MONASH DEBATING
Review

Proudly supported by

 MONASH University

MONASH
DEBATING
Review

TECHNOLOGY

MONASH
DEBATING
review

THE START OF HISTORY

PHILIP BELESKY
Royal Melbourne Institute of Melbourne
Victoria University of Wellington

CHUAN-ZHENG LEE
Stanford University
Victoria University of Wellington
University of Auckland

Introduction

If software is eating the world then debating was swallowed a long time ago. Hand-tabbing is rare; in its place we find a range of digital tab systems, each making ever-increasing strides in their allocation algorithms, user interfaces, online accessibility, and data entry methods. Yet, despite all these advances, released tabs—even from major tournaments—are doomed to die as websites expire and record-keepers fade from circuits.

We believe tabs have immense value. To individuals they are public recognition of hard-won achievements. To the debating community they are historic records and a typically-unrealised means of understanding how to improve our competitive practices.

In this article, we propose a standardised tab archive format which would allow anyone to download, store, and process tournament data produced by any tab system. Unlike web tables or Excel spreadsheets, a standardised and open format could store complex information that would enable a diverse range of applications to innovate independently of tab systems. A simple application could be a website that presents the complete historical records of events such as WUDC, Australs, or the APDA circuit. But much more is possible: these archives could be used to more easily analyse gender and regional diversity over time, create institution records of achievement, attempt to understand possible causes of adjudicator bias, generate new forms of rankings, or create a comprehensive motion bank that performs balance analysis.

The data for all these applications surely exists. But it is hard to find—typically lost in the deepest corners of a since-retired tab director's computer. This inaccessibility, we argue, is the biggest impediment to preserving and analysing tab data. A standard format

would remove this barrier: archives that are easy to exchange are easy to preserve, and analysts could use the same tools to inspect data regardless of the tab system used in each particular tournament.

Background

There are many tab systems. Some run only on web servers, on PCs, or as Excel spreadsheets. In other domains—such as publishing documents or working with photos—standardised formats enable users to view and edit documents produced by different applications. For example, PDF files, originally an Adobe creation, can now be created by and opened in countless applications on both computers and mobile phones (e.g. Adobe Reader, Google Chrome); similarly, JPEG images work with digital cameras, mobile phones, web browsers, and photo editors alike. This interoperability is key to document sharing and creating software 'ecosystems' wherein developers create applications upon a common platform or standard.
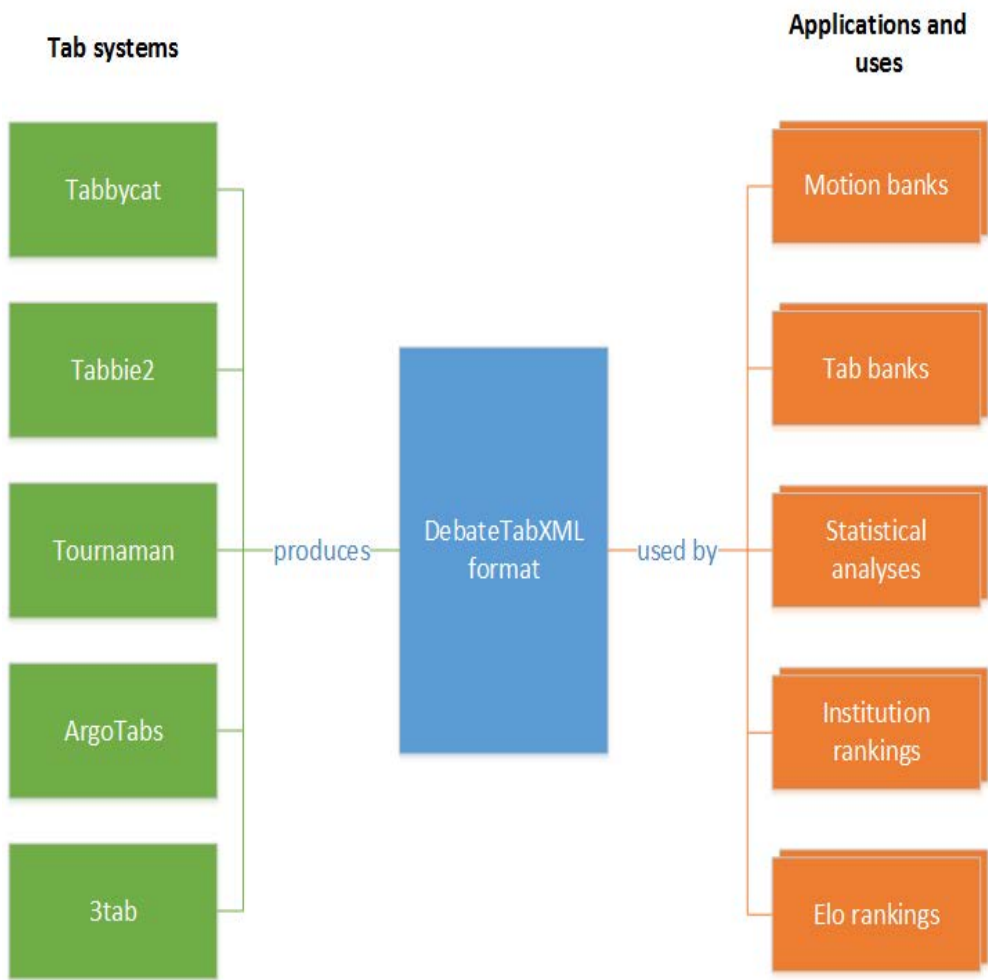
At present, there is no standardised format for recording tab data. Several tab systems do present interfaces that display tournament records in public, most notably the web pages that are published upon the tab's release. However, the visual presentation of tabs as tables is only a small sample of the underlying data recorded by most tab systems. These tables are great for presenting speaker and team rankings to users, but they miss a great deal of information, such as matchups, motion selections and individual adjudicators' speaker scores, that can be used to provide further insight into a tournament.

If we feel that this data is important enough to preserve and analyse, a standardised format is highly desirable. Without a common structure for this information it is extremely difficult to use the same tools or applications on data from different tournaments, even if it is provided by all tab systems.

Some newer tab systems can operate multiple tournaments from a single installation, even allowing for complex correlations to be made over across tournaments. These systems should be commended, but their limitations as a historic archive and comprehensive data set should be recognised. Firstly, no one tab system currently caters to all popular debating formats, meaning each can only cover a subset of the global debating community. Even if a single software could cover all popular formats, tab software is relatively short-lived, as its developers are typically unpaid volunteers who cease development as they retire from circuits. Moreover, the lifespan of tab data is typically short, as tournament-specific websites are not paid for in the long term, tab masters retire, tab data becomes incompatible with future software versions, and web services retire or implement backwards-incompatible changes.

Aims

The basic aim of the standard would be to make it easier for interested debaters to use tab data to archive tabs, run statistical analyses, and create tools such as motion banks. We envision there will, eventually, be a diverse range of applications to help us understand and improve our sport. We do not profess to imagine all such applications—others will undoubtedly have more ideas than we do—but examples could include motion banks, analyses of adjudicator bias or gender equity, records of institutional achievement, and institutional or speaker ranking tables. Proposing or detailing the applications and analyses themselves are beyond the scope of this project. Instead, the goal of the standard is to lay a foundation that radically reduces barriers to performing these activities. With this in mind, we present three major means of achieving this aim: to decouple tab systems from tab data, improve accessibility to tab data and promote the longevity of tab data.

Tab systems

Tabbycat

Tabbie2

Tournaman

ArgoTabs

3tab

produces

DebateTabXML format

used by

Applications and uses

Motion banks

Tab banks

Statistical analyses

Institution rankings

Elo rankings

Decoupling tab systems from uses of tab data. With a common format, applications and analyses would be able to use tab data irrespective of which tab system it comes from as long as the data is exchanged in a manner compliant with the standard. In this way, we decouple the tab systems that provide the data from the applications and analyses that use it.

This decoupling would mean developers of tab systems and other applications would only have to implement one format to be compatible with all other applications. Similarly, analysts would be able to apply the same procedure to data from different tab systems. More broadly, it would allow each tab system and each application using tab data to progress independently of each other, unhindered by difficulties in importing and standardising data.

Improving accessibility. There are two facets to the accessibility of tab data. The more obvious part is to make it easier for interested debaters to find. Of course, the mere existence of a standard does not make this happen. First, we hope tab systems will implement a function to export archive files consistent with this standard. Then, we rely on tab directors making this file widely available, or the tab system making it available through a public interface. Defining a format for this file is the first step: if everyone uses the same standard, it is clear what everyone should implement.

The second aspect of accessibility is what data is available. Currently, tables of speaker scores ("speaker tabs") are routinely available, but more detailed information is harder to find—the type that would be useful in statistical analysis, or would enable functions like filtering motions by topic area, region and balance. The standard will provide a mechanism to make this richer data more available.

Promoting longevity. Our third aim is to facilitate preservation of tab data. The standard itself does not make this happen; however, if tab data can be more easily exchanged, more people will have copies of it and it will be less likely to be lost from the community.

Longevity therefore stems from accessibility, and we envisage a number of applications would support this aim. The most straightforward would be an archival website to which tab directors upload their archives after a tournament. A folder in any file-sharing service would also suffice, but a specialised website could allow users to search or filter for particular tournaments. Another application might be a website that tracks institution rankings over time or with different metrics. These projects will become practical with the introduction of this standard.

Design Principles

This article is not the place for an in-depth discussion of the technical details required to define a robust standard. However, there are several guiding principles which we believe

will help guide discussion and demonstrate the viability of this endeavour.

The standard should admit any tournament structure and debate format. The world has a rich variety of debating formats, varying in aspects such as the number of speakers and teams and whether panels submit single or multiple ballots. Tournaments also have a wide array of structures: some have language categories or a novice break, some allow hybrid teams.

Despite these differences, the participants in many of these formats comprise an international community. Analyses and applications that work across formats and tournament structures therefore have great value and the standard should support this. The flipside is that application developers will need to handle all these cases or detect when a file is not relevant to them. The standard should aim to facilitate this.

The standard should allow for a complete record of the tournament, but be flexible in what information it requires. Most tab systems retain every scoresheet in the tournament. This is valuable information for statistical analysis and debater development, as it provides more detailed information about each debate and enables more complex correlations to be made between speakers, scores, adjudicators, motions, and teams. We propose that the standard allow for archives to optionally include all of this information. This includes, for example, motions (particularly in formats that allow a choice of motions), participating institutions, speaker positions and scores given by individual adjudicators (where adjudicators complete individual ballots). It should also include meta-information about the tournament: when and where it was held, what style it was in, and whether language and novice categories were used.

At the same time, very little information should be required for a file to achieve compliance with the standard. Not all tab systems store the same data: some don't take motions, and some discard scores given by individual adjudicators, storing only the average. Secondly, there is potential for demographic and other fields to be added for statistical purposes, for example, gender, region, or years' experience debating. While this data is useful, tournaments do not necessarily collect it. Thirdly, it is not the intention of this standard to enforce openness of information, merely to facilitate it through a common format. Adjudication cores should retain the ability to set tournament policies without reference to this standard. The only information made mandatory by the standard should be what is technically necessary for the archive to be a coherent record of the tournament's results.

The standard will therefore need to include fields that are sometimes redundant. For example, although the speaker tab can always be generated from the bank of all scoresheets, a tournament that does not release scoresheets should still be able to release an archive containing just the speaker tab.

We note, again, that there is a trade-off here: the more flexible the standard, the more mindful application developers will need to be that not all information can be assumed to be present. We believe that this is acceptable if it means more tournaments can release their archives in a common format.

The standard should be extensible. The needs of the debating community have changed with time, and will continue to do so: consider, for example, the recent advent of information slides. To ensure the longevity of these archives, the standard therefore needs to be able to evolve to add new fields, while remaining compatible with previous versions of the standard. It is this need that informs our suggestion that XML or JSON formats be used. However, extensibility will also need to be kept in mind as we formulate the structure of archive files.

We recognize that spreadsheet-based formats such as CSV files would be more useful in some applications, primarily, data analysis. However, such formats are not naturally extensible and detailed tab data does not lend itself well to a tabular format. As we explain in the section below, we envisage a straightforward tool that could easily generate CSV files from these XML files.

The standard, as far as possible, should not need to be centrally managed. A major strength of open standards is that anyone can implement them without permission and be assured of compatibility with other applications. In a community reliant on short-term enthusiastic volunteers, and a project reliant on the intersection of debaters and programmers, the standard would work best if it did not need to be actively managed by designated individuals.

Inevitably there may be some need for central management. It is desirable, for example, that institutions are identified by consistent and unique codes, which requires the debating community to agree on what those codes are, or at least have a system by which they can be 'reserved'. We hope in principle to minimize aspects such as this that require central coordination.

The standard should be amenable to existing tab system data models. Since a tab archive and a tab database have different purposes, we cannot expect the archive to be a direct export of a tab system's database. Nonetheless, we should consider the data models used by existing tab systems as a way to make implementation of the standard easier.

<u>Implementation and Challenges</u>

We envision tab systems would implement a post-tournament export feature which generates an archive file for any given tournament. Web-based systems would make the archive available for download; stand-alone programs would save this file somewhere for the tab director to make public.

Although the standard itself would be XML- or JSON-based, we imagine that there will be libraries written in each of the major programming languages, to provide an interface for processing tab archive files. Obvious targets include Python and PHP for web applications, and Python and R for statistical analysis. For analysts who use spreadsheets to process data, we suggest that there would be a software tool: users would specify which fields they wish to extract from an archive, and the tool would generate a CSV file with the appropriate columns.

Part of our proposal anticipates that allowance may be made in the standard to optionally include demographic information, such as gender, age groups, region. Since we also anticipate that these archive files will be made publicly available, this raises concerns about the privacy of said information. While we do not believe it is the place of a technical standard to dictate how users should navigate such issues, it may be useful for it to provide some guidance. We hope to consult with others in the debating community about what is best practice in this area.

Next Steps

The first step is for the debating community to agree on the technical details of the open standard. We believe this is best achieved through a process of consensus-building that is public and accessible to all interested parties, as performed in other open standards consortiums such as the World Wide Web Consortium (W3C) or the Internet Engineering Task Force (IETF). To this end we would like to welcome anyone interested to join us on the project's GitHub page, at github.com/TheAgoraProject/dta-spec where we will begin working through the initial details in the coming months. (Users will need a GitHub account.)

Once this is done, implementation of tools using the standard begins. Tab system developers would implement an "export" feature by which an archive file compliant with the standard is generated. We would also write software tools linked to the standard: interface libraries for major programming languages, a tool to extract particular information into CSV files, and perhaps a tool to convert web tables into compliant archive files, so that we can add data from past tournaments to this ecosystem. This is a fair amount of work; it will not happen overnight. Realistically, libraries will be written when they are needed. We hope programmers in the debating community will be interested in contributing to one of these parts of the project.

Conclusion

If the steps detailed above are taken, we believe that tab data will become much more valuable to the debating community in the coming decades. More accessible tab data

would open the door to more enduring historical records. Richer uses of this data would no longer be hypothetical. An open standard for distributing tab archives would be a significant first step in bringing this to fruition.

Author

Philip Belesky (RMIT University, Australia)
Past affiliations: Victoria University of Wellington, New Zealand
Tabbing: Co-tab director of Australs 2012 and Australian Easters 2015. A developer of Tabbycat, a leading tab system for two-team parliamentary formats.
Adjudicating: Judged open semi-finals at WUDC, ESL grand finals at Australs, and the grand finals of the Oxford and Sydney IVs.

Chuan-Zheng Lee (Stanford University, United States of America)
Past affiliations: Victoria University of Wellington, New Zealand; University of Auckland, New Zealand
Tabbing: Tab director of Australs 2014; co-tab director of Australs 2012; assistant tab director of WUDC 2015. A developer of Tabbycat, a leading tab system for two-team parliamentary formats.
Adjudicating: Judged open quarter-final at WUDC, open and ESL semi-finals at Australs, and octo-final at USUDC.

MONASH ASSOCIATION OF DEBATERS
www.monashdebaters.com

Founded in 1962, the Monash Association of Debaters (MAD) is one of the oldest clubs at Monash University and, with over 500 members, one the largest debating societies in the Southern Hemisphere.

The club has a rich history of success in tournaments across the world. In January 2011, 2012, and again in 2013, MAD won the World University Debating Championships, the only club to win in three consecutive years. MAD have also been champions of the Australasian Debating Championships (Australs) six times in the last fifteen years.

The Association has a strong commitment to promoting debating education and innovation throughout the world. MAD is very proud to have founded the only academic journal to cover the art of debating and is a leader in the production of online training resources for the global debating community. Monash was the first club to implement equity policies which ensure diverse participation at national and international tournaments and was a driving force behind the creation of the Australasian Women's Debating Championships.

MONASH ASSOCIATION OF DEBATERS
SECRETARIAT 2015

President: Andrew Elder
Vice-President: Andrew Belyea-Tate
Secretary: Sam Scott
Treasurer: Shara Teo

MAD is supported by the Office of the Vice Chancellor at Monash University
The Monash Debating Review benefits from that support.

Email us: enquiries@monashdebaters.com Visit us: www.monashdebaters.com