

THE GEDCOM STANDARD

Release 5.5

Prepared by the
Family History Department
The Church of Jesus Christ of Latter-day Saints

2 January 1996

Suggestions and Correspondence:

Email:
gedcom@gedcom.org

Telephone:
801-240-4534
801-240-5225

Mail:
Family History Department
GEDCOM Coordinator—3T
50 East North Temple Street
Salt Lake City, UT 84150
USA

TABLE OF CONTENTS

Introduction	3
Purpose and Content of <i>The GEDCOM Standard</i>	3
Purposes for Version 5.x	3
Modifications in Version 5.5 as a result of the 5.4 (draft) review	4
Changes Introduced or Modified in Draft Version 5.4	5
Changes Introduced in Draft Version 5.3	7
Chapter 1	
Data Representation Grammar	9
Concepts	9
Grammar	9
Description of Grammar Components	13
Chapter 2	
Lineage-Linked Grammar	19
Record Structures of the Lineage-Linked Form	23
Substructures of the Lineage-Linked Form	29
Primitive Elements of the Lineage-Linked Form	37
Compatibility with Other GEDCOM Versions	57
Packaging the GEDCOM Transmission File	59
Sample Lineage-Linked GEDCOM Transmission	60
Chapter 3	
Using Character Sets in GEDCOM	63
8-Bit ANSEL	63
ASCII (USA Version)	64
UNICODE (ISO 10646)	64
Chapter 4	
GEDCOM Product Registration	67
GEDCOM Product Registration	67
Appendix A	69
Lineage-Linked GEDCOM Tag Definition	69
Appendix B	79
Structure Cross-Reference	79
Primitive Cross-Reference	81
Appendix C	85
LDS Temple Codes	85
Appendix D	87
ANSEL Character Set	87
Non-spacing graphic characters	87
Spacing graphic characters	89
APPENDIX E	91
Encoding and Decoding Algorithms for Multimedia Objects	91
Introduction	91
Encoding	91
Decoding	92

Introduction

GEDCOM was developed by the Family History Department of The Church of Jesus Christ of Latter-day Saints (LDS Church) to provide a flexible, uniform format for exchanging computerized genealogical data. GEDCOM is an acronym for **GE**nealogical **D**ata **C**ommunication. Its purpose is to foster the sharing of genealogical information and the development of a wide range of inter-operable software products to assist genealogists, historians, and other researchers.

Purpose and Content of *The GEDCOM Standard*

The GEDCOM Standard is a technical document written for computer programmers, system developers, and technically sophisticated users. It covers the following topics:

- GEDCOM Data Representation Grammar (see Chapter 1)
- Lineage-Linked Grammar (see Chapter 2, beginning on page 19)
- Lineage-Linked GEDCOM Tags (see Chapter 2, beginning on page 19 and Appendix A, 69)
- Cross Reference of Structures and Primitives (see Appendix B beginning on page 79)
- The Church of Jesus Christ of Latter-day Saints' temple codes (see Appendix C, page 85)
- ANSEL Character Codes (see Chapter 3, beginning on page 63, and Appendix D beginning on page 87)

This document describes GEDCOM at two different levels. Chapter 1 describes the lower level, known as the ***GEDCOM data format***. This is a general-purpose data representation language for representing any kind of structured information in a sequential medium. It discusses the syntax and identification of structured information in general, but it does not deal with the semantic content of any particular kind of data. It is, therefore, also useful to people using GEDCOM for storing other types of data, not just genealogical data.

Chapter 2 of this document describes the higher level, known as a ***GEDCOM form***. Each type of data that uses the GEDCOM data format has a specific GEDCOM form. This document discusses only one GEDCOM form: the *Lineage-Linked GEDCOM Form*. This is the form commercial software developers use to create genealogical software systems that can exchange compiled information about individuals with accompanying family, source, submitter, and note records with the Family History Department's FamilySearch Systems and with each other if desired.

This document is available on the internet at:

<ftp://gedcom.org/pub/genealogy>

Purposes for Version 5.x

Earlier versions of *The GEDCOM Standard* were released in October 1987 (3.0) and August 1989 (4.0). Versions 1 and 2 were drafts for public discussion and were not established as a standard.

The 5.x series of drafts includes both the first standard definition of the Lineage-Linked GEDCOM Form and also the first major expansion of the Lineage-Linked Form since its initial use in GEDCOM 3.0. The GEDCOM-compatible products registered as 4.0 systems should still be able to exchange all of the data that was previously handled by their product with GEDCOM 5.x systems. See "Compatibility with Previous GEDCOM Releases," (starting on page 57) for compatibility specifics.

The following are the expanded purposes of Lineage-Linked GEDCOM :

- Simplify the description of the GEDCOM data representation grammar (rules) for ease of understanding. (See Chapter 1, starting on page 9.)
- Standardize the valid contexts in which tags, values, and pointers appear in the Lineage-Linked GEDCOM Form. (See Chapter 2, starting on page 19.) The Lineage-Linked GEDCOM Form should not be confused with other GEDCOM forms, which apply the basic GEDCOM data format but use different tag, value, and pointer combinations for other purposes.
- Define new data representations for supporting information such as sources, source citations, repositories, submitter records, submission records, and notes. (See Chapter 2, page 19, for GEDCOM representation of these support structures as used by the lineage-linked grammar.)
- Define a generic event structure.
- Define a way of associating individuals one to another. This is accomplished through a pointer which points from one individual record to another with a user-defined relationship description placed subordinate to this pointer. This feature is not a substitute for handling direct family relationships. Direct family relationships are represented by the FAMC and FAMS pointers.
- Add a product version number and a GEDCOM form and version number to the HEADer record structure.
- Define DATE modifiers (FROM, TO, ABT, BEF, AFT, BET) and more rigorously define the regular date format.
- Define an integration of multimedia within GEDCOM.

Modifications in Version 5.5 as a result of the 5.4 (draft) review

- Added tags for storing detailed address pieces under the address structure.
- Added nickname and surname prefix name pieces to the personal name structure.
- Added subordinate source citation to the note structure.
- Changed the encoding rules and the structure for including embedded multimedia objects.
- Added a RIN tag to the record structures. The RIN tag is a record identification assigned to the record by the source software. Its intended use is to allow for automated access to that record upon receipt of return transactions or other reconciliation processes.
- The meaning of a GEDCOM tag without a value on its line depends on its subordinate context for any assertions intended by the researcher. For example, In an event structure, a subordinate DATE and/or PLACe value imply that an event happened. However, a subordinate NOTE or SOURce context by themselves do not imply that the event took place. For a researcher to indicate that an event took place without knowing a date or a place requires that a Y(es) value be added to the event tag line. Using this convention protects GEDCOM processors which may remove (prune) lines that

have no value and also no subordinate lines. A N(o) value must not be used on an event tag line to assert that the event never happened. This requires the definition of a different tag.

- Returned the calendar escape sequence to support alternate calendars.
- The definition of the date value was refined to include many of the potential ways in which a person may define an imprecise date in a free form text field. Systems which guide users through a date statement should not result in such a precise way of stating an imprecise date. For example, if software was to estimate a marriage date based on an algorithm involving the birth date of the couple's first child, hardly needs to say "EST ABT 1881".
- The following tags were added:
ADR1, ADR2, CITY, NICK, POST, SPFX

Changes Introduced or Modified in Draft Version 5.4

Some changes introduced in GEDCOM draft version 5.4 are not compatible with earlier 5.x draft forms. Some concepts have been removed with the intent to address them in a future release of GEDCOM.

The following features are either new or different:

- The use of the SCHEMA has been eliminated. Although the schema concept is valid and essential to the growth of GEDCOM, it is too complex and premature to be implemented successfully into current products. Implementing it too early could cause developers to spend a great deal of resources programming something that would be outdated very quickly. Object definition languages are likely to contribute to meeting these needs.
- The EVENT_RECORD context has been deleted. This context was intended to support the evidence record concept in the Lineage-Linked GEDCOM Form, which ended up being more complicated than first supposed. Understanding the difference between the role of a source record and the role of a so-called evidence record requires further study.
- Non-standard tags (see <NEW_TAG>, page 50) can be used within a GEDCOM transmission, provided that the first character is an underscore (for example _NUTAG). Non-standard tags should be used only when structured information cannot be represented using existing context. Using a Note field is a more universal way of transmitting genealogical data that does not fit into the standard GEDCOM structure.
- The SOURCE_RECORD structure was simplified into five basic sections: data or classification, author, title, publication facts, and repository. The data or classification section contains facts about the data represented by this source and is used to analyze the collection of sources that the researcher used. The author, title, publication facts, and repository sections provide free-form text blocks that inform subsequent researchers how to access the source data that the original researcher used.
- The <<SOURCE_CITATION>> structure is placed subordinate to the fact being cited. It is generally best if the source citation contains only information specific to the fact being cited and then points to the more general description of the source, defined in a SOURCE_RECORD. This reduces redundancy, provides a way of controlling the GEDCOM record size, and more closely represents the normalized data model.

Systems that represent sources using the AUTHor, TITLE, PUBLication, and REPOsitory descriptions *can and should always pass this information in GEDCOM using the SOURCE record* pointed to by the <<SOURCE_CITATION>>. Systems that do not represent source information in these categories should provide the following information as unstructured text using the tags, TITL, AUTH, PUBL, and REPO, respectively, within the text:

- A descriptive title of the source
 - Who created the work
 - When and where was it created
 - Where can it be obtained or viewed
-
- Some attributes of individuals such as their EDUCation, OCCUpation, RESIdence, or nobility TITLE need to be described using a date and place. Therefore, the structure to describe the attributes was formatted to be the same as for describing events. That is, these attributes are further defined using a date, place, and other values used to describe events. (See <<INDIVIDUAL_EVENT_STRUCTURE>>, page 31.)
 - The LDS ordinance structure was extended to include the place of a living LDS ordinance. The TYPE tag line was changed to a STATus tag line. This allows statements such as *BIC*, *canceled*, *Infant*, and so forth to be removed from the date line and be added here under the STATus tag. (See <LDS_(*ordinance*)_DATE_STATUS>, page 45, 46) where (*ordinance*) represents any of the following: BAPTISM, ENDOWMENT, CHILD_SEALING, or SPOUSE_SEALING.
 - Previous GEDCOM 5.x versions overloaded the FAMC pointer structure with subordinate events which connected individual events and an associated family. An adoption event, for example, was shown subordinate to the FAMC pointer to indicate which was the adoptive family. The sealing of child to parent event (SLGC) was also shown in this manner. GEDCOM 5.4 recognizes that these are events and should be at the same level as the other individual events. To show the associated family, a subordinate FAMC pointer is placed subordinate to the appropriate event. (See <<INDIVIDUAL_EVENT_STRUCTURE>> page 31 and LDS_INDIVIDUAL_ORDINANCE at page 32.)
 - The date modifier (**int**) was added to the date format to indicate that the associated date phrase has been interpreted and the interpretation follows the **int** prefix in the date field. The date phrase is also included in the date value enclosed in parentheses. (See <DATE_APPROXIMATED>, page 40.)
 - The <AGE_AT_EVENT> primitive definition now includes the key words *STILLBORN*, *INFANT*, and *CHILD*. These words should be interpreted as being an approximate age at an event. (See <AGE_AT_EVENT>, page 37.)
 - The family event context in the FAMily record now allows the ages of both the husband and wife at the time of the event to be shown. (See FAM_RECORD page 24)
 - The <<PERSONAL_NAME_STRUCTURE>> structure now allows name pieces to be specifically identified as subordinate parts of the name line. Most products will not use subordinate name pieces. A nickname can now be included on the name line by enclosing it in double quotation marks. *Note:* Systems using the subordinate name parts must still provide the name structure formed in the same way specified for <NAME_PERSONAL> (see page 48.)

- A submission record was added to GEDCOM to enable the sending system to transmit information which will enable the receiving system to more appropriately process the GEDCOM data. The format currently designed for the submission record was created specifically for TempleReady™ system and for GEDCOM files being downloaded from Ancestral File™. (See SUBMISSION_RECORD, page 27.)
- A RESTRICTION (RESN) tag and a <RESTRICTION_NOTICE> primitive were added to the INDIVIDUAL_RECORD context. This allows some records in Ancestral File to be marked for privacy (indicating some personal information is not included) and some records to be marked as locked (indicating that Ancestral File will not make changes to the record without authorization from an assigned record steward).
- The following tags are no longer used in the Lineage-Linked Form:
ARVL, BROT, BUYR, CEME, CNTC, CPLR, DEFM, DPRT, EDTR, FIDE, FILM, GODP, HDOH, HEIR, HFAT, HMOT, INFT, INDX, INTV, ISA, ISSU, ITEM, LABL, LCCN, LGTE, MBR, NAMS, NAMR, OFFI, ORIG, OWNER, PERI, PORT, PWIF, PUBR, RECO, SELR, SEQU, SERS, SIBL, SIGN, SIST, SITE, TXPY, XLTR, WFAT, WITN, WMOT, AUDIO, IMAGE, PHOTO, SCHEMA, VIDEO
- The following tags were added:
BLOB, CTRY, CREM, FCOM, GIVN, NPFX, NSFX, OBJE, PEDI, RELA, RESI, RESN, SUBN, SURN, STAT

Changes Introduced in Draft Version 5.3

Version 5.3 introduced the following changes to the GEDCOM standard:

- An address structure was defined.
- A new tag for marital status (MSTA) at the time of an event was added to the event structure. (This was removed in version 5.4.)
- A mechanism for creating user-defined tags was added. These were defined in a SCHEMA definition in the header record of 5.3. (SCHEMA was removed in version 5.4.)
- The Unicode standard (ISO 10646) was introduced as an additional character set. (This was reduced to potential character set in version 5.4. See Chapter 3, page 63.)
- A <<MULTIMEDIA_LINK>> structure was introduced to provide linking and embedding of digitized photo, video, and sound files. (This was modified in version 5.4. See MULTIMEDIA_LINK page 33 and MULTIMEDIA_RECORD page 26)
- The source structure NAME tag, meaning the name of the source in the <<SOURCE_STRUCTURE>>, was changed back to the TITLE tag and is used to show the title of a book, article, or descriptive title of non-titled sources.
- The <<SOURCE_STRUCTURE>> was changed. Usage of CPLR, XLTR, and INFT tags in source substructures was discontinued.

- The FORM {FORMAT} tag was added subordinate to the PLACe and the GEDCom tags in the HEADER record and also subordinate to the PLACe tag in the <<PLACE_STRUCTURE>>. The PLAC.FORM line in the header record indicates that all of the locality names are specified in a consistent hierarchal sequence as specified by the value of the FORM. For example: 2 FORM City, County, State. GEDCOM 5.2 used the TYPE tag, subordinate to the PLAC tag instead of the FORM tag, for this purpose. This provision is for products which have overly structured the place value.

Chapter 1

Data Representation Grammar

Introduction

This chapter describes the core GEDCOM data representation language.

The generic data representation language defined in this chapter may be used to represent any form of structured information, not just genealogical data, using a sequential stream of characters.

Concepts

A GEDCOM transmission represents a database in the form of a sequential stream of related records. A record is represented as a sequence of tagged, variable-length lines, arranged in a hierarchy. A line always contains a hierarchical level number, a tag, and an optional value. A line may also contain a cross-reference identifier or a pointer. The GEDCOM line is terminated by a carriage return, a line feed character, or any combination of these.

The tag in the GEDCOM line, taken in its hierarchical context, identifies the information contained in the line, in the same sense that a field-name identifies a field in a database record. This means that the data is self-defining. Tags allow a field to occur any number of times within a record, including zero times. They also allow the use of different or new fields to be included in the GEDCOM data without introducing incompatibility, because the receiving system will ignore data which it does not understand and process only the data that it does understand.

The hierarchical relationships are indicated by a level number. Subordinate lines have a higher level number. The hierarchy allows a line to have sub-lines, which in turn may have their own sub-lines, and so forth. A line and its sub-lines constitute a context or enclosure, that is, a cluster of information pertaining directly to the same thing. This hierarchical arrangement corresponds with the natural hierarchy found in most structured information.

A series of one or more lines constitutes a record. The beginning of a new record is indicated by a *line whose level number is 0 (zero)*.

In addition to hierarchical relationships, GEDCOM defines the inter-record relationships that allow a record to be logically related to other records, without introducing redundancy. These relationships are represented by two additional, but optional, parts of a line: a cross-reference pointer and a cross-reference identifier. The cross-reference pointer "points at" a related record, which is identified by a required, matching unique cross-reference identifier. The cross-reference identifier is analogous to a primary key in relational database terminology.

Grammar

This chapter defines the grammar for the GEDCOM format. The grammar is a set of rules that specify the character sequences that are valid for creating the expression of the GEDCOM line. The character sequences are described in terms of various combinations of elements (variables and/or constants).

Elements may be described in terms of a set of other elements, some of which are selected from a set of alternative elements. Each element in the definition is separated by a plus sign (+) signifying that both elements are required. When there is a choice of different elements that can be used, the set of alternatives are listed between opening and closing square brackets ([]), with each choice separated by a vertical bar ([alternative_1 | alternative_2]). The user can read the grammar components of the selected element by substituting any sub-elements until all sub-elements have been resolved.

A GEDCOM transmission consists of a sequence of logical records, each of which consists of a sequence of **gedcom_lines**, all contained in a sequential file or stream of characters. The following rules pertain to the **gedcom_line**:

- Long values can be broken into shorter GEDCOM lines by using a subordinate CONC or CONT tag. The CONC tag assumes that the accompanying subordinate value is concatenated to the previous line value without saving the carriage return prior to the line terminator. The CONT assumes that the subordinate line value is concatenated to the previous line, saving the carriage return.
- The beginning of a new logical record is designated by a line whose **level** number is 0 (zero).
- Each new level number must be no higher than the previous line plus 1.
- Logical GEDCOM record sizes should be constrained so that they will fit in a memory buffer of less than 32K. GEDCOM files with records sizes greater than 32K run the risk of not being able to be loaded in some programs. Use of pointers to records, particularly NOTE records, should ensure that this limit will be sufficient. The size of embedded multimedia records can be controlled through chaining MULTIMEDIA_RECORDS (see multimedia record format on p. 26.)
- Any length constraints are given in characters, not bytes. When wide characters (characters wider than 8 bits) are used, byte buffer lengths should be adjusted accordingly.
- Level numbers must be between 1 to 99 and must not contain leading zeroes, for example, level one must be 1, not 01.
- The cross-reference ID has a maximum of 22 characters, including the enclosing at signs (@), and it must be unique within the GEDCOM transmission.
- Pointers to records imply that the record pointed to does actually exist within the transmission. Future pointer structures may allow pointing to records within a public accessible database as an alternative.
- The length of the GEDCOM TAG is a maximum of 31 characters, with the first 15 characters being unique.
- The total length of a GEDCOM line, including leading white space, level number, cross-reference number, tag, value, delimiters, and terminator, must not exceed 255 (wide) characters.
- Leading white space (tabs, spaces, and extra line terminators) preceding a GEDCOM line should be ignored by the reading system. Systems generating GEDCOM should not place any white space in front of the GEDCOM line.

Grammar Syntax

A **gedcom_line** has the following syntax:

gedcom_line:=

level + delim + [xref_id + delim +] tag + [delim + line_value +] terminator

level + delim + optional_xref_id + tag + delim + optional_line_value + terminator

for example:

1 NAME Will /Rogers/

The components used in the pattern above are defined below in alphabetical order. Some of the components are defined in terms of other primitive patterns. The spaces used in the patterns below are only to set them apart and are not a part of the resulting pattern. Character constants are specified in the hex form (0x20) which is the ASCII hex value of a space character. Character constants that are separated by a (-) dash represent any character with in that range from the first constant shown to and including the second constant shown.

alpha:=

[(0x41)-(0x5A) | (0x61)-(0x7A) | (0x5F)]

where:

(0x41)-(0x5A)=A to Z

(0x61)-(0x7A)=a to z

(0x5F)=(_) underscore

alphanum:=

[**alpha** | **digit**]

any_char:=

[**alpha** | **digit** | otherchar | (0x23) | (0x20) | (0x40)+(0x40)]

where:

(0x23)=#

(0x20)=space character

(0x40)+(0x40)=@@

delim:=

[(0x20)]

where:

(0x20)=space_character

digit:=

[(0x30)-(0x39)]

where:

(0x30)-(0x39) = One of the digits 0, 1,2,3,4,5,6,7,8,9

escape:=

[(0x40) + (0x23) + escape_text + (0x40) + non_at]

where:

(0x40)=@

(0x23)=#

escape_text:=

[any_char | escape_text + any_char]

The escape_text is coded to meet the rules of a particular GEDCOM form.

level:=

[digit | level + digit]

(Do not use non-significant leading zeroes such as 02.)

line_item:=

[any_char | escape | line_item + any_char | line_item + escape]

line_value:=

[pointer | line_item]

non_at:=

[alpha | digit | otherchar | (0x23) | (0x20)]

where:

(0x20)=space character

(0x23)=#

null:= nothing

optional_line_value:= line_value + delim

optional_xref_Id:= xref_Id + delim

otherchar:=

[(0x21)-(0x22) | (0x24)-(0x2F) | (0x3A)-(0x3F) | (0x5B)-(0x5E) | (0x60) | (0x7B)-(0x7E) | (0x80)-(0xFE)]

where, respectively:

(0x21)-(0x22)=! "

(0x24)-(0x2F)=\$ % & ' () * + , - . /

(0x3A)-(0x3F)=: ; < = > ?

(0x5B)-(0x5E)=[\] ^

(0x60)=`

(0x7B)-(0x7E)={ | } ~

(0x80)-(0xFE)=ANSEL characters above 127

Any 8-bit ASCII character except control characters (0x00–0x1F), alphanum, space (), number sign (#), at sign (@), _ underscore, and the DEL character (0x7F).

pointer:=

[(0x40) + alphanum + pointer_string + (0x40)]

where:

(0x40)=@

pointer_char:=

[non_at]

pointer_string:=

[null | pointer_char | pointer_string + pointer_char]

tag:=

[alphanum | tag + alphanum]

terminator:=

[carriage_return | line_feed | carriage_return + line_feed |
line_feed + carriage_return]

xref_id:=

[pointer]

Description of Grammar Components

alpha:=

The alpha characters include the underscore, which is used to link word pieces together in forming tag names or tag labels.

any_char:=

Any 8-bit ASCII character except the control characters found in the range of 0x00–0x1F and 0x7F. If an @ is desired as part of the line_value, it must be written in GEDCOM as a double @, i.e., "3 doz. @ \$20.00" must be stored as "3 doz. @@ \$20.00."

delim:=

The **delim** (delimiter), a single space character, terminates both the variable-length **level** number and the variable-length **tag**. Note that space characters may also be present in a **value**.

escape:=

The **escape** is a character sequence in the grammar used to specify special processing, such as for switching character sets or for indicating an inclusion of a non-GEDCOM data form into the GEDCOM structure. The form of the escape sequence is:

@+##**escape_text**+@+**non_at**.

Receiving systems should discard any space character which follows the escape sequences closing at-sign (@). If the character following the escape sequence's closing at-sign (@) is not a space character then it should be kept as a part of the text following the escape. Systems writing escape sequences should always output a space character following the escape sequence.

The specific format of the escape sequence is defined for the specific GEDCOM form being defined.

escape_text:=

The `escape_text` is defined to meet the requirements of a particular GEDCOM form.

level:=

The **level** number works the same way as the level of indentation in an indented outline, where indented lines provide detail about the item under which they are indented. A line at any level L is enclosed by and pertains directly to the *nearest preceding line* at level L-1. The Level L may increase by 1 at most. Level numbers must not contain leading zeroes, for example level one must be (1), not (01).

The enclosed subordinate lines at level L are said to be in the context of the enclosing superior line at level L-1. The interpretation of a **tag** must be in the context of the **tags** of the enclosing line(s) rather than just the tag by itself. Take the following record about an individual's birth and death dates, for example:

```
0 INDI
  1 BIRT
    2 DATE 12 MAY 1920
  1 DEAT
    2 DATE 1960
```

In this example, the expression DATE 12 MAY 1920 is interpreted within the INDI (individual) BIRT (birth) context, representing the individual's birth date. The second DATE is in the INDI.DEAT (individual's death) context. The complete meaning of DATE depends on the context.

Note: The above example is indented according to the level numbers to make the concept more obvious. *In the actual GEDCOM data, the level numbers are lined up vertically, meaning they are the first character(s) of the GEDCOM line.*

Some systems output indented GEDCOM data for better readability by putting space or tab characters between the terminator and the **level** number of the next line to visibly show the hierarchy. Also, some people have suggested allowing extra blank lines to visibly separate physical records. GEDCOM files produced with these features are not to be used when transmitting GEDCOM to other systems.

line_value:=

The **line_value** identifies an object within the domain of possible values allowed in the context of the **tag**. The combination of the **tag**, the **line_value**, and the hierarchical context of the supporting **gedcom_lines** provides the understanding of the enclosed **values**. This domain is defined by a specific grammar for representing a given GEDCOM form. (See Chapter 2, starting on page 19 for Lineage-Linked GEDCOM Form grammar.)

Values whose source information contains illegible parts of the value should be indicated by replacing the illegible part with an ellipsis (...).

Values are generally not encoded in binary or other abbreviation schemes for reducing space requirements, and they are generally constrained to be understandable by a typical user without decoding. This is intended to reduce the decoding burden on the receiving software. A GEDCOM-optimized data compression standard will be defined in the future to reduce space requirements. Meanwhile, users may agree to compress and decompress GEDCOM files using any compression system available to both sender and receiver.

The **line_value** within the context of a tag hierarchy of **gedcom_lines** represents one piece of information and corresponds to one field in traditional database or file terminology.

otherchar:=

Any 8-bit ASCII character except control characters (0x00–0x1F), **alphanum**, space (), number-sign (#), the at sign (@), and the DEL character (0x7F).

pointer:=

A **pointer** stands in the place of the context identified by the matching **xref_id**. Theoretically, a receiving system should be prepared to follow a **pointer** to find **any** needed **value** in a manner that is transparent to the logic of the subsystem that is looking for specific **tags**. This highly flexible facility will probably be used more in the future. For the time being, however, the use of **pointers** is explicitly defined within the GEDCOM form, such as the Lineage-Linked GEDCOM Form defined in Chapter 2 (see page 19).

The **pointer** represents the association between two objects that usually reside in different records. Objects within a logical record can be associated. If this need exists, the pointer record composition contains an exclamation point (!) that separates the parent record's cross-reference ID from the specific substructure's cross-reference ID, which is at some subordinate level to the logical record at level zero. The cross-reference ID of the substructure subordinate to a zero level record, for inter-record associations is always composed of the Record ID number and the Substructure ID number, such as @I132!1@. Including the Record ID number in the pointer that associates objects within a record will allow the GEDCOM processors to build the index only at the record level and then search sequentially for the appropriate substructure cross-reference ID. The parent record ID is assumed when the cross-reference ID begins with a exclamation point (!) signifying an intra-record association.

Complex logical record structures are divided into small physical records to accommodate memory constraints, many-to-many relationships, and independent record creation and deletion.

The **pointer** must match a corresponding unique **xref_id** within the transmission, unless the colon (:) character is present (which will be used in the future as a network reference to a permanent file record). A **pointer** is given instead of duplicating an object, though the logical result is equivalent. An expanded traversal of a record tree includes following the **pointer** to related records to some depth, and splicing those records (logically) into the resultant expanded tree. *Pointers may refer to either records which have not yet appeared in the transmission (forward reference) or to records that have already appeared earlier in the transmission (backward reference).* This arrangement usually requires a preliminary pass to construct a look up table to support random access by **xref_id** during subsequent passes.

tag:=

A **tag** consists of a variable length sequence of **alphanum** characters. All user-defined tags, tags used that have not been defined in the GEDCOM standard, must begin with an underscore character (0x95).

The **tag** represents the meaning of the **line_value** within the context of the enclosing lines, and contributes to the meaning of enclosed subordinate lines. Specific **tags** are defined in Appendix A (starting on page 69). The presence of a tag together with a value represents an assertion which the submitter wishes to communicate to a receiver. A tag with no value does not represent an assertion. If a tag is absent, no assertion is made, for example, no information is submitted. Information of a negative nature (such as knowing positively an event did not occur) is handled through the semantic definition of a tag and accompanying values that assert the information explicitly. It is not represented by absence of a tag.

Although formally defined **tags** are only three or four characters long, systems should prepare to handle user tags of greater length. Tags will be unique within the first 15 characters.

Valid combinations of specific **tags**, **line_values**, **xref_ids**, and **pointers** are constrained by the GEDCOM form defined for representing a given kind of information. (See Chapter 2, starting on page 19, for the Lineage-Linked GEDCOM Form grammar.)

terminator:=

The **terminator** delimits the variable-length **line_value** and signals the end of the **gedcom_line**. The valid terminator characters are:

```
[carriage_return |  
line_feed |  
carriage_return line_feed |  
line_feed carriage_return ]
```

xref_id:=

(See **pointer**, page 15)

The **xref_id** is formed by any arbitrary combination of characters from the **pointer_char** set. The first character must be an alpha or a digit. The **xref_id** is not retained in the receiving system, and it may therefore be formed from any convenient combination of identifiers from the sending system. No meaning is attributed by the receiver to any part of the **xref_id**, other than its **unique** association with the associated record. The use of the colon (:) character is also reserved.

Examples:

The following are examples of valid but unrelated GEDCOM lines:

```
0 @1234@ INDI  
...  
1 AGE 13y  
...  
1 CHIL @1234@  
...  
1 NOTE This is a note field that is  
2 CONT continued on the next line.
```


The first line has a **level** number 0, a **xref_id** of @1234@, an **INDI tag**, and no **value**.

The second line has a **level** number 1, no **xref_id**, an **AGE tag**, and a **value** of 13.

The third line has a **level** number 1, no **xref_id**, a **CHIL tag**, and a **value** of a **pointer** to a **xref_id** named @1234@.

Chapter 2

Lineage-Linked Grammar

Introduction

This chapter describes the specific tag, value, and pointer combinations used for exchanging family-based lineage-linked genealogical information in the GEDCOM format. Lineage-linked data pertains to individuals linked in family relationships across multiple generations. The chapter also addresses specific compatibility issues pertaining to previous Lineage-Linked GEDCOM Form releases and contains a sample lineage-linked GEDCOM transmission.

The Lineage-Linked GEDCOM Form defined in this chapter is based on the general framework of the GEDCOM data representation grammar defined in Chapter 1. Commercial genealogical software systems are invited to use the Lineage-Linked GEDCOM Form to exchange data. It is the only form approved for exchanging data with Ancestral File, TempleReady and other Family History resource files maintained for this purpose.

Organization

The basic description of the **Lineage-Linked GEDCOM Form's grammar** is presented in the following three major sections:

- "Record Structures of the Lineage-Linked Form" (beginning on page 23)
- "Substructures of the Lineage-Linked Form" (beginning on page 29)
- "Primitive elements of the Lineage-Linked Form" (beginning on page 37)

The definition of the tags used in defining the lineage-linked structures are contained in Appendix A.

Symbols Used in Chapter 2

The following symbols are used in Chapter 2:

<<double_angle bracket>>

Indicates a subordinate GEDCOM structure pattern of a record, structure, or substructure is to be substituted in place of the enclosing double angle brackets. The substitute structure pattern is found subordinate to the LINEAGE_LINKED_GEDCOM beginning on page 24 for record pattern definition or in alphabetical order under the "Substructures of the Lineage-Linked Form" section, beginning on page 29.

<Single_angle bracket>

Indicates the name of the appropriate value for this GEDCOM line—<Primitive>. The specific definition of this value is found in alphabetical order in "Primitive Elements of the Lineage-Linked Form," beginning on page 37.

{braces}

Indicates the minimum to maximum occurrences allowed for this structure or line—{Minimum:Maximum}. Note that minimum and maximum occurrence limits are defined

relative to the enclosing superior line. This means that a required line (minimum = 1) *is not required if the optional enclosing superior line is not present*. Similarly, a line occurring only once (maximum = 1) may occur multiple times as long as each occurs only once under its own multiple-occurring superior line.

[Square brackets]

Indicates a choice of one or more options—[Choice of].

| vertical bar |

Separates the multiple choices, for example [Choice 1 | Choice 2].

n level number

A level number which assumes the level number of the line which referenced the substructure name.

+1, +2 ...

A **+1** level number is 1 greater than the level number assumed by the superior **n** level. A **+2** level number is 2 greater, and so forth.

0xHH

Indicates an allowable hexadecimal character value where HH is that value, for example, 0x20 (decimal 32) indicates the space character.

Lineage-Linked Form Usage Conventions

- The order in which GEDCOM lines are written to a GEDCOM file is controlled by the context and level number. When the lines are of equal level number but have a different tag name then the order is not significant. The occurrence of equal level numbers and equal tags within the same context imply that multiple opinions or multiple values of the data exist. The significance of the order in these cases is interpreted as the submitter's preference. *The most preferred value being the first with the least preferred data listed in subsequent lines by order decreasing preference.* For example, a researcher who discovers conflicting evidence about a person's birth event would list the most credible information first and the least credible or preferred items last.

Systems that support multiple fields or structures should allow their users to indicate their preference opinion. Systems that only store single value structures should use the preferred information (the first occurrence listed) and store the remaining information as an exception, preferably *within an appropriate NOTE field* or in some way that the patron has ready access to the less-preferred data when viewing the record.

- Conflicting event dates and places should be represented by placing them in separate event structures with appropriate source citations rather than by placing them under the same enclosing event.
- The Lineage-Linked GEDCOM Form uses the TYPE tag to further classify its superior tag for the viewer. The value portion given by the TYPE tag is not intended to inform a computer program how to process the data. The difference between this value and a note value is that displaying systems should always display the *type value* when they display the data from the associated context. This gives the user some flexibility in further describing the information provided but does not require

the software to recognize and respond to the large variety qualifiers that might be used. For example:

```
1 EVEN
  2 TYPE Awarded BSA Eagle Rank
  2 DATE 1980
```

- For compatibility, support of multiple calendars using the (#D) escape sequence is still part of the GEDCOM 5.5 Lineage-Linked Form. Multiple calendar support is not satisfactory and should be addressed in future GEDCOM releases. As a possible future direction, simplification could be obtained by converting from a day in a given source calendar to a specific sequential day number based on some beginning point in time. The receiving system would then convert from that day number to the calendar day of their choice. The Julian day system is based on a sequential day beginning from the first day of 4713 years B.C. The Julian day system with an extension that allows imprecise dates. Later versions of GEDCOM may consider this approach. This approach would also introduce the idea of the evidence date being shared in the DATE tag and the universal day number would be passed under a separate tag.
- All controlled line_value choices should be considered as case insensitive. This means that the values should be converted to all uppercase or all lowercase prior to comparing. The terms UPPERCASE and UpperCase are considered equal. TAGS are always UPPERCASE.

Record Structures of the Lineage-Linked Form

LINEAGE_LINKED_GEDCOM:=

This is a model of the lineage-linked GEDCOM structure for submitting data to other lineage-linked GEDCOM processing systems. A header and a trailer record are required, and they can enclose any number of data records. Tags from Appendix A (see page 69) must be used in the same context as shown in the following form. User defined tags (see <NEW_TAG> on page 50) are discouraged but when used must begin with an under-score.

0	<<HEADER>>	{1:1}	p.23
0	<<SUBMISSION_RECORD>>	{0:1}	p.27
0	<<RECORD>>	{1:M}	p.24
0	TRLR	{1:1}	

HEADER:=

n	HEAD	{1:1}	
+1	SOUR <APPROVED_SYSTEM_ID>	{1:1}	p.38
+2	VERS <VERSION_NUMBER>	{0:1}	p.55
+2	NAME <NAME_OF_PRODUCT>	{0:1}	p.48
+2	CORP <NAME_OF_BUSINESS>	{0:1}	p.48
+3	<<ADDRESS_STRUCTURE>>	{0:1}	p.29
+2	DATA <NAME_OF_SOURCE_DATA>	{0:1}	p.48
+3	DATE <PUBLICATION_DATE>	{0:1}	p.51
+3	COPR <COPYRIGHT_SOURCE_DATA>	{0:1}	p.39
+1	DEST <RECEIVING_SYSTEM_NAME>	{0:1*}	p.51
+1	DATE <TRANSMISSION_DATE>	{0:1}	p.55
+2	TIME <TIME_VALUE>	{0:1}	p.55
+1	SUBM @XREF:SUBM@	{1:1}	p.55
+1	SUBN @XREF:SUBN@	{0:1}	p.55
+1	FILE <FILE_NAME>	{0:1}	p.44
+1	COPR <COPYRIGHT_GEDCOM_FILE>	{0:1}	p.39
+1	GEDC	{1:1}	
+2	VERS <VERSION_NUMBER>	{1:1}	p.55
+2	FORM <GEDCOM_FORM>	{1:1}	p.44
+1	CHAR <CHARACTER_SET>	{1:1}	p.39
+2	VERS <VERSION_NUMBER>	{0:1}	p.55
+1	LANG <LANGUAGE_OF_TEXT>	{0:1}	p.45
+1	PLAC	{0:1}	
+2	FORM <PLACE_HIERARCHY>	{1:1}	p.51
+1	NOTE <GEDCOM_CONTENT_DESCRIPTION>	{0:1}	p.44
+2	[CONT CONC] <GEDCOM_CONTENT_DESCRIPTION>	{0:M}	

* NOTE:

Submissions to the Family History Department for Ancestral File submission or for clearing temple ordinances **must** use a DESTination of **ANSTFILE** or **TempleReady**.

The header structure provides information about the entire transmission. The SOURCE system name identifies which system sent the data. The DESTINATION system name identifies the intended receiving system.

Additional GEDCOM standards will be produced in the future to reflect GEDCOM expansion and maturity. This requires the reading program to make sure it can read the **GEDC.VERS** and the **GEDC.FORM** values to insure proper readability. The **CHAR** tag is required. All character codes greater than 0x7F **must be converted to ANSEL**. (See Chapter 3, starting on page 63.)

RECORD:=

[
n	<<FAM_RECORD>>	{1:1}	p.24
n	<<INDIVIDUAL_RECORD>>	{1:1}	p.25
n	<<MULTIMEDIA_RECORD>>	{1:M}	p.26
n	<<NOTE_RECORD>>	{1:1}	p.26
n	<<REPOSITORY_RECORD>>	{1:1}	p.26
n	<<SOURCE_RECORD>>	{1:1}	p.27
n	<<SUBMITTER_RECORD>>	{1:1}	p.28
]			

FAM_RECORD:=

n	@<XREF:FAM>@ FAM	{1:1}	
	+1 <<FAMILY_EVENT_STRUCTURE>>	{0:M}	p.30
	+2 HUSB	{0:1}	
	+3 AGE <AGE_AT_EVENT>	{1:1}	p.37
	+2 WIFE	{0:1}	
	+3 AGE <AGE_AT_EVENT>	{1:1}	
	+1 HUSB @<XREF:INDI>@	{0:1}	p.55
	+1 WIFE @<XREF:INDI>@	{0:1}	p.55
	+1 CHIL @<XREF:INDI>@	{0:M}	p.55
	+1 NCHI <COUNT_OF_CHILDREN>	{0:1}	p.39
	+1 SUBM @<XREF:SUBM>@	{0:M}	p.55
	+1 <<LDS_SPOUSE_SEALING>>	{0:M}	p.33
	+1 <<SOURCE_CITATION>>	{0:M}	p.34
	+1 <<MULTIMEDIA_LINK>>	{0:M}	p.33,26
	+1 <<NOTE_STRUCTURE>>	{0:M}	p.33
	+1 REFN <USER_REFERENCE_NUMBER>	{0:M}	p.55
	+2 TYPE <USER_REFERENCE_TYPE>	{0:1}	p.55
	+1 RIN <AUTOMATED_RECORD_ID>	{0:1}	p.38
	+1 <<CHANGE_DATE>>	{0:1}	p.29

The FAMily record is used to record marriages, common law marriages, and family unions caused by two people becoming the parents of a child. There can be no more than one HUSB/father and one WIFE/mother listed in each FAM_RECORD. If, for example, a man participated in more than one family union, then he would appear in more than one FAM_RECORD. The family record structure assumes that the HUSB/father is male and WIFE/mother is female.

The preferred order of the CHILdren pointers within a FAMily structure is chronological by birth.

INDIVIDUAL_RECORD:=

n @XREF:INDI@ INDI	{1:1}	
+1 RESN <RESTRICTION_NOTICE>	{0:1}	p.52
+1 <<PERSONAL_NAME_STRUCTURE>>	{0:M}	p.34
+1 SEX <SEX_VALUE>	{0:1}	p.53
+1 <<INDIVIDUAL_EVENT_STRUCTURE>>	{0:M}	p.31
+1 <<INDIVIDUAL_ATTRIBUTE_STRUCTURE>>	{0:M}	p.30
+1 <<LDS_INDIVIDUAL_ORDINANCE>>	{0:M}	p.32
+1 <<CHILD_TO_FAMILY_LINK>>	{0:M}	p.29
+1 <<SPOUSE_TO_FAMILY_LINK>>	{0:M}	p.35
+1 SUBM @<XREF:SUBM>@	{0:M}	p.55
+1 <<ASSOCIATION_STRUCTURE>>	{0:M}	p.29
+1 ALIA @<XREF:INDI>@	{0:M}	p.55
+1 ANCI @<XREF:SUBM>@	{0:M}	p.55
+1 DESI @<XREF:SUBM>@	{0:M}	p.55
+1 <<SOURCE_CITATION>>	{0:M}	p.34
+1 <<MULTIMEDIA_LINK>>	{0:M}	p.33,26
+1 <<NOTE_STRUCTURE>>	{0:M}	p.33
+1 RFN <PERMANENT_RECORD_FILE_NUMBER>	{0:1}	p.50
+1 AFN <ANCESTRAL_FILE_NUMBER>	{0:1}	p.38
+1 REFN <USER_REFERENCE_NUMBER>	{0:M}	p.55
+2 TYPE <USER_REFERENCE_TYPE>	{0:1}	p.55
+1 RIN <AUTOMATED_RECORD_ID>	{0:1}	p.38
+1 <<CHANGE_DATE>>	{0:1}	p.29

The individual record is a compilation of facts, known or discovered, about an individual. Sometimes these facts are from different sources. This form allows documentation of the source where each of the facts were discovered.

The normal lineage links are shown through the use of pointers from the individual to a family through either the FAMC tag or the FAMS tag. The FAMC tag provides a pointer to a family where this person is a child. The FAMS tag provides a pointer to a family where this person is a spouse or parent. The <<CHILD_TO_FAMILY_LINK>> (see page 29) structure contains a FAMC pointer which is **required to show any child to parent linkage** for pedigree navigation. The <<CHILD_TO_FAMILY_LINK>> structure also indicates whether the pedigree link represents a birth lineage, an adoption lineage, or a sealing lineage.

Linkage between a child and the family they belonged to at the time of an event can also optionally be shown by a FAMC pointer subordinate to the appropriate event. For example, a FAMC pointer subordinate to an adoption event would show which family adopted this individual. Biological parent or parents can be indicated by a FAMC pointer subordinate to the birth event. The FAMC tag can also optionally be used subordinate to an ADOPTION, or BIRTH event to differentiate which set of parents were related by adoption, sealing, or birth.

Other associations or relationships are represented by the ASSOCIATION tag. **The person's relation or association is the person being pointed to.** The association or relationship is stated by the value on the subordinate RELA line. For example:

```
0 @I1@ INDI
  1 NAME Fred/Jones/
  1 ASSO @I2@
    2 RELA Godfather
```

MULTIMEDIA_RECORD:=

```
n @XREF:OBJE@ OBJE {1:1}
  +1 FORM <MULTIMEDIA_FORMAT> {1:1} p.48
  +1 TITL <DESCRIPTIVE_TITLE> {0:1} p.43
  +1 <<NOTE_STRUCTURE>> {0:M} p.33
  +1 BLOB {1:1}
    +2 CONT <ENCODED_MULTIMEDIA_LINE> {1:M} p.43
  +1 OBJE @<XREF:OBJE>@ /* chain to continued object */ {0:1} p.26
  +1 REFN <USER_REFERENCE_NUMBER> {0:M} p.55
    +2 TYPE <USER_REFERENCE_TYPE> {0:1} p.55
  +1 RIN <AUTOMATED_RECORD_ID> {0:1} p.38
  +1 <<CHANGE_DATE>> {0:1} p.29
```

Large whole multimedia objects embedded in a GEDCOM file would break some systems. For this purpose, large multimedia files should be divided into smaller multimedia records by using the subordinate OBJE tag to chain to the next <MULTIMEDIA_RECORD> fragment. This will allow GEDCOM records to be maintained below the 32K limit for use in systems with limited resources.

NOTE_RECORD:=

```
n @<XREF:NOTE>@ NOTE <SUBMITTER_TEXT> {1:1} p.54
  +1 [ CONC | CONT ] <SUBMITTER_TEXT> {0:M}
  +1 <<SOURCE_CITATION>> {0:M} p.34
  +1 REFN <USER_REFERENCE_NUMBER> {0:M} p.55
    +2 TYPE <USER_REFERENCE_TYPE> {0:1} p.55
  +1 RIN <AUTOMATED_RECORD_ID> {0:1} p.38
  +1 <<CHANGE_DATE>> {0:1} p.29
```

REPOSITORY_RECORD:=

```
n @<XREF:REPO>@ REPO {1:1}
  +1 NAME <NAME_OF_REPOSITORY> {0:1} p.48
  +1 <<ADDRESS_STRUCTURE>> {0:1} p.29
  +1 <<NOTE_STRUCTURE>> {0:M} p.33
  +1 REFN <USER_REFERENCE_NUMBER> {0:M} p.55
    +2 TYPE <USER_REFERENCE_TYPE> {0:1} p.55
  +1 RIN <AUTOMATED_RECORD_ID> {0:1} p.38
  +1 <<CHANGE_DATE>> {0:1} p.29
```

SOURCE_RECORD:=

n	@<XREF:SOUR>@ SOUR	{1:1}	
+1	DATA	{0:1}	
+2	EVEN <EVENTS_RECORDED>	{0:M}	p.44
+3	DATE <DATE_PERIOD>	{0:1}	p.41
+3	PLAC <SOURCE_JURISDICTION_PLACE>	{0:1}	p.54
+2	AGNC <RESPONSIBLE_AGENCY>	{0:1}	p.52
+2	<<NOTE_STRUCTURE>>	{0:M}	p.33
+1	AUTH <SOURCE_ORIGINATOR>	{0:1}	p.54
+2	[CONT CONC] <SOURCE_ORIGINATOR>	{0:M}	p.54
+1	TITL <SOURCE_DESCRIPTIVE_TITLE>	{0:1}	p.53
+2	[CONT CONC] <SOURCE_DESCRIPTIVE_TITLE>	{0:M}	p.53
+1	ABBR <SOURCE_FILED_BY_ENTRY>	{0:1}	p.53
+1	PUBL <SOURCE_PUBLICATION_FACTS>	{0:1}	p.54
+2	[CONT CONC] <SOURCE_PUBLICATION_FACTS>	{0:M}	p.54
+1	TEXT <TEXT_FROM_SOURCE>	{0:1}	p.54
+2	[CONT CONC] <TEXT_FROM_SOURCE>	{0:M}	p.54
+1	<<SOURCE_REPOSITORY_CITATION>>	{0:1}	p.35
+1	<<MULTIMEDIA_LINK>>	{0:M}	p.33,26
+1	<<NOTE_STRUCTURE>>	{0:M}	p.33
+1	REFN <USER_REFERENCE_NUMBER>	{0:M}	p.55
+2	TYPE <USER_REFERENCE_TYPE>	{0:1}	p.55
+1	RIN <AUTOMATED_RECORD_ID>	{0:1}	p.38
+1	<<CHANGE_DATE>>	{0:1}	p.29

Source records are used to provide a bibliographic description of the source cited. (See the <<SOURCE_CITATION>> structure, page 34, which contains the pointer to this source record.)

SUBMISSION_RECORD:=

n	@XREF:SUBN@ SUBN	{1:1}	
+1	SUBM @XREF:SUBM@	{0:1}	p.55
+1	FAMF <NAME_OF_FAMILY_FILE>	{0:1}	p.48
+1	TEMP <TEMPLE_CODE>	{0:1}	p.54
+1	ANCE <GENERATIONS_OF_ANCESTORS>	{0:1}	p.44
+1	DESC <GENERATIONS_OF_DESCENDANTS>	{0:1}	p.44
+1	ORDI <ORDINANCE_PROCESS_FLAG>	{0:1}	p.50
+1	RIN <AUTOMATED_RECORD_ID>	{0:1}	p.38

The sending system uses a submission record to send instructions and information to the receiving system. TempleReady processes submission records to determine which temple the cleared records should be directed to. The submission record is also used for communication between Ancestral File download requests and TempleReady. **Each GEDCOM transmission file should have only one submission record.** Multiple submissions are handled by creating separate GEDCOM transmission files.

SUBMITTER_RECORD:=

n	@<XREF:SUBM>@ SUBM	{1:1}	
+1	NAME <SUBMITTER_NAME>	{1:1}	p.54

+1 <<ADDRESS_STRUCTURE>>	{0:1}	p.29
+1 <<MULTIMEDIA_LINK>>	{0:M}	p.33,26
+1 LANG <LANGUAGE_PREFERENCE>	{0:3}	p.45
+1 RFN <SUBMITTER_REGISTERED_RFN>	{0:1}	p.54
+1 RIN <AUTOMATED_RECORD_ID>	{0:1}	p.38
+1 <<CHANGE_DATE>>	{0:1}	p.29

The submitter record identifies an individual or organization that contributed information contained in the GEDCOM transmission. All records in the transmission are assumed to be submitted by the SUBMITTER referenced in the HEADER, unless a SUBMITTER reference inside a specific record points at a different SUBMITTER record.

Substructures of the Lineage-Linked Form

ADDRESS_STRUCTURE:=

n	ADDR <ADDRESS_LINE>	{0:1}	p.37
	+1 CONT <ADDRESS_LINE>	{0:M}	p.37
	+1 ADR1 <ADDRESS_LINE1>	{0:1}	p.37
	+1 ADR2 <ADDRESS_LINE2>	{0:1}	p.37
	+1 CITY <ADDRESS_CITY>	{0:1}	p.37
	+1 STAE <ADDRESS_STATE>	{0:1}	p.37
	+1 POST <ADDRESS_POSTAL_CODE>	{0:1}	p.37
	+1 CTRY <ADDRESS_COUNTRY>	{0:1}	p.37
n	PHON <PHONE_NUMBER>	{0:3}	p.50

The address structure should be formed as it would appear on a mailing label using the ADDR and ADDR.CONT lines. These lines are required if an ADDRESS is present. Optionally, additional structure is provided for systems that have structured their addresses for indexing and sorting.

ASSOCIATION_STRUCTURE:=

n	ASSO @<XREF:INDI>@	{0:M}	p.55
	+1 TYPE <RECORD_TYPE>	{1:1}	p.51
	+1 RELA <RELATION_IS_DESCRIPTOR>	{1:1}	p.52
	+1 <<NOTE_STRUCTURE>>	{0:M}	p.33
	+1 <<SOURCE_CITATION>>	{0:M}	p.34

CHANGE_DATE:=

n	CHAN	{1:1}	
	+1 DATE <CHANGE_DATE>	{1:1}	p.39
	+2 TIME <TIME_VALUE>	{0:1}	p.55
	+1 <<NOTE_STRUCTURE>>	{0:M}	p.33

The change date is intended to only record the last change to a record. Some systems may want to manage the change process with more detail, but it is sufficient for GEDCOM purposes to indicate the last time that a record was modified.

CHILD_TO_FAMILY_LINK:=

n	FAMC @<XREF:FAM>@	{1:1}	
	+1 PEDI <PEDIGREE_LINKAGE_TYPE>	{0:M}	p.50
	+1 <<NOTE_STRUCTURE>>	{0:M}	p.33

EVENT_DETAIL:=

n	TYPE <EVENT_DESCRIPTOR>	{0:1}	p.43
n	DATE <DATE_VALUE>	{0:1}	p.42/41
n	<<PLACE_STRUCTURE>>	{0:1}	p.34
n	<<ADDRESS_STRUCTURE>>	{0:1}	p.29
n	AGE <AGE_AT_EVENT>	{0:1}	p.37
n	AGNC <RESPONSIBLE_AGENCY>	{0:1}	p.52
n	CAUS <CAUSE_OF_EVENT>	{0:1}	p.38
n	<<SOURCE_CITATION>>	{0:M}	p.34
n	<<MULTIMEDIA_LINK>>	{0:M}	p.33,26

n <<NOTE_STRUCTURE>> {0:M} p.33

FAMILY_EVENT_STRUCTURE:=

[
n [ANUL | CENS | DIV | DIVF] [Y|<NULL>] {1:1}
+1 <<EVENT_DETAIL>> {0:1} p.29
|
n [ENGA | MARR | MARB | MARC] [Y|<NULL>] {1:1}
+1 <<EVENT_DETAIL>> {0:1} p.29
|
n [MARL | MARS] [Y|<NULL>] {1:1}
+1 <<EVENT_DETAIL>> {0:1} p.29
|
n EVEN {1:1}
+1 <<EVENT_DETAIL>> {0:1} p.29
]

INDIVIDUAL_ATTRIBUTE_STRUCTURE:=

[
n CAST <CASTE_NAME> {1:1} p.38
+1 <<EVENT_DETAIL>> {0:1} p.29
|
n DSCR <PHYSICAL_DESCRIPTION> {1:1} p.50
+1 <<EVENT_DETAIL>> {0:1} p.29
|
n EDUC <SCHOLASTIC_ACHIEVEMENT> {1:1} p.53
+1 <<EVENT_DETAIL>> {0:1} p.29
|
n IDNO <NATIONAL_ID_NUMBER> {1:1}* p.49
+1 <<EVENT_DETAIL>> {0:1} p.29
|
n NATI <NATIONAL_OR_TRIBAL_ORIGIN> {1:1} p.49
+1 <<EVENT_DETAIL>> {0:1} p.29
|
n NCHI <COUNT_OF_CHILDREN> {1:1} p.39
+1 <<EVENT_DETAIL>> {0:1} p.29
|
n NMR <COUNT_OF_MARRIAGES> {1:1} p.39
+1 <<EVENT_DETAIL>> {0:1} p.29
|
n OCCU <OCCUPATION> {1:1} p.50
+1 <<EVENT_DETAIL>> {0:1} p.29
|
n PROP <POSSESSIONS> {1:1} p.51
+1 <<EVENT_DETAIL>> {0:1} p.29
|
n RELI <RELIGIOUS_AFFILIATION> {1:1} p.52

+1 <<EVENT_DETAIL>>	{0:1}	p.29
n RESI	{1:1}	
+1 <<EVENT_DETAIL>>	{0:1}	p.29
n SSN <SOCIAL_SECURITY_NUMBER>	{0:1}	p.53
+1 <<EVENT_DETAIL>>	{0:1}	p.29
n TITL <NOBILITY_TYPE_TITLE>	{1:1}	p.50
+1 <<EVENT_DETAIL>>	{0:1}	p.29

]

* *Note:* The usage of IDNO requires that the subordinate TYPE tag be used to define what kind of number is assigned to IDNO.

INDIVIDUAL_EVENT_STRUCTURE:=

[
n [BIRT CHR] [Y <NULL>]	{1:1}	
+1 <<EVENT_DETAIL>>	{0:1}	p.29
+1 FAMC @<XREF:FAM>@	{0:1}	p.55
n [DEAT BURI CREM] [Y <NULL>]	{1:1}	
+1 <<EVENT_DETAIL>>	{0:1}	p.29
n ADOP [Y <NULL>]	{1:1}	
+1 <<EVENT_DETAIL>>	{0:1}	p.29
+1 FAMC @<XREF:FAM>@	{0:1}	p.55
+2 ADOP <ADOPTED_BY_WHICH_PARENT>	{0:1}	p.37
n [BAPM BARM BASM BLES] [Y <NULL>]	{1:1}	
+1 <<EVENT_DETAIL>>	{0:1}	p.29
n [CHRA CONF FCOM ORDN] [Y <NULL>]	{1:1}	
+1 <<EVENT_DETAIL>>	{0:1}	p.29
n [NATU EMIG IMMI] [Y <NULL>]	{1:1}	
+1 <<EVENT_DETAIL>>	{0:1}	p.29
n [CENS PROB WILL] [Y <NULL>]	{1:1}	
+1 <<EVENT_DETAIL>>	{0:1}	p.29
n [GRAD RETI] [Y <NULL>]	{1:1}	
+1 <<EVENT_DETAIL>>	{0:1}	p.29
n EVEN	{1:1}	
+1 <<EVENT_DETAIL>>	{0:1}	p.29
]		

The EVEN tag in this structure is for recording general events or attributes that are not shown in the above <<INDIVIDUAL_EVENT_STRUCTURE>>. The general event or attribute type is declared

by using a subordinate TYPE tag to show what event or attribute is recorded. For example, a candidate for state senate in the 1952 election could be recorded:

```
1 EVEN
  2 TYPE Election
  2 DATE 07 NOV 1952
  2 NOTE Candidate for State Senate.
```

The TYPE tag is also optionally used to modify the basic understanding of its superior event and is usually provided by the user. For example:

```
1 ORDN
  2 TYPE Deacon
```

The presence of a DATE tag and/or PLACe tag makes the assertion of when and/or where the event took place, and therefore that the event did happen. The absence of both of these tags require a Y(es) value on the parent TAG line to assert that the event happened. Using this convention protects GEDCOM processors which may remove (prune) lines that have no value and no subordinate lines. It also allows a note or source to be attached to the event context without implying that the event occurred.

It is not proper to use a N(o) value with an event tag to infer that it did not happen. Inferring that an event did not occur would require a different tag. A symbol such as using an exclamation mark (!) preceding an event tag to indicate an event is known not to have happened may be defined in the future.

LDS_INDIVIDUAL_ORDINANCE:=

```
[
n [ BAPL | CONL ] {1:1}
  +1 STAT <LDS_BAPTISM_DATE_STATUS>           {0:1}   p.45
  +1 DATE <DATE_LDS_ORD>                       {0:1}   p.41
  +1 TEMP <TEMPLE_CODE>                         {0:1}   p.54
  +1 PLAC <PLACE_LIVING_ORDINANCE>             {0:1}   p.51
  +1 <<SOURCE_CITATION>>                       {0:M}   p.34
  +1 <<NOTE_STRUCTURE>>                       {0:M}   p.33
|
n ENDL                                         {1:1}
  +1 STAT <LDS_ENDOWMENT_DATE_STATUS>         {0:1}   p.46
  +1 DATE <DATE_LDS_ORD>                       {0:1}   p.41
  +1 TEMP <TEMPLE_CODE>                         {0:1}   p.54
  +1 PLAC <PLACE_LIVING_ORDINANCE>             {0:1}   p.51
  +1 <<SOURCE_CITATION>>                       {0:M}   p.34
  +1 <<NOTE_STRUCTURE>>                       {0:M}   p.33
|
n SLGC                                         {1:1}
  +1 STAT <LDS_CHILD_SEALING_DATE_STATUS>     {0:1}   p.45
  +1 DATE <DATE_LDS_ORD>                       {0:1}   p.41
  +1 TEMP <TEMPLE_CODE>                         {0:1}   p.54
  +1 PLAC <PLACE_LIVING_ORDINANCE>             {0:1}   p.51
  +1 FAMC @<XREF:FAM>@                       {1:1}   p.55
```


+1 <<SOURCE_CITATION>>	{0:M}	p.34
+1 <<NOTE_STRUCTURE>>	{0:M}	p.33

]

LDS_SPOUSE_SEALING:=

n SLGS	{1:1}	
+1 STAT <LDS_SPOUSE_SEALING_DATE_STATUS>	{0:1}	p.46
+1 DATE <DATE_LDS_ORD>	{0:1}	p.41
+1 TEMP <TEMPLE_CODE>	{0:1}	p.54
+1 PLAC <PLACE_LIVING_ORDINANCE>	{0:1}	p.51
+1 <<SOURCE_CITATION>>	{0:M}	p.34
+1 <<NOTE_STRUCTURE>>	{0:M}	p.33

MULTIMEDIA_LINK:=

[/* embedded form*/		
n	OBJE @<XREF:OBJE>@	{1:1}	p.55
	/* linked form*/		
n	OBJE	{1:1}	p.55
	+1 FORM <MULTIMEDIA_FORMAT>	{1:1}	p.48
	+1 TITL <DESCRIPTIVE_TITLE>	{0:1}	p.43
	+1 FILE <MULTIMEDIA_FILE_REFERENCE>	{1:1}	p.47
	+1 <<NOTE_STRUCTURE>>	{0:M}	p.33
]			

This structure provides two options in handling the GEDCOM multimedia interface. The first alternative (embedded) includes all of the data, including the multimedia object, within the transmission file. The embedded method includes pointers to GEDCOM records that contain encoded image or sound objects. Each record represents a multimedia object or object fragment. An object fragment is created by breaking the multimedia files into several multimedia object records of 32K or less. These fragments are tied together by chaining from one multimedia object fragment to the next in sequence. This procedure will help manage the size of a multimedia GEDCOM record so that existing systems which are not expecting large multimedia records may discard the records without crashing due to the size of the record. Systems which handle embedded multimedia can reconstitute the multimedia fragments by decoding the object fragments and concatenating them to the assigned multimedia file.

The second method allows the GEDCOM context to be connected to an external multimedia file. This process is only managed by GEDCOM in the sense that the appropriate file name is included in the GEDCOM file in context, but the maintenance and transfer of the multimedia files are external to GEDCOM.

NOTE_STRUCTURE:=

[
n	NOTE @<XREF:NOTE>@	{1:1}	p.55
	+1 <<SOURCE_CITATION>>	{0:M}	p.34
n	NOTE [SUBMITTER_TEXT> <NULL>]	{1:1}	p.54
	+1 [CONC CONT] <SUBMITTER_TEXT>	{0:M}	

+1 <<SOURCE_CITATION>> {0:M} p.34
]

PERSONAL_NAME_STRUCTURE:=

n **NAME** <NAME_PERSONAL> {1:1} p.48
 +1 NPFX <NAME_PIECE_PREFIX> {0:1} p.49
 +1 GIVN <NAME_PIECE_GIVEN> {0:1} p.49
 +1 NICK <NAME_PIECE_NICKNAME> {0:1} p.49
 +1 SPFX <NAME_PIECE_SURNAME_PREFIX> {0:1} p.49
 +1 SURN <NAME_PIECE_SURNAME> {0:1} p.49
 +1 NSFX <NAME_PIECE_SUFFIX> {0:1} p.49
 +1 <<SOURCE_CITATION>> {0:M} p.34
 +1 <<NOTE_STRUCTURE>> {0:M} p.33

The name value is formed in the manner the name is normally spoken, with the given name and family name (surname) separated by slashes (/). (See <NAME_PERSONAL>, page 48.) Based on the dynamic nature or unknown compositions of naming conventions, it is difficult to provide more detailed name piece structure to handle every case. The NPFX, GIVN, NICK, SPFX, SURN, and NSFX tags are provided optionally for systems that cannot operate effectively with less structured information. For current future compatibility, all systems must construct their names based on the <NAME_PERSONAL> structure. Those using the optional name pieces should assume that few systems will process them, and most will not provide the name pieces. Future GEDCOM releases (6.0 and later) will likely apply a very different strategy to resolve this problem, possibly using a sophisticated parser and a name-knowledge database.

PLACE_STRUCTURE:=

n **PLAC** <PLACE_VALUE> {1:1} p.51
 +1 FORM <PLACE_HIERARCHY> {0:1} p.51
 +1 <<SOURCE_CITATION>> {0:M} p.34
 +1 <<NOTE_STRUCTURE>> {0:M} p.33

SOURCE_CITATION:=

[
 n **SOUR** @<XREF:SOUR>@ /* pointer to source record */ {1:1} p.55
 +1 PAGE <WHERE_WITHIN_SOURCE> {0:1} p.55
 +1 EVEN <EVENT_TYPE_CITED_FROM> {0:1} p.43
 +2 ROLE <ROLE_IN_EVENT> {0:1} p.53
 +1 DATA {0:1}
 +2 DATE <ENTRY_RECORDING_DATE> {0:1} p.43
 +2 TEXT <TEXT_FROM_SOURCE> {0:M} p.54
 +3 [CONC | CONT] <TEXT_FROM_SOURCE> {0:M}
 +1 QUAY <CERTAINTY_ASSESSMENT> {0:1} p.38
 +1 <<MULTIMEDIA_LINK>> {0:M} p.33,26
 +1 <<NOTE_STRUCTURE>> {0:M} p.33
 | /* Systems not using source records */
 n **SOUR** <SOURCE_DESCRIPTION> {1:1} p.53
 +1 [CONC | CONT] <SOURCE_DESCRIPTION> {0:M}
 +1 TEXT <TEXT_FROM_SOURCE> {0:M} p.54
 +2 [CONC | CONT] <TEXT_FROM_SOURCE> {0:M}

+1 <<NOTE_STRUCTURE>>

{0:M} p.33

]

The data provided in the <<SOURCE_CITATION>> structure is source-related information specific to the data being cited. (See GEDCOM examples starting on page 60.) Systems that do not use SOURCE_RECORDS must use the second SOURce citation structure option. When systems which support SOURCE_RECORD structures encounter source citations which do not contain pointers to source records, that system will need to create a SOURCE_RECORD and store the <SOURCE_DESCRIPTION> information found in the non-structured source citation in either the title area of that SOURCE_RECORD, or if the title field is not large enough, place a "(See Notes)" text in the title area, and place the unstructured source description in the source record's note field.

The information intended to be placed in the citation structure includes:

- A pointer to the SOURCE_RECORD, which contains a more general description of the source.
- Information, such as a page number, on how to find the cited data within the source.
- Actual text from the source that was used in making assertions, for example a date phrase as actually recorded or an applicable sentence from a letter, would be appropriate.
- Data that allows an assessment of the relative value of one source over another for making the recorded assertions (primary or secondary source, etc.). Data needed for this assessment is how much time from the asserted fact and when the source event was recorded, what type of event was cited, and what was the role of this person in the cited event.
 - Date when the entry was recorded in source document, ".SOUR.DATA.DATE."
 - Event that initiated the recording, ".SOUR.EVEN."
 - Role of this person in the event, ".SOUR.EVEN.ROLE".

SOURCE_REPOSITORY_CITATION:=

[

n REPO @XREF:REPO@

{1:1} p.55

+1 <<NOTE_STRUCTURE>>

{0:M} p.33

+1 CALN <SOURCE_CALL_NUMBER>

{0:M} p.53

+2 MEDI <SOURCE_MEDIA_TYPE>

{0:1} p.54

This structure is used within a source record to point to a name and address record of the holder of the source document. Formal and informal repository name and addresses are stored in the REPOSITORY_RECORD. Informal repositories include owner's of an unpublished work or of a rare published source, or a keeper of personal collections. An example would be the owner of a family Bible containing unpublished family genealogical entries. More formal repositories, such as the Family History Library, should show a call number of the source at that repository. The call number of that source should be recorded using a subordinate CALN tag. Systems which do not structure a repository name and address interface should store the information about where the source record is stored in the <<NOTE_STRUCTURE>> of this structure.

SPOUSE_TO_FAMILY_LINK:=

n FAMS @<XREF:FAM>@

{1:1} p.55

+1 <<NOTE_STRUCTURE>>

{0:M} p.33

Primitive Elements of the Lineage-Linked Form

The field sizes show the minimum recommended field length within a database that is constrained to fixed length fields. The field sizes are in addition to the GEDCOM level and tag overhead. GEDCOM lines are limited to 255 characters. However, the CONcAtenation or CONTInuation tags can be used to expand a field beyond this limit. CONT line implies that a new line should appear to preserve formatting. CONC implies concatenation to the previous line without a new line. This is used so that a text note or description can be processed (word wrapped) in a text window without fixed carriage returns. The CONT and CONC tags are being used to extend specified textual values.

ADDRESS_CITY:= {Size=1:60}

The name of the city used in the address. Isolated for sorting or indexing.

ADDRESS_COUNTRY:= {Size=1:60}

The name of the country that pertains to the associated address. Isolated by some systems for sorting or indexing. Used in most cases to facilitate automatic sorting of mail.

ADDRESS_LINE:= {Size=1:60}

Address information that, when combined with NAME and CONTInuation lines, meets requirements for sending communications through the mail.

ADDRESS_LINE1:= {Size=1:60}

The first line of the address used for indexing. This corresponds to the ADDRESS_LINE value of the ADDR line in the address structure.

ADDRESS_LINE2:= {Size=1:60}

The second line of the address used for indexing. This corresponds to the ADDRESS_LINE value of the first CONT line subordinate to the ADDR tag in the address structure.

ADDRESS_POSTAL_CODE:= {Size=1:10}

The ZIP or postal code used by the various localities in handling of mail. Isolated for sorting or indexing.

ADDRESS_STATE:= {Size=1:60}

The name of the state used in the address. Isolated for sorting or indexing.

ADOPTED_BY_WHICH_PARENT:= {Size=1:4}

[HUSB | WIFE | BOTH]

A code which shows which parent in the associated family record adopted this person.

Where:

HUSB = The HUSBand in the associated family adopted this person.

WIFE = The WIFE in the associated family adopted this person.

BOTH = Both HUSBand and WIFE adopted this person.

AGE_AT_EVENT:= {Size=1:12}

[< | > | <NULL>]

[YYy MMm DDDd | YYy | MMm | DDDd |

YYy MMm | YYy DDDd | MMm DDDd |

CHILD | INFANT | STILLBORN]

]

Where:

- > = greater than indicated age
- < = less than indicated age
- y = a label indicating years
- m = a label indicating months
- d = a label indicating days
- YY = number of full years
- MM = number of months
- DDD = number of days
- CHILD = age < 8 years
- INFANT = age < 1 year
- STILLBORN = died just prior, at, or near birth, 0 years

A number that indicates the age in years, months, and days that the principal was at the time of the associated event. Any labels must come after their corresponding number, for example; 4y 8m 10d.

ANCESTRAL_FILE_NUMBER:= {Size=1:12}

A unique permanent record number of an individual record contained in the Family History Department's Ancestral File.

APPROVED_SYSTEM_ID:= {Size=1:20}

A system identification name which was obtained through the GEDCOM registration process. This name must be unique from any other product. Spaces within the name must be substituted with a 0x5F (underscore _) so as to create one word.

ATTRIBUTE_TYPE:= {Size=1:4}

[CAST | EDUC | NATI | OCCU | PROP | RELI | RESI | TITL]

An attribute which may have caused name, addresses, phone numbers, family listings to be recorded. Its application is in helping to classify sources used for information.

AUTOMATED_RECORD_ID:= {Size=1:12}

A unique record identification number assigned to the record by the source system. This number is intended to serve as a more sure means of identification of a record between two interfacing systems.

CASTE_NAME:= {Size=1:90}

A name assigned to a particular group that this person was associated with, such as a particular racial group, religious group, or a group with an inherited status.

CAUSE_OF_EVENT:= {Size=1:90}

Used in special cases to record the reasons which precipitated an event. Normally this will be used subordinate to a death event to show cause of death, such as might be listed on a death certificate.

CERTAINTY_ASSESSMENT:= {Size=1:1}

[0 | 1 | 2 | 3]

The QUAY tag's value conveys the submitter's quantitative evaluation of the credibility of a piece of information, based upon its supporting evidence. Some systems use this feature to rank multiple

conflicting opinions for display of most likely information first. It is not intended to eliminate the receiver's need to evaluate the evidence for themselves.

- 0 = Unreliable evidence or estimated data
- 1 = Questionable reliability of evidence (interviews, census, oral genealogies, or potential for bias for example, an autobiography)
- 2 = Secondary evidence, data officially recorded sometime after event
- 3 = Direct and primary evidence used, or by dominance of the evidence

CHANGE_DATE:= {Size=10:11}
<DATE_EXACT>
The date that this data was changed.

CHARACTER_SET:= {Size=1:8}
[ANSEL | UNICODE | ASCII]
A code value that represents the character set to be used to interpret this data. The default character set is ANSEL, which includes ASCII as a subset. UNICODE is not widely supported by most operating systems; therefore, GEDCOM produced using the UNICODE character set will be limited in acceptance for some time. See Chapter 3, starting on page 63. ASCII contains the character set from 0x0 to 0xFF.

Note: The IBMPC character set is not allowed. This character set cannot be interpreted properly without knowing which code page the sender was using.

COPYRIGHT_GEDCOM_FILE:= {Size=1:90}
A copyright statement needed to protect the copyrights of the submitter of this GEDCOM file.

COPYRIGHT_SOURCE_DATA:= {Size=1:90}
A copyright statement required by the owner of data from which this information was down-loaded. For example, when a GEDCOM down-load is requested from the Ancestral File, this would be the copyright statement to indicate that the data came from a copyrighted source.

COUNT_OF_CHILDREN:= {Size=1:3}
The known number of children of this individual from all marriages or, if subordinate to a family record, the reported number of children known to belong to this family, regardless of whether the associated children are represented in the corresponding structure. This is not necessarily the count of children listed in a family structure.

COUNT_OF_MARRIAGES:= {Size=1:3}
The number of different families that this person was known to have been a member of as a spouse or parent, regardless of whether the associated families are represented in the GEDCOM file.

DATE:= {Size=4:35}
[<DATE_CALENDAR_ESCAPE> | <NULL>]
<DATE_CALENDAR>

DATE_APPROXIMATED:= {Size=4:35}
[

ABT <DATE> |
CAL <DATE> |
EST <DATE>
]

Where:

ABT = About, meaning the date is not exact.
CAL = Calculated mathematically, for example, from an event date and age.
EST = Estimated based on an algorithm using some other event date.

DATE_CALENDAR:= {Size=4:35}
[<DATE_GREG> | <DATE_JULN> | <DATE_HEBR> | <DATE_FREN> |
<DATE_FUTURE>]

The selection is based on the <DATE_CALENDAR_ESCAPE> that precedes the <DATE_CALENDAR> value immediately to the left. If <DATE_CALENDAR_ESCAPE> doesn't appear at this point, then @#DGREGORIAN@ is assumed. No future calendar types will use words (e.g., month names) from this list: FROM, TO, BEF, AFT, BET, AND, ABT, EST, CAL, or INT. When only a day and month appears as a DATE value it is considered a date phrase and not a valid date form.

Date Escape	Syntax Selected
-----	-----
@#DGREGORIAN@	<DATE_GREG>
@#DJULIAN@	<DATE_JULN>
@#DHEBREW@	<DATE_HEBR>
@#DFRENCH R@	<DATE_FREN>
@#DROMAN@	for future definition
@#DUNKNOWN@	calendar not known

DATE_CALENDAR_ESCAPE:= {Size=4:15}
[@#DHEBREW@ | @#DROMAN@ | @#DFRENCH R@ | @#DGREGORIAN@ |
@#DJULIAN@ | @#DUNKNOWN@]

The date escape determines the date interpretation by signifying which <DATE_CALENDAR> to use. The default calendar is the Gregorian calendar.

DATE_EXACT:= {Size=10:11}
<DAY> <MONTH> <YEAR_GREG>

DATE_FREN:= {Size=4:35}
[<YEAR> | <MONTH_FREN> <YEAR> |
<DAY> <MONTH_FREN> <YEAR>]

See <MONTH_FREN> page 40

DATE_GREG:= {Size=4:35}
[<YEAR_GREG> | <MONTH> <YEAR_GREG> |
<DAY> <MONTH> <YEAR_GREG>]

See <YEAR_GREG> page 56.

DATE_HEBR:= {Size=4:35}
[<YEAR> | <MONTH_HEBR> <YEAR> |
<DAY> <MONTH_HEBR> <YEAR>]

See <MONTH_HEBR> page 47

DATE_JULN:= {Size=4:35}
[<YEAR> | <MONTH> <YEAR> | <DAY> <MONTH> <YEAR>]

DATE_LDS_ORD:= {Size=4:35}
<DATE_VALUE>

LDS ordinance dates use only the Gregorian date and most often use the form of day, month, and year. Only in rare instances is there a partial date. The temple tag and code should always accompany temple ordinance dates. Sometimes the LDS_(ordinance)_DATE_STATUS is used to indicate that an ordinance date and temple code is not required, such as when BIC is used. (See LDS_(ordinance)_DATE_STATUS definitions beginning on page 45.)

DATE_PERIOD:= {Size=7:35}
[
FROM <DATE> |
TO <DATE> |
FROM <DATE> TO <DATE>
]

Where:

FROM = Indicates the beginning of a happening or state.

TO = Indicates the ending of a happening or state.

Examples:

FROM 1904 to 1915

= The state of some attribute existed from 1904 to 1915 inclusive.

FROM 1904

= The state of the attribute began in 1904 but the end date is unknown.

TO 1915

= The state ended in 1915 but the begin date is unknown.

DATE_PHRASE:= {Size=1:35}
(<TEXT>)

Any statement offered as a date when the year is not recognizable to a date parser, but which gives information about when an event occurred. The date phrase is enclosed in matching parentheses.

DATE_RANGE:= {Size=8:35}
[
BEF <DATE> |
AFT <DATE> |
BET <DATE> AND <DATE>
]

Where:

AFT = Event happened after the given date.

- BEF = Event happened before the given date.
- BET = Event happened some time between date 1 AND date 2. For example, bet 1904 and 1915 indicates that the event state (perhaps a single day) existed somewhere between 1904 and 1915 inclusive.

The date range differs from the date period in that the date range is an estimate that an event happened on a single date somewhere in the date range specified.

The following are equivalent and interchangeable:

Short form	Long Form
-----	-----
1852	BET 1 JAN 1852 AND 31 DEC 1852
1852	BET 1 JAN 1852 AND DEC 1852
1852	BET JAN 1852 AND 31 DEC 1852
1852	BET JAN 1852 AND DEC 1852
JAN 1920	BET 1 JAN 1920 AND 31 JAN 1920

DATE_VALUE:= {Size=1:35}
 [
 <DATE> |
 <DATE_PERIOD> |
 <DATE_RANGE>
 <DATE_APPROXIMATED> |
 INT <DATE> (<DATE_PHRASE>) |
 (<DATE_PHRASE>)
]

The DATE_VALUE represents the date of an activity, attribute, or event where:
 INT = Interpreted from knowledge about the associated date phrase included in parentheses.

An acceptable alternative to the date phrase choice is to use one of the other choices such as <DATE_APPROXIMATED> choice as the DATE line value and then include the <DATE_PHRASE> as a NOTE value subordinate to the DATE line.

The date value can take on the date form of just a date, an approximated date, between a date and another date, and from one date to another date. The preferred form of showing date imprecision, is to show, for example, MAY 1890 rather than ABT 12 MAY 1890. This is because limits have not been assigned to the precision of the prefixes such as ABT or EST.

DAY:= {Size=1:2}
 dd
 Day of the month, where dd is a numeric digit whose value is within the valid range of the days for the associated calendar month.

DESCRIPTIVE_TITLE:= {Size=1:248}
 The title of a work, record, item, or object.

DIGIT:= {Size=1:1}
 A single digit (0-9).

ENCODED_MULTIMEDIA_LINE:= {Size=1:87}
A line from a multimedia object which has been ENCODED. Each 54 characters of the multimedia object is encoded and stored in a 72 byte value. The encoding algorithm is used to convert binary objects into legal ASCII values which can be transmitted. See the encoding and decoding algorithms that are defined in Appendix E on page 91.

ENTRY_RECORDING_DATE:= {Size=1:90}
<DATE_VALUE>
The date that this event data was entered into the original source document.

EVENT_ATTRIBUTE_TYPE:= {Size=1:15}
[<EVENT_TYPE_INDIVIDUAL> |
<EVENT_TYPE_FAMILY> |
<ATTRIBUTE_TYPE>]
A code that classifies the principal event or happening that caused the source record entry to be created. If the event or attribute doesn't translate to one of these tag codes, then a user supplied code is expected, but will be considered as the generic tag EVEN for source certainty evaluation.

EVENT_DESCRIPTOR:= {Size=1:90}
A descriptor that should be used whenever the **EVEN** tag is used to define the event being cited. For example, if the event was a purchase of a residence, the **EVEN** tag would be followed by a subordinate **TYPE** tag with the value "Purchased Residence." Using this descriptor with any of the other defined event tags basically classifies the basic definition of the associated tag but does not change its basic process. The form of using the **TYPE** tag with defined event tags has not been used by very many products. The **MARR** tag could be subordinated with a **TYPE** tag and **EVENT_DESCRIPTOR** value of Common Law. Other possible descriptor values might include "Childbirth—unmarried," "Common Law," or "Tribal Custom," for example. The event descriptor should use the same word or phrase and in the same language, when possible, as was used by the recorder of the event. Systems that display data from the GEDCOM form should be able to display the descriptor value in their screen or printed output.

EVENT_TYPE_CITED_FROM:= {SIZE=1:15}
[<EVENT_ATTRIBUTE_TYPE>]
A code that indicates the type of event which was responsible for the source entry being recorded. For example, if the entry was created to record a birth of a child, then the type would be BIRT regardless of the assertions made from that record, such as the mother's name or mother's birth date. This will allow a prioritized best view choice and a determination of the certainty associated with the source used in asserting the cited fact.

EVENT_TYPE_FAMILY:= {Size=3:4}
[ANUL | CENS | DIV | DIVF | ENGA | MARR |
MARB | MARC | MARL | MARS | EVEN]

A code used to indicate the type of family event. The definition is the same as the corresponding event tag defined in Appendix A. (See Appendix A, starting on page 69).

EVENT_TYPE_INDIVIDUAL:= {Size=3:4}
[ADOP | BIRT | BAPM | BARM | BASM |
BLES | BURI | CENS | CHR | CHRA |

CONF | CREM | DEAT | EMIG | FCOM |
GRAD | IMMI | NATU | ORDN |
RETI | PROB | WILL | EVEN]

A code used to indicate the type of family event. The definition is the same as the corresponding event tag defined in Appendix A. (See Appendix A, starting on page 69).

EVENTS_RECORDED:= {Size=1:90}

[<EVENT_ATTRIBUTE_TYPE> |
<EVENTS_RECORDED>, <EVENT_ATTRIBUTE_TYPE>]

An enumeration of the different kinds of events that were recorded in a particular source. Each enumeration is separated by a comma. Such as a parish register of births, deaths, and marriages would be BIRT, DEAT, MARR.

FILE_NAME:= {Size=1:90}

The name of the GEDCOM transmission file. If the file name includes a file extension it must be shown in the form (filename.ext).

GEDCOM_CONTENT_DESCRIPTION:= {Size=1:248}

A note that a user enters to describe the contents of the lineage-linked file in terms of "ancestors or descendants of" so that the person receiving the data knows what genealogical information the transmission contains.

GEDCOM_FORM:= {Size=14:20}

[LINEAGE-LINKED]

The GEDCOM form used to construct this transmission. There maybe other forms used such as CommSoft's "EVENT_LINEAGE_LINKED" but these specifications define only the LINEAGE-LINKED Form. Systems will use this value to specify GEDCOM compatible with these specifications.

GENERATIONS_OF_ANCESTORS:= {Size=1:4}

The number of generations of ancestors included in this transmission. This value is usually provided when FamilySearch programs build a GEDCOM file for a patron requesting a download of ancestors.

GENERATIONS_OF_DESCENDANTS:= {Size=1:4}

The number of generations of descendants included in this transmission. This value is usually provided when FamilySearch programs build a GEDCOM file for a patron requesting a download of descendants.

LANGUAGE_ID:= {Size=1:15}

A table of valid latin language identification codes.

[Afrikaans | Albanian | Anglo-Saxon | Catalan | Catalan_Spn | Czech | Danish | Dutch | English | Esperanto | Estonian | Faroese | Finnish | French | German | Hawaiian | Hungarian | Icelandic | Indonesian | Italian | Latvian | Lithuanian | Navaho | Norwegian | Polish | Portuguese | Romanian | Serbo_Croa | Slovak | Slovene | Spanish | Swedish | Turkish | Wendic]

Other languages not supported until UNICODE

[Amharic | Arabic | Armenian | Assamese | Belorussian | Bengali | Braj | Bulgarian | Burmese | Cantonese | Church-Slavic | Dogri | Georgian | Greek | Gujarati | Hebrew | Hindi | Japanese |

Kannada | Khmer | Konkani | Korean | Lahnda | Lao | Macedonian | Maithili | Malayalam |
Mandrin | Manipuri | Marathi | Mewari | Nepali | Oriya | Pahari | Pali | Panjabi | Persian | Prakrit |
Pusto | Rajasthani | Russian | Sanskrit | Serb | Tagalog | Tamil | Telugu | Thai | Tibetan | Ukrainian
| Urdu | Vietnamese | Yiddish]

LANGUAGE_OF_TEXT:= {Size=1:15}
[<LANGUAGE_ID>]

The human language in which the data in the transmission is normally read or written. It is used primarily by programs to select language-specific sorting sequences and phonetic name matching algorithms.

LANGUAGE_PREFERENCE:= {Size=1:90}
[<LANGUAGE_ID>]

The language in which a person prefers to communicate. Multiple language preference is shown by using multiple occurrences in order of priority.

LDS_BAPTISM_DATE_STATUS:= {Size=5:10}
[CHILD | CLEARED | COMPLETED | INFANT | PRE-1970 |
QUALIFIED | STILLBORN | SUBMITTED | UNCLEARED]

A code indicating the status of an LDS baptism and confirmation date where:

CHILD = Died before eight years old.
CLEARED = Baptism has been cleared for temple ordinance.
COMPLETED = Completed but the date is not known.
INFANT = Died before less than one year old, baptism not required.
QUALIFIED = Ordinance request qualified by authorized criteria.
PRE-1970 = Ordinance is likely completed, another ordinance for this person was converted from temple records of work completed before 1970, therefore this ordinance is assumed to be complete until all records are converted.
STILLBORN = Stillborn, baptism not required.
SUBMITTED = Ordinance was previously submitted.
UNCLEARED = Data for clearing ordinance request was insufficient.

LDS_CHILD_SEALING_DATE_STATUS:= {Size=5:10}
[BIC | CLEARED | COMPLETED | DNS | PRE-1970 |
QUALIFIED | STILLBORN | SUBMITTED | UNCLEARED]

BIC = Born in the covenant receiving blessing of child to parent sealing.
CLEARED = Sealing has been cleared for temple ordinance.
COMPLETED = Completed but the date is not known.
DNS = This record is not being submitted for this temple ordinances.
QUALIFIED = Ordinance request qualified by authorized criteria.
PRE-1970 = (See pre-1970 under LDS_BAPTISM_DATE_STATUS on page 45.)
STILLBORN = Stillborn, not required.
SUBMITTED = Ordinance was previously submitted.
UNCLEARED = Data for clearing ordinance request was insufficient.

LDS_ENDOWMENT_DATE_STATUS:= {Size=5:10}
[CHILD | CLEARED | COMPLETED | INFANT | PRE-1970 |
QUALIFIED | STILLBORN | SUBMITTED | UNCLEARED]

A code indicating the status of an LDS endowment ordinance where:

- CHILD = Died before eight years old.
- CLEARED = Endowment has been cleared for temple ordinance.
- COMPLETED = Completed but the date is not known.
- INFANT = Died before less than one year old, baptism not required.
- QUALIFIED = Ordinance request qualified by authorized criteria.
- PRE-1970 = (See pre-1970 under LDS_BAPTISM_DATE_STATUS on page 45.)
- STILLBORN = Stillborn, ordinance not required.
- SUBMITTED = Ordinance was previously submitted.
- UNCLEARED = Data for clearing ordinance request was insufficient.

LDS_SPOUSE_SEALING_DATE_STATUS:= {Size=3:10}
[CANCELED | CLEARED | COMPLETED | DNS | DNS/CAN | PRE-1970 |
QUALIFIED | SUBMITTED | UNCLEARED]

- CANCELED = Canceled and considered invalid.
- CLEARED = Sealing has been cleared for temple ordinance.
- COMPLETED = Completed but the date is not known.
- DNS = This record is not being submitted for this temple ordinances.
- DNS/CAN = This record is not being submitted for this temple ordinances.
- QUALIFIED = Ordinance request qualified by authorized criteria.
- PRE-1970 = (See pre-1970 under LDS_BAPTISM_DATE_STATUS on page 45.)
- SUBMITTED = Ordinance was previously submitted.
- UNCLEARED = Data for clearing ordinance request was insufficient.

MONTH:= {Size=3}
[JAN | FEB | MAR | APR | MAY | JUN |
JUL | AUG | SEP | OCT | NOV | DEC]

Where:

- JAN = January
- FEB = February
- MAR = March
- APR = April
- MAY = May
- JUN = June
- JUL = July
- AUG = August
- SEP = September
- OCT = October
- NOV = November
- DEC = December

MONTH_FREN:= {Size=4}
[VEND | BRUM | FRIM | NIVO | PLUV | VENT | GERM |

FLOR | PRAI | MESS | THER | FRUC | COMP]

Where:

VEND = VENDEMIARE
BRUM = BRUMAIRE
FRIM = FRIMAIRE
NIVO = NIVOSE
PLUV = PLUVIOSE
VENT = VENTOSE
GERM = GERMINAL
FLOR = FLOREAL
PRAI = PRAIRIAL
MESS = MESSIDOR
THER = THERMIDOR
FRUC = FRUCTIDOR
COMP = JOUR_COMPLEMENTAIRES

MONTH_HEBR:= {Size=3}

[TSH | CSH | KSL | TVT | SHV | ADR | ADS |
NSN | IYR | SVN | TMZ | AAV | ELL]

Where:

TSH = Tishri
CSH = Cheshvan
KSL = Kislev
TVT = Tevet
SHV = Shevat
ADR = Adar
ADS = Adar Sheni
NSN = Nisan
IYR = Iyar
SVN = Sivan
TMZ = Tammuz
AAV = Av
ELL = Elul

MULTIMEDIA_FILE_REFERENCE:= {Size=1:30}

A complete local or remote file reference to the auxiliary data to be linked to the GEDCOM context. Remote reference would include a network address where the multimedia data may be obtained.

MULTIMEDIA_FORMAT:= {Size=3:4}

[bmp | gif | jpeg | ole | pcx | tiff | wav]

Indicates the format of the multimedia data associated with the specific GEDCOM context. This allows processors to determine whether they can process the data object. Any linked files should contain the data required, in the indicated format, to process the file data. Industry standards will emerge in this area and GEDCOM will then narrow its scope.

NAME_OF_BUSINESS:= {Size=1:90}

Name of the business, corporation, or person that produced or commissioned the product.

NAME_OF_FAMILY_FILE:= {Size=1:120}
Name under which family names for ordinances are stored in the temple's family file.

NAME_OF_PRODUCT:= {Size=1:90}
The name of the software product that produced this transmission.

NAME_OF_REPOSITORY:= {Size=1:90}
The official name of the archive in which the stated source material is stored.

NAME_OF_SOURCE_DATA:= {Size=1:90}
The name of the electronic data source that was used to obtain the data in this transmission. For example, the data may have been obtained from a CD-ROM disc that was named "U.S. 1880 CENSUS CD-ROM vol. 13."

NAME_PERSONAL:= {Size=1:120}

```
[  
<TEXT> |  
/<TEXT>/ |  
<TEXT> /<TEXT>/ |  
/<TEXT>/ <TEXT> |  
<TEXT> /<TEXT>/ <TEXT>  
]
```

The surname of an individual, if known, is enclosed between two slash (/) characters. The order of the name parts should be the order that the person would, by custom of their culture, have used when giving it to a recorder. Early versions of Personal Ancestral File[®] and other products did not use the trailing slash when the surname was the last element of the name. If part of name is illegible, that part is indicated by an ellipsis (...). Capitalize the name of a person or place in the conventional manner—capitalize the first letter of each part and lowercase the other letters, unless conventional usage is otherwise. For example: McMurray.

Examples:

William Lee (given name only or surname not known)
/Parry/ (surname only)
William Lee /Parry/
William Lee /Mac Parry/ (both parts (Mac and Parry) are surname parts)
William /Lee/ Parry (surname imbedded in the name string)
William Lee /Pa.../

NAME_PIECE:= {Size=1:90}
The piece of the name pertaining to the name part of interest. The surname part, the given name part, the name prefix part, or the name suffix part.

NAME_PIECE_GIVEN:= {Size=1:120}
[<NAME_PIECE> | <NAME_PIECE_GIVEN>, <NAME_PIECE>]
Given name or earned name. Different given names are separated by a comma.

NAME_PIECE_NICKNAME:= {Size=1:30}
[<NAME_PIECE> | <NAME_PIECE_NICKNAME>, <NAME_PIECE>]
A descriptive or familiar name used in connection with one's proper name.

NAME_PIECE_PREFIX:= {Size=1:30}

[<NAME_PIECE> | <NAME_PIECE_PREFIX>, <NAME_PIECE>]

Non indexing name piece that appears preceding the given name and surname parts. Different name prefix parts are separated by a comma.

For example:

Lt. Cmndr. Joseph /Allen/ jr.

In this example **Lt. Cmndr.** is considered as the name prefix portion.

NAME_PIECE_SUFFIX:= {Size=1:30}

[<NAME_PIECE> | <NAME_PIECE_SUFFIX>, <NAME_PIECE>]

Non-indexing name piece that appears after the given name and surname parts. Different name suffix parts are separated by a comma.

For example:

Lt. Cmndr. Joseph /Allen/ **jr.**

In this example **jr.** is considered as the name suffix portion.

NAME_PIECE_SURNAME:= {Size=1:120}

[<NAME_PIECE> | <NAME_PIECE_SURNAME>, <NAME_PIECE>]

Surname or family name. Different surnames are separated by a comma.

NAME_PIECE_SURNAME_PREFIX:= {Size=1:30}

[<NAME_PIECE> | <NAME_PIECE_SURNAME_PREFIX>, <NAME_PIECE>]

Surname prefix or article used in a family name. Different surname articles are separated by a comma, for example in the name "de la Cruz", this value would be "de, la".

NATIONAL_ID_NUMBER:= {Size=1:30}

A nationally-controlled number assigned to an individual. Commonly known national numbers should be assigned their own tag, such as SSN for U.S. Social Security Number. The use of the IDNO tag requires a subordinate TYPE tag to identify what kind of number is being stored.

For example:

n IDNO 43-456-1899

+1 TYPE Canadian Health Registration

NATIONAL_OR_TRIBAL_ORIGIN:= {Size=1:120}

The person's division of national origin or other folk, house, kindred, lineage, or tribal interest.

Examples: Irish, Swede, Egyptian Coptic, Sioux Dakota Rosebud, Apache Chiricawa, Navajo Bitter Water, Eastern Cherokee Taliwa Wolf, and so forth.

NEW_TAG:= {Size=1:15}

A user-defined tag that is contained in the GEDCOM current transmission. This tag must begin with an underscore (_) and should only be interpreted in the context of the sending system.

NOBILITY_TYPE_TITLE:= {Size=1:120}

The title given to or used by a person, especially of royalty or other noble class within a locality.

NULL:= {Size=0:0}

A convention that indicates the absence of any 8-bit ASCII character in the value including the null character (0x00) which is prohibited.

NUMBER:=
[<DIGIT> | <NUMBER>+<DIGIT>]

OCCUPATION:= {Size=1:90}
The kind of activity that an individual does for a job, profession, or principal activity.

ORDINANCE_PROCESS_FLAG:= {Size=2:3}
[yes | no]
A flag that indicates whether submission should be processed for clearing temple ordinances.

PEDIGREE_LINKAGE_TYPE:= {Size=5:7}
[adopted | birth | foster | sealing]
A code used to indicate the child to family relationship for pedigree navigation purposes.
Where:
adopted = indicates adoptive parents.
birth = indicates birth parents.
foster = indicates child was included in a foster or guardian family.
sealing = indicates child was sealed to parents other than birth parents.

PERMANENT_RECORD_FILE_NUMBER:= {Size=1:90}
<REGISTERED_RESOURCE_IDENTIFIER>:<RECORD_IDENTIFIER>
The record number that uniquely identifies this record within a registered network resource. The number will be usable as a cross-reference pointer. The use of the colon (:) is reserved to indicate the separation of the "registered resource identifier" (which precedes the colon) and the unique "record identifier" within that resource (which follows the colon). If the colon is used, implementations that check pointers should not expect to find a matching cross-reference identifier in the transmission but would find it in the indicated database within a network. Making resource files available to a public network is a future implementation.

PHONE_NUMBER:= {Size=1:25}
A phone number.

PHYSICAL_DESCRIPTION:= {Size=1:248}
An unstructured list of the attributes that describe the physical characteristics of a person, place, or object. Commas separate each attribute.
Example:
1 DSCR Hair Brown, Eyes Brown, Height 5 ft 8 in
2 DATE 23 JUL 1935

PLACE_HIERARCHY:= {Size=1:120}
This shows the jurisdictional entities that are named in a sequence from the lowest to the highest jurisdiction. The jurisdictions are separated by commas, and any jurisdiction's name that is missing is still accounted for by a comma. When a PLAC.FORM structure is included in the HEADER of a GEDCOM transmission, it implies that all place names follow this jurisdictional format and each jurisdiction is accounted for by a comma, whether the name is known or not. When the PLAC.FORM is subordinate to an event, it temporarily overrides the implications made by the PLAC.FORM structure stated in the HEADER. This usage is not common and, therefore, not encouraged. It should only be used when a system has over-structured its place-names.

PLACE_LIVING_ORDINANCE:= {Size=1:120}
<PLACE_VALUE>
The locality of the place where a living LDS ordinance took place. Usually only a living LDS baptism place is recorded in this field.

PLACE_VALUE:= {Size=1:120}
[
<TEXT> |
<TEXT>, <PLACE_VALUE>
]
The jurisdictional name of the place where the event took place. Jurisdictions are separated by commas, for example, "Cove, Cache, Utah, USA." If the actual jurisdictional names of these places have been identified, they can be shown using a PLAC.FORM structure either in the HEADER or in the event structure. (See <PLACE_HIERARCHY>, page 51.)

POSSESSIONS:= {Size=1:248}
A list of possessions (real estate or other property) belonging to this individual.

PUBLICATION_DATE:= {Size=10:11}
<DATE_EXACT>
The date this source was published or created.

RECEIVING_SYSTEM_NAME:= {Size=1:20}
The name of the system expected to process the GEDCOM-compatible transmission. The registered RECEIVING_SYSTEM_NAME for all GEDCOM submissions to the Family History Department **must** be one of the following names:

- "ANSTFILE" when submitting to Ancestral File.
- "TempleReady" when submitting for temple ordinance clearance.

RECORD_IDENTIFIER:= {Size=1:18}
An identification number assigned to each record within a specific database. The database to which the RECORD_IDENTIFIER pertains is indicated by the REGISTERED_RESOURCE_NUMBER which precedes the colon (:). If the RECORD_IDENTIFIER is not preceded by a colon, it is a reference to a record within the current GEDCOM transmission.

RECORD_TYPE:= {Size=3:4}
[FAM | INDI | NOTE | OBJE | REPO | SOUR | SUBM | SUBN]
An indicator of the record type being pointed to or used. For example if in an ASSOCIATION, an INDIVIDUAL record were to be ASSOCIATED with a FAM record then:
0 INDI
1 ASSO @F1@
2 TYPE FAM /* ASSOCIATION is with a FAM record.
2 RELA Witness at marriage

REGISTERED_RESOURCE_IDENTIFIER:= {Size=1:25}
This is an identifier assigned to a resource database that is available through access to a network. This is for future GEDCOM releases.

RELATION_IS_DESCRIPTOR:= {Size=1:25}
A word or phrase that states object 1's **relation** is object 2. For example you would read the following as "Joe Jacob's great grandson is the submitter pointed to by the @XREF:SUBM@":
0 INDI
1 NAME Joe /Jacob/
1 ASSO @<XREF:SUBM>@
2 TYPE great grandson

RELIGIOUS_AFFILIATION:= {Size=1:90}
A name of the religion with which this person, event, or record was affiliated.

RESPONSIBLE_AGENCY:= {Size=1:120}
The organization, institution, corporation, person, or other entity that has authority or control interests in the associated context. For example, an employer of a person of an associated occupation, or a church that administered rites or events, or an organization responsible for creating and/or archiving records.

RESTRICTION_NOTICE:= {Size=6:7}
[locked | privacy]
The restriction notice is defined for Ancestral File usage. Ancestral File download GEDCOM files may contain this data.

Where:

locked = Some records in Ancestral File have been satisfactorily proven by evidence, but because of source conflicts or incorrect traditions, there are repeated attempts to change this record. By arrangement, the Ancestral File Custodian can lock a record so that it cannot be changed without an agreement from the person assigned as the steward of such a record. The assigned steward is either the submitter listed for the record or Family History Support when no submitter is listed.

privacy = Information concerning this record is not present due to rights of or an approved request for privacy.

ROLE_DESCRIPTOR:= {Size=1:25}
A word or phrase that identifies a person's role in an event being described. This should be the same word or phrase, and in the same language, that the recorder used to define the role in the actual record.

ROLE_IN_EVENT:= {Size=1:15}
[CHIL | HUSB | WIFE | MOTH | FATH | SPOU | (<ROLE_DESCRIPTOR>)]
Indicates what role this person played in the event that is being cited in this context. For example, if you cite a child's birth record as the source of the mother's name, the value for this field is "MOTH." If you describe the groom of a marriage, the role is "HUSB." If the role is something different than one of the six relationship role tags listed above then enclose the role name within matching parentheses.

SCHOLASTIC_ACHIEVEMENT:= {Size=1:248}
A description of a scholastic or educational achievement or pursuit.

SEX_VALUE:= {Size=1:7}

A code that indicates the sex of the individual:

M = Male

F = Female

SOCIAL_SECURITY_NUMBER:= {Size=9:11}

A number assigned to a person in the United States for identification purposes.

SOURCE_CALL_NUMBER:= {Size=1:120}

An identification or reference description used to file and retrieve items from the holdings of a repository.

SOURCE_DESCRIPTION:= {Size=1:248}

A free form text block used to describe the source from which information was obtained. This text block is used by those systems which cannot use a pointer to a source record. It must contain a descriptive title, who created the work, where and when it was created, and where is source data stored. The developer should encourage users to use an appropriate style for forming this free form bibliographic reference. Developers are encouraged to support the SOURCE_RECORD method of reporting bibliographic reference descriptions.

SOURCE_DESCRIPTIVE_TITLE:= {Size=1:248}

The title of the work, record, or item and, when appropriate, the title of the larger work or series of which it is a part.

For a *published* work, a book for example, might have a title plus the title of the series of which the book is a part. A magazine article would have a title plus the title of the magazine that published the article.

For An *unpublished* work, such as:

- A letter might include the date, the sender, and the receiver.
- A transaction between a buyer and seller might have their names and the transaction date.
- A family Bible containing genealogical information might have past and present owners and a physical description of the book.
- A personal interview would cite the informant and interviewer.

SOURCE_FILED_BY_ENTRY:= {Size= 1:60}

This entry is to provide a short title used for sorting, filing, and retrieving source records.

SOURCE_JURISDICTION_PLACE:= {Size=1:120}

<PLACE_VALUE>

The name of the lowest jurisdiction that encompasses all lower-level places named in this source.

For example, "Oneida, Idaho" would be used as a source jurisdiction place for events occurring in the various towns within Oneida County. "Idaho" would be the source jurisdiction place if the events recorded took place in other counties as well as Oneida County.

SOURCE_MEDIA_TYPE:= {Size=1:15}

[audio | book | card | electronic | fiche | film | magazine |
manuscript | map | newspaper | photo | tombstone | video]

A code, selected from one of the media classifications choices above, that indicates the type of material in which the referenced source is stored.

SOURCE_ORIGINATOR:= {Size=1:248}

The person, agency, or entity who created the record. For a published work, this could be the author, compiler, transcriber, abstractor, or editor. For an unpublished source, this may be an individual, a government agency, church organization, or private organization, etc.

SOURCE_PUBLICATION_FACTS:= {Size=248}

When and where the record was created. For published works, this includes information such as the city of publication, name of the publisher, and year of publication.

For an unpublished work, it includes the date the record was created and the place where it was created. For example, the county and state of residence of a person making a declaration for a pension or the city and state of residence of the writer of a letter.

SUBMITTER_NAME:= {Size=1:60}

The name of the submitter formatted for display and address generation.

SUBMITTER_REGISTERED_RFN:= {Size=1:30}

A registered number of a submitter of Ancestral File data. This number is used in subsequent submissions or inquiries by the submitter for identification purposes.

SUBMITTER_TEXT:= {Size=1:248}

Comments or opinions from the submitter.

TEMPLE_CODE:= {Size=4:5}

An abbreviation of the temple in which LDS temple ordinances were performed. (See Appendix C, page 85.)

TEXT:= {Size=1:248}

A string composed of any valid character from the GEDCOM character set.

TEXT_FROM_SOURCE:= {Size=1:248}

<TEXT>

A verbatim copy of any description contained within the source. This indicates notes or text that are actually contained in the source document, not the submitter's opinion about the source. This should be, from the evidence point of view, "what the original record keeper said" as opposed to the researcher's interpretation. The word TEXT, in this case, means from the text which appeared in the source record including labels.

TIME_VALUE:= {Size=1:12}

[hh:mm:ss.fs]

The time of a specific event, usually a computer-timed event, where:

hh = hours on a 24-hour clock

mm = minutes

ss = seconds (optional)

fs = decimal fraction of a second (optional)

TRANSMISSION_DATE:=	{Size=10:11}
<DATE_EXACT>	
The date that this transmission was created.	
USER_REFERENCE_NUMBER:=	{Size=1:20}
A user-defined number or text that the submitter uses to identify this record. For instance, it may be a record number within the submitter's automated or manual system, or it may be a page and position number on a pedigree chart.	
USER_REFERENCE_TYPE:=	{Size=1:40}
A user-defined definition of the USER_REFERENCE_NUMBER.	
VERSION_NUMBER:=	{Size=1:15}
An identifier that represents the version level assigned to the associated product. It is defined and changed by the creators of the product.	
WHERE_WITHIN_SOURCE:=	{Size=1:248}
Specific location with in the information referenced. For a published work, this could include the volume of a multi-volume work and the page number(s). For a periodical, it could include volume, issue, and page numbers. For a newspaper, it could include a column number and page number. For an unpublished source, this could be a sheet number, page number, frame number, etc. A census record might have a line number or dwelling and family numbers in addition to the page number.	
XREF:=	{Size=1:22}
Either a pointer or an unique cross-reference identifier. If this element appears before the tag in a GEDCOM line, then it is a cross-reference identifier. If it appears after the tag in a GEDCOM line, then it is a pointer. The method of delimiting a pointer or cross-reference identifier is to enclose the pointer or cross-reference identifier within at signs (@), for example, @I123@. A XREF may not begin with a number sign (#). This is to avoid confusion with an escape sequence prefix (@#). The use of a colon (:) in the XREF is reserved for creating future network cross-references and the use of an exclamation (!) is reserved for intra-record pointers. Uniqueness of the cross-reference identifier is required within the transmission file.	
XREF:FAM:=	{Size=1:22}
A pointer to, or a cross-reference identifier of, a fam_record.	
XREF:INDI:=	{Size=1:22}
A pointer to, or a cross-reference identifier of, an individual record.	
XREF:NOTE:=	{Size=1:22}
A pointer to, or a cross-reference identifier of, a note record.	
XREF:OBJE:=	{Size=1:22}
A pointer to, or a cross-reference identifier of, a multimedia object.	
XREF:REPO:=	{Size=1:22}
A pointer to, or a cross-reference identifier of, a repository record.	
XREF:SOUR:=	{Size=1:22}

A pointer to, or a cross-reference identifier of, a SOURCE record.

XREF:SUBM:= {Size=1:22}

A pointer to, or a cross-reference identifier of, a SUBmitter record.

XREF:SUBN:= {Size=1:22}

A pointer to, or a cross-reference identifier of, a SUBmissioN record.

YEAR:= {Size=3:4}

A numeric representation of the calendar year in which an event occurred.

YEAR_GREG:= {Size=3:7}

[<NUMBER> | <NUMBER>/<DIGIT><DIGIT>]

The slash "/" <DIGIT><DIGIT> a year modifier which shows the possible date alternatives for pre-1752 date brought about by a changing the beginning of the year from MAR to JAN in the English calendar change of 1752, for example, 15 APR 1699/00. A (B.C.) appended to the <YEAR> indicates a date before the birth of Christ.

Compatibility with Other GEDCOM Versions

GEDCOM compatibility is measured on a per tag basis, and depends on how similar the data models are for the two different communicating products and on how consistently they understood and complied with the GEDCOM Standard. A few inconsistencies in the use of specific tags also crept into different releases of the standard itself, due to lack of foresight or inadvertent errors. Within these limits, GEDCOM compatible products can exchange data based on GEDCOM 2.0, 3.0, 4.0, and 5.x. Of course, newer GEDCOM releases significantly extend the data model for which the newer tag contexts will not be supported by older products. Some products have introduced their own variations into their GEDCOM form. This will likely provide unique problems with compatibility.

The following are areas in which incompatibilities may arise:

- **Source Structure:**

The SOURce structure was not supported by GEDCOM in versions before 5.x. However, some products, prior to GEDCOM 5.x, developed a SOURce structure for citing sources. These structures varied from product to product, which affects how source citations are interpreted. Products based on 5.x GEDCOM, prior to GEDCOM 5.4, may have used the more detailed source structure suggested by the previous 5.x versions. Older systems already handling sources will need to be modified. GEDCOM 5.x draft products are encouraged to update their programs to *The GEDCOM Standard 5.5* as soon as possible.

- **FAMC Pointer:**

The INDI.FAMC structure has been modified a lot since GEDCOM 4.0. In previous GEDCOM 5.x versions the FAMC structure may contain subordinate adoption events and/or LDS sealing to parent events. See the compatibility implications concerning the LDS sealing to parent event treated in the "LDS Ordinances Dates" in the next paragraph.

- **LDS Ordinance Dates:**

The structure for LDS sealing of child to parent was changed in previous GEDCOM 5.x draft versions from the FAM.CHIL.SLGC structure to the INDI.FAMC.SLGC structure. This was to allow access to child sealing information without having to follow a pointer to the family record. Personal Ancestral File 2.31 writes the sealing date in the FAM.CHIL.SLGC structure but reads this information from either format. A new major release of Personal Ancestral File will change to the newer approach.

GEDCOM 5.4 places the sealing of child to parent event at the same level as all of the other events that are subordinate to the INDIVIDUAL tag. If a system is keeping track of which family the individual is sealed to, then a FAMC pointer is additionally inserted subordinate to the SLGC event tag that points to the sealed-to-family.

To accommodate previous GEDCOM imports, systems handling the LDS ordinance events should look for the child sealing information in either INDI.SLGC (see LDS_INDIVIDUAL_ORDINANCE page 32, INDI.FAMC.SLGC or FAM.CHIL.SLGC structures. Ancestral File exports did not separate the temple code from the ordinance date. Ordinance dates down-loaded from Ancestral File may contain an ordinance date followed by a two digit temple code rather than a separate temple code line.

GEDCOM 4.x systems used certain key words as part of the ordinance dates. GEDCOM 5.x separated these codes from the dates and specified that they should be values of a subordinate STATUS tag. Previous GEDCOM 5.x implementations may have implemented this feature using a TYPE tag instead of the STATUS tag. (See <LDS_(ordinance)_DATE_STATUS>, page 45, 46.)

- **Adoption Events:**

In GEDCOM 5.x, the ADOPTION event was moved from the FAM.CHIL structure to the INDI.FAMC.ADOP structure, it also appears in the INDI.ADOP structure. In GEDCOM 5.4 the ADOPTION event appears only as an individual event which optionally contains a FAMC pointer to the adoptive family. Subordinate to this pointer is another ADOPTION tag which indicates whether the HUSB or WIFE in the pointed at family was the adoptive parent (see <ADOPTED_BY_WHICH_PARENTS> primitive on page 37). Pedigree navigation is provided only by <<CHILD_TO_FAMILY_LINK>> structure found on page 29.

- **Codes in Event Date:**

Some applications, such as Personal Ancestral File, pass key words as part of certain event dates. Some of these key words were *INFANT*, *CHILD*, *STILLBORN*, etc. These have to do with being an approximate age at an event.

In this version of GEDCOM, the information has been removed from the date value and specified by an <AGE_AT_EVENT> key word value which indicates a descriptive age value at the time of the enclosing event. (See <AGE_AT_EVENT>, page 37.) For example:

```
1 DEAT
  2 DATE 13 MAY 1984
  2 AGE STILLBORN
```

meaning this person died at age approximately 0 days old.

```
1 DEAT
  2 DATE 13 MAY 1984
  2 AGE INFANT
```

meaning this person died at age less than 1 year old.

- **Multiple Names:**

GEDCOM 5.x requires listing different names in different NAME structures, with the preferred instance first, followed by less preferred names. However, Personal Ancestral File and other products that only handle one name may use only the last instance of a name from a GEDCOM transmission. This causes the preferred name to be dropped when more than one name is present. The same thing often happens with other multiple-instance tags when only one instance was expected by the receiving system. PAF and other products that handle one name only may drop the preferred name under this arrangement.

- **Alias Names:**

One or two systems used the ALIAS tag for representing multiple names. This form is not supported in the GEDCOM Standard version 5.5.

- **Event Structure:**

The address structure, as part of the place structure, provided more detail than desired for the PLACe structure. Therefore it was removed from beneath the place structure and added to the <<EVENT_DETAIL>> structure at the same level as PLACe. The SITE tag was also eliminated

from the PLACe structure since the site of an event is really part of an address, such as Primary Children's Hospital, 100N Medical Drive, Salt Lake City, Utah.

- **Supplemental Attributes or Facts:**

Sometimes other attributes or facts are used to describe an individual's actions, physical description, employment, education, places of residence, etc. These are not generally thought of as events. However, they are often described like events because they were observed at a particular time and/or place. GEDCOM 5.x lists these attributes under the <<INDIVIDUAL_ATTRIBUTE_STRUCTURE>> on page 31 and allows them to be recorded in the same way as events. The attribute definition allows a value on the same line as the attribute tag. In addition, it allows a subordinate date period, place and/or address, etc. to be transmitted, just as the events are. Previous versions, which handled just a tag and value, can be read as usual by handling the subordinate attribute detail as an exception.

Packaging the GEDCOM Transmission File

The GEDCOM transmission is normally created on a DOS or Macintosh[®] compatible diskette. The DOS filename extension is (.GED). Macintosh filenames do not use file extensions.

When the GEDCOM file is too large to fit on a single diskette, the file is divided after any whole-line (last character is the **terminator**), and the DOS filename extension becomes (G##) where (##) is (00) for the second disk, (01) for the third, and so forth. For Macintosh filenames, append the two digits to the subsequent filenames in parentheses. (See the example below.) This allows the receiving software to ensure that disks are read in the correct sequence.

Given that the user-supplied portion of the file name is SMITH, then the complete filenames for a three-disk transmission would be:

<u>Disk</u>	<u>DOS Filename</u>	<u>Macintosh Filename</u>
1	SMITH.GED	SMITH
2	SMITH.G00	SMITH(00)
3	SMITH.G01	SMITH(01)

The required GEDCOM HEADer record appears only on the first disk and the required TRLR (trailer) record appears only on the last disk and **must be followed by the terminator**.

User-Defined Tags

We do not encourage the use of user-defined GEDCOM tags. Applications requiring the use of non-standard tags should define them with a leading underscore so that they will not conflict with future GEDCOM standard tags. Systems that read user-defined tags must consider that they have meaning only with respect to a system contained in the HEAD.SOUR context.

Escape Sequence Format for the Lineage-Linked Form

The Lineage-Linked GEDCOM Form no longer uses the **escape** sequence feature provided in the GEDCOM grammar.

Sample Lineage-Linked GEDCOM Transmission

The example below shows how some of these value types appear in a valid GEDCOM lineage-linked transmission. The example is a sample transmission of genealogical information about three individuals who are members of the same family—father, mother, and child. In the example, "Joe/Williams/" is the value specified by the tag NAME under the INDI tag for the record (@3@). Other values in other lines, such as the birth date and place, provide additional information about Joe Williams. The value (@4@) specified by the FAMC tag is a pointer to the FAM_RECORD (@4@) of which Joe Williams is a child. Included also in this transmission example are three other record types: a source record, a submitter record, and a repository record. These records are pointed to from within other records in the transmission. This shows how pointer values can be used in creating Lineage-Linked GEDCOM Form.

Example: (**Indentation and bolding are added for readability only.**)

```
0 HEAD
  1 SOUR PAF
    2 VERS 2.1
  1 DEST ANSTFILE
  1 SUBM @5@
  1 SUBN @8@
  1 GEDC
    2 VERS 5.4
    2 FORM Lineage-Linked
  1 CHAR ANSEL
0 @1@ INDI
  1 NAME Robert Eugene/Williams/
  1 SEX M
  1 BIRT
    2 DATE 02 OCT 1822
    2 PLAC Weston, Madison, Connecticut
    2 SOUR @6@
      3 PAGE Sec. 2, p. 45
      3 EVEN BIRT
      4 ROLE CHIL
  1 DEAT
    2 DATE 14 APR 1905
    2 PLAC Stamford, Fairfield, CT
  1 BURI
    2 PLAC Spring Hill Cem., Stamford, CT
  1 RESI
    2 ADDR 73 North Ashley
      3 CONT Spencer, Utah UT84991
    2 DATE from 1900 to 1905
  1 FAMS @4@
  1 FAMS @9@
0 @2@ INDI
  1 NAME Mary Ann/Wilson/
  1 SEX F
  1 BIRT
    2 DATE BEF 1828
    2 PLAC Connecticut
  1 FAMS @4@
0 @3@ INDI
  1 NAME Joe/Williams/
  1 SEX M
```

1 BIRT
 2 DATE **11 JUN 1861**
 2 PLAC **Idaho Falls, Bonneville, Idaho**
 1 FAMC @4@
 1 FAMC @9@
 2 PEDI **Adopted**
 1 ADOP
 2 FAMC @9@
 2 Date **16 MAR 1864**
 1 SLGC
 2 FAMC @9@
 2 DATE **2 OCT 1987**
 2 TEMP **SLAKE**
 0 @4@ FAM
 1 MARR
 2 DATE **DEC 1859**
 2 PLAC **Rapid City, South Dakota**
 1 SLGS
 2 DATE **14 JUN 1975**
 2 TEMP **SLAKE**
 1 HUSB @1@
 1 WIFE @2@
 1 CHIL @3@
 0 @5@ SUBM
 1 NAME **Reldon /Poulson/**
 1 ADDR **1900 43rd Street West**
 2 CONT **Billings, MT 68051**
 1 PHON **(406) 555-1232**
 0 @6@ SOUR
 1 DATA
 2 EVEN **BIRT, DEAT, MARR**
 3 DATE **FROM Jan 1820 TO DEC 1825**
 2 PLAC **Madison, Connecticut**
 2 AGNC **Madison County Court, State of Connecticut**
 1 TITL **Madison County Birth, Death, and Marriage Records**
 1 ABBR **VITAL RECORDS**
 1 REPO @7@
 2 CALN **13B-1234.01**
 2 MEDI **Microfilm**
 0 @7@ REPO
 1 NAME **Family History Library**
 1 ADDR **35 N West Temple Street**
 2 CONT **Salt Lake City, Utah**
 2 CONT **UT 84150**
 0 @8@ SUBN
 1 SUBM @5@
 1 FAMF **Reg Poulson Family**
 1 TEMP **SLAKE**
 0 @9@ FAM
 1 HUSB @1@
 1 CHIL @3@
 0 TRLR

The following is an example of a SOURCE_CITATION subordinate to the birth event being cited that does not contain a pointer to a SOURCE_RECORD. (This is not encouraged.)

0 INDI

1 NAME Fred /Jones/

1 BIRT

2 DATE 14 MAY 1812

2 PLAC Tonbridge, Kent, England

2 SOUR Waters, Henry F., Genealogical Gleanings in England: Abstracts of

3 CONC Wills Relating to Early American Families. 2 vols., reprint 1901, 1907.

3 CONC Baltimore: Genealogical Publishing Co., 1981.

3 CONC Stored in Family History Library book 942 D2wh; films 481,057-58

3 CONC Vol 2, page 388.

Chapter 3

Using Character Sets in GEDCOM

Introduction

GEDCOM needs to accommodate different character sets to facilitate the sharing of genealogical data in different languages. To minimize the number of differing standards, we have chosen to have each system convert its usage to ANSEL, and eventually to UNICODE.

In January 1991, a Unicode Consortium was founded to promote the use of the Unicode standard, which accommodates most all characters in one character set. (See the section "Unicode," on page 64.) The Unicode Consortium has agreed with the ISO 10646 standard to merge, and Unicode will be a subset of the ISO 10646 international character encoding standard.

Currently, it is difficult to handle the two- and four-character code sequences (wide characters). Therefore, until multi-byte handling becomes more common, ANSEL will be used to represent Latin-based characters.

The GEDCOM Standard does not address the implementation methods for multilingual processing, such as keyboard arrangements, sorting sequences, or character and graphic representations (font styles, proportional spacing, and so forth) on the CRT or printers. However, the Unicode standard has defined formatting characters that will indicate the direction of the text presentation and other text formatting character code.

Systems using code pages to support diacritical characters must convert all characters above character codes 128 to its ANSEL representation for that code page.

Most of the genealogy systems developed so far use ASCII, ANSEL, or both. ANSEL accommodates the set of Latin-based languages, as explained below.

8-Bit ANSEL

The 8-Bit ANSEL (American National Standard for Extended Latin Alphabet Coded Character Set for Bibliographic Use, Z39.47-1985 copyright) is the preferred character set for GEDCOM. It is used for all transmissions of information unless another character set is specified.

Using this character set standard makes it possible to preserve the full integrity of the language by providing a method of using the standard ASCII character set and supplementing it with both non-spacing character modifiers (diacritic) as well as spacing special characters.

Note: Non-spacing means that the diacritic is printed without advancing the device's print position. The character being modified is then printed in the same position, resulting in a combined image of both the character and the diacritic(s).

Storing ANSEL requires storing the non-spacing graphic character(s) preceding the ASCII character that the diacritic is to modify. The ANSEL standard specifies an extended 8-bit configuration (above 128) to represent the spacing and non-spacing graphic characters that make up most of the Latin based

languages. ANSEL is a super-set of ASCII. The standard ASCII characters including the control characters are preserved.

ANSEL is known by two other names:

- ANSI Z39.47-1985
- American Library Association character set, used in library systems worldwide, including the MARC (Machine-Readable Catalog) format.

A description of the codes for the ANSEL character set has been reproduced with permission and is included with the printed version of *The GEDCOM Standard*. The description of ANSEL codes is not included in the electronic version. This description may be purchased from—

American National Standards Institute
1430 Broadway
New York, N.Y. 10018

The description of the ANSEL character set standard includes the following:

- An 8-Bit Code Table showing the ASCII and extended ANSEL codes
- An explanation or legend of these codes
- A chart that identifies the ANSEL Non-spacing Graphic Characters
- A chart that identifies the ASCII Control Characters
- A chart that identifies the ASCII Graphic Characters

Character set codes 0 through 127 are the same for 8-Bit ANSEL and 8-Bit ASCII (USA version—ANSI 8-Bit). Character set codes 128 through 255 are unique to the ANSEL character set.

ASCII (USA Version)

When a language does not *need* diacritic characters or other special characters, and if you are not transmitting binary data, you will find it convenient to use ASCII (8-bit USA version) if your computer already supports it. This is a standard of the American National Standards Institute (ANSI). Most of the basic printable characters of ANSEL and ASCII (USA version—ANSI 8-Bit) are identical.

UNICODE (ISO 10646)

The Unicode standard is a new character code designed to encode text for storage in computer files. It is a subset of the upcoming ISO 10646 standard. The design of the Unicode standard is based on the simplicity and consistency of today's prevalent character code set, extended ASCII code set, but goes far beyond ASCII's limited ability to encode only the Latin alphabet: the Unicode encoding provides the capacity to encode most all of the characters used for written languages throughout the world. In order to accommodate the many thousands of characters used in the international text, the Unicode standard uses a 16-bit code set instead of extended ASCII's 8-bit code set. This expansion provides codes for approximately 65,000 characters. The Unicode standard assigns each character a unique 16-bit value, and does not use complex modes or escape codes to specify modified characters or special cases. UNICODE may adopt a 32-bit code to represent characters which should allow for all character representations. The text representation of the Unicode 16-bit numbers is U+0041 which is assigned to the letter A, 65 decimal. The Unicode standard includes the Latin alphabet used for English, the Cyrillic alphabet used for Russian, the Greek, Hebrew, and Arabic alphabets. Other alphabets used in countries

across Europe, Africa, the Indian subcontinent, and Asia, such as Japanese Kana, Korean Hangul, and Chinese Bopomofo are included. The largest part of the Unicode standard is devoted to thousands of unified character codes for Chinese, Japanese, and Korean ideographs. (See "The Unicode standard", vol. 1 and 2, published by Addison-Wesley Publishing, for character code standards.)

The Unicode character set environment should eventually contain a set of character for all languages. If the Unicode environment is used to produce a GEDCOM transmission, the header record would also be in Unicode, requiring receiving systems to determine whether the transmission is Unicode or ASCII before they could interpret the GEDCOM header. This would be done by reading the first two bytes of the transmission. If the first two bytes are 0x30 and 0x20 then the transmission will be in either ASCII or ANSEL as determined by the header record. If the first two bytes are 0x30 and 0x00 then the transmission should be processed as a Unicode transmission. (Different platforms may reverse the position of the null byte, in which case the test would be for 0x00 and 0x30.)

How to Change Character Sets

The character set for an entire transmission is specified in the character set line of the header record.

The example below shows the specification in the header record:

```
Lvl Tag Value  
0 HEAD  
1 SOUR PAF  
2 VERS 2.1  
1 DEST ANSTFILE  
1 CHAR ANSEL
```

The character set change remains in effect until the TRLR record is encountered at the end of the transmission.

UNICODE character set should be used for multi-language support as soon as operating systems begin providing adequate storage and display support.

For more information about character sets, see the following:

- Extended Latin Alphabet Coded Character Set for Bibliographic Use. American National Standards (ANSI), Z39.47, 1985.
- "8-Bit ASCII—Structure and Rules." American National Standards (ANSI) X3.134.1–198x.
- "7-Bit and 8-Bit ASCII Supplemental Multilingual Graphic Character Set (ASCII Multilingual Set)" (manuscript). American National Standards (ANSI), X3.134.2–198x.
- "The Unicode standard", vol. 1 and 2, published by Addison-Wesley Publishing.

Chapter 4

GEDCOM Product Registration

GEDCOM Product Registration

Developers of GEDCOM compatible products should register their product with the Family History Department GEDCOM coordinator.

Registration means that:

- The registered GEDCOM product is listed in Family History Department publications and on the internet at gedcom.org as being GEDCOM 5.5 compatible.
- The product's sample output file has been reviewed according to the established standard, and exceptions have been reported.
- The developer will be informed of future GEDCOM issues.
- Submissions to Family History Department resource files will be accepted from users of registered products.

To register a GEDCOM product, a developer must send the following information to the GEDCOM coordinator:

- A file containing a small sample of the product's GEDCOM output for evaluation. All of the fields that the product manages must be included in this data so that it can be tested for compatibility with other developers' products.
- A proposed unique SOURce name that identifies the *product*, not the company. This identifier should be included in the GEDCOM header record as a value to the SOUR tag. This name can have up to 40 characters, can have mixed upper and lower case, and cannot have embedded spaces. Use either an underscore (_) to connect multiple words or else a combination of upper and lower case letters (for example, FamilyRecords or Family_Records, **not** Family Records). The Family History Department will ensure uniqueness within the first 10 characters of this name.
- Optionally a copy of the GEDCOM product with installation procedures and a text file containing relevant technical documentation about the product's GEDCOM implementation. The product will be protected and used only for GEDCOM output verification and in providing some support to the submitter or researcher of information using the product.

The sample output file, the evaluation results, and the developer's technical description will be publicly available through the internet at gedcom.org.

Send product registration information to:

EMAIL:
gedcom@gedcom.org

MAIL:
Family History Department
GEDCOM Coordinator—3T
50 East North Temple Street
Salt Lake City, UT 84150

TELEPHONE (USA):
801-240-4534
801-240-5225

Appendix A

Lineage-Linked GEDCOM Tag Definition

Introduction

Appendix A is a glossary of the tags used in the Lineage-Linked GEDCOM Form. These tags are used in a hierarchical structure to describe individuals in terms of their families, names, dates, places, events, roles, sources, relationships. Control information and other kinds of data intended for computer processing is also included. (An example of the tags used in the Lineage-Linked Form begins on page 60.) To ensure all transmitted information in the Lineage-Linked GEDCOM is uniformly identified the standardized tags cannot be placed in any other context than shown in Chapter 2. It is legal to extend the context of the form, but only by using user-defined tags which must begin with an underscore. This will not violate the lineage-linked GEDCOM standard unless the context for the grammar of the Lineage-Linked GEDCOM Form is violated. The use of the underscore in the user tag name is to signal a non-standard construct is being used. This notifies the reading system of a discrepancy and will avoid future conflicts with tags that may be standardized in subsequent GEDCOM releases.

Lineage-Linked GEDCOM Tag Definitions

This section provides the definitions of the standardized GEDCOM tags and shows their formal name inside of the {braces}. The formal names are not used in place of the tag. Full understanding must come from the context in which the tag is used.

ABBR {ABBREVIATION}:=

A short name of a title, description, or name.

ADDR {ADDRESS}:=

The contemporary place, usually required for postal purposes, of an individual, a submitter of information, a repository, a business, a school, or a company.

ADR1 {ADDRESS1}:=

The first line of an address.

ADR2 {ADDRESS2}:=

The second line of an address.

ADOP {ADOPTION}:=

Pertaining to creation of a child-parent relationship that does not exist biologically.

AFN {AFN}:=

A unique permanent record file number of an individual record stored in Ancestral File.

AGE {AGE}:=

The age of the individual at the time an event occurred, or the age listed in the document.

AGNC {AGENCY}:=

The institution or individual having authority and/or responsibility to manage or govern.

ALIA {ALIAS}:=

An indicator to link different record descriptions of a person who may be the same person.

ANCE {ANCESTORS}:=

Pertaining to forbearers of an individual.

ANCI {ANCES_INTEREST}:=

Indicates an interest in additional research for ancestors of this individual. (See also DESI, page 72.)

ANUL {ANNULMENT}:=

Declaring a marriage void from the beginning (never existed).

ASSO {ASSOCIATES}:=

An indicator to link friends, neighbors, relatives, or associates of an individual.

AUTH {AUTHOR}:=

The name of the individual who created or compiled information.

BAPL {BAPTISM-LDS}:=

The event of baptism performed at age eight or later by priesthood authority of the LDS Church. (See also BAPM, next)

BAPM {BAPTISM}:=

The event of baptism (not LDS), performed in infancy or later. (See also BAPL, above, and CHR, page 71.)

BARM {BAR_MITZVAH}:=

The ceremonial event held when a Jewish boy reaches age 13.

BASM {BAS_MITZVAH}:=

The ceremonial event held when a Jewish girl reaches age 13, also known as "Bat Mitzvah."

BIRT {BIRTH}:=

The event of entering into life.

BLES {BLESSING}:=

A religious event of bestowing divine care or intercession. Sometimes given in connection with a naming ceremony.

BLOB {BINARY_OBJECT}:=

A grouping of data used as input to a multimedia system that processes binary data to represent images, sound, and video.

BURI {BURIAL}:=

The event of the proper disposing of the mortal remains of a deceased person.

CALN {CALL_NUMBER}:=

The number used by a repository to identify the specific items in its collections.

CAST {CASTE}:=

The name of an individual's rank or status in society, based on racial or religious differences, or differences in wealth, inherited rank, profession, occupation, etc.

CAUS {CAUSE}:=

A description of the cause of the associated event or fact, such as the cause of death.

CENS {CENSUS}:=

The event of the periodic count of the population for a designated locality, such as a national or state Census.

CHAN {CHANGE}:=

Indicates a change, correction, or modification. Typically used in connection with a DATE to specify when a change in information occurred.

CHAR {CHARACTER}:=

An indicator of the character set used in writing this automated information.

CHIL {CHILD}:=

The natural, adopted, or sealed (LDS) child of a father and a mother.

CHR {CHRISTENING}:=

The religious event (not LDS) of baptizing and/or naming a child.

CHRA {ADULT_CHRISTENING}:=

The religious event (not LDS) of baptizing and/or naming an adult person.

CITY {CITY}:=

A lower level jurisdictional unit. Normally an incorporated municipal unit.

CONC {CONCATENATION}:=

An indicator that additional data belongs to the superior value. The information from the CONC value is to be connected to the value of the superior preceding line without a space and without a carriage return and/or new line character. Values that are split for a CONC tag must always be split at a non-space. If the value is split on a space the space will be lost when concatenation takes place. This is because of the treatment that spaces get as a GEDCOM delimiter, many GEDCOM values are trimmed of trailing spaces and some systems look for the first non-space starting after the tag to determine the beginning of the value.

CONF {CONFIRMATION}:=

The religious event (not LDS) of conferring the gift of the Holy Ghost and, among protestants, full church membership.

CONL {CONFIRMATION_L}:=

The religious event by which a person receives membership in the LDS Church.

CONT {CONTINUED}:=

An indicator that additional data belongs to the superior value. The information from the CONT value is to be connected to the value of the superior preceding line with a carriage return and/or new line character. Leading spaces could be important to the formatting of the resultant text. When importing values from CONT lines the reader should assume only one delimiter character following the CONT tag. Assume that the rest of the leading spaces are to be a part of the value.

COPR {COPYRIGHT}:=

A statement that accompanies data to protect it from unlawful duplication and distribution.

CORP {CORPORATE}:=

A name of an institution, agency, corporation, or company.

CREM {CREMATION}:=

Disposal of the remains of a person's body by fire.

CTRY {COUNTRY}:=

The name or code of the country.

DATA {DATA}:=

Pertaining to stored automated information.

DATE {DATE}:=

The time of an event in a calendar format.

DEAT {DEATH}:=

The event when mortal life terminates.

DESC {DESCENDANTS}:=

Pertaining to offspring of an individual.

DESI {DESCENDANT_INT}:=

Indicates an interest in research to identify additional descendants of this individual. (See also ANCI, page 70.)

DEST {DESTINATION}:=

A system receiving data.

DIV {DIVORCE}:=

An event of dissolving a marriage through civil action.

DIVF {DIVORCE_FILED}:=

An event of filing for a divorce by a spouse.

DSCR {PHY_DESCRIPTION}:=

The physical characteristics of a person, place, or thing.

EDUC {EDUCATION}:=

Indicator of a level of education attained.

EMIG {EMIGRATION}:=

An event of leaving one's homeland with the intent of residing elsewhere.

ENDL {ENDOWMENT}:=

A religious event where an endowment ordinance for an individual was performed by priesthood authority in an LDS temple.

ENGA {ENGAGEMENT}:=

An event of recording or announcing an agreement between two people to become married.

EVEN {EVENT}:=

A noteworthy happening related to an individual, a group, or an organization.

FAM {FAMILY}:=

Identifies a legal, common law, or other customary relationship of man and woman and their children, if any, or a family created by virtue of the birth of a child to its biological father and mother.

FAMC {FAMILY_CHILD}:=

Identifies the family in which an individual appears as a child.

FAMF {FAMILY_FILE}:=

Pertaining to, or the name of, a family file. Names stored in a file that are assigned to a family for doing temple ordinance work.

FAMS {FAMILY_SPOUSE}:=

Identifies the family in which an individual appears as a spouse.

FCOM {FIRST_COMMUNION}:=

A religious rite, the first act of sharing in the Lord's supper as part of church worship.

FILE {FILE}:=

An information storage place that is ordered and arranged for preservation and reference.

FORM {FORMAT}:=

An assigned name given to a consistent format in which information can be conveyed.

GEDC {GEDCOM}:=

Information about the use of GEDCOM in a transmission.

GIVN {GIVEN_NAME}

A given or earned name used for official identification of a person.

GRAD {GRADUATION}:=

An event of awarding educational diplomas or degrees to individuals.

HEAD {HEADER}:=

Identifies information pertaining to an entire GEDCOM transmission.

HUSB {HUSBAND}:=

An individual in the family role of a married man or father.

IDNO {IDENT_NUMBER}:=

A number assigned to identify a person within some significant external system.

IMMI {IMMIGRATION}:=

An event of entering into a new locality with the intent of residing there.

INDI {INDIVIDUAL}:=

A person.

LANG {LANGUAGE}:=

The name of the language used in a communication or transmission of information.

LEGA {LEGATEE}:=

A role of an individual acting as a person receiving a bequest or legal devise.

MARB {MARRIAGE_BANN}:=

An event of an official public notice given that two people intend to marry.

MARC {MARR_CONTRACT}:=

An event of recording a formal agreement of marriage, including the prenuptial agreement in which marriage partners reach agreement about the property rights of one or both, securing property to their children.

MARL {MARR_LICENSE}:=

An event of obtaining a legal license to marry.

MARR {MARRIAGE}:=

A legal, common-law, or customary event of creating a family unit of a man and a woman as husband and wife.

MARS {MARR_SETTLEMENT}:=

An event of creating an agreement between two people contemplating marriage, at which time they agree to release or modify property rights that would otherwise arise from the marriage.

MEDI {MEDIA}:=

Identifies information about the media or having to do with the medium in which information is stored.

NAME {NAME}:=

A word or combination of words used to help identify an individual, title, or other item. More than one NAME line should be used for people who were known by multiple names.

NATI {NATIONALITY}:=

The national heritage of an individual.

NATU {NATURALIZATION}:=

The event of obtaining citizenship.

NCHI {CHILDREN_COUNT}:=

The number of children that this person is known to be the parent of (all marriages) when subordinate to an individual, or that belong to this family when subordinate to a FAM_RECORD.

NICK {NICKNAME}:=

A descriptive or familiar that is used instead of, or in addition to, one's proper name.

NMR {MARRIAGE_COUNT}:=

The number of times this person has participated in a family as a spouse or parent.

NOTE {NOTE}:=

Additional information provided by the submitter for understanding the enclosing data.

NPFX {NAME_PREFIX}:=

Text which appears on a name line before the given and surname parts of a name.

i.e. (**Lt. Cmndr.**) Joseph /Allen/ jr.

In this example Lt. Cmndr. is considered as the name prefix portion.

NSFX {NAME_SUFFIX}:=

Text which appears on a name line after or behind the given and surname parts of a name.

i.e. Lt. Cmndr. Joseph /Allen/ (**jr.**)

In this example jr. is considered as the name suffix portion.

OBJE {OBJECT}:=

Pertaining to a grouping of attributes used in describing something. Usually referring to the data required to represent a multimedia object, such an audio recording, a photograph of a person, or an image of a document.

OCCU {OCCUPATION}:=

The type of work or profession of an individual.

ORDI {ORDINANCE}:=

Pertaining to a religious ordinance in general.

ORDN {ORDINATION}:=

A religious event of receiving authority to act in religious matters.

PAGE {PAGE}:=

A number or description to identify where information can be found in a referenced work.

PEDI {PEDIGREE}:=

Information pertaining to an individual to parent lineage chart.

PHON {PHONE}:=

A unique number assigned to access a specific telephone.

PLAC {PLACE}:=

A jurisdictional name to identify the place or location of an event.

POST {POSTAL_CODE}:=

A code used by a postal service to identify an area to facilitate mail handling.

PROB {PROBATE}:=

An event of judicial determination of the validity of a will. May indicate several related court activities over several dates.

PROP {PROPERTY}:=

Pertaining to possessions such as real estate or other property of interest.

PUBL {PUBLICATION}:=

Refers to when and/or were a work was published or created.

QUAY {QUALITY_OF_DATA}:=

An assessment of the certainty of the evidence to support the conclusion drawn from evidence.

REFN {REFERENCE}:=

A description or number used to identify an item for filing, storage, or other reference purposes.

RELA {RELATIONSHIP}:=

A relationship value between the indicated contexts.

RELI {RELIGION}:=

A religious denomination to which a person is affiliated or for which a record applies.

REPO {REPOSITORY}:=

An institution or person that has the specified item as part of their collection(s).

RESI {RESIDENCE}:=

The act of dwelling at an address for a period of time.

RESN {RESTRICTION}:=

A processing indicator signifying access to information has been denied or otherwise restricted.

RETI {RETIREMENT}:=

An event of exiting an occupational relationship with an employer after a qualifying time period.

RFN {REC_FILE_NUMBER}:=

A permanent number assigned to a record that uniquely identifies it within a known file.

RIN {REC_ID_NUMBER}:=

A number assigned to a record by an originating automated system that can be used by a receiving system to report results pertaining to that record.

ROLE {ROLE}:=

A name given to a role played by an individual in connection with an event.

SEX {SEX}:=

Indicates the sex of an individual--male or female.

SLGC {SEALING_CHILD}:=

A religious event pertaining to the sealing of a child to his or her parents in an LDS temple ceremony.

SLGS {SEALING_SPOUSE}:=

A religious event pertaining to the sealing of a husband and wife in an LDS temple ceremony.

SOUR {SOURCE}:=

The initial or original material from which information was obtained.

SPFX {SURN_PREFIX}:=

A name piece used as a non-indexing pre-part of a surname.

SSN {SOC_SEC_NUMBER}:=

A number assigned by the United States Social Security Administration. Used for tax identification purposes.

STAE {STATE}:=

A geographical division of a larger jurisdictional area, such as a State within the United States of America.

STAT {STATUS}:=

An assessment of the state or condition of something.

SUBM {SUBMITTER}:=

An individual or organization who contributes genealogical data to a file or transfers it to someone else.

SUBN {SUBMISSION}:=

Pertains to a collection of data issued for processing.

SURN {SURNAME}:=

A family name passed on or used by members of a family.

TEMP {TEMPLE}:=

The name or code that represents the name a temple of the LDS Church.

TEXT {TEXT}:=

The exact wording found in an original source document.

TIME {TIME}:=

A time value in a 24-hour clock format, including hours, minutes, and optional seconds, separated by a colon (:). Fractions of seconds are shown in decimal notation.

TITL {TITLE}:=

A description of a specific writing or other work, such as the title of a book when used in a source context, or a formal designation used by an individual in connection with positions of royalty or other social status, such as Grand Duke.

TRLR {TRAILER}:=

At level 0, specifies the end of a GEDCOM transmission.

TYPE {TYPE}:=

A further qualification to the meaning of the associated superior tag. The value does not have any computer processing reliability. It is more in the form of a short one or two word note that should be displayed any time the associated data is displayed.

VERS {VERSION}:=

Indicates which version of a product, item, or publication is being used or referenced.

WIFE {WIFE}:=

An individual in the role as a mother and/or married woman.

WILL {WILL}:=

A legal document treated as an event, by which a person disposes of his or her estate, to take effect after death. The event date is the date the will was signed while the person was alive. (See also PROBate, page 76.)

Appendix B
Structure Cross-Reference

STRUCTURE	STRUCTURE/RECORD	
<<ADDRESS_STRUCTURE>>	EVENT_DETAIL:=	p. 29
<<ADDRESS_STRUCTURE>>	HEADER:=	p. 23
<<ADDRESS_STRUCTURE>>	REPOSITORY_RECORD:=	p. 26
<<ADDRESS_STRUCTURE>>	SUBMITTER_RECORD:=	p. 28
<<ASSOCIATION_STRUCTURE>>	INDIVIDUAL_RECORD:=	p. 25
<<CHANGE_DATE>>	FAM_RECORD:=	p. 24
<<CHANGE_DATE>>	INDIVIDUAL_RECORD:=	p. 25
<<CHANGE_DATE>>	MULTIMEDIA_RECORD:=	p. 26
<<CHANGE_DATE>>	NOTE_RECORD:=	p. 26
<<CHANGE_DATE>>	REPOSITORY_RECORD:=	p. 26
<<CHANGE_DATE>>	SOURCE_RECORD:=	p. 27
<<CHANGE_DATE>>	SUBMITTER_RECORD:=	p. 28
<<CHILD_TO_FAMILY_LINK>>	INDIVIDUAL_RECORD:=	p. 25
<<EVENT_DETAIL>>	FAMILY_EVENT_STRUCTURE:=	p. 30
<<EVENT_DETAIL>>	INDIVIDUAL_ATTRIBUTE_STRUCTURE:=	p. 30
<<EVENT_DETAIL>>	INDIVIDUAL_EVENT_STRUCTURE:=	p. 31
<<FAMILY_EVENT_STRUCTURE>>	FAM_RECORD:=	p. 24
<<FAM_RECORD>>	RECORD:=	p. 24
<<HEADER>>	LINEAGE_LINKED_GEDCOM:=	p. 23
<<INDIVIDUAL_ATTRIBUTE_STRUCTURE>>	INDIVIDUAL_RECORD:=	p. 25
<<INDIVIDUAL_EVENT_STRUCTURE>>	INDIVIDUAL_RECORD:=	p. 25
<<INDIVIDUAL_RECORD>>	RECORD:=	p. 24
<<LDS_INDIVIDUAL_ORDINANCE>>	INDIVIDUAL_EVENT_STRUCTURE:=	p. 31
<<LDS_SPOUSE_SEALING>>	FAM_RECORD:=	p. 24
<<MULTIMEDIA_LINK>>	EVENT_DETAIL:=	p. 29
<<MULTIMEDIA_LINK>>	FAM_RECORD:=	p. 24
<<MULTIMEDIA_LINK>>	INDIVIDUAL_RECORD:=	p. 25
<<MULTIMEDIA_LINK>>	SOURCE_CITATION:=	p. 34
<<MULTIMEDIA_LINK>>	SOURCE_RECORD:=	p. 27
<<MULTIMEDIA_LINK>>	SUBMITTER_RECORD:=	p. 28
<<MULTIMEDIA_RECORD>>	RECORD:=	p. 24
<<NOTE_RECORD>>	RECORD:=	p. 24
<<NOTE_STRUCTURE>>	ASSOCIATION_STRUCTURE:=	p. 29
<<NOTE_STRUCTURE>>	CHANGE_DATE:=	p. 29
<<NOTE_STRUCTURE>>	CHILD_TO_FAMILY_LINK:=	p. 29
<<NOTE_STRUCTURE>>	EVENT_DETAIL:=	p. 29
<<NOTE_STRUCTURE>>	FAM_RECORD:=	p. 24
<<NOTE_STRUCTURE>>	INDIVIDUAL_RECORD:=	p. 25
<<NOTE_STRUCTURE>>	LDS_INDIVIDUAL_ORDINANCE:=	p. 32
<<NOTE_STRUCTURE>>	LDS_SPOUSE_SEALING:=	p. 33
<<NOTE_STRUCTURE>>	MULTIMEDIA_LINK:=	p. 33
<<NOTE_STRUCTURE>>	MULTIMEDIA_RECORD:=	p. 26
<<NOTE_STRUCTURE>>	PERSONAL_NAME_STRUCTURE:=	p. 34
<<NOTE_STRUCTURE>>	PLACE_STRUCTURE:=	p. 34
<<NOTE_STRUCTURE>>	REPOSITORY_RECORD:=	p. 26
<<NOTE_STRUCTURE>>	SOURCE_CITATION:=	p. 24
<<NOTE_STRUCTURE>>	SOURCE_RECORD:=	p. 27
<<NOTE_STRUCTURE>>	SOURCE_REPOSITORY_CITATION:=	p. 35
<<NOTE_STRUCTURE>>	SPOUSE_TO_FAMILY_LINK:=	p. 35
<<PERSONAL_NAME_STRUCTURE>>	INDIVIDUAL_RECORD:=	p. 25
<<PLACE_STRUCTURE>>	EVENT_DETAIL:=	p. 29
<<RECORD>>	LINEAGE_LINKED_GEDCOM:=	p. 23

STRUCTURE

STRUCTURE/RECORD

<<REPOSITORY_RECORD>>	RECORD:=	p. 24	
<<SOURCE_CITATION>>	ASSOCIATION_STRUCTURE:=		p. 29
<<SOURCE_CITATION>>	EVENT_DETAIL:=	p. 29	
<<SOURCE_CITATION>>	FAM_RECORD:=		p. 24
<<SOURCE_CITATION>>	INDIVIDUAL_RECORD:=		p. 25
<<SOURCE_CITATION>>	LDS_INDIVIDUAL_ORDINANCE:=	p. 32	
<<SOURCE_CITATION>>	LDS_SPOUSE_SEALING:=	p. 33	
<<SOURCE_CITATION>>	NOTE_RECORD:=	p. 26	
<<SOURCE_CITATION>>	NOTE_STRUCTURE:=		p. 33
<<SOURCE_CITATION>>	PERSONAL_NAME_STRUCTURE:=		p. 34
<<SOURCE_CITATION>>	PLACE_STRUCTURE:=		p. 34
<<SOURCE_CITATION>>	SOURCE_CITATION:=		p. 24
<<SOURCE_CITATION>>	SOURCE_RECORD:=		p. 27
<<SOURCE_RECORD>>	RECORD:=	p. 24	
<<SOURCE_REPOSITORY_CITATION>>	SOURCE_RECORD:=		p. 27
<<SPOUSE_TO_FAMILY_LINK>>	INDIVIDUAL_RECORD:=		p. 25
<<SUBMISSION_RECORD>>	LINEAGE_LINKED_GEDCOM:=		p. 23
<<SUBMITTER_RECORD>>	RECORD:=	p. 24	

Primitive Cross-Reference

PRIMITIVE	STRUCTURE/RECORD	
<ADDRESS_CITY>	ADDRESS_STRUCTURE:=	p. 29
<ADDRESS_COUNTRY>	ADDRESS_STRUCTURE:=	p. 29
<ADDRESS_LINE>	ADDRESS_STRUCTURE:=	p. 29
<ADDRESS_LINE1>	ADDRESS_STRUCTURE:=	p. 29
<ADDRESS_LINE2>	ADDRESS_STRUCTURE:=	p. 29
<ADDRESS_POSTAL_CODE>	ADDRESS_STRUCTURE:=	p. 29
<ADDRESS_STATE>	ADDRESS_STRUCTURE:=	p. 29
<ADOPTED_BY_WHICH_PARENT>	INDIVIDUAL_EVENT_STRUCTURE:=	p. 31
<AGE_AT_EVENT>	EVENT_DETAIL:=	p. 29
<AGE_AT_EVENT>	FAM_RECORD:=	p. 24
<ANCESTRAL_FILE_NUMBER>	INDIVIDUAL_RECORD:=	p. 25
<APPROVED_SYSTEM_ID>	HEADER:=	p. 23
<ATTRIBUTE_TYPE>	EVENT_ATTRIBUTE_TYPE:=	p. 43
<AUTOMATED_RECORD_ID>	FAM_RECORD:=	p. 24
<AUTOMATED_RECORD_ID>	INDIVIDUAL_RECORD:=	p. 25
<AUTOMATED_RECORD_ID>	MULTIMEDIA_RECORD:=	p. 26
<AUTOMATED_RECORD_ID>	NOTE_RECORD:=	p. 26
<AUTOMATED_RECORD_ID>	REPOSITORY_RECORD:=	p. 26
<AUTOMATED_RECORD_ID>	SOURCE_RECORD:=	p. 27
<AUTOMATED_RECORD_ID>	SUBMISSION_RECORD:=	p. 27
<AUTOMATED_RECORD_ID>	SUBMITTER_RECORD:=	p. 28
<CASTE_NAME>	INDIVIDUAL_EVENT_STRUCTURE:=	p. 31
<CAUSE_OF_EVENT>	EVENT_DETAIL:=	p. 29
<CERTAINTY_ASSESSMENT>	SOURCE_CITATION:=	p. 34
<CHANGE_DATE>	CHANGE_DATE:=	p. 29
<CHARACTER_SET>	HEADER:=	p. 23
<COPYRIGHT_GEDCOM_FILE>	HEADER:=	p. 23
<COPYRIGHT_SOURCE_DATA>	HEADER:=	p. 23
<COUNT_OF_CHILDREN>	FAM_RECORD:=	p. 24
<COUNT_OF_CHILDREN>	INDIVIDUAL_ATTRIBUTE_STRUCTURE:=	p. 30
<COUNT_OF_MARRIAGES>	INDIVIDUAL_ATTRIBUTE_STRUCTURE:=	p. 30
<DATE>	DATE_APPROXIMATED:=	p. 40
<DATE>	DATE_LDS_ORD:=	p. 41
<DATE>	DATE_PERIOD:=	p. 41
<DATE>	DATE_RANGE:=	p. 41
<DATE>	DATE_VALUE:=	p. 42
<DATE_APPROXIMATED>	DATE_VALUE:=	p. 42
<DATE_CALENDAR>	DATE:=	p. 39
<DATE_CALENDAR_ESCAPE>	DATE_CALENDAR:=	p. 40
<DATE_CALENDAR_ESCAPE>	DATE:=	p. 39
<DATE_EXACT>	CHANGE_DATE:=	p. 29
<DATE_EXACT>	PUBLICATION_DATE:=	p. 51
<DATE_EXACT>	TRANSMISSION_DATE:=	p. 55
<DATE_FREN>	DATE_CALENDAR:=	p. 40
<DATE_FUTURE>	DATE_CALENDAR:=	p. 40
<DATE_GREG>	DATE_CALENDAR:=	p. 40
<DATE_HEBR>	DATE_CALENDAR:=	p. 40
<DATE_JULN>	DATE_CALENDAR:=	p. 40
<DATE_LDS_ORD>	LDS_INDIVIDUAL_ORDINANCE:=	p. 32
<DATE_LDS_ORD>	LDS_SPOUSE_SEALING:=	p. 33
<DATE_PERIOD>	DATE_VALUE:=	p. 42
<DATE_PERIOD>	SOURCE_RECORD:=	p. 27
<DATE_PHRASE>	DATE_VALUE:=	p. 42

PRIMITIVE**STRUCTURE/RECORD**

<DATE_RANGE>	DATE_VALUE:=	p. 42
<DATE_VALUE>	ENTRY_RECORDING_DATE:=	p. 43
<DATE_VALUE>	EVENT_DETAIL:=	p. 29
<DAY>	DATE_EXACT:=	p. 40
<DAY>	DATE_GREG:=	p. 40
<DAY>	DATE_FREN:=	p. 40
<DAY>	DATE_HEBR:=	p. 41
<DAY>	DATE_JULN:=	p. 41
<DESCRIPTIVE_TITLE>	MULTIMEDIA_LINK:=	p. 33
<DESCRIPTIVE_TITLE>	MULTIMEDIA_RECORD:=	p. 26
<DIGIT>	NUMBER:=	p. 50
<DIGIT>	YEAR_GREG:=	p. 56
<ENCODED_MULTIMEDIA_LINE>	MULTIMEDIA_RECORD:=	p. 26
<ENTRY_RECORDING_DATE>	SOURCE_CITATION:=	p. 34
<EVENT_ATTRIBUTE_TYPE>	EVENT_TYPE_CITED_FROM:=	p. 43
<EVENT_ATTRIBUTE_TYPE>	EVENTS_RECORDED:=	p. 44
<EVENT_DESCRIPTOR>	EVENT_DETAIL:=	p. 29
<EVENT_TYPE_CITED_FROM>	SOURCE_CITATION:=	p. 34
<EVENT_TYPE_FAMILY>	EVENT_ATTRIBUTE_TYPE:=	p. 43
<EVENT_TYPE_INDIVIDUAL>	EVENT_ATTRIBUTE_TYPE:=	p. 43
<EVENTS_RECORDED>	SOURCE_RECORD:=	p. 27
<EVENTS_RECORDED>	EVENTS_RECORDED:=	p. 44
<FILE_NAME>	HEADER:=	p. 23
<GEDCOM_CONTENT_DESCRIPTION>	HEADER:=	p. 23
<GEDCOM_FORM>	HEADER:=	p. 23
<GENERATIONS_OF_ANCESTORS>	SUBMISSION_RECORD:=	p. 27
<GENERATIONS_OF_DESCENDANTS>	SUBMISSION_RECORD:=	p. 27
<LANGUAGE_ID>	LANGUAGE_OF_TEXT:=	p. 45
<LANGUAGE_ID>	LANGUAGE_PREFERENCE:=	p. 45
<LANGUAGE_OF_TEXT>	HEADER:=	p. 23
<LANGUAGE_PREFERENCE>	SUBMITTER_RECORD:=	p. 28
<LDS_BAPTISM_DATE_STATUS>	LDS_INDIVIDUAL_ORDINANCE:=	p. 32
<LDS_CHILD_SEALING_DATE_STATUS>	LDS_INDIVIDUAL_ORDINANCE:=	p. 32
<LDS_ENDOWMENT_DATE_STATUS>	LDS_INDIVIDUAL_ORDINANCE:=	p. 32
<LDS_SPOUSE_SEALING_DATE_STATUS>	LDS_SPOUSE_SEALING:=	p. 33
<MONTH>	DATE_EXACT:=	p. 29
<MONTH_FREN>	DATE_FREN:=	p. 40
<MONTH_HEBR>	DATE_HEBR:=	p. 41
<MULTIMEDIA_FILE_REFERENCE>	MULTIMEDIA_LINK:=	p. 33
<MULTIMEDIA_FORMAT>	MULTIMEDIA_LINK:=	p. 33
<MULTIMEDIA_FORMAT>	MULTIMEDIA_RECORD:=	p. 26
<NAME_OF_BUSINESS>	HEADER:=	p. 23
<NAME_OF_FAMILY_FILE>	SUBMISSION_RECORD:=	p. 27
<NAME_OF_PRODUCT>	HEADER:=	p. 23
<NAME_OF_REPOSITORY>	REPOSITORY_RECORD:=	p. 26
<NAME_OF_SOURCE_DATA>	HEADER:=	p. 23
<NAME_PERSONAL>	PERSONAL_NAME_STRUCTURE:=	p. 34
<NAME_PIECE>	NAME_PIECE_NICKNAME:=	p. 49
<NAME_PIECE>	NAME_PIECE_GIVEN:=	p. 49
<NAME_PIECE>	NAME_PIECE_SURNAME:=	p. 49
<NAME_PIECE>	NAME_PIECE_PREFIX:=	p. 49
<NAME_PIECE>	NAME_PIECE_SUFFIX:=	p. 49
<NAME_PIECE>	NAME_PIECE_SURNAME_PREFIX:=	p. 49
<NAME_PIECE_GIVEN>	PERSONAL_NAME_STRUCTURE:=	p. 34

PRIMITIVE**STRUCTURE/RECORD**

<NAME_PIECE_NICKNAME>	PERSONAL_NAME_STRUCTURE:=	p. 34
<NAME_PIECE_PREFIX>	PERSONAL_NAME_STRUCTURE:=	p. 34
<NAME_PIECE_SUFFIX>	PERSONAL_NAME_STRUCTURE:=	p. 34
<NAME_PIECE_SURNAME>	PERSONAL_NAME_STRUCTURE:=	p. 34
<NAME_PIECE_SURNAME_PREFIX>	PERSONAL_NAME_STRUCTURE:=	p. 34
<NATIONAL_ID_NUMBER>	INDIVIDUAL_ATTRIBUTE_STRUCTURE:=	p. 30
<NATIONAL_OR_TRIBAL_ORIGIN>	INDIVIDUAL_ATTRIBUTE_STRUCTURE:=	p. 30
<NOBILITY_TYPE_TITLE>	INDIVIDUAL_ATTRIBUTE_STRUCTURE:=	p. 30
<NUMBER>	YEAR_GREG:=	p. 56
<OCCUPATION>	INDIVIDUAL_ATTRIBUTE_STRUCTURE:=	p. 30
<ORDINANCE_PROCESS_FLAG>	SUBMISSION_RECORD:=	p. 27
<PEDIGREE_LINKAGE_TYPE>	CHILD_TO_FAMILY_LINK:=	p. 29
<PERMANENT_RECORD_FILE_NUMBER>	INDIVIDUAL_RECORD:=	p. 25
<PHONE_NUMBER>	ADDRESS_STRUCTURE:=	p. 29
<PHYSICAL_DESCRIPTION>	INDIVIDUAL_ATTRIBUTE_STRUCTURE:=	p. 30
<PLACE_HIERARCHY>	HEADER:=	p. 23
<PLACE_HIERARCHY>	PLACE_STRUCTURE:=	p. 34
<PLACE_HIERARCHY>	PLACE_VALUE:=	p. 51
<PLACE_LIVING_ORDINANCE>	LDS_INDIVIDUAL_ORDINANCE:=	p. 32
<PLACE_LIVING_ORDINANCE>	LDS_SPOUSE_SEALING:=	p. 33
<PLACE_VALUE>	PLACE_LIVING_ORDINANCE:=	p. 51
<PLACE_VALUE>	PLACE_STRUCTURE:=	p. 34
<PLACE_VALUE>	SOURCE_JURISDICTION_PLACE:=	p. 54
<POSSESSIONS>	INDIVIDUAL_ATTRIBUTE_STRUCTURE:=	p. 30
<PUBLICATION_DATE>	HEADER:=	p. 23
<RECEIVING_SYSTEM_NAME>	HEADER:=	p. 23
<RECORD_TYPE>	ASSOCIATION_STRUCTURE:=	p. 29
<RECORD_IDENTIFIER>	PERMANENT_RECORD_FILE_NUMBER:=	p. 50
<REGISTERED_RESOURCE_IDENTIFIER>	PERMANENT_RECORD_FILE_NUMBER:=	p. 50
<RELATION_IS_DESCRIPTOR>	INDIVIDUAL_RECORD:=	p. 25
<RELIGIOUS_AFFILIATION>	INDIVIDUAL_ATTRIBUTE_STRUCTURE:=	p. 30
<RESPONSIBLE_AGENCY>	EVENT_DETAIL:=	p. 29
<RESPONSIBLE_AGENCY>	SOURCE_RECORD:=	p. 27
<RESTRICTION_NOTICE>	INDIVIDUAL_RECORD:=	p. 25
<ROLE_DESCRIPTOR>	ROLE_IN_EVENT:=	p. 53
<ROLE_IN_EVENT>	SOURCE_CITATION:=	p. 34
<SCHOLASTIC_ACHIEVEMENT>	INDIVIDUAL_ATTRIBUTE_STRUCTURE:=	p. 30
<SEX_VALUE>	INDIVIDUAL_RECORD:=	p. 25
<SOCIAL_SECURITY_NUMBER>	INDIVIDUAL_ATTRIBUTE_STRUCTURE:=	p. 30
<SOURCE_CALL_NUMBER>	SOURCE_REPOSITORY_CITATION:=	p. 35
<SOURCE_DESCRIPTION>	SOURCE_CITATION:=	p. 34
<SOURCE_DESCRIPTIVE_TITLE>	SOURCE_RECORD:=	p. 27
<SOURCE_FILED_BY_ENTRY>	SOURCE_RECORD:=	p. 27
<SOURCE_JURISDICTION_PLACE>	SOURCE_RECORD:=	p. 27
<SOURCE_MEDIA_TYPE>	SOURCE_REPOSITORY_CITATION:=	p. 35
<SOURCE_ORIGINATOR>	SOURCE_RECORD:=	p. 27
<SOURCE_PUBLICATION_FACTS>	SOURCE_RECORD:=	p. 27
<SUBMITTER_NAME>	SUBMITTER_RECORD:=	p. 28
<SUBMITTER_REGISTERED_RFN>	SUBMITTER_RECORD:=	p. 28
<SUBMITTER_TEXT>	NOTE_RECORD:=	p. 26
<SUBMITTER_TEXT>	NOTE_STRUCTURE:=	p. 33
<TEMPLE_CODE>	LDS_INDIVIDUAL_ORDINANCE:=	p. 32
<TEMPLE_CODE>	LDS_SPOUSE_SEALING:=	p. 33
<TEMPLE_CODE>	SUBMISSION_RECORD:=	p. 27

PRIMITIVE**STRUCTURE/RECORD**

<TEXT>	DATE_PHRASE:=	p. 41
<TEXT>	NAME_PERSONAL:=	p. 48
<TEXT>	PLACE_VALUE:=	p. 51
<TEXT>	TEXT_FROM_SOURCE:=	p. 54
<TEXT_FROM_SOURCE>	SOURCE_CITATION:=	p. 34
<TEXT_FROM_SOURCE>	SOURCE_RECORD:=	p. 27
<TIME_VALUE>	CHANGE_DATE:=	p. 29
<TIME_VALUE>	HEADER:=	p. 23
<TRANSMISSION_DATE>	HEADER:=	p. 23
<USER_REFERENCE_NUMBER>	FAM_RECORD:=	p. 24
<USER_REFERENCE_NUMBER>	INDIVIDUAL_RECORD:=	p. 25
<USER_REFERENCE_NUMBER>	MULTIMEDIA_RECORD:=	p. 26
<USER_REFERENCE_NUMBER>	NOTE_RECORD:=	p. 26
<USER_REFERENCE_NUMBER>	REPOSITORY_RECORD:=	p. 26
<USER_REFERENCE_NUMBER>	SOURCE_RECORD:=	p. 27
<USER_REFERENCE_TYPE>	FAM_RECORD:=	p. 24
<USER_REFERENCE_TYPE>	INDIVIDUAL_RECORD:=	p. 25
<USER_REFERENCE_TYPE>	MULTIMEDIA_RECORD:=	p. 26
<USER_REFERENCE_TYPE>	NOTE_RECORD:=	p. 26
<USER_REFERENCE_TYPE>	REPOSITORY_RECORD:=	p. 26
<USER_REFERENCE_TYPE>	SOURCE_RECORD:=	p. 27
<VERSION_NUMBER>	HEADER:=	p. 23
<WHERE_WITHIN_SOURCE>	SOURCE_CITATION:=	p. 34
<XREF:FAM>	CHILD_TO_FAMILY_LINK:=	p. 29
<XREF:FAM>	FAM_RECORD:=	p. 24
<XREF:FAM>	INDIVIDUAL_EVENT_STRUCTURE:=	p. 31
<XREF:FAM>	LDS_INDIVIDUAL_ORDINANCE:=	p. 32
<XREF:FAM>	SPOUSE_TO_FAMILY_LINK:=	p. 35
<XREF:INDI>	FAM_RECORD:=	p. 24
<XREF:INDI>	INDIVIDUAL_RECORD:=	p. 25
<XREF:NOTE>	NOTE_STRUCTURE:=	p. 33
<XREF:NOTE>	NOTE_RECORD:=	p. 26
<XREF:OBJE>	MULTIMEDIA_LINK:=	p. 33
<XREF:OBJE>	MULTIMEDIA_RECORD:=	p. 26
<XREF:REPO>	REPOSITORY_RECORD:=	p. 26
<XREF:SOUR>	SOURCE_CITATION:=	p. 34
<XREF:SOUR>	SOURCE_RECORD:=	p. 27
<XREF:SUBM>	FAM_RECORD:=	p. 24
<XREF:SUBM>	INDIVIDUAL_RECORD:=	p. 25
<XREF:SUBM>	RELATION_IS_DESCRIPTOR:=	p. 52
<XREF:SUBM>	SUBMITTER_RECORD:=	p. 28
<YEAR>	YEAR_GREG:=	p. 56
<YEAR>	DATE_JULN:=	p. 41
<YEAR>	DATE_FREN:=	p. 40
<YEAR>	DATE_HEBR:=	p. 41
<YEAR_GREG>	DATE_EXACT:=	p. 29
<YEAR_GREG>	DATE_GREG:=	p. 40

Appendix C
LDS Temple Codes
(as of June 1995)

TEMPLE	ABBREVIATIONS	TEMPLE	ABBREVIATIONS
ALBERTA	ALBER AL	SEATTLE, WA	SEATT SE
APIA SAMOA	APIA AP	SEOUL, KOREA	SEOUL SO
ARIZONA	ARIZO AZ	ST. GEORGE, UT	SGEOR SG
ATLANTA GA	ATLAN AT	ST. LOUIS, MISSOURI	SLOUI —
BOGOTA COL.	BOGOT BG	STOCKHOLM, SWDN	STOCK ST
BOISE ID	BOISE BO	SWISS	SWISS SW
BOUNTIFUL UT	BOUNT —	SYDNEY, AUST	SYDNE SD
BUENOS AIRES	BAIRE BA	TAIPEI, TAIWAN	TAIPE TP
CHICAGO IL	CHICA CH	TOKYO, JAPAN	TOKYO TK
COCHABAMBA, BOLIVA	COCHA —	TORONTO, CAN.	TORNO TR
DALLAS, TX	DALLA DA	VERNAL, UT.	VERNA —
DENVER, CO	DENVE DV	WASHINGTON, DC	WASHI WA
ENDOWMENT HOUSE	EHOUS EH		
FRANKFURT	FRANK FR		
FREIBERG	FREIB FD		
GUATAMALA	GUATE GA		
GUAYAQUIL, ECUADOR	GUAYA GY		
HARTFORD, CONN	HARTF —		
HAWAII	HAWAI HA		
HONG KONG	HKONG —		
IDAHO FALLS, ID	IFALL IF		
JOHANNESBURG, S.A.	JOHAN JO		
JORDAN RIVER, UT	JRIVE JR		
LAS VEGAS, NV	LVEGA LV		
LIMA, PERU	LIMA LI		
LOGAN, UT	LOGAN LG		
LONDON	LONDO LD		
LOS ANGELES, Ca	LANGE LA		
MADRID, SPAIN	MADRI —		
MANILA, PHILIPPINES	MANIL MA		
MANTI, UT	MANTI MT		
MEXICO CITY	MEXIC MX		
MT. TIMPANOGAS, UT	MTIMP —		
NASHVILLE, TENN	NASHV —		
NAUVOO	NAUVO —		
NEW ZEALAND	NZEAL NZ		
NUKU'ALOFA, TONGA	NUKUA TG		
OAKLAND, CA	OAKLA OK		
OGDEN, UT	OGDEN OG		
ORLANDO, FL	ORLAN —		
PAPEETE, TAHITI	PAPEE TA		
PORTLAND, OR	PORTL PT		
PRESIDENT'S OFFICE	POFFI —		
PRESTON, ENG	PREST —		
PROVO, UT	PROVO —		
PVRECIFE, BRAZIL	RECIF —		
SALT LAKE, UT	SLAKE SL		
SAN DIEGO, CA	SDIEG SA		
SANTIAGO, CHILE	SANTI SN		
SANTO DOMINGO, D.R.	SDOMI —		
SAO PAULO, BRAZ	SPAUL SP		

Appendix D

ANSEL Character Set

The following tables show the spacing and non-spacing diacritic characters that are contained in the ANSEL set required by the languages supported by GEDCOM 5. This table was added to give help to those receiving the GEDCOM standard on disk. *The graphic characters shown are not always accurate, however the name of the diacritic and the decimal equivalent should agree with the ANSEL standard.*

- C/R column refers to the column and row of the American National Standard Z39.47-1985 table showing the ANSEL character graphic and its 8 bit binary representation.
- wrcode column shows the Wordperfect (code page and character number, for example 1,2) chosen as the closest representation of the diacritic as shown in Wordperfect 5.1 Appendix P.
- Dec column shows to the decimal equivalent for that diacritic as is used in the ANSEL character set. The hexadecimal equivalent is obtained from converting the C/R column in to a character hexadecimal number, for example 14/10 converts to EA hex or 234 dec.
- Name column gives the english name of the diacritic.
- example of use column shows an example of words using this diacritic. For the non-spacing diacritic, this mark appears before the character in which it should be superimposed.

ANSEL Non-spacing graphic characters 8-bit characters (required for languages supported)

C/R	wrcode	Dec	Graphic	Name	example of use
14/1	1,0	225	`	grave accent	règle
14/2	1,6	226	´	acute accent	está
14/3	1,3	227	^	circumflex accent	même
14/4	1,2	228	~	tilde	niño
14/5	1,8	229	-	macron	gājējs
14/6	1,22	230	˘	breve	altă
14/7	1,15	231	·	dot above	žaba
14/8	1,7	232	¨	umlaut (diaeresis)	öppna
14/9	1,19	233	ˇ	hacek	vždy
14/10	1,14	234	°	circle above (angstrom)	hår
14/13	1,10	237	¸	high comma, off center	rozdel'ovac
14/14	1,16	238	˝	double acute accent	időszaki
15/0	2,15	240	¸	cedilla	ça

C/R	wrcode	Dec	Graphic	Name	example of use
15/1	2,17	241	Ꞁ	right hook	vietaꞀ
15/6	2,7	246	—	underscore	<u>samar</u>
15/7	2,16	247	ꞁ	left hook	darziꞁa
15/14	1,9	254	’	high comma, centered	gèotermika

ANSEL Spacing graphic characters
8-bit

C/R	wrcode	Dec	Graphic	Name	example of use
10/1	1,152	161	Ł	slash L—uppercase	Łódź
10/2	1,80	162	Ø	slash O—uppercase	Øst
10/3	1,78	163	Ð	slash D—uppercase	Ðuro
10/4	1,88	164	Þ	thorn—uppercase	Þann
10/5	1,36	165	Æ	ligature AE—uppercase	Ægir
10/6	1,166	166	Œ	ligature OE—uppercase	Œuvre
10/7	1,6	167	´	miagkii znak	fakul´tet
10/8	1,1	168	·	middle dot	novel·la
10/14	1,11	174	’	alif	Un’ yusho
11/0	2,11	176	‘	ayn	fa‘il
11/1	1,153	177	ł	slash l—lowercase	rozbił
11/2	1,81	178	ø	slash o—lowercase	høj
11/3	1,79	179	đ	slash d—lowercase	đavola
11/4	1,89	180	þ	thorn—lowercase	þann
11/5	1,37	181	æ	ligature ae—lowercase	skæg
11/6	1,167	182	œ	ligature oe—lowercase	œuvre
11/8	1,24	184	ı	dotless i—lowercase	masalı
11/9	4,11	185	£	British pound	£5.00
11/10	1,87-86	186	ð (Ð)	eth	verður
12/3	4,23	195	©	copyright mark	©1993
12/5	4,8	197	¿	inverted question mark	¿Que
12/6	4,7	198	¡	inverted exclamation mark	¡Esta
12/15	1,23	207	ß	Es Zet	Preußen

APPENDIX E

Encoding and Decoding Algorithms for Multimedia Objects

Introduction:

Embedded multimedia objects in GEDCOM requires special handling. These objects are normally represented by binary files containing characters which interfere with data transmission protocols. This document describes how the binary characters are encoded for transmission and then decoded to rebuild the multimedia file.

The algorithm for encoding binary images, compatible with GEDCOM transmission, is similar to an encoding scheme that would be used in creating a hexadecimal representation, but it uses a base-64 number representation rather than base-16 number representation.

Encoding:

This algorithm is for converting multimedia images represented in binary numbers into a collection that does not contain any of the ASCII control characters. This conversion eliminates the occurrence of special characters such as the "@" which has special meaning to GEDCOM.

The encoding routine converts a binary multimedia file segment of from 1 to 54 bytes in length into an encoded GEDCOM line value of 2 to 64 bytes in length. This encoded value becomes the <ENCODED_MULTIMEDIA_LINE> used in the MULTIMEDIA_RECORD (see page 26.)

The algorithm accomplishes its goal using the following steps:

1. Each 3 bytes (24 bits) of the binary 1 to 54 character segment is divided into four (6-bit) values. Each of these (6-bit) values are converted into an (8-bit) character making a character whose hexadecimal representation is between 0x00 and 0x3F (0 to 63 decimal.)
2. Each of the 4 new characters represents an Encoding key which is used to obtain the new replacement character from an Encoding Table included in this appendix.
3. Exception processing may be required in processing the last 3 byte chunk of the 1 to 54 character segment, which may consist of 0, 1, or 2 bytes:

- | <u>Retrieved</u> | <u>Action</u> |
|------------------|--|
| a. 0 bytes | Pad the last 3 characters with 0xFF. The conversion is complete. |
| b. 1 byte: | Pad last two bytes with 0xFF then complete steps 1 and 2 above. |
| c. 2 bytes: | Pad last byte with 0xFF then complete steps 1 and 2 above. |
5. Repeat until all characters in the received line value has been substituted. The return value of new encoded characters should contain from 4 to 72 characters. The length of the return value will always be a multiple of 4.

Decoding:

The Decoding routine converts the encoded line value back into the original binary character multimedia file segment.

The decoding algorithm can be accomplished in the following steps:

1. Each encoded multimedia line segment is divided into sets of 4 (8-bit) characters.
2. Each of these characters becomes a decoding key used to look up a corresponding character from the Decoding Table. A new (24-bit) group is formed by concatenating the low-order 6 bits from each of the 4 characters obtained from the decoding table.
3. Divide this new 24 bit group created by step 2 into three (8-bit) characters and concatenate them into the stream of characters being built as the decoded results.
4. Processing ends when the 0xFF padded bytes are encountered.

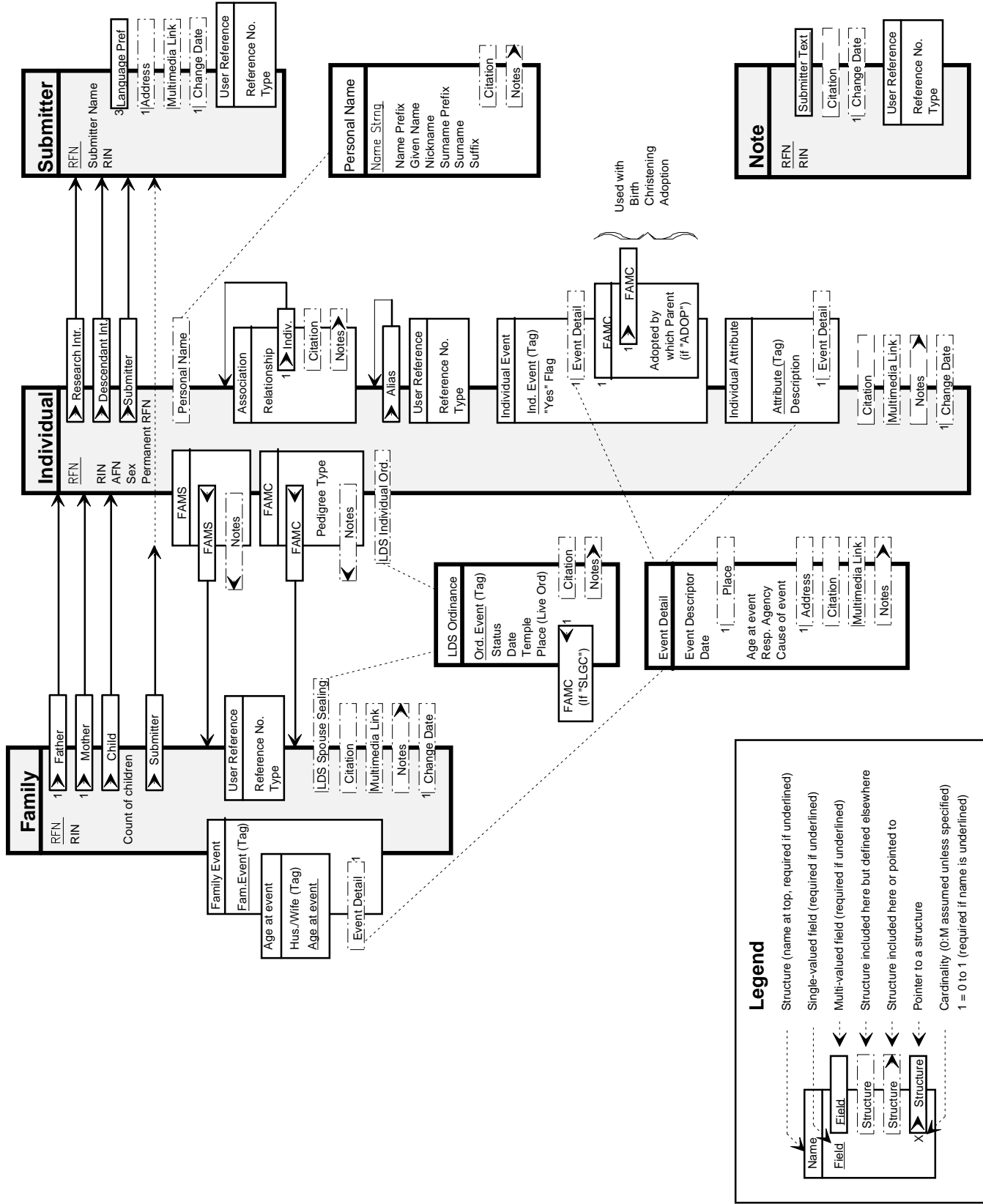
Encoding Table

<u>Encoding Key</u>	<u>Replacement Character</u>	<u>Encoding Key</u>	<u>Replacement Character</u>
0x00	0x2E .		----
0x01	0x2F /	0x26	0x61 a
0x02	0x30 0	0x27	0x62 b
0x03	0x31 1	0x28	0x63 c
0x04	0x32 2	0x29	0x64 d
0x05	0x33 3	0x2A	0x65 e
0x06	0x34 4	0x2B	0x66 f
0x07	0x35 5	0x2C	0x67 g
0x08	0x36 6	0x2D	0x68 h
0x09	0x37 7	0x2E	0x69 i
0x0A	0x38 8	0x2F	0x6A j
0x0B	0x39 9	0x30	0x6B k
	----	0x31	0x6C l
0x0C	0x41 A	0x32	0x6D m
0x0D	0x42 B	0x33	0x6E n
0x0E	0x43 C	0x34	0x6F o
0x0F	0x44 D	0x35	0x70 p
0x10	0x45 E	0x36	0x71 q
0x11	0x46 F	0x37	0x72 r
0x12	0x47 G	0x38	0x73 s
0x13	0x48 H	0x39	0x74 t
0x14	0x49 I	0x3A	0x75 u
0x15	0x4A J	0x3B	0x76 v
0x16	0x4B K	0x3C	0x77 w
0x17	0x4C L	0x3D	0x78 x
0x18	0x4D M	0x3E	0x79 y
0x19	0x4E N	0x3F	0x7A z
0x1A	0x4F O		
0x1B	0x50 P		
0x1C	0x51 Q		
0x1D	0x52 R		
0x1E	0x53 S		
0x1F	0x54 T		
0x20	0x55 U		
0x21	0x56 V		
0x22	0x57 W		
0x23	0x58 X		
0x24	0x59 Y		
0x25	0x5A Z		

Decoding Table

<u>Decoding Key</u>	<u>Replacement Character</u>	<u>Decoding Key</u>	<u>Replacement Character</u>
0x2E	.	0x00	
0x2F	/	0x01	
0x30	0	0x02	
0x31	1	0x03	
0x32	2	0x04	
0x33	3	0x05	
0x34	4	0x06	
0x35	5	0x07	
0x36	6	0x08	
0x37	7	0x09	
0x38	8	0x0A	
0x39	9	0x0B	
0x3A - 0x40	not valid		
0x41	A	0x0C	
0x42	B	0x0D	
0x43	C	0x0E	
0x44	D	0x0F	
0x45	E	0x10	
0x46	F	0x11	
0x47	G	0x12	
0x48	H	0x13	
0x49	I	0x14	
0x4A	J	0x15	
0x4B	K	0x16	
0x4C	L	0x17	
0x4D	M	0x18	
0x4E	N	0x19	
0x4F	O	0x1A	
0x50	P	0x1B	
0x51	Q	0x1C	
0x52	R	0x1D	
0x53	S	0x1E	
0x54	T	0x1F	
0x55	U	0x20	
0x56	V	0x21	
0x57	W	0x22	
0x58	X	0x23	
0x59	Y	0x24	
		0x5A	Z
		0x5B - 0x60	not valid
		0x61	a
		0x62	b
		0x63	c
		0x64	d
		0x65	e
		0x66	f
		0x67	g
		0x68	h
		0x69	i
		0x6A	j
		0x6B	k
		0x6C	l
		0x6D	m
		0x6E	n
		0x6F	o
		0x70	p
		0x71	q
		0x72	r
		0x73	s
		0x74	t
		0x75	u
		0x76	v
		0x77	w
		0x78	x
		0x79	y
		0x7A	z

GEDCOM 5.5 DATA MODEL CHART



GEDCOM 5.5 DATA MODEL CHART

