TSMAD26/DIPWG5-11.7F

Paper for consideration by TSMAD

Roles in S100

Submitted by:	Jeppesen / SNPWG
Executive Summary:	This paper discusses associations and the use of names for association ends (role names) in S-100, provides examples, and proposes to simplify product specifications by providing for the systematic use of default names.
Related Documents:	(1) S-100 Ed. 1.0.0
Related Projects:	(1) S-100

1 Introduction / Background

At TSMAD 25 it was noted that current S-100 prescriptions for roles have the potential to increase the size of the feature catalogue and potentially the size of datasets. Paper TSMAD25-4.3.2 proposed making roles optional and using “source” and “target” as default roles. Both were agreed at TSMAD 25.

S-100 and ISO 19103 include certain conventions for default role names. Also, ISO 19109 does not distinguish between feature and information types but treats both as feature types. S-100 replaces the feature association of ISO 19109 by “feature associations” and “information associations.” Feature associations are required to have roles specified at both association ends and information associations to specify a role at one end and hold the role at one association end fixed.

This paper attempts to clarify the resulting situation concerning roles in S-100. It provides examples of the use of roles. It combines the S-100 and ISO conventions with the TSMAD 25 decisions into a system of defaults for the use of roles in product specifications.

2 Terms and Abbreviations

FAL	Convention on Facilitation of International Maritime Traffic
GFM	General Feature Model (S-100 or ISO 19109)
NPUBS	Nautical Publications
XML	Extensible Markup Language
XSLT	Extensible Stylesheet Language Transforms

3 References

ISO 19103: Geographic Information – Conceptual Schema Language. (2002).

ISO 19109: Geographic Information – Rules for Application Schema. (2005).

ISO 19501: ISO/IEC 19501:2005. Information technology — Open Distributed Processing — Unified Modeling Language (UML) Version 1.4.2. (2005).

ISO 19505-1: Information technology — Object Management Group Unified Modeling Language (OMG UML) — Part 1: Infrastructure. (2012).

ISO 19505-2: Information technology — Object Management Group Unified Modeling Language (OMG UML) — Part 2: Superstructure. (2012).

S-100: Universal Hydrographic Data Model, Edition 1.0.0, January 2010.

S-101: IHO Electronic Navigational Chart Product Specification – Phase 4, June 2012.

S-101-DCEG: S-101 Data Classification and Encoding Guide, working version, February 2013. URL: http://www.iho.int/mtg_docs/com_wg/TSMAD/TSMAD26/DCEG/S101_Data_ClassificationAndEncodingGuide_Working_SubWG.pdf (retrieved 22-04-2013).

TSMAD25-4.3.2: Revisions and Extensions to S-100 Edition 1.0.0, 2013. Paper at TSMAD 25, Jan. 2013. URL: http://www.iho.int/mtg_docs/com_wg/TSMAD/TSMAD25/TSMAD25-4.3.2_S-100_GapsandExtensions-Final.pdf (22 Jan. 2013).

4 Discussion

4.1.1 Roles names in UML, ISO standards, and S-100

The version of UML referenced by ISO 19103 (UML 1.4.2 - ISO 19501:2005) states that the rolename is optional but not suppressible (§ 5.43.2.6). The newer version of UML (version 2.4.1, ISO 19505-1:2012 & 19505-2:2012) provides that the name of an association end (i.e., the role name) is both optional and suppressible (§ 7.3.3).

ISO 19103 (§ 6.7) states:

> *All associations shall have cardinalities defined for both association ends. At least one role name shall be defined. If only one role name is defined, the other will by default be inv_rolename.*
>
> *If an association is navigable in a particular direction, the model shall supply a "role name" that is appropriate for the role of the target object in relation to the source object. Thus in a two-way association, two role names will be supplied. The default role name is "the<target class name>" in which the target class is referenced from the source class (this is the default name in many UML tools). Association names are principally for documentation purposes.*

Part 1 of S-100 (§ 1-4.7.2) also states "if only one role name is defined the other role will by default be inv_rolename."

4.1.2 Practical considerations

The typing of associations as feature associations or information associations in S-100 and its encodings implicitly carries the semantics of specifying whether the link is to an information data object or a feature data object.

In many contexts, explicitly providing role names for both association ends contributes nothing to explaining the domain model.

- The name at one end is often no more than the inverse of the name at the other and basically linguistic candy compared to "inv_roleName" (e.g., compare "marks" and "is marked by" to "marks" and "inv_marks").
- Requiring a name at each end of each association adds clutter and complicates layout in UML diagramming tools.
- In the scope of a domain (as the term is used by the registry) the semantics of the relationship between two classes is often unambiguous from the identities of the classes and a convention for default name, i.e., "the<target class name>" suffices.
- Requiring a role code to be included in the encoding for every association increases the size of data sets and may contribute nothing to the semantics.

On the other hand, if there are multiple associations between two classes either association labels or role names are needed to distinguish associations. Since S-100 does not use association labels that leaves names. Here again naming one role of the two roles often suffices.

It is possible that an encoding format may use role names or a corresponding role code as in the ISO 8211 encoding in Part 10a[1]. Others may not. In the first iteration of the Marine Protected Area (MPA) test dataset developed by SNPWG role names are not encoded in the XML datasets, nor were they specified in the mappings of source data to the application model, yet it was possible to both understand the data and generate sample text output from it. On the other hand, other encoding formats may need a role name at each association end.

[1] The ISO 8211 encoding in Part 10a provides subfields role code (RLCD) and Association Code (ASCD) for feature association fields (FEAS), and subfield information role (IROL) for information association fields (INAS). The latter provides for a default value of 65535 if not available.

4.1.3 Navigability and roles

UML 1.4.2 provides “An arrow may be attached to the end of the path to indicate that navigation is supported toward the classifier attached to the arrow.” Arrows may be shown whenever navigation is supported in a given direction or suppressed in general with only exceptional situations being shown.

In UML 2.4.1 a navigability arrow is indicative rather than prescriptive. A navigability arrow indicates whether instances at one end can be accessed efficiently from instances at the other ends of the association. Further “If an end is not navigable, access from the other ends may or may not be possible, and if it is, it might not be efficient. Note that tools operating on UML models are not prevented from navigating associations from non-navigable ends.”

Bi-directional navigability is useful in some encodings. XSLT has the notion of the ‘current node’ which sets the context for template matching and execution. If there is a link encoded only in one direction in an XML dataset, getting information from inside the source object while the target object is the current node may be inefficient and/or complicate the code[2]. For example, if a MarineProtectedArea feature has a pointer to an Authority object but not vice versa, extracting information about “which authority administers this MPA” is efficient but answering the question “which MPAs does this authority administer?” is inefficient.

Whether and how this affects roles depends on whether role names are required for encoding associations. If the encoding format requires role names, a convention for role names (“inv_rolename” or “the<target class name>” as in ISO 19103) suffices from the technical point of view. Individual product specifications can define conventions for roles and codes, as needed by encoding and delivery modes.

The question of roles and role names should ideally be separate from the navigability of association ends.

5 Examples

Example 1 - Default role names

No name for an association end need be mentioned if the semantics of the application schema are such that the relationship is obvious without using an explicit name. This can be so if there is only one relationship possible in the context of the product specification, or there is a relationship which is overwhelmingly likely. For example, in the association between an Authority and a ContactDetails object in the NPUBS domain, the obvious relationship is that the ContactDetails are the contact details for the authority. The default association end names “theContactDetails” and “theAuthority” suffice and need not be explicitly shown. Defaults are also suitable for the Authority/ServiceHours association.

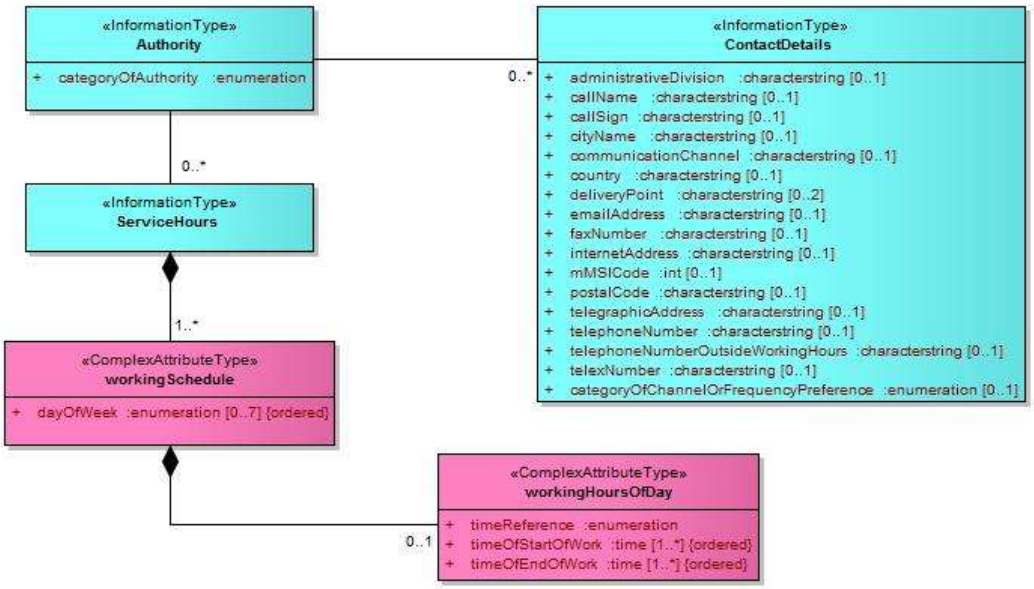

Figure 1. Associations without explicit role names

[2] Obviously this depends on what type of output is being produced – a graphical output where the primary output is geometry which has attached notes, or a text organized by categories.

Navigability on the association should either be unspecified or the UML 2.4.1 convention applied, about navigability being indicative of efficiency rather than prescriptive.

There may be situations where the semantics of the association are not clear. In such cases at least one association end should be named. In general the rule of thumb should be “if in doubt whether an explicit role name is needed, provide one”.

Example 2: Role name suppressed at one association end

Where an explicit role name is considered to clarify or disambiguate the relationship, it can be provided and the inverse role either explicitly named or the default name inv_<roleName> used. The default can be suppressed in the diagram.

Figure 2. Association with role name at one end

Example 3: Role names at both association ends

The current S-101 Data Classification and Encoding Guide [S-101-DCEG] uses master/slave relationships for aids to navigation consisting of structure and equipment objects (see Annex A of this document). While the S-101 application schema is not available, a master/slave or structure/equipment association (i.e., role names like “master” and “slave,” or “structure” and “equipment”) is reasonable considering S-100 § 5-4.2.4 which used master/slave as an example of an association with two roles (and even if S-101 does it differently other data products may have similar circumstances). Further, since some feature classes may play the role of either master or slave role names at both ends of the association are required at least for model clarity though strictly speaking the “inv_rolename” convention works here too. The figure below shows a **hypothetical** model.

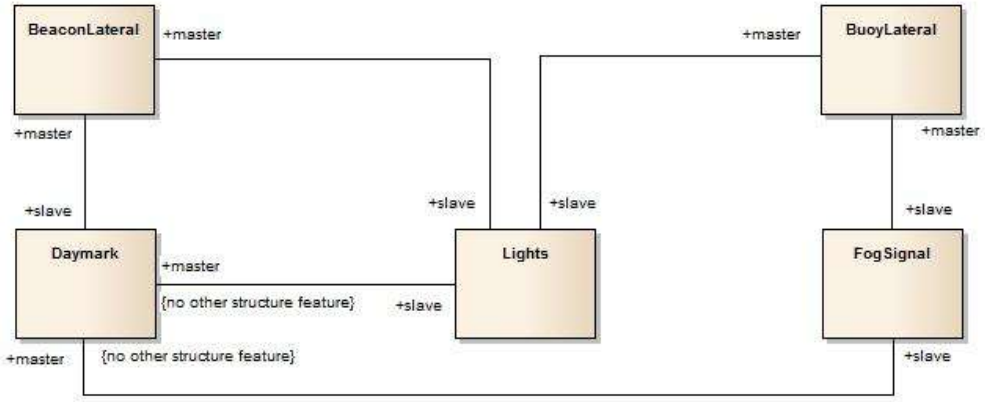

Figure 3. Master/slave relationships for navigation aids in UML

Example 4: Role names for disambiguation

One example of ambiguity is a situation in which an application schema has more than one association relationship between two classes; in that case at least one role must be given for each association, so the two can be distinguished. The figure below shows the possible relationships between information types Regulations, Restrictions, Recommendations, and NauticalInformation on the one hand and subsets of vessels, as defined by vessel dimensions, cargo and performance characteristics in a class named Applicability. Leaving out the names *excludesVessels* and *includesVessels* would make the associations of the Applicability class ambiguous.

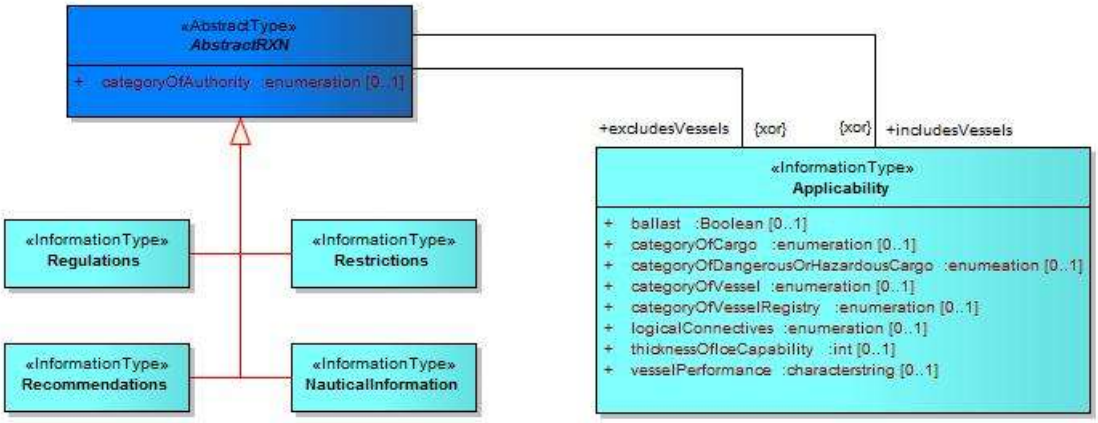

Figure 4. Multiple associations between a pair of classifiers

The relationships are inherited from a class named AbstractRXN which generalizes the Regulations, Restrictions, Regulations, and NauticalInformation classes.

This approach grows increasingly unwieldy as the number of associations between the two classes increases. This is particularly true if there is more than one kind of relationship. It fails if the association is characterized by an attribute of type other than enumeration (e.g. a real number).

Example 5: Association classes

A better way of modeling the case in Example 3 is to use an association class. In the figure below the association class InclusionType models the relationship between Applicability and Regulations, etc. (again by inheritance from class AbstraxctRXN).

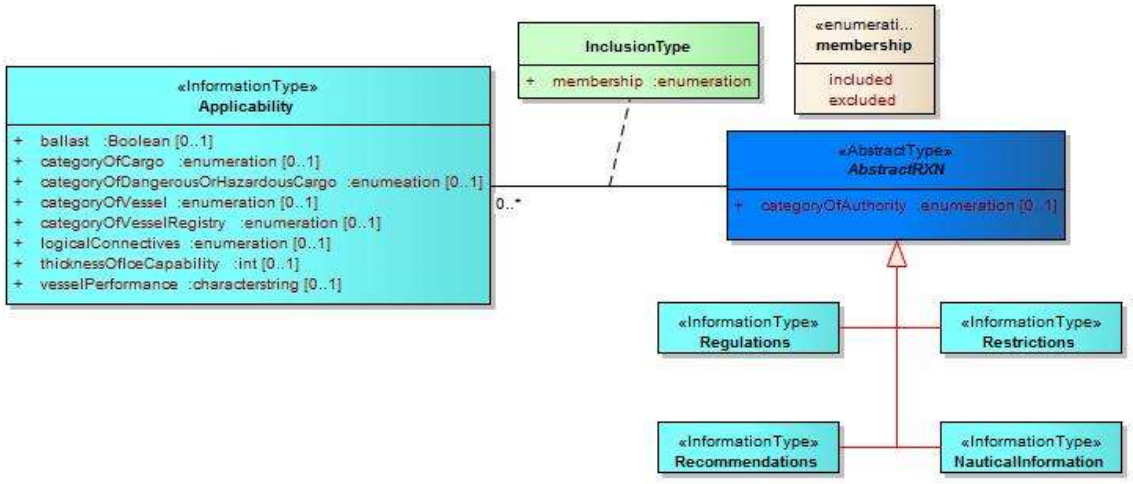

Figure 5. Association class

Explicit or default names can be used on the association ends as needed. The attributes of an association class should be limited to those that characterize the relationship. Association classes can be used even if the relationship is characterized by an attribute which is not of an enumeration type (e.g., a real number).

Example 6: Navigability and roles in information associations

Navigability should be indicated when it expresses relevant domain model semantics. The figure below shows a hypothetical partial application schema modeling FAL information, e.g., for use in generating IMO standard forms or alternate reporting formats. This demonstrates a situation where specifying navigability may be appropriate, in that (application software) navigation from a data object representing a Port to the data object representing arrival information for the *previous* Port of Call may not be possible (e.g., ports being in different countries and therefore arrival notices being in different databases, or because of considerations of privacy, commercial secrets, or national security).

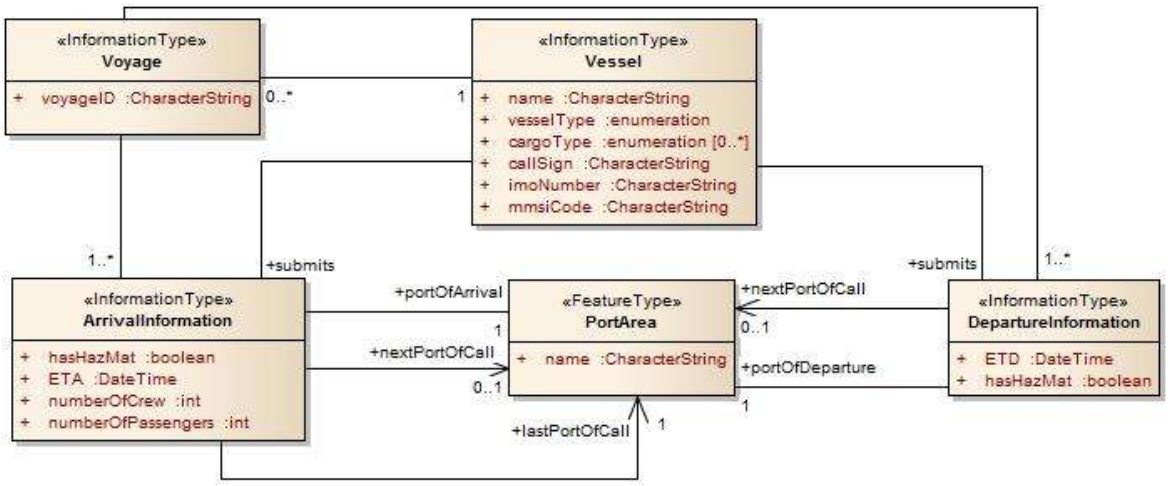

Figure 6. Application schema for FAL information

This example also shows a situation where the 'wrong' end of an information association has a role name, i.e., the "feature" end of a feature/information association. In general which end should receive a name, or whether both ends should, would depend on the perspective of the application domain – an application schema intended for an application domain utilizing graphical chart- or map-like display may be communicative enough naming the information ends of feature/information associations, but an application schema for domains dealing primarily in streamed information or text-based presentations may need to switch the focus to the information rather than the geographic feature and prefer to use role names at the other end.

6 Recommendations

The following specific recommendations are made.

- Product specifications should be allowed to use default role names if the application domain's semantics permit. The ISO 19103 conventions for default role names should be adopted.
- Since the ISO conventions result in individual role names for each and every association end, the defaults "source" and "target" can also be added as approved by TSMAD 25. The default "both" is replaced by "associatedWith", the latter being more meaningful.
- Information associations should permit navigability in both directions and naming of both association ends. The role "additionalInformation" should be a default for associations to information types, instead than fixed (Clause 3-5.4.4, Edition 1.0.0). Retaining it maintains compatibility with Edition 1.0.0. Product specifications can adopt more restrictive conventions, e.g., S-101 can fix the role at one end of an information association, but other product specifications need more flexibility.

The effect is to suppress role names which do not contribute to the semantics or clarity of an UML diagram but keep names available in the form of defaults, should they be needed for encoding format specifications or implementations. Information associations are made more flexible.

Changes to S-100 including the draft text for a clause describing default roles are redlined in the accompanying markup of Part 3.

7 Justification and Impacts

Justification: The changes to the GFM were needed to increase flexibility for modeling especially for nautical publications and domains other than ENCs, simplify application schema design, reduce data storage requirements, and bring S-100 into closer alignment to ISO 191xx standards. Failure to include them will make the development of application schemas more difficult and increase data volumes.

Impacts: Impacts on data products, production, and applications are minimal since S-100 data products are not yet being produced. Part 5 of S-100 Edition 1.0.0 and the FC XML schema will have to be updated (updates to Part 5 are proposed separately). Product specifications should describe any conventions they use for roles. For example, S-101 may state "the role additionalInformation is fixed for

associations to information types and means that additional information is available for a named type" (though this would actually be redundant since the modified § 3-5.4.4 states the role additionalInformation is the default for these associations).

8 Actions Requested

TSMAD is invited to:

- endorse the recommended changes to S-100

Annex A. Extracts from working draft of S-101 DCEG

Extracts from § 17.1 Geo features forming parts of navigational aids

Aids to navigation are composed of fixed or floating structure features carrying equipment features.

The most common structure features are: **Beacon Cardinal, Beacon Isolated Danger, Beacon Lateral, Beacon Safe Water, Beacon Special Purpose/General, Buoy Cardinal, Buoy Installation, Buoy Isolated Danger, Buoy Lateral, Buoy Safe Water, Buoy Special Purpose/General, Bridge, Building, Crane, Daymark, Floating Dock, Fortified Structure, Fishing Facility, Hulk, Light Float, Light Vessel, Landmark, Mooring/Warping Facility, Offshore Platform, Pile, Pontoon, Pylon/Bridge Support, Obstruction, Shoreline Construction, Wreck**.

Equipment features consist of: **Daymark, Fog Signal, Light, Radar Station, Radio Station, Retroreflector, Radar Transponder Beacon, Signal Station Traffic, Signal Station Warning**.

Extracts from § 17.2 Relationships

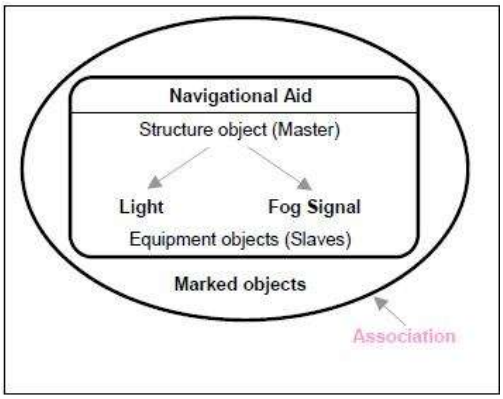

Figure 7. Master/slave relationship

When the navigational aid contains a structure feature (from the list at clause 17.1), this feature must be the master feature, and the equipment features must be the slaves. Note that **Daymark** may be a master feature or a slave feature; where a navigational aid contains a **Daymark** and there is no other base structure (which can serve as the master feature) indicated on the source, the **Daymark** feature should be encoded as the master feature.

When the nature of the base structure on land is unknown or there is no structure feature, one of the equipment features may be chosen as the master feature, giving priority to a **Light** feature, if one exists. Alternatively, a **Pile** feature of type point or a **Beacon Special Purpose/General** feature may be encoded as the structure feature at the same position as the equipment features. When the nature of the base structure in the water is unknown, an ECDIS Base Display feature (see S-52, Annex A, clause 13.2), e.g. **Pile** feature of type point or a **Beacon Special Purpose/General** feature, must be encoded as the structure feature at the same position as the equipment features.

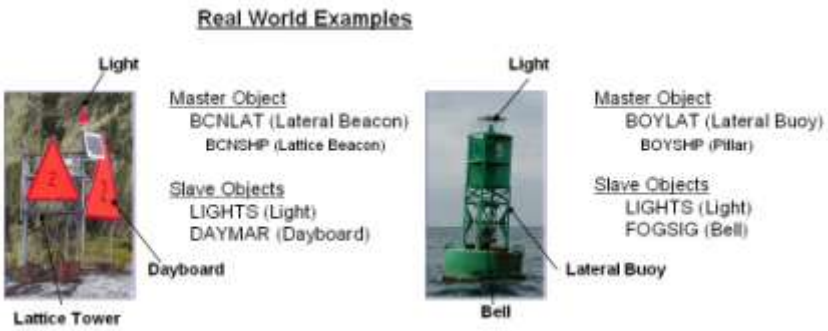

Figure 8. Real world example of master/slave relationship