# Knowledge Base Version Reintegration

**Landon T. Detwiler, MSc, Cornelius Rosse, MD, DSc**
**Structural Informatics Group,**
**Departments of Biological Structure and Medical Education and Biomedical Informatics,**
**University of Washington, Seattle, WA 98195**

*Given two versions of a knowledge base (KB), independently modified, we investigated the problem of incorporating changes made to one KB version into the other. We have implemented a system that will perform such a reintegration, autonomously, using predetermined user preferences. This effort has lead to a greater insight into the version reintegration problem and has highlighted those areas where user intervention would be the most beneficial in a semi-autonomous system.*

The Foundational Model of Anatomy (FMA) is a reference domain ontology which represents the structural phenotype of the human body[1]. As part of a multidisciplinary collaborative effort, the DARPA Virtual Soldier Project (VSP), we are augmenting this model with the addition of physiologic and pathologic information. This combined knowledge source is known as the Virtual Soldier Knowledge Base (VSKB). Such augmentations are performed in parallel on copies of the FMA while the master copy continues to evolve. The motivation for this work arises from the need to reintegrate the content of the evolving master copy into the VSKB.

The reintegration scenario is illustrated in Figure 1 where the original KB A is copied and that copy is modified to become B while the original is independently modified and becomes A'. We wish to form some B' which is B augmented with the changes that have been made to A. This process is then repeated and a new A'' and B'' evolve from A' and B' respectively and are reintegrated to form B''', and so on.

The prototype system, referred to here as "KB Reintegrator", was constructed specifically for reintegrating knowledge bases developed in the Protégé frame-based knowledge modeling system. Although we designed KB Reintegrator for merging versions of any Protégé KB, we do require that the KB meet certain preconditions, like having frame identifiers that are unique and immutable across KB versions. We also require a snapshot of the original knowledge base from the time of the version fork.

We chose to create a system that requires 3 ontologies to perform a reintegration. The source KB contains the insertions/updates/deletions that we wish to propagate, the target KB is the knowledge base that we wish to update, and the original KB is a snapshot of the KB from which the source and target originated. Our reasoning for this choice is illustrated by the following example: Given two aligned frames, one in the source KB, one in the target KB, if the target frame contains an attribute that the source does not, what should we do? It could be that this attribute once existed in the source, but has been deleted, in which case we may wish to propagate this deletion to the target. Or, it could be that this value has simply been added to the target, in which case we probably want to keep it. This ambiguity can be resolved by referring to the original knowledge base.

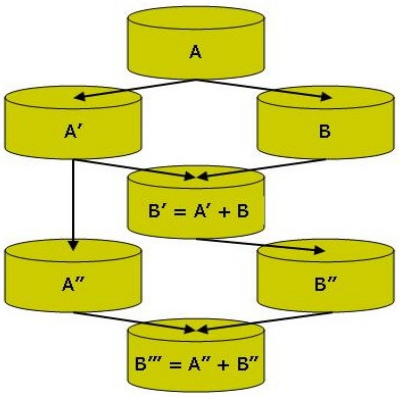


**Figure 1: Reintegration Scenario**

While designing and implementing KB Reintegrator, we encountered, learned from, and came up with reasonable solutions to many version reintegration issues similar to the one mentioned above. In the end we found that it was possible to build an autonomous system that made broad user preference assumptions and did the right thing most of the time. In the future we hope to improve the system by incorporating user feedback at critical decision points.

**Acknowledgements**

This work was supported by the DARPA Virtual Soldier Project.

**References**

1. Rosse, C., Mejino, J. L. V. Jr. A Reference Ontology for Bioinformatics: The Foundational Model of Anatomy. *J Biomed Informatics*, 2003; 36:478-500.

FMA web: http://fma.biostr.washington.edu