

ГРИГОРЯН Л.А.

ПОЛЬЗОВАТЕЛЬСКИЙ ИНТЕРФЕЙС ДЛЯ ИЗВЛЕЧЕНИЯ ХИМИКО-СТРУКТУРНОЙ ИНФОРМАЦИИ ИЗ СИСТЕМАТИЧЕСКОГО НАЗВАНИЯ ОРГАНИЧЕСКОГО СОЕДИНЕНИЯ

*Всероссийский институт научной и технической информации Российской академии наук (ВИНИТИ РАН)
Москва, Российская Федерация*

В ВИНИТИ РАН разработан пользовательский интерфейс «Номенклатурный Генератор», предназначенный для автоматического извлечения химико-структурной информации из систематического названия органического соединения, данного в номенклатуре ИЮПАК.

Ключевые слова: *пользовательский интерфейс, химическая номенклатура, органическое соединение, структурный граф*

Одним из многообещающих направлений компьютерной лингвистики является приложение ее методов и понятийного аппарата к задачам в области химической науки, примером чему служит известная задача извлечения информации о структуре органического вещества из его систематического названия. В самом деле, корректно составленное название органического соединения позволяет получить посредством интеллектуального анализа достаточно полное представление о строении его молекулы, т. е. о числе и характере составляющих ее атомов, перечне межатомных связей, их кратности, валентности каждого атома и т. д., и в конечном итоге построить модель структурного графа органического соединения. Взаимно однозначное соответствие между названием и структурой химического соединения является опорным принципом современной химической номенклатуры и делает указанную задачу принципиально разрешимой.

Теоретической основой для решения данной задачи послужили работы А.М. Цукермана, Г.Г. Стецюры, Г.Э. Влэдуца (СССР), Э. Гарфилда, Дж. Уалгрена (США) и др. [4–5, 20–22, 24, 31, 34]. Первые опыты практической реализации данных теоретических построений с использованием компьютерных технологий были предприняты на заре информационной эры в 1950-е–1960-е гг. По мере развития компьютерной техники в последующие десятилетия совершенствовались и подходы к решению данной задачи, представленные в отечественной науке работами М.М. Ланглебен, Э.А. Давыдовой, В.В. Бондаря [1–2, 6–11, 13–19, 23], а за рубежом – исследованиями Д.И. Кук-Фокса, Р. Клингера, Р.А. Сэйла и др. [25–30, 32–33]. Общим звеном этих подходов являлось рассмотрение химической номенклатуры в качестве лингвистической системы, своего рода искусственного языка, а названий индивидуальных органических соединений – как сложных слов этого языка, строящихся из осмысленных компонентов по определенным правилам, четко формализованным в химической номенклатуре и, следовательно, поддающимся алгоритмизации.

В ВИНИТИ РАН автором настоящей статьи под руководством В.В. Бондаря и Е.Г. Винокурова разработан пользовательский интерфейс «Номенклатурный Генератор», представляющий собой программную оболочку, ориентированную на извлечение химико-структурной информации из вводимого пользователем систематического названия органического соединения и восстановление согласно этой информации молекулярно-структурного графа данного соединения [12].

«Номенклатурный Генератор» предназначен для построения структурных графов по названиям химических соединений, относящихся к ряду разделов органической химии, в т. ч.:

- алифатических соединений;
- моноциклических соединений;
- соединений с функциональными группами, именуемых по правилам заместительной номенклатуры;
- соединений, именуемых по правилам заместительной номенклатуры («а»-номенклатуры);
- гетероциклов, именуемых по правилам расширенной системы Ганча-Видмана;
- основных ароматических соединений;
- некоторых соединений, сохраняющих тривиальные названия.

Ход действия алгоритма данной программной оболочки включает в себя следующие этапы:

- 1) морфемный анализ вводимого пользователем названия химического соединения;
- 2) соотнесение введенного названия химического соединения с определенным типовым шаблоном;
- 3) считывание соответствующей данному шаблону химико-структурной информации из встроенных в программную оболочку служебных лингвохимических словарей;
- 4) свертка шаблона в единый метасимвол по правилам алгоритмизированной укорачивающей грамматики G с одновременным построением структурного графа на основе считанной из словарей информации;

5) проверка полученного структурного графа на валидность;

6) вывод результата на экран.

Для реализации поставленной задачи были составлены специализированные словари химических морфем, в т. ч. химико-структурный словарь, мультиязычный словарь и словарь омонимов.

Химико-структурный словарь содержит информацию о более чем 1000 химических морфем, используемых международной номенклатурой ИЮПАК. Морфемы в словаре сгруппированы по семантико-синтаксическим типам в зависимости от их роли в составе названия химического соеди-

нения: морфемы одного типа обозначают углеродные каркасы молекулы химического соединения, другого – обозначают характер и кратность химической связи между атомами внутри молекулы, третьего – соответствуют гетероатомам, замещающим атомы углерода в составе углеродного каркаса молекулы и т. д. Химико-структурные данные, однозначно характеризующие каждую морфему, помещены в соответствующих полях словаря, организованного в формате электронной базы данных.

Пример таблицы основного словаря можно видеть на рис. 1:

| MorphemeName | MorphemeType | VerticeNur | BondNur | Vertices | Bonds |
|--------------|--------------|------------|---------|-----------------------------------|--|
| мет | Root | 1 | 0 | C 4 | - |
| эт | Root | 2 | 1 | C 3, C 3 | 1 2 1 |
| проп | Root | 3 | 2 | C 3, C 2, C 3 | 1 2 1, 2 3 1 |
| бут | Root | 4 | 3 | C 3, C 2, C 2, C 3 | 1 2 1, 2 3 1, 3 4 1 |
| пент | Root | 5 | 4 | C 3, C 2, C 2, C 2, C 3 | 1 2 1, 2 3 1, 3 4 1, 4 5 1 |
| гекс | Root | 6 | 5 | C 3, C 2, C 2, C 2, C 2, C 3 | 1 2 1, 2 3 1, 3 4 1, 4 5 1, 5 6 1 |
| гепт | Root | 7 | 6 | C 3, C 2, C 2, C 2, C 2, C 2, C 3 | 1 2 1, 2 3 1, 3 4 1, 4 5 1, 5 6 1, 6 7 1 |

Рисунок 1. Фрагмент таблицы основного словаря в редакторе баз данных Microsoft Access

Мультиязычный словарь содержит эквиваленты каждой морфемы химико-структурного словаря на русском и английском языках и допускает расширение на французский, немецкий, испанский, итальянский и прочие языки, для которых предусмотрены адаптации номенклатуры ИЮПАК.

Словарь синонимов описывает соответствия, главным образом, между т.н. тривиальными (т. е. сложившимися исторически) и систематическими (т. е. установленными согласно правилам номенклатуры) химическими морфемами и названиями химических соединений.

Так, например, тривиальное название «Ацетон» соответствует химическому соединению со структурой, представленной на рис. 2. Это же соединение по правилам номенклатуры имеет альтернативное систематическое название «Пропан-2-он».

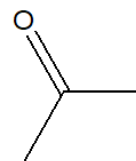


Рисунок 2. Ацетон, он же – Пропан-2-он

Разработанный в ВИНТИ РАН пользовательский интерфейс «Номенклатурный Генератор» может применяться научными работниками для получения структурных графов по названию органического соединения при подготовке макетов научных статей и патентов изобретений, работе с базами структурных данных по различным разделам химии и фармакологии, а также может использоваться как тренажер в учебном процессе [3].

ЛИТЕРАТУРА

1. Бондарь, В.В. Укорачивающая грамматика на основе обновленной классификации морфем химической номенклатуры, используемая в программном комплексе «Номенклатурный Генератор» / В.В. Бондарь, Е.Г. Винокуров, Л.А. Григорян // НТИ. – 2014. – Серия 2, № 7. – С. 6-15.
2. Бондарь, В.В. Комплекс программ визуализации молекулярных структурных графов органических соединений / В.В. Бондарь [и др.] // Научная визуализация. – 2018. – Т. 10, № 3. – С. 133-143.
3. Винокуров Е.Г. Обучающая программа для генерации структурного графа по названию органического соединения / Е.Г. Винокуров [и др.] // Актуальные проблемы химико-технологического образования. Разработка образовательных программ в соответствии с требованиями ФГОС ВО 3+-. 17 межвуз. учеб.-метод. конф.: материалы конф. – М.: РХТУ им. Д.И. Менделеева, 2016. – С. 98.

4. Влэдуч Г.Э. Некоторые вопросы научной информации в области химии / Г.Э. Влэдуч – М.: Институт научной информации АН СССР, 1958. – 134 с.
5. Влэдуч Г.Э. Об эксперименте по машинному переводу названий химических органических соединений в формулы / Г.Э. Влэдуч, И.Я. Паппэ // НТИ, 1962, № 3. – С. 16-17.
6. Григорян Л.А. Программа перевода систематических названий химических соединений в молекулярные графы (расширение на номенклатуру Ганча-Видмана); Всеросс. ин-т науч. и техн. информации / Л.А. Григорян. – М., 2006. – 19 с. – Ил. – Библиогр.: 10 назв. – Рус. – Деп. в ВИНТИ РАН 25.08.06, № 1103-В 2006.
7. Григорян Л.А. Автоматическое порождение структуры по названию химического соединения // Компьютерная лингвистика и интеллектуальные технологии: Труды Международной конференции «Диалог-2006» (Бекасово, 31 мая – 4 июня 2006 г.), – М.: Изд-во РГГУ, 2006. – С. 119–123.
8. Григорян Л.А. Разработка словарей морфем химической номенклатуры // Вестник РГГУ, № 8 (130), серия «Филологические науки. Языкознание» / Л.А. Григорян // Московский лингвистический журнал, т. 16, 2014. – С. 139–149.
9. Григорян Л.А. Программа перевода систематических названий химических соединений в молекулярные графы (расширение на заменительную номенклатуру) / Л.А. Григорян, В.В. Бондарь, И.Б. Немировская // НТИ. Серия 2, Информационные процессы и системы. – 2006. – № 3. – С. 21-25.
10. Григорян Л.А. Программа перевода систематических названий химических соединений в молекулярные графы (расширение на класс ароматических соединений) / Л.А. Григорян, В.В. Бондарь, И.Б. Немировская // НТИ. Серия 2, Информационные процессы и системы. – 2006. – № 8. – С. 21-26.
11. Григорян Л.А. Трехмодульный программный комплекс для перевода русскоязычного названия органического соединения в структурный граф / Л.А. Григорян, Е.Г. Винокуров, В.В. Бондарь // Успехи в химии и химической технологии: сб. науч. тр. Том XXVIII, № 1 (150). – М.: РХТУ им. Д.И. Менделеева, 2014. – С. 117-120.
12. Григорян Л.А., Винокуров Е.Г., Бондарь В.В., Марголин Л.Н., Фарафонов В.В., Королева Л.М. Программный комплекс «Номенклатурный Генератор», предназначенный для преобразования названий органических соединений в MOL-формат, отражающий структуру молекулярного графа. Свидетельство о государственной регистрации программы для ЭВМ № 2014619365. Правообладатель ФГБУН ВИНТИ РАН (RU); заявка № 2014617378; дата поступления 25.07.2014; дата гос. регистрации в Реестре программ для ЭВМ 15.09.2014.
13. Давыдова Э.А. Разработка проблемы перевода названий органических соединений с одного естественного языка на другой : дис. ... канд. техн. наук : 05.25.05 / Э.А. Давыдова. – М.: ВИНТИ, 1986.
14. Ланглебен М.М. О синтезе названий химических соединений / М.М. Ланглебен // НТИ, 1965, № 10. – С. 18-24.
15. Ланглебен М.М. К лингвистическому описанию номенклатуры органической химии / М.М. Ланглебен // НТИ, 1967, № 1. – С. 13-22.
16. Ланглебен М.М. Опыт приспособления лингвистических понятий и лингвистической терминологии к описанию искусственного языка / М.М. Ланглебен // Информационные поисковые системы и автоматическая обработка научно-технической информации, 1967. – С. 170-224.
17. Ланглебен М.М. Структура номинативных сочетаний в специальном фрагменте русского химического языка : дис. канд. филол. наук / М.М. Ланглебен. – М.: ВИНТИ, 1970. – 257 с.
18. Ланглебен М.М. Опыт построения метаязыка для описания квазилингвистической семиотической системы // Сб. «Исследования по математической лингвистике, математической логике и информационным языкам» / М.М. Ланглебен // под ред.: Д.А. Бочвар, Ю.А. Шрейдер; АН СССР. – М.: Наука, 1972. – С. 96-146.
19. Мешалкин В.П. Морфосинтаксический алгоритм компьютерного анализа систематических названий базовых и модифицированных алифатических соединений / В.П. Мешалкин [и др.] // Теоретические основы химической технологии. – 2017. – Т. 51. – № 5. – С. 605-610.
20. Сейфер А.Л. Об алгоритме преобразования названия комплексного соединения, данного в рациональной номенклатуре, в линейную формулу / А.Л. Сейфер, В.С. Штейн // Сообщ. лаборатории электромоделирования. Вып. 1. – М.: Институт научной информации АН СССР, 1960. – 172 с.
21. Сейфер А.Л. Использование ЭВМ для перевода названий комплексных соединений в формулы / А.Л. Сейфер, В.С. Штейн, С.С. Шурова // Журнал неорганической химии, 1960, т. IV, вып. 8, с. 27–29.
22. Стецюра Г.Г. Автоматический перевод названия химических органических соединений в формулы / Г.Г. Стецюра, А.М. Цукерман // НТИ, 1962. – № 3. – С. 17-19.
23. Уткина Е.А. Программа перевода названий химических соединений в систематической номенклатуре в молекулярные графы (для некоторых важных классов органических соединений) / Е.А. Уткина // НТИ. Серия 2. Информационные процессы и системы, 2000. – № 3. – С. 24-36.
24. Цукерман А.М. Об автоматизации перевода названия химических органических соединений в стандартную форму и структурных формул в систематические наименования / А.М. Цукерман, Г.Г. Стецюра // Сообщ. лаборатории электромоделирования. Вып. 1. – М.: Институт научной информации АН СССР, 1960, с. 241.
25. Cooke-Fox D.I. [et al.]. The Concise Connection Table: Collected definitions with extensions for stereochemistry and saccharides. – Journal of Chemical Information and Computer Sciences, 1992, vol. 32, № 5, pp. 556-559.

26. **Cooke-Fox D.I.** [et al.]. Computer translation of IUPAC systematic organic chemical nomenclature. 4. Concise connection tables to structure diagrams. *Journal of Chemical Information and Computer Sciences*, 1990, vol. 30, № 2, pp. 122-127.
27. **Cooke-Fox D.I.** [et al.]. Computer translation of IUPAC systematic organic chemical nomenclature. 5. Steroid nomenclature. *Journal of Chemical Information and Computer Sciences*, 1990, vol. 30, № 2, pp. 128-132.
28. **Cooke-Fox D.I., Kirby G.H., Rayner J.D.** Computer translation of IUPAC systematic organic chemical nomenclature. 1. Introduction and background to a grammar-based approach. *Journal of Chemical Information and Computer Sciences*, 1989, vol. 29, pp. 101-105.
29. **Cooke-Fox D.I., Kirby G.H., Rayner J.D.** Computer translation of IUPAC systematic organic chemical nomenclature. 2. Development of a formal grammar. *Journal of Chemical Information and Computer Sciences*, 1989, vol. 29, pp. 106-112.
30. **Cooke-Fox D.I., Kirby G.H., Rayner J.D.** Computer translation of IUPAC systematic organic chemical nomenclature. 3. Syntax analysis and semantic processing. *Journal of Chemical Information and Computer Sciences*, 1989, vol. 29, pp. 112-118.
31. **Garfield E.** Chemico-linguistics: "Computer translation of chemical Nomenclature" // *Nature*, 1961, vol. 192, № 4798. – 274 p.
32. **Klinger R.** [et al.]. Detection of IUPAC and IUPAC-like chemical names. – *Bioinformatics*, 2008, vol. 24, № 13, pp. i268-i276.
33. **Sayle R.** Foreign language translation of chemical nomenclature by computer. *Journal of Chemical Information and Modeling*, 2009, vol. 49, № 3, pp. 519-530.
34. **Wahlgren J.H.** Linguistics analysis of Russian chemical terminology // *International Conference on Machine Translation and Applied Language Analysis*, 1961. – London, 1962, pp. 250-263.

REFERENCES

1. **Bondar V.V., Vinokurov E.G., Grigoryan L.A.** Shortening grammar based on the renewed classification of chemical nomenclature morphemes for use in the «Nomenclature Generator» software complex // *Automatic Documentation and Mathematical Linguistics*, 2014, Vol. 48, No. 4, pp. 199-207.
2. **Bondar V.V., Vinokurov E.G., Meshalkin V.P., Grigoryan L.A.** Visualisation of the organic compounds' structural graphs generated upon the user's query // *Scientific Visualization*, 2018, Vol. 10, No. 3, pp. 133-143.
3. **Vinokurov E.G., Bondar V.V., Meshalkin V.P., Grigoryan L.A.** Training program for generating a structural graph by the name of an organic compound // *Actual problems of chemical and technological education. Development of educational programs in accordance with the requirements of the Federal State Educational Standard in 3++*. 17th Interuniversity educational and methodological conference: materials of the conference. – M.: D. Mendeleev University of Chemical Technology of Russia, 2016, p. 98.
4. **Vleduts G.E.** Some problems of scientific information in the field of chemistry. – M.: Institute of Scientific Information of the Academy of Sciences of the Soviet Union, 1958. – 134 p.
5. **Vleduts G.E., Pappé I.Ya.** On the experiment on the machine translation of the names of chemical organic compounds into formulas // *NTI*, 1962, No. 3, pp. 16-17.
6. **Grigoryan L.A.** Program for the translation of systematic names of chemical compounds into molecular graphs (extension to the Ganch-Widman nomenclature). M., 2006. – 19 p. – Deposited in VINITI RAS 25.08.06, No. 1103-B 2006.
7. **Grigoryan L.A.** Automatic generation of a structure by the name of a chemical compound // *Computational linguistics and intellectual technologies: Proceedings of the International Conference «Dialog-2006»* (Bekasovo, May 31 – June 4, 2006). M.: RSUH, 2006, pp. 119-123.
8. **Grigoryan L.A.** Development of dictionaries of morphemes of chemical nomenclature // *Bulletin of the Russian State University*, No. 8 (130), series «Philological sciences. Linguistics» / *Moscow Linguistic Journal*, 2014, vol. 16, pp. 139-149.
9. **Grigoryan L.A., Bondar V.V., Nemirovskaya I.B.** Program for the translation of systematic names of chemical compounds into molecular graphs (extension to substitute nomenclature) // *NTI. Series 2, Information processes and systems*, 2006, No. 3, pp. 21-25.
10. **Grigoryan L.A., Bondar V.V., Nemirovskaya I.B.** Program for translating systematic names of chemical compounds into molecular graphs (extension to the class of aromatic compounds) // *NTI. Series 2, Information processes and systems*, 2006, No. 8, pp. 21-26.
11. **Grigoryan L.A., Vinokurov E.G., Bondar V.V.** A three-module software package for translating the Russian-language name of an organic compound into a structural graph // *Successes in chemistry and chemical technology: Collection of scientific papers*, vol. 28, No. 1 (150). M., D. I. Mendeleev Russian State Technical University, 2014, pp. 117-120.
12. **Grigoryan L.A., Vinokurov E.G., Bondar V.V., Margolin L.N., Farafonov V.V., Koroleva L.M.** The «Nomenclature Generator» software package designed to convert the names of organic compounds into MOL-format representing the structure of the molecular graph. Certificate of state registration of the computer program No. 2014619365. Copyright holder VINITI RAS; application No. 2014617378; date of receipt 25.07.2014; date of state registration in the Register of computer programs 15.09.2014.
13. **Davydova E.A.** Development of the problem of translating the names of organic compounds from one natural language to another: dissertation for the degree of Candidate of Technical Sciences: 05.25.05. M.: VINITI, 1986.
14. **Langleben M.M.** On the synthesis of names of chemical compounds // *NTI*, 1965, No. 10, pp. 18-24.
15. **Langleben M.M.** To the linguistic description of the nomenclature of organic chemistry // *NTI*, 1967, No. 1, pp. 13-22.
16. **Langleben M.M.** Experience of adaptation of linguistic concepts and linguistic terminology to the description of an artificial language // *Information search systems and automatic processing of scientific and technical information*, 1967, pp. 170-224.

17. **Langleben M.M.** Structure of nominative combinations in a special fragment of the Russian chemical language: dissertation for the degree of Candidate of Philological Sciences. M., VINITI, 1970, 257 p.
18. **Langleben M.M.** The experience of constructing a metalanguage to describe a quasi-linguistic semiotic system // Collection «Studies in mathematical linguistics, mathematical logic and information languages» / edited by D.A. Bochvar, Yu.A. Schreider; Academy of Sciences of the Soviet Union. M., Nauka Publ., 1972, pp. 96-146.
19. **Meshalkin V.P., Bondar V.V., Vinokurov E.G., Grigoryan L.A.** Morphosyntactic algorithm for computer analysis of systematic names of basic and modified aliphatic compounds // Theoretical Foundations of Chemical Technology, 2017, vol. 51, No. 5, pp. 605-610.
20. **Seifer A.L., Stein V.S.** On the algorithm for converting the name of a complex compound given in rational nomenclature into a linear formula // Message of the electromodeling laboratories. Issue 1. M., Institute of Scientific Information of the Academy of Sciences of the Soviet Union, 1960, p. 172.
21. **Seifer A.L., Stein V.S., Shurova S.S.** Using computers to translate the names of complex compounds into formulas // Journal of Inorganic Chemistry, 1960, vol. 4, issue 8, pp. 27-29.
22. **Stetsyura G.G., Tsuckerman A.M.** Automatic translation of the names of chemical organic compounds into formulas // NTI, 1962, No. 3, pp. 17-19.
23. **Utkina E.A.** Program for translating the names of chemical compounds in the systematic nomenclature into molecular graphs (for some important classes of organic compounds) // NTI. Series 2. Information Processes and Systems, 2000, No. 3, pp. 24-36.
24. **Tsuckerman A.M., Stetsyura G.G.** On automating the translation of the names of chemical organic compounds into standard form and structural formulas into systematic names // Message of the electromodeling laboratories. Issue 1. M., Institute of Scientific Information of the Academy of Sciences of the Soviet Union, 1960, p. 241.
25. **Cooke-Fox D.I., Ewart J.F., Kirby G.H., Lord M.R., Rayner J.D.** The Concise Connection Table: Collected definitions with extensions for stereochemistry and saccharides // Journal of Chemical Information and Computer Sciences, 1992, vol. 32, No. 5, pp. 556-559.
26. **Cooke-Fox D.I., Kirby G.H., Lord M.R., Rayner J.D.** Computer translation of IUPAC systematic organic chemical nomenclature. 4. Concise connection tables to structure diagrams // Journal of Chemical Information and Computer Sciences, 1990, vol. 30, No. 2, pp. 122-127.
27. **Cooke-Fox D.I., Kirby G.H., Lord M.R., Rayner J.D.** Computer translation of IUPAC systematic organic chemical nomenclature. 5. Steroid nomenclature // Journal of Chemical Information and Computer Sciences, 1990, vol. 30, No. 2, pp. 128-132.
28. **Cooke-Fox D.I., Kirby G.H., Rayner J.D.** Computer translation of IUPAC systematic organic chemical nomenclature. 1. Introduction and background to a grammar-based approach // Journal of Chemical Information and Computer Sciences, 1989, vol. 29, pp. 101-105.
29. **Cooke-Fox D.I., Kirby G.H., Rayner J.D.** Computer translation of IUPAC systematic organic chemical nomenclature. 2. Development of a formal grammar // Journal of Chemical Information and Computer Sciences, 1989, vol. 29, pp. 106-112.
30. **Cooke-Fox D.I., Kirby G.H., Rayner J.D.** Computer translation of IUPAC systematic organic chemical nomenclature. 3. Syntax analysis and semantic processing // Journal of Chemical Information and Computer Sciences, 1989, vol. 29, pp. 112-118.
31. **Garfield E.** Chemico-linguistics: «Computer translation of chemical Nomenclature» // Nature, 1961, vol. 192, No. 4798, 274 p.
32. **Klinger R., Kolářik C., Fluck J., Hofmann-Apitius M., Friedrich C.M.** Detection of IUPAC and IUPAC-like chemical names // Bioinformatics, 2008, vol. 24, No. 13, pp. i268-i276.
33. **Sayle R.** Foreign language translation of chemical nomenclature by computer // Journal of Chemical Information and Modeling, 2009, vol. 49, No. 3, pp. 519-530.
34. **Wahlgren J.H.** Linguistics analysis of Russian chemical terminology // International Conference on Machine Translation and Applied Language Analysis, 1961. London, 1962, pp. 250-263.

GRIGORYAN L.A.

USER'S INTERFACE FOR EXTRACTION OF THE CHEMICAL STRUCTURE INFORMATION FROM THE SYSTEMATIC NAME OF ORGANIC COMPOUND

All-Russian Institute of Scientific and Technical Information of the Russian Academy of Sciences

The user's interface «Nomenclature Generator» for extraction of the chemical structure information from the systematic name of organic compound represented according to IUPAC nomenclature is developed at the All-Russian Institute for Scientific and Technical Information of Russian Academy of Sciences.

Keywords: *user's interface, chemical nomenclature, organic compound, molecular graph*

Григорян Лев Арменович – к.т.н., научный сотрудник Всероссийского института научной и технической информации Российской академии наук (ВИНИТИ РАН), Москва, Россия

Grigoryan Lev Armenovich – candidate of technical sciences, researcher in the All-Russian Institute of Scientific and Technical Information of the Russian Academy of Sciences

E-mail: Levgr2@yandex.ru