NDAC/LO/140
1991 November 29

Proposed format for exchange of astrometric catalogues (Version 1)

L. Lindegren
Lund Observatory

Introduction

This note describes a tape format suitable for the exchange of astrometric (and other) data for single stars. It is primarily intended for the comparison of successive sphere solution results obtained in FAST and NDAC. However, it may also be useful for transferring corresponding data to/from INCA and TDAC.

Because of the relatively large quantity of data (some 30 Mb for a full catalogue of 120 000 stars), high-density magnetic tape (6250 bpi) seems to be the only practicable medium. Moreover, to permit transfer between different kinds of computer, formatted (ASCII) data must be used. The format described below tries to include all the relevant information with a sufficient degree of accuracy (corresponding to a small fraction of the expected statistical errors), yet with reasonable compactness.

The data are arranged in logical records having a fixed length of 232 characters (bytes). There is a single *header record*, followed by one *data record* per star (or IC entry), identified by its HIC number. There is no requirement on the completeness of the catalogue (i.e., any subset of it may be exchanged), nor on the ordering of the data records. (It is envisaged that any computer program that will use the catalogue must store the data as an unformatted, direct-access file, using the HIC number either directly as the record number, or via an index table kept in the internal memory.)

The logical records are written on tape in blocks of 100 records (23200 characters). All blocks must have the same length, i.e., the last block may have to be padded with empty records to the full length. The contents of the 'empty' records may in fact be undefined, since the total number of valid data records is specified in the header record.

In the tables below 'POS' refers to the character positions in the record, while 'FMT' is the Fortran descriptor of the datum. The 'NAME' column contains a mnemonic designation which may be used for the corresponding Fortran (character, integer or real) variable; it is included here mainly as an easy reference to the various items.

Blank fields are only permitted (when relevant) in the alphanumeric fields (`A*`). All numeric fields (`I*`, `F*`) must be assigned a value. Undefined numeric items are set to 0, except the colour index, which is set to 99 if unknown.

The present Version 1 of the proposed format is circulated for comments. It is expected that a second version will be needed before the first exchange of data can take place.

Header record format

The header record is the first logical record in the first block on the tape file. The first two items (`lrec` and `lblk`) are included to permit physically reading the tape without knowing the detailed format. The version number (`idvers`) will be incremented for each new version of this format. The record and block lengths may be different in future versions, but the intention is never to change the format of the first three items.

HEADER RECORD			
NAME	POS	FMT	DESCRIPTION
`lrec`	1–6	`I5,1X`	= 232 Length of a logical record in bytes
`lblk`	7–12	`I5,1X`	= 23200 Length of a physical record (block) on the tape
`idvers`	13–16	`I3,1X`	= 1 Version number of this format
`nstars`	17–24	`I7,1X`	Number of stars (entries) in the catalogue
`source`	25–40	`A16`	Identification of the source of the data; should typically begin with 'FAST', 'NDAC', etc
`date`	41–52	`A12`	Date of production of the catalogue (or tape); preferred format is `YYYY.MM.DDbb` (`b` = blank space)
`rframe`	53–60	`A7,1X`	Reference frame: 'EQU2000' or 'ECL2000'
`remark`	61–232	`A172`	Remark (may be blank, i.e., filled with spaces)

Data record format

Following the header record, there are `nstar` data records as described in the table below. Most of the items are self-explanatory. The descriptions apply to the case of equatorial coordinates (`rframe` = 'EQU2000'), but the format is identical in ecliptical coordinates (making everywhere the substitutions $\alpha \leftarrow \lambda$ and $\delta \leftarrow \beta$).

Although the output catalogues will normally have a single epoch for the positions of all the stars, it is proposed to give individual epochs in the data records. There are two reasons for this. Firstly, some stars in the IC (namely, those without a known proper motion) have their positions referred to an earlier epoch, which is then required for a correct comparison with the Hipparcos results. Secondly, it is not entirely clear that a single common epoch will be the best choice for the output catalogues, depending on the actual observing histories of the stars and possible modifications to the programme in the course of the mission. It seems advisable, therefore, to keep the option of having different epochs for different stars. Note however that the epoch can only be given to the nearest 0.01 yr. A few stars have very significant proper motion in 0.005 yr. Thus it must be ensured that the stated epoch is the exact epoch for the given position coordinates, and not a rounded approximation to the actual epoch. The unit for the epoch (and proper motion) is the Julian year, equal to 365.25×86400 s.

The position coordinates `alpha`, `delta` are expressed in radians. It does not matter if they are in the interval $[-\pi, \pi]$ or $[0, 2\pi]$; the only requirement is that they are numerically less than 10 (to avoid format overflow).

The correlation coefficients are always in the range $[-1, +1]$.

DATA RECORD			
NAME	POS	FMT	DESCRIPTION
idstar	1–6	I6	Star identification number (HIC)
alpha	7–20	F14.10	α = Barycentric right ascension at epoch [rad]
delta	21–34	F14.10	δ = Barycentric declination at epoch [rad]
parlax†	35–44	F10.2	π = Parallax [mas]
pma†	45–54	F10.2	$\mu_\alpha \cos\delta$ = Proper motion in α [mas/yr]
pmd†	55–64	F10.2	μ_δ = Proper motion in δ [mas/yr]
radvel†	65–71	F7.1	Radial velocity [km/s]
epoch	72–78	F7.2	Epoch in years from J2000 (e.g., −9.00 for J1991.0)
hpmag	79–85	F7.3	H_p = Hipparcos magnitude [mag]
btmvt	86–92	F7.3	$B_T - V_T$ = Colour index in Tycho system [mag] (set to 99.00 if unknown)
nobs†	93–96	I4	Number of observations (RGC's) on the star used in the solution
npar†	97–98	I2	Number of parameters solved for the star = 0, 2 (α,δ), 3 (α,δ,π), 4 $(\alpha,\delta,\mu_\alpha,\mu_\delta)$ or 5
istat(1:4)†	99–122	4I6	Status variables (see below)
sigma(1:5)†	123–162	5F8.2	$\sigma_\alpha \cos\delta$, σ_δ, σ_π, $\sigma_{\mu\alpha} \cos\delta$, $\sigma_{\mu\delta}$ (in this order) = Standard deviations in [mas] and [mas/yr]
corr(1:10)†	163–232	10F7.3	$((\rho_{ij}, i = 1, j-1), j = 2, 5)$ = Correlation coefficients for α, δ, π, μ_α, μ_δ, taken in this order (i.e., $\rho_{\alpha,\delta}$, $\rho_{\alpha,\pi}$, $\rho_{\delta,\pi}$, $\ldots$, $\rho_{\mu\alpha,\mu\delta}$)

† set to 0 if undefined or irrelevant for the catalogue

The integer status variables `istat` give information about the status of the star in the sphere solution. They may be defined differently in FAST and NDAC, depending on the available information and what is considered desirable to communicate. The integers must fall between −9999 and 99999 to avoid format overflow. Also, they should be defined in such a way that a meaningful selection of stars can be made by taking subsets in which the status variables fall within certain limits.

The status variables to be used for the NDAC sphere solutions are:

$$\texttt{istat}(1) = \pm\text{run number}$$
$$\texttt{istat}(2) = 100 \times \text{gaussianized GOF}$$
$$\texttt{istat}(3) = 100 \times \Delta\chi^2$$
$$\texttt{istat}(4) = \text{(spare)}$$

The run number is a sequential numbering of all NDAC sphere solutions. Positive/negative sign is used for primary/secondary stars. The second variable is a χ^2-type goodness-of-fit, transformed to be approximately $N(0,1)$ for well-behaved stars. Stars with $\texttt{istat}(2) > 400$ (say) may be suspected binaries. A small $\Delta\chi^2$ indicates that the slit errors are not uniquely determined.

FAST may choose to define an entirely different set of status variables. In any case, a short description of how to interpret them should be given with the catalogue.