GENERAL AGREEMENT ON

TARIFFS AND TRADE

Restricted
IDB/W/2
18 April 1988
Special Distribution

INTEGRATED DATA BASE

PROJECT STRUCTURE.

The development of the Integrated Data base (IDB), based on the project proposed hereafter, is organized into three parts. This paper only presents the Data Base Preparation (Part I); Information Services (Part II) and Other Data Base developments (Part III) will be presented in separate papers.

Part I. Data base preparation.

This part is focused on the work involved in setting out the detailed requirements for the creation of a data base at the tariff line level. The scope of PART I includes the following activities:

(1) System design: of the applications needed to prepare the information submitted by the participating countries for loading into the data base. In brief, for each domain of data received - tariffs, statistics and quantitative restrictions - a set of programmes must be designed to check, correct, convert and format the information to be loaded into the tariff line data base.

(2) Design of the tariff line data base: dictionary of all data elements (submitted or generated by the Secretariat), data volume estimation, logical and physical organization of the files, definition of access needs (key elements, file links, control), security requirements and users communication interface.

(3) Data base maintenance: a distinct major application will deal with the data base update matters, including the specifications of the basic query programmes needed for testing.

(4) System test: phase which will take place after all individual programmes are specified, coded and tested. This phase will consist of assembling those programmes by logical sub-systems (conversion, update, etc.) and testing the global functionning of all parts.

(5) Data base unload: this last phase in Part I contains the specification of the programmes required to unload data from the data base in order to satisfy the information exchange requirements.

Part II. Information services.

This part will contain a detailed analysis of the data base usage, focused on the information services required:

(1) On-line services: definition of the necessary programmes to produce screen and hard-copy outputs through menu-driven applications.

(2) Data transfer: these facilities will complement the feature mentioned in Part I (5) above. This service will consist of a set of programmes in which the tariff line information will be selected and manipulated according to selection criteria and processing rules.

(3) Batch services: definition of the applications needed to produce **large** volume printed outputs - analytical reports - based on tariff line information.

Part III. Other data base developments.

The third part of the IDB project will contain a description of possible developments of the data base. Some of these developments will depend on the types of analyses which will be necessary to carry out either on-line or in batch.

While the data base in its present form, i.e. at the tariff line level, enables users to retrieve tariff line information according to given criteria, it can be expected that analyses showing aggregated data will entail higher processing costs.

For that purpose, it might be necessary to envisage creating analytical files where useful aggregates (tariff averages, tariff profiles, etc.) could be compiled and stored. These files could for instance, be compiled at the 6-digit and or 4-digit level of the HS nomenclature. These files could also serve the purpose of keeping historical series which might prove difficult to keep at the tariff line level (because of changes in nomenclature).

Other developments might also concern additional data which would be requested in the future.

TABLE OF CONTENTS

INPUT TO THE DATA BASE.

The information to be stored in the Gatt Integrated Data Base comes from two main sources: countries participating in the exercise and Gatt divisions responsible for data collection and preparation. The input data flows and the corresponding processing phases are described in the following pages where the specific characteristics of each domain of information are considered. A data inventory, by domain, is also given in following pages. These inventories are limited to the information submitted by the participating countries.

INPUT DOMAIN: Trade Statistics.

Trade statistics are sent to Gatt, on a yearly basis, from central governmental institutions in charge of the collection of these data from regional offices and for their consolidation at national level e.g. Ministry of Commerce, Board of Customs and Statistical Offices. The period of reference can vary from country to country: the data can be collected on the basis of a fiscal or a calendar year.

This statistical information contains import values and quantities, by country of origin, according to the Harmonized System classification. The number of statistical items varies from 5000 to 15000 and the average number of origin countries per item can be estimated at 15. The format of the statistical file is given in Part A of the Secretariat "Technical Note" of 25 February 1988.

The following processing steps are performed:

Data reception The tape and any documentation received are registered: date, provenance, volume identification number and other relevant information are recorded. The tape contents are then copied onto a Gatt-owned volume. Any necessary conversion will take place at this stage e.g. introduction of standard record length, block size, label.

Data analysis The files are analysed in terms of data elements' formats and contents. A quantitative report giving the number of records, the number of items, the maximum number of origin countries per item is printed. Another report gives all existing values of all data elements such as the HS chapters coverage, the national partner codes, the national quantity units used etc., and the total import value in national currency. Based on the above "inventories", the corresponding Gatt Reference files are adjusted to take into account new elements or new reference codes values recorded in the national submissions.

Microfiching In order to have a complete and detailed picture of the submission, all records are formatted and "printed" on COM[1]. These fiches are necessary when a precise reference to the original submission is needed.

Data editing This step checks all data elements recorded in the submissions against the Reference files. In addition, erroneous data (or data for which Gatt did not create any standard) can be mass-replaced with correct values wherever possible. The statistical item numbers used in this file are checked against the nomenclature used in the IDB duty rates file. If applicable, programmes will be used to modify the import statistics nomenclature to reflect changes occurring during the year of reference. These programmes will allow for statistical number changes, merge or split of statistical items. If the tariff treatment code is recorded on the submitted file, a verification of this code and the relevant tariff information of the IDB duty rates file will be performed. If the tariff treatment code is not recorded,

1 COM Computer Output Microfiche; one fiche contains 270 pages; reduction rate 1:48.

a set of programmes will be run to apply the correct treatment code on a product/partner basis. On request, a listing in which all errors are identified is printed. A copy of the report is sent to the country which has provided the data. These editing steps are performed as many times as necessary to solve all the problems.

Data conversion The data are converted from national formats and codes to Gatt standards for country codes and quantity unit codes. Imports in national currency are converted to US dollar values using a yearly average exchange rate from the IMF (or received from the reporting country). For every statistical item, the following records are generated: a total for all countries of origin, a total for GATT members, a total of Customs Unions (C.U.) member countries (the definition of C.U. being adapted to the period of reference), and a series of trade aggregate records consisting of an accumulation of individual partner's data according to their tariff treatments. The layout of the file resulting from this process is identical to the format of one of the statistical output files and corresponds to a standard format which can be used for data exchange between the participants.

Pre-load verification

Several reports summarizing the trade information are printed; they are necessary for taking the decision of loading the data in IDB:

- total trade, broken down into HS chapters,
- total trade by country of origin,
- binding status report.

These reports are produced in national currency and US dollars and, when applicable, for the agriculture and industry sectors.

Database loading

The loading procedure stores the information in the data base and performs the functions required by the Adabas[2] software which manages the IDB (data compression, creation of inverted list files necessary for direct accesses to the data etc..). Details of the input statistical file elements are given in the following pages.

Processing follow-up

A database log file will be updated with date and time of processing and the processing phase name. Periodically, or on request, a progress report on data received and processing phases will be printed and routed to participating countries.

2 Adabas: A Data Base Management System from Software AG company; this sofware will be used for the IDB system.

Data Inventory for Trade Statistics

The trade file contains trade statistics by country of origin for each tariff item (or statistical item). At present, only import statistics are recorded but the format of the file is designed to allow, if necessary, the submission of export statistics.

Import statistics are recorded in value and volume and codes are attached to each country of origin to identify the tariff treatment applicable to its supplies. Values are recorded in the currency used by the national authorities in the collection of import statistics for customs purposes. Volumes are reported in one or, where applicable, in two units of quantity.

Entities	Definitions

Trade File key

The data elements of the trade file key identify the reference year, the country submitting the trade statistics, the direction of trade, the tariff item number (or the statistical item number), the trading partner country and the type of tariff treatment.

Direction of trade

Import statistics are identified by the code "1" recorded in this one-digit element. If necessary, export statistics will be identified by the code "2" and statistics on re-exports will be identified by the code "3". At present, only the code "1" is used since only import statistics are requested in the submissions for the IDB.

Tariff item number

The tariff item number is a field of 21 characters identifying the tariff item, if statistics are submitted in the tariff nomenclature. Statistics can also be submitted in the statistical detail and in such cases, the tariff item number contains the statistical item number.

The tariff item number recorded in the trade file is the same as the tariff item number recorded in the tariff and non-tariff files. If the statistical nomenclature is used in the trade file, the same statistical nomenclature should be used in the tariff and non-tariff files.

The reporter code, the reference year and the tariff item number are the only key elements which allow the integration of all information recorded in the various IDB files.

Trading partner country

This element is a field of 5 characters identifying the trading partner country code. The "partner code" recorded by national authorities can be alphabetic or numeric. It will be converted to the GATT standard 3-digit ISO code, before the file is loaded into the data base.

Type of tariff treatment

This element is a 1-character code identifying the type of tariff treatment to which imports from the trading partner country are entitled for the tariff item concerned. If, for a given item, the partner benefits from m.f.n. treatment, the code takes the value "0". If, for another item, the country benefits from GSP treatment, the code takes the value "4", etc.

This code is necessary to identify the product coverage of preferential arrangements or of GSP schemes. The tariff treatment code also serves as a link between each country of origin and the corresponding tariff duty rate applicable, as recorded in the "coded" tariff file. In effect, the first digit of the code of the duty rate type applicable is equal to the tariff treatment code.

Import Partner data.

The import partner data contain information in value and volume by partner. The import data are recorded in the following fields:

Type of tariff relation

The type of tariff relation contains a 1-digit code attached to each country of origin to identify whether or not the country is signatory to a preferential arrangement or, is listed in the GSP scheme as a GSP beneficiary, etc. For a given trading partner, the relation code is the same throughout the file, irrespective of the tariff item. This code allows, for example, the identification of total imports from GSP beneficiary countries, from m.f.n. origins, etc.

To identify GSP beneficiaries, three different relation codes are proposed. Relation code "4" is attached to GSP countries which, in the absence of GSP preferences are entitled to m.f.n tariffs. Relation code "5" is attached to GSP countries which, in the absence of GSP preferences, could either be entitled to preferential tariffs under special arrangements or to m.f.n. tariffs. Finally, relation code "6" is attached to GSP countries which, in the absence of GSP preferences, are entitled to general tariffs (i.e. tariffs higher than m.f.n. tariffs).

The following table shows the possible values which the treatment code can take according to each type of relation:

TREATMENT :		m.f.n 0	FTA 1	Zone 2	Other pref. 3	GSP only 4	GSP & pref. 5	General 8	Unspec. 9
RELATION :									
m.f.n.	0	Y	N	N	N	N	N	N	N
FTA	1	Y	Y	Y	Y	N	N	N	N
Zone	2	Y	N	Y	Y	N	N	N	N
Other pref.	3	Y	N	N	Y	N	N	N	N
GSP/m.f.n.	4	Y	N	N	N	Y	N	N	N
GSP/pref.	5	Y	Y	Y	Y	Y	Y	N	N
GSP/general	6	N	N	N	N	Y	N	Y	N
General	8	N	N	N	N	N	N	Y	N
Unspecified	9	N	N	N	N	N	N	N	Y

C.i.f. value of imports in national currency

This element contains the c.i.f.[3] value of imports in national currency or in US dollars, if the US dollar is used for recording customs statistics. If the c.i.f value of imports is not available, this field should be left blank and the element "customs value" should be used. The imports value can be submitted in units of currency, in thousands of units etc. The unit used should be selected so that total imports can be recorded in a 13-digit element.

3 C.i.f. for Cost, Insurance and Freight.

Customs value of imports in national currency

This element contains the customs value of imports in national currency or in US dollars, if the US dollar is used for recording customs statistics. If the customs value is the c.i.f. value, this field should be left blank. The imports value can be submitted in units of currency, in thousands of units, etc. The unit used should be selected so that total imports can be recorded in a 13-digit field.

Primary unit of quantity

This element contains the primary quantity unit used for recording import volumes. If the duty rate(s) attached to the item is (are) specific, mixed or compound, the unit of quantity to be recorded in this element should, if possible, correspond to the unit used to collect customs duties. If, for a given item, there is no quantity data recorded, this element should be left blank.

Import volume in the primary quantity unit

This field contains the import volume expressed in the unit recorded in the primary unit code. If, for a given item, there is no quantity data available, this field should be left blank.

Supplementary unit of quantitiy

This element contains, where applicable, a supplementary unit of quantity in which import volumes are recorded. Where the duty rate(s) attached to the item is (are) expressed in a form using two units of quantity (e.g. 1 S per KG plus 2 S each) the supplementary unit recorded should, as far as possible, correspond to the second unit used for the collection of customs duties. If there is no supplementary quantity available, this element should be left blank.

Import volume in the supplementary unit

This element contains the import volume expressed in the unit recorded in the supplementary quantity unit code. If, for a given item, there is no supplementary quantity data available, this element should be left blank.

Customs duty collection value

This element contains the value of the customs duty collected, expressed in national currency or in US dollars, if the US dollar is used for the purpose of customs collection. If this information is not available, this element should be left blank.

INPUT DOMAIN: Tariff information.

In the formats proposed by the Secretariat in its 'Technical note' dated 24th February 1988, Part B (i) and (ii), the tariff information is organized in two files: one file contains "coded" data on tariffs and another one contains "textual" data. This 2-file submission has been retained for the following main reasons:

(1) The nature of information is different i.e. the "coded" data set will be used in calculation programmes for analytical reports (averages, frequency distributions etc..) and the "textual" data set will be mainly used to improve the presentation of tariff information on tariff line reports.

(2) For the preparatory work required in capitals, it seems appropriate to structure the tariff information in two groups, considering that a large amount of work to prepare the "textual" information was already done in the capitals, in the context of the Harmonized System common data base.

(3) The 2-file submission will probably reduce the amount of information to be submitted to the Secretariat and consequently simplify the data maintenance procedures, since the "textual" part contains descriptive and historical data which are very stable. In contrast, the "coded" part is subject to more frequent updates such as modifications of ad valorem equivalents (AVEs) of specific rates.

(4) The "textual" file format is designed to allow the preparation of data on a "element number" basis, i.e. the work of data collection and programming can be easily organized in capitals.

Basically, the required processing steps are the same as the ones described for Import Statistics:

Data reception The tapes and any documentation received for the "coded" and "textual" parts of the tariff information are registered: date, provenance, volume identification number and other relevant information are recorded. When applicable, any partial submission of "textual" tariff information will be merged with data already received. The tape contents are then copied onto a Gatt-owned volume. Any necessary conversion will take place at this stage e.g. introduction of standard record length, block size, label.

Data analysis The two files are analysed in terms of data elements' formats and contents. A quantitative report giving the number of records, the number of tariff lines, the number of tariff items and other information needed in subsequent processing steps is printed. Another report gives all existing values of all data elements such as the HS chapters coverage, the duty type codes (MFN, GSP, Other pref. etc..), the national partner codes used for recording bilateral AVEs, the range of validity periods etc... Based on the above "inventories", the corresponding Gatt Reference files are adjusted to take into account new elements or new reference codes recorded in the national submissions.

Microfiching In order to have a complete and detailed picture of the submission, for both tariff files, all records are formatted and "printed" on COM (computer output microfiches). These fiches are necessary when a precise reference to the original submission is needed.

Data editing This step checks all data elements recorded in the submissions against the IDB Reference files. In addition, erroneous data (or data for which Gatt did not create any standard) can be mass-replaced with correct values wherever possible. A matching programme checks nomenclatures used in the "textual" tariff information file against the "coded" file. On request, a listing in which all errors are identified is printed. A copy of the report is sent to the country which has provided the data. These editing steps are performed as many times as necessary to solve all the problems.

Data conversion The data are converted from national formats and codes to Gatt standards for partner codes, quantity unit codes, INRs[4] and Gatt Articles. The record layouts of the files resulting from this process are identical to the format of the tariff output files and correspond to a standard format which can be used for data exchange between the participants.

Pre-load verification

Several reports are printed before taking the decision to load the information into the IDB:

Data analysis

A summarized report showing a quantitative file structure (no. of items, no. of records, etc..) and a list of all code values is printed. It is similar to the one produced after the data reception step, except that at this stage, the information is expressed in standard formats and codes. Such a report allows year-to-year country comparisons and country-to-country comparisons.

Nomenclature comparison

This step compares the nomenclature used in the data submission file with that of the previous year stored in the IDB. The corresponding report shows all differences in the two nomenclatures.

Microfiching

Production of microfiches (or a printed report) containing, in detail, all "coded" and "textual" information to be loaded into the IDB.

Database loading

The loading procedure stores the information in the database and performs the functions required by the Adabas software which manages the IDB. Details of the input tariff files' elements are given in the following pages.

Processing follow-up

A database log file will be updated with date and time of processing and the processing phase name. Periodically, or on request, a progress report on data received and processing phases will be printed and routed to participating countries.

---

[4] INR : Initial Negotiating Right.

Data Inventory for Tariff Information

1. CODED PART OF THE TARIFF FILE.

The "coded" part of the tariff file contains customs tariff duty rates for each tariff line. Duty rates are broken down according to the type of tariff treatment (m.f.n., free-trade-area, GSP etc.) and, where applicable, by individual partner if a particular duty rate applies to a given country or to a group of countries.

Duty rates are recorded in percentage, in a normalized form, for specific, mixed or compound duty rates, or in text form. Various codes are attached to the duty rate to identify the nature of the duty (ad valorem, specific, etc.), the GATT binding status of the duty, etc.

Entities	Definitions

Coded Tariff File Key

The coded tariff file key identifies the reference year, the country submitting the data, the tariff line number, the type of measure (tariff measure in this file), the country to which the rate applies, the type of duty rate (m.f.n., GSP, etc.), the period during which the duty is in force and the record number.

Reporter of the tariff file

This element contains the code identifying the country submitting the information. The code will be converted to the 3-digit ISO country code used in the IDB.

Tariff line number

The tariff line number is an element of 25 characters used to record the tariff item number (first 21 characters), the non-tariff item suffix (positions 22 and 23) and the tariff item suffix (positions 24 and 25).

Tariff item number

The tariff item number is the same as the tariff item number used in the import statistics file. If imports were submitted in the detail of statistical items, the tariff item number recorded in the tariff file is the statistical item number.

Tariff item suffix

The tariff item suffix is recorded in positions 24 and 25 of the tariff line number. The suffix is used in cases where the customs tariff is more detailed than the statistical nomenclature, i.e. statistics are not available for some tariff sub-headings.

The following example illustrates the manner in which the tariff item suffix could be used:

IMPORT STATISTICS tariff item number		CUSTOMS TARIFF tariff item number	suffix
07134000_____Lentils		07134000_____	00
		average rate 10 %	
	In packages over 5 KG	07134000_____	01
		rate FREE	
	Other	07134000_____	02
		rate 20 %	

Validity period of the duty rate

This element is used to record the starting date and the end date of the period during which the duty is in force.

Start date of the validity period

This element of 6 characters contains the date of entry into force of the rate. This element would normally be left blank, unless the rate entered into force in the course of the reference year. The date recorded shows the last two digits of the year (positions 1 and 2), the month (positions 3 and 4) and the day (last two digits).

End date of the validity period

This element of 6 characters contains the date on which the rate ceases to be in force. The end date is normally left blank, unless the rate ceases to be in force in the course of the reference year. The end date is recorded in the same manner as the start date.

Type of measure

The type of measure is a 2-digit code used to distinguish a tariff measure from a non-tariff measure. The code "10" identifies a tariff measure and should be attached to all elements recorded in the coded tariff file.

Partner entitled to the duty rate

This element contains the code of the country or group of countries to which the duty is applicable. If this code is left blank, the duty rate applies to all countries recorded in the import statistics file, for the corresponding tariff item, which have a tariff treatment code equal to the first digit of the duty rate type. For example, if, for duty type "41", no partner code is recorded, duty "41" applies to all countries recorded in the import file, for the corresponding tariff item, having a tariff treatment code equal to "4" (GSP treatment).

Type of duty rate

The type of duty rate is a 2-digit code identifying the scope of each duty rate recorded. The first digit of the code is the tariff treatment code as recorded in the import statistics file. The second digit of the code is to be defined by national authorities according to their own needs. Thus, m.f.n. duties are numbered 01 through 09, FTA duties are numbered 11 to 19, etc.

With regard to m.f.n. duties, it is proposed that duty type "01" be used for recording the GATT consolidated duty rate; that duty type "02" be used for recording the legal or statutory rate; and that duty type "03" be used for recording the effectively applied rate under temporary legislation.

If the m.f.n. duty is unbound, there would be no duty type "01" recorded. If the m.f.n. rate is bound at the level of the statutory rate, duty type "01" would be recorded but not the duty type "02" (which is the same as the rate recorded under duty type "01"). Effectively applied rate "03" would only be recorded if different from rate "01" or "02".

Record number

There are three types of record numbers:

- Record 10: duty rate and relevant codes
- Record 20: normalized coding of specific duties
- Records 30-39: textual information on the duty

RECORD 10: duty rate and relevant codes

The record number 10 contains the rate in percentage, with three decimal places, and various codes to indicate: the nature of the duty (ad valorem, specific...), whether or not the rate was estimated, whether or not the duty is bound, and, for preferential duties, whether or not there are limitations. The information is recorded in the following elements:

Nature of the duty rate

This one-character code indicates that the duty is: ad valorem (blank), specific ("S"), compound ("C"), mixed ("M"), variable ("V") or unclassified ("O").

DUTY EXAMPLES OF THE VARIOUS NATURES:

```
- Ad valorem duty: 10 % of the value of the goods
- Specific duty :  1.00 $ per KG
                   1.00 $ per KG + 2.00 $ each
- Compound duty :  1.00 $ per KG + 10 % ad valorem
- Mixed duty :     1.00 $ per KG or 10 % ad valorem,
                   whichever is lower
- Variable duty : duty comprising a variable element
```

Duty rate in percentage

This six-digit element contains the percentage rate of the duty with three decimal places. Zero rates are recorded as "000000" and percentage rates which cannot be calculated are recorded as "999999".

Estimation of the ad valorem equivalent

This code indicates if the rate is: estimated by the national authorities (code "E"), estimated by the GATT Secretariat (code "C"). Code "A" indicates that the rate is an average of sub-items' rates. If the rate is not estimated, the code is left blank.

Duty rate binding status

This code indicates whether the duty is bound at the prevailing rate (code blank), at a ceiling rate (code "C") or at different levels (code "D"). Code "U" indicates that the duty is unbound.

Binding coverage

If the tariff line is partially bound this code contains the letter "X". Otherwise the code is left blank.

Preferential rate limitation status

In some instances the preference under GSP or other preferential arrangements is granted within value or quantitative limitations. If the limitation is a ceiling or equivalent limitation, this code contains the letter "C". If the limitation is a quota or equivalent limitation, this code contains the letter "Q". If there is no limitation, the code is left blank.

Preferential rate coverage

If the tariff line is partially covered by the preference, this code contains the letter "X". Otherwise, the code is left blank.

RECORD 20: Normalized Coding of specific and other duties

Specific, compound and mixed rates are recorded as percentage ad valorem equivalents (AVE) in record "10". In this record, these duties are recorded in a manner which allows their identification in computer programmes for the calculation of bilateral AVE's or of AVE's by group of countries. The normalized code also allows for the printing of these duties in analytical reports, in a standard format.

The normalized code comprises three parts. Each part of the code offers the possibility of recording one ad valorem rate and one specific duty. The two types of duties are linked by a plus or minus sign (for recording compound duties) or by an operand "OR" (for recording mixed duties).

Detailed examples illustrating the system for recording the various duties in the normalized code are shown in the Technical Note by the Secretariat in which formats for data submissions were circulated on 24 February 1988 (see Part B(i) page 6).

The data elements of each of the three parts of the normalized code are recorded in the following elements:

Normalized code ad valorem rate

This 6-digit element contains the ad valorem rate part of a compound duty or of a mixed duty. The rate is recorded with three decimal places (as in record "10"). For specific duties, which do not comprise an ad valorem rate part, this element is left blank.

Normalized code first operand

This one-character element contains the letter "R" for "OR", the letter "P" for "PLUS" or the letter "M" for "MINUS". This operand links the ad valorem rate part of the compound or mixed duty, to the specific rate part of the duty. If the ad valorem rate part of the duty is blank, this operand should be blank; otherwise, it should be filled in with one of the letters mentioned above.

Normalized code amount of currency units

This thirteen-digit element contains the amount of currency units to be collected per unit of quantity imported. The amount to be collected is recorded as an eight-digit number with five decimal places. The currency unit used in the normalized code should be the same throughout the file. The currency should be the same as the currency used for recording import statistics but the unit could be different (e.g. the unit could be cents in the normalized code and thousands of dollars in the statistics).

Normalized code unit of quantity

This five-character element contains the code of the unit of quantity used in the duty for collecting duties. The quantity unit code used in the normalized code should, as far as possible, be the same as the code used in import statistics. This code is converted to the GATT standard quantity unit code.

Normalized code second operand

This one-character element contains the letter "R" for "OR", "P" for "PLUS", "M" for "MINUS", "H" for "WHICHEVER IS HIGHER" or "L" for "WHICHEVER IS LOWER".

If this operand contains the letter "O", "P" or "M", the second part of the normalized code should contain the other part of the duty, which is implied by the operand.

The second and third parts of the normalized code have the same format as the first part, but the logic to determine the possible values of the operands in the last two parts of the code is of course different from that of the first part (see examples in the formats, Part B(i), page 5).

RECORDS 30-39: Textual information on the duty rate

These records contain any textual information on a duty rate which, for example, could not be coded in the format of the normalized code. Each record contains 70 characters and up to ten lines of text can be coded in record numbers 30-39.

Data Inventory for Tariff Information (continued)

2. TEXTUAL PART OF THE TARIFF FILE.

The "textual" tariff file contains product descriptions for each tariff line recorded in the coded tariff file and, for GATT bound tariff lines, all information submitted in the loose-leaf schedule of concessions. Product descriptions are recorded in two different sets of records. One set contains the full legal text of the product description, another set contains an abbreviated text of the product description.

Some of the records of the textual tariff file are of variable length. For that reason, all records are preceded by a five-digit element containing the total number of characters in the record, including the "record length" field itself. This element would not be required if the computer used to create the file is IBM or IBM compatible.

Entities	Definitions

Textual Tariff File Key

The key of the textual tariff file contains the following elements: the tariff line number, the element number, and the sub-element number.

Tariff line number

The tariff line number is an element of 25 characters used to record the tariff item number (first 21 characters), the non-tariff item suffix (positions 22 and 23) and the tariff item suffix (positions 24 and 25).

In this file the non-tariff item suffix is always blank and the tariff item suffix is the same as the tariff item suffix recorded in the coded tariff file.

Tariff item number

The tariff item number is the same as the tariff item number used in the import statistics file and in the coded tariff file and in the non-tariff files.

Tariff item suffix

The tariff item suffix is recorded in positions 24 and 25 of the tariff line number. The suffix is used in cases where the customs tariff is more detailed than the statistical nomenclature, i.e. statistics are not available for some tariff sub-headings. The tariff item suffix recorded in this file is the same as the tariff item suffix recorded in the coded tariff file (see example of the use of the suffix, shown in the coded tariff file documentation).

Element number

The element number is a two-digit number identifying the different types of information recorded in the textual tariff file. Elements "02" through "07" correspond to the column numbers of the loose-leaf schedule and contain the same information as reported in the respective columns of the loose-leaf schedule.

Sub-element number

The sub-element number is a three-digit number identifying, in some elements, the record number. The sub-element number is necessary in elements "02", "07", "09" and "11". In other elements the sub-element number is left blank.

Text structuring record (ELEMENT 01)

The text structuring record was used in the past for recording tariff heading descriptions which were to be retrieved in the correct sequence (see document TAR/W/47, pages 20 and 21, for a detailed explanation of the structuring system). This record was used in cases where the tariff nomenclature was not sufficiently structured.

Since the HS nomenclature is structured at the two, four, five and six-digit level, it is possible that this structuring record will not be needed in the IDB.

Full Legal Description (ELEMENT 02)

The full legal description is recorded in element "02", sub-elements 001 through 999. The sub-element number gives the sequence of the lines of text recorded in this element. Each line of text contains 70 characters.

A "level" number can be assigned to each description to distinguish descriptions of headings (not associated to a duty rate) from descriptions of tariff lines (associated to a duty rate). If the file contains descriptions of headings, the "level" of the heading description should be equal to the number of digits used the tariff number ("02" for a chapter heading, "04" for 4-digit heading, etc.). Descriptions of tariff lines should all have a "level" equal to "00". If no headings are recorded, the level should be left blank.

Each text line of the description can be indented by indicating in the two-digit "indent" element the number of blank characters to be left between the left margin and the text.

Information on Present Concession (ELEMENT 03)

Element "03" contains textual information on the present bound rate, as reported in the column "3" of the loose-leaf schedule. This textual information would be used if the loose-leaf schedule was to be printed from the textual tariff file. If this element was not recorded, the duty rate recorded in the coded tariff file would be used.

For element "03", the sub-element number should be blank. The textual information is recorded in a variable length field of which the first four digits are used to indicate the length of the field.

Legal Instrument(s) Containing the Present Concession (ELEMENT 04).

Element "04" contains the abbreviation(s) of GATT instrument(s) as reported in the column "4" of the loose-leaf schedule. The list of abbreviations to be used for coding the GATT instruments is shown in document TAR/W/6.

Each instrument is recorded using 10 characters. Up to 99 different instruments can be coded. The number of instruments recorded is indicated in the 2-digit field "number of occurrences". The sub-element number of this element should be left blank.

Initial Negotiating Right (INR) on Present Concession (ELEMENT 05).

Element "05" contains the abbreviated name(s) of the country(ies) holding an INR on the present concession, as reported in column "5" of the loose-leaf schedule. The list of abbreviations to be used for coding the country names is shown in documents TAR/W/5 and Add.1.

Each country is recorded using 5 characters. Up to 99 different countries can be coded. The number of countries recorded is indicated in the 2-digit field "number of occurrences". The sub-element number of this element should be left blank.

Legal Instrument(s) Containing the First Concession (ELEMENT 06)

Element "06" contains the abbreviation(s) of GATT instrument(s) through which a concession was first established, as reported in the column "6" of the loose-leaf schedule. The list of abbreviations for GATT instruments is shown in document TAR/W/6.

Each instrument is recorded using 10 characters. Up to 99 different instruments can be coded. The number of instruments recorded is indicated in the 2-digit field "number of occurrences". The sub-element number of this element should be left blank.

Initial Negotiating Right(s) on Previous Concessions (ELEMENT 07)

Element "07" contains the abbreviated name(s) of GATT instruments and country(ies) holding an INR on previous concession(s), as reported in column "7" of the loose-leaf schedule. Previous concessions are recorded in text form at the end of each record in a variable length field of which the first four digits are used to indicate the length of the field. Each record is identified by the sub-element number as follows:

Sub-element	GATT instrument	No. of occur.	INR	X	...	Field length	Rate ---->
001	A/49	03	CA	X			
			GB				
			JP		...	0005	10.0%
002	blank	01	BR			0004	5.0%
...							
999							

In the above example, the first record (sub-element "001") shows an earlier concession of 10% on which three countries are holding an INR. The concession was established through the GATT instrument "A/49". Country "CA" holds an INR on part of the item. The second record (sub-element "002") shows an earlier concession of 5% on which country "BR" holds an INR. The instrument is not coded in this second record. This means that the concession was established through the "first instrument" coded in element "06". Physically, the above two sub-elements would be recorded as follows:

```
TARIFF LINE NUMBER    07001A/49      03CA   XGB    JP     000510.0%
TARIFF LINE NUMBER    07002          01BR    00045.0%
```

Reference to INR(s) on Previous Concessions (ELEMENT 08)

Element "08" contains a textual reference to the existence of INR(s) on previous concessions. This element was used in the past, in cases where INR(s) could not be precisely identified. The reference indicated that INR(s) on previous concessions existed. The reference is recorded in a variable length field. The sub-element number of this element should be left blank.

Abbreviated Product Description (ELEMENT 09)

Element "09" contains a maximum of 3 sub-elements, corresponding to 3 lines of text of 70 characters, for recording the abbreviated product description. The layout of this element is the same as the layout of the full product description (see element 02 description), however, the two fields "level" and "number of blanks for indentation" should be left blank in this element.

Footnote Referring to any Element of the File (ELEMENT 11)

This element "11" contains footnotes which can refer to any element of the textual tariff file. As shown in the format of the file, an element (01-09) can be sub-divided into: sub-elements, fields and occurrence number. The footnote can refer to an element, a sub-element, a field number and an occurrence number as follows:

Footnote key		Data element referred to:					
Element	Sub-element	Sub-element	Field no.	Occur. no.	Serial no.	Footnote length	text -->
11	002					005	Note1
11	007	002	2	02		005	Note2
11	006			01	0001		

In the example above, the first line shows the coding of a footnote referring to element "02", the legal description. The second line shows a footnote referring to the second field of element "07": INR's. Within this field, the footnote refers to the second occurrence. The third line shows a footnote referring to the first occurrence of the GATT instruments recorded in element "06"; in this case however, the footnote is recorded in a separate file under the serial number "0001". Physically, the above three footnotes would be recorded as follows:

```
TARIFF LINE NUMBER     11002            005Note1
TARIFF LINE NUMBER     11007C02202      005Note2
TARIFF LINE NUMBER     11006     010001
```

INPUT DOMAIN: Quantitative Restrictions.

At present, the notifications on Quantitative Restriction measures are received in Gatt Secretariat, on paper, via the country delegations. In the document "Technical Note by the Secretariat" on 25 February 1988, two formats have been proposed to automate the data collection process in capitals. Part C (i) and Part C (ii) contain record formats for the "coded" and "textual" parts of the QRs file. The "textual" part information submission is only necessary for measures applicable to part of tariff lines. The processing steps required from data reception to load into the IDB can be summarized as follows:

Data reception The tapes and any documentation received for the "coded" and "textual" parts of the Quantitative Restrictions data are registered: date, provenance, volume identification number and other relevant information are recorded. When applicable, tape containing additional information for updating will be processed and merged with data already received. The tape contents are then copied onto a Gatt-owned volume. Any necessary conversion will take place at this stage e.g. labelling, introduction of standard record length, block size.

Data analysis The two files are analysed in terms of data elements' formats and contents. A quantitative report giving the number of records, the number of tariff lines, the number of tariff items and other information needed in subsequent processing steps is printed. Another report gives all existing values of all data elements such as the HS chapters coverage, the measure codes, the national country codes used to define the affected countries, the range of validity periods etc... Based on the above "inventories", the corresponding Gatt Reference files are adjusted to take into account new reference codes recorded in the national submissions.

Microfiching In order to have a complete and detailed picture of the submission, for both QR files, all records are formatted and "printed" on COM (computer output microfiches).

Data editing This step checks all data elements recorded against the Reference files. In addition, erroneous data (or data for which GATT did not create any standard) can be mass-replaced with correct values wherever possible. Two matching programmes check nomenclatures used in the "coded" QR file against the tariff file and the "textual" file is matched with the "coded" QR file: discrepancies are printed. On request, a listing in which all errors are identified is printed. A copy of the report is sent to the country which has provided the data. Thess editing steps are performed as many times as necessary to solve all the problems.

Data conversion The data are converted from national formats and codes to Gatt standards for measure codes, country codes and Gatt article numbers. The record layouts of the files resulting from this process are identical to one of the possible standard formats of the IDB output files.

Pre-load verification
Several data analysis reports are printed before loading the information into the IDB. They summarize the files structure according to various selection criteria.

Database loading
The loading procedure stores the information in the database. Details of the input Quantitative Restrictions files' elements are given in the following pages.

Processing follow-up
A database log file will be updated with date and time of processing and the processing phase name. Periodically, or on request, a progress report on data received and processing phases will be printed and routed to participating countries.

Data Inventory for Quantitative Restrictions.

1. CODED PART OF THE NON-TARIFF FILE.

The coded non-tariff file contains, at this stage, information on quantitative restrictions. The nomenclature used in this file is the same as the nomenclature used in the import statistics file and in the tariff files.

For each quantitative restriction, the file contains the GATT body to which the restriction was notified, the symbol(s) of the GATT document in which the restriction was notified, the GATT article(s) to which the restriction refers and textual information concerning the restriction.

Entities	Definitions

Coded Non-Tariff File Key

The key of the coded non-tariff file contains the following elements: the year of reference, The code of the country or Customs Union (C.U.) maintaining the restriction, the code of the member country of a C.U. maintaining the restriction, the tariff line number, the type of restriction, the restriction code, the code of the country or group of countries affected by the restriction, the period of validity of the restriction and the record number.

Country or C.U. maintaining the restriction

This element contains the code identifying the country or the C.U. maintaining the restriction. The code will be converted to the 3-digit ISO country code used in the IDB.

C.U. member country maintaining the restriction

This element contains, if applicable, the code identifying a C.U. member country maintaining the restriction. The code will be converted to the 3-digit ISO country code used in the IDB.

Tariff line number

The tariff line number is an element of 25 characters used to record the tariff item number (first 21 characters), the non-tariff suffix (positions 22 and 23 of the tariff line number) and the tariff item suffix (positions 24 and 25 of the tariff line number). In this file, the tariff item suffix should be left blank.

Tariff item number

The tariff item number is the same as the tariff item number used in the import statistics file and in the tariff files. If imports were submitted in the detail of statistical items, the tariff item number recorded in this file is the statistical item number.

Non-tariff item suffix

The non-tariff item suffix is recorded in positions 22 and 23 of the tariff line number. The suffix is used in cases where the restriction covers part of a tariff item, i.e. statistics are not available for some non-tariff subheadings.

Period of validity of the restriction

This element is used to record the start date and the end date of the period during which the restriction is in force. The start date and end date are coded using six digits, to indicate the last two digits of the year, the month (01-12), and the day (01-31). If both dates are left blank or if the end date is left blank, the restriction is considered as being in force.

Type of restriction

At this stage, the non-tariff file will contain only quantitative restrictions. In the draft formats circulated on 24 February 1988, it was proposed that QR's on imports be identified by the type "11" and that QR's on exports be identified by the type "61".

Country or country group affected

This element contains the code identifying the country or the group of countries affected by the restriction. This code will be converted to the 3-digit ISO country code which is used in the IDB.

Restriction code

In the draft formats, it was proposed that QR's be grouped by types. The restriction codes proposed used a 3-digit element where the first digit identified the type of QR: "1" for licenses, "2" for quotas, "3" for prohibitions, etc.

Record number

The record number identifies four types of information attached to the restriction.

- Record "10" contains various codes attached to the restriction.

- Records 20 through 29 contain symbols of GATT documents in which the restriction was notified.

- Records 30 through 39 contain GATT article numbers to which the restriction refers.

- Records 40 through 49 contain textual information concerning the restriction.

RECORD 10: Codes attached to the QR .

Record "10" contains the following codes: a 2-character code identifying the GATT body to which the restriction was notified and a one-character code to identify a restriction applicable to part of a tariff line ("X" to indicate a partial coverage, blank otherwise).

The code proposed to identify the GATT body is a two-letter code: "AG" to identify the Committee on Trade in Agriculture, "QR" to identify the Group on Quantitative Restrictions and Other Non-Tariff measures.

It will be possible to record other codes, if necessary.

RECORDS 20-29: GATT document(s) in which the QR was notified

GATT documents in which the QR was notified are coded in record numbers 20 through 29 (one document symbol per record). Document symbols are recorded by the Secretariat. The code adopted is a four-digit code. Document symbols corresponding to each four-digit code are recorded in a separate reference file.

RECORD 30-39: GATT article(s) to which the QR refers

GATT articles to which the QR refers are coded in record numbers 30 through 39 (one GATT article per record). GATT articles are coded in text form on a 70-character element. In some cases, textual comments are recorded in this element. GATT articles and textual comments are converted to a alphanumeric standard code, before the file is loaded into the IDB.

Textual information on the restriction is coded in record numbers 40 through 49 (one line of text per record). Each line contains 70 characters.

Data Inventory for Quantitative Restrictions (continued).

2. TEXTUAL PART OF THE QUANTITATIVE RESTRICTIONS FILE.

The "textual" non-tariff file contains product descriptions for tariff items which are partially covered by a restriction. Thus, the tariff line numbers recorded in this file should all have a number in the non-tariff suffix recorded in positions 22 and 23 of the tariff line number. This suffix indicates that the item is a sub-heading for which no import statistics are available in the statistical file. If an item is fully covered by the measure, and therefore with no suffix or with a suffix equal to "00", the product description of the item can be retrieved from the textual tariff file which contains descriptions for all tariff items.

As in the textual tariff file, product descriptions can be recorded in two different sets of records. The first set of records (element 02) contains full product descriptions which can be recorded in 999 lines of text. The second set of records (element 09) contains abbreviated product descriptions which can be recorded in three lines of text.

Entities	Definitions

Textual Non-Tariff File Key

The key of the textual non-tariff file contains the following elements: the reference year, the code of the country or Customs Union (C.U.) maintaining the restriction, the code of the member country of a C.U.) maintaining the restriction, the tariff line number, the element number, and the sub-element number (sequence of the lines of text).

Tariff line number

The tariff line number is an element of 25 characters used to record the tariff item number (first 21 characters), the non-tariff item suffix (positions 22 and 23) and the tariff item suffix (positions 24 and 25).

In this file the tariff item suffix is always blank.

Tariff item number

The tariff item number (first 21 characters of the tariff line number) is the same as the tariff item number used in the import statistics file, in the tariff files and in the coded non-tariff file.

Tariff item suffix

The non-tariff item suffix is recorded in positions 22 and 23 of the tariff line number. The suffix is used in cases where the restriction covers part of a tariff item, i.e. statistics are not available for those sub-headings.

Element number

The element number is a two-digit number identifying the type of description. Element "02" contains the full text of the product description; element "09" contains the abbreviated description.

Sub-element number

The sub-element number is a three-digit number identifying the sequence of the lines of text of the product description. The full text description is recorded in sub-elements 001 to 999. The abbreviated description is recorded in sub-elements 001 to 003.

Full Product Description (ELEMENT 02)

The full product description is recorded in element "02" sub-elements 001 through 999. Each line of text contains 70 characters.

A "level" number can be assigned to each description to distinguish descriptions of headings (not associated to a duty rate) from descriptions of tariff lines (associated to a duty rate). If the file contains descriptions of headings, the "level" of the heading description should be equal to the number of digits used in the tariff number ("02" for a chapter heading, "04" for 4-digit heading, etc.). Descriptions of tariff lines should all have a "level" equal to "00". If no headings are recorded, the level should be left blank.

Each text line of the description can be indented by indicating in the two-digit "indent" element the number of blank characters to be left between the left margin and the text.

Abbreviated Product Description (ELEMENT 09)

Element "09" contains a maximum of 3 sub-elements, corresponding to 3 lines of text of 70 characters, for recording the abbreviated product description. The layout of this element is the same as the layout of the full product description (see element 02 description), however, the two fields "level" and "number of blanks for indentation" should be left blank in this element.

DATA BASE CONSIDERATIONS.

This section of the IDB Project Definition document contains an overview of the following points:

A. Data storage and volumes:

Countries' data files.

Gatt Reference files.

Other IDB-related files.

Volumes of information.

B. Data base organization:

Access requirements.

Production data base versus working data base.

Maintenance sub-system.

C. Access control.

D. Hardware and software matters.

E. Information services.

A. DATA STORAGE AND VOLUMES.

This section describes the various information stored in the IDB at the tariff line level. The scope of the data covers three main sets of information:

- Data files, received from the participating countries;
- Gatt Reference files, created and maintained by the Secretariat, which are used for data validation, standardization and reports presentation;
- Various files set up by the Secretariat to handle other IDB-related matters.

(1). Countries' data.

Processing steps applied on the information received from countries participating in the IDB have been outlined in the previous section of this document. Decision to load the data is based on the results of these steps. The data are stored according to the three domain of information:

IMPORT STATISTICS.

Trade statistics submitted by all participants are merged into one data base file. In addition to the material received, the Secretariat will generate the partner aggregates required for the various reports production. Import values are stored in national currencies as well as in US dollars. The file sorting sequence is:

Reference year, Reporting country code (including C.U. member, where applicable), Statistical item number and Partner country (or partner aggregate) code, treatment code.

TARIFF INFORMATION.

Tariff data submitted by all participants are merged into two data base files:

One data base file is dedicated to the tariff "coded" part (see Formats, part B(i)). The file sorting sequence is:

Reference year, Reporting country code, Tariff line number, Validity period, Type of measure, Partner country, Type of duty rate and record number.

The record number identifies different types of data relating to the duty rate:

- Record number 10 contains all the codes attached to a duty rate and this record is mandatory.
- Record number 20 is optional and contains duty rate expressions in form of a normalized coding for specific and other rates
- Records 30 to 39 are optional and contain textual information

Another data base file is dedicated to the tariff "textual" part (see Formats, part B(ii)). The file sorting sequence is:

Reference year, Reporting country code, Tariff line number, Element number and Sub-element number

Before loading the information into the data base, some of the elements will be re-formatted in order to have only fixed-length record layouts, which will better suit the requirements of data display and reporting.

QUANTITATIVE RESTRICTIONS.

This domain contains data submitted by all participants and is organized in a similar manner to Tariff information:

One data base file is dedicated to the QR "coded part" (see Formats, part C(i)). The file sorting sequence is:

Reference year, Country maintaining the measure (including C.U. member, where applicable), Tariff line number, Validity period, Type of measure, Country affected, Measure code and Record number.

The record number identifies different types of data relating to the quantitative restriction:

- Record number 10 contains all the codes attached to a quantitative restriction.

- Records 20-29 are used to store references (maximum ten) to Gatt documents in which the measure was notified.

- Records 30-39 contain indications of the grounds and justification of the measure (maximum ten Gatt Articles).

- Records 40-49, which are optional, contain comments related to the measure. These comments give supplementary information: on the product to which the measure applies, on the measure itself or on any other part of the notification.

The second data base file for this domain contains information on the "textual" part of the QRs (see Formats, part C(ii)). Entries in this file are required every time a quantitative restriction applies to part of a tariff item. The file contains abbreviated or full descriptions of the affected product parts. The file sorting sequence is:

Reference year, Country maintaining the measure (including C.U. member, where applicable), Tariff line number, Element number and Text line sequence.

(2). IDB reference files.

This group of data base files contains data required primarily during the execution of processing steps, from the reception of data submissions to the loading of the information into the IDB. Throughout the steps, the reference files are used to validate national data and to convert codes and values used in capitals to common units and standard codes. Secondly, these files are used in information displays and reports: they contain texts/names associated to codes in order to give a more intelligible representation of the coded data.

COUNTRY CODES FILE.

Contains information on geographical definitions used throughout the system.

Country code This 3-digit code is used for all geographical entities such as reporting country, partner country, country maintaining quantitative restrictions, country affected by measures and country entitled to Initial Negotiating Rights. The coding system follows ISO recommendations. This code is also used to identify trade aggregate totals by type of relation and tariff treatment.

Gatt member code.
This code is used for the aggregation of trade from Gatt members.

Country abbreviations.
Names are recorded in three abbreviations (2 characters, 3 characters and 8 characters).

Country name. ISO country name.

Language code The language code (English, French, Spanish) is used to print/display reports documentation in the Gatt official language.

REPORTING COUNTRY FILE

Contains information related to countries submitting data to the IDB.

Year IDB Reference year

Reporter code This code identifies the countries which submitted information; six characters are needed:

Country (or C.U.) code 3-digit ISO code.

C.U. member This digit ISO code identifies information reported by an C.U. member.

Information status
Series of codes which indicates for which domains the reporter concerned submitted information to the IDB, such as tariffs (coded, textual), trade statistics (imports, exports), quantitative restrictions (coded, textual), etc..).

Tariff treatments
Series of codes which represent a list of tariff treatments applied by the reporter.

HS indicator Indicates whether or not the national nomenclature is based on the Harmonized System.

Exchange rate Exchange rate, for the reference year, of national currency unit per one U.S. dollar, where applicable.

Imports currency
Unit of national currency used in the submission of import values.

Duties currency Unit of national currency used in rate expressions in the normalized coding of the rate.

Currency adjustment factor
A factor relating different currency units used in import statistics and in normalized coding of duty rates.

Supplementary quantity signal
Code indicating if supplementary units have been reported in import statistics.

Other codes Used for documentation.

System of trade Code for type of trade: special trade, general trade, etc..

Valuation basis Code for valuation basis: f.o.b, c.i.f. etc..

Duty assessment Code for specific rates assessment: net weight, gross weight etc..

Type of year Indicates, if required, the basis for recording statistics: 'fiscal' or 'calendar' year basis.

NATIONAL COUNTRY CODES FILE

Contains the geographical coverage used in national data submissions for trade partners in import statistics and for affected countries in non-tariff measures files.

Year Reference year.

Reporter code Gatt reporting country code.

National country code
Geographical code used in national submissions for recording partner (or affected) countries.

Partner code 3-digit ISO code.

Partner type of relation
Type of relation code applicable to each partner for the reference year concerned.

LDC indicator Code attached, where applicable, to least developed countries entitled to differentiated preferential treatment.

INITIAL NEGOTIATING RIGHTS FILE.

Contains the list of codes used to identify countries entitled to INRs.

INR code Used in "textual" tariff information file to indicate partner countries holding INRs. Refer to document TAR/W/5 and Add.1 for a list of code values.

Partner code Gatt partner code associated to the INR code linking INR holders to import statistics and QRs.

INR country name
INR holder name.

CUSTOMS UNIONS MEMBER COUNTRIES DEFINITION.

Contains, for each period of reference, the list of partner codes defining C.U. members. Different C.U. codes are used to reflect changes in the C.U. definition.

C.U. code Gatt country code assigned to one C.U. definition.

Period Start date and end date applicable for the definition (year only).

C.U. members list
C.U. members for the period: country codes stored in ascending order.

QUANTITY UNIT FILE.

Contains a list of all quantity unit codes used in the IDB.

Unit code Gatt 3-digit quantity unit code.

Unit abbreviation
3-letter quantity unit abbreviation.

Unit name Quantity unit name.

NATIONAL QUANTITY UNIT FILE.

Contains, for each reporting country, a list of national quantity units associated to Gatt quantity unit codes (note that the information is not related to the year of reference).

Reporting country
Gatt 3-digit reporting country code.

National quantity unit
Unit code used in trade statistics and in the normalized coding of the "coded" tariff file.

Gatt unit Gatt 3-digit quantity unit code.

MEASURE CODES FILE.

Contains a definition of non-tariff measures.

Type of measure 2-digit code indicating the nature of the restriction (e.g. Licensing, Prohibition).

Measure code This 3-digit code is used to define non-tariff measures.

Measure symbol The symbol which is associated with the measure code (i.e BQ, AL, STR..).

Measure designation
Text to be associated with the measure.

NATIONAL MEASURE CODES FILE.

Contains the national definitions of non-tariff measures reported in data submissions.

Year Reference year.

Reporter code Gatt reporting country code (including C.U. member).

National measure code
Any symbol used in national submissions for recording QRs information.

Measure code 3-digit Gatt measure code as listed in the MEASURE file.

GATT LEGAL INSTRUMENTS FILE.

This file is used for the validation of "textual" tariff information file (elements 04, 06, 07).

Instrument symbol
Code for legal instruments. Refer to document TAR/W/6 for a list of symbols.

GATT DOCUMENTS FILE.

This file is used for the validation of QRs information contained in record numbers 20-29.

Document code Gatt numeric code.

Gatt document symbols
As they appear in Gatt documentation.

GATT ARTICLES FILE.

This file is used for the validation of QRs information contained in record numbers 30-39.

Article code Gatt alphanumeric code.

Gatt Article symbol
As they appear in Gatt documentation.

HARMONIZED SYSTEM NOMENCLATURE FILE.

This file is used for the validation of all data reported according to the Harmonized System nomenclature. The file contains descriptions applicable to the various structure levels of the HS.

HS number HS 6-digit number (or 5 digits, 4 digits or 2 digits).

Language code Used to identify English, French and Spanish descriptions.

HS descriptions Abbreviated text attached to each HS number.

CODEBOOK.

This file is used for the validation of all codes recorded in submissions.

Code name Data name assigned to the various codes such as BINDING, NATURE, MEASURE. The code name is used to access the codebook file.

Occurrences Number of possible values for a code.

Code length Number of characters necessary to represent a code.

Code values List of all possible correct values for a given code.

PROCESSING MESSAGES FILE

Contains a list codes and associated messages that can be issued by IDB programmes.

Message code Numeric 4-digit code.

Message type E = error, resulting in record drop; W = warning, the record is kept; I = information, etc..

Text message

(3). IDB related files.

The following files are necessary for the IDB system:

LOG FILE.

This file contains information stored by the conversions programmes such as:

Programme name

Processing date and time

Reporting country name

Domain processed (Tariffs, Statistics, QRs)

Information control such as:

Total value (national currency, U.S. dollars)
Number of records in and out
Number of tariff lines and tariff items
Data analysis report, when applicable
Etc..

This information is made available to the IDB users to answer queries on the data availability by reporter and progress in the work involved in the conversion of the submissions.

COUNTRY NOTES.

This file is organized by reference year, reporting country and domain of information. It contains information on all matters related to the countries' submissions and gives detailed explanations concerning the problems encountered in the conversion steps as well as on specific points of methodology for the work done in the Secretariat. This file is available to the IDB users.

(4). Volumes of information.

The following volumes of data to be stored in the IDB are estimates. They are based on the experience drawn from the Tariff Study and from the Harmonized System common data bases exercises. Numbers are given for one reporting country and one reference year.

Import statistics.

Tariff (or statistical) items : 10000 national HS-based codes.
Partner codes (plus partner aggregates): 20 country codes.
Total number of records: 200000; average length: 100 characters; 20 million characters.

Tariff information.

Coded tariff file:

Tariff items: 10000 national HS-based codes.
Duty types: 1 MFN rate and 2 preferential rates.
Partner code(s): 1 or 2 partner codes.
Validity period: 1 period.
Record number: 2 types of records (record no. 10 plus one optional record).

Total number of records: 60-100000; average length: 70 characters; 5 to 7 million characters.

Textual tariff file:

Tariff items : 8000 national HS-based codes (bound items only).
Elements per tariff item: 3.

Total number of records: 25000; average length: 100 characters; 2-3 million characters.

Quantitative restrictions.

Coded QR file:

Tariff items : 3000 national HS-based codes.
Type of measure: 1 type, QR applied to imports.
Number of measures per item: 1 measure.
Record number: 2 types of records (no. 10 and 30).

Total number of records: 6000; average length: 70 characters; 4 million characters.

Textual QR file:

Tariff items: 500 national HS-based codes (X items only).
Elements per tariff item: 1 element 02 - abbreviated description.
Number of text lines: 3.

Total number of records: 1500; average length: 100 characters; 150 thousands of characters.

TOTAL SPACE REQUIREMENTS : 30 million characters per reporter/year.

B. DATA BASE ORGANIZATION.

Access to the information.

The data identified as key fields in the various files are used to access and link the IDB components:

KEY FIELDS	COUNTRY FILES: Tariff information Coded	Tariff information Textual	Statistics	QRs information Coded	QRs information Textual
Reference year	Y	Y	Y	Y	Y
Reporting country	Y	Y	Y	Y	Y
Tariff item no.	Y	Y	Y	Y	Y
- Tariff suffix	Y	Y	N	N	N
- QR suffix	N	N	N	Y	Y
Validity period	Y	N	N	Y	N
Type of measure	N	N	Y	Y	N
Partner country	Y	N	Y	Y	N
Duty rate type	Y	N	Y	N	N
Measure code	N	N	N	Y	N
Record/element no.	Y	Y	N	Y	Y

KEY ELEMENTS	REFERENCE FILES: Cty	Rep	Nat cty	INR	FTA	Unit	Nat. unit	Inst.	Docs.	Art.	HS	National measures
Ref. year	N	Y	Y	N	Y	N	N	N	N	N	N	Y
Reporter	Y	Y	Y	N	Y	N	Y	N	N	N	N	Y
HS nos.	N	N	N	N	N	N	N	N	N	N	Y	N
Partner	Y	N	Y	Y	Y	N	N	N	N	N	N	N
CHECKING ELEMENTS												
Tariffs:												
- normal.coding							Y					
- elmt. 04/06/07								Y				
- elmt. 05/07				Y								
Statistics:												
- qty units							Y					
QRs:												
- measure code												Y
- documents									Y			
- articles										Y		

Data base maintenance sub-system.

(1) Organization.

When the files submitted by a participant are ready to be loaded into the data base, the participant concerned will receive a copy of the data analysis reports produced at the various stages of the verification and conversion processes. At the same time, the files will be loaded into a "working" data base. After receiving authorization by the country concerned, the files will be transferred to a "production" data base and made available to users. Users would have access to the "production" data base only.

It can be expected that the submissions would not contain all elements of information which were requested. Missing elements would be added progressively either on the basis of further partial submissions from national authorities or by the Secretariat, in the framework of technical assistance, with the help of national experts. Missing information would be added, step by step, to the "working" data base. Updated versions of the files would be transferred regularly to the "production" data base.

(2) Update steps.

(a) Data capture: the update transactions are created by Secretariat staff members in charge of the reporting country concerned for the three domains: tariffs, statistics and non-tariff measures. The data entry programmes will be designed when all IDB data base matters are defined. The design of these programmes should take the following considerations into account:

- Efficient support of the organization set up for the group in charge of the data preparation work.
- Best use of update techniques in liaison with the data base design.

(b) Data checking: all information entered through update transactions will be checked against Gatt Reference files. Data which do not pass through this checking process will be dropped and listed on error reports.

(c) Update processing: update entries will be applied to the "working" data base according to the transaction specifications. It is envisaged to set up two separate update systems: an on-line procedure for all "textual" information (Tariff and QR files, textual parts) and a batch procedure for other information. A separate on-line update system will maintain all the Gatt Reference files. All update transactions will be listed by reporting country and domain. These reports will show the 'before' and 'after' data status. The update reports could be made available, on request, to the reporting country.

(d) Post-update verification programme: when all the update work for a reporter is complete, a post-update programme will be run to detect inconsistent information (logical checking among the files will be applied at this stage). The necessary corrections will be introduced via the update programme. The country concerned will receive a series of reports detailing the files' contents for their verification. When permission is received from national authorities, the information will be transferred to the "production" data base and be made available to users.

C. ACCESS CONTROL.

Users' access to the IDB information will be controlled through passwords, by type of operations (read-write or read-only modes). The access restrictions will be installed through built-in security features of Adabas, the data base management system used for the IDB. Security features implementation and maintenance will be placed under the IDB administrator.

As mentioned previously in the "Data base maintenance sub-system" section, there will be two versions of the IDB: a "production" data base and a "working" data base. Users will have the ability to access the "production" data base only. This possibility will, in fact, guarantee the information stability for a period of time and avoid the users being affected by delays in problems solving and by the maintenance activities (e.g. programme errors, load procedures, data backups and recoveries, etc.).

D. HARDWARE AND SOFTWARE USED FOR THE IDB.

Hardware

The IDB information and programmes will reside on the equipment of the International Computing Centre in Geneva (ICC). The Gatt Secretariat devices (terminals, control units) are linked to the Centre through modems and telephone lines. For more details about the ICC hardware and services, see note below. Personal computers located in delegations or in capitals, and connected to the ICC, could also be used as terminal units to access directly the IDB. This type of equipment could also be used, in local mode, to process IDB data down loaded from the ICC computer.

Software

The IDB project will be developed using a Data Base Management System:

Adabas from Software AG company. Other products of this company based on Adabas technology, will be used:

Natural, a fourth generation programming language,

Super-Natural, a programme and data views generator for non-professional data processing users,

Predict, an integrated data dictionary,

Natural Security System, to define security requirements.

IBM PL/1 language will be used for coding functions not adapted to fourth generation technology. These programmes will be developed using IBM PDF Dialog Manager.

Note on the International Computing Centre

The Secretariat can provide the following detailed information about the International Computing Centre:

a. Services provided by the Centre.

b. Hardware configuration:

- Processing units
- Main storage units
- Direct Access and magnetic tape facilities
- Teleprocessing facilities

c. The CALL/ICC Interactive Network:

- ICC's Electronic Mail service
- On-line Help and News services
- Access to commercial or international organization-owned data bases

d. Procedure to access the ICC through packet switched networks.

e. Operational Policy of the ICC:

- Services availability
- Procedure to Authorise Usage of the ICC Facilities

E. INFORMATION SERVICES.

This section outlines the initial possibilities which will be made available to IDB users for data dissemination. These services will use the "production" data base as input.

Machine-readable output.

(1) Magnetic tape: the data base unload procedure will offer several file formats for the distribution of data on tape. One of these formats will correspond to the format of the files used in the loading procedure which will be described in the Data Base Design documentation. Other formats will be defined after users' requirements are communicated to the Secretariat.

The tapes which will be distributed will have IBM standard labels and the recording density could be 1600 or 6250 bits per inch.

(2) Diskettes: this is another possible magnetic medium for IDB information distribution. For information requests, where a limited volume of data is involved (e.g data produced from a list of selection criteria), the Secretariat could produce micro-computer diskettes. The possible data arrangements should be defined in view of precise users' requirements and would have to take into account some of the many micro-computer file formats. It would be very helpful to know which software packages the participants intend to use on their micro-computing equipment.

On-line users.

Users connected to the ICC (via a dedicated telephone line or commercial networks) will have the ability to access the information stored in the IDB, using Gatt interactive query programmes. In the same manner, batch jobs could also be prepared and submitted to the ICC. Corresponding jobs output would be printed locally or at the ICC and routed to the users' location. The required local equipment is a computer terminal (3270-type) or a micro-computer which emulates this type of terminal and, optionally, a terminal printer or a printer hooked up to a micro-computer. Note that if a micro-computer is available, the users could directly down load selected IDB information on their local storage devices (hard-disk or diskettes).