

XII. LINGUISTICS*

Prof. R. Jacobson
Prof. A. N. Chomsky
Prof. M. Halle

Dr. G. H. Matthews
L. M. Kampf
J. L. Katz
T. M. Lightner

P. Postal
Carlota S. Smith
D. Walker

A. DETERMINERS

This report is a discussion of restrictive and appositive relative clauses with respect to determiners. I assume that relative clauses are adjoined to the determiner of a noun phrase and attempt to answer certain questions implied by such an assumption.

The grammar must allow only correct adjoining of a relative clause to a noun phrase, and must provide conditions that sentences must meet to undergo the relative transformation. These provisions can be made in the following manner: determiners (in all sentences) will be produced with optional relative markers, the markers resulting from a classification of determiners according to what relatives may be adjoined to them. The relative transformations will be written so that they are applicable only to two sentences that share a noun phrase; the term "share" means that the substantives of the noun phrases are identical and that the determiners have the same relative markers.

When determiners are classified according to what relative clauses may be adjoined to them, the results refer to an intuitive classification of them with respect to definiteness and indefiniteness. Definiteness is correlated with appositive, or B-relative, clauses, and indefiniteness with restrictive, or A-relative, clauses. Relative markers are optional for noun phrases in common sentences¹; A-relatives are obligatory (under certain conditions) in predicate sentences.

If the relative transformation is applicable to two sentences, it has no direction; that is, either may be adjoined to the shared noun phrase in the other as a relative clause. I shall give examples of sentences containing determiners of each category.

A-relatives only: any, a, every, each, some, all

A-rel Any book about linguistics is interesting.

B-rel *Any book, which is about linguistics, is interesting.

B-rel *Every traveler, who has visited Venice, wants to return to Venice.

B-relatives only: the -generic; proper names

A-rel *John who is a nice man always helps old ladies across the street.

B-rel John, who is a Boy Scout, enjoys helping old ladies across the street.

B-rel The dog, which is a mammal, is carnivorous.

A-rel *The dog which is a mammal is carnivorous.

*This work was supported in part by National Science Foundation.

(XII. LINGUISTICS)

A-relatives and B-relatives: a, the (and variants)

A-rel An Indian, who knows the country well, has offered to guide us.

B-rel An Indian who knows the country well has offered to guide us.

A-rel The boys who knew the country best were dubious.

B-rel He pointed to a dog, who was looking at me hopefully.

The different classes may be thought of as representing points along a scale of definiteness and indefiniteness:

A rel indefinite	A rel, B rel specified	B rel unique
any, a every, etc.	a the	the proper names

Both 'a' and 'the' are in the intermediate category, accepting A- and B-relative clauses, which I have called specified. Usually 'a' is considered to be indefinite and 'the' definite; this traditional distinction will in fact be necessary at other points, although it is not called into play here.

The phrase structure rules to produce determiners with their appropriate optional relative markers are:

NP → Det + Subs.

Det → $\begin{cases} \text{Name} + (\text{B})^n \\ \text{Specified} + (\text{A})^n + (\text{B})^n \\ \text{Indefinite} + (\text{A})^n \end{cases}$

Name → $\begin{cases} \text{the} + \text{A}_{\text{sg}}^f \\ \text{U} \end{cases}$

Specified → (Predet) + $\begin{matrix} \text{a} \\ \text{the} \end{matrix}$

Indefinite → $\begin{cases} \text{Generic} \\ \text{Predet} \end{cases}$

Generic → $\begin{cases} \text{a} \\ \text{any} \end{cases}$

Predet → $\begin{cases} \text{each} \\ \text{every} \\ \text{some} \\ \text{all} \end{cases}$

The relative transformation will be applicable only when sentences share a noun phrase, with share having the meaning stated above – substantives and relative markers identical. This means that although two noun phrases need not be identical to undergo the relative transformation, they must be of the same degree of definiteness (that is, they must be able to accept the same relative markers). For instance, a noun phrase

with unique 'the' – accepting only B-relatives – may not be adjoined to one with specified 'the' – accepting both A- and B-relatives.

The cat is a mammal.	GT _{Rel} not applicable
The cat scurried under the bed.	
The cat is a mammal.	GT _{Rel} applicable
The cat is a quadruped.	
The cat scurried under the bed.	GT _{Rel} applicable
The cat belongs to Kate.	

Generalized Transformation A- or B-relative may be applied indefinitely many times to a noun phrase with determiner and appropriate relative marker. A succession of either A- or B-relative clauses may occur with a given marker, and, if both markers are picked, a succession of A-relatives followed by a succession of B-relatives may occur; but an A-relative clause (to the matrix NP) may not follow a B-relative clause, as the examples indicate.

A-rels: The men who were exposed to radiation who worked part-time are still healthy, but those who worked full-time are not.

B-rels: I know a skilled engineer, who is a graduate of M.I.T., who is good at golf, who is from the right background, who can't get admitted to the country club.

A-rels and B-rels: The mysterious knight, who was a prince in disguise, rode off into the distance.

B-rels and A-rels: *The kitten, which I put in the attic, that is crying, is sick.

1. Sketch of Relative Transformations² to be Applied in Order

a. A-Relative

S ₁ :	X	Y	(A)	(B)	<u>Z W N</u>	V
	1	2	3	4	5	6
S ₂ :	7	8	9	10	11	12

- Conditions:
- (i) 3 ≠ null
 - (ii) 3 4 5 = 9 10 11
 - (iii) 2...5 is NP
 - (iv) if Y is Predet + Art and S₁ is a predicate sentence, then S₂ is a common sentence.
wh + 12 replaces 3

(XII. LINGUISTICS)

b. B-Relative

Description same as above

Conditions: (i) 4 ≠ null

(ii) 3 4 5 = 9 10 11

(iii) 2...5 is NP; , +wh + 12 replaces 4

2. Generics

Generic determiners may be either indefinite -a, any- or unique -the. The indefinite generic determiners accept only A-relative clauses; the uniques, like names, accept only B-relative clauses.

Both 'any' and 'a' are treated as indefinite generic determiners, since they share not only the restriction as to relative clauses, but also a transformational restriction as to tense. The verbs of sentences with 'any' or generic 'a' cannot have the past tense, as the sentences below indicate. (I have not covered selectional restrictions on generics.)

Cats are mammals.

A man who would make so vile a pun would not scruple to pick a pocket.

* Any owl saw poorly in daylight.

A dog is vigilant.

* A dog was vigilant.

A cat may look at a king.

* A cat will have been an independent pet.

* Any dinosaur was a mammal.

Sentences with unique generic determiners differ from those above in three respects: there are no tense restrictions; 'the' has only the singular affix; the determiner accepts only B-relative clauses.

The dinosaur was a mammal.

Not generic: The dinosaurs were mammals.

The automobile, which was introduced in 1905, is no longer a novelty.

The doctor is highly respected in American small towns.

The difference between unique and indefinite generics is only partly captured by the classification here. In many, but not all, cases, unique and generic determiners are interchangeable (except, of course, when tense restrictions are involved). It would, at least, be necessary to state the selectional restrictions on generics, and the conditions under which the two kinds of generic are not interchangeable, to adequately characterize them.

A few sentences occur in which noun phrases with generic 'the' have restrictive relatives embedded in them. These are either counterexamples or exceptions to the classification given above, according to which generic 'the' accepts only appositive relatives. Examples:

The early bird catches the worm.

The poor man is better than the liar.

It is convenient and, I think, defensible to treat such sentences as having a different pattern. They may be accounted for by treating the noun phrase as Det + Noun, rather than Det + Adj + Noun. (The adjective and noun are pronounced together – earlybird, poorman – in such sentences, rather than according to the normal Adj + Noun pattern.)

3. Predicates

I shall now discuss predicate sentences (by predicate sentences I mean sentences of the form 'NP is Adj', and 'NP is NP'. The two will be differentiated when necessary). Predicate sentences differ from common sentences with respect to relative clauses so that, by making distinctions between the two, one obtains a (partial) characterization of predicate sentences.

It will be helpful to restate the occurrence of relative clauses in common sentences: In common sentences relative clauses may be embedded in noun phrases in either subject or object position. The embedding of a relative clause is never obligatory, and with one exception (discussed below) there are no restrictions on the type of sentence that may be embedded. The determiner of a given noun phrase decides whether A-relative, B-relative, or both, may be embedded in it.

In predicate sentences the embedding of A-relatives to noun phrases is obligatory under certain conditions, which turn on the definiteness of the determiner and upon whether the noun phrase has subject or object position.

a. Determiners may occur in any sequence in common sentences but, as many people have noticed, this is not the case with predicate sentences. The determiner of a predicate noun phrase may not be more definite than the determiner of the subject noun phrase, for example:

The janitor is a philosopher.

The janitor is the philosopher.

A janitor is a philosopher.

* A janitor is the philosopher.

* My friend is the janitor.

(This restriction must be observed in complex sentences also.)

(XII. LINGUISTICS)

b. The definite determiners that accept A-relatives – 'a' and 'the' – have obligatory A-relative markers under certain conditions when they occur in predicate sentences.

When 'the' occurs in a predicate noun phrase, it has an obligatory A-relative marker, that is, an A-relative clause must be embedded in it. For instance, sentences 1-3 are incomplete; when A-relative clauses are embedded in the predicate noun phrases (4-6) the sentences are no longer incomplete:

1. John is the linguist.
2. The janitor is the philosopher.
3. My sister is the doctor.
4. John is the linguist who spoke at the meeting.
5. The janitor is the philosopher with whom I discuss ethical problems.
6. My sister is the doctor who cured Allan.

When 'a' occurs in the subject NP of a predicate sentence, it has an obligatory A-relative marker. (Because of the restriction of Indef...Det sequences in predicate sentences, only sentences of the form 'AN is Adj' and 'AN is AN' occur when the subject has 'a'.) Thus sentences 7-9 are incomplete, but when A-relative clauses are embedded in their subjects (10-12) they are no longer incomplete:

7. A magazine was out-of-date.
8. A girl is a secretary at M.I.T.
9. A ball is striped.
10. A magazine that I found on the table was out-of-date.
11. A girl that I know is a secretary at M.I.T.
12. A ball that Jane has is striped.

Sentences with generic 'a' do not need A-relative clauses for completion:

13. A ball is round.
14. A cat is a mammal.
15. A ball is striped.
16. A child is happy.

The last two are odd for extralinguistic reasons: neither are true – all balls are not striped; all children are not happy. Therefore they will not be read as generic sentences and need A-relative clauses for completion.

Although adjoined predicate sentences may precede the common sentence, the obligatory relative marker that occurs with predicate sentences must have a common sentence adjoined to it. The incomplete sentences below (1 and 2) are still incomplete when predicate sentences are adjoined to them (3-6), both are complete when common sentences are adjoined to them (7 and 8).

- Incomplete: 1. John is the linguist.
 2. A girl is a secretary at M.I.T.
 3. John is the tall linguist.
 4. John is the linguist who is a mathematician.
 5. A pretty girl is a secretary at M.I.T.
 6. A girl who is a musician is a secretary at M.I.T.
- Complete: 7. John is the tall linguist I told you about.
 8. A girl that George knows is a secretary at M.I.T.

One may perhaps interpret these points by saying that predicate sentences are not (with 'a' in subject, 'the' in predicate) self-contained units. When they need completion, a self-contained unit — a common sentence — must be embedded in them.

B-relative markers occur optionally in common sentences with both specified determiners, whether they have subject or object position. In predicate sentences, however, B-relatives may not be embedded in a predicate NP with 'a'. (They may occur with subject NP's with 'the' or after an obligatory A-relative with 'a' or 'the'.) The predicate NP must stay within a narrow compass of definiteness.

*He is an anthropologist, who studies exotic Indian tribes.

*He is an anthropologist, whom I met at a party last week.

To form NP's in predicate sentences, then, one must make relatively fine distinctions with respect to definiteness, for example, 'the' and 'a' behave differently; and one must also know the role of the NP in the given sentence. (Neither of these considerations affect NP's in common sentences.) 'The' can stand alone in the subject of a predicate sentence, but it requires an A-relative when it occurs with a predicate noun. 'A' — less definite than 'the' — requires an A-relative with the subject of a predicate sentence and cannot have a B-relative (which is associated with definiteness) with a predicate noun.

Subject and predicate NP's in predicate sentences must be at certain (different) degrees of definiteness, while common sentences impose no restrictions of this kind on NP's. One may conclude that 'is' is less definite than a common verb, or, conversely, that common verbs contribute a certain element of definiteness to their subjects or objects.

Under the conditions stated above, then, NP's in predicate sentences must have A-relative clauses embedded in them. This, of course, differentiates predicate sentences from common sentences, in which the embedding of A-relatives is always optional. Moreover, common sentences must be embedded in obligatory A-relative markers.

c. The preceding sections have shown that predicate sentences must be differentiated from common sentences in the production of relative clause markers. The transformations that embed sentences as relative clauses, delete relative markers, and effect

(XII. LINGUISTICS)

order changes must also recognize predicate sentences as distinct from common sentences. With respect to relative clauses predicate sentences are distinct at all three stages: the formation of relative markers, the embedding of relative clauses, and the relevant transformations after the embedding of relative clauses.

Common sentences may be embedded as A-relative clauses, whether the shared noun has subject or object position in them. But predicate nominative sentences can be adjoined as A-relatives only if the shared noun has subject position. For instance: all but sentence 5 of the examples below may be embedded in sentence 1 as A-relative clauses. Sentences 2 and 3 are common sentences, 4 and 5 are predicate nominative sentences; in 4, the shared noun has subject position; in 5, the shared noun is the predicate noun.

- | | |
|-----------------------------|-------------------------|
| 1. I know a man. | (1 and 2) I know a man |
| 2. George knows a man. | whom George knows. |
| 3. The man knows George. | (1 and 3) I know the |
| | man who knows George. |
| 4. The man is an eccentric. | (1 and 4) I know a man |
| | who is an eccentric. |
| 5. George is a man. | (1 and 5) *I know a man |
| | whom George is. |

When restrictive relative clauses have been formed, the relative pronoun may be deleted unless the shared noun phrase has subject position in the constituent sentence. For instance:

- I know a man whom George knows.
I know a man George knows.
but I know a man who knows George.
*I know a man knows George.

Since predicate nominative sentences may become restrictive relatives only if the shared noun has subject position, they are not subject to deletion.

Only predicate adjective sentences are subject to the order-change transformation that follows deletion of relative markers, under conditions that I have stated elsewhere.³

The relative pronoun may be deleted from an appositive relative clause only if the constituent sentence is a predicate sentence, as the following examples show.

- I met a doctor, whom George knows, at the country club.
*I met a doctor, George knows, at the country club.
I met the doctor, who is an inveterate gambler, near the roulette wheels.
I met the doctor, an inveterate gambler, near the roulette wheels.

Several additions must be made to the rules already proposed if the grammar is to

account for predicate sentences, as well as common sentences. To the phrase structure rules that produce determiners with their appropriate relative markers two context-restricted rules will be added:

1. Art \rightarrow a + A in context [is]
2. Art \rightarrow the + A in context [is]

These provide that A-relative clauses must be embedded in the determiner 'a' when it is the subject of a predicate sentence, and in 'the' when it is a predicate noun.

Condition (iv) of the A-relative transformations provides correctly that common sentences be adjoined to complete the noun phrase in the predicate sentence.

4. Transformational Restrictions

Between certain transformations there is a mutual block, that is, if one is applied the other may not be, and vice versa. Such a relationship exists between negation and the adjunction of B-relative clauses. If the object of a verb has a B-relative clause, the sentence may not be negated, and if the sentence has undergone negation, B-relative clauses may not be adjoined to object or adverbial-object noun phrases. This situation can be simply handled by ordering the rules, and specifying that the later transformation may not be applied if the earlier one has been. Some examples follow:

- A-rel * He didn't write a novel, which was published by McGraw-Hill.
 He didn't eat the mango Elinor left for him.
 * He didn't eat the mango, which Elinor bought for him yesterday.
 * He didn't eat the mango, which was overripe.
 * I did not buy the rug, which would have been perfect for the hall.
 * I did not sleep on the air mattress, which belongs to the Halls.

This restriction only applies to objects (and certain adverbials). Thus the following negated sentences, whose subjects have B-relative clauses, are grammatical:

- The American contestant, who broke training, did not win the race.
 Jane, who is a friend of mine, did not recommend Kate.
 The novel, which was published by McGraw-Hill, was not a success.

The restriction is related to the fact that sentences are negated in the predicate in English. The only way to negate an entire sentence is to use an impersonal construction, such as "It was not the case that..." (Therefore subjects of negated sentences can have B-relative clauses, but objects cannot.) The scope of a negative extends throughout the predicate of a sentence, and B-relative clauses are outside the scope of the main verb – the keystone of the predicate.

(XII. LINGUISTICS)

The subject of a sentence can be negated by the special forms 'no', 'none'. The former is the form for negating indefinites, and therefore, like indefinites, cannot have B-relative clauses. It also applies to determinate 'a' – here, too, the distinction as to definiteness must go below the relative-clause categories of the first section.

1. An antique dealer wanted to buy the furniture.
2. *No antique dealer wanted to buy the furniture.
3. *No antique dealer, who had any sense, wanted to buy the furniture.

The restriction as to B-relative clauses applies to the other affectives which have been mentioned as such by Klima. For instance:

4. *John rarely spades his flower plot, which he planted last spring.
5. *We never go to the opera house, which is in Boston.

The mutual block that exists between negation and B-relative transformations also exists between B-relative and certain question transformations, namely, those whose scope applies to the entire predicate of a sentence. Thus the question transformation that brings the verb to the front of a sentence may not be applied if the object has a B-relative clause, and vice versa.

B-rel in subject

6. *Did John, who is a well-known artist, write a novel?
7. *Did he eat the mango, which Elinor bought last week?
8. *Did he paint a mural, which hangs in the Hotel Prado?
9. *Did he write a novel, which was published by McGraw-Hill?

The subject of a sentence with a B-relative clause in the object may not be questioned:

- 7a. *Who ate the mango, which Elinor bought last week?
- 8a. *Did he paint a mural, which hangs in the Hotel Prado?
- 9a. *Who wrote a novel, which was published by McGraw-Hill?

It seems reasonable to conclude that when the question transformation is applied to the subject of a sentence, its scope (unlike that of negation) extends to the whole sentence.

The object of a sentence may be questioned if the subject has a B-relative clause. This is hardly surprising: the scope of the object would not be expected to extend to the subject. Thus,

- What kind of house did Jane, who is a well-known architect, build?
Did the accident, which was quite serious, upset them?

Carlota S. Smith

Footnotes

1. Common sentences are sentences in which a transitive or intransitive verb, rather than 'is', etc., occurs as the main verb.

2. The transformation must be stated in much more detail (cf. unpublished paper by E. S. Klima) according to the structure of the constituent sentence and the role played in it by the shared NP. For simplicity, since these points are not relevant here, I assume that the shared NP is the subject of the constituent sentence in both relative transformations. The reason for condition (iv) in 1 will be given below.

3. C. S. Smith, A class of complex modifiers (to be published in Language).

B. NOTE ON CYCLICALLY ORDERED RULES IN THE RUSSIAN CONJUGATION

Consider the following forms of the Russian verb:

	<u>Infinitive</u>	<u>Imperative</u>	<u>1 sg Pres-Fut</u>	<u>3 pl Pres-Fut</u>	<u>3 sg Pres-Fut</u>	<u>Perfect Passive Participle</u>
1.	s, id, ét,	s, id, í	s, ižú	s, id, át	s, id, ít	(pra) s, ížín
2.	plat, ít,	plat, í	plač, ú	plát, at	plát, it	(za) plač, in
3.	živ, ít,	živ, í	živl, ú	živ, út	živ, ít	(a) živl, ón
4.	pisát,	p, iší	p, išú	p, íšut	p, íšít	
5.	plákat,	pláč,	pláč, u	pláč, ut	pláč, it	
6.	p, éč,	p, ik, í	p, ikú	p, ikút	p, ič, ot	p, ič, ón
7.	lgát,	lg, í	lgú	lgút	lžót	
8.	v, ist, í	v, id, í	v, idú	v, idút	v, id, ót	(za) v, id, ón
9.	trónut,	trón,	trónu	trónut	trón, it	
10.	žít,	živ, í	živú	živút	živ, ót	

With the exception of the two infinitives v, ist, í and žít,, where special consonant truncation occurs, each set of forms exhibits alternations. These alternations are of two types: transitive softening, and bare softening.

In undergoing transitive softening consonantal phonemes are replaced according to the following table:

t	k	s	x	d	g	st	sk	zd	zg	p	f	b	v	m	l	r
č	č	š	š	ž	ž	šč	šč	žž	žž	pl,	fl,	bl,	vl,	ml,	l,	r,

In undergoing bare softening a nonsharp (unpalatalized) consonant or liquid is replaced by its sharp (palatalized) cognate; that is, t by t, d by d, etc.

(XII. LINGUISTICS)

The constituent structure rules of the Russian verbal inflection are, in part, as follows:

1. Verbal form \rightarrow VF + (Reflexive) + #
2. VF \rightarrow

$$\left\{ \begin{array}{l} \text{Basic Verbal Stem + } \left\{ \begin{array}{l} \text{Past} \\ \text{Perfect Active Participle} \\ \text{Perfect Passive Participle} \end{array} \right\} + \left\{ \begin{array}{l} \text{Gender} \\ \text{Plural} \end{array} \right\} \\ \\ \text{Extended Verbal Stem + } \left\{ \begin{array}{l} \text{Perfect Gerund} \\ \text{Infinitive} \end{array} \right\} \\ \\ \left\{ \begin{array}{l} \text{Imperative} \\ \text{Person} \end{array} \right\} + \left\{ \begin{array}{l} \text{Sg} \\ \text{Pl} \end{array} \right\} \\ \text{Present Gerund} \\ \\ \left\{ \begin{array}{l} \text{Present Active Participle} \\ \text{Present Passive Participle} \end{array} \right\} + \left\{ \begin{array}{l} \text{Gender} \\ \text{Plural} \end{array} \right\} \end{array} \right.$$
3. Extended Verbal Stem \rightarrow Basic Verbal Stem + Present
4. Basic Verbal Stem \rightarrow

<u>s, id, é</u>	<u>pláka</u>	<u>v, od</u>	
<u>plat, i</u>	<u>p, ok</u>	<u>trónu</u>	
<u>živ, í</u>			
<u>p, isa</u>	<u>lgá</u>	<u>živ</u>	etc.

In Russian grammars it is usual to distinguish two stems for every verb: the Present-Future stem and the Past-Infinitive stem. It has been shown by Jakobson¹ that it is possible to predict one of these two from the other; for example, it is always possible to predict the shorter stem from the longer. The Basic Verbal Stem in the present description corresponds to the stem from which all forms of the verb may be predicted.

5. C \rightarrow C + = in env. X $\left[\begin{array}{l} +\text{voc} \\ -\text{cons} \end{array} \right]$ + Present + Imperative

where C is a velar consonant: k g x, and X contains no vowel if C is followed by a vowel. This rule accounts for the fact that velar consonants are preserved in the imperative, although by Rules C5 and 13 (see below) they should have undergone transitive softening. Rule 5 can naturally be extended to handle such exceptions as the verb tkáť, "to weave," as well as such present gerunds as b, ir, ig, á.

6. Present \rightarrow

$$\left\{ \begin{array}{l} \underline{i} \text{ in env. } X \left\{ \begin{array}{l} \underline{i} \\ \underline{e} \\ \underline{ja} \\ \underline{\check{c}a} \\ \underline{\check{s}a} \\ \underline{\check{z}a} \\ \underline{\check{\check{c}a}} \end{array} \right\} + \text{---} \\ \\ \underline{o} \end{array} \right.$$

There are some exceptions to this rule; for example, spa takes i, r^vža takes o, etc.

$$7. \text{ Person} \rightarrow \begin{cases} \underline{1} \\ \underline{2} \\ \underline{3} \end{cases}$$

$$8. \text{ Past} \rightarrow \underline{1}, \text{ in env. } \underline{\quad} + \text{Pl} \\ \rightarrow \underline{1}$$

$$9. \text{ Imperative} \rightarrow \underline{i}$$

but see Rule C4

$$10. \begin{array}{ll} 1 + \text{sg} \rightarrow \underline{u} & 1 + \text{pl} \rightarrow \underline{m} \\ 2 + \text{sg} \rightarrow \underline{s} & 2 + \text{pl} \rightarrow \underline{t,i} \\ 3 + \text{sg} \rightarrow \underline{t} & 3 + \text{pl} \rightarrow \begin{cases} \underline{at} \\ \underline{ut} \end{cases} \text{ in env. } + i + \underline{\quad} \end{array}$$

$$11. \text{ Perfect Passive Participle} \rightarrow \begin{cases} \underline{n} \text{ in env. } X \underline{a(j)} + \underline{\quad} \\ \underline{t} \text{ in env. } X \begin{pmatrix} \underline{j} \\ \underline{n} \\ \underline{m} \\ \underline{o} \\ \underline{u} \\ \underline{r} \end{pmatrix} + \underline{\quad} \\ \underline{o} + \underline{n} \end{cases}$$

$$12. \text{ Infinitive} \rightarrow \begin{cases} \underline{t,i} \text{ in env. } XC + \\ \text{where } X \text{ contains no stressed vowels} \\ \text{and } C \text{ is a nonvelar consonant} \\ \underline{t}, \end{cases}$$

Given these rules, the forms quoted at the beginning of this report would be presented by the following strings, where the parenthesization indicates the immediate constituent structure:

$$1. ((\underline{s, id, \acute{e}} + \underline{t,}) + \#) \quad ((\underline{s, id, \acute{e}} + \underline{i}) + \underline{i}) + \#) \quad ((\underline{s, id, \acute{e}} + \underline{i}) + \underline{u}) + \#) \\ ((\underline{s, id, \acute{e}} + \underline{i}) + \underline{at}) + \#) \quad ((\underline{s, id, \acute{e}} + \underline{i}) + \underline{t}) + \#) \quad ((\underline{pro} = \underline{s, \acute{i}d, e} + \underline{o} + \underline{n}) + \#)$$

$$2. ((\underline{plat, \acute{i}} + \underline{t,}) + \#) \text{ etc. like (1)}$$

$$3. ((\underline{\acute{z}iv, \acute{i}} + \underline{t,}) + \#) \text{ etc. like (1)}$$

$$4. ((\underline{p, is\acute{a}} + \underline{t,}) + \#) \quad ((\underline{p, is\acute{a}} + \underline{o}) + \underline{i}) + \#) \quad ((\underline{p, is\acute{a}} + \underline{o}) + \underline{u}) + \#) \\ ((\underline{p, \acute{i}sa} + \underline{o}) + \underline{ut}) + \#) \quad ((\underline{p, \acute{i}sa} + \underline{o}) + \underline{t}) + \#) \text{ }^X$$

$$5. ((\underline{pl\acute{a}ka} + \underline{t,}) + \#) \text{ etc. like (4) except for stress which is fixed on stem vowel.}$$

(XII. LINGUISTICS)

6. ((p,ók + t,) + #) ((p,ok + = + ó) + i) + #) (by Rule 5)
 (((p,ok + ó) + u) + #) (((p,ok + ó) + ut) + #) (((p,ok + ó) + t) + #)
 ((p,ok + ó + n) + #) ^x)

7. ((lgá + t,) + #) etc. like (6) except for stress, which is fixed on stem vowel.

8. ((v,od + t,i) + #) ((v,od + ó) + i) + #) like (6)

9. ((trónu + t,) + #) like (5)

10. ((živ + t,) + #) (((živ + ó) + i) + #) like (6)

The following rules apply in a cyclical fashion to the smallest constituent containing no parentheses:

C1. Transitive softening takes place in the env.

$$X \left\langle \begin{array}{l} +\text{voc} \\ -\text{cons} \end{array} \right\rangle \text{---} \langle + \rangle \left[\begin{array}{l} +\text{voc} \\ -\text{cons} \\ -\text{flat} \end{array} \right] + \left[\begin{array}{l} +\text{voc} \\ -\text{cons} \\ +\text{flat} \end{array} \right]$$

where at least one of the entities enclosed in $\langle \rangle$ must be present. The polnoglasié stems of the type kolót, can be handled here by a simple modification.

$$C2. \left[\begin{array}{l} +\text{voc} \\ -\text{cons} \end{array} \right] \rightarrow [+accented] \text{ in env. } \left[\begin{array}{l} +\text{voc} \\ -\text{cons} \\ +\text{accented} \end{array} \right] + \text{---}$$

$$C3. \left[\begin{array}{l} +\text{voc} \\ -\text{cons} \end{array} \right] \rightarrow \emptyset \text{ in env. } \text{---} + \left[\begin{array}{l} +\text{voc} \\ -\text{cons} \end{array} \right]$$

$$C4. \underline{i} \rightarrow \emptyset \text{ in env. } X + \text{---} + \#$$

where X contains an accented vowel and does not end in two consonants.

$$C5. [+cons] \rightarrow [+sharp] \text{ in env. } X \text{---} \left\{ \begin{array}{l} + \\ + = + \end{array} \right\} \left\{ \begin{array}{l} \left[\begin{array}{l} +\text{voc} \\ -\text{cons} \\ +\text{flat} \end{array} \right] + Y \\ \left[\begin{array}{l} +\text{voc} \\ -\text{cons} \\ -\text{flat} \end{array} \right] \end{array} \right\}$$

where $Y \neq 0$.

C6. Erase innermost parentheses and return to C1.

Applying the cyclical rules to the forms represented in example 1 on the previous page, we obtain the following results:

The first cycle applies only to the innermost parentheses.

	s, id, é + t,	s, id, é + i	pro+=+s, íd, e + o + n
Rule C1			pro+=+s, íže + o + n
Rule C2		s, id, é + í	
Rule C3		s, id, + í	pro+=+s, íž + o + n
Rule C4			
Rule C5		s, id, + í	
Rule C6	s, id, é + t, + #	s, id, + í + i s, id, + í + u s, id, + í + at s, id, + í + t	pro+=+s, íž + o + n + #

Later cycles apply only to the forms in the middle column.

	s, id, + i + í	s, id, + í + u	s, id, + í + at	s, id, + í + t
Rule C1		s, iž + í + u		
Rule C2	s, id, + í + í	s, iž + í + ú	s, id, + í + át	
Rule C3	s, id, + í	s, iž + ú	s, id, + át	
Rule C4				
Rule C5	s, id, + í			s, id, + í + t
Rule C6	s, id, + í + #	s, iž + ú + #	s, id, + át + #	s, id, + í + t + #

No rules apply in the third cycle.

The examples in (2) and (3) work in parallel fashion and will not be discussed further. The examples in (4) and (5), except for the infinitive, undergo transitive softening in the first cycle by virtue of satisfying the condition for Rule C1; that is, $\underline{p, isa} + \underline{o} \rightarrow \underline{p, iša} + \underline{o}$ $\underline{pláka} + \underline{o} \rightarrow \underline{pláča} + \underline{o}$. Since all further steps are simple, we shall only illustrate them with the help of the imperative form (($\underline{pláka} + \underline{o}$) + \underline{i}) + #).

	<u>First Cycle</u>	<u>Second Cycle</u>	<u>Third Cycle</u>
Rule C1	pláča + o		
Rule C2			
Rule C3	pláč + o	pláč + i	
Rule C4			pláč + #
Rule C5			
Rule C6	pláč + o + i	pláč + i + #	

In (6) and (7) we work out the derivation for

(XII. LINGUISTICS)

	((lgá+o)+u)+#)	((lgá+ó)+t)+#)	((p,ok+=+ó)+i)+#)	((p,ok+ó+n)+#)
C1				
C2	lgá + ó	lgá + ó		
C3	lg + ó	lg + ó		
C4				
C5				p, ok, + ó + n
C6	lg + ó + u	lg + ó + t	p, ok + = + ó + i	p, ok, + ó + n + #
C1				
C2	lg + ó + ú		p, ok, + = + ó + í	
C3	lg + ú		p, ok, + = + í	
C4				
C5		lg, + ó + t	p, ok, + = + í	
C6	lg + ú + #	lg, + ó + t + #	p, ok, + = + í + #	

No rules apply in the third cycle. Observe that = is introduced in the imperative ((p,ok += + ó) + i) + #) by Rule C4.

To obtain the phonetically correct forms, we require a number of additional rules which apply after the cycle.

$$P1 \quad \begin{bmatrix} -\text{voc} \\ +\text{cons} \\ +\text{comp} \\ +\text{grave} \\ +\text{sharp} \end{bmatrix} \rightarrow [-\text{grave}] \text{ in env. } \text{---} X$$

where $X \neq += + Y$.

That is, palatalized velars undergo transitive softening except before =. Observe that this rule turns $\underline{\text{lg}}, + \underline{\text{o}} + \underline{\text{t}} \rightarrow \underline{\text{l}}\check{\text{z}} + \underline{\text{o}} + \underline{\text{t}}$, but does not affect $\underline{\text{p,ok}}, + = + \underline{\text{o}} + \underline{\text{i}}$.

$$P2 \quad \begin{bmatrix} -\text{voc} \\ +\text{cons} \\ +\text{comp} \\ +\text{grave} \end{bmatrix} \rightarrow [+sharp] \text{ in env. } \text{---} \left\{ \begin{matrix} + \\ + = + \end{matrix} \right\} \begin{bmatrix} +\text{voc} \\ -\text{cons} \\ -\text{grave} \end{bmatrix}$$

That is, velars are palatalized before front vowels.

$$P3 \quad \begin{bmatrix} +\text{voc} \\ -\text{cons} \\ -\text{comp} \\ -\text{diff} \\ +\text{flat} \end{bmatrix} \rightarrow \begin{bmatrix} -\text{grave} \\ -\text{flat} \end{bmatrix} \text{ in env. } \begin{bmatrix} +\text{cons} \\ +\text{sharp} \end{bmatrix} \text{---} \begin{bmatrix} -\text{voc} \\ +\text{cons} \\ +\text{comp} \\ -\text{grave} \\ +\text{sharp} \end{bmatrix}$$

That is, $\underline{\text{o}} \rightarrow \underline{\text{e}}$ in env. palatalized consonant or liquid --- $\left\{ \begin{matrix} \check{\text{c}} \\ \check{\text{z}} \end{matrix} \right\}$.

$$\text{P4} \quad \left[\begin{array}{l} +\text{voc} \\ -\text{cons} \\ -\text{diff} \\ -\text{accented} \end{array} \right] \rightarrow \left[\begin{array}{l} +\text{diffuse} \\ -\text{grave} \end{array} \right] \quad \text{in env.} \quad X \quad \left\{ \begin{array}{l} \left[\begin{array}{l} -\text{voc} \\ +\text{cons} \\ +\text{comp} \\ -\text{grave} \end{array} \right] \\ \left[\begin{array}{l} +\text{cons} \\ +\text{sharp} \end{array} \right] \end{array} \right\} \text{---}$$

That is, $\left\{ \begin{array}{l} \text{o} \\ \text{a} \\ \text{e} \end{array} \right\} \rightarrow \underline{\text{i}}$ after palatals and palatalized consonants and liquids.

$$\text{P5} \quad \left[\begin{array}{l} +\text{voc} \\ -\text{cons} \\ +\text{diff} \\ -\text{flat} \end{array} \right] \rightarrow \left[+\text{grave} \right] \quad \text{in env.} \quad \left[\begin{array}{l} +\text{cons} \\ -\text{sharp} \end{array} \right] \text{---}$$

That is, $\underline{\text{i}} \rightarrow \underline{\text{i}}$ in position after "hard" consonants and liquids.

M. Halle

Footnote

^xThe assignment of stress to unaccented stems is handled by a special set of rules, not discussed here.

References

1. R. Jakobson, *Word* 4, 155-167 (1960).

C. SOME REMARKS ON CHOMSKY'S CONTEXT-FREE LANGUAGES

1. Introduction

This report is devoted to the examination of several families of subsets of a free monoid that arise rather naturally when generalizing some definitions of classical analysis to the noncommutative case. These families contain, in particular, the regular events of Kleene and the context-free languages of Chomsky.

The main tool is the so-called formal power series with integral coefficients in the noncommutative variates $x \in X$.

By definition, such a formal power series, r , is a mapping that assigns to every word $f \in F(X)$, (where $F(X)$ is the free monoid generated by X) a certain positive or negative integral "weight" $\langle r, f \rangle$, the coefficient of f in r . Thus, in fact, a formal power series is just an element of the free module with basis $F(X)$.

In fact, if instead of considering only a subset F' of $F(X)$ we specify a process producing its words, it seems natural to count how many times each of them is obtained and the formal power series is the tool needed for handling this more detailed information.

Of course, with this interpretation we only get positive power series, i. e., power

(XII. LINGUISTICS)

series in which every coefficient $\langle r, f \rangle$ is non-negative. The general case may be thought of as being associated with two processes and then the coefficient of f is the difference between the number of times f is obtained by each of these processes.

In any case, we shall define the support of r as the subset $F'_r = \{f \in F : \langle r, f \rangle \neq 0\}$.

The power series form a ring $\bar{A}(X)$ with respect to the following operation:

multiplication by an integer: the coefficient of f in nr is simply $n\langle r, f \rangle$

addition: $\langle r+r', f \rangle = \langle r, f \rangle + \langle r', f \rangle$, for all f

multiplication: $\langle rr', f \rangle = \sum \langle r, f' \rangle \langle r', f'' \rangle$, where the sum is extended to all factorizations $f = f'f''$.

It is clear that when r and r' are positive, the support $F'_{r+r'}$ of $r+r'$ is just the union of the supports of r and of r' ; similarly, the support of rr' is the set product $F'_r F'_{r'}$. For arbitrary r the interpretation is more complicated.

It is convenient to introduce a topology in $\bar{A}(X)$ in order to be able to define the limit of a sequence. Among the many possibilities that are available the simplest one is based upon the following definition of the distance: $\|r-r'\| = 1/n$ if and only if $\langle r, f \rangle = \langle r', f \rangle$ for every word $f \in F$ of degree ("length") strictly less than n and $\langle r, f \rangle \neq \langle r', f \rangle$ for at least one $f \in F$ of degree n .

Thus, $\|r-r'\| = 0$ if $\langle r, e \rangle \neq \langle r', e \rangle$, where e is the empty word and $\|r-r'\| = 0$ if $r = r'$.

It is easily checked that $\|r-r'\| \leq \sup(\|r-r''\|, \|r'-r''\|)$ for any $r, r', r'' \in \bar{A}(X)$, and that the addition and multiplication are continuous. The norm $\|r\|$ of r is just $\|r-0\|$. Clearly, $\|r\| = 1/n$, where n is the smallest integer such that $\langle r, f \rangle \neq 0$ for some f of degree (= length) n . Thus r has a finite norm if and only if $\langle r, e \rangle \neq 0$.

We now introduce the important notion of an inverse.

By definition $r \in \bar{A}(X)$ is invertible if $r' = e-r$ has a finite norm, i. e., if $\langle r, e \rangle = 1$.

If this is so, the infinite sum $e + \sum_{n>0} r'^n = r''$ satisfies the identity $r'' - r''r' = r'' - r'r'' = e$, i. e., $r''r = rr'' = e$.

This suggests the notation $r'' = r^{-1}$ and, since r'' is invertible, one can also construct $(r'')^{-1}$.

It is easily verified that $(r'')^{-1} = r$, and thus there is no inconvenience in considering the infinite sum r'' as the inverse r^{-1} of r . It is worth noting that if r_1 is a positive element with finite norm, then $(e-r_1)^{-1}$ is positive and has as its support the subset $F^*_{r_1} = \bigcup_{n>0} (F_{r_1})^n$ in Kleene's notation.

Thus we are able to interpret all of the usual set theoretic operations except for complementation and intersection.

With respect to the first, we can observe that by construction the formal power series $(e - \sum_{x \in X} x)^{-1}$ is equal to $\Sigma\{f : f \in F(X)\}$.

Consequently, if we associate with the subset F' of F the formal power series

$r_{F'} = \sum_{f \in F'} f$ (i. e., the power series with $\langle r_{F'}, f \rangle = 1$ if $f \in F' = 0$, otherwise) the support of $\left(e - \sum_{x \in X} x \right)^{-1} - r_{F'}$ is precisely the complement of F' in F .

With respect to the intersection, we can define a Hadamard product which associates with any $r, f \in \bar{A}(X)$ the new power series $r \otimes r'$, defined by $\langle r \otimes r', f \rangle = \langle r, f \rangle \langle r', f \rangle$ for all f . Clearly, the support of $r \otimes r'$ is the intersection of the supports of r and r' .

However, the Hadamard product is no longer an elementary operation and this may explain why some otherwise reasonable families of subsets are not closed under intersection (cf. below).

2. Relation with Ordinary Power Series

This can be expressed in a loose way by saying that ordinary power series are obtained from the elements of $\bar{A}(X)$ by disregarding the order of the letters in the words $f \in F$. Formally, let α be a bijection (one-to-one mapping onto) $X \rightarrow \bar{X}$. An ordinary power series \bar{r} in the variates $\bar{x}_i \in \bar{X}$ is an infinite sum $\bar{r} = \sum a_{n_1 n_2 \dots n_m} \bar{x}_1^{n_1} \bar{x}_2^{n_2} \dots \bar{x}_m^{n_m}$ extended to all the monomials $\bar{x}_1^{n_1} \bar{x}_2^{n_2} \dots \bar{x}_m^{n_m}$.

We can consider that any such r (with integral coefficients $a_{n_1 n_2 \dots n_m}$) is the image by the homomorphism α of at least one $r \in \bar{A}(X)$ by defining $a_{n_1 n_2 \dots n_m}$ as the sum of $\langle r, f \rangle$ extended to all of the words $f \in F(X)$ containing the letters x_1 n_1 times; the letters x_2 n_2 times \dots etc.; i. e., to all words f such that $\alpha f = \bar{x}_1^{n_1} \bar{x}_2^{n_2} \dots \bar{x}_m^{n_m}$, where α is the homomorphism sending $F(X)$ onto the free commutative monoid generated by \bar{X} . It is trivial that $\alpha(r_1 \pm r_2) = \alpha r_1 \pm \alpha r_2$; $\alpha r_1 r_2 = \alpha r_1 \alpha r_2 = \alpha r_2 \alpha r_1$; $\alpha(r_1^{-1}) = (\alpha r_1)^{-1}$ identically.

Also, when X contains a single letter no difference need be made between formal (noncommutative) and ordinary (commutative) power series.

Since the theory of ordinary power series is an extremely well-developed chapter of mathematics, the existence of the homomorphism α may at times be used for the study of the formal power series and of their support. The discussion of some elementary examples of this approach is, in fact, the main content of this report.

3. The Algebraic Elements of $\bar{A}(X)$

In ordinary calculus, one usually considers as especially elementary functions the polynomials, the rational functions, and the algebraic functions.

By definition, a polynomial is the function represented by an ordinary power series with only finitely many nonzero coefficients; a rational function is the quotient of two polynomials; an algebraic function is a function of the variates with the property that

(XII. LINGUISTICS)

it satisfies identically some algebraic relation expressed by a polynomial.

For example, $\frac{\bar{x}_1 \bar{x}_2^2}{1 - \bar{x}_2^2}$ is a rational function of \bar{x}_1 and \bar{x}_2 , and the function r of the commutative variates \bar{x}_1 and \bar{x}_2 that is such that $\bar{x}_1 \bar{x}_2 \bar{r}^2 - \bar{r} + 1 = 0$ identically is an algebraic function.

We can imitate this hierarchy by introducing the following definitions: an element $r \in \bar{A}(X)$ is a polynomial if its support is a finite set; an element $r \in \bar{A}(X)$ is rational if it can be obtained from the generators $x \in X$ as a finite expression using only the sum, the product, and the inversion (of invertible elements).

It is clear that the polynomials form a subring of $A(X)$. Indeed, this ring is what is usually called the free ring generated by X.

In a similar manner, the set $\bar{R}(X)$ of the rational elements is a ring, i. e., it is closed under addition, subtraction, and multiplication. Furthermore, it is closed under inversion (of invertible elements). In fact, $\bar{R}(X)$ is the smallest subring of $\bar{A}(X)$ closed under this last operation and containing X .

It is easily verified that for any $r \in \bar{R}(X)$ the "Abelianized" ordinary power series $\bar{r} = \alpha r$ represents a rational function.

Consider, for instance, the formal power series r with $\langle r, f \rangle = 1$ if and only if $f = x_2^7 x_1^{3+2n_1} x_2^{3+2n_2} x_1 \dots x_2 x_1^{3+2n_m} x_2^5$, and $\langle r, f \rangle = 0$, otherwise. This series r belongs to $\bar{R}(X)$ because r is equal to $x_2^6 (e - x_2 x_1^3 (e - x_1^2)^{-1})^{-1} x_2^5$, and αr can be reduced to the quotient of two polynomials by writing

$$\begin{aligned} \alpha r &= \bar{x}_2^6 (e - \bar{x}_2 \bar{x}_1^3 (e - \bar{x}_1^2)^{-1})^{-1} \bar{x}_2^5 \\ &= \bar{x}_2^6 ((e - \bar{x}_1^2)^{-1} (e - \bar{x}_1^2 - \bar{x}_2 \bar{x}_1^3))^{-1} \bar{x}_2^5 \\ &= (e - x_1^2) (e - x_1^2 - x_2 x_1^3)^{-1} x_2^{11} \\ &\quad x_2^{11} (1 - x_1^2) \\ &= \frac{x_2^{11} (1 - x_1^2)}{1 - x_1^2 - x_1^2 x_2} \end{aligned}$$

The family of all subsets of F that can be the support of a rational element of $\bar{A}(X)$ has been defined elsewhere.⁷ It is not difficult to verify that it is closed under union, intersection, set product, and Kleene's star operation.

Having recalled these facts, we proceed to the definition of an algebraic element of $\bar{A}(X)$.

For this purpose, we consider a finite set $\boxed{\cdot}$ of m new elements ξ_i , and we denote by $\bar{\sigma}$ an m -tuple of polynomials σ_ξ in the (noncommutative) variates $y \in Y = X \cup \boxed{\cdot}$ that satisfy the condition that $\langle \sigma_\xi, e \rangle = \langle \sigma_\xi, \xi' \rangle = 0$ for all $\xi, \xi' \in \boxed{\cdot}$.

Now let W denote the set of all m -tuples $w = (w_1, w_2, \dots, w_m)$ of elements of $\bar{A}(X)$. We consider $\bar{\sigma}$ as a mapping of W into itself by defining the coordinate $\bar{\sigma}w_\xi$ of the transformed vector $\bar{\sigma}w$ as the element of $\bar{A}(X)$ obtained by replacing in the polynomial $\bar{\sigma}_\xi$ every symbol ξ' by the corresponding coordinate $w_{\xi'}$, of w_ξ .

For instance, if $\sigma_{\xi_1} = x_1\xi_2x_2$; $\sigma_{\xi_2} = x_1x_2 + x_1\xi_1x_2\xi_2$; and if w is the vector $(3x_1 - x_2x_1, x_2^2 + 2x_3^4)$, the w coordinates of $\bar{\sigma}w$ are

$$\bar{\sigma}w_{\xi_1} = x_1(x_2^2 + 2x_3^4) x_2 = x_1x_2^3 + 2x_1x_3^4x_2$$

$$\bar{\sigma}w_{\xi_2} = x_1x_2 + x_1(3x_1 - x_2x_1) x_2(x_2^2 + 2x_3^4)$$

$$= x_1x_2 + 3x_1^2x_2^3 + 6x_1^2x_2x_3^4 - x_1x_2x_1x_2^3 - 2x_1x_2x_1x_2x_3^4.$$

It is clear that $\bar{\sigma}$ is a continuous mapping in the sense that if $w, w' \in W$ are such that $\|w - w'\|_\xi \leq 1/n$ for each $\xi \in \bar{\square}$ (i. e., for short, if $\|w - w'\| \leq 1/n$, then $\|\bar{\sigma}w - \bar{\sigma}w'\| \leq 1/n$).

Indeed, the relation $\|w - w'\| \leq 1/n$ expressed the fact that the coefficients $\langle w_\xi, f \rangle$ and $\langle w'_\xi, f \rangle$ are equal for every coordinate $\xi \in \bar{\square}$ and for every word f of degree $\leq n$. Since the coefficient of every word of degree n in the polynomial in the letters $x \in X$ obtained by the substitution $\xi' \rightarrow w_{\xi'}$, or $\xi' \rightarrow w'_{\xi'}$, in σ_ξ depends only on the terms of lower degree, the result is a simple consequence of the definition.

In fact, because of our hypothesis on $\bar{\sigma}$, a stronger result can be proved when w and w' satisfy the supplementary condition that $\langle w_\xi, e \rangle = \langle w'_\xi, e \rangle = 0$ for all ξ . Then, obviously, this last condition is still verified for $\bar{\sigma}w$ and $\bar{\sigma}w'$ (because $\langle \sigma_\xi, e \rangle = 0$). Furthermore (because $\langle \sigma_\xi, \xi' \rangle = 0$), we can conclude from $\|w - w'\| \leq 1/n$ that $\|\bar{\sigma}w - \bar{\sigma}w'\| \leq 1/n + 1$. This, again, is a direct consequence of the fact that the coefficients of the terms of degree $n+1$ of $\bar{\sigma}w$ are determined univocally by the coefficients of the terms of degree $\leq n$ of w .

Let us now consider the infinite sequence $w, w_1, \dots, w_n, \dots$, where $w_0 = (0, 0, \dots, 0)$ and $w_{n+1} = \bar{\sigma}w_n$. By applying our previous remarks and using induction, we can easily show that for all n and $n' > 0$ we have $\|w_n - w_{n+n'}\| \leq 1/n$. Consequently, we have proved that $w = \lim_{n \rightarrow \infty} w_n$ is a well-defined element of W and that $\lim_{n \rightarrow \infty} \bar{\sigma}w_n - w_n = 0$. This suggests that we speak of w as of a solution of the system of equations $\xi = \sigma_\xi$ (i. e., $w = \bar{\sigma}w$), since, in fact, for each ξ , w_ξ is equal to the formal power series in the $x \in X$ obtained by replacing in σ_ξ each ξ' by the coordinate $w_{\xi'}$.

We shall say, accordingly, that w_ξ is an algebraic element of $\bar{A}(X)$. Because of our definition of $\bar{\sigma}$, any w has a finite norm (i. e., $\langle w_\xi, e \rangle = 0$). This restriction would be artificial; we shall denote by $\bar{S}(X)$ the set of all formal power series that is the sum of a polynomial and of a coordinate w_ξ , defined above, for some suitable finite set of polynomials $\bar{\sigma}$, or, as we prefer to say, by a set of "equations" $\xi = \sigma_\xi$.

(XII. LINGUISTICS)

It is not difficult to verify the fact that $\overline{S}(X)$ is a ring closed under the formation of inverses (of invertible elements). Indeed, let r and r' be obtained as the coordinates w_ξ and w'_ξ of the solutions w and w' of the equations $w = \overline{\sigma}w$ and $w' = \overline{\sigma}'w'$. For the sake of clarity, we assume that $\overline{\sigma}$ and $\overline{\sigma}'$ are defined by two disjoint sets $\overline{\sigma}$ and $\overline{\sigma}'$ of m and m' elements, and we consider $\overline{\sigma}''$ the union of $\overline{\sigma}$, $\overline{\sigma}'$ and of a new letter ξ'' . Then, if we denote by σ'' the direct sum of $\overline{\sigma}$ and $\overline{\sigma}'$, it is clear that the new equation $\xi'' = \sigma_\xi + \sigma'_\xi$, determines $w''_{\xi''} = r+r'$. Similarly, the equation $\xi'' = \xi\xi'$ determines $w''_{\xi''} = rr'$.

In order to get $(e-r)^{-1} = e + \sum_{n>0} r^n$ it is enough, for instance, to add the new equation $\xi'' = \xi\xi'' - \sigma$.

As a final remark it may be pointed out that (as for rational elements) the homomorphism α sends the algebraic elements of $\overline{A}(X)$ onto the Taylor series of the ordinary algebraic functions. These last series are easily proved to converge in some small enough domain around 0. Let us also mention that $\overline{S}(X)$, as defined constructively here, can also be shown to be identical to the set of all formal power series with integral coefficients that satisfy a set of equations of the type $w = \overline{\sigma}w$, described above, provided, of course, that such solutions exist.

Example 1.

$$\begin{aligned} \text{Let } \sigma_{\xi_1} &= x_1 \xi_1 x_2 + x_1 x_2 \\ \sigma_{\xi_2} &= \xi_1 \xi_2 + x_1 \xi_1 x_2 + x_1 x_2. \end{aligned}$$

Since the first equation involves only ξ_1 , it can be solved for its own sake, and one easily obtains $r = w_{\xi_1} = \sum_{n>0} x_1^n x_2^n$. Then the second equation gives

$$w_{\xi_2} = r w_{\xi_2} + r, \text{ that is, } w_{\xi_2} = r(e-r)^{-1}.$$

Thus, by definition, a word f belongs to the support of w_{ξ_2} if and only if it can be

factorized as a product $\begin{pmatrix} n_1 & n_1 \\ x_1 & x_2 \end{pmatrix} \begin{pmatrix} n_2 & n_2 \\ x_1 & x_2 \end{pmatrix} \dots \begin{pmatrix} n_m & n_m \\ x_1 & x_1 \end{pmatrix}$ of words belonging to the support of r .

Since, trivially, this factorization is unique, we always have $\langle w_{\xi_2}, f \rangle = 0$ or 1 .

Example 2.

$$\text{Let } \overline{\sigma}_{\xi_1} = x_1 \xi_1 x_2 \xi_1 + x_1 x_2 \xi_1 + x_1 \xi_1 x_2 + x_1 x_2.$$

After setting $r = e + \xi_1$ we get the simpler form $r = x_1 r x_2 r + e$, instead of the equation $\xi_1 = \overline{\sigma}_{\xi_1}$. Again, $\langle r, f \rangle = 0$, or 1 ; with $\langle r, f \rangle = 1$ if and only if

- 1^o. f contains as many x_1 as x_2 .
- 2^o. any left factor f' of f contains at least as many x_1 as x_2 .

Since the equation can also be written in the form $r = (e - x_1 r x_2)^{-1}$, it follows that every $f \in F_r$ has one and only one factorization as a product of words belonging to the support F' of $x_1 r x_2$.

F' is closely related to the well-formed formulas in Lukasiewicz' notation because f belongs to F' if and only if it satisfies 1^0 and, instead of 2^0 , condition 3^0 . Any factor f' of f contains, strictly, more x_1 than x_2 , unless $f' = e$ or $f' = f$.

Let us now observe that $x_1 r x_2$ satisfies the equation $x_1 (e - x_1 r x_2)^{-1} x_2 = x_1 r x_2$. Taking the homomorphic image as a and writing $\bar{r}' = a(x_1 r x_2)$, we get the ordinary equation $\bar{x}_1 \bar{x}_2 (1 - \bar{r}')^{-1} = \bar{r}'$; i. e., $\bar{r}'^2 - \bar{r}' + \bar{x}_1 \bar{x}_2 = \theta$.

By construction, the ordinary power series r' takes the value θ for $x_1 x_2 = 0$ and thus, as is well known,

$$\bar{r}' = \frac{1 - \sqrt{1 - 4\bar{x}_1 \bar{x}_2}}{2} = \frac{1}{2} \sum_{n>0} (-\bar{x}_1 \bar{x}_2)^n \left[\begin{matrix} 1/2 \\ n \end{matrix} \right]$$

where $\left[\begin{matrix} 1/2 \\ n \end{matrix} \right]$ is the binomial coefficient.

Because $\langle x_1 r x_2, f \rangle = 0$ or 1 , we can conclude that $(-1)^n \left[\begin{matrix} 1/2 \\ n \end{matrix} \right]$ is the number of distinct words of degree $2n$ in the support of $x_1 r x_2$.

The reader may notice that our present computation is exactly the one used in the classical problem of the return to equilibrium in coin-tossing games.

Example 3.

Let \square be the union of ζ, η and of ξ_i ($i=1, \dots, 2m$) and agree that $\xi_{i+m} = \xi_i$, when $i = i'+m$. Let $X = \{x_i\} \ i = 1, 2, \dots, 2m$, and consider the $2m$ equations

$$\xi_i = x_i x_{i+m} + x_i \left(\zeta + \zeta^2 + \zeta \eta \zeta - \sum_{j=1}^{2n} x_j (e + \eta) x_{j+m} \right) x_{i+n}$$

$$\zeta = \sum_{i=1}^{2m} \xi_i; \quad \eta = \zeta + \zeta \eta.$$

Simple transformations reduce these to standard form, and it can be proved that $\langle e + \eta, f \rangle = 0$, or 1 with $\langle e + \eta, f \rangle = 1$ if and only if f belongs to the kernel K of the homomorphism ϕ , which sends $F(X)$ onto the corresponding free group (with $(\phi x_i)^{-1} = \phi x_{i+n}$).

After performing the homomorphism a , we compute the value of $a\eta = u(t)$ for $\bar{x}_1 = \bar{x}_2 = \dots = \bar{x}_{2m} = \frac{t}{2m}$. By construction, $u(t)$ is the generating function of the recurrent event consisting in the return to K , and $u(1)$ is the probability that a random word ever belongs to K when the letters $x_i \in X$ are produced independently with constant probability $1/2m$.

We find that $u = u(t)$ is defined by the quadratic equation $(4m^2 - t^2)u^2 - 4m^2u + 2mt^2 = 0$, which is in agreement with similar results of Kesten³ to which we refer for a more

(XII. LINGUISTICS)

explicit interpretation of $u(t)$.

4. Some Subfamilies of $\bar{S}(X)$

It seems natural to distinguish in $\bar{S}(X)$ the subset $\bar{S}_1(X)$ of those elements that are obtained when each σ_ξ of $\bar{\sigma}$ has the special form

$\sigma_\xi = f + \sum f' \xi' f''$, where $f, f', f'' \in F(X)$, and the summation is over any finite set of triples (f', ξ', f'') (with, eventually, the same ξ' occurring several times; i. e., when each σ is linear in the variates $\xi \in \bar{\Sigma}$).

Within $\bar{S}_1(X)$ itself we shall distinguish the special case $S_0(X)$ for which $\sigma = f + \sum \xi' f''$; i. e., only one-sided linear equations are considered.

Clearly, after taking the homomorphic image as a , both $\bar{S}_1(X)$ and $\bar{S}_0(X)$ collapse onto the ring of the ordinary rational functions but, at the level of $\bar{A}(X)$, the sets from $\bar{S}_0(X)$ form only a very restricted subset of $\bar{S}_1(X)$, as we shall see.

A second principle of classification is provided by the restriction that every coefficient in the polynomials σ_ξ is non-negative.

Under this hypothesis, the same is true of the power series w_ξ , and, correspondingly, we obtain three subsets (in fact, three semirings) which we denote $\bar{S}^+(X)$, $\bar{S}_1^+(X)$, and $\bar{S}_0^+(X)$. It is to be stressed that the converse is not true. Indeed, it is quite easy to display examples of formal power series having only non-negative coefficients that belong to $\bar{S}_0(X)$, but not even to $\bar{S}^+(X)$.

A priori the inclusion relations shown in Fig. XII-1 hold. Here, $P_0(X)$ and $P_0^+(X)$

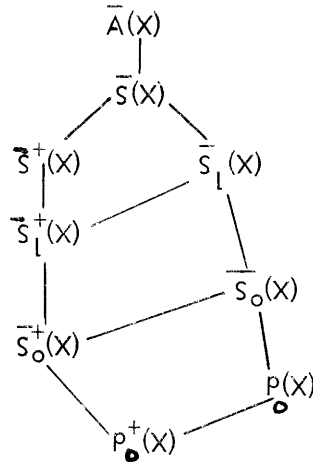


Fig. XII-1.

denote the polynomials and the positive polynomials, respectively. Insofar as the corresponding supports are concerned, three theorems summarize the results.

THEOREM I. (Ginsburg-Rice). The family of the supports of the elements of $\bar{S}^+(X)$ is identical to the family \mathcal{C} of Chomsky's context-free languages.

THEOREM II. (Chomsky). The family of the supports of the elements of $\bar{S}_0^+(X)$ is identical to the family \mathcal{R}_0 of Kleene's regular events.

THEOREM III. The family of the supports of the elements of $\bar{S}_0(X)$ is identical to the family \mathcal{R} of the sets of words accepted by an automaton of the type \mathcal{A} (i. e., it is identical to the family of the supports of the rational elements of $\bar{A}(X)$).

In order to prove Theorem I we need to alter slightly Chomsky's definition and we propose

DEFINITION. A context-free grammar is given by

- i. Two disjoint finite sets \bar{X} and X ;
- ii. A finite set G of pairs (ξ, g) , where $\xi \in \bar{X}$, $g \in F(X \cup \bar{X})$, $g \neq e$, $g \notin \bar{X}$.
- iii. A distinguished element $\xi_0 \in \bar{X}$.

The language $D_X(\xi_0, G)$ produced by G is the intersection $F(X) \cap D(\xi_0, G)$, where $D(\xi_0, G)$ is the smallest subset of $F(X \cup \bar{X})$ which is such that $\xi_0 \in D(\xi_0, G)$ and $g_1 \xi' g_2 \in D(\xi_0, G)$, and $(\xi', g) \in G$ implies $g_1 g g_2 \in D(\xi, G)$. In the usual terminology, \bar{X} (resp. X) is the nonterminal (resp. terminal) vocabulary, and G is the grammar; our definition departs from Chomsky's by the easily met restriction $g \notin \bar{X}$ for each rule (ξ, g) of G .

With this notation the equivalence of \mathcal{C} with the set of all supports F_r^+ : $r \in \bar{S}^+(X)$ is trivial.

Let G be given, and define for each $\xi \in \bar{X}$ the polynomial σ_ξ as the sum $\sum g$ extended to all g so that $(\xi, g) \in G$.

If we interpret the support of w_ξ as the set $D_X(\xi, G)$, it is clear that any equation $w_\xi = \bar{\sigma}w$ can be interpreted as describing $D_X(\xi, G)$ as the union of the sets $D_X(g, G)$ ($(\xi, g) \in G$) obtained by replacing in g every letter ξ' by a terminal word $f \in D_X(\xi', G)$.

Conversely, let us assume that $\bar{\sigma}$ is such that $\langle \sigma_\xi, g \rangle \geq 0$ for all $\xi \in \bar{X}$ and $g \in F(X, \bar{X})$.

By introducing enough new variates ξ' , we can find $\bar{\sigma}'$ which is such that $\langle \sigma_{\xi'}^1, g \rangle = 0$ or 1, and the new polynomials $\sigma_{\xi'}^1$ reduce to old polynomials σ_ξ when the new variates ξ' are identified with the old ones in a suitable manner. Furthermore, for every new ξ' (corresponding to the old variate ξ) we add an equation $\sigma_{\xi'}^1$, identical to σ_ξ .

Thus the original w_ξ is equal to a sum $\sum w_{\xi'}^1$, (with $w' = \bar{\sigma}'w'$) and $\bar{\sigma}'$ can be associated with a grammar in a unique fashion, since $\langle \sigma_{\xi'}^1, g \rangle = 0$ or 1.

This interpretation throws some light on the other families. Thus, $\bar{S}_1^+(X)$ corresponds to the family \mathcal{C}_1 of the context-free languages in which every rule has the form $(\xi; f'\xi'f'')$ or (ξ, f) with $f, f', f'' \in F(X)$.

In turn, $\bar{S}_0^+(X)$ is obtained by restricting the rules to have the form $(\xi, \xi'f)$ or (ξ, f) with $f \in F(X)$.

Observe now that, in any case, the coefficient $\langle w_\xi, f \rangle$ of the word f expresses the number of distinct factorizations of f according to the rule of grammar. Thus, for

(XII. LINGUISTICS)

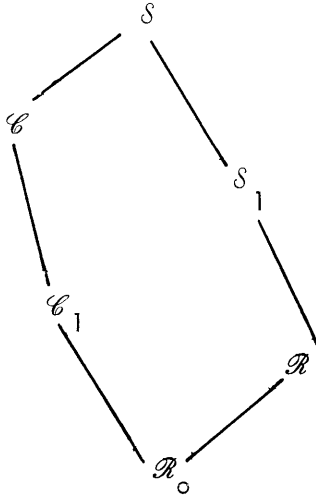


Fig. XII-2.

any two $r, r' \in \bar{S}^+(X)$, the support of $r-r'$ consists precisely of those words that have a different number of factorizations in the two grammars associated with r and r' , respectively.

Reciprocally, given any $r'' \in \bar{S}(X)$, it is easy to prove that $r'' = r-r'$ for at least one pair $r, r' \in \bar{S}^+(X)$, and the same is true for \bar{S}_1 and \bar{S}_1^+ or for \bar{S}_0 and \bar{S}_0^+ .

Summarizing our remarks, we obtain (on top of the family of the finite subsets) the six families illustrated in Fig. XII-2. Here, \mathcal{S} and \mathcal{S}_1 correspond to $\bar{S}(X)$ and $\bar{S}_1(X)$, respectively. In order to prove that these six families are all different and do not enjoy further inclusion relations, it would be enough to

build three subsets, say F_1, F_2, F_3 of $F(X)$ having the following properties:

- $R_1 \in \mathcal{R}, F_1 \notin \mathcal{C}$
- $F_2 \in \mathcal{C}_1, F_2 \notin \mathcal{R}$
- $F_3 \in \mathcal{C}, F_3 \notin \mathcal{S}_1$.

I am not able to construct a set such as F_3 ,¹⁰ but there exists an F_4 which is such that $F_4 \in \mathcal{C}$ and $F_4 \notin \mathcal{C}_1$. Thus the only possible diagrams apart from Fig. XII-2 are Fig. XII-3a and 3b. Again, there is no further inclusion relation. In fact, it seems

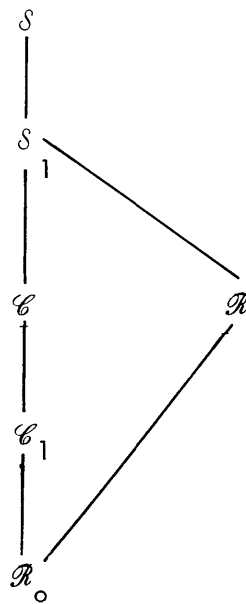


Fig. XII-3a.

or

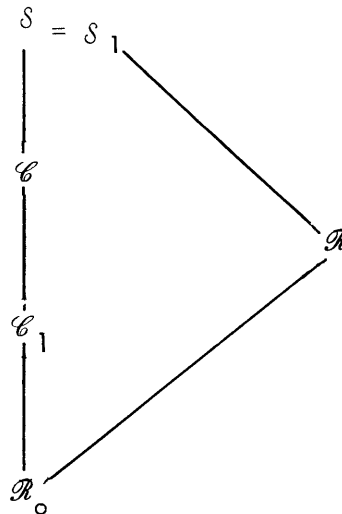


Fig. XII-3b.

most unlikely that $\mathcal{C} \subset \mathcal{S}_1$, and the original scheme probably represents the true situation.

The counterexamples F_2 and F_3 are very simple:

Let $F_2 = \{x_1^n x_2 x_1^n : n \geq 0\}$. This set is produced by $G = \{(\xi, x_1 \xi x_1), (\xi, x_2)\}$, and thus $F_2 \in \mathcal{C}_1 \subset \mathcal{C}$. On the other hand, it is known that F_2 does not belong to \mathcal{R} .⁸

Let $F_3 = \{x_1^n x_2 x_2^{n'} : n, n' \geq 0, n \neq n'\}$. It is known⁸ that $F_3 \in \mathcal{R}$, and it is not difficult to show that F_3 does not belong to \mathcal{C} because of the relatively simple structure of a grammar G which produces infinite sets of the form $\{x_1^n x_2 x_1^{n'} : n, n' \text{ linked by a certain relation}\}$.

Indeed, as the reader can verify, any set of this type is a finite union of finite sets and of sets having the form:

$\{x_1^{n+N} x_2 x_1^{n'+N'} : N, N', n, n' \geq 0; n \equiv 0 \pmod{p}, n' \equiv 0 \pmod{p'}\}$ for some integers N, N', p, p' .

(The proof is based upon the fact that, when X has a single letter, \mathcal{C} reduces to the family of regular events.)

For the construction of F_4 we need a more explicit description of \mathcal{C}_1 :

F' belongs to \mathcal{C}_1 if and only if there exist

(a) A finite set Y ;

(b) Two mappings ϕ and Φ from $F(Y)$ to $F(X)$ that are a homomorphism and an anti-isomorphism (i. e., $\Phi g g' = \Phi g' \Phi g$);

(c) A regular event $G' \subset F(Y)$ that is such that $F' = \{\phi g \Phi g : g \in G'\}$.

The proof of this statement follows the same lines as Chomsky's proof¹ of the fact that the support of any $r \in \overline{\mathcal{S}}_0^+(X)$ is a regular event.

The same technique, of course, is valid for the more general case of $\overline{\mathcal{S}}_1(X)$ (with the obvious modifications) and it displays every element of \mathcal{C}_1 obtained by the three following steps:

1. Taking the words g from some regular event on $F(\square)$;

2. Forming the products $g \xi^* \tilde{g}$, where ξ^* is a new symbol, and \tilde{g} is the "mirror image" of g ;

3. Making a transduction θ of \tilde{g} and of g into $F(X)$, and erasing ξ^* .⁷

Let us now return to our problem. For any $f \in F(X)$ ($X = \{x_1, x_2\}$) let λf denote the difference between the number of times x_1 and x_2 appear in f .

We claim that F_4 belongs to \mathcal{C} and not to \mathcal{C}_1 , where $F_4 = \{f : \lambda f = 0; \lambda f' > 0 \text{ for all proper left factors of } f\}$.

The first part of the claim has already been verified (Example 2).

Let us now observe that if $F' \in \mathcal{C}_1$ is such that for all integers $n < 0$ there exists a $g \in F(\square)$ which is such that $g_1 g g_2 \in G'$ for some g_1, g_2 , and that $\lambda \phi g < n$, then $F' \neq F_4$. Indeed, since G' is a regular event, there exists a finite set of pairs (g_1, g_2) which are such that for any $g \in F(\square)$ either $F(\square) g F(\square) \cap G'$ is empty, or else

(XII. LINGUISTICS)

$g_1 g_2 \in G'$ for some of these pairs. Thus, under this hypothesis, we can find that $f = \phi g_1 g_2 \in F'$ which is such that its left factor $f' = \phi g_1$ satisfies $\lambda f' < 0$, and thus $f \in F'$ and $f \notin F_4$.

It follows that if $F'' \in \mathcal{C}_1$ is contained in F_4 , we can find a large enough integer n' which is such that no $f \in F''$ has a factorization $f = f_1 f_2 f_3 f_4$ with $\lambda f_1 > n$; $\lambda f_1 f_4 = 1$; $\lambda f_1 f_2 f_3 > n$ (and $\lambda f_1 f_2 f_3 f_4 = 0$ because by hypothesis $f \in F_4$). Since clearly F_4 contains such words, we have proved that $F'' \in \mathcal{C}_1$ and $F'' \subset F_4$ implies $F'' \neq F_4$; that is, $F_4 \notin \mathcal{C}_1$.

These remarks can be pictorially expressed by saying that the words of F'' have, at most, one arbitrarily high peak. It follows from the definition of F_4 that this last set contains words having an arbitrary number of arbitrarily high peaks. Thus, incidentally, we have proved the stronger result that \mathcal{C} is different from the family of subsets obtained from \mathcal{C}_1 by closure under a finite number of set products or set unions.

5. Some Miscellaneous Remarks

a. As an easy source of counterexamples we could consider the special case of X reduced to a single element because then no difference exists between commutative and noncommutative power series.

The results known thus far contribute to the statement that in this case \mathcal{C} and \mathcal{S}_1 are equivalent to \mathcal{R}_0 . No result is known for \mathcal{S} .

However, although the proofs that $\mathcal{C} = \mathcal{R}_0$ and that $\mathcal{S}_1 = \mathcal{R}$ are quite easy, the proof that $\mathcal{R} = \mathcal{R}_0$ is a rather deep theorem of Skolem.⁸ Nonetheless, the fact that when $X = \{x\}$ any $r \in \overline{\mathcal{S}}(X)$ is the Taylor series of some ordinary algebraic function of n allows us to construct simple families of sets that cannot belong to \mathcal{S} .

A rather general instance is the family of the infinite sets $\left\{ x^{N_1}, x^{N_2}, \dots, x^{N_m}, \dots \right\}$ which have the property that $\lim_{m \rightarrow \infty} \frac{N_{m+1}}{N_m}$ is infinite (i. e., which have the property that the ratio N_{m+1}/N_m exceeds for some finite m any prescribed finite value).

In order to prove that no set of this type belongs to \mathcal{S} we consider any $r \in \overline{\mathcal{S}}(X)$ ($X = x$). Without loss of generality we may assume that $\langle r, e \rangle = 0$. By definition, $r = a_0 + \sum_{i=2}^m a_i r^i$, where m is finite and the a_i 's are polynomial in x . By comparing the two members of the equation, we see that for each n $\langle r, x^n \rangle$ must be equal to a linear combination with fixed coefficients of sums of the type $\sum \left\langle r, x^{n_1} \right\rangle \left\langle r, x^{n_2} \right\rangle \dots \left\langle r, x^{n_{m'}} \right\rangle$ extended to all representations of $n-h$ as a sum $n_1 + n_2 + \dots + n_{m'}$, where $h \geq 1$ is bounded by the degrees of the a_i 's, and m' are bounded by the degree m of the equation. It follows that if N is such that $\langle r, x^{N+k} \rangle = 0$ for $0 \leq k < mN$, then $\langle r, x^{n'} \rangle = 0$ for all $n' \geq N$; i. e., r is a polynomial.

Since the condition imposed on the set $\left\{ x^{N_1}, x^{N_2}, \dots \right\}$ amounts to the existence of at least one such N for every finite m , our contention is proved.

A similar method could be applied to show that $\{x^{n^2} : n > 0\}$ does not belong to \mathcal{S} .

b. Our next example shows that the intersection problem even for so restricted a family as \mathcal{C}_1 is an undecidable one.

Let $\overline{\Sigma} = \{\xi\}$, $X = \{a, b, c\}$ and consider the two grammars:

$$G = \{(\xi, a\xi a), (\xi, b\xi b), (\xi, c)\}$$

$G' = \{(\xi, c), (\xi, f_i \xi f_i') : i \in I\}$, where $(f_i, f_i') (i \in I)$ is an arbitrary set of pairs of elements from $F = F(a, b)$.

The language $D_X(\xi, G)$ is a special instance of Chomsky's mirror-image languages and there exists an $f \in D_X(\xi, G) \cap D_X(\xi, G')$ if and only if one can find a finite sequence $i_1, i_2, i_3, \dots, i_n$ of indices such that the word $f_{i_1} f_{i_2} \dots f_{i_n}$ is equal to the mirror image of $f_{i_n}' f_{i_{n-1}}' \dots f_{i_2}' f_{i_1}'$. Thus clearly the intersection problem for G and G' is equivalent to the classical correspondence problem of Post⁵ and since this last one is undecidable, our contention is proved.

c. It may be mentioned that other principles could be used for distinguishing interesting subsets of words. For example, Ginsburg and Rice² have shown that \mathcal{C} contains as a proper subset the family \mathcal{C}' corresponding to the case in which the set of equations $w = \bar{\sigma}w$ has the following property which these authors call the "sequential property": There exists an indexing $\xi_1, \xi_2, \dots, \xi_m$ of the variates $\xi \in \overline{\Sigma}$ which is such that for all j the polynomial $\sigma \xi_j$ does not involve the variates $\xi_{j'}$, with $j' > j$.

In Chomsky's terminology this means that no $\xi_{j'}$ ($j' > j$) appears in a word g that is such that $(\xi_j, g) \in G$. (Then, clearly the rewriting process must be started from $\xi_0 = \xi_m$).

Another possibility is to consider the subset $\bar{S}^1(X)$ of those $s \in \bar{S}(X)$ that are such that $\langle s, f \rangle = 0$, or 1, for all f .

It has been shown by Parikh⁴ that there exist sets of words in \mathcal{C} (in fact, in the closure of \mathcal{C}_1 by finite union and set product) which cannot be the support of an $s \in \bar{S}^1(X)$ having this property.

In our notation, Parikh's example is described as follows:

$$\sigma_{\xi_0} = \xi_1 \xi_2 + \xi_4 \xi_3; \quad \sigma_{\xi_1} = x_1 \xi_1 x_1 + x_1 \xi_2 x_1; \quad \sigma_{\xi_2} = x_2 + x_2 \xi_2;$$

$$\sigma_{\xi_3} = x_2 \xi_3 x_2 + x_2 \xi_4 x_2; \quad \sigma_{\xi_4} = x_1 + x_1 \xi_4.$$

From this reasoning we deduce the following equations in which, for short, w_i denotes the coordinate of w whose index is ξ_i :

$$w_0 = w_1 w_2 + w_4 w_3;$$

$$w_1 = x_1 (w_1 + w_2) x_1$$

$$w_3 = x_2 (w_3 + w_4) x_2$$

$$w_2 = x_2 + x_2 w_2$$

$$w_4 = x_1 + x_1 w_4.$$

These equations can easily be solved because they are "sequential" in the sense of

(XII. LINGUISTICS)

Ginsburg and Rice. Indeed, the last equation can be written $w_4 - x_1 w_4 = x_1$. That is, $(e-x_1) w_4 = x_1$ and we have $w_4 = (e-x_1)^{-1} x_1 = \sum_{n>0} x_1^n$.

Similarly, $w_2 = (e-x_2)^{-1} x_2 = \sum_{n>0} x_2^n$. Thus $w_3 = x_2 w_3 x_2 + \sum_{n>0} x_2 x_1^n x_2$, and, consequently, $w_3 = \sum_{n>0} \sum_{m>0} x_2^m x_1^n x_2^m$; $w_1 = \sum_{n>0} \sum_{m>0} x_1^m x_2^n x_1^m$. Thus we finally obtain

$$w_0 = \left(\sum_{n>0} \sum_{m>0} x_1^m x_2^n x_1^m \right) \left(\sum_{n'>0} x_1^{n'} \right) + \left(\sum_{m'>0} x_1^{m'} \right) \left(\sum_{n>0} \sum_{m>0} x_2^n x_1^m x_2^n \right)$$

$$= \sum x_1^m x_2^n x_1^{m'} x_1^{n'} r(m, n, m', n')$$

The last summation is, after all, quadruples (m, n, m', n') of positive integers, and the coefficient $r(m, n, m', n')$ has the following values:

$$\begin{aligned} r(m, n, m', n') &= 0 \text{ if } m \neq m' \text{ and } n \neq n' \\ &= 1 \text{ if } m \neq m' \text{ and } n = n' \\ &= 1 \text{ if } m = m' \text{ and } n \neq n' \\ &= 2 \text{ if } m = m' \text{ and } n = n'. \end{aligned}$$

The fact that this coefficient is equal to 2 for certain words exactly measures the "ambiguity" of the grammar. It would be interesting to give examples in which this grammatical ambiguity is unbounded.

I mention that conversely the following process gives elements $s \in \bar{S}^+(X)$ with $\langle s, f \rangle = 0$, or 1.⁹

Let ϕ be a homomorphism of $F(X)$ into a finite monoid H (i. e., let us consider a finite automaton), and β a mapping that assigns to every pair $(h, h') \in (H, H)$ an integer $\beta(h, h')$. For any word $f \in F(X)$ let $\beta^* f$ be the sum $\sum \beta(\phi f_1, \phi f_2)$ extended to all factorizations $f = f_1 f_2$ of f , and say that f is accepted if and only if $\beta^* f$ does not belong to a prescribed finite set Z' of integers.

Then the formal sum $s = \sum f'$ extended to all f' which are not accepted (i. e., $s = \sum \{f' : \beta^* f' \in Z'\}$) belongs to $\bar{S}^+(X)$.

An equivalent definition⁸ is: Let u be a representation of $F(X)$ by finite integral matrices u_f and assume that there exists a constant K which is such that for all words f the value $(u_f)_{1, N}$ of the $(1, N)$ entry of u_f is, at most, equal to K times the degree (length) of f .

Then the set of all f with the property that $(u_f)_{1, N} \neq 0$ is the set of the words accepted by an algorithm of the type described above (and reciprocally). As an auxiliary result, we have shown that the complement of a set F' belonging to the simplest subfamily of \mathcal{R} which is different from \mathcal{R}_0 belongs itself to the far higher family \mathcal{C} . In general, the complement of a set from F' does not.

Trivially, this construction applies to sets of words defined by the condition that some linear function of the number of times each letter $x \in X$ appears in them has a

given value. It is quite remarkable that the sets defined by two or more such constraints (for instance, the sets of words which contain the same number of times x_1 , x_2 and x_3 or the set $\{x_1^n x_2^n x_3^n = n \geq 0\}$) do not seem to have any relation to \mathcal{C} .

I conclude these rather disconnected remarks by an interesting construction of \mathcal{C} which is due to Parikh and which can also be applied to $S(X)$.

6. Parikh's Construction⁴

Let us consider a grammar G satisfying the usual conditions and extend to a homomorphism $A \rightarrow A$ the mapping $j: \boxed{\cdot} \rightarrow A$ defined by $j\xi = \sum \{g: (\xi, g) \in G\}$.

For any $g \in F(X \cup \boxed{\cdot})$, the support of g is the set of all words which can be derived from g by the application of one rule of G to each of the occurrences of a symbol $\xi' \in \boxed{\cdot}$. Every element of this set has either a strictly larger total degree (length) than f or the same total degree but a strictly larger partial degree in the variates $x \in X$. Thus the supports of the elements $f, jf, j^2f, \dots, j_f^n, \dots$ are all disjoint. Their union, say F' , is a subset of the set $D^*(f, G)$ of all words derivable from f .

Of course, F' is, in general, different from $D^*(f, G)$ because of the extra condition that every $\xi' \in \boxed{\cdot}$ is rewritten at each step. However, when considering only the intersection $D^*(f, G) \cap F(X) = D_{\mathbf{X}}^*(f, G)$ we have $F' \cap F(X) = D^*(f, G) \cap F(X)$, since in order to get an element $f \in F(X)$ we have to rewrite each $\xi \in \boxed{\cdot}$ at least once at one time or another.

Let us now denote by u the sum $\sum \{\xi: \xi \in \boxed{\cdot}'\}$ for any subset $\boxed{\cdot}'$ of $\boxed{\cdot}$. The element $t = u + \sum_{n>0} j^n u$ belongs to A , as we have seen, and it satisfies the Schröder-like equation $u + jt = t$.

Conversely, we can write $t = (\epsilon - j)^{-1} u$, where ϵ is the identity mapping $A \rightarrow A$. Let δ_0 denote the retraction $A(X \cup \boxed{\cdot}) \rightarrow A(X)$ induced by $\delta_0 \xi = 0$ for each $\xi \in \boxed{\cdot}$; (a retraction is a homomorphism that allows a subset invariant and sends everything else into this subset; here the subset is that of the words not containing a single $\xi \in \boxed{\cdot}$.)

Example. $\boxed{\cdot} = \{a, \beta\}$; $X = \{a, b\}$; $\boxed{\cdot}' = \{a\}$
 $G = \{(a, aa\beta), (a, a), (\beta, b)\}$.

We have

$$ja = aa\beta + a$$

$$j\beta = b.$$

Thus $u = a$; $ju = a + aa\beta$;

$$j^2 u = (a+aa\beta)(a+aa\beta) b = a^2 b + aa\beta b + aa\beta ab + aa\beta aa\beta b$$

$$j^3 u = a(a+aa\beta)(a+aa\beta) bb + (a+aa\beta)(a+aa\beta) bab + ((a+aa\beta)^2 b)^2 b \\ = a^3 b^2 + a^2 ba^2 b^2 + \text{terms of degree } \geq 1 \text{ in the } \xi \in \boxed{\cdot}, \text{ etc.}$$

The support F'_t of $\delta_0 t$ is the set of the well-formed formulas in Lukaciewicz notation.

M. P. Schützenberger

(XII. LINGUISTICS)

References

1. N. Chomsky, On certain formal properties of grammars, *Information and Control* 2, 137-167 (1959).
2. S. Ginsburg and H. G. Rice, Two Families of Languages Related to Algol, Technical Memorandum, System Development Corporation, Santa Monica, California, January 1961.
3. H. Kesten, *Trans. Am. Math. Soc.* 92, 336-354 (1959).
4. R. J. Parikh, Language-generating devices, Quarterly Progress Report No. 60, Research Laboratory of Electronics, M.I.T., January 15, 1961, pp. 199-212.
5. E. L. Post, A variant of a recursively unsolvable problem, *Bull. Am. Math. Soc.* 52, 264-268 (1946).
6. M. O. Rabin and D. Scott, Finite automata and their decision problems, *IBM J. Res. Develop.* 3, 114-125 (1959).
7. M. P. Schützenberger, A remark on finite transducers, *Information and Control* 4, 185-196 (1961).
8. M. P. Schützenberger, On the definition of a class of automata, *Information and Control* 4, 245-270 (1961).
9. M. P. Schützenberger, Un problème de la théorie des automates (Séminaire Dubreil Pisot, Paris, 1959-1960).
10. (Added in Proof): Since this paper was written I have obtained a set F_3 and, consequently, the true diagram is Fig. XII-2.