



© CL 2001 Darmstadt

Dr. Christoph Lübbert
Viktoriastraße 36
D-64293 Darmstadt

Tel: 06151 422298, T-Mob: 0171 2045811,
christoph.luebbert@t-online.de,
www.cl-diesunddas.de

Darmstadt, 23.06.2014, V4f
Änderungen gegenüber V2, V3: Kürzungen für den Vortrag

„Verfeinerung“ der Ontology-Sprache durch die FBA-Sprache

© C. Lübbert, V4f, Juni 2014

Danksagung:

Peter Zahn, Thomas Zeh, Karl Erich Wolff und Hermann Bense danke ich herzlich für Kommentare, Hinweise und Korrekturen bei der Ausarbeitung.

Inhalt

1	Vorbemerkung	3
1.1	Was ist für mich der Zweck eines Konzepts „Informatic Ontology“?	6
1.2	Anschauungsbeispiel aus Wikipedia	7
2	O-Schema	13
3	Instanziierung	18
3.1	Instanziierung in Klassen	18
3.2	Was ist die volle Menge <i>aller</i> Attribute, einer Instanz <i>g</i> ?	22
3.3	Instanziierung in Relationship-Types	23
3.4	Vererbung von Relationship-Types entlang der Klassentaxonomie	24
3.5	Instanziierung in Attributen	28
3.6	Was in den mir bekannten O-Schema-Definitionen auffällt	29
4	Bezug zur FBA (Formale Begriffsanalyse)	30
4.1	Formaler Kontext für die Beziehung zwischen Instanzen und Attributen	30
4.2	Eine philosophisch-historische Anmerkung	38

4.3 Berücksichtigung der <i>Attributwerte</i> durch Bildung eines mehrwertigen Kontextes	40
4.4 Formale Kontexte der Relationship-Types	41
4.5 „Verfeinerung“ einer Ontology im Sinne der FBA	44
5 Schlussbemerkung	46
6 Literatur	48

1 Vorbemerkung

Als ich vor etwa drei Jahren zu diesem erlauchten „**Ontologie-Arbeitskreis**“ stieß, glaubte ich, man werde sich in diesem Kreis u.a. ausführlich mit den derzeit kursierenden **Strukturansätzen zum Konzept „informatische Ontology“** beschäftigen; und da der Ontologie-Arbeitskreis der **h_da** glücklicherweise eng mit dem **FBA-Seminar** (Formale Begriffsanalyse) der **TUD** zusammenarbeitet, hoffte ich, Informatiker und Mathematiker würden sich im Lauf der Zeit in diesem Kreis auch fachsprachlich ein bisschen näher kommen.

Davon habe ich in den letzten 2-3 Jahren aber recht wenig verspürt: Eine Annäherung zwischen „Ontology-Sprache“ und „FBA-Sprache“ hat nach meinem Empfinden bei „uns“ noch nicht so recht stattgefunden. Ich empfand viele Diskussionen so, als „redete man in Randbezirken um den heißen Brei herum“!

Ich versuche in dieser Notiz **zu vermitteln** zwischen der **Informatiker-Sprache** (zur Beschreibung eines Konzepts **Ontology**) und der **Mathematiker-Sprache** (insbes. in der **Formalen Begriffsanalyse – FBA**). Die im Titel angegebene „Verfeinerung“ besteht darin, dass man bei Anwendung der FBA-Sprache **Strukturen** im Konzept „Ontology“ erkennt, über welche die mit „Ontology“ beschäftigten Informatiker sich m.E. bislang wohl noch nicht ganz einig sind.

In meiner Notiz unterscheide ich zwischen der **Struktur** eines Konzepts „Ontology“ und ihren vielfachen **Anwendungsmöglichkeiten** („Deutungen“). In Dieser Notiz geht es mir hauptsächlich um die **Struktur**. Deutungen ziehe ich nur in einfachen Beispielen als Anschauungsmaterial heran.

Das scheint für manche Ontologen erstaunlicherweise ein ungewohntes Vorgehen zu sein: *H. Herre / Leipzig, [19]*, stellt in seiner „*General Formal Ontology*“ die mathematischen, also **strukturbestimmenden** Hilfsmittel, auf gleiche Stufe mit ihren vielen Deutungsmöglichkei-

ten. Z.B. stellt er die **Mengennotation**, zusammen mit seinen ominösen „Urelementen“, auf die gleiche Stufe „*Entity*“, obwohl mit der Mengennotation auf sog. „*Entities*“ ja höchstens *in einer Deutung referiert* werden kann, wogegen „Urelemente“ etwas aus einer gewissen außermathematischen *Deutung* der Mengenlehre sind, die man in den Anfängen der Mengenlehre zu benötigen glaubte, die aber heute als obsolet und hinderlich für die Mengennotation gilt, weil sie den **sehr nützlichen** Gebrauch des Terminus „leere Menge“ behindert. Das ergibt für mein Empfinden in *Herre's* System ein **unerträgliches Durcheinander!**

Dies sage ich auch als **WARNUNG** für die Nicht-Mathematiker und die Nicht-Informatiker in unserem erlauchten Kreis: Es geht mir heute **nicht** um *schon wieder und noch einmal eine weitere* „philosophische Betrachtung“ zu dem, was man „Ontologie“ oder eines ihrer Teilaspekte nennen könnte, sondern es geht mir um einen **Vermittlungsversuch zwischen zwei Fachsprachen**, welche beide die Beschreibung sehr ähnlicher Informationsstrukturen zum Ziel haben.

1.1 Was ist für mich der Zweck eines Konzepts „Informatic Ontology“?

Sie implementiert ein **Wissensgebiet** (*knowledge domain*), so dass sie dem menschlichen User oder auch einem IT-System (das dazu in der Lage ist) etwas intelligentere Fragen als nur in Form von Stichworten à la Google verständlich und konsistent beantworten kann.

Eine **Ontology** besteht nach meiner Anschauung aus zwei Teilen:

- den „**Systeminterna**“ und
- der „**Userschnittstelle**“.

Auf die „Userschnittstelle“ (ich halte sie für den *anspruchsvolleren* Teil) will ich hier *nicht* eingehen; ihr Kern kann mit den Hilfsmitteln der **FBA** (Formale Begriffsanalyse) aber leicht dargestellt werden durch zwei „**Formale Kontexte**“: einen für die umgangssprachlichen Ausdrücke der „Objekte“ des Wissensgebiets, einen für die umgangssprachlichen Ausdrücke der „Beziehungen“ (zwischen Objekten) im Wissensgebiet; vgl. auch den etwas unvollständigen Ansatz von *Pickert* [2].

Die **Systeminterna** – auf die beschränke ich mich hier – bestehen aus zwei Ebenen: (A) dem **Ontology-Schema** (O-Schema) und (B) der **Instanziierung**.

Zunächst (Kap.2 - Kap.3) gehe ich auf die mir bekannte **Informatiker-Sprache** für die *Systeminterna* einer „**Ontology**“ ein. Zur metasprachlichen Ontology-Beschreibung benutze ich, falls angebracht, die sehr hilfreiche Notation von

Hermann Bense:

Präfix „[^]“ für den Namen einer „**Klasse**“,

Präfix „[.]“ für den Namen eines „**Attributs**“

Präfix „<>“ für den Namen eines „**Relationship-Type**“

Präfix „>“ für den Namen einer „**Instanz**“ oder eines „**Attributwertes**“.

Danach (ab Kap.4) führe ich die Grundtermini der **FBA (Formale Begriffsanalyse)** wiederholt bei einigen Anwendungsmöglichkeiten der FBA auf die Ontology-Sprache ein und zeige, dass die FBA-Sprache die diversen Gedanken der Ontology-Informatiker präzisieren kann. Damit aber sind die Anwendungsmöglichkeiten von FBA im Konzept „Ontology“ noch längst nicht ausgeschöpft.

1.2 Anschauungsbeispiel aus Wikipedia

Zur Einstimmung die **Abb.1** aus Wikipedia unter dem Strichwort „Ontologie (Informatik)“ [10].

Die obere Ebene: „**O-Schema**“ (in Abb.1 als „Ontologie“ bezeichnet).

Die untere Ebene: „**Instanziierung**“ (in Abb.1 als „Instanzen“ bezeichnet).

Im O-**Schema** der Abb.1 sehen wir:

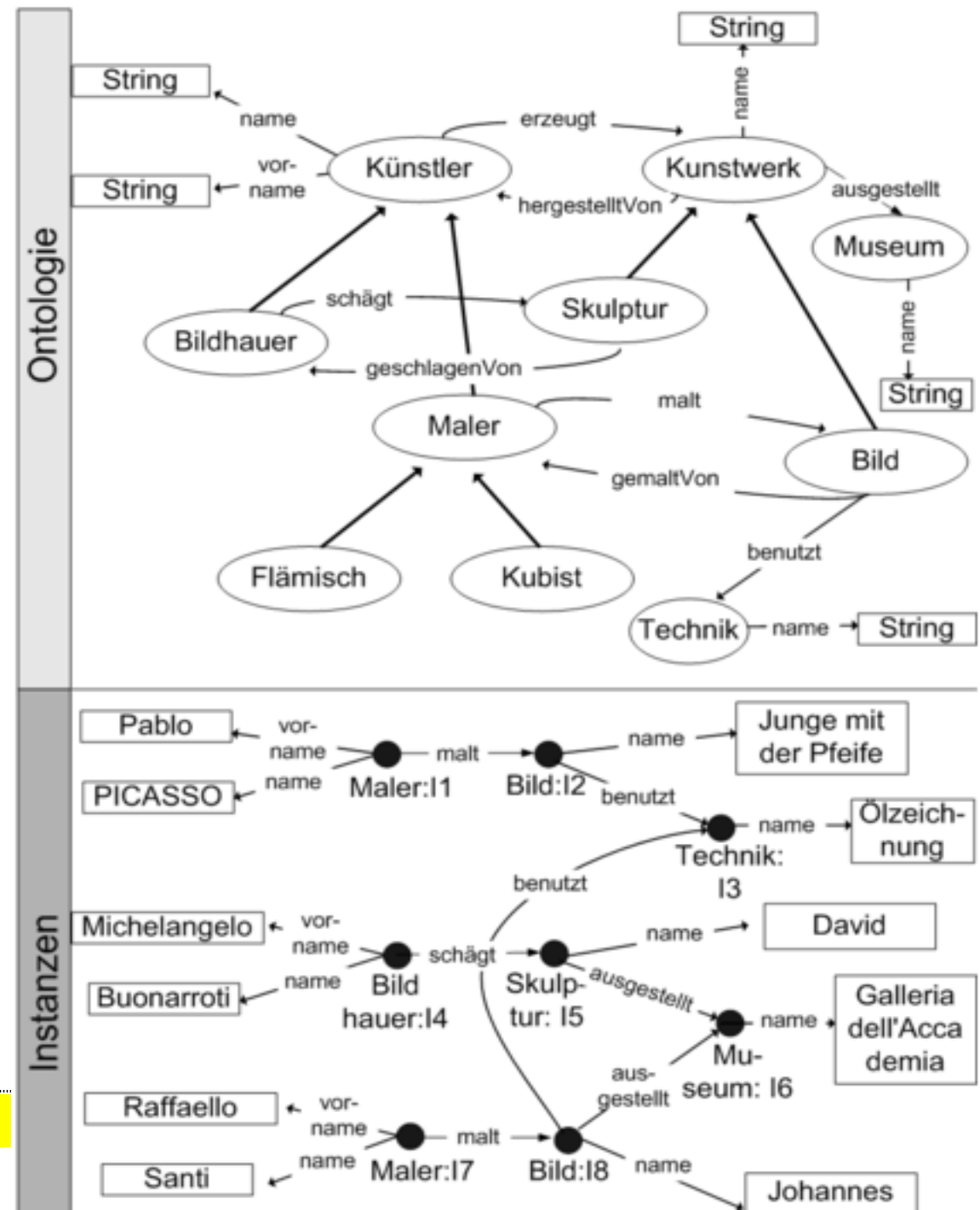
„**Klassen**“ als *Ovale* gezeichnet; „**Attribute**“, als *rechteckige Kästchen* gezeichnet, in denen das Wort „String“ steht. Klassen-Ovale sind mit Attribut-Kästchen durch dünnere gerade und u.a. mit „name“, „vorname“ beschriftete Linien verbunden.

Das soll heißen, dass den Klassen gewisse Attribute zugeordnet sind. „String“ bedeutet, dass bei *Instanziierung* in das betreffende Leerfeld ein Attribut**wert** (Merkmal einer gewissen Instanz) im Format „String“ einzutragen ist (es könnte bei Bedarf auch ein numerisches Format sein).

Einige Klassen sind verbunden durch dicke gerade Pfeile **von unten nach oben**. Zum Beispiel ist die Klasse \wedge Maler pfeilverbunden mit der Klasse \wedge Künstler. Das soll heißen, dass z.B. \wedge Maler eine *Unterklasse* von \wedge Künstler sein soll. Zugleich soll es heißen, dass die Oberklasse ihre **Attribute** an die Unterklassen **vererbt** (also von oben nach unten, *entgegen* der Pfeilrichtung). Damit wird auch die **O-
Schema-Konvention** klar, dass das Attribut-Kästchen einer Oberklasse nicht noch einmal bei ihren Unterklassen eingetragen werden muss, weil es ja nach unten vererbt wird. Die dicken Pfeile von unten nach oben visualisieren die sog. „**Klassentaxonomie**“, auch „**Klassenhierarchie**“ genannt. Mathematisch wird diese „dicke-Pfeile-Struktur“ schlicht eine „**Ordnung**“ genannt.

Schließlich sind einige Klassen paarweise durch geschwungene Pfeile verbunden. Zum Beispiel ist \wedge Künstler verbunden mit \wedge Kunstwerk, und der Pfeil ist beschriftet mit $\langle \rangle$ erzeugt. „Invers“ dazu ist auch ein Pfeil $\langle \rangle$ hergestelltVon zu sehen. Das soll einfach den allgemeinen Satz „Künstler erzeugen Kunstwerke“ (bzw. invers dazu: „Kunstwerke werden von Künstlern hergestellt“) wiedergeben. Die Elemente „ $\langle \rangle$ erzeugt“, „ $\langle \rangle$ hergestelltVon“ nenne ich (binäre) „**Relation-schipp-Types**“ (im Wikipedia-Text wird dafür das Wort „Relation“ verwendet, dieses Wort vermeide ich zunächst.)

Abb.1 (aus Wikipedia) →



Kritische Anmerkung zur O-Schema-Darstellung in Abb.1:

Die Abbildung aus Wikipedia ist meinem Empfinden nach ***mit Vorsicht zu genießen***.

a) Attributevererbung und Identifizierbarkeit bei Klassen: Die Konvention ist, dass ein von oben nach unten vererbtes Attribut nicht noch einmal bei der Unterklasse als Attribut-Kästchen angehängt werden muss. Z.B. **erbt** die Unterklasse \wedge Maler von ihrer Oberklasse \wedge Künstler die Attribute „name:....“ und „vorname: ...“; dasselbe erfolgt bei der Unterklasse \wedge Bildhauer. Diese Konvention wird in Abb.1 **richtig** wiedergegeben.

Aber: \wedge Maler hat, außer den geerbten, keine weiteren Attribute. In was unterscheidet sich die Klasse \wedge Maler von der Klasse \wedge Künstler denn nun systemintern? – Nur in der Beschriftung der Ovale? – Ist diese Beschriftung Teil der Systeminterna des O-Schemas?? Das bleibt im Wikipediabeispiel unklar! Ich meine: An das Oval \wedge Maler müsste mindestens ein weiteres **Attribut**-Kästchen (etwa mit dem Attributnamen „Malereigenschaft:....“ oder „MalerInfo:....“) angehängt werden, das von den ans Oval \wedge Künstler angehängten **verschieden** sein sollte. Andernfalls hätte \wedge Maler **genau dieselben** Attribute wie \wedge Künstler, und man sähe nicht den **Grund** ein, warum in diesem O-Schema das \wedge Maler-Oval formal überhaupt vom \wedge Künstler-Oval unterschieden werden sollte! Denn die Worte „Künstler“ *im Oval* für die Klasse \wedge Künstler und „Maler“ *im Oval* für die Klasse \wedge Maler sind ja nur Beschriftungen zur Lesbarkeit der graphischen Darstellung des O-Schemas. Soll damit aber eine für die Klasse **spezifische Information** gemeint sein, so ist ein dazu geeignetes **Attribut**-Kästchen anzulegen; **das ist ja der Sinn und Zweck von Attributen!**

b) In Abb.1 bleibt auch die Struktur **Relationship-Types** und deren **Vererbung** unklar. Die Gesamtheit der in Abb.1 benutzten Relationship-Types ist die Menge {<>erzeugt, <>hergestelltVon, <>ausgestellt, <>schlägt, <>geschlagenVon, <>malt, <>gemaltVon, <>benutzt}.

Aber: Die Relationship-Types <>malt, <>gemaltVon, <>schlägt, <>geschlagenVon haben einen anderen Charakter als die Relationship-Types <>erzeugt, <>hergestelltVon, <>ausgestellt und <>benutzt. Diesen Unterschied erläutern wir wieder an einem Beispiel: <>malt ist hier offensichtlich als eine *Spezialisierung*, als ein „**Unter-Relationship-Type**“ von <>erzeugt gemeint. In der Interpretation der gemeinten Objekte heißt das: „*Jeder Maler ist ein Künstler, jedes Bild ist ein Kunstwerk*“. Die instanziiierende Aussage „*Der Maler1 malt das Bild1*“ ist eine Spezialisierung der Aussage „*Der Künstler1 erzeugt das Kunstwerk1*“, nämlich: „*Der Künstler1 (welcher der Maler1 ist) erzeugt (also malt) das Kunstwerk1 (welches das Bild1 ist)*“.

Andererseits ist „<>erzeugt“ in der Abb.1 kein Unter-Relationship-Type eines allgemeineren Relationship-Types. <>erzeugt ist vielmehr ein, wie ich vorläufig sagen will, „**maximaler**“ Relationship-Type, von dem die Unter-Relationship-Types <>malt und <>schlägt – und einige andere in Abb.1 nicht genannte – durch **Vererbung abgeleitet** sind. Würde man die **Vererbung** entlang der **Klassentaxonomie** ernst nehmen und entsprechend der Konvention verfahren, dass vererbte Elemente *nicht* noch einmal bei der Unterklasse eingezeichnet werden müssen, so wäre die Erwäh-

nung der abgeleiteten Unter-Relationship-Types <>malt, <>schlägt im O-Schema der Abb.1 **überflüssig**. Wollte man andererseits **alle** möglichen Unter-Relationship-Types von <>erzeugt im O-Schema vermerken, so müsste man dies in Abb.1 bei **allen** zusammengehörigen Klassenpaaren unterhalb ^Künstler und ^Kunstwerk tun und zum Beispiel auch einen Unter-Relationship-Type zwischen ^Kubist und ^Bild – in das O-Schema mit aufnehmen. So aber, wie es in Abb.1 dargestellt ist, erscheint das System der Relationship-Types ziemlich **willkürlich und unsystematisch**.

Die Instanziierungsebene in Abb.1:

Die Beispiel-Eintragungen in der Instanziierungsebene (untere Ebene der Ontology) verstehen sich von selbst; ich gehe hier nicht besonders darauf ein.

Nach dieser Einführung am Beispiel der Abb.1 sind m.E. im Folgenden nun doch ein paar strukturerläuternde und präzisierende Worte zu sagen. Ich gehe aus von den Termini „**Klasse**“, „**Klassentaxonomie**“, „**Attribut**“, „**Relationship-Type**“ und „**Vererbung**“. Für die *Instanziierung* kommen noch die Termini „**Instanz**“ und „**Attributwert**“ hinzu.

Diese Termini seien im Folgenden etwas genauer untersucht, ohne dass ich mich in den *Jungle* der in der Informatik üblichen vielen Abkürzungen begeben will.

2 O-Schema

In einem O-Schema, wie ich es kenne, hat man:

(1) Ein geordnetes System $(\mathbb{C}, \leq_{\mathbb{C}})$ von "**Klassen**". (Manchmal sagt man statt „Klasse“ abstrakter auch „Begriff“ (engl.: „*concept*“); ich verwende hier das Wort „**Klasse**“, um nicht mit dem später einzuführenden FBA-Terminus „*formaler Begriff*“ durcheinander zu kommen).

Die Menge \mathbb{C} ist die Gesamtheit der **Klassen**, die im O-Schema auftreten. Sie ist geordnet durch eine sog. **Klassen-Taxonomie** (Klassenhierarchie); Mathematiker sagen einfach „**Ordnung**“ dazu.

DEF.: Eine binäre Beziehung \leq auf einer Menge M heißt eine (nicht-strikte) **Ordnung** genau dann, wenn sie die folgenden drei Eigenschaften hat:

Reflexivität: $x \leq x$ für alle $x \in M$

Antisymmetrie: wenn $x \leq y$ und $y \leq x$, dann $x=y$ für alle $x, y \in M$

Transitivität: wenn $x \leq y$ und $y \leq z$, dann $x \leq z$ für alle $x, y, z \in M$.

Die zugehörige strikte Form „ $<$ “, definiert als $x < y$:gdw. $x \leq y$ und $x \neq y$, ist irreflexiv und transitiv; daraus folgt, dass sie auch azyklisch ist. Je nach Bedarf verwende ich für „Ordnung“ einmal die nicht-strikte Form „ \leq “, ein andermal die strikte Form „ $<$ “.

Gilt $c_1 <_{\mathbb{C}} c_2$ für zwei Klassen (eines O-Schemas), so heißt c_1 eine (echte) **Unterklasse** von c_2 , bzw. c_2 eine (echte) **Oberklasse** von c_1 .

Die Klassentaxonomie $<_{\mathbb{C}}$ ist in Abb.1 **monohierarchisch** (d.h. jede Klasse hat

höchstens einen oberen Nachbarn). Allgemein darf eine Klassentaxonomie aber auch **polyhierarchisch** sein, d.h. eine Unterklasse darf mehrere obere Nachbarn haben. Das geordnete System $(\mathbb{C}, <_{\mathbb{C}})$ der Klassen ist sozusagen das **Grundgerüst** im O-Schema: An ihm wird alles andere angehängt und entlang der Klassentaxonomie nach unten vererbt.

(2) Ein System \mathbb{A} von "**Attributen**":

Jede Klasse $c \in \mathbb{C}$ ist zu beschreiben durch eine Menge **att**(c) von **Attributen**, $\text{att}(c) \subseteq \mathbb{A}$. Der Operator **att**: $\mathbb{C} \rightarrow \text{Pot}\mathbb{A}$ ordnet jeder Klasse c eindeutig ihre Attributmenge att(c) zu. Das Attribut**mengensystem** $[\mathbb{A}] := \{\text{att}(c) \subseteq \mathbb{A} \mid c \in \mathbb{C}\}$ wird später eine zum Klassensystem \mathbb{C} korrespondierende Rolle spielen.

Ein "Attribut" kann man sich vorstellen als eine Art "Formular".

Beispiele: .Familiename: ; .Vorname:; .geboren am:; .Geburtsort: sind z.B. Attribute einer Klasse \wedge Person. Bei Instanziierung wird in das Leerfeld „:.....“ ein gewisser "Wert" eingetragen. Man muss jedenfalls zwischen „Attribut“ und „Attributwert“ unterscheiden! Hier geht es zunächst nur um die *Attribute selbst*, *nicht* um ihre Werte.

(3) **Vererbung**: Attribute **vererben** sich von oben nach unten entlang der Klassentaxonomie $<_C$, d.h.: ist $c1 <_C c2$ und hat die Klasse $c2$ das Attribut $a2 \in \mathbb{A}$, so hat auch $c1$ das Attribut $a2$. Die Unterklasse ist „**spezieller**“ als die Oberklasse; $c1$ darf daher daneben auch noch ein Attribut $a1$ haben, das die Oberklasse $c2$ nicht hat. Die Attributmenge $\text{att}(c)$ einer Klasse c kann man daher aufteilen in
(2.0) $\text{att}(c) = \text{att}_{\text{vererbt}}(c) \cup \text{att}_{\text{neu bei } c}(c)$.

Im polyhierarchischen Fall dürfen zwei unvergleichbare Klassen auch eine gemeinsame Unterklasse c haben. Dann erbt c die Attributmengen **beider** Oberklassen. In diesem Fall muss c kein weiteres eigenes Attribut aufweisen, d.h. $\text{att}_{\text{neu bei } c}(c)$ darf $=\emptyset$ sein, denn c hat ja auch in diesem Fall **mehr** Attribute als jede einzelne ihrer Oberklassen. Wenn aber c genau **einen** oberen Nachbarn hat, sollte $\text{att}_{\text{neu bei } c}(c) \neq \emptyset$ sein; andernfalls wäre c *nicht spezieller* als ihr oberer Nachbar und somit von dem oberen Nachbarn formal **nicht unterscheidbar!** Es gibt m.E. **keinen Grund**, im O-Schema zwei Klassen c, d zu unterscheiden, wenn ihre Attributmengen $\text{att}(c), \text{att}(d)$ gleich sind. **Das ist ja der Sinn und Zweck von Attributen.** Wir können also formulieren:

(2.1) $\text{att}(c) = \text{att}(d)$ genau dann, wenn $c=d$.

(2.2) $c1 <_C c2$ genau dann, wenn $\text{att}(c2) \subset \text{att}(c1)$ $\text{att}(c2)$ echte Teilm.von $\text{att}(c1)$.

Anmerkung: Hier gibt es *Widerspruch* von Seiten einiger Informatiker. Sie meinen, die Regeln (2.1), (2.2) schränken die Freiheit des Ontology-Designers zu sehr ein. Sie wollen zulassen, dass eine **echte** Unterklasse c_1 von c_2 **die-selbe** Attributemenge wie die Oberklasse c_2 haben darf: $\text{att}(c_1) = \text{att}(c_2)$. Da frage ich mich aber: durch welche systeminternen Parameter unterscheiden sich dann die Klassen im O-Schema, und warum sind diese verdeckt??

- (4) "**Relationship-Types**": Das O-Schema hat meist auch ein System **RELT** von binären "**Relationship-Types**" auf dem System \mathbb{C} der Klassen.
- Ich wähle hier extra das Wort „Relationship-Type“ und **nicht** das Wort "Relation" (wie in Wikipedia), weil ich letzteren Terminus für „**Relation im mathematischen Sinne**“ reservieren will. Hierzu erst Genaueres, wenn wir uns mit der Instanziierung beschäftigen.
 - **Außerdem beschränke ich mich hier der Einfachheit halber auf nur binäre Relationship-Types**, – so verfahren auch *Meadche et.al.*[12] und Pickert [2],
 - obwohl man Ähnliches auch für mehr als 2-stellige Relationship-Types machen kann.
- Zu jedem binären Relationship-Type r gehören zwei Klassen c, d und damit eine **relationale Teilkonfiguration $c r d$** des O-Schemas. Wir haben also einen Operator $\text{relt}: \mathbf{RELT} \rightarrow \mathbb{C} \times \mathbb{C}$. Dabei kommt es auf die Reihenfolge an: $c r d$ ist nicht dasselbe wie $d r c$! In Abb.1 ist zum Beispiel „ $\wedge \text{Maler} \langle \rangle \text{malt} \wedge \text{Bild}$ “ eine solche O-Schema-Teilkonfiguration, aber $\wedge \text{Bild} \langle \rangle \text{gemaltVon} \wedge \text{Maler}$ ist davon

zu unterscheiden: es ist die „inverse“ Teilkonfiguration. Es darf zwischen *demselben* Klassenpaar auch zwei verschiedene Relationship-Types geben: $\text{rel}(r) = \text{rel}(s) = (c, d)$. Es darf aber auch eine Klasse *mit sich selbst* kombiniert werden: $\text{rel}(t) = (c, c)$; Beispiel:

$\text{^Maler} \langle \rangle \text{ist_befreundet_mit} \text{^Maler}$. (In Abb.1 nicht vorgesehen.)

Auch Relationship-Types können **vererbt** werden: ist z.B. $c1 <_C c$ und $d1 <_C d$, und besteht zwischen c, d der Relationship-Type r , also die Teilkonfiguration **$c \ r \ d$** , so ergibt sich daraus zwischen $c1, d1$ ein sog. „**Unter-Relationship-Type**“ $r1$, der von r **abgeleitet** ist durch **Einschränkung** auf die Unterklassen. (Genaueres erst in **Kap.3.4.**)

Wir haben dann die relationale Teilkonfiguration **$c1 \ r1 \ d1$** . Man sollte $r1$ aber **nicht** mit demselben Namen wie r bezeichnen, denn $r1$ ist ja eine **Spezialisierung** von r .

Beispiel (vgl. Abb.1): $\text{^Künstler} \langle \rangle \text{erzeugt} \text{^Kunstwerk}$; $\text{^Maler} \langle \rangle \text{malt} \text{^Bild}$: Hier ist $\langle \rangle \text{malt}$ ein echter Unter-Relationship-Type des Relationship-Types $\langle \rangle \text{erzeugt}$ und nicht mit ihm zu identifizieren, weil er nur zwischen Malern und deren Bildern gemeint ist, wogegen $\langle \rangle \text{erzeugt}$ allgemeiner zwischen Künstlern und deren Kunstwerken gemeint ist!

Ist $r1$ ein Unter-Relationship-Type von r , so schreiben wir: $r1 <_{\text{RELT}} r$.

Merkregel: Bei **Vererbung** entlang der Klassentaxonomie nach unten werden die Attributmengen $\text{att}(c)$ „größer“, die abgeleiteten Relationship-Types „kleiner“ (spezieller). Jedenfalls wird durch die **Vererbung** entlang der Klassentaxonomie \leq_C von

oben nach unten auch auf dem System **RELT** aller Relationship-Types eine **Ordnung** \leq_{RELT} induziert: $(\text{RELT}, \leq_{\text{RELT}})$ ist wieder ein **geordnetes** System. Man könnte die Ordnung \leq_{RELT} die „**Relationship-Taxonomie**“ nennen. Sie wird bei Informatikern nicht immer hervorgehoben, ist aber m.E. genauso relevant wie die Klassen-Taxonomie $\leq_{\mathbb{C}}$, sofern man überhaupt Relationship-Types in der Ontology berücksichtigt.

Mathematisch Genaueres zur Relationship-Taxonomie \leq_{RELT} lässt sich erst dann einfach ausdrücken, wenn wir die Instanziierung eingeführt haben. Daher sei auf **Kap.3.4** vertröstet.

3 Instanziierung

3.1 Instanziierung in Klassen

Die „**Instanziierung**“ ermittelt zu jeder Klasse $c \in \mathbb{C}$ die Menge $\underline{c} := \text{inst}(c)$ ihrer in der Ontology ermittelbaren **Ausprägungen** – „**Instanzen**“. Manche sagen statt „Instanz“ auch „Entity“, aber ich bleibe beim Wort „Instanz“, da das Wort „Entity“ in anderen Datenmodellierungen wie etwa ERM (Entity Relationship Modelling) benutzt wird, die ich hier nicht beschreibe.

Der Operator **inst**: $\mathbb{C} \rightarrow \text{Pot}\mathbf{G}$ ordnet jeder Klasse c eindeutig die Menge $\text{inst}(c)$ ihrer Ausprägungen (Instanzen) zu, wobei \mathbf{G} die Gesamtheit aller *in der Ontology ermittelbaren* Instanzen sei.

Ähnlich wie bei den Attributen gehe ich von folgenden Regeln aus:

(3.1) $c = d$ gdw. $\text{inst}(c) = \text{inst}(d)$, u. dies genau dann, wenn $\text{att}(c) = \text{att}(d)$.

(3.2) $c <_C d$ gdw. $\text{inst}(c) \subset \text{inst}(d)$, u. dies genau dann, wenn $\text{att}(d) \subset \text{att}(c)$.

Merkregel: Bei *Vererbung* nach unten entlang der Klassentaxonomie $<_C$ werden die Klassen c spezieller, die Instanzenmengen $\text{inst}(c)$ demnach *kleiner* und die Attributmengen $\text{att}(c)$ demnach *größer*.

Begründung: Ist c eine echte Unterklasse von d (also $c <_C d$), so sind alle Instanzen von c auch Instanzen von d : $\text{inst}(c) \subseteq \text{inst}(d)$. Wenn aber zwei Klassen c, d *dieselbe* Ausprägungsmenge hätten, $\text{inst}(c) = \text{inst}(d)$, so hätten sie auch dieselben Attribute, $\text{att}(c) = \text{att}(d)$; gemäß **(2.1)** wären sie also bei der Instanziierung *nicht unterscheidbar*, sollten also auch im O-Schema *identifiziert* werden: $c = d$!
Je „kleiner“ (im Sinne der Klassentaxonomie $<_C$) eine Klasse c ist, desto *spezieller* ist sie, und desto *größer* ist ihre Attributmenge $\text{att}(c)$ (vgl. **(2.2)**); je *spezieller* aber die Klasse ist, desto weniger Instanzen kann sie (in der gemeinten Ontology) haben.
Beispiel: „*Alle Maler sind Künstler, aber nicht alle Künstler sind Maler!*“

Anmerkung: Auch diese beiden Regeln werden – soviel ich gesehen habe – bei Darstellung einer Klassentaxonomie nicht immer berücksichtigt, und das kann zu Schwierigkeiten oder sogar zu Inkonsistenzen im O-Schema führen.

Anmerkung: Eine Instanz g kann zu mehreren Klassen gehören. Ist etwa c eine Unterklasse von d , so gehört eine Instanz der Klasse c , wie oben gesagt, auch zur Klasse d : $g \in \text{inst}(c)$ und $g \in \text{inst}(d)$. Sind

andererseits c, d unvergleichbar im Sinne der Klassentaxonomie \leq_C , so kann es trotzdem eine Instanz g geben mit $g \in \text{inst}(c)$ und $g \in \text{inst}(d)$, d.h. die *Ausprägungsmengen* unvergleichbarer Klassen müssen zueinander i.allg. nicht fremd sein (es kann $\text{inst}(c) \cap \text{inst}(d) \neq \emptyset$ sein)!

Beispiel aus Abb.1: Die Instanz `>Pablo_Picasso` ist der Klasse `^Maler` zugeordnet; da Picasso aber auch ein paar Skulpturen gemacht hat, könnte (sofern diese in die Ontology aufgenommen sind), `>Pablo_Picasso` auch als eine Instanz der Klasse `^Bildhauer` gelten.

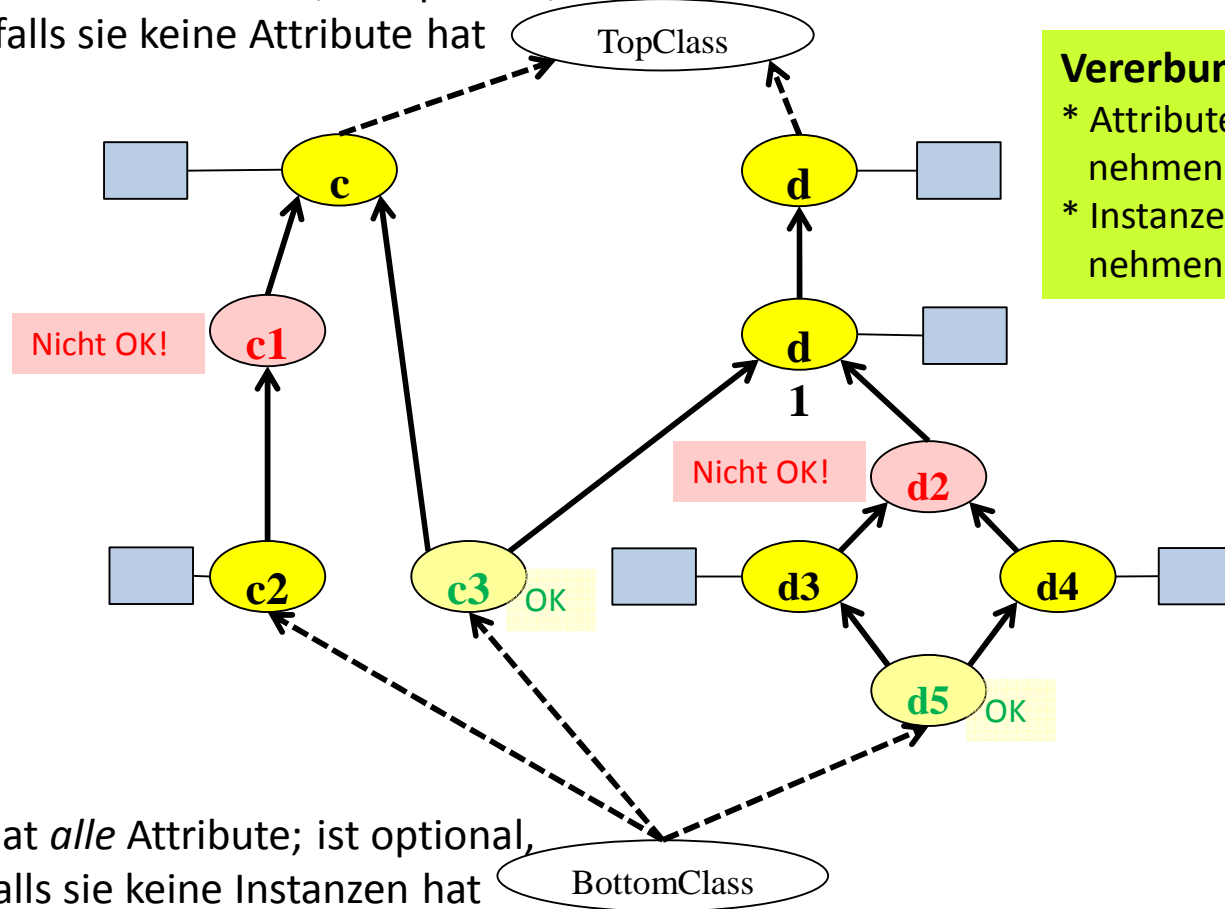
Anmerkung-#: Sind nun c, d zwei *unvergleichbare* Klassen mit $\text{inst}(c) \cap \text{inst}(d) \neq \emptyset$, und haben sie beide noch keine unteren Nachbarn, so besteht m.E. **ein Grund dazu**, einen gemeinsamen unteren Nachbarn e in die Taxonomie $<_C$ einzufügen mit
 $\text{inst}(e) = \text{inst}(c) \cap \text{inst}(d)$ und $\text{att}_{\text{vererbt}}(e) = \text{att}(c) \cup \text{att}(d)$.

– Auch diese Regel wird in einem O-Schema oft nicht beachtet. Das führt zu Unvollständigkeits der Klassentaxonomie $<_C$ bzw. der Gesamtheit C aller im O-Schema zu berücksichtigenden Klassen. Im Beispiel der Abb.1 aus Wikipedia kommt das aber nicht zum Tragen.

Abb.2 veranschaulicht noch einmal die Regeln (2.1), (2.2), (3.1) und (3.2), sowie die erwähnten Regelverletzungen, wobei Relationship-Types noch nicht berücksichtigt sind.

Abb.2: O-Schema ohne Relationship-Types

hat *alle* Instanzen; ist optional,
falls sie keine Attribute hat



Vererbung:

- * Attributmengen (von Klassen) nehmen von oben nach unten zu.
- * Instanzenmengen (von Klassen) nehmen von oben nach unten ab.

Klassen c1, d2: Nicht OK, da sie nur je einen oberen Nachbarn, aber keine eigenen Attribute haben.

Klassen c3, d5: OK. Sie haben zwar keine eigenen Attribute, erben aber von je zwei oberen Nachbarn.

3.2 Was ist die volle Menge *aller* Attribute, einer Instanz g ?

Alle Attribute einer Klasse c kommen auch allen Instanzen $g \in \text{inst}(c)$ zu.

Nehmen wir eine bestimmte Instanz $g \in \text{inst}(c)$ an; innerhalb der Klasse c werden der Instanz g – wie allen anderen Instanzen von c auch – nur die Attribute der Menge $\text{att}(c)$ zugeordnet. Das müssen aber nicht notwendig schon alle der Instanz g insgesamt zugeordneten Attribute sein! Hat nämlich c eine echte Unterklasse d ($d <_C c$), zu der g ebenfalls als Ausprägung gehört, so ist wegen (2.2) $\text{att}(d)$ größer als $\text{att}(c)$, also kommen der Instanz g mehr als nur die Attribute von c zu. Um *alle* der Instanz g zugeordneten Attribute zu bekommen, muss man die (im Sinne der Taxonomie $<_C$) **kleinste** Unterklasse unterhalb c wählen, die noch g als Instanz hat. Sei d_{\min} die **kleinste** Unterklasse von c mit $g \in \text{inst}(d_{\min})$. Dann ist $\text{att}(d_{\min})$ die **volle** Menge *aller* der Instanz g zugeordneten Attribute, und zu *jedem* $a \in \text{att}(d_{\min})$ bekommt die Instanz g je einen definierten und bestimmten **Attributwert** $a(g)$. Diese volle Attributmenge einer Instanz g will ich vorläufig mit „**att**[g]“ bezeichnen:

$$\mathbf{att}[g] := \{a \in \mathbb{A} \mid \text{es gibt } c \in \mathbb{C} \text{ mit } a \in \text{att}(c) \text{ und } g \in \text{inst}(c)\} \subseteq \mathbb{A}.$$

Anmerkung: Ist g eine bestimmte Instanz, so muss sie *nicht für jedes* beliebige Attribut $a \in \mathbb{A}$ überhaupt einen definierten Attributwert $a(g)$ haben. g hat nur **Attributwerte** für die Attribute der Menge $\text{att}[g]$, also der Attribute aller derjenigen

Mengen $\text{att}(c)$, für deren Klassen $g \in \text{inst}(c)$ gilt. – Gegenbeispiel (vgl. Abb.1): Ein Attribut der Klasse $\wedge \text{Museum}$ sei „gegründet_im_Jahr:“. Der Instanz >Pablo_Picasso wird man wohl nicht dieses Attribut zuschreiben; also gibt es auch keinen Wert dieses Attributs für >Pablo_Picasso .

3.3 Instanziierung in Relationship-Types

Beispiel: Sei $c1 := \wedge \text{Person}$, $i1 := \text{>Max} \in \text{inst}(c1)$, $c2 := \wedge \text{Wohnort}$, $i2 := \text{>Darmstadt} \in \text{inst}(c2)$, $r := \langle \rangle \text{wohnt_in}$ mit $\text{relt}(r) = (c1, c2)$. Dann lautet die relationale

(i) **Teilkonfiguration $c1$ r $c2$:** $\wedge \text{Person} \langle \rangle \text{wohnt_in} \wedge \text{Wohnort}$,
und eine Instanziierung ergibt z.B. die

(ii) **Aussage:** $\text{>Max} \circ \langle \rangle \text{wohnt_in} \text{>Darmstadt}$ („Max wohnt in Darmstadt“).

Darin unterscheide ich den **Relationship-Type** $\langle \rangle \text{wohnt_in}$ der Konfiguration (i) von der **Aussageform** $\circ \langle \rangle \text{wohnt_in}$ der **Aussage** (ii). Denn Relationship-Types sind Komponenten des **O-Schemas**, aber Aussagen sind Ausprägungen der **Instanziierung**. Den Terminus „Aussageform“ für $\circ \langle \rangle \text{wohnt_in}$ habe ich aus der Mathematik entlehnt. Er vermittelt sozusagen zwischen Relationship-Type und Aussage.

3.4 Vererbung von Relationship-Types entlang der Klassentaxonomie

Hier wollen wir die Ordnungsbeziehung \leq_{RELT} präzisieren. Besteht ein Relationship-Type r zwischen den Klassen c, d , also $\text{relt}(r) = (c, d)$, und ist $\circ r$ die zu r gehörige **Aussageform**, so ist

$$(3.1) \quad \underline{r} := \text{instR}(r) := \{(g, h) \in \mathbf{G} \times \mathbf{G} \mid g \circ r h\}$$

die Menge der **Ausprägungen** von r (also die Menge der Instanzenpaare (g, h) für welche die Aussage $g \circ r h$ gilt). \underline{r} ist eine „**Relation im mathematischen Sinne**“, also eine gewisse **Teilmenge** des Mengenprodukts $\mathbf{G} \times \mathbf{G}$, wobei \mathbf{G} die Gesamtheit der in der Ontology ermittelbaren **Instanzen** sei.

Sind nun c_1, d_1 irgendwelche Unterklassen von d, c (d.h. $d_1 \leq_c c, d_1 \leq_c d$), so wird durch r ein Relationship-Type r_1 zwischen c_1 und d_1 **induziert**, dessen Ausprägungsmenge die *Teilrelation*

$$(3.2) \quad \underline{r_1} := \text{instR}(r_1) = \underline{r} \cap (\text{inst}(c_1) \times \text{inst}(d_1))$$

ist. Für sie gilt dann natürlich:

$$(3.3) \quad \underline{r_1} = \text{instR}(r_1) \subseteq \underline{r} = \text{instR}(r) \quad \underline{r_1} \text{ Teilmenge von } \underline{r}.$$

Wir können sagen: Der Relationship-Type r wird entlang der Klassentaxonomie \leq_C nach unten durch *Einschränkung* auf die Unterklassen c_1, d_1 **vererbt** auf einen Unter-Relationship-Type r_1 .

Vgl. das Beispiel in Abb.1: $\langle \rangle_{\text{malt}}$ in der Teilkonfiguration „ \wedge Maler $\langle \rangle_{\text{malt}} \wedge$ Bild“ ist ein Unter-Relationship-Type von \wedge erzeugt, welcher in der darüber liegenden Teilkonfiguration „ \wedge Künstler $\langle \rangle_{\text{erzeugt}} \wedge$ Kunstwerk“ besteht.

Die Mengeninklusion in **(3.3) definiert** geradezu die Ordnungsbeziehung $r_1 \leq_{\text{REL T}} r$ zwischen Relationship-Types: **Dadurch wird das System $(\text{REL T}, \leq_{\text{REL T}})$ zu einer geordneten Menge.** Die strikte Form $<_{\text{REL T}}$ nenne ich die **Relationship-Taxonomie** im O-Schema. Sie ist m.E. **genau so relevant** wie die Klassen-Taxonomie $<_C$, sofern Relationship-Types im O-Schema auftreten.

Def.: Ein Relationship-Type r zwischen zwei Klassen c, d heiÙe „**maximal**“, wenn er **keinen** oberen Nachbarn (im Sinne der Ordnung $<_{\text{REL T}}$) hat, d.h. wenn er nicht von oben her (im Sinne der Klassentaxonomie $<_C$) vererbt ist.

Zwei verschiedene *maximale* Relationship-Types r, s sind folglich stets *unvergleichbar* im Sinne der Relationship-Taxonomie $\leq_{\text{REL T}}$, d.h. keiner ist ein Unter-Relationship-Type des anderen. Die **maximalen** Relationship-Types sind sozusagen die „**wichtigen**“ oder „**relevanten**“ im O-Schema, denn alle anderen sind Unter-Relationship-Types von ihnen und können durch Vererbung aus ihnen abgeleitet werden. – Zum Stichwort „relevante Relation“ vgl. auch Lübbert [18].

Die Gesamtheit **RELT** aller Relationship-Types des O-Schemas besteht also aus zwei Teilen:

$$(3.4) \quad \mathbf{RELT} = \mathbf{RELT}_{\max} \cup \mathbf{RELT}_{\text{vererbt}}$$

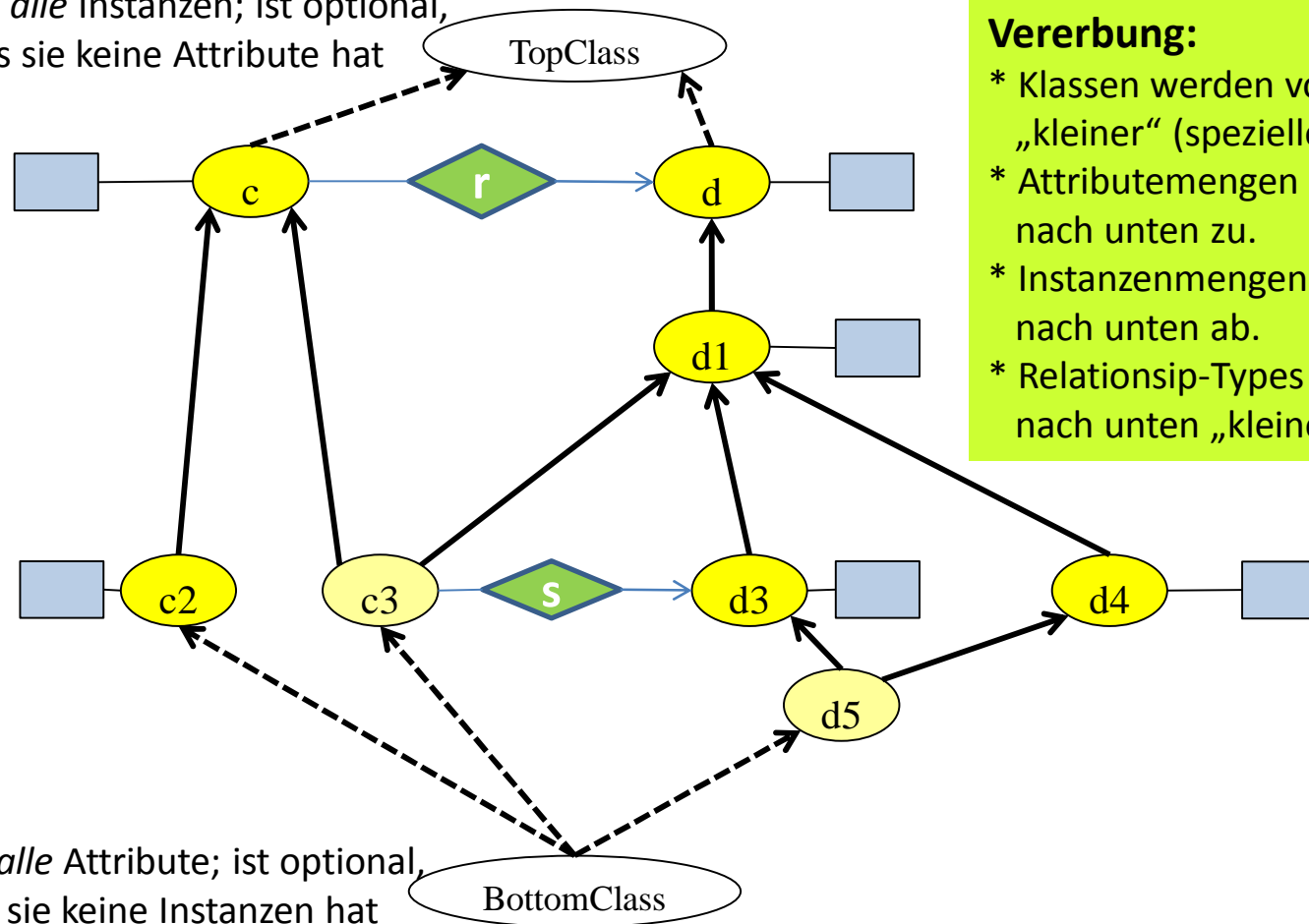
Vorschlag: Aus Gründen der Übersichtlichkeit würde ich in der graphischen Darstellung eines O-Schemas nur die *maximalen* Relationship-Types einzeichnen und ihre eventuell sehr vielen Vererbungen in der Graphik weglassen. Ganz analog macht man das ja auch bei den Attributen. In der **Abb.1** wäre der Eintrag des Relationship-Types $\langle \rangle_{\text{malt}}$ (zwischen \wedge_{Maler} und \wedge_{Bild}) dann *überflüssig*.

Dazu geben wir noch einmal eine Veranschaulichung in **Abb.3**. Dieses Bild ist entstanden aus Abb.2, indem wir die „Fehler“ der Abb.2 beseitigt haben, und danach zwei verschiedene *maximale* Relationship-Types $\langle \rangle_r$, $\langle \rangle_s$ eingeführt haben mit den relationalen Teilkonfigurationen

$$\wedge_c \langle \rangle_r \wedge_d \quad \text{und} \quad \wedge_{c3} \langle \rangle_s \wedge_{d3}.$$

Abb.3: O-Schema
mit zwei **maximalen(!)** Relationship-Types $\langle \rangle r$, $\langle \rangle s$

hat *alle* Instanzen; ist optional,
falls sie keine Attribute hat



Vererbung:

- * Klassen werden von oben nach unten „kleiner“ (spezieller).
- * Attributmengen nehmen von oben nach unten zu.
- * Instanzenmengen nehmen von oben nach unten ab.
- * Relationship-Types werden von oben nach unten „kleiner“ (spezieller).

hat *alle* Attribute; ist optional,
falls sie keine Instanzen hat

Beachte noch einmal:

(a) Gemäß unserem Vorschlag sind weder die bei Vererbung entstehenden Unter-Relationship-Types von $\langle \rangle r$ noch die von $\langle \rangle s$ im O-Schema der **Abb.3** eingezeichnet.

(b) Da $\langle \rangle r$ und $\langle \rangle s$ beide *maximal* (und damit in der Ordnung \leq_{RELT} unvergleichbar) sein sollen ist $\langle \rangle s$ **nicht** etwa ein Unter-Relationship-Type von $\langle \rangle r$; vielmehr bestehen zwischen **denselben** beiden Klassen c_3, d_3 zwei **verschiedene** (eventuell sogar unvergleichbare) Relationship-Types: der eine ist von $\langle \rangle r$ vererbt, der andere ist $\langle \rangle s$.

3.5 Instanziierung in Attributen

Zu jedem Attribut $a \in \mathbb{A}$ gehört die Menge $\underline{a} := \text{val}(a)$ seiner **Attributwerte** – statt „Attributwert“ sagen einige auch "*literal value*", vgl. *Maedche* [12]. Die Termini „Attribut“ und „Attributwert“ muss man unterscheiden – ganz analog dazu, dass man „Klasse“ von „Instanz“ unterscheidet!

Ist g eine Instanz der Klasse c , $g \in \text{inst}(c)$, und a ein Attribut der Klasse c , $a \in \text{att}(c)$, so kommt g ein eindeutig bestimmter **Attributwert** von a zu, den ich mit $a(g)$ bezeichne. Es ist also $a(g) \in \underline{a} = \text{val}(a)$.

3.6 Was in den mir bekannten O-Schema-Definitionen auffällt

- (1) Klassen werden bei Informatikern *anders* behandelt als ihre Attribute. – Was ist der Grund für diese Asymmetrie? Insbesondere scheint das System \mathbb{A} der Attribute keine Ordnung zu tragen, wie das im taxonomisch geordneten Klassensystem $(\mathbb{C}, <_{\mathbb{C}})$ der Fall ist.
 - (2) Die Regeln (2.1), (2.2) und (3.1), (3.2) werden von manchen Informatikern als zu restriktiv angesehen und daher nicht immer eingehalten.
- ⇒ Die in (1) erwähnte Asymmetrie löst sich auf, und die in (2) erwähnten Regeln rechtfertigen sich, wenn man in die Sprache der **FBA (Formale Begriffsanalyse)** übersetzt. Das ist Thema des nächsten Kapitels.

4 Bezug zur FBA (Formale Begriffsanalyse)

In Kap.4 und Kap.5 führe ich die Grundkonzepte der **Formalen Begriffsanalyse (FBA)**, vgl. *Ganter / Wille* [1], ein – nicht mathematisch-abstrakt wie in [1], sondern *wiederholt* an ein paar Anwendungsbeispielen im Bereich des bisher beschriebenen Konzepts „Ontology“. Durch die Wiederholungen hoffe ich, dass die Grundkonzepte der FBA auch für Informatiker eingängiger werden.

4.1 Formaler Kontext für die Beziehung zwischen Instanzen und Attributen

Betrachten wir vom O-Schema zunächst nur das Stück

$$\mathbf{O}_{\text{Tax}} := (\mathbb{C}, <_{\mathbb{C}}, \mathbb{A}, \text{VEB})$$

und lassen das System **RELT** der Relationship-Types noch weg. Der Vererbungsmechanismus VEB bezieht sich dann nur auf Attribute. Wir lassen hier auch die bei Instanziierung notwendige Zuordnung von Attribut**werten** zu Instanzen noch weg, sondern beachten nur, wann bei gegebener Instanz g und gegebenem Attribut $a \in \mathbb{A}$ ein Attributwert $a(g)$ überhaupt definierbar ist und wann nicht.

In der **FBA** hat man für jene ganze – bisher im Kap.2 + 3 recht umständliche – Beschreibung des Stücks \mathbf{O}_{Tax} eine etwas kompaktere Notation, die ich nun

vorstelle, und die den Zusammenhang zwischen Klassen (bzw. deren Instanzen) und Attributen **und deren Vererbung** etwas präziser und vor allem „**symmetrischer**“ darstellt, als es in der Ontology-Sprache zurzeit üblich ist:

Sei **G** die **Gesamtheit aller Instanzen**, die aus Klassen des O-Schemas bei der gewählten Instanziierung **inst** gewonnen werden können, und sei **A** wieder die **Gesamtheit aller Attribute**.

I sei die **Aussageform**, die angibt, ob einer Instanz $g \in \mathbf{G}$ das Attribut $a \in \mathbf{A}$ überhaupt zukommt. Für die Aussage „ $g \mathbf{I} a$ “ wollen auch wir sagen: „die Instanz g hat das Attribut a “. Falls $g \mathbf{I} a$ zutrifft, ist der Attribut**wert** $a(g)$ eindeutig definiert und ermittelbar; der Wert $a(g)$ selbst kommt hier aber in diesem FBA-Anwendungsschritt **noch gar nicht ins Spiel**.

Im Jargon der Mathematiker gehört nun zur Aussageform **I** die "**Relation**" **I** im *mathematischen* Sinne: das ist die Menge

$$\mathbf{I} := \{(x, y) \in \mathbf{G} \times \mathbf{A} \mid x \mathbf{I} y\};$$

es ist eine gewisse **Teilmenge** des Mengenprodukts $\mathbf{G} \times \mathbf{A}$. Die Aussage " $g \mathbf{I} a$ " kann nun auch in der Form „ $(g, a) \in \mathbf{I}$ “ ausgedrückt werden.

Mit den Mengen **G**, **A** und **I** bekommen wir in der **FBA** den sogenannten **formalen Kontext**

$$(4.1) \quad K(\mathbf{I}) := (\mathbf{G}, \mathbf{A}, \mathbf{I})$$

In der FBA-Sprache nennt man nun die Elemente von \mathbf{G} (also die Instanzen) "**Ge-
genstände**", die Elemente von \mathbf{A} (also die Attribute) "**Merkmale**". Die **Inzidenz-
relation** \mathbf{I} induziert zwei Abbildungen $\uparrow\mathbf{I}: \text{Pot}\mathbf{G} \rightarrow \text{Pot}\mathbf{A}$ und $\downarrow\mathbf{I}: \text{Pot}\mathbf{A} \rightarrow \text{Pot}\mathbf{G}$

("Galois-Verbindung zwischen \mathbf{G} und \mathbf{A} "), definiert für beliebige Teilmengen $X \subseteq \mathbf{G}$, $Y \subseteq \mathbf{A}$ durch

$$(4.2) \quad X^{\uparrow\mathbf{I}} := \{a \in \mathbf{A} \mid \text{für alle } g \in X: g \mathbf{I} a\} = \bigcap_{g \in X} \{g\}^{\uparrow\mathbf{I}}, \text{ mit } \{g\}^{\uparrow\mathbf{I}} := \{a \in \mathbf{A} \mid g \mathbf{I} a\} = \text{att}[g],$$

$$(4.3) \quad Y^{\downarrow\mathbf{I}} := \{g \in \mathbf{G} \mid \text{für alle } a \in Y: g \mathbf{I} a\} = \bigcap_{a \in Y} \{a\}^{\downarrow\mathbf{I}}, \text{ mit } \{a\}^{\downarrow\mathbf{I}} := \{g \in \mathbf{G} \mid g \mathbf{I} a\}.$$

In Worten: $X^{\uparrow\mathbf{I}}$ ist die Menge der Attribute (Merkmale), die *allen* Instanzen (Gegenständen) von X gemeinsam sind; und $Y^{\downarrow\mathbf{I}}$ ist die Menge der Instanzen (Gegenstände), die *allen* Attributen (Merkmalen) von Y gemeinsam sind.

$\{g\}^{\uparrow\mathbf{I}}$ ist gerade die volle Attributemenge $\text{att}[g]$ der Instanz g . Symmetrisch (oder „dual“) dazu ist $\{a\}^{\downarrow\mathbf{I}}$ die Menge aller Instanzen $g \in \mathbf{G}$, die das Attribut a haben.

DEF.: Ein sog. „**formaler Begriff**“ zum Kontext $K(\mathbf{I})$ ist ein **Mengenpaar** (A, B) mit $A \subseteq \mathbf{G}$, $B \subseteq \mathbf{A}$, das den beiden Bedingungen

$$(4.4) \quad A^{\uparrow\mathbf{I}} = B, B^{\downarrow\mathbf{I}} = A$$

genügt. A heißt der **Umfang** ("Extension"), B der **Inhalt** ("Intension") des formalen Begriffs.

Zum formalen Kontext $K(\mathbb{I})$ gehört die Menge $\underline{\mathbb{B}}(\mathbb{I})$ all dieser formalen Begriffe; sie trägt in natürlicher Weise eine **Ordnung**, deren strikte Form mit $<_i$ und deren nicht-strikte Form mit \leq_i bezeichnet sei und so definiert ist: Sind (A, B) und (A', B') formale Begriffe in $\underline{\mathbb{B}}(\mathbb{I})$, so definiert man

(4.5) $(A, B) \leq_i (A', B') : \text{gdw. } A \subseteq A' , \text{ und das ist gleichwertig zu } B' \subseteq B.$

Gilt $(A, B) <_i (A', B')$, so heißt (A, B) ein formaler **Unterbegriff** von (A', B') bzw. (A', B') ein formaler **Oberbegriff** von. Die Ordnung \leq_i besagt mit (4.5): **je größer der Umfang, desto kleiner der Inhalt**, und je kleiner der Umfang, desto größer der Inhalt – ganz im Sinne der intuitiven Vorstellung von der "Extension" und der "Intension" eines "Begriffs".

Die geordnete Menge $(\underline{\mathbf{B}}(\mathbf{I}), \leq_{\mathbf{I}})$ heißt der **Begriffsverband** zum Kontext $\mathbf{K}(\mathbf{I})$; sie hat eine **sehr schöne** FBA-Struktur: nämlich die eines sog. **vollständigen Verbandes**¹.

Fassen wir alle Umfänge von $\underline{\mathbf{B}}(\mathbf{I})$ in einer Menge $\mathbf{U}(\mathbf{I})$ und alle Inhalte in einer Menge $\mathbf{J}(\mathbf{I})$ zusammen, so ist sowohl $\mathbf{U}(\mathbf{I})$ als auch $\mathbf{J}(\mathbf{I})$ durch die Teilmengeninklusion \subseteq geordnet, und die geordneten Mengen $(\mathbf{U}(\mathbf{I}), \subseteq)$, $(\mathbf{J}(\mathbf{I}), \subseteq)$ sind sogar zuein-

¹ DEF.: Ein „**vollständiger Verband**“ ist eine geordnete Menge $(\underline{\mathbf{B}}, \leq)$, in der für **jede** Teilmenge $X \subseteq \underline{\mathbf{B}}$ das „Supremum“ $\sup X := \text{Min}\{y \in \underline{\mathbf{B}} \mid x \leq y \text{ für alle } x \in X\}$ = „kleinste obere Schranke von X “ und das „Infimum“ $\inf X := \text{Max}\{y \in \underline{\mathbf{B}} \mid y \leq x \text{ für alle } x \in X\}$ = „größte untere Schranke von X “ existieren und Elemente der Menge $\underline{\mathbf{B}}$ sind $(\sup X \in \underline{\mathbf{B}}, \inf X \in \underline{\mathbf{B}})$ (Das ist allgemein *nicht für jede* Ordnung der Fall!). Insbesondere gilt: $\inf \underline{\mathbf{B}} = \sup \emptyset =$ *kleinstes* Element von $\underline{\mathbf{B}}$, $\sup \underline{\mathbf{B}} = \inf \emptyset =$ *größtes* Element von $\underline{\mathbf{B}}$. (Beachte, dass, nach mathematischer Konvention, die leere Menge \emptyset ja Teilmenge einer beliebigen Menge ist!). Hiermit kann man auf dem Verband $(\underline{\mathbf{B}}, \leq)$ zwei binäre Verknüpfungen \wedge, \vee einführen durch $x \wedge y := \inf\{x, y\}$, $x \vee y := \sup\{x, y\}$ für beliebige $x, y \in \underline{\mathbf{B}}$. Sie haben für alle $x, y, z \in \underline{\mathbf{B}}$ folgende Eigenschaften:

(K)	Kommutativität:	$x \vee y = y \vee x,$	$x \wedge y = y \wedge x$
(As)	Assoziativität:	$x \vee (y \vee z) = (x \vee y) \vee z,$	$x \wedge (y \wedge z) = (x \wedge y) \wedge z$
(Ap)	Absorption:	$x \vee (x \wedge y) = x \wedge (x \vee y) = x$	
(Id)	Idempotenz:	$x \vee x = x \wedge x = x$	
(Vt)	\leq -Verträglichkeit:	$x \wedge y \leq x \leq x \vee y,$	$x \wedge y \leq y \leq x \vee y.$

anderer duale **vollständige Verbände**, geben also beide *zusammen* die Struktur des Begriffsverbandes $(\underline{\mathbf{B}}(\mathbf{I}), \leq_I)$ wieder.

Zurück zur Informatiker-Sprache. Setzen wir im Konzept „Ontology“ die Regeln **(2.1)**, **(2.2)**, **(3.1)**, **(3.2)** und die **Anmerkung-#** in Kap.3.1 voraus, so ersieht man:

- Jeder Umfang $A \in \mathbf{U}(\mathbf{I})$ ist gerade die Ausprägungsmenge $\text{inst}(c)$ der Instanzen einer Klasse $c \in \mathbf{C}$.
- Jeder Inhalt $B \in \mathbf{J}(\mathbf{I})$ ist gerade die Menge $\text{att}(c)$ der Attribute einer Klasse c .
- Damit ist das taxonomisch geordnete Klassensystem (\mathbf{C}, \leq_C) dargestellt durch den **vollständigen Verband** $(\mathbf{U}(\mathbf{I}), \subseteq)$.
- Das geordnete System $([\mathbf{A}], \leq_A)$ der Attributmengen wird durch den **vollständigen Verband** $(\mathbf{J}(\mathbf{I}), \subseteq)$ dargestellt. Das System $[\mathbf{A}]$ trägt also eine **Ordnung** \leq_A , die ebenfalls durch die Mengeninklusion ausgedrückt wird. Die Ordnung \leq_A ist wegen (4.5) „dual“ zur Klassentaxonomie \leq_C . Man könnte \leq_A als die **„Attribute-Taxonomie“** bezeichnen.

Bemerkenswert ist also, dass in einem O-Schema, das den Regeln (2.1), (2.1), (3.1), (3.2) und der **Anmerkung-#** in Kap.3.1 gehorcht, das taxonomisch geordnete Klassensystem (\mathbf{C}, \leq_C) dual-isomorph ist zum geordneten Attributmengen-

system $([\mathbb{A}], \leq_A)$, und dass das Stück $\mathbf{O}_{\text{Tax}} := (\mathbb{C}, \leq_C, \mathbb{A}, \text{VEB})$ des O-Schemas durch den formalen Kontext

$K(\mathbf{I}) = (\mathbf{G}, \mathbb{A}, \mathbf{I})$ bzw. durch den formalen Begriffsverband

$(\underline{\mathbf{B}}(\mathbf{I}), \leq_I)$ bzw. durch die beiden zueinander "dualen" vollständigen Verbände

$(\mathbf{U}(\mathbf{I}), \subseteq) \sim (\mathbb{C}, \leq_C)$, $(\mathbf{J}(\mathbf{I}), \subseteq) \sim ([\mathbb{A}], \leq_A)$

dargestellt werden kann. Wir haben also für das Stück \mathbf{O}_{Tax} des O-Schemas folgenden Zusammenhang zwischen Informatiker- und Mathematiker-Sprache:

- Jede im O-Schema auftretende Klasse $c \in \mathbb{C}$ wird in der FBA dargestellt durch den **Umfang** A eines formalen Begriffs $(A, B) \in \underline{\mathbf{B}}(\mathbf{I})$.
- Jede zu einer Klasse $c \in \mathbb{C}$ gehörige Attributemenge $\text{att}(c) \subseteq \mathbb{A}$ wird in der FBA dargestellt durch den **Inhalt** B eines formalen Begriffs $(A, B) \in \underline{\mathbf{B}}(\mathbf{I})$.
- Jedes zusammengehörige Paar $(c, \text{att}(c))$, bestehend aus einer Klasse $c \in \mathbb{C}$ und ihrer Attributemenge $\text{att}(c) \subseteq \mathbb{A}$, wird in der FBA dargestellt durch einen **formalen Begriff** $(A, B) \in \underline{\mathbf{B}}(\mathbf{I})$ zum formalen Kontext $K(\mathbf{I}) = (\mathbf{G}, \mathbb{A}, \mathbf{I})$.
- Die **Vererbung von Attributen** („Merkmalen“) drückt sich in der FBA aus als Vererbung **von oben nach unten** im Sinne der Begriffsverbandsordnung \leq_I . „Dual“ dazu entspricht dem eine **Vererbung der Instanzen** („Gegenstände“) **von unten nach oben** im Sinne der Begriffsverbandsordnung \leq_I . Das ist eine sehr schöne

„Symmetrie“, zwischen den Instanzen („Gegenständen“) und den Attributen („Merkmalen“), die in der üblichen Ontology-Definition nicht so klar zum Ausdruck kommt wie in der FBA.

- Das gesamte durch die Klassentaxonomie geordnete System $(\mathbb{C}, \leq_{\mathbb{C}})$ der Klassen wird in der FBA dargestellt durch den **Umfangsverband** $(\mathbb{U}(\mathbb{I}), \subseteq)$.
- „Dual“ dazu: Das gesamte – ebenfalls geordnete – *Attributemengensystem* $([\mathbb{A}], \leq_{\mathbb{A}})$ wird in der FBA dargestellt durch den **Inhaltsverband** $(\mathbb{J}(\mathbb{I}), \subseteq)$.

Der größte formale Begriff ist $(\mathbf{G}, \mathbf{G}^{\uparrow})$, der kleinste ist $(\mathbf{A}^{\downarrow}, \mathbf{A})$ (wobei in Anwendungen oft $\mathbf{G}^{\uparrow} = \emptyset$ bzw. oft $\mathbf{A}^{\downarrow} = \emptyset$ ausfällt). Daher: Die größte *Klasse* (oft als „**TopClass**“ bezeichnet und oft weglassen, sofern $\mathbf{G}^{\uparrow} = \emptyset$ ist) wird durch die Menge \mathbf{G} *aller* in der Ontology ermittelbaren Instanzen dargestellt; es gibt aber auch die kleinste *Klasse* (oft als „**BottomClass**“ bezeichnet und oft weglassen, sofern $\mathbf{A}^{\downarrow} = \emptyset$ ist); sie wird durch die Menge \mathbf{A}^{\downarrow} dargestellt, das ist die Menge der Instanzen denen *alle* Attribute von \mathbf{A} zukommen. Die größte *Attributemenge* ist die Menge \mathbf{A} ; es gibt aber auch die kleinste Attributemenge; sie wird durch die Menge \mathbf{G}^{\uparrow} dargestellt, das ist die Menge der Attribute, die *allen* Instanzen zukommt.

Es gibt also eine durchgehende „Symmetrie“ zwischen Klassen und ihren Attributemengen. In der FBA nennt man das die „Dualität“

(zwischen „Gegenständen“ und „Merkmalen“ im Kontext oder auch die „*Dualität*“ zwischen Umfängen und Inhalten im Begriffsverband).

4.2 Eine philosophisch-historische Anmerkung

Es wäre nach meiner Ansicht also auch in einer „Ontology“ besser, „Begriffe“ nicht durch die Klassen c sondern durch die ***Paare*** ($c, \text{att}(c)$) darzustellen! Leider hat sich aber bei Informatikern einseitig und asymmetrisch die „Klasse“ als Begriffsdarstellung eingebürgert, wogegen die Attribute „nur so nebenher vererbt“ werden; und das gibt unnötige Streitereien.

Diese asymmetrische Sichtweise ist uralte: Solange es „philosophische Ontologie“ gibt, hat man in erster Linie auf die „Objekte“/ „Dinge“ gestarrt und ihre „Eigenschaften“ erst in zweiter Linie ernstgenommen. Später hat man sogar angefangen, über die „Dinge an sich“ (also „ohne Eigenschaften“) zu spekulieren. (Erst *David Hume* und *Immanuel Kant* haben mit diesem Unsinn aufgeräumt). Ich nenne das scherzhaft die „***platonisch-mittelalterliche Backsteinmentalität***“. Diese hat natürlich mit der *noch viel älteren* menschlichen Erfahrung zu tun, dass man einen Backstein in die Hand nehmen und sehen bzw. spüren kann, dass er rot / schwer / fest / kalt / usw... ist. Aber seine „Röte“ / sein „Gewicht“ / seine „Festigkeit“ / seine „Kälte“ / etc... kann man nicht in die Hand nehmen, nicht sehen, nicht spüren, denn es sind ja nur nützliche ***Abstraktionen***, die erst im menschlichen Gehirn erzeugt

werden und uns Menschen recht gut unterstützen, **planend** mit unserer Umwelt zurechtzukommen. Beeinflusst aber durch jene „*Backsteinmentalität*“ kamen alte Philosophen – wie etwa der antike **Platon**, und später besonders einige mittelalterlich-christlichen Scholastiker – auf die verrückte Idee, auch „die Röte“ / „das Gewicht“ / „die Festigkeit“ / „die Kälte“ / etc. ... ebenfalls wie „höhere“, evtl. sogar „ewige“ Backsteine zu behandeln. Daraus entstand u.a. die absurde mittelalterliche Streitfrage, ob die sogenannten **Universalialia** (Allgemeinbegriffe wie z.B. „die Röte“, „das Gewicht“, etc. ...) unabhängig von der Wahrnehmung ihrer Ausprägungen „existieren“ oder nicht „existieren“. (Das **Absurde** dabei waren eigentlich weniger jene Universalialia, sondern der **völlig unreflektierte** Gebrauch des Wortes „existieren“! Daran hat ja seit eh und je die sog. „philosophische Ontologie“ gekrank.)

Man sieht: **Das derzeitige informatische, asymmetrische Konzept „Ontology“ ist durch eine lange, nicht besonders originelle Vorgeschichte belastet.**

4.3 Berücksichtigung der Attributwerte durch Bildung eines mehrwertigen Kontextes

Die Inzidenzaussage „ $g \mathbf{I} a$ “ im formalen Kontext $K(\mathbf{I}) = (\mathbf{G}, \mathbf{A}, \mathbf{I})$ besagt nur, dass der Instanz $g \in \mathbf{G}$ ein Attribut $a \in \mathbf{A}$ zugeordnet ist, so dass $a(g)$ ein definierter Attributwert ist. Der Attributwert $a(g)$ selbst ist damit noch nicht ermittelt. $K(\mathbf{I})$ ist sozusagen ein „einfacher“ formaler Kontext; seine Inzidenzrelation ist **binär**. Will man im Stück $\mathbf{O}_{\text{Tax}} = (\mathbf{C}, \leq_{\mathbf{C}}, \mathbf{A}, \text{VEB})$ bei Instanziierung auch die Attributwerte $a(g)$ selbst berücksichtigen (soweit sie definiert sind), so bildet man gemäß dem Vorschlag der FBA [1] einen sog. **mehrwertigen formalen Kontext**

$$K(\mathbf{I}^*) = (\mathbf{G}, \mathbf{A}, W, \mathbf{I}^*),$$

wobei W die Menge aller möglichen *Werte* von Attributen aus \mathbf{A} ist; und \mathbf{I}^* keine binäre, sondern eine aus der binären Relation \mathbf{I} abgeleitete **3-stellige** Relation $\mathbf{I}^* \subseteq \mathbf{G} \times \mathbf{A} \times W$ ist, für die gilt:

$$(g, a, w) \in \mathbf{I}^* \quad :\Leftrightarrow \quad (g, a) \in \mathbf{I} \quad \text{und} \quad a(g) = w \quad (\text{für } g \in \mathbf{G}, a \in \mathbf{A}, w \in W).$$

Nach FBA kann man $K(\mathbf{I}^*)$ in einen „größeren“ aber „einfachen“ Kontext $K(\mathbf{I}^{**}) = (\mathbf{G}, W, \mathbf{I}^{**})$ mit **binärer** Inzidenzrelation \mathbf{I}^{**} umwandeln, indem man \mathbf{G} als „Gegenstandsmenge“ beibehält, aber statt \mathbf{A} die Wertemenge W als „Merkmalsmenge“ nimmt, und \mathbf{I}^{**} definiert durch

$(g, w) \in \underline{\mathbf{I}}^{**} : \Leftrightarrow$ es gibt ein Attribut $a \in \mathbf{A}$, so dass $a(g)$ definiert ist und den Wert $a(g) = w$ hat.

Da nun $K(\underline{\mathbf{I}}^{**})$ wieder ein „einfacher“ formaler Kontext ist, gehört zu ihm ein formaler Begriffsverband $\underline{\mathbf{B}}(\underline{\mathbf{I}}^{**})$. Bei den formalen Begriffen von $\underline{\mathbf{B}}(\underline{\mathbf{I}}^{**})$ bestehen die Umfänge nun aus Mengen von Instanzen, die Inhalte aus Mengen von Attributwerten.

4.4 Formale Kontexte der Relationship-Types

Nun betrachten wir das volle O-Schema $\mathbf{O}_{\text{full}} := (\mathbf{C}, \leq_{\mathbf{C}}, \mathbf{A}, \mathbf{RELT}, \text{VEB})$, beziehen also die Menge \mathbf{RELT} der Relationship-Types mit ein.

Nehmen wir im O-Schema einmal zwei Klassen $c, d \in \mathbf{C}$ her, die durch einen Relationship-Type $r \in \mathbf{RELT}$ verbundenen sind, $\text{relt}(r) = (c, d)$. Es bestehe also die relationale Teilkonfiguration $\mathbf{c} \mathbf{r} \mathbf{d}$.

Beispiel für eine Teilkonfiguration $\mathbf{c} \mathbf{r} \mathbf{d}$ (vgl. Abb.1): $\wedge \text{Künstler} \langle \rangle \text{erzeugt} \wedge \text{Kunstwerk}$.

Ich verwende im Folgenden die etwas kürzere Notation „ \underline{c} “ statt „ $\text{inst}(c)$ “ für die Menge aller Instanzen der Klasse c .

Sind $\underline{c} := \text{inst}(c)$, $\underline{d} := \text{inst}(d)$ die Instanzenmengen der Klassen c, d und ist $\circ r$ die zum Relationship-Type r gehörige Aussageform, und bilden wir daraus die mathematische Relation

$$\underline{r} := \{(g, m) \mid g \in \underline{c}, m \in \underline{d}, g \text{ ? } m\} \subseteq \underline{c} \times \underline{d},$$

so bekommen wir gemäß FBA den **formalen Kontext**

$$K(\underline{r}) := (\underline{c}, \underline{d}, \underline{r}).$$

$K(\underline{r})$ stellt die Teilkonfiguration **c r d** des O-Schemas dar. Die „**Gegenstände**“ von $K(\underline{r})$ sind die Instanzen der Klasse c , die „**Merkmale**“ sind die Instanzen der Klasse d . Die Inzidenzrelation \underline{r} induziert wieder die beiden Abbildungen

$\uparrow r: \text{Pot}_{\underline{c}} \rightarrow \text{Pot}_{\underline{d}}$, $\downarrow r: \text{Pot}_{\underline{d}} \rightarrow \text{Pot}_{\underline{c}}$, („Galois-Verbindung“), definiert für beliebige Teilmengen $X \subseteq \underline{c}$, $Y \subseteq \underline{d}$ durch:

$$X^{\uparrow r} := \{y \in \underline{d} \mid \text{für alle } x \in X: x \text{ ? } y\}, \quad Y^{\downarrow r} := \{x \in \underline{c} \mid \text{für alle } y \in Y: x \text{ ? } y\}.$$

Ein **formaler Begriff** des Kontextes $K(\underline{r})$ ist gemäß FBA-Sprache ein Mengenpaar (A, B) mit $A \subseteq \underline{c}$, $B \subseteq \underline{d}$, das die Bedingungen

$$A^{\uparrow r} = B, \quad B^{\downarrow r} = A,$$

erfüllt. A heißt der **Umfang**, B der **Inhalt** des formalen Begriffs (A, B) .

In Worten: **Der „Inhalt“ B besteht aus denjenigen Instanzen von d , die mit *allen* Instanzen des „Umfangs“ A in der Beziehung ? stehen ; der „Umfang“ A besteht aus denjenigen Instanzen von c , die mit *allen* Instanzen des „Inhalts“ B in der Beziehung ? stehen.**

Das klingt kompliziert, weil es ganz allgemein formuliert ist. Im konkreten Fall einer „kleinen“ Ontology ergibt sich oft etwas ganz Einfaches.

Beispiel (vgl. Abb.1): Klassen $c := \hat{\text{Maler}}$, $d := \hat{\text{Bild}}$, Relationship-Type $r := \langle \rangle \text{malt}$. $\underline{c} = \hat{\text{Maler}}$ ist dann die Menge aller Malerinstanzen, $\underline{d} = \hat{\text{Bild}}$ ist die Menge aller Bildnstanzen, die in der betreffenden Ontologie ermittelbar sind, und $\underline{r} = \langle \rangle \text{malt} = \{(x, y) \mid \text{>}x \text{ } \langle \rangle \text{malt } \text{>}y\}$ ist die zum Relationship-Type $\langle \rangle \text{malt}$ gehörige **Relation**. Eine Ausprägung davon wäre zum Beispiel:

$\text{>Pablo_Picasso } \langle \rangle \text{malt } \text{>Junge_mit_der_Pfeife}$. Nehmen wir nun bei der gewählten Instanziierung an, Picasso habe alle seine Bilder selbst gemalt (ohne Beteiligung von Schülern), dann bildet das Mengenpaar (A, B) mit

$$A := \{\text{>Pablo_Picasso}\} = B^{\downarrow r} \quad (\text{eine nur 1-elementige Menge!}),$$

$$B := \text{Menge aller Bilder von Pablo Picasso} = A^{\uparrow r},$$

einen **Formalen Begriff** zum formalen Kontext $(\hat{\text{Maler}}, \hat{\text{Bild}}, \langle \rangle \text{malt})$.

Der zum formalen Kontext $K(\underline{r})$ gehörige **Begriffsverband** sei mit $(\underline{B}(\underline{r}), \leq_r)$ bezeichnet. $\underline{B}(\underline{r})$ trägt in natürlicher Weise eine **Ordnung** „ \leq_r “, die definiert ist durch

$$(X, Y) \leq_r (X', Y') : \Leftrightarrow X \subseteq X', Y' \subseteq Y \quad \text{für formale Begriffe } (X, Y), (X', Y') \in \underline{B}(\underline{r}).$$

Im Kontext $(\underline{c}, \underline{d}, \underline{r})$ wird die Klasse c durch den Umfang \underline{c} des *größten* Begriffs $(\underline{c}, \underline{c}^{\uparrow r})$, die Klasse d durch den Inhalt \underline{d} des *kleinsten* formalen Begriffs $(\underline{d}^{\downarrow r}, \underline{d})$ dargestellt. (Die Umfänge / Inhalte der formalen Begriffe „zwischen“ dem kleinsten und dem größten stellen sozusagen „Verfeinerungen“ der Klassen c, d dar, gehören aber selbst nicht mehr zum ursprünglichen Klassensystem \mathbb{C} .)

So kann man **jede** relationale Teilkonfiguration $\mathbf{c r d}$ des O-Schemas $[c, d \in \mathbb{C}, r \in \mathbf{RELT}, \text{rel}(r)=(c,d)]$ durch einen **formalen Kontext** $K(\underline{r}) = (\underline{c}, \underline{d}, \underline{r})$, und damit durch einen **Begriffsverband** $(\underline{\mathbf{B}}(\underline{r}), \leq_r)$, darstellen.

4.5 „Verfeinerung“ einer Ontology im Sinne der FBA

In der üblichen Definition einer Ontology begnügt man sich mit nur **zwei** Ebenen: Der oberen Ebene des **O-Schemas** und der unteren Ebene der **Instanziierung**. Setzt man, wie eben beschrieben, FBA-Methoden ein, so bekommt man eine „**Verfeinerung**“ in **drei** Ebenen im folgenden Sinne:

- Die **Oberste Ebene** bildet das Stück $\mathbf{O}_{\text{Tax}} := (\mathbb{C}, \leq_{\mathbb{C}}, \mathbb{A}, \text{VEB})$ des **O-Schemas**. \mathbf{O}_{Tax} wird durch den formalen Kontext $K(\mathbf{I}) = (\mathbf{G}, \mathbb{A}, \mathbf{I})$ bzw. dessen Begriffsverband $(\underline{\mathbf{B}}(\mathbf{I}), \leq_{\mathbf{I}})$ dargestellt, wobei die formalen Begriffe gerade die Paare $(c, \text{att}(c))$ aus Klassen $c \in \mathbb{C}$ und ihren Attributemengen $\text{att}(c) \subseteq \mathbb{A}$ sind. Dabei erfolgt die **Vererbung der Attribute** „von oben nach unten“ entlang der Klassentaxonomie $<_{\mathbb{C}}$ und, symmetrisch („dual“) dazu, erfolgt die „**Vererbung**“ der **Instanzen** „von unten nach oben“ entlang $<_{\mathbb{C}}$.
- Eine **mittlere Ebene** berücksichtigt alle **Relationship-Types** der Menge **RELT**. Diese Ebene wird durch alle formalen Kontexte $K(\underline{r}) = (\underline{c}, \underline{d}, \underline{r})$ bzw. die

formalen Begriffsverbände $(\underline{\mathbf{B}}(\underline{r}), \leq_r)$ dargestellt, wobei $r \in \mathbf{REL T}$ je ein Relationship-Type mit $\text{relt}(c, d)$ ist, der gewisse Paare $c, d \in \mathbb{C}$ zu einer relationalen Teilkonfiguration $\mathbf{c} \mathbf{r} \mathbf{d}$ verbindet. Der *Umfang* des jeweils \leq_r -größten formalen Begriffs von $(\underline{\mathbf{B}}(\underline{r}), \leq_r)$ repräsentiert die Klasse c , der *Inhalt* des jeweils \leq_r -kleinsten formalen Begriffs von $(\underline{\mathbf{B}}(\underline{r}), \leq_r)$ repräsentiert die Klasse d . Die *Inzidenzrelation* $\underline{r} = \{(g, h) \in \underline{c} \times \underline{d} \mid g \circ r h\}$ repräsentiert den *Relationship-Type* r .

Anmerkungen zur mittleren Ebene:

- (i) Diese „mittlere Ebene“ sehe ich als eine „Verfeinerung“ an, weil hier als „Mitteldinge“ zwischen dem Bereich der Klassen/Attribute und dem Bereich der Instanzen/Attributwerte die (aus Instanzen bestehenden) Umfänge und Inhalte bezüglich der **Relationship-Types** untersucht werden können.
- (ii) Wie schon in **Kap.3.4** erwähnt, sind in dieser mittleren Ebene hauptsächlich die formalen Kontexte der „**maximalen**“ Relationship-Types relevant. Die Kontexte der andern Unter-Relationship-Types werden daraus als „Teilkontexte“ durch **Vererbung** abgeleitet.
- (iii) Diese mittlere Ebene kann man sogar durch **einen einzigen** formalen Kontext $K(\mathbf{i}) := (\mathbf{F}_{(\mathbf{REL T})}, \mathbf{REL T}, \mathbf{i})$ darstellen, wobei $\mathbf{F}_{(\mathbf{REL T})}$ die **Vereinigung** aller Begriffsverbände $\underline{\mathbf{B}}(\underline{r})$ ($r \in \mathbf{REL T}$) ist, und die Inzidenzrelation \mathbf{i} einfach definiert ist durch $(A, B) \mathbf{i} r \Leftrightarrow (A, B) \in \underline{\mathbf{B}}(\underline{r})$ für $(A, B) \in \mathbf{F}_{(\mathbf{REL T})}$, $r \in \mathbf{REL T}$. $K(\mathbf{i})$ nenne ich einen „**F-Compound Context**“; die formalen Begriffe dazu nenne ich „**F-Compounds**“.

Vgl. *Lübbert* [18]. Sowohl das System der $(\mathbf{F}_{(\text{REL T})}, \leq_F)$ der „Gegenstände“ als auch das System $(\text{REL T}, \leq_{\text{REL T}})$ der „Merkmale“ sind in natürlicher Weise geordnet. Die Ordnung \leq_F ist durch die Ordnungen \leq_r ($r \in \text{REL T}$) der Begriffsverbände $(\underline{\mathbf{B}}(r), \leq_r)$ induziert; die Ordnung $\leq_{\text{REL T}}$ ist die oben beschriebene **Relationship-Taxonomie**. Der F-Compound Context und sein F-Compound-Begriffsverband dienen hauptsächlich dazu, sich klar zu machen, dass nicht nur einzelne Instanzpaare (x, y) sondern auch formale Begriffe $(A, B) \in \mathbf{F}_{(\text{REL T})}$ zu *verschiedenen* Relationship-Types $r, s \in \text{REL T}$ in Beziehung stehen können.

- Die **unterste Ebene** bildet, wie gehabt, die **Instanziierung**. Diese wiederum wird gemäß FBA dargestellt durch den **mehrwertigen Kontext** $\mathbf{K}(\mathbf{I}^*) = (\mathbf{G}, \mathbf{A}, W, \underline{\mathbf{I}}^*)$ oder auch durch den daraus gebildeten „einfachen“ formalen Kontext $\mathbf{K}(\mathbf{I}^{**}) = (\mathbf{G}, W, \mathbf{I}^{**})$ und dessen Begriffsverband $(\underline{\mathbf{B}}(\mathbf{I}^{**}), \leq_{\mathbf{I}^{**}})$.

5 Schlussbemerkung

Ich habe hier zu schildern versucht, dass die **Struktur** des (mir bekannten) informatischen Konzepts „Ontology“ symmetrischer und mathematisch präziser mit Hilfe der **FBA-Sprache** formuliert werden kann, wodurch die Implementierungspraxis der Informatiker eine wohldefinierte mathematische Basis erhält. Mit Kenntnis dieser Basis darf der Implementierer dann im Bedarfsfalle von der mathematischen Grundstruktur abweichen, weiß aber dann genau um die Probleme, die eventuell durch diese Abweichung entstehen können.

Ich meine, Informatiker könnten sich besser über die verschiedenen Varianten ihres Konzepts „Ontology“ einigen – und wären darüber hinaus „auf der sicheren Seite“ in der Implementierungspraxis – wenn sie sich etwas ernsthafter die dazugehörigen **mathematischen Grundlagen** zu Gemüte führen und außerdem genauer zwischen „**Struktur**“ und „**Deutung**“ (Anwendungsmöglichkeiten) **unterscheiden** würden, als wenn sie sich nur an alten philosophischen und „metaphysik-geschwängerten“ Anschauungen aus Antike und Mittelalter orientieren.

Über das in dieser Notiz Angedeutete hinaus gibt es aber noch mehrere andere Möglichkeiten, wie man die konventionelle Definition eines Konzepts „Ontology“ mit **FBA**-Methoden „verfeinern“ oder „erweitern“ könnte. Insbesondere bietet sich mit FBA-Methoden der Entwurf einer „Ontology“ an, wo die Relationship-Types nicht nur binär sondern auch **mehr als 2-stellig** sein dürfen (also mehr als nur je zwei Klassen betreffen) – vgl. etwa [6], [7], [28]. Darauf kann ich aber der Kürze halber hier nicht weiter eingehen.

6 Literatur

- [1] **B. Ganter / R. Wille: „Formale Begriffsanalyse“ (FBA);** Springer 1996.
[Die am meisten benutzte Referenz]
- [2] G. Pickert: „Einführung in Ontologien“, Humboldt-Universität Berlin, Feb. 2011. http://www.dbis.informatik.hu-berlin.de/dbisold/lehre/WS0203/SemWeb/artikel/2/Pickert_Ontologien_final.pdf
(diente zusammen mit [12] als Motivation zur Abfassung dieser Note)
- [3] W. Bartussek, B. Humm, A. Reibold, T. Zeh (Ontologie-Arbeitskreis der Hochschule Darmstadt): „Zur Definition von ‚Ontologie‘ in den Informationswissenschaften“, Oktober 2010 / Mai 2011, Hochschule Darmstadt (h_da).
- [4] Autorenkollektiv der Ontologie-Arbeitsgruppe der Hochschule Darmstadt (h_da): „Was ist eigentlich Ontologie“, Jan.2012.
- [6] T. Zeh: „Darstellung von Texten in den drei Formalismen Power Context Families (PCF), Relational Data System (RDS) und Entity Relational System (ERS) anhand des Fallbeispiels „Biographie von Albert Einstein“, Hochschule Darmstadt, 13.10.2010.
- [7] K. E. Wolff: „Relational Semantic Systems, Power Context Families, and Concept Graphs, Hochschule Darmstadt, 2009,
<http://www.fbm.fh-darmstadt.de/home/wolff>

- [8] K. E. Wolff: „Relational Scaling in Relational Semantic Systems“, Hochschule Darmstadt, 2009, <http://www.fbmn.fh-darmstadt.de/home/wolff>
- [9] Wikipedia: „Vererbung (Programmierung)“, Version 20.11.2011
- [10] Wikipedia: „Ontologie (Informatik)“, Version Nov. 2013
- [12] A. Maedche & V. Zacharias: „Clustering Ontology-based Metadata in the Semantic Web“, FZI Research Center for Information Technologies at the University of Karlsruhe, Research Group WIM, 2002/2003 (?)
(diente zusammen mit [2] als Motivation zur Abfassung dieser Note)
- [13] Nicola Guarino: „Formal Ontology, Conceptual Analysis and Knowledge Representation“, 1995 (?)
- [14] John F. Sowa: “KR Ontology”, <http://www.jfsowa.com/ontology/> , *last modified:* Nov.2010
- [16] M. Kreuzer / S. Kühling: „Logik für Informatiker“. Vlg. Pearson Studium, 2006.
- [17] C. Lübbert: „Mehrwertige Logiken“, V11, Dez.2011, www.cl-diesunddas.de
- [18] C. Lübbert: „Ontologie-Definition auf Basis von FBA“, V3.14, Darmstadt, Aug. 2012

- [19] H. Herre et.al.: „General Formal Ontology (GFO)“, Part I, Basic Principles. Leipzig, July 2006. www.onto-med.de
- [20] Barry Smith: „Against Fantology“, (in Experience and Analysis), Vienna, 2005
- [25] Stanford Encyclopedia of Philosophy: „Mereology“, 2003 / 2009. <http://plato.stanford.edu/entries/mereology/>
- [26] Stanford Encyclopedia of Philosophy: “States of Affaires”, März 2012 <http://plato.stanford.edu/entries/states-of-affairs/>
- [28] R. Wille: „Conceptual Structures of Multicontexts“, FB Mathematik, TU Darmstadt, 1996

Mit freundlichem Gruß

Christoph Lübbert