Proceedings of the First International Symposium: Category Theory Applied to Computation and Control, published by the Mathematics Department and the Department of Computer and Information Science, University of Massachusetts at Amherst, 1974.

STRUKTURELLE VERWANDTSCHAFTEN VON SEMI-THUE-SYSTEMEN

Günter Hotz
Angewandte Mathematik
und Informatik

Universität des Saarlandes
66 Saarbrücken, W-Germany

Semi-Thue-Systems and Chomsky Grammars (C.L.) are associated with free monoidal categories (X-categories) in a natural way [Ho,Be]. We specialize the weak equivalence of C.L. by introducing different relationships $R(\tilde{F})$ based on the structure of the associated X-categories. These relationships are generated by chains of functors from a functor category $\tilde{F}$. The decisi problem if two grammars are $R(\tilde{F})$-equivalent is strongly connect to the question if products in $\tilde{F}$ exist and if functors $\phi \in \tilde{F}$ ar surjective. In this paper we report on some results which we have gained about this topics in the last years.

Semi-Thue-Systeme (STS) haben durch die Arbeiten von N. Chomsky für die Linguistik eine hervorragende Bedeutung erlangt. Der Ansatz zur Beschreibung der Syntax von Programmiersprachen in der Backus-Notation stellt vielleicht eine noch wichtigere Anwendung dieser Systeme dar. Durch Chomsky haben die (STS) eine kombinatorische Klassifikation erfahren, an die sich eine ausführliche Diskussion des Sprachumfangs dieser Sprachklassen anschloß. Ein großer Teil der Ergebnisse ist negativer Art: Es existieren zur Lösung der interessierenden Probleme keine universellen Algorithmen. Neben dem Wortproblem hat die Frage nach der Gleichheit zweier Chomsky-Sprachen (schwache Äquivalenz) eine besondere Aufmerksamkeit erfahren. Resultate über die strukturelle Verwandtschaft verschiedener Sprachen existieren kaum. Selbst bei der Herleitung von Normalformentheoremen für Grammatiken hat man sich mit der Feststellung der schwachen Äquivalenz der zugehörigen Sprachklassen begnügt. Strukturelle Aussagen enden im allgemeinen mit der Diskussion der Mehrdeutigkeitsfragen der Grammatiken.

Die Bedeutung dieser Fragestellung ist aber eminent wichtig. Findet doch die Interpretation der gesprochenen Sätze großteils, und die der Programme einer Programmiersprache fast ausschließlich, über die durch die Grammatiken definierte Ableitungsstruktur des Satzes, bzw. des Programms statt. Dies ist die Motivation der Untersuchungen, über die hier kurz berichtet werden soll.

Hierbei gehen wir auf Normalformtheoreme nicht ein, sondern verweisen in diesem Zusammenhang auf eine jüngst erschienene Arbeit [Ha] und das Buch [Ho-Cl], das die Herstellung von Normalformen unter kategoriellem Aspekt betrachtet.

Da es uns im folgenden vorwiegend auf die Herausarbeitung der Ideen ankommt, werden wir weder größte Allge-

meinheit noch Vollständigkeit anstreben.
Sei $S = (A,P,\rightarrow)$ ein (STS), d.h. A eine endliche Menge, A^* freies Monoid über A und $P \subset A^* \times A^*$ ebenfalls endlich. Die Relation P wird in bekannter Weise zu der Relation $\rightarrow$ auf A^* fortgesetzt. Gilt für $w,v \in A^*$ die Beziehung $w \rightarrow v$, dann kann man nach allen Verfahren F fragen, $w \rightarrow v$ aus P zu konstruieren, Bezeichnung: $w \overset{F}{\rightarrow} v$. Jede solche Konstruktion F kann durch einen Ausdruck in o und x über P beschrieben werden, indem man definiert

$$w \overset{F}{\rightarrow} v,\ v \overset{G}{\rightarrow} u$$

$$\Rightarrow w \overset{GoF}{\rightarrow} u$$

und

$$w \overset{F}{\rightarrow} w',\ v \overset{G}{\rightarrow} v'$$

$$\Rightarrow wv \overset{GxF}{\rightarrow} w'v.$$

Rechnet man mit diesen Ausdrücken unter Verwendung der Regel

$$(F_1 \times F_2) \circ (G_1 \times G_2)$$
$$= (F_1 \circ G_1) \times (F_2 \circ G_2).$$

Falls $F_1 \circ G_1$ und $F_2 \circ G_2$ definiert sind, dann bilden die Klassen, die man durch Zusammenfassung der Ausdrücke erhält, die durch Anwendung dieser Rechenregel ineinander überführt werden können, eine "freie" monoidale Kategorie [Mc 1], die wir als X-Kategorie [Ho 1] $\mathcal{F} = \mathcal{F}(S)$ bezeichnen. S und $\mathcal{F}$ bestimmen sich gegenseitig eindeutig. In $\mathcal{F}$ ist die "Struktur von S" eingefangen. Unter strukturellen Aussagen über S verstehen wir Aussagen über $\mathcal{F}$.
Seien nun etwa S_1 und S_2 (STS) und $\mathcal{F}_1$ und $\mathcal{F}_2$ die zugehörigen X-Kategorien. S_1 und S_2 kann man als <u>elementar</u> <u>verwandt</u> bezeichnen, wenn gilt:
$\mathcal{F}_1$ ist eine Unter-X-Kategorie von $\mathcal{F}_2$: $(\mathcal{F}_1 \subset \mathcal{F}_2)$.

Eine weitere <u>elementare</u> Verwandtschaft kann in der Existenz von Funktoren

$$\phi : \mathcal{F}_1 \rightarrow \mathcal{F}_2$$

bestehen, das heißt in strukturerhaltenden Abbildungen von $\mathcal{F}_1$ in $\mathcal{F}_2$.

Diese Begriffe werden durch in der Kategorientheorie übliche Differenzierungen verfeinert. Wir formulieren unsere Resultate, ohne solche zusätlichen Forderungen im einzelnen anzugeben.

Verwendet man Semi-Thue-Systeme zur Definition formaler Sprachen

$L(S,T,s) = \{w \in T^* \mid s \to w\}$

mit

$T \subset A, s \in A - T, P\ (A - T)^* \times A^*$,

so wird man an die obigen Verwandtschaftsbeziehungen noch Forderungen stellen, die s und T^* betreffen.

Wir definieren nun ohne vollständig zu sein

$L = (S,T,s)$ und $L' = (S',T',s')$ heißen <u>verwandt</u>, wenn es eine Kette $L = L_1, L_2, \ldots, L_k = L'$ gibt, so daß L_i und L_{i+1} oder L_{i+1} und L_i für $i = 1,\ldots,k - 1$ elementar verwandt sind im obigen Sinne.

Es stellt sich die Frage:

<u>Problem 1</u>

Unter welchen Voraussetzungen an S und die elementaren Verwandtschaften kann man die Verwandtschafts-Ketten verkürzen?

Unter welchen Voraussetzungen sind stets Verkürzungen auf Längen $\leq$ m möglich, wo m unabhängig von den Paaren L,L' ist.

<u>Problem 2</u>

Man gehe das Problem 1 unter dem Gesichtspunkt der effektiven Konstruierbarkeit oder - noch schärfer - unter dem Gesichtspunkt der Komplexität an.

Es liegen zu beiden Problemen einige Resultate vor, die jetzt kurz geschildert werden.

Man schränkt die Fragestellung auf einseitig lineare (STS) ein, dann kann man zeigen:

Ist

$L(S,T,s) = L(S',T,s')$,

dann sind

$L = (S,T,s)$ und $L' = (S',T',s)$ verwandt. Es gilt darüber hinaus:

Es gibt unter dieser Voraussetzung stets eine Kette $L = L_1,\ldots,L_9 = L'$, so daß $L_i = L_{i+1}$ und L_i elementar verwandt zu $L_{i\pm 1}$ ist [Ho 2].

Es liegen Ergebnisse vor, wann gewisse elementare Verwandtschaften effektiv entscheidbar sind [Ber], [Schn], [Ho 3]. Als Beispiel werde ein Satz aus [Ber] angegeben, der sich auf die Klasse der kontextfreien Systeme bezieht.
Für die Menge der Funktoren $\phi : F_1 \to F_2$, die durch Fortsetzung einer Abbildung $\phi' : P_1 \to F_2$ gewonnen werden, ist es entscheidbar, ob ϕ surjektiv oder injektiv ist.
Dieser Satz stellt eine Verallgemeinerung eines Resultats aus [Schn] dar, wo der Satz nur für nicht "erweiternde" Funktoren bewiesen wurde und eines Satzes aus [Ho 3], wo dieser Satz unter Einschränkung auf lineare Systeme gezeigt wurde. Für kontextsensitive Systeme sind diese Eigenschaften nicht mehr generell entscheidbar.
Klassen von Elementarverwandtschaften, für die ähnlich scharfe Sätze wie der aus [Ho 2] zitierte Satz beweisbar sind, werden in [Ba-Ho] angegeben.

Die "Verkürzungssätze" sind Sätze über die Existenz von Produkten in Funktor-Kategorien der X-Kategorien. Diese Sätze stehen in engstem Zusammenhang mit der Frage nach der Zerlegbarkeit von Systemen S. Hierzu wurden Kriterien in [Cl-Wa] und [Schn-Wa] sowie [Wa] gewonnen. Ohne die Bedingung des Nichterweiterns an die Funktoren führen Produktkonstruktionen aus der Kategorie heraus. Man erhält Sprachen L mit card $P = \infty$.

REFERENCES

[Ba,Ho] Bartholomes F. und Hotz G.: "Homomorphismen und Reduktionen linearer Sprachen", Lecture Notes in Operations Research and Mathematical Systems, Vol. 32, 143 S., (1970).

[Be] Benson D.: "Syntax and Semantics: A Categorical View", Inf. and Control 17, 145-160, (1970).

[Ber] Bertsch E.: "Surjectivity of Functors on Grammars", Berichte des Sonderforschungsbereichs Elektronische Sprachforschung.

[Cl-Wa] Claus V. und Walter H.: "Zerlegungen von Semi-Thue-Systemen", Computing 4, 107-124, (1969).

[Gr] Greibach S. A.: "A New Normal Form Theorem For Contextfree Phrase Structure Grammars", Journal ACM 12, 45-52, (1965).

[Ha] Harrison M. A.: "On Covers and Precedence Analysis", GI-Jahrestagung Hamburg, 1973; Ed. G. Goos und J. Hartmanis: Lecture Notes in Computer Science Vol. 1, 1-17, Springer-Verlag, Berlin, Heidelberg, New York.

[Ho-Cl] Hotz G. und Claus V.: "Automatentheorie und Formale Sprachen III", Bibliographisches Institut Mannheim, 240 S., (1971).

[Ho 1] Hotz G.: "Algebraisierung des Syntheseproblems von Schaltkreisen" I und II; Elektronische Informationsverarbeitung und Kybernetik (EIK) 1, 185-231, (1965).
Hotz G.: "Eindeutigkeit und Mehrdeutigkeit formaler Sprachen", (EIK) 2, 235-246, (1966).

[Ho 2] Hotz G.: "Übertragung automatentheoretischer Sätze auf Chomsky-Sprachen", Computing 4, 30-42, (1969).

[Ho 3] Hotz G.: "Reduktionssätze über eine Klasse formaler Sprachen mit endlich vielen Zuständen", Math. Zeitschrift 104, 205-221, (1968).

[Ho 4] Hotz G.: "Erzeugung formaler Sprachen durch gekoppel Ersetzungen", 4. Colloquium über Automatentheorie, München 1967. Ed. F. L. Bauer und K. Samelson, verlegt durch TU München.

[Mcl] MacLane S.: "Kategorien - Begriffssprache und mathem Theorie", Springer-Verlag, Berlin, Heidelberg, New York, 295 S., (1972).

[Schn] Schnorr C.-P.: "Transformational Classes of Grammars Inf. and Control 14, 252-277, (1969).

[Schn-Wa] Schnorr C.-P. und Walter H.: "Pullbackkonstruktionen bei Semi-Thue-Systemen", (EIK) 5, 27-36, (1969).

[Vo 1] Vollmerhaus W.: "Die Zerlegung von kontextfreien Semi-Thue-Systemen mit Anwendung auf das Analyseproblem kontextfreier Sprachen", Beiträge zur Linguistik und Informationsverarbeitung, 12, (1967).

[Vo 2] Vollmerhaus W.: "Über die Zerlegung von freien X-Kategorien", 4. Colloquium über Automatentheorie, München 1967, Ed. F. L. Bauer und K. Samelson, verlegt durch TU München.

[Wa 1] Walter H.: "Pullbackkonstruktionen bei Semi-Thue-Systemen", 4. Collquium über Automatentheorie, 'ünchen 1967, verlegt durch TU München.

[Wa 2] Walter H.: "Verallgemeinerte Pullbackkonstruktionen bei Semi-Thue-Systemen und Grammatiken", (EIK) 6, 239-254, (1970).

[Wa 3] Walter H.: "Einige Topologische Aspekte der syntakti· schen Analyse und Übersetzung bei Chomsky-Grammatikeı erscheint in J. Comp. and System Science.