EBLA User's Manual

by Brian E. Pangburn

EBLA Version 1.0

Created January 7, 2004
Updated February 25, 2004
Updated December 26, 2014

Table of Contents

List of Figures

List of Tables

Introduction

This is the user's manual for the Experience-Based Language Acquisition (EBLA) software system.

EBLA is an open computational framework for visual perception and grounded language acquisition. EBLA can watch a series of short videos and acquire a simple language of nouns and verbs corresponding to the objects and object-object relations in those videos. Upon acquiring this protolanguage, EBLA can perform basic scene analysis to generate descriptions of novel videos.

The general architecture of EBLA is comprised of three stages: vision processing, entity extraction, and lexical resolution. In the vision processing stage, EBLA processes the individual frames in short videos, using a variation of the mean shift analysis image segmentation algorithm to identify and store information about significant objects. In the entity extraction stage, EBLA abstracts information about the significant objects in each video and the relationships among those objects into internal representations called entities. Finally, in the lexical acquisition stage, EBLA extracts the individual lexemes (words) from simple descriptions of each video and attempts to generate entity-lexeme mappings using an inference technique called cross-situational learning. EBLA is not primed with a base lexicon, so it faces the task of bootstrapping its lexicon from scratch.

While there have been several systems capable of learning object or event labels for videos, EBLA is the first known system to acquire both nouns and verbs using a grounded computer vision system.

EBLA was developed as part of Brian E. Pangburn's dissertation research in the Department of Computer Science at Louisiana State University.

More information on EBLA is available from http://www.greatmindsworking.com and http://sourceforge.net/projects/ebla/

Installation

The latest version of the EBLA software system is available from http://sourceforge.net/projects/ebla/. It requires Java SE version 1.7 or later available from http://www.java.com.

To install EBLA, simply download the latest release file to the desired folder/directory and type:

```
jar -xf ebla-1.0.jar
```

The file, readme.txt, located in the EBLA installation directory contains additional details on installing the EBLA software system and the video data set.

Running EBLA

Starting with version 0.6.0-alpha, EBLA is run via a graphical user interface (GUI) developed using Java Swing. The GUI utilizes the SwingSet open source toolkit for making the standard Swing controls database-aware. More information on SwingSet is available from http://swingset.sourceforge.net

To run EBLA on a Windows machine, change to the "scripts" subdirectory of the EBLA installation directory and type:

```
winrun.bat
```

To run EBLA on a Linux machine, change to the EBLA installation directory and type:

```
sh linrun
```

EBLA Menus

The EBLA software system contains five menus: **File**, **Edit**, **Utilities**, **Reports**, and **Help**.

The **File** menu (figure 1) the fields listed in table 1:

Figure 1: EBLA File Menu

Table 1: EBLA File Menu Items

Menu Item	Description
Login	Connects to the currently configured ebla_data PostgreSQL database.
Logout	Disconnects from the currently configured ebla_data PostgreSQL database.
Exit	Exits the EBLA software system.

The **Edit** menu (figure 2) contains the fields listed in table 2:

Figure 2: EBLA Edit Menu

Table 2: EBLA Edit Menu Items

Menu Item	Description
Parameters	Loads the Parameter Screen for configuring a set of experiences (videos and/or animations).
Experiences	Loads the Experience Screen for editing the experiences available to EBLA.
Attributes	Loads the Attribute Screen for editing the perceptual attributes available to EBLA.

The **Utilities** menu (figure 3) contains the fields listed in table 3:
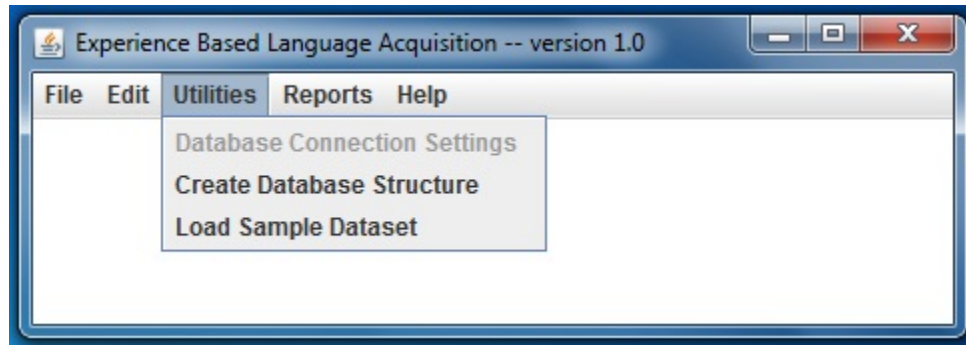
Figure 3: EBLA Utilities Menu

Table 3: EBLA Utilities Menu Items

Menu Item	Description
Database Connection Settings	Loads the Database Connection Screen for configuring the H2 database.
Create Database Structure	Creates the EBLA database table structure in a new/empty database.
Load Sample Dataset	Loads a sample dataset into a database with no records.

The **Reports** menu (figure 4) fields listed in table 4:

Figure 4: EBLA Reports Menu

Table 4: EBLA Reports Menu Items

Menu Item	Description
Data Files	Provides a dialog box with information on the location of the delimited files that contain the results for each session. *Note that integrated reporting is not yet available, but is planned for a future release.*

The **Help** menu (figure 5) contains the fields listed in table 5:
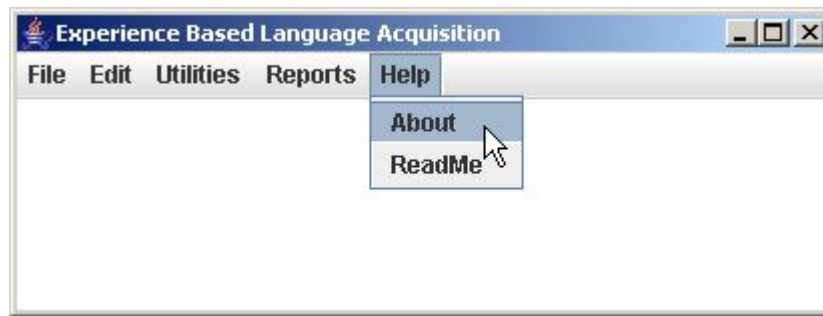
Figure 5: EBLA Help Menu

Table 5: EBLA Help Menu Items

Menu Item	Description
About	Provides a dialog box with information about the current version of EBLA.
ReadMe	Loads the EBLA ReadMe file.

Database Connection Screen

The **Database Connection Screen** (figure 6) allows the user to specify the database connection info for the EBLA H2 database and contains the fields listed in table 6:
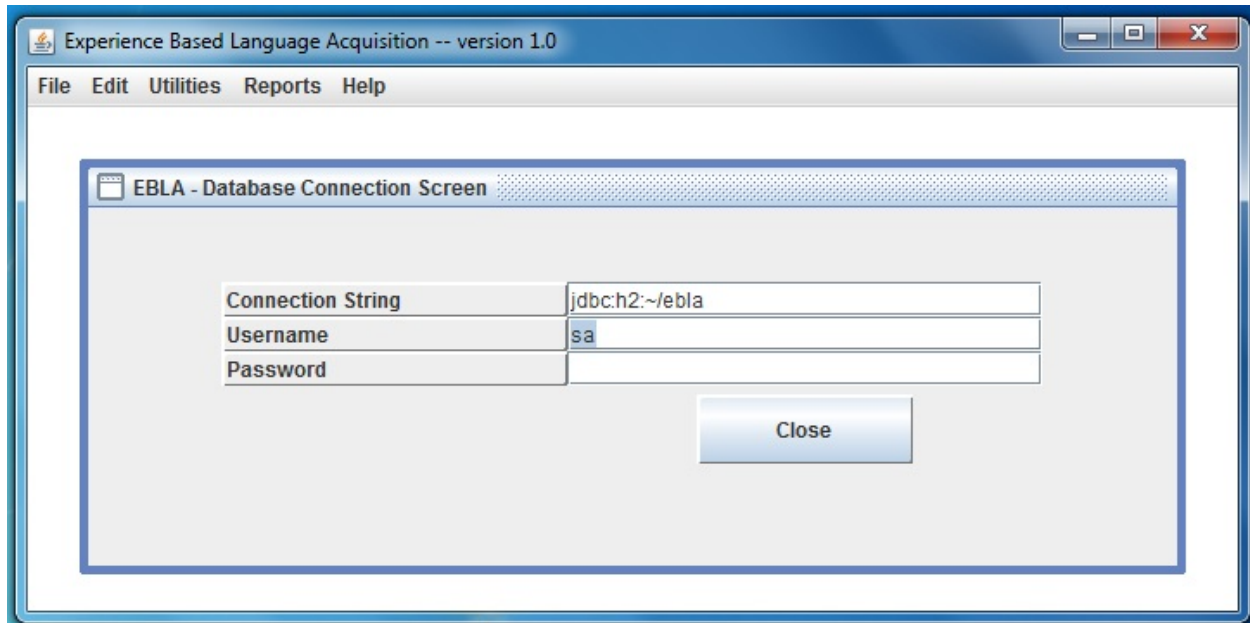
Figure 6: EBLA Database Connection Screen

Table 6: EBLA Database Connection Fields

Field Name	Description
Connection String	H2 database connection string (default "jdbc:h2:~ebla"
Username	H2 database username (default "sa")
Password	H2 database password (default ***none***)

When you start EBLA for the first time, the software will create the database structure and give you the option to load a sample dataset.

If you would like more information on the H2 embedded database, please visit http://www.h2database.com.

Parameter Screen

The **Parameter Screen** (figures 7a, 7b, 7c, & 7d) allows the user to specify a basic set of parameter for controlling the EBLA software system. In general, a single set of parameters determines how the EBLA vision system will operate and where the images generated by the vision system will be stored. Each set of parameters has a set of experiences (videos/animations) associated with it forming a complete package to be processed during a given EBLA calculation session. This way, the vision system can be adjusted for different groups of experiences.
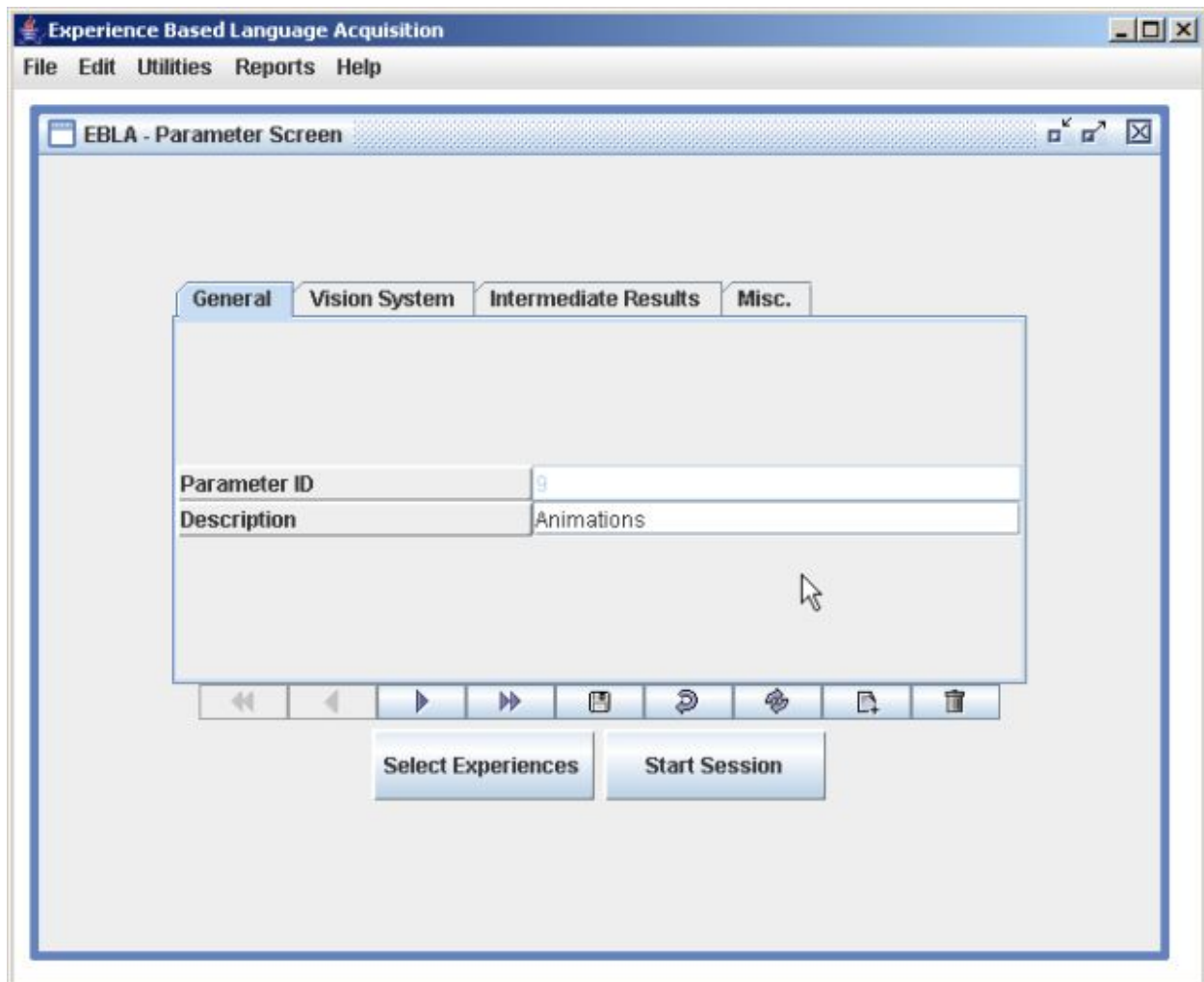
Figure 7a: EBLA Parameter Screen – General Tab

The **Parameter Screen** is divided into three vertical sections: tabs, a record navigation bar, and buttons. The tabs are used for data entry and are described in detail below. The record navigation bar is used to traverse, add, and delete sets of parameters in the database. It can also be used to undo editing changes, commit new changes (note that changes for the current record are always committed by default if the user navigates to another record or if the screen loses the focus), and refresh the screen to reflect any changes made to the underlying database by other users. The "Select Experiences" button is used to select which experiences are associated with

the current set of parameters. The "Start Session" button is used to launch a calculation/processing session.

The **General Tab** (figure 7a) is used to display/edit summary information for each parameter set. It contains the fields listed in table 7a.

Table 7a: Parameter Screen – General Tab Fields

Field Name	Default Value	Description
Parameter ID	n/a	Underlying database record ID for the current set of parameters (editing not allowed).
Description	n/a	Short descriptive name for the current set of parameters.

The **Vision System Tab** (figure 7b) is used to display/edit vision system settings for each parameter set. It contains the fields listed in table 7b.
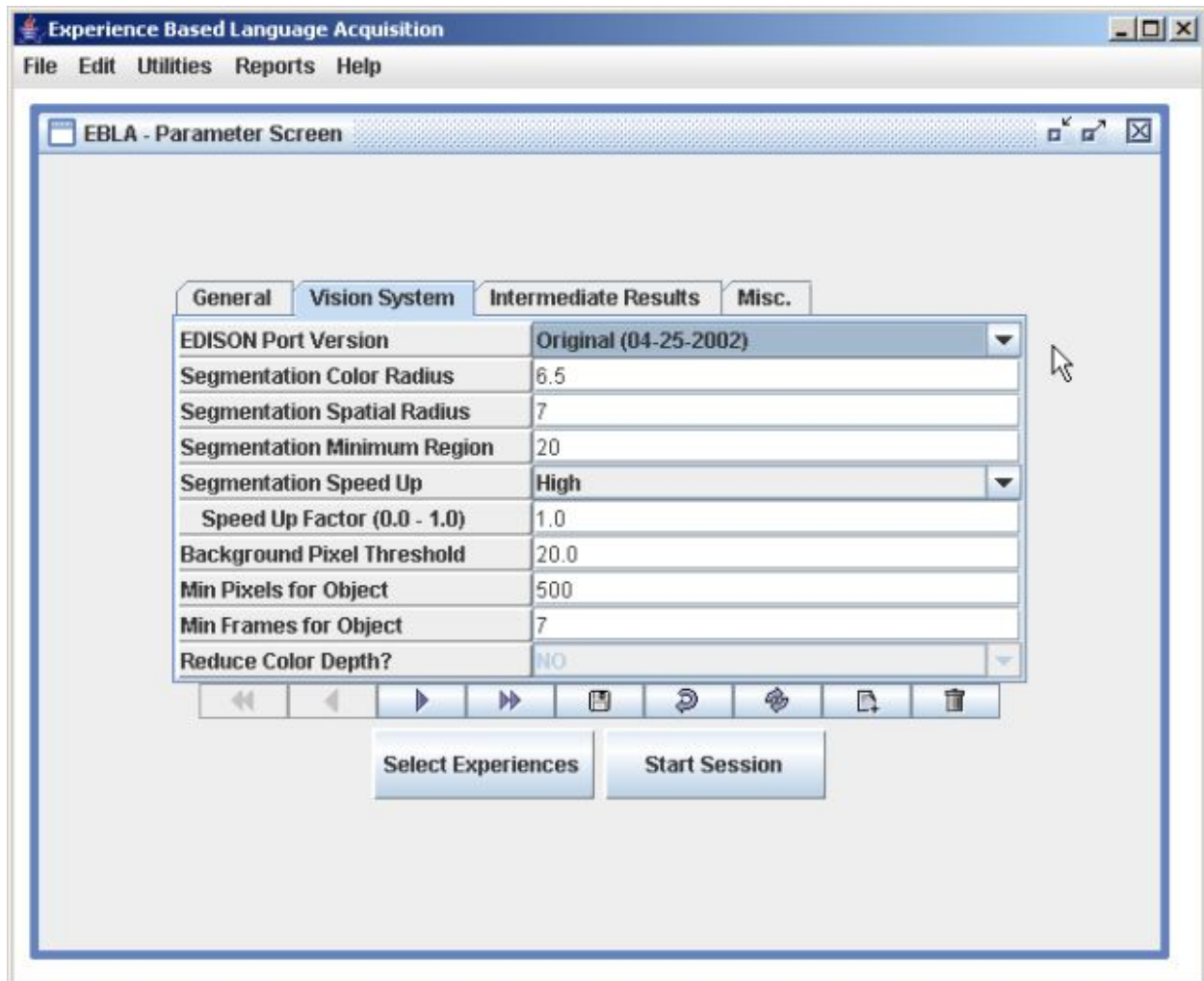

Figure 7b: EBLA Parameter Screen – Vision System Tab

Table 7b: Parameter Screen – Vision System Tab Fields

Field Name	Default Value	Description
EDISON Port Version	Original (04-25-2002)	Version of EDISON port to use for image segmentation. Current versions include "Original" (based on the 04-25-2002 release of EDISON) and "Revised" (based on the 04-14-2003 release of EDISON)
Segmentation Color Radius	6.5	Color or range resolution – determines the extent to which colored regions are merged together. A lower color radius leads to oversegmentation where objects are broken into many small colored regions, and a higher color radius leads to undersegmentation where distinct objects are merged together.
Segmentation Spatial Radius	7	Spatial resolution - defines a spatial search window for the mean shift computations. Spatial resolution has the greatest impact on the execution time. Higher values require more processing time.
Segmentation Minimum Region	20	The minimum area of a segmented region (in pixels). area in pixels. Generally, smaller values lead to small regions in the middle and on the edge of objects where glare and shadows are detected.
Segmentation Speed Up	Medium	There are three speed-up options (no-speed up, medium-speed up, and high-speed up). Generally, the higher the level of speed up, the lower the segmentation quality. The default speed up setting of medium generally produces good results that are not significantly different from the no-speedup setting. The high speed up option takes an addition parameter to control speed up versus quality.
Speed Up Factor	0.5	For the high-speed up option determines the balance of speed up versus quality. Values range from 0.0 (higher quality) to 1.0 (higher speed up).
Background Pixel Threshold	20%	Percentage of total image pixels that an object must contain to be considered part of the background rather than a significant object (0 - 100).
Min Pixels for Object	500	Minimum number of pixels that constitute a significant object. Regions with fewer pixels will be discarded by the vision system.
Min Frames for Object	7	Minimum number of consecutive frames that an object must appear in to be considered a significant object (helps to eliminate noise and shadows).
Reduce Color Depth	No	Indicates whether or not the vision system should reduce the RGB color depth of each significant region from 16,777,216 to 27. *This feature is currently disabled because object "color" is converted to grayscale values between 0 and 255.*

More information on the image segmentation algorithm and its parameters is available from the following papers:

1. D. Comanicu, P. Meer: "Mean shift: A robust approach toward feature space analysis." *IEEE Trans. Pattern Anal. Machine Intell.*, **24**, 603-619, May 2002.
2. P. Meer, B. Georgescu: "Edge detection with embedded confidence." *IEEE Trans. Pattern Anal. Machine Intell.*, **23**, 1351-1365, December 2001.
3. C. Christoudias, B. Georgescu, P. Meer: "Synergism in low-level vision." 16*th International Conference on Pattern Recognition*, Quebec City, Canada, August 2002, vol. IV, 150-155.

All of the above papers are available from: http://coewww.rutgers.edu/riul/research/robust.html

The **Intermediate Results Tab** (figure 7c) is used to display/edit settings for the storage of images generated by EBLA's vision system. It contains the fields listed in table 7c.
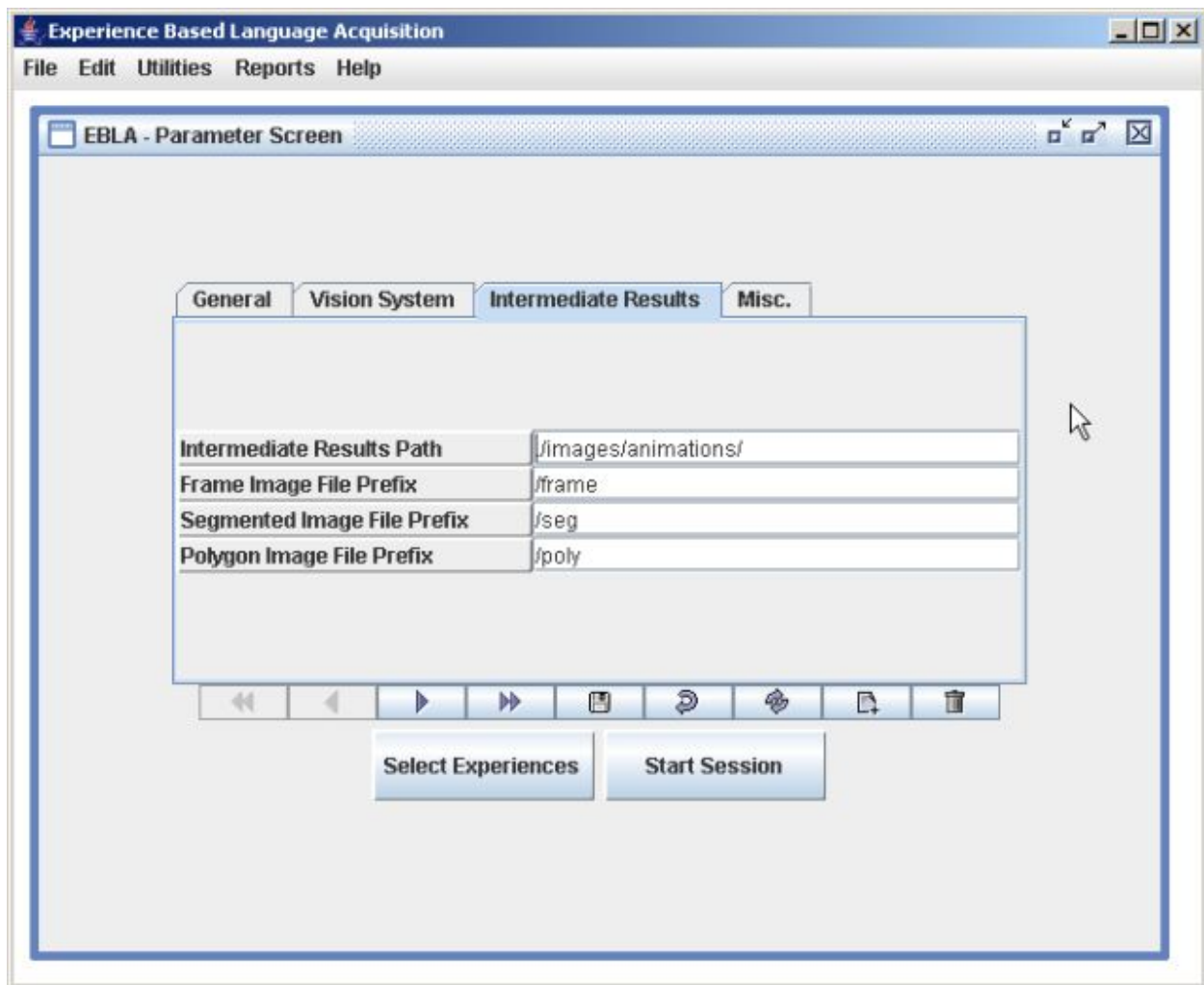

Figure 7c: EBLA Parameter Screen – Intermediate Results Tab

11

Table 7c: Parameter Screen – Intermediate Results Tab Fields

Field Name	Default Value	Description
Intermediate Results Path	n/a	Path for storing images produced while processing experiences (videos/animations). A subdirectory will be created for each experience.
Frame Image File Prefix	n/a	Filename prefix for original frames extracted from each experience.
Segmented Image File Prefix	n/a	Filename prefix for segmented images created for each extracted frame.
Polygon Image File Prefix	n/a	Filename prefix for polygon images created for each extracted frame.

The **Misc. Tab** (figure 7d) is used to display/edit miscellaneous settings for each parameter set. It contains the fields listed in table 7d.

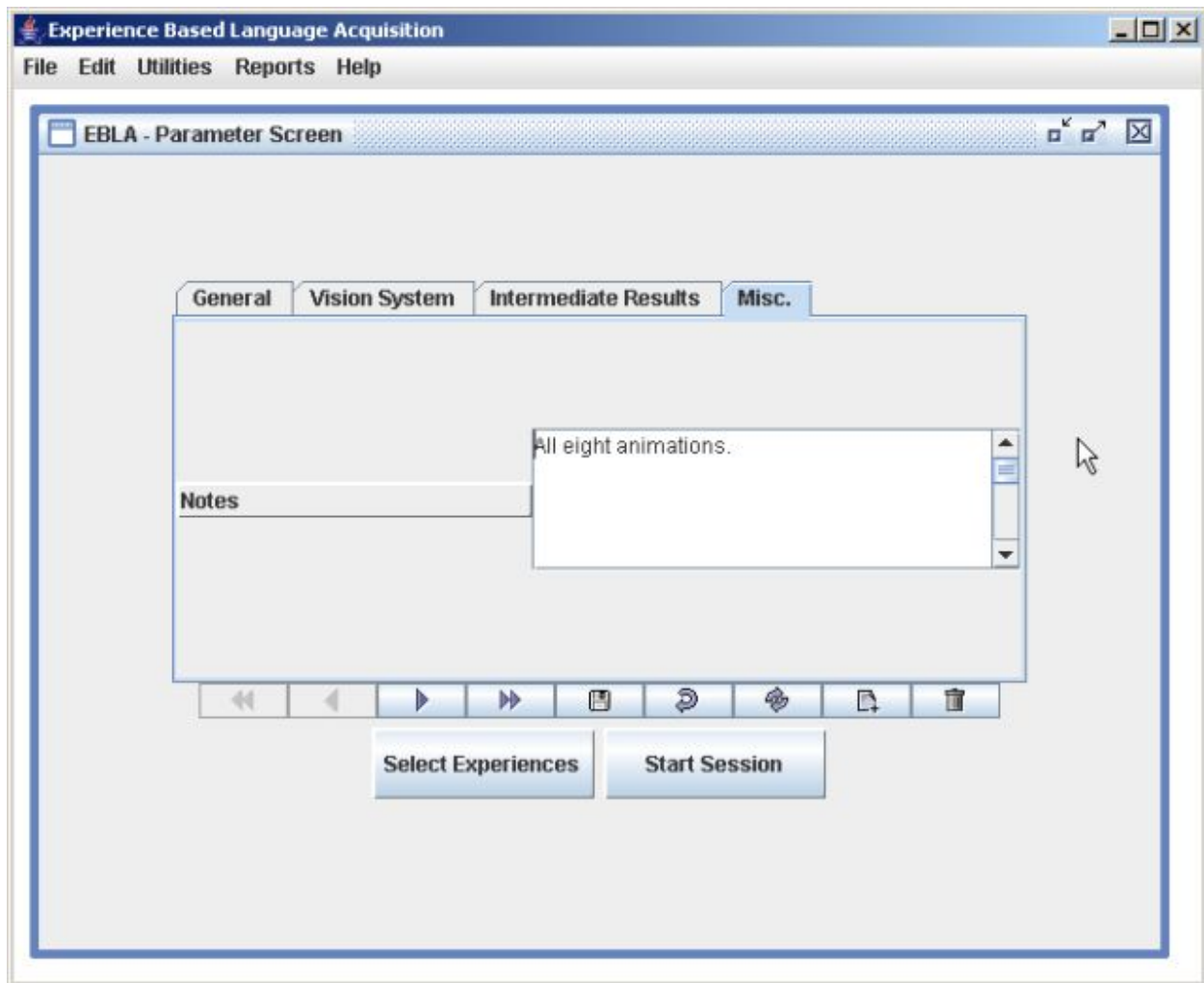Figure 7d: EBLA Parameter Screen – Misc. Tab

Table 7d: Parameter Screen – Misc. Tab Fields

Field Name	Default Value	Description
Notes	n/a	Notes regarding the current set of parameters.

Experience Screen

The **Experience Screen** (figure 8) allows the user to specify the settings for each experience to be processed by the EBLA Software System. An experience is typically a video or animation containing a basic set of objects interacting in some way (e.g. a hand picking up a ball). Each experience is given a basic protolanguage description (e.g. "hand pickup ball) and is delivered to EBLA as an AVI file. EBLA attempts to detect objects and object-object relations in each video using its vision system and then attempts to correlate those objects/relations with the individual words in each protolanguage description.

The **Experience Screen** is divided into three vertical sections: data entry, record navigation, and buttons. The data entry fields are listed in table 8. The record navigation bar is used to traverse, add, and delete experiences in the database. It can also be used to undo editing changes, commit new changes (note that changes for the current record are always committed by default if the user navigates to another record or if the screen loses the focus), and refresh the screen to reflect any changes made to the underlying database by other users. The "Close" button is used to close the screen.
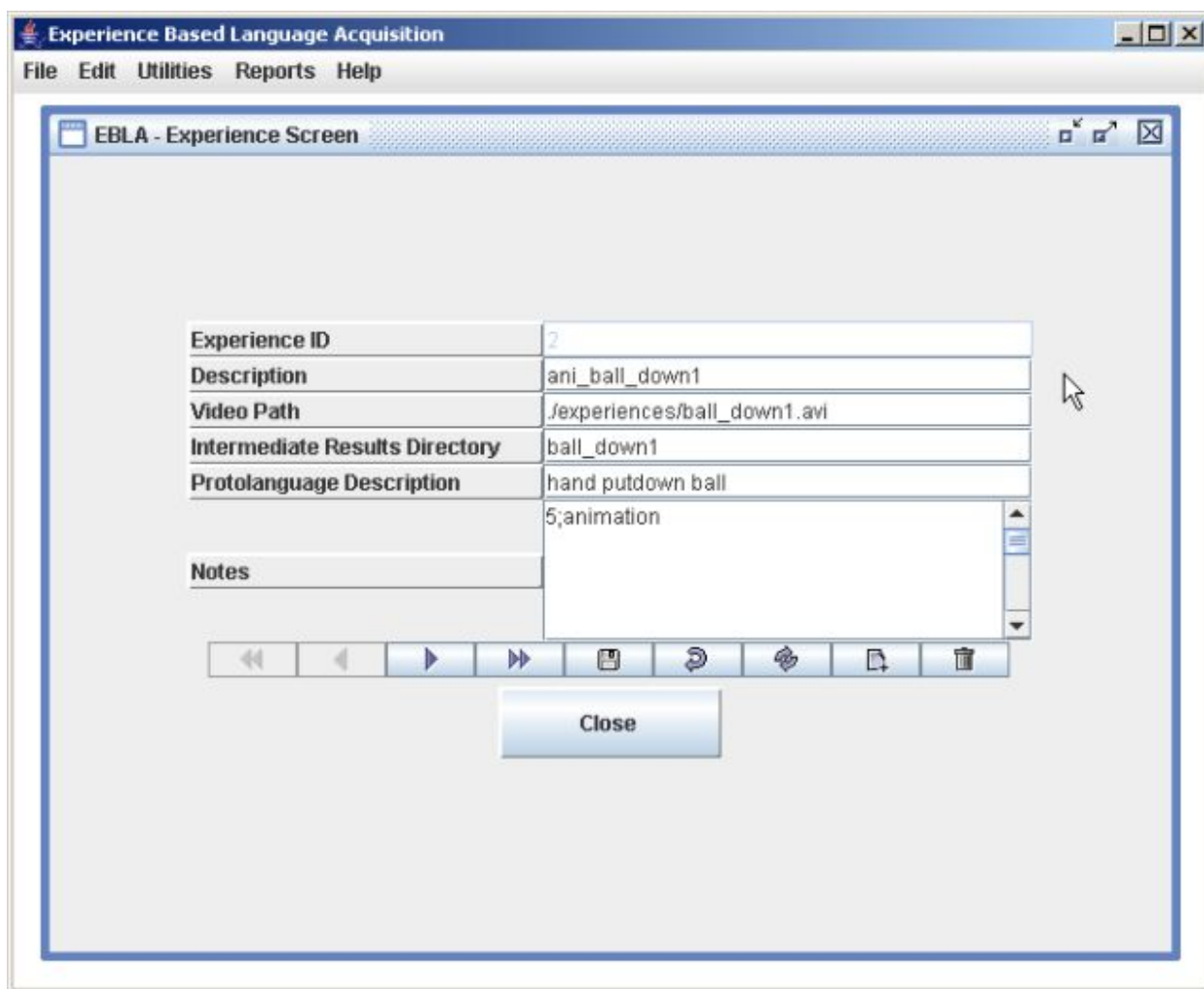

Figure 8: EBLA Experience Screen

14

Table 8: Experience Screen Fields

Field Name	Default Value	Description
Experience ID	n/a	Underlying database record ID for the current experience (editing not allowed).
Description	n/a	Short descriptive name for the current experience.
Video Path	n/a	Full or relative path and filename for the AVI file containing the video/animation for the current experience.
Intermediate Results Directory	n/a	Name of subdirectory to create for storing images produced for storing the current experience. It will be placed in the "Intermediate Results Path" specified on the Parameter Screen.
Protolanguage Description	n/a	Set of space-delimited lexemes (words) describing the current experience.
Notes	n/a	Notes regarding the current experience.

Attribute Screen

The **Attribute Screen** (figure 9) allows the user to specify the attributes that EBLA calculates for each detected object or object-object relation. There are currently five object attributes and seven relation attributes that are hard-coded into the EntityExtractor EBLA calculation class. Since these attributes are hard-coded in the current version, the **Attribute Screen** is only useful for browsing, enabling, and/or disabling the twelve attributes. Future versions of EBLA will hopefully allow users to define their own calculation classes that can be loaded at runtime using the Class.forName() construct.

The **Attribute Screen** is divided into three vertical sections: data entry, record navigation, and buttons. The data entry fields are listed in table 9. The record navigation bar is used to traverse, add, and delete attributes in the database. It can also be used to undo editing changes, commit new changes (note that changes for the current record are always committed by default if the user navigates to another record or if the screen loses the focus), and refresh the screen to reflect any changes made to the underlying database by other users. The "Close" button is used to close the screen.

Figure 9: EBLA Attribute Screen

16

Table 9: Attribute Screen Fields

Field Name	Default Value	Description
Attribute ID	n/a	Underlying database record ID for the current attribute (editing not allowed).
Description	n/a	Short descriptive name for the current attribute.
Include Attribute?	(Yes)	Determines whether or not the current attribute is calculated for each object or object-object relation.
Attribute Type	(Object)	Determines whether the current attribute is calculated for objects or object-object relations. *This feature is currently disabled the EBLA attribute calculations are hard-coded.*
Attribute Calculation Class	n/a	Java class name for custom attribute calculation class that can be loaded at runtime. *This feature is currently disabled because dynamic attribute calculations have not yet been implemented.*
Notes	n/a	Notes regarding the current experience.

Select Experiences Screen

The **Select Experiences Screen** (figure 10) allows the user to specify which experiences are to be processed for each set of parameters. The screen is divided into two sections: experience selection and buttons. The experience selection section consists of a list of all available experiences (left), a list of selected experiences for the current parameter set (right) and buttons (middle) to add/remove experiences from the list of selected experiences. The "Close" button is used to close the screen.
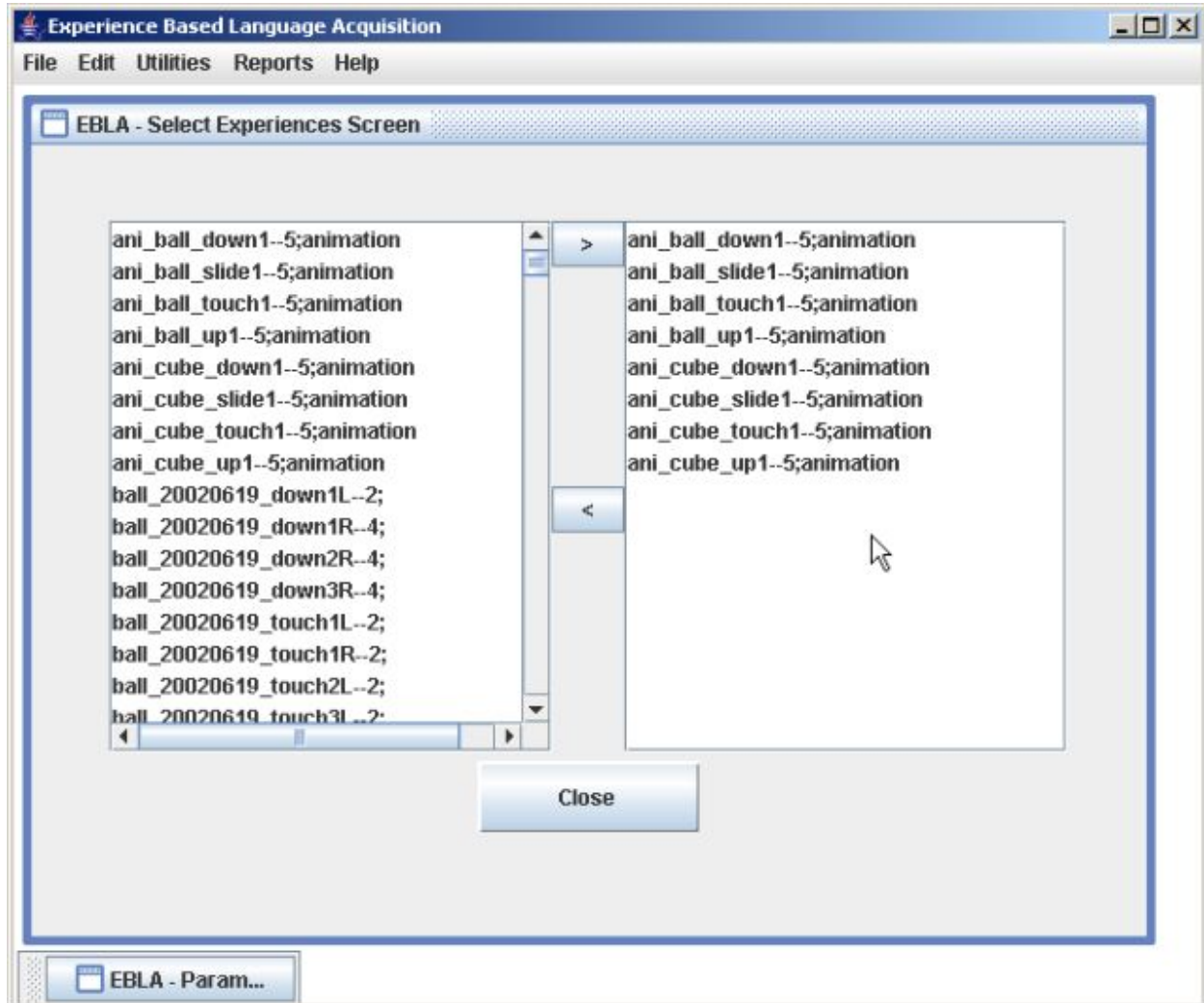

Figure 10: EBLA Select Experiences Screen

Session Screen

The **Session Screen** (figures 11a, 11b, 11c, & 11d) allows the user to specify the settings for an EBLA calculation session. During a given calculation session, EBLA, will perform video processing (as needed), entity extraction, and lexical resolution. While the video processing stage of EBLA only needs to be performed once for a given set of parameters, the entity extraction and lexical resolution stages can be run multiple times to evaluate a variety of settings.

The **Session Screen** is divided into two vertical sections: tabs and buttons. The tabs are used for data entry and are described in detail below. The "Start EBLA" button is used to start the EBLA calculation session using the specified values and the "Close" button is used to close the screen.

The **General Tab** (figure 11a) is used to display/edit summary information for each calculation session. It contains the fields listed in table 11a.

Figure 11a: EBLA Session Screen – General Tab

Table 11a: Session Screen – General Tab Fields

Field Name	Default Value	Description
Description	n/a	Short descriptive name for the current calculation session.
Log To File?	(Yes)	Indicates whether EBLA should log calculation engine messages to the screen or to a log file.
Randomize Experiences?	(Yes)	Indicates whether EBLA should retrieve experiences from the ebla_data database in random order (Yes) or in the order they were entered (No).
Regenerate Images?	(No)	Indicates whether EBLA should rerun the vision processing stage on all experiences for the current parameter set and thus reproduce all of the intermediate images or only perform video processing as needed. *Note that the video processing stage is very processor intensive and can take hours (or even days for an older computer) to complete.*
Display Detailed Messages?	(No)	Indicates whether EBLA should produce detailed messages while processing.

The **Entity Recognition Tab** (figure 11b) is used to display/edit entity recognition settings for each calculation session. It contains the fields listed in table 11b. Since every field in this tab is related to the "Minimum Standard Deviation," a term coined for performing entity comparisons in EBLA, some further explanation is warranted.

In EBLA, both object and relation candidate entities are compared to existing entities using average attribute values. Generally if all average attribute values for a candidate entity are within a single standard deviation the average attribute values for an existing entity, there is said to be a match and the candidate entity is merged with the existing entity. Otherwise the candidate entity becomes a new entity. In practice, a single standard deviation is far too restrictive for some attributes because of their minimal variance. To address this, the concept of a "minimum standard deviation" is introduced. It is simply defined as a percentage of a given candidate entity's average attribute values. If the actual standard deviation is less than the specified minimum, the specified minimum is used for comparing the candidate entity's attribute values to those of existing entities.
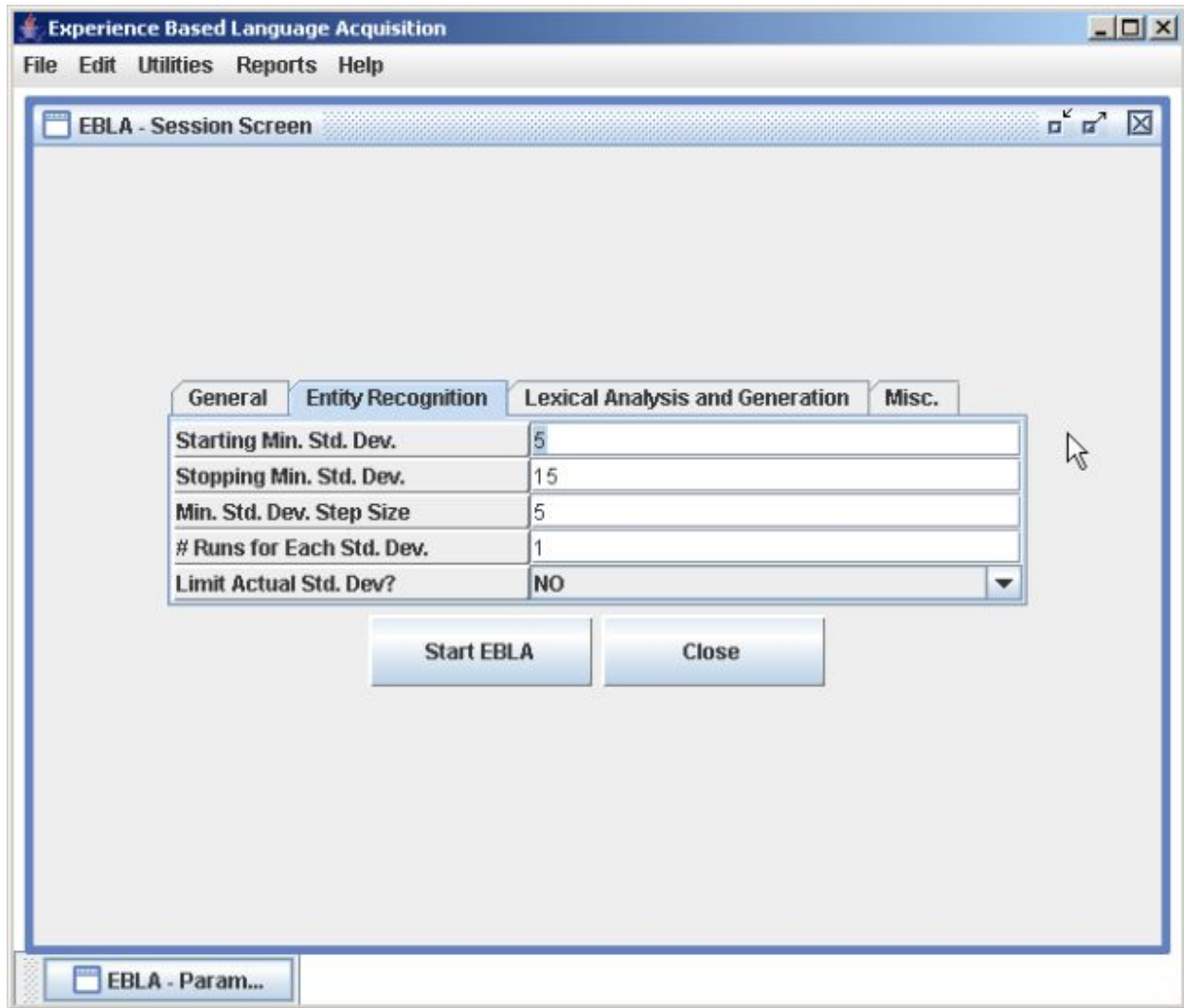
Figure 11b: EBLA Session Screen – Entity Recognition Tab

Table 11b: Session Screen – Entity Recognition Tab Fields

Field Name	Default Value	Description
Starting Min. Std. Dev.	5	The lowest (first) minimum standard deviation to use for comparing entity attribute values.
Stopping Min. Std. Dev.	5	The highest (last) minimum standard deviation to use for comparing entity attribute values.
Min. Std. Dev. Step Size	5	The step size to use for increasing from the starting minimum standard deviation to the stopping minimum standard deviation.
# Runs for Each Min. Std.	1	Number of times to process all experiences for each minimum standard deviation.
Limit Actual Std. Dev?	(No)	Indicates whether or not the actual calculated standard deviation for the current entity attribute should be limited to the current minimum standard deviation.

The **Lexical Analysis and Generation Tab** (figure 11c) is used to display/edit settings for the lexical analysis and generation settings for each calculation session.. It contains the fields listed in table 11c.
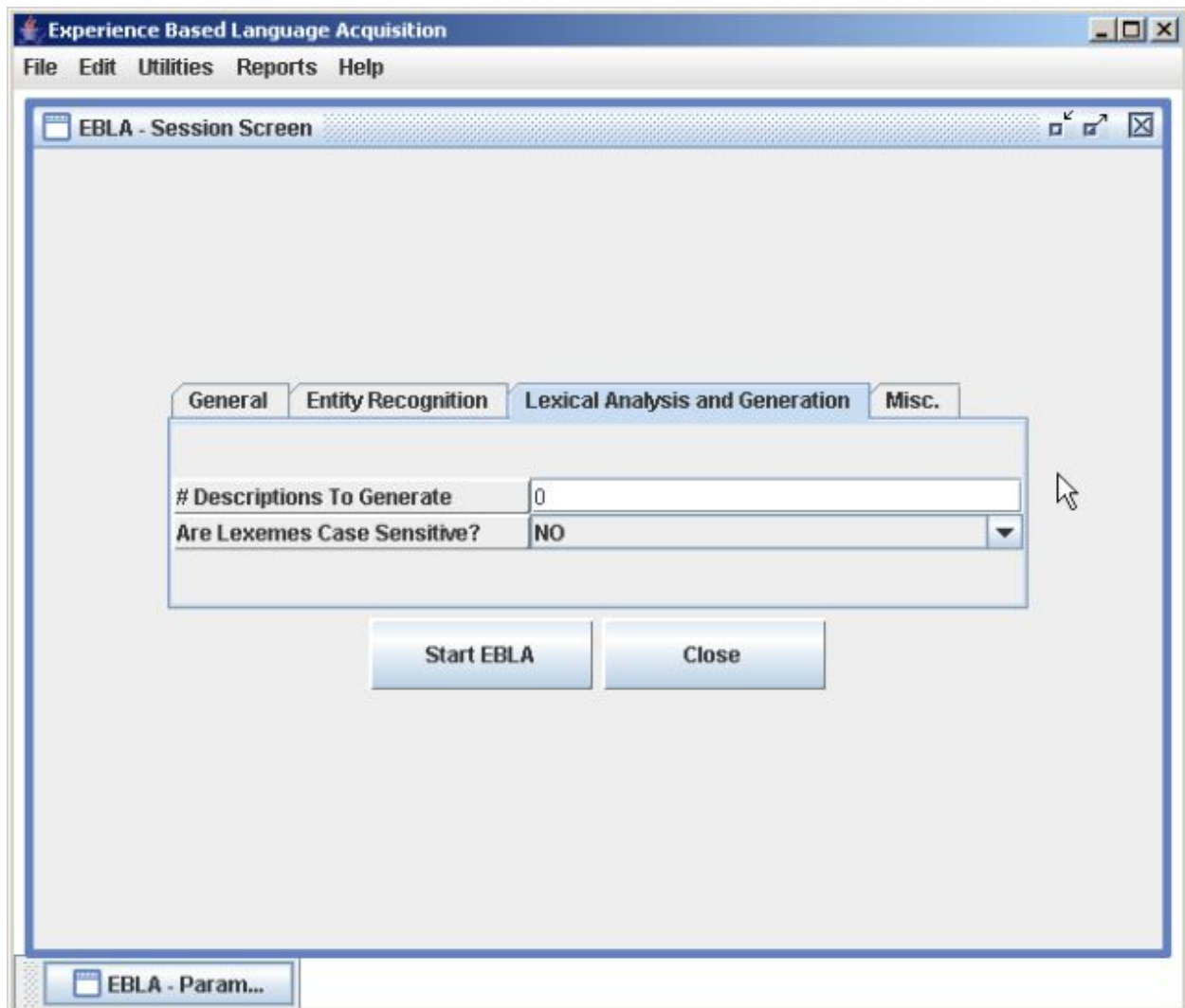

Figure 11c: EBLA Session Screen – Lexical Analysis and Generation Tab

Table 11c: Session Screen – Lexical Analysis and Generation Tab Fields

Field Name	Default Value	Description
# Descriptions to Generate	0	Indicates for how many experiences EBLA should attempt to generate descriptions rather than use the available protolanguage description for standard lexical analysis.
Are Lexemes Case Sensitive?	(No)	Indicates whether or not the lexemes in the protolanguage descriptions provided for each experience should be treated as case sensitive.

The **Misc. Tab** (figure 11d) is used to display/edit miscellaneous settings for each calculation session. It contains the fields listed in table 11d.
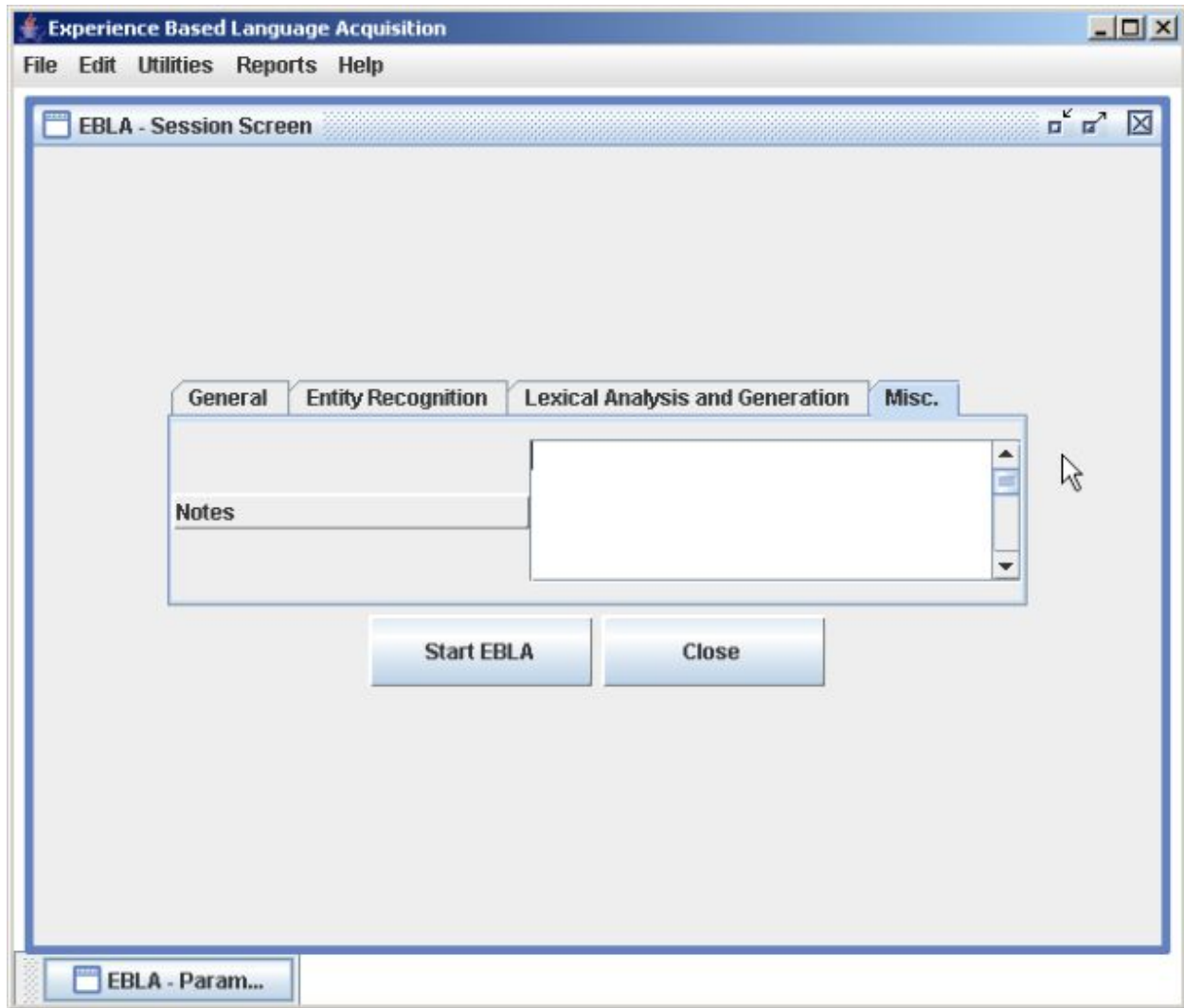
Figure 11d: EBLA Session Screen – Misc. Tab

Table 11d: Session Screen – Misc. Tab Fields

Field Name	Default Value	Description
Notes	n/a	Notes regarding the current calculation session.

Calculation Status Screen

The **Calculation Status Screen** (figure 12) allows the user to monitor the progress of the EBLA calculation engine. If vision processing is required, each frame of each experience will be displayed in the left-most image pane as it is ripped. Then, after each frame is analyzed, the original frame will be displayed along with the segmented image and a polygon image. The polygon images contain polygon traces of all significant objects along with bounding rectangles (red) and centroid markers (black). Following any vision processing, progress bars are provided to track the entity extraction and lexical resolution runs for the specified range of minimum standard deviation values. The user may cancel EBLA processing at any time by pressing the "Cancel" button at the bottom of the screen. There may be a small pause following a cancel request while EBLA confirms that the ebla_data database is in a consistent state before terminating.
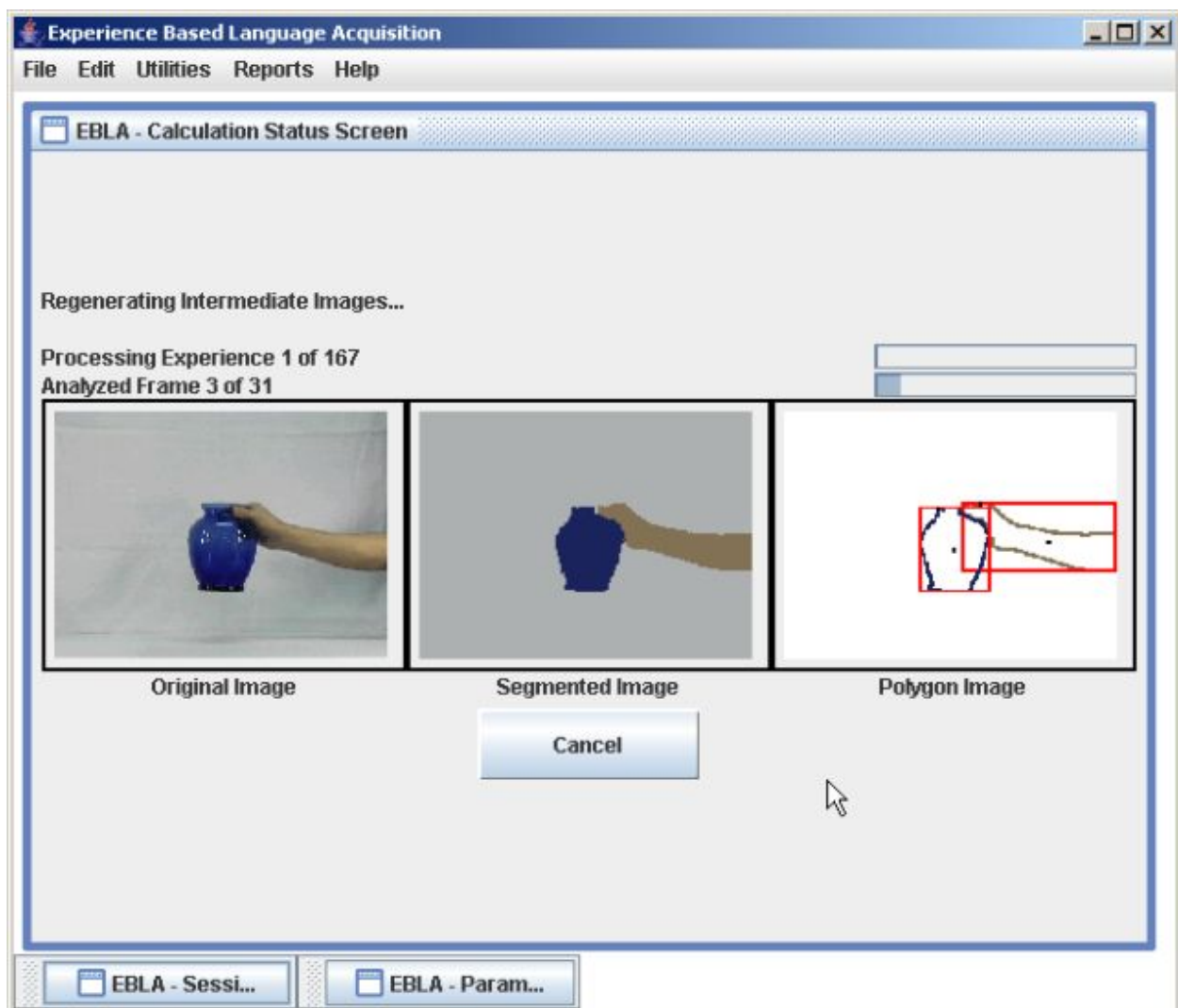

Figure 12: EBLA Calculation Status Screen

Reports

Integrated reporting has not yet been implemented in EBLA. To analyze calculation results in EBLA, users can either extract data directly from the ebla_data database or make use of the delimited text results files generated for each calculation session. The files (session_###_performance.ssv, session_###_mappings.ssv, and session_###_descriptions.ssv) are stored in the EBLA installation directory. More information about the contents of these files is available in the file, readme.txt.

If you are planning to extract data directly from the ebla_data database, the relationships for the various tables are show in figure 13.
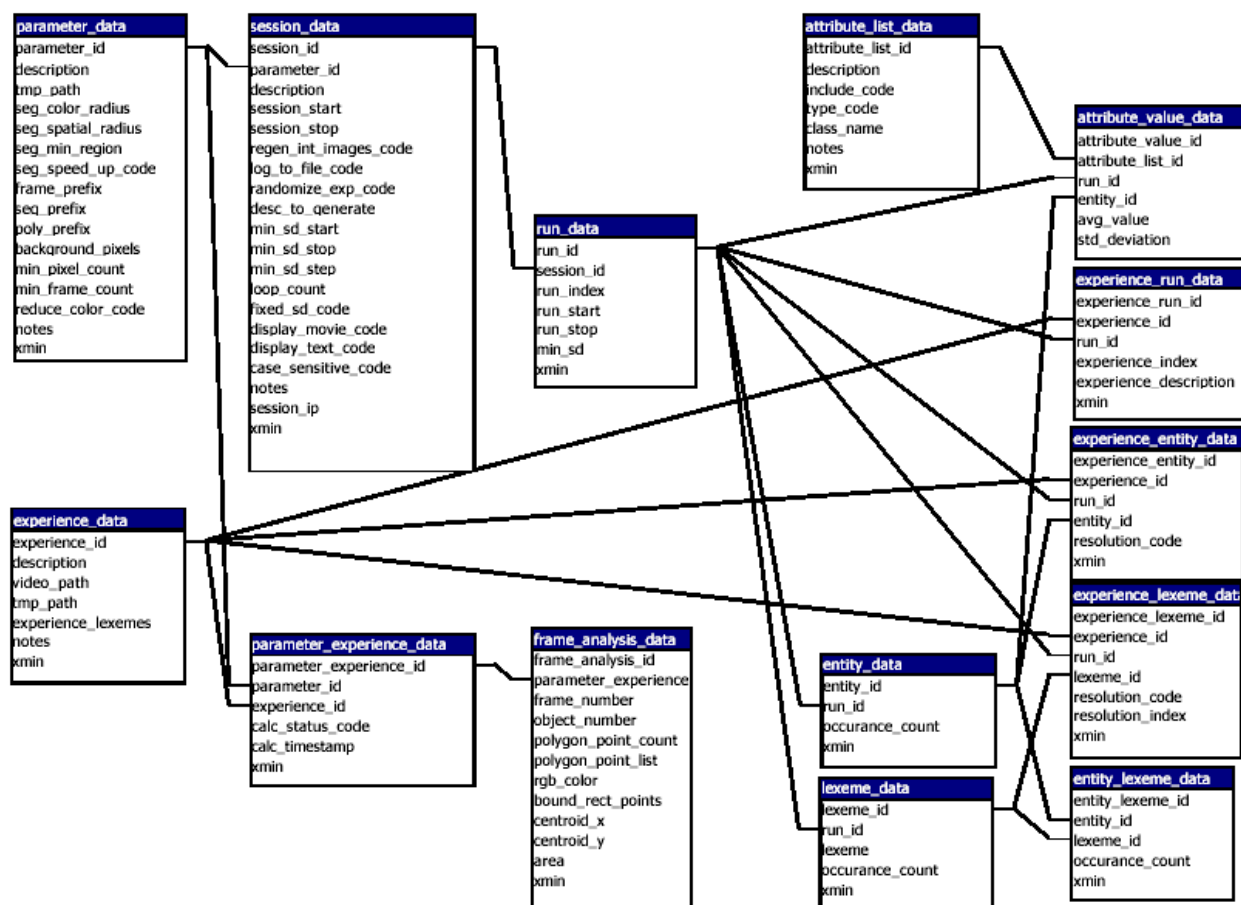

Figure 13: Table Relationships for the ebla_data Database