CHAPTER 5

A COMPARATIVE STUDY OF RST AND saṅgati

This chapter discusses the concept of saṅgati. The usage of RST and saṅgatis in terms of discourse parsing is studied, and the comparisons based on it are presented in this chapter.

5.1 saṅgatis- AN OVERVIEW

Sanskrit literature has many nuggets that could be applied to modern linguistic applications. One such nugget is the concept of saṅgati. saṅgatis induce the desire to know what is being said next in a text; or, expositions of Vedic texts are organized in the form of sūtra (statements), adhikaraṇa (sub topic), pāda (section), adhyāya (chapter) and śāstra (whole content). sūtra express content in crisp, short statements; adhikaraṇa is the organization of a set of related sutras on a sub-topic. A set of adhikaraṇas forms a pāda, and a set of pādas forms an adhyāya. sūtra being cryptic in nature, need to be explained.

The explanation is normally organized at the level of adhikaraṇa. An adhikaraṇa is said to have five components, namely, the subject of discussion, doubt/ambiguity in understanding the subject, saṅgati indicating coherence/continuity for this discussion, the opponent’s view and the proponent’s (proposed) view. Of these, saṅgati is explained at various levels. At the sutra level, in terms of how this sutra is related to the previous sutra; at the adhikaraṇa level as to how this sutra is relevant to the adhikaraṇa, at the
Pāda level as to how it is relevant to that pāda, and so on. Similarly, saṅgati is discussed between adhikaraṇas, and between pādas as well. There are many types of saṅgati relations that have been used in the expository text of the āśtrās. A partial list of saṅgati relations has been identified manually, and is used in this research. The list of saṅgatis considered, their equivalent meaning in English, and the type of texts linked by saṅgatis are listed in Table 5.1 (Srivenkatesha Subba Yajva Shastri 1934).

**Table 5.1 Definition of saṅgatis**

<table>
<thead>
<tr>
<th>S. No</th>
<th>saṅgati</th>
<th>English Meaning</th>
<th>Texts Linked by saṅgati</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Text&lt;sub&gt;1&lt;/sub&gt;</td>
</tr>
<tr>
<td>1</td>
<td>Upodghāta</td>
<td>Introduction</td>
<td>Introduction of a scenario</td>
</tr>
<tr>
<td>2</td>
<td>apavāda</td>
<td>Exception</td>
<td>A scenario</td>
</tr>
<tr>
<td>3</td>
<td>ōkṣepa</td>
<td>Objection</td>
<td>A scenario</td>
</tr>
<tr>
<td>4</td>
<td>prāśangika</td>
<td>Related</td>
<td>A scenario</td>
</tr>
<tr>
<td>5</td>
<td>uttōma</td>
<td>Arises</td>
<td>A scenario</td>
</tr>
<tr>
<td>6</td>
<td>sthirakaraṇa</td>
<td>Strengthen</td>
<td>A scenario</td>
</tr>
<tr>
<td>7</td>
<td>cūtidesika</td>
<td>Transference</td>
<td>Scenario&lt;sub&gt;1&lt;/sub&gt;</td>
</tr>
<tr>
<td>8</td>
<td>Pratyavasthana</td>
<td>Re-instate</td>
<td>Scenario&lt;sub&gt;1&lt;/sub&gt;</td>
</tr>
<tr>
<td>9</td>
<td>drṣṭanta</td>
<td>Example</td>
<td>A scenario</td>
</tr>
<tr>
<td>10</td>
<td>pratydaharaṇa</td>
<td>Counter-Example</td>
<td>A scenario</td>
</tr>
<tr>
<td>11</td>
<td>Anantara</td>
<td>Sequence</td>
<td>Scenario&lt;sub&gt;1&lt;/sub&gt;</td>
</tr>
<tr>
<td>12</td>
<td>visēṣa</td>
<td>Special Case</td>
<td>A scenario</td>
</tr>
</tbody>
</table>
It can be observed from the table that similar to RST–based discourse relations, saṅgatis also link two different text units. The proposed research makes use of the terms, nucleus and satellites, to denote and distinguish the text units, that convey the salient information and the additional information respectively, as was done in RST. Furthermore, it could be observed that saṅgatis, namely, upodghāta and anantara, are equivalent to the discourse relations, Preparation and Sequence respectively. There are a few saṅgatis that are similar to the discourse relations but not equivalent. The saṅgati, uttāna is similar to the discourse relation, Antithesis, while the saṅgatis, Pratyudhāraṇa and apavāda are similar to the discourse relation, Contrast.

This sort of equivalence and similarities has led to the idea of using saṅgatis for discourse parsing, in union with RST.

In the next section, each saṅgati is explained with an example and a comparison has been made with the RST discourse relations.

5.2 saṅgatis–EXAMPLES

upodghāta

The upodghāta saṅgati links the introductory part of any text, to its respective explanatory part in the text. Figure 5.1 illustrates the usage of upodghāta saṅgati, using a text describing cancer.

![Figure 5.1 Usage of upodghāta saṅgati](image)

Cancer is a broad group of various diseases, all involving unregulated cell growth.

Text containing details of Cancer

upodghāta(intr oduction)
It can be observed that the text that contains the introductory details of the following text becomes the nucleus and the latter becomes the satellite. This observation is similar to the discourse relation, *Preparation*, which also prepares the reader by providing necessary details of the following text.

**pratyavasthana, anantara and prāsaṅgika**

The saṅgatis, namely, *anantara* and *prāsaṅgika* have been explained in union with the *pratyavasthana* saṅgati, in order to make the understanding of the *pratyavasthana* saṅgati better. The saṅgati, *anantara* links the texts that are in sequence; the saṅgati, *prāsaṅgika* links texts that are related but not necessarily in sequence. The saṅgati, *pratyavasthana* links the text that deviates from the topic and the text that returns to the topic in focus. Figure 5.2 and Example 5.1 explain the usage of these three saṅgatis.

**Example 5.1**

Ram bought a new car (Sen₁). The brand name of the car is Ford Figo (Sen₂). The founder of Ford Company is Mr. Henry Ford (Sen₃). The colour of Ram’s car is red (Sen₄).

![Figure 5.2 Usage of anantara, prāsaṅgika and pratyavasthana saṅgatis](image-url)
It can be seen from the above figure that the sentences, Sen₁ and Sen₂ are in sequence and therefore linked by the anantara saṅgati. The sentence, Sen₃ deviates from the topic, “Ram’s new car” but is related to the previous sentence, Sen₂. Hence, Sen₃ and Sen₂ are linked by prāsaṅgika saṅgati. The sentence, Sen₄ turns to the topic in focus, Ram’s new car. Hence, Sen₄ is connected to Sen₃ using the pratyavasthana saṅgati.

Both the anatara and prāsaṅgika saṅgatis are multi-nuclear, wherein both the text units are nuclei, and there are no satellites. In the pratyavasthana saṅgati, the text that aids in resuming the topic in focus is the nucleus and the other becomes the satellite.

pratyavasthana and prāsaṅgika are unique saṅgatis and RST does not have equivalent discourse relations. The RST-based discourse relation, Sequence is equivalent to the anantara saṅgati.

This kind of scenario as illustrated in Example 5.1, can be seen in a text containing dialogues and argumentations.

ātidesīka

The ātidesīka saṅgati refers to using a previously discussed concept, to which some modifications are presented in the current context. The concept may refer to any entity, event or action. ātidesīka can be thought of as an inheritance or transference property. Figure 5.3 and Example 5.2 illustrate the usage of ātidesīka saṅgati.

Example 5.2

Maruti Swift has a 1.2L engine, and provides 16K/L mileage. Dzire is a Maruti Swift with a hatch-back, and is a sedan.
It can be observed that the ātidesīka saṅgati illustrates the inheritance of the properties of the car brand, Maruthi Swift in the car brand, Maruthi Swift Dezire. The text that contains the information inherited from some other text becomes the nucleus, and the other becomes the satellite.

ātidesīka is a unique saṅgati and does not have an equivalent RST-based discourse relation. The usage of ātidesīka saṅgati can benefit QA systems for answering definitional and factoid type of QA systems efficiently. If a QA system is given a query, “what are the similarities or differences between the Maruthi Swift and Maruthi Swift Dezire cars?”, the texts linked by the ātidesīka saṅgati can be retrieved as the answers to the query.

uttāna

The uttāna saṅgati indicates a new issue that arises amidst a topic on focus. The dialogue given in Example 5.3 shows the usage of uttāna saṅgati.

Example 5.3

Utterance 1 : By constructing a dam in that village, we solved the drinking water problem completely.

Utterance 2 : We may have to face many problems, in constructing a dam in a village. Sufficient power supply may not be available on hand, and we need to spend more on that. Also, floods and heavy rains may ruin the crops. Will it be a good idea to construct the dam in a nearby town?
*uttāna* saṅgati links Utterance 1 and Utterance 2 shown in Example 5.3. It can be seen that the *uttāna* saṅgati links the text that states the drawbacks of constructing a dam in a village with the text that puts forth an idea of constructing a dam.

The *uttāna* saṅgati is multi-nuclear, where both the text units contain salient information. The discourse relations, namely, Antithesis and Contrast may seem to be similar to the *uttāna* saṅgati, but both these discourse relations link texts that convey quite opposite views, but the *uttāna* saṅgati may not necessarily link such kinds of texts, as the text that expresses the new issue may not always be opposite to the other linked texts. Hence, *uttāna* is also a unique saṅgati.

The IR, QA and summary generation systems can be benefitted by the usage of *uttāna* saṅgati, when a different viewpoint of any entity needs to be identified.

**sthirkaraṇa**

The *sthirkaraṇa* saṅgati indicates a text that expresses the importance of or strengthens a situation in focus. Example 5.4 explains the usage of *sthirkaraṇa*.

**Example 5.4**

Text 1 : The Government has launched a self-help groups scheme for women across the country.

Text 2 : About one crore women could be benefitted using this scheme.

It can be seen that, Text 2 strengthens the viewpoint mentioned in Text1. Text 1 becomes the nucleus, and Text 2 that contains the strengthening fact of Text 1 becomes the satellite.
The discourse relations, *Justification and Motivation* in RST may seem to be similar to the sthirakaraṇa saṅgati, as *Justification* supports the writer’s view expressed in the nucleus text, and the *Motivation* provides the motivating reasons for the facts stated in the nucleus. The facts represented in the satellites of the discourse relations, *Justification* and *Motivation* may not always strengthen the facts stated in the nucleus. Hence, the saṅgati, sthirakaraṇa is also a unique saṅgati, and does not have any equivalent or similar RST-based discourse relation.

*sthirakaraṇa* can help to answer various types of queries in IR and QA systems. A few types are given below.

- What are the advantages of x?
- What are the benefits of x?
- What are the positive sides of x?
- What is the importance of x?

**dṛṣṭanta and pratyudharaṇa**

The *dṛṣṭanta* saṅgati illustrates an example scenario pertaining to the main topic, whereas, *pratyudharaṇa* saṅgati indicates a text that contains a counter example. Example 5.5 explains the usage of these two saṅgatis.

**Example 5.5**

Text1 : In Tamil Nadu, power cuts occur frequently due to shortage of power supply.

Text 2 : In the villages of Tamil Nadu, power shut down is implemented for six hours, and in cities for two hours per day.

Text 3 : Software industries and Call centres in Tamil Nadu barely face a power cut.
Text 2 is an example scenario of Text 1, and is linked by the \textit{drśṭanta} saṅgati. Text 3 is a counter example scenario of Text 1, and is linked by the \textit{pratyudharaṇa} saṅgati. Both \textit{drśṭanta} and \textit{pratyudharaṇa} saṅgatis are multi-nuclear.

RST handles a similar situation, by using the \textit{Contrast} and \textit{Elaboration} relations. But \textit{Contrast} and \textit{Elaboration} relations need not always handle an example and counter example scenario. Hence, both \textit{drśṭanta} and \textit{pratyudharaṇa} saṅgatis are unique saṅgatis.

Examples can be seen in all kind of texts. Counter examples are mostly seen in arguments. These saṅgatis can find their use in argumentative analysis, QA and IR systems.

\textit{visēṣa}

The \textit{visēṣa} saṅgati explains the speciality of some entity. Example 5.6 explains the usage of \textit{visēṣa} saṅgati.

\textbf{Example 5.6}

Text\textsubscript{1} : Tirunelveli is a district in Tamil Nadu.

Text\textsubscript{2} : Tirunelveli is best known for the mouth-watering sweet, “halwa”.

In the above example, Text\textsubscript{2} is connected to Text\textsubscript{1} by \textit{visēṣa} saṅgati. The text containing the speciality becomes the satellite, and the other becomes the nucleus.

RST handles such situations through the \textit{Elaboration} relation. Again, the \textit{Elaboration} relation need not convey the speciality always. Hence, \textit{visēṣa} saṅgati is also a unique saṅgati. This saṅgati too can be used by IR and QA systems.
apavāda

The *apavāda* saṅgati indicates an exception scenario. Example 5.7 illustrates the usage of *apavāda* saṅgati.

**Example 5.7**

Text 1 : All birds can fly.

Text 2 : Penguins cannot fly.

Text 2 is an exception scenario of Text1. This saṅgati is also very common in all kind of texts. *apavāda* saṅgati is a multi-nuclear saṅgati. The *Contrast* relation in RST may seem similar to *apavāda* saṅgati, but a *Contrast* relation does not always mean an exception scenario.

ākṣepa

The *ākṣepa* saṅgati indicates an objection. Example 5.8 describes the usage of *ākṣepa* saṅgati.

**Example 5.8**

Text 1 : Boys have fared better than girls in the entrance examinations. They study better.

Text 2 : Not necessarily. Girls have done better than boys in the board examinations.

*ākṣepa* links Text 1 and Text 2 given in the above example, and is common in dialogues and arguments. *ākṣepa* is also a multi-nuclear saṅgati.

RST deals with this situation through the *Contrast* relation, but the *Contrast* relation need not always mean an objection.
5.3 **SUMMARY**

From the above discussion it can be observed that both RST and saṅgatis are capable of handling the same situation, but saṅgatis add another level of detail, and appear to be capable of linking a situation in a more focused manner, when compared to RST. In other words, a discourse relation overlaps with many saṅgatis. For instance, a situation handled by the single discourse relation, *Elaboration* is handled by different saṅgatis, namely, dṛśṭanta, pratyudharaṇa or viśēṣa depending on the nature of the situation.

On the other hand, there are many RST based discourse relations that are not present in saṅgatis. Discourse relations such as, *Concession, Evidence, Motivation, Restatement, Summary, Condition, Evaluation, Non Volitional Cause and Results, Volitional Cause and Results, Otherwise, Purpose, Solutionhood, Conjunction, Disjunction and Joint* do not have an equivalent saṅgati among the ones considered. Table 5.2 shows the list of unique saṅgatis and RST relations. Table 5.3 shows the list of equivalent saṅgatis and RST relations, and Table 5.4 shows the list of similar saṅgatis and RST relations.

**Table 5.2 Unique saṅgatis and RST Discourse Relations**

<table>
<thead>
<tr>
<th>Unique saṅgatis</th>
<th>Unique RST discourse Relations</th>
</tr>
</thead>
<tbody>
<tr>
<td>ātidesṭika</td>
<td>-</td>
</tr>
<tr>
<td>prāsangika</td>
<td>-</td>
</tr>
<tr>
<td>prathyavasthana</td>
<td>-</td>
</tr>
<tr>
<td>-</td>
<td>Evaluation</td>
</tr>
<tr>
<td>-</td>
<td>Concession</td>
</tr>
<tr>
<td>-</td>
<td>Background</td>
</tr>
</tbody>
</table>
Table 5.3 Equivalent saṅgatis and RST Discourse Relations

<table>
<thead>
<tr>
<th>saṅgatis</th>
<th>RST based Discourse Relations</th>
</tr>
</thead>
<tbody>
<tr>
<td>upodghāṭa</td>
<td>Preparation</td>
</tr>
<tr>
<td>anantara</td>
<td>Sequence</td>
</tr>
</tbody>
</table>

Table 5.4 Similar saṅgatis and RST Discourse Relations

<table>
<thead>
<tr>
<th>saṅgatis</th>
<th>RST based Discourse Relations</th>
</tr>
</thead>
<tbody>
<tr>
<td>uttōna</td>
<td>Antithesis</td>
</tr>
<tr>
<td>Pratyudharaṇa</td>
<td>Contrast</td>
</tr>
<tr>
<td>apavīda</td>
<td>Contrast</td>
</tr>
<tr>
<td>sthīrakaraṇa</td>
<td>Justification</td>
</tr>
</tbody>
</table>

Another difference between discourse relations and saṅgatis is that, most of the discourse relations can normally be identified by discourse markers, but saṅgatis need a lot of semantic processing to be identified. Furthermore, for clausal level discourse analysis, discourse relations are more effective than saṅgatis. saṅgatis go well beyond the clause level.

The aim of the proposed research is to make use of the unique qualities of both these ancient and modern text representations, and construct an efficient discourse parser.