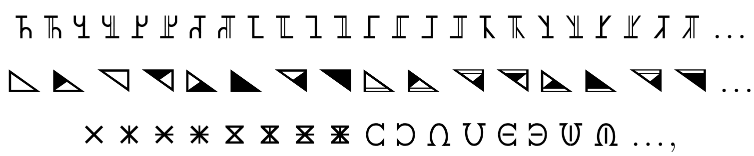
Was ist und was soll UpMath?

Harald Upmeier, Matthias Graefenhan

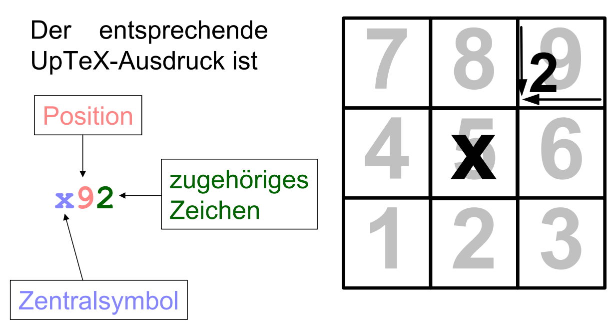
Während die erste Hälfte des 20. Jahrhunderts in der Wissenschaft als Epoche der Physik gilt, kann die zweite Hälfte durchaus als Zeitalter der Mathematik bezeichnet werden. In dieser Blütezeit wurden nicht nur fundamentale Probleme z.B. in der Zahlentheorie gelöst, sondern auch neue Gebiete axiomatisch begründet und mit der klassischen Mathematik verbunden. Als Konsequenz ist der inhaltliche und methodische Umfang der Mathematik seit den 50er Jahren explosionsartig angestiegen.

Wie ist es möglich, bei dieser Erweiterung des mathematischen Wissens einen Gesamtüberblick zu behalten, wie er für die mathematische Forschung nötig ist? Diesem Anliegen dient das UpMath-Projekt[1], eine virtuelle Datenbank mathematischer Einzeldokumente, die im Gegensatz zu den üblichen Internet-Enzyklopädien eine vollständige, also nicht nur stichwortartige, Darstellung des jeweiligen Wissensgebietes anstrebt.

Die erste Stufe des UpMath-Systems ist die neuartige Symbolik: Die üblichen Buchstaben werden kaum benutzt und auf Worte wird fast ganz verzichtet. Stattdessen gibt es als TrueType Schriftart ein doppelt-spiegelungsinvariantes Alphabet aus etwa 5000 Symbolen



welche in alle Richtungen (jeweils 8 Positionen) in zwei-dimensionaler Weise kombiniert werden können.



[1] http://www.mathematik.uni-marburg.de/~upmath/

Dieser Prozess kann baukastenartig iteriert werden. Die speziell für UpMath entwickelte Sprache UpTeX[2] ermöglicht es, die besondere Notation in kompakter Form mit einer herkömmlichen Tastatur einzugeben und für das bekannte Satzsystem LaTeX aufzubereiten. Anschließend werden daraus Dateien im Portable Document Format[3] (PDF) erzeugt. Die einzelnen Übersetzungsschritte sind für den Benutzer transparent, d.h. nach einem Klick wird ihm das übersetzte Dokument angezeigt. Will man die komplette Sammlung virtuell veröffentlichen, werden all ihre Dokumente neu übersetzt und eine Übersichtswebsite generiert. Diese vollständige Neuübersetzung dient dazu, ein einheitliches Layout zu gewährleisten. Globale Einstellungen, wie etwa Abstände der Zeichen untereinander oder feine Änderungen der Symbole selbst, werden so bei allen Dokumenten wirksam. Trotz der großen Anzahl der Dokumente wird die Dauer des Übersetzungsprozesses durch Ausnutzung von Technologien moderner Prozessoren, wie Hyperthreading[4] und Multi-Core-Architekturen[5], relativ kurz gehalten.

Als zweite Stufe erhält man sogenannte „Ikone“ , welche die Grundlage von mathematischen Formeln und Sachverhalten bilden, beispielsweise

$$\llcorner \triangledown^{\cup}_{m}\mathbb{R}^{n} \triangle_{\infty} \lrcorner$$

zusammengesetzt aus $\llcorner\ \lrcorner\ \mathbb{R}\ \cup$ und den allgegenwärtigen Dreiecken $\triangle\ \triangledown$. Diese Ikone bedeuten in allen Dokumenten stets das gleiche mathematische Objekt und werden nur einmal erklärt. Dadurch wird bei aller Ausführlichkeit die Darstellung kurz und elegant gehalten, und die Beziehung zwischen verschiedenen Dokumenten ist schon an der Geometrie der Ikone zu erkennen, etwa bei wichtigen Dualitätsrelationen: Zeilenvektoren $\llcorner^{\cdot}$ und Spaltenvektoren $_{\cdot}\urcorner$ werden durch Matrizen $_{\cdot}\lrcorner^{\cdot}$ verbunden:

$$_{\cdot}\lrcorner^{\cdot}\,{}_{\cdot}\urcorner = {}_{\cdot}\urcorner' \qquad \llcorner^{\cdot}\,{}_{\cdot}\lrcorner^{\cdot} = \llcorner'^{\cdot}$$

Ein weiteres Beispiel: Der Hilbertsche Nullstellensatz lautet in der UpMath-Symbolik:

$$\hat{\lrcorner} = \overbrace{\hbar \diagup \lrcorner} \diagdown \underbrace{\hbar \triangle_{\flat} \mathbb{F}}$$

Hingegen wird der Rungesche Approximationssatz der Funktionentheorie so formuliert:

$$\hat{\hbar} = \mathbb{C} \triangledown_{0} \overbrace{\hbar \diagdown \hbar \triangle_{\varpi} \mathbb{C}}$$

Schon graphisch wird deutlich, dass es sich hier um ähnliche, zueinander duale Sachverhalte handelt. Die herkömmliche Formulierung dagegen ist mehr an Fachausdrücken

[2] Siehe auch Präsentation „UpTeX“.
[3] http://www.adobe.com/products/acrobat/adobepdf.html
[4] http://www.intel.com/technology/hyperthread/
[5] http://de.wikipedia.org/wiki/Multicore-Prozessor

(Radikal, Ideal, Spektrum,...) orientiert, welche in der Regel bei jedem Auftreten neu erklärt werden.

In der dritten Stufe werden die Ikone zu mathematischen Dokumenten (= virtuelle Buchseiten) verbunden, welche jeweils nur einem einzigen Thema gewidmet sind und mit einem charakteristischen Namen („Idol“) versehen werden – auf Basis des Unicode-Systems, bei dem mehrere tausend Zeichen zur Verfügung stehen. Die sinnvolle Ausarbeitung dieser Dateienhierarchie ist ähnlich schwierig wie die mathematishe Arbeit an den Dokumenten selbst. Die in der Mathematik üblichen Themenüberschneidungen werden in „homogene“ Bestandteile zerlegt und in separaten Dokumenten gespeichert. Ein Beispiel: Der Logarithmus ist die „Umkehrfunktion“ der Exponentialfunktion, beide Funktionen werden aber nicht in einem Dokument behandelt, sondern jeweils für sich; die enge Beziehung zwischen beiden wird erst durch Vernetzung hergestellt. Bei den Idol-Zeichenketten stehen Vokale für die „Größe“ der behandelten Objekte (etwa E = endlich-dimensional, U = unendlich-dimensional), Akzente geben die Kategorisierung (Menge, Gruppe usw.) an, die Konsonanten bezeichnen den Zahlbereich (etwa R = reelle Zahlen, C = komplexe Zahlen) und die Unicode-Sonderzeichen werden für Zusatzstrukturen (Einheitskugel, Extremrand usw.) benutzt. Einige Beispiele für (sehr einfache) Idole:

Ú²r	=	unendlich-dimensionaler Hilbertraum über den reellen Zahlen
éc°	=	Einheitskugel eines komplexen Vektorraumes endlicher Dimension
rÏtá°	=	Schnitte eines reellen Geradenbündels über dem Torus.

So wie die Ikone baukastenartig aus den Symbolen zusammengesetzt sind, ermöglicht die Separierung der Dokumente mit eindeutigen Idolnamen eine vielfältige Vernetzung der Einzelseiten mit anderen in sich homogenen Seiten ohne Duplikationen. Die Idole sind also die Bausteine des Gesamtsystems. Ein Schema aus 24 Idolen ist etwa:

$!ek$ L ⧖ L$^\sharp$	$Ke!$	Ke C ⋉ ⅂	ke L	$ke!$ ⅂:L ⋉ ^{L}◺$\mathbb{K}^{\mathbb{N}}$	$Ke!$
$!eK$	$ke!$ L$^\sharp$ ⧖ L	ek ⅂	Ke L ⋊ C	$eK!$	$Ke!$ $\mathbb{K}_{\mathbb{N}}$◹L ⋊ ⅂:L
$!Ek$ ℂ| ^{L}◺$\mathbb{K}^{\mathbb{N}}$	$KE!$	EK C ⋉ ^{L}◺$\mathbb{K}^{\mathbb{N}}$	kE $\mathbb{K}_{\mathbb{N}}$◹L	$EK!$ ^{L}◺$\mathbb{K}^{\mathbb{N}}$ ⧖ $\mathbb{K}_{\mathbb{N}}$◹L	$kE!$
$!EK$	$kE!$ ℂ|$\mathbb{K}_{\mathbb{N}}$◹L	Ek ^{L}◺$\mathbb{K}^{\mathbb{N}}$	KE $\mathbb{K}_{\mathbb{N}}$◹L ⋊ C	$Ek!$	$KE!$ $\mathbb{K}_{\mathbb{N}}$◹L ⧖ ^{L}◺$\mathbb{K}^{\mathbb{N}}$

Insgesamt könnte man UnMath also als ein virtuelles Buch mit (voraussichtlich) etwa 10.000 Seiten bezeichnen, geschrieben in Worten (Ikone) einer Schrift aus 5000 Buchstaben (Symbole) und mit Seitenüberschriften (Idole) aus Unicode-Zeichen. Im Gegensatz zu realen Büchern ist das Projekt dynamisch, durch Erstellung neuer Dateien oder durch Einführung eines neuen Symbols oder Vernetzungstyps, welche automatisch in allen bereits bestehenden Dokumenten ausgeführt wird. Die einzige „menschlich-verbale“ Komponente ist eine beschreibende Textzeile zu Beginn eines jeden Dokuments, die auch in der Gesamtübersicht (Overview) gezeigt wird. Auf lange Sicht soll diese Textzeile wegfallen, das Idol reicht zur Erkennung des Dokuments aus und braucht nicht mehr verbalisiert zu werden.

Nun zur Frage „Was soll UpMath?“. Historisch gesehen hat es immer wieder Versuche einer Gesamtdarstellung der Mathematik gegeben. Besonders bekannt ist das Monumentalwerk von „Bourbaki“[6], einer anonymen Gruppe französischer Spitzenmathematiker. Auch neuartige Symboliken wurden wiederholt vorgeschlagen; so hatte der mathematische Logiker Giuseppe Peano ein eigenes Schriftsystem, welches aber nur im engsten Schülerkreis verstanden wurde. UpMath benutzt stattdessen „physikalische“ Grundideen (Lagrange-Formalismus und Quantisierung) als übergreifende Ordnungsprinzipien auch in der Mathematik. Konkret bedeutet dies, dass fast jedes Dokument als (im Sinne der Physik)

$$\begin{array}{rcl} \text{klassisch} & = & \text{nichtlinear, reell, endlichdimensional} \\ \text{quantisiert} & = & \text{linear, komplex, unendlichdimensional} \end{array}$$

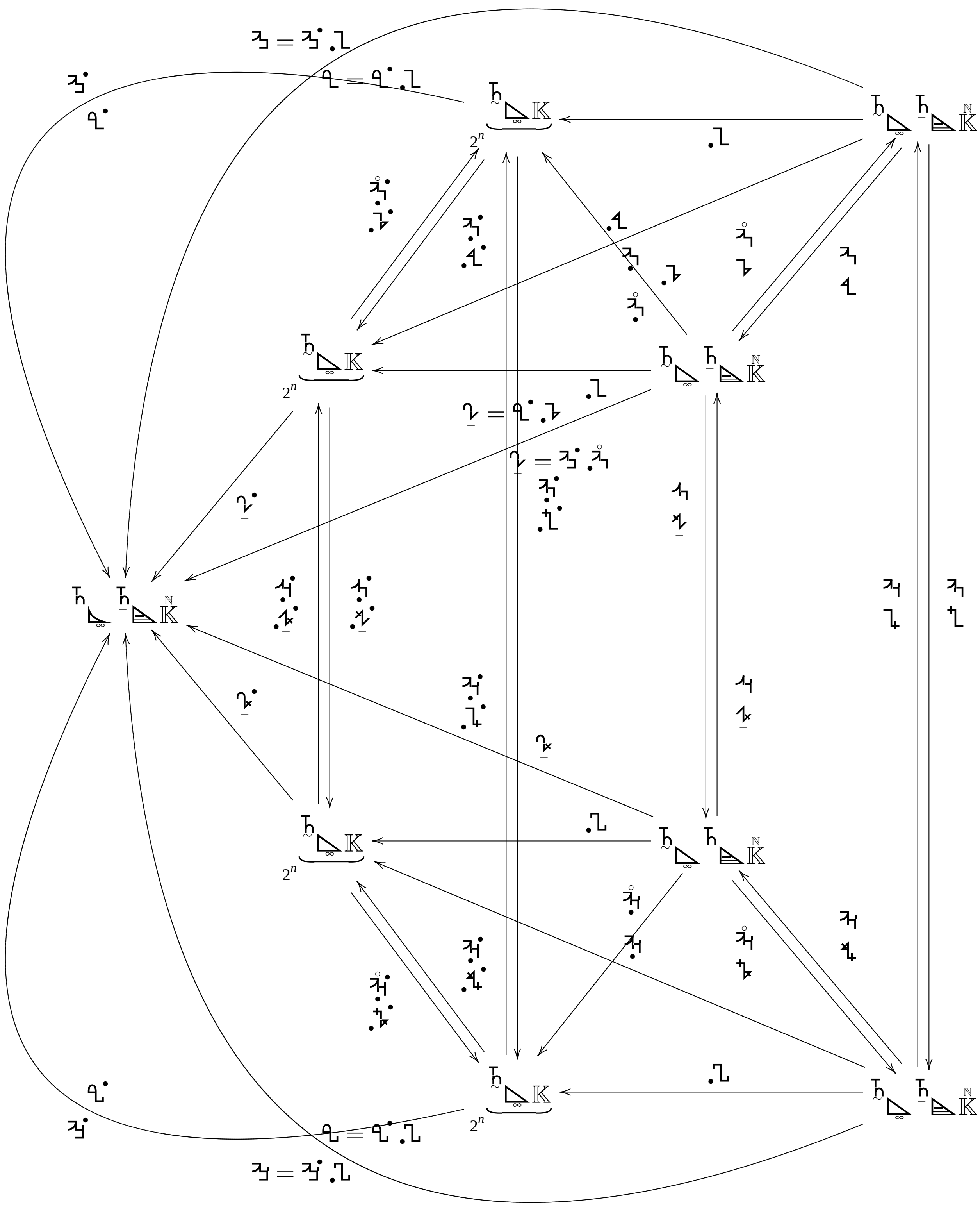
eingeteilt wird. Ein Beispiel aus der Funktionentheorie: Der Cauchy-Integralsatz ist ein „klassischer“ Satz, da er invariant unter komplexer Konjugation ist und daher „reell“ aufzufassen ist (als Spezialfall des Satzes von Stokes). Hingegen gehört die Cauchy-Integralformel, die eine komplexe Struktur voraussetzt, zur „Quantisierung“. Inzwischen werden selbst zahlentheoretische Sätze als Erhaltungssätze im physikalischen Sinne gedeutet. In der visionären Stringtheorie[7], die eine Hauptmotivation zur Arbeit an UpMath ist, ist die Verbindung von Mathematik und Physik vollständig realisiert.

UpMath ist aber nicht nur ein fachwissenschaftliches System, sondern spiegelt auch moderne Medien-Trends wider: Das UpMath-Kompendium existiert ausschließlich in digitalisierter Form (ein Ausdrucken aller Dokumente ist wegen des großen Umfangs nicht praktikabel und wegen der ständigen Änderungen auch nicht wünschenswert), und ästhetische Aspekte wie Symmetrie, Graphik usw. sind bei der ikonographischen Erstellung der Dokumente genauso wichtig wie der mathematische Inhalt – gemäß der Überzeugung, dass konsequente Ästhethisierung auch der mathematischen Wahrheit am nächsten kommt. Die Fotogalerie „Flowers and UpMath“[8] von Dr. Helga Upmeier benutzt UpMath -Symbole als graphischen Hintergrund. Die Kombination zu Ikonen liefert etwa das folgende Diagramm aus der Differentialgeometrie:

[6] http://www.bourbaki.ens.fr/

[7] Brian Greene: The Elegant Universe: Superstrings, Hidden Dimensions, and the Quest for the Ultimate Theory, ISBN 0393058581

[8] http://www.mathematik.uni-marburg.de/~upmeier/Gallery/



Besonders interessant ist nun die geplante vierte Stufe, nämlich die Kombination von Idolen (jeweils für ein mathematisches Thema stehend) zu komplexeren Gebilden. Ähnlich wie bei der 8-fachen Verbindung von Symbolen geht es hier darum, die Dokumente und ihre Beziehungen untereinander auszuwerten. Um etwa die „logische Vergangenheit“ eines Ikons zu bestimmen, wird es schrittweise in seine Bestandteile (die zunächst selber noch Ikone niederer Komplexität sein können) aufgelöst, bis die Elementarsymbole erreicht sind. Dies alles durchläuft verschiedene Dokumente, so dass man eine, auch ästhetisch reizvolle, stammbaumartige Kaskade von Idolen erhält, erzeugt von dem ausgewählten Ikon. Die Komplexität einer Vorlesung oder eines fachlichen Aufsatzes wird durch diese Interdependenzen reflektiert.

Zur Realisierung solcher Anwendungen müssen inhaltliche Suchfunktionen geboten werden, d.h. mehr als eine reine Zeichenkettensuche; es soll Verweise auf andere Dokumente geben, beispielsweise zur Begründung eines Beweisschrittes; eine Gruppierung der Dokumente soll zur Bereitstellung von Material für Lehrveranstaltungen möglich sein; und es soll die Möglichkeit geben, die Dokumente beim Zugriff automatisch zu übersetzen. Übersetzung bedeutet in diesem Zusammenhang zum einen, die Dokumente mit erklärenden Texten zu ergänzen und ggf. in klassischer mathematischer Notation auszugeben, um so einen breiteren Zugang zu ermöglichen, zum anderen die Wahl der Sprache der Ergänzungstexte. (Diese nachträgliche Verbalisierung steht allerdings im Widerspruch zur UpMath-Philosophie und wird daher nur extern angeboten). Bei der aus vielen Programmen bekannten Zeichenkettensuche muss eine exakte, buchstabenweise Übereinstimmung mit dem Suchbegriff vorliegen, damit eine Textstelle gefunden wird. Die besondere UpMath-Notation ermöglicht jedoch die automatisierte Auswertung der Struktur eines Ausdrucks. So kann sowohl eine semantische Suche realisiert werden, als auch das Finden verwandter oder aufeinander aufbauender Dokumente.

All diese Funktionen sind von dem momentanen – für die derzeitige Verwendung bewährten – System, dass auf HTML und PDF aufbaut, nicht effizient oder gar nicht zu leisten. Deshalb wird zukünftig ein Server das UpMath-Archiv verwalten. Zugriff auf die oben beschriebenen Funktionen erhält man über einen Client, dessen Installation und Wartung durch ClickOnce Deployment[9] für den Anwender ähnlich einfach wird wie die Benutzung der bestehenden Website, jedoch deutlich mehr Komfort und Möglichkeiten bietet. Je nach Zielgruppe können zusätzlich Schnittstellen für Clients auf verschiedensten Plattformen, beispielsweise mobilen Geräten, zur Verfügung gestellt werden.

Neben der Nutzung der Dokumente als Teile einer Vorlesung, als Abschnitte einer wissenschaftlichen Abhandlung oder auch als systematische Einordnung von Konferenz-Notizen macht deren freie Kombinierbarkeit weitergehende, auch medienorientierte, Anwendungen denkbar, etwa die folgende: Selbst bei bester Organisation aller Dokumente kann es nicht ausbleiben, dass manche gute Ideen in der Datenbank untergehen und schwierige Rechnungen und Beweise doppelt gemacht werden, obwohl sie bereits im Dateiensystem vorhanden sind. Viele Forscher kennen dieses Gefühl der Duplizierung. Ein Programm

[9] http://msdn2.microsoft.com/en-us/library/wh45kb66(en-US,VS.80).aspx

zur automatischen Anzeige von Redundanzen, als Teil der oben erläuterten umfassenden Dokumentenverwaltung, ist in Vorbereitung. Es gibt aber auch eine psychologische Dimension: Analog zum in Chemie-Hörsälen als permanenter Informations-Hintergrund präsenten Periodensystems der Elemente, könnte man zur Erhöhung des Bewusstsein für die Gesamtstruktur und die inneren Zusammenhänge von UpMath auf großen Flachbildschirmen nach dem Zufallsprinzip Ketten von geöffneten UpMath-Dokumenten aufrufen und als optischen Hintergrund wirken lassen, ähnlich wie in manchen amerikanischen Haushalten, wo der Fernseher niemals ausgeschaltet wird. Der ikonographische Stil von UpMath ist gut geeignet für diese Art des stetigen Aufrufens von Dokumenten, welche wiederum die Dynamik (kontinuierliches Aktualisieren) ermöglicht.

Langfristig bietet die Vernetzung von UpMath-Dokumenten noch mehr Möglichkeiten, welche in Analogie zum Film illustriert werden sollen: Die Kombinierbarkeit von Szenen aus verschiedenen Filmen scheitert schon am Auftreten verschiedener Schauspieler. Die Einheitlichkeit der UpMath-Symbole (Ikonisierungsprinzip) besagt in diesem Zusammenhang, dass es etwa 5000 „typisierte“ Schauspieler gäbe (ähnlich wie bei Cartoons oder im Trickfilm), so dass die Abfolge von Filmszenen nur noch den Regeln der Dramaturgie und Logik unterworfen ist. Im Prinzip könnte man so Filme herstellen, deren Ende offen ist und jedesmal vom Computer nach dem Zufallsprinzip neu bestimmt wird. Die „Idolisierung“ in diesem Bild wäre die Zerlegung von Filmszenen in Grundbestandteile („Sonnenuntergang“ oder verschiedene Arten von „Abschied“). Im Bereich des Films klingt dies alles recht abstrus, aber in der modernen Musik gibt es durchaus computersynthetisierte Kompositionen, die aus musikalischen Elementarharmonien frei kombiniert werden, klassisch tonal oder in Schönbergs 12-Tonmusik. Die Mathematik ist ebenfalls ein (gigantisches) Regelwerk und sollte, bei genügender Strukturierung, ähnliche Kombinationsmöglichkeiten bieten. Ob dabei wirklich einmal ein neues mathematisches Ergebnis herauskommt oder „nur“ ein besseres Verständnis bestehender Mathematik, soll hier offen bleiben. Auf jeden Fall versteht sich UpMath als Beitrag in dieser Richtung, ähnlich wie Tendenzen der Künstlichen Intelligenz oder die „Wissenschaftszeremonien“ in Hermann Hesses Glasperlenspiel, wo ebenfalls ästhetische Form und geistiger Inhalt gleichberechtigt sind. Abschliessend sei erwähnt, dass UpMath langfristig in ein umfassenderes System UpCult, welches auch andere Wissensgebiete wie UpNote und UpChem umfasst, eingebunden werden soll.

Prof. Dr. Harald Upmeier, FB Mathematik und Informatik, Universität Marburg
http://www.mathematik.uni-marburg.de/~upmeier/

Matthias Graefenhan, FB Mathematik und Informatik, Universität Marburg
http://www.mathematik.uni-marburg.de/~graef/