γ_i.

2. $\delta(q, \varepsilon, Z) = \{ (q_1, \gamma_1), \dots , (q_n, \gamma_n) \}$

   Der Lesekopf liest nichts und rückt daher auch nicht nach rechts, sonst bleibt alles ident zu 1, d.h. ein Zustandswechsel ohne ein Eingabesymbol zu konsumieren.
   <u>Achtung</u>: Bedeutet <u>NICHT</u> das Eingabeband ist leer!

<u>Achtung</u>: $\delta(q, a, \varepsilon)$ mit ε in dritter Position ist nicht erlaubt!

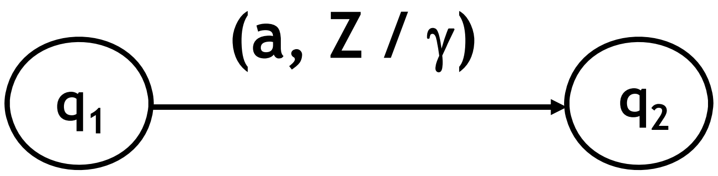
Übergangsdiagramm eines KA

Definition: Das Übergangsdiagramm für einen KA ist ein gerichteter Graph, der wie folgt definiert ist:

- Die Knoten des Graphen entsprechen den Zuständen des KA.
- Eine mit **$(a, Z\,/\,\gamma)$** bezeichnete Kante von Zustand **q_1** nach Zustand **q_2** beschreibt **$(q_2,\gamma) \in \delta(q_1, a, Z)$**.

Der Anfangszustand wird durch einen mit **'Start'** markierten Pfeil angegeben und die Endzustände werden mit einem **doppelten Kreis** markiert.

Bsp.:



Konfiguration und Überführungsrelation

Eine **Konfiguration** eines Kellerautomaten $KA = (Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ ist ein Tripel $(q, w, \gamma) \in Q \times \Sigma^* \times \Gamma^*$ mit

- q aktueller Zustand,
- w zu verarbeitendes Restwort rechts vom Lesekopf,
- γ aktueller Kellerinhalt.

Eine **Überführungsrelation** $\vdash_{KA} \subseteq (Q \times \Sigma^* \times K^*) \times (Q \times \Sigma^* \times K^*)$ ist wie folgt definiert:

- $(q_1, aw, Z\gamma) \vdash_{KA} (q_2, w, \gamma_1\gamma)$ genau dann, wenn $(q_2,\gamma_1) \in \delta(q_1, a, Z)$

  mit $q_1, q_2 \in Q$, $a \in \Sigma \cup \{\varepsilon\}$, $w \in \Sigma^*$, $Z \in \Gamma$ und $\gamma, \gamma_1 \in \Gamma^*$

Die Relation $\vdash_{KA}^{*}$ ist die **reflexive und transitive Hülle** von $\vdash_{KA}$.

Akzeptierte Sprache

Definition: Sei KA = $(Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ ein Kellerautomat. Dann ist

$$\mathbf{L(KA)} = \{ w \in \Sigma^* \mid (q_0, w, Z_0) \vdash^*_{KA} (q, \varepsilon, \gamma) \text{ mit } q \in F \text{ und } \gamma \in \Gamma^* \}$$

die von KA durch **Endzustand** akzeptierte Sprache.

Man beachte:

a) Ein Kellerautomat akzeptiert ein Wort nur, wenn es **zur Gänze** gelesen wurde und man sich in einem Endzustand befindet (analog zu einem endlichen Automaten).

b) Ein leeres Kellerband ist in dieser Definition nicht gefordert.

Anm.: Dieser Akzeptanz-Ansatz wird in der VO und UE verwendet und wird bei den Beispielen durch die Angabe "durch Endzustand" verdeutlicht.

Alternative Sprachakzeptanz-Regeln

Sei KA = $(Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ ein Kellerautomat. Dann ist

1) Sprache L(KA), die von KA <u>durch Endzustand akzeptiert</u> wird, ist

   $L(KA) = \{ w \in \Sigma^* \mid (q_0, w, Z_0) \vdash^*_{KA} (q, \varepsilon, \gamma) \text{ mit } q \in F \text{ und } \gamma \in \Gamma^* \}$

2) Sprache N(KA), die von KA <u>durch leeres Kellerband akzeptiert</u> wird, ist

   $N(KA) = \{ w \in \Sigma^* \mid (q_0, w, Z_0) \vdash^*_{KA} (q, \varepsilon, \varepsilon) \text{ mit } q \in Q \}$

3) Sprache LN(KA), die von KA <u>durch Endzustand und leeres Kellerband akzeptiert</u> wird, ist

   $LN(KA) = \{ w \in \Sigma^* \mid (q_0, w, Z_0) \vdash^*_{KA} (q, \varepsilon, \varepsilon) \text{ mit } q \in F \}$

Sprachen und Kellerautomaten

Sei $L \subseteq \Sigma^*$ eine formale Sprache. Dann sind folgende Aussagen äquivalent:

1) Es gibt eine kontextfreie Grammatik G mit L = L(G).

2) Es gibt einen Kellerautomaten KA mit L = L(KA). (*Endzustand*)

3) Es gibt einen Kellerautomaten KA mit L = N(KA). (*leeres Kellerband*)

4) Es gibt einen Kellerautomaten KA mit L = LN(KA). (*Endzustand+leeres K.*)

Beispiel 1:

Gesucht ist ein Kellerautomat für die Sprache L = $\{ vv^R \mid v \in \{0, 1\}^* \}$, wobei v^R das gespiegelte Wort von v bezeichnet, i.e. Palindrome gerader Länge über {0,1}, der <u>durch Endzustand akzeptiert</u>.

Grammatik G=({S}, {0,1}, { S $\rightarrow$ 0S0|1S1|ε }, S) mit L(G)=L.

Für diese Sprache gibt es **keinen** deterministischen Kellerautomaten!

Strategie: v wird eingelesen und eine Kopie im Keller gespeichert, die Mitte muss geraten werden, und v^R wird mit den Symbolen im Keller verglichen.

KA = $(Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$

$\Sigma = \{0, 1\}$, $Q = \{q_0, q_1, q_2\}$, $\Gamma = \{0, 1, Z_0\}$, $F = \{q_2\}$

Überführungsfunktion **δ** :

Bearbeitung des ersten Symbols von v – wird in den Keller gegeben:

1. $\delta(q_0, 0, Z_0) = \{(q_0, 0Z_0)\}$
2. $\delta(q_0, 1, Z_0) = \{(q_0, 1Z_0)\}$

Beispiel 1 (Fortsetzung): δ

Alle weiteren Symbole von v werden im Keller abgespeichert:

3. $\delta(q_0, 0, 0) = \{(q_0, 00)\}$
4. $\delta(q_0, 0, 1) = \{(q_0, 01)\}$
5. $\delta(q_0, 1, 0) = \{(q_0, 10)\}$
6. $\delta(q_0, 1, 1) = \{(q_0, 11)\}$

Ist die Mitte des Eingabewortes erreicht, wechselt man nichtdeterministisch mit einem ε-Übergang in den Zustand q_1, wobei der Keller nicht verändert wird. Die Mitte wird erraten:

7. $\delta(q_0, \varepsilon, Z_0) = \{(q_1, Z_0)\}$
8. $\delta(q_0, \varepsilon, 0) = \{(q_1, 0)\}$
9. $\delta(q_0, \varepsilon, 1) = \{(q_1, 1)\}$

Beispiel 1 (Fortsetzung): δ

Im Zustand q_1 gilt die Annahme, dass die Mitte des Wortes erreicht wurde und die Übereinstimmung von v mit v^R geprüft werden muss. Die Eingabesymbole und Kellersymbole müssen übereinstimmen und der Keller wird abgebaut, d.h. bei Übereinstimmung wird eine pop-Operation durchgeführt:

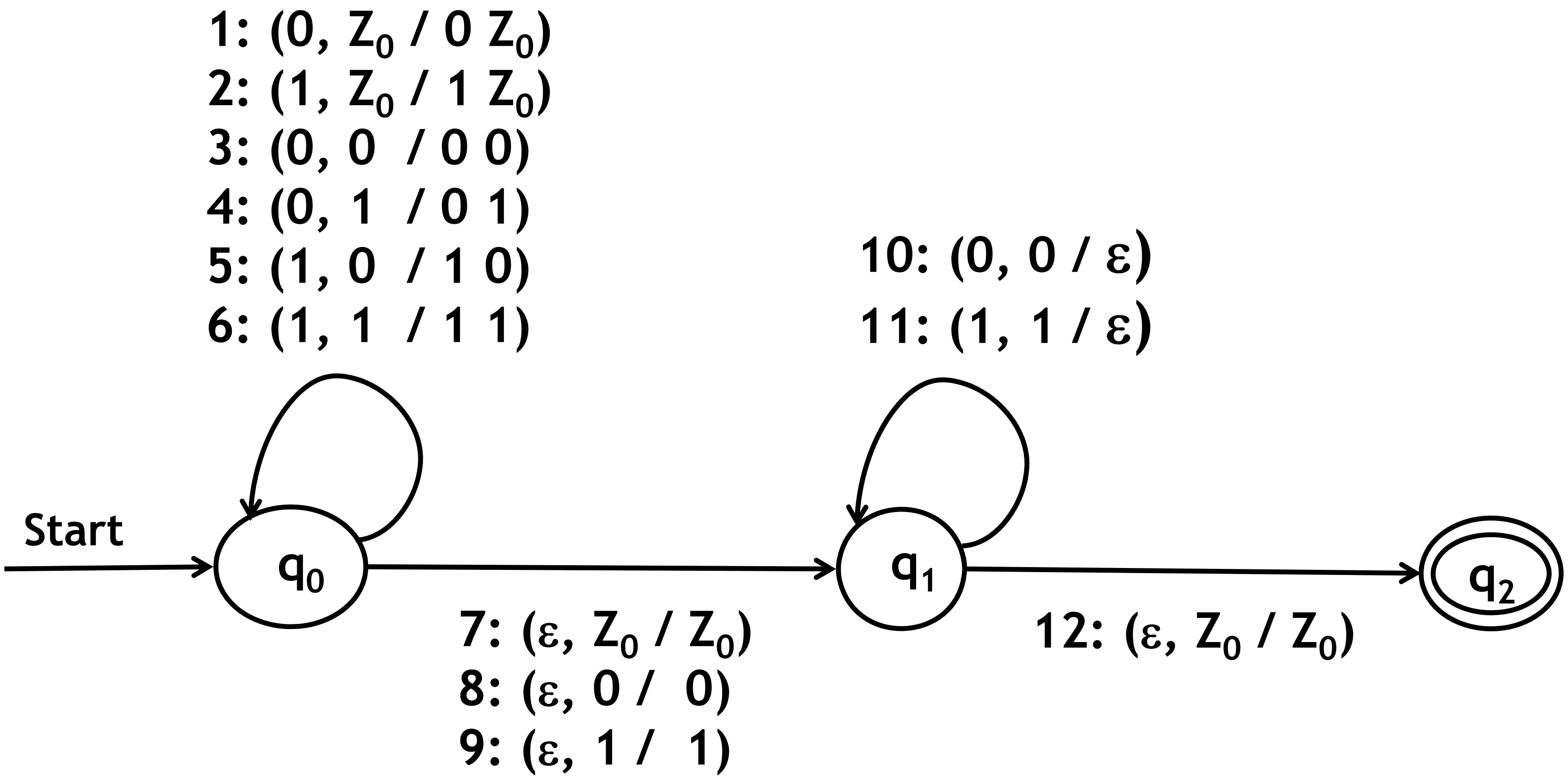
10. $\delta(q_1, 0, 0) = \{(q_1, \varepsilon)\}$
11. $\delta(q_1, 1, 1) = \{(q_1, \varepsilon)\}$

Stößt man auf das Startsymbol des Kellers, geht man in den Endzustand über:

12. $\delta(q_1, \varepsilon, Z_0) = \{(q_2, Z_0)\}$

Das Wort wird im Endzustand akzeptiert, wenn es zur Gänze gelesen wurde, sonst nicht. Das Kellerband ist in obiger Lösung nicht leer. Wird (12) durch $\delta(q_1, \varepsilon, Z_0) = \{(q_2, \varepsilon)\}$ ersetzt um das Startsymbol des Kellers zu löschen, ist der Keller bei Akzeptierung leer.

Beispiel 1 (Fortsetzung): Übergangsdiagramm



Beispiel 1 (Fortsetzung): Analyse

Nichtdeterministische Analyse des Eingabewortes **1111**:

a) $(q_0, 1111, Z_0) \vdash (q_0, 111, 1Z_0) \vdash (q_0, 11, 11Z_0) \vdash$
$(q_1, 11, 11Z_0) \vdash (q_1, 1, 1Z_0) \vdash (q_1, \varepsilon, Z_0) \vdash (q_2, \varepsilon, Z_0)$

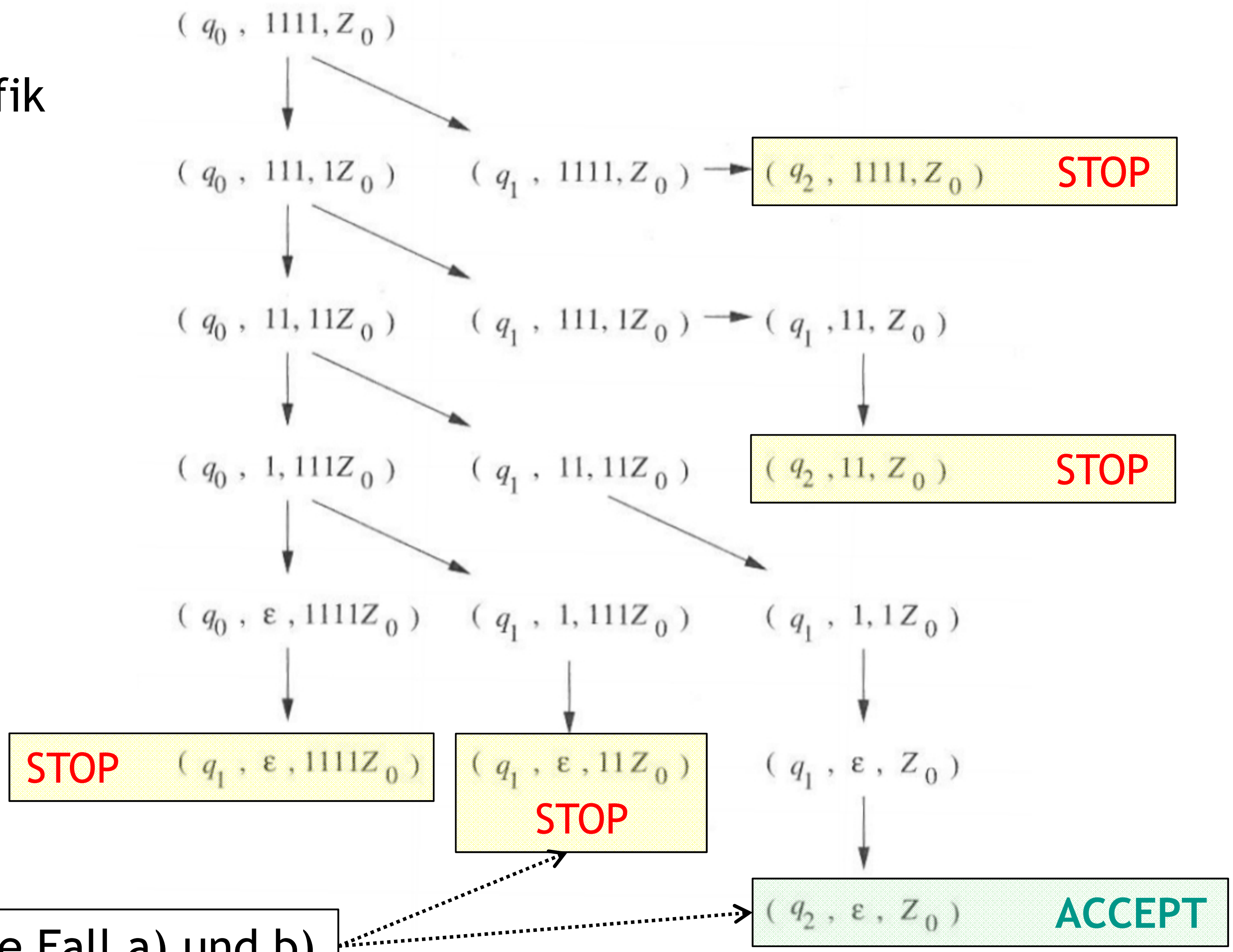
Der Kellerautomat akzeptiert das Eingabewort **1111** durch Endzustand.

b) $(q_0, 1111, Z_0) \vdash (q_0, 111, 1Z_0) \vdash (q_0, 11, 11Z_0) \vdash$
$(q_0, 1, 111Z_0) \vdash (q_1, 1, 111Z_0) \vdash (q_1, \varepsilon, 11Z_0)$

Der Kellerautomat erreicht bei Eingabewort **1111** keinen Endzustand.

Beispiel 1 (Fortsetzung): Analysemöglichkeiten

Anm.:
Pfeile in Grafik entsprechen
⊢ Relation



vorige Folie Fall a) und b)

Deterministischer KA

Bei einem det. KA gibt es für jede Konfiguration $(q, aw, Z\gamma)$ höchstens **eine** Folgekonfiguration.

Definition: Ein Kellerautomat KA = $(Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ ist **deterministisch**, falls für alle $q \in Q$, $a \in \Sigma$, $Z \in \Gamma$ gilt, dass

1. $| \delta(q, a, Z) | \leq 1$

2. $| \delta(q, \varepsilon, Z) | \leq 1$

3. falls $\delta(q, \varepsilon, Z) \neq \emptyset$, dann $\delta(q, a, Z) = \emptyset$ für alle $a \in \Sigma$

   d.h. in Zustand q mit Kellersymbol Z hat man gar keine andere Möglichkeit als im Übergangsdiagramm die ε-Kante zu wählen.

Beispiel 2:

Gesucht ist ein deterministischer Kellerautomat für die Sprache $L = \{ vcv^R \mid v \in \{0, 1\}^*\}$, wobei v^R das gespiegelte Wort von v ist, i.e. Palindrome ungerader Länge über {0,1} mit Mittemarkierung c, der <u>durch Endzustand akzeptiert</u>.
Grammatik $G=(\{S\}, \{0,1,c\}, \{ S \rightarrow 0S0|1S1|c \}, S)$ mit L(G)=L.
Strategie ident zu Beispiel 1.

$KA = (Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$
$\Sigma = \{0, 1, c\}$, $Q = \{q_0, q_1, q_2\}$, $\Gamma = \{0,1,Z_0\}$ (auch möglich: $\Gamma = \{X,Y,Z_0\}$), $F = \{q_2\}$

Überführungsfunktion δ:
Bearbeitung des ersten Symbols von v:

1. $\delta(q_0, 0, Z_0) = \{(q_0, 0Z_0)\}$
2. $\delta(q_0, 1, Z_0) = \{(q_0, 1Z_0)\}$

Bearbeitung der folgenden Symbole von v:

3. $\delta(q_0, 0, 0) = \{(q_0, 00)\}$
4. $\delta(q_0, 0, 1) = \{(q_0, 01)\}$
5. $\delta(q_0, 1, 0) = \{(q_0, 10)\}$
6. $\delta(q_0, 1, 1) = \{(q_0, 11)\}$

Beispiel 2 (Fortsetzung)

Bei Auftreten von c geht der Automat in den Zustand q_1 über, der Keller bleibt unverändert:

7. $\delta(q_0, c, Z_0) = \{(q_1, Z_0)\}$
8. $\delta(q_0, c, 0) = \{(q_1, 0)\}$
9. $\delta(q_0, c, 1) = \{(q_1, 1)\}$

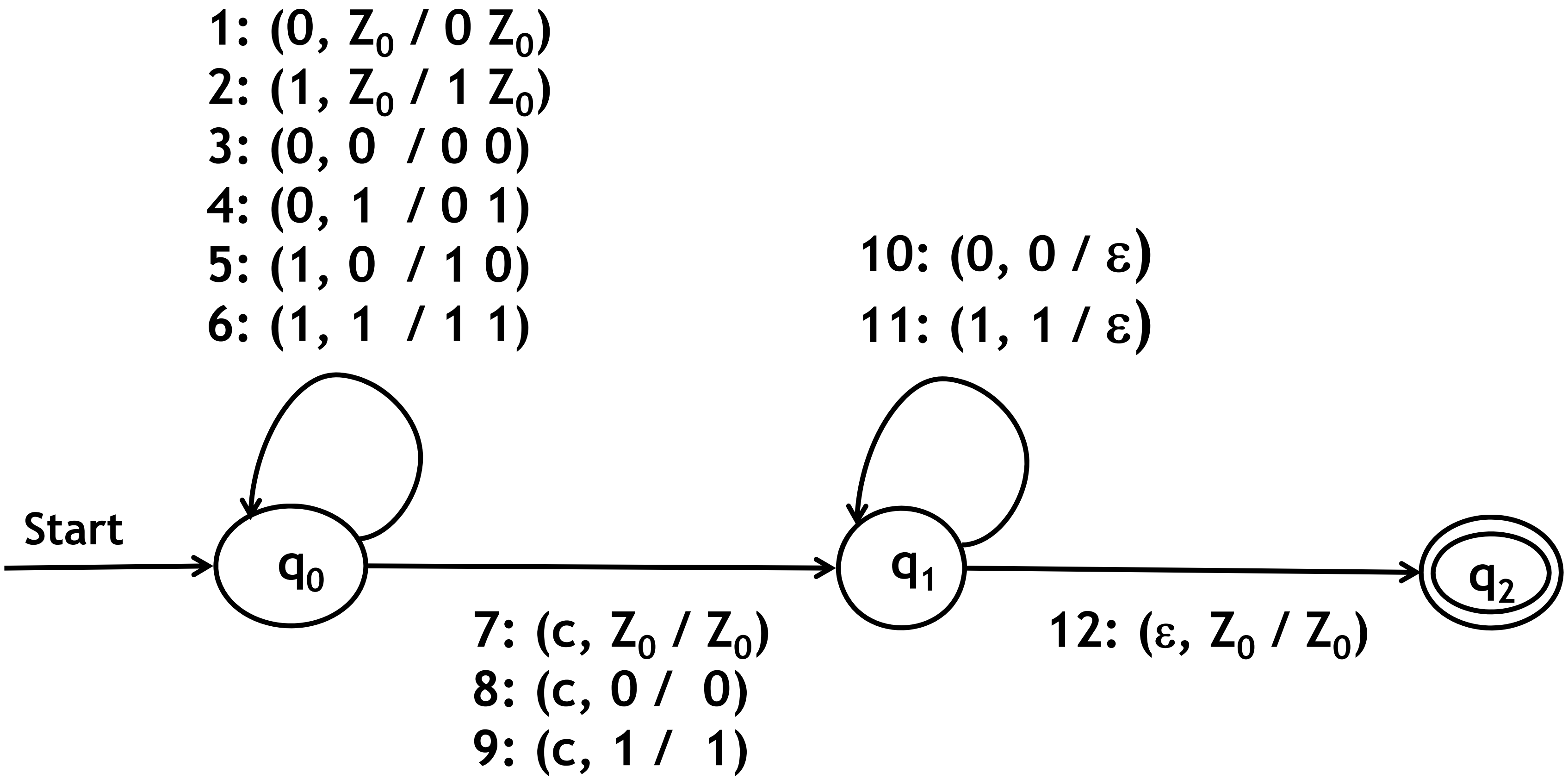
Im Zustand q_1 wird die Spiegelung überprüft, d.h. Eingabesymbole und Kellersymbole müssen übereinstimmen:

10. $\delta(q_1, 0, 0) = \{(q_1, \varepsilon)\}$ (es werden die obersten Kellersymbole gelöscht)
11. $\delta(q_1, 1, 1) = \{(q_1, \varepsilon)\}$

Bei Startsymbol des Kellers geht der Automat in den Endzustand über:

12. $\delta(q_1, \varepsilon, Z_0) = \{(q_2, Z_0)\}$

Beispiel 2 (Fortsetzung): Übergangsdiagramm



Beispiel 2 (Fortsetzung): Analyse

Analyse des Eingabewortes **01c100**:

$(q_0, 01c100, Z_0) \vdash (q_0, 1c100, 0Z_0) \vdash (q_0, c100, 10Z_0) \vdash$

$(q_1, 100, 10Z_0) \vdash (q_1, 00, 0Z_0) \vdash (q_1, 0, Z_0) \vdash$

$(q_2, 0, Z_0)$

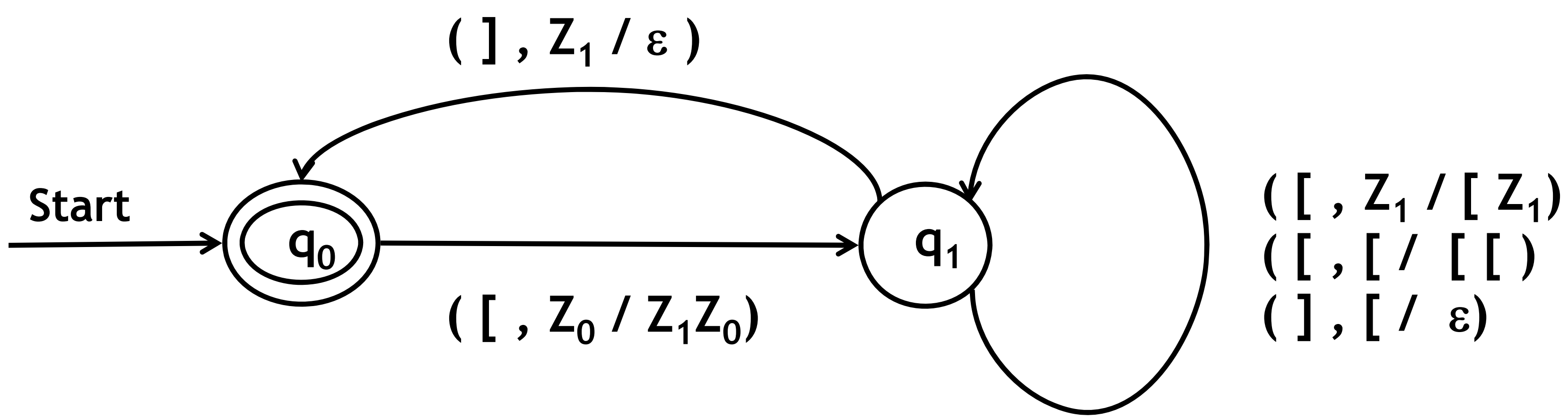
Der Kellerautomat akzeptiert das Eingabewort **01c100** nicht, da das Eingabewort nicht zur Gänze abgearbeitet wurde.

Beispiel 3:

Gesucht ist ein deterministischer Kellerautomat, der wohlgeformte Klammer-ausdrücke <u>durch Endzustand akzeptiert</u>, die durch die KFG G = ({S}, {[,]}, { S $\rightarrow$ [S]S, S$\rightarrow\varepsilon$ }, S) definiert ist.

KA = (Q, Σ, Γ, δ, q_0, Z_0, F)
Σ = { [,] }, Q = {q_0, q_1}, Γ = { [, Z_1, Z_0}, F = {q_0}

Überführungsfunktion δ:



Abschließende Anmerkungen

Theorem: Sei KA=$(Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ ein Kellerautomat, dann existiert eine kontextfreie Grammatik G mit L(G)=L(KA).

Theorem: Sei G=(N,Σ,P,S) eine kontextfreie Grammatik, dann existiert ein Kellerautomat KA mit L(KA)=L(G).

Man beachte: $L_{det\text{-}KA} \subset L_{nichtdet\text{-}KA}$

Nichtdeterministische Kellerautomaten sind mächtiger als deterministische Kellerautomaten:

- nichtdet. KA akzeptieren alle kontextfreie Sprachen
- det. KA akzeptieren alle regulären Sprachen, aber bloß eine Teilmenge der kontextfreien Sprachen
- det. KA sind für Parser von Compilern wichtig, d.h. die Syntax von Programmiersprachen wird i.a. so gewählt, dass sie von einem det. KA analysiert werden kann