A Multilingual Browser for the UN/EDIFACT Communication Standard

Christian Huemer[+] and A Min Tjoa[++]
[+]Institute of Applied Computer Science and Information Systems,
University of Vienna, Austria, email: ch@ifs.univie.ac.at
[++]Institute of Software Engineering, Technical University of Vienna,
Austria, email: amt@ifs.tuwien.ac.at

Abstract

Communication processes in the commercial sectors are increasingly directed towards computer technology of data exchange. This trend leads to the standardization of the data exchange. Among the several proposed standards, the UN/EDIFACT communication standard seems to become the de facto standard in the future. However, the general take-up of EDIFACT would not happen until a large number of companies is participating in the electronic data interchange. Thus, the EDIFACT standard must be introduced to each possible client, preferably in his / her mother tongue. This paper summarizes the problems and illustrates the necessary steps when implementing a multilingual information system for the UN/EDIFACT standard.

1. Introduction

In the field of administration, commerce and transport using faxes, electronic forms etc. over human/human or computer/human interfaces are straight forward, but slow and expensive because someone has to interpret and rekey the exchanged information. In addition to that, as soon as a human has to retype the information, the error rate increases. The solution to the speed, cost and error problems is to use machine/machine interfaces [1]. A precondition to use such interfaces is the definition of a standard interchange structure.

In the past decade various different standards have been proposed, but they are either limited to a certain branch of industry (e.g. ODETTE for the automobile industry) or build just a national solution (e.g. ANSI X.12 in the USA). Since 1988 the United Nations (UN) are developing the EDIFACT (electronic data interchange for administration, commerce and transport) standard to meet the requirements of an internationally valid general business standard [2]. This international standard [3] includes the rules on the application level for the structuring of user data and of associated service data in the interchange of messages in an open environment. Beside the syntax the EDIFACT standard covers also the definition of data elements (= the data information as basic component for message types), segments (= a functionally related set of data elements) and message types (structured representation of the full information on an electronic business transaction).

Since new message types are developed and definitions of existing message types are changing in the course of time, EDIFACT can be considered as a dynamic standard. The complete documentation of the EDIFACT guidelines of a period of time is included in an UN/EDIFACT directory which comprises the message type directory, the segment type directory, the composite data element type directory, the simple data element type directory and the code list directory. At the moment the EDIFACT directories are published by the UN either on paper or in ASCII format. However, at present the UN are developing the UN/EDIFACT Directory Definition Message (DIRDEF) which will allow the transmission of an UN/EDIFACT Directory set (or parts thereof) in EDIFACT syntax [4]. Another restriction concerning the widespread of EDIFACT is the fact, that companies of non English speaking countries have difficulties in overcoming the language barrier, since EDIFACT messages have so far only be developed and published in English. EDIFACT users in these countries have to work either with the English original message or with a - if there is any - translation made by their national standard institute (terminology center). Therefore, it was our intention to provide a multilingual EDIFACT terminology information system (METIS) which allows a flexible browsing through the UN/EDIFACT directories in the preferred language.

The rest of the paper is organized as follows. Section 2 deals with the complete functionality that METIS will offer. Section 3 describes the necessary procedures and problems caused by the import of new UN/EDIFACT messages into a multilingual environment. In Section 4 we present the object-oriented data model which METIS is based on.

Section 5 concerns the implementation of the browser function. We conclude with a short summary and with a prospect of future work to be done.

2. Specification of the browser

METIS (Multilingual EDIFACT Terminology Information System) is a browsing tool for MS Windows to navigate through the UN/EDIFACT standard directories in a very flexible and clearly arranged manner. METIS allows an access at each level of the EDIFACT hierarchy (messages, segments, composite data elements and single data elements) and offers links between these levels (see Fig. 1). Furthermore, METIS provides multilingual browser with respect of two facts. On the one hand the user can specify in which language he (or she) wants to be guided and on the other hand it is flexible enough to present the description of the EDIFACT standard in various languages.

In order to browse through an UN/EDIFACT directory it first has to be included into the METIS database. According to the basic EDI idea it is desirable to receive the UN/EDIFACT directory descriptions on-line from the various UN/EDIFACT reference databases (e.g. UN/EDIFACT reference database in Geneva, reference database of DIN in Germany) and to convert it into an adequate format for the database import. Therefore, METIS must have a gateway to the reference databases which is based on the DIRDEF message.

3. Import of new UN/EDIFACT directories

As mentioned above an UN/EDIFACT directory to be included has to be described in DIRDEF format (see Fig. 2), which is the only format the METIS converter is able to process. The METIS converting tool builds files in a relational format, by transforming each kind of segment into a separate file (table) and in the case of a 1-to-many-relationship - which is expressed by a following segment or segment group at the next level - by adding the primary key of the dominating segment to the table of the subordinate segments.This means, that the primary key `MessageId` of the table which is built from the segment `Message type`

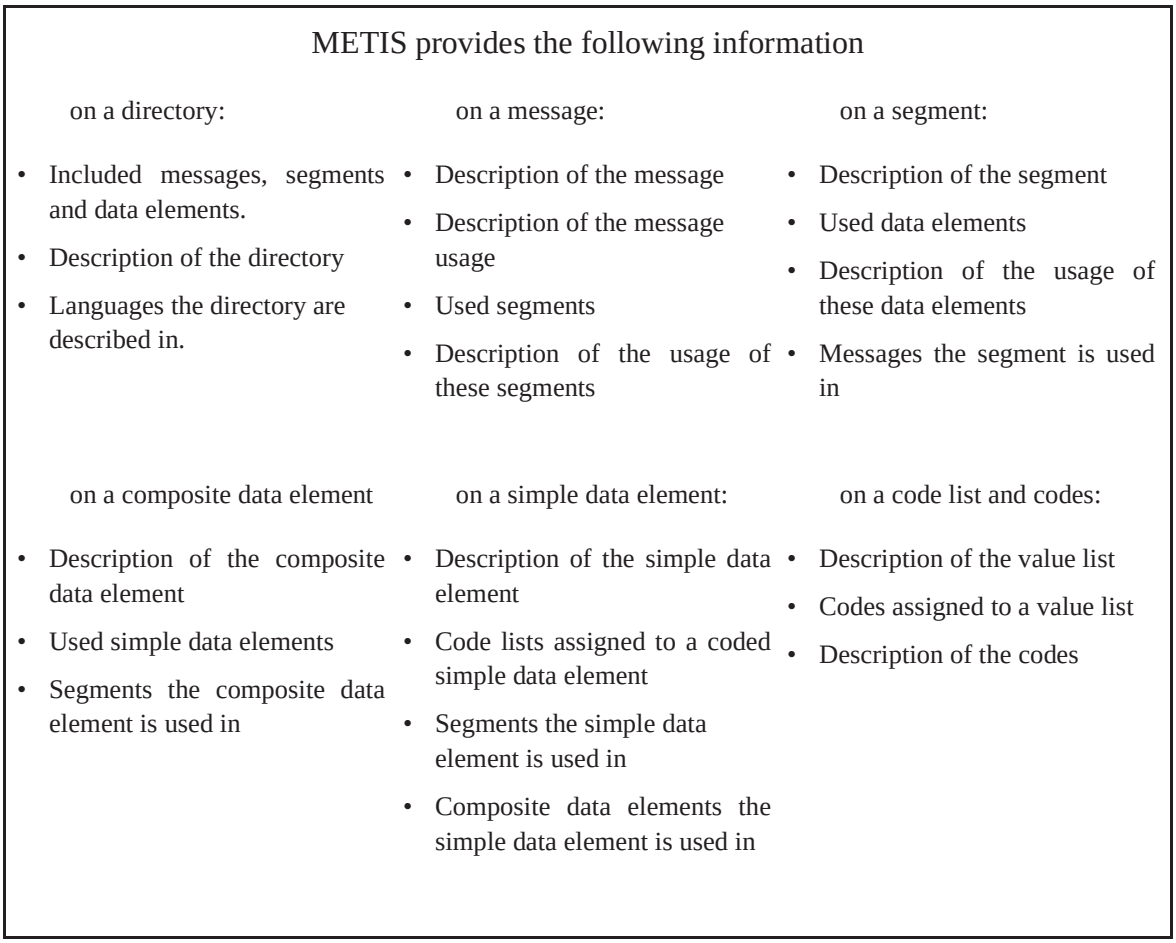


Fig. 1

`identification`, represents together with the `line number` the primary key for the table built from the following segment `Free text`.

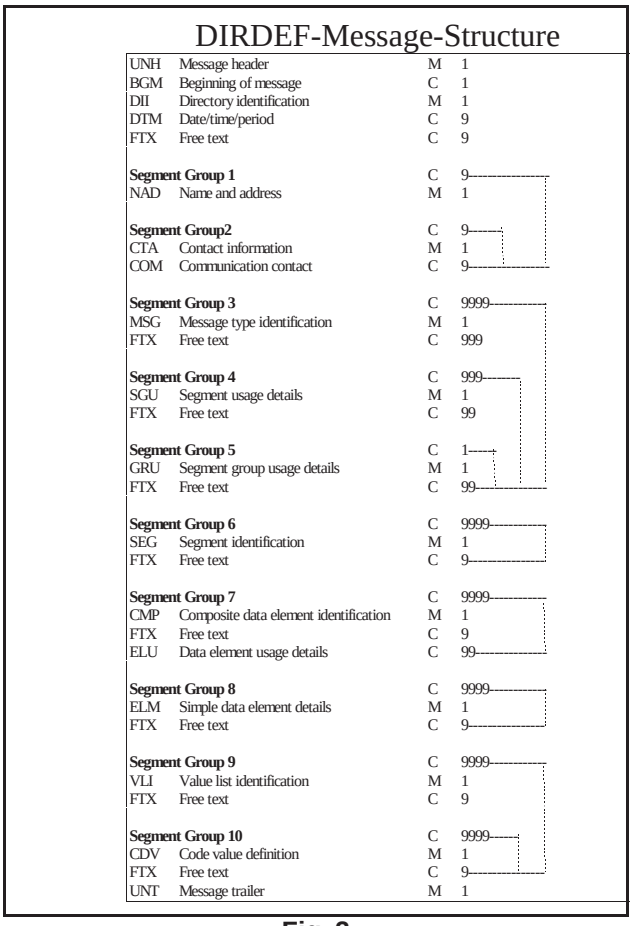


Fig. 2

Having performed the converting procedure one might think that the resulting files are ready for the input into the METIS database. Since the current DIRDEF format allows the description of an UN/EDIFACT directory in only one language, this would be true in a monolingual environment or in the case of storing each language version of a specific UN/EDIFACT directory independently of the other language versions of the same directory. For the purpose of saving storage capacity it would be desirable to store only the language specific attributes (mainly names and descriptions in free text format) in each language whereas storing the structure of an UN/EDIFACT directory only once. This assumes that DIRDEF messages describing the same UN/EDIFACT directory are identical except the language specific descriptions. Inasmuch as an UN/EDIFACT directory is an international standard, this should be quite usual. Nevertheless, in practice many of the non English DIRDEF messages, received from the national reference databases did not correspond to the standard definition. The inconsistencies are due to the facts, that the national reference databases are just establishing a report generator for the relatively new DIRDEF message and that the national terminology centers often adapt the international standard to the specific needs in their countries.

For the above mentioned reason it is a prerequisite, that the firstly received version of an UN/EDIFACT directory goes conform with the standard. Hence, it is advisable to firstly import the English DIRDEF message from a reliable EDIFACT reference database, like the CEBIS database in Brussels. Any further incoming DIRDEF message describing an already imported UN/EDIFACT directory in another language has to pass a consistency check. The appropriate tool removes all the structure information of the files resulting from the conversion process which are not included in the files of the original (English) DIRDEF message. In other words, the consistency checker removes all tupels from the new files which have a primary key not included in the original files.

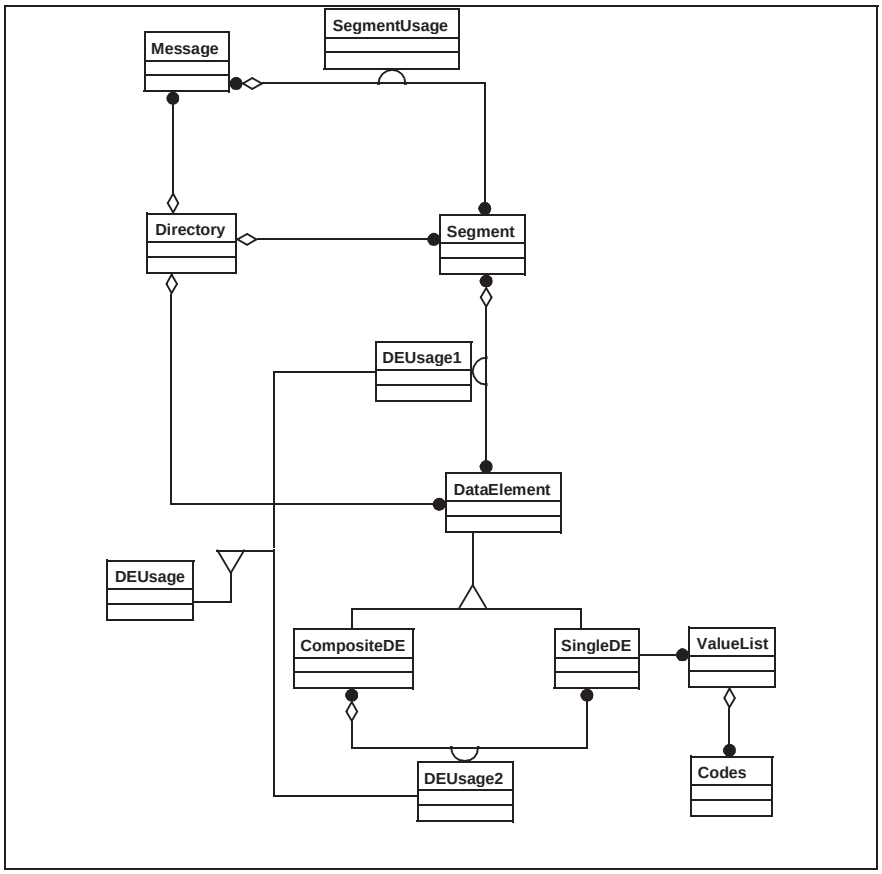
4. The METIS data model and database

One of the most crucial tasks in our project was the definition of the METIS data model, because it has to reflect the structure of the basic EDIFACT elements (messages, segments, data elements and code lists) and the openness to include new languages. In addition to that, the complexity and expressiveness of the data model considerably effects the functionality of the browser.

Based on the above mentioned criteria we decided to use an object-oriented data model for two main reasons: The hierarchical structure between the basic EDIFACT elements can be expressed more efficiently than in a relational model and it permits the extension of further languages in a flexible manner. Furthermore, the object-oriented model allows the definition of clear interfaces to the browser.

The simplified design - following the notion of the Object Modeling Technique (OMT) [5] - of the METIS data model is presented in Fig. 3:

The central class of the METIS object model is the class describing a directory. Each directory covers several messages, segments and data elements. A message consists of a set of segments and their interrelationship is further described in the class SegmentUsage. A segment is again built up by a number of composite and simple data elements. The concrete usage of a data element within a segment is described in the class DEUsage1. The class DEUsage2 defines how a simple data element is used in an composite data element. In the case of a coded simple data element one or more code (value) lists are assigned to the data element. The object model presented in Fig. 3 does not include the multilingual extension which is illustrated in Fig. 4. by the



Fir. 3

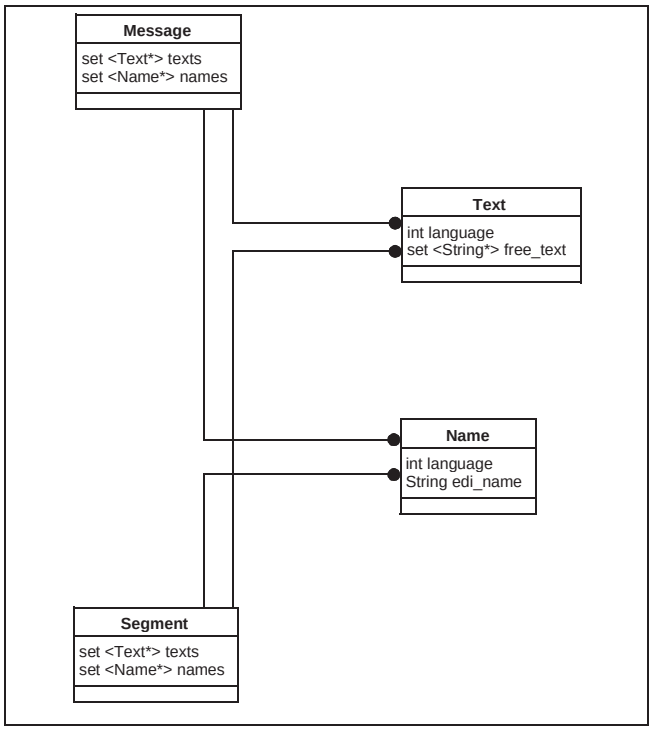


Fig. 4

means of the classes Message and Segment.

The multilingual attributes of a class - names and describing texts - are modeled as references to a set of a separate class Name or Text, respectively. Each member of the set corresponds to a description of an instance in another language.

The above described data model has been implemented in the object-oriented database system POET 2.1 [6]. POET is a precompiler for C++.

5. Implementation of the browser

Since METIS is designated to be used on PCs, we have decided to implement the browser on top of Microsoft Windows to ensure a high user acceptance by a user-friendly guidance. To take advantage of the POET interfaces and for reasons of the simple extensibility, the browser also follows the object-oriented paradigm and is implemented in C++.

The browser can be characterized as follows. First of all the user has to decide the language to be guided. The user can navigate only through one directory at one point of time. If he decides to change the directory, the old directory will be closed. Once having opened a directory and chosen a starting language in which the EDIFACT data should be displayed,

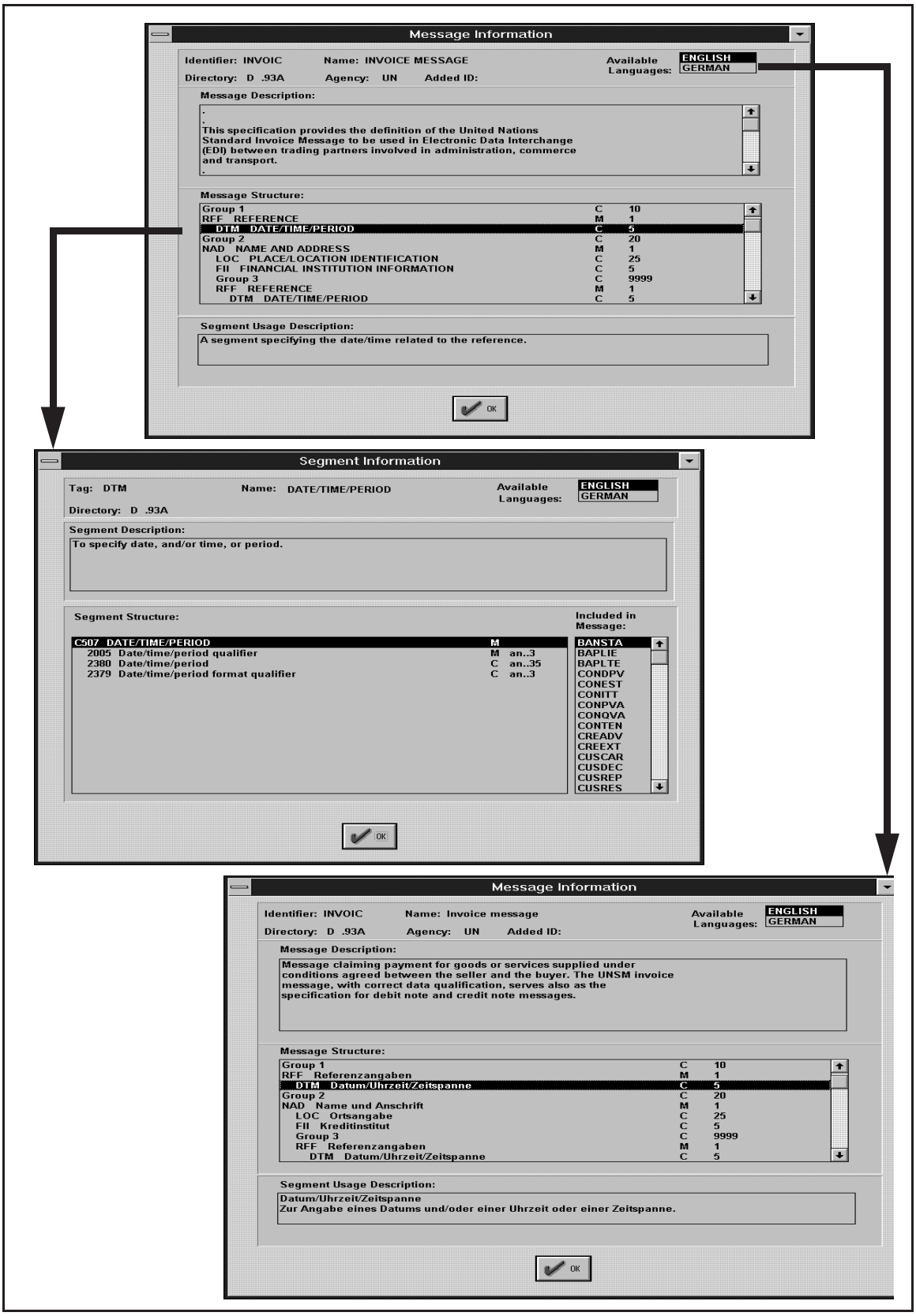

Message Information
Identifier: INVOIC
Name: INVOICE MESSAGE
Directory: D .93A
Agency: UN
Added ID:
Available Languages:
ENGLISH
GERMAN
Message Description:
This specification provides the definition of the United Nations Standard Invoice Message to be used in Electronic Data Interchange (EDI) between trading partners involved in administration, commerce and transport.
Message Structure:
Group 1 C 10
RFF REFERENCE M 1
DTM DATE/TIME/PERIOD C 5
Group 2 C 20
NAD NAME AND ADDRESS M 1
LOC PLACE/LOCATION IDENTIFICATION C 25
FII FINANCIAL INSTITUTION INFORMATION C 5
Group 3 C 9999
RFF REFERENCE M 1
DTM DATE/TIME/PERIOD C 5
Segment Usage Description:
A segment specifying the date/time related to the reference.
OK
Segment Information
Tag: DTM
Name: DATE/TIME/PERIOD
Directory: D .93A
Available Languages:
ENGLISH
GERMAN
Segment Description:
To specify date, and/or time, or period.
Segment Structure:
Included in Message:
C507 DATE/TIME/PERIOD M
2005 Date/time/period qualifier M an..3
2380 Date/time/period C an..35
2379 Date/time/period format qualifier C an..3
BANSTA
BAPLIE
BAPLTE
CONDPV
CONEST
CONITT
CONPVA
CONQVA
CONTEN
CREADV
CREEXT
CUSCAR
CUSDEC
CUSREP
CUSRES
OK
Message Information
Identifier: INVOIC
Name: Invoice message
Directory: D .93A
Agency: UN
Added ID:
Available Languages:
ENGLISH
GERMAN
Message Description:
Message claiming payment for goods or services supplied under conditions agreed between the seller and the buyer. The UNSM invoice message, with correct data qualification, serves also as the specification for debit note and credit note messages.
Message Structure:
Group 1 C 10
RFF Referenzangaben M 1
DTM Datum/Uhrzeit/Zeitspanne C 5
Group 2 C 20
NAD Name und Anschrift M 1
LOC Ortsangabe C 25
FII Kreditinstitut C 5
Group 3 C 9999
RFF Referenzangaben M 1
DTM Datum/Uhrzeit/Zeitspanne C 5
Segment Usage Description:
Datum/Uhrzeit/Zeitspanne
Zur Angabe eines Datums und/oder einer Uhrzeit oder einer Zeitspanne.
OK

Fig. 5

the user can open the message directory, the segment directory, the composite data element directory or the simple data element directory. Depending on the user's choice the content of the respective directory will be displayed to select a specific element for a closer inspection. The representation of a specific element covers the attributes describing the element, a description in free text, the EDIFACT elements of the next level which built up the specific element and vice versa those EDIFACT elements of a higher level which include the specific element. The two latter functions describing the structure of the EDIFACT hierarchy, are equipped with links. A single click on a subordinate element displays the usage of the subordinate element within the specific element and a double click on a superior/subordinate element opens the description of the respective element. In addition to that, a click on one of the available languages opens the description of the same EDIFACT element but in the selected language (See Fig. 5).

Owing to the browser component of METIS, the user is able to get any information concerning the UN/EDIFACT communication standard (See Fig. 1) in the most flexible manner. This means he or she must not read the whole standard to get an answer to a very specific question and he or she can receive the answer in his or her preferred language. METIS is a tool for the newcomer to UN/EDIFACT who wants to get an overview of the possibilities of the electronic data interchange in the commercial sector. Furthermore, METIS is especially designed for the EDI-application programmer who has to keep in touch with the latest version of the standard to adopt the application programs to the current interchange structure. Last but not least it is an aid to the EDI-coordinator who has to decide which documents are to be exchanged electronically.

6. Summary and future work

The current implementation of METIS permits a flexible browsing over the UN/EDIFACT communication standard, but it is limited concerning the applicability in practice due to the following reasons which are out of our range. Since until now one cannot yet receive DIRDEF messages on-line, one has to work with diskettes. The main part of the various national EDIFACT reference databases does not support the export of DIRDEF messages or - if so - provides only a rudimentary description of the complete standard directory. Nevertheless, we are convinced that in the near future the description of the UN/EDIFACT directories will be made available on-line via the DIRDEF message. At this moment METIS can save as a useful and general tool for EDI-users using EDIFACT.

Future work might concern the extension of METIS to a multilingual EDIFACT editing tool to build subsets of existing messages or to create completely new messages. This enables the adoption of a message design to the business needs of the user´s company. The self-created revisions could be translated into the DIRDEF format. Since in future EDIFACT converters might accept DIRDEF as input format, messages based on the self-created revisions could be processed with the company's converter. Furthermore, the specification of the new revision could be transmitted to the EDIFACT partner company, who also might include it in its converter. This means a consequent extension of the basic EDI-idea, because each EDI agreement between two companies on the interchange structure will be based on EDIFACT.

Acknowledgment

METIS is a joint project of AUSTRIA PRO, TERMNET and the University of Vienna. The support by FFF (Austrian Foundation for Research Applied to Industry) is gratefully acknowledged.

References

[1] Ken Steel
Matching Functionality of Interoperating Applications:
Another Approach to EDI Standardisation
ISO/IEC JTC1/ SC30 IT11/7/94-108, 1994

[2] J. Berge
The EDIFACT Standards
NCC Blackwell, 1991

[4] United Nations, Economic and Social Council
UN/EDIFACT Syntax Rules (ISO 9735)
International Organization for Standardization, Geneva 1988

[4] United Nations, Economic and Social Council
Trade Data Interchange Protocols
Directory Definition Message (TAG 52.315)
Geneva, 1993

[5] J. Rumbaugh, M. Blaha, W. Premerlani, F. Eddy, W. Lorensen
Object-Oriented Modelling and Design
Prentice-Hall 1991

[6] Programmer´s & Reference Guide, POET 2.1
POET Software Corporation, Santa Clara 1993 -94