

## Vereinfachte Darstellung von Begriffen und Relationen am Bildschirm

Die voll automatisierte Generierung von Diagnosecodes benötigt ein Expertensystem, das Freitexte (Nominalphrasen) interpretieren kann. Zu diesem Zweck müssen die in den Worten enthaltenen Sinninhalte analysiert und dargestellt werden. Anschliessend generiert ein Kalkulus die Codes.

Gegenüber anderen Expertensystemen zeichnet sich ein solches System dadurch aus, dass das Hauptgewicht auf den *Datenstrukturen* liegt, und nicht etwa auf dem Kalkulus; denn die Durchführung der Codezuweisung ist einfach - schwieriger hingegen ist es, die komplexe Information, die in den Worten der natürlichen Sprache steckt, in ein systematisches Datenmodell zu integrieren. Erste Versuche nahmen tatsächlich an, dass je ein Begriffsinhalt mit je einem Wort verknüpft sei, eine solch naive Sicht ist heute mit Sicherheit überholt. Die Notation der Begriffe und ihrer Relationen ist die eigentliche Aufgabe.

Bekannt sind die Notationsweisen der Prädikatenlogik, die Conceptual Graphs nach J. F. Sowa [2,3], die GRail des GALEN-Projektes [1] oder - in jüngerer Zeit - der Vorschlag des W3-Consortiums (Tim Berners-Lee) zur Notation eines RDFs (Resource Description Frameworks) [7]. Wir haben für die Aufgabe der Codezuweisung 1989 eine Notation in Begriffspartikeln [4] entwickelt. Die dabei gewonnenen Erfahrungen führten uns 1996 zur Darstellung von *Begriffsmolekülen*, mit denen sowohl komplexe Datenstrukturen wie vielfach verzweigte Regeln auf einfache Weise notiert werden können [5]. In der vorliegenden Arbeit erklären wir die Grundlagen dieser Notation und vergleichen sie mit einer anderen modernen Begriffsrepräsentation, den Conceptual Graphs.

von H.R.Straub, H.Mosimann und N.Frei

### 1. Der zweifache Anspruch an die Begriffsnotation

Bei der Entwicklung unseres Textinterpreters stellten wir folgende Anforderungen an die Begriffsnotation:

1.	Die Notation muss alle nötigen Strukturen nachbilden können.
2.	Die Zahl der formalen Elemente soll dabei möglichst klein gehalten werden.
3.	Die Notation muss eindeutig sein (ein Inhalt immer nur auf eine Weise darstellbar).
4.	Die Begriffsrepräsentation muss erweiterbar sein ( <i>keine „closed world“</i> ).
5.	Sie muss am Bildschirm für den Menschen einfach - <i>intuitiv</i> - lesbar sein.
6.	Sie soll möglichst <i>dicht</i> sein, d.h. pro Bildschirmfläche möglichst viel Inhalt zeigen.

Tab. 1: Forderungen an eine moderne Begriffsrepräsentation

Diese Forderungen entspringen einem *doppelten Anspruch* an die Begriffsnotation, nämlich sowohl für die *Maschine* als auch für den *Menschen* einfach lesbar zu sein. Für die Maschine muss die Notation mathematisch klar umsetzbar sein. Für den Menschen kommt der zusätzliche Anspruch, dass die Regeln schnell und einfach lesbar sind; dies ist umso wichtiger, je grösser und komplexer die Regelbasis des Expertensystems ist. Ein mathematisch noch so perfektes System, das die Bedingungen 4 bis 6 nicht erfüllt, muss versagen, da es ab einer gewissen Grösse nicht

mehr wartbar ist. Deshalb sind die letzten drei Punkte so wichtig.

Der von uns entwickelte semantische Interpreter (SIP) zur Codezuweisung verfügt über eine Notation, welche die genannten Bedingungen erfüllt. Dies war möglich durch die Einführung von *Begriffsmolekülen*, welche die Zweidimensionalität des Bildschirms benützen um komplexe Begriffsstrukturen abzubilden.

### 2. Begriffsatome und Begriffsmoleküle

Begriffe sind die Vorstellungsinhalte, die wir mit den Worten verbinden.

Doch nicht die Begriffe allein spielen eine Rolle: Zur Formulierung von Wissen gehört entscheidend, wie die Begriffe miteinander verknüpft sind. Die *Relationen* - d.h. die Verbindungen zwischen den Begriffen - enthalten das eigentliche Wissen und müssen in jeder Begriffsrepräsentation explizit darstellbar sein.

In unserer Darstellung sehen wir die Begriffe als unteilbare Einheiten, d.h. als *Atome*. Alles, was über die Begriffe ausgesagt wird, wird in ihren Relationen zu anderen Begriffen ausgesagt. Auch die Relationen eines Begriffs zu seinen Unterbegriffen und zu seinen Teilen werden als Relationen *ausserhalb* des Begriffs dargestellt. Dies hat den Vorteil, dass wir den Begriff als *unveränderliche* Entität führen können, auch wenn wir die Wissensbasis sukzessive erweitern. Was verändert wird ist nicht der Begriff, sondern das Bündel an Relationen, das über den Begriff bekannt ist.

Wenn sich Begriffe verknüpfen, bilden sie Cluster. Diese Cluster stellen wir am Bildschirm auf eine strikt reglierte Weise dar. So wie die Atome der Chemie Bindungsstellen haben, über die sie sich mit anderen Atomen verbinden können, haben unsere Begriffsatome genau definierte Bindungsstellen, über die sie mit genau definierten anderen Atomen Bindungen eingehen können. Die entstehenden Begriffscluster nennen wir *Begriffsmoleküle*. Mit den Begriffsmolekülen können insbesondere die *Kontexte* der Regeln klar dargestellt werden.

### 3. Begriffsatome: Gleichzeitig Klasse und Ausprägung

Um mit Begriffen systematisch umzugehen, werden sie *klassifiziert*. Dies führt zu zwei Arten von Werten, nämlich einerseits zu den Klassen (denen die Begriffe zugeteilt werden) und andererseits den Ausprägungen, d.h.

#### Kontaktinformationen:

Hans Rudolf Straub  
Belsitostrasse 32  
CH - 8645 Jona  
Email: [straub@meditext.ch](mailto:straub@meditext.ch)

den Werten innerhalb der Klasse, welchen die eingegebenen Begriffe darstellen. Die beiden Formen – Klasse und Ausprägung – werden z.B. bei Datenbanken sehr klar unterschieden, indem die Klasse der Datenbankspalte entspricht und die Ausprägung dem Wert eines bestimmten Feldes in der Spalte. In einer Spalte „Diagnose“ sind „Angina pectoris“, „Bronchuskarzinom“ und „Lungen-Tb“ mögliche Werte. Klassen und Ausprägungen spielen aber nicht nur bei Datenbanken eine Rolle, sondern bei allen Arten von Kategorisierungen und somit auch bei allen höher entwickelten Wissensrepräsentationen. Bei den Conceptual Graphs (CG) nach Sowa kann ein Begriff z.B. aus einem Klasse-Ausprägungs-Paar bestehen<sup>1</sup>:

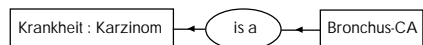


Abb. 1: Klasse und Ausprägung, mit Conceptual Graphs dargestellt

Der Begriff im linken Rechteck von Abb. 1 ist ein solches Klasse-Ausprägungs-Paar. Mit dem Doppelpunkt wird die Relation von Klasse (Typ) zu Ausprägung notiert.

Die Information von Abb. 1 kann mit Begriffsmolekülen sehr einfach dargestellt werden. Wir verketteten dabei Ausprägungen und Klassen über mehrere Stufen:

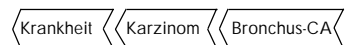


Abb. 2: Die Information von Abb. 1 als Kette von 3 Begriffsatomen

Der Begriff „Karzinom“ ist in Abb. 2 gleichzeitig Klasse und Ausprägung, nämlich Klasse für seine Unterbegriffe (hier für „Bronchus-CA“) und Ausprägung für seine Oberbegriffe (hier für „Krankheit“). Wir sprechen in diesem Zusammenhang von der *Bifazialität* des Begriffsatoms. Jeder Begriff kann – je nach Konstellation – Klasse oder Ausprägung sein, oder auch beides zusammen. Dies ist bei konventionellen Begriffsrepräsentationen nicht möglich, sondern Klasse (Typ) und Ausprägung sind bei ihnen klar getrennt.

In Kapitel 1 haben wir die Forderung aufgestellt, dass das semantische

Netz (die Begriffsrepräsentation) eine „offene Welt“ darstellen soll, und deshalb an beliebiger Stelle erweiterbar sein muss. Die Bifazialität der Begriffsatome kommt dem entgegen, indem die Begriffsketten an beliebiger Stelle geöffnet und Zwischenbegriffe eingefügt werden können – ein Vorgang, der mit konventionellen Datenbank-Methoden nicht so problemlos verläuft:

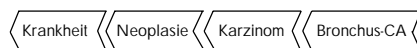


Abb. 3: Das Begriffsmolekül von Abb. 2 - um einen Zwischenbegriff erweitert

Abbildungen 2 und 3 machen übrigens deutlich, wie die Zahl der formalen Elemente mit Begriffsmolekülen reduziert wird: Wir kommen mit einem formalen Element aus, während der Conceptual Graph von Abb. 1 mit Rechteck, Oval, Pfeil und Doppelpunkt deren vier brauchte.

#### 4. Implizite Darstellung der Relatoren

Die Tatsache, dass wir für die Begriffsmoleküle mit einem formalen Element auskommen, liegt in einer Art „Trick“ begründet. Der in Conceptual Graph von Abb. 1 mit einem Oval dargestellte Relator kommt nämlich bei den Begriffsmolekülen ebenfalls vor, allerdings in einer impliziten Form: immer wenn zwei Begriffe auf der gleichen Zeile nebeneinander stehen, bedeutet das, dass zwischen ihnen eine „is-a“-Relation besteht. Abb. 2 meint also: „Ein Bronchus-CA, das ein Karzinom ist, das eine Krankheit ist“.

Der „is-a“-Relator ist für eine Begriffsrepräsentation sehr fundamental: er verbindet Oberbegriff und Unterbegriff. Das ist nun genau die Beziehung, die auch zwischen Klasse (Typ) und Ausprägung besteht. Deshalb kann sowohl der (ovale) „is-a“-Relator wie der Doppelpunkt von Abb. 1 bei Begriffsmolekülen auf die gleiche Weise dargestellt werden.

Die Bedeutung der Relation zwischen zwei Begriffen ergibt sich bei Begriffsmolekülen auf eine sehr einfache Weise aus ihrer relativen Position: Immer wenn z.B. zwei Begriffe

nebeneinander stehen, ist der Begriff links der Oberbegriff (oder die Klasse) und der Begriff rechts der Unterbegriff (oder die Ausprägung der Klasse). Diese hierarchische („is-a“) Beziehung kann sich über beliebig viele Stufen ausdehnen, Im Sinne eines wirklichen offenen Systems können dabei links und rechts, aber auch an beliebiger Stelle in der Mitte der Kette neue Glieder angefügt werden. Die entstehende Darstellung ist einfach lesbar und platzsparend. Neben der „is-a“-Relation werden bei Begriffsmolekülen auch alle anderen Relation implizit, d.h. durch die relative Position der Begriffsatome dargestellt. Dazu gleich mehr.

#### 5. Die zwei Grundrelationen: Hierarchie und Attribution

In Begriffsmolekülen lassen sich alle<sup>2</sup> Verknüpfungen auf zwei Grundrelationen zurückführen:

1. Die hierarchische Relation (=“is-a“) wird horizontal dargestellt
2. Die attributive Relation (=“has-a“<sup>3</sup>) wird vertikal dargestellt.

Beide Relatoren sind implizit, sie werden am Bildschirm somit nicht gezeichnet, sondern durch die Positionsbezüge der zu verknüpfenden Begriffe dargestellt. Im Hintergrund, d.h. im Datenmodell und für die Kalkulation der Codierungen, werden die Relationen genau verrechnet.

Beide Beziehungen sind asymmetrisch, d.h. die beiden verknüpften Begriffe können nicht gegeneinander ausgetauscht werden. Die Relation enthält also eine Richtung, die nicht umgekehrt werden kann. Diese Richtung wird am Bildschirm durch die Links-Rechts-Achse definiert:

	Links	Rechts
Hierarchie	Oberbegriff	Unterbegriff
Attribution	attributionierter Begriff	Attribut

Tab. 2: die zwei Grundrelationen sind asymmetrisch

Haben wir bisher mit den Hierarchien der Abb. 2 und 3 nur lineare Ketten von Begriffen dargestellt, können wir mit den Attributen auch verzweigte Begriffsmoleküle zeichnen:

<sup>1</sup> Begriffe (concepts) werden bei den CGs mit Rechtecken, Relationen mit Ovalen dargestellt. Ein Begriff kann, muss aber nicht aus einem Klasse-Ausprägungs-Paar bestehen. Der Begriff rechts in Abb.1 ist ein einfacher Begriff.

<sup>2</sup> Der Vollständigkeit halber sei erwähnt, dass für die „part-of“-Relation ein dritter Relator eingeführt werden kann. Mit Begriffsmolekülen können die „part-of“-Verhältnisse logisch sehr klar dargestellt werden, wie wir in einer weiterführenden Arbeit zeigen werden.

<sup>3</sup> Sprachlich korrekt müsste es „has-Attribute:“ heißen. Wir lassen es bei der salopperen Form.

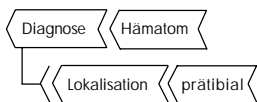


Abb. 4: Ein verzweigtes Begriffsmolekül mit einer „has-a“-Relation

Der Begriff „Diagnose“ in Abb. 4 hat ein Attribut, nämlich „Lokalisation“. Mit diesem Attribut ist er über eine attributive Relation verbunden, die mit dem Haken unter „Diagnose“ dargestellt ist. Eine konventionelle Begriffsrepräsentation würde den Sachverhalt z.B. so darstellen:

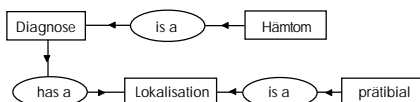


Abb. 5: Der Inhalt von Abb. 4 in konventioneller Notation

Abb. 5 enthält nicht nur mehr Elemente als Abb. 4, sie ist in ihrer räumlichen Anordnung auch nicht normiert, und deshalb räumlich nicht stabil. Abb. 6 zeigt eine mögliche Variante:

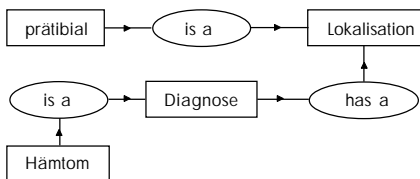


Abb. 6: Variante von Abb.5 (konventionelle Notation)

## 6. Vorteile der strikten räumlichen Anordnung innerhalb des Begriffsmoleküls

Für Begriffsmoleküle ist die räumliche Anordnung immer vom Inhalt der Begriffe her gegeben. Die normierte räumliche Anordnung in Begriffsmolekülen hat für den Wissensbasis-Bearbeiter nicht nur den Komfort einer Gewohnheit. Die Lesbarkeit und Bearbeitungsqualität wird durch die strenge Systematik der Begriffsmoleküle erhöht:

1. Auf jeder Zeile finden sich ausschließlich Begriffe aus einer Hierarchie (d.h. ausschließlich Begriffe aus der gleichen semantischen Dimension, s. [6,8] und Abb. 10)
2. Sobald die Zeile gewechselt wird, wird auch die semantische Dimension gewechselt. Z.B. enthält in Abb. 5 die erste Zeile Angaben zur Diagnose, d.h. zur Pathologie, die zweite Zeile Angaben zur Lokalisation – zwei semantisch klar unterschiedliche Qualitäten, die

zwei verschiedenen semantischen Dimensionen entsprechen.

3. Die Anzahl Zeilen zeigt somit immer die Anzahl der semantischen Dimensionen auf.
4. Links auf jeder Zeile findet sich immer der allgemeinste Begriff der entsprechenden semantischen Dimension (z.B. ist Diagnose allgemeiner als Trauma oder als Hämatom, Lokalisation allgemeiner als prätibial oder Unterschenkel).
5. Links oben findet sich immer die Wurzel des gesamten Begriffsmoleküls, d.h. derjenige Begriff, dem sich alle anderen im Molekül auf die eine oder andere Weise subsumieren. Zugleich ist die oberste Zeile diejenige semantische Dimension, der sich die semantischen Dimensionen der anderen Zeilen unterordnen.
6. Da ein Molekül mit allen seinen Verästelungen (siehe z.B. Abb. 10) immer eine genau reglierte Baumstruktur hat, lässt es sich von einer Maschine (einem Computerprogramm) auf eine einfache Weise systematisch verarbeiten.

Der letzte Punkt zeigt, wie eine Darstellung, die der menschlichen Lesbarkeit entgegenkommt, auch die Lesbarkeit durch die Maschine verbessern kann.

## 7. Umwandlung der benannten Relatoren in unbenannte

Begriffsmoleküle kommen vollständig ohne benannte Relatoren aus. Die beiden Relatoren der Begriffsmoleküle, der hierarchische und der attributive, sind allein durch ihre Stellung erkennbar und brauchen am Bildschirm nicht mit einem Namen gekennzeichnet zu werden. Die übrigen benannten Relatoren werden in eine Kombination von attributiver und hierarchischer Relation umgeformt:

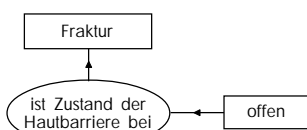


Abb. 7: Conceptual Graph mit benanntem Relator

Der benannte Relator wird in eine Kombination von 1 Konzept und 2 Grundrelatoren umgeformt:

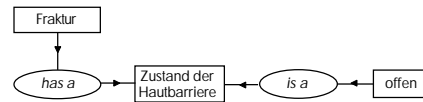


Abb. 8: Information der Abb. 7 mit ausschliesslicher Verwendung der beiden Grundrelatoren

Diese Information wird dann mit einem Begriffsmolekül geschrieben:

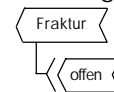


Abb. 9: Information von Abb. 8, in einem Begriffsmolekül dargestellt

Abb. 9 ist kürzer und besser lesbar als Abb. 7 oder 8. Allerdings ist das Konzept „Zustand der Hautbarriere“ ausgelassen. Dies ist zulässig, weil bei Begriffsmolekülen alle Bindungsstellen – auch wenn sie nicht benannt sind – im gesamten semantischen Netz eindeutig sind. In der aktuellen Zürcher Wissensbasis hat z.B. der Begriff Fraktur sieben unterschiedliche und nur ihm zugehörige attributive Bindungsstellen. An eine dieser Bindungsstellen binden ausschliesslich die beiden Begriffe „offen“ und „geschlossen“, an eine andere die beiden Begriffe „intraartikulär“ und „extraartikulär“. Obwohl die Bindungsstellen unbenannt sind, haben sie doch eine eindeutige Identität, und der Bearbeiter sieht die inhaltliche Bedeutung der Bindungsstelle sofort an den verknüpften Begriffen – wie z.B. in Abb. 9. Dies hat folgende Vorteile:

1. Der Bearbeiter muss sich keine Gedanken über die geeigneten Namen der Relatoren machen (Diese sind nämlich selten völlig zufriedenstellend und oft etwas umständlich, wie Abb. 7 zeigt). Die Arbeitszeit der Namenswahl wäre eine unproduktive Zeit, die nicht der Notation von Wissen dient. Unvermeidlich wäre zudem die gelegentliche Vergabe von unterschiedlichen Namen für die gleiche Relation, was zusätzliche Verwirrung schaffen würde. Bei Begriffsmolekülen treten solche Erschwernisse nicht auf.
2. Die Darstellung mit Begriffsmolekülen (Abb. 9) braucht am Bildschirm weniger Platz.
3. Sie ist deshalb schneller lesbar.
4. Es kann dadurch insgesamt mehr Inhalt auf einen Blick aufgenommen werden.

- Der Inhalt ist trotzdem immer sehr klar.
- Auch die Schlüsse, die der Computer aus den Begriffsmolekülen zieht, sind immer eindeutig und beziehen sich nur auf eine ganz bestimmte Bindungsstelle – so wäre sie benannt. (Dies daher, weil im Hintergrund, die Bindungsstelle eindeutig identifiziert ist)

Ein weiterer, in der Tabelle nicht aufgeführter Vorteil hat mit semantischen Grundlagenüberlegungen zu tun. Die beiden Basisrelatoren, der hierarchische und der attributive, entsprechen nämlich den beiden grundsätzlichen Beziehungen, die zwei Werte innerhalb des semantischen Raumes zu einander einnehmen können (s. Kap. 3.3 in [5]). Zudem weisen Begriffe in Begriffsmolekülen (Abb. 10) und die Objekttypen der OOP eine verblüffende Verwandtschaften auf (s. Kap. 8 in [5]).

Ganz abgesehen von solchen theoretischen Überlegungen zeigt die praktische Arbeit, dass Begriffsmoleküle ohne benannte Relatoren funktionieren, dass sie sehr gute, präzise Resultate ergeben und es erlauben, grosse, komplexe Wissensbasen übersichtlich zu warten.

## 8. Mehrfach verzweigte Begriffsmoleküle

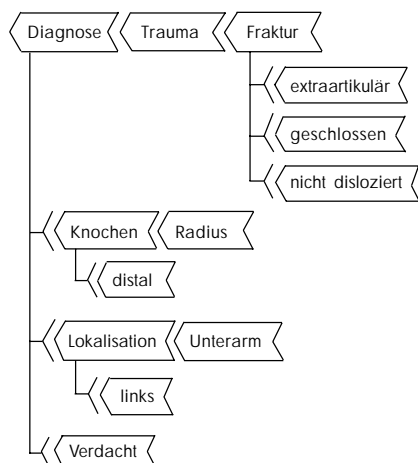


Abb. 10: Ein durchschnittlich verzweigtes Begriffsmolekül

Abb. 10 zeigt das Zusammenspiel der hierarchischen und der attributiven Relationen in einer konkreten Diagnose. Das Beispiel enthält 10 Zeilen, und somit 10 Hierarchien (bzw. Dimensionen oder semantische Achsen). Die Wurzel des Begriffsmoleküls ist der Begriff „Diagnose“. Der

„Diagnose“ sind auf drei Achsen drei direkte Attribute zugeordnet: „Knochen“ (implizite Achse: Organ), „Lokalisation“ (implizite und explizite Achse: Lokalisation) und „Verdacht“ (implizite Achse: Certainty). Der Begriff „Fraktur“ ist ein Unterbegriff von „Diagnose“. Er bildet semantisch einen eigenen Fokus [6] mit eigenen Attributen, nämlich: „extraartikulär“, „geschlossen“ und „nicht disloziert“. Trauma ist ein *bifazialer* Zwischenbegriff (s. Kap. 3). „Knochen“ bildet wie „Fraktur“ einen eigenen Fokus mit einem eigenen Attribut, nämlich „distal“. Wir finden drei hierarchische Ketten: „Fraktur ist ein Trauma, ist eine Diagnose“, „Radius ist ein Knochen“ und „Unterarm ist eine Lokalisation“. Falls der Leser findet, dass die Beschreibung in diesem Abschnitt etwas schwierig zu lesen sei, stimme ich ihm zu; ich bevorzuge die visuelle Darstellung von Abb. 10, welche die Information einfach und klar präsentiert.

Abb. 10 könnte aus folgender Nominalphrase gewonnen sein: „Verdacht auf distale, geschlossene, extraartikuläre und nicht dislozierte Radiusfraktur rechts“. Die Phrase könnte natürlich auch ganz anders formuliert sein. Abgesehen von der Eindeutigkeit der Form hat die Darstellung von Abb. 10 gegenüber der Nominalphrase folgende Vorteile:

- Die impliziten Inhalte sind rekonstruiert („Unterarm“, „Knochen“). Implizite Inhalte können entscheidend sein für die Codevergabe oder die Abfrage in einem Data-Warehouse.
- Die Verknüpfungen sind viel klarer. So gehört „distal“ zum Komplex „Knochen<Radius“ und nicht zum Komplex der Fraktur. Ohne eine solche klare mehrdimensionale und multifokale Strukturierung der zugrundeliegenden Datenstruktur kann die darauf aufsetzenden Inferenzmaschine für die Analyse der Nominalphrasen und die Codegenerierung nicht funktionieren. Die Strukturierung hilft zudem dem Wissensingenieur entscheidend bei der Übersicht über die Wissensbasis. Mensch und Maschine profitieren beide gleichermassen von der eindeutigen logischen Struktur der Begriffsmoleküle.

## 9. Darstellung von Verarbeitungsregeln

Abb. 10 zeigt Begriffe in einer bestimmten Anordnung und stellt einen momentanen Zustand von Information dar. Ein solcher Zustand (Status) wird durch eine Verarbeitungsregel verändert. Die Regel bewirkt, dass ein bestimmter Vorzustand (ein „if“) in einen Folgezustand (ein „then“) übergeführt wird. Die Regel führt somit einen Status in einen anderen Status über; sie wirkt dynamisch.

Da es sich bei unseren Zuständen und Regeln immer und ausschliesslich um Information handelt, können wir von *Infostatik* und *Infodynamik* sprechen. Ganz allgemein gesagt: Das *Datenmodell* entspricht der Infostatik, die Regeln der Textanalyse (die *Algorithmen*) entsprechen der Infodynamik. Der Input für die Inferenzmaschine (den semantischen Interpretier) ist ein Status, ebenfalls der Output und alle Zwischenzustände bei der Verarbeitung.

Begriffsmoleküle wie in Abb. 10 dienen der Darstellung von Zuständen (Infostatik). Für die Darstellung der Regeln, d.h. für die Algorithmen (Infodynamik) braucht es nur wenige zusätzliche formale Elemente, die zeigen, welche Veränderungen in den Begriffsmolekülen durchgeführt werden. Vier einfache formale Elemente sind ausreichend, drei explizite und ein implizites. Es handelt sich um *Operatoren*, die den einzelnen Atomen in den Begriffsmolekülen zugeordnet werden.

	vorher	nachher
positiv	if (bland, weiss)	then add (grün)
negativ	if not (blau)	then remove (rot)

Tab. 3: Die vier Operatoren für semantische Regeln

Eine Regel wird geschrieben, in dem ein Begriffsmolekül mit solchen Operatoren versehen wird. Die Regel von Abbildung 11 sagt z.B. aus, dass ein Karzinom als Neoplasie maligne ist. Die Unterstreichung des Begriffs „maligne“ kennzeichnet den „then-add“-Operator. Die Begriffe „Neoplasie“ und „Karzinom“ sind ohne Operator geschrieben und somit Bedingungen („if“-s). Durch die Möglichkeit, verzweigte Ketten zu bilden, können die Bedingungen für die Anwendung der Regeln präzise gesteuert werden. In kleinen, gut kontrollier

baren Regelschritten wird sukzessive der Input interpretiert, bis zur Codierung. Die Erfahrung zeigt, dass auch ein komplexes Regelwerk auf diese Weise übersichtlich und effizient gestaltet werden kann.

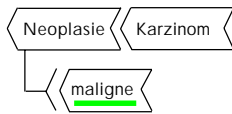


Abb. 11: Eine einfache Regel mit einem „then-add“-Operator (am Bildschirm ist die Unterstreichung grün).

## Referenzen:

[1] Rector, A et al, 1995: A Terminology Server for Medical Language and Medical Information Systems. *Methods of*

*Information in Medicine*, 34. S. 147-157.

[2] Sowa, J.F. 1984: *Conceptual Structures: Information Processing in Mind and Machine*. Reading: Addison-Wesley, 481 S.

[3] Sowa, J.F. 2000: *Knowledge Representation: Logical, Philosophical and Computational Foundations*. Pacific Grove: Brooks/Cole, 594 S.

[4] Straub, H.R, 1994: Wissensbasierte Interpretation, Kontrolle und Auswertung elektronischer Patientendossiers. – In: Kongressband der IX. Jahrestagung der SGMI. Schweizerische Gesellschaft für Medizinin-

formatik, Nottwil, S. 81-87.

[5] Straub H.R, 2001.: *Das interpretierende System - Wortverständnis und Begriffsrepräsentation in Mensch und Maschine*, mit einem Beispiel zur Diagnose-Codierung. Z/I/M-Verlag, Wolfertswil, 176 S.

[6] Straub H.R. 2001: *Four Different Types of Classification Models*. – In: Grütter R. (Hrsg.), 2001: *Knowledge Media in Health Care: Opportunities and Challenges*. Hershey / London: Idea Group Publishing, S. 58 – 82. (im Druck).

[7] <http://www.w3.org/RDF/>

[8] <http://www.meditext.ch/texte/infogehalt.htm>