PAUL BARAN
1425 East Dorothy Lane
Fullerton, California

Oct 7, 57

Mr. R. J. Solomonoff
Technical Research Group
New York 3, N. Y.

Dear Mr. Solomonoff:

This writer is presently engaged in graduate study at the University of California at Los Angeles and would appreciate receiving a copy of your paper entitled, "An Inductive Inference Machine", a privately circulated report, used as a reference to your paper of the same title presented at the 1957 National IRE Convention. As I wish to discuss your IRE paper in detail at a seminar Oct. 15th if possible a most rapid reply would be appreciated.

Sincerely yours,

Paul Baran

Paul Baran

rcvd Sat nite
replied Sunday Oct 13,57
sent Airmail with letter
asked report to be returned
in ~ 1 month

1425 E. Dorothy Lane
Fullerton, Calif.

October 27, 1957

Mr. R. J. Solomonoff
26 Boylston St.
Cambridge 38, Mass.

Dear Mr. Solomonoff,

My most appreciative thanks to your airmailing a copy
of your paper "An Inductive Inference Machine" to me.
Because of a delayed schedule, it has arrived in more than
adequate time to be of excellent use.

The seminar that I mentioned is one on digital computers
in general with the session of interest on advanced "switching
theory" in particular.   The whole area of self-organizing
logic computation seems to be a most active one and my
appetite was whetted by the first sentence of your paper--
"A machine is described which is designed to operate as human
beings seem to."

I have thoroughly enjoyed reading your complete paper,
and would like to congratulate you on your invention of a
most interesting and novel approach, although I have some
practical type reservations on things like the memory capacity
and some other details.

I might mention that in contrast to the enjoyment of
reading your complete paper, I thought the condensation that
appeared in the IRE National Convention Record came a bit
short of adequately describing your ideas.   Ah, but I guess
such is the price of condensing 55 pages into 5½ pages.

I shall return your paper as soon as I have an
opportunity to have it duplicated.

Many thanks,

Paul Baran

Paul Baran

9-6 57

Dear Solomonoff,

I am now losing hope of being able to see you in N.Y.C. May I ask you then to send me some reprints of your papers — I do not see any of the IRE Publications.

As for my recent work, a ½ book, to be published by Presses Universitaires de France, has been translated by a student of A. Oettinger, and must be available thru Toni. Another lengthy paper will be translated by IBM, and be available thru M. Kochen, at IBM Lab. in Ossining, N.Y.

Next year, & 1957-8 I am joining the University of Lille, in Northern France.

Yours

Mandelbrot

Best address at present:

17, rue des Petites Ecuries
Paris, France

December 24, 1957

Dr. Warren S. McCulloch
Research Laboratory for Electronics
Massachusetts Institute of Technology
Cambridge, Massachusetts

Dear Warren,

Ray Solomonoff was leaving for a holiday trip to New York on Monday, but before going, he prepared the enclosed informal statement of his interests. He asked me to add whatever supplementary information I thought might be useful. Ray is definitely seeking support for his researches in inductive inference. At present, he is operating on his savings, though this cannot long continue.

It is our assessment that one of the early pay-offs of his work in inductive inference will be in information retrieval -- in the area of machine assignment of descriptors or index terms. However, it will take a lot more preliminary work before he is really ready to do such sophisticated operations on text.

The current status of his work is at about the point indicated in his IRE paper, "An Inductive Inference Machine". Refer to the items in Section 4 of the enclosed reprint. It is still very much in the pencil-and-paper, basic thinking stage. This fall, he has been working particularly on the problems of assignment of utility or priority to the n-gram prediction patterns and pattern transformations. He has also been learning about information retrieval. He has a good deal of careful thinking yet to do before some simple tests are ready for machine programming.

About this machine programming: some estimates at NBS indicated that exact programming of this approach would be very slow on existing machines. However, it is my speculation that some of my random pattern coding could give a good probabilistic computation of the n-grams and transformations, and if so, it would speed things up immensely. Also, it would have features that probably are closely related to the way the brain works.

In general, I find Ray's work the most interesting and fundamental of the current work I am familiar with on artificial intelligence.

Sincerely yours,

Calvin.

Calvin N. Mooers

encl. IRE reprint
RJS letter

Dec 22, 1957

For Warren McCulloch :

The main problem I am interested in, is the Artificial
Intelligence problem. The particular approach I have
chosen, is ~~through~~ the analysis of inductive inference, and
attempts to mechanize the inductive inference process.
Perhaps the simplest way in which inductive inference
techniques are applicable to the information retrieval
problem, is through the ~~automated~~ assignment of descriptors
or subject classifications to documents by machine —
using various types of feedback from the information user
to optimize descriptor assignment.

As an example, let us give the machine a set
of documents, and with each document, a set of
descriptors or subject classifications that have been given
to them by human analysis. Using this set of
document, descriptor pairs, the machine will
formulate various hypotheses about the relationships
between the descriptors and the words and phrases
in the documents. These hypotheses will then be
tested using the sample set of descriptor-document
pairs. The successful hypotheses will be retained
in the machine memory, and used to construct
new promising trial hypotheses for testing.

Some methods

~~of~~ of formulating trial hypotheses have ~~been~~ ~~to some extent~~ been devised and tested for ~~the~~ simple arithmetic problems, and have been tested to a much lesser extent for prediction of English grammatical sequences.

Techniques of inductive inference, are also useful in information retrieval systems more complex than those using descriptor or subject classifications. In all cases, the systems can be optimized with respect to feedback from the information user and/or with respect to feedback from "authorities" who are given more apriori weight than the ordinary library user.

— Ray Solomonoff —

January 16, 1958

Dr. Burton Adkinson
National Science Foundation
Washington 25, D. C.

Dear Burton:

Enclosed is an Addenda to Ray Solomonoff's Proposal for Sponsored
Research on Inductive Inference that we left with you before Christmas.
This Addenda fills in certain gaps in the presentation. In particular,
it describes in some detail just what are the next steps that he proposes
and the cost amounts. The text of this Addenda follows right along after
the text of the proposal. If you should require additional copies of the
complete proposal, we should be glad to furnish them to you at your request.

We realize that the technical evaluation of proposals of this
character is quite difficult. For this reason, Ray suggests that any
of the following people, who are familiar with his work on inductive
inference to some extent, could give their expert comments. The first
of these is Dr. Peter Elias at Massachusetts Institute of Technology
who had personally reviewed Ray's work and has a current request for
Ray to prepare a paper on his theory of inductive inference for publication
in Elias' new international journal on information theory. Several
people on Elias' review panel probably also looked over Ray's work, and
Elias could furnish their names. Dr. Claude E. Shannon, who is now at
the Center for Advanced Study in Behavioral Sciences, Stanford, California,
is also quite familiar with the work. However, since Shannon is deluged
with letters from all over, it is difficult to get an answer out of him.
Dr. Marshall C. Yovits, Office of Naval Research, Washington, D. C.
has recently familiarized himself with Ray's work. In this same category
is Mr. Russell A. Kirsch at the National Bureau of Standards. At the
Lincoln Laboratories, Bedford, Massachusetts, Dr. Marvin Minsky and
Mr. Roland Silver are both quite familiar with his work. We trust that
these suggestions will be of some assistance to you.

In order to see how things are going, we hope to telephone your
office at the end of next week. I shall be glad to furnish any additional
information as needed.

Sincerely,

Calvin

Calvin N. Mooers
Proprietor

CNM:mw
encl. addenda

January 30,1958

Dr. Marshall C. Yovits
Information Systems Branch Code 437
Office of Naval Research
Washington 25, D. C.

Dear Marshall:

Talking to Lea Bohnert yesterday reminded me that I really do owe you a letter. Since our trip to Washington, Ray has prepared a little more information in the form of an Addenda to his proposal, a copy of which I believe he left with you. Thus I am sending herewith a complete copy of his proposal as it is in the present state including the Addenda. The additional information that you will find in it will be of assistance in making clearer just what Ray has in mind for research on inductive inference with a slant towards information retrieval. In my opinion, the importance of "inductive inference machines" or "artificial intelligence" as applied to information retrieval cannot be underestimated. From my long background in working out both the theory and practise of information retrieval, I can certainly say that this is the key to the long-range future developments.

In our talk at the Computer Conference, you asked that I write to you a little bit of what my particular interest in the way of future work is. As you know, I have been working on behalf of the Bureau of Standards, and have so far gotten two papers out as the result of my studies. The first (No. 111) "Some Mathematical Ideas Needed for a Retrieval Theory" is a very general discussion for the assistance of non-mathematicians. The second (No. 114) "A Mathematical Theory of Language Symbols in Retrieval (Part I)" is a more advanced paper. I am extending and revising it in a paper to be presented at the International Conference on Scientific Information this fall. The results of my studies of the last year have covered a good deal more ground than has been boiled out into these particular papers. In order to give a little indication of what is still to come, I have prepared an informal prospectus of what seemed to be some of the papers that need to be written in this general area. I am enclosing a copy of it, but please regard it as a private communication to your office at this time. This "Prospectus of Papers in Preparation" describes a number of points of theory that need filling in for the development of a foundation theory

for information retrieval. What I feel is needed is a solid theoretical basis for the information retrieval systems we now have, and once these postulates are brought out into the open so we can examine them, these postulates and systems will invariably direct us to more elaborate conceptions of retrieval that can well be put into use. For another thing, the development of a competent retrieval theory will be most valuable in guiding the work on the application of inductive inference machines. In the early stage of their development and application, these machines will not be able to do the whole problem of discovering internally and empirically how to do retrieval. They will have to be guided by certain general principles. Thus we need a very good general theory to form the skeleton on which inductive inference machines o f a simple character can work. It is in this sense that I am particularly desirous of having the opportunity to work in loose association with Solomonoff. There should be an area of easy informal contact between these two endevours, without warping any of them seriously in the direction of the other at too early a stage.

It is my desire to direct my work in the next year or two along the general course pointed out by the "Prospectus of Papers in Preparation". (Of course there will be other topics that will arise as my thinking progresses.) One of the very good possibilities, for which the time begins to appear to be ripe, is to mould some of this output into a form suitable for a book on the theory of information retrieval systems. However, I have taken no formal steps in this direction. Support for my work, for the period after April of this year, is in a very fluid state. From the scholarly standpoint, there would be some very real advantages in not having to slant my thinking to the peculiar problems of the Patent Office.

Along the above line of interest as marked out in the "Prospectus of Papers in Preparation" and in particular with a bias towards an inclination to mediate between advanced theoretical work in information retrieval and the problems of inductive inference, there may be an area suitable for the delineation of a research proposal with ONR. I should appreciate any informal comments you may have.

Lea Bohnert mentioned the discussions with you in regard to plans involving her and Bar-Hillel and the matter of "putting an institutional roof over the head of the project". Though I know there are certain limitations, might I just advance the suggestion that Zator Company be considered as a possibility. I have the space here and Cambriage is a desirable center to work from.

Both Ray and I would appreciate your informal reactions to our several prospective lines of activity. Give my regards to Gordon Goldstein.

Sincerely,

Calvin N. Mooers
Proprietor

CNM:mw
encl. Solomonoff Proposal, Mooers Prospectus of Papers

CONVAIR

A DIVISION OF GENERAL DYNAMICS CORPORATION

February 2, 1959

Dr. Raymond Salmonoff
Zator Company
Cambridge, Massachusetts

Dear Dr. Salmonoff:

The SIAM News Letter inidcates that you present-
ed a talk on the subject of "Mechanization of Inductive
Inference", on December 10, 1958. I am very much
interested in this subject and I wonder if you have
any written material which I might borrow on either a
short term or permanent loan basis so that I may study
your contribution.

Thank you very much for your cooperation.

Sincerely,

Lawrence J. Fogel
Head, Reliability Group
Mail Zone 6-141

LJF/ms

Sent him Dort Rap.
and IRE paper
· will sent 2 or 3 long. papers
when available.

February 9, 1959

Dr. L. J. Fogel, Head
Reliability Group
Mail Zone 6-141
Convair Div. General Dynamics Corp.
San Diego 12, California

Dear Dr. Fogel:

I am enclosing a preprint of "An Inductive Inference Machine" and an earlier, longer report of identical title. My more recent work is on pattern discovery through generalized languages which treats the simple arithmetic learning problems in a more general way. I shall send you reports on some of this work in the near future. I would appreciate any comments or questions that you might have on these papers.

Sincerely,

R. J. Solomonoff

RJS: mw
encl. IRE paper
        Dartmouth report

CONVAIR

A DIVISION OF GENERAL DYNAMICS CORPORATION

February 24, 1959

Dr. Raymond Salmonoff
Zator Company
140 1/2 Mount Auburn St.
Cambridge 38, Massachusetts

Dear Dr. Salmonoff:

Thank you so much for sending all the information relative to the Inductive Inference Machine. It certainly presented me with much which requires careful study and detailed analysis. Let me assure you that I will find the time to study the contents of your papers in the near future, and I am sure that this effort will be well spent.

I certainly appreciate your cooperation.

Sincerely,

Lawrence J. Fogel
Head, Reliability Group
Mail Zone 6-141

LJF/ms

April 9, 1959

Dr. Noam Chomsky
Institute for Advanced Study
Princeton, New Jersey

Dear Dr. Chomsky:

Enclosed are two papers, giving two methods for discovering the grammars of phrase structure languages, in a training situation similar to that which you and Miller used for finite state grammar discovery.

I would be much interested in any questions or comments that you may have on these papers.

Recently, I have been working on methods for discovering the grammars of phrase structure languages, in which a "teacher" is not available. Ordinarily, the solution is not unique and probabilistic a priori hypotheses about the nature of the grammar must be made.

Another topic of much recent interest, upon which I hope to soon publish a report and/or paper is the generalization of phrase structure languages to what one might call "phrase structure patterns". These "phrase structure patterns" include multidimensional patterns, ordinary transformation languages, some kinds of arithmetic learning, and certain kinds of mechanical translation as special cases. The kind of mechanical translation given in "The Mechanization of Linguistic Learning" (pp. 10-15) is one simple example. Simple though it is, it is much more powerful than the mechanical translation scheme demonstrated by Georgetown University in 1954 on an IBM 701.

The importance of this particular generalization of phrase structure languages, is that the grammars of many of them can be discovered by methods substantially the same as those used for ordinary phrase structure languages.

With regard to phrase structure languages, I have two questions that you may be able to help me with.

1) Do phrase structure languages form a Boolean algebra -- or, equivalently, is the complement of a phrase structure language also a phrase structure language? You and Miller have shown this for finite state languages. This question is, to some extent, relevant to the problem of discovering "approximate grammars" -- i. e. simple grammars that fit a given corpus optimumly -- using a suitable fidelity criterion.

2) Is a language employing context dependent substitution (e. g. $abc \rightarrow adec$) in its grammar expressible as an ordinary terminal phrase structure language that does not have context dependent substitution in its grammar? Or, more briefly, does context dependent substitution of this type expand the domain of phrase structure languages? In

section 3.4 of "Three Models . . ." you seem to imply that there is no expansion of domain, but it is not entirely clear as to whether this is what you meant.

I had been discussing some of my work with G. Miller recently. He suggested that you might be much interested in some kinds of English sentences that I have found, which are not describable by phrase structure grammars. These are sentences containing the word "respectively". An example is: "George, Jane, Bill and Maria phoned his, her, his and her parents, respectively." The important things are that the proper names and the possessive pronouns are correlated, and are in the same order; also the number of subjects (four in the above example) must be arbitrarily increasable. If they were in reverse order, we could express this with a phrase structure grammar. It may be that examples of the use of "respectively" in languages other than English, can be more easily found, which cannot be expressed as phrase structure languages. This is because in English we correlate words and their modifiers by sex, number and tense only -- while in other languages, many genders and cases may exist.

I do not know if the above sentence will convince anyone that English is not a phrase structure language. It is a fairly artificial sentence -- and is unlikely to be used, ordinarily. Also, they would be rejecting a phrase structure description of English for the wrong reasons! As you point out, the phrase structure description of English is usually unacceptable, because it is a far more complex description of English than a transformational description -- not because counter examples to phrase structure description are so common.

Could you send me reprints of the following papers:

     1) "Finite State Languages", Information and Control, May 1958, pp. 91-112.

     2) "Semantic Considerations in Grammar: Monograph 8".

     3) "Systems of Syntactic Analysis", Journal of Symbolic Logic 18.242-56(1953).

I will send you a copy of the paper on generalizing phrase structure languages, as soon as it is available.

                             Most sincerely,

                             R. J. Solomonoff

RJS:mw

April 9, 1959

Professor Benoit Mandelbrot
18 Rue du Douanier
Paris 14, France

Dear Professor Mandelbrot:

Enclosed are two reports on how to discover the grammars of phrase structure languages in a suitable training situation. They require, however, a "teacher" -- and I have been trying to devise a method of grammar discovery that can proceed from a large sample of grammatically correct sentences, without a teacher. The solution is certainly not unique, and one must make some probabilistic a priori assumptions on the possible forms of the grammar. I would be much interested in any questions or comments that you may have relevant to these papers.

The main reason I have been interested in these problems, is that I have been able to generalize the concept of "phrase structure language". The generalization (which may be called "phrase structure patterns") includes multidimensional patterns, transformational languages, and certain types of mechanical translation. A simple example of the latter is treated in "The Mechanization of Linguistic Learning", section 6. In particular, the arithmetic learning that I had been working on sometime ago, can all be done far more elegantly by phrase structure pattern discovery. In many (though not all) cases, the "grammars" of phrase structure patterns can be discovered using the same methods as are applicable to phrase structure languages.

I think it likely that phrase structure patterns cover an enormous range of pattern types, and that by devising methods for discovering their "grammars" one can perform inductive inferences on this large class of patterns.

Also in this direction, I have devised several practical routines for determining whether an arbitrarily selected sequence of words is an acceptable sentence in a phrase structure language of known grammar. If the sentence is acceptable, the routines will also give all possible methods of dividing the sentence into phrases. There will be more than one such method if the sentence is at all ambiguous.

I will send you papers and/or reports on the generalization of phrase structure languages, and on phrase structure analysis, as soon as they are available.

For the last year I have been working for the Zator Company. We have obtained a (unclassified) contract from the Air Force, in which I work on "Inductive Inference". Other than occasional reports, there are no strings attached. The contract extends for at least another year and probably longer, if things look promising. A particular advantage of the situation is the rather liberal travel funds. I'll be giving a paper "A New Method

for Discovering the Grammars of Phrase Structure Langauges" at the International Con-
ference on Information Processing in Paris June 13-23, 1959. I shall, in all liklihood,
be staying at the Hotel Delavigne, 1 Rue Casimir Delavigne, Paris VI. Could you write
me your phone number or numbers?

Also could you send me reprints of the following papers:

1) "Structure Formelle des Textes et Communication: Deux Etudes", "Word",
   10. 1-27 (1954).

2) "A Note on a Law of Berry and Insistence Stress", Information and Control, Sept. 1957.

3) "Linguistique Statistique Macroscopique I" (1957).

4) "Linguistique Statistique Macroscopique II" (1957).

I will see you in June

Sincerely,

R. J. Solomonoff

RJS:mw

April 10, 1959

Dr Zelig Harris
Department of Linguistics
University of Pennsylvania
Philadelphia, Pa.

Dear Dr. Harris:

Enclosed are reprints of reports on two methods for discovering the grammars of phrase structure languages in suitable training situations, involving a "teacher". I would be very much interested in any questions or comments that you may have on these papers.

I am now working on methods of grammar discovery in which a "teacher" is not necessary -- only a large set of sentences in the language of interest, and some broad probabilistic knowledge of the likelihood of various grammar types is needed.

I have been somewhat successful in generalizing the concept of phrase structure language, so that transformational languages are included in this generalization. In many cases of these generalized phrase structure languages, routines of grammar discovery may be successfully used which are similar to those devised for ordinary phrase structure languages. A paper on these generalized phrase structure languages will be sent to you, as soon as it is available.

These concepts can probably be extended to cover pieces of English text consisting of more than one sentence, and so I am much interested in your work on discourse analysis. Could you send me a reprint of "Discourse Analysis", Language, 28.1-30 (1952), as well as reprints and/or references to any other work of yours that you think is relevant?

Most sincerely,

R. J. Solomonoff

RJS:mw
encl. 124
     125

THE INSTITUTE FOR ADVANCED STUDY

PRINCETON, NEW JERSEY

this becomes a ⊗ PS sentence if we us th. notation $1^{(n)} \wedge$ for integers.

"Is 5 of th. 4 people were killed" grammatical?

*The 3 boys were called Boris, Morris and **Angus**.*

Also, changing this to 6, matter changes choice.

Also : (4/3) of the 5 people were killed, and [th. remaining (people were) / the person who was not was / those who were not, were] injured

I think this type of s is somehow related to something **April 14, 1959**
I heard at Bell Labs. (The "John came into th. room --- " one is closer I think)

Also : Today I bought 5 ~~shou~~ shirts and 3 ties and yesterday, I bought exactly
Dear Mr. Solomonoff, *the same thing, 5 shirts and 3 ties.*

SCHOOL OF MATHEMATICS

Thanks very much for sending me the two papers on phrase structure languages, just received in this morning's mail. I've only had a chance to thumb through them, but they look quite intriguing. I'm delighted to learn that you are working on these problems and that you are now in Cambridge. I hope very much that we can get together and talk about these questions when I return, in August or September.

Just at MIT MT.summer

Your example of a sentence (rather, set of sentences) that go beyond phrase structure in English is very neat. I have ~~has~~ been looking unsuccessfully for such examples for several years. Have you mentioned it in print, or unpublished reports, etc., anywhere, or do you expect to? I would like to use it as an example in the book I am now working on, and am wondering what sort of reference to give. I think it is very much worthwhile looking into the question of how drastically particular forms of description, e.g., phrase structure, fail in natural language -- e.g., ~~how ismith~~ what structural properties of natural language cause one type of description to be more successful than another, etc. Please let me know of any further work of yours along these lines. *— The Eng. s's with "Mail" in them e.g. John came into th. room 2 hours ago, and 2 hours after he came in he will ~~will~~ leave. [or and a ⅟₂ hr. after he came, he left.]*

Also give him Astia and conference dates for th. 2 reports.

See "comment a "on page 2 (?)

As to your two specific questions, I have the answer to only one. I have worked on the question whether the complement of a phrase structure language (with respect to a fixed vocabulary) is a phrase structure language. Unless I am missing something, it's not an easy problem. I can't construct a counter-example, and haven't been able to get anywhere near proving it.. In the course of working on this problem, which I think is crucial, I did get some other results about phrase structure languages which will appear in a paper in Information and Control sometime --~~ha~~haven't gotten proofs yet (i.e., galley proofs). One result has to do with your second question, whether context dependent substitutions expand the domain of phrase structure languages. The answer is, much to my surprise, at least, that it does, considerably. In fact, when I wrote Three Models... I thought I had a proof that there was no extension, and on p.119, last paragraph, I stated that L_3 (the language with sentences ~~thamxesondxhaxhfnsa~~ xx, where x is a string of a's and b's) was beyond the bounds of ~~nxha~~ phrase structure description. In fact, it is only beyond the bounds of context independent phrase structure description, but can be generated by a ~~langxagexwith~~ grammar with contextual dependencies.

Ask for reference ccc or this, so I can refer to it. Also — just how is it done?

I think th. order av of applic. of rules "is impt. only in cont. dep. PSL's. I would like to see how its done.

This is surprising

This suggests → that my (anti-phrase structure) English s's can be done by context-dep. subs.

Section 3.4 of Three Models seems to be OK, though. This has to do with order of ~~xmkmx~~ application of rules, rather than with context dependence. The only mistake I know of in that paper is the one about L_3. I do not, incidentally, have a natural

Better "John arrived ~~when~~ 2 hrs ago, stayed for an hour and then left." v.s. *mention "error" in "Mech. of Lang. learning" that involves L_3 of page 13.*

$1^{(n)}\wedge$ notation may.

I think → This may be not an X
I know how to
do it with an XL!

						101

— Also try th. binary counting lang.:

what about X X X ?   etc.

example of a language beyond the bounds of phrase structure description
when rules with context dependence are allowed, though it is easy to show
that the languages that can be described in this way are a proper subset of
the decidable languages. There's plenty of room for work of all sorts here.

— mention "theorem" about all FSL's have a finite no. of loopless s's.
— but "X X X" lang. has only a finite no. of loopless s's! — so th. theorem isn't terribly useful.

I do not have my reprints with me here in Princeton, except for the Finite
State ~~Models~~ Languages one ~~with~~ which I'll send along. I'll see if I can
dig up some copies of the ~~others~~ when I return to Cambridge.

Incidentally, the ~~fexe~~ reference to that Pattern Conception paper should))
have Miller's name first. I was added as an author only out of courtesy.

— th. decidability may be a consequence of chom's "ordering of
rules" — so there may be a simple way to generate all
s's of length < n, for any n.

Sincerely,

Noam Chomsky
Noam Chomsky

I prefer to regard English as a PSL with a very large no. of rules in its grammar.
Th. fact that Eng. is ~~better~~ more compactly expressed as an XL, means that th. large
no. of PSL rules may be expressed more compactly. Th. PSL as rules are a meta lang.
I want to use a meta meta lang. to express th'l set of PSL rules in more compact form.   finite (or infinite)
[give example of multiplication using meta-meta PSL].   realizable
Th. reason why I want to regard XL's in this way is that it suggests always to mechanize
~~learning~~ of th. lang. — i.e. in 2 steps.

Mention that my method of MT of PSL's may be th. best that is doable without context dep. subst'd

"Constituent Str. analysis"
Does ~~immediate Constituent Analysis~~ "include CDS ?
of "immediate constituent Analysis"
Other than s's of th. type I've given, do you know of any English s's that can be
generated by PSL's with CDS, but not without CDS?

→ To show this: say ~~since~~ we have a cont. dep. Grammar G, and ~~rather~~
we want to see if α is in th. lang. gen. by G: For G₂, (the ~~counter~~ same as
G₁, but without th. context dep. constraints — so G₂ has a lang. that includes that of G₁)
find if α is in L₂; find ~~th~~ all possl. derivations of α in L₂ (there are
a finite no of them) see if any of these derivations conform with G₁. If
any do, α is in L₁, if not, it isn't. This / suggests modifications
analysis
of th. normal PSL parsing schemas, that mite work with CDS. PSL's.

April 28, 1959

Dr. Noam Chomsky
Institute for Advanced Study
Princeton, New Jersey

Now (. May 62) MIT
UN 4 6900 EX 3221

Dear Dr. Chomsky:

Although I had suspected that phrase structure languages with context dependent substitution were significantly different from ones without context dependent substitution, I certainly didn't think that they could go as far as expressing the L_3 of p. 119 of "Three Models. . .". If you can do the language x x, it is very likely that you can express English sentences with "resprctively" in them by a similar device. Could you tell me how you did x x? Can you do x x x, or x x x x, etc.? Also, could you tell me the title of your paper in Information and Control with this result in it -- I may want to refer to it in the paper that I'm working on now.

Is the following language expressible as a phrase structure language with context dependent substitution?

a b c d, a a c e, a b b a b c d d c d, etc.

An arbitrary string of a's and b's followed by the same string in which c is substituted for a, and d substituted for b.

A fairly simple language that is a) decidable, b) is not a context independent phrase structure language is:

1|1 ; 11|10 ; 111|11 ; 1111|100 ; 11111|101 etc.

Each sentence consists of a number of 1's, followed by a vertical bar, followed by the binary notation for the number of 1's to the left of the bar. There are certainly no "cycles" on any of the sentences -- not even context dependent cycles. I'm using "cycles" in the sense used in "A New Method for Discovering the Grammars of Phrase Structure Languages". This test is enough to show it is not a context independent phrase structure language -- but I really don't know any criteria for the context dependent kind.

Another similar language that is not expressible as a context independent phrase structure language is:

1|1 ; 11|1111 ; 111|111111111 etc.

or $1^{(f(n))}$ where $f(n) = n^2$ or many other integral functions of integers

To the right of the bar is a sequence of n ones. To the left of the bar is a sequence of n^2 ones.

With regard to the sentences with "respectively" in them: I haven't referred to this in any reports or papers -- just verbally at the M.I.T. MT seminar. You might refer to it as a "personal communication" if you want to use it in your new book.

I have some other, more complex examples. Practically any sentence with arithmetic in it makes context independent phrase structure description impossible. Some examples:

1) John arrived 2 hours ago, stayed 1 hour and then left.

2) John arrived 2 hours ago, will stay 3 hours and then leave.

The tense of the verbs "to stay" and " to leave" depends on whether 1 is less than 2, and whether 3 is more than 2.

3) 4 of the 5 people were friends and the person who was not was a relative.

4) 3 of the 5 people were friends and the people who were not were relatives.

5) The 3 boys were called Morris, Boris and Angus.

In 5) we can use a context independent phrase structure grammar, if we use the notation 1Λ for 1, 11Λ for 2, 111Λ for 3, etc. This enables us to match up the 1's with the boys in reverse order. We can also do the same thing with 3) and 4): The basic sentence for 3) is:

$\alpha 1\Lambda$ of the $\alpha 111\Lambda$ people were friends and the person who was not was a relative.

with cycle 1, 1 at the α, α positions. However, if we don't use this notation 3), 4) and 5) don't seem to be expressible as phrase structure languages. By using suitable notations for numbers, 1) and 2) can be made into phrase structure languages -- e. g. John arrived (A+ B) hours ago, stayed B hours, and then left. Here A and B may be any number notations one wants to use just so A and B are both > 0. The notation is very artificial, and not integrated with the previous notation. This grammar is all right for generating acceptable sentences, but not very good for deciding if a proposed sentence is acceptable. E. g., if one were given the sentence "John arrived (3+7) hours ago, stayed 5 hours, then left", one would have to convert the 3 7 hours to 5+5 -- an operation that seems external to the grammar.

The above five examples of sentences that are not usually phrase structure languages (unless suitable number notations are used) are best regarded, I think, as an example of sentences in which the grammatical forms are intimately involved with what we ordinarily

regard as "meaning". Essentially, one has to do some computing to determine if a sentence is grammatically correct. Other examples can be found which involve human kinship -- in which the facts expressed in the sentence enable one to determine that two people referred to in the sentence are identical, or that a certain person is female rather than male, or that a certain set of people is plural rather than singular. Other factual data in a sentence may be used to determine sex, and thereby control "grammatical" decisions.

A sentence may be "grammatically indeterminate", in the sense of the gender of a particular object being unknown. In French, "Donnez le moi" is or is not grammatical, depending on whether "le" refers to a masculine or a feminine object.

In general, grammars that do take "meaning" into account, and produce only sentences that are "true" (in the sense of following logically from some stated set of axioms) will be fairly complex -- certainly many of them will describe "undecidable" languages.

This is one of the areas in which "grammaticalness", "meaningfulness", and "logical reasonableness" begin to overlap somewhat, and perhaps it is necessary to make arbitrary boundaries.

Getting back to context dependent substitution -- if one starts with a set of kernel sentences that are formed by a context independent phrase structure language, then transformations of the kind described in "Three Models . . ." Section 5 do not increase the scope of the language very much. About all the transformations could do would be to allow identical repetitions of certain phrase types -- as in the language x x, (L_3 of Section 2 of "Three Models . . ."), or x x x, etc. I don't believe that they could even do sentences with "respectively" in them (though I'm not absolutely sure of this). Since repetitions aren't often (if at all) used as part of grammars of ethnic languages, this seems to be no great loss. If we don't use transformations with repetitions, then the transformations add nothing to the original context independent phrase structure language -- except, perhaps, a more compact means of expressing the grammar rules.

However, since context dependent substitutions do appear to be important ("transformational" rules 15, 16, 17, and $21^1/2$ of pp. 112-113 "Syntactic Structures" are really context dependent substitutions rather than "transformations" in the sense of "Three Models . . ." Section 5), it is reasonable to ask if transformations on a set of kernels from a context dependent phrase structure language does enlarge the grammar at all. Are such transformational languages "decidable"? (If we allow only context dependent substitutions to form the kernel sentences I think the resultant transformational language is decidable -- but I have't worked out a really rigorous proof.) If it is possible to get repetitions with a context dependent phrase structure language, it is quite conceivable that transformations will not expand the domain of context dependent phrase structure languages.

I'm glad to hear that you'll be back in Cambridge in the fall, and look forward to seeing you then.

Sincerely,

R. J. Solomonoff

RJS:mw

April 28, 1959

Dr. Noam Chomsky
Institute for Advanced Study
Princeton, New Jersey

Dear Dr. Chomsky:

Although I had suspected that phrase structure languages with context
dependent substitution were significantly different from ones without context
dependent substitution, I certainly didn't think that they could go as far as
expressing the L_3 of p. 119 of "Three Models. . .". If you can do the lan-
guage x x, it is very likely that you can express English sentences with "res-
prctively" in them by a similar device. Could you tell me how you did x x?
Can you do x x x, or x x x x, etc.? Also, could you tell me the title of
your paper in Information and Control with this result in it -- I may want
to refer to it in the paper that I'm working on now.

Is the following language expressible as a phrase structure language with
context dependent substitution?

a b c d, a a c e, a b b a b c d d c d, etc.

An arbitrary string of a's and b's followed by the same string in which c is
substituted for a, and d substituted for b.

A fairly simple language that is a) decidable, b) is not a context inde-
pendent phrase structure language is:

1 | 1 ; 11 | 10 ; 111 | 11 ; 1111 | 100 ; 11111 | 101 etc.

Each sentence consists of a number of 1's, followed by a vertical bar, followed
by the binary notation for the number of 1's to the left of the bar. There are
certainly no "cycles" on any of the sentences -- not even context dependent
cycles. I'm using "cycles" in the sense used in "A New Method for Discovering
the Grammars of Phrase Structure Languages". This test is enough to show it is
not a context independent phrase structure language -- but I really don't know
any criteria for the context dependent kind.

Another similar langugge that is not expressible as a context independent
phrase structure languages is:

1 | 1 ; 11 | 1111 ; 111 | 111111111 etc.

To the right of the bar is a sequence of n ones. To the left of the bar is a sequence of n^2 ones.

Withrregard to the sentences with "respectively" in them: I haven't referred to this in any reports or papers -- just verbally at the M I T MT seminar. You might refer to it as a "personal communication" if you want to use it in your new book.

I have some other, more complex examples. Practically any sentence with arithmetic in it makes context independent phrase structure description impossible. Some examples:

1) John arrived 2 hours ago, stayed 1 hour and then left.

2) John arrived 2 hours ago, will stay 3 hours and then leave.

The tense of the verbs "to stay" and " to leave" depends on whether 1 is less than 2, or 3 is more than 2.

3) 4 of the 5 people were friends and the person who was not was a relative.

4) 3 of the 5 people were friends and the people who were not were relatives.

5) The 3 boys were called Morris, Boris and Angus.

In 5) we can use a context independent phrase structure grammar, if we use the notation 1Λ for 1, 11Λ for 2, 111Λ for 3, etc. This enables us to match up the 1's with the boys in reverse order. We can also do the same thing with 3) and 4): The basic sentence for 3) is:

$\alpha 1\Lambda$ of the $\alpha 11\Lambda$ people were friends and the person who was not was a relative.

with cycle 1, 1 at the α, α positions. However, if we don't use this notation 3), 4) and 5) don't seem to be expressible as phrase structure languages. By using suitable notations for numbers, 1) and 2) can be made into phrase structure languages -- e. g. John arrived (A B) hours ago, stayed B hours, and then left. Here A and B may be any number notations one wants to use just so A and B are both > 0. The notation is very artificial, and not integrated with the previous notation. This grammar is all right for generating acceptable sentences, but not very good for deciding if a proposed sentence is acceptable. E. g., if one were given the sentence 'John arrived (3 7) hours ago, stayed 5 hours, then left", one would have to convert the 3 7 hours to 5 5 -- an operation that seems external to the grammar.

The above five examples of sentences that are not usually phrase structure languages (unless suitable number notations are used) is best regarded, I think, as an example of sentences in which the grammatical forms are intimately involved with what we ordinarily

regard as "meaning". Essentially, one has to do some computing to determine if a sentence is grammatically correct. Other examples can be found which involve human kinship -- in which the facts expressed in the sentence enable one to determine that two people referred to in the sentence are identical, or that a certain person is female rather than male, or that a certain set of people is plural rather than singular. Other factual data in a sentence may be used to determine sex, and thereby control "grammatical" decisions.

A sentence may be "grammatically indeterminate", in the sense of the gender of a particular object being unknown. In French, "Donnez le moi" is or is not grammatical, depending on whether "le" refers to a masculine or a feminine object.

In general, grammars that do take "meaning" into account, and produce only sentences that are "true" (in the sense of following logically from some stated set of axioms) will be fairly complex -- certainly many of them will describe "undecidable" languages.

This is one of the areas in which "grammaticalness", "meaningfulness", and "logical reasonableness" begin to overlap somewhat, and perhaps it is necessary to make arbitrary boundaries.

Getting back to context dependent substitution -- if one starts with a set of kernel sentences that are formed by a context independent phrase structure language, then transformations of the kind described in "Three Models . . ." Section 5 do not increase the scope of the language very much. About all the transformations could do would be to allow identical repetitions of certain phrase types -- as in the language $x\,x$, (L_3 of Section 2 of "Three Models . . ."), or $x\,x\,x$, etc. I don't believe that they could even do sentences with "respectively" in them (though I'm not absolutely sure of this). Since repetitions aren't often (if at all) used as part of grammars of ethnic languages, this seems to be no great loss. If we don't use transformations with repetitions, then the transformations add nothing to the original context independent phrase structure language -- except, perhaps, a more compact means of expressing the grammar rules.

However, since context dependent substitutions do appear to be important ("transformational" rules 15, 16, 17, and $21\frac{1}{2}$ of pp. 112-113 "Syntactic Structures" are really context dependent substitutions rather than "transformations" in the sense of "Three Models . . ." Section 5), it is reasonable to ask if transformations on a set of kernels from a context dependent phrase structure language does enlarge the grammar at all. Are such transformationsl languages "decidable"? (If we allow only context dependent substitutions to form the kernel sentences I think the resultant transformational language is decidable -- but I have't worked out a really rigorous proof.) If it is possible to get repetitions with a context dependent phrase structure language, it is quite conceivable that transformations will not expand the domain of context dependent phrase structure languages.

I'm glad to hear that you'll be back in Cambridge in the fall, and look forward to seeing you then.

Sincerely,

R. J. Solomonoff

RJS:mw

THE INSTITUTE OF RADIO ENGINEERS

INCORPORATED

PROFESSIONAL GROUP CORRESPONDENCE

**PLEASE ADDRESS
REPLY TO**

Dr. Howard E. Tompkins
The Moore School of E. E.
200 S. 33rd Street
Phila. 4, Pa.

July 7, 1959

Mr. R. J. Solomonoff
Zator Company
140½ Mount Auburn Street
Cambridge 38, Mass.

Dear Mr. Solomonoff:

      Thank you for copies of several re-
ports on methods for discovering the grammars of phrase
structure languages in a suitable training situation.
If Dr. Tompkins has any comments, criticisms or questions
regarding the reports, he will get in touch with you early
in August when he returns from his vacation.

Sincerely,

Helen B. Yonan

Helen B. Yonan
Secretary to Dr. H. Tompkins

hby