ISO JCT1 SC32 WG2
June 2012 Berlin
Metamodel for Fact Based Information
Model Registration

Baba Piprani - baba@metaglobal.ca
Serge Valera - serge.valera@esa.int
The Fact Based Modelling WG

- Initiated in October 2009
- Main objective is to specify means to exchange fact based oriented conceptual data models
- WG organization
  - Convenor S. Valera F, ESA
  - Secretary I. Lemmens B
  - Method representatives
    - CogNIAM S. Nijssen NL
    - FCO/IM D. Smeets NL
    - DOGMA R. Meersman B
    - ORM T. Halpin AU
  - Experts
    - C. Heath AU
    - J.P. Koster NL
    - M. Curland US
    - V. Morgante AU
    - Y. Tang CN
    - H. Balsters NL
    - P. Bollen NL
    - L. Bruil NL
    - A. le Cat NL
    - N. de la Cruz US
    - K. Evans UK
    - B. Piprani CA
    - R. Schmaal NL

www.factbasedmodelling.org
The Fact Based Modelling WG, cont.

- **FBM tools include**

- **Toward ISO**
  - 24/10/2011 – WG2N1612_FBM_WD01
    - 1\(^{st}\) Draft containing a NORMA view of the FBM Specification
  - 19/04/2012 – WG2N1640_FBM_WD02
    - 2\(^{nd}\) Draft containing a UML representation of means to register FBM models
Toward Standardization of FBM

- **ECSS** – European Cooperation for Space Standardization
  - E-10 System Engineering, Space System Data Repository – ECSS-E-TM-10-23A
- **ISO**
  - Registering Fact Based Information Models
  - ...
- **OMG**
  - Semantic Information Modelling for Federation RFP
- **ESA**
  - Software Engineering – Fact Based Modelling and Databases
  - Specifying at Global level using fact based modelling
- **CCSDS** – Consultative Committee for Space Data Systems
FBM focuses in specifying data requirements using “Natural Language” e.g. Plain English, in a “Controlled way”.

Modelling with FBM implies
- Fully specifying all data requirements i.e. addressing semantics, focusing on stakeholder’s needs
- be free of logical and physical implementation features

FBM provides the means to reach the 100% principle expressed in ISO TR9007

The “Metamodel for Fact Based Information Model Registration” shall be able to formally capture the specification of the knowledge for interoperability purpose
4.34

**Fact**

A proposition that is taken to be true by the relevant community whereby the proposition is represented by a predicate with an associated ordered set of objects.

1. **SOHO** was launched on the 2nd of December 1995.
2. **SOHO** was launched on the 2nd of December 1996.
3. **Herschel** was launched on the 14th of May 2009.
4. **Planck** was launched on the 14th of May 2009.
5. ...
Facts are used to illustrate a specification but more important to validate the proposed model.

Modelling facts allows verifying the adequacy of the model to capture the truth but also to detect invalid statements.

1. SOHO was launched on the 2\textsuperscript{nd} of December 1995.
2. SOHO was launched on the 2\textsuperscript{nd} of December 1996.
3. Herschel was launched on the 14\textsuperscript{th} of May 2009.
4. Planck was launched on the 14\textsuperscript{th} of May 2009.
5. ...
4.37

**fact type**

*type, each of whose instances are facts that express the same kind of information*

1. `<satellite>` was launched on `<launch date>`
2. The operations of `<satellite>` have been extended until `<end of life>`
Fact Types define which facts are considered to be in scope.

*A fact type is a populatable construct, a typed construct, generalizing the facts on the basis of common properties.*

*Fact types can be atomic elementary or existential or compound.*

1. `<satellite>` was launched on `<launch date>`
2. The operations of `<satellite>` have been extended until `<end of life>`
4.55

**predicate reading (syn. Fact Type Form)**

*sentence with placeholders for object(s) that expresses a logical predicate*

**NOTE** A predicate must have at least one predicate reading.

1. `<satellite>` was launched on `<launch date>`
   but also for example
   translation:
2. `<satellite>` a été lancé le `<date de lancement>`
   Reverse reading:
3. On `<launch date>`, `<satellite>` was launched
Predicate Readings define the mechanism to be used to communicate about the facts and rules about the populations of facts using domain-specific (stakeholders) terminology.

1. `<satellite> was launched on <launch date>`
   but also for example
   translation:
   1. `<satellite> a été lancé le <date de lancement>`

Reverse reading:
1. On `<launch date>`, `<satellite> was launched`
Validation rules (VR)

Identify the validation rules (a.k.a. the constraints, the integrity rules) that restrict the fact population to those that are considered useful, the set of facts and the transitions between fact populations.

- Validation rules have the function to restrict the populations permitted by the fact types to useful ones.

  Constraints such as uniqueness, mandatory, frequency, ring, value comparison, equality, exclusion, subset, value, cardinality constraints
4.71

**uniqueness constraint**

*constraint over a sequence of one or more roles that requires that in each state of the fact base, each instantiation of that role sequence occurs only once*

1. *Each* `<satellite>` *was launched on* **at most one** `<launch date>`.
2. *rules applying to the reverse reading direction:* **On** `<launch date>`, **zero or more** `<satellite>` *are launched.*
4.46

**mandatory role constraint**

Constraint requiring that each instance in the population of a given object type must play the constrained role or at least one of the constrained roles (inclusive-or)

1. Each `<product>` has some `<product name>`. 
4.40

**frequency constraint**

A constraint that restricts, for each state of the fact base, the number of times any given sequence of objects that instantiate the constrained role sequence appears in the population of that role sequence.

1. *Each* `<OTS hardware procurement payment plan>` consists of *at most 2* `<payments>`.
4.73

**value comparison constraint**

constraint that specifies how the values of instances of two roles with co-roles played by the same object are related by one of the following comparison operators: \(<, \leq, >, \geq\)

1. The <starting date> of each <Project> shall be *earlier than its <ending date>*.
4.74

value constraint
constraint that specifies the possible values for the instances of a role or a value type

1. The <risk level> is one of "low", "medium" or "high".
4.30

exclusion constraint

set-comparison constraint that specifies that, for each state of the fact base, the populations of the constrained sequences of role occurrences must be mutually exclusive i.e. do not overlap.

1. For each <procured product>, at most one of the following holds:
   • that <procured product> is some <off the shelf product>;
   • that <procured product> is some <developed product>.
4.29

**equality constraint**

*set-comparison constraint that specifies that, for each state of the fact base, the populations of the constrained sequences of role occurrences must be equal*

1. For each `<invoice>` that `<invoice> <is paid> if and only if that <invoice>` has some `<payment date>`.

```
+---------------------------+--------------------------+-----------------------------+------------------------+
|                        | Equality_Constraint      |   -comparing_equality_constraint | 0..1                    |
|                        |                          |                                |                        |
|                        |   -compared_compared_role_list_occurence | 2..*                      |
+---------------------------+--------------------------+-----------------------------+------------------------+
```

**Compared_Role_List_Occurrence**
4.68

subset constraint

set-comparison constraint that specifies that, for each state of the fact base, the population of a sequence of one or more role occurrences must be a subset of the population of another compatible sequence of role occurrences

1. If some <procured product> is for <project>, then that <procured product> is part of some <project> related <product tree>.
4.9

**cardinality constraint**

A constraint on an object type or role that determines the possible number of instances that the object type or the role may contain for any given state of the fact base.

1. For each `<product tree>`, exactly one `<product tree element>` has no `<parent product tree element>`.
4.62

ring constraint

logical constraint between two type-compatible role occurrences that specifies how the populations of these role occurrences may be related

EXAMPLES irreflexive, asymmetric, intransitive, antisymmetric and acyclic

1. No <product tree element> may cycle back to itself via one or more traversals through <product tree element> is child of <product tree element>.
   i.e. <product tree> shall be acyclic and intransitive
4.23
derived fact type
fact type, each of whose instances is a derived fact

4.22
derived fact
derived fact
fact that is deduced from other facts by means of a derivation rule

NOTE A fact that is not derived is an asserted fact.

4.20
derivation rule
rule that specifies how to derive instances of a derived fact type or semi-derived fact type from other facts
Identify the rules required to derive (to calculate) new information (i.e. new facts) on the basis of existing information (i.e. existing and derived facts).

*Derivation rules describe how to derive new facts on the basis of existing facts.*

1. The development cost of `<satellite>` is `<satellite development cost>`.

   **Derivation rule:**

   The `<satellite development cost>` of `<satellite>` is equal to the sum of all `<satellite>` related `<subsystem development cost>` + the sum of all `<satellite>` related `<system activity cost>`.
• Any questions?
WHAT is DATA, WHAT is INFORMATION

- DATA
  "Data" on its own carries no meaning.
  
  Examples:
  "Val Thorens", "2300"

- INFORMATION
  In order for "data" to become "information", it must be interpretable, that means extended with Semantics [meaning the conceptual data model], e.g. "what the nouns refer to, what the verb means".
  
  Examples:
  "Val Thorens" refers to the highest ski resort in Europe.
  "2300 is meant to be 2300 meters above sea level"
  The missing verb is "is located at"
WHAT is KNOWLEDGE

• Knowledge
  Knowledge is when we know everything about the validation associated with the information.

Examples:

  Val Thorens is a ski resort in Europe!
  True → Val Thorens is located in the Alps.

  Val Thorens is the highest ski resort in Europe!
  True → If one populates all ski resorts located in Europe together with their altitude, it is easy to derive and validate that Val Thorens is the highest ski resort in Europe.
Can FBM models be instantiated in ISO 19763 part 12?
Type-Instance constructs

- FBM has only one type-instance construct: the fact type
- 19763 part 12 currently provides for three type-instance constructs:
  - the entity type
  - the attribute type
  - the relationship type

None of these three is able to express the fact type, there is a mismatch.
Terms and definitions are different between 19763-12 and FBM

- FBM supports n-ary fact types.
- 19763-12 Entity type and attributes only maps binary relationships. They cannot handle unary and n-ary relationships without artificial constructs.

- FBM has fact communication patterns (expressing facts) and rule communication patterns (expressing fact types)
- 19763-12 does not.

Note The Editors note1 on page 1 of 19763-12 comes to a similar conclusion.
Constraints

- FBM provides the functionality as available in set theory w.r.t. subset, equality and exclusion.

- 19763-12 only permits the referential integrity (a very special case of subset) and exclusion on association ends.
ISO TR9007

- FBM is based on the ISO TR9007 100% and Conceptualization principles.
- The 100% principle requires much more expressive power in integrity rules than currently possible in 19763-12.
Arity on roles

- FBM permits any arity in the fact type.
- 19763-12 does not permit unary associations.
Type of constraints over sets and roles

- FBM has a full set of integrity rules to satisfy the ISO TR9007 100% principle (provided that derived fact types can be freely expressed, see next slide)

- 19763-12 cannot represent among others the non-overlap constraint, cardinality on combinations (e.g. There are at maximum two subjects awarded per Nobel category per year; there are at maximum three Nobel Laureates per Nobel category per year) and the ring constraints.
Recognition of derived concepts

- One of the ways to satisfy the 100% principle is to apply the engineering rule to replace a complicated expressions with a series of related but easy to understand expressions. This requires the concept of derived fact type as fully supported by FBM.

- This concept is currently not represented in 19763-12.
Recognition of objectified concepts

- FBM offers the subject matter expert the option to include nominalization (objectification, reification) in their information models.
- In 19763-12 this functionality is not represented.
Extra Motivation

by Prof. G.M. Nijssen
Why is it recommended to have FBM in ISO 19763?

The answer to this question can be given in a historical perspective...
Historical perspective (1)

1963

- Up till 1963 automatic business data processing was based on sequential record processing. The distance between business processes and IT was quite large.
- In 1963 Charlie Bachman introduced IDS (Integrated Data Store) at GE. From then on business facts could be encoded both as attributes of a record as well as through pointer structures using hardware addresses of a record instance.
1971

- In 1971 CODASYL (Conference on Data Systems Languages) issued a standard for DDL and DML. It was basically an extension of IDS of 1963.
- The result was that there were then three IT fact encoding mechanisms:
  1. Attributes in a record instance
  2. Pointer structures between record instances and
  3. Repeating groups within a record.
1971

- CODASYL DDL and DML (as well as IMS of IBM, a hierarchical data structure system with about the same expressives as CODASYL DDL) were then considered major steps forward in the IT solution space.

- Ted Codd published his famous article in the Communications of the ACM about the relational model.

- A few years later the historical debate between Charlie Bachman and Ted Codd took place.
1974

- Establishment of IFIP (International Federation of Information Processing) WG 2.6 Databases.
- WG 2.6 concentrated from the very beginning on true conceptual aspects, fully independent of IT implementations but based on natural language to make models understood by the business and on logic to be able to apply more productive techniques.
1978-1987

- Several persons involved in IFIP WG 2.6 started a wider cooperation in ISO. The work was performed during the ten year period 1978 through 1987.

- In 1987 the ISO TR9007 was published: Concepts and Terminology for the Conceptual Schema and the Information Base.

- The intent was to produce a purely conceptual modeling language; for practical consensus reasons an ER dialect was also included.
Mid and late 1980’s

- Many ER dialects were introduced and OO was being accepted as the big (IT) step forward.

Early 1990’s

- OMG came with the proposal for a standard for the various ER dialects to be combined with some OO aspects: UML was born.
- UML became a fantastic success in terms of followers. It became the de-facto modeling standard, strongly biased by IT.
Second decade of the 21st century

- The need for much better specifications independent of IT signed off by the business and expressed in a language business understands, became clear.

FBM 2012 - The re-birth of ISO TR9007
Today (2)

• There is an increasing demand for systems to *interoperate* by exchanging data. For these data exchanges to be meaningful it is essential that the business information requirements that are met by the data stored in these systems are understood so that suitable data exchange mechanisms can be developed.

• Business information requirements, including the semantics of the information, need to be represented by conceptual data models before any software development and associated logical and physical database schema.
• This new part of ISO/IEC 19763 intends to provide a generic framework for registering these information models using the Fact Based Modelling approaches.

• This standard treats ISO TR9007 (1987) as its base.

• The combination of the ISO TR9007 100% principle and the ISO TR9007 Conceptualization principle make a new dimension possible in conceptual modeling, namely the automatic transformation of a model expressed in one conceptual notation into another conceptual notation.
• The ISO TR9007 opens the way to a productive approach that is able to trap all errors as early as possible in the conceptual specification and have the subject matter experts involved in the validation as early as possible, long before program code development and testing.

• The ISO TR9007 based architecture permits true 100% conceptual modelling and transformation.
Overview of the FBM Protocol

by Prof. G.M. Nijssen
### What is Fact Based Modeling? (1)

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>What</td>
<td>assumed to be known</td>
</tr>
<tr>
<td>is</td>
<td>assumed to be known</td>
</tr>
<tr>
<td>Fact</td>
<td>see next few slides</td>
</tr>
<tr>
<td>Based</td>
<td>assumed to be known</td>
</tr>
<tr>
<td>Modeling</td>
<td>the process that results in a Conceptual Schema (ISO TR9007)</td>
</tr>
</tbody>
</table>
54. What is Fact Based Modeling? (2)

1. Fact Based Modeling (FBM)
   uses only 1 type-instance construct, namely: fact type and fact.

2. FBM
   is a protocol which systematically uses concrete examples for developing the model as well as for immediate verification and completeness validation of the model.
What is a fact? By example (1)

1. Mozart visited Italy in 1769.
3. Verdi visited France in 1853.
4. Mozart was born in Austria.
5. Mozart died at the age of 35.
6. Verdi was born in Italy.
7. Verdi died at the age of 87.
8. The capital of Austria is Vienna.
9. The capital of Italy is Rome.
10. Bach was born in Germany.

These examples show 10 individual facts (or: fact instances)
56. What is a fact? **By example (2)**

1. The Nobel Prize in Physiology or Medicine 2010 was awarded to Robert G. Edwards "for the development of in vitro fertilization".

2. The Nobel Prize in Physiology or Medicine 2006 was awarded jointly to Andrew Z. Fire and Craig C. Mello "for their discovery of RNA interference - gene silencing by double-stranded RNA".

3. The Nobel Prize in Physiology or Medicine 1966 was divided equally between to Peyton Rouse "for his discovery of tumour-inducing viruses" and Charles Brenton Huggins "for his discoveries concerning hormonal treatment of prostatic cancer".

These examples show 3 individual facts (or: fact instances)
57. What is a fact? By example (3)

1. The Nobel Prize in Physiology or Medicine 2002 was awarded jointly to Sydney Brenner, H. Robert Horvitz and John E. Sulston “for their discoveries concerning genetic regulation of organ development and programmed cell death”.

2. The Nobel Prize in Physiology or Medicine 1977 was divided, one half jointly to Roger Guillemin and Andrew V. Schally “for their discoveries concerning the peptide hormone production of the brain” and the other half to Rosalyn Yalow “for the development of radioimmunoassays of peptide hormones”.

3. For the year 1942 no Nobel Prize was awarded in Physiology or Medicine.
58. What is a fact? By definition

“A fact is a proposition that is taken to be true by the relevant community.”
59. An example in FBM and UML (1)

Situation 1 in FBM:

<Famous composer> visited <Country> in <Year>.

Situation 1 in UML:
Situation 2 in FBM

1: <Famous composer> was born in <Country>.

2: <Famous composer> died at age <Age>.

3: <Country> has <City> as its capital.

4: <Famous composer> visited <Country> in <Year>.
61. An example in FBM and UML (3)

Situation 2 in UML
62. Comparison of FBM and UML based information models

<table>
<thead>
<tr>
<th>FBM</th>
<th>UML</th>
</tr>
</thead>
<tbody>
<tr>
<td>Natural language based focusing on specification</td>
<td>Software technology based focusing on implementation</td>
</tr>
<tr>
<td>Logic based</td>
<td></td>
</tr>
<tr>
<td>Extensive protocol</td>
<td>No known protocol</td>
</tr>
</tbody>
</table>
### How does FBM relate to OWL?

<table>
<thead>
<tr>
<th>FBM</th>
<th>OWL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Natural language based</td>
<td>Only unary and binary sentences</td>
</tr>
<tr>
<td>Logic based</td>
<td>Logic based</td>
</tr>
<tr>
<td>Extensive protocol</td>
<td>No known protocol</td>
</tr>
</tbody>
</table>
### How does FBM relate to relational?

<table>
<thead>
<tr>
<th>FBM</th>
<th>Relational</th>
</tr>
</thead>
<tbody>
<tr>
<td>Natural language based</td>
<td>Mathematics based (n-tuples)</td>
</tr>
<tr>
<td>Logic based</td>
<td>Logic based</td>
</tr>
<tr>
<td>Extensive protocol</td>
<td>Normalization</td>
</tr>
</tbody>
</table>
65. Preview on UML metaclasses of WG2N1640_FBM_WD02

In the following slides the most important UML metaclasses of WG2N1640_FBM_WD02 will be illustrated.

Also some slides have been incorporated to describe to what extent ISO TR9007 is in the process of adoption at OMG.
66. Fact Based Modeling using the FBM protocol

Fact Based Modeling using the FBM protocol to specify a Conceptual Domain Model (CDM)

Experience the methodology by utilizing a concise but relevant example:

*Number of different years famous composers visited various countries*
67. From facts to the Conceptual Domain Model

A. More than 99.9% of all business communication consists of facts.

B. Users, management and subject matter experts are very familiar with such facts.

C. FBM is so far the only protocol (methodology) that takes these facts as a solid starting point for a protocol to develop the Conceptual Domain Model in a way that the subject matter expert and the conceptual modeler can productively work together, including quality control. Hence this is a learnable and repeatable process.

D. The steps from the familiar fact to the corresponding cornerstone of the Conceptual Domain Model, the fact type, will subsequently be illustrated.
68. **Facts: the solid starting point**

FBM assumption 1: Concrete facts make up more than 99.9% of all business communication.

**Number of different years famous composers visited various countries:**

<table>
<thead>
<tr>
<th></th>
<th>Italy</th>
<th>Belgium</th>
<th>The Netherlands</th>
<th>France</th>
<th>Russia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mozart</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Verdi</td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

*Presenting facts in the form of a spreadsheet is common practice to most business users.*
69. **Facts: the solid starting point**

FBM assumption 1: Concrete facts make up more than 99.9% of all business communication.

**Number of different years famous composers visited various countries:**

<table>
<thead>
<tr>
<th></th>
<th>Italy</th>
<th>Belgium</th>
<th>The Netherlands</th>
<th>France</th>
<th>Russia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mozart</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Verdi</td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

**Famous composer**

Definition: Person who writes music and is well-known for his work.
70. Facts: the solid starting point

FBM assumption 1:
Concrete facts make up more than 99.9% of all business communication.

Number of different years famous composers visited various countries:

<table>
<thead>
<tr>
<th></th>
<th>Italy</th>
<th>Belgium</th>
<th>The Netherlands</th>
<th>France</th>
<th>Russia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mozart</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Verdi</td>
<td></td>
<td></td>
<td>3</td>
<td></td>
<td>2</td>
</tr>
</tbody>
</table>

**Country**

**Definition:** Political division of a geographical entity, a sovereign territory, most commonly associated with the notions of state or nation and government.
71. Facts: the solid starting point

FBM assumption 1:
Concrete facts make up more than 99.9% of all business communication.

Number of different years famous composers visited various countries:

<table>
<thead>
<tr>
<th></th>
<th>Italy</th>
<th>Belgium</th>
<th>The Netherlands</th>
<th>France</th>
<th>Russia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mozart</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Verdi</td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

**Number of different years**
Definition: Mathematical sum of all the non-overlapping years.
For example, if Mr. Smith visited France once in 1881, twice in 1883, and finally stayed in France from 1886 through 1888, he visited France during 5 different years.
72. **Facts: the solid starting point**

FBM assumption 1:
Concrete facts make up more than 99.9% of all business communication.

**Number of different years famous composers visited various countries:**

<table>
<thead>
<tr>
<th></th>
<th>Italy</th>
<th>Belgium</th>
<th>The Netherlands</th>
<th>France</th>
<th>Russia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mozart</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Verdi</td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

**Visit**

*Definition:* To be in a specific **Country** at a certain moment in time, where the **Country** is not the **Country** of birth.
### 73. Verbalization of facts

<table>
<thead>
<tr>
<th></th>
<th>Italy</th>
<th>Belgium</th>
<th>The Netherlands</th>
<th>France</th>
<th>Russia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mozart</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Verdi</td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

**Verbalize to a colleague:**

Mozart visited Italy in 5 different years.

Mozart visited Belgium in 3 different years.

Verdi visited France in 3 different years.
74. Determine variable and constant positions in facts

Are the following equivalent?

- Mozart visited Italy in 5 different years.
- Mozart visited Belgium in 3 different years.
- Verdi visited France in 3 different years.

and:

- Mozart visited Italy in 5 different years.
- Mozart visited Belgium in 3 different years.
- Verdi visited France in 3 different years.

The Subject Matter Expert answers: “YES!”
75. Determine variable and constant positions in facts

- Mozart visited Italy in 5 different years.
- Mozart visited Belgium in 3 different years.
- Verdi visited France in 3 different years.

**Constant parts**
76. Determine variable and constant positions in facts

Variable parts

Mozart visited Italy in 5 different years.
Mozart visited Belgium in 3 years.
Verdi visited France in 3 years.
77. Further qualify the variable positions (if necessary)

Conceptual domain modeler:

“Do you prefer to use the following representation, in which we have added further qualification:

The famous composer Mozart visited the country Italy in 5 different years.”
78. Abstract from fact instances to fact types

Given the previously introduced facts,

<table>
<thead>
<tr>
<th>Composer</th>
<th>Action</th>
<th>Country</th>
<th>Years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mozart</td>
<td>visited</td>
<td>Italy</td>
<td>5</td>
</tr>
<tr>
<td>Mozart</td>
<td></td>
<td>Belgium</td>
<td>3</td>
</tr>
<tr>
<td>Verdi</td>
<td></td>
<td>France</td>
<td>3</td>
</tr>
</tbody>
</table>

the Conceptual Domain Modeler proposes the associated fact type:

<Famous composer> visited <Country> in <Number of years> different years.

The correspondence between a fact and a fact type is a cornerstone of the FBM Conceptual Domain Modeling protocol.
79. Abstract from fact instances to fact types

Given the previously introduced facts,

<table>
<thead>
<tr>
<th>Composer</th>
<th>Action</th>
<th>Country</th>
<th>Years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mozart</td>
<td>visited</td>
<td>Italy</td>
<td>5</td>
</tr>
<tr>
<td>Mozart</td>
<td></td>
<td>Belgium</td>
<td>3</td>
</tr>
<tr>
<td>Verdi</td>
<td></td>
<td>France</td>
<td>3</td>
</tr>
</tbody>
</table>

different years.

the Conceptual Domain Modeler proposes the associated fact type:

<Famous composer> visited <Country> in <Number of years> different years.
## 80. Abstract from fact instances to fact types

Given the previously introduced *facts*,

<table>
<thead>
<tr>
<th>Famous composer</th>
<th>visited</th>
<th>Country</th>
<th>in</th>
<th>Number of years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mozart</td>
<td>visited</td>
<td>Italy</td>
<td>in</td>
<td>5</td>
</tr>
<tr>
<td>Mozart</td>
<td></td>
<td>Belgium</td>
<td>in</td>
<td>3</td>
</tr>
<tr>
<td>Verdi</td>
<td></td>
<td>France</td>
<td>in</td>
<td>3</td>
</tr>
</tbody>
</table>

the Conceptual Domain Modeler proposes the associated *fact type*:

< Famous composer > visited < Country > in < Number of years > different years.

**NOTE:** Here one clearly sees the similarities between the fact instances and the fact type.
81. **Abstract from fact instances to fact types**

Given the previously introduced *facts*,

<table>
<thead>
<tr>
<th>Famous composer</th>
<th>Visited</th>
<th>Country</th>
<th>In</th>
<th>Number of years</th>
<th>Different years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mozart</td>
<td>visited</td>
<td>Italy</td>
<td>in</td>
<td>5</td>
<td>different years.</td>
</tr>
<tr>
<td>Mozart</td>
<td></td>
<td>Belgium</td>
<td></td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Verdi</td>
<td></td>
<td>France</td>
<td></td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

the Conceptual Domain Modeler proposes the associated *fact type*:

<**Famous composer**> visited <**Country**> in <**Number of years**> different years.
82. Abstract from fact instances to fact types

Given the previously introduced facts,

<table>
<thead>
<tr>
<th>Famous composer</th>
<th>Visited</th>
<th>Country</th>
<th>In</th>
<th>Number of years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mozart</td>
<td>visited</td>
<td>Italy</td>
<td>in</td>
<td></td>
</tr>
<tr>
<td>Mozart</td>
<td></td>
<td>Belgium</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>Verdi</td>
<td></td>
<td>France</td>
<td></td>
<td>3</td>
</tr>
</tbody>
</table>

the Conceptual Domain Modeler proposes the associated fact type:

< Famous composer > visited < Country > in < Number of years > different years.
83. Abstract from fact instances to fact types

Given the previously introduced facts,

<table>
<thead>
<tr>
<th>Composer</th>
<th>Visited</th>
<th>Country</th>
<th>In</th>
<th>Years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mozart</td>
<td>visited</td>
<td>Italy</td>
<td>in</td>
<td>5</td>
</tr>
<tr>
<td>Mozart</td>
<td>&quot;</td>
<td>Belgium</td>
<td>&quot;</td>
<td>3</td>
</tr>
<tr>
<td>Verdi</td>
<td>&quot;</td>
<td>France</td>
<td>&quot;</td>
<td>3</td>
</tr>
</tbody>
</table>

the Conceptual Domain Modeler proposes the associated fact type:

<Famous composer> visited <Country> in <Number of years> different years.

Variable names proposed by the Conceptual Domain Modeler to the Subject Matter Expert
84. Populate the fact type to verify the common understanding

The Conceptual Domain Modeler populates the fact type with the 3 fact instances which were verbalized previously:

<table>
<thead>
<tr>
<th>Famous composer</th>
<th>Country</th>
<th>Number of years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mozart</td>
<td>Italy</td>
<td>5</td>
</tr>
<tr>
<td>Mozart</td>
<td>Belgium</td>
<td>3</td>
</tr>
<tr>
<td>Verdi</td>
<td>France</td>
<td>3</td>
</tr>
</tbody>
</table>

Conceptual Domain Modeler:

“Do these 3 facts correspond in meaning with the original 3 facts?”
85. **Determine the integrity rules**

The next step in developing the Conceptual Domain Model is to determine the integrity rules, starting with the most essential one: **the uniqueness constraint**.

In a protocolled manner the Conceptual Domain Modeler and the Subject Matter Expert discuss the various combinations of facts.

The outcome of this dialog is the establishment of the uniqueness constraint (or constraints) for the fact type.
86. Determine the integrity rules

<table>
<thead>
<tr>
<th>Fact ID</th>
<th>Famous composer</th>
<th>Country</th>
<th>Number of years</th>
</tr>
</thead>
<tbody>
<tr>
<td>RF</td>
<td>Mozart</td>
<td>Italy</td>
<td>5</td>
</tr>
<tr>
<td>1</td>
<td>Mozart</td>
<td>Italy</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>Mozart</td>
<td>Germany</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>Beethoven</td>
<td>Italy</td>
<td>5</td>
</tr>
</tbody>
</table>

The Subject Matter Expert answers: “NO!”

Conceptual Domain Modeler:
“Can the Reference Fact (RF) and fact 1 occur at the same time?”
87. Determine the integrity rules

<table>
<thead>
<tr>
<th>Fact ID</th>
<th>Famous composer</th>
<th>Country</th>
<th>Number of years</th>
</tr>
</thead>
<tbody>
<tr>
<td>RF</td>
<td>Mozart</td>
<td>Italy</td>
<td>5</td>
</tr>
<tr>
<td>1</td>
<td>Mozart</td>
<td>Italy</td>
<td>6</td>
</tr>
</tbody>
</table>

Conclusion:

No duplicates are permitted for the combination of the variables (roles) “Famous composer” and “Country” in any population of this fact type.

In other words: the variables “Famous composer” and “Country” are the independent (determining) variables and the variable “Number of years” is the dependent (determined) variable.
88. Determine the integrity rules

<table>
<thead>
<tr>
<th>Fact ID</th>
<th>Famous composer</th>
<th>Country</th>
<th>Number of years</th>
</tr>
</thead>
<tbody>
<tr>
<td>RF</td>
<td>Mozart</td>
<td>Italy</td>
<td>5</td>
</tr>
<tr>
<td>1</td>
<td>Mozart</td>
<td>Italy</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>Mozart</td>
<td>Germany</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>Beethoven</td>
<td>Italy</td>
<td>5</td>
</tr>
</tbody>
</table>

Conceptual Domain Modeler:
“Can the Reference Fact (RF) and fact 2 occur at the same time?”

The Subject Matter Expert answers: “YES!”
89. Determine the integrity rules

<table>
<thead>
<tr>
<th>Fact ID</th>
<th>Famous composer</th>
<th>Country</th>
<th>Number of years</th>
</tr>
</thead>
<tbody>
<tr>
<td>RF</td>
<td>Mozart</td>
<td>Italy</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>Mozart</td>
<td>Germany</td>
<td>5</td>
</tr>
</tbody>
</table>

Conclusion:

No integrity rule needs to be added to allow for the combination of these facts.
Determine the integrity rules

<table>
<thead>
<tr>
<th>Fact ID</th>
<th>Famous composer</th>
<th>Country</th>
<th>Number of years</th>
</tr>
</thead>
<tbody>
<tr>
<td>RF</td>
<td>Mozart</td>
<td>Italy</td>
<td>5</td>
</tr>
<tr>
<td>1</td>
<td>Mozart</td>
<td>Italy</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>Mozart</td>
<td>Germany</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>Beethoven</td>
<td>Italy</td>
<td>5</td>
</tr>
</tbody>
</table>

Conceptual Domain Modeler:
“Can the Reference Fact (RF) and fact 3 occur at the same time?”

The Subject Matter Expert answers: “YES!”
91. **Determine the integrity rules**

<table>
<thead>
<tr>
<th>Fact ID</th>
<th>Famous composer</th>
<th>Country</th>
<th>Number of years</th>
</tr>
</thead>
<tbody>
<tr>
<td>RF</td>
<td>Mozart</td>
<td>Italy</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>Beethoven</td>
<td>Italy</td>
<td>5</td>
</tr>
</tbody>
</table>

**Conclusion:**

No integrity rule needs to be added to allow for the combination of these facts.
92. **Concept Definitions**

The function of the Concept Definitions in the Conceptual Domain Model is to define every term or group of terms:

1. That are used in expressing the fact instances;
2. For which the assumption is made that their meaning may not be fully clear to the intended audience.

Ultimately, the goal is to achieve complete disambiguation of the meaning of the terms used in the Conceptual Domain Model.
93. Concept Definitions for our Conceptual Domain Model

**Famous composer**
Definition: Person who writes music and is well-known for his work.

**Country**
Definition: Political division of a geographical entity, a sovereign territory, most commonly associated with the notions of state or nation and government.

**Number of years**
Definition: Mathematical sum of all the non-overlapping years.
For example, if Mr. Smith visited France once in 1881, twice in 1883, and finally stayed in France from 1886 through 1888, he visited France during 5 different years.

**Visit**
Definition: To be in a specific Country at a certain moment in time, where the Country is not the Country of birth.
94. **Noteworthy quotes (1)**

“We only communicate by referencing **shared concepts** with **agreed meanings** using **common symbols**. We can write these down and use them.”

(excerpt from email message from Cory Casanave, sent to OMG Semantic Information Modeling for Federation (SIMF) team, April 28, 2012)
95. **Noteworthy quotes (2)**

"The conceptual data model is a comprehensive and exhaustive specification of system level data."

96. Noteworthy quotes (3)

"Model driven S/W engineering is key to manage development costs."