A Word and Its Rules

Inner Mongolia University

Nasun-urt

As"deep learning"and"neural network" has become the mainstream technology of natural language processing today, language resources of many small languages in the world are relatively deficient, and they can not meet the processing pattern based on the "big data".In this case, it is necessary to consider the injection of language knowledge to expect the realization of the understanding and processing of language at different levels through "deep learning".Therefore, it is decisive to pay attention to the grammatical and semantic characteristics of different languages and sum up the rules for the natural language processing.

Mongolian language is a typical agglutinate language, and its word formation and configuration are all realized by attaching various supplementary components to the stem.The grammatical meaning of a Mongolian word can only be expressed with phrases or sentences in most western languages and oriental languages, and this kind of changes of the words in real Mongolian text constitute about 82% of all words. If these changes are not taken into account and each word is only understood by its stem meaning, it would be impossible to correctly handle the entire text. As a special case of the language knowledge description, we have summarized the grammar rules of a Mongolian word ᠵᠠᠷᠤᠬᠤ . It is worth mentioning that the rules are only at the lexical change level and its related parts of one semantic item of the word stem , and the other more semantic items and the relevant rules are not exhaustively described.

The Mongolian word "ᠶᠠᠪᠤ go " and its relevant rules

ᠶᠠᠪᠤᠭᠤᠯᠵᠠᠭᠠᠴᠢᠬᠠᠵᠠᠢ (They) have let (all people) go.

---YABV/Ve2+GVL/Fe11+JAGA/Fe5+CIHA/Fi21+JAI/Fs11

---goYABV/Ve2(stem-imperative form-second person)+GVL/Fe11（causative voice）+JAGA/Fe5（multiple voice）+CIHA/Fi21（perfective aspect）+JAI/Fs11（past tense-statement）

ᠶᠠᠪᠤ YABV/Ve2
ᠶᠠᠪᠤᠵᠠᠢ YABV/Ve2+JAI/Fs11
ᠶᠠᠪᠤᠭᠤᠯ YABV/Ve2+GVL/Fe11
ᠶᠠᠪᠤᠭᠤᠯᠵᠠᠢ YABV/Ve2+GVL/Fe11+JAI/Fs11
ᠶᠠᠪᠤᠵᠠᠭᠠ YABV/Ve2+JAG_A/ Fe5
ᠶᠠᠪᠤᠵᠠᠭᠠᠵᠠᠢ YABV/Ve2+JAGA/Fe5+JAI/Fs11

ᠷᡆᠵᠠᠨ᠋ YABV/Ve2+CIH_A/Fi21
ᠷᡆᠵᠠᠨᠠᠷᡁ?YABV/Ve2+CIHA/Fi21+JAI/Fs11
ᠷᡆᠵᠠᠨᠠᠷ᠋ YABV/Ve2+GVL/Fe11+JAG_A/Fe5
ᠷᡆᠵᠠᠨᠠᠷYABV/Ve2+GVL/Fe11+JAGA/Fe5+JAI/Fs11
ᠷᡆᠵᠠᠨᠠᠷ᠋YABV/Ve2+GVL/Fe11+CIH_A/Fi21
ᠷᡆᠵᠠᠨᠠᠷYABV/Ve2+GVL/Fe11+CIHA/Fi21+JAI/Fs11
ᠷᡆᠵᠠᠨᠠᠷ᠋YABV/Ve2+JAGA/Fe5+CIH_A/Fi21
ᠷᡆᠵᠠᠨᠠᠷYABV/Ve2+JAGA/Fe5+CIHA/Fi21+JAI/Fs11
ᠷᡆᠵᠠᠨᠠᠷ᠋YABV/Ve2+GVL/Fe11+JAGA/Fe5+CIH_A/Fi21
ᠷᡆᠵᠠᠨᠠᠷYABV/Ve2+GVL/Fe11+JAGA/Fe5+CIHA/Fi21+JAI/Fs11

The pronouns in co-occurrence with the verb ᠷᡆ
ᠪᡠ ᠨᠠᠮ᠋ ᠪᠢᠳᠠ᠋ BI/Rb11NAM_A/Rb12   BIDE/Rb13
ᠴᡳ ᠴᠢᠮ᠋ ᠲᠠᠨᠠᠷ CI/Rb21CIM_A/Rb22 TANAR/Rb23
ᠲᠠᠷᠠ᠋ ᠲᠠᠭᠦᠨ ᠲᠠᠳᠠᠨᠠᠷ TERE/Rb31TEGUN/Rb32   TEDENER/Rb33

The following is B0 Rule Set (the subject or the agent is the second person singular or plural, or singular plural).
ᠴᡳ ᠲᠠᠨᠠᠷ ᠷᡆ
CI/Rb21|| TANAR/Rb23→YABV/Ve2
ᠴᡳ ᠲᠠᠨᠠᠷ (ᠨᠠᠮ᠋ ᠪᠢᠳᠠ᠋ ᠲᠠᠭᠦᠨ ᠲᠠᠳᠠᠨᠠᠷ) ᠷᡆᠵᠠᠷ
CI/Rb21 || TANAR/Rb23(NAM_A/Rb12 || BIDE/Rb13 || TEGUN/Rb32 || TEDENER/Rb33)
→YABV/Ve2+GVL/Fe11
ᠲᠠᠨᠠᠷ ᠷᡆᠵᠠᠷ᠋
TANAR/Rb23→YABV/Ve2+JAG_A/Fi5
ᠴᡳ ᠲᠠᠨᠠᠷ ᠷᡆᠵᠠᠷ᠋
CI/Rb21 TANAR/Rb23→YABV/Ve2+CIH_A/Fi21
ᠲᠠᠨᠠᠷ (ᠪᠢᠳᠠ᠋ ᠲᠠᠳᠠᠨᠠᠷ) ᠷᡆᠵᠠᠨᠠᠷ᠋
TANAR/Rb23(BIDE/Rb13TEDENER/Rb33) →YABV/Ve2+GVL/Fe11+JAG_A/Fe5
ᠴᡳ ᠲᠠᠨᠠᠷ (ᠨᠠᠮᠪᠢᠳᠠ᠋ · ᠲᠠᠭᠦᠨᠲᠠᠳᠠᠨᠠᠷ) ᠷᡆᠵᠠᠨᠠᠷ᠋
CI/Rb21TANAR/Rb23 (NAM_A/Rb12BIDE/Rb13TEGUN/Rb32TEDENER/Rb33)
→YABV/Ve2+GVL/Fe11+CIH_A/Fi21
ᠲᠠᠨᠠᠷ ᠷᡆᠵᠠᠨᠠᠷ᠋
TANAR/Rb23 YABV/Ve2+JAGA/Fe5+CIH_A/Fi21
ᠲᠠᠨᠠᠷ (ᠪᠢᠳᠠ᠋ ᠲᠠᠳᠠᠨᠠᠷ) ᠷᡆᠵᠠᠨᠠᠷ᠋
TANAR/Rb23(BIDE/Rb13TEDENER/Rb33)→YABV/Ve2+GVL/Fe11+JAGA/Fe5+CIH_A/Fi21

The following is A0 Rule Set

ᠪᡠ ᠪᠢᠳᠠ᠋ ᠴᠢᠨᠠᠷ ᠲᠠᠷᠠᠳᠠᠨᠠᠷ ᠷᡆᠵᠠᠷ
BI/Rb11BIDE/Rb13CI/Rb21TANAR/Rb23TERE/Rb31 TEDENER/Rb33→YABV/Ve2+JAI/Fs11
ᠪᡠ ᠪᠢᠳᠠ᠋ ᠴᠢᠨᠠᠷ ᠲᠠᠷᠠᠳᠠᠨᠠᠷ (ᠨᠠᠮ᠋ ᠪᠢᠳᠠ᠋ ᠴᠢᠮ᠋᠋ ᠲᠠᠨᠠᠷ ᠲᠠᠭᠦᠨ ᠲᠠᠳᠠᠨᠠᠷ) ᠷᡆᠵᠠᠷ
BI/Rb11BIDE/Rb13CI/Rb21TANAR/Rb23TERE/Rb31 TEDENER/Rb33
(NAM_A/Rb12BIDE/Rb13CIM_A/Rb22TANAR/Rb23TEGUN/Rb32TEDENER/Rb33)

→YABV/Ve2+GVL/Fe11+JAI/Fs11

ᠪᠢᠳᠡᠭᠦᠨ ᠲᠠᠨᠠᠷ ᠲᠧᠳᠡᠨᠧᠷ

BIDE/Rb13TANAR/Rb23TEDENER/Rb33→YABV/Ve2+JAGA/Fe5+JAI/Fs11

ᠪᠢ ᠪᠢᠳᠡ ᠴᠢ ᠲᠠᠨᠠᠷ ᠲᠧᠳᠡᠨᠧᠷ

BI/Rb11BIDE/Rb13CI/Rb21TANAR/Rb23TERE/Rb31

TEDENER/Rb33→YABV/Ve2+CIHA/Fi21+JAI/Fs11

ᠪᠢᠳᠡᠭᠦᠨ ᠲᠠᠨᠠᠷ (ᠪᠢᠳᠡ ᠲᠠᠨᠠᠷᠪᠢᠳᠡᠭᠦᠨ) ᠲᠧᠳᠡᠨᠧᠷ

BIDE/Rb13TANAR/Rb23TEDENER/Rb33(BIDE/Rb13TANAR/Rb23TEDENER/Rb33)→YAB

V/Ve2+GVL/Fe11+JAGA/Fe5+JAI/Fs11

ᠪᠢᠳᠡᠭᠦᠨ ᠲᠠᠨᠠᠷ ᠲᠧᠳᠡᠨᠧᠷ

BIDE/Rb13TANAR/Rb23TEDENER/Rb33→YABV/Ve2+JAGA/Fe5+CIH_A/Fi21+JAI/Fs11

ᠪᠢᠳᠡᠭᠦᠨ ᠲᠠᠨᠠᠷ (ᠪᠢᠳᠡ ᠲᠠᠨᠠᠷᠪᠢᠳᠡᠭᠦᠨ)ᠲᠧᠳᠡᠨᠧᠷ

BIDE/Rb13TANAR/Rb23TEDENER/Rb33(BIDE/Rb13TANAR/Rb23TEDENER/Rb33)

→YABV/Ve2+GVL/Fe11+JAGA/Fe5+CIH_A/Fi21+JAI/Fs11

……

      Although these rules appear very complicated, there are certain laws and large coverage.We provide these rules to the computer through the training set of machine learning and other various channels so as to make up the deficiencies brought about by the "sparse data"of a small language, to improve the accuracy of machine learning , and to make the "learning""deeper".