

Mögliche Zeichen für die Verständigung mit dem Backofen sind also die Ziffern 0, 1, 2, 3, 4, 5, 6, 7, 8 und 9 sowie die Symbole \sim , \sqsubset , \varnothing , Υ , \vdash und \rightarrow . Durch die Zeichenfolge ' $\vdash 115 \rightarrow 1230 \varnothing 160$ ' wird der gewünschte Ablauf zum Garen des Kartoffelgratins beschrieben.

Alphabet

Die endliche Menge einzelner **Zeichen**, die in einer formalen Sprache verwendet werden, heißt **Alphabet** oder **Zeichenvorrat** dieser Sprache und wird mit Σ bezeichnet. Im Beispiel ist $\Sigma = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, \sim, \sqsubset, \varnothing, \Upsilon, \vdash, \rightarrow\}$.

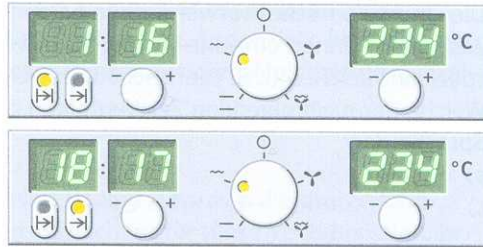
Durch Aneinanderreihen mehrerer Zeichen des Alphabetes wird ein **Wort** (auch **Zeichenkette** genannt) gebildet. Ein korrektes Wort für die Zubereitung des Kartoffelgratins lautet ' $\vdash 115 \rightarrow 1230 \varnothing 160$ '.

Formale Sprachen

Nicht alle möglichen Kombinationen von Zeichen eines Alphabetes bilden korrekte Wörter einer bestimmten Sprache. Durch die Angabe **formaler Regeln** wird die Menge aller zulässigen Zeichenketten als **formale Sprache** definiert.

Am Beispiel des oben beschriebenen Backofens könnten Regeln für die Bildung von Wörtern z.B. wie folgt lauten:

- Auf jedes der Zeichen \vdash bzw. \rightarrow müssen jeweils ein bis vier Ziffern aus 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 folgen.
- Genau eines der Zeichen \sim , \sqsubset , \varnothing , Υ gefolgt von einer Kombination aus einer der Ziffern 0, 1, 2 mit zwei Ziffern aus 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 muss enthalten sein.



Syntax und Semantik

Die Gesamtheit der Regeln, in denen festgelegt wird, welche Zeichenketten zu einer Sprache gehören, bezeichnet man als **Syntax** dieser Sprache. In der oben beschriebenen Sprache ist z.B. die Zeichenkette ' $\vdash 999999 \Upsilon$ ' syntaktisch falsch, da die Regeln der Sprache nicht eingehalten werden. Dagegen erhält man mit ' $\sqsubset 215 \vdash 35 \rightarrow 817$ ' ein nach obigen Regeln syntaktisch korrekt gebildetes Wort.

Obwohl die Syntax des Wortes ' $\vdash 876 \rightarrow 9897 \sim 234$ ' korrekt ist, das Wort also zur oben definierten Sprache gehört, ist eine sinnvolle Interpretation hier nicht möglich, da 876 und 9897 nicht als Uhrzeit verstanden werden können. Allgemein gibt die **Semantik** einer Zeichenkette die Bedeutung dieses Wortes wieder. Nach Austausch einiger Ziffern erhält man mit ' $\vdash 116 \rightarrow 1817 \sim 234$ ' ein Wort, dessen Semantik umgangssprachlich mit „1h und 16min lang wird die Grilltemperatur auf 234°C gestellt, sodass um 18.17 Uhr der Backvorgang beendet ist“ beschrieben werden kann.

Für die Definition einer Sprache aus allen Wörtern, die für den Backofen korrekt sind, müssen demnach die beiden obigen Regeln um weitere Regeln ergänzt werden.

Bei der Kommunikation mit natürlichen Sprachen können Zweideutigkeiten und Missverständnisse auftreten. Dies kann man durch die Verwendung einer **formalen Sprache** verhindern, die als Menge zulässiger Zeichenketten über einem Alphabet definiert wird. Die Regeln zur Bildung zulässiger Zeichenketten stellen die **Syntax** einer Sprache dar. Die Bedeutung eines Wortes nennt man auch **Semantik**.

