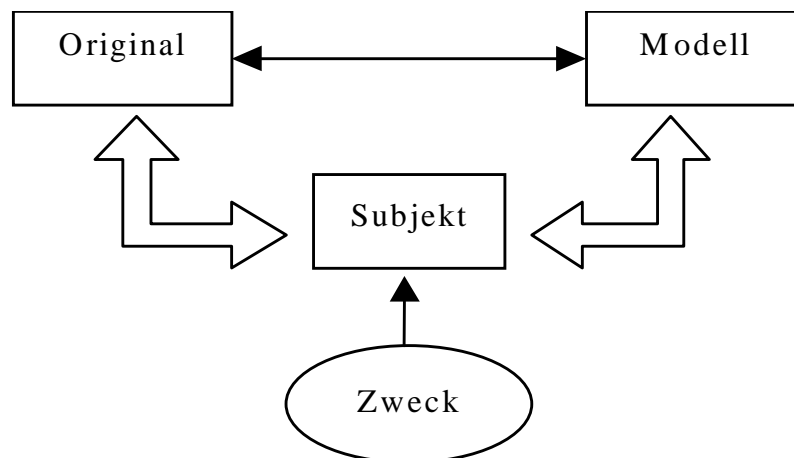


## Prozeß der Modellbildung



## Einige Modelle

### in der Chemie

- Atomvorstellung
- Modellsubstanzen, -systeme
- Experimente mit Modellcharakter
- räumliche Modelle von Molekülen
- funktionale Repräsentation technischer Prozesse
- "Berg" der Aktivierungsenergie
- Diagramme
- chemische Symbole
- Formelsprache
- Reaktionsgleichungen
- Formalismus des chemischen Gleichgewichts
- Kinetikmodelle der Physikalischen Chemie
- chemische Strukturformeln

### in der Informatik

- Automatenmodell
- Registermaschine
- Turingmaschine
- Roboter
- Pipe, Baum, Liste, Stack
- Vererbung, Vater, Söhne,
- Diagramme
- Programmiersprachen
- Piktogramme
- Bit, Information
- mathematische Logik
- interaktive, ereignisgesteuerte künstliche "Weltmodelle"
- formale Sprachen und Übersetzer

## Rechnergestützte Deutung der Nomenklatur chemischer Stoffe

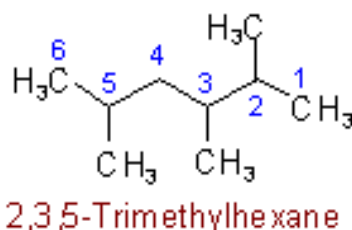
### *Fachübergreifendes Projekt zur Modellbildung im Informatikunterricht*

In der frühen Chemie und Alchemie erfolgte die Benennung chemischer Stoffe zumeist anhand phänomenologischer Merkmale (Bsp. Buttersäure), persönlicher Daten der Entdecker (Bsp. Germanium) oder anderen willkürlichen Kriterien. Diese sogenannte Trivialnamen erschwerten zunehmend die Kommunikation der Chemiker, da der Trivialname nur unzureichend Informationen zu den chemischen Eigenschaften des Stoffes lieferte.

Eine systematische und allgemeinverbindliche Namensgebung (Nomenklatur) - basierend auf dem mächtigen Prinzip der chemischen Formel - wurde Ende des 18. Jahrhunderts entworfen, auf internationaler Ebene organisiert und seit 1993 von der International Union of Pure and Applied Chemistry (IUPAC) weitergeführt.

Das Regelwerk der IUPAC beschreibt die Übersetzung von chemischen Strukturformeln in eine Sprechsprache und umgekehrt.

Bsp.



Bisher ist eine umfassende Systematisierung der Nomenklatur nur in der organischen Chemie gelungen. Die Empfehlungen der IUPAC lassen zudem zu einer chemischen Strukturformel einige verschiedene Namen zu. Umgekehrt kann der Chemiker jedoch aus jedem systematischem Stoffnamen unter Anwendung des Regelwerkes die exakte Strukturformel eindeutig ableiten und damit Informationen über die Eigenschaften des Stoffes gewinnen.

Der Prozess der Übersetzung des Fachausdrucks in die Strukturformel ist für den Chemiker eher eine Fleißaufgabe und birgt i. d. R. kaum neue Erkenntnisse.

Problemstellung:

???

### **Materialien:**

Arbeitsanleitung

IUPAC-Nomenklatur (englisch) <http://www.acdlabs.com/iupac/nomenclature/>

Literatur zu den informatischen Grundlagen

Hinweise für den Lehrenden (inkl. Einordnung des Themas in den Unterricht)

**Problemstellung:**

Basierend auf der IUPAC-Nomenklatur soll ein (leicht erweiterbares) Programm konstruiert werden, das in der Schule vorkommende Stoffnamen der organischen Chemie in entsprechende Strukturformeln umsetzt, die ggf. auch grafisch dargestellt werden können.

**1) Auseinandersetzung mit der IUPAC - Nomenklatur (Chemie ab Jgst. 11)**

Homologe Reihen:

Alkane:

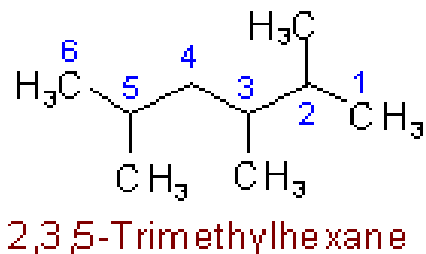
Methan	$\text{H}_3\text{C}-\text{H}$
Ethan	$\text{H}_3\text{C}-\text{CH}_3-\text{H}$
Propan	$\text{H}_3\text{C}-\text{CH}_2-\text{CH}_3-\text{H}$
...	...

Alkanole: ...  $-\text{O}-\text{H}$

Halogenderivate: ...  $-\text{Cl}$

Alkene, Alkine, Aldehyde, Ketone, Carbonsäuren, Ester, ...

Bsp.



Vereinfachung: 2, 3, 5 - Trimethylhexane → CC[C]CC[C]C[C]C

**Anmerkung:**

**Im Chemieunterricht wird die IUPAC-Nomenklatur nicht gesondert eingeführt, sondern im Anwendungszusammenhang erschlossen.**

⇒ mögliche Lernziele aus Sicht des Chemieunterrichts:

- LZ-CU 1: Schulung des systematischen Umgangs mit der Nomenklatur
- LZ-CU 2: Verdeutlichung der Mächtigkeit der chem. Formelsprache
- LZ-CU 3: Herausstellen von Beziehungen zwischen verschiedenen Modellen

IUPAC Nomenclature - Netscape

File Edit View Go Communicator Help

Bookmarks Location: <http://www.acdlabs.com/iupac/nomenclature/>





**advanced chemistry development** **main menu**

## IUPAC Nomenclature of Organic Chemistry












[SEARCH](#)

---

### Recommendations 1979

-  [A. Hydrocarbons](#)
-  [B. Fundamental Heterocyclic Systems](#)
-  [C. Characteristic Groups Containing Carbon, Hydrogen, Oxygen, Nitrogen, Halogen, Sulfur, Selenium, and/or Tellurium](#)
-  [List of Radical Names](#)

### Recommendations 1993

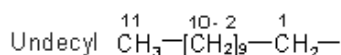
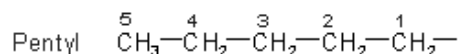
-  [Preamble](#)
-  [R-0 Introduction](#)
-  [R-1 General Principles of Organic Nomenclature](#)
-  [R-2 Parent Hydrides and their Derived Substituent Groups](#)
-  [R-3 Characteristic \(Functional\) Groups](#)
-  [R-4 Guide to Name Construction](#)
-  [R-5 Applications to Specific Classes of Compounds](#)
-  [R-6 Name Interpretation](#)
-  [R-7 Stereochemical Specification](#)
-  [R-8 Isotopically Modified Compounds](#)
-  [R-9 Appendix](#)

This HTML reproduction of Sections A, B and C of IUPAC "Blue Book" is as close as possible to the published version. See Document Done

17 Heptadecane 80 Octacontane  
 18 Octadecane 90 Nonacontane  
 19 Nonadecane 100 Hectane  
 20 Icosane 132 Dotriacontahectane  
 21 Henicossane

**1.2** - Univalent radicals derived from saturated unbranched acyclic hydrocarbons by removal of hydrogen from a terminal carbon atom are named by replacing the ending "-ane" of the name of the hydrocarbon by "-yl". The carbon atom with the free valence is numbered as 1. As a class, these radicals are called normal, or unbranched chain, alkyls.

*Examples to Rule A-1.2*



*See Recommendations'93 R-2.2.1*

**Next:**

[Saturated Branched-chain Compounds and Univalent Radicals](#)  
[Unsaturated Compounds and Univalent Radicals](#)  
[Bivalent and Multivalent Radicals](#)

## 2) Transfer der Formelsprache in ein Anschauungsmodell (Übersetzertechnik)

- Überprüfung der Eingabe einer Formel auf ihre Gültigkeit (Analyse)
- Übersetzung der Eingabezeichenfolge in eine Zielsprache (Codeerzeugung)
- Darstellung der in der Zielsprache beschriebenen Struktur (Visualisierung)

### Bsp. 2, 3, 5 - Trimethylhexane

#### Quellsprache

Alphabet = {1,2,...,9,-,a,b,...,z, }

kontextfreie Grammatik für verzweigte, nicht-zyklische Alkane

	:=	IUPAC-Regel	im Bsp.
<Formula>	<Prefix><MainChain><Infix>	R-6.0	
<Infix>	ε		
<MainChain>	<SUAC>ane	A-1.1	hexane
<SUAC>	meth   eth   prop   but   pent   hex	R-9.1: Table 19(a) , A-1.	
<SUAC-Radical>	<SUAC>yl	A-1.2	methyl
<Prefix>	<Pos> - <Substituent>   <Substituent>   ε	A-2.1, 2.2, 2.25	2,3,5-trimetyl
<Substituent>	<BNT><SUAC-Radical>	A-2.2 (s. u.)	trimethyl
<BNT>	mono   di   tri   tetra   penta   hexa	R-2.2.1 Table 11	tri
<Pos>	<Number>   <Number>, <Pos>		2,3,5
<Number>	1   2   3   4   5   6		

SUAC = saturated unbranched-chain compound = gesättigte unverzweigte Kette, BNT = basic numerical term

Grundsymbole (Token) der Quellsprache als regulärer Ausdruck:

1 | 2 | 3 | 4 | 5 | 6 | m(ono|eth) | di | t(ri|etra) | p(rop|ent(a|ε)) | hex(a|ε) | eth | but | ane | -

#### Zielsprache:

	:=	Semantik
<Code>	( C ([<Code>] <sup>3</sup> ) <sup>*</sup> )	

(<Code><sup>3</sup>)<sup>\*</sup> ist eine Vereinfachung für <Code>|<Code><Code>|<Code><Code><Code>|ε

Bsp. → **CC[C]CC[C]C[C]C**

oder in VRML (Virtual Reality Modelling Language)

```

.....
Carbon {
  Carbon {
    Carbon { }
  }
  Carbon {
    Carbon { }
  }
  Carbon {
    Carbon {
      Carbon { }
    }
  }
  Carbon {
    Carbon {
      Carbon {
        Carbon { }
      }
    }
  }
}
.....

```

# 1. Ziel: (lexikalische Analyse)

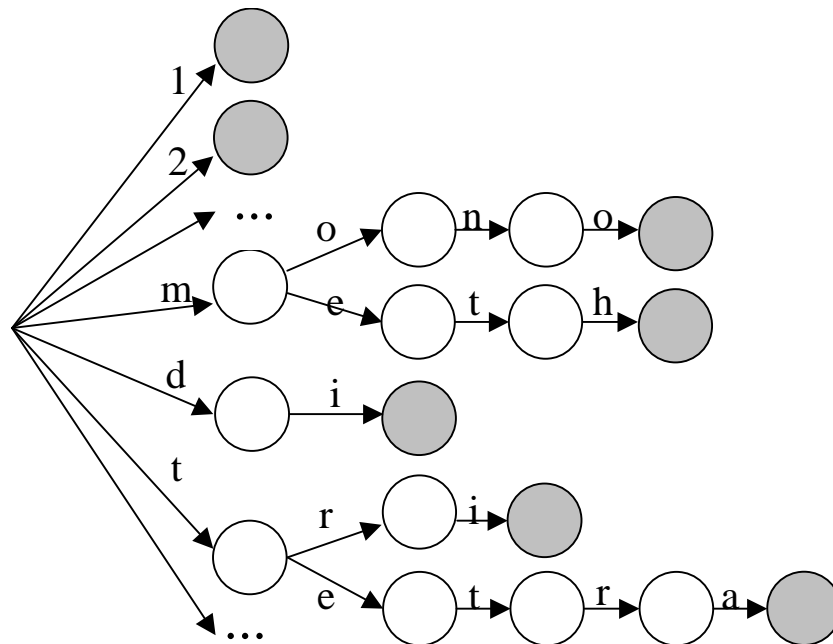
Übersetzung einer Folge von Zeichen aus der Quellsprache in eine Folge von Token der Quellsprache, um bedeutungsmäßig zusammengehörige Zeichen zusammenzufassen.

Bsp. 2, 3, 5 - Trimethylhexane → 2 , 3 , 5 : tri meth yl hex ane

Lösungsansatz:

Ein Programm (Scanner) liest das Quellprogramm zeichenweise ein und gibt die dazugehörige Tokenfolge aus. Realisiert wird die lexikalische Analyse mit einem endlichen Automaten. Ein Token wird erkannt, wenn der Automat vom Startzustand in den Endzustand gelangt.

Bsp.

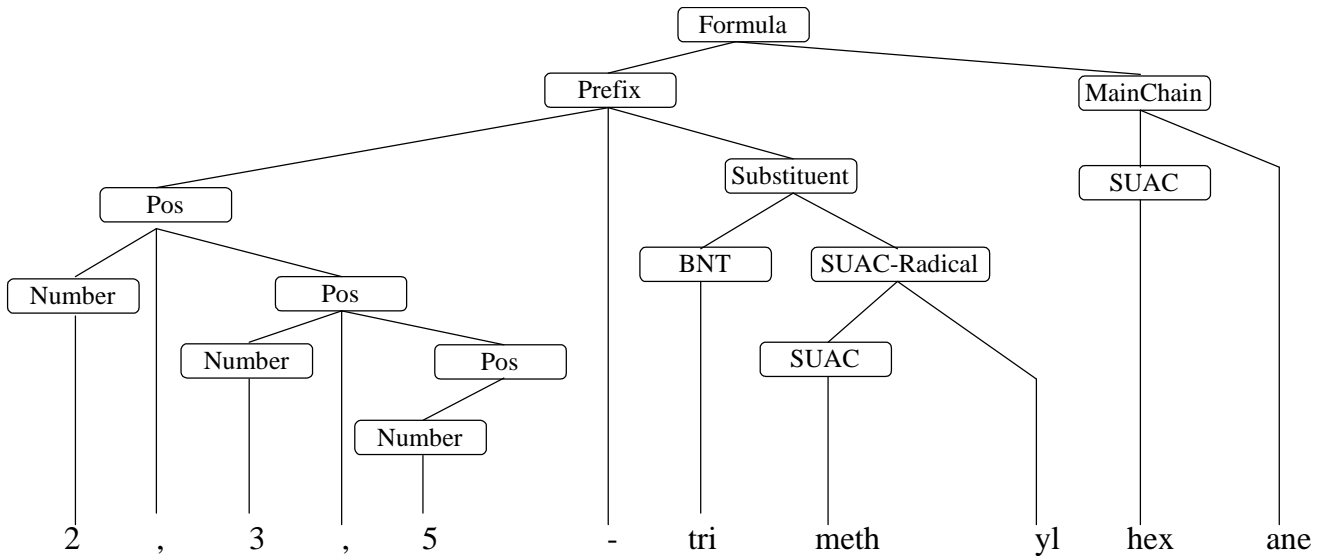


Scanner können sehr umfangreich werden. Die Informatik kennt jedoch Verfahren, um nicht-deterministische und deterministische endliche Automaten derart ineinander umzuwandeln, dass sie minimiert werden [s. Hopcroft, 13ff].

## 2. Ziel (syntaktische Analyse)

Überprüfung einer Folge von Token auf ihre syntaktische Korrektheit gemäß der aufgestellten Grammatik der Quellsprache und Überführung in einen Ableitungsbaum.

Bsp. 2 , 3 , 5 - tri meth yl hex ane →



Lösungsansatz:

Ausgehend von der ersten Grammatikregel  $\text{Formula} := \dots$  versucht der sogenannte Parser den Ableitungsbaum rekursiv aufzubauen. Für jedes Token enthält der Parser einen Programmabschnitt, in dem abhängig vom nächstfolgenden Token eine Ableitung gemäß der Grammatik erfolgt.

Bsp.  $\langle \text{SUAC-Radical} \rangle := \langle \text{SUAC} \rangle \text{yl} \rightarrow$

```

procedure SUAC_Radical;
begin
    SUAC;
    if top(tokenFolge) = "yl" then
        yl_erkannt und pop(tokenFolge);
    else
        syntaxfehler_in_Formel;
    end
end
    
```

Innerhalb der Prozeduren werden zusätzlich Anweisungen eingebaut, die den Baum aufbauen.

Es können jedoch diverse Probleme auftreten:

Bsp.  $\langle \text{Prefix} \rangle := \langle \text{Pos} \rangle - \langle \text{Substituent} \rangle$  und  $\langle \text{Prefix} \rangle := \langle \text{Substituent} \rangle$

Bsp.  $\langle \text{Prefix} \rangle := \langle \text{Substituent} \rangle$  und  $\langle \text{Prefix} \rangle := \epsilon$

### 3. Ziel (semantische Analyse)

- 1) Überprüfung des Ableitungsbaums auf Fehler, die nicht in der Grammatik beschrieben wurden, aber im Kontext als Zusatzbedingungen "vereinbart" worden sind.

Bsp. 2,3,5 - tri ...

Findet sich im Baum ein BNT-Knoten tri, muss überprüft werden, ob im linken Ast des Vor-Vorgängers entsprechende Positionen verzeichnet sind (kontextsensitiver Fehler, *Attributierung*)

- 2) Vorbereitung der Code-Erzeugung

Der Baum wird umstrukturiert, so dass den einzelnen Positionen die entsprechenden Substituenten zugeordnet werden, um die Code-Erzeugung zu erleichtern.

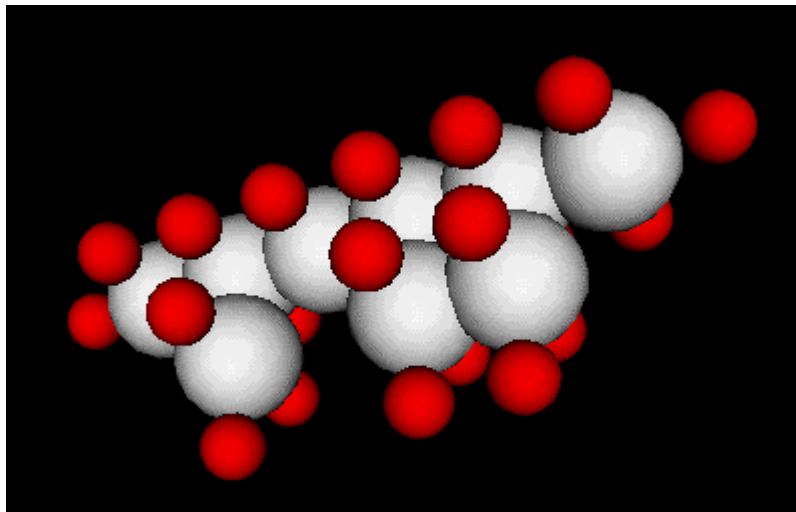
### 4. Ziel (Code-Erzeugung für die Zielsprache)

Ausgehend von einem vorstrukturiertem Ableitungsbaum wird der Code der Zielsprache systematisch erzeugt.

Bsp.

→ **CC[C]CC[C]C[C]C** o. ä.

### 5. Visualisierung des Codes (z. B. mit VRML)



Zu allen obigen Punkten existieren unterschiedliche Lösungsvarianten, die sich insbesondere in ihrer Komplexität unterscheiden. Die Auswahl des Lösungsansatzes bleibt den jeweiligen Gruppen vorbehalten. Die Vereinbarung der Schnittstellen und die Aufstellung der Spezifikation muss daher sorgfältig und präzise erfolgen.



**Literatur:**

- [Becker] Becker, H.-J.: "Fachdidaktik Chemie". Aulis Verlag Deubner, Köln 1992
- [Hopcroft] Hopcroft, J.E.: "Einführung in die Automatentheorie, formale Sprachen und Komplexitätstheorie". 2. Aufl., Addison-Wesley, Bonn 1990
- [InfoDuden] Schülerduden Informatik
- [IUPAC] IUPAC Nomenclature of Organic Chemistry (Internet, CD-ROM)  
<http://www.acdlabs.com/iupac/nomenclature/>
- [Kastens] Kasten, U.: "Übersetzerbau". Handbuch der Informatik Bd. 3.3. Oldenbourg Verlag
- [Pfeifer] Pfeifer, P.: "Konkrete Fachdidaktik Chemie". Oldenbourg, München 1992
- [UNESCO] <http://www.ge-dip.etat-ge.ch/cptic/unesco/en/app3ga3.html>
- [Vollmer] Vollmer, G.: "Sprache und Begriffsbildung im Chemieunterricht". Diesterweg, Frankfurt a. M., 1980