Syntax und Semantik:

Repräsentation und Bedeutung von Informationen

Der Informationsbegriff

Die Informatik ist eine junge und sich stürmisch entwickelnde Disziplin. Entsprechend schwer ist es, eine Definition dieses Gebiets vorzunehmen, die nicht alle zehn Jahre aktualisiert werden muss. Manfred Broy verwendet in seinem Buch Informatik die folgende Begriffsbestimmung:

Informatik ist die Wissenschaft, Technik und Anwendung der maschinellen Verarbeitung, Speicherung und Übertragung von Information.

Zur Beantwortung der Frage, wie man das Wort *Information* in der Informatik zu verstehen hat, muss man unter den vielen Aspekten, die diesen Begriff prägen, zwei genauer betrachten:

1. Information hat eine äußere Gestalt. Man bezeichnet diese auch als Darstellung oder Repräsentation der Information. Dabei muss die Darstellung einer Information bestimmte formale Regeln berücksichtigen, die sogenannte Syntax.

2. Der Kern der Information ist die aus der Darstellung abgeleitete Bedeutung, ihr Inhalt, die abstrakte Information. Diese Seite der Information nennt man die Semantik.

So wird z.B. ein digitales Bild durch eine in einem rechteckigen Schema angeordnete Kollektion von Farbwerten (Zahlen aus einem bestimmten Bereich) repräsentiert. Der Inhalt ist das Bild, das man erhält, wenn jeder Farbwert in dem Schema durch ein entsprechend gefärbtes kleines Quadrat (Pixel) ersetzt wird.

Um das Zusammenspiel beider Seiten genauer zu beschreiben, kann man sogenannte *Informationssysteme* (siehe Broy) einführen. Ein solches System besteht aus einer Menge R von (zulässigen) Repräsentationen, einer Menge A von abstrakten Informationen und einer Interpretation $I : R \longrightarrow A$, also einer Funktion, die jeder Repäsentation die entsprechende Information zuordnet.

Bevor wir verschiedene Beispiele von Informationssystemen besprechen, müssen wir einige grundlegende Begriffe und Fakten über natürliche Zahlen, Rekursion, vollständige Induktion und formale Sprachen einführen.

Natürliche Zahlen und vollständige Induktion

Wir setzen voraus, dass die natürlichen Zahlen mit den Grundoperationen Addition und Multiplikation aus der Schule bekannt sind. Es ist aber wichtig zu wissen, dass man all die bekannten Fakten aus wenigen elementaren Grundvoraussetzungen, den sogenannten Peano–Axiomen ableiten kann. Diesen Axiomen definieren die Menge der $\mathbb{N}$ der natürlichen Zahlen wie folgt.

Definition: Die Menge $\mathbb{N}$ der natürlichen Zahlen ist die kleinste Menge, die das Element 0 enthält, und für jedes $n \in \mathbb{N}$ auch einen eindeutig bestimmten Nachfolger $s(n)$ enthält, wobei zwei verschiedene natürliche Zahlen auch verschiedene Nachfolger haben müssen und 0 kein Nachfolger einer natürlichen Zahl ist.

Auch wenn das auf den ersten Blick verwirrend scheinen mag, ist das Prinzip dahinter sehr einfach: Den Nachfolger von 0 nennen wir 1 und wegen $0 \in \mathbb{N}$ ist auch 1 eine natürliche Zahl, ebenso die 2 als Nachfolger der 1, die 3 als Nachfolger der 2 usw. Man kann damit jede natürliche Zahl von der 0 aus durch eine endliche Anzahl von Aufrufen der Nachfolgerfunktion erreichen und der Nachfolger $s(n)$ der Zahl n ist die Zahl, die man üblicherweise als $n + 1$ schreibt.
Beim rein axiomatische Ansatz geht man den umgekehrten Weg, man definiert die Addition mit Hilfe des Nachfolgers. Für eine beliebige (aber fest gewählte) natürliche Zahl m wird die Addition einer weiteren natürlichen Zahl n zu m durch die folgenden zwei Regeln festgelegt:

(1)	$m + 0 := m$	(das ist der Fall $n = 0$)
(2)	$m + s(n) := s(m + n)$	(damit wird die Addition des Nachfolgers $s(n)$ auf die Addition von n zurückgeführt)

Man sollte sich wieder an einem Beispiel (sagen wir mit $m = 5$) klar machen, dass sich diese Definition mit den bekannten Fakten deckt:

$$\begin{array}{lllll}
5+0 & = 5 & & & \\
5+1 & = 5+s(0) & = s(5+0) & = s(5) & = 6 \\
5+2 & = 5+s(1) & = s(5+1) & = s(6) & = 7 \\
5+3 & = 5+s(2) & = s(5+2) & = s(7) & = 8 \\
\ldots\ldots & & & &
\end{array}$$

Das Prinzip, das dieser Definition der Addition zu Grunde liegt, nennt man *Rekursion*. Eine rekursive Definition besteht immer aus einer *Verankerung* für einen (oder mehrere) Anfangswert(e) und einer *Rekursionsregel*, mit der man die Operation für größere Werte auf die Opeartion mit kleineren Werten (in unserem Fall mit dem direkten Vorgänger) zurückführt.

Unter der Voraussetzung, dass man die Addition für beliebige natürliche Zahlen ausführen kann, lässt sich die Multiplikation auf ähnliche Weise rekursiv definieren:

(1)	$m \cdot 0 := 0$	(Verankerung)
(2)	$m \cdot s(n) := (m \cdot n) + m$	(Rekursionsregel)

In der axiomatischen Theorie kann man die (empirisch) bekannten Assoziativ-, Kommutativ- und Distributivgesetze mathematisch beweisen. Als Beweismethode verwendet man die *vollständige Induktion*, eine Technik, deren Grundidee große Ähnlichkeit zur Rekursion hat. Der wesentliche Unterschied besteht im Ziel: Geht es bei der Rekusion darum, eine Definition zu geben, die für alle natürlichen Zahlen gültig ist, soll mit der Induktion eine Aussage bewiesen werden, die für alle natürlichen Zahlen gültig ist. Um das nachzuweisen, zeigt man zuerst, dass die Aussage für einen (oder mehrere) Anfangswert(e) gültig ist - diesen Teil nennt man *Induktionsanfang* oder *Induktionsverankerung* - und man beweist danach mit dem sogenannten *Induktionsschritt*, dass aus der Gültigkeit der Aussage für eine beliebige (aber fest gewählte) Zahl $n \in \mathbb{N}$ auch die Gültigkeit der Aussage für $n + 1$ folgt. Die Gültigkeit für n nennt man *Induktionsvoraussetzung* und die Gültigkeit für $n+1$ die *Induktionsbehauptung*. Da die Induktionsbeweise der oben genannten Rechenregeln teilweise schwierig sind (in einigen Fällen erfordern sie eine Doppelinduktion bei der im Induktionschritt bezüglich n ein Induktionsbeweis bezüglich m verschachtelt ist), wollen wir uns hier mit einem einfacheren Beispiel begnügen, wobei die Rechenregeln als bekannt (also bereits bewiesen) vorausgesetzt werden.

Aussage: Für jede natürliche Zahl n gilt $0 + 1 + 2 + \ldots + n = \frac{n\cdot(n+1)}{2}$.

Wir führen für die linke Seite der Gleichung die symbolische Schreibweise $\sum_{i=0}^{n} i$ ein und führen einen Beweis mit vollständiger Induktion.

Induktionsanfang: Für $n = 0$ gilt $\sum_{i=0}^{0} i = 0 \quad = \frac{0\cdot(0+1)}{2}$

Induktionsvorauss.: $\sum_{i=0}^{n} i = \frac{n\cdot(n+1)}{2}$

Induktionsbeh.: $\sum_{i=0}^{n+1} i = \frac{(n+1)\cdot(n+2)}{2}$

Induktionsschritt:

$$\begin{aligned}
\sum_{i=0}^{n+1} i &= \sum_{i=0}^{n} i + (n+1) && \text{(Summand abtrennen)}\\
&= \frac{n\cdot(n+1)}{2} + (n+1) && \text{(nach Induktionvorauss.)}\\
&= \frac{n\cdot(n+1)+2\cdot(n+1)}{2} && \text{(Bruchrechnung)}\\
&= \frac{(n+2)\cdot(n+1)}{2} && \text{(nach Distributivgesetz)}\\
&= \frac{(n+1)\cdot(n+2)}{2} && \text{(nach Kommutativgesetz)}
\end{aligned}$$

$\square$

Man kann vollständige Induktion auch verwenden, um zu beweisen, dass eine bestimmte Aussage für alle natürlichen Zahlen ab einer Zahl n_0 gültig ist - in diesem Fall muss der Induktionsanfang für n_0 gezeigt werden. Ob man beim Induktionsschritt die Induktionsvoraussetzung für die Zahl n und die Induktionsbehauptung für die Zahl $n + 1$ formuliert oder ob man für die Voraussetzung $n - 1$ und für die Behauptung n wählt, ist eine reine Geschmacksfrage. Eine weitere Variante dieser Beweismethode ist die verallgemeinerte Induktion. Während man in der bisher besprochenen Variante den Fall n auf den Fall $n - 1$ (oder $n + 1$ auf n) zurückführen musste, kann man bei der verallgemeinerten Induktion die Behauptung für n nachweisen, indem man die Gültigkeit für alle Werte zwischen 0 (bzw. zwischen dem Anfangswert n_0) und $n - 1$ voraussetzt. Ein typisches Beispiel dafür ist die Aussage, dass man jede natürliche Zahl $n \geq 2$ als ein Produkt von Primfaktoren darstellen kann.

Eine weitere Beweisvariante wird strukturelle Induktion genannt. Dabei geht es um Aussagen, die für alle Vertreter mit einer bestimmten Struktur gelten sollen, wobei jeder einzelne Vertreter einen Rang in Form einer natürlichen Zahl hat. Beispielsweise könnte man Listenstrukturen betrachten bei denen der Rang einer konkrete Liste durch die Länge der Liste definiert ist. Eine Aussage über alle Listen ist dann wahr, wenn sie für die leere Liste gilt (das ist die einzige Liste der Länge 0) und aus der Gültigkeit der Aussage für alle Listen der Länge n auch die Gültigkeit der Aussage für alle Listen der Länge $n+1$ abgeleitet werden kann.

Formale Sprachen

Eine einfache und gebräuchliche Form der Repräsentation von Information sind sogenannte Wörter (bevorzugter Begriff in der theoretischen Informatik) oder Strings (bevorzugter Begriff in Programmiersprachen). Das sind Sequenzen von (nicht durch Kommata getrennten) Symbolen aus einem Alphabet Σ. Die folgenden Stichpunkte fassen die wichtigsten damit in Verbindung stehenden Definitionen zusammen:

- Ein *Alphabet* ist eine endliche, nichtleere Menge von Symbolen (Zeichen), wie z.B. $\Sigma = \{0, 1\}$ oder $\Sigma = \{0, 1, 2, \ldots, 9\}$ oder $\Sigma = \{a, b, c, \ldots, z\}$.
- Ein *Wort* (*String*) über einem Alphabet Σ ist eine endliche, möglicherweise leere Folge von Symbolen aus Σ dargestellt als $w = x_1 x_2 \ldots x_n$, wobei $x_i \in \Sigma$ für $i = 1, \ldots, n$. In diesem Fall ist n die Länge des Worts w, als Formel $n = |w|$.
- Das leere Wort wird mit ε bezeichnet. Natürlich gilt $|\varepsilon| = 0$ und man muss $\varepsilon \notin \Sigma$ voraussetzen, weil man anderenfalls das leere Wort nicht von dem aus dem Zeichen ε gebildeten Wort der Länge 1 unterscheiden könnte.
- Mit Σ^* bzw. mit Σ^+ und Σ^k bezeicht die Menge aller Wörter über Σ , bzw. die Menge aller nichtleeren Wörter und die Menge aller Wörter der Länge k, d.h.
  - $\Sigma^0 = \{\varepsilon\}$
  - $\Sigma^k = \{x_1 x_2 \ldots x_k \mid x_i \in \Sigma \text{ für alle } i = 1, \ldots k\ \}$
  - $\Sigma^* = \bigcup_{k=0}^{\infty} \Sigma^k \quad \text{und} \quad \Sigma^+ = \bigcup_{k=1}^{\infty} \Sigma^k$
- Die Operation des Verknüpfens von zwei Wörtern durch einfaches Hintereinanderschreiben nennt man *Konkatenation.* Sie wird durch das Operationszeichen $\circ$ symbolisiert, d.h. für $w = x_1 \ldots x_n \in \Sigma^*$ und $w' = x'_1 \ldots x'_m \in \Sigma^*$ gilt
  $$w \circ w' \;=\; x_1 \ldots x_n x'_1 \ldots x'_m \quad \text{und} \quad w \circ \varepsilon \;=\; \varepsilon \circ w \;=\; w.$$

Durch Verwendung der Konkatenation kann man die folgende (alternative) Definition der Menge Σ^* einführen:

$$\Sigma^* \;=\; \{\varepsilon\} \;\cup\; \Sigma \circ \Sigma^*$$

In dieser Definition fallen zwei ungewöhnliche Punkte auf. Erstens ist die Operation $\circ$ für die Verknüpfung von Wörtern definiert, aber hier wird sie auf Mengen von Wörtern angewendet. Zweitens taucht die Menge, die definiert werden soll auf der

rechten Seite wieder auf.
Das erste Problem wird durch eine naheliegende und in der Mathematik häufig verwendete Konstruktion gelöst, mit der man eine Operation, nennen wir sie $\otimes$, zur Verknüpfung von Objekten, nennen wir diese x und y, auf Mengen von Objekten, nennen wir sie X und Y oder auf Verknüpfungen von einem Objekt mit einer Menge übertragen kann:

$$X \otimes Y \;=\; \{\, x \otimes y \mid x \in X \;\text{ und }\; y \in Y \,\} \;\text{ und }\; x \otimes Y \;=\; \{\, x \otimes y \mid y \in Y \,\}.$$

Der zweite Punkt erinnert an eine rekursive Definition, in der die zu definierende Funktion sich selbst aufruft. Damit ergibt sich für die obige Definition von Σ^* (die zunächst nur ein formaler Ausdruck ist) die folgende Interpretation:

Σ^* ist die kleinste Menge mit den folgenden zwei Eigenschaften:

1. $\varepsilon \in \Sigma^*$
2. $\underbrace{\text{Wenn } x \in \Sigma \text{ und } w \in \Sigma^*,}_{\text{Prämisse}}\; \underbrace{\text{so ist auch } x \circ w \in \Sigma^*}_{\text{Konklusion}}$

Da sich verschiedene Varianten von rekursiven Definitionen wie ein roter Faden durch die Vorlesung ziehen werden, geben wir hier noch eine alternative und speziell auf die obige Beschreibung von Σ^* zugeschnittene Definition der Länge von Wörtern an: Die Länge $|w|$ eines Wortes $w \in \Sigma^*$ ist wie folgt definiert:

1. $|\varepsilon| = 0$
2. Wenn $x \in \Sigma$ und $w \in \Sigma^*$, dann ist $|x \circ w| = 1 + |w|$

Zur Längenbestimmung des Wortes $aba \in \{a, b\}^*$ wendet man dreimal die Regel 2 und einmal die Regel 1 an:

$$\begin{aligned} |aba| &= \underbrace{|a \circ ba| = 1 + |ba|}_{\text{Regel 2}} \\ &= \underbrace{1 + |b \circ a| = 1 + (1 + |a|)}_{\text{Regel 2}} \\ &= \underbrace{2 + |a \circ \varepsilon| = 2 + (1 + |\varepsilon|)}_{\text{Regel 2}} \\ &= \underbrace{3 + |\varepsilon| = 3 + 0}_{\text{Regel 1}} = 3 \end{aligned}$$

Definition: Eine *formale Sprache* über einem Alphabet Σ ist eine Teilmenge $L \subseteq \Sigma^*$. Unter dem *Wortproblem* von L versteht man die Frage, für ein beliebiges Wort $w \in \Sigma^*$ zu entscheiden, ob es zu L gehört oder nicht.

Zahldarstellungen als Beispiel von Informationssystemen

Dezimaldarstellung von natürlichen Zahlen

Natürliche Zahlen werden traditionell im Dezimalsystem repräsentiert, zumindest in unserem Kulturkreis - die Römer haben es bekanntlich anders gemacht. Als Alphabet Σ verwendet man dabei die Menge der Ziffern von 0 bis 9. Da es nicht üblich ist, die Darstellung einer Zahl mit einer Null zu beginnen (außer bei der Null selbst), wird die Menge R der zulässigen Repräsentationen wie folgt definiert:

$$R = \{0\} \cup \{1, 2, \ldots, 9\} \circ \{0, 1, \ldots, 9\}^*$$

Die Menge A der abstrakten Informationen besteht aus den natürlichen Zahlen, d.h. $A = \mathbb{N}$ und für die Interpretation I muss man neben der richtigen Interpretation der Ziffern durch die entsprechenden (abstrakten) natürlichen Zahlen auch schon die Grundrechnenoperationen voraussetzen. Um die Regel für die Interpretation möglichst einfach zu halten ist es üblich, die allgemeine Form einer Zahldarstellung als Wort $a_n a_{n-1} \ldots a_2 a_1 a_0$ zu schreiben:

$$I(a_n\, a_{n-1} \ldots a_1\, a_0) = \sum_{i=0}^{n} a_i \cdot 10^i = a_n \cdot 10^n + a_{n-1} 10^{n-1} + \ldots + a_1 \cdot 10^1 + a_0 \cdot 10^0$$

Dabei sind die Potenzen wieder rekursiv definiert durch $10^0 = 1$ und $10^{k+1} = 10 \cdot 10^k$.

Binärdarstellung und k–näre Darstellungen von natürlichen Zahlen

Das traditionelle Dezimalsystem, dessen Bedeutung nur daraus zu erklären ist, dass die Natur uns mit zehn Fingern ausgestattet hat, bekam mit der Entwicklung von Rechenmaschinen und Computern einen wichtigen Konkurrenten, das binäre Zahlensystem. In diesem System sind nur die Ziffern 0 und 1 zugelassen und da man auch hier führende Nullen verhindern will, kann die Menge R der zulässigen Repräsentationen durch $\{0\} \cup 1 \circ \{0, 1\}^*$ beschrieben werden. Die Interpretation einer durch $a_k a_{k-1} \ldots a_1 a_0$ dargestellten Zahl (mit $a_k = 1$ und $a_i \in \{0, 1\}$ für $i = 0, 1, \ldots k-1$) erfolgt analog zur Dezimaldarstellung, aber mit der 2 an Stelle der 10:

$$I(a_n\, a_{n-1} \ldots a_1\, a_0) = \sum_{i=0}^{n} a_i \cdot 2^i = a_n \cdot 2^n + a_{n-1} 2^{n-1} + \ldots + a_1 \cdot 2^1 + a_0 \cdot 2^0$$

Allgemein kann man für eine beliebige ganzzahlige Basis $k \geq 2$ den Ziffernbereich $\{0, 1, \ldots, k-1\}$ festlegen und jeder Ziffernfolge $a_n a_{n-1} \ldots a_1 a_0$ die Zahl

$$z = \sum_{i=0}^{n} a_i k^i = a_n k^n + a_{n-1} k^{n-1} + \ldots + a_1 k + a_0$$

zuordnen. Man spricht von der k-nären Darstellung der Zahl z. Um von der Darstellung zur dargestellten Zahl zu kommen, muss also nur eine Summenformel ausgewertet werden. Die größte Zahl, die man k-när mit n Ziffern darstellen kann ist $k^n - 1$.

Eine besondere Rolle spielt die Basis $k = 8$ bei deren Verwendung das sogenannte Oktalsystem entsteht und die Basis $k = 16$, die dem sogenannten Hexadezimalsystem zu Grunde liegt. Da man in Hexadezimalsystem Ziffer–Symbole für die Zahlen von 0 bis 15 benötigt, werden die Ziffern des Dezimalsystems durch die Großbuchstaben von A bis F ergänzt, d.h. $\Sigma = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E, F\}$.

Berechnung der Darstellung von Zahlen im k–nären System

Grundlage der Bestimmung der k-nären Darstellung einer gegeben Zahl z ist die ganzzahlige Divsion durch k mit Rest, die vom Schulverfahren zur "schriftlichen" Division bekannt ist. Der Fakt, dass für jedes $z \in \mathbb{N}$ und jeden positiven Teiler $k \in \mathbb{N}$ zwei eindeutig bestimmte Zahlen $q, r \in \mathbb{N}$ existieren, so dass $z = q \cdot k + r$ und $r \in \{0, 1, \ldots, k - 1\}$ erfüllt sind, liefert übrigens auch ein schönes Beispiel für einen Beweis mit verallgemeinerter Induktion. Der ganzzahlige Quotient q und der Rest r bei ganzzahliger Division von z durch k werden symbolisch durch $\lfloor \frac{z}{k} \rfloor$ bzw. $z \texttt{ mod } k$ bezeichnet.

Für die k–näre Darstellung $a_n a_{n-1} \ldots a_1 a_0$ von z muss offensichtlich $z \texttt{ mod } k$ mit dem gesuchten a_0 übereinstimmen, denn alle anderen Summanden in $\sum_{i=0}^{n} a_i k^i$ sind durch k teilbar. Darüber hinaus muss die Folge der noch zu bestimmenden Ziffern $a_n \ldots a_1$ die Zahl $\lfloor \frac{z}{k} \rfloor$ darstellen. Man kann also rekursiv fortfahren bis der ganzzahlige Quotient 0 wird.

```
Pseudocode zur Bestimmung der k-naeren Darstellung von z
if (z==0) rep = 0 else rep = ε
                Kommentar: ε ist das leere Wort
while z > 0
      a = z mod k
      rep = a∘rep
      z = ⌊z/k⌋
return rep
```

Bei der Binärdarstellung besteht die Restbestimmung nur aus der Fallunterscheidung, ob der aktuelle Wert von z gerade (Rest 0) oder ungerade (Rest 1) ist.

Für die Grundrechenarten mit Zahlen im k–nären System kann man die in der Schule gelernten Schemata zur "schriftlichen" Addition, Subtraktion, Multiplikation und Division anwenden, aber mit dem Unterschied dass sich die einzutragenen Ziffern nicht als Rest des Operationsergebnisses beim Teilen durch 10 sondern beim Teilen durch k ergeben und dass man zur Bestimmung der Überträge das Operationsergebnis nicht ganzzahlig durch 10 sondern durch k teilt.

Multiplikationsmethode der russischen Bauern

Diese Methode hat auf den ersten Blick wenig mit Binärdarstellungen zu tun: Um zwei natürliche Zahlen k und m zu multiplizieren trägt man beide in die erste Zeile einer zweispaltigen Tabelle ein. Für die jeweils nächste Zeile verdoppelt man die linke Zahl und teilt die rechte ganzzahlig (d.h. mit Abrunden) durch 2. Man bricht ab, wenn auf der rechten Seite eine 1 steht und streicht anschließend alle Zeilen mit einem geraden Wert in der rechten Spalte. Von den verbleibenden Zeilen werden die Werte in der linken Spalte addiert und die so erhaltene Summe ist das Produkt $k \cdot m$. Im nachfolgenden Schema wird dieser Algorithmus am Beispiel der Multiplikationsaufgabe $26 \cdot 41$ demonstriert. Daneben stehend das Schema zur schriftlichen Multiplikation (Schulmethode) dieser Zahlen in Binärdarstellung und man kann beobachten, dass die gestrichenen Zeilen genau zu den Nullen in der Binärdarstellung des rechten Faktors korrespondieren. Man kann also sagen, dass die russischen Bauern für ihre Multiplikation die Binärdarstellung genutzt haben, ohne sie wirklich zu kennen.

Multiplikationsmethode der russischen Bauern		Schulmethode mit Binärdarstellung
Aufgabe: 41 x 26		Aufgabe: 1 0 1 0 0 1 x 1 1 0 1 0
~~41~~	~~26~~	~~0 0 0 0 0 0~~
82	13	1 0 1 0 0 1
~~164~~	~~6~~	~~0 0 0 0 0 0~~
328	3	1 0 1 0 0 1
656	1	1 0 1 0 0 1
1066		1 0 0 0 0 1 0 1 0 1 0

Darstellung von gebrochenen Zahlen

Prinzipiell kann man zwei Darstellungsarten unterscheiden: Die von Dezimalbrüchen bekannte Darstellungsart mit Nachkommastellen und die Darstellung durch ganzzahlige Zähler und Nenner. Für die erste Variante geht man wieder von den bekannten Formeln für die Dezimaldarstellung aus und kann analoge Regeln für die Darstellung von gebrochenen Zahlen bezüglich einer anderen Basis ableiten. Da wir die Darstellung des ganzzahligen Anteils (also des Teils vor dem Komma bzw. dem Dezimalpunkt) schon ausführlich besprochen haben, konzentrieren wir uns jetzt auf die Nachkommastellen, die den echt gebrochenen Anteil beschreiben.

Ein Dezimalbruch $0, a_1 a_2 \dots a_n$ repräsentiert die rationale Zahl

$$r = \sum_{i=1}^{n} a_i \cdot 10^{-i} = \frac{a_1}{10} + \frac{a_2}{100} \dots + \frac{a_n}{10^n} = \frac{a_1 10^{n-1} + a_2 10^{n-2} + \dots + a_n}{10^n}$$

Das kann man unmittelbar auf eine beliebige Basis $b > 1$ übertragen. In diesem Fall müssen alle a_i aus dem Bereich $\{0, 1, \dots, k-1\}$ sein und $0, a_1 a_2 \dots a_n$ repräsentiert die Zahl

$$r = \sum_{i=1}^{n} a_i \cdot k^{-i} = \frac{a_1}{k} + \frac{a_2}{k^2} \dots + \frac{a_n}{k^n} = \frac{a_1 \cdot k^{n-1} + a_2 \cdot k^{n-2} + \dots + a_n}{k^n}$$

Um für eine gegebene Zahl r aus dem Bereich $0 < z < 1$ die Darstellung zur Basis k zu berechnen, kann man die folgende Methode verwenden:

```
Pseudocode zur Bestimmung der k-naeren Darstellung von r (0< r< 1)
rep = 0,
while r > 0
      a = ⌊k · r⌋
      rep = rep ∘ a
      r = k · r - a
return rep
```

Zu beachten ist, dass der hier dargestellte Algorithmus eventuell nie terminiert. Es gibt zwei Wege, das zu ändern. Einerseits kann man die Anzahl der zu berechnenden Stellen vorher festlegen und in der `while`–Schleife einen Zähler mitlaufen lassen. In diesem Fall erhält man aber nur eine Näherung der darzustellenden Zahl. Andererseits kann man (wieder in Analogie zu Dezimalbrüchen) den Fakt nutzen, dass für jede rationale Zahl r die Darstellung entweder abbricht (also endlich ist) oder irgendwann periodisch wird. Man erkennt den Zeitpunkt, wann man die Periode erstmalig durchlaufen hat daran, dass sich der aktuelle Wert von `r` zum ersten Mal wiederholt.

Wie wir später sehen werden, ist die gerade besprochene Verallgemeinerung von Dezimalbrüchen auf andere Basen weniger gebräuchlich. In der Praxis spielt neben dem Datentyp der Gleitkommazahl, der sich unmittelbar aus der Dezimalbruchdarstellung ableiten lässt, die Darstellung einer rationalen Zahl durch einen ganzzahligen Zähler $a \in \mathbb{Z}$ und positiv–ganzzahlingen Nenner $b \in \mathbb{N}^+$ eine wichtige Rolle. Insbesondere wird diese Art der Repräsentation dann verwendet, wenn man exakt rechnen will.

Die durch das Paar (a, b) repräsentierte Zahl wird als rationale Zahl (Bruch) $\frac{a}{b}$ interpretiert. An diesem Informationssystem wird zum ersten Mal deutlich, dass die Darstellung eines Objekts nicht eindeutig sein muss. Zum Beispiel repräsentieren die Paare $(2, 6)$ und $(3, 9)$ die gleiche rationale Zahl, nämlich $\frac{2}{6} = \frac{3}{9} = \frac{1}{3}$.

Definition: Ist (R,A,I) ein Informationssystem mit der Interpretation $I : R \longrightarrow A$, dann nennt man zwei Repräsentationen $r_1 \in R$ und $r_2 \in R$ *semantisch äquivalent*, wenn sie die gleiche Interpretation haben, d.h. wenn $I(r_1) = I(r_2)$.

Wie man leicht sieht, gibt es für jeden Bruch einen Standardrepräsentanten bei dem Zähler und Nenner nicht mehr gekürzt werden können. Geht man von einem beliebigen Repräsentanten (a, b) aus, kann der Standardrepräsentant dadurch bestimmt werden, dass der größte gemeinsame Teiler c von a und b berechnet wird und beide Werte durch c geteilt werden.