CCS Talk Edinburgh

1.
2. **The Computerized Comprehensibility System: An Aid for Comprehensible Writing of Technical Material**
3. **An Example of Applied Cognitive Science**
4. **Problem: Low Comprehensibility of Technical Text**
5. **Sample Early Draft of Training Material**
6. **Sample of a "Mature" Technical Manual**
7. **Solution: A Computer-Based Aid**
8. **Basic Idea - an Artificial Copy Editor**
9. **Previous Writing Aid Systems**
10. **WWB: Example of *prose* program output**
11. **WWB: Example of *style* program output**
12. **CRES: Example output**
13. **Limitations of Previous Systems**
14. **Basic Concept of the Computerized Comprehensibility System (CCS)**
15. **Complex Reading in Isolation**
16. **Complex Reading in Conjunction with a Task can be Overwhelming**
17. **Only Simple Processes should be Needed for Reading during a Task**
18. **Example Comprehensibility Rules from the Psycholinguistics Research Literature**
19. **Why is the System Feasible?**
20. **Would Using the System Help?**
21. **Basics of How CCS Works**
22. **Structure of CCS**
23. **Example of HGSL Specification**
24. **Parsing Output**
25. **Representation of Sentence Content is Syntactically-Tagged Semantic Structure**
26. **Propositional Analysis Output**
27. **Example Production Rules for Comprehensibility Criticisms**
28. **Criticisms Currently Supplied**
29. **Example of CCS Output - 1**

The Computerized Comprehensibility System: An Aid for Comprehensible Writing of Technical Material

David Kieras

University of Michigan

Presented at

Human Communication Research Center

University of Edinburgh

August, 1992

Research Supported by

Office of Naval Research

and

Navy Personnel Research & Development Center

An Example of Applied Cognitive Science

Cognitive Science is overlap between
Cognitive Psychology
- Empirical results, theory

Artificial Intelligence
- Methods, techniques

Use both to help solve a practical problem
A "real" problem with no good existing solutions
A domain that is well-developed in both disciplines

Solve problem in a way that contributes to the science
Make it a test of the underlying theory
Be open to influence from the real world
Ensure that practical products are scientifically valuable

Problem: Low Comprehensibility of Technical Text

Huge quantities of technical text produced
Equipment documentation
Training materials

Material is generally poorly written and hard to use
Even after considerable effort and expense

How can comprehensibility of technical text be improved?
Guidelines, procedures don't seem to help

Editing for comprehensibility is very difficult
Editing is a different skill than reading
Very difficult for human editors to detect problems
• Due to automaticity of reading processes
Having domain knowledge makes problems harder to spot
Patricia Wright's work on high rate of editorial errors

Sample Early Draft of Training Material

INFORMATION

Common U.S. Navy practice in main boiler repair has been the removal of blocks of main generating bank tubes to determine condition of the tubes and need for more extensive sampling and/or renewal. Because of access required for removal and replacement of tubes, the block has generally been 10 tubes wide and completely through the bank. The selection of a specific block has normally been based on ships' tube renewal sheets (if available and of sufficient time length coverage), access to previously plugged tubes, visible waterside pitting as an indication of service life, fireside conditions of tubes and/or fireside deposit accumulations, refractory renewals, and in some cases suspected or visible problem areas in specific boiler designs, if known.

Sample of a "Mature" Technical Manual

2-4-3. PRIMARY POWER MODE. The primary power mode is a cage mode wherein initial application of power to SINS is accomplished. The primary power mode is entered when the PRIMARY POWER (MODE SELECTOR) pushbutton of the NCCP is pressed. During the primary power mode, the platform is course leveled by the pendulous leveling resolvers and course aligned in azimuth by the DEPTH and HEADING data converter monitor drawer. The platform will drive to the indicated heading when a cage mode is selected. The platform temperature alarm circuits are activated, causing the platform temperature alarm lamp to flash until the binnacle temperature is within its operating range. The gyro bottoming circuits and alarms are deactivated. The velocity meter and gyro pump power supply is turned on. The power relays in the navigation console connect 115v 400-Hz 3-phase power to the SINS power supplies and 115v 60-Hz 3-phase power to the SINS blowers. MARDAN memory precision power is also applied in the primary power mode.

Solution: A Computer-Based Aid

Significant piece of the editorial process can now be computerized
System can *detect* comprehensibility problems
  • Alleviate writer's automaticity
System could not correct problems
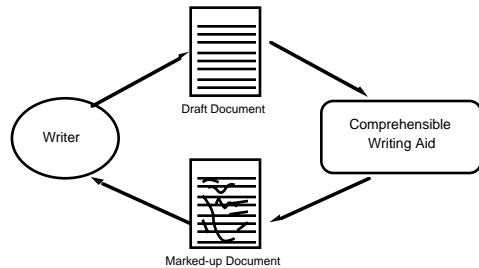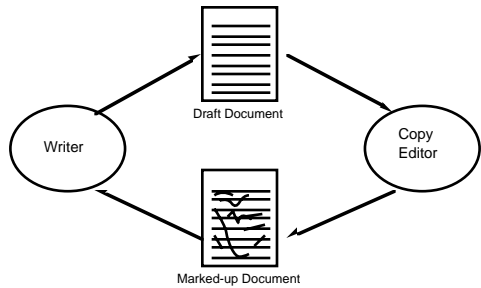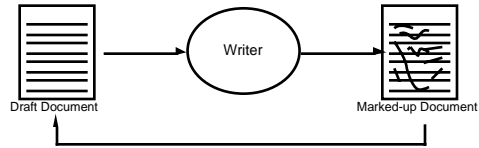 • Still need the writer!

An editorial tool for improving comprehensibility
Input is a document
Output is a critique of comprehensibility

A comprehension research tool and foundation for models of comprehension
More later

Basic Idea - an Artificial Copy Editor

Previous Writing Aid Systems

Writer's Work Bench (WWB)
Based on clever word-classification algorithm
Not integrated
General, statistical feedback, rather than a mark-up

Computerized Readability Editing System (CRES)
Developed by a Navy training materials laboratory
• Peter Kincaid (Kincaid-Flesch readability formula)
Integrated, Navy-specific word lists
Produces a mark-up

Epistle
Developed at IBM from computational linguistics work
Oriented to criticism of grammar in business letters
Good feedback to writer
Not widely available - ran on large mainframe
Newer PC version may be available

WWB: Example of *prose* program output

SENTENCE STRUCTURE
Passives

This text contains a much higher percentage of passive verbs (44%) than is common in good documents of this type (22%). A sentence is in the passive voice when its grammatical subject is the receiver of the action.

PASSIVE: The ball was hit by the boy.

When the doer of the action in a sentence is the subject, the sentence is in the active voice.

ACTIVE: The boy hit the ball.

The passive voice is sometimes needed.

1.   to emphasize the object of the sentence,
2.   to vary the rhythm of the text, or
3.   to avoid naming an unimportant actor.

EXAMPLE: The appropriations were approved.

Although passive sentences are sometimes needed, psychological research has shown that they are harder to comprehend than active sentences. Because of this, you should transform as many of your passives to actives as possible. You can use the style program to find all your sentences with passive verbs in them, by typing the following command when this program is finished.
style -p filename

WWB: Example of *style* program output

readability grades:
(Kincaid) 10.3 (auto) 11.0 (Coleman-Liau) 11.7 (Flesch) 11.2 (54.1)
sentence info:
no. sent 217 no. wds 3971
av sent leng 18.3 av word leng 4.94
no. questions 0 no. imperatives 2
no. nonfunc wds 2400 60.4%   av leng 6.35
short sent (<13) 25% (57) long sent (>28)   10% (21)
longest sent 48 wds at sent 196; shortest sent 4 wds at sent 97
sentence types:
simple 53% (115) complex 26% (56)
compound 12% (25) compound-complex 10% (21)
word usage:
verb types as % of total verbs
tobe 37% (162) aux 11% (48) inf 21% (90)
passives as % of non-inf verbs 19% (66)
types as % of total
prep 11.6% (460) conj 4.0% (159) adv 3.5% (138)
noun 29.9% (1188) adj 18.2% (722) pron 3.2% (128)
nominalizations 2% (78)
sentence beginnings:
subject opener: noun (42) pron (14) pos (1) adj (21) art (80) tot 73%
prep 9% (20) adv 6% (12)
verb 1% (2) sub_conj 9% (19) conj 0% (1)
expletives 2% (5)

CRES: Example output

```
Under no circumstances should any person
reach within or enter the enclosure for
the purpose of servicing or adjusting
the equipment without presence or
<assistance><*AID/HELP*> of another
person capable of <rendering>
<*GIVING/MAKING*> aid.$$032$$
Do not depend upon door switches or
interlocks for protection; but always
shut down motor generators or other
equipment.  Under no circumstances
should any access gate, door, or other
safety interlock switch be removed,
<!! PASSIVE VOICE !!> <short-circuited>,
or <tampered> with in any way by other
than authorized maintenance personnel,
nor should <reliance> be <placed>
<*PUT/*><!! PASSIVE VOICE !!> upon the
interlock swithches for removing
<voltagesfrom> the <equiptment.>$$043$$

    GRADE LEVEL = 14.0
       (Based on DoD Readability Standard)

WORDS NOT ON COMMON WORD LIST

WORD                FREQ
equiptment            1
reliance              1
short-circuited       1
tampered              1
voltagesfrom          1
```

REPLACEMENT WORDS SUGGESTED

LONG SENTENCE (32 WORDS)

PASSIVE VOICE

UNCOMMON WORD

KEYING ERROR, MISSPELLED WORD

READABILITY GRADE LEVEL

Limitations of Previous Systems

Based on "writer's wisdom" or traditions
But what *really* makes material hard to understand?

Emphasize readability measures (or grammaticality) only
Ample evidence that readability measures are inaccurate or misleading

Do not analyze semantic or referential content
Many problems from poor terminology, excessive amounts of content

Are not sensitive to organization of text
Could put sentences in reverse order and get the same results

Basic Concept of the Computerized Comprehensibility System (CCS)

Simulate some simple comprehension processes
"Dumb" Parsing
Simple Reference
Immediate semantics
Basic referential coherence

Criticize if simple processes can't handle the material
Generate a markup of the text

Base criticisms on specific results from
Psycholinguistics
Cognitive psychology of comprehension
Text linguistics

Simple comprehension processes should be adequate in reading-to-do situations
Reading and task processing interact
Doing both can lead to information-processing overload
Readers who are doing a task should be able to read just with their simplest, most automatic reading processes

---

Complex Reading in Isolation

Just Reading - I can handle it.

User

Comprehension Processes

Simple

Complex

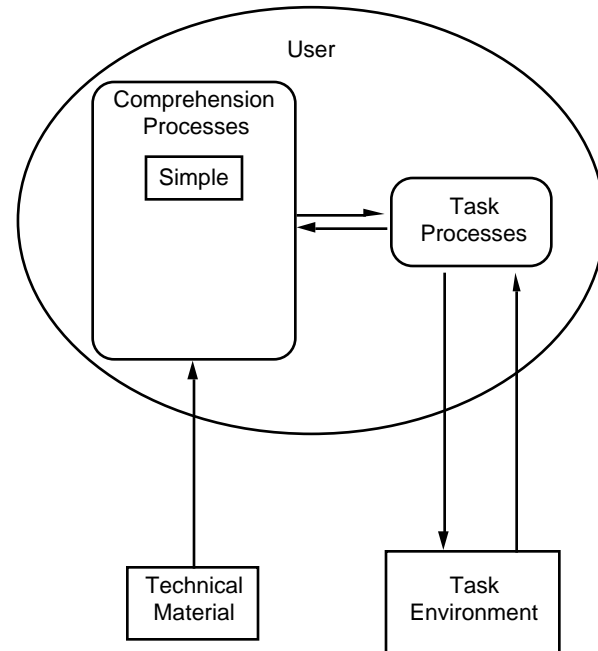*Really* Complex

Technical Material
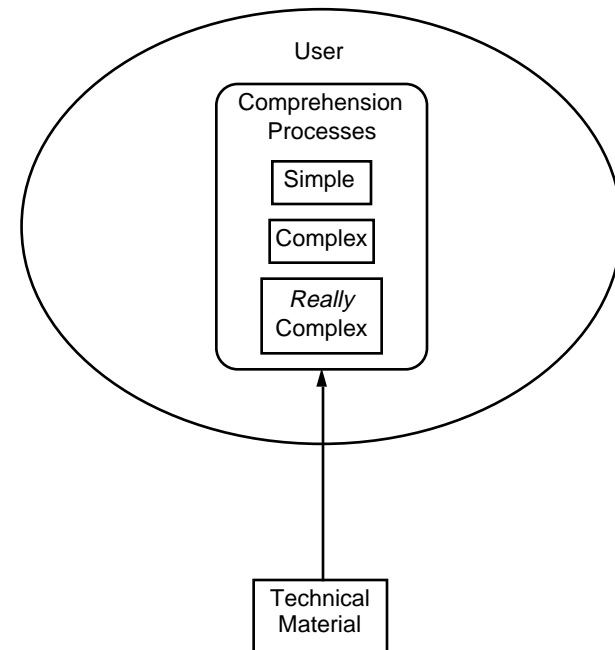
Complex Reading in Conjunction with a Task can be Overwhelming

Ouch!!! I'm trying to get a job done, and I have to wade through this!!!
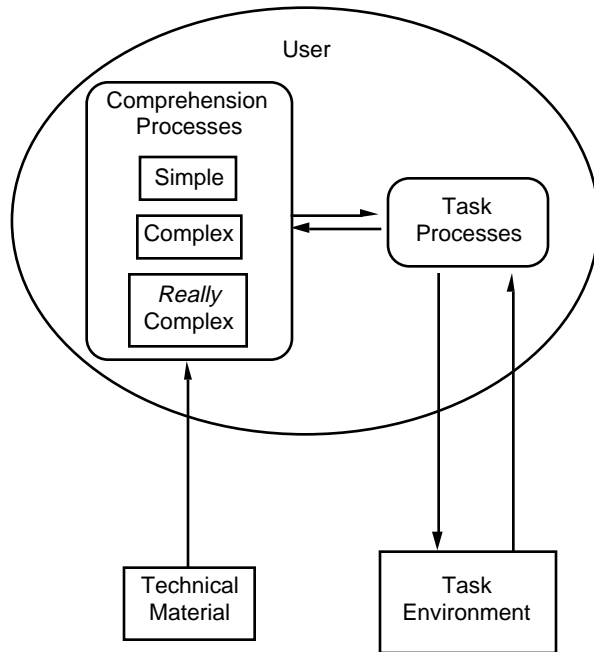
Only Simple Processes should be Needed for Reading during a Task

Much better! All I have to do is follow the instructions!

Example Comprehensibility Rules from the Psycholinguistics Research Literature

Active is better than passive.
(Tannenbaum & Williams, 1968)

A pronoun should refer to the subject of the previous sentence.
(Frederiksen, 1979)

Relative clauses should begin with a relative pronoun.
(Hakes & Foss, 1970)

If the topic of the passage is the logical object,
then passive is better than active.
(Perfetti & Goldman, 1975)

Temporarily changing the subject impedes processing.
(Lesgold, Roth, Curtis, & Riley, 1979)

Refer to an object in the same way as it was previously referred to;
even a synonym slows processing.
(Yekovich, Walker, & Blackman, 1979)

Refer to an object that was either explicitly mentioned previously,
or is strongly implied by the previous text.
(Haviland & Clark, 1974)

Indefinite determiners should be used only on textually new items.
(de Villiers, 1974)

Connective words (e.g., however) improve comprehension.
(Haberlandt & Kennard, 1981)

Why is the System Feasible?

Natural Language Processing is notoriously difficult!

Why is this any easier?

Military technical material should be relatively simple
Military material supposed to meet 9th RGL
Technical material should be clear and direct

System only has to emulate a poor reader, not a good one
- "Artificial Stupidity"

Consequently, don't have to solve the hard problems
Don't have to parse everything
 • If can't parse, neither can reader!
Don't have to integrate everything
 • If hard to do, something is wrong with the text!
Don't have to use much domain knowledge
 • If necessary, then system is impractical!
 • Can't depend on readers having much domain knowledge!

Open question: How smart does the system have to be in order to be useful?
If too stupid, is just an annoyance
 • Like many extant criticism programs
If has to be too smart, not practical to develop
 • Mostly a problem in parser coverage

Would Using the System Help?

Empirical results on improving text are mixed
Document redesign, revision don't always help
When it does, not clear why

The task demands may be critical, difficult to control
Ample experience that manuals can be improved
Kieras study using a mock-up manual

Only way to tell is with a real field test
May be done, since delivered to the Navy
Some work underway with Bruce Britton

Basics of How CCS Works

Standard components
ATN parser
- Compiled from High-level Grammar Specification Language (HGSL)
- Uses chart-parsing logic to increase speed
- Very limited use of registers
- Outputs a parse tree
ACT-style semantic structures
- Built directly from parse tree
  Represent immediate propositional content
- Syntactically-tagged
  Production rules can recognize combinations of syntax and semantic features
Production rules
- Generate criticisms
- Perform inference and integration functions
Simple reference done by special module
- Can resolve a reference to a previously mentioned object by either match on surface noun phrase, or match on any subset of propositions in the noun phrase
- Criticize if referent can't be identified by simple reference resolution process

Comprehension based on given/new contract model
Identify given referents in each sentence
Add new content to representations of given referents

Structure of CCS
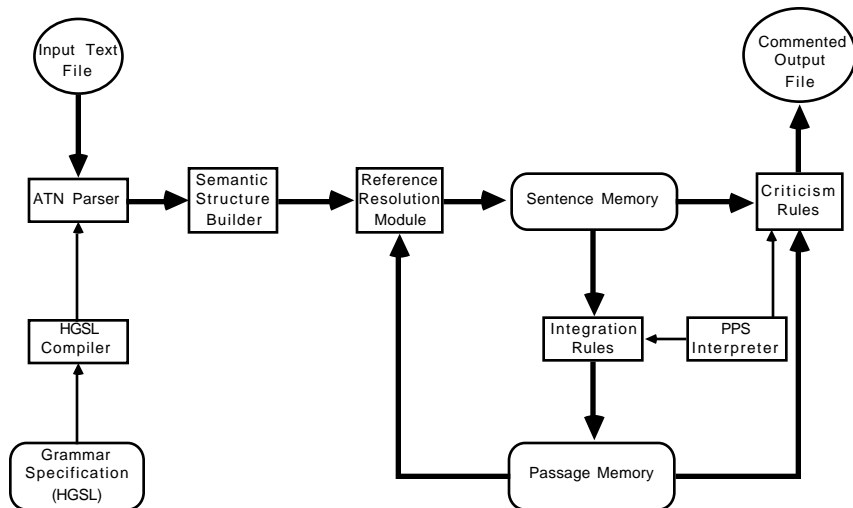
Example of HGSL Specification

```
(NET-DEF $DECLARATIVE-STATEMENT :SCOPED-REGISTERS
(?PERSON-NUMBER)
 (- ($INITIAL-VERBMOD - (!\,)) $NP $CLSPRED))


(NET-DEF $COPPHR
 ($COPSEQ - (!TO $COPSEQ) $COPCOMP) )

(NET-DEF $COPSEQ
 (((- (ADV) MODAL - (ADV / NEG)
  ((HAVE-INF - (ADV / NEG) BE-PSP - (- (ADV / NEG) BE-PRP))
  / (BE-INF - (- (ADV / NEG) BE-PRP))))
  / (- (ADV) HAVE-FORM - (ADV / NEG) BE-PSP - (BE-PRP))
  / (- (ADV / NEG) BE-FORM - (- (ADV / NEG) BE-PRP)))
     - (ADV / NEG)))

(NET-DEF $COPCOMP-SIMPLE :SCOPED-REGISTERS (?PERSON-NUMBER)

 (($ADJ / $PREPPHR / $NP / (!THAT $STATEMENT)) - (ADV)))
```

Parsing Output

If the PF Indicator does not flash, then notice that there is a malfunction.

Parsing Transitions: 1375

```
($STATEMENT
 ($STATEMENT-COMPLEX
  ($PROP-CONJ-A-B-STATEMENT (PROP-CONJ-A-B IF)
              ($STATEMENT-SIMPLE
               ($DECLARATIVE-STATEMENT
                ($NP
                 ($NP-SIMPLE (DEFDET THE) (NOUN PF)
                      (NOUN INDICATOR)))
                ($CLSPRED
                 ($CLSPRED-SIMPLE
                  ($VERBPHR
                   ($VERBPHR-SIMPLE
                    ($VERBPHR-ACTIVE
                     ($VERBSEQ-ACTIVE (MODAL DOES)
                      ($VERBMOD (NEG NOT))
                      ($VERB-INF (VERB-INF FLASH)))))))))
              |,|
              ($STATEMENT-SIMPLE
               ($IMPERATIVE-STATEMENT
                ($INITIAL-VERBMOD (ADV THEN))
                ($IMPERATIVE-VERBPHR
                 ($VERB-INF (VERB-INF NOTICE))
                 ($VCOMP-ACTIVE
                  ($NP
                   ($NP-STATEMENT (RELPRN THAT)
                        ($THERE-SUBJECT-STATEMENT
                         THERE ($COPSEQ (BE-FORM IS))
                              ($NP
                               ($NP-SIMPLE
                                (NDEFDET
                                A)
                                (NOUN
                                MALFUNCTION)))))))))))
  \.)
```

Representation of Sentence Content is Syntactically-Tagged Semantic Structure

Concept: Preserve relevant information about the original surface form of the sentence

Semantic (propositional) representation of sentence content is tagged with correspond surface structure objects
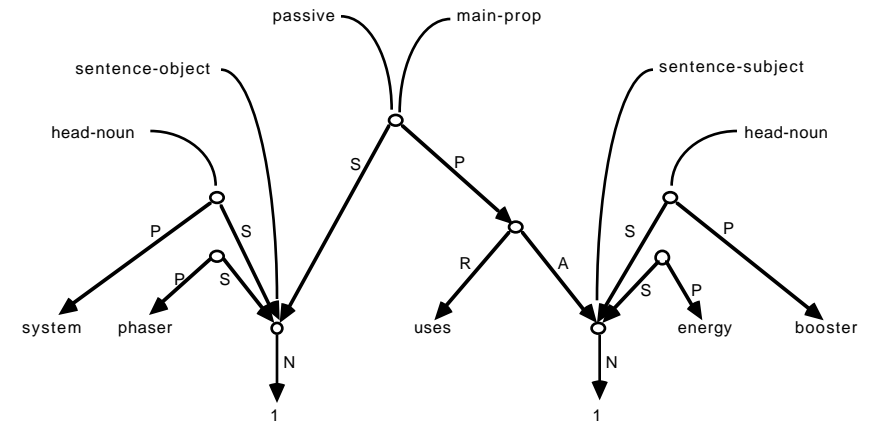E.g., passive sentence is normalized to active representation, but original form of main proposition noted

Useful for text integration
E.g., how surface subject is related to passage topic

Allows easy criticism of both propositional content and syntactic form
E.g., passive sentence surface subject is not the topic

The energy booster is used by the phaser system.

Propositional Analysis Output

If the PF Indicator does not flash, then notice that there is a malfunction.

ACT Propositions:
((S PROP1 REF1) (P PROP1 PF) (S PROP2 REF1) (P PROP2 INDICATOR)
(S PROP3 REF1) (P PROP3 PRED1) (R PRED1 FLASH) (A PRED1
UNSPECIFIED) (S PROP4 PROP3) (P PROP4 FALSE)
(S PROP7 THE_READER) (P PROP7 PRED2) (R PRED2 NOTICE)(A PRED2
PROP6) (S PROP6 REF2) (P PROP6 EXISTS)
(S PROP5 REF2)(P PROP5 MALFUNCTION)
(S PROP8 PRED2) (P PROP8 THEN)
(S PROP9 PROP7)(P PROP9 PRED3) (R PRED3 IF) (A PRED3 PROP3))

ALL Propositions:
((TAG PROP7 SENTENCE-MAIN)
(TAG REF1 COMPLETELY-NEW) (TAG REF1 NEW-REFERENT)(TAG REF1
NP-REFERENT)
(TAG PROP1 NP-PROP) (S PROP1 REF1) (P PROP1 PF)
(TAG PROP2 NP-PROP) (S PROP2 REF1) (P PROP2 INDICATOR) (TAG
PROP2 HEAD-NOUN)
(TAG REF1 DEFINITE) (NUMBER REF1 SINGULAR) (GENDER REF1 N)
(SRF REF1 (PF INDICATOR)) (HNF REF1 INDICATOR) (NP-PROP-NUMBER
REF1 2)
(TAG REF1 STATEMENT-SUBJECT)
(S PROP3 REF1) (P PROP3 PRED1) (R PRED1 FLASH)(A PRED1
UNSPECIFIED)
(S PROP4 PROP3) (P PROP4 FALSE)
(TAG PROP3 STATEMENT-MAIN) (TAG THE_READER
STATEMENT-SUBJECT)
(S PROP7 THE_READER) (P PROP7 PRED2) (R PRED2 NOTICE) (A PRED2
PROP6)
(TAG REF2 STATEMENT-SUBJECT) (S PROP6 REF2) (P PROP6 EXISTS)
(TAG REF2 NEW-INDEFINITE) (TAG REF2 NEW-REFERENT) (TAG REF2
NP-REFERENT)
(TAG PROP5 NP-PROP) (S PROP5 REF2) (P PROP5 MALFUNCTION) (TAG
PROP5 HEAD-NOUN)
(TAG REF2 INDEFINITE) (NUMBER REF2 SINGULAR) (GENDER REF2 N)
(SRF REF2 (MALFUNCTION)) (HNF REF2 MALFUNCTION)
(NP-PROP-NUMBER REF2 1)
(S PROP8 PRED2) (P PROP8 THEN)
(TAG PROP7 IMPERATIVE)(TAG PROP7 STATEMENT-MAIN)
(S PROP9 PROP7) (P PROP9 PRED3) (R PRED3 IF) (A PRED3 PROP3)
(TAG PROP3 SUBORDINATE-CLAUSE)
(TAG PROP7 SUPERORDINATE-CLAUSE) (TAG SENTENCE-PROP-NUMBER

Example Production Rules for Comprehensibility Criticisms

```
(BadPassive
IF
    ((CONTROL GOAL DO CRITICISM)
    (SM TAG ?MAIN-PROP PASSIVE)
    (SM TAG ?MAIN-PROP STATEMENT-MAIN)
    (SM TAG ?STATEMENT-SUBJECT STATEMENT-SUBJECT)
    (NOT (PM TAG ?STATEMENT-SUBJECT DISCOURSE-TOPIC)))
THEN
    ((PRINT-MSG "BAD PASSIVE")))


(GoodPassive
IF
    ((CONTROL GOAL DO CRITICISM)
    (SM TAG ?MAIN-PROP PASSIVE)
    (SM TAG ?MAIN-PROP STATEMENT-MAIN)
    (SM TAG ?STATEMENT-SUBJECT STATEMENT-SUBJECT)
    (PM TAG ?STATEMENT-SUBJECT DISCOURSE-TOPIC))
THEN
    ((PRINT-MSG "GOOD PASSIVE")))
```

Criticisms Currently Supplied

Terminology
    Failure to use simple terminology
    Inconsistent terminology
    Ambiguous terminology

Grammatical difficulty
    No systematic check of grammatical correctness
    If can be parsed, grammatical difficulty is ok
    If fail to parse
- attempt to find constituents and proceed
- warn that grammar might be too difficult

Amount of Content
    Noun phrases too big
    Too much new content within a sentence

Sentence structure
    Telegraphic style - sentence fragments (except in headings)
    Lack of which/that in relative clauses
    Object relatives (subject relatives preferred)
    Self-embedding constructions
    Negation, especially in top-level clause

Coherence
    Failure to follow given/new contract for each sentence
    New referent mentioned in slot usually reserved for old
    Failure to use simple referential coherence forms
    Referent is non-trivial to determine
    Passive voice misused or not used when appropriate
    Pronoun references pointed out

Topic Structure
    Current topic tracked
    Final structure presented

Example of CCS Output - 1

The oil is stored in the lube oil sump.

Sentence No. 11

QUESTIONABLE-NEW-REFERENT
Check: Can your reader tell what you are referring to:
REF27 (LUBE OIL SUMP)

AMBIGUOUS-REFERENT
Check and rephrase if incorrect:
Assuming that OIL is REF13 OIL,
  ... other possibilities: (REF9) (OIL)

NO-SIMPLE-FORM
Try introducing these with the simpler wording in this sentence:
REF13 (OIL)

INAPPROPRIATE-PASSIVE
Using the passive voice in the clause about REF13 OIL
is inappropriate because it is not about a current topic.
This can be very hard to understand - try to rephrase into the active voice.

TOPIC-CHANGE
You may need to rewrite to avoid incoherence.
~~~~~~

Example of CCS Output - 2

The idle strainer must not be put into service when the cap is not properly secured.

Sentence No. 23

NEGATED-MAIN-CLAUSE
Try to rewrite into positive form.

MULTIPLE-NEGATION
This sentence has multiple negatives, which can be very hard to understand.
You should rewrite it into positive form.

SUBJECT-NOT-KNOWN
Clause subject REF59 CAP should be previously mentioned.

QUESTIONABLE-NEW-REFERENT
Check: Can your reader tell what you are referring to:
REF59 CAP

INAPPROPRIATE-PASSIVE
Rephrase passive clause about REF50 (IDLE STRAINER).

INAPPROPRIATE-PASSIVE
Rephrase passive clause about REF59 CAP.

TOPIC-CHANGE
You may need to rewrite to avoid incoherence.
~~~~~~

Simulated Usage Test

A simple trial
not realistic, but encouraging

Writer
a talented undergraduate engineering student

Original Material
A sample of actual draft training material
119 lines long, about 1.5 pages
Readability Score = 13.06

Revision Process
Obtained system outputs, made revisions in response
Total of 5 revisions

Equipment
Apollo DN4000 (4 MIPS, 12 MB)
Franz Allegro Lisp, all code compiled

Sample of Original Training Material

Common U.S. Navy practice in main boiler repair has been the removal of blocks of main generating bank tubes to determine condition of the tubes and need for more extensive sampling and/or renewal. Because of access required for removal and replacement of tubes, the block has generally been 10 tubes wide and completely through the bank. The selection of a specific block has normally been based on ships' tube renewal sheets (if available and of sufficient time length coverage), access to previously plugged tubes, visible waterside pitting as an indication of service life, fireside conditions of tubes and/or fireside deposit accumulations, refractory renewals, and in some cases suspected or visible problem areas in specific boiler designs, if known.

Sample of Revised Training Material

Common Navy practice in main boiler repair has been the removal of a block of boiler tubes to check their condition. Based on this removed block, it can be decided whether the tubes need to be replaced or further checked. The removed block has usually been 10 tubes wide and through the entire bank of tubes. The block has been this large because of the access required for the removal and the replacement of the tubes. Normally, the selection of a specific block has been based on a number of factors. These factors include the ship's tube renewal sheets, access to tubes that were plugged before, visible damage from wear, fireside conditions of tubes, fireside deposit buildup, refractory renewals, and sometimes problem areas in specific boiler designs.

Original vs. Revised Sample

Original

Size
- 119 lines long, about 1.5 pages
- 51 sentences found by system

System Performance
- 20/51 sentences could be parsed (40%)
- 92 min. cpu time

System Output
- About 168 referents defined
- About 350 propositions defined
- About 230 comments (multiple item comments counted as one)
- 31 pages of output

Results
- Due to many grammatically unanalyzable sentences, relatively few comments made
- Readability Score = 13.06

Revised

Size
- 108 lines long, 1.5 pages approximately
- 70 sentences found by system

System Performance
- 44/70 could be parsed (63%)
- 58 min. cpu time

System Output
- About 215 referents defined
- About 495 propositions defined
- About 489 comments (multiple item comments counted as one)
- 40 pages of output

Results
- Seems much more clear
- Readability Score = 8.41

Earlier Versions of the System

Demonstration System
Constructed from earlier simulation model components
Very limited parser
Could generate some sophisticated criticisms

Parsing Workbench System
Extending ATN grammar is the main problem
Graphical interface in Interlisp-D environment
Allow ATNs to be displayed, observed, modified easily
Extended grammar using real samples from NPRDC
But grammar development was unexpectedly difficult

Grammar Development System
Even with Parsing Workbench, large ATN was too difficult
A High-level Grammar Specification Language (HGSL)
- John Mayer
A large grammar built using NPRDC and other samples
Convergence of coverage in this domain
Conclusion: Have a practically useful grammar for parsing

Current State of the System

Supported by NPRDC program on Authoring Instruction Materials
Fielding advanced document preparation software for Navy Training Writers

Portable implementation
Run on UNIX workstations
COMMON LISP implementation

Current trial version
Provide a limited number of comprehensibility criticisms
Accepts text files with interleaved document formatting commands
Works well enough for actual writers to try out
Parsing still needs work
- Many parsing failures are trivial, could be corrected
- System could comment on partially parsed input better

An evaluation study underway with Bruce Britton
Comparison of fixed original and revised version of a complex passage
- Original passage produces identifiable and serious failures of comprehension
Original was modified so that all sentences would parse adequately
- Minimal other changes
Revision was changed to respond to CCS criticisms
Recall data collected, currently being analyzed

Experimentation with semantics-based reference resolution
Making CCS "smarter" about some kinds of reference
- When a referent is implied by another

Example of CCS "Stupidity" about Implied Referents

The F-16 aircraft is a high-performance fighter.
~~~~~~
The wings have a swept-back configuration.

The main proposition of this sentence is PROP9:
- REF3 WINGS has relation HAVE
  to REF4 (SWEPT-BACK CONFIGURATION).

NO-KNOWN-REFERENTS
This sentence does not appear to refer to anything previously mentioned,
and so readers may not understand how it relates to the rest of the material.
Be sure that the sentence directly and clearly refers to a previous item.

QUESTIONABLE-NEW-REFERENT
These items were referred to as if the reader already knows about them,
but they could not be matched with something previously introduced:
REF3 WINGS
Check: Can your reader easily figure out what you are referring to?
...

Experimentation on Semantics-Based Reference Resolution of Implied Referents

Common CCS criticism: Questionable New Referent
Writer refers to an apparently new object with a definite noun phrase
Form implies that reader should know it
But can't be matched against a previously mentioned item

Often, the new referent seems obviously implied by previously mentioned one

Is domain knowledge really required, or is more general semantic knowledge adequate?
Everybody knows that airplanes have wings!
- True domain knowledge: Airplane engine throttles have "quadrants"

Many reference resolution inferences can be done using simple semantic relations
- subset/superset
- part-of/has-part
- member-of/has-member

Miller's WordNet Database

Miller's project - a sort of on-line super-thesaurus
Also sponsored by ONR and NPRDC

A large semantic lexicon
E.g. about 30K noun concepts
Includes, nouns, adjectives, verbs

Words grouped into synonym sets
Synonym set corresponds to a concept or type node in a semantic net

Simple semantic relations between the synonym sets
subset/superset, part-whole, member, opposite, similar, etc.

Database is publicly available

Sample of WordNet Database (reformatted)

(^N-AIRCRAFT (AIRCRAFT) <M ^N-FLEET >P ^N-SKELETON3 >> ^N-HELICOPTER
>> ^N-GLIDER >P ^N-FUEL_GAUGE >> ^N-DRONE3 >P ^N-CABIN2 >P ^N-COCKPIT2
>> ^N-LIGHTER-THAN-AIR_CRAFT >> ^N-AIRPLANE >P ^N-AIRCRAFT_ENGINE
<< ^N-VEHICLE)

(^N-AIRCRAFT_CARRIER (AIRCRAFT_CARRIER CARRIER FLATTOP
ATTACK_AIRCRAFT_CARRIER)
>P ^N-FLIGHT_DECK >P ^N-ARRESTER << ^N-WARSHIP)

(^N-AIRCRAFT_ENGINE (AIRCRAFT_ENGINE) <P ^N-AIRCRAFT << ^N-ENGINE2)

(^N-AIRFOIL (AIRFOIL AEROFOIL) >> ^N-WING6 >> ^N-VERTICAL_TAIL
>> ^N-STABILIZER >> ^N-RUDDER >> ^N-ROTOR_BLADE >> ^N-FLAP5
>> ^N-ELEVATOR >> ^N-HORIZONTAL_STABILIZER
>> ^N-AILERON << ^N-DEVICE2)

(^N-AIRLINE2 (AIRLINE) << ^N-TRANSPORTATION_SYSTEM)

(^N-AIRLINE (AIRLINE) << ^N-HOSE3)

(^N-AIRLINER (AIRLINER) >P ^N-SEAT5 >P ^N-GALLEY << ^N-AIRPLANE)

(^N-AIRLOCK (AIRLOCK AIR_LOCK) << ^N-CHAMBER2)

(^N-AIR_PASSAGE (AIR_PASSAGE AIR_DUCT AIRWAY) >P ^N-VENT2 >> ^N-UPCAST >>
^N-SNORKEL2 >> ^N-DOWNCAST << ^N-DUCT2)

(^N-AIRPLANE (AIRPLANE AEROPLANE PLANE) >P ^N-WING6 >P ^N-WINDSHIELD >>
^N-TURBOJET >> ^N-SEAPLANE >P ^N-RADOME >> ^N-PROPELLER_PLANE
>P ^N-POD2 >> ^N-MONOPLANE >P ^N-LANDING_GEAR >> ^N-JET3 >P ^N-FUSELAGE >>
^N-FIGHTER4 >P ^N-ESCAPE_HATCH >P ^N-COWL >> ^N-BOMBER
>> ^N-BIPLANE >> ^N-AMPHIBIAN >> ^N-AIRLINER << ^N-AIRCRAFT)

(^N-AIRPLANE_PROPELLER (AIRPLANE_PROPELLER AIRSCREW PROP)
<P ^N-PROPELLER_PLANE << ^N-PROPELLER)

(^N-AIRFIELD (AIRFIELD LANDING_FIELD) >P ^N-TAXIWAY >P ^N-RUNWAY >>
^N-AUXILIARY_AIRFIELD >P ^N-APRON2 >> ^N-AIRSTRIP >> ^N-AIRPORT
<P ^N-TRANSPORTATION_SYSTEM << ^N-FACILITY5)

Augmented Version of CCS with Implied Reference Resolution

Experimental approach:
No guidance or constraint by context or noun phrase
modifiers; search for relations between head nouns only.
Of course, many inappropriate connections will be found.
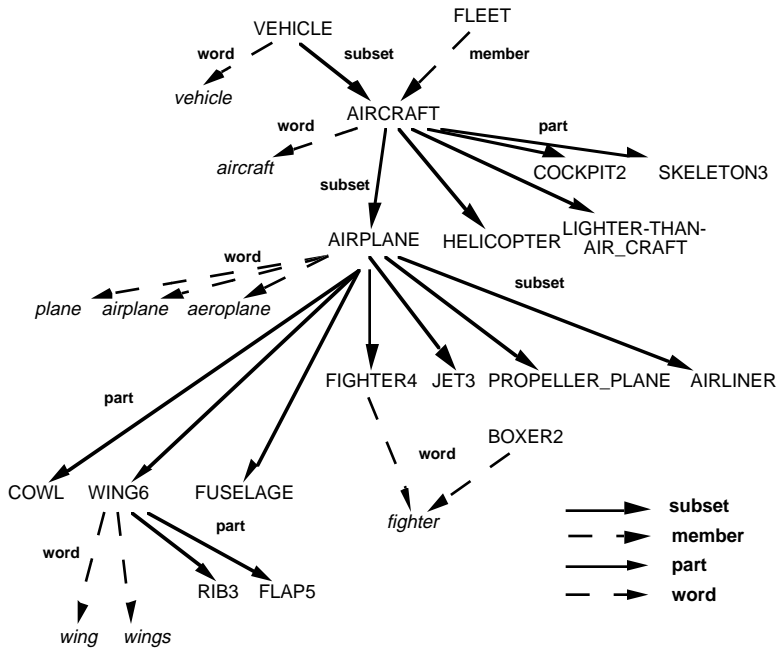Will correct connections be found?

Use reduced version of WordNet database based on CCS lexicon
Complete database is too large for convenient testing

Processing:
If simple reference resolution fails on a definite noun phrase,
apply semantics-based resolution.
Find relations between questionable new referent and
previous referents.
If appropriate relation present, the new referent is implied by
previous referent; do not criticize.

Relevant Semantic Network for Aircraft Example

Example of Semantics-Based Reference Resolution about Aircraft

The F-1 is an aircraft.
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
The wing is long.
IMPLIED-PART
Assuming that these newly introduced items are part of previously mentioned items:
New REF3 WING (concept: ^N-WING6) is part of REF2 AIRCRAFT (concept: ^N-AIRCRAFT)
Check: Is this correct?
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
The flaps are big.
IMPLIED-PART
Check: Is this correct:
New REF5 FLAPS (concept: ^N-FLAP5) is part of REF3 WING (concept: ^N-WING6)
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
The airplane is expensive.
IMPLIED-SUBSET
Assuming that these newly introduced items refer to previously mentioned items:
New REF6 AIRPLANE (concept: ^N-AIRPLANE) is included by, and refers to, REF2 AIRCRAFT
(concept: ^N-AIRCRAFT)
Check: Is this correct?
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
The fighter is essential.
IMPLIED-SUBSET
Check: Is this correct:
New REF7 FIGHTER (concept: ^N-FIGHTER4) is included by, and refers to, REF6 AIRPLANE
(concept: ^N-AIRPLANE)
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
The vehicle has wheels.
IMPLIED-SUPERSET
Assuming that these newly introduced items refer to previously mentioned items:
New REF8 VEHICLE (concept: ^N-VEHICLE) includes and refers to REF2 AIRCRAFT
(concept: ^N-AIRCRAFT)
Check: Is this correct?

Trial of Augmented CCS

Tested on three long passages
Britton's history of the air war in Vietnam
T-38 Flight Control from T-38 Flight Manual
Ejectors from Machinist's Mate 2 & 3 Rate Training Manual

How many questionable new referents were resolved through connections in the semantics?
Definite noun phrases with no antecedent referent that can be determined using simple reference resolution

Results:
Many connections found
Most were incorrect
- Expected from experimental approach used
- Most connections were spurious, due to lack of constraint
- E.g., *make-up feed valve* in context of steam engine condenser system is matched to British term *valve* for vacuum tube as related to *condenser* as an electronic component (obsolete term)

Most correct ones could be found by simpler lexical mechanisms
- E.g., direct synonym => concept representation

In spite of lack of constraint, very few correct connections found in technical passages
- E.g., one of few cases: *Airplane* => *wing flaps*
- But not *switch* => *position*

Conclusions from Semantics-Based Reference Resolution Experiment

Technically feasible to incorporate simple semantic information and inference rules for using it
Encourages further work in comprehension modeling

WordNet database, although very broad, is not consistently deep enough to help in technical passages
E.g. for T-38 passage:
- *airplane* => *flaps* in database
- *airplane* => *control stick* not in database

Other passage types may require much more complex inferences than those based on simple semantic relations
E.g. *war* => *enemy* in Britton's historical passage
Seems to require action-participant relation

Original no-knowledge approach is impressively good
To do better, considerable domain-specific knowledge would be required

How can we build useful semantic databases for real domains?

Scientific Contribution: Testing the Theory

Basic concept that sentence and text syntax are important
Can the system adequately characterize important properties of comprehensibility even though it is mostly syntactically-driven?
Division of labor - do as much as possible with syntax, save knowledge-based inference for the hard problems - a good idea?

Value of standard mechanisms and theoretical concepts
Examples
- Production rules for comprehension processing
- Topic signaling role of topic sentence, surface subject position

Can these do an adequate job?
What do we know from where they fail?

Scientific Contribution: Influence from the Real World

Nature of real reading and real materials
Predominance of descriptive and procedural text
- rather than heavily schematic stories

Massive knowledge-based inference shouldn't be the basic everyday process
Task-driven reading is the key, but it is very poorly understood

Things we need scientifically
Only a fraction of laboratory research addresses real-world materials and problems
- E.g. many studies of non-existent multiply-embedded sentences
- Few studies of terminology, complex reference

Still lack a good characterization of sentence complexity

Scientific Contribution:
Research Value of Product

Software is usable in research

Variations on system can provide analytic tools
Analyze structure and content of text rigorously and automatically

Foundation for advanced models of complex forms of comprehension
Has a reasonable parser, so input can be actual experimental materials
Has slots for additional processes such as complex inference
A portable, well-disciplined software environment

System is a Foundation for Models of Text Comprehension

What it has already
Facilities for parsing input sentences in text
Basic reference resolution of text integration
Simple topic tracking mechanism
Provision for inference rules
- E.g. used in semantics-based reference resolution

What is does not have
Significant knowledge base
Substantial knowledge-driven comprehension mechanisms
- Can beyond experimental simple semantic relations
Basic structure of system could incorporate knowledge-based mechanisms after parsing stage
- No feasible way to use knowledge during parsing with this system

Examples of Possible Research Applications

Analyze text for propositional content
A standard, labor-intensive process in comprehension research
CCS uses Anderson ACT style - close relationship to Kintsch propositions

Classify sentence form or complexity
E.g. How often does center embedding occur?

Analyze referential structure of text
Distribution of types of reference
Where inference is required
Distance between antecedent and reference

Analyze knowledge requirements of text
Where is semantic knowledge critical?
Where is domain knowledge needed?

Re-implement Kintsch & van Dijk model, but take actual text as input
Could develop materials that system can parse before using in an experiment

Front-end for model of learning procedures from text (Susan Bovair)
Above examples

Environment for exploring inference and knowledge in comprehension
"Slot" for production rules that could compare long-term memory contents with sentence and passage content
- E.g. WordNet database mechanisms