ANVIL: Annotation of Video and Language Data 5.0

By Michal Kipp, DFKI

Reviewed by Ning Tan & Jean-Claude Martin, LIMSI-CNRS

1. OVERVIEW. Anvil is a tool for the annotation of video data. The aim of Anvil is to support the manual encoding of phenomena that occur in the video on multiple layers throughout a timeline. The official web site notes that while it was “originally developed for gesture research in 2000, Anvil is now being used in many research areas including human-computer interaction, linguistics, ethology, anthropology, psychotherapy, embodied agents, computer animation, and oceanography” (Kipp 2011).

![Anvil tool interface]

**Figure 1:** An example of the annotation of posture using the Anvil tool (Clavel et al. 2009; Tan et al. 2010)

2. GETTING STARTED. The Anvil user usually goes through the following steps when annotating videos. Some steps might not be required in some cases. For example, the user might be satisfied with a specification file (also called coding scheme) that has been written and provided by someone else.
2.1 SPECIFICATION. The user has to define a coding scheme before starting the annotations. The user directly creates this scheme within an XML editor or with the help of a graphical user interface provided by Anvil (this is an update from the previous version). The structure of the coding scheme is based on the definitions of tracks, attributes, and values. In Anvil, a track (also called tier, layer, or level in other tools) is a container for one type of information (e.g., gesture phase). Each track contains a number of elements that can be edited and filled with attribute-value pairs. Annotations within a single track may never overlap. Tracks can be of three types, called primary, singleton, and span. Anvil 5.0 adds a track relationship that enables the user to encode the larger entity first (e.g., a gesture) and then subdivide this entity (e.g., into gesture phases) on a separate “subdivision” track. Tracks can be combined into groups according to their semantic properties (e.g., linguistic tracks or gesture-related tracks).

The user can insert coding instructions and documentation at the specification stage so that coders can access them during coding. Each track attribute that is documented will have a small question mark next to it; clicking on the question mark will bring up the documentation for attributes and corresponding labels. The user can also let Anvil generate a suite of HTML files of the specification instructions using the pull-down menu Edit > Create Coding Manual. Any browser can then be used to view these pages.

2.2 DATA IMPORT. Anvil allows importing annotation files that have been produced with PRAAT (TextGrid file) or in ELAN (another annotation tool) so that the user can view, modify, and analyze these data in Anvil.

ANVIL also supports data produced in a motion capture (mocap) lab in the form of BVH files. It allows them to be opened and viewed as 3D motion capture data by reconstructing a 3D skeleton with the usual 3D camera controls. It also provides features for setting video and motion capture in synchrony. Mocap allows visualization of velocity using circles that are located orthogonal to the motion vector and show the peak of the motion phase. Optional features for mocap visualization allow the user to specify information about the mocap files, for example, which joints correspond to hips, wrists (left/right), or thorax.

2.3 ANNOTATION. Using the pop-up menus, one can define new track elements. When coding, the user can right-click the mouse over the Annotation Board to select options like start annotation and end annotation. For annotating single time points, Anvil offers the new primary point track type, which is complimentary to the usual interval sequences. Several features make the annotation process more efficient, such as keyboard short-cuts: F1 posi-
tions the green line, F2 creates an empty element; F3 creates element plus input dialog. Pressing the SPACE bar starts and stops the video. The user can also add bookmarks to noteworthy positions in the video by choosing Bookmarks > add bookmark. The bookmark is then signaled with a small yellow triangle on the top timeline.

2.4 ANALYSES. Anvil is capable of various kinds of statistical analyses for the verification of hypotheses and intercoder reliability studies:

a) **Frequency distributions of labels** for a single track can be viewed by right-clicking the track and choosing Show Histograms (figure 3). This also works for multiple annotation files, which can be accessed using the Project Tool (under Tools in the main menu). In managing projects, the user can browse, search, and analyze a sequence of several annotation files, provided that they point to the same specification.

![Figure 3: Track analysis: histogram of the values assigned to an attribute](image3)

b) **The Transition diagram** (Figure 4) analyzes the temporal relationships between different attributes within the same track in terms of estimated transition probabilities.

![Figure 4: Transition diagram: transitions between the values assigned to an attribute](image4)
c) **Attribute value association analysis** (chi-square and/or entropy-based) provides the analysis of temporal relations between two attributes that belong to different tracks. The selected attribute must be a ValueSet attribute. The type of relations is defined based on the timeline: equals, before, meets, overlaps, starts, finishes, or during. Also, the user can set a Tolerance value (e.g., 0.5 second) to extend the association’s probabilities. The association analysis results will be represented by tables of mutual information, difference matrix, expected values, and contingency table (Figure 5).

![Figure 5: Mutual information, difference matrix, expected values, and contingency table.](image)

**d) Intercoder agreement measures.** Anvil allows the user to compute coding agreement using kappa and corrected kappa. It works with annotation files and project files (figure 6), and the matrix content can be conveniently transferred to Excel (or other software). Anvil uses Cohen’s kappa for quantifying the level of agreement (figure 7). This statistical method is appropriate for testing whether agreement exceeds chance levels for binary and nominal ratings. Kappa is widely used to measure the degree of agreement between two independent coders. The extension of kappa to three or more coders has traditionally involved measuring pairwise agreement among all possible pairs of coders.

---

1 The mutual information of two random variables is a quantity that measures the mutual dependence of the two variables (http://en.wikipedia.org/wiki/Mutual_information).

2 A contingency table (or cross-tab) displays the frequency distribution of the variables in a matrix format (http://en.wikipedia.org/wiki/Contingency_table).

3 The mutual information of two random variables is a quantity that measures the mutual dependence of the two variables (http://en.wikipedia.org/wiki/Mutual_information).
The kappa coefficient $\kappa$ measures pairwise agreement among two coders making category judgments. It is generally thought to be a more robust measure than a simple percent agreement calculation, since it takes the observed categories' frequencies as given, and corrects for expected chance agreement.

3. LIMITATIONS. Anvil currently supports .avi (Intel Indeo 5.04, Intel Indeo 5.1, Intel Indeo R3.2) and .mov (Cinepak, Component Video, H-261, H263, foto jpeg, raw without compressor). Videos of type .mts or .avi (DIvX5) cannot be directly opened with Anvnt and need to be converted to the accepted formats.

4. CONCLUSION.

Pros

- Timeline-based annotations
- Data importation from common tools ELAN and PRAAT
- Graphical interface to specify coding scheme
- Useful keyboard shortcuts for coding
- Primary point track type
- Subdivision track type
- Export data to WEKA or frame by frame for advanced statistical analysis
- Visual diagrams/histograms
- Visualization and annotation of motion capture files
Cons
- Some video formats/codecs cannot be imported and have to be converted to the accepted formats

Primary Function
- Timeline-based annotations
- Specification of coding schemes
- Integration with annotation files of other tools
- Track analysis

Platforms
- Windows XP/Vista
- Linux
- Mac

Open Source
Available from http://www.anvil-software.de/: e-mail download@anvil-software.de with subject Anvil Download and provide the following information:
- Your name
- Your research institute
- Some words on your research and how Anvil would be used
- How did you hear of Anvil (Internet, colleague, publication, etc.)?
- Your computer’s operating system (Windows XP/Vista, Linux, Mac)

Reviewed Version: Anvil 5.0 beta 15

Application Size: 30.7 MB for entire directory

Documentation: Anvil manual is accessible from the Anvil graphical interface

Ning Tan
ntan@limsi.fr

Jean-Claude Martin
martin@limsi.fr