

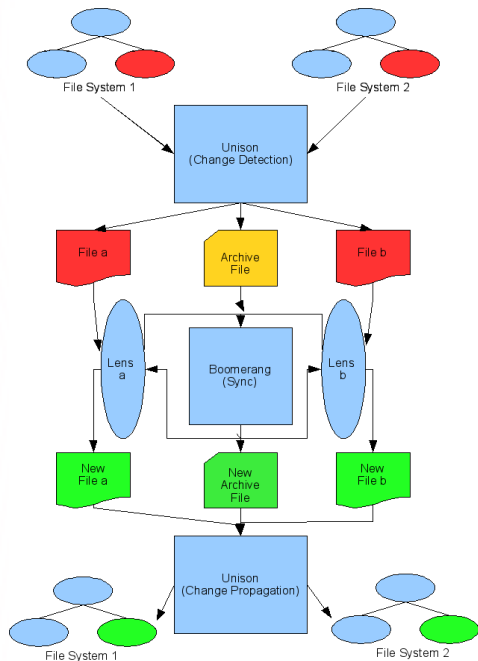
# Universal Data Synchronizer for Personal Calendars

Hui Du  
Faculty Advisor: Benjamin C. Pierce

## Abstract:

- Create an extensible multi-platform data synchronizer that handles personal calendar data from multiple machines and applications using the bidirectional language Boomerang

## System Structure:



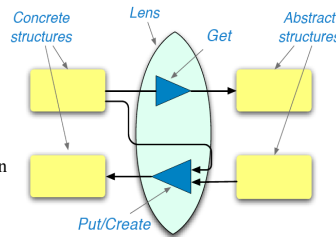
- Unison detects any change in a file, and Boomerang reconciles changes within a file. Unison also stashes the archive file.

## Augmenting Unison:

- The original Unison required both copies of a file to be identical, which is undesirable for the implementation of partial sync
- To allow for partial sync, the structure of the Unison archive was changed to include an "ArchivePair" data structure, which holds the data for a pair of files
- By default, if only one copy of a file changes, the unchanged file is overwritten. For files marked for sync by Boomerang, this action is changed to call out to Boomerang.

## What is a Lens?

- A lens converts string data from one format to another. It has operations that allow the conversion to go in both directions, eliminating the need to create separate conversion tools for each direction of transformation.



Foster, J Nathan and Pierce, Benjamin C. *Boomerang Manual*

- Boomerang is a bidirectional language for creating lenses

## The iCal to CSV Lens

- The most important lens implemented in my system converts iCal/vCal format files to CSV format. The CSV format based on the format used for Outlook exports, with the addition of a UID field. Here is sample conversion:

### iCal:

```

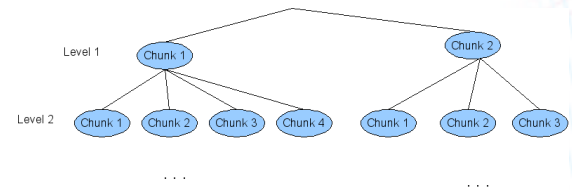
BEGIN:VCALENDAR
VERSION:2.0
METHOD:PUBLISH
BEGIN:VEVENT
DTSTART;TZID="Eastern Time (US & Canada)":20080903T183000
DTEND;TZID="Eastern Time (US & Canada)":20080903T213000
LOCATION:COHN 204
SUMMARY:PHILO55601: EXISTENTIALISM
DESCRIPTION:
UID:D2161A4B-5056-9301-4BE4F2A2CC1315F7@wharton.upenn.edu
PRIORITY:5
END:VEVENT
BEGIN:VEVENT
DTSTART;TZID="Eastern Time (US & Canada)":20080904T103000
DTEND;TZID="Eastern Time (US & Canada)":20080904T120000
LOCATION:JMHI 360
SUMMARY:FNCE219002: INTL FINANCIAL MARKETS
DESCRIPTION:
UID:D2161A6A-5056-9301-4B45ED7143B1C743@wharton.upenn.edu
PRIORITY:5
END:VEVENT
END:VCALENDAR
    
```

	A	B	C	D	E	F	V	
Subject								UID
PHILO55601: EXISTENTIALISM		Start Date	Start Time	End Date	End Time	All day event	Show time as	20080903T183000
		09/03/08	06:30:00 PM	09/03/08	09:30:00 PM	FALSE		20080903T213000
FNCE219002: INTL FINANCIAL MARKETS		09/04/08	10:30:00 AM	09/04/08	12:00:00 PM	FALSE		20080904T103000
								20080904T120000

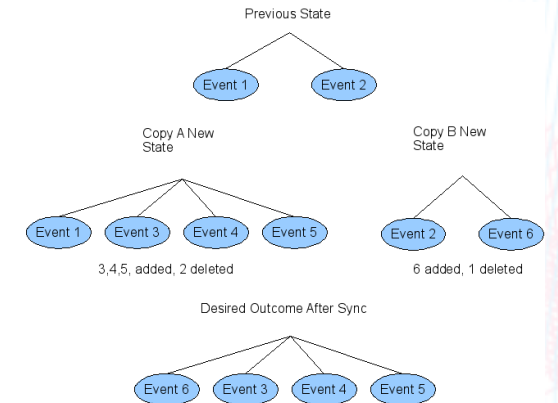
- All events are synchronized in the CSV format
- Using the lens structure, it becomes simple to specify events that one does not want synchronized. In my implementation, this is done by placing <NoSync> at the beginning of the event description.

## Augmenting Boomerang Sync Functionality

- All synchronization is done using three inputs:
  - O – the archive file (stored in CSV format) that represents the state of the system at the end of last sync. This is a "blank" file if no last sync exists.
  - A – the first copy of the file
  - B – the second copy of the file
- Boomerang uses a system called "chunking" to create a tree for syncing (each chunk is a block of text):



- By default, boomerang syncs each level using diff3 sync and matching by key, but for calendar data it is useful to sync chunks as if they were unordered, such as when adding and deleting events:



- Diff3 sync would detect a conflict between the added events rather than achieving the desired outcome, so a new sync method was implemented.
- An option in boomerang there was added to specify any specific chunk as ordered (default) or unordered
- In syncing using the CSV format, events are unordered, and individual fields are ordered