

18. Kureychik V.V., Kureychik V.I. Bionicheskiy poisk pri proektirovanii i upravlenii [Search inspired by natural systems, for the design and management], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2012, No. 11 (136), pp. 178-183.
19. Zaporozhets D.Yu., Zaruba D.V., Kureichik V.V. Hybrid bionic algorithms for solving problems of parametric optimization, *World Applied Sciences Journal*, 2013, No. 23 (8), pp. 1032-1036.
20. Kureychik V.V., Zaporozhets D.Yu. Rovey algorithm v zadachakh optimizatsii [Swarm algorithm in optimisation problems], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2010, No. 7 (108), pp. 28-32.
21. Sandro V.P., Jyrko C.M., José R.S. Weighted Partition Consensus via Kernels, *Pattern Recognition*, 2010, Vol. 43(8), pp. 2712-2724.
22. Karpov V.E. Vvedenie v rasparallelvanie algoritmov i programm [Introduction to parallel algorithms and software], *Komp'yuternye issledovanie i modelirovanie* [Computer research and modeling], 2010, Vol. 2, No. 3, pp. 231-272.

Статью рекомендовал к опубликованию д.т.н., профессор А.Г. Коробейников.

Кравченко Юрий Алексеевич – Южный федеральный университет; e-mail: yakravchenko@sfedu.ru; 347928, г. Таганрог, пер. Некрасовский, 44; тел.: 88634371651; кафедра систем автоматизированного проектирования; доцент.

Нацкевич Александр Николаевич – e-mail: natskevich.a.n@gmail.com; кафедра систем автоматизированного проектирования; аспирант.

Kravchenko Yuriy Alekseevich – Southern Federal University; e-mail: yakravchenko@sfedu.ru; 44, Nekrasovskiy lane, Taganrog, 347928, Russia; phone: +78634371651; the department of computer aided design; associate professor.

Natskevich Alexander Nikolaevich – e-mail: natskevich.a.n@gmail.com; the department of computer aided design; graduate student.

УДК 004.912

В.С. Корнилов, В.М. Глушань, А.Ю. Лозовой

ОЦЕНКА КАЧЕСТВА МАШИННОГО ПЕРЕВОДА ТЕКСТА С ИСПОЛЬЗОВАНИЕМ МЕТОДА АНАЛИЗА НЕЧЕТКИХ ДУБЛИКАТОВ*

Статья посвящена разработке способов оценки качества и улучшения результатов машинного перевода. Машинный перевод рассматривается как полностью автоматический перевод текста на основе правил. Существующие методы и средства машинного перевода имеют как преимущества так и недостатки, заключающиеся в потере семантической целостности при переводе одного и того же текста с одного естественного языка на другой естественный язык, в итоге результат перевода в большинстве случаев некорректен. Целью работы является создание системы автоматической корректировки машинного перевода, результатом работы которой будет текст на уровне публикации. Научная новизна заключается в использовании процедуры получения обратного перевода, его сравнения с оригинальным текстом для численной оценки качества машинного перевода, а также поиска несоответствий с применением системы выявления нечетких дубликатов в параллельных корпусах и последующей их корректировки. В настоящее время существует широкий набор методов и приложений для оценки машинного и автоматизированного переводов с использованием обработки параллельных корпусов. Недостатком данных методов является невозможность отследить ошибки в конкретном параллельном корпусе. Существует широкий набор методов для анализа совпадений в неструктурированном тексте, применяемых для поиска плагиата в различных сферах. Классические методы анализа текста подразделяются на синтаксические методы анализа последовательностей, со-

* Работа выполнена при финансовой поддержке РФФИ (проект № 15–01–05669).

стоящих из символов, слов, словосочетаний или предложений и методы лексического анализа и определения контрольных сумм слов, словосочетаний, предложений и абзацев. Данные методы рассмотрены на предмет возможности их использования в системе автоматической корректировки машинного перевода текстов. Проанализированы перспективы использования данного метода.

Машинный перевод; качество перевода; обратимость перевода; поиск плагиата; нечеткий дубликат; автоматизированное редактирование перевода.

V.S. Kornilov, V.M. Glushan, A.Y. Lozovoy

ASSESSMENT OF THE QUALITY OF MACHINE TRANSLATION TEXT WITH USE A METHOD OF ANALYSIS OF FUZZY DUPLICATES

This article is devoted to problems of evaluation of machine translation quality. Modern methods and means of fully automatic machine translation is far from perfect, because when you translate a text from one natural language to another natural language, the semantic component is often lost in the translation and the result is often incorrect. The aim of the authors is to develop applications for automatic correction of machine translation to the level of publication. Currently, there are many methods and applications for evaluation of machine and automated translation involving the processing of parallel corpora. The disadvantage of these methods is the inability to track errors in a specific parallel corpus. The authors propose to use the results of reverse translation to evaluate the machine translation quality and to search the discrepancies, and comparing it with the original text with use the system identification of fuzzy duplicates, to search for inconsistencies in the parallel corpora and the subsequent adjustments. The modern methods of analysis of fuzzy duplicates in unstructured text that used to search for plagiarism are described. Classical methods of analysis of the text are divided into syntactic methods of analysis of sequences, consisting of characters, words, phrases or sentences and methods of lexical analysis and definition of the checksum words, phrases, sentences and paragraphs. These methods are reviewed for their use in the system of adjustments of machine translation. The novelty of this work is to develop schemes of work machine translator with automatic correction of discrepancies that identified by the described method. The prospects of using this method are analyzed.

Machine translation; quality of translation; the reverse translation; search for plagiarism; fuzzy duplicate; computer-aided editing of translations.

Введение. Проблема качества перевода текстов была интересна людям с самых первых дней существования перевода и речи в целом. Требования к переводу и оценке его качества, возникают еще в античные времена, а также у многих писателей и философов, обращавшихся к проблемам перевода [1]. Несмотря на более чем полувековые усилия современные методы и средства полностью автоматического машинного перевода далеки от совершенства. Они имеют существенный недостаток, выражающийся в потере смысла и некорректности результата при переводе одного и того же текста с одного естественного языка на другой естественный язык [2]. Оценка качества перевода определяется из соответствия оригиналу или «эквивалентность перевода», либо с точки зрения достигаемого с помощью перевода коммуникативного эффекта или его прагматической ценности [3]. Данные подходы весьма субъективны, поскольку конкретная оценка формируется в зависимости от установки оценивающего и намерений самого переводчика. Что касается машинного перевода, численная оценка его качества является решающей для проверки корректности алгоритма работы программы-переводчика [1].

1. Постановка задачи. Целью статьи является разработка приложения для автоматической корректировки машинного перевода до уровня публикации. В этой связи предлагается использовать ряд существующих разработок с необходимыми дополнениями для качественной численной оценки машинного перевода путем обратного перевода, предложенного в [1], и сравнения оригинального текста и обратного перевода методом анализа нечетких дубликатов. Это возможно, так

как поиск и анализ нечетких дубликатов является схожей задачей интеллектуальных систем в современном информационном пространстве. Данный метод позволяет сопоставлять без участия человека синтаксические и морфологические параметры переводимого текста, а также анализировать и корректировать параллельные корпуса в соответствии с обнаруженными несоответствиями.

Для реализации предлагаемого метода подходит правило-ориентированный машинный перевод (англ.: Rule-Based Machine Translation-RBMT). В данном типе перевода существуют правила для анализа исходного предложения, правила для передачи представления на стадии анализа, и, наконец, существуют правила для формирования целевого предложения от переданного материала.

2. Обзор основных работ по исследуемой тематике. Целесообразность использования обратного перевода для оценки самого перевода впервые была высказана в [4], однако здесь не предполагалось использование автоматизированного сравнения оригинального текста с обратным переводом.

Идея использовать средства машинного перевода в прямом и обратном направлении с последующим сравнением оригинала и обратного перевода для оценки качества самого перевода изложена в ранее опубликованной авторами статье [1]. В данной работе предполагается использование модели «оригинальный текст – перевод – обратный перевод» для достижения обратимости перевода вне зависимости от самого алгоритма перевода. Оценка качества перевода предполагает сравнение оригинального текста с обратным переводом с получением численного значения схожести текстов.

В работе [1] также предлагается оценивать качество перевода с помощью модели обратимого «черного ящика» на предмет их обратимости и возможности получения не искаженного обратного перевода. В качестве меры оценки предлагается использовать количество корректирующих операций до получения обратным машинным переводчиком перевода, идентичного оригинальному тексту.

В настоящее время существует множество методов оценки машинного перевода, автоматизированного перевода и методов сравнения нечетких дубликатов текста. Рассмотрим некоторые из них.

2.1. Методы оценки машинного перевода. Опишем основные методы, применяемые для оценки машинного перевода:

а) метрика BLEU [5, 6] для статистического МП, рассчитывающая соотношение переводимых параллельных корпусов к соответствующим фразам на целевом языке, вводимом в поисковик;

б) Amazon Mechanical Turk [7] (краудсорсинговый сервис от компании Amazon), экспертные системы для оценки переводов, где большое количество анонимных пользователей оценивает машиннопереведенные тексты, полученные разными системами, в которых на выходе получается некоторый нормализованный коэффициент сравнения вариантов перевода, выполненных различными системами;

в) оценка в процессе перевода и редактирования текста с применением типовых инструкций (предполагает участие постредактора); схема метода, представленная на рис. 1, не предполагает обратный машинный перевод, что является продолжением идеи, изложенной в работе [3].



Рис. 1. Процесс перевода с участием постредактора (TAUS 2010)[8]

2.2. Методы сравнения нечетких дубликатов текста. В настоящее время методы сравнения нечетких дубликатов неструктурированного текста развиваются в следующих направлениях [9]:

- а) синтаксические методы, выполняющие анализ последовательностей, состоящих из символов, слов, словосочетаний или предложений (расстояние Хэмминга [10], расстояние Левенштейна [11], расстояние Джаро-Винклера [12]);
- б) лексические методы, основанные на определении контрольных сумм для слов, словосочетаний, предложений, абзацев и последующий анализ (метод шинглов в различных модификациях [13], метод весовых коэффициентов терминов [14]).

3. Гипотеза исследования. Машинный перевод является эффективным способом перевода текстов большого объема. Автоматизированный перевод используется многими бюро перевода, соответственно, разработка приложения для автоматической коррекции машинного перевода имеет серьезные конкурентные преимущества. Адекватная система машинного перевода должна обладать максимальной обратимостью.

Машинный переводчик должен обладать набором правил [15] для преобразования текста по правилам треугольника Ваукуа [16], который также называется пирамидой Ваукуа, представленной на рис. 2.

Пирамида показывает, что перевод требует обработки на многих уровнях. Левая сторона треугольника является восходящей вверх, а правая сторона является спускающейся вниз. В левом углу упоминается исходный язык в правом углу – целевой язык. Когда мы поднимаемся вверх по левой стороне, мы проводим анализ различных видов с входным предложением. Эта обработка входного предложения может включать один или несколько этапов или же все из следующих этапов:

- а) морфологический анализ;
- б) определение частей речи (англ.: Part of speech – POS);
- в) идентификация группы существительного и глагола (также называется неглубокий разбор или беглый разбор);
- г) разбор с последующим введением семантики;
- д) дискурсивный вывод в виде сопутствующих ссылок;
- е) прагматический анализ.

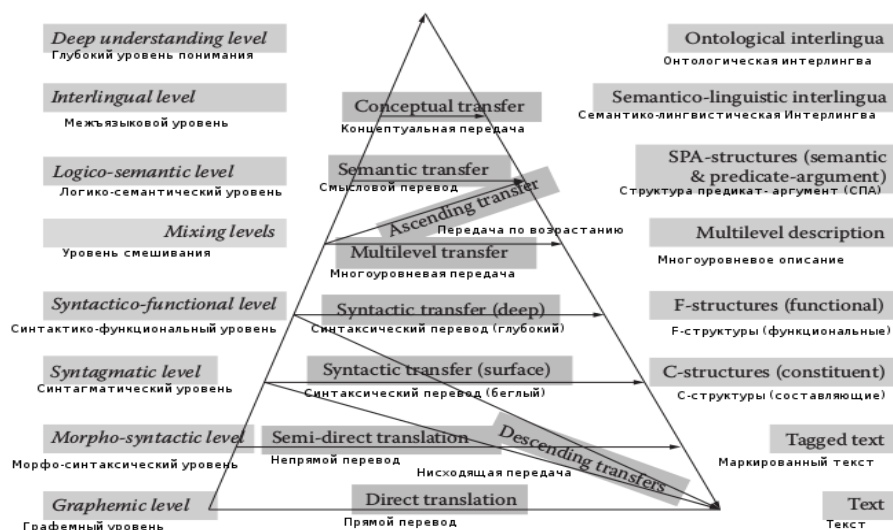


Рис. 2. Пирамида Ваукуа

Основная гипотеза исследования состоит в том, чтобы анализировать процесс перевода на всех этапах преобразования и устанавливать несоответствия при прямом и обратном переводе, с последующей корректировкой. Примером реализации данного метода может быть схема, представленная на рис. 3.

Информация в тексте сравнивается на графемном, морфосинтаксическом и логико-семантическом уровнях. В качестве модели для принятия решения о корректировке выбрана модель на основе установления сходства ситуаций (поиск по образцу), предполагающую сравнение признаков, описывающих ситуацию. При этом подразумевается независимость признаков (свойств, показателей) [17]. Процесс корректировки может осуществляться до достижения необходимой схожести оригинального текста и обратного перевода, либо до достижения заданного максимального числа операций, либо ограничиваться временем.

Сравнение на графемном уровне может быть выполнено при помощи иерархического алгоритма diff [18], в основе которого лежит вычисление расстояния Хемминга [19].

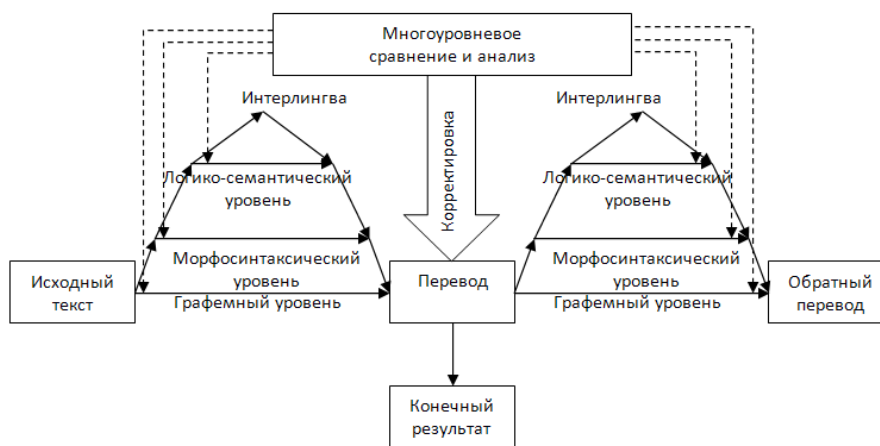


Рис. 3. Сравнение и корректировка результатов машинного перевода

Сравнение на морфосинтаксическом уровне предполагается выполнять при помощи системы морфо-синтаксического анализа Treeton и анализатора Treeval [20];

Логико-семантическое сравнение предполагается производить при помощи алгоритма оценивания знаний студентов в Интеллектуальной системе [21].

4. Подходы к реализации метода. Современные исследования в области анализа переводимых текстов на лексическую целостность показали, что компьютеры успешно справляются с простыми частями речи и устойчивыми выражениями, но допускают несоответствия при переводе падежей, оборотов речи, логико-смысловых атрибутов предложения при их построении. Редактирование переведенного текста подразумевает профессиональную корректировку обнаруженных погрешностей либо разработку приложений автоматической корректировки машиннопереведенного текста. В связи с этим методика улучшения машинного перевода развивается в двух направлениях: постредактирование машинного перевода и доработка алгоритмов самого перевода. Общая схема работы системы автоматической оценки качества и корректировки машинного перевода для абзаца текста представлена на рис. 4.



Рис. 4. Система автоматической оценки качества и корректировки машинного перевода

В данной системе участие человека в процессе перевода ограничивается выбором текста и заданием количества корректирующих операций.

Практическая реализация разрабатываемого метода предполагает:

а) создание развернутого алгоритма анализа нечетких дубликатов машинного перевода для выявления несоответствий отдельных элементов оригинального текста и обратного машинного перевода;

б) разработку развернутого алгоритма принятия решений о корректировке перевода, в случае установления несоответствий отдельных элементов оригинального текста и обратного машинного перевода;

в) написание тестового приложения для верификации метода и алгоритмов на примере современных онлайн-переводчиков;

г) создание интернет-приложения, дополняющего современные онлайн-переводчики, корректирующее результаты их работы в автоматическом режиме.

Заключение. Работа была посвящена проблемам оценки и улучшения качества машинного перевода. Авторами был предложен метод улучшения качества машинного перевода на основе анализа нечетких дубликатов оригинального текста и обратного перевода. Данный метод не имеет аналогов и является новым междисциплинарным направлением на стыке машинного перевода и поиска дубликата-

тов текста. Важность работы заключается в том, что современные бюро переводов не примут на работу специалиста, который умеет переводить тексты только в одном направлении. Поэтому автоматические переводчики становятся необходимым инструментарием для широкого круга специалистов.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Корнилов В.С., Глушань В.М.* Критерии численной оценки качества машиннопереведенного текста // Информационные технологии, системный анализ и управление – ИТСАУ-2016: Сборник трудов XIV Всероссийской научной конференции молодых ученых, аспирантов и студентов, 16-19 ноября 2016 г. Т. 1. – Таганрог: Изд-во ЮФУ, 2016. – С. 170-175.
2. *Яшина Л.И.* Качество автоматического перевода текстов // Актуальные проблемы лингвистики – 2015: Материалы Международной научно-практической конференции студентов, аспирантов и молодых ученых, 15 апреля / отв. ред. Х.С. Шагбанова. – Тюмень: ТюмГНГУ, 2015. – 448 с.
3. *Гуреева Л.В., Козьмина Н.А.* Концепции перевода в контексте современных лингвистических исследований // Молодой ученый. – 2015. – № 11 (91). – 1148 с.
4. *Цвиллинг М.Я., Туровер Г.Я.* О критериях оценки перевода // Тетради переводчика. – 1978. – № 15. – 32 с.
5. *Alexandra Antonova, Alexey Misyurev* Building a Web-based parallel corpus and filtering out machinetranslated text // Proceedings of the 4th Workshop on Building and Using Comparable Corpora, 49th Annual Meeting of the Association for Computational Linguistics, Portland, Oregon, 24 June 2011. – P. 136-144.
6. *Kishore Papineni, Salim Roukos, Todd Ward, Wei-Jing Zhu* BLEU: a Method for Automatic Evaluation of Machine Translation // Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL), Philadelphia, July 2002. – P. 311-318.
7. *Ellen Cushing.* Dawn of the Digital Sweatshop. – East Bay Express article, August 01, 2012. – 12 p.
8. *Serge Gladkoff* Moses: from out of the box to industry quality level in three months // Language business innovation, TAUS user conference Portland (OR), USA 4-6 October 2010. – 20 p.
9. *Дербенев Н.В., Козлюк Д.А., Никитин В.В., Толчеев В.О.* Экспериментальное исследование методов выявления нечетких дубликатов научных публикаций // Машинное обучение и анализ данных. – 2014. – Т. 1, № 7. – 232 с.
10. *Галатенко А.В., Галатенко В.В.* О расстоянии Хэмминга между почти всеми функциями алгебры логики // Фундаментальная и прикладная математика. – 2009. – Т. 15, № 5. – С. 43-47.
11. *Хромов Н.А.* К задаче выявления нечетких дубликатов для обнаружения плагиата в научных публикациях и отчетных материалах // Конференции на факультете Физико-математических и естественных наук РУДН, Информационно-телекоммуникационные технологии и математическое моделирование высокотехнологичных систем 2012. – М.: Изд-во РУДН, 2012. – С. 41-49.
12. *Рубцов Д.Н., Барахнин В.Б.* Выявление дубликатов в разнородных библиографических источниках // Вестник НГУ. Серия: Информационные технологии. – 2009. – Т. 7, № 3. – С. 430-438.
13. *Зеленков Ю.Г., Сегалович И.В.* Сравнительный анализ методов определения нечетких дубликатов для Web-документов // Труды 9-ой Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные коллекции» – RCDL2007. Переславль-Залесский. – 2007. – С. 16-25.
14. *Загорюлько Ю.А., Саломатина Н.В., Серый А.С., Сидорова Е.А., Шестаков В.К.* Выявление нечетких дубликатов при автоматическом формировании тематических коллекций документов на основе Web-публикаций // Вестник Новосибирского государственного университета. Серия: Информационные технологии. – 2013. – Т. 11, Вып. 4. – С. 59-70.
15. *Семенова А.В., Курейчик В.М.* Обзор и анализ состояния проблемы обработки текстовой информации в системах машинного перевода // Информатика, вычислительная техника и инженерное образование. – 2014. – № 2 (17). – С. 76-93.

16. *Pushpak Bhattacharyya* Machine Translation. – CRC Press Taylor&Francis Group, 2015. – 323 p.
17. Глушань В.М., Карелин В.П. Использование математических моделей принятия решений в интеллектуальных САПР // Известия ЮФУ. Технические науки. – 2007. – № 2 (77). – С. 96-103.
18. Луцив Д.В., Кознов Д.В. Иерархический алгоритм DIFF при работе со сложными документами // Системное программирование. – 2012. – Т. 7. – С. 105-114.
19. Глушань В.М., Карелин В.П., Кузьменко О.Л. Нечеткие модели и методы многокритериального выбора в интеллектуальных системах поддержки принятия решений // Известия ЮФУ. Технические науки. – 2009. – № 4 (93). – С. 51-63.
20. Старостин А.С., Мальковский М.Г. Алгоритм синтаксического анализа, используемый в системе морфосинтаксического анализа «Treetop» // Труды международной конференции «Диалог 2007». – М.: Изд-во МГУ, 2007. – С. 516-524.
21. Комарницкая О.И., Комарницкая И.И. Метод семантического сравнения нечеткой информации при проверке текстов на наличие плагиата // Философские проблемы информационных технологий и киберпространства. Прикладные аспекты информационных технологий. – Пятигорск: Изд-во ПГЛУ, 2015. – № 2 (10). – С. 127-139.

REFERENCES

1. Kornilov V.S., Glushan' V.M. Kriterii chislennoy otsenki kachestva mashinoperevedennogo teksta [Criteria for numerical evaluation of the quality of machine translated text], *Informatsionnye tekhnologii, sistemnyy analiz i upravlenie – ITSAU-2016: Sbornik trudov XIV Vserossiyskoy nauchnoy konferentsii molodykh uchenykh, aspirantov i studentov, 16-19 noyabrya 2016 g.* [Information Technology, System Analysis and Management – ITSAU-2016. Proceedings of XIV all-Russian scientific conference of young scientists, postgraduates and students, 16-19 November 2016]. Vol. 1. Taganrog: Izd-vo YuFU, 2016, pp. 170-175.
2. Yashina L.I. Kachestvo avtomaticheskogo perevoda tekstov [The quality of the automatic translation of texts], *Aktual'nye problemy lingvistiki – 2015: Materialy Mezhdunarodnoy nauchno-prakticheskoy konferentsii studentov, aspirantov i molodykh uchenykh, 15 aprelya* [Actual problems of linguistics - 2015: materials of the International scientific-practical conference of students, graduate students and young scientists, April 15], executive editor Kh.S. Shagbanova. Tyumen': TyumGNGU, 2015, 448 p.
3. Gureeva L.V., Koz'mina N.A. Kontseptsii perevoda v kontekste sovremennykh lingvisticheskikh issledovaniy [The concept of translation in the context of contemporary linguistic research], *Molodoy uchenyy* [Young Scientist], 2015, No. 11 (91), 1148 p.
4. Tsvilling M.Ya., Turover G.Ya. O kriteriyakh otsenki perevoda [About the criteria for assessing the translation], *Tetrad' perevodchika* [Interpreters' notebooks], 1978, No. 15, 32 p.
5. Alexandra Antonova, Alexey Misyurev Building a Web-based parallel corpus and filtering out machinetranslated text, *Proceedings of the 4th Workshop on Building and Using Comparable Corpora, 49th Annual Meeting of the Association for Computational Linguistics, Portland, Oregon, 24 June 2011*, pp. 136-144.
6. Kishore Papineni, Salim Roukos, Todd Ward, Wei-Jing Zhu BLEU: a Method for Automatic Evaluation of Machine Translation, *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL), Philadelphia, July 2002*, pp. 311-318.
7. Ellen Cushing Dawn of the Digital Sweatshop. East Bay Express article, August 01, 2012, 12 p.
8. Serge Gladkoff Moses: from out of the box to industry quality level in three months, *Language business innovation, TAUS user conference Portland (OR), USA 4-6 October 2010*, 20 p.
9. Derbenev N.V., Kozlyuk D.A., Nikitin V.V., Tolcheev V.O. Eksperimental'noe issledovanie metodov vyyavleniya nechetkikh dublikatov nauchnykh publikatsiy [Experimental study of methods for identifying near duplicates of scientific publications], *Mashinnoe obuchenie i analiz dannykh* [Machine learning and data analysis], 2014, Vol. 1, No. 7, 232 p.
10. Galatenko A.V., Galatenko V.V. O rasstoyanii Khemminga mezhdu pochti vsemi funktsiyami algebrы logiki [On the Hamming distance between almost all functions of the algebra of logic], *Fundamental'naya i prikladnaya matematika* [Fundamental and applied mathematics], 2009, Vol. 15, No. 5, pp. 43-47.

11. *Khromov N.A.* К задаче выявления нечетких дубликатов для обнаружения плагиата в научных публикациях и отчетных материалах [To the problem of identifying near duplicates for detection of plagiarism in scientific publications and reports], *Konferentsii na fakul'tete Fiziko-matematicheskikh i estestvennykh nauk RUDN, Informatsionno-telekommunikatsionnye tekhnologii i matematicheskoe modelirovanie vysokotekhnologichnykh sistem 2012* [Conferences at the Faculty of Physics and Mathematics and Natural Sciences, PFUR, Information and Telecommunication Technologies and Mathematical Modeling of High-Tech Systems 2012]. Moscow: Izd-vo RUDN, 2012, pp. 41-49.
12. *Rubtsov D.N., Barakhnin V.B.* Выявление дубликатов в различных библиографических источниках [Identification of duplicates in various bibliographic sources], *Vestnik NGU. Seriya: Informatsionnye tekhnologii* [Vestnik NSU. Series: Information technology], 2009, Vol. 7, No. 3, pp. 430-438.
13. *Zelenkov Yu.G., Segalovich I.V.* Сравнительный анализ методов определения нечетких дубликатов для Web-документов [Comparative analysis of methods for determining near duplicates for Web-documents], *Trudy 9-oy Vserossiyskoy nauchnoy konferentsii «Elektronnye biblioteki: perspektivnye metody i tekhnologii, elektronnye kolleksii» – RCDL2007* [Proceedings of the 9th All-Russian Scientific Conference "Digital Libraries: Advanced Methods and Technologies, Digital Collections" - RCDL'2007]. Pereslavl'-Zalesskiy, 2007, pp. 16-25.
14. *Zagorul'ko Yu.A., Salomatina N.V., Seryy A.S., Sidorova E.A., Shestakov V.K.* Выявление нечетких дубликатов при автоматическом формировании тематических коллекций документов на основе Web-публикаций [Identification of near duplicates in the automatic formation of thematic collections of documents based on Web publications], *Vestnik Novosibirskogo gosudarstvennogo universiteta. Seriya: Informatsionnye tekhnologii* [Vestnik of Novosibirsk state University. Series: Information technology], 2013, Vol. 11, Issue 4, pp. 59-70.
15. *Semenova A.V., Kureychik V.M.* Обзор и анализ состояния проблемы обработки текстовой информации в системах машинного перевода [Review and analysis of the state of the problem of processing textual information in machine translation systems], *Informatika, vychislitel'naya tekhnika i inzhenernoe obrazovanie* [Informatics, computer science and engineering education], 2014, No. 2 (17), pp. 76-93.
16. *Pushpak Bhattacharyya* Machine Translation. CRC Press Taylor&Francis Group, 2015, 323 p.
17. *Glushan' V.M., Karelin V.P.* Использование математических моделей принятия решений в интеллектуальных САПР [Use of mathematical models of decision-making in intelligent CAD], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2007, No. 2 (77), pp. 96-103.
18. *Lutsiv D.V., Koznov D.V.* Иерархический алгоритм DIFF при работе со сложными документами [Hierarchical diff algorithm for working with complex documents], *Sistemnoe programmirovaniye* [System Programming], 2012, Vol. 7, pp. 105-114.
19. *Glushan' V.M., Karelin V.P., Kuz'menko O.L.* Нечеткие модели и методы многокритериального выбора в интеллектуальных системах поддержки принятия решений [Near models and methods of multi-choice in intellectual decision support systems], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2009, No. 4 (93), pp. 51-63.
20. *Starostin A.S., Mal'kovskiy M.G.* Алгоритм синтаксического анализа, исполняемый в системе морфосинтаксического анализа «Treton» [Algorithm of parsing used in the system of morphosyntactic analysis "Treton"], *Trudy mezhdunarodnoy konferentsii «Dialog 2007»* [Proceedings of the international conference "Dialogue 2007"], Moscow: Izd-vo MGU, 2007, pp. 516-524.
21. *Komarnitskaya O.I., Komarnitskaya I.I.* Метод семантического сравнения нечеткой информации при проверке текстов на наличие плагиата [Method of semantic comparison of near information when checking texts for plagiarism], *Filosofskie problemy informatsionnykh tekhnologiy i kiberprostranstva. Prikladnye aspekty informatsionnykh tekhnologiy* [Philosophical problems of information technologies and cyberspace. Applied aspects of information technologies]. Pyatigorsk: Izd-vo PGLU, 2015, No. 2 (10), pp. 127-139.

Статью рекомендовала к опубликованию д. филол. н., профессор В.В. Миргородская.

Корнилов Василий Сергеевич – Южный федеральный университет; e-mail: koresh-jr@yandex.ru; 347928, г. Таганрог, пер. Некрасовский, 44; кафедра САПР; аспирант.

Глушань Валентин Михайлович – e-mail: gluval07@rambler.ru; кафедра САПР; профессор.

Лозовой Алексей Юрьевич – e-mail: lozovoy@sfedu.ru; кафедра ИЯ; доцент.

Kornilov Vasilii Sergeevich – Southern Federal University; e-mail: koresh-jr@yandex.ru; 44, Nekrasovskiy, Taganrog, 347928, Russia; the department of computer aided design; postgraduate student.

Glushan Valentin Mihailovich – e-mail: gluval07@rambler.ru; the department of computer aided design; professor.

Lozovoy Aleksey Yurievich – e-mail: lozovoy@sfedu.ru; the department of foreign languages; associate professor.

УДК 004.023

Ю.А. Кравченко, И.О. Курситыс

**КОМБИНИРОВАННЫЙ ПОДХОД К РЕШЕНИЮ ЗАДАЧИ
РАСПРЕДЕЛЕНИЯ РЕСУРСОВ***

Данная статья посвящена решению задачи распределения вычислительных ресурсов в компьютерных системах. Общей смысл задачи распределения ресурсов – распределить ограниченный ресурс между потребителями оптимальным образом. В работе сформулирована постановка задачи и приведена ее математическая модель. Задача распределения вычислительных ресурсов относится к классу NP-сложных, что обуславливает невозможность или ограниченность применения традиционных средств математического анализа для ее решения. Разработанный авторами подход основан на использовании имитационного моделирования и аппарата сетей Петри для распределения ресурсов между процессами. Так как распределяемый ресурс является вычислительным, главным критерием оптимизации принято время, затрачиваемое на исполнение всех процессов, использующих ресурсы, а для моделирования задачи используются временные сети Петри. Комбинированный подход к решению задачи заключается в совместном применении аппарата сетей Петри для моделирования задачи, а также генетических алгоритмов, успешно показавших себя в решении NP-сложных задач. Для оптимизации модели применены известные генетические алгоритмы и модифицированные генетические операторы кроссинговера и мутации. Авторами разработан программный продукт – подсистема распределения ресурсов. В статье отражены основные функции подсистемы, ее модульная структура. Построена и проанализирована информационная модель подсистемы на основе стандарта IDEF0, отражающая информационные процессы, протекающие в подсистеме. Проведены вычислительные эксперименты на тестовых примерах (бенчмарках), доказывающие эффективность и перспективность разработанного подхода.

Распределение ресурсов; CALS-технологии; имитационное моделирование; сети Петри; генетические алгоритмы.

Yu.A. Kravchenko, I.O. Kursitys

**COMBINED APPROACH FOR SOLVING RESOURCE ALLOCATION
PROBLEM**

The paper is devoted to solving the computative resource allocation problem. In general, the resource allocation problem involves allocating the limited resource among processes in optimal way. Authors formulated the problem statement and its mathematical model. The resource allocation

* Работа выполнена при финансовой поддержке РФФИ (проект № 17-07-00446).