TOWARD a COMPLETE PHYLOREFERENCING LANGUAGE

T. Michael Keesey
keesey@gmail.com

A phyloreference is a statement indicating a taxon within a phylogenetic context. A common use for phyloreferences is in phylogenetic definitions, which tie taxonomic names to taxa via such statements. Several conventions for writing phyloreferences have been proposed, but most only cover a few “standard” forms (node-, branch-, and perhaps apomorphy-based clades) without the capacity to represent more “exotic” forms (e.g., ancestor-based clades and qualified/modified references). In order to build a complete phyloreferencing language, the mathematical underpinnings of phylogenetic contexts must be clarified. A phylogenetic context may be modeled as a directed, acyclic graph, wherein nodes model taxonomic units and directed edges model immediate descent. Higher taxa are modeled as unions of nodes. A phyloreferencing language must minimally allow for certain classes of entity: Boolean values, sets (including taxa, relations, and the empty set), and lists (including graphs and functions). It must also minimally allow for basic operations related to logic, set theory, and graph theory. Higher structures such as declarations and piecewise constructs must also be possible. With these as a basis, functions related to phylogeny can be defined: maximal, minimal, predecessor union/intersection, successor union/intersection, exclusive predecessors, synapomorphic predecessors, clade, crown clade, and total clade. I show how such a language may be used to represent various types of phyloreference, both “standard” and “exotic”.