The MS Exchange Personal Addressbook Format

Author: Hans Dijkema
Email: h.dijkema@hum.org

After reengineering for several weeks, I've found out the following about the MS Exchange (Outlook) .PAB format. Various sources on the internet point out that the .PAB format turns out to be a format used by MAPI. What I have found out is the following:

The Main index

There's a pointer in the file at 0x000000c4 to a main index. This main index is an index of indices, or just an index of messages. This depends on something. What?

- The size of the .PAB file?
- The number of groups of addresses in the index file?

I don't know. In each case: The main index begins with:

0x00000008 − This seems to be an index of indices.
0x00000004 − This seems to be an index of messages.

The index of indices consists of three longs per entry:

1. The first long being the start−long ***S*** of a subindex. This is some control number. If the first long of the subindex doesn't match this long, it's probably not a subindex.
2. The second long is the end−long ***E*** of the subindex. Now we know when to stop reading the subindex.
3. The third long is the pointer to the subindex.

Reading (sub)indices (index of messages)

Each subindex consists of tuples of three longs. On each subindex the program expects the first long to be the same as the start long ***S*** (as stated above). After that we can start reading the subindex. We read until the end is reached, i.e. the second long ***E*** in the main index. This is the first long of the tuple.

Now the second long in the tuple is always the pointer to a data−record. Now we're getting where we want to be.

The usage of the third long in the tuple is unknown to me.

Reading the records

A record should begin with 0xbcec**** to be meaningfull. **** is the offset in the record to an index of variable fields. These are the fields we want to extract because they usually contain strings with information (names, email addresses, etc.). It consists of

shorts (i.e. 16 bit integers). The first **short** being the number of fields.

There’s also something like 0x7cec****. I don’t know what it is, but I expect it to be some index or so.

The index of fields consists always of 0x000c 0x0014, being what? 0x0014 turns to be an offset in the record to record types in most cases. What follows is a list of numbers that seem to describe record types. The index of fields only points to variable types like strings and arrays. The type of fields are in there. There’s a MAPI header file that has the kind of fields in it (see tags.h). I found out the following:

1. The types of fields are there as the lower word of the first of a pair of longs.
2. The upper word indicates extra information about what datatype we’re talking about.
3. The next long seems to indicate the sorting order of the types. They are ordered on type (ascending) but the entries in the ’index of fields’ are orderd on this long. That’s what I’ve found out.

   It looks like the index of offsets is an index to data that has no static size. So, we’ll have to recognize the data types we’re talking about. The most important will probably be:

T_MS_ARRAY	0x1100	An array of any type
T_MS_STRING	0x1e	This is a string

   The function to sort out of the datatypes are the ones we want to recognize is as follows:

   $$isUsed(type) = ((type = \mathrm{T\_MS\_STRING}) \vee (type \;\mathrm{and}\; \mathrm{T\_MS\_ARRAY} \neq 0))$$

   An other important issue: Although the order long word, describes the order of the entries in the record. The order of the field types also has a context. Business address type is the same as address type (home address type). The context of home or business bla blah could determine the correct interpretation. But I don’t know how to make the difference between these fields. They often have the same field−type number, although they are used differently. I don’t know how to determine the context.

   The next **short** in the index definitly points to the first offset of a real record. The field pointed to, ends at the next **short** (offset). So this is an array of fields.

Conclusion

Well, that’s about it. Follow the recipe and read the .PAB format. Or just use the code provided in pablib etc. See the *kmailcvt* project.

Appendix − Currently known field types for the .PAB format.

```
/***************************************************************************
                             pabtypes.h  -  description
                             -------------------
    begin                : Wed Aug 2 2000
    copyright            : (C) 2000 by Hans Dijkema
    email                : kmailcvt@hum.org
 ***************************************************************************/

/***************************************************************************
 *                                                                         *
 *   This program is free software; you can redistribute it and/or modify  *
 *   it under the terms of the GNU General Public License as published by  *
 *   the Free Software Foundation; either version 2 of the License, or     *
 *   (at your option) any later version.                                   *
 *                                                                         *
 * The information in this header file was reengineered from the various   *
 * .PAB files I used to test the .PAB conversion.                          *
 *                                                                         *
 * It turns out that the values below are the type information we searched *
 * for in the .PAB file. So we can use them to read the format.            *
 *                                                                         *
 ***************************************************************************/


#ifndef PABTYPES_H
#define PABTYPES_H 1.0

/*
 *  MS Windows tags as reengineered from an MS Exchange .PAB and
 *  Outlook .PAB file.
 */

/////////////////////////////////////////////////////////////////////////////

#define MS_GIVEN_NAME                        0x3a13
#define MS_GIVEN_NAME_1                      0x3a45
#define MS_GIVEN_NAME_2                      0x3a47
#define MS_GIVEN_NAME_3                             0x3a4f
#define MS_GIVEN_NAME_4                             0x3001
#define MS_GIVEN_NAME_5                             0x3a20
#define SET_MS_GIVEN_NAME                                           \
               MS_GIVEN_NAME,MS_GIVEN_NAME_1,MS_GIVEN_NAME_2,              \
               MS_GIVEN_NAME_3,MS_GIVEN_NAME_4,MS_GIVEN_NAME_5

/////////////////////////////////////////////////////////////////////////////

#define MS_EMAIL                                    0x3a56
#define MS_EMAIL_1                                  0x3003
#define SET_MS_EMAIL                                                \
               MS_EMAIL,MS_EMAIL_1

/////////////////////////////////////////////////////////////////////////////

#define MS_FIRSTNAME                                0x3a06
#define SET_MS_FIRSTNAME                                            \
               MS_FIRSTNAME

/////////////////////////////////////////////////////////////////////////////

#define MS_LASTNAME                                 0x3a11
#define SET_MS_LASTNAME                                                    \
               MS_LASTNAME


/////////////////////////////////////////////////////////////////////////////

#define MS_MIDDLENAME                               0x3a44
#define SET_MS_MIDDLENAME     \
               MS_MIDDLENAME

/////////////////////////////////////////////////////////////////////////////

#define MS_TITLE                                    0x3a17
#define SET_MS_TITLE          \
               MS_TITLE

/////////////////////////////////////////////////////////////////////////////

#define MS_ADDRESS                                  0x3a15
#define MS_ADDRESS_1                                0x3a29
#define MS_ADDRESS_2                                0x3a59
#define SET_MS_ADDRESS                \
               MS_ADDRESS, MS_ADDRESS_1, MS_ADDRESS_2
```

```
/////////////////////////////////////////////////////////////////////////////

#define MS_ZIP                                      0x3a5b
#define MS_ZIP_1                                    0x3a2a
#define SET_MS_ZIP            \
               MS_ZIP, MS_ZIP_1

/////////////////////////////////////////////////////////////////////////////

#define MS_STATE                                    0x3a28
#define MS_STATE_1                                  0x3a5c
#define SET_MS_STATE          \
               MS_STATE, MS_STATE_1

/////////////////////////////////////////////////////////////////////////////

#define MS_TOWN                                             0x3a27
#define MS_TOWN_1                                   0x3a59
#define SET_MS_TOWN           \
               MS_TOWN, MS_TOWN_1

/////////////////////////////////////////////////////////////////////////////

#define MS_COUNTRY                                  0x3a26
#define MS_COUNTRY_1                                0x3a5a
#define SET_MS_COUNTRY                \
               MS_COUNTRY, MS_COUNTRY_1

/////////////////////////////////////////////////////////////////////////////

#define MS_TEL                                      0x3a08
#define MS_TEL_1                                    0x3a09
#define MS_TEL_2                                    0x3a1a
#define MS_TEL_3                                    0x3a1b
#define MS_TEL_4                                    0x3a1f
#define MS_TEL_5                                    0x3a1d
#define MS_TEL_6                                    0x3a2d
#define MS_TEL_7                                    0x3a2f
#define SET_MS_TEL            \
               MS_TEL,MS_TEL_1,MS_TEL_2,MS_TEL_3,MS_TEL_4,  \
               MS_TEL_5,MS_TEL_6,MS_TEL_7

/////////////////////////////////////////////////////////////////////////////

#define MS_MOBILE                                   0x3a1c
#define MS_MOBILE_1                                 0x3a1e
#define MS_MOBILE_2                                 0x3a21
#define SET_MS_MOBILE         \
               MS_MOBILE,MS_MOBILE_1,MS_MOBILE_2

/////////////////////////////////////////////////////////////////////////////

#define MS_FAX                                      0x3a23
#define MS_FAX_1                                    0x3a24
#define MS_FAX_2                                    0x3a25
#define MS_FAX_3                                    0x3a2c
#define SET_MS_FAX            \
               MS_FAX,MS_FAX_1,MS_FAX_2,MS_FAX_3

/////////////////////////////////////////////////////////////////////////////

#define MS_ORG                                      0x3a16
#define SET_MS_ORGANIZATION   \
               MS_ORG

/////////////////////////////////////////////////////////////////////////////

#define MS_DEP                                      0x3a18
#define SET_MS_DEPARTMENT     \
               MS_DEP

/////////////////////////////////////////////////////////////////////////////

#define MS_COMMENT                                  0x3004
#define SET_MS_COMMENT                \
               MS_COMMENT

/////////////////////////////////////////////////////////////////////////////

#define SET_NOT_USED          \
               0x3002,        \
               0x300b,        \
               0x3a2e,        \
               0x3a30,        \
               0x3a19
               // 3002 probably address type
               // 300b some sort of key
               // 3a2e secretary tel number
               // 3a30 name of secretary
               // 3a19 office location


/////////////////////////////////////////////////////////////////////////////

/*
 * HP Openmail as reengineered from the X.400 .PAB file.
 */

/////////////////////////////////////////////////////////////////////////////

#define HP_OPENMAIL_JOB                            0x672b
#define HP_OPENMAIL_ORGANIZATION                   0x6728
#define HP_OPENMAIL_DEPARTMENT                     0x6729
#define HP_OPENMAIL_SUBDEP                         0x672b
#define HP_OPENMAIL_LOCATION_OF_WORK               0x672a

/////////////////////////////////////////////////////////////////////////////

#endif
```