CHAPTER 6

REALIZATION

6.1 INTRODUCTION

Traditionally, the formal properties of the words and sentences of a language includes morphology, which describes the ways in which words are formed from smaller units or other words, and syntax, which describes how words combine into sentences. It sometimes also includes phonology and semantics. A central branch of modern linguistics is generative grammar, which seeks to provide precise formal descriptions of the grammatical systems of languages and to develop a theory of universal grammar: a set of general statements about the structure of human languages (Migo Otani 1988). Transformational grammar is a form of generative grammar which makes use of operations known as 'transformations', which systematically indicate the links between various types of sentence and derive one type from the other.

6.2 TYPES OF GRAMMAR

All the various types of grammars are also used for building NLG systems.
6.2.1 Generative grammar

The theory that for each language a set of rules can be formulated capable of 'generating' the infinite number of possible sentences of that language and providing them with the correct structural description; a set of rules in such a grammar (Robin Elhadad 1992).

6.2.2 Transformational grammar

(also transformational-generative grammar) a grammar consisting of a set of rules which generate all and only the grammatical sentences of a language. Some of these rules are phrase-structure rules, which, when applied to the lexicon of the language, generate the logical deep structure of its sentences. Others are transformational rules, which generate the surface structure of sentences from their deep structure, and which account for logically related structures, for example active and passive sentences. Noam Chomsky in Syntactic Structures first described transformational grammar theoretically.

6.2.3 Finite state grammar

A form of grammar in which sentences are characterized in terms of the transitions of an automaton from one state to another.

6.2.4 Context-free grammar

A form of phrase structure grammar in which each rule holds for a specific category regardless of context: hence, more fully, context-free phrase structure grammar.
6.2.5  Phrase structure grammar

Any form of generative grammar consisting only of phrase structure rules. Hence any grammar which assigns to sentences a type of structure that can be represented by a single phrase structure tree.

6.2.6  Head driven Phrase Structure Grammar:

The head constituent of a phrase is a central notion. Otherwise, the grammar borrows freely from previous unification-based grammars. The treatment of syntactic categories, syntactic features, and some of the principles are from generalized phrase structure grammar (Gazdar et al., 1985). A number of the lexical rules are similar to those in lexical-functional grammar (Kaplan & Bresnan, 1982).

6.2.7  Tree Unification Grammar

Tree unification grammar (Popowich, 1989a, 1989b, 1993) resembles head-driven phrase structure grammar in certain respects. Both formalisms use signs as information structures and provide facilities for establishing the relationships between signs. Tree unification grammar incorporates a greater degree of lexicalization, using one rule as opposed to the four or five rules described in Pollard and Sag (1987, 1994). Head-driven phrase structure grammar also relies on the use of universal grammar principles which apply in conjunction with the grammar rules.
Phrase structure tree

A tree diagram which shows the division of a form into successively smaller constituents and labels each as belonging to one or more categories.

Tree diagram

Any branching diagram in which different branches are connected only at a point of origin, and all are connected, directly or indirectly, to one node which is the origin of the whole: e.g. a 'family tree' which displays the genetic classification of languages, a phrase structure tree, a dependency tree (Ward N 1990, 1992, 1994).

6.2.8 Context-sensitive grammar

A form of phrase structure grammar which is not subject to the restriction that defines a context-free grammar: hence, more fully, context-sensitive phrase structure grammar. A context-sensitive language is a formal language that can be generated by a context-sensitive grammar.

6.2.9 Generalized Phrase Structure Grammar

Formal model of syntax developed by G. J. M. Gazdar, and subsequently by others, from the end of the 1970s. Basically a phrase structure grammar, but with additional devices, in particular the use of metarules and slash categories, that removed the need, assumed in most theories of generative grammar at the time, for transformations. A passive sentence, for example, is characterized by phrase structure rules derived in part, by a higher-level rule,
from those that characterize an active; hence there is no need for a transformation that relates such sentences directly. In an interrogative like Who can you see?, who is related, by a phrase structure rule, to a phrase that must contain a null noun phrase. This removes the need for a transformation (e.g. move [alpha]) which would move who from the object position.

### 6.2.10 Case Grammar

A theory developed within generative grammar by the US linguist Charles Fillmore, c. 1970. The theory was intended to capture the constant deep semantic relations found, especially for noun phrases, in different syntactic functions and employs the traditional term case to name such relations. In the sentences "I split the log with my axe", "My axe split the log", "The log split", and "The log was split by my axe", the semantic role of the log is patient (affected by the action) and that of my axe is instrument (used to perform the action). Influenced by formal logic, case grammar characterized the underlying structure of a sentence as having two parts: features such as tense, interrogation, and negation, which relate to the sentence as a whole; the verb and those arguments that accompany it, representing together the basic proposition denoted by the sentence. The arguments are expressed by noun phrases that have different semantic roles, such as agent, patient, instrument. These roles, termed cases, can be represented in constructions in various syntactic functions (as shown above). Although linguists have appreciated the insights of case grammar, the theory has not been further developed because of difficulties in determining and defining the set of cases, and because of problems in formalizing the model in a comprehensive description. References continue to be made to argument-like cases in grammars, but variations appear in the number and names of cases. Among the cases commonly referred to be: agent,
patient (also objective, affected), recipient, instrument, experiencer, source, goal, location/locative, path. The term case derives from its use in inflectional languages for the variant forms of a word that relate to syntactic functions: such inflectional forms as nominative, accusative, and genitive in Latin.

6.2.11 Dependency grammar

A dependency grammar is a generative grammar by which the structures of sentences are represented by dependency trees. Thence generally of other grammars that assign dependency relations, whether or not the structures are restricted to that form.

6.2.12 Systemic Grammar

Model of functional syntax developed by Halliday from the late 1950s. The basic idea is that any act of communication realizes a set of choices: thus e.g. the utterance of She went out realizes, among others, the choice of a declarative structure. Each choice is at a certain level in a hierarchy of ranks: e.g. the choice of declarative is at clause level. It is also related to other choices on a scale of delicacy or detail: e.g. the choice of interrogative instead of declarative would entail a further choice between polar interrogative and wh-interrogative. Each individual set of choices forms a system: thus polar interrogative and wh-interrogative form one system, declarative and interrogative form or are part of another. A grammar will accordingly describe the systems of a language, the relations between them, and the ways in which they are realized, to a level of detail at which all remaining choices are between open sets of lexical units (Fawcett Tucker 1993, Wanner L 1992, 1994, Busemann S, 1993).
6.3 GRAMMAR FOR SOFTWARE DOCUMENTATION

While generating software document, the phrases obtained from the planner phase are realized using extended BNF grammar (Uma G.V. Kanimozhi N.S.M. Geetha T.V. 1999).

Phrases Generated from various patterns

1. Entity Relation Entity
2. Entity Attribute
3. Process Request Entity/Process/Data store
4. Process sub process
5. Process Reply Entity/Process/Data store
6. Sub Process request Entity/Process/Data store
7. Sub Process reply Entity/Process/Data store
8. Initial state event trigger precondition
9. Precondition Post condition next state
10. Entity cost Expected cost
11. Entity effort expected effort
12. Process data
13. Data process
14. Sub process process
15. Entity Entity
16. Entity Data Entity
17. Entity constraint (Technical/Environmental)
18. Process Constraint (Technical/Environmental)
19. Sub Process Constraint (Technical/Environmental)
20. Entity operator Entity
21. Entity process Entity
22. State1 before state2
23. Entity Initial state final state
24. State action
25. Entity process data

Grammar for generating functional description from organized DFD:

\(<\text{Process}\> = \text{Consists of (}\langle\text{sub process}\rangle\) * /}
\gets \langle\text{I/P}\rangle \text{ from } \{ \langle\text{data store}\rangle/\langle\text{process}\rangle/\langle\text{entity}\rangle\}/
gives\langle\text{O/P}\rangle \text{ to } \{ \langle\text{data store}\rangle/\langle\text{process}\rangle/\langle\text{entity}\rangle\}
\langle\text{I/P}\rangle=\langle\text{data}\rangle \langle\text{O/P}\rangle = \langle\text{I/P}\rangle/\langle\text{changes}\rangle
\langle\text{data}\rangle=\langle\text{structured data}\rangle/\langle\text{individual data}\rangle
\langle\text{structured data}\rangle = \text{case [description]}/\langle\text{system [description]}\rangle/\langle\text{structural [description]}\rangle/\langle\text{specification [format]}\rangle/\langle\text{state [description]}\rangle

Case description ::= list of features
System description ::= Functions and Components
Structural description = skeletal
Specification format = Static/behavioural
\langle\text{Individual datum}\rangle = \langle\text{Constraint}\rangle/\langle\text{Requirement}\rangle/
\langle\text{Symptom}\rangle/\langle\text{Observable}\rangle/
\langle\text{Variable}\rangle/\langle\text{Value}\rangle/

Grammar for generating information description
\langle\text{Entity}\rangle = \text{Contains(}\langle\text{attributes}\rangle\) has a relation of \langle\text{Relation}\rangle with\langle\text{entity}\rangle \text{ by } \langle\text{Cardinality}\rangle
\langle\text{attributes}\rangle = \langle\text{value}\rangle \langle\text{type}\rangle \langle\text{Range}\rangle
Grammar for generating Behavioural description

State = process has { initial state / intermediate state of / state } when the event occurs it triggers gets new state process has final state of

While generating software document, thematic roles have to be considered in order to produce complete document without any semantic ambiguity.

Thematic Roles

In predicate calculus, they can be represented by dyadic predicates. In conceptual graphs, conceptual relations that link the concept of a verb to the concepts of the participants in the occurrent expressed by the verb represent them. In frame systems, slots in the frame for the corresponding verb represent them. All those notations are equivalent ways of representing the links between a process and its participants.

In the KR ontology, the thematic roles are classified as subtypes of Participant, which is further subdivided by two pairs of distinctions: determinant or immanent and source or product.

Graph representation of the subtypes of Participant.
The same subdivision is represented by the matrix

<table>
<thead>
<tr>
<th>Source</th>
<th>Product</th>
</tr>
</thead>
<tbody>
<tr>
<td>Determinant</td>
<td>Initiator</td>
</tr>
<tr>
<td>Immanent</td>
<td>Resource</td>
</tr>
</tbody>
</table>

Matrix representation of the subtypes of Participant

Every participant is an entity that plays some role in a process. In natural languages, those distinctions are expressed by grammatical markers such as prepositions and case markers, which link the verbs that express the processes to the nouns that express the participants. In logic, relations or predicates that link the symbols that identify the processes to the symbols that identify the participants express those distinctions.

- A **determinant** participant determines the direction of the process, either from the beginning as the initiator or from the end as the goal.
- An **immanent** participant is present throughout the process, but does not actively control what happens.
- A **source** must be present at the beginning of the process, but need not participate throughout the process.
- A **product** must be present at the end of the process but need not participate throughout the process.
After analyzing and summarizing various systems of case relations or thematic roles, Harold Somers (1987) organized them in a matrix with four types of participants at the top and six categories of verbs along the side. In the 24 boxes of the matrix, Somers had some boxes with duplicate role names and some boxes with two roles that were distinguished by other properties: ±animate, ±physical, ±dynamic, or ±volitional. In using Somers's classification, Judith Dick (1991) applied the roles to conceptual graphs as a knowledge representation for legal arguments.

Stimulated by the work of Julius Moravcsik (1991) and James Pustejovsky (1995), Sowa (1996) related the four columns of the Somers-Dick matrix to Aristotle's four causes or aitia, as described in the Metaphysics:

- **Initiator** corresponds to Aristotle's efficient cause, "whereby a change or a state is initiated" (1013b23).
- **Resource** corresponds to the material cause, which is "the matter or the substrate (hypokeimenon)" (983a30).
- **Goal** corresponds to the final cause, which is "the purpose or the benefit; for this is the goal (telos) of any generation or motion" (983a32).
- **Essence** corresponds to the formal cause, which is "the essence (ousia) or what it is (to ti èn einai)" (983a27).

The four terms initiator, resource, goal, and essence better describe the participants of an action than the traditional translations for Aristotle's four causes.
Table 6.1 Thematic roles as sub types – the four types of participants

<table>
<thead>
<tr>
<th>Action</th>
<th>Initiator</th>
<th>Resource</th>
<th>Goal</th>
<th>Essence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Agent,</td>
<td>Instrument</td>
<td>Result,</td>
<td>Patient,</td>
</tr>
<tr>
<td></td>
<td>Effector</td>
<td></td>
<td>Recipient</td>
<td>Theme</td>
</tr>
<tr>
<td>Process</td>
<td>Agent,</td>
<td>Matter</td>
<td>Result,</td>
<td>Patient,</td>
</tr>
<tr>
<td></td>
<td>Origin</td>
<td></td>
<td>Recipient</td>
<td>Theme</td>
</tr>
<tr>
<td>Transfer</td>
<td>Agent,</td>
<td>Instrument,</td>
<td>Experiencer,</td>
<td>Theme</td>
</tr>
<tr>
<td></td>
<td>Origin</td>
<td>Medium</td>
<td>Recipient</td>
<td></td>
</tr>
<tr>
<td>Spatial</td>
<td>Origin</td>
<td>Path</td>
<td>Destination</td>
<td>Location</td>
</tr>
<tr>
<td>Temporal</td>
<td>Start</td>
<td>Duration</td>
<td>Completion</td>
<td>PointInTime</td>
</tr>
<tr>
<td>Ambient</td>
<td>Origin</td>
<td>Instrument,</td>
<td>Result</td>
<td>Theme</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Matter</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The options and duplications in the boxes of the table indicate that further distinctions can be made. The option of Agent or Effectors as the initiator of an action is determined by the distinction of a voluntary initiator (Agent) or involuntary initiator (Effectors). The duplication of the Agent role for actions, processes, and transfers indicates implicit interactions between the types of verbs and the types of participants.

But according to the hierarchy of participants, Result < Goal < Product, and Patient < Essence < Product. Since Product is a common super type, the initial interpretation could have the label Prod. The resulting
representation would be a single conceptual graph that expressed exactly the same information as the original sentence without making any assumptions about the immanent or determinant status of the pie:

But Agent < Initiator < Source, Effectors < Initiator < Source, and Instrument < Resource < Source. Since all three of the more specialized types of participants are subtypes of Source, a single CG with the Source relation would express the equivalent information in the original sentence.

The first term in each entry is the name of the role, such as Agent. Following the symbol < is the supertype, such as Initiator. Then comes an abbreviation, such as Agnt, followed by category restrictions on the concept type of the verb (Act) and the concept type of the participant (Animate).

Agent < Initiator; Agnt(Act,Animate).
An active animate entity that voluntarily initiates an action.

Beneficiary < Recipient; Benf(Act,Animate).
A recipient that derives a benefit from the successful completion of the event.

Completion < Goal; Cmpl(TemporalProcess,Physical).
A goal of a temporal process.

Destination < Goal; Dest(SpatialProcess,Physical).
A goal of a spatial process.

Duration < Resource; Dur(State,Interval).
A resource of a temporal process.
Effector $<$ Initiator$;$ Efct(Entity, Entity).
An active determinant source, either animate or inanimate, that initiates an action, but without voluntary intention.

Experiencer $<$ Goal$;$ Expr(State, Animate).
An active animate goal of an experience.

Instrument $<$ Resource$;$ Inst(Act, Entity).
A resource that is not changed by an event.

Location $<$ Essence$;$ Loc(Physical, Physical).
An essential participant of a spatial nexus.

Matter $<$ Resource$;$ Matr(Act, Substance).
A resource that is changed by the event.

Medium $<$ Resource; Med(Transfer, Physical).
A physical resource for transmitting information, such as the sound of speech or the electromagnetic signals that transmit data.

Origin $<$ Initiator$;$ Orgn(Process, Physical).
A passive determinant source of a spatial or ambient nexus.

Path $<$ Resource; Path(Process, Place).
A resource of a spatial nexus.
An essential participant that undergoes some structural change as a result of the event.

PointInTime $<$ Essence$;$ PTim(Physical, Time).
An essential participant of a temporal nexus.
Recipient < Goal; Rcpt(Act,Animate).
An animate goal of an act.
Result < Goal; Rslt(Process,Entity).
An inanimate goal of an act.
Start < Initiator; Strt(Entity,Time).
A determinant source of a temporal nexus.
Theme < Essence; Thme(Situation,Entity).
An essential participant that may be moved, said, or experienced, but is not structurally changed.

The above shows some sample branches of the ontology under Agent and Theme. Doer, for example, has a subtype Driver, which has more specific subtypes. In principle, any of the thematic roles could be subdivided further to show distinctions that might be significant in some culture or domain of interest. Other thematic roles listed in Section B.4 could also be subdivided further to represent the participants of specific concept types: Sayer < Agent; Senser < Agent; Addressee < Recipient; Experienced < Theme; Moved < Theme; Said < Theme.

Although the thematic roles represent a linguistically important class of ontological categories, their common super type is also represented. Therefore, Role would include many types that are not directly associated with verbs.
6.4 CONCLUSION

While realizing any software document, document should be grammatically correct and it should also mention about the roles involved such as participant, process etc.