Cross-Matching of Very Large Catalogs

Maxim Martynov[1], Daniil Bodryagin[1]

[1] *RI Mykolaiv Astronomical Observatory*

Introduction

Cross-identification is a powerful tool which is used for solving many astrometric and astrophysical problems. Number of programs and Web services has been developed for this in frames of Virtual Observatory (VO). Unfortunately, their functionality and options are limited by:

- the list of directories which are available for identification ;

- strictly specified object identification algorithms;

- certain restrictions and difficulties in identification and loading of large (hundreds of thousands or more stars) custom object list.

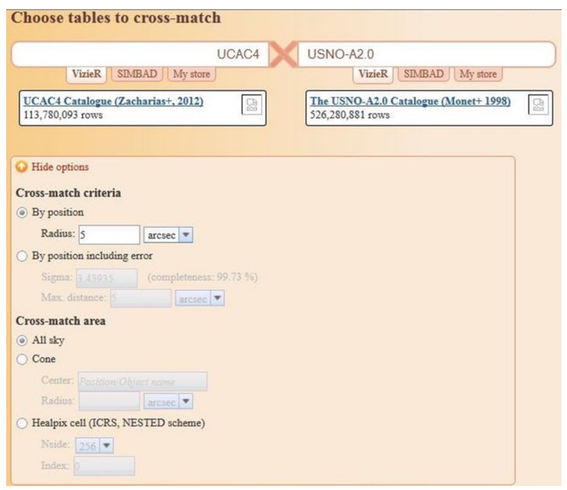
Catalogs & VO

VizieR	14,000 catalogs
CDS	500 Tb
UCAC4	110M records
XPM	300M
2MASS	450M
USNO A2.0	500M
SDSS8	800M

AMK	100K
Zod	3M

CDS X-Match Service

http://cdsxmatch.u-strasbg.fr/xmatch



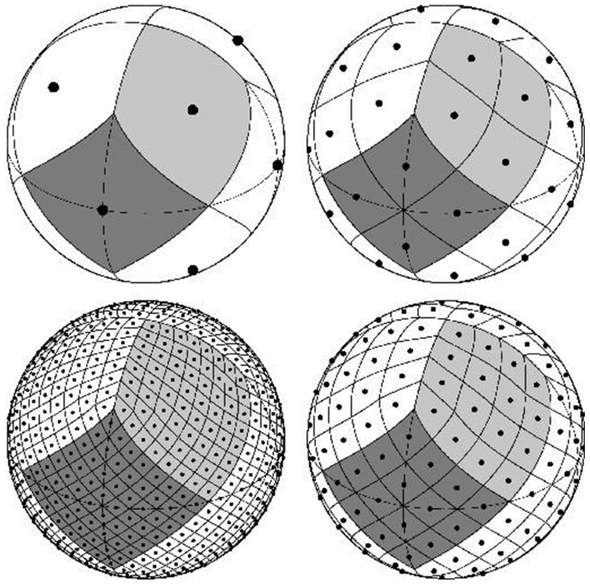
"ACrId" (Astronomical CRoss-IDentification)

Program of extra-large astronomical catalogs pixelation and cross-identification "ACrId" (Astronomical CRoss-IDentification) is used to divide the celestial sphere on small plots of equal area (pixels) for the following appropriate grading of individual pixels of modern star catalogs and lists, as well as to find in them shared objects (cross-matching, cross-identification). The software package "ACrId" allows you to connect an arbitrary number of directories with any format and carry out multi-stage objects cross-identification .

HEALPix

Hierarchical Equal Area isoLatitude Pixelization of a sphere

http://healpix.jpl.nasa.gov



NSIDE	numbers	resolution
1	12	~59º
2	48	~29º
64	49 152	~55′
128	196 608	~27′
8192	805 306 368	~26″

Data Organization

Linux; ReiserFS; 300 Gb (CMC-11, UCAC3,4, USNO-A2.0, XPM, of 4 Tb)

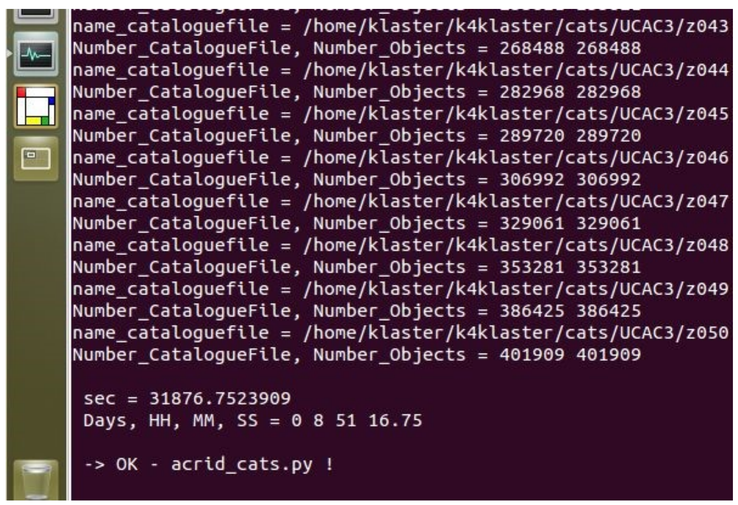
Pixelation of a new catalog does not require new routines or procedures. It does not even require any programming skills at all: it is enough to fill a standard description form. Usually there are three files:

1) General catalog description;

2) Separate catalog files list (original, before pixelation);

3) Format description files (one, if the format is common for all catalog files).

Fragments of UCAC-4 catalog description files (1 and 3):

```
path_to_source    = /home/klaster/k4klaster/cats/UCAC4/u4b/
default_pars      = ucac4_parameters.cfg
list_of_catsfiles = ucac4_list.cfg
briefly           = - u4 - - ucac4
original          = 1
eqinox            = 2000.
epoche            = 2000.,individual
is_pm             = 1
...
file_format       = bin
heading           = 0
# "шапки" в бинарном файле, разумеется, нет
byteorder         = little-endian
bin_string        = IIHHBBBBBBBBHHhhbbIhhhBBBBBBhhhhhhbbbbBIBBIHI
bin_N_byte        = 78
N_str             = 1
separator         = -
id_cat            = 1 43          -         *1          1          182895114  -          rnm
RA2000            = 1 1           mas       *1          14         1295999997 -          ra
SPD2000           = 1 2           mas       *1          57955      647908447  -          spd
e_RA              = 1 8           mas       +128        1          255        255        sigra
e_DE              = 1 9           mas       +128        1          255        255        sigdc
epocheRA          = 1 13          year      *0.01,+1900 1947.07    2004.38    -          cepra
epocheDE          = 1 14          year      *0.01,+1900 1947.00    2004.38    -          cepdc
pmRA              = 1 15          mas/yr    *0.1        -3260.3    3276.7     0          pmrac
pmDE              = 1 16          mas/yr    *0.1        -3157.4    3276.7     0          pmdc
e_pmRA            = 1 17          mas/yr    +128,*0.1   0.3        50         50         sigpmr
e_pmDE            = 1 18          mas/yr    +128,*0.1   0.3        50         50         sigpmd
Bmag_APASS        = 1 29          mag       *0.001      -1.087     20         20         apasm
Vmag_APASS        = 1 30          mag       *0.001      -1.440     20         20         apasm(2)
```

Pixelation process example for UCAC-3, catalog files from "z043" to "z050" (OC Linux, Ubuntu 12.4):