Paper for Consideration by the S-100WG

Order of Records in ISO8211

Submitted by:	SevenCs GmbH – Holger Bothien
Executive Summary:	Introduction of general rules for the order of records in the ISO8211 encoding
Related document(s):	S-100 Part 10a

Introduction / Background

S-57 has specified the order of records in the ENC product specification.

> The order of data in each base or update cell file is described below:
> Data set file
> Data set general information record
> Data set geographic reference record (for EN application profile)
> Vector records
> Isolated nodes (SG3D)
> Isolated nodes (SG2D)
> Connected nodes
> Edges
> Feature records
> Meta features
> Geo features (ordered from slave to master)
> Collection features
>
> This order of records will enable the import software to check that the child record exists each time the parent record references it (i.e. it will already have read the child record so it will know if it exists or not).

It is important to understand the general rule described in the last paragraph:

Everything that is referenced must exist

For the deletion we have the following rule:

A referenced record must not be deleted

The means that the referencing record must be deleted first or the reference must be removed by a modification.

For the EN profile the given order is in line with the rule but for the ER profile (updates) the rule is broken. This is because objects that are deleted cannot be referenced in a subsequent record. (e.g., an edge that is deleted will be removed from the spatial list of a feature).

The problem can be solved in software, by collecting the objects to be deleted and delete them after all other changes are made. Nevertheless, the idea of providing the data to allow a 'one pass' loop is broken. Though tests (555a and 555b) in S-58 exists it was observed that not all producers comply with the rule even for base data.

In S-100 the situation is slightly more complicated due to the enhanced data model. But the intention is still to follow the rule.

Analysis/Discussion

Unfortunately, the S-100 standard is a little weak in describing how that can be achieved.

1. ***S-100 10a-3.1*** states:
   The grouping of records into ISO/IEC 8211 files is considered application specific and is, therefore, described in the relevant product specification.
   Leaving the grouping (order?) to the product specification means that for different products separate implementations must be made. But the standard should be implemented not the product specification. It seems to be necessary to give at least a general guidance on the record order to help implementors to write their software. This guidance must also cover base data sets as well as updates. Product specifications further limit the generic rule e.g. by only allowing certain record types but not changing the general order.

2. Three record types allow references to records of the same type: Information Type records, Composite Curve records, and Feature Type records.
   a. **Information Type records**
      In S-100 10a-5.3.1 states:
      Information types may reference other information types. For this encoding it is important that an information type record must be stored prior to any record that references this record.
      This is a sufficient rule for the order of Information Type records.
   b. **Composite Curve records**
      A similar statement exists in S-100 10a-5.8.1
      Composite curves can have other composite curves as components. In this case the record of the component must be stored prior to the record which references the component.
      This sufficient as well.
   c. **Feature Type records**
      In S-100 10a-5.10.1 only this paragraph is somehow related to the order:
      Note: When updating associations to other records, the other records must already exist in the target (base data or added by the appropriate update record).
      This is too generic in the scope of Feature type records.

3. When records are deleted by an update file the order should guarantee that every record that is going to be deleted is not referenced by any other record. This can be achieved by using the reverse order as for the insertion of records, e.g., parents will be deleted before children.

4. Modifying records are not critical in their order but for the sake of consistency the same order as for insertion should be used. Some modification will require multiple records for the same record with increasing record version numbers. It is important that the order of such records must be in the order of the record versions.

Summarizing the analysis:

- A general rule for the record order is missing
- A specific order that all products must follow is missing too
- Some rules for specific record types exist but not all of them are appropriate.
- Updates are not considered so far

Conclusions

The following proposals will solve the problem:

The sentence

The grouping of records into ISO/IEC 8211 files is considered application specific and is, therefore, described in the relevant product specification.

Should be removed from the clause S-100 10a-3.1

A new clause **Order of Records** should be inserted after **S-100 10a-3.6 Descriptive fields.**

It should read:

The order of records will enable the import software to check that a referenced record exists each time it is referenced.

Exists means either:

- The record is inserted in this data set file prior to the record that reference it
- Or it is inserted by the base data set file or an earlier update file and not deleted between the insertion and the record that reference it

---

In addition, when a record is going to be deleted it must not be referenced by any other record.

The record order is:

---

1. **Data Set General Information record**
2. **Data Set Coordinate Reference System record**
3. **Information Type records** (RUIN=Insert) (for the order inside this group see the encoding rules for Information Type records)
4. **Point records** (RUIN=Insert)
5. **Multi Point records** (RUIN=Insert)
6. **Curve records** (RUIN=Insert)
7. **Composite Curve** records (RUIN=Insert) (for the order inside this group see the encoding rules for Composite Curve records)
8. **Surface records** (RUIN=Insert)
9. **Feature Type records** (RUIN=Insert) (for the order inside this group see the encoding rules for Feature Type records)
10. **Information Type records** (RUIN=Modify)
11. **Point records** (RUIN=Modify)
12. **Multi Point records** (RUIN=Modify)
13. **Curve records** (RUIN=Modify)
14. **Composite Curve records** (RUIN=Modify)
15. **Surface records** (RUIN=Modify)
16. **Feature Type records** (RUIN=Modify)
17. **Feature Type records** (RUIN=Delete) (reverse order as for Insert)
18. **Surface records** (RUIN=Delete)
19. **Composite Curve records** (RUIN=Delete) (reverse order as for Insert)
20. **Curve records** (RUIN= Delete)
21. **Multi Point records** (RUIN= Delete)
22. **Point records** (RUIN= Delete)

> 23. **Information Type records** (RUIN= Delete) (reverse order as for Insert)

> Note that product specifications can omit entries they don't use but not change the order. They might further define a more specific order within each group if the general rule regarding references is not broken.
>
> If several records for the modification of one record are required in one dataset, they must be using increasing record version numbers and the order must be according to that numbers.

The paragraph:

Note: When updating associations to other records, the other records must already exist in the target (base data or added by the appropriate update record).

should be removed from **S-100 10a-5.10.1** and the following paragraph should be inserted:

Feature types may reference other Feature types. For this encoding it is important that a Feature Type record must be stored prior to any record that references this record.

Recommendations

To increase clarity and consistency over several product specifications it is recommended that the S-100 Part 10a is amended as proposed.

Action Required by the S-100WG

The S-100WG is invited to:

     a.   Note this paper
     b.   Discuss the paper
     c.   Revise the proposals where necessary
     d.   Endorse the proposals