МИНИСТЕРСТВО НА ОБРАЗОВАНИЕТО И НАУКАТА НА РЕПУБЛИКА БЪЛГАРИЯ
КИРИЛО-МЕТОДИЕВСКИ НАУЧЕН ЦЕНТЪР ПРИ БЪЛГАРСКА АКАДЕМИЯ НА НАУКИТЕ
ИЖЕВСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ ИМ. М. Т. КАЛАШНИКОВА
НАУЧНОЕ СООБЩЕСТВО "ПИСЬМЕННОЕ НАСЛЕДИЕ"
DIGITAL MEDIEVALIST SCHOLARLY COMMUNITY
ФОНДАЦИЯ „УСТОЙЧИВО РАЗВИТИЕ НА БЪЛГАРИЯ "

Писменото наследство и информационните технологии

El'Manuscript–2014

Материали от V международна научна конференция
Варна, 15–20 септември 2014 г.

София · Ижевск
2014

Сборникът е издаден с финансовата подкрепа на Министерството на образованието и науката на Република България по процедура за подкрепа на международни научни форуми.

Отговорни редактори: проф. дфн В. А. Баранов
доц. д-р В. Желязкова
д-р А. М. Лаврентьев

Редактори: Нели Ганчева, Веселка Желязкова (български текст)
О. В. Зуга, В. А. Баранов (руски текст)
Кевин Хокинс (Kevin Hawkins) (английски текст)

Писменото наследство и информационните технологии [Текст] : материали от V международна науч. конф. (Варна, 15–20 септември 2014 г.) / отг. ред. В. А. Баранов, В. Желязкова, А. М. Лаврентьев. — София ; Ижевск, 2014. — 448 с.

Сборникът съдържа материали от конференция, посветена на разработването и създаването на съвременни средства за съхраняване, описване, обработка, анализ и публикуване на ръкописни и старопечатни книжовни паметници и исторически извори, а също и на въпросите за подготвянето на електронни ресурси в областта на хуманитаристиката и тяхното използване в научните изследвания и преподаването.

© Кирило-Методиевски научен център — БАН, 2014
© Ижевский государственный технический университет им. М. Т. Калашникова, 2014
© Авторски колектив, 2014
© Лилия Тошкова — графичен дизайн на корицата, 2014

ISBN 978–954–9787–25–2

Поиск фрагментов мелодий в знаменном песнопении

М. В. Даньшина, А. Ю. Филиппович

Знаменные песнопения, знаменная нотация, поиск в древнерусских рукописях, музыкальные информационные технологии, семиография

Search of Fragments of Melodies in Znamenny Chants

Marina Danshina, Andrey Philippovich

Many texts of medieval books have already been digitized, and digitization of medieval Russian musical manuscripts using the znamenny notation is being conducted. When researching the chants contained in these books, the researcher often needs to query the signs that they contain, for example, to find all fragments of the manuscript in which a certain melody is found. This paper describes the realization of two options of search: strict and mild. The search algorithm is given and the some results of the function's performance are shown.

В настоящее время электронный формат рукописи более предпочтителен для исследователей, поэтому многие тексты древних книг уже переведены в электронный вид. Помимо обычных книг, в электронные форматы переводятся и другие сохранившиеся источники древней культуры — музыкальные рукописи, которые были записаны с помощью знаменной нотации (рис. 1).

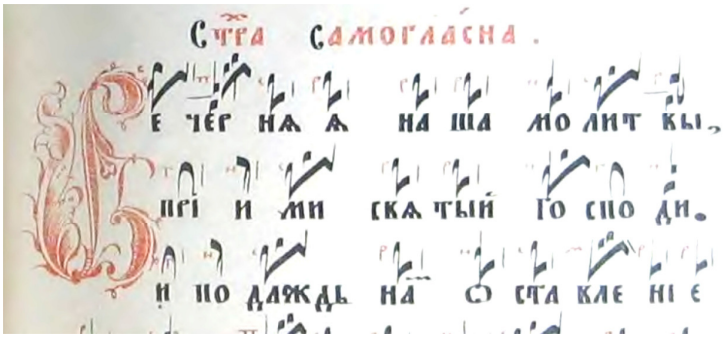


Рис. 1. Пример знаменного песнопения [Круг пения 1884]

Особенностью этих песнопений является то, что они записаны не привычными нотами на пяти линейках, а с помощью специальных знаков, которые называются “крюками” или “знаменами” [Разумовский; Смоленский 1888]. Общее количество знамен достаточно велико, и оно менялось с течением времени. В

© М. В. Даньшина, А. Ю. Филиппович, 2014

нашем исследовании выделено и закодировано с помощью специально разработанного шрифта 212 различных знамен.

В рамках проекта “Компьютерная семиография” [КС] проводятся исследования древнерусских музыкальных рукописей с целью разработки структуры и методов обработки знаменных песнопений [Даньшина И., Даньшина М. 2010], вопросов визуализации и разработки электронного издания семиографических песнопений [Даньшина И., Даньшина М. 2009], статистического и семантического анализа музыкальных рукописей [Голубева, Филиппович А. 2013а, 2013б; Голубева (Даньшина) 2012; Филиппович А., Голубева 2012], исследования методов машинного перевода знаменных песнопений в линейную нотацию [Даньшина М., Филиппович А. 2013] и др. В 2011–2013 годах исследования были поддержаны грантом Российского гуманитарного научного фонда (РГНФ) № 110412025в “Автоматизированная система научных исследований в области компьютерной семиографии (АСНИ КС)”.

Задача поиска знамен

База данных АСНИ “Компьютерная семиография” содержит записи из трех древнерусских музыкальных рукописей — “Круг древнего знаменного церковного пения” [Круг 1884], “Ирмологий” [Ирмологий] и “Сборник попевок Соловецкого собрания” [Сборник попевок]. При работе с песнопениями, содержащимися в этих книгах, исследователю часто требуется выполнить поиск по знаменам, которые в них содержатся. Например, найти все фрагменты рукописи, в которых встречается определенная мелодия. Формально поиск можно представить как функцию следующего вида:

$$\{R_1\}_n = \varphi_1 (M,Kr,Str),$$

где M — рукопись, Kr — критерии поиска, Str — поисковая строка, φ_1 — функция поиска, $\{R_1\}$ — результат поиска, n — количество результатов.

Другой важной задачей является сравнительный анализ мелодий в двух рукописях для выявления соответствий между ними. Эту операцию также можно представить как поиск, но уже по двум источникам:

$$\{R_2\}_n = \varphi_2 (M_1,M_2,Kr,Str),$$

где M_1 и M_2 — рукописи, по которым производится поиск, Kr — критерии поиска, Str — поисковая строка, φ_2 — функция сравнительного поиска, $\{R_2\}$ — результат поиска, n — количество результатов.

Для осуществления этих операций были разработаны специальные веб-сервисы, которые позволяют осуществлять *строгий* и *нестрогий* поиск. В первом случае система ищет все знамена в строго заданной последовательности. Во втором случае пользователь сам настраивает критерии, в том числе со значениями “учитывать последовательность знамен”, “включать в результат поиска фрагменты мелодии, в которых встретились не все искомые знамена”.

Структура базы данных

Песнопения в исследуемых книгах представлены в двух форматах — “крюковая рукопись” (мелодии записаны только знаменами) и “двоезнаменник” (мелодии записаны как в знаменной нотации, так и линейной). Изначально рукописи вводились с помощью текстового редактора MS Word, а затем конвертировались с помощью специальной программы в формат XML. В настоящее время все записи песнопений преобразованы и хранятся в реляционной базе данных.

Каждому знамени рукописи соответствует одна строка в таблице БД, в которой дополнительно содержится метаинформация: *слог*, с которым знамя встретилось, *помета*, *ноты*, в случае если рассматриваемая рукопись является двоезнаменником, а также *страница* и *порядковый номер* в песнопении.

От формата XML было принято решение отказаться, так как он накладывает ограничение на объем исходного файла и на скорость обработки запроса, несмотря на то, что определить рядом стоящие знамена в песнопении при его использовании проще, чем в случае с реляционной базой данных. Для повышения эффективности поиска было принято решение добавить в структуру таблицы базы данных дополнительное поле, которое будет содержать идентификаторы соседних знамен. Это расширяет возможности поиска, однако накладывает ограничение на максимальную длину поисковой последовательности (экспериментально выбрано значение 5).

Поисковая форма

В данный момент визуализация знамен на сайте осуществляется с помощью специально разработанного шрифта “Andrew Semio”, в котором каждое знамя представлено отдельным символом. Для выбора знамени был создан специальный виджет, с помощью которого можно быстро найти нужное знамя (рис. 2). В проекте используется две его разновидности, которые отличаются возможностью сохранять одно или несколько знамен в одном поле ввода.

Алгоритм поиска

Рассмотрим алгоритм поиска в случае, когда пользователь задал критерий “строгий поиск”. В данном случае ищутся все записи БД с первым символом искомой последовательности (одна запись в базе данных соответствует одному знамени). Далее в специально подготовленном поле, в котором хранятся следующие 5 знамен, ищутся оставшиеся символы поисковой строки. Если все знамена совпадают, то исходное знамя попадает в результаты поиска.

Input: поисковой запрос q, q=(q1, q2,…, qn), n<6
Output: найденные знамена Res

1. first_znam = q[0];
2. znams ← find_znams(first_znam) ;
3. **for** z in znams **do**

```
4.    for i=1,n do
5.          If q[i] != z[i-1]:
6.                is_result = False;
7.    if is_result:
8.          Res ← z;
9.  return Res;
```

В случае выбранного критерия “нестрогого поиска” сначала выполняется “строгий” поиск и далее к полученным результатам добавляются результаты неполных совпадений искомой последовательности.

Рис. 2. Виджет для выбора последовательности знамен

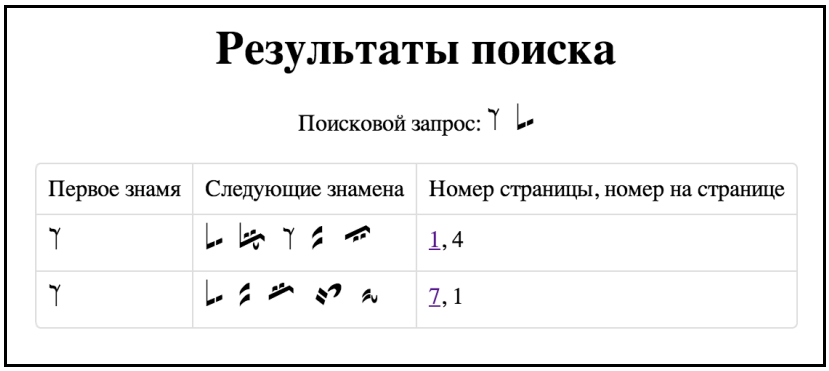
Результаты поиска

После выполнения поиска пользователю выдается список страниц, на которых были найдены искомые знамена (рис. 3). При этом найденные фрагменты подсвечиваются на странице результатов.

В ходе доработки сервиса удалось преодолеть ограничение на максимальную длину поиска в 5 знамен с помощью специального приема (дополнительной функции), которая разбивает искомую последовательность знамен на части и далее использует уже существующую функцию поиска, объединяя полученные результаты и выделяя из них идущие подряд последовательности. Для удобства пользователей в веб-сервис была также добавлена возможность выделения на странице просмотра песнопений нескольких знамен и выполнения поиска по ним.

В настоящее время осуществляется опытная эксплуатация нового сервиса, доступного по адресу http://compsemiografy.ru/project/ipsm/dvoz/. В перспективе

планируется развить поисковые процедуры, добавив в них функции перевода из современной линейной нотации, использования текста песнопений и других.



Результаты поиска

Поисковой запрос:

Первое знамя	Следующие знамена	Номер страницы, номер на странице
		1, 4
		7, 1

Рис. 3. Страница с результатами поиска по двоезнаменнику

Источники

Ирмологий — Двузнаменный ирмологий последней четверти XVII века // ГБЛ, собр. В. Ф. Одоевского № 18.

Круг пения 1884 — Круг церковнаго древняго знаменнаго пения в шести частях // Общество любителей древней письменности. СПб., 1884.

Сборник попевок — Сборник попевок (Кокизник) Соловецкого собрания середины XVII века // РНБ. Солов. собр. № 752/690.

Литература

Голубева (Даньшина) 2012 — *Голубева (Даньшина) И. В.* Семантика в музыкальных знаковых системах // Сборник тезисов и статей Российско-Германской молодежной дистанционной научной школы “Актуальные и перспективные направления создания систем, обеспечивающих семантический анализ данных в режиме реального времени” (27 сентября 2012 г.). М.: МИСиС, 2012. С. 105–108.

Голубева, Филиппович А. 2013а — *Голубева И. В., Филиппович А. Ю.* Синтаксический анализ музыкальных текстов // Доклады по компьютерным наукам и информационным технологиям. № 2: Доклады всеросс. научн. конф. “Анализ Изображений, Сетей и Текстов (АИСТ 2013)” (Екатеринбург, 4–6 апреля 2013 г.). М.: Национальный Открытый Университет “ИНТУИТ”, 2013. С. 184–192.

Голубева, Филиппович А. 2013б — *Голубева И. В., Филиппович А. Ю.* Синтаксический анализ музыкальных текстов // Новые информационные технологии в автоматизированных системах. 2013. № 16. С. 257–262.

Даньшина И., Даньшина М. 2009 — *Даньшина И. В., Даньшина М. В.* Визуализация и разработка электронного издания семиографических песнопений // Научная школа для молодых ученых “Компьютерная графика и математическое моделирование (Visual Computing)”: тезисы и доклады. М., 2009. С. 89–105.

Даньшина И., Даньшина М. 2010 — *Даньшина И. В., Даньшина М. В.* Структура и обработка древнерусских певческих рукописей // Сб. тезисов докл. “Печатные средства информации в современном обществе (к 80-летию МГУП)”. Секция “Электронные средства информации в современном обществе”. М. 2010. С. 19–21.

Даньшина М., Филиппович А. 2013 — *Даньшина М. В., Филиппович А. Ю.* Применение методов машинного перевода для анализа древнерусских музыкальных рукописей // Доклады по компьютерным наукам и информационным технологиям. № 2: Доклады всеросс. научн. конф. “Анализ Изображений, Сетей и Текстов (АИСТ 2013)” (Екатеринбург, 4–6 апреля 2013 г.). М. Национальный Открытый Университет “ИНТУИТ”, 2013. С. 77–84.

КС — *Компьютерная семиография* [Электронный ресурс]. Режим доступа: http://compsemiografy.ru/, свободный.

Разумовский — *Разумовский Д. В.* Церковное пение в России (Опыт историко-технического изложения): В 3-х вып. М., 1867, 1868, 1869. 368 с. (URL: http://www.seminaria.ru/raritet/razum_history.htm).

Смоленский 1888 — *Смоленский С. В.* Азбука знаменного нения старца Александра Мезенца. Казань, 1888. 143 с. (URL: http://www.seminaria.ru/raritet/azb_mezen.htm).

Филиппович А., Голубева 2012 — *Филиппович А. Ю., Голубева И. В.* Исследование синтаксиса семиографических песнопений // Изв. высш. учебн. заведений. Проблемы полиграфии и издательского дела. М.: Изд-во МГУП, 2012. № 6. С. 147–163.