

Inhaltsverzeichnis

1. Einleitung	11
2. Automatische Generierung von Texten	15
2.1. Zum Begriff " <u>semantische Repräsentation</u> "	16
2.1.1. Motivation	16
2.1.2. Exkurs: Linguistik und Künstliche Intelligenz (KI)	18
2.1.3. Frames als semantische Repräsentation	19
2.2. Generierung als Auswahlproblem	20
2.3. Andere Arbeiten über Generierung	21
2.3.1. Exkurs: Schanks Theorie der konzeptuellen Dependenz	22
2.3.2. Goldmans BABEL	23
2.3.3. Incremental Procedural Grammar (IPG)	24
2.3.4. Generierung von Japanisch aus CD	24
2.3.5. KAFKA, der Generator des XCALIBUR-Projekts	25
2.3.6. Jacobs' PHRED	25
2.3.7. Generierung mit "direkter Methode" in XPLAIN	26
2.3.8. Generierung in HAM-ANS	27
2.3.9. Appelts KAMP und TELEGRAM	28
2.3.10. McKeowns TEXT	28
2.3.11. Gabriels YH: Wie schreibt man sorgfältig?	29
2.3.12. Generieren mit PROLOG	30
2.3.13. McDonalds MUMBLE	32
2.4. Verarbeitung von Zeitungstexten in der KI-Forschung	34
3. Generierung in einem Übersetzungsprojekt	37
3.1. Japanisch/deutsche Titelübersetzung in SEMSYN	37
3.2. Der Gesamtaufbau des SEMSYN-Systems	40
3.3. Die Transformation des semantischen Netzes	42
3.3.1. Das semantische Netz als Datenstruktur	42
3.3.2. Die Interpretation des semantischen Netzes	43
3.3.3. Vom Netz zur Frame-Beschreibung	45
3.3.4. Sonderfälle	48
3.4. Generierung aus Frame-Beschreibungen	48
3.4.1. Die Notation der Frame-Sprache	49
3.4.2. Konzeptuelle Charakterisierung der Frame-Sprache	49
3.4.3. Entscheidungen bei der Generierung	50
3.4.4. Die globale Kontrollstruktur	51
3.4.5. Der Formalismus zur linguistischen Beschreibung	51
3.4.6. Die Sprache der Titel	53
3.5. SUTRA-S - die morpho/syntaktische Komponente des Generators	55

4. Zur Implementierung des Generator-Kerns	57
4.1. Vorgaben für die Implementierung des Generators	57
4.2. Exkurs: Objektorientierte Programmierung	59
4.2.1. Sichtweisen der objektorientierten Programmierung	59
4.2.1.1. Abstrakter Datentyp + "procedural attachment"	59
4.2.1.2. Die Sicht der Programmierung	60
4.2.1.3. Objektorientierte Wissensrepräsentation	60
4.2.2. Objektorientierte Programmierung im FLAVOR-System	61
4.3. Die objektorientierte Implementation des Generator-Kerns	63
4.3.1. Die Grundhierarchie	63
4.3.2. Die Instantiierung von Frame-Beschreibungen	64
4.3.3. Ein Beispiel	65
4.3.4. Ergänzungen zur Instantiierung	68
4.3.5. Die Nachricht :TRY-TO-REALIZE	69
4.3.6. Zur Lexikalisierung semantischer Symbole	69
4.4. Die Klasse CASE-SCHEMA	71
4.4.1. Instanzen ohne fokussierte Rolle	71
4.4.1.1. Die Realisierung als Nominalgruppe (:NG)	71
4.4.1.2. Die .MANNER-Rolle bei der NG-Realisierung	72
4.4.1.3. Einheitliche Behandlung sonstiger Rollen	73
4.4.1.4. Die Realisierung eines Kasusrahmens als Satz (:CLAUSE)	74
4.4.2. Realisierung von Instanzen mit fokussierter Rolle	76
4.4.2.1. Fokussierte Rollen bei der Realisierung als :CLAUSE	76
4.4.2.2. Fokussierte Rollen bei der Realisierung als Nominalgruppe	77
4.4.2.3. Fokussierte :OBJECT-Rolle	77
4.4.2.4. Fokussierte :AGENT-Rolle	78
4.4.2.5. Andere fokussierte Rollen	78
4.4.2.6. Die Stilpräferenz *PREFER-RELATIVE-CLAUSES*	78
4.4.2.7. Die Methode :TRY-NG-WITH-NOMINALIZED-ACT-REALIZATION	79
4.4.2.8. Weitere Rollen	81
4.4.3. Die Behandlung der Rollen in CASE-SCHEMA	82
4.4.3.1. Die Rolle :AGENT	83
4.4.3.2. Die Rolle :OBJECT	83
4.4.3.3. Die Rolle :MANNER	84
4.4.3.4. Die Rollen :INSTRUMENT und :METHOD	84
4.4.3.5. Die Rolle :SOURCE	85
4.4.3.6. Die Rolle :GOAL	85
4.4.3.7. Die Rollen :TIME-PERIOD und :TIME-POINT	86
4.4.3.8. Die Rolle :PLACE	86
4.4.3.9. Die Rolle :ELEMENTS	87
4.4.3.10. Die Rolle :MODALITY	87
4.4.4. Unterklassen von CASE-SCHEMA	88
4.4.4.1. Die Klasse COMPARE	88
4.4.4.2. Die Klasse OBJECT-AS-PREP-OBJ	89
4.4.4.3. Die Klasse GOAL-TO-DIROBJ&OBJECT-TO-PREP-OBJ	89
4.4.4.4. Die Klasse OBJECT-TO-SUBJ	90
4.5. Die Klasse CONCEPT-SCHEMA	91
4.5.1. Die initiale IRS-Struktur	91
4.5.2. Die Rolle NAME	91
4.5.3. Die Rolle :SPECIALIZE	92
4.5.3.1. Terminaler :SPECIALIZE-Füller	92

4.5.3.2. Nicht-terminale Füller der :SPECIALIZE-Rolle	94
4.5.3.3. Bildung eines Nominalkompositums	94
4.5.3.4. Anschluß als Possessivattribut	94
4.5.3.5. Sonderfälle	95
4.5.4. Die Rolle :ATTRIBUTES	95
4.5.5. Die Rolle :CONCERN	96
4.5.6. Die Rolle :POSSESSOR	97
4.5.7. Die Rolle :ROLE	98
4.5.8. Die Rolle :ELEMENTS	98
4.5.9. Die Rolle :PLACE	99
4.5.10. Die Rolle :QUANTIFIER	99
4.5.11. Die Rolle :MODALITY	101
4.6. Unterklassen von RELATION-SCHEMA	101
4.6.1. CONCERN-REL	101
4.6.2. *PHRASE-SEQUENCE*	101
4.6.3. *APPENDIX-STRUCT*	102
4.6.4. ENUMERATION	102
4.6.4.1. Die Reihenfolge bei der Realisierung	103
4.6.4.2. Erhaltung möglicher Lesarten	104
4.6.4.3. Enumerierte Adjektive	104
4.6.4.4. Koordinierte Sätze	105
4.6.5. ENUMERATION-OR	105
4.6.6. ISA-REL	105
4.6.7. PROPERTY-REL	106
4.6.8. PURPOSE-REL	107
4.6.9. SCOPE-REL	109
4.6.10. LOCATION-DESCR	111
5. Die Entwicklung des Textgenerators SEMTEX	113
5.1. Untersuchungen an Zeitungstexten	116
5.1.1. Die Fragestellung	116
5.1.2. Exkurs: Textsorten in Zeitungstexten	116
5.1.3. Inhaltliche Charakterisierung der untersuchten Texte	117
5.1.4. Exkurs: Wie Arbeitsmarktmeldungen zustandekommen	118
5.1.5. Wissensquellen beim Textverstehen	119
5.1.6. Elemente eines Lesermodells	119
5.1.7. Die Strukturierung von Texten	120
5.1.8. Schemata beim Verstehen von Argumentation	122
5.1.9. Schemata für Erklärungen	123
5.1.10. Schemata für Interpretationen	125
5.1.11. Rhetorische Schemata	126
5.2. Rekonstruktion und Paraphrasierung vorgefundener Textpassagen	127
5.2.1. Die Sprache zur Inhaltsdarstellung	127
5.2.2. Erweiterungen des Generatorkerns	128
5.2.2.1. Definition anwendungsspezifischer Alias-Namen für Rollen	129
5.2.2.2. Das Schema CONTINUING	130
5.2.2.3. Realisierungsmöglichkeiten der Klasse ALTHOUGH	130
5.2.2.4. Realisierungsmöglichkeiten der Klasse BECAUSE-OF	132
5.3. AMEX ~ von Daten zum Textplan	133
5.3.1. Die Arbeitsmarkt-Situation	133
5.3.2. Die Schnittstelle zur Textgenerierung	134
5.3.3. Wissen über temporale Beziehungen	137

5.3.4. Wissen über Veränderungen	139
5.3.5. Die Elemente des Textplans	141
5.3.5.1. Das Format für Elemente des Textplans	141
5.3.5.2. Arbeitsmarktgrößen	141
5.3.5.3. Veränderungen	141
5.3.5.4. Wertangaben	143
5.3.5.5. Vergleiche	143
5.3.5.6. Zeitangaben	144
5.3.5.7. Rhetorische Strukturen	145
5.4. SEMTEX – vom Textplan zum Text	147
5.4.1. Die Rolle des Kontexts	147
5.4.2. Die Objekte von SEMTEX	148
5.4.2.1. Die Klasse *TEXT*	148
5.4.2.2. Die Klasse *CONTEXT*	149
5.4.3. Die Verwaltung des Kontexts	149
5.4.3.1. LEXICAL-CHOICES	149
5.4.3.2. EXPRESSED-ROLES	151
5.4.3.3. REFERABLE-ROLES	152
5.4.4. Die Behandlung von Tempus und Zeitangaben	153
5.4.4.1. Die Realisierung von Monaten	153
5.4.4.2. Die Bestimmung der grammatischen Zeit	155
5.4.4.3. Variation des Sprechzeitpunkts	156
5.4.5. Kontextsensitive Realisierung von Arbeitsmarktgrößen	157
6. Zusammenfassung, Vergleiche, Ausblick	159
6.1. Vergleich der beiden Anwendungen	159
6.2. SEMTEX im Vergleich	161
6.2.1. Die Rolle des Fokus	161
6.2.2. Strukturen für Hilfe-Texte	162
6.2.3. Pläne und Motive – oder: Macbeth nacherzählt	165
6.3. Erweiterungen des Systems	167
6.3.1. Das Gesamtziel weiterer Arbeiten	167
6.3.2. Ausschalten potentieller Fehlerquellen	168
6.3.3. Begonnene und geplante Arbeiten	170
6.3.3.1. IRS-Sprache und SUTRA-S	170
6.3.3.2. Semantisch/deutsche Lexika	171
6.3.3.3. Generatorkern	172
6.3.3.4. Textgenerierung	172
6.3.3.5. Textplanung	173
6.4. Anwendung des Systems auf andere Gebiete	173
6.4.1. GEOTEX	173
6.4.2. Hilfesysteme	173
6.4.3. Generierung aus EUROTRA-Strukturen	174
Anhang A. Formale Syntax für Instantiierte Realisierungsschemata (IRS)	175
Anhang B. IRS-Strukturen zu verschiedenen Nominalgruppen aus Titeln	179
Anhang C. Ein Vergleich zwischen IRS- und F-Strukturen	185

Anhang D. Verzeichnis der verwendeten Implementationen	187
Anhang E. Zur Lexikalisierung semantischer Symbole	189
Anhang F. IRS-Tools	193
Anhang G. Wie man sich SEMTEX nutzbar macht	197
Literaturverzeichnis	201
Danksagung	209
Lebenslauf	211