



NETWORX Nr. 8

Peter Schlobinski/Michael Tewes

Graphentheoretische Analyse von Hypertexten

(1999)

online verfügbar seit: 1999

URL: <http://www.websprache.uni-hannover.de/networx/docs/networx-8.pdf>

Versionshinweise: keine



NETWORX – Online-Publikationen zum Thema Sprache und Kommunikation im Internet:
<http://www.websprache.uni-hannover.de/networx/>

Graphentheoretisch fundierte Analyse von HyperTexten
von
Peter Schlobinski & Michael Tewes
(Skizze)

Inhalt

1. Fragestellungen und Ausgangsprobleme
2. Analyseschritte
 - 2.1 Aufbau einer Web-Site
 - 2.2 Das Gerüst eines Hypertextes als Graph
 - 2.3 Semantisch-textuelles Netzwerk
3. Zusammenfassung und Perspektiven
4. Literatur
5. Glossar

1. Fragestellungen und Ausgangsprobleme

Was ein Hypertext ist – darüber gibt es unterschiedliche Meinungen und Ansichten. Für Krajewski (1997: 65) verweist das Kompositum Hypertext „einerseits auf alle Funktionen, die ein gewöhnlicher Text liefert, und deutet andererseits eine Ergänzung an, die über den bisherigen Begriff hinausgeht (und damit auch über den gewöhnlichen Text)“. Das, was über den ‚gewöhnlichen‘ Text hinausweist, ist einerseits mit Begriffen wie non-linear, delinear oder rhizomatisch verbunden, andererseits mit dem Begriff der Multimedialität. Wirth (1997: 321) spricht vom Hypertext „als einer rhizomatisch vernetzten Karte“, für Wingert (1996: 112) sind Hypertexte „eher Collagen aus Textstücken, [...] eher Assemblagen aus unterschiedlichem medialen Material wie Text, Bilder, Ton- und Filmsequenzen“. In nicht-digitalisierter Form gibt es Hypertexte im Sinne der vorangestellten Definitionen lange Zeit vor dem Computerzeitalter (in der Theodizee von Georg Wilhelm Leibniz), und in der modernen Literatur wurden Hypertexturen geschaffen, wie in Ulysses von James Joyce, Zettels Traum von Arno Schmidt oder Wenn ein Reisender in einer Winternacht von Italo Calvino. Aber auch Zeitungstexte oder Comics sind partiell ebenso delinear wie Bücher mit Fußnotenverweisen, Zitaten oder einem Register. Eine neue Qualität wird jedoch dem Hypertext im Internet und in Sonderheit in der Netzliteratur oder Hyperfiction zugewiesen, insofern das Labyrinthische den Hypertext charakterisiert: „Im Hypertext ist es schwierig, wenn nicht

Graphentheoretisch fundierte Analyse von Hypertexten

unmöglich, den Beginn eines Textes genau zu definieren, da der Leser zwischen einer Reihe von Anfängen unterscheiden kann. Ähnlich verhält es sich mit dem Ende. Der Leser kann nicht nur verschiedene Punkte wählen, an denen er ein Ende setzen möchte, er kann ebensogut damit fortfahren, an den ursprünglichen Text etwas anzufügen.“ (Gabriel 1997: 73). Schlimmstenfalls ertrinkt der Leser im Meer digitaler Texte, die durch unendliche viele Links miteinander verknüpft sind. Zur Bestimmung von Strukturformen von Hypertexten hat insbesondere der Begriff des Rhizoms für (‘chaotische’) Netzstrukturen Karriere gemacht. Suter (1999) unterscheidet die Internetliteratur in drei Strukturmuster: lineare Struktur, Baumstruktur und Rhizomstruktur. Die Rhizomstruktur kennzeichnet hierbei eine Verflechtung von Netzstrukturen, die nach allen Seiten chaotisch wächst. Die neueren Ansätze zur Definition von Hypertexten bezeichnen wie der in den sechziger Jahren geprägte Terminus ‚Hypertext‘ „eher eine faszinierende Vision über (...) Schreib- und Lesetechniken als ein klares Konzept.“ (Storrer 1999: 34).

Wir wollen im folgenden den metaphorischen Konzepten zur Beschreibung von sog. nicht-linearen Hypertextstrukturen ein Analyseinstrumentarium entgegenstellen, von dem wir glauben, dass es bei der konkreten Analyse von Hypertexten, genauer: Hypertexten im Internet hilfreich ist. Es geht uns weder darum, komplexe Hypertexte zu analysieren noch eine Theorie zur Bestimmung von Hypertexten zu präsentieren, sondern allein darum, Analyse Kriterien zu entwickeln, mit deren Hilfe Delinearität präziser beschrieben werden kann als in den wenig ausgearbeiteten Konzepten wie sie insbesondere bei der Beschreibung von Netzliteratur vorliegen. Hierfür scheint es uns sinnvoll, auf graphentheoretische Modellierungen zurückzugreifen; Ansätze hierzu finden sich bei Lobin (1999) und bei Vesper (1998: 40 und 88ff.), wo ‚Baupläne‘ graphisch abgebildet als lineare Folge, Baumstruktur und Plenarstruktur behandelt werden.

Einen Hypertext, wie er im Internet vorliegt, – im folgenden HyperText (HT) – werden wir unter zweierlei Aspekten zu betrachten: (i) unter dem strukturell-medialen Aspekt, nämlich als Link-Struktur, und (ii) unter einem inhaltlichen Aspekt, und zwar durch die sprachlichen, im Besonderen durch die semantisch-textuellen Verknüpfungen. Hierunter sind semantisch-textuelle Relationen im weitesten Sinne zu verstehen, die Topik-Progression ebenso beinhalten wie Marker der Textkohärenz und Textkohäsion, also die Schnittstellen zu Syntax und Pragmatik miteinbeziehen. Die Link-Struktur wollen wir als das Gerüst eines HyperTextes $G(HT)$ betrachten, die semantisch-textuellen Verknüpfungen als das semantisch-

Graphentheoretisch fundierte Analyse von Hypertexten

textuelle Netzwerk eines HyperTextes Sem(HT). Gerüst und semantisch-textuelles Netzwerk eines HyperTextes können unabhängig voneinander (als Graphen) analysiert werden und stellen unterschiedliche Analyseebenen dar, wenn auch beide Strukturierungen partiell verknüpft werden können.

2. Analyseschritte

Wir wollen uns im folgenden in erster Linie mit der Analyse der Link-Struktur unter graphentheoretischen Aspekten befassen, wofür zunächst kurz in den Aufbau einer Web-Site eingeführt wird, da dies für die dann folgenden Präzisierungen als Verständnishintergrund vorausgesetzt werden muss. Wir gehen davon aus, dass der Leser mit dem WWW, Web-Sites und Browsern vertraut ist (s. sonst Runkehl/Schlobinski/Siever 1998 und <http://www.websprache.uni-hannover.de/>). Für die Analyse semantisch-textueller Netzwerke und Textrelationen gibt es eine Reihe linguistischer Ansätze, weshalb wir dies nur am Rande behandeln und soweit, wie es für die Verknüpfung beider Strukturebenen von Relevanz ist.

2.1 Aufbau einer Web-Site

Eine Web-Site besteht aus einer Menge von HTML-Seiten. Eine Web-Site ist nicht mit einer Web-Seite zu verwechseln, letztere bezeichnet eine einzelne WWW-Seite, erstere alle Seiten eines Webprojektes (z.B. sprache@web). Die Titelseite einer Web-Site wird Homepage (üblicherweise >index.htm<) genannt (s. auch Abb.1) oder aber auch die Web-Site einer Person wird als Homepage bezeichnet und hat folglich eine entsprechende URL. Nehmen wir an, wir wollen eine Websprache-Site erstellen. Ausgehend von der Startseite, der Homepage, wollen wir für zwei Personen (Peter und Max) eine eigene Seite anlegen, auf der wiederum ein Foto von Peter bzw. Max verankert werden soll. Hierfür ist es sinnvoll, eine Verzeichnisstruktur zu erstellen (s. Abb. 1).

Graphentheoretisch fundierte Analyse von Hypertexten

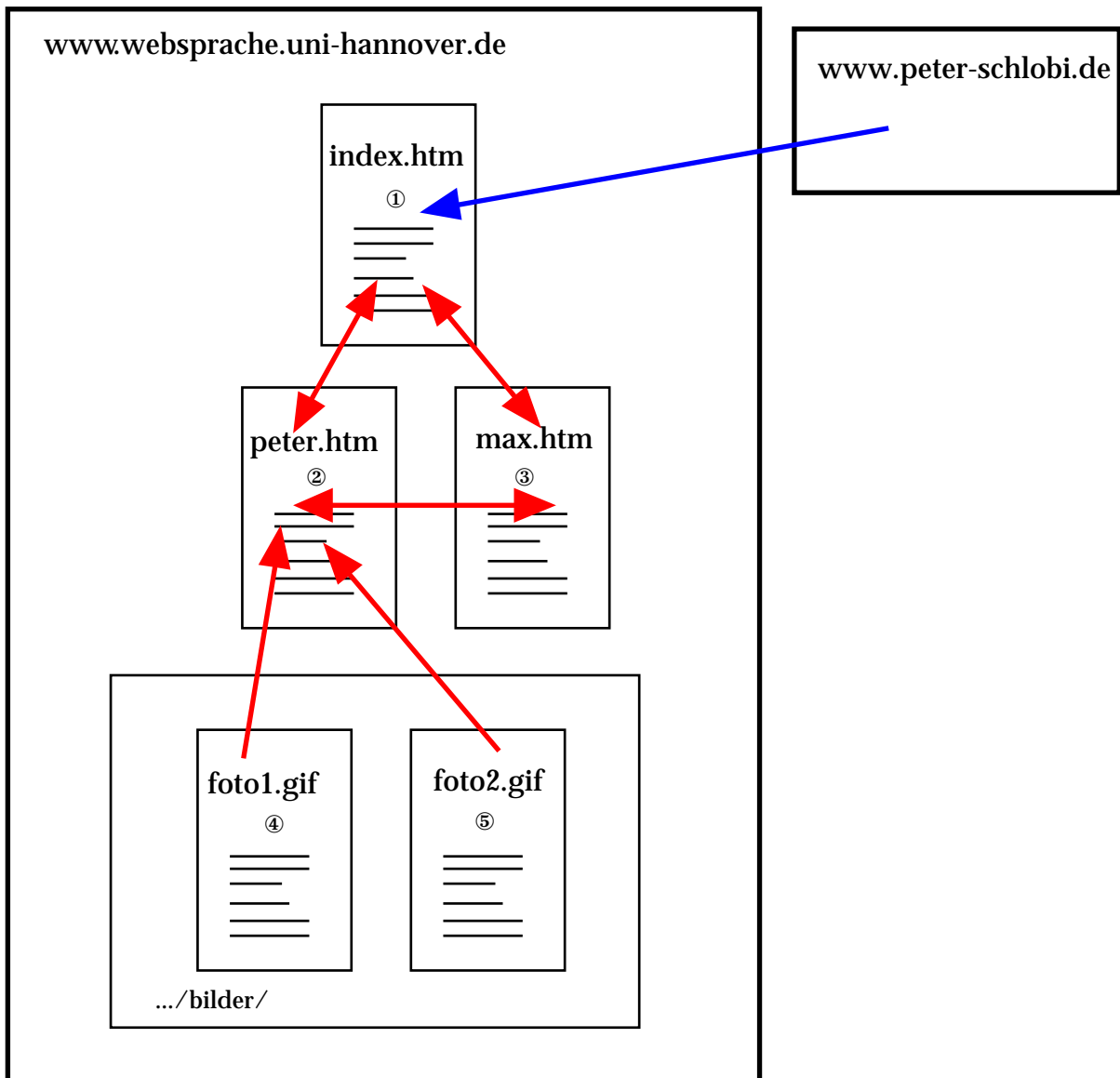


Abb. 1: Aufbau einer Site

Auf dem Websprache Server mit der URL: `www.websprache.uni-hannover.de/` legen wir ein Hauptverzeichnis für alle Textdateien und ein Unterverzeichnis für alle Fotos an. Damit ein User von der Homepage zu Peter bzw. Max gelangen kann, muss ein entsprechender Link gesetzt werden, und will er von dort aus zu den Fotos, muss ebenso ein Link zu dem jeweiligen Bilddokument im Unterverzeichnis `/bilder/` gelegt werden.

Graphentheoretisch fundierte Analyse von Hypertexten

Es handelt sich in beiden Fällen um Links, die innerhalb der eigenen Site verwendet werden. Diese Links nennt man relative Links, es handelt sich also um verzeichnisinterne Links (in Abb. 1 rot gezeichnet). Damit Leser einer Site möglichst schnell zum Ausgangspunkt zurückkehren können, haben Web-Sites häufig auf jeder einzelnen Seite eine Zurück-Funktion auf die Index-Seite.

Links, die von einer Site auf eine andere verweisen, genauer: die eine komplette URL-Angabe beinhalten, werden als absolute Links bezeichnet (in Abb. 1 blau gezeichnet). Der Link von der Peter-Schlobi-Homepage zu der Websprache-Homepage ist ein absoluter Link.

Um eine Web-Site lesen zu können, braucht man einen Browser. Ein Browser legt durch bestimmte Funktionen einen Pfad durch eine oder mehrere Web-Sites an, was für unsere graphentheoretische Betrachtung von Wichtigkeit ist:

1. Zurück (-Funktion): kehrt zur zuletzt aufgerufenen Seite zurück;
2. Vor (-Funktion): zeigt die nachfolgend aufgerufenen Seite an;
3. Anfang (-Funktion): ruft die als Standard eingerichtete Web-Seite auf;
4. Lesezeichen (-Funktion): ermöglicht den direkten Zugriff auf Web-Seiten;
5. Protokoll (-Funktion): ermöglicht Zugriff auf zuvor gebrowste Seiten.

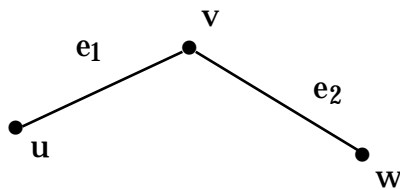
2.2 Das Gerüst eines Hypertextes als Graph

Wir sind davon ausgegangen, dass die Link-Struktur eines Hypertextes als sein Gerüst betrachtet werden kann. Dies soll im folgenden präzisiert und graphentheoretisch modelliert werden.

1. Das Gerüst eines Hypertextes $G(HT)$ wird als ein ungerichteter Graph G insofern betrachtet, als Punkte in $G(HT)$ als Knotenpunkte bzw. Knoten u, v in G definiert sind (allgemein $v_k, k \in \mathbb{N}^+$) und eine Kante $e_k (e_k, k \in \mathbb{N}^+)$ die Verbindung zwischen den Knoten u und v ist. G besteht aus einer Knoten- und einer Kantenmenge, die jeweils alle auftretenden Knoten v_k und Kanten e_k umfasst.

Graphentheoretisch fundierte Analyse von Hypertexten

(1)

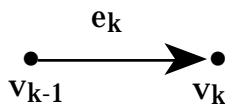


Sprungmarken sind Knoten, von denen aus zu anderen Stellen einer Web-Site Verbindungen hergestellt werden.

Eine Visualisierung von Gerüststrukturen (als ungerichtete Graphen) bietet das ausgezeichnete Programm Hyperbolic Tree (<http://www.inxight.com/>). Dieses zur Navigation in hierarchischen Strukturen entwickelte Programm ist hervorragend zur Analyse von G(HT)-Strukturen geeignet.

2. G ist eine Menge von Kantenfolgen von v_0 nach v_k . In einer Kantenfolge $KF = (v_0, e_k, v_k)$ heißt der Knoten v_0 Anfangsknoten von KF , der Knoten v_k Endknoten von KF . Eine Kantenfolge, die jeden Knoten nur einmal, enthält ist ein Weg. In einem Weg kann also keine Kantenfolge wiederholt werden.

(2)



Ein Link ist eine Kantenfolge wie in (2) mit einer Kantenlänge $L = 1$, die sich aus dem Abstand der Knotenindizes ergibt.

Zentral für unsere Argumentation ist die Tatsache, dass wir Linkstrukturen als einen (virtuellen) Weg begreifen, unabhängig von der Strukturierung in HTML. D.h., die Link-Struktur, wie sie sich aus dem Quell-Code ergibt, ist nicht identisch mit der Beschreibung der Gerüst-Struktur, wie sich im Folgenden zeigen wird.

Bucher (1999: 22f.) wählt einen Ansatz, der wie folgt lautet: Link L (z.B. ein Button, ein Ausdruck) verknüpft A (z. B. eine Site, einen Textabschnitt) mit B (z. B. eine Site, eine Abbildung) im Hinblick auf C (z. B. ein Thema, eine Funktion). Gegenüber der Sichtweise, nach der ein Link A mit B verknüpft, ist in unserer Perspektive die Verknüpfung (Relation) zwischen A und B

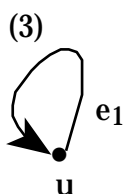
Graphentheoretisch fundierte Analyse von Hypertexten

ein Link, wobei A eine konkrete Absprungstelle ist (sei es ein Wort, eine Phrase, ein Button, ein Icon etc.) und abstrakt ein Anfangsknoten, so wie B ein konkreter Zielpunkt ist (sei es eine Site, Seite, Textstelle, Abbildung etc.), der abstrakt einen Endknoten darstellt. Der Ausgangspunkt ist immer eindeutig definiert über die Absprungmarke, die über das Anklicken aktiviert wird, der Zielpunkt ist letztlich immer eine Seite bzw. Teil einer Seite, sei es die Indexseite einer Site oder eine Textstelle, Mailadresse auf einer Seite. Das, was Bucher als ‚im Hinblick auf C‘ begreift, wird in unserem Ansatz unter das semantisch-textuelle Netzwerk subsumiert. Da die Link-Struktur Teil des HyperTextes ist (s. auch oben), kann diese unter semantisch-pragmatischen Gesichtspunkten analysiert werden wie andere Strukturen des HyperTextes auch.

3. G ist ein Graph, für den gilt: (i) es gibt keine Schlingen und (ii) keine adjazenten Kanten im Anfangspunkt.

Def.: Eine Kante e heißt mit einem Knoten u inzident, wenn u ein Endpunkt von e ist.

Def.: Zwei Kanten e und f, die mit einem gemeinsamen Knoten inzidieren, heißen adjazent. Wenn zwei (oder mehrere) Kanten von G ein- und dieselben Endknoten haben, so werden diese Kanten adjazent im Endknoten genannt. Wenn zwei (oder mehrere) Kanten von G dieselben Anfangsknoten haben, so werden diese Kanten adjazent im Anfangsknoten genannt. Eine Kante mit identischen Endknoten wird als Schlinge bezeichnet.



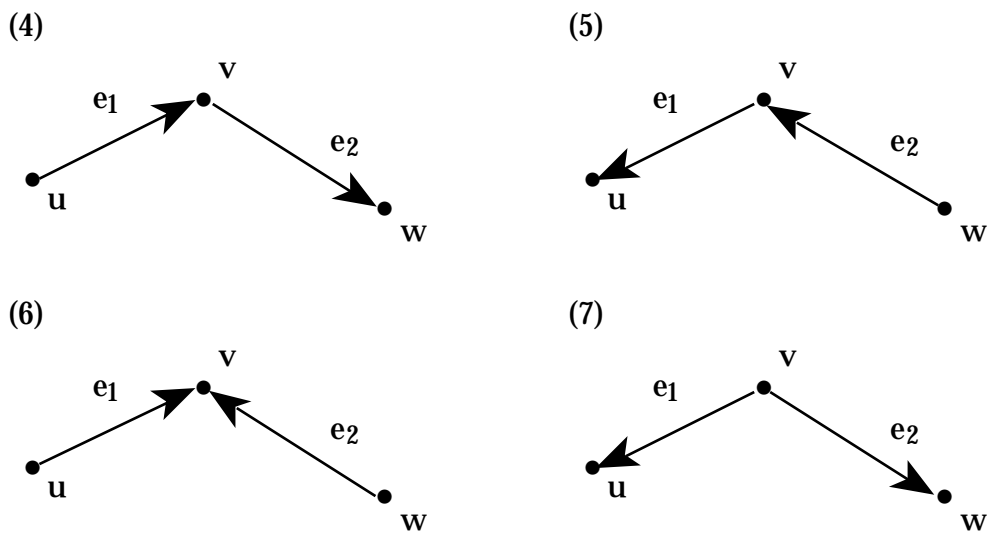
Schlingen können theoretisch auftreten, kommen praktisch aber nicht vor, da ein Selbstverweis in der Regel wenig Sinn macht; wenn auf eine Textmarke geklickt und auf diese (zurück)gesprungen wird, liegt ein reflexiver Link vor. Wird indes von einer Textmarke A auf eine andere Textmarke B gesprungen, so betrachtet wird dies als Link, auch wenn bzgl. der Seite Reflexivität vorliegt.

Es treten jedoch Pseudo-Schlingen auf, wenn beispielsweise auf einer Seite per Klick auf ein Symbol z.B. ein Bild erscheint. Das Bild liegt indes – wie

Graphentheoretisch fundierte Analyse von Hypertexten

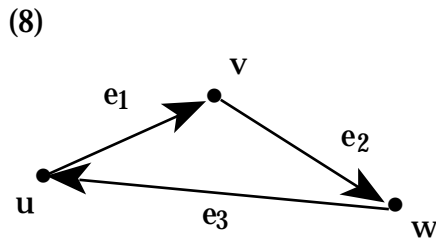
wir in 2.1 dargestellt haben – in einem (Unter)Verzeichnis und als eine andere Datei vor, sodass hier ein Link wie in (2) existiert, wenn es auch für den User erscheint, als läge Reflexivität vor. Dies ist häufig dann der Fall, wenn ohne die Image-Funktion gebrowst wird, aber einzelne Buttons, Bilder etc. per Klick geöffnet werden. Es scheint, als sei die Kantenlänge gleich Null, da kein Weg zurückgelegt wird, sondern sich ein Symbol ‚öffnet‘, in gewisser Weise expandiert wird.

In (4) liegt ein Weg $W_1 = u e_1 v e_2 w$ vor, in (5) ein Weg $W_2 = w e_2 v e_1 u$. Graph (6) stellt den Fall dar, dass von zwei Sprungmarken jeweils ein Link auf einen Zielpunkt (z.B. eine Homepage-Seite) gesetzt sind, also ein Weg mit zwei adjazenten Kanten in v . Der Fall (7) hingegen ist ausgeschlossen, da es nicht möglich ist, von einer Adresse aus zwei oder mehr Links auf einen Zielpunkt zu setzen. Parallele Kanten im Anfangspunkt sind also theoretisch/technisch nicht möglich.



4. Sind zwei Knoten u und v eines Graphen G gegeben, so heißt eine u - v -Kantenfolge geschlossen, wenn $u = v$ ist, oder offen, wenn $u \neq v$. In (4) liegt beispielsweise eine offene Kantenfolge vor, in (8) hingegen eine geschlossene mit $W = (u, e_1, v, e_2, w, e_3, u)$.

Graphentheoretisch fundierte Analyse von Hypertexten



Geschlossene Graphen sind im Internet von Bedeutung, da das Gerüst verlinkter Texträume einen geschlossenen Graphen darstellen kann, was insbesondere für Netzliteratur gilt.

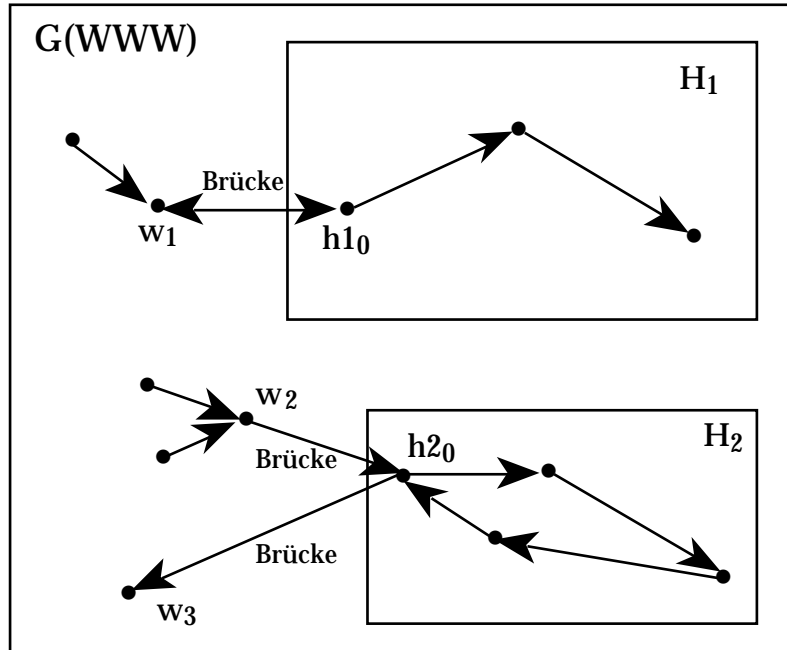
5. Def.: Zwei Knoten, die durch eine Kante verbunden sind, werden als benachbart bezeichnet. In (8) sind alle Knoten Nachbarn voneinander. Ein Link ist demnach eine Kantenfolge zweier benachbarter Knoten.

6. WWW ist ein HyperText, dessen Gerüst wir als einen Graph $G(\text{WWW})$ betrachten. Eine Komponente von W ist ein Untergraph H mit einer benachbarten h_0 - w_k -Folge, wobei h_0 ein Anfangsknoten in H und w ein Knoten in W und nicht in H ist. Es gibt Komponenten mit genau einer h_0 - w_k -Folge. Da jede Site mindesten einen Knoten h_0 hat, hat sie mindestens einen Anfang. (Wenn auf eine Seite gesprungen wird und mehrere Sprungstellen vorliegen, so wird die erste bzw. oberste Sprungstelle als Knoten h_0 angesetzt.)

Jeder h_0 - w -Link bzw. w - h_0 -Link stellt eine Brücke dar zwischen Textraum H und anderen Teilen des WWW (s. 9). Es gibt aber auch Brücken zwischen w_k und h_i ($i \neq 0$). In (9) liegt eine isomorphe (umkehrbare) Brücke vor zwischen $G(\text{WWW}) \setminus H_1$ und H_1 vor, da der Weg $W_1 = (w_1, h_1)$ umkehrbar ist in $W_2 = (h_1, w_1)$, also ein Link von w_1 auf h_1 und von h_1 auf w_1 besteht. Die Brücken zwischen H_2 und $G(\text{WWW}) \setminus H_2$ sind hingegen nicht-isomorph.

Graphentheoretisch fundierte Analyse von Hypertexten

(9)

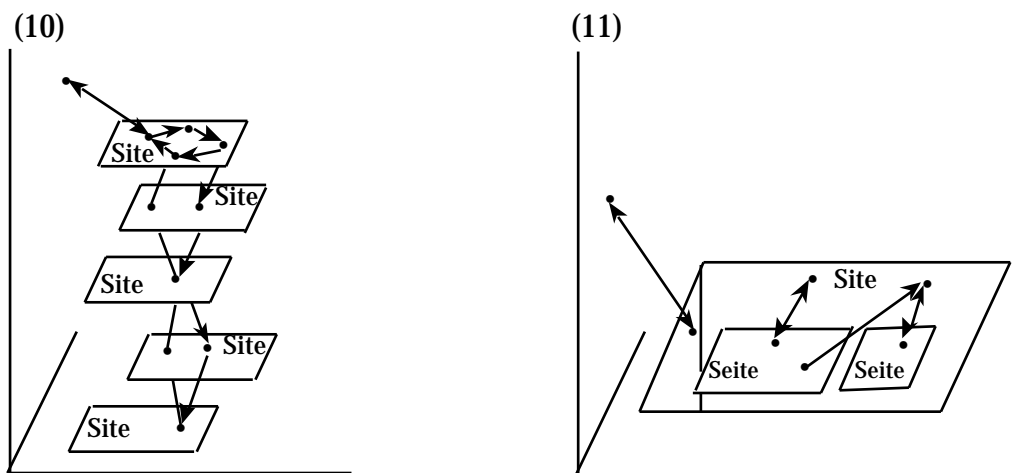


Es gibt zudem Brücken zwischen Untergraphen in H_x und Knoten in H_x , die nicht Teil des Untergraphen sind (analog zu oben).

7. Der Grad der Vernetzung, also die Multiplexität eines Graphen, läßt sich über den Grad eines Knoten ermitteln. In h_{10} aus (9) sind zwei Kanten adjazent im Endknoten und eine Kante ist adjazent im Anfangsknoten; in h_{20} sind zwei Kanten adjazent im Endknoten und zwei Kanten adjazent im Anfangsknoten. Im ersten Fall ist die Anzahl der mit h_{10} inzidenten Kanten 3, im zweiten Fall 4. Allgemein gilt: Der Grad $d(v)$ von v entspricht der Anzahl der mit v inzidenten Kanten (von G); der Grad der Adjazenz im Anfangsknoten $a(v)$ entspricht der Anzahl der mit v adjazenten Kanten im Anfangsknoten, der Grad der Adjazenz im Endknoten $e(v)$ entspricht der Anzahl der mit v adjazenten Kanten im Endknoten. Also: $a(h_{10}) = 1$; $e(h_{10}) = 2$; $a(h_{20}) = 2$; $e(h_{20}) = 2$. Wir können nunmehr die Multiplexität unterschiedlicher Graphen H_x berechnen, indem wir die Grade aller Knoten eines Graphen H_i aufsummieren und mit der Summe aller Grade eines Graphen H_j vergleichen. Je höher diese Summe der Grade, desto multiplexer der Graph.

Graphentheoretisch fundierte Analyse von Hypertexten

8. In Abhängigkeit von den Verknüpfungsebenen sind Graphen n-dimensional strukturiert. In (10) haben wir Verknüpfungsebenen auf einzelnen Sites, die untereinander verknüpft sind. Das Verknüpfungspotential auf einer Hierarchiestufe betrachten wir als 2-dimensional. In (10) haben wir fünf Sites, die site-extern und somit dreidimensional verknüpft sind. In (11) bilden die Seiten Ebenen, die seiten-extern mit der Site verknüpft sind (Erhöhung der Dimension), die wiederum site-extern verknüpft ist (erneute Erhöhung der Dimension).

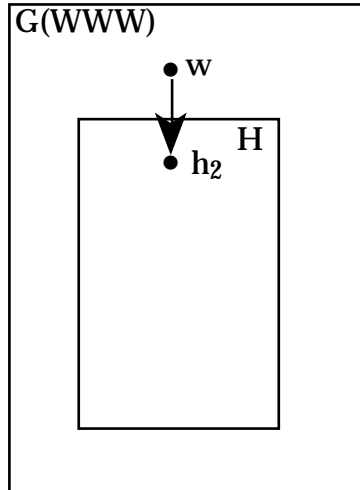


Die Basisstrukturen, genauer: das Gerüst kann nunmehr beschrieben und klassifiziert werden, wobei Pseudo-Schlingen in die Analyse integriert sind. Ein minimaler Graph ist dann gegeben, wenn ein Knoten h_0 aus der w - h_0 -Kantenfolge vorliegt, der den Ausgangspunkt zu einem oder weiteren Knoten bildet. Werden zwei Knoten (h_0, h_i) ($i > 0$) miteinander verbunden, so liegt eine Kante/Linie in H vor. Wir wollen im folgenden einzelne Graphen mit ihren Brücken betrachten.

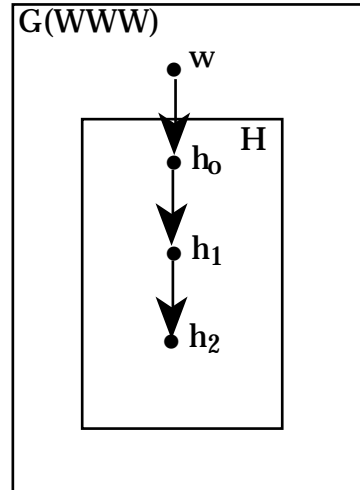
In (12) liegt ein Gerüst mit einem Knoten (h_0) vor. Der Weg führt in eine Sackgasse, aus H führt kein Weg mehr heraus – ausser über das Browsen (s.u.). Strukturen dieser Art finden sich bei ‚linearen‘ Texten, die in das Netz gestellt werden, ohne dass weitere Verknüpfungen stattfinden, z.B. Texte zum Herunterladen. In (13) liegt ein Baumgraph vor, ein zusammenhängender azyklischer Graph. Der Baum des Gerüsts hat drei Knoten, von dem einer (h_0) durch eine Brücke mit einem Knoten w verbunden ist.

Graphentheoretisch fundierte Analyse von Hypertexten

(12)

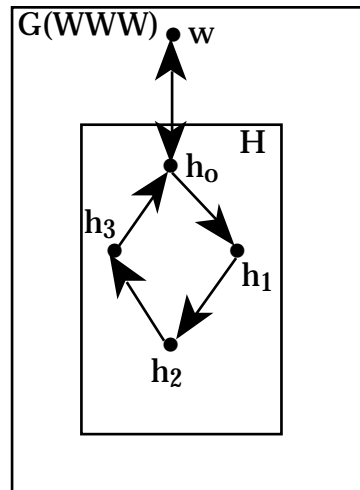


(13)

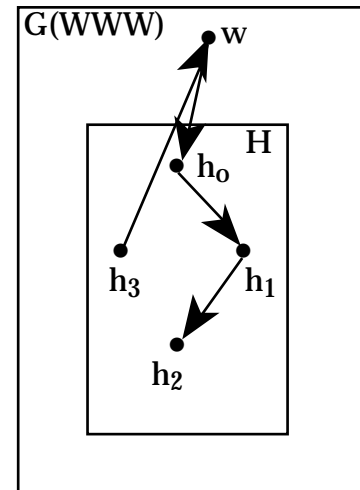


In (14) ist ein Ringgraph in H über eine isomorphe Brücke mit $G(WWW)$ – H verbunden, in (15) liegt ein zusammenhängender Graph vor, in dem über zwei Brücken ($W_1 = w h_0$ und $W_2 = h_3 w$) der Untergraph H mit einem tableauexternen Knoten verbunden ist.

(14)



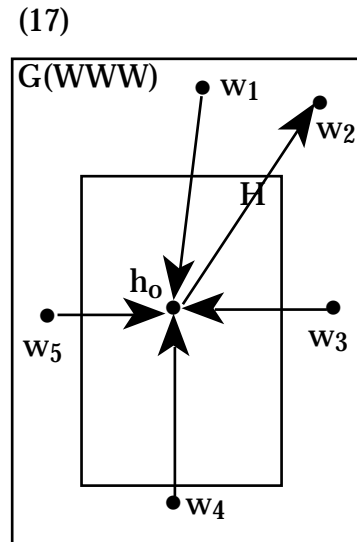
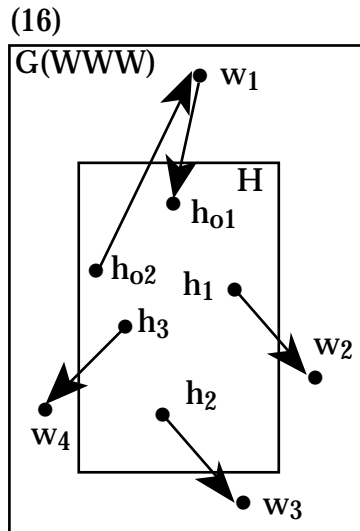
(15)



In (16) stellt das Gerüst einen nicht-zusammenhängenden Graphen dar, z.B. ein Dokument, das allein über seitenexterne Links mit anderen Seiten, Sites verbunden ist. (17) zeigt einen multiplexen Graphen mit einem Grad $d(h_0) = 5$, es bestehen vier Links von außen auf das Dokument (externe Links =

Graphentheoretisch fundierte Analyse von Hypertexten

Brücken) und ein Weg von H weg, also eine Brücke von h_0 zu einem externen Knoten.



Texte (Seiten) einer Site, verschiedene Sites lassen sich hinsichtlich ihres Gerüsts analysieren und klassifizieren. Für Hyperfiction (Netzliteratur) lassen sich folgende Feststellungen treffen bzw. Hypothesen aufstellen:

(a) Netzliteratur hat immer einen Anfang, genauer: mindestens eine $u-h_0$ -Kantenfolge.

(b) Netzliteratur bildet in der Regel einen geschlossenen Textraum insofern als ein Untergraph vorliegt mit einer Brücke.

(c) Das Gerüst eines Textes besteht aus Baum- und/oder Ringgraphen (vgl. Graph 13 bzw. 14). Kommentar: Netzliteraturen lassen sich hinsichtlich ihrer Gerüste analysieren, differenzieren und typologisieren! Hier wären in einem ersten Schritt Hyperfictiondokumente nach den vorliegenden Analyse Kriterien zu analysieren. Aus der empirischen Analyse ließen sich prototypische Strukturen für unterschiedliche HyperTextsorten etc. gewinnen.

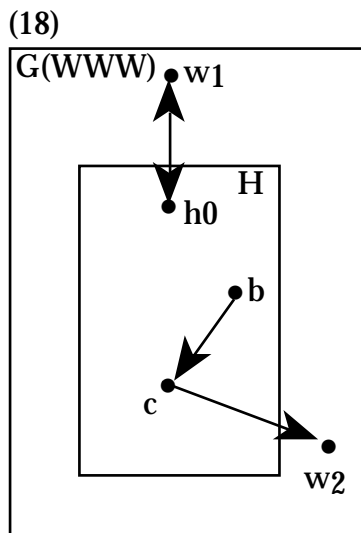
9. Ein Gerüst gibt mögliche Pfadstrukturen, beim Browsen werden konkrete Pfade auf der Folie eines Gerüsts realisiert; aus allen möglichen Wegen wird ein Weg realisiert. Der realisierte Benutzerpfad ist einerseits temporal-linear, andererseits strukturell durch das Gerüst vordefiniert. Da beim Browsing in einer Kantenfolge die Knoten v_{k-1} v_k benachbart sind, und die Menge aller

Graphentheoretisch fundierte Analyse von Hypertexten

benachbarten Knoten einer Browsing-Kantenfolge symmetrisch sind, gilt auch $v_k v_{k-1}$.

Kommentar: Die benachbarten Knoten können übersprungen werden, es gilt aber prinzipiell die Nachbarschaftsbedingung. Wichtig ist, dass beim Browsen eine ausgezeichnete Kantenfolge entsteht, die immer umkehrbar ist und somit zum Ausgangspunkt zurückführt, auch wenn ein Graph in einer Sackgasse endet. Zudem kann immer auf die Startseite gesprungen werden.

Zentrale Navigationselemente sind die Vor- und Zurückfunktion, die iterativ angewendet werden können. Nehmen wir an, wir haben eine Seite mit folgender Link-Struktur.



Aus der Link-Struktur können unterschiedliche Wege beschriftet werden, wobei wir iterativ-isomorphe Relationen durch $a \rightarrow b$ darstellen wollen. $a \rightarrow b$ besagt, dass zwischen den Knoten a und b beliebig oft hin- und hergesprungen werden kann. Auch beim Browsen werden Untergraphen über Brücken erreicht. Unter anderem sind folgende Wege möglich: $W_1 = (w_1 \rightarrow h_0)$, $W_2 = (w_1 \rightarrow h_0 \rightarrow b \rightarrow c \rightarrow w_2)$, $W_3 = (w_1 \rightarrow h_0 \rightarrow c \rightarrow w_2)$ usw. Zudem kann von jedem Knoten aus zu der Startseite gesprungen werden und die Pfade werden im Browser protokolliert, sodass durch Klicken auf die Protokolladressen zusätzlich Knotenpunkte erreicht werden können, z.B. $W_4 = (c, w_1)$.

Ein Browser ist also ein Navigationsinstrument durch das komplex strukturierte Gerüst. Aus dem Navigationsverhalten von Usern, das sich in

Graphentheoretisch fundierte Analyse von Hypertexten

konkreten Browserpfaden niederschlägt, lassen sich Rückschlüsse auf die Rezeption von Hypertexten ziehen, genauer als über AdClick-Raten und PageImpressions.

2.3 Semantisch-textuelles Netzwerk

Unabhängig vom Gerüst bestehen zwischen Inhalten von Dokumenten im Netz semantisch-textuelle Relationen, wobei das Gerüst auch unter semantisch-textuellen Beziehungen analysiert werden kann, die mit Link-Strukturen verbunden, aber semantisch-textuelle Relationen nicht notwendigerweise mit Gerüststrukturen gekoppelt sind. Wir wollen dies an einem einfachen Beispiel verdeutlichen.

Nehmen wir an, auf einer Site befindet sich eine Seite wie die folgende:

<p>Klaus sperrte die Tür auf, aber seine Ehefrau sperrte sie sofort wieder zu. „Genug!“, schrie Klaus und mit _____ dem <u>linken Fuß</u> eintrat ER *krachsplitter* die <u>Tür</u>.</p>

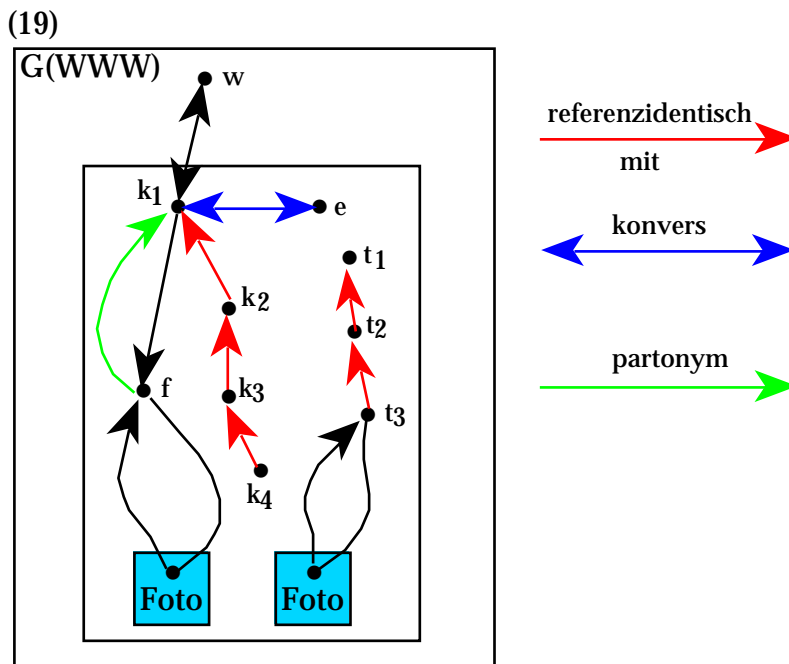
Der Text weist folgende Link-Struktur auf: Wenn wir auf die Textmarke **Klaus** klicken, dann springt der Cursor auf die Textstelle mit linken Fuß, bei Klick auf linken Fuß wird ein Foto mit dem Fuß von Klaus geöffnet, bei Klick auf Tür wird das Foto einer zersplitterten Tür geöffnet. (Die Fotos wollen wir als Pseudo-Schlingen in den Graphen der Seite einbetten, s. 19). Unabhängig von der Link-Struktur können wir den Text nach semantischen Relationen analysieren und die Relationen wiederum graphisch darstellen. Wir wollen folgende grundlegende und eindeutig zu definierende Sinnrelationen betrachten (im einzelnen s. Lutzeier 1995: 71ff.)

- (a) Konversen-Relation zwischen Klaus (Ehemann) und (seiner) Ehefrau (blauer Pfeil) unter dem Aspekt ‚Verwandtschaft‘;
- (b) Partonymie-Relation zwischen Klaus und (seinem) linken Fuß (grüner Pfeil) unter dem Aspekt ‚Körperteil‘.

Graphentheoretisch fundierte Analyse von Hypertexten

Andere Sinnrelationen wie die Reversivitäts-Relation zwischen auf- und zusprechen lassen wir außer acht. Zusätzlich wollen wir noch eine diskurssemantische Relation betrachten, und zwar (c) die Relation der Topikontinuirung, also zu einer NP referenzidentische, anaphorische Ausdrücke, seien sie 1. nominal oder 2. pronominal (roter Pfeil). In unserem Beispiel wird Klaus wiederholt und pronominal weitergeführt (seine, ER), auf die Tür wird pronominal und anschließend nominal referiert. Auf das Problem, dass das Possessivpronomen einerseits der Topikkontinuirung dient, andererseits die Konversenrelation determiniert – es ist nicht irgendeine Ehefrau, sondern die von Klaus gemeint – gehen wir der Einfachheit halber nicht ein.

Eine graphische Darstellung des Gerüsts und des semantisch-textuellen Netzwerkes findet sich in (19). Die sprachlichen Elemente, die durch semantisch-textuelle Relationen verbunden sind, sind hier ebenfalls als Knoten definiert: die Elemente in der Topikprogression von ‚Klaus‘ durch k_1 {Klaus}, k_2 {seine}, k_3 {Klaus}, k_4 {Er} von ‚die Tür‘ durch t_1 {Tür}, t_2 {sie}, t_3 {Tür}, ‚die Ehefrau‘ durch e und ‚der linke Fuß von Klaus‘ durch f .



Interessant ist hier, dass der Link von ‚Klaus‘ auf seinen linken Fuß‘ ebenfalls mit die Partonymie-Relation gekoppelt ist, dass also in den Knoten k_1 und f jeweils die Knoten aus dem Gerüst (k'_1 , f') und dem semantisch-

Graphentheoretisch fundierte Analyse von Hypertexten

textuellen Netzwerk (k_1, f) verschmolzen sind. (So integrierte Graphen weisen auch adjazente Kanten im Anfangspunkt auf.) Hier wäre weiter zu fragen und zu untersuchen, wie Link-Strukturen und semantisch-textuelle Relationen zusammenhängen. Unter graphentheoretischen Gesichtspunkten ist interessant und hervorzuheben, dass Graphen unter bestimmten Voraussetzungen verschmolzen und vereinigt werden können; die Knoten- und Kantenmengen des Gerüsts und des semantischen Netzwerkes können vereinigt werden.

In der Linguistik liegt eine Reihe von Ansätzen vor, semantische Netze und Textrelationen zu beschreiben. Die Wahl des jeweiligen Ansatzes hängt von theoretischen Vorannahmen ab. Für die HyperTextanalyse ist Voraussetzung, dass ein relationaler Ansatz gewählt wird, in dem die durch die klar definierte Relationen verbundenen Elemente klar definiert sind (was trivial erscheint, in der Praxis aber keineswegs so trivial ist).

3. Zusammenfassung und Perspektiven

Wir haben einen Ansatz gewählt, in dem die Analyse von HyperTexten als die Analyse von Strukturen von Hypertexten begriffen wird. Hinsichtlich der Strukturanalyse haben wir grundsätzlich unterschieden zwischen der Analyse der Link-Struktur (Analyse des Gerüsts eines HT) und der Analyse semantisch-textueller Strukturen. In beiden Fällen wählten wir einen relationalen Ansatz, der sich graphentheoretisch fundieren läßt. Die graphentheoretische Fundierung bringt eine Reihe von Vorteilen mit sich:

- (1) Begriffe wie Delinearität, Baumstruktur, Rhizom lassen sich operativ fassen.
- (2) Die Graphenanalyse läßt sich formalisieren, d.h. mathematisch formulieren. (So lassen sich Graphen als Matrizen darstellen.)
- (3) Die Modellierung ist empirisch überprüfbar.
- (4) Mediale und textuelle Struktur lassen sich aufeinander beziehen.

Mengen von Texten sowie Textsegmente lassen sich graphentheoretisch analysieren (Unter- und Übergraphen). Damit läßt sich die Frage, was nun ein Hypertext im WWW sei, umformulieren in die Frage, welche Strukturen und Strukturebenen im WWW auftreten. Hierzu liegen bisher kaum Analysen vor und hier sollten unserer Meinung nach konkrete Analysen ansetzen. Sites, Dokumente, sog. Hyperfiction sollten einer konkreten Strukturanalyse und ihre Resultate einer Typologisierung unterzogen werden. Es wird sich unserer Meinung nach erweisen, dass zum einen sehr

Graphentheoretisch fundierte Analyse von Hypertexten

heterogene Strukturmuster vorliegen, die mit wenigen Metaphern wie Linie, Baum und Rhizom nicht ausreichend beschrieben werden können, zum anderen, dass bestimmte HyperTextsorten mit bestimmten Strukturmustern verbunden sind. Es wird empirischen Untersuchungen vorbehalten sein, dies zu veri- bzw. zu falsifizieren.

4. Literatur

Bucher, Hans-Jürgen (1999). „Die Zeitung als Hypertext. Verstehensprobleme und Gestaltungsprinzipien für Online-Zeitungen.“ In: Text im digitalen Medium. Linguistische Aspekte von Textdesign, Texttechnologien und Hypertext Engineering. Hrsg. von Lobin, Henning. Opladen, S. 9-32.

Clark, John und Derek Allan Holton (1994). Graphentheorie. Grundlagen und Anwendung. Heidelberg.

Deleuze, Gilles und Guattari, Félix (1977). Rhizom. Berlin.

Krajewski, Markus (1997). „Spür-Sinn. Was heißt einen Hypertext lesen?“ In: Soziologie des Internet. Handeln im elektronischen Web-Werk. Hrsg. von Lorenz Gräf und Markus Krajewski. Frankfurt am Main, S. 60-78.

Gabriel, Norbert (1997). Kulturwissenschaft und Neue Medien. Wissensvermittlung im digitalen Zeitalter. Darmstadt.

Lobin, Henning (1999). „Linguistische Repräsentation komplexer Inhalte für die hypermediale Wissensvermittlung.“ In: Text im digitalen Medium. Linguistische Aspekte von Textdesign, Texttechnologien und Hypertext Engineering. Hrsg. von Lobin, Henning, Opladen, S. 155-177.

Lutzeier, Rolf Peter (1995). Lexikologie. Ein Arbeitsbuch. Tübingen.

Runkehl, Jens, Schlobinski, Peter und Siever, Torsten (1998). Sprache und Kommunikation im Internet. Überblick und Analysen. Opladen.

Storrer, Angelika (1999). „Kohärenz in Text und Hypertext.“ In: Text im digitalen Medium. Linguistische Aspekte von Textdesign, Texttechnologien und Hypertext Engineering. Hrsg. von Lobin, Henning, Opladen, S. 33-65.

Graphentheoretisch fundierte Analyse von Hypertexten

Suter, Beat (1999). WWW Hyperfiction Liste Deutsch. Online im Internet: URL: <http://www.update.ch/beluga/hypfic.htm> [Stand 13.1.1999]

Vesper, Sebastian (1998). Das Internet als Medium. Auftrittsanalysen und neue Medien. Bardowick.

Wingert, Bernd (1996). „Die neue am Lesen?“ In: Kursbuch Neue Medien. Trends in Wirtschaft und Politik, Wissenschaft und Kultur. Mannheim, S. 112-129.

Wirth, Uwe (1997). „Literatur im Internet. Oder: Wen kümmert's, wer liest?“ In: Mythos Internet. Hrsg. von Stefan Münker und Alexander Roesler. Fr. am Main, S. 319-337.

5. Glossar

Ad-Click-Rate. Verhältnis zwischen Sichtkontakten und AdClicks, der Anzahl der Klicks auf einen Banner.

PageImpressions. Anzahl der Abrufe von WWW-Seiten.

Rhizom (gr. Wurzelstock). Der Begriff stammt von dem Philosophen Gilles Deleuze und dem Psychoanalytiker Félix Guattari. Zentral ist das gegen Chomskysche Binarität gerichtete „Prinzip der Konnexion und Heterogenität. Jeder beliebige Teil eines Rhizoms kann und muß mit jedem anderen verbunden werden.“ (Deleuze/Guattari 1977: 11).

URL (Uniform Resource Locator). Vollständige Angabe einer Internet-Adresse, z.B. <http://www.websprache.uni-hannover.de/>.