



UDK:519.6:004.6

Vazira BEKOVA,
O'zbekiston Milliy universiteti tayanch doktoranti (PhD),
E-mail: Vazirauz316@mail.com

F.f.d.professor, N.Z.Abdurahmonova taqrizi asosida

MASHINA TARJIMASIDAGI MUAMMOLARNI BARTARAF ETISH VA SIFATLI TARJIMA UCHUN MATEMATIK MODELLAR YARATISH

Аннотация

Ushbu maqolada sifatli mashina tarjimasi sohasidagi dolzarb muammolar tadqiq qilingan va ingliz, o'zbek tillarida ot yasovchi qo'shimchalarning xususiyatlari, vazifalari o'rganildi. So'zlarning turlari bo'yicha gaplarning matematik modellari. Ingliz tilida so'zlarni hosil qilishning umumiy matematik modeli. Gaplarni yasashning matematik modellari keltirib o'tilgan.

Kalit so'zlar: Sifatli mashina tarjimasi, sun'iy intellekt, algoritim, matematik model, affiks, prefiks, so'z qo'shimchalari.

УСТРАНЕНИЕ НЕИСПРАВНОСТЕЙ ПРИ МАШИННОМ ПЕРЕВОДЕ И СОЗДАНИЕ МАТЕМАТИЧЕСКИХ МОДЕЛЕЙ ДЛЯ КАЧЕСТВЕННОГО ПЕРЕВОДА

Аннотация

В данной статье изучены актуальные проблемы в области качественного машинного перевода, а также изучены особенности и функции существительных суффиксов в английском и узбекском языках. Математические модели предложений по типам слов. Общая математическая модель словообразования в английском языке. Приводятся математические модели образования предложений.

Ключевые слова: Качественный машинный перевод, искусственный интеллект, алгоритм, математическая модель, аффиксы, префиксы, суффиксы.

MACHINE TRANSLATION TROUBLESHOOTING AND CREATION OF MATHEMATICAL MODELS FOR QUALITY TRANSLATION

Annotation

This article examines current problems in the field of high-quality machine translation, as well as studies the features and functions of noun suffixes in English and Uzbek languages. Mathematical models of sentences by word type. General mathematical model of word formation in English. Mathematical models of sentence formation are presented.

Keywords: High-quality machine translation, artificial intelligence, algorithm, mathematical model, affixes, prefixes, suffixes.

Kirish. Sifatli mashina tarjimasi hozirgi vaqtda til faoliyatining juda dolzarb muammolaridandir. Chunki ommaviy kommunikatsiyalar hajmi doimiy ravishda kengayib bormoqda, tobora ko'proq yangi tillar muloqotga kirishmoqda va til to'siqlarini samarali bartaraf etishning yagona vositasi tarjimadir. Til to'siqlari axborotni tarqatish yo'lidagi asosiy to'siq bo'lib, ularning asosiy qismi texnik taraqqiyot va sivilizatsiya rivojini belgilab beruvchi ilmiy-texnik ma'lumotlarga tegishli. Til to'siqlarini yengib o'tishning boshqa usullari mavjud, ya'ni yagona til va ta'limning turli darajalarida chet tillarini o'rgatishdan iborat.

Tarjima uchun kompyuterdan foydalanish g'oyasi yangi emas va 70 yildan ortiq vaqtdan buyon mavjud. U deyarli birinchi elektron kompyuterlar - paydo bo'lishi bilan paydo bo'ldi, chunki ilgari kompyuterlar deb atalgan. Kompyuterlarning birinchi ilovalaridan biri xabarlarini kodlash va dekodlash bo'lgan. Birinchi kompyuterlarni ishlab chiquvchilar ularni tarjimada qo'llash g'oyasini quyidagicha ifodaladilar: Bir tabiiy til ikkinchi tilning kodi ekanligini tasavvur qiling. Agar biz ushbu kodni shifrlasak, unda bir tildan ikkinchi tilga tarjima qilishimiz mumkin [1].

Shunday qilib, til qandaydir kodga tenglashtirildi. Shu munosabat bilan aytishimiz mumkinki, tilning birinchi modeli qandaydir kodning modeli biroq, bunday model kompyuter yordamida tarjima qilish muammosini hal qila olmadi, chunki til kod birliklari, agar lingvistik obyektlarni shunday deb atash mumkin bo'lsa, kod birliklaridan farqli o'laroq, polisemitikdir. Shunga qaramay, tilga bo'lgan bunday qarash

nafaqat tarjima nazariyasida, balki amaliy va nazariy tilshunoslikda ham katta rol o'ynaydi.

Hozirgi kunga qadar yaratilgan mashina tarjimalari ko'p tilli va turli sohalar bo'yicha bizga manfaatli til uchun tarjima jarayonini amalga oshiradi. Mashina tarjimasidan biz ilmiy sohada foydalanadigan bo'lsak, albatta bir necha yildan buyon bir ta'qiqqa duch kelamiz bu ham bo'lsa tarjima jarayonida qo'pol so'zlar va kirish tilidagi matn mazmuni, chiqish tilida albatta o'z manosiini yo'qotgan va mazmun jihatdan to'g'ri kelsa-da muammoning asl mazmunini bermayotganini sezishimiz mumkin. Shu mu'ammolarni biz bugungi kunda bartaraf etishda sun'iy intellektdan foydalandik va mashina tarjimasida ancha mazmunli tarjimaga erishdik, lekin hali ham mashina tarjimasidagi muammolar yetarli darajada hal bo'lmaganligini ko'rishimiz mumkin.

Mashina tarjimasining asosi kompyuter lingvistikasi deb qaralyapti lekin har bir tabiiy til (NL) matematik jihatdan tuzilmagan va rasmiylashtirilmagan komponentlardan tashkil topgan murakkab tizimdir. Ю.Н. Марчук [2, 3] kitoblarida matematiklar va tilshunolar tomonidan mashina tarjimasi sohasidagi modelning turli tushunchalarini batafsil yoritib berilgan. U shuningdek, tarjima yozishmalariga asoslangan mashina tarjimasi modellaridan birini tasvirlaydi. [3] muallifning fikriga ko'ra, "bu modelga kiritilgan nazariy tamoyil ma'lum bir til juftligida ishlaydigan tarjimonning harakatlarini takrorlashdir. Tarjimon izchil ravishda iboradan iboraga o'tib, matn mazmuni haqida qandaydir taxminiy fikrni o'z ongida shakllantiradi, so'ngra bu fikrni lingvistik vositalar bilan taqqoslaydi, tarjima ekvivalentlarini tanlaydi va uchta

turdagi tarjima mosligini qidiradi: ekvivalent, variant va transformatsion". Har qanday mashinali tarjima tizimining vazifasi NL da kiritilgan matnning ma'nosini "dekodlash" va uni tarjimon tizimiga tushunarli rasmiylashtirilgan shaklda taqdim etishdir. Keyinchalik, tizim ushbu matnni tarjima qilishi va natijani yetkazishi kerak foydalanuvchi uchun kiritilgan matnga semantik manoli matn kelib chiqishi lozim. Bunga turli NL grammatikasini rasmiylashtirish va aniqlash orqali erishish mumkin modellashtirishga imkon beruvchi tillarning konstruksiyalari Novikov I.A.ning xulosasida ham tasdiqlangan. [4]. NL, xususan, ingliz tili bo'yicha olib borilgan tadqiqotlar [1] NL qoidalariga asoslangan modellar ma'lum mantiqiy formulalar shaklida ma'noni aks ettirishi mumkinligini ko'rsatadi.

Sun'iy intellekt. Bugungi kunda tarjimoni mexanizatsiyalash deb nomlanuvchi mashina tarjimasining rivojlanishi (MT) so'nggi yillarda katta ilmiy qiziqish mavzusiga aylandi. MT tizimlari, birinchi navbatda, sun'iy intellekt va neyron tarmoq texnologiyalarini qo'llash tufayli ajoyib yutuqlarga erishdi neyron mashina tarjimasi (NMT), shubhasiz, tarjimoni yanada qulay va samarali qilgan bo'lsada, ularda bir nechta ilmiy muammolarni ko'rishimiz mumkin.

Bu muammolardan birinchisi sifatli mashina tarjimasiga ta'luqli bo'lsada muammoning to'liq yechimini bermaydi. Sezilarli yaxshilanishlarga qaramay, SMT tizimlari ko'pincha inson tarjimasiga mos tarjimoni keltirib chiqarmaydi. Shu nuqtai nazardan sifatli mashina tarjimasini yaratish uchun harakat ishlari olib borilmoqda. Axloqiy mulohazalar, ayniqsa, xolis yoki mazmun jixatidan qo'pol tarjimalar ehtimolini ko'rishimiz mumkin.

II. So'zlarning turlari bo'yicha matematik modellari

Ingliz tilidagi so'z yasashning leksik tahlili shuni ko'rsatdiki, ingliz tilidagi so'z yasashni quyidagilardan iborat bo'lishi mumkin: predlog, old qo'shimchalar, o'zak, so'z yasovchi qo'shimchalar, shakl yasovchi qo'shimchalar va so'z o'zgartiruvchi qo'shimchalar. Ushbu strukturaga ko'ra, shuningdek ingliz tilining mantiqiy-lingvistik modellari va semantik berilganlar bazalariga ko'ra, ingliz tilida so'zlarni hosil qilishning umumiy matematik modeli quyidagicha ifodalanadi: [1]

$$L1h30(D, T, K(C, P, G, M, N, F), A1, A2, A3) = \downarrow \$[i, 1-52]Di \oplus \downarrow \$[j, 1-96]Tj \oplus \$[i1, 1-h0]Ki1(C, P, G, M, N, F) \oplus \downarrow \$[j1, 1-86]A1j1 \oplus \$[j2, 1-93]A2j2 \oplus \downarrow \$[j3, 1-2]A3j3$$

Bu yerda o'zak so'z (K) gapning istalgan bo'lagi bo'lishi mumkin - ot, sifat, fe'l, olmosh, ravish va son. Ushbu o'zak so'zlarga predlog, prefiks va turli xil qo'shimchalar qo'shilishi so'zlarning shakllanishini beradi. O'zgaruvchi h0 o'zagi bo'lgan so'zlarning umumiy miqdorini, $h30 = \{\min(93h0), \max(79852032h0)\}$ - esa ushbu qonuniyat asosida hosil bo'ladigan so'zlarning minimal va maksimal miqdorini beradi. [1]

III. O'zbek tilidagi gaplarning matematik-modellari va natijasi

Matematik model tizimni matematik izohlash uchun ishlatiluvchi abstrakt model bo'lib, ma'lum bir hodisa va jarayonni matematik formula va bog'lanishlar orqali tushuntirib beradi. Shu jumladan tabiiy tillarni (o'zbek, ingliz) matematik modeli qurishga harakat qildik, chunki tabiiy tillarni bir tildan boshqa tilga tarjima qilishimizda matematik modelning o'rni juda katta ahamiyatga ega. Bu borada M.X. Hakimov [1] monografiyasida so'zlar va gaplar uchun umumiy matematik model keltirilgan shu matematik model asosida gap turlari bo'yicha so'zlarning va gaplarning matematik modellari qurilgan.

Darak gaplarning matematik-modeli

$$\$[h40, 1/h41](C[h40]) \oplus \$[h40, 1/h41](C[h40]) \oplus \$[h42, 1/h43](G[h42])$$

$$\begin{aligned} & \$[h40, 1/h41](C[h40]) \oplus \$[h42, 1/h43](G[h42]) \oplus \\ & \$[h42, 1/h43](G[h42]) \oplus \$[h37, 1/26](F[h37]) \oplus \\ & \$[h40, 1/h41](C[h40]) \oplus \$[h40, 1/h41](C[h40]) \oplus \\ & \$[h42, 1/h43](G[h42]) \oplus \\ & \$[h39, 1/h28](N[h39]) \oplus \$[h40, 1/h41](C[h40]) \oplus \\ & \$[h44, 1/h45](P[h44]) \oplus \$[h42, 1/h43](G[h42]) \oplus \\ & \$[h40, 1/h41](C[h40]) \oplus \$[h40, 1/h41](C[h40]) \oplus \\ & \$[h40, 1/h41](C[h40]) \oplus \$[h42, 1/h43](G[h42]) \oplus \\ & \$[h40, 1/h41](C[h40]) \oplus \$[h39, 1/h28](N[h39]) \oplus \\ & \$[h42, 1/h43](G[h42]) \oplus \$[j, 1/48](D[j]) \oplus \\ & \$[h40, 1/h41](C[h40]) \oplus \$[h44, 1/h45](P[h44]) \oplus \\ & \$[h40, 1/h41](C[h40]) \oplus \$[h42, 1/h43](G[h42]) \oplus \\ & \$[h42, 1/h43](G[h42]) \oplus \$[h40, 1/h41](C[h40]) \oplus \\ & \$[h42, 1/h43](G[h42]) \oplus \\ & \$[h40, 1/h41](C[h40]) \oplus \$[h44, 1/h45](P[h44]) \oplus \\ & \$[h40, 1/h41](C[h40]) \oplus \$[h42, 1/h43](G[h42]) \oplus \\ & \$[i, 1/h13](Y[i]) \oplus \$[h40, 1/h41](C[h40]) \oplus \\ & \$[h40, 1/h41](C[h40]) \oplus \$[h42, 1/h43](G[h42]) \oplus \\ & \$[h39, 1/h28](N[h39]) \oplus \$[h39, 1/h28](N[h39]) \oplus \\ & \$[h40, 1/h41](C[h40]) \oplus \$[i, 1/h13](Y[i]) \oplus \\ & \$[h40, 1/h41](C[h40]) \oplus \$[h39, 1/h28](N[h39]) \oplus \\ & \$[h40, 1/h41](C[h40]) \oplus \$[h42, 1/h43](G[h42]) \oplus \\ & \$[i, 1/h13](Y[i]) \oplus \$[h40, 1/h41](C[h40]) \oplus \\ & \$[h39, 1/h28](N[h39]) \oplus \$[h42, 1/h43](G[h42]) \oplus \\ & \$[h40, 1/h41](C[h40]) \oplus \$[h42, 1/h43](G[h42]) \oplus \\ & \$[h40, 1/h41](C[h40]) \oplus \$[h44, 1/h45](P[h44]) \oplus \\ & \$[h42, 1/h43](G[h42]) \oplus \$[h44, 1/h45](P[h44]) \oplus \\ & \$[h40, 1/h41](C[h40]) \oplus \$[h42, 1/h43](G[h42]) \oplus \end{aligned}$$

Yuqoridagi matematik-model bo'yicha har bitta algoritm uchun quyidagi misollarni keltirishimiz mumkin:

Rasul yong'oqni yoqtirmaydi.

Alisher ishtirok etishdan bosh tortdi.

Anora to'rtta samolyot chiptasini sotib oldi.

Kecha Zilola omadli edi.

Anvar bolalar bog'chasini tark etadi.

Odil bugun xarid qilish uchun ketdi.

Karim baxtsiz hodisaga uchradi.

O'tkaziladigan joy yopiq qoladi.

Dengiz tinch.

Zuhra hikoya qiladi va Ramazon o'yinchoqlar yig'adi.

Har kuni ertalab Hilola va Doston birga nonushta qilishadi, lekin sukunat asta-sekin davom etadi.

Pardani yopib qo'yganimga shamol qattiq esdi, baland ovoz eshitildi.

Qo'shimcha yasashning matematik modeli:

Qo'shimcha yasashda o'zak so'z ot, son, ergash gap va sifatosh bo'lishi mumkin. Bu radikallarga qo'shilganda zarrachalar, old qo'shimchalar, qo'shimchalar va oxirlardan iborat so'zlar qo'shimcha hosil qiladi. Ikki ergash gap birga kelgan hollarda yangi ergash gap yasaladi. Matematik olti turdagi mantiqiy-lingvistik modellar asosida qo'shimchalarni xulosa qilish modeli, ba'zi o'zgarishlardan so'ng, quyidagicha ifodalanadi:

- 1) noun-forming : (Ot yasovchi) -er, -dom, -ness, -ation, -ity, -age, -ance, -ence, -ist, -hood, -ship, -ment etc;
- 2) adjective-forming (Sifat yasovchi) -able, -ible, -uble, -al, -ian, -ese, -ate, -ed, -ful, -ive, etc;
- 3) numeral-forming (Son yasovchi) -teen, -th, -ty etc;
- 4) verb-forming (fe'l yasovchi) -ate, -en, -ify, -ize etc;
- 5) adverb-forming (Ergash gap yasovchi) -ly, -ward, -wise etc.

Ch31(C(T),K(C),C(S),C(O),K(P),K(G),K(M),U,W) =
 $(\downarrow \$[i,1-16]C(Ti) \oplus \$[j,1-h1]K(Cj) \oplus \$[i,1-120]C(Si) \oplus$
 $\downarrow \$[j,1-120]C(Sj1) \oplus \downarrow \$[i,1-28]C(Oi2)) \vee (\$[j,1-h2]K(Pj2) \oplus$
 $\$[i,3-1-h1]K(Ci3) \oplus \$[i,4-1-120]C(Si4) \oplus \downarrow \$[i,1-28]C(Oi5)) \vee$
 $(\$[i,1-h1]K(Ci) \oplus \$[j,1-28]C(Oj)) \vee (\downarrow \$[j,1-h2]K(Pj2) \oplus$
 $\$[j,3-1-h1]K(Cj3) \oplus \downarrow \$[j,4-1-h3]K(Gj4) \oplus \downarrow \$[i,1-120]C(Si) \oplus$
 $\downarrow \$[i,1-28]C(Oi1)) \vee (\$[i,1-h5]K(Mi2) \oplus \$[i,3-1-h3]K(Gi3)) \vee$
 $(\$[i,1-85]Ui \oplus \$[j,1-h1]K(Cj)) \vee (\$[j,1-h1]K(Cj2) \oplus$
 $\$[i,1-2]Wi1 \oplus \$[i,2-1-h1]K(Ci2))$

Sifatli mashina tarjimasini yaratishda bugungi kun mu'ammolaridan biri bu ikki tildagi so'z qo'shimchalarini moslash algoritimi va matematik modelini yaratishdan iborat. Qo'shimchalar ingliz va o'zbek tillarida so'z yasashning unumli usuli hisoblanadi. Ingliz tili hind-yevropa tillari oilasiga, o'zbek tili esa oltoy tillariga mansub. Affiksatsiya barcha tillarda hosila affikslar yordamida so'z yasash usullaridan biridir. Affikslar bog'langan morfemaning bir turi bo'lib, ular yakka holda turolmaydi va o'z ma'nosini olish uchun asosiy so'z bilan birga kelishi kerak. Ingliz tili, o'zbek tili grammatikasi va morfologiyasida affiksatsiya so'zga morfema- yoki affiks qo'shib, shu so'zning boshqa shaklini yoki boshqa ma'noli yangi so'z hosil qilish; affiksatsiya eng

keng tarqalgan solishtirma tillarda yangi so'z yasash usuli. Affiksning ikki asosiy turi - old qo'shimcha, qo'shimcha qo'shilish va qo'shimcha, qo'shimcha qo'shilish, affikslarning klasterlari esa murakkab so'zlarni hosil qilish uchun ishlatilishi mumkin. Affiks - a grammatikaning so'z elementi so'zning ma'nosi yoki shaklini o'zgartirish uchun ishlatiladi va prefiks yoki qo'shimcha shaklida keladi. Prefikslar "un-", "self-" va "re-" kabi misollarni o'z ichiga oladi, qo'shimchalar esa ingliz tilida "-hood", "-ing" yoki "-ed" kabi tugatish elementlari shaklida keladi. Asosiy qiyoslangan tillarda qo'shimchalarning vazifasi nutqning bir qismini boshqasidan hosil qilish, ikkinchi darajali vazifasi esa lug'aviy ma'noni o'zgartirishdan iborat. Affikslar keng umumlashgan ma'noga ega bo'lib, butun so'z tomonidan ma'lum bir turkumga yetkaziladigan tushunchadir. Affiksatsiya prefiks va qo'shimchalarga bo'linadi. So'z o'zagidan oldin qo'shilgan hosila morfemalarga old qo'shimchalar, so'z o'zagidan keyin qo'shilgan hosila morfemalar esa qo'shimchalar deyiladi. Ingliz, o'zbek tillarida qo'shimchalar turli gap bo'laklarini hosil qiladi. Bu tillarda ot yasovchi qo'shimchalar juda ko'p.

Xulosa. Ushbu maqolada ingliz, o'zbek tillarida ot yasovchi qo'shimchalarning xususiyatlari, vazifalari o'rganildi va ingliz tili bilan o'zbek tilidagi so'z yasovchi qo'shimchalar, predlog va iboralarni moslash algoritimi va matematik modelini yaratish bo'yicha kelgusi ishlarimizda keltirib o'tamiz. Bu ikki til o'rtasida ilmiy matinni o'z ma'no mazmunini yo'qotmagan xolda tarjima qilishimiz ya'ni sifatli mashina tarjimasiga erishishimiz uchun albatta kirish tilidagi inglizcha matn chiqish tilidagi sifatli o'zbekcha matnga aylanishi uchun biz shu yuqorida keltrib o'tilgan usulda algaritim va matematik modelini yaratishimiz lozim.

ADABIYOTLAR

1. Хакимов М.Х. Технология моноязыкового моделируемого компьютерного переводчика Тошкент 2009.
2. Дудников А.В. Современный русский язык. М. «Высшая школа», 1990. 424 с. 2. Марчук Ю.Н. Модели перевода. М., Издательский центр «Академия», 2010. 176 с.
3. Марчук Ю.Н. Компьютерная лингвистика. М., АСТ: Восток-Запад, 2007. 318 с.
4. Новиков И.А. Семантика текста и ее формализация. М. «Наука», 1983. 216 с.
5. Хакимов М.Х. Расширяемый входной язык математического моделирования естественного языка для многоязычной ситуации машинного перевода. ЎзМУ хабарлари. № 1, 2009. С. 75-80.
6. Хакимов М.Х. Математические модели узбекского языка. ЎзМУ хабарлари. № 3, 2010. С. 187-191.
7. Better Translation for Better Communication 1983. Machine Translated by Google