**ӘОЖ 004.912**

**МРНТИ**

**М.А.Самбетбаева1, Г.Қалман2**

16D070300 –Ақпараттық жүйелер мамандығы бойынша философия докторы (PhD)

Л.Н.Гумилев атындағы Еуразия ұлттық университеті, Қазақстан.

21-курс Доктарант

Л.Н.Гумилев атындағы Еуразия ұлттық университеті, Қазақстан

(guljamal14@gmail.com )

**Табиғи тілге арналған мультиагенттік жүйе**

**Андатпа**

Бұл мақалада мультиагенттік жүйенің әдістерін қолдана отырып, табиғи тілді өңдеуге арналған қазіргі зерттеу жобасы сипатталған.

Табиғи тілдерді өңдеу (NLP) көптеген қосымшаларға ие, мысалы, мәліметтер базасының интерфейсі, машиналық аударма. Білім алуды және есеп беруды абстракция. Табиғи тілдермен жұмыс жасауда бірнеше тәсіл қолданылады. Бұл мақалада көп агентті жүйенің әдістерін қолдана отырып, табиғи тілді енгізуді түсінуге арналған қазіргі зерттеу жобасы сипатталған. Жүйе лексикалық-құрылымдық тәсіл және танымдық-құрылымдық тәсілді біріктіре өтырп жасалған. Лексикалық құрылымда барлық лингвистикалық ақпараты бар агенттер ретінде қарастырылған. Танымдық құрылым 6 модульден тұрады: Сөйлеу-мәтін, мәтін-сөйлеу, морфологиялық,семантикалық, дискурсивті және сұрау анализаторлары. Бұл агенттер пайдаланушының кірісін білдіретін агенттердің ішкі жиындарын құру үшін бір-бірімен өзара әрекеттеседі. Сұрау анализаторының агенті пайдаланушының сұрақтарына жауап беру үшін кіріс агенттерінің қосалқы қауымдастықтарымен өзара әрекеттесетін қосалқы агенттер қауымдастығын жасайды. Мақала қысқаша кіріспеден басталады көп агентті жүйелер және табиғи тілді өңдеу. Жобаның сипаттамалары, оның әртүрлі модульдері және әр модульдің жұмыс істеу барысы көрсетілген.

*Түйін сөздер*: агенттер, мультиагентті жүйе, морфологиялық талдау, семантикалық талдау.

*Кіріспе*

Табиғи тілді өңдеудің негізгі мақсаты - компьютерді пайдаланушы енгізген деректерді түсіну. Түсіну - бір көріністен (пайдаланушы енгізуінен) екіншісіне (ішкі көрініс) ауысу ретінде анықталуы мүмкін. Бұл трансформация әр түрлі кезеңдерді қамтиды: сөйлеуді мәтінге айналдыру, морфологиялық талдау, синтаксистік талдау, семантикалық талдау және прагматикалық талдау. Осы кезеңдердің соңында кіріс сөйлемдерінің компьютерлік семантикалық түсіндірмесі жасалады. Содан кейін бұл интерпретацияны әртүрлі қосымшалар қолдана алады. Мұндай қосымшаларға мысалы, машиналық аударма, мәліметтер базасының интерфейсі және сұрақтарға жауап беру жүйесі жатады. Көптеген факторлар түсіну мәселесінің күрделенуіне ықпал етеді. Картография түрі- негізгі фактор болып табылады.

Бірінші кезең-сөйлеуді мәтінге түрлендіру, ол пайдаланушының сөзін кіріс мәтініне түрлендіреді. Табиғи тілді өңдеу жүйесінің екінші кезеңі-пайдаланушы (немесе синтаксистік анализатор) мұнда сөздердің реттілігі, сөздердің бір-бірімен қалай байланысты екенін көрсететін құрылымға айналады. Сөздік немесе лексика пайдаланушы енгізген деректердегі сөздерді анықтау үшін қолданылады. Сөздік көлемін азайту үшін тек сөздердің түбірлері сақталады.

Морфологиялық анализатор кіріс мәтіндік сөздерді, сөздік сөздерге жасыру үшін қолданылады. Сондықтан морфологиялық талдау, сөздік көлемін азайтуда маңызды болып саналады.

Табиғи тілді өңдеудің үшінші кезеңі- семантикалық талдау. Бұл кезеңде, кіріс сөйлем құрылымының семантикалық көрінісі құрылады. Әрі қарай, дискурстық талдау жүргізіледі. Пайдаланушының енгізілуіне жауап- дискурстың семантикалық көрінісі негізінде жасалады. Нәтижесінде мәтінді сөйлеуге түрлендіру модулінің көмегімен, сойлеуге түрленеді.

Мультиагентті жүйелер технологиясы - бағдарламалық жүйелерді жобалауды енгізудің жаңа парадигмасы болып жарияланды. Агенттер - бұл өз пайдаланушыларының атынан дербес әрекет ететін күрделі компьютерлік бағдарламалар. Бұл ашық және таратылған ортада күрделі мәселелерді шешуге арналған [2]. Табиғи тілді өңдеу, бірге жұмыс істей алатын бірнеше агенттерді қажет етеді. Мультиагенттік жүйе - бұл әр мәселені шешушінің жеке мүмкіндіктерінен немесе білімінен тыс мәселелерді шешу үшін өзара әрекеттесетін, бағдарламалық жасақтама агенттерінің байланысқан желісі.

*Жүйе туралы түсінік*

Жүйені түсіну үшін пайдаланушы енгізген деректерді (сөйлем немесе абзац) сөйлеу түрінде қабылдайды. Содан кейін пайдаланушы пайдаланушының енгізуіне қатысты бірнеше сұрақтарды енгізе алады. Жүйе бұл сұрақтарға жауап беруі керек. Бұл кіріс туралы түсініктерді көрсетеді. 1-суретте мультиагент түсіну жүйесінің схемасы. Жүйе 8 модульден тұрады. Лексикалық құрылымдық модульі ( лексикалық агенттердің білім базасы), 6 танымдық құрылымдық модуль (сөйлеу-мәтін, морфологиялық, семантикалық, дискурсивті және сұрау анализаторлары) және пайдаланушы интерфейсі. Тезис модульдерінің толық сипаттамасы көрсетілген.



1-сурет-мультиагенті жүйесі

*Пайдаланушы интерфейсі*

Пайдаланушы интерфейсі - түсіну жүйесі мен пайдаланушы арасындағы байланысты жеңілдету үшін қажет. Пайдаланушы жүйемен сөйлесіп жатқанда немесе сұрақтар қойса, жүйе сұрақтарға жауап береді, немесе қателер туралы хабарлайды. Тиісті мәтіндер төртінші терезеде көрсетіледі. Бұл терезелер-пайдаланушы кірісін көрсетуге арналған енгізу терезесі, пайдаланушы сұрақтарын көрсетуге арналған сұрау терезесі, жүйелік жауапты көрсетуге арналған жауап терезесі және қате туралы хабарламаларды көрсетуге арналған қате терезесі. Жүйеде әр түрлі қателерді анықтауы керек, олар танылмаған сөзді, грамматикалық сөйлем қатесін, семантикалық сөйлем қатесін, шешілмеген түсініксіз сұрақты қамтуы мүмкін. Бұл қателер пайдаланушыға ұсынылады. Кейбір сұрақтарға жауап келесідей болуы мүмкін: "жауап беруге болатын ақпарат жоқ". Осы терезелерден басқа, пайдаланушыға әр агенттің шығысын мультиагенттік жүйеде көрсетуге мүмкіндік беретін ашылмалы мәзір бар. Осылайша, пайдаланушы әр агентті бақылай алады.

*Сөйлеуді мәтінге айналдыру агенті*

Пайдаланушы сөйлемді, абзацты немесе сұрақты сөйлеу түрінде айта алады. Сөйлеуді мәтінге түрлендіру агенті кіріс деректерін мәтінге түрлендіреді. Түрлендірілген мәтін кіріс терезесінде немесе сұрау терезесінде көрсетіледі. Сонымен қатар, пайдаланушы сөйлемді, абзацты немесе сұрақты мәтін түрінде енгізе алады. Бұл агент әр пайдаланушы үшін дауыстық профильді қолдайды.

*Морфологиялық талдау агенті*

Морфологиялық талдау агенті: кіріс сөйлемін негізгі формадағы сөздер тобына (етістіктер, зат есімдер, сын есімдер, үстеулер және т.б.) түрлендіреді. Сонымен қатар, бұл агент олардың лингвистикалық ақпаратын жинайды (түрі, уақыты,) . Бұған когнитивті агент (морфологиялық анализатор) және лексикалық агенттердің (префикстер мен жұрнақтар) өзара әрекеттесуі арқылы қол жеткізіледі. Морфологиялық талдаушы агент - сөздердің негізгі нысандарын сәйкестендіру, олардың лингвистикалық ақпаратын алу және талдау жүргізу үшін аффикстер агенттерін пайдалану үшін білім базасына (сөздік агенттеріне) жүгінеді. Бір сөз үшін морфологиялық анализатор агенті бірнеше негізгі формаларды бөліп алып, оларды бағалай алады.

*Семантикалық талдау агенті*

Сөйлемге арналған сөздер тобын және олардың лингвистикалық ақпаратын ескере отырып, бұл агент басқа агенттерімен (зат есімдер мен етістіктер) өзара әрекеттеседі . Зат есімдер мен зат есімдердің атрибуттарын өңдеуде сын есімдер, үстеулер, есімдіктер және т.б. қолданылады.

*Білім базасы*

Сөздік агенттер лексикалық құрылымда ұсынылған. Онда агент ретінде табиғи тіл лексикасы және олар туралы барлық лингвистикалық ақпарат бар. Білім базасында агенттердің көрінісі атрибуттарды көрсетеді.

*Дискурстық талдау агенті*

Көптеген ұсыныстар агенттерінің қосалқы қауымдастықтарын ескере отырып, дискурстық анализатор агенті сөйлемдер арасындағы айырмашылықтарды шешеді. Бұл осы сөйлемдердегі есімдіктер мен сөйлемдер негізінде жасалады.

*Сұрау тілін тандаушы агент*

Бұл білім базасындағы "сұраулы сөздерінің" агенттерімен және мәселені шешу агенттердің қосалқы қауымдастығын құру арқылы жасалады. Сұрау агенті жауаптарды алады және оларды қолданушыға ұсынады.

*Мәтінді сөйлеуге түрлендіру агенті*

Жүйе пайдаланушының сұрақтарына жауаптарды немесе қате туралы хабарламаларды көрсетеді. Бұл тұжырымдардың екеуі де мәтін немесе сөйлеу түрінде ұсынылған. Мәтінді сөйлеуге түрлендіру агенті - шығыс мәтінді сөйлеуге түрлендіреді.

*Агенттер арасындағы өзара байланыс*

Табиғи тілді түсінудің мультиагентті жүйесі оның танымдық құрылымдық модульдері (морфологиялық, семантикалық, дискурсивті және сұрау анализаторлары) мен лексикалық құрылымдық модуль (лексика) агенттер арасындағы өзара әрекеттесуді көрсетеді. Cемантикалық агент сөйлемдегі етістіктерді анықтау үшін етістік агенттерімен өзара әрекеттеседі. Негізгі етістік агент атаулы агенттермен өзара әрекеттеседіде, олар пайдаланушының ішкі қауымдастығына ұсынады. Дискурсивті агент әр түрлі сөйлемдерге қабілетті болу үшін есімдік пен пропозициямен байланысады. Содан кейін үш координаталық, бағыныңқы және екілік қатынастар агенттері толық абзацты құрайтын ішкі қатынастардың құрылымын жасайды. Сұранысты талдау агенті сұрау агентін құру үшін пайдаланушының ішкі қауымдастығымен өзара әрекеттеседі.

Бұл ішкі қауымдастық дұрыс жауапты алу үшін пайдаланушы енгізу ішкі қауымдастықтарының құрылымымен өзара әрекеттеседі.



2-сурет агенттер арасындағы байланыс

*Қорытынды*

Табиғи тілді түсінуді, тілдегі мәтіннің мағынасын көрсететін интемдік көрініске бейнелеу ретінде қарастыруға болады. Бұл мақалада мультиагентті жүйені қолдана отырып, жүйені түсінуге арналған қазіргі зерттеу жобасы ұсынылған. Жүйе лексикалық-құрылымдық тәсіл мен танымдық-құрылымдық тәсілді біріктіреді. Танымдық құрылым 6 модульден тұрады: сөйлеу, мәтін, морфологиялық, семантикалық, дискурсивті және сұрақ анализаторлары. Бұл агенттер пайдаланушы енгізуін білдіретін агенттердің ішкі жиындарын құру үшін бір-бірімен өзара әрекеттеседі. Сұрау анализаторының агенті пайдаланушыға жауап беру үшін, кіріс агенттерінің қосалқы қауымдастықтарымен өзара әрекеттесетін, қосалқы агенттер қауымдастығын жасайды.

 Бұл жүйенің кейбір модульдері жүйе ретінде талдау жасалынып іске асырылды. Жеке қосымша ретінде ағымдағы жұмыс пайдаланушы интерфейсі, мен жүйелік модульдер арасындағы байланысқа бағытталған.

**ПАЙДАЛАНЫЛҒАН ӘДЕБИЕТТЕР ТІЗІМІ**

1. Processing," in Advances in Artificial Intelligence: 12th Brazilian Symposium on Artificial Intelligence, J. Wainer & A. Carvalho Pfeirer (Editors), Springer-Verlag, Berlin - Heidelberg, pp. 312-322. 1995.
2. http://www-2.cs.cmu.edu/-sofiagents/multi.html, The Intelligent Software Agents Lab, Camegie Mellon Univeristy.
3. V. Pallotta, "A logical framework for designing robust distributed NLP applicatons," EUROLAN 2001, Intemational Workshop on Modular Programming for NLP, Iasi, Romania, 2001.
4. Репьёв А.В. Адаптивная система обучения работников станций. // Научная сессия МИФИ-2006. Сборник научных трудов. М.:МИФИ, 2006. Т.З. С. 170-171.
5. Luck M., McBurney P., Preist C. Agent technology: Enabling next generation computing a roadmap for agent based computing. - Southampton, UK: AgentLink, 2003.
6. Иванов C.H., Кудинов B.A. Об архитектуре компьютерных обучающих систем на основе мультиагентных систем // Ученые записки: Электронный научный журнал Курского государственного университета. 2010. – №1.
7. L. Wollin, "Microsoft Agent Part 1: Animating Office Applications with Microsoft Agent," Microsoft Corporation, April 2001.
8. S. Walczak, "Knowledge Acquisition and Knowledge Representation with Class: the Object-Oriented Paradigm," Expert Systems with Applications, 15, pp. 235-244. 1998.
9. M. Aref & S. Desiraju, "An Object-Oriented Approach for Discourse Analysis," Proceedings of the 15th International Conference on Computer Applications in Industry and Engineering, San Diego, California, PP 7-11, 2003.

**REFERENCES**

1. 1Processing," in Advances in Artificial Intelligence: 12th Brazilian Symposium on Artificial Intelligence, J. Wainer & A. Carvalho Pfeirer (Editors), Springer-Verlag, Berlin - Heidelberg, pp. 312-322. 1995.
2. http://www-2.cs.cmu.edu/-sofiagents/multi.html, The Intelligent Software Agents Lab, Camegie Mellon Univeristy.
3. V. Pallotta, "A logical framework for designing robust distributed NLP applicatons," EUROLAN 2001, Intemational Workshop on Modular Programming for NLP, Iasi, Romania, 2001.
4. Repyev A.V. Adaptive training system for station workers. // Scientific session of MEPHI-2006. Collection of scientific works. Moscow: MEPHI, 2006. T. Z. S. 170-171.
5. Luck M., McBurney P., Preist C. Agent technology: Enabling next generation computing a roadmap for agent based computing. - Southampton, UK: AgentLink, 2003.
6. Ivanov C. H., Kudinov B. A. About the architecture of computer training systems based on multi-agent systems // Uchenye zapiski: Electronic Scientific Journal of Kursk State University. 2010. – №1.
7. L. Wollin, "Microsoft Agent Part 1: Animating Office Applications with Microsoft Agent," Microsoft Corporation, April 2001.
8. S. Walczak, "Knowledge Acquisition and Knowledge Representation with Class: the Object-Oriented Paradigm," Expert Systems with Applications, 15, pp. 235-244. 1998.
9. M. Aref & S. Desiraju, "An Object-Oriented Approach for Discourse Analysis," Proceedings of the 15th International Conference on Computer Applications in Industry and Engineering, San Diego, California, PP 7-11, 2003.

**Г.Қалман1, М.А.Самбетбаева2**

ЕНУ им.Л.Гумилева докторант 1 курса, г. Нур-Султан, Республика Казахстан.

(guljamal14@gmail.com )

доктор философии (PhD) по специальности 6D070300 – Информационные системы, г. Нур-Султан, Республика Казахстан

**Мультиагентная система для естественного языка**

Ключевые слова: агенты, мультиагентная система, морфологический анализ, семантический анализ.

**G.Kalman1,M.A.Sambetbayeva2**

1L. Gumilyov ENU 1st year doctoral student, Nur-Sultan, Republic of Kazakhstan

2Doctor of philosophy (PhD) in 6D070300 –Information Systems, L. Gumilyov ENU, Nur-Sultan, Republic of Kazakhstan

**A Multi-Agent System for Natural Language Understanding**

Keywords: agents, multi-agent system, morphological analysis, semantic analysis.

**Мақала авторлары туралы ақпарат:**

**М.А.Самбетбаева-**6D070300 –Ақпараттық жүйелер мамандығы бойынша философия докторы (PhD), Л.Н.Гумилев атындағы Еуразия ұлттық университеті, Нұр-Сұлтан қ., Қазақстан.

**М.А.Самбетбаева-** доктор философии (PhD) по специальности 6D070300 – Информационные системы, г. Нур-Султан, Республика Казахстан.

**M.A.Sambetbayeva-**Doctor of philosophy (PhD) in 6D070300 –Information Systems, L. Gumilyov ENU, Nur-Sultan, Republic of Kazakhstan.

**Қалман.Г**- 1-курс Доктарант Л.Н.Гумилев атындағы Еуразия ұлттық университеті, , Нұр-Сұлтан қ., Қазақстан.

**Калман.Г**- ЕНУ им.Л.Гумилева докторант 1 курса, г. Нур-Султан, Республика Казахстан.

**Kalman.G**-L. Gumilyov ENU 1st year doctoral student, Nur-Sultan, Republic of Kazakhstan.