

# База данных двойных звезд (BDB)

© П.В. Кайгородов

pasha@inasan.ru

© О.Ю. Малков

Институт астрономии РАН

malkov@inasan.ru

© Д.А. Ковалева

dana@inasan.ru

## Аннотация

Представлено описание разрабатываемой в Институте астрономии РАН Базы данных двойных звезд (Binary star DataBase, BDB). BDB включает в себя информацию о физических и астрометрических параметрах двойных звезд различных типов – визуальных, спектроскопических, затменных и др. Рассматриваются основные проблемы, связанные с агрегацией в единой базе данных информации о различных типах двойных, полученной из разнородных источников данных – астрономических каталогов и обзоров. Описана структура разрабатываемой базы данных, методика преобразования данных из форматов различных каталогов в формат BDB, а также основные методы кросс-идентификации объектов, упоминающихся в более чем одном источнике.

## 1 Введение

Основным методом исследования в физике является эксперимент. В случае астрофизических исследований, экспериментальная проверка заключается в сравнении выводов, следующих из проверяемой гипотезы, с результатами наблюдений. Временные масштабы, на которых протекают явления, изучаемые астрофизикой, как правило, очень велики и астрофизические явления могут быть исследованы только статистически, на основе анализа наблюдений множества объектов, находящихся на разных стадиях одного и того же процесса. Кроме того, невозможность прямой экспериментальной проверки астрофизических гипотез требует тщательного сравнения их выводов с данными, накопленными за всю историю наблюдений. Ключевой особенностью, отличающей астрономические данные от результатов физических экспериментов, также является их уникальность – измерения, проведенные в

процессе наблюдений, не могут быть повторены с точным соблюдением исходных условий. Все это приводит к необходимости хранения и обработки большого количества наблюдательного материала, что невозможно без его систематизации и создания специальных инструментов для работы с ним.

Основной проблемой при работе с астрономическими данными является их разнородность, поскольку исследователи, занимающиеся их непосредственным получением, сосредотачиваются, как правило, на решении неких узких задач, ограничиваясь измерением одного или нескольких параметров, являющихся предметом данного конкретного обзора. Другой проблемой является многовариантность полученных данных, поскольку наблюдения одних и тех же объектов могут быть проведены различными группами, с использованием различных методик и с разной точностью даже в том случае, когда измеряемые параметры идентичны. При этом сопоставление полученных результатов с результатами других групп может быть неполным, либо не проводиться вовсе. Наконец, третьей проблемой является отслеживание изменений, поскольку новые наблюдения могут как подтверждать результаты, полученные ранее, так и дополнять и уточнять их. При этом результаты любой обработки, проведенной для предыдущей версии данных, должны быть, соответственно, скорректированы.

## 2 Астрономические данные

Развитие средств наблюдения, появление автоматизированных телескопов, в том числе орбитальных, а также перевод в цифровую форму данных, полученных за более чем трехвековую историю астрономических наблюдений, привело к накоплению огромного количества наблюдательного материала, исследование и обработка которого превратилось в отдельную задачу.

Астрономические данные могут быть представлены в виде ресурсов различного вида и разных типов организации [1]. Так, в частности, можно выделить такие типичные виды ресурсов астрономических данных, как астрономические обзоры, астрономические каталоги и астрономические базы данных.

- Термин «астрономический обзор» может относиться как к наблюдательной программе, так и к типу организации астрономических данных. В этом смысле астрономическим обзором принято называть коллекцию астрономических данных об объектах, обладающую полнотой по координатам и по блеску включаемых объектов (т.е. включающую все астрономические объекты в определенном участке неба, имеющие блеск в заданных пределах). Астрономические обзоры могут представлять данные как в табличном виде, так и в виде изображений.
- Астрономический каталог – упорядоченная коллекция большого количества (обычно редуцированных, т.е. подвергнутых обработке) данных определенного типа, собранных с определенной целью, и представляемых, как правило, в табличном виде. Значительная часть существующих астрономических каталогов включена в коллекцию Страсбургского международного центра астрономических данных (CDS, <http://cdsarc.u-strasbg.fr/viz-bin/Cat>).
- Астрономические базы данных представляют собой тематические коллекции астрономических данных с более высокими уровнями структурирования. Они могут содержать данные разных видов (численные данные, изображения, библиографическую информацию, и пр.)

Примером одной из современных широко используемых астрономических баз данных может служить база данных астрономических каталогов VizieR (<http://vizier.u-strasbg.fr/viz-bin/VizieR>), позволяющая проводить комплексный поиск астрономических данных в любой выборке из почти 10 тысяч каталогов, содержащихся в CDS. В качестве примера астрономического каталога может быть приведен объединенный каталог переменных звезд (ОКПЗ, [2]) – один из наиболее востребованных мировым научным сообществом российских специализированных астрономических ресурсов. ОКПЗ содержит информацию обо всей известных переменных звездах (десятки тысяч объектов) и является наиболее полным и авторитетным источником информации об астрономических объектах этого типа.

Все каталоги, находящиеся в архиве CDS, имеют стандартный формат описания и представления данных, позволяющий включать каталог в систему VizieR. Кроме того, в рамках проекта Международной Виртуальной Обсерватории (МВО, <http://www.ivoa.net>) разработан и применяется XML-стандарт для представления и обмена табличными данными в астрономии

VOTable [3], особенностью которого является возможность работы с большими массивами данных и использования в распределенных вычислениях. В частности, одним из сервисов базы данных VizieR является представление результатов поиска по включенным в нее каталогам в формате VOTable.

Наиболее распространенным форматом для хранения и передачи изображений в архивах астрономических данных является формат Flexible Image Transport System (FITS) [4], одобренный Международным астрономическим союзом. FITS был разработан специально для научных данных и потому включает в себя метаданные, описывающие информацию о фотометрической и пространственной калибровке, вместе с метаданными исходного изображения, которые хранятся в удобочитаемой заголовке, формата ASCII. Поддержка FITS доступна во множестве языков программирования, которые используются для научных работ: C, Fortran, Java, Perl, Python, S-Lang и IDL. Офис Поддержки FITS в НАСА (<http://fits.gsfc.nasa.gov/iaufwg/iaufwg.html>) утверждает список библиотек и платформ, которые в текущее время поддерживают FITS.

По ряду причин таблицы с астрономическими данными чаще всего представлены в виде текстовых файлов ASCII. Формат этих таблиц обычно описан в отдельном файле (как правило, с именем ReadMe), где указывается, в каких позициях строк находятся данные, соответствующие тому или иному столбцу таблицы, их формат (обычно в виде форматных спецификаций языка фортран), единицы измерения, а также дополнительные данные, необходимые для правильной интерпретации содержимого таблицы. Часто астрономический каталог содержит несколько файлов с таблицами и один файл ReadMe, где приводится описание для всех файлов, а также общую информацию о каталоге, авторах, ссылки и пр. Структура файла ReadMe, как правило, соответствует некоему общепринятому стандарту (введенному CDS [5]) и может быть прочитана автоматическими средствами.

При работе исследователей с каталогами могут решаться, как правило, задачи следующих видов:

- поиск в каталоге информации об определенном объекте;
- создание выборки данных об объектах с определенными характеристиками;
- поиск корреляционных соотношений между различными параметрами, представленными в каталогах.

При решении поисковых задач требуется найти определенный объект по его известным характеристикам (идентификатору, координатам, либо

другим свойствам) в каталоге (обзоре, базе данных), где он вероятно или наверняка содержится. С целью облегчения задачи поиска объекта по идентификатору значительная часть астрономических ресурсов предоставляет несколько (как правило, не менее двух) идентификаторов для каждого объекта. Однако значительное разнообразие астрономических идентификаторов делает проблему кросс-идентификации во многих случаях весьма трудоемкой. Задача отождествления объекта по координатам, как правило, требует привлечения дополнительной наблюдательной информации для исключения риска ложных отождествлений.

Для решения многих астрономических задач требуется исследовать свойства определенных выборок астрономических объектов. Это могут быть объекты одного типа (например, переменные звезды, принадлежащие к определенному классу переменности, скажем, цефеиды), или объекты с одинаковыми или сходными наблюдательными характеристиками (например, с величиной красного смещения  $Z$  больше 2), или объекты, определенным образом локализованные в пространстве (например, расположенные в ограниченном участке небесной сферы). Для создания таких выборок из каталога выбираются строки, содержащие в соответствующих позициях значения требуемых параметров, ограниченные в соответствии с задачей исследователя.

### 3 Двойные звезды

Двойные звезды весьма многочисленны: более половины (по некоторым оценкам - более 90%) звезд образуются в двойных и кратных системах. Кроме того, двойные звезды чрезвычайно важны для определения фундаментальных (определяющих эволюционный сценарий) характеристик звезд: масс, радиусов и температур. В частности, масса звезды (параметр, практически в одиночку определяющий всю звездную эволюцию) принципиально не может быть определена из наблюдений одиночных звезд (единственное исключение – Солнце). Таким образом, двойные звезды - единственный источник данных для получения фундаментальных соотношений между звездными параметрами (соотношение «масса-светимость», соотношение «масса-радиус» и т.п.). Эти соотношения необходимы в астрономии для оценки параметров звезд, а также являются ключом к решению вопроса об образовании двойных и одиночных звезд (т.н. история звездообразования – распределение по разным параметрам звезд, образующихся в данной звездной системе). Таким образом, исследование двойных звезд является одной из наиболее актуальных задач современной астрофизики.

Двойственность звезд проявляется по-разному и может быть обнаружена, например, из визуальных позиционных наблюдений (компоненты наблюдаются как две отдельных звезды), из фотометрических наблюдений (компоненты пары сливаются в один объект, блеск которого периодически меняется из-за орбитального движения, приводящего к затмениям в системе), из спектроскопических наблюдений (орбитальное движение компонентов проявляется из эффекта Доплера - периодического смещения линий в спектре звезды), из рентгеновских наблюдений (регистрируется излучение от нагретого до сверхвысоких температур вещества, перетекающего с одного компонента на другой) и пр.

Комплексное исследование двойных звезд затруднено тем, что эти разные ансамбли двойных звезд:

- исследуются, как правило, разными научными коллективами (таким образом, публикуемые базы данных и каталоги содержат сведения лишь о двойных определенного типа);
- зачастую представляют весьма различные семейства и в пространстве параметров (так, почти все визуальные двойные - слабые маломассивные пары, а почти все спектроскопические двойные – наоборот, яркие массивные пары).

Как следствие, результаты исследований различных типов двойных звезд перекрываются весьма слабо. В результате до сих пор остается открытым такой фундаментальный вопрос как уникальность истории звездообразования, а именно: должны ли мы считать, что образование широких и тесных двойных подчиняется разным законам, притом, что существуют наблюдательные и теоретические аргументы как «за», так и «против»? Далее, получаемые для разных типов двойных калибровочные зависимости между фундаментальными параметрами звезд «сшиваются» между собой зачастую весьма неувренно.

Для проверки астрофизических гипотез, касающихся двойных звезд, необходимо получение репрезентативных выборок объектов, параметры которых удовлетворяют некоторым критериям, зачастую довольно сложным. Например, для верификации модели аккреционных дисков звезд типа SU UMa может потребоваться выборка из звезд, находящихся на стадии полуразделенной звезды с компактным компонентом (эволюционный статус), имеющих массу главного компонента порядка массы Солнца, с орбитальным периодом не более 3-х часов, демонстрирующих наличие супервспышек и имеющих наклон экваториальной плоскости к лучу зрения не более 30 градусов. Кроме того, поскольку выборка может

делаться для составления программы наблюдений, необходимо, чтобы все эти звезды были доступны для наблюдения из определенной точки в заданный период времени, и их звездные величины не выходили за пределы чувствительности используемых приборов. Имеющиеся на данный момент инструменты не позволяют проводить такие выборки автоматически, что приводит к необходимости создания специального инструментария для работы с астрономическими данными, касающимися двойных звезд.

#### 4 База данных двойных звезд

Разрабатываемая в настоящий момент в Институте Астрономии РАН база данных двойных звезд (Binary star DataBase, BDB) будет содержать данные о порядка 100 000 звездных систем с кратностью от 2 до 22. Данные, содержащиеся в BDB, планируется получить путем объединения множества каталогов, содержащих информацию о двойных и кратных звездах разных типов: визуальных двойных, спектральных, астрометрических, интерферометрических, спектроскопических, фотометрических и т.д. Из этих каталогов планируется извлечь как наблюдательные данные о координатах, собственных движениях, периодах, переменности, звездных величинах, а также астрофизические параметры – эволюционный статус, отношение масс и т.п. Данные из базы могут быть получены по запросам, включающим один или несколько индексируемых параметров, либо по идентификатору, принадлежащему какой-либо из поддерживаемых систем идентификации.

BDB разрабатывалась изначально в обсерватории Безансона (Франция), в 2008 году ее разработка была перенесена в Москву. Одновременно с переносом структура и интерфейс базы данных были существенно переработаны. Основные цели и идеи, лежащие в основе разработки BDB, изложены в [6–8]. Структура базы данных BDB схематично показана на Рис. 1. Объект, хранящийся в базе данных, может принадлежать к одному из следующих классов: Системы, Пары и Компоненты, образующих иерархию за счет связей. Системы являются наиболее высокоуровневыми объектами иерархии, каждая система ссылается на одну или несколько входящих в нее Пар. Пара представляет собой ровно два гравитационно связанных объекта, каждый из которых может быть Парой (более тесной) или Компонентом. Пара ссылается, как правило, на один или два компонента, в зависимости от имеющихся данных (в каталогах в некоторых случаях может отсутствовать информация об одном из компонентов Пары). Компонент соответствует отдельной звезде, входящей в Пару.

Каждый объект, хранящийся в базе данных,

содержит информацию о соответствующей ему сущности. Так, Пара может содержать информацию о размере большой полуоси, периоде, типе пары (визуальная двойная, спектроскопическая, фотометрическая и др.), а также прочие параметры, относящиеся к двойной звезде, как к целому. Компонент же содержит информацию о массе, эволюционном статусе, светимости и т.п. данные, относящиеся к отдельной звезде. Кроме того, каждый объект может иметь один или несколько идентификаторов в разных системах идентификации, соответственно каждый объект ссылается на одну или несколько записей в таблице идентификаторов.

Информация, содержащаяся в BDB, скомпилирована из множества различных каталогов, содержащих разные данные. Часть этих данных может пересекаться, при этом информация из одного каталога может противоречить информации из другого. Прежде всего это связано с различием в точности и методике измерений, однако также может быть следствием ошибок, содержащихся в исходных каталогах. Выбор между несколькими значениями одного и того же параметра, предлагаемыми различными каталогами часто не может быть сделан автоматически, поскольку причина расхождения часто вызвана природой самого объекта. Например, светимость переменных звезд может варьироваться в пределах нескольких звездных величин как периодическим, так и аperiodическим образом, соответственно, измерения, проведенные в разное время могут давать существенно различный результат. Соответственно, чтобы избежать потери важных данных, в базе данных BDB хранятся одновременно все версии данных, относящихся к одной и той же сущности (системе, паре или компоненту). При этом число хранимых объектов, соответствующих, например, некой паре, будет равно числу каталогов, в которых эта пара упоминается.

Для связи между различными версиями одного объекта используются идентификаторы, а также кросс-идентификационная таблица, также входящая в состав BDB. Например, два объекта, ссылающихся на один и тот же идентификатор всегда являются версиями одного и того же объекта. Кроме того, поскольку каждый объект может ссылаться на несколько идентификаторов, возможна кросс-идентификация между объектами, не ссылающимися непосредственно на один и тот же идентификатор. Действительно, если объект А ссылается на идентификаторы id1 и id2, объект В на id2 и id3, а объект С на id3, то все они являются версиями одного объекта. Кросс-идентификационная таблица используется в случае, когда информация о соответствии идентификаторов друг другу получена не из

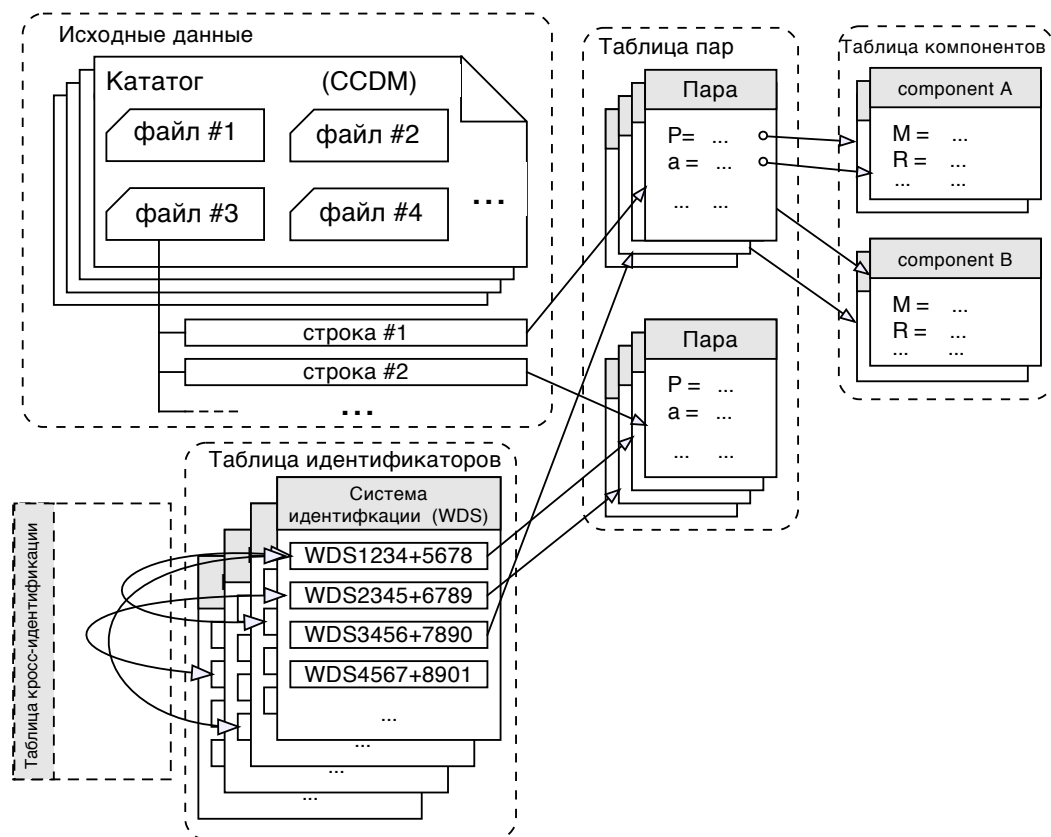


Рис. 1: Структура базы данных BDB

каталогов, а из каких-либо внешних источников, не входящих в BDB.

Многовариантность астрономических данных может приводить к ситуациям, когда один и тот же объект в различных источниках относится к различным классам. Так, например, компонент пары, присутствующей в одном из источников, может сам фигурировать как пара в других источниках. Это связано с тем, что при повышении точности наблюдений часть объектов, считавшихся ранее точечными, разделяются на компоненты и входят в каталоги уже как пары. Одним из принципов построения BDB является сохранение максимально большого количества информации о двойных звездах, что приводит к необходимости хранения противоречивых данных. В случае кросс-идентификации объектов, относящихся к различным классам (например – Компонента и Пары) будут сохранены обе версии объекта, а наличие таблиц кросс-идентификации позволяет отображать всю найденную информацию об объекте по запросу пользователя.

Соответственно, при поиске по заданным пользователем критериям будут найдены все объекты, параметры хотя бы одной из версий которых удовлетворяют заданному критерию. Однако, поскольку все версии одного и того же объекта связаны, в поисковой выдаче будут пока-

заны все версии найденных объектов, а не только те, что удовлетворяют критериям поиска. Это позволяет исследователю получить более полный объем информации и уже самостоятельно провести отбор необходимых данных в соответствии с собственными потребностями.

## 5 Реализация BDB

Создаваемая база данных реализована на основе свободно распространяемой СУБД PostgreSQL, доступ к которой осуществляется через веб-интерфейс, написанный на языке Python с использованием свободно распространяемого фреймворка Nagare (<http://www.nagare.org>). Фреймворк Nagare основан на использовании продолжений (continuations), что позволяет создавать веб-приложения, не заботясь о сохранении состояния программы (содержимого переменных и полей объектов) между перезагрузками страницы. Это существенно облегчает и ускоряет разработку, а также упрощает отладку сложных веб-приложений. Интегрированная в Nagare система ORM (object-relational mapping) Elixir (<http://elixir.ematia.de/>), основанная на SQLAlchemy (<http://www.sqlalchemy.org/>) позволяет работать с записями в базе данных как с

объектами. Это упрощает доступ к данным и позволяет абстрагироваться от конкретной реализации SQL. Кроме того, фреймворк Nagare позволяет прозрачно интегрировать в веб-страницы элементы, построенные на технологии Ajax, что дает возможность обновлять часть информации на странице без ее полной перезагрузки, снижая нагрузку на сервер и делая интерфейс более отзывчивым на действия пользователя. Также фреймворк Nagare включает в себя модуль прозрачной трансляции кода, написанного на Python в код Javascript, исполняемый на стороне клиента. При помощи такой трансляции становится возможным вынести, например, проверку корректности вводимых данных в клиентскую часть, исполняемую непосредственно в браузере пользователя.

## 6 Источники данных BDB

Наполнение BDB данными представляет собой очень сложную задачу и включает в себя следующие этапы:

1. Подбор источников, содержащих нужную информацию
2. Построение алгоритмов, преобразующих данные выбранных источников в формат, соответствующий структуре BDB
3. Преобразование данных и пополнение BDB
4. Кросс-идентификация новых данных
5. Выборочная проверка и тестирование новых данных

На этапе 1 основной трудностью является выбор необходимых источников информации, поскольку существует множество различных каталогов и обзоров, содержащих информацию о двойных звездах. В них, как правило, применяются различные системы идентификации, кроме того, они имеют, как правило, разную структуру.

На этапе 2 главной проблемой является то, что данные, необходимые для пополнения BDB, могут не присутствовать в отобранных источниках в явном виде. Например, один и тот же параметр может быть дан в различных единицах измерения в различных каталогах, либо для его получения требуется проведение вычислений, подчас довольно нетривиальных. Кроме того, большинство каталогов не было рассчитано на машинную интерпретацию и часть важных данных может быть дана в них в виде текстовых пояснений. Таким образом, для извлечения нужной информации зачастую необходимо построение сложных эвристических

алгоритмов, анализирующих как табличные данные фиксированного формата, так и текстовую информацию.

Этап 3 проводится автоматически, но, учитывая большой объем данных и сложность применяемых алгоритмов, может быть весьма затратным в плане вычислительных ресурсов.

Этап 4 также требует нетривиального подхода, поскольку, несмотря на то, что кросс-идентификация частично проведена авторами используемых каталогов и обзоров, она не всегда свободна от ошибок. Имея доступ к большому количеству источников, возможно реализовать нахождение случаев ошибочной кросс-идентификации, однако это требует создания довольно сложных эвристических алгоритмов и специального инструментария для анализа данных. Кроме того, возможно проведение кросс-идентификации объектов по различным критериям, таким, как близость координат, собственных движений, светимостей и т.п., однако такой подход также требует создания весьма сложных алгоритмов. В рамках BDB планируется ввести собственную систему идентификации, покрывающую все известные классы двойных и кратных звезд, а также создать таблицы кросс-идентификации, устанавливающие соответствие между идентификаторами BDB и другими системами идентификации.

На этапе 5 возможна как автоматическая, так и полу-автоматическая верификация данных. Для автоматической проверки требуется разработка критериев «корректности» вычисляемых значений. В случае обнаружения ошибочных данных требуется коррекция алгоритмов, применяемых на этапах 2-4.

Процесс наполнения BDB данными выглядит следующим образом:

1. Скачивание архива с данными и его распаковка. В настоящий момент для удобства отладки реализована возможность автоматического скачивания данных с сервера CDS по CDS-идентификатору каталога.
2. Распознавание содержимого файла ReadMe и определение формата данных, содержащихся в текстовых файлах таблиц. В случае отсутствия или нестандартного формата файла ReadMe возможно задание формата вручную. Пока реализована поддержка импорта данных только из текстовых таблиц.
3. Описание алгоритмов, выделяющих индексированные данные (такие, как координаты, светимости, собственные движения и пр.) из текстовых файлов. Алгоритмы (в общем случае) представляют собой фрагменты кода на языке Python, которые впоследствии

- будут скомпилированы и исполнены в процессе распознавания каждой строки. Для ввода кода предусмотрен специальный web-интерфейс, позволяющий автоматизировать наиболее рутинные действия и предоставляющий подсказки, облегчающие написание кода. Также в интерфейсе реализована возможность сохранения промежуточных версий введенного кода, что существенно упрощает отладку и тестирование.
4. Пробное распознавание первых 100 строк таблицы. Этот шаг служит проверки кода и вывода диагностической информации в случае наличия ошибок.
  5. Полное распознавание таблицы, проводится автоматически и может (для больших каталогов) занимать до часа времени. В процессе распознавания web-интерфейс отображает степень завершенности задачи, ожидаемое время окончания и число найденных ошибок. В случае необходимости распознавание может быть прервано. Кроме того, процесс распознавания прерывается автоматически в случае нахождения большого числа ошибок (>1000).
  6. Обзор ошибок распознавания (если такие были найдены). В случае необходимости возможен возврат к шагам 2-5.

При наличии в каталоге нескольких таблиц шаги 2-6 повторяются для каждой из них. После распознавания всех таблиц возможно проведение выборочной ручной проверки правильности данных при помощи web-интерфейса, либо автоматическая проверка при помощи скриптов, обращающихся напрямую к СУБД.

## 7 Заключение

В работе описаны цели, процесс создания и содержание Базы данных двойных звезд (Binary star DataBase, BDB). В процессе разработки BDB были решены следующие задачи:

1. Создан список оригинальных каталогов и баз данных, содержащих эмпирические данные для двойных и кратных звезд различных типов.
2. Проведена большая работа по созданию, уточнению и доработке схем классификации двойных звезд различных типов, как по наблюдательным проявлениям, так и по эволюционному статусу. Разработана методика сквозной системы идентификации BDB, пригодная для идентификации двойных и кратных звезд всех известных типов.

3. Проведен анализ задач, для решения которых необходимо построение базы данных двойных и кратных звезд. На основе проведенного анализа разработана схема базы данных.
4. Разработан web-интерфейс для импорта данных в BDB, проведен тестовый импорт нескольких каталогов двойных звезд.
5. Разработан прототип пользовательского web-интерфейса для поиска данных в BDB.

База данных двойных звезд будет являться мощным инструментом для исследования этого типа астрономических объектов, так как основная цель ее создания - интеграция и обеспечение доступа к данным обо всех типах двойных звезд. Каталоги двойных звезд различных типов будут включены в BDB, с базами данных (как специализированными, так и общего назначения) будут установлены связи. Тестирование бета-версии BDB намечено на осень-зиму 2012.

## Благодарности

Работа выполнена при поддержке грантов РФФИ 10-02-00426, 10-07-00342, 11-02-00076, 12-02-00047, 12-07-00528, при поддержке Федерального агентства по науке и инновациям (грант 02.740.11.0247), Программы Президиума РАН «Поддержка ведущих научных школ» (грант НШ-3602.2012.2), а также в рамках реализации ФЦП «Научные и научно-педагогические кадры инновационной России» на 2009 - 2013 годы.

## Список литературы

- [1] Д.А. Ковалева, О.Б. Длужневская. Российские ресурсы данных и их интеграция в структуру Международной Виртуальной Обсерватории. В сб.: «Электронные библиотеки: перспективные методы и технологии, электронные коллекции». Труды XIII Всероссийской научной конференции RCDL'2011, Воронеж, 19-22 октября 2011 года. Ред. Л.А.Калининченко, А.В.Сычев, С.А.Ступников. – ВГУ: 2011, с. 317
- [2] N.N. Samus, O.V. Durlevich, E.V. Kazarovets, N.N. Kireeva, E.N. Pastukhova, A.V. Zharova, et al. General Catalog of Variable Stars (GCVS database, Version 2012Feb), CDS: B/gcvs, <http://www.sai.msu.su/gcvs/gcvs/>
- [3] F. Ochsenbein et al., VOTable Format Definition Version 1.2, IVOA Recommendation 2009-11-30, 2009,

<http://www.ivoa.net/Documents/VOTable/20091130/REC-VOTable-1.2.html>

- [4] W.D. Pence, L. Chiappetti, C.G. Page, R.A. Shaw and E. Stobie, Definition of the Flexible Image Transport System (FITS), version 3.0, *Astron. Astrophys.* 2010, 524, A42
- [5] F. Ochsenbein, Standards for Astronomical Catalogues, Version 2.0, 2000, <http://cdsweb.u-strasbg.fr/vizier/catstd/catstd-3.1.htm>
- [6] O. Malkov, E. Oblak, B. Debray, The Binary Star Database – BDB, 2009, in ADASS XVIII Conference, eds. David Bohlender, Daniel Durand, Patrick Dowler, Québec, Nov 2008, *ASP Conf. Ser.*, 411, 442
- [7] Kaygorodov P., Debray B., Kolesnikov N., Kovaleva D., Malkov O., The new version of Binary star database (BDB), *Baltic Astronomy*, 21, N 3, 2012, in press
- [8] О.Ю. Малков, П.В. Кайгородов, Э. Облак, Б. Дебрэ, БДБ: база данных о двойных звездах, *Динамика сложных систем*, 2010, 4, №2, 48

## **Binary star database (BDB)**

Pavel Kaygorodov, Oleg Malkov, Dana Kovaleva

### **Abstract**

BDB is developed at the Institute of Astronomy RAS, and contains information on physical and positional parameters of binary stars, belonging to various observational types: visual, spectroscopic, eclipsing, etc. Problems of information aggregation on binaries of various types (obtained from heterogeneous sources of data - astronomical catalogues and surveys) aggregation in a unified database are considered in the paper. Database structure, principles of data transformation from original catalogs' formats into the BDB format, as well as basic cross-identification methods for multi-source objects are also described.