OASIS SET TC
Discussion on Dale Moberg’s Comments
DM: GS1, UBL, and OAGIS documents will all eventually be semantically harmonized with the CCTS information model and with the CC Library.

DM: Harmonization means that document structure will be mapped into CCTS information model patterns and that semantic primitives for messages will be found in the CC Library.

Discussion Questions:
- OASIS SET TC gives all these correspondences in an automated way. Why not use SET Harmonized Ontology for semantic harmonization of CCTS?
- Note however that there will always be irreconcilable differences between their schemas. To what extent can harmonization be realized without changing the GS1, UBL and OAGIS schemas?
- Would these standards consider to become more CCTS conformant in the future? What will happen to backward compatibility of their schemas in such a case?
- Further comments?
DM:
- SET has proposed an “upper ontology” of OWL DL assertions, with content taken from the CCTS model of business information.
- The upper ontology is combined with additional assertions describing information structures from messages that are defined by the message standards bodies (GS1, UBL, and OAGIS).
- The ontologies are combined and “classified” to produce a completed set of asserted and inferred OWL claims, containing many class equivalences.
- These equivalences form the basis for maps between information elements within the overall documents.
- The resulting maps are not normally complete, and “heuristic” rules are used to derive additional class equivalences until all “corresponding” parts of documents are connected.
- For a given pair of documents, it may still be that some parts in either document have no parts in the other; each document may, for example, make use of some semantic primitives not found in the other.
Given that GS1, UBL and OAGIS are seeking to harmonize against the CC Library, there should be a number of correlations based not on the CCTS models, but on the low level maps of document parts to core components. David Webber leveraged these “couplets” to discover equivalences through the UID dictionary cross-references which he then adds to his transforms or CAM templates. In the METU approach, UBL CityName.Name correlates with CCL CityName.Text which correlates with GS1 city.Text. These correlations are derived using an additional equivalence, Name.Type = Text.Type that allows some surface conventions to be abstracted away. My question is whether the need for such rules might be diminished by using a more direct correlation stemming from the UIDs being equal?

- UID Matching is lexical
- SET TC Matching uses semantics
- Further comments?
Comment 3

- More generally, the framework provided in the current draft seems to combine a very constrained DL approach to knowledge representation and inference, with several escape mechanisms to maneuver around various obstacles tied to both the ontology’s limited content (insufficient to derive various needed equivalences) and the inference engines constraints on inferences.

- The need for Jess based augmentation of inferences and special rules that have sufficient content to derive additional equivalences seems to be “a patch” for something that points to an inadequacy in the CCTS model, the CC Libraries (which may not be adequately leveraged), or perhaps in the constraints on inference rules. Maybe what we are seeing is that there is a need for a richer ontology in CCTS? Or more information in the CC library?

- More research may be necessary in this area. However, if the document standards base their components on the components from the CC Library, it will help a lot with the interoperability?

- Further comments?
If we are to add more general rules and inferences to overcome OWL limitations, I think SWRL could be considered as at least connected to the OWL technology in some ways. In addition, the Pellet reasoner is said to support SWRL reasoning, so that demonstrations could still be run within Protégé, if that is the preferred development environment. SWLR can be stored in an OWL file, apparently, and then the additional equivalences would come out of OWL syntax and DL reasoners with greater inferential powers.

Agree. However in our trials with Pellet, we get “out of memory” errors

Further comments?
However, the fact that these additional rules have not come from an analysis of the content of the CCTS models, the specifics about GS1 or UBL or OAGIS information structures, or the CC library suggests that the project falls short as a proof of concept for UN/CEFACT technologies in producing maps.

At the very least, the precise details need to be captured with more care about just what can be derived (without the heuristic patches) and what those equivalences or inclusions can accomplish with respect to generating maps between documents.

We captured all the semantics of GS1, UBL and OAGIS that is conformant to CCTS.

The specifics of GS1, UBL and OAGIS are there and cannot be ignored when mapping them one another.

More insight into the process can be gained through more use cases.

Further comments?