Introduction

Punjabi is an Indo-Aryan language, widely spoken in the Punjab region of south Asia. The grammar of the Punjabi language is the study of the word order, case marking, verb conjugation, and other morphological and syntactic structures of the Punjabi language. Punjabi has a canonical word order of SOV (Subject-Object-Verb). It has postpositions rather than preposition. Morphological analyzer and morphological synthesizer are two essential and basic tools for building any language processing application for a natural language. Morphological analysis means to study the internal structure of the words of a language. A Morphological analyzer gives the morph analysis of a word i.e. for a given word a morphological analyzer will return its root word and word class along with other grammatical information depending upon its word class. Like for nouns it will provide gender, number, and case information and for verbs it will provide tense, phase etc. in addition to this. Morphological synthesizer does exactly the reverse of it, i.e. given a root word and grammatical information it will generate the word form of that root word.

(I) Morphological Analysis

Morphological analysis is essential for Punjabi it has a rich system of inflectional morphology as like other Indo-Aryan family languages. This paper presents the strategy for morphological analysis and synthesis for Punjabi language using paradigm approach. In linguistics, a morpheme is the smallest component of word, or other linguistic unit, that has semantic meaning. The morphological analyzer segment a word to a sequence of morphemes (root/stem and affixes), tag the part-of-speech (POS) of those morphemes and identify the morpho – syntactic relation between them. It provides the grammatical information depending upon the word category. For nouns it will provide gender, number, and case information and for verbs, it will be tense, aspects, and moods. The general format of morphological analysis of Punjabi is word = root/stem+suffix. All the word forms of Punjabi words are kept in the lexicon along with their root and other grammatical information

For Example

(1) Rama called Sita
rAmA ne sIwA nUm bulAiA
ਰਾਮ ਨੇ ਸੀਤਾ ਨੂੰ ਬੁਲਾਇਆ

(1) Noun Morphology: Punjabi distinguishes two genders, two numbers, and five cases of direct, oblique, vocative, ablative, and locative/instrumental. The latter two cases are essentially now vestigial, the ablative occurs only in the singular, in free variation with oblique case plus ablative postposition, and the locative/instrumental is confined to set adverbal expressions.
Samples Of different forms:
1. I did the work with my own hands.
   मैं अपने हाथीं काम किया
   mEz ApaNe haYWIz kaMma kIwA
2. He started from Delhi .
   उसे दिल्ली थिया
   uha xiYlIoz caYliA
3. He is at home .
   उसे घर में है
   uha Gara hI hE
4. He was expelled from the area .
   उसे इलाक़े विचारे कालू दिया
   uha ilAke viYcoz kaYDa xiYwAuha ilAke
table 2. Inflection of number and case on noun

<table>
<thead>
<tr>
<th>Gender</th>
<th>case</th>
<th>Affix</th>
<th>singular</th>
<th>plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>Masculine</td>
<td>Direct</td>
<td>e</td>
<td>muMdA</td>
<td>muMde</td>
</tr>
<tr>
<td>Masculine</td>
<td>Oblique</td>
<td>iAz</td>
<td>muMde</td>
<td>muMdiAz</td>
</tr>
</tbody>
</table>
(2) Verb Morphology

Verbs are the most complex part of speech in Punjabi. The Punjabi verbal system is largely structured around a combination of aspect and tense/mood. For example, ਜਾਂਦਾ ‘go’, ਕਰ ‘do’, ਸੁਣ ‘listen’, ਹੱਸ ‘laugh’, ਲਿਖ ‘write’ are all verbs. In sentences, these single word verbs are occasionally used and frequently lengthier phrases are used as compound verbs. For example,

1. ਜਾਂਦਾ ਹੈ ‘he goes is for present tense,
2. ਜਾਂਦਾ ਸੀ ‘he was going’ is for past tense, and
3. ਜਾਵੇਗਾ ‘he will go’ is for future tense. Please note that for future tense, no auxiliary verb is used.

Inflectional class wise grouping of the forms (or paradigm) of verb ਜਾਂਦਾ ‘go’ are provided here,

<table>
<thead>
<tr>
<th>Gender</th>
<th>case</th>
<th>Affix</th>
<th>singular</th>
<th>plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feminine</td>
<td>Direct</td>
<td>IAz</td>
<td>kudZI</td>
<td>kudZIAz</td>
</tr>
<tr>
<td>Feminine</td>
<td>Oblique</td>
<td>IAz</td>
<td>kudZI</td>
<td>kudZIAz</td>
</tr>
</tbody>
</table>

A Punjabi verb root is always in second person singular form in imperative mood. In a Punjabi sentence, verbs must agree with the subject or object of the sentence in terms of gender, number, and person.

We are interested only in simple verbs as they consist of one token and therefore can be captured morphologically. In contrast, conjunct and compound verbs can only be identified by the analysis of multiple words and mostly have a semantic role. Therefore, they should be handled in syntax. Simple verbs can be transitive or intransitive. Verbs inflect for tense/mood (subjunctive, perfective, imperfective), person (first, second, third), number and gender. Punjabi verbs take forms for future tense only, and for present and past tense, appropriate auxiliary verb form (for present or past tense) is used. For example

1. ਜਾਂਦਾ ਹੈ ‘he goes is for present tense,
2. ਜਾਂਦਾ ਸੀ ‘he was going’ is for past tense, and
3. ਜਾਵੇਗਾ ‘he will go’ is for future tense. Please note that for future tense, no auxiliary verb is used.

(3) Adjective Morphology

Adjective is a word that qualifies a noun. Most of the adjectival forms in Punjabi will alter due to the influence of its gender and number distinction. The semantic information of head noun and subject will have a major role in the
formation of adjectives in Punjabi. Adjectives may be divided into declinable and indeclinable categories. Declinables are marked, through termination, for the gender, number, case of the nouns they qualify.

Indeclinable adjectives are completely invariable, and can end in either consonants or vowels (including A and I)

(4) Pronoun Morphology

Pronoun is a word used instead of a noun and refers to a person or a thing that has already talked about. Pronoun can take the place of a noun in a sentence and function as a noun. In Punjabi, we have six types of pronouns:

1. Personal Pronoun: These are used to refer to first, second, and third person in sentences. ਮੈਂ ‘I’ is used for the first person and ਤੂੰ ‘you’ for the second person. For third person, demonstrative pronouns ਉਹ ‘that’ and ਇਹ ‘this’ are used. All the personal pronouns can equally be used for both the genders and change forms based on gender and number.

Personal pronouns take different forms for two cases – direct and oblique. For direct and oblique case, the same forms are applicable for the first and second person, but for the third person ਉਸ ‘us’ and ਇਸ ‘is’ are used.

<table>
<thead>
<tr>
<th>Gender</th>
<th>Case</th>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>Masculine</td>
<td>Direct</td>
<td>caMg_A</td>
<td>caMg_e</td>
</tr>
<tr>
<td></td>
<td>Oblique</td>
<td>caMg_e</td>
<td>caMg_iAM</td>
</tr>
<tr>
<td>Feminine</td>
<td></td>
<td>caMg_I</td>
<td>caMg_IAM</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gender</th>
<th>Case</th>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>Masculine</td>
<td>Direct</td>
<td>caMg_A</td>
<td>caMg_e</td>
</tr>
<tr>
<td></td>
<td>Oblique</td>
<td>caMg_e</td>
<td>caMg_iAM</td>
</tr>
<tr>
<td>Feminine</td>
<td></td>
<td>caMg_I</td>
<td>caMg_IAM</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gender</th>
<th>Case</th>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>Masculine</td>
<td>Direct</td>
<td>caMg_A</td>
<td>caMg_e</td>
</tr>
<tr>
<td></td>
<td>Oblique</td>
<td>caMg_e</td>
<td>caMg_iAM</td>
</tr>
<tr>
<td>Feminine</td>
<td></td>
<td>caMg_I</td>
<td>caMg_IAM</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Reflexives</th>
<th>Punjabi</th>
</tr>
</thead>
<tbody>
<tr>
<td>Singular</td>
<td>myself yourself himself herself itself</td>
<td>Apez ApaNe Apa</td>
</tr>
<tr>
<td>Plural</td>
<td>ourselves yourselves themselves</td>
<td>Apez</td>
</tr>
</tbody>
</table>
2. Reflexive Pronouns: indicates that the person who realizes the action of the verb is the same person who receives the action.

3. Indefinite Pronouns: An indefinite pronoun does not refer to any specific person, thing or amount. It is vague and “not definite”. Some typical indefinite pronouns are ਕੋਈ koI, ਕੁਝ kuJa, ਸਾਰੇ sAre, ਅਨੇਕ aneka, ਬਹੁਤ bahuwa etc. Indefinite pronouns change forms with number and case. Some indefinite pronouns can only be used for singular number, e.g. ਕੋਈ koI, and some are applicable only for plural number, e.g. ਕਈ kaI, ਸਭ saBa. Indefinite pronouns are commonly used in two case forms – direct and oblique. Oblique case form of ਕੋਈ koI is ਕਿਸੇ kise.

Examples:
(a) I was thoroughly ashamed of myself

ਮੈਂ ਆਪਣੇ-ਆਪ ਤੋਂ ਪੂਰੀ ਤਰ੍ਹਾਂ ਸ਼ਰਮਿੰਦਾ ਸੀ

(b) This is for all.

ਇਹ ਸਭ ਲਈ ਹੈ

Punjabi adverbs are unmarked and normally do not inflect. Examples of them are: ਛੇਤੀ nAla (quickly), ਹੌਲੀ hOlZI (slow), etc.

(5) Postpositions in Punjabi

Punjabi Postpositions are similar to prepositions in English. These link noun, pronoun, and phrases to other parts of the sentence. Some Punjabi postpositions are ਤੋਂ tō, ਨੇ ਨ੍ਰੂਮ nē nūm (generally used with objects), and ਤੋਂ ਫੌਮ ਫੌਮ tōm ‘from’ always immediately follow the noun. ਤੋਂ nē and ਨ੍ਰੂਮ nūm lose their tone and become suffixes after certain pronouns, like ਮੈਨੂੰ mainū, ਤੈਨੂੰ tainū, ਹੌਲੂੰ hOlZI, ਇਹਨੂੰ ihnū, ਇਹਨੇ ihnē, ਇਹਨੇ ihnē etc.

(II) Synthesizer for Punjabi

Morphological generator and lexicon plays the major role in development of Machine Translation system for Punjabi. Morphological generation is the reverse of morphological analysis. Morphological analyzer is designed to analyse the constituents of words and it segments the word into stem and inflectional markers. The morphological synthesizer combines a stem with its suffixes based on the syntax of a language. Morphological synthesizer requires all inflected word forms are to be organised into paradigms and which are defined by set of syntactic rules.

(1) Paradigm

A paradigm is a complete set of related word forms associated with a root word. All inflected forms of verb, noun, adjective, adverb and pronoun are classified into certain paradigm class based on their morphological behaviour. For example, ਮੂੰਮਾਦੴਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐائبCd belongs to one paradigm class and ਮੂੰਮਾਦੴਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐਐائبCdbelongs to another paradigm class in noun.
Example for inflection List:
muMdA, muMde, muMde, muMdiAz
kudZI, kudZIAz, kudZI, kudZIAz

In the above example two number markers (singular and plural) and two case markers are the feature values for noun formation. Thus 4 word formations of a root word can be obtained from one paradigm. Any word with similar inflections as that of the word kudZI can be put in to same paradigm class. For example muMdA and GodZA can belong to same paradigm class as both of them have same word formations.

The above inflections can be observed in following examples:

a. ਮੁੰਡਾ ਜਾ ਰਿਹਾ ਹੈ <> muzdA jA rihA hE <> Boy is going.

b. ਮੁੰਡੇ ਜਾ ਰਹੇ ਹਨ <> muzde jA rihe hana <> Boys are going.

c. ਇਹ ਮੁੰਡੇ ਦਾ ਪੈਨ ਹੈ <> iha muMde xA pEna he <> This is boy's pen

d. ਮੈਂ ਉਹਨਾਂ ਮੁੰਡਿਆਂ ਨੂੰ ਬੁਲਾਇਆ <> mEz uhanAz muMdiAz nUM bulAiA <> I called those boys.

Adjectives in Punjabi have variations only for words ending with “_A”, so a single paradigm has been created for two genders. The forms are generated by considering direct/oblique forms along with gender and number information.

Eg: (Masculine) caMgA, caMge, caMge, caMgiAz;

(Feminine) caMgI, caMgIAz, caMgI, caMgIAz

In this way we can generalise the relationship between different roots having same word formation. Verbs and pronouns can also be classified like this by taking features which makes classification easier.

(2) Postpositions in Punjabi

Postpositions in Punjabi are suffixes which occur immediately after nouns and establish some grammatical relations between the nouns and the verbs of sentences. Prepositions in English language will be transformed to post positions in Punjabi. The semantic distribution of a single preposition will be varying in different context due to the influence of nouns and main verbs that follow. Prepositions in English can be directly translated and attached to the Punjabi nouns as postpositions. For Example

a. Radha went with Gopi
   rAXA gopI xe nAla gal
   ਰਾਧਾ ਗੋਪੀ ਨਾਲ ਗਈ

b. The kid is playing with letters.
   baccA aYKarAz nAla Keda rihA hE
   ਬੱਚਾ ਅੱਖਰਾਂ ਨਾਲ ਖੇਡ ਰਿਹਾ ਹੈ

Here, English preposition “with” has been disambiguated with two meanings in Punjabi, “xe nAla” and “nAla”. Hence, suitable addendum has been substituted in Punjabi for “with” by comparing the English sentence with its Punjabi translation. These types of word formations are possible by adding some rules in Text generator.
(III) Summary

Morphological analysis and synthesis play major role in Machine translation. In morphologically rich language like Punjabi word formation is more complex. Morphological analysis of Punjabi gives an idea about suffixes and inflections in word formation of Punjabi. Morphological analysis segments a word and identifies the grammatical information. This information can be transformed into generalized rules for Punjabi Text Generation. Paradigm classification is one of the generalized methods for Text Generation.