

## **Об одном подходе к автоматизации текстологического исследования (нечеткая генеалогическая классификация)**

**С. В. Шпирко**

*Текстология, генеалогическая классификация, нечеткие множества,  
отношение порядка, нечеткая классификация*

### **On One Approach to Automation in Textology Research (A Fuzzy Genealogical Classification)**

**Sergey Shpirko**

We present a new approach for genealogical classification in textology. We formally introduce a preference relation between pairs of texts, which allows us to apply Fuzzy Set Theory for genealogical classification. We propose an algorithm of fuzzy classification of texts and illustrate it using an example.

В работе [Шпирко, 2014] предпринята попытка построения формального алгоритма для генеалогической классификации списков (вариантов текста древнего произведения), что, на взгляд автора, является актуальной целью. Необходимым условием такой классификации является построение функции отношения (в данном случае, отношения предпочтения) между списками, которое будет проводиться в два этапа.

Первый этап заключается в анализе всех возможных разночтений и их хронологическом упорядочивании. То есть, для любых произвольно выбранных чтений одного и того же места в текстах списков мы можем указать - какое из них более старшее, а какое из них более новое. Данный этап подразумевает наличие эксперта, который и будет проводить соответствующую оценку. В качестве аналогии данному этапу можно указать на процесс выбора основного списка (задание “стрелы времени”) в работе [Бородкин, 1977].

Второй этап связан уже с упорядочиванием списков относительно друг друга. Можно поставить задачу отыскания минимального элемента, как это делается в классической теории принятий решений. Необходимо заметить, что каждый список - это своего рода “слоеный пирог”, имеющий разнородный и разновременный состав. И упорядочивание списков возможно лишь нечетким образом. То есть, мы можем говорить, что один список предшествует другому лишь с некоторым порогом достоверности. Отсюда понятно, что в рамках классической теории принятия решений данная задача упорядочивания становится неопределенной, поскольку непонятно, как искать в этом случае минимальный элемент.

Теория нечетких множеств, как будет показано, дает положительный ответ на данный вопрос. На основе функции предпочтения, построенной на первом этапе, автор строит нечеткое отношение предпочтения между списками.

Далее, с помощью развитого аппарата теории нечетких множеств [Мелихов, 1990] строится формальный алгоритм классификации. В результате работы данного алгоритма проводится не только упорядочивание (нечеткое) списков друг относительно друга, но и выделение классов (нечетких) эквивалентных списков. Здесь под эквивалентностью понимается внутриклассовая минимизация неоднородности списков. При этом степень достоверности регулируется заданным порогом.

Предложенный алгоритм тестировался для классификации четырех списков полной редакции “Царева государева послания во все его Российское царство” Ивана IV Грозного [Лурье, 1954]. Многие выводы, сделанные в [Лурье, 1954], совпали с результатами работы предложенного алгоритма нечеткой классификации.

В частности:

- выделен класс эквивалентности, состоящий из двух списков (полная редакция, согласно [Лурье, 1954]);
- установлена независимость данного класса списков от третьего списка;
- установлено, что появлению списков данного класса предшествовало создание четвертого списка.

### Литература

- Бородкин, 1977 - Бородкин Л.И., Милов Л.В. О некоторых аспектах автоматизации текстологического исследования (Закон Судный людем)// Математические методы в историко-экономических и историко-культурных исследованиях: сборник статей. Москва. 1977.С.235-280.
- Мелихов, 1990 - Мелихов А.Н., Бернштейн Л.С., Коровин С.Я. Ситуационные советующие системы с нечеткой логикой. Москва. 1990. С. 272.
- Шпирко, 2014 – Шпирко С.В. Применение теории нечетких множеств к задаче генеалогической классификации в текстологическом исследовании// Историческая информатика. Барнаул. 2014. № 2 (6). в печати.
- Лурье, 1954 - Лурье Я.С. Новые списки “Царева государева послания во все его Российское царство”// Труды отдела древнерусской литературы. Москва-Ленинград. 1954.С.305-310.