Understanding 3D File Format Conversions

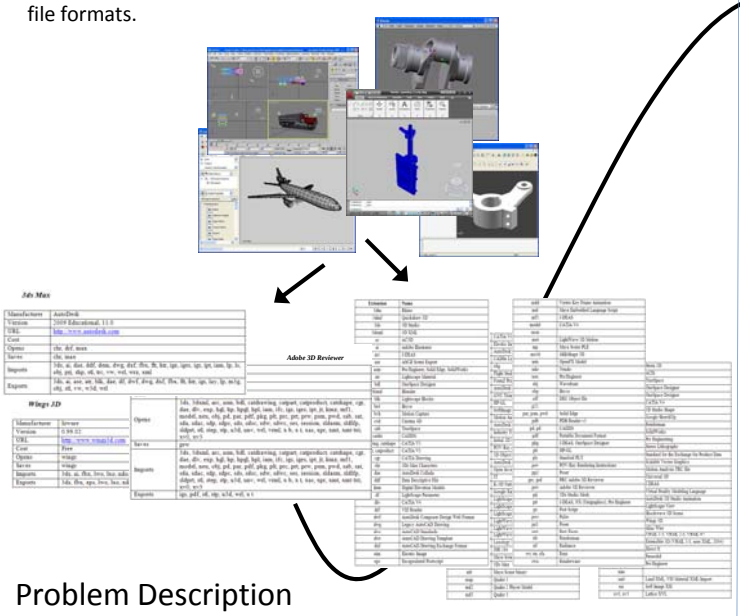
Kenton McHenry and Peter Bajcsy
National Center for Supercomputing Applications (NCSA), University of Illinois at Urbana-Champaign (UIUC)
{kmchenry, pbajcsy}@ncsa.uiuc.edu

Abstract

We have designed a framework for measuring information loss as one converts from one 3D file format to another. Different formats store different pieces of information, thus moving from one format to another sometimes involves dropping information and/or converting the 3D content itself. In addition, software vendors have unique implementations of file importers/exporters which must be considered when converting between formats. In light of this, the information loss obtained from a format conversion becomes a function of both the specification of the file formats involved and the software used to perform the conversion. We introduce a directed graph of conversions called an I/O-graph. This data structure captures the inputs and outputs of various input/output operations (e.g. the import/export operations of 3D software packages). We use the IO-graph to drive an extensible conversion system, NCSA Polyglot, which based on the data stored in the graph calls the relevant third party packages in order to perform a conversion from a source format to a target format. Given a set of 3D files we are then able to assign numerical values to the IO-graphs conversion paths by comparing the original and resulting content.

Introduction

When a software vendor creates a new application for working with 3D models it often creates its own file format to store that 3D content as well. The result of this practice is a world in which 3D content is difficult to share due to the fact that it is dispersed among hundreds of different file formats.



Problem Description

•Motivation
- Many formats are proprietary/closed, change over time and/or are no longer supported.
- There is a need to convert files to an open/standardized format that is likely to have long life and would preserve the content.

•Challenges
- Different formats store different 3D information and/or use different information representations (e.g. facet, CSG, B-Rep).
- Moving to another format can result in omitting information that is not supported or approximating information that is stored in a different way (e.g. tessellation).
- Software applications support only subsets of all possible file format conversions and have unique implementations of file importers/exporters (i.e. may not support the exact format specification).

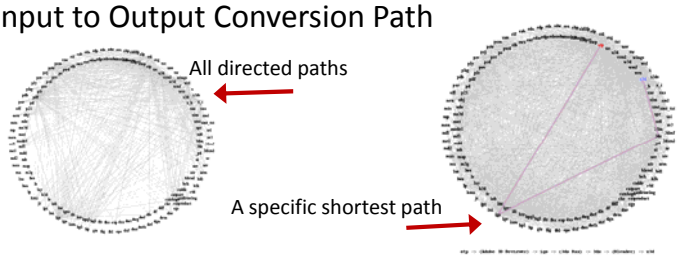
Approach

- Combine available 3rd party applications to perform conversions over union of supported formats.
- Compare 3D content before and after conversions to estimate information loss.

For more information: URL: http://isda.ncsa.uiuc.edu/NARA/

Acknowledgments
This research was partially supported by a National Archive and Records Administration supplement to NSF PACI cooperative agreement CA #SCI-9619019.

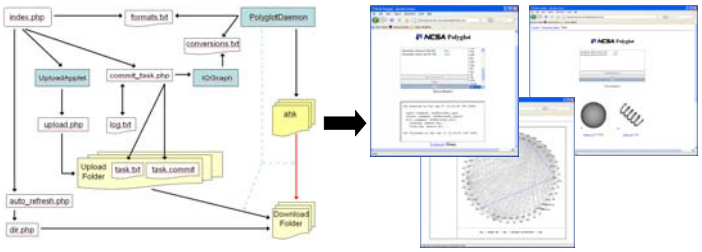
Input to Output Conversion Path



•Input to output conversion paths (a chain of imports/exports of various applications are captured by an Input/Output directed graph (I/O-Graph).
•I/O-Graph allows for the chaining of conversions to accommodate a source and target format not directly supported by any one application.
•If multiple paths between formats choose shortest path (i.e. with the fewest involved software applications).

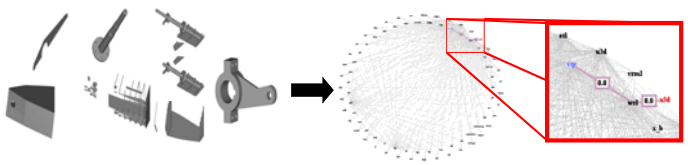
3D File Format Conversions using NCSA Polyglot

•Polyglot uses I/O-Graph and AutoHotKey scripts to perform desired conversions and a simple web-based user interface to interact with the I/O-Graph and 3D visualization, and for uploading and downloading 3D files.



Measuring Information Loss

•Convert a set of files across many paths in the I/O-graph
•Compare 3D content in before/after files via user selected measures
•Add information retention as edge weights in I/O-Graph

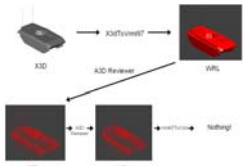


Experimental 3D Test Files

Edge Weights Based on Information Loss Measurements

Conclusions

Polyglot allows us to select
•conversion paths based on weights (e.g. Dijkstra's graph search algorithm)
•file format based on high information retention from many other formats



Simple Example of 100% Information Loss