

Format d'un fichier Heredis

Loïc Fejoz

29 juillet 2003

Table des matières

1	Notations	3
2	Structure du fichier	4
2.1	Structure de l'en-tête	4
2.2	Liste des tables	4
2.2.1	Particules des noms	5
2.2.2	Liens	5
2.2.3	Liens avec Médias	5
2.2.4	Sources	5
2.2.5	Liens sources-événements	5
2.2.6	adresses	5
2.2.7	Lieux	6
2.2.8	Noms	6
2.2.9	Evenements	6
2.2.10	Individus	6
2.2.11	Médias	6
2.2.12	Unions	7
3	Structure des tables	8
3.1	en-tête	8
3.2	Table des Tailles	8
3.3	Table des IDs	8
3.4	Comment retrouver un élément connaissant son ID ?	9
4	Format des éléments	10
4.1	Adresse	10
4.2	Evenement	11
4.2.1	Constantes	11
4.2.2	Structure	12
4.3	Individu	13
4.3.1	Constantes	13
4.3.2	Structure	14
4.4	Lieu	15

4.5	Lien	15
4.5.1	Constantes	15
4.5.2	Structure	16
4.6	Lien vers une source	16
4.7	Lien vers un média	16
4.8	Média	17
4.9	Nom	17
4.10	Sosa	17
4.11	Source	18
4.11.1	Constantes	18
4.11.2	Structure	18
4.12	Union	19
4.13	IdxDoc	19
4.14	TBLink-IdxDE	19
4.15	TBLink-IdxVERS	19
4.16	TBEvent-Idx	19
4.17	Date	19
4.18	Dates de création et de modification	20

Chapitre 1

Notations

STRING	Chaine de caractère terminée par un caractère nul.
STRING[X]	une chaine contenant au maximum X caractères
INT	entier sur 4 bytes.
SHORT	entier sur 2 bytes.
CHAR	entier sur 1 byte.
...	le reste des données

Par défaut, tout est en little-endian sauf indication contraire.

Sans indication, les entiers sont non signés.

Les valeurs sont en décimales sauf les nombres notés '0x??' qui sont en hexa-décimal.

Chapitre 2

Structure du fichier

Un fichier est constitué d'une en-tête suivi de tables.

Apparemment il n'y a pas d'ordre pour l'apparition des tables. De même toutes ne sont pas forcément présentes et ceci varie d'une version à une autre (comme par exemple la table des Sosas).

2.1 Structure de l'en-tête

L'en-tête du fichier fait 2380 bytes dont :

en 0x00 :	CHAR	0xCO (Code café ?!)
	CHAR	0xDE
	CHAR	0xCA
	CHAR	0xFE
...		
en 0x7C :	STRING[0x20]	Nom de la généalogie
...		
en 0x9C :	STRING[0xFF]	Commentaire
...		
en 0x4C :	INT	dernière ID attribué
...		
en 0x19C :	STRING[0x20]	Nom du premier champ utilisateur
	STRING[0x20]	Tag gedcom utilisé lors de l'export
	idem pour les 9 autres champs	
...		
en 0x92C :	STRING[0x20]	Version de Heredis.
...		

2.2 Liste des tables

On peut trouver les tables présentées dans le tableau suivant. Néanmoins, elles ne figurent pas nécessairement toutes dans un fichier.

2.2.1 Particules des noms

Les tables correspondent à la liste des particules que reconnaît Heredis.

PartNameTbl	Table des particules
PartNameTbl-ItemSize	Table des tailles des éléments de PartNameTbl

2.2.2 Liens

La table contient les liens entre un individu ou un événement et un individu.

TBLINK	Table des liens (cf. 4.5)
TBLINK-IDList	Table des IDs de TBLink
TBLINK-ItemSize	Table des tailles pour TBLink
TBLINK-IdxDE	ID de, ID vers, ID du lien (cf. 4.14)
TBLINK-IdxVERS	ID vers, ID de, ID du lien (cf. 4.15)

2.2.3 Liens avec Médias

Liens entre une source ou un individu et un média.

TBMedia-IdxOwner	?
TBMedia-IdxOwner-ItemSize	Table des tailles pour TBMedia-IdxOwner
TH5LinkMedia-IDList	Table des IDs pour TBMedia-IdxOwner

2.2.4 Sources

TH5Doc	Table des sources (cf. 4.11)
TH5Doc-IDList	Table des IDs de TH5Doc
TH5Doc-ItemSize	Table des tailles pour TH5Doc

2.2.5 Liens sources-événements

TH5LinkDoc	Table des liens sources - Evenements (cf. 4.6)
TH5LinkDoc-IDList	Table des IDs de TH5LinkDoc
TH5LinkDoc-ItemSize	Table des tailles pour TH5LinkDoc
TBDoc-IdxDoc	contenant les 3 IDs d'un liens (cf. 4.13)

2.2.6 adresses

TH5TableDicoAdresses	Table des adresses (cf. 4.1)
TH5TableDicoAdresses-IDList	Table des IDs de TH5TableDicoAdresses
TH5TableDicoAdresses-ItemSize	Table des tailles pour TH5TableDicoAdresses

2.2.7 Lieux

TH5TableDicoLieux	Table des Lieux (cf. 4.4)
TH5TableDicoLieux-IDList	Table des IDs de TH5TableDicoLieux
TH5TableDicoLieux-ItemSize	Table des tailles pour TH5TableDicoLieux
TBPlace-IdxPlace	contient les IDs dans l'ordre d'apparition dans le dictionnaire des lieux
TBPlace-IdxCode-Match	?
TBPlace-IdxPlace-Match	?
TBNom-IdxCode	?

2.2.8 Noms

TH5TableDicoNoms	Table des Noms (cf. 4.9)
TH5TableDicoNoms-IDList	Table des IDs de TH5DicoNoms
TH5TableDicoNoms-ItemSize	Table des tailles pour TH5TableDicoNoms

2.2.9 Evenements

TH5TableEvenements	Table des événements (cf. 4.2)
TH5TableEvenements-IDList	Table des IDs de TH5TableEvenements
TH5TableEvenements-ItemSize	Table des tailles pour TH5TableEvenements
TBEvent-Idx	Evenement dans l'ordre des IDs des individus (cf. 4.16)

2.2.10 Individus

TH5TableIndividus	Table des individus (cf. 4.3)
TH5TableIndividus-IDList	Table des IDs de TH5TableIndividus
TH5TableIndividus-ItemSize	Table des tailles pour TH5TableIndividus
TBPeople-IdxAlpha	IDs des individus en ordre alphabétique croissant
TBPeople-IdxAlphaR	IDs des individus en ordre alphabétique décroissant
QuickList	ID fils, ID père, ID mère
QuickList-Childs	ID parent, ID enfants
SOSATbl	Table des sosa (cf. 4.10)
CcheckedPeoples	une liste d'IDs d'individus

2.2.11 Médias

TH5TableMedias	Table des médias (cf. 4.8)
TH5TableMedias-IDList	Table des IDs de TH5TableMedias
TH5TableMedias-ItemSize	Table des tailles pour TH5TableMedias

2.2.12 Unions

TH5TableUnion	Table des unions (cf. 4.12)
TH5TableUnion-IDList	Table des IDs de TH5TableUnion
TH5TableUnion-ItemSize	Table des tailles pour TH5TableUnion
TBUnion-IdxHusb	ID du mari, ID de l'union
TBUnion-IdxSpouse	Id de la femme, ID de l'union

Chapitre 3

Structure des tables

3.1 en-tête

L'en-tête d'une table est constituée de 80 bytes constitués ainsi :

CHAR	0xCO (Code café ? !)
CHAR	0xDE
CHAR	0xCA
CHAR	0xFE
STRING[40]	Nom de la table
STRING[12]	Taille d'un item
STRING[12]	Nombre d'items
STRING[12]	Taille des données de la table (en bytes)

3.2 Table des Tailles

Les tables dont les noms se terminent notamment par 'ItemSize' contiennent l'offset en bytes du déplacement à effectuer entre le début des données de la table et le début de l'enregistrement. Le premier enregistrement est donc toujours zéro.

Chaque élément est un INT.

3.3 Table des IDs

Les tables dont les noms se terminent notamment par 'IDList' contiennent le lien ID - index dans la table. Ces tables semblent contenir les IDs dans l'ordre croissant...

Chaque item est donc constitué ainsi :

INT	ID
INT	Index

3.4 Comment retrouver un élément connaissant son ID ?

1) chercher dans la table des IDs l'index de cet élément (recherche linéaire ou dichotomique, etc.)

2) si la taille des éléments est variable alors trouver l'offset dans la table des Tailles de l'index-ième élément ainsi que son suivant afin d'avoir la taille. L'accès est direct puisque le n-ième élément est à la position $n * \text{taille}$ d'un item.

3) extraire de la table principale les bytes compris entre les deux offsets précédents et les décoder ou bien extraire les bytes à la position $n * \text{taille}$ d'un item.

Chapitre 4

Format des éléments

4.1 Adresse

INT	ID
INT	date de création (cf. 4.18)
INT	date de modification (cf. 4.18)
INT	ID de l'union
INT	ID du mari
INT	ID de la femme
SHORT	booléen correspondant à "Toutes ces informations sont privées"
CHAR	?
STRING	Contact
STRING	adresse ligne 1
STRING	adresse ligne 2
STRING	Code postal
STRING	Ville
STRING	Pays
STRING	téléphone
STRING	Fax
STRING	Email
STRING	Adresse Web
STRING	Région
...	

4.2 Evenement

4.2.1 Constantes

4	naissance
8	baptême
12	décès
16	diplôme
6	inhumation
24	acquisition
34	adoption
1	autre Baptême
9	baptême adulte
2	bar Mitzvah
3	bas Mitzvah
5	bénédictio
10	confirmation
35	crémation
25	décoration
13	émigration
26	en vie
20	homologation de testament
17	immigration
18	naturalisation
19	ordination
32	première communion
23	profession
7	recensement
30	résidence
21	retraite
27	service militaire
31	testament
33	titre
28	vente d'un bien
29	voyage
0	baptême LDS
11	confirmation LDS
14	dotatio
22	liens parental LDS
61	union
54	divorce
68	mariage religieux
55	demande de divorce
66	adoption

67	anulation du mariage
58	bans
59	contrat de mariage
60	certification de publication des bans
64	domicile
57	évènement
15	évènement
56	fiançailles
65	séparation
63	lien conjugal LDS

4.2.2 Structure

INT	ID
INT	date de création (cf. 4.18)
INT	date de modification (cf. 4.18)
INT	ID de l'item associé (Individu ou Union)
CHAR	Type (cf. 4.2.1)
INT	ID du lieu
CHAR	?
CHAR	calendrier 1 ¹
CHAR	calendrier 2
CHAR	modificatif 1
CHAR	modificatif 2
CHAR	jour 1
CHAR	jour 2
CHAR	mois 1
CHAR	mois 2
CHAR	modificatif 3
CHAR	partie de l'année 1
CHAR	selecteur de l'année 1
CHAR	partie de l'année 2
CHAR	selecteur de l'année 2
CHAR	Heure
CHAR	Minute
16 CHARS	?
STRING	Note
STRING	Subdivision du lieu
STRING	Nom de l'évènement
SHORT	?
STRING	age sur l'acte
CHAR	Acte à rechercher
CHAR	?
CHAR	?

4.3 Individu

4.3.1 Constantes

signatures

- 0 Signe peut-être
- 1 Sait signer
- 2 Ne signe pas

Enfant

- 0 enfant légitime
- 1 enfant naturel
- 2 enfant reconnu
- 3 enfant légitimé
- 4 enfant trouvé
- 5 enfant adopté
- 6 enfant adultérin
- 7 enfant mort-né
- 8 filiation non connue

4.3.2 Structure

INT	ID
INT	date de création (cf. 4.18)
INT	date de modification (cf. 4.18)
INT	ID du père
INT	ID de la mère
INT	ID du nom de famille
INT	?
STRING	prénom
STRING	Activité
STRING	Genre (sexe)
STRING	Note
STRING	Numéro
STRING	Champ utilisateur 1
STRING	Champ utilisateur 2
STRING	Champ utilisateur 3
STRING	Champ utilisateur 4
STRING	Champ utilisateur 5
STRING	Champ utilisateur 6
STRING	Champ utilisateur 7
STRING	Champ utilisateur 8
STRING	Champ utilisateur 9
STRING	Champ utilisateur 10
SHORT	?
SHORT	Sans descendance
CHAR	Signature (cf. 4.3.1)
CHAR	Enfant (cf. 4.3.1)
CHAR	Marqué
SHORT	?
CHAR	confidentiel
STRING	Suffixe
STRING	Surnom
STRING	Titre
...	

4.4 Lieu

INT	ID
INT	date de création (cf. 4.18)
INT	date de modification (cf. 4.18)
INT	ID du lieu principal (ID du lieu auquel ce lieu est rattaché).
STRING	Ville
STRING	Code
STRING	Département
STRING	Région
STRING	Pays
...	

4.5 Lien

4.5.1 Constantes

1	Aïeul(e) -> Descendant(e)
2	Ami(e) -> Ami(e)
3	Beau-frère/Belle-soeur -> Beau-frère/Belle-soeur
4	Beau-père/Belle-mère -> Gendre/Bru
5	Lien consanguin
6	Cousin/Cousine -> Cousin/Cousine
7	Doublon ? -> Doublon ?
8	Frère/Soeur -> Frère/Soeur
9	Jumeau/Jumelle -> Jumeau/Jumelle
10	Testataire/Testatrice -> Héritier/Héritaire
11	Oncle/Tante -> Neveu/Nièce
12	Parent(e) -> Parent(e)
13	A reconnu -> Reconnu par
14	Tuteur/Tutrice -> Sous tutelle
15	Autre Lien
17	Déclarant
18	Officier d'état civil
19	Officiant Religieux
20	Parrain/Marraine
21	Présent(e)
22	Témoin
23	Autre Lien

4.5.2 Structure

INT	ID
INT	date de création (cf. 4.18)
INT	date de modification (cf. 4.18)
INT	ID de l'origine (individu ou événement ou union)
INT	ID de l'extrémité (individu)
STRING	Note
SHORT	?
CHAR	type
...	

4.6 Lien vers une source

INT	ID
INT	date de création (cf. 4.18)
INT	date de modification (cf. 4.18)
INT	ID de l'événement
INT	ID de la source
STRING	Note
...	

4.7 Lien vers un média

INT	ID
INT	date de création (cf. 4.18)
INT	date de modification (cf. 4.18)
INT	ID du propriétaire (Individu, Nom, Union, etc.)
INT	ID du média
INT(?)	principal (le média est-il le média principal de l'individu ?)
...	

4.8 Média

INT	ID
INT	date de création (cf. 4.18)
INT	date de modification (cf. 4.18)
STRING	répertoire où est situé le fichier
STRING	nom du fichier
CHAR	?
STRING	Commentaire
INT	?
CHAR	partie de l'année (cf. 4.17)
CHAR signé	partie de l'année
SHORT	?
SHORT	?
...	Miniature au format jpeg

4.9 Nom

INT	ID
INT	date de création (cf. 4.18)
INT	date de modification (cf. 4.18)
INT	ID du nom principal (auquel est rattaché le nom en question)
STRING	Nom
...	

4.10 Sosa

INT	ID de l'individu
INT (ou suite de CHAR) en big-endian	: Numéro Sosa

4.11 Source

4.11.1 Constantes

0	non trouvé
1	original
2	copie
3	photocopie
4	transcription
5	numérisation
6	microfilm
7	demandé
8	extrait
10	introuvable
11	cdrom
12	audio
16	carte
17	journal
13	livre
14	magazine
15	manuscrit
18	pièce tombale
19	video
20	souvenir
21	internet

4.11.2 Structure

INT	ID
INT	date de création (cf. 4.18)
INT	date de modification (cf. 4.18)
CHAR	?
STRING	Origine
STRING	Document
STRING	Cote
STRING	Archivage
SHORT	Nature (cf. 4.11.1)
SHORT	?
STRING	Note
STRING	Nom
...	

4.12 Union

INT ID
INT date de création (cf. 4.18)
INT date de modification (cf. 4.18)
INT ID du mari
INT ID de la femme
INT ?
STRING Note
...

4.13 IdxDoc

INT ID de l'événement
INT ID de la source
INT ID du lien source-événement

4.14 TBLink-IdxDE

INT ID de l'individu pointé par le lien
INT ID de l'individu ou de l'événement à l'origine du lien
INT ID du lien entre les deux

4.15 TBLink-IdxVERS

INT ID de l'individu ou de l'événement à l'origine du lien
INT ID de l'individu pointé par le lien
INT ID du lien entre les deux

4.16 TBEvent-Idx

INT ID de l'individu ou de l'union
INT ?
INT type de l'événement (cf. 4.2.1)
INT ID de l'événement

4.17 Date

Une date Heredis est toujours représentée par deux dates qui peuvent être égales. De plus deux indicateurs renseignent sur le type de calendrier ainsi que sur les interprétations des dates.

Ainsi, à partir de calendrier1, calendrier2, modificatif1, modificatif2, jour1, jour2, mois1, mois2, modificatif3, partie de l'année 1, selecteur de l'année 1, partie de l'année 2, selecteur de l'année 2 :

année1 = 256 * selecteur de l'année 1 + partie de l'année 1

année2 = 256 * selecteur de l'année 2 + partie de l'année 2

On forme alors les dates :

jour1, mois1, année1

jour2, mois2, année2

modificateur 1 et 2 :

1 après

2 avant

3 entre ... et ...

4 vers

5 calculée

6 estimée

calendrier 1 et 2 valent :

0x4a JULIEN

0x00 GREGORIAN

0x48 HEBRAIQUE

0x52 REPUBLICAN

4.18 Dates de création et de modification

Visiblement, Heredis enregistre pour tous les éléments principaux, sa date de création et sa date de modification sous la forme d'un entier depuis le 30 décembre 1899 !

Pourquoi le 30 décembre 1899 est une bonne question ! Peut-être que mon logiciel de calcul est faux mais ce qui est sur, c'est que le jour n°37829 correspond au 27 juillet 2003...