

Jedermann kennt natürliche Sprachen als Mittel zur Kommunikation zwischen Menschen. Eindeutiges gegenseitiges Verstehen ist nur dann gewährleistet, wenn sich die Kommunikationspartner auf eine gemeinsame Sprache einigen und die zu deren Gebrauch festgelegten Regeln einhalten.

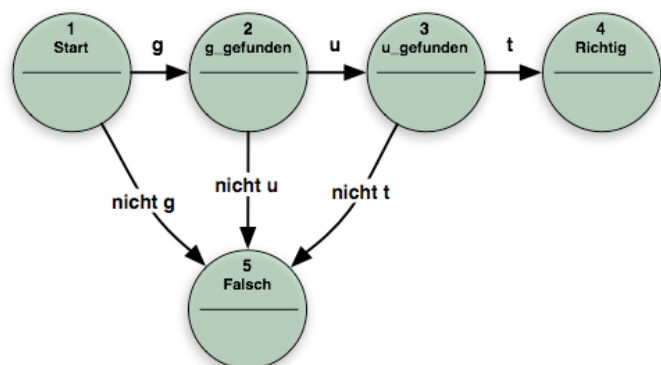
Die Verständigung zwischen Mensch und Maschine stellt ebenfalls einen Kommunikationsprozess dar, der ähnlichen Regeln unterliegt.



In der theoretischen Informatik ist eine **formale Sprache** eine abstrakte Sprache, bei der im Unterschied zu konkreten Sprachen oft nicht die Kommunikation im Vordergrund steht, sondern die mathematische Verwendung. Eine formale Sprache besteht aus einer bestimmten Menge von Symbolketten („Worte“ der Sprache), die aus einem Symbolvorrat („Alphabet“) zusammengesetzt werden können. Wir werden uns verschiedene Möglichkeiten anschauen, formale Sprache zu definieren, also eine endliche (und verständliche) Beschreibung dieser Sprachen zu finden. So ist zum Beispiel schon die Sprache aller Wörter, die nur aus dem Buchstaben „a“ bestehen, unendlich groß („a“, „aa“, „aaa“, ...). Wir untersuchen den strukturellen Aufbau formaler Sprachen und lernen die Begriffe Syntax und Semantik einer Sprache zu unterscheiden.

Formale Sprachen spielen eine wichtige Rolle in der Informatik, da sie sich im Gegensatz zur natürlichen Sprache gut zur Verarbeitung am Computer eignen. So stellt z.B. jede Programmiersprache eine formale Sprache dar, die genau festlegt, wann Quelltext ein gültiges Programm darstellt und wann nicht (Syntax), und was dies dann für die Ausführung des Programmes bedeutet (Semantik). Dies ist notwendig, um genaue Aussagen, z.B. über Korrektheit und Effizienz, treffen zu können. Die Methoden, die solches erst möglich machen, sind Gegenstand dieses Moduls.

An vielen Stellen z.B. in der Programmierung muss man eine Eingabe validieren oder in Texten bestimmte Muster finden. Häufig lässt sich dies mit sogenannten regulären Sprachen erledigen. Am einfachsten und verständlichsten lassen sich diese mit regulären Ausdrücken angeben und in praktisch allen Programmiersprachen gibt es Unterstützung dafür. Wir möchten den Weg vom Angeben eines regulären Ausdrucks zum eigentlichen Berechnen, ob der gegebene Text zur Sprache gehört, nachvollziehen. Dazu wählen wir folgenden Weg: Der reguläre Ausdruck wird in einen nichtdeterministischen Automaten umgewandelt. Dies erfolgt rekursiv anhand der Struktur des regulären Ausdrucks. Danach wird der Automat determinisiert, d.h. in einen deterministischen Automaten umgewandelt. Dieser erlaubt eine effiziente und einfache Überprüfung, ob ein Wort zu unserer Sprache gehört.



1 Endlicher Automat, der das Wort "gut" erkennt

Danach werfen wir noch einen kurzen Blick auf nicht reguläre, sondern kompliziertere Sprachen und betrachten dort das Konzept einer (kontextfreien) Grammatik. Auch diese sind in der Praxis relevant, da

sich der Syntax vieler Programmiersprachen als kontextfreie Grammatik formulieren lässt.

Der Abschluss bildet die Erkenntnis, dass nicht jede formale Sprache berechenbar ist: es gibt formale Sprachen, für die es keinen Algorithmus geben kann, der entscheidet ob ein Wort zu der Sprache gehört oder nicht.

Voraussetzungen:

Interesse an Mathematik und deren Anwendung in der Informatik.

Es wird eine besonders hohe Bereitschaft erwartet, sich mit den Themen selbstständig auseinander zu setzen. Zwischen den Treffen müssen die besprochenen Inhalte intensiv nachbereitet werden. Eine regelmäßige Teilnahme an angegebenen Terminen ist notwendig.

Die gemeinsame Abschlusspräsentation durch Vorträge und Poster findet am Samstag, 04. Juli 2020 am International Department in Karlsruhe statt.

Teilnehmerzahl: höchstens 20

Ort: Bunsengymnasium,
Humboldtstr. 23, Heidelberg

Betreuer: Moritz Lichter, TU Kaiserslautern
Gisela Döbbeling, Kursleiterin Heidelberg
Dr. Oliver Rudolph, Kursleiter Heidelberg

Email: doebbeling@hector-seminar.de,
rudolph@hector-seminar.de

Geplante Termine:

- **Do, 23.04.2020, 16⁰⁰-18¹⁵ Uhr**
- **Mi., 29.04.2020, 16⁰⁰-18¹⁵ Uhr**
- **Do, 07.05.2020, 16⁰⁰-18¹⁵ Uhr**
- **Mi., 20.05.2020, 16⁰⁰-18¹⁵ Uhr**
- **Do, 28.05.2020, 16⁰⁰-18¹⁵ Uhr**
- **Do, 18.06.2020, 16⁰⁰-18¹⁵ Uhr**
- **Do, 25.06.2020, 16⁰⁰-18¹⁵ Uhr**
- **Mi., 01.07.2020, 16⁰⁰-18¹⁵ Uhr**
- **Do, 02.07.2020, 16⁰⁰-18¹⁵ Uhr**
- **Sa, 04.07.2020, Modulfest
International Department KIT**
- **Do, 09.07.2020, 15⁰⁰-17⁰⁰ Uhr
Nachtreffen**