



Formale Sprachen: E-Mail-Adresse



Gültige E-Mail-Adressen

- jenni@zirb.nich.de
 - jenni.zirbnich@simsoftlab.net
 - jz@ssl.net
 - f112358@4leonardo.ch
-
- Beschreibe den strukturellen Aufbau einer E-Mail-Adresse

Erste Strukturregel

- **E-Mail-Adresse = Benutzerkennung "@" Domain "." TLD**
- Die E-Mail-Adresse besteht aus einer Benutzerkennung gefolgt von einem @, anschließend wird die Domain angegeben. Nach der Domain kommt ein . und zum Schluss die Top-Level-Domain.
- Bestimmte Bestandteile der E-Mail-Adresse sind bereits vollständig definiert; andere bedürfen noch weiterer Erklärung. Benenne diese entsprechenden Bestandteile.

Nichtterminale und Terminale

- Symbole, die nicht weiter ersetzt werden müssen nennen wir Terminalsymbole oder kurz **Terminale**.
- Symbole, die noch weiterer Erklärung bedürfen und erst mittels weiterer Erklärungen ersetzt werden müssen (Platzhalter), nennen wir Nichtterminalsymbole oder **Nichtterminale**.
- Das erste Nichtterminal, bei dem mit der Erzeugung einer gültigen Zeichenkette begonnen wird, hier die E-Mail-Adresse, nennt man **Startsymbol**

Zweite Strukturregel

- Falls sich ein Nichtterminal in verschiedene mögliche Bestandteile weiter aufteilen lässt, kann bei den Regeln durch ein „**oder**“ (mit dem Zeichen „**|**“) diese **Fallunterscheidung** darstellen.
- Dies ist ein sogenanntes Metazeichen. Diese gehören nicht zur erzeugten Sprache, sondern dienen der Sprachdefinition.
- Wie könnte man nun die Benutzerkennung genauer beschreiben?

• Benutzerkennung = Einzelzeichen | Einzelzeichen Einzelzeichen
| Einzelzeichen Einzelzeichen Einzelzeichen | ...

Zweite Strukturregel (Alternative)

- Für eine Wiederholung gibt es das folgendes Metazeichen { ... }
 - Das zu wiederholende Nichtterminal oder Terminal wird zwischen die Klammern geschrieben.
 - Dabei darf der Inhalt der Klammer **keinmal**, **einmal** oder **beliebig oft** vorkommen/wiederholt werden.
- Benutzerkennung = Einzelzeichen { Einzelzeichen }

Weitere Regeln zur Benutzerkennung

- Für Benutzerkennung sind Kleinbuchstaben, Zahlen und als Sonderzeichen Punkt, Ausrufezeichen, Binde- und Unterstrich erlaubt.

- Einzelzeichen = Buchstabe | Ziffer | "." | "!" | "-" | "_"
- Buchstabe = "a" | "b" | "c" | ... | "z"
- Ziffer = "0" | "1" | "2" | ... | "9"

- Das Aufzählen aller 26 Buchstaben des Alphabets wäre sehr umfangreich, weshalb das Metazeichen ... verwendet wird.

Weitere Metazeichen

- Option **[]**
 - ["0"] "1"
 - -> Ergebnisse: „01“ oder „1“
-
- Gruppierung **()**
 - ("M" | "H") ("aus" | "und")
 - -> Ergebnisse: „Hund“, „Haus“, „Mund“, Maus“

Restlichen Regeln

- Setze die restlichen Regeln gemäß der Vorgaben um:
 - Die (Internet-)Domain setzt sich zusammen aus mindestens einer Unterdomain.
 - Jede Unterdomain darf Buchstaben, Zahlen und den Bindestrich enthalten. Die minimale Zeichenkettenlänge ist zwei; an erster und letzter Stelle ist kein Bindestrich erlaubt.
 - Die Top-Level-Domain darf nur aus Buchstaben zusammengesetzt sein: mindestens zwei, höchstens vier.

E-Mail-Adresse

- E-Mail-Adresse = Benutzerkennung "@" Domain "." TLD
- Benutzerkennung = Einzelzeichen { Einzelzeichen }
- Einzelzeichen = Buchstabe | Ziffer | "." | "!" | "-" | "_"
- Buchstabe = "a" | "b" | "c" | ... | "z"
- Ziffer = "0" | "1" | "2" | ... | "9"
- Domain = Unterdomain { "." Unterdomain }
- Unterdomain = (Buchstabe | Ziffer) { Buchstabe | Ziffer | "-" }
(Buchstabe | Ziffer)
- TLD = Buchstabe Buchstabe [Buchstabe] [Buchstabe]

Erweiterte Backus-Naur-Form

- Syntaxregeln können mithilfe der **erweiterten Backus-Naur-Form (EBNF)** formuliert werden.
- In der EBNF sind folgende **Metazeichen** zugelassen:
 - | für eine **Alternative**
 - [] für eine **Option**
 - { } für eine **Wiederholung**
 - () für eine **Gruppierung**
 - (... für eine **Aufzählung**)

Aufgabe

- Für Datumsangaben gibt es unterschiedliche Formate. Eine Möglichkeit wäre z. B. „6. Dezember 2010“.
- Begründe, warum es sich bei der Menge aller Datumsangaben dieser Form um eine formale Sprache handelt.
- Lege mithilfe der Regeln der EBNF die Syntax der Datumsangaben fest.
- Ergänze deine Regeln aus so, dass optional zum Datum noch eine fünfstellige Zeitangabe (Stunden und Minuten durch einen Doppelpunkt getrennt) hinzugefügt werden kann.

Autokennzeichen

- Aufbau eines Kennzeichens:
 - 1 bis 3 Buchstaben kennzeichnen die Stadt oder Landkreis
 - Trennzeichen: wir benutzen das Zeichen -
 - 1 bis 2 Buchstaben; anschließend Trennzeichen
 - Ein- bis vierstellige Zahl ohne führende Nullen
 - Zum Schluss gegebenenfalls ein E, falls es sich um ein Elektroauto oder ein H, falls es sich um ein historisches Auto handelt.
- Definiere mittels EBNF die formale Sprache der Autokennzeichen.

